

AIRLINES DATA ANALYTICS FOR AVIATION INDUSTRY

Team ID: PNT2022TMID16657

```
[4] import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

[6] data = pd.read_csv('/content/airports.csv')
data.drop(["id"], axis=1, inplace=True)
data.head()
```

	ident	type	name	latitude_deg	longitude_deg	elevation_ft	continent	iso_country	iso_region	municipality	scheduled_service	gps_
0	00A	heliport	Total Rf Heliport	40.070801	-74.933601	11.0	NaN	US	US-PA	Bensalem	no	
1	00AA	small_airport	Aero B Ranch Airport	38.704022	-101.473911	3435.0	NaN	US	US-KS	Leoti	no	0
2	00AK	small_airport	Lowell Field	59.947733	-151.692524	450.0	NaN	US	US-AK	Anchor Point	no	0
3	00AL	small_airport	Epps Airpark	34.864799	-86.770302	820.0	NaN	US	US-AL	Harvest	no	0
4	00AR	closed	Newport Hospital & Clinic	35.608700	-91.254898	237.0	NaN	US	US-AR	Newport	no	

```
[7] data.describe()
```

	latitude_deg	longitude_deg	elevation_ft
count	35703.000000	35703.000000	28537.000000
mean	32.146054	-33.209686	1217.982093
std	22.808783	82.179302	1529.671183
min	-89.989444	-169.511018	-210.000000
25%	30.753430	-92.494900	223.000000
50%	38.223701	-74.932899	728.000000
75%	44.843299	13.367192	1450.000000
max	82.750000	179.259167	22000.000000

Handling missing values

```
[8] data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35704 entries, 0 to 35703
Int64Index: 0 to 35703 in steps of 1
dtypes: object (12), float64 (3)
```

```
[7]
75%    44.843299    13.367192    1450.000000
max    82.750000    179.259167    22000.000000
```

Handling missing values

```
[8] data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 35704 entries, 0 to 35703
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype  
---  --
0   ident                 35703 non-null  object  
1   type                 35703 non-null  object  
2   name                 35703 non-null  object  
3   latitude_deg         35703 non-null  float64  
4   longitude_deg        35703 non-null  float64  
5   elevation_ft         28537 non-null  float64  
6   continent            16238 non-null  object  
7   iso_country          35623 non-null  object  
8   iso_region           35703 non-null  object  
9   municipality         33286 non-null  object  
10  scheduled_service    35703 non-null  object  
11  gps_code             22685 non-null  object  
12  iata_code            4111 non-null   object
```

memory usage: 4.6+ MB

```
[9] data.isnull().sum()

ident                1
type                1
name                1
latitude_deg        1
longitude_deg        1
elevation_ft        7167
continent          19466
iso_country          81
iso_region           1
municipality        2418
scheduled_service    1
gps_code           13019
iata_code           31593
local_code          15641
home_link           33532
wikipedia_link       30101
keywords            28669
dtype: int64
```

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

[ ] df=pd.read_csv("/content/airports.csv")

[ ] df.continent.unique()

array([nan, 'OC', 'AF', 'AN', 'EU', 'AS', 'SA'], dtype=object)

[ ] df=df.replace("no",0)
df=df.replace("yes",1)
#df = df.drop(['continent'],axis=1)
print(df.tail())
```

	id	ident	type	name	latitude_deg	\
14992	39834	CA-0188	closed	Kilometer 176 Airport	56.866665	
14993	39835	CA-0189	closed	Kimsquit Airport	52.900002	
14994	39836	CA-0190	closed	Kincaid Airport	49.666668	
14995	39837	CA-0191	closed	King Christian Airport	77.763338	
14996	39838	CA-0192	closed	Kivitoo Airport	67.933334	

	longitude_deg	elevation_ft	continent	iso_country	iso_region	\
14992	-106.150002	NaN	NaN	CA	CA-SK	
14993	-127.083336	NaN	NaN	CA	CA-BC	

ibm.ipynb ☆

File Edit View Insert Runtime Tools Help Last edited on 12 October

Comment Share

+ Code + Text

Connect Editing

```
[ ] df=df.replace('NaN',0)
df=df.replace('OC',1)
df=df.replace('AF',2)
df=df.replace('AN',3)
df=df.replace('EU',4)
df=df.replace('AS',5)
df=df.replace('SA',6)
print(df)
```

	id	ident	type	name	\
0	6523	00A	heliport	Total Rf Heliport	
1	323361	00AA	small_airport	Aero B Ranch Airport	
2	6524	00AK	small_airport	Lowell Field	
3	6525	00AL	small_airport	Epps Airpark	
4	6526	00AR	closed	Newport Hospital & Clinic Heliport	
...	
14992	39834	CA-0188	closed	Kilometer 176 Airport	
14993	39835	CA-0189	closed	Kimsquit Airport	
14994	39836	CA-0190	closed	Kincaid Airport	
14995	39837	CA-0191	closed	King Christian Airport	
14996	39838	CA-0192	closed	Kivitoo Airport	

	latitude_deg	longitude_deg	elevation_ft	continent	iso_country	\
0	40.070801	-74.933601	11.0	NaN	US	
1	38.704022	-101.473911	3435.0	NaN	US	
2	59.947733	-151.692524	450.0	NaN	US	
3	34.864799	-86.770302	820.0	NaN	US	