Project review

- Team: ISZ_3
- Authors: Michał Święciło, Dawid Bogon
- Review by: Michał Sołek, Marek Krupa

1. Problem formulation [0-5 pts]:
   - Is the problem clearly stated

     Yes, the problem describes level of happiness in Somerville [1/1]
   - What is the point of creating model, are potential use cases defined

     Yes, the point and use cases are described [1/1]
   - Where do data come from, what does it contain

     Yes, data comes from website of City Somerville which contains many variables [1/1]
   - DAG has been drawn?

     Yes [1/1]
   - Confounding (pipe, fork, collider) were described

     Yes [1/1]
2. Data preprocessing [0-2 pts]:
   - Is preprocessing step clearly described

     Yes, all steps are briefly described [1/1]
   - Reasoning and types of actions taken on dataset have been described

     Yes, all data was cleared and sorted [1/1]
3. Model [0-4 pts]:
   - Are two different models specified

     Yes, ordinal logistic regression model with extension of more params [1/1]
   - Is difference between two models explained

     Yes [1/1]
   - Is the difference in the models justified (e.g. does adding additional parameter makes sense)

     Yes, adding more params lead to more accurate model [1/1]
   - Are models sufficiently described (what are formulas, what are parameters, what data are required)

     Yes, they are well described in model description [1/1]
4. Priors [0-4 pts]:
   - Is it explained why particular priors for parameters were selected

Yes, it is described [1/1]

- Have prior predictive checks been done for parameters (are parameters simulated from priors make sense)

  Yes, checked all the params [1/1]
- Have prior predictive checks been done for measurements (are measurements simulated from priors make sense)

  Yes, it is a partly overlap between real and simulated data [1/1]
- How prior parameters were selected

  Yes, they were selected logically due to the fact of lack of information [1/1]

5. Posterior analysis (model 1) [0-4 pts]:
   - Were there any issues with the sampling? If were what kind of ideas for mitigation were used

     Yes, there is a warning, but it is well described [1/1]
   - Are the samples from posterior predictive distribution analyzed

     Yes, they are described [1/1]
   - Are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided)

     Yes, they are partially consistent, but it is justified [1/1]
   - Have parameter marginal distributions been analyzed (histograms of individual parameters plus summaries, are they diffuse or concentrated, what can we say about values)

     Yes, there are histograms with descriptions [1/1]
6. Posterior analysis (model 2) [0-4 pts]:
   - Were there any issues with the sampling? If were what kind of ideas for mitigation were used

     No warnings or errors [1/1]
   - Are the samples from posterior predictive distribution analyzed

     Partially, they are plotted without description [0.5/1]
   - Are the data consistent with posterior predictive samples and is it sufficiently commented (if they are not then is the justification provided)

     Partially, data and simulated values are much more consistent without comments [0.5/1]
   - Have parameter marginal distributions been analyzed (histograms of individual parameters plus summaries, are they diffuse or concentrated, what can we say about values)

     Yes, there are histograms with descriptions [1/1]

7. Model comparison [0-4 pts]
    - Have models been compared using information criteria

      Yes, WAIC and LOO criteria [1/1]
    - Have result for WAIC been discussed (is there a clear winner, or is there an overlap, were there any warnings)

      Yes, it is briefly discussed [1/1]
    - Have result for PSIS-LOO been discussed (is there a clear winner, or is there an overlap, were there any warnings)

      Yes, it is briefly discussed [1/1]
    - What the model comparison discussed? Do authors agree with information criteria? Why in your opinion one model is better than another

      Yes, comparison is discussed. Model with additional param is better [1/1]

SUM: 26/27 -> 96,3%