1. Name, Contact info: Beth Birhanu (bbirhanu@regis.edu)
2. Title of the project: Automated Specific Response System for Email
3. High level description of the project: what question or problem are you addressing?

The project aims to create an automated response system that generates specific, predefined replies to emails, facilitating quick and relevant responses to common queries.

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

* Text classification using supervised learning.
* Natural language processing for feature extraction
* Text generation for automated response creation

1. Data: Brief description of data. How big do you expect the data will be? Is the 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?

The dataset is the Enron Email Dataset featuring about 0.5 million messages from senior management. For practicality, a 10,000-email subset will be used, which is a manageable size and allows for a focused approach without being overwhelming.

1. 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 NLP techniques for text preprocessing and feature extraction, with machine learning methods like Naive Bayes or LSTM for response prediction and generation. The goal is to derive a model that understands email context and suggests specific appropriate replies.

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

Challenges include the intricacies of natural language understanding, the complexities of preprocessing data to fit modeling needs, and generating suitable responses. The preprocessing phase is particularly critical as it involves cleaning the data, handling missing or incorrect information, and converting unstructured email text into a structured form suitable for analysis. These might be addressed through:

* Iterative modeling and evaluation to refine understanding and response accuracy.
* Employing advanced NLP libraries and techniques for robust text preprocessing.
* Utilizing pre-trained models where applicable to improve efficiency.
* Adopting a modular approach in the preprocessing pipeline to iteratively address and test for data quality and readiness for modeling.

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

Week 3: Set up environment, begin data preprocessing.

Week 4: Continue preprocessing, initial model training.

Week 5: Develop response generation logic, test models.

Week 6: Model evaluation, maybe begin report drafting.

Week 7: Final model refinements, complete report draft.

Week 8: Finalize report, prepare presentation.