

Session # 12

Innovative Data Collections to Analyze North Carolina's Airport Activities

By

John Karnowski, P.E. and Marsha Dale Anderson

Street Smarts
3400 McClure Bridge Road
Duluth, Georgia 30096
770-813-0882 (voice)
770-813-0688 (fax)
jkarnowski@freewwwweb.com
manderson@atlanta.com

For
Transportation Research Board
Sixth National Conference on Transportation Planning
for Small and Medium-Sized Communities

September 16-18, 1998

Spokane, WA

ABSTRACT

The North Carolina Department of Transportation, Airport Division, contracted with the University of North Carolina, Charlotte (UNCC) and Street Smarts to collect origin-destination (O-D) data at all 14 North Carolina airports. UNCC was responsible for the project management and data analysis. Street Smarts was charged with the task of collecting the data.

When it comes to [data collection](#) of any kind, public officials care most about one thing, "How can we get the most useable data for the money we have budgeted?" It is a reasonable question that needs an innovative answer. There are really three elements in that question that demand attention: quality of data, quantity of data, and cost of getting that information. The North Carolina Airport [O-D study](#) was an exercise in achieving an optimal balance of all three.

PROJECT SCOPE

A crucial step in satisfying a public agency is defining a clear scope up front that is beneficial to all parties. For data collection, the scope is as large as the budget allows. The negotiating process is familiar to anyone in the public or private sector:

- The client has a number of records needed to address statistical validity requirements
- The data collection company gives a price for that quantity under a certain set of assumptions
- The client revises its number to reduce the cost and still maintain confidence in the ultimate results
- The data collector revises its price under the same set of assumptions
- The client can't reduce its cost and either 1) walks away from the project thinking there is no way to get the analysis done, or 2) hires someone who will work for the money and risk getting sub-par data, or 3) helps develop a new methodology for collecting the data
- The data collector either 1) walks away thinking there is no profit in the job, or 2) becomes more efficient, creative and innovative in its methodology
- Assuming both work together to reach a solution, the project proceeds using the new methodology.

This type of negotiation could apply to just about any public/private job. In the case of the North Carolina Airport Study, the data collectors (Street Smarts) already had the experience of working with new technology to obtain more information at a substantially reduced cost.

The Airport Division wanted to determine in what manner and for what reason people were traveling by air throughout North Carolina. An [airport](#) O-D study is the most common way of determining this information. However, it was felt that a meaningful study would require interviews at all 14 commercial airports in the state over 4 seasons of a year which ended up being 87 days of surveying. The airport planners who would eventually use the information set a target minimum of 6,500 records.

To put this figure into perspective, some agencies have paid up to \$150 per record for this type of information. *Source: SBIR Final Report, produced by Street Smarts for FHWA, July 1997.*

The common method for collecting Airport O-D information is to:

1. Position interviewers at the airport departure and arrival gates
2. Interview passengers in the waiting area or baggage claim area
3. Record the responses on pre-made form
4. Code the responses into a database at the home office
5. Spot check the data entry
6. Clean up responses and delete "bad" records

It became clear early on that some things would have to be different from the “common” methodology.

PLANNING FOR THE STUDY

Survey “Form”

NCDOT Airport Division and UNCC were charged with developing the questionnaire which would be used in the passenger interviews. The length of the survey form is a factor in the number of records that can be obtained. Every question adds a little more time to the survey and reduces the number of passengers who can be interviewed.

The questions generally followed national guidelines used by USDOT and the FAA for origin-destination studies. However, there were also questions pertaining specifically to North Carolina and its visitors. A response to a question might branch to another set of questions before continuing. And, every [survey](#) would have to have the name of the airport, the airline surveyed, and a date/time stamp.

Figure 1 shows the survey, as it would appear on paper. However, a paper form was not the preferred method of collecting the data.

Figure 1. NC Airport Survey Questionnaire

<p>1. How did you arrive into the airport today?</p> <p><input type="checkbox"/> Transfer from other flight city</p> <p><input type="checkbox"/> Flew here (arriving flight)</p> <p><input type="checkbox"/> Drove self</p> <p><input type="checkbox"/> Given a ride</p> <p><input type="checkbox"/> Public transit</p> <p><input type="checkbox"/> Taxi/Limo</p> <p><input type="checkbox"/> Shuttle service</p> <p><input type="checkbox"/> Rental car</p> <p><input type="checkbox"/> Other</p> <p><input type="checkbox"/> Don't know</p> <hr/> <p>2a. Which airport did you start your trip from? <i>Arriving passengers only!</i></p> <hr/> <p>2b. If you came by private car, was the vehicle...</p> <p><input type="checkbox"/> Not parked, but driven away</p> <p><input type="checkbox"/> Parked at meter</p> <p><input type="checkbox"/> Parked in an airport parking lot</p> <p><input type="checkbox"/> Parked in a commercial lot</p> <p>3. What type of place was your starting location?</p> <p><input type="checkbox"/> Your workplace</p> <p><input type="checkbox"/> Other workplace/business</p> <p><input type="checkbox"/> Your home</p> <p><input type="checkbox"/> Other home</p> <p><input type="checkbox"/> Hotel/motel</p> <p><input type="checkbox"/> Other</p> <p>4. What is your starting location?</p> <p>City/County _____</p> <p>Zip _____</p> <p>Intersection _____</p> <p>Address _____ Num: _____</p> <p>Street: _____</p> <p>5. What is the purpose of this trip?</p> <p><input type="checkbox"/> Business</p> <p><input type="checkbox"/> Commute to/from work</p> <p><input type="checkbox"/> Visit friend/relative</p> <p><input type="checkbox"/> Personal activities</p> <p><input type="checkbox"/> Recreation/vacation</p> <p><input type="checkbox"/> Other</p> <hr/> <p>6. Which airport will you end your trip?</p> <hr/> <p>7. What type of place is your ending location?</p> <p><input type="checkbox"/> Your workplace</p> <p><input type="checkbox"/> Other workplace</p> <p><input type="checkbox"/> Your home</p> <p><input type="checkbox"/> Other home</p>	<p>8. How many people are in your group?</p> <hr/> <p>9. How often do you make this round trip between these two place?</p> <table style="width: 100%; border: none;"> <tr> <td style="width: 50%;"><input type="checkbox"/> <1/year</td> <td style="width: 50%;"><input type="checkbox"/> 1-2/week</td> </tr> <tr> <td><input type="checkbox"/> 1/year</td> <td><input type="checkbox"/> 3-4/week</td> </tr> <tr> <td><input type="checkbox"/> 2/year</td> <td><input type="checkbox"/> 5+/week</td> </tr> <tr> <td><input type="checkbox"/> 3-5/year</td> <td><input type="checkbox"/> Other</td> </tr> <tr> <td><input type="checkbox"/> 6-12/year</td> <td><input type="checkbox"/> Don't know</td> </tr> <tr> <td><input type="checkbox"/> 1-3/month</td> <td></td> </tr> </table> <p>10. How will you get to your final location from the airport?</p> <p><input type="checkbox"/> Driving self</p> <p><input type="checkbox"/> Getting a ride</p> <p><input type="checkbox"/> Public transit</p> <p><input type="checkbox"/> Taxi/limo</p> <p><input type="checkbox"/> Shuttle service</p> <p><input type="checkbox"/> Rental car</p> <p><input type="checkbox"/> Don't know</p> <p><input type="checkbox"/> Other</p> <hr/> <p>11. What improvements would you like to see:</p> <hr/> <p>12. If you are a visitor, what are your daily average expenditures at this location?</p> <p><input type="checkbox"/> \$1-50</p> <p><input type="checkbox"/> \$51-150</p> <p><input type="checkbox"/> \$151-200</p> <p><input type="checkbox"/> \$200+</p> <p>13. Gender</p> <p><input type="checkbox"/> Male</p> <p><input type="checkbox"/> Female</p> <p>14. Age Category</p> <table style="width: 100%; border: none;"> <tr> <td style="width: 33%;"><input type="checkbox"/> <18</td> <td style="width: 33%;"><input type="checkbox"/> 18-20</td> <td style="width: 33%;"><input type="checkbox"/> 21-30</td> </tr> <tr> <td><input type="checkbox"/> 31-40</td> <td><input type="checkbox"/> 41-50</td> <td><input type="checkbox"/> 51-60</td> </tr> <tr> <td><input type="checkbox"/> >60</td> <td></td> <td></td> </tr> </table> <p>15. How many people live in your household?</p> <hr/> <p>16. What is the total before tax income of your household (in \$1,000's)?</p> <table style="width: 100%; border: none;"> <tr> <td style="width: 33%;"><input type="checkbox"/> <\$20</td> <td style="width: 33%;"><input type="checkbox"/> \$20-39</td> <td style="width: 33%;"><input type="checkbox"/> \$40-59</td> </tr> <tr> <td><input type="checkbox"/> \$60-79</td> <td><input type="checkbox"/> \$80-99</td> <td><input type="checkbox"/> \$100-119</td> </tr> <tr> <td><input type="checkbox"/> \$120-139</td> <td><input type="checkbox"/> \$140+</td> <td></td> </tr> </table>	<input type="checkbox"/> <1/year	<input type="checkbox"/> 1-2/week	<input type="checkbox"/> 1/year	<input type="checkbox"/> 3-4/week	<input type="checkbox"/> 2/year	<input type="checkbox"/> 5+/week	<input type="checkbox"/> 3-5/year	<input type="checkbox"/> Other	<input type="checkbox"/> 6-12/year	<input type="checkbox"/> Don't know	<input type="checkbox"/> 1-3/month		<input type="checkbox"/> <18	<input type="checkbox"/> 18-20	<input type="checkbox"/> 21-30	<input type="checkbox"/> 31-40	<input type="checkbox"/> 41-50	<input type="checkbox"/> 51-60	<input type="checkbox"/> >60			<input type="checkbox"/> <\$20	<input type="checkbox"/> \$20-39	<input type="checkbox"/> \$40-59	<input type="checkbox"/> \$60-79	<input type="checkbox"/> \$80-99	<input type="checkbox"/> \$100-119	<input type="checkbox"/> \$120-139	<input type="checkbox"/> \$140+	
<input type="checkbox"/> <1/year	<input type="checkbox"/> 1-2/week																														
<input type="checkbox"/> 1/year	<input type="checkbox"/> 3-4/week																														
<input type="checkbox"/> 2/year	<input type="checkbox"/> 5+/week																														
<input type="checkbox"/> 3-5/year	<input type="checkbox"/> Other																														
<input type="checkbox"/> 6-12/year	<input type="checkbox"/> Don't know																														
<input type="checkbox"/> 1-3/month																															
<input type="checkbox"/> <18	<input type="checkbox"/> 18-20	<input type="checkbox"/> 21-30																													
<input type="checkbox"/> 31-40	<input type="checkbox"/> 41-50	<input type="checkbox"/> 51-60																													
<input type="checkbox"/> >60																															
<input type="checkbox"/> <\$20	<input type="checkbox"/> \$20-39	<input type="checkbox"/> \$40-59																													
<input type="checkbox"/> \$60-79	<input type="checkbox"/> \$80-99	<input type="checkbox"/> \$100-119																													
<input type="checkbox"/> \$120-139	<input type="checkbox"/> \$140+																														

Survey Instrument/Tool

Based on research done by Street Smarts for the USDOT, Federal Highway Administration, a handheld, pen-based computer was the preferred survey instrument.

A pen computer is light enough to be carried around all day. Most small pen-based computers are under 2 pounds. The one employed in the North Carolina O-D study was 20 ounces. The computer is powered by three AA batteries. The small battery power means lightweight and a streamlined CPU.

Some handheld computers are not pen-based. They are small notebook computers in a clam shell design, with a small keyboard. That design makes it uncomfortable to see and hold in one hand all day. Most pen-based computers have just a screen, a few hot key buttons and use a pen for all entry.

The computer is programmed with the same questions as the paper form with additional intelligence. The major difference in the recording of responses is the computer guides the surveyor through the interview. The order of the questions is laid out for the surveyor.

The automatic time stamping of records was found to be a great advantage. The automatic time stamping allowed the user to concentrate fully on the interviewee.

The program has intelligence built into it so it “knows” what to do after a particular response. In the North Carolina Airport Study, if a respondent arrived at the airport via personal automobile, the next questions pertained to where the vehicle was parked. Obviously, if the person came by taxi, his/her vehicle was not parked at the airport. This type of branching is easily accomplished using the computer but can be difficult to train using a paper form. (For an example of a confusing, paper-form branching scheme, refer to any IRS tax form.)

A computer can also check the data that are entered before the interviewee walks away. The most common error in paper-form surveys is incomplete records. There is nothing that requires a surveyor to complete an entire form or specific “minimum” required fields. The Airport O-D survey requires four fields to be entered before a record can be considered “complete” and then saved.

Logic checking data entry is impractical and virtually impossible using a paper method. The Airport O-D Survey checks for logic errors such as no like starting and ending locations. Had it been necessary to the client, cities and zip codes could have been checked against one another. Other pen-based, data collection programs have more logic checking needs than the Airport O-D survey.

The most cost-saving feature of computer, field-data entry is the storage of responses in an electronic database at the time the data is collected. This greatly reduces post-processing and speeds up analysis. In the Airport O-D survey, all records were written to

a .dbf file and sent to the client. The analyst had to lookup zip codes for some long entry fields; but, other than that, he/she was able to start the analysis.

In addition to the time saving, automatic entry eliminates coding errors from duplication of data. The data is already in electronic form and does not require another processing step that often introduces error. This improves accuracy in analytical or automated processing.

The computers were usually found to be easy to use. A computer like one shown in Figure 2 was used to collect the data in the North Carolina Airport Study.

Figure 2. Survey Instruments



Survey Schedules

One of the most difficult challenges faced in the North Carolina Airport study was conformity to the schedule dictated by UNCC. Through a random generator process and a required number of surveys at each airport, the number of days and specific days of the week were fixed. That is, interviews would be collected in Asheville on a Thursday and Sunday during the Fall, a Monday and Tuesday in the Winter, and only a Friday in the Spring. The challenge was to make the most efficient use of the surveyor by moving from one airport on one side of the state to a different airport the next day.

The problem: how to manage 23-25 trained surveyors, in 14 cities with variable days, over an 7 or 8 day period, with 20 pen-based computers? The easy solution would have to be fly everywhere. However, the budget didn't allow for such an easy solution. In fact, the budget didn't allow for full-time employees to perform the surveying; college students had to be used. One added complication was the availability of the students: e.g., Joe could work on Tuesdays, Thursdays and all weekends except the 2nd weekend in February.

Each surveyor had to be trained. Each one had to have hotel accommodations. Some surveyors needed plane flights; some needed car rentals; some had other special needs. Supervisors had to be trained and assigned to a group of surveyors for quality, good public relations, and problem solving.

Specific destination cities were determined by UNCC and were based on a random generator corresponding to a flight frequency. That is, a Tuesday in Charlotte might include four flights to New York City, three to Atlanta, one to Spartanburg, SC, one to Anchorage, AK, etc. This threw yet another kink in an already messy chain.

It either takes a sophisticated computer program to perform the necessary scheduling with all the previously discussed constraints, or, a lot of costly man-hours. The reality of origin-destination surveys is that there are fixed costs which are not accounted for in a purely per record cost. Half the number of records is not half the cost.

CONDUCTING THE SURVEY

Once actually in the field and surveying, the interviews were quite easy. The pen-based computer tool made collecting the data simple. The computer flowed through the questions, prompting interviewers for the next question, recording responses with little or no typing and generating interest from other [potential] interviewees.

Approaching Interviewees

Approaching a passenger is not always an easy task. All people naturally gravitate to those they are most comfortable being around. But that type of attraction can create a bias in the results. Therefore, it was necessary to instruct all surveyors to make an effort to approach all types of people. For example, if Joe is single and likes to talk to women, Joe would

have to force himself to talk to men. Similarly, Ahmed is Islamic and is culturally uncomfortable speaking to women; he, too, must avoid bias and talk to some women. The supervisor monitored potential sources of bias and took steps to correct them.

Introduction

Passengers were approached in departure/arrival gates and at baggage claim areas and asked something like: “Hi! I’m doing a study for the North Carolina Department of Transportation. How did you get to this airport today?” [the first survey question] The credibility is established from the start – North Carolina Department of Transportation. And, this is not a survey; it’s a study! Also note, the passenger is not asked if he/she wants to participate; the study just begins.

It is important to note, all surveyors had special permission from airport management to conduct the interviews. All surveyors wore badges with their name in large letters and NCDOT, UNCC and Street Smarts in plain view. At the Charlotte-Douglas Airport, the surveyor had to don badges issued by the airport with his/her picture on the badge.

Interview

The questions are shown in the paper version of the form in Figure 1. Some questions were harder to ask than others. For example, no one liked asking Household Income because many passengers did not like answering the question. And while it is humorous to talk about, answering the Gender question by observation was not always straightforward.

NCDOT and UNCC defined a minimum acceptable set of questions which would constitute a “complete” record. Those were 1) How did you get to this airport today? 2) Where did you come from [City and Zip]? 3) What is your purpose for traveling? 4) What is your final airport city? These questions could be asked very quickly if a passenger was about to board a flight or arrived on a flight and was headed out the door. Observed responses such as gender, age, group size, and possibly others could be coded once the passenger left the interview area.

Entry

Data entry on pen-based computers is most efficient if all acceptable responses are present on the computer form. One tap of a pen marks an entry and one can continue. However, there are “long entry” fields which require creative programming of the survey instrument in order to make entry fast and accurate.

In the Airport O-D Survey, all origins and destinations are cities. Respondents who arrive at an airport via ground transportation had to supply additional location information. A zip code, intersection, or place name was sufficient. These responses required some time to code.

To lessen the coding time on long entry fields, intelligent pick lists were created. For the ground transportation origin city, all cities and counties in North Carolina were sorted and indexed. The user simply had to start “typing” the first letters of the name and four cities starting with those letters appeared in a pick list. The same algorithms are used in most on-line help screens on personal computers; but instead of subject, cities are the pick items.

The requirements of the study only required zip code entry. In fact, if the zip code was known, no other information was necessary for analysis. However, in order to logic check the records, city and zip were required of the surveyors. If a city was known but not the zip (as in the case of a traveling salesman), the surveyor would “type” the name of a place (hotel, convention center, resort, etc.) and general location. Intersections or other place names were acceptable. Due to the vast number of streets in the state, it was decided not to provide a pick list of streets.

As stated previously, there were a minimum number of fields that constituted a complete record. The program would not let a record be saved without these fields. All of the fields could have been required and then programmed as such.

LESSONS LEARNED FROM THE SURVEY METHODOLOGY

An origin-destination study of the magnitude of the North Carolina Airport study has its challenges; a lot was learned.

Random Cities Hurt The Way Data Can Be Collected

The scope of the project was set with the assumption that surveys could be conducted at cities based on their geographic proximity to one another. The randomness of cities constrained the way data could be collected. The scheduling was more difficult, more costly, and much less efficient.

Students

Another lesson that is not too unfamiliar to the data collection community is that students are economical and smart but their schedules are difficult to work around. Over 50 students were trained to conduct the study and fewer than ten worked more than four days during the four seasons of the project. Classes, projects, homework, exams, and personal activities had to be accommodated.

Supply and Demand

At the small airports, the number of records depends more on supply of passengers than on the speed of interviews. The expectations of the NC DOT and UNCC were based on the ability to continually collect records like at the larger airports. The number of records statistically needed at the small airports required more survey days in order to interview more passengers.

Smile When You Say That

A positive, friendly attitude really helped. Surveyors found that passengers were much more receptive to those types of people. While this is not a major revelation, it is an important reminder.