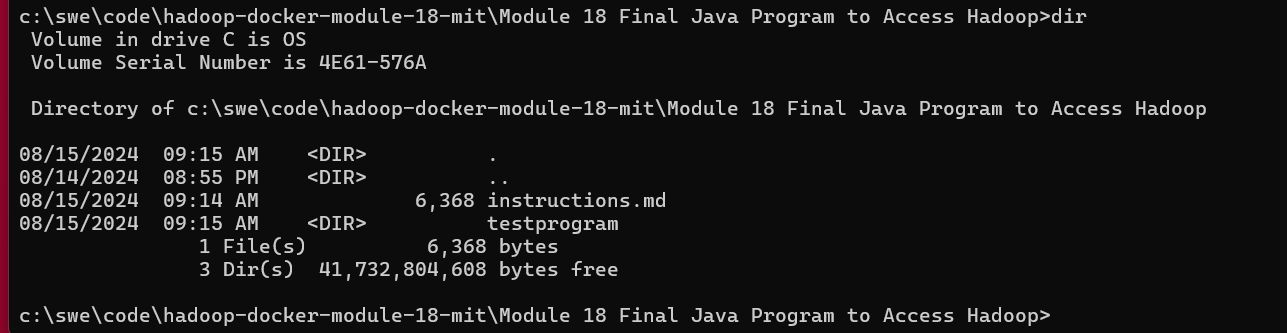
**Submission Instructions:**

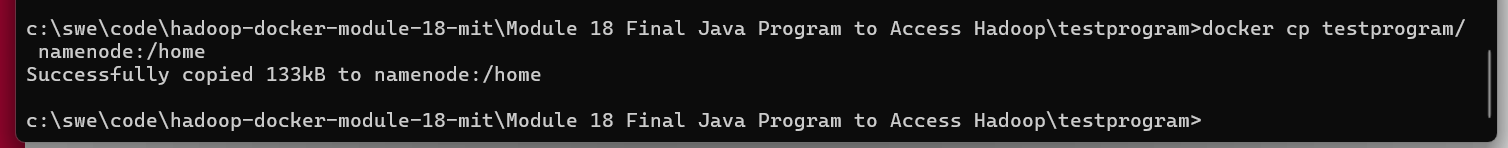
Your submission for this assignment should be a Word document that includes the following screenshots, each labeled for the step that the screenshot represents:

**Part 1: Ingesting Data into the HDFS**

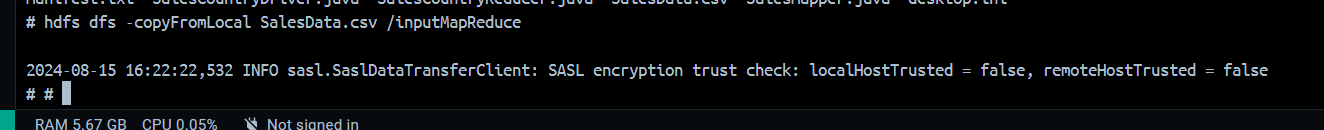
1. Provide a screenshot to show that you unzipped the testprogram.zip file.



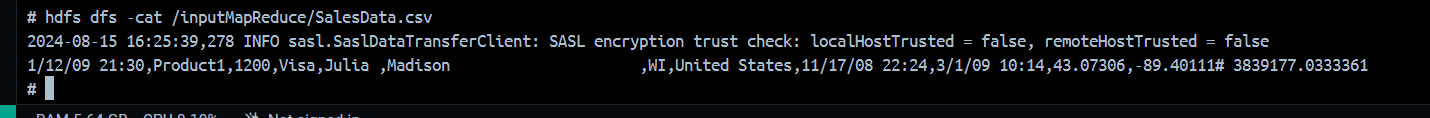
1. Provide a screenshot to show that you copied the testprogram folder into the home *directory* of the Hadoop namenode.



1. Provide a screenshot to show that you successfully copied the SalesData.csv file into the inputMapReduce folder.



1. Provide a screenshot to show that the file has been copied using the HDFS cat command.



**Part 2: Performing MapReduce — Aggregation Sales by Country**

1. Provide a description of the three Java files listed below that are used to apply the MapReduce framework on the Hadoop database:
   * SalesCountryDriver.java

The SalesCountryDriver class sets up a Hadoop MapReduce job by configuring the job name, output key/value types, and specifying the mapper and reducer classes. This setup is essential for running a MapReduce job that processes sales data per country. The class leverages Hadoop's JobClient and JobConf to manage and configure the job, ensuring that the appropriate classes and data types are used for the MapReduce process.

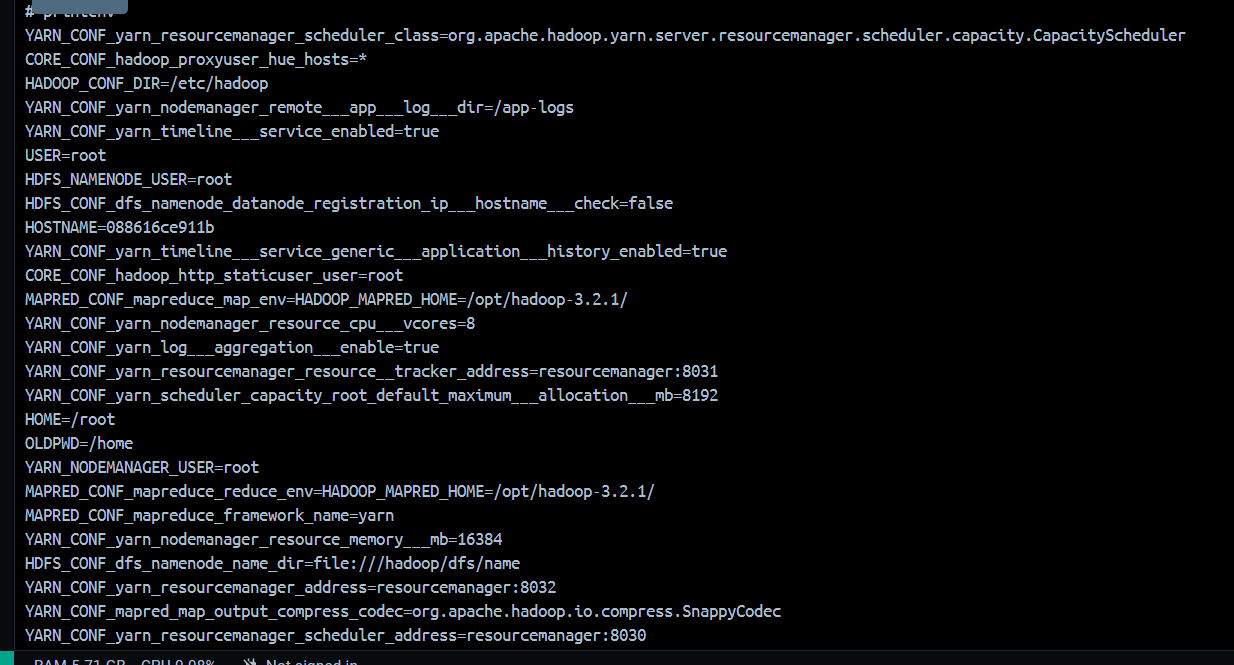
* + SalesMapper.java

The SalesMapper class reads input data line by line, splits each line by commas, and emits the 8th field of each line as the key with a constant value of 1. This is typically used to count occurrences of the 8th field in the input data.

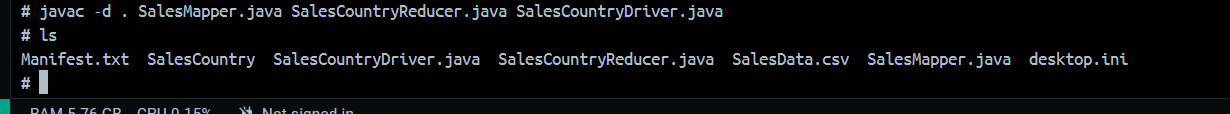
* + SalesCountryReducer.java

The SalesCountryReducer class processes the key-value pairs produced by the mapper phase, where the key is a country and the value is a count. It sums up the counts for each country and emits the total count for each country. This is typically used to count the total occurrences of sales per country.

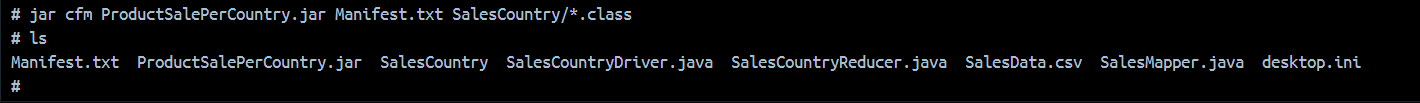
1. Provide a screenshot to show that you have successfully defined the environment variables in the namenode CLI.



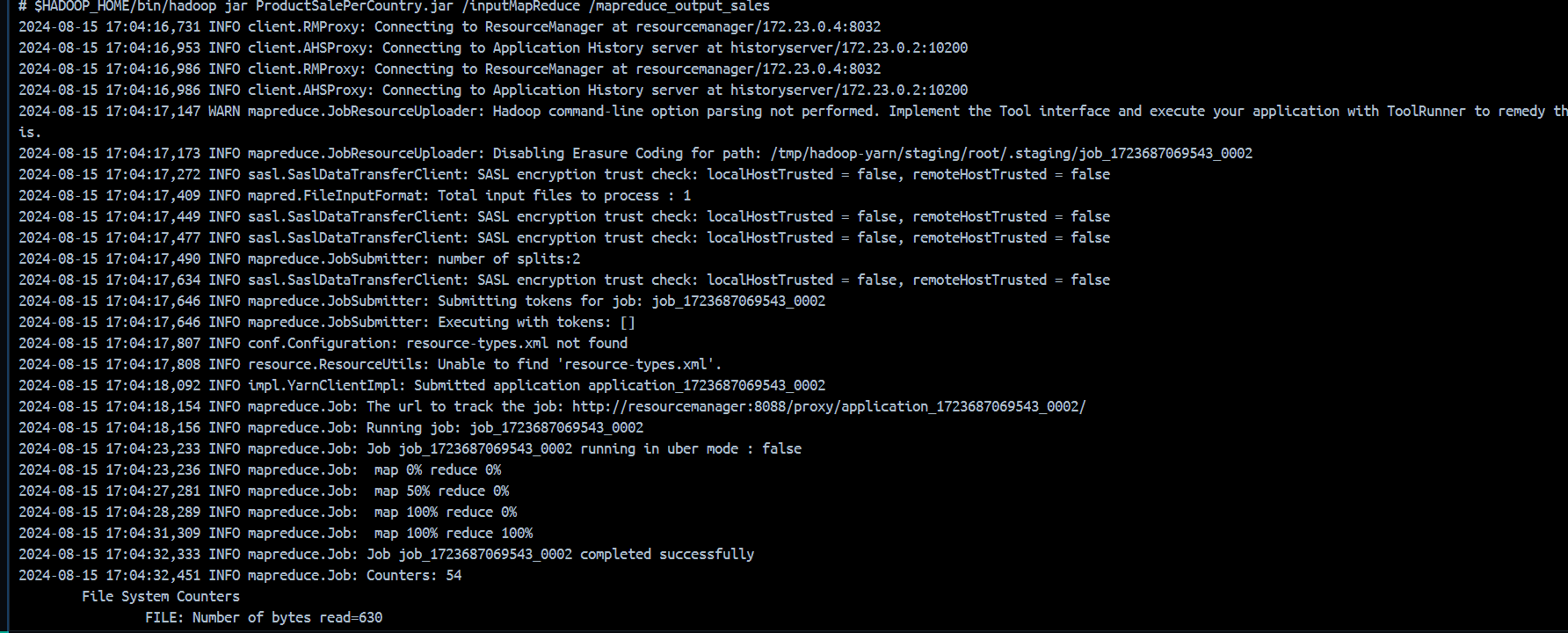
1. Provide a screenshot to show that you successfully compiled the Java files in the SalesCountry folder.



1. Provide a screenshot to show that you successfully created a jar file from the compiled Java code.



1. Provide a screenshot to show that you successfully ran the MapReduce operation to distribute the analysis of the data.



1. Provide a screenshot to show that you successfully visualized the content of the part-00000 file inside the mapreduce\_output\_sales folder.

