

# Class 08

## DATA1220-55, Fall 2024

Sarah E. Grabinski

2024-09-16

## Homework 2

- ▶ Describing numerical distributions: modality, skew, outliers
- ▶ Describing numerical distributions: appropriate summary statistics
- ▶ Matching numerical distributions to their summary statistics, reading a boxplot
- ▶ Calculating proportions from a contingency table

## Homework 2

- ▶ Instructions (`homework2_instructions.pdf`), a Quarto markdown template (`homework2_template.qmd`), and an example HTML output (`homework2_example.html`) are available for download under Chapter 2 on the Modules page in Canvas.
- ▶ Upload **TWO** (2) documents to Homework 2 on the Assignments page in Canvas by **Friday 9/20/2024** by **6:00pm**: `homework2_yourlastname.qmd` and `homework2_yourlastname.html`
- ▶ Video walk-through of Homework 2 under Tutorials on the Modules page in Canvas. Make sure you're caught up on the video walk-through of homework 1.

# Late Policy

“This homework is due by 6:00pm on Friday, 9/20/24. No credit will be lost for assignments received by 7:00pm to account for issues with uploading. 10% of the points will be deducted from assignments received by 9:00am on Saturday, 9/21/24.

Assignments turned in after this point are only eligible for 50% credit, so it benefits you to turn in whatever you have completed by the due date.”

# How can I get help with homework?

- ▶ ***Read the textbook.*** Many of you are asking for additional examples. Luckily, there are tons we didn't go over in the textbook.
- ▶ ***Look at the homework early.*** I can see in Canvas that many students didn't download the documents until 1-2 days before it was due. That's not a lot of time to get help.
- ▶ ***Ask a question on our Campuswire class feed.*** I'm only one person, and I may not be able to give you a prompt answer. However, the 20+ other people in the class might be able to.
- ▶ ***Come to office hours.*** I will be available after class today (Monday 9/23/2024) and Wednesday 9/25/2024 from 2:30pm - 4:00pm. If you cannot make it, reach out to me to try and schedule an appointment.

## Last Time...

- ▶ Contingency tables: counts and proportions (frequencies)
- ▶ Visualizing frequencies: bar plots, mosaic plots
- ▶ Describing numerical relationships: linear vs nonlinear, strong vs weak
- ▶ Visualizing 3+ variables

# Chapter 3 Objectives

- ▶ Define probability, random processes, and the law of large numbers
- ▶ Describe the sample space for disjoint and non-disjoint outcomes
- ▶ Calculate probabilities using the General Addition and Multiplication Rules
- ▶ Create a probability distribution for disjoint outcomes

# Defining Probability

What does the word ***probability*** mean to you?



# Defining Probability

What does the word ***probability*** mean to you?

*“Highly likely”*

# Defining Probability

What does the word ***probability*** mean to you?

*“Highly likely”*

*“Probably”*

# Defining Probability

What does the word ***probability*** mean to you?

*“Highly likely”*

*“Probably”*

*“About even”*

# Defining Probability

What does the word ***probability*** mean to you?

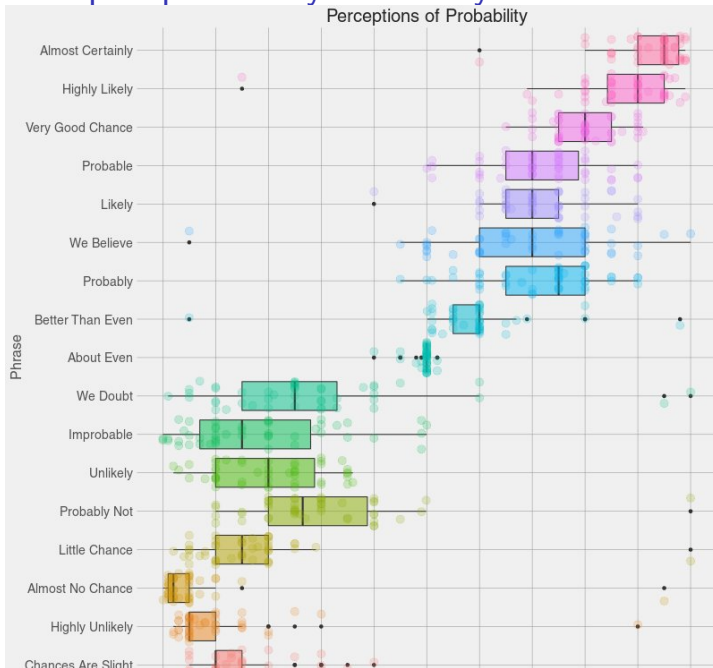
*“Highly likely”*

*“Probably”*

*“About even”*

*“Almost no chance”*

# People interpret probability differently



# So what is probability?

## ! Frequentist Definition

The proportion of times that a particular outcome would occur if we observed a random process an infinite number of times.

▶ A **random process** is one where you know which outcomes are possible (i.e. the **sample space**) but you don't know which outcome comes next

▶ *Examples: coin toss, die roll, stock market*