Bank of England (BoE) Employer's Project

Team 6: Project Plan and Scope

Bank of England

Team members:

Alex Joseph

Manpreet Minhas

Prabha Raveendran

Rafa Abdelazim

Vinod Thomas Thoppil

Background/context and clearly defined problem statement:

One role of the Bank of England(BoE) is to provide stability and reassurance to financial markets through speeches at public events. BoE wants to know if trends in speeches correlate with observed events and economic indicators as well as whether the sentiment of given speeches can be used to predict market behaviour. More specifically, BoE's Data Strategy & Implementation Division would like answers to the following questions:

- Has the sentiment of central bank speeches changed over time? If so, how has it changed?
- How the sentiment of BoE's speeches correlates with key events such as:
    - base rate changes
    - publication of the Monetary Policy Report and Financial Stability Report/Review
    - other events or trends that may be relevant or interesting.
- How speech sentiment correlates with key economic indicators of the UK, such as:
    - GDP growth
    - inflation
    - labour market statistics (e.g., unemployment and wages)
    - any other economic indicators that may be relevant or interesting.
- Can speech sentiment assist in predicting market behaviour?
- Are there other findings from the analysis that are of interest to BoE?
- What are the potential reasons for any correlations discovered above? How have these conclusions been drawn?

Additional questions to potentially enhance analysis:

- How does the sentiment of central bank speeches affect stock market indices (e.g., FTSE 100, S&P 500)?
- What is the relationship between speech sentiment and volatility in the markets?
- What is the relationship between speech sentiment and the trade balance of imports and exports? If exports > imports then a positive indicator, but if exports < imports then a negative indicator.
- How does the sentiment of BoE speeches affect the UK exchange rate relative to other countries?

Team roles and ways of working:

Through our team charter and virtual meetings, we have engaged in discussions about our individual strengths. We assigned two team members to each role, ensuring that each role had a primary and a secondary strength. This approach allowed one team member to handle the task at hand while the other provided peer review, thereby optimizing our output.

Here is a summary of the team members and their respective strengths:

Team Members	Primary Strength	Secondary Strength
Alex	Data wrangling	Data visualisation
Manpreet	Model development	Correlation analysis
Prabha	Recommendation	Data wrangling
Rafa	Visualisation	Model development
Vinod	Correlation analysis	Documentation

The work distribution was covered in 4 main sections which include:

- Data acquisition and cleaning
- Model development
- Correlation analysis
- Documentation and presentation

To facilitate this, we created a Gantt chart (*see Appendix A*) to provide a detailed breakdown of the work distribution and to address the business questions pertinent to this project. This also allowed us to allocate the necessary resources for each step, taking into consideration our primary and secondary strengths and help measure our progress.

For communication, we have established the following plans:

- **Platform**: WhatsApp group-chat for day-to-day communication.
- **Regular Check-ins**: Weekly teams meetings to discuss progress, address concerns and realign if needed.
- **Asynchronous Collaboration** *(allow members to work based on their schedule*)**:**
  - Google Docs for document collaboration
  - Canvas group pages to share work
  - Github repository to share code

Project Plan:

1. Initiation Phase:

- **Objective Definition**: Collectively agreeing on the project's purpose, aligning on the business problem, and defining success criteria.
- **Stakeholder Identification**: Although this is an academic project, it's vital to recognise who might benefit from our findings, such as fellow students, course instructors or even financial institutions.

2. Planning Phase:

- **Task Breakdown**: Data collection, pre-processing, sentiment analysis, correlation analysis and report/presentation drafting.
- **Resource Allocation**: Assign specific roles based on team strengths and ensure every member has the tools and resources needed.
- **Milestone Creation**: Define key milestones.
- **Risk Management**: Identify potential roadblocks and create contingency plans such as a project scope creep (*see Appendix A*).

3. Execution Phase:

- **Data Collection & Cleaning**: All members jointly source speeches.
- **Sentiment & Correlation Analysis**: Data Modeler & NLP Specialist leading this stage, supported by the Economic Correlation Analyst to provide context.
- **Continuous Feedback Loop**: Ensure progress against milestones and address challenges.

4. Monitoring & Control Phase:

- **Progress Tracking**: Using a Gantt chart to monitor progress and through open communication via WhatsApp and use of shared documents.
- **Quality Assurance**: Periodic reviews of the work done
- **Stakeholder Communication**: Keep stakeholders informed about progress and challenges.

5. Closure Phase:

- **Final Review & Refinement**: A final team review after draft report completion to ensure findings align with project objectives and success criteria.
- **Documentation**: Of data, models and project artifacts was archived for future use.
- **Feedback Session**: Discuss project outcomes and lessons learned for future endeavours.

Our project roadmap decisions encompassed:

- **Tool selection**: (Canvas for task management, Google Docs and GitHub for real-time collaboration)
- **Sentiment analysis** approach: Involving pre-trained NLP models and fine-tuning for accuracy
- **Data source Limitation:** To ensure integrity, and bi-weekly feedback check-ins for agility.

These decisions were driven by three core pillars:

1. **Efficiency** for time and resource optimization
2. **Accuracy** for high-quality analysis
3. **Collaboration** to keep the team in sync while working toward project goals.

Please see project roadmap and Gantt chart in appendix for key milestones and timeline of deliverables.

Analytical approach:

1. **Data Collection + Cleaning (1 week)**:

    - **Web scraping**: To obtain and store data sets for economic indicators (e.g., GDP and inflation).
    - **Identification**: Exploratory data analysis(EDA) to identify missing values, duplicates, outliers, invalid entries, or inconsistencies in the speech and economic indicators data.
    - **Standardisation**: Uniformity in data formatting, especially in dates, currency, text formatting and data types.
    - **Normalisation**: Adjust text to a consistent case (e.g., lowercase) and remove elements like stop words, punctuation, or any non-relevant metadata. Check for spelling errors and inconsistent units.
    - **Handling Missing Data**: Determine if missing values should be imputed, interpolated, or removed, based on their nature and impact on analysis.
    - **Programs Used**: Excel to check spelling errors in speeches. Python with libraries such as Pandas for data manipulation, Matplotlib & Seaborn for visualisation during EDA and Beautiful Soup for web scraping.

2. Data Analysis (8 days):

- **Sentiment Extraction** : Utilise NLP techniques to extract sentiment scores from speeches – categorising them as positive, negative, or neutral.
- **Time Series Analysis** : Align sentiment scores with dates and events to observe trends over time.
- **Correlation Analysis** : Map sentiment scores against market behaviours and economic indicators to find patterns or correlations.
- **Predictive Modelling** : Attempt to construct models that can potentially predict market behaviour based on past speech sentiments.
- **Programs Used**: Python with libraries such as NLTK or spaCy for NLP tasks, Scikit-learn for machine learning, and TensorFlow or PyTorch if deep learning models are deemed necessary.

Appendix

a. Project Roadmap
b. Gantt chart

a. Project Roadmap

b. Gantt chart

Team six			Week 1							Week 2							Week 3							Week 4							Week 5							Week 6							Week 7
Task	Description	Duration (days)	Resources	D1	D2	D3	D4	D5	D6	D7	D1	D2	D3	D4	D5	D6	D7	D1	D2	D3	D4	D5	D6	D7	D1	D2	D3	D4	D5	D6	D7	D1	D2	D3	D4	D5	D6	D7	D1	D2	D3	D4	D5	D6	D7	D1
1	Project Plan and scope	7	5								Assignment 1 submission: Project Scope and Plan							Initial Meeting with BoE													Presentation day with BoE							Assignment 2 submission: Initial Recommendation pitch			Presentation day with BoE				Assignment 3 & 4: Final Report + presentation and Individual reflection
2	Import and clean provided data set (central bank speeches)	4	1 (+1 peer review)																																										
3	Web scraping to extract, store and clean GDP growth data	2	1 (+1 peer review)																																										
4	Web scraping to extract, store and clean inflation data	2	1 (+1 peer review)																																									
5	Web scraping to extract, store and clean labour market statistics data	2	1 (+1 peer review)																																									
6	Web scraping to extract, store and clean stock market indicies data	2	1 (+1 peer review)																																									
7	Web scraping to extract, store and clean volatility data	2	1 (+1 peer review)																																									
8	Web scraping to extract, store and clean Trade Balance data	2	1 (+1 peer review)																																									
9	Web scraping to extract, store and clean Exchange rate data	2	1 (+1 peer review)																																									
10	Conduct sentiment analysis of central bank speeches	2	1 (+1 peer review)																																									
11	Conduct Exploratory Data Analysis for initial insights	4	5																																									
12	Has the sentiment of central bank speeches changed over time? If so, how has it changed?	2	1 (+1 peer review)																																									
13	How the sentiment of BoE's speeches correlates with key events	2	2 (+1 peer review)																																									
14	How speech sentiment correlates with key economic indicators of the UK	2	2 (+1 peer review)																																									
15	Can speech sentiment assist in predicting market behaviour?	3	5																																								
16	How does the sentiment of central bank speeches affect stock market indices (e.g., FTSE 100, S&P 500)?	2	1 (+1 peer review)																																								
17	What is the relationship between speech sentiment and volatility in the markets?	2	1 (+1 peer review)																																								
19	What is the relationship between speech sentiment and the trade balance of imports and exports?	2	1 (+1 peer review)																																								
20	How does the sentiment of BoE speeches affect the UK exchange rate relative to other countries?	2	2 (+1 peer review)																																								
21	Project scope creep allowance	5	5																																								
22	Work on actions/feedback given by BoE and facilitator + submit slides to canvas	6	5																																								
23	Prepare for final presentation	2	5																																								
24	Finalise report, submit presentation slides + individual reflection	3	5																																								
25																																												
26																																												
27																																												
28																																												
29																																												
30																																												
Total resources being used				5	5	5	5	5	5	5	1	5	5	5	5	5	5	5	5	5	5	5	5	0	5	5	5	5	5	5	5	5	5	5	5	5	5	1	5	5	5	5	5	5	1

Key	
	D1 = Mon
	D2 = Tues
	D3 = Wed
	D4 = Thurs
	D5 = Fri
	D6 = Sat
	D7 = Sun
	Resources = no. of people assigned to task