

Análise Estatística em Marketing Digital

Orientador: Adriano Kamimura Suzuki

Aluno: Roger Gregório Marcondes



Projeto estabelecido

01

Problema

Crescimento e disseminação de **conteúdos relevantes**;

Informações disponíveis, em quantidade e velocidade;

Volatilidade de informações.

02

Consequências

Mudança acelerada de **produtos e processos**;

Uso de **novas ferramentas**.



03

Solução

Atingir o cliente em potencial, no exato momento em que ocorre a busca de um produto, serviço ou qualquer outra informação.

04

Metodologia

Regressão Linear

Variável resposta:
Faturamento

Variáveis explicativas:
Alcance, Frequência, Impressões, Cliques, Taxa de Cliques entre outras.

05

Resultado

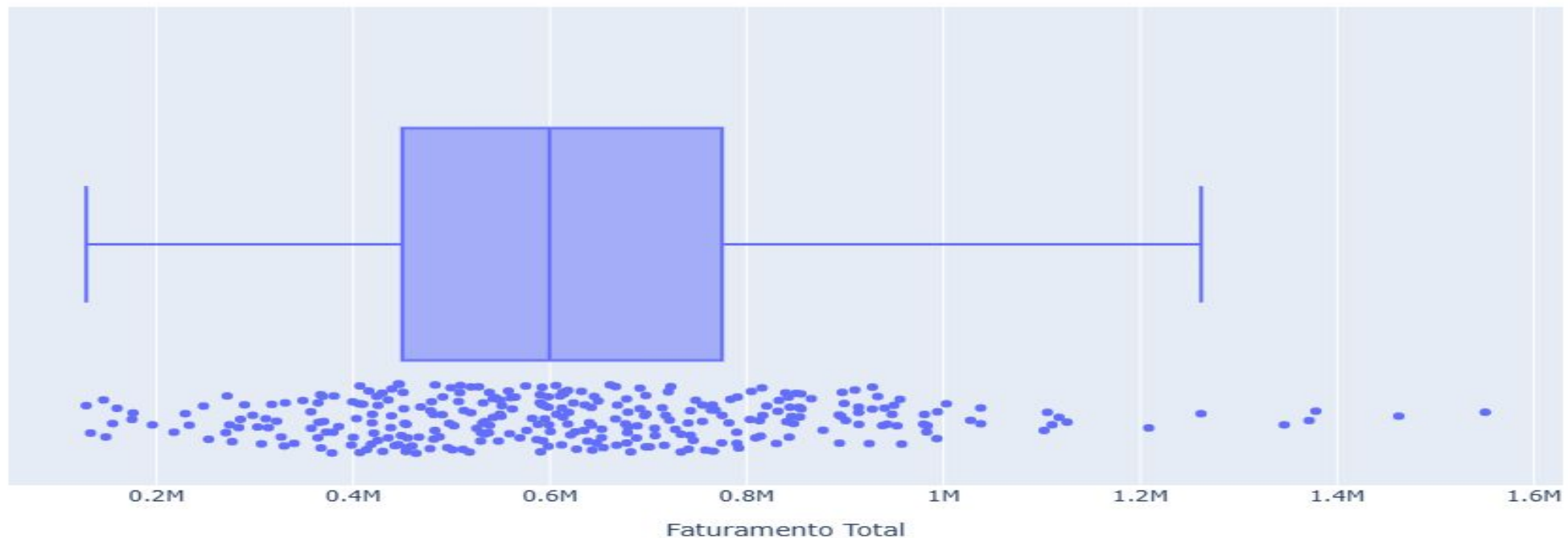
Avaliar quais canais e métricas de mídia vão impactar de forma positiva na variável resposta, com esse modelo podemos **prever o resultado futuro** de acordo com as alterações de investimentos ao longo do tempo.

Análise de Regressão Múltipla

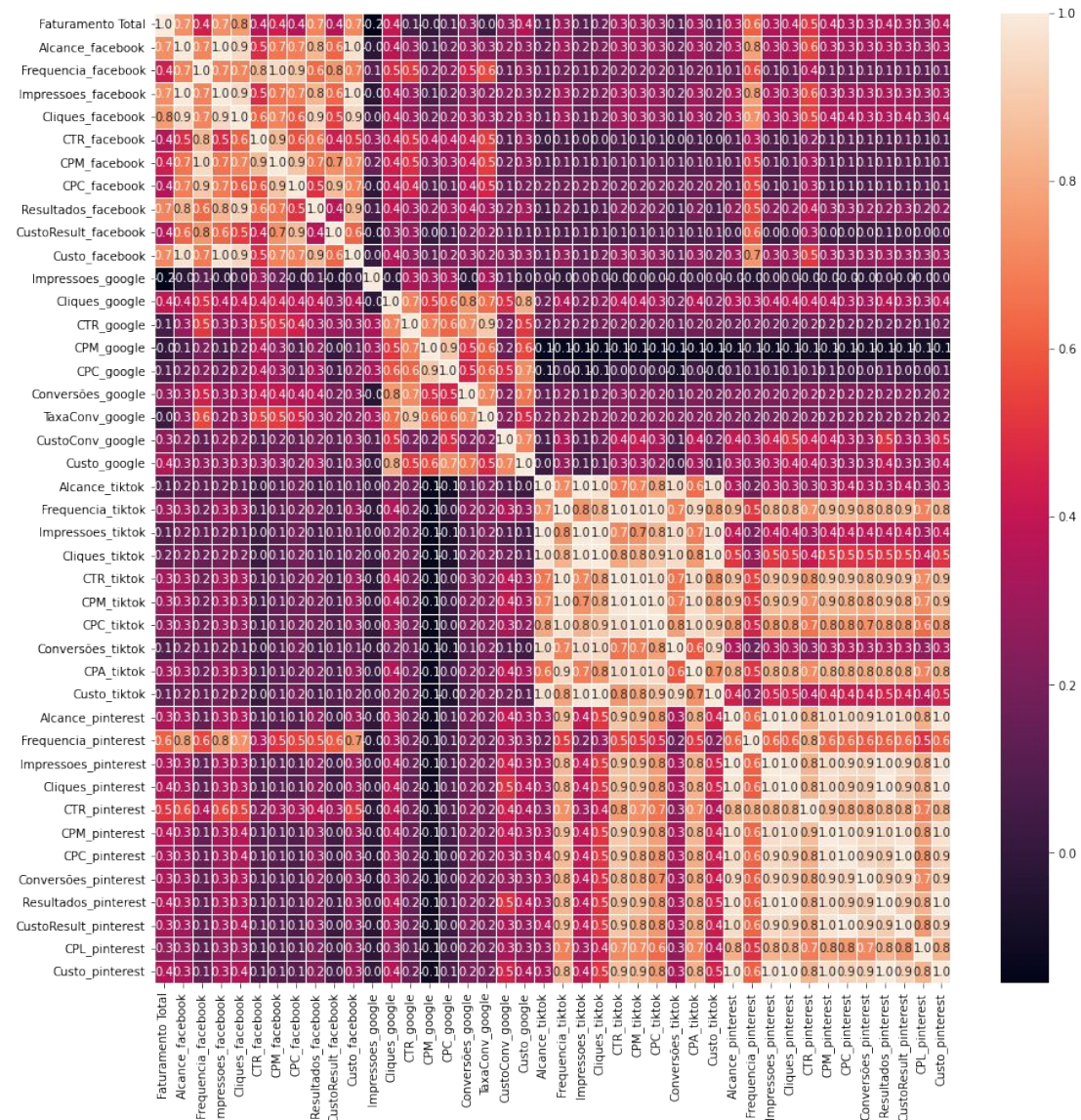
Etapas da Análise		
	01 Pré-Processamento	Limpeza dos dados: Remoção de valores nulos, padronização das métricas nos distintos canais de divulgação
	02 Análise Descritiva	Descrever as principais tendências nos dados existentes
	03 Definição de Modelos	Com base em correlação selecionar métricas apropriadas , com potencial de descrever a variável dependente
	04 Avaliação de Modelos	Utilizando K-Fold evitar overfit e underfit utilizando R², AIC e BIC
	05 Predições	Com o conjunto de teste realizar predições e avaliar a capacidade do modelo de predizer novas amostras
	06 Teste de Hipótese	Sobre modelo
	07 Conhecimento	Conclusão sobre o estudo

Análise Descritiva

Variável	Média	Desvio Padrão	Mínimo	Máximo	Qtd.
Faturamento	R\$623.516,20	R\$ 243.130,1	R\$ 128.677,6	R\$ 1.550.524,00	304



Análise Descritiva



Definição de Modelos

Simplificação do Modelo:
Reduzir Correlação entre variáveis independentes

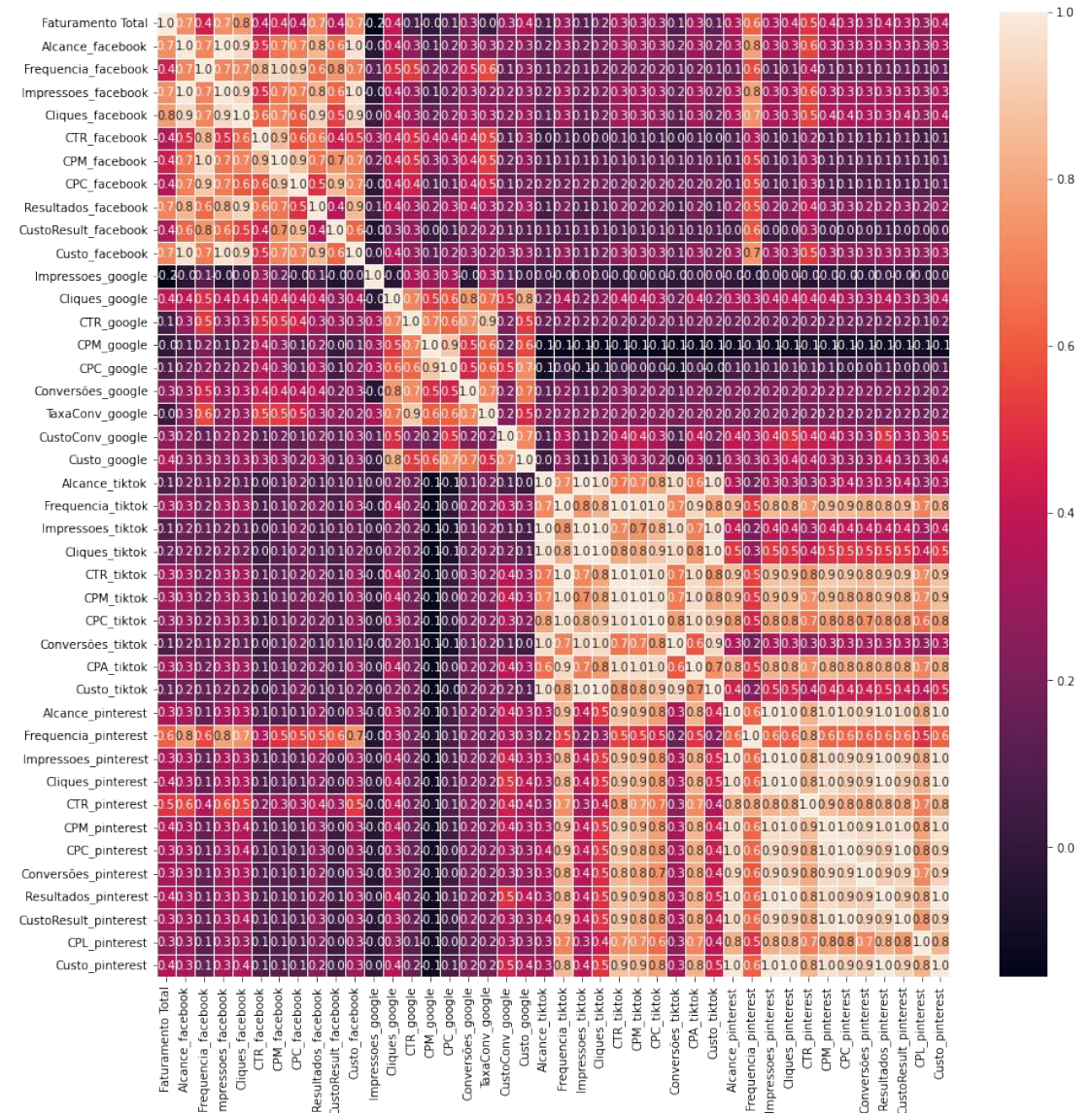
EXEMPLOS:

$$\text{CTR} = \frac{\text{Cliques}}{\text{Impressões}}$$

A taxa de cliques é a proporção de usuários que clicam em um link específico em relação ao número total de usuários que visualizam.

$$\text{Frequência} = \frac{\text{Impressões}}{\text{Alcance}}$$

A Frequência é a proporção de usuários já alcançados que visualizam ou interagem com o conteúdo disseminado.



Avaliação de Modelos

Fator de inflação de variância (em inglês VIF)						Intervalo de Confiança		
Primeira Rodada		Sexta Rodada		Décima Sexta Rodada		p-valores	0,025	0,975
Alcance_facebook	14017.88	Impressoes_facebook	176.46	Cliques_google	8.51	0,000	233,901	278,342
Impressoes_facebook	12383.83	Cliques_facebook	79.19	Conversões_google	6.64	0,006	22,242	132,399
Cliques_facebook	1227.7	Custo_facebook	73.65	Cliques_facebook	2.49	0,000	263,268	485,459
CTR_facebook	681.48	CPC_facebook	47.56					
CPC_facebook	605.47	TaxaConv_google	45.56					
CPM_facebook	574.69	Resultados_facebook	44.42					
Custo_facebook	496.33	Custo_google	31.6					
Frequencia_facebook	451.28	CustoResult_facebook	19.82					
CTR_google	228.41	Cliques_google	19.46					
CPM_google	216.13	Conversões_google	17.11					
TaxaConv_google	149.34	CustoConv_google	11.72					
CPC_google	144.06	Impressoes_google	3.45					
CustoResult_facebook	114.38							
Custo_google	83.95							
Resultados_facebook	78.57							
Cliques_google	54.4							
CustoConv_google	54.21							
Conversões_google	43.97							
Impressoes_google	10.61							

Avaliação de Modelos

Base de Teste

MODELO



R^2

Explicabilidade
do modelo na
base de teste



MSE

Erro
quadrático
médio

Parâmetros β

Métrica_Canal	Coef.	p-value	0.025	0.975
Cliques_facebook	256.12	0.000	233.901	278.342
Cliques_google	77.32	0.006	22.242	132.399
Conversões_google	374.36	0.000	263.268	485.459

Teste de Hipótese

Hipótese Nula (H0): Coeficiente não representativo

$$\beta=0$$

$$p>0,05$$

Hipótese Alternativa (H1): Coeficiente representativo

$$\beta\neq 0$$

$$p\leq 0,05$$

	coef	std err	t	P> t	[0.025	0.975]
const	3.254e+05	1.98e+04	16.465	0.000	2.86e+05	3.64e+05
x1	256.1218	11.267	22.732	0.000	233.901	278.342
x2	374.3636	56.333	6.646	0.000	263.268	485.459
x3	77.3208	27.928	2.769	0.006	22.242	132.399

Parâmetros β

Métrica_Canal	Coef.	p-value	0.025
Cliques_facebook	256.12	0.000	Rejeitamos H0
Cliques_google	77.32	0.006	Rejeitamos H0
Conversões_google	374.36	0.000	Rejeitamos H0

Avaliação de Modelos

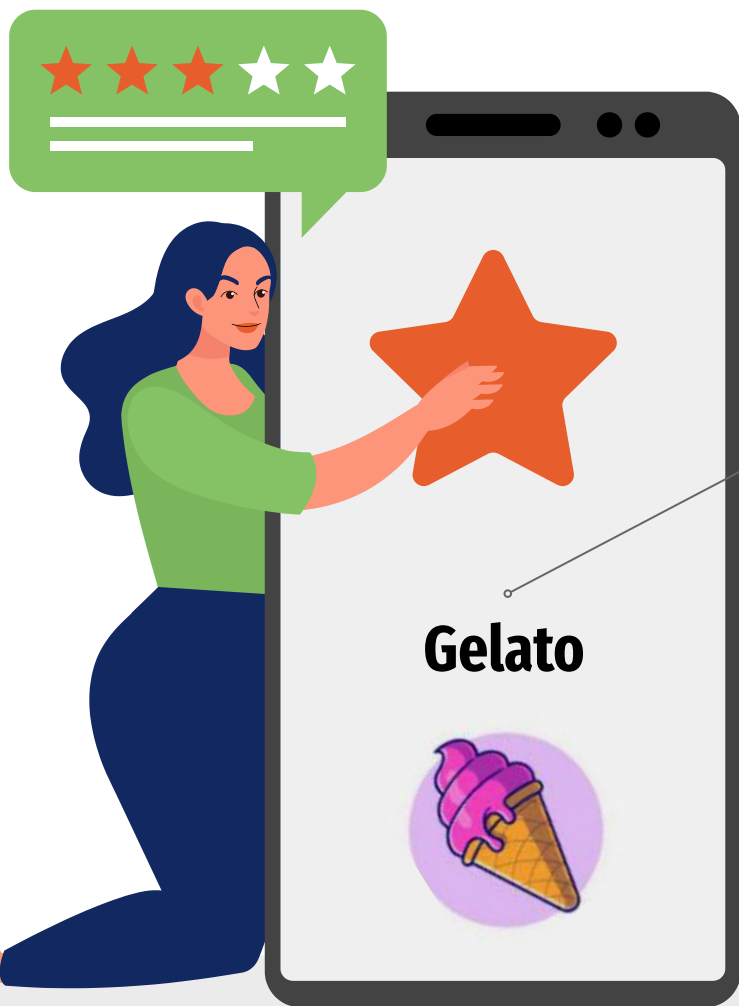
Variável	Média	Desvio Padrão	Mínimo	Máximo
Faturamento	R\$ 689.483,10	R\$ 212.245,20	R\$ 330.955,30	R\$ 1.550.524,0
Cliques_facebook	789,42	644,81	35	3137
Conversões_google	296,86	141,46	31	1285,54
Cliques_google	656,33	293,37	85	1748

Predição

Amostra mês de outubro para predição



Conhecimento



01 Tendências

O produto possui sazonalidade.

02 Relação

Verificou-se que algumas métricas já utilizadas para desenvolver campanhas, possuem relevância.

03 Rejeição

Métricas que apresentam resultado negativo demonstraram falta de relação no modelo.

04 Coleta de dados

Existe dificuldade de filtro de informação sobre faturamento.

Bibliografia

- Organização Pan-Americana de Saúde. Folha informativa sobre COVID-19. Organização Pan- Americana de Saúde. Disponível em: <<https://www.paho.org/pt/covid19>>. Acesso em: 20 de jun. de 2021.
- Amplifica Digital. Guia definitivo do marketing digital B2b. Equipe Amplifica Digital. Disponível em: <<https://amplificadigital.com.br/blog/importancia-do-marketing-digital/>>. Acesso em: 20 de jun. de 2021.
- Sorrentino, J. Guia de estatística – Regressão linear. Ecommerce Brasil. Disponível em: <<https://www.ecommercebrasil.com.br/artigos/guia-de-estatistica%E2%80%8A-%E2%80%8Aregressao-linear/>>. Acesso em: 20 de jun. de 2021.