Homework 04

DSO 545: Statistical Computing and Data Visualization

Fall 2019

Due Date: Friday September 27, 2019 (EOD at 11:59pm)

Instructions

- Use Python to answer all questions (Feel free to use any IDE or Jupyter notebook)
- Submit your Python file to blackboard with comments when needed
- USC won't tolerate any kind of cheating
- Good luck

Case 01: Heavenly Chocolates

Heavenly Chocolates manufactures and sells quality chocolate products and its plant and retail store located in Saratoga Springs, New York. Two years ago, the company developed a website and began selling its products over the Internet. Website sales have exceeded the company's expectations, and management is now considering startegies to increase sales even further. To learn more about the website customers, a sample of 50 Heavenly transacations was selected from previous month's sales. Data showing the day of the week each transaction was made, the type of browser the customer used, the time spent on the website, the number of web pages viewed, and the amount spent by each of the 50 customers are contained in the files names `HeavenlyChocolates.csv`.

Heavenly Chocolates would like to use the sample data to determine whether online shoppers who spend more time and view more pages also spend more money during their visit to the website. The company would also like to investigate the effect that the day of the week and the type of browser have on sales.

Variables

Variable	Explanation
Customer	customer's id
Day	the day of the week each transaction was made
Browser	the type of browser the customer used
Time	the time spent on the website (in minutes)
Pages_Viewed	the number of web pages viewed
Amount_Spent	the amount spent by each (in dollars)

1. **(2 points)** Using the techniques we have learned in class so far (i.e. DON'T use groupby() method), create a dataframe that shows the number of visits to Heavenly Chocoloate's website using different browsers. Your output should be similar to the following table.

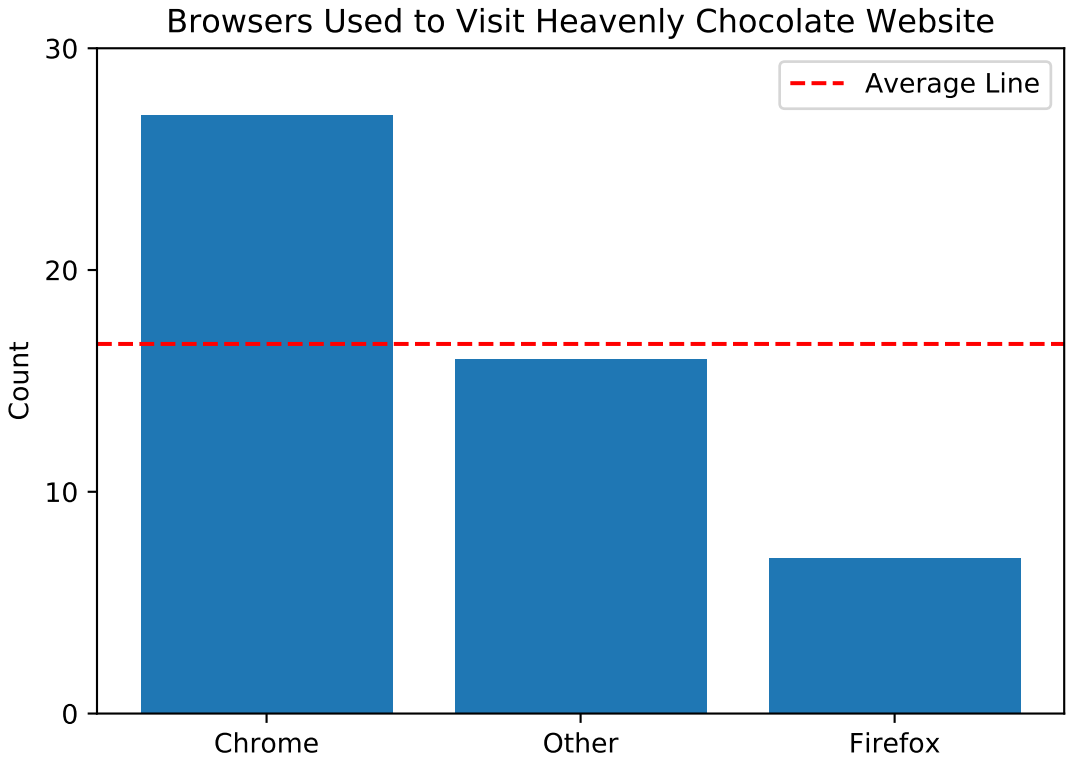
Browser	Count
Chrome	27
Other	16
Firefox	7

2. **(2 points)** Create a duplicate of the following graph to show which browser is used more by the

Heavenly Chocolate website visitors? The dashed line represents the mean of count visits for all browsers in the dataset.

```
## <BarContainer object of 3 artists>
```

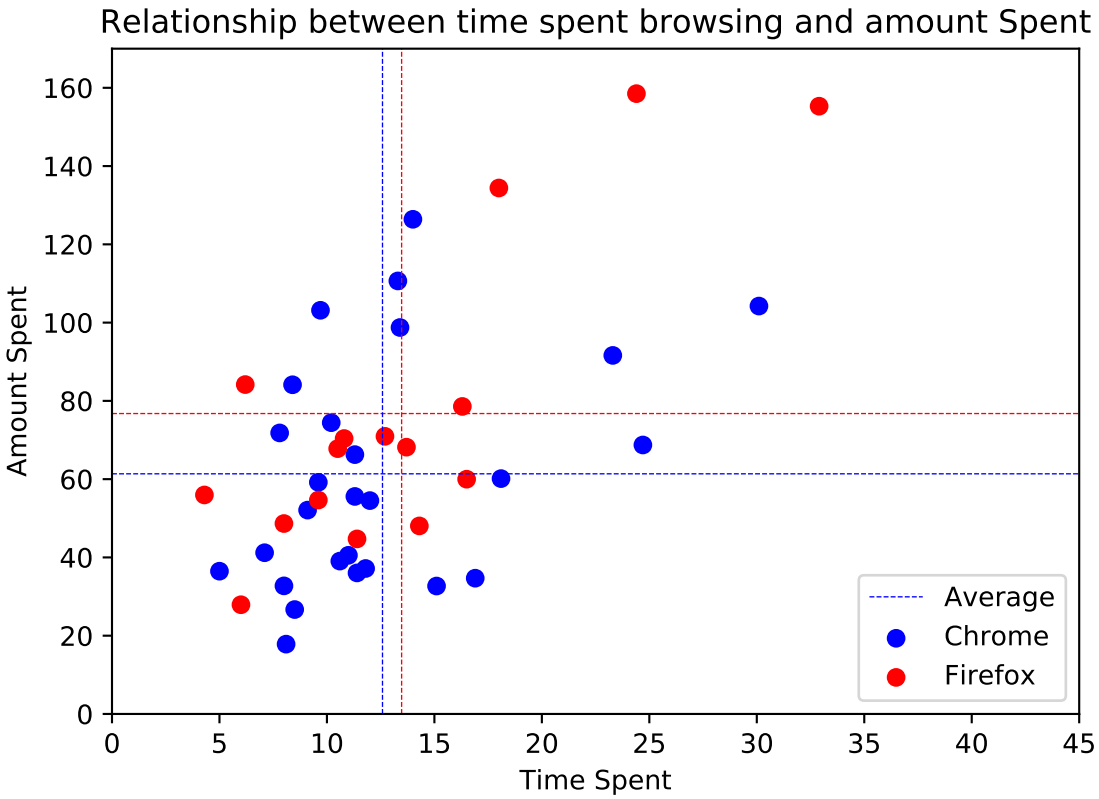
```
## [<matplotlib.axis.YTick object at 0x123d80990>, <matplotlib.axis.YTick object at 0x123d800d0>, <matpl
```



3. **(2 points)** Create an duplicate of the following scatterplot that shows the relationship between the time spent browsing, and the amount spent buying using either Firefox or Chrome browswers. What does this plot tell you about the company's revenues using both channels?

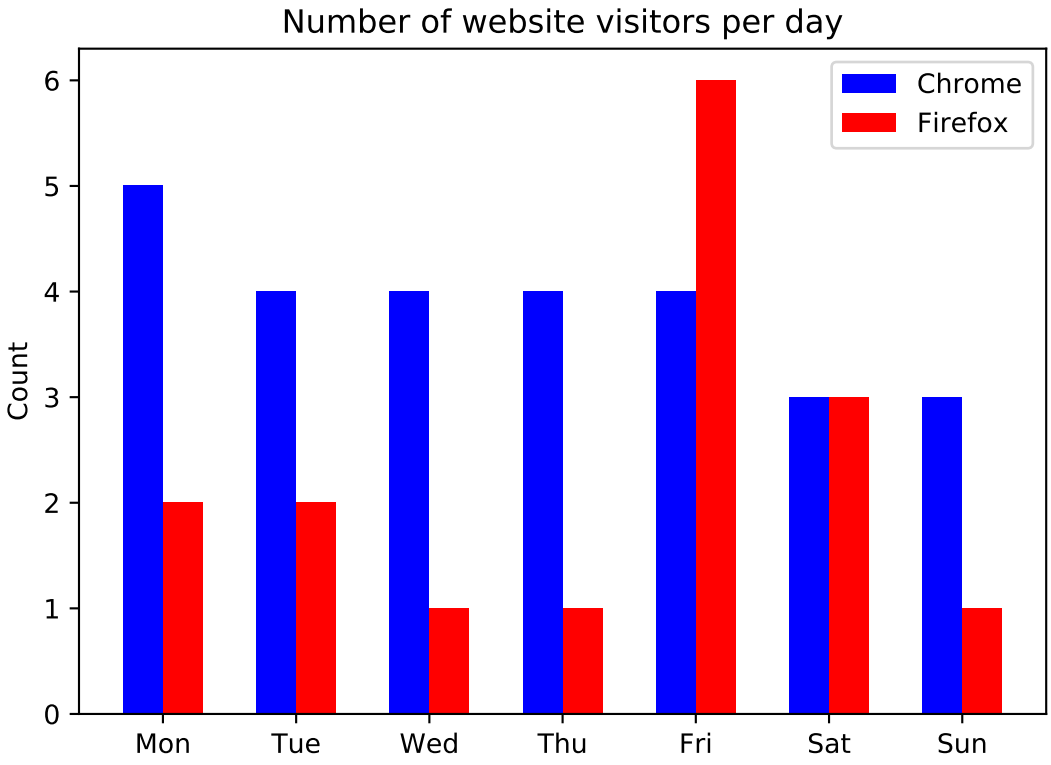
```
## (0, 170)
```

```
## (0, 45)
```



4. **(2 points)** Create the following chart which shows the number of website visitors per day for people using Chrome or Firefox.

```
## <BarContainer object of 7 artists>
## <BarContainer object of 7 artists>
## ([<matplotlib.axis.XTick object at 0x12713d590>, <matplotlib.axis.XTick object at 0x127141c50>, <mat
## [Text(0, 0, 'Mon'), Text(0, 0, 'Tue'), Text(0, 0, 'Wed'), Text(0, 0, 'Thu'), Text(0, 0, 'Fri'), Text
```



Case in Point: Graph Analysis for Consulting and Case Interviews

5. **(2 points)** Read pages 13-20 (excluding Radar Chart), and create a powerpoint presentation that discusses the main points in these pages (max of 6 slides).

PS. I have intentionally left out the section on Pie Charts for personal reasons. :D