Analysis Report:

By Shraddha Shaligram

**INTRODUCTION**:

The data set that we are working on is a classification data set which is used to differentiate the economy class air-ticket price from the premium economy class air-ticket price. Several other factors are provided in the data set which are closely related to the airline industry.

The aim of this analysis report is to find out which factors exactly contribute toward to difference in the prices of the premium economy class and the economy class air tickets.

I have listed down, in the form of pointers, different observations and conclusions that I have come to after the analysis of the airline industry data set.Following are the inferences that I have made:

* The British Airline has the maximum frequency of occurrence in the data set.It has the maximum number of aeroplanes with a different set of factor values for every aeroplane.
* The boxplot graphs of Airlines vs Premium air-ticket cost and Airlines vs Economy class air-ticket cost have a similar hyperbolic curve shape which suggests that the difference between the maximum and minimum ticket cost (for both classes) is dependent upon the airline type.
* The maximum difference between the minimum and maximum relative air-ticket cost is for British Airline.
* The Airline factor is statistically related to the economy class air ticket price, the premium economy class air ticket price and the relative price of both the classes from the correlation tests for the same.
* From the linear regression model between the flight duration and the relative price between the two classes, due to the p-value(<2.2e-16) being <0.05, we conclude that the fight duration factor is very much significantly important in determining the difference in the prices of the economy class and the premium economy class air tickets.
* The seats in the economy class are statistically related to the price of the economy class air ticket, as per the adjoining regression model and correlation table.
* The seats in the premium economy class are statistically related to the price of the premium economy class air ticket, as per the adjoining regression model and correlation table.
* The difference in the number of seats in the economy class and the premium economy class does contribute significantly to the difference in the air ticket cost of the economy class and the premium economy class, since the p-value =0.00173 as per the adjoining linear regression model.
* The difference in the pitch of the economy class seat and the premium economy class seat does contribute significantly to the difference in the prices of the air tickets of the economy class and the premium economy class, as the p-value = 0.0117 from the adjoining linear regression model.
* The difference in the width of the economy class seat and the premium economy class seat does contribute significantly to the difference in the prices of the air tickets of the economy class and the premium economy class, as the p-value = 0.023 from the adjoining linear regression model.
* Surprisingly, the month is positively correlated to the difference in the prices of the economy class and premium economy class air tickets, from the adjoining correlation table and test. As the months pass by from July to October the difference in the air tickets of the two economy classes goes on increasing.
* From the correlation test between the quality factor and the difference in the prices of the premium economy and economy class air tickets, we conclude that the quality is positively correlated with the later with a p-value = 0.01165 and a correlation coefficient =0.1172
* The Aircraft factor is negatively correlated to the difference in the prices of the economy and premium economy class air tickets due to negative correlation coefficient. With a p-value =0.6862, the Aircraft factor is not a significant contributor to the latter.