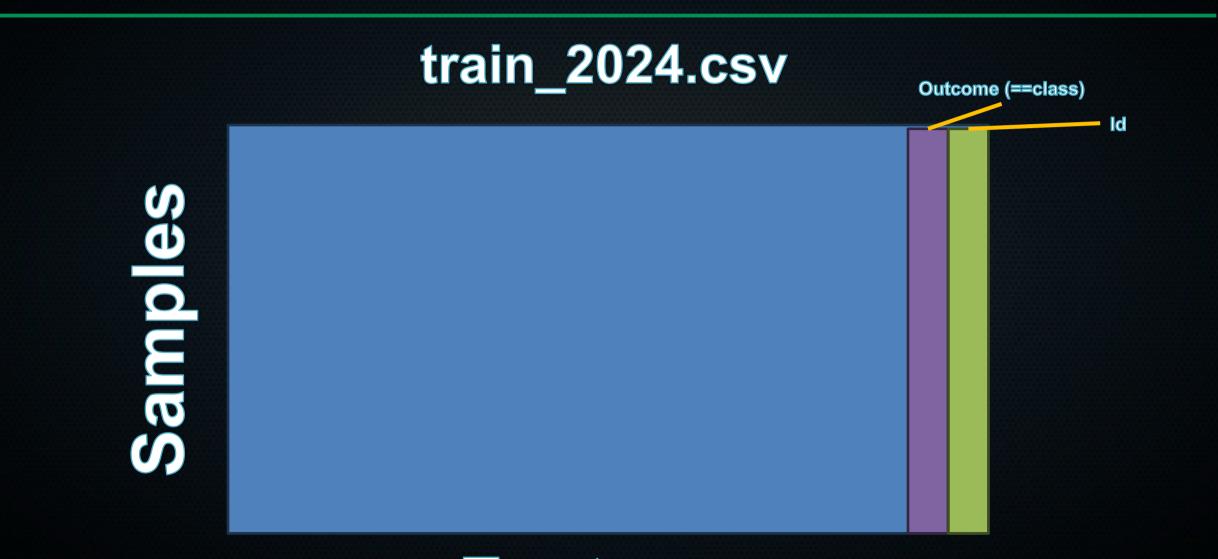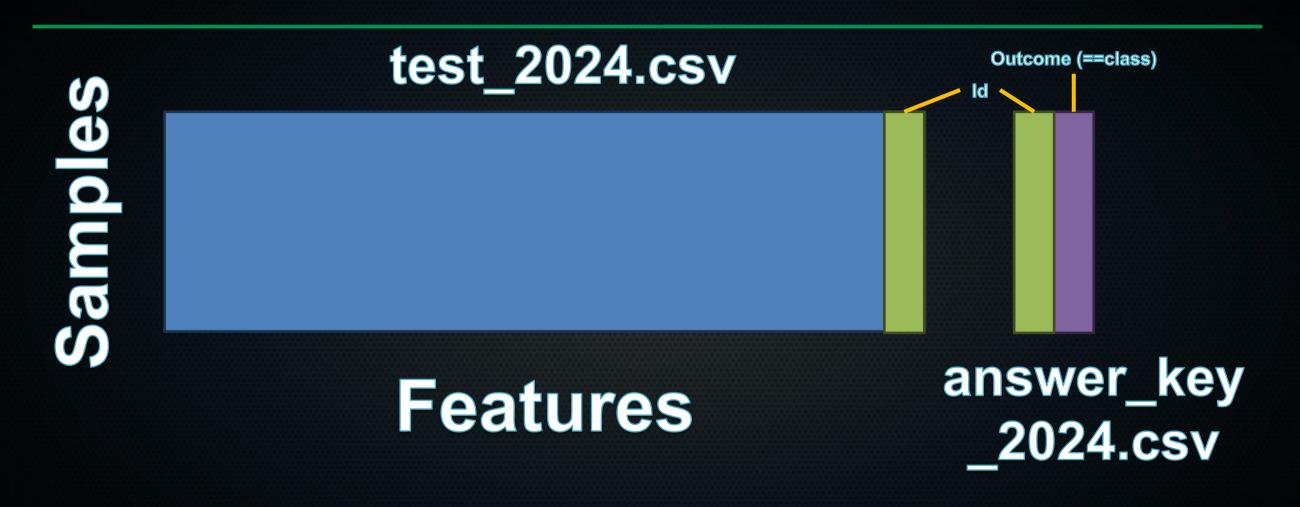# Project introduction

- Work in teams of 2 (working alone allowed, if one person is left out then a group of 3 is allowed as well)
- Follow the instructions in the Project folder of the GitHub repository (.docx file). Data available from 12:00 today.
- Briefly:
  - 10 guided sklearn-based exercises on a dataset. Building different classifiers, doing proper cross-val, etc.
  - Make one final classifier that does something different: more hyperparam optim. ; different algorithm ; ???
- Only at the very end test a few of the classifiers on a held-out test set (answer_key_2024.csv; test_2024.csv)
- Hand in Jupyter Notebook by mail, at the latest Wednesday the 13th at 15:00

Project introduction

Some details

- Don't care about proper naming or very good organisation (especially in the age of ChatGPT or CodeStar of Claude 3 rewrites)
- *Do care* that I can understand it and that what should be in the Jupyter notebook is in there. This is all explicitly stated.
- If there are problems with the Keras wrappers (SciKeras or installing TensorFlow seeing as Keras 3 is quite new) then Question 10 will be skipped. Especially if many have issues. However, for now, this is on a per-team basis and I need to be e-mailed about the difficulties encountered.

Some details

- I will not be here tomorrow as I am paranymph at a PhD defence. There's three avenues of getting help at your disposal:
- - E-mail me at dieter_stoker@hotmail.com. I can respond in the afternoon.
  - Tomorrow from 13:15-17:00 Marc will be there in Daltonlaan 500 to answer questions. He is 1. brilliant and 2. has the answers so all in all that should be helpful.
  - Ask Marc for my phone number and SMS me in dire circumstances. No WhatsApp though. Old-school, baby.

Some details

- Want to work together/independently and merge using GitHub? Great! Be sure to install nbdev as instructed.
- I used to use Kaggle for this. If you find mention of it: disregard it. You just hand in the notebook to me.
- Any questions?