Experiment 3

Name : Sneha Savarkar **UID: 2019120055** **Branch : EXTC**

Aim: To apply F-test for hypothesis testing using SAS software.

Dataset:

The dataset I use was the SAS application's built-in car dataset. There are 428 rows and 15 columns in the dataset.

HO: mean horsepower of all the types of cars is same

Ha: mean horsepower of all the types of cars is different

Code:

```
PROC SQL;

CREATE TABLE WORK.query AS

SELECT Make , Model , 'Type'n , Origin , DriveTrain , MSRP , Invoice , EngineSize , Cylinders , Horsepower , MPG_City , MPG_Highway , Weight , Wheelbase , 'Length'n
FROM SASHELP.CARS;

RUN;

QUIT;


PROC DATASETS NOLIST NODETAILS;

CONTENTS DATA=WORK.query OUT=WORK.details;

RUN;


PROC ANOVA DATA = WORK.query;

CLASS type;

MODEL Horsepower = type;

RUN;


PROC ANOVA DATA = WORK.query;

CLASS type;

MODEL horsepower = type;

MEANS type / tukey lines;
```

RUN;

PROC PRINT DATA=WORK.details;

RUN;

Output:

The ANOVA Procedure

Class Level Information		
Class	Levels	Values
Type	6	Hybrid SUV Sedan Sports Truck Wagon

Number of Observations Read	428
Number of Observations Used	428

Fig.1

The ANOVA Procedure

Dependent Variable: Horsepower

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	367645.296	73529.059	16.90	<.0001
Error	422	1835852.095	4350.360		
Corrected Total	427	2203497.390			

R-Square	Coeff Var	Root MSE	Horsepower Mean
0.166846	30.55196	65.95726	215.8855

Source	DF	Anova SS	Mean Square	F Value	Pr > F
Type	5	367645.2957	73529.0591	16.90	<.0001

Fig.2

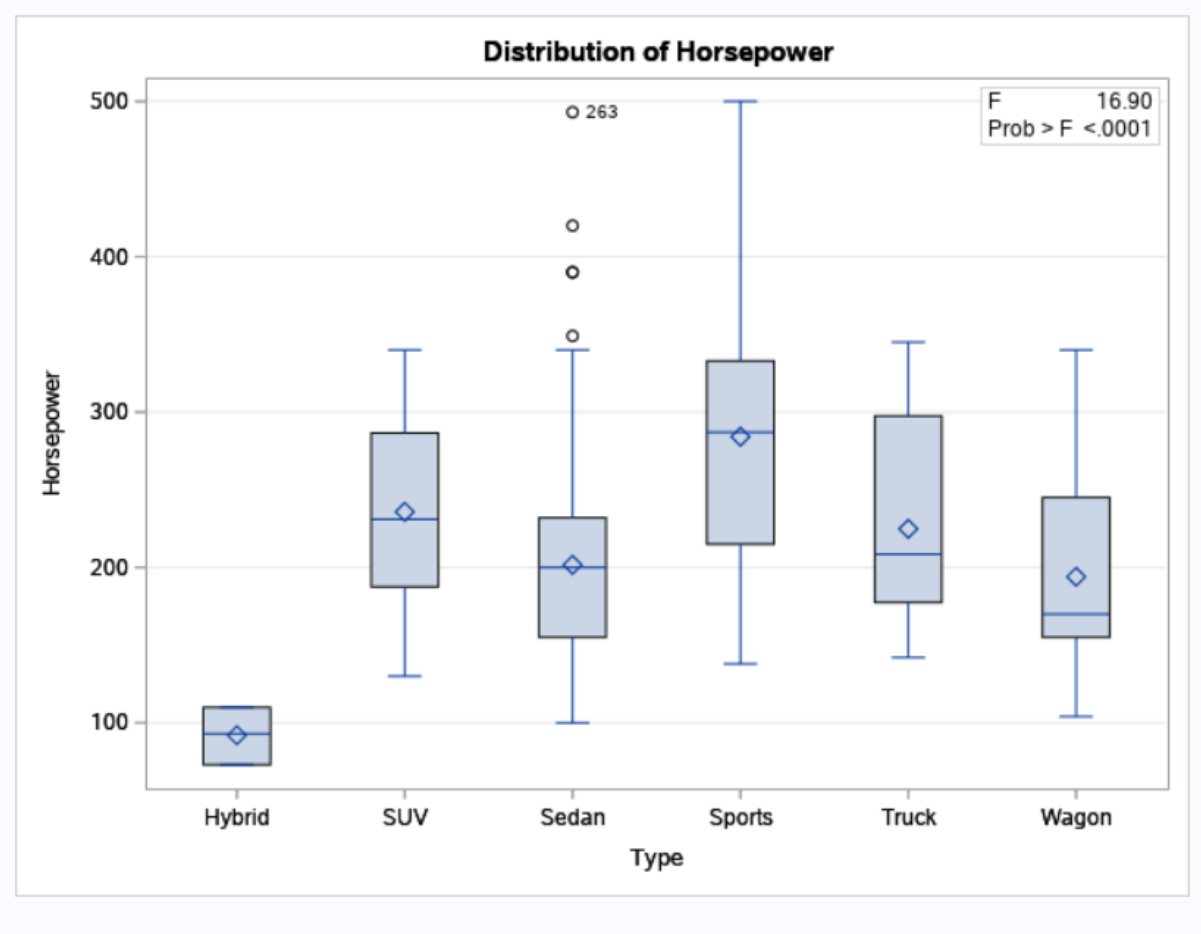
Distribution of Horsepower
F 16.90
Prob > F <.0001
263
500
400
300
200
100
Horsepower
Hybrid
SUV
Sedan
Sports
Truck
Wagon
Type

Fig.3

The ANOVA Procedure

Tukey's Studentized Range (HSD) Test for Horsepower

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than REGWQ.

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.

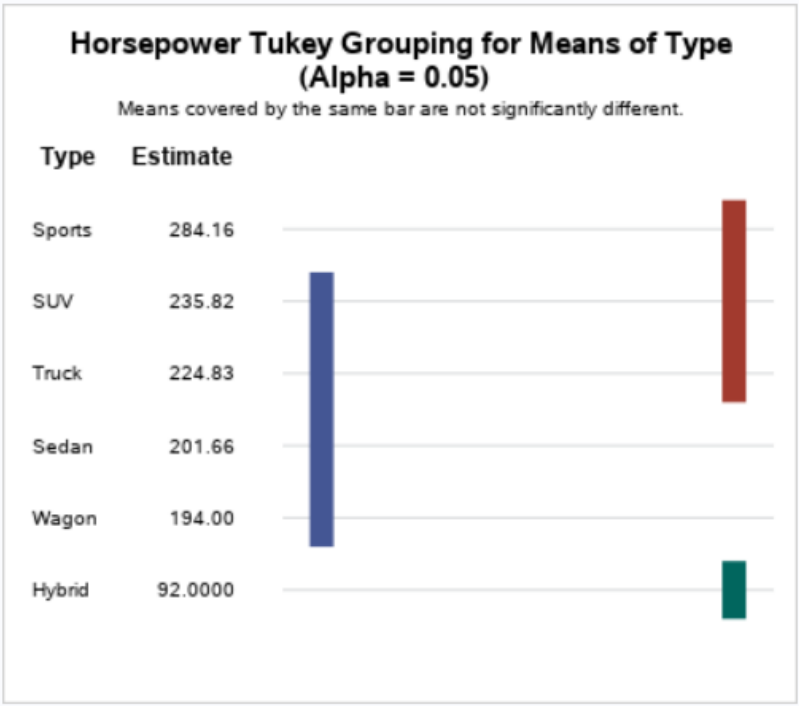

Fig.4

Conclusion:

From the above table the p-value from the f-value is <0.001 indicating that the p-value is less than 0.01. As a result, we can conclude that our null hypothesis is invalid. That is to say, the mean of all car types is different.

We can back up this claim by looking at Fig. 4, which contains the entry 'Minimum significant difference.' This entry with the number 73.069 validates our conclusion.