Cowry Data Scientist Interview Task

Objective

These task assesses your ability to:

- Frame and solve complex behavioural data problems.
- Design and defend a rigorous, scalable methodology.
- Extract and communicate actionable insights from structured and unstructured data.
- Balance statistical sophistication with interpretability and business impact.

You will respond to two real-life client briefs. Your submission will be presented in a 1-hour session:

- 45 minutes to walk through your approach, models, insights, and recommendations.
- 15 minutes for Q&A and technical discussion.

Prepare all of your slides in Google Slides. Provide figures to illustrate your findings throughout. Provide all code/notebooks in advance of your interview.

Task 2: Understanding Impact of a Behaviourally-Optimised Call Script on Customer Perceptions

Context: Applied NLP, Behavioural Analysis, and Experimentation

You have conducted a field experiment testing a new decision-support script for call-centre agents. The aim was to improve customer outcomes when choosing broadband packages - simplifying decisions, building trust, and reducing decision regret.

Your data consists of:

- Treatment and control groups (with/without the new script),
- Open-text feedback from post-call surveys,
- Metadata including date and whether the customer is from the VOLT customer segment

Your Challenge

You are tasked with understanding:

1. How customers define "good" service, and how the new script shifted those definitions.
2. What aspects of service (e.g. clarity, empathy, agent personality) drive sentiment.
3. Whether the new script systematically changed perceptions or emotional tone - particularly for high-value segments like VOLT.

Your Tasks

1. Preprocessing

- Exclude any responses from February (simulate a temporal filter).
- Segment the data into:
    - Non-VOLT customers
    - VOLT-only subgroup

    - Treatment vs Control conditions within each.

2. Topic Modelling & Theme Discovery

Use methods of your choice e.g.:

- LDA / BERTopic / NMF for topic extraction from the open text.
- Open-Source LLMs for zero-shot theme extraction or summarisation.

For each group (general / VOLT), answer:

- What are the top latent topics mentioned?
- How do topic distributions differ by treatment?
- What percentage of comments mention agent personality, clarity, or reassurance?

3. Sentiment Analysis

- Quantify sentiment using a technique of your choice
- Compare:
    - Sentiment distribution between treatment vs control
    - Strength of sentiment for VOLT vs general sample
- What topics co-occur with negative or positive sentiment?
- Any shifts in emotional tone driven by the new script?

4. Modelling + Interpretation

(Optional but strongly encouraged)

logistic regression, likely most interpretable while remaining effective

- Train a classification model to predict positive vs negative sentiment using extracted topics or LLM embeddings.
    - Identify key explainers of sentiment
- Evaluate whether treatment assignment is predictive of topic use or sentiment, controlling for segment

Deliverables

1. Slides including:
    - Key behavioural themes that define good customer service.
    - Impact of the new script on topic prevalence and sentiment
    - Recommendations for refining agent scripts based on linguistic patterns and psychological insights.
2. Python notebooks demonstrating your pipeline (bonus for modular, reusable).

Evaluation Criteria

Your submission will be assessed based on the following:

- Rigour and scalability of methodology (e.g., modular code, reusable NLP pipeline)
- Integration of behavioural science frameworks (e.g., trust, decision ease, motivation)
- Depth and clarity of insights (not just surface-level topic identification)
- Ability to isolate causal or meaningful group differences (e.g., treatment vs control, VOLT vs general)
- Use of interpretable models and transparency of assumptions
- Appropriate use of LLMs with clear justification and critique including pros/cons of scale-up (cost, explainability, hallucination risk)
- Consideration of potential response bias or noise in text data
- Clearly note assumptions, limitations, and ethical risks (e.g., over-interpreting LLM-extracted themes)
- Professional communication: clarity, structure, prioritisation of insights in slides
- Actionability of recommendations for script refinement and deployment
- Optional: inclusion of modelling or sentiment drivers to enhance insight depth