Stephen\_Re\_Assignment\_18

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

1. Provide a screenshot to show that you unzipped the testprogram.zip file.  
   A screenshot of a computer program

   Description automatically generated
2. Provide a screenshot to show that you copied the testprogram folder into the home *directory* of the Hadoop namenode.  
   
3. Provide a screenshot to show that you successfully copied the SalesData.csv file into the inputMapReduce folder.  
   A black background with white text

   Description automatically generated
4. Provide a screenshot to show that the file has been copied using the HDFS cat command.  
   A screen shot of a computer

   Description automatically generated

**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

Driver code for the Hadoop job. it is used to define

* + 1. the parameters of the job
    2. Mapper and reducer classes
    3. Input and output formats
    4. Output data types
    5. mapper and reducer names
  + SalesMapper.java

Mapper code for the Hadoop job

* + 1. From the input file, convert each line to a string
    2. Split the string on the comma
    3. For each new occurrence of a country name (position 7 in the line), assign the number 1
  + SalesCountryReducer.java

Reducer code for Hadoop job. counts the number of times a country occurs in the csv file.

* + 1. From the mapper output, iterate through each line key (country name) and value( 1) pair
    2. For each key encountered, += the value

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

   Description automatically generated
2. Provide a screenshot to show that you successfully compiled the Java files in the SalesCountry folder.  
   A computer screen with red arrows pointing to a black background

   Description automatically generated
3. Provide a screenshot to show that you successfully created a jar file from the compiled Java code.  
   A black screen with white text

   Description automatically generated
4. Provide a screenshot to show that you successfully ran the MapReduce operation to distribute the analysis of the data.  
   A screenshot of a computer screen

   Description automatically generated
5. Provide a screenshot to show that you successfully visualized the content of the part-00000 file inside the mapreduce\_output\_sales folder.  
   A computer screen shot of a black screen

   Description automatically generated