WC_Mapper.java
package com.javatpoint;

```java
        import java.io.IOException;
        import java.util.StringTokenizer;
        import org.apache.hadoop.io.IntWritable;
        import org.apache.hadoop.io.LongWritable;
        import org.apache.hadoop.io.Text;
        import org.apache.hadoop.mapred.MapReduceBase;
        import org.apache.hadoop.mapred.Mapper;
        import org.apache.hadoop.mapred.OutputCollector;
        import org.apache.hadoop.mapred.Reporter;




        public class WC_Mapper extends MapReduceBase implements
Mapper<LongWritable,Text,Text,IntWritable>{

        private final static IntWritable one = new IntWritable(1);
private Text word = new Text();

        public void map(LongWritable key, Text value,OutputCollector<Text,IntWritable>
output, Reporter reporter) throws IOException{



        String line = value.toString();

        StringTokenizer  tokenizer = new StringTokenizer(line);


        while (tokenizer.hasMoreTokens()){
                word.set(tokenizer.nextToken());
                output.collect(word, one);

}
        }

        }
```
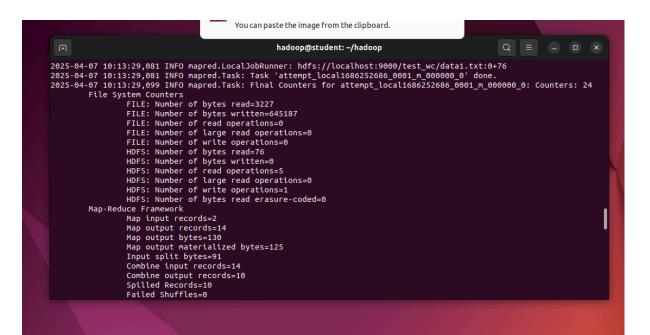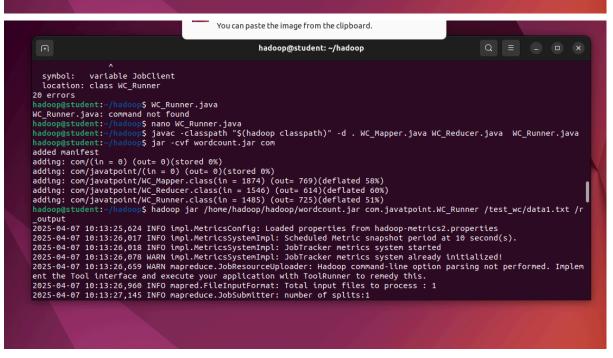
WC_Reducer.java

```java
package com.javatpoint;
    import java.io.IOException;
    import java.util.Iterator;
    import org.apache.hadoop.io.IntWritable;
    import org.apache.hadoop.io.Text;
    import org.apache.hadoop.mapred.MapReduceBase;
    import org.apache.hadoop.mapred.OutputCollector;
    import org.apache.hadoop.mapred.Reducer;
    import org.apache.hadoop.mapred.Reporter;

    public class WC_Reducer  extends MapReduceBase implements
Reducer<Text,IntWritable,Text,IntWritable> {

    public void reduce(Text key, Iterator<IntWritable>
values,OutputCollector<Text,IntWritable> output, Reporter reporter) throws IOException {

    int sum=0;
    while (values.hasNext()) {
    sum+=values.next().get();
    }


    output.collect(key,new IntWritable(sum));

    }
    }
```

WC_Runner.java

```java
    package com.javatpoint;

    import java.io.IOException;
    import org.apache.hadoop.fs.Path;
    import org.apache.hadoop.io.IntWritable;
    import org.apache.hadoop.io.Text;
    import org.apache.hadoop.mapred.FileInputFormat;
    import org.apache.hadoop.mapred.FileOutputFormat;
    import org.apache.hadoop.mapred.JobClient;
    import org.apache.hadoop.mapred.JobConf;
    import org.apache.hadoop.mapred.TextInputFormat;
    import org.apache.hadoop.mapred.TextOutputFormat;

    public class WC_Runner {

    public static void main(String[] args) throws IOException{
    JobConf conf = new JobConf(WC_Runner.class);
```

```java
        conf.setJobName("WordCount");

                conf.setOutputKeyClass(Text.class);
                conf.setOutputValueClass(IntWritable.class);
        conf.setMapperClass(WC_Mapper.class);
                conf.setCombinerClass(WC_Reducer.class);
        conf.setReducerClass(WC_Reducer.class);
                conf.setInputFormat(TextInputFormat.class);
        conf.setOutputFormat(TextOutputFormat.class);

            FileInputFormat.setInputPaths(conf,new Path(args[0]));
                FileOutputFormat.setOutputPath(conf,new Path(args[1]));

                JobClient.runJob(conf);
        }

}
```

```
hadoop@student: ~/hadoop

2025-04-07 10:13:29,081 INFO mapred.LocalJobRunner: hdfs://localhost:9000/test_wc/data1.txt:0+76
2025-04-07 10:13:29,081 INFO mapred.Task: Task 'attempt_local1686252686_0001_m_000000_0' done.
2025-04-07 10:13:29,099 INFO mapred.Task: Final Counters for attempt_local1686252686_0001_m_000000_0: Counters: 24
        File System Counters
                FILE: Number of bytes read=3227
                FILE: Number of bytes written=645187
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=76
                HDFS: Number of bytes written=0
                HDFS: Number of read operations=5
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=1
                HDFS: Number of bytes read erasure-coded=0
        Map-Reduce Framework
                Map input records=2
                Map output records=14
                Map output bytes=130
                Map output materialized bytes=125
                Input split bytes=91
                Combine input records=14
                Combine output records=10
                Spilled Records=10
                Failed Shuffles=0
```

```
hadoop@student: ~/hadoop

              ^
  symbol:    variable JobClient
  location: class WC_Runner
20 errors
hadoop@student:~/hadoop$ WC_Runner.java
WC_Runner.java: command not found
hadoop@student:~/hadoop$ nano WC_Runner.java
hadoop@student:~/hadoop$ javac -classpath "$(hadoop classpath)" -d . WC_Mapper.java WC_Reducer.java  WC_Runner.java
hadoop@student:~/hadoop$ jar -cvf wordcount.jar com
added manifest
adding: com/(in = 0) (out= 0)(stored 0%)
adding: com/javatpoint/(in = 0) (out= 0)(stored 0%)
adding: com/javatpoint/WC_Mapper.class(in = 1874) (out= 769)(deflated 58%)
adding: com/javatpoint/WC_Reducer.class(in = 1546) (out= 614)(deflated 60%)
adding: com/javatpoint/WC_Runner.class(in = 1485) (out= 725)(deflated 51%)
hadoop@student:~/hadoop$ hadoop jar /home/hadoop/hadoop/wordcount.jar com.javatpoint.WC_Runner /test_wc/data1.txt /r
_output
2025-04-07 10:13:25,624 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-04-07 10:13:26,017 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-04-07 10:13:26,018 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2025-04-07 10:13:26,078 WARN impl.MetricsSystemImpl: JobTracker metrics system already initialized!
2025-04-07 10:13:26,659 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implem
ent the Tool interface and execute your application with ToolRunner to remedy this.
2025-04-07 10:13:26,960 INFO mapred.FileInputFormat: Total input files to process : 1
2025-04-07 10:13:27,145 INFO mapreduce.JobSubmitter: number of splits:1
```
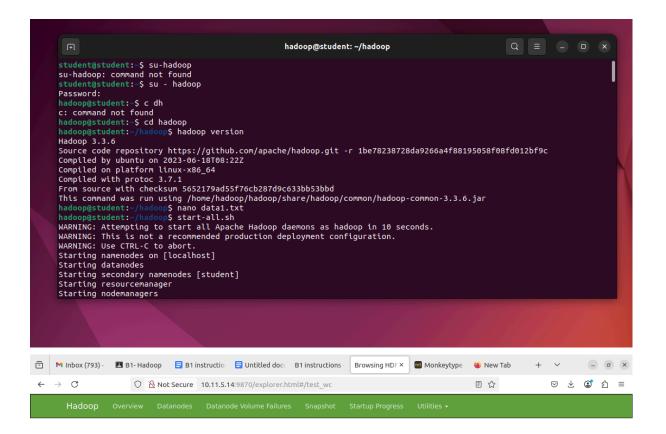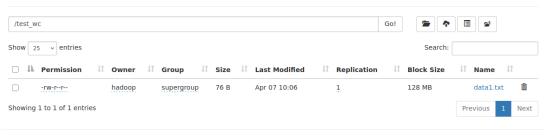
```
                    ^
 symbol:    variable JobClient
 location: class WC_Runner
20 errors
error: compilation failed
hadoop@student:~/hadoop$ java WC_Runner
Error: Could not find or load main class WC_Runner
Caused by: java.lang.ClassNotFoundException: WC_Runner
hadoop@student:~/hadoop$ WC_Runner.java
WC_Runner.java: command not found
hadoop@student:~/hadoop$ nano WC_Runner.java
hadoop@student:~/hadoop$ javac -classpath "$(hadoop classpath)" -d . WC_Mapper.java WC_Reducer.java
hadoop@student:~/hadoop$ WC_Runner.java
WC_Runner.java: command not found
hadoop@student:~/hadoop$ javac WC_
error: Class names, 'WC_', are only accepted if annotation processing is explicitly requested
1 error
hadoop@student:~/hadoop$ javac WC_R
WC_Reducer.java  WC_Runner.java
hadoop@student:~/hadoop$ javac WC_Runner.java
WC_Runner.java:5: error: package org.apache.hadoop.fs does not exist
        import org.apache.hadoop.fs.Path;
                                 ^
WC_Runner.java:6: error: package org.apache.hadoop.io does not exist
        import org.apache.hadoop.io.IntWritable;
```

```
 command 'ifconfig' from deb net-tools (1.60+git20181103.0eebece-1ubuntu5)
Try: apt install <deb name>
hadoop@student:~/hadoop$ ifconfig
enp2s0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 10.11.5.14  netmask 255.255.0.0  broadcast 10.11.255.255
        inet6 fe80::8b65:e0b0:1988:2fad  prefixlen 64  scopeid 0x20<link>
        ether a4:1f:72:8a:94:10  txqueuelen 1000  (Ethernet)
        RX packets 83554  bytes 67239715 (67.2 MB)
        RX errors 0  dropped 307  overruns 0  frame 0
        TX packets 34330  bytes 7849050 (7.8 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 1000  (Local Loopback)
        RX packets 2602  bytes 302824 (302.8 KB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 2602  bytes 302824 (302.8 KB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

hadoop@student:~/hadoop$ ^C
hadoop@student:~/hadoop$ nano WC_Mapper.java
hadoop@student:~/hadoop$ nano WC_Reducer.java
hadoop@student:~/hadoop$ Nano WC_Runner.java
```

```
student@student:~$ su-hadoop
su-hadoop: command not found
student@student:~$ su - hadoop
Password:
hadoop@student:~$ c dh
c: command not found
hadoop@student:~$ cd hadoop
hadoop@student:~/hadoop$ hadoop version
Hadoop 3.3.6
Source code repository https://github.com/apache/hadoop.git -r 1be78238728da9266a4f88195058f08fd012bf9c
Compiled by ubuntu on 2023-06-18T08:22Z
Compiled on platform linux-x86_64
Compiled with protoc 3.7.1
From source with checksum 5652179ad55f76cb287d9c633bb53bbd
This command was run using /home/hadoop/hadoop/share/hadoop/common/hadoop-common-3.3.6.jar
hadoop@student:~/hadoop$ nano data1.txt
hadoop@student:~/hadoop$ 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 [student]
Starting resourcemanager
Starting nodemanagers
```

Hadoop   Overview   Datanodes   Datanode Volume Failures   Snapshot   Startup Progress   Utilities

Browse Directory

/test_wc   Go!

Show 25 entries   Search:

	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
☐	-rw-r--r--	hadoop	supergroup	76 B	Apr 07 10:06	1	128 MB	data1.txt	🗑

Showing 1 to 1 of 1 entries   Previous   1   Next

Hadoop, 2023.