

Marketing Analytics Spring 2020 Final Project

Writing is an important part of any analysis work. Learning to convey your findings – both visually (e.g., graphs) and with words – is important.

The final project requires each student to perform data analysis tasks/research of a **marketing data set** selected by the student using **SAS Studio** tool. Each student will produce **a Project Proposal** and **a Final Project Paper** (which is a written report detailing the analysis techniques and the findings of the project) **with a SAS syntax file**.

Data Set: The data set must be obtained from an external publicly available source. That you will provide a direct link to. The data set **must be related to marketing** (e.g. consumers, profits, sales, costs, etc.) and there must be enough data to perform necessary analyses. Here are some links where you may find good datasets:

- <https://toolbox.google.com/datasetsearch>
- <https://aws.amazon.com/fr/datasets/>
- <https://www.kaggle.com/datasets>
- <http://archive.ics.uci.edu/ml/index.php>

Analyses: You must use SAS Studio that we learn and practice in this course to conduct **all** your analyses. In your analyses, you should cover **at least three** of the analyses topics we talk about in this course. However, covering 4 to 5 topics is recommended in order to gain excellent points in **Completeness and Thoroughness (Please see GRADING at the last page)**.

They include:

- Data preprocessing: how to clean, aggregate, match the raw data sets and how to transform, clean different variables
- Descriptive analysis: summary/descriptive statistics for your data
- Data visualization: different tables, charts, graphs to help audience better understand your analyses and your findings
- Statistical analysis: hypothesis testing with assumptions and limitations, testing for differences between groups and for predictive relationships
- Predictive analysis: predictive models (such as linear regression) with their assumptions and limitations

Paper: The paper should not be too lengthy but needs to be long enough so that I understand what the data and analyses are, the conclusions you make, and know how you arrived at the conclusions. **5 to 10 pages** (DOUBLE spaced, font size 12, excluding visualizations, dataset, code/spreadsheets and other attachments or reference material) should be enough.

SAS syntax file: When you submit your final project paper, you should also submit a SAS syntax file called “*YourLastName_YourFirstName_Project_sassyntax*” separately, which include the codes/syntax you generate/use to conduct your project. If you fail to submit this file or the syntax could not reflect the work you have done for your project paper, you may receive a failing score for your project.

Citation format: You must include proper citation in your paper. You are allowed to use the citation formatting that you prefer for this project. Your paper will be checked by NYU Turnitin tool for

plagiarism. If you fail to use proper citation, or the paper you submit contain more than 30% exact wording from other sources, you may receive a failing score for your project.

DETAILED INSTRUCTIONS

Use the respective assignment for your submissions

1. Final Project Proposal – Due Date: Monday, April 6, 2020, 11:55pm

You will need to turn in a 1-page paper (SINGLE spaced) proposal (Word or PDF) called “YourLastName_YourFirstName_project_proposal”. The proposal should have the following 5 sections:

- **Background:** 1 paragraph description of the overall problem you will solve
- **Purpose:** 1 paragraph that begins: “In this paper, ...” (Provide the goals of what you will accomplish in the paper and how.)
- **Data:** 1 paragraph telling the data source, and important features of the data (links to the data set, sample size, who was sampled, year collected, covariates you will use)
- **Analysis plan:** 1-2 paragraphs with your analysis plan.
- **Discussion:** 1 paragraph on how your results will help answer the question/problem you posed in “purpose.”

2. Final Project Paper + SAS syntax file – Due/Present Date: Monday, May 4, 2020, 11:55pm

You will need to turn in your paper (DOUBLE spaced, font size 12, Word or PDF) called “YourLastName_YourFirstName_Final_Project”. I have offered one way to organize your paper below, but feel free to do what works best for you. You need to make sure that the organization is easy to follow.

- **Introduction/background**—A brief (approximately one page) general description of the problem, including:
 - Why the problem is of interest—you might refer to previous studies using this, or similar data sets.
 - A brief summary of the methods utilized—tell me what methods you have utilized, e.g., “I perform a significance hypothesis testing to examine the differences of cereals across the shelf on which they are displayed.” Detailed information about the methods should be left for a later section.
 - The main results of your analysis.
- **Data/Data Preprocessing:**
 - Information about the data set (e.g, the source of the data, how many observations, number of variables, etc.)
 - If you make any change to the raw data set, you should describe how you change it and why you change it as part of data preprocessing (e.g. you delete the observations with missing values, you select random samples in the data set, etc.)

- **Exploratory analyses** —This section should provide the reader (me) with graphical and numerical summaries and results of the data, paying special attention to summaries that provide evidence for the results you've mentioned in the introduction.
 - Descriptive statistics of the variables you are interested in
 - Data visualizations of the variables you are interested in
 - You should explain both the processes and the meanings of conducting the descriptive statistics and visualizations
- **Methods** —The methods section should expand the description of the methods used. Topics that should be covered in this section include:
 - If appropriate, explicitly define any tests/models used in your analysis (e.g., significance tests, linear regression, etc.). Make sure to state why you think the method(s) is (are) appropriate for the data.
 - Discuss and evaluate the assumptions of the method(s) used. If your data does not quite conform to the assumptions, make note of it, and discuss the implications.
 - State the hypotheses you are testing and also state which testing procedure(s) you are using. For example, "I perform a hypothesis test to determine whether the mean response vector varies by display shelf, with the p-value statistic."
- **Detailed Results**—This section expands the explanation of the results, and includes, where appropriate, tables and figures providing evidence for the conclusions you've stated. You can also report any secondary results you've found.
- **Discussion**—Summarize the findings one last time, paying close attention to the limitations of the analysis. You can share thoughts with the reader about how you might expand the study, improve on the model you've used, and what are the long-term implications of the findings.

GRADING

1. Final Project Proposal (2% in Class Participation – 100-points scale)
 - Total 5 sections, 20 points of each section's contents
2. Final Project Paper + SAS syntax file (20% of course grade – 100-points scale)
 - 1) **Clarity, 15 points.** If I scratch my head and ask myself, "what are they trying to say here?" a few times when reading the paper, then it is probably not very clear. Delete long sentences with complex structure in favor of ones that are relatively short and easy to understand.
 - 2) **Appropriateness of the data and analyses, 30 points.** This is the most important piece of the project. When evaluating this portion of the project I will be asking myself, "Is this what I would have done?" "Is there a better way to perform this analysis or data visualizations?"
 - 3) **Ability to explain the results and draw the correct conclusions, 25 points.** You have used the right method(s) for data analysis. Did you explain the results well and clearly? Did you use that method correctly to draw conclusions?
 - 4) **Completeness and Thoroughness, 25 points.** The project should include **at least three** of the analyses topics we talk about in this course. Also, the paper should have proper citation inside the texts, appendix, and references. Did you perform a complete analysis? Was an adequate exploratory data analysis performed? Is there something in the data that you failed to discuss? Were all of the model assumptions discussed and evaluated? Were limitations of the method(s) discussed?
 - 5) **The wow factor, 5 points.** Extremely well-written papers will be rewarded. Also, if you **use SAS Studio** to conduct any task/analysis/research beyond what we learned in the class, you will be rewarded. For example, you find some new SAS codes we did not talk in class to conduct additional related analysis, etc. Did the student go beyond the call of duty in the analysis? Is the paper extremely well-written? Did the student suggest ways to extend the work or how the analysis could be improved?