

iFood CRM Data Analyst Case

Vladimir Oliveira

Agenda

1. Contexto
2. Desafio
3. Desenvolvimento da solução
4. Possíveis abordagens

Contexto

Perspectiva de lucro nos próximo 3 anos não são promissoras, uma forma de inverter essa situação é focar nas campanhas de marketing.

Desafio

Problema

- Campanhas de marketing pouco otimizadas, gerando prejuízo

Causas

- Falta de uma abordagem quantitativa

Solução

- Propor uma segmentação de clientes com base em dados históricos de compras

Desenvolvimento da solução

Descrição dos dados

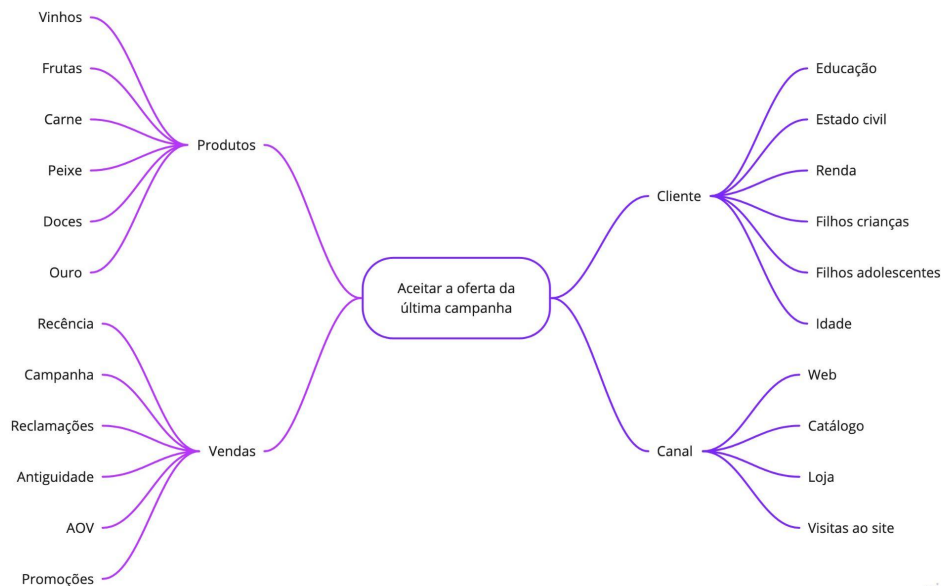
Número de linhas: 2240
Número de colunas: 29

	attributes	mean	median	std	min	max	range		skew	kurtosis	
Descriptive Statistics	Year_Birth	1.968,81	1970	11,98	1893	1996	103	-	0,35	0,72	
	Income	51.687,46	51075	25.603,63	0	666666	666666		6,33	147,94	
	Kidhome	0,44	0	0,54	0	2	2		0,64	-	0,78
	Teenhome	0,51	0	0,54	0	2	2		0,41	-	0,99
	Recency	49,11	49	28,96	0	99	99	-	0,00	-	1,20
	MntWines	303,94	173,5	336,52	0	1493	1493		1,18		0,60
	MntFruits	26,30	8	39,76	0	199	199		2,10		4,05
	MntMeatProducts	166,95	67	225,66	0	1725	1725		2,08		5,52
	MntFishProducts	37,53	12	54,62	0	259	259		1,92		3,10
	MntSweetProducts	27,06	8	41,27	0	263	263		2,14		4,38
	MntGoldProds	44,02	24	52,16	0	362	362		1,89		3,55
	NumDealsPurchases	2,33	2	1,93	0	15	15		2,42		8,94
	NumWebPurchases	4,08	4	2,78	0	27	27		1,38		5,70
	NumCatalogPurchases	2,66	2	2,92	0	28	28		1,88		8,05
	NumStorePurchases	5,79	5	3,25	0	13	13		0,70	-	0,62
	NumWebVisitsMonth	5,32	6	2,43	0	20	20		0,21		1,82

Desenvolvimento da solução

Mapa mental de hipóteses

A ideia aqui é listar qualitativamente o que aparenta ser uma entidade relevante no problema e quais os fatores que as influenciam. Serve para dar clareza ao formular hipóteses.



Desenvolvimento da solução

Hipóteses da análise exploratória

Aqui crio hipóteses sobre a influência dos atributos na variável resposta baseadas em conhecimento de negócio, ainda não pautadas em dados.

Durante o desenvolvimento, utilizo a Análise Exploratória para abordar o problema de duas formas:

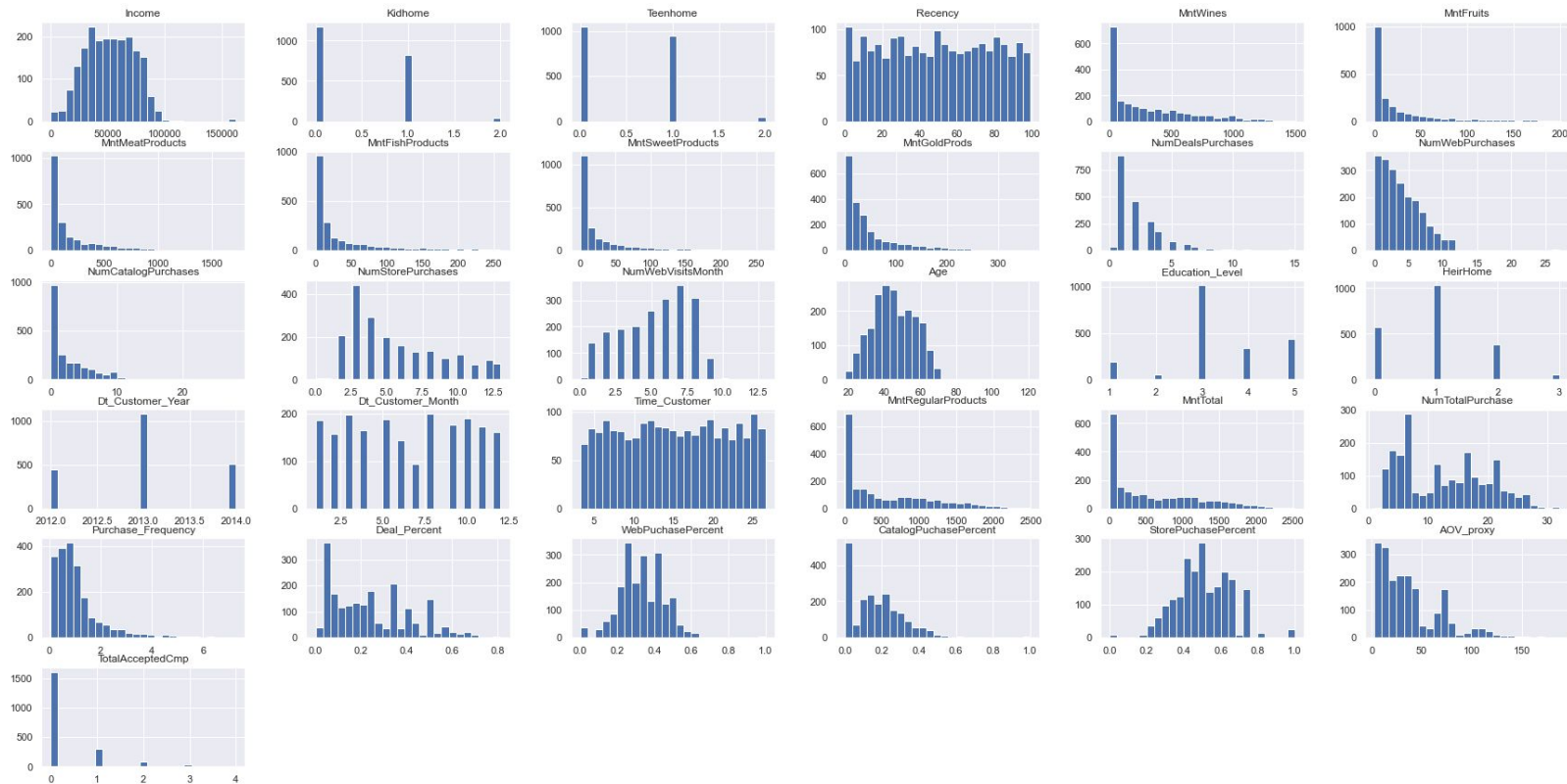
- Confirmando ou negando crenças de negócio
- Trazendo novos insights não necessariamente relacionados às hipóteses

Lista de hipóteses

- H1. Clientes com maior renda são MAIS propensos a aceitar a oferta da campanha
- H2. Clientes com mais filhos são MENOS propensos a aceitar a oferta da campanha
- H3. Clientes casados são MENOS propensos a aceitar a oferta da campanha
- H4. Clientes com maior nível de educação formal são MAIS propensos a aceitar a oferta da campanha
- H5. Clientes mais velhos são MENOS propensos a aceitar a oferta da campanha
- H6. Clientes que visitam mais o site são MAIS propensos a aceitar a oferta
- H7. Clientes que fazem mais compras são MAIS propensos a aceitar a oferta
- H8. Clientes que fazem mais compras via web são MAIS propensos a aceitar a oferta
- H9. Clientes que gastam mais são MAIS propensos a aceitar a oferta
- H10. Clientes que gastam mais em Ouro são MAIS propensos a aceitar a oferta
- H11. Clientes que gastam mais em Vinho são MAIS propensos a aceitar a oferta
- H12. Clientes que já reclamaram são MENOS propensos a aceitar a oferta da campanha
- H13. Clientes que compraram em outras campanhas são MAIS propensos a aceitar a oferta da campanha
- H14. Clientes mais antigos são MAIS propensos a aceitar a oferta da campanha
- H15. Clientes que fizeram compras mais recentes são MENOS propensos a aceitar a oferta da campanha
- H16. Clientes com maior AOV são MAIS propensos a aceitar a oferta da campanha
- H17. Clientes que compram mais em promoções são MAIS propensos a aceitar a oferta da campanha

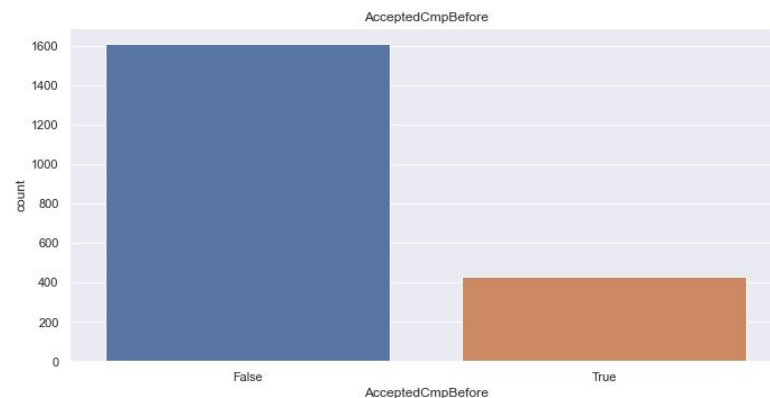
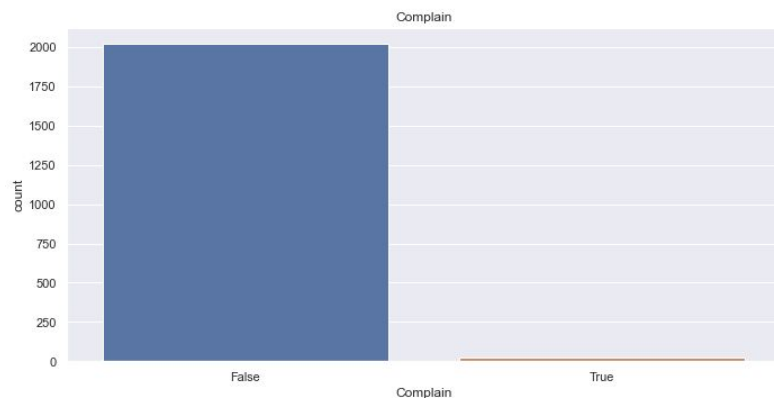
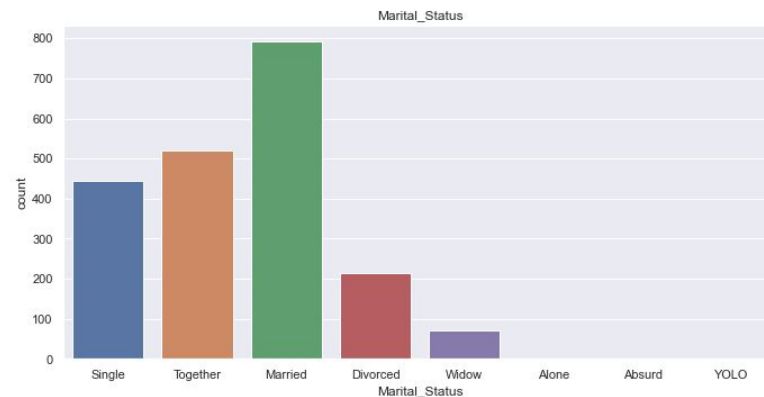
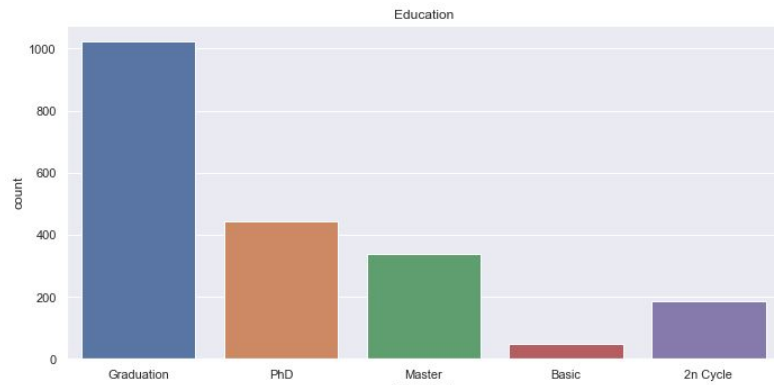
Desenvolvimento da solução

Análise exploratória: variáveis numéricas



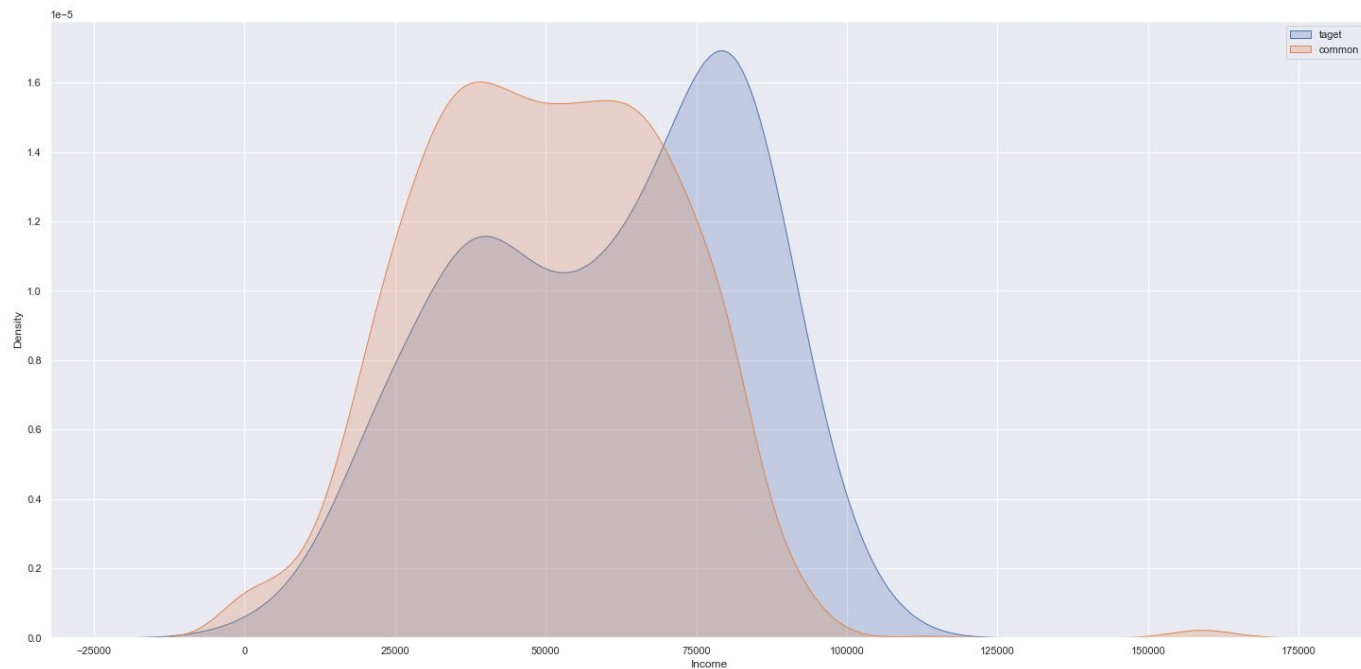
Desenvolvimento da solução

Análise exploratória: variáveis categóricas



H1. Clientes com maior renda são MAIS propensos a aceitar a oferta da campanha

[VERDADEIRO]

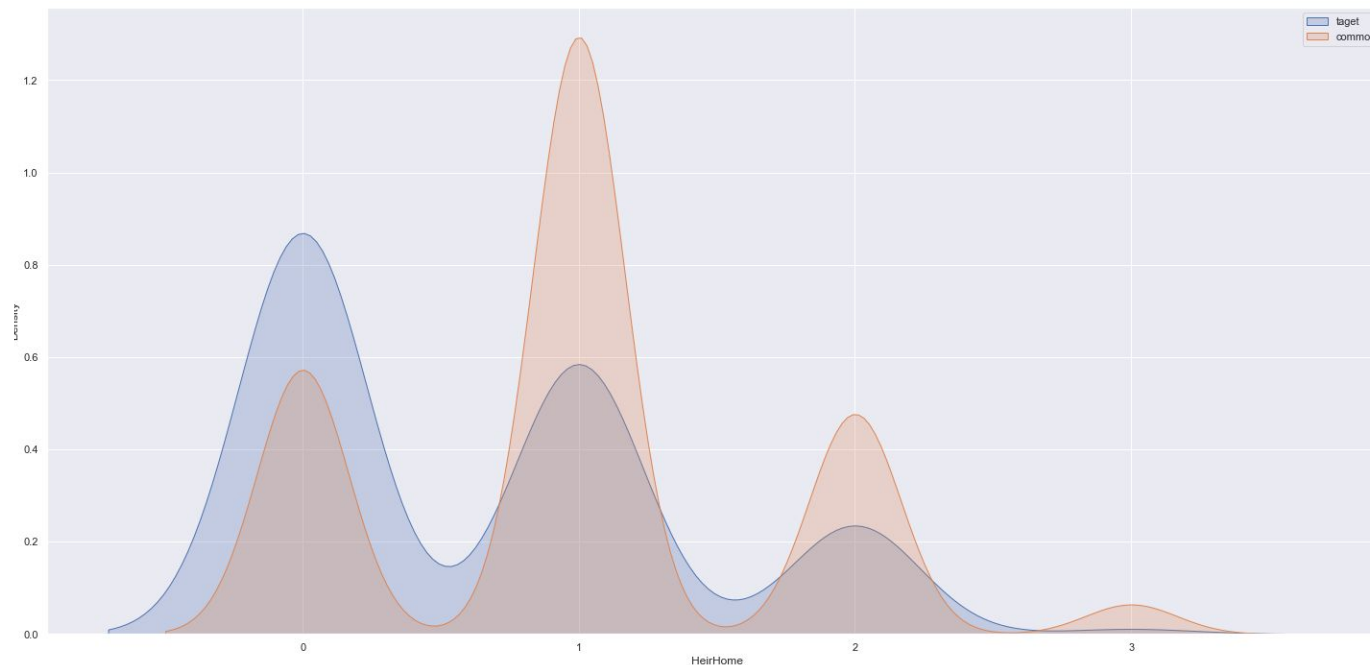


Mediana de Income para target:
64090

Mediana de Income para common:
49991

H2. Clientes com mais filhos são MENOS propensos a aceitar a oferta da campanha

[VERDADEIRO]



Média de HeirHome para target:
0.64

Média de HeirHome para common:
1.01

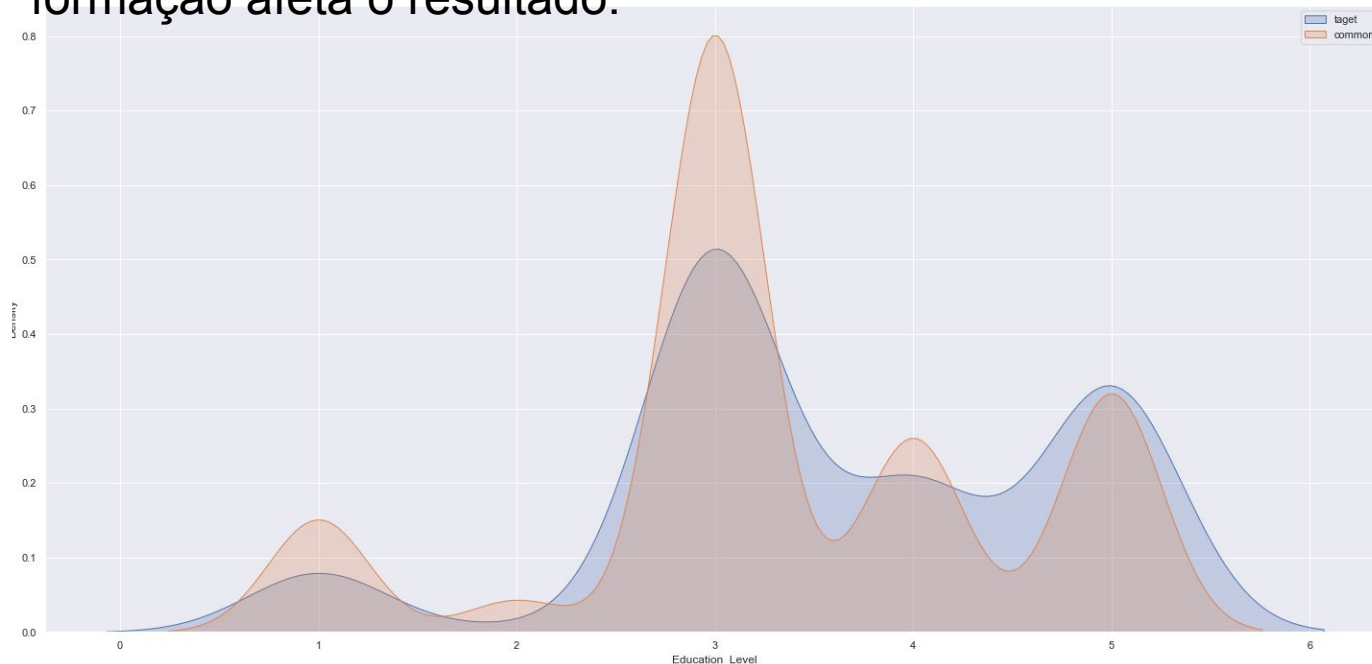
H3. Clientes casados são MENOS propensos a aceitar a oferta da campanha

[FALSO] Clientes sozinhos (Single, Widow, Divorced) são mais propensos a aceitar a oferta.

Marital_Status	All	TRUE	FALSE
Divorced	213	20%	80%
Married	792	12%	88%
Single	444	23%	77%
Together	520	11%	89%
Widow	71	27%	73%
All	2043	15%	85%

H4. Clientes com maior nível de educação formal são MAIS propensos a aceitar a oferta da campanha

[FALSO] As distribuições são bem parecidas, não é seguro concluir que a formação afeta o resultado.



Mediana de Education Level
para target:
3.00

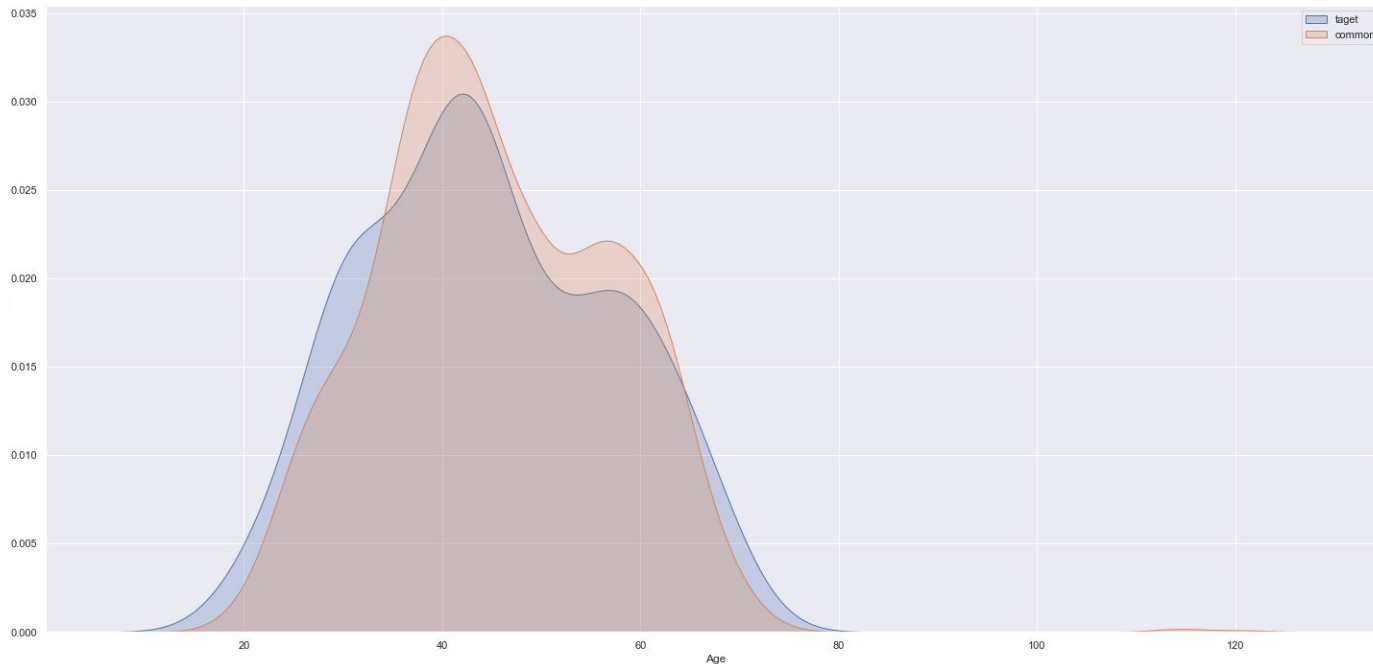
Mediana de Education Level
para common:
3.00

Média de Education Level para
target:
3.61

Média de Education Level para
common:
3.35

H5. Clientes mais velhos são MENOS propensos a aceitar a oferta da campanha

[FALSO] As distribuições são bem parecidas, não é seguro concluir que a formação afeta o resultado.



Mediana de Age para target:
43.00

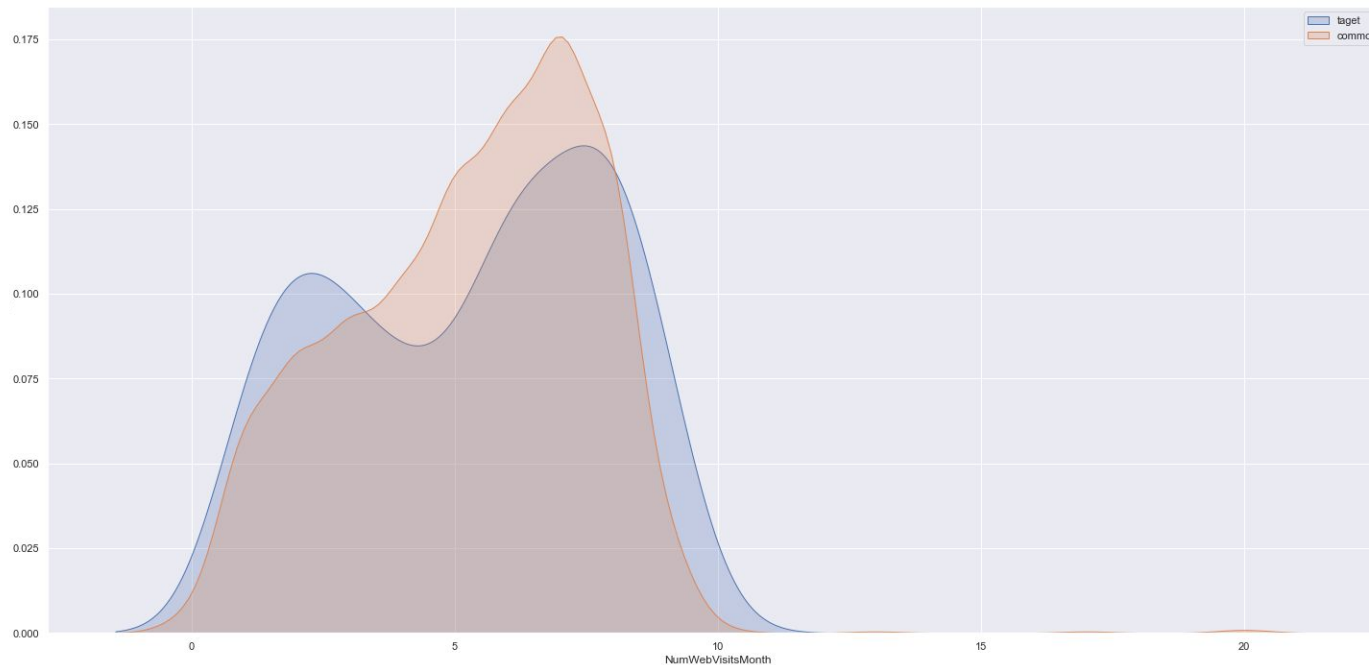
Mediana de Age para common:
44.00

Média de Age para target:
44.50

Média de Age para common:
45.37

H6. Clientes que visitam mais o site são MAIS propensos a aceitar a oferta

[FALSO] As distribuições são bem parecidas, não é seguro concluir que a formação afeta o resultado.



Mediana de NumWebVisitsMonth
para target:
6.00

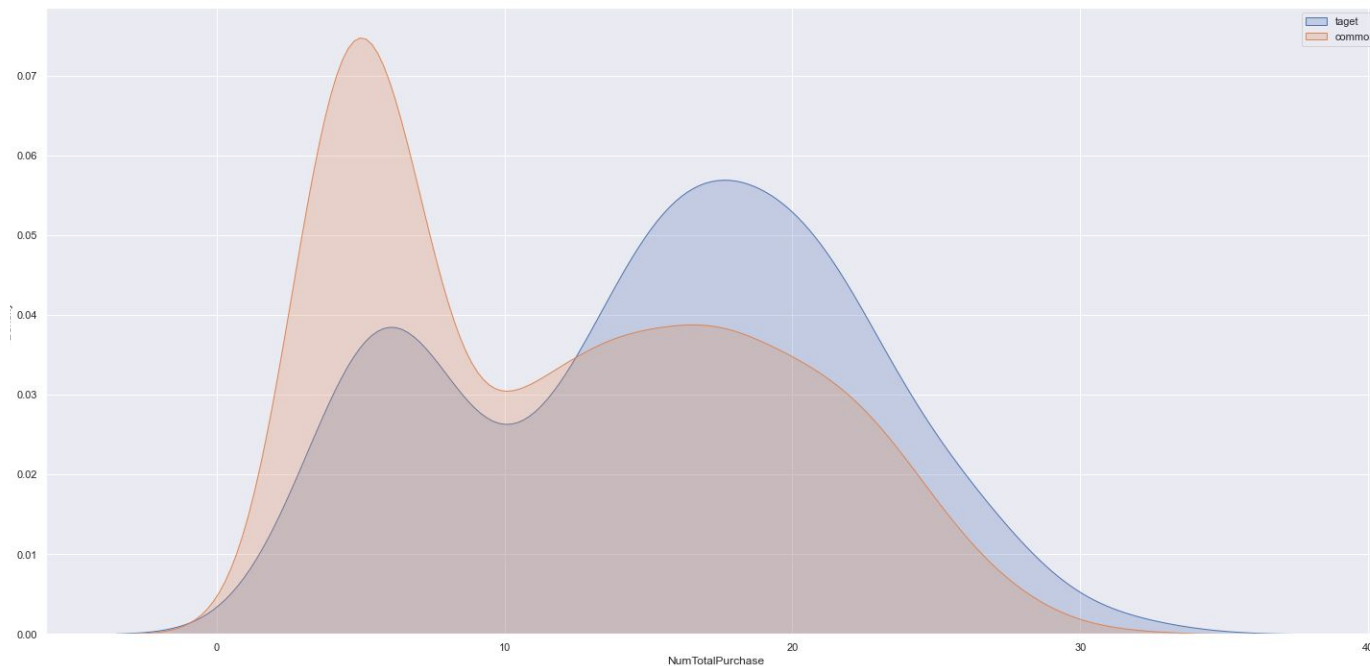
Mediana de NumWebVisitsMonth
para common:
6.00

Média de NumWebVisitsMonth
para target:
5.30

Média de NumWebVisitsMonth
para common:
5.29

H7. Clientes que fazem mais mais compras são MAIS propensos a aceitar a oferta

[VERDADEIRO]



Mediana de NumTotalPurchase
para target:
16.00

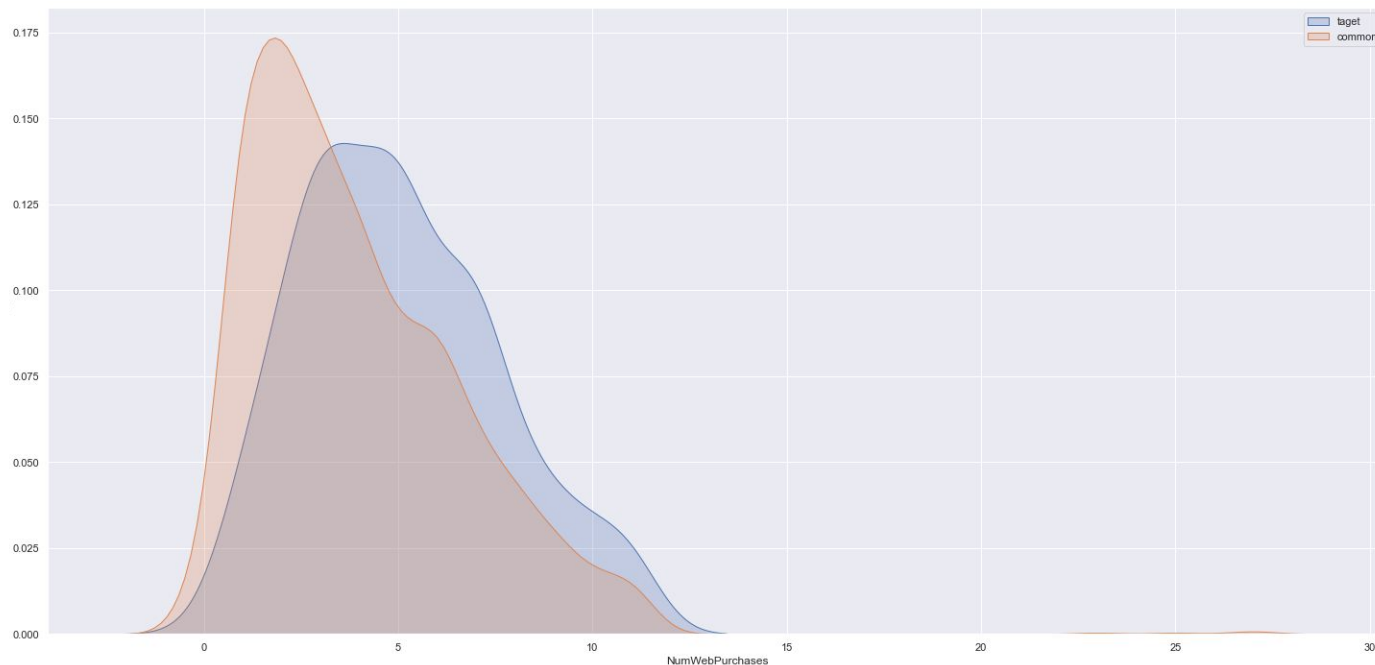
Mediana de NumTotalPurchase
para common:
11.00

Média de NumTotalPurchase para
target:
15.32

Média de NumTotalPurchase para
common:
12.10

H8. Clientes que fazem mais mais compras via web são MAIS propensos a aceitar a oferta

[VERDADEIRO]



Mediana de NumWebPurchases
para target:
5.00

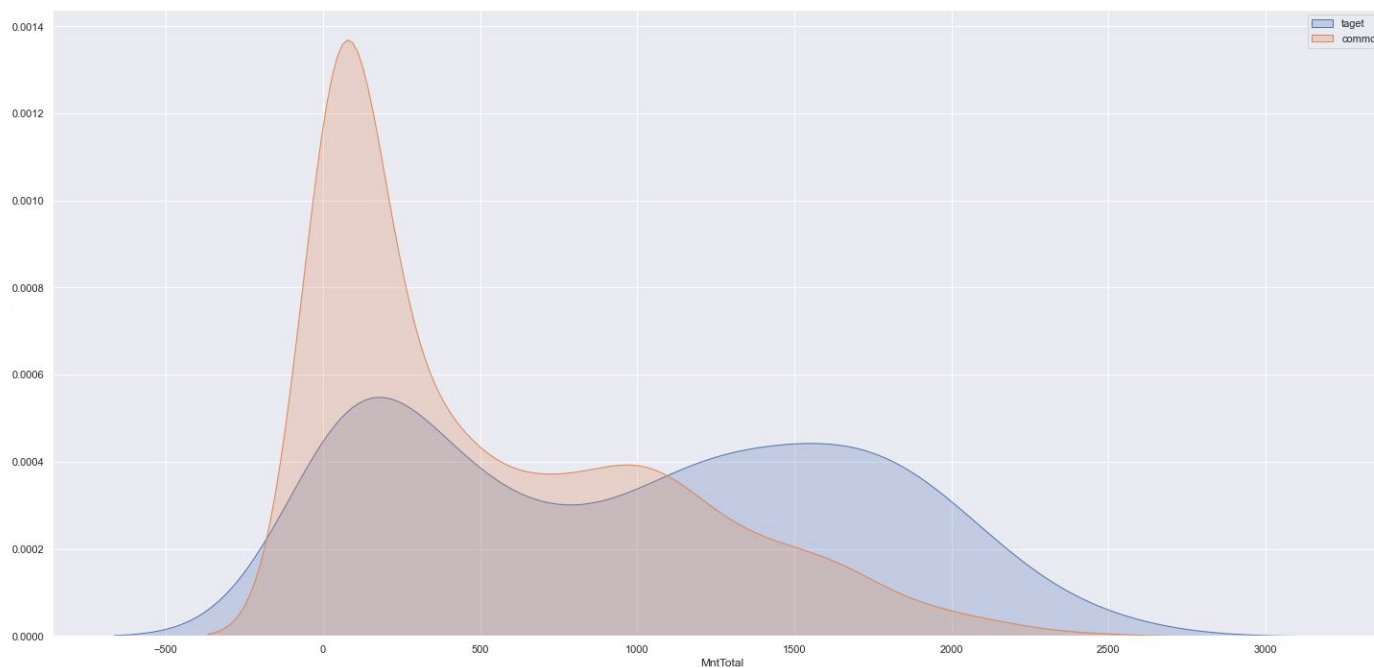
Mediana de NumWebPurchases
para common:
3.00

Média de NumWebPurchases para
target:
5.07

Média de NumWebPurchases para
common:
3.95

H9. Clientes que gastam mais são MAIS propensos a aceitar a oferta

[VERDADEIRO]



Mediana de MntTotal para
target:
1062.00

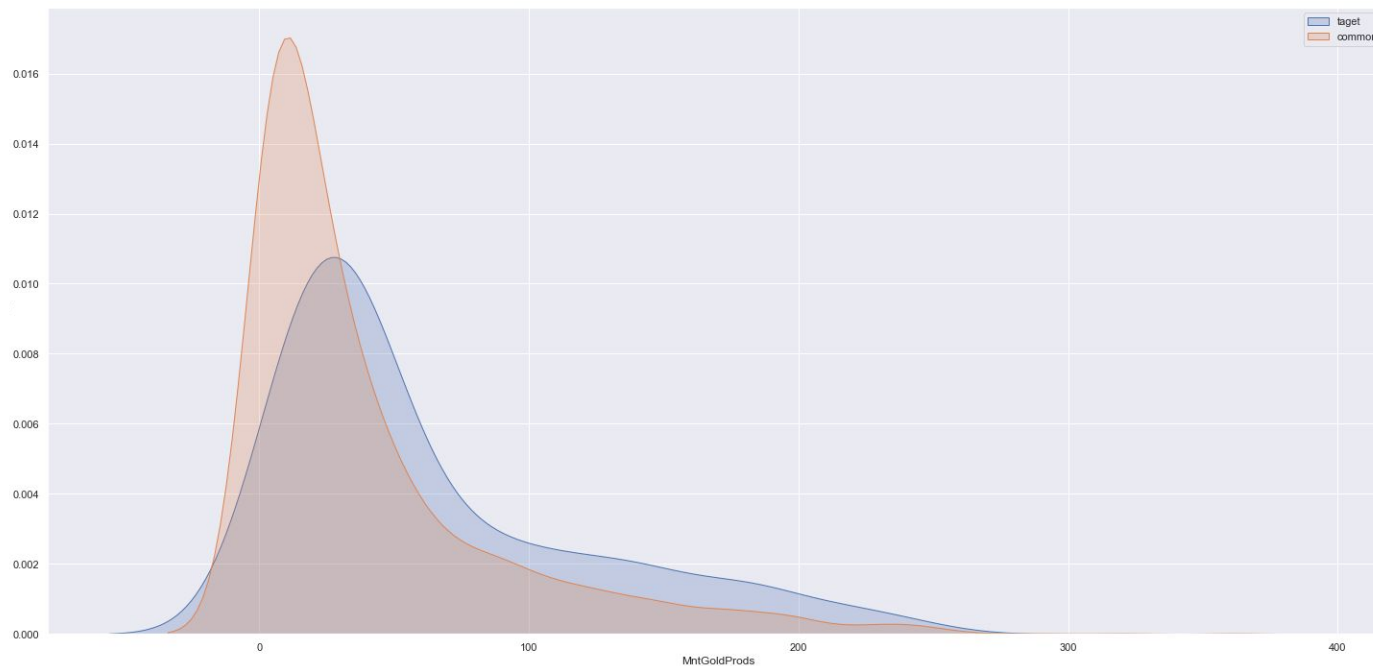
Mediana de MntTotal para
common:
318.50

Média de MntTotal para target:
984.77

Média de MntTotal para common:
542.08

H10. Clientes que gastam mais em Ouro são MAIS propensos a aceitar a oferta

[VERDADEIRO]



Mediana de MntGoldProds para
target:
40.00

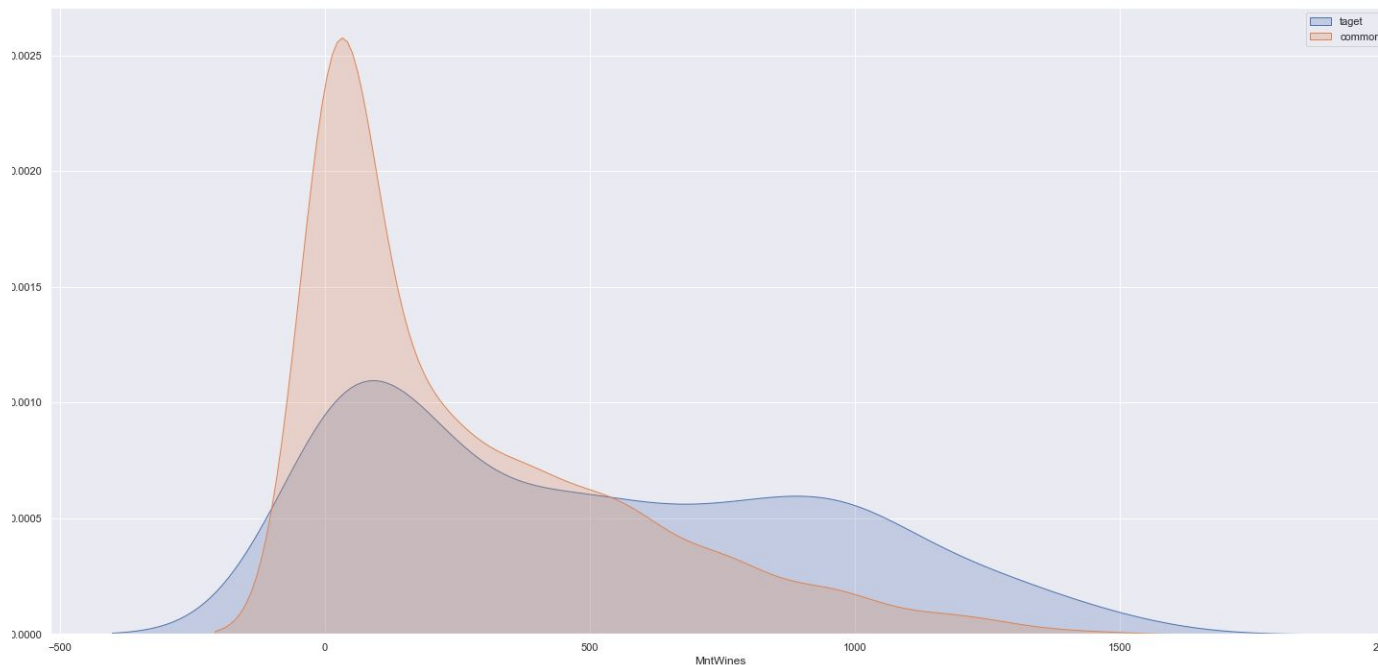
Mediana de MntGoldProds para
common:
21.00

Média de MntGoldProds para
target:
62.97

Média de MntGoldProds para
common:
40.66

H11. Clientes que gastam mais em Vinho são MAIS propensos a aceitar a oferta

[VERDADEIRO]



Mediana de MntWines para
target:
448.00

Mediana de MntWines para
common:
157.50

Média de MntWines para target:
495.82

Média de MntWines para common:
271.47

H12. Clientes que já reclamaram são MENOS propensos a aceitar a oferta da campanha

[FALSO] A reclamação parece não afetar o aceite da oferta.

Complain	All	TRUE	FALSE
FALSE	2023	15%	85%
TRUE	20	15%	85%
All	2043	15%	85%

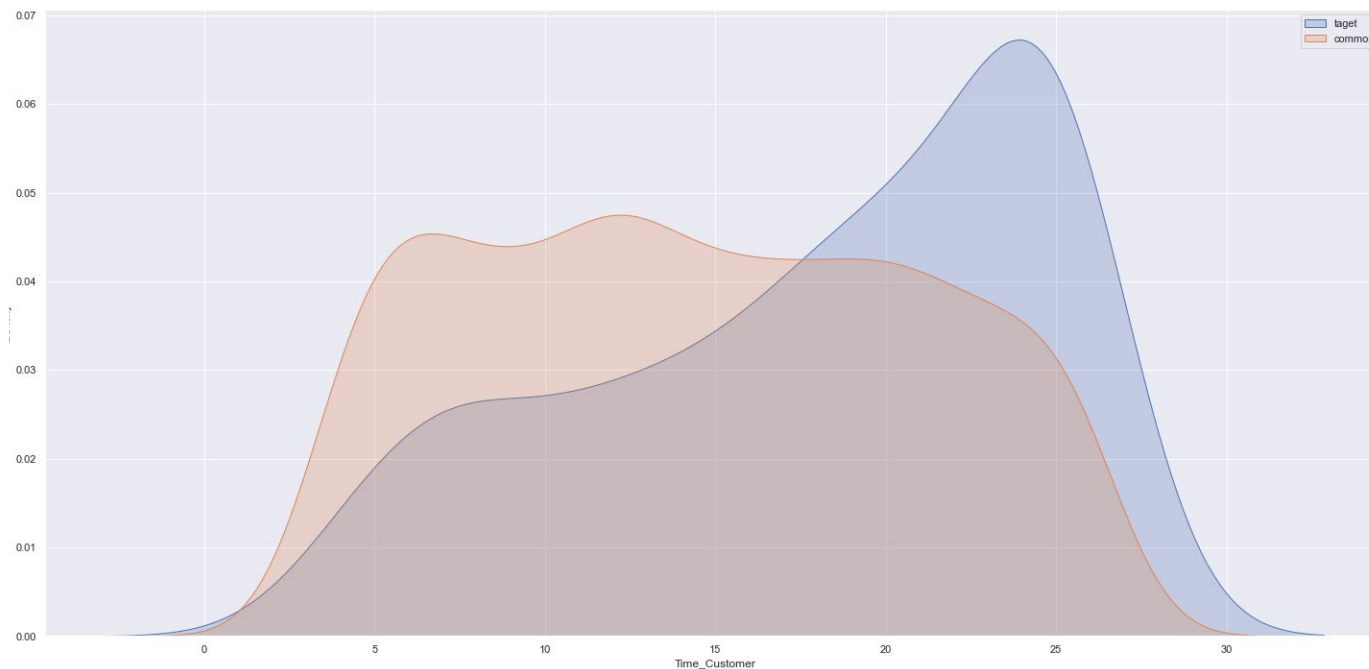
H13. Clientes que compraram em outras campanhas são MAIS propensos a aceitar a oferta da campanha

[VERDADEIRO]

AcceptedCmpBefore	All	FALSE	TRUE
FALSE	1613	92%	8%
TRUE	430	60%	40%
All	2043	85%	15%

H14. Clientes mais antigos são MAIS propensos a aceitar a oferta da campanha

[VERDADEIRO]



Mediana de Time Customer para
target:
19.43

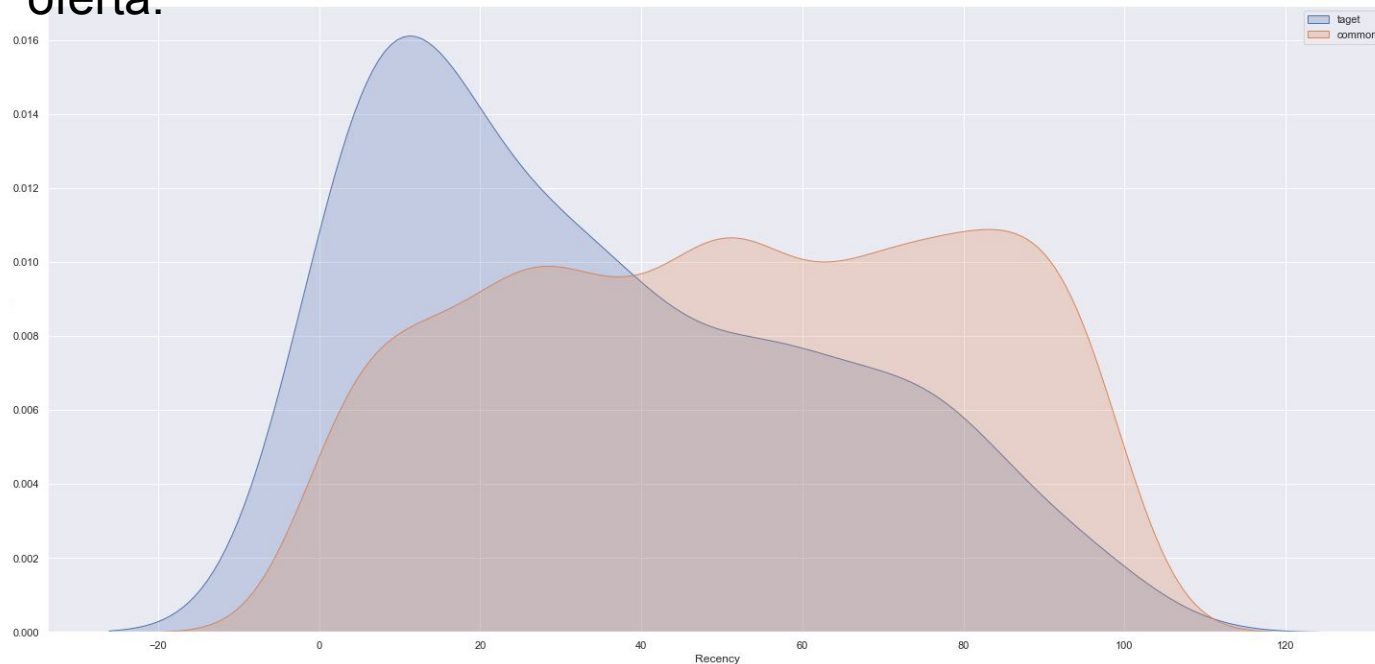
Mediana de Time Customer para
common:
14.17

Média de Time Customer para
target:
18.02

Média de Time Customer para
common:
14.47

H15. Clientes que fizeram compras mais recentes são MENOS propensos a aceitar a oferta da campanha

[FALSO] Clientes com compras mais recentes são MAIS propensos a aceitar a oferta.



Mediana de Recency para
target:
30.00

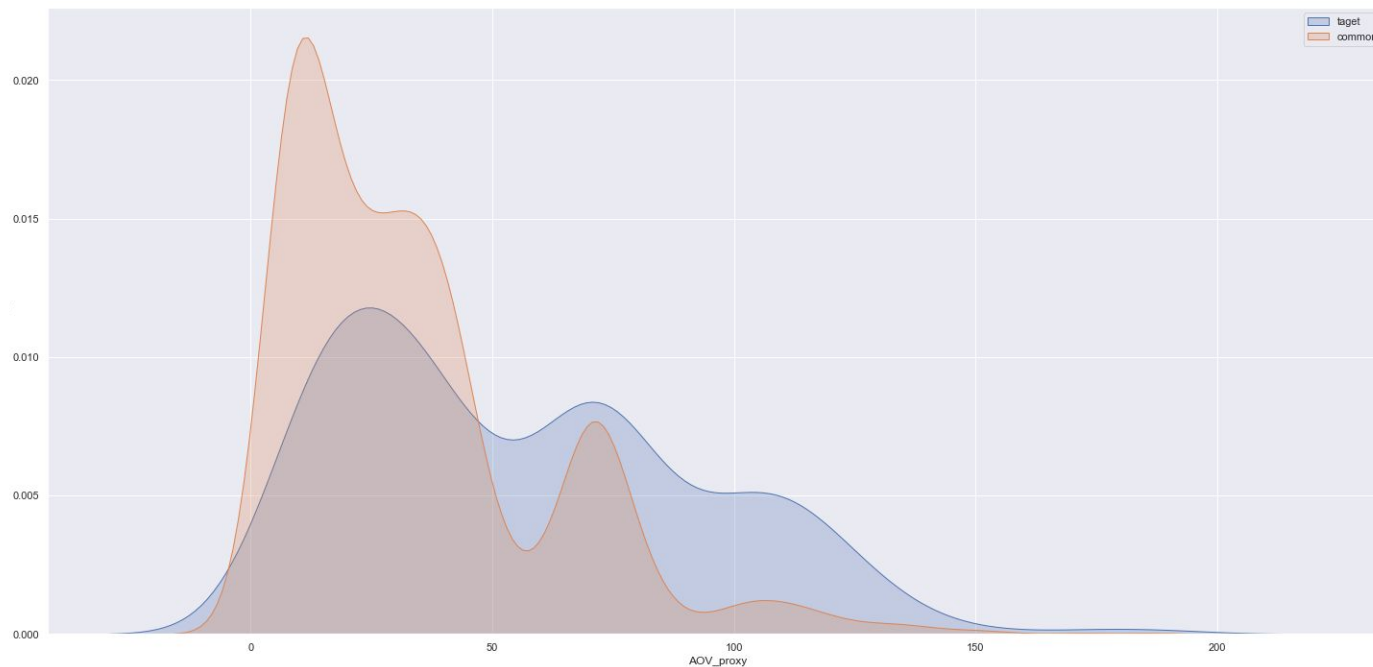
Mediana de Recency para
common:
52.00

Média de Recency para target:
35.16

Média de Recency para common:
51.46

H16. Clientes com maior AOV são MAIS propensos a aceitar a oferta da campanha

[VERDADEIRO]



Mediana de AOV proxy para
target:
47.27

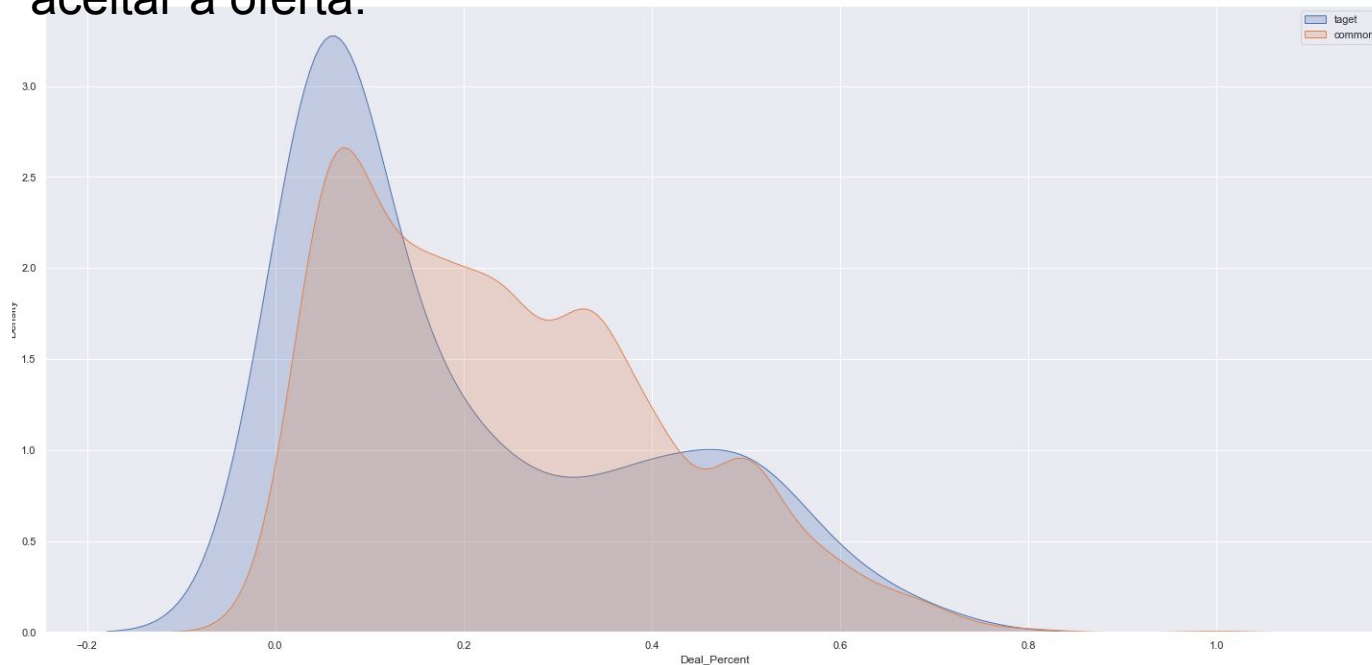
Mediana de AOV proxy para
common:
27.90

Média de AOV proxy para
target:
56.25

Média de AOV proxy para
common:
35.17

H17. Clientes que compram mais em promoções são MAIS propensos a aceitar a oferta da campanha

[FALSO] Clientes que compram MENOS em promoções são MAIS propensos a aceitar a oferta.



Mediana de Deal Percent para
target:
0.12

Mediana de Deal Percent para
common:
0.22

Média de Deal Percent para
target:
0.21

Média de Deal Percent para
common:
0.25

Insights

Clientes de MAIOR RENDA tendem a comprar mais Vinho e Carne e menos Produtos de Ouro.



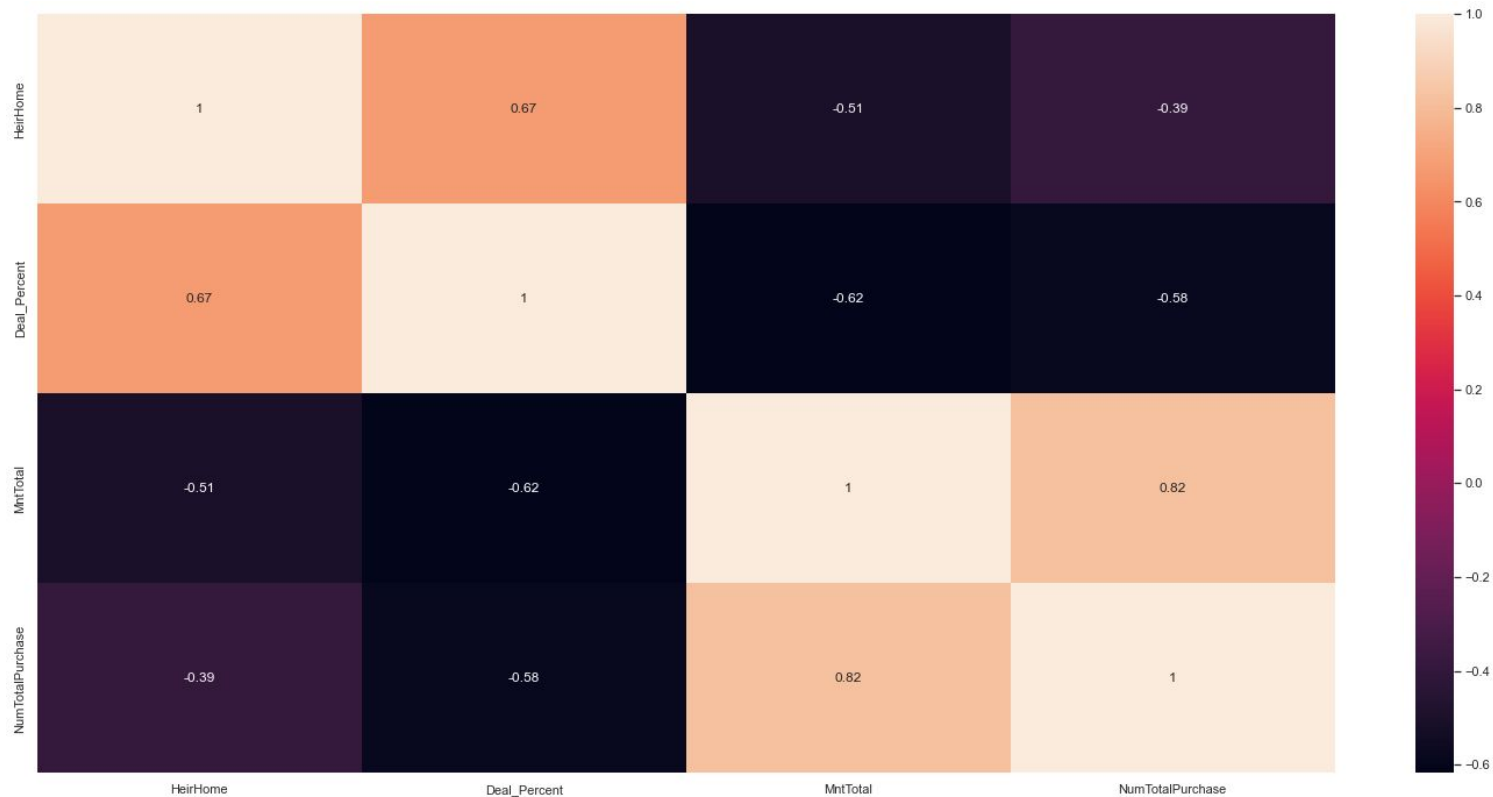
Insights

Clientes que compram mais Vinho parecem comprar mais via catálogo e aproveitar menos promoções



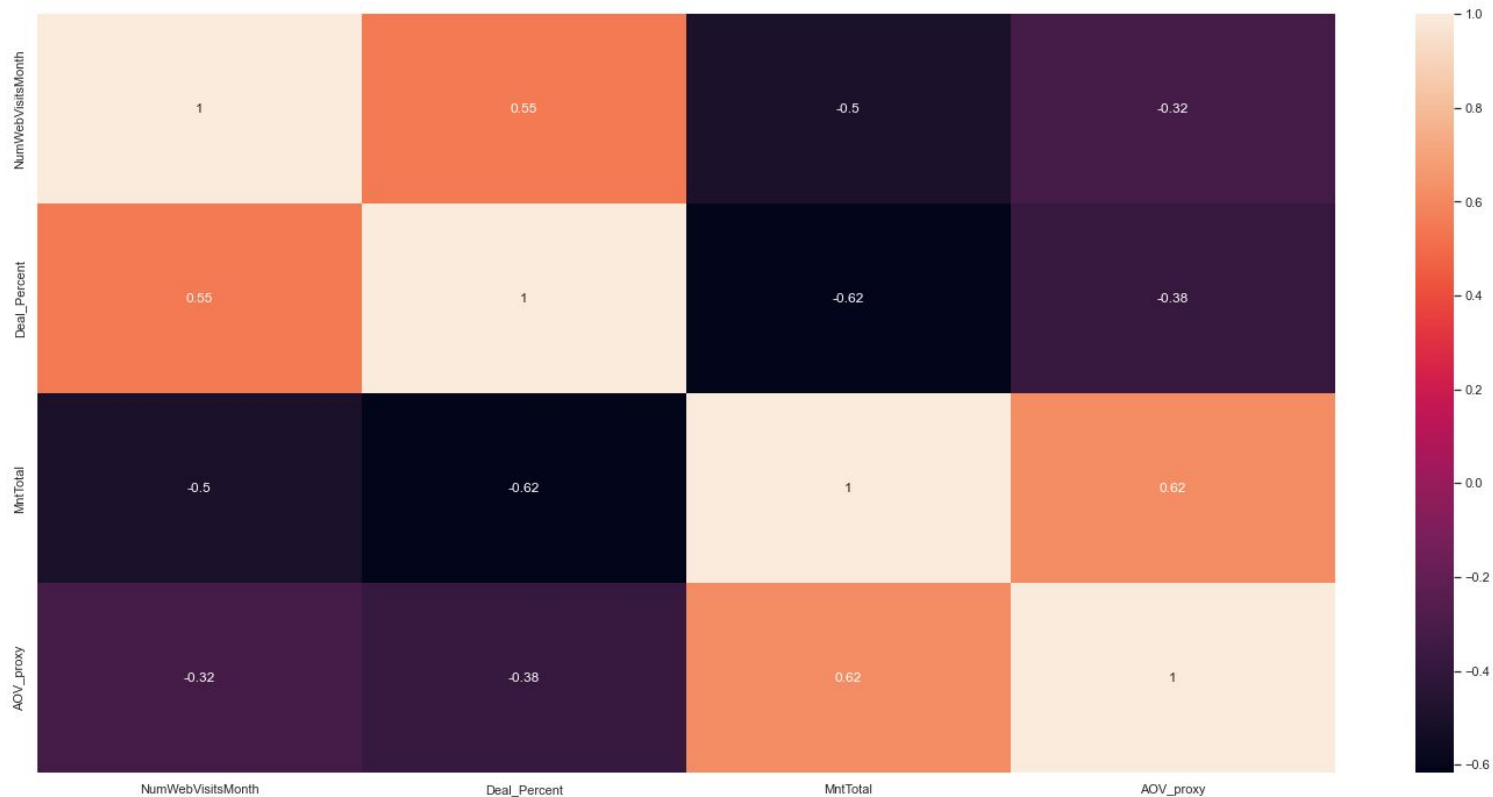
Insights

Clientes com mais crianças tendem a comprar menos e aproveitar mais promoções



Insights

Clientes que compram mais pela Web e que fazem mais visitas ao site aproveitam mais promoções e gastam menos



Possíveis abordagens

Uma forma de otimizar a próxima campanha é direcionar os gastos para alguns grupos evidentes de clientes:

- Clientes com AOV mais alto, geralmente compradores de vinho e carne via catálogo
- Clientes mais antigos, frequentes e/ou que converteram em outras campanhas, provavelmente quem faria parte de um programa de fidelidade
- Clientes que compram via web, mas sem desconto
- Evitar clientes que fazem compras do dia-a-dia para casa, como casados, com família e que compram mais em promoções