

spark-pi

March 21, 2021

```
[1]: #  
# Licensed to the Apache Software Foundation (ASF) under one or more  
# contributor license agreements. See the NOTICE file distributed with  
# this work for additional information regarding copyright ownership.  
# The ASF licenses this file to You under the Apache License, Version 2.0  
# (the "License"); you may not use this file except in compliance with  
# the License. You may obtain a copy of the License at  
#  
# http://www.apache.org/licenses/LICENSE-2.0  
#  
# Unless required by applicable law or agreed to in writing, software  
# distributed under the License is distributed on an "AS IS" BASIS,  
# WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.  
# See the License for the specific language governing permissions and  
# limitations under the License.  
#  
  
import sys  
from random import random  
from operator import add  
  
from pyspark.sql import SparkSession
```

```
[2]: if __name__ == "__main__":  
    """  
        Usage: pi [partitions]  
    """  
    spark = SparkSession\  
        .builder\  
        .appName("PythonPi")\  
        .getOrCreate()  
  
    partitions = 2  
    #int(sys.argv[1]) if len(sys.argv) > 1 else 2  
    n = 100000 * partitions  
  
    def f(_):
```

```
x = random() * 2 - 1
y = random() * 2 - 1
return 1 if x ** 2 + y ** 2 <= 1 else 0

count = spark.sparkContext.parallelize(range(1, n + 1), partitions).map(f).
↪reduce(add)
print("Pi is roughly %f" % (4.0 * count / n))

spark.stop()
```

Pi is roughly 3.145240