

**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).

The screenshot displays the Databricks web interface. At the top, the Databricks logo is on the left, a search bar in the center, and the user name 'VandersonLopes' on the right. Below the search bar, the notebook title 'Transformação dos arquivos importados do site PAINEL DE OBRAS' is shown, along with 'main' branch and 'Python' language. A yellow banner indicates the notebook is in a 'MERGE' state. The left sidebar contains navigation options like 'New', 'Workspace', 'Recents', 'Catalog', 'Workflows', 'Compute', 'SQL', 'SQL Editor', 'Queries', 'Dashboards', 'Alerts', 'Query History', 'SQL Warehouses', 'Data Engineering', 'Job Runs', 'Data Ingestion', 'Delta Live Tables', 'Machine Learning', 'Playground', 'Experiments', and 'Features'. The main area shows three code blocks, each with a play button and a line number (2, 3, and 4). Each block contains a Spark SQL query to select data from a specific table in the 'hive\_metastore.default' database, followed by a loop that iterates over the columns, renames them by replacing spaces with underscores, and then writes the result to a new table with the mode 'overwrite'.

```
2
df = spark.sql("SELECT * FROM hive_metastore.default.painel_rj")

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

4
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
```

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



2

```
df = spark.sql("SELECT * FROM hive_metastore.default.painel_rj")

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
```



4

```
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
```