Individual Project

Statistics 506, Fall 2018 (./index.html)

- About
- Due Dates
- Instructions
- Graded Elements

About

The final course project is intended as an opportunity for you to showcase what you've learned this semester. It is worth up to 100 points, however, you have the three options described below:

Options:

1. *Option 1.* Complete the course project as described. In this case each of the two research questions you pose and answer is worth 50 points and your final course grade is based on the percentage of points earned out of 1000 total points as described in the syllabus.

2. *Option 2.* Complete the course project as described, with the exception that you pose and answer only a single research question. In this case, your project is worth 50 points and your final course grade is based on the percentage of points earned out of 950 total points.

3. *Option 3.* Pass on the course project. In this case, your final course grade is based on the percentage of points earned out of 900 total points.

You will need to commit to one of the options above by Thursday Dec 13.

Due Dates

1. *Thursday December 13, 10am* - Respond to the survey on Canvas to elect one of the options above. If election option 1 or 2, you will also need to state your research questions though some modification is acceptable as you work on the solution.

2. *Sunday December 16, 12 noon* - Submit the course project as a pdf to canvas.

Instructions

The data for this project is related to that used for Problem Set 4, Question 3. The Medicare provider utilization and payment data is available here (https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/Medicare-Provider-Charge-Data/Physician-and-Other-Supplier.html). You may make use of any of the available years, 2012-2016. You may also wish to make use of the Medicare Enrollment numbers by state, available under the 'Downloads' section for each year.

For the project, you should pose *two* clearly-defined research questions relating to the public use medicare data sets above. You will then answer each question yourself, using two or more data files to support your arguments. You must use *at least two* data sets in your solutions if electing option 1. If electing option 2, you may use a single data set. For option 1, you may either pose two questions related to the same group of two or more of these data sets, or pose one question based on a single data set and a second on another. The second question may not be a repeat of the first applied to a second data set.

Your questions should be limited in their overall scope, similar to a single question from one of the problem sets. Your answer to each question - including any text, tables, and graphs - should be a single typed page, submitted as a pdf via Canvas.

This is an individual project and you should not purposefully pose the same question as another student. You may, however, consult with your peers regarding techniques or presentation details.

You should also submit the code you write to answer your questions. Your code can use any of the software packages we have learned this semester. Code should follow the style guidelines and be reasonably concise and efficient.

Your report should be divided into four sections:

- Introduction: Clearly state the research question.

- Methods: Describe the data elements and statistical techniques used.

- Results: Report your findings and provide evidence in the form of graphs, tables, and factual statements that support your answer to the question.

- Discussion: Summarize your key findings as they relate to the question you posed in the introduction.

Graded Elements

Each of the following elements will be graded on a 10 point scale.

1. Quality and challenge of question posed.
2. Appropriateness of statistical techniques used.
3. Relevance of results to the research question posed.
4. Clarity of exposition and professional appearance of final report.
5. Coding style and readability.