STEPS TO FOLLOW:

TASK1
- To run the code which is in the Jar file for task1 write the below command
  - `./spark-submit --class Task1 /Path/of/your/Jar/File/ /Path/of/your/input_file /Path/of/your/output/folder/`
  - `My Jar file name is Task1.jar so the path will be as /Path/where/your/jar/file/Task1.jar`
  - `For the input file /Path/where/the/input/file/survey_results_public.csv`
  - `For the output file` **please note** `it will create a folder and inside folder there will be a file which will start with "part".`
  - `For output file for me it was as I named` **folder as output** `so it will be something like: Path/where/the/output/folder/`**output**
  - **P.S.** `inside the output folder there will be a csv file which starts with "part"`
  - **E.g.** `-- ./spark-submit --class Task1 /Users/IdeaProjects/Task1/out/artifacts/Task1_jar/Task1.jar "/Users/Desktop/DataMining/stack-overflow-2018-developer-survey/survey_results_public.csv" "/Users/Desktop/DataMining/stack-overflow-2018-developer-survey/output"`

TASK2
- To run the code which is in the Jar file for task2 write the below command
  - `./spark-submit --class Task2 /Path/of/your/Jar/File /Path/of/your/input_file /Path/of/your/output/file.csv`
  - `My Jar file name is Task1.jar so the path will be as <Path/where/your/jar/file/Task1.jar`
  - `For the input file /Path/where/the/input/file/survey_results_public.csv`
  - For output file `/Path/where/the/output/file/output.csv`
  - **P.S.** Here you will name the **csv file**
  - **E.g.** `./spark-submit --class Task2 /Users/IdeaProjects/Task1/out/artifacts/Task1_jar/Task1.jar "/Users/Desktop/DataMining/stack-overflow-2018-developer-survey/survey_results_public.csv" "/Users/Desktop/DataMining/stack-overflow-2018-developer-survey/output.csv"`

TASK3
- To run the code which is in the Jar file for task1 write the below command
  - `./spark-submit --class Task3 /Path/of/your/Jar/File/ /Path/of/your/input_file /Path/of/your/output/folder/`
  - `My Jar file name is Task1.jar so the path will be as /Path/where/your/jar/file/Task1.jar`
  - `For the input file /Path/where/the/input/file/survey_results_public.csv`

- For the output file **please note** it will create a folder and inside folder there will be a file which will start with "part".
- For output file for me it was as I named **folder as output** so it will be something like: Path/where/the/output/folder/**output1**
- **P.S.** inside the output folder there will be a csv file which starts with "part"
- **E.g.** ./spark–submit ––class Task3 /Users/IdeaProjects/Task1/out/artifacts/Task1_jar/Task1.jar "/Users/Desktop/DataMining/stack–overflow–2018–developer–survey/survey_results_public.csv" "/Users/Desktop/DataMining/stack–overflow–2018–developer–survey/output1"