

Scala Array take array and print its maximum and minimum:

```

bmscsece@bmscsece-HP-Elite-Tower-800-G9-Desktop-PC: ~
25/05/27 14:54:30 WARN Utils: Your hostname, bmscsece-HP-Elite-Tower-800-G9-Desktop-PC resolves to a loopback address: 127.0.1.1; using 10.124.2.95 instead (on interface eno1)
25/05/27 14:54:30 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.spark.unsafe.Platform (file:/opt/spark/jars/spark-unsafe_2.12-3.0.3.jar) to constructor java.nio.DirectByteBuffer(long,int)
WARNING: Please consider reporting this to the maintainers of org.apache.spark.unsafe.Platform
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
25/05/27 14:54:31 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
Setting default log level to "WARN".
To adjust logging level use sc.setLogLevel(newLevel). For SparkR, use setLogLevel(newLevel).
Spark context Web UI available at http://10.124.2.95:4040
Spark context available as 'sc' (master = local[*], app id = local-1748337874675).
Spark session available as 'spark'.
Welcome to

  ____      __
 / ___ |__ /  __/
/ /___/ __/ /  /
/____/___/_/  /_

version 3.0.3

Using Scala version 2.12.10 (OpenJDK 64-Bit Server VM, Java 11.0.22)
Type in expressions to have them evaluated.
Type :help for more information.

scala> val data = Array(10, 5, 20, 8, 15, 2, 25)
data: Array[Int] = Array(10, 5, 20, 8, 15, 2, 25)

scala> val dataRDD = sc.parallelize(data)
dataRDD: org.apache.spark.rdd.RDD[Int] = ParallelCollectionRDD[0] at parallelize at <console>:126

scala> val maximum = dataRDD.max()
maximum: Int = 25

scala> val minimum = dataRDD.min()
minimum: Int = 2

scala> println(s"Maximum element: ${maximum}")
Maximum element: 25

scala> println(s"Minimum element: ${minimum}")
Minimum element: 2

scala> █

```

```
Welcome to  

 _   _          _  
| |_|_|_      | |_||_|_||_|_||_|_  
/_/_/_/_    /_/__/_/_/_/_\_____\n        version 3.0.3
```

Using Scala version 2.12.10 (OpenJDK 64-Bit Server VM, Java 11.0.22)
Type in expressions to have them evaluated.
Type :help for more information.

```
scala> val data = Array(10, 5, 20, 8, 15, 2, 25)
data: Array[Int] = Array(10, 5, 20, 8, 15, 2, 25)
```

```
scala> val dataRDD = sc.parallelize(data)
dataRDD: org.apache.spark.rdd.RDD[Int] = ParallelCollectionRDD[0] at parallelize at <console>:26
```

```
scala> val maximum = dataRDD.max()
maximum: Int = 25
```

```
scala> val minimum = dataRDD.min()
minimum: Int = 2
```

```
scala> println(s"Maximum element: ${maximum}")
Maximum element: 25
```

```
scala> println(s"Minimum element: ${minimum}")
Minimum element: 2
```

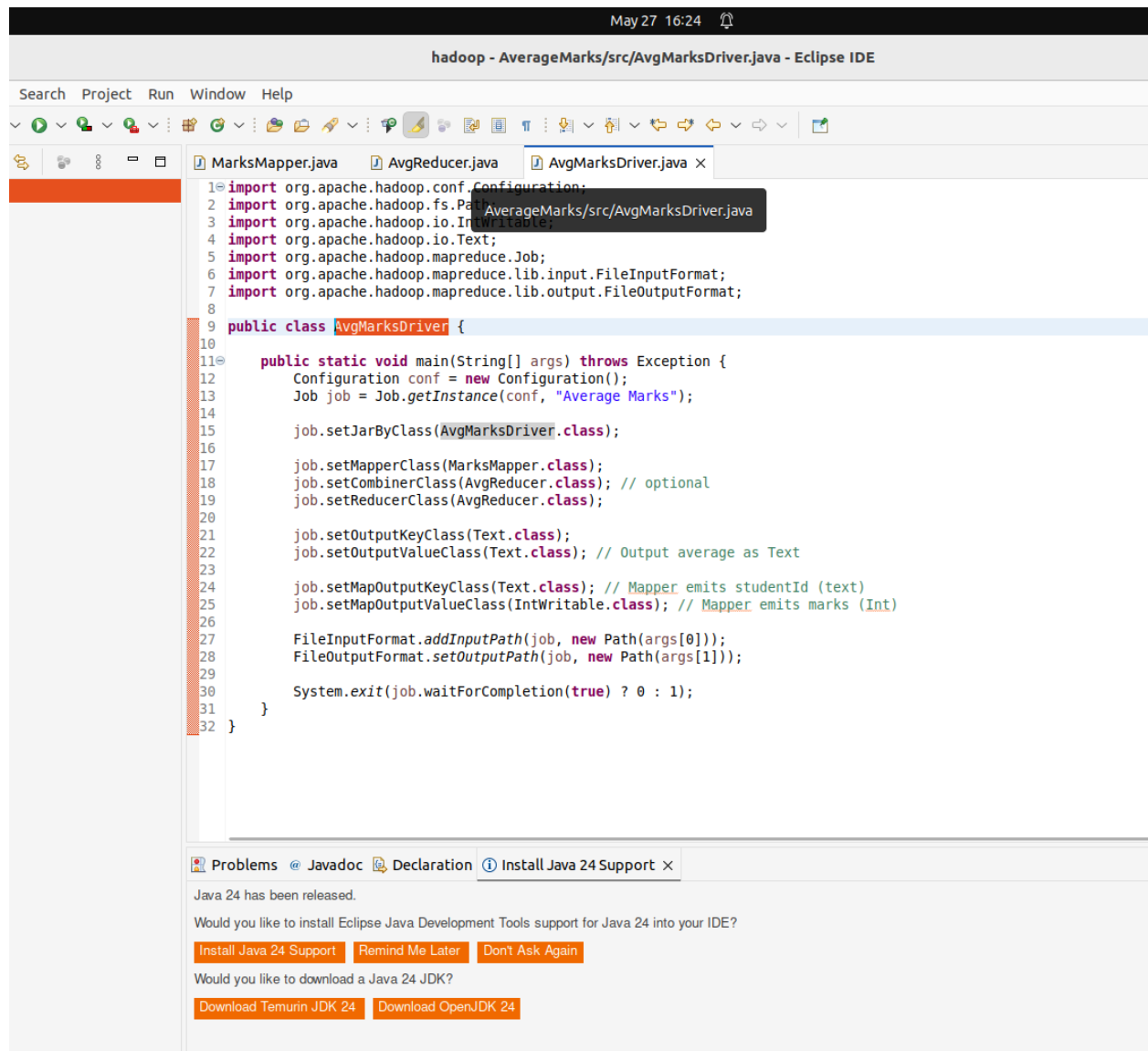
Hadoop program for average marks of student

```
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC: ~/Desktop
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ 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@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ cd Desktop
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ nano student_marks.txt
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~/Desktop$ cd ~
hadoop@bmscsecse-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]
localhost: namenode is running as process 6230. Stop it first and ensure /tmp/hadoop-hadoop-namenode.pid file is empty before retry.
Starting datanodes
localhost: datanode is running as process 6409. Stop it first and ensure /tmp/hadoop-hadoop-datanode.pid file is empty before retry.
Starting secondary namenodes [bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC]
bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC: secondarynamenode is running as process 6701. Stop it first and ensure /tmp/hadoop-hadoop-secondarynamenode.pid file is empty before retry.
Starting resource manager
bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC: resourcemanager is running as process 6983. Stop it first and ensure /tmp/hadoop-hadoop-resourcemanager.pid file is empty before retry.
Starting node managers
localhost: nodemanager is running as process 7145. Stop it first and ensure /tmp/hadoop-hadoop-nodemanager.pid file is empty before retry.
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ jps
6230 NameNode
17702 org.eclipse.equinox.launcher_1.6.1000.v20250227-1734.jar
6983 ResourceManager
7145 NodeManager
19289 Jps
6409 DataNode
6701 SecondaryNameNode
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -mkdir /shu
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -copyFromLocal /home/hadoop/Desktop/student_marks.txt /shu/tet.txt
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop jar /home/hadoop/Desktop/avgmarks.jar AvgMarksDriver /shu/tet.txt /shu/out1
2025-05-27 16:03:30,210 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2025-05-27 16:03:30,244 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2025-05-27 16:03:30,244 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2025-05-27 16:03:30,296 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface and execute your application.
2025-05-27 16:03:30,341 INFO Input.FileInputFormat: Total input files to process : 1
2025-05-27 16:03:30,371 INFO mapreduce.JobSubmitter: number of splits:1
2025-05-27 16:03:30,427 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1821707493_0001
2025-05-27 16:03:30,427 INFO mapreduce.JobSubmitter: Executing with tokens: []
2025-05-27 16:03:30,477 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2025-05-27 16:03:30,477 INFO mapreduce.Job: Running job: job_local1821707493_0001
2025-05-27 16:03:30,477 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2025-05-27 16:03:30,480 INFO output.PathOutputCommitterFactory: No output committer factory defined, defaulting to FileOutputCommitterFactory
2025-05-27 16:03:30,481 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2025-05-27 16:03:30,481 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failed: true
2025-05-27 16:03:30,481 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter
2025-05-27 16:03:30,527 INFO mapred.LocalJobRunner: Waiting for map tasks
2025-05-27 16:03:30,527 INFO mapred.LocalJobRunner: Starting task: attempt local1821707493_0001_m_000000_0
```

Final output:

```
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -ls /shu/out1
Found 2 items
-rw-r--r-- 1 hadoop supergroup 0 2025-05-27 16:18 /shu/out1/_SUCCESS
-rw-r--r-- 1 hadoop supergroup 8 2025-05-27 16:18 /shu/out1/part-r-00000
hadoop@bmscsecse-HP-Elite-Tower-800-G9-Desktop-PC:~$ hadoop fs -cat /shu/out1/part-r-00000
1 90.67
2 82.67
3 81.67
4 67.50
```

Driver:



May 27 16:24

hadoop - AverageMarks/src/AvgMarksDriver.java - Eclipse IDE

Search Project Run Window Help

MarksMapper.java AvgReducer.java AvgMarksDriver.java x

```
1 import org.apache.hadoop.conf.Configuration;
2 import org.apache.hadoop.fs.Path;
3 import org.apache.hadoop.io.IntWritable;
4 import org.apache.hadoop.io.Text;
5 import org.apache.hadoop.mapreduce.Job;
6 import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
7 import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
8
9 public class AvgMarksDriver {
10
11     public static void main(String[] args) throws Exception {
12         Configuration conf = new Configuration();
13         Job job = Job.getInstance(conf, "Average Marks");
14
15         job.setJarByClass(AvgMarksDriver.class);
16
17         job.setMapperClass(MarksMapper.class);
18         job.setCombinerClass(AvgReducer.class); // optional
19         job.setReducerClass(AvgReducer.class);
20
21         job.setOutputKeyClass(Text.class);
22         job.setOutputValueClass(Text.class); // Output average as Text
23
24         job.setMapOutputKeyClass(Text.class); // Mapper emits studentId (text)
25         job.setMapOutputValueClass(IntWritable.class); // Mapper emits marks (Int)
26
27         FileInputFormat.addInputPath(job, new Path(args[0]));
28         FileOutputFormat.setOutputPath(job, new Path(args[1]));
29
30         System.exit(job.waitForCompletion(true) ? 0 : 1);
31     }
32 }
```

Problems Javadoc Declaration Install Java 24 Support x

Java 24 has been released.

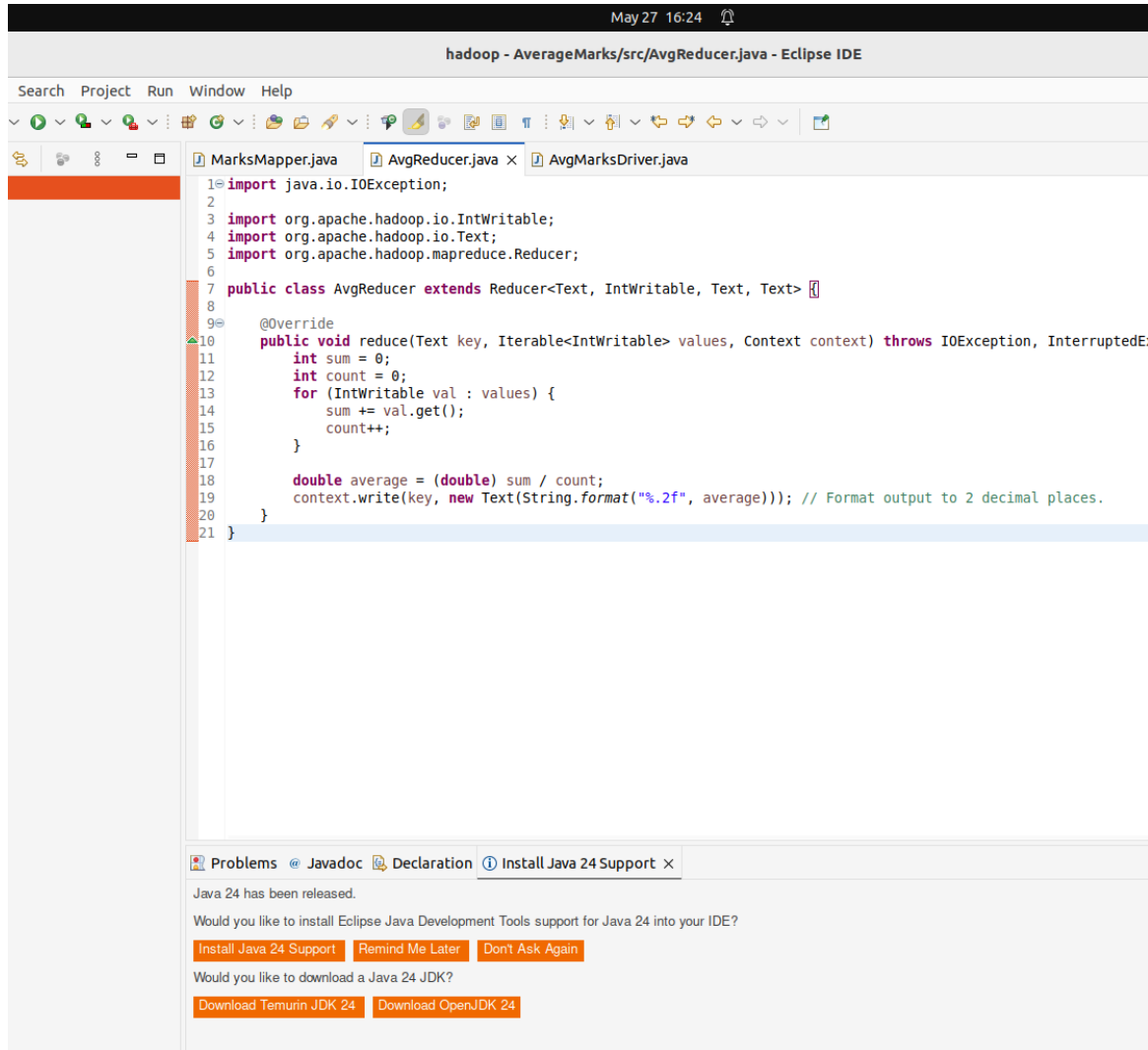
Would you like to install Eclipse Java Development Tools support for Java 24 into your IDE?

Install Java 24 Support Remind Me Later Don't Ask Again

Would you like to download a Java 24 JDK?

Download Temurin JDK 24 Download OpenJDK 24

Reducer code :



```
1 import java.io.IOException;
2
3 import org.apache.hadoop.io.IntWritable;
4 import org.apache.hadoop.io.Text;
5 import org.apache.hadoop.mapreduce.Reducer;
6
7 public class AvgReducer extends Reducer<Text, IntWritable, Text, Text> {
8
9     @Override
10    public void reduce(Text key, Iterable<IntWritable> values, Context context) throws IOException, InterruptedException {
11        int sum = 0;
12        int count = 0;
13        for (IntWritable val : values) {
14            sum += val.get();
15            count++;
16        }
17
18        double average = (double) sum / count;
19        context.write(key, new Text(String.format("%.2f", average))); // Format output to 2 decimal places.
20    }
21 }
```

Java 24 has been released.

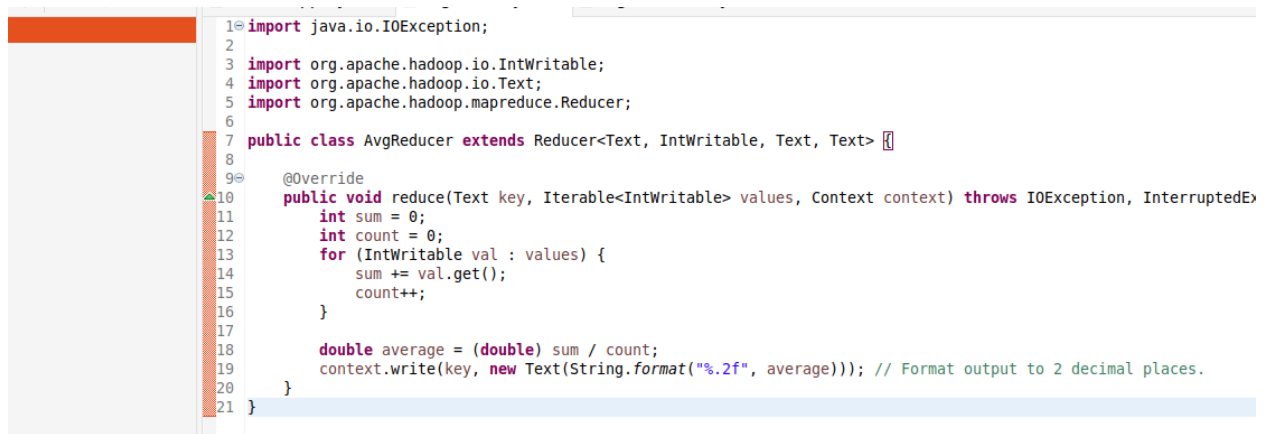
Would you like to install Eclipse Java Development Tools support for Java 24 into your IDE?

[Install Java 24 Support](#) [Remind Me Later](#) [Don't Ask Again](#)

Would you like to download a Java 24 JDK?

[Download Temurin JDK 24](#) [Download OpenJDK 24](#)

Mapper



```
1 import java.io.IOException;
2
3 import org.apache.hadoop.io.IntWritable;
4 import org.apache.hadoop.io.Text;
5 import org.apache.hadoop.mapreduce.Reducer;
6
7 public class AvgReducer extends Reducer<Text, IntWritable, Text, Text> {
8
9     @Override
10    public void reduce(Text key, Iterable<IntWritable> values, Context context) throws IOException, InterruptedException {
11        int sum = 0;
12        int count = 0;
13        for (IntWritable val : values) {
14            sum += val.get();
15            count++;
16        }
17
18        double average = (double) sum / count;
19        context.write(key, new Text(String.format("%.2f", average))); // Format output to 2 decimal places.
20    }
21 }
```

