Interim GitHub Submission: Brent Oil Price Change Point Analysis

This submission addresses Task 1 of the project, laying the foundation for the analysis of Brent oil prices. It includes an interim report outlining the planned workflow, assumptions, and strategy, along with a structured dataset of key events and the proposed GitHub repository organization.

Interim Report (Task 1: Laying the Foundation for Analysis)

1. Overall Challenge and Project Objectives

The primary objective of this project is to analyze Brent crude oil prices using Bayesian change point analysis to detect structural breaks. Aim to investigate how major geopolitical events including political decisions, conflicts, sanctions, and OPEC policy changes influence these price fluctuations. The ultimate goal is to deliver clear, data driven insights that can guide investment strategies, policy development, and operational planning for stakeholders in the energy sector.

2. Data Analysis Workflow

The analysis of Brent oil prices will follow a systematic, phased approach to ensure thoroughness, accuracy, and the generation of actionable insights.

Phased Approach:

- **Phase 1: Data Acquisition & Preparation (Current Focus - Task 1):** This involves gathering the historical Brent oil price data (May 20, 1987, to September 30, 2022) and compiling a comprehensive dataset of relevant geopolitical and economic events. Key steps include data cleaning, handling missing values, and verifying data integrity.
- **Phase 2: Exploratory Data Analysis (EDA) & Feature Engineering:** A deep dive into the time series properties of the Brent oil price data will be conducted, examining trends, stationarity, and seasonality. Event data will be engineered into suitable features for modeling (e.g., binary indicators, lagged effects).
- **Phase 3: Model Selection & Implementation (Task 2):** Bayesian change point models, specifically using PyMC3, will be selected and implemented to identify structural shifts in the data.
- **Phase 4: Model Validation & Interpretation (Task 2):** Rigorous validation of the models will be performed. Detected change points will be interpreted in the context of compiled event data, assessing the nature, magnitude, and duration of price responses.
- **Phase 5: Insight Generation & Communication (Task 3):** Analytical findings will be translated into clear, actionable intelligence, tailored for diverse stakeholders and disseminated through formal reports and an interactive dashboard.

Detailed Steps for Data Ingestion and Initial Assessment:

The Brent oil price data will be loaded, ensuring correct parsing of 'Date' and 'Price' columns. The 'Date' column will be converted to a datetime object and set as the time series index. Initial data quality checks will verify completeness and identify anomalies.

Integrating Geopolitical and Economic Event Data:

A structured dataset of 10-15 key geopolitical and economic events, with approximate start dates, will be compiled. This dataset will provide crucial context for detected change points. Events will be represented as features (e.g., binary indicators) for integration into the models.

Modeling and Interpretation Framework:

Bayesian change point models will be applied to identify statistically significant shifts in the mean, variance, or other parameters of the Brent oil price time series. Detected change points will be correlated with the compiled events, examining temporal proximity and the nature of the price response. Where feasible, the magnitude and duration of event impacts will be quantified.

3. Assumptions and Limitations

Core Assumptions:

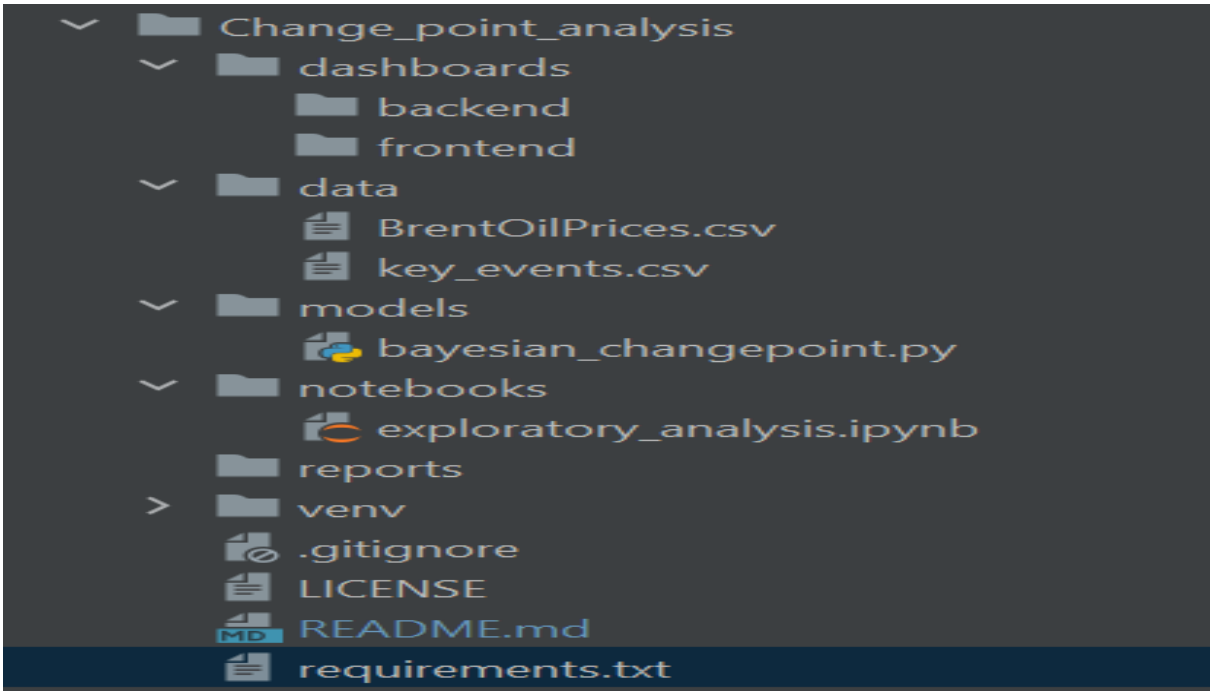
- **Data Reliability:** The provided Brent oil price data is assumed to be accurate and representative.
- **Event Significance:** The selected geopolitical and economic events are assumed to be significant enough to potentially induce structural breaks in oil prices.
- **Model Applicability:** Bayesian change point models are considered appropriate for detecting shifts in the underlying statistical parameters of the time series.

Distinguishing Statistical Correlation from Causal Impact:

It is crucial to differentiate between statistical correlation and causal impact. While change point analysis can identify structural breaks in the data that *coincide* with specific events, this temporal association does not automatically imply direct causation. The global oil market is influenced by numerous simultaneous factors, including economic growth, OPEC policies, and other geopolitical developments. For example, the 9/11 attacks initially saw a price increase, but prices then dropped significantly due to demand concerns, illustrating a complex interplay rather than a simple cause-and-effect relationship.

4. Proposed GitHub Repository Structure

To facilitate collaboration and reproducibility, the GitHub repository will be organized as follows:



5. Communication Methods

- **Interactive Dashboard:** Develop a dashboard to visualize insights and allow users to explore data interactively.
- **Comprehensive Report:** Create a formal report summarizing the methodology, findings, and recommendations tailored for different stakeholders.

This interim report lays the groundwork for the subsequent phases of the project, ensuring clarity and a focused approach as we move forward with the analysis of Brent oil prices and their relationship with significant geopolitical events.