```
-recomendo super
                                                                                                                                                                                          confiâvel parceiro
                      parabéns 2
data reodu
                                                                                                                                                                                                                   i tudo perfeito
                                                                                                                                                                                                                                          cortina policina
                                                                                                                                                                                                                                     tuncionou o silino
                                                                                                                                                                                                                                         Trace President for Tailor
               desor a
                                                                                                                                                                                                                                                ótima compra g
               E
                                                                                                                                                                                                                                                        r aguardando
                                                                                                                                                                                                                                          gerasboa compra
                                                                                                                                                                                                                                   recomento
          antes prazo
                                                                                                                                                                                                                   custo beneficio
                        bem boa boa
                                                                          ótimo ótimo pecico capa de la cap
```

```
regulação Tormecedor que la
Total 100 cm
  entrega ok
  frete desproporcion
adorei atenção
  lannister
 problema
mota au gostei
  entregue
entregue
 super rápida
# entrega antes
original oricincia otimo
entrega prazo
rapida ***
  satisfeito
 entrega rapida
desproporcional frete ut
babet ballede
 chegou pee 👸 i
  errado o
 or pulsuand 2
    bom
  rapidez
      Tapete
 tudo contorne
  recebi 🚉
 regular 👭 🖔
  troca recomendad
     eficiente
```

```
otimo otimo
    dez simples
certeza ....
 atendimento ....
prazo project months
 otimos produto nota
  Lagar loja
  vendedor,""
  maravilhoso
 ainda recebi
 Otudo boa a
    >= lindo
   legal -- 3
 9 pontual
  recent produt
entregue prazo
perfeito
  show empresa
2 comprer
porte chegou
     otima ___

    entrega perfeita

 rápido *******
 praticidade praticidade
```

--- luminária





400 pessoas na Olist e as 400 palavras mais usadas nos reviews \o/

Metodologia Descritivo

- Para este desafio e qualquer outro em negócios, busco me basear em 3 frentes; utilizar a metod. data science unido à minha experiência em business e o método científico em si.
- Meu raciocínio fluirá desta forma:
 - 1. Entendimento do Negócio
 - 2. Problemas a serem resolvidos
 - 3. Requisitos, Levantamento e entendimento dos dados
 - 4. Transformação dos dados
 - 5. Modelagens e Hipóteses
 - 6. Validação e Estratégia
 - 7. Deploy de algum modelo
 - 8. Avaliação/Feedback

Os dois notebooks contendo a estruturação das análises estão no mesmo diretório desta apresentação.

Entendimento do Negócio

Estamos interpretando o negócio como sendo a área de atuação no time de Marketing

 Acreditamos que a necessidade seja gerar valor com dados estruturados e não estruturados de n fontes, para suportar experimentos e validar hipóteses

Problemas a serem resolvidos

- Ingerir os dados do banco afim de gerar valor e insights que possam suportar experimentos na área de marketing focado em incremento de conversão e/ou receita.
 - Exemplo:
 - Aumentar a satisfação dos clientes aumentaria a conversão ?
 - Qual categoria de produto é a mais problemática ?
 - Poderíamos prever clientes insatisfeitos e realizar testes com eles ?
 - Que ganho teríamos acelerando o processo de descoberta de clientes insatisfeitos ?

 Seguimos a primeira dica do teste e comecei pela análise dos comentários, depois de entender o data schema, fiz alguns tratamentos, entre eles:

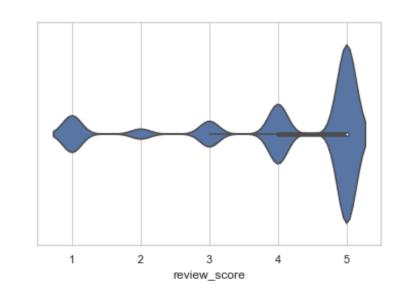
- Comentários e títulos preenchidos em campo de data;
- Gravar comentarios e títulos de review nos campos corretos;
- Etc;
- Dados filtrados, iniciamos a exploração;

• Separamos a análise em 2:

- Clientes com produto entregue sem comentário no review
- Comentarios variados.

Clientes com produto entregue e sem comentário no review

■ A maior parte dos reviews atingiu nota máxima.



■ 76% dos reviews atingiram nota 4 ou superior.

Nota	Percentual
5	57.42%
4	19.2%
1	11.8%
3	8.2%
2	3.2%

- 88% dos reviews está sem título
- 56% dos reviews não tem comentário

Como seria de se esperar, menos dias até a entrega resultam em score maior.

Relação entre dias até entrega e Score do Review



dias_ate_entrega
22.254663
16.850394
13.540489
11.600768
10.135684

- Compras que receberam review com nota máxima levam em média 10 dias para serem entregues, enquanto aquelas que tiveram a pior nota são entregues em quase 30 dias.
- São apenas 3 dias de intervalo na entrega que separam os reviews nota 5 dos nota 3.

Clientes com produto entregue e sem comentário no review

As compras que são recebidas antes da data estimada tendem a receber um review melhor.

no eixo y, quanto mais dias o produto é recebido antes do prazo. maior tende a ser o score do review

Relação entre data estimada, entrega e Score do Review



 O prazo de entrega médio passado para todos os clientes é de 23 dias.

	review_score	delta_estimado_entregue
0	1	-2.167862
1	2	-8.057743
2	3	-10.974599
3	4	-12.371973
4	5	-13.290389

■ Podemos perceber acima que encomendas que chegam até 12 dias antes do prazo recebem score 4 ou maior.

■Conclusão:

- clientes que recebem o produto em até 11 dias depois da compra ficam satisfeitos
- clientes que recebem o produto em até 11 dias antes do prazo acabar ficam satisfeitos.
- Poderíamos lançar uma hipótese e teste A/B para Growth:
- Aumentar a data de entrega estimada, aumentaria o score dos reviews ? Assim, produtos com reviews mais altos teriam melhor conversão ?



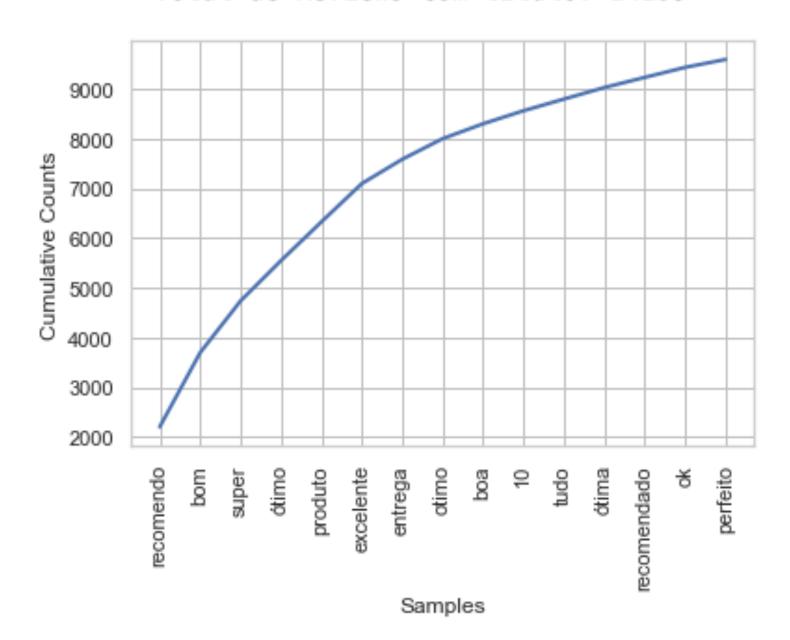
Análise geral dos comentários

- Separamos a análise textual dos reviews olhando para os títulos e para o conteúdo:
- + de 70% dos reviews com título, levam essas 15 palavras:

'recomendo', 'bom', 'super', 'ótimo'....

TÍTULOS MAIS COMUNS NOS REVIEWS

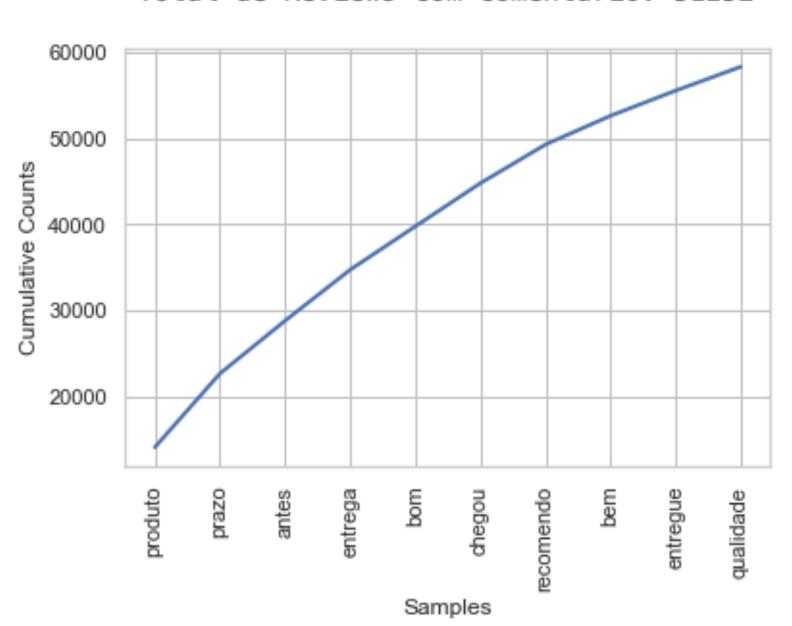
Total de Reviews com título: 14189



- Estas são as 10 palavras mais citadas no conteúdo dos comentários:
- 'prazo', 'antes', 'bom', 'entrega'...

PALAVRAS MAIS COMUNS NOS REVIEWS

Total de Reviews com comentário: 51252

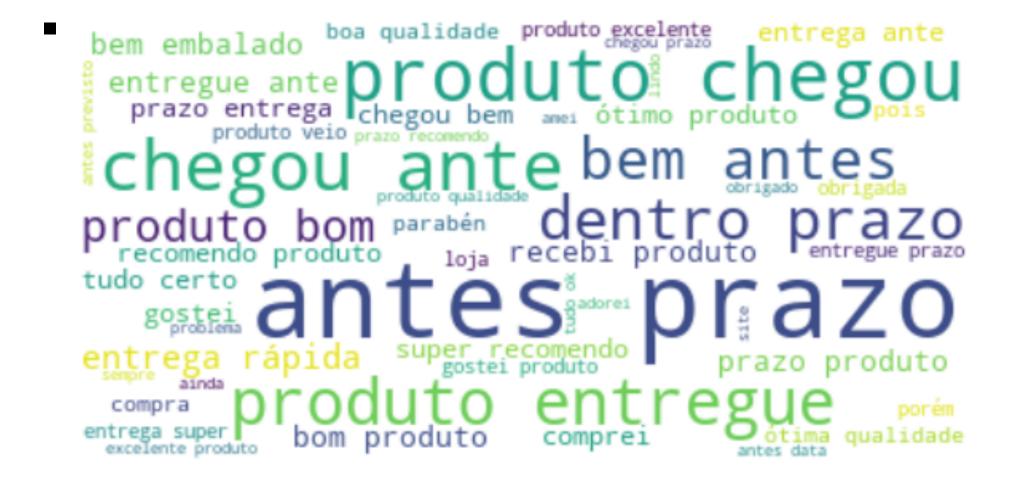


olist

Análise dos Reviews

Análise geral dos comentários

Wordcloud para melhor percepção dos comentários:



- Fizemos uma análise também dos comentários negativos (Clientes insatisfeitos somaram 15% das análises)
- As principais causas se referiram à:
 - Pedido não entregue no prazo...
 - Recebimento do pedido de forma parcial.
 - Esta análise está no notebook
- Foi criado um algoritmo para classificar os reviews em: Review Positivo e Review Negativo. A apresentação deste está no final.

Análise de Produtos

- Foram separadas as vendas com score abaixo de 3
- As vendas foram agrupadas por categoria e por estado
- Foram contadas as categorias de produto por estado com Score baixo

- Categorias mais propensas a insatisfação:
 - Cama Mesa & Banho
 - Acessórios de Informática
 - Moveis e Decoração
- SP, RJ e MG são os estados com mais reclamações

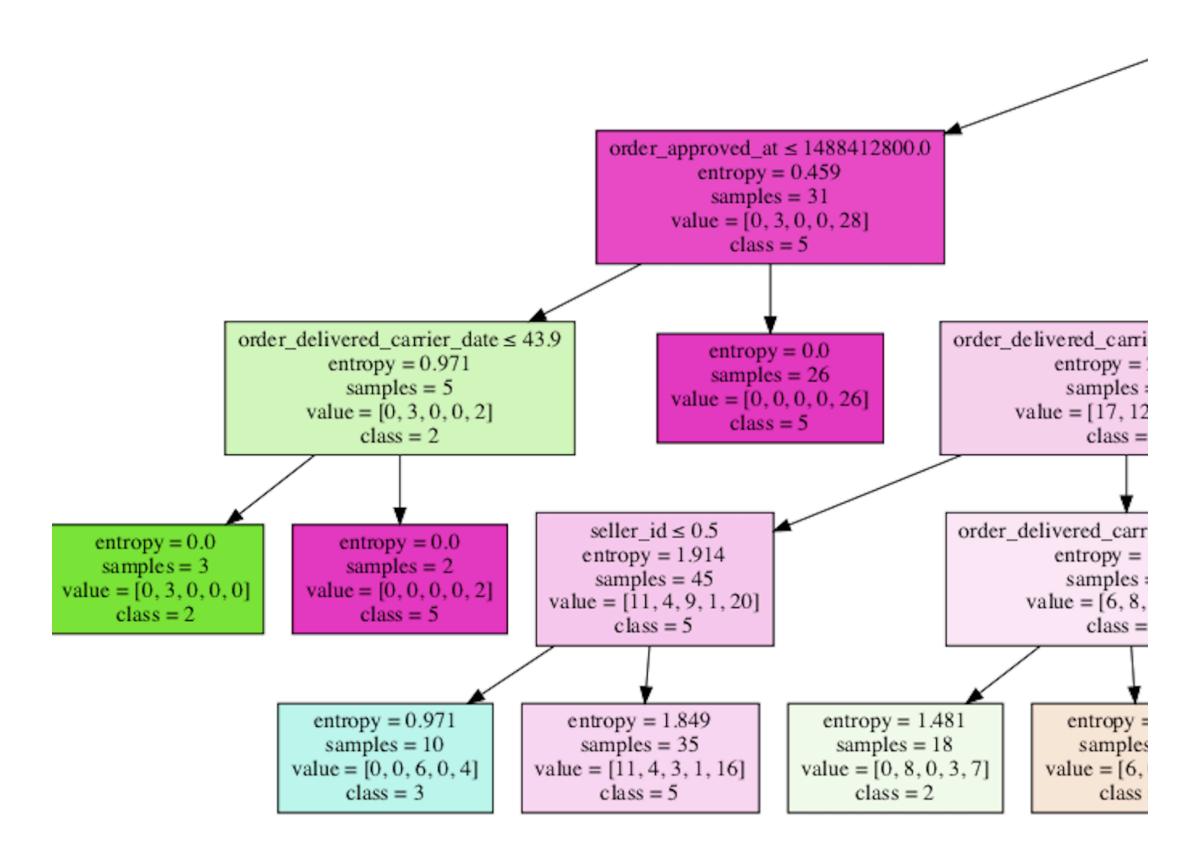
- Sabemos que Móveis e acessórios são itens grandes, e que entrega atrasada gera review baixo.
- Como experimento, o sistema poderia dar preferência na exibição para o cliente, de lojistas próximos à sua cidade.

Nivel de satisfação por categoria de produto utilidades_domesticas telefonia fixa lablets_impressao_imagem sinalizacao_e_seguranca seguros e servicos relogios_presentes portateis_cozinha_e_preparadores_de_alimentos portateis_casa_forno_e_cafe perfumaria pcs pc_gamer papelaria moveis_quarto noveis_decoracao market_place malas acessorios livros_tecnicos fivros_interesse_geral livros_importados la_cuisine instrumentos_musicais informatica acessorios tria_comercio_e_negocios fraldas_higiene ferramentas_jardim fashion_underwear_e_moda_praia fashion_roupa_masculina fashion_roupa_infanto_juvenil fashion_roupa_feminina fashion_esporte fashion_calcados fashion_bolsas_e_acessorios eletroportateis eletrodomesticos_2 eletrodomesticos dvds_blu_ray construcao_ferramentas_seguranca construcao_ferramentas_jardim construcao_ferramentas_iluminacao construcao_ferramentas_ferramentas construcao ferramentas construcao consoles_games dimatizacao casa_construcao casa_conforto_2 casa_conforto ama_mesa_banho brinquedos beleza_saude automotivo audio artigos_de_natal artigos_de_festas artes_e_artesanato alimentos_bebidas alimentos agro_industria_e_comercio

Prevendo a Insatisfação

Modelo de árvore de decisão

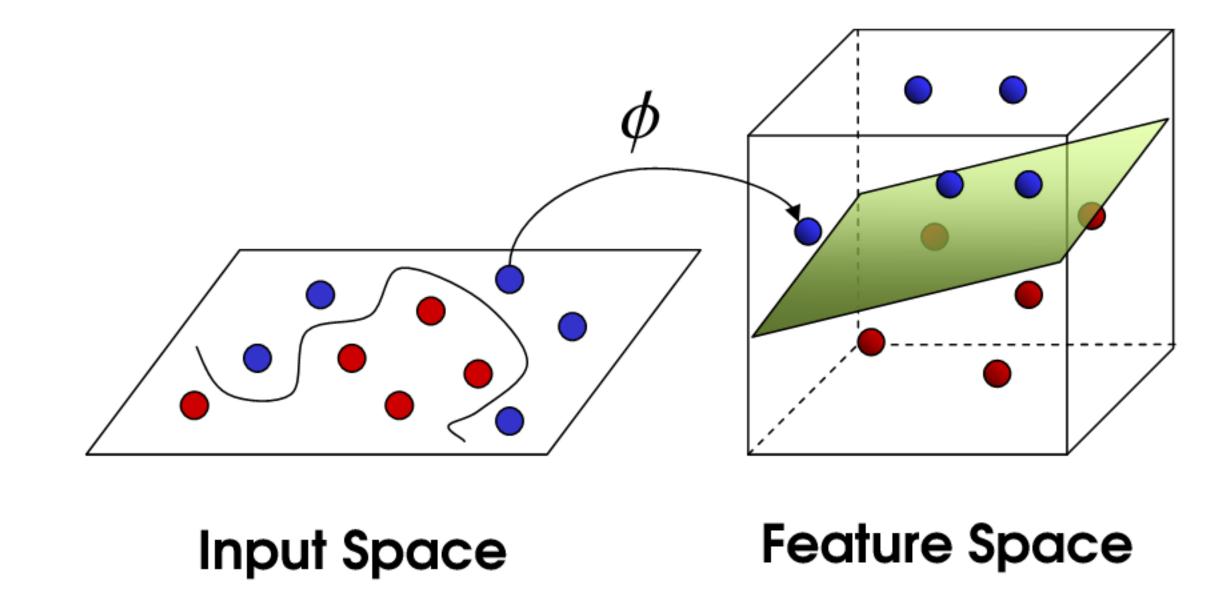
- Buscando ter como base possíveis futuros desafios no time de Marketing, criamos um algoritmo para prever clientes insatisfeitos.
- Foram usadas 14 das features fornecidas no teste
- As features tem dados da compra do cliente, cidade, infos do produto, vendedor e etc.
- Entre testes e ajustes o modelo atingiu uma acurácia de 58% na previsão do Score das compras efetuadas
 - Em um teste o modelo previu 746 possíveis reviews negativos (considerei aqueles com nota menor que 3)
 - Este Algoritmo pode ajudar o time a se antecipar perante possíveis detratores e criar gatilhos que surpreendam esses clientes com uma experiência positiva e que os desencorajem a postar reviews negativos



"Medindo a Febre"

Modelo de análise de sentimento com SVM

- Para complementar o projeto envolvendo prever clientes insatisfeitos antes mesmo de receber o produto
- Desenvolvemos também um modelo para classificar os reviews
- Foi utilizada a técnica Support Vector Machine
- Entre testes e ajustes o modelo atingiu uma acurácia de 92% na classificação de reviews positivos e 91% para reviews negativos
 - Este Algoritmo pode ajudar o time a mais rapidamente classificar comentários negativos e tomar ações corretivas.



- Categorias mais propensas a insatisfação:
 - Cama Mesa & Banho
 - Acessórios de Informática
 - Moveis e Decoração
 - Clientes que recebem o produto em até 11 dias antes do prazo acabar ficam satisfeitos.

- Como experimento, o sistema poderia dar preferência na exibição para o cliente, de lojistas próximos à sua cidade.
- SP, RJ e MG são os estados com mais reclamações

- Clientes que recebem o produto em até 11 dias depois da compra ficam satisfeitos.
- Aumentar a data de entrega estimada, aumentaria o score dos reviews ? Assim, produtos com reviews mais altos teriam melhor conversão ?
 - Algoritmo para prever os Insatisfeitos e Classificar os Reviews