



41%

ATIVIDADES
1 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



15.3k xp

a

01

Analizando assinantes

Transcrição

A empresa Alucar possui uma Newsletter mensal sobre descontos e promoções. Depois de dois anos da criação dessa Newsletter, foi solicitada uma análise para descobrir o comportamento do número de assinantes dela, se tem aumentado ou diminuído.

Faremos essa análise. Para começar nosso documento, digitaremos uma hashtag ou cerquilha (#) e escreveremos o título "Alucar - Analizando assinantes da newsletter".

Já temos nossos arquivos em "Files", bastando clicar na seta à esquerda da tela do Colab. Para poder ler o arquivo "newsletter.csv", vamos primeiro armazená-lo em uma variável. Chamaremos essa variável de "assinantes". Também escreveremos o cabeçalho para visualizar os 5 primeiros itens desse arquivo.

```
assinantes = pd.read_csv('newsletter_alucar.csv')  
assinantes.head()
```

COPIAR CÓDIGO

Teremos lido, armazenado na variável e pedido para ler os primeiros registros. Existirá uma coluna para o mês. Sabemos que há 2 anos de Newsletter, desde que a empresa foi aberta. Outra coluna será para a quantidade de assinantes que tivemos a cada mês.



41%

ATIVIDADES
1 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



15.3k xp

a

Agora, descobriremos se os assinantes tem crescido ou diminuído. Porém, não nos esqueceremos de fazer uma análise prévia dos dados para saber, por exemplo, que tipos de dados temos. Como estamos trabalhando com Time Series, nosso mês deve ser tipado como *datetime*. Escreveremos: `assinantes.dtypes`.

Veremos que nosso mês será do tipo *object*, então faremos a devida alteração. Antes disso, veremos primeiro qual é a quantidade de linhas e colunas e a de dados nulos, gerando prints.

```
print('Quantidade de linhas e colunas', assinantes.shape)
print('Quantidade de dados nulos', assinantes.isna().sum().sum())
```

COPIAR CÓDIGO

Nos será devolvido que em 2 anos de registros haverá 24 linhas (referentes aos meses) e 2 colunas, e 0 dados nulos. Portanto, foram realmente colhidas informações para todos os meses.

Converteremos o tipo do mês: `assinantes['mes'] = pd.to_datetime(assinantes['mes'])`.

Para conferir se a conversão foi executada, podemos digitar `assinantes.dtypes`. Agora, o mês já estará tipado como *datetime64*.

Com o conhecimento que já adquirimos no DataFrame do Alucar, vamos fazer o `diff` para descobrir o aumento e a aceleração do número de assinantes.

```
assinantes['aumento'] = assinantes['assinantes'].diff()
assinantes['aceleracao'] = assinantes['aumento'].diff()
```



41%

ATIVIDADES
1 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODOS
NOTURNOS



15.3k xp



COPIAR CÓDIGO

Então, com o `diff` dos assinantes, descobrimos o aumento. Com o `diff` do aumento, descobrimos a aceleração. Para conferir, escrevemos `assinantes.head` e teremos o resultado de que no mês 1 ainda não tínhamos nenhum assinante. No mês 2 teremos 10 assinantes, portanto o aumento foi de 10, porque o cálculo foi $10 - 0$.

No mês 3, estaremos com 21 assinantes. $21 - 10 = 11$. Será 11 nosso aumento, e 1 nossa aceleração, pois calculamos 11 menos os 10 referentes ao aumento do mês 2, e assim por diante.

Teremos um plot de comparação para exibir os 3 itens. Se escrevermos `plot_` e pressionarmos o "Tab" o programa já fornecerá a função de comparação `plot_comparacao(x, y1, y2, y3, dataset, titulo)`. Completaremos:

```
plot_comparacao('mes', 'assinantes', 'aumento', 'aceleracao', assinantes, 'Análise de  
assinantes da newsletter')
```

Pressionaremos "Shift + Enter" para executar e os gráficos funcionarão.

O primeiro, de assinantes, crescerá inicialmente, depois esse aumento reduzirá, mas a tendência será de crescimento de qualquer forma.

O gráfico de aumento dos assinantes teve um boom, mas chegando num determinado período de tempo, um pouco antes do mês 4 de 2018, o aumento começa a cair.



41%

ATIVIDADES

1 de 6

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



15.3k xp

a

A aceleração começa positiva, entretanto, se torna negativa muito significativamente. No mês 10 de 2018, chega quase a -5 antes de se recuperar um pouco.

Com esses gráficos, veremos que decompondo as vendas poderemos encontrar casos muito interessantes. Se tivéssemos só o gráfico da análise de assinantes, acharíamos que a tendência seria ao crescimento expressivo, ainda que isso diminua.

Fazendo a decomposição, vemos que de fato houve um crescimento, mas depois o aumento do número de assinantes cai de forma notória. Também ocorre desaceleração dos assinantes da newsletter num determinado momento, havendo aceleração negativa, ainda que ela tenha se iniciado positiva.

É muito importante podermos ver isso com clareza por meio dos gráficos de comparação