

Métricas do período após o blog (valores médios)

transacoes_site
21.289,85
↓ -7.3%

usuarios_site
115.498,43
↑ 12.5%

taxa_conversao_site
18,51%
↓ -17.2%

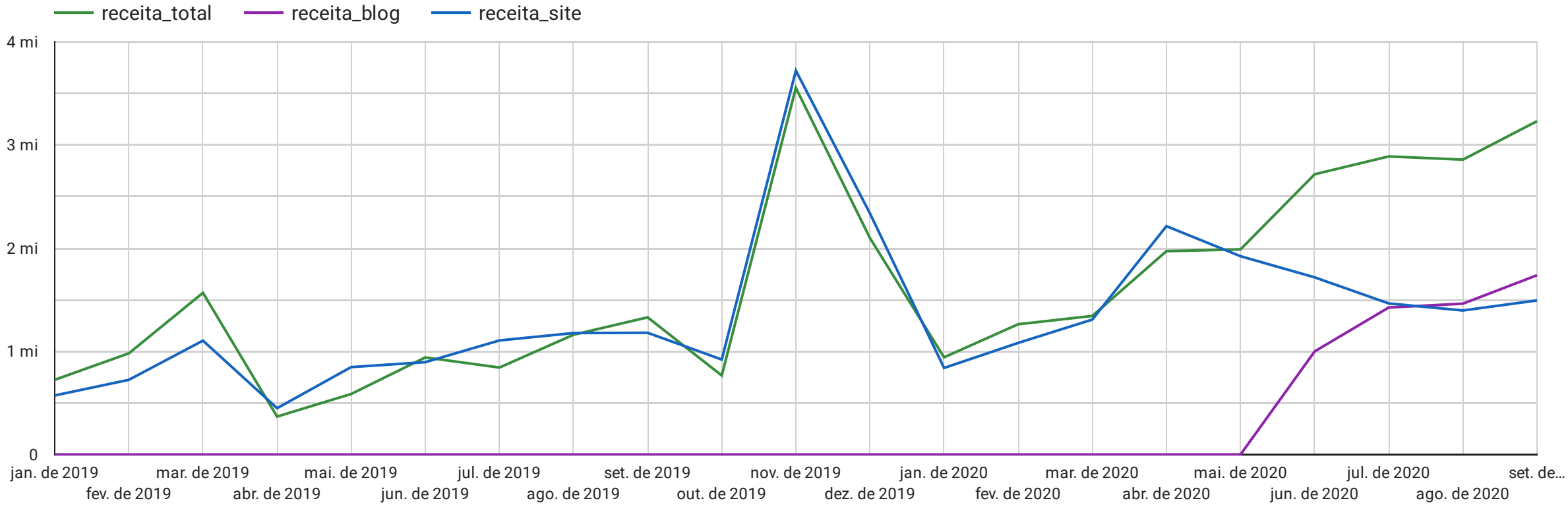
receita
2.921.583,47
↑ 77.7%

transacoes_blog
2.767,34
↑ N/A

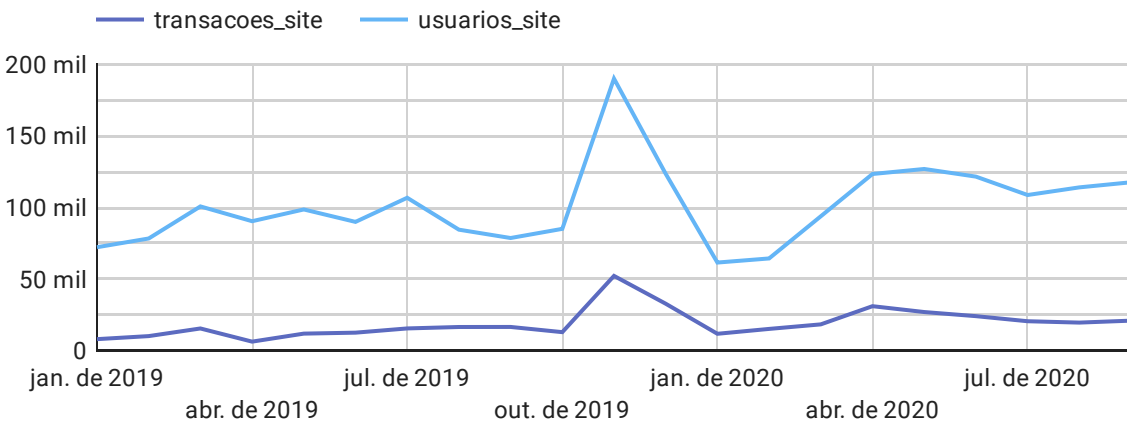
usuarios_blog
7.541,5
↑ N/A

taxa_conversao_blog
43,99%

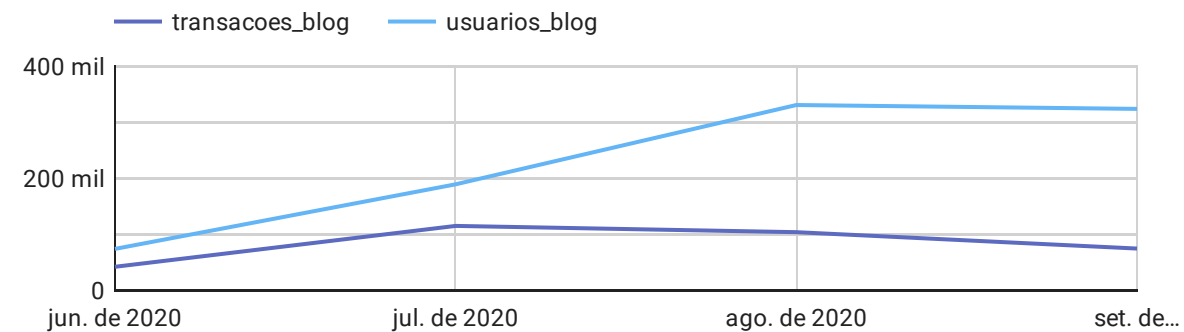
Média da receita por mês



Site: média de transações e usuários



Blog: média de transações e usuários



PROBLEMA DE NEGÓCIO E MÉTODO DE ANÁLISE

Este projeto foi desenvolvido para auxiliar a empresa X nas tomadas de decisões envolvendo o blog da marca. O objetivo é entender a influência que esse novo canal, criado há cerca de 3 meses, está exercendo na receita da empresa.

Assim, neste relatório, apresentamos uma análise exploratória dos dados, oferecendo possíveis insights para a resolução do problema de negócio. Para atender esse objetivo buscaremos:

- Detectar padrões e desvios nos dados;
- Identificar as variáveis mais relevantes e de que forma se relacionam entre si;
- Encontrar tendências e formular hipóteses;
- Realizar inferências estatísticas básicas;

CONTEXTO

Status:

A empresa já possui um site de vendas que funciona há um tempo;
A empresa criou um blog com conteúdo há aproximadamente 3 meses;

Perguntas iniciais:

Qual o impacto que o blog tem gerado na receita da empresa desde a sua criação? Aumentou?
É possível saber se é viável para a empresa (em termos de receita) manter ou não o blog ativo?
Qual canal exerce maior impacto na receita? Blog? site? Os dois?

Base de dados a ser utilizada na resolução do problema:

Data
O número total de pessoas que acessaram o blog por dia;
O número de pessoas que acessaram o site e não passaram pelo blog por dia;
O número total de compras de pessoas que passaram pelo blog por dia;
O número total de compras de pessoas que não passaram pelo blog por dia;
Receita das compras que passaram ou não pelo blog por dia receita.

ANÁLISE EXPLORATÓRIA

Esboço inicial

No nosso primeiro passo manipulando a base de dados, analisamos e destacamos as variáveis principais, com as quais iremos trabalhar. Assim, no próprio arquivo excel, montamos um esboço com algumas informações gerais das variáveis. Dividimos os dados em duas variáveis: “antes do blog” e “depois do blog”, já que essa é a linha temporal que marca o nosso problema. O resultado foi essa tabela ao lado:

O primeiro padrão encontrado foi um aumento de mais de 1,5 milhão na receita após a criação do blog. E isso pode ser constatado tanto nos valores totais quanto nas médias por mês. Como o período de tempo antes do blog é maior, consequentemente o valor da receita também será mais alto. Por isso, para obter valores mais ‘manipuláveis’, utilizamos a média desses dois períodos como base de análise para as variáveis “usuários site”, “usuários blog” e “receita”. Além disso, fizemos um recorte dos últimos quatro meses da empresa funcionando apenas com o site e comparamos com os meses depois do blog, para trabalharmos com variáveis dentro do mesmo período. Como resultado, ainda obtemos um aumento na receita no período do blog, tanto com a comparação da média dos valores como com a comparação entre os recortes.

Porém, apenas com essas informações não é possível afirmar se esse aumento foi decorrente da existência do blog, principalmente porque a maior fonte de receita ainda está nas transações do site, como podemos observar na tabela acima. Por outro lado, esses valores atrelados as taxas de conversão e as taxas de crescimento, dão um forte indício de que o blog teve um impacto significativo na receita. Para obter um resultado mais próximo dessa confirmação ainda analisaremos mais detalhadamente alguns dados, como a média da receita antes e depois do blog, as transações, os usuários e o valor médio por transação de ambos os canais.

ANTES DO BLOG						
Variável	Total	Média	taxa de conversão	Receita	Total	Média
usuarios_site	50.838.280,51	98.333,23	18,82%	Receita	681.233.277,38	1.644.073,61
transacoes_site	9.568.651,96	18.508,03				

4 MESES ANTES DO BLOG						
Variável	Total	Média	taxa de conversão	Receita	Total	Média
usuarios_site	12.421.605,40	102.657,90	22,37%	sem blog	R\$ 198.932.906,99	1.317.665,91
transacoes_site	2.778.693,11	22.964,41				

DEPOIS DO BLOG						
Variável	Total	Média	taxa de conversão	Receita	Total	Média
usuarios_site	14.080.804,75	115.498,40	18,43%	com blog	R\$ 356.433.182,91	2.921.583,47
transacoes_site	2.597.365,98	21.289,89				
usuarios_blog	920.062,86	7.541,50	36,69%			
transacoes_blog	337.616,53	2.767,35				

Taxa de crescimento após o blog		
usuários_site		13%
transações_site		-7%
receita último trimestre		78%

ANÁLISE EXPLORATÓRIA

Antes do blog x depois do blog: Média da receita por mês

O próximo passo foi analisar a base de dados a partir de outras ferramentas de análise que favorecem maior detalhamento e visualização gráfica. Utilizamos a linguagem de programação Python, juntamente com as bibliotecas Pandas, Matplotlib, Plotly e outras.

Nesse novo ambiente, reproduzimos o esboço do excel, com mais detalhes e possibilidade de analisar visualmente os dados. Para isso, primeiramente, geramos um gráfico apenas para verificar a evolução da receita total ao longo do tempo.



Confirmando nossa verificação anterior, podemos enxergar um aumento na receita após a inclusão do blog. Por outro lado, a receita já estava seguindo um percurso crescente anteriormente. Então, mais uma vez, ainda precisamos de mais dados para confirmar nossas hipóteses. Outro detalhe interessante é que houve um pico (outlier) no faturamento no final do ano de 2019 que está destoando da média. Será preciso descartar esse pico para que não influencie negativamente nos nossos resultados.

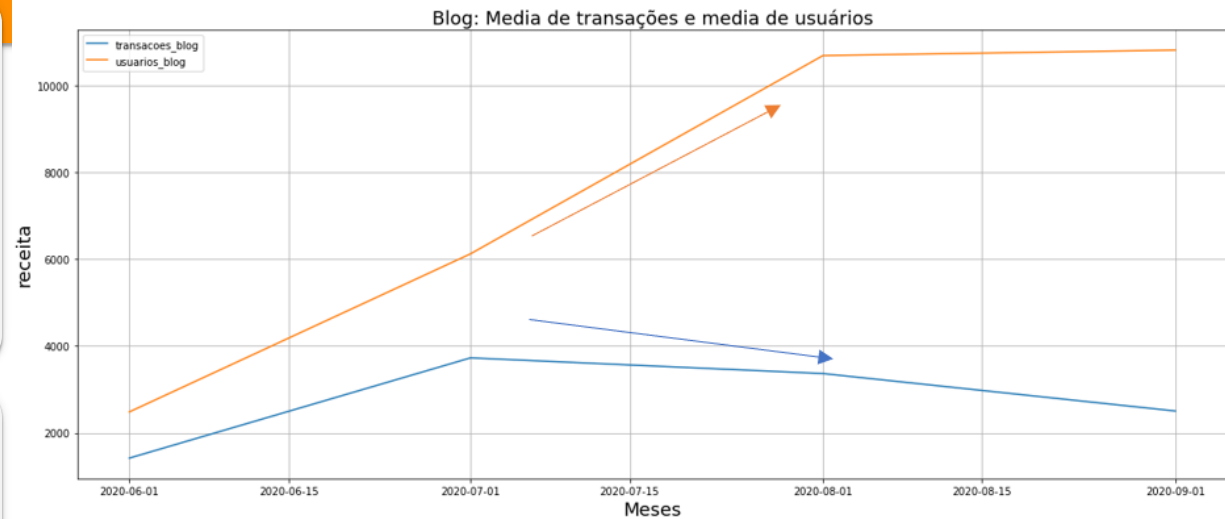
O próximo passo foi dividir o dataframe (base de dados) em dois, os quais chamamos de `df_site` (contendo apenas dados da receita, das transações e dos usuários do site no período antes da criação do blog) e `df_blog` (contendo apenas dados da receita, das transações e dos usuários no período depois da criação do blog).

ANÁLISE EXPLORATÓRIA

BLOG: Média de usuários x media das transações

Conforme vimos no gráfico anterior, ignorando o outlier da receita no final de 2019, pode-se perceber que a receita estava seguindo um curso crescente, que foi potencializado com a chegada do blog. Entretanto, para investigar melhor essa relação, foi preciso fragmentar um pouco o nosso percurso e analisar a relação entre as variáveis transações x usuários para ambas as bases de dados (df_site e df_blog). Isso porque esse aumento na receita poderia ser decorrente das transações do site e não do blog. Assim, inicialmente, analisamos as variáveis “transações_blog” e “usuários_blog”, pra tentar encontrar alguma correlação. Podemos observar isso no gráfico ao lado:

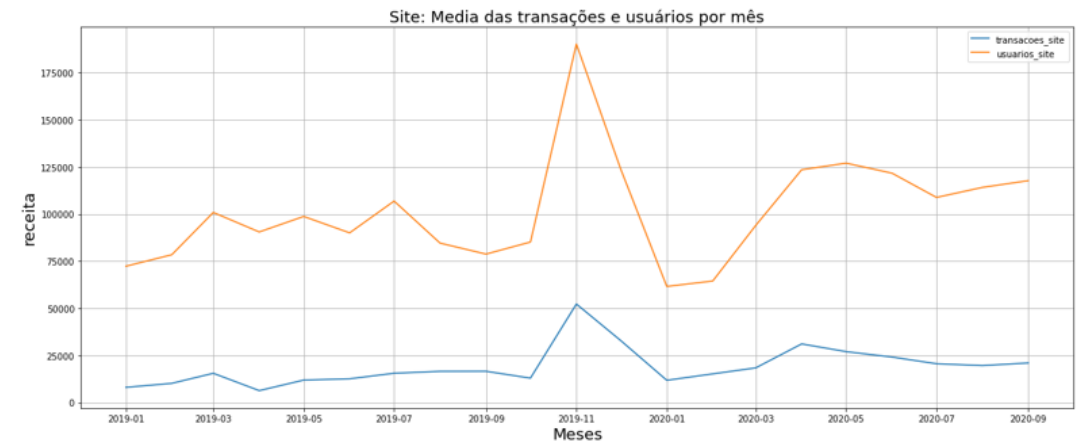
Podemos observar que o número de transações do blog estava crescendo e o número de usuários estava acompanhando esse crescimento. Porém, a partir do mês 07, houve uma queda no número de transações que acabou quebrando essa correlação transações x usuários. Seria interessante verificar com a equipe responsável o que houve nesse período que pode ter ocasionado essa queda, pois apenas a partir desses dados não foi possível identificar a raiz do problema. Um fato curioso é que apesar dessa baixa nas transações, o blog ainda acompanhou o aumento da receita.



SITE: Media de usuários x media das transações

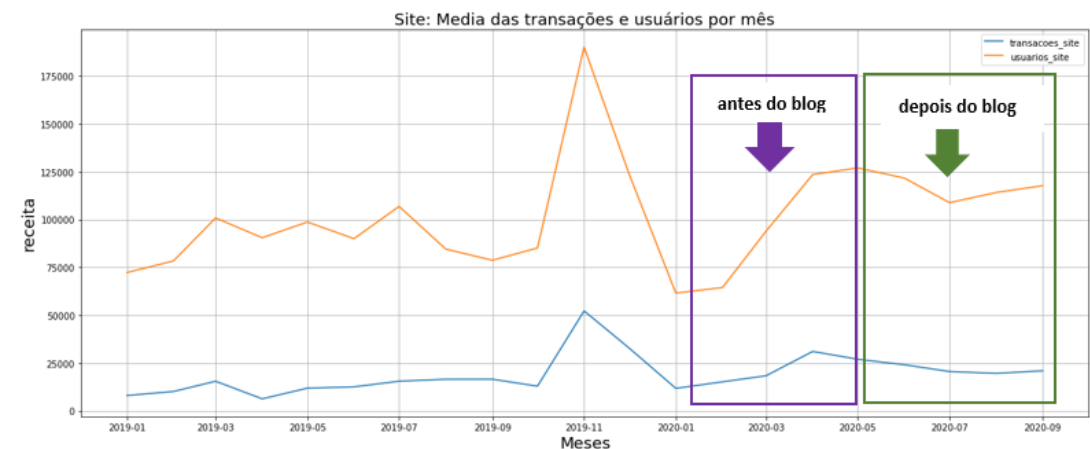
Diferentemente do que aconteceu no blog, a média de usuários do site seguiu acompanhando a média das transações, conforme podemos observar no gráfico seguinte.

Entretanto, com a chegada do blog, esse cenário mudou um pouco. Comparando os últimos quatro meses antes do blog com os três meses após a sua criação, podemos observar que houve uma queda sutil na quantidade de transações e um leve crescimento no número de usuários. De acordo com os dados obtidos inicialmente no esboço, houve uma taxa de crescimento de 13% no número de usuários e uma queda de 7% no número de transações no site.



Algumas perguntas poderiam ser feitas para entender essa ocorrência e, a partir disso, poderíamos levantar algumas hipóteses. Dentre elas:

- 1) A empresa atualmente conta com dois canais de relacionamento com o cliente. Assim, alguns usuários, podem estar realizando as transações diretamente no blog, sem passar pelo site. Porém, não dispomos de informações suficientes para confirmar isso. Uma das investigações a serem feitas para validar essa hipótese seria, por exemplo, fazer um levantamento de dados históricos dos usuários que realizavam transações pelo site e passaram a fazê-las pelo blog.
- 2) A empresa pode ter aumentado o valor dos produtos nesse período. Isso justificaria a alta na receita, apesar da queda nas transações.



ANÁLISE EXPLORATÓRIA

Afinal, qual o impacto das transações do blog na receita total da empresa?

Para se chegar a esse resultado, primeiramente, seria necessário identificar a receita de cada transação do blog. Entretanto, não é possível encontrar esse valor a partir da nossa base de dados, já que não existem receitas específicas do site ou do blog, contamos apenas com a variável “receita”, que corresponde a valores advindos de ambos os canais.

Por outro lado, podemos contornar isso e obter uma informação aproximada acerca da influência do blog na receita total, já que possuímos os dados das transações do site antes da criação do blog. Assim, encontrando o valor médio de cada transação do site, podemos obter o valor da receita do site em qualquer período. Esse valor por transação no site pode ser obtido a partir do seguinte cálculo:

Valor por transação site = Média da receita do site / Média das transações do site

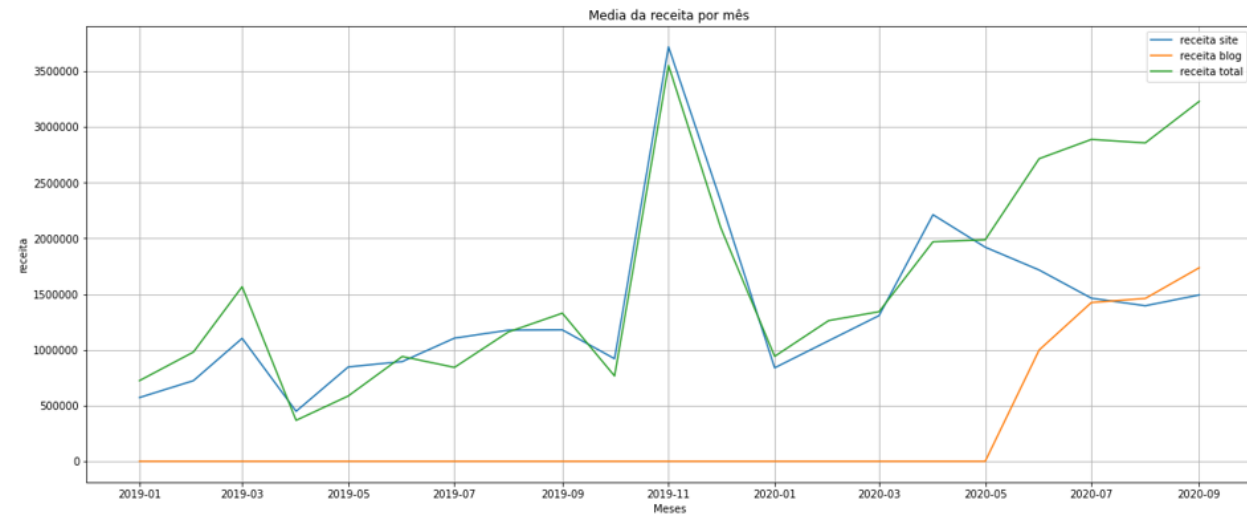
Feito isso, podemos inferir o valor por transação do blog e, consequentemente, a média da receita dele. Chegamos a esse valor a partir dos cálculos:

- 1) **Receita = Receita site + Receita blog**
- 2) **Receita blog = Receita total – Receita site**

O resultado foi o seguinte:

- **Valor médio por transação do site: 71,19**
- **Valor médio por transação do blog: 508,05**

Assim, a partir disso, conseguimos, finalmente, obter o valor individual aproximado de cada um desses canais (blog e site) e enxergar o impacto deles na receita total.



Conforme podemos observar, no período do blog, a receita do site teve uma queda, porém, a receita total continuou aumentando, o que é surpreendente, já que o site conta com mais transações do que o blog, considerando a correlação transação e receita. Por outro lado, observamos que o valor médio por transação do blog é mais alto do que o do site, o que poderia justificar esse impacto na receita. Para fortalecer essa hipótese, podemos olhar para a taxa de conversão

- **TAXA DE CONVERSÃO SITE: 37%**
- **TAXA DE CONVERSÃO BLOG: 18%**

Ou seja, embora o blog conte com menos usuários e transações do que o site, conforme observamos anteriormente, consegue ter uma taxa de conversão de vendas maior. Esse fator aliado às outras variáveis já analisadas, “transações”, “usuários” e “receita”, leva a crer que o blog tem sido um canal importante para a receita da empresa.

CONCLUSÃO

Com os dados explorados até o momento, obtemos os seguintes resultados:

- Houve um aumento de 1,6 milhão na receita total após a criação do blog ;
- Comparando os últimos 4 meses antes do blog até período de 4 meses após a sua criação, houve um aumento de 78% na receita total;
- O valor médio por transação do site é R\$ 71,19, e o valor médio por transação do blog é de 508,05;
- A taxa de conversão de vendas do blog é de 37%, enquanto a do site é de 19%;
- O número de transações do blog estava crescendo juntamente com o número de usuários, porém, a partir do mês 07 houve uma queda no número de transações;
- Após a chegada do blog o número de transações no site caiu 7% e o número de usuários cresceu 13%