

Practice 4: ToolRunner to pass parameters

Eclipse package/project `practice_4`

Java files: `MR2ToolRunner.java`

Data: `shakespeare`

In this Exercise, you will implement a driver using ToolRunner.

ToolRunner code resembles code you wrote/used in Practice_3. You will tweak that code so that your program now implements a Tool.

Pass a parameter programmatically

1. Modify the driver - AvgWordLength.java

Set a new Boolean configuration parameter called `caseSensitive` . Add the following code to your driver:

```
conf.setBoolean("caseSensitive", false)
```

2. Add a setup() method to LetterMapper.java

Mapper contains a method called `setup()` . It is used to set parameters for the `map()` method. It runs only once before there are any calls to the `map()` method.

In `setup()` get the value of the configuration parameter `caseSensitive`, and use it to set a member variable indicating whether to do case sensitive or case insensitive processing as follows:

```
@Override
public void setup(Context context) {
    Configuration conf = context.getConfiguration();
    caseSensitive = conf.getBoolean("caseSensitive", false);
}
```

3. Modify the map() method in LetterMapper.java

In the map() method, choose whether to do case sensitive processing (leave the letters as-is), or insensitive processing (convert all letters to lower-case) based on the caseSensitive member variable.

Pass a parameter as a runtime parameter

Test your code twice in Eclipse, once passing false and once passing true.

When set to true, your final output should have both upper and lower case letters; when false, it should have only lower-case letters. The RunConfiguration arguments should look like this:

```
-D caseSensitive=true data/shakespeare output/toolrunner
```

Test passing both true and false to confirm the parameter works correctly. □

If you like you can also run against the cluster by building the jar with Maven Install and running thus:

```
$ hadoop jar <name of practice.jar> hints.MR2ToolRunner \  
-DcaseSensitive=true <hdfs location of shakespeare> <output dir>
```

This is the end of the Exercise