

Universidade de São Paulo
Instituto de Ciências Matemáticas e de Computação

Marketing Mix Modeling: uma aplicação digital

Mariane Romildo dos Santos

São Carlos - SP
Julho/2020

Resumo

Escolher uma combinação de mídias ideal era uma tarefa complicada para o *marketing* tradicional e é ainda mais difícil para o digital devido à grande quantidade de opções disponíveis. *Marketing mix modeling* é uma ferramenta criada para que os estrategistas de *marketing* tradicional possam entender a contribuição de diversos esforços no sucesso de sua empresa. Neste trabalho a aplicação da ferramenta MMM será feita para um cenário exclusivo de mídias digitais a fim de encontrar a melhor combinação de mídias e a melhor distribuição dos investimentos futuros.

Palavras-chave: *Marketing* digital; *Marketing mix modeling*; Regressão; Previsão.

Lista de Figuras

3.1	Proporção de receita atribuída às mídias.	3
3.2	<i>Ticket</i> médio por mídia.	3
3.3	Modelo de crescimento logístico	4
3.4	Modelo de crescimento logístico para a mídia A.	5
3.5	Gráfico de envelope para o modelo final.	5
3.6	Divisão do orçamento entre as mídias antes e depois da análise.	6

Sumário

Lista de Figuras	v
Sumário	vi
1 Introdução	1
2 Metodologia	1
3 Resultados e discussões	2
3.1 Análise descritiva	2
3.2 Relação investimento X receita	3
3.3 Modelo final	4
3.4 Divisão sugerida de orçamento	6
4 Conclusão	6
Referências Bibliográficas	8

1 Introdução

Com a crescente utilização dos meios digitais para a realização de compras e contratação de serviços dos mais variados tipos, o *marketing* digital se tornou uma parte indispensável para se ter e manter um negócio bem sucedido. Existem diversas estratégias e essas devem estar alinhadas com o objetivo e o momento do negócio. Para empresas novas no mercado e que possuem como objetivo inicial a divulgação da marca, são mais indicadas as campanhas de divulgação para usuários que ainda não conhecem a marca; para as empresas que já possuem uma base de cliente e estão interessadas em fidelizar essa base, são indicadas campanhas de aproximação da marca com o cliente por meio de produção de conteúdo e *posts* em redes sociais; já para empresas que tem como objetivo aumentar a receita são indicadas as campanhas de *remarketing* que são direcionadas para clientes que tem uma alta chance de finalizar uma compra.

Para diferentes objetivos temos diferentes tipos de campanhas que podem ser veiculadas. Dentro de um mesmo objetivo temos diferentes mídias que são mais apropriadas para veicular essas campanhas. Por exemplo, para campanhas em sites de busca temos como opções veicular anúncios em *Google* ou *Bing*.

Para a construção de uma estratégia global e efetiva de *marketing* devemos definir bem os objetivos, escolher as mídias que mais se alinham à eles e distribuir o orçamento entre as mídias de forma que o retorno seja maximizado. Devido a importância e dificuldade dessa tarefa foi criada uma ferramenta para auxiliar os profissionais do *marketing* tradicional em uma época que as mídias digitais tinham uma pequena participação no *mix* de mídias (LATENTVIEW, 2019).

Com isso, o objetivo desse trabalho é aplicar o modelo de *marketing mix* para um cenário mais atual e, conseqüentemente, digital, propor uma combinação adequada de mídias e maximização do retorno por meio da distribuição ótima do orçamento entre elas.

2 Metodologia

Os modelos MMM relacionam uma métrica que mede o sucesso do negócio com métricas que possivelmente tem uma relação de causalidade com a primeira. As variáveis mais comumente utilizadas nos modelos são os investimentos em cada tipo de *marketing*, variáveis que indiquem promoções que foram feitas, mudanças de preço do produto ou serviço, variáveis macroeconômicas, além das componentes de séries temporais (CHAN and PERRY, 2017).

É sabido que, tanto para o *marketing* digital quanto para o tradicional, em algum ponto de investimento ou exposição de um anúncio o retorno para de aumentar (CLAIRVOYANT, 2018). Isso acontece porque o número de potenciais consumidores já foi completamente atingido pelo anúncio. Sendo assim, temos que transformar as variáveis de investimento para considerar a real relação delas com a métrica de sucesso.

Outra transformação necessária é considerar a defasagem dos investimentos, já que algumas ações de *marketing* não surtem efeito imediato.

Para estudar a relação entre as variáveis escolhidas utilizaremos um modelo linear generalizado com resposta gama, modelo adequado para ser aplicado em dados positivos e assimétricos (PAULA, 2004).

$$Y_i \sim G(\mu_i, \phi)$$

$g(\mu_i)$ é a função que faz a ligação entre a média e o preditor linear η_i :

$$g(\mu_i) = \eta_i$$

O vetor \mathbf{x} é a parte sistemática do modelo:

$$\eta_i = \mathbf{x}_i^T \beta$$

Para a avaliação e seleção do modelo, utilizaremos o critério de informação bayesiano (BIC), que é dado por:

$$BIC = -2\log f(x_n|\theta) + p \log n \quad ,$$

$f(x_n|\theta)$ é o modelo escolhido, p é o número de parâmetros a serem estimados e n é o número de observações da amostra.

A significância dos parâmetros será importante para o ajuste do modelo e interpretação dos resultados já que pode indicar que o efeito do investimento de alguma mídia na variável de sucesso é igual a zero.

A multicolinearidade é um problema bastante encontrado em modelagens com dados de *marketing* digital. Esse problema pode impactar as estimativas dos parâmetros do modelo, sendo assim uma verificação importante a se fazer é tentar diagnosticá-lo através do VIF (*Variance Inflation Factor*) que é calculado por

$$VIF_j = \frac{1}{1 - R_j^2}$$

3 Resultados e discussões

Para todas as etapas da análise será utilizado o *software R*, desde a extração dos dados até a avaliação dos modelos.

Os dados levantados são de um comércio eletrônico de grande porte que faz investimento em 8 mídias diferentes, sendo 2 de *search* (veiculadas em *sites* de busca), 2 de *display* (veiculadas em *sites* que oferecem espaços para *banners*) e 4 de *remarketing* (impactam clientes que estão muito próximos de finalizar a compra). Foram obtidos dados diários de investimento de cada mídia, a receita diária de cada mídia e receita total do site do período de janeiro de 2017 a agosto de 2019, portanto são 974 observações. O objetivo do cliente ao solicitar a análise do *mix* de mídias, é diminuir a quantidade de mídias de *remarketing*.

3.1 Análise descritiva

Analisando o ano de 2019, temos que as 8 mídias pagas analisadas são responsáveis por cerca de 60% da receita do *site*, sendo 28% proveniente de somente uma mídia, como pode ser observado na Figura 3.1. Apesar disso, as mídias com maior participação na receita possuem menor *ticket* médio, ou seja, menor receita por compra como pode ser visto na Figura 3.2.

As mídias de *search* recebem em média 35% do investimento mensal, as de *remarketing* cerca de 12% enquanto as de *display* recebem os 53% restantes. Cerca de 80% das conversões acontecem no mesmo dia que o cliente teve o primeiro contato com alguma mídia.

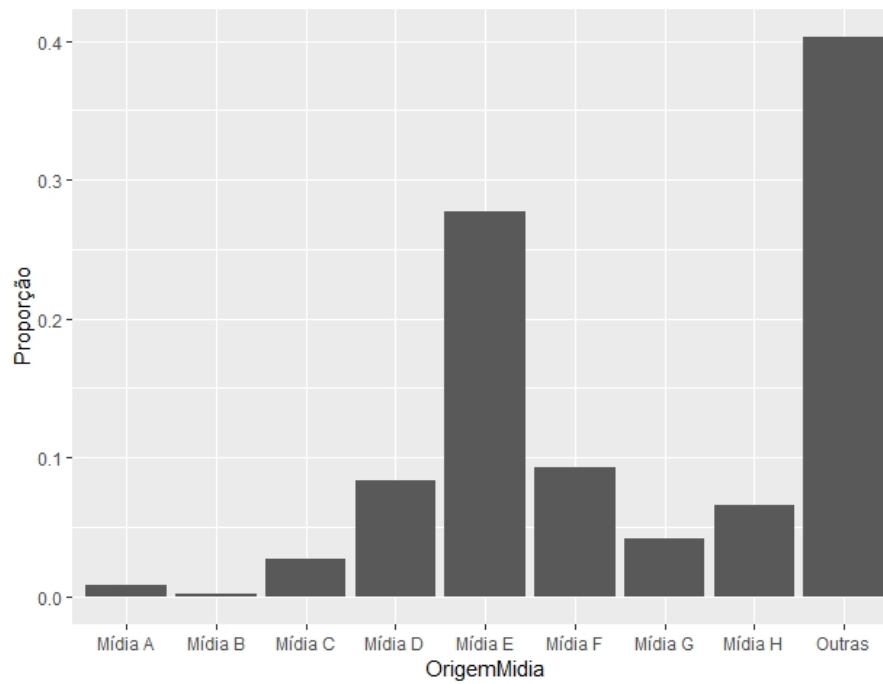


Figura 3.1: Proporção de receita atribuída às mídias.

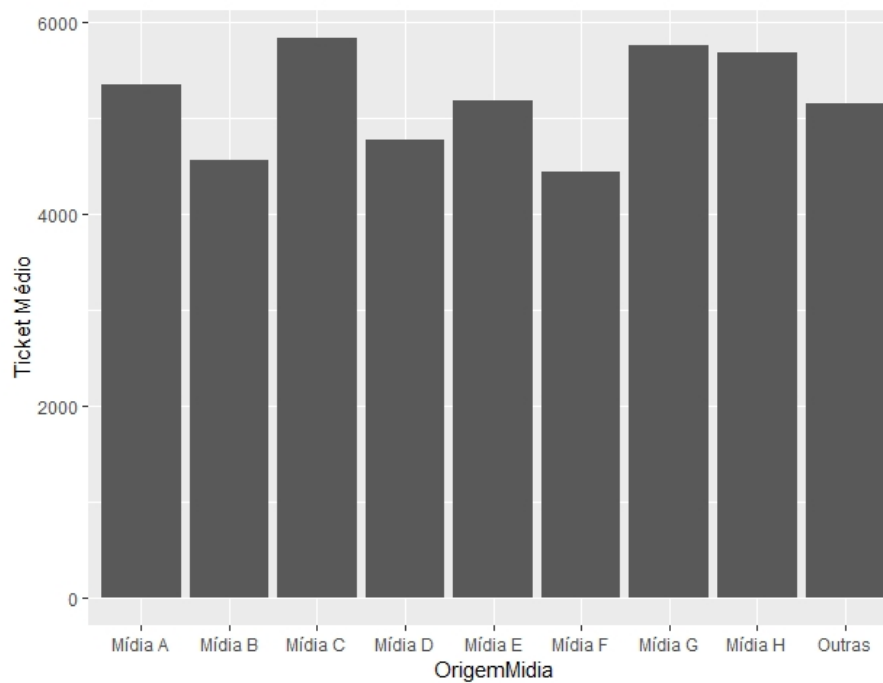


Figura 3.2: *Ticket* médio por mídia.

3.2 Relação investimento X receita

Como foi dito na Seção 2, é esperado que, em algum momento, mesmo com o aumento do investimento a receita pare de crescer. Para verificar se esse comportamento em forma de "S" aparece nas mídias estudadas, foram ajustados modelos de crescimento logístico que são dados por:

$$Y = \frac{\alpha}{1 + e^{\frac{\beta - x}{\gamma}}},$$

em que α representa o valor máximo de receita que a curva atinge, β é o ponto de inflexão da curva e γ regula a inclinação da curva, como pode-se observar na Figura 3.3.

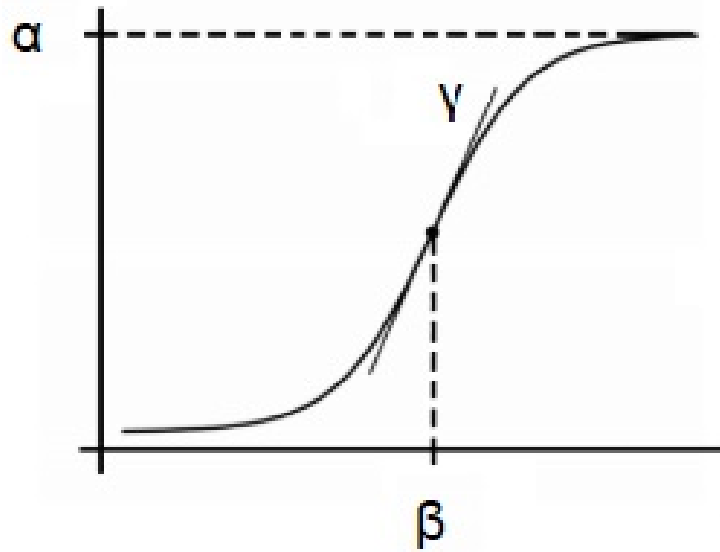


Figura 3.3: Modelo de crescimento logístico

Intervalos de investimento ideal para cada mídia foram gerados considerando o ponto de inflexão, β , e o ponto no qual a média móvel do ROI incremental é igual a média dos retornos. Sendo ROI a razão entre receita e investimento e ROI incremental a razão entre a diferença entre $Receita_{t+1}$ e $Receita_t$ e a diferença de $Investimento_{t+1}$ e $Investimento_t$.

Na Figura 3.4, temos a representação do modelo de crescimento logístico aplicado para cada combinação diária de investimento e receita da mídia A. As barras verticais representam o intervalo obtido.

3.3 Modelo final

Após levantamento de possíveis variáveis com o solicitante da análise, desconsideramos a relação com variáveis macroeconômicas, o preço também não foi considerado já que a empresa oferece diversos produtos. As variáveis incluídas no modelo foram os investimentos em cada mídia.

Apesar de ter sido proposto na metodologia a utilização de um modelo linear que tinha como regressoras os investimentos das mídias transformados, após a avaliação do conjunto de dados um modelo linear generalizado com variável resposta seguindo uma distribuição gama foi mais adequado. Foram ajustados modelos com função de ligação identidade, log e inversa, pela avaliação do BIC o modelo escolhido foi com função de ligação identidade.

Ao analisar os resíduos dos modelos ajustados, foram identificados pontos influentes através da distância de Cook e pontos discrepantes através do teste de Bonferroni. Após a análise e retirada de alguns pontos, o modelo ajustado apresentou resíduos aleatoriamente distribuídos e um gráfico de envelope dentro das bandas de confiança como pode-se observar na Figura 3.5.

A multicolinearidade também foi avaliada e, apesar de ser um problema recorrente em dados de *marketing* digital, não foi detectada no modelo. Os valores do VIF ficaram entre 1,1 e 1,7.

Com a interpretação do modelo final tivemos uma mídia de *remarketing* para a qual o co-

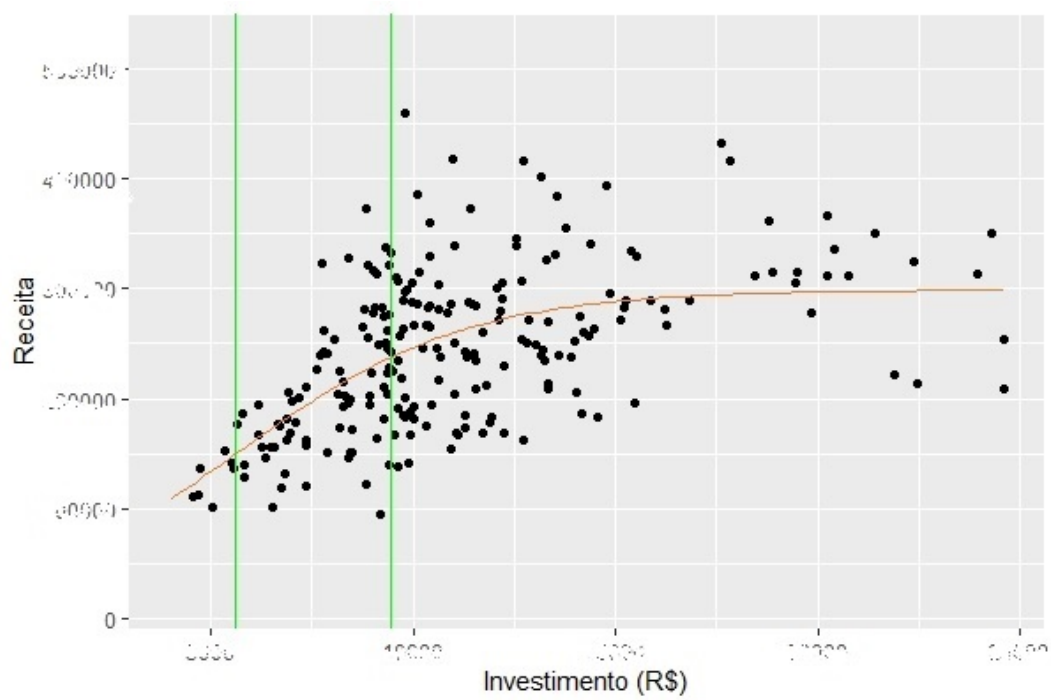


Figura 3.4: Modelo de crescimento logístico para a média A.

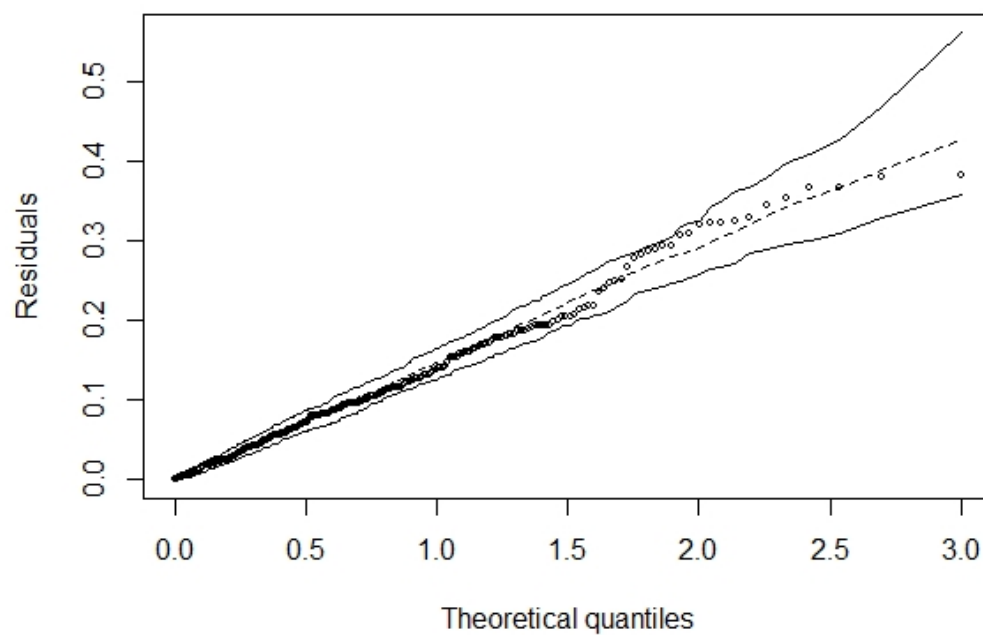


Figura 3.5: Gráfico de envelope para o modelo final.

eficiente não foi estatisticamente diferente de 0 ($\alpha = 0,05$). Essa resposta está alinhada com uma necessidade do cliente ao solicitar a análise, que era de diminuir o número de mídias de *remarketing*.

3.4 Divisão sugerida de orçamento

Para a divisão ótima do orçamento foi utilizado o pacote *lpSolveAPI* que fornece funções para resolver problemas de otimização linear. Para a aplicação é necessário definir duas componentes: uma função objetivo e as restrições para os valores das variáveis.

Neste caso, a função objetivo era dada pela combinação dos coeficientes resultantes do modelo final:

$$Receita = \beta_1 Midia_A + \beta_2 Midia_B + \dots + \beta_8 Midia_H ,$$

Os intervalos obtidos na Seção 3.2 foram utilizados como restrições do algoritmo de otimização, além da restrição que definia o valor máximo diário de investimento.

Como resultados da aplicação do algoritmo nas condições definidas anteriormente, temos um ROI (Receita/Investimento) 24,8% maior do que a média histórica. Na Figura 3.6 pode-se comparar a distribuição do orçamento que era utilizada anteriormente e a distribuição proposta através da otimização, a proporção das duas mídias mais importantes para a conta não mudou muito, a realocação entre as mídias de apoio foi maior.

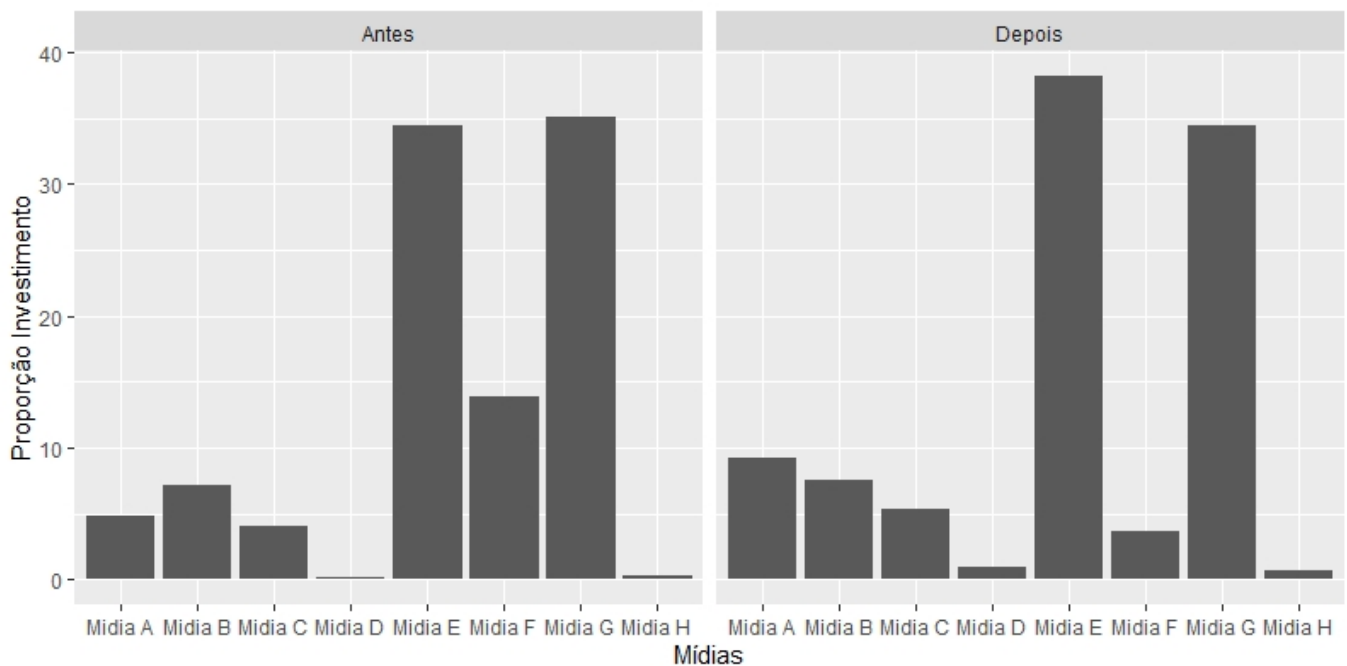


Figura 3.6: Divisão do orçamento entre as mídias antes e depois da análise.

4 Conclusão

A aplicação de modelos MMM para um cenário digital, no contexto desse trabalho, foi diferente do que sugerem diversas metodologias encontradas. Não houve dificuldade para a obtenção dos dados e a realidade do marca estudada parece não ter relação com a situação atual da economia. Outra diferença encontrada foi quanto ao período de dados que seria analisado, no contexto tradicional as empresas sugerem que sejam avaliados mais de 2 anos de histórico, enquanto no contexto digital é sugerido que a janela de dados seja diminuída já que as plataformas de anúncios e a utilização da *internet* são muito dinâmicas e sofrem mudanças muito frequentemente (FACEBOOK, 2018).

Os resultados foram apresentados para o cliente e as respostas obtidas através das análises vão bastante de encontro do que é acompanhado no dia a dia dos responsáveis pela conta. Planeja-se que as sugestões sejam aplicadas e testadas em janeiro de 2020, após o período sazonal de final de ano.

Como possíveis próximos passos temos a aplicação da mesma técnica em algum grupo de interesse. Por exemplo, podemos avaliar a distribuição de orçamento entre as mídias para os cenários de *desktop* e *mobile* separadamente; outro exemplo é avaliar por região do país.

Referências Bibliográficas

- [CHAN and PERRY, 2017] Chan, D.F., Perry, M. *Challenges and Opportunities in Media Mix Modeling*, 2017. Disponível em: <https://ai.google/research/pubs/pub45998>. Acesso em: 10 de Setembro de 2019.
- [CLAIRVOYANT, 2018] Clairvoyant. *Market Mix Modeling (MMM) — Introduction, Methodology, and Use case*, 2018. Disponível em: <https://blog.clairvoyantsoft.com/market-mix-modeling-mmm-introduction-methodology-and-use-case-dc5ae68820f8>. Acesso em: 18 de Setembro de 2019.
- [FACEBOOK, 2018] Facebook IQ. *Measuring Facebook accurately in marketing mix models*, 2019. Disponível em: <https://www.facebook.com/business/news/insights/considerations-for-creating-modern-marketing-mix-models>. Acesso em: 13 de Outubro de 2019.
- [LATENTVIEW, 2019] Latentview *A complete guide to Marketing Mix Modeling*, 2019. Disponível em: <https://www.latentview.com/marketing-mix-modeling/>. Acesso em: 13 de Setembro de 2019.
- [PAULA, 2004] Paula, G. A. (2004). *Modelos de regressão: com apoio computacional*. São Paulo: IME-USP.