

Weather Data Analysis System

CEIS110
Faith Burnett
DeVry University

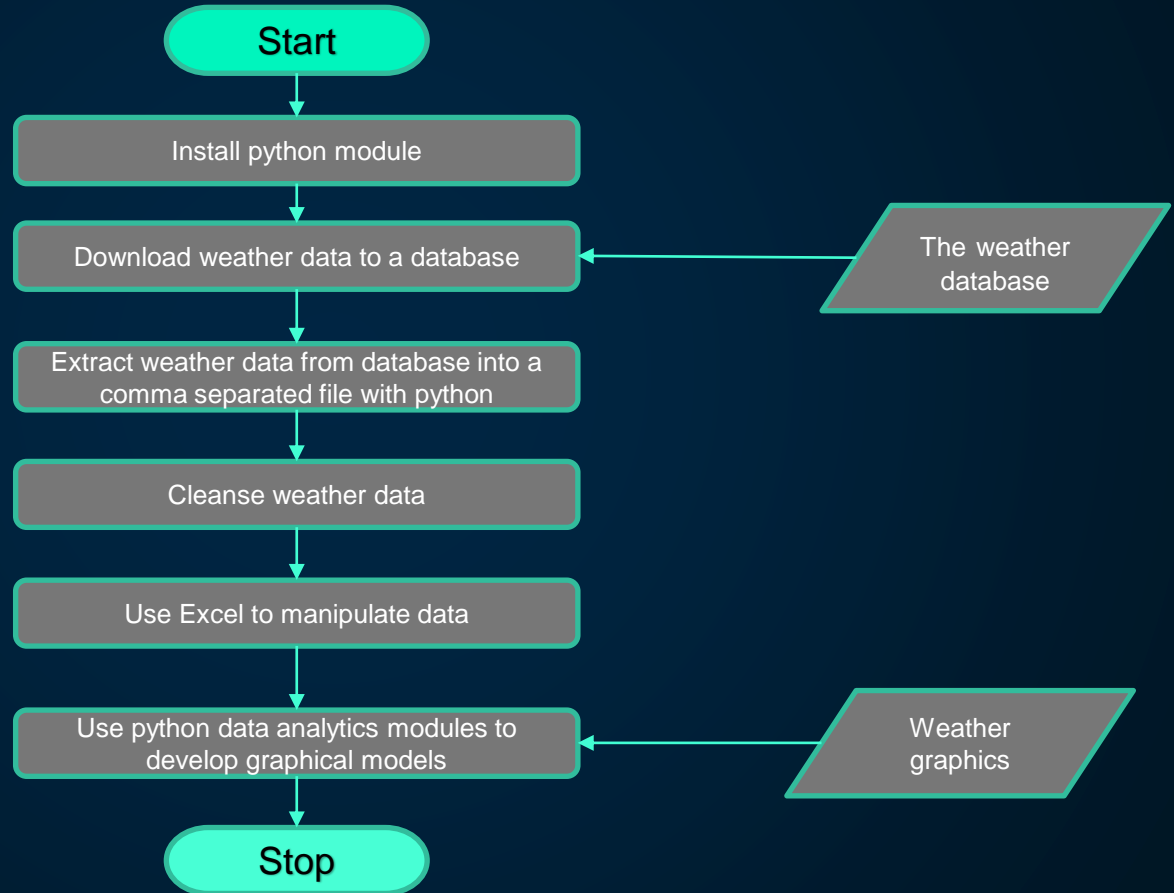
Introduction

- Develop a software system to upload data, insert into database, and process
- Analyze data to make predictions
- Visualize data by using programming to create charts

Design and Library Setup

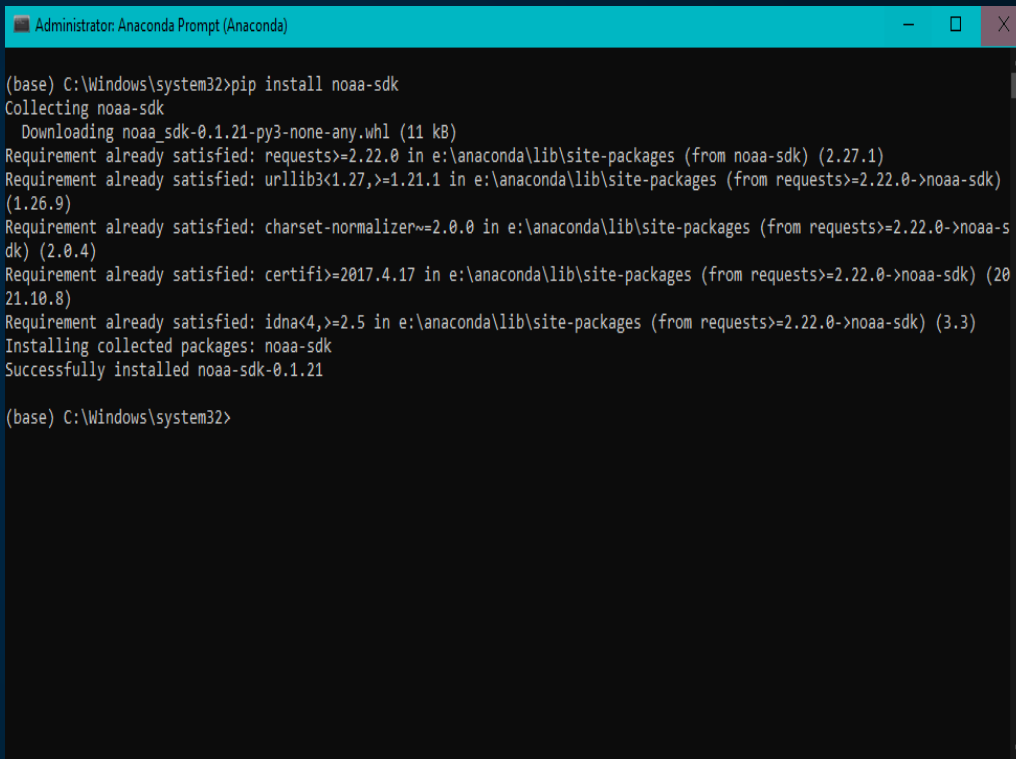
- Create a flowchart showing the input, processing, and output
- Install noaa-sdk

Flowchart



Library install

- Perform pip install noaa-sdk to install library
- Verify noaa-sdk was successfully installed



```
Administrator: Anaconda Prompt (Anaconda)

(base) C:\Windows\system32>pip install noaa-sdk
Collecting noaa-sdk
  Downloading noaa_sdk-0.1.21-py3-none-any.whl (11 kB)
Requirement already satisfied: requests>=2.22.0 in e:\anaconda\lib\site-packages (from noaa-sdk) (2.27.1)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in e:\anaconda\lib\site-packages (from requests>=2.22.0->noaa-sdk) (1.26.9)
Requirement already satisfied: charset-normalizer~=2.0.0 in e:\anaconda\lib\site-packages (from requests>=2.22.0->noaa-sdk) (2.0.4)
Requirement already satisfied: certifi>=2017.4.17 in e:\anaconda\lib\site-packages (from requests>=2.22.0->noaa-sdk) (2021.10.8)
Requirement already satisfied: idna<4,>=2.5 in e:\anaconda\lib\site-packages (from requests>=2.22.0->noaa-sdk) (3.3)
Installing collected packages: noaa-sdk
Successfully installed noaa-sdk-0.1.21

(base) C:\Windows\system32>
```

Creating Weather Database

- Write a program BuildWeatherDb.py that downloads recent weather data and stores it in a database
- Verify that the database was successfully created
- Verify that the weather database is stored in the same location as BuildWeatherDb.py

```

6         """
7     #Purpose: Build weather database from NOAA data
8     #Name: Faith Burnett
9     #Date: 08/11/2022
10    # See https://pypi.org/project/noaa-sdk/ for details on noaa sdk package used
11
12    #Purpose: Build weather database from NOAA data
13    #Name: Faith Burnett
14    #Date: 08/11/2022
15    # See https://pypi.org/project/noaa-sdk/ for details on noaa_sdk package used
16
17    from noaa_sdk import noaa
18    import sqlite3
19    import datetime
20
21    # parameters for retrieving NOAA weather data
22    zipCode = "63366" # change to your postal code
23    country = "US"
24    #date-time format is yyyy-mm-ddThh:mm:ssZ, times are Zulu time (GMT)
25    #gets the most recent 14 days of data
26    today = datetime.datetime.now()
27    past = today - datetime.timedelta(days=14)
28    startDate = past.strftime("%Y-%m-%dT00:00:00Z")
29    endDate = today.strftime("%Y-%m-%dT23:59:59Z")
30
31    #create connection - this creates database if not exist
32    print("Preparing database...")
33    dbFile = "weather.db"
34    conn = sqlite3.connect(dbFile)
35    #create cursor to execute SQL commands
36    cur = conn.cursor()
37
38    #drop previous version of table if any so we start fresh each time
39    dropTableCmd = "DROP TABLE IF EXISTS observations;"
40    cur.execute(dropTableCmd)
41
42    #create new table to store observations
43    createTableCmd = """ CREATE TABLE IF NOT EXISTS observations (
44        timestamp TEXT NOT NULL PRIMARY KEY,
45        windSpeed REAL,
46        temperature REAL,
47        relativeHumidity REAL,
48        windDirection INTEGER,
49        barometricPressure INTEGER,
50        visibility INTEGER,
51        textDescription TEXT
52    ); """
53    cur.execute(createTableCmd)
54    print("Database prepared")
55
56    # Get hourly weather observations from NOAA Weather Service API
57    print("Getting weather data...")
58    n = noaa.NOAA()
59    observations = n.get_observations(zipCode,country,startDate,endDate)
60
61    #populate table with weather observations
62    print("Inserting rows...")
63    insertCmd = """ INSERT INTO observations
64        (timestamp, windSpeed, temperature, relativeHumidity,
65        windDirection, barometricPressure, visibility, textDescription)
66        VALUES
67        (2, 2, 2, 2, 2, 2, 2, 2) """

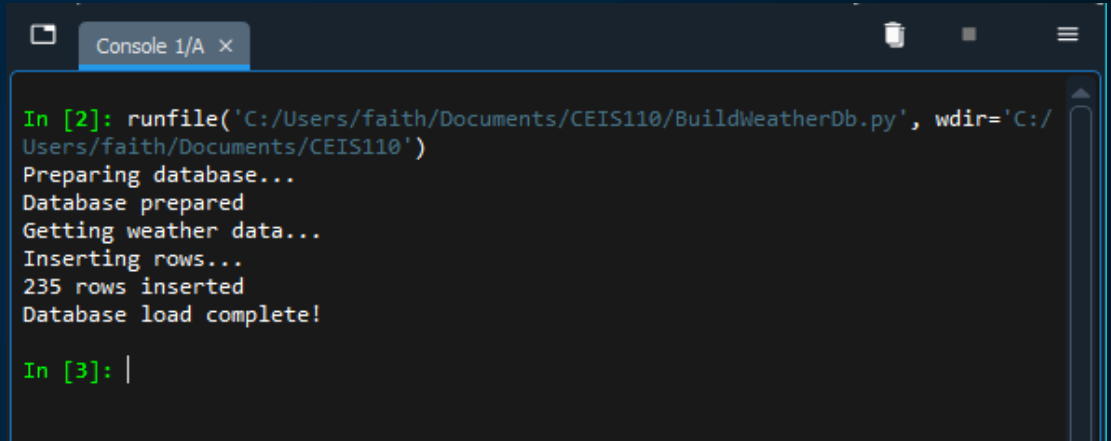
```

Build weather database

- Build database using your own zip code
- Run program to create database

Python console

- Database successfully created

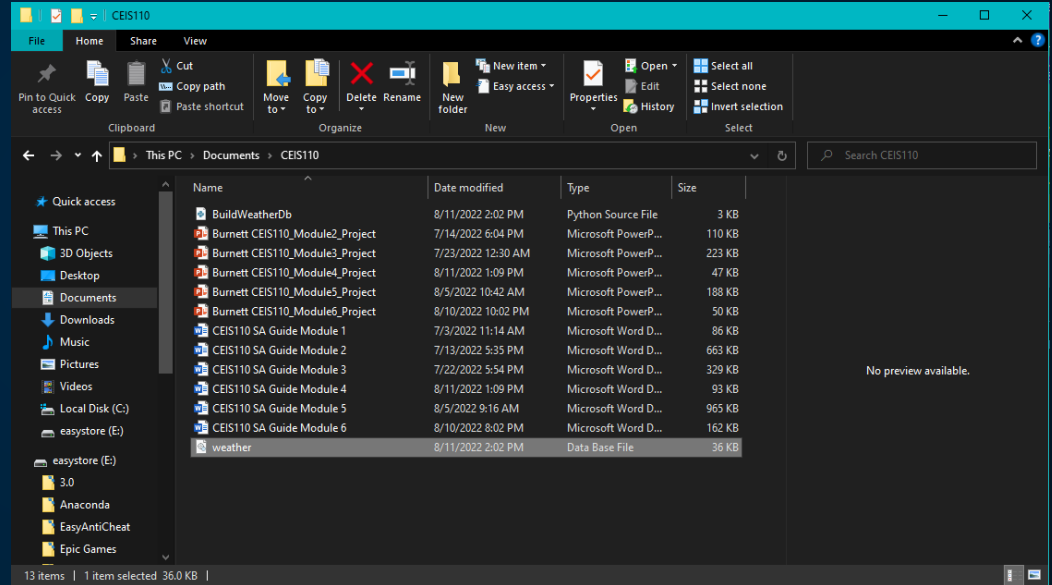


```
Console 1/A x
In [2]: runfile('C:/Users/faith/Documents/CEIS110/BuildWeatherDb.py', wdir='C:/
Users/faith/Documents/CEIS110')
Preparing database...
Database prepared
Getting weather data...
Inserting rows...
235 rows inserted
Database load complete!

In [3]: |
```


Location of Database

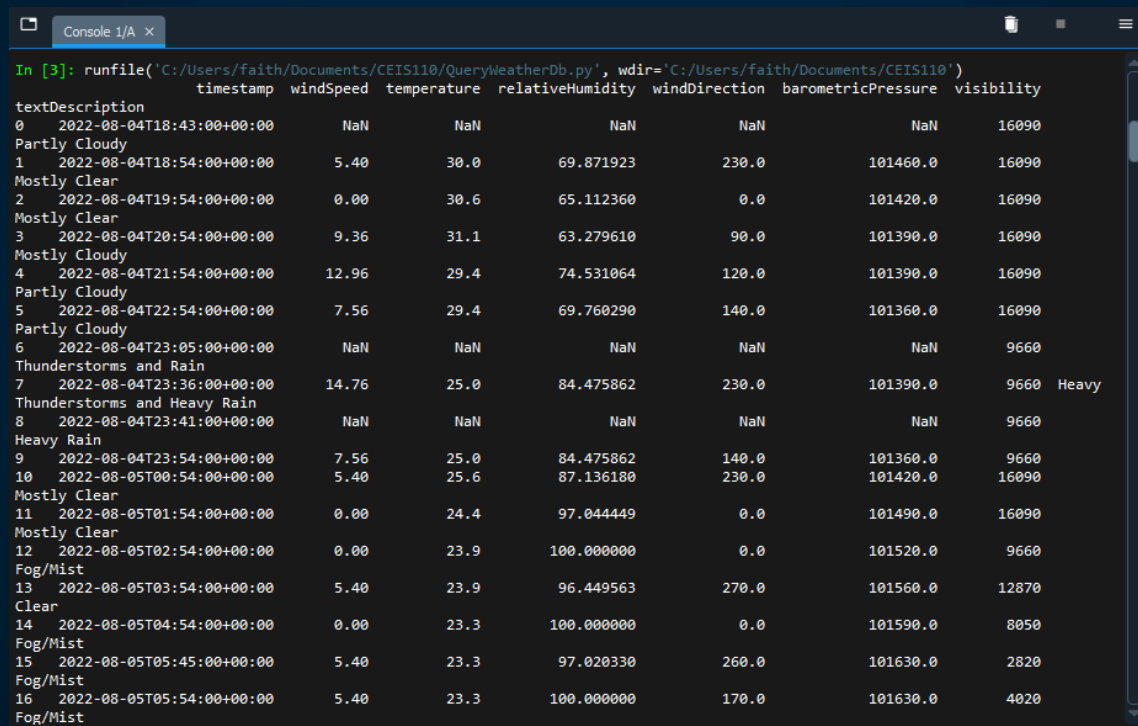
- weather.db in proper location with BuildWeatherDb.py



Querying the database

- Create a program to issue query commands to database
- Display results in console

Query all data

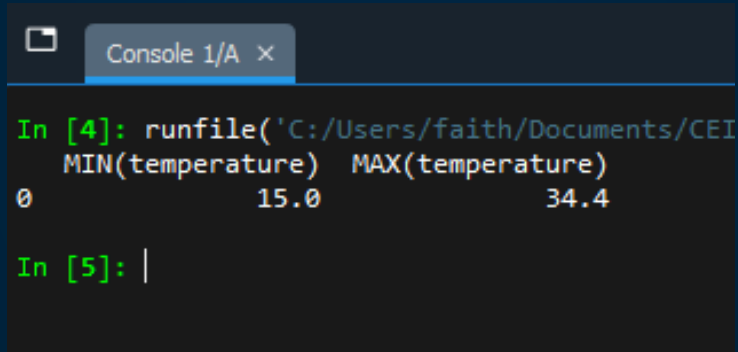


The screenshot shows a Jupyter Notebook console window titled "Console 1/A x". It displays the output of a `runfile` command that reads a CSV file into a pandas DataFrame. The DataFrame contains 17 rows of weather data, with columns for timestamp, wind speed, temperature, relative humidity, wind direction, barometric pressure, and visibility. The data is presented in a table format with alternating light and dark rows. The first row (index 0) has missing values (NaN) for wind speed, temperature, relative humidity, wind direction, and barometric pressure, and a visibility of 16090. The second row (index 1) has a wind speed of 5.40, temperature of 30.0, relative humidity of 69.871923, wind direction of 230.0, barometric pressure of 101460.0, and visibility of 16090. The third row (index 2) has a wind speed of 0.00, temperature of 30.6, relative humidity of 65.112360, wind direction of 0.0, barometric pressure of 101420.0, and visibility of 16090. The fourth row (index 3) has a wind speed of 9.36, temperature of 31.1, relative humidity of 63.279610, wind direction of 90.0, barometric pressure of 101390.0, and visibility of 16090. The fifth row (index 4) has a wind speed of 12.96, temperature of 29.4, relative humidity of 74.531064, wind direction of 120.0, barometric pressure of 101390.0, and visibility of 16090. The sixth row (index 5) has a wind speed of 7.56, temperature of 29.4, relative humidity of 69.760290, wind direction of 140.0, barometric pressure of 101360.0, and visibility of 16090. The seventh row (index 6) has missing values (NaN) for wind speed, temperature, relative humidity, wind direction, and barometric pressure, and a visibility of 9660. The eighth row (index 7) has a wind speed of 14.76, temperature of 25.0, relative humidity of 84.475862, wind direction of 230.0, barometric pressure of 101390.0, and visibility of 9660. The ninth row (index 8) has missing values (NaN) for wind speed, temperature, relative humidity, wind direction, and barometric pressure, and a visibility of 9660. The tenth row (index 9) has a wind speed of 7.56, temperature of 25.0, relative humidity of 84.475862, wind direction of 140.0, barometric pressure of 101360.0, and visibility of 9660. The eleventh row (index 10) has a wind speed of 5.40, temperature of 25.6, relative humidity of 87.136180, wind direction of 230.0, barometric pressure of 101420.0, and visibility of 16090. The twelfth row (index 11) has a wind speed of 0.00, temperature of 24.4, relative humidity of 97.044449, wind direction of 0.0, barometric pressure of 101490.0, and visibility of 16090. The thirteenth row (index 12) has a wind speed of 0.00, temperature of 23.9, relative humidity of 100.000000, wind direction of 0.0, barometric pressure of 101520.0, and visibility of 9660. The fourteenth row (index 13) has a wind speed of 5.40, temperature of 23.9, relative humidity of 96.449563, wind direction of 270.0, barometric pressure of 101560.0, and visibility of 12870. The fifteenth row (index 14) has a wind speed of 0.00, temperature of 23.3, relative humidity of 100.000000, wind direction of 0.0, barometric pressure of 101590.0, and visibility of 8050. The sixteenth row (index 15) has a wind speed of 5.40, temperature of 23.3, relative humidity of 97.020330, wind direction of 260.0, barometric pressure of 101630.0, and visibility of 2820. The seventeenth row (index 16) has a wind speed of 5.40, temperature of 23.3, relative humidity of 100.000000, wind direction of 170.0, barometric pressure of 101630.0, and visibility of 4020.

```
In [3]: runfile('C:/Users/faith/Documents/CEIS110/QueryWeatherDb.py', wdir='C:/Users/faith/Documents/CEIS110')
```

| | timestamp | windSpeed | temperature | relativeHumidity | windDirection | barometricPressure | visibility |
|----|---------------------------|-----------|-------------|------------------|---------------|--------------------|------------|
| 0 | 2022-08-04T18:43:00+00:00 | NaN | NaN | NaN | NaN | NaN | 16090 |
| 1 | 2022-08-04T18:54:00+00:00 | 5.40 | 30.0 | 69.871923 | 230.0 | 101460.0 | 16090 |
| 2 | 2022-08-04T19:54:00+00:00 | 0.00 | 30.6 | 65.112360 | 0.0 | 101420.0 | 16090 |
| 3 | 2022-08-04T20:54:00+00:00 | 9.36 | 31.1 | 63.279610 | 90.0 | 101390.0 | 16090 |
| 4 | 2022-08-04T21:54:00+00:00 | 12.96 | 29.4 | 74.531064 | 120.0 | 101390.0 | 16090 |
| 5 | 2022-08-04T22:54:00+00:00 | 7.56 | 29.4 | 69.760290 | 140.0 | 101360.0 | 16090 |
| 6 | 2022-08-04T23:05:00+00:00 | NaN | NaN | NaN | NaN | NaN | 9660 |
| 7 | 2022-08-04T23:36:00+00:00 | 14.76 | 25.0 | 84.475862 | 230.0 | 101390.0 | 9660 Heavy |
| 8 | 2022-08-04T23:41:00+00:00 | NaN | NaN | NaN | NaN | NaN | 9660 |
| 9 | 2022-08-04T23:54:00+00:00 | 7.56 | 25.0 | 84.475862 | 140.0 | 101360.0 | 9660 |
| 10 | 2022-08-05T00:54:00+00:00 | 5.40 | 25.6 | 87.136180 | 230.0 | 101420.0 | 16090 |
| 11 | 2022-08-05T01:54:00+00:00 | 0.00 | 24.4 | 97.044449 | 0.0 | 101490.0 | 16090 |
| 12 | 2022-08-05T02:54:00+00:00 | 0.00 | 23.9 | 100.000000 | 0.0 | 101520.0 | 9660 |
| 13 | 2022-08-05T03:54:00+00:00 | 5.40 | 23.9 | 96.449563 | 270.0 | 101560.0 | 12870 |
| 14 | 2022-08-05T04:54:00+00:00 | 0.00 | 23.3 | 100.000000 | 0.0 | 101590.0 | 8050 |
| 15 | 2022-08-05T05:45:00+00:00 | 5.40 | 23.3 | 97.020330 | 260.0 | 101630.0 | 2820 |
| 16 | 2022-08-05T05:54:00+00:00 | 5.40 | 23.3 | 100.000000 | 170.0 | 101630.0 | 4020 |

Query Minimum and Maximum Temperatures



The screenshot shows a Jupyter Notebook console window titled "Console 1/A". It displays the execution of a query. The first line shows the command `runfile('C:/Users/faith/Documents/CEI`. The second line shows the output of the query, which is a table with two columns: `MIN(temperature)` and `MAX(temperature)`. The first row of data shows the values `0` and `34.4` respectively. The third line shows the prompt `In [5]: |`.

```
In [4]: runfile('C:/Users/faith/Documents/CEI
MIN(temperature) MAX(temperature)
0                15.0                34.4

In [5]: |
```

Query all clear days

Console 1/A x

```
In [5]: runfile('C:/Users/faith/Documents/CE:
```

| | temperature | windSpeed | textDescription |
|----|-------------|-----------|-----------------|
| 0 | 27.8 | 5.40 | Clear |
| 1 | 27.2 | 14.76 | Clear |
| 2 | 26.1 | 9.36 | Clear |
| 3 | 25.0 | 5.40 | Clear |
| 4 | 21.7 | 0.00 | Clear |
| 5 | 18.9 | 5.40 | Clear |
| 6 | 16.1 | 0.00 | Clear |
| 7 | 15.0 | 0.00 | Clear |
| 8 | 16.7 | 0.00 | Clear |
| 9 | 17.0 | 0.00 | Clear |
| 10 | 17.2 | 0.00 | Clear |
| 11 | 17.8 | 0.00 | Clear |
| 12 | 18.3 | 5.40 | Clear |
| 13 | 18.9 | 0.00 | Clear |
| 14 | 25.0 | 9.36 | Clear |
| 15 | 26.7 | 11.16 | Clear |
| 16 | 27.2 | 12.96 | Clear |
| 17 | 27.2 | 12.96 | Clear |
| 18 | 20.0 | 0.00 | Clear |
| 19 | 20.6 | 0.00 | Clear |
| 20 | 21.7 | 0.00 | Clear |
| 21 | 22.2 | 0.00 | Clear |
| 22 | 24.4 | 9.36 | Clear |
| 23 | 25.0 | 14.76 | Clear |
| 24 | 25.6 | 5.40 | Clear |
| 25 | 26.1 | 0.00 | Clear |
| 26 | 27.2 | 14.76 | Clear |
| 27 | 31.7 | 14.76 | Clear |
| 28 | 26.7 | 12.96 | Clear |
| 29 | 25.6 | 12.96 | Clear |
| 30 | 24.4 | 5.40 | Clear |
| 31 | 24.4 | 7.56 | Clear |
| 32 | 24.4 | 7.56 | Clear |
| 33 | 25.6 | 0.00 | Clear |

Extracting Data

- Create a program to query temperature and humidity
- Convert Celsius to Fahrenheit
- Store Celsius, Fahrenheit, and humidity data in a csv file
- Using the csv file create a graph to analyze data

Extracting temperature and humidity

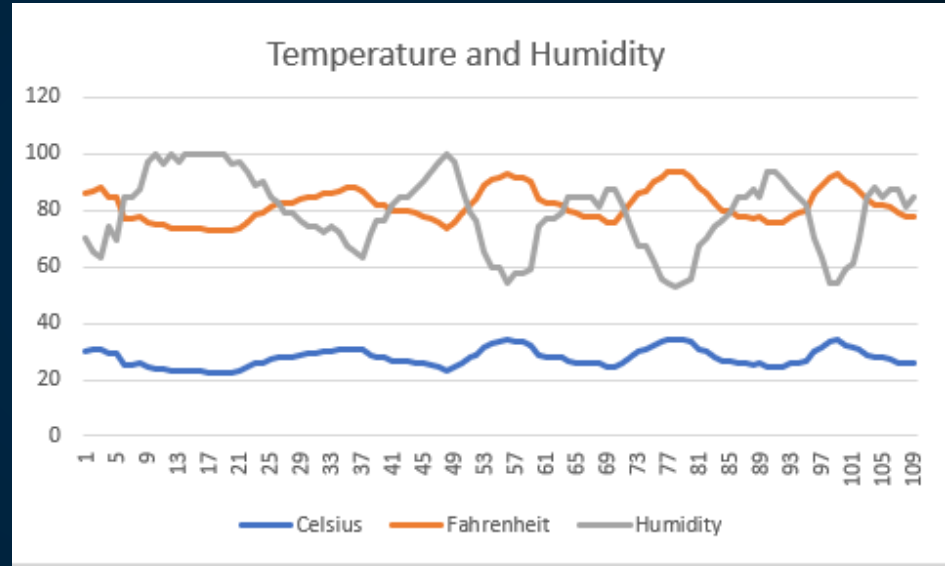
```
C:\Users\faith\Documents\CEIS110\ExtractTempHumidity.py
ExtractTempHumidity.py* X

1  # -*- coding: utf-8 -*-
2  """
3  Created on Thu Aug 11 14:41:54 2022
4
5  @author: faith
6  """
7  #Purpose: Extract temperature, humidity data from weather database into CSV file
8  #Name: Faith Burnett
9  #Date: 08/11/2022
10 # Run BuildWeatherDB.py to build weather database before running this program
11
12 import sqlite3
13
14 #convert Celsius temperature to Fahrenheit
15 def convertCtoF(tempC):
16     return (tempC*9.0/5.0) + 32.0
17
18 #file names for database and output file
19 dbFile = "weather.db"
20 output_file_name='formatdata.csv'
21
22 #connect to and query weather database and
23 dbFile = "weather.db"
24 conn = sqlite3.connect(dbFile)
25 #create cursor to execute SQL commands
26 cur = conn.cursor()
27 selectCmd = """ SELECT temperature, relativeHumidity FROM observations
28                  ORDER BY timestamp; """
29 cur.execute(selectCmd)
30 allRows = cur.fetchall()
31 #limit the number of rows output to half
32 rowCount = len(allRows)//2 # double slash does integer division
33 rows = allRows[:rowCount]
34
35 #write data to output file
36 with open(output_file_name,"w+") as outf:
37     outf.write('Celsius,Fahrenheit,Humidity')
38     outf.write('\n')
39     for row in rows:
40         tempC = row[0]
41         if tempC is None: #handle missing temperature value
42             outf.write(',,')
43         else:
44             tempF = convertCtoF(tempC)
45             outf.write(str(tempC)+',')
46             outf.write(str(tempF)+',')
47             humidity = row[1]
48             if humidity is None: #handle missing humidity value
49                 outf.write('\n')
50             else:
51                 outf.write(str(humidity)+'\n') #print data to file separated by commas
52
53
```

| | A | B | C |
|----|---------|------------|----------|
| 1 | Celsius | Fahrenheit | Humidity |
| 2 | 30 | 86 | 69.87192 |
| 3 | 30.6 | 87.08 | 65.11236 |
| 4 | 31.1 | 87.98 | 63.27961 |
| 5 | 29.4 | 84.92 | 74.53106 |
| 6 | 29.4 | 84.92 | 69.76029 |
| 7 | 25 | 77 | 84.47586 |
| 8 | 25 | 77 | 84.47586 |
| 9 | 25.6 | 78.08 | 87.13618 |
| 10 | 24.4 | 75.92 | 97.04445 |
| 11 | 23.9 | 75.02 | 100 |
| 12 | 23.9 | 75.02 | 96.44956 |
| 13 | 23.3 | 73.94 | 100 |
| 14 | 23.3 | 73.94 | 97.02033 |
| 15 | 23.3 | 73.94 | 100 |
| 16 | 23.3 | 73.94 | 100 |
| 17 | 23.3 | 73.94 | 100 |
| 18 | 22.8 | 73.04 | 100 |
| 19 | 22.8 | 73.04 | 100 |
| 20 | 22.8 | 73.04 | 100 |
| 21 | 22.8 | 73.04 | 96.42062 |
| 22 | 23.3 | 73.94 | 97.02033 |
| 23 | 24.4 | 75.92 | 93.59895 |
| 24 | 26 | 78.8 | 88.75895 |

Retrieve and convert data to csv format

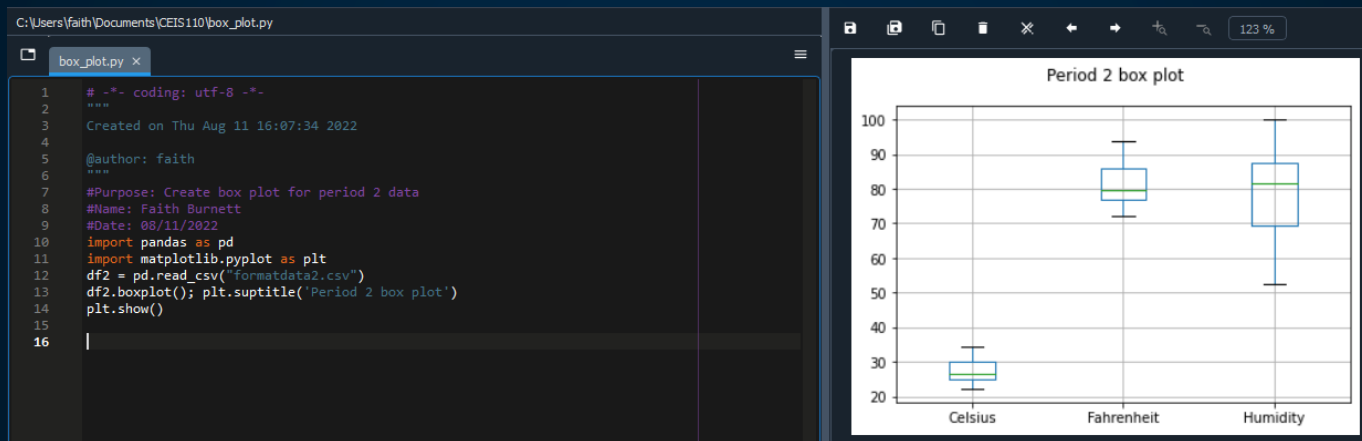
Temperature and humidity chart



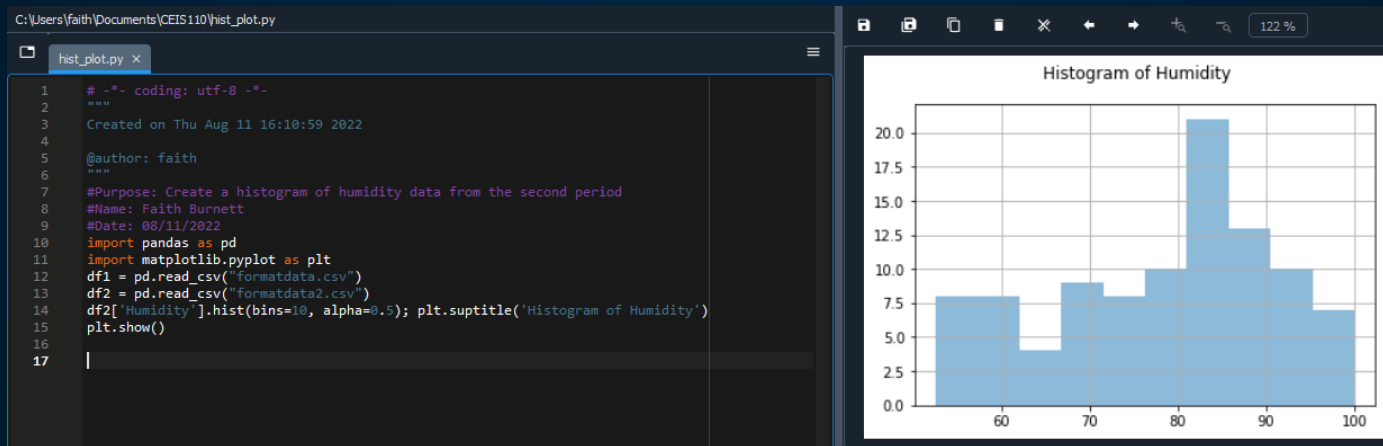
Develop graphical models and interpret results

- Create a box plot, line plot, histogram plot and scatter plot using temperature and humidity data from csv files
- Develop a prediction based off recent data

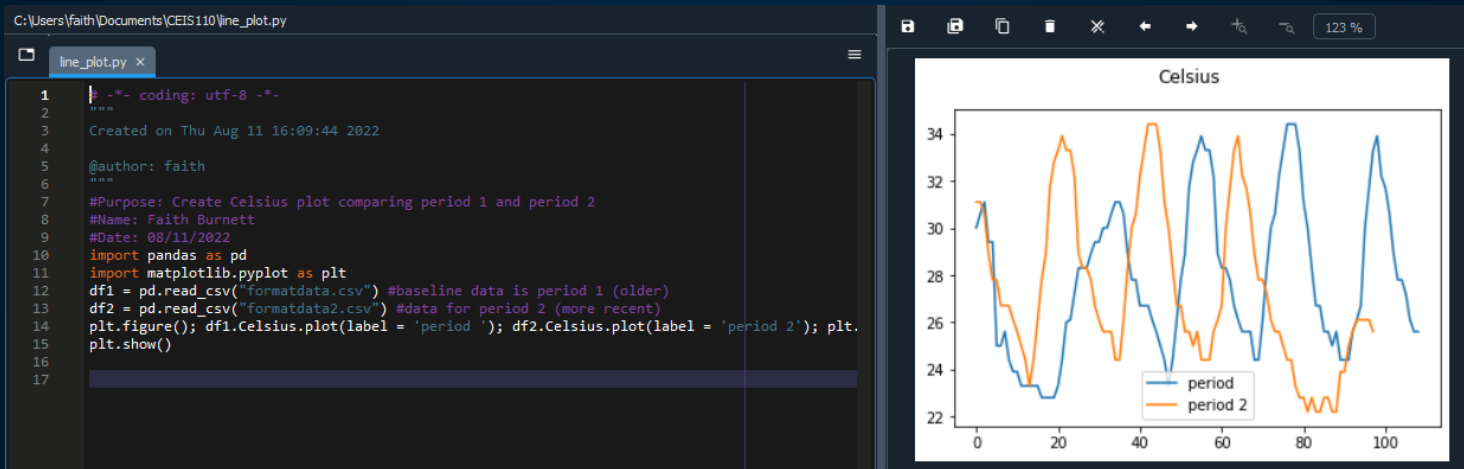
Plot 1



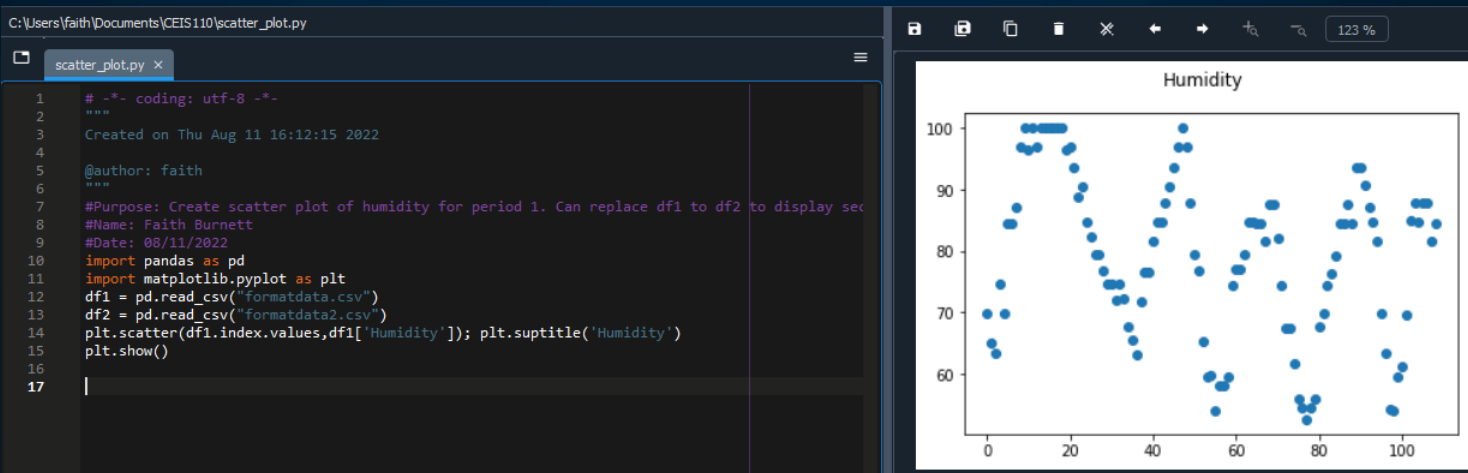
Plot 2



Plot 3

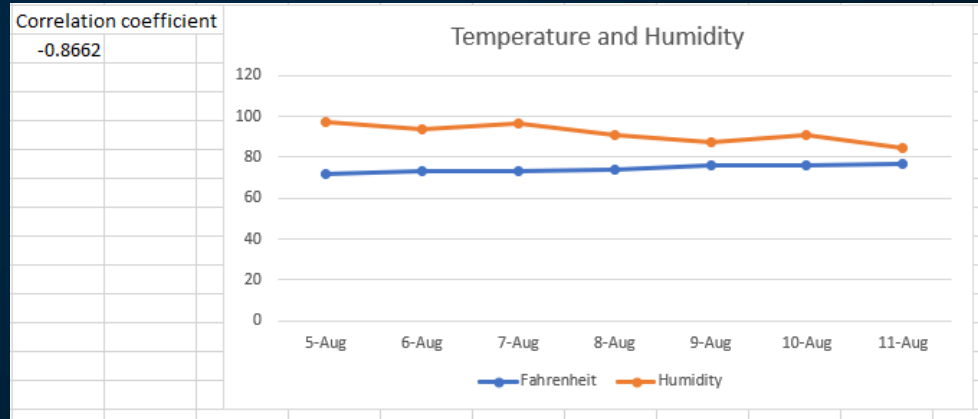
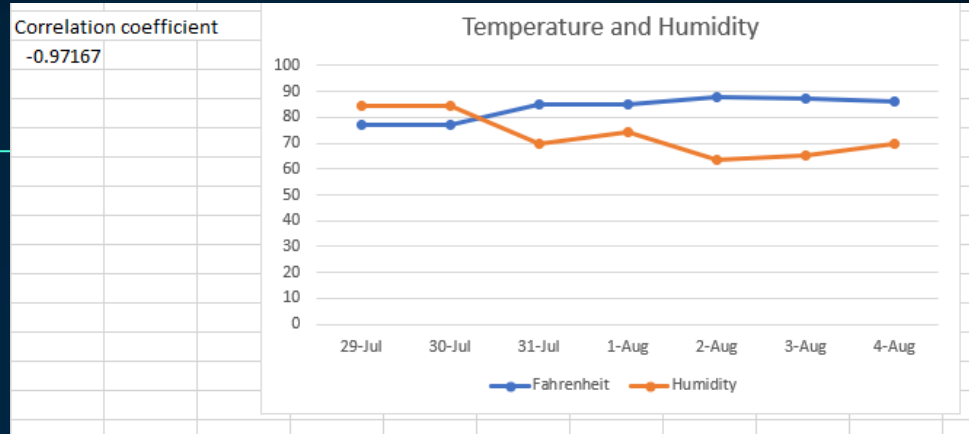


Plot 4



Prediction

Based on the 2 weeks of data gathered I would predict that the temperature will rise leading to lower humidity.



My Own Analysis

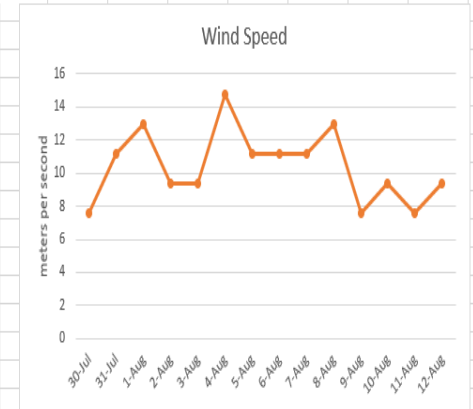
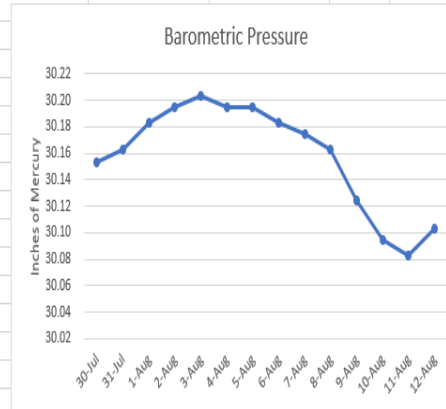
- Find the answer to my own question
- Use correlation formula to validate claims
- Provide plots to visualize claims

Correlation Between Barometric Pressure and Wind Speed

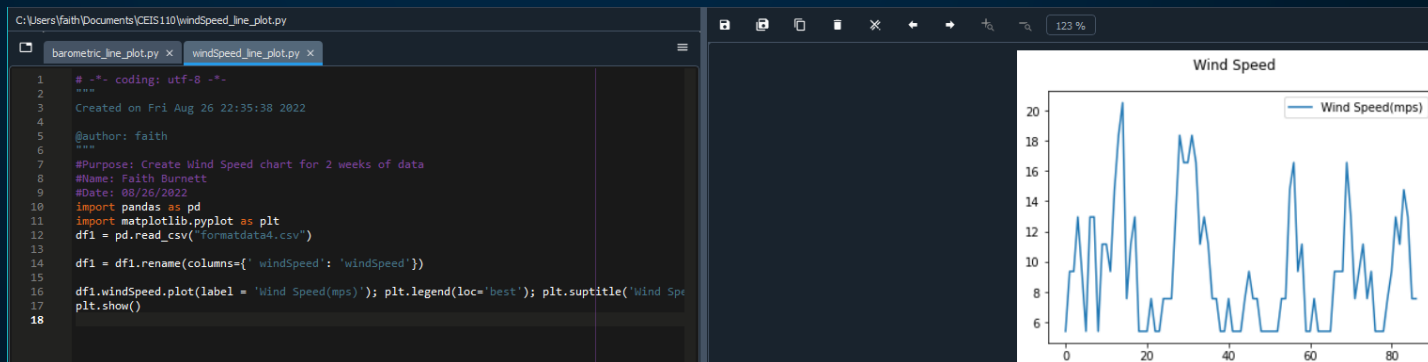
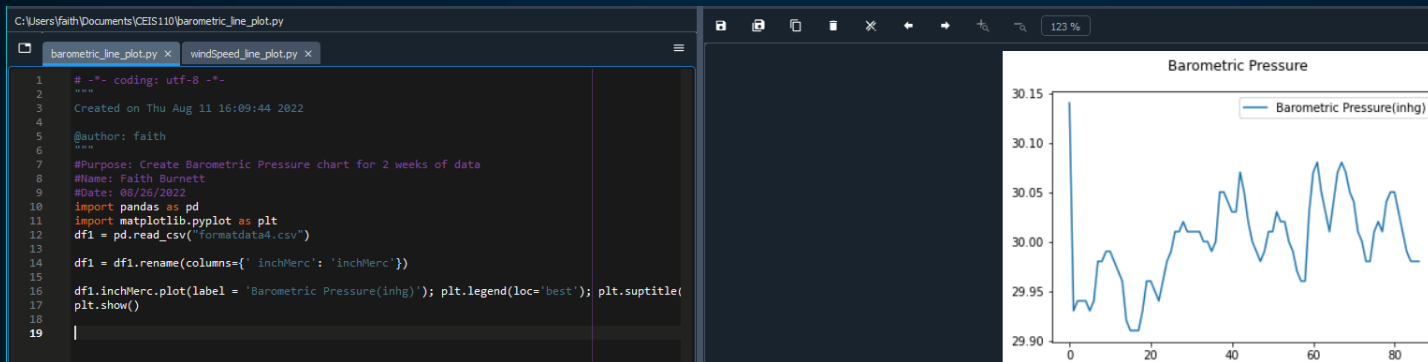
What is the correlation between barometric pressure and wind speed?

Using the correlation function we can see that there is a negative correlation meaning that when one goes up the other goes down.

| Week | barometri | Constant | InchMerc | windSpee | MinimumW | MaximumW | Minimum Inch Merc | Maximum Inch Merc |
|--------|-----------|----------|----------|----------|-------------|--------------|-------------------|-------------------|
| 30-Jul | 1021 | 33.86 | 30.15 | 7.56 | 5.4 | 20.52 | 29.91 | 30.20 |
| 31-Jul | 1021.3 | | 30.16 | 11.16 | | | | |
| 1-Aug | 1022 | | 30.18 | 12.96 | Correlation | -0.042352058 | | |
| 2-Aug | 1022.4 | | 30.19 | 9.36 | | | | |
| 3-Aug | 1022.7 | | 30.20 | 9.36 | | | | |
| 4-Aug | 1022.4 | | 30.19 | 14.76 | | | | |
| 5-Aug | 1022.4 | | 30.19 | 11.16 | | | | |
| 6-Aug | 1022 | | 30.18 | 11.16 | | | | |
| 7-Aug | 1021.7 | | 30.17 | 11.16 | | | | |
| 8-Aug | 1021.3 | | 30.16 | 12.96 | | | | |
| 9-Aug | 1020 | | 30.12 | 7.56 | | | | |
| 10-Aug | 1019 | | 30.09 | 9.36 | | | | |
| 11-Aug | 1018.6 | | 30.08 | 7.56 | | | | |
| 12-Aug | 1019.3 | | 30.10 | 9.36 | | | | |
| | 1020 | | 30.12 | 5.4 | | | | |
| | 1020.3 | | 30.13 | 5.4 | | | | |
| | 1020.7 | | 30.14 | 5.4 | | | | |
| | 1021.3 | | 30.16 | 14.76 | | | | |
| | 1021.7 | | 30.17 | 9.36 | | | | |
| | 1021.7 | | 30.17 | 5.4 | | | | |
| | 1021.7 | | 30.17 | 5.4 | | | | |
| | 1021 | | 30.15 | 5.4 | | | | |
| | 1019.7 | | 30.12 | 9.36 | | | | |
| | 1019.3 | | 30.10 | 11.16 | | | | |
| | 1019.7 | | 30.12 | 12.96 | | | | |
| | 1019.7 | | 30.12 | 12.96 | | | | |



Python Plots of Barometric Pressure and Wind Speed



Conclusion

- Implemented a weather data analysis system
- Stored downloaded information inside a database to retrieve later for analysis
- Performed analysis on selected tables from database
- Created plots to visualize data
- Made predictions for future dates based off data collected
- Produced a question of my own to find a correlation between other data sets

Challenges

- Query commands
- Trying to produce my own question about the data
- Making predictions without more data to analyze

Skills Learned

- Python
- Project organization using flowcharts, and IPO
- Data analysis
- Designing, coding, and testing programs

THANKS!

Does anyone have any question?

faithburnett@outlook.com

CREDITS

- Presentation template by Slidesgo