

Data Analytics Project Overview

Cali Wilkerson

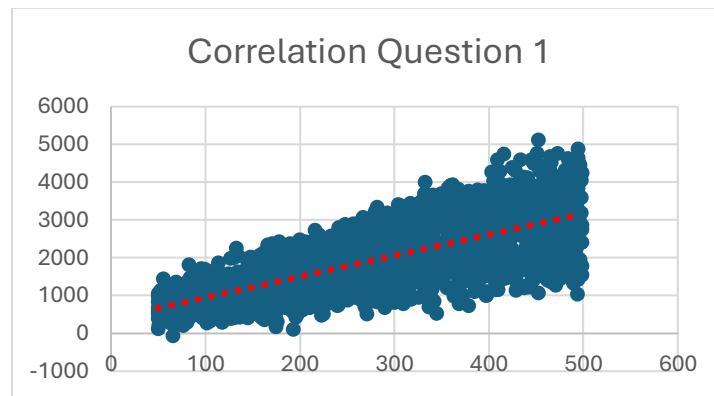
Coffee Shop Daily Revenue Prediction Dataset:

<https://www.kaggle.com/datasets/himelsarder/coffee-shop-daily-revenue-prediction-dataset?resource=download>

This data set lists data from 2,000 different types of coffee shops. From smaller local coffee shops to large, populated shops that spend large amounts on their marketing campaigns. The data provided helps to give insight into how different factors such as customer behavior and business activities affect the coffee shop's performance. The data list includes the number of customers per day, average order value, operating hours per day, number of employees, money spent on marketing per day, location foot traffic, and daily revenue.

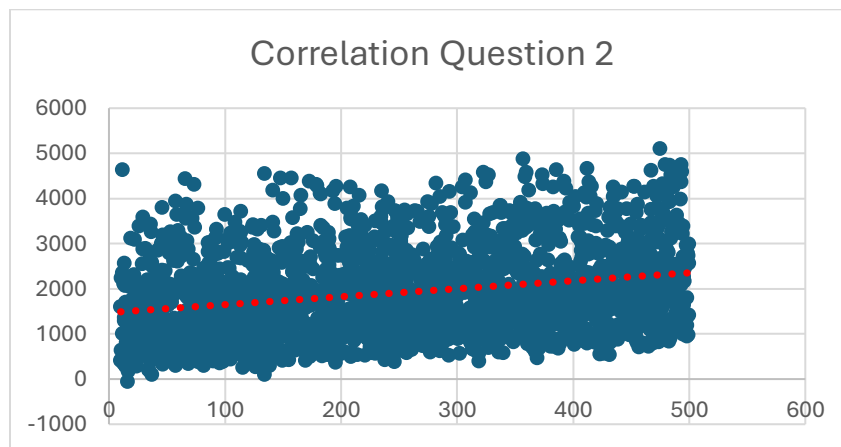
Question 1: What is the correlation between customer count and daily sales? (Complex analysis)

Using the =CORREL () function to find the relationship between the customer count and the daily sales, the results indicated a strong relationship. The result was .736460699, which means they are basically directly affected by each other. If one goes up, then the other will go up as well.



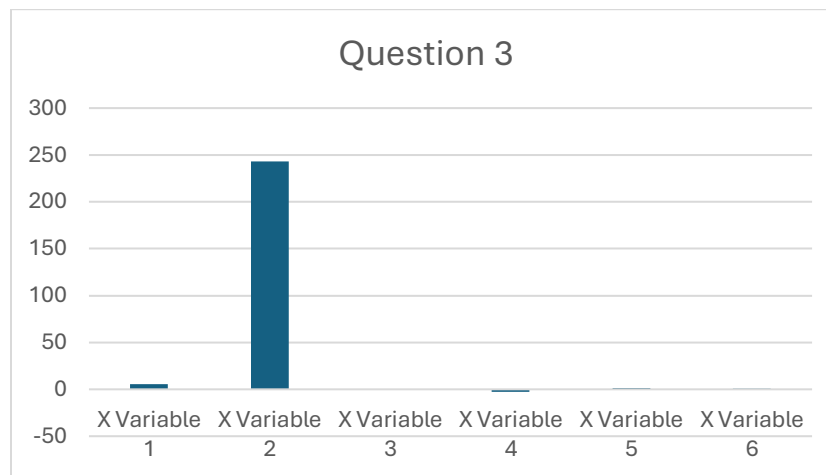
Question 2: Does the amount of money spent on advertising and the revenue have any correlations? (complex analysis)

Using the =CORREL () function, the results showed that the correlation between the amount of money spent on marketing and the revenue is .25481245. This means there is a slight positive correlation. If the marketing costs increase, the revenue will as well, but the relationship between the two is not very strong.



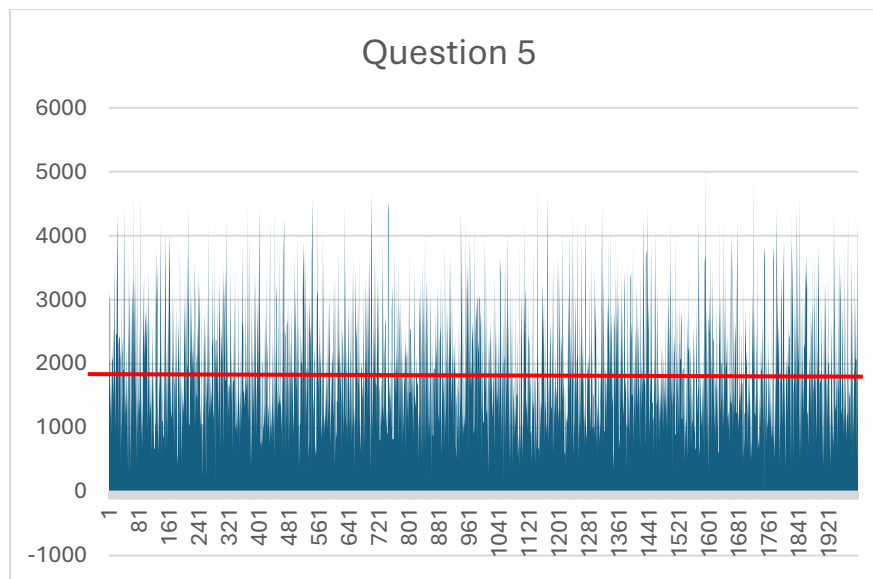
Question 3: Which factor affects the revenue the most? (Complex analysis)

After completing a regression analysis on the dataset, the average order value affected the coffee shop revenue the most because its coefficient was 243.2842435 which is the out of all the X variables which means they have a very strong correlation with each other.



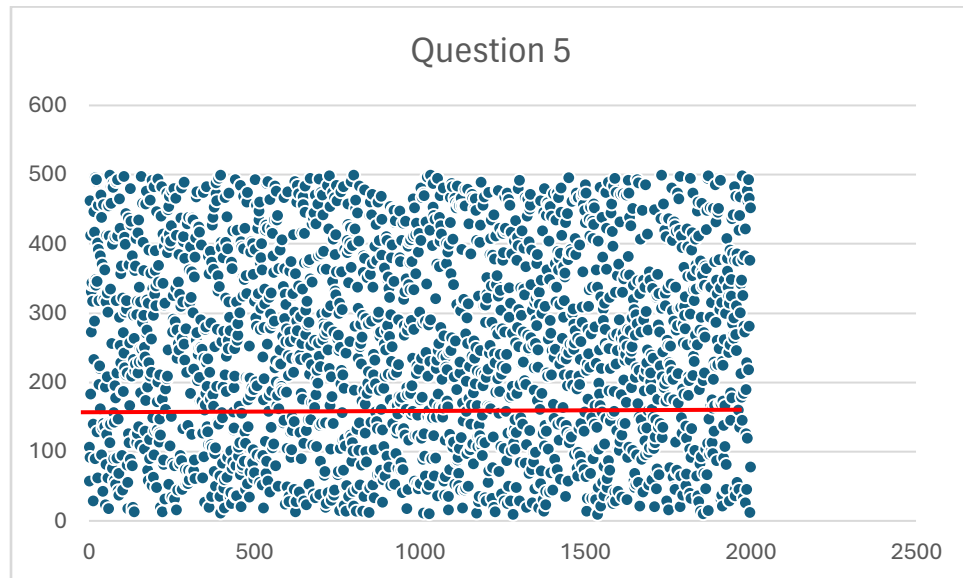
Question 4: What would the average of the daily revenue column be? (Simple analysis)

To find the average daily revenue the `=AVERAGE ()` function was used, and the result was \$1,917.33.



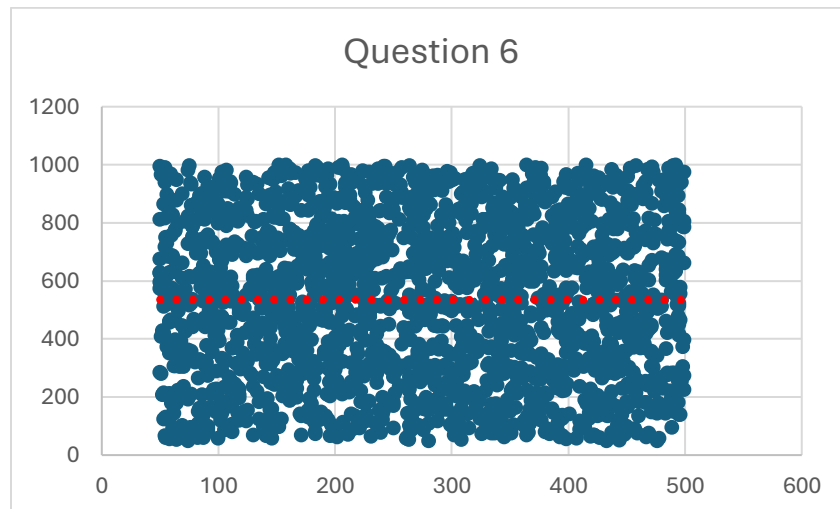
Question 5: What is the standard deviation of the marketing spend per day? (Simple analysis)

The =STDEV.P() function was used to solve this and the standard deviation of the marketing spend per day came out to be \$141.10.



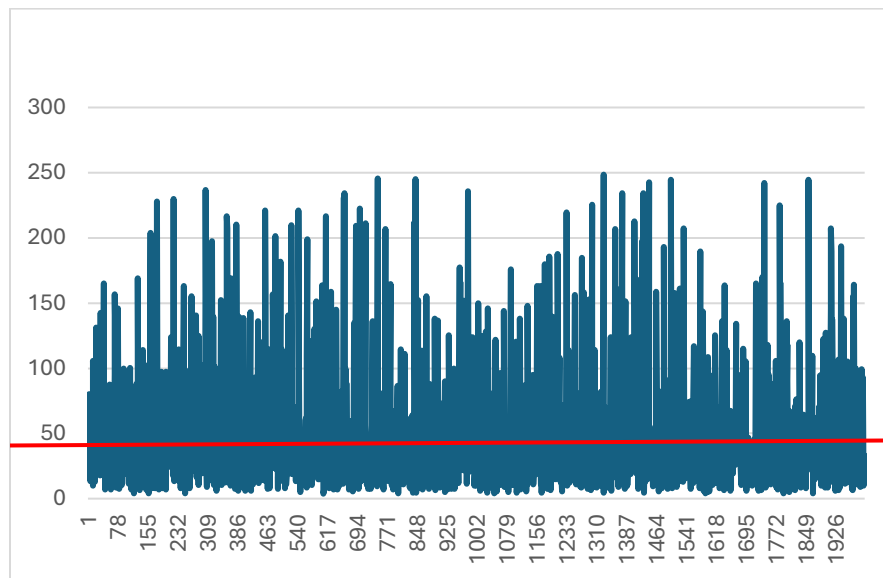
Question 6: Compare the foot traffic to the number of customers per day. Does higher foot traffic correlate to the number of customers?

Using the =CORREL () function it was found that there was a $-.00088$ correlation, so there's a slight negative correlation between the foot traffic and the number of customers.



Question 7: What is the average ratio of the number of customers to the number of employees?

To solve this, the number of customers was divided by the number of employees and then found the average of all the answers. The result was 47.9364 average customers per employee



Conclusion

Doing this data analysis project showed how to use the regression analysis along with the average function, standard deviation, correlation and determine the ratio between two rows of data. This project also revealed how much information you can gather just from one dataset. Learning this information is so important as technology keeps becoming apart of our daily lives and career as excel is used in many jobs.

These findings can assist managers in many different aspects of their job. For example, the findings from question seven would be helpful when managers are hiring employees and determining if they are under or overstaffed because they can see how they operate with certain customers to employee ratios. Another example could be the findings from question 5 helping the manager keep track of how the amount they're spending on marketing is impacting their company and the success of their marketing campaigns.

Some data that could have helped to provide so much more information could be the number of sales during different quarters of the year/ the period of time this data was taken because then you would be able to get more specific and see what the busiest time of year is. Even more specifically, they could have also provided the data for the time of day on the foot traffic and number of customers throughout different parts of the day.

Overall, this project presented many learning opportunities and can be used in marketing, operating, keeping track of finances, but also has so many more capabilities. Excel can be an extremely useful tool for employees, managers and even just in your daily life. While having more information is always useful, so much can be done with a few stats about a company.