Quilmes Data Science Exercise

Denis Trosman - 30/11/2023

Github repository:

Index

Objective

Create a lift analysis for Quilmes' marketing campaign

The Commercial Intelligence team oversees the creation of marketing campaigns that involve promoting particular products. They reached out to the Data Science team to assess the effectiveness of their marketing initiatives. In order to gauge this, they've requested the calculation of Lift. In this context, Lift represents the percentage increase in sales attributed to the execution of these marketing campaigns, relative to a scenario where no such campaigns were conducted.

$$Lift = \frac{Experiment\ sales\ avg - Control\ sales\ avg}{Control\ sales\ avg}$$

Data

Six datasets were delivered on 30/11/2019

- *Clients*: clients id with channel and region information - (49,458 x 3)
- *Channel segmentation*: channel id information - (38 x 2)
- *Blacklist*: clients that should not be considered - (5,937 x 1)
- *Coupons*: coupons sales and brands information - (805,757 x 5)
- *Sales:* sales amount by month since 2022-01 to 2023-06 - (+1.5M x 5)

All of them were used to construct a solid dataset to develop a lift number for each channel desired.

Analysis and cleansing

After applying the filters below, a final dataset was created for each desired channel.

Drop duplicates

Filter by blacklist

Filter by date > 2023 and +3 months of data

Select desired channels

Keep clients with sales data

Channel	Size	Campaign	No campaign
Tradicional	5,851	3,046	2,805
Autoservicio	6,742	3,761	2,981
Kioscos	1,272	688	584

Results

The marketing campaign had a positive impact on sales for all considered channels

If clients had more than 5 sales using coupons, then they were considered as campaign users. This flag was used to divide experiments and control groups.

Sales average amount was used to build the lift. **The campaign showed a 67% increase on sales average for tradicional channel, 60% for autoservicio, and 37% for kioscos.**

Besides, a t-test statistical check was constructed to confirm the difference between the control and experiment groups.

Channel	Lift	T-test P value
Tradicional	0.67	0.00
Autoservicio	0.60	0.00
Kioscos	0.37	0.00