Fun With Statistics

Establishing Transportation Statistics Quality Standards

A Research Proposal

By
Todd Litman
Victoria Transport Policy Institute
23 July 2007

Abstract
Accurate statistics are essential for transportation research, policy analysis and planning. Currently, transportation-related statistics are collected, but there is little consistency between data sets making it difficult (often impossible) to compare information between different agencies, jurisdictions and time periods. This paper reviews some current transportation data sets, provides general recommendations for establishing quality standards for transportation-related statistics, and proposes a research program for setting such standards.

Submitted for presentation at the 2008 Transportation Research Board Annual Meeting
Paper 08-0557

Words: 3,500 plus 1 figure x 250 = 3,750
**Introduction**

Many people have the misimpression that statistics (quantitative data) are boring. They couldn’t be more wrong. Properly evaluated statistics can provide exciting information. They are an extension of our senses, allowing people to understand what occurs in distant places, at scales beyond our comprehension, and in ways that we cannot direct perceive. Working with statistics can be fun and exciting.

For example, that New York County averages 4.4 traffic deaths per 100,000 residents is by itself not particularly interesting, but it becomes a little more interesting if compared with the 38.2 death rate in Yadkin Co, North Carolina, and much more interesting if traffic fatality rates of numerous counties are compared relative to some other factor, such as a sprawl indicator, as in Figure 1, which indicates that land use patterns affect traffic death rates, suggesting that smart growth policies may be a traffic safety strategy. This has important implications for traffic safety, land use planning and public health.

**Figure 1**

Annual Traffic Death Rate (Ewing, Schieber and Zegeer, 2003)

*The ten communities with the lowest sprawl rating have about a quarter of the per capita annual traffic fatality rates of the most sprawled communities.*

Transportation statistics have a wide variety of impacts on individuals, communities, the economy and the environment. They affect policy and planning decisions, from how facilities are designed and how money is spend, to how problems are defined and how solutions are evaluated. For example, traffic congestion is typically considered a problem if a particular area has more congestion delay than other comparable areas. Similarly, a particular mode, road or community is considered dangerous if its crash rates are higher than average. What impacts are measured and how they are measured can determine whether a particular situation is considered a problem, how that problem is defined, and what solutions are considered as a response. Statistics are the basis for rational decision-making, and are particularly important for sustainable transport planning (Litman, 2005).
Current transportation statistics are criticized for various reasons:

- Statistics are often incomplete, consisting of a limited set of information needed for analysis and planning.
- Some statistics are of questionable accuracy, based on inadequate methods or sample size (Schipper, et al., 1993).
- Transportation decision-making is skewed in favor of modes and impacts that are easy to measure at the expense of those that are more difficult to measure. In particular, motor vehicle travel tends to receive more attention than nonmotorized travel, and mobility tends to receive more attention than accessibility, because of the ease of measuring their impacts (Litman, 2003).
- Statistics are incompatible between different agencies, jurisdictions and time periods, making it very difficult to compare conditions, evaluate relationships and track trends.
- Statistics that do exist are often unavailable except to a limited audience, or they are made available in a format that is difficult to work with.
- There is seldom independent review and reporting of data quality.

As a result, a considerable amount of valuable research is constrained by inadequate statistics. This is not simply a question of increasing available resources, considerable time and money is already devoted to collecting and evaluating transportation statistics, the primary issue considered here is the lack of overall coordination and quality standards. As a result, the statistics that are collected are often of limited use, resulting in huge wasted resources. Considerable benefits can be achieved by improving the coordination of existing transportation statistics gathering activities.

This issue should be of concern to anybody who uses transportation statistics or is affected by the analysis of such statistics…which is everybody. Described more positively, everybody should be interested in finding ways to improve the quality of transportation statistics so decision-makers can better evaluate transportation impacts, problems and potential solutions.

This paper defines statistical quality factors, reviews the current state of transportation-related statistics, defines statistic quality factors, and recommends a research program for establishing standards for transportation-related statistics.
Defining Statistical Quality
This paper is concerned with the quality of transportation-related statistics. This includes the following factors:

- **Accuracy.** The methods used to collect statistics must be suitably accurate.
- **Transparency.** The methods used to collect statistics must be accessible for review.
- **Comprehensiveness.** An adequate range of statistics should be collected to allow various types of analysis. This should be disaggregated in various ways, including by geographic area (particularly by urban region), mode and vehicle type and demographic group.
- **Frequency.** Data should be collected regularly, which may be quarterly, annually, or ever several years, depending on type.
- **Consistency.** The range of statistics, their definitions and collection methodologies should be suitably consistent between different jurisdictions, modes and time periods.
- **Availability.** Statistics should be readily available to users. As much as possible, data sets should be available free on the Internet in spreadsheet or database format.

The following are categories of transportation-related data. These should be disaggregated in various ways, including by geographic location, demographic group, mode, time period, travel conditions, etc., depending on the type of statistics.

- Transportation facilities (roads, railroads, ports, airports and parking facilities) and services (transit, air travel, shipping, etc.)
- Expenditures on transportation facilities and services, and user fee revenues.
- Vehicles ownership, registrations, fuel efficiency and taxes.
- Driver license availability (by age, income, location, etc.)
- Vehicle use (mileage by each vehicle class).
- Personal travel activity (person-kilometers, person-trips, time devoted to travel, by mode, including nonmotorized modes).
- Freight travel activity (tonne-kilometers by mode).
- Energy consumption by mode, fuel type, plus fuel prices and taxes.
- Pollution emissions, by type of pollutant, vehicle type and travel conditions.
- Climate change emissions.
- Traffic accidents and fatalities.
- Household expenditures on transportation.
- Demographic factors (age, income, disabilities).
- Land use factors (urbanization, population and employment density, land use mix, etc.)
- Parking supply.
- Telephone and Internet availability.
Current Transportation-Related Statistics Sets
Below is a summary of some major transportation-related statistics sets.

International

Cross National Time Series (www.scc.rutgers.edu/cnts/about.cfm)
The Cross-National Time-Series Data Archive (CNTS) offers a comprehensive listing of international and national country data facts, going back to 1815. Transportation-related statistics include: area and population, urbanization, highway vehicles, and railroad data.

EarthTrends Searchable Database (www.earthtrends.wri.org)
The EarthTrends database provides city- and country-level indicators on road networks, private vs. public transport use, vehicle fleets, road traffic, and fuel prices. It uses statistics from the International Road Federation, the World Bank and the UNHabitat's Global Urban Indicators Database.

International Road Federation (www.irfnet.org)
The IRF publishes World Road Statistics, a global compilation of road and vehicle statistics compiled from official sources within national statistics offices and national road administrations in more than 200 countries.

Millennium Cities and Mobility In Cities Database
The International Association of Public Transport (www.uitp.com) has compiled the Mobility In Cities Database (www.uitp.com/rome2005/RMR/en/pics/MCD-en.pdf), which contains than 200 indicators from 100 cities for various years. Statistics include:

- Population
- Land use factors
- Vehicle supply
- Road and public transit networks
- Vehicle travel, personal travel and mode split
- Accidents
- Energy consumption and pollution
- Transportation expenditures

OECD Transport Statistics (www.oecd.org)
Organization for Economic Cooperation and Development produces a variety of transportation statistics, including the OECD Factbook (www.sourceoecd.org/factbook), which provides information on passenger transport, freight transport and road accidents for 30 countries, including transportation statistics in spreadsheet format (http://dx.doi.org/10.1787/353365538624), and the International Road Traffic and Accident Database, which provides comprehensive crash statistics. The OECD Online Information Services (OLISnet) is a program to improve and broaden information flows among OECD countries, including a section on transport statistics.

World Bank (www.worldbank.org)
The World Bank Development Indicators provides some information on transportation (road networks, vehicle ownership, freight transport) as available for all countries in the world. It also includes other data on population and productivity in the World Development Indicators (www.worldbank.org/data/wdi2003/index.htm) and various specialized databases (http://econ.worldbank.org/resource.php?type=18).
Europe

European Union
The European Union’s *Energy and Transport in Figures 2006* ([http://ec.europa.eu/dgs/energy_transport/figures/pocketbook/2006_en.htm](http://ec.europa.eu/dgs/energy_transport/figures/pocketbook/2006_en.htm)), provides energy and transport statistics for the Member States of the European Union and other European countries (the countries that have applied to join the European Union and the countries of the European Free Trade Association). Data are available in spreadsheet format. The European Commission’s Eurostat Online Database ([www.europa.eu.int](http://www.europa.eu.int)) provides statistical information on European countries, including vehicle ownership, personal and freight travel, population and income. The report, *Passenger Mobility in Europe* summarizes and compared passenger transport data from 20 European countries.

European Commission, DGVII ([www.europa.eu.int/comm/dgs/energy_transport](http://www.europa.eu.int/comm/dgs/energy_transport))
The European Commission, DGVII, web page contains transport statistics for EU countries, some countries in Eastern Europe, with comparisons with the U.S. and Japan. It publishes *EU Energy and Transport In Figures*, which provides comprehensive information on transportation in the European Union, and other documents on transportation.

European Conference of Ministers of Transport ([www.oecd.org/cem/stat](http://www.oecd.org/cem/stat))
The European Conference of Ministers of Transport website provides information on various forms of transportation, including time series data on road, rail and marine transport networks and use, and transport safety.

European Environment Agency ([www.eea.eu.int](http://www.eea.eu.int))
This international organization provides information on European vehicle emissions and emission reduction strategies.

United States

Bureau of Transportation Statistics ([www.bts.gov](http://www.bts.gov))
The U.S. Bureau of Transportation Statistics (BTS), a diversion of the U.S. Department of Transportation, serves as the lead agency in developing and coordinating intermodal transportation statistics. It provides comprehensive data collection, analysis, and reporting for a variety of government agencies, and produces a wide range of reports and databases. The annual *National Transportation Statistics* report provides annual information on vehicle ownership, vehicle travel, personal travel, freight travel, crashes, energy consumption, pollution emissions, and other transport activities and impacts. TransStats ([www.transtats.bts.gov](http://www.transtats.bts.gov)) is a searchable index with more than 100 transportation-related databases of various modes — with many social and demographic data sets that are commonly used in transportation analysis.
Highway Statistics ([www.fhwa.dot.gov/ohim](http://www.fhwa.dot.gov/ohim))
The U.S. Federal Highway Administration’s annual *Highway Statistics* reports provide information on roadway conditions, travel and expenditures in the U.S. The major sections are listed below, with most data provided by state and national totals. This is a model for the standardized collection and dissemination of transportation statistics.

- Motor Fuel: consumption of petroleum and alternative fuels, prices and taxes.
- Motor Vehicles: registrations, mileage, fuel efficiency and taxes.
- Driver Licensing: driver licenses by age group.
- Highway Finance: vehicle user fee receipts, roadway expenditures, toll revenues, and transit expenditures, each grouped by local, state and federal level of government.
- Roadway Extent, Characteristics, and Performance: road supply, vehicle traffic and fatalities by roadway category (highway, arterial, collector, etc.) and location (urban or rural).
- International Comparison: compares various transportation data with Japan, France, Germany, Sweden, United Kingdom, Canada, Mexico and the United States.

Census Bureau ([www.census.gov](http://www.census.gov))
The U.S. Census Bureau collects and distributes a variety of data about the people and economy of the United States, including information on population, income, trade, transportation activities, journey to work, transportation industry employment and production. The *American Community Survey* ([www.census.gov/acs](http://www.census.gov/acs)), which replaces the census long form, provides community-level demographic and economic data, including income, employment rates, commute travel time and mode split. The 2001 *American Housing Survey* has commute distance, time and mode statistics. The *Journey to Work* report provides information on commute patterns and their shifts between 1990 and 2000.

National Household Travel Survey ([www.bts.gov/nhts](http://www.bts.gov/nhts))
The National Household Travel Survey (formerly known as the Nationwide Personal Transportation Survey [NPTS] and the American Travel Survey [ATS]) is a household-based travel survey, conducted every five years that collects information on daily, local trips and on long-distance travel in the United States. Summary data and analysis are presented in reports, and users can access the complete database through an Internet interface that performs searches, sorting and other types of analysis.

Department of Energy ([www.doe.gov](http://www.doe.gov))

Regional Planning Agencies ([www.narc.org](http://www.narc.org))
Most regional planning agencies (called Metropolitan Planning Organizations or MPOs) and state and provincial transportation agencies collect data on road, rail and transit networks, personal and freight transport activity, expenditures and crashes. This information is increasingly available through the Internet.
Canada

Transport Canada (www.tc.gc.ca)
Transport Canada provides a variety of statistical information on vehicles, personal travel, transport safety and freight transport in Canada.

Statistics Canada (www.statscan.ca)
Statistics Canada is a federal agency that collects, analyzes and distributes objective information on Canada’s population, resources, economy, society and culture, including statistics on road, rail, air, marine, vehicles, journey to work, trade, freight, consumer expenditures, and other transportation related factors. The Canadian Vehicle Survey provides quarterly data on vehicle ownership and use. The National Private Vehicle Use Survey (NaPVUS) provides data on fuel consumption and travel activity.

Transportation Association of Canada (www.tac-atc.ca)
The Transportation Association of Canada is a non-profit association of transportation stakeholders in government, private industry, and educational institutions. It produces various information resources, including Urban Transportation Indicators Database and the Directory Of Canadian Transportation Data Sources.

Britain

Transport Statistics Great Britain (www.dft.gov.uk/transtat)
The British Department for Transport, Transport Statistics website provides the following national and regional transportation data:

- Personal Travel, including information from the National Travel Survey.
- Public Transport, including tables using information collected from bus and coach operators.
- Road Vehicles, including stock and first registrations.
- Road Traffic, including traffic and traffic speeds by region.
- Roads, including road lengths and road condition by region.
- Safety, including regional casualty figures for different types of road users.
- Freight, including goods moved and lifted by origin and destination.
- Air, including passenger and freight movements by region.

UK National Travel Survey (NTS)
(www.statistics.gov.uk/ssd/surveys/national_travel_survey.asp)
The NTS, performed each decade, details of the travel habits of British residents and shows how these vary by factors such as age, gender, car ownership and location.

Australia

Australian Bureau of Transport and Regional Economics (www.btre.gov.au)
Australian Transportation Statistics provides comprehensive statistics on roadway, rail, marine and air transport in Australia. The Transport Indicators Homepage provides a variety of statistics on Australian transport services and activities.
Recommended Research Program
There are, to my knowledge, no recognized formal standards of quality that would apply to transportation-related statistics, nor any international program to establish such standards. A number of organizations deal with a portion of data quality issues, including various organizations which collect statistics on specific travel activities and impacts (road building, accidents, pollution emissions) and professional organizations which manage travel surveys and vehicle traffic counts, but no single organization is responsible for collecting and disseminating the full range of statistics needed among multiple jurisdictions around the world.

The following research program is recommended to begin improving the quality of transportation statistics.

1. Perform a comprehensive review of existing transportation-related data sets and data collection activities.

2. Consult with data collection experts and users to identify best current practices with regard to factors such as collection methods, comprehensiveness and availability.

3. Develop a set of appropriate standards for transportation-related statistics, which clearly define minimum requirements for accuracy, transparency, comprehensiveness, frequency, consistency and availability.

4. Work to have these standards adopted by all major transportation research organizations and their members, including TRB, WCTRS, OECD, World Bank, etc.

5. Develop an auditing program through which individual agencies engaged in collecting transportation-related statistics can have their programs evaluated by peers to assess quality and recommend improvements.

The ultimate goal of this program is to raise the overall standard of data collection and to create a framework so that data are consistent, allowing statistics to be compared between different jurisdictions, agencies and time periods. This could provide tremendous value to society at relatively low cost.
Conclusions
Transportation-related statistics are widely gathered, but their quality is highly variable, ranging from good to abysmal, and even the best data sets are incompatible with those created in other jurisdictions and agencies. This is a huge waste of resources and a lost opportunity for improving our ability to understand important trends and relationships.

It would be difficult to understate the value of improving the quality and consistency of transportation-related statistics. Each year transportation statistics are used in hundreds of millions of dollars worth of research which affects tens of billions of dollars in planning decisions, which affects hundreds of billions of dollars worth of transport activity. The quality of these statistics is highly variable, some data sets are good but many are poor. As a result, decision-makers often fly blind, unable to perceive important trends and relationships.

This paper outlines a research program for establishing minimal standards for transportation-related data which can be adopted by transportation research organizations and their members. These standards would then become the basis for evaluating and improving data collection and dissemination. This effort is consistent with good management practices (good measurement is the basis for good management) and sustainable transportation planning. To my knowledge, no effort such as this is currently underway.

This will not be an easy project, but if successful, it could provide huge value to future data users, which ultimately include everybody on earth.
References And Resources For More Information

*Bureau of Transportation Statistics (www.bts.gov)* provides comprehensive transportation information for the United States.


*International Association of Public Transport (www.uitp.com).*


*National Household Travel Survey* (2001), BTS, USDOT (www.bts.gov/nhts) provides information on a series of travel surveys performed every five years in the United States.


OECD (annual reports), *International Road Traffic and Accident Database*, Organization for Economic Cooperation and Development (www.bast.de/htdocs/fachthemen/irtad//english/we2.html).
OECD (annual reports), *OECD Factbook*, Organization for Economic Cooperation and Development ([www.sourceoecd.org/factbook](http://www.sourceoecd.org/factbook)).

Alan Pisarski (2006), *Commuting In America III*, Transportation Research Board ([www.trb.org](http://www.trb.org)).


*TrafficLinq* ([www.trafficlinq.com](http://www.trafficlinq.com)) is an extensive directory of links covering issues regarding road traffic and transportation. It covers about 1,000 web sites world wide, and has an option to scan all transportation sites with one query.

*Transport Geography on the Web* ([www.people.hofstra.edu/geotrans](http://www.people.hofstra.edu/geotrans)) is an Internet resource to promote access to transport geography information, including articles, maps, and datatsets.

*TransStats: The Intermodal Transportation Database* ([www.transtats.bts.gov](http://www.transtats.bts.gov)).


