# Analysis of observed temperature vs normal

by Kjell Svensson,  2020-02-28

*Project URL: [https://github.com/k-svensson/Temperature-analysis](https://github.com/k-svensson/Temperature-analysis)*

This is a little project to demonstrate the following:

- using API to get data from an online service
- data wrangling in Python and Power BI
- visualizing data and conclusions in Power BI

## Project background

I have a Fine Offset WH2900 weather station, purchased in August 2019, which is more than accurate enough for my needs and it uploads measurements continuously to Weather Underground (from now on, "WU"). The station is also sold under several other brands/models, such as PanTech WH2900 and Ambient Weather WS-2902.

WU has changed their data policy over time which makes it more difficult to get access to your own data. Since I had no intention to lose this information, I wanted to download it regularly to store on a local disk where **I** decide what to do with it.

I found that using their API, I could write a Python script that gets all the data for a particular day and saves that to disk. This script is run every day using the Windows task scheduler.

In order to do some visualizing of this data, I decided to get some answers to these questions:

- The winter of 2019/20 has been very mild in Oslo, but *how* mild has it been?
- How many days have been over or under the "normal" temperature?
- How do my observations at home compare to the official station at Blindern?
- Which days had the most extreme differences at home, compared to the normal curve?

And let me emphasize that this is not meant to a contribution to the current climate debate, but more my being curious of how I could get data off the web, and to compare that to other observations and long-term averages.

I could probably have done this project as a Jupyter notebook, but I wanted to see what I could do with Power BI to visualize and present the data.

# Methodology

Data needed:

1. Normal temperature per day for the official station at Blindern
2. Actual temperature observed for the period 1 September 2019 to 25 February 2020.
3. Temperature observations for my weather station at Galgeberg, Oslo

Sources for the data:

1. Meteorologisk Institutt eKlima service, [www.met.no](http://www.met.no)
2. [www.met.no](http://www.met.no)
3. Weather Underground, [https://www.wunderground.com/dashboard/pws/IOSLO1862](https://www.wunderground.com/dashboard/pws/IOSLO1862), api.weather.com

*Definition of daily "normal temperature": The daily average temperature for the period 1961-1990.*

To-do:

- Get data from met.no
- Get data from WU, and convert JSON to CSV files
- Make the WU data the same granularity (daily) as the data from met.no
- Load and transform data into Power BI
- Remove unnecessary rows and columns from the tables
- Create calculated measures to use in visuals
- Create clear visuals and KPIs to illustrate findings
- Formulate a conclusion

# Results

- Over the whole period, the daily average deviation is +2.8 degrees C.
- There have been many more days that are warmer than days that are colder, compared to normal temperature for that date.
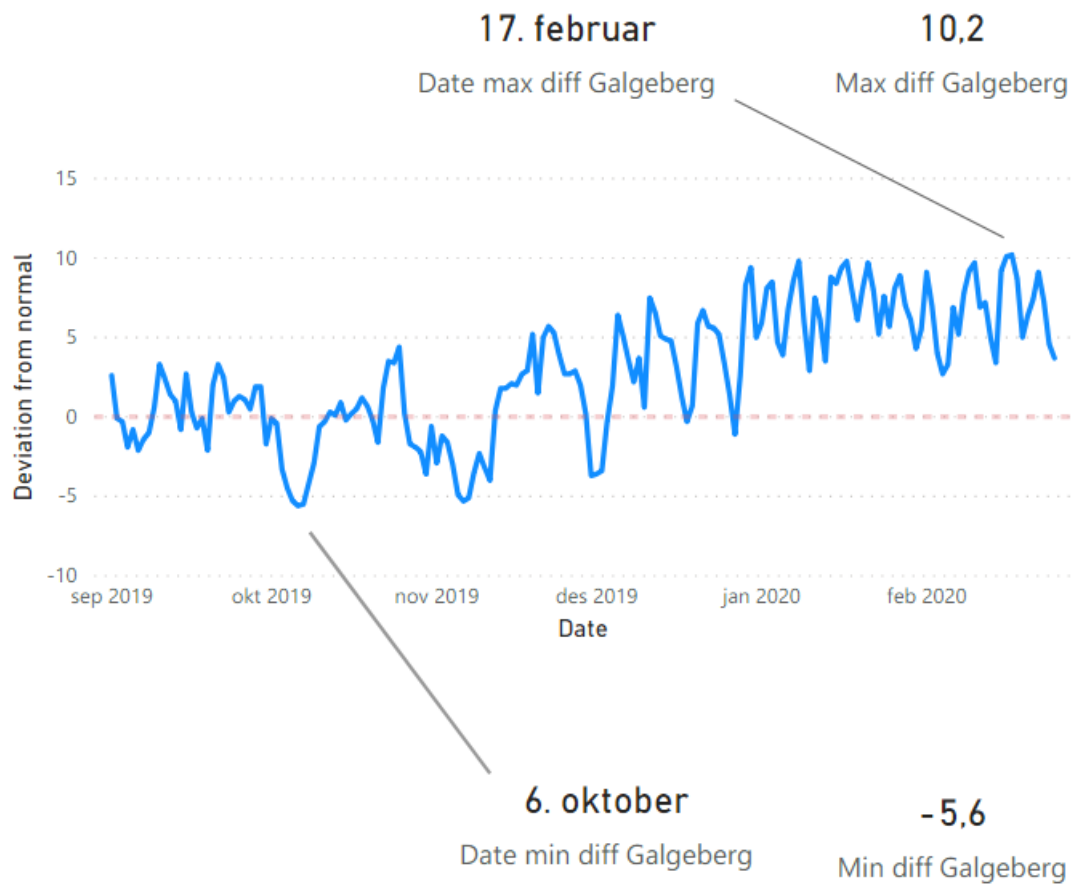- The deviation for warm days is greater than for cold days
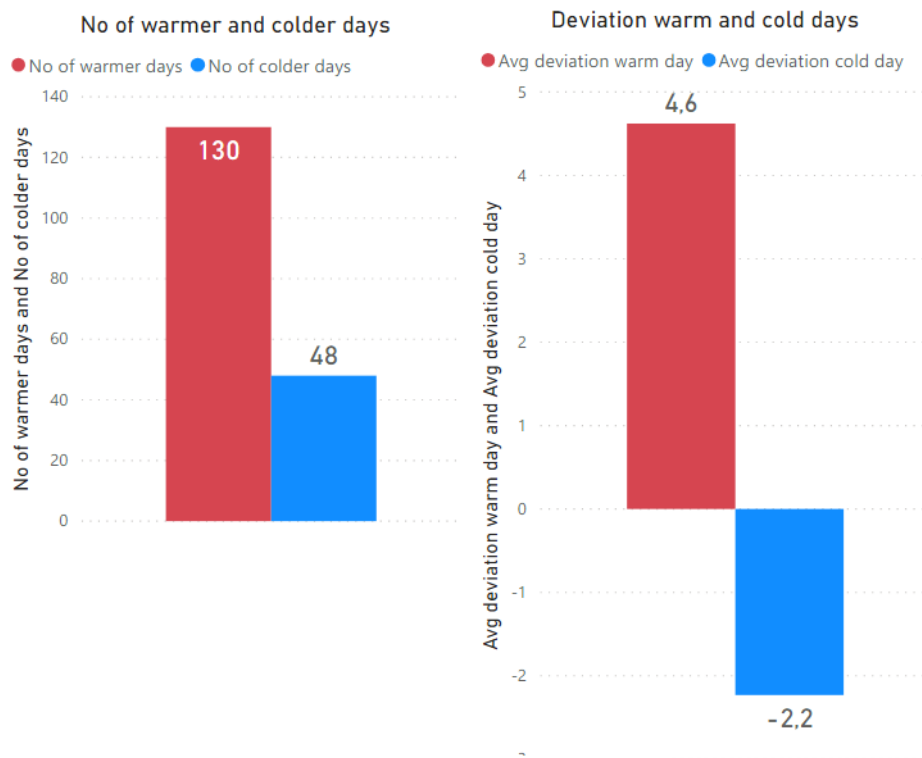


**2,77**
Overall daily deviation

- My personal station reports temperatures that are around 1 degree warmer than Blindern.
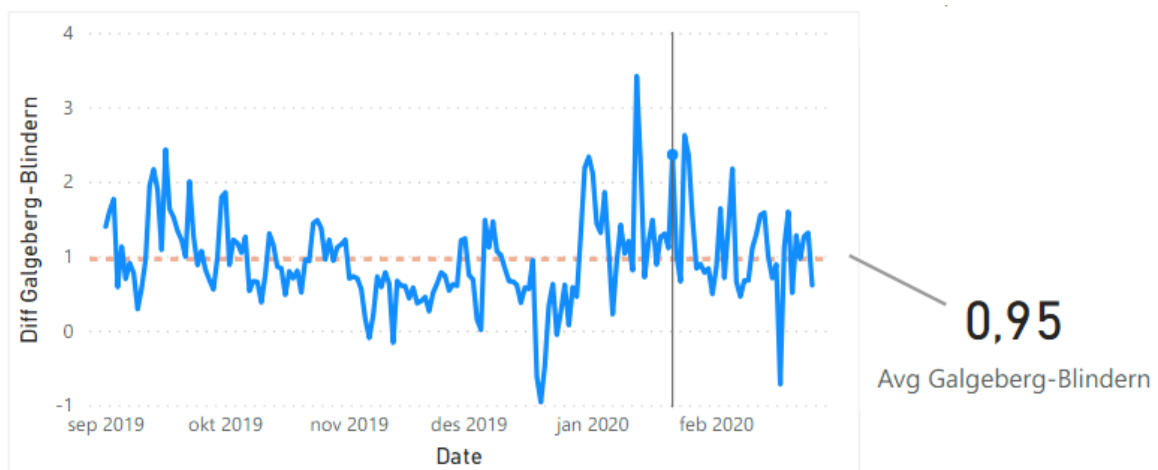
- The largest positive deviation at home occurred on 17th February, the largest negative deviation was on the 6th October.

**17. februar**

Date max diff Galgeberg

**10,2**

Max diff Galgeberg



**6. oktober**

Date min diff Galgeberg

**−5,6**

Min diff Galgeberg

# Discussion

As anyone who has lived in Oslo during this winter, the fact that it has been very mild comes as no surprise. Almost +3 degrees deviation from the norm does explain why there has been very little snow and a snow cover of zero for most of the period. From early December, the deviation has been around +5 degrees most of the time.



Daily deviation from normal temperature
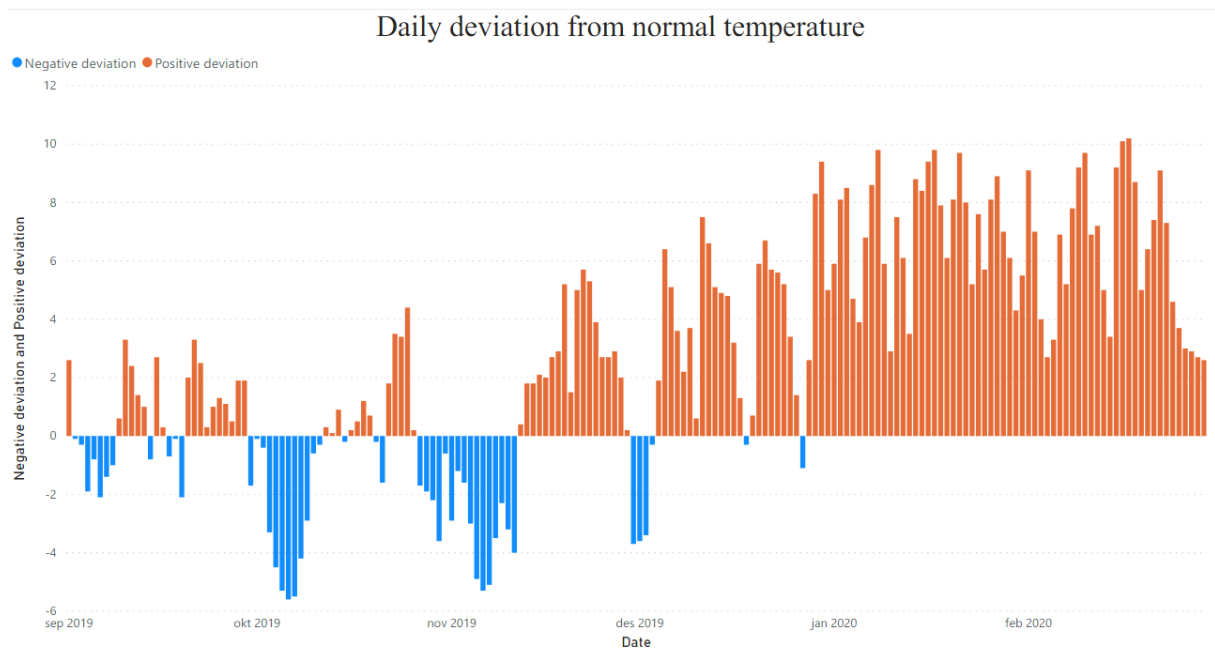
The observations at home show a similar pattern to Blindern, which also is to be expected as the distance between the two is only approximately 5 km. We have noticed that temperatures reported from Blindern seem to be lower than ours, and the stats show that the average difference is 1 degree. This can have several reasons, for example surrounding buildings, elevation, less suitable location for the home station, wind direction etc.

The Python code is included for reference and is only meant to be run once, by me.
It is by no means elegant or efficient but that wasn't my goal for this exercise.

There is not much data missing, so I have chosen to simply disregard that issue. Possible causes of data loss include wifi problems, but thankfully there has been none of that during this period!

## Possible extension of the analysis

Since WU records numerous measurements, we could extend the comparison to air pressure, precipitation, wind, etc. For this demo, however, I have decided to limit the scope to just temperature.

# Conclusion

During the period September 2019 to February 2020, there has been more than twice as many "warm" days as "cold" ones, and those warm days are also a lot warmer than usual, +4.6 degrees above average.

This gives an average temperature difference over the whole period of +2.8 degrees C, which is ***significantly above the normal*** and this explains why we haven't really had any snow in downtown Oslo this winter. Not a good year to sell winter sports equipment.