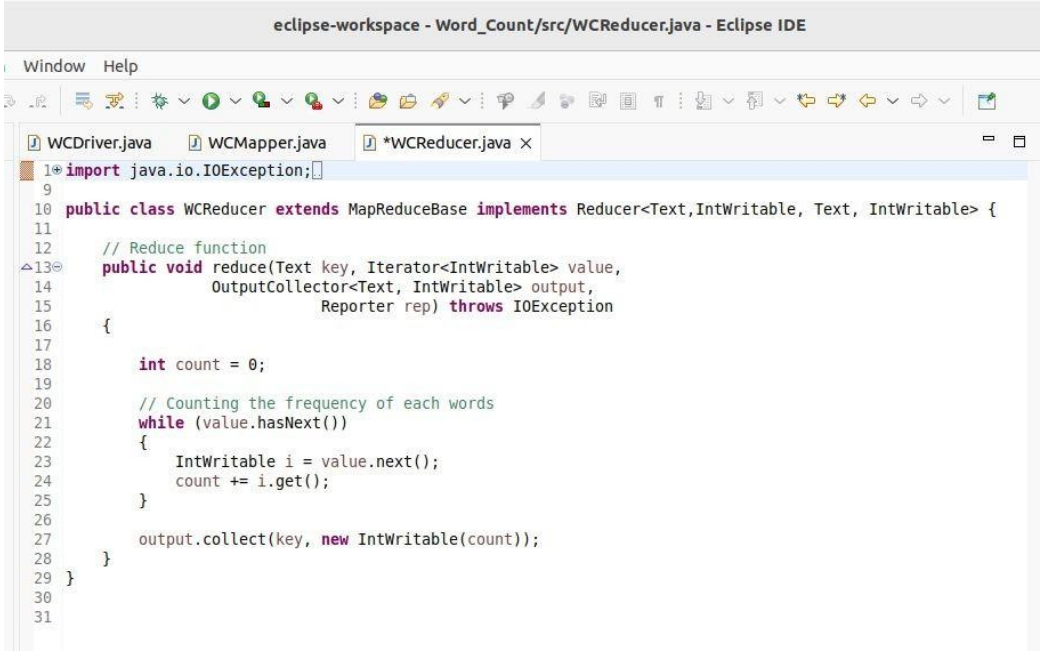
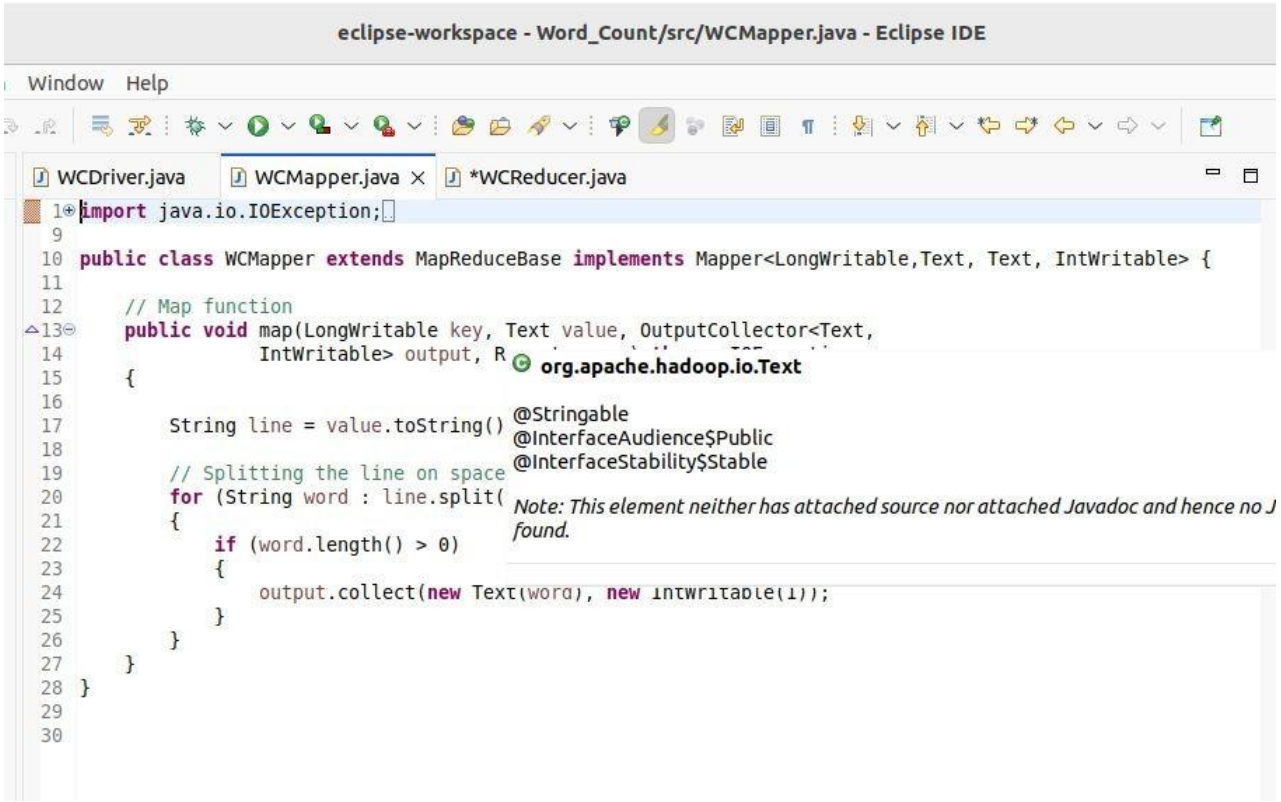
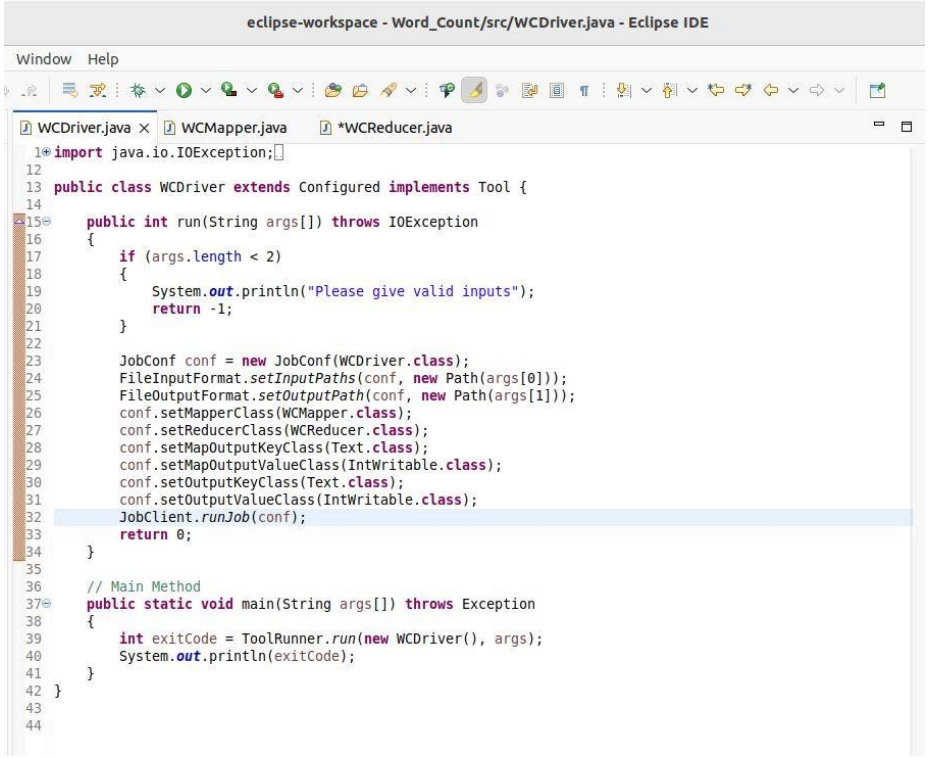
Eclipse - jar files creation (WCReducer)



Eclipse - jar files creation (WCMapper)



Eclipse - jar files creation (WCDriver)



```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [bmscecse-HP-Elite-Tower-800-G9-Desktop-PC]
Starting resourcemanager
Starting nodemanagers
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ su hduser
su: user hduser does not exist or the user entry does not contain all the required fields
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ jps
7890 ResourceManager
8054 NodeManager
8488 Jps
7259 DataNode
5468 org.eclipse.equinox.launcher_1.6.1000.v20250227-1734.jar
7117 NameNode
7517 SecondaryNameNode
```

hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC: ~/Desktop

hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC: ~/Desktop

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
localhost: namenode is running as process 7117.  Stop it first and ensure /tmp/hadoop-hadoop-namenode.pid file is empty before retry.
Starting datanodes
localhost: datanode is running as process 7259.  Stop it first and ensure /tmp/hadoop-hadoop-datanode.pid file is empty before retry.
Starting secondary namenodes [bmscecse-HP-Elite-Tower-800-G9-Desktop-PC]
bmscecse-HP-Elite-Tower-800-G9-Desktop-PC: secondarynamenode is running as process 7517.  Stop it first and ensure /tmp/hadoop-hadoop-secondarynamenode.pid file is empty before retry.
Starting resourcemanager
resourcemanager is running as process 7890.  Stop it first and ensure /tmp/hadoop-hadoop-resourcemanager.pid file is empty before retry.
Starting nodemanagers
localhost: nodemanager is running as process 8054.  Stop it first and ensure /tmp/hadoop-hadoop-nodemanager.pid file is empty before retry.
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ jps
7890 ResourceManager
12325 Jps
8054 NodeManager
7259 DataNode
7117 NameNode
7517 SecondaryNameNode
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop fs -mkdir /shuraim
```

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop fs -copyFromLocal /home/hadoop/Desktop/file1.txt /shuraim/test.txt
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop jar /home/hadoop/Desktop/WordCount.jar WCDriver /shuraim/test.txt /shuraim/output
2025-04-29 15:31:03,072 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-04-29 15:31:03,114 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-04-29 15:31:03,114 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2025-04-29 15:31:03,121 WARN impl.MetricsSystemImpl: JobTracker metrics system already initialized!
2025-04-29 15:31:03,186 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
2025-04-29 15:31:03,248 INFO mapred.FileInputFormat: Total input files to process : 1
2025-04-29 15:31:03,273 INFO mapreduce.JobSubmitter: number of splits:1
2025-04-29 15:31:03,342 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local457507589_0001
2025-04-29 15:31:03,342 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-04-29 15:31:03,407 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2025-04-29 15:31:03,408 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-04-29 15:31:03,408 INFO mapreduce.Job: Running job: job_local457507589_0001
2025-04-29 15:31:03,409 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
2025-04-29 15:31:03,411 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-04-29 15:31:03,412 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-04-29 15:31:03,458 INFO mapred.LocalJobRunner: Waiting for map tasks
2025-04-29 15:31:03,459 INFO mapred.LocalJobRunner: Starting task: attempt_local457507589_0001_m_000000_0
2025-04-29 15:31:03,471 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-04-29 15:31:03,471 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
2025-04-29 15:31:03,478 INFO mapred.Task:  Using ResourceCalculatorProcessTree : [ ]
2025-04-29 15:31:03,481 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/shuraim/test.txt:0+89
2025-04-29 15:31:03,489 INFO mapred.MapTask: numReduceTasks: 1
2025-04-29 15:31:03,521 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
2025-04-29 15:31:03,521 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
2025-04-29 15:31:03,521 INFO mapred.MapTask: soft limit at 83886080
2025-04-29 15:31:03,521 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
2025-04-29 15:31:03,521 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
2025-04-29 15:31:03,523 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
2025-04-29 15:31:03,561 INFO mapred.LocalJobRunner:
2025-04-29 15:31:03,561 INFO mapred.MapTask: Starting flush of map output
2025-04-29 15:31:03,561 INFO mapred.MapTask: Spilling map output
2025-04-29 15:31:03,561 INFO mapred.MapTask: bufstart = 0; bufend = 169; bufvoid = 104857600
2025-04-29 15:31:03,561 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26214320(104857280); length = 77/6553600
2025-04-29 15:31:03,564 INFO mapred.MapTask: Finished spill 0
2025-04-29 15:31:03,569 INFO mapred.Task: Task:attempt_local457507589_0001_m_000000_0 is done. And is in the process of committing
2025-04-29 15:31:03,571 INFO mapred.LocalJobRunner: hdfs://localhost:9000/shuraim/test.txt:0+89
2025-04-29 15:31:03,571 INFO mapred.Task: Task 'attempt_local457507589_0001_m_000000_0' done.
2025-04-29 15:31:03,574 INFO mapred.Task: Final Counters for attempt_local457507589_0001_m_000000_0: Counters: 23
        File System Counters
                FILE: Number of bytes read=4049
                FILE: Number of bytes written=642420
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=89
```

```
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop fs -ls  /shuraim/output
Found 2 items
-rw-r--r--   1 hadoop supergroup          0 2025-04-29 15:31 /shuraim/output/_SUCCESS
-rw-r--r--   1 hadoop supergroup         69 2025-04-29 15:31 /shuraim/output/part-00000
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ hadoop fs -cat  /shuraim/output/part-00000
are     1
brother 1
family  1
hi      1
how     5
is      4
job     1
sister  1
you     1
your    4
hadoop@bmscecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$
```