The Department of Transportation's DRAFT Report Implementing OMB's Information Dissemination Quality Guidelines

WHAT IS THE PURPOSE OF THIS POSTING? In accordance with The Office of Management and Budget's (OMB) Guidelines (for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies) implementing Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (P.L. 106-554), each Federal agency subject to the Paperwork Reduction Act (PRA) (44 U.S.C. Chapter 35) must prepare a draft report presenting the agency's guidelines and explaining how the guidelines ensure the quality of disseminated information. Each agency must also post its draft guidelines on the agency's website to provide an opportunity for public comment. Accordingly, the Department is requesting comments on its draft information dissemination quality guidelines.

WHEN SHOULD COMMENTS BE SUBMITTED? You should submit comments by May 31, 2002.

WHERE SHOULD COMMENTS BE SUBMITTED? You may file comments using the Internet by logging in on DOT's Dockets Management System (DMS) website at **http://dms.dot.gov**. Please follow the online instructions for submitting an electronic comment and for reviewing all comments on line. Once received, a notification receipt will be forwarded to you. You may fax your comments to the DMS at (202) 493-2251. You may also submit your comments by mail or in person by sending your comments to the U. S. Department of Transportation (DOT), Office of Dockets and Media Management to the Docket Clerk, Docket No. OST-2002-11996, 400 Seventh Street, SW., Room PL- 401, Washington, DC 20590-0001. If you would like the Department to acknowledge receipt of your written comments, you must include a self-addressed stamped postcard on which the following statement is made: "Comments on Docket OST-2002-11996." The Docket Clerk will date stamp the postcard prior to returning it to you via the U.S. mail. Comments should identify the docket number. Written comments should be submitted in duplicate.

The Office of Dockets and Media Management is open for examination and copying, at the above address, from 10:00 a.m. to 5 p.m., Monday through Friday, except Federal holidays. All comments received will be available for inspection at the above address.

Please note that due to current mail security procedures affecting U.S. Postal Service delivery to Government offices, commenters may find it advantageous to use an alternative method (the internet, fax, or professional delivery service) to submit comments to the Docket and ensure their timely receipt at the U.S. Department of Transportation.

WHO SHOULD BE CONTACTED FOR FURTHER INFORMATION ABOUT THESE DRAFT GUIDELINES? Vanester M. Williams, Office of the Chief

Information Officer, U. S. Department of Transportation; 202-366-1771 (not a toll-free call) or by email at vanester.williams@ost.dot.gov. For inquiries on the Department's administrative mechanisms for persons to seek correction of information, please contact Robert Ashby, Office of the General Counsel, U. S. Department of Transportation; 202-366-9306 (not a toll-free call) or by email at bob.ashby@ost.dot.gov. For inquiries on the guidelines concerning statistical disseminated information, contact Dr. Patrick Flanagan, Bureau of Transportation Statistics, U.S. Department of Transportation; 202-366-4168 (not a toll-free call) or by email at pat.flanagan@bts.dot.gov.

REQUEST FOR COMMENTS: The Department of Transportation has long been committed to providing high quality information to the public and to basing regulatory and other decisions on the best information available. These draft guidelines do not represent a new or changed policy on the Department's part, but rather a reaffirmation of that existing commitment. The draft guidelines begin an evolutionary process of incorporating in the Department's activities specific information quality standards and procedures developed by the Office of Management and Budget to ensure continued high information quality. The language of the draft guidelines makes a basic point very clear: these are not regulations or legally binding mandates. Much of the material is in a "discussion" format designed to inform the public and DOT staff of the thinking process the Department intended to use in addressing data quality issues (see, for instance, the discussion of the "influential information" concept).

While much in the draft guidelines is self-explanatory, we would call interested persons' attention to the following provisions.

The Department of Transportation and its organizations disseminate a wide variety of information. The Department's information products include such things as compilations of statistical data, technical and guidance manuals, reports generated in our safety and grant programs, web sites containing safety and economic information about transportation industries, and consumer information. Consistent with the statute and the OMB guidelines, DOT intends that the substantive information we produce will be subject to these guidelines and will be reviewed before dissemination to ensure its quality. However, there are some types of information that these guidelines would not cover.

Section IV concerns types of information that are not covered by the draft guidelines. Most are based on items that the OMB guidelines specifically do not cover (you may refer to OMB's Guidelines for the types of information not covered in their guidelines). In some cases (e.g., archival records, adjudicatory processes) the Department has provided additional explanation or examples of the kind of materials that we would exempt from coverage. In other cases (e.g., internal manuals and material presented to Congress) we would add categories of information that the OMB guidelines do not specifically exempt. In the latter case, we believe it would be duplicative and counterproductive to subject some materials submitted to Congress, such as testimony or legislative information, since the political process provides a means to evaluate and, if necessary, call for revision of such materials. The Department seeks comment on the

issue of coverage of the guidelines, as well as any other subjects that the guidelines should or should not cover.

Section III mentions an important concept that may not be immediately obvious to persons reading the OMB guidelines for the first time. As Dr. John Graham, Director of the OMB Office of Information and Regulatory Affairs (OIRA) and others have pointed out in meetings about the information quality guidelines, the standards for data quality that apply directly to Federal agencies also apply, at least indirectly, to outside parties who supply information to the Department. If the Department is to rely on technical, scientific, or economic information submitted by, for example, a commenter to a proposed rule, that information would need to meet appropriate standards of objectivity and utility. Numbers submitted by a commenter as the basis for a regulatory decision – which the Department would necessarily disseminate as part of a rulemaking issuance – should meet data quality standards no less than in the case of information the Department itself generates. If they do not, then the Department's ability to rely on them would probably be reduced. The Department seeks comment on the effects of these guidelines on information provided by outside parties and on how the pre-dissemination review and administrative corrections processes could most constructively apply to this information.

A special instance of this issue concerns information submitted by state and local governments to DOT organizations in the context of grant programs. Typically, state and local agencies that receive Federal funds provide information to the Department that the Department uses to evaluate the grant programs and, in some cases, to allocate funds. The Department may not have control over the production or presentation of this information, and often there may be no alternatives to using it in DOT's information products. The Department is considering exempting information of this type from the guidelines, and seeks comment on this point.

In the administrative correction process, DOT will make extensive use of the internet-accessible DMS. All requests for correction would come, in the first instance, to the DMS, whether electronically or in hard copy. By docketing requests for correction and subsequent DOT responses in the DMS, the Department will ensure the transparency of the request and response process. The DMS will also electronically notify DOT organizations of pending requests. In addition, filing requests with DMS will allow other interested parties to comment about or make requests with respect to an information issue. For example, suppose DOT publishes a study indicating that 75 percent of a certain kind of accident is caused by a component of a motor vehicle. Manufacturers of that component request correction of the study. Alerted to the request by the DMS posting, vehicle manufacturers could respond within 30 days. The Department seeks comment on this process.

Section VIII (c) lists items that persons requesting correction of information should include in their requests. As the guidelines state, these items are not intended to be mandates. However, the DOT may not be able to respond in a timely fashion, or at all, in the absence of needed information concerning a request. We seek comments on what the appropriate contents of correction requests should be.

The Department need not respond substantively to frivolous, repetitive, or stale requests for correction. Nor should DOT have to respond substantively to requests that concern information not covered by the guidelines or from a person whom the information does not affect. Section VIII (d) contains a series of "filters" designed to ensure that DOT's resources are not expended in responding to requests of this kind. The Department can request information regardless of these filters. At the same time, we recognize that, in the interest of good customer service, we should respond to relevant requests in a timely fashion. The guidelines contain 45-day and 90-day timeframes for various responses to requests. Are these time frames appropriate?

For the most part, the OMB guidelines and these draft guidelines contemplate that DOT will make a written response to the requester within the time frames noted above. However, there are some circumstances in which there is an existing process to respond to concerns expressed about the DOT's information. The OMB guidelines encourage agencies to make use of existing processes in a flexible way, tailored to their programs. When there is a sound existing process, (such as a process that provides opportunities for public participation in making an agency decision), DOT organizations are asked not to duplicate that process with a separate request response mechanism.

For example, when an agency issues a notice of proposed rulemaking (NPRM), it typically describes in the preamble the basis for its proposed regulatory provisions, which may include technical or scientific studies and a regulatory evaluation. In so doing, it disseminates these studies or evaluations, within the meaning of these guidelines. The public comment process can, and often does, generate views from interested persons about the soundness of the underlying information. If someone submits a request for correction pertaining to a document cited in an NPRM, DOT would treat it procedurally like a comment to the rulemaking, responding to it in the preamble of the final rule or a subsequent document such as a Supplemental Notice of Proposed Rulemaking (SNPRM), rather than through the separate request response mechanism of these guidelines. The content of the response would address the issues of the document's compliance with the information quality principles of the OMB and DOT guidelines. (DOT could choose to make an earlier correction, if warranted, assuming so doing would not delay the issuance of the final rule.) This approach would also apply to other processes involving a structured opportunity for public participation on a proposed document before a final document is issued, such as a draft environmental impact statement (EIS), an air quality conformity determination, or a Section 4(f) determination under the Department of Transportation Act.

In the case of a correction request pertaining to a final rule, EIS, etc., DOT would first evaluate whether the request pertained to something on which a requester could have commented at the proposed stage. For example, suppose an NPRM or draft EIS cited a study as a basis for DOT's proposed decision. Requesters had the opportunity to review the study and comment or request administrative correction at the proposed stage. In this situation, DOT would not be required to consider an administrative correction request at the final rule or EIS stage pertaining to the same material. The structured opportunity to

be heard on the issue had come and gone, and DOT's interests in finality and avoiding uncertainty with respect to its decisions would outweigh the requester's interest in reopening the issue after the final document had been issued.

On the other hand, with respect to new information appearing for the first time in a final rule or EIS, DOT would consider a request for correction. The Department would not stay the final action involved. However, if it appeared that the information that was the subject of the request did not comply with the guidelines, and that, as a result, the final document was materially flawed, DOT would treat the matter as a request for reconsideration. In such cases, the Department would use any already existing mechanisms and procedures to reconsider corrections, such as the process to petition for a new rule or to request a Supplemental EIS. The submission of a request for correction by itself does not in any way affect the finality of a decision of the Department.

We believe that this approach serves the purposes of the guidelines, affords an opportunity for correction of any material that does not comply with the guidelines, yet does not duplicate effort or interfere with the orderly progress of DOT's work. We seek comment on this approach.

The OMB guidelines call for requesters to have an opportunity for reconsideration or appeal from an initial agency action on a request for correction that does not satisfy them. One important issue surrounding this opportunity is how to choose the official(s) who perform this function. Ideally, the decision maker would be someone who knows the subject matter well but who is far enough removed from involvement with previous agency actions in the matter to be fully objective. Finding both qualities in the same person can be difficult. The draft language asks DOT organizations to use their best judgment to find a good balance between the two considerations. We seek comment on whether, at least in some cases (e.g., those involving "influential" information), it would be a good idea to have a panel of people making the reconsideration decision (e.g., someone from the operating administration who made the initial decision plus two others from other parts of DOT who are familiar with the types of issues involved). (Refer to Section XI for a more complete discussion of the term "influential".)

With respect to the discussion of influential information, DOT is using, as a tentative criterion for the economic effect of information, the same $100 million figure that Executive Order 12866 uses for determining whether a rule is economically significant. The Department seeks comment on whether this figure is appropriate for this purpose, or whether some other number or combination of numbers (e.g., different numbers for small and large entities) would be better.

In general, the Department solicits suggestions on other ways these guidelines can be improved to meet the objectives of the statute and the OMB guidelines.

TABLE OF CONTENTS

DRAFT GUIDELINES FOR ENSURING AND MAXIMIZING THE QUALITY, OBJECTIVITY, UTILITY AND INTEGRITY OF INFORMATION DISSEMINATED BY THE DEPARTMENT OF TRANSPORTATION (DOT).

I. WHAT IS THE BACKGROUND AND PURPOSE OF THESE GUIDELINES?

The Department of Transportation (DOT) is drafting guidelines to implement Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (PL 106-554). The Office of Management and Budget (OMB) has issued Government-wide guidelines under Section 515, which direct Federal agencies subject to the Paperwork Reduction Act (44 U.S.C. Chapter 35) to establish and implement written procedures to ensure and maximize the quality, utility, objectivity and integrity of the information that they disseminate. DOT's guidelines will apply to a wide variety of substantive information dissemination activities in order to meet basic information quality standards set forth by Section 515. The purpose of these guidelines is to provide a framework under which DOT will provide affected persons an opportunity to seek and obtain correction of information maintained and disseminated by DOT that does not comply with these guidelines. DOT designated the Departmental Chief Information Officer (CIO) as the senior official responsible for DOT compliance with these guidelines**.** Heads of Departmental Organization are responsible for ensuring compliance with these Departmental guidelines. Final guidelines will be effective October 1, 2002.

II. WHO ARE THE DOT ORGANIZATIONS TO WHICH THESE GUIDELINES APPLY?

These guidelines apply to the following DOT organizations.

Office of the Secretary (OST)
Bureau of Transportation Statistics (BTS)
Office of the Inspector General (OIG)
Federal Aviation Administration (FAA)
Federal Highway Administration (FHWA)
Federal Motor Carrier Safety Administration (FMCSA)
Federal Railroad Administration (FRA)
Federal Transit Administration (FTA)
Maritime Administration (MARAD)
National Highway Traffic Safety Administration (NHTSA)
Research and Special Programs Administration (RSPA)
Saint Lawrence Seaway Development Corporation (SLSDC)
Transportation Administrative Service Center (TASC)
Transportation Security Administration (TSA)
United States Coast Guard (USCG)

DOT organizations may adopt further guidance, consistent with these guidelines, applying to the specifics of their programs and information products.

III. WHAT ARE THE SCOPE, NATURE AND LEGAL EFFECT OF THESE GUIDELINES?

a. The Scope: These guidelines apply to certain information (both statistical and non-statistical) disseminated by DOT on or after October 1, 2002, regardless of when the information was first disseminated. Guidelines apply to all media (printed, electronic, or in other form). As is the intent of OMB's guidelines, DOT's guidelines will focus primarily on the dissemination of substantive information (i.e., reports, studies, summaries) rather than information pertaining to basic agency operations.

The standards of these guidelines apply not only to information that DOT generates, but also to information that other parties provide to DOT, if the other parties seek to have the Department rely upon or disseminate this information or the Department decides to do so. For example, suppose that a trade association, in commenting on a proposed rule, supplies a scientific or technical study or an economic analysis in support of its position on what the final rule should say. In order for DOT to rely on this information in a subsequent DOT dissemination of information (e.g., as part of the basis cited for decisions in the final rule), the quality of the trade association's information would have to be consistent with these guidelines.

The types of Departmental information not subject to these guidelines are outlined in Section IV.

b. Nature and Legal Effect: These guidelines are suggestions, recommendations, and policy views of DOT. They are not intended to be, and should not be construed as, legally binding regulations or mandates. These guidelines are intended only to improve the internal management of DOT and do not create any right or benefit, substantive or procedural, enforceable at law or equity, by any party against the United States, its agencies (including the Department of Transportation or any DOT organization), officers, or employees, or any person.

IV. WHAT TYPES OF INFORMATION ARE NOT SUBJECT TO THESE GUIDELINES?

Below is a sample of the various types of information not covered:

a. Distribution of information that is limited to government employees, agency contractors or grantees;
b. Intra- or inter-agency use of sharing of government information; responses to requests under FOIA, Privacy Act, the Federal Advisory Committee Act or other similar laws;
c. Predisclosure Notification to Submitters of Confidential Commercial Information (FHWA Notice N1320.6);

d. Distribution limited to correspondence with individuals or persons (regardless of media, ex. electronic mail);
e. Archival records disseminated by Federal agency libraries or similar Federal data repositories; (e.g., inactive or historical materials in DOT libraries and other data collections - including bibliographies or responses to reference requests pertaining to such materials);
f. Public filings (e.g. material filed by DOT or non-DOT parties in DOT Dockets or by DOT in agency dockets);
g. Information relating to subpoenas and adjudicatory processes (For the purpose of these guidelines, these processes would include, but are not limited to:
    1. Court litigation (e.g., briefs and attachments, or other information submitted to a court);
    2. Administrative law enforcement proceedings;
    3. Civil rights and personnel complaints and reviews; (e.g., under Titles VI and VII of the Civil Rights Act; the Americans with Disabilities Act; Sections 501, 504 and 508 of the Rehabilitation Act of 1973; Title IX of the Education Amendments of 1972, and Disadvantaged Business Enterprise matters);
    4. Debarment and suspension matters, 49 CFR Part 29 (Federal-aid contracts) and 48 CFR Part 9 (direct contracts);
    5. Merit System Protection Board matters (Sections 7511, 7543, and 70701 of Title 5, United States Code);
    6. Matters before the Board for Correction of Military Records (USCG) (Section 1552 of Title 10, United States Code and Part 52 of Title 33, Code of Federal Regulations (1996));
    7. USCG Commandant decisions on Appeal and Review of Suspension and Revocation Proceedings, Sections 7702 and 7704, Title 46, United States Code)
    8. Locomotive engineer certification matters, 49 CFR Part 240, Subpart E-Dispute Resolution procedures;
h. Hyperlinks to information that others disseminate (as well as paper-based information from other sources referenced but not adopted or endorsed by DOT);
i. Views or opinions, where the presenter makes it clear that what is being offered is someone's opinion rather than fact or the Department's views;
j. Information presented to Congress (as part of the legislative or oversight processes (e.g., testimony of DOT officials, information or drafting assistance provided to Congress in connection with pending or proposed legislation) that is not simultaneously disseminated to the public;
k. Press releases and other information of an ephemeral nature, advising the public of an event or activity of a finite duration - regardless of medium; and
l. Procedural, operational, policy, and internal manuals prepared for the management and operations of DOT that are not primarily intended for public dissemination. This includes personnel notices such as vacancy announcements.

V. WHAT GUIDELINES ARE DOT ORGANIZATIONS IMPLEMENTING FOR GENERAL INFORMATION?

DOT has traditionally utilized standards, policies, and other operational guidelines to ensure the quality of all its disseminated information. Incorporating these guidelines further reinforces DOT's commitment of meeting higher standards of quality prior to disseminating information to the public.

To ensure compliance with these guidelines, each DOT organization has appointed a data quality official who will serve as the liaison for implementing these guidelines within its organization. Each DOT organization is responsible for modifying appropriate information technology and administrative policies to comply with these guidelines.

OMB's guidelines define "quality" as an encompassing term comprising utility, objectivity, and integrity. Therefore, the guidelines sometime refer to these four statutory terms, collectively, as "quality." At a minimum, a basic standard of quality will be ensured and established for all DOT information prior to its dissemination. In addition, on-going disseminated information will be reviewed on a regular basis to ensure all information is current and complies with these guidelines. Specifically, DOT will set the following standards at levels appropriate to the nature and timeliness of substantive information to be disseminated.

a. Utility: DOT organizations will assess the usefulness of the information to be disseminated to the public. The originating office will continuously monitor the information needs and develop new sources or revise existing methods, models, and information products where appropriate;

b. Objectivity: DOT organizations will ensure disseminated information is substantively accurate, clear, complete, and presented in an unbiased manner. The originating office will use reliable data sources and sound analytical techniques. To the extent possible and consistent with confidentiality protections, the originating office will identify the source of disseminated information so that the public can assess whether the information is objective;

c. Integrity: DOT organizations will ensure information is protected from unauthorized access, corruption or revision (i.e., make certain disseminated information is not compromised through corruption or falsification). The Department is highly protective of information collected under pledges of confidentiality. To ensure integrity of information disseminated electronically, the departmental CIO has implemented an aggressive Information Technology Security Program (ITSP). This Program, which complies with the computer security provisions of the Paperwork Reduction Act covers all DOT information, data and resources collected, stored, processed, disseminated, or transmitted using DOT information systems, to include the physical facilities in which the information, data and resources are housed. The ITSP applies to all DOT employees, specifically

those in positions of public trust, contractors, subcontractors and other users of DOT information technology and related resources. For example, Departmental policy, as stated in the Department's Information Resources Management Manual (DIRMM), states that all DOT web managers are responsible for establishing appropriate security safeguards for ensuring the "integrity" of the information disseminated on the web sites and complying with The Privacy Act of 1974, as amended, to ensure appropriate disclosure of information.

DOT is also subject to several other statutory requirements to protect the information it gathers and disseminates. These include, but are not limited to: The Paperwork Reduction Act of 1995, the Government Information Security Reform Act; OMB Circular A-130 (Management of Federal Information Resources, dated 12/12/85, revised 11/28/00); 49 U.S.C. 111(i) (BTS establishment) and the Trade Secret Act (18 U.S.C. 1905)

d. Accessibility: In 2001, the Departmental CIO issued a policy to ensure accessibility to all persons. This policy applies to all Departmental electronic and information technology developed, procured, maintained, or used by DOT organizations on or after June 21, 2001 unless covered by one of the following conditions: 1) micro-purchases made prior to January 1, 2003 (FAR 39.204(a); or 2) conformity would impose an undue burden on the agency (FAR 39.204(e) and 36 CFR 1194.2).

DOT organizations will ensure that all disseminated information (including electronic and information technology media) is accessible to all persons (see DOT Section 508 Policy Statement).

VI. WHAT ADDITIONAL GUIDELINES ARE DOT ORGANIZATIONS IMPLEMENTING FOR STATISTICAL INFORMATION?

The 1991 Intermodal Surface Transportation Efficiency Act (ISTEA) created the Bureau of Transportation Statistics (BTS) within DOT. Among other things, it made BTS responsible for: "issuing guidelines for the collection of information by the Department of Transportation required for statistics … in order to ensure that such information is accurate, reliable, relevant, and in a form that permits systematic analysis." (49 U.S.C. 111 (c)(3))

A parallel requirement for developing guidelines emerged in the Paperwork Reduction Act of 1995. It tasked the Office of Management and Budget (OMB) to "develop and oversee the implementation of Government wide policy, principles, and guidelines concerning statistical collection procedures and methods; statistical data classification; statistical information presentation and dissemination; timely release of statistical data; and such statistical data sources as may be required for the administration of Federal programs." (44 U.S.C. 3504 (e)(3))

The Department's Bureau of Transportation Statistics (BTS) has established additional guidelines covering the Department's statistical programs. These guidelines (for statistical information) are based on structured planning, sound statistical methods and the principle of openness. Structured planning maintains the link between user needs and data system design. Sound statistical methods produce information (data and analysis results) that conforms to that design. Openness ensures that users of statistical information can easily access and interpret the information.

Each section of the statistical guidelines (Section VI a-e below) begins with a statement of principles, which contain definitions, assumptions, and rules or concepts governing action. The *principles* are followed by *guidelines,* which are specific recommended actions. Finally, each section concludes with *references and examples*.

To access DOT's Statistical Data Quality Guidelines click on section links below or refer to Appendix A of this report for a complete version of these guidelines.

a. Planning Data Systems
   - Data System Objectives
   - Data Requirements
   - Methods to Acquire Data
   - Determine Sources of Data
   - Sample Design or 100% Collection

b. Collection of Data
   - Forms
   - Data Collection Operations
   - Response Rates

c. Processing Data
   - Data Editing and Coding
   - Handling Missing Data
   - Production of Estimates and Projections
   - Data Analysis and Interpretation

d. Dissemination of Information
   - Presentation of Data
   - Source and Accuracy Statements
   - Documentation
   - Revisions

e. Evaluating Information Quality
   - Pre-Dissemination Review
   - Data Quality Assessments

VII. WHAT PROCESSES DOES DOT UTILIZE TO ENSURE INFORMATION QUALITY BEFORE IT IS DISSEMINATED?

a. Each DOT organization will conduct a pre-dissemination review on all information it disseminates on or after October 1, 2002**.** During this review, each DOT organization may utilize internal peer reviews and other review mechanisms to ensure the quality of all disseminated information. The costs and benefits of using a higher quality standard or a more extensive review process will be considered in deciding the appropriate level of review and documentation. At a minimum, DOT organizations will:

1. Allow adequate time for reviews, consistent with the level of standards required for the type of information to be disseminated. Consult with others (e.g., other DOT organizations, the public, State governments, etc…) that have a substantial interest in the proposed dissemination of the information.

2. Verify compliance with these guidelines (i.e., utility, objectivity, integrity and accessibility requirements) as well as other DOT organization specific guidance/procedures;

3. Indicate (for record keeping purposes) whether proposed information is considered "influential" and if so, what additional standards will be applied to ensure its quality. Refer to section XI of these guidelines for the term "influential" as it applies to DOT's information;

4. Ensure that the entire report fulfills the intentions stated and that the conclusions are consistent with the evidence;

5. Indicate origin of data (when including data from an external source); and

6. Beginning on the effective date of these guidelines, each DOT organization will include the following Notice on all substantive disseminated information products subject to these guidelines:

    - **Information Quality Notice**:
      This information is subject to the Department of Transportation's information quality guidelines (cio.ost.dot.gov/policy.xxx).

7. Ensure that each program office can provide additional data on the subject matter of any covered information it disseminates.

VIII. WHAT ARE DOT'S PROCEDURES CONCERNING REQUESTS FOR CORRECTION OF INFORMATION?

a. Who may request a correction of information from the Department?

Any person who is affected by information that the Department has disseminated on or after October 1, 2002 may request that the Department correct that information.

An affected person(s) is one who may benefit from or be harmed by the disseminated information.

b. How and where are requests for correction of information submitted?

The Department's Docket Management System (DMS) will be the primary point of entry for requests for correction of information DOT has disseminated. Interested persons should use an electronic form for electronic submission to the Department.

Incoming requests for correction, requests for reconsideration and DOT organizational responses will be posted on the DMS website. DMS will electronically notify designated data quality officials in DOT organizations that a request for correction/reconsideration is pending. In the event that DOT staff receives requests for correction/reconsideration by another means (e.g., mail to a DOT office), the staff will refer the request to the Office of Dockets and Media Management for inclusion in the DMS.

Requests for correction of information/or request for reconsideration may be made via the on-line form which can be accessed from the DMS or the Department's customer support web site. (Access DOT's online request form here.) Although a completed on-line form is preferred, DOT will also respond to request in written form, by letter, fax or e-mail to the following address:

U. S. Department of Transportation (DOT)
Office of Dockets and Media Management
SUBJECT: Request for Correction of Information
Room PL-401
400 7th Street, S.W.,
Washington, DC 20590
Fax number: 202/366-7202

Note: For a request for reconsideration, you must include a reference to your initially assigned DOT docket number.

c. What You Should Include in a Request for Correction of Information?

In keeping with the non-regulatory nature of these guidelines, this guidance-for the content of requests for correction of information - is not intended to constitute a set of legally binding requirements. However, DOT may be unable to process, in a timely fashion or at all, requests that omit one or more of the requested elements. DOT will attempt to contact and work with requesters to obtain additional information when warranted.

The following information is also provided on the electronic form mentioned in section VIII (b) above.

1. You should include a statement that a request for correction of information is submitted under DOT's Information Dissemination Quality Guidelines;

2. You should include your name, mailing address, fax number, or e-mail address, telephone number and organizational affiliation, if any); DOT will not respond to anonymous requests;

   Privacy Act Statement: DOT is authorized to obtain certain information under Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (Public Law No. 106-554, codified at 44 U.S.C. § 3516, note). Information (as identified in Section VI) will be needed to process requests and allow DOT to reply accordingly. This information is needed to respond to your request and initiate follow-up contact with you if required. Please do not send us your Social Security Number. You are advised that you do not have to furnish the information but failure to do so may prevent your request from being processed. The information you furnish is almost never used for any purpose other than to process and respond to your request. However, we may disclose information to a congressional office in response to an inquiry made on your behalf, to the Department of Justice, a court, other tribunal when the information is relevant and necessary to litigation, or to a contractor or another Federal agency to help accomplish a function related to this process;

3. You should describe how the information in question affects you (e.g., how an alleged error harms you, and/or how the correction will benefit you);

4. You should clearly identify the report, data set, or other document that contains the information you want the Department to correct. Please be as specific as possible, include such identifying characteristics as title, date, how information was received (e.g., web-accessed, etc.);

5. You should clearly identify the specific information that you want the Department to correct. Please be as specific as possible, include such identifying characteristics as the name of DOT agency that originated the data, title, date, etc**.** For example, you should not rely solely on general statements that allege some type of error. Requests for information that are specific and provide evidence to support the need for correction will likely be more persuasive than requests that are general, unfocused, or simply indicate disagreement with the information in question;

6. You should specify, in detail, why you believe the information in question is inconsistent with the Department's information quality guidelines (i.e., how the information fails to meet standards of integrity, utility, and/or objectivity);

7. You should specify your recommendations for what corrections DOT should make to the information in question and reasons for believing that these recommended corrections would make the information consistent with the DOT's information quality guidelines;

8. In a case where the Department has not designated a report, data set, or document as being subject to these information quality guidelines, and you believe it should be, you should specify why the information should be subject to the guidelines; and

9. You should include any documentary evidence you believe is relevant to your request (e.g., comparable data or research results on the same topic).

d. <u>May the Department reject a request for correction of information?</u>

Once the appropriate data quality official has received your request for correction of information from DMS, he/she will review your request and answer the following questions to determine if your request for correction is a valid request:

1. <u>Did DOT (as opposed to some other person or organization) actually disseminate the information you are requesting to be corrected?</u>

2. <u>Did DOT disseminate this information recently (i.e., within one year of your request), or does it have a continuing significant impact on DOT projects or policy decisions or on important private sector decisions?</u>

3. <u>Are you a person affected by the information in question?</u>

4. <u>Is the information you are requesting DOT to correct exempt from these Guidelines (see Section IV)?</u>

5. Is your request frivolous or not germane to the substance of the information in question?

6. Has DOT responded previously to a request that is the same or substantively very similar?

7. With respect to information in a final rule, final environmental impact statement, or other final document on which there was an opportunity for public comment or participation, could interested persons have requested the correction of the information at the proposed stage?

If the DOT organization determines that the answer to Question d1, d2, or d3 is "no" or that the answer to Question d4, d5, d6 or d7 is "yes," DOT will reject your request.

If DOT rejects your request on these grounds, the DOT organization will send a written response explaining why. Normally, the DOT organization will send this response within 45 days of receiving your request. The DOT organization will file this response in the DMS.

If the DOT organization does not reject your request on these grounds, it will consider the request on its merits.

e. What determinations does the Department make concerning a request for correction of information?

[*Requesters should be aware that they bear the "burden of proof" with respect to the necessity for correction as well as with respect to the type of correction they seek*].

If the DOT organization considers your request on its merits, it will make the following determinations: 1) Whether information subject to the DOT information quality guidelines complies with the guidelines; and 2) In the case of information that the Department did not designate as subject to the information quality guidelines (see Section IV of these procedures), whether the information should be subject to the guidelines.

If information subject to the DOT information quality guidelines does not comply with the guidelines, the Department is not required to change, or in any way alter, the content or status of information simply based on the receipt of an inquiry or request for correction. For example, DOT need not withdraw an information product from a website just because a request for correction has been received with respect to it.

Except with respect to information covered in VIII (d) (7) of these procedures, the DOT organization provides a written response directly to the requester,

including the Department's determinations on the items listed in paragraph VIII (e)(1) above and the reasons for these determinations.

The DOT organization will normally issue this response within-90-calendar days of receiving the request. If the DOT organization's response will take significantly longer than this period, the DOT organization will provide a written explanation to the requester, which will also be filed in DMS.

When the DOT organization determines that a correction of the information is warranted, revisions/corrections to the information in question will begin as quickly as practicable. However, the Department's budget, resources, and priorities, as well as the complexity of the correction task itself, may result in DOT actually taking this corrective action within a reasonable time after the Department has made the determination that a correction is appropriate.

The receipt of a request, or the consideration of a request by the Department, does not result in staying or changing any action of the Department, or the withdrawal or suspension of any information product. By itself, a request for correction does not affect the finality of any decision of a DOT organization.

f. How does the Department process requests for correction concerning information on which the Department has sought public comment?

This section concerns requests for correction concerning information on which a DOT organization has sought public comment (e.g., a notice of proposed rulemaking (NPRM), studies cited in an NPRM, a regulatory evaluation or cost-benefit analysis pertaining to the NPRM; a draft environmental impact statement; a proposed policy notice or aviation order on which comment has been sought; a request for comments on an information collection subject to the Paperwork Reduction Act).

The DOT organization's response to the request for correction will normally be incorporated in the next document it issues in the matter concerning which it had sought comment (e.g., in the case of an NPRM, the preamble to the final rule), DOT may choose to provide an earlier response, if doing so is appropriate and will not delay the issuance of the final action in the matter. Once again, the DOT organization will place their response in the DMS. As stated above in VIII (d), a DOT organization may reject a request for correction with respect to information in a final document if there was an opportunity for public comment or participation and interested persons could have requested the correction of the information at the proposed stage.

g. May a requester obtain reconsideration of a DOT organization's decision on a request for the correction of information?

You may request reconsideration under this section only if you have requested a correction of information under these guidelines, and you are not satisfied with the DOT organization's response.

IX. HOW DOES THE DEPARTMENT PROCESS REQUESTS FOR RECONSIDERATION OF DOT ORGANIZATION DECISIONS ON REQUESTS FOR CORRECTION OF INFORMATION?

You should send your request in the same manner, and to the same address, as provided in section VIII of these procedures. When completing the electronic DMS form, you need only complete Section II-"Request for Reconsideration".

You should request reconsideration within 30 days of the date you received the DOT organization's decision on your original request for correction.

If there is an existing process for reconsidering a particular sort of information disseminated by DOT, the DOT organization will make use of that process. For example, if the information relates to a final rule a DOT organization has issued, and the DOT organization has an existing process for handling requests for the reconsideration of a final rule, the DOT organization would use that process. If the information relates to a final EIS, the DOT organization may handle the request as though it were a request for a Supplemental EIS.

In the absence of an existing applicable reconsideration process, the DOT organization will designate a reconsideration official. This official should be someone who can offer objectivity (i.e., was not involved in making the decision on the original request for correction or in producing the underlying information) and who has a reasonable knowledge of the subject matter. The official can either be within the DOT organization to which the request for reconsideration pertains or in another DOT organization. In appropriate cases, the Department may designate a panel of officials to perform this function.

The reconsideration official will determine if additional corrective action is needed. This determination may pertain to the specific correction that is appropriate in a given case as well as to the issue of whether correction is merited at all. The reconsideration official will issue a written response to the requester stating the reasons for the decision.

The DOT organization will normally issue this response within 90 calendar days of receiving the request for reconsideration. If the DOT organization's response will take significantly longer than this period, the DOT organization will provide

a written explanation to the requester. This response will be filed in the DMS. The reconsideration official's determination will also be filed in DMS.

X. WHAT ARE THE DEPARTMENT'S REPORTING REQUIREMENTS?

The Departmental Office of the Chief Information Officer will:

a. Revise draft report (which provides DOT's draft guidelines; explains how these guidelines will meet OMB's standards; and details agency's administrative mechanisms) after considering public comment and submit to OMB for review by July 1, 2002;

b. Publish notice of availability of final report, incorporating any changes (after OMB's review) and post final report on agency website by October 1, 2002;

c. Issue final information quality guidelines by October 1, 2002; and

d. Provide annual reports to OMB (which will include the number and nature of inquiries received concerning agency compliance to these guidelines as well as how inquiries were resolved) beginning January 1, 2004.

XI. WHAT ARE THE DEFINITIONS ASSOCIATED WITH THESE GUIDELINES?

DOT has adopted the definitions of terms set forth in The Office of Management and Budget's Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies. The following information explains further the way that DOT uses some of these terms.

a. **Influential:** The following discussion is intended as guidance for DOT officials and other interested persons in determining whether scientific, financial, or statistical information is influential within the meaning of OMB's guidelines. This definition is important because it determines the level of scrutiny afforded to information.

The OMB guidelines define "influential" information as information that the agency reasonably can determine "will have or does have a clear and substantial impact on important public policies or important private sector decisions." The guidelines assign to each agency the task of defining this term in ways appropriate to the agency and its various programs.

DOT emphasizes that to be influential; information must have a clear and substantial impact. A clear and substantial impact, first of all, is one that the agency is firmly convinced has a high probability of occurring. If it's merely

arguable that an impact will occur, or if it is a close judgment call, then the impact is probably not clear and substantial. To determine that there is a clear and substantial impact, the agency must have greater certainty than would be the case for many ordinary factual determinations. The impact must be on "important" public policy or private sector decisions. Even if information has a clear and substantial impact, it is not influential if the impact is not on a public or private decision that is important to policy, economic, or other decisions.

OMB's guidelines' definition of this term applies only to scientific, financial, or statistical information. The definition does not address other types of information, no matter how important the information may seem to be. It should also be noted that the definition applies to "information" itself, not to decisions that the information may support. Even if a decision or action by DOT is itself very important, a particular piece of information supporting it may or may not be "influential."

In rulemaking, influential information is scientific, financial, or statistical information that can reasonably be regarded outcome determinative with respect to one or more key issues in a significant rulemaking, as that term is defined in Executive Order 12886. The "outcome determinative" part of this standard reflects the "clear and substantial impact" language in the OMB guidelines language. The reference to key issues on significant rules reflects the "important" public policy language of the guidelines.

In non-rulemaking contexts, DOT will consider two factors – breadth and intensity – in determining whether information is influential.

Every decision DOT makes based on disseminated information is important to someone. That does not mean that disseminated information used for each decision is influential, as the term is used in the guidelines.

In determining whether information is influential, DOT organizations should consider whether the information affects a broad range of parties. Information that affects a broad, rather than a narrow, range of parties (e.g., an entire industry or a significant part of an industry, as opposed to a single company) is more likely to be influential.

DOT organizations will also consider whether the information has an intense impact. Information that has a low cost or modest impact on affected parties is less likely to be influential than information that can have a very costly or crucial impact. In considering whether information has a high-intensity impact, DOT organizations will establish and use as a benchmark the $100 million figure used to determine whether a rule is economically significant.

In most cases, information that has an intense impact on a broad range of parties should be regarded as influential. Information that affects a broad range of parties, with a low-intensity impact, or information that affects a narrow range of parties, with a high-intensity impact, may or may not be influential.

DOT organizations may designate certain classes of information as "influential" or not in the context of their specific programs. Absent such designations, DOT organizations will determine whether information is influential on a case-by-case basis, using the principles articulated in these guidelines.

The "influential" designation is intended to be applied to information sparingly. DOT organizations should not designate information products or types of information as influential on a regular or routine basis. Nor should DOT organizations actually place an "influential" label in the title page or text of an information product.

b. **Reproducibility.** Documented methods are capable of being used on the same data set to achieve a consistent result. For more information on this term, please refer to OMB's guidelines.

c. **Dissemination.** As provided in OMB's guidelines, these guidelines apply only to information disseminated on or after October 1, 2002. The fact that an information product that was disseminated by DOT before this date is still maintained by the Department (e.g., in DOT's files, in publications that DOT continues to distribute on a website) does not make the information subject to these guidelines or to the request for correction process.

   For example, suppose that DOT first issued a study in 1999. The study is relied upon in a 2000 DOT organization publication, and the DOT organization makes the study available on its website. This study is not subject to these guidelines or to the request for correction process. However, if DOT issues a notice of proposed rulemaking in 2003 that relies on the same study, then it becomes subject to these guidelines - because it then has been disseminated (or, one might say "re-disseminated") after October 1, 2002.

d. **Departmental organizations.** Offices within the Office of the Secretary, Operating Administrations (OA), offices, divisions, and comparable elements of the DOT.

e. **Departmental Chief Information Officer (CIO).** The Departmental CIO is the senior management official responsible for the DOT Information Dissemination Quality Program.

f. **Data Quality Administrator (DQA)**. Designated representative in the Office of the CIO (S-80) responsible for compiling agency reports and serving as agency liaison to OMB.

g. **Data Quality Official (DQO)**. The DQO serve as the point of contact for the Departmental CIO/Data Quality Administrator and will be responsible for implementing these guidelines within their organization.

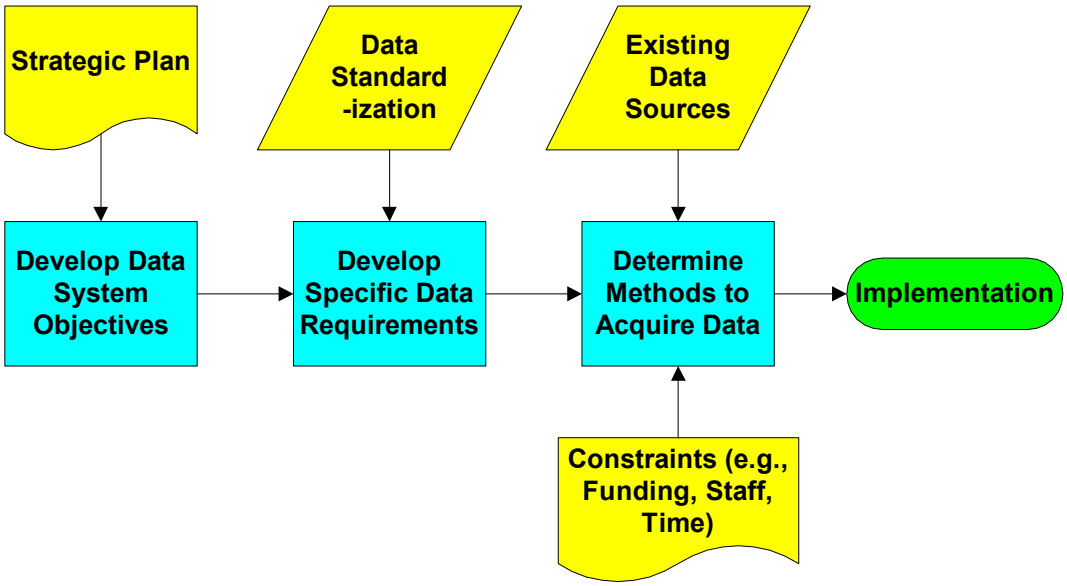
h. **Docket Management System (DMS).** DMS is an electronic, image-based database in which all DOT docketed information is stored for easy research, and retrieval.

i. **Docket.** A docket is an official public record. DOT publishes and stores on-line information about proposed and final regulations, copies of public comments on proposed rules, and related information in the DMS. DOT uses this docketed material when making regulatory and adjudicatory decisions, and makes docketed material available for review by interested parties. Specific documents covering the same issues are stored together in a docket.

j. **Transparency:** Providing a full description of information, methods, assumptions, and sources of error.

VI a. Planning Data Systems

A data system produced within a DOT agency is inseparably linked to that organization's strategic plan. The data is compiled to measure success toward a goal, satisfy an external user need (which should also be a goal), or used as a tool necessary to perform work toward a goal. Data system planning consists of three stages: development of objectives for the system, translation of those objectives into data requirements, and planning of the top-level methods that will be used to acquire the data.



VI a.1: Data System Objectives

Principles

- "Objectives" of the data system describe what federal programs and external users will accomplish with the information.
- There is an assumption that the sponsoring organization's strategic plan is current and contains all of its goals and objectives, including those relative to the creation of the data system.
- The first step to guaranteeing that the data system will help fulfill strategic plans is to create data system objectives that are traceable to the strategic plan goals and objectives.
- Clear objectives in specific terms, identifying data users, and key questions to be answered by the data system are more likely to guide the system development to produce the results required.
- Just as strategic plans change over time, the objectives of the data system will need to change over time to meet new requirements.
- Users will benefit from knowing the objectives that guided the system design.

APPENDIX A

Guidelines

- Every data system objective should be traceable to the goals and objectives in the sponsoring organization's strategic plan.
- The system sponsor should develop the data system objectives in partnership with critical users and stakeholders.
- The objectives should indicate each major need that will be fulfilled by the system and the data users associated with that need.
- Objectives should indicate every key question that the data will answer.
- Objectives should be written in terms of what will be accomplished with the data, not in terms of the data itself.
- The objectives should be used in the data system design, creation, and maintenance.
- The objectives should be documented and clearly posted with the data.
- The sponsor should review the objectives periodically to ensure that the data collection system continues to meet evolving user needs.

Examples and Additional References

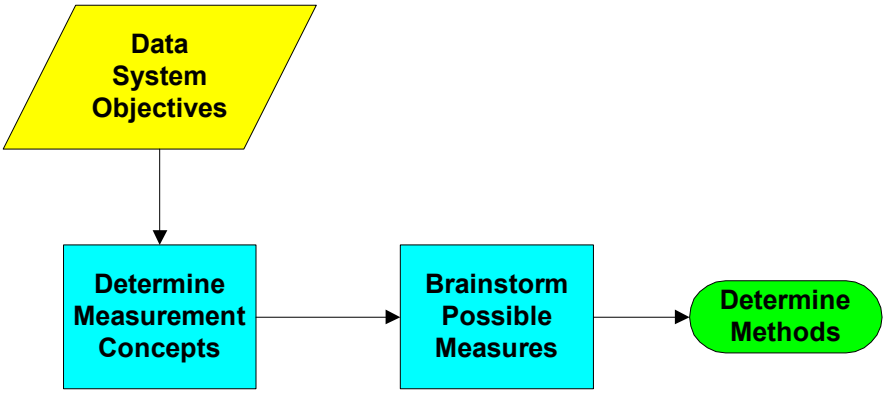
- Fatality Analysis Reporting System (FARS):

  To provide an overall measure of highway safety, to help suggest solutions, and to help provide an objective basis to evaluate the effectiveness of motor vehicle safety standards and highway safety programs.
  To reduce the number of motor vehicle crashes and deaths on our nation's highways . . . FARS data are critical to understanding the characteristics of the environment, traffic way, vehicles, and persons involved in the crash.

- Highway Pavement Monitoring System (HPMS)

  To provide state and national level measures of the overall condition of the nation's open-to-traffic, public road systems.
  To provide a measure of state and national use of open-to-traffic, public road systems.

VI a.2: Data Requirements



Principles

- Before deciding on specific data elements for a data system or how to acquire them, the data system objectives need to be translated to "measurement concepts," from which data requirements will be derived.
- This link from data system objectives to measurement concepts to data requirements will help to ensure "relevance" of the resulting data to users and strategic plans.
- A "measurement concept" is a characteristic of individual units (e.g., people, businesses, objects, events) or groups of units (e.g., people or businesses in a city or state, cars or trains in the United States, actions at an airport or on highways). Examples: The level of success stopping illicit drug smuggling into the U.S. over maritime routes, or the safety of people and pedestrians on the highways of the U.S.
- Information on the link from the data back to measurement concepts will help data users interpret data from the system and statistics derived from it.
- Measurement concepts related to objectives can be outcomes that will change when objectives are achieved, outputs from agency accomplishments related to an objective, efficiency concepts, inputs, and quality of work.
- The next step is to identify data requirements for possible measurement of each measurement concept.
- The use of standard names, variables, numerical units, and codes allow data comparisons across databases.
- In addition to measurement concepts and data requirements that are directly related to strategic plans, additional data may be required for possible cause and effect analysis. In theory, these should be part of the original objectives.
- If we try to limit the amount of information collected, it will reduce burden on the data suppliers and make collecting the data easier.

Guidelines

- The first stage is to take each objective and determine what "concept" you need to measure. Each objective should have one or more measurement concepts. Every characteristic or attribute of the target group or groups that are the focus of the objective should be covered by one or more measurement concepts.
- The measurement concepts should be those characteristics which, when changing in a favorable way, indicate achievement or progress toward achievement of an objective.
- Once you have the measurement concepts, the next step will be to develop data requirements needed to quantify it.
- There is usually more than one way to quantify a measurement concept. All reasonable measures should be considered without regard to source or availability of data. The final data choices will be made in the "methods" phase based on ease of acquisition, constraining factors (e.g., cost, time, legal factors), and accuracy of available data.
- The data requirements for each type of data should include required accuracy, timeliness, and completeness. The accuracy should be based on the amount of change that you need to be able to detect between two data elements at different times or between two groups, or the maximum error that is tolerable in the primary estimator.
- When selecting possible data, an effort should be made to consider standardization with other databases. First, you should consider measures used for similar concepts in other DOT databases. Second, consider measures for similar concepts in databases outside DOT (e.g., The Census).
- Coding standards such as the North American Industry Classification System (NAICS) codes should be used where appropriate. Such standardization leads to "coherence" across datasets.
- Needless to say, if law or regulation requires a specific measure be implemented, it should be included (it should also be in the objectives).
- For each data element, determine what the accuracy, timeliness and completeness required to make the data useful.
- Examine the data requirements for variables that have limited or no defined need and eliminate questionable data requirements to reduce data provider burden.

Examples and Additional References

- FARS:

  The objective "To provide an overall measure of highway safety" leads to the measurement concept of "The safety of people and pedestrians on the highways of the U.S." which leads to data requirements for counts of fatalities, injuries, and motor vehicle crashes on U.S. highways and streets. The fatalities for a fiscal year should be as accurate as possible (100% data),

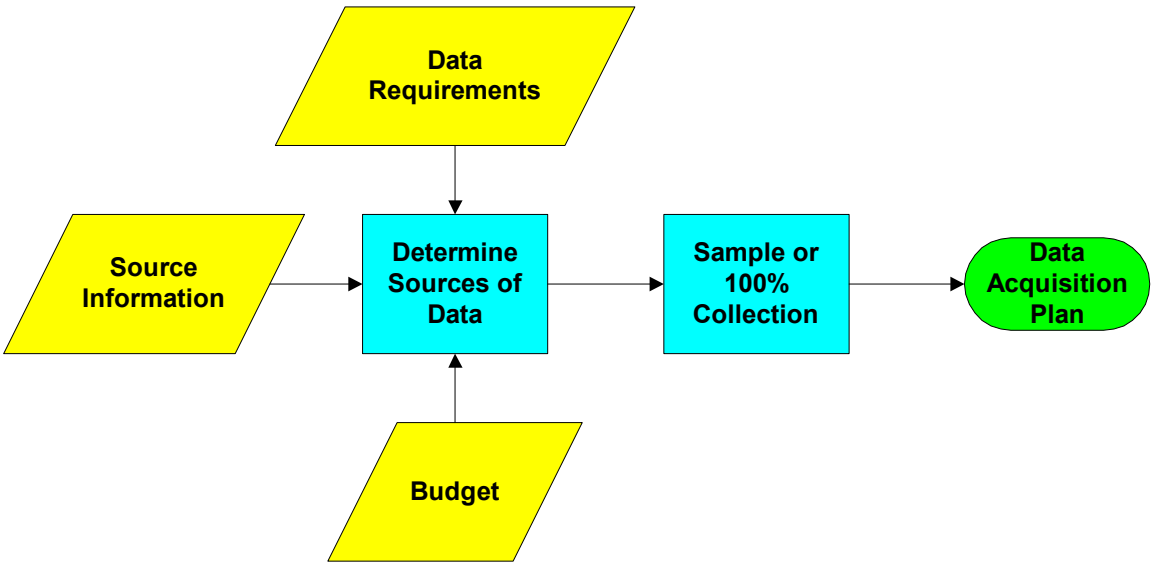
available within three months after the end of the fiscal year, and as complete as possible. The injuries and crashes for the fiscal year totals should have a standard error of no more than 50,000, be available within three months after the end of the fiscal year, and have an accident coverage of 90 percent.

- HPMS:

  The objective "to provide a measure of highway road use" leads to the measurement concept of "the annual volume of vehicles on state and interstate roads," which in turn leads to data requirement for estimates of annual vehicle-miles traveled accurate at the state level to within 10 percent at 80 percent confidence.

VI a.3: Methods to Acquire Data

Given data requirements for a wide range of possible measures, the next phase is to consider the realities associated with gathering the data to construct estimates and perform analysis. After looking at the ease of data acquisition, complexity of possible acquisition approaches, budget restrictions, and time considerations, the list of possible measures is likely to be reduced to a more reasonable level. First, you should consider possible sources of data and then the process of acquiring it.



The more critical data needs invariably require greater accuracy. This in turn usually leads to a more complex data collection process. As the process gets more complex, there is no substitute for expertise. If you do not have the expertise for a complex design in-house, you should consider acquiring the expertise by contacting an agency that specializing in statistical data collection like the Bureau of Transportation Statistics or getting contractual support.

VI a.4: Determine Sources of Data

Principles

- Some data systems have data collections that are specified by law (e.g., 46 USC 6102 specifies a marine casualty reporting system). That limits the collection planning to working out the physical details.
- Use of existing data is by far the most efficient (i.e., cheapest) approach to data acquisition. Sources of existing data can be current data systems or administrative records.
- Administrative records are data that are collected for some administrative reason (e.g., state driver's license records, social security records). Administrative records may also be used as information helpful in designing a data collection system (e.g., sampling lists, stratification information)
- Another approach, less costly than developing a new data collection system, is to determine if existing data collections are available to tailor to your needs. The sponsor may be willing to change the collection process to gather data for your needs.
- The "target group" is the group of all people, businesses, objects, or events that you want information about. One possible source is to go directly to the entire "target group" (100%) or a sample of them.
- "Proxies" are another possible source of data. They are people or businesses that have knowledge about the target group, such as experts, investigators, observers, or service providers (e.g., doctors). These may be existing sources of data or ones created by your collection process.

Guidelines

- Research whether government and private data gathering systems already have data that meet the data requirements. Consider surveys, reporting systems, and administrative records.
- If existing data meets some but not all of your requirements, determine whether the existing data collection system can be altered to meet your needs. For example, another agency may be willing to add to or alter their process in exchange for financial support.
- A primary consideration in whether to gather data from the target group or proxies is the access to those sources; all of those sources. A 100% data gathering would obviously need access to the entire target group. A sample approach will not include the entire target group, but all members must have a probability of selection, or the sampling will not necessary be representative of the target group.

Examples and Additional References

- <To Be Developed>

VI a.5: Sample Design or 100% Collection

Principles

- For large target groups, data collection from 100% of the target group is usually the most accurate approach, but is rarely feasible due to cost, time, and other resource restrictions. It also is often far more accurate than the data requirements demand.
- A "probability sample" is an efficient way to automatically select a data collection representative of the target group with the accuracy determined by the size of the sample.
- For a sample to be truly representative of the target group, the sample selection process must ensure that all members of the target group have a non-zero probability of selection. In reality, when the target group is large, and complex this may not be feasible, but the design should attempt to do so to the greatest extent possible.
- Sampling lists (also known as frames) of the target group are required to select the sample. Availability of such lists is often a restriction to the method used in data collection.
- Multistage sampling designs may be necessary or advisable based on the physical circumstances. For example, geographically sampling counties, and then sampling the licenses from within those counties, is a method for taking a manageable sample of U.S. driver's license information.
- Samples are often stratified. Stratification is the separating of the sample list into groups and then sampling from each group as a separate act. It reduces the variability of estimators for the target group and ensures the accuracy of selected subgroups of the target group.
- Some principles for determining sample size:
  - The size of the sample required for a data collection is a fairly complex issue. It should be based on the accuracy (within a specified level of confidence) indicated in the data requirements. Since the data requirements specify accuracies for all of the measurements, you need to decide which data elements are most important.
  - One basic formula for sample size is based on the accuracy requirement of a target group mean derived from the formula for the 95% confidence interval, $\bar{x} \pm 2\sigma / \sqrt{n}$. So, if the required accuracy is ± 5, then the sample size should be $n = \frac{4\sigma^2}{25}$. The estimated value for σ^2 is usually based on past information or knowledge of similar data or studies. It should also be adjusted for "design effect" for stratified and other design complexities. Similarly, if an accuracy requirement exists for detecting a change between estimates from two sets of data over time, a similar formula for that can be derived.
  - Of course, the sample size is also very dependent upon funding.

  - Some of the initially sampled cases may not yield useful data (due to ineligibility, nonresponse, lost data, etc.). In anticipation, sample sizes are usually adjusted upward to compensate.
  - Sample sizes within strata are often allocated from the total sample size using a formula referred to as Neyman allocation (see Cochran 1977) with adjustments for minimum stratum sample sizes.
- Two alternatives for sample selection within each stratum are simple random samples or probability proportional to size (PPS) sampling, where the size is some variable, known for each item, which is related to what you are estimating. Systematic approaches may also be used for convenience and to invoke an additional form of stratification. A systematic sample is drawn by selecting every k^{th} item, where 1/k is the required sampling fraction. Once again, a statistician or other sampling expert should perform the selection.

Guidelines

- For a large target group, unless a 100% collection is required by law or implied by accuracy requirements, you should use a probability sample.
- The sample design should give all members of the target group a non-zero probability of being represented in the sample.
- Use a multi-stage sample, when appropriate to simplify the collection process or reduce variability.
- Stratify during each stage of the sample selection to reduce variability of estimators and ensure accuracy of estimation for sub-groups.
- Determine sample size formulas to ensure data requirements for accuracy are met using appropriate formulas.
- Allocate sample to strata to minimize variability for target group estimators, taking into account minimum sample sizes for sub-group estimation
- Use an appropriate random method to sample according to the design.
- A statistician or other sampling expert should be involved in the design of a data collection involving sampling.

Examples and Additional References

- Cochran, William G., *Sampling Techniques* (3^{rd} Ed.), New York: Wiley, 1977.
- <To Be Developed>

VI b. Collection of Data

VI b.1: Forms

Principles

- Forms are sets of questions or requests for information designed to elicit information from data suppliers. Forms are the interface between the data system sponsor and the data supplier. As such they play a central role in many data collection systems, and can significantly affect data quality.
- Questionnaire and form design is a well-developed field of expertise, based on extensive experience and cognitive testing. Many lessons have been learned in the laboratory and actual use of forms about spacing of information, use of fonts, separation of different types of information, how to ensure the data supplier provides only the appropriate information, use of color, and proper wording.
- Screen design for automated data collection follows similar methods.
- On most survey questionnaires and some data forms, the answer to a question or information entered in one place will determine if the next question or entry is filled in or "skipped" over. For example, a question asking, "Do you own a car?" may be followed by several questions about the car. If you answer "No" to the first question, the form will tell you to skip over the follow-on questions. Such a design element is called a "skip pattern."

Guidelines for Form Content and Instructions

- Each data element should support program goals and not unnecessarily duplicate information available elsewhere.
- Wording and concepts should be designed for the data suppliers, who may not use the same wording or concepts as the data collection sponsors. Any testing of the form should involve subjects as similar to the actual data providers as possible.
- Time reference periods and units of response should be clear to the data supplier. As far as possible, choose questions, time reference periods, and response categories that are compatible with the record-keeping practices.
- Response categories should be mutually exclusive and exhaustive. However, avoid a large number of categories in one question. (data suppliers will tend to the top of the list)
- Text fields (e.g., fill in the blank) should be used sparingly. Standardizing and coding text fields is a time-consuming, difficult, and error-prone task. Text fields that are not coded are difficult to use analytically.
- Questions should be as self-explanatory as possible, reducing the need for additional instructions.
- To the extent feasible, instructions should be short, clear, and appear as close to question as possible. Try to avoid placing instructions on a separate sheet or on the reverse of page with the questions.

- Instructions should include a clear, specific statement of who is required to file the form, when it should be filed, and where it should be sent. The phone number for questions from the data supplier should be prominently displayed.

Guidelines for Form Layout

- The introduction on the form should provide the title or subject of the data collection effort, identify the sponsor, explain the purpose, and cite the authority for data collection. If applicable, the introduction should indicate what confidentiality protection applies.
- Questions should be sequenced in a manner that is logical for the data supplier. Provide visual cues (e.g., titles or headings for each section) in forms, whether the forms are paper-based, displayed on a laptop screen, or Internet-based.
- Skip patterns (see definition under principles) should be clearly identified, and should ensure that only appropriate questions are asked of each respondent. Complex skip patterns should be avoided, especially for self-administered forms.
- Certification statement and signature block should be included when appropriate.
- Prior to the use of a form in actual data collection, it is often helpful to conduct a "pretest" of the collection process using the form. If possible, the test should be conducted in the actual data collection environment with actual data suppliers. A pretest can identify poor wording or ordering of questions or entries, errors in form layout or instructions, and problems caused by the data suppliers' inability to answer the questions or make the entries properly. Pretesting can also suggest additional response categories to pre-code onto the form.

Examples and Additional References

- <To be developed>

VI b.2: Data Collection Operations

Principles

- Data collection includes all the processes involved in acquiring data.
- Data collection operations can have a high impact on the ultimate data quality.
- The data collection method (e.g., mail, telephone, internet) should be appropriate to the complexity and amount of data requested, as well as to the funding level and amount of time available.
- The extent of the data collection complexity is also dependent upon the importance of the data quality and the level of response that is required.

Guidelines

- Computer-assisted information collection can result in more timely and accurate information at less cost. Initial development costs will be higher, and much more lead time will be required to develop, program, and test the data collection

system. However, since data can be edited during data entry, the lag between data collection and data availability will be reduced. Computer assisted collection should be considered when the collection is repetitive over a long period of time making the gains in quality and data processing time worth the expense.

- Electronic data collection requires effective control systems to ensure the security of data transmission and handling. Safeguards should be put in place to prevent loss of information (and the resultant loss in quality) due to system failures or human errors.
- If data entry is performed without a real-time edit capability, 100 percent verification should be required, and a 99 percent or better data entry accuracy rate (as measured on a key stroke basis) should be maintained for all entry operations.
- In data collection systems that use automated edits, documentation should be available specifying how error messages are to be handled, including a description of edit override procedures.
- Data entry documentation should also describe what records are to be maintained on the frequency of file changes and edit overrides. This information can be used to improve data collection forms and instructions.

Examples and Additional References

- <To be developed>

VI b.3: Response Rates

Principles

- Some missing data occurs in almost any data collection effort. Unit nonresponse occurs when a report that should have been received is completely missing or is received and cannot be used (e.g., garbled data, missing key variables). Item nonresponse occurs when data are missing for one or more items in an otherwise complete report.
- The extent of unit nonresponse can sometimes be difficult to determine. If a report should be filed whenever a certain kind of incident occurs, then non-reporters can only be identified if other data sources are available to be crosschecked. On the other hand, if companies are required to file periodic reports, the previous period can provide a list of the expected reporters for the current periods.
- Response rates indicate how successful the data collection system is at obtaining the responses it is designed to collect. A high response rate helps ensure that results are representative of the target population.
- Weighted response rates reflect the true impact of nonresponse. Differences between weighted and unweighted response rates also indicate whether data suppliers who responded were dissimilar from those who did not respond.

Guidelines

- Data collection programs should be conducted in a manner that encourages high rates of response.
- All data collection programs require some follow-up of missing reports and data items. The data system sponsor should plan follow-up to achieve acceptable response rates. Different follow-up methods should be evaluated to decide which method will produce the best results
- Data collection staff should follow-up large units first, possibly at the risk of missing smaller units. The degree to which complete response is pursued is based on budget and time constraints, fitness for use requirements and the risk of nonresponse bias.
- For missing data items the survey sponsor should distinguish between: critical items, which must be followed up (items legally required, likely to be influential, or perhaps used to measure DOT performance); important items, which would be sought if already following up on a critical item; and any remaining items.
- Both unit and item nonresponse contribute to bias of estimates when non-reporters differ from reporters in the affected items. It may be possible to link to external data sources (for example, other administrative data files). However, it is easier to compare the characteristics of the reporters and the non-reporters to see how they differ for known characteristics.
- Unit response rates (number of data supplier forms received divided by the number that should have been received) should be calculated and monitored. Useful response performance measures include the response rate by the specified receipt deadline, the response rate by the cutoff date for initial information dissemination, and the "final" response rate.
- Weighted response rates should be reported along with unweighted response rates.
- Unit response rates and response rates for key items should be reported to users of the data as part of the overall considerations of data quality.

Examples and Additional References

- <To be developed>

VI c. Processing Data

VI c.1: Data Editing and Coding

Principles

- Data editing is the application of checks that identify missing, invalid or inconsistent entries or that point to data records that are potentially in error.
- Typical data editing includes range checks, validity checks, and consistency checks (comparing answers to related questions).
- For numerical data, "outliers," especially moderate outliers, are not necessarily bad data. But, they are worth looking at for errors.
- Editing is a final inspection-correction method. It is almost always necessary, but data quality is better achieved much earlier in the process through clarity of definitions, forms design, data collection procedures, etc.
- "Statistical edits" are methods for examining statistical properties of the data to detect more subtle errors. They include examining distributions of variables for outliers, distribution anomalies, scatter plots of two related variables, and examining ratios of related variables or one variable over time.
- Coding, in general, is the addition of codes to the data set that indicate collection information or convert collected data, such as text data, into a form more useful during data analysis.
- Many coding schemes have been standardized. For example, the Federal Information Processing Standards (FIPS) have codes for Countries, States, Counties, Local Areas, Metropolitan Areas, and Occupations and the North American Industry Classification System (NAICS) codes.
- Automated coding and automation-assisted coding reduce coding error significantly.

Guidelines

- Editing should be performed early in the data collection process, if possible. Editing should also be one of the first steps in processing data that has been collected. Even if edits were used during data entry, not all problems may have been detected or resolved.
- The best approach to editing is to make as many editing decisions as possible in advance and automate it. Reliance on manual intervention in editing should be avoided, since it may introduce human error.
- Specifying edits requires both knowledge of the data files and knowledge of the subject matter. One person is not likely to have both kinds of knowledge, so a working group may be required to develop an editing system.
- Missing values should be clearly identified on the data file. They should have an unmistakable code indicating that it is a missing value and possibly why it is missing. Avoid using blanks or zeros, to indicate missing data—zeros, in particular, may be confused with legitimate reports.

- Consider the application of statistical edits. They can be difficult to develop, but can be very useful for detecting erroneous reporting of quantities (e.g., distance versus airtime).
- Avoid the overuse of outliers edits. Outliers can be very informative for analysis. Over-editing can lead to severe biases resulting from fitting data to implicit models imposed by the edits. Rapid industry changes, e.g., resulting from deregulation, could be missed if an agency follows an overly restrictive editing regimen that rejects large changes.
- Accuracy has to be balanced against timeliness in implementing an editing system. Prioritize time spent editing. Following up too many edits with data suppliers can raise data processing costs significantly (both budget and time) without leading to significant gains in data quality.
- A separate field containing a code (i.e., a "flag") should be added to the edited data file indicating which variables have been edited and by what editing method.
- Monitor editing performance by examining the number of edit flags generated, the number of items in a report with one or more edit flags, the number of items with flags that required contacting data suppliers, and the number of edit flags overridden.
- Reapply edits to records to which corrections were made to ensure that no further errors were inadvertently introduced.
- Editing should be designed very carefully in situations where the data system sponsor has less control over the data collection process.
- Codes should be added to the file for missing data and nonresponse, as discussed in sections 4.2 and 3.3, respectively.
- Text information that will be used for analysis should be converted to codes.
- Standard codes should be used whenever possible.
- Automated coding and automation assisted coding should be considered to reduce coding error.

Examples and Additional References

- <To be developed>

VI c.2: Handling Missing Data

Principles

- Referring to section 3.3 principles, the handling of missing data covers both unit nonresponse and item nonresponse.
- Untreated, missing data can introduce serious error into estimates. Frequently, there is a correlation between the characteristics of those missing and variables to be estimated. This results in biased estimates. For this reason, it is best to employ adjustments to mitigate this damage.
- The processing of missing data is usually done after editing, since editing often identifies "bad" data that should be regarded as missing.

- One method used primarily for unit nonresponse is weighting adjustments. All cases, including the missing cases, are put into classes using variables known for both types. Within the classes, the weights for the missing cases are evenly distributed among the non-missing cases.
- Imputation is a process that substitutes values for missing or inconsistent reported data. Such substitutions may be strongly implied by known information or derived as a statistical estimate.
- If missing data are used in estimates of totals, zero values are implicitly imputed for the missing data. These implicit imputations of zero cause the estimates of totals to underestimate the true totals.
- Ratios (or averages) using missing data implicitly impute the overall ratio for the missing data. Estimated ratios will be biased to the extent that the unknown ratio for the missing reports differs from that of the reports obtained.
- Due to the presence of imputation "flags," users can either use the imputed values or deal with the missing data themselves.

Guidelines

- An information product may contain estimates of totals or ratios, even though a significant portion of the values is missing and not imputed for a data item. To avoid the risk of misleading data users in such cases, report the item nonresponse rates and describe the risks of missing data.
- Unit nonresponse should normally be adjusted by a weighting adjustment as described above. One exception to this may be a data collection that collects the same information from the same subject over a relatively short time frame.
- The decision on when and how to impute (see definition above) for missing data is a complex statistical issue that should be made with input from missing data experts.
- The simplest form of imputation is logical imputation where other data collected in the same case, past data, or administrative data imply the correct missing value with near certainty. This method should be used for imputation if available.
- If a logical method of imputation is not available, then a statistical imputation method of estimation such as modeling, hot-deck substitution, or maximum likelihood should be applied.
- A separate filed containing a code (i.e., a flag) should be added to the imputed data file indicating which variables have been imputed and by what method.
- The method of imputation or weight adjustment should be fully documented and summarized in the data system's source and accuracy statement. Retain the unimputed and imputed values of the record's fields to evaluate the impact of imputation, which should also be reported in the source and accuracy statement.
- Some legal or administrative data systems may severely limit missing data correction. Since that may have severe consequences on various forms of analysis and estimation, the data system sponsor should inform the users of that limitation.

Examples and Additional References

- <To be determined>

VI c.3: Production of Estimates and Projections

Principles

- "Derived" data items are additional case-level data that is either directly calculated from other data collected (e.g., # of days from two dates), added from a separate data source (e.g., the weather on a given date), or some combination of the two (e.g., give the departing and arriving airports, calculating distance from an external source). Derived data is a way to enhance the data set without increasing respondent burden or significantly raising costs.
- An "estimate" is an approximation of some characteristic of the target group, like the average age, constructed from the data.
- A "projection" is a prediction of an outcome from the target group.
- Estimates from samples should be calculated taking the sample design into account. The most common way this is done is weighted averages using weights based on the design.
- Estimates of sampling error for an estimate will give an indication of the precision of the estimate.

Guidelines

- Use derived data to enhance the data set without additional burden on data suppliers.
- Weights should be used in all estimates from samples. Weights give the number of cases in the target group that each case represents, and are calculated as the inverse of the sampling probability.
- Consider adjusting weights for nonresponse as discussed in section 4.2.
- If you have "control" totals for the population or for subpopulations, you should consider a ratio adjustment of the weights or a form of raking to reduce the variance.
- Consider ratio estimates when you have preliminary estimates (e.g., partial year data) and you want to estimate the final value (e.g., whole year data), using known whole year data or past data.
- Sampling error estimates should accompany any estimates from samples.
- Sampling errors should be calculated taking the sample design in account. For more complex sample designs, use replicated methods (e.g., jackknife, successive differences) incorporating the sample weights. Consult with a sampling expert.
- When disseminated large data files from samples, consider disseminating generalized variances to allow users to calculate approximate variances for a wide range of estimates.
- Use auxiliary data whenever possible to improve the reliability of the estimates.

Examples and Additional References

- Cochran, William G., *Sampling Techniques* (3rd Ed.). New York: Wiley, 1977.
- <To be developed>

VI c.4: Data Analysis and Interpretation

Principles

- Data analysis starts with questions that need to be answered. Most of the questions are the general questions that stem from the data system objectives (see section 2.) of one or more data systems, refined to some specific area. Analyses should be designed to focus on answering the key questions rather than showing all data results from a collection.
- Analysis methods are designed around probability theory allowing the analyst to separate indications of information from uncertainty.
- Careful planning of complex analyses involving concerned parties will often ensure a successful result.
- For analysis of data collected using complex sample designs, such as surveys, the design must be taken into account when determining data analysis methods (e.g., use weights, replication for variances).
- Just because an analysis shows a significant relationship between two variables, does not mean the relationship is cause and effect. Adding additional related variables to the analysis will improve the likelihood that the true causal factors will become evident.
- A repeated analysis using data from a second source (if available) can strengthen the conclusions from the analysis, corroborating the results.
- 100% data collections do not have sampling error, though they are usually measuring a random phenomenon (e.g., highway fatalities).
- Data collected at sequential points in time often requires analysis with time series methods to account for inter-correlation.
- Interpretation is based on the analysis method assumptions, taking the limitation of the data and the analysis into account.
- Interpretation should take into account the stability of the process being analyzed. If the analysis interprets something about a process, but the process has been altered significantly since the data collection, the analysis results may have limited usefulness in decision making.
- Given the wide range of possible data analysis methods and the complexities of model selection, distributional assumptions, and statistical interpretation, important or complex analyses would be best designed and/or performed by a data analysis expert.

Guidelines

- The data analysis should begin with identifying the questions that need to be answered, specify the statistical methods that will be used, and conclude with analysis results and interpretation.
- All statistical methods used should be justifiable by statistical derivation or reference to statistical literature.
- Model assumptions should be made using distributions appropriate for the data being analyzed and models recognized in the particular application as most appropriate. Appropriateness should be justifiable by derivation or reference.
- For more complicated analysis projects, an analysis plan should be developed. Subject matter experts should review the plan to ensure that the analysis is relevant to the questions that need answering. Data analysis experts should also review the plan (even if written by one) to ensure proper methods are used.
- Results of the analysis should be written to focus on the questions that are answered, identify the methods used (along with the accompanying assumptions) with derivation or reference, and include limitations of the analysis.
- Any analysis of data collected using a complex sample design should incorporate the sample design into the methods via weights and changes to variance estimation (e.g., replication).
- Data analysis for causality or relationship between two or more variables should also involve other related variables to assist in the interpretation. For example, an analysis may find a relationship between race and travel habits. That analysis should probably include income, education, and other variables that vary with race. A subject matter expert should choose the related variables.
- If a second source of similar data is available, the analysis should corroborate the results using that second source.
- If a data analysis involves a series of statistical tests, the point of significance (often called the alpha value) should be adjusted. For example, many people use a "$p > 0.05$" as the determining criteria for significance. If five such tests were performed, that value (using one method) would be reduced to 0.01.
- Results from analysis of 100% data should not include tests or confidence intervals that are based on a sampling concept. Any test or confidence interval should use a measure of the variability of the underlying random phenomenon (e.g., regression standard error).
- A complex analysis of data collected at sequential time points should be done with time series methods.
- The interpretation of the analysis results should comment on the stability of the process analyzed.
- The results should indicate the importance of the data systems used in the analysis to the conclusions and possible actions that may be taken with the results.
- For more complex analyses, a statistician or other data analysis expert should design the analysis, implement the analysis, or both. Also, the more important the analysis, the more expertise and research time should be allowed to ensure accurate results.

- If the analysis produces useful results, the report of the analysis should acknowledge the importance of the data system in answering such questions and its continued existence to do so in the future.
- The analysis report should always contain a statement of the limitation including coverage and response limitations (e.g., not all private transit operators are included in the National Transit Database; any analysis should take this into account).

Examples and Additional References

- Choices of data analysis methods include descriptive statistics for each variable, a wide range of graphical methods, comparison tests, multiple linear regression, logistic regression, analysis of variance, nonparametric methods, nonlinear models, Bayesian methods, control charts, data mining, cluster analysis, and factor analysis (this list is not meant to be exhaustive).
- <To Be Developed>

VI d. Dissemination of Information

VI d.1 Presentation of Data

Principles

- Disseminating organizations may either summarize information for users (in tabular or graphical form) or release microdata for users to further analyze.
- In either case, information should be clearly presented to users, and users should be informed about the source(s) of the information presented.
- As far as possible, tables and graphs should be interpretable as stand-alone products in case they become separated from their original context.
- Making microdata available can enhance the usefulness of the information, and can assist the public in determining whether results are reproducible. However, microdata should not be released in violation of existing protections of privacy or proprietary information.

Guidelines for Tabular and Graphical Presentation

- All tables and graphs should include a title and a source.
- Titles for tables and graphs should be clearly worded and answer three questions: what (data presented), where (geographic area represented), and when (date covered by data).
- The source for a table or graph should contain one or more entries with references to the sources for the information presented. Reference to the primary source, e.g., the name of the database, is preferred.
- Since databases may be updated frequently, the "as of" date for the database should also be noted. Including the date of the database can help a user understand why estimates may not be exactly reproducible.

- In addition to a source statement, tables may require footnotes (referring to particular table entries), abbreviation symbol definitions, and notes (referring to the table in general) to help the reader understand the information presented.
- Unrounded numbers should be used in calculations if at all possible. If it is necessary to add, multiply, or divide numbers that have been rounded to different significant digits, the result can only be stated in terms of the number with the fewest significant digits.

Guidelines for Micro data Releases

- If possible, microdata should be released in formats (e.g., comma-separated values, common spreadsheet or database formats) that can be easily accessed by users with generally available software.
- Microdata releases should be accompanied by file layouts, definitions for codes used on the file, and a source and accuracy statement (in text, HTML, or other easily accessed format).
- Policies should be developed for archiving data for future secondary analysis.

Examples and Additional References

- Bureau of Transportation Statistics, *The BTS Style Guide for Authors and Editors,* [forthcoming 2002]
- Energy Information Administration, Chapter 5 "Tables and Graphs" in *EIA Publishing Style Guide* (2nd ed.), http://www.eia.doe.gov/neic/pubstyle/contents.htm (December 1997).
- Energy Information Administration, *EIA Guidelines for Statistical Graphs* (5th ed.), http://www.eia.doe.gov/neic/graphs/preface.htm.

VI d.2: Source and Accuracy Statements

Principles

- An effective information dissemination program provides not only the data, but also information about the sources, strengths, and weaknesses of the data.
- Source and accuracy statements help users to better understand the information disseminated and to assess whether the information is suitable for their intended use.
- Information about the limitations of the data should be made available for every data release, regardless of release format.

Guidelines

- All information products should have a source and accuracy statement containing basic information about the data sources used in the analysis, major known data limitations, and a contact for further information. The amount of detail reported with a particular information product depends on the nature of the product and its

intended use. If an abbreviated source and accuracy statement is provided, users should be told how to obtain the full source and accuracy details (e.g., on a website).

- The source description for an information release should include a full description of the purpose of the program, as well as the methods and assumptions used for data collection, processing, and reporting.
- The accuracy description for an information release presents what is known (and not known) about the quality and relevance of the data. It should include a description of each major source of error, the magnitude of the error source (if available), and any other known limitations of the data. Include findings of applicable validation and verification studies.
- If sampling was used, source statements should include a description of the sample design (sampling frame construction, weighting, and selection procedures). Accuracy statements should include procedures for calculating sampling errors and the estimates obtained.

Examples and Additional References

- Bureau of Transportation Statistics, "Source and Accuracy Compendium," http://www.bts.gov/statpol/SAcompendium.html.
- General Accounting Office, *Performance Plans: Selected Approaches for Verification and Validation of Agency Performance Information*, GAO/GGD-99-139 (July 1999).
- Office of Management and Budget, *Statistical Policy Working Paper 31: Measuring and Reporting Sources of Error in Surveys* (July 2001).

VI d.3: Documentation

Principles

- Statistical information should be reproducible, meaning that the documented methods are capable of being used on the same data set to achieve a consistent result.
- The more the influential the information is, the greater the need for clear, full, and complete documentation to facilitate reproducibility.
- A substantial amount of this material is typically available on these topics in the data collection specifications. The difficult task for the data producer is to synthesize the available material into a short technical report.

Guidelines

- Organizations should maintain documentation on obtaining, processing, and disseminating data that is sufficient to allow a qualified third party to reproduce the information product.
- Reproducibility of data is an indication of transparency in research design and methods. Replication is not required prior to each dissemination.
- <To be developed>

Examples and Additional References

- <To be developed>

VI d.4: Revisions

Principles

- Users should have the most current information, and should be aware when information may be, or has been, revised.

Guidelines for Revision Control

- Data systems should document each update to the database, indicating what the update was, whether the changes were reviewed, and the person authorizing the change
- Database home screens should indicate the date of the most recent update.
- Update history should be available to the user.

Guidelines for Revising Information After Dissemination

- Major errors in disseminated information should be corrected as soon as possible. Users should be informed through errata sheets, website notices, or other means, that the information has been changed.
- Due to the needs of users for the information, some data may routinely be published before data collection and validation are complete. These data should be clearly labeled as preliminary estimates and users should be informed about differences among preliminary, revised, and final estimates. When final data replace the preliminary estimates, revised values should be clearly identified.
- If changes are needed in data items or collection procedures for data series, the disseminating organization should provide adjustment methods, such as crosswalks and bridge studies, so that trend analysis is possible.
- Data system managers should conduct periodic reviews of the magnitude of changes between preliminary and final estimates. These reviews should be made available to users on the web.

Examples and Additional References

- <To be developed>

VI e. Evaluating Information Quality

VI e.1: Pre-Dissemination Review

Principles

- Information quality should be part of each step in the development of the information; the pre-dissemination review verifies the quality of the information to be disseminated for the public.
- The pre-dissemination review checks data quality and quality of information accompanying the data (i.e., metadata) for openness.

Guidelines

- Pre-dissemination review should verify compliance with the guidelines for dissemination of information (i.e., presentation, presence of source and accuracy statements, adequacy of documentation, and labeling of revisions).
- Pre-dissemination review should check for errors in the product (e.g., consistency of figures used in the text, tables and charts, verification of the accuracy of external data, and simple arithmetic). The review should also check consistency with other data sources, such as information from other data systems or from the same data system for previous time periods.
- For reports, the pre-dissemination review should ensure that the rest of the report fulfills the intentions stated in the introduction, and that the conclusions are consistent with the evidence.
- The reviewers should include at least one other knowledgeable person not involved in preparing the product for public dissemination. Such a reviewer is more likely to uncover unclear or misleading items in the information proposed for dissemination.
- Reviewers should be recognized as making a significant contribution to information quality. For proper planning and management of resources, the reviews should be scheduled well in advance, and the reviewers should be kept informed of any schedule changes.
- Since not all information products require the same level of review, organizations should classify products according to the required level of review. More attention is needed for products such as analysis reports and first releases of data series, especially when the content is deemed likely to be influential. Less attention would be required for recurring data releases, for summaries or highlights of information previously disseminated, or for information not likely to be influential.

Examples and Additional References

- <To be developed>

VI e.2: Data Quality Assessments

Principles

- "Data quality assessments" is a cumulative term for data quality audits or special studies of the data in the data system or some aspect of the data collection process.
- Data quality audits are comprehensive reviews of sources of error other potential quality problems in the data. They permit users to make more informed use of the data, and also help the data system sponsor to improve data quality.
- Data quality audits consist of a review of data system design, data collection procedures, procedures for processing data, and dissemination practices. It concludes with recommendations and suggestions for data quality improvements.
- Data quality assessment studies can include special data collections to assess any aspect of the data system, comparisons of the data with similar data from other systems, analysis of the data or a special collection to estimate aspects of nonsampling error (e.g., nonresponse, coverage, measurement error).

Guidelines

- Since data users do not have the same access to or exposure to information about the data system that its sponsors have, the data system sponsors should make the initial data quality assessment.
- Data quality assessments should be undertaken periodically to ensure that the quality of the information disseminated meets requirements.
- Data quality assessments should be used as part of a data system redesign effort.
- Data users, including secondary data users, should be consulted to suggest areas to be assessed, and to provide feedback on the usefulness of the data products.
- Knowledgeable persons not involved in preparing the information data for public dissemination may bring a useful perspective to an assessment team.
- Findings and results of a data quality assessment should always be documented. The issues investigated may continue to be relevant years later.
- Factors to consider in determining the appropriate level of data quality evaluation include whether the information is likely to be influential, the uses and users of the data, the potential for error and its impact on the use of the data, the variation in quality over time, and the potential for improvement of quality, efficiency or productivity.

Examples and Additional References

- General Accounting Office, *Performance Plans: Selected Approaches for Verification and Validation of Agency Performance Information*, GAO/GGD-99-139 (July 1999).

APPENDIX B

U.S. Department of Transportation
Request to Seek Correction of Information

*** THIS IS A SAMPLE FORM ONLY ***

In keeping with the non-regulatory nature of these guidelines, this guidance (for the content of requests for correction of information) is not intended to be legally binding requirements. However, DOT may be unable to process, in a timely fashion or at all, requests that omit one or more of the requested elements. DOT will attempt to contact and work with requestors to obtain additional information.

If you need further clarification of the questions below, please refer to the main guidelines.

SECTION I: REQUEST FOR CORRECTION

Contact Information:

First Name:	
Last Name:	
Email:	
Organization/Company:	
Phone:	
Fax:	
Title:	
Street Address:	
City:	
State/Province:	
Postal Code:	
Country:	

Describe how the information in question affects you (i.e., how an alleged error harms you, and/or how the correction will benefit you):

Publication Information

Clearly identify the report, data set, or other document that contains the information you want the Department to correct.

DOT Agency	BTS
Publication/Report Title	
Date of report:	
How did you receive the information? If found on a website, please indicate the URL:	

Clearly identify the specific information that you believe needs correction:

Specify, in detail, why you believe the information fails to meet standards of integrity, utility, and objectivity.

Specify your recommendations for what corrections DOT should make to the information in question and reasons for believing that these recommended corrections would make the information consistent with the DOT's information quality guidelines.

In a case where the Department has not designated a report, data set, or document as being subject to these information quality guidelines, and you

believe it should be, you should specify why the information should be subject to the guidelines; and include any documentary evidence you believe is relevant to your request (e.g., comparable data or research results on the same topic).

How would you like us to contact you? Mail

SECTION II: REQUEST FOR RECONSIDERATION

☐ Please check this box if this is a request for reconsideration and no more than 30 days has elapsed from the date you received DOT's response to your request for correction. (If this is a request for correction, please complete Section I of this form. You do not need to complete Section II.)

Please provide the following information relating to the request for correction submitted to the Department:

1. DOT Docket Number: OST- 2002 -

2. Date request for correction submitted:/ 2002 / Jan / 1

3. How did you submit request? Mail

4. Please provide a detailed explanation of why you are dissatisfied with DOT's response.

Privacy Statement | SUBMIT | RESET

U.S. Department of Transportation
Request to Seek Correction of Information
Privacy Act Statement

DOT is authorized to obtain certain information under Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (Public Law No. 106-554, codified at 44 U.S.C. § 3516, note). Information (as identified in Section IV) will be needed to process requests and allow DOT to reply accordingly. This information is needed to respond to your request and initiate follow-up contact with you if required. Please do not send us your Social Security Number. You are advised that you do not have to furnish the information but failure to do so may prevent your request from being processed. The information you furnish is almost never used for any purpose other than to process and respond to your request. However, we may disclose information to a congressional office in response to an inquiry made on your behalf, to the Department of Justice, a court, other tribunal when the information is relevant and necessary to litigation, or to a contractor or another Federal agency to help accomplish a function related to this process.