Joseph Silva Jr

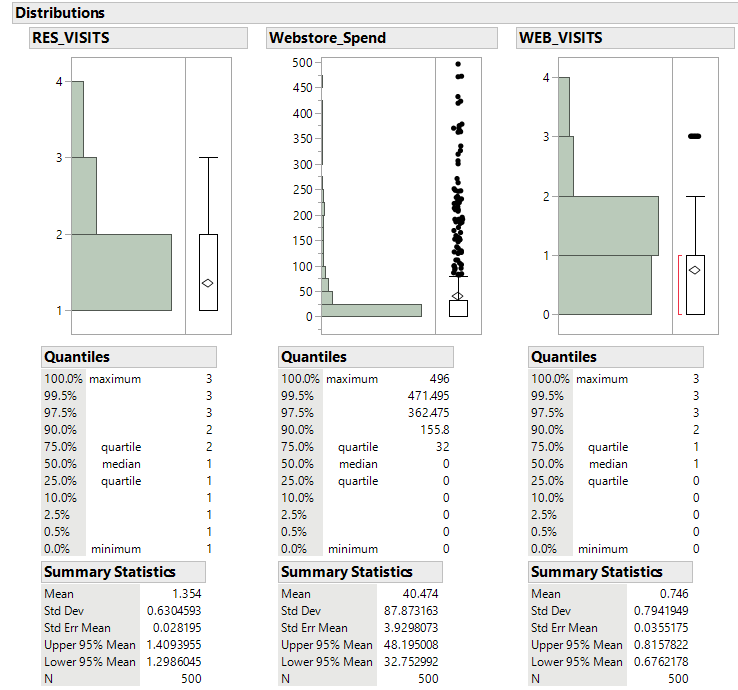
10/17/21

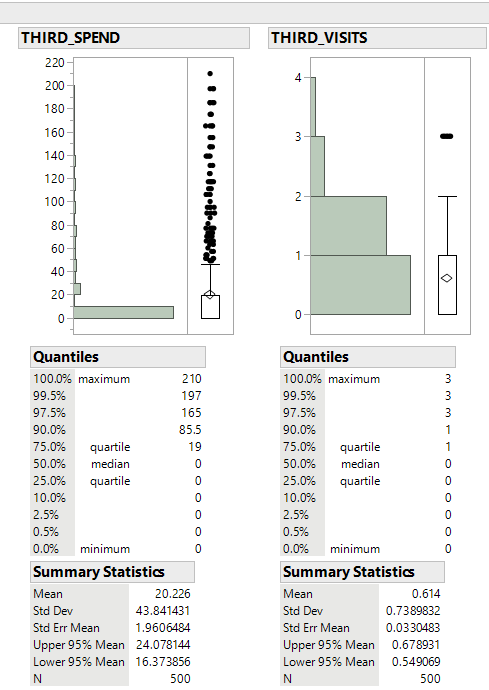
SNHU

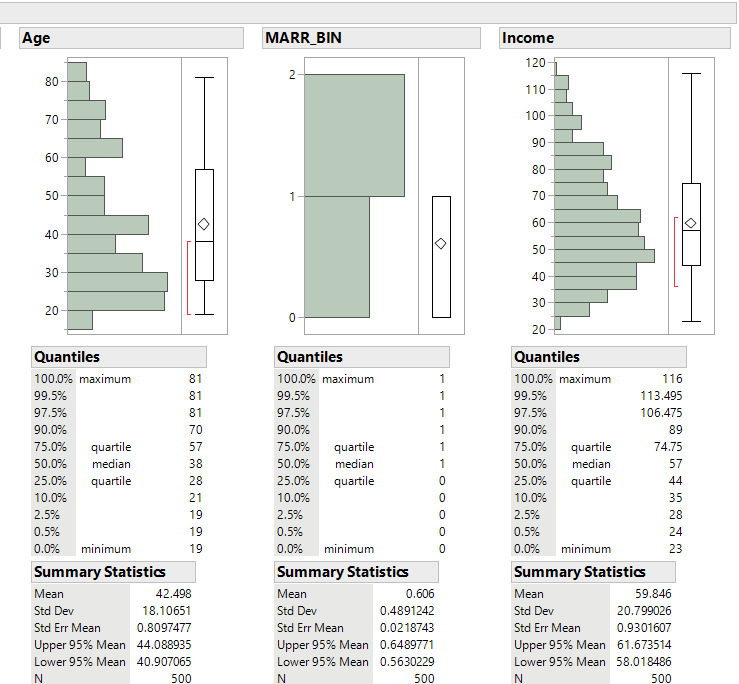
DAT 220: Milestone 4

**Display and Interpretation:**

Data Set Survey Visualizations:

****

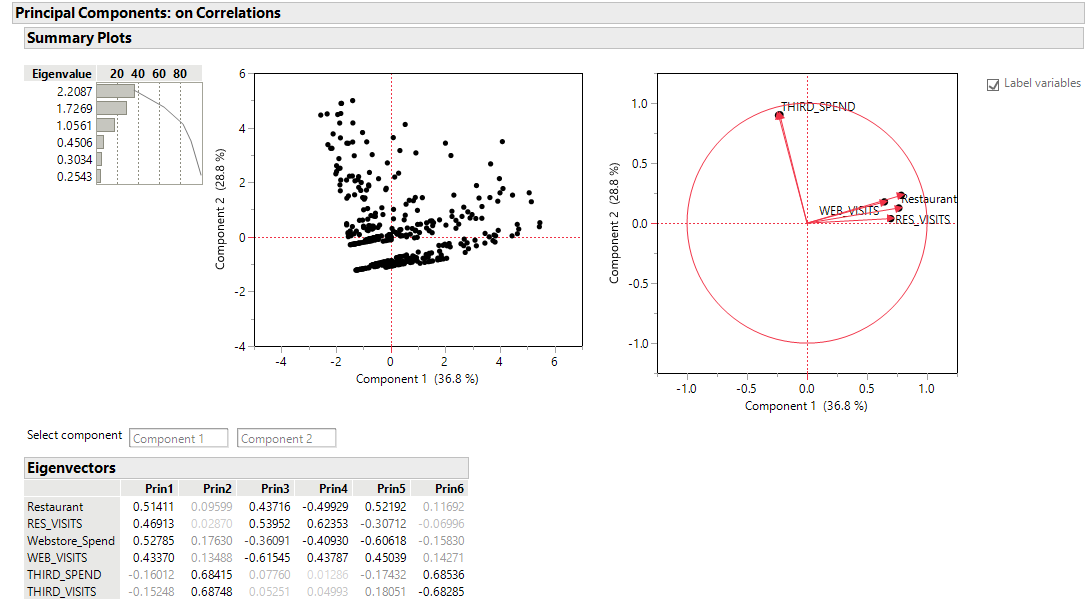
****

****

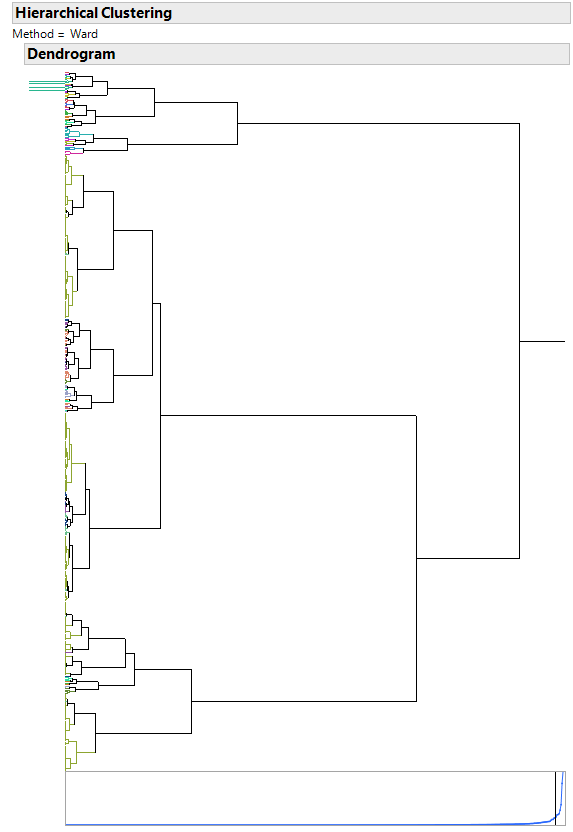
With these graph designs, we can see a visualization of the data taken for the survey. We can see customers tend to go to the webstore to but merchandise instead of using third-parties or visit the restaurants. The webstore can be shown to be the main source of revenue for Bubba Gump. I can also see 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.

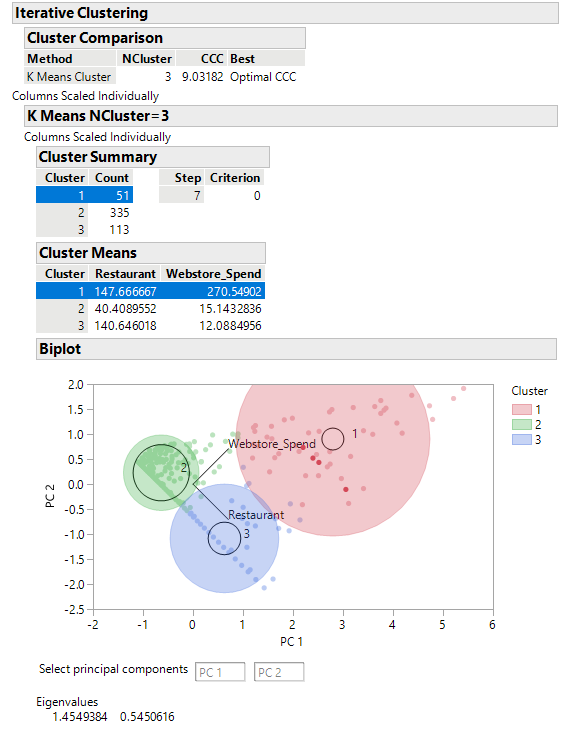
Correlations and Associations:

****

****

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.

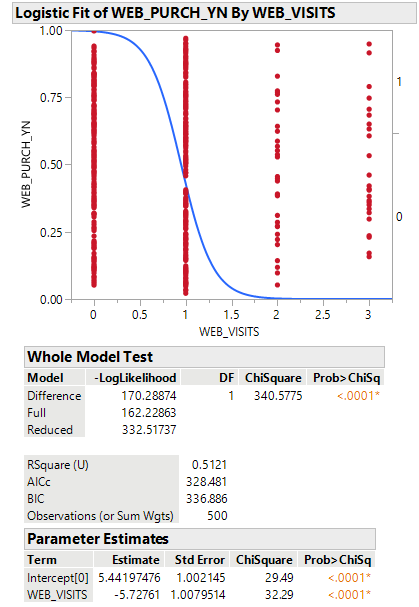




I generated a Hierarchical Cluster and a K Means Cluster using the variables of Web store spend and Restaurant (Spend). The natural amount of clusters developed from these two variables is 19 clusters in the dendogram. While looking at the K Means Cluster data, we can see the web store spends generates the more purchases than the restaurant spends from the data collected by the customers.



I compared the Web store spending variable to two different characteristics of customers. The characteristics were the age and income of the customer. Looking at the graphs, we can see there are not many grouping of data points and they do not appear to line up with the linear fit line. We can also see in the summary of fit the RSquare values for both graphs. The RSquare results show how close the data is related to the fitted regression line. The RSquare is measured from 0% to 100% with 0% being a variability having no explanation or relation to the response data, while 100% means the variability has an explanation or relation to the response data means. Looking at the RSquare values, I believe age and income have no relations with the variable Web store spending.



Comparing the Web visit and the web purchase variables, we can see the graph visualization. We can see a lot of the customers do not purchase items from the web store because they do not visit the web store. Next, the graph shows there is a 50% chance a customer will purchase an item when they visit the web store for the first time. According to the graph, most customers’ data retrieved for the survey are in the area of visiting the web store either 0 to 1 time. However, looking at the graph we can see that the few customers that visit the web site 2+ times usually have a 100% web store purchase.

**Validity, Reliability, and Limitations:**

The graphs and explanations above allow us to see the validity, reliability, and limitations of these visualizations whether it be the graphs or the numbers. The graphs show how certain characteristics did not affect web store purchases by using the Rsquare numbers. We were also able to see the reliability by comparing web store and in-store purchases revenue. From prior data, I was able to tell the revenue from the webstore was higher than in-store purchase. The graphs and numbers were able to show my hypothesis was correct. Next, the limitations of these graphs were shown when I discovered information was limited by comparing the variables of web store purchases vs the number of visits to the webstore by customers. The data was limited in showing the number of purchases were made during those visits by each customer.

**Resulting Decision Influence:**

I would display the results to a customer or business by showing a comparison of each of their forms of revenue. This comparison will allow a creation of a list showing the lowest and highest forms of revenue. I will then be able to break down which areas need to be cut or modified and which areas need to be pushed forward. These break downs in data will allow companies to reallocate their spending to benefit the company to prevent any further loss and to increase any profit.

**Visual Evaluation:**

While looking at the data graph visualizations, I was able to determine to see if any variables affecting other variables such as the age of customer vs the web store purchases from the customer. I was able to also see the data comparing the web store purchase vs in-restaurant purchases from customers and the data showed the web store purchases brought in more revenue to Bubba Gump over in-restaurant purchases. In the graph with the variable of web store visits, I believe there was an error when it came to a web store purchase from a customer because it showed a 100% purchase after 2+ visits, but the data needs to be able to show if a customer purchased an item on the web store for each visit or a purchase was made for the web store visits in total. A ¼ purchase out of 4 visits is a big difference compared to a purchase being made for each visit of the 4 visits.

**Next Steps:**

The next steps to help Bubba Gump restaurant would be separating the areas that obtain profit from lowest to highest. I would then break down each path of revenue into further detail such as comparing customers’ characteristics with the in-restaurant purchases like it was done with the web store purchases. Breaking down these areas of revenue, the details from the data will be able to show where the company needs to modify, pursue, or dismantle the areas affecting revenue.