Eight Sleep Take-Home Project

By: Stephen Kim

A. Executive Summary

Recommendations

1a. Allocate capital (**$4,645,618)** in the following way for Sep - Dec '23

Channel	**Suggested Allocation ($)**	**Average ROAS**
Facebook	$1,858,247.20	3.56
TikTok	$371,649.44	3.00
Google	$1,951,159.56	17.74
YouTube	$464,561.80	1.36

- Google is the clear winner but search is different in that users are further down the purchasing lifecycle and have higher propensity to buy. It is to be expected that ROAS from search is highest.
- TikTok's ROAS is higher than YouTube, which is surprising but we need to take into account that total spent for TikTok was less than 1 percent.
- See Figure 1 in Appendix for ROAS distribution curve.
- Cross-referenced projected revenues with another forecasting tool, Meta's Prophet.

The above suggested allocation would yield a higher revenue of **$6,016,968.59** in the months of Sep- Dec '23

- We lose money in Facebook but more than make up those losses through Google, YouTube, and TikTok.
- Google is the clear winner in ROAS but search advertising should always be complemented with other channels. In order to avoid channel saturation, there should be distributed capital allocation across channels (see Appendix Figure 2 for saturation curve)

Commented [1]: This is a good callout, we hypothesize that Google Ads is less incremental than other digital channels

Commented [2]: Right, I suppose further nuanced analysis would be: If someone searches for 'Eight Sleep' chances are, they are going to buy it anyway. But if say they search for Mattresses or Fitness or Recovery and they discover Eight Sleep and convert, that would be very valuable. It would be interesting to see what we are doing for Keyword Optimization under the hood.

Commented [3]: We lose money in Facebook but more than make up those losses through Google, YouTube, and TikTok. - Then why allocate spend to FB?

Current vs. Suggested

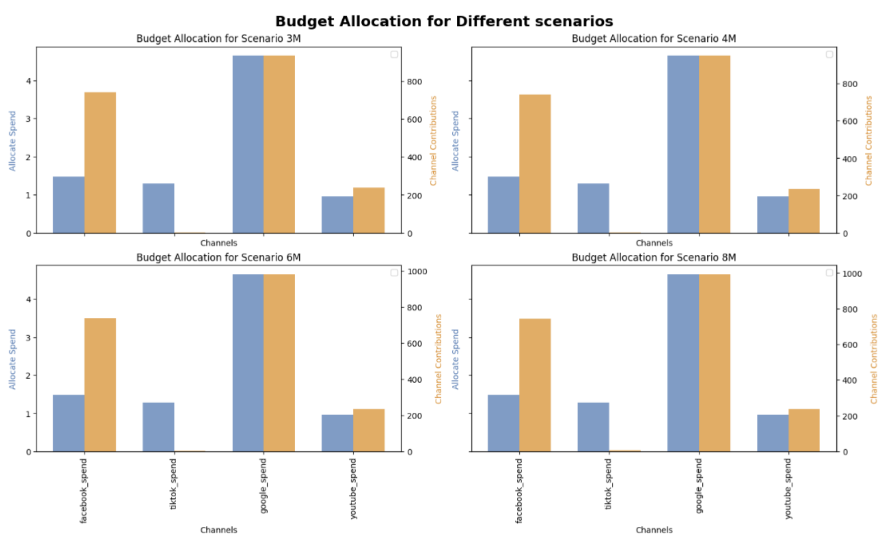
Channel	Revenue (Current)	Revenue (Suggested)	Delta
Facebook	$7,730,817.59	$6,609,147.91	($1,121,669.68)
Google	$28,832,722.73	$34,621,371.33	$5,788,648.60
Tiktok	$348,362.64	$1,116,615.91	$768,253.27
YouTube	$988,472.04	$1,570,208.43	$581,736.39
Total	**$37,900,374.99**	**$42,547,860.83**	**$6,016,968.59**

1b. Consider tweaking the marketing budget

- After running multiple simulations with current MMM and different marketing budget scenarios, we can see there are marginal gains when increasing the budget.
- The graph indicates that boosting the budget beyond an spend level greater than 3 Million induces extremely marginal changes in the potential outcome. Therefore, one can use the budget detailed in scenario three as a cap for our budget.

Commented [4]: Can you explain the difference between these 4 charts? at face value they look the same

Commented [5]: You are spot on. We simulated Budget Allocation of 3M, 4M, 6M, and 8M and the differences are marginal. Orange is our returns (they call it contributions), and blue is our spend. As you can see, we do not get much returns the more we spend. Companies like Bolt have also had this problem and have opted for a more 'custom' method in which it scales better. See here: https://github.com/pymc-labs/pymc-marketing/pull/945

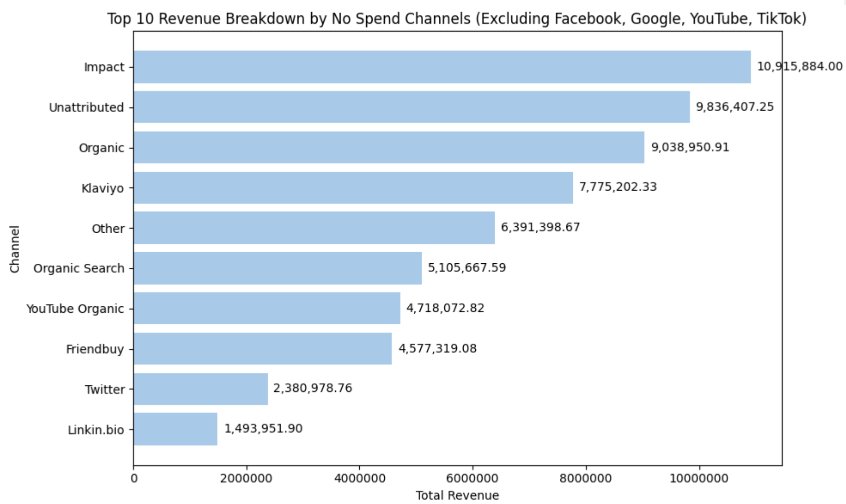


1c. Continue investing in organic platforms and channels

- While many channels are excluded from our MMM, they are still bringing in revenue. We should investigate how customers are engaging on channels such as **Impact, Unattributed, Organic, Other, Klaviyo** and how it affects customers' decisions.
- Specifically for organic channels, customers are clearly engaging and purchasing, proving that these channels are important touchpoints within the users' paths to purchase.
- See Appendix figure 3 for full list

Commented [6]: How would you go about understanding the relationship of the digital channels spend with organic? What model would you use to understand this?

Commented [7]: We can start with Exploratory Data Analysis and see if there is a correlation, chances are there is positive corr. between digital ch. spend with organic engagement. But as we all know, correlation is not causation. For that, we'd have to go into A/B testing if we have the infrastructure setup and enough randomized users (+100,000) in an ideal scenario. Alternatively we can resort to quasi-experimental techniques like user surveys, sentiment analysis, ANOVA, propensity matching etc. I'd also like to delve into the identity graph and Multi-touch attribution model currently in motion to further study the relationship amongst these variables.

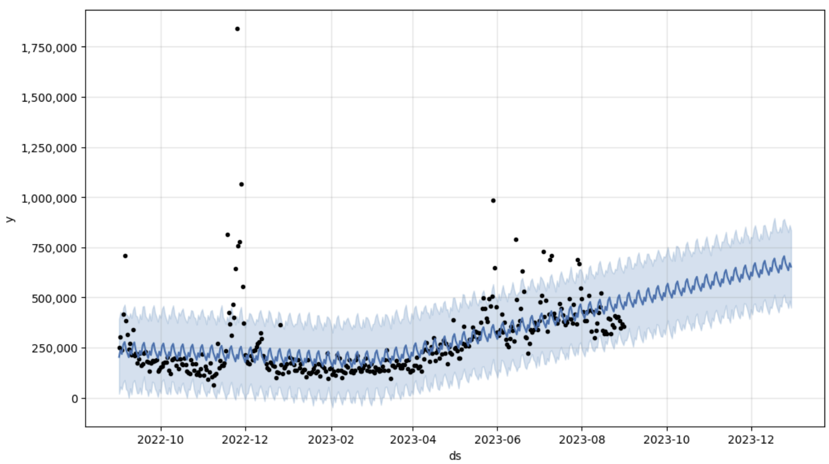


B. Data Science Deep-dive

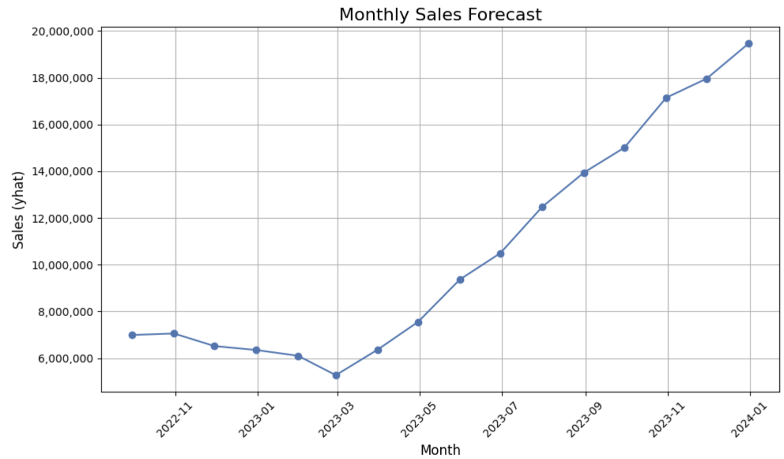
Exploratory Data Analysis

2a. Forecasting

I used Facebook's Prophet to forecast revenue for the months of Sep - Dec '23 with the available data. Below chart shows us on a daily cadence. As you can see it is projected to go up in those months.



Here is another view that is monthly. Roughly, revenue forecasts are: $14mm Sep '23, $15mm Oct '23, $17mm Nov '23, $18mm Dec '23. Each data point is available on google drive as prophet_forecasts.csv.



2b. Aggregate

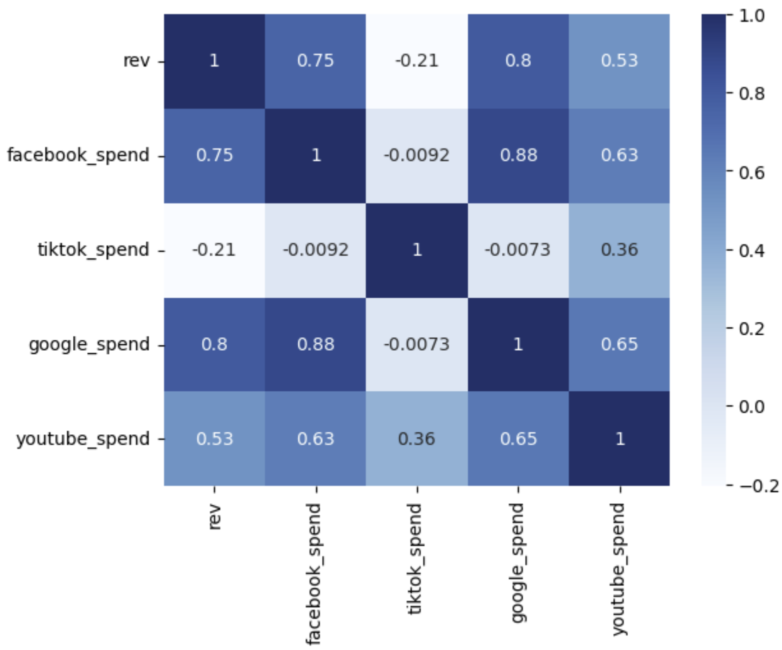
The budget for 2022 is $10,168,617.38 and the budget for 2023 is $13,219,202.59. Information I was given was *“Assume +30% (1.3x) overall marketing budget year-over-year across these months.”* Hence, the budget for Sept- Dec ‘23 is **$4,645,618.28** since the budget for Sept - Dec ‘22 was $3,573,552.52.

Blended numbers show almost a 10:1 revenue to spend ratio. It is a healthy sign that during the period of Aug ‘22 to Sep ‘23, 8S has been growing well. Comparable series C companies have a revenue to spend ratio of 2:1, showing it is ahead of its peers. This is not TRUE Revenue to Spend since COGS and other relevant costs are not factored in.

Total Revenue	**Total Spend**	**Revenue to Spend Ratio**
$98,465,466.05	$10,168,617.38	9.68

Looking at the correlation plot, we can see

- Google spend is positively correlated to **revenue** (0.8) as is Facebook (0.75), and Youtube (0.53)
- TikTok (-0.21) is negatively correlated
- While correlation is not causation, it can still provide directional insights



2c. Per Channel

We should continue to invest in organic channels where there is attributed revenue but no spend such as ‘Organic Search’ and ‘Facebook Organic’. The data shows that customers are engaging with these organic channels (whether searching the 8S instagram profile, or clicking on an organic search result) at some point in their path to purchase, so it is important to ensure that these touchpoints are appealing and encourage customers to go down the purchase funnel.

I started looking at Top Spend Channels.

Channels Rev to Spend

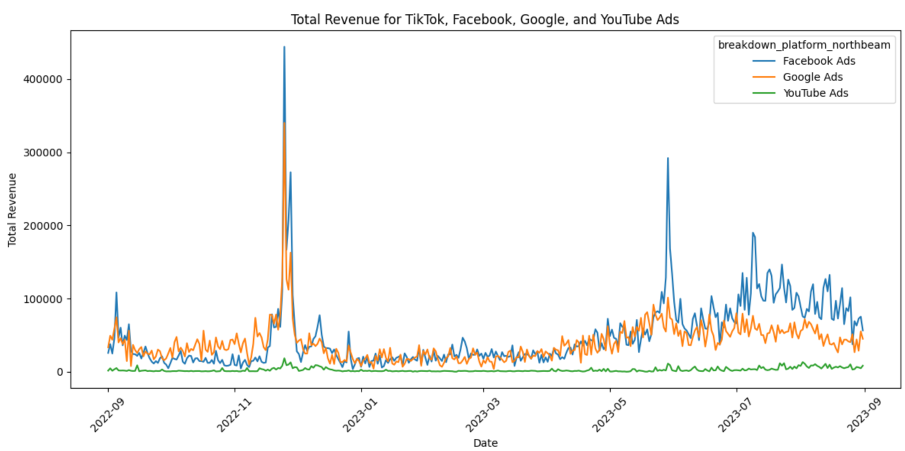
Breakdown P..	Rev	Spend
Facebook Ads	17,770,745	4,942,361
Google Ads	14,318,791	3,817,452
Impact	10,915,884	0
Unattributed	9,836,407	0
Organic	9,038,951	0
Klaviyo	7,775,202	0
Other	6,391,399	0
Organic Search	5,105,668	0
YouTube Organic	4,718,073	0
Friendbuy	4,577,319	0
Twitter	2,380,979	0
Linkin.bio	1,493,952	0
YouTube Ads	959,754	1,319,614
Excluded	897,794	0
Facebook Organic	578,946	0
TikTok	307,169	89,190
Instagram Organic	284,804	0
Influencer	198,541	0
Other Email	184,989	0
Reddit	135,661	0
LinkTree	131,931	0
LinkedIn Ads	130,254	0
Pinterest	77,398	0
Instagram Shop	64,798	0
Snapchat Ads	63,960	0
ActiveCampaign	52,447	0
Transactional	32,002	0
Rakuten	16,845	0
Discount Site	8,423	0
Yahoo Gemini	4,638	0
Taboola	3,714	0
Criteo	3,033	0
Outbrain	2,526	0
Microsoft Ads	2,468	0

I then determined the top 4 are relevant, are shown below. This is the reasoning for including them into our MMM. As you can see Facebook is leading our spend at 48%, Google at 37, Youtube at 12% and Titkok almost 1%.

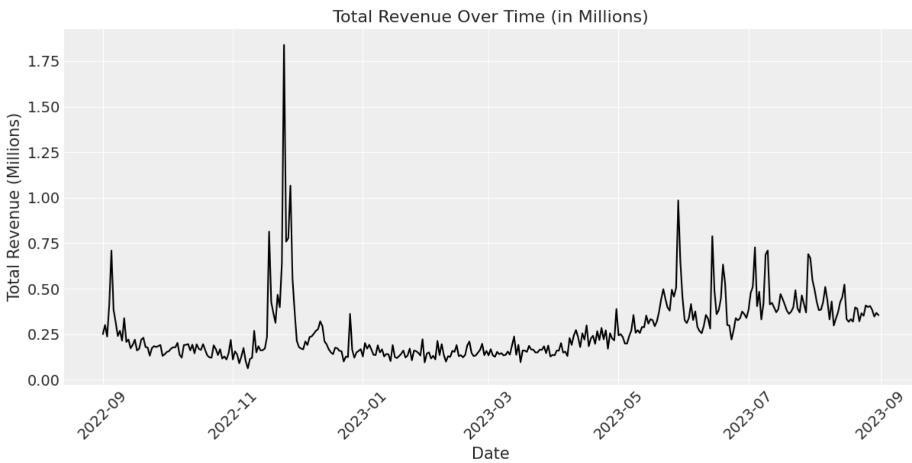
```
Top 4 Total Spend per Channel (Descending Order):
 breakdown_platform_northbeam
Facebook Ads    4.942361e+06
Google Ads      3.817452e+06
YouTube Ads     1.319614e+06
TikTok          8.919037e+04
Name: spend, dtype: float64

Top 4 Spend Share per Channel (Descending Order):
 breakdown_platform_northbeam
Facebook Ads    0.486041
Google Ads      0.375415
YouTube Ads     0.129773
TikTok          0.008771
Name: spend, dtype: float64
```

I also looked at revenue over time for these 4 channels. Tiktok is not even at the same scale as the other 3, but does not mean it's not relevant.



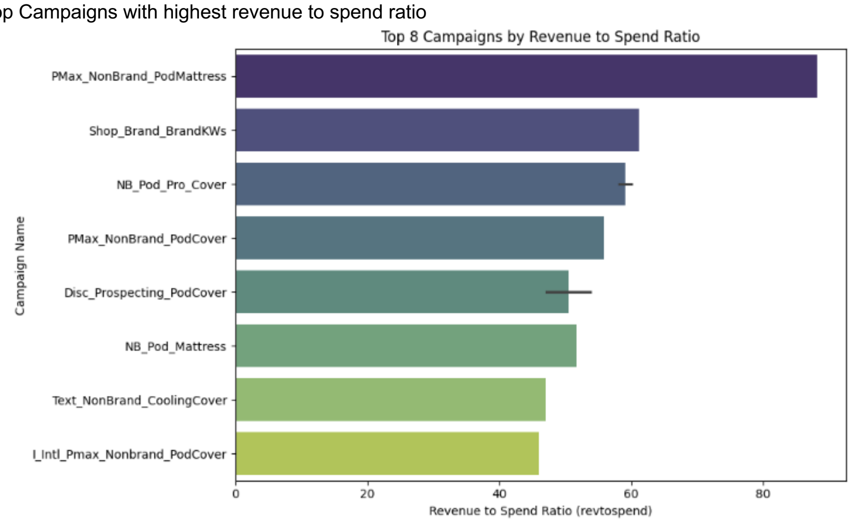
Below is the same chart but aggregated view.



2d. Per Campaign

We should take learnings from both the top performing and worst performing campaigns (targeting strategy, audience strategy, pricing, etc.) to increase overall performance.

Top Campaigns with highest revenue to spend ratio

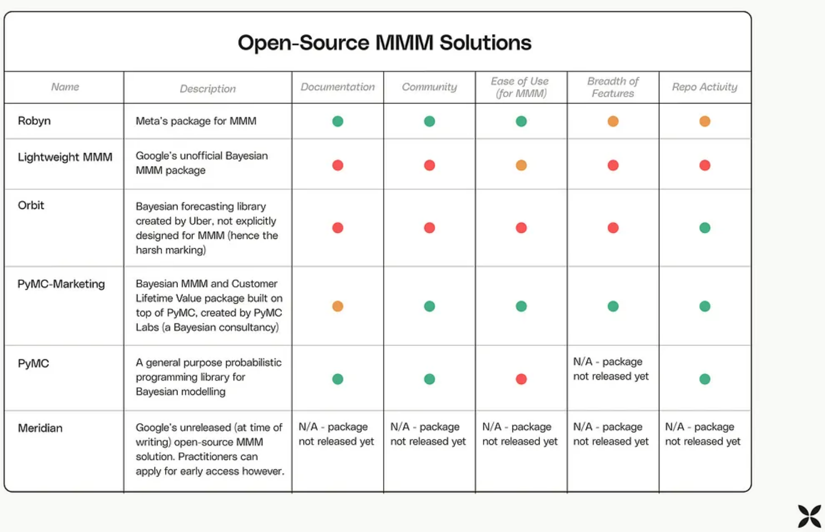


Worst Campaigns with no revenue despite ad spend

campaign_name	spend	rev	revtospend
I_CA_Text_NonBrand_CoolingCover	771.26	0.00	0.00
Text_Blog_SweatySleep	703.11	0.00	0.00
UK_Pmax_Nonbrand_PodCover	866.68	0.00	0.00
AUS_Text_NonBrand_CoolingCover	542.46	0.00	0.00
I_CA_Text_NonBrand_CoolingCover	937.81	0.00	0.00
I_UK_Text_NonBrand_CoolingCover	694.79	0.00	0.00
AUS_Text_NonBrand_CoolingCover	758.25	0.00	0.00
I_CA_Text_NonBrand_CoolingCover	580.41	0.00	0.00
NB_Pod_Pro_Cover	642.05	0.00	0.00
NB_Pod_Mattress	451.44	0.00	0.00

Marketing Mix Model

Marketing mix models help allocate capital better in a privacy-first era in which third party cookies are being deprecated. There are multiple open source solutions and the below table highlights some pros and cons. I've selected **PyMC** and **PyMC-marketing** for its simplicity and comprehensive robustness. Google's new Bayesian model called Meridian (invite only) is also something that can be tested.



Open-Source MMM Solutions

Name	Description	Documentation	Community	Ease of Use (for MMM)	Breadth of Features	Repo Activity
Robyn	Meta's package for MMM	🟢	🟢	🟢	🟠	🟠
Lightweight MMM	Google's unofficial Bayesian MMM package	🔴	🔴	🟠	🔴	🔴
Orbit	Bayesian forecasting library created by Uber, not explicitly designed for MMM (hence the harsh marking)	🔴	🔴	🔴	🔴	🟢
PyMC-Marketing	Bayesian MMM and Customer Lifetime Value package built on top of PyMC, created by PyMC Labs (a Bayesian consultancy)	🟠	🟢	🟢	🟢	🟢
PyMC	A general purpose probabilistic programming library for Bayesian modelling	🟢	🟢	🔴	N/A - package not released yet	🟢
Meridian	Google's unreleased (at time of writing) open-source MMM solution. Practitioners can apply for early access however.	N/A - package not released yet	N/A - package not released yet	N/A - package not released yet	N/A - package not released yet	N/A - package not released yet

Commented [8]: If you were to productionalize this model, what do you think the output looks like? How often would you re-run?

What we are trying to get to is: How would we implement the findings from an MMM into our budget planning

Commented [9]: Output could either be the suggested breakdown of the current budget in $ values or percentage, or some composite metric we devise, such as (how much until channel saturation in percentage), We'd have to implement ML Ops for this to work. I suggest we go with MLFlow, which will give us a snapshot of all the inputs and results for reproducibility and data version control. Qonto has a great series of pieces here: https://medium.com/qonto-way/marketing-measurement-series-marketing-mix-modeling-at-qonto-part-vi-7bb9805076ba

3a. Model Assumptions

For our MMM, we will be using a Bayesian regression model. While a traditional regression model is a good baseline and can show which variables affect sales, there are some limitations to it because it assumes a linear relationship. However, because the correlation between sales and marketing is not linear, the Bayesian model is a more comprehensive model that can account for additional variables such as adstock and saturation effects and are comprised of a prior.

The below traditional regression formula is decent, but we can do better:

$$\boxed{\text{Sales} = \beta o + \beta 1 * x1 + \beta 2 * x2}$$

For validation of the Bayesian model, we can do a train/test split and measure error rates such as MAE. The main parameters to consider are

- Alpha: adstock effect: we use the **Geometric Adstock (more here)** with a max lag of 8
- Lambda: we use a **Logistic Saturation (more here)**
- Sigma: we use our prior sigma defined as follows (we must scale it since all the spend values are different), we can create priors from business knowledge. We do so by calculating the total spend per channel, to ultimately get a spend_proportion. The prior sigma is defined as **priori_sigma = halfnormal_scale *n_channels *spend_proportions**

```
[199]: total_spend_per_channel = gdf[['facebook_spend', 'tiktok_spend', 'google_spend', 'youtube_spend']].sum(axis=0)
       total_spend_per_channel

[199]: facebook_spend    4.942361e+06
       tiktok_spend      8.919037e+04
       google_spend      3.817452e+06
       youtube_spend     1.319614e+06
       dtype: float64

[203]: spend_proportion = total_spend_per_channel / total_spend_per_channel.sum()
       spend_proportion

[203]: facebook_spend    0.486041
       tiktok_spend      0.008771
       google_spend      0.375415
       youtube_spend     0.129773
       dtype: float64

[205]: HALFNORMAL_SCALE = 1/ np.sqrt(1- 2 / np.pi)

[207]: n_channels = 4

[209]: prior_sigma= HALFNORMAL_SCALE * n_channels * spend_proportion

[211]: prior_sigma.tolist()

[211]: [3.2251649838917307,
        0.058201664407954115,
        2.491099129892152,
        0.8611211816893876]
```

Furthermore, we will be adding trend and seasonality into the model given that it is a time series.

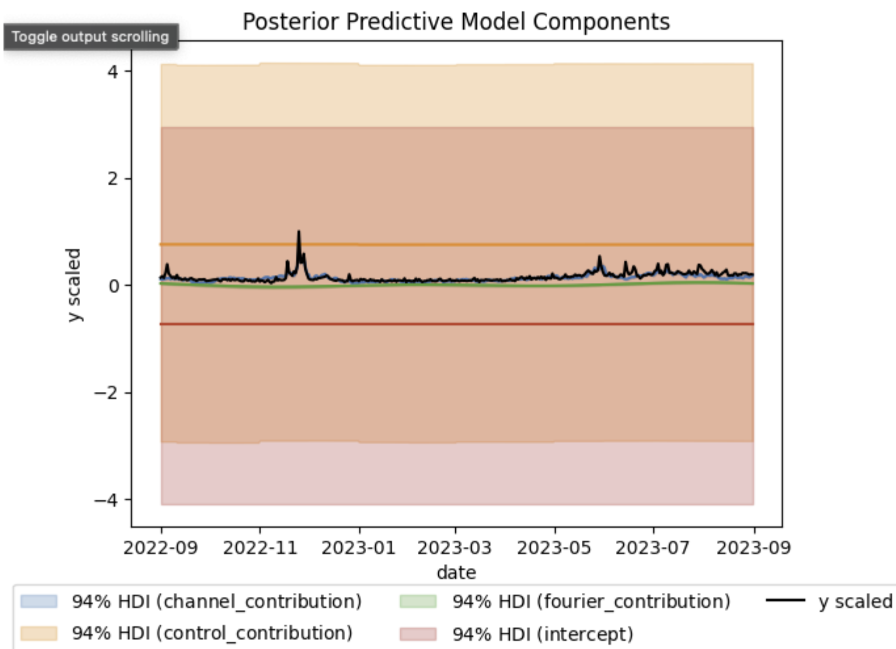
Other model considerations: I've played around with other considerations like events. For instance, adding events like launch of pod3, pod3 summer refresh, intro of subscription service, were considered but not deemed very relevant.

Commented [10]: Nice callouts - how did you find out about Pod 3 launch/subscription?

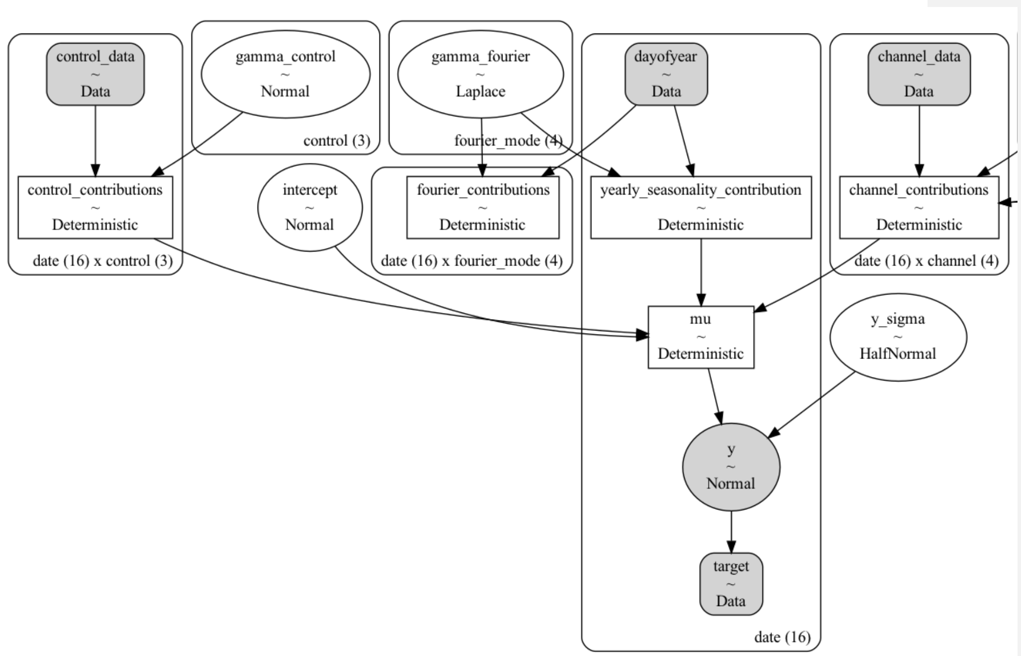
Commented [11]: I just did research on the product rollout and launches. See Figures 4,5,6.

3b. Model Interpretation

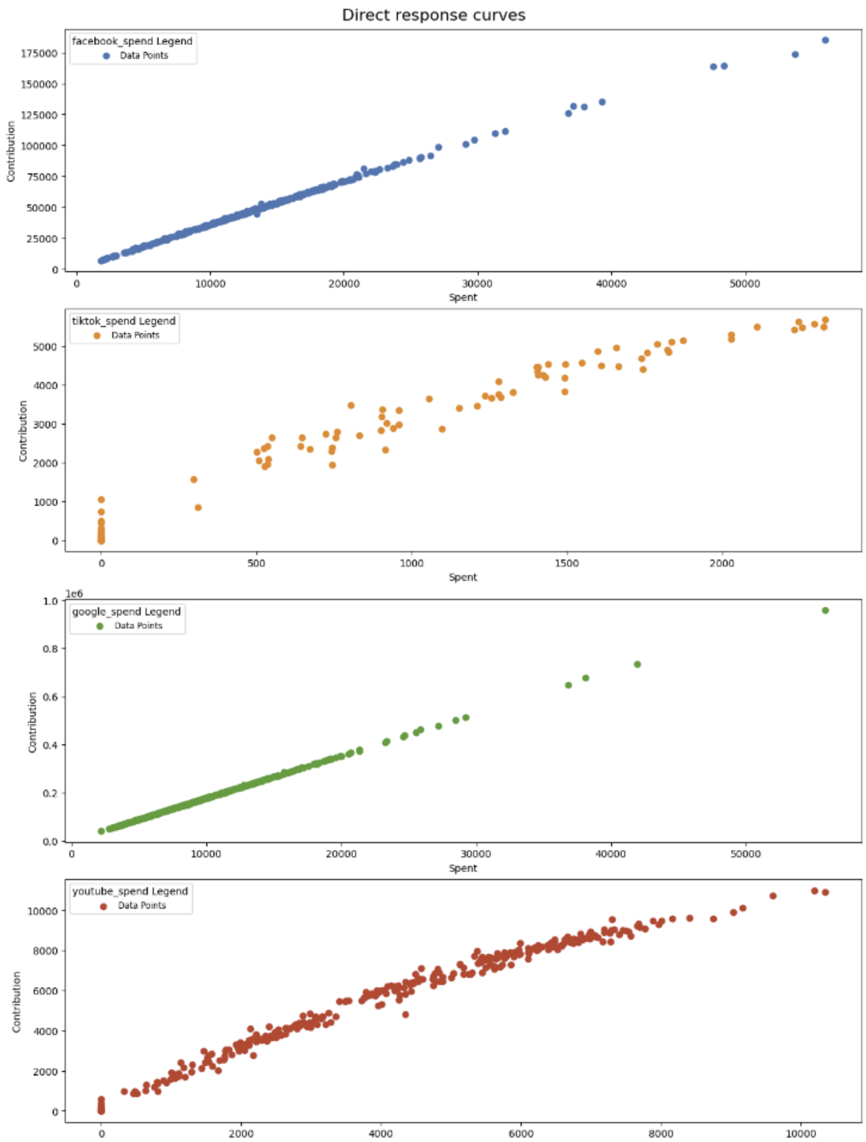
The chart below shows contribution levels for each of the components.



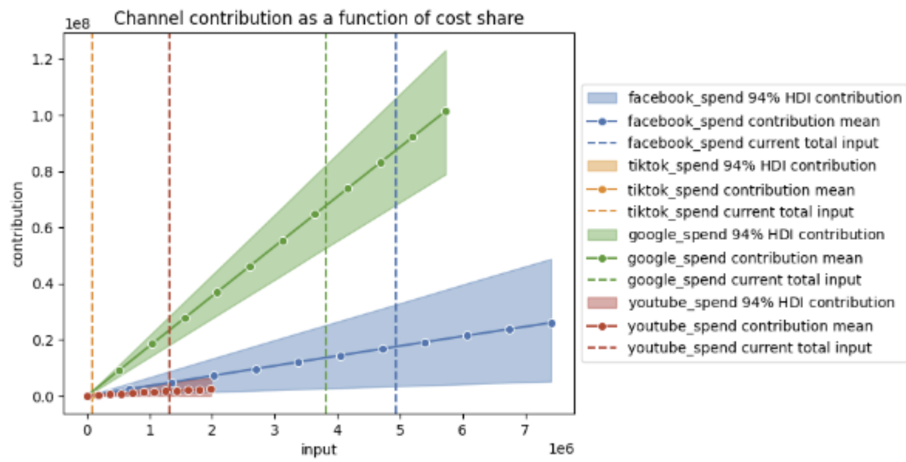
This chart shows us the inner workings of the Bayesian model. As you can see, there are many parameters that go into the model that can be further tweaked. How do all the different variables relate to each other to affect sales? Link



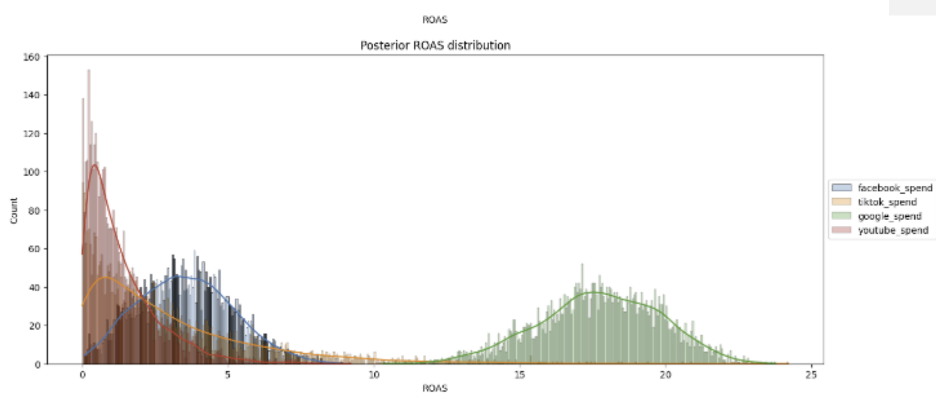
Now we look at the contribution curves below.As you can see, the **saturation effects** are visible as there is a plateauing effect. Facebook and Google however seem more linear so it seems to suggest that contribution will continue to increase as the spend increases.


Direct response curves
facebook_spend Legend
Data Points
Contribution
Spent
tiktok_spend Legend
Data Points
Contribution
Spent
1e6
google_spend Legend
Data Points
Contribution
Spent
youtube_spend Legend
Data Points
Contribution
Spent

Furthermore, Channel contributions as a function of cost share gives us another angle of the same information.



Here, we look at the ROAS (Return on Ad Spend) Distribution. We can assess the effectiveness of each of these different channels, factoring all the time series components from the model. We can interpret that Google is the best one, followed by facebook, tiktok, and youtube. Youtube maxes out at a ~2 ROAS, which is good but not GREAT.



Below is the same ROAS data in a dataframe. As you can see, Youtube peaked at a 1.35 ROAS, Tiktok at a 3.0, facebook at a 3.56, and Google at a 17.74.

[292]:

channel	chain	draw	roas
facebook_spend	1.5	499.5	3.556657
google_spend	1.5	499.5	17.743998
tiktok_spend	1.5	499.5	3.004487
youtube_spend	1.5	499.5	1.351991

Below is looking at the same data but through a different lens which can help us answer questions probabilistically. For example, what is the probability that this channel is above the break-even point (1.0)? For example, Tiktok's average ROAS is 3.00 but at the 95th percentile is going to be 0.06 to 11.11. This means that there is a chance that TikTok will perform below the break-even point. We can also use this logic to assess budget allocation depending on risk appetite.

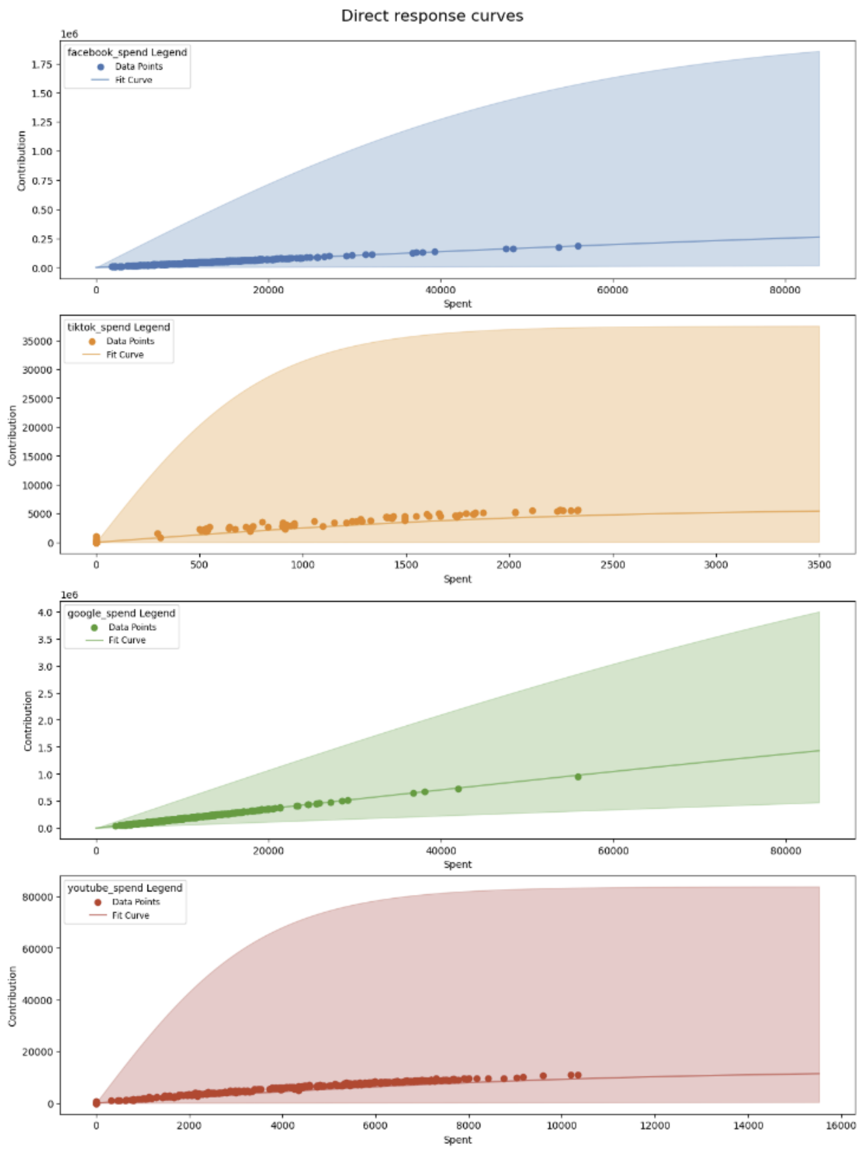
Commented [12]: Knowing this, what experiment would you set up to re-test TikTok?

Commented [13]: I'd start the EDA with understand our user distribution for TikTok, age, demographics, average duration of engagement, what kind of ads (design artifacts they are seeing) and then conduct AB Testing. I'd also leverage marketing team's domain expertise and current MTA model, identity graph, and others. If we cannot AB test, we can deploy causal inference techniques as mentioned above.

[356]:

channel	count	mean	std	min	2.5%	50%	97.5%	max
facebook_spend	4000.0	3.556657	1.622441	0.072765	0.697627	3.498522	6.901654	8.915051
google_spend	4000.0	17.743998	2.082701	10.315904	13.484459	17.783358	21.541903	23.763280
tiktok_spend	4000.0	3.004487	3.030163	0.006498	0.060873	2.103601	11.105197	24.214393
youtube_spend	4000.0	1.351991	1.258511	0.001669	0.038589	0.974795	4.704678	9.208689

Below the response curves, we can fit using sigmoid (or michaelis-menten)



3c. Budget Allocation with Model

Budget allocation is done with PyMC Marketing's open source library. Companies like Bolt have also adopted this technique. It leverages scipy's Scipy's SLSQP optimization. Exact breakdown of the budget was : 40% Meta, 42% Google, 8% TikTok, 10% YouTube.

- Budgeting module is not an exact science so I optimized for diversity without compromising revenue or growth.
- We've developed budget allocation strategies tailored to meet each specific objective we aim to achieve, rather than solely focusing on maximizing the total contribution of the target variable.
- It is important to consider the current marketing strategy. Are we trying to target a younger demographic via Tiktok? Do they have the disposable income to spend $3-4K in the Pod? What kind of audience overlap is there across different channels and are we taking this into account when optimizing performance?

C. Next Steps

- Incorporate enhancements to the model to create a more 'bespoke' model
  - Use most recent data and more of it to feed model and re-run since current sample size is too small
  - Factor in important events like price increases, launch of pod 3,4 3-summer-refresh, subscription, discounts (Figure 4,5,6)
  - Incorporate marketing team's subject matter expertise into prior beliefs of Bayesian model and re-calibrate
  - Most MMM models are done weekly, so re-run with weekly model and experiment with order of adstock and saturation
- Compare results with different packages
  - Google has a new library called Meridian, which is invite only. You have to formally apply to test it.
- Build self-service application for cross functional stakeholders with Streamlit
  - This will allow non-DS employees to use a UI to conduct simulations (nice to have)
- Leverage other models like MTA, Churn, Incrementality, Uplift models to complement MMM.
- Set up model ops with ml flow to keep track of ongoing model versioning and results

Appendix

Sources:

https://bolt.eu/en/blog/budgeting-with-bayesian-models-pymc-marketing/

https://developers.google.com/meridian/docs/basics/about-the-project
https://github.com/google/lightweight_mmm
https://research.google/pubs/bayesian-methods-for-media-mix-modeling-with-carryover-and-shape-effects/
https://storage.googleapis.com/gweb-research2023-media/pubtools/3806.pdf
https://github.com/pymc-labs/pymc-marketing?tab=readme-ov-file
https://www.pymc-marketing.io/en/stable/notebooks/mmm/mmm_example.html#roas
https://www.pymc-marketing.io/en/stable/notebooks/mmm/mmm_budget_allocation_example.html
https://zaai.ai/mmm-bayesian-framework-for-marketing-mix-modeling-and-roas/

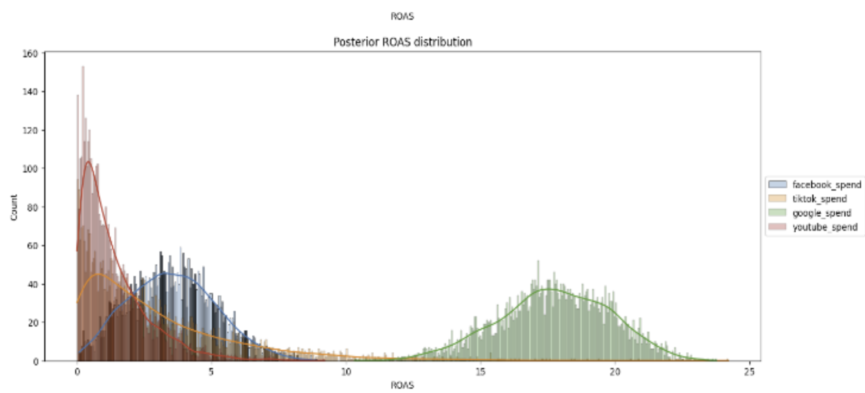


Figure 1

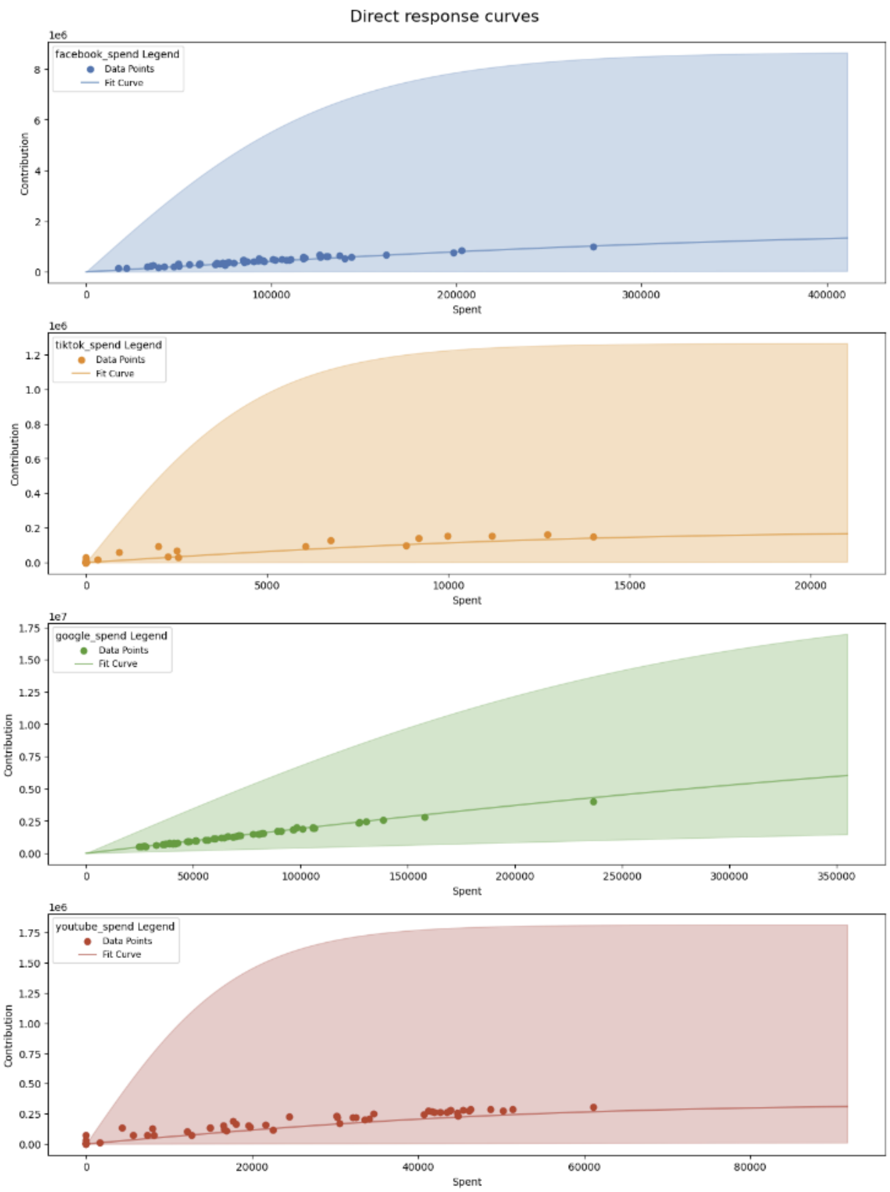

Direct response curves
facebook_spend Legend
Data Points
Fit Curve
tiktok_spend Legend
google_spend Legend
youtube_spend Legend
Contribution
Spent

Figure 2

	Platform	Total Revenue
4	Facebook Ads	17,770,745.16
7	Google Ads	14,318,791.14
8	Impact	10,915,884.00
30	Unattributed	9,836,407.25
17	Organic	9,038,950.91
12	Klaviyo	7,775,202.33
19	Other	6,391,398.67
18	Organic Search	5,105,667.59
33	YouTube Organic	4,718,072.82
6	Friendbuy	4,577,319.08
29	Twitter	2,380,978.76
15	Linkin.bio	1,493,951.90
32	YouTube Ads	959,753.92
3	Excluded	897,794.39
5	Facebook Organic	578,946.25
27	TikTok	307,168.82
10	Instagram Organic	284,804.04
9	Influencer	198,541.19
20	Other Email	184,988.85
24	Reddit	135,660.91
13	LinkTree	131,930.98
14	LinkedIn Ads	130,253.77
22	Pinterest	77,398.47
11	Instagram Shop	64,798.03
25	Snapchat Ads	63,959.79
0	ActiveCampaign	52,446.51
28	Transactional	32,002.41
23	Rakuten	16,845.41
2	Discount Site	8,422.65
31	Yahoo Gemini	4,638.38
26	Taboola	3,714.25
1	Criteo	3,033.36
21	Outbrain	2,525.95
16	Microsoft Ads	2,468.12

Figure 33

Author: **Editorial** August 03, 2023

At Eight Sleep, the pursuit of product excellence never stops. We've made some impactful improvements to Pod 3 since the original launch in July 2022, with continuous software and hardware improvements that have rolled out behind the scenes.

If you've purchased a Pod in the past months, you'll be delighted to learn about some of our recent updates, including:

Enhanced durability

We've updated some materials on the Pod Cover, as well as some of the Hub components to protect units from the daily wear & tear. These new materials have shown to be up to 2X more resistant to punctures and have been rigorously tested to simulate 10 years of use. These improvements will increase the lifetime of the product, while maintaining our desired aesthetic and comfortable feel.

Figure 4

Products [edit]

	Alexandra Zatarain
Headquarters	New York City, United States
Products	Smart mattress
Website	eightsleep.com

Eight Sleep's flagship product is the Pod mattress.[7] The company has also integrated its offerings with Amazon's Echo devices.[8][9] In 2024, Eight Sleep launched the four generation of the Pod, Pod 4.[10]

In 2021, *The Strategist* described the company's Pod mattress as the only product of its type able to provide perceptible temperature control.[11] A more recent review from that website, however, noted that "the Eight Sleep's temperature-controlling features are easily found elsewhere for cheaper."[12]

In February 2023, Eight Sleep started to require a paid subscription (with an annual cost ranging from $180 to $288) to access most of the Pod's functionality, including sleep tracking and variable temperature control.

Figure 5

Author: **Matteo Franceschetti** February 02, 2023

We're honored to introduce Pod 3, a new version of our award-winning sleep technology. From the start, we set out to create the best systems in the world to power sleep optimization, and with the launch of Pod 3 we are taking a giant leap forward.

Three years after the launch of the original Pod, it is rewarding to see how our innovations are improving people's lives, with clinically validated results that prove better recovery night after night.

Clinical data have shown that the technology behind Pod 3 can improve sleep quality index by up to 32%, increase deep sleep by up to 34%, and increase heart rate variability (HRV) by up to 19%. While still maintaining the award-winning features of the Pod 2 Pro, the new Pod 3 includes enhanced technology and an improved user experience with the following enhancements:

Figure 6