**Steps to consider to run the project**

1. Create the following directories(You will find it in the submitted packages) on cloudera Desktop
   1. ProjectInputFiles
   2. ProjectJarFiles
2. If you need input file, Put it inside the “input” directory under its respective package name in the directory step1-a. The package names for each problem are listed below.
   * + Part\_1A, Part\_1B, Part\_1C, Part\_1D, Part\_2, Part\_3, Part\_4
3. Open terminal
4. Sudo su hdfs
   1. Run the following batch file to process the MapReduce for the problem we want just by giving the package name as an argument as follows
      * Bash-4.1$ Desktop/Execute Part\_1A
5. Exit from hadoop environment to local environment
   1. Similarly Run the following batch file to copy the MapReduce Outputs to their respective local directory
      * [cloudera@quickstart ~ ] Desktop/CopyOutputToLocal Part\_1A
6. The output files could be found in the ProjectInputFiles/{Package-Name}/Output

Note: each java files can be found in the Ecliplse project name under the directory “BigDataCourseProject” with in the submitted package (BigData-Project-Group1).