```
spark
```

SparkSession - in-memory

SparkContext

Spark UI

Version
v3.4.0
Master
local[*]
AppName
Spark - Lab 3

```
with open("view_data_lab3.csv", 'r') as file:
    lines = file.readlines()

header = lines[0].strip().split('\t')
rows = [line.strip().split(',') for line in lines[1:]]
items = sc.parallelize(rows) \
          .filter(lambda item: 200000 <= int(item[0]) < 230000) \
          .map(lambda item: ((item[3], item[4]), 1)) \
          .reduceByKey(lambda a, b: a + b) \
          .map(lambda x: (x[0][0], (x[1], 1))) \
          .reduceByKey(lambda a, b: (a[0] + b[0], a[1] + b[1])) \
          .mapValues(lambda x: x[0] / x[1]) \
          .map(lambda x: (x[1], x)) \
          .sortByKey(False).map(lambda x: x[1]) \

for i, j in items.take(5):
  print(i, j)
```

```
7.5E+14 64.0
7.46E+14 8.0
7.503E+14 7.166666666666667
8.00001E+11 5.6
8.4843E+14 5.4
```