Explanations about why I changed the subject and dataset.

Due to the problems in the first dataset I chose (Kaggle stress detection dataset), I chose to change the subject and the datasets. Firstly, I knew from the beginning that it was most probably a synthetic dataset (it had no source in Kaggle) and I was aware of some of the problems (stated in the report). But when I discovered that the dataset had 30 different scores per person in the Big Five personality test results (this should be a constant across at least 30 consecutive days) and the range did not coincide with what is usually used, I had a very hard time analysing the result in a meaningful manner. Therefore I chose a dataset with real world data.