

# Crafting with Data

Reality, Illusions, Truth & the Future

---

Instructor: Rob Faludi

# Plan for Today

---

- Discovery Seeker pilot results
- Exercise One: Families
- Data Gathering Redux
- Exploratory Data Analysis
- Readings & Assignments

Discovery Seeker: Pilot Data

# Exercise One: Families

---

- How many children in your immediate family?
- Average in US is about 2.0
- How do we compare?

# Present data gathering redux

---

- Remind us what data you looked at and what
- What changes did you make from your first pilot?
- How were the results affected?



# Exploratory Data Analysis

---

- many wells, many routes
- state of mind
- data as a source of new ideas
- skepticism for summaries
- openness to new patterns and views
- not by statistical summary alone, that's not open

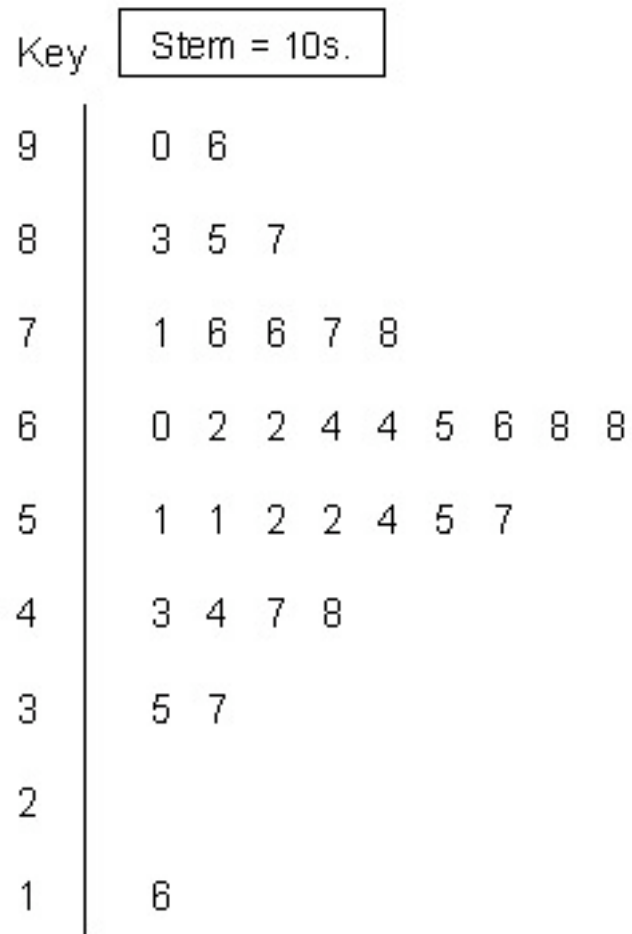
# EDA

---

- Smooth vs. Rough
- inductive methods
- re-expression, scales
- visualization
- resistant or robust statistics
- outliers, include or exclude?

# Visualization: Stem & Leaf

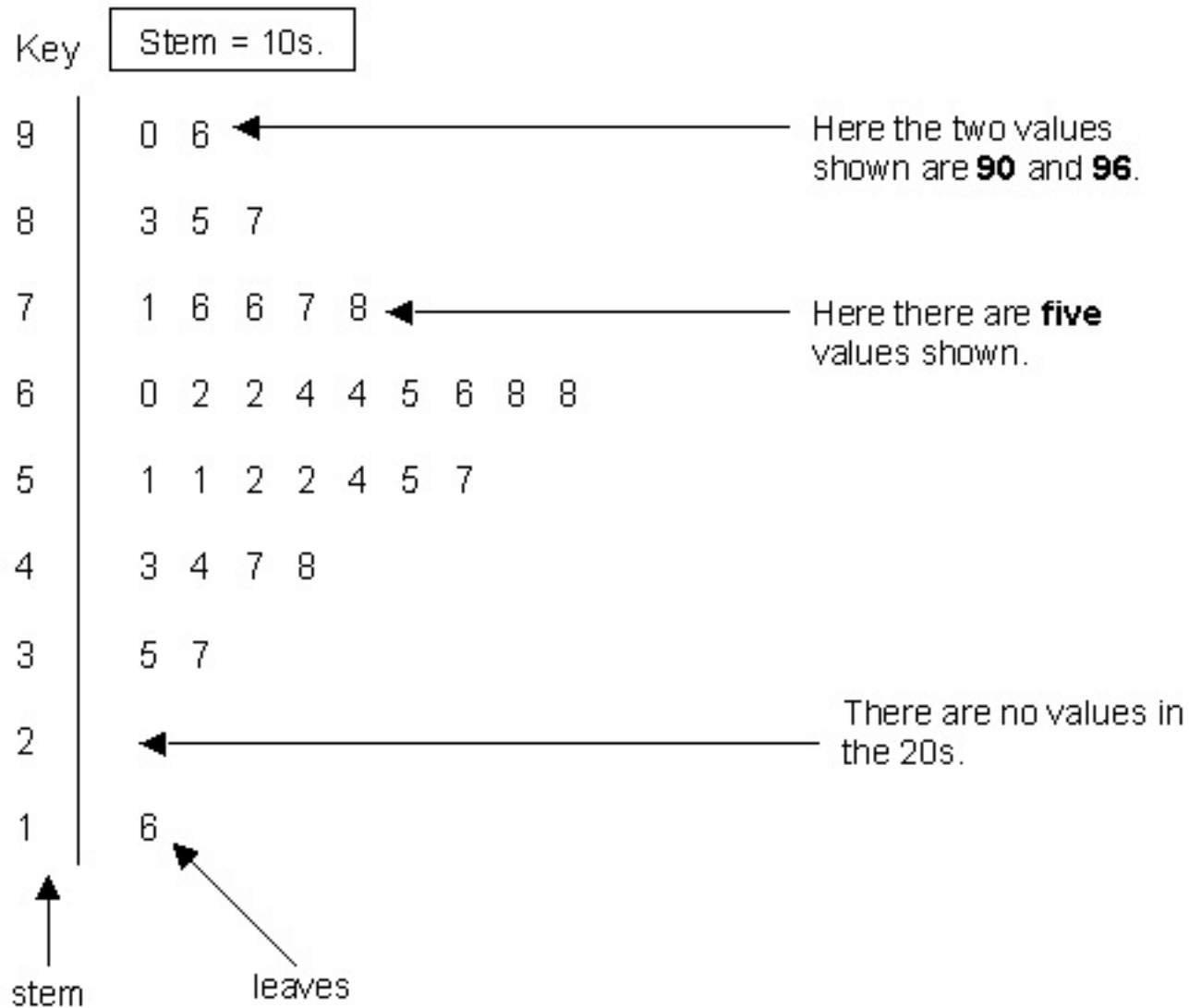
---





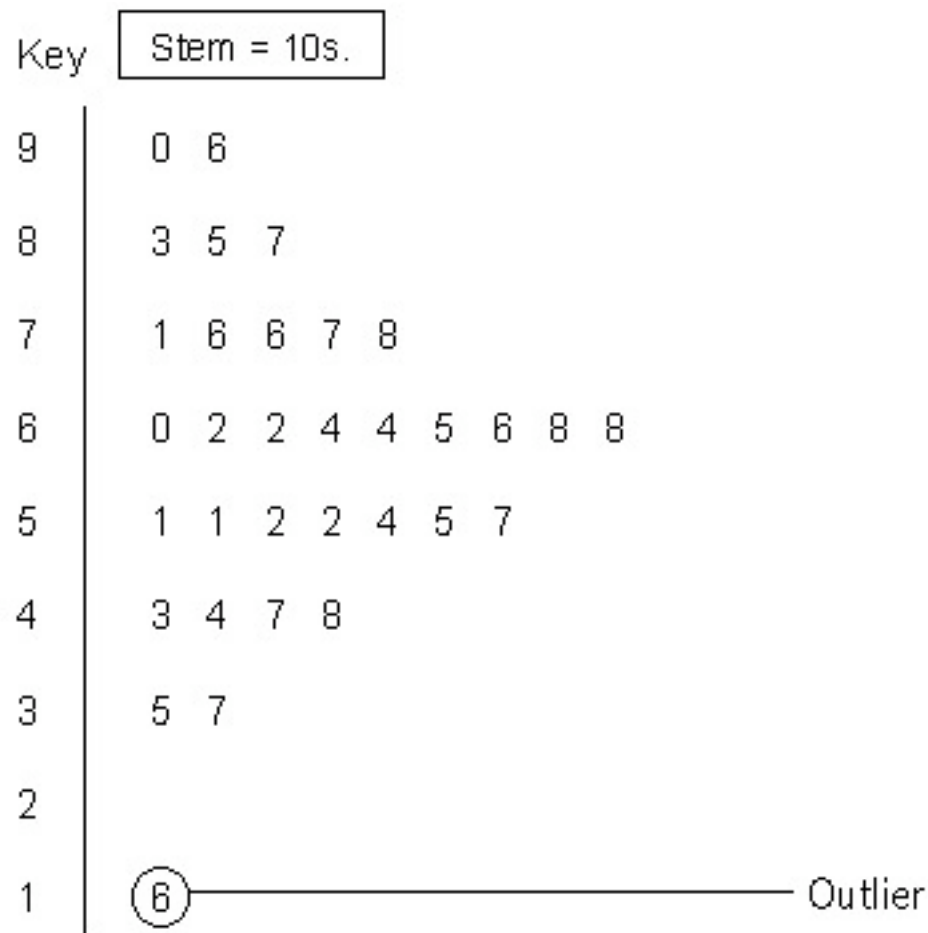
# Visualization: Stem & Leaf

---



# Visualization: Stem & Leaf

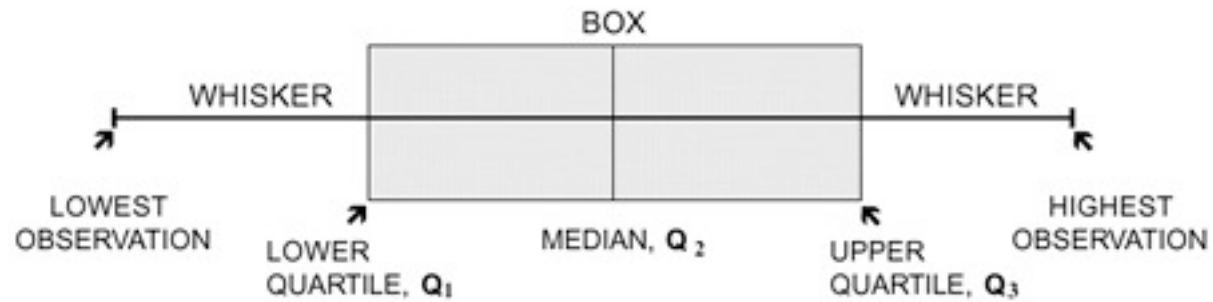
---



Total number of student marks = 33

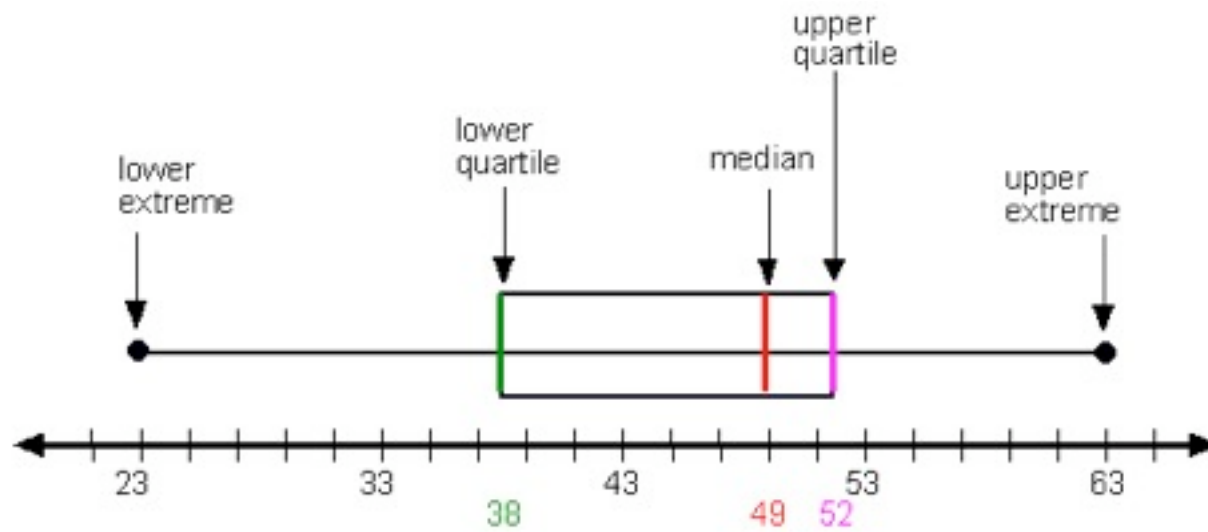
# Visualization: Box & Whiskers

---

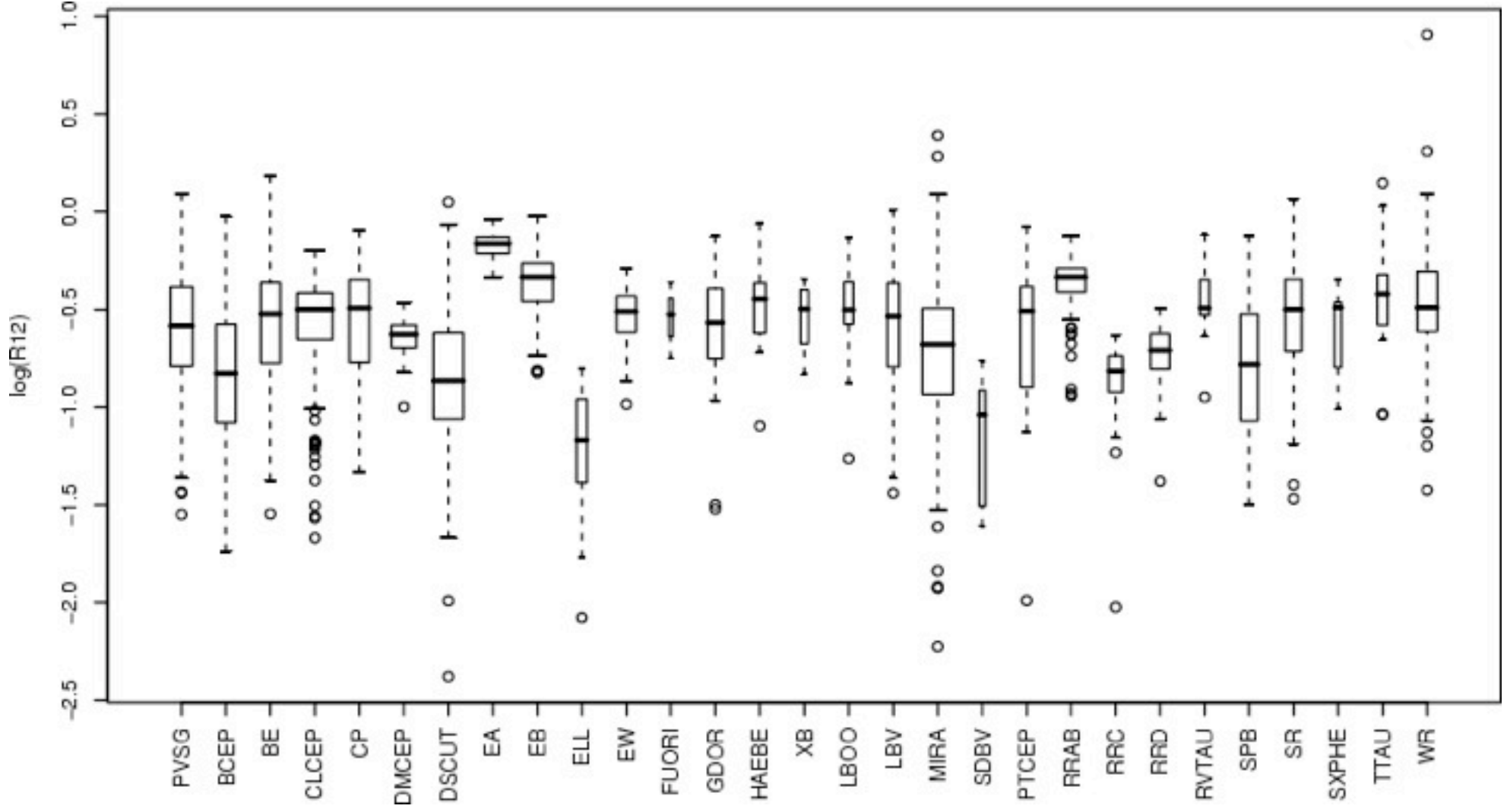


# Visualization: Box & Whiskers

---

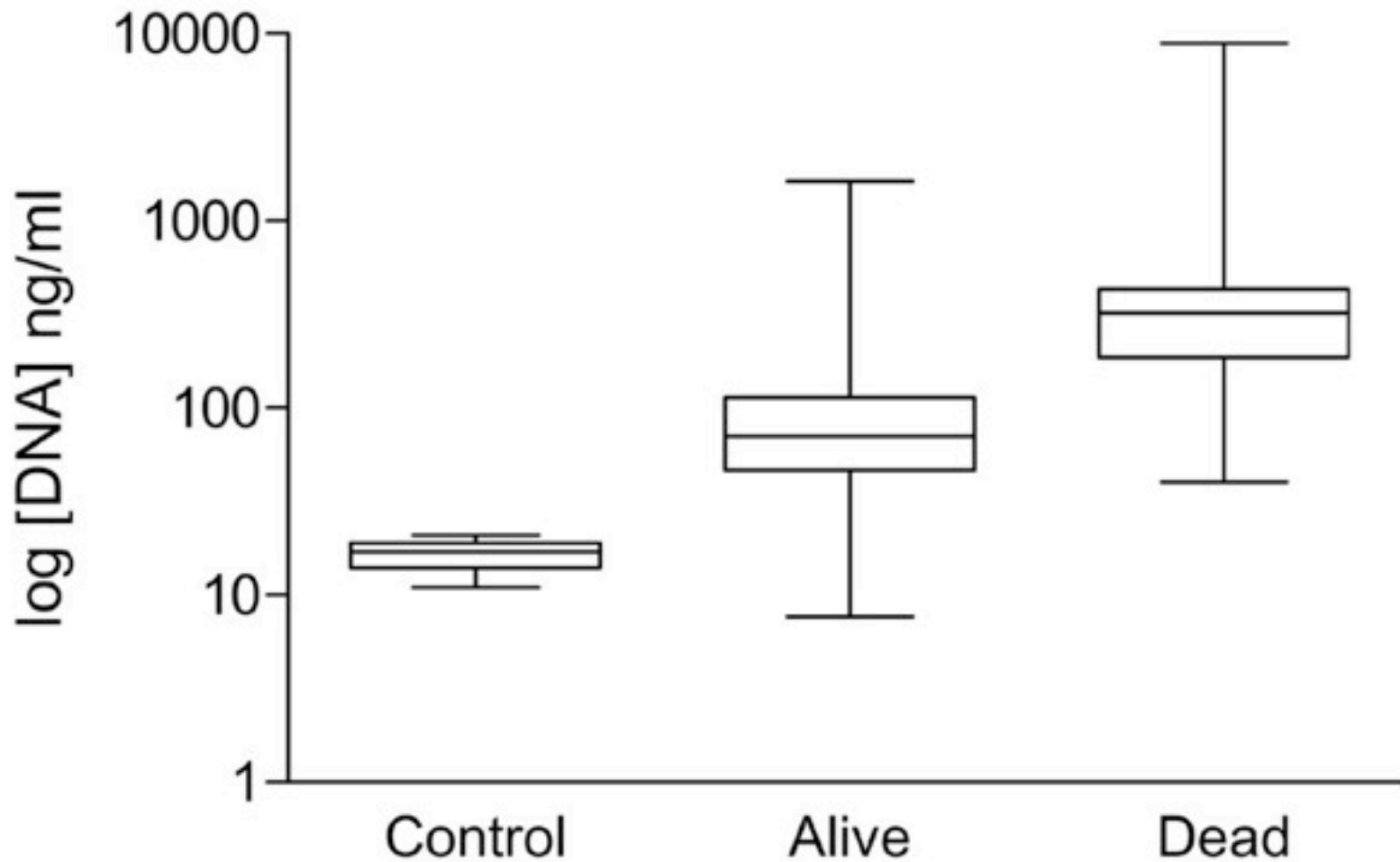


# Visualization: Box & Whiskers



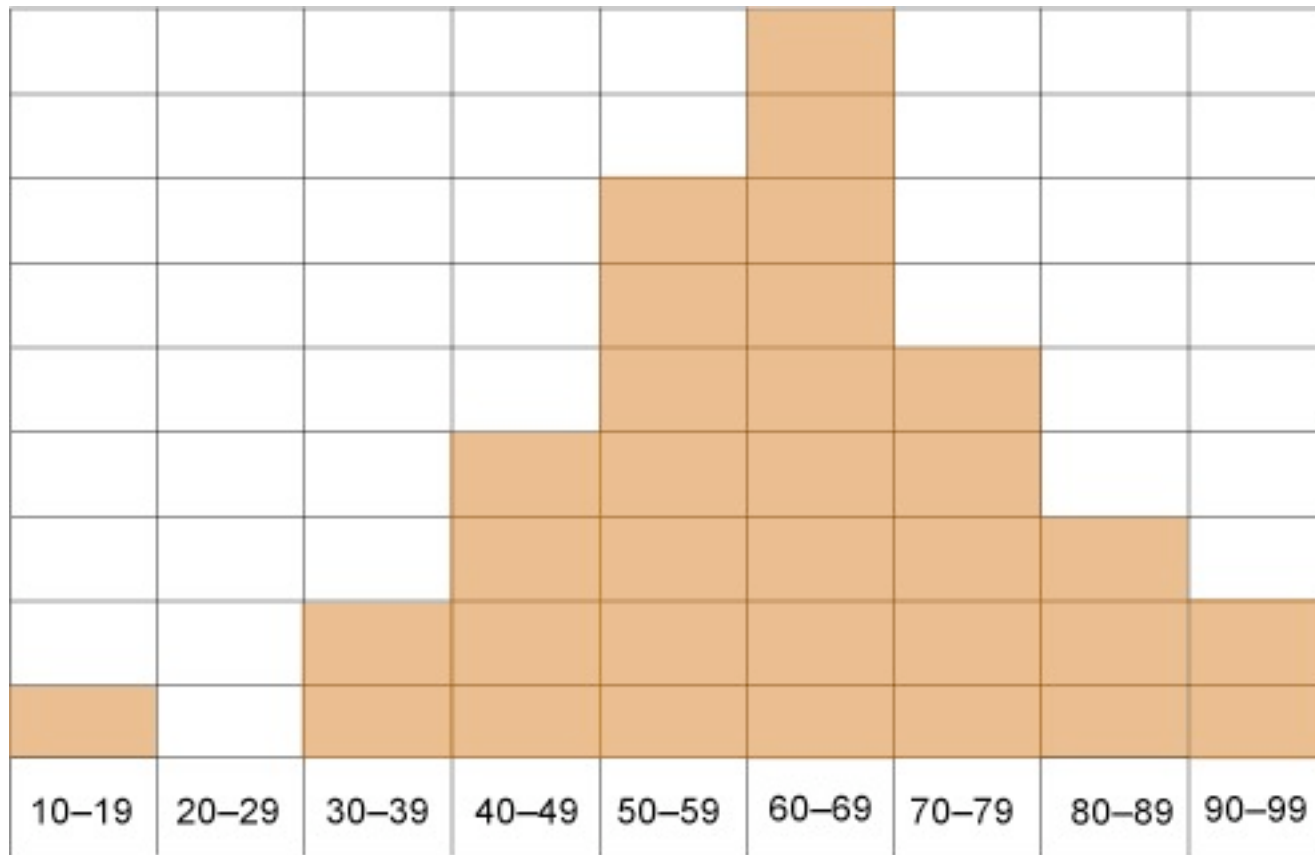
# Visualization: Box & Whiskers

---



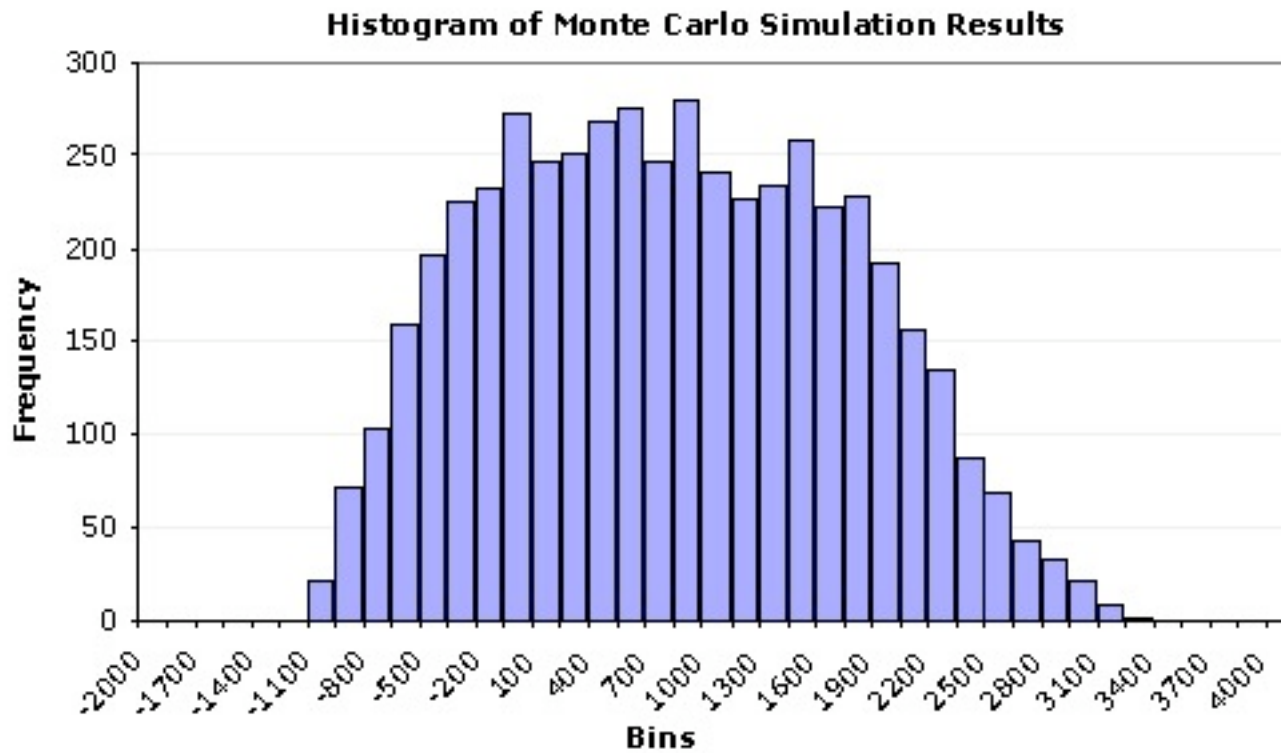
# Visualization: Histograms

---



# Visualization: Histograms

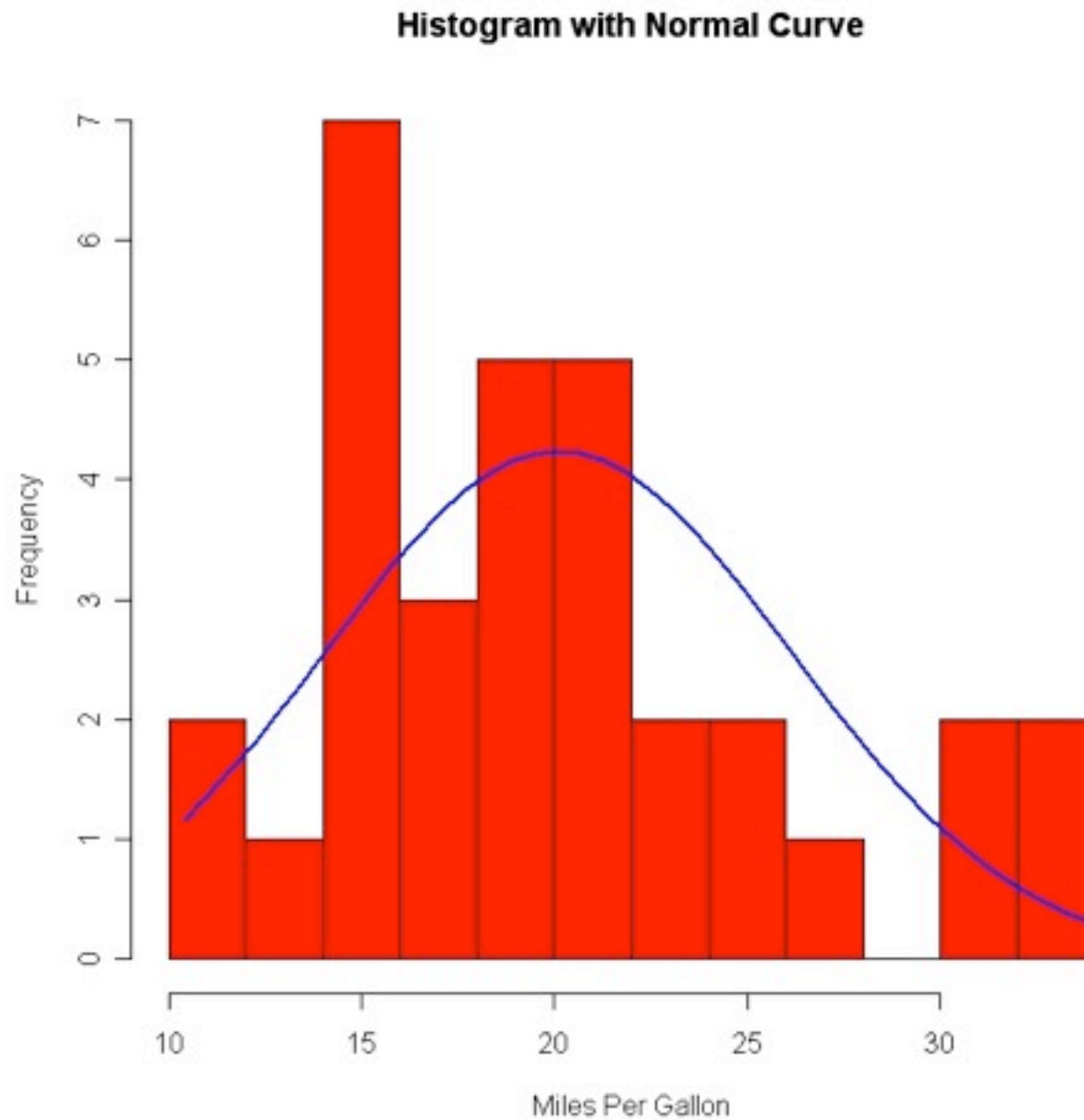
---





# Visualization: Histograms

---



# Readings and Assignments

---

- Readings

- The Canon, by Natalie Angrier — Chapter 2: *Probabilities*

- Assignments

- Discovery Seeker: improve your technique, gather more data and explore it. You'll share your improvements and exploration findings at the next critique.
- Toss a coin 100 times and write down the results OR imagine tossing a coin 100 times and write down the results. You'll be assigned to one task.