



# Lab 1 - Attribution Modelling

Digital and Social Media Strategies

Fall 2024

## **Assignment Information**

This formative assignment is designed to help you build and revise your knowledge of Marketing Attribution Models. This material was discussed in Lecture 1. There are two components:

1. Revision Questions: Short answer questions relating to the lecture content.
2. Hands On: Use R to write and/or run short code snippets to do marketing attribution analysis.

This question document, the dataset for this assignment, any additional information about the data and an R script for you to write your code are available on Canvas as a zip file "lab\_01\_attribution.zip." Download this repository and unzip on your computer in a location where you are keeping files for this class.

We will provide solutions to the coding parts of the assignment via Canvas on Friday after the final Lab Section of the week has concluded.

**You do not need to submit this assignment for grading.**

## **Learning Goals**

By the end of this assignment, you will be able to:

- Use rule based- and data-driven touch attribution models to assign conversion credits to a set of advertisements
- Assess the pros and cons of alternative based attribution models
- Implement attribution models in R, explain their output and compare their relative performance

## Revision Questions

The following information is relevant for Questions 1 through 6. Consider the following customer's journey:



1. Use last touch attribution to assign conversion credits to the ads above.
2. Use first touch attribution to assign conversion credits to the ads above.
3. Identify and explain the 3 shortcomings of the first- and last-touch attribution strategies you implemented above.
4. Use a linear attribution model to attribute the conversions to each ad above.
5. Explain the conceptual differences and similarities between the position-based- and time-decay- attribution models from the linear attribution model you implemented above?
6. Use a position based attribution model to assign conversion credits to the touchpoints above. In your answer explain the weighting strategy you choose and why it is appropriate.
7. How does Shapley Value attribution differ from the attribution models you have described above? Explain intuitively.
8. Use the conversions data on slide 39 of the Lecture to compute the number of Google Ad conversions using the Shapley Value approach.
9. Assess the accuracy of the following statement:

*"Advertising attribution models do a good job of measuring the incremental effect of each advertising channel on sales"*

When making your assessment, be sure to define what the incremental effect is, and explain why you decide this statement is accurate or inaccurate.

# Hands On: Doing Marketing Attribution Modelling

We will work with an advertising attribution dataset. The data contain information from a set of users and a detailed sequence of ads they saw, the date they saw an ad, whether the user converted (i.e. bought something after seeing the ad), and if they converted it details the users spending measured in Euros.

Your task is to put some of the theory learned in the “Marketing Attribution Models” lecture to work to do attribution modelling using the R programming language.

## Data & Variables

The data provided has been artificially generated, but closely reflects what a real dataset a marketer tasked with doing digital advertising attribution would receive.

The variables included in the data are:

- **cookie** - a unique identifier that identifies one device (which we assume is one user).
- **date** - the date that a cookie was exposed to an ad, formatted as YYYY-MM-DD
- **time** - the time of day the ad was seen, formatted as HH:MM:SS
- **interaction** - is the exposure an “impression” or did a “conversion” occur after seeing the ad?
- **conversion** - did the cookie buy something immediately after seeing this ad?
  - 0: No, 1: Yes
- **conversion\_value** - how many Euros were spent when a conversion took place.
- **channel** - what type of ad did the consumer see

We assume that you have R and RStudio installed on your computer. Instructions for doing this are posted on Canvas in the “Computer Labs” module. We will also assume that you have installed the necessary R packages to work through the example. The script that comes with this worksheet has some example code to help you install a package.

The following links provide you with some additional instruction in the form of short videos from YouTube and some notes. These links are also available on Canvas.

- [General Tips on working with R and RStudio](#)
- [Using RStudio Projects](#)
- [Project Oriented Workflows](#)

We also assume that you have worked through Data Carpentry’s “[Data Analysis and Visualization with R for Social Scientists](#)” material (Sections 1 through 6) prior to starting this exercise.

### Task 1: Opening R and Starting an RStudio Project

- A. Open R.
- B. Start a new RStudio project in the folder “lab-01-attribution” that you have already downloaded unzipped (see page 1).
- C. Open the script file “attribution\_analysis.R”.

## **Task 2: Loading Data into R & Developing an Understanding of the Data**

- A. Load the data into R.
- B. How many rows are there in the data?
- C. How many columns are there in the data?
- D. How many unique users are in the data.  
*Hint: Each user has their own identifier in the “cookie” column of the data.*
- E. How many unique marketing channels are there in the data? Print the unique values in the console.
- F. How many unique days are in the data set?

## **Task 3: Selecting Data for Analysis**

- A. Create a new dataset “conversions” that stores only the rows where a user has converted.
- B. How many users converted in the data?  
*Hint: You want to use the result from B combined with the skills you developed in Task 2 to answer this question.*

## **Task 4: Last Touch Attribution – Doing it Manually**

- A. Count the number of conversions attributed to each marketing channel based on last touch attribution.  
*Hint: Use the conversions dataset you created in Task 3*
- B. What is the dollar value of conversions attributed to each marketing channel based on last touch attribution?
- C. Count the number of conversions per marketing channel for each calendar day in the data set based on last touch attribution. Store the results in a dataset called “daily\_conversions\_lta”.
- D. Create a plot that visualises the dollar value of conversions for each channel.

## **Task 5: Attribution Modelling – Using a Package for Marketing Attribution**

The R package “ChannelAttribution” can compute many attribution models for us. These exercises work through how to use the package to get results.

- A. Run the code written in the R script that transforms your dataset to the structure that the ChannelAttribution package needs.
- B. The function `heuristic_models()` computes three rule based attribution models. Run the code, and report which rule based models are returned.
- C. The function `markov_model()` estimates a Markov attribution model. Run the code and report the attributed sales and revenue from the model.

## **Task 6: Shapley Attribution**

There are no packages to compute a Shapley Value based attribution model in R. Lachlan has written a function called ``shapley_attribution()`` that computes conversions per advertising channel.

- A. Run the code to compute attributed conversions.

## **Task 7: Comparing Attribution Modelling Alternatives**

Compare the results from the different attribution models. Which would you choose? Why?