

1. SPARK

1.1 Single record lookup

```
spark2-submit --class sparkAssignmentRDDfetch.SparkRDD1 \  
  
--master yarn --deploy-mode client /home/ec2-user/sparkAssignmentRDDfetch-  
0.0.1-SNAPSHOT.jar /user/ec2-user/spark_assignment/input2/yellow_tripdata_*  
/user/ec2-user/output_spark/sparkrdd_fetchrecord2
```

1.2 Filter records with rate code id 4

```
spark2-submit --class sparkAssignmentRDD2.saprkRDD2 \  
  
--master yarn --deploy-mode client /home/ec2-user/sparkAssignmentRDD2-0.0.1-  
SNAPSHOT.jar /user/ec2-user/spark_assignment/input2/yellow_tripdata_*  
/user/ec2-user/output_spark/sparkrdd2_fetchrecord1
```

1.3 Group by records based on payment type

```
spark2-submit --class sparkPayment1.topPayment \  
  
--master yarn --deploy-mode client /home/ec2-user/sparkPayment1-0.0.1-  
SNAPSHOT.jar /user/ec2-user/spark_assignment/input2/yellow_tripdata_*  
/user/ec2-user/output_spark/sparkrdd3_fetchrecord8
```

2. MAP REDUCE

2.1 Single record lookup

```
hadoop jar sparkq1.jar sparkAssignmentMR.fetchDriver /user/ec2-  
user/spark_assignment/input2/yellow_tripdata_* /user/ec2-  
user/output_spark/sparkq1_output2
```

2.2 Filter records with rate code id 4

```
hadoop jar sparkq2.jar sparkAssignmentMRRate.fetchDriverRatecodeID /user/ec2-  
user/spark_assignment/input2/yellow_tripdata_* /user/ec2-  
user/output_spark/sparkq2_output1
```

2.3 Group by records based on payment type

```
hadoop jar sparkq3.jar sparkAssignmentMRpayment.driverPayment /user/ec2-  
user/spark_assignment/input2/yellow_tripdata_* /user/ec2-  
user/output_spark/sparkq3_output1
```