**The development of the resulting visualisations**

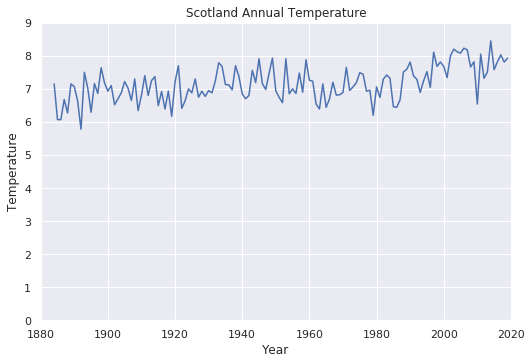
Instroduction

Due to global warming has been getting worse since the last few decades, I try to find out what influence it brings to the snowline in Scotland by visualising the temperature (Spencer 2016) and the snowline data (Met Office 2020). My assumption is that the snowline of Scotland has become higher during the last couples of decades. To confirm the assumption, I asked myself 4 questions and answered these questions by analysing and visualising the data mainly using python with the packages of numpy, pandas, seaborn, matplotlib.pyplot, seaborn and plotly.

Body

* How the air temperature changed in Scotland from 1884 to 2019?

First, I needed to present the change of the temperature to see whether it increased or decreased and to what extent it changed. It is an appropriate way to use the mean temperature of each year to show the tendency of the temperature change. The prototype visualisation was illustrated by the line plot (Figure 1).

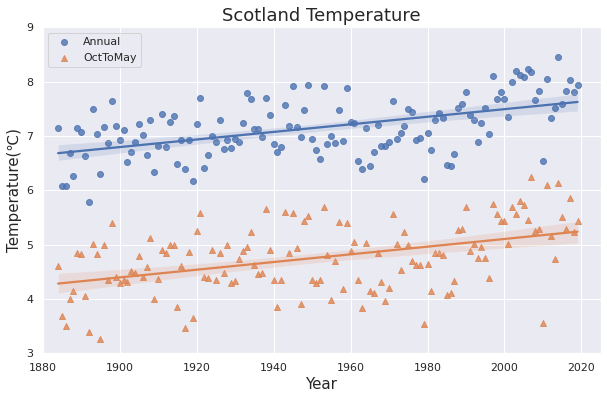
Figure 1. Prototype visualisation of Scotland annual temperature

As the data is related to the time, the first idea came up in my mind was the line plot. However, it seems even rather than any changes in the temperature. Then I improved this chart with three improvements (Figure 2).

1. Changing the line plot into the scatter plot with a regression line.

There was no need to show the line among the data points because it should focus on the change from 1884 to 2019 instead of the change of each year. And adding the regression line can give explicit information to readers to perceive the tendency of the temperature change.

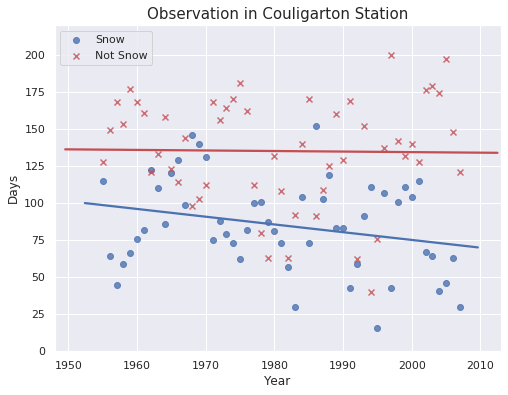
1. As the dataset of snowline was recorded in winter, usually from October to May (ref), I added this dataset in the same diagram to show how the temperature changed in the same period.
2. I changed the plot with the y-axis started from 3 rather than 0. Because, in these two datasets, no single figure of temperature is under 3 and the total change from 1884 to 2019 is not huge. Make the y-axis started from 3 can higher the “data-ink ratio” (Alberto 2012, pp. 64-66) with a bit distortion not misleading readers, which is a better way to present these data.

Figure 2. Improved visualisation of Scotland annual temperature.

After understanding the temperature had increased from Figure 2, I kept answering three more further questions. But first of all, I had to analyse some stations’ data. My initial idea was to choose the stations with about 40-50 years’ records (Spencer 2016, p. 7) in order to come out with a more precise analysis result. But the official document does not label the names of those stations. Given that, I chose the top 3 biggest size record files, which are Couligarton, ForrestLodge and LochVennachar. The following visualisations take the Couligarton station’s data for example.

* Did the change of the temperature influence the snow day percentage?

The prototype question was to study the influence on the number of snow day. However, I found there was something wrong in this analysis after generate the scatter plot (Figure 3).

 Figure 3. Prototype of the second question.

As we can see the Figure 3, the number of snow day had decreased while the number of no snow day did not change. This is not reasonable and it suggests that the reason may be the population of each year’s data is different. Then I tried to solve this problem with the percentage of the snow day and no snow day (Figure 4).

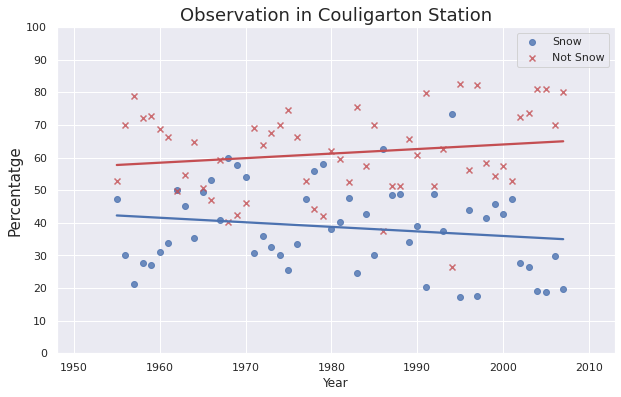
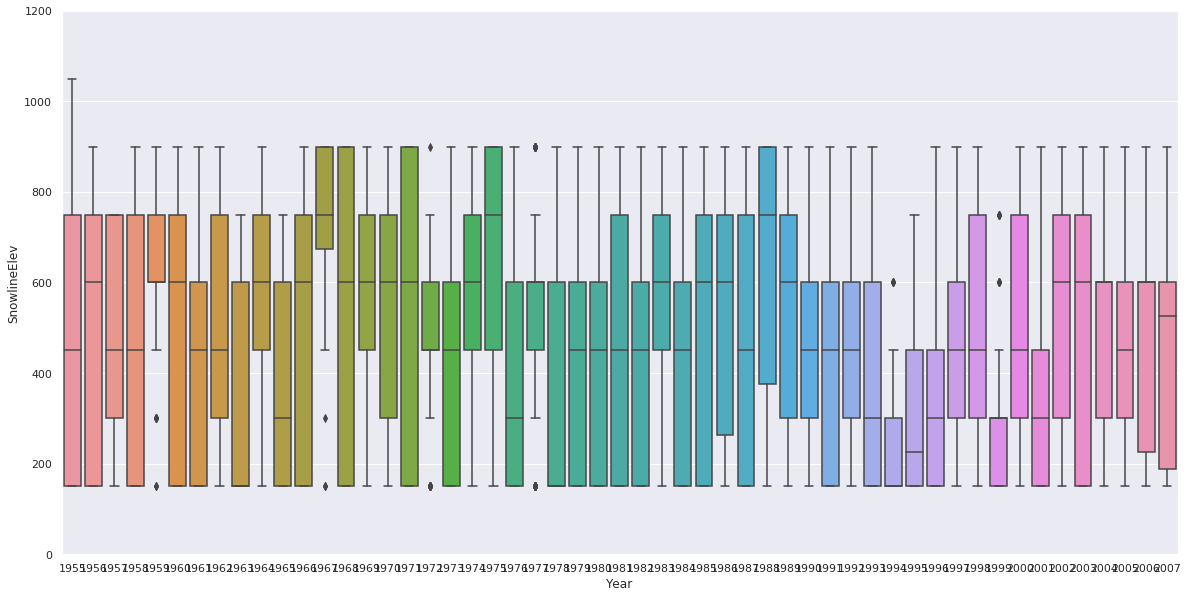
 Figure 4. Improved visualisation of the second question.

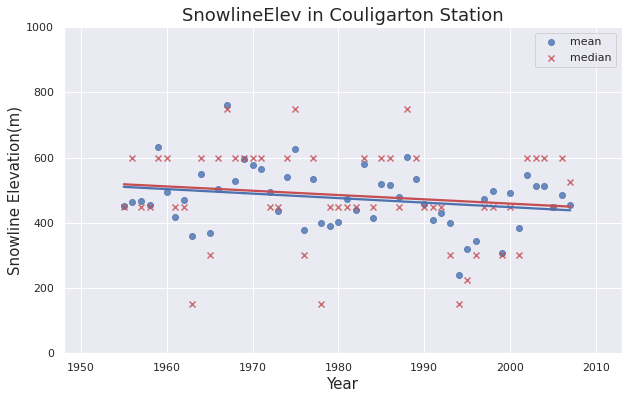
Figure 4 illustrates clearly to the readers that the snow day percentage had been decreasing between 1955 and 2007. It may suggest that the increasing temperature have an impact on the percentage of the snow day.

* Whether the snowline was higher or not?

I used box plot (Figure 5) as the prototype visualisation for this question because I considered that the box plot can show the distribution and the median of the snowline to let readers understand the difference and the tendency of the snowline among years.

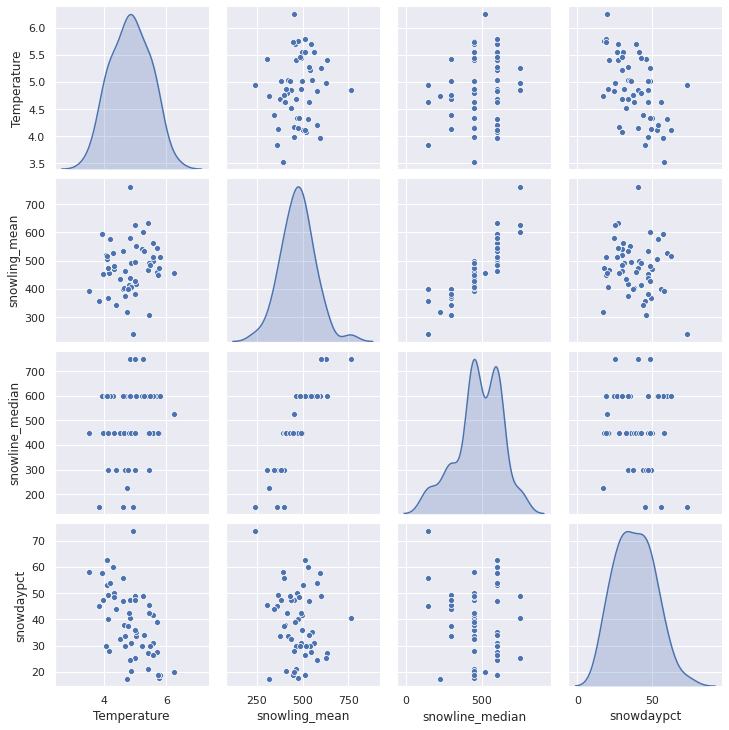
Figure 5. Prototype of the third question.

It seems that there is little difference in the distribution of each year’s snowline dataset because they are not exactly continuous data, grouped into 150 m bands from 0 to 1200 m (Spencer 2016, p. 2). Also, it is hard to perceive the change of the snowline by reading the median figures in this diagram. After that, I presented the data with scatter plot (Figure 6), which can tell readers the snowline had been lower in a few seconds.

Figure 6. Improved visualisation of the third question.

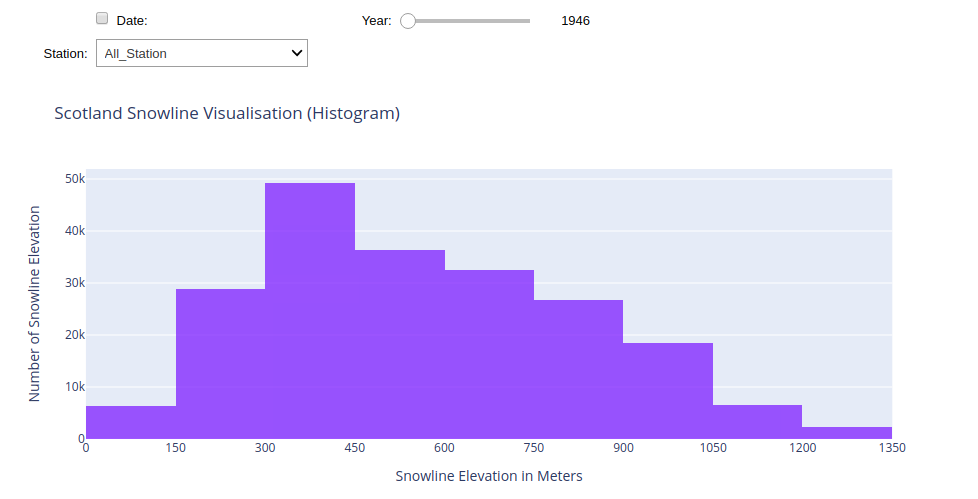
* Is there a correlation between temperature and snowline data?

I generated a pairplot to see if there is a correlation between temperature and snowline (Figure 7).

Figure 7. The visualisation of the fourth question.

As we can see above, the data points are distributed in a circle, which means there is no correlation between temperature and snowline. But, the diagram shows that there is a weak minus correlation between temperature and the percentage of the snow day.

At the end of the analysis, I made an interactive histogram (Figure 8) to give readers an opportunity to explore further information in snowline distribution.

Figure 8. Interactive histogram of snowline elevation in meters

Conclusion

After analysing and visualising the data, I found that there is a correlation between the increasing temperature and the decreasing percentage of the snow day, instead of the change of snowline.

Reference:

Alberto, C. 2012. The Functional Art: An introduction to information graphics and visualization [Website version]. America: New Riders. Available at: https://learning.oreilly.com/library/view/the-functional-art/9780133041187/ch06.html#page\_114 [Accessed: 8 May 2020].

Met Office, 2020. UK and regional series. Available at: https://www.metoffice.gov.uk/research/climate/maps-and-data/uk-and-regional-series [Accessed: 17 March 2020].

Spencer, M. 2016. Snow Survey of Great Britain: transcribed data for Scotland, 1945 to 2007 NERC Environmental Information Data Centre. https://doi.org/10.5285/caf989a5-82d7-4db7-b6ff-c0475fdae07e