Subject: Proposal for Testing the Hypothesis on Customer Churn at PowerCo

Dear Sir/Ma'am,

I hope this email finds you well. Following our discussion on PowerCo's customer churn issue, Me and my team have framed the problem using the **5-step Data Science methodology**, with a focus on understanding the drivers behind customer attrition — particularly around price sensitivity, service satisfaction, and renewable energy preferences.

Below, we outline our proposed approach to test the hypothesis that **customer churn is influenced by factors such as pricing, service quality, and preference for clean energy**, and how we can leverage data science to identify actionable insights.

1. Business Understanding
PowerCo has observed higher-than-expected customer churn rates and wants to understand the underlying reasons. The business goal is to *identify key drivers of churn* and develop targeted strategies to improve retention.

2. Problem Framing
Hypothesis:
Customers are more likely to switch providers due to price sensitivity, dissatisfaction with service (e.g., outage response), or lack of access to green energy options.

Type of Problem:
This is a **predictive classification task** — predicting whether a customer will churn (leave) based on historical data and behavioral patterns.

3. Key Drivers of Churn (Hypotheses)
Based on initial analysis and industry knowledge, we believe the following factors may influence churn:
- **Price sensitivity** – Are customers leaving due to higher prices compared to competitors?
- **Service satisfaction** – Do frequent outages or poor support lead to dissatisfaction?
- **Preference for clean energy** – Are customers switching to providers offering greener alternatives?
- **Contract type & tenure** – Are month-to-month customers more likely to leave?
- **Demographics & location** – Is churn concentrated in certain regions or customer segments?

4. Steps to Test the Hypothesis

Step 1: Data Collection
We will need access to the following datasets from PowerCo:
- **Customer-level data:** Churn status, monthly charges, total charges, contract type, payment method, usage history, demographics, and location.
- **Service data:** Outage frequency, resolution time, complaint history, CSAT scores.
- **Energy plan data:** Whether the customer is enrolled in a green energy plan.
- **Historical discount/pricing changes** – if available.
- **Competitor pricing data**– optional but useful for context.

Step 2: Feature Engineering
- Create new features like price change over time, average outage duration per customer, or satisfaction score.
- Encode categorical variables and normalize numerical features.

Step 3: Exploratory Data Analysis (EDA)
- Analyze trends in churn over time, by region, and across demographic groups.
- Use visualizations to explore relationships between churn and potential drivers.

Step 4: Predictive Modeling
- Train models (e.g., logistic regression, random forest) to predict churn likelihood.
- Evaluate model performance using accuracy, precision, recall, and AUC-ROC.
- Interpret feature importance to understand which factors most influence churn.

Step 5: Actionable Insights
- Identify high-risk customers who might benefit from targeted offers (e.g., discounts, green plans).
- Recommend operational improvements (e.g., better outage response).

5. Visualizations and Techniques to Investigate Impact of Churn Drivers
- **Bar charts** showing churn rate by contract type, payment method, or green plan enrollment.
- **Boxplots** comparing monthly charges for churn vs non-churn customers.
- **Time series plots** to visualize churn trends over months/years.
- **Heatmaps** of churn by location or income bracket.
- **Survival curves** to estimate time-to-churn by customer segment.
- **Feature importance plots** to highlight top predictors of churn.

Next Steps & Timeline
- **Week 1–2:** Request and preprocess data from PowerCo.
- **Week 3:** Conduct EDA and begin modeling.
- **Week 4:** Finalize model, generate insights, and prepare report for client.

We're ready to move forward once we receive confirmation and data access from the client.

Thank you,
Best regards,
Maliha Sanjana
Data Science Intern