

Análise de Dados empresa
E-commerce

Cenário

Uma empresa indiana de e-commerce quer fazer uma análise do seu desempenho do ano de 2018 e 2019 com o intuito de criar um perfil de cliente mais preciso, mas para isso é preciso investir na empresa e o mesmo não tem certeza se vale a pena realizar o investimento. Por isso que descobrir como está a saúde do negócio quando a lucro para realizar o investimento.

Plano de análise

Qual é a meta da empresa?

3% de aumento do faturamento de um ano para o outro.

Qual o objetivo da análise?

Descobrir a saúde monetária da empresa e responder se vale a pena investir.

De onde obterei os dados?

Do site Kangle

link: <https://www.kaggle.com/datasets/benroshan/ecommerce-data?select=List+of+Orders.csv>

Quais serão as planilhas analisadas?

List_orders,
Orders_details,

Quais serão as ferramentas utilizadas?

Phyton e biblioteca pandas, numpy, matplotlib e seaborn.

Qual será forma a fornece a credibilidade a qualidade dos dados?

1. Limpeza inicial dos dados.
2. Testes estatísticos referente a distribuição dos dados.

Sumário

- i. Teste estatístico de distribuição
- ii. Análise por questões
 - i. Qual o valor que representa 3% de crescimento de lucro ao ano?
 - ii. Quais são os períodos mensais do ano de 2018 e 2019 que batem a meta?
 - iii. Qual a proporção dos top 10 clientes em relação aos demais do ano 2019?
E qual impacto no lucro?
 - iv. Top 7 estados que aparecem no ano de 2018 e 2019?
 - v. Como está o lucro de produtos por categoria no ano de 2018 e 2019?
- iii. Descobertas e sugestões

Teste Estatístico de distribuição

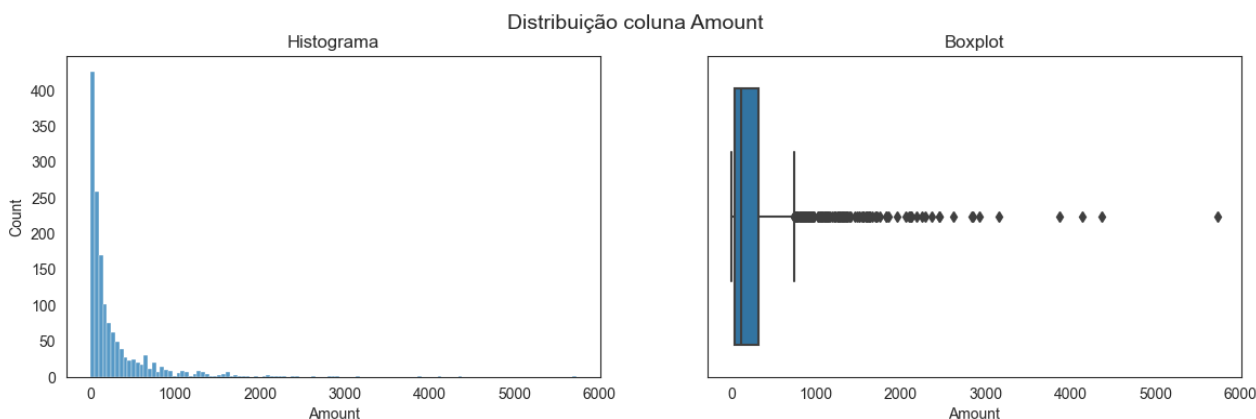
Sobre o teste:

O mesmo é feito somente com as variáveis numéricas, pois busca de descobrir a normalização dos dados. Isso é se os dados estão distribuídos da forma que deviam.

Por que é importante?

Evita erros de compreensão tanto nas amostras como nas perguntas estatística que podem ser feitas posteriormente.

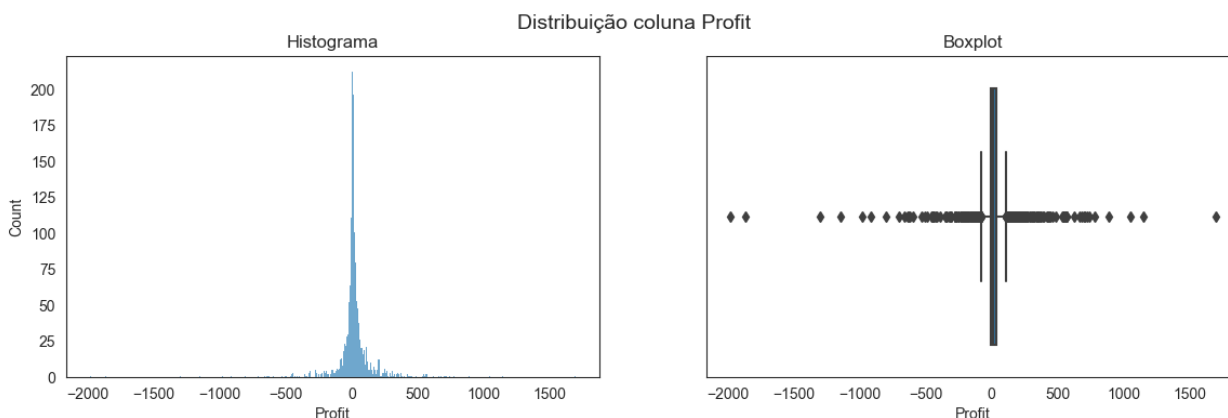
Coluna amount



Sobre:

- É visível que se tratando somente da distribuição dos dados o valor inicial de outlier é de 1000.
- Temos uma concentração de valores em grupos com frequências entre 50 e 600 unidades.
- Deve-se ter cautela com os valores referenciando os outliers, pois os mesmos têm pouca expressividade quando diz respeito a proporção total.

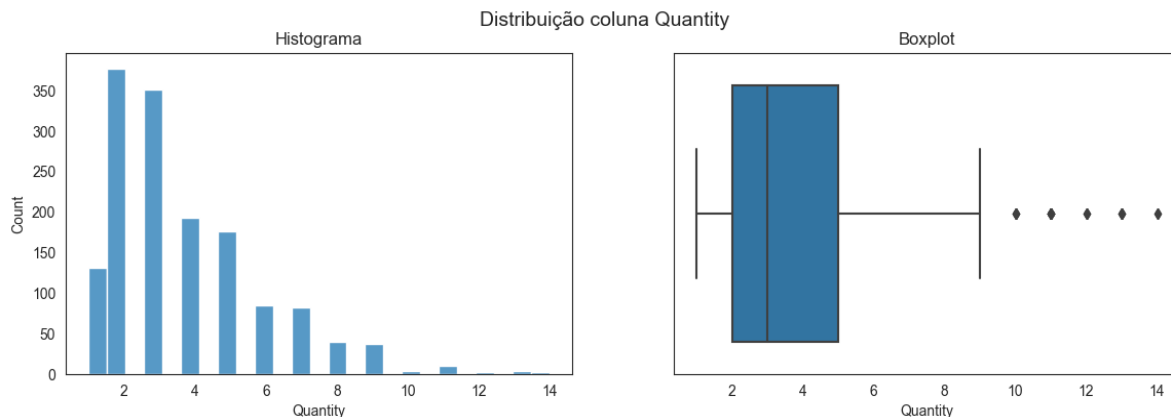
Coluna profit



Sobre:

- O mais interessante são os valores negativos que podem indicar um problema na qualidade dos dados ou que a empresa de fato passou por períodos sem lucro algum.
Obs: A coluna já veio com o seu nome indicando lucro, pois isso estou considerando esse aspecto.
- Considerando que isso pode ser um erro na qualidade dos dados, talvez uma opção seria contabilizar os dados a partir de um valor mínimo negativo ou do valor zero.

Coluna quantity



Sobre:

- i. Esse espaçamento entre as barras no histograma indica que temos uma variação de valores fixo, isso é não tem números decimais, o que não é problema já que a coluna se refere a quantidade de produtos.
- ii. Podemos considerar abaixo do número 9 os outliers e por agir com cautela em decisões que levam em conta cliente que obtiveram mais de 9 produtos, já que temos uma pouca concentração dos mesmos.

Análise de desempenho por questões

Do que se trata?

De questões feita em relação ao desempenho da empresa, com o objetivo de descobrir um perfil de cliente mais preciso.

1. Qual o valor que representa a meta de 3% de crescimento de lucro ao ano?

O valor que representa 3% da meta no ano de 2018 e de R\$-53.00 negativo, pois a maior parte do ano a empresa trabalhou em dívida. E no ano de 2019 R\$ 772.62 aqui o valor já é positivo, pois a empresa se recuperou.

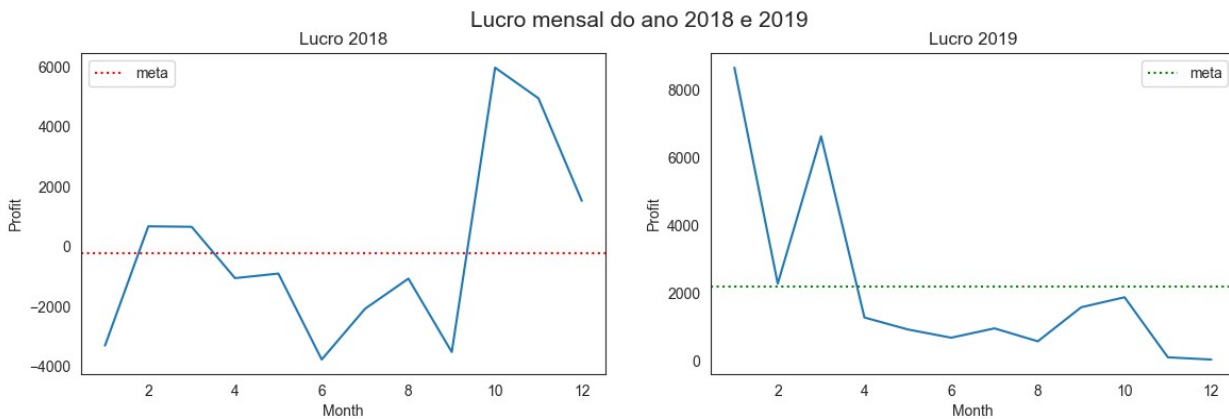
	lucro total	valor 3% anual	Meta Anual + 3%	Meta Mensal + 3%
2018	-1799.0	-53.00	-1745.03	-145.41
2019	25754.0	772.62	26562.62	2210.51

Figura 1: Tabela de valores

Porque os valores da meta são tão baixos?

Essa empresa indiana é microporte, ou seja, entre tem entre 1 e 11 funcionários apenas. Sua atuação é pela internet e teve um início não rentável, como será explicado nos períodos mensais.

2. Quais são os períodos mensais do ano 2018 e 2019 bateram a meta?



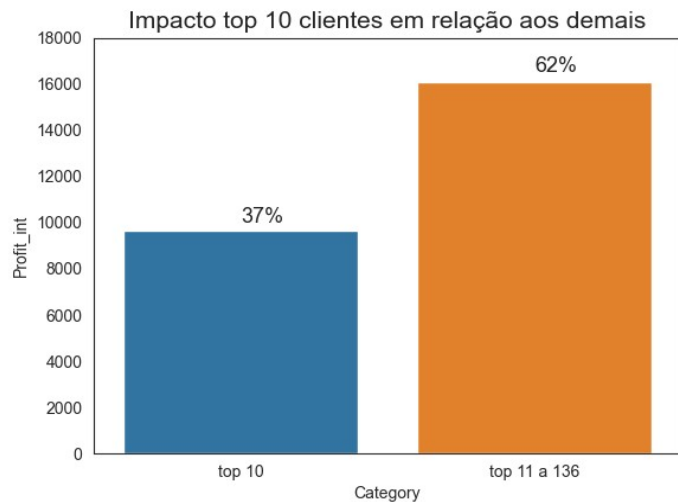
No ano de 2018 temos 84% dos meses negativos ou abaixo de um valor ideal para ser considerado lucro, diante disso os meses que foi obtido um valor ideal ou bateu a meta e outubro, novembro e dezembro. Já no ano de 2019 temos 76% dos meses abaixo da meta, mas apesar disso estão com valores positivos diferente do ano anterior.

3. Qual a proporção dos top 10 clientes em relação os demais? E qual impacto no lucro?

Esse pequeno grupo de 10 clientes tem um impacto financeiro referente a 37% de todo o lucro. E os 126 clientes restantes dividem os 62% do lucro obtido pela empresa.

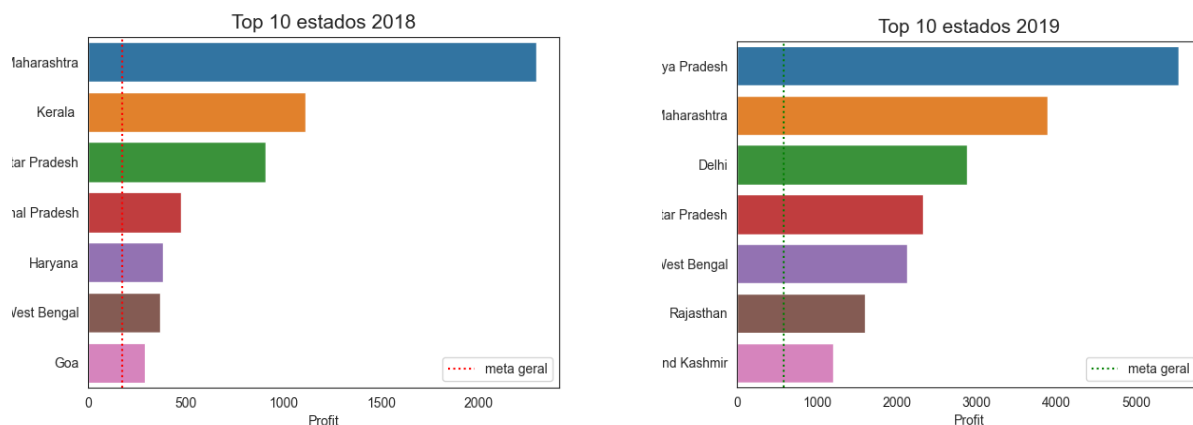
Vale apenas criar uma partição só para esse pequeno grupo?

Sim, pois tem um grande impacto no lucro. E não, pois e menos de 10% do número total de cliente e vai exigir uma maior complexidade.



4. Top 7 estados quem aparecem em 2018 e 2019?

O valor percentual referente a 3% de crescimento dos top 7 estados no ano de 2018 e de R\$ 175,26 e no ano de 2019 R\$ 586,86.



Somente 3 estados reaparecem no ano de 2019, mas todos eles forneceram um lucro muito maior. Além de estarem melhores posicionadas do que no ano anterior.

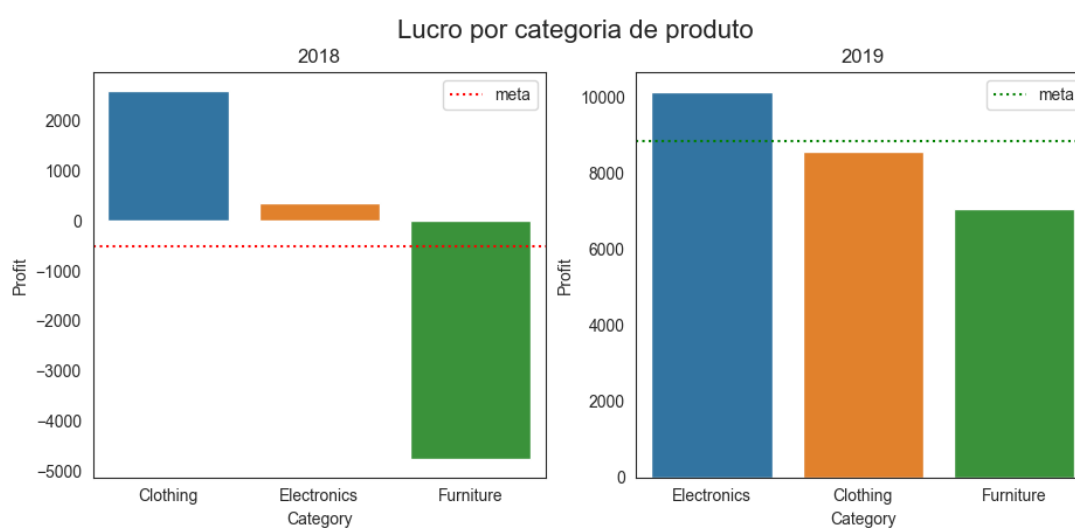
Estados: Maharashtra, Uttar Pradeshe e West Bengal.

Porque temos estados com lucro positivo no ano de 2018?

Pois, mesmo tendo o lucro total negativo ainda tivemos 36% do total de estados que obtiveram um valor acima de 300 ou positivo. O que se consideramos somente os top 7 conseguimos um valor positivo.

5. Como está o Lucro de produtos por categoria no ano de 2018 e 2019?

O valor percentual referente a 3% de crescimento por produto e de R\$-518,00 no ano 2018 e de R\$ 8841,62 no ano de 2019.



No ano 2018 a categoria Clothing e Eletronics ambas conseguem bate a meta, mas o impacto negativo da categoria furniture foi tão grande que suprimiu os valores positivos. Ocasionalmente em uma meta mínima mas negativa e grande ano de perda.

Já no ano de 2019 todas as categorias tiveram valores positivos muito expressivos (6x maior), o que gerando uma meta muito acima do prevista e ainda assim a batendo.

Descobertas

1. Quanto a lucro o ano 2018 o resultado foi péssimo. A ponto do lucro total ser negativo, mas em 2019 tivemos uma grande recuperação.
2. Quanto aos produtos a maior parte da concentração do lucro advêm que poucos produtos vendidos para muitos clientes (especificamente menos de 9 produtos por cliente).
3. Quanto a distribuição de lucro mensal temos um grande crescimento exponencial no mês 9 de 2018 até o mês 4 de 2019.
4. Quanto à proporção de clientes em relação ao lucro foi concluído que os top10 clientes representam 40% de lucro juntando o ano de 2018 e 2019.
5. Quanto a distribuição de lucro por estado, apesar do lucro total negativo em 2018 temos mais de 36% do número de estados com lucro positivo ou acima de R\$300,00.

Sugestão

Esperar obter uma consistência de lucro mensal e de produtos no próximo ano de pelo menos 6 meses, principalmente os produtos de categoria furnitere.

Mas caso queira realiza o investimento já, recomendo realiza outra análise de produtos específicos da categoria funiture já que teve um grande impacto no lucro e um acompanhamento de lucro trimestral, se possível tente realiza uma previsão de lucro.