SUB TASK 3

Key Findings:

1. **Price Sensitivity vs. Churn for Energy:**
  - Median price sensitivity is similar between churned and retained customers, suggesting no significant average difference.
  - Moderate box heights indicate moderate variability in price sensitivity within both groups.
  - Longer upper whisker for retained customers implies potential outliers with extreme price sensitivity among them.

2. **Price Sensitivity vs. Churn for Power:**
  - Median price sensitivity is comparable between churned and retained customers, indicating no substantial average difference.
  - Both groups exhibit tall box heights, signifying substantial variability in price sensitivity within each group.
  - Short upper whiskers for both groups suggest fewer potential outliers.

Suggestions for Data Augmentation:

To enhance our analysis and modeling, we recommend augmenting the dataset with additional data sources:

1. **Customer Demographics:** Include demographic data such as age, gender, income, and geographical location for each customer. This will allow us to assess whether customer characteristics influence price sensitivity.

2. **Competitor Pricing Data:** Obtain data on pricing strategies and offers from competitors in the energy market. This information will help us evaluate PowerCo's pricing competitiveness.

3. **Customer Behavior History:** Collect historical data on customer behavior, such as previous switching patterns, customer inquiries, and complaints. Understanding past behaviors can aid in predicting future churn.

4. **Weather Data:** Incorporate weather-related data, including temperature, seasonal variations, and climate patterns. Weather can impact energy consumption and customer behavior.

5. **Economic Indicators:** Integrate economic indicators such as inflation rates, employment data, and economic growth. Economic factors can influence customer decisions related to energy consumption and churn.

6. **Social Media Sentiment:** Analyze sentiment data from social media platforms related to PowerCo and competitors. Customer sentiment can provide insights into customer satisfaction and dissatisfaction.

Open Source Datasets:

Consider leveraging publicly available open-source datasets to enrich our analysis:

1. **Energy Consumption Data:** Explore government or utility-provided datasets on energy consumption patterns by region or industry.

2. **Census Data:** Utilize publicly available census data to gain insights into population demographics and geographic trends.

3. **Weather APIs:** Access weather data through open APIs or datasets provided by meteorological organizations.

4. **Economic Data Repositories:** Government agencies often publish economic data that can be accessed for analysis.

5. **Social Media APIs:** Utilize APIs from social media platforms to collect sentiment data.

These additional data sources will provide a more comprehensive view of customer behavior and market dynamics, enabling us to refine our predictive modeling and churn analysis.