

Komuravellyv1@udayton.edu

1015237060

Installed Hadoop 2.7.1 with Spark 2.1.2

Started Hadoop, Spark on top of hadoop.

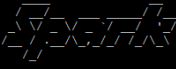
```
hduser@ip-172-31-29-213: ~
hduser@ip-172-31-29-213:~$ cd /usr/local/hadoop/etc/hadoop
hduser@ip-172-31-29-213:/usr/local/hadoop/etc/hadoop$ ls
capacity-scheduler.xml  hadoop-env.sh          httpfs-env.sh          kms-env.sh             mapred-env.sh          ssl-client.xml.example
configuration.xml       hadoop-metrics2.properties  httpfs-log4j.properties  kms-log4j.properties  mapred-queues.xml.template  ssl-server.xml.example
container-executor.cfg  hadoop-metrics.properties  httpfs-signature.secret  kms-site.xml           mapred-site.xml        yarn-env.cmd
core-site.xml           hadoop-policy.xml         httpfs-site.xml         log4j.properties      masters                yarn-env.sh
hadoop-env.cmd          hdfs-site.xml            kms-acls.xml            mapred-env.cmd         slaves                 yarn-site.xml

hduser@ip-172-31-29-213:/usr/local/hadoop/etc/hadoop$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [ip-172-31-29-213.us-west-2.compute.internal]
hduser@ip-172-31-29-213.us-west-2.compute.internal's password:
ip-172-31-29-213.us-west-2.compute.internal: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hduser-namenode-ip-172-31-29-213.out
172.31.16.8: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser-datanode-ip-172-31-16-8.out
172.31.29.75: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser-datanode-ip-172-31-29-75.out
Starting secondary namenodes [ip-172-31-17-109.us-west-2.compute.internal]
ip-172-31-17-109.us-west-2.compute.internal: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-hduser-secondarynamenode-ip-172-31-17-109.out
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-hduser-resourcemanager-ip-172-31-29-213.out
172.31.29.75: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hduser-nodemanager-ip-172-31-29-75.out
172.31.16.8: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hduser-nodemanager-ip-172-31-16-8.out
hduser@ip-172-31-29-213:/usr/local/hadoop/etc/hadoop$ start-yarn.sh
starting yarn daemons
resourcemanager running as process 1798. Stop it first.
172.31.16.8: nodemanager running as process 1518. Stop it first.
172.31.29.75: nodemanager running as process 1524. Stop it first.
hduser@ip-172-31-29-213:/usr/local/hadoop/etc/hadoop$ jps
2132 Jps
1525 NameNode
1798 ResourceManager
hduser@ip-172-31-29-213:/usr/local/hadoop/etc/hadoop$ cd
hduser@ip-172-31-29-213:~$ $SPARK_HOME/sbin/start-all.sh
starting org.apache.spark.deploy.master.Master, logging to /usr/local/spark/logs/spark-hduser-org.apache.spark.deploy.master.Master-1-ip-172-31-29-213.out
172.31.16.8: starting org.apache.spark.deploy.worker.Worker, logging to /usr/local/spark/logs/spark-hduser-org.apache.spark.deploy.worker.Worker-1-ip-172-31-16-8.out
172.31.29.75: starting org.apache.spark.deploy.worker.Worker, logging to /usr/local/spark/logs/spark-hduser-org.apache.spark.deploy.worker.Worker-1-ip-172-31-29-75.out
hduser@ip-172-31-29-213:~$ jps
1525 NameNode
1798 ResourceManager
2158 Master
2223 Jps
hduser@ip-172-31-29-213:~$
```

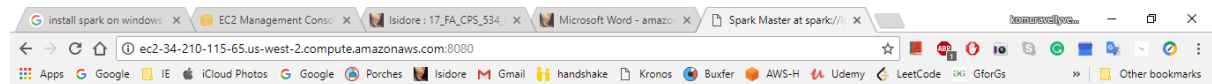
```
hduser@ip-172-31-29-213:~$ jps
1525 NameNode
1798 ResourceManager
2158 Master
2223 Jps
hduser@ip-172-31-29-213:~$ hadoop -fs ls
Error: No command named '-fs' was found. Perhaps you meant 'hadoop fs'
hduser@ip-172-31-29-213:~$ hadoop fs -ls
Found 2 items
drwxr-xr-x - hduser supergroup 0 2017-12-04 21:43 kmeansv
drwxr-xr-x - hduser supergroup 0 2017-12-04 21:57 venkat
hduser@ip-172-31-29-213:~$ hadoop fs -put /home/ubuntu/iris.txt /user/hduser/venkat
put: '/home/ubuntu/iris.txt': No such file or directory
hduser@ip-172-31-29-213:~$ hadoop fs -put /home/hduser/iris.txt /user/hduser/venkat
hduser@ip-172-31-29-213:~$ hadoop fs -lsr
lsr: DEPRECATED: Please use 'ls -R' instead.
drwxr-xr-x - hduser supergroup 0 2017-12-04 21:43 kmeansv
drwxr-xr-x - hduser supergroup 0 2017-12-05 04:15 venkat
-rw-r--r-- 2 hduser supergroup 576 2017-12-04 21:57 venkat/input_kmeans.txt
-rw-r--r-- 2 hduser supergroup 4700 2017-12-05 04:15 venkat/iris.txt
hduser@ip-172-31-29-213:~$
```

Komuravellyv1@udayton.edu
1015237060

```
hduser@ip-172-31-29-213: ~  
6.7,2.5,5.8,1.8,Iris-virginica  
7.2,3.6,6.1,2.5,Iris-virginica  
6.5,3.2,5.1,2.0,Iris-virginica  
6.4,2.7,5.3,1.9,Iris-virginica  
6.8,3.0,5.5,2.1,Iris-virginica  
5.7,2.5,5.0,2.0,Iris-virginica  
5.8,2.8,5.1,2.4,Iris-virginica  
6.4,3.2,5.3,2.3,Iris-virginica  
6.5,3.0,5.5,1.8,Iris-virginica  
7.7,3.8,6.7,2.2,Iris-virginica  
7.7,2.6,6.9,2.3,Iris-virginica  
6.0,2.2,5.0,1.5,Iris-virginica  
6.9,3.2,5.7,2.3,Iris-virginica  
5.6,2.8,4.9,2.0,Iris-virginica  
7.7,2.8,6.7,2.0,Iris-virginica  
6.3,2.7,4.9,1.8,Iris-virginica  
6.7,3.3,5.7,2.1,Iris-virginica  
7.2,3.2,6.0,1.8,Iris-virginica  
6.2,2.8,4.8,1.8,Iris-virginica  
6.1,3.0,4.9,1.8,Iris-virginica  
6.4,2.8,5.6,2.1,Iris-virginica  
7.2,3.0,5.8,1.6,Iris-virginica  
7.4,2.8,6.1,1.9,Iris-virginica  
7.9,3.8,6.4,2.0,Iris-virginica  
6.4,2.8,5.6,2.2,Iris-virginica  
6.3,2.8,5.1,1.5,Iris-virginica  
6.1,2.6,5.6,1.4,Iris-virginica  
7.7,3.0,6.1,2.3,Iris-virginica  
6.3,3.4,5.6,2.4,Iris-virginica  
6.4,3.1,5.5,1.8,Iris-virginica  
6.0,3.0,4.8,1.8,Iris-virginica  
6.9,3.1,5.4,2.1,Iris-virginica  
6.7,3.1,5.6,2.4,Iris-virginica  
6.9,3.1,5.1,2.3,Iris-virginica  
5.8,2.7,5.1,1.9,Iris-virginica  
6.8,3.2,5.9,2.3,Iris-virginica  
6.7,3.3,5.7,2.5,Iris-virginica  
6.7,3.0,5.2,2.3,Iris-virginica  
6.8,3.2,5.9,2.3,Iris-virginica  
6.5,3.0,5.2,2.0,Iris-virginica  
6.2,3.4,5.4,2.3,Iris-virginica  
5.9,3.0,5.1,1.8,Iris-virginica  
hduser@ip-172-31-29-213:~$
```

```
hduser@ip-172-31-29-213: /usr/local/spark/bin  
6.3,2.8,5.1,1.8,Iris-virginica  
6.1,2.6,5.6,1.4,Iris-virginica  
7.7,3.0,6.1,2.3,Iris-virginica  
6.3,3.4,5.6,2.4,Iris-virginica  
6.4,3.1,5.5,1.8,Iris-virginica  
6.0,3.0,4.8,1.8,Iris-virginica  
6.9,3.1,5.4,2.1,Iris-virginica  
6.7,3.1,5.6,2.4,Iris-virginica  
6.9,3.1,5.1,2.3,Iris-virginica  
5.8,2.7,5.1,1.9,Iris-virginica  
6.8,3.2,5.9,2.3,Iris-virginica  
6.7,3.3,5.7,2.5,Iris-virginica  
6.7,3.0,5.2,2.3,Iris-virginica  
6.3,2.5,5.0,1.9,Iris-virginica  
6.5,3.0,5.2,2.0,Iris-virginica  
6.2,3.4,5.4,2.3,Iris-virginica  
5.9,3.0,5.1,1.8,Iris-virginica  
hduser@ip-172-31-29-213:~$ cd  
hduser@ip-172-31-29-213:~$ cd /usr/local  
hduser@ip-172-31-29-213:/usr/local$ ls  
bin etc games hadoop include lib man sbin share spark src  
hduser@ip-172-31-29-213:/usr/local$ cd spark  
hduser@ip-172-31-29-213:/usr/local/spark$ ls  
bin conf data examples jars LICENSE licenses logs NOTICE python R README.md RELEASE sbin yarn  
hduser@ip-172-31-29-213:/usr/local/spark$ cd bin  
hduser@ip-172-31-29-213:/usr/local/spark/bin$ pyspark  
Python 3.5.2 (default, Nov 23 2017, 16:37:01)  
[GCC 5.4.0 20160609] on linux  
Type "help", "copyright", "credits" or "license()" for more information.  
Setting default log level to "WARN".  
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).  
17/12/05 04:17:57 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable  
17/12/05 04:18:07 WARN metastore.ObjectStore: Failed to get database global_temp, returning NoSuchObjectException  
Welcome to  
 version 2.1.2  
Using Python version 3.5.2 (default, Nov 23 2017 16:37:01)  
SparkSession available as 'spark'.  
>>>
```

Komuravellyv1@udayton.edu
1015237060



Spark Master at spark://ip-172-31-29-213.us-west-2.compute.internal:7077

URL: spark://ip-172-31-29-213.us-west-2.compute.internal:7077
REST URL: spark://ip-172-31-29-213.us-west-2.compute.internal:6066 (cluster mode)
Alive Workers: 2
Cores in use: 2 Total, 0 Used
Memory in use: 2.0 GB Total, 0.0 B Used
Applications: 0 Running, 0 Completed
Drivers: 0 Running, 0 Completed
Status: ALIVE

Workers

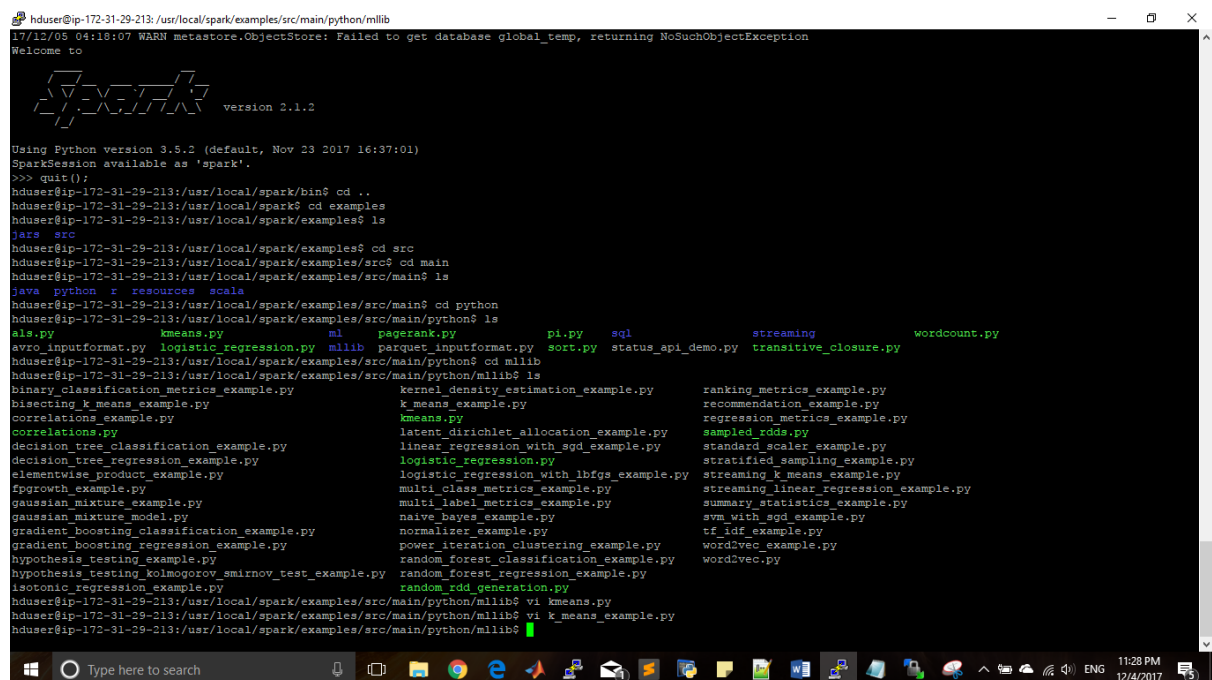
Worker Id	Address	State	Cores	Memory
worker-20171205040939-172.31.16.8-45943	172.31.16.8:45943	ALIVE	1 (0 Used)	1024.0 MB (0.0 B Used)
worker-20171205040939-172.31.29.75-40193	172.31.29.75:40193	ALIVE	1 (0 Used)	1024.0 MB (0.0 B Used)

Running Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
----------------	------	-------	-----------------	----------------	------	-------	----------

Completed Applications

Application ID	Name	Cores	Memory per Node	Submitted Time	User	State	Duration
----------------	------	-------	-----------------	----------------	------	-------	----------



1015237060

```
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python/mllib
from __future__ import print_function

# $example on$
from numpy import array
from math import sqrt
# $example off$

from pyspark import SparkContext

# $example on$
from pyspark.mllib.clustering import KMeans, KMeansModel
# $example off$

if __name__ == "__main__":
    sc = SparkContext(appName="KMeansExample") # SparkContext

    # $example on$
    # Load and parse the data
    data = sc.textFile("/user/hduser/venkat/iris.txt")
    parsedData = data.map(lambda line: array([float(x) for x in line.split(' ')]))

    # Build the model (cluster the data)
    clusters = KMeans.train(parsedData, 2, maxIterations=10, initializationMode="random")

    # Evaluate clustering by computing Within Set Sum of Squared Errors
    def error(point):
        center = clusters.centers[clusters.predict(point)]
        return sqrt(sum([x**2 for x in (point - center)]))

    WSSSE = parsedData.map(lambda point: error(point)).reduce(lambda x, y: x + y)
    print("Within Set Sum of Squared Error = " + str(WSSSE))

    # Save and load model
    clusters.save(sc, "target/org/apache/spark/PythonKMeansExample/KMeansModel")
    sameModel = KMeansModel.load(sc, "target/org/apache/spark/PythonKMeansExample/KMeansModel")
    myModelPath = "/user/hduser/kmeans_output"
    clusters.save(sc, myModelPath)
    sameModel = KMeansModel.load(sc, myModelPath)

# $example off$

sc.stop()

-- INSERT --
```

```
hduser@ip-172-31-29-213: /usr/local/spark/bin$
version 2.1.2

Using Python version 3.5.2 (default, Nov 23 2017 16:37:01)
SparkSession available as 'spark'.
>>> quit();
hduser@ip-172-31-29-213: /usr/local/spark/bin$ cd ..
hduser@ip-172-31-29-213: /usr/local/spark$ cd examples
hduser@ip-172-31-29-213: /usr/local/spark/examples$ ls
src
hduser@ip-172-31-29-213: /usr/local/spark/examples$ cd src
hduser@ip-172-31-29-213: /usr/local/spark/examples/src$ cd main
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main$ ls
java  python  r  resources  scala
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main$ cd python
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python$ ls
als.py      kmeans.py      ml      pagerank.py      pi.py      sql      streaming      wordcount.py
avro_inputformat.py  logistic_regression.py  mllib  parquet_inputformat.py  sort.py  status_api_demo.py  transitive_closure.py
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python$ cd mllib
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python/mllib$ ls
binary_classification_metrics_example.py      kernel_density_estimation_example.py      ranking_metrics_example.py
bisecting_k_means_example.py                  k_means_example.py                       recommendation_example.py
correlations_example.py                       kmeans.py                                regression_metrics_example.py
correlations.py                               latent_dirichlet_allocation_example.py     sampled_rdds.py
decision_tree_classification_example.py         linear_regression_with_sgd_example.py      standard_scaler_example.py
decision_tree_regression_example.py             logistic_regression.py                    stratified_sampling_example.py
elementwise_product_example.py                 logistic_regression_with_lbfgs_example.py  streaming_k_means_example.py
fpgrowth_example.py                            multi_class_metrics_example.py            streaming_linear_regression_example.py
gaussian_mixture_example.py                    multi_label_metrics_example.py            summary_statistics_example.py
gaussian_mixture_model.py                     naive_bayes_example.py                   svm_with_sgd_example.py
gradient_boosting_classification_example.py     normalizer_example.py                    tf_idf_example.py
gradient_boosting_regression_example.py         power_iteration_clustering_example.py      word2vec_example.py
hypothesis_testing_example.py                  random_forest_classification_example.py    word2vec.py
hypothesis_testing_kolmogorov_smirnov_test_example.py
isotonic_regression_example.py                  random_rdd_generation.py
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python/mllib$ vi kmeans.py
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python/mllib$ vi k_means_example.py
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python/mllib$ vi k_means_example.py
hduser@ip-172-31-29-213: /usr/local/spark/examples/src/main/python/mllib$ cd
hduser@ip-172-31-29-213: /usr/local/spark$ cd /usr/local/spark
hduser@ip-172-31-29-213: /usr/local/spark$ cd bin
hduser@ip-172-31-29-213: /usr/local/spark/bin$ spark-submit --master spark://ip-172-31-29-213:7077 /usr/local/spark/examples/src/main/python/ml
lib/k_means_example.py /user/hduser/venkat/iris.txt /user/hduser/output/
```

Komuravellyv1@udayton.edu
1015237060

```
hduser@ip-172-31-29-213: /usr/local/spark/bin
17/12/05 04:35:24 INFO storage.DiskBlockManager: Created local directory at /tmp/blockmgr-46c48128-f21f-42aa-8e04-a98a32417fa
17/12/05 04:35:24 INFO memory.MemoryStore: MemoryStore started with capacity 413.9 MB
17/12/05 04:35:24 INFO spark.SparkEnv: Registering OutputCommitCoordinator
17/12/05 04:35:24 INFO util.log: Logging initialized $2626ms
17/12/05 04:35:24 INFO server.Server: jetty-9.2.12-SMPSHOT
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@5217f3d0(/jobs,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@37ebc9d8(/jobs/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@293bb8a5(/jobs/job,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@2416a51(/jobs/job/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@6fa590ba(/stages,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@6e9319f(/stages/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@72e34f77(/stages/stage,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@7bf9b098(/stages/stage/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@389adfd(/stages/pool,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@77307458(/stages/pool/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@1fc0053e(/storage,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@290b1b2e(/storage/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@47874b25(/storage/rdd,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@33617539(/storage/rdd/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@5c177f9e(/environment,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@5db4c359(/environment/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@209775a9(/executors,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@18e7143f(/executors/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@f9b7332(/executors/threadDump,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@74cec793(/executors/threadDump/json,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@6efce9e(/static,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@4f8969b0(/,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@1bdf8190(/api,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@192cf2f7(/jobs/job/kill,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO handler.ContextHandler: Started o.s.j.s.ServletContextHandler@8a589a2(/stages/stage/kill,null,AVAILABLE,@Spark)
17/12/05 04:35:24 INFO server.ServerConnector: Started Spark@20464746(HTTP/1.1)(0.0.0.0:4040)
17/12/05 04:35:24 INFO server.Server: Started @2853ms
17/12/05 04:35:24 INFO util.Utils: Successfully started service 'SparkUI' on port 4040.
17/12/05 04:35:24 INFO ui.SparkUI: Bound SparkUI to 0.0.0.0, and started at http://172.31.29.213:4040
17/12/05 04:35:24 INFO spark.SparkContext: Added JAR file:/usr/local/spark/examples/jars/spark-examples_2.11-2.1.2.jar at spark://172.31.29.213:37632/jars/spark-examples_2.11-2.1.2.jar with timestamp 1512448524733
17/12/05 04:35:24 INFO spark.SparkContext: Added JAR file:/usr/local/spark/examples/jars/scopt_2.11-3.3.0.jar at spark://172.31.29.213:37632/jars/scopt_2.11-3.3.0.jar with timestamp 1512448524737
17/12/05 04:35:24 INFO executor.Executor: Starting executor ID driver on host localhost
17/12/05 04:35:24 INFO util.Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port 37607.
17/12/05 04:35:24 INFO netty.NettyBlockTransferService: Server created on 172.31.29.213:37607
17/12/05 04:35:24 INFO storage.BlockManager: Using org.apache.spark.storage.RandomBlockReplicationPolicy for block replication policy
17/12/05 04:35:24 INFO storage.BlockManagerMaster: Registering BlockManager BlockManagerId(driver, 172.31.29.213, 37607, None)

17/12/05 04:35:27 INFO server.ServerConnector: Stopped Spark@20464746(HTTP/1.1)(0.0.0.0:4040)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@8a589a2(/stages/stage/kill,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@192cf2f7(/jobs/job/kill,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@1bdf8190(/api,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@4f8969b0(/,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6efce9e(/static,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@74cec793(/executors/threadDump/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@f9b7332(/executors/threadDump,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@18e7143f(/executors/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@209775a9(/executors,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@5db4c359(/environment/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2c177f9e(/environment,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@33617539(/storage/rdd/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@47874b25(/storage/rdd,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@290b1b2e(/storage/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@1fc0053e(/storage,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@77307458(/stages/pool/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@389adfd(/stages/pool,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@7bf9b098(/stages/stage/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@72e34f77(/stages/stage,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6e9319f(/stages/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@6fa590ba(/stages,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@2416a51(/jobs/job/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@293bb8a5(/jobs/job,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@37ebc9d8(/jobs/json,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO handler.ContextHandler: Stopped o.s.j.s.ServletContextHandler@5217f3d0(/jobs,null,UNAVAILABLE,@Spark)
17/12/05 04:35:27 INFO ui.SparkUI: Stopped Spark web UI at http://172.31.29.213:4040
17/12/05 04:35:27 INFO spark.MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
17/12/05 04:35:27 INFO memory.MemoryStore: MemoryStore cleared
17/12/05 04:35:27 INFO storage.BlockManager: BlockManager stopped
17/12/05 04:35:27 INFO storage.BlockManagerMaster: BlockManagerMaster stopped
17/12/05 04:35:27 INFO scheduler.OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
17/12/05 04:35:27 INFO spark.SparkContext: Successfully stopped SparkContext
17/12/05 04:35:27 INFO util.ShutdownHookManager: Shutdown hook called
17/12/05 04:35:27 INFO util.ShutdownHookManager: Deleting directory /tmp/spark-a2f5106c-47fc-4e0b-9e1d-ba6464f04567
hduser@ip-172-31-29-213: /usr/local/spark/bin$
```


Komuravellyv1@udayton.edu
1015237060

```
hduser@ip-172-31-29-213: ~  
hduser@ip-172-31-29-213:~$ hadoop fs -lsr  
lsr: DEPRECATED: Please use 'ls -R' instead.  
drwxr-xr-x  - hduser supergroup          0 2017-12-04 21:43 kmeansv  
drwxr-xr-x  - hduser supergroup          0 2017-12-05 04:41 output  
drwxr-xr-x  - hduser supergroup          0 2017-12-05 04:41 output/kmeans_output  
drwxr-xr-x  - hduser supergroup          0 2017-12-05 04:41 output/kmeans_output/data  
-rw-r--r--  2 hduser supergroup          0 2017-12-05 04:41 output/kmeans_output/data/_SUCCESS  
-rw-r--r--  2 hduser supergroup    936 2017-12-05 04:41 output/kmeans_output/data/part-00000-59c878da-699e-4acf-8e00-0a44bac84e32.snappy.parquet  
-rw-r--r--  2 hduser supergroup 1439 2017-12-05 04:41 output/kmeans_output/data/part-00001-59c878da-699e-4acf-8e00-0a44bac84e32.snappy.parquet  
-rw-r--r--  2 hduser supergroup    936 2017-12-05 04:41 output/kmeans_output/data/part-00002-59c878da-699e-4acf-8e00-0a44bac84e32.snappy.parquet  
-rw-r--r--  2 hduser supergroup 1439 2017-12-05 04:41 output/kmeans_output/data/part-00003-59c878da-699e-4acf-8e00-0a44bac84e32.snappy.parquet  
-rw-r--r--  2 hduser supergroup 1455 2017-12-05 04:41 output/kmeans_output/data/part-r-00000-4b458e79-6603-4342-b53b-fcf0652959c1.snappy.parquet  
drwxr-xr-x  - hduser supergroup          0 2017-12-05 04:41 output/kmeans_output/metadata  
-rw-r--r--  2 hduser supergroup          0 2017-12-05 04:41 output/kmeans_output/metadata/_SUCCESS  
-rw-r--r--  2 hduser supergroup    80 2017-12-05 04:41 output/kmeans_output/metadata/part-00000  
drwxr-xr-x  - hduser supergroup          0 2017-12-05 04:15 venkat  
-rw-r--r--  2 hduser supergroup    576 2017-12-04 21:57 venkat/input_kmeans.txt  
-rw-r--r--  2 hduser supergroup 4700 2017-12-05 04:15 venkat/iris.txt  
hduser@ip-172-31-29-213:~$
```

```
data_clusters.txt - Notepad  
File Edit Format View Help  
cluster : [97.11, 19.56] Data Points : [[85.97, 44.19], [91.48, 38.27], [93.81, 26.96], [86.22, 2.05], [90.13, 43.54], [83.19, 36.04], [84.08, 25.5], [81.71, 38.52], [90.35, 1.14], [91.06, 44.65], [86.72, 43.08], [80.25, 29.09], [89.43, 14.7], [77.47, 34.4], [83.72, 17.56], [92.83, 1.66], [97.56, 32.24], [87.01, 32.89], [83.18, 12.81], [81.59, 6.04], [78.32, 24.59], [95.1, 33.72], [95.5, 20.22], [91.8, 19.71], [83.61, 1.22], [97.51, 13.05], [80.23, 35.87], [88.64, 38.94], [98.31, 44.26], [76.37, 21.42], [84.46, 23.79], [98.55, 22.11], [94.98, 4.62], [91.85, 25.8], [85.32, 46.84], [80.99, 18.27], [78.8, 36.14], [84.64, 45.25], [78.5, 42.68], [87.31, 47.21], [79.66, 5.9], [86.96, 20.95], [96.95, 38.42], [87.67, 20.13], [84.59, 40.23], [91.99, 17.89], [93.56, 35.6], [85.6, 14.7], [87.61, 15.57], [84.73, 13.91], [93.06, 15.66], [92.79, 27.29], [79.03, 19.26], [97.07, 49.79], [89.81, 25.63], [93.37, 29.49], [99.49, 12.93], [99.16, 55.16], [95.21, 41.82], [96.27, 15.4], [82.5, 7.73], [97.15, 38.13], [92.42, 23.43], [86.74, 31.51], [97.56, 14.01], [99.0, 50.76], [96.53, 14.62], [84.95, 33.57], [83.76, 27.12], [77.07, 12.88], [95.33, 24.24], [83.66, 17.26], [97.71, 3.33], [83.92, 13.56], [95.76, 38.52], [93.53, 17.18], [94.56, 9.86], [83.82, 5.55], [77.9, 30.2], [86.96, 10.01], [80.78, 18.5], [96.57, 43.46], [97.92, 21.52], [75.38, 15.69], [98.62, 33.84], [87.6, 29.41], [99.54, 7.01], [86.3, 43.76], [98.18, 29.79], [85.03, 16.31], [82.39, 36.47], [87.85, 23.41], [86.38, 35.35], [96.44, 16.67], [81.13, 41.48], [96.13, 33.33], [90.37, 8.75], [96.67, 44.07], [86.91, 36.04], [89.09, 31.39], [77.95, 30.37], [75.58, 28.04], [84.78, 33.14], [91.16, 28.24], [89.58, 22.46], [86.75, 19.72], [94.77, 1.82], [78.91, 33.48], [84.63, 1.58], [76.96, 33.99], [88.7, 15.21], [96.31, 42.15], [95.49, 42.48], [81.68, 26.78], [92.6, 47.27], [83.38, 29.39], [93.56, 35.64], [98.08, 50.53], [78.96, 38.2], [81.93, 22.14], [95.3, 37.68], [82.15, 2.29], [98.24, 36.23], [86.73, 19.7], [75.92, 25.16], [95.0, 42.78], [91.63, 18.47], [83.92, 41.33], [77.72, 16.62], [88.99, 42.86], [84.11, 3.97], [82.18, 9.98], [87.72, 2.42], [92.86, 50.67], [76.74, 34.59], [83.24, 43.65], [76.54, 18.21], [84.08, 15.92], [94.2, 39.41], [91.56, 48.7], [99.41, 40.99], [82.27, 40.42], [91.77, 3.73], [79.29, 21.41], [91.21, 31.63], [88.42, 37.65], [89.23, 24.76], [80.86, 13.04], [87.9, 36.93], [98.31, 7.03], [96.58, 38.23], [81.09, 36.78], [80.59, 12.1], [75.55, 16.96], [93.56, 31.71], [92.44, 45.79], [85.52, 25.54], [79.78, 34.37], [94.51, 46.81], [86.09, 30.1], [91.36, 1.24], [88.8, 14.91], [81.06, 27.53], [84.19, 9.51], [91.68, 24.92], [78.13, 18.81], [95.03, 23.64], [78.46, 38.5], [93.6, 39.06], [81.97, 14.39], [95.82, 44.6], [88.26, 9.57], [83.39, 23.65], [76.62, 29.55], [85.53, 7.59], [97.04, 24.99], [90.13, 4.11], [79.72, 18.66], [82.89, 33.13], [92.88, 49.49], [94.34, 7.55], [87.21, 24.56], [86.81, 5.34], [98.0, 35.83], [90.81, 2.09], [97.98, 19.63], [91.6, 10.05], [91.74, 28.99], [95.64, 39.76], [92.61, 28.2], [94.63, 17.63], [95.14, 51.84], [98.48, 17.3], [88.82, 19.55], [92.77, 32.33], [79.88, 35.58], [92.25, 3.97], [89.5, 1.2], [78.3, 22.47], [92.77, 41.76], [84.34, 28.37], [79.3, 37.65], [78.41, 23.55], [97.56, 44.4], [86.59, 18.11], [98.23, 4.92], [92.07, 48.96], [91.63, 42.24], [86.71, 23.4], [79.88, 6.68], [90.36, 12.55], [96.0, 12.15], [79.92, 27.56], [93.99, 32.22], [93.82, 22.93], [92.62, 41.43], [88.97, 35.33], [83.98, 21.91], [93.07, 2.25], [78.61, 27.34], [94.85, 35.46], [83.25, 6.34], [87.23, 4.96], [92.7, 29.25], [82.68, 15.64], [98.61, 6.87], [94.32, 44.31], [91.33, 15.26], [79.53, 16.44], [98.9, 17.98], [78.88, 23.5], [94.23, 26.24], [88.87, 1.08], [85.84, 2.69], [88.04, 16.68], [78.1, 14.24], [89.29, 31.43], [84.27, 37.42], [92.84, 6.15], [79.7, 24.74], [92.76, 33.29], [99.52, 2.67], [91.26, 34.79], [85.36, 10.28], [94.02, 40.9], [78.84, 41.89], [87.46, 48.28], [94.54, 39.95], [94.38, 14.54], [93.22, 37.65], [82.39, 34.34], [99.01, 44.7], [78.88, 28.77], [88.23, 15.54], [79.33, 30.49], [95.8, 19.38], [96.95, 23.1], [93.39, 20.98], [84.58, 30.77], [79.62, 6.68], [82.1, 15.06], [90.64, 41.38], [88.62, 16.65], [80.65, 10.03], [77.9, 10.6], [88.86, 14.03], [86.58, 10.24], [98.01, 26.09], [88.74, 17.52], [75.59, 17.11], [84.25, 40.04], [94.42, 5.54], [83.31, 22.33], [83.2, 44.8], [83.13, 18.08], [92.49, 20.34], [87.6, 33.5], [79.95, 38.31], [84.88, 34.34], [92.46, 43.09], [85.1, 22.17], [90.34, 24.09], [87.72, 22.01], [88.86, 39.81], [77.72, 21.8], [87.85, 25.86], [97.7, 13.11], [88.73, 48.94], [87.75, 33.27], [89.07, 11.92], [95.86, 12.58], [98.64, 7.16], [95.96, 10.29], [96.65, 4.07], [79.35, 26.88], [95.7, 22.88], [91.93, 19.36], [96.47, 27.89], [88.23, 44.9], [85.37, 39.19], [83.03, 38.1], [95.52, 44.83], [99.52, 50.86], [92.36, 16.8], [91.19, 36.04], [91.6, 42.84], [93.67, 17.69], [99.17, 55.22], [77.9, 22.69], [96.2, 24.78], [88.33, 23.17], [81.95, 14.75], [96.92, 25.13], [84.58, 12.47], [84.9, 10.64], [91.6, 44.6], [95.72, 16.55], [96.67, 4.54], [97.22, 43.92], [89.26, 27.5], [97.86, 23.33], [83.75, 5.98], [78.9, 24.02], [94.62, 18.97], [97.04, 44.86], [90.71, 45.38], [85.71, 34.9], [89.88, 48.81], [89.5, 9.85], [82.15, 16.23], [87.93, 46.49], [80.32, 41.19], [91.67, 16.95], [90.78, 26.25], [96.52, 13.47], [79.28, 24.95], [98.21, 24.3], [92.32, 23.7], [75.24, 19.84], [86.2, 11.9], [93.9, 41.75], [86.96, 31.01], [81.59, 13.78], [91.16, 30.46], [75.06, 20.23], [99.8, 19.47], [91.0, 14.77], [79.27, 21.31], [97.26, 15.45], [81.63, 17.36], [84.86, 36.86], [95.97, 44.46], [99.56, 42.83], [84.17, 2.61], [92.12, 50.6], [84.19, 45.19], [93.44, 27.7], [90.43, 5.78], [92.02, 7.83], [82.42, 2.79], [76.98, 37.7], [95.07, 39.75], [84.99, 26.62], [79.14, 11.82], [92.83, 22.68], [89.55, 25.88], [97.57, 5.93], [91.09, 2.03], [77.41, 21.21], [89.8, 48.38], [98.06, 44.61], [99.86, 36.47], [78.2, 30.03], [95.0, 36.25], [94.14, 9.59], [97.52, 29.0], [99.08, 18.66], [83.56, 30.36], [96.68, 14.21], [89.14, 10.11],
```

Komuravellyv1@udayton.edu

1015237060

C:\Users\vkumravelly\Downloads\part-r-00000 - Notepad++

File Edit Search View Encoding Language Settings Macro Run Plugins Window ?

new 1 wordcount.py wordcount.py MLkmeans.py kmeans.py part-r-00000

```
1 5.179999999999999,3.145333333333333,2.274666666666666,0.5640000000000001 [5.1,3.5,1.4,0.2,Iris-setosa] [5.4,3.9,1.3,0.4,Iris-setosa] [5.1
2 6.506666666666666,2.962666666666667,5.242666666666668,1.833333333333333 [6.6,3.0,4.4,1.4,Iris-versicolor] [7.0,3.2,4.7,1.4,Iris-versico
3
```

Normal text file length: 4,650 lines: 3 Ln: 1 Col: 1 Sel: 0 | 0 Unix (LF) UTF-8 INS

Type here to search 11:42 PM 12/4/2017