

- A seguinte análise exploratória de dados foi proveniente da base de dados previamente tratada em Python, após testes estatísticos para garantir a qualidade e segurança dos dados.
- A tabela df é, portanto, um dataframe criado pelo pipeline em Python, demonstrando a versatilidade e capacidade da função de carga de dados através de scripts.
- As features foram previamente preparadas e tratadas no momento ETL, como segue figura abaixo:
- Demais tratativas foram feitas em Power Query

```
import pandas as pd
import numpy as np
import matplotlib as plt
import seaborn as sns
from datetime import date

df = pd.read_excel('G:/Meu Drive/L3 challenge/01 Artifact/data/sample.xls')

#
colstriper = ['Row ID', 'Country', 'Postal Code']
df = df.drop(colstriper, axis = 1)

# Podemos criar também uma feature que irá definir o custo de cada produto, através de (Sales - Cost = Profit)
df['Cost'] = round(df['Sales']-df['Profit'],2)

# Criamos uma feature para definir o valor unitário de cada produto para um estudo mais específico de preços.
df['Unit Price'] = round(df['Sales']/df['Quantity'],2)

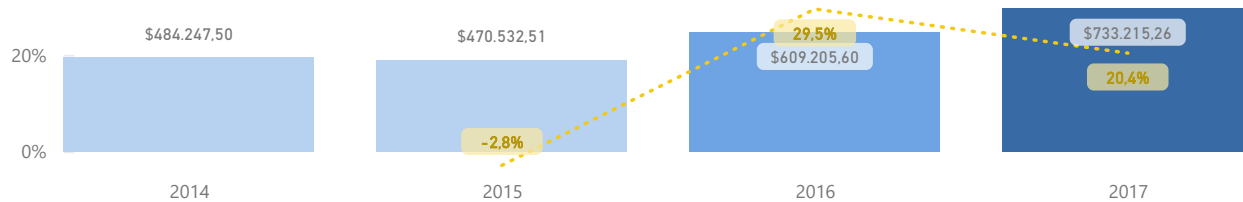
# Função para arredondar a feature 'Profit' para 2 casas decimais e melhor aspecto do modelo.
df['Profit'] = round(df['Profit'],2)

# Criamos um feature para calcular a diferença, em dias, da data de envio e data do pedido para possível estudo e feito
# transformação do objeto originalmente 'datetime' para objeto numérico (int64)
df['Date Diff'] = df['Ship Date'] - df['Order Date']
df['Date Diff'] = (df['Date Diff'] / np.timedelta64(1,'D')).astype(int)

# O objeto criado, ou dataset final já tratado, está pronto para consumo na ferramenta de visualização.
df.head(5)
```

TOTAL VENDAS E CRESCIMENTO (%) ANUAL

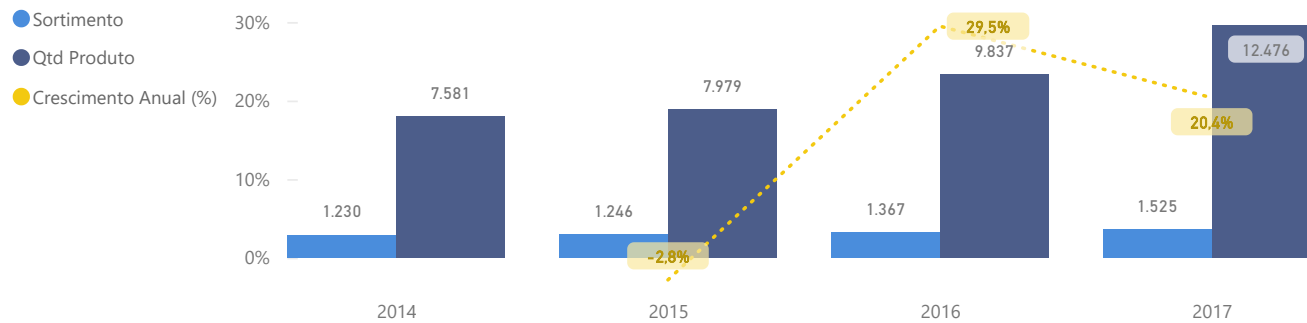
● Total Vendas ● Crescimento Anual (%)



Resumo de crescimento anual de vendas mostra constante crescimento.

DESEMPENHO DE PRODUTOS POR ANO

- Sortimento
- Qtd Produto



Relação entre o Sortimento (que é a quantidade de SKUs no portfólio) e o volume de produtos vendidos no ano. Apesar do ligeiro crescimento em 2015, há uma queda de vendas, a qual recupera fôlego nos anos seguintes.

DESEMPENHO MENSAL TOTAL VENDAS POR ANO

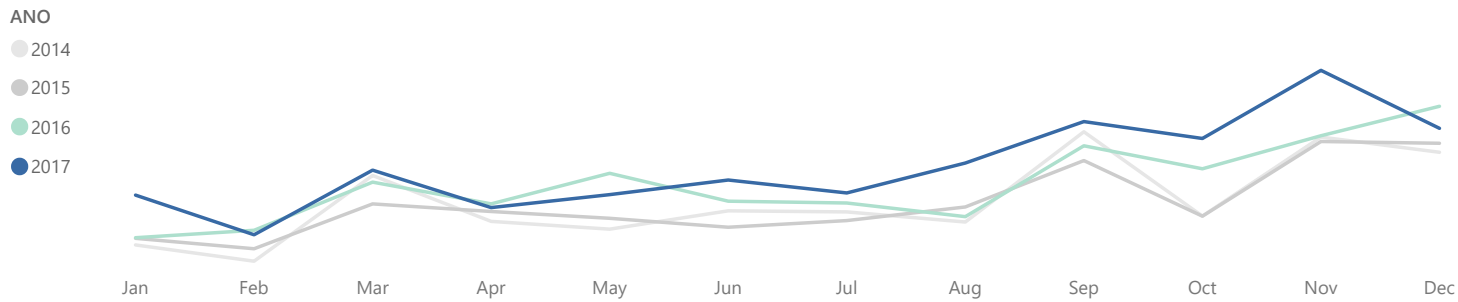
ANO

2014

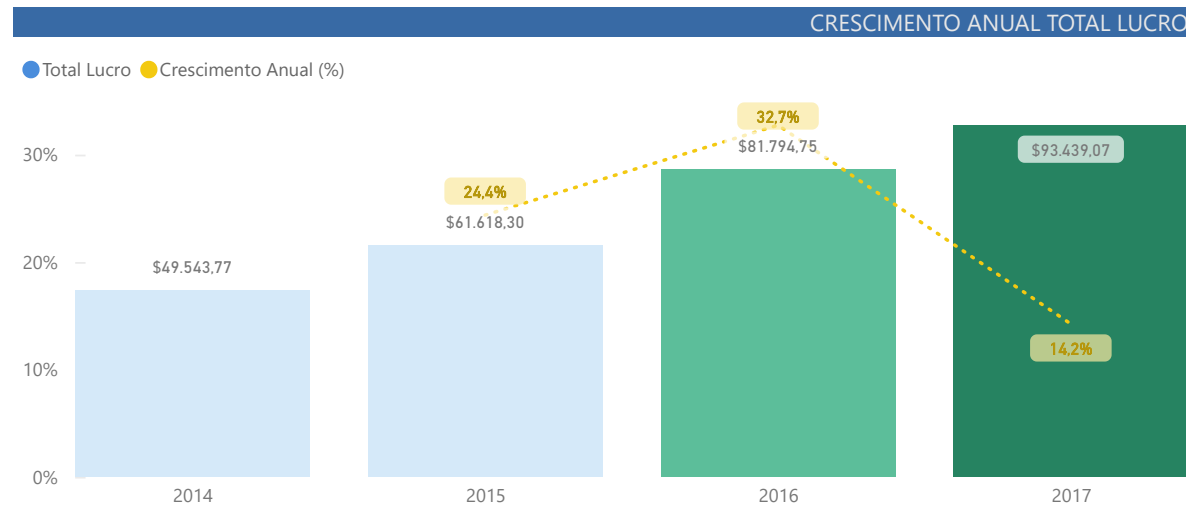
● 2015

● 2016

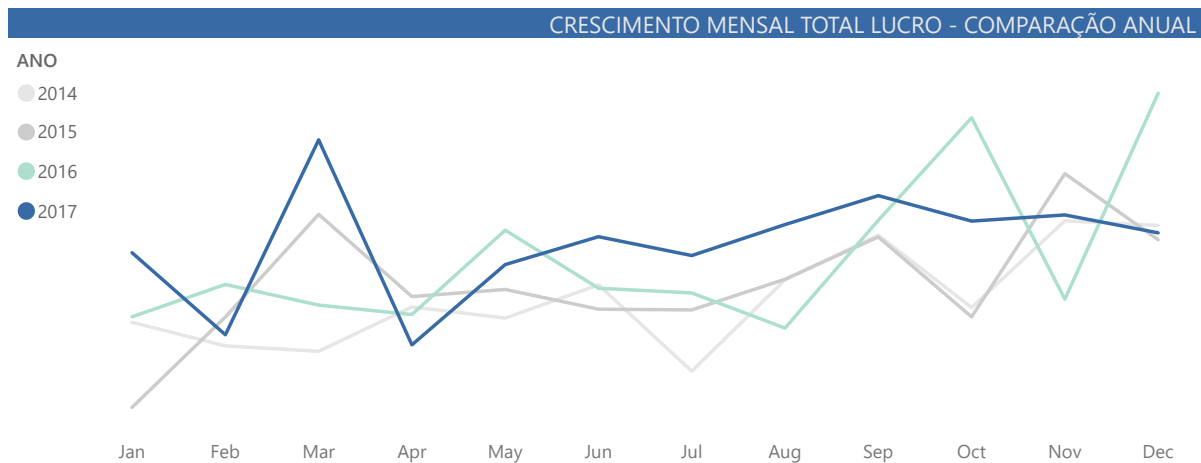
● 2017



Com o comportamento mensal de vendas, podemos definir uma certa sazonalidade nos períodos de Março, Outubro e Dezembro, com elevada queda em Novembro e determinar estratégias para este comportamento.



O Lucro Anual também apresentou bom crescimento. Podemos ver que no último ano, em 2017, o houve uma maior constância do crescimento de lucro.

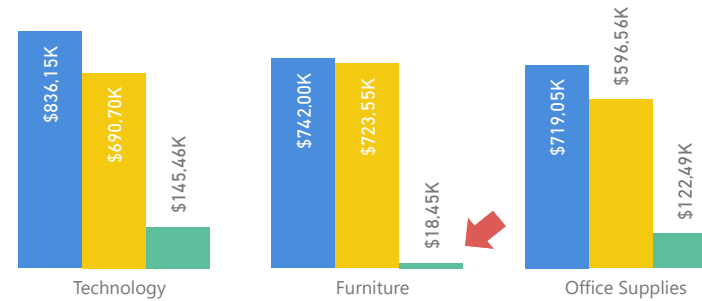


Resumo das principais métricas
de desempenho ao longo dos
anos medidos

RESUMO DESEMPENHO ANUAL				
	2014	2015	2016	2017
01.01 - TOTAL VENDAS	\$484.247,50	\$470.532,51	\$609.205,60	\$733.215,26
01.03 - TOTAL LUCRO	\$49.543,77	\$61.618,30	\$81.794,75	\$93.439,07
01.04 - TOTAL CUSTO	\$434.703,86	\$408.914,25	\$527.410,78	\$639.776,29
01.02 - TOTAL QTD PRODUTO	7.581	7.979	9.837	12.476
01.06 - SORTIMENTO	1.230	1.246	1.367	1.525
01.07 - QTD CLIENTES	595	573	638	693

DESEMPENHO GERAL POR CATEGORIA

● Total Vendas ● Total Custo ● Total Lucro

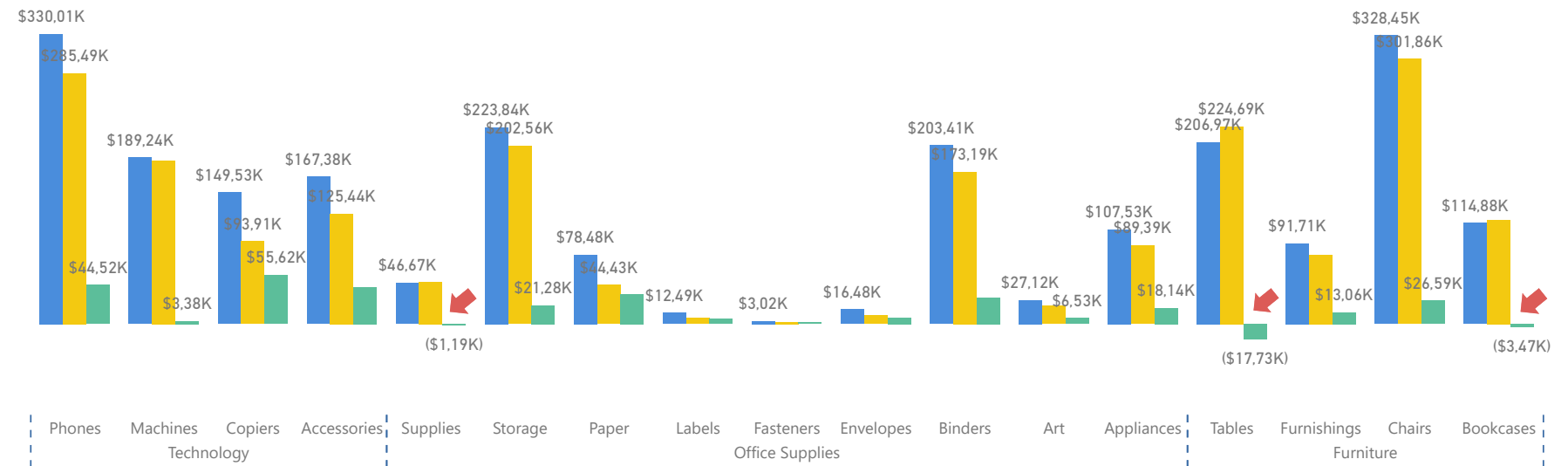


Ao analisar sob a perspectiva de categorias de produtos, pode-se notar um baixo desempenho do lucro na categoria 'Furniture'

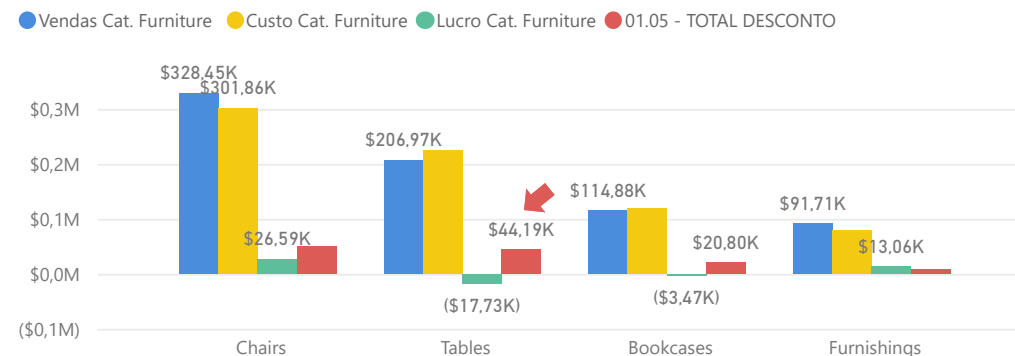
Quando fazemos um drill-down na categoria, podemos observar que algumas sub-categorias apresentam baixo desempenho, entre elas "Tables" e "Bookcases"

DESEMPENHO GERAL POR SUB-CATEGORIA

● Total Vendas ● Total Custo ● Total Lucro



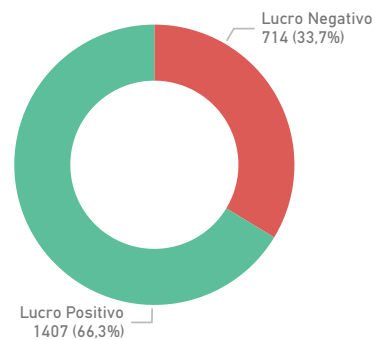
DESEMPENHO POR SUB-CATEGORIA (FURNITURE)



Ao analisar as ordens da Categoria "Furniture", podemos ver que 33,6% apreñsetam baixo desempenho e este ato se propaga ao longo das sub-categorias e apresenta um padrão durante todos os anos..

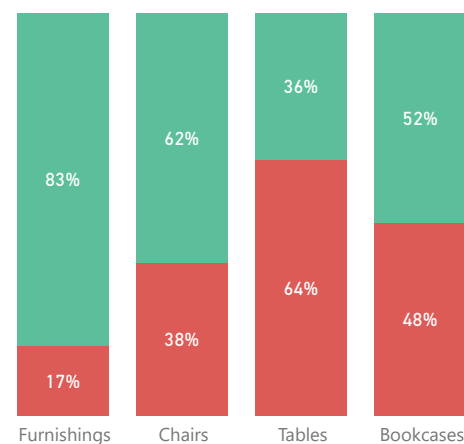
É possível detectar um desempenho negativo de lucro na sub-cat "Tables", devido ao ao custo excessivo.

COMPOSIÇÃO TIPO DE LUCRO



CLASSIFICAÇÃO DO LUCRO (POSITIVO / NEGATIVO)

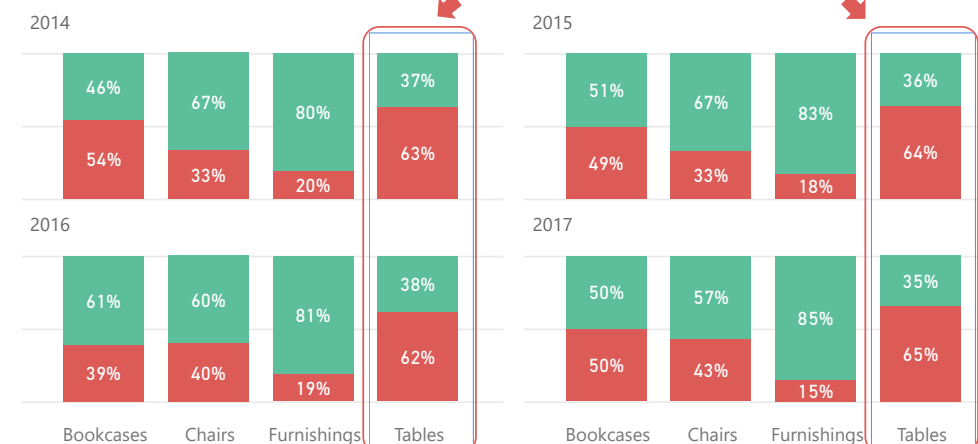
Class ● Lucro Negativo ● Lucro Positivo



CLASSIFICAÇÃO DO LUCRO (POSITIVO / NEGATIVO) - COMPARAÇÃO ANUAL

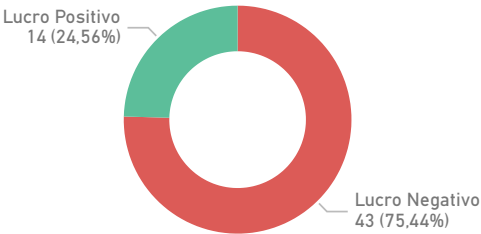
Class ● Lucro Negativo ● Lucro Positivo

A sub-categoria se apresenta problemática



57
SUB-CAT TABLES - SORTIMENTO

DESEMPENHO DE PRODUTOS (SUB-CAT. TABLES)



RQuanto olhamos para dentro da sub-categoria "Tables", podemos notar que 75,4% dos SKUs apresnetam baixo desempenho, endossando o levantamento anterior.

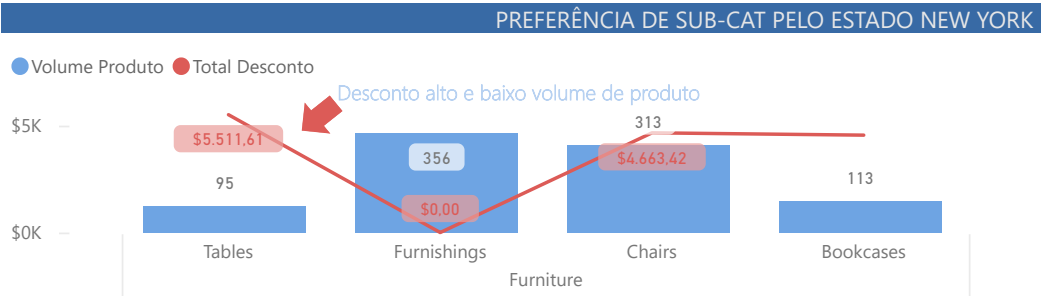
TOP 10 LUCRO MAIS BAIXO - DESEMPENHO POR SKUS - SUB-CAT. TABLES

Product ID	Product Name	Qtd Clientes Interessados	Total Vendas	Total Custo	Total Lucro	Preço Médio Unit.	Variância Preço Médio Unit.
FUR-TA-10000198	Chromcraft Bull-Nose Wood Oval Conference Tables & Bases	5	\$9.917,64	\$12.793,76	(\$2.876,11)	396,71	7.771,39
FUR-TA-10001889	Bush Advantage Collection Racetrack Conference Table	7	\$9.544,73	\$11.479,13	(\$1.934,40)	275,74	4.241,76
FUR-TA-10001950	Balt Solid Wood Round Tables	4	\$6.518,75	\$7.719,81	(\$1.201,06)	357,19	7.974,49
FUR-TA-10004289	BoxOffice By Design Rectangular and Half-Moon Meeting Room Tables	3	\$1.706,25	\$2.854,69	(\$1.148,44)	113,02	26,55
FUR-TA-10004154	Riverside Furniture Oval Coffee Table, Oval End Table, End Table with Drawer	5	\$4.446,18	\$5.593,58	(\$1.147,41)	200,80	1.974,79
FUR-TA-10004256	Bretford "Just In Time" Height-Adjustable Multi-Task Work Tables	4	\$5.634,90	\$6.599,10	(\$964,20)	323,49	5.553,35
FUR-TA-10002958	Bevis Oval Conference Table, Walnut	8	\$6.942,07	\$7.798,10	(\$856,02)	189,21	2.170,95
FUR-TA-10002533	BPI Conference Tables	5	\$2.241,87	\$3.037,84	(\$795,98)	92,01	759,28
FUR-TA-10003954	Hon 94000 Series Round Tables	5	\$7.404,50	\$8.085,71	(\$681,22)	248,79	1.964,97
FUR-TA-10003238	Chromcraft Bull-Nose Wood 48" x 96" Rectangular Conference Tables	6	\$4.297,64	\$4.909,22	(\$611,59)	367,32	7.758,08
Total		51	\$58.654,52	\$70.870,94	(\$12.216,43)	257,71	13.311,81

Podemos ver o Bottom 10 Produtos com lucro mais baixo e determinar os SKUs mais problemáticos. Uma interessante análise, é a Variância do Preço Médio Unitário. Isso pode significar violação de preço de venda e altas taxas de desconto.

DESEMPENHO POR ESTADO - SUB-CAT. TABLES							
State	Qtd Clientes	Volume Produto	Total Vendas	Total Custo	Total Lucro	Diferença (%) Custo por Vendas	Total Desconto
California	66	281	\$45.430,23	\$45.733,81	(\$303,58)	0,67%	\$9.086,05
Washington	21	109	\$25.062,32	\$20.186,34	\$4.875,98	19,46%	\$0,00
New York	26	95	\$13.779,02	\$18.314,68	(\$4.535,65)	32,92%	\$5.511,61
Texas	31	118	\$15.760,66	\$17.977,34	(\$2.216,68)	14,06%	\$4.728,20
North Carolina	8	49	\$9.681,73	\$13.366,00	(\$3.684,25)	38,05%	\$3.872,69
Virginia	12	63	\$14.042,47	\$11.503,79	\$2.538,68	18,08%	\$0,00
Illinois	17	65	\$6.550,67	\$10.860,43	(\$4.309,76)	65,79%	\$3.275,34
Pennsylvania	14	65	\$8.052,19	\$10.640,96	(\$2.588,76)	32,15%	\$3.220,87
Ohio	16	49	\$7.887,11	\$10.602,45	(\$2.715,34)	34,43%	\$3.154,85
Tennessee	8	34	\$6.214,36	\$8.877,77	(\$2.663,40)	42,86%	\$2.485,74
Florida	10	45	\$5.484,97	\$7.978,10	(\$2.493,12)	45,45%	\$2.468,24
Arizona	9	26	\$2.095,25	\$6.277,22	(\$2.281,89)	57,11%	\$1.007,69
Total	261	1.241	\$206.965,53	\$224.691,18	(\$17.725,60)	8,56%	\$44.192,20

Por fim, podemos ver o desempenho de "Tables" sob a perspectiva de região: New York apresenta um Custo muito elevado e um desempenho de lucros muito negativo ao relacionar a quantidade de clientes com os outros estados.



Podemos ver pelo comportamento de compras em New York, a sub-categoria "Tables" é a menos querida em relação ao volume vendido e com uma alta taxa de desconto, sacrificando a contribuição das vendas

- Esta análise teve o objetivo de investigar as causas e raízes e encontrar argumentos que pudesse explicar o teste estatístico feito previamente, identificando o Lucro com valor negativo.
- O aspecto abordado foi o foco em uma Sub-Categoria específica que pode explicar o fato e gera ações corretivas.

- Ao verificar o desempenho de vendas de uma forma geral, pôde-se verificar que a Categoria "Furniture" apresentava baixo desempenho;
- Ao executar drill-down, foi possível detectar que a Sub-Categoria "Tables" apresentava desempenho inferior aos demais, tornando-se motivo e foco da investigação.
- Ao mudar a perspectiva, realizando um cross entre região e sub-cat, foi possível detectar estado com maior incidência de anomalia, destacando alto custo e altas taxas de descontos.
- Pode-se presumir que devido a baixa preferência por "Tables" em New York, as altas taxas de desconto elevaram muito o custo, sacrificando a margem e se tornando prejuízo.
- A investigação poderá direcionar os analistas responsáveis pela reestruturação da estratégia de vendas dessa categoria nessa região e tentar recuperar a contribuição