```
from pyspark.sql import SparkSession
from pyspark.sql import functions as F
from pyspark.sql.functions import col, when
import sys, os
import matplotlib.pyplot as plt

# Order of object SparkSession
spark = SparkSession.builder.appName("Pyspy").getOrCreate()
sc = spark.sparkContext
df = spark.read.csv("/content/ember_v2/train.csv", header=True, inferSchema=True)
#df.show()

shape = (df.count(), len(df.columns))
print(shape)
df.show(1)
```

```
(1000000, 79)
+---+---+---
```