Name: Milin Desai

Subject: Data Visualisation (DSC 465)

Date: 04/12/2020

Direction on how I used Tableau:

Uploaded Datasheet, made modification needed in the data set then drag them in to the column and raw as per the requirement. Used colour option, different plot option s from show me and I think almost same for all graph for question 1, 2 and 3.

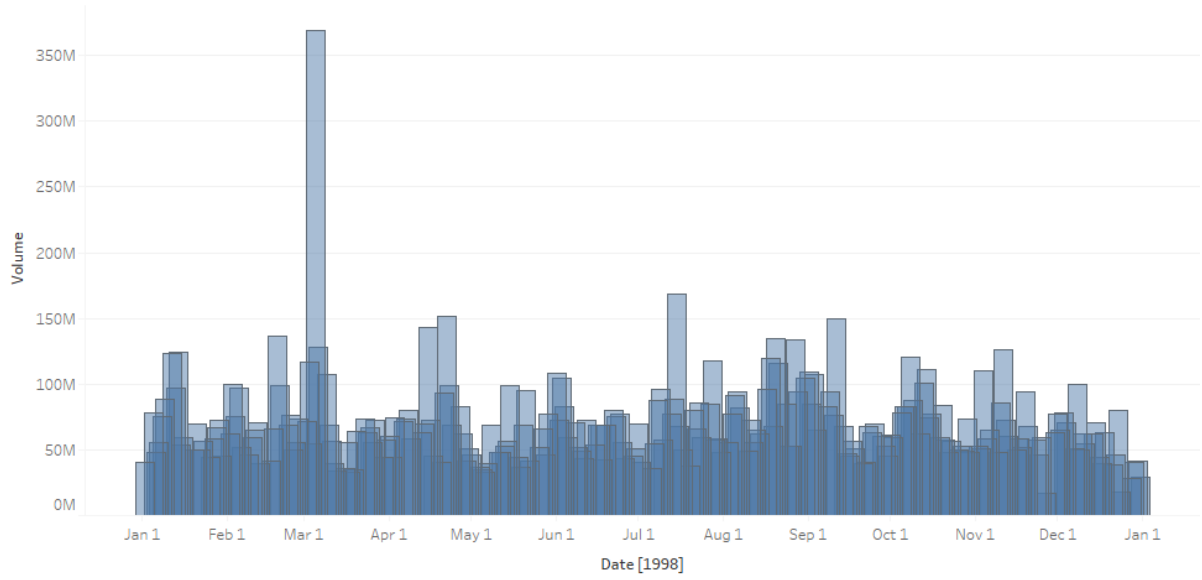
Question 1)

Answer:

I have used Tableau for this question you can find the graph attached as below

From the graph you can see, closing dates from Intel stocks of 1998 from Whole year January to January. You can see from the graph that closing price was low in the starting which fluctuating up and down between month February to September then after closing price of Intel Tax increase until the end of the year with minor fluctuations.

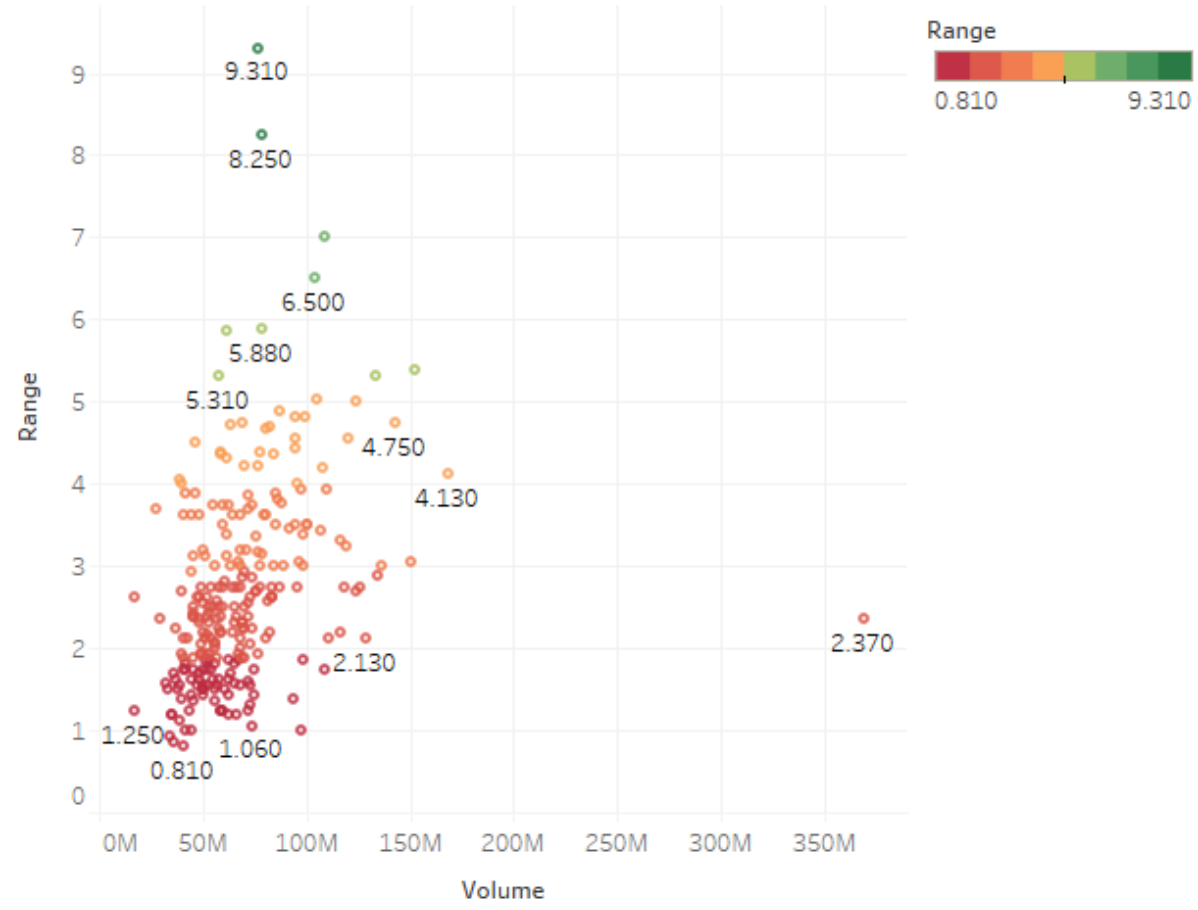
1 B)



The plot of sum of Volume for Date.

From the graph above you can see volume vs exact dates from Intel stock price. From the graph you can see that volume of shared stock on particular date time. You can see that on March 1 it was highest among all the volume of day which is almost 360 million.

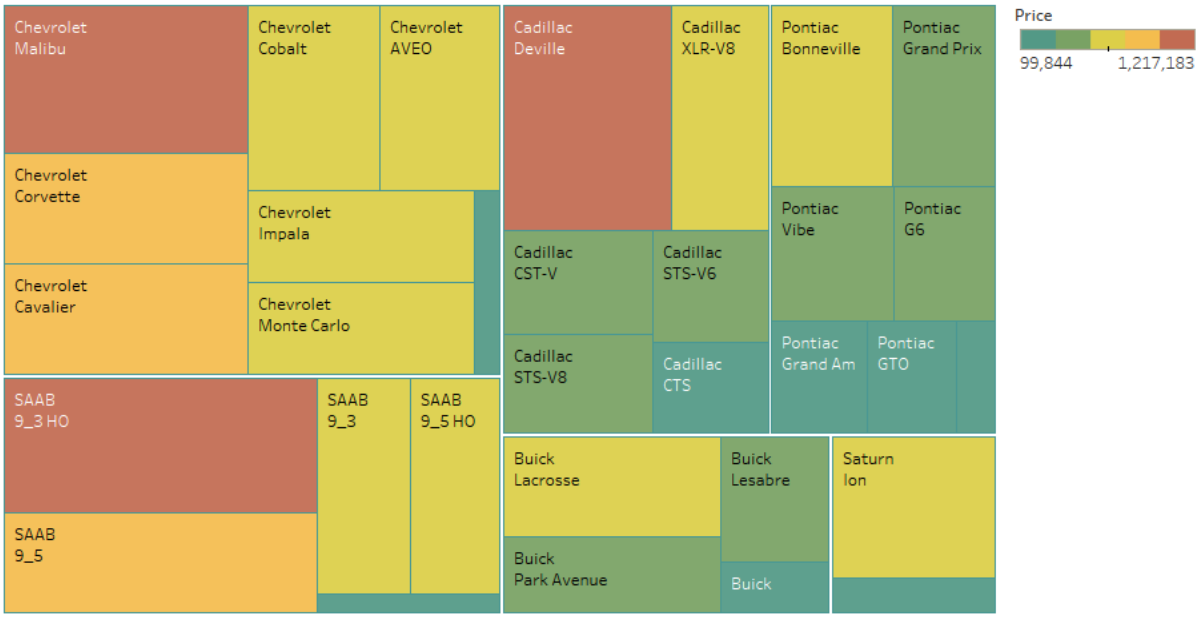
1.C)



Volume vs. Range. Colour shows Range.

➔ This is the volume vs Range, where range is difference of high and low as you can see in the graph that you can see that small range is high volume where there is very less high range example in the low range. Where for high volume there is only on example of high volume in low range.

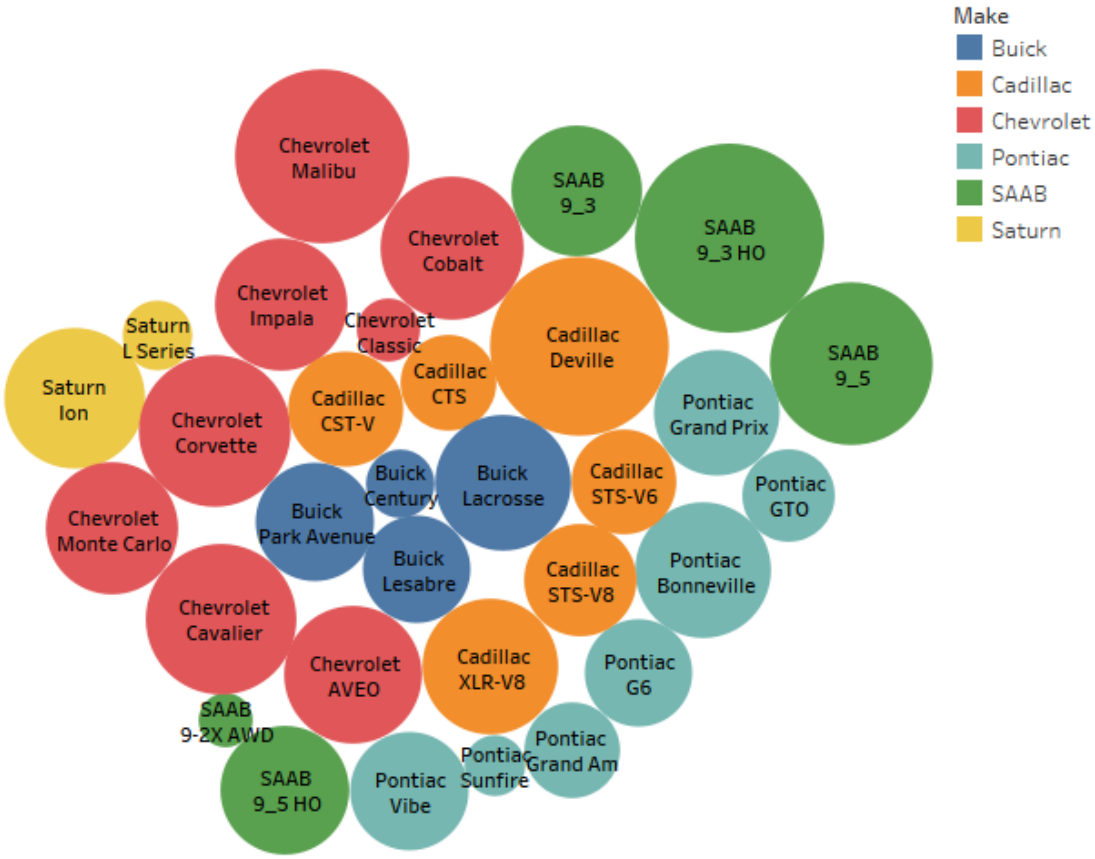
2.A) Price variation in car by model and company



Make and Model. Colour shows sum of Price. Size shows sum of Price. The marks are labelled by Make and Model.

- ➔ You can see in the graph that price of different car with name and model which you can easily decode with colour code.
- ➔ With green colour it's a cheap car and with red are most expensive. You go to the price colour code to know about difference price of the Cars Company and model wise.

2.B) Bubble plot for dominent car company



Make and Model. Colour shows details about Make. Size shows sum of Price. The marks are labelled by Make and Model.

➔ You can see this bubble and tree graph for make and model which shows details about make of the car. Also size shows that sum of price. Where same colour code is a car of same companies with different model.

2.C) Difference between two graphs:

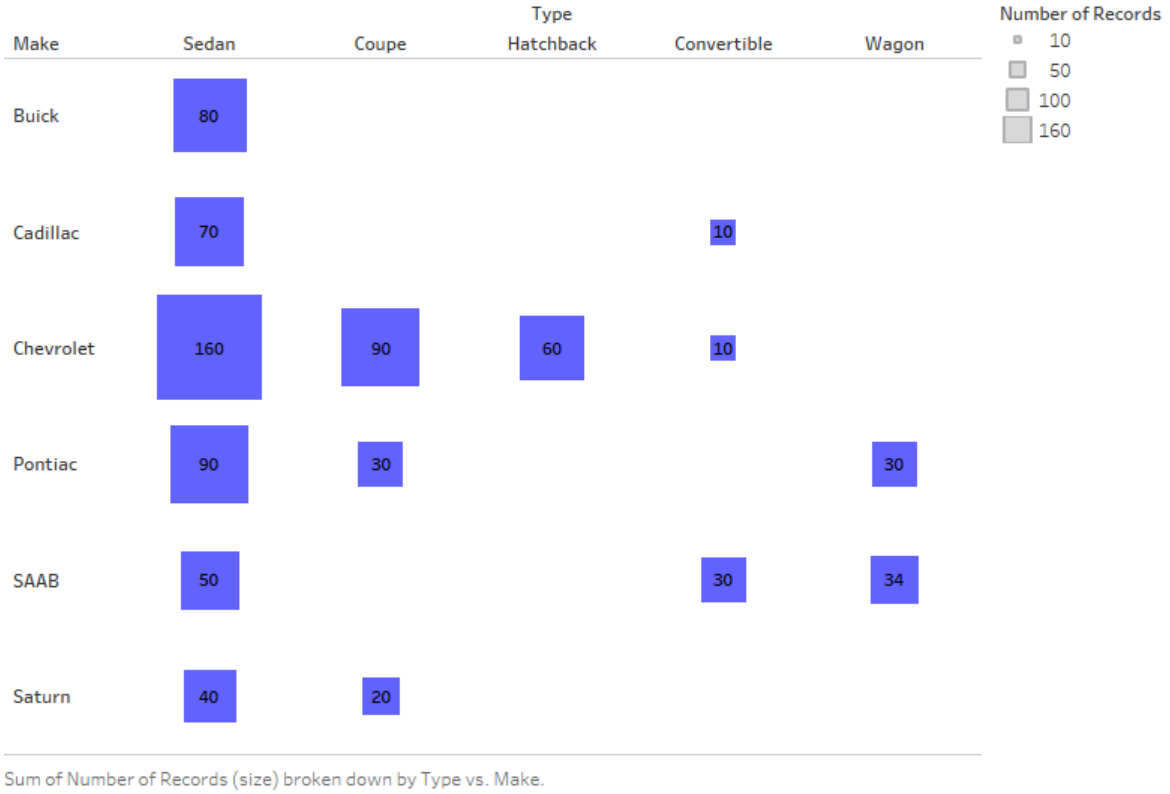
➔ Graph, first we need to understand why we use graph. To change data into visual mode now here our motive is that customer or the audience can understand all the information in less time yet it can be impactful and attractive.

➔ Both graphs satisfy that condition but here we have to make it more efficient how one can understand easily and infer more from the graph if you see graph which is bubble graph. it is easy to understand, you can see that group of the car from same company as they share same colour. And price of the model-type by the size of the bubble

➔ Where In the first chart Prize of the car is shown by colour if it is red colour than it is costly car where totally opposite if the colour is green and they are couple variance of price as colour varies.

- ➔ So those are the main difference between the bubble graph and the other graph. In my opinion I found bubble graph more unique and easy to understand where other graph is detailed information of the dataset.
- ➔ So as above we have explained difference between the two graphs.

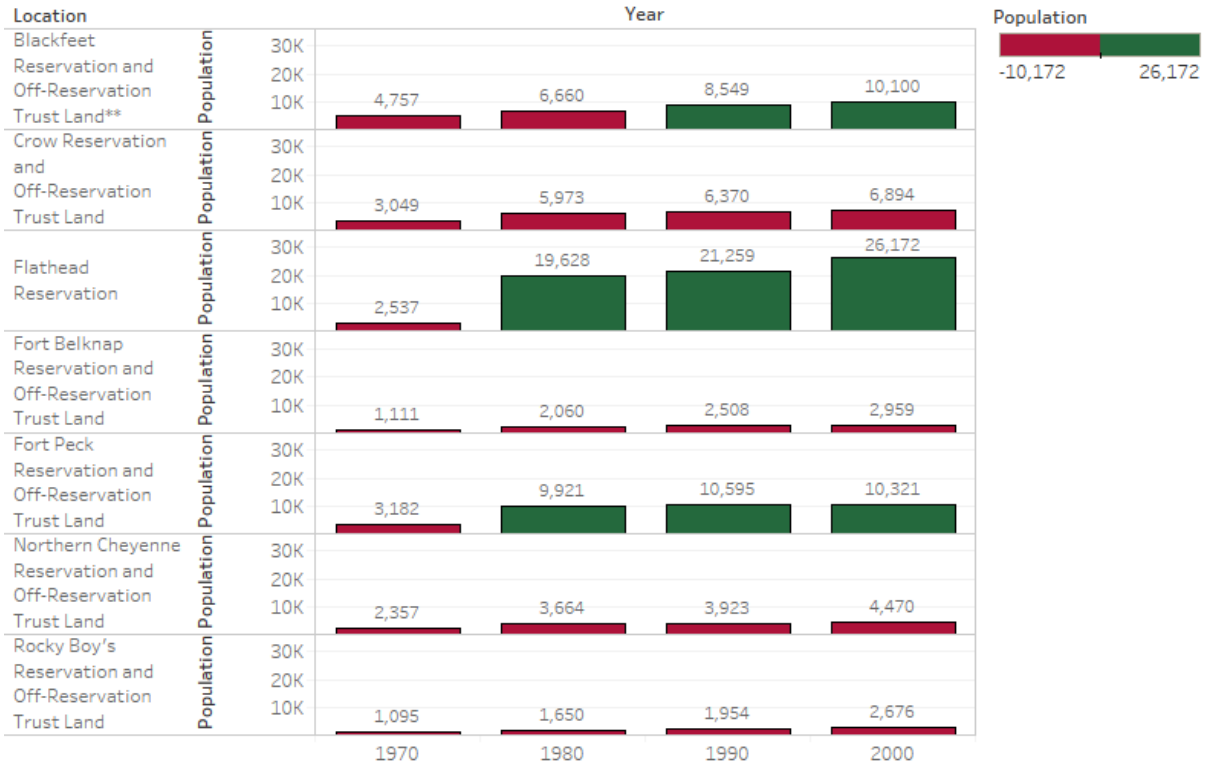
2.D) Number of cars by body type



Sum of Number of Records (size) broken down by Type vs. Make.

- ➔ From the graph you can say that sedan is most selling car in almost every company maybe be the reason is everyone like sedan than the other car type. Where people prefer less convertible cars even that type of car looks eye catch and attractive. But because of some reason people less that car.
- ➔ Where you can see the approximately 160 records of selling sedan can with Chevrolet that makes it best seller company for sedan type cars.

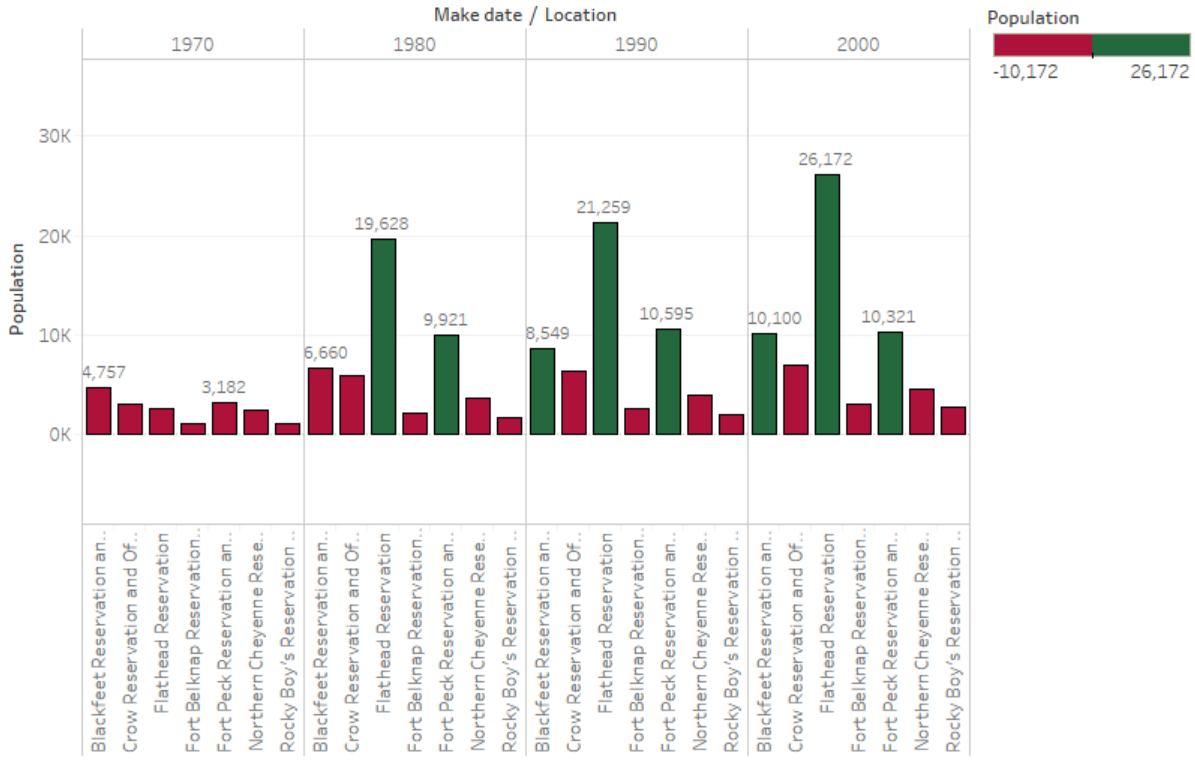
3.A) Population growth over years



Population for each Year broken down by Location. Colour shows Population. The view is filtered on Location, which excludes Montana.

- ➔ From the graph you can infer that, population growth of all reservation over 1970 to 2000. This is shown in one row for all the year so you can compare with other reservation population category.
- ➔ Where I have shown green where reservation population is more than 8000 and less for less population.

3.B) Growth over each year



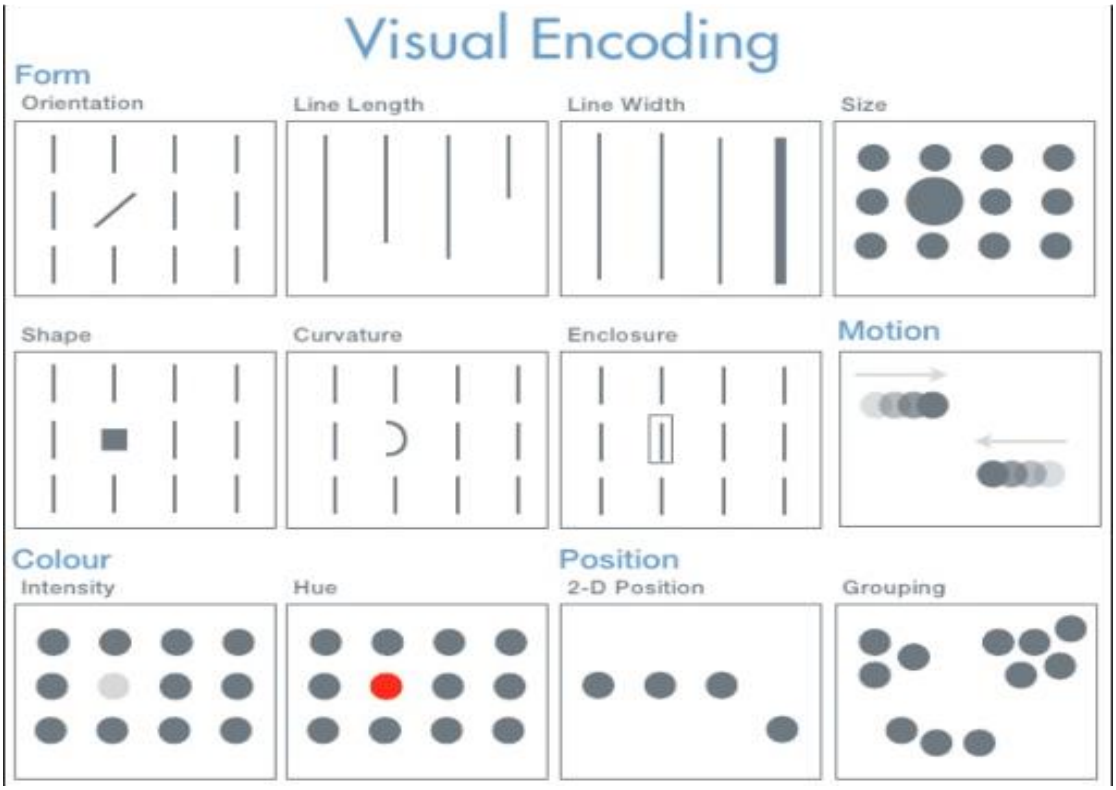
Population for each Location broken down by Year of Make date. Colour shows Population. The view is filtered on Location, which excludes Montana.

- ➔ Where this graph is little is different than the 3.A) where you can track population growth of all reservation population for individual years. Same for population you can see that number in the right side corner number above 8k I have noted as green.
- ➔ Hence, from the graph you can understand the reservation population form different years in 1970-2000.

Question 4:

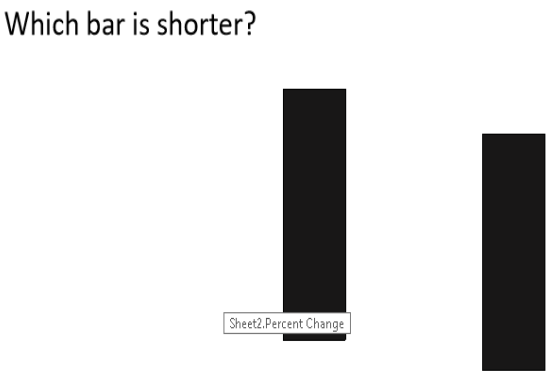
A) Pre-attentive:

- ➔ By pre-attentive attributes we mean we can distinguish or we can see the difference in the visual data with in fraction of seconds. Typically, we mean around 200ms it is that easy to catch no need to think a lot, very simple patterns. You can see some example of pre-attentive attributes as below:

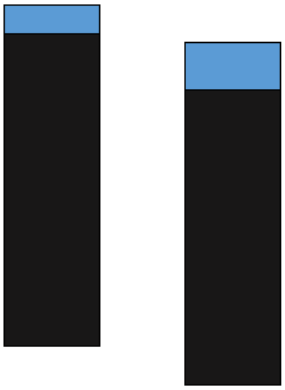


B) Weber's Law:

- ➔ Weber's law says that difference between two stimuli is to proportion to the magnitude to the stimuli, In addition he also says that human perceive easily if you use percent change.
- ➔ one of the important example we have seen in the class was, bar and the line. So when numerical value is not aligned it is difficult to judge the data, which one is short but if you use percentage remove from the line you can understand and judge the data. And can jump to conclusion that which one is shorter and vice versa.
- ➔ However, It is really important to start the numerical data with the same baseline preferably zero so that it is easy for audience or say human perception can easily judge that. Also that will be better for audience.



➔ It is really hard to say which one is short because both dosent share same baseline, if they have baseline zero then that will be really easy to guess.



➔ But human understand by percent change.

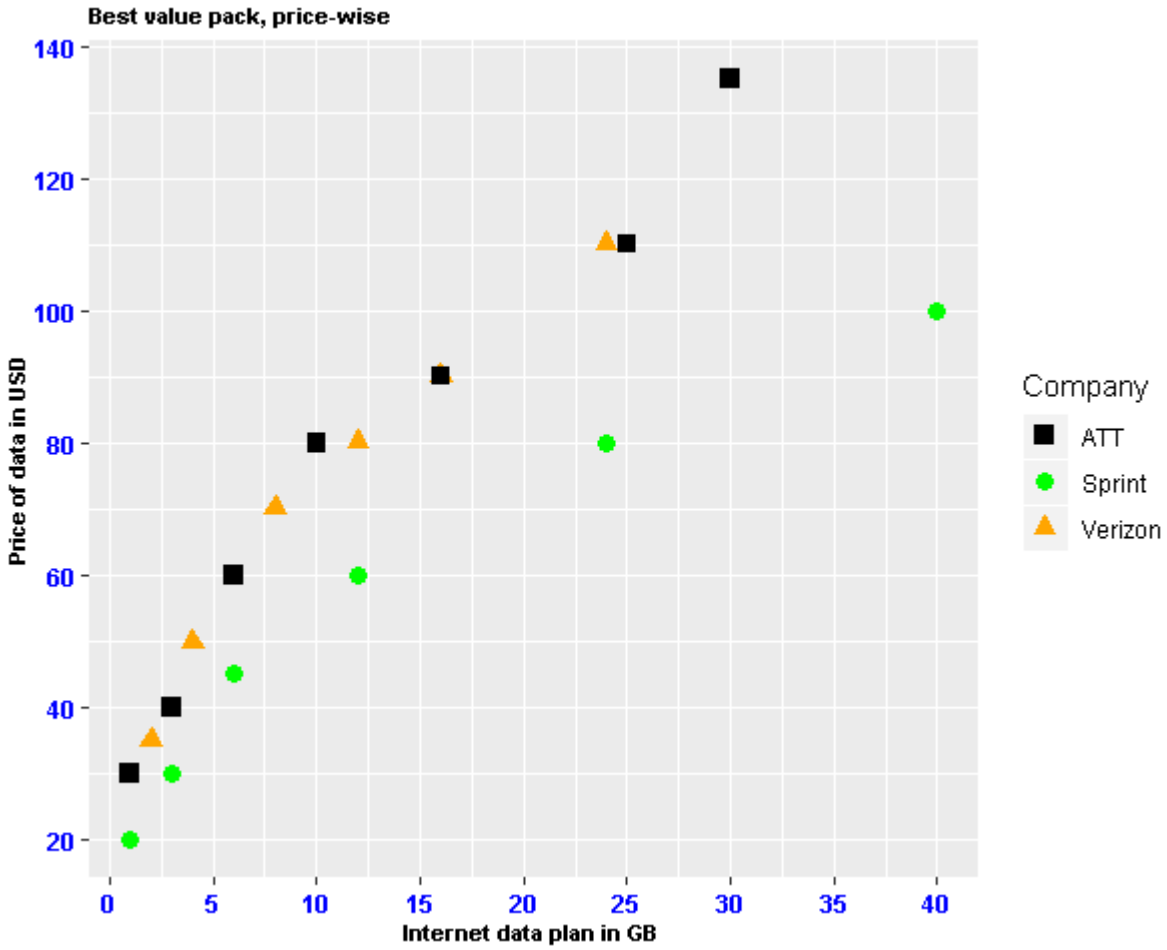
Question 5:

Graph 1: Scatter plot for company data plan price wise:

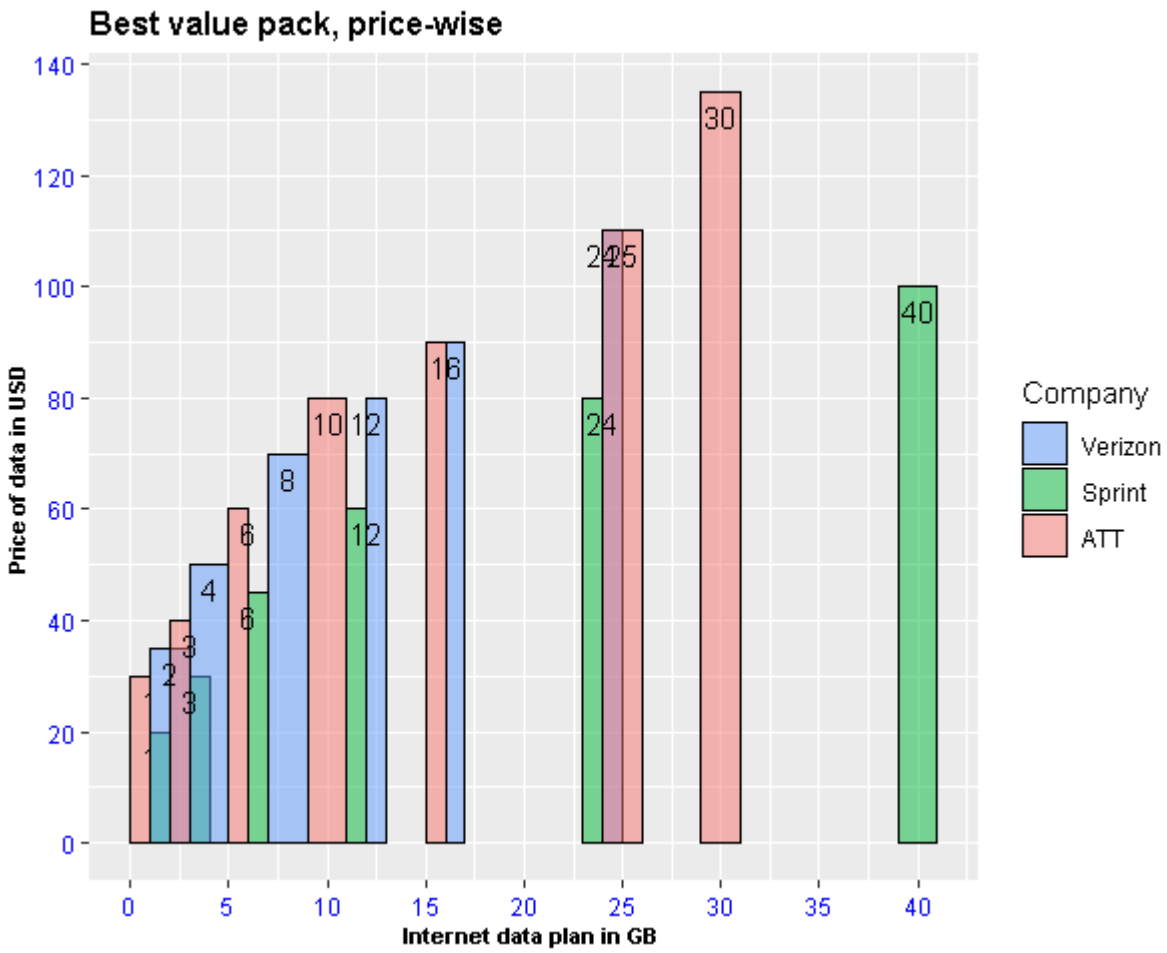
➔ Price, we all know whenever we buy something we budget it or price plays important factor in the purchasing so here I am making graph of different data plan in GB for different company where price you can check it from left side always.

➔ You can see sprint plans are comparatively cheaper than the ATT and Verizon. also more GB data in less price

➔ I have used different shapes for companies, and have also scale for the graph in the increment of 20 for price and scale of 5 for data. I have also change axis appearance and title of the graph as well. Code for this graph I have submitted at the end of the document. Mainly use geom_point for scatter plot and geom_col for bar graph.



Graph 2 Bar Plot for best value pack, price wise:



- ➔ For this graph I've add label as well for data plan, here I have kept colour changes from the scatter plot just to differentiate graph. Where, Cookbook suggest to keep the colour same. But I am making this exception.
- ➔ From the bar graph you can see many data plans are overlapping in the starting range of data plan as you go further for high range and large data plan there are large price difference you can see in the graph.
- ➔ In addition, For bar plot graph you can find the data plan you want from the plot and for interested plan you can check the company and price for the same plan .
- ➔ So second graph make it easier for people who are more interested in data rather than price
- ➔ As above we have explained two graphs for selecting best data plan in lowest price from different companies. The graph explains that very clearly.

R Code:

Code for Question 5.A)

```
#Loading libraries
library(tidyverse)
library(ggplot2)
library(dplyr)
library(mosaic)
library(lubridate)
library(gcookbook)
library(MASS)
#cellplans dataplan price for diffrent companies
cellPlans = data.frame(
  c("ATT", "Sprint", "Verizon", "ATT", "Sprint",
    "Verizon", "ATT", "Sprint", "Verizon", "ATT",
    "Verizon", "Sprint", "Verizon", "ATT",
    "Verizon", "Sprint", "ATT", "ATT", "Sprint"),
  c(1, 1, 2, 3, 3, 4, 6, 6, 8, 10, 12, 12, 16, 16,
    24, 24, 25, 30, 40),
  c(30, 20, 35, 40, 30, 50, 60, 45, 70, 80, 80, 60,
    90, 90, 110, 80, 110, 135, 100))


names(cellPlans) = c("Company", "DataGB", "Price")
head(cellPlans, 5)
#creating scatter plot for Data in GB vs Price for Individual company all together
ggplot( data=cellPlans, aes(x=DataGB, y=Price, label = Price,shape=Company,
color=Company) ) +
  geom_point(size=3) +
  xlab("Internet data plan in GB ")+  #naming X and Y axis
```

```
  ylab("Price of data in USD") +
  scale_shape_manual(values = c(15,16,17)) + #allloting diffrent shapes for diffrent companies
  scale_colour_manual(values = c("Black", "green", "orange"))+ #alloting colors to identify diffrent company's plan and price in the graph
  scale_y_continuous(breaks = seq(0,150,20))+ #increasingscale numbers so audience can easily understand data
  scale_x_continuous(breaks = seq(0,50,5))+
  # Ashtetic touch to the title, font, colors and lable of graph
  ggtitle("Best value pack, price-wise")+ #giving title
  theme(
    axis.title.x = element_text(color = "Black", size = 8, face="bold"),
    axis.text.x = element_text(color = "Blue", face="bold"),
    axis.title.y = element_text(color = "Black", size = 8, face="bold"),
    axis.text.y = element_text(color = "Blue", face="bold"),
    plot.title = element_text(color = "Black", size = 8, face="bold")
  )
```

5.B) Code for question 5.B:

```
#Loading libraries
library(tidyverse)
library(ggplot2)
library(dplyr)
library(mosaic)
library(lubridate)
library(gcookbook)
library(MASS)
#cellplans dataplan price for diffrent companies
```

```
cellPlans = data.frame(
  c("ATT", "Sprint", "Verizon", "ATT", "Sprint",
    "Verizon", "ATT", "Sprint", "Verizon", "ATT",
    "Verizon", "Sprint", "Verizon", "ATT",
    "Verizon", "Sprint", "ATT", "ATT", "Sprint"),
  c(1, 1, 2, 3, 3, 4, 6, 6, 8, 10, 12, 12, 16, 16,
    24, 24, 25, 30, 40),
  c(30, 20, 35, 40, 30, 50, 60, 45, 70, 80, 80, 60,
    90, 90, 110, 80, 110, 135, 100))

names(cellPlans) = c("Company", "DataGB", "Price")
head(cellPlans, 5)

#making recipes for catchy bar graph
ggplot(cellPlans, aes(x = DataGB, y = Price, fill = Company)) +
  geom_col(position = "dodge",color="black", width = 2, alpha = .5 )+
  geom_text(aes(label = DataGB), vjust = 1.5, color="Black")+
  guides(fill = guide_legend(reverse = TRUE))+
  xlab("Internet data plan in GB ")+
  ylab("Price of data in USD") +
  scale_y_continuous(breaks = seq(0,150,20))+
  scale_x_continuous(breaks = seq(0,50,5))+
  #ylim(0,140)+
  #xlim(0,50)+
  # geom_text(vjust = 1)+
  ggtitle("Best value pack, price-wise")+
```

```
theme(
  axis.title.x = element_text(color = "Black", size = 8, face="bold"),
  axis.text.x = element_text(color = "Blue"),
  axis.title.y = element_text(color = "Black", size = 8, face="bold"),
  axis.text.y = element_text(color = "Blue"),
  plot.title = element_text(color = "Black", size = 12, face="bold")
)
```