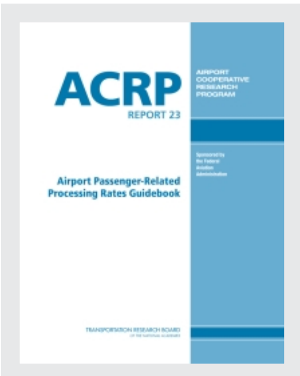


Airport Passenger-Related Processing Rates Guidebook

DETAILS

117 pages | | PAPERBACK
ISBN 978-0-309-43557-4 | DOI 10.17226/22990

AUTHORS

AIRPORT COOPERATIVE RESEARCH PROGRAM

ACRP REPORT 23

Airport Passenger-Related Processing Rates Guidebook

Michael Cassidy
INNOVATIVE DECISIONS INCORPORATED
Arlington, VA
and
MARYMOUNT UNIVERSITY
Arlington, VA

AND

Joseph Navarrete
HNTB CORPORATION
Arlington, VA

Subject Areas
Planning and Administration • Aviation

Research sponsored by the Federal Aviation Administration

TRANSPORTATION RESEARCH BOARD

WASHINGTON, D.C.
2009
www.TRB.org

AIRPORT COOPERATIVE RESEARCH PROGRAM

Airports are vital national resources. They serve a key role in transportation of people and goods and in regional, national, and international commerce. They are where the nation's aviation system connects with other modes of transportation and where federal responsibility for managing and regulating air traffic operations intersects with the role of state and local governments that own and operate most airports. Research is necessary to solve common operating problems, to adapt appropriate new technologies from other industries, and to introduce innovations into the airport industry. The Airport Cooperative Research Program (ACRP) serves as one of the principal means by which the airport industry can develop innovative near-term solutions to meet demands placed on it.

The need for ACRP was identified in *TRB Special Report 272: Airport Research Needs: Cooperative Solutions* in 2003, based on a study sponsored by the Federal Aviation Administration (FAA). The ACRP carries out applied research on problems that are shared by airport operating agencies and are not being adequately addressed by existing federal research programs. It is modeled after the successful National Cooperative Highway Research Program and Transit Cooperative Research Program. The ACRP undertakes research and other technical activities in a variety of airport subject areas, including design, construction, maintenance, operations, safety, security, policy, planning, human resources, and administration. The ACRP provides a forum where airport operators can cooperatively address common operational problems.

The ACRP was authorized in December 2003 as part of the Vision 100-Century of Aviation Reauthorization Act. The primary participants in the ACRP are (1) an independent governing board, the ACRP Oversight Committee (AOC), appointed by the Secretary of the U.S. Department of Transportation with representation from airport operating agencies, other stakeholders, and relevant industry organizations such as the Airports Council International-North America (ACI-NA), the American Association of Airport Executives (AAAE), the National Association of State Aviation Officials (NASAO), and the Air Transport Association (ATA) as vital links to the airport community; (2) the TRB as program manager and secretariat for the governing board; and (3) the FAA as program sponsor. In October 2005, the FAA executed a contract with the National Academies formally initiating the program.

The ACRP benefits from the cooperation and participation of airport professionals, air carriers, shippers, state and local government officials, equipment and service suppliers, other airport users, and research organizations. Each of these participants has different interests and responsibilities, and each is an integral part of this cooperative research effort.

Research problem statements for the ACRP are solicited periodically but may be submitted to the TRB by anyone at any time. It is the responsibility of the AOC to formulate the research program by identifying the highest priority projects and defining funding levels and expected products.

Once selected, each ACRP project is assigned to an expert panel, appointed by the TRB. Panels include experienced practitioners and research specialists; heavy emphasis is placed on including airport professionals, the intended users of the research products. The panels prepare project statements (requests for proposals), select contractors, and provide technical guidance and counsel throughout the life of the project. The process for developing research problem statements and selecting research agencies has been used by TRB in managing cooperative research programs since 1962. As in other TRB activities, ACRP project panels serve voluntarily without compensation.

Primary emphasis is placed on disseminating ACRP results to the intended end-users of the research: airport operating agencies, service providers, and suppliers. The ACRP produces a series of research reports for use by airport operators, local agencies, the FAA, and other interested parties, and industry associations may arrange for workshops, training aids, field visits, and other activities to ensure that results are implemented by airport-industry practitioners.

ACRP REPORT 23

Project 03-02
ISSN 1935-9802
ISBN 978-0-309-11805-7
Library of Congress Control Number 2009938322

NOTICE

The project that is the subject of this report was a part of the Airport Cooperative Research Program conducted by the Transportation Research Board with the approval of the Governing Board of the National Research Council. Such approval reflects the Governing Board's judgment that the project concerned is appropriate with respect to both the purposes and resources of the National Research Council.

The members of the technical advisory panel selected to monitor this project and to review this report were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the project. The opinions and conclusions expressed or implied are those of the research agency that performed the research, and while they have been accepted as appropriate by the technical panel, they are not necessarily those of the Transportation Research Board, the National Research Council, or the Federal Aviation Administration of the U.S. Department of Transportation.

Each report is reviewed and accepted for publication by the technical panel according to procedures established and monitored by the Transportation Research Board Executive Committee and the Governing Board of the National Research Council.

The Transportation Research Board of the National Academies, the National Research Council, and the Federal Aviation Administration (sponsor of the Airport Cooperative Research Program) do not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the clarity and completeness of the project reporting.

Published reports of the

AIRPORT COOPERATIVE RESEARCH PROGRAM

are available from:

Transportation Research Board
Business Office
500 Fifth Street, NW
Washington, DC 20001

and can be ordered through the Internet at

http://www.national-academies.org/trb/bookstore

Printed in the United States of America

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

C O O P E R A T I V E   R E S E A R C H   P R O G R A M S

CRP STAFF FOR ACRP REPORT 23

Christopher W. Jenks, *Director, Cooperative Research Programs*
Crawford F. Jencks, *Deputy Director, Cooperative Research Programs*
Michael R. Salamone, *ACRP Manager*
Theresia H. Schatz, *Senior Program Officer*
Eileen P. Delaney, *Director of Publications*
Maria Sabin Crawford, *Assistant Editor*

ACRP PROJECT 03-02 PANEL
Field of Policy and Planning

Gloria G. Bender, *TransSolutions, Fort Worth, TX* (Chair)
Derrick K.Y. Choi, *Xchange Architects, Brookline, MA*
Anthony Dockery, *Metropolitan Washington Airports Authority, Dulles, VA*
Michael T. Drollinger, *Port of Seattle Aviation Planning, Seattle, WA*
Michael D. Floyd, *JACOBS Global Buildings, Advance Planning Group, Atlanta, GA*
Lloyd A. McCoomb, *Greater Toronto Airports Authority, Toronto, ON*
Wenbin Wei, *San Jose State University, San Jose, CA*
Elisha Novak, *FAA Liaison*
Christine Gerencher, *TRB Liaison*

FOREWORD

By Theresia H. Schatz
Staff Officer
Transportation Research Board

ACRP Report 23: Airport Passenger-Related Processing Rates is a guidebook that provides user-friendly guidance on how to best collect accurate passenger-related processing data for evaluating facility requirements to promote efficient and cost-effective airport terminal design.

Often, significant amounts of data are required in accurate planning and design of airport terminal facilities. Furthermore, the dynamic nature of the airport industry necessitates an understanding of how rapidly changing passenger characteristics, processing technologies, and security protocols impact terminal development. This Guidebook will be of assistance and value to airport operators, planners, designers, and other stakeholders in planning these and future airport terminal facilities.

Planning future airline passenger terminals and assessing existing terminals typically involve the determination of facility requirements. These requirements may be derived by various methods ranging from simple rules of thumb to sophisticated simulation models. However, all methods require data on airline passenger volumes and the rates at which these passengers can be served at ticket counters, baggage check-in, passenger security screening, and other processing points. Passenger processing rates are influenced by many factors, including the type of airline service (e.g., domestic, trans-border, and long-haul international), type of travel (e.g. business or leisure), amount of baggage, and size of party. Recent developments, including the growth of low-cost carriers, increased security, and the increased use of Internet and self-service devices, raise doubts about the validity of data collected in the past.

The objective of this research was to provide guidance on how best to collect passenger-related processing rate data. The research was conducted by HNTB Corporation in association with Innovative Decisions Incorporated, and produced a user-friendly guidebook to identify the best methods of determining accurate passenger-related processing rates. The guidebook covers such topics as confirming the need for data, designing a methodology, defining team roles and responsibilities, selecting an appropriate sampling strategy, choosing a data recording method, establishing staffing requirements, managing staff, and analyzing data.

The results of this research are complementary to the variety of other airport terminal related projects under ACRP's recent research initiatives and are meant to provide a set of coordinated and cooperative guidance tools to assist airport operators and their planning teams. An understanding of best methods to collect accurate data is essential in providing functional and efficient, yet premier airport terminal facilities.

CONTENTS

CHAPTER 1

Introduction and Overview

This chapter describes why the Airport Processing Rates Guidebook was developed; the Guidebook's relationship to other relevant ACRP research and its potential benefits; and finally, the Guidebook's organization and how to access information.

1.1 Background of the Guidebook

Commercial aviation continues to be the most dynamic in the transportation industry. Strong, long-term growth, continuously evolving technologies, changes in passenger processing protocols, and a general shift toward defining passengers as customers have resulted in continued expansion and redevelopment for most major metropolitan airports. Additionally, passenger-related processing variations exist among individual airports due to local conditions such as whether an airport serves primarily business or leisure customers, short-haul or long-haul markets, low cost or legacy carriers, local culture, climate, etc. Not surprisingly, this volatile environment drives the need for airport passenger-related processing data collection initiatives. Data might be collected, for example, during the inventory process of an airport terminal master plan or in support of developing a master plan. Government agencies such as the Transportation Security Administration (TSA) often perform audits of processes to gauge customer service. Airlines gather data to assess how to modify procedures.

These evolving needs demand significant investments, yet these needs come at a time of financially constrained resources and economic uncertainty. In this environment, aviation industry practitioners are becoming increasingly aware of the importance of maximizing the value of the research initiatives they sponsor. A properly designed and executed study can return substantial benefits for the investment and contribute meaningfully to decision making. A poorly designed study will likely not only preclude data useful to others, but may result in unwarranted and unjustified conclusions. The reader may consider the "garbage-in garbage-out" (GIGO) model, attributed to George Fuechsel, as an apt analogy.[1]

This guidebook reflects the increasing emphasis being placed on the application of sound practice from which evidenced-based decision making can be built and which can be seen in many other industries.

[1] http://it.toolbox.com/wiki/index.php/GIGO

1.1.1 The Importance of "Good" Data

The following three arguments are for those skeptical about the benefits of expending resources in pursuit of technically rigorous airport processing-rate research:

1. Research results will help inform airport specific planning questions and decisions, thereby reducing risk. Given the substantial costs of airport renovation and expansion, for example, well-conducted studies can help prevent the error of over-building or under-building facilities.
2. Well-collected processing rate data will be of value at all stages of facility development, including planning (where the data can help determine future facility requirements), design (where various layouts can be evaluated), and implementation (where the performance of various resources can be monitored).
3. Local data collection increases the likelihood of stakeholder "buy-in" to the results. While research findings can often be generalized beyond a specific initiative, confirmatory, local research increases the likelihood that a study's results are credible.[2]

1.1.2 Guidebook Impetus

The impetus for the *Airport Passenger-Related Processing Rates Guidebook* was a charge by industry experts in 2005 through the newly formed ACRP of TRB, a division of the National Academies. In particular, the twofold charge was to do the following:

- Prepare a "unified" large-scale database of existing information on passenger-related processing rates to help in airport planning; and
- Prepare a guidebook that would help practitioners collect this type of data in a manner that was statistically appropriate and would be comparable with data collected from other airports and time periods.

A panel consisting of public and private sector industry leaders was formed to guide the project (designated ACRP Project 03-02); TRB issued a request for proposals in the spring of 2006. In the summer of 2006, the panel selected HNTB Corporation (along with Innovative Decisions, Incorporated, as a sub-consultant) to conduct the research. The project began in the fall of 2006.

1.1.3 Efforts To Date

The first task, developing a unified database of existing passenger processing rate data, was divided into several steps: canvassing airports and consultants to identify available data, collecting existing data, determining the technical and statistical feasibility of creating a "unified" database, and finally, providing a recommendation to the panel as to the appropriateness of building the database.

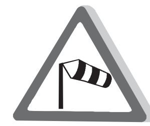
The Research Team collected and reviewed extant databases, and in the fall of 2007, documented its findings to-date in an interim report. In October 2007, the Research Team met with the Panel to discuss its findings and make recommendations as to future work. A final copy of this report is available from TRB upon request.

The principle recommendation to the panel was that the creation of a large-scale, unified database using the files collected as part of this study, while technically feasible, was methodologically flawed. Simply, wide variation and missing documentation across the data files argued against meaningful aggregation. The importance of creating a guidebook that might serve as an impetus and framework for developing such a unified database became apparent: codifying ter-

[2] Parkin, R. T. (2004). Communications with research participants and communities: foundations for best practices. *Journal of Exposure Analysis and Environmental Epidemiology*, 14(7), 516–523. Retrieved Feb. 3, 2009, from ProQuest Science Journals database. (Document ID: 984455131).

minology, suggesting standard methods, and so forth, might facilitate data sharing and aggregation across studies. The panel concurred with the Research Team's recommendation and worked with the Research Team to revise the scope of the study. Development of the Guidebook became the primary emphasis.

1.2 Potential Benefits of the Guidebook



This Guidebook reflects the current best practices in the industry. A guiding principle in its development has been to create a document that is of practical value to the practitioner and grounded in methodological rigor.

There are both immediate and potential long-term benefits to acceptance and use of this guidebook by practitioners.

- In the near-term, it will help users define the purpose of their data collection effort (even determining whether it's necessary or even feasible);
- Second, adhering to its recommendations will help ensure that data are collected and analyzed in a statistically appropriate manner, reducing the risk of making significant development decisions based on faulty data. As data are collected across time, the opportunity to identify meaningful patterns will emerge; and
- Finally, depending on the thoroughness of the data collection efforts, it may be possible to anticipate the impact of future procedures, technologies, and protocols before they are put into practice.

In many instances, the Guidebook provides concrete advice rooted in commonly accepted research practices. In other instances, when a sufficient research base does not exist from which to draw recommendations, the decision was made to include suggestions and *heuristics,* or "rules of thumb." When a recommendation is based on anecdotal evidence, however, it is clearly noted. The alternative, to limit recommendations to those of indisputable quality, would have left the user with too many unanswered questions to be useful.

This version of the Guidebook is a first step. Project sponsors expect that over time the Guidebook will evolve to reflect new knowledge and to adjust to emerging technologies and innovations pertinent to collecting data in the airport terminal environment. Its adoption should make practical the development of a unified database, providing planners and others with a robust set of data for multiple airports, processes, and time periods.

1.2.1 Complement to Other ACRP Research

This Guidebook is focused on providing guidance for collecting airport terminal processing data through *observation.* There are also many key pieces of data that can only be obtained through input provided by users (primarily passengers). Under a separate ACRP research project (ACRP Project 03-04, "Guidebook for Airport-User Survey Methodology"), another guidebook is being developed to help aviation practitioners conduct airport user surveys, and it complements this guidebook by focusing on how to gather data through intercept/interview survey methods.

1.3 The Guidebook's Structure and Organization

The Research Team anticipates the potential audience for this Guidebook to be broad and diverse. To accommodate the needs of this diverse audience, the Guidebook has been structured in a way to permit it to be accessed in different ways based on differing user needs. You can, for example, read a chapter as a largely self-contained unit, or proceed directly to a specific topic. Correspondingly, one with little knowledge of airport processes can read a detailed description of each, while someone else might read a brief overview, assuming a given level of knowledge.

1.4 Content

Chapter 2, "Passenger-Related Processes Overview," provides a general overview of airport passenger-related processes. This review may be beneficial to those who are unfamiliar with these activities. Check-in, security screening, Federal Inspection Service (international arrivals), baggage claim, passenger boarding and deplaning processes, as well as concession use and restroom use are all covered, and trends are also described.

Chapter 3, "Defining the Research: Purpose, Focus, and Potential Uses," is designed to place the need for an airport passenger-related processing rate study requiring data collection in context. It begins with identifying roles, relationships, and responsibilities of stakeholders; it then provides guidance to help the reader determine the right questions to ask, starting with, is a study even needed. The chapter then describes different types of research methods.

Chapter 4, "Designing the Methodology," introduces the concept of metrics and levels of measurement; it then applies these concepts to airport terminal data collection. The concepts of *entities* (e.g., passengers and luggage), *resources* (e.g., self-serve kiosks, ticket agents, security checkpoint walk-through metal detectors, etc.), and *processes* (e.g., checking in for a flight or passing through a security checkpoint) are explained, and recommendations for developing specific operational definitions are provided.

Chapter 5, "Sampling Techniques for Airport Data Collection," provides an overview of the sampling process and recommendations for scheduling airport data collection events, focusing on identifying peak periods of activity at various airport terminal elements. Guidance on determining sample size, minimizing statistical effort, and trading between the benefit and cost of various sampling plans are addressed.

Chapter 6, "Developing the Action Plan," recognizes that collecting a robust, useable, and statistically defensible data set requires a well thought-out plan that ensures sufficient lead time, thorough preparation, the establishment of a well-trained and professional data collection team, and ongoing coordination with stakeholders. Chapter 6 provides step-by-step guidance in establishing roles and responsibilities, discusses the pros and cons of various sources of staffing, offers advice on choosing an appropriate data recording technology, and provides a model schedule countdown that can be used as a framework for specific studies. Finally, the chapter provides practical guidance for determining staffing levels and schedules.

Chapter 7, "Managing and Implementing Data Collection," reviews the importance of site visits, reconnaissance, field testing, and team training. Example case studies are also provided.

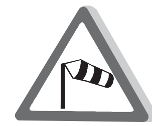
Chapter 8, "Summary," briefly summarizes key factors involved in planning and implementing Airport processing data collection projects.

The Appendices include a glossary of terms; an overview of issues relevant to analyzing and displaying quantitative information in tabular and graphical formats; samples of training and orientation materials; and more detailed information on sampling and statistics than that presented in the body of the Guidebook.

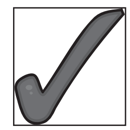
1.4.1 Useful Icons

Background information is identified by the "thinker" graphic shown. Given that the guidebook was prepared for several types of users, (e.g., airport directors, architects, planners), read sections marked with this graphic if, for example, you have little or no familiarity with the topic, are seeking clarification for something presented in a recommendation for practice, or simply have an interest in the subject.

The windsock symbol is used to highlight content in the Guidebook that is strongly suggested for the user to read because it is largely concerned with content that might be in conflict with commonly held assumptions, or argues for the importance of a particular topic.



Recommendations for practice are identified with a check mark.



When a new term is first introduced, it is presented in italics along with its definition. Select suggestions, such as the one on the right, have been inserted throughout the guidebook, and are shown in shaded boxes. Appendix A of the Guidebook includes a glossary of terms.

Do not assume that everyone defines terms similarly. Select or create clear unambiguous **operational definitions.**

CHAPTER 2

Passenger-Related Processes Overview

This chapter provides an overview of the most common passenger-related processes that occur at an airport. Each process is described, and emerging trends are briefly considered.

2.1 Passenger-Related Processes: Overview

Exhibit 2-1 illustrates the general flow of passengers and their bags through the airport terminal environment. While this Guidebook addresses a number of factors pertinent to airport planning, emphasis is on the five primary passenger-related functions defined in Phase 1 of this study. These are the following:

1. Passenger check-in and ticketing (i.e., obtaining a boarding pass and checking bags);
2. Passenger security screening (i.e., the security screening of passengers and carry-ons);
3. *Federal Inspection Services* (FIS) (i.e., processes routinely conducted by U.S. Customs and Border Protection (USCBP) on passengers and bags as they enter the U.S.);
4. Baggage claim (i.e., the transferring of bags from aircraft and displaying them at a claim device for passenger pickup); and
5. Enplaning/Deplaning (i.e., the loading and unloading of passengers and bags from aircraft).

This chapter includes a relatively high-level overview of each process and issues pertinent to research on passenger rates. The intent is to help provide a context for those with varying degrees of familiarity with the environment within which processing-rate research is conducted. Chapter 7 is focused on the practice of data collection and contains a number of specific recommendations related to each process. Exhibit 2-1 summarizes the general processes for departing passengers.

2.2 Check-in and Ticketing Process

2.2.1 Introduction

The check-in process comprises two primary functions: obtaining a boarding pass and "dropping off" any checked bags intended for transport in the *belly hold* of the aircraft. A number of sub-functions might also be executed during the check-in process, including reserving/changing seats, obtaining upgrades, and presenting appropriate documents for international travel. This section briefly describes several dimensions relevant to the check-in and ticketing process, and presents an overview of recent and emerging trends related to this process. Exhibit 2-2 shows a typical domestic passenger check-in process. The initial decision upon reaching the terminal hinges on whether a passenger has already checked in for his or her flight. A passenger's fare class, number of checked bags, and travel itinerary (i.e., domestic or international) also affects how he or she is processed. Given that the flow-chart permits several alternative sequences, the order in which each dimension is described is somewhat arbitrary.

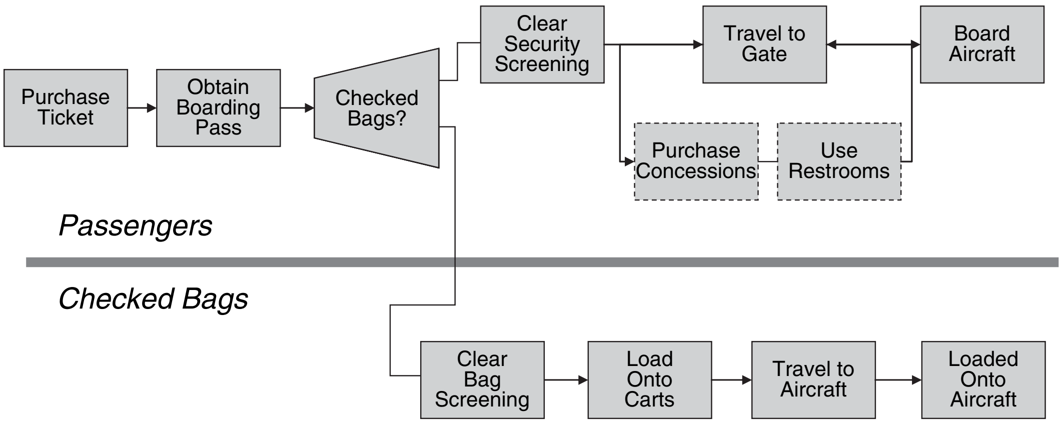


Exhibit 2-1. Generalized airport passenger process (departing passengers).

2.2.2 Check-In Method

All passengers must check in for their flight, which, at a minimum requires obtaining a boarding pass. Today, a significant number of passengers have already checked in for their flights electronically prior to reaching the terminal. Often, these passengers will bypass the airline ticketing/check-in hall and proceed directly to the security checkpoint. Passengers who have not previously checked in can do so at the curb with the assistance of a skycap (this is a common check-in method for passengers with bags to check). Other passengers may decide to use a self-serve kiosk or visit an airline agent at the counter.

2.2.3 Number of Checked Bags

Another relevant aspect of check-in is the process of checking bags. Typically, when an estimate of the number of bags checked is relevant to the research, a member of the study team

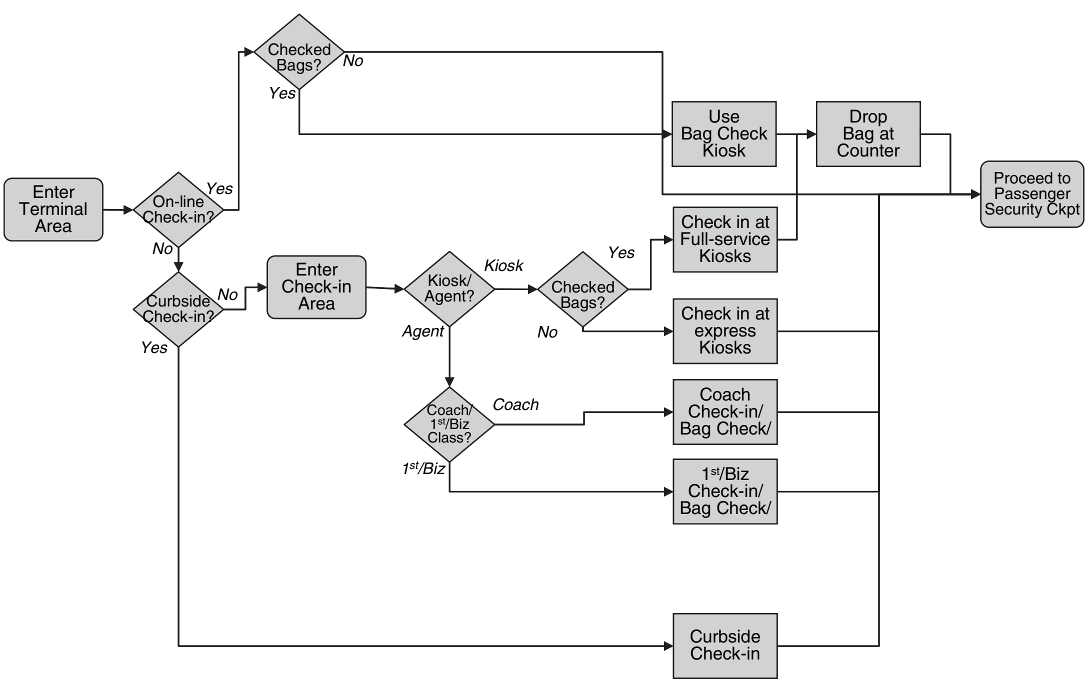


Exhibit 2-2. Domestic passenger check-in process.

counts the number of bags placed on a *bag well* at a staffed check-in counter. The agent then tags the bag(s) with a *bag tag,* lacing it either on a *take-away belt* or returning it to the passenger to be conveyed to a TSA screening area in the check-in lobby.

2.2.4 Fare Class

Airlines frequently provide separate check-in facilities for different classes of passengers (e.g., first class, business class, frequent fliers, etc.), although implementation varies considerably. Airlines might establish a separate queue for these passengers who are taken by the next available agent (i.e., a separate counter/agent is not provided); alternatively, the airline might provide a separate queue and agent(s) restricted to serving only first class or business class passengers; or, in some instances, the airline might provide an entire check-in area for first-class or business class passengers, physically separated from the main check-in lobby to which coach class passengers are directed. Based on anecdotal evidence, there might be a difference between coach and first class processing rates, although this has not been rigorously tested.

2.2.5 Domestic vs. International

Check-in can also be differentiated on the basis of whether the flight is classified as *domestic* or *international.* On average, international flight check-in takes longer than domestic check-in. In general, you can determine if passengers are domestic or international by observing the queue the passengers enter (assuming it is signed for international passengers), or by noting whether the passenger presents a passport.

2.2.6 Check-in and Ticketing: Recent Trends

Historically, passengers were processed for flight check-in either at a ticket counter or at curbside. The choice of where to check in was generally driven by itinerary (domestic or international) and class (first class or coach). The ticket counter was a central location for all functions: check-in, seat selection, boarding passes, and checking luggage. In the late 1990s, the check-in process began to change in substantial ways, driven by new technologies, as well as the airlines' goals of processing passengers more efficiently and at a lower cost. Today, most passengers use some form of electronic check-in (either checking in online or using a self-serve kiosk at an airport or a remote location). The check-in process continues to evolve in response to three principal objectives:

- Increasing decentralized services,
- Providing more disaggregated services, and
- Enhancing customized services.

2.2.7 Check-in Decentralization

Check-in decentralization is typified by the airlines' introduction of self-serve kiosks at non-traditional locations within an airport (for example, in parking garages, at the curb, or on the concourse). In addition to increasing customer convenience, adding kiosks in remote locations also reduces congestion in the check-in lobby. The emergence of online check-in redefines decentralization beyond the physical limits of an airport. Passengers are now able to check-in at remote locations. For example, passengers can check in for return flights while checking in at their origin airport for their outbound flight (provided that it is within 24 hours of the outbound trip).

2.2.8 Check-in Disaggregation

Disaggregation refers to the separation of activities during check-in. For example, even if passengers obtain their boarding pass off-airport or at an airport kiosk, they must still interact with an agent to drop off checked bags. While the separation of these two activities (passenger check-in and baggage check-in) reduces staffing requirements and helps expedite the overall process, it complicates data collection.

2.2.9 Check-in Customization

Technology has given passengers more options. For example, passengers might select express kiosks if they are *not* checking bags or a full service kiosk if they *are* checking bags. Another example is the installation of flush-mounted kiosks at counters. When a counter is so equipped, agents can offer varying degrees of customer assistance depending on a passenger's level of experience and the complexity of an individual transaction. From a data collection standpoint, this trend has blurred the distinction between "self-serve" and "traditional full-serve" service.

An emerging example of customization is the potential use of electronic boarding passes, currently being tested by Continental Airlines and TSA. By displaying a bar code on a cell phone or Personal Digital Assistant (PDA), passengers can enter security checkpoints and board aircraft without having to handle a paper boarding pass. Should this technology be broadly adopted, the number of passengers visiting the ticketing lobby will likely decrease.

Other examples of emerging customization include the following:

- *Common-use-self-service* (CUSS) kiosks, which are more often found outside the United States, are becoming more popular domestically.
- Check-in at hotels—printing boarding passes at hotels.
- Roving agents with carts and hand-held devices who can check in customers.
- International online check-in.

2.3 Passenger Security Screening

While the check-in and ticketing process is focused on efficiently processing people and baggage onto an aircraft, security screening is concerned with reducing potential safety threats and doing so with maximum efficiency.

As shown in Exhibit 2-3, two types of *entities* are processed concurrently in the security screening phase: passengers and their belongings. Prior to stepping through a *metal detector,* passengers enter a divesting area where they place their carry-on items on a feed belt that transports the items through an x-ray machine. Upon clearing the metal detector and entering the secure side of the airport, passengers step toward a composure area to collect their belongings. Should an alarm sound or a suspicious item be observed, the passenger and/or his/her belongings would move to

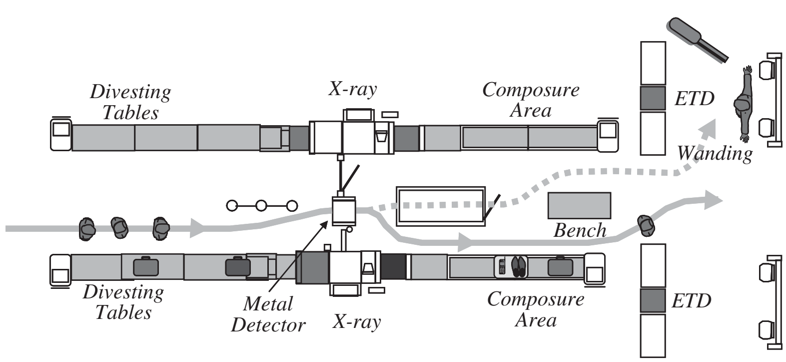


Exhibit 2-3. Passenger security screening process.

a secondary screening area. At this location, passengers would be "*wanded*" with a portable metal detector, and/or their bags would be hand inspected or be further analyzed with trace detection procedures.

The overall capacity of a checkpoint is determined by the slower of these two processes. At a typical checkpoint during a busy period, for example, the walk-through metal detector is unused while the adjoining belt, transporting personal effects and carry-on items through the x-ray machine, is functioning continuously, or operating at capacity.

Overall *security screening checkpoint capacity* is measured by counting the number of passengers going through a metal detector during a period of constant demand (i.e., when there is a queue). Screening of carry-on baggage includes x-raying/scanning, *trace detection,* and *hand inspection.* For passengers, screening is done during the initial walk through the metal detector (or *explosive detection trace portal*), "*wanding,*" and *pat-downs.*

Both the security protocols and technologies are continually evolving at the checkpoint, and passenger response to these changes is ongoing as well.

People screened at a checkpoint might potentially be classified in one or more ways. For example, they could be categorized as airport employees, airline employees, or passengers. Unless these pedestrian types have their own screening lanes, it is often difficult to classify pedestrians by type. Further complicating matters is that some people in the employee line are actually non-revenue passengers, and when these groups are mixed, it is nearly impossible to assign a specific process time.

2.4 Federal Inspection Services (FIS)

Passengers arriving at an airport from an international destination are generally processed at an FIS facility.[1] In 2003, the U.S. Immigration and Naturalization Service and U.S. Customs functions, as well as the functions of some other agencies, were combined into *Customs and Border Protection* (CBP) under the Department of Homeland Security (DHS).

Exhibit 2-4 illustrates the general FIS process. Typically, passengers entering the U.S. go directly to the FIS inspection area upon deplaning. There, passengers submit travel documents for review by a CBP agent (passport control). Visitors are also photographed and *finger-scanned.*

Concurrent with passenger inspection, international passengers' luggage is offloaded and transported to baggage claim devices within the FIS. Some passengers may undergo additional processing at a CBP agent's request. After processing, passengers typically retrieve their luggage and proceed to a CBP agent for customs processing. At this stage, passengers present forms they have completed pertaining to items being brought into the U.S. This step in the process may involve bag inspection or payment of a *duty.* Finally, upon exiting the FIS, passengers may be re-screened prior to entering the domestic portion of the terminal.

Data collection for FIS processing includes such elements as inspection throughput rates, bag inspection demand and processing rate differences based on factors such as party size, and so forth.

[1] Passengers arriving from several Canadian cities, Caribbean locations, and Shannon, Ireland, can be "pre-cleared" at the origin city and do not need to pass through an FIS in the United States.

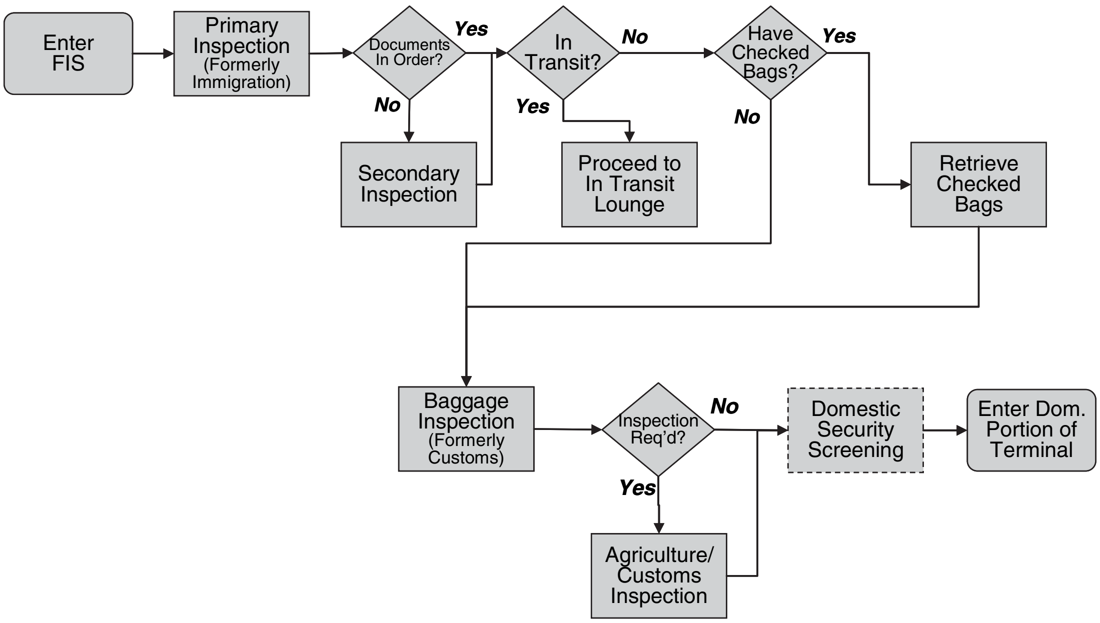


Exhibit 2-4. General FIS process (passengers).

2.5 Baggage Claim

The baggage claim process typically varies by the size and activity levels of airports. At smaller airports, for example, baggage delivery is done "*through the wall.*" At larger airports, bags are often transported from the aircraft and offloaded onto baggage claim devices for passenger retrieval.

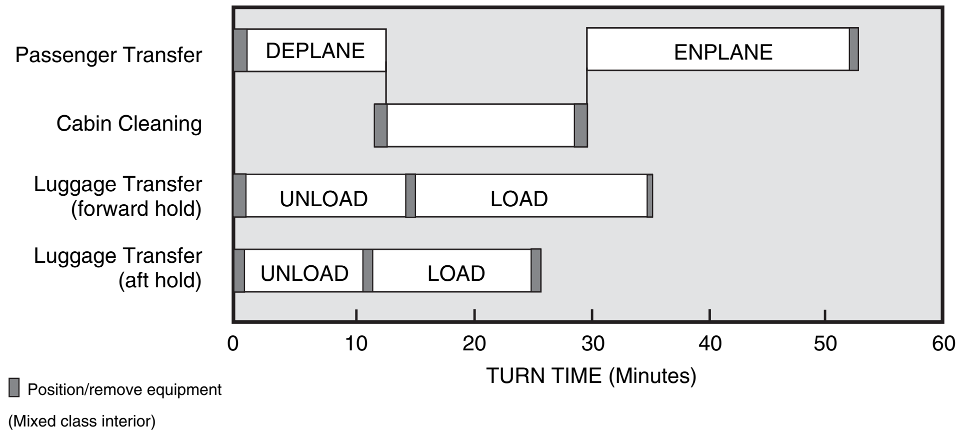
In general, the farther a bag has to travel between the arriving aircraft and the baggage claim device, the longer it will take for a bag to appear at the carousel. A less obvious phenomenon is that for airlines performing hub operations at an airport, priority often appears to be given to baggage intended for connecting flights. As a result, delivery of local bags might be delayed.

The key items typically observed at bag claim include bag feed rates onto the carousel, number of bags that can be presented on a carousel, and number of passengers in the active claim area.

2.6 Passenger Enplaning/Deplaning Rates

A key factor influencing operational efficiency is passenger enplaning and deplaning rates (i.e., the time needed for passenger boarding and offloading, respectively). These rates are influenced by a number of factors including airline pre-boarding practice; boarding technology (e.g., loading bridge, air stairs, mobile lounge); aircraft type and configuration (seat layout, number of aisles, aisle width, door positioning and width, and number of doors, amount of overhead space); and passenger mix (party sizes and make up, number and size of carry-ons, etc.). These variables may be used to identify strata to help isolate their impact on passenger enplaning and deplaning rates.

As reflected in Exhibit 2-5, a flight's overall turn time—the time between an aircraft's arrival at a gate and its pushback from the gate—is influenced by other activities including unloading



Source: *The Role of Computer Simulation in Reducing Airplane Turn Time,* Boeing Commercial Airplane Group.

Exhibit 2-5. Typical aircraft turn time (Boeing 757-200 with 201 passengers).

and loading luggage, fueling the aircraft, servicing the aircraft (e.g., cabin cleaning, galley servicing), and pre-flighting the aircraft, etc.

2.7 Concessions

A good concessions program can help enhance customer service while also providing a potentially significant source of income for airports. The long-term trend is for an increasing share of an airport's total square-footage to be allocated for concessions. The following three factors appear to be driving this phenomenon.

1. Airlines are increasingly cutting-back on amenities, such as in-flight meals, as part of a ticket charge.
2. The average trip length has increased, increasing the likelihood that passengers will purchase food in the terminal for in-flight consumption.
3. Finally, largely in response to increased security measures, passengers are allocating more time at airports prior to their scheduled departure which, in turn, increases the time they have for visiting concessions.

Exhibit 2-6 lists some common types of concessions that can found at airports. Smaller airports may only feature vending machines, while mid-sized airports might have snack food concessions and newsstand/gift shops. Large international airports will also have the greatest variety of concessions, including duty free shops and currency exchange.

While attention usually focuses on customer service and expenditures, concession activity can still be examined from the perspective of a process (e.g., efficiency and layout).

2.8 Restrooms

Terminal restroom facilities have become more crowded as aircraft size and passenger loads have increased. And while demand for air travel has increased overall, the percentage of women passengers has increased more rapidly. Given that many facilities were constructed when a disproportionate number of business travelers were male, female restroom demand might well exceed capacity, creating an extra burden on both facilities and female travelers.

Food	Services	Sundries
Vending Machines	Rental Car	Newsstand/Gift Shop
Snack Food/Coffee Counter	Shoeshine	Duty Free
Fast Food Restaurants	Post Office	Specialty Retail
Sit Down Restaurants	Business Service Center	
	Bank/ATM	
	Currency Exchange	
	Airline Club	

Exhibit 2-6. Typical airport concessions.

Members of the Research Team have collected temporal demand data (i.e., the number of people entering restroom facilities over a given time period); data on the amount of time people spend in restrooms (without being tied to a particular activity);[2] as well as data on simultaneous demand (i.e., the number of people who want to use a particular type of fixture—urinals, stalls, and sinks—at any given moment in time). Chapter 7 includes suggestions for capturing restroom use data.

[2] The duration measurement began when a person entered a restroom and terminated when the same person exited the restroom.

CHAPTER 3

Defining the Research: Purpose, Focus, and Potential Uses

Chapter 3 identifies roles, relationships, and responsibilities of stakeholders. It examines principal steps involved in planning an airport passenger-rate data collection effort. It begins with the question of whether the potential benefits of the proposed effort outweigh the anticipated cost; describes different types of research (i.e., exploratory, descriptive, inferential); summarizes the questions each type addresses; and notes the ends to which the data might be used.

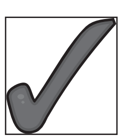
3.1 Roles and Responsibilities

When an airport data collection event is first mentioned, it invariably raises numerous questions: Who is asking for the data? How will it be used? What's the budget? What's the schedule? What kind of resources can be made available? Without answers to these fundamental questions, the success of your research is in jeopardy. This section will help the researcher establish the role of key stakeholders and their interrelationships within the team.

Many entities can sponsor a data collection study, including airports, airlines, manufacturers, and various agencies. Likewise, there are many ways of managing and staffing the event and promoting involvement with stakeholders. There are therefore myriad ways of organizing a study. Exhibit 3-1 is an example of how a study could be arranged with the airport as the sponsor.

3.1.1 Client/Sponsor

For airports, oversight is guided by a board, commission, or an authority consisting of appointed or elected officials. While these agencies typically provide oversight to airport management and approve long-term plans and large capital expenditures, usually it is the airport director or manager who makes day-to-day decisions. Depending on the size of the airport, there may be several departments, each having its own manager. In such cases, passenger terminal-related studies would typically fall within the purview of the planning and/or engineering department and would be managed by its director.



Regardless of the affiliation of the project sponsor(s), it is essential that the following questions be answered clearly and unambiguously as they pertain to the sponsor at the beginning of any project:

- Who has primary responsibility for defining the questions the study is intended to address?
- What preference does this person or group have regarding ongoing involvement with the project?
  - What information would they like to receive, in what format, and with what frequency?
  - Who should be the principal point-of-contact (POC) on the client's side for questions that might emerge related to the study's focus, direction, etc.?

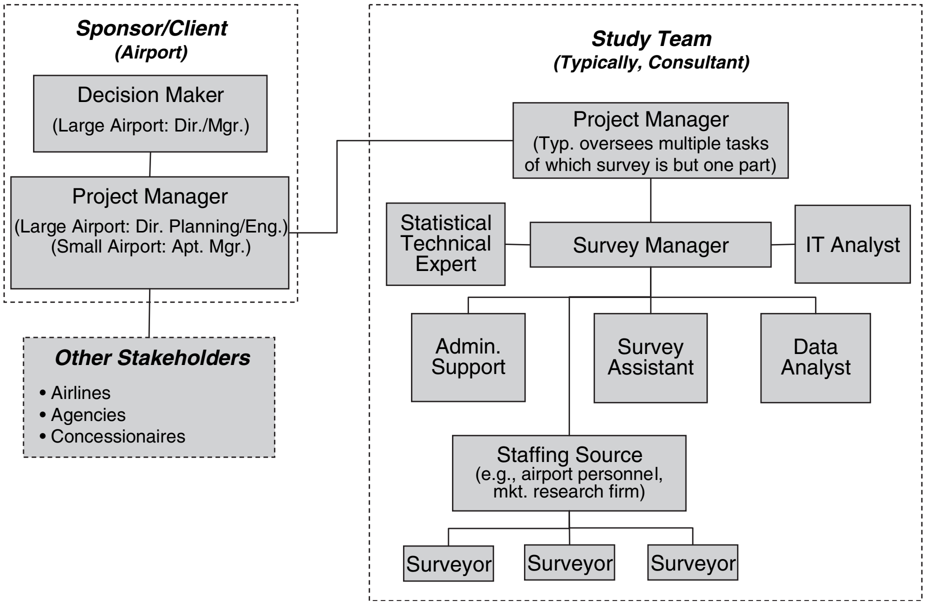


Exhibit 3-1. Typical sponsor and study team roles (assuming an airport is the sponsor).

- Who is the designated project manager, and what information would he or she like to receive, in what format, and with what frequency?
- If the person given responsibility for day-to-day issues pertaining to access, authorizations, etc. is different from the project manager, who is that person, and what is the scope of issues he or she is authorized to address?
- If problems or obstacles arise in implementing the study, and the project manager is not able or authorized to resolve them, what is the chain of persons through which the issue should be escalated?

3.1.2 Study Team

The size of the study team will depend on the team's depth and organization, and the size, duration, and complexity of the study itself. For a typical medium- to large-scale study, the roles listed in the following sections are the most typical. Multiple roles might be assumed by a single person or distributed across multiple persons. Titles vary as well, but the functions are largely universal.

Project Manager

The project manager is typically a mid-level to senior person who has the long-term, day-to-day relationship with his or her client counterpart. The need for the passenger-related processing rate study may initially originate from discussions between the project manager and those within the airport or airline.

Survey Manager

The survey manager is usually a mid-level staff person. His/her role on the project would be to oversee the day-to-day management of the data processing rate study, including leading the development of the scope, schedule, and budget; developing the team; and assigning roles and responsibilities. The survey manager would have the responsibility of ensuring the survey goals were adequately defined and met.

Research and Statistical Expert

A person(s) with expertise in research methodology and quantitative/statistical analysis should be consulted to develop, or provide comments and recommendations about, the overall methodology, the sampling plan, and so forth. Most of this person's input would occur at the project's initiation. A distinction is sometimes drawn in the consulting literature among different approaches to consulting. One such approach, generally referred to as *process consultation* might be of particular appeal.[1] When acting in this role, the consultant not only provides technical expertise related to the specific project, but also works with the client to develop expertise. This arrangement has the goal of, over time, reducing the reliance on the consultant.

Survey Assistant

The survey assistant has primary responsibility for assisting the survey project manager and secondarily to assist others on the project team throughout the duration of the study. Typically, this staff person will be at a junior level. The degree of assistance this person can provide is based on his/her level of education and current skill sets.

Data Analyst

The data analyst should not only be well-versed in technical analysis, but should also have a strong familiarity with the airport terminal environment. This person could be a terminal or airport planner or aviation architect. The analyst is often largely responsible for documenting results, and responsibilities might extend to presenting findings to the client.

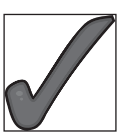
Administrative Support

Data collection efforts are inherently complex and, as such, often require a significant level of coordination and administration. The staff person serving this function would be responsible for such things as making travel plans, scheduling visits to the airport's security office, buying supplies, shipping and receiving materials, scheduling meetings, preparing invoices and contracts, and editing/proofing the report.

Data Collection Staff

For small studies (e.g., small airports where only a few functional elements are being observed for a limited time period), airport/airline or consultant staffing may be used. For larger studies, typically examining multiple functional elements of a medium or large airport over a multi-day period, a market-research firm is frequently employed. The data collection staff reports directly to the survey manager.

3.2 Is the Study Needed?



While the need for data collection is often justifiable, the benefit of validating the need, and avoiding what might be a costly, and possibly unjustified, effort well exceeds the relatively minor cost of pausing to consider a few basic questions (see Appendix C for more information). Exhibit 3-2 illustrates these questions.

3.3 Research Fundamentals

This section summarizes a number of fundamental issues and terms related to the research process. (Additional detail is included in Appendix C.)

[1] Schein, E. H. (1999). Process Consultation Revisited: Building the Helping Relationship. NY: Addison Wesley.

Question	Things to Consider
Are there data available that might help answer the research question?	Have relevant data been collected at this airport in the past that might be used rather than collecting new data? Might you be able to get data from another airport similar in key ways to this airport? Might access to the data be blocked due to proprietary or security issues? Sometimes the data are perceived to be so sensitive that the "owner" of the data may not give permission to share it.
What role will the results play in the decision being considered?	Has the decision already been made, and the data are being collected to legitimize the decision? Is there anything to suggest that the study is an attempt to "prove" something true or false? To what extent will the decision makers be persuaded by the results?
What will the decision makers accept as credible evidence?	Before collecting data, make certain that the research plan will result in data that the sponsors will accept. It is better to learn beforehand, for example, that the proposed sampling plan does not meet the sponsor's criteria for rigor.
What is the cost of the potential investment that the data will help inform? What is the cost of conducting the research?	Does the benefit equal or outweigh the cost? Cost should be considered not only in economic terms, but as safety, inconvenience, and so forth.

Exhibit 3-2. Considerations to determine need for data collection.

Research is a dynamic process with both *deductive* and *inductive* dimensions. This differs in some ways from what some present as the "traditional" approach to research, i.e., that theory drives hypothesis testing. Sometimes it does, but sometimes it doesn't work this way.

3.3.1 Theory, Hypotheses, and Evidence

The word "theory" often implies a formal set of laws, propositions, variables, and the like, whose relationships are clearly defined. A related implication is that theory may not be particularly germane to the everyday world of work.

This view of theory is not incorrect, but neither is it complete. While theory can be abstract and complex in its detail, it does not necessarily have to be abstract, complex, or formal. It can be thought of more broadly and simply as an explanation of "how the world works." For example, an organization might develop a mission or a value statement (or both); engrave the words in a medium intended to last millennia; and prominently display the statement in the workplace with the intent of communicating to all its perspective clients on issues pertinent to its view. In

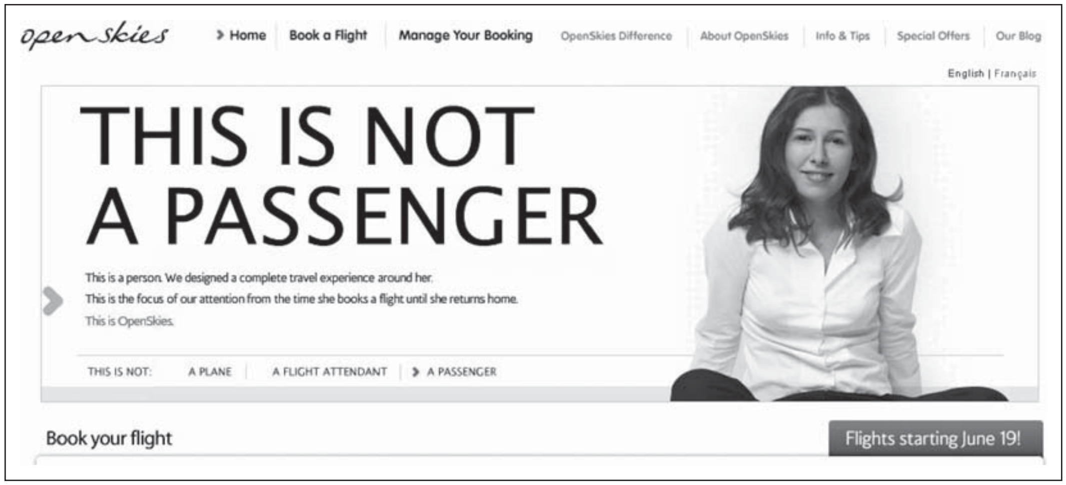


Exhibit 3-3. OpenSkies advertisement.

2008, British Airways announced a new venture: OpenSkies. The "theory" OpenSkies used to define its clients is reflected in its advertising as shown in Exhibit 3-3.

So, how does this relate to airport processing rate studies? It relates in the following two ways:

Question key assumptions, even if they seem to be "common sense," by checking with **informants,** looking at the literature, etc.

1. The published research literature may well contain formal theories relevant to what data to collect and how to collect it. For example, Appendix B includes a bibliography of recent research articles related to passenger and baggage processing in airports. It is intended to illustrate the scope and diversity of research available on a given topic. Before embarking on an investigation, review the literature to see how it might enhance the quality of the planned research. The Internet provides access to numerous sources for such scholarly documents.
2. Informally, the key decisions about how to go about collecting data are grounded in assumptions about how things work (i.e., one's own theory). For example, you might choose to collect passenger security screening data between 6:00 a.m. and 8:00 a.m. on a Monday because your experience is that this time period reflects peak checkpoint activity. While this "theory" may be correct in some circumstances, it may also be wrong in others. For example, at many vacation-oriented airports, the peak at the checkpoint occurs in the late morning due to check-out times at hotels.

Another common view of research is of the stereotypical scientist, objectively testing hypotheses (or an "educated guess") arising from theory. Exhibit 3-4 reflects this general approach to research.

This is certainly one way in which research proceeds, but, similar to theory, it is not the only way. Before considering an "evidence first" approach, we wish to mention a variation on the traditional approach displayed in Exhibit 3-4 that has been gaining dominance in recent years. In particular, this is a *confidence interval* (*CI*) approach rather than a hypothesis driven approach. In a hypothesis driven approach, the researcher's primary interest is in testing a population parameter, and uses a sample drawn from the population. When the researcher takes a CI approach, the intent is to calculate an interval within which the population parameter is likely

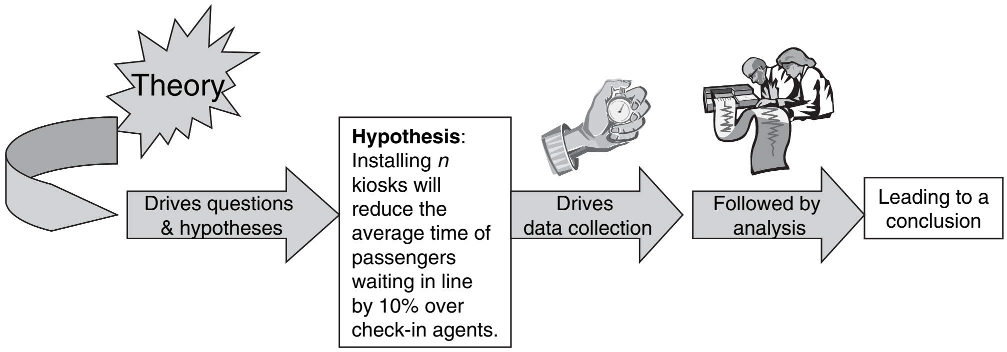


Exhibit 3-4. Hypothesis driven approach.

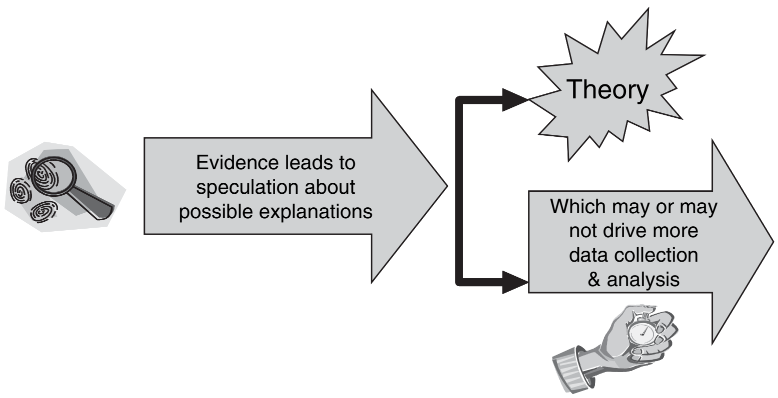


Exhibit 3-5. Bayesian approach.

to fall. Hypotheses are stated before data collection; CIs are calculated after data are collected.[2] In conducting passenger-processing rate research in airport environments, the CI approach is going to be the most appropriate in most instances.

A markedly different approach to those described above is shown in Exhibit 3-5. In contrast to beginning with a theory and then collecting evidence to test the theory or estimate a population parameter within some CI, this approach begins with evidence for which one seeks potential explanations, or "theories" to explain the evidence. This approach is subsumed under the broad heading of Bayesian Law, so named after the 18th Century English clergyman, Thomas Bayes, credited with developing the approach. Depending on where one begins can result in potentially dramatic conclusions (see Exhibit 3-6).

This is important because limiting oneself to a particular perspective of how research should be conducted and how data ought to be gathered may impose unnecessary constraints. What is important is that the research is executed systematically and with rigor. The documented ways in which science proceeds are often idealized: portraying what is inherently a very dynamic and nonlinear process as logical and linear.

3.3.2 Research Questions and Purposes

A basic issue in research is specifying the question the research will help answer. Penning a specific question also helps in determining what approach might be best used in seeking an

[2] While these approaches are presented here as mutually exclusive, they might be integrated in practice.

If your intent is to…	And take action based on…	Use…	Example
Test a hypothesis regarding a population parameter	Whether you reject or fail to reject the null hypothesis	Hypothesis testing approach	The proportion of coach passengers checking in more than 60 min prior to scheduled departure is 80% $H_A : p > .80$[3] $H_0 : p \leq .80$[4]
Estimate a population parameter	The confidence interval selected	CI approach	Plus or minus 5%, what is the average time coach passengers check in prior to scheduled departure?
Determine the likelihood of an event given some evidence	The calculated probability	Bayesian approach	What is the probability that a passenger's carry on-luggage will be subject to secondary security screening given that the passenger is boarding an international flight?

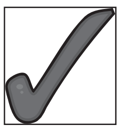


Exhibit 3-6. Research approaches.

answer. One classic text in research methodology[5] suggests that a research question should express a relationship between two or more *variables,* and it should imply an *empirical* approach, that is, it should lend itself to being measured using data. A variable is, not surprisingly, something that can vary, or assume different values.

In the next section, illustrative questions are given, categorized by the purpose of research with which they are best matched. The five research purposes are presented as the following:

1. Explore,
2. Describe,
3. Test,
4. Evaluate, and
5. Predict.

The distinctions among these purposes are not absolute, nor are they necessarily exclusive of one another. A research initiative might be directed at answering questions with multiple purposes. Indeed, this is but one of many ways of classifying research. In addition, the reader whose practice lies primarily in the arena of modeling and simulation might note their absence from this list. Although modeling and simulation applications require input data, for example, to generate distributions and parameters for use as stochastic varieties in modeling, the techniques used to collect data are largely independent of specific applications (such as simulation and modeling). Those issues unique to modeling are beyond the scope of this guidebook.

Explore (Exploratory Research)

Exploratory research is sometimes defined as "what to do when you don't know what you don't know." Its aim is discovery and to develop an understanding of relevant variables and their interactions in a real (field) environment. Exploratory research, as such, is appropriate when the

[3] This is the research, or Alternative, hypothesis. It reads: The proportion is greater than 80%.
[4] This is the null hypothesis. It is what is tested, and reads: The proportion is less than or equal to 80%.
[5] Kerlinger F. & Lee, H. (2000). *Foundations of Behavioral Research, 4th ed.* NY: Harcourt Brace.

problem is not well defined. For example, passenger complaints about signs within a facility might prompt the following exploratory question:

- "Where should signage be located to minimize passenger confusion?"

As another example, if a new security checkpoint configuration is proposed, it may be too novel to rely on variables used in other studies. The question, therefore, might then be the following:

- "How does a given alternative security checkpoint configuration affect capacity?"

This type of research is often qualitative rather than quantitative. That is, it employs verbal descriptors of observations, rather than counts of those observations (see Appendix C for more information).

Describe (Descriptive Research)

Descriptive research, as the name implies, is intended to describe phenomena. While descriptive research might involve collecting qualitative data by asking open-ended questions in an interview, it typically employs quantitative methods resulting in reporting frequencies, calculating averages, and the like.

The following two questions illustrate the nature of descriptive research. Each implies that the relevant variables have been identified as well as the conditions under which the data should be collected.

- "What is the average number of passengers departing on international flights on weekday evenings in July at a given airport?"
- "How many men use a given restroom at a particular location at a given time?"

Test (Experimental and Quasi-experimental Research and Modeling)

Often, the intent of the research is not simply to describe something, but to test the impact of some intervention. In an airport environment, such research might be initiated to evaluate the relative effectiveness of a security screening technology in accurately detecting contraband. It is similar in approach to research conducted to assess the relative effectiveness of an experimental drug in comparison to a control (*placebo*) or another drug. Variables are often *manipulated* and *controlled.* This research lies largely outside the scope of this guidebook and, as such, will not receive much attention. Examples of questions that might be asked in this type of research include the following:

- "What is the impact of posting airline personnel near check-in waiting lines on the average passenger waiting time?"

In addition to the classic "experiment," simulation modeling might be used, employing representative data to help answer questions such as the following:

- "What would be the impact on processing time of a new security measure being considered?"
- "How many agents are needed to keep passenger waiting time below an average of 10 min?"

Evaluate (Evaluative Research)

Sometimes, the intent of the research is to assess performance against some standard or stated requirement. Basically, evaluation research is concerned with seeing how well something is working, with an eye toward improving performance, as illustrated by the following two questions:

- "Is the performance of a given piece of equipment in the field consistent with manufacturer's specifications?"
- "On average, what proportion of passengers waits in a security checkpoint line longer than the 10-minute maximum threshold specified by an airline?"

Research Purpose	Characteristics
Explore	Primary purpose: to better define or understand a situation. Data will help answer the research question. The benefit of conducting the research justifies the cost. Qualitative data are recorded, using observation.
Describe	Primary purpose: to provide descriptive information about something.
Test	Primary purpose: to assess the impact of a proposed change in procedure or policy.
Evaluate	Primary purpose: to assess performance against requirements.
Predict	Primary purpose: to consider possible future circumstances with the purpose of being better prepared for emerging trends.

Exhibit 3-7. Summary of research types.

Predict

Finally, research might be initiated to attempt to predict or anticipate potential emerging patterns before they occur. This is related to *environmental scanning,* insofar as it represents a deliberate attempt to monitor potential trends and their impact. For example, in the early 1970s, one might have posed the following question:

- "What would be the impact of an increase in the number of women in the workforce on airport design?"

There are numerous documented approaches to answering questions such as these. While well beyond the scope of this guidebook, here is one as illustrative: scenario planning. This method involves convening persons with relevant expertise to identify those areas that might most impact the industry (e.g., regulation, fuel costs, demographic changes), and then to systematically consider what the best, worst, and might likely scenarios might be. The principal value of such an approach is that it facilitates deliberate consideration of future trends, and in so doing, presumably leaves people better prepared.

When the goal of the research is to predict, data from multiple sources might be sought. The scenario planning example relies, to an extent, on the judgments of experts. Probabilities can also be drawn from historical data to help identify patterns and trends. Exhibit 3-7 is a summary of the key characteristics of each research type.

3.4 Developing the Research Plan

Large research studies, particularly when funding is being requested, often require the researchers to adhere to a specific set of technical requirements. The Research Team is aware that the *ad hoc* and short timeline of many airport-planning research efforts makes developing a "formal" research plan impracticable. Nonetheless, even though you might not have the "luxury" of developing such a plan, there are benefits to considering the issues described in this section, as well as documenting basic information. The following are the three major elements the Research Team believes worth documenting, regardless of the size of the research endeavor.[6]

1. Goals or aims.
2. Background and significance.
3. Research design and methods.

Each is described in the sections that follow.

3.4.1 Goals or Aims

Specify the question the research is intended to help answer or the specific purpose of the research. The experience of having to translate an intended purpose into words can help clarify your intent. In addition, a written statement can serve as a way of ensuring that your understanding of the purpose of the research is consistent with that of the sponsor and other stakeholders. Two examples follow:

Statement of Purpose—Example 1

The purpose of this study is to aid decision makers in determining if extending the dwell time of the airport's automated guideway transit system (AGTS) vehicles from 30 sec to 35 sec at the Concourse C station might improve overall system capacity by providing more boarding time for passengers.

Statement of Purpose—Example 2

The goal of this study is to provide airport management with recent data showing the percentage of arriving flights whose first checked bag reaches the claim device within the airport's goal of 15 min.

3.4.2 Background and Significance

Document what is already known, and specify how the proposed research initiative will add to this knowledge. Consider a "devil's advocate" perspective by asking what the consequences of not doing the research might be.

3.4.3 Research Design and Methods

In this section, describe how you will go about collecting and analyzing data. Additional information about these issues, including sampling strategies and sample size, is presented in Chapter 5 and in Appendix C.

The research plan does not need be lengthy. It should, however, capture key information that, were it not documented and those familiar with the research were not available, would be difficult to ascertain.

[6] This section is partly based on guidelines published by the Agency for Healthcare Research and Quality, Department of Health and Human Services. http://www.ahrq.gov/fund/esstplan.htm.

CHAPTER 4

Designing the Methodology

This chapter begins with a general treatment of metrics and levels of measurement. Airport specific metrics are then considered, and entities, resources, and processes are defined as they relate to the airport environment.

4.1 Metrics Overview

The term *metric* is often used synonymously with the terms *measure* and *indicator*. A metric is an indicator insofar as it points to something else (as in the way a person uses an index finger for pointing). It also connotes performance. Simply, data are collected with a purpose—to assess how something is performing, should perform, etc. The amount of time to process a given type of passenger, for example, might be an indicator of service quality, or efficiency.

Metrics are sometimes categorized as being counted or measured. One *counts* bags, passengers, number of domestic flights, etc. These largely objective measures are relatively easy to deal with, and are the focus in this Guidebook. By contrast, one may *measure* passenger satisfaction, agent courteousness, and other attributes, but these concepts are abstract and, as such, more difficult to assess. They are largely outside the scope of this Guidebook.

At the risk of stating the obvious, error is unavoidable if there is inconsistency in how fundamental terms are defined or interpreted. Ambiguity will invariably result when collecting data in the airport terminal environment if terms are not fully defined. For example, Exhibit 4-1 shows three acceptable definitions for the term “passenger.” Note that while no definition is inherently correct or incorrect, failure to define the term will likely not only result in confusion but, ultimately, in a serious reduction in the value of the collected data.

An *operational definition* specifies precisely how counts or measurements will be made. Developing consistent operational definitions not only *reduces* the likelihood of miscommunication among persons associated with a particular data collection effort, but it *increases* the likelihood that data collected at different locations and at different times can be aggregated and compared.

Do not assume that everyone defines terms similarly. Select or create clear unambiguous **operational definitions.**

As noted in Chapter 1, the impetus for developing this Guidebook was the realization that while a large amount of data were available, definitions of the entities and other issues (e.g., time, location) were often missing; when definitions *were* documented, they were frequently defined differently in different studies. These limitations made comparisons across data sets impracticable.

The metrics described in this Guidebook should be sufficient for many airport passenger-related data collection efforts. Inevitably, however, you will encounter situations for which the research questions will require you to develop additional metrics. When this occurs, it is sug-

Definition	Example	Number of Passengers
Someone who traveled by air within the previous 6 months.	Mr. Jones made one round-trip flight from Dallas to Florida.	1
A person traveling by air to a destination.	Mr. Jones flew from L.A. to New York and back again.	2
A person getting on a flight.	Mr. Jones flew round trip from Dallas to Orlando, making a connection in Atlanta in both directions.	4

Exhibit 4-1. Definitions of "passenger."

gested that you document *how* the metric was defined. Doing so, and then sharing that definition with others, will increase the likelihood of eventually having a common data base useful to all persons and organizations concerned with improving airport performance.

4.2 Levels of Measurement

Before considering airport-specific metrics, consider another perspective on data—the level of measurement at which the data is collected. Consider this because the level at which you collect data impacts the calculation of sample size. It also might restrict the types of analysis that are permissible.

Based on work by Stevens,[1] social and behavioral psychologists customarily distinguish among four measurement levels or categories: nominal, ordinal, interval, and ratio.[2] (As a pneumonic, you might consider the word *NOIR,* as in a bottle of Pinot Noir wine as shown in Exhibit 4-2). Each measurement level is briefly described below.

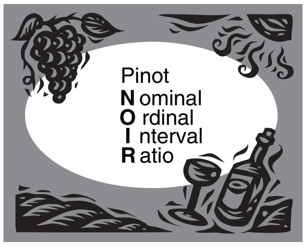


Exhibit 4-2. Use the pneumonic "NOIR" to remember nominal, ordinal, interval, and ratio.

4.2.1 Nominal

These are categorical data that can be counted. For example, if you count the number of business class passengers, you are collecting nominal level data. If, for simplicity, you decided to code men with the number 1 and women with the number 2, these numbers are arbitrary. You might have coded men as 0 and women as 6; the frequencies are meaningful, but the numerical substitute for a label has no quantitative meaning. The late New York Yankee Mickey Mantle wore the number 7. If you went to a sporting event, you might receive a program that matches each number with the name of a team's member. Imagine that you were counting passengers' bags, and you classified them as small, medium, and large. As you noted each bag, you would classify it by size and increase by one the number of elements in the group by which you classified it.

4.2.2 Ordinal

When data are at an ordinal level, the numbers, rather than being arbitrary as with nominal level data, can be meaningfully sequenced.

The number reflects the location of an entity in an ordered sequence. In a marathon race, for example, the number "1" represents the person who was the first to complete the race. Coding passengers by flight class is another example of ordinal level data. For example, first class passengers

[1] Stevens, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.) *Handbook of Experimental Psychology.* New York: Wiley.
[2] Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement Theory for the Behavioral Sciences.* San Francisco, CA: W. H. Freedman and Company.

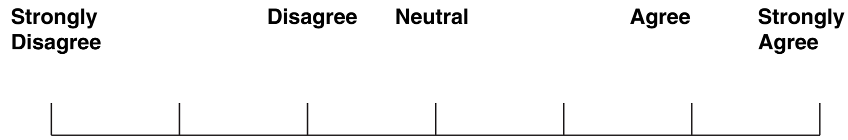


Exhibit 4-3. Ordinal agreement scale.

might be coded as 3, business class passengers as 2, and coach class passengers as 1. The numbers might correspond to the average ticket price. To use another example, an attitude survey of passengers might employ a standard agreement scale for a set of items as shown in Exhibit 4-3. While the order from left to right reflects increasing agreement, there is no justification that the distances among the labels (the intervals) are equivalent.

4.2.3 Interval

These data are measured using a consistent scale (i.e., the interval between the values 1 and 2 is the same as the interval between 22 and 23), but the scale lacks a true zero value. That the value 0° on a Celsius scale is equivalent to 32° on a Fahrenheit scale reflects that the scales are arbitrary in regard to a true zero. As such, it would be incorrect to state that 80° Fahrenheit is twice as hot as 40° Fahrenheit, because there is no absolute zero. To use another example, differences among calendars (e.g., the Jewish calendar, the Chinese calendar, and the Gregorian calendar) reflect differences in starting points. Practically, there is no unambiguous point at which time began, or was equal to 0.

4.2.4 Ratio

Ratio level data have the same properties as interval data, but have one additional property: the presence of an absolute zero. Temperature is an often used example to distinguish interval and ratio scales. If you have $4 in your pocket and someone else has $2, you have twice as much money as the other person. This is ratio data and permits one to generate ratios. In the example given about a marathon in the previous section, were the data collected at an ordinal level, you would have noted each person's name or ID number in the order in which they finished the race. Were you collecting ratio data, you might record the time each runner took to go from the start to the finish line. Someone who took a total of 10 min to complete the race would have taken half as much time, for example, as someone for whom 20 min were needed to run the entire race.

Exhibit 4-4 presents common summary statistics that are permissible for each level of measurement. Explanations for each of the summary measures are presented in Appendix C of the Guidebook.

Measurement Level	Location Measures			Dispersion Measures	
	Mode	Median	Mean	Range	Standard Deviation
Nominal	✓				
Ordinal	✓	✓		✓	
Interval	✓	✓	✓	✓	✓
Ratio	✓	✓	✓	✓	✓

Exhibit 4-4. Appropriate use of summary statistics by measurement level.

4.3 Introduction to Airport Specific Metrics

This Guidebook is largely concerned with *entities* (objects to be counted such as passengers and baggage); *resources* (such as an agent or ticket kiosk); and *processes* (a function performed on an entity by a resource).

In turn, measurements taken for entities and resources relate to *processing rates* and *processing times.* A *processing rate* is the number of entities (e.g., passengers, bags, etc.) processed by a single resource (e.g., metal detector, agent, x-ray machine, etc.) in a given unit of time (e.g., minutes and seconds). A processing rate can only be expressed from the perspective of a resource.

Processing time is the duration of a transaction or other process for a particular entity and resource. Processing time can be expressed from the perspective of either the entity or the resource; these two times may or may not be equal. Average processing time is a summary statistic of the mean time required for all entities to be processed within a given time period.

For processing rate analysis, researchers must distinguish among and define *entities, resources,* and *processes.* As an example, Exhibit 4-5 illustrates entities and resources for the passenger check-in process. The two most common entities in the airport terminal environment are passengers and bags. Elements are processed by a resource—something performing a function or process on the entity. Common resources include self-serve check-in kiosks, ticketing agents, metal detectors and x-ray machines, escalators, and others. Typical processes include checking in for a flight and the screening of a passenger or a bag.

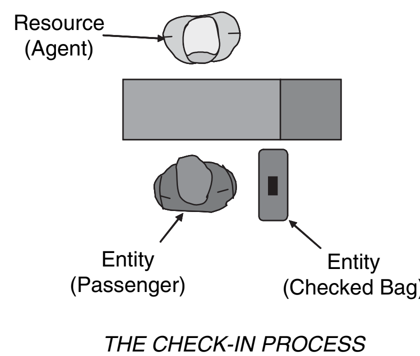


Exhibit 4-5. Illustration of entity and resource in the check-in process.

Failing to operationally define entities, resources, and processes at the start of a research project may result in problems. For example, one might ask if the entity should be defined as a passenger or a travel party. And if the data are collected by party, what definition of party should be used? (Travel parties may be self-defined, while airlines and CBP both have differing definitions.) Should the passenger security screening function be viewed as one overall process or should it be disaggregated between people and bags? This will be addressed in the following sections.

4.4 Defining Entities in the Airport Terminal Environment

As noted previously, it is important to define key units and measurements when collecting data. The two most common entities in an airport terminal are passengers and bags. Each of these is examined below.

4.4.1 Defining Passengers and Travel Parties

Depending on the function being analyzed, passengers are often processed at the airport in a cluster or a group, referred to as a *travel party*. The cost of recording data for both travel party and individuals can be relatively low, while the value of the additional data might be relatively high. In addition, resources usually process entities in parties.[3] The question of how best to define a travel party can be complicated because airlines, CBP, and others have specific (and often different) definitions of what

> Regardless of whether processing rates or other results will be presented in terms of travel parties or by individual passengers, we recommend recording data for both travel party and individuals.

[3] If gross processing rates are being calculated (i.e., counting the total number of entities simultaneously processed by multiple resources over an extended period of time), party size information would be difficult to determine.

constitutes a party. Additionally, the members of the party themselves may have their own, different definition of what constitutes a party. The following sections consider how the definition of party differs by process.

Defining Passengers and Travel Parties at the Check-in (Ticketing) Counter

From the airline's perspective, a travel party is generally defined as those traveling on one unique *Passenger Name Record (PNR)*. These passengers may or may not be related as members of the same family, and the party can be of any size. Airlines might treat a large group (e.g., school class trip) as one party (i.e., process everyone at a particular counter location), even though members were in separate PNRs. It should be noted, however, that it can be difficult to distinguish who is a member of a travel party and who is merely a well-wisher.

Recommendation. Record both the size of the party as well as the number of bags checked.

Defining Passengers and Travel Parties at the Security Screening Checkpoint

Typically, passengers are processed individually at a security checkpoint. An exception is babies or children-in-arms. If you choose to record the amount of time it takes passengers to complete the entire checkpoint process, you will also have to account for the impact of party members waiting for others in the party to complete the process before proceeding to the gate.

Recommendation. Unless you are monitoring passengers as they go through the entire screening process, collect data on an individual passenger basis.

Defining Passengers and Travel Parties at the Federal Inspection Service (International Arrivals) Area

At the FIS, a party is defined as persons who both belong to a single family and who live in the same household. Persons who do not belong to the same family but who live in the same household are processed separately. The party size is easy to determine by counting the number of people simultaneously presenting themselves to a CBP agent during a unique transaction.

Recommendation. Define the party size as the number of people simultaneously approaching a CBP agent at passport control.

By counting the total number of passengers exiting primary inspection in a given period and dividing this by the number of active agents, an overall throughput capacity can be estimated. Collecting this data can serve to validate individual processing times.

Defining Passengers and Travel Parties at Other Terminal Functions

At a gate, aircraft boarding and deplaning rates are typically expressed in passengers per minute. For this reason, and because it is often difficult to collect travel party information solely through observation when monitoring aircraft boarding and deplaning rates, it is considered to not be critical to collect party size information.

At a bag claim device, processing information about passengers is also not expressed in terms of parties. It is difficult, however, to determine whether a person at a claim device is a passenger or a meeter/greeter.

Finally, it should be noted that a group of passengers may have their own definition(s) of a travel party which can conflict with those described above. For example, three unrelated business people traveling together may define themselves as a party of one, two, or three depending on factors such as whether they traveled to the airport together and the degree to which they share the same travel itinerary.

Entity Focus	Recommendation
Airline check-in	Use a definition that matches the airline's PNR. Record number of bags and size of party.
Security check-in	Count individuals, not parties.
Federal Inspection Service	Define as the number of people simultaneously approaching a CBP agent at passport control.
Other	Use professional judgment.

Exhibit 4-6. Recommended travel party definitions.

Recommendation. For those processes that do not lend themselves to a formal definition of party size, use your own professional judgment (assuming you cannot interview the party).

Dealing with Large Groups

Dealing with large parties (typically, those having 10 or more passengers and which receive special processing procedures) is primarily limited to check-in facilities. Procedures for dealing with large parties vary by airline, station, party size, and agent preferences. For example, at the counter of an airline with a hub at the airport, a special counter may be set up for handling large parties, while at a small airport, these parties might be processed along with other passengers. In a recent passenger survey at a large international airport, only about four percent of passengers identified themselves as traveling in parties of 10 or more.[4] It is not clear how this would translate into the percentage of PNRs with 10 or more passengers. Regardless of party size, in the check-in environment, one agent typically processes all the passengers listed on one PNR.

Recommendation. When collecting check-in time data, document not only the size of the party, but also how the airline handled the check-in process for large parties. This practice ultimately will help determine how (or even whether) to include the check-in data for large groups. For airports which frequently experience large parties (e.g., vacation destinations), it is recommended to collect data specifically when large parties are likely to occur and to obtain sufficient data to make suitable inferences regarding performance.

Collecting Data about Passengers as Entities—Conclusion

There is no single best definition of a party that is appropriate across all airport-related processes. It is recommended that you use the standard party definition when observing functions for which a traditional definition is known (for example, at ticketing), and that you use professional judgment in other situations (e.g., at baggage claim). For some functions, party size may not be critical (e.g., aircraft boarding and deplaning rates), and in others, party size information may or may not be needed, depending on the level of analysis being done (e.g., passenger security screening). Regardless of how you define party, clearly document the definition used to avoid ambiguity later. A summary of recommendations for defining parties is shown in Exhibit 4-6.

4.4.2 Defining Bags and Other Items as Entities

Belongings that passengers take with them when they travel represent a second major type of entity to be considered. As with varying definitions of party size, the ways in which belongings are defined, classified, and processed also vary depending on the process of interest.

[4] 2007. Peak Week Survey Results, Hartsfield-Jackson Atlanta International Airport.

Defining Bags and Other Items at the Airline Check-in (Ticketing) Counter

At the check-in counter, a distinction is made between *carry-on* bags (those brought by the passenger into the aircraft cabin) and *checked* bags (those which are stored in the *belly hold* of the aircraft). Since the number of bags checked impacts the overall length of a party's check-in time, it is important to record this information. In fact, many airlines provide "express" kiosks for passengers not checking bags. Determining whether a bag is being checked can be made through observation, as a passenger is observed handing a bag to an agent for tagging. (Note that bags might be checked at the gate; as such, do not record this information at another location—for example at baggage claim—if the purpose is to determine how many bags are being checked on a flight.)

Recommendation. Record the number of checked bags as well as party size. Document baggage information at the source of baggage check-in.

Defining Bags and Other Items at the Passenger Security Screening Checkpoint

At the security checkpoint, the TSA is concerned with items in carry-on luggage and items on one's person. The current procedure (at the time of developing the Guidebook) is that all items in a person's possession are x-rayed while the passenger proceeds through a *metal detector*.

The most common way of defining an item at the checkpoint, and the approach used by the TSA, is to classify it as either a *bag* or a *bin* (which can contain several individual items). Using this definition, any item that is laid directly on the x-ray feed belt is considered an item. For example, if a passenger placed a carry-on bag on the feed belt and then placed two bins down (one containing his jacket and shoes and one containing his laptop, its case, and his cell phone), and finally a basket for his loose change, the total number items (bags/bins) would be four.

Another option is to define an item as anything laid down by the passenger for inspection, regardless of whether it is contained in a bin. Using the same passenger just described, counting this way would result in a total of seven items. Exhibit 4-7 compares how items on the x-ray belt would be counted and grouped using these two methods. In this case, however, clear definitions of what constitutes an "item" must be provided. For example, is a pair of shoes one item or two? Will loose change be counted as one item or will each coin be counted?

By observing the specific, individual items going through the x-ray, it is easier to see the impact of TSA requirements and passenger response on checkpoint capacity as they change over time. For example, the impact of the recent restriction on liquids may not be seen as clearly with the "bags/bins" definition; however, adhering to a practice of routinely counting *all* items, would result in data that would enable one to analyze and explain a decrease in checkpoint throughput resulting from the additional time needed for passengers to divest a plastic bag prior to security,

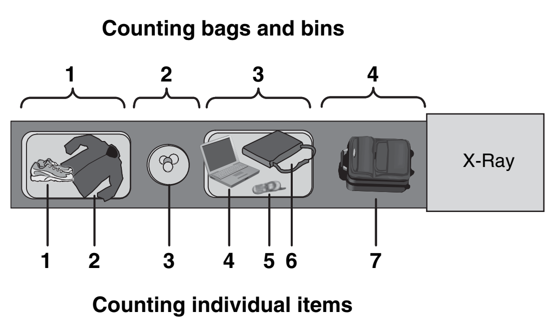


Exhibit 4-7. Two methods of counting x-rayed items at the security checkpoint.

Entity Focus	Recommendation
Airlines	Define as carry-on or checked. Wait until bags are tagged by an agent before classifying the bag as a carry-on element or one intended to be conveyed in the aircraft belly.
Security check-in	Each bag counts as one. The contents of each tray may be counted as one entity, or each item contained within the bin as a single entity.
FIS	Agricultural or customs.

Exhibit 4-8. Baggage definitions.

return it to the carry-on following inspection, as well as the increase in time needed for agents to scan these items.

Recommendation. We recommend that both operational definitions be used when collecting TSA checkpoint x-ray processing rates (i.e., bags/bins and individual items), unless precluded by time or resource constraints.

Defining Bags and Other Items at a Federal Inspection Service (International Arrivals)

At the FIS, passengers enter the passport control area with only their carry-on luggage. After passport control, they then retrieve their checked bags and proceed to declaration and baggage inspection. If passengers have items that need inspection, these items are typically grouped as customs or agricultural, and agents usually provide separate inspection stations for these two groupings—both employing different inspection processes and equipment.

Recommendation. When observing the FIS in the passport control area, use the classification scheme employed by the U.S. CBP. That is, classify items to be inspected into one, or both of the following:

- **Customs** (typically, checking bags for items and their proper declaration), or
- **Agricultural** (looking for contraband animal/vegetable products).

You can estimate the average number of checked bags per party by observing individual parties retrieving bags from FIS bag claim carousels.

Collecting Data about Bags and Other Entities—Conclusion

A summary of recommendations for defining baggage is shown in Exhibit 4-8.

4.5 Defining Resources

This section examines how resources might be addressed on a process by process basis.

4.5.1 Defining Resources at Airline Check-in

There are several resources in the ticketing process, including self-serve kiosks, agents, scales (to weigh bags), document "*spitters,*" and electronic readers. Unfortunately (for the researcher

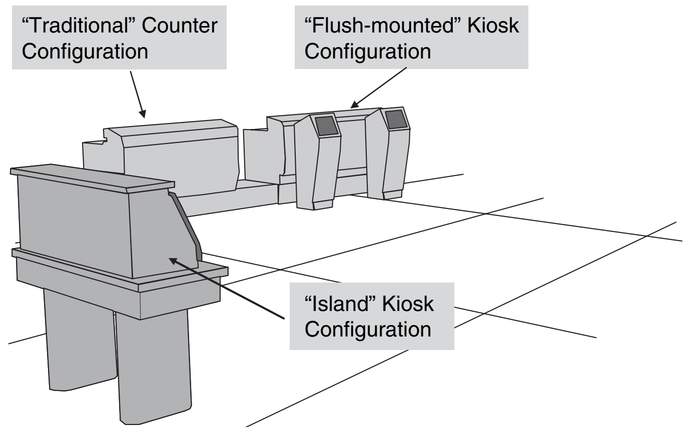


Exhibit 4-9. Ticket counter layouts.

at least), definitions of resources are somewhat blurred in the check-in process, as reflected in the use of *flush-mounted kiosks,* where a passenger may engage either one of the resources (kiosk or agent) or both of them. Many passengers use online check-in and avoid these resources altogether. Exhibit 4-9 displays two types of kiosks in relation to a traditional counter design.

For example, a passenger may check in using the kiosk, ignoring the agent nearby. Or, a passenger may step up to the counter, interact directly and exclusively with the agent, but block the kiosk for use by other passengers. Alternatively, the passenger may begin with an agent, who then walks the passenger through the kiosk process, or, conversely, start at the kiosk but end up needing agent assistance.

In terms of disaggregation, a passenger may check-in at a kiosk and then walk to a designated bag drop area at the counter. As such, the researcher must define the process, a topic considered in Section 4.6.

4.5.2 Defining Resources at the Passenger Security Screening Checkpoint

At a gross level, the passenger security checkpoint consists of lanes in which passengers first divest their belongings and place them on a belt for x-ray screening then proceed through a metal detector to be tested for metals. Traditionally, each lane has consisted of one x-ray machine and one walk-through metal detector. More recently, as increased scrutiny of bags has slowed x-ray throughput, some checkpoints have a pair of x-ray machines serving only one metal detector. These two layouts are shown in Exhibit 4-10.

To help reduce confusion, it is recommended that a checkpoint lane be defined as an x-ray device, not a walk-through metal detector.

Recently, TSA has begun installing *explosive detection trace portals* to screen people. Passengers stand inside portals while puffs of air are blown over them; dislodged particles are then sucked into the top of the machine and analyzed for explosives. Once clear, the passenger can step out of the device. TSA also uses a manual trace detection process on bags. With this process, the outside of a bag is swiped and the sample is placed in a machine for analysis. Finally, TSA also uses physical inspection of people and bags; in this instance, the resource is the agent.

4.5.3 Defining Resources at the FIS (International Arrivals)

The first resource arriving international passengers encounter is the CBP agent at passport control. These agents are typically positioned in booths with a queuing area in front of them. In

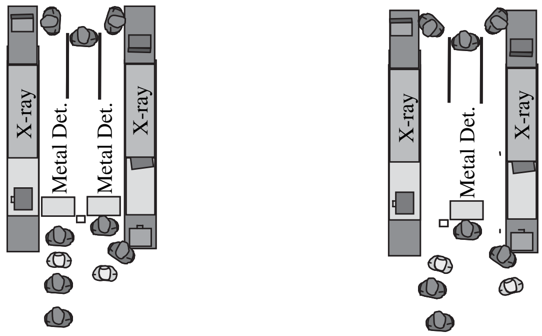


Exhibit 4-10. Two layouts of x-ray devices and metal detectors. Traditional two-lane arrangement with two x-ray devices and two metal detectors (left) compared to a two-lane arrangement with just one metal detector.

most instances, booths are assigned to handle either United States' residents or visitors (who take much longer to process). At busier airports, some booths may be reserved for crew and diplomats. For any formal study of FIS processing rates, it is desirable to collect separate observations for residents and visitors. While there is secondary processing of arriving international passengers, usually of those whose paperwork is not in order, there will probably not be a need to collect observations of these resources due to their infrequency of use.

Once clearing passport control, passengers claim their luggage at carousels. These carousels are considered to be resources for presenting bags for passenger claiming.

Upon claiming their bags, passengers proceed to their third resource, which is primary customs processing. Depending on the volume of passengers, there are usually one to three agents who scan passengers' forms.

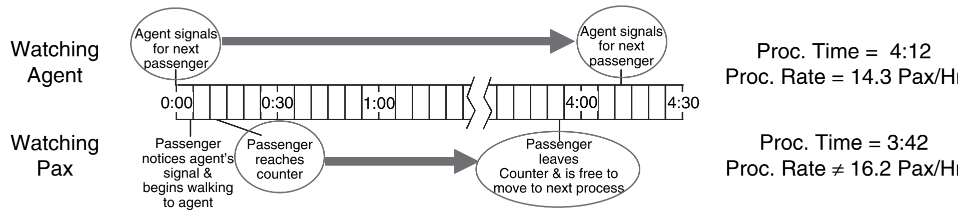
Should passengers have items requiring further inspection, they are directed to a more formal customs and agricultural inspection area. Both of these processes have specific protocols and machinery.

4.6 Defining Processes

A fundamental question is whether one should examine a process from the perspective of the entity or the resource. For many processes, the length of the process can differ depending on which perspective is used.

To take an example, refer to Exhibit 4-11. In preparing this guidebook, one of the Research Team members calculated processing time in two ways: (1) using a resource perspective and (2) using an entity perspective. As shown, a travel party may spend 3.5 min with an agent and then proceed to the security checkpoint. From the agent's perspective, however, there may be an additional 0.3 min of processing time required to complete that transaction and before the agent can process another passenger, and beyond that, it may take an additional 0.2 min before the next party at the front of the queue is aware of the available agent and make its way to them.

At the outset, define whether the perspective of a **resource** or an **entity** will be used.



Source: Observation made at Washington Dulles, fall 2007.

Exhibit 4-11. The processing time from the perspective of the entity and the resource is often different.

To determine processing rates (i.e., the number of entities that a resource can process in a given unit of time), one must define the processing time from the perspective of the resource. As shown in the exhibit, if one were to calculate processing rates using the processing time from perspective of the entity (passenger), one would erroneously conclude that the processing rate was 13 percent higher.

In summary, the importance of documenting how the data are being collected is emphasized. In the following section, the processes and their functional elements most commonly studied at airports are described.

4.6.1 Airline Check-in Related Processing

With the advent of new technologies, the airline check-in process has become disaggregated, decentralized, and amorphous—at least from a data collection perspective. (For more information on this see Section 2.2.6.) For this reason, it is essential that the researcher rigorously define the check-in functions being observed.

In the airline check-in area (and at curbside check-in), several passenger transactions are common. In descending order of frequency, these include the following:

- Checking in (obtaining a boarding pass);
- Checking/dropping a bag;
- Rebooking (due to missing a flight, or having a flight canceled);
- Upgrading/changing seats;
- Purchasing a ticket; and
- Obtaining information or an airline pass to go through security.

In most instances, when observing check-in activity, only the first two processes (checking in and checking or dropping a bag) are typically defined, most likely because they are the most common and the easiest to distinguish through observation.

4.6.2 Passenger Security Screening Processing

As noted previously, there are two broad processes occurring at a security screening checkpoint—the screening of passengers and the screening of their bags. Secondary screening activities include wandings and pat-downs of passengers and hand inspection and trace detection of bags.

Most processing rate studies at the checkpoint focus on overall throughput and use passenger counts as they pass through the metal detector as a proxy for the overall capacity of checkpoint, since passengers do not leave a checkpoint without their belongings.

4.6.3 FIS (International Arrivals) Processing

The key processes at the FIS include passport control and primary customs. Other processes include baggage claim, secondary customs inspection and agricultural inspection. In addition, some passengers must pay a duty. Finally, at some FIS facilities, passengers recheck bags for connecting flights and often must proceed through a TSA checkpoint to enter the domestic portion of the terminal.

4.6.4 Baggage Processing

Processes related to checked bags include their transfer to and from the aircraft and their inspection by TSA.

For outbound bags, a skycap or agent prints and applies a “bag tag” and weighs the bag. The agent then places the bag on a take-away belt for screening by TSA and sorting at the airline’s outbound bag make up area. At that point, they are loaded onto bag carts and transferred to the aircraft, where they are then placed in the plane’s belly hold.

Upon a flight’s arrival, checked bags are offloaded, placed on bag carts and driven to the terminal to be placed on bag claim devices for retrieval by passengers.

CHAPTER 5

Sampling Techniques for Airport Data Collection

This chapter discusses general sampling strategies and procedures, including calculating sample size, minimizing statistical error, and trading between the benefit and cost of various sampling plans. Specific guidance is provided on when to schedule airport data collection events, with a focus on determining the peak period.

5.1 Sampling: Introduction

Sampling is a subject broad in scope and quite often detailed in its technical content. Many books have been written that consider relatively narrow sampling sub-topics. This chapter attempts to limit the topic to a relatively few, key issues. More detailed information is included in the Technical Appendix C.

5.1.1 Populations & Samples

A *population* is often defined as the set of all elements of interest, or a set about which inferences are drawn.[1] In airport planning, you might be interested in populations such as the following:

- Passengers checking in for domestic flights on Monday mornings,
- Passenger security processing times,
- Oversized bags on international flights, and
- Persons using boarding area restrooms during peak periods.

While some research uses a *census* approach, in which each and every element in a population is assessed, this approach is generally neither practical nor efficient for airport planning studies. Rather, *samples*, or subsets of a population, are used to make inferences about the population. For example, an average wait time of a sample of passengers checking in using a kiosk is a stand-in or proxy estimate of the average amount of time across all persons checking in this way. A "good" sample is one in which you can be confident that it is a relatively accurate approximation of the population value. Scientifically conducted political surveys, for example, with sample sizes of 2,000 or 3,000 people can predict within a margin of error of about 2 or 3 points how millions of people will vote in a national election. But more is not always better! A poorly done sample may lead to profoundly inaccurate conclusions about the population of interest. So, what distinguishes a sample that provides a relatively accurate description of a population from one that leads to faulty conclusions? The next section attempts to answer this question, as well as addressing how to select samples of entities and resources, and when and where to sample.

[1] Clapham, C. & Nicholson, J. (2005). *Concise dictionary of mathematics.* NY: Oxford University Press.

5.1.2 Scientific and Convenience Samples

A *scientific sample* employs random selection to control for *bias*. A *convenience sample*, by contrast, is one characterized by the relative availability of the sample elements. For example, assume that your goal is to sample 50 passengers going through security screening, and your plan is to record the amount of time each passenger took to complete screening. The recommended approach would be to use some random method of selecting the 50 passengers you observe, such as systematic random sampling (see Section 5.4.2). To save time, however, you decide to time the first 50 people in line. The following are biases this method might introduce:

> To reduce bias, use tactics that maximize randomness.

- The passengers at the front of the line likely arrived earlier than those at the end of the line, and, as such, might be different in some meaningful way, such as being better prepared, or, by contrast, people who do not often fly.
- If the people at the front of the line are more experienced travelers, they may proceed more quickly, knowing the "drill."
- Vacationers who travel infrequently, however, and arrive well in advance of the flight's departure, may be less savvy and, as such take more time going through security.

5.1.3 Defining the Population

While defining the population of interest might appear to be so pedestrian a subject as not to warrant formal attention, experience shows that this critical first step is often not properly done. It can be tricky as well. For example, an airport director charges you with the following task:

"Find out what the average transaction time is for business class check-in on domestic flights during weekday mornings."

> Defining the research question helps to define the population to be studied.

While the task seems reasonably straightforward, it begs a fundamental question, i.e., what is driving the request for the data? Is the director interested in assessing performance against quality goals? Is he or she considering installing a new technology for check-in processing, or might the underlying reason be a question about the utility of adding staff during that time period? Determining the specific impetus might impact how the population from which the sample is drawn is defined. It might impact, for example, where data will be collected, whether parties or individual passengers will be observed, etc.

The director's charge also suggests several potentially meaningful differences. The wording implies that business class differs from other classes of flight, domestic differs from international, weekdays differ from weekends, and mornings differ from other times of the day. If these distinctions are meaningful, the implication is that there is less variation in transaction time within each of these groups than there is between or among the groups. If this is the case, it argues for using a *stratified sampling* approach. The benefit is that it enables the researcher to make specific generalizations for each group such as, for example, being able to comment on the average transaction time for business class passengers.

5.2 Sampling Strategies

5.2.1 Stratified Sampling

Exhibit 5-1 is an example of the impact of stratified and non-stratified sampling on assertions that can justifiably be made. This assumes that stratification has been done prior to data collec-

Stratified	Not Stratified
We are 95% confident that the average transaction time for business class passengers on weekday domestic flights is between 2.8 and 3.2 min.	We are 95% confident that the average transaction time for passengers on weekday domestic flights is between 2.8 and 3.2 min. The average time recorded for coach and business classes were, respectively, 3.5 and 2.25 min.

Exhibit 5-1. Example of appropriate assertions under stratified and non-stratified samples.

tion. Note that the statement associated with the stratified sample contains a specific inference about business class passengers, whereas, inference from the non-stratified sample is to *all* passengers flying during the stated times. It is legitimate to describe subgroups in a non-stratified sample, but not to draw from them generalizations to a more narrowly defined population, assuming no post-stratification.

5.2.2 Simple Random Sampling[2]

Stratified sampling within the context of defining a population has been addressed. Before addressing the particulars of *how* to sample, several sampling strategies of potential value when collecting airport processing data will be briefly described. Each strategy presented is a form of *probability sampling,* i.e., each element that can be sampled has a specific chance or likelihood of being selected.[3] Within those parameters, it assumes that being selected for inclusion in the sample or not being selected is a *random event.*

In a simple random sample, each element has the same likelihood (probability) of being selected as all other elements. If 100 bags, for example, were on a conveyor belt, each bag would have a 1/100 chance of being included in the sample.

5.2.3 Stratified Random Sampling

As discussed in the previous section, the population is sorted into "subpopulations." In essence, this approach treats each subpopulation as population unto itself.

5.2.4 Systematic Random Sampling

This is a variation on simple random sampling which is practically more manageable in the field. Every kth element, separated by the same number, is selected. For example, if you wished to select a sample of 100 bags from a population of 1,000 bags, you would divide 1,000 by 100 to get 10. This means that every 10th bag would be selected. To avoid bias, one randomly selects a starting point. If the number 32 were randomly selected as the starting point, the first bag included in the sample would be the 32nd bag in the 1,000-bag population. The next would be

[2] Some of the content of this section is based on Sahai & Khurshid's (2002) *Pocket Dictionary of Statistics.*

[3] Technically, one does not sample from a population per se, but from a *sample frame.* For example, a schedule of flights for a given airport at a given time is a sample frame from which selected flights might be selected. The use of sample frames is unavoidable, but they can introduce error when, for instance, the actual flights deviate from the scheduled flights. When reference is made to sampling from a population, we acknowledge that the sampling is actually from a sample frame.

bag 42, then bag 52, and so forth. In October 2008, Washington, D.C., Metro announced that Transit Police would begin randomly searching riders' bags using such a system.

5.2.5 Cluster Sampling

Clusters are groupings of units, such as flights, gates, or terminals. Cluster sampling is used when it is difficult to define a formal list of elements (for example, a list of all passengers arriving at an airport in a given time period) from which one could draw a random sample. However, since each arriving passenger must enter the airport at a gate, one could group passengers into clusters by gates. In this instance, a random sample of gates would be drawn, and the characteristics of all arriving passengers exiting those gates would be collected. This is known as a *one-stage* approach. (Additional information about cluster sampling is included in Appendix C).

5.2.6 Time Periods as Sampling Units

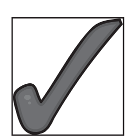
As noted previously, in order to assign a measure of data reliability to a sample, each unit must have an equal (or at least known) chance of being included. Since there are often many instances where a list of items cannot be obtained, another method might have to be employed. The following case study is an example of a survey where a sample frame of flights may not be the optimal way of efficiently collecting data.

Time Periods as Sampling Units—Case Study

At one airport, there has been a desire to estimate the average time between a flight's arrival at the gate and the appearance of its first bag on the bag claim carousel. This data is to be collected in such a way as to make inferences not only for the airport as a whole, but by airline and concourse. The first year, a stratified random sample of arriving flights was drawn and data collectors were assigned to "meet" the flight to record its arrival time and then proceed to bag claim to watch for its bags at the assigned carousel. While this method appeared to be a sound approach, in practice, it resulted in many missed flights and resurveys. This is because flights often reached the gate before their scheduled arrival time, much later than their scheduled arrival time, or at a gate different from the one listed on the monitor.

To overcome these challenges, a different sampling approach was used—that of making time periods a sampling unit. In this example, a 7-day survey period was divided into 126 1-hour time periods (7 days × 18 hrs per day—for practical reasons, only hours that had at least two percent of the day's arrivals were included) producing a sample frame. Then using proportional sampling, the probability of drawing a particular 1-hour time period during the week would be the same as the percentage of arriving flights scheduled to arrive during that same 60-min period. Separate sub-frames for each concourse/airline were drawn, creating a stratified sample. A random sample of concourse/time periods was then drawn. During the selected hour, the data collector could proceed to the concourse and wait for the next aircraft to come in. The data from that flight would then be included in the database. The result of using this method was that there was significantly less makeup work due to missed flights.

5.3 Sampling Steps



The following is an overview of the principal steps (see Exhibit 5-2) involved in designing a sample. While the steps are presented in sequence, experience shows that the steps are often iterative and the process dynamic.

Sampling Design Step	Examples	Comments
1. Define the population	Domestic passengers checking-in during peak weekday mornings in autumn.	If you wish to generalize to specific groups, consider using a stratified sample. (See Section 5.1.3.)
2. Define sampling units	Individual passengers who obtain a boarding pass.	A sampling unit is that which is counted or measured.
3. Define what is to be reported.	Count of the number of passengers obtaining a boarding pass. Average time to complete a transaction.	When reporting frequencies and averages, different formulas are used. (See Section 5.4.3.)
4. Identify the level of error you can tolerate.	The sample count should be plus and minus 5 persons of what is the true (and unknown) population count. The sample average should be plus and minus 20 seconds of the true population average.	This is a measure of precision. It reflects the width of the confidence interval. (See Section 5.4.3.)
5. Specify the desired level of confidence.	We wish to be 95 percent confident.	By convention, confidence levels range between 90% and 99%, with 95% being the most common. (See Section 5.4.5.)
6. Estimate the variance in the population.		Population data always vary to some extent. (See Section 5.4.2.)
7. Estimate the required sample size.		This is an iterative process. While one cannot change the observed or estimated variance in a population, one can manipulate confidence levels and tolerable error.
8. Select sampling strategy	Simple random sampling. Stratified random sampling.	

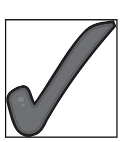
Exhibit 5-2. Steps to sample design.

5.4 Error

Error in research is the difference between what actually is (i.e., its true state), and what it is reported to be true. While error is inevitable and impossible to eliminate, some error can be reduced.

Random error is inevitable, but systematic error can be reduced through careful planning and execution.

Broadly, there are two types of errors: *random error* and *systematic error*. Random error is generally attributed to chance alone, whereas systematic error reflects inadequacies in how data are collected. If a regularly traveled flight took an average of 75 min, for example, sometimes it will take a little longer and on other occasions it will take less time. To an extent, the difference between the actual duration of the flight and the average (what is expected) is due to chance.



Consider doing the following when first collecting data:

1. During the pilot (see Section 5.4.2), collect a small amount of data (e.g., 20 observations);
2. Perform a rough calculation of the mean and standard deviation. (If you plan on reporting frequencies see Section 5.4.3).

3. If the standard deviation appears to be relatively large (as a rule of thumb, consider it large if it is 10 percent or more of the average value), ask yourself if there is something similar about those observations that were particularly large or small. For example, do the longer transaction times seem to be associated with party size or number of checked bags? Did the relatively brief transaction times seem to be related to persons apparently traveling on business? Granted, this distinction may not be verifiable, but if you think there may be a pattern, or something that helps explain variation, incorporate it into your sampling plan by recording the distinguishing information. Using the examples above, for example, might suggest that party size be recorded, while the number of checked bags might be ignored, or that the suspected purpose of travel (e.g., business or pleasure) be documented.
4. If the questions just asked both had affirmative answers, for example, consider capturing one more piece of data—whether the passenger appeared to be on business or pleasure travel.

5.4.1 Systematic Error

Systematic error can be attributed to one or more causes, and often those causes are within the researcher's ability to detect and correct. Exhibit 5-3 on the following page presents common reasons for systematic error, and proposes ways of avoiding such errors.

5.4.2 Calculating Error

Variance is a measure of the average dispersion around a mean.[4] How does one estimate the population variance? Sometimes, but not usually, the population standard deviation is known. Square it and you have the variance. When this unusual situation does not exist, there are several approaches to developing the estimate.

1. A previous similar study may have been conducted, and the variance calculated in that study could be used.
2. A small pilot study could be conducted, and the variance calculated from those data could be used.
3. If nothing is available, select a conservative value based on judgment and a very limited amount of data. For example, if the average time passengers wait in line is 4 min, you could conservatively guess that the standard deviation is the relative large value of 3 min. If this approach is taken, you can test your guess by calculating the actual standard deviation and variance for the real data.
4. Another approach, as described by van Bell is to make approximately 15 observations and then divide the range by the number of observations, an approach that is sufficiently robust to accommodate distributions both normal and substantially different from normal, (e.g., uniform).

5.4.3 Variance

Steps 4, 5, and 6 in Exhibit 5-2 concern the amount of error one is willing to tolerate, the *confidence* one wishes to have, and *variance* in the population. Each of these issues needs to be considered to calculate sample size. In particular, consider if and how the data's level of measurement influences what is and is not permissible to calculate, and whether you are using a point estimate or a confidence interval approach. These topics are considered in some detail in Appendix C.

> The standard deviation is the square root of the variance.

[4] Additional information on the topic of variance is included in Appendix C.

5.4.4 Confidence Levels & Hypothesis Testing

Researchers often specify a *research hypothesis,* or a statement of the way the researcher *believes* things to be, and a *null hypothesis.* The null hypothesis is what is tested. The null hypothesis essentially asserts that there is no relationship between the variables. For example, if you suspected that, on average, people spent less time in check-in using kiosks rather than transactions with agents, you might state the hypotheses as follows:

Research Hypothesis: On average, people spend less time in kiosk transactions than in agent transactions.

Null Hypothesis: On average, people spend the same amount of time or more time in kiosk transactions than in agent transactions.

A given confidence level reflects the probability of rejecting the null hypothesis when in fact it is true. It is customarily set at 0.95, meaning that the researcher is 95 percent confident in a decision to reject the null hypothesis. Rejecting a null hypothesis which is in reality true is referred to as *Type I error.*

Exhibit 5-3 presents an example of how these error types might arise in an Airport environment. Here, the decision to expand an airport facility is based on anticipated growth in demand. One cannot know with certainty what will happen in the future, the "true" state, but one can hypothesize and decide that growth will or will not occur. In two of the four possible combinations, the decision is accurate. The other two combinations, however, represent error. In this instance, a *Type I* error occurs when it is hypothesized that there *will* be growth, but it does not happen. The result is an increase in debt but not in revenue. If there *is* growth, however, but the decision is that growth will not happen and, as such, the facility is not expanded, the consequences might be missed revenue, a decrease in customer service, etc. *Power* is the other side of the confidence level issue. Power reflects the probability that one rejects the research hypothesis when indeed it is true. This is also known as a *Type II* error.

5.4.5 Tolerable Error

This is the amount of error one is willing to tolerate in the sample estimate. For example, you might assert that the true average time spent in kiosk transactions is equal to the calculated sample mean, plus or minus a given amount of time. If the time were equal to one minute, for instance, you could state, "95% confidence that the true population mean is equal to the sample mean of 4.5 min plus or minus one min."

	The "True" State	
Decision	**No Growth**	**Growth**
No Growth	Correct Decision (Do not expand)	Type II error (missed revenue, missed opportunities, customer service is a choke point, etc.)
Growth	Type I error (straddled with debt & no revenue)	Correct Decision (Expand)

Exhibit 5-3. Type I and type II errors, aviation example.

5.4.6 Random Error

There is always some inherent amount of unexplained variation in what is observed. Not every person takes precisely 4.5 min to complete a transaction, nor will every registered Democrat vote for a Democratic candidate in an election. While one can't control for this variation, it is important to estimate it when planning for sample size.

5.4.7 Trading Benefits and Costs

In a perfect world, resource constraints would not be a problem. As such, one could minimize the possibility of making an error in incorrectly rejecting the null hypothesis, or incorrectly failing to reject the null hypothesis when it is indeed true. Given the reality of a resource-constrained world, what criteria might you use to weigh the benefits of a larger sample against its cost?

It is proposed that you weigh the potential consequences of each type of error. Whereas, for example, a five percent risk of a key part failure would not be acceptable, an error rate of five percent is assessing passenger satisfaction might be acceptable. Unfortunately, there are no absolute rules. Customarily, an error of five percent is acceptable in many of the sciences. Formulas and additional detail are included in Appendix C, and there are numerous websites that provide basic sample size calculators.

5.5 Calculating Sample Size[5]

This section offers guidance on *how* to calculate sample size for reporting proportions and means. As noted elsewhere, the approach for calculating sample size for simple random samples may, under certain conditions, be used for systematic and stratified sampling. In particular, for systematic random sampling, the order of elements from which the sample is drawn must be random, and, assuming poststratification is not used, each stratum is treated as a separate population.

The flow diagram presented as Exhibit 5-4 identifies some key questions and issues associated with the sample design.

The formulas presented in this section will yield appropriate sample size calculations for both proportions and means (see Exhibits 5-5 through 5-8). They are appropriate, as well, for systematic random samples assuming that the sampling elements from which the sample is drawn are randomly ordered. Situations in which one cannot assume such random ordering is beyond the scope of this guidebook. One text describes it as ". . . a formidable problem."[6]

5.5.1 Calculating Sample Size When Reporting Proportions

When proportions are to be reported, you need to estimate the incidence of the event in the population, specify the level of error you are willing to tolerate, and identify the desired confidence level.

Example

In planning for redesign of a passenger check-in area, an airline is considering redesigning the space to better meet the demands of an increasing number of international passengers.[7]

---

[5] The formulas presented are for determining approximate rather than exact estimates insofar as they do not require knowledge of the size of the population (N). The formulas presented assume that the population is sufficiently large and can be treated as an *infinite* population. When working with relatively small populations, a finite correction factor needs to be used to compensate for not replacing the sampling units.

[6] Levy, P. & Lemeshow, S. (1999). *Sampling of populations: Methods and applications.* NY: Wiley.

[7] For the purposes of the example, assume that these data are not available elsewhere.

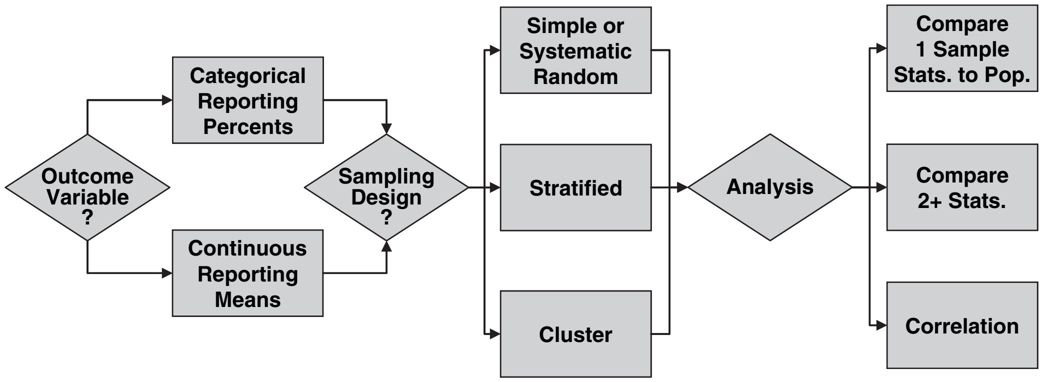


Exhibit 5-4. Flow diagram.

$$n = \frac{Z^2 p(1-p)}{e^2}$$

n = sample size.
Z = the level of confidence desired (see Exhibit 5-6).
p = the estimated proportion of what you wish to sample that is present in the population.
$1 - p$ = the estimated proportion of what you wish to sample that is not present in the population.
e = the level of error one is willing to tolerate.

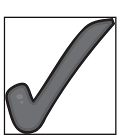
Exhibit 5-5. Formula to determine sample size when reporting proportions.

Confidence Level	*Z value*
90%	1.645
95%	1.96
99%	2.58

Exhibit 5-6. Confidence level and corresponding Z values.

$$n = \frac{(1.96^2)(0.30)(0.70)}{0.05^2}$$

Exhibit 5-7. Example of determining sample size for reporting proportions.



If the proportion of international passengers exceeds 30 percent, renovation will be initiated. The project sponsors indicate that they would like to be 95 percent confident that the proportion observed in the sample is +/– 0.05 (five percent) of the true population proportion.

In Exhibit 5-7, the value 1.96 is used to represent the 95 percent confidence level. (Referring to Exhibit 5-6, note that where the desired confidence level is set at 99 percent, the value 2.58 would be used.) The proportion of international passengers that would trigger renovation is 30 percent. Hence, p is set at 0.30; $1 - p$ therefore must be 0.70. Finally, the numerator represents the tolerable error. Solving the equation suggests a sample size of 323.

Using this sample size would permit the researcher to assert, "we are 95 percent confident that the true proportion of international passengers is plus or minus 5 percent of the proportion observed in the sample."

Had the confidence level been set at 99 percent, the required sample size would have increased to 559.

When calculating sample size for a proportion, do the following:

1. Trade off desired rigor and cost constraints by modifying the confidence level and/or the amount of tolerable error. First, change the error level and then, if necessary, change the confidence level.
2. If you have no basis for estimating the proportion, set p as 0.50.

$$n = \frac{Z^2\sigma^2}{e^2}$$

n = sample size.
Z = the level of confidence desired.
σ^2 = the Greek letter sigma. Sigma squared is referred to as variance. Similar to p and $1 - p$ in the formula for proportions, the variance is a measure of dispersion (random error), or unexplained variance in the distribution.
e = the level of error one is willing to tolerate.

Exhibit 5-8. Sample size formula for reporting means.

5.5.2 Calculating Sample Size When Reporting Means (Averages)[8]

When means are to be reported, the estimated variance, rather than the estimated proportion, is used in the calculation. Use the formula in Exhibit 5-8 when conducting a simple random sample with the intent of reporting averages. The structure is identical to the formula in Exhibit 5-5.

Example

An airport is interested in determining the average transaction time for domestic passengers interacting with an agent during check-in. The project sponsors would like to be 95 percent confident in the estimate (sample average) obtained from the sample. In addition, they would like the estimate to be within five seconds (error) of the true population mean.

Given:

$Z = 1.96$
$e = 5$ seconds
$\sigma = 35$ seconds

The sample size is calculated to be 188 transactions.

Thus far, essentially the same information needed for the proportion formula exists. What is missing is a measure of variation. In the proportion formula, variation was operationalized as an estimate of the proportions in the population. When using the formula for calculating sample size when averages are to be reported, another estimate must be used, namely, the variance. For the moment, where the variance estimate comes from will be ignored.

Appendix C contains information for calculating sample size when assessing change, examining relationships, situations involving varying sample costs, and when testing for differences in averages.

5.6 Determining When to Schedule a Data Collection Event

One characteristic of a good sample is that it is a fair approximation of the population it represents. This pertains not only to entities but to time periods. For example, were you to choose the late afternoon and early evening on December 23 to collect data on passenger processing rates at Chicago O'Hare International Airport, it would obviously give you very different results compared to those obtained on a Monday morning in September.

[8] vanBell, G. *Statistical rules of thumb.* (2002), NY: Wiley.

Airport passenger-related activity varies by season/month of the year, day of the week, and time of day. For this reason, it is important to give consideration to the scheduling of your data collection event(s). Generally, it is best to collect data during a *peak period* for the following two reasons:

1. Most planning is focused on providing adequate peak hour capacity;
2. A queue and constant demand for resources are required to calculate processing times.

While there is some variation in defining peak periods across functions, the commonalities are much stronger. As such, determining peak periods from a general perspective is considered, noting, as appropriate, differences unique to specific functions.

As shown in Exhibit 5-9, the aviation industry has many sources of data that, while not perfect, are considered useful for determining peak periods of activity.

Independent of breadth of time for which you are interested in determining a peak period, the following are basically only two sources to go to for help:

- Data—presumably "objective" in how it was captured, and likely collected for another purpose; and
- People—from whom subjective insights and judgments are solicited.

Sources to Determine Peak Periods

Source	Peak Month	Busy/Peak Day	Peak Hour
Airport Mgt. Statistics			
--Landing Reports	X		
--Monthly Activity Stats.	X		
--Parking Counts (1)	X	X	X
--Roadway Traffic Counts (2)	X	X	X
U.S. DOT T-100	X		
Official Airline Guide (3) (OAG)	X	X	X
TSA Wait Time Data (4)		X	X
Informant (5)			
--Local Airline Station Mgrs.	X	X	X
--Local TSA Manager	X	X	X
--Local US CBP Manager	X	X	X

Notes: (1) Primarily reflect travel patterns of residents.
(2) Roadway counts can be affected by other activity centers (e.g., nearby construction area or employee parking lot).
(3) Must also account for load factor variation and percent originations.
(4) From TSA website (www.tsa.gov/travelers/waittime.shtm); high wait times may not only be due to peak demand but insufficient staffing.
(5) Informants can provide both proprietary, quantitative records and qualitative/anecdotal input; the latter should be used with caution.

Exhibit 5-9. Sources determining peak periods of activity.

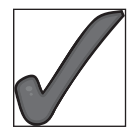
People, the latter source, are sometimes formally referred to as *informants:* persons with knowledge specific to a unique situation. Consider informants as a potentially valuable source of data. (For more guidance on using informants, see Section 5.8.)

The next section provides guidance on how to establish an airport's peak month, busy day, and peak hour, including which data sources to use and how to use them effectively to determine peak periods for your particular area of interest.

5.6.1 Identifying the Peak Month

In general, most functional elements that are typically examined for an airport processing rate study tend to have a common peak month. (This should be verified for your airport, however.) The following sections provide discussion of the usefulness of various databases to determine the peak month.

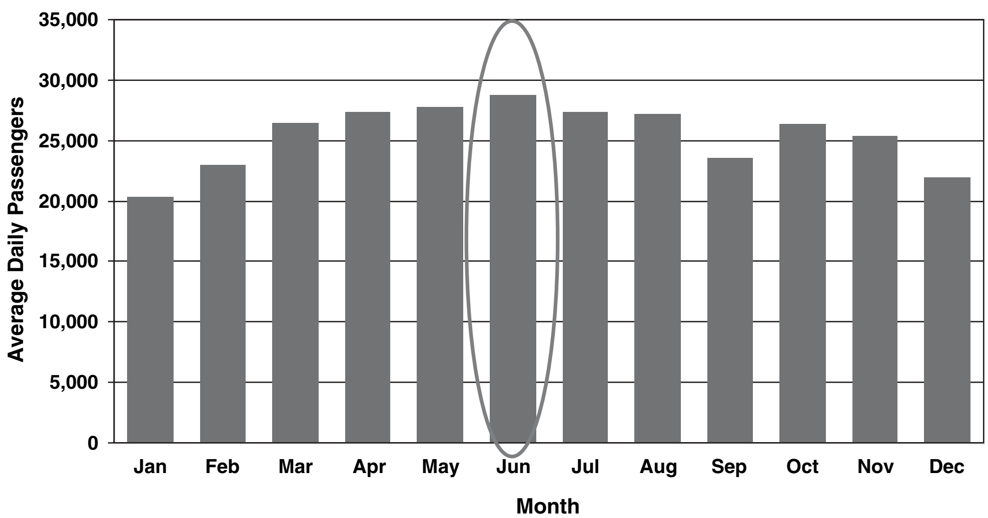
Airport Management Statistics



Nearly every airport with scheduled passenger service keeps records of monthly passenger and aircraft activity. These reports, most commonly prepared by the finance department, are available from airport management or from the airport's web-site. Collected by the airlines, they include monthly data related to passenger enplanements and deplanements, freight and mail tonnage, and aircraft landings and takeoffs. Domestic and international activity may be collected separately. It is recommended to use these statistics for determining the peak month for overall passenger and aircraft operations activity using the process described below.

To determine the peak month, review monthly activity data for the most recent 3-year period. In general, the peak month occurs in the summer. For some markets, however, particularly vacation destinations such as Orlando, the peak may occur in the spring or other time of the year. Exhibit 5-10 shows average daily enplanements by month at Washington Reagan National Airport for 2007. As shown, June was the peak month.

The following is an evaluation of other various sources for peak month data.



Source: Metropolitan Washington Airports Authority; HNTB analysis.

Exhibit 5-10. Average daily departing passengers by month at Reagan National Airport, 2007.

U.S. DOT Form 41 Data/T100

These data report monthly traffic data by carrier and city-pair and can be used to determine peak month activity.

Origin-Destination Data

These data represent a 10 percent sample of aviation activity (both passengers and cargo). The benefit of Origin-Destination (O&D) data is that it excludes connecting traffic. However, there are several disadvantages. First, while they represent actual passenger origins and destinations, they are prone to error due to inaccurate/incomplete reporting by the airlines; however, this can be largely overcome by obtaining the data from a third party source that rectifies and "cleans" the data—although at a cost. In addition, O&D data are only available by quarter, not by month, so they can only be used to get a general sense of when an airport is busy.

Official Airline Guide

The Official Airline Guide (OAG) is the sole database of all scheduled commercial airline flights worldwide. Typical *OAG* data include published and scheduled operator (airline), aircraft type, seats, scheduled arrival and departure times, origin and destination (and downline stops), and effective/discontinued dates. A sample raw schedule pull from OAG is shown in Exhibit 5-11.

OAG data can only be used to examine trends in *scheduled* activity and only of flights and seats. While in general, airlines schedule more flights and seats during busy periods, competitive pressures, aircraft positioning requirements, and other factors can sometimes distort when peaks occur. OAG data can also be useful for very short-range forecasts of activity—up to six months. Beyond that, the level of uncertainty increases significantly.

The cost of a "schedule pull" varies in a largely linear fashion by the number of fields included and the number of records, so that a schedule pull for a large commercial airport can be significantly more expensive than for a less busy airport, costing several thousand dollars. For additional information, contact an OAG customer service agent or sales manager at (630) 515-5300 or at custsvc@oag.com. Note that the OAG is potentially useful in identifying peak days and hours as well.

Other Sources

The airlines, TSA, and CBP collect passenger activity data and therefore have historic information to help determine peak periods or to serve as a back-check for surveyed data. However, unless the study is being conducted for the particular entity, it is unlikely they will share proprietary or security-related information.

Finally, there is usually an abundance of anecdotal information that comes from talking with airport staff, on-site government agencies, and tenants (informants). As noted elsewhere, be cautious and seek confirmation or disconfirmation from multiple persons.

carrier	fltno	depapt	arrapt	arrctry	deptim	arrtim	days	genacft	inpacft	seats	domint	efffrom	effto	sad	acft_owner
DL	5113	ABE	ATL	US	0600	0800	1	CRJ	CRJ	50	DD	20080707	20080707	OH	OH
DL	4171	ABE	ATL	US	0630	0836	1234567	CRJ	CRJ	50	DD	20080616	20080706	EV	EV
DL	4171	ABE	ATL	US	0630	0836	1234567	CRJ	CRJ	50	DD	20080708	20080727	EV	EV
DL	4171	ABE	ATL	US	0630	0836	1234567	CRJ	CRJ	50	DD	20080729	20080818	EV	EV
DL	4171	ABE	ATL	US	0650	0856	1234567	CRJ	CRJ	50	DD	20080819	20080901	EV	EV
DL	4915	ABE	ATL	US	0650	0856	1234567	CRJ	CRJ	50	DD	20080902	20090615	EV	EV

Exhibit 5-11. Sample OAG "schedule pull" database.

5.6.2 Identifying Peak Days

At the daily level, a distinction should be made between departing and arriving passengers. The researcher should therefore decide whether his or her area of focus is largely driven by departing passengers (e.g., airline check-in and security), arriving passengers (e.g., FIS or bag claim), or a combination of the two (e.g., curbside activity). Exhibit 5-12 shows day-to-day local departing and arriving passenger activity at Hartsfield-Jackson Atlanta International Airport over a one-week period in late June 2008.

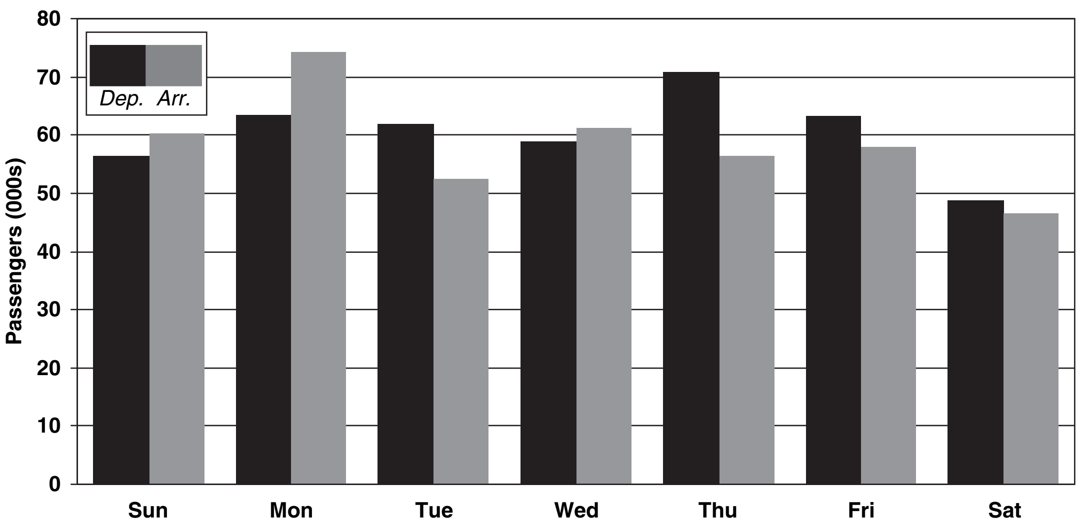
As shown, the peak day for departing passengers was Thursday, while the peak day for arriving passengers was Monday.

Determining peak days of the week is more difficult than determining peak months because most readily-available sources do not typically keep data on a daily basis. The data above were obtained by a one-week survey at the entrances and exits to the airport's secure areas. In most instances, time and resource constraints will preclude you from obtaining this type of data.

The optimal way of obtaining daily counts of passengers would be to obtain historical data from local airline station managers and/or local TSA or CBP officer. But, as noted in the previous section, they may be unwilling to provide this data. They may be willing, however, to review their data themselves and provide general guidance.

An examination of OAG data by day can be useful; however, as noted previously, there are competitive and operational factors that tend to mute day-to-day schedule variations. (For example, on a typical weekday, airlines will have to schedule the same number of departing flights as arriving flights, even though on any given weekday there may be more outbound passengers than inbound passengers, historically.)

One can also examine TSA wait time data, which is available on their website. One can choose an airport, day of week and hour of the day, and the average and maximum wait time at various



(1) Counts of people entering and exiting secure side of airport, respectively.

Source: Hartsfield-Jackson Atlanta International Airport 2008 Peak Week Survey; HNTB analysis.

Exhibit 5-12. Counts of originating and terminating passengers by day of week in June 2008 at Hartsfield-Jackson Atlanta International Airport.

checkpoints will be displayed. Assuming that the days with the highest wait times are the busiest, a busy day can be identified. (It should be noted, however, that high wait times may also be the result of insufficient staffing at the checkpoint.)

5.6.3 Defining Peak Hours

As with peak day activity, peak hour activity at various functional elements will be largely determined by whether it serves primarily departing passengers, arriving passengers, or both.

Exhibit 5-13 shows hourly counts of local departing and arriving passengers on a Wednesday in early July 2008 at Hartsfield-Jackson Atlanta International Airport.

As shown, local departing passenger activity peaked early in the morning (between 6 a.m. and 10 a.m.); a secondary departing passenger peak occurred in mid-afternoon. For arriving passengers activity was generally low until about 2 p.m. The peak for arriving passengers occurred between 7 a.m. and 9 a.m.

Sources of Hourly Data

As with day-of-week data, the optimal sources for hourly data include local airline station managers and TSA. Again, however, there will likely be a reluctance to share this type of data, leaving the researcher with less-than-optimal sources.

One could examine TSA checkpoint wait time data, which are available on their website; however, as noted previously, delays may be due to insufficient staffing.

An OAG schedule pull of departing and arriving aircraft and seats will be of some benefit. A caution should be noted here in that, for airports with significant amounts of connecting passenger activity, an hour-by-hour of summary of seats will show multiple peaks across the day, directly corresponding to the "banks" of flights operated by the hub airline. This will make it difficult to pick the bank or banks with the most local passengers (if the element of study is in fact primarily affected only by local passengers).

Nevertheless, as there are really no other readily-available data sources, an OAG schedule pull is often the best source.

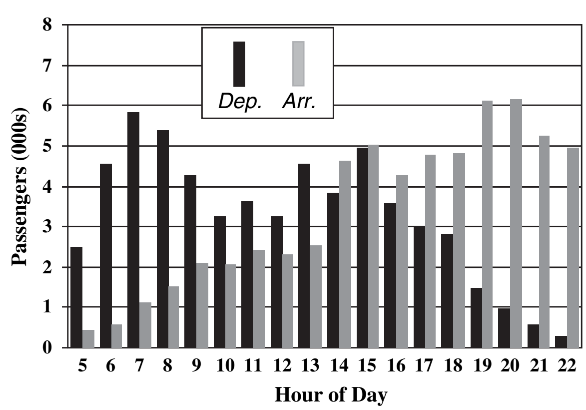


Exhibit 5-13. Hourly distribution of local departing and arriving passengers at Hartsfield-Jackson Atlanta International Airport on Wednesday, July 2, 2008.

Using an OAG Schedule Pull to Determine the Peak Hour for Various Functional Elements

At the outset, it must be understood that the published times listed in an OAG schedule pull are the times when a flight is scheduled to depart or arrive. If the peak hour for departing seats is between 8 a.m. and 9 a.m., do not assume that this is when you should be collecting data at ticketing, for example, because passengers will already be at the gate boarding their flights. Rather, adjustments must be made to the OAG schedule to anticipate when particular functional elements will peak in activity as described below.

Adjustments to Departing Seat Schedule

Exhibit 5-14 shows the cumulative percentage of local passengers arriving at a terminal by hours and minutes before their flight is scheduled to depart. The information comes from actual survey data collected from one small, regional east coast airport and one large, international east coast airport in the fall of 2005. Patterns at other airports may vary from those presented here.

Note that the time interval is strongly influenced by whether a passenger is traveling on a domestic or international itinerary and if the flight is leaving early in the morning; therefore, separate exhibits are provided for these scenarios. In addition, each graph also shows separate curves for data gathered at a large international airport and a small/regional airport; a median curve (dotted line) is also provided. In general, passengers tend to allow less time for making their flight early in the morning, when traveling domestically, and at small airports. Conversely, passengers

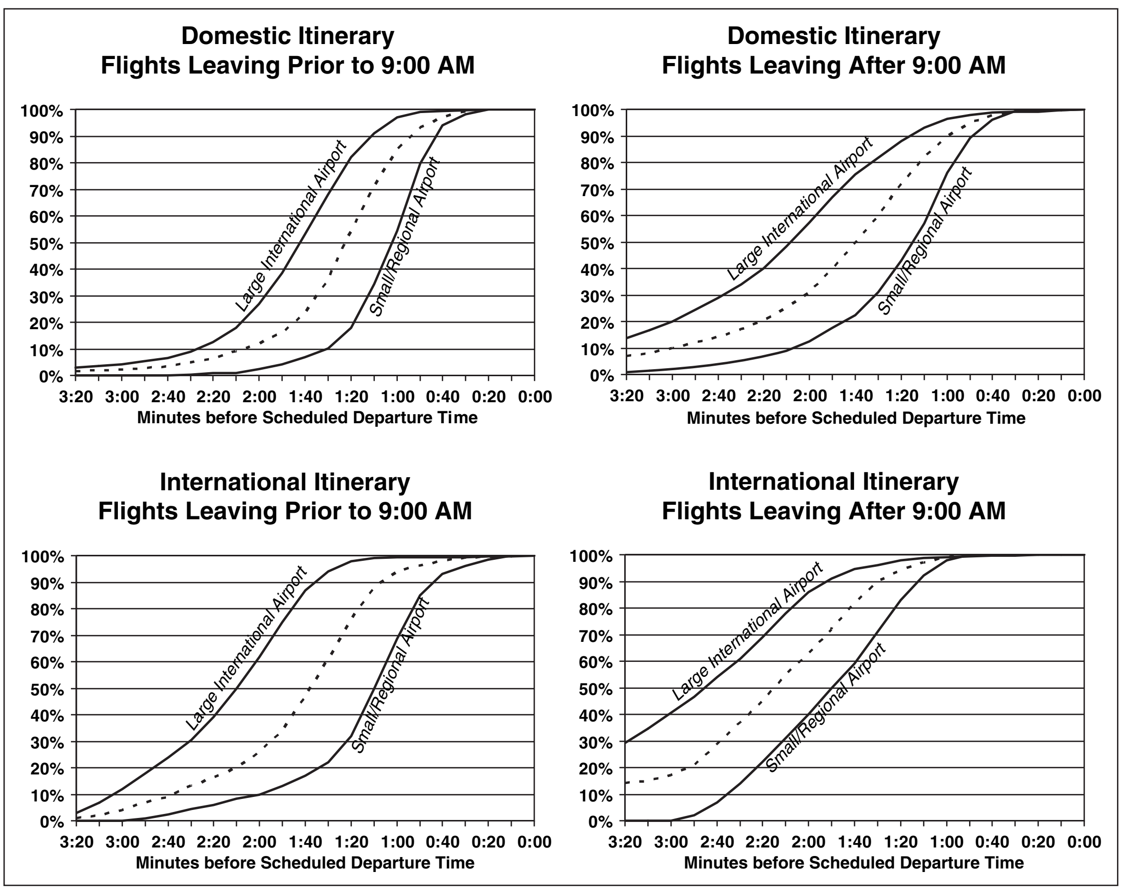


Exhibit 5-14. Departing passenger time of arrival at terminal relative to scheduled departure time (cumulative).

tend to allow more time for processing when their flight leaves later in the day, when traveling internationally, and when using a large airport.

For example, to estimate when the morning peak period would begin at the American Airlines ticket counter at Hartsfield-Jackson Atlanta International Airport, for example, plot out scheduled seat departures (Exhibit 5-15). Note that these data reflect that the morning peak hour, in terms of seat departures, is between 6:15 a.m. and 7:15 a.m. Recognizing that most passengers are traveling on a domestic itinerary and that these flights are departing before 9:00 a.m., examine the curve for a large international airport in Exhibit 5-14 to see that these passengers traveling under these conditions have a median time of arrival at the terminal of about 1 hour and 40 min. Therefore, the recommended time to begin the actual data collection would be 4:35 a.m. (It should be noted that the airline may not even open the counter until 5:00 a.m.; however, if one is measuring demand, it may be appropriate to begin monitoring at about 4:30 a.m. as that is when passengers will likely begin showing up.)

A similar process can be used to estimate when peaks would occur at other terminal elements, depending on their location within a departing passenger's flow through the terminal. For example, if one were looking at processing rates at a security checkpoint, one might slightly reduce the time factor assumed compared to those listed above, recognizing that this process is closer to a flight's actual departure time.

Adjustments to Arriving Seat Schedule

Processes most closely related to arriving seat schedules include bag claim, restroom utilization, and FIS processing. As with departing seat schedules, those for arriving seats may need to be adjusted, albeit less dramatically. For instances, at a small airport, only 5 min may elapse between a flight's arrival and the time its passengers reach a bag carousel. At a large airport, it might take 15 min before passengers reach bag claim.

Adding a "Cushion"

Lastly, it is wise to bracket your peak hour estimate by at least 30 min (preferably one hour) on either side. This will help reduce the chance of missing the peak—either because of a misapplied assumption, inherent variability between scheduled times and actual times (particularly

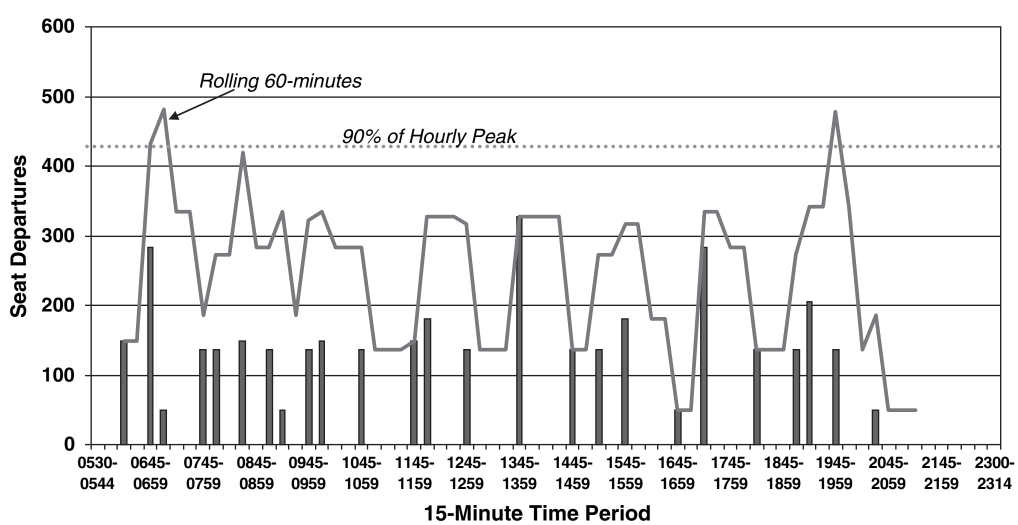


Exhibit 5-15. American Airlines scheduled seat departures per 15-min increments, Hartsfield-Jackson Atlanta International Airport on July 18, 2007.

with arriving flights), and because it gives you a chance to see and describe how a peak builds and wanes.

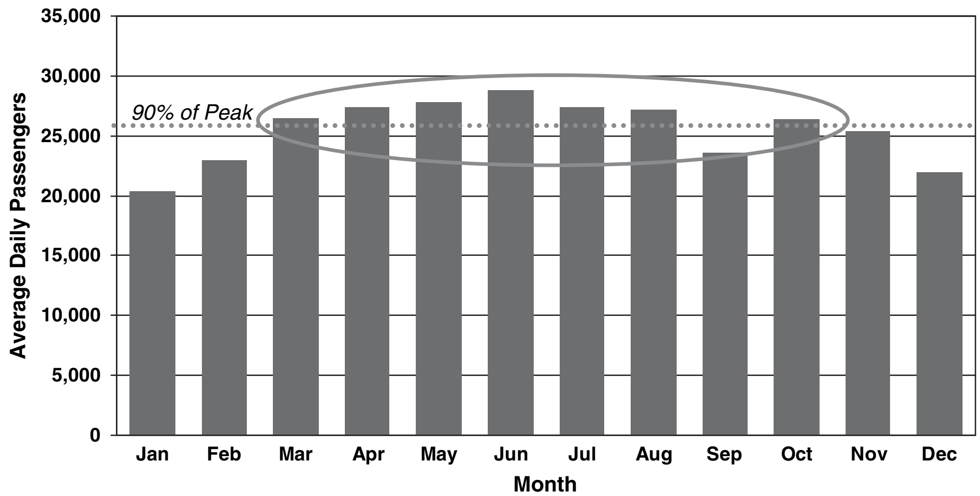
Peak Hour Rules-of-Thumb

There are many factors that affect when peaking occurs at airports; however, when there is a lack of data, the following rules of thumb might be useful to helping establish the peak hour for various terminal elements. While some of these are based on "hard data," many are based on surveyor experience.

1. Local departing passenger activity peaks in the early morning at most airports. At *hub airports,* the busiest morning period in terms of departing flights is usually between 7 a.m. and 10 a.m. (resulting in ticketing and security activity peaking between 5 a.m. and 8 a.m.), while at spoke airports, the busiest morning period begins about one hour earlier—between 6 a.m. and 9 a.m. (resulting in ticketing and security processing peaking around 4:30 a.m. and 7:30 a.m.).
2. A second, more spread-out departure peak begins in the mid-afternoon and lasts through early evening.
3. Local arriving passenger activity is generally light until mid-afternoon, peaking in the early evening.
4. FIS selection of a time of day should be based on knowledge of when peak activity is anticipated at the FIS facility. Historically, the traditional peak period for international arrivals has been the mid- to late-afternoon, reflecting the predominance of trans-Atlantic travel. Beginning in the 1980s, traffic between Asia began to grow rapidly. Since these flights typically arrive in the late morning, United States airports with nonstop service to Asia also see international arrival activity increase at that time as well. In the 1990s, Central American and South American markets began growing at a faster-than-average rate. Peak arrival times for flights from these markets have a less definite diurnal arrival pattern. Overall, therefore, while the typical busy period at most FIS facilities remains the mid-afternoon, the geographic location of the study airport and the mix of international markets served by that airport require the researcher to examine actual schedules.
5. At O&D airports (i.e., those without a significant number of connecting passengers) or airlines with only "spoke" service at hub airports, the peak hour for most functions can be determined by examining the timing of scheduled seat departures (for check-in, security, and other departing passenger-related functions) and arrivals (for FIS, baggage claim, and other arrival-related functions) and making adjustments for the anticipated amount of time between a passenger's reaching the terminal and the flight's scheduled departure or arrival time.
6. For hub airports (and airlines with hub-type activity at the station of interest), actual local passenger (i.e., originating and terminating passengers) peaking activity it is more difficult to directly tie peaks in originating passenger activity to peaks in seat departures and arrivals (as might be obtained from an OAG schedule pull) because the percentage of connecting passengers will vary by bank.

5.7 Proxies for Absolute Peak

While it is often desirable to gather data during the peak month or peak hour, there may be overall project scheduling or resource constraints that would make that impractical. In these instances, it is recommended to choose a time period with at least 90 percent of peak activity. Using the Reagan National monthly statistics as shown in Exhibit 5-16, one can see that six other months were at least 90 percent as busy as the peak month, June. It should be noted, however, that while the *absolute number* of passengers is similar to that seen in the peak month, passenger travel characteristics may be different, which could affect results.



Source: Metropolitan Washington Airports Authority; HNTB analysis.

Exhibit 5-16. Assuming 90 percent of peak period activity is a way to expand survey opportunities.

5.8 The Role of Informants

This section looks at the role informants can play in helping not only determine when to schedule data collection events, but also in helping identify those factors that signal salient differences, or might help explain how entities, resources, and processes are different from one another and hence need to be recorded in data collection.

To use a basic example, assume that your task is to estimate the average amount of time agents spend with customers. Passengers and their travel characteristics will differ in an infinite number of ways, but which attributes might be relevant in explaining how long passengers spend with airline check-in agents? Gender, hair color, height, and so forth are obviously irrelevant, but what is relevant? To learn what is relevant requires knowledge and perspective that can only come from experience. Persons with such knowledge and experience are sometimes referred to as *informants.*

An informant is a person who "knows what is going on" at a given airport, or a particular process at that airport, and is willing to share relevant information with you.

5.8.1 Example of Use of Informants

The following, drawn from the experience of a member of the Research Team, illustrates the value of informants. Annually, international arriving passenger-related processing rate data are gathered at a particular busy airport during one day of the week in the peak month. Having collected these data over several years, and using anecdotal information as well as an analysis of scheduled seat arrivals, the researcher traditionally picked Saturday afternoons as a busy period. In preparation for this annual effort, he reviewed the scheduled seat arrival data for the week and day of interest and noticed that the peak had seemingly shifted from the afternoon to the early evening. Further, a review of TSA wait-time data at the international arrival checkpoint through which arriving passengers must pass prior to entering the domestic portion of the terminal showed significant delays on Thursday afternoons, suggesting that Thursday, not Saturday was the peak day for international arrivals. He was skeptical that Thursday would be a busy day for international arrivals, based on his assumption that most people would want to return from an international destination on a weekend.

To confirm that the day of the week and the time of day had shifted, the researcher discussed the seat arrival and TSA wait time data with a representative of the company that handles international flights at the airport. During the discussion he learned that the peak time of day had, in fact, shifted to early evening; however, the busiest days of the week were still Saturday and Sunday. The informant interpreted the longer wait times at the checkpoint on Thursdays likely to be a function of lower TSA staffing levels on Thursdays, and not reflective of a shift from the weekend. The researcher kept Saturday as the day of data collection, but shifted the time period into the evening to capture the new peak.

Use **informants**; they can help you identify issues of which you may not even be aware, as well as verifying the accuracy of what you have learned through other sources (e.g., past reports, databases).

5.8.2 Summary

While humans are proficient in making meaningful discriminations, there is a large body of research literature that people are also quite adept at inferring patterns when there are none. A related finding is that increased experience is often accompanied by an increase in self-confidence in the accuracy of judgments. For example, experienced law enforcement personnel will often be very confident in their ability to tell when someone is lying or telling the truth; unfortunately, their performance is usually about as poor as those with no experience—essentially equivalent to flipping a coin.

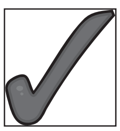
In summary, it is strongly recommended to supplement "hard" data obtained from reports and databases with informants' observations. Given human limitations, however, we recommend conferring with more than a single informant.

CHAPTER 6

Developing the Action Plan

While it's easy to "just go out and collect data," the acquisition of a robust, useable, and statistically defensible data set requires a well thought-out plan to ensure sufficient lead time, thorough preparation, and coordination with many parties. This chapter describes a recommended action plan which provides a process and schedule designed to ensure a successful data collection effort.

6.1 Staffing



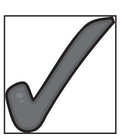
In this section, key staffing roles are considered as the pros and cons of filling these positions in different ways.

6.1.1 Roles and Responsibilities

Typically, large studies involving the collection and analysis of processing rate data will have a person managing these efforts who usually reports to the project manager of the overall study. Exhibit 6-1 shows the typical hierarchical structure for both the client and the research team.

6.1.2 Technical and Social Dimensions of Roles in the Study Team

If you have ever worked as a member of a group charged with completing a task, you could likely evaluate the experience from two vantage points: technical and social. Technical criteria might encompass aspects such as the extent to which the group's output matched the task requirements, the efficiency with which the task was completed, how well the product was received by the client, and so forth. From a social perspective, you may have negative memories of personal attacks, accusations of incompetence, yelling, etc., or positive memories of how members might have put the needs of the group above individual needs. Not surprisingly, behaviors technical and social in nature influence one another. As such, the topic warrants some consideration. Additional information is presented in Appendix C.



A few suggestions follow:

Clearly identify roles and responsibilities. Within the study team, for example, this may be facilitated by documenting the role of each team member, as illustrated in Exhibit 6-2, and then meeting to clarify potential ambiguities. One way of doing this is to think of a situation that might arise and posing it to the group with your interpretation of the limits of authority, sequencing, etc.

Communicate each team member's competence and relative experience. This, as well, might be done at a meeting conducted by the study team's project manager. So doing might help initiate a request from one team member to another who is known to have encountered a similar

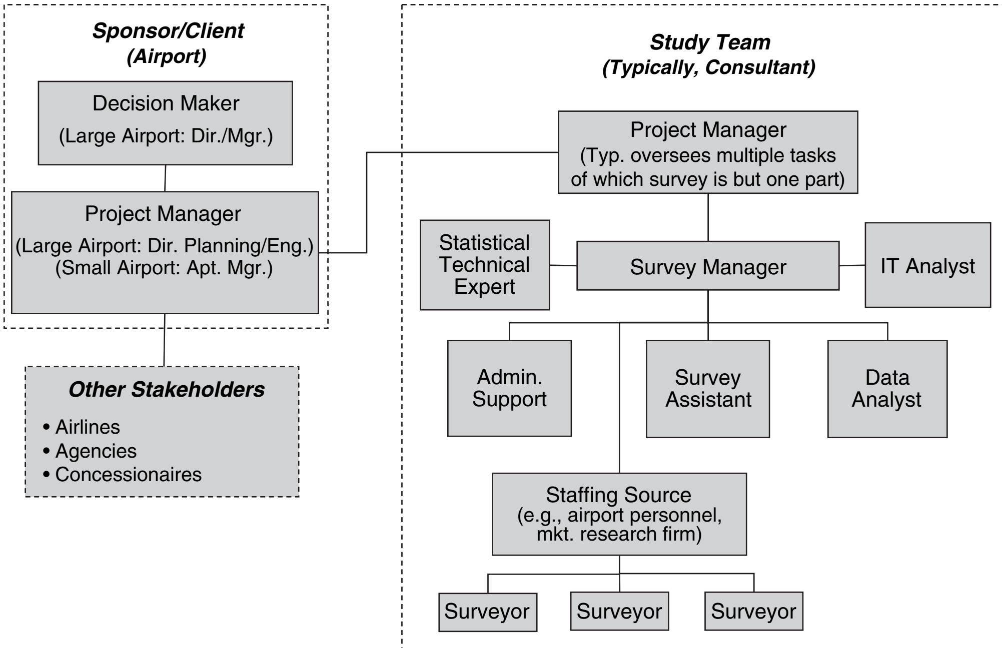


Exhibit 6-1. Sample processing rate study organizational chart (assuming airport as sponsor).

situation and may offer new insights. It also legitimizes suggestions made by persons whose background provides credibility.

Encourage cognitive conflict. Since cognitive conflict can often have a positive impact, encourage those on the project with the appropriate background to suggest alternative ways of implementing the project.

6.1.3 Relationship Between the Study Team and Sponsors/Clients

The assumption is that the study team hired to conduct the research are external consultants, i.e., they are not regular employees of the sponsor, or what are often referred to as internal consultants. While there are somewhat unique advantages and disadvantages to being internal consultants, this Guidebook limits consideration to external consultants.

Just as conflict within the study team may have positive or negative implications, the relationship among members of the study team, and the project manager's relationship with the sponsor or client can sometimes differentiate between a perception of great success or abysmal failure. Lacey[1] proposes generic stages through which external consultants pass in their relationship to the client organization. Exhibit 6-3 builds on the issues Lacey raises as pertinent for the first two stages when working as external consultants. All of this is scalable given the size of airport.

6.1.4 Who Will Staff the Survey?

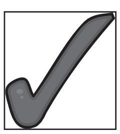
The subject of staffing a data collection event must consider the qualities of individual staff as well as the sources of staff. These must be weighed against the specifications and constraints of

[1] Lacey, M. Y. (1995). Bridging the cultural gap in management consulting research. *International Journal of Cross Cultural Management, 8,* 41–57.

Title/Function	Position in Organization	Role in Study
Project Manager	Senior to mid-level person. Usually managing an overall project (e.g., on-call tasks, or master plan) of which airport survey effort is only one component.	Receives recommendation from survey data and determines fit with overall project goals.
Survey Manager	Mid-level staff person, preferably with survey field experience.	Directs overall survey study, including: o Ensures survey goals are specified o Works with technical experts to determine sampling plan o Coordinates with data collection staffing provider to ensure staffing. o Oversees/reviews analysis and documentation. o Presents results to Project Manager
Statistical Technical Expert	Mid-level go-to person for developing statistical methods.	Provides guidance on sampling plan and analytic approaches.
Survey Assistant	Typcially a junior staff person.	Supports survey manager.
Data Analyst	Mid-level staff person who specializes in technical analysis; should be familiar with airport terminal planning.	Analyzes/synthesizes data; presents results to study team project manager.
IT Analyst/Support	Mid-level staff person who is familiar with technical and programming aspects of PDAs and management of large data files.	Develops software for data collection, processing, and analysis.
Administative Support	Support personnel.	Provides administative support to study team (e.g., preparing/ printing forms, editing reports, shipping, organizing data and support materials.
Data Collection (Survey) Staff	Often a market/research company.	Responsible for providing staff

(1) Depending on study size and existing organizational structures, some roles may be combined or arranged differently.

Exhibit 6-2. Study team roles. [1]



Stage	Key Issues
Entering	Distinguish client roles. [For example, the person who is funding the project might not be the Principal Point of Contact (POC); indeed, he or she might be displeased if you approach him or her rather than the POC on key issues of the project's implementation]. Build relationships with the clients. Try to understand the unique jargon of the organization. Assist the client in articulating what he or she really needs rather than what they might say they want.
Contracting	Document your understanding of the project's purpose, key milestones, funding arrangements. Clarify in writing how confidentiality will be treated. Recognize that your approach and the ethical implications of what you propose might result in a termination of the relationship.

Exhibit 6-3. First two stages of client-consultant relationship.

your study, including the size of the staff, the duration of the data collection period, and the amount of time prior to the start of data collection that is available. This section reviews these considerations and also lists possible sources for staffing your data collection team.

Qualities of a Good Data Collector

Many qualities that are desirable for a data collector in an airport environment are also desirable for any job position. These attributes, as well as those that are of particular benefit to data collection are listed as the following, moving from general to specific qualities:

- Has a good work ethic;
- Is a team player;
- Is able to follow directions;
- Has demonstrated accuracy and attention to detail;
- Is able to focus attention for extended periods;
- Can be held directly accountable;
- Has the ability to work long and non-traditional hours (i.e., early morning, evenings and weekends);
- Is able to think on his/her feet;
- Has a vested interest in the outcome of the research;
- Has at least a general understanding of airport passenger processing; and
- Has security badging for the subject airport.

Sources of Staffing

There are several options for staffing data collection events. The most common include the following:

- Airport staff;
- Consultant staff;
- College students and interns;
- Temporary agencies; and
- Market research firms.

Airport Staff. Sources of airport staff can include personnel from management, planning and engineering, operations, and/or other offices. Airport staff are already familiar with the facility, have proper security clearances (i.e., they are “badged”), and presumably have an allegiance to their airport. These benefits are most readily seen for studies of shorter duration and that have short lead times. For large studies of long duration, however, it may be difficult to find a large pool of available airport staff.

Fellow Consultant Employees. If you are a consultant, you could choose to staff your data collection event with fellow employees. In many cases, you might already have a good working relationship with them; most are likely familiar with working in an airport environment and possibly even with the specific study; and some may already have security clearances. Challenges may be finding sufficient staff for longer, larger studies. Also, they may have high hourly rates which must be weighed against the likely benefit of their providing higher quality data than that gathered via other staffing sources.

College Students and Interns. The mere fact that they are going to college suggests that students and interns are academic/research oriented and able to observe and record information. In addition, one big advantage of using college students and interns is that they are relatively inexpensive. If you have recruited the students or interns from academic departments such as management science, operations research, or aviation management, they may be personally motivated to work in the field to gain experience. Airport interns will also likely have been badged. College students, however, will need to go through a security check and badging process, increasing the lead time needed prior to data collection.

Temporary Agencies. Temporary agencies provide significant flexibility, staffing for many kinds of jobs and for time periods ranging from 1 day to several years. Often, people in the "pool" are looking for variety in their career and an opportunity to earn extra money working outside of the typical nine-to-five environment. Lastly, temporaries come at a relatively low hourly rate.

The benefits listed, however, are often offset by several factors. Depending on policy, it may be necessary to bid the data collection portion of the project, and doing so can add a couple of months to the study. Obtaining security clearances will also increase the lead time needed by at least a couple of weeks. Finally, it must be added that there is probably less incentive for temporary agency personnel to produce solid, accurate results.

Market Research Firms. Market research firms exist to collect data needed to make informed decisions. Because this is what they do, they have an incentive for ensuring your project goes smoothly and that you get what you want compared to other temporary agencies. Another advantage is that they usually have a large pool of employees who are experienced in such tasks as record-keeping, taking quotas, using data collection forms, interviewing people, and so forth. Personnel often enjoy working non-traditional hours and being away from an office. Depending on the arrangement, the management staff of the market-research firm can handle the scheduling of employees, freeing you to focus on other project issues; this is especially beneficial for larger studies of long duration. Finally, market research firms can provide services at lower hourly rates compared to using fellow employees of a consulting team.

Nevertheless, there are some disadvantages to using market research firms. The first is, as with using a temporary agency, it may be necessary to bid out the data collection portion and to get market-research staff badged, adding significant lead time. Second, while they may be familiar with making observations in general, many employees may not be familiar with the airport environment. Finally, as with temporary agencies, their lower hourly rate and unfamiliarity with the overall project may lead to them collecting less accurate data compared to data collected using other staffing methods.

Summary

Identifying the most appropriate staffing source and individuals for your data collection event is contingent upon how much lead time is available, your budget, the size and duration of the study, and its complexity. Several staffing sources exist and the researcher must consider the factors outlined above when making a selection.

6.1.5 How Many Staff are Required?

Staffing levels are primarily a function of the following factors:

1. The number of observations to be collected per event;
2. The complexity of the data to be collected;
3. The general length of a unique process (transaction time);
4. The anticipated duration of the data gathering event; and
5. The layout of the facility in which the observations will be made.

In an attempt to reduce costs, you may be tempted to give data collection staff the responsibility of monitoring too much activity than can be done accurately. The result is often poor-quality data. The following section offers some guidelines for sizing a data collection effort.

6.1.6 Sizing Guidelines

Review the Total Amount of Observations

As a rule of thumb, assume that it will take more resources to collect 10,000 observations than 1,000 observations, and that the relationship between observations and resources is largely linear (i.e., a straight line best captures the relationship.)

Review the Amount of Data to be Collected for Each Record

The more elements being recorded per unique observation, the fewer observations can be made per staff person. For example, if surveyors are merely timing check-in transactions, they can likely observe three or four transactions simultaneously. However, if they are asked to record party size, number of checked bags, and determine whether the party is traveling on an international or domestic itinerary, they may only be able to observe one or two simultaneous transactions.

Review the General Length of a Unique Transaction

The longer a unique transaction takes, the more surveyors will be required to collect an adequate sample. For example, if a transaction takes about 2 min and a surveyor can monitor one transaction at a time, he will collect about 30 records per hour. If, however, a transaction takes twice as long, he will only be able to get about 15 records per hour.

Review the Duration of the Data-Gathering Event

Assume that survey personnel will be able to remain in position for a maximum of approximately three hours unless there are mitigating circumstances. For outdoor events in conditions of extreme weather, for example, reduce the length of the observation period. Allow staff approximately one short break (10 to 15 min) every hour or two (assuming the loss of some data would be acceptable). To minimize loss of key data, schedule these breaks, when possible, when you anticipate less activity. If the loss of data is not tolerable, schedule at least one additional person to relieve or "spot" others so they can take breaks. Finally, to counter boredom, try to rotate staff across locations. (A disadvantage of rotating staff is the fact that they may require more training and time to familiarize themselves with a new setting.)

6.1.7 Lines-of-Sight and Site Layout

Two often neglected issues pertinent to determining staff requirements are (1) what observers are physically capable of given the average person's field of vision, and (2) the layout of the site.

While the human field of vision is about 180°, the area where binocular fusion (the range of stereoscopic vision) is possible is between 120° and 140° (see Exhibit 6-4). You can approximate this angle by extending your arms away from the body midway between straight ahead and parallel with one's shoulders. More importantly, however, the *area of focus* is limited to only about three degrees. This means that for detailed observation, the data collector must pivot his or her head to observe multiple activities (e.g., observing several agents as they check in passengers). Based on these experiences, it is recommended that a data collector be limited to making observations within an area of no more than about 60° to 70°. Beyond this range, the likelihood of missing pieces of information or even whole observations will increase to intolerable levels.

Exhibit 6-5 shows two hypothetical airline check-in areas, both with four self-serve kiosks placed together and parallel to the adjoining circulation area. In the first case (left image), an observer can stand far enough away to minimize the need to pivot her head to see the entire check-in area. In the second case, the narrower kiosk area means a single observer cannot stand far enough away to get all four kiosks in the same field of view. In this instance, two people might be needed.

Another consideration is a site's layout with regard to obstructions. As shown in Exhibit 6-6, the layout of resources (in this example, kiosks) and impediments within lines-of-sight, such as

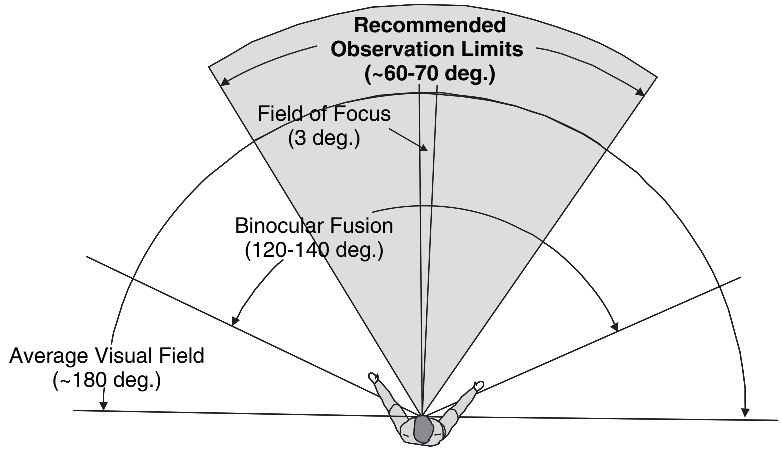
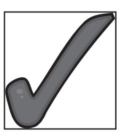


Exhibit 6-4. Recommended observation limits for data collection.

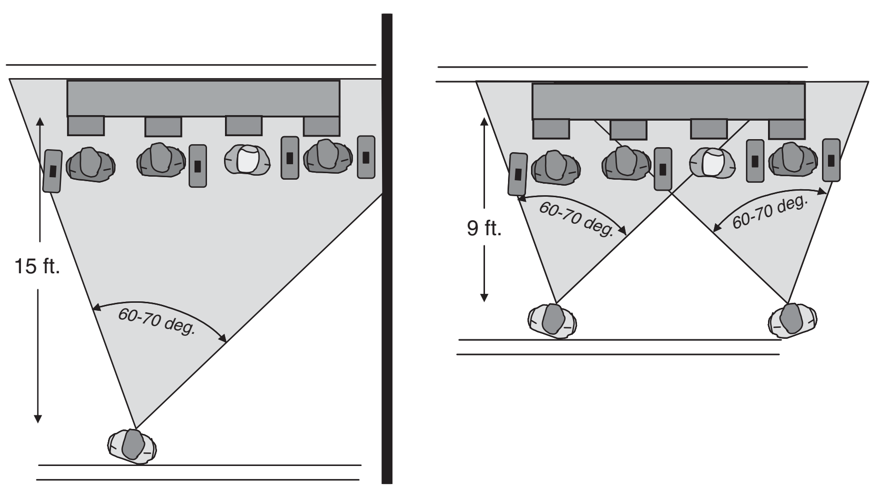
queues and columns, may require more staff. Likewise, the recording of various elements in a process may require the data collector to move from one location to another (for example, following a passenger from a kiosk to the bag drop area), again limiting the number of transactions that can be monitored at the same time.

Challenges associated with site layout and lines-of-site are most easily prevented by conducting a pilot survey at the site under similar conditions.

6.1.8 Planning for Staffing Contingencies

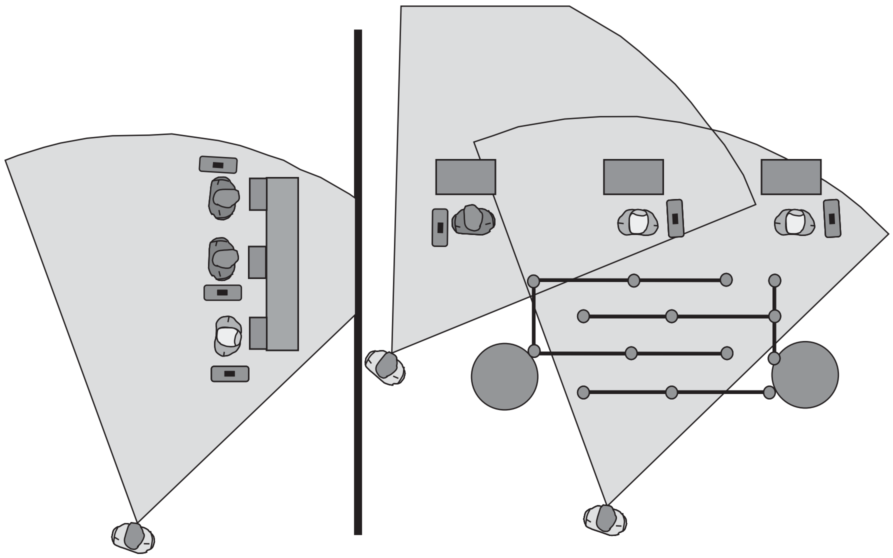


Finally, to counter random problems that inevitably arise, plan to overstaff each event by at least 10 percent. If, for example, you anticipate the need for 10 people, arrange for 11 or 12 to show up. If all proceeds without mishap, the additional staff can be deployed to collect additional data, assist in quality assurance, or provide additional rest breaks to others.



Although the kiosk configuration is the same in both situations, the observer on the left is able to see all four positions within his field of view, while the narrower circulation area on the right results in the need for two observers to capture activity.

Exhibit 6-5. Impact of field of view on staffing.



One person can observe three kiosks in the left configuration, while with the layout on the right, two people would likely be required.

Exhibit 6-6. Impact of site layout/obstructions and field of view on staffing.

6.1.9 Staff Accountability

Employing temporary personnel, regardless of their source, may bring problems of staff accountability and, in turn, the quality of the data collected. For example, personnel may simply fail to show up for an assignment, or may be late in arriving. In addition, staff may not be as vigilant as desired in accurately capturing the needed data. As suggested in Section 7.2, orientating staff to the purpose of the research, the importance of maintaining accurate records, etc., should help improve accountability. The following are a few suggestions:

1. Quite simply, there is truth in the adage “you get what you pay for.” If at all possible, hire quality personnel, preferably from a reputable market-research firm which has its bottom line at risk if its staff performs poorly during your study.
2. Document a realistic set of requirements as criteria for selecting candidates. At the simplest level this might entail specifying that staff are comfortable filling out forms and are familiar with the typical airport environment.
3. You should specify that staff will need to stand for 3 hours at a time and possess sufficiently good eyesight to observe events occurring at a distance of 30 ft. For more sophisticated requirements, such as proficiency in using a particular device for recording observations, specify the requirements in the purchase order or contract with the company providing staff.
4. Confirm that each employee has been properly screened (i.e., can operate within the security parameters required) to reduce the chance of having undependable staff on your study.
5. Finally, do not hesitate to remove individuals who are not dependable or do not seem to be performing at the desired standard. Spending time at the beginning to define performance requirements, and selecting persons on the basis of those requirements, will reduce the likelihood of having to address poor performance.

6.1.10 Summary

In determining the number of staff required, take the following into account:

- How much data is required for each record, and how many persons are needed to capture those data?

- How long will the data gathering events last, and how long can you reasonably expect staff to function without rest breaks?
- Do the layout and line-of-sight issues impact the number of staff required?

6.2 Coordination

A key component of a successful data collection event is good coordination: ensuring that each staff member knows what he or she should be doing, when and where the task should be executed, and how his or her tasks relate to the work performed by others. Ultimately, it is the project manager's responsibility to ensure that happens. Issues related to coordination with various groups are addressed in this section.

6.2.1 Coordination with Research Team

All too often, the data collection team is told to "go get some data about . . ." and is not given guidance as to the specific data needs. Ideally, coordination with the research team will be continual in the weeks leading up to the actual data collection event. In particular, there will be ongoing coordination with the project sponsor in defining and refining the research question and the associated data collection requirements.

6.2.2 Coordination with Various Airport Entities

During one large-scale airport data collection event a few years ago, several traffic counting personnel were stationed alongside an airport road early one weekday morning. Airport police pulled up, told them not to make "any sudden moves," and began questioning them about their activity, eventually threatening to take them in for additional questioning. In actuality, the airport police had been notified of the event and the surveyors had with them copies of signed letters from a senior airport manager, identification, and typical traffic counting instruments. After a few phone calls, the situation eventually defused. The lesson learned is that coordination with the airport is essential, and that even when properly planned and executed, problems can still arise.

Data collection events are often executed by a consultant under contract to an airport or airline. It is recommended that the data collection project manager contact the airport/airline project manager (usually someone in the planning department) daily to discuss field work progress, and discuss any issues that have come up that the airport project manager may be able to correct. Encourage the airport's representative to communicate with airport operations and airport police throughout the data collection effort about all planned data collection events including their schedule, location, and purpose.

6.2.3 Coordination with Other Stakeholders

Airports are complex systems and responsibilities are widely dispersed. As such, you may need to extend coordination to groups such as the TSA, the airlines, commercial vehicle providers, and CBP, to improve overall awareness of the field work and, in turn, minimize the risk of potential disruptions. Collaborate with the field project manager and the airport's project manager, to identify and engage the support and cooperation of other potentially relevant stakeholder groups.

6.3 Data Recording Methods

There are two basic options for capturing data: paper and pencil (PAPI) and electronic handheld devices (personal digital assistants, or PDAs). Both are considered in this section.

6.3.1 Paper and Pencil

Stopwatches, clipboards, and paper and pencil have long been the tried-and-true method for collecting data. The PAPI method has several advantages and disadvantages over electronic data gathering means.

The following are **advantages** of the PAPI method:

- In general, once introduced to how the data are to be collected, most people will not need special technical training on using the PAPI method;
- A traditional 8½ × 11-in. paper is usually large enough to allow several rows and columns of data to be entered for several records, often allowing multiple events to be monitored simultaneously and recorded on one sheet;
- The forms can be easily seen in moderate to very bright lighting conditions;
- In the field, errors can be quickly corrected by erasing or strikeout;
- Alpha characters (especially notes) can easily be entered to clarify observations; and
- In the field, the forms can easily be edited to change the kind of information being recorded.

The following are **disadvantages** of the PAPI method:

- Data recorded on paper almost always have to be manually entered into a computer database, requiring additional time (and cost);
- It can be hard to read someone's handwriting, resulting in misread information;
- As data are entered, typographical errors can be made;
- The amount of paper generated is roughly linearly related to the amount of data collected, often resulting in added shipping costs and data management/storage challenges;
- The PAPI method (which typically includes a stopwatch) does not lend itself to timing very short events—i.e., those lasting only a few seconds; and
- If an insufficient number of sheets is provided to a surveyor, some observations may be lost (although the back side of a form or scrap pieces of paper can be used in an emergency).

6.3.2 Electronic

Three electronic data collection options are considered below: PDAs, video monitoring, and video data capture software.

Personal Digital Assistants

While personal digital assistants (PDAs) were introduced in the 1990s, their use was not popularized until a decade later. On the basis of the datasets the Research Team reviewed for this project, as well as anecdotal evidence, PDAs appear to play a role in many airport passenger-related processing rate data collection events. While notebook or tablet computers may also be options, they have not been considered given their current lack of suitability for situations where data collectors have to move around or must stand for long periods.

A typical PDA has a touch screen for entering data and a memory card for storing data. When used in the field, the PDA is powered by a rechargeable battery.

In general, it is recommended that your choice should be driven by the particular data collection event. The following are some advantages and disadvantages of PDAs, based on personal experience and a review of the current literature.

The following are some **advantages** of PDAs:

- They often result in more completed records than PAPI insofar as the software can be programmed to "force" surveyors to enter items;

- Limits can be placed on input fields (e.g., fields requiring a party size will not take an alpha character);
- The initial costs for hardware and programming are usually offset by reduced labor costs, especially in situations where the same type of data is repeatedly collected. (The cost-recovery period will vary based on numerous factors, including the number of PDAs acquired, their unit cost, the cost to program, the amount of data collected, and the number and scheduling of data collection events.)
- Directly downloading data from a PDA eliminates manual coding errors that can occur with manual data entry;
- They result in higher accuracy when quick processes are recorded (i.e., those lasting only a few seconds); and
- The amount of data capable of being stored on a PDA can be equivalent to thousands of sheets of paper, reducing shipping, handling, and storage issues.

The following are some noted **disadvantages** of PDAs:

- Specialized IT staff are usually required to program the devices and develop procedures to download and save backups of data;
- Some personnel may be unfamiliar with the technical aspects of using PDAs and will need additional training;
- If many forms and locations are available on a device, care must taken during the programming/user interface layout phase to minimize the chance of users erroneously hitting buttons that could result in the device to "think" it's at a different location or is collecting different data;
- It can be difficult to edit records in the field;
- Several studies noted technical difficulties in entering, storing, and retrieving data from PDAs which compromised the study;
- If it is decided to change how data are being recorded in the field, it can be very cumbersome to edit the fields on the fly;
- Their small screens limit the number of data fields that can be displayed at any given time;
- Their small screens can also make it difficult to enter data from multiple observations occurring simultaneously;
- Their screen may not be easily visible in bright sunlight;
- Some studies reported lost/stolen PDAs resulting in lost data, although this was not experienced directly;
- When using PDAs, typing text strings can be cumbersome, causing the surveyor to miss some events;
- Extra devices are required due to recharging periods, download periods, and tight schedules (for example, if one data collection event ends at 2:30 p.m. and the second begins at 3:00 p.m., you might need extra PDAs); and
- While the price of good, rugged PDAs has decreased significantly, hardware costs obviously increase as the number of devices increases.

Video Monitoring

Video can be collected using existing cameras, or employing special setups, ranging from one camera to a network of cameras, whose use is limited to the duration of the study.

The following are **advantages** of video monitoring:

- Video images can be analyzed at times that are convenient to the observer;
- Using rewind and fast-forward functions permits the observer to examine transactions multiple times, as well as allowing him or her to skip past "dead" time; and
- It can serve as an archive of the actual transactions.

The following are **disadvantages** of video monitoring:

- The camera is usually in a fixed position and focal length, limiting coverage, all desired locations may not be monitored;
- The limited field of view prohibits the observer from seeing events that are upstream, downstream, or directly adjacent to the activity being monitored;
- Objects may either be permanently or temporarily blocking a portion of the field of view, affecting the "robustness" of specific observations;
- Approval to install cameras may be time consuming and involve extensive coordination; and
- Video may be deliberately or accidentally overwritten by someone in the agency overseeing the storage.

Video Data Capturing Software

Recently, a new technology, typically called data capturing, has become available. This technology interprets video and automatically creates a data set of activity without human involvement. While the technology appears promising, an initial review of the current state-of-the-art suggests it may not perform satisfactorily in a crowded airport environment. Monitor the technology as it evolves in sophistication.

6.3.3 Summary and Recommendations

In most instances, the kind of data collected for passenger-related processing analysis and the field environment in which these data are collected (airport terminals) lend themselves very well to the use of either PAPI or PDA technology as long as standard protocols are followed.

Video recording may be useful as transactions can be viewed at convenient times, may be reviewed repeatedly to confirm what actually took place, and act as a formal way of documenting events. Nevertheless, because the events take place in three dimensions and are affected by activities upstream and downstream, it is our opinion that video recording is inappropriate in many situations.

Video capture technology, it is believed, still has some significant hurdles to overcome before becoming a simple, reliable, and inexpensive method.

The series of graphics in Exhibit 6-7 below provides general, relative guidance as to the factors to be evaluated in selecting between PAPI and PDA.

In summary, limiting the choice between PAPI and PDAs, researchers are encouraged to consider the following factors:

- In general, the PAPI data collection method is most appropriate for short, one-time, simple data collection efforts. As complexity, sample size, and number of data collection occasions

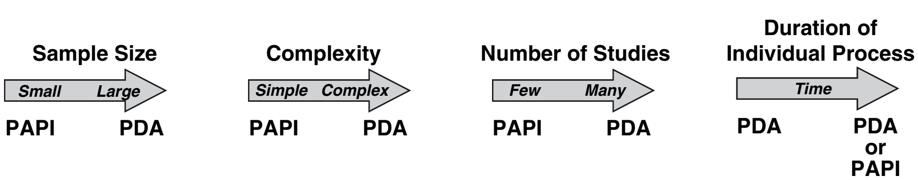


Exhibit 6-7. Relative guidance on selecting between PAPI and PDA for recording data.

increases, the value of PDAs increases. In addition, PDAs perform especially well for monitoring processes of extremely short duration (e.g., escalator boardings).

- The tradeoffs between PAPI and PDAs are focused around weighing the anticipated cost of entering data compared to the cost of programming the PDAs.

6.4 Countdown to Data Collection

While no two research efforts are alike, even those of comparable purpose, scope, and complexity, there are some general planning guidelines that will help. Be aware that you may need as much as 2 months of preparation, from the start of initial planning until the first day of data collection. To summarize what was considered earlier in this chapter, key variables that will affect the schedule include the following:

- The size of the data collection effort (e.g., the number of airports and areas with each airport being observed, the duration of events, the number of days, the level of detail of individual data points).
- The anticipated level of coordination (level of airport bureaucracy, number of stakeholders involved, etc.).
- Prior experience in conducting similar efforts (e.g., Can data sheets from similar efforts be revised or do new sheets need to be created from scratch?).

Exhibit 6-8 presents a general timeline for planning the data collection effort beginning at 8 weeks.

Week		Activities & Tasks
8	✓	Confirm data requirements (sample Size, items of interest, process definitions, etc.).
	✓	Finalize survey dates.
7	✓	Identify staffing source(s) (e.g., market research firm, Airport staff, etc.).
	✓	Prepare draft scope, budget, and schedule.
6	✓	Initiate preparation of draft data collection sheets.
5	✓	Initiate coordination with Stakeholders (Operations, Police, TSA, CBP, Airlines, etc.).
	✓	Conduct site visit (identify potential observation points, diagram layout, etc.).
	✓	Initiate process for obtaining security badges (if needed).
4	✓	Field test data entry forms (electronic or hard copy).
	✓	Confirm staffing requirements per data collection shift.
	✓	Prepare detailed data collection schedule.
3	✓	Refine data collection entry forms as needed.
	✓	Prepare detailed staffing schedule.
	✓	Identify and secure space for data collection operations within airport.
2	✓	Obtain data collection materials (stop watches, tally-counters, clipboards, pens/pencils, etc.)
	✓	Conduct follow-up coordination with stakeholders.
	✓	Design/print survey tags.

Exhibit 6-8. Pre-planning: 2 months out.

Day	Activities & Tasks
5	✓ Duplicate data collection forms.
4	✓ Assemble observer packages, sorted by day, data collection event, and location.
3	✓ Ship all materials to site.
2	✓ Set up field office.
1	✓ Conduct training and orientation for data collection team.
Day 0 – Day n	✓ Implement data collection.

Exhibit 6-9. Immediate pre-planning and execution.

Exhibit 6-9 identifies key activities usually performed the week immediately prior to the start of data collection, as well as the actual data collection.

6.5 Post-Data Collection

Exhibit 6-10 outlines post-data collection functions.

Week	Activities & Tasks
1	✓ Begin data entry (if not collected electronically).
2	✓ Begin data editing and "cleaning," eliminating incomplete records, correcting coding errors, etc. ✓ Decide how to handle outliers.[2]
3	✓ Initiate data analysis.
4	✓ Begin formal documentation and report preparation.

(1) Outliers are data values that seem to be part of another distribution based on their relative difference from other values.

Exhibit 6-10. Post-data collection functions.

CHAPTER 7

Managing and Implementing Data Collection

This chapter provides guidance on managing day-to-day field operations during a data collection event. Airports are unique places for collecting data and a failure to adequately take this factor into your study will seriously compromise your results. In Chapter 7, the proposed action plan and strategies are tested through site visits and field testing. Successful implementation of the final action plan is met through team preparation. Finally, an overview of two data collection events is presented as a checklist.

7.1 Understanding the Environment

In Chapter 6, the importance of coordinating research efforts with airport personnel is emphasized. The world, however, is imperfect, and even when one does solicit and get cooperation, problems can still arise. At one airport, for example, the dominant carrier refused to permit data collectors access to its gate hold areas, essentially crippling the study. The purpose of the study had been communicated to all station managers through memos from a high-level airport administrator, and all other airlines had signed off on the study. Eventually, a face-to-face meeting with airline representatives, airport management, and the lead consultant was required to obtain cooperation.

There are many other examples of the unique challenges to data collection in an airport. Many of these challenges intensified after the September 11, 2001, terrorist attacks, as airports and the agencies working in them become more sensitive to potential security threats. This section addresses how to better understand the dynamics at an airport that can impact the study, as well as how to reduce the likelihood of unanticipated problems. In particular, consider both the bureaucratic and the physical environment of airports.

7.1.1 The Bureaucratic and Physical Environment

In many ways, an airport is like a microcosm of a small city. Its functions are myriad, and authority is spread across multiple agencies and staff with differing reporting relationships. As such, coordination must be maintained with those with functional responsibility for the following:

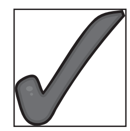
- Airport management (those not directly involved in the study);
- TSA;
- Airport police;
- Airport operations;
- CBP;
- Airlines;
- Transit agency (if transit station is on-site); and
- Ground transportation providers.

The amount of communication and coordination with these entities can vary significantly depending on factors such as the following:

- Existing relationships among the organizations;
- Security threat levels;
- Their prior experience with similar data collection events; and
- Local, regional, and national policies (relevant to national agencies).

Largely related to security issues emerging after September 11, 2001, protocols and policies may change substantially over time at both a local and national level. For example at one large airport, prior to September 11, 2001, data collection personnel could enter security along with passengers, well-wishers, and greeters. The next year (after the terrorist attacks), the FAA/TSA required personnel to present a letter from airport management with the person's name on it. The following year, personnel were required to have a security badge. It is important to monitor changing policies and practices, and to prepare for changes sufficiently in advance of data collection. (See Chapter 6 for scheduling recommendations.)

7.1.2 Site Visit and Reconnaissance



As the one responsible for collecting the data, the data collection project manager should visit the functional element(s) to be studied. The purpose of this site visit is to observe processing, clarify the particular data elements to be collected, and confirm how special situations, if any, should be handled. Most importantly, these initial observations allow the researcher to see the elements in their setting. The following are some recommendations for the site visit:

1. Invite the researcher (end user) to accompany you. This will not only contribute to a common understanding between the two of you but may also result in modifying elements of the project's purpose and scope.
2. If possible, schedule the visit for a similar time period to the one anticipated for the actual data collection event.
3. Meet with airport staff to identify potential modifications to the existing layout prior to the scheduled data collection.
4. To the extent possible, pre-test the data collection protocol, data recording instruments, and so forth. Is there a situation that wasn't initially anticipated? Do the data entry sheets/templates need to be modified? Will additional staff be needed to cover the desired locations?

The next section provides a recommended field testing approach.

7.1.3 Field Testing

Research projects are costly efforts. While you can try to anticipate contingencies in your planning, a formal walk-through before collecting usable data will permit you to make needed modifications to forms, schedules, and procedures, thereby increasing the likelihood of gathering reliable data. What has been proposed in this Guidebook is grounded in a substantial and multidisciplinary body of research. The concept in its most basic form is that, before you can trust data, you need to validate that you are collecting the "right" data (validity) and doing so in the right manner (reliability). After Deming,[1] the approach is sometimes referred to as **Plan, Do, Study, Act** which means the following in the context of this chapter:

- **Plan** involves specifying the performance criteria for data collection methods and materials (see Chapter 3);

---

[1]Deming, W. E. (1986). *Out of the Crisis.* MIT Center for Advanced Engineering Study.

- **Do** refers to implementing these methods and materials on a trial basis;
- **Study** relates to evaluating performance in light of requirements, and;
- **Act** is using the feedback to make changes for improvement.

In research, triangulation involves use of multiple researchers, multiple methods, and multiple sources to enhance the credibility of the findings.

You might reasonably think of this as a calibration activity. The approach often also incorporates a triangulation method that is commonly used in research to enhance confidence. In both surveying and navigation, triangulation is used to determine coordinates and distance to a point using two other known reference points of a triangle. In research methodology, triangulation implies using multiple sources of information, multiple methods, and/or multiple researchers. In an organizational setting in which one is studying employee satisfaction, for example, interviews may be collected with persons from different units and at different hierarchical levels (multiple sources). To supplement interviewing as a data gathering strategy, the researchers might also review employee turnover statistics as well as survey employees (multiple methods). Finally, as an illustration of multiple researchers, two or more researchers may participate in analyzing focus group data, independently reviewing the data and then meeting to identify where the researchers' conclusions converge and diverge. As an analogy, if it looks like a duck, quacks like a duck, and walks like a duck, the conclusion that it is a duck is more persuasive because more than one attribute was examined.

A Generic Field Testing Process

Passenger check-in is used to illustrate the process, but it can be adapted to other data gathering foci.

Schedule Field Test Session. Arrange for all persons involved in the pilot test to meet approximately 1.5 hours prior to the actual start.

If possible, orientation should be done on site (i.e., at the study airport), in a conference room facility capable of accommodating the entire survey staff. An airport representative should make welcoming remarks noting the airport's significance to the region and the importance of the particular study.

An example of a training outline and agenda, as well as sample orientation materials is included in Appendix C of this Guidebook.

Field Test Orientation. Define the purpose of conducting a field test and refer participants to that section of the handout related to field testing. Depending on the number of persons in the team, consider having two persons collect the same data on the same elements at the same time. In other words, both persons observe and record identical data for purposes of calibration. Review task particulars and times. Synchronize watches, and direct the team to the field-test location.

Initiate Field Test. Allow approximately 15 to 20 min for data collection. While the team is collecting data, observe each person, looking specifically for any factor that might impact that person's ability to collect those specific data for which he or she has responsibility. Note any unanticipated problems or difficulties that the staff is having. At the end of the period, meet with the team to debrief. In particular:

- Ask each team member to comment on any questions, problems, etc., that arose during the trial.
- Review the logic of the observation strategy for possible change.
- Have each person run a quick calculation of averages, ranges, and standard deviations and have each person report the results.

Review Field Test Results. Review the results in regard to the following:

- Consistency in data where expected. For example, if two persons were making identical observations but of different check-in lines, are the averages relatively similar? Are there relatively large standard deviations for each of these observers? If so, is there something different between the two lines that suggest that you should not treat them alike? If so, examine the ranges to see if it might be a function of one or two extreme values.
- Consistency across observers. If you chose to have two persons making identical observations for part of the trial, do the observations match? While a small amount of observation is normal, you should not expect to find wide variation. If there is, what were the two persons doing differently? Why are they not calibrated?
- Consistency in data where not expected. If you have opted to stratify based on the assumption that groups will differ in some way (e.g., business and first class passengers will, on average, be unalike in regard to time to check-in), have these differences emerged? If the averages and standard deviations are relatively similar, should you consider ignoring the stratification strategy and treat them as a single group (i.e., combining business class and first class)?

Spend a few minutes by yourself to consider if you want to make any changes. Then reconvene the team, communicate any changes, and direct them to begin actual data gathering.

7.2 Team Preparation and Training

To help ensure a successful data collection event, a staff training session will be necessary. It is recommended that this be broken out into three components:

1. Overview of study and role of data collection,
2. Procedural/management issues, and
3. Actual training.

7.2.1 Conduct Team Briefing

While some staff preparation might have already occurred during the planning phase, it is still important to communicate to the assembled data collection team the overall purpose of the study, i.e., the actions that might arise as a result of the data. This helps to place the frequently mundane task of data collection in the greater context of the study. There is evidence from the research literature that performance is enhanced when workers understand not only how to perform a task but why the task is being done and how it fits into a larger framework.

7.2.2 General Requirements and Logistics

It is important to stress to staff that the data must be collected at the designated times (typically, peak times) and the importance of meeting their agreed upon responsibilities, i.e., that if even just one person doesn't show up, the entire event can be compromised. Once the scheduled period is gone, it's too late, particularly if the period was a peak time. Finally, given that other data collection events may have been scheduled, it is often very difficult to make up the missed one. Regardless, even a highly motivated and well prepared group of people will, from time to time, encounter unexpected and unavoidable circumstances. Emphasize the importance of contacting the supervisor at the first indication of a potential conflict.

Give each data collector a handout which summarizes the purpose of the study, his/her role, contact information of other team members (including key airport contacts), a detailed schedule, and instructions for data collection events in which he/she would be participating (see Appendix C for examples).

The amount of time spent on each of these items and when they are covered will depend on the size and complexity of the event, its schedule, and experience of the personnel.

For a large study, for example, with many inexperienced personnel, it may be appropriate to hold the introductory session the day before the actual data collection event. At this first session, one could outline the general study and how the collected data will be used to make a recommendation. One could then review procedural issues, such as when to report, how to dress, who key airport and study team members are and how to get in touch with them, and how to record data. The next item on the agenda would be a walk-through of the airport, including visits to the sites where data collection will take place.

7.2.3 Specific Data Collection Training

The next item on the agenda would be the actual training. It is recommended that this occur as near as possible to the actual data collection event—preferably, immediately prior to it. Sufficient time should be allowed (at least 30 min, preferably, more) to allow staff to enter sample data and iron out the exact process. During this time, staff can also be placed in position. For example, surveyors monitoring skycap transactions at the curb would be taken outside, those monitoring kiosks would be shown the best location for monitoring those transactions, and those monitoring agents would be placed in their positions. Any personnel acting as "floaters" (i.e., surveyors who will be giving breaks to the primary surveyors) would need to understand the data collection process and location of each position.

7.3 Data Collection Case Studies

The previous chapters reviewed typical airport passenger-related processes, provided guidance on defining the research and developing sampling plans, and gave direction on determining when to sample and how to staff the survey. The previous sections of this chapter focused on reconnaissance, field testing, and training. In the next section of Chapter 7, a typical data collection event at an airline check-in counter is described. Finally, Sections 7.3.2 and 7.3.3 provide guidance for collecting data at bag claim and restrooms. Appendix E provides example templates for recording data using paper and pencil.

7.3.1 Collecting Data at an Airline Check-in Facility

This section includes a one-day data collection event at a moderately sized airport where check-in processing rates are being estimated. Based on an understanding of the project (See Chapter 3), the end users wish to compare processing rates among three airlines (versus simply obtaining an aggregate number). In addition, a distinction is to be made between first class and coach customers as well as those using express and full service e-ticket kiosks.

The peak hour, in terms of seat departures, for Airline A is between 6:30 a.m. and 7:30 a.m.; the peak hour for Airline B is between 7:00 a.m. and 8:00 a.m.; the peak hour for Airline C is between 8:30 a.m. and 9:30 a.m. Assuming passengers will reach the check-in counter about 90 min prior to their scheduled departure time (Section 5.6.3), it is estimated that the peak period of activity at the check-in counter will actually be between 5:00 a.m. and 6:00 a.m. for Airline A, 5:30 a.m. and 6:30 a.m. for Airline B, and between 7:00 a.m. and 8:00 a.m. for Airline C. Allowing for the desire to also collect demand profiles at each counter, one hour before and after each scheduled peak is applied. This means that the data collection schedule would be as follows:

- Airline A from 4:00 a.m. to 7:00 a.m.
- Airline B from 4:30 a.m. to 7:30 a.m.
- Airline C from 5:00 a.m. to 8:00 a.m.

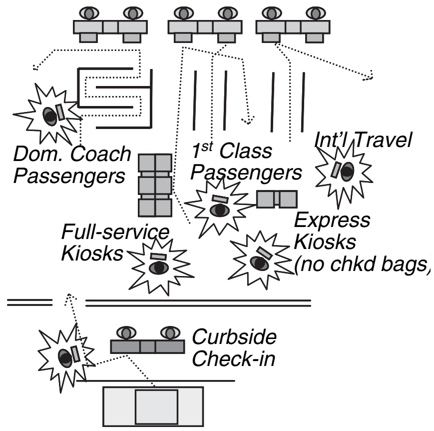


Exhibit 7-1. Airline check-in staffing locations for case study.

Sample sizes are calculated (Section 5.5), a sampling plan is developed, and a data collection technology and staffing method are selected (Section 6.1.4 and 6.3). A general, preliminary action plan is generated, and refined after a reconnaissance visit and field testing effort (Section 7.1.3).

Based on the scope of the effort, the various sources for staffing (airport staff, consultant staff, temporary agency staff, or market/research staff) are reviewed and a preferred source is identified.

Recognizing the early start times, the employees are asked to come the previous day for training and a walk through of the airline check-in facility (Section 7.2).

On the day of the survey, the survey personnel arrive about 30 min prior to the scheduled start time. The project manager distributes survey materials; positions staff (as shown in Exhibit 7-1); and reviews procedures.

Once data collection begins, the project manager allows 10–20 min to transpire and then revisits each surveyor to review data and address any remaining questions.

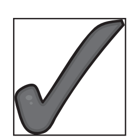
At the conclusion of the event, survey materials are collected, and surveyors are queried as to whether they observed anything out-of-the-ordinary that would warrant documentation.

7.3.2 Collecting Data for a Bag Transfer Timing Study

This section provides guidance from measuring the interval between an aircraft's arrival at the gate and the appearance of its bags at the bag claim device.

There are two methods that can be used. The first is having each surveyor collect data for a particular flight. In this instance, the data collector would travel to the gate to record an aircraft's arrival time; he/she would then head directly to the bag claim area to await the arrival of the first bag from that flight.

A second option would be to have two data collectors—one in the concourse and one at the claim area. The two would then relay information about a particular flight via cell phone. With this setup, however, it is imperative that their watches be synchronized for accuracy.



The passenger boarding time includes getting passengers onto the aircraft and seated and also having all their carry-ons stowed. Unless the survey effort is being undertaken directly by or for an airline, data collection will not occur onboard the sample aircraft, but rather at the gate; therefore, the definition of the start and finish time for boarding will, by necessity, be different.

When monitoring a flight, confirm the flight number and origin by asking a nearby gate agent or a deplaning passenger.

In this instance, each discrete activity should be recorded by time, expressed in hours, minutes, and seconds. Therefore, a PDA that can "time-stamp" their occurrence would be ideal. Items to record include the following:

- Time when gate door opens for boarding;
- Time of pre-board announcement;
- Time when each pre-board passenger enters the loading bridge;
- Time of first-class boarding announcement;
- Time when each first-class passenger enters the loading bridge;
- Time of each successive boarding announcement (whether by row or zone);
- Time when each passenger enters loading bridge;
- Time of final boarding announcements;
- Time when gate door closes; and
- Time when aircraft door closes (or when loading bridge is pushed away from aircraft).

If a PDA is not used, the time when each passenger enters the loading bridge should be noted, until the last passenger's entrance time is recorded.

Enplaning rates vary by aircraft size, composition of passengers, destination (e.g., domestic or international), and individual airline practices. It is recommended that your sampling plan take these variables into consideration, and if separate analyses of these factors are desired, the sampling plan should include stratification (Section 5.2.1).

For drawing a sample of departing flights only, the most straightforward method is to obtain an OAG schedule pull and randomly draw flights based on the strata you have identified.

For arriving flights, it is often much less efficient to identify sampled flights from an OAG due to the greater variability between a flight's scheduled and actual arrival times, and because there often seems to be a general lack of information regarding a particular flight's arrival gate, especially at large airports. For these reasons, it is recommended that a different sampling plan be used to obtain deplaning rates (see Section 5.2.6).

7.3.3 Collecting Restroom Use Data

Based on previous restroom data collection and planning efforts, it appears that restroom facility planning (in particular determining the number of fixtures required to accommodate demand) is most closely tied to simultaneous demand (i.e., how many fixtures are occupied at any given moment). A method for collecting this demand is described below.

It is recommended that collecting simultaneous demand data for restroom use be accomplished by teaming a male and female surveyor and having this pair enter the respective restrooms at regular intervals to record the number of fixtures occupied at that time and the number of people in queue. (An initial inventory of fixtures—including those which are out of order—needs to be conducted prior to the survey.) It is recommended that the observations be obtained at intervals no greater than 10 min because longer periods would increase the chance of missing the peak. Determining the actual time period for the survey should be based on looking at scheduled seat arrivals and seat departures, with more weight given to the peak arrival hour, since the deplaning process results in greater surges of restroom demand. It is recommended that the event begin about 1 hour before the anticipated peak hour and conclude 1 hour after the anticipated peak hour to minimize the chance of missing the actual peak. For example, if the scheduled peak hour is between 2:30 p.m. and 3:30 p.m., the data collection period should begin at 1:30 p.m. and end at 4:30 p.m.

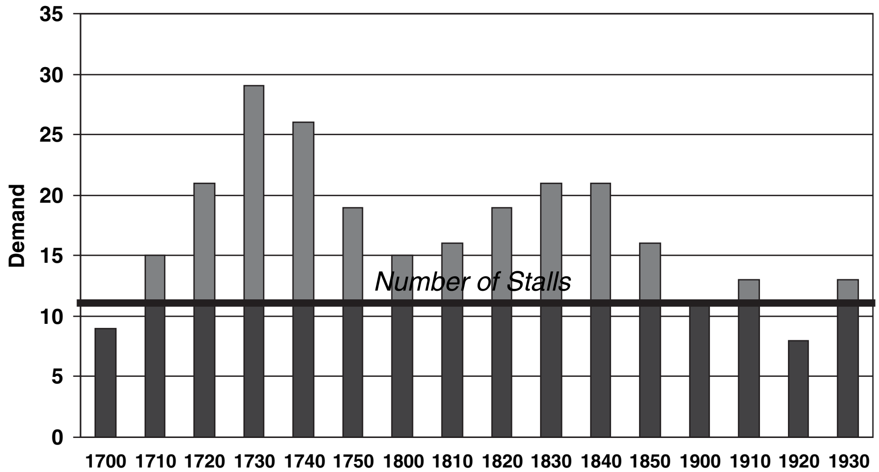


Exhibit 7-2. Example of simultaneous demand for women's restroom facilities.

Exhibit 7-2 shows an actual example of simultaneous demand results at a women's restroom at a large international airport. As shown, a capacity of 11 fixtures is clearly inadequate to accommodate the surges of demand, which reached 29 women at about 5:30 p.m. This resulted in a queue of more than 15 women.

CHAPTER 8

Summary

This chapter presents a highlight of the key factors that will lead to a successful airport data collection effort.

8.1 The Importance of Good Data

Recognizing the significant financial investment often required for terminal or process improvements, it is beneficial to support the decisions and planning with data that has been gathered in a technically rigorous manner. Doing so will also increase the likelihood of stakeholder "buy-in."

8.2 Typical Airport Passenger Processes

The most common processes observed at an airport include passenger check-in/ticketing, passenger and bag security screening, FIS (international arrivals processing), baggage claim, and airport boarding and deplaning. While some processes have stayed fairly static, others are characterized as dynamic due to evolving technology, security protocols, and changing passenger characteristics and behavior.

8.3 Confirm Need for Data and Feasibility of Data Collection

Although data collection costs are minor relative to construction costs, they are not immaterial. Additionally, time constraints may make the execution of a data collection effort very difficult. Finally, from a logistical perspective, there may be challenges to gathering data on some passenger-related processes due to security or proprietary issues. For these reasons, one should always explore whether extant data might already exist and consider the degree to which the data would be used in decision making in order to weigh the potential benefit of conducting the study against its cost.

8.4 Study Team

Airport data collection efforts are organized around a study team, including a project manager, survey manager, research and statistical expert, survey assistant, data analyst, administrative support, and data collection staff. Depending on the size and duration of the study and the qualifications of team members, some staff members may have multiple roles.

8.5 Airport Specific Metrics

When analyzing processes, one is largely focused on entities (typically, passengers and bags); resources (such as an agent or ticket kiosk); and processes (a function performed on an entity by a resource). The processing rate is defined as the number of entities processed by a single resource in a given unit of time and can only be expressed from the perspective of the resource. A processing time is the duration of a transaction from the perspective of either the entity or the resource. Failure to operationally define entities, resources, and processes at the start of a research project will limit the value of the collected data.

8.6 Sampling Techniques for Airport Data Collection

While there are many sampling approaches that can be used in gathering data at an airport, the overarching goal is to ensure the sample reflects the population of interest to allow the research to make inferences about the population. The challenge then becomes avoiding convenience samples (where individuals are included because they're available) and instead ensuring that each element in the sample has an equal (or at least known) chance of being selected.

The first step in obtaining a representative sample is defining the study population. This is accomplished by identifying the entities, resources, and processes of interest, the location, and the time period. When the desire is to obtain processing rate information, data collection is often scheduled during a peak period. Peak periods can be identified by examining airport activity statistics, published flight schedules, or asking informants (i.e., knowledgeable airport, airline, or government personnel).

The sample size is primarily based on balancing the amount of risk or error one is willing to tolerate relative to available resources (time and money). If one is interested in subgroups of passengers or processes, samples will need to be stratified, increasing the sample size and cost.

8.7 Action Plan

The action plan includes defining roles and responsibilities within the study team, framing relationships between the study team and sponsors/clients, determining who will staff the survey, determining how many staff are needed, selecting the data collection method, and scheduling specific data gathering periods.

There are several options for data collection staffing, including enlisting airport personnel, employing consultant staff, using college interns, hiring temporary staff, and using a market-research firm. Each has its advantages and disadvantages.

A typical study will need at least two months lead time, although this would increase directly with the complexity and magnitude of the data collection effort.

There are several data collection technologies, including basic paper and pencil; hand-held devices (PDAs); video monitoring; and video capturing software. The advantages and disadvantages of each method should be considered relative to the type, complexity, and amount of data being collected.

Attention should be given to providing the appropriate level of staff. Consideration should be given to not only the duration, complexity, and amount of data being collected, but to other issues, including facility layout, lines-of-sight, and anticipating contingencies.

8.8 Managing and Implementing Data Collection

It is strongly recommended that, to answer many of the questions above, a site visit and pilot test be done prior to actual data collection. This will allow refinement of the action plan.

To help ensure a successful data collection event, a staff training session is also necessary. During the training, the study purpose should be overviewed, procedural and management issues should be addressed, and actual training should be conducted.

Finally, during the actual data collection effort, the management team should monitor data collectors to make sure procedures are being followed and address any unforeseen issues that may arise.

APPENDIX A

Glossary

Airport Cooperative Research Program (ACRP)—Program authorized by Congress and sponsored by the FAA with the goal of developing near-term, practical solutions to problems faced by airport operators.

Bag Claim Device—Typically, a mechanical device designed to hold and display checked luggage for passengers to claim upon arriving at their destination airport.

Bag Tag—A tag placed on each piece of luggage checked-in by a passenger for tracking purposes. The tag usually displays the origin and destination airport, airline, flight number, a record identifier, and a UPC (bar code).

Bag Well—Area on either side of ticket counter used to hold, weigh, and tag checked luggage.

Belly Hold—Portion of aircraft below the passenger compartment frequently used to store luggage and cargo.

Common Use Self Service (CUSS)—Typically, device used to allow passengers to perform a particular function (e.g., check in for a flight), regardless of the airline on which they are traveling.

Customs and Border Protection (CBP), U.S.—Agency under the U.S. Department of Homeland Security (DHS) with the priority mission of keeping terrorists and their weapons out of the United States. It also has a responsibility for securing and facilitating trade and travel while enforcing United States' regulations, including immigration and drug laws.

Deplane (Deplanement)—Act of getting off an aircraft; passengers getting off an aircraft.

Domestic Travel—Typically, that portion of air travel within the borders of a particular country; may also include travel to pre-clear destinations within Canada and the Caribbean.

Duty—A tax on items brought into a country.

Enplane (Enplanement)—Act of boarding an aircraft; passenger getting on an aircraft.

Entity—Person (e.g., passenger) or object (e.g., luggage) being processed by a resource

Fare Class—Typically, premium or first class tickets and less expensive coach tickets.

Federal Aviation Administration (FAA)—Agency under the U.S. Department of Transportation, responsible for ensuring the safety of and the promotion of the aviation industry.

Note: This glossary provides definitions in the context of the aviation industry and associated activities.

Federal Inspection Service (FIS)—Facility operated by U.S. CBP, designed to process arriving international passengers and their luggage.

Finger Scan—Digital method of reading and recording fingerprints.

Hand Inspection—Process whereby TSA or CBP agent visually inspects an item and its contents; typically done as part of *secondary processing.*

International Travel—Typically, that portion of air travel outside the borders of a particular country.

Kiosk (Self-serve)—Small, stand-alone device which provides a service through an interactive computer screen.

Metal detector—A portal-like device used to detect the presence of metallic objects at airport passenger security screening checkpoints.

Master Plan—Document outlining the general, long-term development strategy for a facility to meet projected activity.

Non-revenue Passenger—Typically, an airline passenger working for the airline industry.

Official Airline Guide (OAG)—Provides a database for scheduled airline activity; available in hard copy (monthly) or electronically.

Operational Definition—Definition that specifies precisely how counts or measurements will be made. A description of a discrete function performed by a resource on an entity; the more precise the definition, the better the resulting data.

Originations—Passengers who are beginning their air travel at an airport, having arrived by some form of ground transportation.

Party (Travel)—An individual or group of passengers traveling together; definition (and therefore size) of travel party can vary among agencies, activities, etc.

Passenger Security Screening Checkpoint (PSSCP)—Operated by TSA, a screening checkpoint examines both passengers and their carry-on belongings for items that are banned from the passenger compartment of a commercial aircraft.

Pre-clear Airport—An international airport where passengers headed for the United States can go through the CBP process, thereby avoiding processing upon landing at their United States destination.

Pre-flighting—The process of inspecting and preparing a flight for departure—traditionally referring to the work of the flight crew.

Primary Inspection—Refers to the initial, general level of examination of passengers and/or their belongings; if something suspicious is found, the passenger or item is taken for *secondary inspection* processing.

Processing Rate—Number of entities that a single resource can process in a given unit of time.

Processing Time—Time interval between the beginning of a process on one entity and the beginning of a process on the next entity, assuming a constant rate of demand and a queue.

Protocol—A formal, prescribed procedure or process.

Resource—Typically a person (e.g., an agent) or mechanical device (e.g., kiosk or x-ray machine) performing a specific function on an entity.

Revenue Passenger—Passenger paying a fare on a flight; includes passengers traveling on redeemed frequent flier miles.

Secondary Inspection—Refers to a more detailed, thorough examination of passengers and/or their belongings as a result of an initial alarm or suspicion during the *primary inspection* process.

Take-Away Belt—Automated belt located behind most airline check-in counters designed to transport checked bags to a sorting facility for outbound luggage.

Terminations—Passengers who are ending their air travel at an airport and are leaving by some form of ground transportation. (Also, *destinations.*)

Through-the-Wall—At some smaller airports, checked bags are returned to passengers using a simple method of transferring bags from the airside to the claim area via secure openings in the wall.

Trace Detection—Process whereby a TSA or CBP agent wipes the exterior and handles of a bag with a special cloth and submits the sample into a special device used to test for explosive residue; typically done as part of *secondary inspection.*

Transportation Research Board (TRB)—Part of the nonprofit National Research Council; provides leadership in transportation innovation and progress through research and information exchange.

Ticket Counter/Check-in Counter—Portion of airport terminal where departing passengers purchase tickets, check in for flights, change itineraries, etc.

Transportation Security Administration (TSA)—Responsible for protecting the United States' transportation system; operates under the DHS.

Turn-time—The time interval between an aircraft's arrival at the gate and its departure. Typically refers to the minimum time needed to prepare an arriving aircraft for its outbound flight.

Wanding—Process whereby a TSA agent passes a hand-held *metal detector* around a passenger to check for metal objects; typically done as part of *secondary screening* process.

X-ray Device—Device used for viewing through solid objects to identify other objects that may lie underneath.

APPENDIX B

Analyzing and Reporting Data

B.1 Data Analysis: Introduction

Researchers analyze data for a number of reasons:

- To describe, summarizing data in accordance with previously defined requirements.
- To explore, searching for unanticipated findings, and a "story" the data might be revealing.
- To examine relationships or potential correlations.
- To compare groups.
- To develop and use prediction models.
- To test for hypothesized relations among the data.

Regardless of the reasons for analysis, a fundamental goal of data analysis is to reduce relatively large numbers of data values into a more succinct form, facilitating interpretation and reporting.

Consider Exhibit B-1. The first exhibit (B-1a) contains a number of observations in the order in which they were collected. The second (B-1b) presents the same data, sorted in ascending order.

The third exhibit (B-1c) presents the same data, grouped by range. Finally, the bar chart (B-1d), shows the number of observations per group. While transforming the data from an unsorted state to a sorted state helps somewhat in discerning meaning from the data, summarizing it in tabular and graphical format substantially helps. While a set of only seven observations, however, may not gain much from such reduction, examine Exhibit B-8. In this instance, approximately 23,400 data values have been summarized.

Summarizing the data using a frequency table (Exhibit B-2) permits one to see quite quickly that the region most represented is the North Central (35.9 percent), and that the regions with the lowest representation are New England and Pacific, at 2.1 and 2.2 percent, respectively. The same data are shown graphically in Exhibit B-3.

Tables and graphics can effectively help translate raw data into usable information by virtue of condensing and summarizing data. As common to most tools, however, there is a potential risk: the use of tables and graphs also makes it easy for researchers to unintentionally bias results, and for those of ethically questionably character, to intentionally misrepresent. The remainder of this chapter presents select recommendations for how to be effective in communicating results, and how to become more critical in evaluating what is presented *to* you.

B.2 Frequency Distributions

Data that are collected using a nominal level of measurement can be represented in frequency tables using the same categories in which the data were collected. For example, Exhibit B-4 summarizes gender data coded as male or female.

Unsorted
15
53
19
22
46
31
38

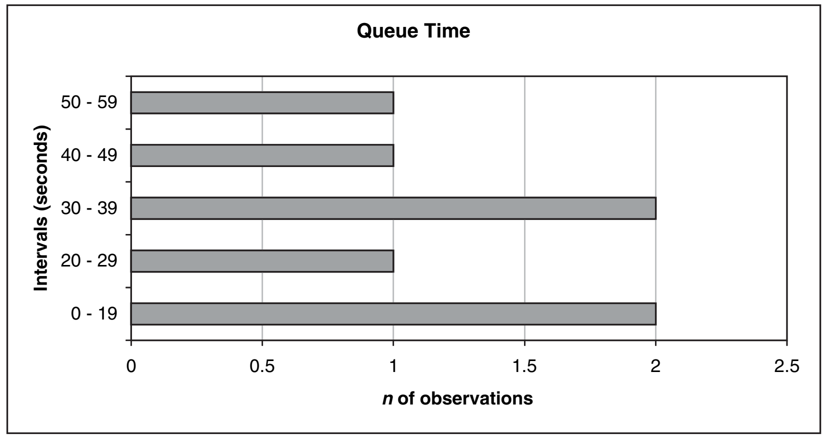
B-1a

Sorted
15
19
22
31
38
46
53

B-1b

Range	Freq.	Pct.	Cum. Pct.
0-19	2	28.6	28.6
20-29	1	14.3	42.9
30-39	2	28.6	71.4
40-49	1	14.3	85.7
50-59	1	14.3	100

B-1c



B-1d

Exhibit B-1. Data ordering and presentation.

Exhibit B-5, however, presents a different situation. Here, data were collected at a *scale* level of measurement, i.e., a stopwatch was used, and a recording of the time, rounded to the nearest second, was made. Whereas in Exhibit B-4 any given datum could reasonably only assume one of two states—male or female—the approximately 14,500 data points reflected in Exhibit B-5 could be organized in any number of ways. As represented here, less than 4 percent of the times were recorded to be less than 50 sec. In contrast, nearly 62 percent of the recorded times took between 51 sec and a minute.

Consider now the same data, categorized differently, shown in Exhibit B-6. With questionable ethics, a researcher presenting the data in this manner could legitimately assert something to the effect of: "Approximately 66 percent of all wait times were one minute or less." The moral is to maintain a healthy skepticism about the scheme the researcher chose for aggregating data. A common abuse occurs when data from rating scales are presented (e.g., satisfaction, quality). Responses labeled as "very satisfied" and "satisfied" are sometimes collapsed to obscure that relatively few respondents might have reported that they were indeed *very* satisfied.

Region

	Frequency	Valid Percent	Cumulative Percent
Middle Atlantic	1513	6.5	6.5
Mountain	4510	19.3	25.7
New England	502	2.1	27.9
North Central	8399	35.9	63.8
Pacific	519	2.2	66.0
South Atlantic	4377	18.7	84.7
South Central	3578	15.3	100.0
Total	23398	100.0	

Exhibit B-2. Frequency and percent of observations by region.

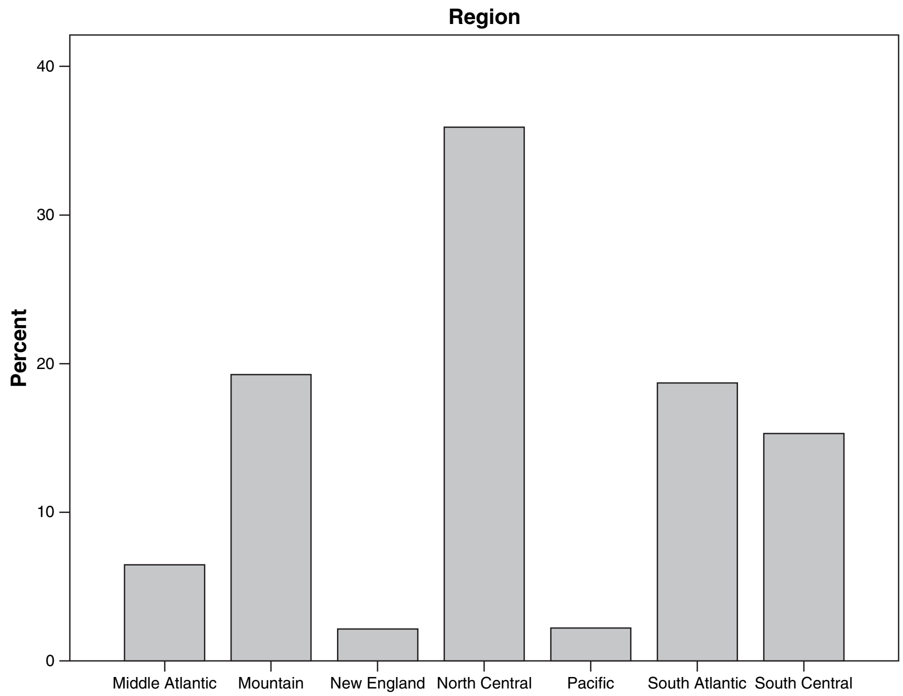


Exhibit B-3. Percent of observations by region.

Gender

	Frequency	Percent	Valid Percent	Cumulative Percent
Female	8821	60.6	60.6	60.6
Male	5729	39.4	39.4	100.0
Total	14550	100.0	100.0	

Exhibit B-4. Gender data.

	Frequency	Percent	Cumulative Percent
30 seconds or less	20	0.1%	0.1%
31 - 40 seconds	180	1.2%	1.4%
41 - 50 seconds	350	2.4%	3.8%
51 - 60 seconds	9000	61.9%	65.6%
61 - 75 seconds	1500	10.3%	75.9%
> 75 seconds	3500	24.1%	100.0%
Total	*14550*	*100.0%*	

Exhibit B-5. Restroom observations arranged by 10-second time groups.

	Frequency	Percent	Cumulative Percent
60 seconds or less	9550	65.6%	65.6%
61 - 75 seconds	1500	10.3%	75.9%
> 75 seconds	3500	24.1%	100.0%
Total	*14550*	*100.0%*	

Exhibit B-6. Restroom observations arranged by irregular time groups.

CEO	$ 400,000.00
President	$ 350,000.00
VP	$ 300,000.00
Worker 1	$ 80,000.00
Worker 2	$ 80,000.00
Worker 3	$ 75,000.00
Worker 4	$ 70,000.00
Worker 5	$ 55,000.00
Worker 6	$ 50,000.00
Worker 7	$ 40,000.00
Worker 8	$ 30,000.00
Worker 9	$ 25,000.00
Worker 10	$ 22,000.00
Average	$ 121,307.69
Standard Deviation	$ 128,199.59
Median	$ 70,000.00

Exhibit B-7. Annual employee salaries.

B.3 Averages and Medians

Exhibit B-7 presents fictitious data reflecting the annual salaries of 13 persons employed at a hypothetical company. Let's assume that during the course of interviewing a prospective employee, an unscrupulous interviewer, when asked about the average salary of persons employed by the company, responds "On average our employees earn more than $120,000 annually." Technically, this is an accurate response. Ethically however it has some problems. The reason the average salary is calculated to be in excess of $121,000 is that the average, or mean, is sensitive to extreme values, and the salaries of the CEO, the president, and the VP are extreme relative to the salaries of all other employees. If every employee earned $121,307.69, the average salary would be $121,307.69. If one reports only an average value the person to whom it is reported has no way of assessing how well that value reflects all the other values for which it stands. To be useful, at least two values need to be reported: the mean and standard deviation. As addressed in Chapter 5, the standard deviation is a measure of how much values, on average, are dispersed around the mean. The standard deviation of $128,199.59 shown in Exhibit B-7 suggests that, on average, salaries vary around the mean on average of plus or minus about $121,000.

A more accurate reflection of salary in the company is the *median,* which is the value midway between the highest and the lowest values. Whenever a distribution is markedly skewed, that is, lacking symmetry by leaning to the right or the left, the median provides a better summary statistic for how the data values cluster together.

So, if you want to consider yourself wealthy, attend a meeting at which Bill Gates is present, and estimate the average salary of those attending the meeting. On average, the average salary of those in the group, including your own, will likely exceed $1 million. *Voila:* now you are a multimillionaire.

B.4 Correlation

Correlation was referred to in the sampling section of the Guidebook. It is a way of quantifying two aspects of the relationship between variables: the strength of their association, and the directionality of the relationship. While methods have been developed for looking at relationships between variables at different levels of measurement, the Guidebook will limit consideration to the most commonly known statistic, r, sometimes referred to as Pearson's r, or Pearson's

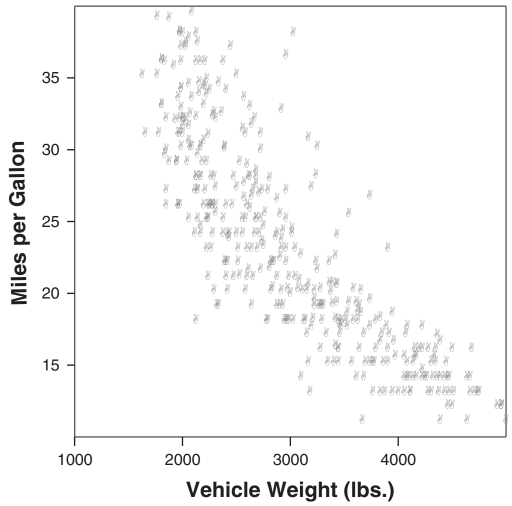


Exhibit B-8. Plot of miles per gallon and vehicle weight.

product moment correlation coefficient. It is named for the British statistician, Karl Pearson in the early years of the 20th century.[1]

Pearson's r can assume any value between −1 and +1. A graphical representation of the relationship between two variables, MPG and weight in pounds, of automobiles manufactured in the 1970s is shown as Exhibit B-8[2].

Not surprisingly, there is a negative relationship between a vehicle's weight and the average gas mileage it achieves. The *r* value for these data is −0.83, a relatively strong, albeit negative correlation: as automobile weight increases, gas mileage decreases.

B.4.1 Correlation and Causation

A colleague relayed the following, likely apocryphal, story. She was teaching an introductory statistics course at a small university in Ohio. While addressing the topic of correlation one of her students reported on an interesting finding. He had discovered a surprising relationship between annual pig iron production in the birth rate of pigs raised in the state. This story is similar to that related by Duckworth (2004) concerning a statistically significant correlation between the number of births in a town and the number of stork nests. In the birth-rate story, one might infer some meaningful relationship in so far as the larger the town is, the higher the birth rate, and, in turn, the greater the number of chimney stacks. Such sites also are an apparently desirable location for storks to build their nests. Simply, be wary about drawing causal linkages. As the popular asserts, "correlation does not imply causation."

B.5 Misleading Graphical Displays

Exhibits B-9 and B-10 represent two line graphs. The data used to construct the graphs are the same. There is, however, one difference. Note that the scale in Exhibit B-9 ranges from 64 to 84, whereas the scale in Exhibit B-10 ranges from 0 to 100.

[1] Clapham, C. & Nicholson, J. (2005). *Concise dictionary of mathematics, 3rd edition.* New York: Oxford University Press,
[2] Adapted from SPSS V. 15.

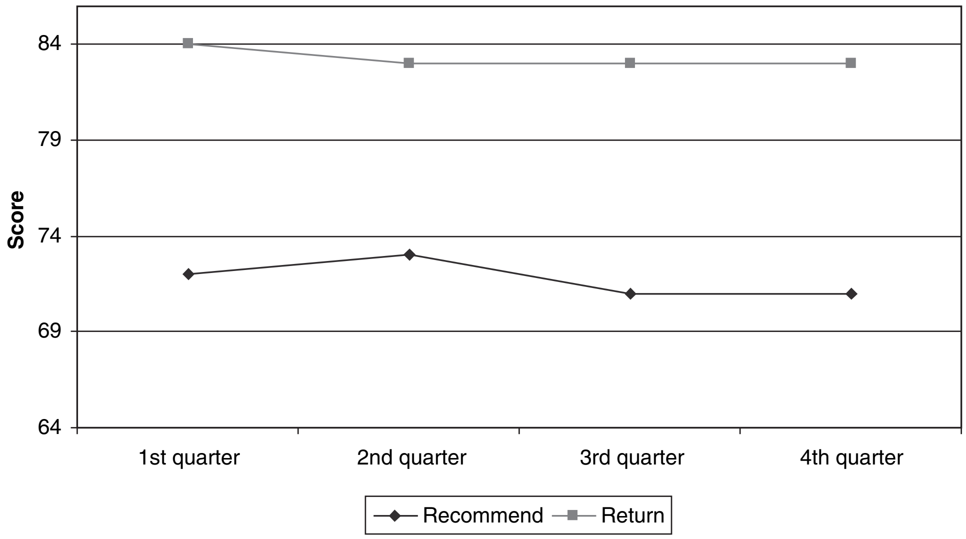


Exhibit B-9. Data graphed using inappropriate scale.

While we have eliminated identifying information, Exhibit B-9 is a facsimile of what was provided to a client by a consultant who had been secured to measure the performance of the Company's website. The scale employed by the consultant was 100 points, yet interestingly the consultant chose to plot the data using a somewhat compressed scale.

Exhibit B-10 plots the same data using a 100 point scale. You'll notice that the lines in Exhibit B-10 are essentially straight, indicating no change in average performance across time. Whether Exhibit B-9 was created in ignorance or with intent to misrepresent, is not clear. The lesson, however, is hopefully clear.

A related issue is incorrectly inferring meaningful variation when indeed the variation is simply random or "noise." Tracking changes in the Dow Jones industrial index on an hourly basis may be interesting, but the vantage is so close that it obscures the proverbial forest for the trees. For example, a decrease in passenger waiting times based on data collected over a brief

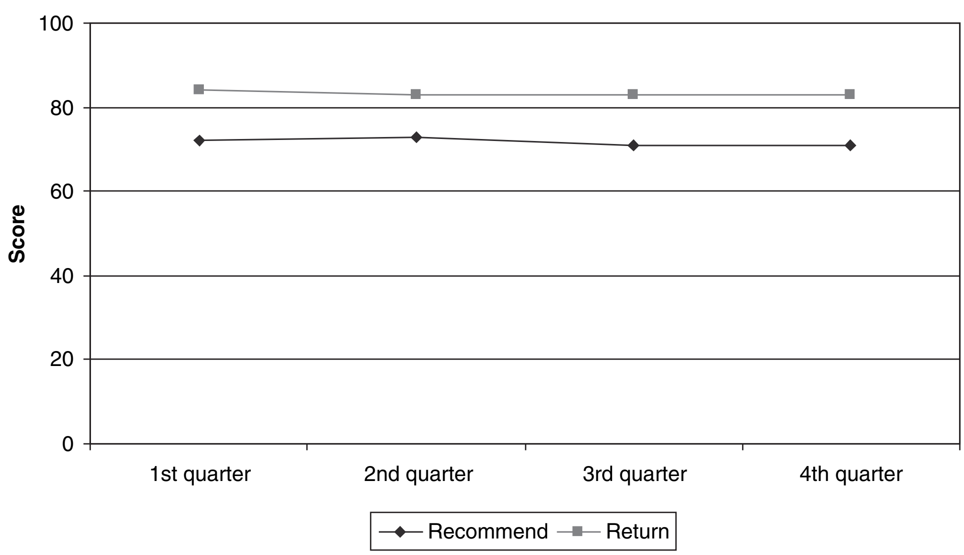


Exhibit B-10. Data graphed using appropriate scale.


U.S. Debt Expected To Soar This Year

$2 Trillion Increase May Test Federal Ability to Borrow

By Lori Montgomery
Washington Post Staff Writer

With President-elect Barack Obama and congressional Democrats considering a massive spending package aimed at pulling the nation out of recession, the national debt is projected to jump by as much as $2 trillion this year, an unprecedented increase that could test the world's appetite for financing U.S. government spending.

For now, investors are frantically

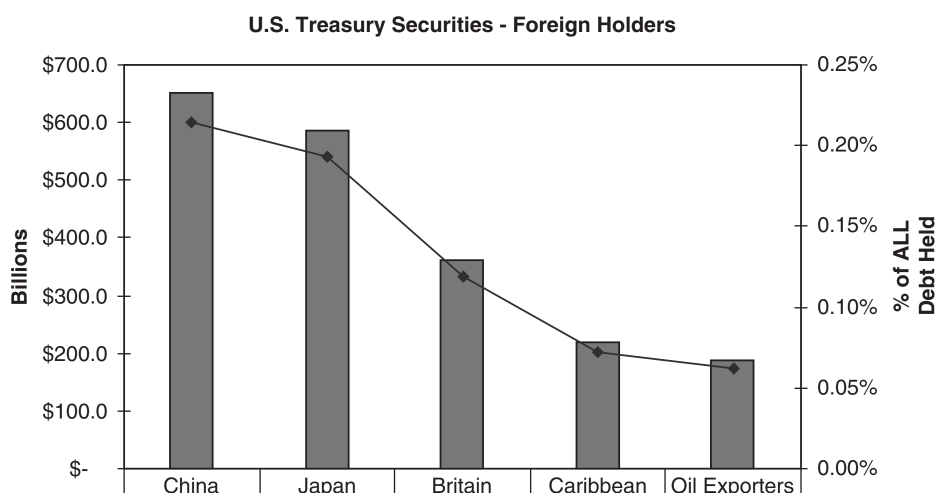
Source: The Washington Post

Exhibit B-11. Example of complex chart.

and statistically inadequate sample might result in an unjustifiable conclusion that a process is genuinely improving.

Exhibit B-11 comes from the January 9, 2009, *Washington Post.* The height of each stack of money is intended to represent the debt, in billions, associated with five geographic areas. The pie chart is intended to represent the proportion of debt for the five geographical regions as well as all other holders of debt.

The same data are presented in Exhibit B-12. While decidedly less embellished than Exhibit B-11, the graphical representation is more straightforward, and, in turn, less suspect to misinterpretation.



	China	Japan	Britain	Caribbean	Oil Exporters
Amount	$652.9	$585.5	$360.2	$219.5	$187.7
Percent	0.21%	0.19%	0.12%	0.07%	0.06%

Exhibit B-12. Simplified presentation of data.

Accuracy	Perceptual Feature
Most Accurate	Position along common scale
↑	Position along identical, non-aligned scales
	Length
↓	Slope
	Angle
	Area
	Volume
Least Accurate	Color hue; saturation; density

Exhibit B-13. Relative accuracy of various graphic methods.

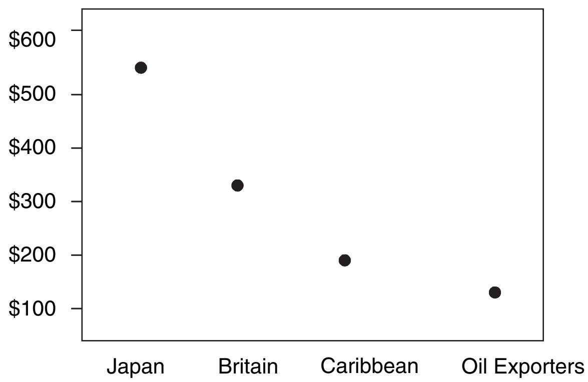


Exhibit B-14. Data on a common scale.

While there is admittedly disagreement among those who conduct research into the graphical display of data,[3] and, as such, no recommendation is without some controversy, the following are recommendations based on *The Elements of Graphing Data.*[4] Based on a number of studies assessing people's accuracy in discriminating proportions, Cleveland rank ordered perceptual factors in descending order of accuracy, as documented in Exhibit B-13.

To illustrate, Exhibit B-14 is an example of data presented on a common scale. In this instance, the data are the same as for four of the geographic regions shown in Exhibit B-12. Note that the information can be captured by a single point aligned with the scale on the *y* axis.

The same data are represented in Exhibit B-15. Here, however, the viewer is required to make a judgment about length.

Exhibit B-16 depicts three pie-chart variations. To interpret a pie-chart, the viewer is tasked with judging variations in angle, a task prone to inaccuracy as suggested in Cleveland's research.

In the bottom left representation, only the labels are shown, not the percentages. The reader might consider how the accuracy of his or her estimate of the actual percentages for Britain and the Caribbean as shown in the top graphic.

---

[3] Spense, I. (2005). No humble pie: The origins and usage of a statistical chart. *Journal of Educational and Behavioral statistics,* 40,4, 353–368.

[4] Cleveland, W. S. (1985). *The elements of graphing data.* Pacific Grove, CA: Wadsworth.

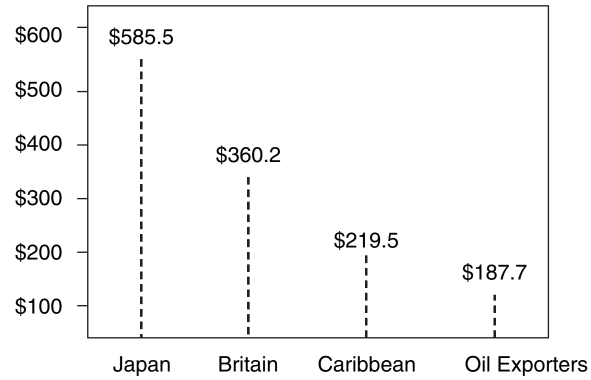


Exhibit B-15. Data on a fixed scale.

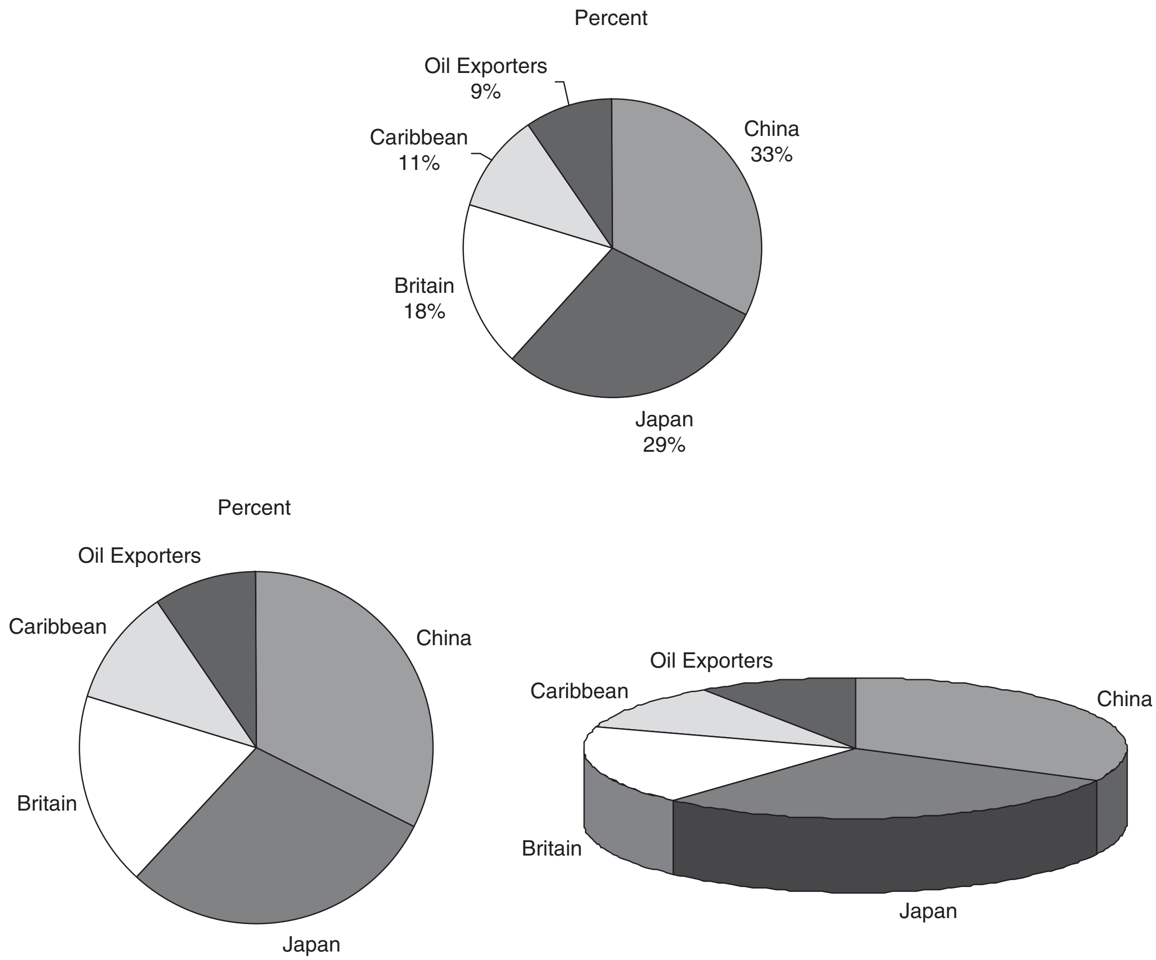


Exhibit B-16. Pie chart examples.

Whereas the percentages do convey substantially more information than the pie-chart without values, one might reasonably ask what value the graphic itself adds to the information.

The final pie-chart is presented as an egregiously poor graphic, providing the reader with a 3-dimensional rendering that interferes with information.

B.6 Threats to Internal Validity

Internal validity is a measure of the extent to which a given outcome can be ascribed to a particular reason.

Let's assume that you are planning to evaluate the value of a leadership training program marketed by a vendor, the results of which will drive your decision whether or not to purchase the training for deployment to all of your employees. What factors, independent of the training itself, might impact the results? Let's say that you decide to place your "best" employees into the program to really give it "a run for its money." A problem arises, however, in that these employees may not be typical. As an alternative, you consider an open enrollment. Given a limited number of seats in the program, however, those who rush to enroll may also not be typical employees. In both situations something may be "getting in the way," potentially moderating the outcome. When the stakes are high, for example when medications are prescribed to a patient, potential interactions are often emphasized. (Think about all those stickers adorning your prescription medicine the next time you go to your medicine cabinet.)

What's the bottom line? Be wary of people and organizations who tout universal solutions. They may not be telling you the whole story. While printed ads for weight loss products often contain a disclaimer, in very small print, noting that the loss of weight described in the accompanying testimonials may not be typical results, we have not encountered any consultants or service companies that take a similar approach. Another lesson is to be very careful in analyzing data to not attribute some outcome to a particular cause.

APPENDIX C

Technical Appendix

C.1 Is the Study Needed? (Additional Considerations)

A member of the Research Team regularly teaches a graduate course in research methods; most enrollees are working professionals, many with considerable experience. One of the course's requirements is to develop a research plan to answer a question of the student's choosing. Semester after semester, virtually all students begin with a faulty assumption, i.e., the question can not be answered with existing data, and if collecting data is warranted, a new data collection instrument must be developed. Often, even if the proposed study is justified, students begin by defining the proposed study's purpose as proving that a particular practice works or does not work; advocacy research, however, is not defensible.

While these assumptions are inappropriate within the context of scientific research, they are actually quite understandable. Formal research requires us to approach the world in a way that is, for most people, unnatural. When we interact with others in normal work or social situations, for example, we generally accept what another asserts as "true," not requiring that evidence be presented or that the methodology used in reaching conclusions be defined and justified. If, as an illustration, your spouse comes home after a particularly trying day at work stating that "my boss is a complete jerk," your relationship might not be affected if you accept the assertion as accurate, and provide emotional support. Asking for evidence, scrutinizing the evidence as potentially biased, etc., might have a different, and likely not particularly pleasant, impact on your relationship. Formal research, however, operates with a different set of rules than those used in daily living.

C.2 Sampling (Additional Considerations)

C.2.1 Stratification

Stratified sampling is useful when meaningful differences exist within a population, but there is a downside: stratifying increases sample size, with concomitant increases in time and money. The importance of stratifying is also reflected in the extent to which the data collected might reasonably be generalized to other circumstances. Distinguishing, for example, between a certain type of space usage at a hub airport and an O&D airport implies that the two are so different from one another that treating them the same would result in an increase in unexplained variation.

Exhibit C-1 is a checklist to help you consider potentially relevant factors to consider in planning your sample. Keep in mind that only differences that impact the metric should be selected. Identify what makes the elements different from one another.

If you determine that stratification is appropriate, how to stratify can be done in several ways. There are statistical methods that might be used to define the number of strata to use, but these

Dimension	Example Factors to Consider in Stratifying Samples
Entities	• Passengers • Well-wishers • Bags
Where observations will be made	• Check-in • Security screening • At gate during boarding • At gate during deplaning • Bag claim • Hub airline terminal/non-hub terminal
When observations will be made	• Season • Month • Day of week • Hour(s) of day
What are the characteristics of event?	• Domestic flight • International flight • Hub airline flight • Spoke airline flight • Departing flight • Arriving flight • Restroom usage • Concession usage

Exhibit C-1. Example factors to consider in planning your sample.

are beyond the scope of this Guidebook. We suggest that you use experience and professional judgment in selecting strata.

If the sampling is not stratified by class, you can describe the average transaction times for the different classes, but additional work needs to be done if you wish to make a generalization about each group. This additional work is referred to as *post-stratification.*

C.2.2 Two-Stage Cluster Sampling

When a random sampling technique is used to select a subset of elements within a cluster, the technique is referred to as *two-stage cluster sampling.* Ideally, clusters should be microcosms of a population, reflecting the heterogeneity within it. A departure gate at a given terminal, for example, might be defined as a cluster insofar as it seems to represent the variation in type of passengers, baggage, processing times, etc. that one might expect at any similar gate. As an illustration, the World Health Organization (WHO) employs a variation of cluster sampling in evaluating immunization coverage across geographic areas (Hoshaw-Woodard, S. 2001).[1]

C.2.3 Error

Before concluding that the variation observed is simply due to chance, consider that your categories may be too large or ambiguous to explain the variation. In other words, some of the variation you see in the data might not be due to chance; it may be explainable if you refine your approach. To illustrate, when the Research Team reviewed some passenger check-in data from one airport, we noticed that often the expected value (mean or average) was not a very good statistic to represent a range of values. Attempting to make the data as "homogenous" as possible, we limited the times recorded to those transactions of one passenger, checking two

[1]Hoshaw-Woodard, S. (2001). *Description and comparison of the methods of cluster sampling and lot quality assurance sampling to assess immunization coverage.* Geneva: Department of Vaccines and Biologicals, World Health Organization.

Observations	*Line 1*	*Line 2*	*Line 3*
1	3.01	3.0	7.0
2	3.01	3.3	0.8
3	3.01	2.7	5.5
4	3.01	3.2	1.5
5	3.01	3.0	0.5
Average Processing Time	3.01	3.04	3.05
Standard Deviation	0	0.23	2.99

Exhibit C-2. Sample data set.

bags, on the same type of airline. The data were collected over 4 days in September 2004. The average processing time for the 75 cases or so was 3.06 min. The standard deviation, however, a measure of dispersion around the mean was 1.95 min. Simply, average variation around the mean of 3 min was approximately 1 min to 5 min.

Exhibit C-2 depicts fictitious data for transaction times in minutes for three queue lines. For Line 1, note that every transaction took exactly 3.01 min. As such, the average is 3.01, and, since there is no variation, the standard deviation is 0. Line 2 has a similar average (3.04), but notice that the standard deviation is 0.23 min. Line 3 represents a more extreme case. While the mean is essentially the same as for the other two lines, (3.05), the standard deviation (2.99) is almost as large as the mean. The reason is that the processing times recorded for line 3 varied widely, from a low of 0.5 min to 7.0 min.

While the data from Line 1 is unrealistic, the times recorded for Line 2 are somewhat realistic. There is some variation, but not much. In other words, this is likely random variation. Line 3, however, is another matter. When two of the processing times took 5.5 min or longer, and two took less than a minute, attributing the wide range to chance is likely not the best approach.

C.2.4 Calculating Error

From each value is subtracted the average value and each of these deviations is then squared. When the squared deviations are summed, the result is the sum of the squared deviations around the mean. To calculate the average deviation, this sum is divided by $n - 1$.[2] While this is a fine and useful measure of average variation around the mean, the number is no longer in the same units of measurement. For example, if the average time were 75 sec and the variance were 16, this is not 16 seconds. To get back to the original units of measurement, one must undo the squaring by taking the square root of the variance. The square root of 16 is 4 sec. Exhibit C-3 lists potential sources of error and potential solutions.

Within Excel, a number of statistical functions are available. As illustrated in Exhibit C-4, descriptions and the format of the formulas is provided.

In Section 5.4.2, reference was made to an approach described by van Bell (2002). In particular, the recommendation is to make approximately 15 observations and then divide the range by the number of observations, an approach that is sufficiently robust to accommodate distributions both normal and substantially different from normal (e.g., uniform). Exhibit C-5 illustrates this approach for a sample of five observations.

[2]The reason for using n – 1 rather than 1 is a bit complicated. Technically, each time an average is calculated it constrains one way in which the values are free to vary. Since a mean is being calculated here, one degree of freedom is subtracted, or n – 1.

Source of Error	Comments	Potential Solutions
Design		
o Failure to collect the appropriate data or incorrect data	o Sponsor has not adequately defined study purpose.	o Document the principal reason(s) the data are being collected.
	o Sponsor has not evaluated which data might best help answer the research question.	o Subject proposed approach to review by experts.
	o Sponsor has not adequately communicated with the research team.	o Benchmark against other studies of acknowledged quality.
o Ambiguous terminology	o Don't assume that common term will mean the same to all.	o Develop operational definitions.
Methodology		
o Sampling design problems	o See Section 5.3.	o Use randomized sampling techniques.
		o Ensure that the sample size is adequate to support anticipated output.
o Instrumentation problems	o At a most fundamental level, the mechanism for collecting the data, regardless of the medium, should have a one-to-one correspondence to the type and levels of data defined in the research design.	o Audit the data collection instrument for completeness and accuracy.
Implementation		
o Personnel	o Personnel should be adequately prepared to ensure that their understanding of what to do, when to do it, and how to do it are consistent with the research plan.	o Ensure that you have an adequate number of properly prepared people to complete the task. (See Chapter 6.)
	o Even with the proper preparation, needed equipment, time to complete the task, space to do so, etc., human error is unavoidable.	o Ensure that you have an adequate number of properly prepared people to complete the task. (See Chapter 6.)

Exhibit C-3. Potential sources of error and potential solutions.

The *range* is the difference between the largest and the smallest values. In this case, the range is 52 – 38, or 14. Since there are 5 observations, the estimated *standard error (SE)* is 2.8. An estimate of a sample standard deviation (SD) can then be obtained by multiplying the estimated SE by the square root of the sample size. The square root of the sample size, 5, is 2.2. The product of 2.2 and the SE (2.8) is approximately equal to 6.3. The estimated variance, therefore, or the square of the SD, is equal to 39.7.

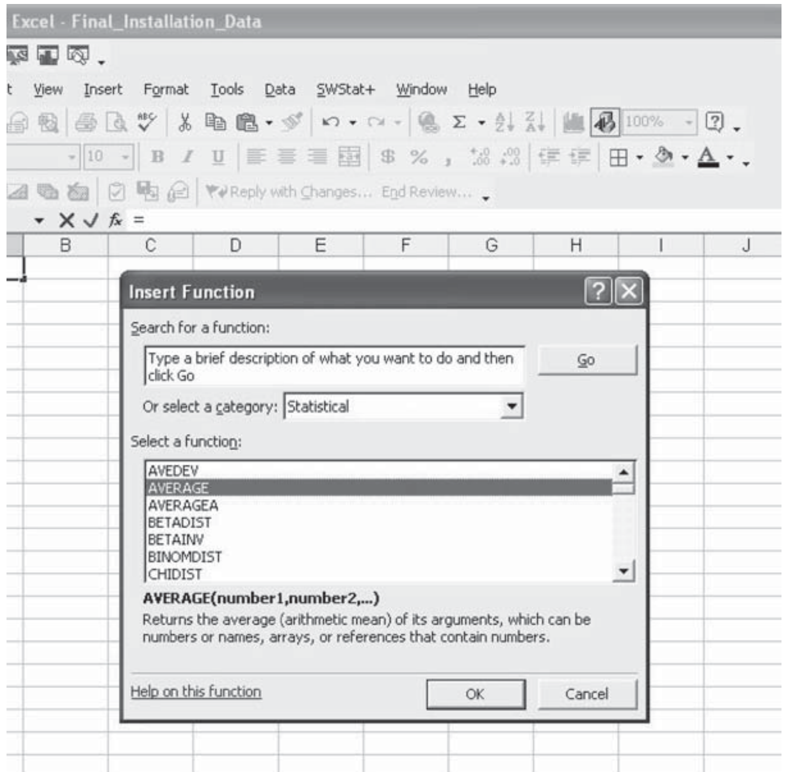


Exhibit C-4. Formulas in Excel.

C.2.5 Variance

A sample is a proxy or a surrogate for a population. Since only those elements selected for the sample will be analyzed, rather than all elements of the population, the results of the analysis from a sample will, to one extent or another, differ from what would result from analysis of the entire population. The actual, but unknown, attributes of a population are referred to as *population parameters,* and include, for example, a population's mean, median, standard deviation, and so forth. Not only is a sample a stand-in for a population, but the results of the analysis, known as sample statistics, are estimates of the population's parameters.

As described in Chapter 5, the choice of metric relates to whether you are principally interested in reporting frequencies or averages. If you are counting the number of persons traveling

Observations	Values (in seconds)
1	38
2	46
3	52
4	40
5	43

Exhibit C-5. Sample of five observations for example of using the Van Bell method of estimating population variance.

	Flight Type	
Count per time	Domestic (1)	International (2)
T_1	125	40
T_2	153	35
T_n	119	45
Total	**397**	**120**

Exhibit C-6. Nominal level data.

domestic and international, for example, you will report frequencies and percentages (*nominal* level data). Exhibit C-6 is an example of nominal level data which would be reported as frequencies or percentages. Approximately 77 percent of all passengers counted across the time periods were flying on domestic flights, while approximately 33 percent were flying internationally. The numerical codes to the right of each heading might be what would be entered when the data are recorded.

Note that while one could legitimately report the *mode* (most frequently occurring), which in this case is domestic, or 1, it would not be appropriate to report the average. The average flight type would be 1.23, calculable but meaningless. Were you measuring the amount of *time* persons waited in line, however, you would likely be calculating the average time in queue across a number of people (interval or scale level data).

In these examples, a single value is calculated to represent some aspect or parameter of a distribution. For example, 77 percent of passengers are flying domestic; the average time in queue is 3.4 min. This single numerical value, calculated from sample data, is called a *point estimator* of a population parameter (e.g., the actual percentage in a population flying domestic, or the actual average wait time of a population). Another approach to estimating population values from samples is the use of *confidence intervals.* Rather than calculating a single, point value, a confidence interval defines the lower and upper bounds within which one has a given level of confidence that, were repeated samples taken, the interval would include, or capture, the true population mean. Using flight time as an example, one might calculate an interval ranging between 2.8 min and 4.0 min, with a 95 percent confidence that the true population mean would fall within this interval.

The sample statistic of the mean or average ($\overline{X}$ pronounced X-Bar) is an estimator of the population mean parameter, μ (the Greek letter MU). How well the sample statistics actually represent population parameters is a function of the three elements previously identified: tolerable error, confidence, and random (or unexplained) variance. Exhibit C-7 represents these factors graphically using an Influence Diagram. The diamond shaped box, sample size, represents the desired outcome. The four square boxes represent factors over which you have control. The circle, labeled variation, reflects inherent variation in what you are observing. It is something over which there is really no control, but is rather a chance event. Brief descriptions of each follow.

C.2.6 Error Types: Legal Example

Exhibit C-8 depicts the two types of errors one can make using a legal situation. The true state of nature is that the person did or did not commit a crime. Given that the person asserts her innocence, and there is no perfect knowledge but only evidence that points to her having committed

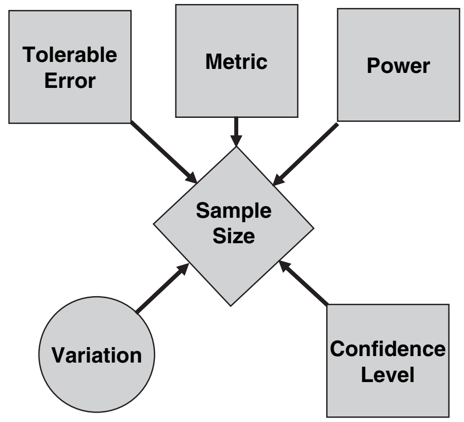


Exhibit C-7. Factors affecting sample size.

	The "True" State	
Verdict	**Innocent**	**Guilty**
Not Guilty	Correct Decision	Type II error
Guilty	Type I error	Correct Decision

Exhibit C-8. Example of Type I and Type II errors.

the crime, she is brought to trial. The legal use of the term "not guilty" is interesting insofar as it reflects that we cannot with absolute certainty assert that she is innocent. The best that we can do is to declare the evidence insufficient to find her guilty. She may be innocent, but one cannot with certainty assert this.

C.2.7 Calculating Sample Size: Reporting Proportions

$$n = \frac{Z^2 p(1-p)}{e^2}$$

Exhibit C-9. Formula to determine sample size when reporting proportions.

The formula shown in Exhibit C-9 might look a bit intimidating, but it is actually pretty straightforward, and is built using the terms introduced earlier in this chapter.

Each element of the formula is defined below:

n = sample size
Z = the level of confidence desired. (See Exhibit C-10.)
p = the estimated proportion of what you wish to sample that is present in the population
$1 - p$ = the estimated proportion of what you wish to sample that is not present in the population
e = the level of error one is willing to tolerate

Confidence Level	*Z value*
90%	1.645
95%	1.96
99%	2.58

Exhibit C-10. Confidence level and corresponding Z values.

C.2.8 Calculating Sample Size When Assessing Change

One might be interested in assessing whether modification of an existing process reduced average processing time by a target goal. If the target goal was set at 20 percent, and the estimated variability in observations was 30 percent, the sample required would be approximately 30 as calculated below. Note that the coefficient of variation (CV) is the ratio of the standard deviation to the mean.

$$n = \frac{8(CV)^2}{(PC)^2}\left[1 + (1 - PC)^2\right]$$

PC = proportionate change in means
CV = Coefficient of variation

$$CV = \frac{\sigma_0}{\mu_0}$$

$$n = \frac{8(0.30)^2}{(0.20)^2}\left[1 + (1 - 0.20)^2\right] = 29.52 \approx 30$$

Exhibit C-11 shows illustrative sample sizes as a function of the size of the change one is attempting to detect, and the CV.

C.2.9 Calculating Sample Size When Assessing Relationships

In the previous example, sample size was calculated based on the detection of a change in an average value. In some instances, one might be interested in detecting a given correlation. A correlation coefficient is a measure of the strength and direction of a linear relationship between two variables, and is typically expressed as a value ranging from −1.0 to 1.0. This is approached using the following formulas:

$$n = \frac{16}{\Delta^2}$$

where

$$\Delta = \frac{1}{2}\ln\frac{1+\rho}{1-\rho}$$

where:

ln = natural log of a number
The ln of the constant e (2.71828183) = 1
ρ (rho) = the correlation to be detected

CV	Percent Change in Mean						
(in Pct)	5	10	15	20	30	40	50
5	16	4	2	1	1	1	1
10	61	15	7	4	2	1	1
15	137	33	14	8	3	2	1
20	244	58	25	14	6	3	2
30	548	131	56	30	12	7	4
40	975	232	98	53	22	11	7
50	> 1,000	362	154	82	34	17	10
75	> 1,000	815	345	185	75	39	23
100	> 1,000	> 1,000	613	328	133	68	40

CV = Coefficient of Variation
Two-sample test, two-sided alternative hypothesis; Type I error = 50; Power = 80

Exhibit C-11. Sample size for assessing change.

To illustrate, what sample size is needed to detect a correlation of 0.40? Using the aforementioned approach, approximately 41 observations would be needed for a single sample. For two samples, the numerator would be 16.

$$1.4/0.6 = 2.33$$

$$\ln(2.33) = 0.8458$$

$$0.5 * 0.8458 = 0.4423$$

$$n = \frac{8}{0.4423^2} \approx 41$$

Sample Size: Stratified Samples
Unweighted
Weighted
Post stratification

Sample Size: Cluster Sampling
1 stage
2-stage

C.2.10 Calculating Sample Size When Varying Sample Costs

In some situations, the cost of varying types of sample might vary considerably. As such, the formula shown below can be used to select sample size.

$$\frac{n_1}{n_0} = \sqrt{\frac{c_0}{c_1}}$$

Essentially, one selects sample sizes that are inversely proportional to the square root of the cost of the observations.

As an example, assume that the cost per observation for the first sample is $160, and the cost per observation for the second sample is approximately $40. Using the formula shown above, 160/40 = 4, the square root of which is 2. As such one would use twice as many observations for the second group as for the first.

C.2.11 Calculating Sample Size When Testing for Differences in Averages

Assumptions:

- **Two averages of two groups are being compared.**
- **Two-sided alternative/research hypothesis.** This means that the null hypothesis is that the means are equal to one another. By contrast, sometimes one uses a one-sided, or one-tailed, approach, proposing in the research hypothesis that results will be in a given direction. The following are examples of each.
  - **Two tailed research hypothesis:** The average ticket processing times for first class and coach passengers are different (or not equal).
  - **One tailed research hypothesis:** The average ticket processing time for first class passengers is less than the average processing time for coach passengers.
- **Normal distribution.** The population distribution resembles the "bell curve." See Exhibit C-12. The curve shown at the top is a *normal distribution.* In this example, the mean is 100, and the

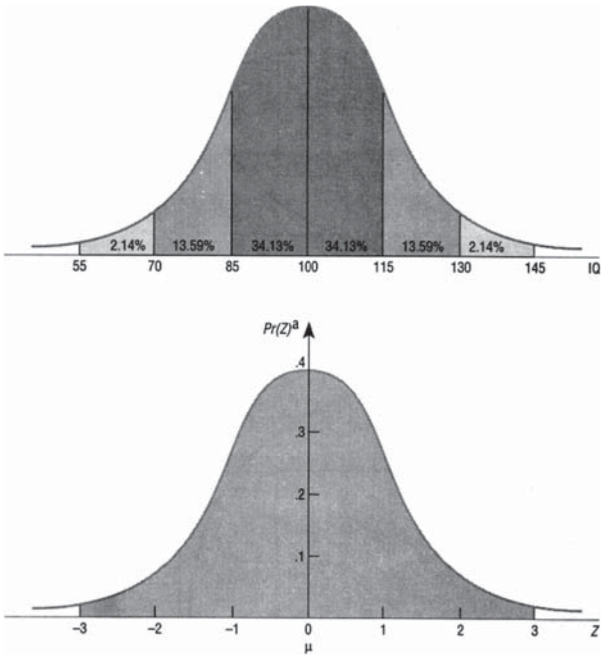


Exhibit C-12.

standard deviation is 15. The areas are approximations. For example, approximately 68% of the observations of a random, continuous variable x would be expected to fall between 85 and 115. The curve shown at the bottom is a standardized normal distribution. When standardized, the data are converted from their original measurement units to standard units in which the mean is always equal to 0, and the standard deviation is always equal to 1.

– **Homogeneity of variance.** This means that the variation of values in the two distributions is assumed to be approximately the same.
– **The variance of the population is known.**

Exhibit C-13 shows the appropriate numerator to use under different circumstances in the following formula. The values shown in bold reflect the customary practice of setting Power at .80.

		Numerator for sample size equation
Type II Error β	Power 1 - β	2 Sample
0.50	0.50	8
0.20	**0.80**	**16**
0.10	0.90	21
0.05	0.95	26
0.025	0.975	31

Exhibit C-13. Appropriate numerator to use.

$$n = \frac{16}{\Delta^2}$$

where

$$\Delta = \frac{\mu_0 - \mu_1}{\sigma} = \frac{\delta}{\sigma} \text{ (Variance known), or}$$

$$\Delta = \frac{4}{\sqrt{n}} \text{ or } \mu_0 - \mu_1 = \frac{4\sigma}{\sqrt{n}} \text{ (Variance unknown)}$$

Two sided alternative hypothesis; Type I error, $\alpha = 0.05$

For example, in comparing two groups, with power set as 0.80, hypothesizing no difference in the group means, and with variances approximately equal, e.g. 35%, a sample size of approximately 131 would be required.

$$n = \frac{16}{.35^2} = \frac{16}{0.1225} \approx 131 \text{ per group}$$

If variance is unknown

Detectable standardized difference in, for example, a two-sample case, is approximately 4 divided by the square root of the number of observations per sample.

APPENDIX D

Sample References from Research Literature

Correia, A. R., & Wirasinghe, S. C. (2008). Analysis of level of service at airport departure lounges: User perception approach. *Journal of Transportation Engineering, 134*(2), 105.

Correia, A. R., Wirasinghe, S., & de Barros, A. G. (2008). A global index for level of service evaluation at airport passenger terminals. *Transportation Research. Part E, Logistics & Transportation Review, 44*(4), 607.

Daniel, J. I., & Harback, K. T. (2008). (When) do hub airlines internalize their self-imposed congestion delays? *Journal of Urban Economics, 63*(2), 583.

Drexl, A., & Nikulin, Y. (2008). Multicriteria airport gate assignment and Pareto simulated annealing. *IIE Transactions, 40*(4), 385.

Fayez, M. S., Kaylani, A., Cope, D., Rychlik, N., & Mollaghasemi, M. (2008). Managing airport operations using simulation. *Journal of Simulation, 2*(1), 41.

Feng, Q. (2007). On determining specifications and selections of alternative technologies for airport checked-baggage security screening. *Risk Analysis, 27*(5), 1299.

Fodness, D., & Murray, B. (2007). Passengers' expectations of airport service quality. *The Journal of Services Marketing, 21*(7), 492.

Hsu, C., & Chao, C. (2005). Space allocation for commercial activities at international passenger terminals. *Transportation Research. Part E, Logistics & Transportation Review, 41E*(1), 29.

Lin, M. H. (2006). Hub-airport competition: Connecting time differentiation and concession consumption*. *Australian Economic Papers, 45*(4), 299.

Rhoades, D. L., & Waguespack, B., Jr. (2008). Twenty years of service quality performance in the US airline industry. *Managing Service Quality, 18*(1), 20.

Seneviratne, P. N., & Martel, N. (1995). Space standards for sizing air-terminal check-in areas. *Journal of Transportation Engineering, 121*(2), 141.

Tseng, K., Ho, J., & Liu, Y. (2008). A study on the performance evaluation of major international airports in the world. *Journal of Modeling in Management, 3*(1), 71.

Waguespack, B., Jr, Rhoades, D. L., and Tiernan, S. (2007). An investigation into airline service quality performance between U.S. legacy carriers and their EU competitors and partners. *Journal of Air Transportation, 12*(2), 59.

Whalen, W. T., Carlton, D. W., Heyer, K., & Richard, O. (2008). A solution to airport delays. *Regulation, 31*(1), 30.

Zhang, T., Ouyang, Y., & He, Y. (2008). Traceable air baggage handling system based on RFID tags in the airport. *Journal of Theoretical and Applied Electronic Commerce Research, 3*(1), 106.

APPENDIX E

Example Training and Orientation Materials

HARTSFIELD-JACKSON ATLANTA INTERNATIONAL AIRPORT

PEAK WEEK 2008 TERMINAL OBSERVATIONS

PRIMER

Background

Hartsfield-Jackson Atlanta International Airport (H-JAIA) is the busiest airport in the world. Last year, 89 million passengers used the airport and 994,000 flights landed or took off. The City is currently in the midst of a multi-year, multi-billion dollar development effort designed to provide the facilities needed to meet long-term growth.

A significant data collection effort is being undertaken which consists of a departing passenger survey, MARTA survey, other surveys, and terminal observations. The surveys and observations are being conducted in July 2008 (a busy month) and will involve surveying passengers on approximately 280 departing flights and observing various functional elements of the terminal to obtain data on processing rates and quality of service. A staff of more than 120 people will be participating with up to 40 surveyors and data collectors working simultaneously.

Field Objective

The objective of Peak Week 2008 is shown below:

***** To accurately collect the data needed to help long-term planning efforts. *****

General Procedures

Our success at meeting this objective will require all team members to:

Bring a good attitude—Depending on your assignment, you may be here early in the morning or late in the evening or on weekends; you may be stuck out in a place where the sun is beating down on you; you may be asked to do something as simple as click a tally counter for three hours. Take it all in stride and roll with it! :)

Remember to bring your security ID badge—If you've been asked to get a badge, remember to bring it; if you don't have your badge, you may not be able to participate and you won't get paid.

Bring something to write with—In fact, bring several things to write with.

Wear appropriate clothing—Wear comfortable shoes and clothes that are professional but washable. Some events are outside—and it's hot out there! Also, minimize jewelry and accessories because all that stuff has to be removed and screened when going through security. It will just slow you and everyone else down.

Be on time—The days and start/stop times of each event have been carefully coordinated to occur at a scheduled peak period and we have no cushion for makeup. If you're late, you may end up holding up dozens of your teammates or cause a teammate to have to double-up and do your share of the work, possibly jeopardizing that portion of the study. Give yourself time for highway traffic backups, MARTA delays, time to find a parking space (lots are often full or even closed in the summer), using the restroom, etc. Also, it takes time to do training and to position staff. If you're late, you won't know where to go because the rest of the team will already be at the site and you won't know what to do. And someone will have to go back and get you and position you and train you. (And that someone is usually me, and I'm getting old and tired!)

If you can't make it or are going to be late, call your supervisor as soon as you see trouble brewing—Your supervisor's primary responsibility to me is to ensure that the number of people they committed to providing are there at the specified time and ready to go. By telling your supervisor.

Listen to and explicitly follow directions—If you misunderstand what you're asked to do, all your data may end up being unusable. Don't be afraid to ask for clarification. Also, for most of the observations, there is a precedence for how the data are collected, so it would be a mistake for you to start doing something differently. However, if what you're being asked to do doesn't make any sense to what's happening on the ground, trust your instincts, make adjustments, and note accordingly on the form.

Write neatly—Remember, someone has to translate your data into a spreadsheet. Hard-to-read writing slows down data entry and may be misinterpreted; impossible-to-read writing means your data won't be useable.

Make sure you know how to use the PDA—This year, we're using handheld data collection devices (personal digital assistants or PDAs) to collect a lot of our data. The devices have multiple programs on them, so be sure you're using the right one.

Put your name, date, company name, and location on all forms—If we don't know where and when your data was collected, or even what exactly you were timing or counting, we can't use it. But if your name and company are on the form, and we need you to clarify something, we'll be able to get in touch with you at a later date.

Pick a good spot for your observations—one with a good line-of-sight—Don't be afraid to fine-tune your position when recording data. If someone gets in your way, move or ask the person to move. You must be able to see whatever you're supposed to be observing, otherwise, your error rate and amount of lost data will go up.

Minimize distractions when recording data—Don't use your cell phone, do your homework, or engage in conversations with others.

If you're timing things, use your stopwatch and make sure to record the time in hours, minutes, and seconds and specify AM or PM—Don't use your cell phone because it may not be in synch with everyone else's stopwatches. Also, most cell phones don't have seconds on them, so you will not be able to effectively time short-lived events. Don't use an analog watch/clock (i.e., with a face and hands)—How do you time a car stopping for 45 seconds with a minute hand?

If you're making tallies (𝍸), record them the right way.

Feel free to jot down occasional notes—For example, if there's a big line of people waiting to be checked-in and the line is blocking people from going in and out of doors, make a note of it. These kinds of notes help pepper the report with interesting pieces of information.

Stay for the duration of the scheduled data collection event—If you're scheduled to stay at your post until 2:30 p.m., stay until 2:30 p.m. and then head back to the office.

If you're scheduled to be relieved and your replacement is late, try to stay at your post if at all possible.

Coordinate with neighboring teammates to ensure you're not double-counting something or not counting something you're supposed to.

Promptly return to designated meeting place at conclusion of data collection event—The longer your data is "out there," the higher the probability it will get misplaced. Also, your clipboard, stop watch, tally counter, or PDA may be needed for another event.

HARTSFIELD-JACKSON ATLANTA INTERNATIONAL AIRPORT

PEAK WEEK 2008 TERMINAL OBSERVATIONS

Contact Information

Name:________________________ **Name:**________________________

Company:_____________________ **Company:**_____________________

Phone:_______________________ **Phone:**_______________________

Name:________________________ **Name:**________________________

Company:_____________________ **Company:**_____________________

Phone:_______________________ **Phone:**_______________________

Name:________________________ **Name:**________________________

Company:_____________________ **Company:**_____________________

Phone:_______________________ **Phone:**_______________________

Name:________________________ **Name:**________________________

Company:_____________________ **Company:**_____________________

Phone:_______________________ **Phone:**_______________________

Name:________________________ **Name:**________________________

Company:_____________________ **Company:**_____________________

Phone:_______________________ **Phone:**_______________________

APPENDIX F

Sample Data Sheets

Sample Airline Check-in Form
(To collect individual transaction data)

Name: ______________ Company: ______________ Date: ______________

Airport: ______________ Airline: ______________ Fare Class: ______________

Domestic/International: ______________ Counter Position: ______________

Obs.	Party Size	Checked Bags	Reach Agent	Leave Agent	Notes
1			: :	: :	
2			: :	: :	
3			: :	: :	
4			: :	: :	
5			: :	: :	
6			: :	: :	
7			: :	: :	
8			: :	: :	
9			: :	: :	
10			: :	: :	
11			: :	: :	
12			: :	: :	
13			: :	: :	
14			: :	: :	
15			: :	: :	
16			: :	: :	
17			: :	: :	
18			: :	: :	
19			: :	: :	
20			: :	: :	
21			: :	: :	
22			: :	: :	
23			: :	: :	
24			: :	: :	
25			: :	: :	

Observations must be made sequentially and only when there is a queue.

Sample Airline Check-in Form
(To collect gross processing rate data)

Name: ____________________ Company: ____________________ Date: ____________________

Airport: ____ Airline: ____________________

Dom./Int'l/Mixed: ____ Fare Class: ____________________

(Use Only When Queue is Observed)

No. of Agents Monitored	Start Time	End Time	Tot'l Pax Processed	Notes
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		
	: :	: :		

Sample Security Screening Form
(To collect individual Processing Time data)

Name: ____________ Company: ____________ Date: ____________

Airport: ____________ Location ____________ Lane: ____________

Conditions: ____________

Obs.	Party Size	Bags/Bins	1st Thru Metal Detector	Last Thru Metal Detector	E=Employee C=Crew P=Passenger	Notes
1			: :	: :		
2			: :	: :		
3			: :	: :		
4			: :	: :		
5			: :	: :		
6			: :	: :		
7			: :	: :		
8			: :	: :		
9			: :	: :		
10			: :	: :		
11			: :	: :		
12			: :	: :		
13			: :	: :		
14			: :	: :		
15			: :	: :		
16			: :	: :		
17			: :	: :		
18			: :	: :		
19			: :	: :		
20			: :	: :		
21			: :	: :		
22			: :	: :		
23			: :	: :		
24			: :	: :		
25			: :	: :		

Observations must be made sequentially and only when there is a queue.

Sample Form for Determining Number and Types of Items X-rayed

Passenger Security Screening Detail
Items Per Person

Surveyor: ______________________ **Firm:** ______________________

Date: ______________________

Obs.	Passenger/ Employee	NO (0) ITEMS	Carry-on/ Backpack	Purse/ Hand Bag	Laptop	Laptop Case	Coat/Jacket Hat/Gloves	Belt Shoes (Pr)	Watch/ Jewelry	Cell Phone/ PDA/ Camera/ Electronics	Wallet/ Keys/ Change	ZIPLoc Bag	Other
1													
2													
3													
4													
5													
6													
7													
8													
9													
10													
11													
12													
13													
14													
15													

Directions: For each person approaching security station: 1) Identify whether passenger (or meeter/greeter or well-wisher) or employee; 2) Note how many of each items the person places on the x-ray belt (a pair of shoes should be counted as one item). If person places no items, note that on sheet.

FIS SURVEY
Immigration Processing Rates (Individual Parties)

Name: ______________________ Date: ____________________

Agent A

☐ U.S. Residents
☐ Visitors
☐ Crew

Obs.	Party Size	Reach Agent
1		: :
2		: :
3		: :
4		: :
5		: :
6		: :
7		: :
8		: :
9		: :
10		: :
11		: :
12		: :
13		: :
14		: :
15		: :
16		: :
17		: :
18		: :
19		: :
20		: :
21		: :
22		: :
23		: :
24		: :
25		: :

Agent B

☐ U.S. Residents
☐ Visitors
☐ Crew

Obs.	Party Size	Reach Agent
1		: :
2		: :
3		: :
4		: :
5		: :
6		: :
7		: :
8		: :
9		: :
10		: :
11		: :
12		: :
13		: :
14		: :
15		: :
16		: :
17		: :
18		: :
19		: :
20		: :
21		: :
22		: :
23		: :
24		: :
25		: :

Observations must be made sequentially, and with queue.

Restroom Occupancy Counts

Date: ______________________ Name: ______________________ Concourse: __________ ☐ Male ☐ Female

Location 1: ______________________

Available Fixtures

Urinals __________
Stalls __________
Sinks __________

Occupied Fixtures

Time	Counts: Urinals	Stalls	Sinks	Queue
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				

Location 2: ______________________

Available Fixtures

Urinals __________
Stalls __________
Sinks __________

Occupied Fixtures

Time	Counts: Urinals	Stalls	Sinks	Queue
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				

Location 3: ______________________

Available Fixtures

Urinals __________
Stalls __________
Sinks __________

Occupied Fixtures

Time	Counts: Urinals	Stalls	Sinks	Queue
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				
:				

Abbreviations and acronyms used without definitions in TRB publications:

AAAE	American Association of Airport Executives
AASHO	American Association of State Highway Officials
AASHTO	American Association of State Highway and Transportation Officials
ACI–NA	Airports Council International–North America
ACRP	Airport Cooperative Research Program
ADA	Americans with Disabilities Act
APTA	American Public Transportation Association
ASCE	American Society of Civil Engineers
ASME	American Society of Mechanical Engineers
ASTM	American Society for Testing and Materials
ATA	Air Transport Association
ATA	American Trucking Associations
CTAA	Community Transportation Association of America
CTBSSP	Commercial Truck and Bus Safety Synthesis Program
DHS	Department of Homeland Security
DOE	Department of Energy
EPA	Environmental Protection Agency
FAA	Federal Aviation Administration
FHWA	Federal Highway Administration
FMCSA	Federal Motor Carrier Safety Administration
FRA	Federal Railroad Administration
FTA	Federal Transit Administration
IEEE	Institute of Electrical and Electronics Engineers
ISTEA	Intermodal Surface Transportation Efficiency Act of 1991
ITE	Institute of Transportation Engineers
NASA	National Aeronautics and Space Administration
NASAO	National Association of State Aviation Officials
NCFRP	National Cooperative Freight Research Program
NCHRP	National Cooperative Highway Research Program
NHTSA	National Highway Traffic Safety Administration
NTSB	National Transportation Safety Board
SAE	Society of Automotive Engineers
SAFETEA-LU	Safe, Accountable, Flexible, Efficient Transportation Equity Act: A Legacy for Users (2005)
TCRP	Transit Cooperative Research Program
TEA-21	Transportation Equity Act for the 21st Century (1998)
TRB	Transportation Research Board
TSA	Transportation Security Administration
U.S.DOT	United States Department of Transportation