

siMMuLator

AN OPEN SOURCE METHOD TO GENERATE DATA FOR MMMs



Agenda

- 01 Background on Marketing Mix Models (MMMs)
- 02 siMMULATOR: A Solution to MMM Validation
- 03 Example of siMMULATOR in Action

01 Background on Marketing Mix Models (MMM)

What are MMMs?

Marketing Mix Models (MMM) are data-driven statistical analysis tools used by marketers to understand what the ROI of their advertising spend on various channels are. This helps inform future investment decisions.

MMMs:

- Are built on aggregate data -- so they are privacy-friendly and resilient
- Work for all businesses and are highly scalable

MMMs usually incorporate data for many advertising channels in addition to non-advertising factors such as macroeconomic trends.



Despite being commonly used, validating MMMs is very difficult because:

Lack of ground-truth data

In a typical model, we would compare the model's predictions to actual data. However, in an MMM the "true" ROI of each advertising channel is not known, so advertisers do not have a comparison to gauge model performance.

Time series data

Since MMMs use time-series data, this means that future predictions are tied to values from the past. Since the different data points are connected to one another, it makes it difficult for us to decide which data to select for our hold out data set.

Requires many years of data

MMMs are usually built on weekly or monthly sales data. Newer advertisers may not have sufficient data points to make an accurate model or do model validation.

One solution to this problem is to create a simulated data set.

02 siMMuMulator: A Solution to MMM Validation

An open-source R-package that lets users simulate data from scratch.

This is how it works:

1. Users input characteristics about their business and their context.
2. siMMMulator adds statistical noise on top of these inputs
3. Then, siMMMulator simulates ad activity and aggregates data into a format for MMMs.

The result is a data set where we know what the ground truth ROI is. We can use this data to see how closely our MMMs can recover our ground-truth ROIs.

Advantages of siMMMulator



Private

All data generated is not real. All inputs to the generation and data generated are done on the users' local computer. No data is sent anywhere.



Flexible

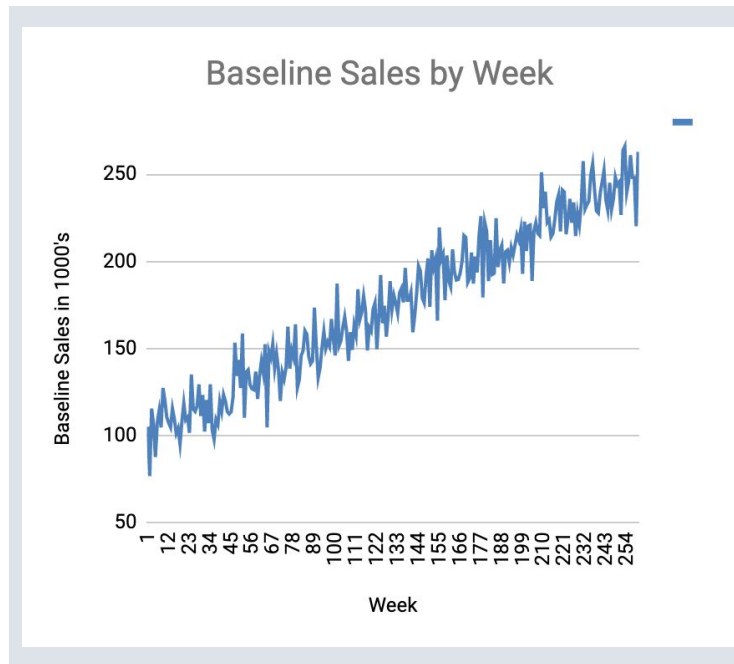
Users can simulate data that best matches their business scenario or test out hypothetical scenarios.



True ROI

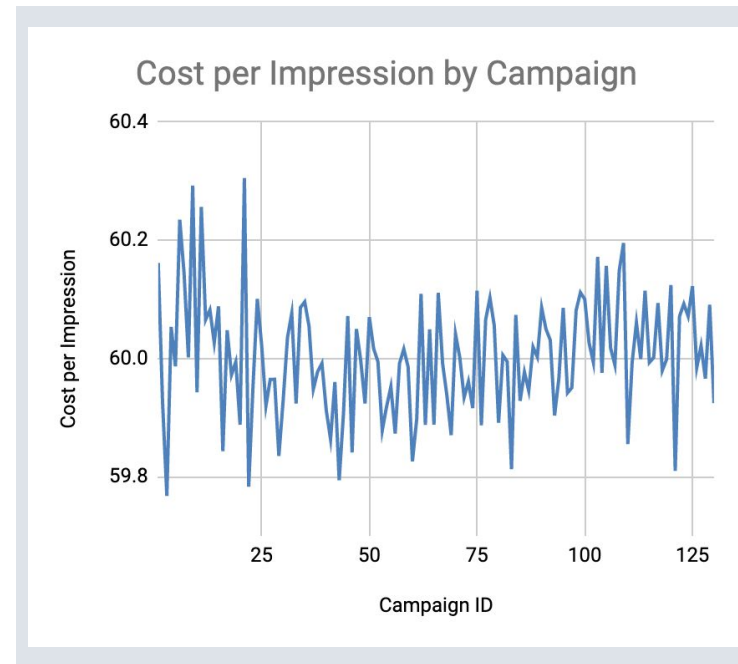
Ground-truth outcomes are known because we provide the inputs to generate true ROI.

Some inputs that users can change:



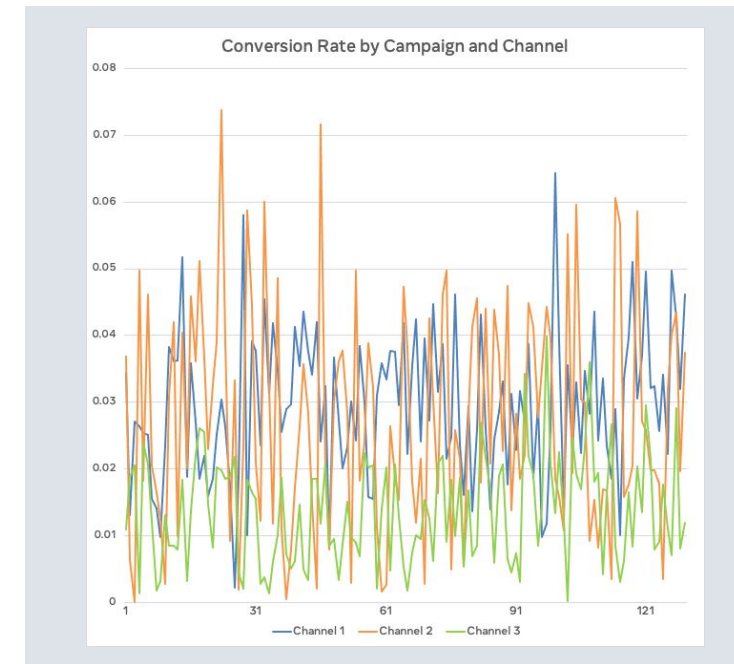
This graph is for illustration purposes only. Individual results may vary.

Baseline sales and its growth over time



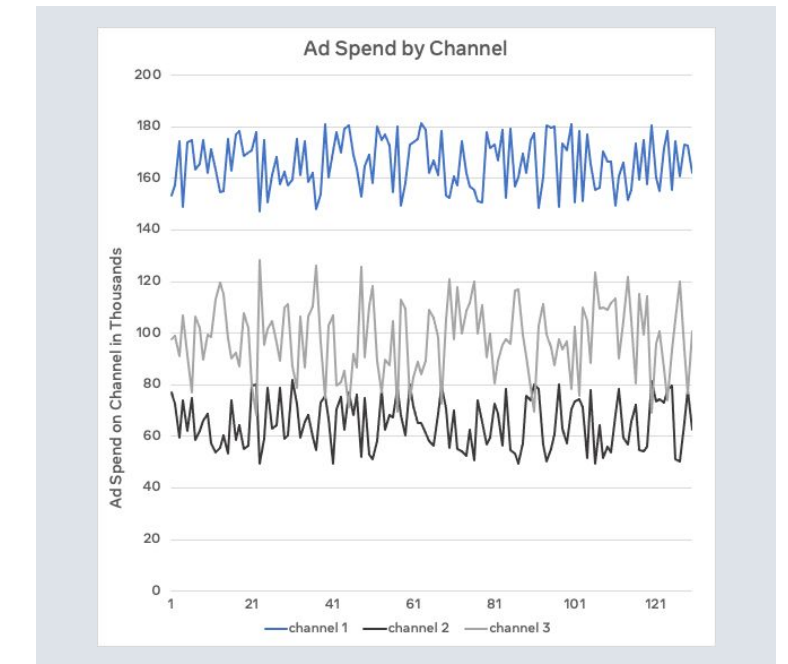
This graph is for illustration purposes only. Individual results may vary.

Average CPM or CPC on a given channel



This graph is for illustration purposes only. Individual results may vary.

Average conversion rate on a given channel



This graph is for illustration purposes only. Individual results may vary.

Spend on a given channel and campaign

siMMMulator provides a data set with ground-truth, enabling users to:

1

Validate and compare accuracy of various MMMs

2

Quantify the value of a particular innovation to MMMs

3

Simulate various scenarios and see their impact on MMMs (e.g. How sensitive is our MMM to changes in CPAs, CVRs, etc.?)

Limitations of siMMMulator

Dependence on users' inputs

How closely siMMMulator reflects reality depends on users' inputs and assumptions. If the user provides inputs that are very different from reality, the resulting simulation will give a good view of accuracy for the given inputs but may not match reality.

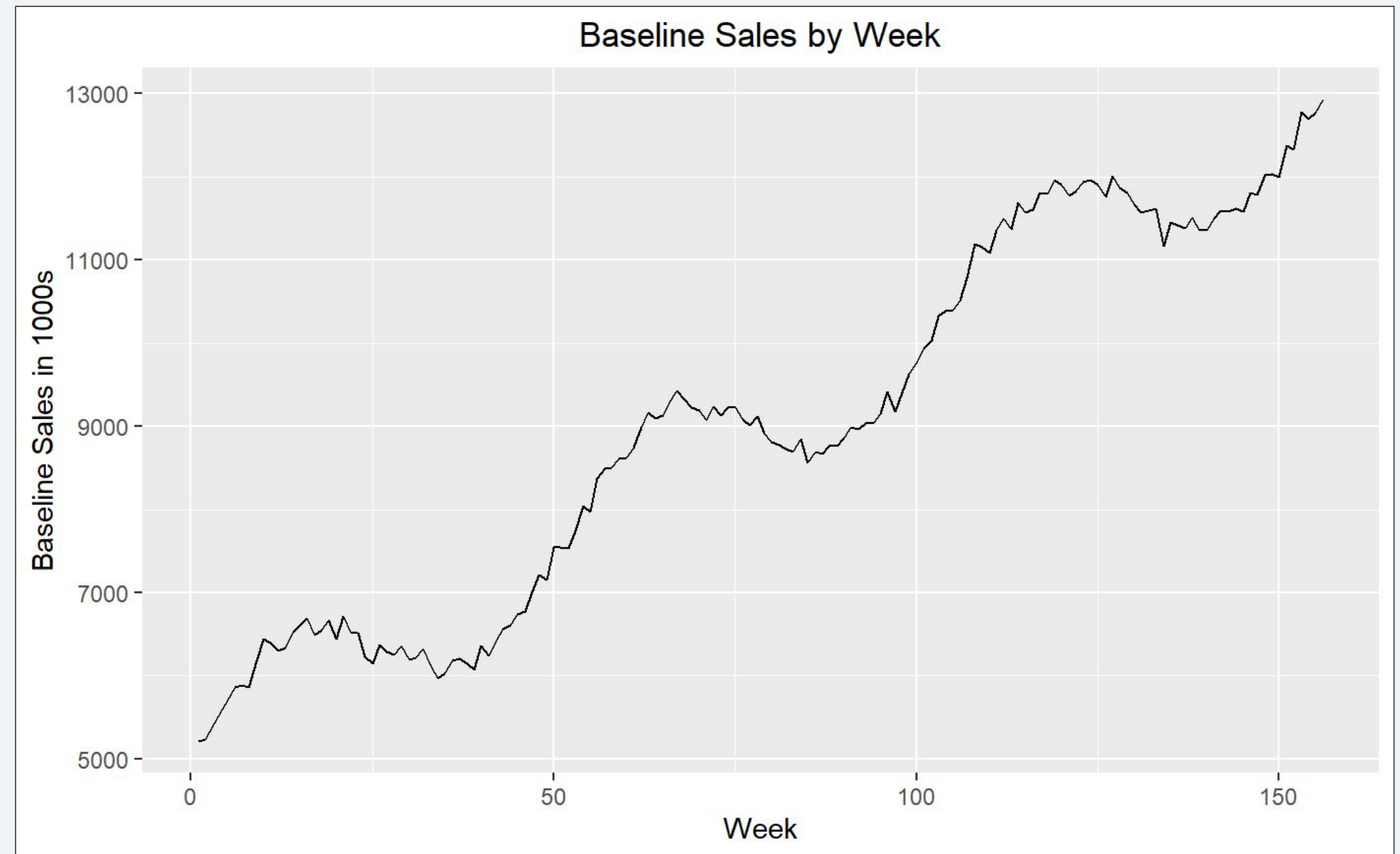
Limited advanced features

Users may have to code for themselves more advanced features (e.g. interaction terms between variables, some weeks where media is off)

03 Example of siMMULATOR in Action

Example of siMMMulator in action

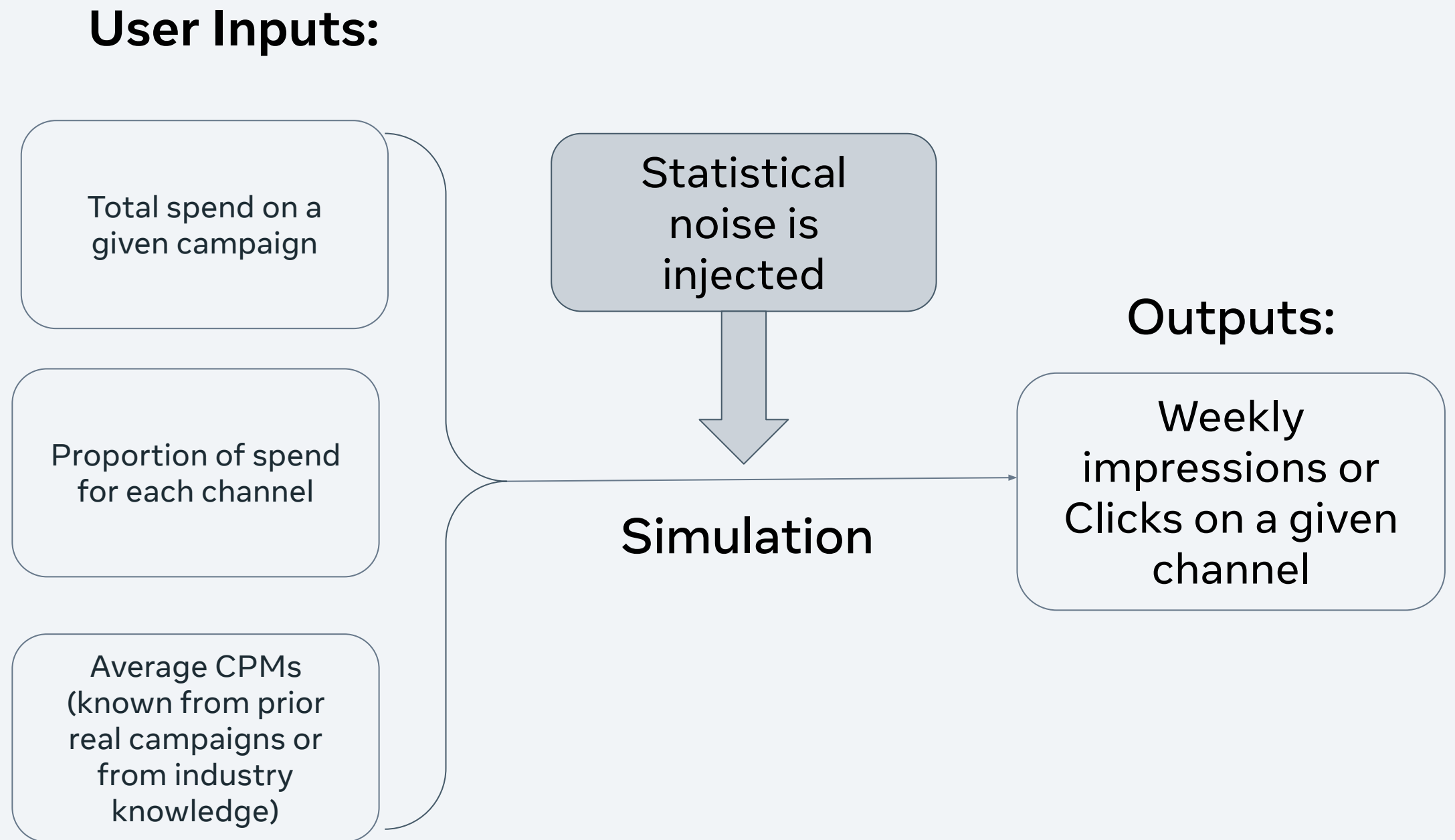
- 1 Users input in quantities (number of years, trend of sales, strength of seasonality) to simulate baseline sales



This graph is for illustration purposes only. Individual results may vary.

Example of siMMMulator in action

2-3 Users input in quantities to simulate ad spend and ad activity



The chart on the right has simplified the process for illustrative purposes.

We follow a similar procedure to simulate the rest of the data:

4

Simulate
Conversion
Rates

5

Transform
Media Variables
(Apply adstock
and diminishing
returns)

6

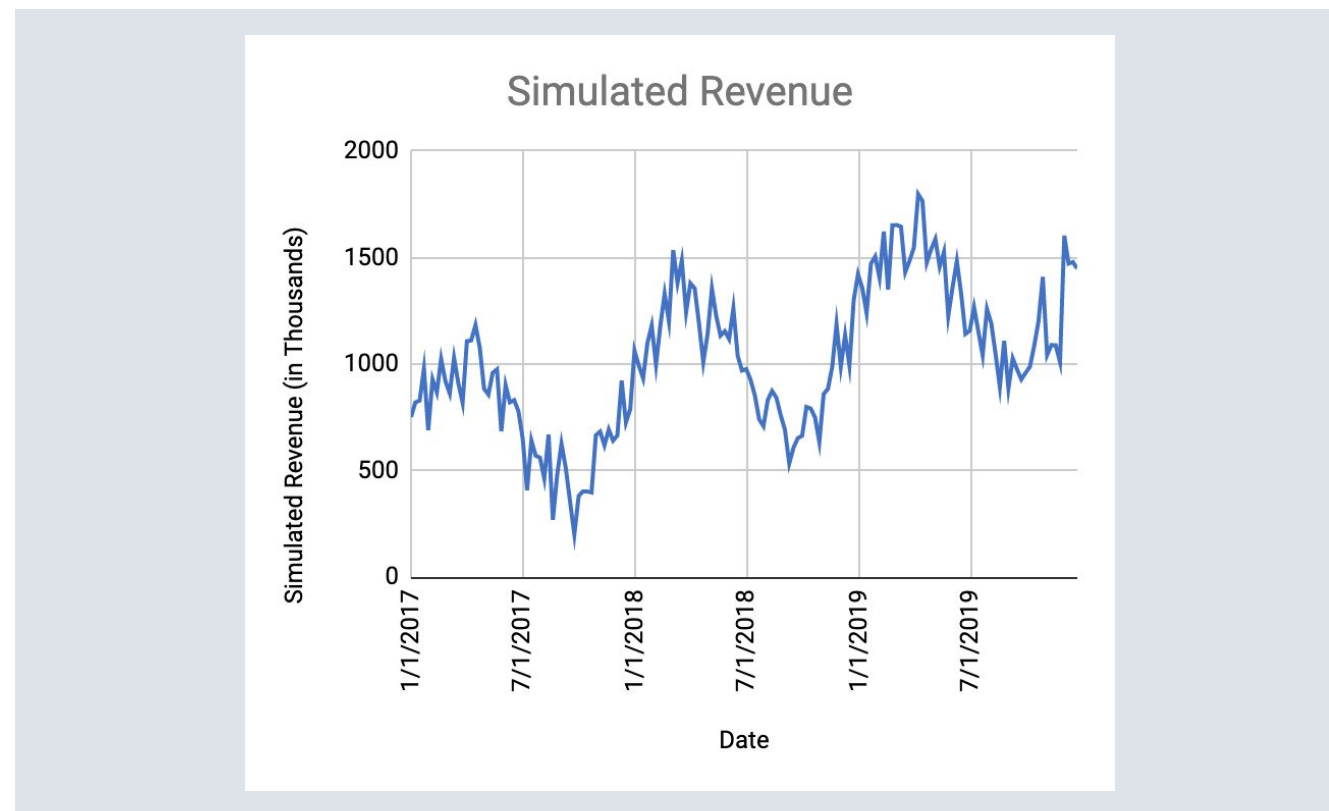
Calculate # of
conversions

7

Aggregate data
and calculate
ground-truth
ROIs

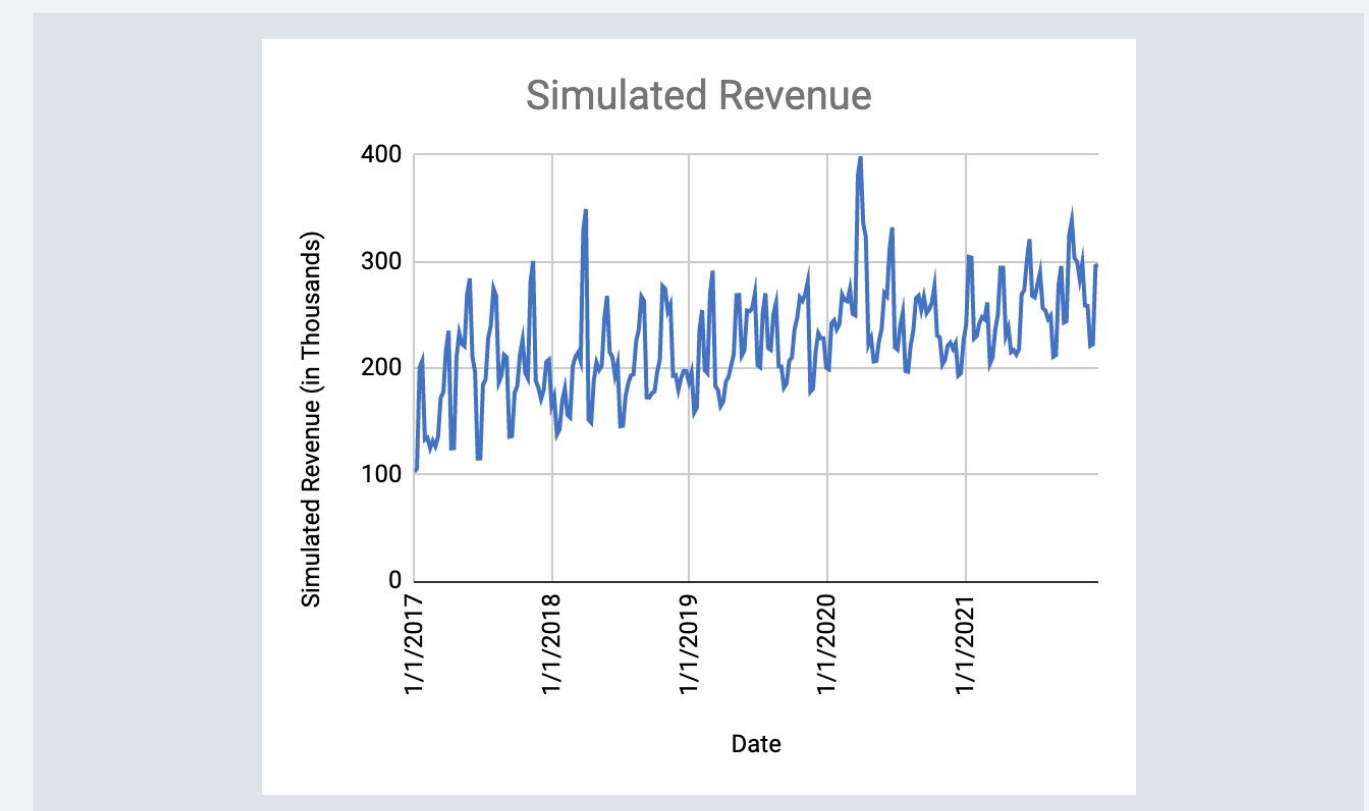
03 EXAMPLE OF SIMMMULATOR IN ACTION

Different Data Sets that Can be Generated



This graph is for illustration purposes only. Individual results may vary.

We can get different data sets by putting in different inputs. For example, the data set on the left is more driven by seasonality than the data on the right.



This graph is for illustration purposes only. Individual results may vary.

03 EXAMPLE OF SIMMMULATOR IN ACTION

Building community: This is an open-source project. We are hoping to work with the industry to make the code better. Please submit requests for new features, bug reports, and any code that you created to use with siMMMulator that you found useful.

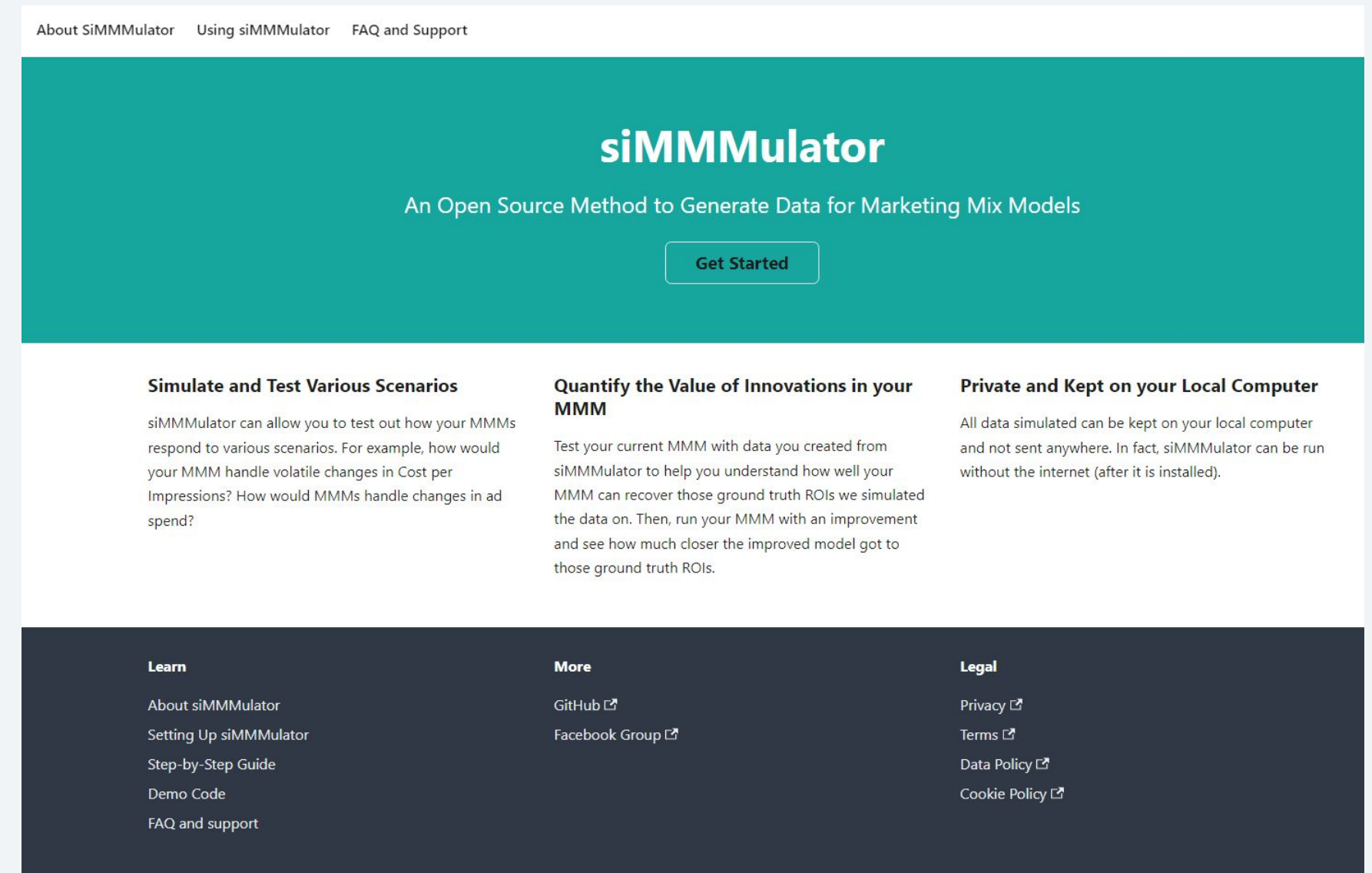
Getting started: Visit our website and download the R-package to get started.

Microsite with Quick Start Guide:

<https://facebookexperimental.github.io/siMMMulator>

siMMMulator Github Repository:

<https://github.com/facebookexperimental/siMMMulator>



The screenshot shows the homepage of the siMMMulator project. At the top, there is a navigation bar with links for "About SiMMMulator", "Using siMMMulator", and "FAQ and Support". The main header features the project name "siMMMulator" in a large, bold font, followed by the tagline "An Open Source Method to Generate Data for Marketing Mix Models" and a prominent "Get Started" button. Below the header, the page is divided into three columns of content:

- Simulate and Test Various Scenarios:** This section explains that siMMMulator allows users to test how their MMMs respond to various scenarios, such as changes in Cost per Impressions or ad spend.
- Quantify the Value of Innovations in your MMM:** This section describes how users can test their current MMM with data generated from siMMMulator to understand how well it can recover ground truth ROIs, and how an improved model can get closer to those ground truth ROIs.
- Private and Kept on your Local Computer:** This section highlights that all simulated data is kept locally and not sent anywhere, emphasizing privacy and offline functionality.

At the bottom of the page, there is a dark footer with three columns of links:

- Learn:** About siMMMulator, Setting Up siMMMulator, Step-by-Step Guide, Demo Code, and FAQ and support.
- More:** GitHub and Facebook Group (both with external link icons).
- Legal:** Privacy, Terms, Data Policy, and Cookie Policy (all with external link icons).

siMMMulator website homepage



THANK YOU