

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
ESCOLA DE NEGÓCIOS
BIG DATA E INTELIGÊNCIA ANALÍTICA

DOUGLAS BRANCO PESSANHA LOPES

ATIVIDADE PRÁTICA DA DISCIPLINA DE INTELIGÊNCIA ANALÍTICA EM
NEGÓCIOS

Atividade Teórico Prática(ATP)
apresentada ao curso de
graduação Big Data e
Inteligência Analítica da
Pontifícia Universidade Católica
do Paraná, como requisito
parcial para aprovação na
disciplina de
Big Data e Inteligência Analítica
em Negócios
Prof-tutor orientador: Lucas
Emanuel Silva e Oliveira.

CURITIBA

2021

ÍNDICE

1. INTRODUÇÃO.....	3
2. DADOS E INFORMAÇÕES ESCOLHIDAS.....	3
3. METODOLOGIA DE TRATAMENTO DOS DADOS E FERRAMENTAS DE ANÁLISE	4
4. RELAÇÃO E CORRELAÇÃO CAUSADAS.....	5
5. DIAGNÓSTICO PARA A EMPRESA E AÇÕES A SEREM TOMADAS.....	7
6. CONSIDERAÇÕES FINAIS.....	7

1. INTRODUÇÃO

Este trabalho tem como objetivo realizar a análise de dados de uma empresa com a finalidade de uma melhor compreensão da sua estrutura organizacional e de suas atividades para propor soluções de melhoramentos que possibilitem ganhos em seus resultados.

Foi escolhido um restaurante localizado em um dos estados do sul do país para que fosse realizada a análise. As informações sobre o restaurante foram encontradas no site [kaggle.com/datasets](https://www.kaggle.com/datasets) e o nome do estabelecimento não foi divulgado. Desta forma, embora tenha sido possível tirar conclusões interessantes dos dados coletados, as análises se limitam ao arquivo obtido, não sendo possível uma extrapolação a partir de outras características da empresa.

2. DADOS E INFORMAÇÕES ESCOLHIDAS

Os dados foram coletados pela própria empresa. Inicia-se em 14/02/2018 e estende-se até 06/03/2020. Basicamente tratam-se de informações sobre o número de vendas e as características do dia, conforme descreveremos abaixo. Não há nenhum campo sem valor. Com elas, podemos verificar quais são os fatores que influem na venda e buscarmos uma correlação entre eles e o número de vendas.

Segue nome e descrição de cada coluna do dataset utilizado:

'DATA': Informa a data em que foram feitas as vendas.

'VENDAS': Número de vendas realizadas naquele dia.

'FDS': Informa se o dia é um fim de semana ou não.

'DS': Oposto do 'FDS', ou seja, dia de semana.

'DATA_FESTIVA'. Não há muita informação do que é considerada uma data festiva ou não. Também é um dado booleano.

'VESPERA_DATA_FESTIVA'. Sem descrição mais detalhada, obviamente é o dia anterior a 'DATA_FESTIVA'.

'POS_DATA_FESTIVA': Também sem descrição detalhada, refere-se aos dias após as datas festivas.

'DATA_NAO_FESTIVA': Informa se a data é ou não festiva, não faz detalhamento do que seriam datas festivas.

'FERIADO': Traz a informação se o dia seria feriado ou não. Não existe explicação de quais tipos de feriado foram considerados.

'NAO_FERIADO': Oposto do 'FERIADO'.

'IS_SEMANA_PAGAMENTO': Como no Brasil o pagamento geralmente é mensal, aqui temos a informação da semana de pagamento, também um valor booleano.

'SEMANA_DE_NAO_PAGAMENTO': O oposto de semana de pagamento.

'BAIXA_TEMPORADA': Outra coluna sem uma explicação precisa.

'ALTA_TEMPORADA': Oposto de 'BAIXA_TEMPORADA'.

'QTD_CONCORRENTES': Apresenta a quantidade concorrentes, variando entre 07 e 13.

'PRECIPITAÇÃO': Demonstra os milímetros de chuva durante o dia.

'TEMPERATURA': Apresenta a temperatura, como não é apresentado somente um valor, concluiu-se ser a temperatura média.

'UMIDADE': Coluna com a umidade do ar.

3. METODOLOGIA DE TRATAMENTO DOS DADOS E FERRAMENTAS DE ANÁLISE

Não existe a informação do valor das vendas, então não foi possível retirar qualquer análise a partir disto. Entretanto, existe uma informação bastante relevante, que é a quantidade de vendas feitas por dias. Deste modo, buscou-se a análise das relações existentes entre as vendas e os demais dados fornecidos pela empresa. Foi então observado se havia alguma alteração nas vendas dependendo de outros fatores coletados. Para isso, utilizou-se ferramentas básicas de estatística.

Os dados foram fornecidos em um arquivo csv e tratados na linguagem python usando a biblioteca pandas. Foram feitas análises estatísticas usando essa ferramenta. Também foi utilizado o PowerBI para gerar visualizações das informações.

Sobre os dados booleanos, verificou-se os valores mínimos, máximo, desvio padrão, valores no primeiro, segundo e terceiro quartis, além da média. Nenhum deles apresentava grandes outliers ou dados que pudessem ser muito discrepantes. O desvio padrão não era muito grande em nenhum deles, revelando uma proximidade da média no conjunto dos dados. Em seguida foi calculada a média de vendas dentro de cada uma das categorias apresentadas sendo possível observar como estas afetavam o total de vendas.

A respeito da precipitação e temperatura, com base na distribuição dos dados nos quartis, foi calculado a média e feito análise semelhante aos dados booleanos. Também tomou-se o cuidado de verificar a existência de valores que pudessem atrapalhar na análise, não foram encontrados.

Com relação aos dados dos concorrentes. Foi feita regressão linear verificando a correlação entre concorrentes e número de vendas.

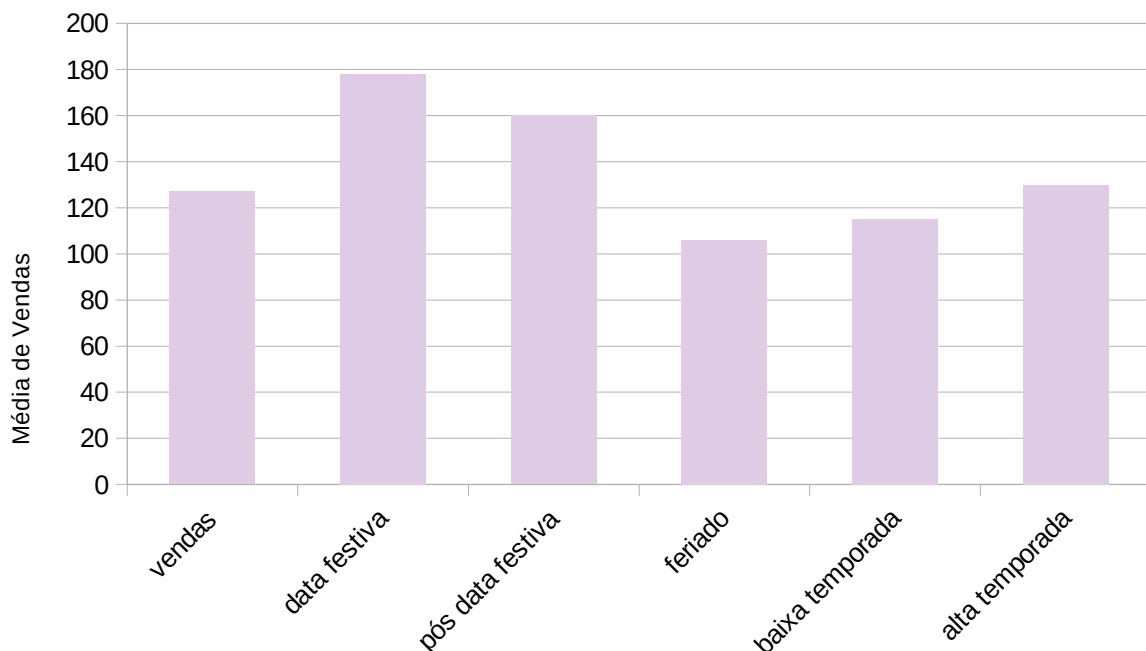
4. RELAÇÃO E CORRELAÇÃO CAUSADAS

A média total de vendas no período analisado é de 127 por dia. Em dias de semana a média de vendas é 126 enquanto nos fins de semana é de 132, sendo muito pequena a variação em relação a média, este fator interfere muito pouco nas vendas feitas pelos restaurante.

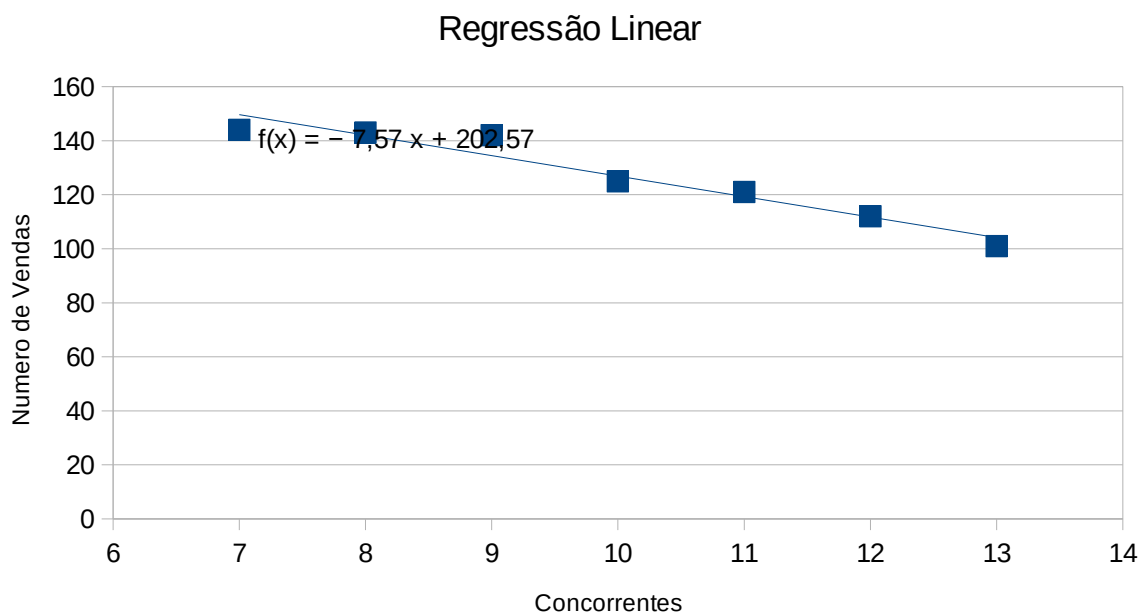
É observado um aumento significativo nos dias de véspera de datas festiva e nas datas festivas, a média é de 160 na véspera, e nas datas festivas, o número de vendas é 178. Na data véspera há um aumento de 23% nas vendas e nas datas festivas uma aumento de 41%. Após a data festiva ocorre uma queda significativa, sendo a média 109 vendas.

A semana de pagamento pouco interfere nas vendas do restaurante. Nas demais semanas, a média de venda é de 127, enquanto na semana de pagamento é de 128. Em baixa temporada acontece uma queda significativa nas vendas, a número baixa em quase 10%.

Tanto a precipitação quanto a temperatura pouco influem nas vendas. Quando a precipitação é 0, algo que corresponde a mais de 50% dos dias, a média é de 129 vendas. Em situações em que a precipitação é maior que 2,3mm, correspondendo a 25% dos dias, são vendidas em média 125 refeições. Quando a temperatura é menor que 22,3 graus, ou seja em 25% dos dias, são vendidas 130 marmitas, quando maior que 30 graus, são vendidas 124. Há uma correlação negativa entre temperatura e venda de marmitas, entretanto a variação é pequena, como existe um limite natural para o aumento e queda de temperatura, acaba não sendo algo muito importante.



Sobre o número de concorrentes, há uma influência grande sobre a venda de refeições e conseqüentemente sobre o negócio. O gráfico abaixo mostra a reta de regressão linear.



Como pode ser observado, com 7 concorrentes, o estabelecimento vendia em média 144 refeições, ao aumentar os concorrentes o numero de refeições cai linearmente chegando ao mínimo de 101 com 13 concorrentes.

5. DIAGNÓSTICO PARA A EMPRESA E AÇÕES A SEREM TOMADAS

As principais quedas verificadas nas vendas ocorrem na baixa temporada e quando surgem novos concorrentes. Ambos os fatores são externos ao negócio e não tem relação direta com a qualidade do produto, atendimento ou preço empregado pelo comerciante. Entretanto, sabendo que objetivamente há uma alteração devido a estes fatores, algumas atitudes podem ser pensadas para diminuir a variação.

Sobre a queda sazonal provocada em baixa temporada, o empreendedor pode investir em sistema de delivery. O cliente pode optar por utilizar aplicativos existentes, com Ifood ou Uber eats para este serviço. Como aumento da área de abrangência da empresa, pode-se compensar a perda sazonal.

A respeito dos impactos da concorrência, pode ser buscada uma fidelização do clientes. Convênios com empresas da região podem ser feitos, promovendo descontos e também podem ser oferecidos descontos a clientes que paguem várias refeições adiantadas.

A forma de verificar a eficácia das ações é pelo aumento do número de vendas.

6. CONSIDERAÇÕES FINAIS

Os dados fornecidos foram suficientes para entender alguns fatores que afetam as vendas feitas no estabelecimento. A partir deles foi possível propor melhorias aos negócios que podem ser medidas através do aumento das vendas. Considera-se que as informações apresentadas neste relatório são suficientes para propor boas alternativas ao negócio, além de atender os requisitos das disciplinas.