

# Exploring Sensor Data

By the end of this activity, you will be able to:

- 1. Identify the major components in semi-structured data from a weather station
- 2. Create plots of weather station data

**Step 1. Open a terminal shell and activate your virtual environment.** Open your local terminal shell and go to your *big-data-2/* directory. Activate your virtual environment using the command that corresponds to your operating system (adjust the command if you named your virtual environment differently).

Windows:

```
1 .\big-data-2-env\Scripts\Activate
```

macOS:

```
1 source big-data-2-env/bin/activate
```

**Step 2. Go to sensor directory.** With you virtual environment activated go to your *sensor* directory.

PS C:\Users\[redacted]\Desktop\coursera\big-data-2\sensor>

Run *ls* to see the sensor data files and scripts:

Directory: C:\Users\[redacted]\Desktop\coursera\big-data-2\sensor		
Mode	LastWriteTime	Length Name
----	-----	-----
-a----	12/11/2023 2:33 PM	944 plot-data.py
-a----	12/13/2023 1:46 PM	256 stream-data.py
-a----	12/13/2023 2:14 PM	1586 stream-plot-data.py
-a----	4/19/2016 3:30 PM	6119195 wx-data.txt
-a----	3/14/2016 2:59 PM	1018 wxt-520-format.txt

**Step 3. View weather station data.** Twenty-four hours of weather station data is in the file *wx-data.txt*. View this data by running *more wx-data.txt*:

```
1 more wx-data.txt
```



Each line in the file contains a timestamp in the first column and a set of measurements in the second column. These column are separated by a tab. The measurements are also a set of columns separated by commas.

The timestamp is the number of seconds since January 1, 1970. By looking at the successive timestamps in the file, we can see that the measurements arrive about every second.

The measurements are prefixed with *R0*, *R1*, *R2*, etc. Each R-value holds a different set of measurements. For example, *R1* has *Dn*, *Dm*, *Dx*, etc. Additionally, the different types are measured at different frequencies. *R1* is measured every second, but *R0* and *R2* are much less frequent.

**Step 4. Look at key for measurements.** We can learn what the measurement fields mean by looking at the measurements key file. Open a new terminal and *cd* to *Downloads/big-data-2/sensor*. Run *more wxt-520-format.txt*:

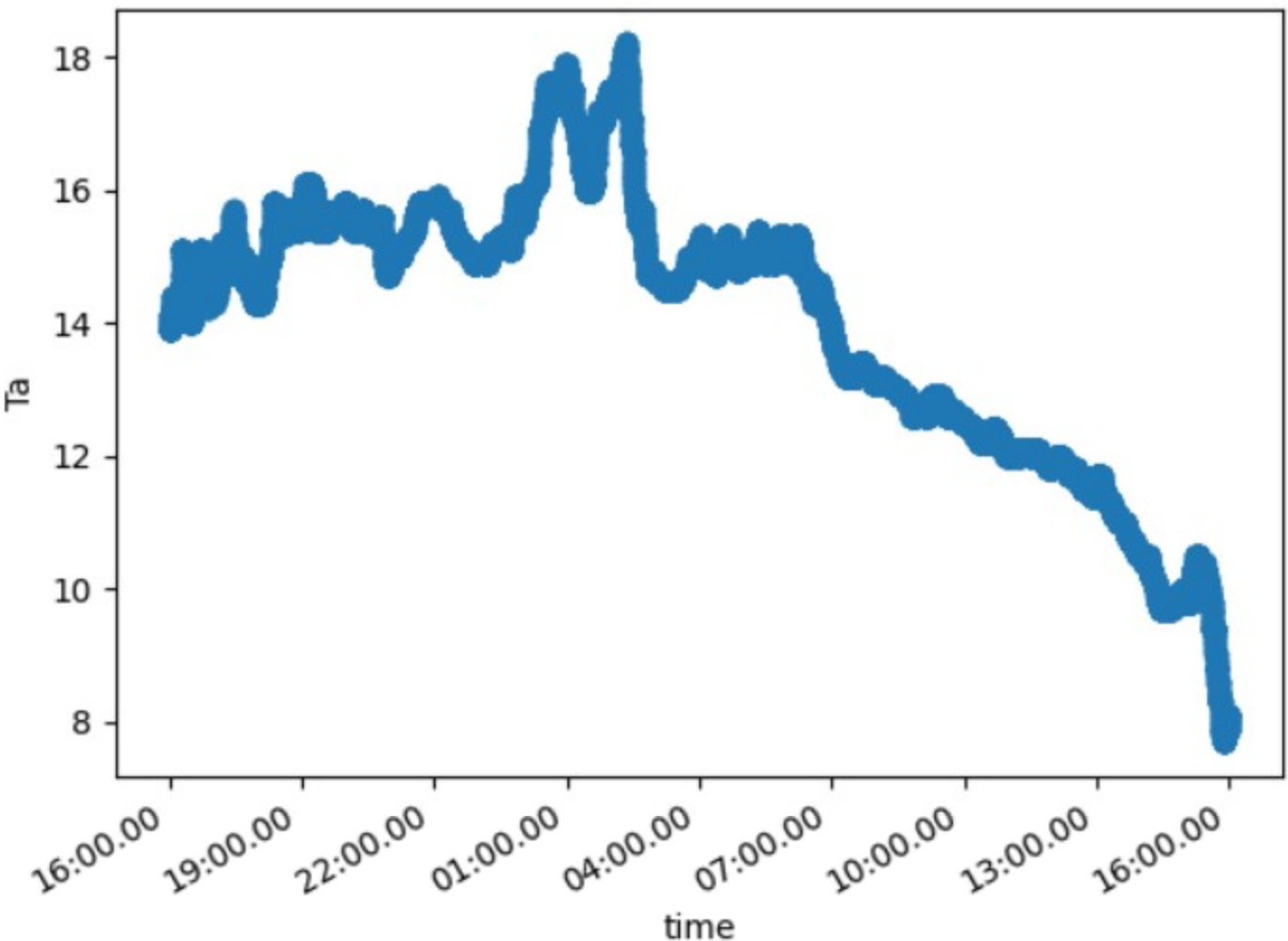
```
1 more wxt-520-format.txt
```

```
Sn      Wind speed minimum m/s, km/h, mph, knots #,M, K, S, N
Sm      Wind speed average m/s, km/h, mph, knots #,M, K, S, N
Sx      Wind speed maximum m/s, km/h, mph, knots #,M, K, S, N
Dn      Wind direction minimum deg #, D
Dm      Wind direction average deg #, D
Dx      Wind direction maximum deg #, D
Pa      Air pressure hPa, Pa, bar, mmHg, inHg #, H, P, B, M, I
Ta      Air temperature °C, °F #, C, F
Tp      Internal temperature °C, °F #, C, F
Ua      Relative humidity %RH #, P
Rc      Rain accumulation mm, in #, M, I
Rd      Rain duration s #, S
Ri      Rain intensity mm/h, in/h #, M, I
Rp      Rain peak intensity mm/h, in/h #, M, I
Hc      Hail accumulation hits/cm2, hits/in2, hits #, M, I, H
Hd      Hail duration s #, S
Hi      Hail intensity hits/cm2h, hits/in2h, hits/ h #, M, I, H
Hp      Hail peak intensity hits/cm2h, hits/in2h, hits/ h #, M, I, H
Th      Heating temperature °C, °F #, C, F
Vh      Heating voltage V #, N, V, W, F2
Vs      Supply voltage V V
Vr      3.5 V ref. voltage V V
```

This file provides a list of what each field means. For example, *Ta* is the Air temperature, and *Dn* is the Wind direction minimum.

**Step 5. Plot data.** Run *python plot-data.py wx-data.txt Ta* to plot the air temperature:

```
1 python3 ./plot-data.py wx-data.txt Ta
```



You can plot other measurements in the file by using the field name as the last argument. For example, use *Pa* to plot the air pressure:

```
1 python3 ./plot-data.py wx-data.txt Pa
```

**Step 6. Deactivate your virtual environment.** Run *deactivate* to deactivate the virtual environment.

```
1 deactivate
```