

# **STAT 3011 Discussion 007**

**Week 3**

Talha Hamza  
University of Minnesota

**Spring 2026**

# Lab 3: R Cheat Sheet

---

## Sampling:

- `sample(x, size)` Simple random sample from x
- `set.seed(#)` Make sampling reproducible

## Summary Statistics:

- `mean(x)` Sample mean
- `sd(x)` Sample standard deviation
- `na.rm = TRUE` Ignore missing values

## Tables:

- `table(CategoricalVariable)` Frequency table

# Lab 3: R Cheat Sheet Continued

---

## Indexing / Subsetting:

- `data[rows, columns]` Select rows and columns
- `data$Variable` Select one column
- `c()` Multiple indices
- `:` Index range

## Visualizations:

- `hist(NumericalVariable)` Histogram
- `boxplot(NumericalVariable)` Boxplot
- `boxplot(Numerical ~ Categorical)` Side-by-side boxplots

## Useful:

- Add labels: `main=""`, `xlab=""`, `ylab=""`
- Larger samples give more stable statistics.
- Use plots to understand distribution shape.

# Questions?