Course Project

Course Project Schedule:

Week 1: Milestone 1 Due (Team Information/Communication Plan)
Week 2: Milestone 2 Due (Data Selection and Project Proposal) & Peer Review
Week 3: Peer Review
Week 4: Milestone 3 Due (Preliminary Analysis)
Week 5: Peer Review
Week 6: Peer Review
Week 7: Milestone 4 Due (Project Presentation & Status)
Week 8: Peer Review
Week 10: Milestone 5 Due (Final project paper and presentation) & Peer Review (Due Saturday!)

Project Overview:

The project will be done in teams of 3 students max or independently.

For students working in teams, it is expected that all team members will contribute equally and that everyone will take the opportunity to learn from each other. There will be 6 peer reviews due throughout the term evaluating team members for those working in a group. For students working independently, there will be a separate review to complete based on another classmates' independent project.

Students will identify a business problem to address through predictive analytics. The goal is to select appropriate models and model specifications and apply the respective methods to enhance data-driven decision making related to the business problem.

Students will identify potential use of predictive analytics, formulate the problem, identify the right sources of data, analyze data, and prescribe actions to improve not only the process of decision making but also the outcome of decisions.

Some online resources that contain interesting data.

- Nature Scientific Data
- US EPA
- Human Development Reports
- National Center for Health Statistics
- US Census Bureau (econ, population, geographic, health data)
- Bureau of Justice Statistics
- Statistics Resources Online
- USA Federal Government Data
- Google Public Datasets

There are no restrictions on what data you use, except you cannot use a data source used for course assignments.

You will want to choose a large enough dataset with some variable so that you can demonstrate the concepts and methods learned in this course. Your project is required to use a combination of both Python and R. It is up to you to determine which tool you use when, but you must use both.

You will turn in your project at the end of the term with a paper, presentation, and supporting technical documentation.

Find the details for each of the milestones and the peer reviews below:

Milestone Deliverables:

- Milestone 1 – Team Information/Team Communication/Work Plan (Week 1)
  - The first step is to determine if you are doing the project independently or as a team. Teams can have 3 students max. Based on this decision, follow the below instructions for submission.
  - Team Project Requirements:
    - Submit team member names
    - Determine your method of communication as a team (Slack, Teams, Discord, Skype, etc.)
    - Provide high level plan for tackling the project as a team
  - Independent Project Requirements:
    - If you are working independently, craft a rough project plan for how you will accomplish each milestone and task, what your project work schedule will be like, etc.
    - **If you are working independently, you need to create a thread in the discussion board "Course Project Deliverables" – you will use this thread the remainder of the course for other independent team members to be able to review your work. Failure to do this, will result in your inability to complete a peer review.**
  - Submit the requirements to the assignment link provided. This information will be used for the instructor to create a group within Blackboard that will have a location for assignment submission and a group discussion board. If you are working independently, you will still have a group created for you, but will only be visible by you and the instructor.
  - This is the only milestone that won't be submitted via your group.
  - Start discussing your topic and dataset as a team, as this will be due the following week.
- Milestone 2 – Data Selection and Project Proposal (Week 2)
  - Data selection and your project proposal are due this week. While you might determine to add additional data sources as the project progresses, you should have a good idea of your initial dataset by this milestone.
  - The project proposal should include the following information:
    - Introduction
      - Background
      - Problem Statement
      - Scope
      - Document Overview
    - Preliminary Requirement (summary of what is required to proceed, data required, business processes that need to be in place, assumptions, etc.)
      - Technical Approach
      - Data sources or plan for data
      - Analysis

                - Requirement Development (What would you need technically in place for this to work, do you have to build anything, initial work that would need to be done)
                - Model Deployment
                - Testing and Evaluation
            - Expected Results
            - Execution and Management of Project
                - Project Plan
                - Project Risk
    - The proposal should be 5-7 pages long, double spaced. Margins should be 1-inch top, bottom, left, and right. Use any font that is suitable for a professional paper and use 12-point type. Remember it is just the initial proposal – your findings might take you in a different direction for the final submission.
    - This should be submitted through the group assignment submission regardless if it is an independent project or multi-person group.
    - **If you are working independently, you will also need to reply to your thread in the discussion board "Course Project Deliverables" with this assignment– you will use this thread the remainder of the course for other independent team members to be able to review your work.**
- Milestone 3 – Preliminary Analysis (Week 4)
    - The Preliminary analysis should follow the following format as this is the format of your final paper.
        - Abstract
        - Intro/background of the problem
        - Methods
        - Results
        - Discussion/conclusion
        - Acknowledgments
        - References
    - You should be able to complete some of the intro/background of the problem, methods, preliminary results, and discussion. I understand that some of this information can change over the next few weeks, but it is better to start your Master Doc now instead of waiting until the end of the semester.
    - This should be submitted through the group assignment submission regardless if it is an independent project or multi-person group.
    - **If you are working independently, you will also need to reply to your thread in the discussion board "Course Project Deliverables" with this assignment– you will use this thread the remainder of the course for other independent team members to be able to review your work.**
- Milestone 4 – Project Presentation & Status (Week 7)
    - In this milestone, each team should present the intermediate results of their project. A deck of slides that describes the steps of the analysis up to this point should be prepared. The goal of this presentation is to describe the contributions toward the completion of the project, and to discuss the issues, challenges encountered during the work, and future plans to complete the analysis. When building this presentation, assume an executive level audience (8-15 slides max) and that they have minimal knowledge of the topic.
    - This should be submitted through the group assignment submission regardless if it is an independent project or multi-person group.

    - **If you are working independently, you will also need to reply to your thread in the discussion board "Course Project Deliverables" with this assignment– you will use this thread the remainder of the course for other independent team members to be able to review your work.**
- Milestone 5 – Final project paper and presentation (Week 10)
    - Course Presentation:
        - Each team should prepare a presentation to describe the results of the analytics project and use of the concepts and methods taught throughout the course. Presentations should be 15-20 minutes in length. The presentation will be recorded and submitted through Blackboard. The presentation is due by Thursday at midnight. Audio, not video is required.
    - Course Final Paper:
        - The final paper should be 8-10 pages double spaced in length not including figures and tables. The executive summary should be no longer than one typewritten page, describing the conclusions of your data analysis to a non-technical audience. It should be intelligible to a person who does not know data mining or machine learning techniques. Suppose you are talking to your boss or to a friend who is not familiar with statistical terminology and data science methods. This can be seen as the executive summary/introduction of your report.
        - The technical report should follow. The technical report should include an intro/background of the problem, methods, results, discussion/conclusion and acknowledgments, references, in that order. Clearly state the problem you have chosen to investigate. List the resources you used to come up with the project and reference all sources you used to complete the project. This section is intended for a technical audience and must be written in a clear organized fashion.
    - Margins should be 1-inch top, bottom, left, and right. Use any font that is suitable for a professional paper and use 12-point type. The final paper is due by Saturday at midnight.
    - **If you are working independently, you will also need to reply to your thread in the discussion board "Course Project Deliverables" with this assignment– you will use this thread the remainder of the course for other independent team members to be able to review your work.**

Peer Reviews

- Peer reviews are due in Weeks 2, 3, 5, 6, 8, and 10 and are NOT submitted via your group
- Team Projects
    - You will complete the Team Peer Review Form
    - It is an evaluation of the contribution of the other team members, overview of issues, risks, etc.
    - Submit via assignment link
- Individual Projects
    - You will communicate with classmates doing the project independently, either via the discussion board or through Slack, etc.
    - You can work with the same classmate the entire course, or you can interact with different peers the entire course.
    - You will complete the Individual Peer Review Form – this form will ask you to assess the project topic, the classmates plan, provide input on risks/issues/status, etc. The goal of this review is to provide a second set of eyes

to help your classmate. During Week 3 & 6, feel free to reach out to your classmates to ask questions in order to complete your review since there will not be a direct milestone to review. You are also free to share your thoughts with the student via the discussion board since this will be submitted via the assignment board.

Remember – your GitHub repository can act as a portfolio for potential employers! I would highly suggest using this to submit your work or add your milestones to. You are welcome to submit a link to your GitHub for these project milestones. Make sure your GitHub is not private.