

# Applied Statistical Analysis

EDUC 6050

Week 2

Finding clarity using data

# Today

1. Working with Data
2. Overview of Statistics
3. Intro to Statistical Terminology
4. Intro to Jamovi

**Reading**

# **Data in Spreadsheets**

# Data in Spreadsheets

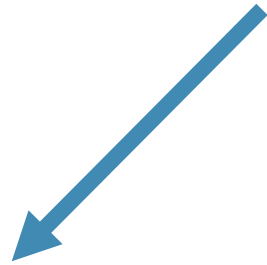
- 2 Be **Consistent**
- 3 Choose **good names** for things
- 4 Write dates as YYYY-MM-DD
- 6 Put just **one thing in a cell**
- 7 Make it a **rectangle**
- 8 Create a **data dictionary**

# Data and Statistics

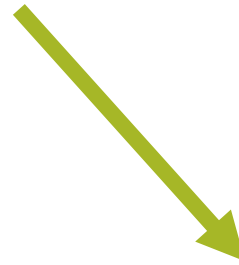
**Statistics** helps us understand  
our data

# Data and Statistics

**Statistics** helps us understand  
our data



Summarize the  
data easily

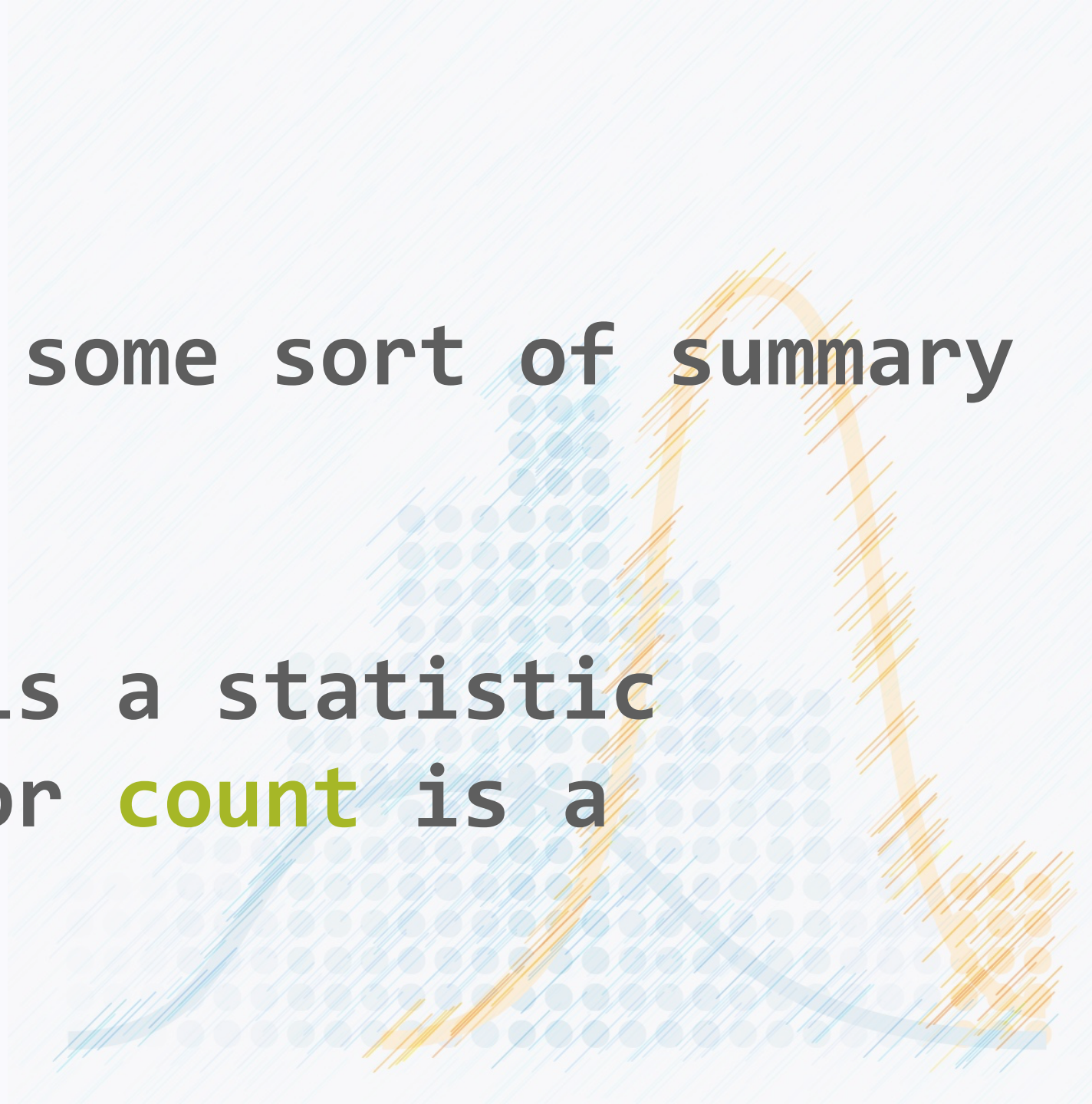


Ask questions about  
what the data mean

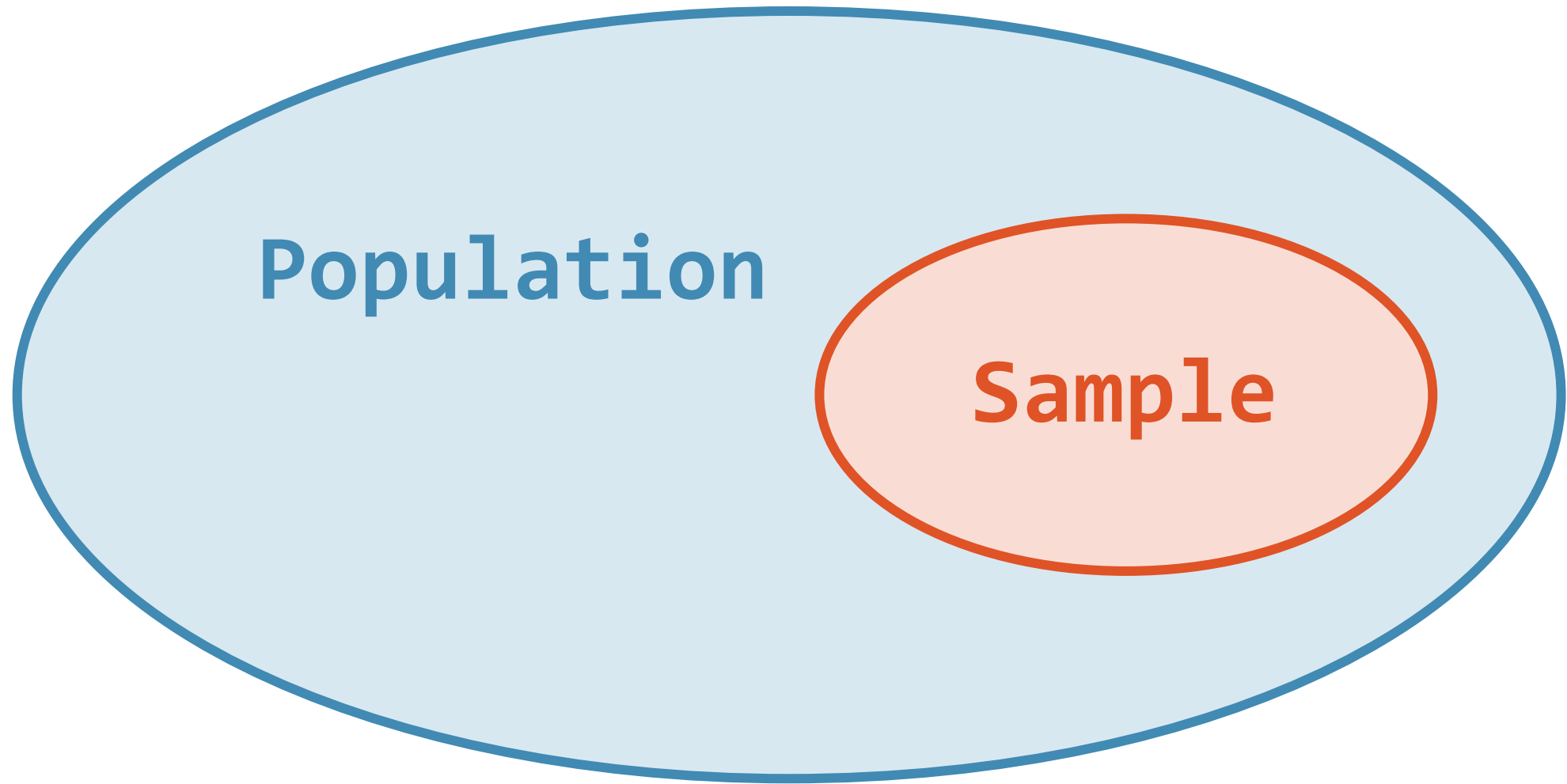
# Statistics

A **statistic** is some sort of summary of the data

- The **average** is a statistic
- A **frequency** or **count** is a statistics



# The Vocabulary of Statistics





# The Vocabulary of Statistics

## Descriptive Statistics

Describing the data that  
you have (your sample)

## Inferential Statistics

Understanding what your data  
say about the population

# The Vocabulary of Statistics

**Independent  
Variables**



**Dependent  
Variables**

“predictors” or “IV”

These are the variables  
that we think are  
causing or influencing  
the outcome

“outcomes” or “DV”

These are the variables  
that we think are caused  
by an independent  
variable

# The Vocabulary of Statistics

## Hypothesis Testing (Inferential Statistics)

“Null Hypothesis Significance Testing”

Gives us an idea about what the population may look like based on our sample (accounts for **sampling error**) = “significance”

# The Vocabulary of Statistics

## Hypothesis Testing (Inferential Statistics)

“Null Hypothesis Significance Testing”

## Effect Sizes

“Magnitude of the effect”

Tells us how big the effect is = “meaningfulness”

**Break Time**

# Why Learn Statistics?

It is the **language** of understanding data (and data are everywhere)

- Allows you to **complete your thesis!**
- Helps you **communicate** with other data people you work with
- Gives you power to **convince stakeholders with evidence**
- Opens up **job** opportunities

# Scales of Measurement

"The way a variable is measured determines the kinds of statistical procedures that can be used" (pg 10)

Want measures that:

1. Are reliable
2. Are valid
3. Are meaningful
4. Have a high degree of information

# Scales of Measurement

## 4 General Types (see pg. 11)

<i>Scale</i>	<i>Definition</i>	<i>What the scale allows you to do</i>
<b>Nominal</b>	Categories based on qualitative similarity (no order to the categories)	Count the number of things in the categories
<b>Ordinal</b>	Like nominal, but the categories can be ranked	Count and rank the number of things in each category
<b>Interval</b>	Quantify how much of something	Count, rank, and quantify how much of something there is (zero does not mean there's nothing)
<b>Ratio</b>	Quantify how much of something (zero means there is none of that thing)	Count, rank, and quantify how much of something there is with a meaningful zero



# Scales of Measurement

## 4 General Types (see pg. 11)

<i>Scale</i>	<i>Definition</i>	<i>What the scale allows you to do</i>
<b>Nominal</b>	Categories based on qualitative similarity (no order to the categories)	Count the number of things in the categories
<b>Ordinal</b>	Like nominal, but the categories can be ranked	Count and rank the number of things in each category
<b>Interval</b>	Quantify how much of something	Count, rank, and quantify how much of something there is (zero does not mean there's nothing)
<b>Ratio</b>	Quantify how much of something (zero means there is none of that thing)	Count, rank, and quantify how much of something there is with a meaningful zero

# Scales of Measurement

4 General Types (see pg. 11)

Scale	Definition	What the scale allows you to do
Nominal	Categories based on qualitative similarity (no order to the categories)	Count the number of things in the categories
Ordinal	Like nominal, but the categories can be ranked	Count, rank, and quantify how much of something there is in each category
Interval	Quantify how much of something	Count, rank, and quantify how much of something there is (zero does not mean there's nothing)
Ratio	Quantify how much of something (zero means there is none of that thing)	Count, rank, and quantify how much of something there is with a meaningful zero

Increasing degree of information

# Scales of Measurement

4 General Types (see pg. 11)

Scale	Definition	What the scale allows you to do
Nominal	Categories based on qualitative similarity (no order to the categories)	Count the number of things in the categories
Ordinal	Like nominal, but the categories can be ranked	Count and rank the number of things in each category
Interval	Quantify how much of something	Count, rank, and quantify how much of something there is (zero does not mean there's nothing)
Ratio	Quantify how much of something (zero means there is none of that thing)	Count, rank, and quantify how much of something there is with a meaningful zero

Let's think of some examples of each type

# Scales of Measurement

These lie on a spectrum from qualitative to quantitative



# Scales of Measurement

## Discrete

Cannot be broken  
down into smaller  
units

*Number of siblings,  
racial groups, have the  
disease or not*

## Continuous

Can be broken into  
smaller units

*Time to finish an exam,  
height of a person*

**Break Time**

# Graphing Data

A VERY IMPORTANT part of data analysis

It is useful for both:

1. Understanding patterns in the data
2. Communicating results in a much more meaningful way

Takes some practice

# Some Types of Data Graphics

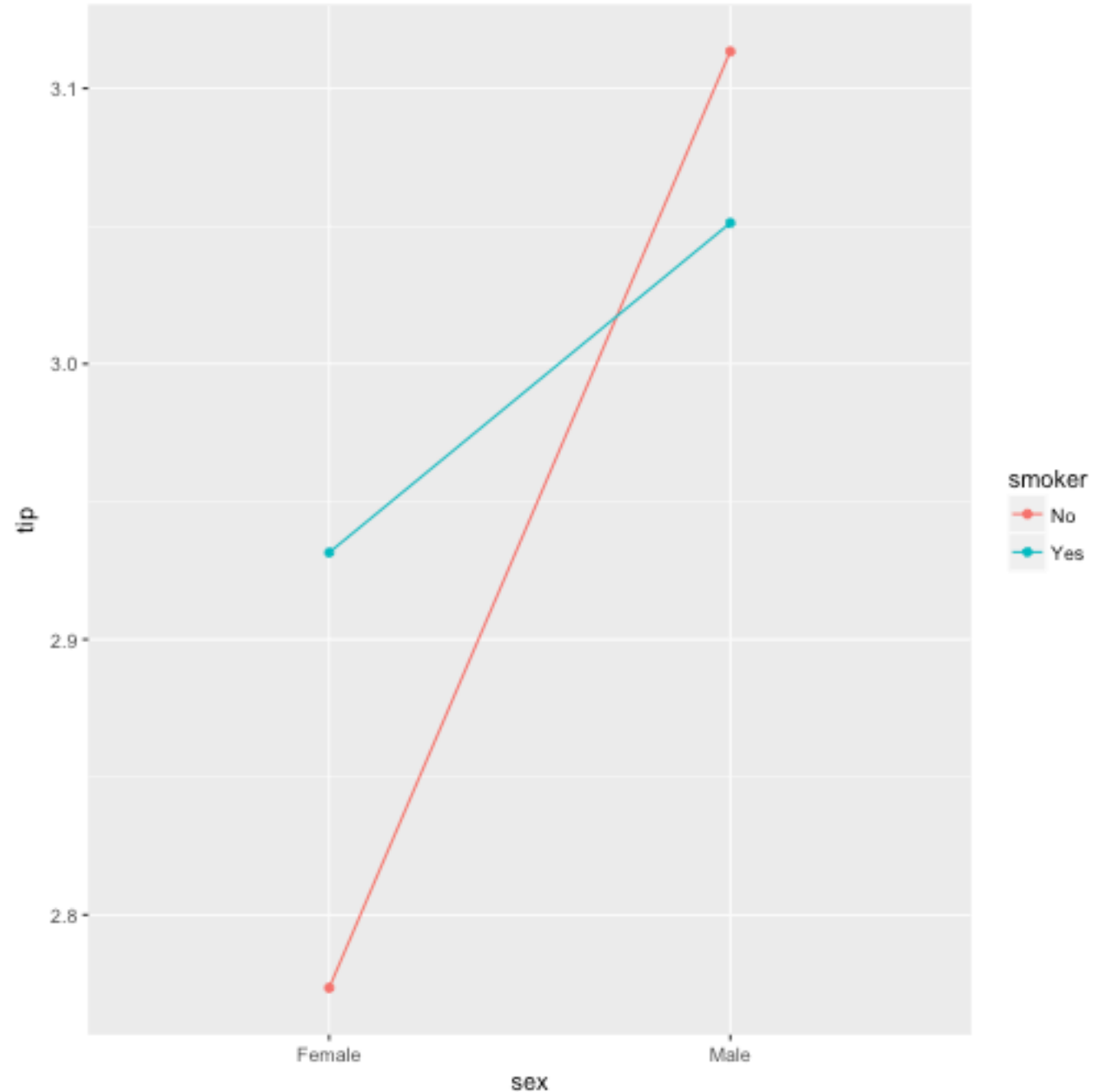
Each provide different insights into the data

1. Line Graphs
2. Bar Graphs and Histograms
3. Scatterplots
4. Boxplots



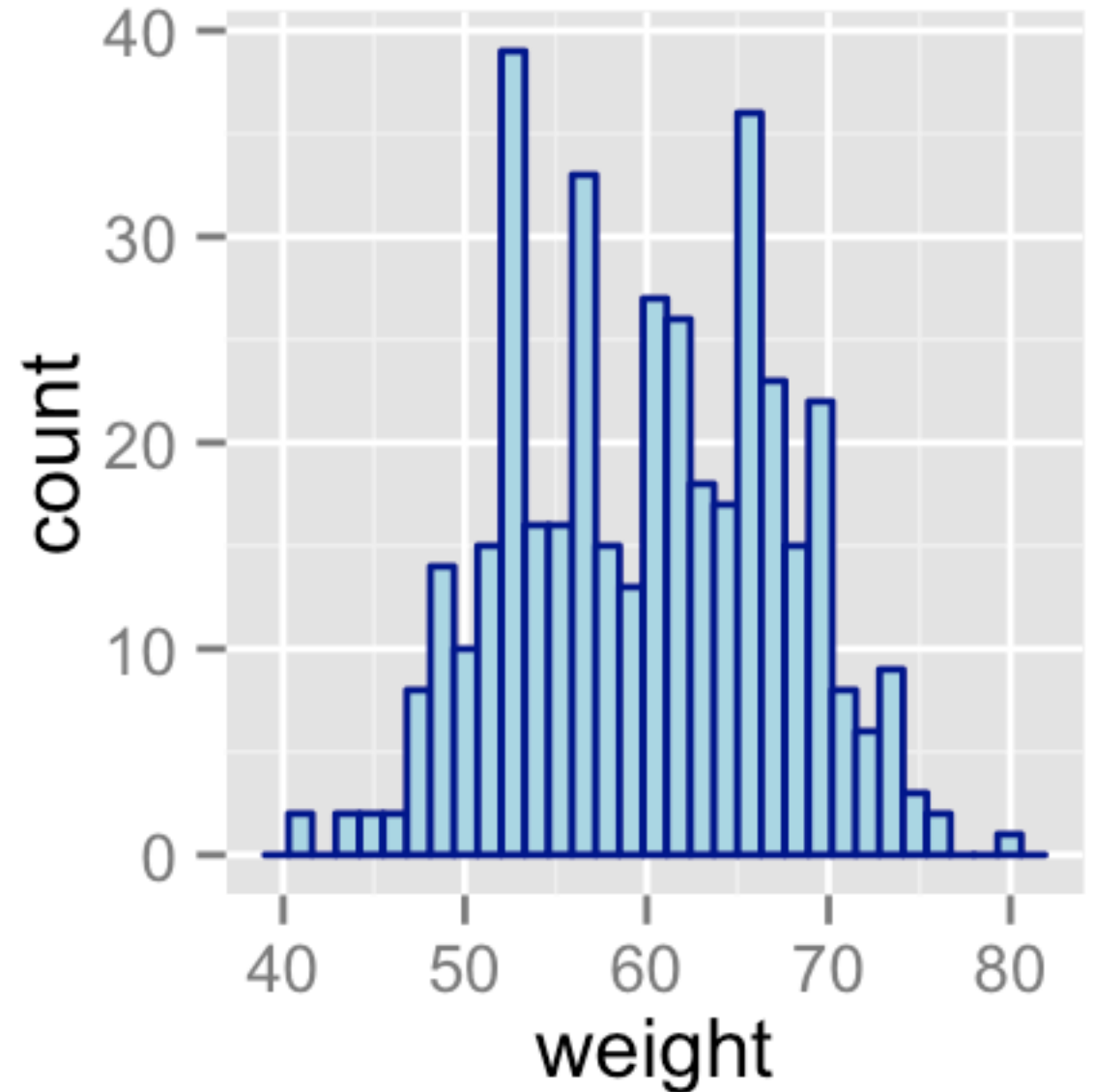
# Line Graphs

Generally shows trends and patterns across groups



# Bar Graphs and Histograms

These help us understand distributions and frequencies



# Bar Graphs and Histograms

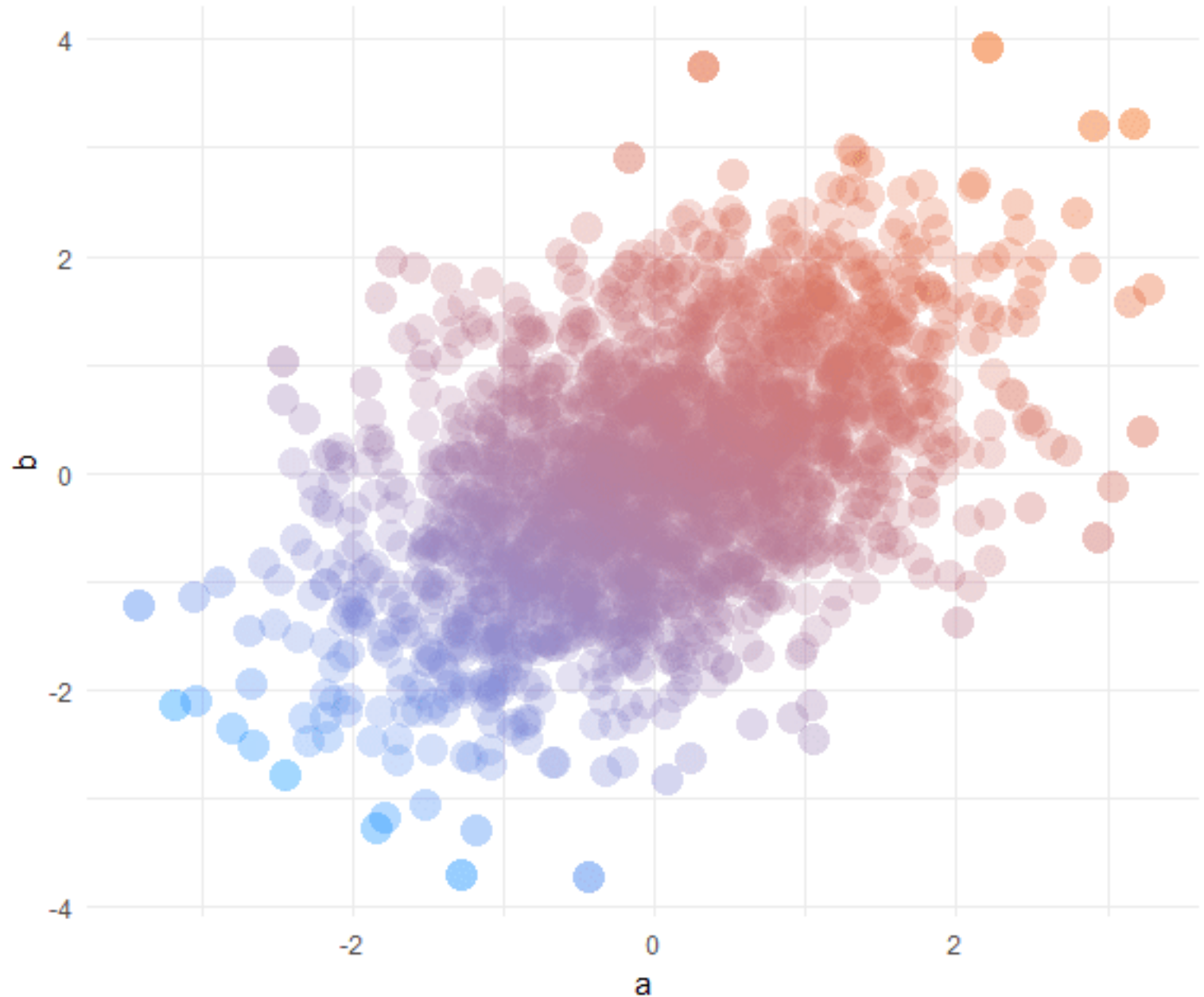
These help us  
understand  
distributions and  
frequencies



Symmetric vs. Asymmetric  
Unimodal vs. Multimodal  
Short-tailed vs. long-tailed

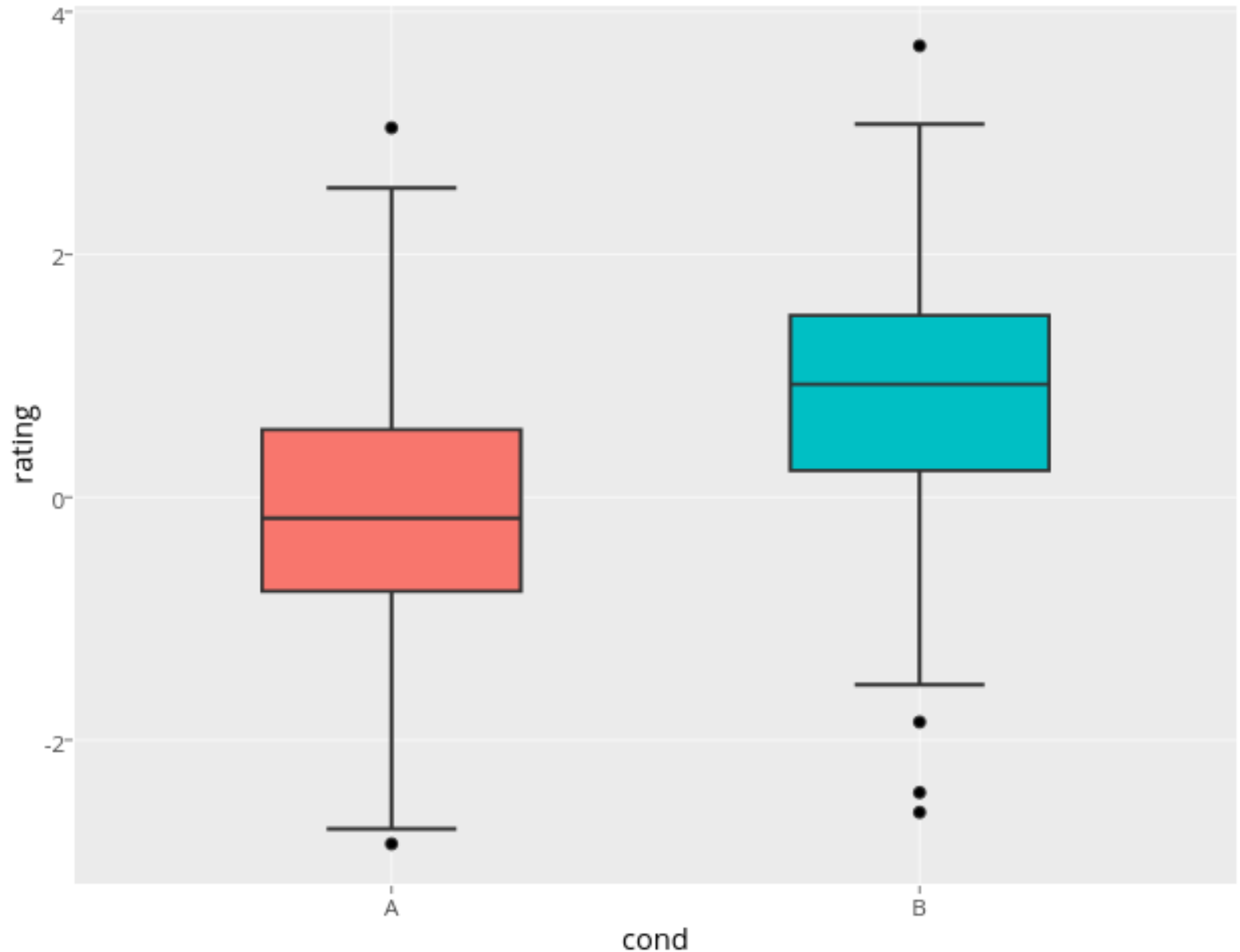
# Scatterplots

Show us how  
two (or more)  
variables are  
related



# Boxplots

Show us the range and where most values are for a variable (usually across groups)



# Frequency Tables

Tables can also be very valuable to understand patterns in the data

Level	Frequency	Percent	Cumulative Percent
A	10	25.0%	25.0%
B	5	12.5%	37.5%
C	20	50.0%	87.5%
D	5	12.5%	100%

Questions?

# Next week:

1. Statistics terminology (Hypothesis, IV and DV, Measurement, Validity and Reliability, Correlation and Experimentation, Distributions, Central Tendency and Variability)
2. Chapters 1, 2, and 3 in Book
3. Start looking for articles