KEY FIND OUTS FROM CLIENT DATASET

1. Using distribution of categorical data like channel sales and origin up, we found out that, which channel is being used more and add campaign contributing to customer retention.
2. The distribution of 'electricity consumption' and 'forecasted consumptions' over 12 months shows a right-skewed distribution, indicating that most customers have lower consumption values, with a few customers having very high consumption.
3. Using categorical data by grouping them with respect to churn, we found out which channel sales and campaign are prone to churn.
4. Used stacked bar plot to displayed, no of percentage (%) of companies at verge of churn and not churn
5. Used pair plot with kind=’scatter’, to show linear relation ship between dependent and independent variable, and found following
6. Data is right skewed
7. No linear relationship with churn(‘Target’), need to transform the data
8. Then used boxplot for outlier detection and found out that , many of numerical features have outliers

SUGGESTIONS

1. Competitor price data - perhaps a client is more likely to churn if a competitor has a good offer available?
2. Average Utilities prices across the country - if PowerCo’s prices are way above or below the country average, will a client be likely to churn?
3. Client feedback - a track record of any complaints, calls or feedback provided by the client to PowerCo might reveal if a client is likely to chur