Name-Divya gatkal

Roll no- 313

div - C (C1 batch)

PRN- 202201040043

_____

ASSIGNMENT NO: 5

```python
import pandas as pd
df=pd.read_csv("/content/Book1 Weather forecast (1).csv")
df
```

index	2006-04-01 00:00:00.000 +0200	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Loud Cover	Pressure (millibars)	Daily Summary
0	2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.472222222	7.388888889	0.89	14.1197	251	15.8263	0	1015.13	Partly cloudy throughout the day.
1	2006-04-01 02:00:00.000 +0200	Partly Cloudy	rain	9.355555556	7.227777778	0.86	14.2646	259	15.8263	0	1015.63	Partly cloudy throughout the day.
2	2006-04-01 03:00:00.000 +0200	Mostly Cloudy	rain	9.377777778	9.377777778	0.89	3.9284	204	14.9569	0	1015.94	Partly cloudy throughout the day.
3	2006-04-01 04:00:00.000 +0200	Partly Cloudy	rain	8.288888889	5.944444444	0.83	14.1036	269	15.8263	0	1016.41	Partly cloudy throughout the day.
4	2006-04-01 05:00:00.000 +0200	Mostly Cloudy	rain	8.755555556	6.977777778	0.83	11.0446	259	15.8263	0	1016.51	Partly cloudy throughout the day.
5	2006-04-01 06:00:00.000 +0200	Partly Cloudy	rain	9.222222222	7.111111111	0.85	13.9587	258	14.9569	0	1016.66	Partly cloudy throughout the day.
6	2006-04-01 07:00:00.000 +0200	Partly Cloudy	rain	7.733333333	5.522222222	0.95	12.3648	259	9.982	0	1016.72	Partly cloudy throughout the day.
7	2006-04-01 08:00:00.000 +0200	Partly Cloudy	rain	8.772222222	6.527777778	0.89	14.1519	260	9.982	0	1016.84	Partly cloudy throughout the day.
8	2006-04-01 09:00:00.000 +0200	Partly Cloudy	rain	10.82222222	10.82222222	0.82	11.3183	259	9.982	0	1017.37	Partly cloudy throughout the day.
9	2006-04-01 10:00:00.000 +0200	Partly Cloudy	rain	13.77222222	13.77222222	0.72	12.5258	279	9.982	0	1017.22	Partly cloudy throughout the day.
10	2006-04-01 11:00:00.000 +0200	Partly Cloudy	rain	16.01666667	16.01666667	0.67	17.5651	290	11.2056	0	1017.42	Partly cloudy throughout the day.
11	2006-04-01 12:00:00.000 +0200	Partly Cloudy	rain	17.14444444	17.14444444	0.54	19.7869	316	11.4471	0	1017.74	Partly cloudy throughout the day.
12	2006-04-01 13:00:00.000 +0200	Partly Cloudy	rain	17.8	17.8	0.55	21.9443	281	11.27	0	1017.59	Partly cloudy throughout the day.
13	2006-04-01 14:00:00.000 +0200	Partly Cloudy	rain	17.33333333	17.33333333	0.51	20.6885	289	11.27	0	1017.48	Partly cloudy throughout the day.
14	2006-04-01 15:00:00.000 +0200	Partly Cloudy	rain	18.87777778	18.87777778	0.47	15.3755	262	11.4471	0	1017.17	Partly cloudy throughout the day.
15	2006-04-01 16:00:00.000 +0200	Partly Cloudy	rain	18.91111111	18.91111111	0.46	10.4006	288	11.27	0	1016.47	Partly cloudy throughout the day.
16	2006-04-01 17:00:00.000 +0200	Partly Cloudy	rain	15.38888889	15.38888889	0.6	14.4095	251	11.27	0	1016.15	Partly cloudy throughout the day.
17	2006-04-01 18:00:00.000 +0200	Mostly Cloudy	rain	15.55	15.55	0.63	11.1573	230	11.4471	0	1016.17	Partly cloudy throughout the day.
18	2006-04-01 19:00:00.000 +0200	Mostly Cloudy	rain	14.25555556	14.25555556	0.69	8.5169	163	11.2056	0	1015.82	Partly cloudy throughout the day.
19	2006-04-01 20:00:00.000 +0200	Mostly Cloudy	rain	13.14444444	13.14444444	0.7	7.6314	139	11.2056	0	1015.83	Partly cloudy throughout the day.
20	2006-04-01 21:00:00.000 +0200	Mostly Cloudy	rain	11.55	11.55	0.77	7.3899	147	11.0285	0	1015.85	Partly cloudy throughout the day.
21	2006-04-01 22:00:00.000 +0200	Mostly Cloudy	rain	11.18333333	11.18333333	0.76	4.9266	160	9.982	0	1015.77	Partly cloudy throughout the day.
22	2006-04-01 23:00:00.000 +0200	Partly Cloudy	rain	10.11666667	10.11666667	0.79	6.6493	163	15.8263	0	1015.4	Partly cloudy throughout the day.
23	2006-04-10 00:00:00.000 +0200	Mostly Cloudy	rain	10.2	10.2	0.77	3.9284	152	14.9569	0	1015.51	Partly cloudy throughout the day.
24	2006-04-10 01:00:00.000 +0200	Partly Cloudy	rain	10.42222222	10.42222222	0.62	16.9855	150	15.8263	0	1014.4	Mostly cloudy throughout the day.

Show 25 ∨ per page   1 2 3 4 5 6

```python
import pandas as pd

# Load the CSV file into a DataFrame
data = pd.read_csv('/content/Book1 Weather forecast (1).csv')

# Display the first few rows of the DataFrame
print(data.head())

# Perform data analysis tasks on the DataFrame
# For example, you can calculate basic statistics
print("\nData Statistics:")
print(data.describe())

# filter and subset the data based on specific conditions
# filtered data for temperature greater than 20 degrees
filtered_data = data[data['Temperature (C)'] > 20]
print("\nFiltered Data (Temperature > 20):")
print(filtered_data)

# other analysis tasks like grouping, aggregating, plotting, etc.
# grouping by 'Precip Type' and calculating the average temperature
grouped_data = data.groupby('Precip Type')['Temperature (C)'].mean()
print("\nAverage Temperature by Precip Type:")
print(grouped_data)
```

OUTPUT:

```
2006-04-01 00:00:00.000 +0200 Summary Precip Type Temperature (C) \
0 2006-04-01 01:00:00.000 +0200 Partly Cloudy rain 9.472222
1 2006-04-01 02:00:00.000 +0200 Partly Cloudy rain 9.355556
2 2006-04-01 03:00:00.000 +0200 Mostly Cloudy rain 9.377778
3 2006-04-01 04:00:00.000 +0200 Partly Cloudy rain 8.288889
4 2006-04-01 05:00:00.000 +0200 Mostly Cloudy rain 8.755556

Apparent Temperature (C) Humidity Wind Speed (km/h) \

0 7.388889 0.89 14.1197
1 7.227778 0.86 14.2646
2 9.377778 0.89 3.9284
3 5.944444 0.83 14.1036
4 6.977778 0.83 11.0446
      Wind Bearing (degrees) Visibility (km) Loud Cover Pressure (millibars) \
0 251 15.8263 0 1015.13
1 259 15.8263 0 1015.63
2 204 14.9569 0 1015.94
3 269 15.8263 0 1016.41
4 259 15.8263 0 1016.51

Daily Summary

0 Partly cloudy throughout the day.
1 Partly cloudy throughout the day.
2 Partly cloudy throughout the day.
3 Partly cloudy throughout the day.
4 Partly cloudy throughout the day.

Data Statistics:

Temperature (C) Apparent Temperature (C) Humidity \
count 149.000000 149.000000 149.00000
mean 10.927442 9.776137 0.74094
```

```
std 4.110883 5.173100 0.14792
min 3.722222 1.494444 0.36000
25% 7.688889 5.294444 0.66000
50% 9.911111 9.377778 0.75000
75% 13.838889 13.838889 0.85000
max 21.183333 21.183333 0.99000

          Wind Speed (km/h) Wind Bearing (degrees) Visibility (km) Loud Cover \
count 149.000000 149.000000 149.000000 149.0
mean 14.853817 217.328859 11.086957 0.0
std 7.554826 95.024424 3.262636 0.0
min 0.322000 0.000000 1.223600 0.0
25% 9.804900 152.000000 9.982000 0.0
50% 14.248500 215.000000 10.899700 0.0
75% 20.447000 309.000000 14.168000 0.0
max 32.167800 359.000000 15.874600 0.0

Pressure (millibars)
count 149.000000
mean 1010.461409
std 4.514743
min 1003.570000
25% 1005.970000
50% 1010.390000
75% 1014.490000
max 1017.740000

Filtered Data (Temperature > 20):
            2006-04-01 00:00:00.000 +0200 Summary Precip Type Temperature (C) \
37 2006-04-10 14:00:00.000 +0200 Mostly Cloudy rain 20.044444
38 2006-04-10 15:00:00.000 +0200 Mostly Cloudy rain 21.050000
39 2006-04-10 16:00:00.000 +0200 Mostly Cloudy rain 21.183333
40 2006-04-10 17:00:00.000 +0200 Mostly Cloudy rain 20.116667
41 2006-04-10 18:00:00.000 +0200 Mostly Cloudy rain 20.216667

Apparent Temperature (C) Humidity Wind Speed (km/h) \
37 20.044444 0.40 28.3682
38 21.050000 0.40 26.9031
39 21.183333 0.37 25.6956
40 20.116667 0.40 25.3092
41 20.216667 0.36 18.1125

        Wind Bearing (degrees) Visibility (km) Loud Cover Pressure (millibars) \
37 170 9.9820 0 1012.22
38 187 10.3523 0 1011.44
39 179 9.9820 0 1010.52
40 162 9.9820 0 1009.83
41 161 10.3523 0 1009.26

Daily Summary
37 Mostly cloudy throughout the day.
38 Mostly cloudy throughout the day.
39 Mostly cloudy throughout the day.
40 Mostly cloudy throughout the day.
41 Mostly cloudy throughout the day.

Average Temperature by Precip Type:
Precip Type
rain 10.927442
Name: Temperature (C), dtype: float64
```

```python
import matplotlib.pyplot as plt
import seaborn as sns

# Set the style for Seaborn plots
sns.set(style="darkgrid")

# Plot 1: Line plot of Temperature over time
plt.figure(figsize=(12, 6))
plt.plot(data.index, data['Temperature (C)'])
plt.title('Temperature over Time')
plt.xlabel('Index')
plt.ylabel('Temperature (C)')
plt.show()

# Plot 2: Bar plot of Summary
plt.figure(figsize=(12, 6))
sns.countplot(data=data, x='Summary')
plt.title('Summary')
plt.xlabel('Summary')
plt.ylabel('Count')
plt.xticks(rotation=45)
plt.show()

# Plot 3: Box plot of Temperature by Precip Type
plt.figure(figsize=(12, 6))
sns.boxplot(data=data, x='Precip Type', y='Temperature (C)')
plt.title('Temperature by Precip Type')
plt.xlabel('Precip Type')
plt.ylabel('Temperature (C)')
plt.show()

# Plot 4: Histogram of Humidity
plt.figure(figsize=(12, 6))
sns.histplot(data=data, x='Humidity', bins=10)
plt.title('Humidity Distribution')
plt.xlabel('Humidity')
plt.ylabel('Count')
plt.show()

# Plot 5: Scatter plot of Temperature vs. Wind Speed
plt.figure(figsize=(12, 6))
sns.scatterplot(data=data, x='Temperature (C)', y='Wind Speed (km/h)')
plt.title('Temperature vs. Wind Speed')
plt.xlabel('Temperature (C)')
plt.ylabel('Wind Speed (km/h)')
plt.show()

# Plot 6: Violin plot of Apparent Temperature by Summary
plt.figure(figsize=(12, 6))
sns.violinplot(data=data, x='Summary', y='Apparent Temperature (C)')
plt.title('Apparent Temperature by Summary')
plt.xlabel('Summary')
plt.ylabel('Apparent Temperature (C)')
plt.xticks(rotation=45)
```

```python
plt.show()

# Plot 7: Bar plot of Average Temperature by Precip Type
plt.figure(figsize=(12, 6))
sns.barplot(data=grouped_data.reset_index(), x='Precip Type', y='Temperature (C)')
plt.title('Average Temperature by Precip Type')
plt.xlabel('Precip Type')
plt.ylabel('Average Temperature (C)')
plt.show()

# Plot 8: Heatmap of Correlation Matrix
plt.figure(figsize=(10, 8))
correlation = data.corr()
sns.heatmap(correlation, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()

# Plot 9: Pairplot of select columns
plt.figure(figsize=(12, 12))
columns = ['Temperature (C)', 'Apparent Temperature (C)', 'Humidity', 'Wind Speed
(km/h)', 'Visibility (km)']
sns.pairplot(data[columns])
plt.title('Pairplot of Select Columns')
plt.show()

# Plot 10: Pie chart of Precip Type distribution
plt.figure(figsize=(6, 6))
precip_type_counts = data['Precip Type'].value_counts()
plt.pie(precip_type_counts, labels=precip_type_counts.index, autopct='%1.1f%%')
plt.title('Precip Type Distribution')
plt.show()
```
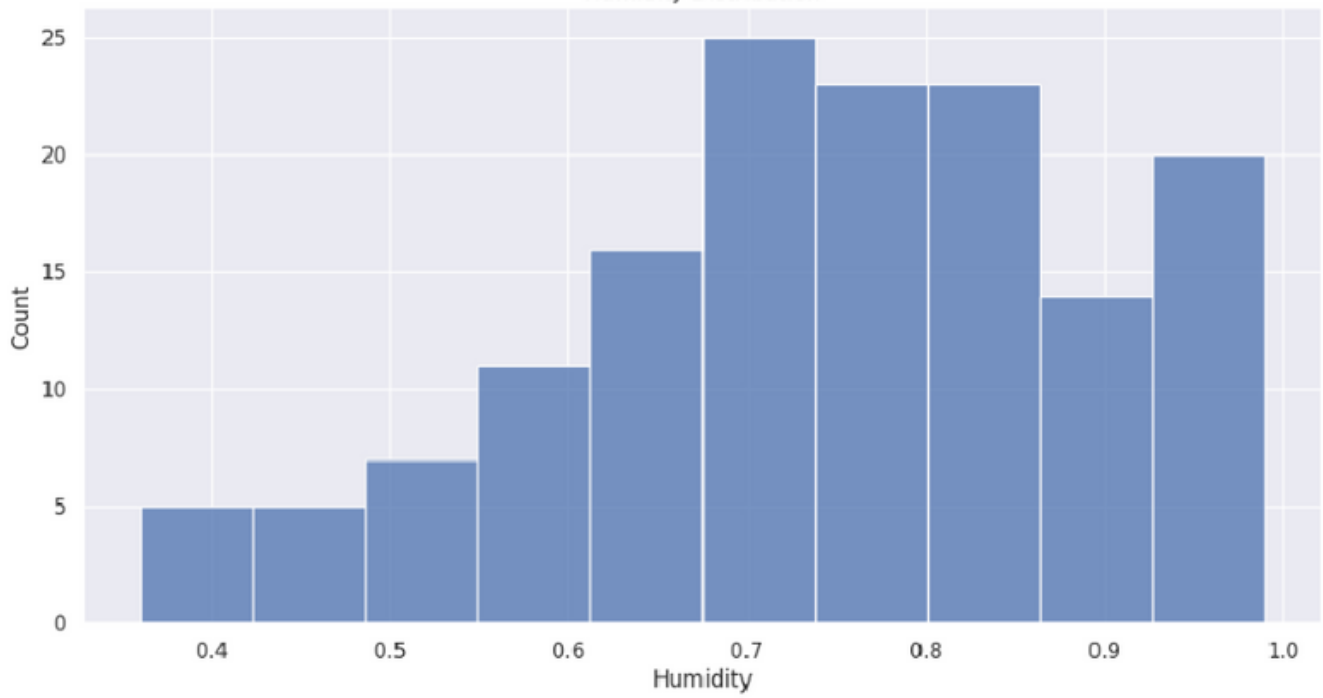
Temperature over Time
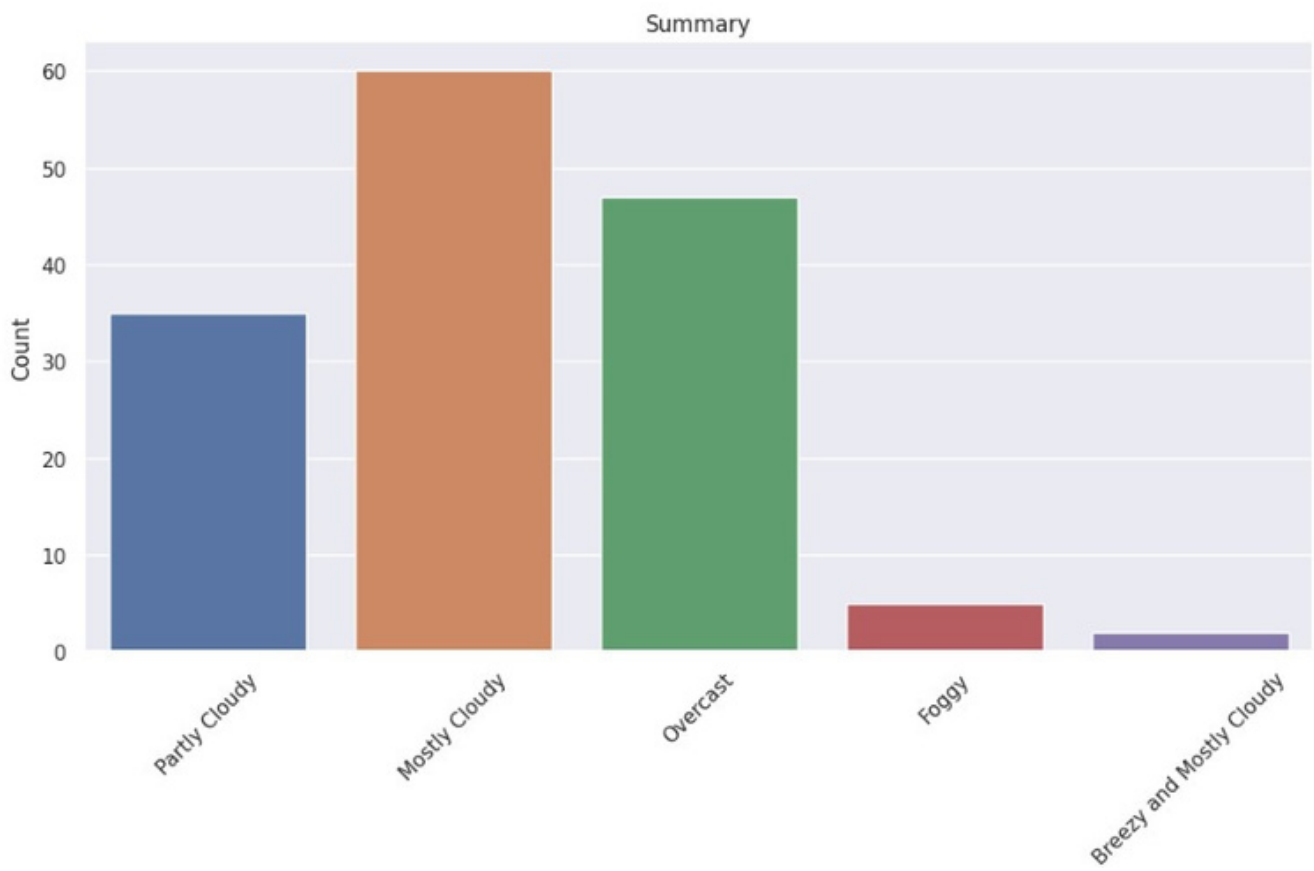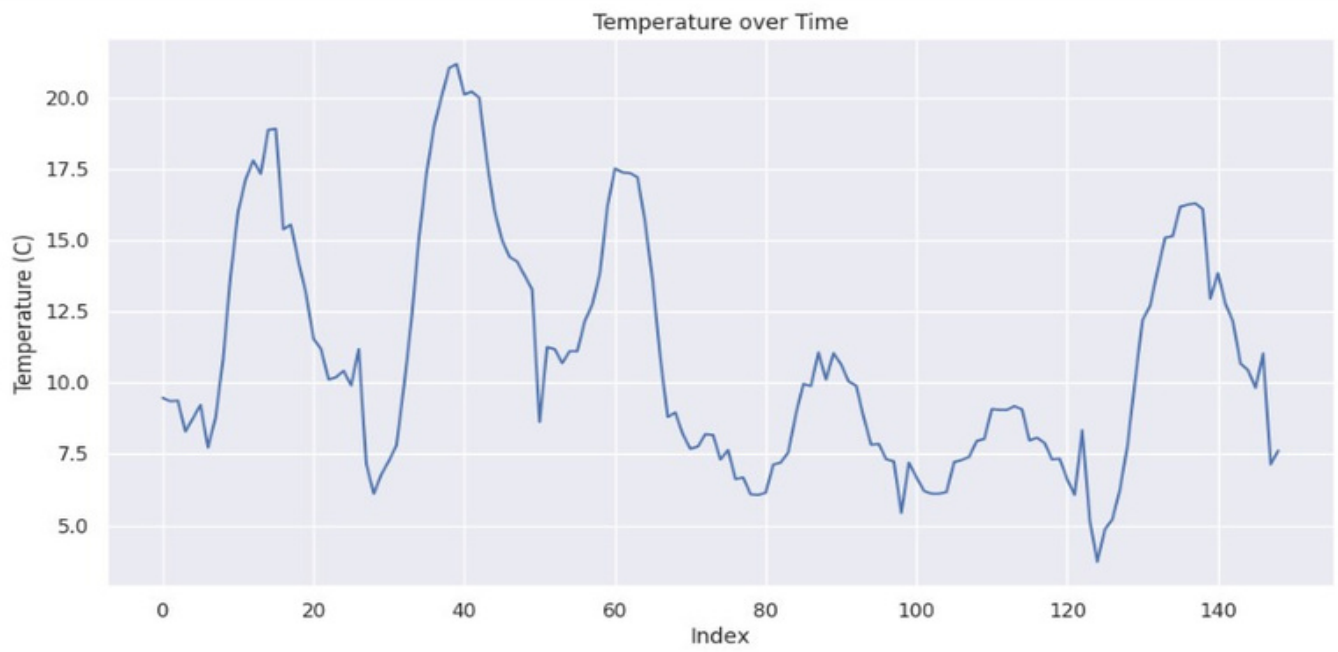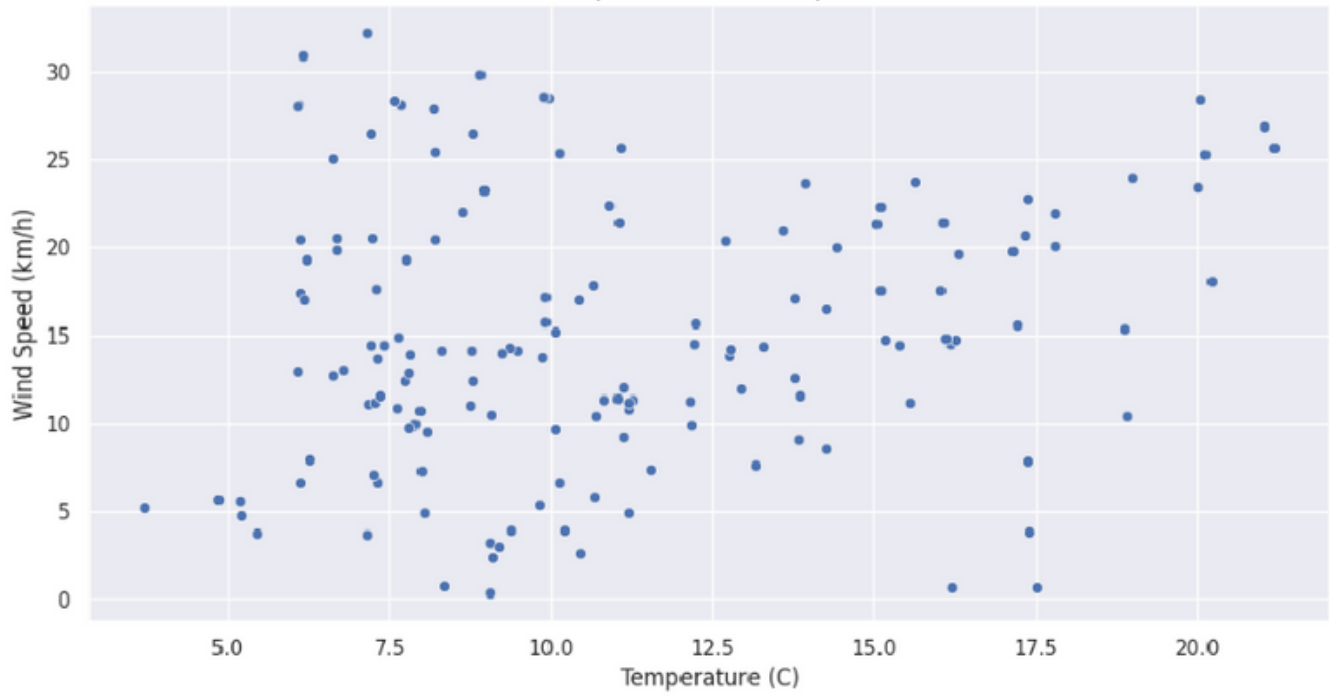
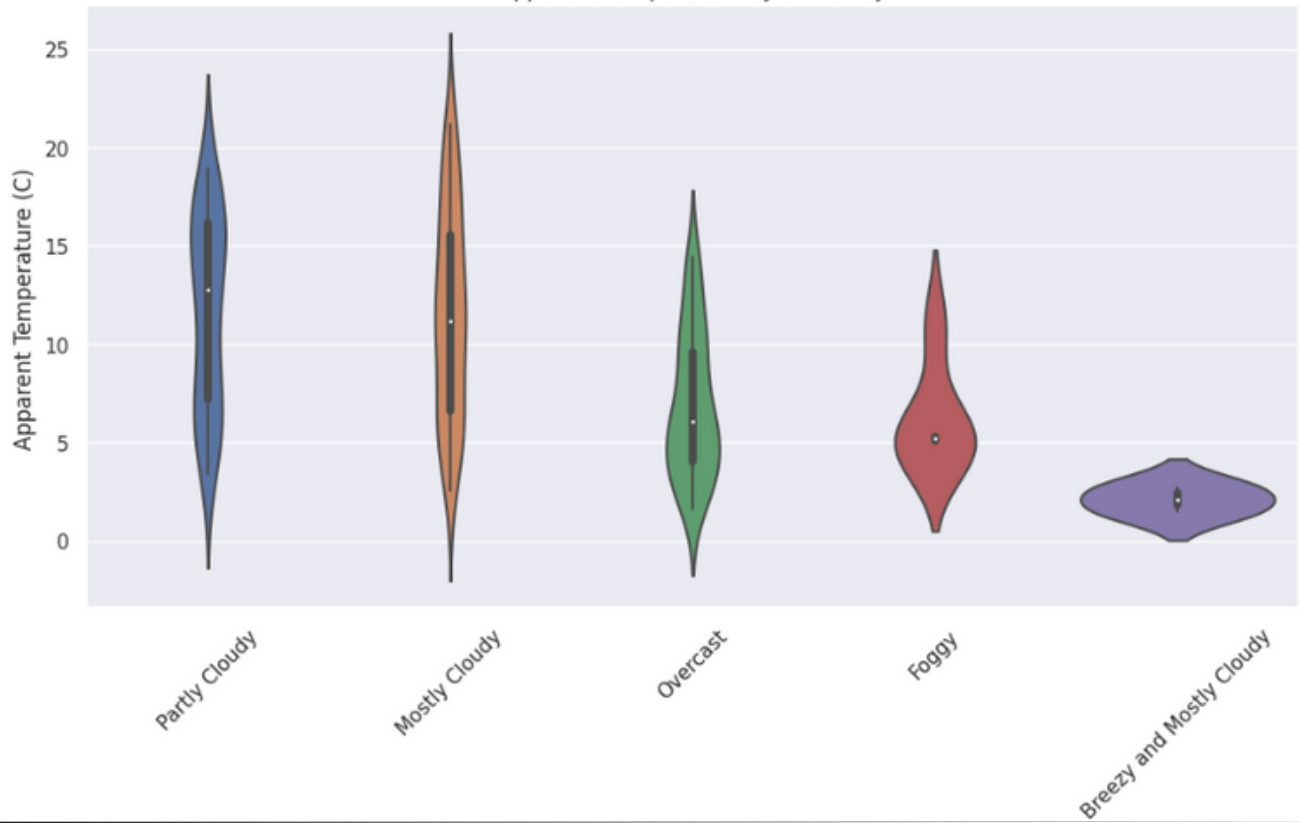

Summary

Summary

Temperature by Precip Type
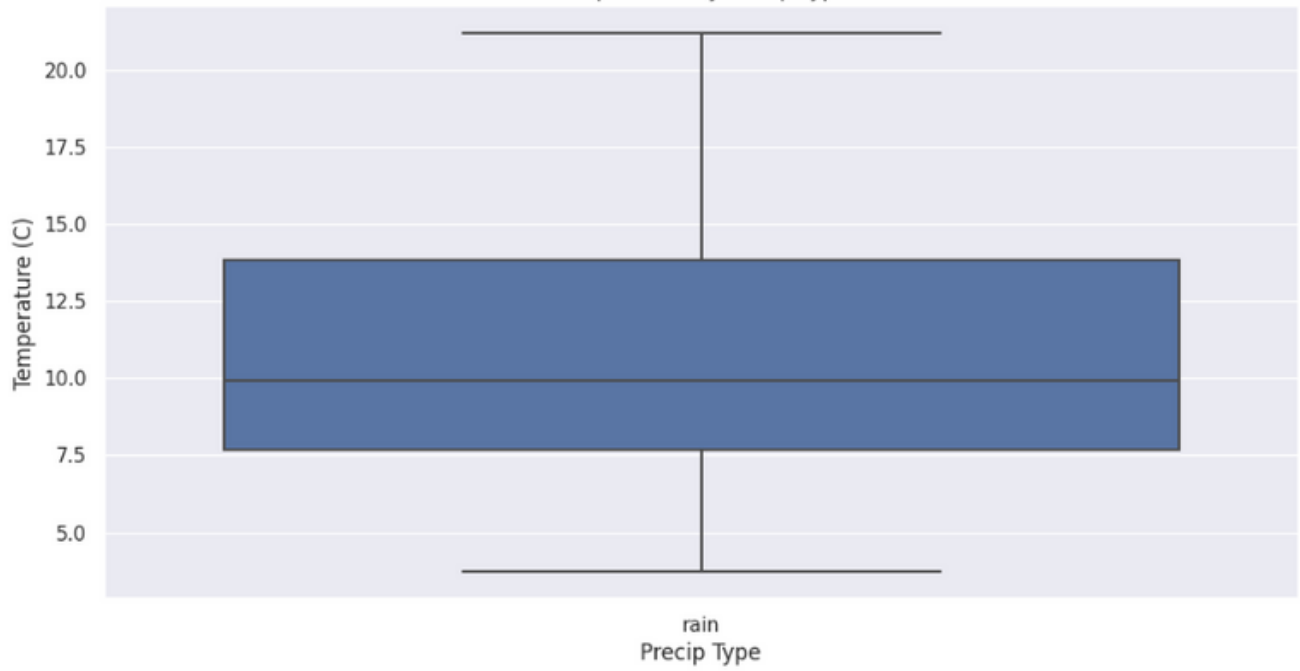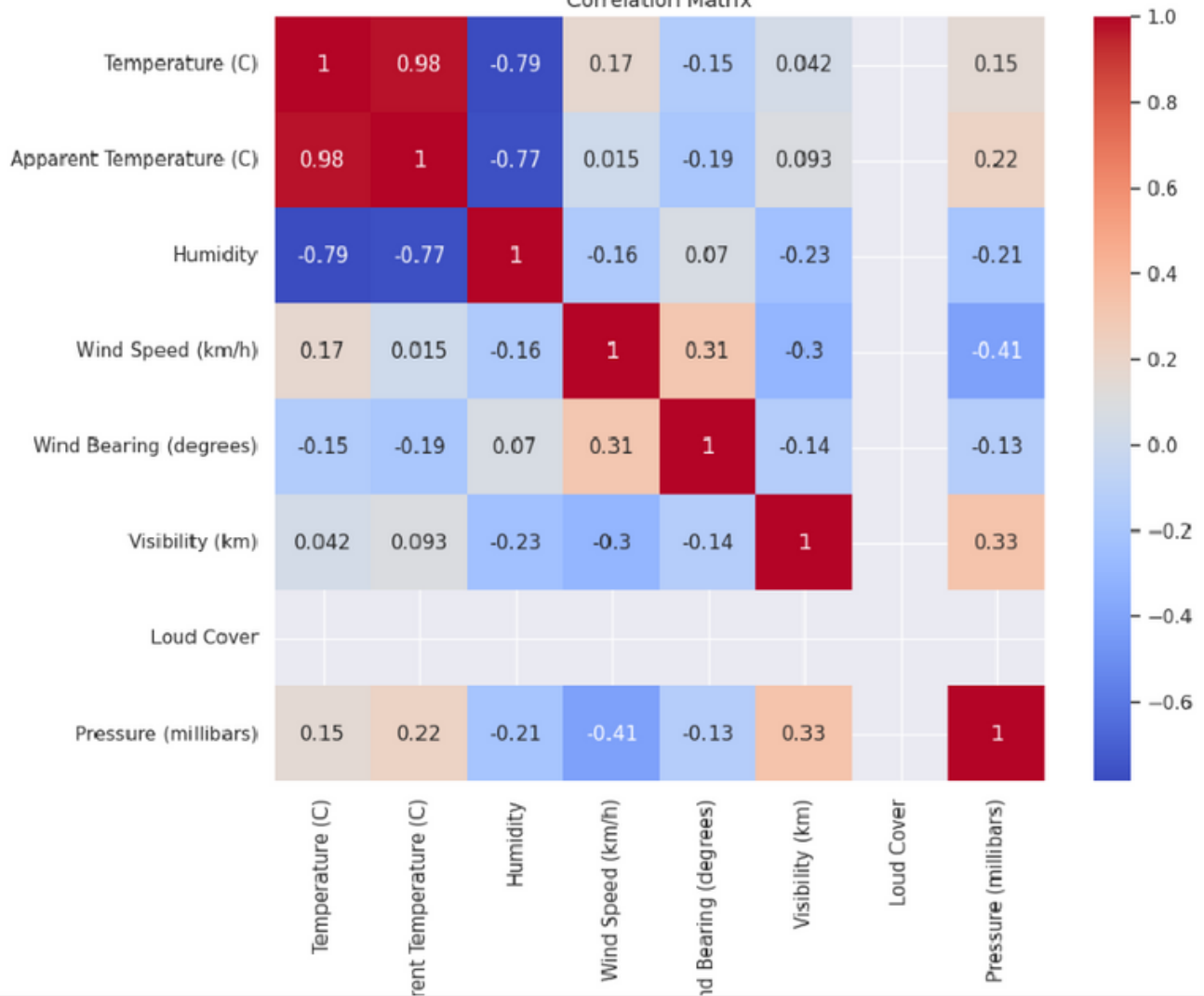

Humidity Distribution

Temperature vs. Wind Speed

Apparent Temperature by Summary

Average Temperature by Precip Type

Correlation Matrix

Pairplot of Select Columns

Precip Type Distribution

rain 100.0%