Instructions – project work

- Choose a data set and make all the steps of Bayesian data analysis workflow listed below
- Project outcome is a Python or R notebook similar to notebooks in
  - BDA R demos https://github.com/avehtari/BDA_R_demos/tree/master/demos_rstan
  - BDA Python demos https://github.com/avehtari/BDA_py_demos/tree/master/demos_pystan
  - Stan case studies http://mc-stan.org/users/documentation/case-studies.html
  - StanCon case studies http://mc-stan.org/users/documentation/case-studies.html (some of these notebooks are for a bigger projects, but reflect still the basic idea of a notebook presentation)
- The submitted notebooks need to illustrate the knowledge of the Bayesian workflow.
- The notebooks have to include
  - Description of the data, and the analysis problem
  - Description of at least two models, for example:
    - non-hierarchical and hierarchical
    - linear and non-linear
    - variable selection with many models
  - Non-informative priors and description of the prior choices
  - Stan code
  - How Stan model is run
  - Convergence diagnostics (Rhat, divergences, ESS)
  - Posterior predictive checking
  - Model comparison (e.g. with loo)
  - Predictive performance assessment if applicable (e.g. classification accuracy)
  - Sensitivity analysis with respect to prior choices
  - Discussion of problems, and potential improvements

Peergrade rubric

Part of the questions are used to check that the minimal requirements of the project work are included. Most of the questions are for giving feedback to other students. The received feedback and your response to that will be discussed in the evaluation meeting.

- Can you open the notebook?
  - yes
  - no
- Is there an introduction?
  - There is no clear introduction
  - The introduction touches on the main topic
  - The introduction states the main topic and provides an overview of the notebook
  - The introduction is inviting, presents an overview of the notebook. Information is relevant and presented in a logical order.

- Do you have any suggestions on how to improve the introduction?
- Is there a conclusion?
  - There is no clear conclusion
  - A conclusion is included
  - The conclusion is clear

  Describe in your own words what is the main conclusion of the data analysis in this notebook?
- The structure and organization of the notebook
  - The notebook lacks a clear data analysis story
  - The notebook attempts to tell a coherent data analysis story but lacks some focus and clarity.
  - The notebook presents a clear cohesive data analysis story
  - The notebook presents a clear cohesive data analysis story, which is enjoyable to read
- Overall, what did you think of the structure and organization of the notebook? Name at least one way your peer could improve structure and organization.
- Accuracy of use of statistical terms
  - There are numerous errors in use statistical terms
  - There are some errors in use of statistical terms
  - Statistical terms are used accurately but sometimes lack clarity
  - Statistical terms are used accurately and with clarity
- Description of the data, and the analysis problem
  - yes
  - no
  - Did you get a sense of what is the data and the analysis problem when they were first introduced? Where and how might the author make the model description more clear?
- Are there more than one model
  - yes
  - no
  - Was it easy to find the list of the models?
- Description of the models
  - yes
  - no
  - Did you get a sense of what the models are? Where and how might the author make the model description more clear?

- Description of the prior choices
  - No priors or improper priors (e.g. uniform on unconstrained parameter) used
  - Priors listed but not justified
  - Priors are listed and justified
- Is Stan code included?
  - yes
  - no
- Is the code for how Stan model is run included?
  - yes
  - no
- Is Rhat convergence diagnostics included?
  - No
  - Yes, but no discussion what can be concluded from the shown Rhat values
  - Yes, with discussion what can be concluded from the shown Rhat values
- Are HMC specific convergence diagnostics (divergences, tree depth) included?
  - No
  - Yes, but no discussion what can be concluded from the shown values
  - Yes, with discussion what can be concluded from the shown values
- Is effective sample size diagnostic (ESS) included?
  - No
  - Yes, but no discussion what can be concluded from the shown values
  - Yes, with discussion what can be concluded from the shown values
- Is there posterior predictive checking?
  - No
  - Yes, but no discussion what can be concluded from the shown checks
  - Yes, with discussion what can be concluded from the shown checks
- Is there a discussion of problems and potential improvements ?
  - No
  - Yes
- Choose something you like about the notebook and explain why you like it.
- If you were to go back and redo your own notebook after reading this submission, what would you change?

- If the student were to complete this project work again, what could they change, to make it overall better?