***Real Stats* – Chapter 2**

**True or False Questions:**

1. **True** or False: The standard deviation measures how widely dispersed the values of the observations are.
2. **True** or False: If an analysis cannot be replicated, it cannot be trusted.
3. **True** or False: Replication files are used by researches to verify another researcher’s projects results.
4. **True** or False: Statistical projects should document both the data and the methods used to arrive at the conclusion.
5. True or **False**: The statistical package R is easier to use than Stata.

**Short Answer Question:**

1. Please list the descriptive statistics that are used to better understand the data.

**Answer: Mean, number of observations, standard deviation, minimum, maximum.**

1. Explain the importance of keeping replication files and properly documenting the data/research.

**Answer: Important because it allows others to check the robustness of the results and therefore helps verify if a study was done in an accurate manner.**

1. Describe/list some of the information that generally goes into a codebook.

**Answer: Type of data, source of data, variable information, and individual/organization that collected the data**.

1. What is the purpose of plotting the data?

**Answer: A useful first step in order to help identify any patterns or anomalies in the data prior to moving forward with further analysis.**

1. Compare and contrasts the basic benefits/drawbacks of Stata and R.

**Answer: Stata is more user friendly in the beginning, however once you get a hang of it, R could be easier to use because the coding is more direct and isn’t hidden behind a backend/GUI.**