Subject: Exploring Price sensitivity and SME Churn - Initial Analysis

Hi Associate Director,

I hope this email finds you well. I've spent some time thinking about how we could approach our task, which is to dig deeper into the hypothesis that customer churn is driven by price sensitivities. Here are my initial thoughts and steps needed to test this hypothesis.

Initial Hypothesis Formulation
The hypothesis is that price changes exert influence on customer churn. To assess this theory, we can compose the problem as a predictive modeling task where we aim to predict which customers are more likely to churn based on their price sensitivities.

Steps to draw out the Hypothesis:

1. Understanding the Churn:
Let's start by clearly defining what constitutes churn for PowerCo. It could be a situation in the context of SME, where the SME's terminate the contracts or significantly reduce energy usage.

2. Data Collection:
Gather historical data on SME customer interactions, including:
- Billing information (price changes over time)
- SME energy usage patterns
- Customer type (SME, corporate, residential)
- Customer demographics(size of the business, location)
- Contract length(duration of their relationship with PowerCo)

3. Feature Engineering:
Identify key features that could influence churn, with a focus on price sensitivity:
- Percentage change in price:(current price - previous month's price / previous month's price)
- Monthly average spending:(total spend / number of months)
- Contract length:(current month - contract start month)
- Customer feedback or complaints related to pricing.

4. Data Cleaning:
Pre-processing the data:
- Impute missing values using average or forward-fill for time series data.
- Remove outliers using statistical methods.
- Ensure data types are consistent and suitable for analysis.
- Perform ETL(extract, transform & loading) operations.

5. Exploratory Data Analysis (EDA):
Explore the data:
- Visualise how price changes align with SME churn rates
- Explore correlation between variables
- Examine the distribution of monthly average spending.
- Investigate the relation between contract length & churn.
- Look for patterns or trends in churn behaviour related to SME customers.

6. Model Selection:
Choose an appropriate predictive model for SME customer Churn:

- Logistic regression , decision tree, and random forest are good starting points; as logistic regression models are for interpretability & suitability for classification tasks like churn prediction.
- Consider the trade-off between model complexity and interpretability.

7. Model Training and Evaluation:
Split the data into training(80 %) and testing(20 %) sets:

- Train the logistic regression model on a training set.
- Evaluate its performance on the testing set using metrics like accuracy, precision, recall, and F1- score.

8. Interpretability:
Use model interpretability techniques:

- Examine coefficients of the logistic regression model to understand feature importance.
- Focus on the coefficients of the price change feature to assess its impact on prediction.

9. Testing the 20% Discount:
Apply the trained model on the 1st working day of every month:

- Identify SMEs predicted to churn.
- Offer a 20% discount to these SMEs for the upcoming month.
- Monitor churn rates in the subsequent month to assess the discount's effectiveness.

Data Requirements:
For this analysis, we would need access to historical customer data, including billing records, usage patterns, customer demographics, and any relevant feedback or complaints.

Next Steps:
I suggest we schedule a meeting to discuss these initial thoughts further and gather additional insights from the team. It would be beneficial to ensure we have a comprehensive understanding of the data and the feasibility of the proposed approach.

I hope this provides a clearer picture of the solution for each stepLooking forward to your feedback.

Best regards,
Tejal Mhatre
Data Science Intern