

Final project

- ▶ This is a small gene expression data set for mouse. It contains the values of 24 genes (rows) in $n = 40$ samples (columns). The first column corresponds to the name of these 24 genes. In this data analysis, we will use the gene, Mapk1, as the response variable. The remaining $p = 23$ genes are used as the predictors.
- ▶ **Turn in written report + R programs by December 2 (2016), 6:00 pm via blackboard.** Please make sure to write down the names and Netid for all group members on your report.
- ▶ When your group submit the report via blackboard, I recommend you to upload the file only through one student's account. This makes me easier to know this is the right version of your group report, in case students in the same group may submit different versions. Again, students in the same group will receive the same score.
- ▶ The main report should be no more than 8 pages (can be less) in normal format with font size 12. Note that figures and tables that are used to support your conclusions should be part of the report and will count into 8 page limit. All the analysis should be done in R and all programs used for data analysis should also be submitted (no length limit for the R program).

Tasks

- ▶ Identify important genes and describe their association with the target gene Makp1. Find model that is as simple as possible but still has decent performance.
- ▶ As a minimum requirement, the report should contain the model and method you use, and interpretation of results and conclusions.

Grades based on:

- 1 whether the model and method address the problem
- 2 clarity of presentation
- 3 result interpretation
- 4 reproducibility of results
- 5 justify your method is more appropriate than the others for this data set
- 6 originality of approach

The project counts 50 points. If the above points 1-4 (mostly the minimum requirements) are correctly addressed, you will get 46 points. The remain 4 points will depend on whether your method can incorporate some aspects of points 5 and 6.

- ▶ Report that is late receives 20% off within 24 hours late and 0 after that. **There will be no exceptions for this rule!**
- ▶ You should not discuss the project with anyone not in your group. Good Luck!