



CTL Featured Assignment Form

Your name & email: Priya Kohli, pkohli@conncoll.edu

Name & number of course: Advanced Regression Techniques, MAT 207.

Name of assignment: Data Analysis Project

Class size: 27

Last time class was taught & when it will be taught again: Fall 2014, Fall 2016.

Description of the assignment (e.g., when you give the assignment, if you scaffold the assignment):

The purpose of this assignment is to give students hands-on experience for data analysis in an applied course like regression. On the first day of class we discuss the major requirements for the project as mentioned in the syllabus. During the first week of classes students form a group of three to work together on a collaborative project for the semester. Each group then selects a topic, either from the list of resources provided in the class or other areas that the students have interest in. The range of topics that students have worked on includes diverse areas like biology, social sciences, economics, environmental issues, physics, and hydrology among others. At the beginning of second week each group sends me their selected topic with the related dataset and I go over them to provide students feedback about the depth and relevance of the topic as per the project requirements. By the end of second week each group has a final topic for their project. At this point, I ask students to submit a small description of their topic along with the major objectives (mostly as one of the homework problems). This allows students to think about the questions they want to address through the analysis before actually doing any statistical analysis. The project then evolves throughout the semester as we learn class material which students apply to their group project. The assignment requires making three in-class presentations done by one member of group each time and submitting a final project report. These steps are hierarchical because students use what they learned from previous presentation (through analysis, comments, suggestions, and questions) as they work on the next step of analysis. After every presentation, I send the graded presentations to each group with comments and suggestions for improvements and students include these comments in their final report.

Presentation 1: Exploratory and Simple Linear Regression

Once the students have developed some background in basic regression they begin to explore the implications of using regression methods on their project. Usually at the end of first month, one member from each group makes a fifteen minutes presentation. Each group introduces their topic with background information and motivation for the choice of particular topic. Then they present a simple linear regression analysis using graphical and numerical techniques learned during the first month of class. Students present the summaries, show how the model performed, and also presents ways to assess the goodness of the model they fitted to the data. Because there are several groups, students get a chance to learn about many diverse applications of the methodology and they ask thoughtful questions from each other. This leads to a dynamic environment during the presentations which often continues in future lectures also. In future classes I relate the topics to the datasets that students are using for their group project.

Presentation 2: Multiple Linear Regression and Hypothesis Testing

The next step for students is to formulate a complete regression model for their project to present the overall fit and discuss the interpretation of their findings in the context of their topic. Another important aspect of this part of the project is variable selection in which students learn how to screen the important predictors from the set of available predictors to construct a parsimonious and statistically efficient model. Doing this requires fitting many different models and testing the statistical significance of regression coefficients before coming up with the final model. This makes students think critically and they always ask questions in class or office hours arising from their own experience while trying to find a good model.

Presentation 3: A Complete Regression Analysis for “chosen” Topic.

The final presentation is usually done in last week of classes. In this round the students have to show how their regression analyses have addressed the questions proposed at the beginning of the project. This involves key modeling steps like regression diagnostics, transformation, and testing of regression assumptions. At the end, each group provides a complete picture of their project with real applications arising from the objectives they have met through analysis.

On the last day of classes students also submit a final report explaining the work they have done on the topic throughout the semester and the report is usually restricted to approximately five pages. This allows students to learn how to present their findings in the form of a research paper and they also include the suggestions made in their presentations in this report.

Learning goals of the assignment & briefly how they relate to the goals of the course:

The major goal of this semester-long assignment is to make learning a two-way procedure, including transfer of information to students through lectures as a first step and then giving them a chance to “own” the subject material by applying what they have learned to a real-problem. This is crucial for true understanding of the data analysis techniques. Through this assignment, the students get an opportunity to be engaged critically with the course material by making connections to the life outside of the classroom through the diverse applications covered by different groups. The topics presented by different groups are mostly related to various different areas from both science and everyday life. It is very evident that more such connections the students make the more interested and curious they become in learning the material. This group-oriented project also provides an opportunity for students to learn how to work collaboratively and how to research background material to motivate others for their project. In addition, students learn how to cogently present their findings in a meaningful fashion using the facts they found which definitely enhance their critical engagement and understanding of the course material along with their interaction skills.

What you like about the assignment & why you think it is effective:

This assignment challenges students in many different ways, including choice of an interesting topic, finding relevant dataset for the selected area, researching about the background of the topic, learning how to apply the statistical techniques learned in class to a real problem, and then presenting the major findings meaningfully in a limited time framework. All of this requires that the students first understand the modeling techniques to be able to apply them to find useful results. One of the most delightful aspects is that the students rise to the challenge and they work very hard to find meaningful results in the context of the real-life application they are interested in and they take all the necessary help both inside and outside the classroom. I also believe that even though these projects do not always involve “true research” in the sense of discovering new results, but they are “research-like” in the engagement of students and serve as a first step towards the process of conducting research.