Additional Informations:

-There are two datasets which are created when running the notebook:

# expanded\_data1.plz.gz

# undersampled\_data

If you want to test the creating process of the two datasets you need to delete both of them. Otherwise you do not have to do anything because the code for the creation automatically gets skipped if both of them already exist.