Basic Spark Exercise

**First, make sure that you download the spark docker using the docker-compose.yml.**

Then, follow the exercise [here](https://towardsdatascience.com/six-spark-exercises-to-rule-them-all-242445b24565).

Try to download the [file needed for the exercise](https://drive.google.com/file/d/1kCXnIeoPT6p9kS_ANJ0mmpxlfDwK1yio/view?usp=sharing).

If it fails for some reason (it is 6GB), I prepared a subset of the dataset that you should download from *Teams*, instead of the giant 6GB Zip file in the original exercise.