

Using Simulated Data to Validate Marketing Mix Models

MSI 2023 Analytics Conference

 **siMMulator**

 **Meta Open Source**



Jessica Nguyen

Quantitative Researcher, Meta

Agenda

Resurgence in Marketing Mix Models

Validating Marketing Mix Models and
its challenges

Using Simulated Data

Marketing Mix Models (MMMs) have seen a resurgence.

The Rebirth of Marketing Mix Models

-epam

MMMs have seen a resurgence because of:

Why Marketing Mix Modeling is Making a Comeback

-Kinesso

- Increased privacy concerns
- Concerns over causality of some attribution methods
- Increased ability to build models faster and incorporate machine learning

Beyond Multi-Touch: The Resurgence of Marketing Mix Modeling

-bounteous

Source:

“Beyond Multi-Touch: The Resurgence of Marketing Mix Modeling,” bounteous, October 13, 2022, <https://www.bounteous.com/events/2022/10/13/beyond-multi-touch-resurgence-marketing-mix-modeling>

“Why Marketing Mix Modeling is Making a Comeback,” Kinesso, August 5, 2022, <https://kinesso.com/why-marketing-mix-modeling-is-making-a-comeback/>

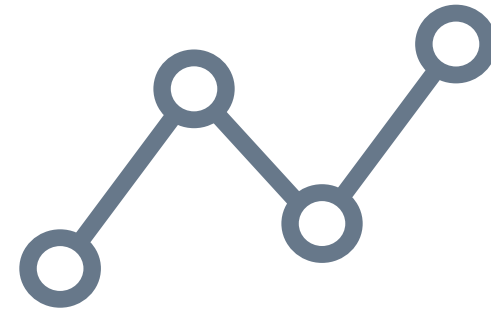
“The Rebirth of Marketing Mix Models,” Francisco Marco-Serrano & Marco De Nooijer, August 5, 2022, <https://www.epam.com/insights/blogs/the-rebirth-of-marketing-mix-models>

MMMs are commonly used, but validating them is difficult because ...



Lack of ground truth data

Advertisers may have limited data with ROIs of various channels that predictions can be compared to



Time series data

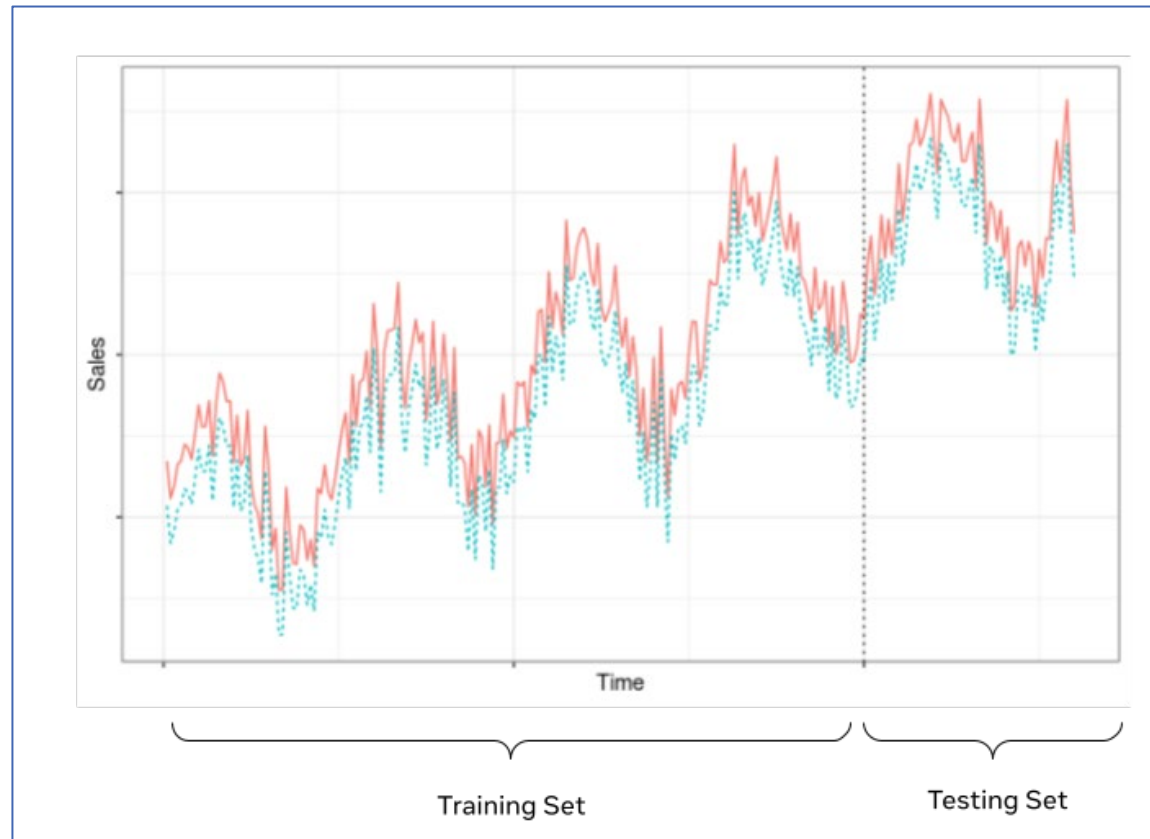
Since past data influences subsequent data, it's hard to pick the right time period for a holdout



Requires many years of data

Newer advertisers may not have sufficient data points to make an accurate model or do model validation

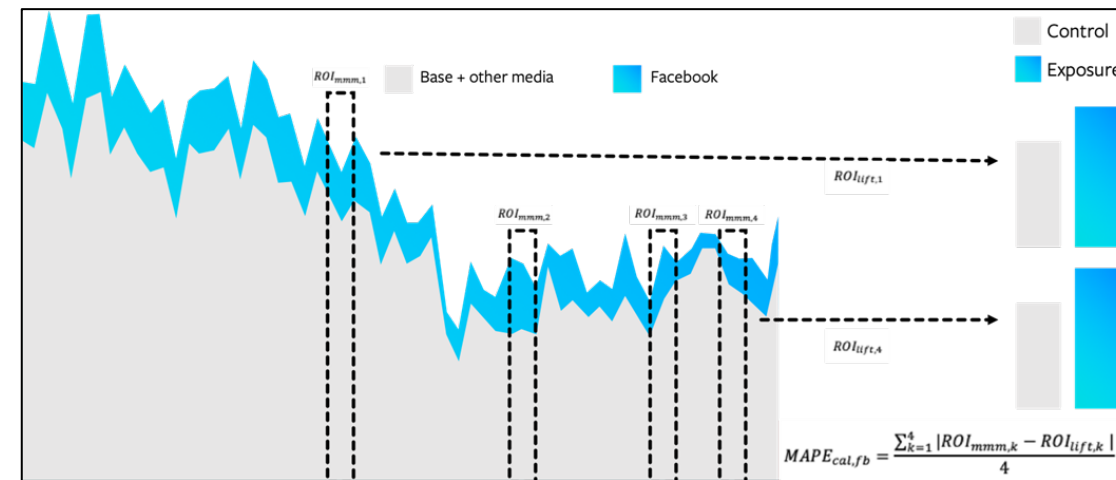
Some methods used to validate MMMs:



Backtesting

Model is trained on a portion of the data and validated on the other portion of data

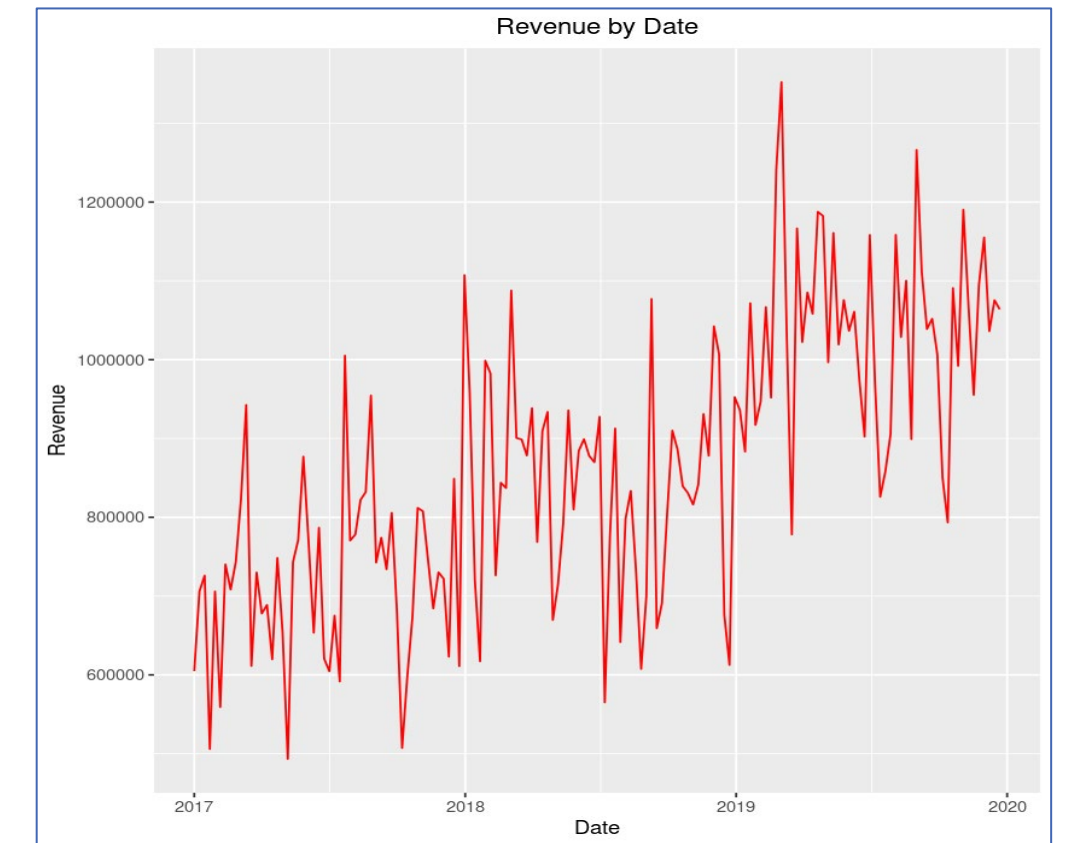
VERY COMMON



Calibrating with experiments

Compare how well models line up with real life experimental results

LESS COMMON



Testing on simulated data

Use assumptions to make data, add some statistical noise, see if our model can recover our assumptions

MSI's MMM initiative: an industry-academic blue ribbon panel to validate and certify MMMs

- An effort between academics and industry partners
- Aims to come up with a way to validate practices and innovations most important to the MMM industry
- MSI firms who wish to participate will contribute to the study financially, by contributing data sets, and/or by asking their suppliers to participate in the initiative
- Firms who participate can get insights, networking, and data

More information [here](#), talk to MSI staff to participate

What is simulated data?

- Simulated data is data that we make up
- Some ways you could simulate data:
 - 1) Take some existing data and add a lot of statistical noise to it
 - 2) Start with some basic assumptions, add noise



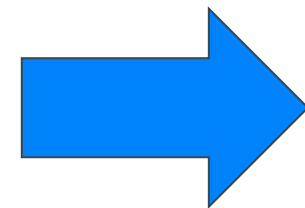
Why is simulated data useful?

Since we created the data, we know what a model or code is supposed to output

Simple example:

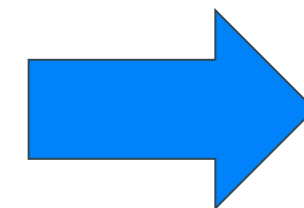
“The cow jumped
over the moon.”

Our “simulated” data



Count the number of **e**'s
in a sentence

Our algorithm

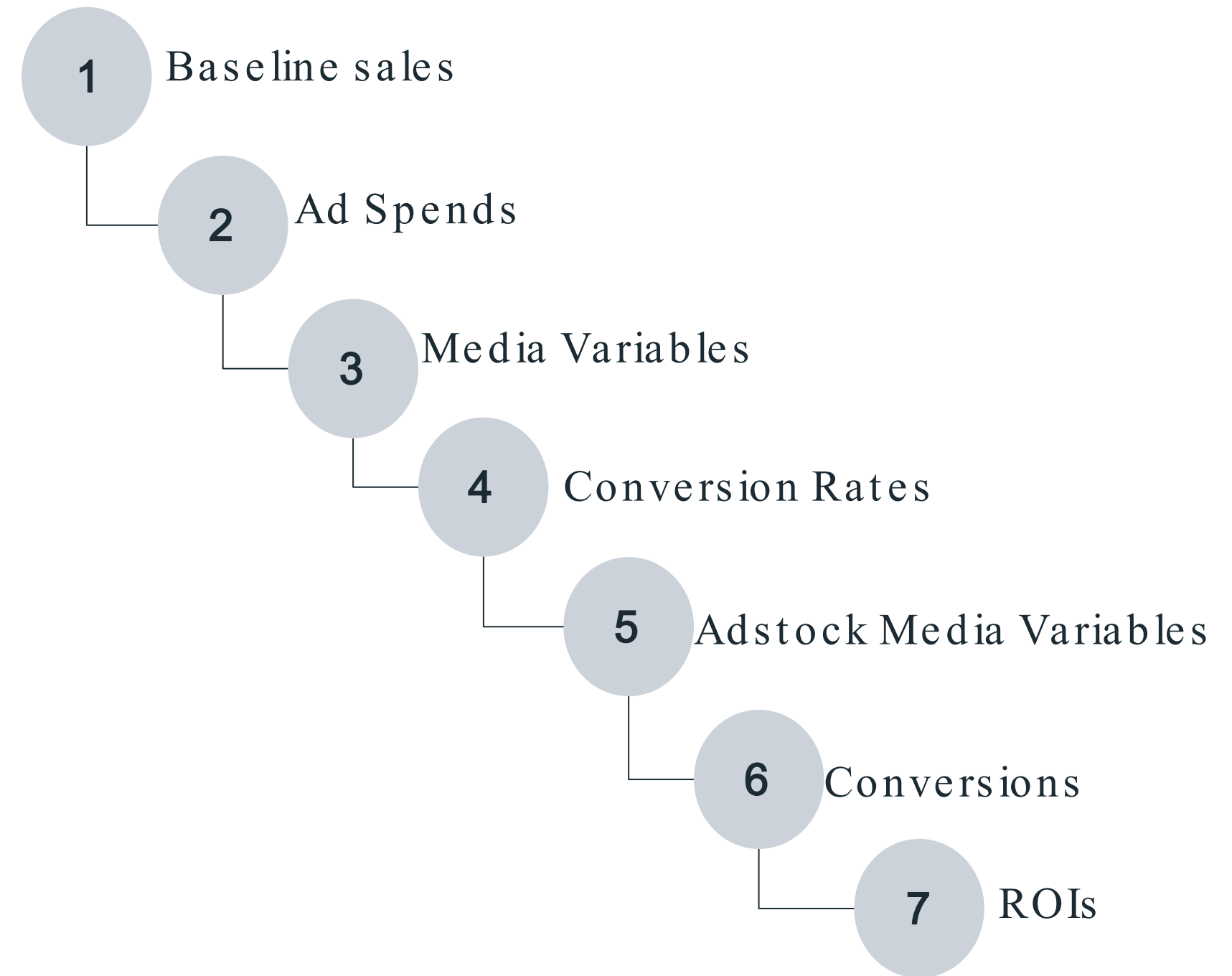


2 e's

Our way to validate:
We can check to see if our
algorithm got the correct
answer

How can simulated data be used to validate Marketing Mix Models?

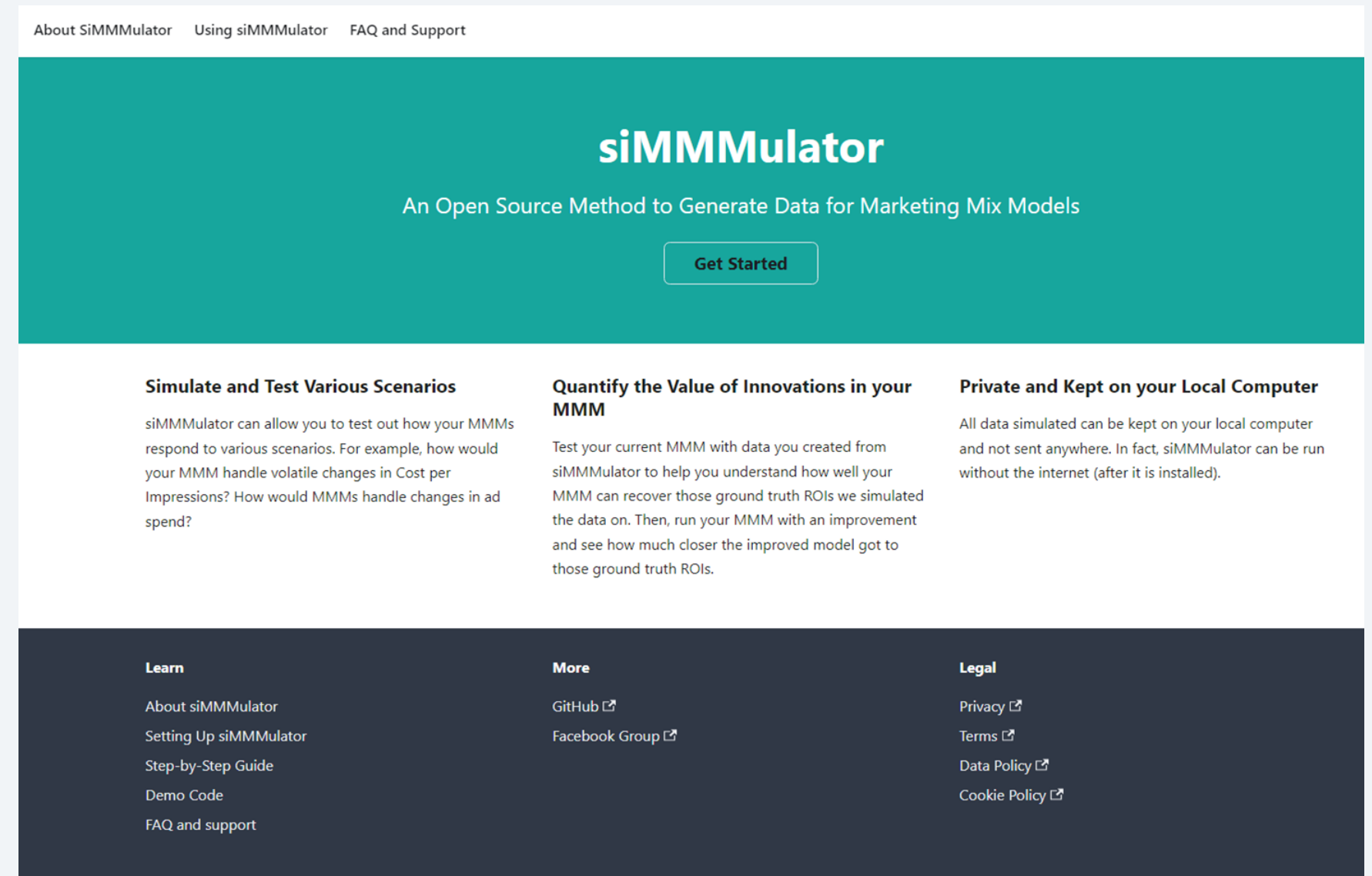
1. Simulate data that ‘recreates’ the advertising process
2. Plug them into MMMs and see how close your MMMs can get to the assumptions you put in
 - a. Simulate various scenarios: see how your MMM responds, what kind of environment your MMMs are sensitive to
 - b. Quantify value of innovations



siMMULATOR:

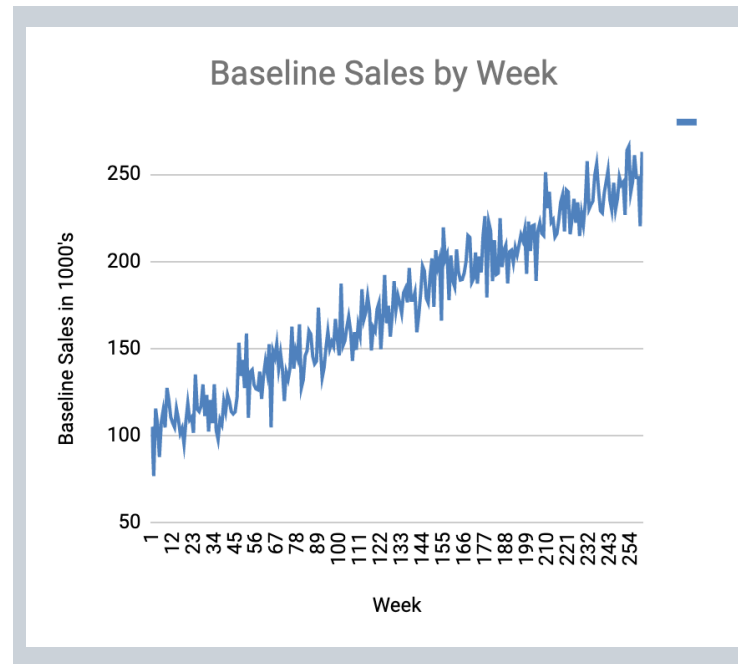
An open-source R package that lets users simulate data from scratch.

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



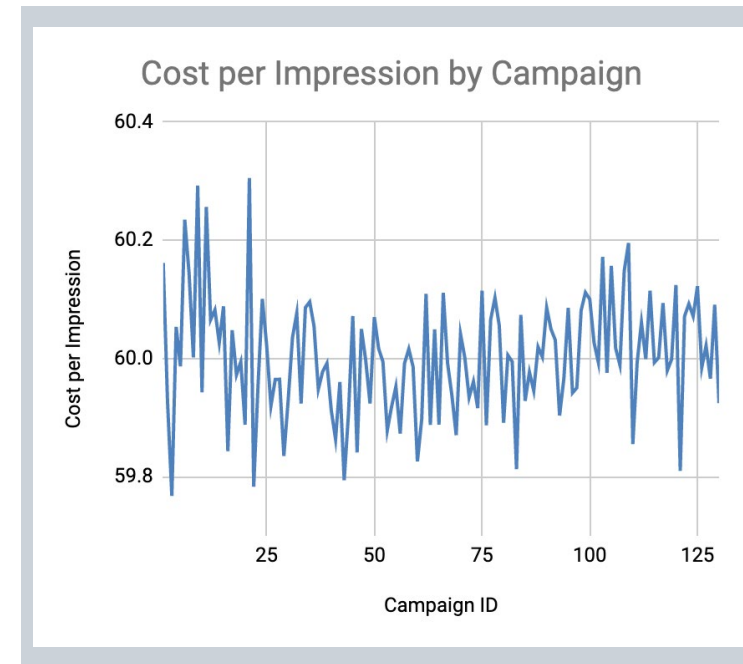
siMMULATOR website homepage

Some inputs that users can change:



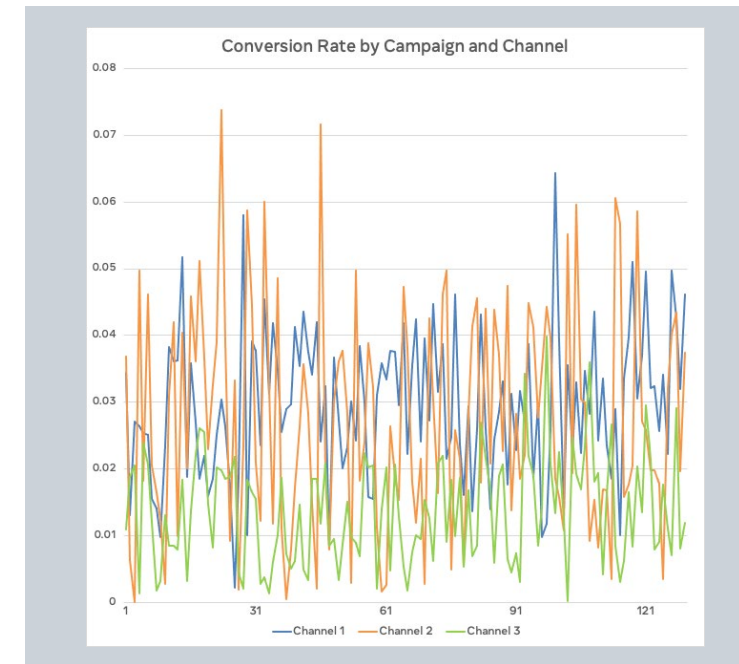
This graph is for illustration purposes only. Individual results may vary.

Baseline sales and its growth over time



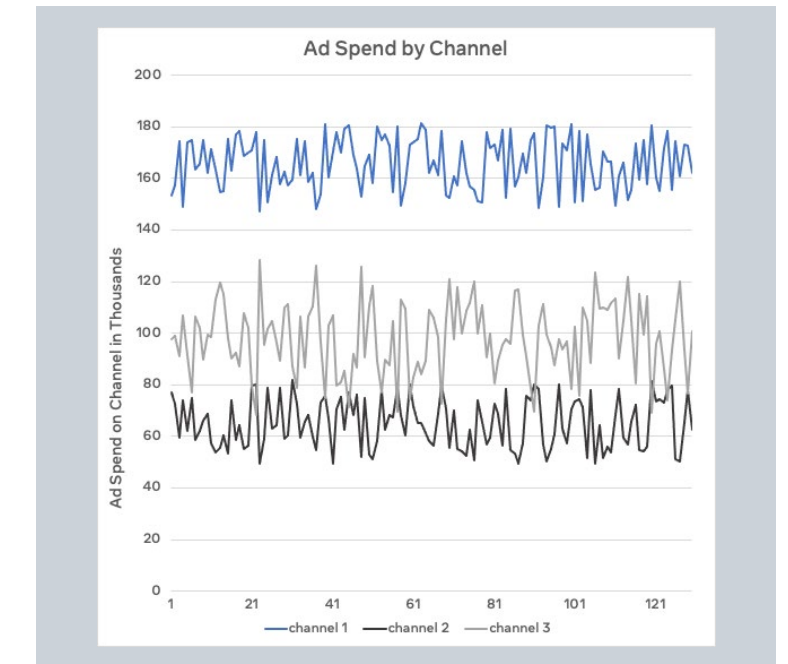
This graph is for illustration purposes only. Individual results may vary.

Average CPM or CPC on a given channel



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

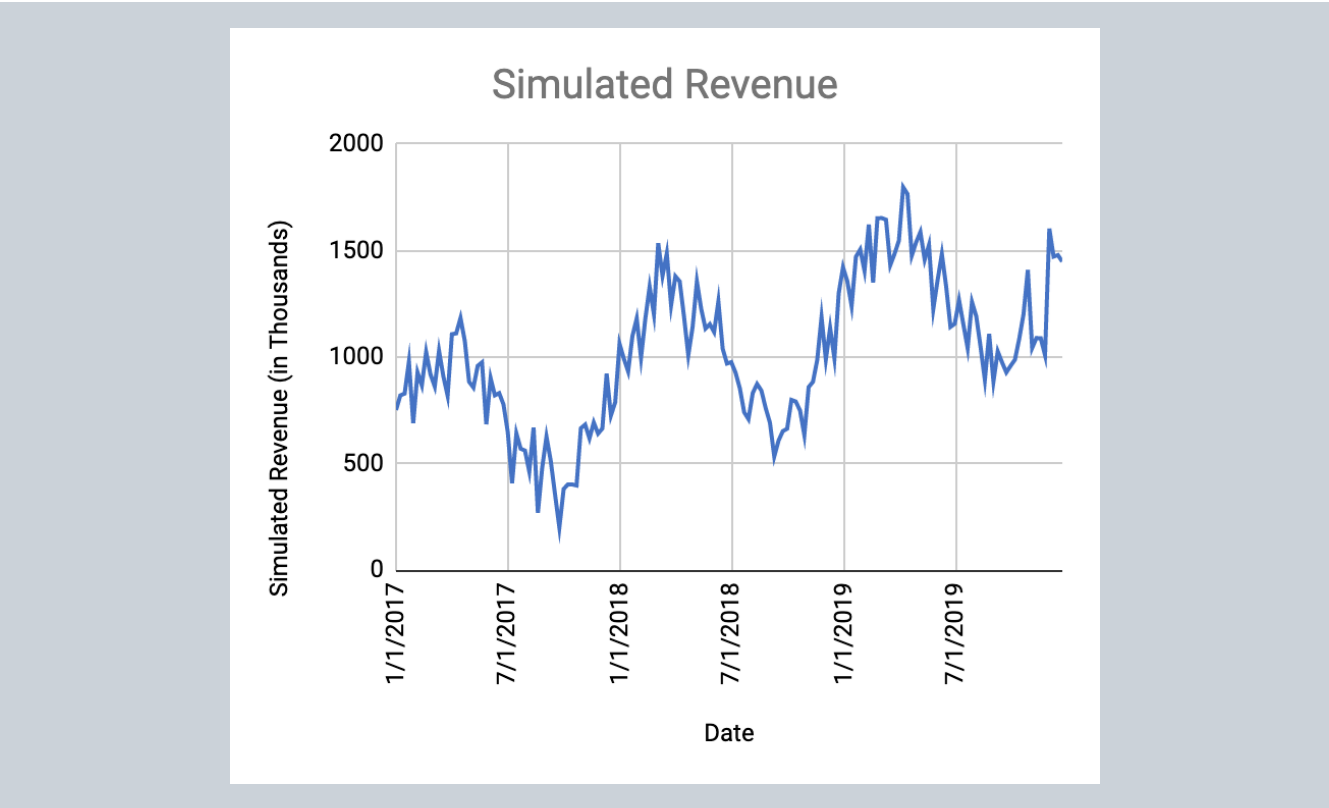
Average conversion rate on a given channel



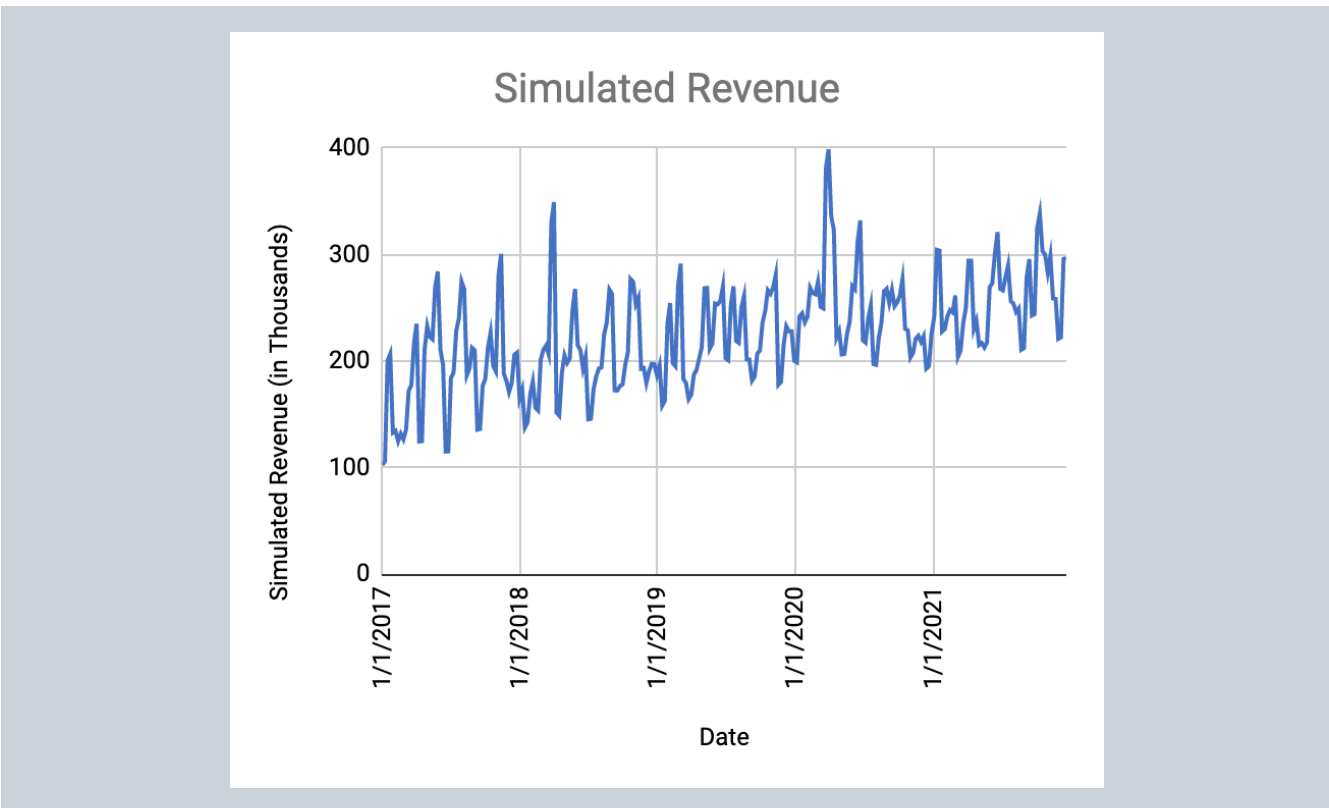
This graph is for illustration purposes only. Individual results may vary.

Spend on a given channel and campaign

Different data sets that can be generated:



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



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

We can get different data sets by putting in different inputs

The data set on the left is more driven by seasonality than the data on the right

Limitations of using simulated data to validate Marketing Mix Models:

Dependence on users'
inputs

How closely simulated data reflects
reality depends on users' inputs and
assumptions

Limited advanced features

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

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.

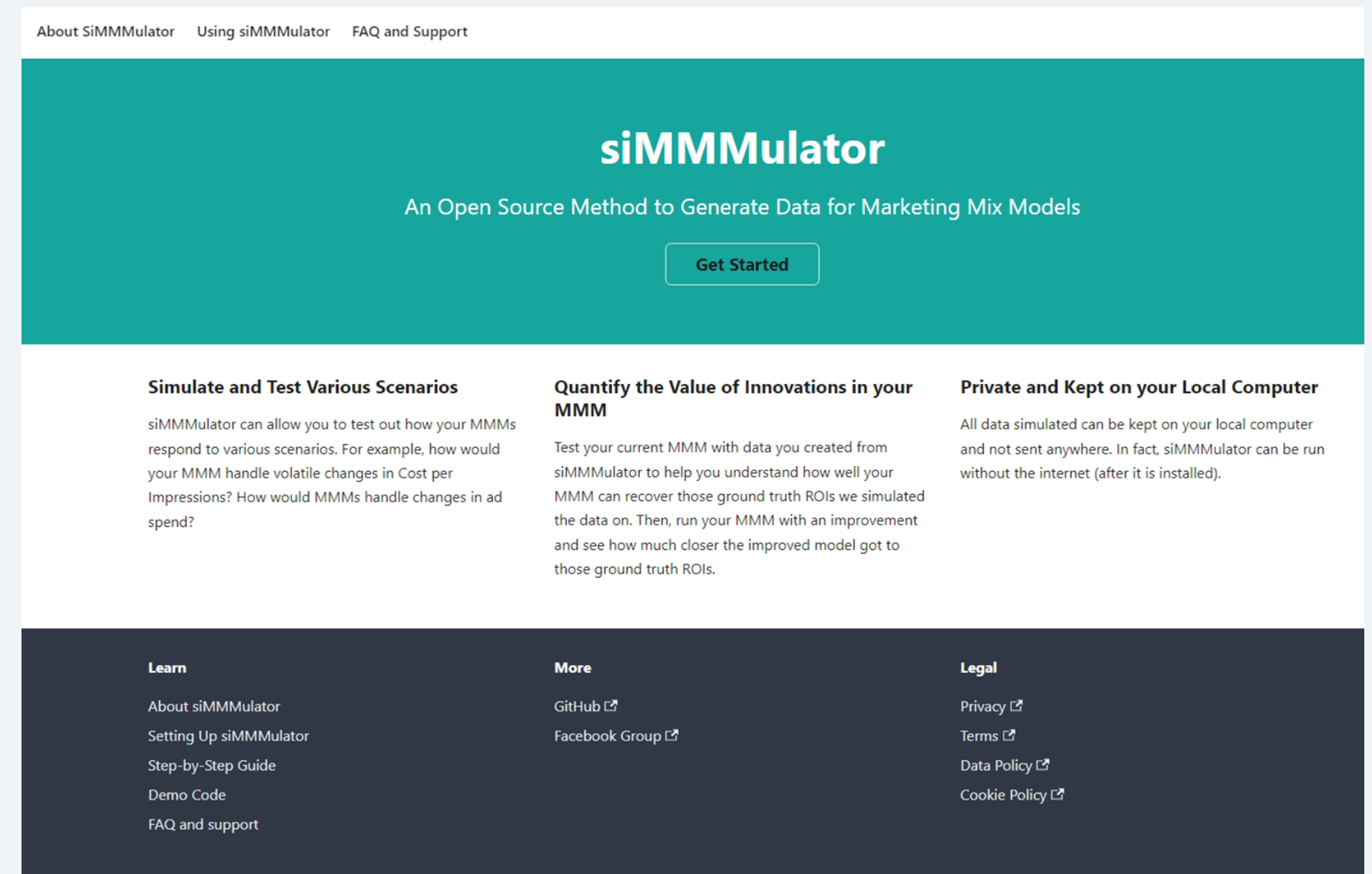
Getting started: Visit our website and download the R -package to get started

Microsite with Quick Start Guide:

<https://tinyurl.com/2u8xkbbw>

siMMMulator Github Repository:

<https://tinyurl.com/2y5mcrnw>

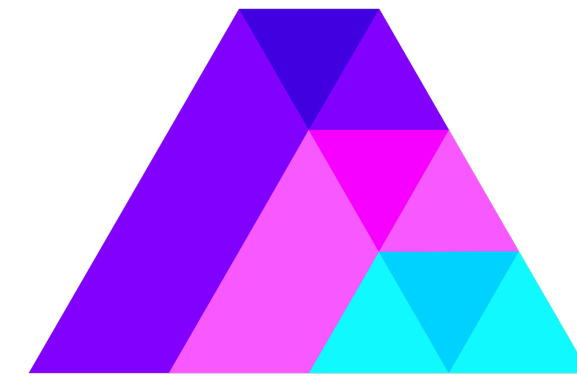


siMMMulator website homepage

Robyn: an open sourced MMM

Industry vision for contemporary Marketing Mix Models

Machine Learning
supported, granular,
automated and
experiment-calibrated
analysis, delivering
insights faster
and on continuous basis



Meta Open Source
Project Robyn

Open source MMM code
for running your own
automated MMM with
experiments



Other case studies

Twigeo (SWEDEN)

Rise Science had 37% higher incremental ROAS driven by digital marketing vs. old model.
Continuous, monthly readouts.

Twigeo
How Twigeo Used Marketing Mix Modeling to Better Capture Marketing ROI for Energy Management App Rise Science

CONTENTS

- 1 The Challenge
- 2 The Solution
- 3 Impact
- 4 Key Takeaways

Twigeo, a full-service digital marketing agency based in Stockholm, Sweden, wanted to provide its client—Rise Science, an energy management app—with better information about the impact of its marketing investments. Specifically, Twigeo wanted to provide Rise Science with its incremental return on investment (ROI) across multiple marketing channels, including Facebook and Instagram, as well as non-digital channels. After careful consideration, the agency opted to partner with the Marketing Science team at Meta in order to offer Rise Science a new measurement approach: marketing mix modeling (MMM) leveraging Robyn, a solution from Meta Open Source.

Coppel (MX)

Five days to implement Robyn model vs. up to three months.

Coppel
Optimizing ad spend with Meta's new Robyn solution

CASE OF SUCCESS

Seeking to know the impact and profitability achieved in each contact channel with its customers, Coppel implemented Robyn, a new Meta Marketing Mix Modeling (MMM) measurement solution. In this way, the Mexican department chain was able to analyze the data of all its campaigns from the last two years, thus getting to know how and where to optimize its advertising investment to achieve better results.

3X more attributed return on ad spend (ROAS) on Robyn implementation vs. previously used model

5 days to implement the Robyn model (compared to traditional modeling which takes up to 3 months)

App-based Gaming Advertisers

Blended results across these 5 businesses saw 47% higher incremental ROAS due to the transparent form of adopting Meta Open Source solutions and adapting them to meet their own business needs, this created trust in calling this their 'source of truth'.

Business Solutions Platforms and Products Inspiration Education and Resources Support

April 13, 2022

How App-Based Businesses Embraced A New, Unified Measurement Standard In A Privacy-First World

Talisa (USA)

Frequent model re-runs, after implementing Robyn, Talisa now has an efficient way to allocate its marketing budget by channel or market on a *quarterly, monthly and weekly basis*.

May 12, 2022

Talisa
JEWELRY

How Talisa Optimized Global Marketing Spend with Marketing Mix Modeling from Meta Open Source

CONTENTS

- 1 The Challenge

New York-based Talisa is a jewelry brand led by husband and wife team Tal Kerret and Lisa Silverstein that specializes in creating personalized fine jewelry.



Install the open-source library from CRAN or Github, and visit Robyn's site to get started

Website with Quick Start Guide:

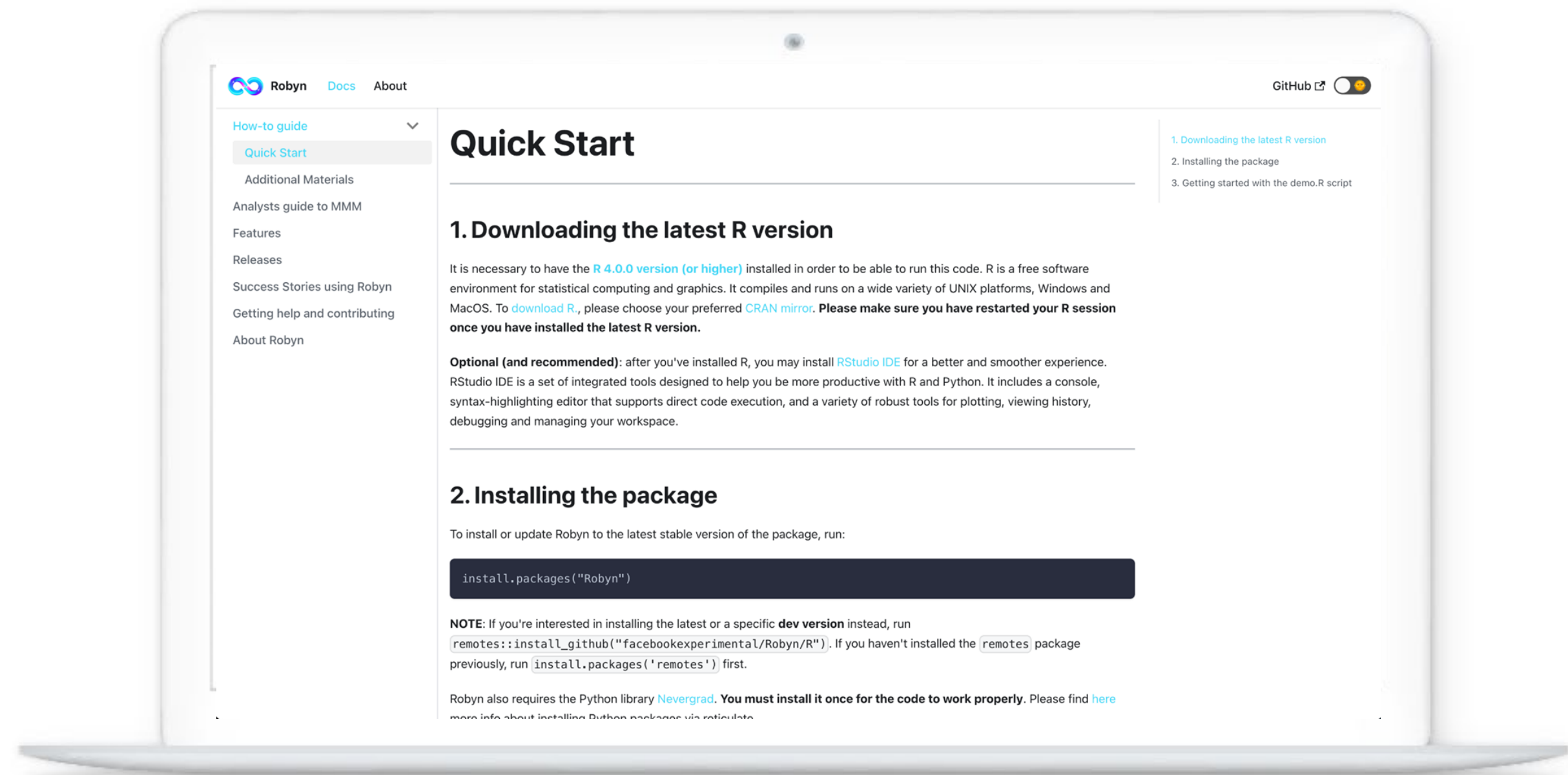
<https://tinyurl.com/ywt8wwus>

Robyn CRAN Repository:

<https://tinyurl.com/yhdvya2u>

Robyn Github Repository:

<https://tinyurl.com/2p9erh4n>



STABLE: `install.packages("Robyn")`

DEV: `remotes::install_github("facebookexperimental/Robyn/R")`