

# **Case – Data Solutions Analyst**

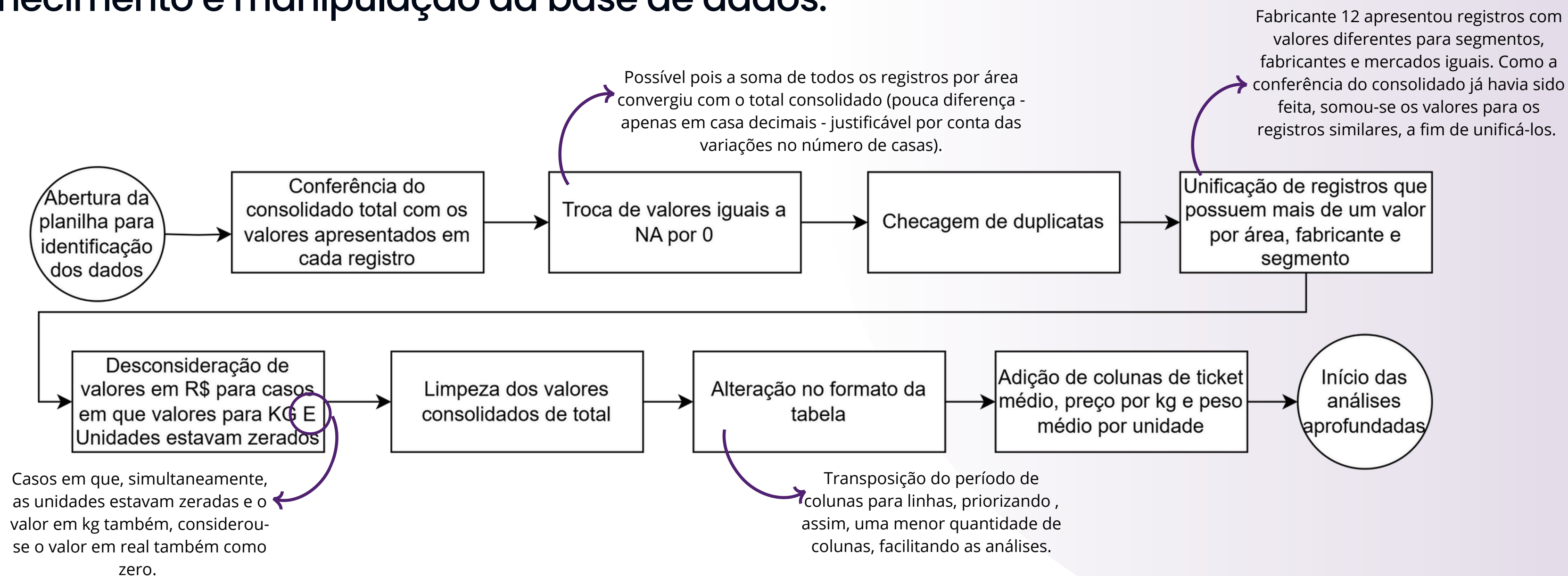


# Enunciado

- Arquivo contendo dados da categoria de Chocolates por um período de dois anos (2015 e 2016)
- Dados abertos por fabricante, segmento e mercado (Total Brasil e 7 áreas)
- 3 variáveis: Vendas em KG, Vendas em R\$ e Vendas em Unidades

- Análise descritiva, associada a construção de um storytelling sobre a categoria de chocolates
- Insights e recomendações de negócio
- Entrega em formato de PowerPoint

## Reconhecimento e manipulação da base de dados:



# Análise dos dados – Visão Geral

Uma representação da base de dados após a manipulação dos dados pode ser vista a seguir:

Mercado	Fabricante	Segmento	Data	KG	R\$	Unidades	Ticket Médio	Preço por KG	Peso Médio por KG
AREA I - INA	FABRICANTE 1	BOMBOM	2015-01-01	6,9	193,5	61,8	3,13	28,04	0,11
AREA I - INA	FABRICANTE 1	BOMBOM	2015-01-01	5,2	149,1	78,3	1,9	28,67	0,06
AREA I - INA	FABRICANTE 1	BOMBOM	2015-01-01	8,1	199,3	50,8	3,92	24,6	0,15
AREA I - INA	FABRICANTE 1	BOMBOM	2015-01-01	10,2	249,3	61,1	4,08	24,44	0,16
AREA I - INA	FABRICANTE 1	BOMBOM	2015-01-01	4,8	130,5	36,4	3,58	27,18	0,13

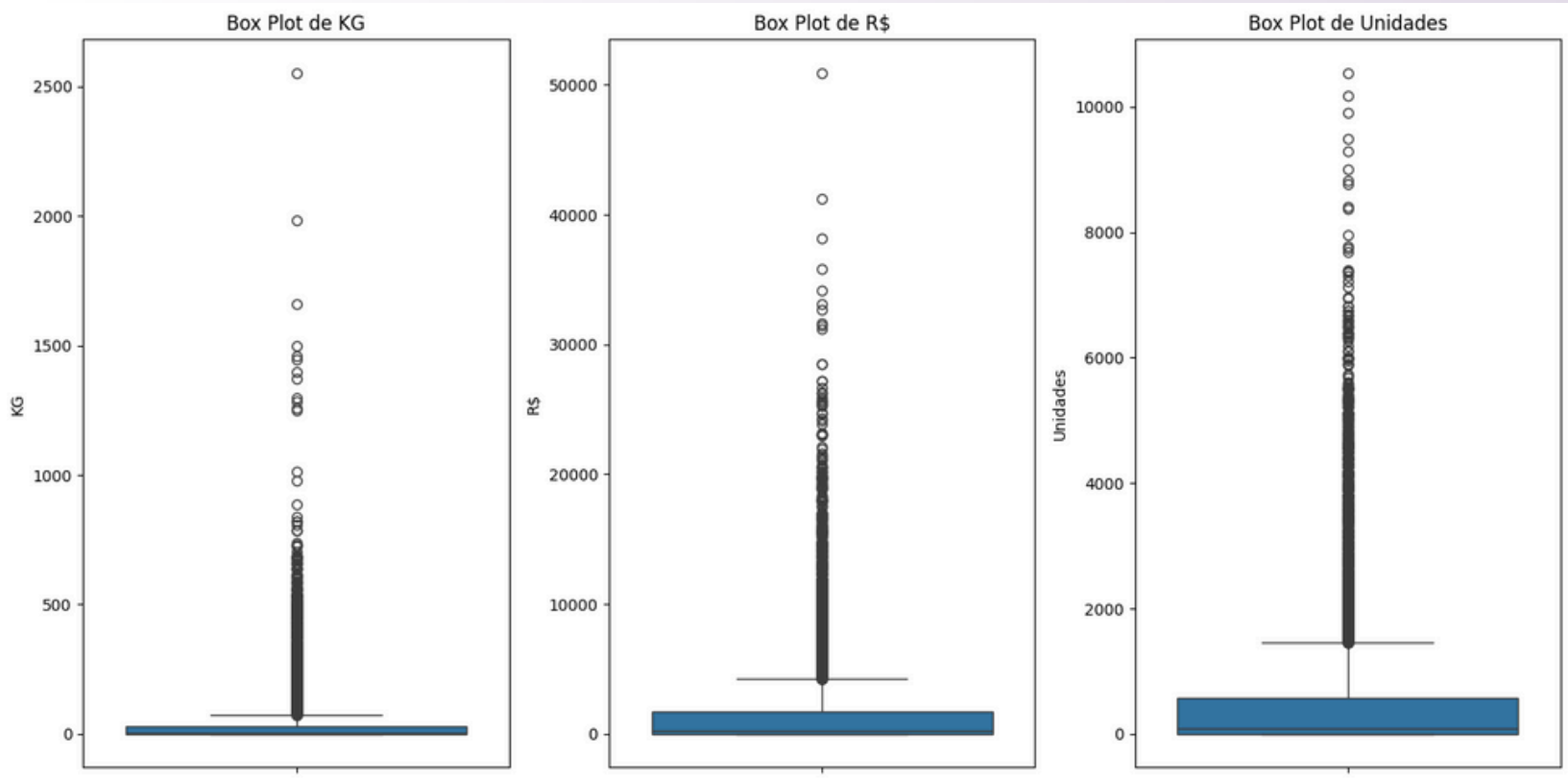
Dando início a análise, o primeiro passo foi ter uma visão de dados estatísticos gerais, como, média, valores mínimos e máximos, divisão entre quartis, para os valores em KG, R\$ e Unidades. Esse resultado pode ser visto a seguir:

	KG	R\$	Unidades
Soma	372916,2	14050800,9	4715621,7
Média	43,89	1653,81	555,04
Desvio Padrão	112,46	3477,86	1083,3
Mínimo	0	0	0
25,00%	0,3	18,2	3,3
50,00%	4,1	223,5	83
75,00%	30,42	1707,97	583,67
Máximo	2553,6	50937,5	10545,1

O quartil médio (mediana) se diferencia bastante da média para as três dimensões. Isso é uma evidência de que existem outliers que destoam bastante dos demais valores. Isso fica ainda mais visível ao analisar os valores máximos para KG, R\$ e Unidades que são bastante elevados, deixando clara uma assimetria entre os dados.

Ainda, o valor máximo em KG e R\$ ocorreu para o segmento de bombom, na área VI, realizado pelo fabricante 12 no período de Abril de 2015, enquanto o valor máximo de unidades foi referente a uma venda de formatos variados no mercado da Área I, em que o fabricante era o número 5, no período de Fevereiro de 2015.

Aprofundando ainda mais no comportamento de dispersão dos dados, plotou-se um gráfico de diagrama de caixa (box plot) para KG, R\$ e unidades:

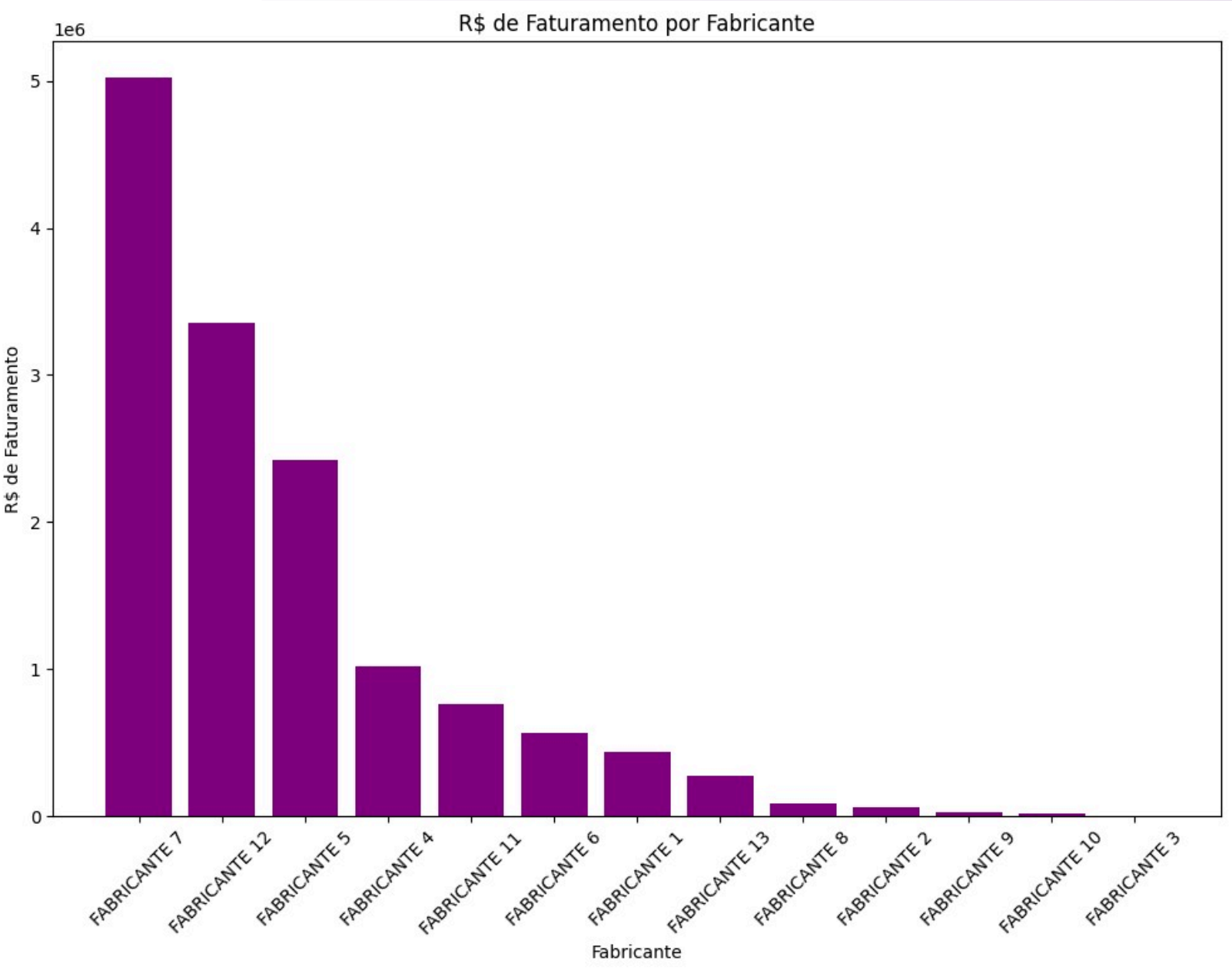
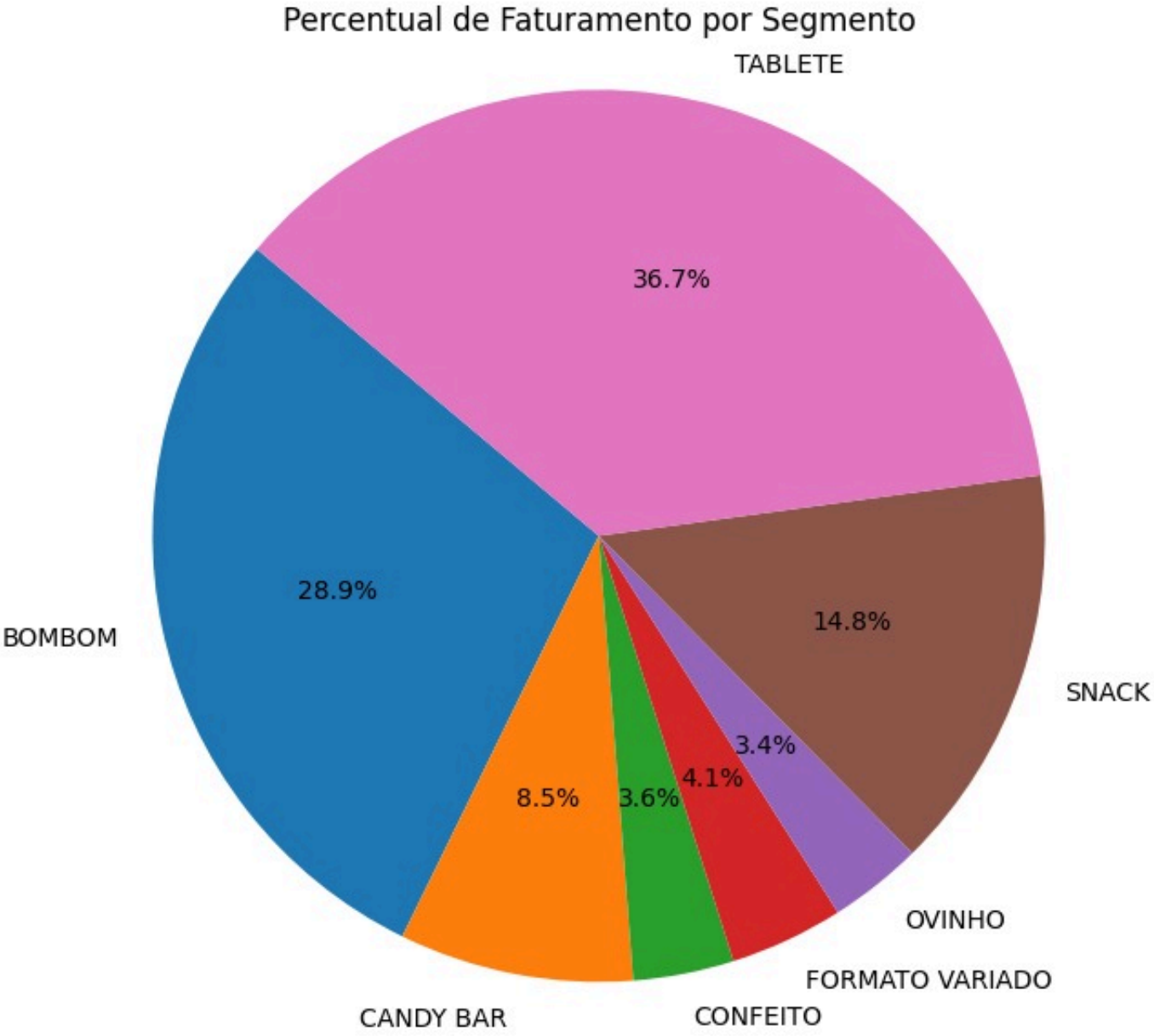


É possível identificar que os valores para unidades vendidas são menos dispersos do que aqueles para KG e R\$. Para estes últimos, fica bastante evidente a questão dos outliers comentada anteriormente, e o grande destaque vai para os valores máximos, também citados acima.

# Análise dos dados – Visão Geral

Para iniciar, fez-se uma análise holística, comparando os dois períodos, em termos de faturamento. Pode-se perceber que 2015 teve um faturamento menor do que 2016, sendo que os valores calculados foram, respectivamente de: R\$ 6.916.129,9 e R\$ 7.134.671.

Visando, então, compreender melhor isto, analisou-se o faturamento por segmento e por fabricante, obtendo-se os seguintes gráficos:

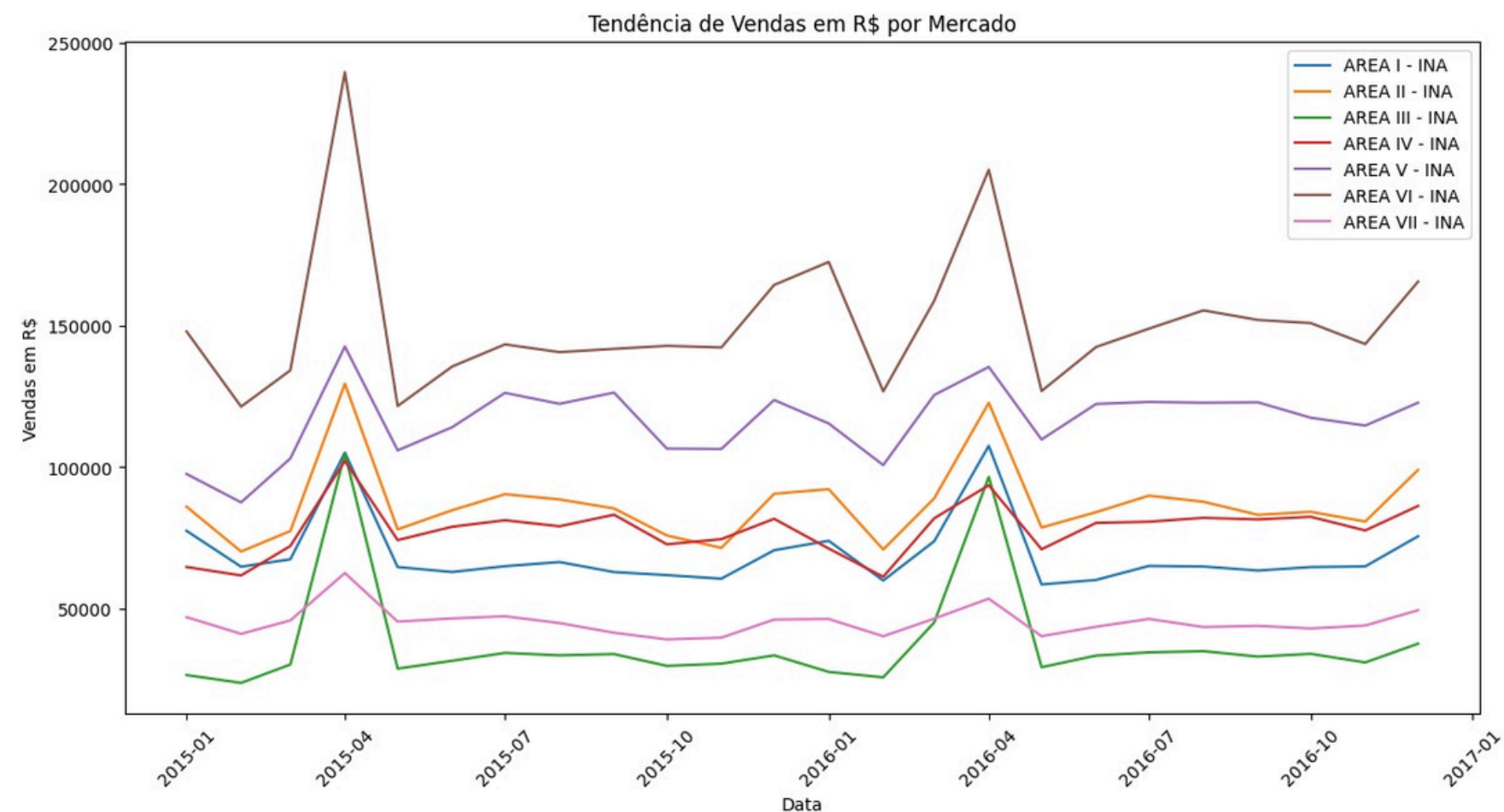


Os principais insights gerados a partir da análise são que os segmentos de tablete, bombom e snacks são os mais representativos, e em relação aos fabricantes, o 7, 12 e o 5 se destacam como os principais, em termos de faturamento. Esses pontos voltarão a ser explorados mais adiante.



# Análise dos dados

Para aprofundar as análises, buscou-se entender o comportamento do mercado em mais detalhes. Para isso, plotou-se um gráfico de linha, utilizando as informações de faturamento mensal e o período, para cada um dos mercados. O resultado pode ser observado a seguir:



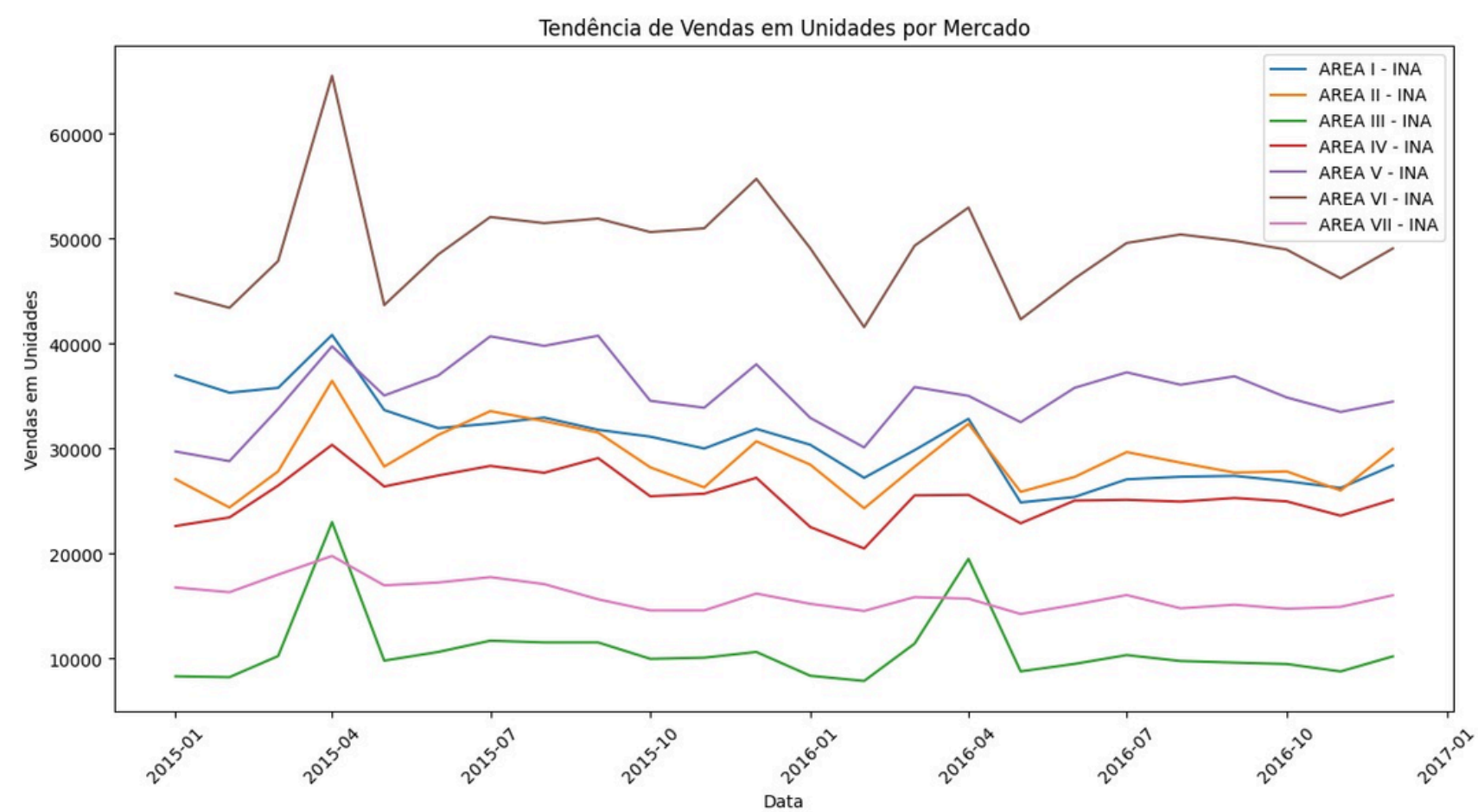
Observou-se, com base neste gráfico, a sazonalidade: é possível notar que os maiores picos ocorrem no mês de Abril, muito provavelmente relacionado a Páscoa; existe um considerável aumento de faturamento nos meses de dezembro, e que Fevereiro é o mês em que se teve o menor valor de vendas em reais. Contudo, é importante ressaltar que este é o mês com menor quantidade de dias, quase 10% a menos do que os demais, fazendo com que, mesmo que a quantidade de vendas por dia fosse igual, nesse mês ainda assim seria observado um decaimento.

Algo que chama bastante a atenção é o destaque da área VI em relação aos demais mercados. Buscou-se, então, avaliar hipóteses para as causas dessa relevância:

- Maior quantidade de unidades vendidas;
- Preço por KG pode ser reduzido neste mercado;
- Variedade de segmentos;
- Quantidade de fornecedores;

Analisou-se, então, o gráfico de unidades vendidas por período, para cada um dos mercados, visando averiguar a primeira hipótese:

# Análise dos dados



Neste gráfico, observa-se que a área VI ainda se destaca com a maior quantidade de vendas em todos os períodos. Entretanto, percebe-se que Dezembro de 2015 teve uma quantidade de vendas maior do que Abril de 2016, mesmo que este último tenha apresentado maior faturamento. Isto é um indicativo de que em Abril são comprados itens com maior peso/unidade ou o preço da categoria chocolates é mais elevado.

Para aprofundar neste ponto, e compreender também a segunda hipótese, do preço ser mais baixo no mercado VI analisou-se, então, o preço médio por KG de cada mercado em diferentes períodos.

Data	Mercado	Preço médio por KG em R\$
2016-10-01	AREA III - INA	48,5
2016-10-01	AREA IV – INA	47,8
2016-09-01	AREA IV – INA	47,5
2016-11-01	AREA III - INA	47,5
2016-11-01	AREA II - INA	47,3
2016-10-01	AREA II - INA	47,2
2016-08-01	AREA IV – INA	46,9
2016-09-01	AREA III - INA	46,8
2016-09-01	AREA II - INA	46,8
2016-08-01	AREA III - INA	46,8

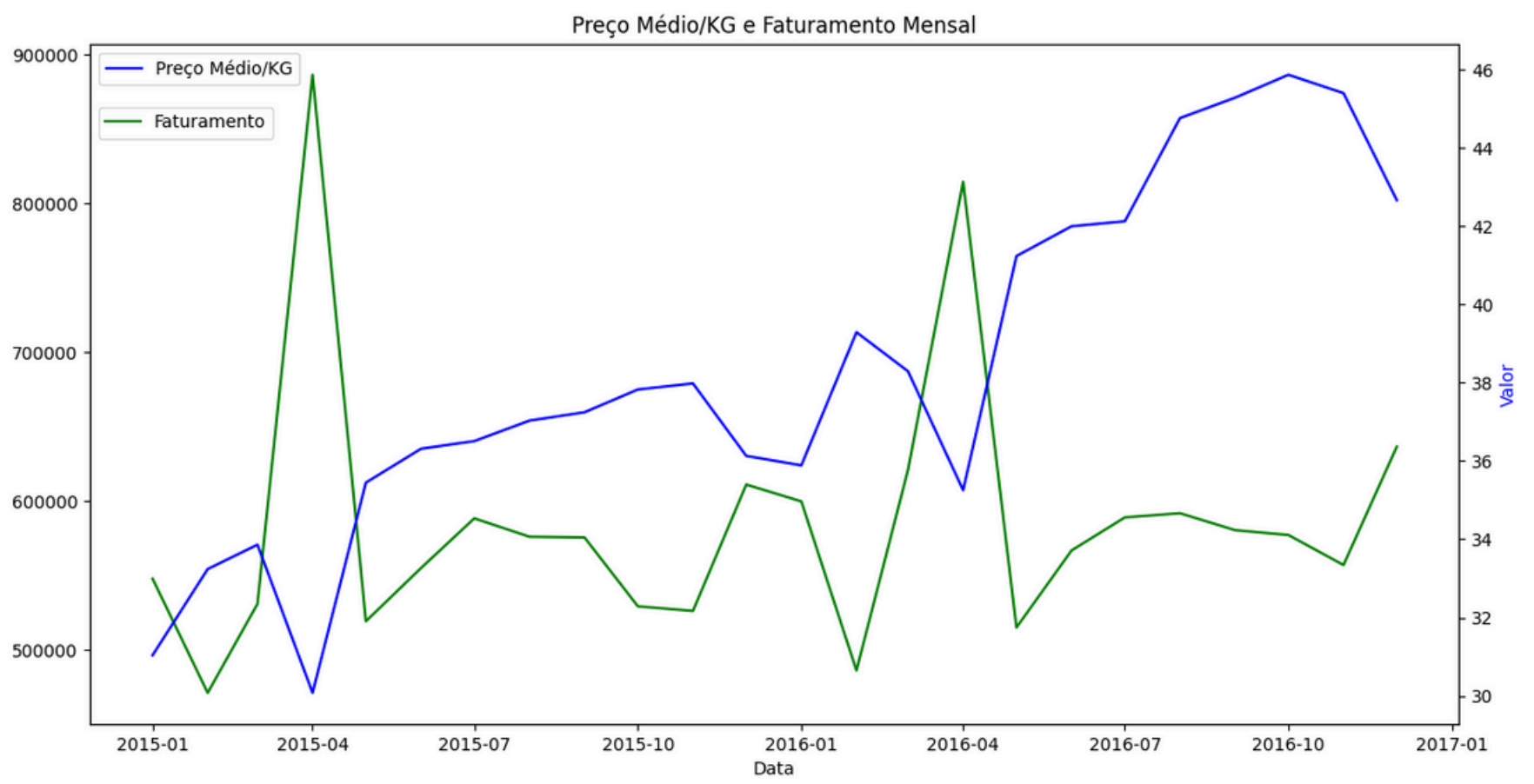
Com esses dados, é possível perceber que o preço médio por KG em 2016 é superior ao de 2015, sendo essa uma das possíveis justificativas para a diferença de faturamento entre os anos. Apesar disso, é importante dizer que esse comportamento é esperado, uma vez que fatores como a inflação, impactam diretamente no preço dos produtos, visto que a Selic passou por grandes aumentos na metade de 2015.

Além disso, nota-se que as áreas II, III e IV se destacam com os maiores preços médios.

Comparando esses valores com o gráfico de unidades vendidas, é possível perceber uma correlação que indica que a quantidade de vendas decaiu enquanto o preço médio por kg aumentou, sendo essa, então, uma possível causa para este declínio nas vendas.

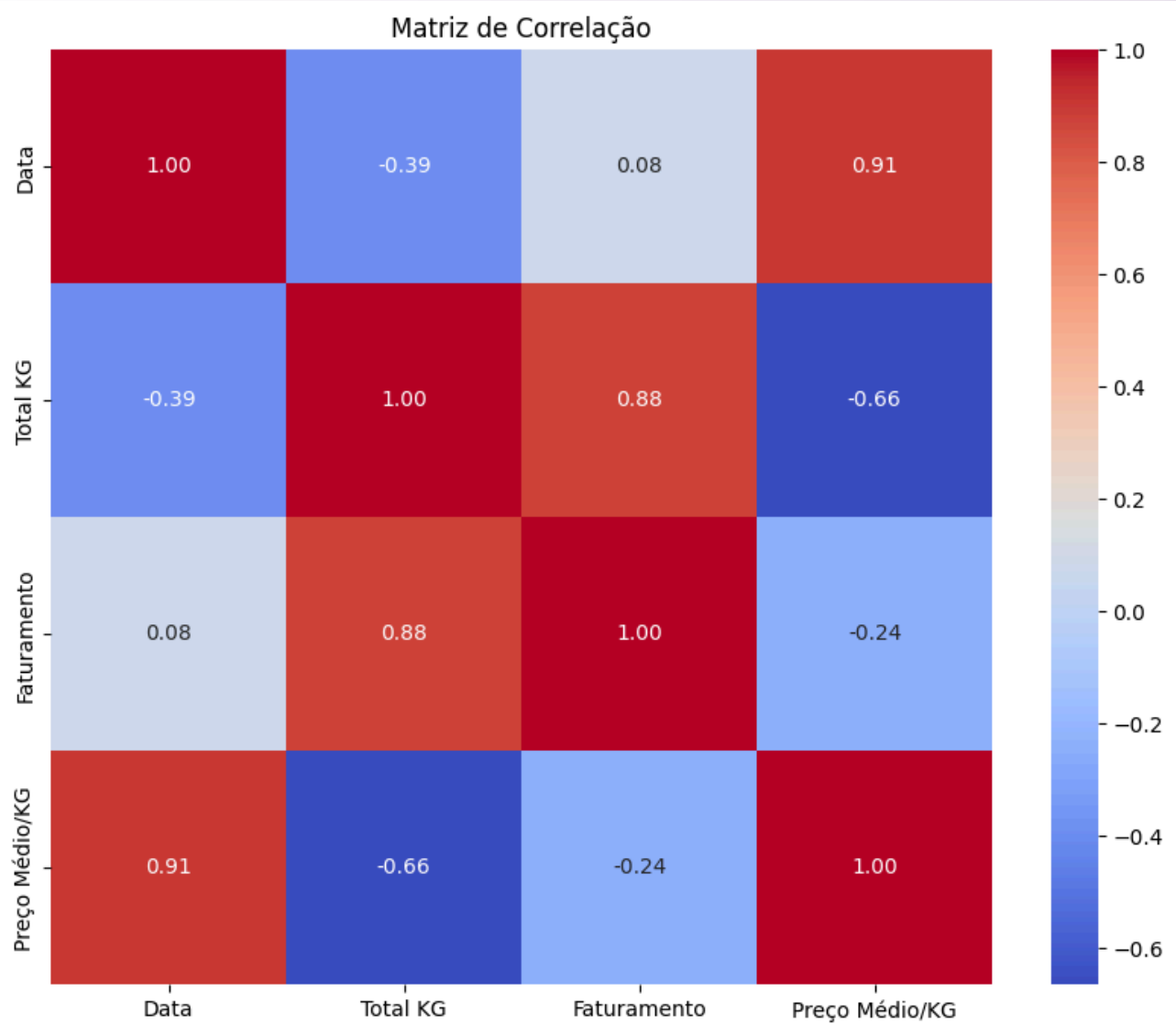
# Análise dos dados

Avaliando, então, esse ponto mais a fundo, plotou-se um gráfico de preço médio por kg vs. faturamento. O resultado obtido é exibido a seguir, onde o faturamento mensal é orientado pelo eixo à esquerda e o eixo do preço médio pode ser visto à direita:



Este gráfico comprova a correlação entre preço médio e faturamento. Os picos de faturamento ocorrem quando o preço médio apresenta um forte declive, e o inverso também ocorre, evidenciando, assim que a quantidade vendida está diretamente relacionada com o preço praticado.

Outras correlações podem ser aferidas pela matriz de correlação a seguir:



Algumas informações interessantes podem ser extraídas, como um forte relacionamento entre Preço Médio por KG e Data, o Total em KG com o faturamento e inversamente proporcional é o Preço médio por KG e o Total de KG, evidenciando que quanto mais caro, menos atrativas são as vendas.



# Análise dos dados

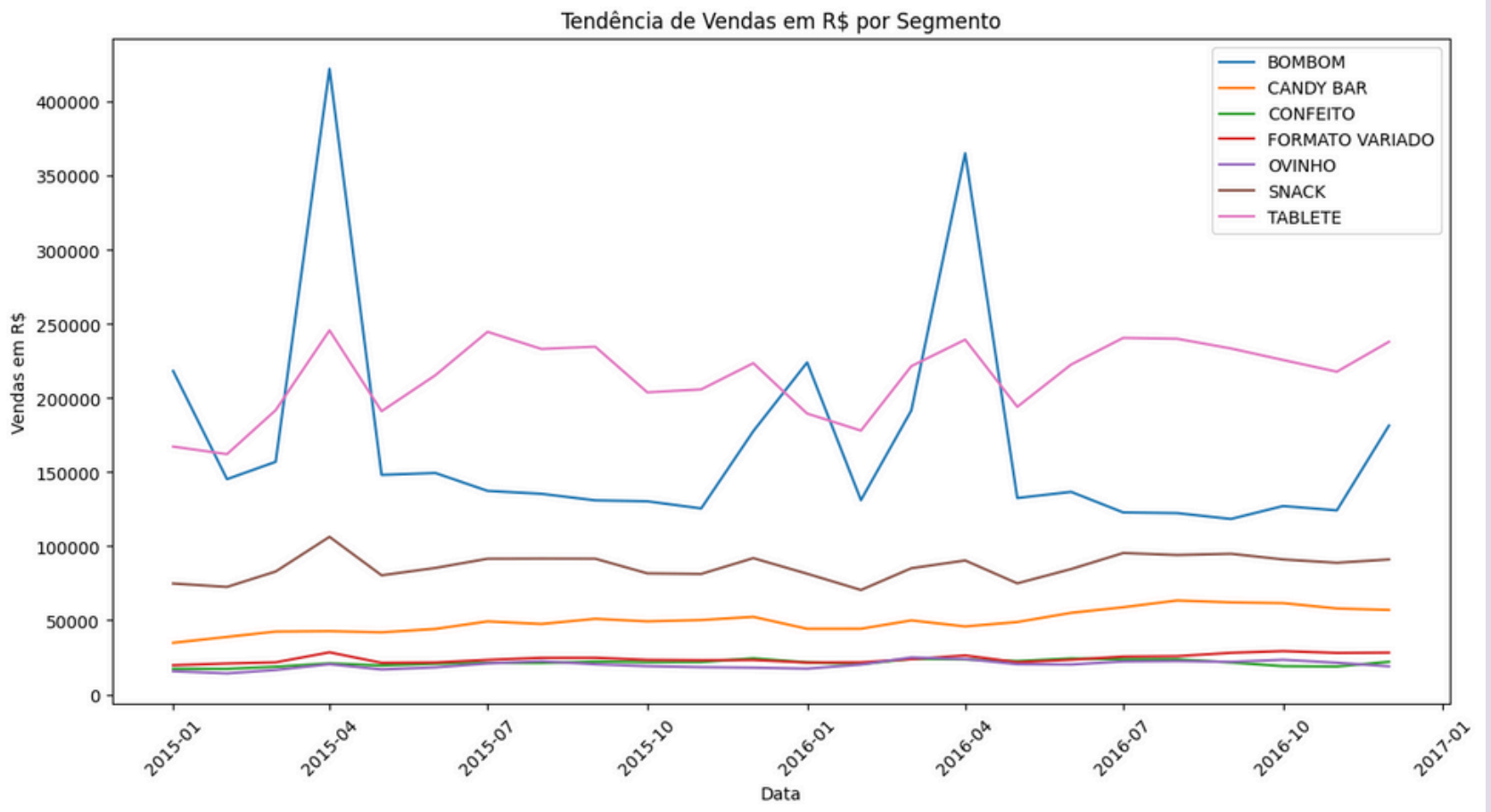
Partiu-se, então, para o próximo passo, que era avaliar a variedade de segmentos:

Mercado	Segmento	Faturamento
AREA VI - INA	TABLETE	1441708,5
AREA V - INA	TABLETE	1035659,8
AREA VI - INA	BOMBOM	959092,8
AREA V - INA	BOMBOM	766632,5
AREA IV - INA	TABLETE	755123,9
AREA II - INA	TABLETE	748680,7
AREA II - INA	BOMBOM	668934,7
AREA I - INA	BOMBOM	611864,7
AREA VI - INA	SNACK	522308,3
AREA V - INA	SNACK	474019,9

Analisando o faturamento por segmento e mercado, é possível notar que os principais destaques são para tablete e bombom, e que ambos os segmentos são bem relevantes para a área VI. Ainda, snacks para a área VI também aparece entre os top 10 faturamentos, evidenciando que a variedade de segmentos para esse mercado é importante.

Vale dizer que todos os segmentos estão presentes em todos os mercados, apenas com faturamentos diferentes, evidenciando a regionalidade.

Para aprofundar ainda mais, é possível perceber que o segmento de bombons atinge um grande pico nos meses de abril e dezembro, já o segmento de tablete é mais consistente e o preferido para consumo diário.



Analisando a questão da representatividade dos fornecedores em relação ao faturamento, é possível perceber que os fabricantes 5, 7 e 12 possuem maior destaque.



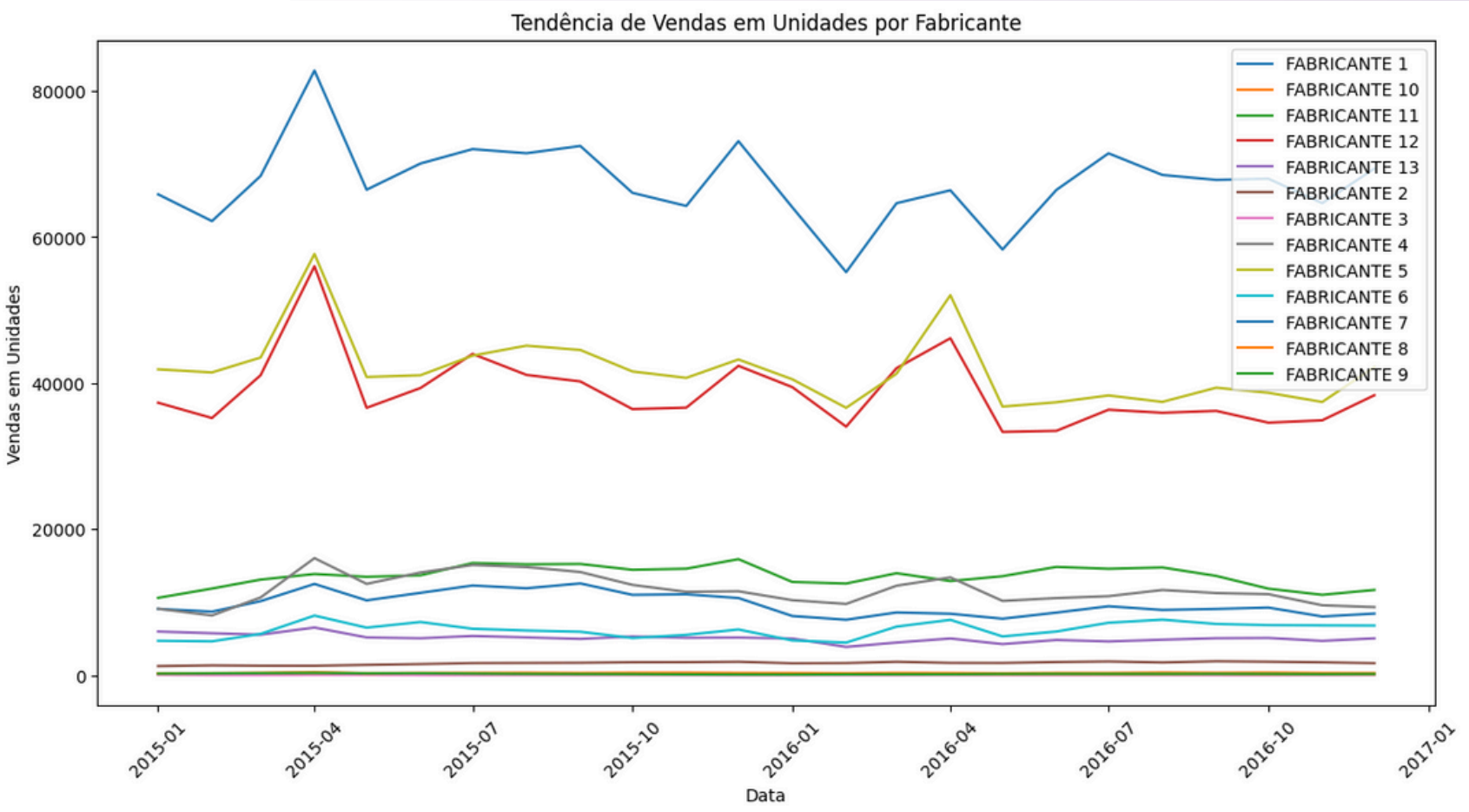
# Análise dos dados

Analizando os maiores faturamentos por mercado e fabricantes, temos o seguinte comportamento:

Mercado	Fabricante	R\$
AREA VI - INA	FABRICANTE 7	1385995,2
AREA V - INA	FABRICANTE 7	1100630,4
AREA VI - INA	FABRICANTE 12	1050684,1
AREA V - INA	FABRICANTE 12	758310
AREA IV - INA	FABRICANTE 7	678566,7
AREA II - INA	FABRICANTE 7	638474,2
AREA I - INA	FABRICANTE 5	606926,1
AREA II - INA	FABRICANTE 5	597857,5
AREA VII - INA	FABRICANTE 7	497718,4
AREA I - INA	FABRICANTE 7	472032,4

É notável também que a relevância dos fabricantes 7 e 12 na área VI é bastante elevada, o que é um forte indicativo de que os fabricantes também são fatores de influência para o destaque da área VI em relação às demais.

As unidades vendidas podem nos mostrar uma visão ainda mais aprofundada do comportamento de cada fabricante. Para isso, foi plotado um gráfico de linhas com as unidades vendidas por fabricante.



Esta análise final nos mostra que o fabricante 7 é o maior player no setor de chocolates, seguido pelos fornecedores 5 e 12. O fornecedor 3 possui um percentual muito baixo, menos de 1% do faturamento total, seguido pelos fornecedores 9 e 10 que estão na faixa de 1% de representação do faturamento.

## Conclusões:

- O faturamento de 2015 foi menor do que 2016, sendo que os valores calculados foram, respectivamente de: R\$ 6.916.129,9 e R\$ 7.134.671, correspondente a 3,1% de aumento. Uma possível justificativa para isso seria o aumento do preço médio por kg para a categoria de chocolates, que eventualmente pode causar uma queda na quantidade de unidades vendidas, conforme foi observado ao analisar essa dimensão.
- Os principais segmentos são tablete, bombom e snacks e os fabricantes 5, 7 e 12 se destacam como os principais, em termos de faturamento.
- São observados picos de vendas em Abril, provavelmente relacionados à Páscoa.
- Por meio do mapa de correlação foi possível perceber que o relacionamento entre preço médio por kg e data; e total em kg e faturamento estão diretamente relacionados. Para inversamente proporcionais, tem-se o preço médio por kg e o total de kg vendidos. Evidenciando então que é uma regra de mercado que a data impacta no preço, e isso por sua vez impacta no faturamento, uma vez que, com o preço mais alto, menores quantidades em kg são vendidas.
- A área VI se destaca em relação às demais, em termos de vendas em R\$, unidades e kg. Pela análise dos dados, os principais fatores que parecem ter influência direta nesse destaque são: a representatividade dos segmentos e os fabricantes.

## Recomendações:

- Clarificar o que está contido em cada categoria (ex.: Ovinhos inclui ovos de Páscoa, ou este está em Formatos Variados?). Assim, seria possível entender e prever melhor os comportamentos de acordo com a época do ano.
- Para alguns produtos, a unidade de medida de quilogramas talvez não seja tão interessante. Por exemplo, para confeitos, a quantidade em kg vendida aparece zerada, e isso pode estar relacionado ao fato de ser um produto leve, em que o peso não é significativo perante a quantidade de casas decimais utilizadas.
- Como alguns fabricantes se destacam muito mais do que outros, seria interessante entender quais são os pontos fortes desses fornecedores, a fim de criar boas práticas e replicar para os demais, garantindo um padrão entre eles. Ou ainda, entender o quanto esses fabricantes destaques conseguiriam absorver as vendas de outros fabricantes minoritários, visando reduzir a quantidade de fornecedores, reduzindo, assim, custos operacionais.
- Compreender as metas de negócio para os próximos anos visando desenhar estratégias assertivas baseadas nos dados analisados (ex.: Se o objetivo for aumentar o faturamento, uma leve redução de preço deve ajudar na quantidade de vendas, visto o comportamento dos últimos anos).

**Obrigado!**

