The BCG project team has identified the potential of understanding churn behaviour through a model, especially focusing on the role of price sensitivity. To address this, an in-depth exploratory data analysis utilizing historical customer and pricing data provided by the client is conducted. The primary objective of this analysis is to discern patterns in customer churn and specifically probe the hypothesis that price sensitivity might be a significant driver behind it.

1. Importing Libraries:

- Libraries for data analysis (e.g., pandas, numpy) and visualization (e.g., matplotlib, seaborn) are imported.

2. Loading Data:

- `client\_data.csv` and `price\_data.csv` are loaded into pandas dataframes.

3. Initial Data Inspection:

- The first and last few rows of both dataframes are displayed.

- Column descriptions are provided for both datasets.

4. Data Types and Transformation:

- The `info()` method is called on both dataframes to check data types.

- Date columns in `client\_data` are converted from object type to datetime type.

5. Descriptive Statistics:

- The `describe()` method is called on both dataframes to view statistical summaries.

- Null values in both dataframes are counted.

- The distribution of the churn variable in `client\_data` is displayed.

6. Data Visualization:

- Multiple functions are defined to aid in plotting various charts.

- A bar plot showing the churning status of customers is displayed.

- Sales channel analysis is visualized using stacked bar plots.

- Consumption metrics are plotted using box plots.

- Histograms display forecast metrics, margins, subscribed power, and other variables.

- Contract type distribution is visualized with a bar plot.

7. Price Data Analysis:

- The number of unique dates and client IDs in `price\_data` is checked.

- Histograms for different price metrics are plotted.

8. Investigating Price Sensitivity:

- The price sensitivity of customers is explored based on data from 2015.

- Average price data is calculated for the entire year, last 6 months, and last 3 months.

- The average price data is merged to create a new dataframe.

9. Final Data Preparation:

- The `client\_data` and modified price data are merged to create a unified dataframe named `churn\_data`.

- The column `price\_date\_x` is dropped.

- The resulting dataframe is saved to a CSV file for further analysis.

10. Conclusion:

- The EDA process is concluded with a note that the next step involves Feature Engineering & Modelling.

This code provides a comprehensive EDA of the given datasets, exploring various facets of the data, and setting the stage for further analyses and modelling.