

```
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

```
from pyspark.sql import SparkSession
```

```
spark=SparkSession.builder.appName("WordCount").getOrCreate()
```

```
rdd=spark.sparkContext.textFile("/content/drive/MyDrive/Input.txt")
```

```
words=rdd.flatMap(lambda line:line.split(" ")).map(lambda word:(word,1))
```

```
word_counts=words.reduceByKey(lambda a,b:a+b)
```

```
for word,count in word_counts.collect():
    print(f"{word}:{count}")
```

Streaming output truncated to the last 5000 lines.

```
ridicule,:2
forfeit!":3
Nizhni.:1
Contributions.:1
Gallicism,"":1
author::1
"Quand:2
regiment,"":1
caustic,:1
target:4
unreasonable,:1
Razumovskis:1
going::1
besides--I:1
Obolenski's:3
Belaya:3
Tserkov:3
returns.":1
brighter.:1
novels.":1
admire,:1
"Forfeit,:2
Moscow'?:1
Natalie,:1
Vera.":1
cruelty...":1
s'accuse,"*:1
*"Who:2
peu:1
Russian?":1
Wittgenstein:2
arsenal::1
jocose:1
Chigirin:1
talks.:1
stormcloud:2
wait?":1
bantering:2
"Satisfactory,:1
satisfactory!:1
Barbara:2
rule.":1
Muttering:1
See!:1
idiotic:3
lockup:1
hair'? :1
(How:1
silly!):1
says.:4
irresolution,:1
Vorontsovo:1
Lobnoe:3
flogged.:1
piteously.:1
Frenchmen.:6
crowd--officials,:1
```

```
spark.stop()
```

