

Marketing Analytics

Group Assignment 1

April 21, 2021

INSTRUCTIONS

- This is a team assignment. Teams must have 4 or 5 students and all team members will receive the same grade.
- Submit your proposal digitally as two files:
 1. A **PDF file** with a report that answers all the questions. The file name should follow the convention: **GA1_GroupXX.pdf** in which **XX** corresponds to the two digits that identify your team.
 - There is no formal page limit for this document. You should answer each of the questions in a **complete but concise** way.
 - Pay attention to the reader. This means you should consider which outputs are important to show to make your point (and answer the question) and which outputs are better left out of the document.
 2. An **R** or **Rmd** file containing a script that contains all the analyses you performed on the data to answer the questions. The analyses for each of the question should be properly identified. The file name should follow the convention: **GA1_GroupXX.R** (or **GA1_GroupXX.Rmd**) in which **XX** corresponds to the two digits that identify your team.

Deadline: Tuesday, April 27 at 23:59.

- Late submissions are not allowed

Warning: The detection of any form of plagiarism in your work means the assignment will be graded with ZERO points.

Analyzing an Email Campaign

Congratulations! You have been hired by Books&Tech, an upcoming online store focused on selling books and technology. Your first task is to assess the performance of an email marketing campaign targeting customers that have performed at least one purchase in the last 12 months in the store.

The campaign had been designed and deployed before you were hired. Your boss has asked you — a marketing analytics specialist — to evaluate both the design of the experiment and the effectiveness of the campaign.

Exploratory Data Analysis and Randomization Checks

Your first task is to better understand the data. You should perform the required exploratory data analyses so that you are confident about the nature of the data and about the nature of the randomization.

1. [4 points] You should start by providing a set of initial descriptive statistics and graphical analyses that can help others better understand your customers. For this question use only data from **before the email campaign**, i.e., only variables that have the role of predictor. You can present summary statistics on the data, some correlational analyses between variables you may be interested in and any potential anomalies or strange findings.
2. [3 points] Next, you should perform randomization checks, i.e., check whether the randomization was adequately performed. In other words, you should answer the question: are customers assigned to each of the treatment conditions comparable? Reflect on the obtained results.

Analysis of the Experiment Design and Results

After the initial exploratory analyses and randomization checks it is time to assess the results of the experiment and assess its design. Please answer each of the following questions.

3. [3 points] Which campaign performed the best overall, the Books version, or the Electronics version?
4. [3 points] How much incremental sales per customer did the Books version of the campaign drive? How much incremental sales per customer did the Electronics version of the campaign drive?
5. [3 points] Is cross-selling a good strategy for Books&Tech? In other words, which audience would you target the Books version to, and the Electronics version to, given the obtained results? Justify your recommendation.
6. [4 points] If you were asked to redesign the experiment, what would be your recommendations? Think of (at least) the following aspects: metrics used, size of the experiment, (number of) treatments, duration.

Information about the campaign

Books&Tech randomly assigned 45,572 customers that purchased in the last 12 months to one of three conditions:

- customers who received an email campaign featuring books.
- customers who received an email campaign featuring electronics.
- customers who did not receive an email campaign.

The results of the campaign were tracked for a period of four weeks. The table below contains a brief description of the obtained data.

VARIABLE	ROLE	TYPE	DESCRIPTION
<code>last_purchase</code>	predictor	numeric	Months since last purchase.
<code>hist_spend</code>	predictor	numeric	Actual dollar value spent in the past year.
<code>books</code>	predictor	binomial	1/0 indicator, 1 = customer purchased from the books section in the past year.
<code>electronics</code>	predictor	binomial	1/0 indicator, 1 = customer purchased from the electronics section in the past year.
<code>pop_density</code>	predictor	multinomial	Use location classified as Urban, Suburban, or Rural.
<code>new_customer</code>	predictor	binomial	1/0 indicator, 1 = New customer in the past twelve months.
<code>device</code>	predictor	multinomial	Describes the devices the customer purchased from in the past year.
<code>treatment</code>	manipulation	multinomial	The e-mail campaign the customer received.
<code>visit_after</code>	outcome	binomial	1/0 indicator, 1 = Customer visited website in the following four weeks.
<code>conversion_after</code>	outcome	binomial	1/0 indicator, 1 = Customer purchased in the following four weeks.
<code>spend_after</code>	outcome	numeric	Actual dollars spent in the following four weeks.