

Stat 216 Project Instructions

The purpose of this project is to give you experience through the six steps of a statistical investigation, from start to finish (see p. 2-3 in our textbook). It gives you the opportunity to explore something you are interested in and conduct statistical research in a collaborative team, similar to what you will experience in most of your careers!

How this factors into your overall grade
The project consists of a proposal, data collection, PowerPoint presentation, and final report, for a total of 15% of your grade broken down as follows:

<table>
<thead>
<tr>
<th>Project Component</th>
<th>% of Project Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project proposal (graded based on completion)</td>
<td>10</td>
</tr>
<tr>
<td>Excel data set (graded based on completion)</td>
<td>10</td>
</tr>
<tr>
<td>Project presentation</td>
<td>30</td>
</tr>
<tr>
<td>Project report and appendix</td>
<td>50</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
</tr>
</tbody>
</table>

You will work in self-selected groups of 2-4 people. All components of the project will be completed and turned in by your group (of up to four people), not individually.

Minimum requirements for project
The minimum requirements for the project are:

1. Your project should involve doing a short literature review to investigate and present prior research work on the topic to both motivate and put your research in context.
2. You must gather the data yourself using a reasonable sampling method or proper experimental design.
   - You should be actively involved in the data collection process. It would be nice for this to be some sort of an experiment, but it doesn’t have to be. If you choose to conduct an observational study, you should be actively engaging or observing the observational unit. For example, online surveys are not actively engaging or observing the observational unit. I have listed a number of project ideas later in this document.
   - If you plan on conducting research that involves direction interaction with human subjects (e.g., in person, mail, email, web, or telephone), all student researchers will need to complete research ethics training through the online CITI training through MSU’s Institutional Review Board website: [http://www.montana.edu/irb/getting_started.html](http://www.montana.edu/irb/getting_started.html). You must make sure your participants know their participation is voluntary.
3. You must report at least one visualization of your data using Tableau, both in the presentation and the report.
4. You must conduct at least two tests of significance using our applets. (They don’t need to be two different kinds of tests, but testing two different data sets---both related to your research topic.)
Timeline of due dates for the project
On or before 11pm on Friday, October 27: Submit a project proposal via D2L Assignments using guidelines below.
On or before 11pm on Friday November 17: You should have your data collected and the data file (in Excel format) submitted via D2L Assignments.
Last week of classes (Dec 4-8): Project presentations in class. Your group must submit your PowerPoint file via D2L Assignments prior to your presentation.
On or before 11pm on Monday December 11: Your final report submitted via D2L Assignments using guidelines below.

In your project proposal you should include:
1. The names of everyone in your group;
2. Your research questions;
3. What makes your research questions interesting or important;
4. How you plan to gather your data;
5. What variables you will collect;
6. What sample size you will try to obtain;
7. What statistical tests you will run.
The project proposal should be limited to at most two pages (12 pt font, single-spaced).

Before you collect your data!
As a first step to conducting your project you should write up an outline of what you plan to do. This should include your research question(s), your data gathering methods (how you plan to conduct the study, your proposed data analysis techniques (both descriptive and inferential), anticipated conclusions and limitations.

Practical Hint: A VERY common mistake on student final projects is to not clearly understand the connection between the data being gathered and the tests being run. For example, what type of data you collect (quantitative or categorical) impacts what analysis you will do. It is better to collect too much than too little.

Report Guidelines
Your paper should consist of no more than 6 pages of concisely written 12 pt font, single-spaced text (tables and graphs are included in the 6-page limit; however the appendix and data set are not). Note: Fewer than six pages is NOT a problem! Some research write-ups will need more room than others. The text should be organized into sections as follows:

1. Title and Authors
2. Abstract
3. Background and Significance
4. Methods
5. Results
6. Discussion/Conclusions
7. References
8. Appendix (a separate file submission and thus not included in the 6 page limit)

Each section should receive a heading.
1. Title and Authors

Give an informative title to your research project and provide your group members names.

Assessment: Does the title give an accurate preview of what the paper is about? Is it informative, specific and precise?

2. Abstract

The abstract provides a brief summary of the entire paper (background, methods, results and conclusions) in no more than 200 words. This allows you approximately 1 sentence (and likely no more than two sentences) summarizing each of the following sections. Typically, abstracts are the last thing you write.

Assessment: Are the main points of the paper described clearly and succinctly?

3. Background and significance

In this section you are providing the background of the research area and arguing why it is interesting and significant. This section relies heavily on literature review (prior research done in this area and facts that argue why the research is important). This whole section should provide the necessary background leading up to a presentation (in the last few sentences of this section) of the research hypotheses that you will be testing in your study. Well-accepted facts and/or referenced statements should serve as the majority of content of this section. Typically, the background and significance section starts very broad and moves towards the specific area/hypotheses you are testing. You should cite your sources in this section. For example, Roadrunners are known for their speed and ability to avoid allurements (Coyote, W. E. 1967).

Assessment:
- Does the background and significance have a logical organization? Does it move from the general to the specific?
- Has sufficient background been provided to understand the paper? How does this work relate to other work in the scientific literature?
- Has a reasonable explanation been given for why the research was done? Why is the work important? Why is it relevant?
- Is the final paragraph a brief description of the hypothesis/goals of the paper?

4. Methods

a. Data collection. Explain how the data was collected/experiment was conducted. Additionally, you should provide information on the individuals who participated to assess representativeness. Non-response rates and other relevant data collection details should be mentioned here if they are an issue. However, you should not discuss the impact of these issues here---save that for the limitations section. Make sure you include enough information in this section so I understand exactly what you did to collect your data.

b. Variable creation. Detail the variables in your analysis and how they are defined (if necessary). For example, if you combined items to create a variable you should describe how.
c. Analytic Methods. Explain the statistical procedures that will be used to analyze your data. E.g. Boxplots are used to illustrate differences in GPA across gender and class standing. Correlations are used to assess the impacts of gender and class standing on GPA.

Assessment: Could the study be repeated based on the information given here? Is the material organized into logical categories (like the ones above)?

5. Results

Typically, results sections start with descriptive statistics, e.g., what percent of the sample is male/female, what is the mean GPA overall, in the different groups, etc. You should include at least one Tableau data visualization here to illustrate these differences. Information presented must be relevant in helping to answer the research question(s) of interest. Typically, inferential statistics (tests of significance and confidence intervals) come next. **DO NOT GIVE SOFTWARE OUTPUT OR APPLET SCREENSHOTS HERE UNLESS THEY ARE APPROPRIATE GRAPHS!** This should look like a peer-reviewed journal article results section. **Tables and figures should be labeled, embedded in the text, and referenced appropriately.** The results section typically makes for fairly dry reading. It does not explain the impact of the findings, it merely highlights and reports statistical information.

Assessment:
- Is the content appropriate for a results section? Is there a clear description of the results?
- Are the results/data analyzed well? Given the data in each figure/table is the interpretation accurate and logical? Is the analysis of the data thorough (anything ignored?)
- Are the figures/tables appropriate for the data being discussed? Are the figure legends and titles clear and concise?

6. Discussion/Conclusions

Restate your objective and draw connections between your analyses and objective. In other words, how did (or didn't) you answer/address your objective. Place these all in the larger scope of previous research on your topic (i.e. what you found from the literature review), that is, how do your findings help the field move forward? Talk about the limitations of your findings and possible areas for future research to better investigate your research question. End with a concluding sentence or two that summarizes your key findings and impact on the field.

Assessment:
- Does the author clearly state whether the results answer the question (support or disprove the hypothesis)?
- Were specific data cited from the results to support each interpretation? Does the author clearly articulate the basis for supporting or rejecting each hypothesis?
- Does the author adequately relate the results of the current work to previous research?

7. References

Assessment: Are the references appropriate and of adequate quality? Are the references cited properly (both in the text and at the end of the paper)?
8. Appendix

In this section you should give a detailed description of what you did (statistically) and how you did it. NOTE: This section is NOT included in the 6 page limit; it can be a separate file submission. Give your formal null and alternative hypotheses here and not in the body of the paper. (Your paper should have your research question and a conclusion which is essentially your alternative hypothesis.) Describe how you used your software or applets to do the analysis, the details of variable creation and any other information about how you conducted the study that may be important that isn’t already in the paper. Software output or applet screenshots will make up a significant portion of this section. THIS (Not the body of the report) IS THE SECTION WHERE YOU CANNOT HAVE TOO MUCH DETAIL! Explain and show me why you did what you did. I should be able to recreate your analysis exactly based on I see in this section.

GRADING

Overall, you will be assessed on the areas outlined above listed as “Assessment”. Additionally, you will be assessed on the following areas with regards to writing quality:

1. Is the paper well-organized? (Paragraphs organized in a logical manner)
2. Is each paragraph well-written? (Clear topic sentence, single major point)
3. Is the paper generally well-written? (Good use of language, sentence structure)
4. Are tables/figures labeled correctly and referenced accordingly
5. Does the entire paper flow and answer the research question(s) sufficiently? Is there extraneous information presented?

Presentation Guidelines

Each group will have 5 minutes to present their project. Every member of the group is expected to speak during the presentation.

Your presentation slides should be a very short, bulleted version of your report. You should have pretty much the same sections:

1. Title and Authors
3. Background and Significance
4. Methods
5. Results
6. Discussion/Conclusions
7. References

Don’t gloss over the descriptive statistics and the methods of your data collection. They are an important part of telling the story. **Include at least one Tableau data visualization!**
Project Ideas

Your project idea should be one that is academic in nature and one you can find prior research about. It can be a replication of a previous study or something totally original. Most likely it should be one that answers a small part of some bigger question.

In this course, you have the tools to conduct inference on the following scenarios:

- One quantitative response variable and one binary explanatory variable (paired or independent).
- One quantitative response variable and one quantitative explanatory variable.
- One binary response variable and one binary explanatory variable.

This does not mean you shouldn’t collect more than two variables on each observational unit or that your categorical variables need to be binary! Data visualizations can include any number of variables. It just limits your confidence intervals and hypothesis tests to questions involving at most two variables.

There will be some links to apps on our Stat 216 webpage that you could use in your study: reaction time apps, memorization apps, and a brain test.

The following are project ideas from past projects – the research questions below cannot be used for your project; your group must come up with your own research question.

- What is associated with brain lateralization? Gender, handiness, math ability, etc.
- Can people memorize objects better if they are shown the objects in pictures or in words?
- Did you just cut me off? See if a person reacts to being cut in line by a specific gender and whether males and females respond differently.
- Do age, gender, major, distraction, etc. affect memory?
- I need to make a call. Find out if male or female students are more likely to lend a phone to a stranger. Does saying “please” help?
- Does listening to different types of music affect a person’s reaction time.
- Do Energy Drinks give you energy? Does consumption of an energy drink allow students to complete more sit ups?
- Does what one consider a serving size differ with gender, age, or athlete/non-athlete.
- Can people accurately estimate the length of a short event like a song snippet?
- Size estimates. Do people estimate the amount of a food item differently if they are given a small spoon or a large spoon, a small bowl or a large bowl, etc.?
- Do people like the taste of water better if they think it costs more?
- Does the amount of candy in a bowl effect whether or not people obey the sign that says take only one?
- Do men and women differ on whether or not they tell someone they have some food on their face?
- Is age related to the time needed to complete and the number wrong in a test to investigate the Stroop effect (color and words)?
- What factors are related to people being able to follow directions when given a short quiz?
When deciding on a project topic, think of one that is of interest to you and will be something that could contribute to some broader topic.

**Bad:** *Are the proportion of female customers larger at Café M or Sola (two coffee shops close to campus)?* Would anyone really care about this? Does this add to some broader topic? Would this ever be published in a national journal?

**A bit better:** *Will regular coffee drinkers complete a puzzle faster than non-coffee drinkers?* This, as you might guess, would be loaded with confounding variables.

**Perhaps a bit better:** *Do gamers have a faster reaction time than non-gamers?* While this one will also have confounding variables, it would seem that gamers should have a better reaction time since that is probably a big part of what they do when playing games. To make this even better you could control for other variables, like a person’s sex.

**Even better:** *Does bowl size affect how much cereal someone will serve themselves?* Now we have a randomized experiment and have the potential to determine cause and effect---this is, of course, assuming it is a nicely controlled randomized experiment.

**How to Work Well as a Team**

It is expected that each team member will contribute equally to the project. All members of the team receive the same grade. If you are having difficulty with a team member who does not contribute (doesn’t show up or is late to meetings, doesn’t offer ideas, etc.), speak to your instructor early in the process. A google search of “how to work well as a team” produces a plethora of suggestions, but some of the top ones are:

- **Set up regular meetings and deadlines (and meet those deadlines!)** where each team member’s contribution is well-defined during the meeting.
- **At each meeting, each team member should have a role, and roles should rotate between team members each meeting:**
  - The “leader” sets the agenda for the meeting and leads the discussion.
  - The “facilitator” keeps the group focused, assures all participate and that none dominate, and intermittently tries to summarize the discussion.
  - The “recorder” is responsible for recording what was discussed in the meeting and then emailing “minutes” from the meeting.
  - The “timekeeper” helps the facilitator and leader move through the agenda and keep the meeting in the set time.
- **Keep communication open, honest, and respectful.**
- **Appreciate that each team member is a unique person from a different background with different viewpoints.** Take advantage of this diversity and allow every team member to contribute.