GROUP ASSIGNMENT
INSTRUCTIONS & GRADING CRITERIA

This document describes the requirements and grading criteria for the RMBA 2771 Fall 2025 group assignment.

The learning objective of this assignment is to test whether the students can apply the methods they learned in class to a real-world data set. It assesses the following skills:

- Can the students identify an ***interesting research*** question and ***derive testable hypotheses***?
- Are the students able to ***select appropriate statistical tests*** to test their hypotheses?
- Can the students apply these statistical tests to a real-world data set and ***implement*** them ***in Python***?
- Can the students interpret their findings and clearly communicate their results in a comprehensive ***report in a Jupyter Notebook***?
- Can the students ***critically reflect on the value they added*** to the project beyond the AI tools they used? (*By now, we assume that all students use AI tools to write their code to some extent.*)

Deliverable:

The following deliverable will be used to evaluate the learning objective:

- **One Jupyter Notebook per team submitted via Moodle**
  (Deadline: 11.10.2025 23:59, one week prior to the final exam).

The Jupyter Notebook should document the analysis and present the findings, along with supporting statistics and figures. The final report should assume its reader has no prior knowledge about the data set and ensure that all findings are reproducible.

The Notebook must include at least the following sections:

1. **Executive Summary** – A short summary that highlights the goal of the project and core findings.
2. **Introduction** – A section that describes (in words) the dataset and variables. This section should clearly state the research question(s) and the related hypotheses and describe a plan to test them.
3. **Exploratory data analysis** – A section that uses different statistical metrics and visualizations to describe the data set and presents the first descriptive insights.
4. **Method 1** – This section should describe why a specific method (e.g., t-test, linear regression, logistic regression, cluster analysis, factor analysis, time series model, or panel regression) is used to test a hypothesis. It should apply the method to the data set, check the most important assumptions, and provide an interpretation of the results obtained (i.e., what did we learn about the hypotheses, and how good is the model).
5. **Method 2** – This section should contain the same information as the previous section, but with another method to test a different hypothesis.
6. **Reflection on use of AI** – This section is dedicated to discussing the use of AI, should detail what AI models were used, for what tasks AI was used, how it was used (e.g., prompt examples), and what value the students contributed beyond the tasks the AI completed (e.g., what instructions were crucial to improve the quality of the project, what approaches did not work, etc.).
7. **Conclusion** – This section should discuss the findings and explain what we learned about the research question. Further, it should discuss the chosen approach's limitations and ways to improve the analysis.

8. References (optional) - List any external sources or documentation used to support your analysis.

Important** The final report should be clean. Remove code cells that do not run (i.e., cells you used to try some code) and internal notes that are only for your team. Add comments to your code and clearly explain your process to the reader. Use **markdown cells** and **headings** to structure your notebook. **Run all cells of your notebook before you submit it to show your results and to see if they run in order!

Grading criteria:

We will use the following criteria to grade each notebook.

Structure of the notebook (10 points)

- **0 points:** One or more sections are missing.
- **5 points:** The report includes all required sections and provides explanations of the most important steps, but is difficult to follow in some sections.
- **10 points:** The report includes all required sections, provides detailed explanations of what has been done, and is easy to follow. All supporting elements (plots/tables) have a title, axis labels, and a one-sentence takeaway (ideally).

Research question and hypotheses (20 points)

- **0 points:** The report misses either a research question or does not develop hypotheses related to the research question.
- **5 points:** The report states a research question, but the hypotheses are poorly developed or not clearly linked to the research question.
- **10 points:** The report presents a research question and hypotheses, but they have notable flaws (e.g., weak connection to the data, limited clarity, or lack of relevance).
- **15 points:** The report states a clear research question and suitable hypotheses, explains why the research question is relevant and interesting, and suggests a somewhat novel angle, but the originality or depth of insight is only partially developed.
- **20 points:** The report states a clear research question and suitable hypotheses, explains why the research question is ***novel, relevant, and interesting***, and presents a novel or original perspective. It derives a strong set of hypotheses well-suited to investigate the question.

Exploratory data analysis (15 points)

- **0 points:** The report does not provide an exploratory data analysis.
- **5 points:** The report only provides basic descriptive statistics with little or no relation to the research question and with major flaws.
- **10 points:** The report presents exploratory data analysis of variables relevant to the research question(s), but with notable flaws (e.g., incomplete coverage, unclear presentation, or limited interpretation of findings).
- **15 points:** The report provides a clear and well-structured exploratory data analysis of all relevant variables, highlighting and explaining interesting findings that directly support the research question(s) and hypotheses.

Method 1 (20 points)

- **0 points:** The report does not provide this section.
- **5 points:** The report conducts some analysis beyond the exploratory data analysis, but the analysis is not clearly related to the research question and hypotheses, or the analysis is not suitable for testing the stated hypotheses.
- **10 points:** The analysis is suitable to test the hypotheses, but has some flaws in how it is executed, reported, or interpreted.
- **15 points:** The analysis is suitable for testing the hypotheses, the most important assumptions are checked, the results are well-reported, and the results are correctly interpreted.
- **20 points:** The analysis is suitable for testing the hypotheses, all important assumptions are thoroughly checked, appropriate robustness checks are reported, the results are well-reported and interpreted, and the method generates *novel insights* from a business perspective.

Method 2 (20 points)

- **0 points:** The report does not provide this section.
- **5 points:** The report conducts some analysis beyond the exploratory data analysis, but the analysis is not clearly related to the research question and hypotheses, or the analysis is not suitable for testing the stated hypotheses.
- **10 points:** The analysis is suitable to test the hypotheses, but has some flaws in how it is executed, reported, or interpreted.
- **15 points:** The analysis is suitable for testing the hypotheses, the most important assumptions are checked, the results are well-reported, and the results are correctly interpreted.
- **20 points:** The analysis is suitable for testing the hypotheses, all important assumptions are thoroughly checked, appropriate robustness checks are reported, the results are well-reported and interpreted, and the method generates *novel insights* from a business perspective.

Reflection on the use of AI (15 points)

- **0 points:** The report does not include any meaningful reflection on the use of AI.
- **5 points:** The report briefly mentions AI use but does not explain how it was applied or what the team contributed beyond the AI output.
- **10 points:** The report provides some details on which AI tools were used and for what tasks, with limited or generic examples. The reflection touches on the team's added value beyond AI but lacks depth or critical analysis.
- **15 points:** The report provides a clear, detailed, and critical reflection on the use of AI. It specifies which AI tools/models were used, for what tasks, and how they were applied (e.g., prompt examples). It discusses both benefits and limitations and clearly explains the team's unique contributions beyond AI (e.g., refining prompts, validating outputs, integrating insights, correcting errors). The reflection shows thoughtful consideration of how AI shaped the project and what originality the team added.

Note: *In addition to the grade we assign you, there will be a peer evaluation on Moodle.*