

Atividade 1 – Projeto Covid19 Brasil

O Brasil continua sofrendo com a pandemia causada pela infecção generalizada pelo vírus Covid-19, também conhecido por Corona vírus. Como você trabalha numa empresa de Data Science voltada a Helthcare, agora seu objetivo é analisar o impacto da doença nos estados brasileiros e correlacionar com suas respectivas características.

Utilizar o arquivo checkpoint1.ipynb para obter os dados do Covid para os estados e carregar os dados do censo de 2010. Faça e responda as seguintes perguntas:

1 – Quantas cidades existem em cada dataframe? (dados_municipios_gerais e dados_municipios_covid)?

2 – Como realizar o join desses dataframes? (realizar o join!! Dica, usar a função pd.merge). Indicar quais cidades não houve o join e retirar das análises.

3 – No dataset após o join, obter o índice de Gini por UF (state). Qual UF chamou atenção? Por que isto ocorreu?

4 – Obtenha o sumário das colunas por UF (state), similar a figura abaixo (apresentar o código):

state	CasosTotais	MortesTotais	MediaESPVIDA	MediaE_ANOSESTUDO	MediaT_ANALF18M	Mediana_IDHM	Mediana_RDPC	DP_RDPC
AC	86644	1776	70.991818	7.752273	26.106818	0.5885	313.705	127.780499
AL	225064	5610	69.367843	8.896961	35.308529	0.5625	233.440	79.124345
AM	410367	13421	70.924677	7.451774	20.889194	0.5640	227.585	98.250134
AP	119540	1877	72.028125	9.035625	13.997500	0.6410	402.440	133.018963
BA	1157427	24835	70.570288	8.453453	25.816427	0.5890	272.860	101.927806

5 – Por meio das análises na tabela obtida na questão 4, pode-se presumir algum resultado interessante? Apresente sua opinião.