Alloy Abasali

Prof. Robert Domanski

CPS 3320 - 01

March 4th, 2022

**Project Assignment #1 documentation**

**1. Background:** a few sentences to put your project idea in context and the overall goal of your project.

My family usually likes spending time together and eating at different restaurants. I was always curious to know if the waiters made good tips and if it makes a difference with the quantity of the bill. I usually tip higher if the bill amount is large; my curiosity is not about the included gratuity but the customer's generosity.

**2. Hypothesis:** the specific hypothesis you tested – something that is quantifiably true or false.

The higher the customer’s bill the higher the tip.

**3. Data:** a clear description of the data that you used, including:

I used a dataset from [www.kaggle.com](http://www.kaggle.com), the data was a representation of ranjeetjain3/seaborn-tips-dataset .

**4.** **What the data represents:**

The data represents tips, total amount of the bill, the day and time tips were collected an what meal from one particular restaurant.

**5. What are the possible biases with the data?**

* Size of the data
* Location data was collected
* Gender of the waiters collecting the tips
* Type of restaurant: casual or formal

**6. What are other potential issues with the data**?

The size of the data might be too small to make a generalized conclusion because it was based on one restaurant and not many.

**4. Analysis/Methodology:** a list of steps that were executed to carry out your analysis.

Step 1: Research to find the dataset for tips in a restaurant

Step 2: Downloading and importing the important data need to prove my hypothesis

Step 3: Used python code to group data headings to get results from data set.

Step 4: Used python libraries to plot the graphs to analyze the data.

Step 5: After all data used was analyzed, conclusion was made

**7. Describe your main computation.** That is, describe in English or Pseudocode the main,

high-level overview of what you’ll calculate.

The main computation was to analyze total of bill vs total tips reported.

**8. Describe your results.**

My hypothesis proved true with this data set; the scatter plot showed that the larger the bill, the larger the tip. Also, other data revealed that tips were higher on weekends and higher at dinner time.

**9. Describe how the end product of your analysis**: does support your claim, or show it is false.

My end product supports my claim because when I used the scatter plot bill diagram to compare the size of the bill to the amount of tip, it displayed that tips were higher with the larger bills.

**10.** **What were your results** : hypothesis supported or refuted? The goal is to convey your findings effectively.

I concluded that my hypothesis was confirmed with this dataset

**11. If your results were expected**: how could you see expanding this (a short 1-sentence idea is

sufficient) to test the limits of your hypothesis.

To expand and test the limits of my hypothesis, I would test If males would tip more than females or couples would tip more than families.