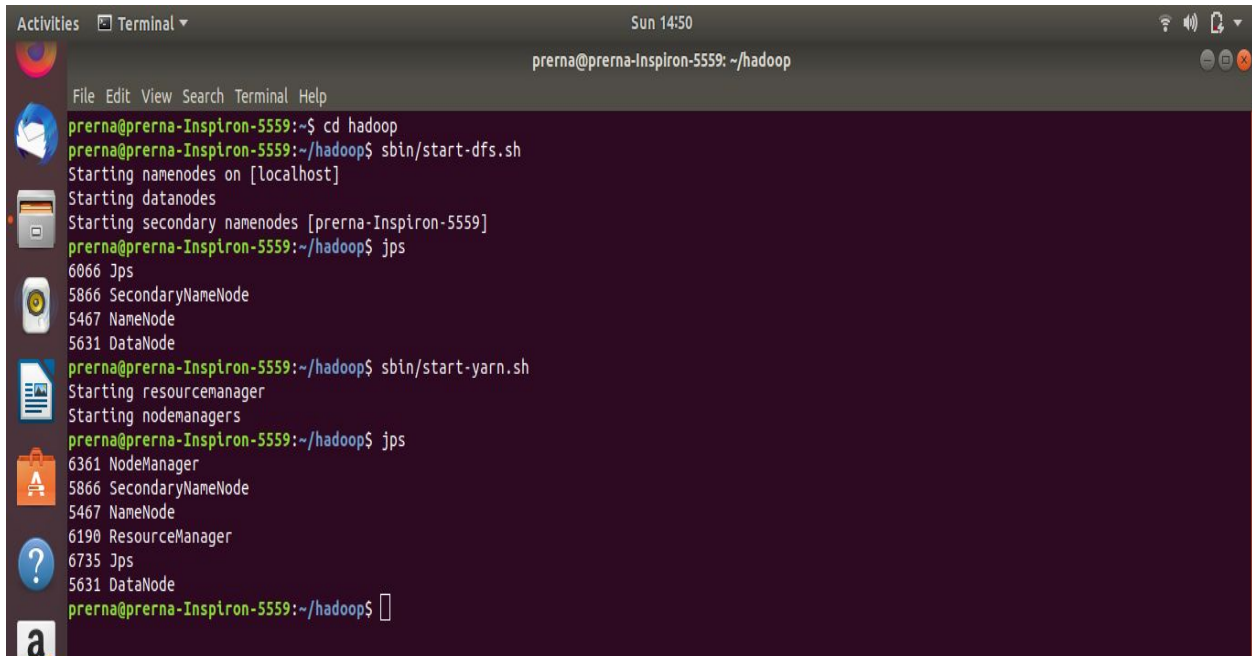


Procedure of Number Sum

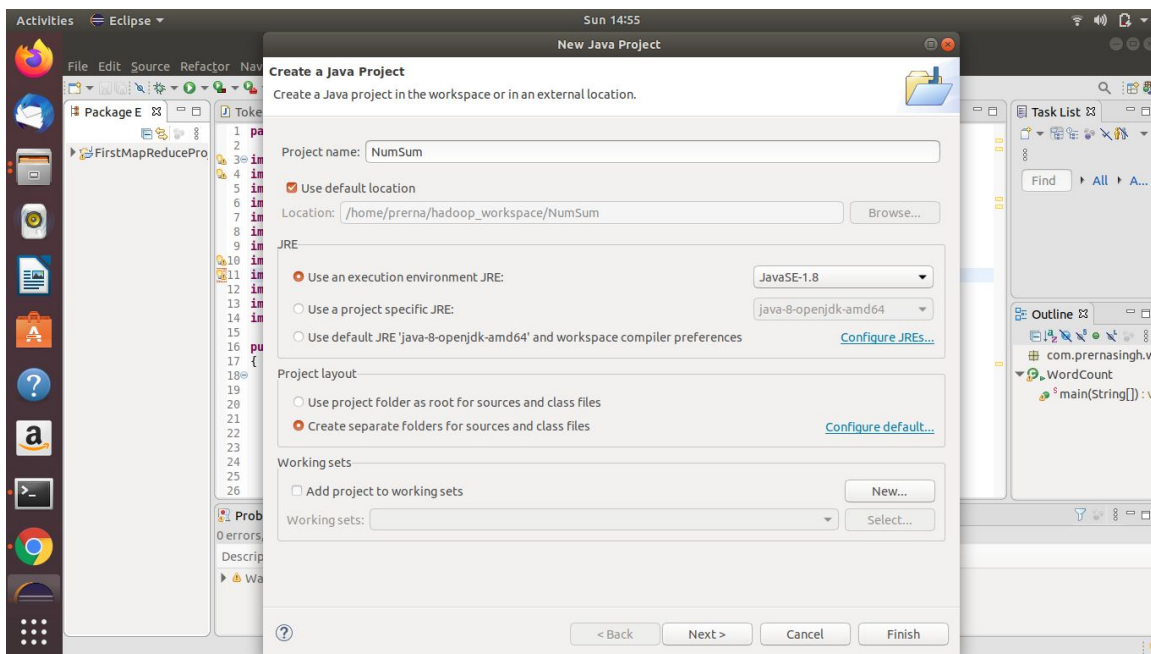
Step 1 : Start hadoop



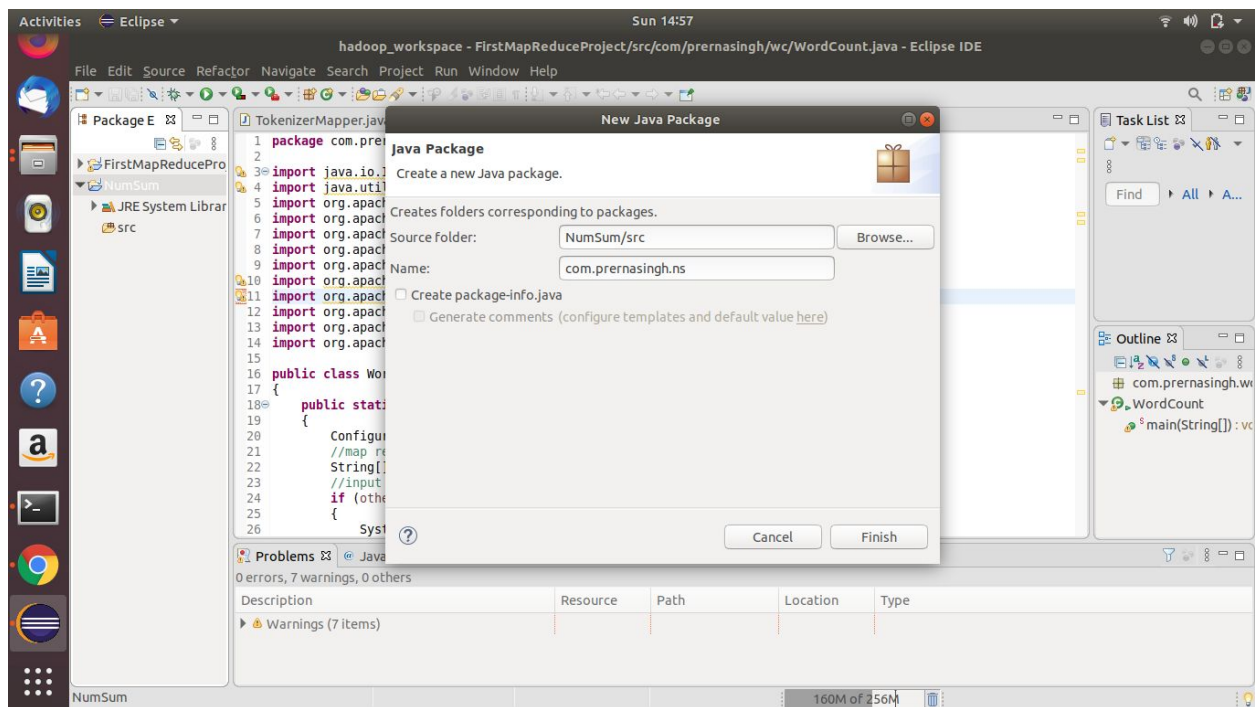
```
prerna@prerna-Inspiron-5559:~$ cd hadoop
prerna@prerna-Inspiron-5559:~/hadoop$ sbin/start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [prerna-Inspiron-5559]
prerna@prerna-Inspiron-5559:~/hadoop$ jps
6066 Jps
5866 SecondaryNameNode
5467 NameNode
5631 DataNode
prerna@prerna-Inspiron-5559:~/hadoop$ sbin/start-yarn.sh
Starting resourcemanager
Starting nodemanagers
prerna@prerna-Inspiron-5559:~/hadoop$ jps
6361 NodeManager
5866 SecondaryNameNode
5467 NameNode
6190 ResourceManager
6735 Jps
5631 DataNode
prerna@prerna-Inspiron-5559:~/hadoop$
```

Step 2: For Java Program

Step 2.1: Create project



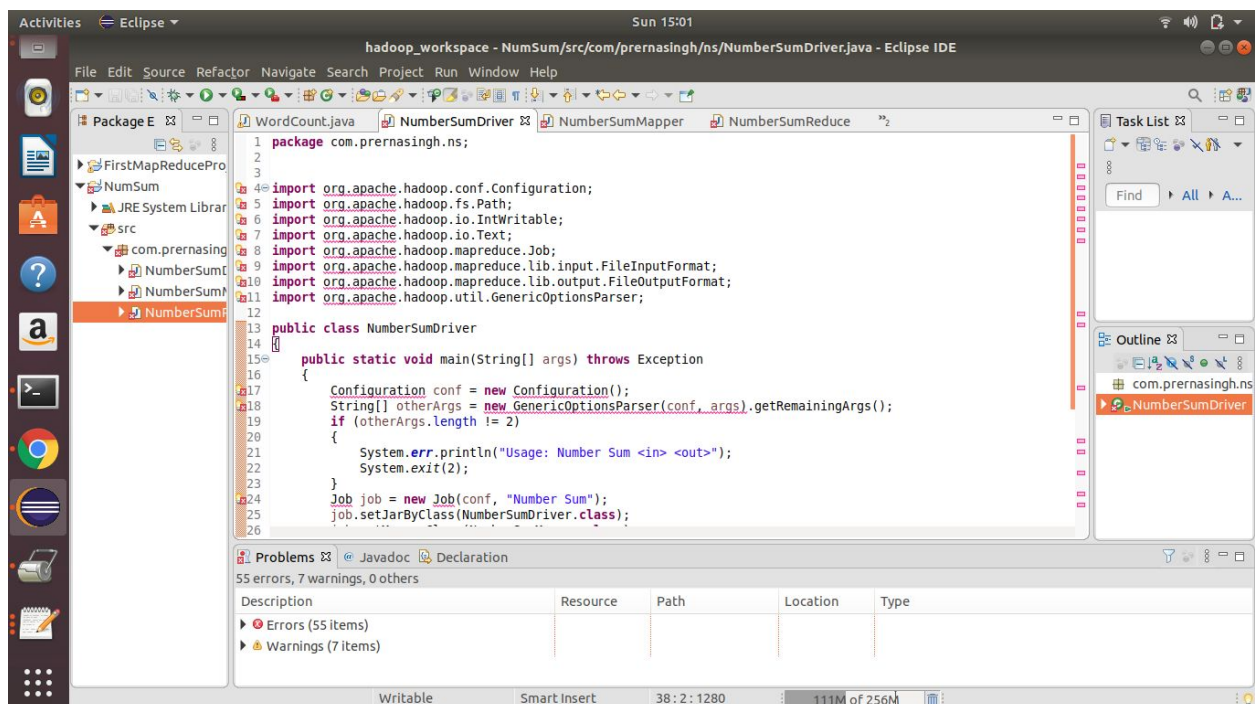
Step 2.2: Create package



Step 2.3: Write Code

<https://github.com/Prerna11071997/NumberSum>

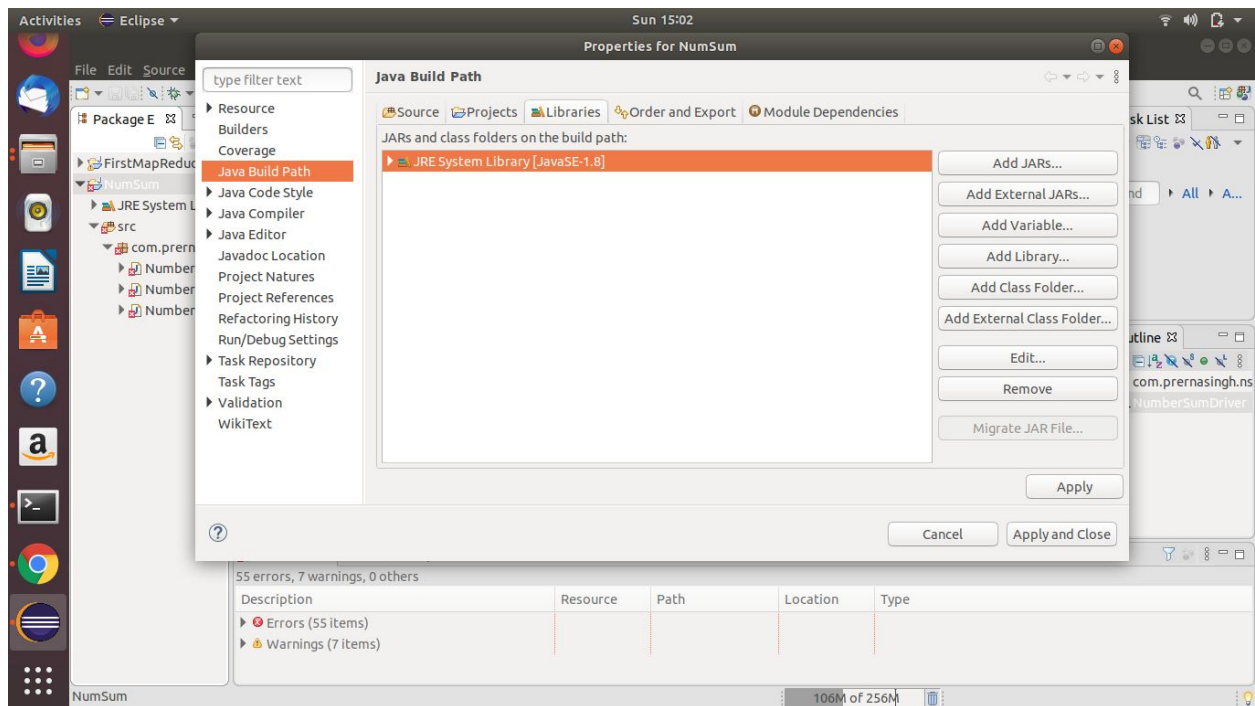
Code in src



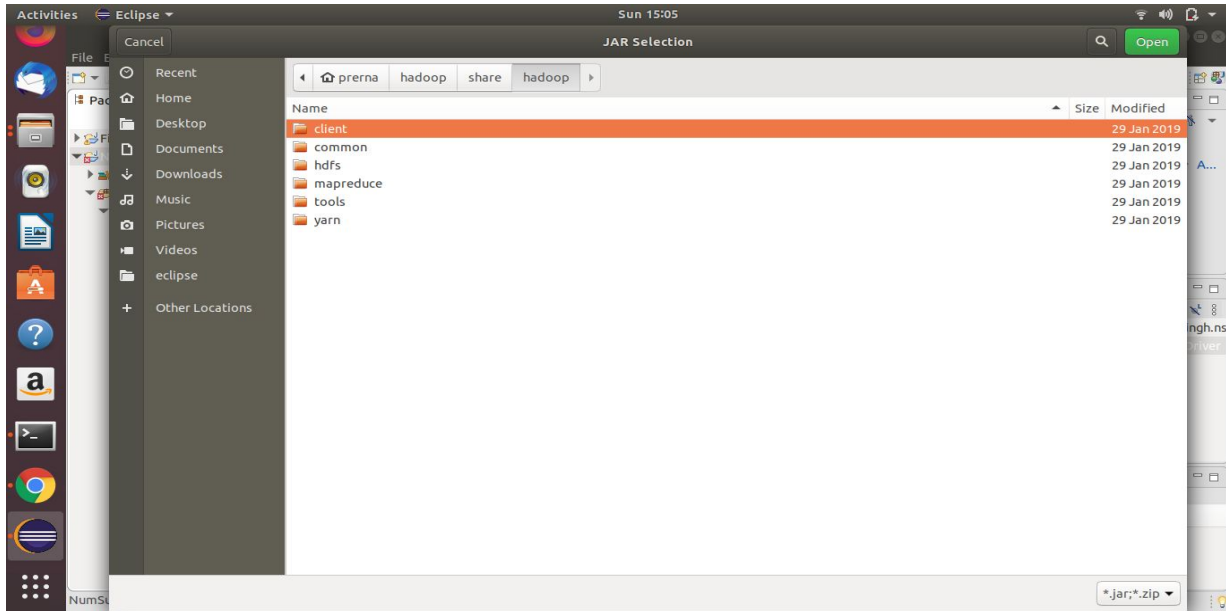
Step 2.4: Build Path

Right click on Project Name and go to Build Path and then config BuildPath

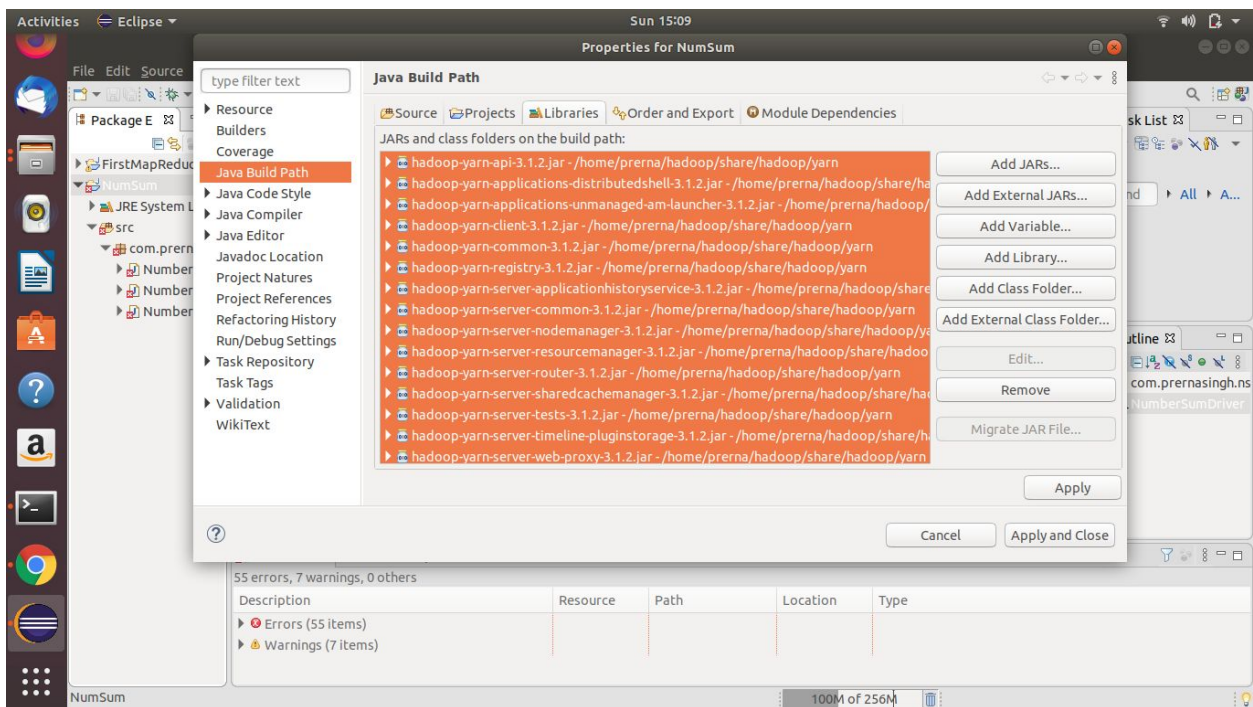
Step 2.5: Add External Jars



Step 2.6: go to hadoop >> share>> hadoop

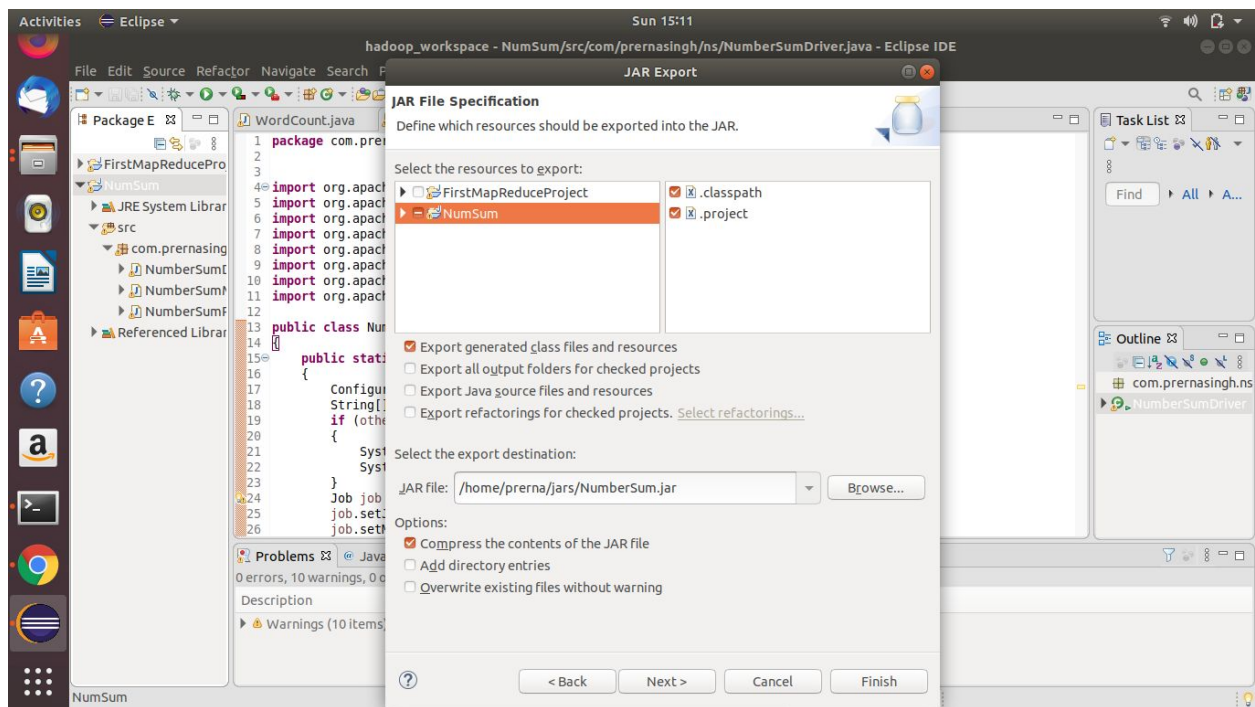


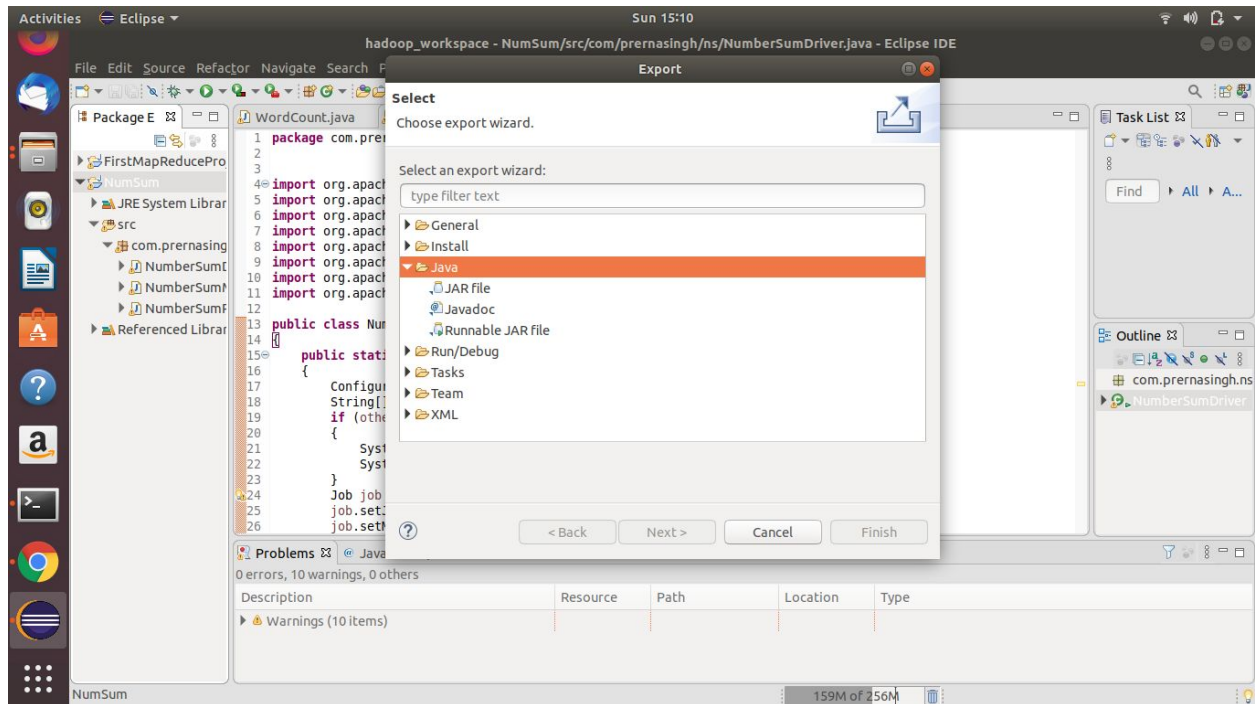
Step 2.7: add client jars ;
common jars;
common>>lib jars ;
yarn jars ;
mapreduce jars;
hdfs jars



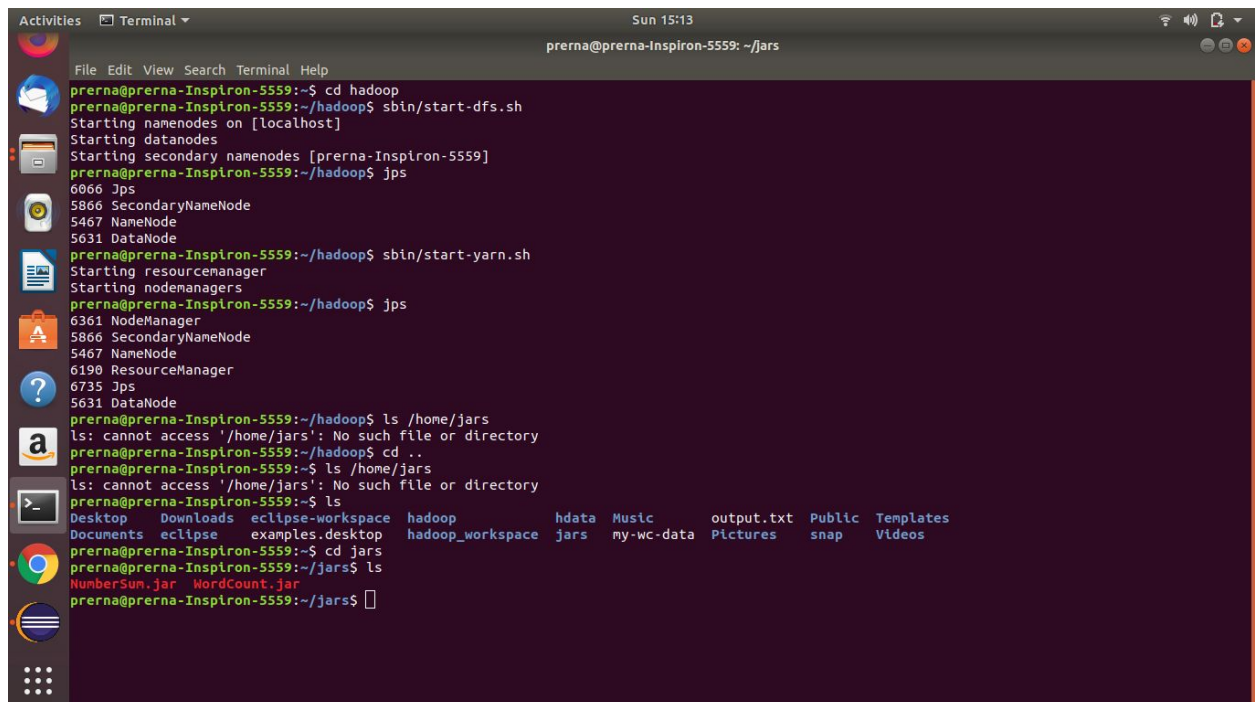
Click on apply and close

Step 2.8: Right click on ProjectName >> Export and select jar file
Give a path name or destination name
Click Finish

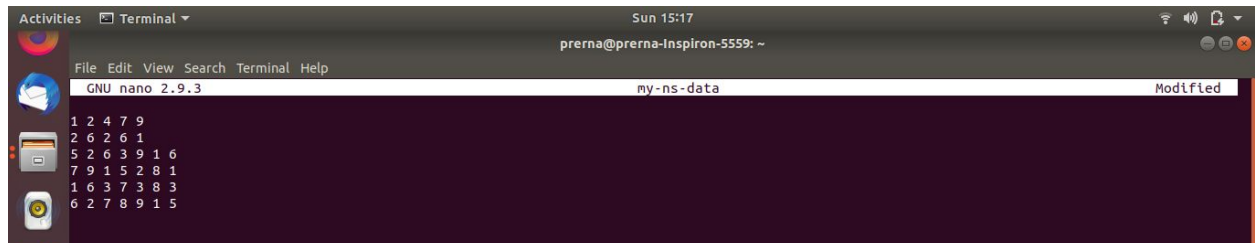




Step 3: Run on terminal

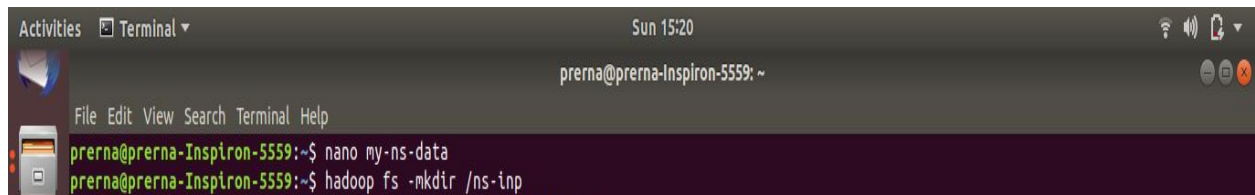


Step 4: Create input file



```
prerna@prerna-Inspiron-5559: ~  
GNU nano 2.9.3 my-ns-data Modified  
1 2 4 7 9  
2 6 2 6 1  
5 2 6 3 9 1 6  
7 9 1 5 2 8 1  
1 6 3 7 3 8 3  
6 2 7 8 9 1 5
```

Step 5: Create directory in hadoop



```
prerna@prerna-Inspiron-5559: ~  
prerna@prerna-Inspiron-5559:~$ nano my-ns-data  
prerna@prerna-Inspiron-5559:~$ hadoop fs -mkdir /ns-inp
```

Step 6: Put data in directory



```
prerna@prerna-Inspiron-5559: ~  
prerna@prerna-Inspiron-5559:~$ nano my-ns-data  
prerna@prerna-Inspiron-5559:~$ hadoop fs -mkdir /ns-inp  
prerna@prerna-Inspiron-5559:~$ hadoop fs -put my-ns-data /ns-inp
```

Step 7: Start the program



```
prerna@prerna-Inspiron-5559: ~  
prerna@prerna-Inspiron-5559:~$ nano my-ns-data  
prerna@prerna-Inspiron-5559:~$ hadoop fs -mkdir /ns-inp  
prerna@prerna-Inspiron-5559:~$ hadoop fs -put my-ns-data /ns-inp  
prerna@prerna-Inspiron-5559:~$  
prerna@prerna-Inspiron-5559:~$ hadoop jar jars/NumberSum.jar com.pernasingh.ns.NumberSumDriver /ns-inp /ns-out
```

```
Activities Terminal Sun 15:26
prerna@prerna-Inspiron-5559: ~
prerna@prerna-Inspiron-5559:~$ hadoop fs -put my-ns-data /ns-inp
prerna@prerna-Inspiron-5559:~$ hadoop jar jars/NumberSum.jar com.prernasingh.ns.NumberSumDriver /ns-inp /ns-out
2020-02-23 15:22:56,302 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2020-02-23 15:22:56,793 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/prerna/.staging/job_1582449516316_0001
2020-02-23 15:22:57,082 INFO input.FileInputFormat: Total input files to process : 1
2020-02-23 15:22:57,301 INFO mapreduce.JobSubmitter: number of splits:1
2020-02-23 15:22:57,488 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1582449516316_0001
2020-02-23 15:22:57,498 INFO mapreduce.JobSubmitter: Executing with tokens: []
2020-02-23 15:22:57,691 INFO conf.Configuration: resource-types.xml not found
2020-02-23 15:22:57,691 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2020-02-23 15:22:57,967 INFO impl.YarnClientImpl: Submitted application application_1582449516316_0001
2020-02-23 15:22:58,105 INFO mapreduce.Job: The url to track the job: http://prerna-Inspiron-5559:8088/proxy/application_1582449516316_0001/
2020-02-23 15:22:58,105 INFO mapreduce.Job: Running job: job_1582449516316_0001
2020-02-23 15:23:06,224 INFO mapreduce.Job: Job job_1582449516316_0001 running in uber mode : false
2020-02-23 15:23:06,227 INFO mapreduce.Job: map 0% reduce 0%
2020-02-23 15:23:11,456 INFO mapreduce.Job: map 100% reduce 0%
2020-02-23 15:23:17,518 INFO mapreduce.Job: map 100% reduce 100%
2020-02-23 15:23:18,545 INFO mapreduce.Job: Job job_1582449516316_0001 completed successfully
2020-02-23 15:23:18,684 INFO mapreduce.Job: Counters: 53
File System Counters
  FILE: Number of bytes read=66
  FILE: Number of bytes written=432839
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=181
  HDFS: Number of bytes written=4
  HDFS: Number of read operations=8
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counters
  Launched map tasks=1
  Launched reduce tasks=1
  Data-local map tasks=1
  Total time spent by all maps in occupied slots (ms)=2770
```

Step 8: See the Output (First file should be success)

```
Activities Terminal Sun 15:23
prerna@prerna-Inspiron-5559: ~
prerna@prerna-Inspiron-5559:~$ hadoop fs -ls /ns-out
Map output bytes=48
Map output materialized bytes=66
Input split bytes=104
Combine input records=0
Combine output records=0
Reduce input groups=1
Reduce shuffle bytes=66
Reduce input records=6
Reduce output records=1
Spilled Records=12
Shuffled Maps =1
Failed Shuffles=0
Merged Map outputs=1
GC time elapsed (ms)=115
CPU time spent (ms)=1340
Physical memory (bytes) snapshot=511709184
Virtual memory (bytes) snapshot=5320900608
Total committed heap usage (bytes)=459800576
Peak Map Physical memory (bytes)=282804224
Peak Map Virtual memory (bytes)=2656976896
Peak Reduce Physical memory (bytes)=228904960
Peak Reduce Virtual memory (bytes)=2663923712
Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=77
File Output Format Counters
  Bytes Written=4
prerna@prerna-Inspiron-5559:~$ hadoop fs -ls /ns-out
Found 2 items
-rw-r--r-- 1 prerna supergroup 0 2020-02-23 15:23 /ns-out/_SUCCESS
-rw-r--r-- 1 prerna supergroup 4 2020-02-23 15:23 /ns-out/part-r-00000
prerna@prerna-Inspiron-5559:~$
```

Step 9: See the output


```
File Output Format Counters
  Bytes Written=4
prerna@prerna-Inspiron-5559:~$ hadoop fs -ls /ns-out
Found 2 items
-rw-r--r--  1 prerna supergroup          0 2020-02-23 15:23 /ns-out/ SUCCESS
-rw-r--r--  1 prerna supergroup          4 2020-02-23 15:23 /ns-out/part-r-00000
prerna@prerna-Inspiron-5559:~$ hadoop fs -cat /ns-out/part-r-00000
174
prerna@prerna-Inspiron-5559:~$
```

Step 10: Stop hadoop

```
prerna@prerna-Inspiron-5559:~$ cd hadoop
prerna@prerna-Inspiron-5559:~/hadoop$ sbin/stop-yarn.sh
Stopping nodemanagers
localhost: WARNING: nodemanager did not stop gracefully after 5 seconds: Trying to kill with kill -9
Stopping resourcemanager
prerna@prerna-Inspiron-5559:~/hadoop$ sbin/stop-dfs.sh
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [prerna-Inspiron-5559]
prerna@prerna-Inspiron-5559:~/hadoop$ jps
9953 Jps
7157 org.eclipse.equinox.launcher_1.5.600.v20191014-2022.jar
prerna@prerna-Inspiron-5559:~/hadoop$
```