Como já tenho o ambiente python em minha máquina, náo foi preciso utilizar o ambiente virtual.

Exercício

Realizaremos uma amostragem, conforme as seguintes especificações:

nome da variável

coluna observada

amostra = df.sampleBy('ticker', fractions = {'HON': 0.03, 'WMT': 0.05, 'BA':0.01})

função de amostra

frações escolhidas

mostramos resultado

amostra.show()

O arquivo .py também foi enviado juntamente com esse arquivo.

```python
Amostragem.py > ...
1    from pyspark.sql import SparkSession
2
3    spark = SparkSession.builder.appName("my_app").getOrCreate()
4
5    df = spark.read.option("delimiter", ",").csv("C:/Users/USER/Downloads/CODE/TOKIO/bigdata/modulo 5/EDA/stocks_2021.csv", header=True, inferSchema=True)
6
7    amostra = df.sampleBy('ticker', fractions = {'HON': 0.03, 'WMT': 0.05, 'BA':0.01})
8
9    amostra.show()
```

```
PROBLEMS    OUTPUT    DEBUG CONSOLE    TERMINAL    SQL CONSOLE

PS C:\Users\USER\Downloads\CODE\TOKIO\bigdata\modulo 5\amostragem de dados> & C:/Users/USER/AppData/Local/Microsoft/WindowsApps/python3.10.exe "c:/Users/USER/Downloads/COI
ados/Amostragem.py"
23/08/15 10:20:37 WARN Shell: Did not find winutils.exe: java.io.FileNotFoundException: java.io.FileNotFoundException: HADOOP_HOME and hadoop.home.dir are unset. -see htt
ms
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
23/08/15 10:20:37 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
+------+------------------+------------------+------------------+------------------+-------+---------+------------+----------+---+
|ticker|              open|              high|               low|             close| volume|dividends|stock splits|      date|ccy|
+------+------------------+------------------+------------------+------------------+-------+---------+------------+----------+---+
|   HON|225.16168009787762|228.11446369587193|225.16168009787762|226.52450561523438|2186500|      0.0|           0|2021-04-14|USD|
|   HON|226.14924210849583|229.75379823683897|226.14924210849583|228.06509399414062|2391600|      0.0|           0|2021-05-10|USD|
|   HON|228.80500603030313|229.72731964305495|226.58349653588851|227.30746459960938|2321300|      0.0|           0|2021-06-03|USD|
|   HON|222.31898675286905|226.87108267727626|222.31898675286905|222.31898675286905|2631800|      0.0|           0|2021-07-14|USD|
|   HON| 230.2727733989021| 231.9289821326574|227.72398702575632|228.42813110351562|3592800|      0.0|           0|2021-07-16|USD|
|   HON|219.68063234430596| 221.9508202720369|219.19273534839354|220.73606872558594|2021700|      0.0|           0|2021-09-15|USD|
|   HON|227.27000427246094|227.91000366210938|225.17999267578125|225.41000366210938|1324800|     0.98|           0|2021-11-10|USD|
|   WMT|146.66762123295769| 147.580562230309|145.99827285569145|146.63809204101562|7791100|      0.0|           0|2021-01-12|USD|
|   WMT|145.50608824342405|145.68326338121142|143.75393787901373| 144.6693878173828|7672400|      0.0|           0|2021-01-14|USD|
|   WMT| 139.0856900227994|139.57002394789876|138.63099349689105|138.97695922851562|6117000|      0.0|           0|2021-05-05|USD|
|   WMT| 140.3265321713795|141.88449509035908| 138.6990983183797| 140.1479034423828|9127700|      0.0|           0|2021-07-19|USD|
|   WMT|140.04868008840064|141.03108510649187|139.72120669588537|140.18760681152344|4337800|      0.0|           0|2021-07-22|USD|
|   WMT|148.40247754175644|149.60763186511247|148.21323705153046| 148.4622344970703|6089700|     0.55|           0|2021-08-12|USD|
|   WMT| 150.9223352511977|151.18128713161917|149.80683152310704|149.84666442871094|5388700|      0.0|           0|2021-08-23|USD|
|   WMT| 146.9184542868296|147.09774035890118| 145.5937859898068|145.93243408203125|7691300|      0.0|           0|2021-08-27|USD|
|   WMT|145.83283075203687|147.40649652513872| 145.434440894522|147.10769653320312|5480200|      0.0|           0|2021-08-30|USD|
|   WMT|149.37854102448688| 150.1454574005334|148.7510778055377|149.18930053710938|4183200|      0.0|           0|2021-11-01|USD|
|   WMT|137.0780842190712|137.88483354418165|134.69766914972124|134.92674255371094|13268600|      0.0|          0|2021-12-02|USD|
|    BA|225.89999389648438|229.1000060351562|221.3000030517578|229.08999633789062|7326400|      0.0|           0|2021-08-03|USD|
+------+------------------+------------------+------------------+------------------+-------+---------+------------+----------+---+
```