Group Project Instruction

This group project requires you to pack all skills you learned from this course (and possibly other DSA courses) to solve a real business problem using effective data management and analytics. You will practice how to solve real-world problems and generate insights using data management tools. You have two options for the project. With either option, you will work either individually or in a group of up to 3 members. The final deliverable is a short research report and a PowerPoint presentation.

Option 1: Data Processing and Analytics

This year, we provide two data analytics topics that you can choose from – healthcare IT or Airbnb listings. Your group only needs to pick one from the following two options.

Option 1A: The central theme of this year's healthcare IT project is: *How does IT transform the healthcare sector?* The instructor has secured a complete dataset of HIMSS 2017, which is a detailed survey of several thousands of healthcare entities' various IT systems in the North America area in 2017. The data dictionary is available here Download here

. Each group will identify a relevant business problem in this context. Then, you will need to identify other datasets that can complement the provided HIMSS dataset, merge and process the data, analyze it, and provide the findings of your study. You can use any data management tools that you are familiar with, although the tools we learned in this class are recommended.

Option 1B: The Airbnb Listings project uses Airbnb listings data to analyze questions related to Airbnb. The Inside Airbnb website (http://insideairbnb.com/get-the-data/Links to an external site.) provides detailed data of Airbnb listings of several major cities, including the listings, calendar, and review information. Utilizing the Inside Airbnb data as the base, you can analyze how Airbnb listings compete and survive, factors determining Airbnb growth, the impacts of Airbnb on related industries and societal issues, etc. These are just examples of questions you can explore. You are encouraged to brainstorm the interesting and novel questions to study. Similarly, You can use any data management tools that you are familiar with, although the tools we learned in this class are recommended.

Suggested research report outline: In your research report, you need to include the following sections:

- **Introduction:** Elaborate on the problem you studied and why it is important (with evidence from practice, media, or phenomena); Clearly state the question of your study;
- **Data:** Describe your data source and data processing approaches;

- **Method:** Describe your method of analyzing the data, which could be but is not limited to descriptive statistics, correlations, data visualizations, regression analysis, and predictive models;
- **Results:** Describe and summarize your results or findings from the data analysis;
- **Conclusion:** Conclude the paper and make managerial or policy recommendations;

Grading Rubrics: Your project and report will be evaluated based on the overall "contribution". A contribution can be understood as the novel knowledge you uncovered from the data analysis, which can be achieved in various ways. For example, you may study a question that nobody has looked at before and find reasonable results that are explainable. You also may focus an "old" question but add nuances such as conditions leading to different impacts or different types of interventions causing different outcomes, etc. Another potential contribution is that your findings have strong recommendations for practitioners.

In terms of grading, reasonable work should satisfy the following:

- The question of your study is reasonably motivated. In other words, you have a decent reason for studying the question;
- Data you collected and used are from valid sources and are able to answer the question of your study;
- The method employed doesn't have obvious errors. In other words, you are drawing the conclusion from an appropriate analysis
- Able to clearly interpret the results and explain why it happens and accordingly offer practical recommendations.

Beyond these general rubrics, studying a novel question or identifying counterintuitive findings challenging our common sense will be a plus to elevate your score.

Option 2: Database Design

Alternatively, you can contact a small business such as a coffee shop or a local restaurant. These small businesses may not have a database but would need one as they grow. You can perform a consulting project to help them structure their raw data into a well-designed database using techniques we learn in this class. Then, identify a business area that the small business can apply data management or analytics to improve performance and build a demo for them.

Suggested research report outline: In your research report, you need to include the following sections:

- **Introduction:** Elaborate on the small business situation and the objective of its data management or analytics project;
- **Data:** Describe the data that the small business has and can be leveraged in data management or analytics;
- **Method:** Describe your design of the data management or analytics solution;
- **Conclusion:** Outline the recommendations and next step of moving it forward;

Grading Rubrics: Your final project will be evaluated by the following:

- the extensiveness of your project scope;
- accuracy and optimization of data management or analytics solution design;
- feasibility and potential of your recommendations to the small business.