Bringing Data to Undergraduate Classrooms

The Social Science Data Analysis Network (SSDAN) and ICPSR's Online Learning Center (OLC)

IASSIST
May 26-29, 2009
Tampere, Finland

John Paul DeWitt (SSDAN) and Lynette Hoelter (ICPSR)
The University of Michigan
Outline

• Quantitative Literacy
  – Why is it important?

• Resources from the Social Science Data Analysis Network (SSDAN)
  – DataCounts!
  – CensusScope

• Resources from ICPSR’s Online Learning Center (OLC)
  – Learning Modules and Data

• Coming soon! (ICPSR and SSDAN Partnerships)
  – Assessment Materials
  – Digital Library (TeachingWithData.org)
Quantitative Literacy

• Importance
  – Students are participants in a democratic society
  – Skills include:
    • Questioning the source of evidence in a stated point
    • Identifying gaps in information
    • Evaluating whether an argument is based on data or opinion/inference/pure speculation
    • Using data to draw logical conclusions

• Student Comfort and Aptitude
  – Over 50% of early undergraduate students report substantial “statistics anxiety” (DeCesare, 2008)
  – Using only one or two learning modules has yielded significant increases in students’ comfort

• Solution: Introduce students to “real world” data early and often
Quantitative Literacy

Quantitative literacy, also known as quantitative reasoning (QR), is a “habit of mind” that can be strengthened considerably during a student’s college years. While curricular opportunities for students to enhance their quantitative literacy skills come from across the curriculum and at all levels of the curriculum, not all students will encounter such courses each and every year. Opportunities for students to develop their QR skills are strongly influenced by the degree to which their major employs these skills. As such, this Quantitative Literacy Rubric does not tie the four levels of competency to the four years of college; rather, it is constructed on a scale in which level 4 indicates exemplary skills; 3 indicates strong skills; 2 indicates limited skills; and 1 indicates very weak skills. Details on the scale are provided for the six quantitative literacy criteria below.

Evaluators are encouraged to assign a zero to any performance that doesn’t meet level one performance.

<table>
<thead>
<tr>
<th>Interpretation</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to explain information presented in a mathematical form (e.g., equations, graphs, diagrams)</td>
<td>Skillfully explains information presented in mathematical form (e.g., equations, graphs, diagrams, tables). Consistently provides clear explanations with no errors.</td>
<td>Competently explains information presented in mathematical form (e.g., equations, graphs, diagrams, tables). Rarely makes errors or gives unclear explanations.</td>
<td>Developing the ability to explain information presented in mathematical form (e.g., equations, graphs, diagrams, tables). Sometimes makes errors or gives unclear explanations.</td>
<td>Attempts to explain information presented in a mathematical form (e.g., equations, graphs, diagrams, tables), but has trouble doing so correctly. Frequently makes errors or gives unclear explanations.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Representation</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to convert relevant information into various mathematical forms (e.g., equations, graphs, or diagrams)</td>
<td>Consistently demonstrates fluency in converting relevant information into various mathematical forms (e.g., equations, graphs, or diagrams, tables). Reliably chooses the best form for the problem at hand</td>
<td>Generally able to convert relevant information into various mathematical forms (e.g., equations, graphs, or diagrams, tables) accurately. Rarely makes errors and almost always chooses the best form for the problem at hand.</td>
<td>Developing the ability to convert relevant information into mathematical forms (e.g., equations, graphs, or diagrams, tables). Sometimes makes errors or uses forms that are not the best for the problem at hand.</td>
<td>Able to identify relevant information, but has difficulty converting it into mathematical forms (e.g., equations, graphs, or diagrams, tables). Frequently makes errors or uses forms that are not the best for the problem at hand.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Calculation</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Successfully complete all of calculations for the task at hand with consistency.</td>
<td>Successfully complete most calculations for the task at hand most of the time, though they may not be able to successfully complete several of the tasks.</td>
<td>Ability to complete successfully calculations for the task at hand is limited. Perhaps the student can do a few of these calculations very well, but others are inconsistently completed and still others cannot be completed at all.</td>
<td>Attempts to complete calculations for the task at hand are rarely and inconsistently successful.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Application / Analysis</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to make judgments based on quantitative analysis of data. Consistently draws appropriate conclusions from the data and recognizes the limits of the analysis used</td>
<td>Makes informed judgments based on quantitative analysis of data. Rarely making errors or drawing unwarranted conclusions.</td>
<td>Makes judgments based on quantitative analysis of data. Sometimes makes errors or draws unwarranted conclusions.</td>
<td>Attempts to make judgments based on quantitative analysis of data. Frequently makes errors or draws unwarranted conclusions.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimation / reasonableness checks</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reality checks</td>
<td>Consistently checks calculated answers for reasonableness; makes good assumptions for estimation problems that involve unknown quantities; performs reality checks on numbers reported by others; as appropriate</td>
<td>Often checks calculated answers for reasonableness; makes good assumptions for estimation problems that involve unknown quantities; performs reality checks on numbers reported by others; as appropriate</td>
<td>Sometimes checks calculated answers for reasonableness; confident in making estimates that require assumptions about unknown quantities; performs reality checks on numbers reported by others; as appropriate</td>
<td>Rarely checks answers for reasonableness; confident in making estimates that require assumptions about unknown quantities; performs reality checks on numbers reported by others; as appropriate</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Communication</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressing a solution so that an audience understands what the solution means</td>
<td>Clearly communicates quantitative information for reader or user, shaping it into an argument, solution, or conclusion as appropriate, using a well-chosen, effective format and placing values in context.</td>
<td>Clearly communicates quantitative information for reader or user, although information may not cohere as an argument, solution, or conclusion, may not be in the most effective format or with necessary context.</td>
<td>Communicates quantitative information for reader or user, but does not constitue a clear or coherent point, chosen format is neither the most effective nor in context.</td>
<td>Attempts to communicate quantitative information for reader or user, but is unsuccessful at making an argument, selecting an appropriate format, or placing in context.</td>
</tr>
</tbody>
</table>

Created by a team of faculty from higher education institutions across the United State.
SSDAN: Background

- Started in 1995
- University-based organization that creates demographic media and makes U.S. census data accessible to policymakers, educators, the media, and informed citizens.
  - web sites
  - user guides
  - hands-on classroom computer materials
SSDAN

• DataCounts! (www.ssdan.net/datacounts)
  – Collection of approximately 85 Data Driven Learning Modules (DDLMs)
  – WebCHIP (simple contingency table software)
  – Datasets (repackaged decennial census and American Community Survey)
  – Target is lower undergraduate courses

• CensusScope (www.censusscopescope.org)
  – Maps, charts, and tables
  – Demographic data at local, region, and national levels
  – Key indicators and trends back to 1960 for some variables
SSDAN: DataCounts!

DataCounts! is an interactive online tool designed to help integrate social statistics into the classroom setting. We have several collections of data, including American Community Survey Data from 2004-2006, the 1990 and 2000 censuses, census trend data from 1950-2000, the General Social Survey, and the Current Population Survey. Each collection contains a wide variety of datasets that can be viewed online with WebCHIP. DataCounts! also houses a collection of teaching modules that have been created by teachers across the country to integrate social science data into their classes.

Please click the "How-To" button in the navigation bar for help using this site and software.

..or Data Driven Learning Modules
SSDAN: DataCounts!

Dataset Browser
From here, browse available datasets. This is the best method to find datasets based on the general category of the data. It’s easy! Just follow the four steps below.

1) In the drop-down box to the right, choose a collection by clicking on the downward pointing arrow then clicking on the desired collection, and click "Submit."

2) Next, choose your dataset by clicking again on the downward pointing arrow.

3) If you are using a geography collection, repeat this process to select your geography.

4) Clicking on "Submit" will take you to the dataset page. You may click "Start Over" to begin again. A glossary of dataset names can be found in our glossary.

These are the most popular WebCHIP Dataset Collections:

<table>
<thead>
<tr>
<th>Collection Name</th>
<th>Description of Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>acs2005i</td>
<td>2005 American Community Survey immigration topics for the nation</td>
</tr>
<tr>
<td>acs2006</td>
<td>2006 American Community Survey for the nation</td>
</tr>
<tr>
<td>cen1990</td>
<td>1990 US census data for the nation</td>
</tr>
<tr>
<td>cen2000</td>
<td>2000 US census data for the nation</td>
</tr>
<tr>
<td>cntrend</td>
<td>Data for several censuses, usually from 1950-2000</td>
</tr>
<tr>
<td>cntrend</td>
<td>The Current Population Survey from 2007 for the nation</td>
</tr>
</tbody>
</table>

Menu for choosing a dataset for analysis

Brief List of available dataset collections
SSDAN: DataCounts!

Submit a Module

After you submit this form you will be able to immediately view a page containing your materials and make changes to
their page. That means you can come back later to finish your work, but keep in mind that leaving this page before you
submit ensures the data. To submit your submission and make changes, you will need a GCRC account. Visit the page
to create an account if you do not have one already. Make sure to use the same email address to create your account
and as this submission form.

You retain all rights to your contribution work and are responsible for referencing other people's work and for
obtaining permission to use any copyrighted material within your contribution. By contributing your work to this
website, you agree to license your work for non-commercial distribution of the material, provided that we attribute
the material to you.

Activity Title

This title should be a concise statement of the main purpose of the activity. It needs to summarize the full content of the
activity on its own as it will show up in places like search returns (e.g., Google) where people won't have any
contextual clues. So, it should convey the idea that this is a teaching activity, what the subject matter is and what
the relevant pedagogical focus is. For example: Solar Radiation Sample Scenario Questions

Author

Name and institution of author(s) of the activity and any other appropriate attribution information. If the page is based
on materials originally created elsewhere that should be cited with attribution given to the original authors and links
provided to the original materials.

For example: The page authored by John Smith, Big State University, based on an original activity by Jane Smith,
Stateville College.

Authorship and Attribution

Title:

Submit a Module

Email

Email addresses of the activity author(s) separated by commas. These will not be displayed in the activity page
but are used for internal tracking.

[example@umich.edu]

Summary

This text should make it clear what the activity is. It should provide an overview of the things that students will do and
the intended outcomes. The description should be concise and completely typically no more than 1-2 very brief
paragraphs.

For example:

In this energy lab, students investigate whether yolke from a heat lamp reflected from without a plate and are
different from those collected from a small white piece of paper. They then test for statistical differences in concepts
using statistical analysis software. Students determine the percentage of eggs that lamp; students compare the combined
class statistics for the two groups of eggs for statistically using one-sample analysis. Students need a related scientific
paper and discuss it in a subsequent class session. Students create a full-length report describing their results using standard
scientific paper formatting. A detailed description of this format and the writing process is provided.

Goals

Describe the concepts and content should students learn from this activity (e.g., DNA replication, the cell cycle,
ionic behavior).

Context for Use

This text should help faculty understand the types of teaching situations for which this activity is appropriate. Important
factors of context include educational level, class size, institution type, etc. Is it lab, lecture, or held exercise or a larger
project? How much time is needed for the activity? Is there special equipment that is necessary? Are these skills or concepts
that students should have already mastered before encountering this activity? How is this activity situated in the course? How
easy (or hard) would it be to adapt the activity for use in other settings?

Grade Level:

Middle School

Submit a Module
SSDAN: DataCounts!

<table>
<thead>
<tr>
<th>Title</th>
<th>Author and Institution</th>
<th>Brief Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>The Explosion of Teenage Motherhood: Myth or Reality</em></td>
<td>Elizabeth Jordan, Kalamazoo College</td>
<td>The focus of this module will be to dispel some misconceptions about teenage motherhood and to introduce students to CHART. Data from 1950 to 1990 will be examined by age, race/ethnicity, education, and poverty ...</td>
</tr>
<tr>
<td><em>U.S. Housing Patterns, Living Arrangements, and Life Chances</em></td>
<td>Diane Kayongo-Male, South Dakota State University</td>
<td>In this assignment you will use U.S. census data to get an overview of housing patterns (e.g., housing) and living arrangements as a way of understanding more about one example of life chances ...</td>
</tr>
<tr>
<td><em>Current and Historic Patterns in the Distribution of Income</em></td>
<td>Jeffrey Lashbrook, SUNY - Brockport</td>
<td>We’ve argued that societal stratification is “both a condition and a process” (Kerckhoff, 2000). The former captures what the distribution of valued resources (e.g., money, education) among other ...</td>
</tr>
<tr>
<td><em>An Analysis of Earnings</em></td>
<td>Jeffrey Lashbrook, SUNY - Brockport</td>
<td>While a much larger percentage of American families are located in the top income bracket in 1990 compared to fifty years ago, there were still slightly over 1/4 of American families with income of $25K or less in ...</td>
</tr>
<tr>
<td><em>Social Structure-Personality: What is the relationship between social class and child-rearing values?</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Faceted browsing to refine the search:
- Appropriate Grade Levels
- Subjects (e.g., Family, Sexuality and Gender)
- Learning Time
SSDAN: DataCounts!

Data Driven Learning Modules are clearly laid out

• Easy to read
• Instructors can quickly identify whether a module would be relevant to a specific course

U.S. Housing Patterns, Living Arrangements, and Life Chances
Diane Kropf-Madle, South Dakota State University

Summary

In this assignment you will use U.S. census data to get an overview of housing patterns (ownership and type of housing) and living arrangements as a way of understanding more about the example of what Molotch referred to as life chances (or, basically, the ability to access resources we need to live a good life). As well, access to housing is considered one of the main features of the American Dream. Unfortunately, as the need for lower-income housing units has doubled over the last three decades, the number of lower-income housing units nationwide has been cut in half. Minneapolis and other cities have established annual goals for increasing the stock of affordable housing in response to a severe shortage of affordable housing.

Learning Goals
Technical data analysis and data analysis skills
As you complete this exercise you will learn how to:
• Access U.S. census data from the SSDAN site
• Set up cross-tabs properly (including the correct way to total percentages to 100)
• Write simple generalizations from selected tables
• Understand basic causal analysis including the difference between independent and dependent variables
• Use a control variable to further examine your cross-tabulated data
• Construct a bar graph using the WebCHIP site
• Get information from CensusScope on housing
• Get information from the U.S. Census Bureau website

Sociological reasoning skills
By completing this exercise you will
• Gain knowledge of how age, race, gender, disability status, and household type affect housing patterns and living arrangements
• Gain knowledge of some of the housing patterns and living arrangements of older citizens
• Gain knowledge of the complex meaning of social diversity in terms of gender, race, marital status, etc.; how cohort differences, and disability as it relates to homeownership, type of housing and living arrangement
• Find out what percentage of American families could afford to purchase a modestly priced home
• Find out whether rent as a percentage of household income increased or decreased from 1989 to 1999
• Compare your analysis based on census data with a report of the National Fair Housing Alliance or with any one of a number of articles listed at the end of this assignment
• Learn what is meant by co-housing and be asked to suggest how it might improve residential living experiences in the U.S.

Content for Use
This exercise was developed for use in an introductory sociology course. The exercise was presented as part of a unit to better understand the concept of “life chances” in relation to achieving one of the components of the American Dream: housing.

Description and Teaching Materials

Example: PDF (per year) (Click to download)
Example: DOC (Click here to download)

Teaching Notes and Tips
This activity uses a three customized data set made from the 1990 Census and guides students through data manipulation using WebCHIP software found at DataCounts. To open WebCHIP with the dataset for the activity, please see instructions and links in the exercise documents under teaching materials. For more information on how to use WebCHIP, see the How To section on DataCounts!

Assessment
Visit DataCounts for assessment tool.

References and Resources
Original Archive Module:
http://www.ssdan.net/datacounts/modules#data&look&census_housing/index.shtml
SSDAN: DataCounts!

- WebCHIP

Commands for selecting variables, creating tables, graphing, and recoding

Basic information about the dataset

Running the “marginals” command shows the categories for each variable and frequencies

Education by Immigration status for Age 25+, weighted 2005 ACS, Frey/M-HITCH-not for commercial purposes

<table>
<thead>
<tr>
<th>Race/Lat</th>
<th>Gender</th>
<th>Age</th>
<th>Imm</th>
<th>Educ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latino</td>
<td>Male</td>
<td>25-34</td>
<td>Native</td>
<td>LTHS</td>
</tr>
<tr>
<td>12.8%</td>
<td>42.9%</td>
<td>35-44</td>
<td>83.3%</td>
<td>159,698,736</td>
</tr>
<tr>
<td>22,663,412</td>
<td>90,586,768</td>
<td>45-54</td>
<td>8.1%</td>
<td>15,005,355</td>
</tr>
<tr>
<td>20,225,166</td>
<td>90,443,952</td>
<td>55-64</td>
<td>7.1%</td>
<td>13,366,670</td>
</tr>
<tr>
<td>8,672,665</td>
<td>39,5,450</td>
<td>65+</td>
<td>18.4%</td>
<td>34,768,524</td>
</tr>
<tr>
<td>1,236,485</td>
<td>1,672,860</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>396,275</td>
<td>51,944,628</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7,892,352</td>
<td>51,366,908</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SSDAN: DataCounts!

Students can quickly run simple cross-tabulations to see distributions and test hypotheses.
SSDAN: DataCounts!

Controlling for an additional variable allows for deeper analysis.

(c) 2006 by Zula Data.
SSDAN

• DataCounts!
  – Collection of approximately 85 Data Driven Learning Modules (DDLMs)
  – WebCHIP (simple contingency table software)
  – Datasets (repackaged decennial census and American Community Survey)
  – Target is lower undergraduate courses

• CensusScope
  – Maps, charts, and tables
  – Demographic data at local, region, and national levels
  – Key indicators and trends back to 1960 for some variables
SSDAN: CensusScope

AN AGING POPULATION

Click on a state to view related charts and data.

Percent of a county's residents aged 65 and over.
- 0.0%-8.6%
- 8.6%-12.8%
- 12.8%-16.2%
- 16.2%-20.7%
- 20.7%-34.7%

America's elderly are a growing population, but one that tends to be concentrated in specific geographical areas. Areas where a high percentage of the population is over the age of 65 include not only retirement hotspots like Florida and the Southwest, but also places like the Great Plains, with a high

New ACS data with improved look & feel coming Fall 2009
SSDAN: CensusScope

- Charts, Trends, and Tables
- All available for states, counties, and metropolitan areas
SSDAN-OLC

• SSDAN’s primary focus is to assist in the dissemination of social data into the classroom with sites like DataCounts! and CensusScope

• Until recently, ICPSR primarily targeted researchers, the OLC now provides a welcomed Instructors’ portal to ICPSR resources and is continuing its outreach to educators
ICPSR: Background

- Founded in 1962 at the University of Michigan, is the largest non-governmental social science data archive in the world.
- ICPSR has special expertise in on-line dissemination of data, the development of metadata standards, and training in quantitative methods to facilitate effective data use.
- ICPSR now has nearly 650 diverse members around the world.
- The ICPSR archive currently holds
  - 7,121 studies consisting of
  - 60,979 datasets with
  - 519,934 files
- The collection grows by 1200-1400 files per year.
- Data are available instantaneously.
ICPSR: OLC

- OLC: www.icpsr.umich.edu/OLC
- Over 40 DDLMs (also called DDLGs or Data Driven Learning Guides)
- Format of DDLGs allows instructors to make decisions on how to incorporate into class
  - Based on lesson plan format
- A tool to help develop classroom lectures and exercises that integrate data early into the learning process.
- Intended for use in introductory-level substantive classes.
- Requires no additional software.
Welcome to the OLC

ICPSR’s Online Learning Center (OLC) supports quantitative literacy in the social sciences by providing an effective and reliable means of bringing secondary data into the classroom. Teaching faculty designed, built, and tested OLC tools. OLC tools and resources assist faculty who strive to help students open the door to the world of statistical literacy and substantive inquiry. Use of the Data-Driven Learning Guides builds numeracy and fosters excitement through exploration of concepts as they apply to the real world.

The Online Learning Center (OLC) is a resource for undergraduate instruction funded by the Inter-university Consortium for Political and Social Research (ICPSR).
How to Use the OLC

- Faculty use of charts in class to introduce topic
- Sending students to the website to work through a DDLG in class or as homework
- Using DDLG as part of larger project
Data-Driven Learning Guides

These guides can be used to enhance teaching of core concepts in the social sciences. In order to be useful to the widest audience, the focal topics are drawn from concepts that are included in standard introductory-level social science textbooks. New guides will continually be added and topic areas expanded.

Browse Guides by Subject

Note that guides are listed in the subject area in which they best fit. Subject areas will be narrowed as more guides are added with the idea that some guides may be cross-listed in multiple sections as appropriate.

* Political Science
  - Accountability
  - Campaigns
  - Civic participation and elections
  - Comparative politics
  - Democracy
  - Demographics
  - Elections
  - Identity politics
  - International relations
  - Linked face
  - Partisanship
  - Political awareness
  - Public opinion and attitudes
  - Representation
  - Social capital
  - State Politics
  - Terrorism
  - Vote choice

* Sociology
  - Acculturation
  - Adolescents
  - Aging
  - Attitudes
  - Crime and deviance
  - Demographics
  - Discrimination
  - Education
Data-Driven Learning Guide

Religion among Teens: A Data-Driven Learning Guide

Goal & Concept

Goal

The goal of this exercise is to examine the relationship between demographic and socio-cultural factors and religiosity among American high school students. Crosstabulations, comparison of means, and graphs will be used.

Concept

Religiosity is the term used to refer to the importance of religion in a person’s life. It includes religious identity, behavior, attitudes, perceptions, and practices.

Research has shown that religiosity has a positive influence on adolescents’ lives and is associated with healthier diets, exercise, sleep habits, and self-esteem, as well as lower rates of alcohol and drug use, early sexual behavior, delinquency, depression, and suicide.

Less is known about how religiosity develops in children and adolescents; however, though research suggests that socio-cultural factors such as racial identity, gender, parental marital situation and geographical location of residence influence the extent of teenagers’ religiosity, religious attendance of religious services, parents’ level of education, household income, and parental presence in the household.

Examples of possible research questions about adolescent religiosity:

- What are the mechanisms by which religious beliefs develop during childhood and how are they related to teenagers’ behaviors and attitudes?
- How are family structure, gender, race, and ethnicity related to adolescents’ religiosity?
- How does the contextual environment (family, peers, school, and community) influence religious beliefs and behaviors in youth, and through these religious beliefs and behaviors influence health risk behaviors in youth?
Religion among Teens: A Data-Driven Learning Guide

Dataset


Data for this exercise come from the Monitoring the Future (12th Grade Survey), 2006 (MTF). The 2006 MTF is part of an ongoing series of cross-sectional data collection designed to explore changes in important values, behaviors, and lifestyle orientations of contemporary American youth. Each year, large, distinct, nationally representative samples of 8th, 10th, and 12th-grade students in the United States are asked to respond to drug use and demographic questions, as well as to additional questions on a variety of subjects, including attitudes toward religion, parental influences, changing roles of women, educational aspirations, self-esteem, exposure to sex and drug education, and violence and crime both in and out of school. Data used for this exercise are restricted to the 2006 12th Grade Survey, Core Data. Funding for MTF is provided by the United States Department of Health and Human Services, National Institute on Drug Abuse. The principal investigators are Lloyd D. Johnston, Jerald G. Bachman, Patrick M. O’Malley, and John E. Schulenberg of the University of Michigan Institute for Social Research, Survey Research Center.

The MTF 12th Grade Survey is a school-based sample designed to represent high school seniors in the contiguous United States. In this exercise, data from the Core dataset (DS1) are used.

This exercise will use the following variables:

- Respondent’s sex (V150)
- Respondent’s race (V151)
- Respondent’s father living in household (V155)
- Respondent’s mother living in household (V156)
- Father’s education (V158)
- Mother’s education (V164)
- Respondent’s attendance at religious services (V169)
- Religion importance to respondent’s life (V170)
Links go directly to the processed SDA functions so there is nothing to break (no additional manipulation needed—no functions to run)
### Frequency Distribution

Cells contain:
- Column percent
- Weighted N

<table>
<thead>
<tr>
<th></th>
<th>MALE: (1)</th>
<th>FEMALE: (2)</th>
<th>ROW TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1: not important</td>
<td>21.1</td>
<td>14.3</td>
<td>17.5</td>
</tr>
<tr>
<td></td>
<td>1,086</td>
<td>820</td>
<td>1,906</td>
</tr>
<tr>
<td>2: little important</td>
<td>25.7</td>
<td>24.2</td>
<td>24.9</td>
</tr>
<tr>
<td></td>
<td>1,320</td>
<td>1,386</td>
<td>2,707</td>
</tr>
<tr>
<td>3: pretty important</td>
<td>27.2</td>
<td>28.4</td>
<td>27.9</td>
</tr>
<tr>
<td></td>
<td>1,401</td>
<td>1,628</td>
<td>3,029</td>
</tr>
<tr>
<td>4: very important</td>
<td>26.0</td>
<td>33.0</td>
<td>29.7</td>
</tr>
<tr>
<td></td>
<td>1,336</td>
<td>1,891</td>
<td>3,227</td>
</tr>
<tr>
<td>COl TOTAL</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>5,143</td>
<td>5,725</td>
<td>10,868</td>
</tr>
</tbody>
</table>

---

**BY 062C04(R):R’S RACE B/W/H**

- **BLACK:**
  - never: 20%
  - rarely: 30%
  - 1-2 per month: 40%
  - once per week or more: 10%

- **WHITE:**
  - never: 25%
  - rarely: 35%
  - 1-2 per month: 40%
  - once per week or more: 10%

- **HISPANIC:**
  - never: 22%
  - rarely: 32%
  - 1-2 per month: 40%
  - once per week or more: 16%
Important pitfalls and limitations that the instructor should address with students

Summary of results

- Things to think about in interpreting the results:
  - When a survey or poll is administered, researchers go to great lengths to ensure that the sample is representative of a target population. If this is done well, they have a higher degree of confidence in saying that the results of the survey are applicable to the population from which the sample was drawn. Since this survey was conducted among American high school 12th graders, the hope is that the results can be generalized to all American 12th graders and not just used as a description of the sample.

- One potential problem when using inference to generalize the results to all American 12th graders is the amount of missing data in each relationship. If there seems to be a large amount of missing data, it is important to try to understand why this might be the case. Sometimes questions are sensitive and respondents choose not to answer. Other times, questions might not be asked to all respondents. One particular quirk in this dataset is that the researchers note that they have "nullified" responses from the western region for the questions about religion, thus increasing the amount of missing data. This does not necessarily mean that the results cannot be generalized, but it does mean that we have some reason to be cautious about making generalizations to all American 12th graders.

- Reading the results: The numbers in each cell of the cross-tabulation tables show the percentage of the respondents who fall into the overlapping categories, followed by the actual number of people that that represents in this sample. The coloring in the tables demonstrates how the observed values in each cell compare to the expected value if there was no association between the two variables. More specifically, if there is no association between variables, then the percentages across each cell in the row would all be equal. The accompanying bar charts display the patterns visually as well.

- In the comparison of means tables, the numbers in the cells represent essentially the same thing - the top number is the mean of the dependent variable for each category of the
Includes not only bibliography, but access to related electronic articles via the local institutions library.
Student Benefits:

• Critical thinking skills
• Increases students’ comfort with quantitative reasoning
• Many schools have focus on quantitative literacy and related skills
  – ASA charge for exposure “early and often”
• Engages students with the discipline more fully
  – Better picture of how social scientists work
  – Prevents some of the feelings of “disconnect” between substantive and technical courses
• Piques student interest
• Opens the door to the world of data
Looking Ahead (ICPSR-SSDAN Partnerships)

- Assessment Tools and Results

Rubric and Examples for Quantitative Literacy in the Social Sciences
Draft: 5/13/2009

Quantitative Literacy Learning Outcomes
1. **Calculation**: Ability to perform mathematical operations.
2. **Interpretation**: Ability to explain information presented in a mathematical form (e.g., tables, equations, graphs, or diagrams).
3. **Representation**: Ability to convert relevant information from one mathematical form to another (e.g., tables, equations, graphs, or diagrams).
4. **Analysis**: Ability to make judgments based on quantitative analysis.
5. **Method Selection**: Ability to choose the mathematical operations required to answer a research question.
6. **Estimation/Reasonableness Checks**: Ability to recognize the limits of a method and to form reasonable predictions of unknown quantities.
7. **Communication**: Ability to use quantitative information to support a conclusion.
8. **Identify/Generate Data**: Ability to identify or generate appropriate information to answer a question.
9. **Research Design**: Understand the links between theory and data.
10. **Confidence**: Level of comfort in performing and interpreting a method of quantitative analysis.
11. **Content Learning Outcomes**: Varies by module/course.

1. Calculation:
   - What is the median in the above set of numbers?
   - What is the mean in the above set of numbers?
   - What is the mode in the above set of numbers?
   - a. 0  
   - b. 1  
   - c. 2  
   - d. 4

   1A:
   - 0, 0, 1, 4, 4

1. B.
   - A student's grade depends on her score on four equally weighted exams. Her average on the first three exams is 92. What must she score on the fourth exam in order to guarantee a final average of at least 90?

1. C.
   - The maternal mortality rate is measured as the number of deaths due to childbirth divided by the number of births. Suppose that the maternal mortality rate remains constant but the birth rate goes down. What will happen to the proportion of female deaths caused by complications of childbirth?
Looking Ahead

• TeachingWithData.org (October 2009)
  – Social Science Pathway to the National Science Digital Library
  – Virtual repository of resources to support the teaching of quantitative social sciences
    • Data-Driven Learning Modules
    • Data Sources
    • Pedagogical Resources
    • Analysis and Visualization Tools
    • Other Resources useful to instructors
  – Collaboration tools
  – Web 2.0 technologies
Acknowledgements

• National Science Foundation
• Inter-university Consortium for Political and Social Research
• Population Studies Center
• Institute for Social Research
• University of Michigan