

# MA213 L2

## L2 : Plotting and Summaries

---

OpenIntro Statistics, 4th Edition

Slides developed by Mine Çetinkaya-Rundel of OpenIntro.

The slides may be copied, edited, and/or shared via the CC BY-SA license.

Some images may be included under fair use guidelines (educational purposes).

## L2 : Plotting and Summaries

- *Previously:* What we learend from last time
- *This time:* What we are learing today
- *Deadlines/Announcements:* Quiz / Exercise / Explain / after work?

# Lesson Plan

- xx min Exploring Data, what type of information...?
- xx min Some basic built-in R plot functions to show data...  
how to summarize?
- xx min Scatterplot, Scatterplots are useful for visualizing the relationship between two numerical variables.
- xx min data manipulation – dplyr syntax?
- xx min ggplot2 packages and
- xx min Data explore by groups

## Learning Objectives

- Use R for Data Management and Exploration: Utilize R to load, pre-process, and explore data through visualization and summarization techniques. [Q1, L1, L2]
- Classify and Analyze Variables: Categorize variables based on their types (e.g., numerical/categorical, continuous/discrete, ordinal), assess their association (positive, negative, or independent), and determine which make sense as explanatory vs. response variables. [Q1, L2]

Core

# Numerical Data

- Scatterplots are useful for visualizing the relationship between two numerical variables.
- Dot plots; Useful for visualizing one numerical variable.
- Sample mean
- Stacked dot plot
- Histograms; and talk about Bin width
- Shape of a distribution : modality (unimodal, bimodal/multimodal, uniform)
- Shape of a distribution : skewness
- Outliers
- Variance and standard deviation
- Median
- Q1, Q3, IQR
- Box plot
-

## Categorical Data

- Contingency Tables; A table that summarizes data for two categorical variables is called a contingency table.
- Bar plot
- Pie charts
- Simulation ??