

# Capstone Project 1 : Project Proposal

Project Title :

“Fly High and Pay for Correct Buy!”

Project Mentor : Debarya Dutta

*What is the Problem to be solved?*

There are many instances where lots of people say that the Flight Prices are unpredictable. It is also observed that Flight prices may vary based on the distance, time, stops, etc.. So the problem that I would be solving is the Prediction of Flight Prices.

*Who is this project beneficial for?*

This project is user beneficial as in for the public. People who fly frequently, people going on a holiday can also be benefited from it price prediction. It is also beneficial for different companies who have to fly their employees or clients.

*What is the data used for the project? And where I received the data from?*

So the data consists of all flight records in India for a span of 6 months and I have extracted the data from Machine Hack website. This is a multivariate data consisting of origin, destination, departure time, arrival time, price, etc.

*Techniques that I would use to solve the problem:*

- 1) Data Wrangling - I will observe the data for all the missing values, incorrect entries and based on the observation I will clean the data for further use.
- 2) Exploratory Data Analysis - I would plot the relation between different parameters in the dataset and based on that I can find the trends for storytelling.

- 3) Inferential Statistics - This is another important aspect based on which we can find out some statistical relations in the dataset. I will applying different testing methods in Inferential Statistics.
- 4) Machine Learning Model - After all the initial work on the dataset it would be time to build a Machine Learning Model to find the Prediction of the Flight ticket Prices. I may use Linear Regression, Random Forest regression.

### *What is the methods I will use to deliver the project?*

I will be using the following tools to work on the project :-

- Jupyter Notebook for the major part of the project.
- I will be using Python version 3.x for writing the code.
- I will use Google doc for submission on the in process project reports.
- Powerpoint for Data Storytelling presentation.

### *References:*

- Springboard Curriculum material
- DataCamp courses
- Google, Stackoverflow, Github, etc.
- And any other Open Resources.