

Question 7 - Using the data from Question 4, write code to analyze the data and answer the following questions Note -

1. Draw plots to demonstrate the analysis for the following questions for better visualizations

2. Write code comments wherever required for code understanding

Insights to be drawn -

- Get all the Earth meteorites that fell before the year 2000
- Get all the earth meteorites co-ordinates who fell before the year 1970
- Assuming that the mass of the earth meteorites was in kg, get all those whose mass was more than 10000kg

Ans:

```
In [47]: import numpy as np
import pandas as pd
import requests
import json
import csv
```

```
In [48]: df = pd.read_csv(r"C:\Users\hrush\Downloads\y77d-th95.csv")
```

```
In [49]: df.head()
```

```
Out[49]:
```

	name	id	nametype	recclass	mass	fall	year	reclat	reclong	geolocation/type	geolocation/coordinates/0	ge
0	Aachen	1	Valid	L5	21.0	Fell	1880-01-01T00:00:00.000	50.77500	6.08333	Point	6.08333	
1	Aarhus	2	Valid	H6	720.0	Fell	1951-01-01T00:00:00.000	56.18333	10.23333	Point	10.23333	
2	Abee	6	Valid	EH4	107000.0	Fell	1952-01-01T00:00:00.000	54.21667	-113.00000	Point	-113.00000	
3	Acapulco	10	Valid	Acapulcoite	1914.0	Fell	1976-01-01T00:00:00.000	16.88333	-99.90000	Point	-99.90000	
4	Achiras	370	Valid	L6	780.0	Fell	1902-01-01T00:00:00.000	-33.16667	-64.95000	Point	-64.95000	

```
In [50]: df.shape
```

```
Out[50]: (1000, 14)
```

```
In [51]: # Get all the Earth meteorites that fell before the year 2000

df["only_year"] = df["year"].apply(lambda x : str(x).split("T")[0].split("-")[0])
df["before_2000"] = df["only_year"].astype(float) < 2000
df.loc[df["before_2000"] == True]["name"]
```

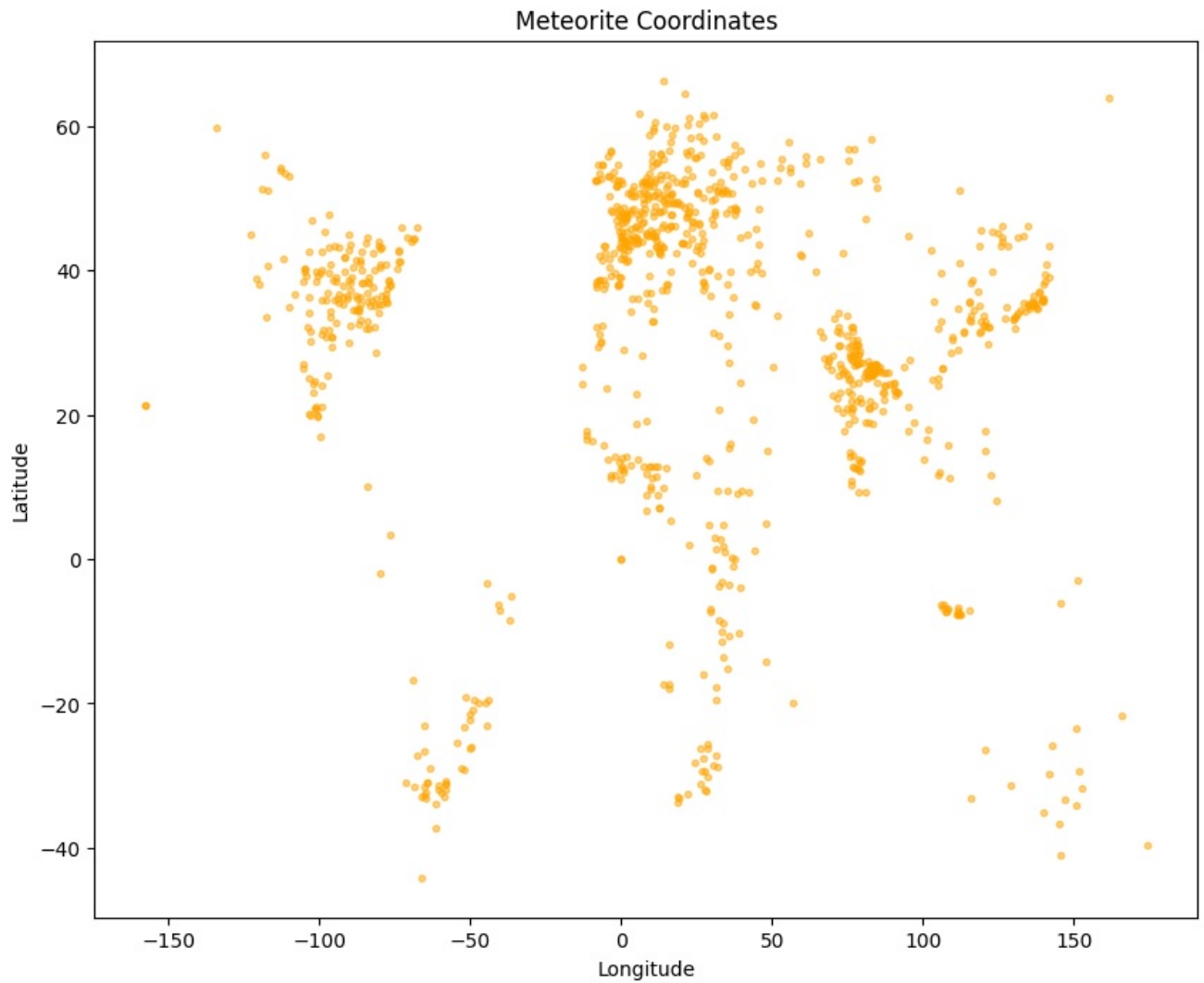
```
Out[51]: 0      Aachen
1      Aarhus
2      Abee
3      Acapulco
4      Achiras
...
994    Timochin
995    Tirupati
997      Tjabe
998    Tjerebon
999    Tomakovka
Name: name, Length: 929, dtype: object
```

```
In [54]: # Assuming that the mass of the earth meteorites was in kg, get all those whose mass was more
# than 10000kg
```

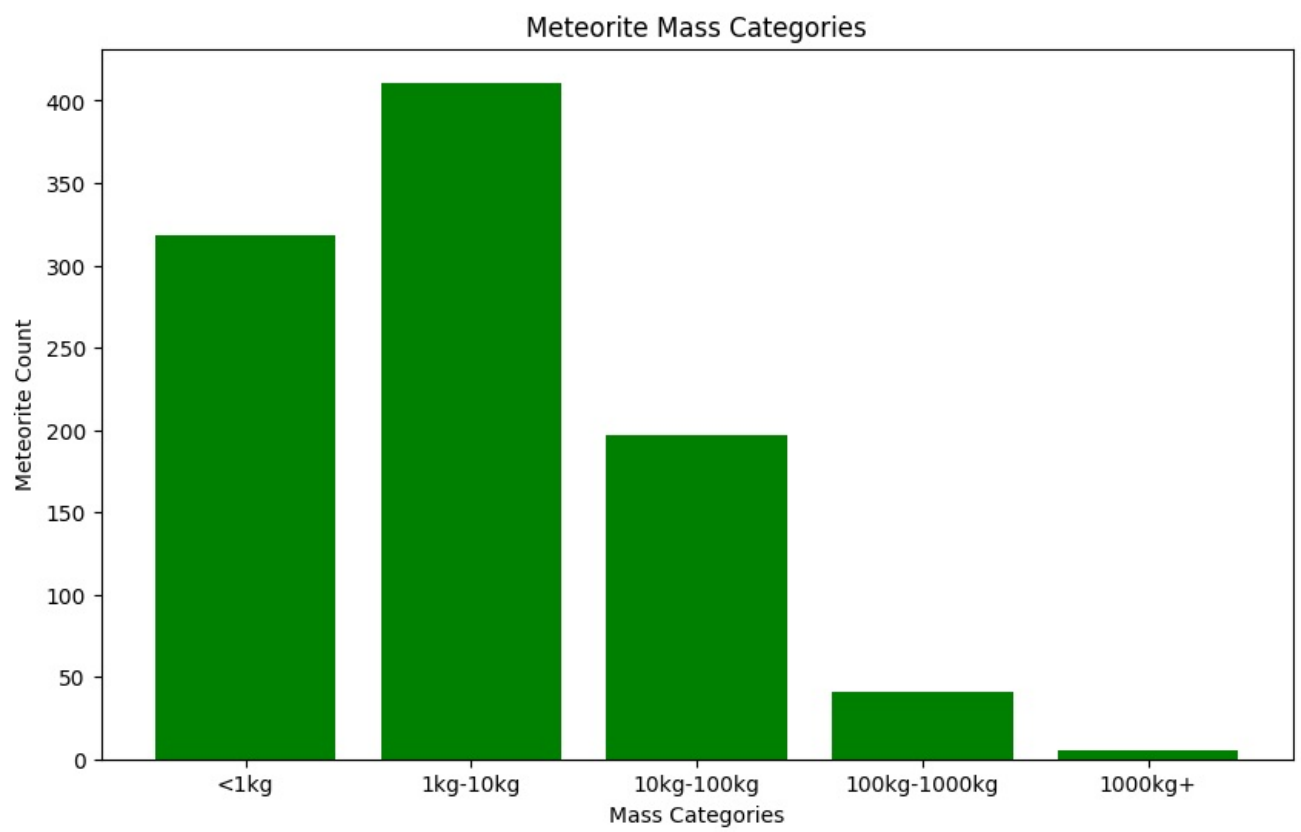
```
df["col"] = df["mass"] > 10000
df.loc[df["col"] == True]["name"]
```

```
Out[54]: 2      Abee
7      Agen
11     Aïr
16     Akyumak
27     Alfianello
...
991    Tieschitz
992    Tilden
994    Timochin
997    Tjabe
998    Tjerebon
Name: name, Length: 243, dtype: object
```

```
In [56]: import matplotlib.pyplot as plt
plt.figure(figsize=(10, 8))
plt.scatter(df['reclong'], df['reclat'], s=10, alpha=0.5, color='orange')
plt.xlabel('Longitude')
plt.ylabel('Latitude')
plt.title('Meteorite Coordinates')
plt.show()
```



```
In [57]: mass_categories = ['<1kg', '1kg-10kg', '10kg-100kg', '100kg-1000kg', '1000kg+']
mass_counts = df.groupby(pd.cut(df['mass'], bins=[0, 1000, 10000, 100000, 1000000, float('inf')])).size()
plt.figure(figsize=(10, 6))
plt.bar(mass_categories, mass_counts, color='green')
plt.xlabel('Mass Categories')
plt.ylabel('Meteorite Count')
plt.title('Meteorite Mass Categories')
plt.show()
```



In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js