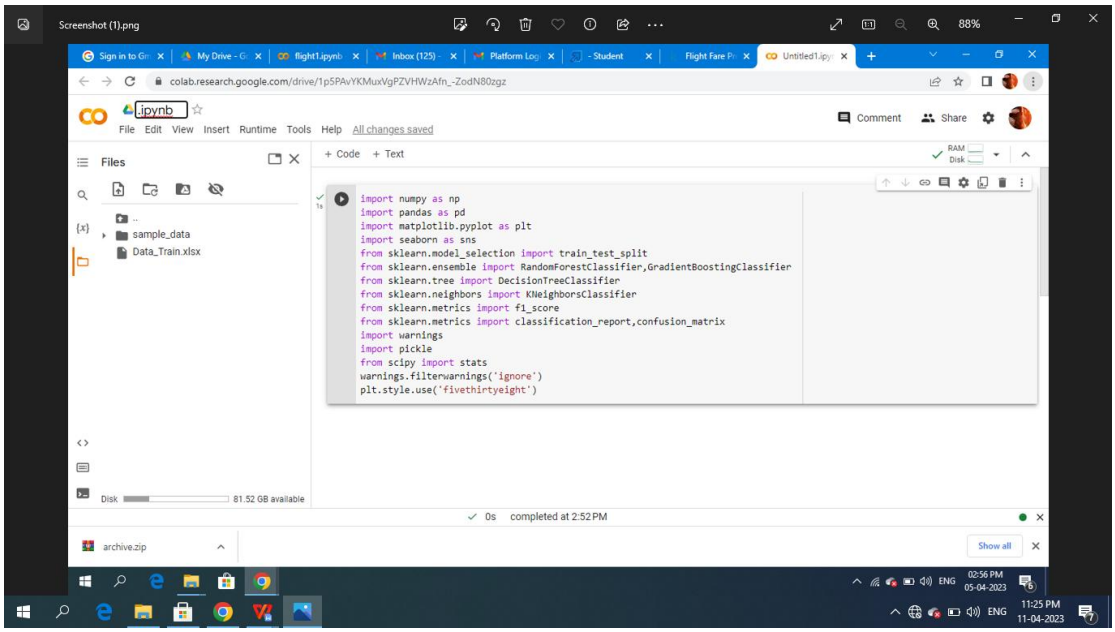
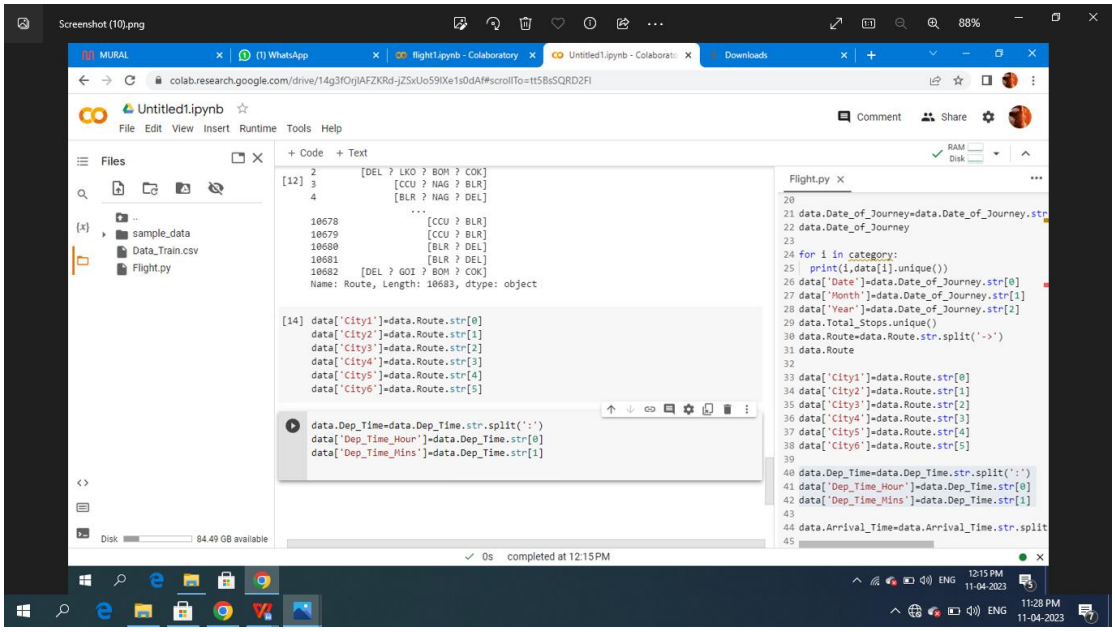
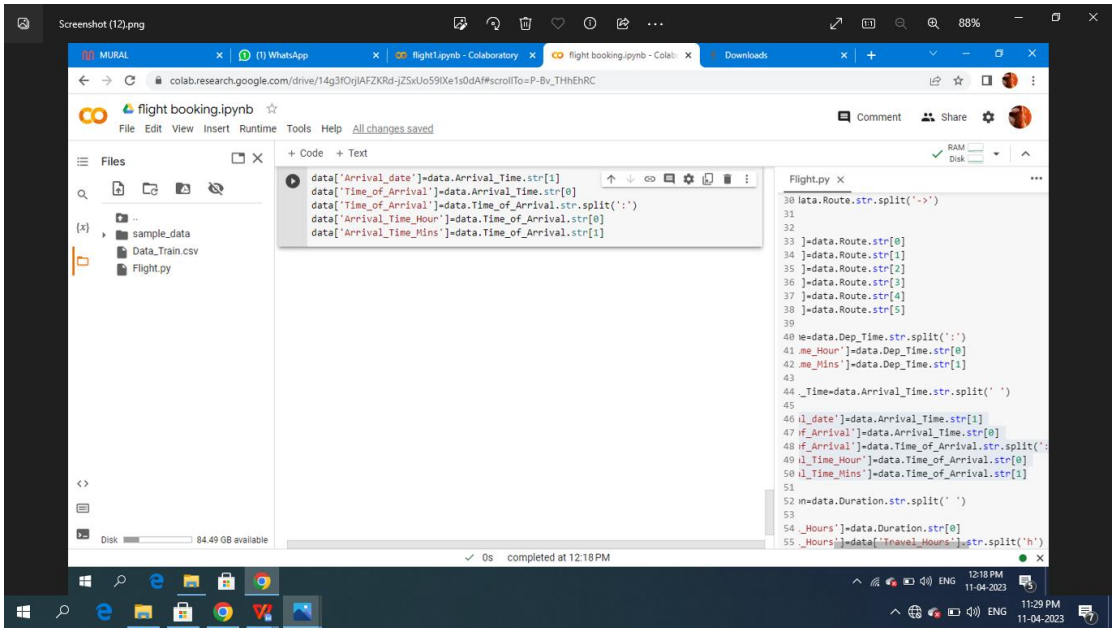
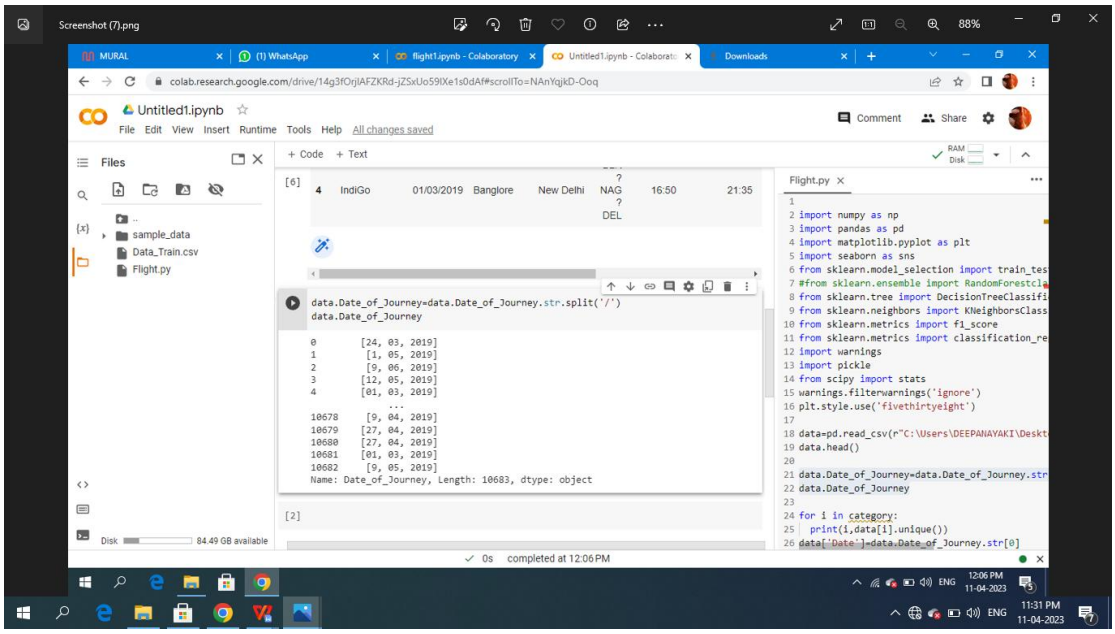
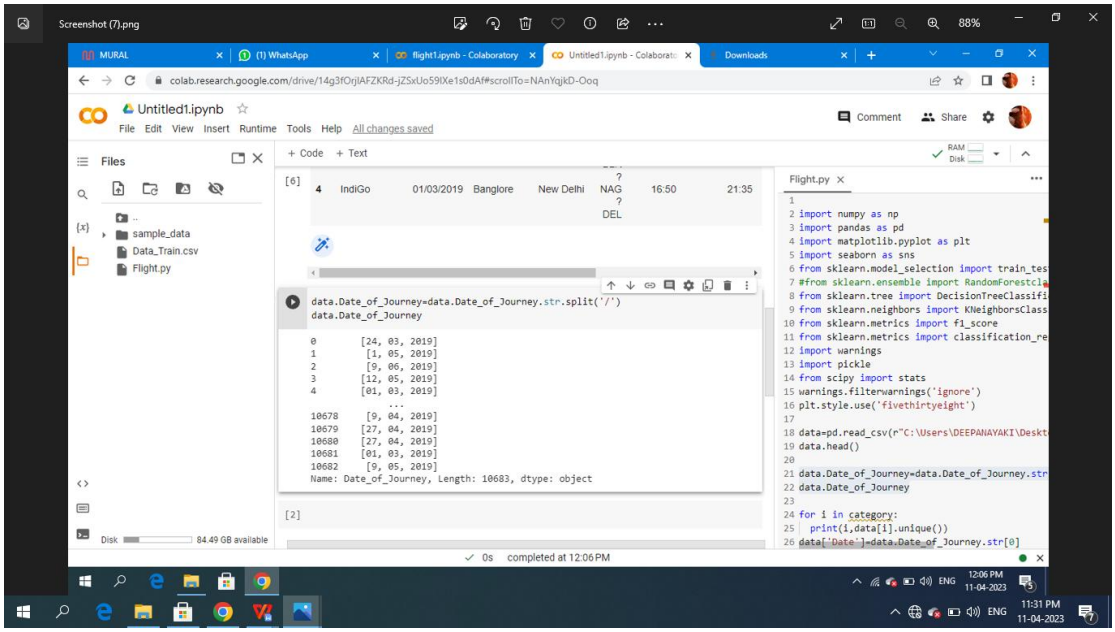
OUTPUTS

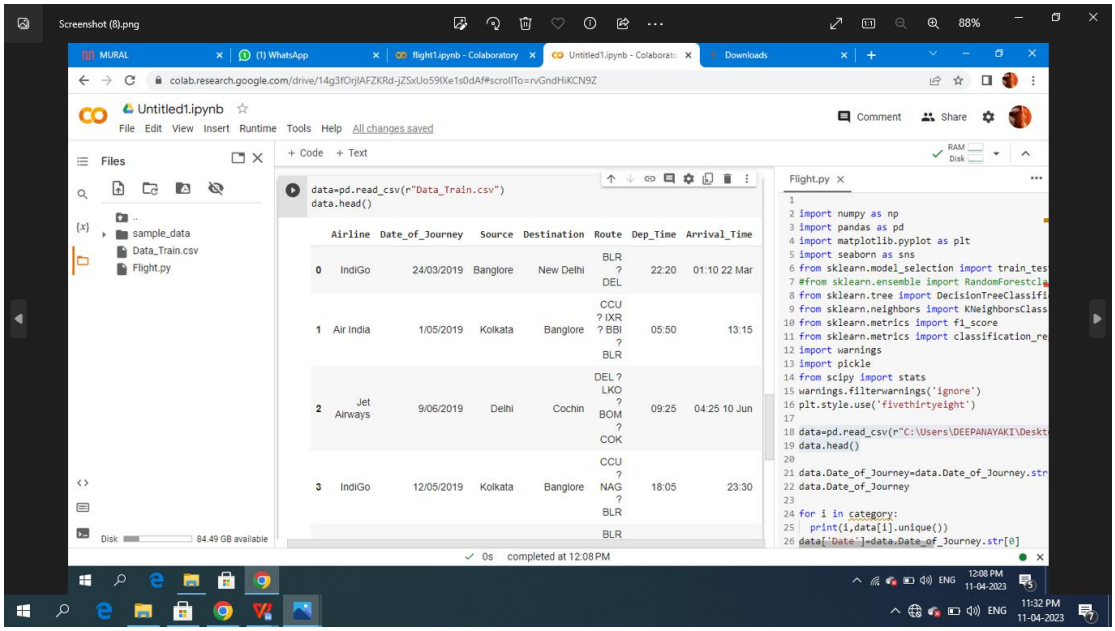

Screenshot (1).png
colab.research.google.com/drive/1p5PAvYKMuxVgPZVHWzAfn_-ZodN80zgz
ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
Data_Train.xlsx
+ Code + Text
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.ensemble import RandomForestClassifier,GradientBoostingClassifier
from sklearn.tree import DecisionTreeClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.metrics import f1_score
from sklearn.metrics import classification_report,confusion_matrix
import warnings
import pickle
from scipy import stats
warnings.filterwarnings('ignore')
plt.style.use('fivethirtyeight')
Disk 81.52 GB available
0s completed at 2:52 PM
archive.zip
Show all

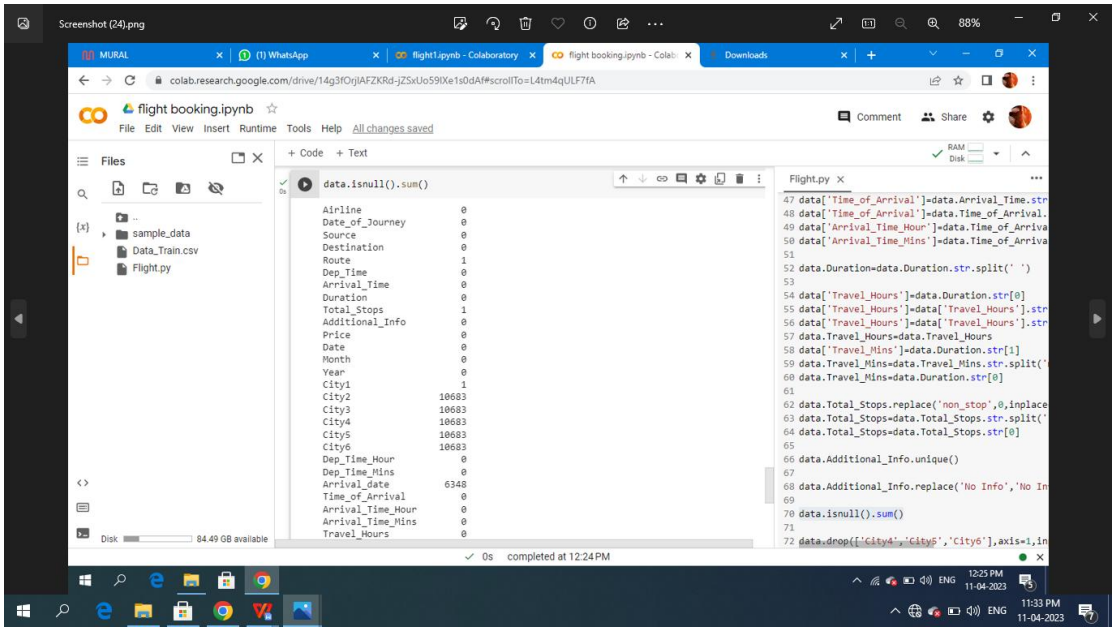

Screenshot (10).png
colab.research.google.com/drive/14g3fOrjlAFZKRd-jZSxUo59IXe1s0dAf#scrollTo=tt5BsSQRD2FI
Untitled1.ipynb
File Edit View Insert Runtime Tools Help
Comment
Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
[12] 2 [DEL ? LKO ? BOM ? COK]
3 [CCU ? NAG ? BLR]
4 [BLR ? NAG ? DEL]
...
10678 [CCU ? BLR]
10679 [CCU ? BLR]
10680 [BLR ? DEL]
10681 [BLR ? DEL]
10682 [DEL ? GOI ? BOM ? COK]
Name: Route, Length: 10683, dtype: object
[14] data['City1']=data.Route.str[0]
data['City2']=data.Route.str[1]
data['City3']=data.Route.str[2]
data['City4']=data.Route.str[3]
data['City5']=data.Route.str[4]
data['City6']=data.Route.str[5]
data.Dep_Time=data.Dep_Time.str.split(':')
data['Dep_Time_Hour']=data.Dep_Time.str[0]
data['Dep_Time_Mins']=data.Dep_Time.str[1]
Flight.py
data.Date_of_Journey=data.Date_of_Journey.str
data.Date_of_Journey
for i in category:
print(i,data[i].unique())
data['Date']=data.Date_of_Journey.str[0]
data['Month']=data.Date_of_Journey.str[1]
data['Year']=data.Date_of_Journey.str[2]
data.Total_Stops.unique()
data.Route=data.Route.str.split('->')
data.Route
data['City1']=data.Route.str[0]
data['City2']=data.Route.str[1]
data['City3']=data.Route.str[2]
data['City4']=data.Route.str[3]
data['City5']=data.Route.str[4]
data['City6']=data.Route.str[5]
data.Dep_Time=data.Dep_Time.str.split(':')
data['Dep_Time_Hour']=data.Dep_Time.str[0]
data['Dep_Time_Mins']=data.Dep_Time.str[1]
data.Arrival_Time=data.Arrival_Time.str.split
Disk 84.49 GB available
0s completed at 12:15 PM

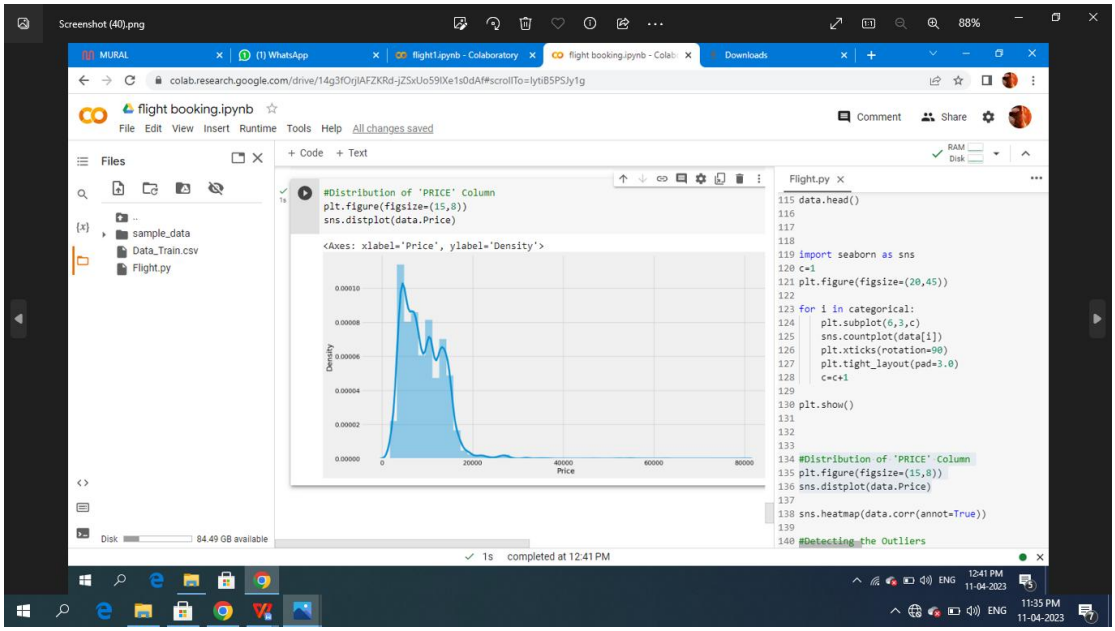

Screenshot (12).png
flight booking.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
data['Arrival_date']=data.Arrival_Time.str[1]
data['Time_of_Arrival']=data.Arrival_Time.str[0]
data['Time_of_Arrival']=data.Time_of_Arrival.str.split(':')
data['Arrival_Time_Hour']=data.Time_of_Arrival.str[0]
data['Arrival_Time_Mins']=data.Time_of_Arrival.str[1]
Disk 84.49 GB available
0s completed at 12:18 PM

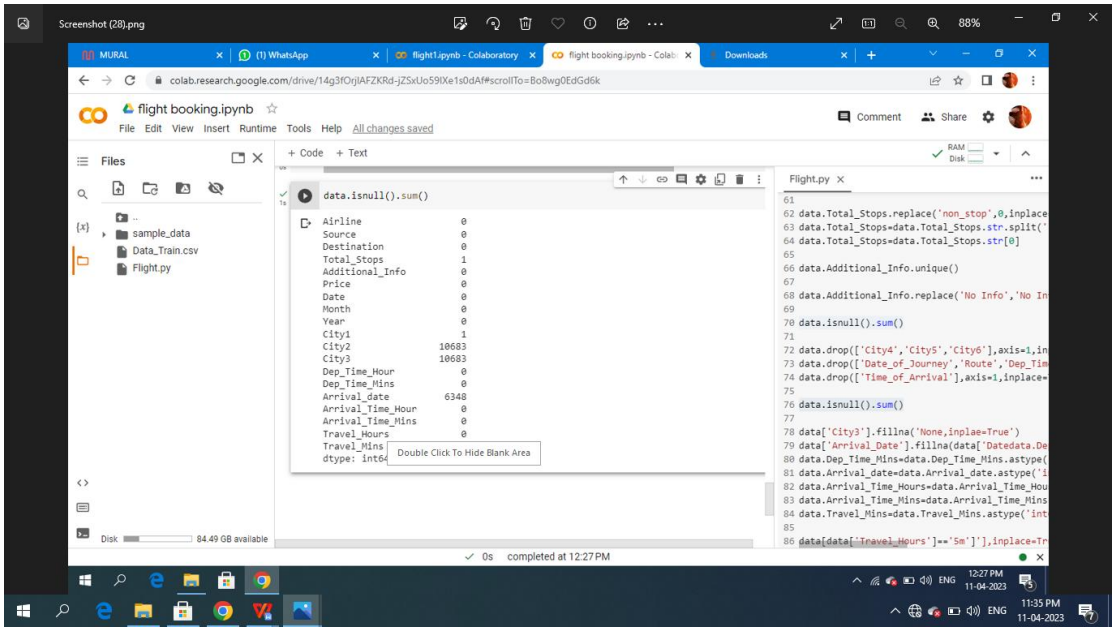

Screenshot (7).png
Untitled1.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
[6]
4 IndiGo 01/03/2019 Banglore New Delhi NAG ? DEL 16:50 21:35
data.Date_of_Journey=data.Date_of_Journey.str.split('/')
data.Date_of_Journey
0 [24, 03, 2019]
1 [1, 05, 2019]
2 [9, 06, 2019]
3 [12, 05, 2019]
4 [01, 03, 2019]
10678 [9, 04, 2019]
10679 [27, 04, 2019]
10680 [27, 04, 2019]
10681 [01, 03, 2019]
10682 [9, 05, 2019]
Name: Date_of_Journey, Length: 10683, dtype: object
[2]
Disk 84.49 GB available
0s completed at 12:06 PM

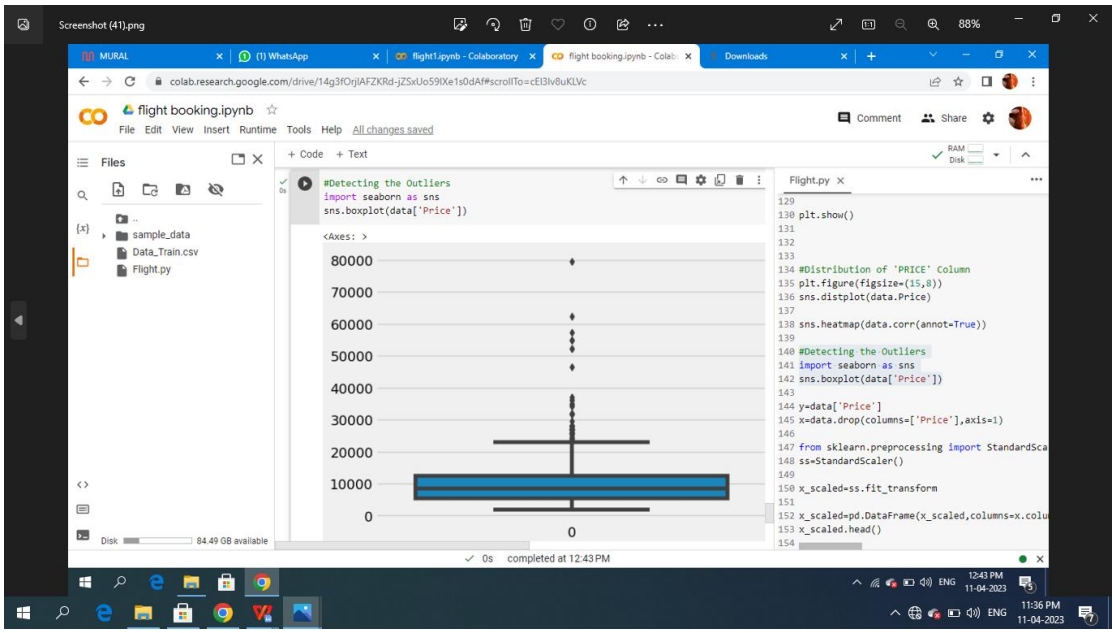

Screenshot (7).png
Untitled1.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
[6]
4 IndiGo 01/03/2019 Banglore New Delhi ? NAG ? DEL 16:50 21:35
data.Date_of_Journey=data.Date_of_Journey.str.split('/')
data.Date_of_Journey
0 [24, 03, 2019]
1 [1, 05, 2019]
2 [9, 06, 2019]
3 [12, 05, 2019]
4 [01, 03, 2019]
10678 [9, 04, 2019]
10679 [27, 04, 2019]
10680 [27, 04, 2019]
10681 [01, 03, 2019]
10682 [9, 05, 2019]
Name: Date_of_Journey, Length: 10683, dtype: object
[2]
Disk 84.49 GB available
0s completed at 12:06 PM
Flight.py
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_tes
#from sklearn.ensemble import RandomForestCl
from sklearn.tree import DecisionTreeClassifi
from sklearn.neighbors import KNeighborsClass
from sklearn.metrics import f1_score
from sklearn.metrics import classification_re
import warnings
import pickle
from scipy import stats
warnings.filterwarnings('ignore')
plt.style.use('fivethirtyeight')
data=pd.read_csv(r"C:\Users\DEEPANAYAKI\Deskt
data.head()
data.Date_of_Journey=data.Date_of_Journey.str
data.Date_of_Journey
for i in category:
print(i,data[i].unique())
data['Date']=data.Date_of_Journey.str[0]
12:06 PM 11-04-2023
11:31 PM 11-04-2023

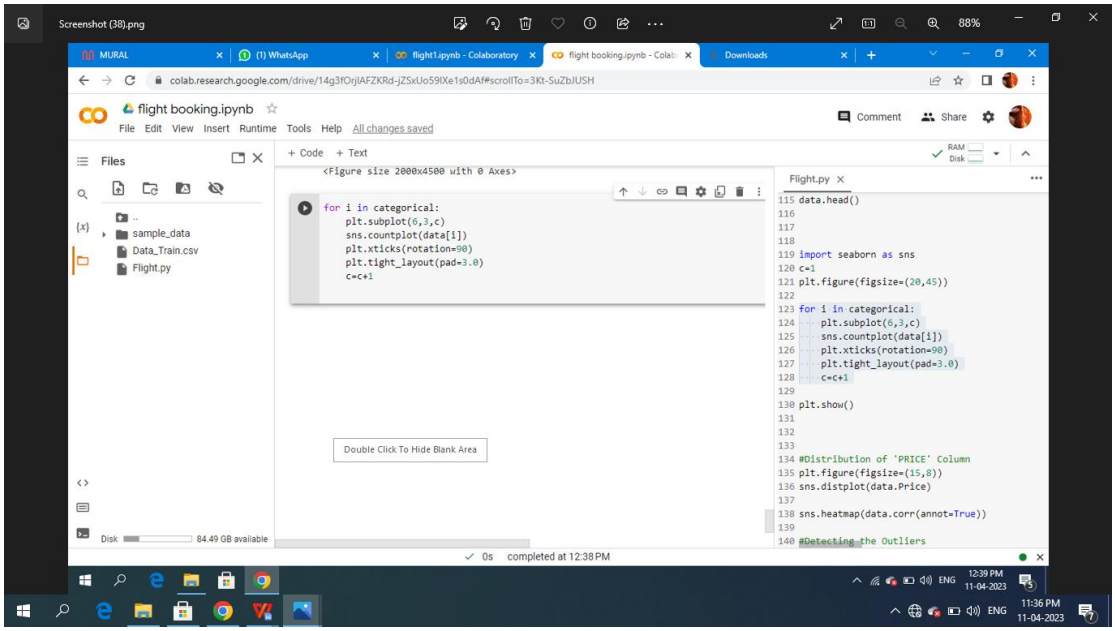

Screenshot (8).png
Untitled1.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
data=pd.read_csv(r"Data_Train.csv")
data.head()
Airline Date_of_Journey Source Destination Route Dep_Time Arrival_Time
0 IndiGo 24/03/2019 Banglore New Delhi BLR ? DEL 22:20 01:10 22 Mar
1 Air India 1/05/2019 Kolkata Banglore CCU ? IXR ? BBI ? BLR 05:50 13:15
2 Jet Airways 9/06/2019 Delhi Cochin DEL ? LKO ? BOM ? COK 09:25 04:25 10 Jun
3 IndiGo 12/05/2019 Kolkata Banglore CCU ? NAG ? BLR 18:05 23:30
Disk 84.49 GB available
0s completed at 12:08 PM
Flight.py
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_tes
#from sklearn.ensemble import RandomForestCl
from sklearn.tree import DecisionTreeClassifi
from sklearn.neighbors import KNeighborsClass
from sklearn.metrics import f1_score
from sklearn.metrics import classification_re
import warnings
import pickle
from scipy import stats
warnings.filterwarnings('ignore')
plt.style.use('fivethirtyeight')
data=pd.read_csv(r"C:\Users\DEEPANAYAKI\Deskt
data.head()
data.Date_of_Journey=data.Date_of_Journey.str
data.Date_of_Journey
for i in category:
print(i,data[i].unique())
data['Date']=data.Date_of_Journey.str[0]
12:08 PM 11-04-2023
11:32 PM 11-04-2023

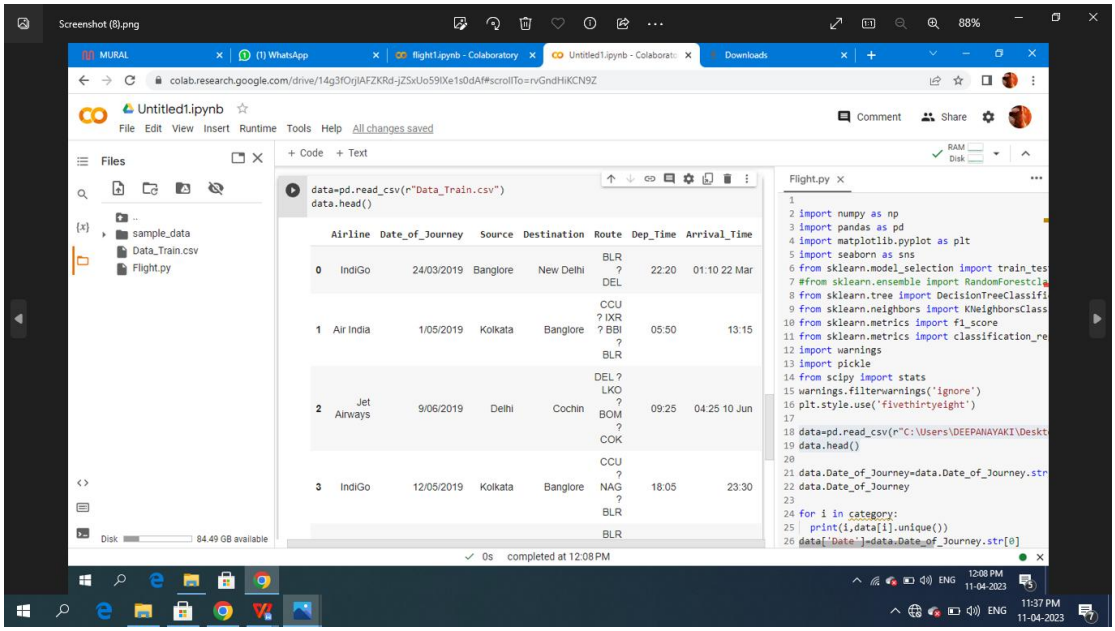

Screenshot (24).png
flight booking.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
data.isnull().sum()
Airline 0
Date_of_Journey 0
Source 0
Destination 0
Route 1
Dep_Time 0
Arrival_Time 0
Duration 0
Total_Stops 1
Additional_Info 0
Price 0
Date 0
Month 0
Year 0
City1 1
City2 10683
City3 10683
City4 10683
City5 10683
City6 10683
Dep_Time_Hour 0
Dep_Time_Mins 0
Arrival_date 6348
Time_of_Arrival 0
Arrival_Time_Hour 0
Arrival_Time_Mins 0
Travel_Hours 0
Disk 84.49 GB available
0s completed at 12:24 PM
Flight.py
data['Time_of_Arrival']=data.Arrival_Time.str
data['Time_of_Arrival']=data.Time_of_Arrival.
data['Arrival_Time_Hour']=data.Time_of_Arriva
data['Arrival_Time_Mins']=data.Time_of_Arriva
data.Duration=data.Duration.str.split(' ')
data['Travel_Hours']=data.Duration.str[0]
data['Travel_Hours']=data['Travel_Hours'].str
data['Travel_Hours']=data['Travel_Hours'].str
data.Travel_Hours=data.Travel_Hours
data['Travel_Mins']=data.Duration.str[1]
data.Travel_Mins=data.Travel_Mins.str.split('
data.Travel_Mins=data.Duration.str[0]
data.Total_Stops.replace('non_stop',0,inplace
data.Total_Stops=data.Total_Stops.str.split('
data.Total_Stops=data.Total_Stops.str[0]
data.Additional_Info.unique()
data.Additional_Info.replace('No Info','No In
data.isnull().sum()
data.drop(['City4','City5','City6'],axis=1,in
12:25 PM 11-04-2023
11:33 PM 11-04-2023


Screenshot (40).png
flight booking.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
#Distribution of 'PRICE' Column
plt.figure(figsize=(15,8))
sns.distplot(data.Price)
<Axes: xlabel='Price', ylabel='Density'>
Density
Price
Disk 84.49 GB available
1s completed at 12:41 PM
Flight.py
data.head()
import seaborn as sns
c=1
plt.figure(figsize=(20,45))
for i in categorical:
plt.subplot(6,3,c)
sns.countplot(data[i])
plt.xticks(rotation=90)
plt.tight_layout(pad=3.0)
c=c+1
plt.show()
#Distribution of 'PRICE' Column
plt.figure(figsize=(15,8))
sns.distplot(data.Price)
sns.heatmap(data.corr(annot=True))
#Detecting the Outliers
12:41 PM 11-04-2023
11:35 PM 11-04-2023


Screenshot (28).png
flight booking.ipynb
data.isnull().sum()
Airline 0
Source 0
Destination 0
Total_Stops 1
Additional_Info 0
Price 0
Date 0
Month 0
Year 0
City1 1
City2 10683
City3 10683
Dep_Time_Hour 0
Dep_Time_Mins 0
Arrival_date 6348
Arrival_Time_Hour 0
Arrival_Time_Mins 0
Travel_Hours 0
Travel_Mins 0
dtype: int64
Double Click To Hide Blank Area
completed at 12:27 PM


Screenshot (41).png
flight booking.ipynb
#Detecting the Outliers
import seaborn as sns
sns.boxplot(data['Price'])
<Axes: >
80000
70000
60000
50000
40000
30000
20000
10000
0
completed at 12:43 PM


Screenshot (38).png
flight booking.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
<Figure size 2000x4500 with 0 Axes>
for i in categorical:
plt.subplot(6,3,c)
sns.countplot(data[i])
plt.xticks(rotation=90)
plt.tight_layout(pad=3.0)
c=c+1
Double Click To Hide Blank Area
Disk 84.49 GB available
0s completed at 12:38 PM
Flight.py
115 data.head()
119 import seaborn as sns
120 c=1
121 plt.figure(figsize=(20,45))
123 for i in categorical:
124 plt.subplot(6,3,c)
125 sns.countplot(data[i])
126 plt.xticks(rotation=90)
127 plt.tight_layout(pad=3.0)
128 c=c+1
130 plt.show()
134 #Distribution of 'PRICE' Column
135 plt.figure(figsize=(15,8))
136 sns.distplot(data.Price)
138 sns.heatmap(data.corr(annot=True))
140 #Detecting the Outliers
12:39 PM 11-04-2023
11:36 PM 11-04-2023


Screenshot (8).png
Untitled1.ipynb
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
sample_data
Data_Train.csv
Flight.py
+ Code + Text
data=pd.read_csv(r"Data_Train.csv")
data.head()
Airline Date_of_Journey Source Destination Route Dep_Time Arrival_Time
0 IndiGo 24/03/2019 Banglore New Delhi BLR ? DEL 22:20 01:10 22 Mar
1 Air India 1/05/2019 Kolkata Banglore CCU ? IXR ? BBI ? BLR 05:50 13:15
2 Jet Airways 9/06/2019 Delhi Cochin DEL ? LKO ? BOM ? COK 09:25 04:25 10 Jun
3 IndiGo 12/05/2019 Kolkata Banglore CCU ? NAG ? BLR
BLR
Disk 84.49 GB available
0s completed at 12:08 PM
Flight.py
2 import numpy as np
3 import pandas as pd
4 import matplotlib.pyplot as plt
5 import seaborn as sns
6 from sklearn.model_selection import train_tes
7 #from sklearn.ensemble import RandomForestCla
8 from sklearn.tree import DecisionTreeClassifi
9 from sklearn.neighbors import KNeighborsClass
10 from sklearn.metrics import f1_score
11 from sklearn.metrics import classification_re
12 import warnings
13 import pickle
14 from scipy import stats
15 warnings.filterwarnings('ignore')
16 plt.style.use('fivethirtyeight')
18 data=pd.read_csv(r"C:\Users\DEEPANAYAKI\Deskt
19 data.head()
21 data.Date_of_Journey=data.Date_of_Journey.str
22 data.Date_of_Journey
24 for i in category:
25 print(i,data[i].unique())
26 data['Date']=data.Date_of_Journey.str[0]
12:08 PM 11-04-2023
11:37 PM 11-04-2023