Regression Project

STAT40x3: Statistical Methods II

Each aspect of the course gives me the chance to evaluate your abilities in different aspects of this course. The Regression Project allows me to evaluate your ability to perform the entire quantitative process, from crafting a well-written research question to the final write-up.

The emphasis, however, will be on proper analysis of your data. This means you will need to determine the appropriate test(s), the assumptions of those tests, and that the assumptions are not violated by the data.

To make things easier on all of us, there are four checkpoints throughout the semester. These checkpoints are spaced throughout so that you should have enough time to complete the necessary activities with time to spare (to tidy things up).

Checkpoints

Checkpoint One	Hypothesis	February 15	5 points
Checkpoint Two	Data	February 22	5 points
Checkpoint Three	Analysis	March 22	10 points
Checkpoint Four	Paper	April 20	25 points
Checkpoint Five	Presentation	April 26	5 points

What follows are the Checkpoints explained in greater detail.

This Checkpoint is usually the first step in any research: Come up with an interesting research question. Research questions are always in the form of a question and are very broad in scope. They give a broad framing for your research.

The hypothesis is a proposed answer to the research question. Were this a full research project, you would use the Literature Review to create your research hypothesis. Note that the research hypothesis is not a null hypothesis, it is the alternative hypothesis.

The research hypothesis is a verbal statement that relates several variables. For this project, I want one dependent variable (continuous) and two independent variables (one continuous, one categorical).

Finally, since you will need to collect the data to test the null hypothesis, you also need to provide the source of this data. If the source is the Internet, provide a link. If you are collecting the data yourself, state this. Finally, I want to know the experimental design. Please explain how the original researcher collected the data.

Checklist

- ☐ Research question
- ☐ Research hypothesis
  - Three variables involved:
    - Dependent variable = continuous
    - Independent variable 1 = continuous
    - Independent variable 2 = discrete
- ☐ Source of data
- ☐ Experimental design

Grading

This Checkpoint will be graded solely on whether you provided all of the necessary parts.

This checkpoint consists of the data. The data needs to be in csv format and emailed to the professor by the due date.

Checklist

☐ Data in a spreadsheet (csv format) emailed to professor

Grading

This Checkpoint will be graded solely on whether you provided all of the necessary parts.

This Checkpoint consists of your analysis of the data in light of the hypothesis. Your analysis must be correct and coherent. All assumptions must be checked. Graphs should be included to illustrate your point. Code also needs to be included, as that allows me to determine if you performed the tests correctly.

This analysis, including the graphs and the code, needs to be handed in to the professor on (or before) the due date. Do not wait until the last minute. Work on this throughout the month, so that you (or we) can deal with issues other than time pressures and procrastination.

Checklist

- ☐ Tests
- ☐ Graphs
- ☐ Code

Grading

This Checkpoint will be graded solely on whether you provided all of the necessary parts.

This Checkpoint has you type up the entire project. There should be the following sections: Introduction, Research Question, Literature Review, Hypothesis, Analysis, Discussion, Conclusion. As usual, write the Conclusion and Introduction last.

The Literature Review section should link the Research Question to the Hypothesis section. It needs to be brief, only long enough to show that the Hypothesis logically answers the Research Question. This section should be about 3 paragraphs.

The Analysis section leads the reader through the analysis you performed, but emphasizing the logic behind the tests you performed and the graphs you made. It must be clear why the tests you ran answered the question, what assumptions the tests have, and how you appropriately tested the assumptions. This section needs to be as long as necessary for you to properly analyze your data in light of the null hypothesis.

The Discussion section actually answers the research question in light of the findings of the analysis (which dealt with the hypothesis). This section should be about 3 paragraphs in length.

Checklist

- ☐ Introduction
- ☐ Research Question
- ☐ Literature Review
- ☐ Hypothesis
- ☐ Analysis
- ☐ Discussion
- ☐ Conclusion

Grading

This Checkpoint will be graded based on how well you fulfill the requirements of each section (as described above). I focus more on how well you explain the analysis and how well you discuss the results than on other section, but all sections are important.

April 26 — Checkpoint Five — [5]

This Checkpoint is the presentation. You have five (5) minutes to present your project. After you present, you will need to answer any and all questions raised by the other class members (and by me). If you want to use PowerPoint, then you need to bring it to class that day.

Checklist

- ☐ Presentation
- ☐ PowerPoint slides (optional)

Grading

This Checkpoint will be graded on how well you present the most important aspects of your paper and how well you stay within the time limit.