Group Assignment Instructions

International DAAD Summer School on Geospatial Data Science
Jena, August 2019

1. **Pick one of the data sets.** Refer to the Data Description file for information on available variables, their measurement units and possible research questions.
2. **Decide what aspect of (predictive) modelling you wish to focus on.** For example, one of the following:
    - Benchmark different statistical and machine-learning techniques in terms of their spatial cross-validation accuracy.
    - Explore the sensitivity of a model's predictive performance to its hyperparameters.
    - Implement a 'self-tuning' version of a machine-learning method such as SVM (i.e. implement an internal spatial cross-validation to identify optimal hyperparameter values).
    - Implement a feature selection method, and explore the sensitivity of a model's predictive performance to its hyperparameters (e.g. variable selection threshold).
    - Find out how sensitive model results (and computing time) are to sample size.
    - Visualize a machine-learning model, i.e. relationships between predictors and response. (Remember that you can only visualize up to three dimensional subspaces of the higher-dimensional feature space…)
3. **Create 2-3 slides briefly describing the objective, methods and results.**