

CEIS110

Programming with Data

```
for object to mirror_mod.mirror_object
    operation == "MIRROR_X":
        mirror_mod.use_x = True
        mirror_mod.use_y = False
        mirror_mod.use_z = False
    operation == "MIRROR_Y":
        mirror_mod.use_x = False
        mirror_mod.use_y = True
        mirror_mod.use_z = False
    operation == "MIRROR_Z":
        mirror_mod.use_x = False
        mirror_mod.use_y = False
        mirror_mod.use_z = True

#selection at the end -add
mirror_ob.select= 1
modifier_ob.select=1
context.scene.objects.active
= bpy.context.selected_object
data.objects[one.name].select

print("please select exactly one")

--- OPERATOR CLASSES ---

bpy.types.Operator):
    X mirror to the selected
    object.mirror_mirror_x"
    mirror X"
```



Introduction



Data is growing

This project uses a cloud system to gather
temperature and humidity data

Then the data is analyzed using
programming data

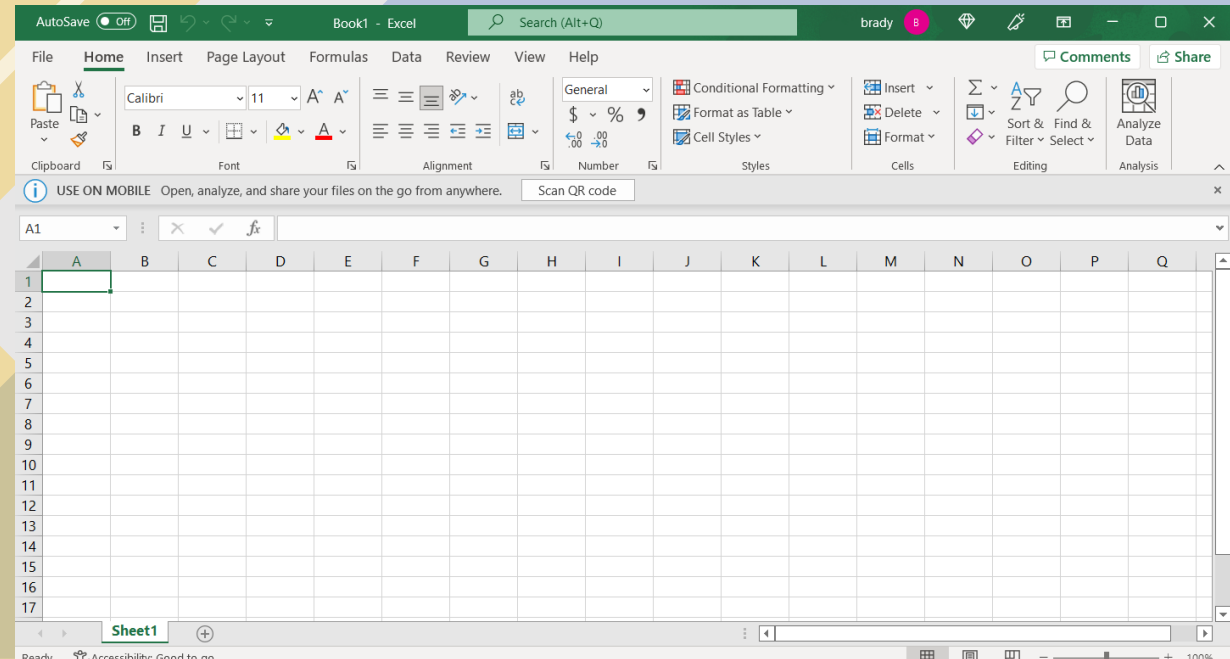
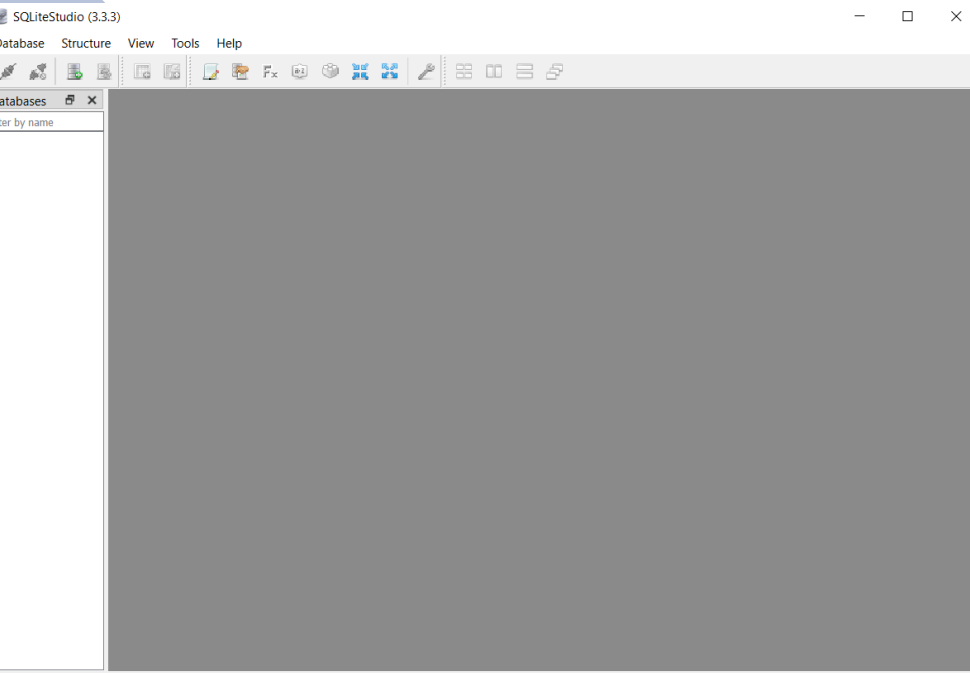
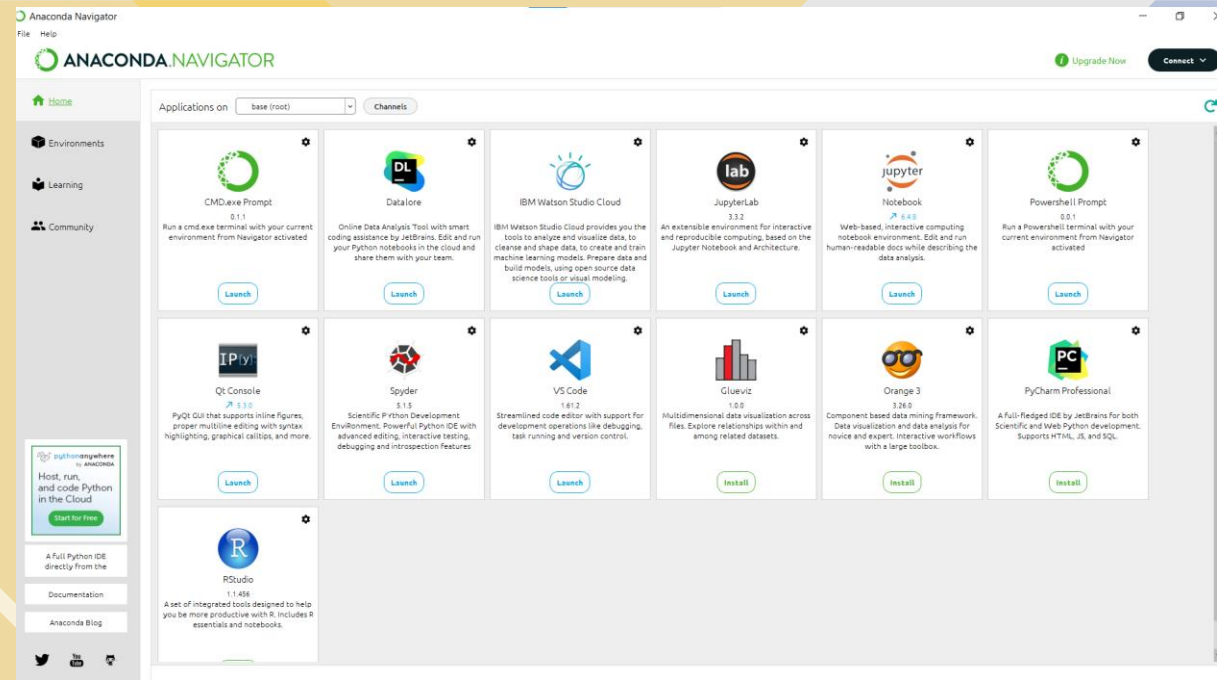


Software Inventory

- Before developing programming with data project, all software must be downloaded and installed
 - Software needed included SQLite studio, Excel, and Python IDE - anaconda
- 
- 

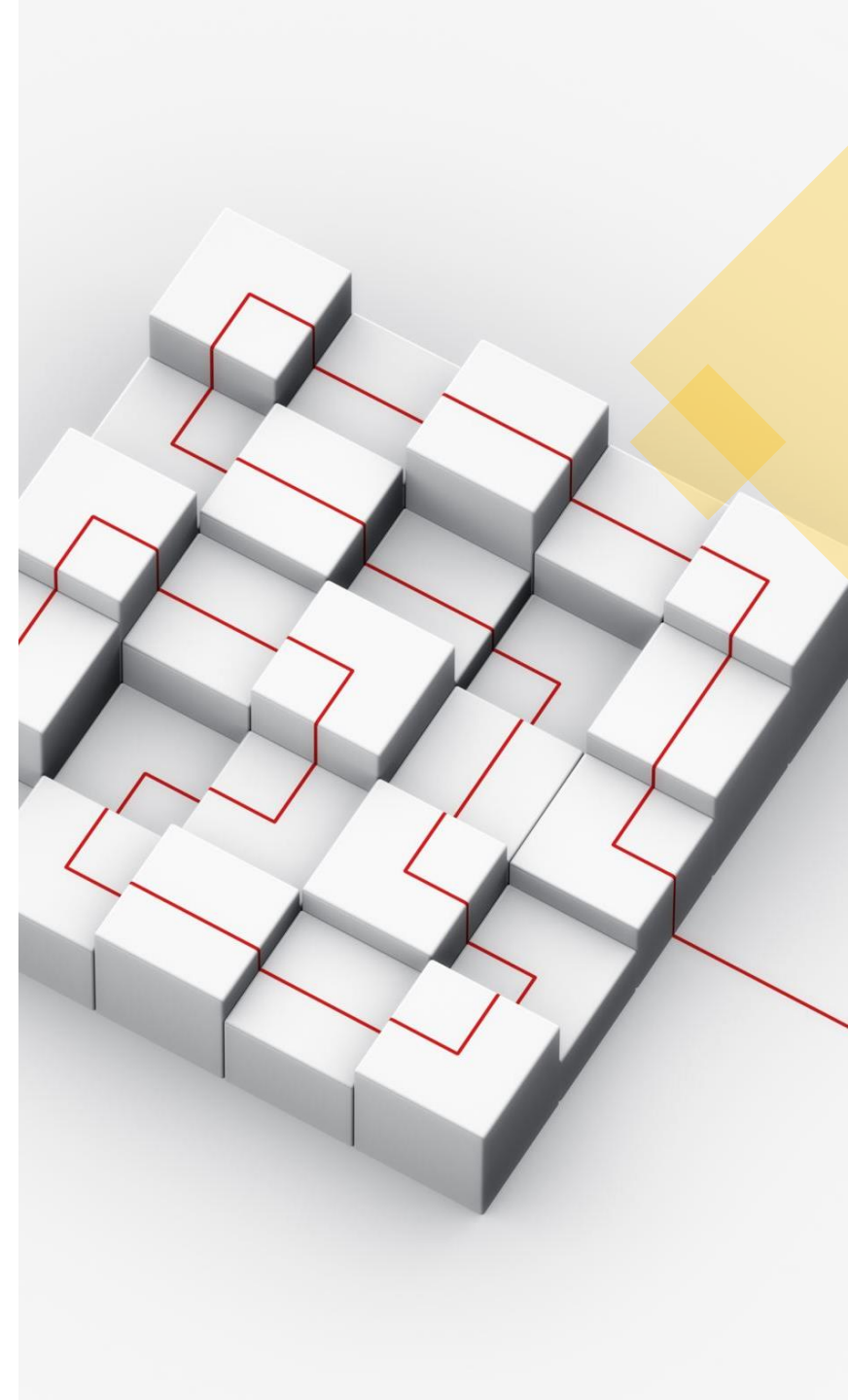
Software:

The Software needed for this project includes Microsoft Excel
SQLite anaconda Navigation
with Spyder



Planning and Design

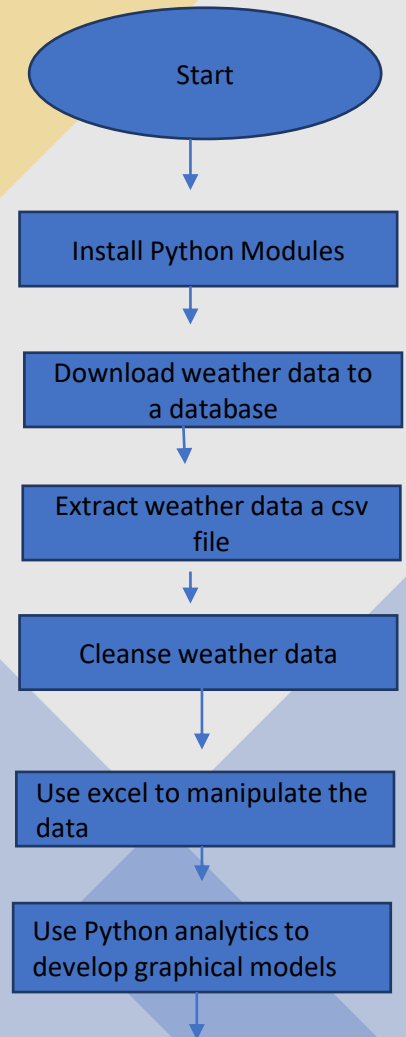
- After performing of the software needed for the project, a plan was created for the data and temperature project.
- To plan the design of the project, a flowchart was generated.
- Planning and design of the project are crucial steps to understanding the development process.



Flowchart

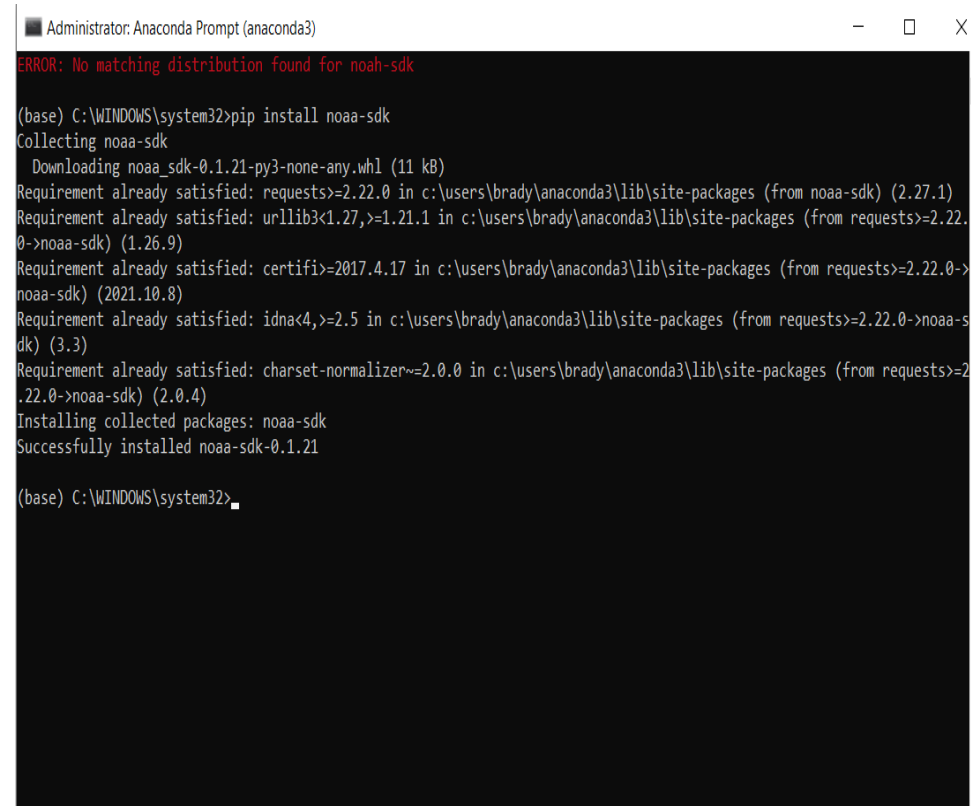
Include the following processes:

- Install python
- Download weather data to a database.
- Extract weather data from database into a comma separated file with python
- Cleanse weather data
- Use Excel to manipulate data
- Use python data analytics modules to develop graphical models



Adding Library

- In order for python to connect to the US government Nation oceanic and Atmospheric Administration (NOAA) weather data service using a cloud-based Application Programming Interface (API), a library module must be installed to use this service.
- The screenshot shows the NOAA-SDK library installed.



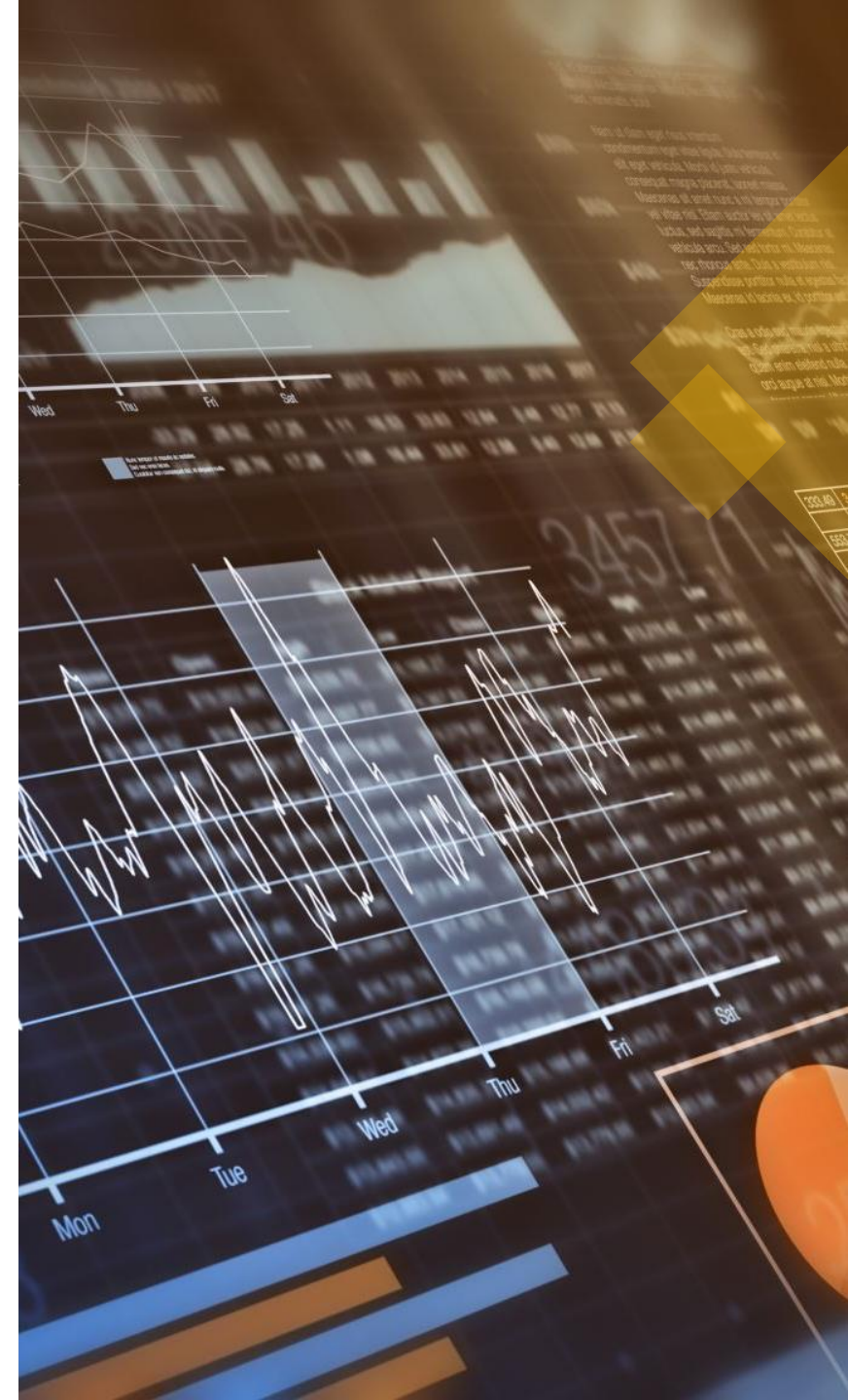
```
Administrator: Anaconda Prompt (anaconda3)
ERROR: No matching distribution found for noah-sdk

(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 c:\users\brady\anaconda3\lib\site-packages (from noaa-sdk) (2.27.1)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in c:\users\brady\anaconda3\lib\site-packages (from requests>=2.22.0->noaa-sdk) (1.26.9)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\brady\anaconda3\lib\site-packages (from requests>=2.22.0->noaa-sdk) (2021.10.8)
Requirement already satisfied: idna<4,>=2.5 in c:\users\brady\anaconda3\lib\site-packages (from requests>=2.22.0->noaa-sdk) (3.3)
Requirement already satisfied: charset-normalizer<=2.0.0 in c:\users\brady\anaconda3\lib\site-packages (from requests>=2.22.0->noaa-sdk) (2.0.4)
Installing collected packages: noaa-sdk
Successfully installed noaa-sdk-0.1.21

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

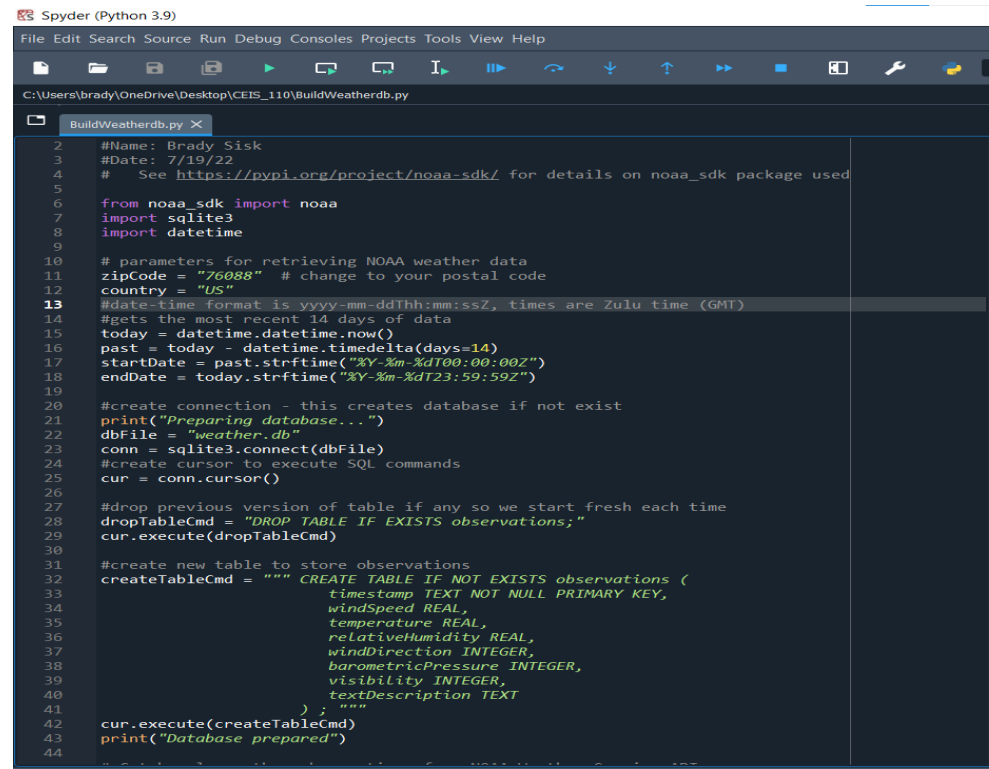

Gathering Temperature and Humidity data

- After planning and design, the code was developed to download a set of weather observations
- This data was stored on local database in table for later analysis



BuildWeatherDb.py

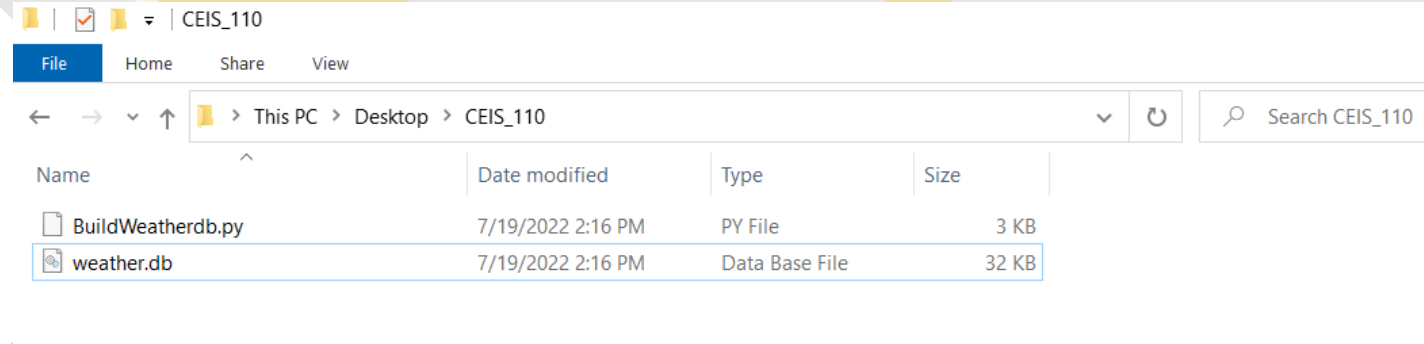
- Screenshot of code in Spyder
- The code will create a table named Observations, the fields timestamp, windspeed, temperature, relativeHumidity, windDirection, barometricPressure, visibility, and textDescription.
- The database will be named weather.db and stored in the same directory as the python code



```
1  #Name: Brady Sisk
2  #Date: 7/19/22
3  # See https://pypi.org/project/noaa-sdk/ for details on noaa_sdk package used
4
5
6  from noaa_sdk import noaa
7  import sqlite3
8  import datetime
9
10 # parameters for retrieving NOAA weather data
11 zipCode = "76088" # change to your postal code
12 country = "US"
13 #datetime format is yyyy-mm-ddTth:mm:ssZ, times are Zulu time (GMT)
14 #gets the most recent 14 days of data
15 today = datetime.datetime.now()
16 past = today - datetime.timedelta(days=14)
17 startDate = past.strftime("%Y-%m-%dT00:00:00Z")
18 endDate = today.strftime("%Y-%m-%dT23:59:59Z")
19
20 #create connection - this creates database if not exist
21 print("Preparing database...")
22 dbFile = "weather.db"
23 conn = sqlite3.connect(dbFile)
24 #create cursor to execute SQL commands
25 cur = conn.cursor()
26
27 #drop previous version of table if any so we start fresh each time
28 dropTableCmd = "DROP TABLE IF EXISTS observations;"
29 cur.execute(dropTableCmd)
30
31 #create new table to store observations
32 createTableCmd = """ CREATE TABLE IF NOT EXISTS observations (
33     timestamp TEXT NOT NULL PRIMARY KEY,
34     windSpeed REAL,
35     temperature REAL,
36     relativeHumidity REAL,
37     windDirection INTEGER,
38     barometricPressure INTEGER,
39     visibility INTEGER,
40     textDescription TEXT
41 ); """
42 cur.execute(createTableCmd)
43 print("Database prepared")
44
```

Weather db File

Screenshot of Windows explorer showing database wether db was created.






Querying the Database

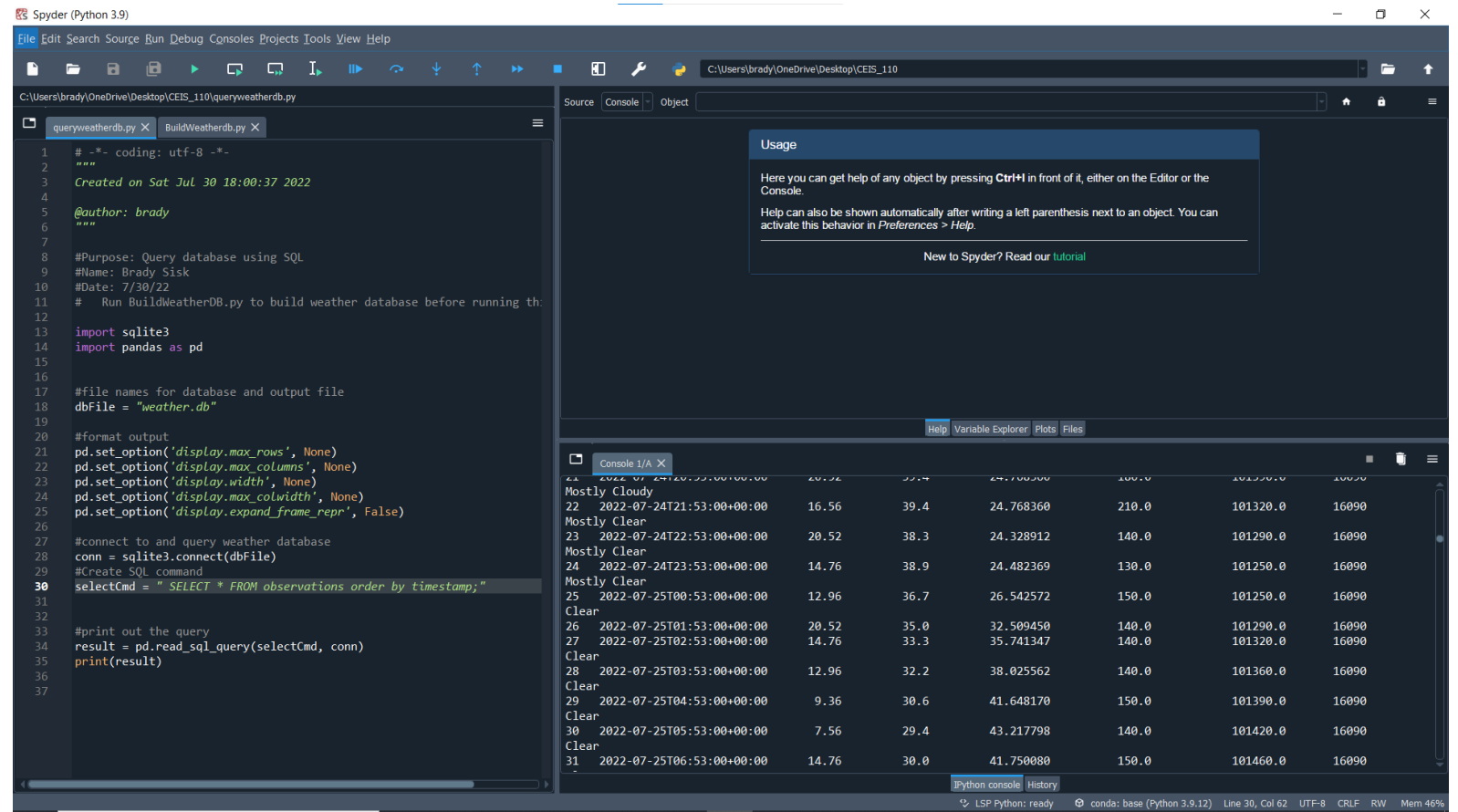


Querying the Database

- Structured Query Language is a programming language used for working with relational database
 - SQLiteStudio was used to query the database and view the results
- 

Query to retrieve all columns and all rows

- The SQL command “select * from observations” was executed to retrieve all rows and columns from the observations table.



The screenshot displays the Spyder Python IDE interface. The left pane shows a Python script named `queryweatherdb.py` with the following content:

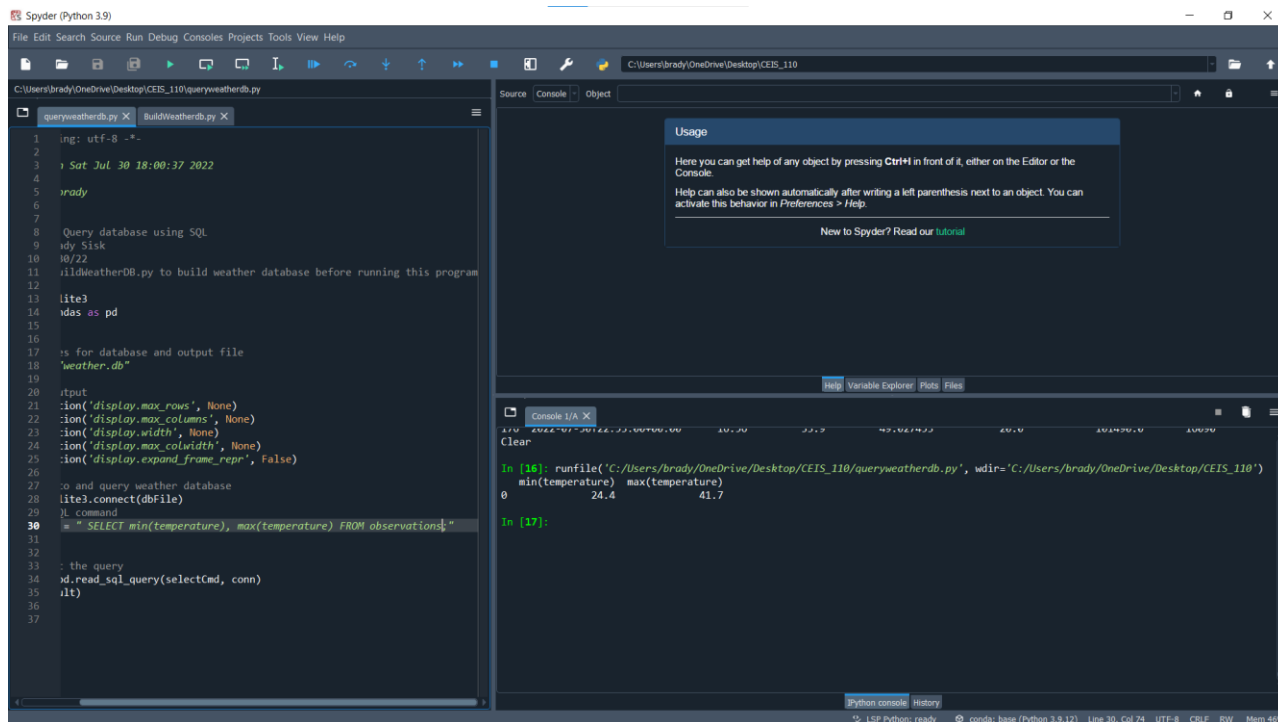
```
1  # -*- coding: utf-8 -*-
2  """
3  Created on Sat Jul 30 18:00:37 2022
4
5  @author: brady
6  """
7
8  #Purpose: Query database using SQL
9  #Name: Brady Sisk
10 #Date: 7/30/22
11 # Run BuildWeatherDB.py to build weather database before running th
12
13 import sqlite3
14 import pandas as pd
15
16 #file names for database and output file
17 dbFile = "weather.db"
18
19 #format output
20 pd.set_option('display.max_rows', None)
21 pd.set_option('display.max_columns', None)
22 pd.set_option('display.width', None)
23 pd.set_option('display.max_colwidth', None)
24 pd.set_option('display.expand_frame_repr', False)
25
26 #connect to and query weather database
27 conn = sqlite3.connect(dbFile)
28 #Create SQL command
29 selectCmd = " SELECT * FROM observations order by timestamp;"
30
31 #print out the query
32 result = pd.read_sql_query(selectCmd, conn)
33 print(result)
```

The right pane shows the console output of the script, displaying a table of weather observations. The table has 7 columns: `timestamp`, `temp`, `humidity`, `wind`, `pressure`, `visibility`, and `clouds`. The output shows 31 rows of data, including weather conditions like "Mostly Cloudy", "Mostly Clear", and "Clear".

timestamp	temp	humidity	wind	pressure	visibility	clouds
2022-07-24T20:53:00+00:00	20.52	39.4	24.768360	210.0	101320.0	16090
2022-07-24T21:53:00+00:00	16.56	39.4	24.768360	210.0	101320.0	16090
2022-07-24T22:53:00+00:00	20.52	38.3	24.328912	140.0	101290.0	16090
2022-07-24T23:53:00+00:00	14.76	38.9	24.482369	130.0	101250.0	16090
2022-07-25T00:53:00+00:00	12.96	36.7	26.542572	150.0	101250.0	16090
2022-07-25T01:53:00+00:00	20.52	35.0	32.509450	140.0	101290.0	16090
2022-07-25T02:53:00+00:00	14.76	33.3	35.741347	140.0	101320.0	16090
2022-07-25T03:53:00+00:00	12.96	32.2	38.025562	140.0	101360.0	16090
2022-07-25T04:53:00+00:00	9.36	30.6	41.648170	150.0	101390.0	16090
2022-07-25T05:53:00+00:00	7.56	29.4	43.217798	140.0	101420.0	16090
2022-07-25T06:53:00+00:00	14.76	30.0	41.750880	150.0	101460.0	16090

- The min and max temperatures were retrieved. These temperatures are captured based on the Celsius scale.

Query to retrieve lowest and highest temperature



The screenshot shows the Spyder Python IDE interface. The left pane displays a Python script named `queryweatherdb.py` with the following content:

```
1 # -*- coding: utf-8 -*-
2
3 # Sat Jul 30 18:00:37 2022
4
5 # brady
6
7
8 # Query database using SQL
9 # by Sisk
10 # 10/22
11 # BuildWeatherDB.py to build weather database before running this program
12
13 # lite3
14 # db as pd
15
16
17 # is for database and output file
18 # 'weather.db'
19
20 # output
21 # ion('display.max_rows', None)
22 # ion('display.max_columns', None)
23 # ion('display.width', None)
24 # ion('display.max_colwidth', None)
25 # ion('display.expand_frame_repr', False)
26
27 # do and query weather database
28 # lite3.connect(dbFile)
29 # IL command
30 # = "SELECT min(temperature), max(temperature) FROM observations;"
31
32
33 # the query
34 # id.read_sql_query(selectCmd, conn)
35 # it)
36
37
```

The right pane shows the console output for the execution of the script:

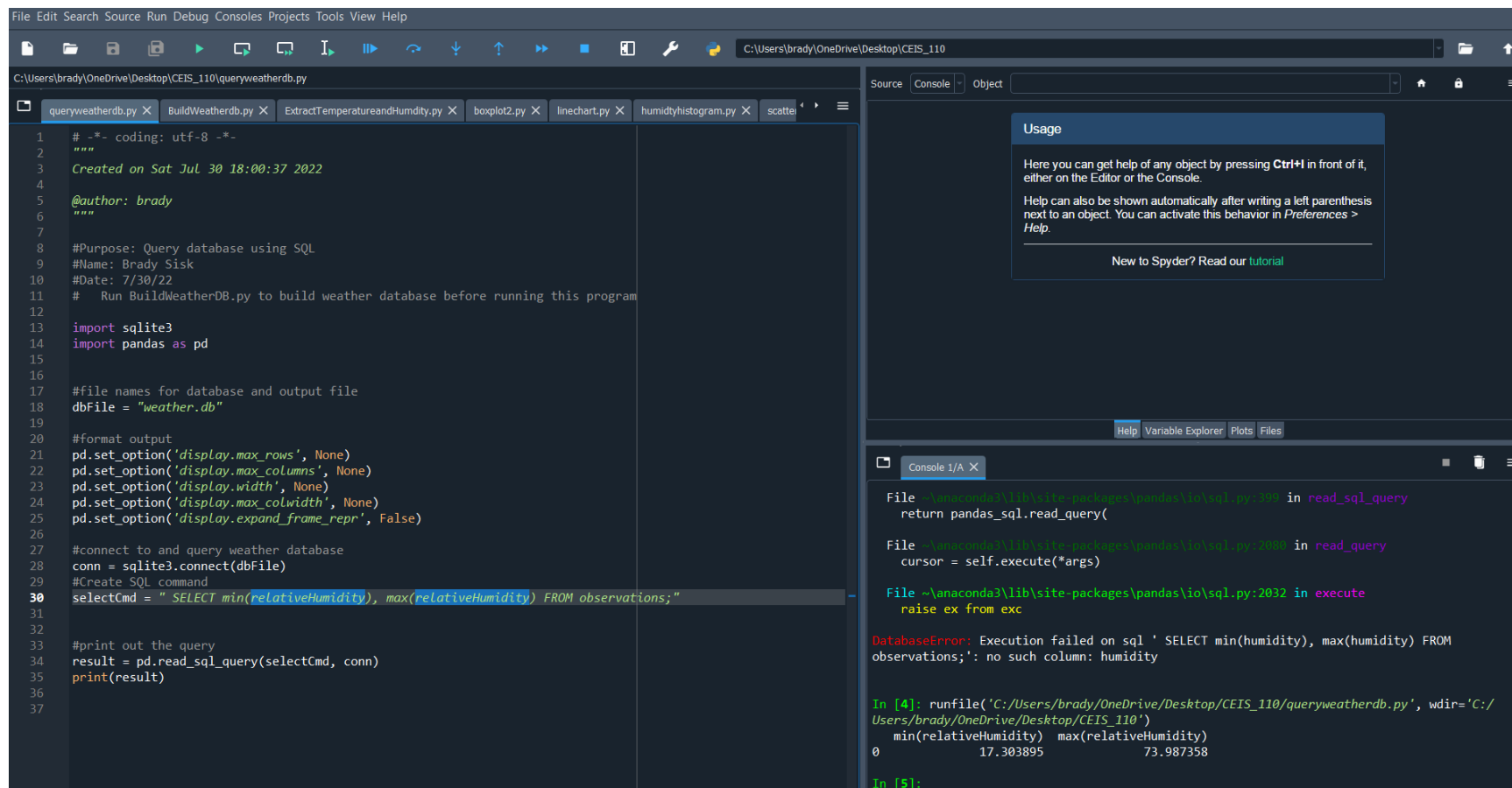
```
In [16]: runfile('C:/Users/brady/OneDrive/Desktop/CEIS_110/queryweatherdb.py', wdir='C:/Users/brady/OneDrive/Desktop/CEIS_110')
min(temperature)  max(temperature)
0                24.4             41.7

In [17]:
```

The status bar at the bottom indicates the current environment: `conda: base (Python 3.9.12)`, `Line 30, Col 74`, `UTF-8`, `CRLF`, `Raw`, `Mem 49%`.

Query to retrieve lowest and highest humidity

- Another query performed to retrieve the lowest and highest relative humidity. The relativeHumidity column was used to retrieve this information.



The screenshot shows the Spyder IDE interface. The main editor displays a Python script named `queryweatherdb.py`. The script includes a header with author information and a purpose statement: "Query database using SQL". It imports `sqlite3` and `pandas`, and sets various display options for `pandas`. The script connects to a database file named `weather.db` and executes a SQL query to retrieve the minimum and maximum relative humidity from the `observations` table. The console output shows the execution of the script, including the SQL query and the resulting data.

```
1  # -*- coding: utf-8 -*-
2  """
3  Created on Sat Jul 30 18:00:37 2022
4
5  @author: brady
6  """
7
8  #Purpose: Query database using SQL
9  #Name: Brady Sisk
10 #Date: 7/30/22
11 # Run BuildWeatherDB.py to build weather database before running this program
12
13 import sqlite3
14 import pandas as pd
15
16
17 #file names for database and output file
18 dbFile = "weather.db"
19
20 #format output
21 pd.set_option('display.max_rows', None)
22 pd.set_option('display.max_columns', None)
23 pd.set_option('display.width', None)
24 pd.set_option('display.max_colwidth', None)
25 pd.set_option('display.expand_frame_repr', False)
26
27 #connect to and query weather database
28 conn = sqlite3.connect(dbFile)
29 #Create SQL command
30 selectCmd = " SELECT min(relativeHumidity), max(relativeHumidity) FROM observations;"
31
32
33 #print out the query
34 result = pd.read_sql_query(selectCmd, conn)
35 print(result)
36
37
```

The console output shows the execution of the script, including the SQL query and the resulting data:

```
File ~\anaconda3\lib\site-packages\pandas\io\sql.py:199 in read_sql_query
return pandas_sql.read_query(

File ~\anaconda3\lib\site-packages\pandas\io\sql.py:200 in read_query
cursor = self.execute(*args)

File ~\anaconda3\lib\site-packages\pandas\io\sql.py:203 in execute
raise ex from exc

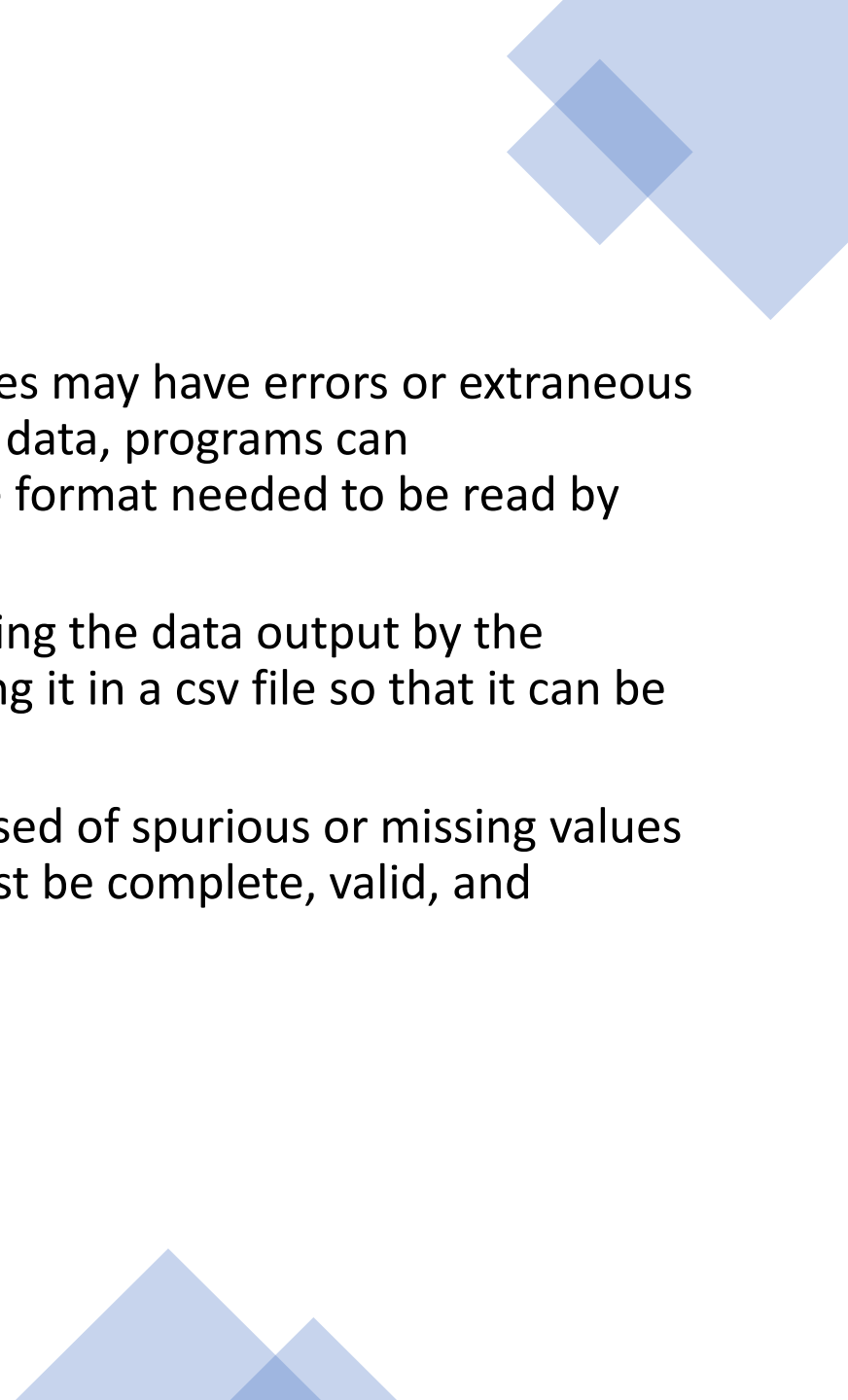
DatabaseError: Execution failed on sql ' SELECT min(humidity), max(humidity) FROM
observations;': no such column: humidity

In [4]: runfile('C:/Users/brady/OneDrive/Desktop/CEIS_110/queryweatherdb.py', wdir='C:/
Users/brady/OneDrive/Desktop/CEIS_110')
min(relativeHumidity)  max(relativeHumidity)
0                    17.303895                73.987358

In [5]:
```

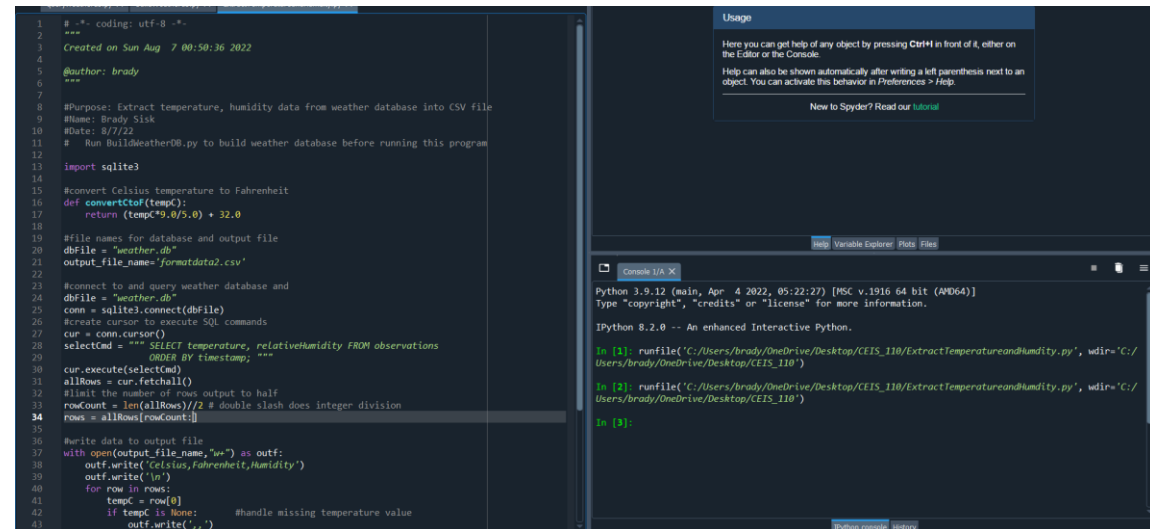



Data Cleansing

- Data output from machines may have errors or extraneous data. When cleansing the data, programs can automatically put it in the format needed to be read by other programs.
 - A python program is reading the data output by the python program and saving it in a csv file so that it can be read by Excel.
 - Often data must be cleansed of spurious or missing values in a dataset. The data must be complete, valid, and standardized.
- 

Extracting Temperature and Humidity using Python code

- The weather.db database may contain null or missing values. The code used retrieves only the temperature and humidity and writes them to a comma separated values (CSV) file. Two files are created – a formatdata.csv and formatdata.2 that each contain half of the rows. Missing or invalid values are not written to the file.



```
1 # -*- coding: utf-8 -*-
2 """
3 Created on Sun Aug 7 00:50:36 2022
4
5 @author: brady
6 """
7
8 #Purpose: Extract temperature, humidity data from weather database into CSV file
9 #Name: Brady Sisk
10 #Date: 8/7/22
11 # Run BuildWeatherDB.py to build weather database before running this program
12
13 import sqlite3
14
15 #convert Celsius temperature to Fahrenheit
16 def convertCtoF(tempC):
17     return (tempC*9.0/5.0) + 32.0
18
19 #file names for database and output file
20 dbFile = "weather.db"
21 output_file_name = "formatdata2.csv"
22
23 #connect to and query weather database and
24 dbFile = "weather.db"
25 conn = sqlite3.connect(dbFile)
26 #create cursor to execute SQL commands
27 cur = conn.cursor()
28 selectCmd = """ SELECT temperature, relativeHumidity FROM observations
29                 ORDER BY timestamp; """
30 cur.execute(selectCmd)
31 allRows = cur.fetchall()
32 #limit the number of rows output to half
33 rowCount = len(allRows)//2 # double slash does integer division
34 rows = allRows[rowCount:]
35
36 #write data to output file
37 with open(output_file_name, "w") as outf:
38     outf.write('Celsius,Fahrenheit,Humidity')
39     outf.write('\n')
40     for row in rows:
41         tempC = row[0]
42         if tempC is None: #handle missing temperature value
43             outf.write(',,')
```

Usage

Here you can get help of any object by pressing `Ctrl+H` in front of it, either on the Editor or the Console.

Help can also be shown automatically after writing a left parenthesis next to an object. You can activate this behavior in `Preferences > Help`.

[New to Spyder? Read our tutorial](#)

Help Variable Explorer Plot File

Console 1/1 X

Python 3.9.12 (main, Apr 4 2022, 05:22:27) [MSC v.1916 64 bit (AMD64)]
Type "copyright", "credits" or "license()" for more information.

IPython 8.2.0 -- An enhanced Interactive Python.

In [1]: runfile('C:/Users/brady/OneDrive/Desktop/CEIS_110/ExtractTemperatureandHumidity.py', wdir='C:/Users/brady/OneDrive/Desktop/CEIS_110')

In [2]: runfile('C:/Users/brady/OneDrive/Desktop/CEIS_110/ExtractTemperatureandHumidity.py', wdir='C:/Users/brady/OneDrive/Desktop/CEIS_110')

In [3]:

Python console history

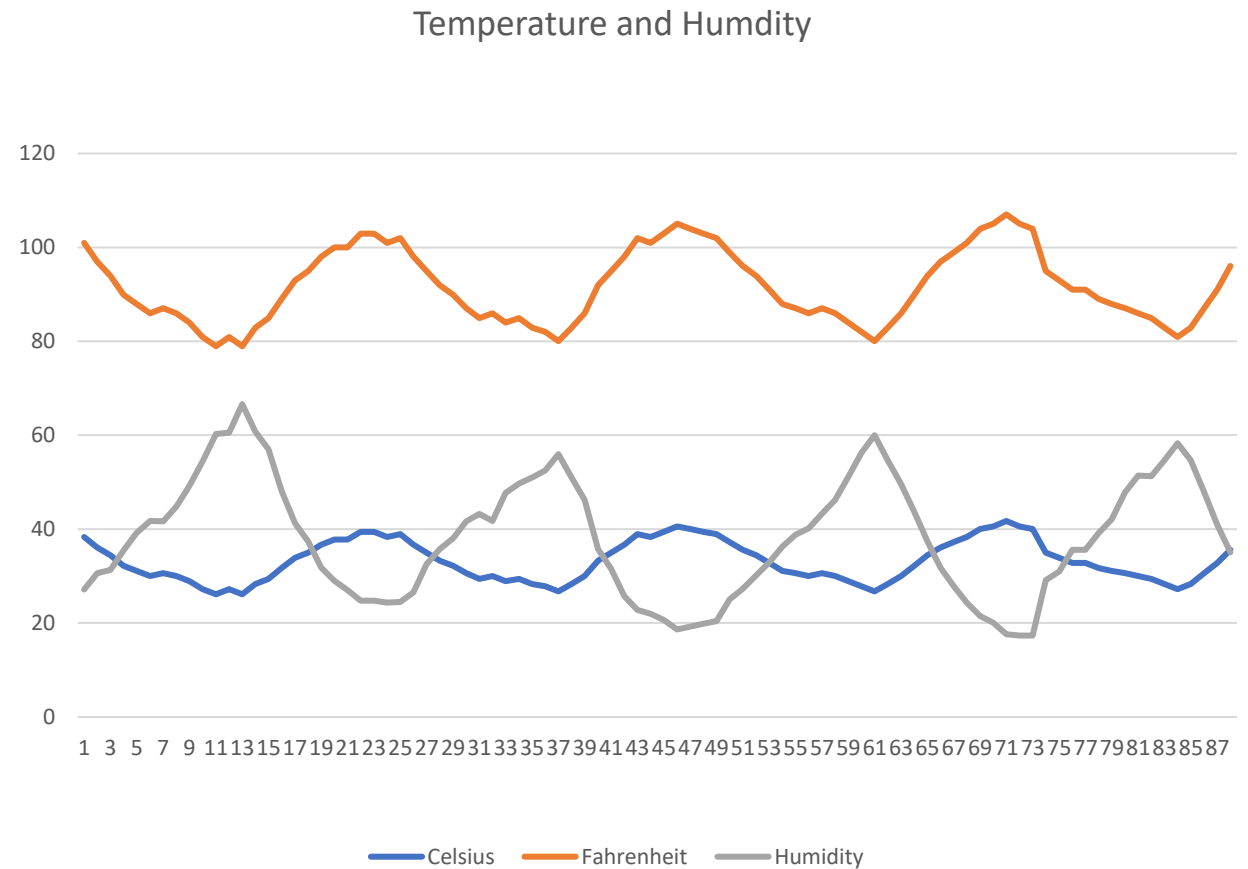
Data Formatted in an Excel Spreadsheet

- The python program created a formatdata.csv file
- This file contains 3 columns: Celsius, Fahrenheit, and Humidity
- Statistics can be performed on this spreadsheet

	A	B	C
1	Celsius	Fahrenheit	Humidity
2	38.3	100.94	27.1349
3	36.1	96.98	30.59447
4	34.4	93.92	31.32396
5	32.2	89.96	35.43829
6	31.1	87.98	39.2029
7	30	86	41.75008
8	30.6	87.08	41.64817
9	30	86	44.78357
10	28.9	84.02	49.25382
11	27.2	80.96	54.38442
12	26.1	78.98	60.2634
13	27.2	80.96	60.5158
14	26.1	78.98	66.60937
15	28.3	82.94	60.76613
16	29.4	84.92	57.01373
17	31.7	89.06	48.15548
18	33.9	93.02	41.24052
19	35	95	37.36307
20	36.7	98.06	31.74628
21	37.8	100.04	28.9657
22	37.8	100.04	27.00366
23	39.4	102.92	24.76836
24	39.4	102.92	24.76836
25	38.3	100.94	24.32891

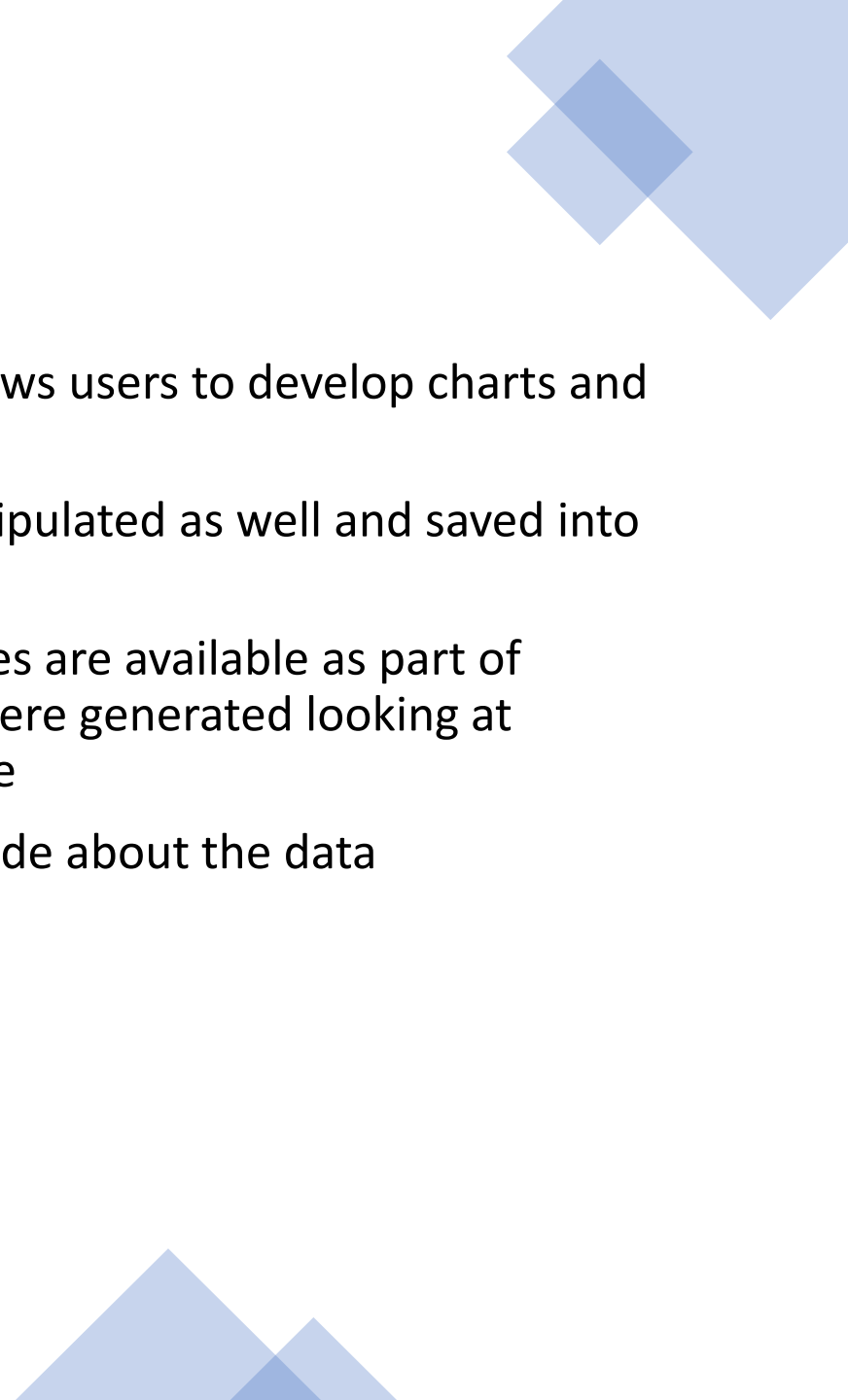
Data Visualization

A line chart was developed in Excel showing the Temperature and Humidity over Period 1



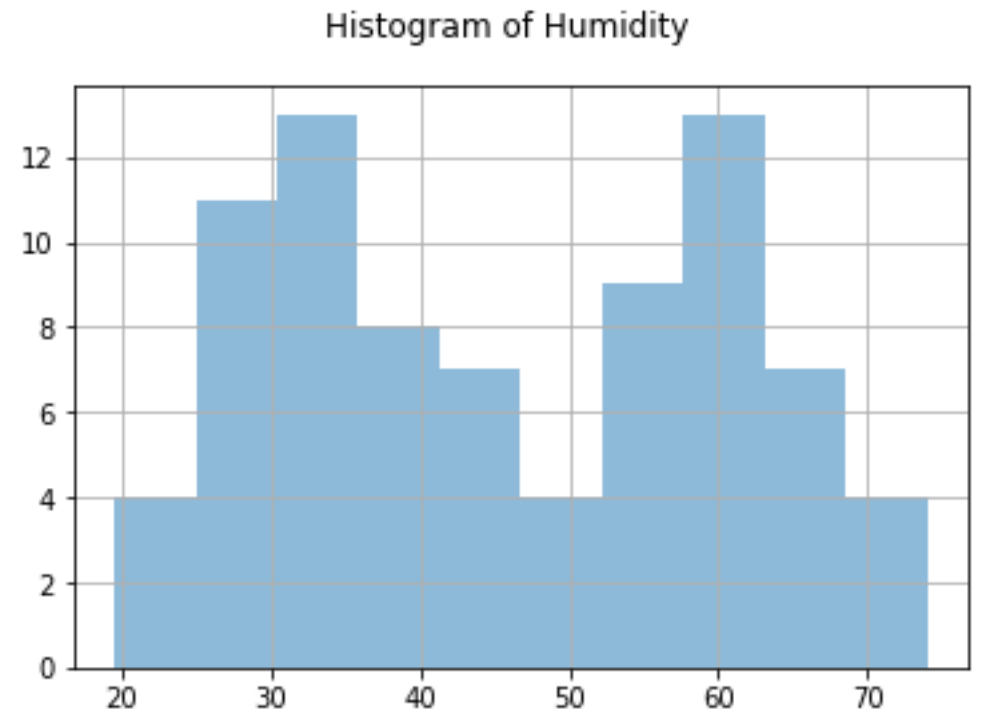


Data Analytics

- Python Data modules allows users to develop charts and graphs depicting data.
 - The data sets can be manipulated as well and saved into tabular format
 - The data analytics modules are available as part of Anaconda Several plots were generated looking at humidity and temperature
 - Then a prediction was made about the data
- 

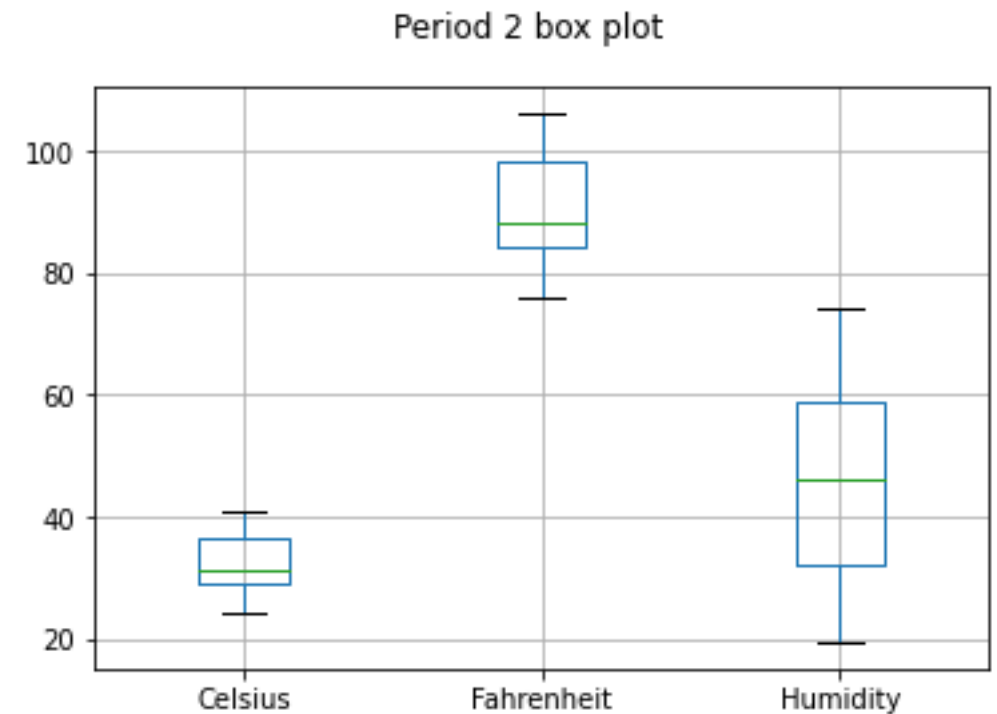
Histogram of Humidity

```
#Purpose: Create a histogram of humidity data from the second
period
#Name: Brady Sisk
#Date: 8/13/22
import pandas as pd
import matplotlib.pyplot as plt
df1 = pd.read_csv("formatdata.csv")
df2 = pd.read_csv("formatdata2.csv")
df2['Humidity'].hist(bins=10, alpha=0.5); plt.suptitle('Histogram
of Humidity')
```



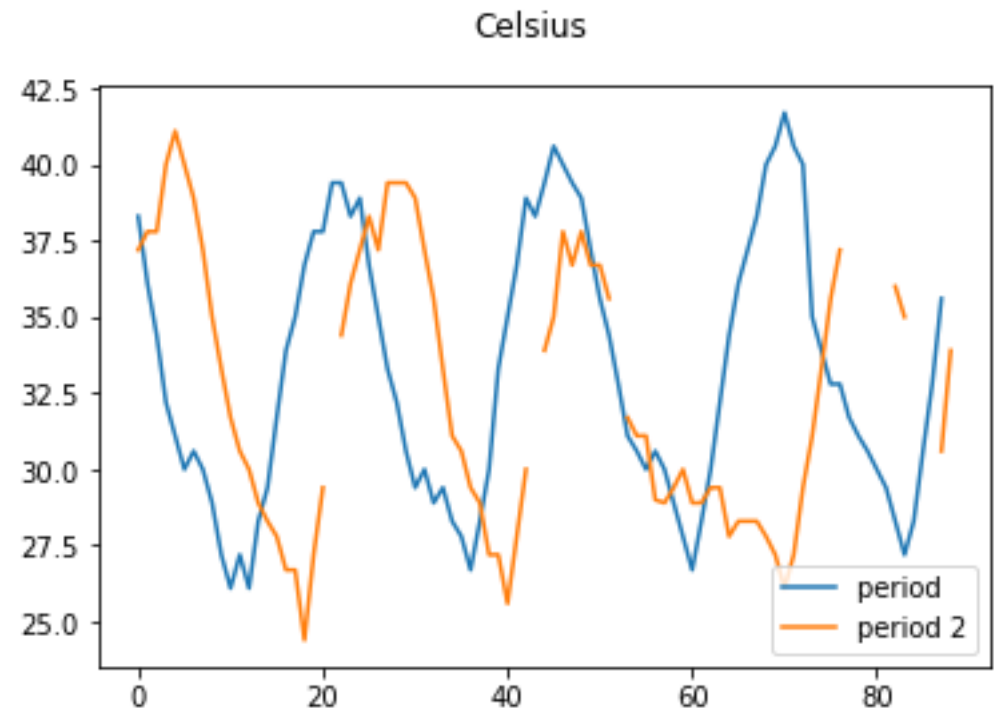
Period 2 Box Plot

```
#Purpose: Create box plot for period 2 data
#Name: Brady Sisk
#Date: 8/13/22
import pandas as pd
import matplotlib.pyplot as plt
df2 = pd.read_csv("formatdata2.csv")
df2.boxplot(); plt.suptitle('Period 2 box plot')
plt.show()
```



Analysis

- Think of your own question and create a chart/graph to answer it
- Your own question:
 - Was there a big difference between period 1 and
 - Period 2?
- Answer supported by Chart:
- No, judging by the chart I would say the periods were very similar

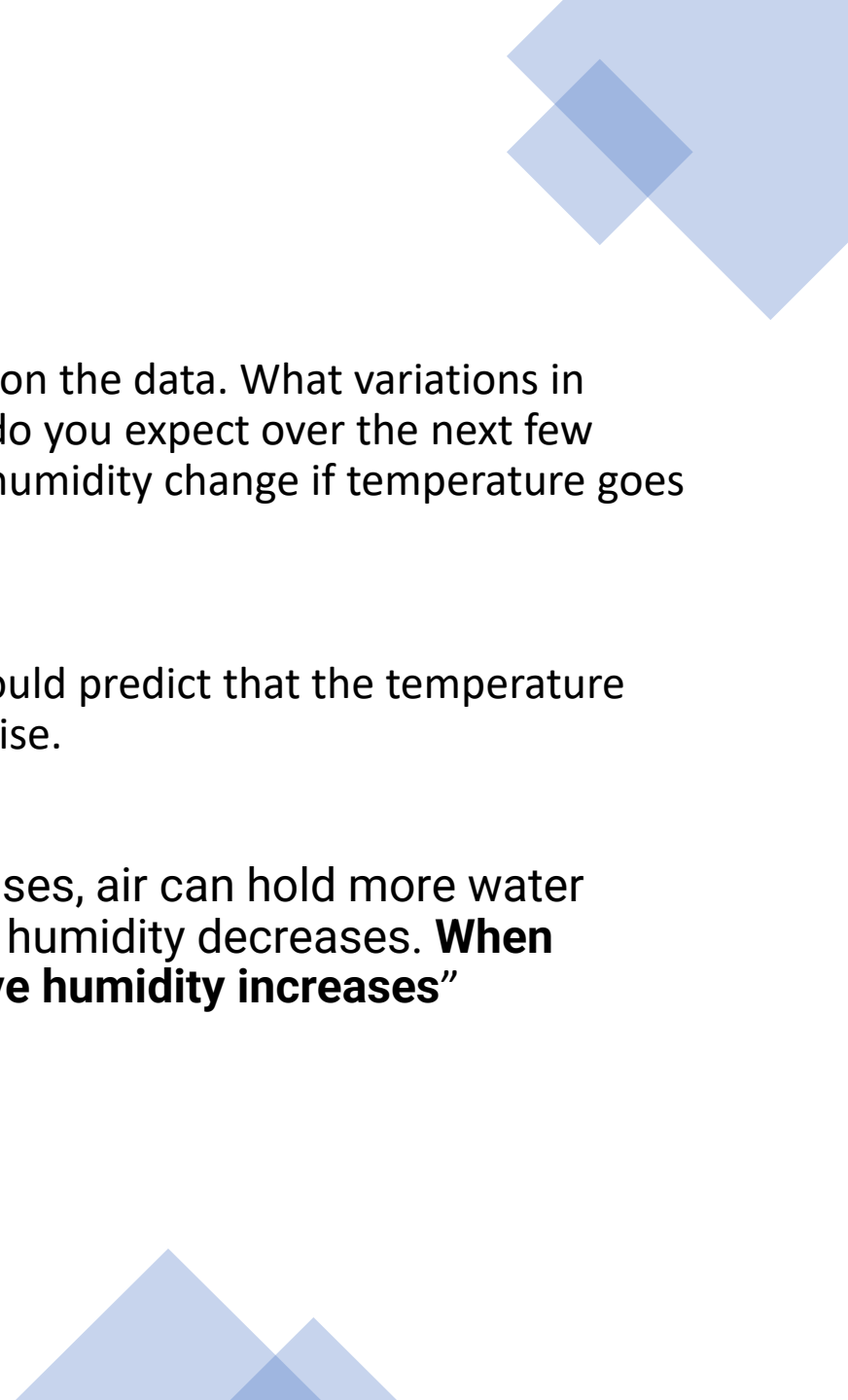


Code

- #Purpose: Create Celsius plot comparing period 1 and period 2
- #Name: Brady Sisk
- #Date: 8/13/22
- import pandas as pd
- import matplotlib.pyplot as plt
- df1 = pd.read_csv("formatdata.csv") #baseline data is period 1 (older)
- df2 = pd.read_csv("formatdata2.csv") #data for period 2 (more recent)
- plt.figure(); df1.Celsius.plot(label = 'period '); df2.Celsius.plot(label = 'period 2'); plt.legend(loc='best'); plt.suptitle('Celsius')
- plt.show()

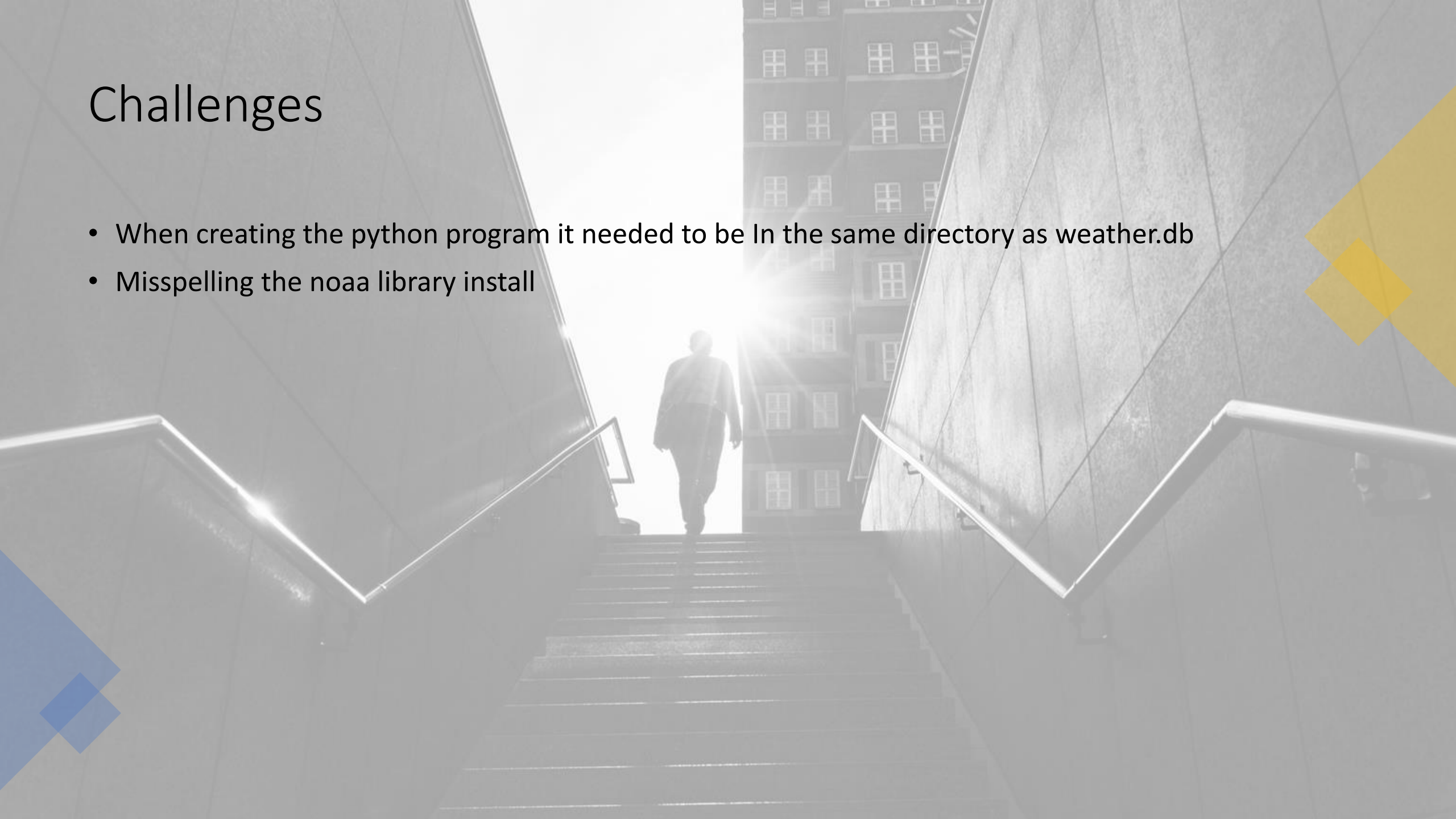


Prediction

- Develop a prediction based on the data. What variations in temperature and humidity do you expect over the next few hours or days? How would humidity change if temperature goes up or down?
 - Over the next few hours I would predict that the temperature will drop and the humidity rise.
 - “As air temperature increases, air can hold more water molecules, and its relative humidity decreases. **When temperatures drop, relative humidity increases**”
- 

Challenges

- When creating the python program it needed to be In the same directory as weather.db
- Misspelling the noaa library install



Career Skills

- Several career skills were gained in this project.
 - Communication – using flowcharts to depict the plan of a project
 - Database Development
 - Programming using python
 - Troubleshooting errors in the code and data cleansing
 - Analysis – Reviewing and graphs to make prediction on the data



Conclusion

- This project covered the fundamental topics of programming with data by using data gathered from cloud service to perform data analytics operations
 - Building this project provided a hands on learning opportunity to put into practice the topics covered in the course.
- 