

Name: Arun Kumar N
Student ID : 2017CBDE030
Spark Assignment

EC2 Data:

Job Run Time

	Single Record Lookup	Filter	Group by Accompanied with Order by
Pig	215s	283s	431s
Spark RDD	95s	16s	110s

EC2 Data:**Total Count of Records in Output(The cells in each column should have the same value)**

	Single Record Lookup	Filter	Group by Accompanied with Order by
Pig	1	31029	5
Spark RDD	1	31029	5

Local File 2MB Data**Total Count of Records in Output(The cells in each column should have the same value)**

	Single Record Lookup	Filter	Group by Accompanied with Order by
Pig	6	21018	4
Spark RDD	6	21018	4



Local File(2MB Data)

	Single Record Lookup	Filter	Group by Accompanied with Order by
Pig	21s	20s	75s
Spark RDD	4611 ms	3698 ms	3765 ms