Attribution and allocation for D2C clothing

Motivation: Our new direct to consumer (D2C) clothing business has been running experiments to understand the best advertising channels for our product. To do this, we've collected a week of data from our website traffic monitoring platform to observe how spend relates to customer conversion (conversion = customer bought the product). For our paid advertising channels we're worried about market saturation: is spending more money on marketing valuable or do we start to see diminishing returns? To test this we ran identical ad campaigns on three different (identical) population samples that only differed by total level of spend: tier 1 ($5,000), tier 2 ($10,000), and tier 3 ($15,000). We're interested in understanding our marketing effectiveness by channel and developing a spending allocation plan for next week's advertising budget.

Instructions

Part 0, optional: Listen to interview with Kevin Frisch, Uber CMO on detecting attribution fraud on behalf of programmatic agency partners: https://www.marketingtodaypodcast.com/194-historic-ad-fraud-at-uber-with-kevin-frisch/

Part 1, attribution: Under five different allocation methods (last touchpoint, first touchpoint, last non-direct touch point, linear, and position-based) calculate the marginal CAC for each of the channels (social, referral, email, paid search, and display) and evaluate effectiveness of each channel.

Partially completed code and data available here:

https://github.com/mallory-archer/Fall2020_attribution_allocation.git

Part 2, allocation: Choose one of the allocation methods and determine how you would spend an ad budget of $11,000 for next week in one target market. (This week we spent $30,000 total on advertising across all platforms running the three experiments in three different markets). Next week we want to allocate the $11,000 budget in $1,000 increments, i.e. $1,000 spend on display adds, $3,000 spend on social ads, etc. Assume that you cannot spend more than $3,000 on any channel. Determine how you would most effectively allocate this budget.

Hint: the data set is large, so as is best practice for data science, I recommend **randomly** selecting observations to test code on while in the development stage, then scaling up the sample until running the code on the full set once you're reasonable assured the code is working as intended

Submission

Submit 5-7 slides as a group to Assignment 3 responding to the analysis request above

On your title slide be sure to include the names of group members and a link to one GitHub repository containing a complete version of the code. Only one version will be reviewed. All are welcome to have the code in their repos – and it's encouraged – but it will not be reviewed. You can work in groups of up to 3 people. Fewer (or individual work) is ok; groups are for your convenience and learning.

Submissions of presentation and code for all groups are due by Friday, 06-NOV-2020 9am Shanghai time.