The first part of the assignment was pretty straightforward as I have experience in using jupyter notebooks for exploratory data analysis.

For this part I loaded and inspected the data using pandas. Then made histograms for each column using matplotlib. For plotting all the columns together, I used MinMaxScaler from scikit learn and made a line plot using matplotlib.

The biggest challenge I faced in part one was in creating the pdf, as pandoc was not installed on my system. This was easily fixed as it’s a common problem when creating pdfs from jupyter notebook.

The second part of the assignment was a bit more challenging as I have not used Streamlit before. I started with creating an app only displaying “hello world” in the browser. Then I created multiple pages for the app using a folder structure and added page navigation. I found that the page navigation didn’t work very well with an app with folder structure, so I changed to using functions to define the pages of the app. I then created one python script for each page which I import in the main streamlit\_app.py file. For me, this way to structure the app made the most sense. Then I filled the pages with the content as described in the task, spending most time on creating the third page. The last thing I did was to implement the caching of the dataframe. I tried to do this in the streamlit\_app.py file but was not able to make this work properly. I then chose to make a script for loading the data with containing a function for loading the csv-file into a dataframe using the @st.caching decorator. I think this will work for caching.

Use of AI: I used ChatGPT for the for-loop creating plots for the individual plots, and for the tip of using MinMaxScaler. In addition, I used ChatGPT for the altair code in the streamlit app.