

File Edit View Search Terminal Help

```
20/06/24 22:27:08 INFO ShutdownHookManager: Detecting block manager /tmp/spark-21149c1-1801-4430-b3d4-288c0c02b3c0
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ bin/spark-submit --class com.prerna.retail.ProductSales ../jars_file/productjar.jar .
./data/Retail_Sample_Data_Set.txt ret-out-prod-01
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
20/06/24 22:27:08 INFO SparkContext: Running Spark version 2.0.2
20/06/24 22:27:09 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
20/06/24 22:27:09 WARN Utils: Your hostname, prerna-Inspiron-5559 resolves to a loopback address: 127.0.1.1; using 192.168.1.10 instead (on in
terface wlp1s0)
20/06/24 22:27:09 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
20/06/24 22:27:09 INFO SecurityManager: Changing view acls to: prerna
20/06/24 22:27:09 INFO SecurityManager: Changing modify acls to: prerna
20/06/24 22:27:09 INFO SecurityManager: Changing view acls groups to:
20/06/24 22:27:09 INFO SecurityManager: Changing modify acls groups to:
20/06/24 22:27:09 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(prerna);
groups with view permissions: Set(); users with modify permissions: Set(prerna); groups with modify permissions: Set()
20/06/24 22:27:09 INFO Utils: Successfully started service 'sparkDriver' on port 37891.
20/06/24 22:27:09 INFO SparkEnv: Registering MapOutputTracker
20/06/24 22:27:09 INFO SparkEnv: Registering BlockManagerMaster
20/06/24 22:27:09 INFO DiskBlockManager: Created local directory at /tmp/blockmgr-7dfe78a6-7ec3-4b40-a51b-5246ba0bf71f
20/06/24 22:27:09 INFO MemoryStore: MemoryStore started with capacity 366.3 MB
20/06/24 22:27:09 INFO SparkEnv: Registering OutputCommitCoordinator
20/06/24 22:27:10 INFO Utils: Successfully started service 'SparkUI' on port 4040.
20/06/24 22:27:10 INFO SparkUI: Bound SparkUI to 0.0.0.0, and started at http://192.168.1.10:4040
20/06/24 22:27:10 INFO SparkContext: Added JAR file:/home/prerna/spark-2.0.2-bin-hadoop2.6/../../jars_file/productjar.jar at spark://192.168.1.10
:37891/jars/productjar.jar with timestamp 1593017830209
20/06/24 22:27:10 INFO Executor: Starting executor ID driver on host localhost
20/06/24 22:27:10 INFO Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port 36867.
20/06/24 22:27:10 INFO NettyBlockTransferService: Server created on 192.168.1.10:36867
20/06/24 22:27:10 INFO BlockManagerMaster: Registering BlockManager BlockManagerId(driver, 192.168.1.10, 36867)
20/06/24 22:27:10 INFO BlockManagerMasterEndpoint: Registering block manager 192.168.1.10:36867 with 366.3 MB RAM, BlockManagerId(driver, 192.
168.1.10, 36867)
20/06/24 22:27:10 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 192.168.1.10, 36867)
20/06/24 22:27:10 WARN SparkContext: Use an existing SparkContext, some configuration may not take effect.
20/06/24 22:27:10 INFO SharedState: Warehouse path is 'file:/home/prerna/spark-2.0.2-bin-hadoop2.6/spark-warehouse'.
20/06/24 22:27:13 INFO FileSourceStrategy: Pruning directories with:
20/06/24 22:27:13 INFO FileSourceStrategy: Post-Scan Filters:
20/06/24 22:27:13 INFO FileSourceStrategy: Pruned Data Schema: struct<value: string>
20/06/24 22:27:13 INFO FileSourceStrategy: Pushed Filters:
20/06/24 22:27:13 INFO MemoryStore: Block broadcast 0 stored as values in memory (estimated size 214.4 KB, free 366.1 MB)
```


File Edit View Search Terminal Help

```
20/06/24 22:27:10 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 192.168.1.10, 36867)
20/06/24 22:27:10 WARN SparkContext: Use an existing SparkContext, some configuration may not take effect.
20/06/24 22:27:10 INFO SharedState: Warehouse path is 'file:/home/prerna/spark-2.0.2-bin-hadoop2.6/spark-warehouse'.
20/06/24 22:27:13 INFO FileSourceStrategy: Pruning directories with:
20/06/24 22:27:13 INFO FileSourceStrategy: Post-Scan Filters:
20/06/24 22:27:13 INFO FileSourceStrategy: Pruned Data Schema: struct<value: string>
20/06/24 22:27:13 INFO FileSourceStrategy: Pushed Filters:
20/06/24 22:27:13 INFO MemoryStore: Block broadcast_0 stored as values in memory (estimated size 214.4 KB, free 366.1 MB)
20/06/24 22:27:13 INFO MemoryStore: Block broadcast_0_piece0 stored as bytes in memory (estimated size 20.5 KB, free 366.1 MB)
20/06/24 22:27:13 INFO BlockManagerInfo: Added broadcast_0_piece0 in memory on 192.168.1.10:36867 (size: 20.5 KB, free: 366.3 MB)
20/06/24 22:27:13 INFO SparkContext: Created broadcast 0 from javaRDD at ProductSales.java:30
20/06/24 22:27:13 INFO FileSourceStrategy: Planning scan with bin packing, max size: 4194304 bytes, open cost is considered as scanning 4194304 bytes.
20/06/24 22:27:13 INFO CodeGenerator: Code generated in 232.454231 ms
20/06/24 22:27:14 INFO deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
20/06/24 22:27:14 INFO deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
20/06/24 22:27:14 INFO deprecation: mapred.task.is.map is deprecated. Instead, use mapreduce.task.ismap
20/06/24 22:27:14 INFO deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
20/06/24 22:27:14 INFO deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
20/06/24 22:27:14 INFO SparkContext: Starting job: saveAsTextFile at ProductSales.java:48
20/06/24 22:27:14 INFO DAGScheduler: Registering RDD 4 (mapToPair at ProductSales.java:32)
20/06/24 22:27:14 INFO DAGScheduler: Got job 0 (saveAsTextFile at ProductSales.java:48) with 1 output partitions
20/06/24 22:27:14 INFO DAGScheduler: Final stage: ResultStage 1 (saveAsTextFile at ProductSales.java:48)
20/06/24 22:27:14 INFO DAGScheduler: Parents of final stage: List(ShuffleMapStage 0)
20/06/24 22:27:14 INFO DAGScheduler: Missing parents: List(ShuffleMapStage 0)
20/06/24 22:27:14 INFO DAGScheduler: Submitting ShuffleMapStage 0 (MapPartitionsRDD[4] at mapToPair at ProductSales.java:32), which has no missing parents
20/06/24 22:27:14 INFO MemoryStore: Block broadcast_1 stored as values in memory (estimated size 9.8 KB, free 366.1 MB)
20/06/24 22:27:14 INFO MemoryStore: Block broadcast_1_piece0 stored as bytes in memory (estimated size 5.2 KB, free 366.1 MB)
20/06/24 22:27:14 INFO BlockManagerInfo: Added broadcast_1_piece0 in memory on 192.168.1.10:36867 (size: 5.2 KB, free: 366.3 MB)
20/06/24 22:27:14 INFO SparkContext: Created broadcast 1 from broadcast at DAGScheduler.scala:1012
20/06/24 22:27:14 INFO DAGScheduler: Submitting 1 missing tasks from ShuffleMapStage 0 (MapPartitionsRDD[4] at mapToPair at ProductSales.java:32)
20/06/24 22:27:14 INFO TaskSchedulerImpl: Adding task set 0.0 with 1 tasks
20/06/24 22:27:14 INFO TaskSetManager: Starting task 0.0 in stage 0.0 (TID 0, localhost, partition 0, PROCESS_LOCAL, 5974 bytes)
20/06/24 22:27:14 INFO Executor: Running task 0.0 in stage 0.0 (TID 0)
20/06/24 22:27:14 INFO Executor: Fetching spark://192.168.1.10:37891/jars/productjar.jar with timestamp 1593017830209
20/06/24 22:27:14 INFO TransportClientFactory: Successfully created connection to /192.168.1.10:37891 after 33 ms (0 ms spent in bootstraps)
20/06/24 22:27:14 INFO Utils: Fetching spark://192.168.1.10:37891/jars/productjar.jar to /tmp/spark-7a5eb18b-4e22-408a-b319-8435de7b48b7/user5
```


File Edit View Search Terminal Help

```
20/06/24 22:27:14 INFO Executor: Running task 0.0 in stage 0.0 (TID 0)
20/06/24 22:27:14 INFO Executor: Fetching spark://192.168.1.10:37891/jars/productjar.jar with timestamp 1593017830209
20/06/24 22:27:14 INFO TransportClientFactory: Successfully created connection to /192.168.1.10:37891 after 33 ms (0 ms spent in bootstraps)
20/06/24 22:27:14 INFO Utils: Fetching spark://192.168.1.10:37891/jars/productjar.jar to /tmp/spark-7a5eb18b-4e22-408a-b319-8435de7b48b7/userFiles-a789f4d3-2078-47f1-a9c0-5638c54734c6/fetchFileTemp1067789246415535096.tmp
20/06/24 22:27:14 INFO Executor: Adding file:/tmp/spark-7a5eb18b-4e22-408a-b319-8435de7b48b7/userFiles-a789f4d3-2078-47f1-a9c0-5638c54734c6/productjar.jar to class loader
20/06/24 22:27:14 INFO CodeGenerator: Code generated in 13.889186 ms
20/06/24 22:27:14 INFO FileScanRDD: Reading File path: file:///home/prerna/data/Retail_Sample_Data_Set.txt, range: 0-10309, partition values: [empty row]
20/06/24 22:27:14 INFO CodeGenerator: Code generated in 14.862754 ms
20/06/24 22:27:14 INFO Executor: Finished task 0.0 in stage 0.0 (TID 0). 1953 bytes result sent to driver
20/06/24 22:27:14 INFO TaskSetManager: Finished task 0.0 in stage 0.0 (TID 0) in 522 ms on localhost (1/1)
20/06/24 22:27:14 INFO TaskSchedulerImpl: Removed TaskSet 0.0, whose tasks have all completed, from pool
20/06/24 22:27:14 INFO DAGScheduler: ShuffleMapStage 0 (mapToPair at ProductSales.java:32) finished in 0.543 s
20/06/24 22:27:14 INFO DAGScheduler: looking for newly runnable stages
20/06/24 22:27:14 INFO DAGScheduler: running: Set()
20/06/24 22:27:14 INFO DAGScheduler: waiting: Set(ResultStage 1)
20/06/24 22:27:14 INFO DAGScheduler: failed: Set()
20/06/24 22:27:14 INFO DAGScheduler: Submitting ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at ProductSales.java:48), which has no missing parents
20/06/24 22:27:14 INFO MemoryStore: Block broadcast_2 stored as values in memory (estimated size 65.7 KB, free 366.0 MB)
20/06/24 22:27:14 INFO MemoryStore: Block broadcast_2_piece0 stored as bytes in memory (estimated size 23.5 KB, free 366.0 MB)
20/06/24 22:27:14 INFO BlockManagerInfo: Added broadcast_2_piece0 in memory on 192.168.1.10:36867 (size: 23.5 KB, free: 366.3 MB)
20/06/24 22:27:14 INFO SparkContext: Created broadcast 2 from broadcast at DAGScheduler.scala:1012
20/06/24 22:27:14 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at ProductSales.java:48)
20/06/24 22:27:14 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks
20/06/24 22:27:14 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0, ANY, 5232 bytes)
20/06/24 22:27:14 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)
20/06/24 22:27:14 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks
20/06/24 22:27:14 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 7 ms
20/06/24 22:27:15 INFO FileOutputCommitter: Saved output of task 'attempt_202006242227_0001_m_000000_1' to file:/home/prerna/spark-2.0.2-bin-hadoop2.6/ret-out-prod-01/_temporary/0/task_202006242227_0001_m_000000
20/06/24 22:27:15 INFO SparkHadoopMapRedUtil: attempt_202006242227_0001_m_000000_1: Committed
20/06/24 22:27:15 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 1808 bytes result sent to driver
20/06/24 22:27:15 INFO TaskSetManager: Finished task 0.0 in stage 1.0 (TID 1) in 118 ms on localhost (1/1)
20/06/24 22:27:15 INFO TaskSchedulerImpl: Removed TaskSet 1.0, whose tasks have all completed, from pool
```


prerna@prerna-Inspiron-5559: ~/spark-2.0.2-bin-hadoop2.6

File Edit View Search Terminal Help

```
20/06/24 22:27:14 INFO MemoryStore: Block broadcast_2_piece0 stored as bytes in memory (estimated size 23.5 KB, free 366.0 MB)
20/06/24 22:27:14 INFO BlockManagerInfo: Added broadcast_2_piece0 in memory on 192.168.1.10:36867 (size: 23.5 KB, free: 366.3 MB)
20/06/24 22:27:14 INFO SparkContext: Created broadcast 2 from broadcast at DAGScheduler.scala:1012
20/06/24 22:27:14 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at ProductSales.java:48)
20/06/24 22:27:14 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks
20/06/24 22:27:14 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0, ANY, 5232 bytes)
20/06/24 22:27:14 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)
20/06/24 22:27:14 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks
20/06/24 22:27:14 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 7 ms
20/06/24 22:27:15 INFO FileOutputCommitter: Saved output of task 'attempt_202006242227_0001_m_000000_1' to file:/home/prerna/spark-2.0.2-bin-hadoop2.6/ret-out-prod-01/_temporary/0/task_202006242227_0001_m_000000
20/06/24 22:27:15 INFO SparkHadoopMapRedUtil: attempt_202006242227_0001_m_000000_1: Committed
20/06/24 22:27:15 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 1808 bytes result sent to driver
20/06/24 22:27:15 INFO TaskSetManager: Finished task 0.0 in stage 1.0 (TID 1) in 118 ms on localhost (1/1)
20/06/24 22:27:15 INFO TaskSchedulerImpl: Removed TaskSet 1.0, whose tasks have all completed, from pool
20/06/24 22:27:15 INFO DAGScheduler: ResultStage 1 (saveAsTextFile at ProductSales.java:48) finished in 0.118 s
20/06/24 22:27:15 INFO DAGScheduler: Job 0 finished: saveAsTextFile at ProductSales.java:48, took 0.837729 s
20/06/24 22:27:15 INFO SparkUI: Stopped Spark web UI at http://192.168.1.10:4040
20/06/24 22:27:15 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
20/06/24 22:27:15 INFO MemoryStore: MemoryStore cleared
20/06/24 22:27:15 INFO BlockManager: BlockManager stopped
20/06/24 22:27:15 INFO BlockManagerMaster: BlockManagerMaster stopped
20/06/24 22:27:15 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
20/06/24 22:27:15 INFO SparkContext: Successfully stopped SparkContext
20/06/24 22:27:15 INFO ShutdownHookManager: Shutdown hook called
20/06/24 22:27:15 INFO ShutdownHookManager: Deleting directory /tmp/spark-7a5eb18b-4e22-408a-b319-8435de7b48b7
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ ls ret-out-prod-01/
part-000000 _SUCCESS
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ cat ret-out-prod-01/part-000000
(Men's Clothing,4030.89)
(Video Games,2573.38)
(Computers,2102.66)
(Women's Clothing,3736.8699)
(Sporting Goods,1952.89)
(Music,2396.4)
(CDs,2644.51)
(Garden,1882.25)
```


File Edit View Search Terminal Help

```
20/06/24 22:27:15 INFO BlockManagerMaster: BlockManagerMaster stopped
20/06/24 22:27:15 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
20/06/24 22:27:15 INFO SparkContext: Successfully stopped SparkContext
20/06/24 22:27:15 INFO ShutdownHookManager: Shutdown hook called
20/06/24 22:27:15 INFO ShutdownHookManager: Deleting directory /tmp/spark-7a5eb18b-4e22-408a-b319-8435de7b48b7
```

```
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ ls ret-out-prod-01/
```

```
part-000000 _SUCCESS
```

```
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ cat ret-out-prod-01/part-000000
```

```
(Men's Clothing,4030.89)
(Video Games,2573.38)
(Computers,2102.66)
(Women's Clothing,3736.8699)
(Sporting Goods,1952.89)
(Music,2396.4)
(CDs,2644.51)
(Garden,1882.25)
(Pet Supplies,2660.83)
(Cameras,2591.27)
(Books,3492.7998)
(Health and Beauty,2467.3198)
(Baby,2034.2297)
(Children's Clothing,2778.2102)
(Toys,3188.18)
(Consumer Electronics,2963.59)
(Crafts,3258.09)
(DVDs,2831.0002)
```

```
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ bin/spark-submit --class com.prerna.retail.StoreSales ../jars_file/salesjar.jar ../data/Retail_Sample_Data_Set.txt ret-out-store-01
```

```
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
```

```
20/06/24 22:27:50 INFO SparkContext: Running Spark version 2.0.2
```

```
20/06/24 22:27:50 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
```

```
20/06/24 22:27:50 WARN Utils: Your hostname, prerna-Inspiron-5559 resolves to a loopback address: 127.0.1.1; using 192.168.1.10 instead (on interface wlp1s0)
```

```
20/06/24 22:27:50 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
```

```
20/06/24 22:27:51 INFO SecurityManager: Changing view acls to: prerna
```

```
20/06/24 22:27:51 INFO SecurityManager: Changing modify acls to: prerna
```

```
20/06/24 22:27:51 INFO SecurityManager: Changing view acls groups to:
```

```
20/06/24 22:27:51 INFO SecurityManager: Changing modify acls groups to:
```


File Edit View Search Terminal Help

(DVDs,2831.0002)

prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6\$ bin/spark-submit --class com.prerna.retail.StoreSales ../jars_file/salesjar.jar ../data/Retail_Sample_Data_Set.txt ret-out-store-01

Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties

20/06/24 22:27:50 INFO SparkContext: Running Spark version 2.0.2

20/06/24 22:27:50 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

20/06/24 22:27:50 WARN Utils: Your hostname, prerna-Inspiron-5559 resolves to a loopback address: 127.0.1.1; using 192.168.1.10 instead (on interface wlp1s0)

20/06/24 22:27:50 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address

20/06/24 22:27:51 INFO SecurityManager: Changing view acls to: prerna

20/06/24 22:27:51 INFO SecurityManager: Changing modify acls to: prerna

20/06/24 22:27:51 INFO SecurityManager: Changing view acls groups to:

20/06/24 22:27:51 INFO SecurityManager: Changing modify acls groups to:

20/06/24 22:27:51 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(prerna); groups with view permissions: Set(); users with modify permissions: Set(prerna); groups with modify permissions: Set()

20/06/24 22:27:51 INFO Utils: Successfully started service 'sparkDriver' on port 35211.

20/06/24 22:27:51 INFO SparkEnv: Registering MapOutputTracker

20/06/24 22:27:51 INFO SparkEnv: Registering BlockManagerMaster

20/06/24 22:27:51 INFO DiskBlockManager: Created local directory at /tmp/blockmgr-22dec21c-087e-414e-90b9-a4072e4f7c38

20/06/24 22:27:51 INFO MemoryStore: MemoryStore started with capacity 366.3 MB

20/06/24 22:27:51 INFO SparkEnv: Registering OutputCommitCoordinator

20/06/24 22:27:51 INFO Utils: Successfully started service 'SparkUI' on port 4040.

20/06/24 22:27:51 INFO SparkUI: Bound SparkUI to 0.0.0.0, and started at http://192.168.1.10:4040

20/06/24 22:27:51 INFO SparkContext: Added JAR file:/home/prerna/spark-2.0.2-bin-hadoop2.6/../../jars_file/salesjar.jar at spark://192.168.1.10:35211/jars/salesjar.jar with timestamp 1593017871814

20/06/24 22:27:51 INFO Executor: Starting executor ID driver on host localhost

20/06/24 22:27:51 INFO Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port 37247.

20/06/24 22:27:51 INFO NettyBlockTransferService: Server created on 192.168.1.10:37247

20/06/24 22:27:51 INFO BlockManagerMaster: Registering BlockManager BlockManagerId(driver, 192.168.1.10, 37247)

20/06/24 22:27:51 INFO BlockManagerMasterEndpoint: Registering block manager 192.168.1.10:37247 with 366.3 MB RAM, BlockManagerId(driver, 192.168.1.10, 37247)

20/06/24 22:27:51 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 192.168.1.10, 37247)

20/06/24 22:27:52 WARN SparkContext: Use an existing SparkContext, some configuration may not take effect.

20/06/24 22:27:52 INFO SharedState: Warehouse path is 'file:/home/prerna/spark-2.0.2-bin-hadoop2.6/spark-warehouse'.

20/06/24 22:27:54 INFO FileSourceStrategy: Pruning directories with:

20/06/24 22:27:54 INFO FileSourceStrategy: Post-Scan Filters:

20/06/24 22:27:54 INFO FileSourceStrategy: Pruned Data Schema: struct<value: string>

20/06/24 22:27:54 INFO FileSourceStrategy: Pushed Filters:

File Edit View Search Terminal Help

```
20/06/24 22:27:51 INFO BlockManagerMaster: Registering block manager BlockManagerId(driver, 192.168.1.10, 37247)
20/06/24 22:27:51 INFO BlockManagerMasterEndpoint: Registering block manager 192.168.1.10:37247 with 366.3 MB RAM, BlockManagerId(driver, 192.168.1.10, 37247)
20/06/24 22:27:51 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 192.168.1.10, 37247)
20/06/24 22:27:52 WARN SparkContext: Use an existing SparkContext, some configuration may not take effect.
20/06/24 22:27:52 INFO SharedState: Warehouse path is 'file:/home/prerna/spark-2.0.2-bin-hadoop2.6/spark-warehouse'.
20/06/24 22:27:54 INFO FileSourceStrategy: Pruning directories with:
20/06/24 22:27:54 INFO FileSourceStrategy: Post-Scan Filters:
20/06/24 22:27:54 INFO FileSourceStrategy: Pruned Data Schema: struct<value: string>
20/06/24 22:27:54 INFO FileSourceStrategy: Pushed Filters:
20/06/24 22:27:55 INFO MemoryStore: Block broadcast_0 stored as values in memory (estimated size 214.4 KB, free 366.1 MB)
20/06/24 22:27:55 INFO MemoryStore: Block broadcast_0_piece0 stored as bytes in memory (estimated size 20.5 KB, free 366.1 MB)
20/06/24 22:27:55 INFO BlockManagerInfo: Added broadcast_0_piece0 in memory on 192.168.1.10:37247 (size: 20.5 KB, free: 366.3 MB)
20/06/24 22:27:55 INFO SparkContext: Created broadcast 0 from javaRDD at StoreSales.java:29
20/06/24 22:27:55 INFO FileSourceStrategy: Planning scan with bin packing, max size: 4194304 bytes, open cost is considered as scanning 4194304 bytes.
20/06/24 22:27:55 INFO CodeGenerator: Code generated in 239.811429 ms
20/06/24 22:27:55 INFO deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
20/06/24 22:27:55 INFO deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
20/06/24 22:27:55 INFO deprecation: mapred.task.is.map is deprecated. Instead, use mapreduce.task.ismap
20/06/24 22:27:55 INFO deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
20/06/24 22:27:55 INFO deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
20/06/24 22:27:55 INFO SparkContext: Starting job: saveAsTextFile at StoreSales.java:47
20/06/24 22:27:55 INFO DAGScheduler: Registering RDD 4 (mapToPair at StoreSales.java:31)
20/06/24 22:27:55 INFO DAGScheduler: Got job 0 (saveAsTextFile at StoreSales.java:47) with 1 output partitions
20/06/24 22:27:55 INFO DAGScheduler: Final stage: ResultStage 1 (saveAsTextFile at StoreSales.java:47)
20/06/24 22:27:55 INFO DAGScheduler: Parents of final stage: List(ShuffleMapStage 0)
20/06/24 22:27:55 INFO DAGScheduler: Missing parents: List(ShuffleMapStage 0)
20/06/24 22:27:55 INFO DAGScheduler: Submitting ShuffleMapStage 0 (MapPartitionsRDD[4] at mapToPair at StoreSales.java:31), which has no missing parents
20/06/24 22:27:56 INFO MemoryStore: Block broadcast_1 stored as values in memory (estimated size 9.8 KB, free 366.1 MB)
20/06/24 22:27:56 INFO MemoryStore: Block broadcast_1_piece0 stored as bytes in memory (estimated size 5.2 KB, free 366.1 MB)
20/06/24 22:27:56 INFO BlockManagerInfo: Added broadcast_1_piece0 in memory on 192.168.1.10:37247 (size: 5.2 KB, free: 366.3 MB)
20/06/24 22:27:56 INFO SparkContext: Created broadcast 1 from broadcast at DAGScheduler.scala:1012
20/06/24 22:27:56 INFO DAGScheduler: Submitting 1 missing tasks from ShuffleMapStage 0 (MapPartitionsRDD[4] at mapToPair at StoreSales.java:31)
20/06/24 22:27:56 INFO TaskSchedulerImpl: Adding task set 0.0 with 1 tasks
20/06/24 22:27:56 INFO TaskSetManager: Starting task 0.0 in stage 0.0 (TID 0, localhost, partition 0, PROCESS_LOCAL, 5972 bytes)
20/06/24 22:27:56 INFO Executor: Running task 0.0 in stage 0.0 (TID 0)
```


File Edit View Search Terminal Help

```
20/06/24 22:27:56 INFO TaskSchedulerImpl: Adding task set 0.0 with 1 tasks
20/06/24 22:27:56 INFO TaskSetManager: Starting task 0.0 in stage 0.0 (TID 0, localhost, partition 0, PROCESS_LOCAL, 5972 bytes)
20/06/24 22:27:56 INFO Executor: Running task 0.0 in stage 0.0 (TID 0)
20/06/24 22:27:56 INFO Executor: Fetching spark://192.168.1.10:35211/jars/salesjar.jar with timestamp 1593017871814
20/06/24 22:27:56 INFO TransportClientFactory: Successfully created connection to /192.168.1.10:35211 after 34 ms (0 ms spent in bootstraps)
20/06/24 22:27:56 INFO Utils: Fetching spark://192.168.1.10:35211/jars/salesjar.jar to /tmp/spark-71cd949a-99e3-419e-8cbd-c7599992bcab/userFiles-33157af5-e875-4afb-a7b0-bf98e1513b76/fetchFileTemp7107598854030102679.tmp
20/06/24 22:27:56 INFO Executor: Adding file:/tmp/spark-71cd949a-99e3-419e-8cbd-c7599992bcab/userFiles-33157af5-e875-4afb-a7b0-bf98e1513b76/salesjar.jar to class loader
20/06/24 22:27:56 INFO CodeGenerator: Code generated in 13.33283 ms
20/06/24 22:27:56 INFO FileScanRDD: Reading File path: file:///home/prerna/data/Retail_Sample_Data_Set.txt, range: 0-10309, partition values: [empty row]
20/06/24 22:27:56 INFO CodeGenerator: Code generated in 12.527225 ms
20/06/24 22:27:56 INFO Executor: Finished task 0.0 in stage 0.0 (TID 0). 1953 bytes result sent to driver
20/06/24 22:27:56 INFO TaskSetManager: Finished task 0.0 in stage 0.0 (TID 0) in 580 ms on localhost (1/1)
20/06/24 22:27:56 INFO TaskSchedulerImpl: Removed TaskSet 0.0, whose tasks have all completed, from pool
20/06/24 22:27:56 INFO DAGScheduler: ShuffleMapStage 0 (mapToPair at StoreSales.java:31) finished in 0.602 s
20/06/24 22:27:56 INFO DAGScheduler: looking for newly runnable stages
20/06/24 22:27:56 INFO DAGScheduler: running: Set()
20/06/24 22:27:56 INFO DAGScheduler: waiting: Set(ResultStage 1)
20/06/24 22:27:56 INFO DAGScheduler: failed: Set()
20/06/24 22:27:56 INFO DAGScheduler: Submitting ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at StoreSales.java:47), which has no missing parents
20/06/24 22:27:56 INFO MemoryStore: Block broadcast_2 stored as values in memory (estimated size 65.7 KB, free 366.0 MB)
20/06/24 22:27:56 INFO MemoryStore: Block broadcast_2_piece0 stored as bytes in memory (estimated size 23.5 KB, free 366.0 MB)
20/06/24 22:27:56 INFO BlockManagerInfo: Added broadcast_2_piece0 in memory on 192.168.1.10:37247 (size: 23.5 KB, free: 366.3 MB)
20/06/24 22:27:56 INFO SparkContext: Created broadcast 2 from broadcast at DAGScheduler.scala:1012
20/06/24 22:27:56 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at StoreSales.java:47)
20/06/24 22:27:56 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks
20/06/24 22:27:56 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0, ANY, 5230 bytes)
20/06/24 22:27:56 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)
20/06/24 22:27:56 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks
20/06/24 22:27:56 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 8 ms
20/06/24 22:27:56 INFO FileOutputCommitter: Saved output of task 'attempt_202006242227_0001_m_000000_1' to file:/home/prerna/spark-2.0.2-bin-hadoop2.6/ret-out-store-01/_temporary/0/task_202006242227_0001_m_000000
20/06/24 22:27:56 INFO SparkHadoopMapRedUtil: attempt_202006242227_0001_m_000000_1: Committed
20/06/24 22:27:56 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 1808 bytes result sent to driver
```


File Edit View Search Terminal Help

```
20/06/24 22:27:56 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at StoreSales.java:47)
20/06/24 22:27:56 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks
20/06/24 22:27:56 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0, ANY, 5230 bytes)
20/06/24 22:27:56 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)
20/06/24 22:27:56 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks
20/06/24 22:27:56 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 8 ms
20/06/24 22:27:56 INFO FileOutputCommitter: Saved output of task 'attempt_202006242227_0001_m_000000_1' to file:/home/prerna/spark-2.0.2-bin-hadoop2.6/ret-out-store-01/_temporary/0/task_202006242227_0001_m_000000
20/06/24 22:27:56 INFO SparkHadoopMapRedUtil: attempt_202006242227_0001_m_000000_1: Committed
20/06/24 22:27:56 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 1808 bytes result sent to driver
20/06/24 22:27:56 INFO DAGScheduler: ResultStage 1 (saveAsTextFile at StoreSales.java:47) finished in 0.153 s
20/06/24 22:27:56 INFO TaskSetManager: Finished task 0.0 in stage 1.0 (TID 1) in 152 ms on localhost (1/1)
20/06/24 22:27:56 INFO TaskSchedulerImpl: Removed TaskSet 1.0, whose tasks have all completed, from pool
20/06/24 22:27:56 INFO DAGScheduler: Job 0 finished: saveAsTextFile at StoreSales.java:47, took 0.930328 s
20/06/24 22:27:56 INFO SparkUI: Stopped Spark web UI at http://192.168.1.10:4040
20/06/24 22:27:56 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
20/06/24 22:27:56 INFO MemoryStore: MemoryStore cleared
20/06/24 22:27:56 INFO BlockManager: BlockManager stopped
20/06/24 22:27:56 INFO BlockManagerMaster: BlockManagerMaster stopped
20/06/24 22:27:56 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
20/06/24 22:27:56 INFO SparkContext: Successfully stopped SparkContext
20/06/24 22:27:56 INFO ShutdownHookManager: Shutdown hook called
20/06/24 22:27:56 INFO ShutdownHookManager: Deleting directory /tmp/spark-71cd949a-99e3-419e-8cbd-c7599992bcab
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ ls ret-out-store-01/
part-000000 _SUCCESS
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ cat ret-out-store-01/part-000000
(Albuquerque,1074.88)
(Pittsburgh,1271.35)
(Memphis,565.21)
(Garland,134.33)
(Seattle,934.39)
(Orlando,160.07)
(Dallas,273.49)
(Newark,39.75)
(Scottsdale,880.26)
(Anaheim,114.41)
(Fort Worth,1128.14)
```


File Edit View Search Terminal Help

part-00000 _SUCCESS

prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6\$ cat ret-out-store-01/part-00000

(Albuquerque,1074.88)
(Pittsburgh,1271.35)
(Memphis,565.21)
(Garland,134.33)
(Seattle,934.39)
(Orlando,160.07)
(Dallas,273.49)
(Newark,39.75)
(Scottsdale,880.26)
(Anaheim,114.41)
(Fort Worth,1128.14)
(Irvine,208.16)
(Fremont,465.66)
(Aurora,117.81)
(Reno,230.84999)
(Stockton,247.18)
(San Antonio,125.35)
(Rochester,75.73)
(Lincoln,712.76996)
(Corpus Christi,25.38)
(Jersey City,465.96)
(Omaha,1811.8899)
(Miami,331.75)
(Portland,108.69)
(San Jose,429.87)
(Raleigh,314.1)
(Austin,1787.88)
(Charlotte,440.11)
(Gilbert,374.00998)
(New York,468.90997)
(Kansas City,1093.66)
(Chicago,146.15)
(Chandler,1648.77)
(Buffalo,483.82)
(Winston-Salem,361.66)
(Detroit,134.89)

File Edit View Search Terminal Help

(Miami,331.75)
(Portland,108.69)
(San Jose,429.87)
(Raleigh,314.1)
(Austin,1787.88)
(Charlotte,440.11)
(Gilbert,374.00998)
(New York,468.90997)
(Kansas City,1093.66)
(Chicago,146.15)
(Chandler,1648.77)
(Buffalo,483.82)
(WinstonSalem,361.66)
(Detroit,134.89)
(San Francisco,260.65)
(Cincinnati,323.37997)
(Long Beach,98.23)
(Los Angeles,37.77)
(Bakersfield,217.79)
(Tulsa,431.95)
(Indianapolis,135.96)
(Atlanta,254.62)
(Wichita,460.38)
(Tucson,367.76)
(Denver,413.21002)
(Columbus,392.5)
(Washington,650.48)
(Chesapeake,676.35)
(Saint Paul,172.06)
(El Paso,103.01)
(Lubbock,1517.0801)
(Anchorage,1086.22)
(Phoenix,955.31)
(Spokane,448.97998)
(Santa Ana,922.30005)
(Baltimore,7.98)
(Greensboro,749.73)
(Cleveland,427.43)

File Edit View Search Terminal Help

(Phoenix,955.31)
(Spokane,448.97998)
(Santa Ana,922.30005)
(Baltimore,7.98)
(Greensboro,749.73)
(Cleveland,427.43)
(Milwaukee,418.91)
(Las Vegas,146.65)
(Lexington,1080.47)
(Norfolk,669.25)
(Arlington,400.08)
(Philadelphia,1166.76)
(Minneapolis,504.41998)
(Houston,1101.95)
(Toledo,605.23)
(North Las Vegas,409.26)
(New Orleans,1118.0801)
(Honolulu,717.57996)
(Fresno,1185.8701)
(Oakland,501.99)
(Boise,481.08997)
(Nashville,765.05)
(Mesa,384.8)
(Tampa,437.19)
(Glendale,220.63)
(Madison,515.48)
(Durham,980.32)
(San Diego,448.92)
(Boston,1114.54)
(Jacksonville,1330.1799)
(St. Petersburg,187.44)
(Fort Wayne,370.55)
(Riverside,1106.01)
(San Bernardino,270.58)
(Virginia Beach,647.67)

prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6\$ bin/spark-submit --class com.prerna.retail.TotalSales ../jars_file/totalsalesjar.jar
../data/Retail_Sample_Data_Set.txt ret-out-total-01
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties

File Edit View Search Terminal Help

```
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ bin/spark-submit --class com.prerna.retail.TotalSales ../jars_file/totalsalesjar.jar
../data/Retail_Sample_Data_Set.txt ret-out-total-01
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
20/06/24 22:28:27 INFO SparkContext: Running Spark version 2.0.2
20/06/24 22:28:28 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
20/06/24 22:28:28 WARN Utils: Your hostname, prerna-Inspiron-5559 resolves to a loopback address: 127.0.1.1; using 192.168.1.10 instead (on in
terface wlp1s0)
20/06/24 22:28:28 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
20/06/24 22:28:28 INFO SecurityManager: Changing view acls to: prerna
20/06/24 22:28:28 INFO SecurityManager: Changing modify acls to: prerna
20/06/24 22:28:28 INFO SecurityManager: Changing view acls groups to:
20/06/24 22:28:28 INFO SecurityManager: Changing modify acls groups to:
20/06/24 22:28:28 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(prerna);
groups with view permissions: Set(); users with modify permissions: Set(prerna); groups with modify permissions: Set()
20/06/24 22:28:28 INFO Utils: Successfully started service 'sparkDriver' on port 42199.
20/06/24 22:28:28 INFO SparkEnv: Registering MapOutputTracker
20/06/24 22:28:28 INFO SparkEnv: Registering BlockManagerMaster
20/06/24 22:28:28 INFO DiskBlockManager: Created local directory at /tmp/blockmgr-2a066729-fd10-45ae-be9d-ab566574195c
20/06/24 22:28:28 INFO MemoryStore: MemoryStore started with capacity 366.3 MB
20/06/24 22:28:28 INFO SparkEnv: Registering OutputCommitCoordinator
20/06/24 22:28:29 INFO Utils: Successfully started service 'SparkUI' on port 4040.
20/06/24 22:28:29 INFO SparkUI: Bound SparkUI to 0.0.0.0, and started at http://192.168.1.10:4040
20/06/24 22:28:29 INFO SparkContext: Added JAR file:/home/prerna/spark-2.0.2-bin-hadoop2.6/../../jars_file/totalsalesjar.jar at spark://192.168.1
.10:42199/jars/totalsalesjar.jar with timestamp 1593017909244
20/06/24 22:28:29 INFO Executor: Starting executor ID driver on host localhost
20/06/24 22:28:29 INFO Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port 42585.
20/06/24 22:28:29 INFO NettyBlockTransferService: Server created on 192.168.1.10:42585
20/06/24 22:28:29 INFO BlockManagerMaster: Registering BlockManager BlockManagerId(driver, 192.168.1.10, 42585)
20/06/24 22:28:29 INFO BlockManagerMasterEndpoint: Registering block manager 192.168.1.10:42585 with 366.3 MB RAM, BlockManagerId(driver, 192.
168.1.10, 42585)
20/06/24 22:28:29 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 192.168.1.10, 42585)
20/06/24 22:28:29 WARN SparkContext: Use an existing SparkContext, some configuration may not take effect.
20/06/24 22:28:29 INFO SharedState: Warehouse path is 'file:/home/prerna/spark-2.0.2-bin-hadoop2.6/spark-warehouse'.
20/06/24 22:28:32 INFO FileSourceStrategy: Pruning directories with:
20/06/24 22:28:32 INFO FileSourceStrategy: Post-Scan Filters:
20/06/24 22:28:32 INFO FileSourceStrategy: Pruned Data Schema: struct<value: string>
20/06/24 22:28:32 INFO FileSourceStrategy: Pushed Filters:
20/06/24 22:28:32 INFO MemoryStore: Block broadcast_0 stored as values in memory (estimated size 214.4 KB, free 366.1 MB)
```



```
20/06/24 22:28:29 INFO BlockManagerMaster: Registered BlockManager BlockManagerId(driver, 192.168.1.10, 42585)
20/06/24 22:28:29 WARN SparkContext: Use an existing SparkContext, some configuration may not take effect.
20/06/24 22:28:29 INFO SharedState: Warehouse path is 'file:/home/prerna/spark-2.0.2-bin-hadoop2.6/spark-warehouse'.
20/06/24 22:28:32 INFO FileSourceStrategy: Pruning directories with:
20/06/24 22:28:32 INFO FileSourceStrategy: Post-Scan Filters:
20/06/24 22:28:32 INFO FileSourceStrategy: Pruned Data Schema: struct<value: string>
20/06/24 22:28:32 INFO FileSourceStrategy: Pushed Filters:
20/06/24 22:28:32 INFO MemoryStore: Block broadcast_0 stored as values in memory (estimated size 214.4 KB, free 366.1 MB)
20/06/24 22:28:32 INFO MemoryStore: Block broadcast_0_piece0 stored as bytes in memory (estimated size 20.5 KB, free 366.1 MB)
20/06/24 22:28:32 INFO BlockManagerInfo: Added broadcast_0_piece0 in memory on 192.168.1.10:42585 (size: 20.5 KB, free: 366.3 MB)
20/06/24 22:28:32 INFO SparkContext: Created broadcast 0 from javaRDD at TotalSales.java:24
20/06/24 22:28:32 INFO FileSourceStrategy: Planning scan with bin packing, max size: 4194304 bytes, open cost is considered as scanning 4194304 bytes.
20/06/24 22:28:33 INFO CodeGenerator: Code generated in 239.500012 ms
20/06/24 22:28:33 INFO deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
20/06/24 22:28:33 INFO deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
20/06/24 22:28:33 INFO deprecation: mapred.task.is.map is deprecated. Instead, use mapreduce.task.ismap
20/06/24 22:28:33 INFO deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
20/06/24 22:28:33 INFO deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
20/06/24 22:28:33 INFO SparkContext: Starting job: saveAsTextFile at TotalSales.java:42
20/06/24 22:28:33 INFO DAGScheduler: Registering RDD 4 (mapToPair at TotalSales.java:26)
20/06/24 22:28:33 INFO DAGScheduler: Got job 0 (saveAsTextFile at TotalSales.java:42) with 1 output partitions
20/06/24 22:28:33 INFO DAGScheduler: Final stage: ResultStage 1 (saveAsTextFile at TotalSales.java:42)
20/06/24 22:28:33 INFO DAGScheduler: Parents of final stage: List(ShuffleMapStage 0)
20/06/24 22:28:33 INFO DAGScheduler: Missing parents: List(ShuffleMapStage 0)
20/06/24 22:28:33 INFO DAGScheduler: Submitting ShuffleMapStage 0 (MapPartitionsRDD[4] at mapToPair at TotalSales.java:26), which has no missing parents
20/06/24 22:28:33 INFO MemoryStore: Block broadcast_1 stored as values in memory (estimated size 9.8 KB, free 366.1 MB)
20/06/24 22:28:33 INFO MemoryStore: Block broadcast_1_piece0 stored as bytes in memory (estimated size 5.2 KB, free 366.1 MB)
20/06/24 22:28:33 INFO BlockManagerInfo: Added broadcast_1_piece0 in memory on 192.168.1.10:42585 (size: 5.2 KB, free: 366.3 MB)
20/06/24 22:28:33 INFO SparkContext: Created broadcast 1 from broadcast at DAGScheduler.scala:1012
20/06/24 22:28:33 INFO DAGScheduler: Submitting 1 missing tasks from ShuffleMapStage 0 (MapPartitionsRDD[4] at mapToPair at TotalSales.java:26)
20/06/24 22:28:33 INFO TaskSchedulerImpl: Adding task set 0.0 with 1 tasks
20/06/24 22:28:33 INFO TaskSetManager: Starting task 0.0 in stage 0.0 (TID 0, localhost, partition 0, PROCESS_LOCAL, 5977 bytes)
20/06/24 22:28:33 INFO Executor: Running task 0.0 in stage 0.0 (TID 0)
20/06/24 22:28:33 INFO Executor: Fetching spark://192.168.1.10:42199/jars/totalsalesjar.jar with timestamp 1593017909244
20/06/24 22:28:33 INFO TransportClientFactory: Successfully created connection to /192.168.1.10:42199 after 33 ms (0 ms spent in bootstraps)
```


File Edit View Search Terminal Help

```
20/06/24 22:28:33 INFO TaskSchedulerImpl: Adding task set 0.0 with 1 tasks
20/06/24 22:28:33 INFO TaskSetManager: Starting task 0.0 in stage 0.0 (TID 0, localhost, partition 0, PROCESS_LOCAL, 5977 bytes)
20/06/24 22:28:33 INFO Executor: Running task 0.0 in stage 0.0 (TID 0)
20/06/24 22:28:33 INFO Executor: Fetching spark://192.168.1.10:42199/jars/totalsalesjar.jar with timestamp 1593017909244
20/06/24 22:28:33 INFO TransportClientFactory: Successfully created connection to /192.168.1.10:42199 after 33 ms (0 ms spent in bootstraps)
20/06/24 22:28:33 INFO Utils: Fetching spark://192.168.1.10:42199/jars/totalsalesjar.jar to /tmp/spark-01e7199f-2496-4e15-99b1-43c886249cf5/userFiles-b433283c-5d86-48b6-95e1-85cf0ae1e3eb/fetchFileTemp2904619657743979480.tmp
20/06/24 22:28:33 INFO Executor: Adding file:/tmp/spark-01e7199f-2496-4e15-99b1-43c886249cf5/userFiles-b433283c-5d86-48b6-95e1-85cf0ae1e3eb/totalsalesjar.jar to class loader
20/06/24 22:28:33 INFO CodeGenerator: Code generated in 13.783722 ms
20/06/24 22:28:33 INFO FileScanRDD: Reading File path: file:///home/prerna/data/Retail_Sample_Data_Set.txt, range: 0-10309, partition values: [empty row]
20/06/24 22:28:33 INFO CodeGenerator: Code generated in 13.411306 ms
20/06/24 22:28:34 INFO Executor: Finished task 0.0 in stage 0.0 (TID 0). 1953 bytes result sent to driver
20/06/24 22:28:34 INFO TaskSetManager: Finished task 0.0 in stage 0.0 (TID 0) in 533 ms on localhost (1/1)
20/06/24 22:28:34 INFO DAGScheduler: ShuffleMapStage 0 (mapToPair at TotalSales.java:26) finished in 0.560 s
20/06/24 22:28:34 INFO DAGScheduler: looking for newly runnable stages
20/06/24 22:28:34 INFO DAGScheduler: running: Set()
20/06/24 22:28:34 INFO DAGScheduler: waiting: Set(ResultStage 1)
20/06/24 22:28:34 INFO DAGScheduler: failed: Set()
20/06/24 22:28:34 INFO DAGScheduler: Submitting ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at TotalSales.java:42), which has no missing parents
20/06/24 22:28:34 INFO TaskSchedulerImpl: Removed TaskSet 0.0, whose tasks have all completed, from pool
20/06/24 22:28:34 INFO MemoryStore: Block broadcast_2 stored as values in memory (estimated size 65.7 KB, free 366.0 MB)
20/06/24 22:28:34 INFO MemoryStore: Block broadcast_2_piece0 stored as bytes in memory (estimated size 23.5 KB, free 366.0 MB)
20/06/24 22:28:34 INFO BlockManagerInfo: Added broadcast_2_piece0 in memory on 192.168.1.10:42585 (size: 23.5 KB, free: 366.3 MB)
20/06/24 22:28:34 INFO SparkContext: Created broadcast 2 from broadcast at DAGScheduler.scala:1012
20/06/24 22:28:34 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at TotalSales.java:42)
20/06/24 22:28:34 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks
20/06/24 22:28:34 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0, ANY, 5235 bytes)
20/06/24 22:28:34 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)
20/06/24 22:28:34 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks
20/06/24 22:28:34 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 7 ms
20/06/24 22:28:34 INFO FileOutputCommitter: Saved output of task 'attempt_202006242228_0001_m_000000_1' to file:/home/prerna/spark-2.0.2-bin-hadoop2.6/ret-out-total-01/_temporary/0/task_202006242228_0001_m_000000
20/06/24 22:28:34 INFO SparkHadoopMapRedUtil: attempt_202006242228_0001_m_000000_1: Committed
20/06/24 22:28:34 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 1808 bytes result sent to driver
20/06/24 22:28:34 INFO TaskSetManager: Finished task 0.0 in stage 1.0 (TID 1) in 124 ms on localhost (1/1)
```


File Edit View Search Terminal Help

```
20/06/24 22:28:34 INFO TaskSetManager: Finished task 0.0 in stage 0.0 (TID 0) in 533 ms on localhost (1/1)
20/06/24 22:28:34 INFO DAGScheduler: ShuffleMapStage 0 (mapToPair at TotalSales.java:26) finished in 0.560 s
20/06/24 22:28:34 INFO DAGScheduler: looking for newly runnable stages
20/06/24 22:28:34 INFO DAGScheduler: running: Set()
20/06/24 22:28:34 INFO DAGScheduler: waiting: Set(ResultStage 1)
20/06/24 22:28:34 INFO DAGScheduler: failed: Set()
20/06/24 22:28:34 INFO DAGScheduler: Submitting ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at TotalSales.java:42), which has no missing parents
20/06/24 22:28:34 INFO TaskSchedulerImpl: Removed TaskSet 0.0, whose tasks have all completed, from pool
20/06/24 22:28:34 INFO MemoryStore: Block broadcast_2 stored as values in memory (estimated size 65.7 KB, free 366.0 MB)
20/06/24 22:28:34 INFO MemoryStore: Block broadcast_2_piece0 stored as bytes in memory (estimated size 23.5 KB, free 366.0 MB)
20/06/24 22:28:34 INFO BlockManagerInfo: Added broadcast_2_piece0 in memory on 192.168.1.10:42585 (size: 23.5 KB, free: 366.3 MB)
20/06/24 22:28:34 INFO SparkContext: Created broadcast 2 from broadcast at DAGScheduler.scala:1012
20/06/24 22:28:34 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at TotalSales.java:42)
20/06/24 22:28:34 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks
20/06/24 22:28:34 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0, ANY, 5235 bytes)
20/06/24 22:28:34 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)
20/06/24 22:28:34 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks
20/06/24 22:28:34 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 7 ms
20/06/24 22:28:34 INFO FileOutputCommitter: Saved output of task 'attempt_202006242228_0001_m_000000_1' to file:/home/prerna/spark-2.0.2-bin-hadoop2.6/ret-out-total-01/_temporary/0/task_202006242228_0001_m_000000
20/06/24 22:28:34 INFO SparkHadoopMapRedUtil: attempt_202006242228_0001_m_000000_1: Committed
20/06/24 22:28:34 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 1808 bytes result sent to driver
20/06/24 22:28:34 INFO TaskSetManager: Finished task 0.0 in stage 1.0 (TID 1) in 124 ms on localhost (1/1)
20/06/24 22:28:34 INFO TaskSchedulerImpl: Removed TaskSet 1.0, whose tasks have all completed, from pool
20/06/24 22:28:34 INFO DAGScheduler: ResultStage 1 (saveAsTextFile at TotalSales.java:42) finished in 0.125 s
20/06/24 22:28:34 INFO DAGScheduler: Job 0 finished: saveAsTextFile at TotalSales.java:42, took 0.883266 s
20/06/24 22:28:34 INFO SparkUI: Stopped Spark web UI at http://192.168.1.10:4040
20/06/24 22:28:34 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
20/06/24 22:28:34 INFO MemoryStore: MemoryStore cleared
20/06/24 22:28:34 INFO BlockManager: BlockManager stopped
20/06/24 22:28:34 INFO BlockManagerMaster: BlockManagerMaster stopped
20/06/24 22:28:34 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
20/06/24 22:28:34 INFO SparkContext: Successfully stopped SparkContext
20/06/24 22:28:34 INFO ShutdownHookManager: Shutdown hook called
20/06/24 22:28:34 INFO ShutdownHookManager: Deleting directory /tmp/spark-01e7199f-2496-4e15-99b1-43c886249cf5
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ ls ret-out-total-01/
```


File Edit View Search Terminal Help

```
20/06/24 22:28:34 INFO DAGScheduler: waiting: Set(ResultStage 1)
20/06/24 22:28:34 INFO DAGScheduler: failed: Set()
20/06/24 22:28:34 INFO DAGScheduler: Submitting ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at TotalSales.java:42), which has no missing parents
20/06/24 22:28:34 INFO TaskSchedulerImpl: Removed TaskSet 0.0, whose tasks have all completed, from pool
20/06/24 22:28:34 INFO MemoryStore: Block broadcast_2 stored as values in memory (estimated size 65.7 KB, free 366.0 MB)
20/06/24 22:28:34 INFO MemoryStore: Block broadcast_2_piece0 stored as bytes in memory (estimated size 23.5 KB, free 366.0 MB)
20/06/24 22:28:34 INFO BlockManagerInfo: Added broadcast_2_piece0 in memory on 192.168.1.10:42585 (size: 23.5 KB, free: 366.3 MB)
20/06/24 22:28:34 INFO SparkContext: Created broadcast 2 from broadcast at DAGScheduler.scala:1012
20/06/24 22:28:34 INFO DAGScheduler: Submitting 1 missing tasks from ResultStage 1 (MapPartitionsRDD[6] at saveAsTextFile at TotalSales.java:42)
20/06/24 22:28:34 INFO TaskSchedulerImpl: Adding task set 1.0 with 1 tasks
20/06/24 22:28:34 INFO TaskSetManager: Starting task 0.0 in stage 1.0 (TID 1, localhost, partition 0, ANY, 5235 bytes)
20/06/24 22:28:34 INFO Executor: Running task 0.0 in stage 1.0 (TID 1)
20/06/24 22:28:34 INFO ShuffleBlockFetcherIterator: Getting 1 non-empty blocks out of 1 blocks
20/06/24 22:28:34 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 7 ms
20/06/24 22:28:34 INFO FileOutputCommitter: Saved output of task 'attempt_202006242228_0001_m_000000_1' to file:/home/prerna/spark-2.0.2-bin-hadoop2.6/ret-out-total-01/_temporary/0/task_202006242228_0001_m_000000
20/06/24 22:28:34 INFO SparkHadoopMapRedUtil: attempt_202006242228_0001_m_000000_1: Committed
20/06/24 22:28:34 INFO Executor: Finished task 0.0 in stage 1.0 (TID 1). 1808 bytes result sent to driver
20/06/24 22:28:34 INFO TaskSetManager: Finished task 0.0 in stage 1.0 (TID 1) in 124 ms on localhost (1/1)
20/06/24 22:28:34 INFO TaskSchedulerImpl: Removed TaskSet 1.0, whose tasks have all completed, from pool
20/06/24 22:28:34 INFO DAGScheduler: ResultStage 1 (saveAsTextFile at TotalSales.java:42) finished in 0.125 s
20/06/24 22:28:34 INFO DAGScheduler: Job 0 finished: saveAsTextFile at TotalSales.java:42, took 0.883266 s
20/06/24 22:28:34 INFO SparkUI: Stopped Spark web UI at http://192.168.1.10:4040
20/06/24 22:28:34 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
20/06/24 22:28:34 INFO MemoryStore: MemoryStore cleared
20/06/24 22:28:34 INFO BlockManager: BlockManager stopped
20/06/24 22:28:34 INFO BlockManagerMaster: BlockManagerMaster stopped
20/06/24 22:28:34 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
20/06/24 22:28:34 INFO SparkContext: Successfully stopped SparkContext
20/06/24 22:28:34 INFO ShutdownHookManager: Shutdown hook called
20/06/24 22:28:34 INFO ShutdownHookManager: Deleting directory /tmp/spark-01e7199f-2496-4e15-99b1-43c886249cf5
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ ls ret-out-total-01/
part-00000 _SUCCESS
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$ cat ret-out-total-01/part-00000
(Total Sales,49585.37)
prerna@prerna-Inspiron-5559:~/spark-2.0.2-bin-hadoop2.6$
```