Chapter 5: Exploring Numerical Data

### Section 5.1: Scatterplots for Paired Data

Define the following terms:

Scatterplot:

Linear Relationship:

Nonlinear Relationship:

#### Notes:

What are the four characteristics of a scatterplot?

What are two variables that would have a horseshoe-shaped association in a scatterplot (∩ or ⌢)?

### Section 5.2: Dot Plots and the Mean

Define the following terms:

Dot plot:

Mean / Average:

Distribution:

Point Estimate:

### Section 5.3: Histograms and Shape

Define the following terms:

Histogram:

What is the tail of a distribution?

How does a right skewed distribution look?

How does a left skewed distribution look?

How does a right symmetric distribution look?

**Density Plot:**

Can can you tell what the mode of a distribution is?

How does a unimodal distribution look?

How does a bimodal distribution look?

How does a multimodal distribution look?

### Section 5.4: Variance and Standard Deviation

What is variability?

What does deviation mean?

What is the variance of a numerical variable? How is it calculated?

What is the standard deviation of a numerical variable? How is it calculated?

What does represent?

### Section 5.5: Boxplots, Quartiles & The Median

Sketch what a boxplot looks like:

What is a median?

What is the inner quartile range (IQR)?

What is the first quartile (Q1)?

What is the third quartile (Q3)?

What do the whiskers of a boxplot correspond with?

How are “outliers” plotted in a boxplot?

#### Notes:

How can you mathematically determine if an observation is an “outlier”?

### Section 5.6: Robust Statistics

What makes a statistic “robust”?

What are some robust statistics?

#### Notes:

What are the two ways to measure the ‘center’ of a distribution? Which one is considered robust to skew/outliers?

What are the three ways to measure the ‘variability’ of a distribution? Which one is considered robust to skew/outliers?

### Section 5.7: Transforming Data

What is a variable transformation?

Why would you decide to transform a variable?

### Wrap-up

What type of plot(s) are appropriate for displaying one quantitative variable?

What type of plot(s) are appropriate for displaying two quantitative variables?

How are variance and standard deviation related?

Fill in the following table with the appropriate notation.

| Summary Measure | Parameter | Statistic |
| --- | --- | --- |
| Mean |  |  |
| Variance |  |  |
| Standard deviation |  |  |

Statistics summarize \_\_\_\_\_\_\_\_\_\_\_\_\_ .

Parameters summarize \_\_\_\_\_\_\_\_\_\_\_\_\_.