# Experiment 3

Sudhanshu Narsude
2019130044
TE COMPS
Data Analytics

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

**Dataset**:

The dataset that I used is inbuilt car dataset available in SAS application. The dataset contains 428 rows with 15 columns.

**H0**: 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 |

## 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 |

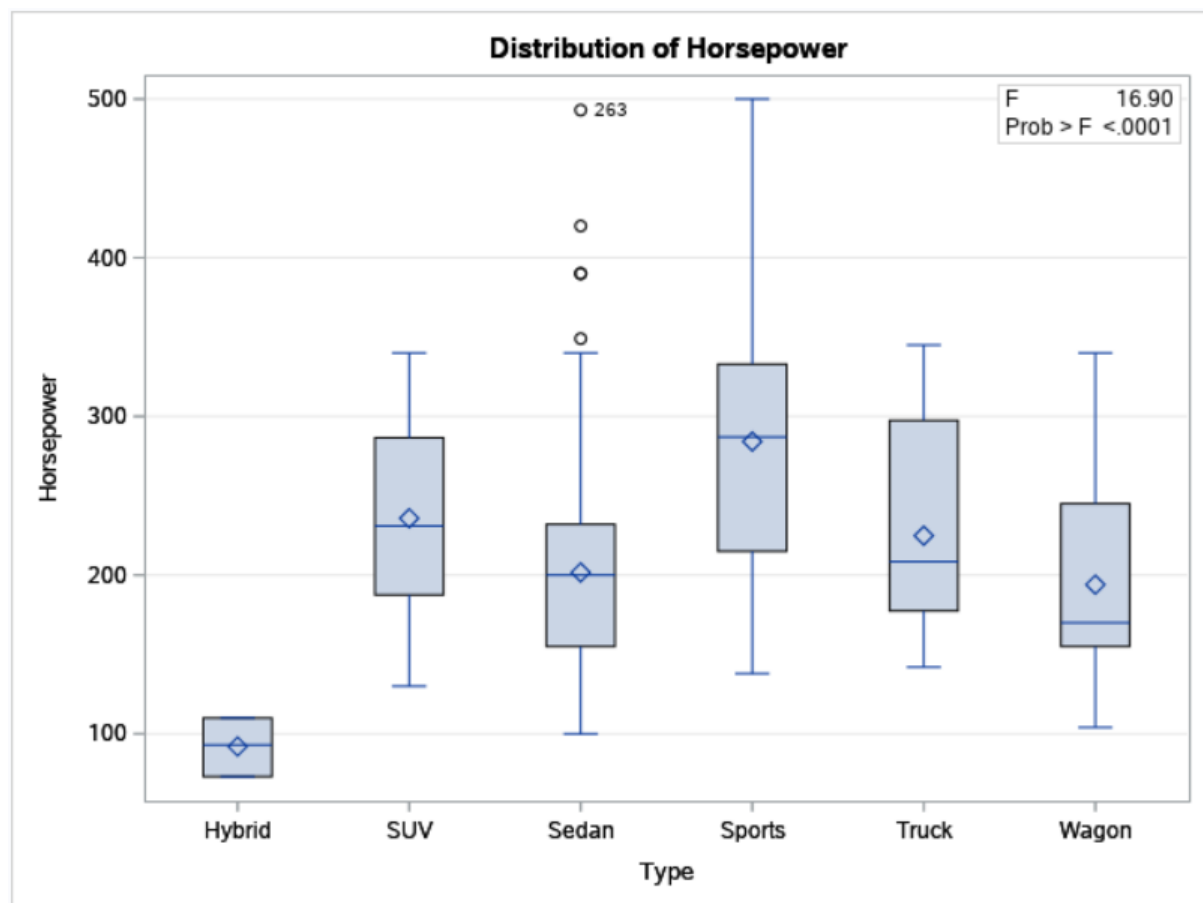| Source | DF | Anova SS | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Type | 5 | 367645.2957 | 73529.0591 | 16.90 | <.0001 |

*Fig 1*

*Fig 2*

| 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.

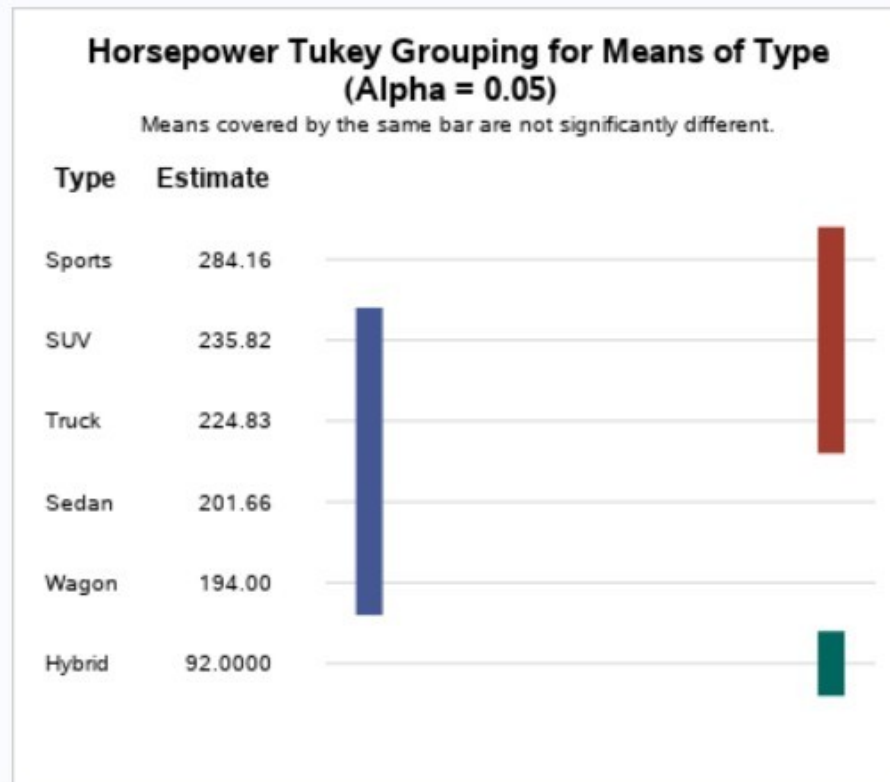| Type | Estimate |
|---|---|
| Sports | 284.16 |
| SUV | 235.82 |
| Truck | 224.83 |
| Sedan | 201.66 |
| Wagon | 194.00 |
| Hybrid | 92.0000 |

*Fig 3*

**Conclusion**:

As shown if the table given in fig 1 above, we have got the p-value from the f-value as $< 0.001$ i.e. p-value is lesser that 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.

To support this inference, we refer to fig. 3 which has the entry called 'Minimum significant difference'. This entry with value 73.069 proves our result to be correct.