

Analyzing and Interpreting Surveys

ENGL 7360: Justice-Oriented Writing Assessment

Professor Mya Poe

Fall 2022

Taught by: Juniper Johnson and Ana Abraham



Northeastern University
NULab for Texts, Maps, and Networks

*Feel free to ask questions at any point
during the presentation!*

Workshop Agenda

- Review: Social science methods + Survey design
- Introduction to NCES & DataLab
- Navigating DataLab
 - Statistical Analysis with PowerStats
 - Table and chart analysis with Tables Library
- Activity: DataLab exploration and discussion

Slides and presentation materials available at:

<https://bit.ly/fa22-poe-surveyanalysis>



Review: Data Ethics + Survey Design



Northeastern University
NULab for Texts, Maps, and Networks

*Feel free to ask questions at any point
during the presentation!*

Big Data & Algorithmic Bias

Big Data: vast amounts of data collected by companies, governments, and other groups about users, often analyzed and used for advertising, marketing, surveillance, etc. with the goal of *predicting individual user behavior based on patterns from the user*, turning our information into a *product*.

Algorithmic bias: perpetuation of biases in algorithmic processes, either through use of data that misrepresents populations or technology design.

“What is counted—like being a man or a woman—often becomes the basis for policymaking and resource allocation. By contrast, what is not counted—like being nonbinary—becomes invisible...”

Catherine D’Ignazio & Lauren Klein, *Data Feminism*, 2020



Critical Questions for Researchers:

- What **information** is being collected and from where? To whom does this data **belong**?
- How is it being **collected**? Do **participants** know that it is collected, how it will be collected, and how will it be used?
- **How** will the data be analyzed? What **biases** and **ideologies** may be implicit in this analysis?
- Who will this research impact? Who will it **benefit**? Who will it potentially **harm**?



Fundamentals of Social Science Research

Variables: sets of attributes or characteristics that can take on different values depending on what is being measured.

- **Ordinal Variables:** variables that can be measured and ordered, either quantitatively or qualitatively.
- **Nominal Variables:** variables without order, often where there is no expected or measured value collected.

Common Inquiry Errors: these are examples of common errors when creating or measuring variables.

- **Inaccurate Observations**
- **Oversgeneralization**
- **Selective Observation**
- **Illogical Reasoning**



Survey Design: Types of Questions

Multiple choice: questions that only have specific answers and the user can only click one. For example, “What is your favorite physical exercise?”

Likert-scale: a type of response scale in which respondents can specify their level of agreement, importance or satisfaction typically in 5 points, (1) strongly agree to (5) strongly disagree.

Checkboxes: questions that only have specific answers and the user can click multiple. For example, “select all the physical exercises you did last week”.

Linear Scale: questions that invite users to choose from a lower to higher number that match their experience. Typically on a scale from 1 to 10; allows a more granular measure of affect and participants are able to express their degrees of response.

Paragraph/Short Answer: open-ended questions that usually invite longer prose. For example, “Describe your favorite gym exercise and how it makes you feel”.



Statistical Analysis

Descriptive Statistics: summary statistics used to quantitatively describe or summarize features of a collection of information or dataset, including frequency, central tendency (mean), variation, etc.

Association: a form of inferential statistics used to make inferences of a sample set of data for a larger populations; includes measures of association like Chi square to show how strongly two variables are related.

Correlation: a form of inferential statistics that is used to show an empirical relationship between two variables like if a change in one variable is associated with change in another. **Remember, correlation does not mean causation!**



Introduction: National Center for Education Statistics



Northeastern University
NULab for Texts, Maps, and Networks

*Feel free to ask questions at any point
during the presentation!*

National Center for Education Statistics (NCES)

IES :: NCES National Center for
Education Statistics

National Center for Education Statistics (NCES) is a federal entity that collects and analyzes data on education in the US and internationally, from early childhood up through post-secondary educational levels.

Examples of some studies and assessments include:

- National Assessment of Educational Progress (NAEP)
- Education Longitudinal Study (ELS)
- National Household Education Survey (NHES)
- Baccalaureate and Beyond (B&B)



Closer Look: NCES Fast Facts

Fast Facts: a series of short, consumable fact sheets composed of “concise information on a range of educational issues” that is drawn from published studies from the NCES.

- [Degree conferred by race/ethnicity and sex](#)
- [Undergraduate graduation rates](#)
- [Student Debt](#)
- [Students with Disabilities](#)
- [Race/ethnicity of college faculty](#)

Questions for Consideration:

- What do you notice about how these fast facts are formatted?
- What sort of patterns do you see in the questions that they are asking?
- Is there anything missing?
- Are there other things you wish you could know based on this information?



NCES DataLab

DataLab is a web-based platform providing access, navigation, analysis, and presentation of data collected by the NCES, including the **PowerStats** and **Tables Library Tools**

Tables Library

- Repository for data tables from all NCES publications
- Allows users to search by keyword, filter by source and topic, and download for Excel or CSV

PowerStats

- Generate statistical analyses and data visualizations
- Create and run percentage distributions, averages, medians, linear and logistic regressions, correlation matrices, etc.



Create a DataLab Account



Login

☐ I accept the usage agreement terms

LOGIN

[Forgot password ?](#)

[Create an account](#)

NCES DATA USAGE AGREEMENT

Under law, public use data collected and distributed by the National Center for Education Statistics (NCES) may be used only for statistical purposes. Any effort to determine the identity of any reported case by public-use data users is prohibited by law. Violations are subject to Class E felony charges of a fine up to \$250,000 and/or a prison term up to 5 years.

NCES does all it can to assure that the identity of data subjects cannot be disclosed. All direct identifiers, as well as any characteristics that might lead to identification, are omitted or modified in the dataset to protect the true characteristics of individual cases. Any intentional identification or disclosure of a person or institution violates the assurances of confidentiality given to the providers of the information. Therefore, users shall:

- Use the data in any dataset for statistical purposes only.
- Make no use of the identity of any person or institution discovered inadvertently, and advise NCES of any such discovery.
- Not link any dataset with individually identifiable data from other NCES or non-NCES datasets.

To proceed you must signify your agreement to comply with the above-stated statutorily based requirements.

Note: NCES data policies for survey results.



Northeastern University
NULab for Texts, Maps, and Networks

Feel free to ask questions at any point during the presentation!

Your Turn!

Create your own DataLab account by going to:

<https://nces.ed.gov/datalab/membership/login>

- 1) Enter in your preferred email address
- 2) Check account for DataLab account verification
- 3) Create your password.

After creating an account, go to DataLab homepage.



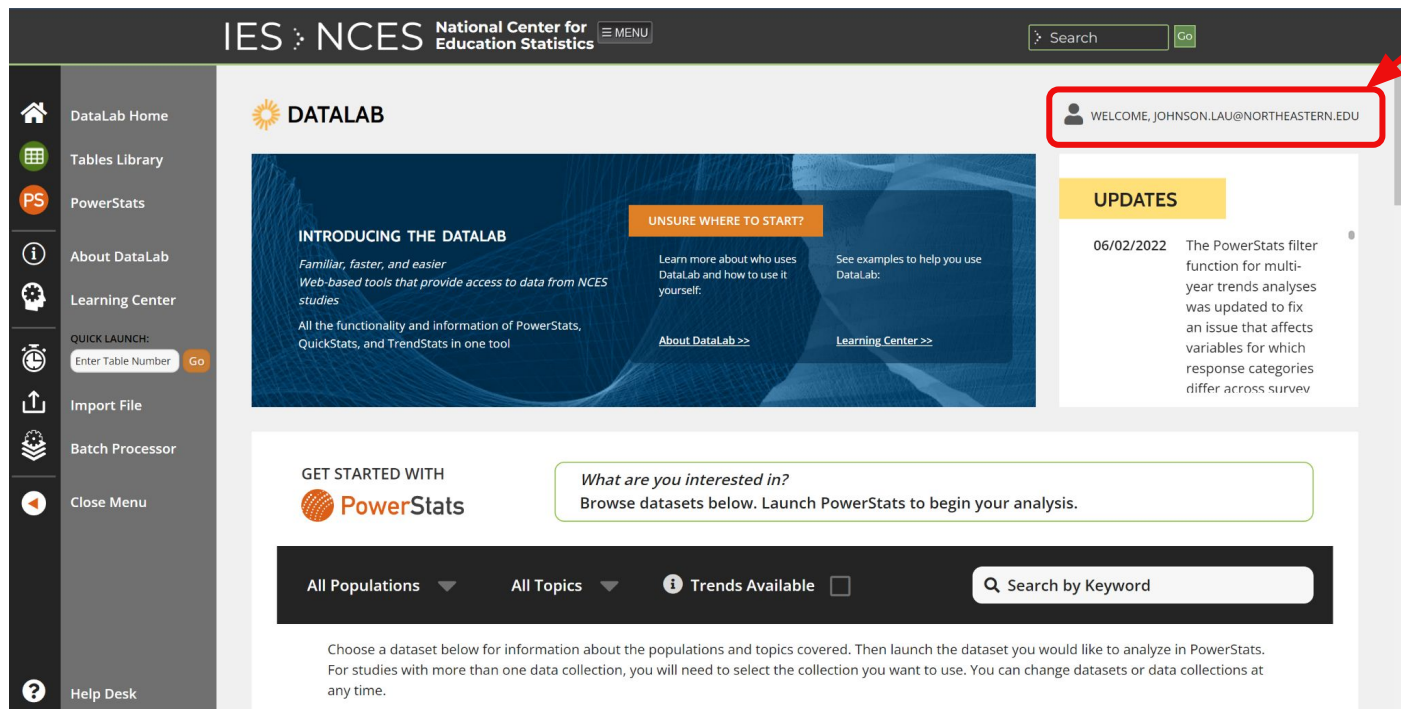
Navigating NCES DataLab



Northeastern University
NULab for Texts, Maps, and Networks

*Feel free to ask questions at any point
during the presentation!*

DataLab Interface



The screenshot shows the DataLab interface with the following elements:

- Header:** IES | NCES National Center for Education Statistics. Search bar with a 'Go' button.
- Left Sidebar:** DataLab Home, Tables Library, PowerStats, About DataLab, Learning Center, QUICK LAUNCH: Enter Table Number (Go), Import File, Batch Processor, Close Menu, Help Desk.
- Main Content Area:**
 - DATALAB:** Introducing the DATALAB (Familiar, faster, and easier Web-based tools that provide access to data from NCES studies). All the functionality and information of PowerStats, QuickStats, and TrendStats in one tool.
 - UNSURE WHERE TO START?:** Learn more about who uses DataLab and how to use it yourself. See examples to help you use DataLab.
 - UPDATES:** 06/02/2022 The PowerStats filter function for multi-year trends analyses was updated to fix an issue that affects variables for which response categories differ across survey.
- User Profile:** WELCOME, JOHNSON.LAU@NORTHEASTERN.EDU (highlighted with a red box and an arrow).
- GET STARTED WITH PowerStats:** What are you interested in? Browse datasets below. Launch PowerStats to begin your analysis.
- Filters:** All Populations, All Topics, Trends Available (checkbox), Search by Keyword.

User Profile:

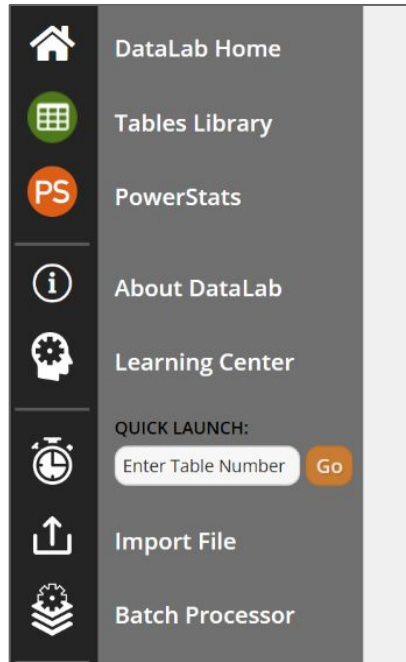
You can change certain settings for your user profile, including preferences for data (confidence and statistical significance) and access any tables that you have saved in the interface.



Northeastern University
NULab for Texts, Maps, and Networks

Feel free to ask questions at any point during the presentation!


DataLab Menu




- **Home:** choose different studies or surveys to explore.
- **Tables Library:** explore tables from NCES publications.
- **PowerStats:** interface to explore data through different statistical analyses.
- **About:** access for FAQs, resources, and tutorials.
- **Learning Center:** tutorials for navigating DataLab interface and examples of analyses.
- **Import:** upload XML or JSON files for tables to edit in DataLab.
- **Batch Processor:** process multiple queries from specific files/surveys.



DataLab Learning Center

 **DATALAB** Learning Center

LOG IN / CREATE ACCOUNT



Welcome to the DataLab Learning Center

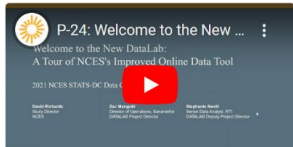
The DataLab Learning Center is a hub for learning materials designed to help analysts understand and use the DataLab tools.

The Analysis Examples and Tutorials section includes examples of the types of analyses that are available in PowerStats. The examples provide step-by-step instructions for running each type of analysis.

The Glossary section includes definition of statistical terms used in the DataLab tools.

Additional learning materials will be added to the Learning Center over time, so check back periodically to view the latest resources. If you have any questions or feedback, please contact the NCES Help Desk at nces.info@rti.org.

VIDEO SPOTLIGHT



Welcome to the NCES DataLab

NCES Study Director David Richards provides an overview and tutorial on the DataLab site.

ANALYSIS EXAMPLES AND TUTORIALS

Users can run the analyses listed below in PowerStats. Clicking on the links will take you to examples of and instructions for each type of analysis.

GLOSSARY

- + Average
- + b
- + Balanced repeated replication (BRR)
- + Base weight
- + Bias
- + Categorical variable
- + Coarsened
- + Confidence interval
- + Continuous variable
- + Correlation matrix
- + Degrees of freedom
- + Dependent variable
- + Design effects
- + Independent variable
- + Jackknife
- + Likelihood ratio
- + Linear regression



DataLab PowerStats: Populations

The screenshot shows the DataLab PowerStats web application. The top navigation bar includes the IES | NCES logo, the text 'National Center for Education Statistics', a 'MENU' button, and a search bar. The left sidebar contains links to DataLab Home, Tables Library, PowerStats, About DataLab, Learning Center, and a 'QUICK LAUNCH' section with a 'Go' button. The main content area is titled 'GET STARTED WITH PowerStats' and includes a box asking 'What are you interested in?' with the instruction 'Browse datasets below. Launch PowerStats to begin your analysis.' Below this is a 'Search by Keyword' bar. A dropdown menu is open under the 'All Populations' tab, listing various population categories: All Populations, Adult Education, Adults, P-12, Children/Students, Schools, Postsecondary, Faculty, Students, Education Longitudinal Study (ELS), High School and Beyond (HSB), High School Longitudinal Study (HSLS), National Education Longitudinal Study of 1988 (NELS:88), and National Household Education Survey: Adult Education Survey (AE-NHES). Each category has a 'Launch' button. A red arrow points from the 'All Populations' dropdown to the 'Populations' text in the adjacent callout box.

IES | NCES National Center for Education Statistics

GET STARTED WITH PowerStats

What are you interested in?
Browse datasets below. Launch PowerStats to begin your analysis.

All Populations All Topics Trends Available Search by Keyword

All Populations
Adult Education
Adults
P-12
Children/Students
Schools
Postsecondary
Faculty
Students
+ Education Longitudinal Study (ELS)
+ High School and Beyond (HSB)
+ High School Longitudinal Study (HSLS)
+ National Education Longitudinal Study of 1988 (NELS:88)
+ National Household Education Survey: Adult Education Survey (AE-NHES)

Launch Launch Launch Launch Launch Launch Launch Launch

Populations:
You can choose which assessment, study, or survey data to investigate based on different populations using this drop down menu.



DataLab PowerStats: Topics

The screenshot shows the DataLab PowerStats interface. At the top, there's a header with 'IES NCES National Center for Education Statistics' and a search bar. Below the header, a sidebar on the left contains navigation links: DataLab Home, Tables Library, PowerStats, About DataLab, Learning Center, QUICK LAUNCH (with a table number input), Import File, Batch Processor, and Close Menu. The main content area is titled 'GET STARTED WITH PowerStats' and includes a prompt: 'What are you interested in? Browse datasets below. Launch PowerStats to begin your analysis.' Below this, there's a section for 'All Populations' and a dropdown menu for 'All Topics'. The 'All Topics' dropdown is open, showing a list of topics: Attendance and Enrollment, Enrollment Intensity Patterns, Early Childhood, Early Childhood Education History, Academic Experiences, Academic Performance, Admissions, Assessments, Field of Study, Outcomes, Persistence and Attainment, and Programs and Courses. To the right of the dropdown, there's a search bar labeled 'Search by Keyword'. Below the dropdown, there's a list of datasets with 'Launch' buttons. A red arrow points from the 'All Topics' dropdown to the 'Major Topics' text box on the right.

IES NCES National Center for Education Statistics

GET STARTED WITH PowerStats

What are you interested in?
Browse datasets below. Launch PowerStats to begin your analysis.

All Populations ▼ All Topics Trends Available ☐ Search by Keyword

Choose a dataset below for studies with more than any time.

- ★ + Baccalaureate and
- ★ + Beginning Postsec
- ★ + Early Childhood Lo
- ★ + Education Longitud
- ★ + High School and Be
- ★ + High School Longit
- ★ + National Education Longitudinal Study of 1988 (NELS:88)
- ★ + National Household Education Survey: Adult Education Survey (AE-NHES)

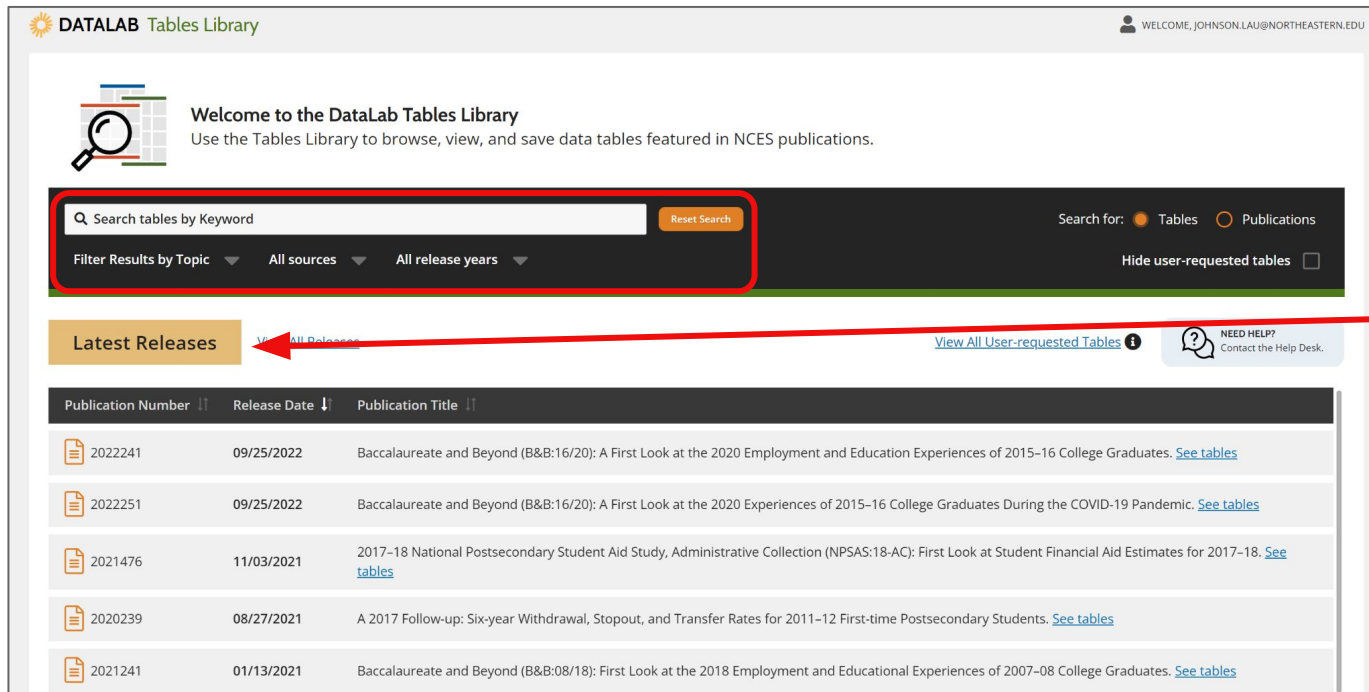
Launch Launch Launch Launch Launch Launch Launch Launch Launch

Major Topics:

You can also choose which assessment, study, or survey data to investigate based on different study topics using this drop down menu.



DataLab Tables Library



DATALAB Tables Library

WELCOME, JOHNSON.LAU@NORTHEASTERN.EDU

Welcome to the DataLab Tables Library
Use the Tables Library to browse, view, and save data tables featured in NCES publications.

Search tables by Keyword [Reset Search](#)

Search for: ☒ Tables ☐ Publications

Filter Results by Topic

[Latest Releases](#) [View All Releases](#) [View All User-requested Tables](#) [NEED HELP? Contact the Help Desk.](#)

Publication Number	Release Date	Publication Title
2022241	09/25/2022	Baccalaureate and Beyond (B&B:16/20): A First Look at the 2020 Employment and Education Experiences of 2015-16 College Graduates. See tables
2022251	09/25/2022	Baccalaureate and Beyond (B&B:16/20): A First Look at the 2020 Experiences of 2015-16 College Graduates During the COVID-19 Pandemic. See tables
2021476	11/03/2021	2017-18 National Postsecondary Student Aid Study, Administrative Collection (NPSAS:18-AC): First Look at Student Financial Aid Estimates for 2017-18. See tables
2020239	08/27/2021	A 2017 Follow-up: Six-year Withdrawal, Stopout, and Transfer Rates for 2011-12 First-time Postsecondary Students. See tables
2021241	01/13/2021	Baccalaureate and Beyond (B&B:08/18): First Look at the 2018 Employment and Educational Experiences of 2007-08 College Graduates. See tables

Tables Library:
You can explore data tables featured in NCES publications in the Tables Library, looking at latest releases by publication title or all releases. You can also search the database by keywords or by topic.



Example Tables

8 Tables Found

Publication (2022251): Baccalaureate and Beyond (B&B:16/20): A First Look at the 2020 Experiences of 2015–16 College Graduates During the COVID-19 Pandemic - [View Publication](#)

1-8 of 8 results

Table Number	Table Title	Data Source	
A-1	PROFESSIONAL EXPERIENCES DURING THE COVID-19 PANDEMIC: Percentage of 2015–16 bachelor's degree earners who said they had various employment and enrollment experiences due to the COVID-19 pandemic, by demographic, enrollment, and employment characteristics in 2020	B&B:16/20	View
A-2	PERSONAL EXPERIENCES DURING THE COVID-19 PANDEMIC: Percentage of 2015–16 bachelor's degree earners who said they had various family and financial experiences due to the COVID-19 pandemic, by demographic, enrollment, and employment characteristics in 2020	B&B:16/20	View
A-3	STUDENT DEBT REPAYMENT DURING THE COVID-19 PANDEMIC: Among 2015–16 bachelor's degree earners who took out federal student loans to pay for undergraduate or graduate education, percentage who had all federal student loans in forbearance in February 2020 and June 2020, and among those in administrative forbearance as of June 2020, percentage who made payments between March 2020 and June 2020, and among those who made payments during administrative forbearance, cumulative amount paid in prior 3.5 months as of February 2020 and June 2020, by demographic, enrollment, and employment characteristics in 2020		
A-4	EMPLOYMENT STATUS 4 YEARS AFTER BACHELOR'S DEGREE COMPLETION: Percentage distribution of employment status 4 years after bachelor's degree completion, by whether 4 years after completing the 2015–16 bachelor's degree was before or during the COVID-19 pandemic and demographic and enrollment characteristics in 2020		
A-5a	EMPLOYMENT CHARACTERISTICS AMONG FULL-TIME WORKERS: Among 2015–16 bachelor's degree earners working full time 4 years after bachelor's degree completion, hours worked per week and annualized pay, by whether 4 years after completing the 2015–16 bachelor's degree was before or during the COVID-19 pandemic and demographic, enrollment, and employment characteristics in 2020		

When you select a publication, you are brought to a list of tables for each publication, including a descriptive title. You can view tables by clicking the **VIEW** button at the right.



Using PowerStats on DataLab



Northeastern University
NULab for Texts, Maps, and Networks

*Feel free to ask questions at any point
during the presentation!*

DataLab Statistical References

There are a number of different statistical analyses available with the PowerStats tool on the DataLab tool. Below is a brief explanation of three of the tools we recommend exploring today with links to step-by-step instructions for setting up each type of analysis:

- [Percentage Distribution](#): a distribution of frequency in which the total frequency equals 100 and individual variable or class frequencies is expressed as proportion to the total frequency ($x/100$ or $x\%$).
- [Averages, Medians, & Percents](#): descriptive statistics that mark central tendency and frequency distribution of a dataset.
- [Percentiles](#): a number or value at which a specified s percentage of data falls below, expressed as the “xth” percentile.



Case Study: Baccalaureate & Beyond

All Populations ▾ All Topics ▾ ⓘ Trends Available ☐

Choose a dataset below for information about the populations and topics covered. Then launch the dataset you would like to analyze in PowerStats. For studies with more than one data collection, you will need to select the collection you want to use. You can change datasets or data collections at any time.

Baccalaureate and Beyond (B&B)

ABOUT B&B

This study of undergraduate students who are graduating seniors collects data on students at the end of their senior year and at one year, four years, and ten years after their senior year. Use this study to learn about students' education and work experiences after completing a bachelor's degree and about the experiences of new elementary and secondary teachers.

[Visit Study Homepage](#)

POPULATION

Bachelor degree recipients who were surveyed at the time of graduation, one year after graduation, four years after graduation, and ten years after graduation

CODEBOOKS

[Variables by Subject](#)
[Variables by Name](#)

KEYWORDS

Outcomes for bachelor's degree recipients, Graduate and professional program access, Labor market experiences, Rates of return on investment in education, Post-baccalaureate education, Teacher preparation, Certifications and licenses, Enrollment while employed

ANALYSIS TYPES AVAILABLE

Percentage Distribution ⓘ
Averages, Medians, & Percents ⓘ
Percentiles ⓘ
Linear Regression ⓘ
Logistic Regression ⓘ
Correlation Matrix ⓘ

B&B collection: 2016/2017 **Launch**

Beginning Postsecondary Students (BPS)

BPS collection: 2012/2017 **Launch**

B&B Study:

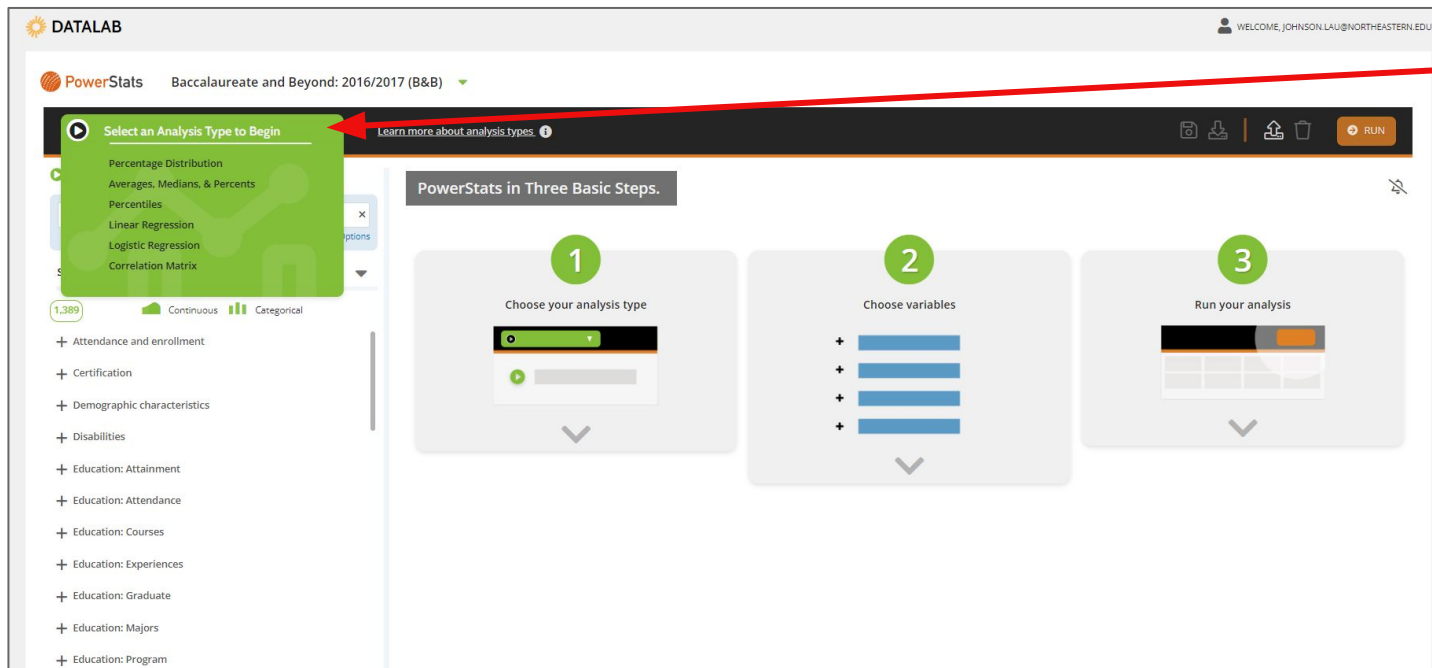
Assessment of undergraduate students during their senior year, then 4 and 10 years after graduation to learn about education and work experiences.

The DataLab interface shows you information on each study's **population**, **keywords**, and **types of analysis available**.

To get started, click on the **LAUNCH** button at the right.



Step 1: Choose your analysis type



The screenshot shows the DATALAB PowerStats interface. At the top, it says 'DATALAB' and 'WELCOME, JOHNSON.LAU@NORTHEASTERN.EDU'. Below that, it says 'PowerStats' and 'Baccalaureate and Beyond: 2016/2017 (B&B)'. A red arrow points to a green box titled 'Select an Analysis Type to Begin' which lists: Percentage Distribution, Averages, Medians, & Percents, Percentiles, Linear Regression, Logistic Regression, and Correlation Matrix. To the right of this box is a link 'Learn more about analysis types'. Below the green box is a list of variables: Attendance and enrollment, Certification, Demographic characteristics, Disabilities, Education: Attainment, Education: Attendance, Education: Courses, Education: Experiences, Education: Graduate, Education: Majors, and Education: Program. The main area of the interface is titled 'PowerStats in Three Basic Steps.' and shows three steps: 1. Choose your analysis type, 2. Choose variables, and 3. Run your analysis. Each step has a corresponding icon and a dropdown arrow.

Step 1: Choose your analysis type.

There are a number of built-in statistical analyses available in the DataLab interface, including **percentile distribution, averages, linear regressions, and correlation matrices.**



Study Variables (by Subject)

[Attendance and enrollment](#)
[Certification](#)
[Demographic characteristics](#)
[Disabilities](#)
[Education: Attainment](#)
[Education: Attendance](#)
[Education: Courses](#)
[Education: Experiences](#)
[Education: Graduate](#)
[Education: Majors](#)
[Education: Program](#)
[Education: Services](#)
[Education: Tests](#)
[Education: Transfer](#)
[Educational expectations](#)
[Employment](#)
[Employment: Benefits](#)
[Employment: Description](#)
[Employment: Employer](#)
[Employment: Future](#)
[Employment: History](#)
[Employment: Satisfaction](#)
[Employment: Search](#)
[Employment: Status](#)
[Family](#)
[Family: Finances](#)
[Finances](#)

[Download HTML](#)

National Center for Education Statistics PowerStats

Study name: *Baccalaureate and Beyond: 2016/2017*

Created on: 11/02/2022 at 23:55:05 from <https://nces.ed.gov/DataLab>

Subject: Attendance and enrollment

Label: Primarily student or employee while concurrently employed and enrolled, between BA completion and June 2017

Name: B1WRKS

Description: Indicates whether the respondent considered himself or herself a student working to meet expenses or an employee who decided to enroll in school, between completion of the respondent's 2015-16 bachelor's degree and June 2017.

Source: B&B:16/17 Interview

Descriptive Statistics:

Value	Percentage	Value label
1	17.48	A student working to meet expenses
2	4.75	An employee who decided to enroll in school
-3	77.78	{Skipped}

Weight used in frequency: (WTA000)

Programming Notes: A student who works to meet expenses is a student who is enrolled full time, but also holds a part time job to earn additional money. An employee who also attends school is a student who considers his or her primary focus to be employment but is attending school in order to further his or her career.

Survey Questions: When you were last enrolled as a student and also working, would you say you were primarily: 1 = A student working to meet expenses; 2 = An employee who decided to enroll in school

Applies To: Respondents who were concurrently working and enrolled between completion of their 2015-16 bachelor's degree and June 2017.

Subject: Certification

Label: First job, within 12 months after BA completion: License required for work

Name: B1LICREQ01ST

Study Variables:

Each of the codebooks for the NCES studies include detailed information for the different variables including:

- **Subject title**
- **Label + Name**
- **Description**
- **Source**
- **Descriptive Statistics:** percentage of values for this single variable
- **Programming notes:** information about when participants were surveyed about this topic.
- **Associated Survey Questions** (including possible answers)



Step 2: Choose your variables.

The screenshot shows the DATALAB interface for the 'Baccalaureate and Beyond: 2016/2017 (B&B)' study. The 'Percentage Distribution' analysis type is selected. In the 'Choose Variables' section, the 'Categorical' filter is active, and a red arrow points from it to the 'Columns' section. The 'Columns' section has a placeholder 'Drag and drop column variable here'. Below it, a table shows the variable structure:

Row Variable Label	Column Category 1	Column Category 2	Column Category 3	To
Category	#	#	#	10
Category	#	#	#	10
Row Variable Label				
Category	#	#	#	10
Category	#	#	#	10

Step 2: Choose your variables.

Each study has many different variables, more than we can look at!

Continuous variables:
Variables with data that is categorized or grouped.

Categorical variables:
Variables that can be measured, but with a theoretically infinite number of values.

Drag your variable to the column and row sections.



Variable Customization: Gender

The screenshot shows a web interface for customizing a variable. At the top, a dark header bar displays 'VARIABLE: Gender, as of B&B:16/17 interview' with a pencil icon and a description: 'Indicates the respondent's identified gender, as of B&B:16/17 interview.' Below this, three tabs are visible: 'USE & CUSTOMIZE' (active), 'DESCRIPTIVE STATISTICS', and 'MORE INFO'. Under the 'USE & CUSTOMIZE' tab, there's a section titled 'Default Categories' with the instruction 'Select categories to add them to your table. Unselected categories will be excluded from your analysis.' It includes 'Select All' and 'Deselect All' buttons. A table lists six categories, each with a checked checkbox, a 'Value' column, and a 'Label' column. To the right of the table, a text box says 'You can also combine categories to make your own.' with a 'Create Your Own Custom Categories' button. At the bottom left, a green arrow points to 'CHANGE VARIABLE USAGE'. At the bottom right, there is a 'Save' button. A red arrow points from the 'Create Your Own Custom Categories' button to the text box on the right.

VARIABLE:
Gender, as of B&B:16/17 interview ✎
Indicates the respondent's identified gender, as of B&B:16/17 interview.

USE & CUSTOMIZE DESCRIPTIVE STATISTICS MORE INFO

Default Categories
Select categories to add them to your table. Unselected categories will be excluded from your analysis.

Select All Deselect All

	Value	Label
<input checked="" type="checkbox"/>	1	Male
<input checked="" type="checkbox"/>	2	Female
<input checked="" type="checkbox"/>	3	Transgender: male-to-female
<input checked="" type="checkbox"/>	4	Transgender: female-to-male
<input checked="" type="checkbox"/>	5	Genderqueer or gender nonconforming
<input checked="" type="checkbox"/>	6	A different gender identity

Categories in brackets are skips or missing values. To learn more, view the descriptive statistics.

CHANGE VARIABLE USAGE Save

You can also combine categories to make your own.

Create Your Own Custom Categories

Variable Customization Prompt

After choosing a variable, a variable usage and customization box will open up on the screen.

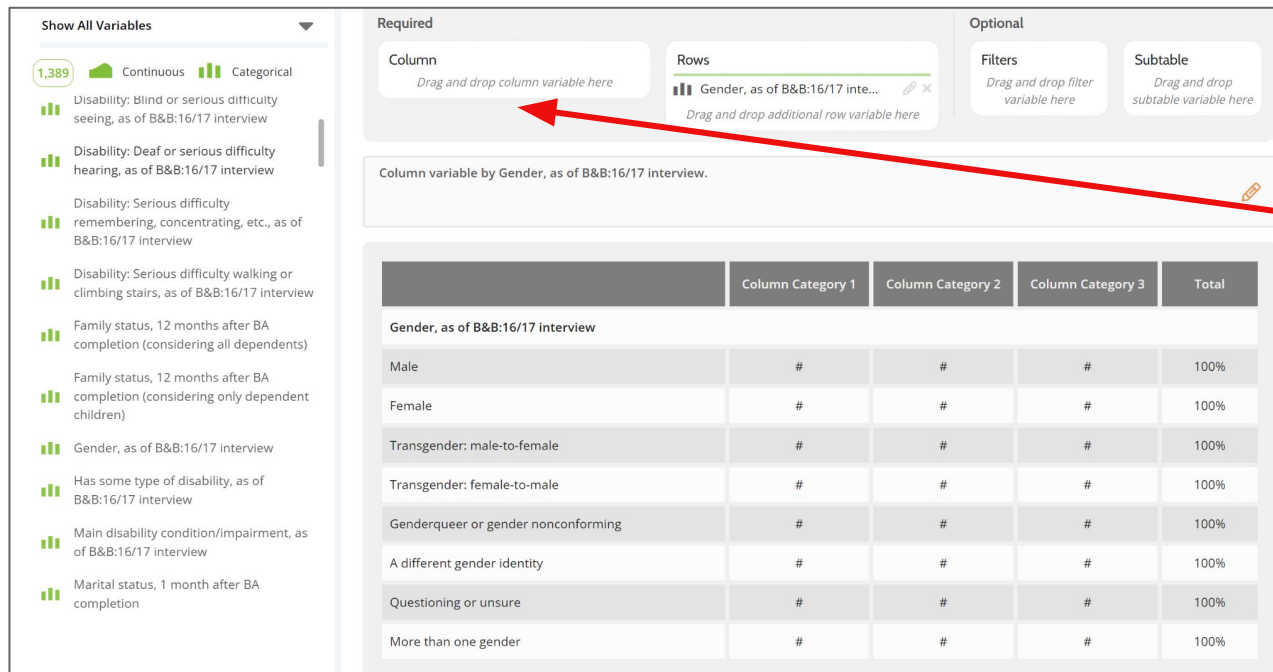
In this window, you can customize your variable. Each variable has different options, but this allows you to change or **customize** the variables beyond the **default values**.

For example, you can choose to leave out certain variable values from your analysis and table, or create new categories of your own to group variable values.

Question: what are some possible benefits or drawbacks of this feature?



Step 2: Choose your variables (cont.)



The interface shows a list of variables on the left and a selection area on the right. A red arrow points from the 'Required' section to the 'Column' variable.

Show All Variables

- 1,389 Continuous Categorical
- Disability: Blind or serious difficulty seeing, as of B&B:16/17 interview
- Disability: Deaf or serious difficulty hearing, as of B&B:16/17 interview
- Disability: Serious difficulty remembering, concentrating, etc., as of B&B:16/17 interview
- Disability: Serious difficulty walking or climbing stairs, as of B&B:16/17 interview
- Family status, 12 months after BA completion (considering all dependents)
- Family status, 12 months after BA completion (considering only dependent children)
- Gender, as of B&B:16/17 interview
- Has some type of disability, as of B&B:16/17 interview
- Main disability condition/impairment, as of B&B:16/17 interview
- Marital status, 1 month after BA completion

Required

Column
Drag and drop column variable here

Rows
Gender, as of B&B:16/17 inte...
Drag and drop additional row variable here

Optional

Filters
Drag and drop filter variable here

Subtable
Drag and drop subtable variable here

Column variable by Gender, as of B&B:16/17 interview.

	Column Category 1	Column Category 2	Column Category 3	Total
Gender, as of B&B:16/17 interview				
Male	#	#	#	100%
Female	#	#	#	100%
Transgender: male-to-female	#	#	#	100%
Transgender: female-to-male	#	#	#	100%
Genderqueer or gender nonconforming	#	#	#	100%
A different gender identity	#	#	#	100%
Questioning or unsure	#	#	#	100%
More than one gender	#	#	#	100%

Step 2: Choose your variables (cont.)

After choosing one of your variables, you will need to choose the second. You can swap the variables between the column or row categories to change how the data will be displayed in the final table. After choosing one variable, it will show you what the format for the table will look like before you run it.



Variable Customization: Field of Study

VARIABLE:
Field of study: Undergraduate (10 categories)

Indicates the respondent's major or field of study, using 10 categories, for the 2015-16 bachelor's degree.

USE & CUSTOMIZE **DESCRIPTIVE STATISTICS** MORE INFO

Default Categories
Select categories to add them to your table. Unselected categories will be excluded from your analysis.

Select All Deselect All

	Value	Label
<input checked="" type="checkbox"/>	1	Computer and information sciences
<input checked="" type="checkbox"/>	2	Engineering and engineering technology
<input checked="" type="checkbox"/>	3	Biological and physical science, science technology, math, and agriculture
<input checked="" type="checkbox"/>	4	General studies and other
<input checked="" type="checkbox"/>	5	Social sciences
<input checked="" type="checkbox"/>	6	Humanities

Categories in brackets are skips or missing values. To learn more, view the descriptive statistics.

CHANGE VARIABLE USAGE Save

You can also combine categories to make your own.
Create Your Own Custom Categories

Variable Customization Prompt

Here is an example of another customization prompt for a second variable, this time for field of undergraduate study.

While having access to the codebook for the study variables is helpful, you can access this information about the name, subject, descriptive statistics, and survey questions for each variables on the other tabs **DESCRIPTIVE STATISTICS** and **MORE INFO**.



Step 3: Run your analysis

The screenshot shows the PowerStats web application interface. At the top, the title is "Baccalaureate and Beyond: 2016/2017 (B&B)". Below the title bar, there's a green button labeled "Percentage Distribution" and a link "Learn more about analysis types". On the right side of this bar, there's an orange button labeled "RUN" which is highlighted with a red box and a red arrow pointing to it. Below the title bar, there's a "Choose Variables:" section with a search bar and a list of variables. The variables are categorized into "Continuous" and "Categorical". The "Categorical" section is expanded, showing a list of variables including "Education: Experiences", "Education: Graduate", "Education: Majors", "2015-16 BA major or field of study same as highest post-bachelor's major or field of study, within 12 months after BA completion", "BA major (detailed), 2015-16", "Field of study: Undergraduate (10 categories)", and "Highest degree enrollment, within 12 months after BA completion: Major or field of study (23 categories)". Below the variable list, there's a "Required" section with two columns: "Column" and "Rows". The "Column" column has a variable "Field of study: Undergraduat...". The "Rows" column has a variable "Gender, as of B&B:16/17 inte...". Below the "Required" section, there's a "Field of study: Undergraduate (10 categories) by Gender, as of B&B:16/17 interview." section. Below this, there's a table with columns for "Computer and information sciences", "Engineering and engineering technology", "Biological and physical science, science technology, math, and agriculture", "General studies and other", "Social sciences", "Humanities", and "Health care fields". The table is titled "Gender, as of B&B:16/17 interview".

Step 3: Run your Analysis.

Once you have selected your variables from the list and formatted them (choose which value options you want or not, if applicable), hit the **RUN** button.

NOTE: you can change the title of the table in the box with the orange edit arrow. Below that is a “draft” preview of what the table will look like, including the units of the numbers expected in each of the cells (% here).



Percentile Distribution: Gender + field of study

PowerStats Baccalaureate and Beyond: 2016/2017 (B&B) ▼

Show: Estimates and Standard Errors ▼

Field of study: Undergraduate (10 categories) by Gender, as of B&B:16/17 interview.

	Computer and information sciences %	Engineering and engineering technology %	Biological and physical science, technology, math, and agriculture %	General studies and other %	Social sciences %	Humanities %	Health care fields %	Business %	Education %	Other applied %	Total
Estimates											
Total	3.4	6.2	12.1	2.4	14.8	9.9	11.5	19.9	4.6	15.0	100%
Gender, as of B&B:16/17 interview											
Male	6.8	11.6	12.6	2.1	11.3	9.7	4.3	24.9	2.2	14.4	100%
Female	1.0	2.2	11.7	2.6	17.3	10.0	17.0	16.4	6.5	15.2	100%
Transgender: male-to-female	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	100%
Transgender: female-to-male	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡	100%
Genderqueer or gender nonconforming	‡	‡	23.6 !	‡	17.6 !	26.2 !	2.4 !!	7.6 !!	‡	19.5 !	100%

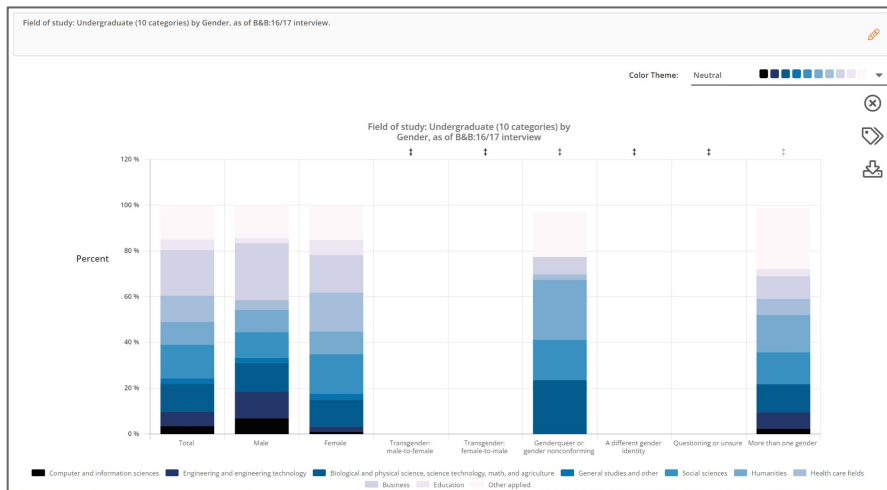
Step 4: Explore your resulting statistical table!

Once you have run your statistical analysis, you will see a large table that has run the calculation for the variables you selected. Feel free to scroll down and explore the information it yielded.

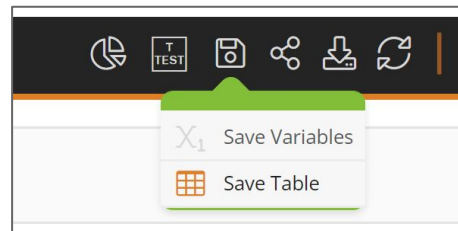


Next Steps: what can I do with this analysis?

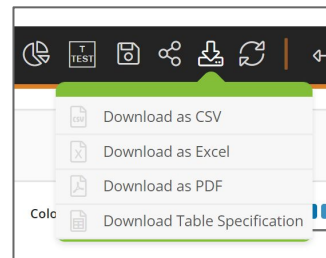
1. **Create charts:** after running an analyses, you can create bar charts, pie charts, or a stacked bar chart from the data.



2. **Save your table to your DataLab library.**



2. **Download your table and analyses:** choose from CSV, Excel, PDF, or other specifications. This includes the confidence levels, relative standard error, etc.



Using DataLab: Exploration and Discussion



Northeastern University
NULab for Texts, Maps, and Networks

*Feel free to ask questions at any point
during the presentation!*

Your Turn!

Use one of the PowerStats options to explore different ways to analyze the Baccalaureate & Beyond NCES study. If you have questions about how to set up these different analyses, see the step-by-step tutorials in the [DataLab Learning Center](#).

Consider the following questions:

1. What variables interest you? How might you find variables in this (or other studies) that connect to your research interests?
2. What kind of questions are you interested in asking about these variables? What comes to mind?
3. Are there any variables that are missing?



Your Turn!

Consider the following additional questions:

1. What does it mean to have caution around statistical analysis?
2. Did you try any of the variable customization features? Why or why not?
3. How might customizing the values for variables help or hinder your analysis or presentation of statistical results?
4. Did you use the table/chart tool? Did being able to present the data in a non-table format help you look at it in a new way?
5. What connections do you see between this tool and the discussions you have had about surveys/assessments so far?
6. How does thinking or looking at data in a database level change what kind of questions or observations you have?



Thank you!

If you have any questions, contact us at: nulab.info@gmail.com

Sign up for office hours at: <http://calendly.com/diti-nu/>

We'd love your feedback! Please fill out a short survey here:

<https://bit.ly/diti-feedback>

Slides & handouts available at: <https://bit.ly/fa22-poe-surveyanalysis>

