Task 2: Change Point Modeling and Insight Generation - Final Report

Part 2.1: Core Analysis

1. Implement the Change Point Model: A Bayesian Change Point detection model was applied to the Brent crude oil price series. To ensure a stationary time series suitable for modeling, the analysis was performed on the daily log returns of the price data. The model was implemented using the PyMC library, with a discrete uniform prior on the change point (`tau`) and normal priors for the mean (`mu`) and half-normal priors for the volatility (`sigma`) before and after the change point. The model successfully converged, as evidenced by all R-hat values being 1.0.

2. Identify Change Points: The model detected a single, statistically significant structural break in the behavior of the Brent crude oil price series. The posterior distribution of the change point (`tau`) was highly concentrated, indicating high certainty about the timing of the shift. The most probable date for this change point, based on the mode of the posterior distribution, is **August 21, 2008**.

3. Associate Changes with Causes: The detected change point of **August 21, 2008**, aligns precisely with the onset of the most intense phase of the 2008 Global Financial Crisis. This period was characterized by unprecedented global market turmoil, which profoundly impacted all financial assets, including commodities like crude oil. This event serves as a strong candidate for triggering the observed changes in oil price behavior. It is hypothesized that the escalating fears of a global recession and the subsequent flight to safety directly contributed to the detected shifts in price dynamics.

4. Quantify the Impact: The change point model provides clear quantitative evidence of the impact of this event:

- **Change in Mean Price Behavior:** The average daily log return of oil prices shifted from a slightly positive value before the change point (mean = 0.0003) to a slightly negative value after the change point (mean = -0.0001). The model found a minor average percentage change in price of -0.04% after the change point. The probability that the mean log return increased was low (23.24%), suggesting a probable decrease or stabilization in average daily returns.
- **Change in Volatility:** The most significant and certain finding is a dramatic increase in market volatility. The model is **100% certain that volatility increased** after the change point. Quantitatively, the daily volatility (`sigma`) is estimated to have shifted from an average of `0.023` before August 21, 2008, to `0.029` after. This jump in volatility is a hallmark of a chaotic and uncertain market environment.

Part 2.2: Advanced Extensions (Optional)

For future work, the analysis could be extended with the following approaches to build a more comprehensive explanatory model:

- **Explore Other Potential Factors:** To move from correlation to a more causal understanding, other data sources could be integrated. Incorporating macroeconomic variables such as **GDP growth rates**, **inflation rates**, and **exchange rates** (e.g., USD/EUR) would allow the model to test whether the observed change point is a standalone event or if it is part of a broader macroeconomic shift. These factors could be included as additional predictors within the PyMC model, allowing for a more nuanced analysis of their impact on price behavior.
- **Consider Advanced Models:** Alternative models could provide different and valuable insights:

    - ✓ A **Vector Autoregression (VAR) model** could be used to analyze the dynamic, multi-directional relationship between oil prices and macroeconomic variables. This would reveal not only how oil prices react to economic news but also how changes in oil prices might subsequently affect economic indicators.
    - ✓ A **Markov-Switching model** could be employed to explicitly define and model different market "regimes" (e.g., a "calm" regime with low volatility and a "turbulent" regime with high volatility). This would allow for a more flexible analysis of market behavior, as it would not be constrained to a single, discrete change point, but rather a probability of switching between these regimes at any given time.