import library

```
In [14]: import pandas as pd
         import numpy as np
         import joblib
         from sklearn.model_selection import train_test_split
         from sklearn.tree import DecisionTreeRegressor
         from sklearn.metrics import mean_squared_error
```

import dataset

```
In [15]: dataset = pd.read_csv('C:/Users/Abirami/Desktop/flight delay.csv')
         dataset.head()
```

Out[15]:

	YEAR	QUARTER	MONTH	DAY_OF_MONTH	DAY_OF_WEEK	UNIQUE_CARRIER	TAIL_NUM	FL_NUM	ORIGIN_AIRPORT_ID	ORIGIN	...	DEP_DEL15	CRS_A
0	2016	1	1	1	5	DL	N836DN	1399	10397	ATL	...	0.0	
1	2016	1	1	1	5	DL	N964DN	1476	11433	DTW	...	0.0	
2	2016	1	1	1	5	DL	N813DN	1597	10397	ATL	...	0.0	
3	2016	1	1	1	5	DL	N587NW	1768	14747	SEA	...	0.0	
4	2016	1	1	1	5	DL	N836DN	1823	14747	SEA	...	0.0	

5 rows × 25 columns