



Projeto de Banco de Dados

etapa 6

Lucas Fernandes André - 182495
Mariana Alves de Sousa - 241201



Contexto e motivação do dataset

Infelizmente, o covid-19 gerou milhões óbitos no mundo inteiro. Devido à pandemia e ao isolamento social, diversas esferas da sociedade foram afetadas. É importante analisar em como esses problemas afetaram a sociedade sob uma ótica econômica nestes últimos anos.

g1**ECONOMIA** **BUSCAR**


Desemprego diante da pandemia bate recorde no Brasil em setembro, aponta IBGE

País encerrou o mês com 13,5 milhões de desempregados, cerca de 3,4 milhões a mais que em maio, o que representa uma alta de 33,1%. Taxa de desemprego ficou em 14%.



Descrição do objetivo do dataset

- Análise comparativa entre os estados brasileiros com relação ao dados de covid-19.
- Análise econômica da sociedade brasileira devido ao covid-19.



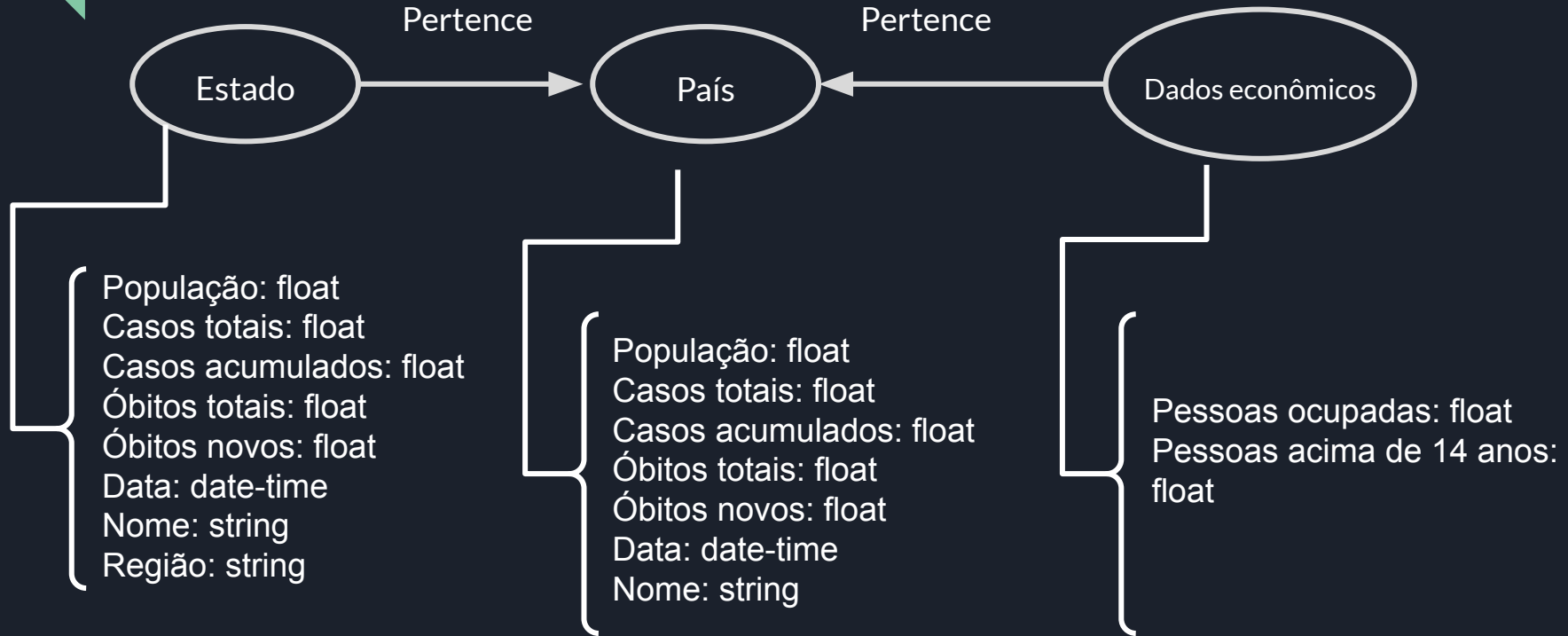
Modelo lógicos

Os modelos lógicos estão na forma de tabela.

Por exemplo:

data	data_mes	regiao	obitosNovos
2021	5	Brasil	59010
2021	5	Centro-Oeste	5137
2021	5	Nordeste	12257
2021	5	Norte	3220
2021	5	Sudeste	29485
2021	5	Sul	9013
2021	6	Brasil	55275
2021	6	Centro-Oeste	5193
2021	6	Nordeste	11095

Modelos Relacional



Fontes de dados

Dados relacionados à quantidade de casos, óbitos e região:

<https://covid.saude.gov.br/>

Detalhar por							
Região	Estado	Município	Região Metropolitana	Todos			
Região	População	Casos Novos	Casos Acumulados	Casos Acumulados 100mi	Óbitos Novos	Óbitos Acumulados	Óbitos Acumulados 100mi
Totais	210.147.125	15.300	21.924.598	10.433	188	610.224	290
Sudeste	88.371.433	4.325	8.555.357	9.681	72	290.057	328
Sul	29.975.984	2.918	4.267.795	14.237	58	96.221	321
Nordeste	57.071.654	5.876	4.876.483	8.544	33	118.399	207
Norte	18.430.980	1.037	1.872.514	10.160	16	46.917	255
Centro-Oeste	16.297.074	1.144	2.352.449	14.435	9	58.630	360



Tratamento de dados

Extração de tabelas .csv que constavam os dados de covid-19 do Brasil.

Integração entre essas tabelas.

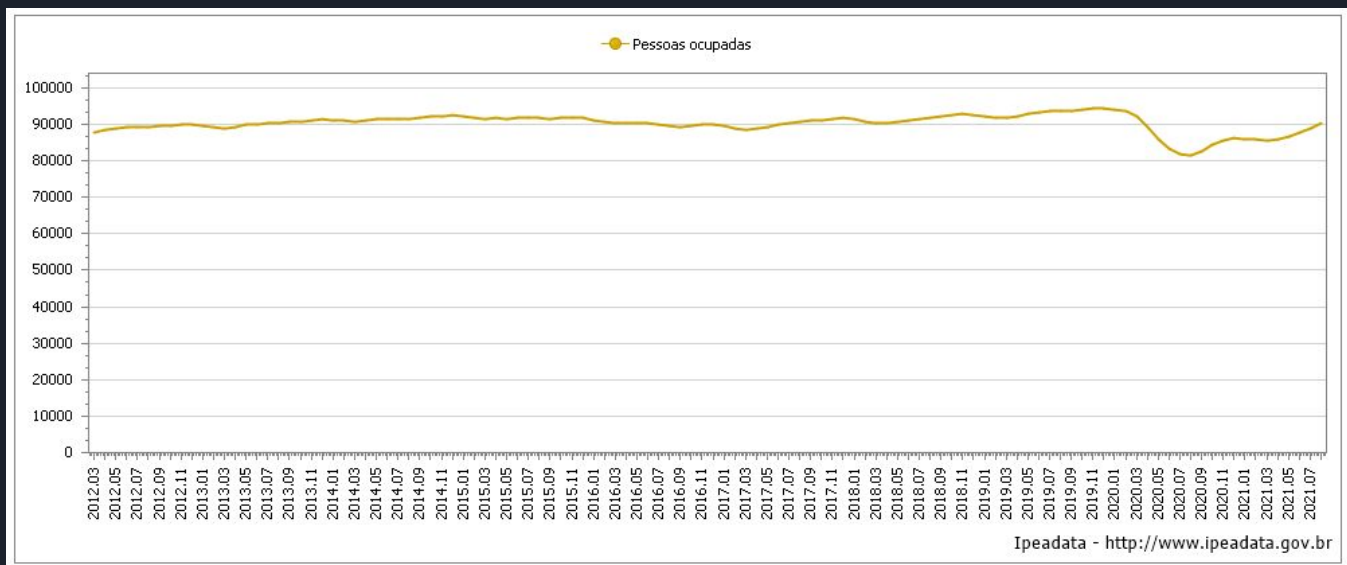
Separação entre regiões do Brasil e escolha de colunas de interesse.

regiao	estado	coduf	data	semanaEpi	populacaoTCU2019	casosAcumulado	casosNovos	obitosAcumulado
Brasil		76	2020-05-25	22	210147125.0	374898.0	11687	23473
Brasil		76	2020-05-26	22	210147125.0	391222.0	16324	24512
Brasil		76	2020-05-27	22	210147125.0	411821.0	20599	25598
Brasil		76	2020-05-28	22	210147125.0	438238.0	26417	26754
Brasil		76	2020-05-29	22	210147125.0	465166.0	26928	27878
Brasil		76	2020-05-30	22	210147125.0	498440.0	33274	28834
Brasil		76	2020-05-31	23	210147125.0	514200.0	15760	29314
Brasil		76	2020-06-01	23	210147125.0	526447.0	12247	29937

Fonte de dados

Dados relacionados à quantidade de pessoas ocupadas (exerce atividade profissional, formal ou informal, remunerada ou não, durante pelo menos 1 hora completa na semana de referência da pesquisa):

<http://www.ipeadata.gov.br/>





Tratamento de dados

Extração de tabelas .csv que constavam o número de pessoas maiores que 14 anos e número de pessoas ocupadas (ambas vezes mil).


Integração entre essas tabelas.

Data	Pessoas de 14 anos ou mais	Pessoas ocupadas
2019.01	170193	92291
2019.02	170395	91880
2019.03	170500	91863
2019.04	170494	92365
2019.05	170615	92947
2019.06	170864	93342
2019.07	170975	93584
2019.08	171123	93631
2019.09	171158	93801
2019.10	171281	94055
2019.11	171401	94416



Pergunta relacionada

- O mês de 2020 que teve maior número de óbitos foi o mesmo mês que teve menor número de pessoas ocupadas?




```
[ ] path1 = '../data/processed/data_economico.csv'  
df = pd.read_csv(path1)
```

```
[ ] menor = df.loc[0, "Pessoas ocupadas"]  
id_menor = 0  
for i in range(len(df)):  
    if df.loc[i, "Pessoas ocupadas"] < menor:  
        menor = df.loc[i, "Pessoas ocupadas"]  
        id_menor = i  
print(df.loc[id_menor, "Data"])
```

2020.08

Agosto de 2020 foi o mês com menos pessoas ocupadas.




```
[ ] path2 = '../data/processed/analiseobito.csv'
    df = pd.read_csv(path2)

[ ] maior = (df.loc[0, "obitosNovos"])
    id_maior = 0
    for i in range(len(df)):
        if (df.loc[i, "obitosNovos"]) > maior and df.loc[i, "data"]==2020:
            maior = (df.loc[i, "obitosNovos"])
            id_maior = i
    print(df.loc[id_maior, "data_mes"])
```

7

Julho de 2020 foi o mês com mais óbitos.



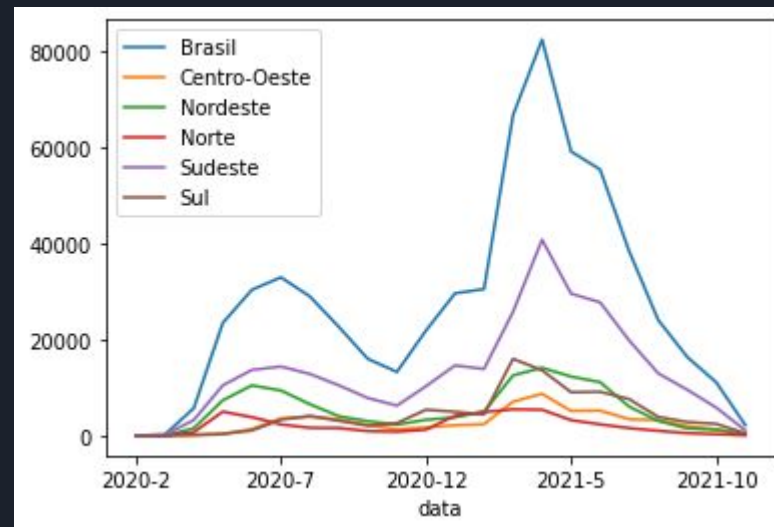
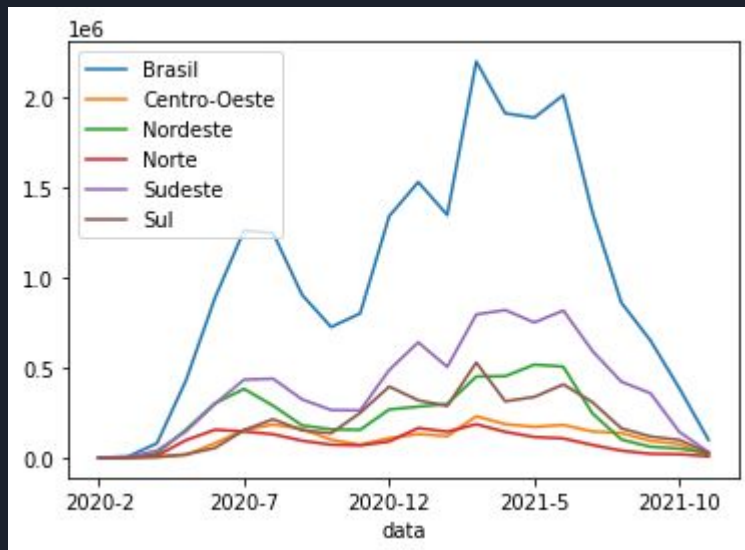
```
[ ] path1 = '../data/processed/data_economico.csv'
    df = pd.read_csv(path1)

[ ] menor = df.loc[0, "Pessoas ocupadas"]
    id_menor = 0
    for i in range(len(df)):
        if df.loc[i, "Pessoas ocupadas"] < menor and df.loc[i, "Data"] != 2020.08:
            menor = df.loc[i, "Pessoas ocupadas"]
            id_menor = i
    print(df.loc[id_menor, "Data"])
```

2020.07

O segundo mês de 2020 com menos pessoas ocupadas coincide com o de mais óbitos, que é julho de 2020.

Comparação de número de casos e número de óbito





Possível conclusão

Devido a necessidade de medidas de distanciamento social, teve uma paralisação temporária em algumas atividades econômicas, o que também influenciou na quantidade de oportunidades disponíveis e até mesmo na decisão das pessoas de procurarem trabalho. Com o relaxamento dessas medidas ao longo do ano de 2020, um maior maior número de pessoas pôde voltar a buscar uma ocupação.



Possíveis perguntas

Quando houve um crescimento no número de ocupados, o número de casos subiu consequentemente?

Entre os estados brasileiros, qual região ficou mais afetada? Nessa região, o número de pessoas maiores de 14 anos é maior?

As regiões com o maior número de casos em determinado mês também são as regiões com maior número de mortes?