Instructions for Proposal

**Business Problem Scenario**

* Clearly articulate the business problem or challenge that your model will address.
* Identify the specific stakeholders who would benefit from your solution.
* Define the primary goals of your project in terms of business outcomes.
* Explain why a machine learning approach is appropriate for this problem.
* Describe the dataset/s you've chosen, including its source, size, and key variables.
* Connect the relevancy of the dataset/s to your business problem.
* Define how you will measure success from both technical and business perspectives.

**Problem Solving Process**

Should include the following:

1. **Data Acquisition and Understanding**  
   * How you will obtain and explore the dataset
   * Initial data quality assessment plan
   * Preliminary visualization strategy
2. **Data Preparation and Feature Engineering**  
   * Data cleaning approach
   * Feature selection/engineering methodology
   * Implementation plan for sci-kit learn Pipeline
3. **Modeling Strategy**  
   * Algorithms you plan to evaluate (minimum 3)
   * Cross-validation strategy
   * Hyperparameter tuning approach
   * Evaluation metrics selection and justification
4. **Results Interpretation and Communication**  
   * How you will translate model results to business insights
   * Visualization plans for model performance and feature importance
   * Strategy for explaining technical concepts to non-technical stakeholders
5. **Conceptual Framework**  
   * Include a flowchart of your proposed solution pipeline
   * Outline dependencies between different stages of your project

**Timeline & Scope**

Create a detailed timeline covering the the time allotted for your capstone project:

* **Dataset finalization and problem formulation - [TIME]**  
  + Dataset acquisition and initial exploration
  + Business problem definition refinement
  + Project repository setup
* **Exploratory Data Analysis - [TIME]**  
  + Comprehensive data profiling
  + Statistical analysis of relationships
  + Creation of informative visualizations
  + Documentation of insights
* **Data Preprocessing - [TIME]**  
  + Data cleaning implementation
  + Feature engineering
  + Pipeline development
  + Data splitting (train/validation/test)
* **Model Development - [TIME]**  
  + Implementation of baseline models
  + Algorithm comparison
  + Hyperparameter tuning
  + Cross-validation
* **Model Evaluation and Refinement - [TIME]**  
  + Final model selection
  + Performance evaluation on test data
  + Business metric calculation
  + Interpretation of results
* **Documentation and Reporting - [TIME]**  
  + Code commenting and cleanup
  + Technical report writing
  + Executive presentation development
* **Final Review and Submission - [TIME]**
  + Quality assurance
  + Video recording
  + Final submission preparation

Identify potential challenges or areas where you might need to conduct additional research or learning