PS 211: Introduction to Experimental Design

Fall 2025 · Section C1

Lecture 1: Course Overview, Intro to Statistics and Research Design

Statistics!

Outline for Today

- Introductions
- Logistics (meeting times, locations)
- Syllabus review
- Tools you'll use
- Course goals & outcomes
- Requirements & grading
- Policies, participation, and help
- Getting started with Slack, R, and R Studio
- Q&A

Instructor

Teaching Fellow

Juneau Wang

PhD Student in Psychological & Brain Sciences

- **Pronouns:** he/him
- **Email:** juneauw@bu.edu
- **Office:** 111 Cummington Mall, Room 242
- **Office Hours:** Tuesdays 12:30 - 2:30 p.m.

Logistics

- **Lecture (C1):** Tue & Thu · 11:00 a.m. – 12:15 pm
- **Discussions:**
  - C2: Wed 12:20 – 1:10 p.m.
  - C3: Wed 1:25 – 2:15 p.m.

All course meetings in CAS 306

Communication logistics

- We will use **Slack** for class communication. Everyone should be added. Please reach out to Kate or Juneau if you are not yet on the Slack workspace.
- Please check Slack regularly for important announcements.
- Please also use the Slack to ask questions about course material!
  - In general, our policy is that if you ask us a question about course material, we will publicly post it (anonymously, if you desire) and the answer on Slack, so that everyone can benefit from the learning opportunity.
- That said, our hope is that you will take advantage of lecture, discussion, and office hours to ask questions **IN PERSON**. If you send us a content-based question via email or slack, we may ask you to come to office hours or ask during lecture/discussion instead.

What is this class?

Introduction to the logic and processes involved in descriptive & inferential statistics for psychology

This class involves math and programming, but the goal of the course is not for you to derive or memorize complex formulas or to become an expert coder.

The goal is for you to strengthen your critical thinking and quantitative reasoning skills, and learn how to conduct statistical analyses and interpret their results **with as little pain as possible.**

Course learning goals

In this course, you will (ideally) learn:

1) How to describe and measure different types of variables.

2) The difference between populations and samples, and the concept of sampling.

3) What it means to test a hypothesis.

4) How to use R to analyze data.

5) How to interpret the results of statistical tests.

6) How to communicate results via graphs, figures, and written text.

7) The limits and risks of statistics, and how to avoid misusing them.

Specific topics we will cover

We will cover:

1) Frequency distributions

2) Measures of central tendency and variability.

3) Sampling and basic probability.

4) The normal distribution, standardization, z-scores.

5) Hypothesis-testing.

6) Confidence intervals, effect sizes, statistical power.

7) T-tests, ANOVA (one-way, two-way, repeated-measures), Correlations, Regressions, Chi-Square tests.

Course Materials

There is *NO TEXTBOOK* for this course

That means it's very important that you come to class, because we will go over all the material you need to know for the exams and assignments.

However, we will use ...

R and R Studio

We will use R and R studio for statistical analyses. Both are free and open-source software, which means you can download and use them without paying anything.

This is **not** a coding class. We do not expect students to have prior programming experience, and we will walk you through everything you need to know. Do not panic!

Course Requirements & Grading

- **Exams (3 × 15% + 1 × 25%) = 70%**

  Exams will involve short-answer + multiple-choice questions.

  - Exams will be *conceptual*.
  - They will not test your knowledge of complex formulas or coding, but rather your understanding of the logic behind statistical tests and your ability to interpret results.
  - A review sheet will be provided; You may write notes on it and use it during the exam.
  - Calculators are allowed, but phones and computers are not. We will let you know ahead of time if a calculator is necessary.
  - Make-ups will be given for excused/extreme circumstances only; Make-ups must be taken within 1 week of the scheduled exam.

Course Requirements & Grading (cont.)

- **Homework Assignments (3 × 3.33%) = 10%**
  - Four assignments; lowest homework grade is dropped.
  - You will answer conceptual questions and use R to analyze and interpret data.
  - R Markdown templates will be provided so you will not have to code from scratch.
  - Homeworks are due 11:59 PM ET on syllabus dates.
  - Submit on Blackboard as HTML files, knitted from R Markdown. (instructions to follow!)
  - Late homeworks will lose 3 points each day they are late; No late work accepted after answers are posted.

Course Requirements & Grading (cont.)

- **Data Write-Ups (1 × 10%) = 10%**
  - You will use R to analyze and interpret a provided dataset.
  - R Markdown templates will be provided so you will not have to code from scratch.
  - You will produce APA-style results & discussion sections (+ figures/tables).
  - The data write-up is due 11:59 PM ET on syllabus dates.
  - You will lose 3 points each day it is late; No late work accepted after answers are posted.

Course Requirements & Grading (cont.)

- **Discussion Section = 10%**
  - In-depth review of class material and semester-long research project.
  - In small groups, you will come up with a research hypothesis, conduct a literature review, and write hypothetical methods, results, and discussion sections to be presented on a poster at the end of the semester.
  - If you miss a discussion section, contact your TF about make-up work.
  - If you miss three or more discussion sections, you will not get full credit.
  - If you miss five or more discussion sections, you will get a 0 for discussion participation.

Attendance

Lecture attendance is highly encouraged but not required

- Missing class limits learning & participation.
- Students who attend lectures (esp. later in semester) generally earn better grades.
- In class we go over problems that prepare you for assessments.
- All lecture slides will be posted right before or after class, but they will be hard to understand without attending.

  > There will be occasional pop quizzes administered in class (lecture and discussion).

  > These short multiple-choice quizzes will enable you to earn bonus points on the next exam.

  > There will be NO make-ups. If you miss class, you miss the opportunity to earn bonus points.

Participation

Please ask questions and engage. Speaking up can feel hard but your classmates will benefit from your questions!

Office hours: If you prefer to ask questions in a one-on-one setting, please come to office hours. We are happy to talk to you!

- That said, our availability to meet outside of scheduled office hours is limited.

Class Policies: Respect

General policies:

We will build a classroom that supports everyone's ability to learn.

We will strive to create a classroom environment that is inclusive of all backgrounds.

Language or behavior that undermines learning will be addressed and, if repeated, may affect grades.

Specific policies:

Bring your laptop to class for note-taking and R demonstrations.

If tech use distracts others and continues after a warning, you may be asked to leave.

Food and drink are okay if not distracting.

Class Policies: Ask for Help

If you're struggling or perform poorly on an exam or assignment:

- Come to office hours and talk to us!
- We will not be returning exams, but we can go over what you got wrong and how to improve.

Take advantage of campus resources, including the:

- **Educational Resource Center (ERC):** tutoring, time management, academic planning
  - http://www.bu.edu/erc

Class Policies: Collaboration and Use of Artificial Intelligence

Collaboration is a huge part of science. Learning how to collaborate effectively is an important skill.

- As such, you are allowed to ask each other for help on homework assignments and data write-ups.
- However, you must each submit your own work. Copying someone else's work is plagiarism.

You may also use AI tools (e.g., ChatGPT) to help you with your code.

- However, use them cautiously. They can be wrong!
- In addition, remember that you will need to understand the statistical concepts, as you will complete the exams on your own without any help.

There will be a place on the assignment to acknowledge your collaborators and any AI tools you used. You will not be penalized for collaboration or use of AI if you disclose it honestly.

Accommodations

- Boston University provides reasonable accommodations for students with documented disabilities.
- Please reach out to Kate and Juneau (via Slack) **as soon as possible** and share your paperwork from disability services.

Pre-course survey

- We have posted a link to a pre-course survey on Slack.
- Please complete it by the end of the day on Friday, September 5 (for a bonus point on Exam 1).
- The survey has two parts:
  - A few questions about your background and comfort with math and programming.
  - A few fun questions to generate a dataset for us to analyze later in the semester.

Questions?

- What can we clarify about the course structure?
- What can we clarify about course policies?

It is my first time teaching this course!

It is actually my first time teaching ever!

Acknowledgments

- Thank you to Dr. Chloe Jordan for sharing her syllabus and lecture slides, which I adapted for this course!

Getting started with Slack

(Slack demo)

Why are we using R?

- R is a free and powerful programming language for data analysis and visualization.
- R is widely used in academia and industry.
- My goal in this course is to introduce you to R. You will be able to do well (and get a 100) in this class *without becoming fluent in R.*
  - You are welcome to go above and beyond the course requirements and challenge yourself to learn more R on your own!
  - I use R in my own research and am always happy to talk to students who want to learn more.

You may even find it fun!

Installing R and R Studio

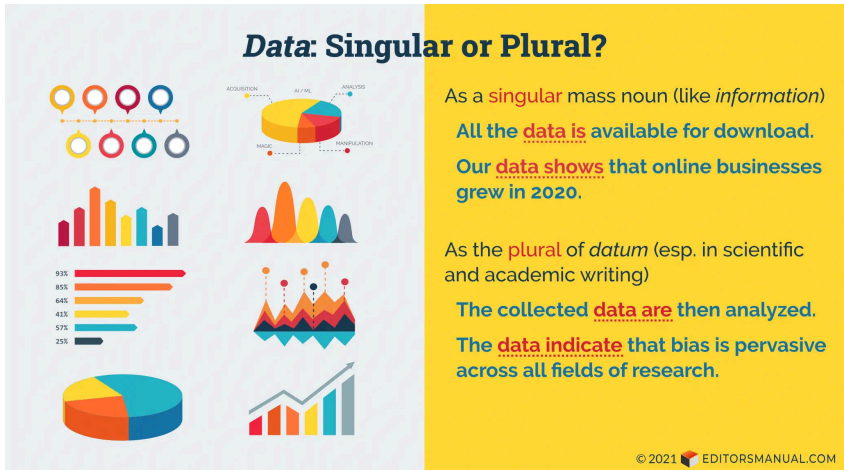
- There are instructions pinned on Slack for how to install R and R Studio on your computer.
- Please follow them carefully as soon as possible.
- If you run into issues, please:
  - Come to office hours.
  - Ask Juneau for help during your discussion section.
  - Post on the Slack to see if a classmate can help you.
- We will go over the basics of R markdown next week when we assign the first homework.

Now let’s get started with statistics!

What is/are data?

Data:

- Generally: values that convey information.
- Numbers, words, observations, etc.
- Used for calculation, reasoning, discussion, decision-making.



What is/are statistics?

Depending on the context, "statistics" can refer to:

- Data.
- Numerical values that summarize data (e.g., mean, median, standard deviation).
- Techniques for the collection, analysis, and interpretation of data.
- The science of creating and applying such techniques.

Why should I care about statistics?

Statistics are not just important for scientists.

Statistics are used constantly in daily life.

- Statistics are used for:
  - Making medical decisions (e.g., screening tests, treatment options)
  - Making financial decisions (e.g., investments, loans)
  - Understanding public policy (e.g., election polls, public health data)
  - Evaluating news and media reports (e.g., understanding graphs, charts, and statistics presented in the media)

Goals of science

Describe

- Observe and/or measure natural phenomena or behavior (one or more variables).

Predict

- Identify relationships between variables.

Explain

- Determine causal relationships between variables.

The scientific method

1. Make an observation

- Notice something in the world.
- Observations can arise from experience, introspection, or reading prior research.

2. Research existing literature

- Review prior research to see what is already known about the topic.
- Identify gaps in knowledge or areas for further investigation.

3. Develop your hypothesis

- Formulate a testable prediction about the relationship between variables.
- In general, good hypotheses are specific and falsifiable.

Example: "Increased study time will lead to higher exam scores."

The scientific method (cont.)

4. Design your study

- Decide how to test your hypothesis.
- Choose a research method (e.g., experiment, survey, observational study).
- Identify your variables (independent, dependent, control).

5. Collect and analyze data

- Gather data according to your study design.
- Use statistical methods to analyze the data and test your hypothesis.
- Determine whether the results support or refute your hypothesis.

6. Draw conclusions and communicate results

- Interpret the results in the context of your hypothesis and existing literature.
- Share your findings through reports, presentations, or publications.

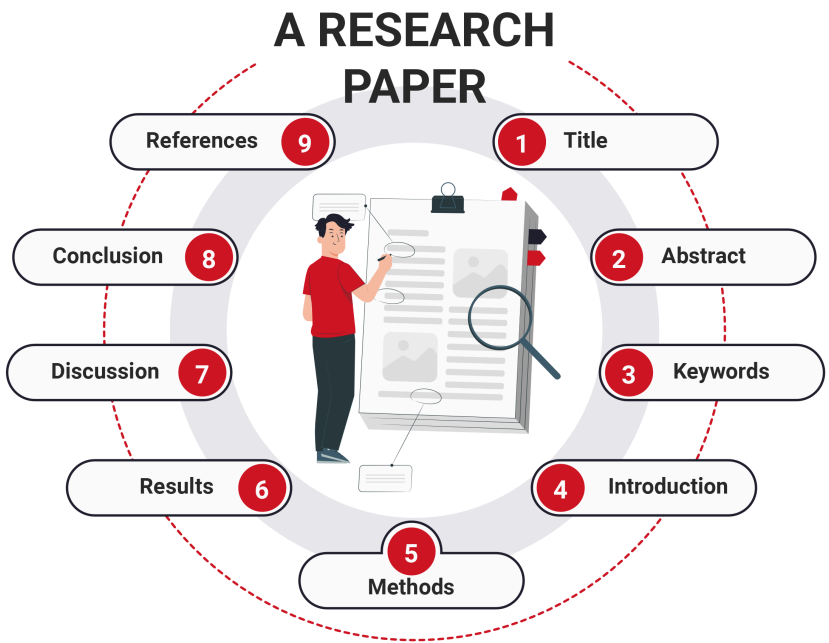
The scientific method (cont.)

In this class, we will focus on some of the key steps of this method, including:

- Developing hypotheses.
- Determining whether study designs are appropriate for testing specific hypotheses.
- Selecting the right statistical tests for analyzing data.
- Interpreting the results of statistical tests to draw conclusions.

The scientific publication process

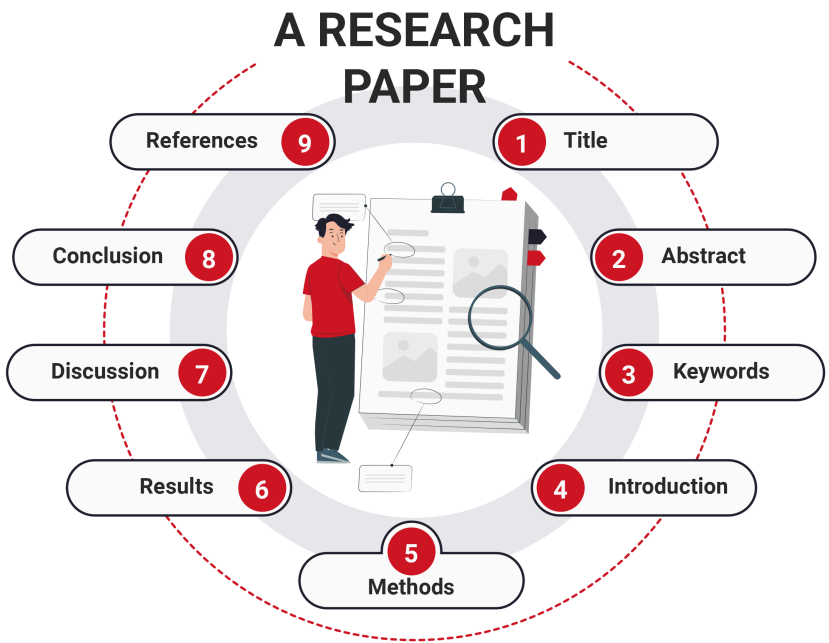
- Once studies are completed, scientists write up their findings in a manuscript.
- Manuscripts are submitted to peer-reviewed journals.
- Other scientists in the field review the manuscript and provide feedback.



The scientific publication process (cont.)

- Authors almost always need to revise their manuscript based on this feedback.
- If the manuscript meets the journal's standards, it is published and becomes part of the scientific literature.
- This is important because it enables other scientists to evaluate, replicate, and build upon the findings.

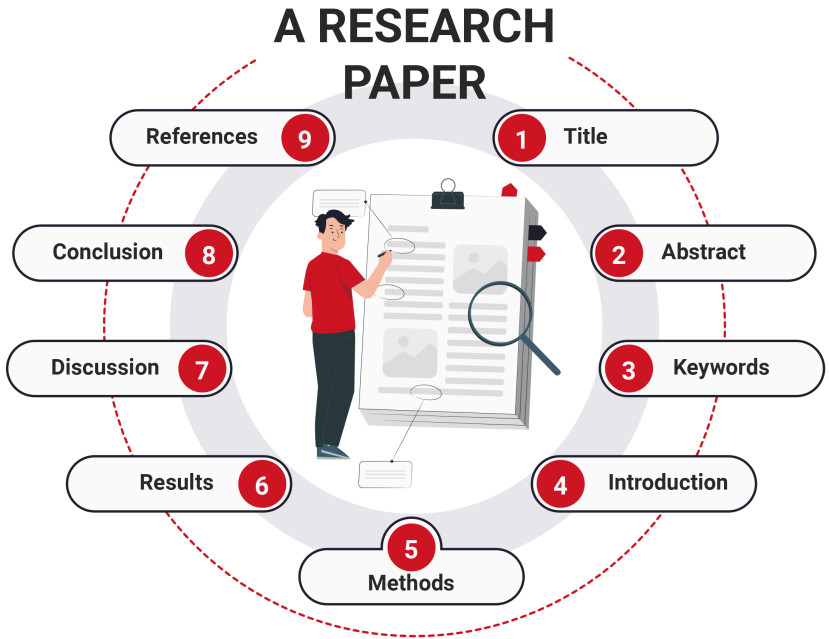
Caveat: Not all published research is useful AND not all useful research is published in traditional journals.



The scientific publication process (cont.)

- Scientific articles often follow a standard structure.
- This structure helps readers quickly find the information they need.
- In this class, we will focus on how researchers put together the methods, results, and discussion sections of a research article.

There are multiple types of scientific publications. We will focus on empirical research articles.



Example of an original research article

Warning: shameless self promotion.

Example of an original research article

Warning: shameless self promotion.

Two types of statistics

1. Descriptive statistics

- Organize, summarize, and communicate numerical information.

2. Inferential statistics

- Use sample data to make inferences about a larger population.
- Enable descriptions, predictions, and explanations to be generalized beyond the specific data collected.

Populations

- The entire group of individuals or items that you want to draw conclusions about.
- Examples: All rats in the world, all adults in the U.S., all students at BU, all voters in an election.

The problem:

- We usually cannot collect data from an entire population.
- Example: Leading up to elections, we cannot ask every single voter who they plan to vote for.

Populations and samples (cont.)

The solution: Samples

- A subset of individuals or items selected from a population.
- Examples: 100 rats from a lab, 1,000 adults surveyed in the U.S., 200 students at BU, 500 voters polled in an election.

Goal:

- Use data from a representative sample to make inferences about the larger population.
- This is where inferential statistics come in!

Practicing the concepts

According to the World Health Organization, approximately 300 million people worldwide suffer from depression.

Do you think the research behind this number is based on a sample or a population? Explain.

Is 300 million a descriptive or inferential statistic? Explain.

Practicing the concepts

According to the World Health Organization, approximately 300 million people worldwide suffer from depression.

Do you think the research behind this number is based on a sample or a population? Explain.

Answer: Sample. It would be nearly impossible to survey every single person in the world. Instead, researchers likely surveyed a representative sample of people and used inferential statistics to estimate the total number of people with depression worldwide.

Is 300 million a descriptive or inferential statistic? Explain.

Practicing the concepts

According to Boston University, 91% of incoming first-years were in the top 10% of their high school class.

Do you think the research behind this number is based on a sample or a population? Explain.

Is 91% a descriptive or inferential statistic? Explain.

Practicing the concepts

According to Boston University, 91% of incoming first-years were in the top 10% of their high school class.

Do you think the research behind this number is based on a sample or a population? Explain.

Answer: Population. This statistic likely includes all incoming first-year students at Boston University, rather than a sample.

Is 91% a descriptive or inferential statistic? Explain.

Transforming observations into variables

Variable

- Any observation that can take on different values.
- Research process begins by transforming observations into variables.

Examples of variables in psychology:

- Responses to surveys or questionnaires (e.g., depression scores, personality traits).
- Behavioral observations (e.g., number of times a rat presses a lever, time spent on a task).
- Physiological measurements (e.g., heart rate, brain activity).

Different types of variables

Qualitative

- Differ in quality (meaning character or kind) not quantity.
- Also called a 'nominal' variable.
- Examples: hair color, gender, species.

Quantitative

- Differ in quantity (meaning amount or number).
- Examples: height, weight, age.

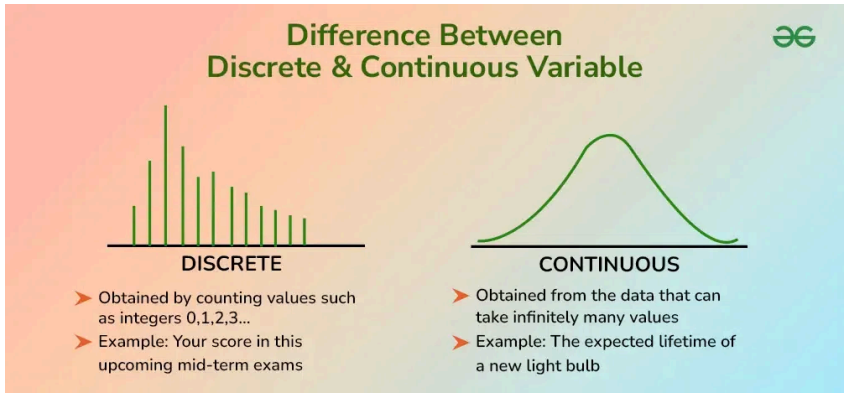
Subcategories of variables

Continuous (quantitative)

- Can assume any value within a range.

Discrete (quantitative or qualitative)

- Can only take on specific values.



Types of discrete variables

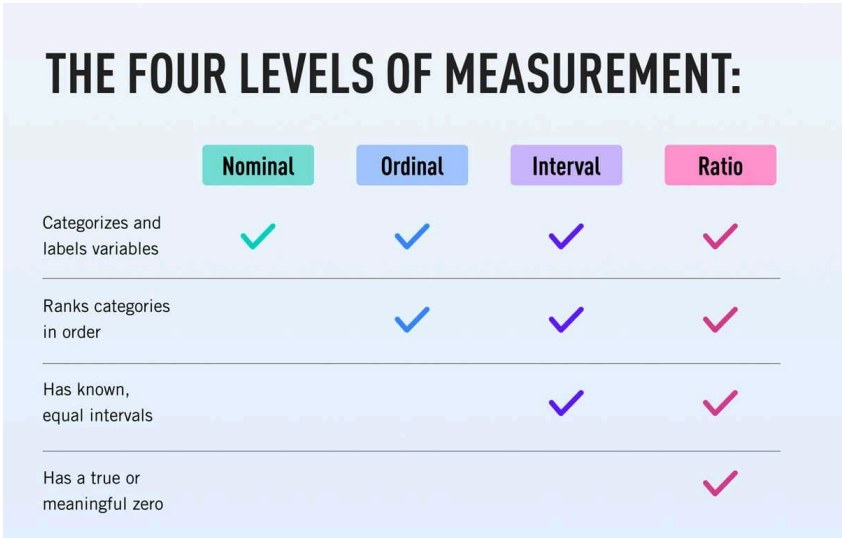
Types of continuous variables

Interval variables

- Used for labeling variables where the intervals between values are meaningful and equal.
- No true zero point (i.e., zero does not indicate the absence of the quantity being measured).

Ratio variables

- Interval variables with a true zero point (i.e., absence of the quantity being measured).



Can you think of examples of each type?

Types of variables: practice

You want to conduct a survey of college students to understand their study habits and academic performance. You collect the following data:

1. Number of hours spent studying per week.
2. Major (e.g., Psychology, Biology, Engineering).
3. GPA (Grade Point Average).
4. Class rank (e.g., Freshman, Sophomore, Junior, Senior).
5. Satisfaction with academic experience (rated on a scale from 1 to 5).

Which of these variables are qualitative and which are quantitative?

Which of these variables are discrete and which are continuous?

Which of these variables are nominal, ordinal, interval, and ratio?

Types of variables: practice

You want to conduct a survey of college students to understand their study habits and academic performance. You collect the following data:

1. Number of hours spent studying per week.
2. Major (e.g., Psychology, Biology, Engineering).
3. GPA (Grade Point Average).
4. Class rank (e.g., Freshman, Sophomore, Junior, Senior).
5. Satisfaction with academic experience (rated on a scale from 1 to 5).

Which of these variables are qualitative and which are quantitative?

Answer: Quantitative: 1, 3, 5; Qualitative: 2, 4.

Which of these variables are discrete and which are continuous?

Answer: Discrete: 2, 4, 5; Continuous: 1, 3.

Which of these variables are nominal, ordinal, interval, and ratio?

Answer: Nominal: 2; Ordinal: 4, 5; Interval: None; Ratio: 1, 3

Putting it all together: practice

You are not sure if you want to take this class, so you design an experiment to see if BU students who take PS 211 are generally happier than students who do not take PS 211. To do this, you email 200 students who took PS 211 last year and 200 students who did not take PS 211 last year. 150 students in each group respond to your email. You collect the following data:

1. Whether or not the student took PS 211 (Yes/No).
2. The student's happiness score (rated on a scale from 1 to 10).
3. The student's major (e.g., Psychology, Biology, Engineering).

What is the population in this study?

What is the sample size of this study?

You determine that the average happiness score for students who took PS 211 is 8.2, while the average happiness score for students who did not take PS 211 is 6.5. Is this a descriptive or inferential statistic?

Putting it all together: practice

You are not sure if you want to take this class, so you design an experiment to see if BU students who take PS 211 are generally happier than students who do not take PS 211. To do this, you email 200 students who took PS 211 last year and 200 students who did not take PS 211 last year. 150 students in each group respond to your email. You collect the following data:

1. Whether or not the student took PS 211 (Yes/No).
2. The student's happiness score (rated on a scale from 1 to 10).
3. The student's major (e.g., Psychology, Biology, Engineering).

What is the population in this study?

Answer: All BU students.

What is the sample size of this study?

Answer: 300 (150 from each group).

You determine that the average happiness score for students who took PS 211 is 8.2, while the average happiness score for students who did not take PS 211 is 6.5. Is this a descriptive or inferential statistic?

Answer: Descriptive statistic. These numbers directly describe the average happiness scores of the two groups in the sample.

Putting it all together: practice

You are not sure if you want to take this class, so you design an experiment to see if BU students who take PS 211 are generally happier than students who do not take PS 211. To do this, you email 200 students who took PS 211 last year and 200 students who did not take PS 211 last year. 150 students in each group respond to your email. You collect the following data:

1. Whether or not the student took PS 211 (Yes/No).
2. The student's happiness score (rated on a scale from 1 to 10).
3. The student's major (e.g., Psychology, Biology, Engineering).

How could you change Question 2 to collect a nominal variable instead of an ordinal variable?

What would be the advantage of collecting a nominal variable for Question 2 instead of an ordinal variable?

How could you change Question 2 to collect a ratio variable instead of an ordinal variable?

Putting it all together: practice

That’s all for this week! See you Thursday!