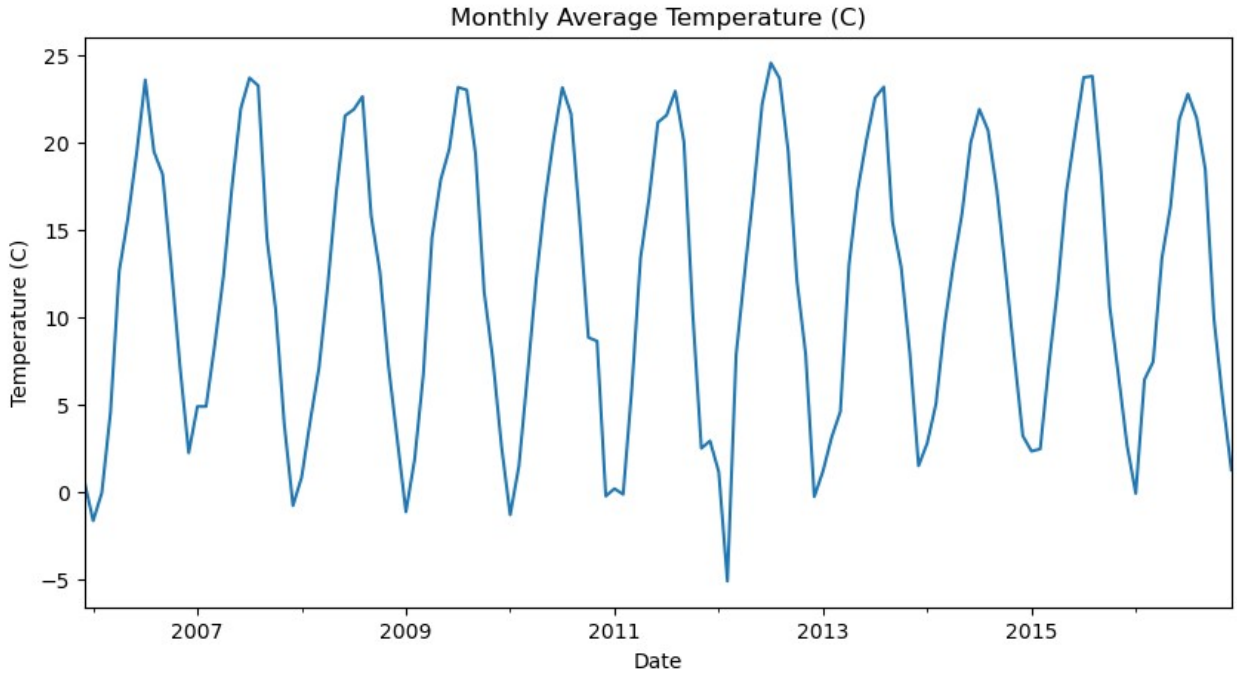
```python
# Anwar_Siraj data science and AI Batch V
# Question 3 (20 marks)

# A dataset containing information about the weather in a city is
available at weatherHistory.csv. Analyze the dataset and identify the
trends in temperature, precipitation, and humidity over time.
Visualize your findings using appropriate charts and graphs.

import pandas as pd

# Load the dataset
weather_file_path = '/Users/cylustariq/Downloads/Hybrid Exam
Paper/weatherHistory.csv'
weather_data = pd.read_csv(weather_file_path)

# Parse dates and set the index to the date column
weather_data['Formatted Date'] =
pd.to_datetime(weather_data['Formatted Date'], utc=True)
weather_data.set_index('Formatted Date', inplace=True)

# Resample to get monthly averages
monthly_averages = weather_data.resample('M').mean()

# Print the monthly averages for temperature and humidity
print("Monthly Averages of Temperature and Humidity:")
print(monthly_averages[['Temperature (C)', 'Humidity']])
```

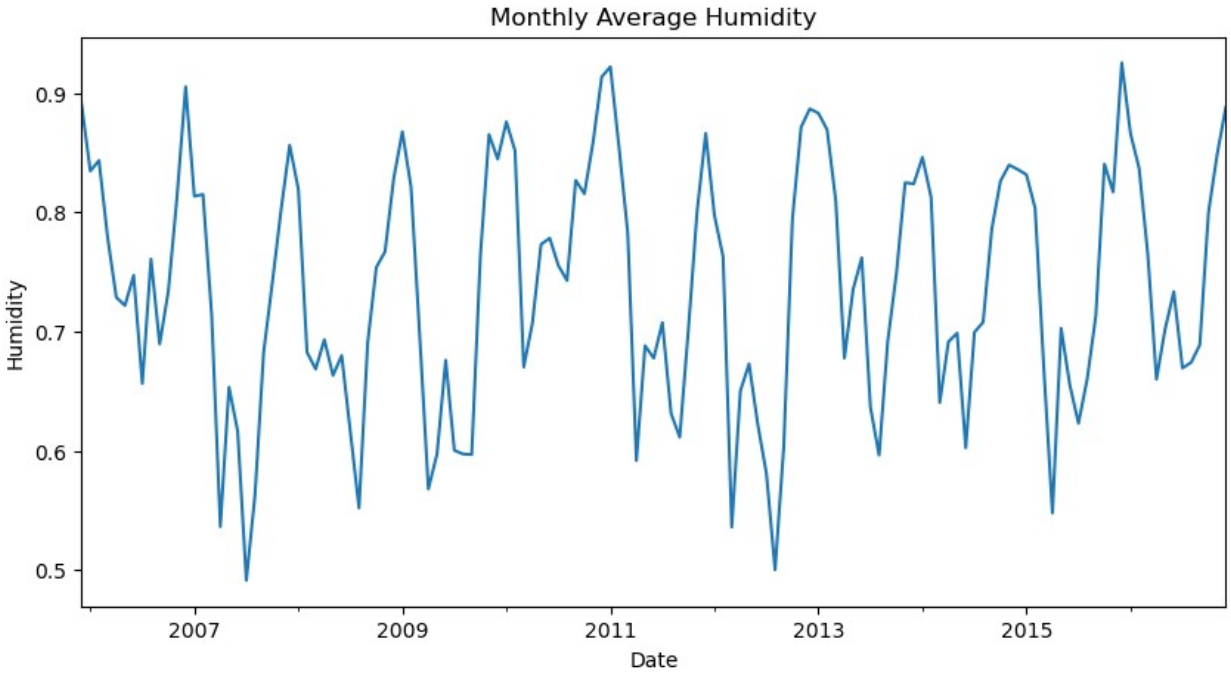
```
Monthly Averages of Temperature and Humidity:
                           Temperature (C)  Humidity
Formatted Date
2005-12-31 00:00:00+00:00         0.577778  0.890000
2006-01-31 00:00:00+00:00        -1.677942  0.834610
2006-02-28 00:00:00+00:00        -0.065394  0.843467
2006-03-31 00:00:00+00:00         4.559274  0.778737
2006-04-30 00:00:00+00:00        12.635031  0.728625
...                                    ...       ...
2016-08-31 00:00:00+00:00        21.420296  0.674046
2016-09-30 00:00:00+00:00        18.467924  0.688833
2016-10-31 00:00:00+00:00         9.893242  0.799906
2016-11-30 00:00:00+00:00         5.282662  0.848472
2016-12-31 00:00:00+00:00         1.239158  0.887981

[133 rows x 2 columns]
```

```python
import matplotlib.pyplot as plt

# Resample to get monthly averages
monthly_data = weather_data.resample('M').mean()

# Plot the monthly average temperature
```

```
plt.figure(figsize=(10, 5))
monthly_data['Temperature (C)'].plot(title='Monthly Average
Temperature (C)')
plt.xlabel('Date')
plt.ylabel('Temperature (C)')
plt.show()
```




```
# Resample to get monthly averages
monthly_data = weather_data.resample('M').mean()

# Plot the monthly average humidity
plt.figure(figsize=(10, 5))
monthly_data['Humidity'].plot(title='Monthly Average Humidity')
plt.xlabel('Date')
plt.ylabel('Humidity')
plt.show()
```

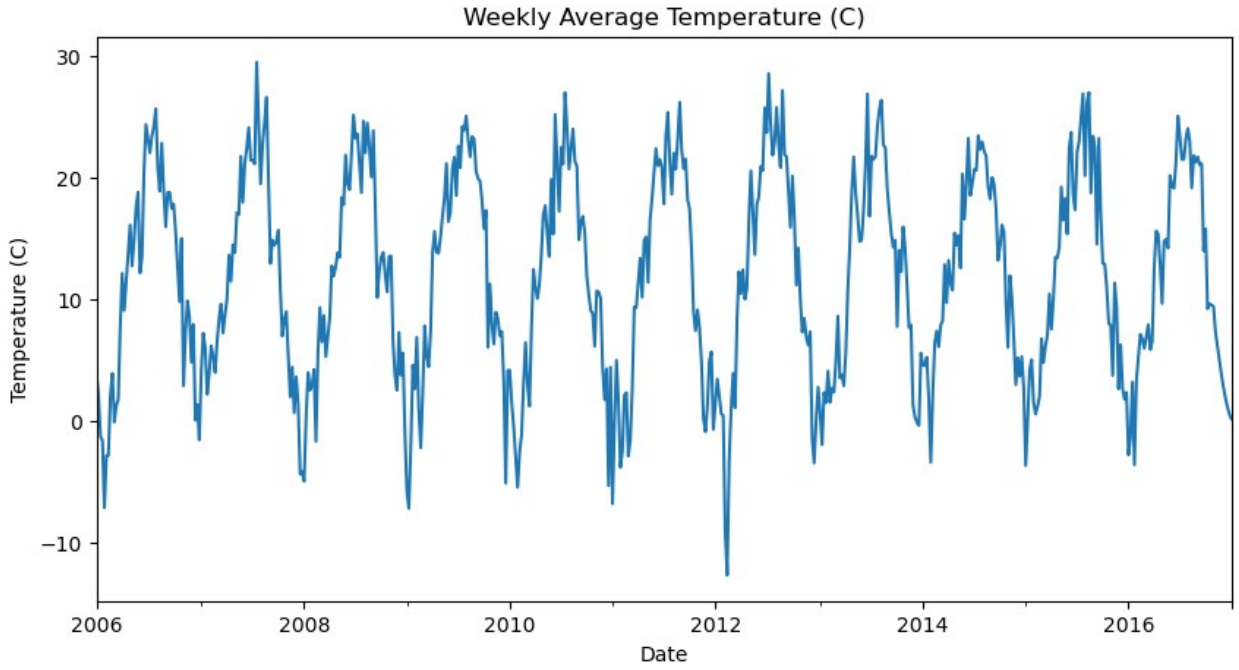



```
# Resample to get weekly averages
weekly_averages = weather_data.resample('W').mean()

# Print the weekly averages for temperature and humidity
print("Weekly Averages of Temperature and Humidity:")
print(weekly_averages[['Temperature (C)', 'Humidity']])
```

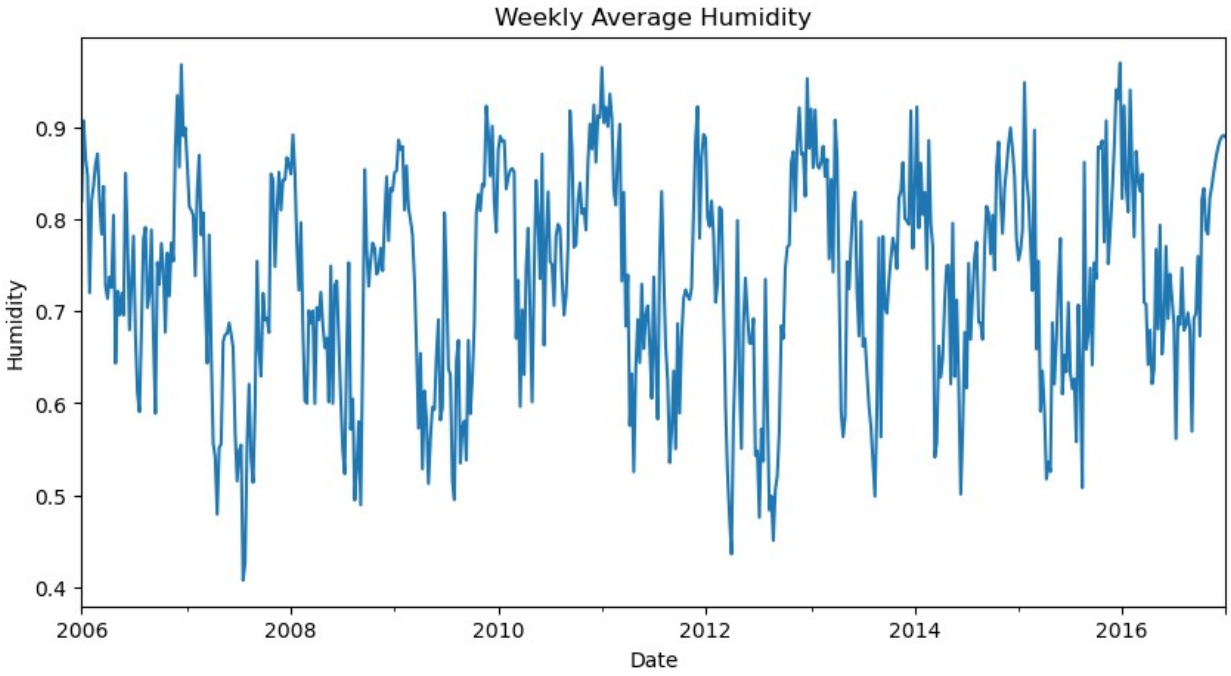
```
Weekly Averages of Temperature and Humidity:
                           Temperature (C)  Humidity
Formatted Date
2006-01-01 00:00:00+00:00         3.935111  0.820000
2006-01-08 00:00:00+00:00         2.107903  0.907500
2006-01-15 00:00:00+00:00        -1.226819  0.866429
2006-01-22 00:00:00+00:00        -1.579563  0.847024
2006-01-29 00:00:00+00:00        -7.093221  0.720119
...                                    ...       ...
2016-12-04 00:00:00+00:00         2.842063  0.877560
2016-12-11 00:00:00+00:00         1.961442  0.885714
2016-12-18 00:00:00+00:00         1.216634  0.890476
2016-12-25 00:00:00+00:00         0.618056  0.891488
2017-01-01 00:00:00+00:00         0.206915  0.889231

[575 rows x 2 columns]
```

```
# Resample to get weekly averages
```

```
weekly_data = weather_data.resample('W').mean()

# Plot the weekly average temperature
plt.figure(figsize=(10, 5))
weekly_data['Temperature (C)'].plot(title='Weekly Average Temperature
(C)')
plt.xlabel('Date')
plt.ylabel('Temperature (C)')
plt.show()
```



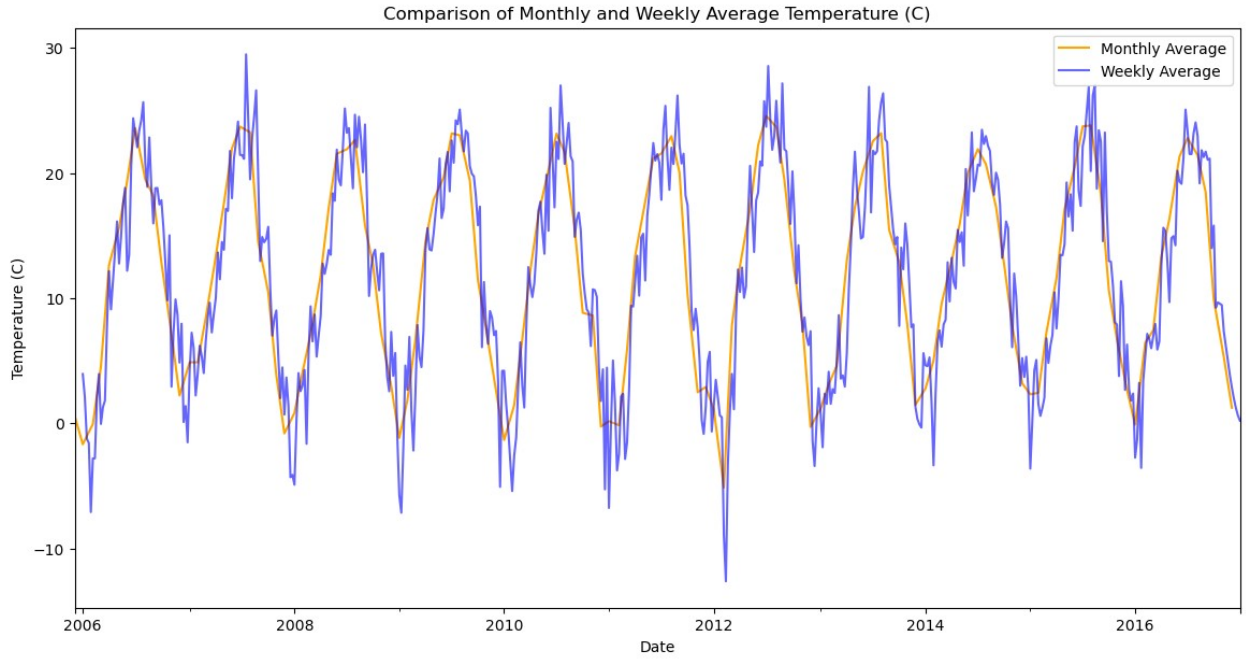

```
# Resample to get weekly averages
weekly_data = weather_data.resample('W').mean()

# Plot the weekly average humidity
plt.figure(figsize=(10, 5))
weekly_data['Humidity'].plot(title='Weekly Average Humidity')
plt.xlabel('Date')
plt.ylabel('Humidity')
plt.show()
```




```
# Resample to get monthly and weekly averages
monthly_averages = weather_data.resample('M').mean()
weekly_averages = weather_data.resample('W').mean()

# Print the monthly and weekly averages for temperature and humidity
print("Monthly Averages of Temperature and Humidity:")
print(monthly_averages[['Temperature (C)', 'Humidity']].head())
print("\nWeekly Averages of Temperature and Humidity:")
print(weekly_averages[['Temperature (C)', 'Humidity']].head())
```

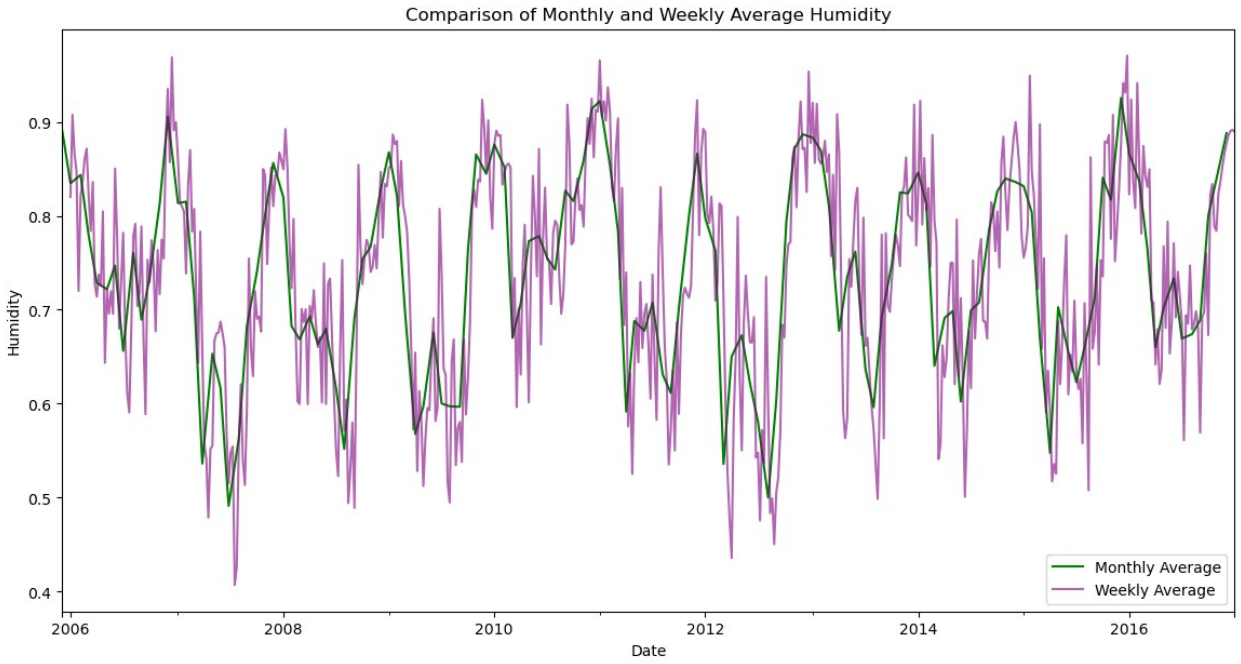
```
Monthly Averages of Temperature and Humidity:
                           Temperature (C)  Humidity
Formatted Date
2005-12-31 00:00:00+00:00         0.577778  0.890000
2006-01-31 00:00:00+00:00        -1.677942  0.834610
2006-02-28 00:00:00+00:00        -0.065394  0.843467
2006-03-31 00:00:00+00:00         4.559274  0.778737
2006-04-30 00:00:00+00:00        12.635031  0.728625

Weekly Averages of Temperature and Humidity:
                           Temperature (C)  Humidity
Formatted Date
2006-01-01 00:00:00+00:00         3.935111  0.820000
2006-01-08 00:00:00+00:00         2.107903  0.907500
2006-01-15 00:00:00+00:00        -1.226819  0.866429
2006-01-22 00:00:00+00:00        -1.579563  0.847024
2006-01-29 00:00:00+00:00        -7.093221  0.720119
```

```
# Resample to get monthly and weekly averages
monthly_data = weather_data.resample('M').mean()
weekly_data = weather_data.resample('W').mean()

# Plot the comparison of monthly and weekly average temperature
plt.figure(figsize=(14, 7))
monthly_data['Temperature (C)'].plot(label='Monthly Average',
color='orange')
weekly_data['Temperature (C)'].plot(label='Weekly Average',
color='blue', alpha=0.6)
plt.title('Comparison of Monthly and Weekly Average Temperature (C)')
plt.xlabel('Date')
plt.ylabel('Temperature (C)')
plt.legend()
plt.show()
```




```
# Resample to get monthly and weekly averages
monthly_data = weather_data.resample('M').mean()
weekly_data = weather_data.resample('W').mean()

# Plot the comparison of monthly and weekly average humidity
plt.figure(figsize=(14, 7))
monthly_data['Humidity'].plot(label='Monthly Average', color='green')
weekly_data['Humidity'].plot(label='Weekly Average', color='purple',
alpha=0.6)
plt.title('Comparison of Monthly and Weekly Average Humidity')
```

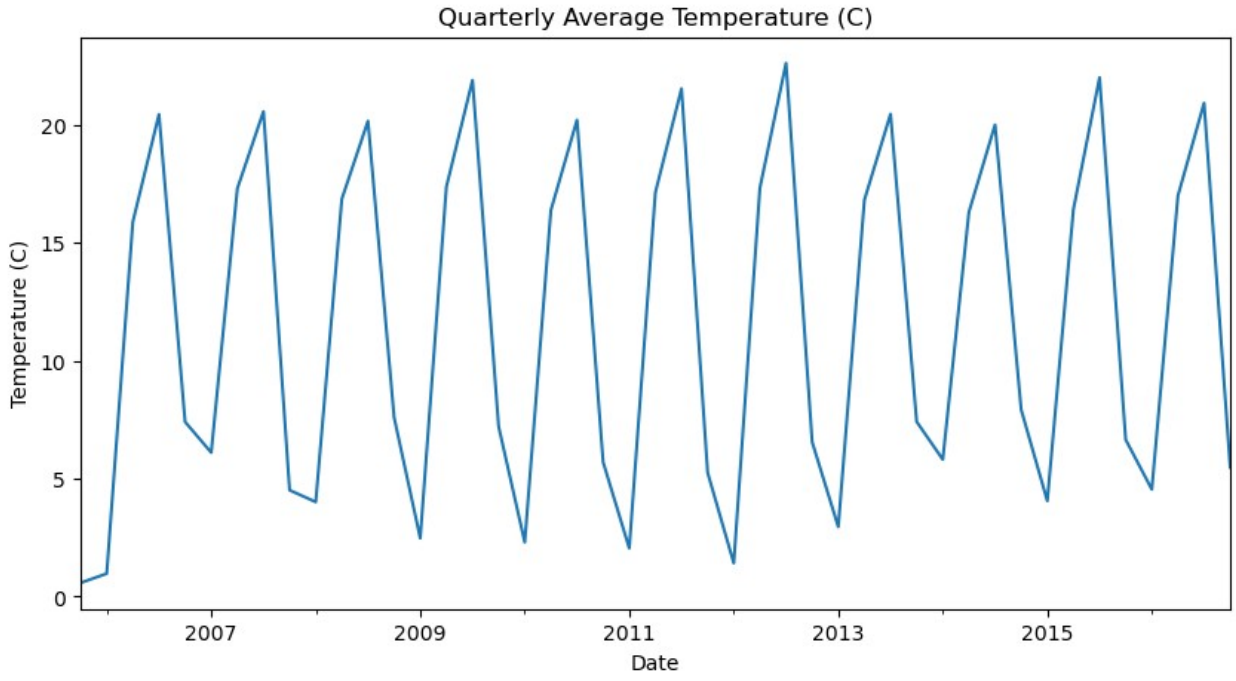
```
plt.xlabel('Date')
plt.ylabel('Humidity')
plt.legend()
plt.show()
```



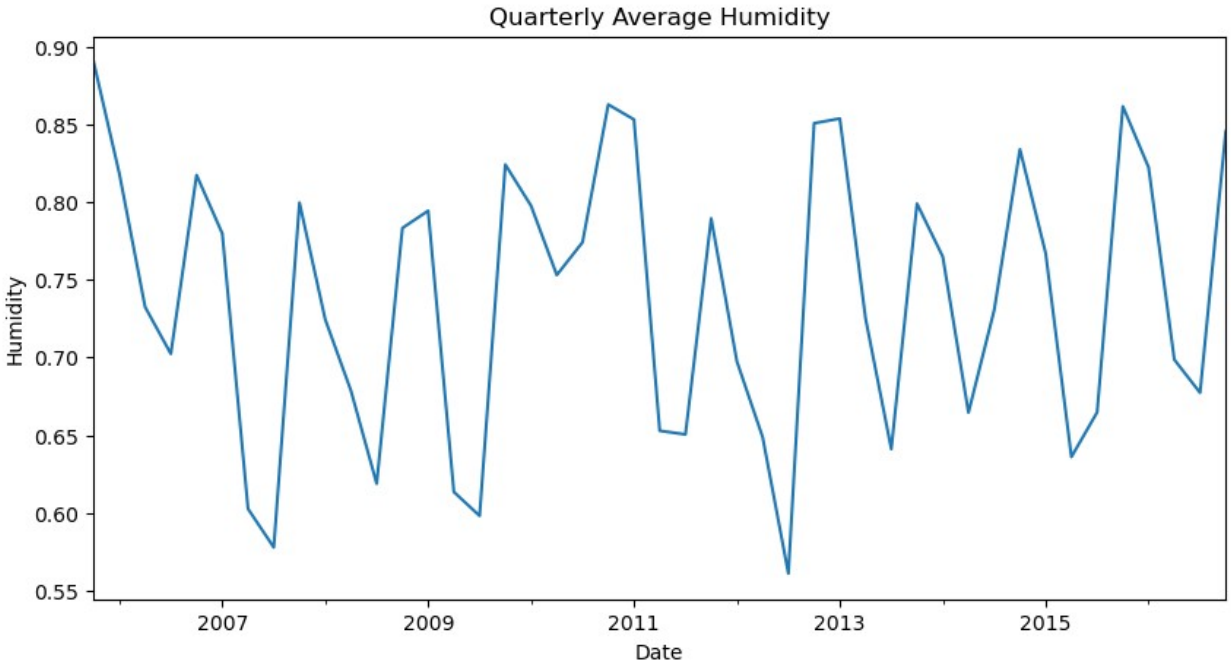

```
# Resample to get quarterly averages
quarterly_averages = weather_data.resample('Q').mean()

# Print the quarterly averages for temperature and humidity
print("Quarterly Averages of Temperature and Humidity:")
print(quarterly_averages[['Temperature (C)', 'Humidity']])
```

```
Quarterly Averages of Temperature and Humidity:
                           Temperature (C)  Humidity
Formatted Date
2005-12-31 00:00:00+00:00         0.577778  0.890000
2006-03-31 00:00:00+00:00         0.972114  0.818120
2006-06-30 00:00:00+00:00        15.877607  0.732399
2006-09-30 00:00:00+00:00        20.436099  0.702305
2006-12-31 00:00:00+00:00         7.403080  0.817296
2007-03-31 00:00:00+00:00         6.102536  0.779657
2007-06-30 00:00:00+00:00        17.293216  0.602596
2007-09-30 00:00:00+00:00        20.559219  0.577817
2007-12-31 00:00:00+00:00         4.508673  0.799529
2008-03-31 00:00:00+00:00         4.007489  0.724418
2008-06-30 00:00:00+00:00        16.853643  0.678558
2008-09-30 00:00:00+00:00        20.154267  0.618773
```

```
2008-12-31 00:00:00+00:00          7.594241  0.783211
2009-03-31 00:00:00+00:00          2.472063  0.794287
2009-06-30 00:00:00+00:00         17.359076  0.613466
2009-09-30 00:00:00+00:00         21.883371  0.598084
2009-12-31 00:00:00+00:00          7.206547  0.824022
2010-03-31 00:00:00+00:00          2.300051  0.797505
2010-06-30 00:00:00+00:00         16.370574  0.752995
2010-09-30 00:00:00+00:00         20.198693  0.774068
2010-12-31 00:00:00+00:00          5.696296  0.862663
2011-03-31 00:00:00+00:00          2.046263  0.853009
2011-06-30 00:00:00+00:00         17.128800  0.652816
2011-09-30 00:00:00+00:00         21.527582  0.650466
2011-12-31 00:00:00+00:00          5.251945  0.789380
2012-03-31 00:00:00+00:00          1.413876  0.697386
2012-06-30 00:00:00+00:00         17.318063  0.648732
2012-09-30 00:00:00+00:00         22.607641  0.560965
2012-12-31 00:00:00+00:00          6.550740  0.850611
2013-03-31 00:00:00+00:00          2.952194  0.853625
2013-06-30 00:00:00+00:00         16.809529  0.725055
2013-09-30 00:00:00+00:00         20.450777  0.641064
2013-12-31 00:00:00+00:00          7.409076  0.798890
2014-03-31 00:00:00+00:00          5.808146  0.764829
2014-06-30 00:00:00+00:00         16.295905  0.664515
2014-09-30 00:00:00+00:00         19.993976  0.730494
2014-12-31 00:00:00+00:00          7.919892  0.833863
2015-03-31 00:00:00+00:00          4.052341  0.766972
2015-06-30 00:00:00+00:00         16.415766  0.635980
2015-09-30 00:00:00+00:00         21.996397  0.664715
2015-12-31 00:00:00+00:00          6.648895  0.861454
2016-03-31 00:00:00+00:00          4.540601  0.822248
2016-06-30 00:00:00+00:00         16.985882  0.698571
2016-09-30 00:00:00+00:00         20.917819  0.677278
2016-12-31 00:00:00+00:00          5.475661  0.845401


# Resample to get quarterly averages
quarterly_data = weather_data.resample('Q').mean()

# Plot the quarterly average temperature
plt.figure(figsize=(10, 5))
quarterly_data['Temperature (C)'].plot(title='Quarterly Average
Temperature (C)')
plt.xlabel('Date')
plt.ylabel('Temperature (C)')
plt.show()
```




```
# Resample to get quarterly averages
quarterly_data = weather_data.resample('Q').mean()

# Plot the quarterly average humidity
plt.figure(figsize=(10, 5))
quarterly_data['Humidity'].plot(title='Quarterly Average Humidity')
plt.xlabel('Date')
plt.ylabel('Humidity')
plt.show()
```




```
import seaborn as sns

# Resample to get monthly, weekly, and quarterly averages
monthly_data = weather_data.resample('M').mean()
weekly_data = weather_data.resample('W').mean()
quarterly_data = weather_data.resample('Q').mean()

# Set up the plotting layout
fig, axes = plt.subplots(nrows=3, ncols=1, figsize=(15, 15),
sharex=True)

# Define titles for subplots
titles = ['Monthly Trends', 'Weekly Trends', 'Quarterly Trends']

# Define data for each subplot
time_data = [monthly_data, weekly_data, quarterly_data]

# Plotting the trends in each subplot
for i, ax in enumerate(axes.flatten()):
    sns.lineplot(ax=ax, data=time_data[i]['Temperature (C)'],
label='Temperature (C)', color='orange')
    sns.lineplot(ax=ax, data=time_data[i]['Humidity'],
label='Humidity', color='blue')
    ax.set_title(titles[i])
    ax.set_ylabel('')
    ax.legend()

# Fine-tune and show the plot
```

```
plt.xlabel('Date')
plt.tight_layout()
plt.show()
```