Spark:

Wget <https://raw.githubusercontent.com/ashaypatil11/hadoop/main/Spark_Hadoop.py>

Wget <https://raw.githubusercontent.com/ashaypatil11/hadoop/main/friends.csv>

hdfs dfs -mkdir spark hdfs dfs -mkdir spark/job\_output hdfs dfs -copyFromLocal Spark\_Hadoop.py spark/Spark\_Hadoop.py hdfs dfs -copyFromLocal friends.csv spark/friends.csv

Pig:

wget <https://raw.githubusercontent.com/ashaypatil11/hadoop/main/ratings.data>

Wget [https://raw.githubusercontent.com/ashaypatil11/hadoop/main/movies.data](https://raw.githubusercontent.com/ashaypatil11/hadoop/main/ratings.data)

hdfs dfs -mkdir pig hdfs dfs -copyFromLocal Spark\_Hadoop.py pig/ratings.data hdfs dfs -copyFromLocal friends.csv pig/movies.item