

# Assignment Report for Shoppers Dataset

## Statistics

**AYA SABRY MOHAMED**

*Cairo University*

**DA ISMAILIA**

**Submitted to  
Dr. Abdeltawab Ahmed Gira**

Jan 16, 2023

# Questions

## Colab link

<https://colab.research.google.com/drive/1KkR8LO-rqjXHzxLvaDFplQEovXHisji0?usp=sharing>

**Use the methods of descriptive statistics to learn about the customers who visit the Heavenly Chocolates website. Include the following in your report.**

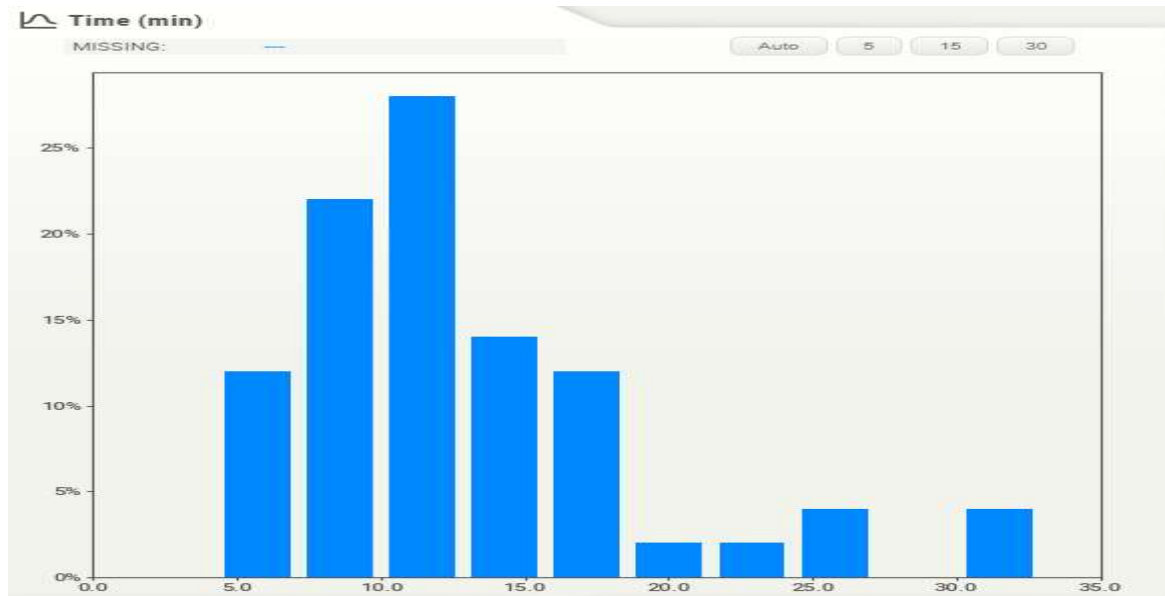
1. Graphical and numerical summaries for the length of time the shopper spends on the website, the number of pages viewed, and the mean amount spent per transaction.

Q1

AYA

```
[17] 1 shoppers.describe()
```

| 1 to 8 of 8 entries |                    |                    |                    | Filter |  | ? |
|---------------------|--------------------|--------------------|--------------------|--------|--|---|
| index               | Time               | Pages              | Amount_spent       |        |  |   |
| count               | 50.0               | 50.0               | 50.0               |        |  |   |
| mean                | 12.809999999999997 | 4.82               | 68.12819999999999  |        |  |   |
| std                 | 6.063313904970126  | 2.0373051428045725 | 32.343757855327844 |        |  |   |
| min                 | 4.3                | 2.0                | 17.84              |        |  |   |
| 25%                 | 8.65               | 3.25               | 45.559999999999995 |        |  |   |
| 50%                 | 11.4               | 4.5                | 62.15              |        |  |   |
| 75%                 | 14.9               | 6.0                | 82.735             |        |  |   |
| max                 | 32.9               | 10.0               | 158.51             |        |  |   |



Discuss what you learn about Heavenly Chocolates' online shoppers from these numerical summaries.

Last month, the average Heavenly Chocolates website visitor spent **13 minutes** browsing, viewed **five pages**, and spent **\$68**.

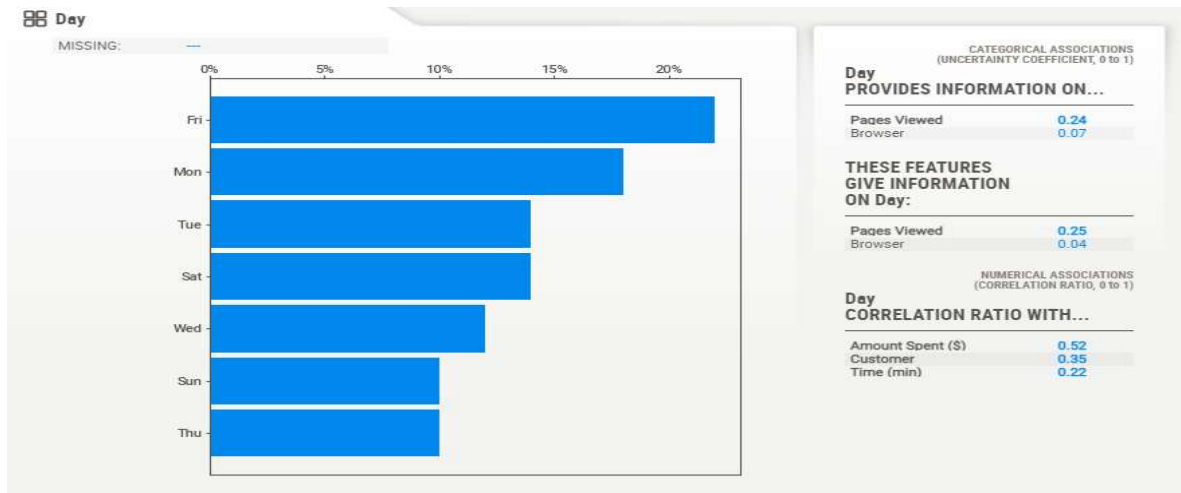
2. Summarize the frequency, the total dollars spent, and the mean amount spent per transaction for each day of week. What observations can you make about Heavenly Chocolates' business based on the day of the week? Discuss.

Q2

AYA

| Day | Average amount spent | No of transaction | Total amount       |
|-----|----------------------|-------------------|--------------------|
| Fri | 85.948181818182      | 11                | 945.430000000001   |
| Mon | 90.3755555555555     | 9                 | 813.38             |
| Sat | 54.105714285714285   | 7                 | 378.74             |
| Sun | 43.63                | 5                 | 218.15             |
| Thu | 58.806000000000004   | 5                 | 294.03000000000003 |
| Tue | 59.26571428571429    | 7                 | 414.86             |
| Wed | 56.97                | 6                 | 341.82             |

Show 25 per page



| TOP CATEGORIES |    |      |
|----------------|----|------|
| Fri            | 11 | 22%  |
| Mon            | 9  | 18%  |
| Tue            | 7  | 14%  |
| Sat            | 7  | 14%  |
| Wed            | 6  | 12%  |
| Sun            | 5  | 10%  |
| Thu            | 5  | 10%  |
| ALL            | 50 | 100% |

- The days with the most transactions were **Monday and Friday**, with **9** and **11** transactions, respectively.
- The least active day of the week was **Sunday**, when only **five** transactions took place.
- Additionally, **Monday and Friday** saw the **highest** average transaction costs, at **\$90.38** and **\$85.95**, respectively.
- **Sunday** is the day with the lowest mean amount spent and lowest total amount spent, with a mean amount spent of **43.63 dollars** and a total amount spent of **218.15 dollars**.
- The corporation should therefore expand its advertising initiatives to reach more clients in light of the day of the week statistics. Provide and offer special deals, discounts, and other promotions, especially on **Thursdays and Sundays**.

3. Summarize the frequency, the total dollars spent, and the mean amount spent per transaction for each type of browser.

Q3

AYA



The screenshot shows a data table interface with a title bar '1 shoppersBrowser'. The table has four columns: 'Browser', 'No of transaction', 'Average amount spent', and 'Total amount'. It contains three rows of data. Below the table, there is a 'Show 25 per page' dropdown and a link to a 'data table notebook'.

| Browser           | No of transaction | Average amount spent | Total amount |
|-------------------|-------------------|----------------------|--------------|
| Firefox           | 16                | 76.763125            | 1228.21      |
| Internet Explorer | 27                | 61.36333333333333    | 1656.81      |
| Other             | 7                 | 74.48428571428572    | 521.39       |

Show 25 per page  
Like what you see? Visit the [data table notebook](#) to learn more about interactive tables.

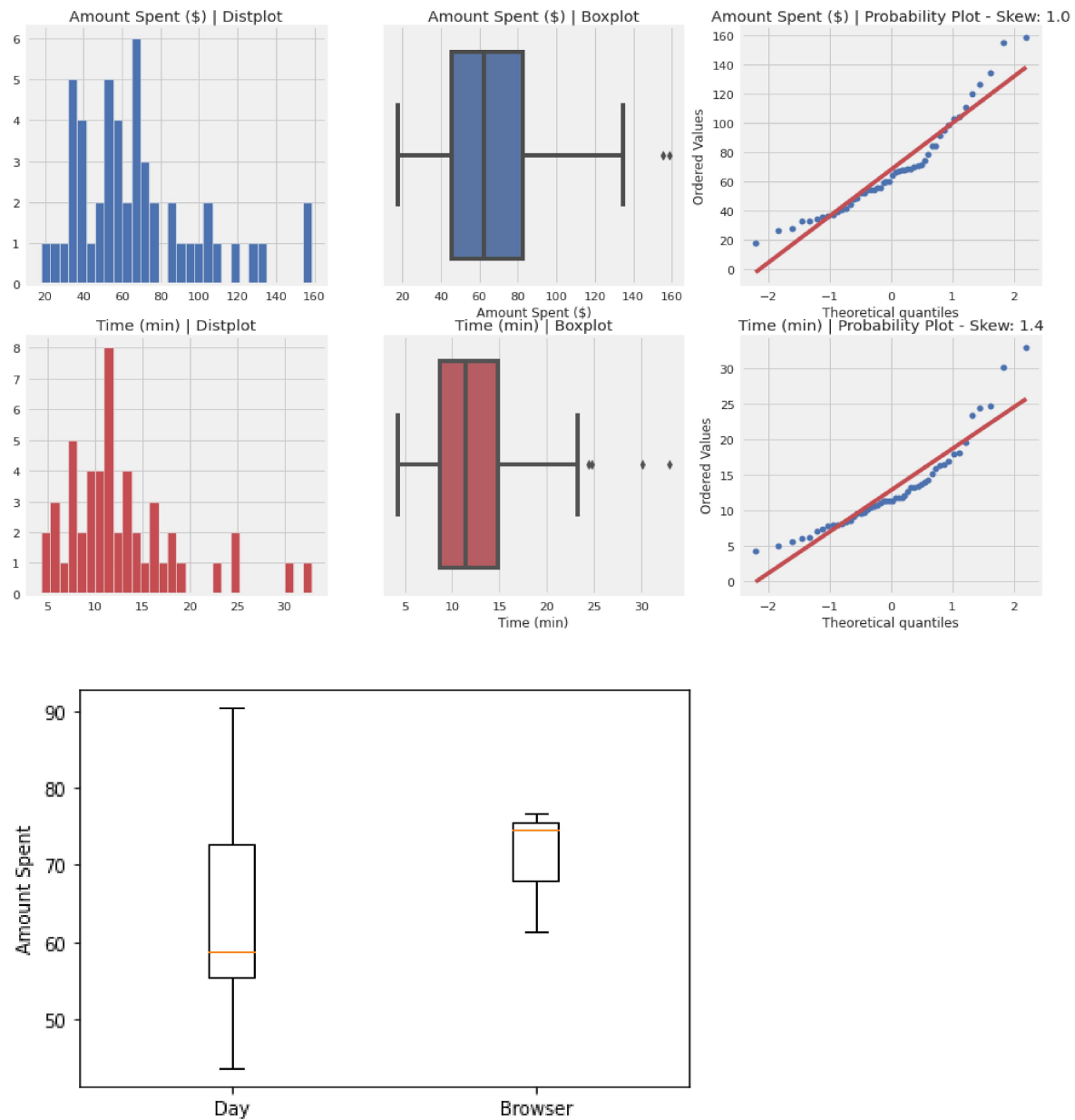
What observations can you make about Heavenly Chocolate's business based on the type of browser? Discuss.

It shows that **Internet Explorer**, out of the various browsers, has the highest frequency, with **27** instances, a total cost of **\$1228.21**, and an average transaction cost of **\$76.66**.

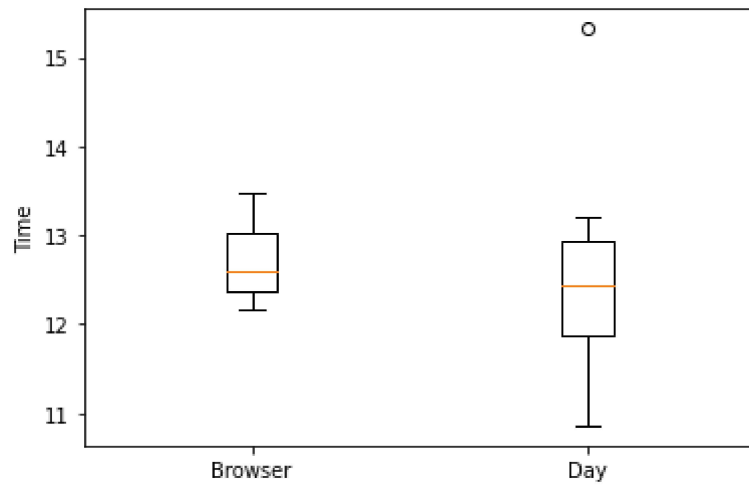
In contrast, **Firefox** has **16** frequencies, a spending total of **\$1656.21**, and an average transaction cost of **\$61.36**.

Another is the user who browses the least, paying **\$521.39** overall, an average of **74.48** each transaction, and having a frequency of just **7**. As a result, marketing strategies that include providing specials and other tactics must be established for **Firefox and other browser users**.

4. Develop a boxplot for each variable and also compare by browser and day . discuss



-We noticed **Right Skew in Amount Spent Average per each Day**, Also in the opposite **Left Skew in Amount Spent Average per each Browser**.

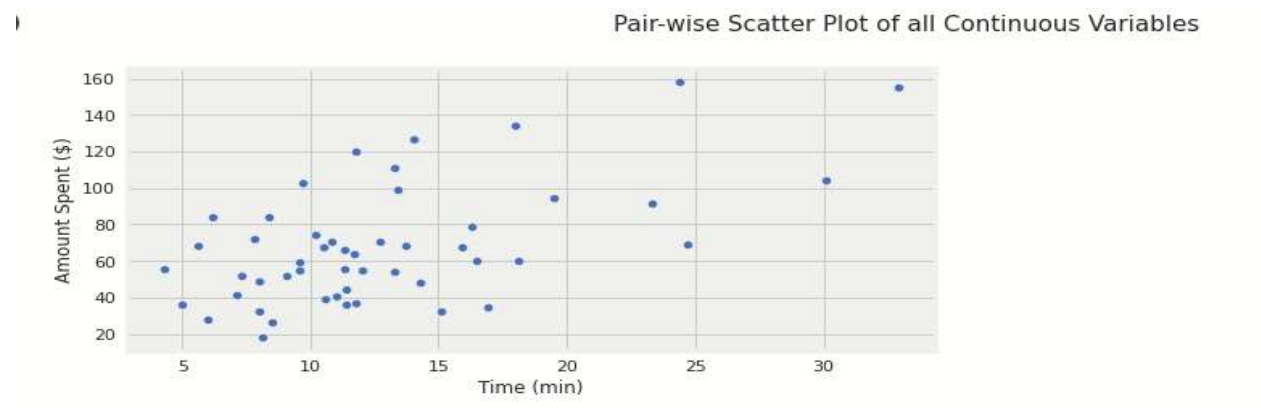


**In Browser Vs Time Skewness:** Then slightly Right skew similarly in **Day Vs Time** But higher than that in Browser



5. Develop a scatter diagram and compute the sample correlation coefficient to explore the relationship between the time spent on the website and the dollar amount spent.

Use the horizontal axis for the time spent on the website. Discuss.



## ▼ Q5 Correlation

Aya

```
1 print(Corr_time_spend)
```

0.580047615980413

The relationship between time spent online and money spent has a correlation coefficient of **0.58**. This suggests a **significant positive association** between **time spent online** and **total spending**. Thus, it may be inferred that, generally speaking, website visitors who stay longer make purchases.

6. Develop a scatter diagram and compute the sample correlation coefficient to explore the relationship between the number of website pages viewed and the amount spent. Use the horizontal axis for the number of website pages viewed. Discuss.

## Q6 Correlation

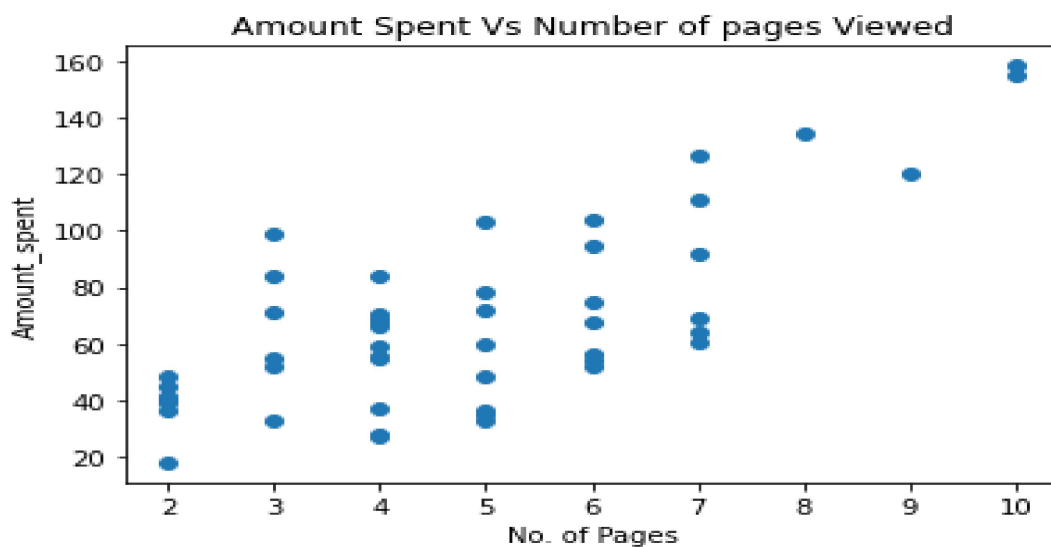
Aya

▶ 1 cell hidden

## Correlation

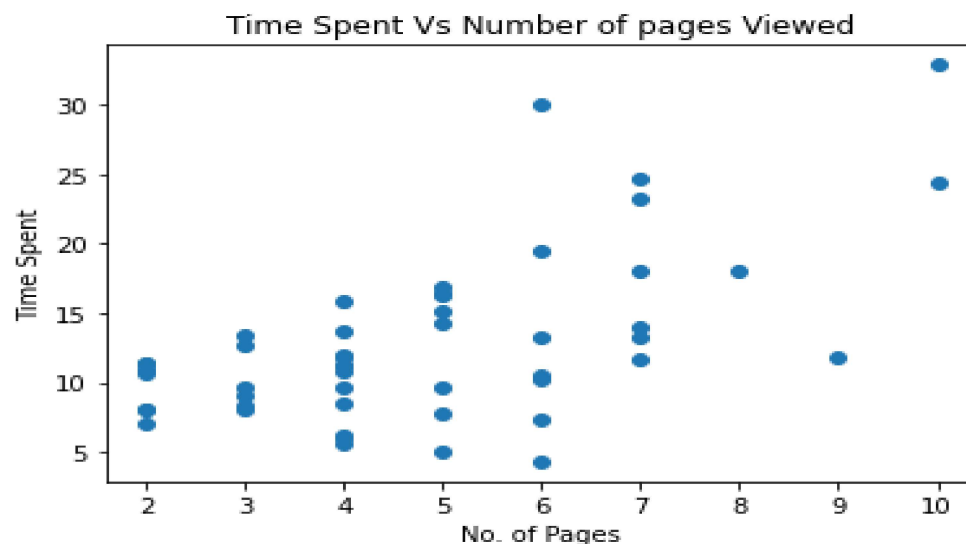
```
[36] 1 print(Corr)
```

0.723666923476961



The correlation coefficient is **0.72**. According to this correlation coefficient, there is a **stronger** connection between the **quantity of website pages visited and the total amount spent**. It demonstrates that when **more people read the pages**, they spend more money.

7. Develop a scatter diagram and compute the sample correlation coefficient to explore the relationship between the time spent on the website and the number of pages viewed. Use the horizontal axis to represent the number of pages viewed. Discuss.



## Q7 Correlation

Aya

1 cell hidden

## Correlation

[39] 1 print(Corr\_time\_pages)

0.5955675236548259

According to this correlation coefficient, there is a **positive correlation (.5956)** between the **amount of time spent and the number of website pages seen**. It demonstrates that the **longer** time spent on a website, the **more pages are visited**. However, there is **less of a relationship** between these two variables than there is between page views and spending.(.72)

According to the data, increased sales are associated with increased page views and time spent. The company should improve the website design so that more customers will visit and view it, resulting in increased sales.

Bar plots for each Continuous by each Categorical variable

