

Data Visualization

eusha54

1. Read and view the first 5 rows of the Data

```
library(datasets)

data(mtcars)
# loading data
head(mtcars, 5)
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2

```
# first 5 rows of data
```

2. type the following to get info about the a variable

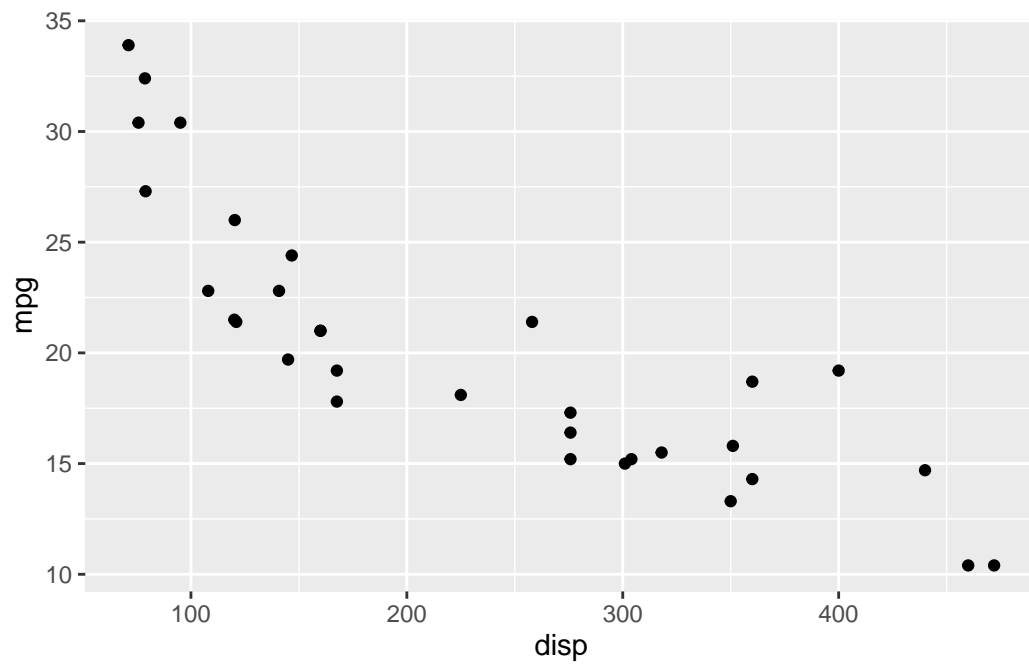
```
# ?mtcars
```

3. use the following code to create a scatterplot of two variables disp and mpg

```
library(ggplot2)
```

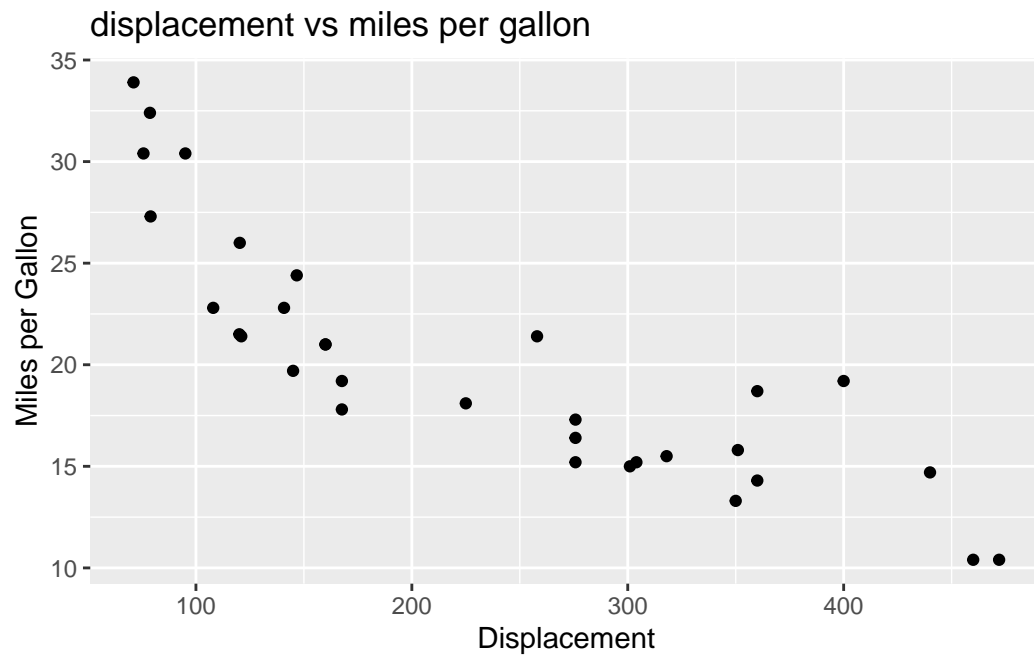
Warning: package 'ggplot2' was built under R version 4.4.2

```
# loading ggplot package  
ggplot(aes(x=disp, y=mpg), data=mtcars) + geom_point()
```



4. add title and labels to the last plot

```
ggplot(aes(x=disp, y=mpg), data=mtcars) + geom_point() + ggtitle("displacement vs miles per ga.
```

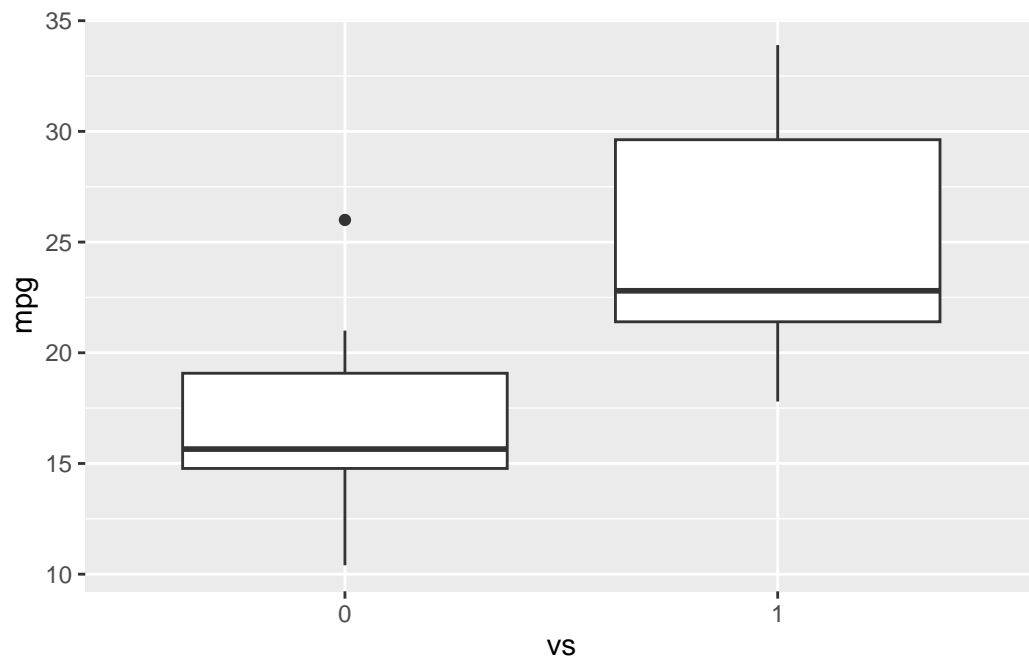


5. Use the following to create a boxplot of the the distribution of mpg for the individual Engine types vs Engine (0 = V-shaped, 1 = straight) To do this you have to make vs a string or factor.

```
#make vs a factor
mtcars$vs <- as.factor(mtcars$vs)

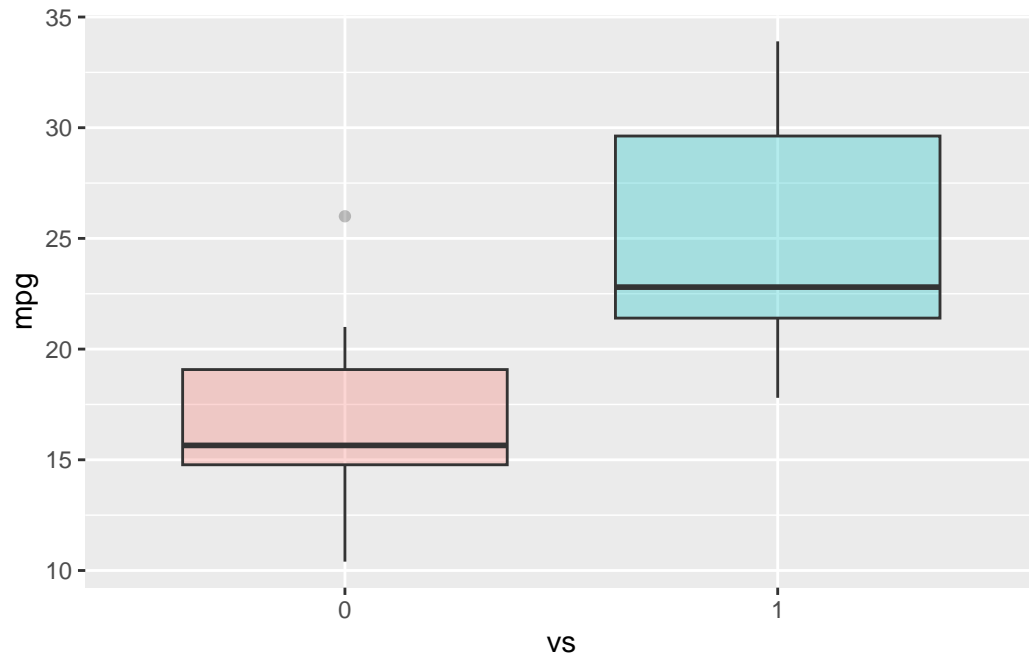
#create boxplot of the distribution for v-shaped and straight Engine

ggplot(aes(x=vs, y=mpg), data = mtcars) + geom_boxplot()
```



6. add color to boxplot

```
ggplot(aes(x=vs, y=mpg, fill=vs), data=mtcars) + geom_boxplot(alpha=0.3) + theme(legend.position="right")
```



7. histogram of single variable.

```
ggplot(aes(x=wt), data=mtcars) + geom_histogram(binwidth = 0.5)
```

