

Data Analysis on Airlines Delays

Goal :

To analyse airlines delay data in USA and curve delays caused in major airlines between busiest routes which will give a fair idea to passengers about frequency of delays in general for a period of one year along with a particular airline and between a particular route. This will help customers for choosing a better airlines and taking a right call based on the weather for a particular region. Also, the same will also help airlines for better management and work on root cause of the delays.

Objective:

To provide general descriptive analysis on delays and to analyse the root cause of delays. To generate Predictive analysis in determining probable delays in defined airlines and airports. To compare between different airlines and major airports with respect to delays and cancellation.

```
In [65]: import unicodcsv #importing csv
import pandas as pd #importing pandas
import numpy as np #importing numpy
import time #importing time
import matplotlib.pyplot as plt #importing matplotlib
import seaborn as sns #importing seaborn
import datetime as dt #importing datetime
from scipy import stats #importing scipy
from scipy.stats import norm
from os import walk
import time
%matplotlib inline
sns.set_style('whitegrid')
```

Source: Airlines Data - Bureau of Transportation Statistics

Data Cleaning of the dataset: