# Contribution of Channels

TBA
Mert Eygi

# Agenda

- Introduction & Objective
- Brief Introduction to Attribution Models
- Dataset Overview & Preprocessing
- Exploratory Data Analysis
- Applied Attribution Model Approach
- Results & Key Findings

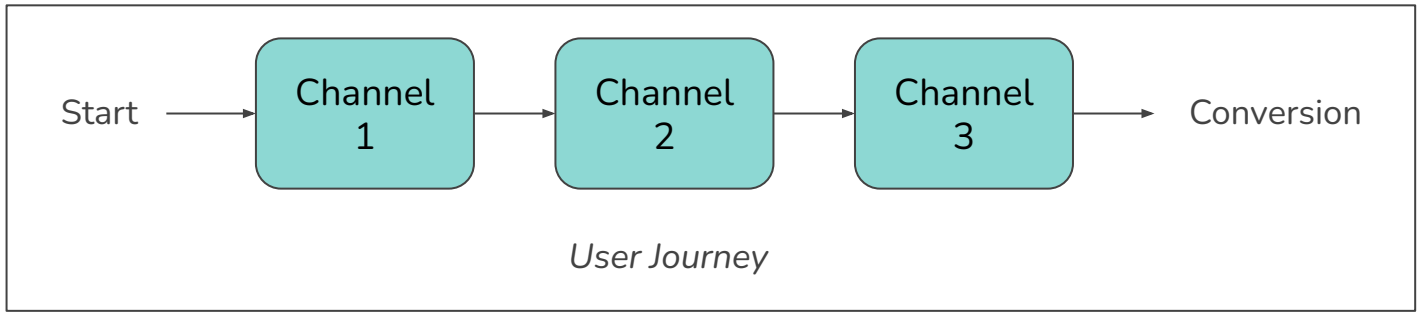
# Introduction & Objective

**Challenge**:

Bidders interact with multiple touchpoints before converting. But how do we know which channels are truly driving conversions?

**Goal**:

The goal of this task is to assess the value of each channel and evaluate the results of the attribution models.



# Brief Introduction to Attribution Models

4 common attribution models have been selected:

- **First Touch**: Attributes a user's conversion entirely to the first channel they interacted with.
- **Last Touch**: Assigns full credit for a user's conversion to the last channel they engaged with.
- **Linear Touch**: Distributes credit equally across all channels involved in the user's journey.
- **Markov Chains**: Uses a probabilistic approach to model the user's journey and determine the impact of each channel.

First Touch

Last Touch

Linear Touch

Markov Chain
(Data-Driven)

# Data Overview & Preprocessing

**Dataset**

- Over 580K cookie records from 240K users, containing impression and conversion events across five channels: Paid Search, Facebook, Instagram, Online Video, and Online Display.

**Data Cleaning & Quality Checks**

- Removed duplicate records and handled missing values.
- Verified that no users had multiple conversions incorrectly recorded.
- Checked for duplicate records at the same time instance for each user.
- Identified and excluded an extremely high conversion value (€10M) from a single Facebook user to prevent skewed analysis.

# Exploratory Data Analysis

## Key Metrics Per Channel at First Glance

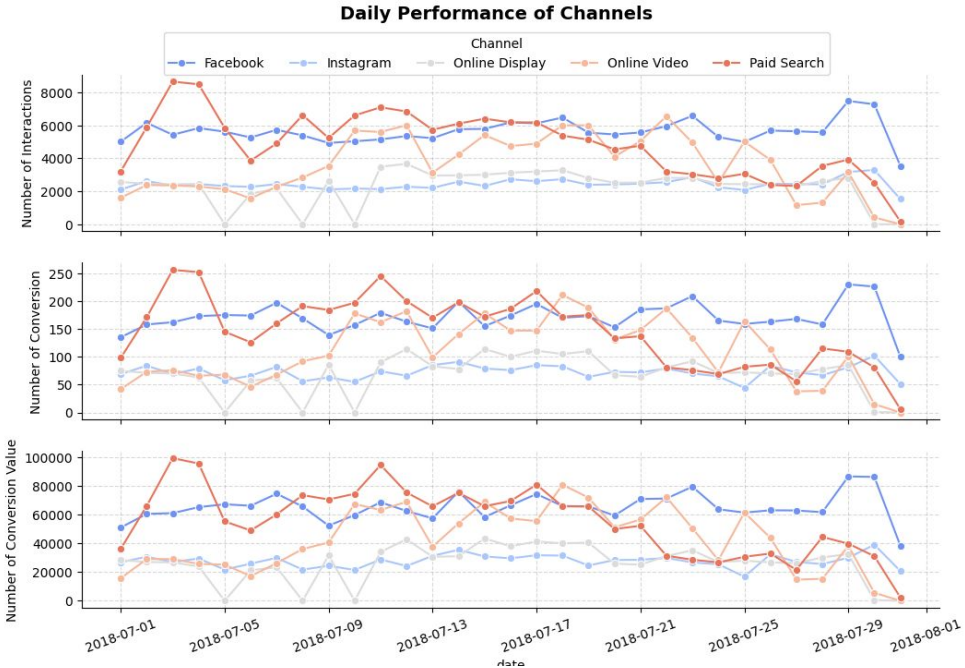
- Conversion rates across all channels are quite similar (around 3%)
- Facebook has the highest number of interactions, conversions and conversion values.

| channel | interactions | num_unique_user | conversion | conversion_value | conv_rate | avg_conv_val | conv_channel_share | conv_value_channel_share |
|---|---|---|---|---|---|---|---|---|
| Facebook | 175,068 | 88,125 | 5,301 | €2,021,357 | 3.03% | €381.3 | 30.1% | 30.1% |
| Paid Search | 150,666 | 88,840 | 4,547 | €1,728,222 | 3.02% | €380.1 | 25.8% | 25.7% |
| Online Video | 110,879 | 40,962 | 3,408 | €1,306,498 | 3.07% | €383.4 | 19.3% | 19.4% |
| Instagram | 75,096 | 49,466 | 2,244 | €856,410 | 2.99% | €381.6 | 12.7% | 12.7% |
| Online Display | 70,883 | 42,604 | 2,139 | €811,208 | 3.02% | €379.2 | 12.1% | 12.1% |
| Total | 582,592 | 240,108 | 17,639 | €6,723,694 | 3.03% | €381.2 | 100.0% | 100.0% |

# Exploratory Data Analysis

## Daily Performance of Channels

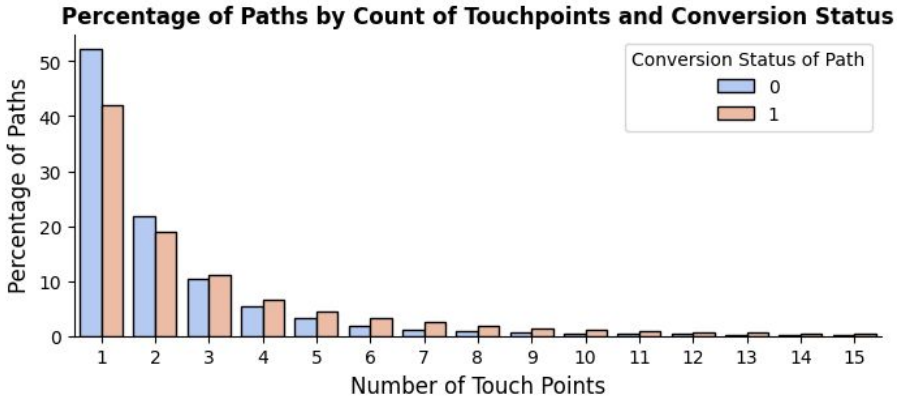
- Facebook and Instagram performance remained stable throughout the month.
- Paid Search traffic and conversions saw a significant decline.
- Online Video traffic peaked mid-month.



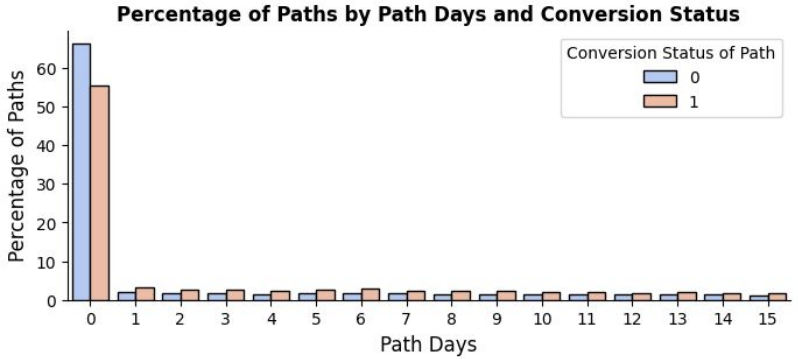
# Exploratory Data Analysis

**Paths**

- Most paths consist of 1 touchpoint (52.3% for no conversion vs. 42% with conversion).
- Conversion rates increase with more touchpoints, rising from 3 touchpoints onward.




- The majority of paths occur within the same day (66% for no conversion vs. 55% with conversion).
- Conversion rates rise with the number of days in a path, starting from 2 days.



# Applied Attribution Model Approach

- Each touch-point is represented in a path.
- The number of paths is equal to the number of unique users, as each user has a single conversion.
- The *'ChannelAttribution'* Python library is used to calculate each channel's contribution, balancing simplicity and effectiveness.
    - *ChannelAttribution.heuristic_models* is used to get 'First-Touch', 'Last-Touch' and 'Linear-Touch'
    - *ChannelAttribution.markov_model* is used to get 'Markov Chains'
- The necessary path data frame is created for the library.
- The results are then compared across each channel.

# Results & Key Findings

- **Facebook** remains the largest contributor **(30%)**, maintaining consistent values across all models.
- **Instagram** gains significant importance **in the Markov Chain model**, rising to become the third most important channel. (20% in Markov Model, %13 in others)
- **Paid Search** and **Online Video** experience the greatest performance decline and **Online Display** has the least impact.
- The Markov Chain model highlights the significance of data-driven and probabilistic attribution models.