# Approach for Final Project

## Northwestern Data Science & Virtualization Bootcamp

1. **Project Management and Coordination** 
   1. **Role**: Project Manager – Brittany
   2. **Ideal Skills**: Leadership, organization, communication
   3. **Responsibilities**:
      1. Oversee the entire project and ensure deadlines are met.
      2. Coordinate between team members to ensure smooth communication.
      3. Compile and finalize documentation, including the GitHub repository and README.
      4. Keep track of the project’s progress and address any roadblocks.
2. **Data Collection and Preparation** 
   1. **Role**: Data Engineer – Alexa
   2. **Ideal Skills**: Python, data manipulation, SQL
   3. **Responsibilities**:
      1. Fetch historical stock data using the yfinance library.
      2. Clean and preprocess the data, including handling missing values and normalizing/standardizing.
      3. Create and manage the SQL database if needed Keep track of the project’s progress and address any roadblocks.
3. **Exploratory Data Analysis (EDA)** 
   1. **Role**: Data Analyst (EDA) - Abby
   2. **Ideal Skills**: Data analysis, visualization, Python, statistical analysis
   3. **Responsibilities**:
      1. Perform EDA to understand the data, identify trends, and uncover insights.
      2. Visualize key trends and patterns using tools like Matplotlib and Seaborn.
      3. Share insights with the team to guide feature engineering and modeling decisions
4. **Model Selection, Training, Evaluation and Optimization**
   1. **Role**: Machine Learning Engineers (Vlad, Soumik, Charly)
   2. **Ideal Skills**: Data analysis, Machine learning Python, model training
   3. **Responsibilities**:
      1. Create new features from the raw data that might improve the predictive power of the models.
      2. Develop technical indicators like moving averages, RSI, etc.
      3. Work closely with the Data Engineer and Data Analyst to ensure the features align with the data and insights
      4. Select appropriate models for predicting stock prices (e.g., ARIMA, Random Forest, LSTM).
      5. Train models on the prepared dataset.
      6. Tune hyperparameters to optimize model performance.
      7. Evaluate model performance using appropriate metrics like R-squared, MAE, MSE.
      8. Iterate on model selection and tuning to improve accuracy.
      9. Document the model evaluation process and share findings with the team.
5. **Visualization and Presentation**
   1. **Role**: Visualization Specialist
   2. **Ideal Skills**: Visualization, design, Python (Plotly/Bokeh), presentation skills.
   3. **Responsibilities**:
      1. Create visualizations that effectively communicate the results and insights from the project.
      2. Develop interactive plots using Plotly or Bokeh if needed.
      3. Work with the Project Manager to prepare presentation materials and visual aids.