

**CSE – 3020**

**Data Visualization**

**Theory DA**

**Name : Anish Desai**

**Reg. No. : 20BCE0461**

**Slot : D1**

**Guided by : Prof. Jyotismita Chaki**



**VIT<sup>®</sup>**

**Vellore Institute of Technology**

(Deemed to be University under section 3 of UGC Act, 1956)

## Question :

Create a dashboard using the MTCARS dataset. Dashboard should contain at least 10 pages. At least two of the pages should contain some interactivity. Write proper interpretation.

## Output :



Page 1 : Chart 1.1

Chart 2.1

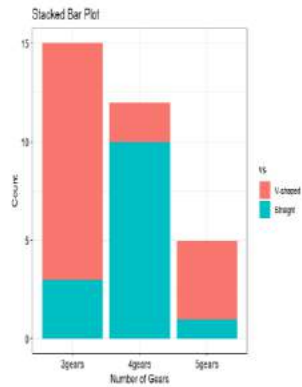
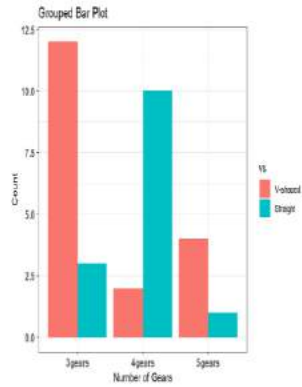
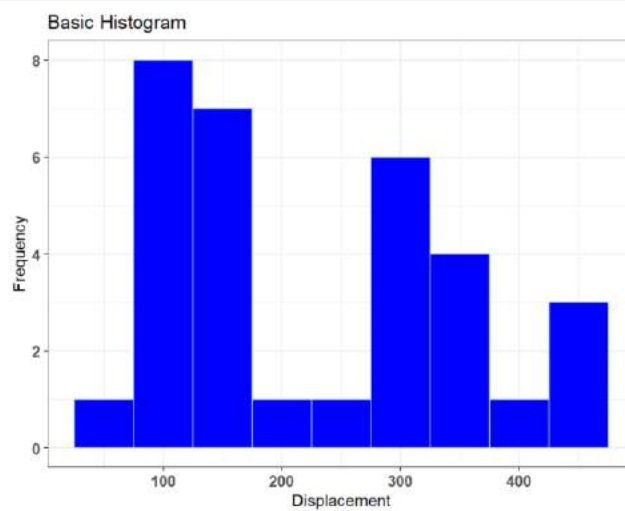


Chart 2.2

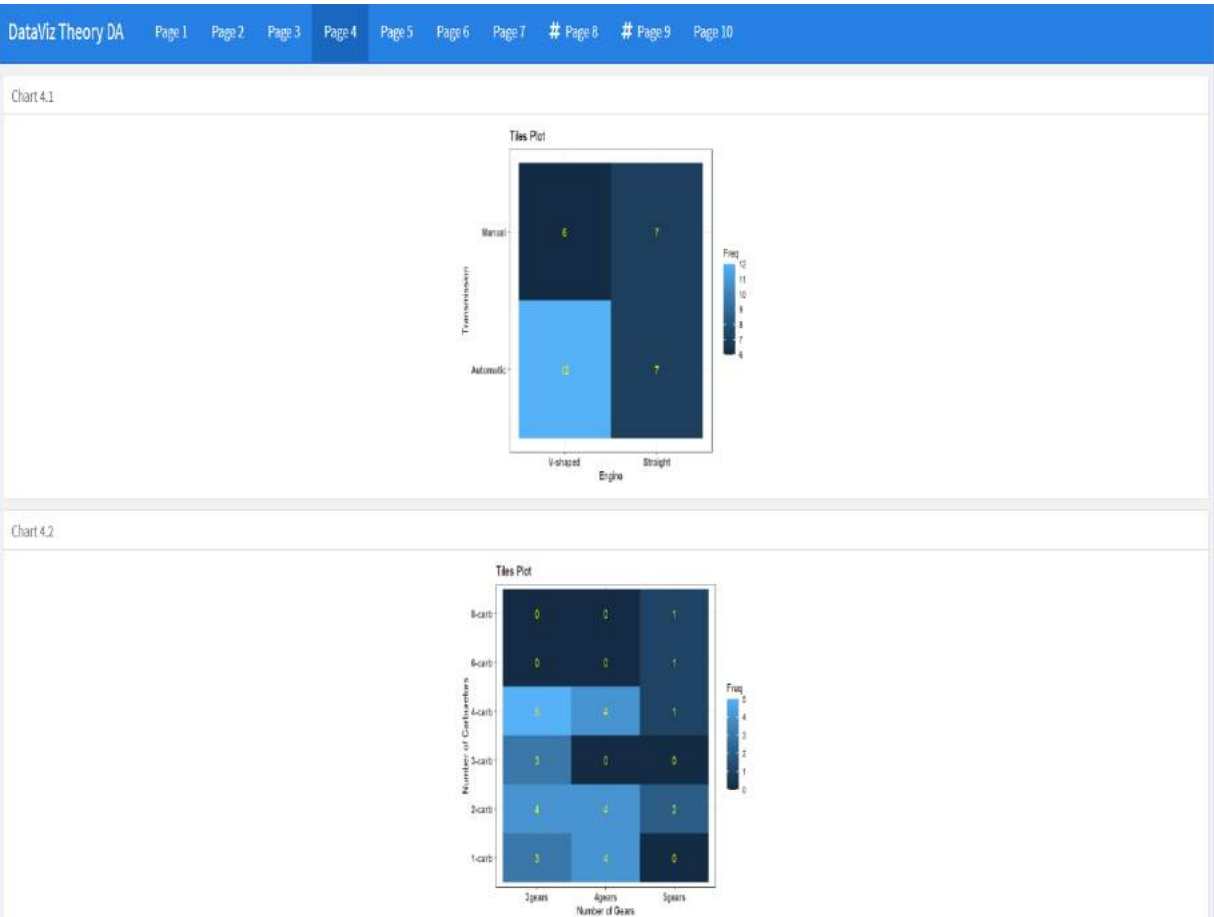


Page 2 : Charts 2.1 and 2.2

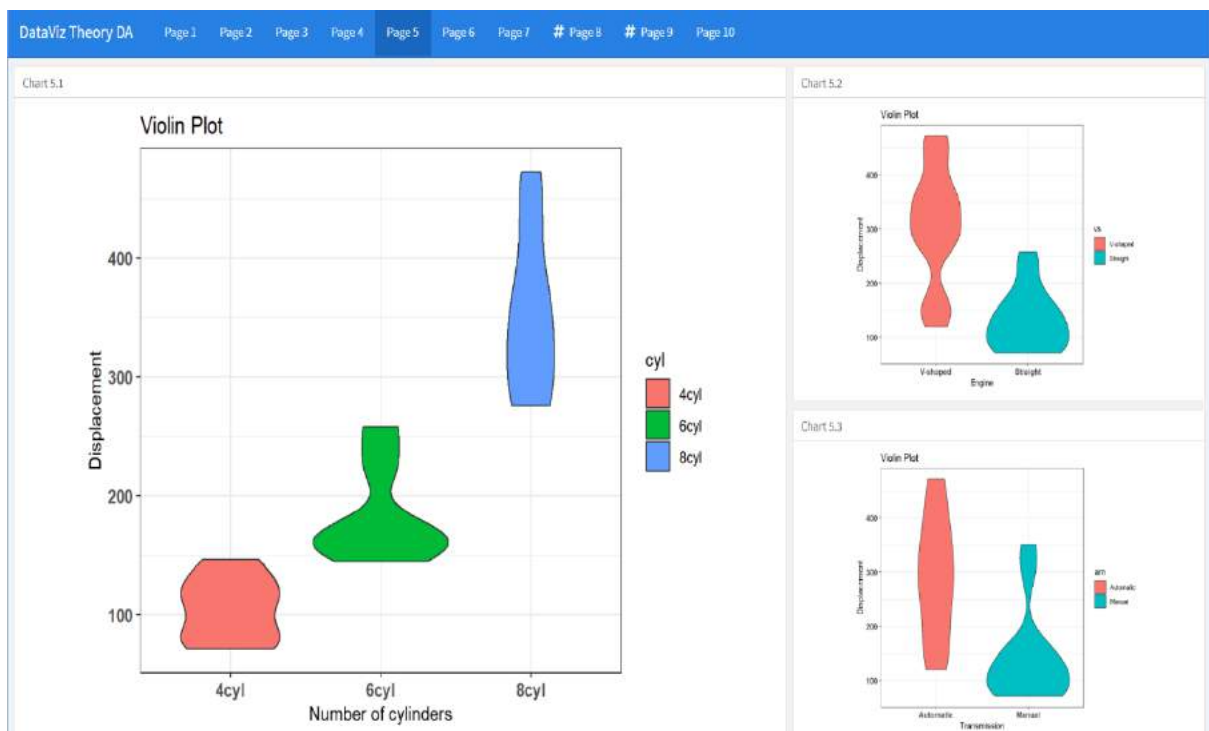
Chart 3.1



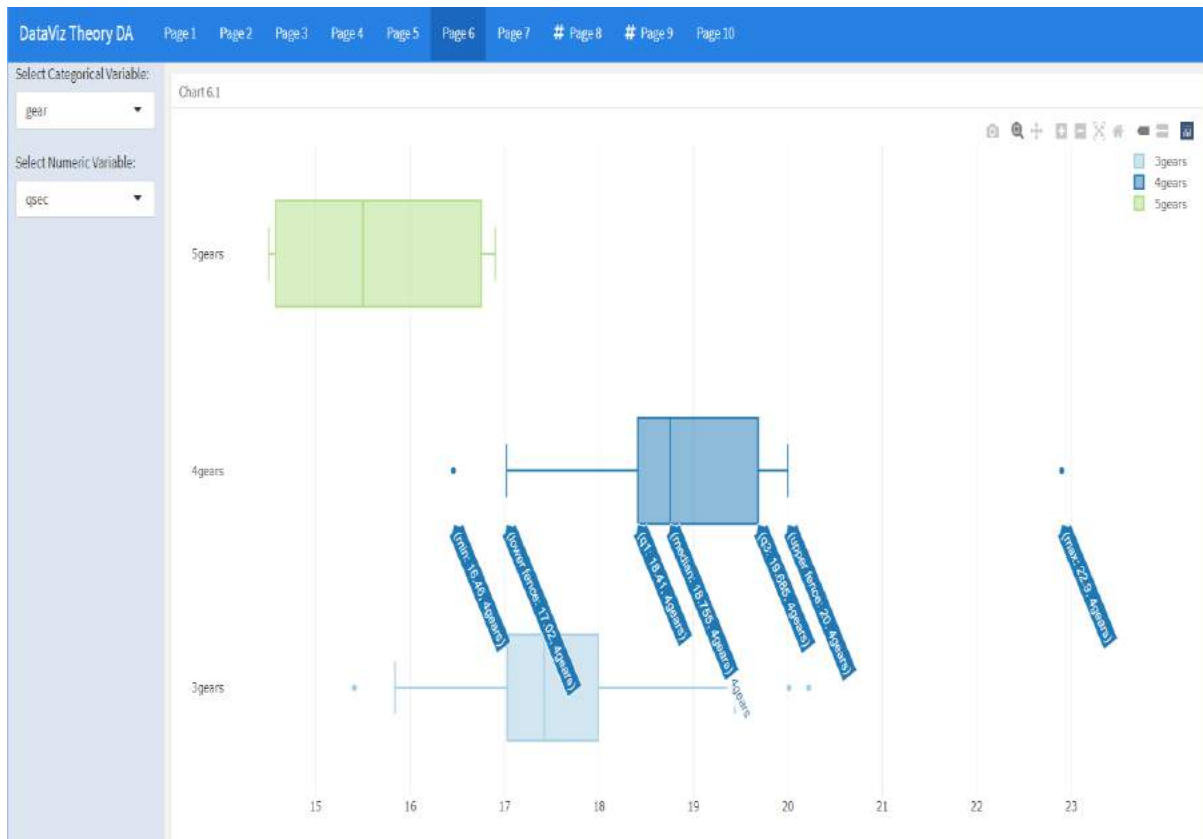
Page 3 : Chart 3.1



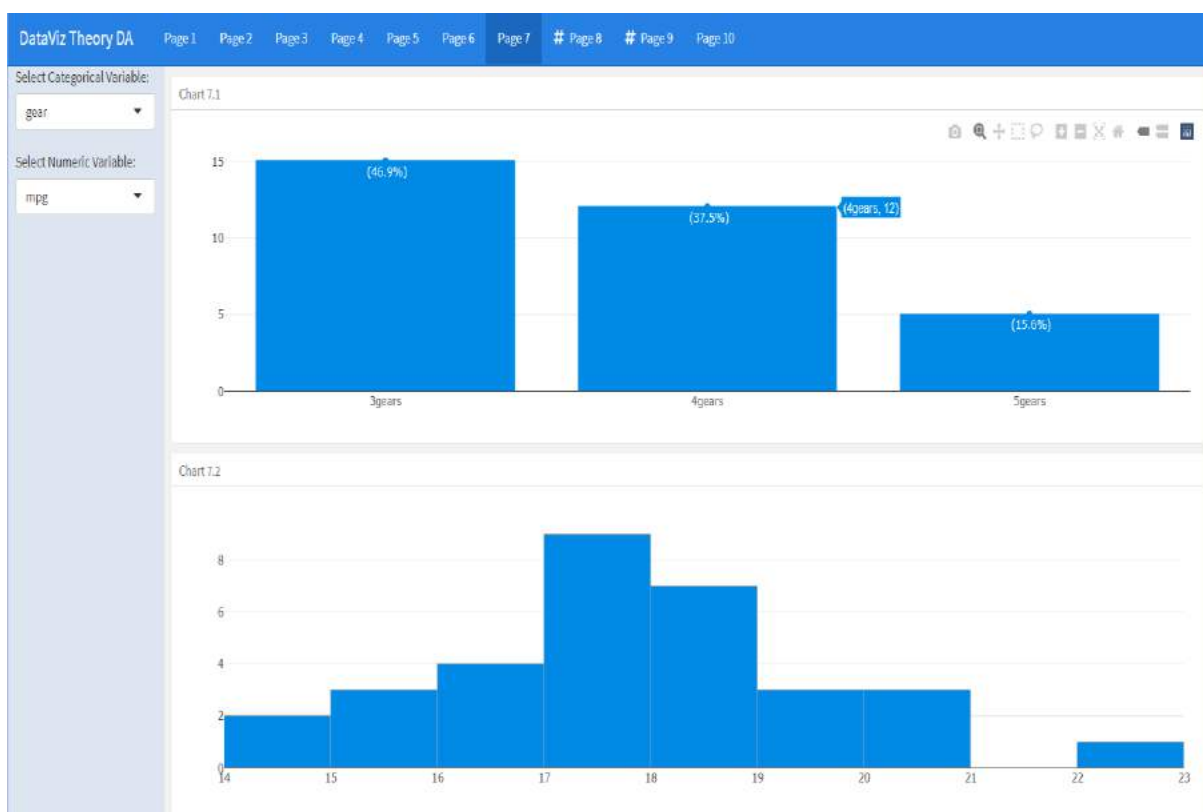
Page 4 : Charts 4.1 and 4.2



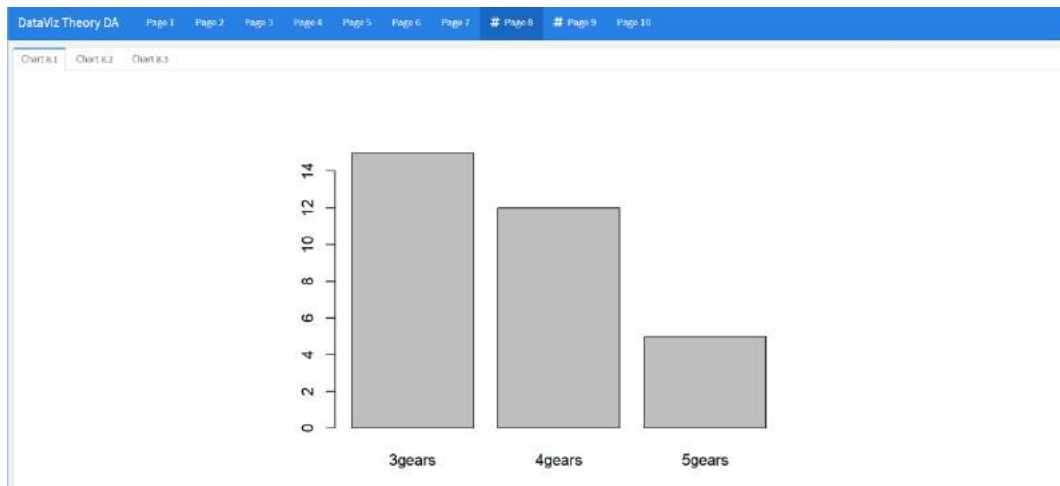
Page 5 : Charts 5.1, 5.2 and 5.3



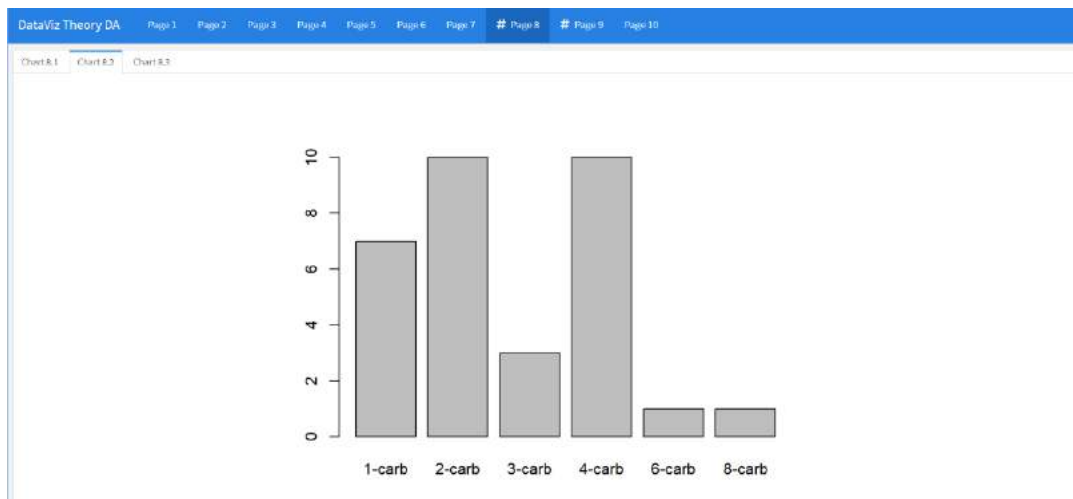
Interactive Sidebar Page 6 : Chart 6.1



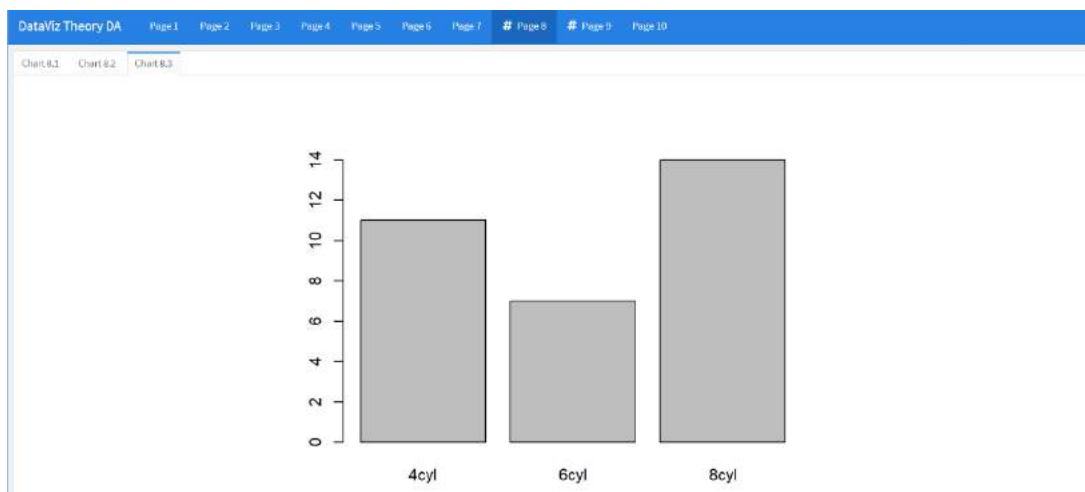
Interactive Sidebar Page 7 : Charts 7.1 and 7.2



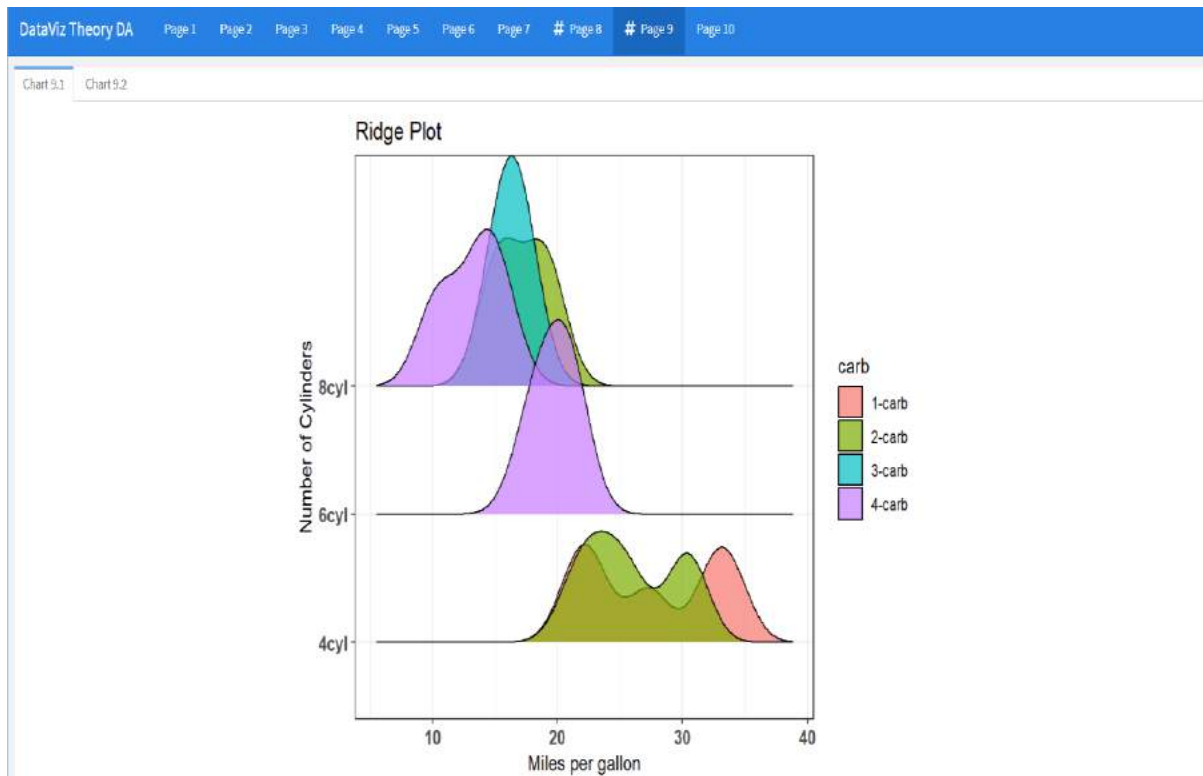
**Interactive Tabset Page 8 : Chart 8.1**



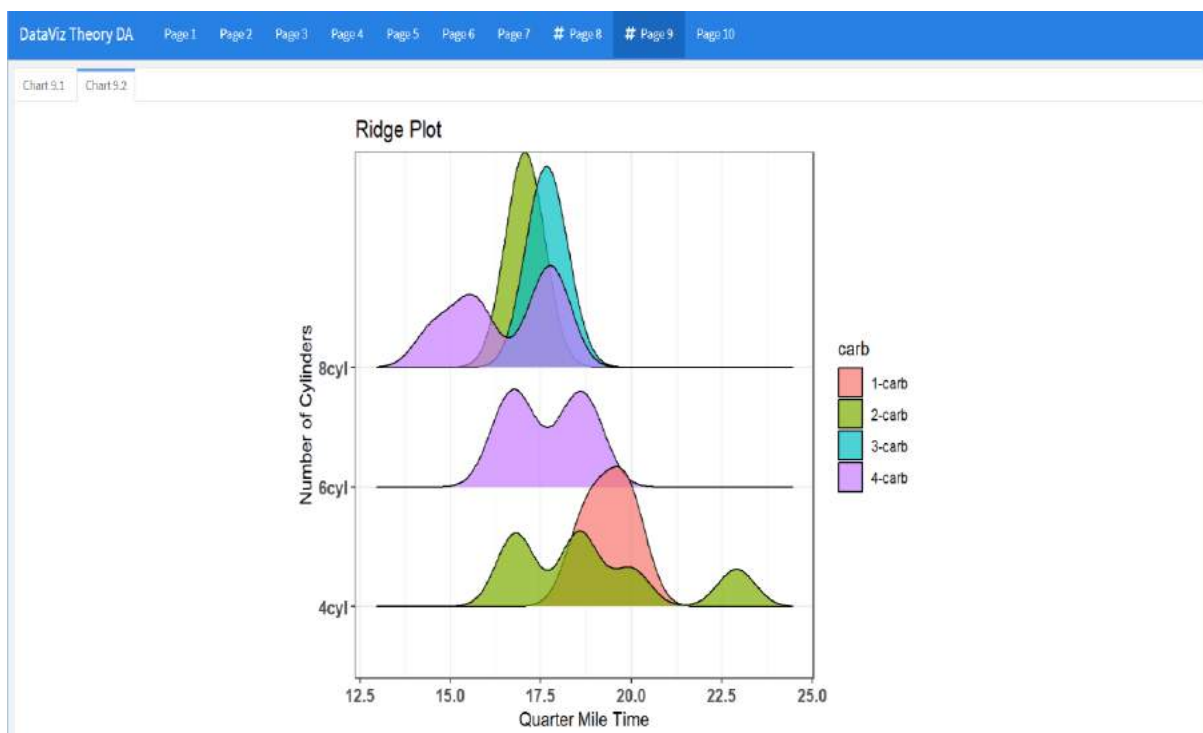
**Interactive Tabset Page 8 : Chart 8.2**



**Interactive Tabset Page 8 : Chart 8.3**

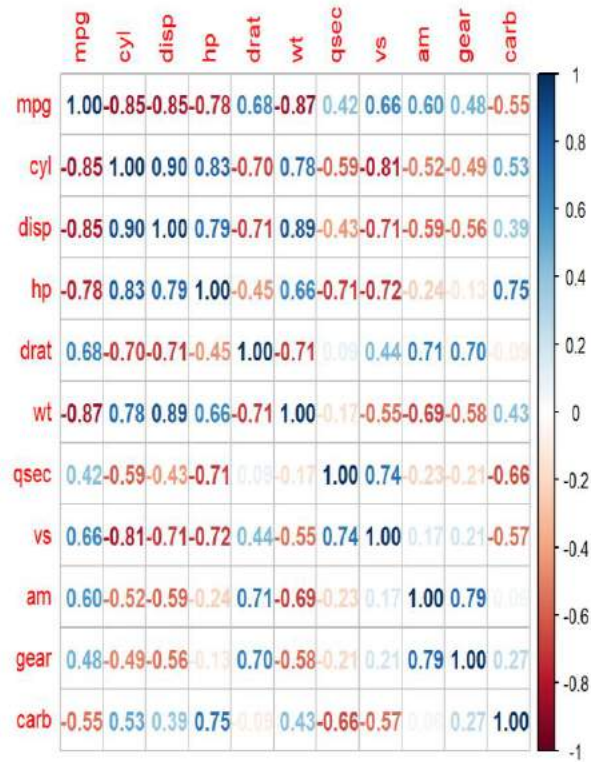


Interactive Tabset Page 9 : Chart 9.1



Interactive Tabset Page 9 : Chart 9.2

Chart 10.1



Page 10 : Chart 10.1



## Interpretation :

### Data Visualization CSE 3020

#### Theory DA

Anish Desai

20BCE0461

DI

#### # Interpretation

Using flexdashboard, many various types of plots have been visualized. For this, features and observations of 'mtcars' dataset has been used.

In the first 5 pages of the dashboard, various kinds of Multivariate analysis graphs have been plotted, such as bar plots - basic, stacked and grouped, histogram, tiles plot, violin plot, etc..

The first page has a layout of only one chart. This is a basic model of flexdashboard. In this chart, a bar plot has been plotted for average displacement against number of gears.

The dataset has cars which have either 3, 4 or 5 gears. The average displacement of cars with 3 gears is the highest ( $\sim 3500$  cu.in) while those of cars with 5 gears is the least ( $\sim 1100$  cu.in).

The second page has a split column layout. The first chart shows a stacked bar plotted for the count against number of gears on the basis of engine. As we can see from the graph, the count of 3-gear cars is the highest while that of 5-gear cars the least. Among the 3-gear cars, those possessing V-shaped engine rather than straight engine are in majority. While it is the same case for 5-gear cars, 4-gear cars are just the opposite with them having straight-shaped engine in majority. The same has been visualized using Grouped Bar Plot. It shows the count against number of gears for each engine separately.

The third page visualizes a histogram for displacement of cars. The displacement of 8 cars is in the range 75-125 cu. in. which is the highest count for a given range. The displacement of majority cars are in the range 75-175 or/and 275-375 cu. in. (25 out of 32 cars have their displacement in this range only).



The fourth page has a split column (chart stack) layout housing two tile plots. The first tile plot is visualized for transmission vs engine. Majority cars have Automatic transmission and V-shaped engine ( $\sim 12$  cars). The second tile plot is visualized for Number of carburetors vs Number of gears. The maximum count of 5 cars is for 3-gear, 4-carb cars. There are no 3-gear 6-carb 8-carb, 4-gear 3, 6, 8-carb and 5-gear 1, 3-carb cars.

The numbers inside the tile shows the count; colour scheme : darker the tile, lesser is the count.

The fifth page has focal chart layout. It consists of three violin plots : No. of cyl vs disp, Engine vs disp, and Transmission vs disp of cars. From chart 5.1, we can infer that most 4-cylinders cars have displacement of around 125, most 6-cyl have around 170 and most 8-cyl cars have displacement around 325. The min. and max. displacements are in increasing order.

From Chart 5.2, we can infer that most V-shaped engine cars have displacement of about 350, the straight-shaped engines have their mode near 100. The V-shaped engine cars have greater min. and max. displacement than straight-shaped engines.

From Chart 5.3, we can infer that most automatic cars have displacement of around 300 cu. in. while most manual cars have of about 100, way less than that of automatic cars.

Pages 6 and 7 are interactive pages of the dashboard. It is achieved using library 'plotly'. Page 6 shows boxplots between any of the categorical and numerical variables selected using the tab beside. On moving the cursor over the boxplots, we can see the values of lower quartile, upper quartile, median, lower range, higher range, interquartile range and outliers. This provides complex statistical analysis in a simpler, easy-to-understand way.



Page 7 has an interactive stack chart layout. The first chart 7.1 is a bar plot which shows the count of categorical variables of the dataset, while the chart 7.2 is a histogram that visualizes the frequency of numerical variables. On moving the cursor over the graphs, the count of each variable (or) numerical range is displayed.

Pages 8 and 9 have an interactive tabset layout.

Page 8 has 3 tabs, each of which shows the count of cars on the basis of number of gears, number of carburetors and number of cylinders respectively.

Page 9 has 2 tabs, each of which visualizes 'miles per gallon' and 'quarter mile time' against the number of cylinders on the basis of number of carburetors.

From Chart 9.1, we can infer that :

1. Most of the 4-cylindereed 2-carb cars have the value of 'miles per gallon' as near 24, while 4-cylindereed 1-carb cars have the value of around 21-22 and 34. The 2-carb graph is a unimodal graph while the latter is a bimodal, thus having two peaks.
2. Most 6-cylindereed 4-carb cars have the value of around 20. (as indicated by peak)
3. Most 8-cylindereed 3-carb cars have the peak at near 15-16 miles per gallon.

The other graphs have ambiguous interpretation. The values of Chart 9.1 indicates the number of miles a car can travel for every gallon of diesel.

Chart 9.2 can be interpreted in the same way as for the previous ridge plot.

The values in Chart 9.2 indicates the time a car takes to reach a quarter mile.



From Page 10, we can infer that a correlation matrix has been visualized.

The matrix shows the values of correlation between all the pairs possible.

A positive correlation indicates that increase in one value causes an increase in the other and vice versa for negative correlation.

A correlation of 1 and -1 indicates perfectly positive and perfectly negative correlation respectively.

A zero correlation indicates that change in one variable value have no effect on the value of the other variable.

Greater the magnitude of correlation coefficient, greater is the effect of change.

Sign of value controls the direction of the effect of change.

-----Thank you-----