

**PSY 3307**

**Review of Common Terms**

**Jonathan A. Pedroza, MS, MA**

**Cal Poly Pomona**

**2021-08-24**

# Agenda

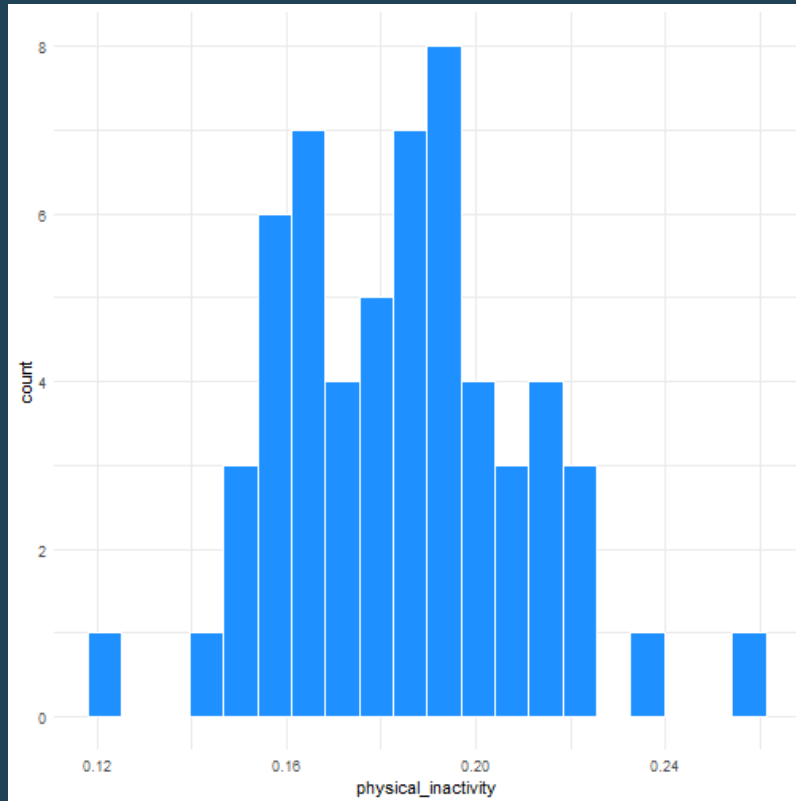
- Steps of Statistics
- Differences in Population & Sample
- Variables & Relationships
- Descriptive & Inferential Statistics
- Statistics & Parameters
- Experimental & Correlational Designs
- Experiments
- Correlational Studies
- Measurement Scales
- Continuous & Discrete Variables

# Steps of Statistics

1. Organize Data
2. Summarize Data
3. Communicate Data
4. Interpret Data

# What does this mean?

## 1. Organize Data



## 1. Summarize Data

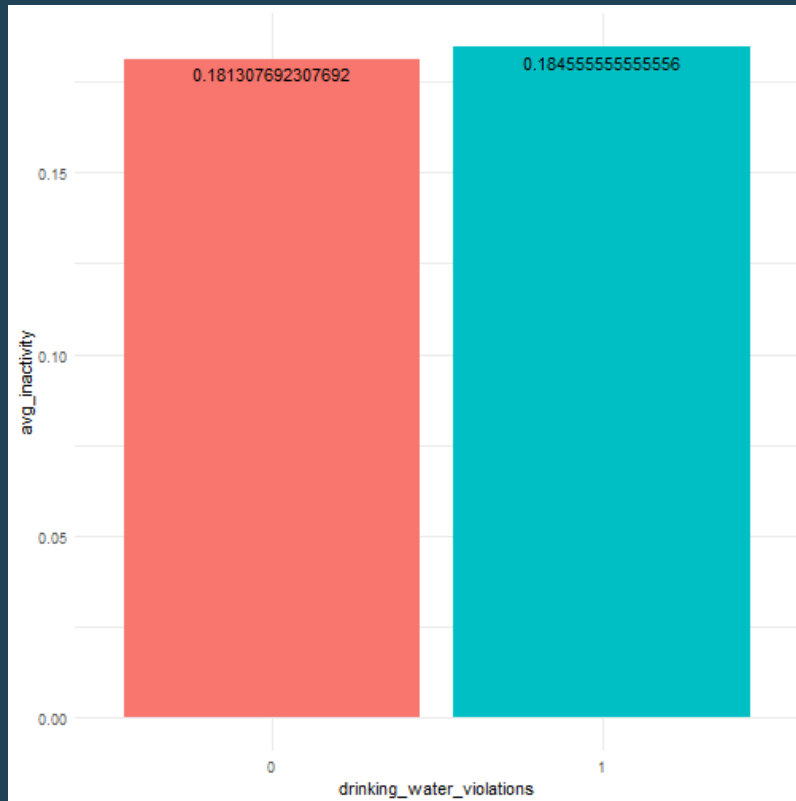
```
##      mean    sd median  min   max skew kurtosis
## X1 0.18 0.03   0.18 0.12 0.26 0.23   0.07
```

## 1. Communicate Data

```
##  
##      Two Sample t-test  
##  
## data:  physical_inactivity by drinking_water_violations  
## t = -0.40336, df = 56, p-value = 0.6882  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
##  -0.01937784  0.01288211  
## sample estimates:  
## mean in group 0 mean in group 1  
##      0.1813077      0.1845556
```

## 1. Interpret Data

"Based on the findings from the statistical test, there were no differences in physical inactivity between counties that had drinking water violations and the counties that did not have drinking water violations.



# Differences in Population & Sample

- **Population** is a large group of individuals, which a law of nature applies
- **Sample** is a group of a given population intended to represent that population
- **Participants** are those measured in a sample. Participants always refers to individuals (e.g., students, children, prisoners)
- A sample is *supposed to generalize* to a given population
- We never use subjects in the social sciences anymore, they are participants



# Real World Examples

Tell me whether or not the examples are a population or a sample. If it is a sample then tell me whether or not it is a representative sample.

1. All adults in the United States
2. All students in PSY 3307-04
3. All female students in PSY representing all students in PSY
4. All adults in California representing adults in the United States
5. All students in PSY 3307-06 representing students taking PSY 3307

# Variables & Relationships

- **Variable** is anything that can produce different scores.
- **Qualitative Variable** is a variable that represents a quality of category
- **Quantitative Variable** is a variable measured in amounts
- Qualitative = Categorical
- Quantitative = Numeric

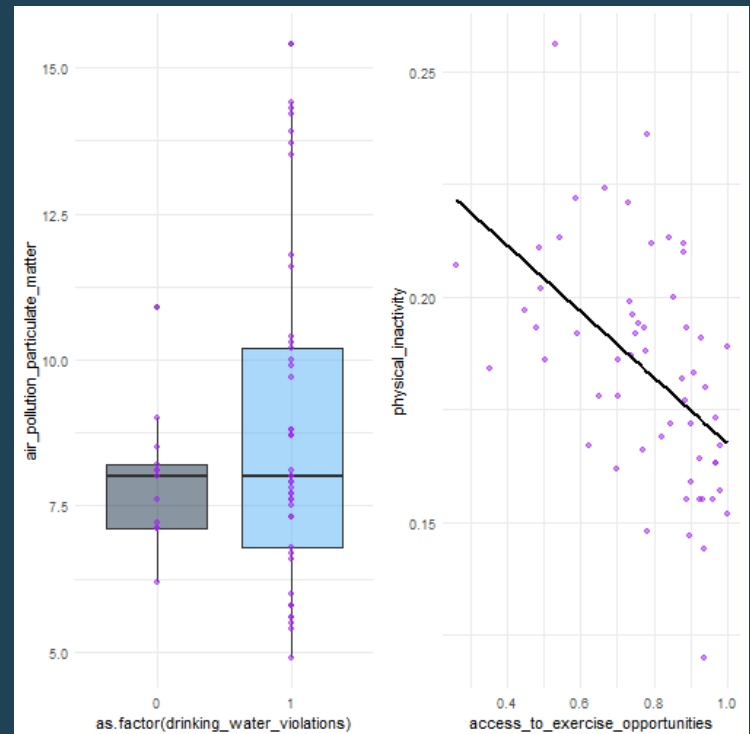
# Qualitative/Categorical or Quantitative/Numeric

1. Height in meters
2. Name
3. Race
4. Gender
5. Test Score
6. Cups of Coffee Drank Today
7. Full Address
8. Number of times I've said "Um"

# Relationships

- **Relationships** are when two variables are related; one variable changes the other changes consistently
- Often use the term *association* when talking about relationships between two variables

```
## `geom_smooth()` using formula 'y ~ x'
```



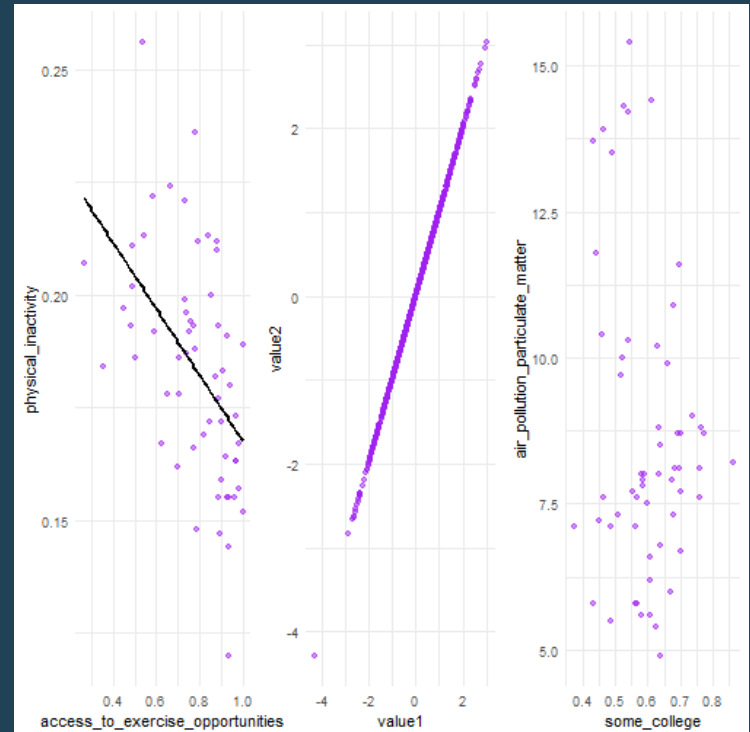
# Look at the actual correlation between the two variables

```
##  
##      Pearson's product-moment correlation  
##  
## data:  data$access_to_exercise_opportunities and data$physical_inactivity  
## t = -4.441, df = 56, p-value = 0.00004264  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
##  -0.6791215 -0.2903227  
## sample estimates:  
##      cor  
## -0.510348
```

# Three Types of Relationships

1. Positive - one variable goes up, the second variable also goes up
2. Negative - one variable goes down, the second variable goes up (or the other way around too)
3. No relationship - random data points everywhere

```
## `geom_smooth()` using formula 'y ~ x'
```



# Descriptive & Inferential Statistics

- **Descriptive Statistics** used for organizing and summarizing your data.
- Tells the reader how the data looks.
- **Inferential Statistics** used to estimate or infer if there is a relationship in your sample data that represents the population studying
- Tells the reader if there is a relationship.

# Examples of Descriptive & Inferential Statistics

1. The average amount of physical inactivity in California counties
2. Finding the populations of each California county
3. The association between parks and physical activity engaged
4. The *difference* in the number of liquor stores between San Diego and Los Angeles counties.
5. The number of males and females in PSY 3307.



# Statistics vs Parameters

- **Statistics** use English letters to get values for a sample
- **Parameter** use greek letters for values of a population

Statistic = Sample

Parameter = Population

# Learning Greek & Roman

$\Sigma = \textit{Sum}$

$\bar{x} = \textit{sample mean}$

$\mu = \textit{population mean}$

$s^2 = \textit{variance}$

$s = \textit{standard deviation}$

$\sigma = \textit{population standard deviation}$

$\sigma^2 = \textit{population variance}$

$r = \text{sample correlation coefficient}$

$\rho = \text{population correlation coefficient}$

$N = \text{Total Observations (Either Sample or Population)}$

$f = \text{Frequency}$

$n = \text{observations for categories in survey}$

# Designs

- **Design** is how a study is organized
1. Experimental (manipulate your independent variable to see what effect it has on the dependent variable)
    - True Experiment
    - Quasi Experiment
    - Natural Experiment
  2. Correlational (don't manipulate anything; just look at the relationship between two variables)
    - Correlation
    - Regression

# Types of Variables

- **Independent Variable (IV)** variable that is manipulated by experimenter to see if causes change in DV
- **Dependent Variable (DV)** behavior that is measured; expected to be influenced by IV; outcome you are interested in
- Let's keep it simple Correlational Design: IV --> DV Experimental Design: Manipulate IV to see if DV changes

# Conditions

- **Conditions** are the categories of the IV to measure DV scores
- Example: Examining test scores of students in a class. Half of the class gets a study guide while the other half is given nothing.

IV = Study guide DV = Test scores Conditions: 1 = Group receiving study guide; 2 = Group receiving nothing (control group)

- For experiments, you can say "IV had an effect on DV" or the "IV caused DV" but for correlational students that is a big NO NO. Every time JP hears/reads a correlational study that describes the effect of IV on DV he cries a little inside.



**NAME THAT VARIABLE/DESIGN!**

# Choose Design, IV, & DV

1. Two randomly chosen groups are created to see who does better on an exam, with one group getting tutoring lessons.
2. A drug trial where two groups are randomly selected; one group gets a placebo and the other gets the drug to see if there are differences in blood pressure.
3. Examining the association between getting a COVID vaccine and being hospitalized for COVID.
4. Giving male and female college students an exam and telling half of the male and female students that they will probably fail to examine the effect of stereotype threat
5. Examining differences between races/ethnicities and physical activity levels.
6. Put college male students in a "prison" to see how guards maintain law and order in a prison setting.
7. Randomly choosing three groups of adults to see their physical activity habits; one group gets a personal trainer, one group gets a workout buddy, and one group gets nothing.



# 4 Types of Measurement Scales

## 1. Nominal

- Categorize individuals (names, race/ethnicity, gender, sex, whether or not you can juggle), categories that are not ordered

## 1. Ordinal

- scores indicate rank order (race order, podiums), has categories that are ordered

## 1. Interval

- scores measure amounts, but zero is not an absolute zero (smallest number possible), has negative numbers, equal amounts separate adjacent numbers

## 1. Ratio

- has a true zero (zero is smallest number), no negative numbers



NAME THAT SCALE!

1. California Counties
2. Temperature measured in Celcius 2b. Temperature measured in Kelvin
3. Olympic Medal Winners (Gold, Silver, Bronze)
4. Money
5. Income Brackets [0-10k, 11-20k, 21-30k, ...]
6. Miles per hour (mpg)
7. 10 most common baby names
8. Grades (A, B, C, D, F)
9. Grades (100-95, 94-90, 89-85, 84-80, ...)
10. Grades (100-0%)

# Continuous & Discrete Variables

- **Continuous** variable that can be measured with decimals.
- Example: you can get a 94.3 on an exam.
- **Discrete** can be categorical (sex, race/ethnicity);, does not allow fractional amounts
- Example: Being a 4th year compared to a 3rd year