PROJECT DEVELOPMENT PHASE

PROJECT DEVELOPMENT DELIVERY OF SPRINT-II

TEAM ID	**PNT2022TMID46908**
PROJECT NAME	**AIRLINES DATA ANALYTICS FOR AVIATION INDUSTRIES**
DATE	**05 NOVEMBER 2022**

Coding:

In [18]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot
import seaborn as sns
```

In [19]:

```
data= pd.read_csv("/airp
data.head()
```

Out[19]:

	id	ident	type
0	6523	00A	heliport
1	323361	00AA	small_airport
2	6524	00AK	small_airport
3	6525	00AL	small_airport
4	6526	00AR	closed

In [20]:

```
data = pd.read_csv("/air
data.drop(["id"], axis=1
data.head()
```

Out[20]:

	ident	type	name
0	00A	heliport	Total Rf Heliport
1	00AA	small_airport	Aero B Ranch Airport
2	00AK	small_airport	Lowell Field
3	00AL	small_airport	Epps Airpark
4	00AR	closed	Newport Hospital & Clinic Heliport

In [21]:

```
data.describe()
```

Out[21]:

	latitude_deg	longitude_
count	67312.000000	67312.000
mean	25.945866	-31.136
std	26.380436	84.227
min	-90.000000	-179.876
25%	11.195161	-93.801
50%	35.437555	-70.799
75%	43.035376	18.963
max	82.750000	179.975

In [22]:
```
data.info()
```

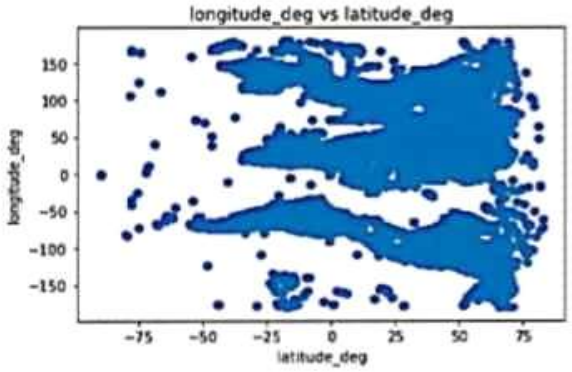
```
RangeIndex: 67312 entries, 0 to 67311
Data columns (total 17 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   ident
67311 non-null  object
 1   type
67312 non-null  object
 2   name
67312 non-null  object
 3   latitude_deg
67312 non-null  float64
 4   longitude_deg
67312 non-null  float64
 5   elevation_ft
54335 non-null  float64
 6   continent
34320 non-null  object
 7   iso_country
67055 non-null  object
 8   iso_region
67312 non-null  object
 9   municipality
61781 non-null  object
 10  scheduled_service
67312 non-null  object
 11  gps_code
42618 non-null  object
 12  iata_code
9244 non-null   object
 13  local_code
32055 non-null  object
 14  home_link
3300 non-null   object
 15  wikipedia_link
10370 non-null  object
 16  keywords
12367 non-null  object
dtypes: float64(3), object(14)
```

In [23]:
```
data.isnull().sum()
```

Out[23]:
```
ident
1
type
0
name
0
latitude_deg
0
longitude_deg
0
elevation_ft          129
77
continent             329
92
iso_country             2
57
iso_region
0
municipality           55
31
scheduled_service
0
gps_code              246
94
iata_code             580
68
local_code            352
57
home_link             640
12
wikipedia_link        569
42
keywords              549
45
dtype: int64
```

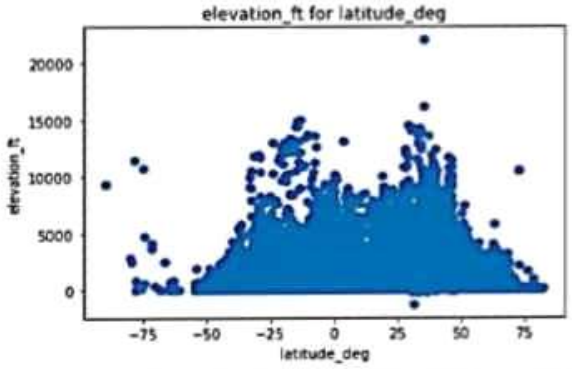
In [24]:

```
plt.scatter(data['latitu
plt.title('longitude_deg
plt.xlabel('latitude_deg
plt.ylabel('longitude_de
plt.show()
```



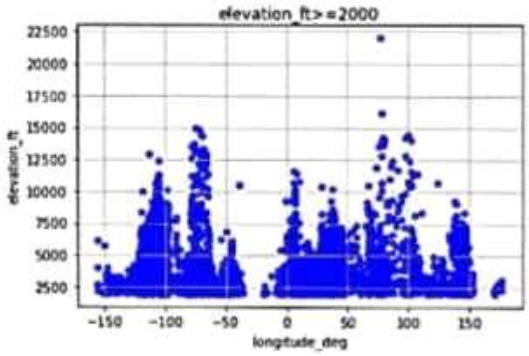
In [25]:

```
plt.scatter(data['latitu
plt.title('elevation_ft
plt.xlabel('latitude_deg
plt.ylabel('elevation_ft
plt.show()
```



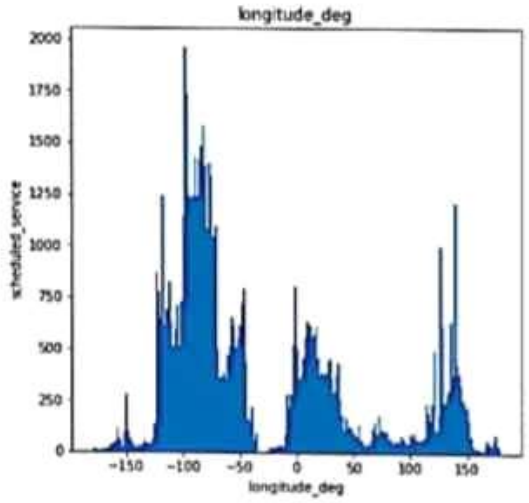
In [26]:

```
data[data.elevation_ft >
plt.xlabel("longitude_de
plt.ylabel("elevation_ft
plt.title("elevation_ft>
plt.grid(True)
plt.show()
```



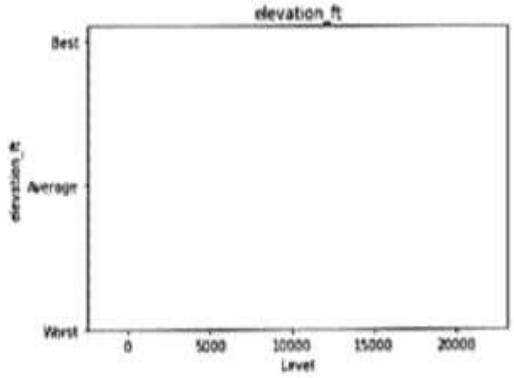
In [27]:

```
data["longitude_deg"].pl
plt.title("longitude_deg
plt.xlabel("longitude_de
plt.ylabel("scheduled_se
plt.show()
```



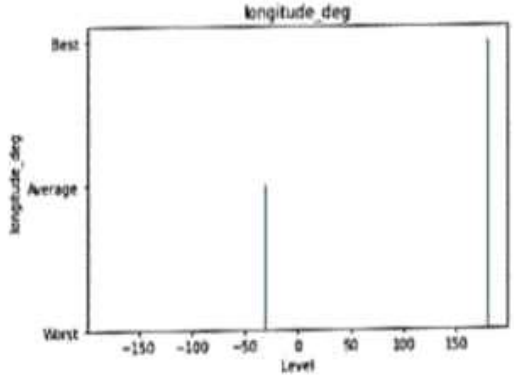
In [28]:

```
p = np.array([data["elev
r = ["Worst","Average",'
plt.bar(p,r)
plt.title("elevation_ft'
plt.xlabel("Level")
plt.ylabel("elevation_ft
plt.show()
```



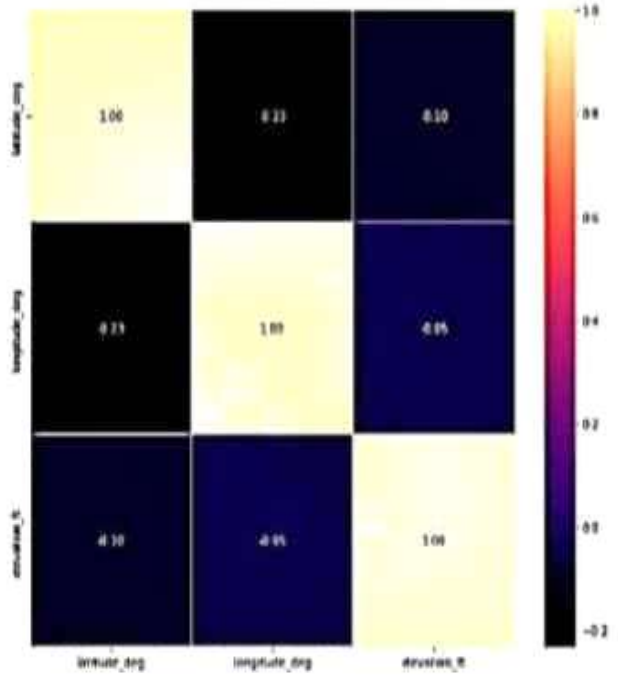
In [29]:

```
g = np.array([data["long
h = ["Worst","Average",'
plt.bar(g,h)
plt.title("longitude_deg
plt.xlabel("Level")
plt.ylabel("longitude_de
plt.show()
```



In [30]:

```
plt.figure(figsize=(10,
sns.heatmap(data.corr(),
plt.show()
```



In [31]:

```
print(data.shape)
n = len(pd.unique(data['
d=len(pd.unique(data['ty
print("name",n,"type",d)
data['scheduled_service'
```

```
(67312, 17)
name 63826 type 7
```

Out[31]:

```
no      63228
yes      4084
Name: scheduled_service,
dtype: int64
```