**Habu Project 1: Optimal marketing mix and budget optimization**

**Project Proposal**

1. **Research Questions:**
2. What is the overall return on investment (ROI[[1]](#footnote-0)) of advertisements in terms of total sales per spending on advertisements over a particular period of time?
3. What is the effectiveness of the various digital marketing channels used in previous marketing campaigns for a chosen company?
   1. Costs on impressions vs. costs on clicks
   2. Sponsored display, sponsored products, sponsored brands, Amazon DSP (Demand-Side-Platform), video ads, etc.
4. What would be the optimized budget allocation for the different digital marketing channels to maximize ROI given an advertising budget?
5. Would the effectiveness of the various digital marketing channels and the corresponding budget optimization strategy be dependent on other factors?
   1. Industry of the company/brand (e.g. necessities vs. luxury goods)
   2. Geographical location of customers (e.g. countries or different US states)
   3. Target audience (e.g. new-to-brand customers vs. repeat customers)
   4. Other demographic factors (depending on availability)
6. **Data Sources**

This project will focus on aggregated data extracted from the Amazon Marketing Cloud (AMC) to study the effectiveness of various digital marketing channels on Amazon and the budget optimization of spending on advertising on Amazon.

Daily/weekly data for a period of 12.5 months for the following variables will be used in a marketing mix model (see **Table 1**). The project will be divided into several phases - starting with only a few variables in a simple model (Phase 1), and gradually moving towards more robust and complex analyses by adding more variables / segmentations (Phase 2) (see **Part C Project Plan** for more details).

**Table 1. Outcome variables, predictors, and control variables in Phase 1 and Phase 2**

|  | **Phase 1** | **Phase 2** |
| --- | --- | --- |
| **Outcome variable** | * Product sales | * Product sales * Number of clicks and impressions * Sales generated from new customers |
| **Predictors** | * Impression cost * Click cost | * Impression cost * Click cost * Split cost by channels   + Sponsored display   + Sponsored products   + Sponsored brands   + Amazon DSP   + Video ads * Bidding price * Event timestamp: hour of the day the ad was shown * Creative setup information |
| **Control Variables** | * Price of the products / discounts | * Price of the products / discounts * Competitor sales * Different segmentations:   + New to brand   + By industry   + By geographical location of customers   + Other demographic factors |

1. **Project Plan**

The main goal of this endeavor is to generate a simple but robust model that can be shared and modified easily. We view replication and explainability as paramount in this project.

We acknowledge the amount of uncertainty that comes from creating statistical models still at an exploratory stage in the market. That is why we divided our project plan into three main stages allowing us to i) gain strong fundamentals in the area we are tackling, ii) develop quickly the simplest model (Phase 1), iii) and finally create a stronger model taking into account rigorous industry requirements in budget optimization (Phase 2). The fourth stage will consist of presenting our results.

1. **Understanding the literature**
   1. Research on Marketing Mix Modeling
   2. Research on Unified Marketing Modeling
   3. Training on Bayesian statistics
   4. Training on statistical modeling with different distributions
2. **Creating our first base model** (Phase 1):
   1. Understanding the data
      1. Selecting variables from the AMC dataset
      2. Exploratory analysis
   2. Setting up the parameters
      1. Outcome variable = Product sales
      2. Predictors = Marketing spend on impressions and clicks, transformed:

* Nonlinear function transformation to account for diminishing returns on spending
* Adstock transformation of channel spend to account for lagged effects of advertising[[2]](#footnote-1)
  + 1. Control variables = Price of the product/discounts
  1. Running the Marketing Mix Model (Google’s LightweightMMM)
     1. Organizing the data for modeling
     2. Checking data quality (checking correlation matrix, checking variances, checking the spend fractions, checking the variance inflation factors)
     3. Training the model
     4. Obtaining media insights (e.g. estimated media and baseline contribution, estimated media channel contribution percentage, estimated media channel ROI, response curves of media channels)
     5. Budget optimization
  2. Validation stage - cross-checking our results with other marketing metrics (e.g. eCPM, eCPC, CPA, etc.)[[3]](#footnote-2)
  3. Interpreting the results
  4. Determining variables to create a more robust model (for Phase 2)
  5. Collecting more data from Habu

1. **Creating our final model** (Phase 2)
   1. Exploring additional Key Performance Indicators (KPIs) as outcome variables - number of clicks and impressions, sales generated from new customers
   2. Selecting the new variables to include and running our MMM following the previous steps from Phase 1 (refer to Phase 2 in **Part B Data Sources** for details)
   3. Running MMM on segmented data according to the different segmentations set out in Phase 2 in **Part B Data Sources**, i.e. new to brand, by industry, etc.
   4. Interpreting the results
2. **Presenting deliverables in a report**

1. One way of calculating ROI is:

   ROI = (Total Ad Revenue - Total Ad Campaign Cost) / # of clicks received for each company [↑](#footnote-ref-0)
2. If Google’s LightweightMMM is deployed, the transformations would be automatically carried out by specifying model\_name = “Adstock” / “Hill-Adstock” / “Carryover”. In this case, we would compare the performances of the three approaches and choose the best performing approach for the remaining analyses. [↑](#footnote-ref-1)
3. Reference: <https://headerbidding.co/calculate-cpm-cpc-cpa-ecpm-ecpc-ecpa-roi/> [↑](#footnote-ref-2)