Data Science Capstone Project

OVERVIEW & PURPOSE

This project will bring together all of the skills you have learned in this class and allow you to showcase your work to the public. It will incorporate data capture, entry, cleaning, transformation, analysis, modeling, and action plan. Most importantly, you will drive business value through the usage of data and demonstrate industry best practices.

REQUIREMENTS

1. Use R as the major tool for work with data
2. Ensure your work is 100% transparent and reproducible
3. Present your work in a high quality fashion through the use of RMarkdown

OBJECTIVES

1. Find a topic for your project
    a. Plenty of data available for analysis
    b. Business case can be made to analyze - i.e. is there an action plan that can impact a business in this industry?
    c. Requires advanced data analysis techniques
    d. Is interesting / creative - i.e. not just following analysis that has been done by thousands of others already online
2. Create a roadmap for completion of your project
    a. Include objectives & goals as well as technical details
    b. This is not set in stone
3. Clearly document all of the following:
    a. Collect / find / enter data
        i. Be careful to ensure you are following best practices to avoid biases
        ii. Ensure ethical data collection and that you are legally allowed to use data
    b. Clean data
    c. Transform / Wrangle data

    d. Complete exploratory data analysis (EDA)
    e. Create hypotheses to test
    f. Analyze data
    g. Model data
        i. Build models to understand data and describe what you are seeing
        ii. Build predictive models
        iii. Build prescriptive models
    h. Show findings
4. Create write-up of all steps
    a. RMarkdown to be used
    b. Will look professional
    c. 100% of project on GitHub
    d. Will be 100% reproducible
        i. Anyone should be able to clone the repository and be able to knit your document and get the exact results you present
    e. Ensure everything contributes to the business goal you set out to accomplish and highlight that throughout the write-up

VERIFICATION

1. Project is due 1-week before class ends
2. Presentation of project will occur on the final day of class
3. Grade (pass / fail) based off of:
    a. 30% weekly checkpoints
    b. 70% final project