

Read Me Document

To run the tasks, follow the below instructions

Task 1

- Run the following command in the terminal

```
./spark-submit --class Task1 /path/to/jar/file/parth_vaghani.jar "/path/to/input  
file/survey_results_public.csv" "/path/to/output/folder"
```

- Output csv file will be generated in the output folder. The name of the output file will be part-*

Task 2

- Run the following command in the terminal

```
./spark-submit --class Task2 /path/to/jar/file/parth_vaghani.jar "/path/to/input  
file/survey_results_public.csv" "/path/to/output/folder/parth_vaghani_task2.csv"
```

- A csv file with the name "parth_vaghani_task2" will be generated in the specified path.

Task 3

- Run the following command in the terminal

```
./spark-submit --class Task3 /path/to/jar/file/parth_vaghani.jar "/path/to/input  
file/survey_results_public.csv" "/path/to/output/folder"
```

- Output csv file will be generated in the output folder. The name of the output file will be part-*