

Assignment 10: Cluster Analysis

1. Use the Sample – Superstore data to
 - a. Construct a scatter plot using Discount and Quantity (with level of detail as Customer). Cluster the data based on (Avg) Discount, and (Avg) Quantity. *The default in the Number of Clusters textbox is "Automatic." Do not change this.* Call the sheet with Clusters "Store Clusters." Change the shape of the hollow circles in the scatter plot to filled or solid circles.
 - b. Create a dimension called Segments using the Clusters Variable created above. Construct a bar chart using Profit Ratio and Segments. Color bars by segment and express thickness using segment size (for this dataset, it is the same as number of records).
2. Use the World Indicators data
 - a. For the year 2012, cluster the data into two Segments GDP, CO2 emissions, Energy Usage, Internet Usage, and Mobile Phone Usage. Construct a bar chart to contrast the Life Expectance Female by Cluster. Represent number of observations in each segment by thickness of the bars.
3. Use JPMorgan Data for this part. To import this data into Tableau, select, "statistical file" option and then select the file. Once the file appears in the Data Source tab, change Connection from Live to "Extract." The dataset contains responses to survey questions on importance and satisfaction for a set of questions. The 12 importance questions were combined using a principle components analysis with oblimin rotation into four components: Customer Service Factor, Variety of Service Factor, Value Factor, and Performance Factor.
 - a. Cluster the data using the four components: Customer Service Factor, Variety of Service Factor, Value Factor, and Performance Factor. Change the shape of the hollow circles in the scatter plot to filled or solid circles. Set number of clusters to 3.
 - b. Rename the clustering variable generated above as "Customer Segments". Construct a bar chart to show the differences between the Customer Segments in terms of Demographic Variables in the data (age, dependent children, education, marital status), and number of records. You could either create a bar chart for each demographic variable or create one for all variables using 'side-by-side bars'

Submission Guidelines

1. Save the Tableau file as a "Tableau Packaged Workbook." File → Save As and then change file format to "Tableau Packaged Workbook". The resulting file should have an extension .twbx.
2. Publish your Tableau workbook to Tableau Public. For this to work, your data must be an Extract not Live.
3. Submit the Tableau packaged workbook and URL for Tableau Public workbook.