

Multi-Touch Attribution Modeling for B2B Marketing

Abstract

This project focuses on analyzing marketing channel effectiveness using multi-touch attribution modeling. A synthetic B2B marketing dataset was created and analyzed using Python to implement First-Touch, Last-Touch, Linear, Time Decay, and Markov Chain attribution models. The results were visualized using Power BI dashboards with interactive KPIs, drilldowns, and attribution comparisons. The study highlights the limitations of single-touch attribution models and demonstrates how data-driven attribution can support better marketing budget allocation and decision-making.

1: Introduction

1.1 Background

In the digital marketing ecosystem, customers interact with multiple marketing channels such as social media, email campaigns, search engines, and paid advertisements before making a purchase or conversion. Traditional attribution models often fail to capture the true contribution of each channel because they assign all credit to a single touchpoint, usually the first or last interaction. This limitation leads to inefficient budget allocation and suboptimal marketing decisions.

Multi-Touch Attribution (MTA) addresses this challenge by distributing conversion credit across all touchpoints in a customer journey. By leveraging data analytics and modeling techniques, organizations can better understand which channels influence customer behavior and optimize their marketing strategies accordingly.

1.2 Problem Statement

Most businesses struggle to accurately measure the effectiveness of individual marketing channels due to the complexity of customer journeys. Relying on single-touch attribution models provides an incomplete picture and may undervalue critical channels that assist conversions indirectly. Therefore, there is a need for a comprehensive attribution framework that fairly evaluates all customer touchpoints.

1.3 Objectives of the Project

The primary objectives of this internship project are:

- To design and implement multiple attribution models using Python
- To compare different attribution techniques and analyze their impact on channel performance
- To visualize attribution results using an interactive Power BI dashboard
- To derive actionable business insights for marketing optimization

1.4 Scope of the Project

This project focuses on analyzing customer journeys using a synthetic dataset that simulates real-world marketing interactions. The scope includes implementing five attribution models, performing comparative analysis, and

building dashboards for decision support. Advanced machine learning-based attribution models are beyond the scope of this project.

2: Literature Review

Marketing attribution has evolved significantly with the growth of digital platforms. Early attribution approaches such as First-Touch and Last-Touch models were simple but often misleading. Research highlights that customers typically engage with multiple channels before conversion, making multi-touch models more reliable.

Linear attribution distributes equal credit to all touchpoints, while Time Decay assigns more weight to recent interactions. Markov Chain models use probabilistic transitions between channels to estimate removal effects, making them more data-driven and robust. Existing studies suggest that Markov models provide deeper insights but require higher data quality and computational effort.

3: Dataset Description

3.1 Data Source

Due to confidentiality and accessibility constraints, a synthetic dataset was created to simulate real-world marketing journeys. The dataset reflects realistic customer interactions across various digital marketing channels.

3.2 Dataset Structure

The dataset contains the following key fields:

- Lead ID: Unique identifier for each customer
- Channel: Marketing channel interacted with
- Timestamp: Date of interaction
- Journey Length: Number of touchpoints in a journey
- Converted: Binary indicator of conversion

- Revenue: Revenue generated from conversion
- Touchpoint Order: Sequence number of interaction

3.3 Data Assumptions

The dataset assumes that all recorded interactions are valid, timestamps are accurate, and each lead follows a linear chronological journey. Revenue is attributed only to converted leads.

4: Tools and Technologies Used

4.1 Python

Python was used for data preprocessing and attribution model implementation due to its flexibility and rich ecosystem of analytical libraries.

4.2 Libraries

- Pandas: Data manipulation and analysis
- NumPy: Numerical computations
- NetworkX: Markov chain modeling

4.3 Power BI

Power BI was used to create interactive dashboards, enabling visualization of attribution results, KPI tracking, and drill-down analysis.

5: Attribution Models Implementation

5.1 First-Touch Attribution Model

This model assigns 100% of the conversion credit to the first interaction channel. It is useful for understanding which channels initiate customer journeys.

5.2 Last-Touch Attribution Model

The last-touch model gives full credit to the final interaction before conversion, emphasizing closing channels.

5.3 Linear Attribution Model

The linear model distributes conversion credit equally across all touchpoints in the journey, treating each interaction as equally important.

5.4 Time Decay Attribution Model

Time decay assigns higher weights to interactions closer to the conversion event, assuming recent engagements have stronger influence.

5.5 Markov Chain Attribution Model

The Markov model analyzes transition probabilities between channels and calculates the removal effect to estimate each channel's contribution. This model provides a data-driven and probabilistic view of attribution.

6: Power BI Dashboard Design

6.1 KPI Overview

Key performance indicators such as total conversions, total revenue, conversion rate, and average journey length are displayed using KPI cards.

6.2 Channel Drilldown Analysis

Interactive matrices allow users to explore customer journeys by channel and timestamp, enabling granular analysis.

6.3 Attribution Comparison

A model selector slicer allows users to switch between attribution models and observe how channel contributions change dynamically.

7: Results and Analysis

The comparison of attribution models revealed significant differences in channel performance. First and Last Touch models heavily favored specific channels, while Linear and Time Decay provided balanced views. The Markov model highlighted the true assisting value of mid-funnel channels that were undervalued in simpler models.

8: Business Insights and Recommendations

Channels like Organic Search and Email Campaigns consistently contributed across models. Paid channels played a strong assisting role rather than direct conversion. Marketing budgets should be allocated using multi-touch insights rather than single-touch models

9: Limitations of the Study

- Use of synthetic data may not capture all real-world complexities
 - Assumes linear customer journeys
 - Does not incorporate offline marketing channels
-

10: Conclusion and Future Scope

This project successfully demonstrates the importance of multi-touch attribution in modern marketing analytics. By combining Python-based modeling with Power BI visualization, the project provides a practical framework for channel performance evaluation.

Future work may include using real-world datasets, incorporating machine learning-based attribution models, and integrating cost data for ROI analysis.

Appendix

- Python code snippets for attribution models
- Power BI dashboard screenshots
- Data dictionary