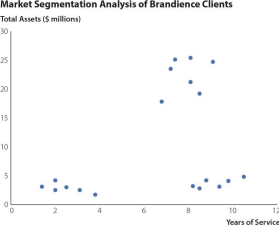
**DS 620 – Week 3 Assignment**

**Ritesh Kumar B**

1. Brandience Marketing LLC provides marketing analytics consulting for clients. For one of its clients, Brandience has been asked to perform a market segmentation study for a business client that provides auditing services to manufacturing customers. The client believes there are two variables of importance that should be used to group similar customers into clusters: Years of Service with the Client and Total Assets. Brandience plans to use a clustering algorithm to group similar customers into different clusters, but before applying the algorithm, Brandience creates a simple scatter chart to plot each customer based on their Years of Service with the Client and Total Assets. The scatter chart created by Brandience follows.



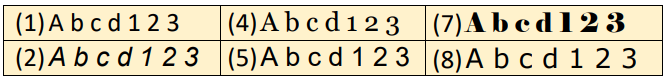
**a. Based on the Gestalt principle of proximity, how many different groups of customers are shown in the scatter chart?**

Based on the scatter plot above we can see that there are 3 different groups/clusters of customers shown.

**b. What other Gestalt principle could be used to reinforce the appearance of the different groups of customers for the audience?**

Gestalt's law of continuity can also be used to reinforce the appearance of the different groups of customers for the audience. This law states that elements in a line or a curve are perceived to be more related than those not on the curve. Here, the points close to each other can be seen as being on a curve. Hence, it can be stated that they are more related.

2. **Serif versus Sans-Serif Fonts**. Consider the fonts that are shown in the table that follows. LO 5





**a. Which of these fonts are considered serif fonts?**

Figure 3,4,6 & 7 are serif fonts

**b. If you are creating text for a data visualization that will mostly be viewed on smartphones and other small mobile devices, would you suggest using the serif fonts for text in the data**

**visualization? Why or why not?**

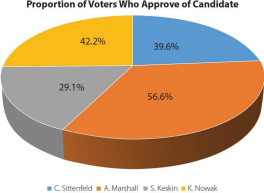
We will use sans serif fonts for text on small screen devices as sans serif typefaces are preferred for data visualization and they are simple and straightforward. It is considered more legible at small sizes, considering that mobile screens are small, this can be scaled easily for mobile screens.

3. Approval Voting Results. Approval voting is a type of voting system in which voters can vote for any number of eligible candidates. A vote for a candidate in this type of system indicates that the voter “approves” of that candidate for the position the candidate is seeking. In the final tally of the election results, the total number of approval votes for each candidate is calculated. The candidate who receives the most approval votes is declared the winner. The figure below displays the results from an approval voting election in which there were four eligible candidates: C. Sittenfeld, A. Marshall, S. Keskin, and K. Nowak. A total of 1218 voters participated in this election. The table below shows the number of votes received by each candidate as well as the proportion of voters who approved of each candidate. LO 4, 6

**Proportion of Voters Who**

**Candidate Number of Votes Received**

**Approve of This Candidate**

C. Sittenfeld 482 39.6% A. Marshall 689 56.6% S. Keskin 354 29.1% K. Nowak 514 42.2% 

a. **The figure displays the election result proportions as a 3D pie chart. What problems are associated with displaying the election results in this way?**

Pie Charts are not recommended for displaying more than 2 - 3 categories, but here we see 4. Humans just naturally aren’t very good at distinguishing differences in slices of a circle, especially not at a glance; we’re much better equipped to notice differences in rectangular shapes. They often distort the information and make it more difficult for decision-makers to understand the messages.

**b. Suggest an alternative type of chart for displaying these data that would be easier for the audience to interpret. Explain why this alternative type of chart would be easier to interpret.**

Bar Graph is a better alternative for this kind of representation as our brains understand information/data better in rectangular shapes. Also, Bar Graphs have an axis for us to help understand/quantify the data. Hence, Bar Graph would be the best alternative for representing this kind of data.

4. **Inbound and Outbound Shipping Costs at Rainbow Camping**. Rainbow Camping makes outdoor equipment for camping and backpacking that it sells online. Because Rainbow Camping mails each of its sales from its distribution center direct to the consumer, it spends considerable funds each month on outbound shipping. Rainbow Camping also has costs related to inbound shipping to get its products from the manufacturers to its distribution center. Rainbow Camping would like to examine its outbound and inbound shipping expenses over the last 12 months (labeled months 1 through 12). These shipping costs are shown in the following table.

**Shipping Costs ($1000s)**

**Month Outbound Inbound**

1 478 324

2 524 274

3 628 224

4 787 292

5 648 348

6 612 309

7 598 378

8 641 287

9 546 292

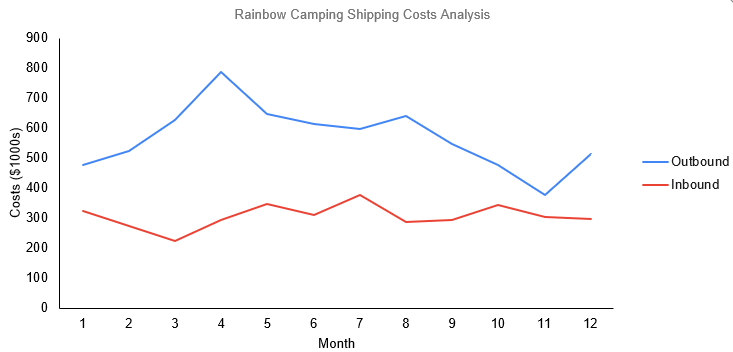
10 476 345

11 378 304

12 512 298

**a. Using the data in the file *RainbowCamping*, a single chart that displays both the outbound and inbound shipping costs for each month on the same chart. Use the title “Rainbow Camping Shipping Costs Analysis” for the chart title, “Costs ($1000s)” for the vertical-axis title and “Month” for the horizontal-axis title. Format the chart title and vertical axis title to make these easy for the audience to interpret. Modify the Minimum and Maximum Bounds of the horizontal axis so that this axis starts at 1 and ends at 12 (Minimum and Maximum Bounds can be found in the Format Axis task pane by clicking on the Axis Options icon and then under Axis Options.) Remove the markers so that only lines are displayed for each of these shipping costs. Delete any unnecessary gridlines in the chart.**

**b. To further minimize the eye travel required by the audience, delete the default legend created with the chart and use text boxes to label the lines in the chart as “Outbound” and “Inbound”.**

****

5. **Funstate Carnivals Staffing Analysis**. Funstate Carnivals operates four large amusement parks, each of which is located in California. Funstate would like to examine the staffing across these four amusement parks. It is most interested in comparing the total number of staff employed at each park, but it is also interested in comparing the gender breakdown of the staffing at each park. The table below shows the total number of staff employed at each park, broken out by gender.

**Total Number of Staff**

**Location Male Female**

Fresno 72 84

Sacramento 89 61

Long Beach 65 84

Oakland 48 51

**a. Using the data in the file *Funstate*, create a stacked bar chart to display these data. Use different colors for Male and Female. Use the chart title, “Funstate Carnivals Staffing Analysis,” choose appropriate axes titles, and include a chart legend to identify Male versus Female in the bar chart. Format the chart to minimize eye travel and remove any unnecessary gridlines to increase the data-ink ratio.**

**b. Funstate would like to show the exact numbers of Male and Female staff at each location on the chart. Add data labels to the stacked bar chart to display the number of Male and number of Female staff at each location. Format the data labels so they are easy for the audience to read.**

**Because of decreased revenues, Funstate wants to implement a policy of no more than 125 total staff members at each location. Draw a line and use a text box to indicate “Staffing Goal = 125” on the stacked bar chart to show this limit. (*Hint*: Click Insert on the Ribbon, then click Shapes in the Illustrations group to add a line and text box.) Which locations currently exceed this staffing limit?**

Chart, waterfall chart

Description automatically generated