Exercise 6 - deadline 7.05.2021 23:59

The dataset for this (and the following exercises) is the Penn World Table dataset.
You can find the latest version of the dataset here:
https://www.rug.nl/ggdc/productivity/pwt/?lang=en

The goal for this exercise is to create and analyse a grouping of countries for two different years. Analyse the grouping and the evolution between the two years.

Exercise 6 continues where 5 left off. Adapt the clustering technique for one (or both) years using the following changes.

The goals of the exercise:
1. Use PCA (Principal Component Analysis) to reduce the number of features. Pay special attention to the need of normalization.
2. Interpret the results for one year using a decision tree. (with the original fields)
3. Compare the results to the clusters from exercise 5.

References:
https://scikit-learn.org/stable/modules/clustering.html

**Important Note:**
The submission should be done individually. Please name your script/notebook accordingly. At the beginning of your script / notebook give the name(s) of all students who worked on the exercise. You may work in small groups (of 2-3) but all students must be confident with the work of others. They should understand and be able to answer detailed questions of the implementation.