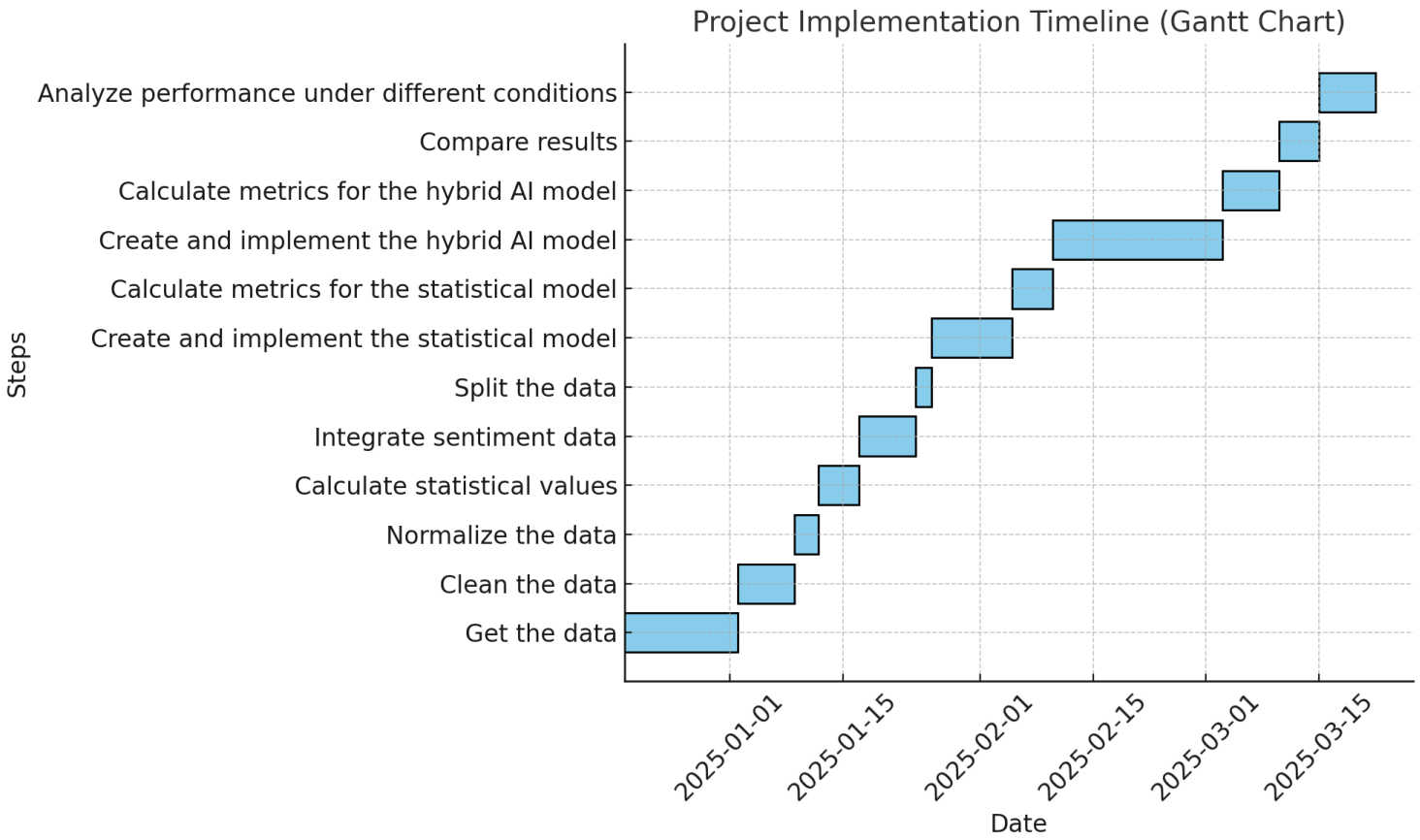
Steps in the Implementation that I can think of:

- **Get the data**:
  - Collect minute-level UK stock market data (2015–2024)
    - Start with **Yahoo Finance** (easily accessible with Python's yfinance).
    - Use **Alpha Vantage** for intraday data (free but with rate limits).
  - Extract sentiment data from financial news articles and reports.
    - Use web scraping or APIs to fetch financial news, earnings reports, or macroeconomic announcements related to the stocks or indices you are focusing on.
    - Clean and structure the text data for sentiment analysis (e.g., remove HTML tags, tokenize text).
    - Use pre-trained models (e.g., VADER, FinBERT) or custom models to analyze the sentiment of the text data.
    - Assign a sentiment score to each piece of news (e.g., positive, negative, or neutral).
    - Map the sentiment scores to the corresponding time periods in your market data.
    - Create a unified dataset for use in statistical and AI models.
- **Clean the data**:
  - Handle missing values, outliers, and inconsistencies.
  - Preprocess sentiment text data for analysis.
- **Normalize the data**:
  - Apply feature scaling to market and technical data.
  - Standardize sentiment scores.
- **Calculate statistical values**:
  - Compute Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Simple Moving Average (SMA), Exponential Moving Average (EMA), Bollinger Bands, Average True Range (ATR), On-Balance Volume (OBV), and Volume Weighted Average Price (VWAP).
- **Integrate sentiment data**:
  - Assign sentiment scores to corresponding time periods in the dataset.
- **Split the data**:
  - Divide into training, validation, and testing sets.
- **Create and implement the statistical model**:
  - Use technical indicators to create a regression or ARIMA-based strategy.
- **Calculate metrics for the statistical model**:
  - Evaluate using metrics like Sharpe Ratio, Sortino Ratio, Maximum Drawdown, Calmar Ratio, etc., and record results.
- **Create and implement the hybrid AI model**:
  - Develop reinforcement learning models (e.g., DQN, PPO) integrated with sentiment data.
- **Calculate metrics for the hybrid AI model**:
  - Evaluate using the same metrics (Sharpe Ratio, Sortino Ratio, Maximum Drawdown, Calmar Ratio, etc.) and record results.
- **Compare results**:
  - Compare profitability and risk management performance between the statistical and hybrid AI models.
- **Analyze performance under different conditions**:
  - Test robustness during stable and volatile periods (e.g., Brexit, COVID-19).



Pardon the fact I don’t fully understand the info but it seems good to me