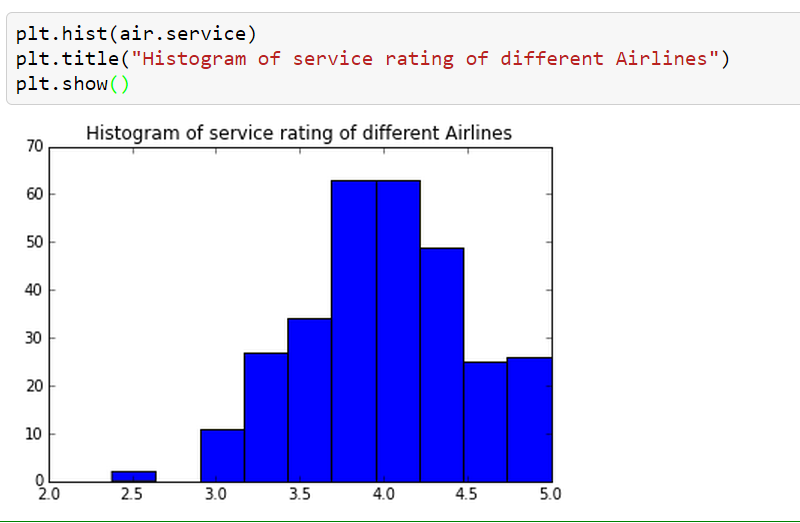
Best Airlines

Introduction about the Project:

The project I have chosen is predicting the best airlines. I have used regression methods to build a model. In this model my main aim was to show if my dataset has any null or missing values and plot the bar charts and histograms for each and every attribute in my dataset.

EDA of my project

I have checked for the null values and plotted the histograms and bar charts for all the attributes in the dataset. The airline ID, price, speed and service are my main categorical variables. I have chosen these variables to determine which airline is the best suitable one for the passengers. Seat, delay time and counter number are assumed to be of least importance because price describes the seat and delay time is based on the climatic conditions and other factors which has nothing to do with the airline management. Counter number is also of least importance. I did not find any null values for the categorical variables I have selected and hence I need not dummy out any of the column in the dataset. Bar charts and histograms are shown for all the variables. A sample bar chart is shown below:



I have shown the minimum, maximum and standard deviation values for few categorical variables.