

04

Faça o que eu fiz na aula

Seguindo o contexto da aula com os dados do TMDB e a biblioteca do seaborn importada, vamos verificar a quantidade de línguas que existem no nosso Dataframe. Vamos contar quantas línguas existem no nosso CSV e além disso iremos fazer uma soma e alocar todas as línguas de filmes que estão em inglês na variável total_de_ingles

```
total_por_lingua = tmdb["original_language"].value_counts()  
total_geral = total_por_lingua.sum()  
total_de_ingles = total_por_lingua.loc["en"]
```

Vamos fazer a subtração do total_geral menos o total_de_ingles para saber quantas línguas temos que não sejam em inglês e mostrar os dois na tela.

```
total_do_resto = total_geral - total_de_ingles  
print(total_de_ingles, total_do_resto)
```

Prosseguindo vamos criar os dados para que sejam válidos vai ter os dados em inglês e outros na coluna línguas e na coluna total vai ter o total de inglês e o de outros. E com isto vamos fazer um Dataframe.

```
dados = {  
    'lingua' : ['ingles', 'outros'],  
    'total' : [total_de_ingles, total_do_resto]  
}  
dados = pd.DataFrame(dados)
```

Iremos fazer um gráfico em barra o nosso X vai ser língua e o Y vai ser nossos total e os nossos dados vai ser dados:

```
sns.barplot(x="lingua", y="total", data = dados)
```

E vamos fazer uma query para mostrar todas as línguas que não sejam em inglês e contar os valores.

```
total_por_lingua_de_outros_filmes = tmdb.query("original_language != 'en'").original_language.value_  
total_por_lingua_de_outros_filmes
```

E fazer outra query para que seja os filmes que não contém a língua original em inglês.

```
filmes_sem_lingua_original_em_ingles = tmdb.query("original_language != 'en'")
```

Feito isto como podemos gerar um gráfico de barras sem a língua original em inglês?

