Question 1:

"How do you model spend carryover?"

Response: Implemented a geometric adstock with normalized retention weights.

$$\text{adstock}_t = \sum_{l=0}^{L-1} w_l \cdot \text{spend}_{t-l}, \quad w_l = \frac{\alpha^l}{\sum_{k=0}^{L-1} \alpha^k}$$

Here, the retention-rate "α" parameters were sampled from a Beta$(2, 2)$ distribution. This distribution was chosen because it provides moderate decay and performs well across different channels. For the maximum-lag "L" parameter, I experimented with values of 6, 8, and 12, and selected 8 as it yielded the best R^2 score. The next step would be to define channel-specific maximum lags and retention rates, however, that would add more complexity to the model.

Question 2:

"Explain your choice of prior inputs to the model."

Response: Priors were initially selected to provide flexibility, and then were updated based on posteriors.

Base revenue: Sampled from a normal distribution that was centered around average historical revenue.

Chanel effectiveness (β): Sampled from a half-normal distribution to ensure media impacts remained non-negative.

Retention rate (α): Sampled from a Beta(2, 2) distribution, which provides medium decay. I also experimented with models with channel-specific distributions, where the channels were sampled with Beta(2, 2), Beta(8, 2) and Beta(2, 8), based on the posterior profile of each channel. However, this approach added additional complexity, so I decided not to use it.

Trend: Sampled from a normal distribution, where the μ parameter was chosen from previous posteriors to allow faster convergence.

Dual seasonality: Both the semi-annual and quarter-annual components were sampled from a normal distribution.

Trend: Sampled from a normal distribution.

Sigma: Sampled from a half-normal distribution to keep the noise positive.

Events: Sampled from normal distributions that were centered around average event revenue.

Question 3:

"How are your model results based on prior sampling vs. posterior sampling?"

Response: The model parameters showed clear convergence after fitting.

Initially, prior samples for channel effectiveness were wide and uncertain. After running the model with data, the posterior distributions narrowed significantly, reflecting updated and more confident estimates.

Question 4:

"How good is your model performing? How do you measure it?"

Response: The model works reasonably well.

To evaluate performance, I used MAPE (Mean Absolute Percentage Error), which was approximately 16.44%, acceptable for gaining a general understanding of the data. The R^2 value was 0.733, indicating that the model explains more than half of the variance in revenue, suggesting that the key drivers are largely captured.

Question 5:

"What are your main insights in terms of channel performance/effects?"

Response: The channel effects/performances show large variation.

The results reveal that channel 2 has the highest β of 29.0, an order of magnitude higher than the other channels. Channels 1 and 7 have moderate β values of 2.9 and 1.5, respectively, while channel 4 has no effect ($\beta = 0$).

Question 6:

"Can you derive ROI estimates per channel?"

Response: The ROI estimates also reflect large variation.

The results reveal one superstar– channel 2, with an ROI of 8.44%, which is significantly higher than the others. This suggests that either channel 2 is truly exceptional or the model is overfitting to its relatively small spend. Meanwhile, channel 1 has an ROI of 0.23% and channel 6 has an ROI of 0.01%. Compared to other channels, they still provide a modest return. However, channels 3 and 7 show ROIs that are essentially zero, with spends of 6915.14 € and 19507.18 € respectively. Given their low returns, these channels should be closely examined, as their performance may not justify the level of investment.