

“

Avaliação
Analista de
Dados

Matheus Mendes Araujo

Preparação do Estudo

- Entendimento do que é esperado na avaliação
- Entendimento do modelo de negócio Thinkmart
- Perguntas de negócio a serem respondida
 - Identificar e criar hipóteses sobre os pontos chaves
 - Avaliar a estrutura e informações da base de dados
 - Gerar perguntas para orientar a exploração dos dados
- Revisar o que foi explorado, entendendo os resultados
 - Verificar possíveis caminhos que gerem novos insights
- Conclusões relevantes e próximos passos

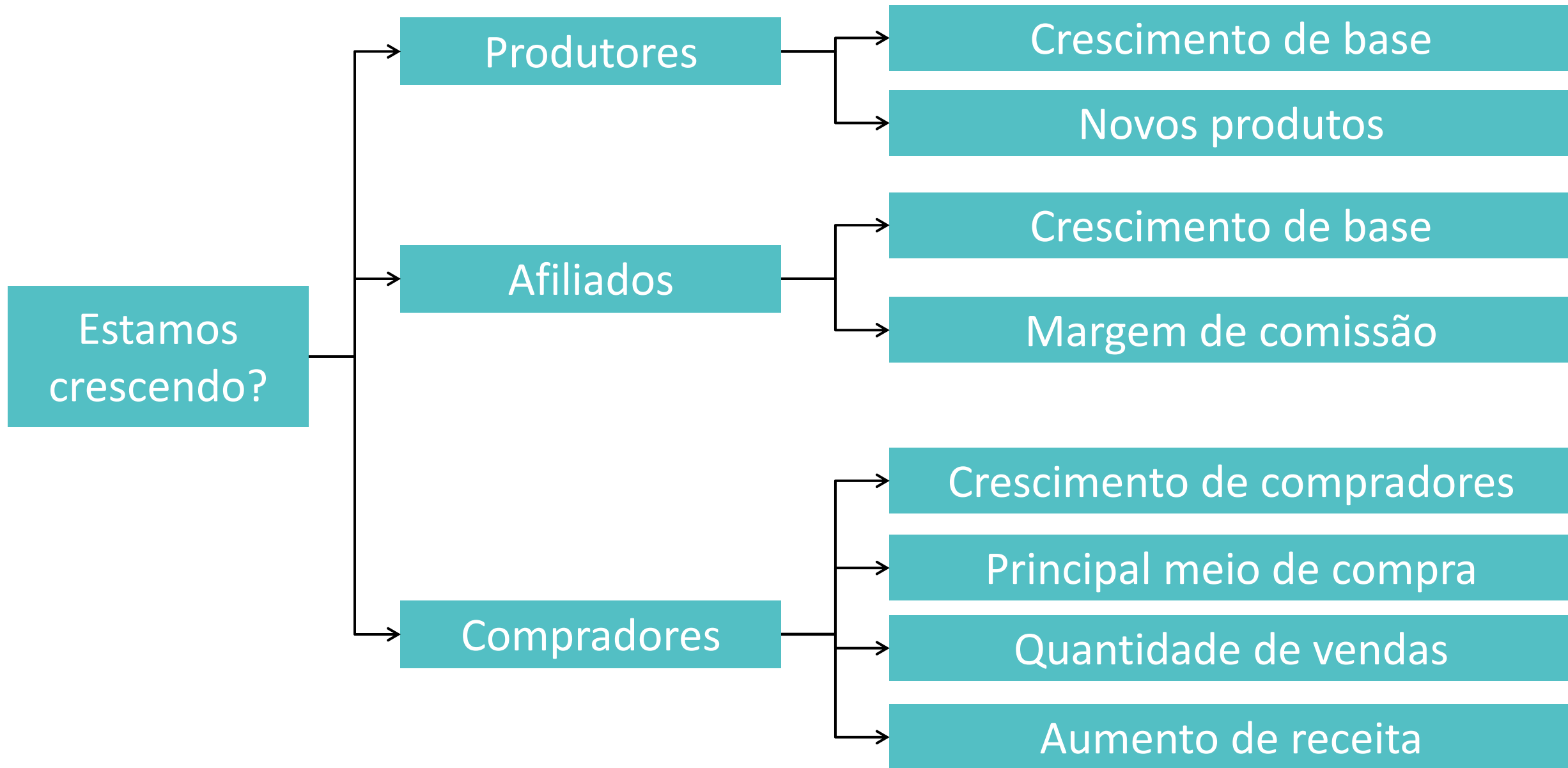
Problemas a serem explorados

1. Como você classifica a saúde do negócio Thinkmart a partir dos dados extraídos?
2. Existe algum padrão ou tendência relevante nos dados?
3. Baseado nos dados extraídos, o time de marketing precisa saber como aumentar a eficiência de suas campanhas. O que você recomendaria?

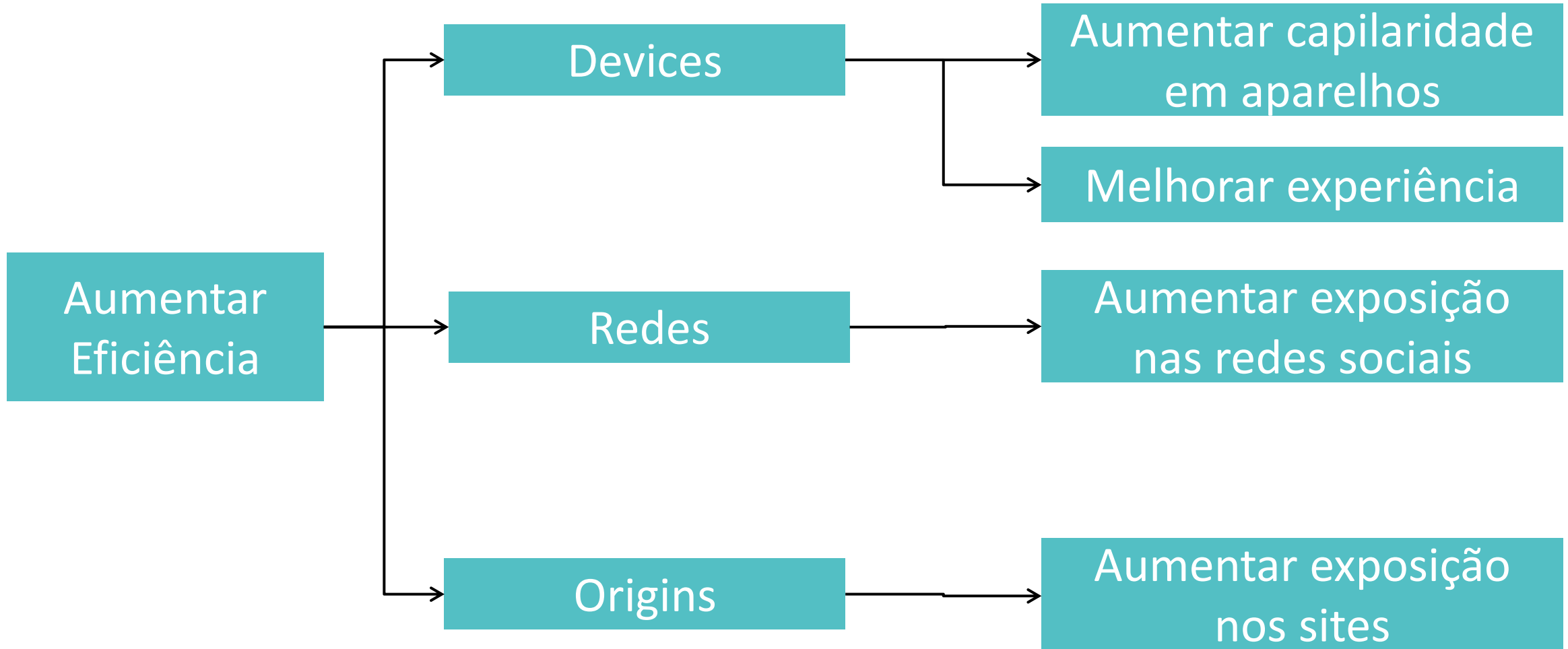
A partir das perguntas mencionadas, foi estruturado um processo de análise utilizando indicadores e hipóteses. Esse processo foi criado com o objetivo de responder às perguntas e aprofundar no negócio.

Para cada pergunta, foram definidos indicadores chave que poderiam ser utilizados para mensurar o desempenho e buscar respostas.

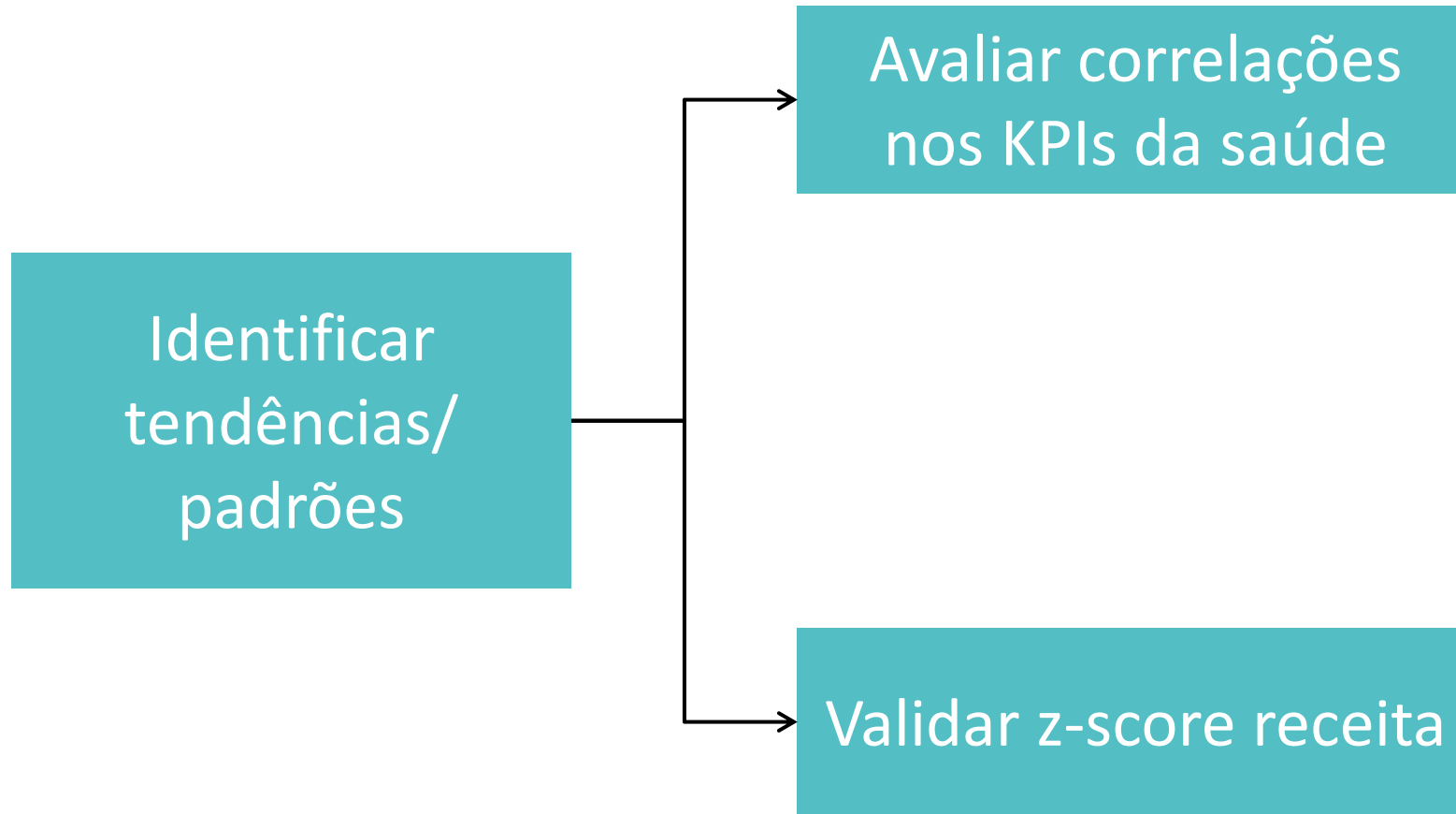
Nosso negócio está saudável?



Como aumentar a eficiência das campanhas?



Quais os padrões e tendências dos dados?



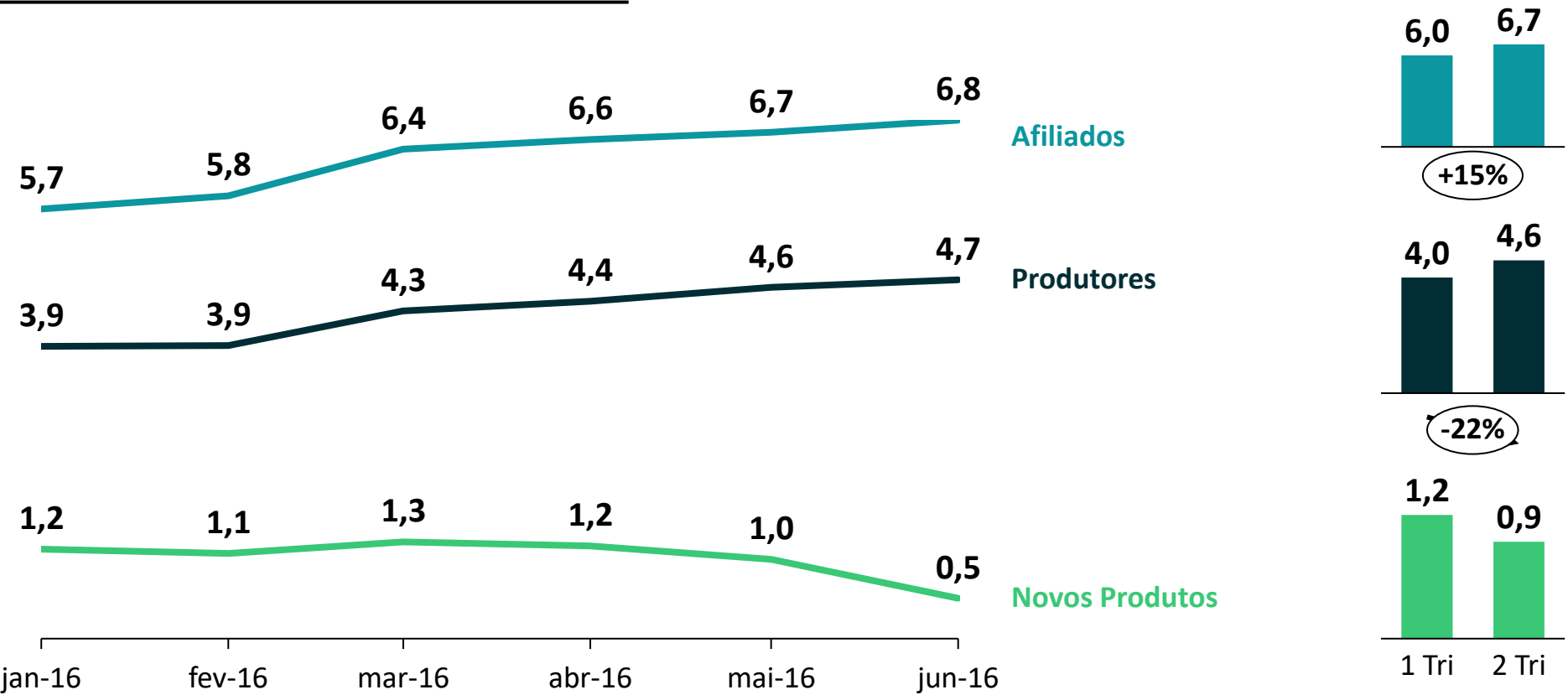
Perguntas iniciais, ponta pé a partir dos modelos

Após validar os campos da base, temos as seguintes perguntas iniciais:

1. Qual foi o crescimento de vendas, base de afiliados e produtores em 2016?
2. Quantos produtos foram lançados em 2016?
3. Qual foi o produto mais vendido em 2016?
4. Qual foi a categoria de produto com maior volume de vendas em 2016?
5. Qual foi o nicho de mercado com maior volume de vendas em 2016?
6. Quais foram os afiliados que geraram mais vendas em 2016?
7. Qual foi a origem do tráfego que gerou mais vendas em 2016?

Crescimento médio de afiliados e produtores de 15% e 12%, respectivamente, analisando os dois trimestres. Queda de 22% na criação de novos produtos no trimestre

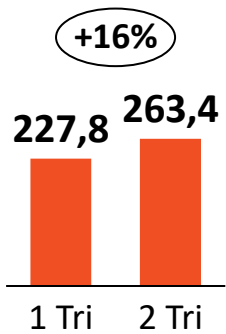
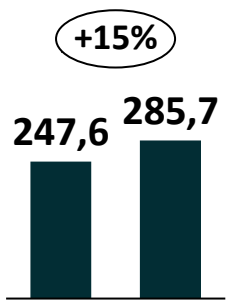
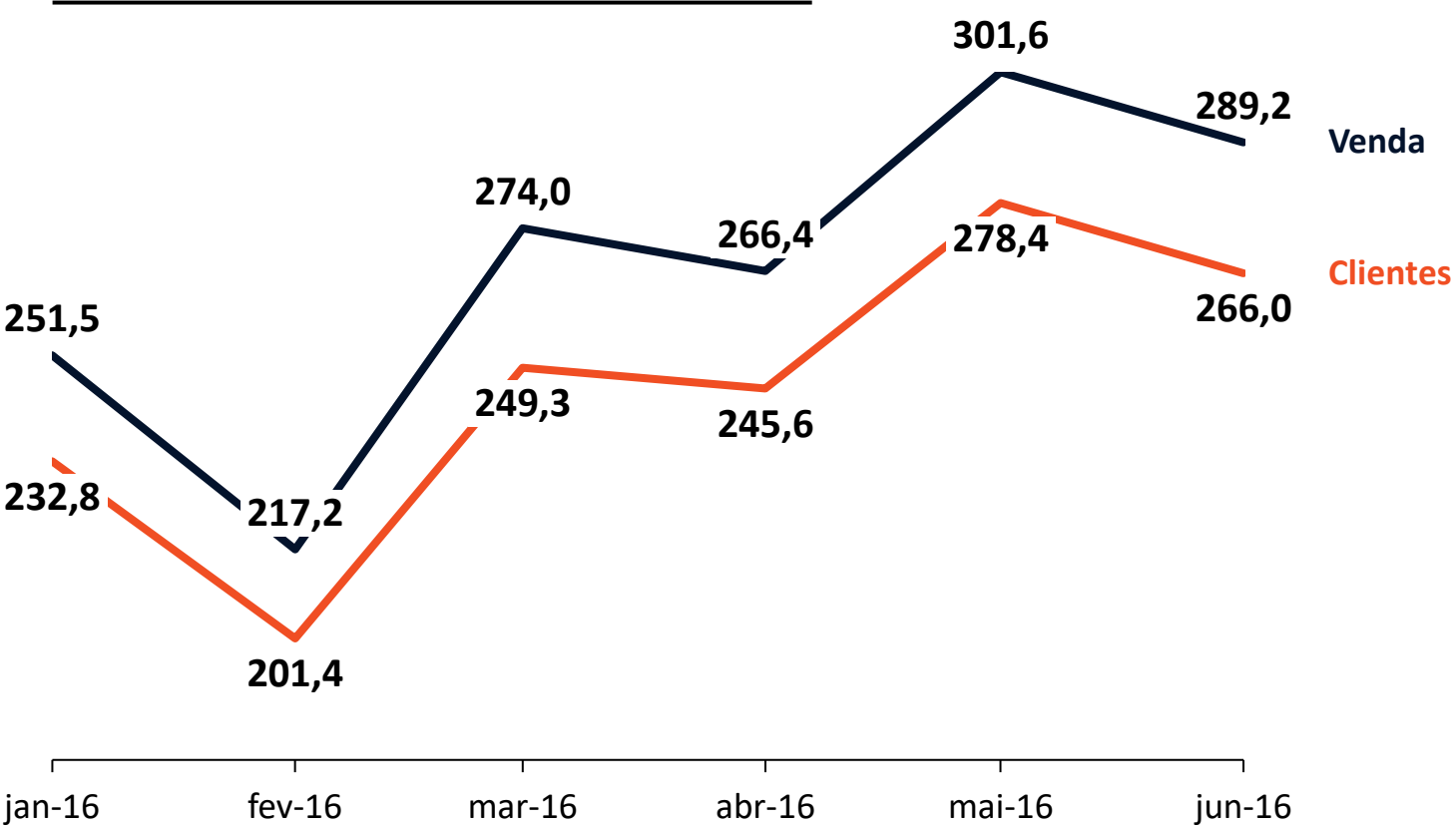
Crescimento de Base (em k):



Obs.: O valor do trimestre é gerado a partir da média dos meses para reduzir o efeito de sazonalidade.

Crescimento médio de vendas e clientes de 15% e 16%, respectivamente, comparando a média dos trimestres

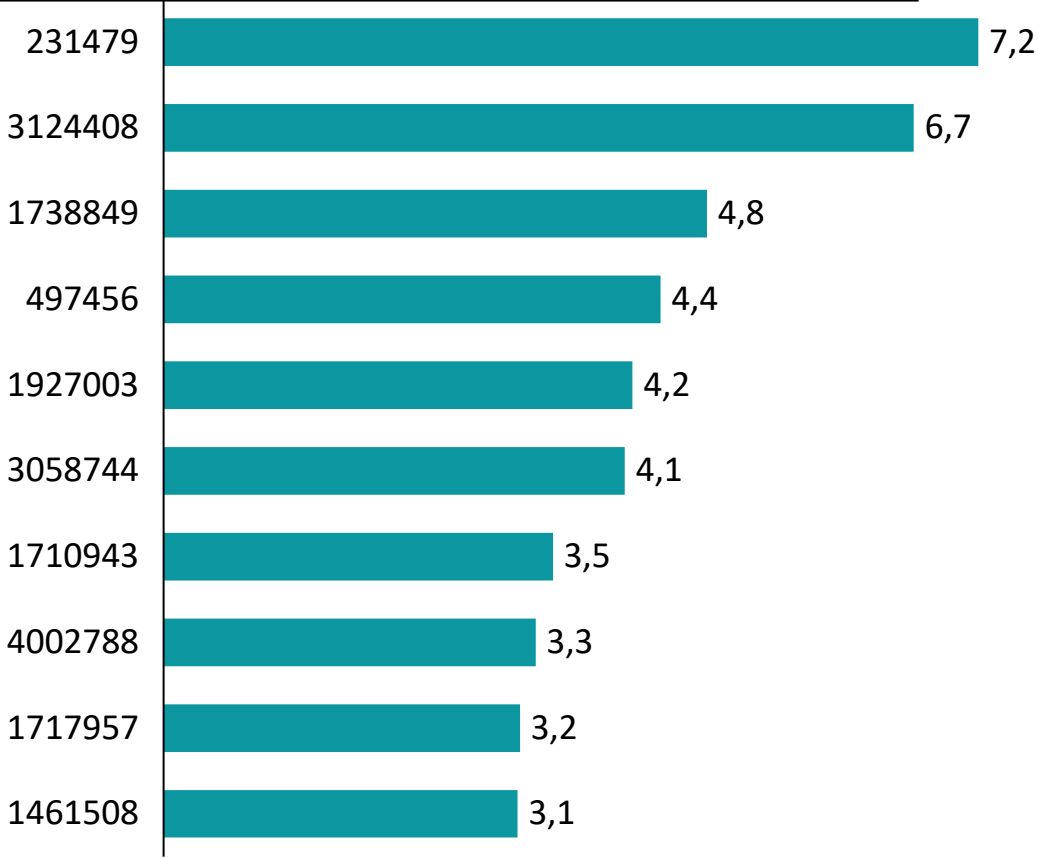
Crescimento de Base (em k):



Obs.: O valor do trimestre é gerado a partir da média dos meses para reduzir o efeito de sazonalidade.

Identificamos os melhores afiliados do semestre. Objetivo: aprofundar na frequência e tipo de produtos vendidos

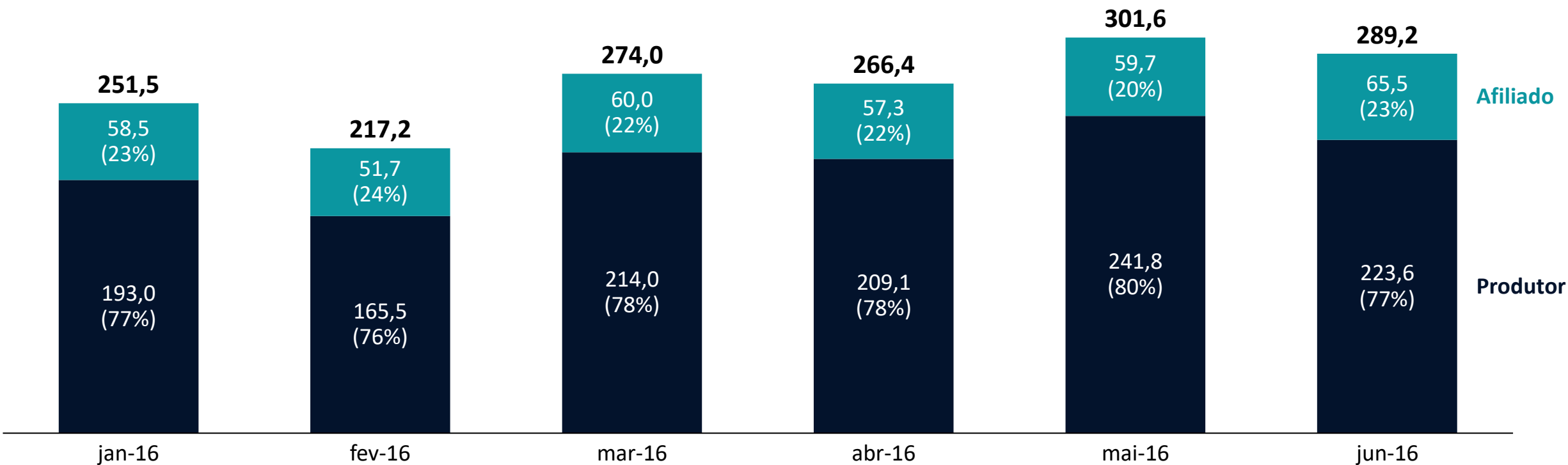
Vendas dos Top Afiliados no semestre (em k):



As vendas feitas por afiliados representam em média 22% do total no mês.

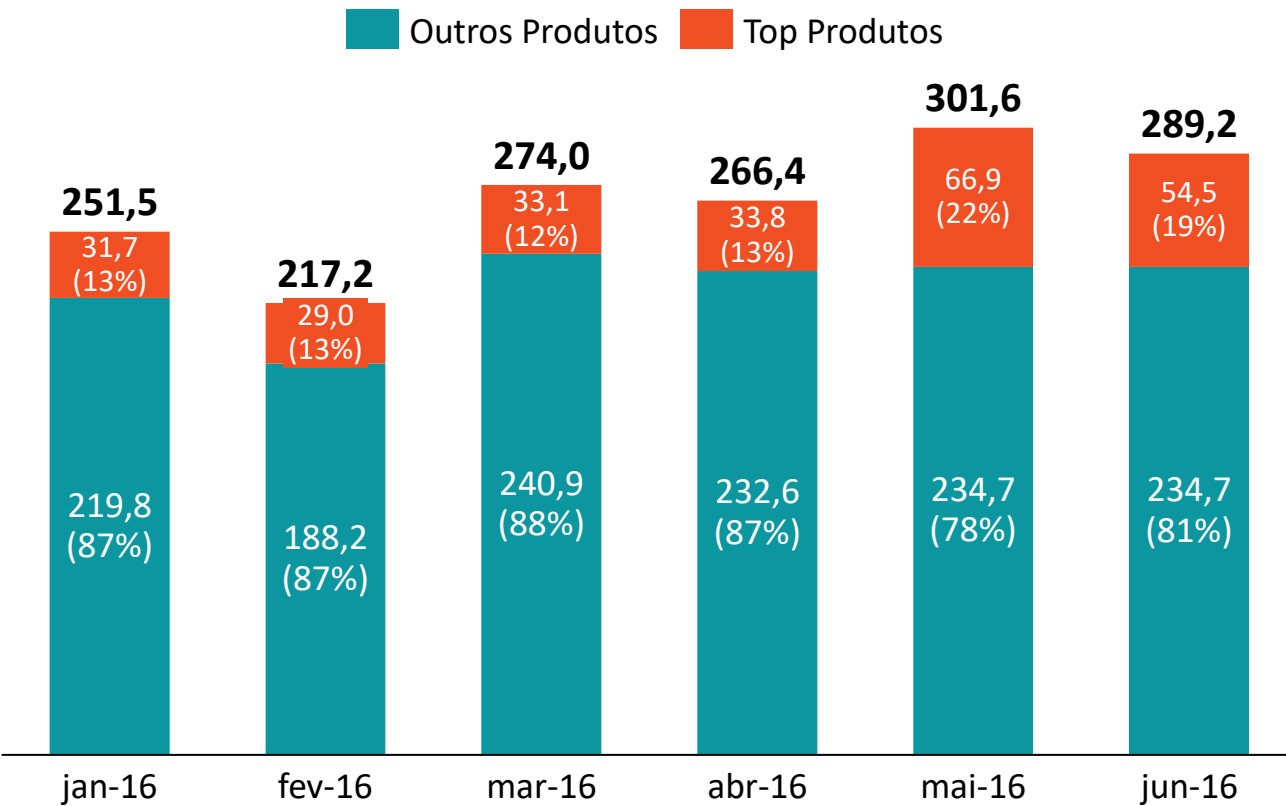
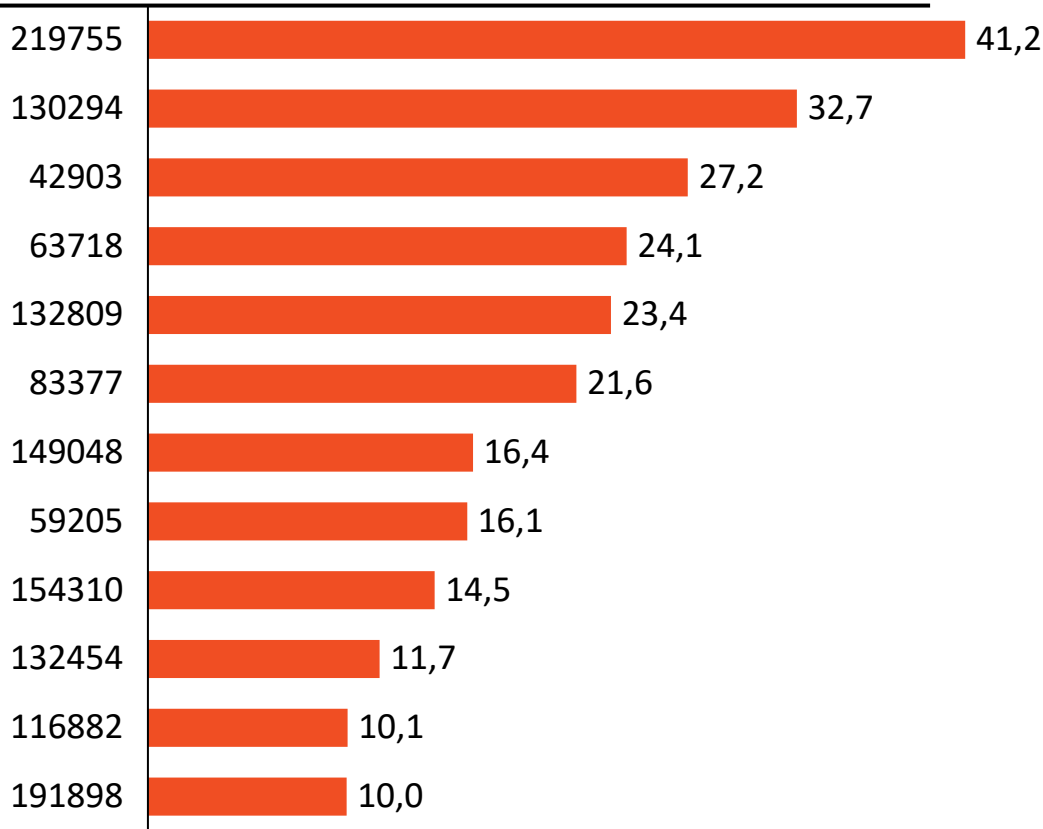
Objetivo: aprofundar quais são os produtos mais vendidos por afiliados

Principal vendedor (em k):



Tops produtos representam em média 15% das vendas, crescendo a sua participação no 2º Trimestre. Uma boa diversificação na carteira de produtos

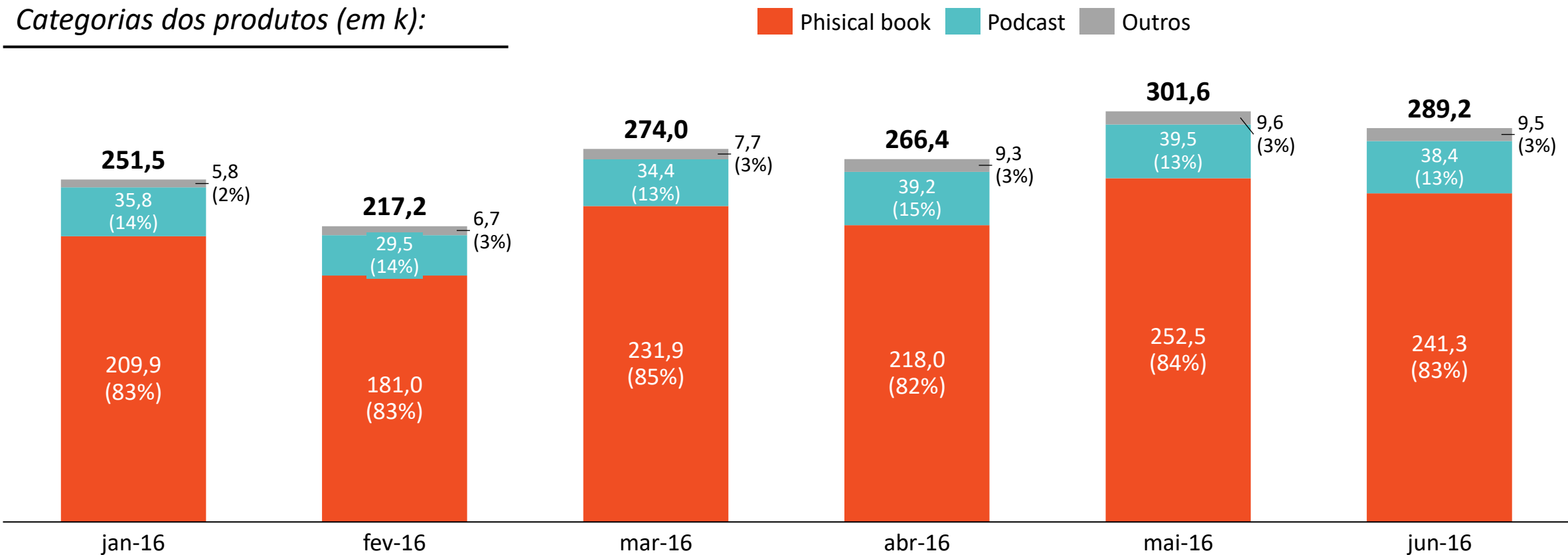
Vendas dos Top Produtos no semestre (em k):



Obs.: Outros – Corresponde aos demais produtos vendidos na plataforma.

97% das nossas vendas são das categorias Phisical book e Podcast, por ordem de grandeza

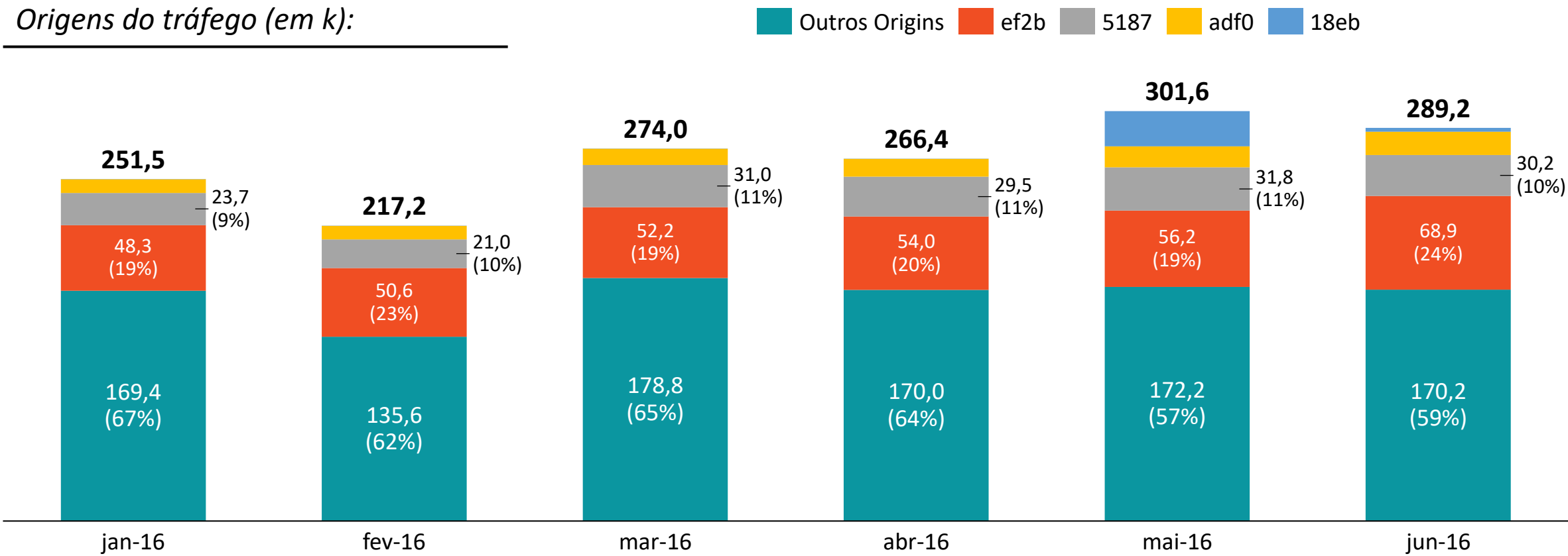
Categorias dos produtos (em k):



Obs.: Outros – Corresponde aos Workshop, eBook, Subscription, In-class course, App, eTicket, Webinar, Video.

Para aumentarmos a eficiência de marketing recomendamos explorarmos os sites: ef2b, 5187, adf0 e 18eb, que representam 37%, em média, das vendas

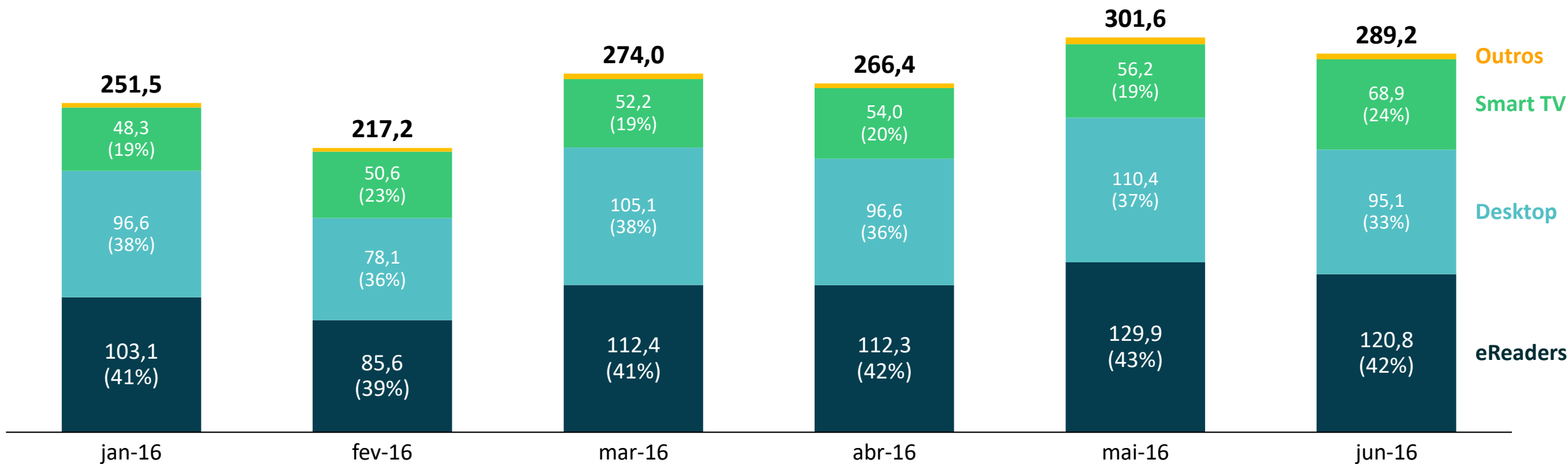
Origens do tráfego (em k):



Obs.: Outros – Corresponde aos demais sites.

eReaders, Desktop e Smar TV são os aparelhos mais utilizados na hora da compra, devemos focar na usabilidade e experiência neles

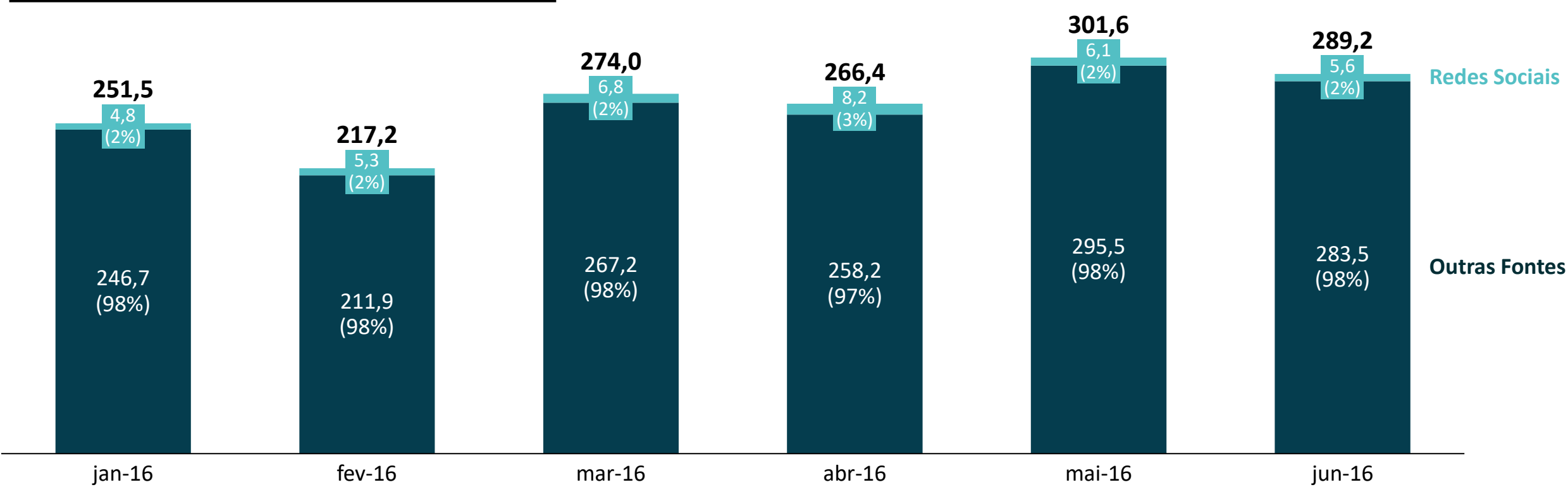
Principais aparelhos de consumo (em k):



Obs.: Outros – Corresponde a cellphone e tablet.

2% das nossas vendas são de sites de redes sociais. Objetivo: entender qual a fricção que temos nesse tipo de ambiente

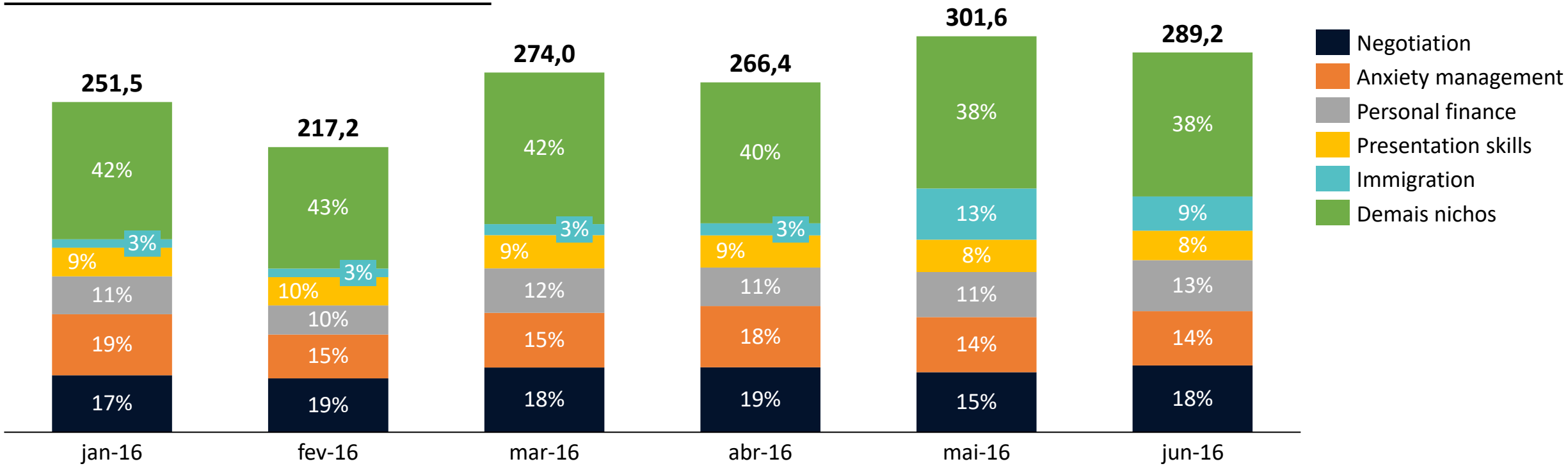
Vendas pelas redes sociais (em k):



Obs.: Outras Fontes – sites que não são redes sociais.

Negociação, gerenciamento de ansiedade e finanças pessoais são os principais nichos de vendas. O nicho de imigração ganhou participação nos últimos meses

Crescimento de Base (#N):

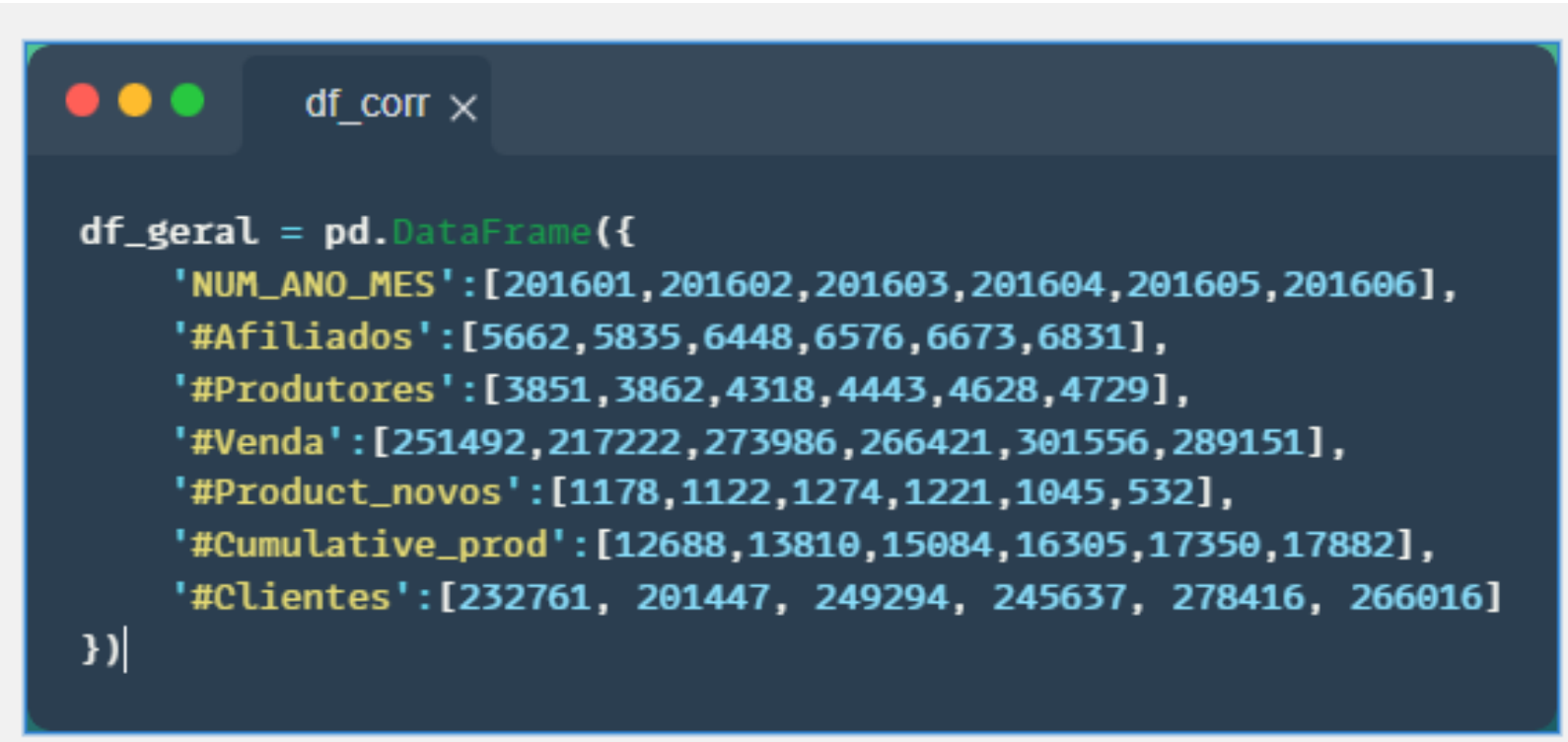


Aprofundamento da exploração

Após avaliar os primeiros resultados, seguimos para aprofundar em alguns tópicos:

1. Qual a correlação entre as seguintes variáveis:
 1. Afiliados/Produtores/Novos Produtos/Vendas/Clientes
2. Como manipular a receita com o z-score do valor real

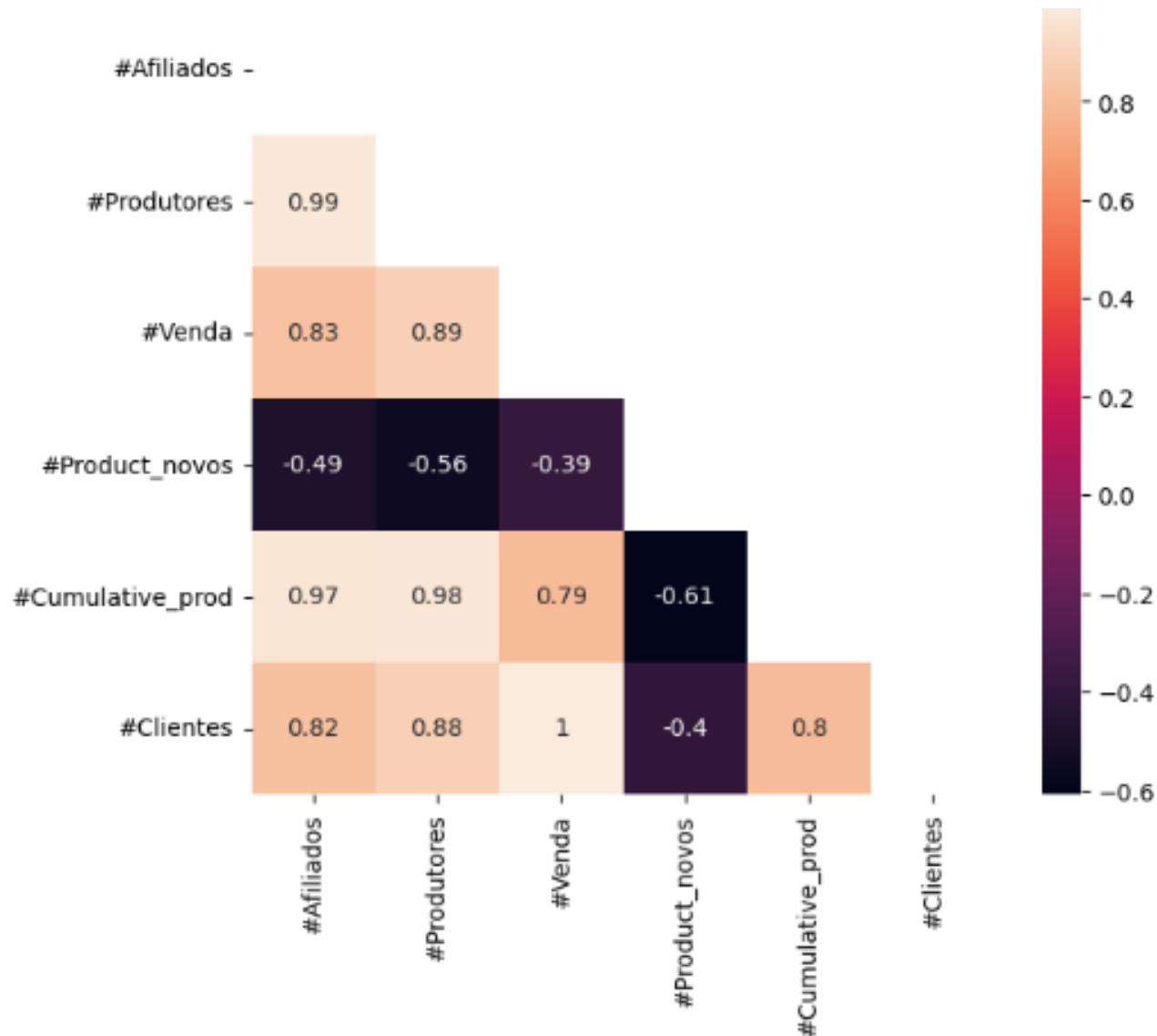
Para validar a correlação entre as variáveis, eu criei um dataframe agrupando-as na granularidade mensal



```
df_geral = pd.DataFrame({  
    'NUM_ANO_MES': [201601, 201602, 201603, 201604, 201605, 201606],  
    '#Afiliados': [5662, 5835, 6448, 6576, 6673, 6831],  
    '#Produtores': [3851, 3862, 4318, 4443, 4628, 4729],  
    '#Venda': [251492, 217222, 273986, 266421, 301556, 289151],  
    '#Product_novos': [1178, 1122, 1274, 1221, 1045, 532],  
    '#Cumulative_prod': [12688, 13810, 15084, 16305, 17350, 17882],  
    '#Clientes': [232761, 201447, 249294, 245637, 278416, 266016]  
})
```

Obs.: Cumulative_prod – Corresponde a quantidade total de produtos acumulados, feito a partir da data de criação.

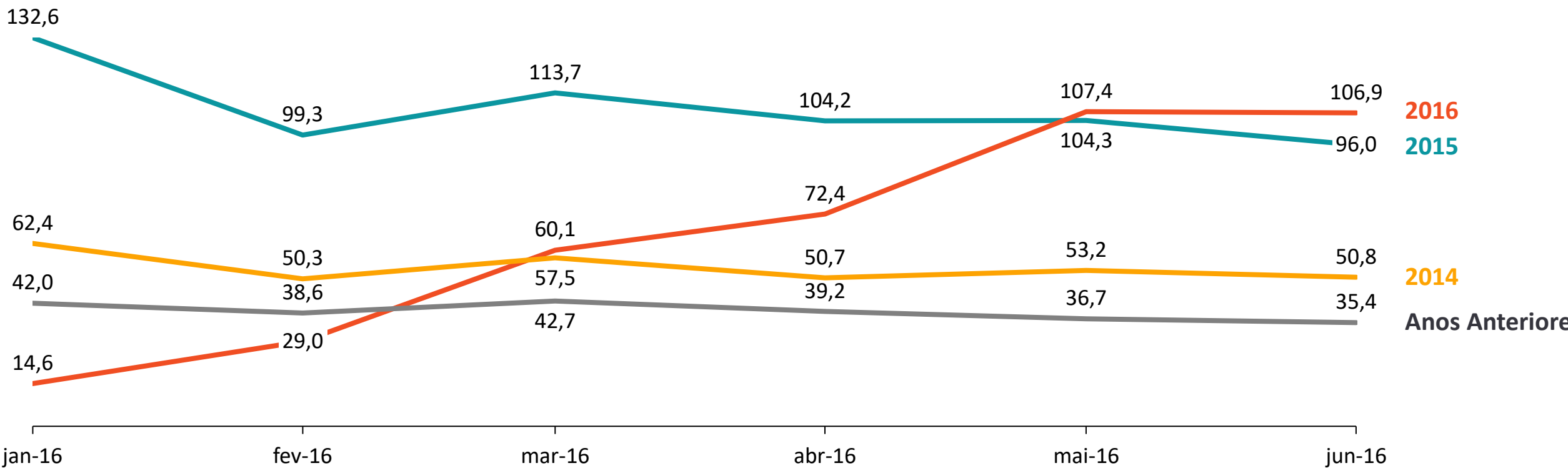
Avaliando a matriz de correlação:



- Cliente influencia diretamente nas vendas e na quantidade total de produtos
- Produtos novos apresentou uma correlação negativa com a quantidade de vendas, dessa forma, irei aprofundar nas vendas feitas pela data de criação do produto
- Produtores e afiliados tem uma relação direta entre si
- O crescimento de produtores influencia diretamente na quantidade total de produtos

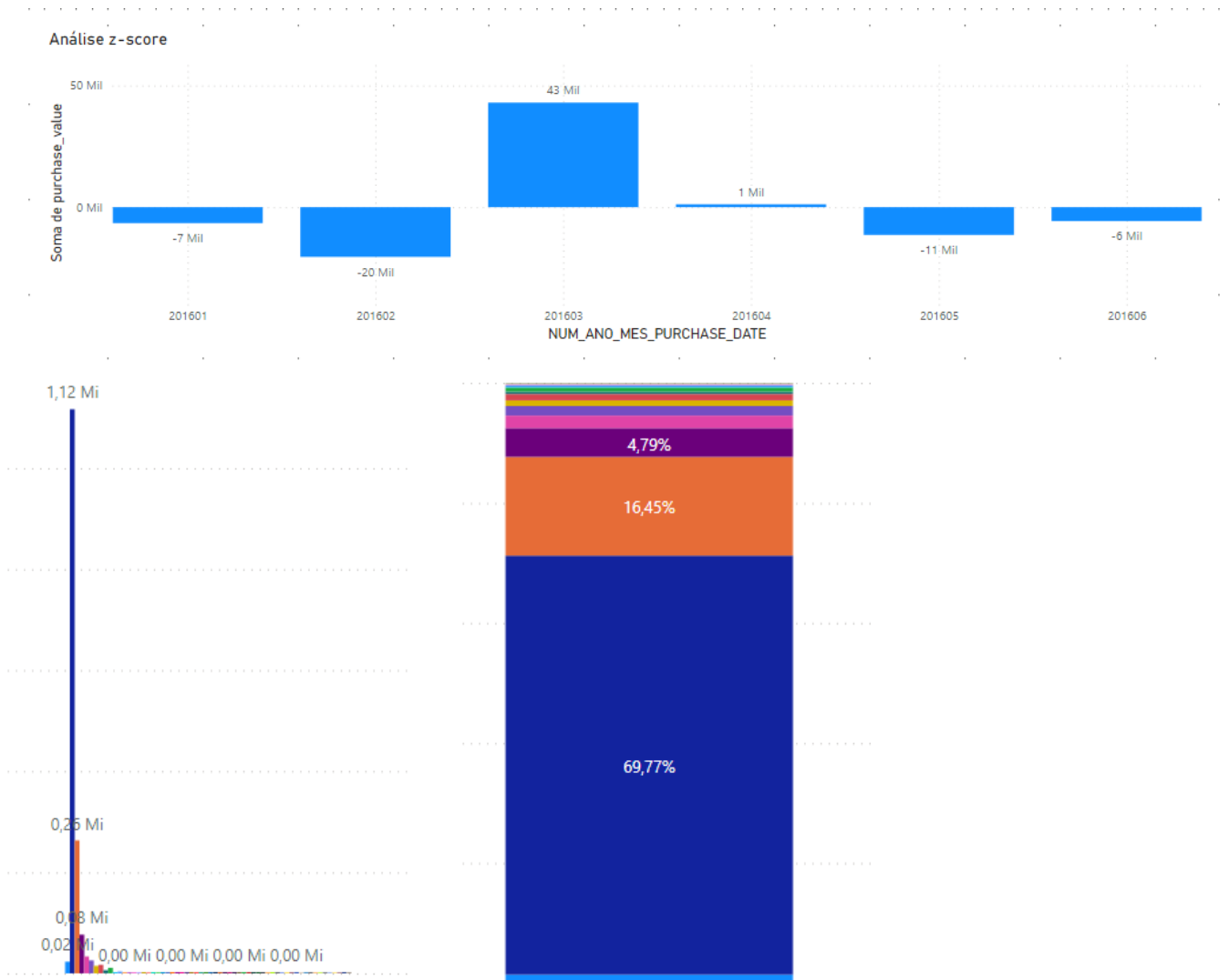
Produtos criados em 2016 se tornaram a principal fonte de venda a partir de maio do mesmo ano

Vendas pelo ano criação do produto (em k):



Obs.: Anos anteriores é composto pelos anos de 2008, 2009, 2010, 2011, 2012 e 2013.

A variável purchase_value(receita) não foi possível avaliar, devido a codificação em z-score



- Z-score ao ser somado não gera uma informação interpretável
- Busquei realizar avaliar a distribuição dos valores de z-score
- Cerca de 70% dos valores de **purchase_value** se encontram no seguinte intervalo $[-0,534;0[$
- Não houve mais considerações que gerasse algum insight

Conclusões

Após o aprofundamento nos dados:

1. Saúde da Thinkmart

1. Estamos crescendo a nossa base em 3 perspectivas: clientes, produtores e afiliados, com uma diversidade de novos produtos, a qual devemos acompanhar a evolução nos próximos meses

2. Para o time de marketing:

1. Recomendo aprofundarmos investimentos nos sites: ef2b, 5187, adf0 e 18eb, focando nos aparelhos eReaders, Desktop e Smar TV
2. Realizar testes A/B para aumentarmos nossa conversão em redes sociais e sites de baixa performance

**Ficou com alguma
dúvida?**