MVP 2 – Engenharia de Dados

Aluno: Vanderson Lopes Felix da Silveira

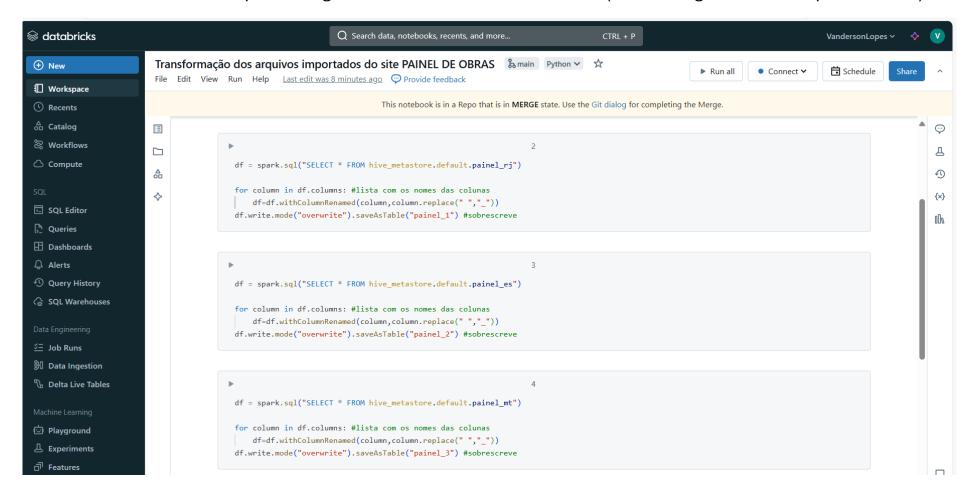
Painel de Obras Públicas do Brasil

Data: 12/07/2024

Evidências do código de transformação dos arquivos importados do site PAINEL DE OBRAS

Continua...

A tela inteira evidenciando que o código no meu notebook do Databricks (ver meu login no canto superior direito).



Continua...

O detalhe do código (mais legível). Foram 3 rotinas ao todo.

Transformação dos arquivos importados do site PAINEL DE OBRAS & main Python ✓ ☆ File Edit View Run Help Last edit was 12 minutes ago Provide feedback This notebook is in a Repo that is in MERGE state. Use the Git dialog for \blacksquare :: 2 df = spark.sql("SELECT * FROM hive_metastore.default.painel_rj") 습 for column in df.columns: #lista com os nomes das colunas ♦ df=df.withColumnRenamed(column,column.replace(" ","_")) df.write.mode("overwrite").saveAsTable("painel_1") #sobrescreve 3 df = spark.sql("SELECT * FROM hive_metastore.default.painel_es") for column in df.columns: #lista com os nomes das colunas df=df.withColumnRenamed(column,column.replace(" ","_")) df.write.mode("overwrite").saveAsTable("painel_2") #sobrescreve df = spark.sql("SELECT * FROM hive_metastore.default.painel_mt") for column in df.columns: #lista com os nomes das colunas df=df.withColumnRenamed(column,column.replace(" ","_")) df.write.mode("overwrite").saveAsTable("painel 3") #sobrescreve