IS5740 Group Project

Semester A 2023-2024

Important Dates:

- Submit **your dataset and business problem** due on **Oct 5**.
- The **project proposal presentation** due on **Week 7** (Presentation slides must be submitted in advance of the presentation).
- **Final report submission** due on **Week 13**.
- **Final presentation** due on **Week 12 & Week 13** (Presentation slides must be submitted in advance of the presentation).

Description:

The objective of this project is for your team to walk through the business decision-making/analytics process using real-life data that is of interest to you. You need to collect real data, understand the context of the data, identify problems from the data, and apply analytics skills to form business insights. You can browse kdnuggets.com, Kaggle.com (or any other online data source you are aware of) for available datasets (see Appendix A). Pick one that is not only interesting, but also understandable to you. The students are free to choose their data source and identify questions within the context of their chosen data. The problems they choose to focus on should be carefully investigated. The deliverables of this project include:

- A presentation of the analytics project (question, analysis, recommendation);
- A detailed report of the project and the recommended solutions.

Process

1. Team Business Problem and Dataset Submission (Due Oct 5):

Please collaborate with your team members to select the business problem you intend to address and clearly specify the dataset you will use to solve this problem. To prevent duplication, please ensure that no two teams choose the same business problem and dataset. In the event of any overlap, priority will be given to the team that initially submitted the information.

Please enter the following information at this link:
https://docs.google.com/forms/d/e/1FAIpQLSce0XvAHZsBFGfeM7cYykJanhkrUNIkScZpGHfnTtFazB-mBQ/viewform?vc=0&c=0&w=1&flr=0

- Company name (Specify business area if the exact company name is not available)
- Business problem to be resolved
- Dataset name
- Data source (a website where data were obtained)

The business problem and dataset should not overlap between teams. If there is any overlap, priority will be given to the team that entered the information first.

2. Project Proposal and Presentation (Due Week7):

A detailed proposal should be submitted before **the Week 7 class via the Canvas.** It should be **five pages long ppt slides excluding a cover page**. The purpose of the interim report is to briefly describe the selected data set (including the background of the data, the structure of the data, definition of variables, etc.), preliminary descriptive analysis of the data, identified problems and questions that you want to further investigate, and how you would like to proceed.

Please think about the following questions while formulating the report:

- What is the business context and background of the project?
- Which question do you want to answer and why is answering the question important?
- Describe the data. What are the summary and preliminary descriptive statistics of the data?
- What methods will you use to analyze the data?
- What solutions can you expect from business analytics?

Proposal Presentation on the 7th Week:

- Your team's presentation should include a brief description of the innovation strategies, which your team will be dealing with (e.g., business area, etc.). I'll make a quick assessment based on this proposal whether this project will be feasible and interesting enough. Just make sure you pick something you are interested in!
- Submit your pptx file to the Canvas by the Week 7.

3. Final Report and Presentation (Due Week 12 ~ Week 13):

Be Creative! No template for the final report and presentation.

1) Analytics

- Get a data set related to the area you are interested in.
    - Links to sources of sample dataset are available in Appendix A.
- Import the data into Excel or other software tools and create relationships if needed.
- Analyze the data (need to identify the questions and use analytics to support the question)
- Any other analysis that you have conducted, regression, what-if analysis, association rule, clustering, etc.
- It is OK if you want to use other software program other than Excel. Please communicate well with the instructor.

2) Written Final Report (Due Week 13):

NOTE: ALL deliverables should be well written, free of typos, and logically organized, printed on A4 paper with 1" margin on all sides, 11-point Times New Roman font size and 1.5-line spacing.

- *Excel Project File*: An innovative, sensible and error-free analytics program using Excel (an excel project file).

- *If you prefer, you can also use other software as long as it can clearly demonstrate your data work.*
- *Written report (no more than 20 pages)*:

  - One cover page (The cover page should be included with project title, group name and group member list, which does not count into the page limit)
  - Main text (excluding the Appendix about analysis procedure and detail results). In the report, you need to cover the following:

    A. Description of data
    B. Problem of interest identified through preliminary analysis
    C. The objectives of your analytics project and the reason(s) to choose them
    D. The techniques/methods used in your analysis
    E. The purposes of adding additional columns. You may need to create additional columns.
    F. Findings from the analytics and recommendations.
    G. (Depending on the topic) The implication for managers, government, etc. To show case how your analytical program support decision making.
    H. Further discussion and additional analytics concerning the suggested solutions.
    I. (Optional) Additional analytics using other packages that you learned by yourself – briefly describe the problems, the tool/methods, and the results.

  - An example – Table of Contents:
    - Introduction (topic chosen and why, the purpose of the analysis, the main issue to be analyzed, and an overview of the report).
    - Data Description
    - Question Description (What you are interested in and Why)
    - Analysis and Findings
    - Managerial Implications
    - Description of the Dashboard (using Tableau or Power BI)
    - Conclusion (summarize the main issues and findings; generally, discuss the implications of the report)
    - References
    - Appendix – Procedure of the analysis, Additional details of analysis, creation of the dashboard, etc.

3) Project Presentation

In this final presentation, the team will present their project findings, including data, analysis and findings, problem and solutions. During the presentation, the team will take a role of consulting team and assume that the audience are the clients. Please highlight the most interesting issues and findings of the project. The presentation may be online or in class. The format will be confirmed later this semester.

Each group will be given **20 minutes to present** and **5 minutes for Q&A** session. Not all team members have to present. Team members can contribute to the project in different ways and their efforts should be reflected in the peer evaluation part. The team need to submit their **presentation slides** with the final report.

Note: All the submissions **MUST be done using CANVAS system**, including project proposal, final report (including the Excel file), and presentation slides**.** *The instructor won't accept any hard copy submission. Late submission will result in a 20% deduction in the group project grade.*

4) **Peer Evaluation**

To encourage team work, individual score will be based on the group performance as well as peer evaluation. Every student is required to submit a **private** evaluation of the performance of other members in (his/her group) in the 12th - 13th week.

Appendix A

Students are strongly encouraged to explore more data sources and identify the most rich and interesting one:

https://www.kaggle.com/datasets
http://www.data.gov/
http://data.worldbank.org/data-catalog
http://data.gov.uk/data
http://data.un.org/
http://www.economicsnetwork.ac.uk/links/data_free.htm
http://gsociology.icaap.org/data.htm
https://archive.ics.uci.edu/ml/datasets.php
https://datasetsearch.research.google.com/
https://snap.stanford.edu/data/
https://data.world/
https://registry.opendata.aws/
https://fred.stlouisfed.org/tags/series
https://www.gapminder.org/data/