```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from importlib import reload
plt = reload(plt)

# Load dataset and parse 'Formatted Date' as datetime
df1 = pd.read_csv(r"C:\Users\Jayditya\Downloads\weather.zip",
parse_dates=["Formatted_Date"])

# Display first few rows of the dataset
df1.head()
```

```
          Formatted_Date          Summary Precip_Type  Temperature
(C)  \
0  2006-04-01 00:00:00+02:00  Partly Cloudy         rain
9.472222
1  2006-04-01 01:00:00+02:00  Partly Cloudy         rain
9.355556
2  2006-04-01 02:00:00+02:00  Mostly Cloudy         rain
9.377778
3  2006-04-01 03:00:00+02:00  Partly Cloudy         rain
8.288889
4  2006-04-01 04:00:00+02:00  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.

columns_order=["Date-Time","Summary","Precip Type","Temperature
(C)","Apparent Temperature (C)",
              "Humidity","Wind Speed (km/h)","Wind Bearing
(degrees)","Visibility (km)","Loud Cover",
              "Pressure (millibars)", "Daily Summary"]

df2=df1.reindex(columns=columns_order)
df3=df2.drop(columns=["Loud Cover"])
df3.head()
```

```
                  Date_Time        Summary Precip_Type
Temperature_(C)  \
0  2006-04-01 00:00:00+02:00  Partly Cloudy        rain
NaN
1  2006-04-01 01:00:00+02:00  Partly Cloudy        rain
NaN
2  2006-04-01 02:00:00+02:00  Mostly Cloudy        rain
NaN
3  2006-04-01 03:00:00+02:00  Partly Cloudy        rain
NaN
4  2006-04-01 04:00:00+02:00  Mostly Cloudy        rain
NaN

   Apparent_Temperature_C  Humidity  Wind_Speed_(km/h)  \
0                7.388889      0.89                NaN
1                7.227778      0.86                NaN
2                9.377778      0.89                NaN
3                5.944444      0.83                NaN
4                6.977778      0.83                NaN

   Wind_Bearing_(degrees)  Visibility_(km)  Pressure_(millibars)  \
0                     NaN              NaN                   NaN
1                     NaN              NaN                   NaN
2                     NaN              NaN                   NaN
3                     NaN              NaN                   NaN
4                     NaN              NaN                   NaN

   Daily_Summary
0            NaN
1            NaN
2            NaN
3            NaN
4            NaN
```

```
df3["Date-Time"]=pd.to_datetime(df3["Date-Time"])
df3.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 96453 entries, 0 to 96452
Data columns (total 11 columns):
 #   Column                   Non-Null Count  Dtype
---  ------                   --------------  -----
 0   Date-Time                0 non-null      datetime64[ns]
 1   Summary                  96453 non-null  object
 2   Precip Type              0 non-null      float64
 3   Temperature (C)          96453 non-null  float64
 4   Apparent Temperature (C) 0 non-null      float64
 5   Humidity                 96453 non-null  float64
 6   Wind Speed (km/h)        96453 non-null  float64
 7   Wind Bearing (degrees)   96453 non-null  int64
 8   Visibility (km)          96453 non-null  float64
 9   Pressure (millibars)     96453 non-null  float64
 10  Daily Summary            96453 non-null  object
dtypes: datetime64[ns](1), float64(7), int64(1), object(2)
memory usage: 8.1+ MB
```

```python
df3["Year"]=pd.DatetimeIndex(df3["Date-Time"]).year
df3["Month"]=pd.DatetimeIndex(df3["Date-Time"]).month_name()
df3["Day"]=pd.DatetimeIndex(df3["Date-Time"]).day
df3.head()
```

```
  Date-Time        Summary Precip Type  Temperature (C)  \
0       NaT  Partly Cloudy         NaN         9.472222
1       NaT  Partly Cloudy         NaN         9.355556
2       NaT  Mostly Cloudy         NaN         9.377778
3       NaT  Partly Cloudy         NaN         8.288889
4       NaT  Mostly Cloudy         NaN         8.755556

   Apparent Temperature (C)  Humidity  Wind Speed (km/h)  \
0                       NaN      0.89            14.1197
1                       NaN      0.86            14.2646
2                       NaN      0.89             3.9284
3                       NaN      0.83            14.1036
4                       NaN      0.83            11.0446

   Wind Bearing (degrees)  Visibility (km)  Pressure (millibars)  \
0                     251          15.8263               1015.13
1                     259          15.8263               1015.63
2                     204          14.9569               1015.94
3                     269          15.8263               1016.41
4                     259          15.8263               1016.51

                         Daily Summary  Year Month  Day
0  Partly cloudy throughout the day.    NaN   NaN  NaN
1  Partly cloudy throughout the day.    NaN   NaN  NaN
2  Partly cloudy throughout the day.    NaN   NaN  NaN
```

```
3  Partly cloudy throughout the day.    NaN    NaN   NaN
4  Partly cloudy throughout the day.    NaN    NaN   NaN
```