Question Python:

① Describe in your own words why Python is used for many machine learning tasks and describe the four major libraries used in machine learning.

② Which two libraries are being used to store the data, what are the differences and when do we use them?

① - Python is one of the most popular programming languages for data-science. It is open-source software and offers a wide community of helpers. Further Python offers many extension libraries which have been build upon lower layer Fortan and C implementations for fast and vectorized operations on multidimensional arrays.

- For ML programming tasks, mostly scikit-learn library is used. scikit comes with a lot of algorithms to be applied on a dataset.
- Numpy: This is a package to deal with numerical data of the same type
- Pandas: A python package to deal with data organised as labels.
- Matplotlib/Seaborn: Packages to deal with plotting data.

② The data load from the csv file will mostly saved on a python Dataframe. Pandas offers good ways to handle the data and to manipulate them for analysis. Numpy is especially popular for Multidimensional arrays and Matrices with lots of Mathematical functions. Numpy is also the fastest option to iterate through loops.