Textual description: Did the team include a textual description of their solution? Is it understandable? Is the described process correct?

Code quality: Is the code readable and concise? Does it use the powerful features of the libraries, or does it re-implement everything from scratch?

Results: Is the final result correct? Are all the assumptions well justified? Are there textual comments/visualizations to convince you of the final result?

Textual description: + Skrev ned hva som var i hver csv. Skrive det hva coden gjør på en veldig kort og forståelig måte. Code in each cell is well described.

Code quality: + Uses built in functions. Not too much code in each cell.

* Writes too long code lines. Doesn’t drop nans in task B. While oneliners looks nice, they are often a lot harder to read and understand when doing code reviews, harder to comment on each step of the code.

Results: +

* Is making negative years positive and years with 5 digits to 4 digits a fair assumption?

Don’t remove nans, just fill them. May have been better to remove the rows as there wasn’t that many of them.

The review:

The comments through out the whole notebook is short, but at the same time manages to describe what they want to do and how they will achieve it. Very easy to read notebook as they have divided the cells into short segments with descriptive comments.

As far as I can see the use of different features from the Pandas library has been used to its fullest and this can be seen through the whole notebook. This makes each task and their cells very short. There is no space for comments in the code and the textual descriptions doesn’t go into detail of how things are done in the actual code. As one line of code that fulfils the objective of the task they are not that easy to read or understand, this can be seen through almost all of task B3.

The results they arrive to makes a lot of sense and reflect the work put down before arriving there. But I can’t help but to think that choosing to flip negative year to positive ones and removing the last digit in years with 5 to get a year with 4 digits in task A is the best way to deal with these anomalies. If they were really unlucky these data point might have polluted the dataset they had even more. Same goes for task B where nans are not removed but filled instead. As I remember there weren’t that many of them so removing them might have been the better option.

This is a group that does more than is asked of them, see the plot of locations of the different stations I task B, and when I read through their work I can’t help but get the feeling of that they know what they are doing.

Grades:

* Textual description: 6
* Code quality: 6
* Results: 6