EXERCISE 1.2*: Ethics and Direction of Machine Learning Programs*

The ClimateWins dataset might not look very biased at first glance, but it is always wise to be cautious and analyze it thoroughly to ensure it is exempt from these.

Regarding PII, the set does not contain this type of information so I believe it's safe to say it is free of bias in this matter.

We have information on weather reports across different points of Europe, this leads to the first questions in relation to whether there are any biases; are all of these weather stations equipped with new updated technology? It can be the case that governments support these matters with different budgets, which can lead to some weather stations having less precise technology. This problem accentuates when we take into consideration that we have data starting around 1800. These 200 years have gone very differently for countries across europe and there is a big probability that some regions developed their technology greater than others, which can ultimately lead to bias in the data collection process.

Analyzing the different variables in the dataset I realized that not all weather stations measure the same aspects of climate, for example the Budapest station doesn't measure wind speed and snow depth while the Dusseldorf one does. This can lead to some of these factors not being taken into account when making decisions about where weather conditions might worsen and cause harm.