

# STAT 206 Final Project Guideline

*Fall 2015*

## Final Project Timeline

Nov. 25 (Wednesday) In Class	Project Proposal Due
December 7 (Monday) 12:00pm	Final Project Report Due

## About the Final Project

Your task is to conduct a start-to-finish statistical analysis of a real data set aided by knowledge and techniques learned in this class (or beyond). This includes formulating questions, building models, performing analysis and diagnostics, interpreting results and wrapping everything up in a paper format report.

## Project Proposal

On the Project Proposal, you should provide the information listed below.

### 1. Group Membership.

You are encouraged to form groups with no more than 3 members. You can also do an individual project, then the group will only have one member. **In the proposal, you need to list all group member(s) and their email address(es). Each group also need to indicate a contact person for future coordinations with the Professor and TA (Jilei Yang).**

### 2. Which Data to Analyze.

Each group need to choose a data set to analyze. I will upload several data sets on the smartsite (data sets and descriptions under “Resources/Project”), or you may use your own data (e.g., from your lab). In the latter case, make sure that the data from yourself is at least as complex (e.g., in terms of number of variables and number of cases) as the data sets provided by me. If you are not sure, you should come to my office hour or make an appointment with me to discuss the data set.

**In your proposal, you need to indicate which data you want to analyze.** If you will use a data set provided by me, you only need to indicate the data set name. If you will use your own data set, you need to provide a short description of your data, including:

- (i) Where do you obtain the data? What the data is about?
- (ii) The number of variables in the data set.
- (iii) Brief descriptions of each variable in the data set (e.g., What are they measuring? Are they categorical or quantitative variables?).
- (iv) The number of observations (cases) in your data set.

### 3. Questions of Interest.

**In the proposal, you should list several questions that you want to answer by analyzing this data set.**

### 4. Plan for Data Analysis.

**In the proposal, you should provide a brief plan for data analysis.** For example: How would you explore the data (e.g., graphs, summaries statistics) ? Which models are you considering? What are potential pitfalls and how would you deal with them? (You may change the actual analysis strategies in your final report.)

### 5. How to Submit The Proposal.

**Each group only need to submit one proposal with the information listed above, as well as a (tentative) project title. Each group need to submit both a hard copy (with the contact person's name highlighted) in class and an electronic copy through the dropbox (submitted by the contact person) on the smartsite by the proposal due date.**

For the hard copy, each member of the group should provide signature to acknowledge that they are aware of their group membership and agree to work together for the final project.

For the electronic copy, please use the name format, "STAT206–XX–YY–ZZ–Proposal", where XX, YY, ZZ stand for group member names (FirstName Last Name) with "XX" being the name of the contact person. For example, a group by John Smith and Jane Green with the John being the contact would have a name "STAT206–JohnSmith–JaneGreen–Proposal". The preferred file format is either a word document or a pdf file.

You will not be graded on your project proposal. It will be used to make sure you have a project of the appropriate scope.

## Project Report

### Page Limits

- The main text (not including appendices and references) of the report should be 4 to 6 pages.
- The report should be prepared using 12pt font, single spacing on US letter paper.
- The report should be prepared by a word processor such as *word* or *latex* rather than handwritten.

## Project Report Format

The project report should include the follows:

1. **Title:** The title of the project and the names of all group members and their email addresses.
2. **Abstract :** A brief summary of the project, questions of interest and findings.
3. **Introduction:** A description of (i) the relevant background of the topic; (ii) questions of interest; and (iii) the motivation of the project i.e., why other people should be interested. You should also briefly describe the data.
4. **Methods and Results:** A detailed description of your analysis, supported by numerical results, figures and tables (figures, tables and R outputs should be included in appendices, but being referred to in the main text as needed).

For example: Describe in details how you reached your final model and then write down the final model and the assumptions, parameters estimates and relevant statistics, diagnostics, etc. You should include an exploratory data analysis step to explore the data before conducting more sophisticated analysis. For each step of your analysis, you should provide a motivation (Why do you do it? What goals you intend to achieve by this analysis?), a justification (Is the analysis reasonable for the data and the goals?), and a description of results (What does the analysis tell you?).

5. **Conclusions and Discussion:** A brief summary of main findings and discussion.

For example: What do your results mean for the questions you were trying to answer? What is your conclusion? How general are your results, to what extent do they apply or not apply? What are the limitations of your analysis? What you might have done to address these limitations? And any other comments and discussions that you deem important.

6. **Appendices:** The following should be included in the appendices (do not include these in the main text):
  - **Appendix 1:** Figures and tables (if any).
  - **Appendix 2:** R Codes and outputs.
7. **References:** papers and other sources that being used in your report.

## How to Submit the Report

**Each group only need to submit one report which follows the above guidelines. Each group need to submit both a hard copy (with the contact person's name highlighted) and an electronic copy through the dropbox on the smartsite (submitted by the contact person) by the report due date.**

On the hard copy, each member of the group should provide signature to acknowledge that they have worked together on the final project and they have agreed to submit the report in its current form.

For the electronic copy, please use the name format, "STAT206-XX-YY-ZZ-Report", where XX, YY, ZZ, etc., stand for group member names (FirstName Last Name) with "XX" being the

name of the contact person. For example, a group by John Smith and Jane Green with John Smith being the contact would have a name “STAT206–JohnSmith–JaneGreen–Report”. The preferred file format is either a word document or a pdf file.