

Secondary Data Analysis Project

Jonathan Gallimore, Embry-Riddle Aeronautical University, Prescott Campus
SF 420

Summary: This activity is designed to give students an opportunity to apply what they have learned in statistics to a real dataset.

Learning Goal: This activity will help students apply what they have learned in statistics to real world data and answer their own research questions. Students will also practice reporting their results in a paper using APA format.

Resources Used:

1. This activity uses the worksheet on page 2, “Secondary Data Analysis Project.”
2. The class will need access to a computer lab with some type of analytical software like SPSS or SAS. If all else fails, Excel can be used but requires adding the “Analysis Toolpak” add-in.
3. Instructors need to find and provide students with step-by-step instructions for the chosen program (e.g., SPSS or Excel). Web searches and YouTube are great for statistical software guides.
4. Use a Learning Management System (LMS) like Canvas, Blackboard, or Angel to submit assignments. Email can be used as a last resort.

Teaching Notes:

1. Think about spacing the due dates for each assignment so that there is time to give students feedback and students have time to respond to feedback. I like having the first assignment due early in the semester to prevent procrastination.
2. Decide on a due date for the final paper. I like having the final paper due on the last day of class. Alternatively, a poster session, presentation, or other method of dissemination could work in lieu of, or in conjunction with, a final paper.
3. I recommend taking the class to the computer lab at least once to work on their project and data analyses. As an expert, the instructor can guide students and answer questions in real time - which is invaluable.
4. Consider taking the students to the library for help finding scholarly sources or invite a librarian to do a brief demonstration of how to find articles in class. Alternatively, making a video can be helpful as well. Screen capture is available in VLC, Office Mix, and numerous other programs. For those that are not computer savvy, I recommend Office Mix because it is free and seamlessly couples with PowerPoint.
5. Students will need some type of writing resources to help them revise their projects. A great book on scientific writing is *Style: Lessons in Clarity and Grace* by Joseph Williams. However, I want to reduce book costs for students, so I do web searches for writing resources and post them in the LMS.
6. I left my point distribution on each assignment, so you can see my weightings. Feel free to change the point weights as needed. Additionally, feel free to use all or part of this resource / assignment.

Secondary Data Analysis Project

Instructions: I want you to apply what you have learned in class to a project. I understand that, like me, many of you are very busy so collecting data may not be feasible. Therefore, I would like you to do a secondary analysis on an existing dataset. Please investigate original, logical research questions. For example, do not investigate the relationship between number of days of sunshine and the amount of precipitation because that relationship is not original and has already been explored. Do investigate something like the type of consumer complaints and the types of businesses receiving these complaints as a function of time. That is, do customer complaints change over time and are there some types of businesses that receive more customer complaints than others?

Below are some resources for datasets. However, if you would like to use another dataset, please talk to me about the dataset you plan to investigate.

<http://www.apa.org/research/responsible/data-links.aspx>

<https://catalog.data.gov/dataset>

<https://www.springboard.com/blog/free-public-data-sets-data-science-project/>

Assignment 1: Your first assignment is to find a dataset. Explore the repositories of datasets (see links above) and download at least 3 dataset you could potentially use for your project. You will submit these datasets as part of your assignment. Next you need to write down and submit at least 3 interesting research questions per dataset that you would like to pursue. Finally, select one dataset and one research question that you would like to pursue for your project, and write me a one-page (minimum) summary of why this research question is worth pursuing and what you think you might find (i.e., expected results). Then speculate about the analysis you would use to investigate your research question. This will require you to look ahead and you must use an analytical strategy that is covered in class. Put another way, use an analysis that is new to you and taught in this class. (10 points)

Assignment 2: Now that you have a research question, you need some background information on your topic. You will need to search the literature about your topic using academic databases. Remember to use keywords in your searches. I recommend the following databases: **PsychInfo, Web of Science, EBSCO, ProQuest Central, JSTOR, and ERIC**. If you do not know how to search databases, I recommend going to the library and asking a librarian for help. I expect you to use journal articles from peer reviewed journals using the databases listed above. Next, jump into the relevant literature on your topic. You need background information, and you need to evaluate if anyone else has already asked your research question. If someone has already asked your research question, then pick another research question. You will need to find at least 7 academic articles on your topic. You need to write a one-page summary for each article and then on the second page, discuss how each article will help you with your project (at least 1 paragraph). (15 points)

Assignment 3: You will write an Introduction, Methods, and Results section in APA format. You will submit your Introduction, Methods, and Results section. I am expecting a reference section and in-text citations. Additionally, please submit a copy of your data output. (30 points)

Introduction: The Introduction describes the state of knowledge on a topic. Typically, you want to start with the big picture and discuss the research on your topic. At the end of the Introduction, you want to describe how there is a gap in the current state of knowledge on your topic and how your research question will fill this gap.

Methods (in past tense): The Methods describe how you collected your data and the key to the Methods section is providing enough information for **Replication**. You will do this in the following three sections: Participants, Materials, and Procedure. The Participants section describes who participated in your study (i.e., your unit of analysis). For example, if cars in Prescott Valley, AZ are my unit of analysis, I describe my sample with enough detail so that someone in Australia can replicate my study. The Materials sections describes all the materials you used to collect your data. Everyday items like a pen or pencil should be mentioned but do not require a deep level of detail. However, non-standard items need to be described in extreme detail. For example, if I built an apparatus to collect data, I would describe it in great detail so that someone in Australia could replicate my study. The Procedure walks the reader through a step by step description of how the data was collected. Normally, I'd say that you should start from when the participant entered the lab but perhaps you should start from the beginning of your observations to the end of your observations.

Results (in past tense): The Results section describes the statistical analyses that were conducted and the Results of those analyses. Please explain your results with simple terms. Use the variables and group names when interpreting your results. Make sure to display your statistical results, central tendencies for each group, and variability for each group. Do not discuss what your Results imply because that is part of the Discussion section.

Assignment 4: There are PDFs in our Learning Management System (LMS) about writing. Read these articles about writing and apply at least 3 techniques from the writing material in Canvas to your Introduction, Methods, and Results sections. You cannot apply every technique, but cohesion, coherence, and actors and actions are good places to start. Revise the Introduction, Methods, and Results sections, add a Discussion section, and submit another draft. (25 points)

Discussion (in present tense): The Discussion section has three “movements” to it. First, describe the implications of your research to the conceptual framework laid in the Introduction. Think about what your research findings mean in terms of the Big Picture. For example, maybe I find that white cars do cut off more cars than any other car color. What does this finding mean in terms of driver safety or maybe the people that pick white cars? The second movement is the standard apology paragraph where you explore alternative hypotheses and how your study was limited. For example, maybe there are more white cars in Prescott Valley than surrounding areas like Chino Valley and Prescott. Therefore, I could have noticed more white cars cutting people off because there are more white cars on the road than any other color. For limitations, I might say something like “This study was limited to one small city in Arizona during one season (Autumn).” External or internal validity issues are common limitations. Finish the second movement by making suggestions for future research. What questions logically follow or would be interesting to know considering your discovery? The third movement for your Discussion section is the conclusion. Start narrow with your study and broaden out to the big picture. For example, I may start with my study found that white cars cut off more cars than any other car color and end with safe driving habits can save lives.

Final Project: Revise your Introduction, Methods, Results, and Discussion sections using at least 3 other techniques from the writing documents in our LMS. (55 points)