Marketing Attribution Analysis Report

Objective: The Jupyter Notebook is designed to perform a marketing attribution analysis using Markov chains. This analysis helps in understanding the impact of different marketing channels on customer conversion.

Methodology:

1. Data Preprocessing:
    - The data is loaded from a CSV file containing various marketing channel touchpoints.
    - All numerical values are converted to strings, and any trailing '.0' is removed to clean the data.
2. Path Creation:
    - A 'Path' column is created by concatenating all the touchpoints for each customer journey, leading up to a conversion event (channel 21).
3. Conversion Attribution:
    - The 'Conversion' column is set to 1 for each path, indicating a conversion has occurred.
    - The data is then grouped by 'Path' to sum up the total conversions for each unique path.
4. Markov Model Implementation:
    - A Markov Model is instantiated with specified parameters such as path feature, conversion feature, and simulation settings.
    - The model is fitted with the preprocessed data to calculate the attribution for each channel.
5. Attribution Results:
    - The results of the Markov Model are saved, showing the total conversions attributed to each channel.
    - Additional attribution methods like First Touch, Last Touch, and Linear are also calculated for comparison.
6. Visualization:
    - The attributions from different methods are visualized using a bar plot, providing a clear comparison of how each channel is valued across different attribution models.

Code Output Explanation:

- The output includes a DataFrame with channels and their corresponding conversion values as per the Markov Model.

- A heatmap is generated to visualize the transition probabilities between channels, indicating the likelihood of a customer moving from one channel to another in their path to conversion.

Conclusion: The notebook effectively demonstrates the application of Markov chains in marketing attribution, providing insights into the contribution of each channel to conversions. The visualizations aid in interpreting the model's results, making it easier to compare the effectiveness of different channels.

Dependencies: To run this notebook, the following Python libraries are required:

- pandas
- seaborn
- matplotlib
- pychattr (for the Markov Model)

Please ensure these libraries are installed in your Python environment before executing the notebook.