Joseph Silva Jr

09/26/21

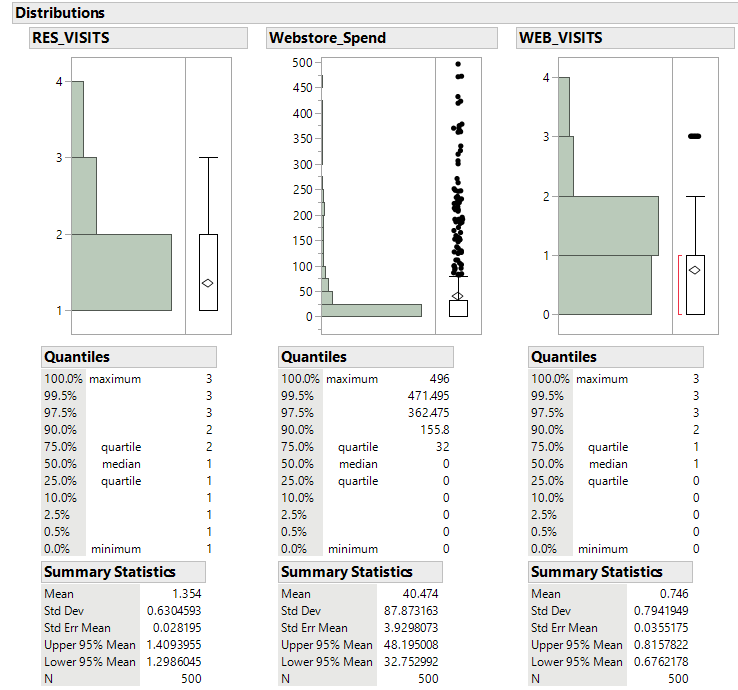
SNHU

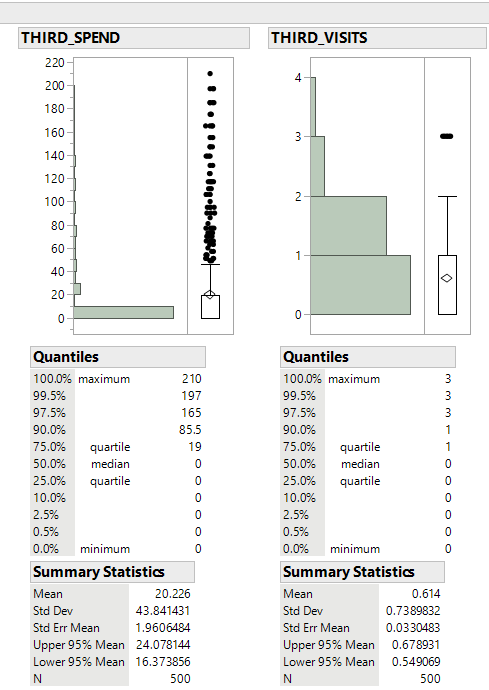
DAT 220: Module 4 Exercise

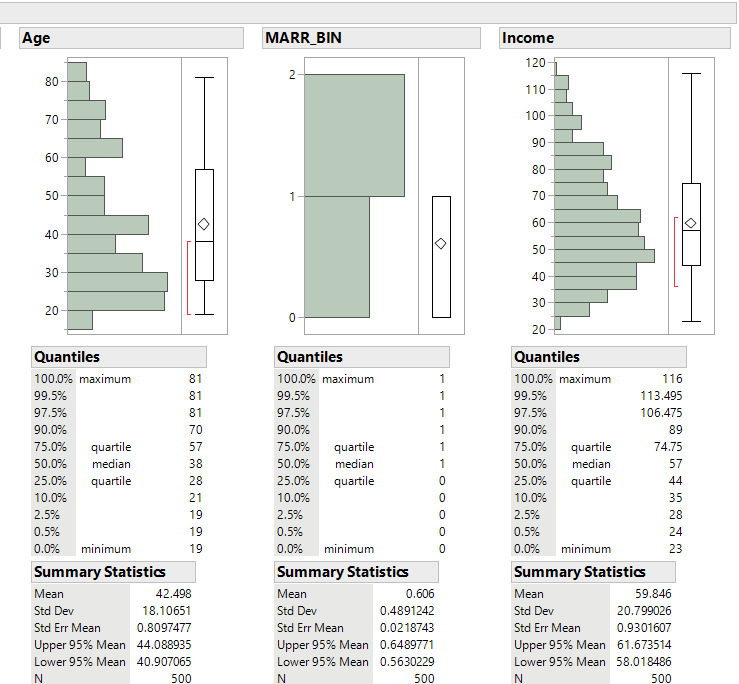
**#1 Sample Data Survey Source:**

The information for this data sample survey was taken from a data warehouse for the Bubba Gump restaurant. The data was information collected from customers of the Bubba Gump restaurant and these customers’ information were stored in the Bubba Gump’s data warehouse. The data stored contained the following customer’s information: first name, last name, city, county, state, zip, restaurant purchase price, restaurant visit, web purchase, web purchase price, web visit, third party spend, third party visit, age, marriage, and income. For this survey, 500 randomly selected customers data was selected from the data warehouse to be analyzed. The survey was to use a small percentage of customers chosen at randomly to use as a model to discover ways to increase the revenue for Bubba Gump’s restaurant.

**#2 Data Set Survey Visualizations:**

****

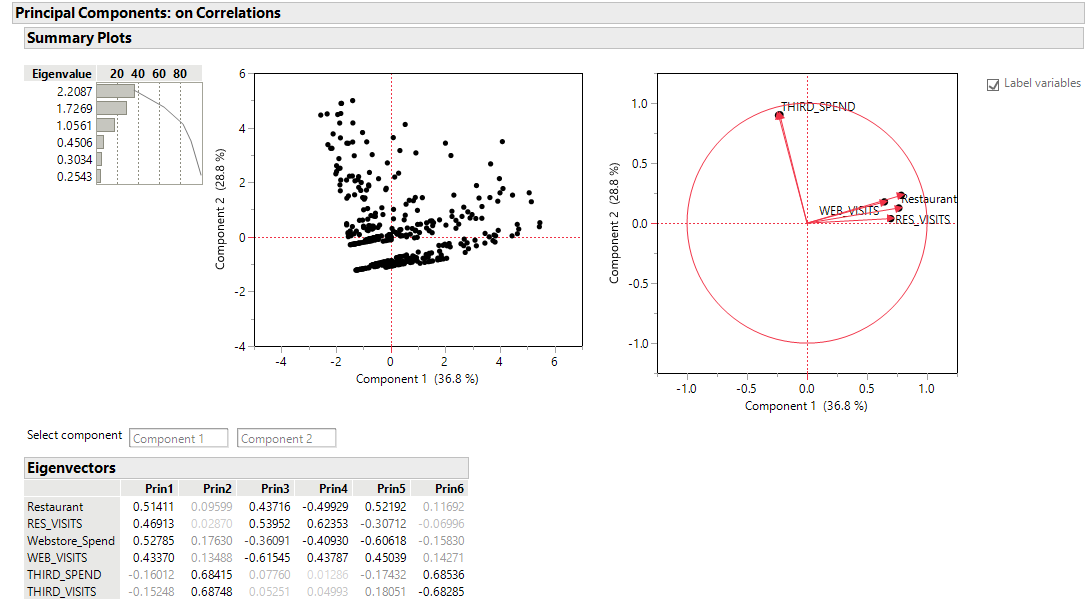
****

****

With these graph designs, we can see a visualization of the data taken for the survey. We can see customers tend to go to the webstore to but merchandise instead of using third-parties or visit the restaurants. The webstore can be shown to be the main source of revenue for Bubba Gump. I can also see majority of the customer only visit Bubba Gump restaurant or webstore mostly once instead of being regular customers, which can be a reason for the decrease of revenue over a certain number of years. We can see a breakdown for the age of the customers that will mostly visit Bubba Gump. We can see the age area of 20-45 for the customers have the highest rate of visitations and we can see a decrease of the visitation rate from customers after the age of 55 or below 20. These types of information can show Bubba Gump about which groups their restaurant is being able to affect whether it be age, income, and/or location.

**#3 Correlations and Associations:**

****

****

We can see in the correlation that the third-party variables whether it be spending, or visits are very low when comparing their relationship with restaurant purchase / visits and webstore purchase / visits. The third-party correlation degree is in the negatives when comparing its relationship to the other variables. The other variables in the restaurant and webstore categories correlation values have a positive degree which shows they work in comparison with each other when it comes to the customer and revenue of the restaurant.

**#4 Data Set Support Analysis:**

Using the data information, Bubba Gump main form of revenue comes from its customers spending money at the Bubba Gump’s webstore over money spent at the restaurant’s location whether it be merchandise or food. When reviewing the data through visualization, I can see an area where a decrease in revenue can be found. The three categories restaurant visit, website visit, and third-party visit have high numbers when the customer visit’s the restaurant or website for the first time but overtime the customer stops coming or gradually comes less back to the restaurant. The lack of customer returns is most likely an issue with the current product being presented to the customer whether it’s the food or merchandise. We can see the food, or the merchandise is not enough to make the customers desire to return. Also, an area the data causes shortcoming in this survey is the detailed area of the product being sold to the customer. I believe the survey needs a to add additional columns to discover which area is lacking in satisfactory level with the customer, which causes the customer from the desire to return. The data does not break down the satisfactory level of the food or merchandise. I can only speculate this information, but I believe the food is below satisfactory level. I can speculate this information since the webstore is Bubba Gump’s main form of revenue instead of the restaurant, but I cannot 100% stand by this statement due to the lack of information.