Crie uma função que receba um Pyspark DataFrame e realize a escrita desse DataFrame no formato JSON com todas as opções de compressão disponíveis na [documentação](https://spark.apache.org/docs/3.3.1/sql-data-sources-json.html). Após a escrita, identifique qual a melhor forma de compressão para os dados apresentados durante o curso.

O código da função apresentada durante o vídeo está disponível abaixo:

def salvando\_arquivo\_comprimido\_json(dataframe):

compressoes = ["none", "bzip2", "gzip", "lz4", "snappy", "deflate"]

caminho\_padrao = "/FileStore/tables/arquivos\_curso/desafio/json\_"

for tipo\_compressao in compressoes:

dataframe.write\

.option("compression", tipo\_compressao)\

.mode("overwrite") \

.format("json") \

.save(caminho\_padrao + tipo\_compressao)

display(dbutils.fs.ls(caminho\_padrao + tipo\_compressao))

COPIAR CÓDIGO

E para executar a função que acabamos de criar executamos o comando:

salvando\_arquivo\_comprimido\_json(df)