

## **BANA 200 Assignment 2**

### **Data Visualization and Regression Analysis**

**Due Wednesday, August 24th on Canvas by 6PM Pacific Standard Time (1AM UTC Time)**  
**50 Points**

#### **Overview:**

One of the most important elements to good business analytics is data visualization: The ability to show important relationships visually and the ability to “tell the story” with graphs and charts is crucial.

The cleaned Excel file “Starbucks HW2 Data.xlsx” contains survey data on a random sample of 6,121 Starbucks Coffee customers. The survey was done in Orange County, CA, and contains the following data:

1. **X1:** Overall, how would you rate the beverages served at Starbucks? - Taste
2. **X2:** Overall, how would you rate the beverages served at Starbucks? - Overall quality
3. **X3:** Overall, how would you rate the beverages served at Starbucks? - Temperature
4. **X4:** Overall, how would you rate the beverages served at Starbucks? - Freshness
5. **X5:** Overall, how would you rate the beverages served at Starbucks? - Presentation
6. **X6:** Overall, how would you rate the beverages served at Starbucks? - Variety
7. **X7:** Overall, how would you rate the food served at Starbucks? - Temperature
8. **X8:** Overall, how would you rate the food served at Starbucks? - Variety
9. **X9:** Overall, how would you rate the food served at Starbucks? - Taste
10. **X10:** Overall, how would you rate the food served at Starbucks? - Overall quality
11. **X11:** Overall, how would you rate the food served at Starbucks? - Presentation
12. **X12:** Overall, how would you rate the food served at Starbucks? - Freshness
13. **X13:** How do you rate the value for the money?
14. **X14:** How would you rate the Starbucks staff along the following dimensions? - Well dressed and appear neat
15. **X15:** How would you rate the Starbucks staff along the following dimensions? - Remembering your name
16. **X16:** How would you rate the Starbucks staff along the following dimensions? - Knowledgeable
17. **X17:** How would you rate the Starbucks staff along the following dimensions? - Personal treatment
18. **X18:** How would you rate the Starbucks staff along the following dimensions? - Polite
19. **X19:** How would you rate the Starbucks staff along the following dimensions? - Remembering your order correctly
20. **X20:** How would you rate the Starbucks staff along the following dimensions? - Friendly/attentive
21. **X21:** How would you rate the Starbucks staff along the following dimensions? - Have your best interest at heart
22. **X22:** How would you rate the Starbucks staff along the following dimensions? - Providing prompt service
23. **satis100:** A customer satisfaction variable that ranges from 0 to 100 points. Customers were asked the following question: “Overall, how satisfied are you with Starbucks? 0 = very dissatisfied; 100 = very satisfied.”

24. **recommend:** “How likely are you to recommend Starbucks to others? 0 = definitely WILL NOT recommend; 10 = definitely WILL recommend.” This variable ranges from 0 to 10.
25. **profits:** Average monthly profits that Starbucks earns on each customer (in US Dollars). Some profit numbers may be negative (i.e. Starbucks loses money on some customers).
26. **ZipCode:** The five digit zip code associated with the customer’s place of residence.
27. **Income:** Estimated annual income of each customer (reported in US Dollars), based on the US Census Bureau Zip Code demographics data.

Variables X1 – X22 are all measured on a 5 point scale (1 = terrible, 2 = poor, 3 = average, 4 = good, 5 = excellent).

Imagine for a moment that you have been hired by the Starbucks’s Corporation as a data scientist. Your job over the next several weeks is to conduct some insightful analysis to help senior management understand more about how to improve customer engagement and profitability. In this second assignment, you will get practice with doing data visualization and data interpretation.

### **Q1 Data Visualization (20 Points)**

Using either R or Excel (you may use the software of your choice for data visualization), open the “Starbucks HW2 Data.xlsx” Excel spreadsheet. Once you have connected/opened/imported the Excel spreadsheet, create two high quality graphs or charts. The charts and graphs can be anything of your choosing, but must contain at least two of the following four variables: {satis100, recommend, profits, income}.

Once you have created your two charts, paste them into a Word document. Then spend some time interpreting them. The focus here is on “telling the story”: What interesting relationships do you see among the variables? What can you conclude? Based on these charts, what recommendations would you give to senior management?

**Hint:** If you can’t find anything interesting, then you should consider trying a different chart or a different set of variables. Choose the charts or graphs that you feel are most compelling or interesting. You are allowed to summarize the data in any way you see fit (e.g. taking averages, grouping the data into bins, creating categories etc.) but you must explain in your write up any data manipulations you performed.

### **Q2 Regression Analysis (15 Points)**

Import the “Starbucks HW2 Data.txt” file into R. Executive management is interested in understanding the impact of factors that affect the average monthly profits of its customers. Run a regression using “profits” as the dependent variable and use “satis100”, “recommend”, and “Income” as the three independent variables.

Report your regression results below (including the regression estimates and significance levels) and answer the following questions:

- a) Do the three predictor variables do a good job of predicting the average monthly profits of each customer? Comment on the p-values and the  $R^2$  value, and interpret the meaning of the  $R^2$  value in this context.
- b) For each 10 point increase in satisfaction (e.g. for a 10 point increase in satis100), by how much do we expect the average monthly profits to go up by? Round your answer to two decimal places and comment on whether you feel like this is a big increase in average monthly profitability or not.
- c) Calculate the predicted average monthly profits for a customer with satis100 = 77, recommend = 8, and Income = \$121,500. Report this predicted profit value below rounded to two decimal places.

### **Q3 Dummy Variables Regression (15 Points)**

Senior management at Starbucks is also very interested in understanding whether failing to meet customers' expectations has a bigger effect on profits than does exceeding customer expectations. In order to test this, you will need to create two new dummy variables from "satis100" and run a regression. First, do the following:

1. Create a dummy variable called "fail" that equals 1 if satis100 < 20 and 0 otherwise. This dummy variable is flagging very dissatisfied customers.
2. Create a dummy variable called "exceed" that equals 1 if satis100 > 80 and 0 otherwise. This dummy variable is flagging highly satisfied customers.
3. Once you have created these two new dummy variables, rerun your regression in Q2 by using profit as the dependent variable, but now use fail, exceed, recommend, and Income as the four independent variables. Be sure to exclude satis100 as one of the predictor variables when you run the regression (you are now using the two dummy variables to take the place of satis100).

Now, paste your table of regression results from R below and answer the following questions:

- a. How many customers in the dataset is Starbucks failing to meet customer expectations? That is, report the number of customers where satis100 < 20.
- b. How many customers in the dataset is Starbucks exceeding customer expectations? Report the number of customers in the dataset where satis100 > 80.
- c. Comment on the regression coefficients for the dummy variables "exceed" and "fail". What seems to have a bigger impact on profitability: Failing to meet customer expectations or exceeding them? Report the expected change in profits when Starbucks exceeds customer expectations (exceed = 1) vs. when Starbucks fails to meet customer expectations (fail = 1) and comment. What advice would you give to senior management?