

# EDA Review: Describing Data Numerically

## Skills:

- Calculate numeric measures of center, spread and position
- Determine which measures of center and spread are most appropriate for a data set
- Determine if a data set contains outliers

# Measuring Center

---

---

Your choices are **mean** or **median**.

Both can be found via 1-Var Stats in the calculator.

Sometimes you might have to read computer output to find these...but the output will be well labeled and needs no special training...

---

---

---

---

---

---

---

---

---

---

# Measuring Spread

---

---

Your primary choices are **standard deviation** and **IQR**. *Range* and *variance* are also available.

Only standard deviation is given directly from the calculator. Be careful to choose the correct one!

---

---

---

---

---

---

---

---

---

---

# Measuring Position

---

---

Standardized scores (*z-scores*), quartiles and percentiles all do this.

Only quartiles are produced directly by the calculator.

---

---

---

---

---

---

---

---

---

---

---

You need to be able to interpret  
all of these measures in the  
context of a problem!

# Choosing Appropriate Measures

---

Look at a graph of the data.

- If the data are skewed/have outliers, then use the median and IQR.
- Otherwise, use the mean and standard deviation.

*This is because the mean and standard deviation are more affected by skewness/outliers.*

---

---

---

---

---

---

---

---

---

---

# Locating Outliers

---

---

Primary Method: Tukey's Rule

Lower Bound:  $Q_1 - (1.5)(IQR)$

Upper Bound:  $Q_3 + (1.5)(IQR)$

---

---

Data outside of these bounds are likely outliers.

---

---

---

---

---

---

---

---

---

---

# Locating Outliers

---

Primary Method: z-scores

Lower Bound:  $\bar{x} - 2 \cdot s$                        $\mu - 2 \cdot \sigma$

Upper Bound:  $\bar{x} + 2 \cdot s$                        $\mu + 2 \cdot \sigma$

---

Data outside of these bounds are likely outliers.

---

---

---

---

---

---

---

---

---

---



# Example

46	51	44	50	33	46	60	41	55	46
42	44	50	54	46	41	48	53	53	

The percentage of juice lost after thawing for 19 different strawberry varieties appeared in the article “Evaluation of Strawberry Cultivars with Different Degrees of Resistance to Red Scale” (*Fruit Varieties Journal* [1991]: 12–17):