

# ST537 – Project Instructions

Arnab Maity

## 1. Instructions

Please follow the instructions below when you prepare and submit your assignment.

- **Include a cover-page** with your project report. It should contain
  - (i) Project Title,
  - (ii) List of authors
  - (iii) Team number
- Assignments should be submitted **using moodle** by the due date specified.
- Neatly typed work should be submitted.
- Submission should be in the **PDF format**.
- The project should be your *own team's* work – no collaboration across teams is permitted.

## 2. Your assignment

Project report should be neatly typed and presented. Your group report should be *no more than 15 pages long* (with 1 inch margin, 12 point or larger Times New Roman font, 1.5 line space). You should include any plots, graphs, or tables that support your statements, but these should be referenced in the text, and placed at the end of the document. Each plot, graph and table should be numbered and must have a proper caption. Any cover page, figures, references and appendix are *not* counted towards the page limit.

Include the following sections in your report:

**1. Introduction:** This section should give a short description of the problem, a brief summary of the study carried out (include study design/sampling methods), mention each of the specific scientific questions you will investigate, an overview of the conclusions of the data analysis, and a short roadmap for the remainder of the report. You can *paraphrase* the materials at the beginning of Section 2 in here, if you wish. Also, include the reference of the the data source here.

**2. Methods:** This section should contain description of your statistical model, data analysis, and results - you may wish to divide the section into subsections. Start by providing visual and/or numeric summaries of the data (e.g., summary of relevant variables, are there any missing values and so on) and relevant discussion/interpretation and how they might impact your modeling choices. Clearly describe any data pre-processing you may have done (e.g., did you omit missing values and how many, did you standardize the data, and so on). *Do not expect the reader to read your raw R code and figure out what you did. It is your responsibility to clearly describe the steps you have taken to prepare the data for analysis.*

Next, clearly describe the mathematical model and/or methods you used to analyze your data. Be sure to define any symbols you use in your models. Specify all the assumptions required to implement your methods of choice. In most situations, there will be multiple methods/models available to fit (e.g., multiple classification methods or multiple choices for mean and covariance structures). Compare at least a few

different methods/models, and investigate their performance using appropriate criteria. Before analyzing the data, be sure to clearly specify how your scientific questions can be answered using the model parameters (e.g., “test for effect of age is equivalent to testing  $H_0 : \beta_2 = 0$  in model . . .”).

Finally, get back to each of the scientific questions you proposed in *Introduction* chapter, present the results for each of the problems, and *interpret* them accordingly. *Do not just say “we reject  $H_0 : \beta_2 = 0$ ”, instead describe this result in the context of the scientific problem.*

In preparing this section, provide all relevant details such as which fitting method you used, which testing procedure you employed, the software used and so on. **Do not** give raw data, raw code or raw output in this section. You may summarize your results using tables and/or figures, but do not just print the raw output.

Also, present possible limitations of your methods, if any. If this section is too long, you can break it into smaller subsections for each question you address.

**3. Conclusions:** This section is a brief summary of your study objectives and data analysis results in terms of the subject matter. Succinctly present the scientific questions again, your findings and interpretation. Discuss if your project goal were met, and if you expected outcome matched with your actual findings. If you have made any additional observations during your analysis, you may present them here as well.

**4. Appendix:** This section contains technical details and other supporting information, if any. All *well commented R code* and the corresponding detailed output (if any) can be included here in an organized format.

**References:** Put any references you may have here. You can omit this section if you do not have any references.