

<https://databricks.com>

Fetch Data from the Open Brewery DB API

```
import requests
import json
import zipfile
import os
from pyspark.sql import SparkSession
from pyspark.sql.types import StructType, StructField, StringType
```

```
# Inicializar Spark Session
spark = SparkSession.builder.appName("BreweryETL").getOrCreate()
```

```
# Passo 1: Buscar dados da API
try:
    response = requests.get("https://api.openbrewerydb.org/breweries")
    response.raise_for_status() # Levantar exceção para erros de resposta
    data = response.json()
except requests.exceptions.RequestException as e:
    print(f"Erro ao buscar dados da API: {e}")
    data = None
```

```
# Passo 2: Definir esquema e criar DataFrame
if data:
    try:
        schema = StructType([
            StructField("id", StringType(), True),
            StructField("name", StringType(), True),
            StructField("brewery_type", StringType(), True),
            StructField("street", StringType(), True),
            StructField("city", StringType(), True),
            StructField("state", StringType(), True),
            StructField("country", StringType(), True),
            StructField("longitude", StringType(), True),
            StructField("latitude", StringType(), True),
            StructField("phone", StringType(), True),
            StructField("website_url", StringType(), True)
        ])

        raw_df = spark.createDataFrame(data, schema=schema)
    except Exception as e:
        print(f"Erro ao criar DataFrame: {e}")
        raw_df = None
```

```
# Passo 3: Salvar DataFrame no DBFS
if raw_df:
    try:
        bronze_path = "/dbfs/tmp/bronze/"
        raw_df.write.mode("overwrite").json(bronze_path)
    except Exception as e:
        print(f"Erro ao salvar DataFrame no DBFS: {e}")
```

%fs
ls /dbfs/tmp/bronze/

Table							
	File path	File name	size	modification time			
1	dbfs/dbfs/tmp/bronze/_SUCCESS	_SUCCESS	0	1721782631000			
2	dbfs/dbfs/tmp/bronze/_committed_1306289423817917245	_committed_1306289423817917245	1455	1721782631000			
3	dbfs/dbfs/tmp/bronze/_committed_4706318761324704560	_committed_4706318761324704560	744	1721776137000			
4	dbfs/dbfs/tmp/bronze/_committed_705453976322867528	_committed_705453976322867528	1466	1721782604000			
5	dbfs/dbfs/tmp/bronze/_committed_vacuum4197618084827725604	_committed_vacuum4197618084827725604	96	1721782604000			
6	dbfs/dbfs/tmp/bronze/_started_1306289423817917245	_started_1306289423817917245	0	1721782630000			
7	dbfs/dbfs/tmp/bronze/_started_705453976322867528	_started_705453976322867528	0	1721782603000			
8	dbfs/dbfs/tmp/bronze/breweries.json/	breweries.json/	0	0			
9	dbfs/dbfs/tmp/bronze/part-00000-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-397-1-c000.js...	part-00000-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-397-1-c000.js...	1731	1721782630000			
10	dbfs/dbfs/tmp/bronze/part-00001-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-398-1-c000.js...	part-00001-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-398-1-c000.js...	1690	1721782631000			
11	dbfs/dbfs/tmp/bronze/part-00002-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-399-1-c000.js...	part-00002-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-399-1-c000.js...	1849	1721782631000			
12	dbfs/dbfs/tmp/bronze/part-00003-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-400-1-c000.js...	part-00003-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-400-1-c000.js...	1674	1721782631000			
13	dbfs/dbfs/tmp/bronze/part-00004-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-401-1-c000.js...	part-00004-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-401-1-c000.js...	1700	1721782631000			
14	dbfs/dbfs/tmp/bronze/part-00005-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-402-1-c000.js...	part-00005-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-402-1-c000.js...	1615	1721782631000			
15	dbfs/dbfs/tmp/bronze/part-00006-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-403-1-c000.js...	part-00006-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-403-1-c000.js...	1648	1721782631000			
16	dbfs/dbfs/tmp/bronze/part-00007-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-404-1-c000.js...	part-00007-tid-1306289423817917245-00f82c5e-4541-4d6d-b071-7644153fdee7-404-1-c000.js...	2265	1721782631000			
16 rows							