



Data Visualization part 2

- quando há muita informação, muitas vezes os gráficos não conseguem passar a mensagem desejada
- isso acontece tanto no gráfico de barras, como no gráfico de pizza, que pode ser visto em:

```
plt.pie(contagem_de_lingua.total, labels=contagem_de_lingua.original_language)
# plotagem usando pyplot do matplotlib, função pie, em português chamamos de pizza
```

- portanto, devemos saber a mensagem a ser passada
- separar e comparar esses dados de uma maneira mais adequada
- associar a essa mensagem o gráfico que melhor lhe representa

Re-organizando os dados

```
total_por_lingua = tmdb.original_language.value_counts() # mostra a contagem de cada língua
total_geral = total_por_lingua.sum() # soma todas as categorias
total_de_ingles = total_por_lingua.loc['en'] # localiza (soma), tudo que encontrar da categoria en
total_do_resto = total_geral - total_de_ingles # subtrai total - categoria en
print(total_de_ingles, total_do_resto)
```

Utilizando um dicionário

```
dados = {  
    'lingua': ['ingles', 'outros'],  
    'total': [total_de_ingles, total_do_resto]  
}  
dados
```

```
pd.DataFrame(dados)  
# transforma o dicionário criado em em Data Frame com o pandas
```

Plotando o Data Frame

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

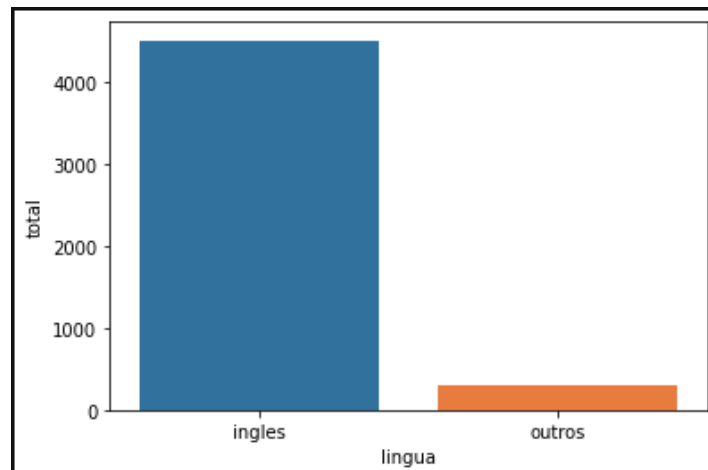


gráfico comparando a quantidade de filmes em inglês, com o total de filmes em outras línguas

- no gráfico acima pode-se visualizar a mensagem clara que há muitos mais filmes na língua inglesa, que em qualquer outra língua