

Please carefully read and follow the general instructions regarding coding assignments. Failing to meet the requirements might lead to penalties. <https://elearn.uef.fi/mod/page/view.php?id=248672>

If you suspect that something is wrong with some task instructions, please contact the lecturer.

If you face persistent issues while working on a task, do ask for help, e.g. during a course meeting or by contacting the lecturer via email.

Task 1. The aim is to apply methods studied in the course to analyse a (spatio-)temporal dataset of your choice.

Choose a (spatio-)temporal dataset, preferably multivariate time-series.

You can download a dataset from a data repository, such as

- the *UCI Machine Learning repository* (<https://archive.ics.uci.edu/>)
- *Kaggle* (<https://www.kaggle.com/datasets>) e.g. Commuter train timetable, Cargo 2000, Parties in NY, etc. access to the datasets requires registration
- the *Finnish Meteorological Institute* (<https://en.ilmatieteenlaitos.fi/download-observations#!/>)
- some *Open Data initiative* (<https://data.worldbank.org/>, <https://www.opendata.fi>, etc.)

or from some other source relevant to a topic of interest to you.

Present the dataset, explain what the variables represent, provide basic statistics and show plots.

Apply methods studied in the course to analyse the data. You should implement at least one of the algorithms (such as, for instance, frequent subsequences mining, wavelet decomposition, dynamic time warping distance computation, HMM training, SAX discretization) or reuse your implementation from a previous assignment.

Explain the different steps of the analysis process, from acquiring the data, preparing it, choosing and applying the methods, to interpreting the results. Remember that the analysis process is often iterative, requiring for instance to make adjustments to the data preparation steps and rerun the subsequent steps.

Discuss what worked and what did not, what difficulties you encountered and what you learnt in the process.

Contact the course instructor if you need guidance regarding the choice of dataset and/or analysis methods or to confirm that they fulfill the requirements.