Weather Dataset

In [21]:
```
import pandas as pd
import numpy as np
import matplotlib.pylab as plt
import seaborn as sns
plt.style.use('ggplot')
```

Putting Data into a Data Frame

In [22]:
```
df = pd.read_csv(r"C:\Users\logap\Downloads\archive\top100cities_weather_data.csv")
```

In [23]:
```
df
```

Out[23]:

	City	Temperature (Celsius)	Wind Speed (m/s)	Latitude	Longitude	Description	Country
0	Paris	10.46	2.06	48.8534	2.3488	clear sky	France
1	Dubai	29.32	3.09	25.2582	55.3047	clear sky	United Arab Emirates
2	Madrid	9.97	3.60	40.4165	-3.7026	clear sky	Spain
3	Tokyo	21.58	2.06	35.6895	139.6917	broken clouds	Japan
4	Amsterdam	8.00	3.58	52.3740	4.8897	clear sky	Netherlands
...	...	...	...	...	...	...	...
95	Punta Cana	26.87	6.69	18.5818	-68.4043	few clouds	Dominican Republic
96	Guilin	24.06	0.97	25.2819	110.2864	clear sky	China
97	Hanoi	28.00	2.49	21.0245	105.8412	clear sky	Vietnam
98	Cairo	19.42	4.63	30.0626	31.2497	clear sky	Egypt
99	Muscat	28.01	2.76	23.6139	58.5922	broken clouds	Oman

100 rows × 7 columns

In [25]:
```
df.head()
```

Out[25]:

	City	Temperature (Celsius)	Wind Speed (m/s)	Latitude	Longitude	Description	Country
0	Paris	10.46	2.06	48.8534	2.3488	clear sky	France
1	Dubai	29.32	3.09	25.2582	55.3047	clear sky	United Arab Emirates
2	Madrid	9.97	3.60	40.4165	-3.7026	clear sky	Spain
3	Tokyo	21.58	2.06	35.6895	139.6917	broken clouds	Japan
4	Amsterdam	8.00	3.58	52.3740	4.8897	clear sky	Netherlands

Data Cleaning

In [26]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 7 columns):
 #   Column                 Non-Null Count  Dtype  
---  ------                 --------------  -----  
 0   City                   100 non-null    object 
 1   Temperature (Celsius)  100 non-null    float64
 2   Wind Speed (m/s)       100 non-null    float64
 3   Latitude               100 non-null    float64
 4   Longitude              100 non-null    float64
 5   Description            100 non-null    object 
 6   Country                100 non-null    object 
dtypes: float64(4), object(3)
memory usage: 5.6+ KB
```

Check for Missing Values

In [27]:
```
df.isnull().sum()
```

Out[27]:

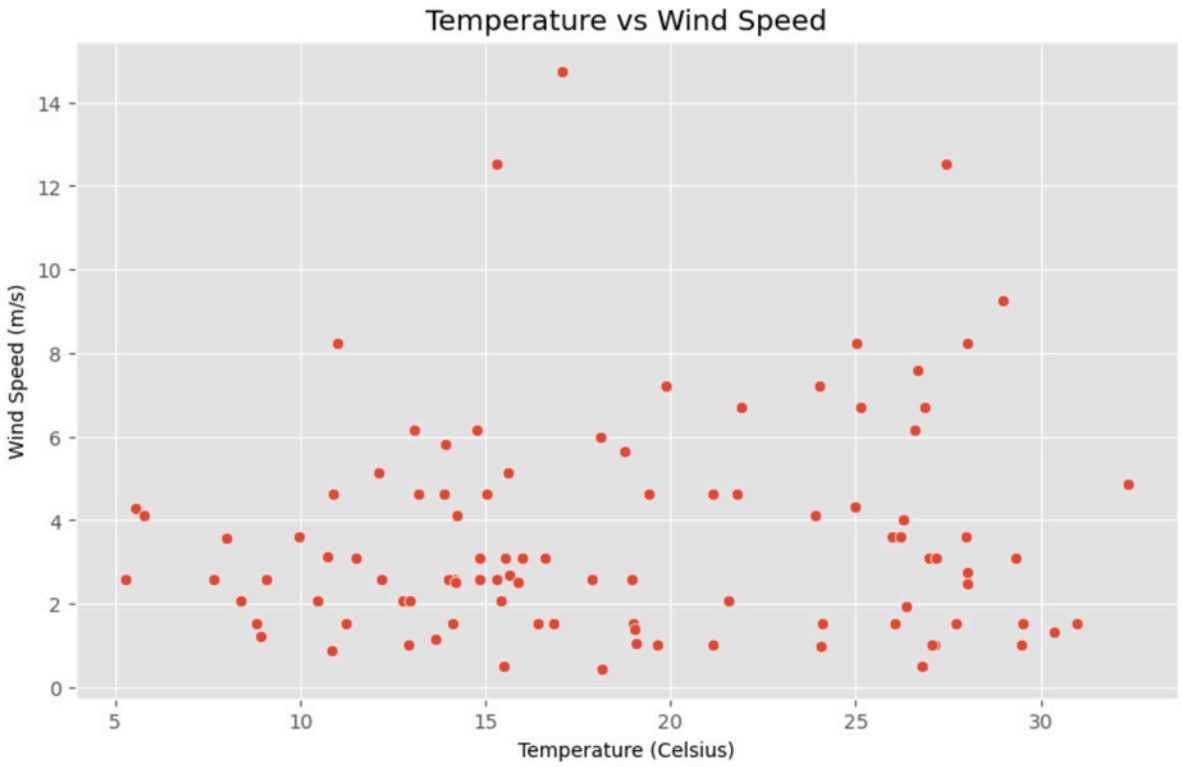
```
City                    0
Temperature (Celsius)   0
Wind Speed (m/s)        0
Latitude                0
Longitude               0
Description             0
Country                 0
dtype: int64
```

Data is clean so we dont need to perform any other method

Data Visualization

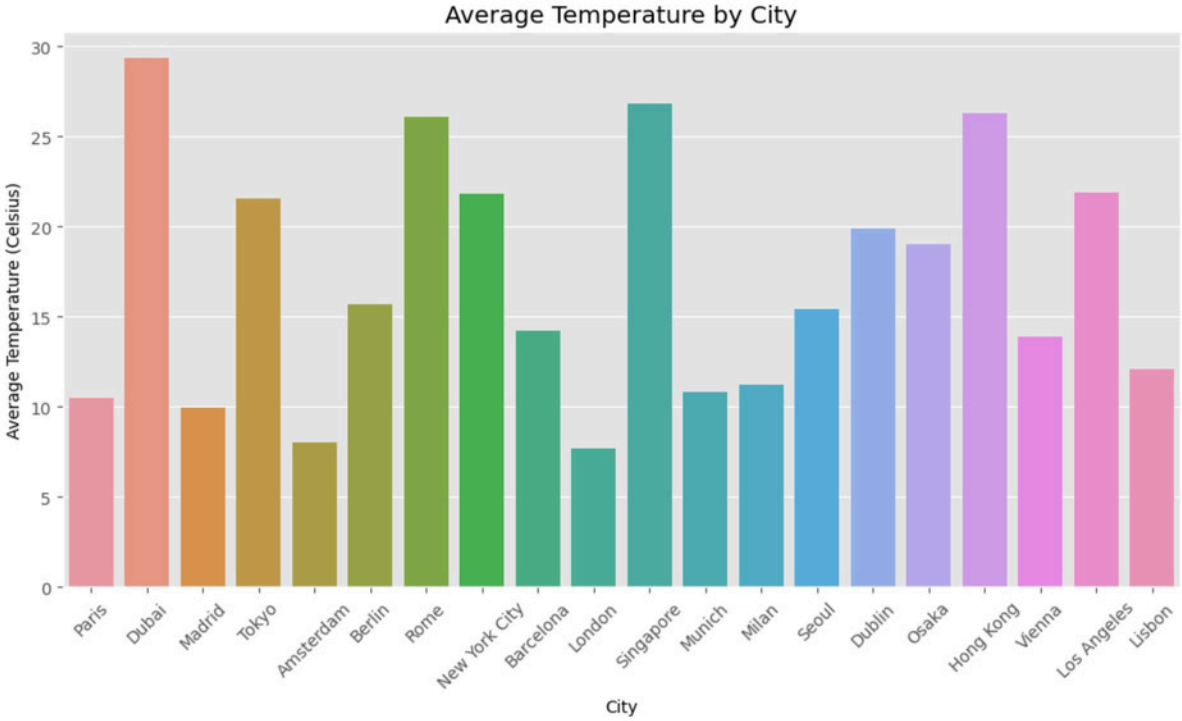
A scatterplot shows the relationship between two quantitative variables measured for the same individuals

In [14]:
```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='Temperature (Celsius)', y='Wind Speed (m/s)', data=df)
plt.title('Temperature vs Wind Speed')
plt.xlabel('Temperature (Celsius)')
plt.ylabel('Wind Speed (m/s)')
plt.show()
```



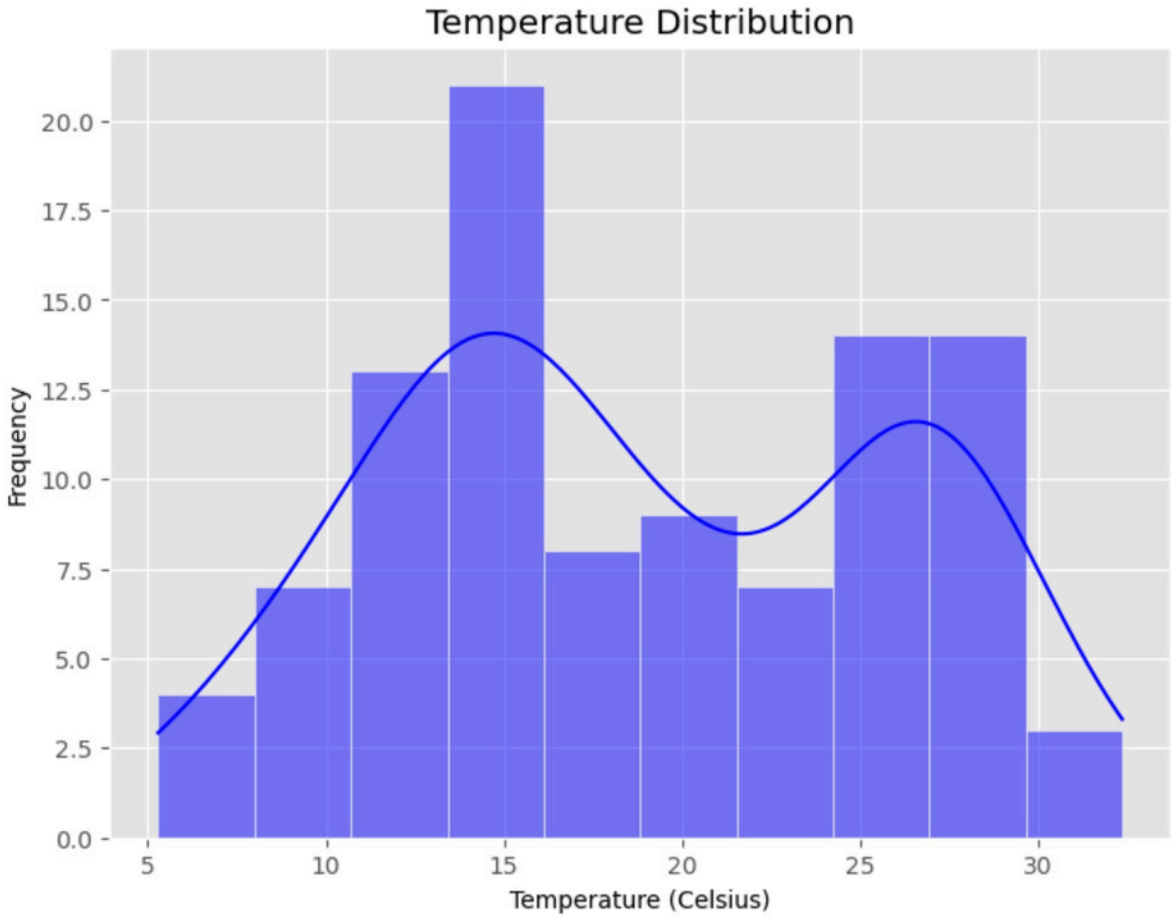
It shows the relationship between a numeric and a categoric variable.

In [28]:
```
plt.figure(figsize=(12, 6))
sns.barplot(x='City', y='Temperature (Celsius)', data=df.head(20))
plt.title('Average Temperature by City')
plt.xticks(rotation=45)
plt.ylabel('Average Temperature (Celsius)')
plt.xlabel('City')
plt.show()
```



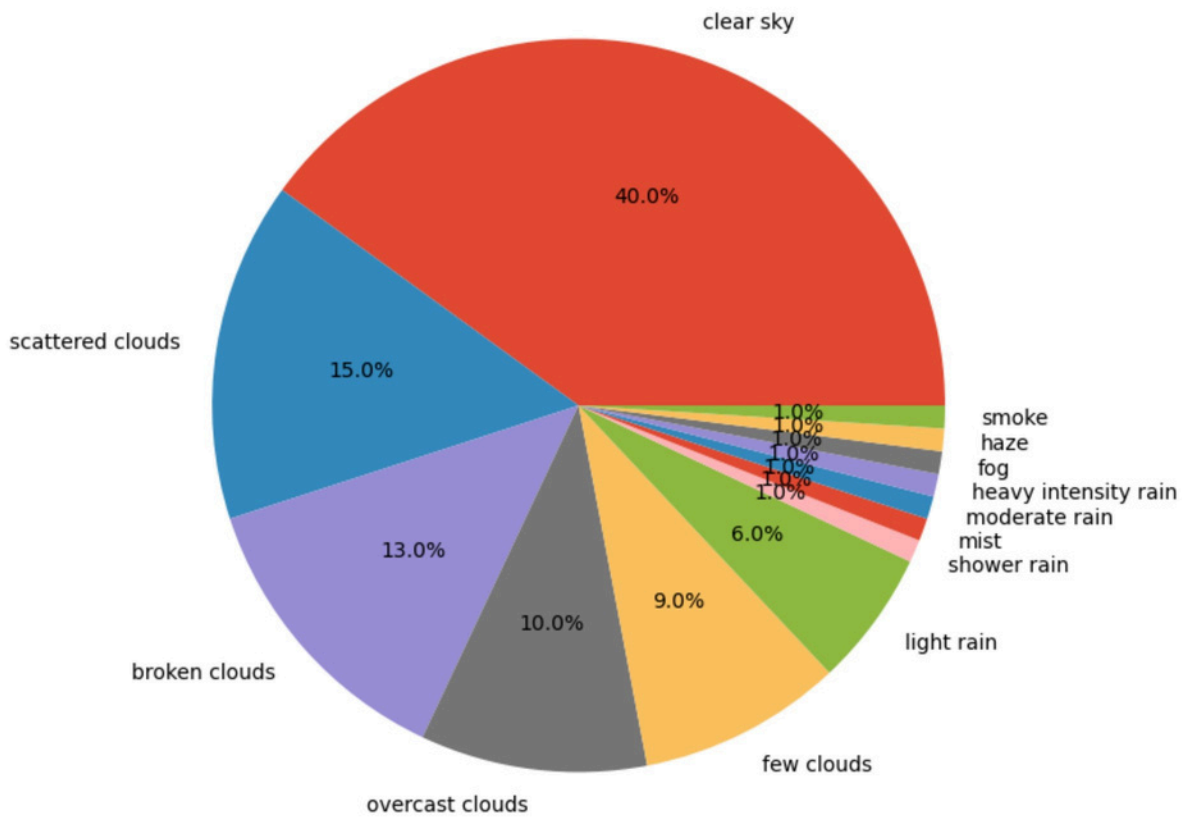
A statistical graph that represents the distribution of a continuous dataset through plotted bars, each representing a particular category or class interval.

In [19]:
```
plt.figure(figsize=(8, 6))
sns.histplot(df['Temperature (Celsius)'], bins=10, kde=True,color='blue')
plt.title('Temperature Distribution')
plt.xlabel('Temperature (Celsius)')
plt.ylabel('Frequency')
plt.show()
```



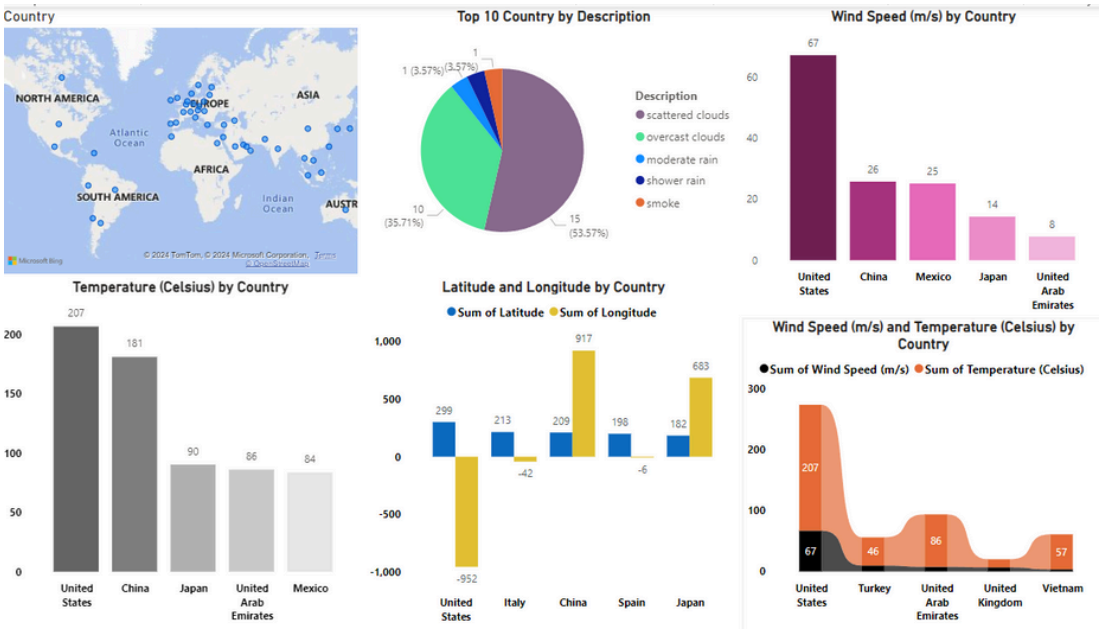
It contains different segments and sectors in which each segment and sector of a pie chart forms a specific portion of the total(percentage).

In [20]:
```
plt.figure(figsize=(8, 8))
df['Description'].value_counts().plot(kind='pie', autopct='%1.1f%%')
plt.title('Weather Description Distribution')
plt.ylabel('')
plt.show()
```

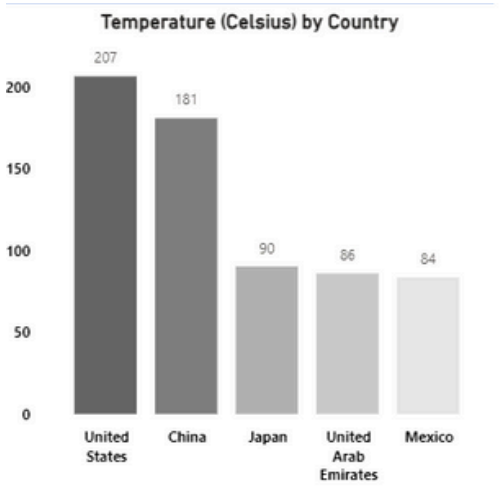

Weather Description Distribution
clear sky
40.0%
scattered clouds
15.0%
broken clouds
13.0%
overcast clouds
10.0%
few clouds
9.0%
light rain
6.0%
shower rain
mist
moderate rain
heavy intensity rain
fog
haze
smoke
1.0%

In [ ]:

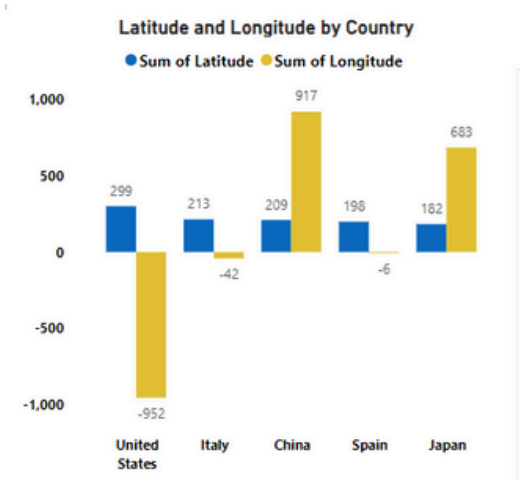
Weather Dashboard



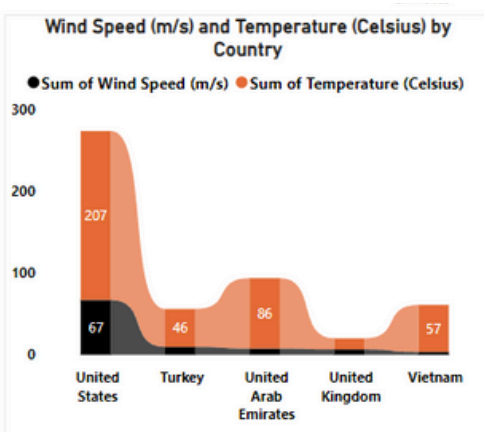
It shows various weather data for different countries, including temperature, wind speed, description (e.g. scattered clouds, overcast clouds, moderate rain) , latitude, longitude and sums. For example, China has overcast clouds and a wind speed of 17 ms. Overall, the Dashboard provides a summary of weather data from various countries.



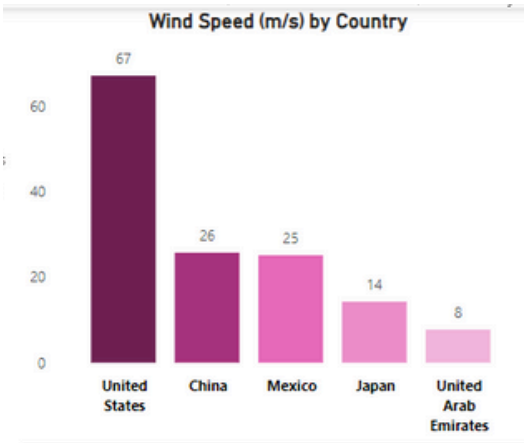
The bar graph shows the average temperature in six countries: United States, China, Japan, Mexico, United Arab Emirates, and an unidentified country. The United States has the highest average temperature at 207 degrees Celsius, while the unidentified country has the lowest at 84 degrees Celsius. It is important to note that the temperature scale is likely Celsius, not Fahrenheit, as 207 degrees Celsius is far above the boiling point of water.



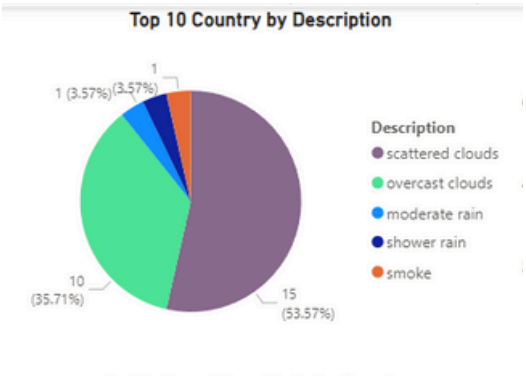
The graph shows the sum of latitude and longitude by country. It likely doesn't represent meaningful geographic data because summing latitude and longitude together doesn't correspond to a standard geographic measurement. The countries with the highest values are the United States, Italy, China, Spain, Japan, and the United Kingdom.



The bar graph likely shows average wind speed and temperature in several countries. The United States has the highest average wind speed (around 10 meters per second) and temperature (around 77 degrees Fahrenheit). It's important to note that the scale might be incorrect as 77 degrees Fahrenheit is unusually high for an average temperature.



The bar graph shows average wind speeds in different countries. The United States has the highest average wind speed at 67 miles per hour, while the United Arab Emirates has the lowest at 5 miles per hour. Countries included are the United States, China, Mexico, Japan, and the United Arab Emirates.



The pie chart shows the percentage of different cloud types observed in various countries. Smoke makes up the largest portion (35.71%) followed by overcast clouds (15.38%) and scattered clouds (3.57%). The data for ten countries is included, but only percentages for the top three are shown.

Conclusion :

The weather dashboard shows a summary of weather conditions for various countries around the world. It includes information on temperature, wind speed, and weather description. North America has the most descriptions (35.71%), while Asia has the most wind speed (53.57%). The United States has the highest temperature (207°C) and wind speed (67 m/s).