

TBR+Matched Markets design colab guide

Author: Marco Longfils

The purpose of this document is to guide you step by step in the process of designing a Matched Markets experiment using the [Python library](#) developed at Google. In the following we will use the acronyms TBR for Time Based Regression and MM for Matched Markets. For a general introduction to TBR and MM, please refer to the [TBR paper](#), the [MM paper](#), and this [introduction](#) to geo experiments. The colab can be found [here](#). **At the end of this document you can find a [glossary](#) of the main terms in this document.**

[Data needed to run the Colab](#)

[client_sales_table](#):

[geo_eligibility_table](#):

[Cell by cell description of the colab](#)

[Data Input](#)

[Cell 1](#)

[Cell 2](#)

[Design with TBR+MM](#)

[Cell 3](#)

[Cell 4](#)

[Cell 5](#)

[Summary of the design and save the pretest data, the geo pairs, treatment and control stores in a trix.](#)

[Cell 6](#)

[Cell 7](#)

[Glossary](#)

Data needed to run the Colab

In order to run the colab, you will need the following table

client_sales_table:

Description: this table is the table containing the response and spend data at the geo level and at daily/weekly frequency. The table should be specified in the colab by using the URL of the google sheet.

Format: Google sheet containing the following mandatory columns: date, response, cost, geo. Any additional column is ok.

Example: [here](#)

geo_eligibility_table:

Description: this table is optional and can be used to constrain the assignments of each geo. The table should be specified in the colab by using the URL of the google sheet.

Format: Google sheet containing the following mandatory columns: geo, control, treatment, exclude. This table is used to constrain the assignment of the geos to the different groups. For example, if the input table looks like the following

geo	control	treatment	exclude
10	0	1	0
11	1	1	1
12	1	1	1
14	0	0	1
15	1	1	1
16	0	1	1

It means that:

- Geo 10 must be included in the experiment, and in particular assigned to the treatment group. This geo CANNOT be excluded or assigned to control.
- Geos 11, 12, and 15 can be assigned to control, to treatment, or it can be excluded from the experiment.
- Geo 14 must be excluded from the experiment.
- Geo 16 can be included in the treatment group or excluded, but it CANNOT be assigned to control.
- Geos that are not listed in the table can be assigned to any group (control, treatment, or excluded, which is equivalent to the case where the geo is listed in the table with values 1,1,1).

Example: see image above.

Cell by cell description of the colab

Data Input

Cell 1

[\[View in colab\]](#)

Description: This cell will install and load the necessary libraries for the Matched Markets library. It can take some time to run this cell.

Action needed: run the cell.

Cell 2

[\[View in colab\]](#)

Description: This cell will load the data needed to run the analysis. You will be asked for permission to access your file in Drive. Follow the instructions on screen which appear in the cell output. You will need to click a link which will open a new tab in the browser and ask access to Google Drive. You will get a verification code which you can copy and paste in the corresponding field in the colab, see below.

Go to the following link in your browser:

https://accounts.google.com/o/oauth2/auth?response_type=

Enter verification code:

The `geo_eligibility_table` is optional, and you need to specify this table only if you want to constrain the assignments for some geos. For example, you need to specify the `geo_eligibility_table` if you want to force the London geo to be in the treatment group.

Action needed: Before running the first cell, make sure to update the Google Sheet url.

Design with TBR+MM

Cell 3

[\[View in colab\]](#)

Description: Select the parameters to be used in the design of an experiment with TBR+MM. Below you can find a short description of all the necessary parameters:

- **confidence_level** indicates the confidence level for testing the hypothesis $H_0: iROAS = 0$ vs. $H_1: iROAS > 0$. The value of this parameter should be a real number between 0 and 1. The default value is 90%.
- **power_level** indicates the power that we would like to achieve in our design. The value of this parameter should be a real number between 0 and 1. The default value is 80%.
- **experiment_duration_in_weeks** length of the experiment measured in weeks. For example, set this parameter to 4 if you want to design an experiment lasting 4 weeks.
- **experiment_budget** is the maximum budget the client is willing to spend on the campaign.
- **alternative_budget** is a list of budgets, separated by a comma, that can be used to “override” the automated budget choices based on the **experiment_budget**. The default value is an empty field.

Example

alternative_budget: " 125000, 150000, 175000"

- **minimum_detectable_iROAS** is defined as the smallest iROAS that the experiment will be able to reliably measure, given the chosen **power_level** and **confidence_level**. For example, if the **confidence_level** used is 90%, and the **power_level** is 80%, then the **minimum_detectable_iROAS** is the smallest value such that we have 80% chance of detecting an effect of that size with a 10% probability of a false positive. The default value of the minimum detectable iROAS is 3, but should be changed depending on the business vertical of the client. For example, for grocery stores we recommend a value of 2.
- **average_order_value** is the mean value in dollars that the client attributes to a transaction/visit/contract. Use a value of 1 if the design is based on sales/revenue or an actual average order value (e.g. 80\$) for a design based on transactions/footfall/contracts. The default value is 1. **N.B.: make sure to use a value of 1 if the response variable of interest is revenue/sales.**

Action needed: update the value of the parameters to the desired levels.

Cell 4

[\[View in colab\]](#)

Description: Summary of the designs given the parameters specified in Cell 3. Each row in the table corresponds to a design in the output of this cell, see the figure below for an example. Orange rows/cells indicate that one or more conditions are not met in the design, e.g. the minimum detectable iROAS of the design is larger than what was set in input. Green rows correspond to the optimal and recommended designs, while rows in beige contain designs which are feasible but require a higher budget. Other metrics may be highlighted in red if they don't meet some criterion (e.g. if the minimum detectable lift in response is too large)

	Budget	Minimum detectable iROAS	Minimum detectable lift in response	Treatment/control/excluded geos	Revenue covered by treatment group	Cost/baseline response	Cost if test budget is scaled nationally
Design							
0	111K	3.75	5.62 %	3 / 6 / 3	5.18 %	1.50 %	1.9M
1	139K	3.0	5.62 %	3 / 6 / 3	5.18 %	1.87 %	2.37M
2	167K	2.5	5.62 %	3 / 6 / 3	5.18 %	2.25 %	2.85M
3	125K	3.34	5.62 %	3 / 6 / 3	5.18 %	1.68 %	2.14M

Note: if a design does not pass one of the diagnostic checks and so does not satisfy one of the assumptions of TBR, we will print a WARNING message in red with some information about the problem, e.g.

WARNING: the design does not pass the A/A test. We do not recommend using the proposed design.

Action needed: run the cell.

Cell 5

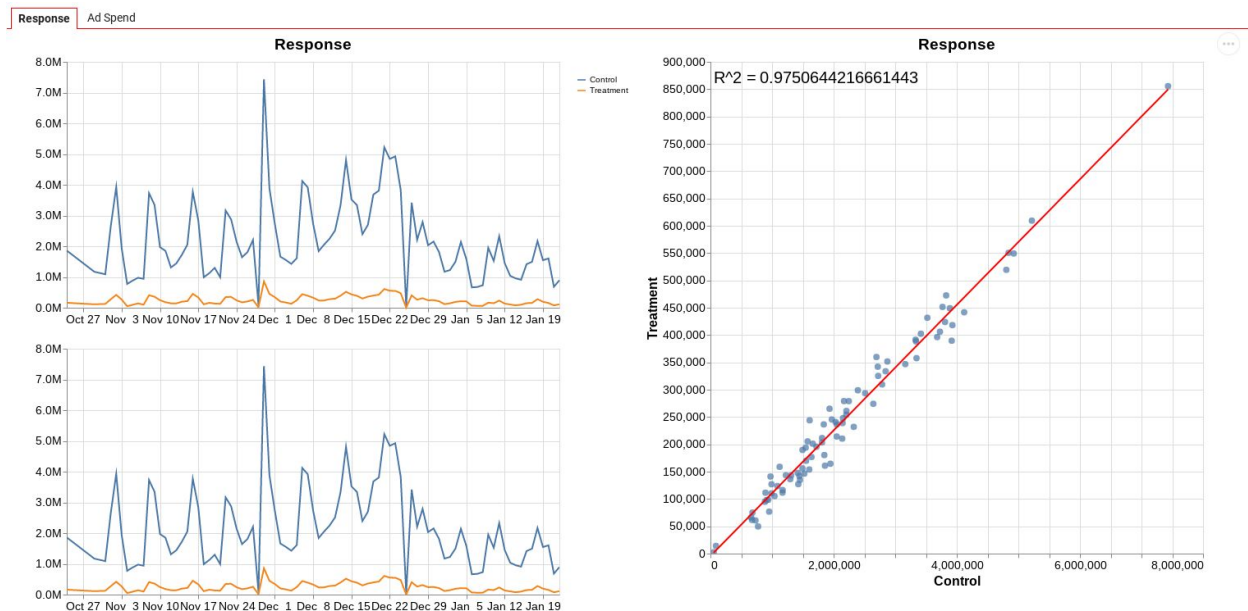
[\[View in colab\]](#)

Description: Select the design that you choose as the final design to be used in the experiment. The design number is the number displayed in the table from Cell 4 as shown in the figure below. The default design is Design 1, which outputs the least expensive design based on the minimal iROAS and power given in input.

	Budget	Minimum detectable iROAS	Minimum lift
Design			
0	126k	6.25	
1	158k	5.0	
2	189k	4.17	
3	1.25M	0.631	

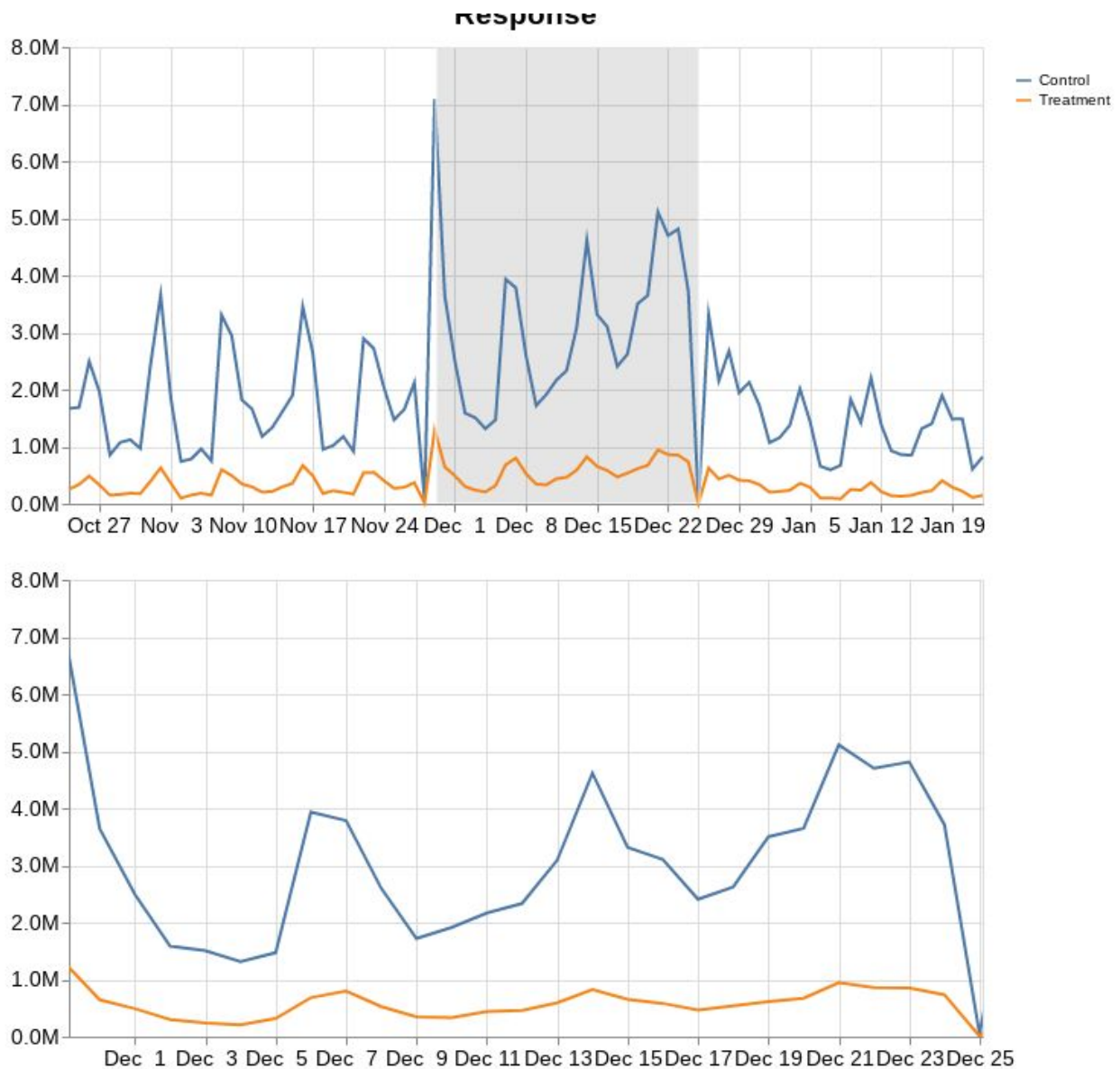
For the chosen design, the aggregated (treatment vs control) time series of the response and cost variables will be shown. The results are presented in two different tabs, one for the response and one for the cost variable. The title of each tab indicates which metric (response or cost) is shown. The scatterplot on the right can give indication of the goodness of the

match between the two groups (control on the x-axis, treatment on the y-axis). Each dot is a single day, the red line is the regression line between treatment and control, and in the top left corner we highlight the R squared.



The top left plot is interactive:

- If you move the cursor on a point on the time series, we will show the value and the day for that point.
- If you select a time window on the top plot, we will show a zoomed in version of the plot in the bottom figure, see below.
- Once a time window is selected, you can shift forward or backward the window by drag and drop.
- Once a time window is selected, you can increase or decrease the size of the window with the mouse wheel (while your mouse cursor is on the gray shaded area).



Action needed: select the design and run the cell.

Summary of the design and save the pretest data, the geo pairs, treatment and control stores in a Google Sheet.

Cell 6

[\[View in colab\]](#)

Description: print the summary of the selected design with information about expected minimum detectable iROAS, and minimum detectable lift as %. This is a useful copy and pasteable summary which can be sent in an email.

Data in input:

- 12 Geos

Output:

The output contains two lists of geos: one for treatment and the other for control

- 3 Geos in treatment and 6 geos control

Baseline store response: \$7.42M for treatment

Cost/baseline = \$111K / \$7.42M ~ 1.5%

Minimum detectable iROAS = 3.75

Minimum detectable lift in % = 5.62 %

The design has Power 80.0+% with Type-I error 10.0% for testing $H_0: iROAS=0$ vs $H_1: iROAS \geq 3.75$

Action needed: run the cell.

Cell 7

[\[View in colab\]](#)

Description: save the design, including the geo-level pretest data (response+cost), treatment and control geos, in the Google Sheet with name specified as input. The Google Sheet is saved in your Google Drive under “My Drive”.

Action needed: Change the name of the file and run the cell.

Glossary

AA test: an AA test can be thought of as an AB test or an experiment in which there is no difference between the two groups (treatment and control). Such a test can be used as a diagnostic tool. Since we know that there is no difference between the groups, if we find a statistically significant difference, it may indicate that there is some confounding factor in the design which would compromise the actual experiment.

Colab: Data analysis tool that combines code, output and descriptive text into one document. You can create a template for data analysis and let other people smoothly run it (as we do for GeoX design and analysis).

Design dataset : this is a synonym for the **client_sales_table**, i.e. the dataset containing the time series of response and cost at the geo level during the time period preceding the experiment.

Geo: region which can be targeted by a campaign and used as the unit during the experiment. Each geo can be treated, controlled, or excluded from the experiment. Examples of geos are the standard DMA (by Nielsen in US) or GMA (Generalized Marketing Areas).

Pretest data: this is a synonym for the **client_sales_table**, i.e. the dataset containing the time series of response and cost at the geo level during the time period preceding the experiment.