

Name: Sreejit Chatterjee

UID: 2019120010

EXTC

Data Analytics

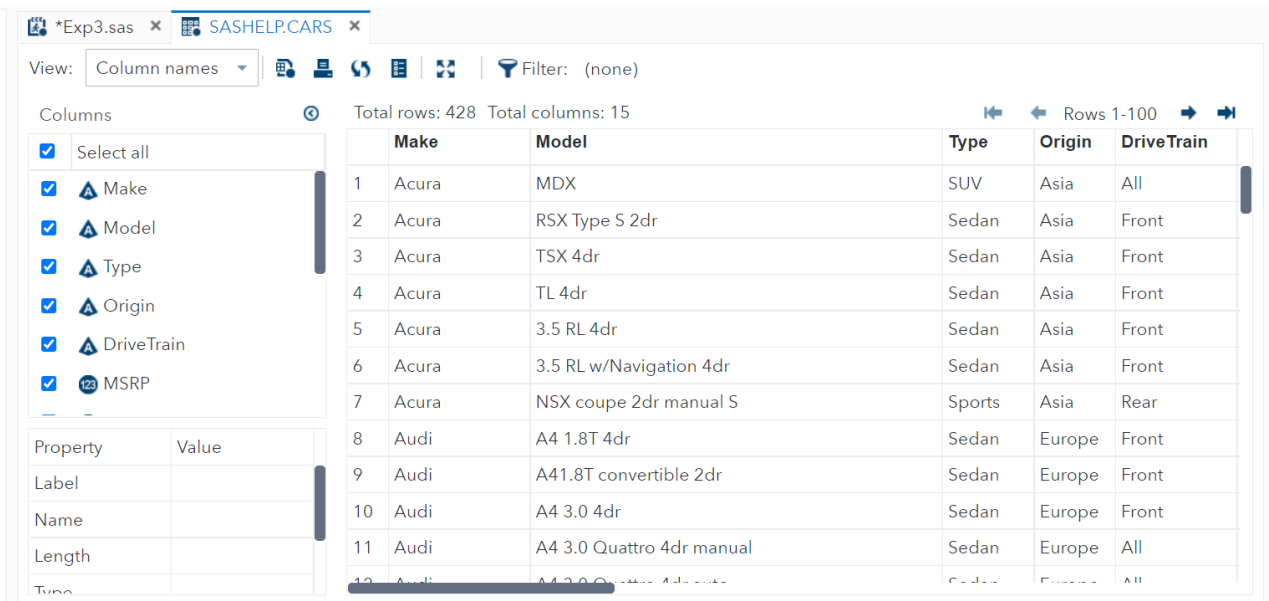
EXPERIMENT 3

AIM: To perform one way Anova with Sas platform

Procedure:

ANOVA stands for Analysis of Variance. In SAS it is done using PROC ANOVA. It performs analysis of data from a wide variety of experimental designs. In this process, a continuous response variable, known as a dependent variable, is measured under experimental conditions identified by classification variables, known as independent variables.

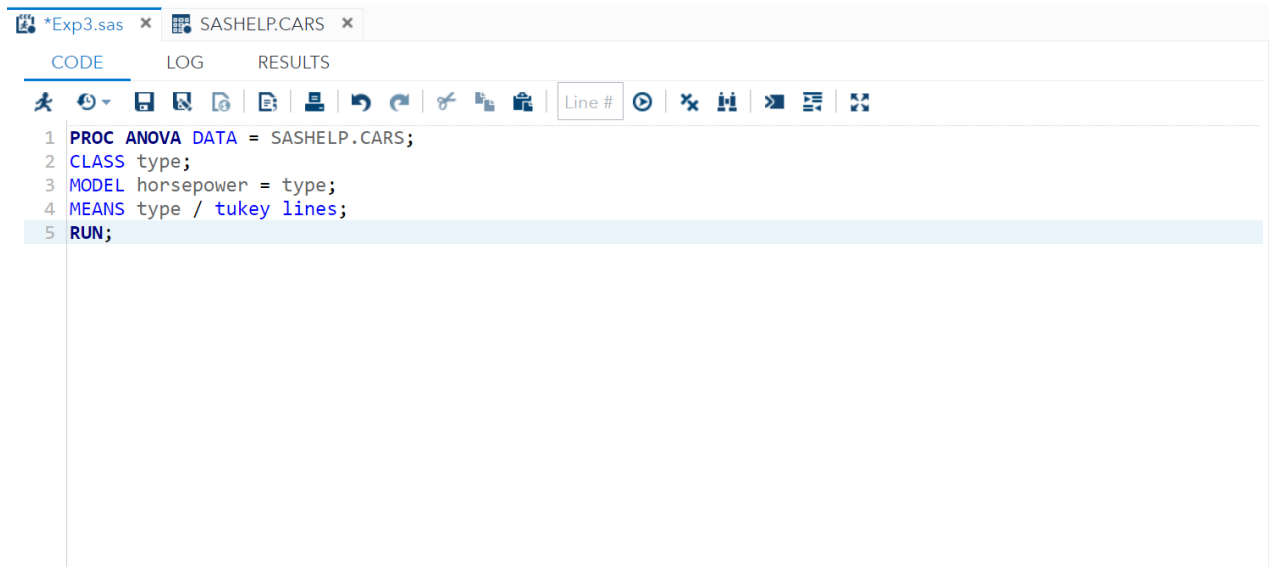
Consider the SASHELP.CARS dataset for one way anova test.



	Make	Model	Type	Origin	DriveTrain
1	Acura	MDX	SUV	Asia	All
2	Acura	RSX Type S 2dr	Sedan	Asia	Front
3	Acura	TSX 4dr	Sedan	Asia	Front
4	Acura	TL 4dr	Sedan	Asia	Front
5	Acura	3.5 RL 4dr	Sedan	Asia	Front
6	Acura	3.5 RL w/Navigation 4dr	Sedan	Asia	Front
7	Acura	NSX coupe 2dr manual S	Sports	Asia	Rear
8	Audi	A4 1.8T 4dr	Sedan	Europe	Front
9	Audi	A4 1.8T convertible 2dr	Sedan	Europe	Front
10	Audi	A4 3.0 4dr	Sedan	Europe	Front
11	Audi	A4 3.0 Quattro 4dr manual	Sedan	Europe	All
12	Audi	A4 3.0 Quattro 4dr manual	Sedan	Europe	All

In the SASHELP.CARS dataset there is dependence between the variables car type and their horsepower. As the car type is a variable with categorical values, we take it as class variable and use both these variables in the MODEL.

```
PROC ANOVA DATA = SASHELP.CARS;  
CLASS type;  
MODEL horsepower = type;  
MEANS type / tukey lines;  
RUN;
```

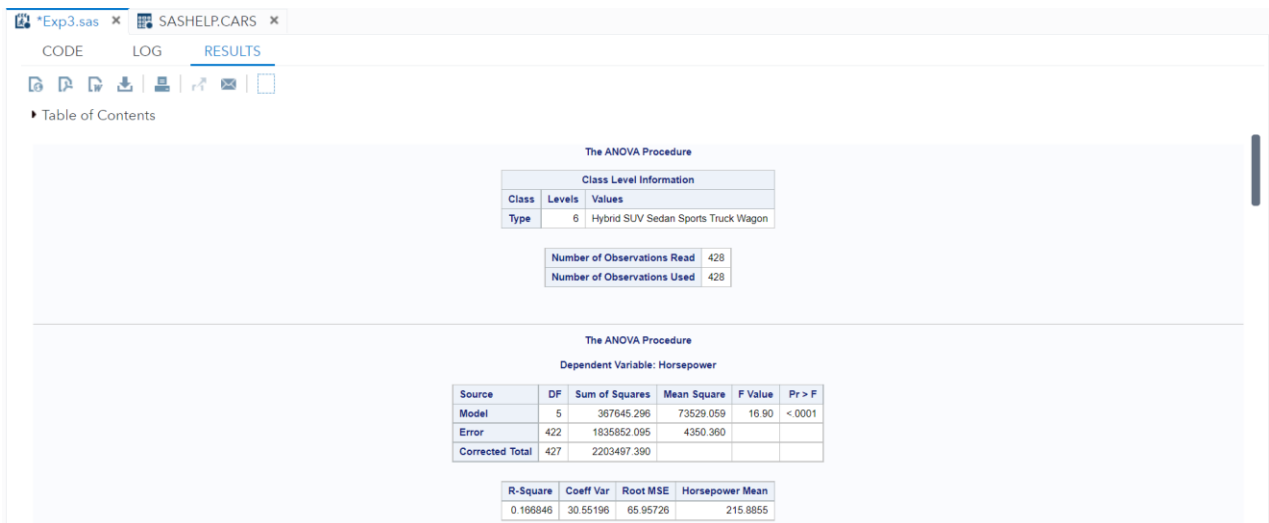


The screenshot displays the SAS Studio interface. At the top, there are two tabs: '*Exp3.sas' and 'SASHELP.CARS'. Below the tabs, there are three main sections: 'CODE', 'LOG', and 'RESULTS'. The 'CODE' section is active, showing a list of five lines of SAS code. The code is as follows:

```
1 PROC ANOVA DATA = SASHELP.CARS;  
2 CLASS type;  
3 MODEL horsepower = type;  
4 MEANS type / tukey lines;  
5 RUN;
```

The fifth line, 'RUN;', is highlighted in blue. To the right of the code editor, there is a toolbar with various icons for file operations, execution, and viewing. A 'Line #' input field is also present.

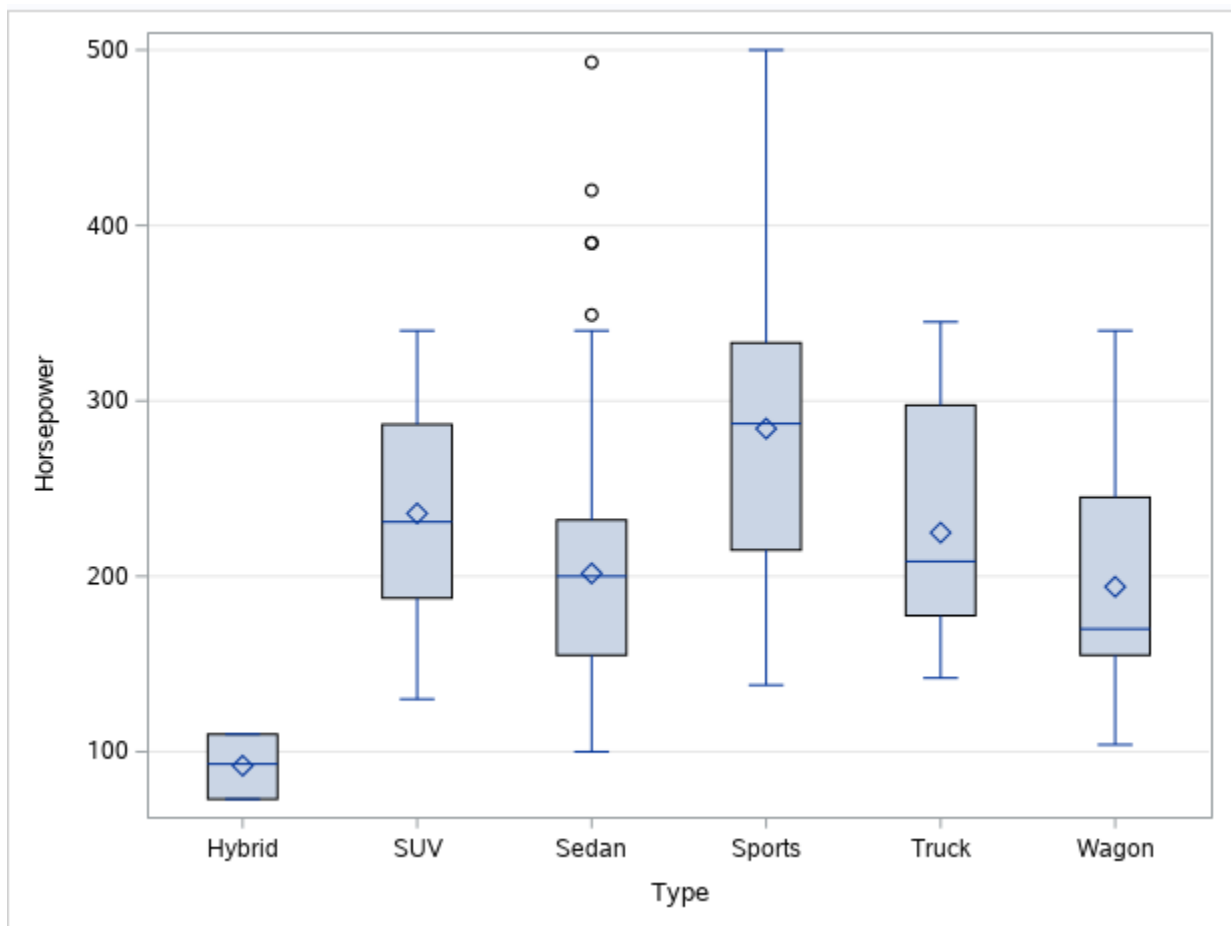
When the above code is executed, we get the following result –



In the above result

The overall F-value is 16.90 and P-value is <.0001 As the P-value is less than 0.5 so it rejects null hypothesis

Box plot

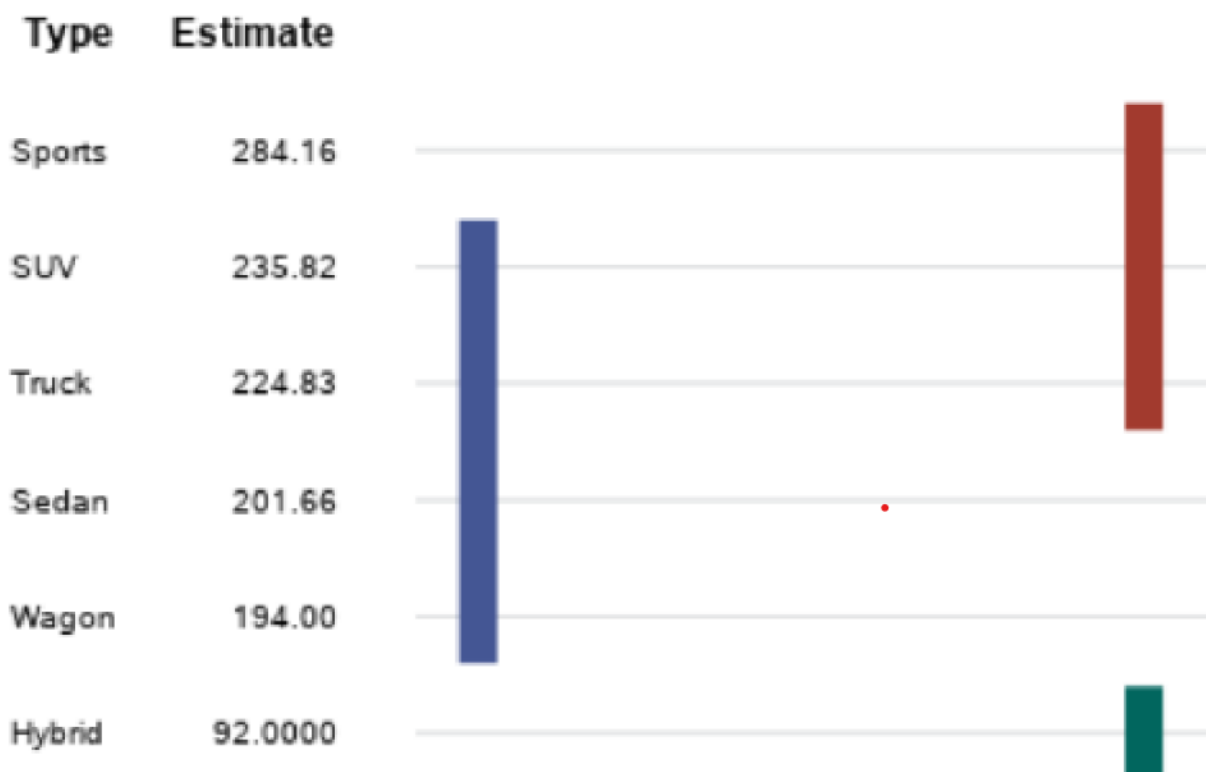


Alpha	0.05
Error Degrees of Freedom	422
Error Mean Square	4350.36
Critical Value of Studentized Range	4.04870
Minimum Significant Difference	73.069
Harmonic Mean of Cell Sizes	13.35634

Note: Cell sizes are not equal.

Horsepower Tukey Grouping for Means of Type (Alpha = 0.05)

Means covered by the same bar are not significantly different.



Conclusion:

As shown if the table given in fig above, we have got the p-value from the f-value as < 0.001 i.e. p-value is lesser than 0.01. Hence with this, we can conclude that our null hypothesis is rejected. That means, mean of all the types of cars are different.