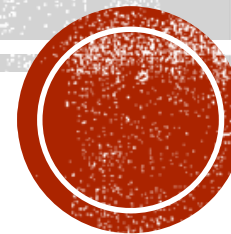


AN INTRO TO NCES DATA

Matthew B. Fuller, Ph.D.

Assistant Professor, Higher Education Leadership, Dept. of
Educational Leadership



NATIONAL CENTER FOR EDUCATION STATISTICS

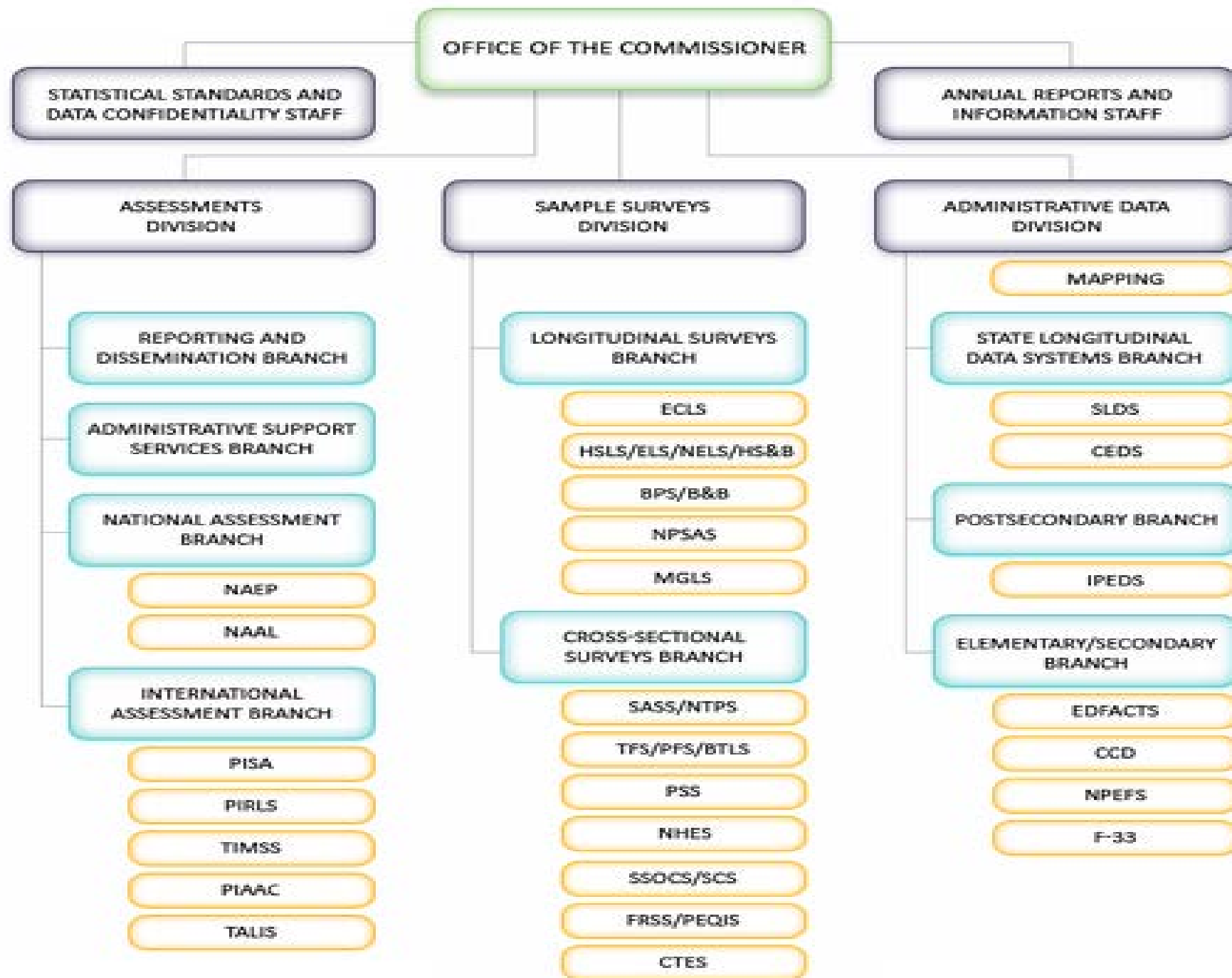
- Established in 1867
- *“That there shall be established at the City of Washington, a department of education, for the purpose of collecting such statistics and facts as shall show the condition and progress of education in the several States and Territories, and of diffusing such information respecting the organization and management of schools and school systems, and methods of teaching, as shall aid the people of the United States in the establishment and maintenance of efficient school systems, and otherwise promote the cause of education throughout the country.”*



NCES MISSION

- Full Mission and goals: <http://www.ed.gov/open/plan/nces>
- The mission of NCES is to collect, analyze, report, and disseminate education information and statistics in a manner that:
 - meets the highest methodological standards;
 - is timely, relevant, and useful to practitioners, researchers, policy makers, and the public;
 - is objective, secular, neutral, and non-ideological; and
 - is free of partisan political influence and racial, cultural, gender, or regional bias.









ALPHABET SOUP

- See handout.
- Don't let the acronyms worry you.
- Start to zero in on your “go to” survey or assessment for you research agenda.
- Some of the naming and groupings can be misleading.
 - An Example: IPEDS is the primary element under the Postsecondary unit.
 - There are, however, at least 9 projects that clearly relate to postsecondary education.



THE BIG FOUR (OR MORE)

- National Assessment of Educational Progress (NAEP) 
- Education Longitudinal Studies 2002 and 2012 (ELS: 2002; ELS:2012) 
- Integrated Postsecondary Education Data System (IPEDS) 
- Postsecondary Education Transcript Collections (PETS) 
- Helpful website: <http://nces.ed.gov/surveys/>



PROVISIONAL DATA, MANIPULATORS, AND PUBLIC ACCESS DATA SETS

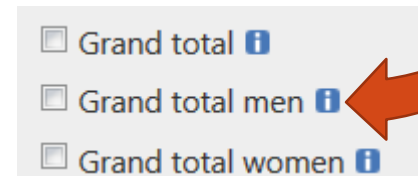
- Most surveys include public data, available to everyone because it contains non-identifiable, non-risky, or aggregate data.
 - They also include restricted use data, which contain identifiable or sensitive data, sometime even down to very detailed, personal data
 - There are restrictions on the use of these data. More on this later.
- Most surveys operate with an embargo or “preparation” period. Data from two years prior are often finalized after institutional or school representatives have had time to update data.
 - Provisional data are more recent, but not updated. Finalized data are updated but less recent (often by a year or two).
 - Check if the project is working on the same “calendar” you intend to work on.



MANIPULATORS AND THE DATA ANALYSIS SYSTEM



- Can you accomplish what you need to with basic, descriptive statistics?
 - Consider a manipulator: Table or Chart builder; Frequency table, DAT, Powerstats.
 - IPEDS offers the Data Analysis System which is a log-in system that allows you to conduct basic statistical analyses such as t-tests and, for limited variables, ANOVA.
 - As this is an online system, recognize its limitations.
- Most questions require us to download data.
- There are simple point and click ways to select data.
- Take a moment to familiarize yourself with the methodological report and definitions.
- Realize it takes years to become familiar with definitions.



A FEW STATISTICAL MATTERS

- Data for very small, easily identifiable subgroup can be swapped.
 - Some scholars question the accuracy or usefulness of this decision.
- Some projects offer weighted variable to provide an opportunity to depict the entire nation's student bodies.
- If weighting is used, some scholars caution against the way NCES weights its variables, suggesting it disfavors larger groups' perspective to normalize the entire groups attitudes toward smaller groups' perspectives.
 - Is your study significantly enhanced by using the weighted variables?
- Data with cells smaller than 3 should not be reported.



EXAMPLES

- A K-12 educational researcher is interested what effect having friends or teachers of similar or different racial backgrounds might influence standardized test performance.
- What data source can be helpful in answering this question?





- Overview
- Policy and Research Issues
- Survey Design
- Questionnaires
- Available Data
- User Manuals
- Reports
- Online Bibliography Search Tool
- Table Library
- Questions?
- Elise Christopher
- Join Newsflash

Questionnaires

Base Year (2002):

- [Student Questionnaire Base Year](#) (215 KB)
- [Student \(Abbreviated\) Questionnaire](#) (81 KB)
- [Parent Questionnaire Base Year](#) (195 KB)
- [Parent \(Abbreviated\) Questionnaire Base Year](#) (229 KB)
- [School Administrator Questionnaire](#) (156 KB)
- [Teacher Questionnaire](#) (245 KB)
- [Facilities Checklist](#) (67 KB)
- [Library Media Center Questionnaire](#) (110 KB)

First Follow-up (2004):

- [Student Questionnaire](#) (154 KB)
- [Student \(Abbreviated\) Questionnaire](#) (119 KB)
- [New Student Supplement](#) (80 KB)
- [Transfer Student Questionnaire](#) (138 KB)
- [Early Graduate Questionnaire](#) (135 KB)
- [Dropout Questionnaire](#) (147 KB)
- [Homeschool Student Questionnaire](#) (171 KB)
- [School Administrator Questionnaire](#) (124 KB)
- [School Administrator \(Abbreviated\) Questionnaire](#) (235 KB)

Second Follow-up (2006):

- [Flow Chart of Instrument Logic](#) (76 KB)
- [Question Stems of Instrument](#) (70 KB)

Third Follow-up (2012):

- [Flowchart of Instrument Logic](#) (202 KB)
- [Sample member questionnaire](#) (265 KB)

For more information on viewing PDF files click here.



Variables on the race of each student's three best friends and three favorite teachers

Variables on standardized test performance from the year of survey administration and from 4 years later

Interested in Methods and Results? Newgent, R., Lee, S., Daniel, A., (2008). Interracial best friendships: Relationship with 10th graders' academic achievement level. *Professional School Counseling, 11*(2), 98-104.



ANOTHER EXAMPLE

- A higher education researcher is interested in the amount of debt students of various backgrounds take on and their loan default rates.
- What data sources can be helpful in this case?
- The National Postsecondary Aid Study provides data on loan encumbrance, loan default, and demographic variables, including a variable on the setting of their high school.
- There is a matter of degree of accuracy and type for every variable.



RESTRICTED USE DATA

- Suppose this same researcher is interested in studying the urbanicity of a college student's high school to determine if rural students take on more or less loan debt than suburban and urban students.
- There is a variable in the [National Postsecondary Student Aid Study - NPSAS](#) called urbanicity. In the public data it simply offers a categorical variable (rural, suburban, and urban). You run descriptive statistics and see that a disproportionate number of schools are coded as urban and suburban. The researcher decides to use a more refined variable: Zip Code, coupled with data from the U.S. Census Bureau that is the definitive means of determining urbanicity.
- To accomplish this a restricted data license holder is needed.



ARE RESTRICTED USE DATA AVAILABLE FOR YOUR STUDY?

- Check:
- <http://nces.ed.gov/pubsearch/licenses.asp>
- Then refer to the [Restricted-Use Data Procedures Manual](#).
- The Institute of Education Sciences (IES) collects survey and research data containing individually identifiable information, which is confidential and protected by federal law.
- But how do you know what data are publically available and what data are available to restricted data use licensees?
 - Two Demos
 - <http://nces.ed.gov/surveys/sass/>
 - <http://nces.ed.gov/surveys/npsas/>



RESTRICTED USE LICENSE HOLDER CRITERIA

- You must be affiliated with an organization, usually a university or research institute.
 - Often this means you will need to be terminally degreed.
- Your organization must submit, through you:
 - An online [Formal Request](#) through the NCES electronic application system
 - You identify the Principal Project Officer (PPO), Senior Official (SO), and Systems Security Officer (SSO), each of whom are responsible for various liabilities during the project.
 - a signed [License document](#),
 - The formal document that will be signed between you and the IES director once all other documents are completed.
 - executed [Affidavits of Nondisclosure](#),
 - You agree to not provide identifiable data, to allow NCES staff to review publications, and to affirm you have read the penalties
 - And a signed [Security Plan Form](#).
 - You develop a plan to maintain data under at least two locks, on a computer not connected to the internet, and with only certain staff who can access the data room.



THE DATA ROOM AT SHSU

- Data and computer are stored separately.
- Behind 4 locks. (Outer Door, Department Door, Office Door, Security Cabinet Door)
- Notified UPD of the security concern
- Only 2 people with access to the room.
 - No custodial staff have access to the room.
 - No student worker staff or faculty can access to room without the license holder's presence
- Computer not connected to the internet; yet updated regularly.
- Appropriate, required signage
- Password security
- Hardware security
- Return data when no longer in use
- Not report any data with less than 3 cases



DISCLOSURE REVIEW AND SITE INSPECTION

- Licensees are required to provide a draft copy of each information product that is based on or uses restricted-use data to the IES Data Security Office for a disclosure review.
- Not a statistical review but often can be if egregious errors are noted.
- Administered by the Statistical Standards Office of NCES
- An inspector will also be allowed to inspect, unannounced, the security procedures for the data analysis room.



PENALTIES

- Alleged violations of the [Privacy Act of 1974](#) or IES-specific laws are subject to prosecution by the United States Attorney after first making reasonable efforts to achieve compliance.
- Any violation of this License may also be a violation of federal criminal law under the Privacy Act of 1974, 5 U.S.C. 552a, and may result in a misdemeanor and a penalty of up to \$5,000.
- Anyone violating the confidentiality provisions of Section 183 of the Education Sciences Reform Act of 2002 (P.L. 107-279), or making an unauthorized disclosure, when using the data shall be found guilty of a ***class E felony*** and can be ***imprisoned up to five years***, and/or ***fined up to \$250,000***.
- Penalties, fines and imprisonment, may be enforced for each occurrence of a specific violation.



WHY SEEK A RESTRICTED USE DATA LICENSE

- You can address some really interesting research questions
- Data sets are powerful and offer large sample sizes
- Data are often collected for you and they are data to which others may not have access
- Their use is noteworthy



TIPS AND ADVICE

- Read for depth, not breadth. The NCES services are so comprehensive. No one can be an expert in all of these projects. Let your interests and research agenda drive you toward a project and, over time, get to know that project.
- Review the questionnaires and the methods manuals for your prime survey.
- Find a key leader on your campus that has experience with datasets.
- Engage them in conversations; you might not need to get the license if someone else on your campus has it already.
- Always ask “Do public data suffice for the best study I can produce?”
 - This question takes into account your time, capacity, needs, and research directions.
- Find a partner or co-author.
- Rely on Restricted-Use License Holders.
- Contact the NCES staff lead for each project. They are invaluable guides.



QUESTIONS? DIALOGUE.

