

Introdução

Temos nesse artigo uma dataset (download pode ser feito no meu github. Link no final do artigo) de informações sociodemográficas e de vendas de uma empresa no ramo alimentício. Essa empresa possui milhares de clientes em todo mundo e atende a quase 1 milhão de clientes por ano. Os principais produtos que eles vendem são vinhos, carnes, frutas, peixes e doces. O grande objetivo desse estudo é fazer uma extensa EDA e criar KPIs para melhor entender as características socioeconômicas dos respondentes (clientes da empresa) para realizarmos recomendações para o time de marketing da empresa.

O arquivo do Power BI, PDF e o dataset encontram-se no meu github. Link no final.

Análise

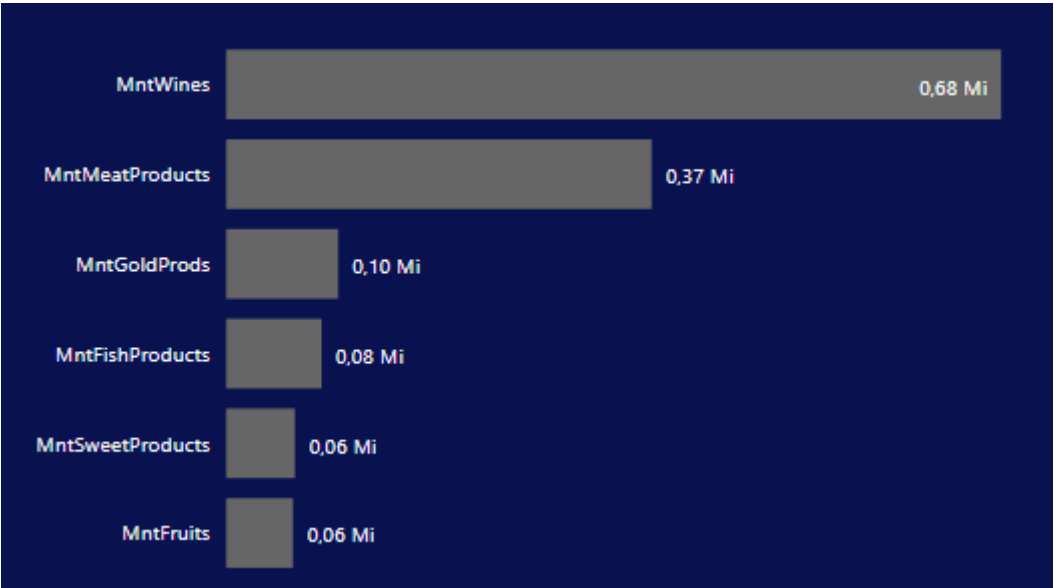
Figura 1 – Parte dos dados

Y3 ID	Y3 Year_Birth	A3 Education	A3 Marital_Status	A3 Income	Y3 MntWines	Y3 MntFruits	Y3 MntMeatProducts	Y3 MntFishProducts	A3 Dt_Customer
1826	1970	Graduation	Divorced	\$84,835.00	189	104	379	111	6/16/14
1	1961	Graduation	Single	\$57,091.00	464	5	64	7	6/15/14
10476	1958	Graduation	Married	\$67,267.00	134	11	59	15	5/13/14
1386	1967	Graduation	Together	\$32,474.00	10	0	1	0	5/11/14
5371	1989	Graduation	Single	\$21,474.00	6	16	24	11	4/6/14
7348	1958	PhD	Single	\$71,691.00	336	139	411	240	3/17/14
4073	1954	In Cycle	Married	\$63,564.00	769	80	252	15	1/29/14
1991	1967	Graduation	Together	\$44,931.00	78	0	11	0	1/18/14
4047	1954	PhD	Married	\$65,324.00	384	0	102	21	1/11/14
9477	1954	PhD	Married	\$65,324.00	384	0	102	21	1/11/14

Na figura 1, temos uma parte dos dados que o dataset contém. Dentre eles, ‘Year Birth’ (idade), ‘Education’, ‘Income’, ‘MntWines’ que seria a quantidade gastos nos últimos 2 anos em produtos de vinho, ‘Dt Customer’ que representa desde quando é cliente. Outro campo importante, mas que não aparece na figura 1 é ‘AcceptedCmd’ que diz se o cliente respondeu positivamente ou não a uma campanha de marketing. No anexo no final do texto, encontra-se os detalhes de todas as ‘features’ do dataset.

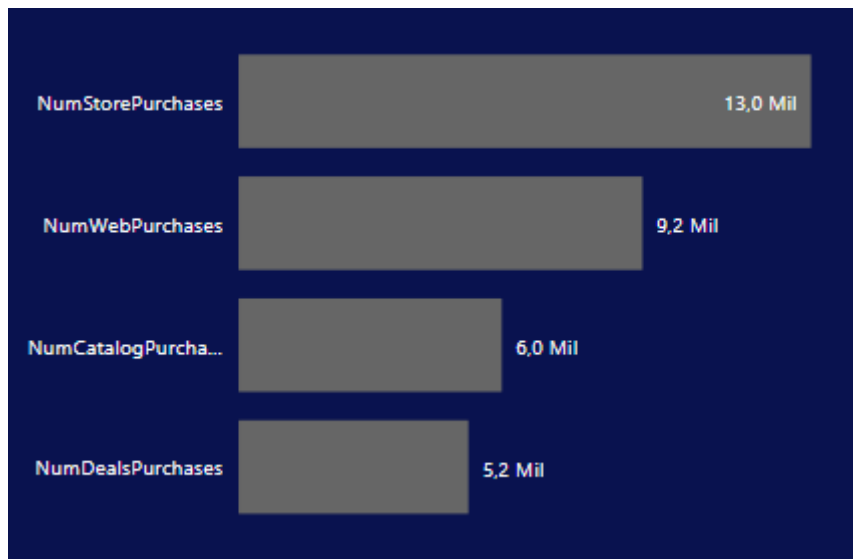
O primeiro indicador é o total gasto em 2 anos. Para isso, basta somar todas as colunas de produtos, como ‘MntWines’, ‘MntFruits’ e etc. Também vamos calcular a média do total gasto com esses produtos. Temos que o total gasto em 2 anos com os produtos foram de 1357 mil dólares e a média de 605,80 dólares.

Figura 2 – Valor gasto por produto



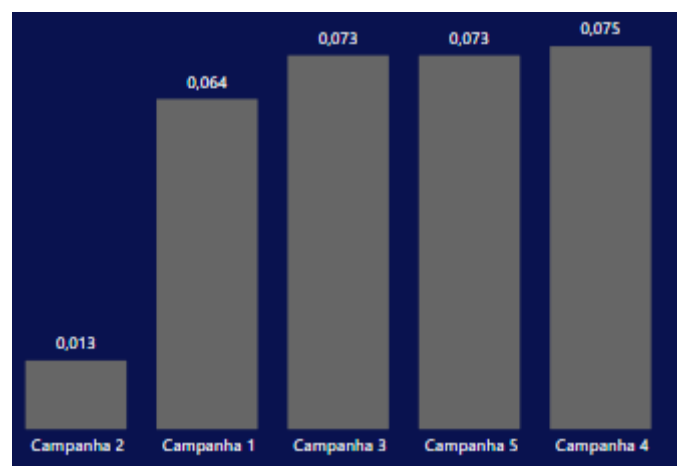
Um dado importante é saber o quanto foi gasto em cada produto para termos uma ideia qual(is) performaram melhor do que os outros. Na figura 2, temos que 'MntWines' e 'MntMeatProducts' foram os produtos que mais venderam nesses 2 anos. Além disso, as suas vendas foram muito maiores do que os dos outros produtos. 'MntWines' e 'MntMeatProducts' representam 77,7 % de todas as vendas. Ou seja, esses 2 produtos são o carro chefe da empresa.

Figura 3 – Valor gasto por canal



Na mesma lógica, podemos fazer um gráfico para analisar quais canais de venda foram os que mais tiveram sucesso. Na figura 3, temos que o número de vendas em lojas físicas, 'NumStorePurchases', e vendas na internet, 'NumWebPurchases', representam 66,4 % das vendas, enquanto as vendas por catálogos de revistas, 'NumCatalogPurchases', e vendas com descontos, 'NumDealsPurchases', representam 33,6 %.

Figura 4- Taxa de aceitação por campanha (%)



Foram feitas 5 campanhas de marketing e foi computado sucesso de cada uma, em outras palavras, se o cliente respondeu a oferta que foi feita, ela é considerada um sucesso (1 na tabela), caso contrário ela não foi bem sucedida (0 na tabela). Vemos que na figura 4, a campanha 2 foi a pior com 1,3 % de taxa de aceitação dentre as 5 campanhas feitas.

Figura 5 – Total gasto por faixa de salário

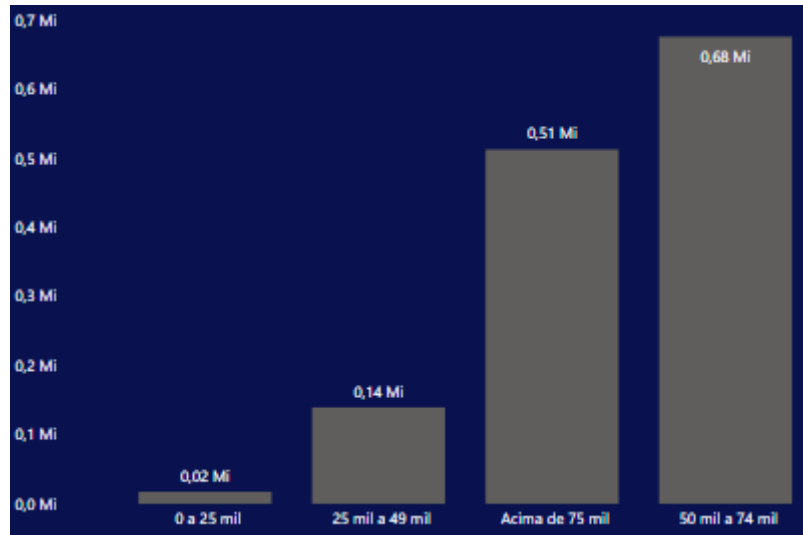
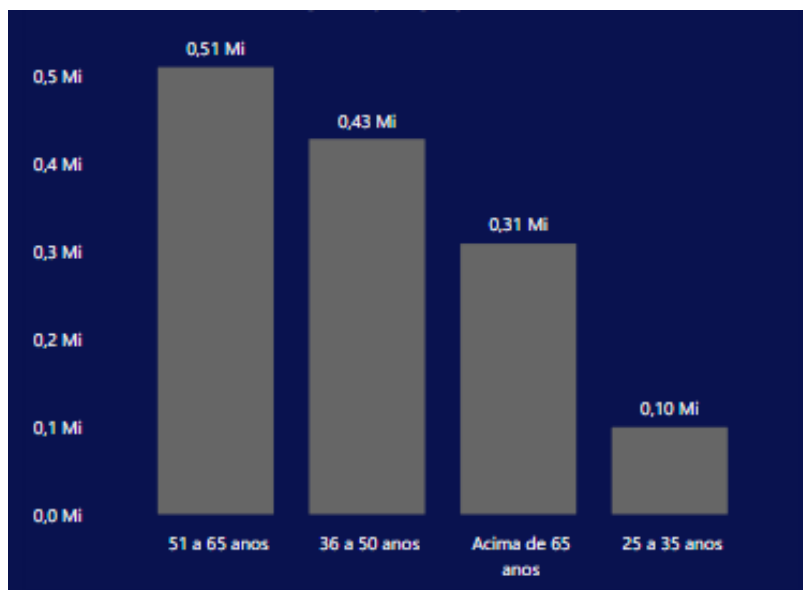
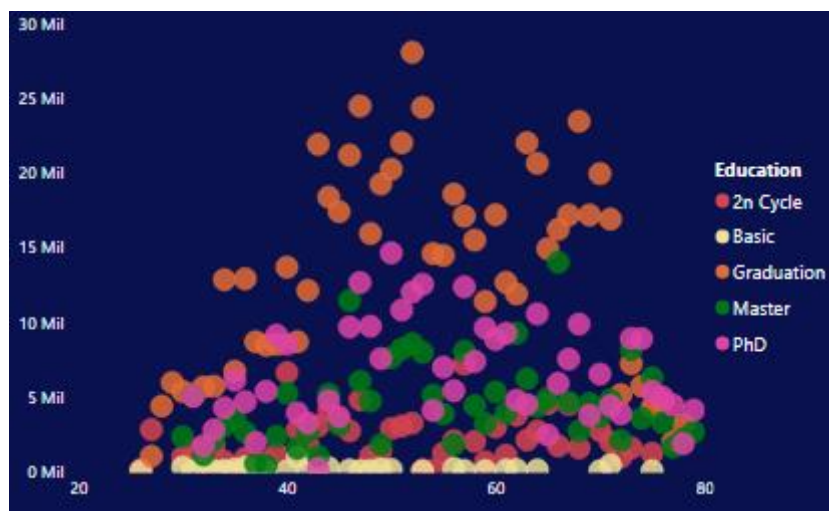


Figura 6 – Total gasto por grupo de idade



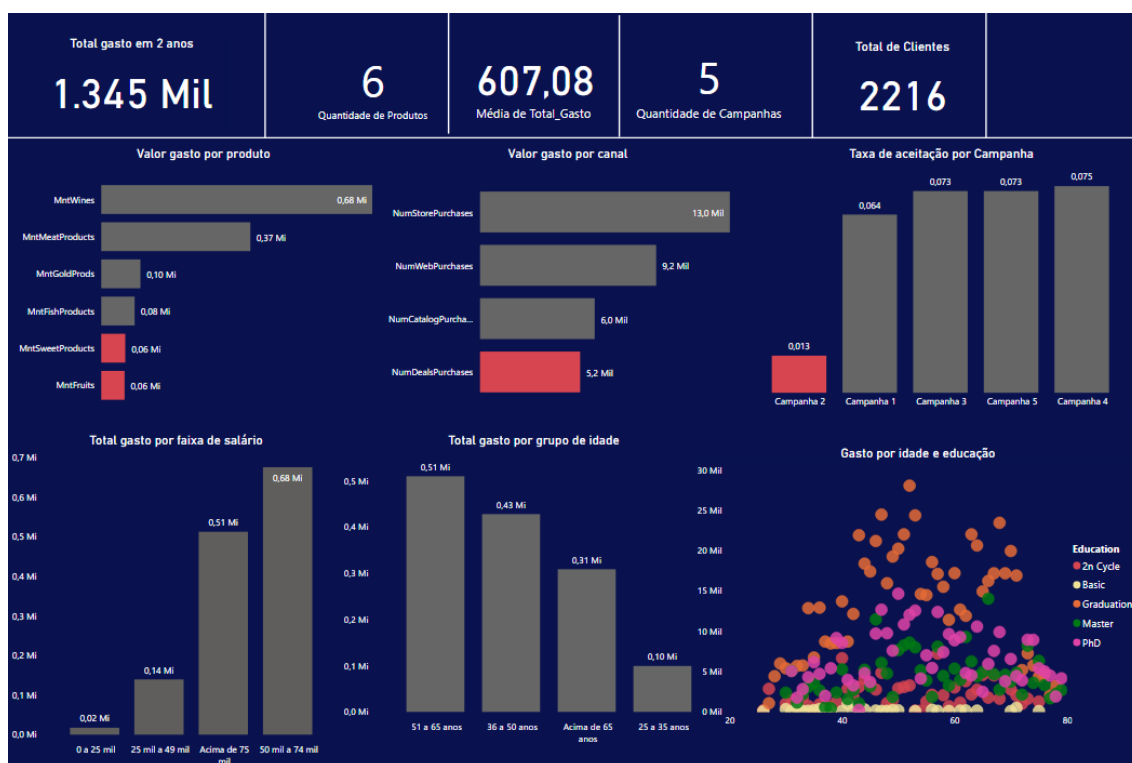
Para termos um entendimento melhor sobre os dados socioeconômicos coletados, podemos separar as faixas de salários e o quanto foi gasto por cada faixa, como visto na figura 5. Temos que a faixa entre 50 a 74 mil e a acima de 75 mil são as que mais gastam com os produtos e eles representam mais de 88 % dos que mais gastam. Já no total gasto por grupo de idade, figura 6, temos que o grupo que menos gasta são os mais jovens, entre 25 a 35 anos, representando apenas 7,4 % do total.

Figura 7 – Gasto por idade e educação



Com um gráfico de dispersão, temos como avaliar a relação entre idade, educação e o total gasto, visto na figura 7. Dela podemos observar que as pessoas com a nível de educação 'Basic' são as que menos gastam, as pessoas com '2nd Cycle' são as que gastam entre 0 e 5 mil e as que com 'Graduation' são as que mais gastam e elas estão mais ou menos concentradas entre as idades 40 a 65-70 anos. Logo em seguida, em segundo lugar no quesito que mais gastam, são as pessoas com 'PhD'.

Figura 8 – Dashboard do Power BI



Na figura 8, temos como ficou a dashboard do Power BI com todos as visualizações feitas.

Conclusões

Com todas as visualizações feitas, podemos tirar as seguintes conclusões. São elas:

- 1- Valor gasto por produto – Vinhos representaram o maior gasto por produto em 2 anos, 55,17 %, se comparado aos demais produtos e foi mais de 1000 % maior que doces e frutas, que teve o menor valor, aproximadamente 4 %.
- 2- Valor gasto por canal – Canal 'NumStorePurchases' teve o maior canal de compras com 64 % do total e foi, aproximadamente, 150 % maior do que 'NumDealsPurchases'
- 3- Grupo dos 51 a 65 anos são os que mais gastam e representam 37,7 %. Logo em seguida temos o grupo dos 36 a 50 anos que representam 32 %. Por fim, quase 70 % dos clientes tem entre 36 a 65 anos.
- 4- Taxa de aceitação por campanha – A campanha 4 foi a mais bem sucedida e sua taxa de aceitação foi, aproximadamente, 477 % maior do que a campanha 2, que teve apenas 1,3 % de taxa de aceitação.
- 5- Gasto por faixa salarial – Clientes na faixa salarial entre 50 a 74 mil são os que mais gastam, representando 50 % dos gastos nos últimos 2 anos.

Recomendações

Com essas conclusões, podemos dar as seguintes recomendações para a empresa:

- 1- Repensar nas vendas dos produtos do tipo doce ('MntSweetProducts') e frutas ('MntFruits')
- 2- Desativar a campanha 2

Anexo

Feature	Description
AcceptedCmp1	1 if customer accepted the offer in the 1st campaign, 0 otherwise
AcceptedCmp2	1 if customer accepted the offer in the 2nd campaign, 0 otherwise
AcceptedCmp3	1 if customer accepted the offer in the 3rd campaign, 0 otherwise
AcceptedCmp4	1 if customer accepted the offer in the 4th campaign, 0 otherwise
AcceptedCmp5	1 if customer accepted the offer in the 5th campaign, 0 otherwise
Response (target)	1 if customer accepted the offer in the last campaign, 0 otherwise
Complain	1 if customer complained in the last 2 years
DtCustomer	data of customer's enrollment with the company
Education	customer's level of education
Marital	customer's marital status
Kidhome	number of small children in customer's household
Teenhome	number of teenagers in customer's household
Income	customer's yearly household income
MntFishProducts	amount spent on fish products in the last 2 years
MntMeatProducts	amount spent on meat products in the last 2 years
MntFruits	amount spent on fruits products in the last 2 years
MntSweetProducts	amount spent on sweet products in the last 2 years
MntWines	amount spent on wines products in the last 2 years
MntGoldProds	amount spent on gold products in the last 2 years
NumDealsPurchases	number of purchases made with discount
NunCatalogPurchases	number of purchases made using catalog
NunStorePurchases	number of purchases made directly in stores
NumWebPurchases	number of purchases made through company's web site
NumWebVisitsMonth	number of visits to company's web site in the last month
Recency	number of days since the last purchase