Marketing Campaign Causal Analysis Summary

By Ifeanyi Enekwa

Problem Statement

- Assess the **causal impact of a marketing campaign** on user **spend** and **churn**.
- Use multiple modeling approaches to capture different aspects of user behavior:
- **Bayesian Regression** for per-user spend modeling
- **CausalImpact (Time Series)** for total spend over time
- **Survival Analysis (Cox Model)** for churn risk analysis

Bayesian Regression Results (Surrogate Model)

Equation:

```
predicted_spend_i = β_0 + β_1 × received_campaign_i + ε
```

Parameter	Estimate	95% Credible Interval	Interpretation
β_0 (Intercept)	144.34	137.65 to 151.08	Average spend without the campaign
β_1 (Campaign Effect)	+84.32	75.02 to 93.49	Increase in spend when campaign is received

Sigma (Noise / Unexplained variance): 92.49 (87.58 to 97.84)

Interpretation:

- The model estimates that **running the campaign increases average spend by \~84 units**.
- Without the campaign, users spend around **144 units** on average.
- With the campaign, spend increases to about **228 units**.
- The model converged well (**Rhat=1.00**, high effective sample size).

Plain English Summary:

- The marketing campaign **significantly increased per-user spend**.

Surrogate Labels (Predicted Spend Data)

- The **surrogate_labels.csv** file contains model-generated predictions of user spend based on campaign exposure.

Campaign Group	Predicted Spend (approximate)
Received (1)	\~300 units
Not Received (0)	\~101 units

Why the Difference?

- The surrogate model assigns higher predicted spend to campaign-exposed users:

For non-campaign users: predicted_spend ≈ baseline (β_0) For campaign users: predicted_spend ≈ baseline + campaign effect ($\beta_0 + \beta_1$)

Note: Some discrepancies between the surrogate model (\~101 baseline, \~300 with campaign) and the Bayesian regression summary (\~144 baseline, +84 uplift) are due to data splits, different models, or noise in predictions.

CausalImpact Analysis (Time-Series Approach)

Script: `causal_inference.R`

Metric	Value
Observed Avg Spend (Post-Campaign)	298
Predicted Avg Spend (No Campaign)	305 (95% CI: 290 to 321)
Absolute Effect (Avg)	–7.2 units
Relative Effect	–2.3%
Probability of Causal Effect	79%
Bayesian p-value	0.206

Interpretation:

- The **time-series model suggests no significant impact on overall spend over time**.
- Although the point estimate shows a slight decrease, the 95% credible interval includes zero.

Plain English Summary:

- There is **no strong evidence** that the campaign changed total spend over time.

- The analysis was based on **pre-post time periods**, not campaign vs control groups.

Survival Analysis (Churn Model)

Script: `survival_model.R`

Metric	Value
Hazard Ratio	0.913 (≈9% churn reduction)
p-value	0.152 (not significant)
95% CI	0.8061 to 1.034

Interpretation:

- The campaign group had **slightly lower churn risk**, but this effect is **not statistically significant**.
- The **hazard ratio of 0.913** suggests about a **9% reduction in churn**, but the confidence interval includes 1.0.

Plain English Summary:

- The campaign **may reduce churn slightly**, but we cannot conclude this with confidence.

Summary of All Results

Analysis	Metric	Result
Bayesian Regression	Spend per user	**Significant increase** (+84 units)
CausalImpact (Time Series)	Spend over time	**Small negative estimate, not significant**
Survival Analysis (Cox Model)	Churn / retention	**Small reduction in churn, not significant**

Key Takeaways

- The **Bayesian regression** shows that the **campaign increased individual user spend** significantly.
- The **CausalImpact time-series analysis** suggests **no significant change in overall spend over time**.
- The **survival model** shows a **small, non-significant reduction in churn risk**.

Next Steps

- Use **Bayesian regression insights** to guide **personalized marketing**.
- Consider further investigation into **heterogeneous treatment effects** (segment-level analysis).
- Use findings to inform **marketing strategy and resource allocation**.

End of Summary