MSDS Data Practicum I – Revised Project Proposal

# Title of the project

Identifying the Markers for Sociopolitical Unrest in the US

# High level description of the project: what question or problem are you addressing?

Are there consistent, broad-based “markers” which can be used to indicate an increased risk of sociopolitical unrest in the United States? Markers of interest include economic factors, educational trends, social chatter, public health concerns, political priorities, and law enforcement activity.

# What type of data science task is it? (some example answers but not limited to)

This effort will involve identifying relationships, predictive analysis, and data visualization.

# Data: Brief description of data. How big do you expect the data will be? Is amount of your data too big or too small? If you're web-scraping or collecting data, how long do you expect to collect the data?

Datasets to be collected include:

* Consumer Price Index (economics)
* Consumer Confidence Index (economics)
* Sentiment analysis (social media)
* Outbreak tracking (public health)
* Correlates of State Policy (politics)
* Violent crime rate (law enforcement)
* Property crime rate (law enforcement)

Ideally, I would like 80 to 100 years’ worth of data; however, I do not expect data to be available over the full timeframe for each of the desired datasets.

# How will you analyze the data? What machine learning methods do you plan to use, and/or what business intelligence aspect do you plan on incorporating?

Data will be analyzed using an unsupervised machine language technique, most likely an association rule learning method such as link analysis given the varied data sources. The R implementation of the Apriori Algorithm seems well-suited for this task.

# Describe any anticipated difficulties and problems. Discuss how you may overcome the problems.

Finding valid data sources over the desired timeframe for each of the anticipated markers will likely be my biggest issue.

# Suggest a timeline for the project. This should be a weekly breakdown of what you plan on doing each week.

My proposed project plan, by week, is detailed below:

A screenshot of a computer

Description automatically generated

Figure 1: Data practicum project timeline (Gantt Chart)

As noted in both the Gantt (above) and table (below) views I expect to spend roughly one week on each major section of the project starting with sourcing data following the submission of my revised project proposal.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **WBS Number** | **Title** | **Start** | **End** | **Duration** | **Prerequisites** |
| 1 | MSDS 692 - Data Practicum I Project | 1/18/24 | 3/10/24 | 1mo 3w 2d |  |
| 1.1 | Week 1 | 1/18/24 | 1/21/24 |  |  |
| 1.1.1 | Course Start | 1/18/24 | 1/18/24 |  |  |
| 1.1.2 | Project Proposal Due | 1/21/24 | 1/21/24 |  | 1.1.1 |
| 1.2 | Week 2 | 1/28/24 | 1/28/24 |  |  |
| 1.2.1 | Proposal Revision Due | 1/28/24 | 1/28/24 |  | 1.1.2 |
| 1.3 | Week 3 | 1/29/24 | 2/4/24 |  |  |
| 1.3.1 | Identify and Source Data | 1/29/24 | 2/2/24 | 1w | 1.2.1 |
| 1.3.2 | Progress Report Due | 2/4/24 | 2/4/24 |  | 1.2.1 |
| 1.4 | Week 4 | 2/5/24 | 2/11/24 |  |  |
| 1.4.1 | Cleanse and Prep Data | 2/5/24 | 2/9/24 | 1w | 1.3.1 |
| 1.4.2 | Progress Report Due | 2/11/24 | 2/11/24 |  | 1.3.2 |
| 1.5 | Week 5 | 2/12/24 | 2/18/24 |  |  |
| 1.5.1 | Explore and Understand Data | 2/12/24 | 2/16/24 | 1w | 1.4.1 |
| 1.5.2 | Progress Report Due | 2/18/24 | 2/18/24 |  | 1.4.2 |
| 1.6 | Week 6 | 2/19/24 | 2/25/24 |  |  |
| 1.6.1 | Model Data | 2/19/24 | 2/22/24 | 4d | 1.5.1 |
| 1.6.2 | Progress Report Due | 2/25/24 | 2/25/24 |  | 1.5.2 |
| 1.7 | Week 7 | 2/23/24 | 3/3/24 |  |  |
| 1.7.1 | Interpret Data | 2/23/24 | 2/29/24 | 1w | 1.6.1 |
| 1.7.2 | Progress Report Due | 3/3/24 | 3/3/24 |  | 1.6.2 |
| 1.8 | Week 8 | 3/1/24 | 3/10/24 |  |  |
| 1.8.1 | Explain Analysis | 3/1/24 | 3/6/24 | 4d | 1.7.1 |
| 1.8.2 | Draft Presentation Due | 3/7/24 | 3/7/24 |  | 1.8.1, 1.7.2 |
| 1.8.3 | Final Presentation Due | 3/10/24 | 3/10/24 |  | 1.8.2 |

# Create GitHub repository for your Practicum project. Add this proposal, begin a ReadMe document, and begin adding your data to your repository. Add a link to your GitHub repository to this document.

Repository – https://github.com/ZeakGhox/MSDS692\_Practicum\_I