**UCS654 - Predictive Analytics Using Statistics**

**Assignment01 - Marks Analysis**

|  |
| --- |
| **General Instructions – Must Read**   * **Submission Due Date:** 30 Jan 2022 | 23:59:59 * **Marks:** 05 (Five) * **Number of Questions:** 02 * **Submission Link:** [**Click Here**](https://forms.gle/zKpcmpkV8LBzbCmb9) * **Submission Guidelines:** You need to submit TWO python files only. * One python (.py) file for 1 | File Name must be <YourRollNum-1>.py | Example: **10155-1.py** * One python (.py) file for 2 | File Name must be <YourRollNum-2>.py | Example: **10155-2.py** * Your program must be run from **command line** only: * **Usages:** python <program.py> <InputDataFile> * **Example:** python 10155-1.py input.csv * **Example:** python 10155-2.py output.csv * Your program must be capable to handle exception (if any) and write to **log file**: * Correct number of parameters (inputFileName). * Show the appropriate message for wrong inputs. * Handling of “File not Found” exception * Input file must contain three columns only. * If any issue with the input record it must be write to a log file * **Note:**   + Multiple submissions are allowed, but **latest submission** will be considered for the evaluation.   + Submission link will open all the time, but only 50% marks will be awarded if you fail to submit with in the due date. No excuse will be consider for the submission.   + **Zero marks** will be awarded for plagiarized code or result. |

1. **Write a python program that converts the input file to output file.** [Input file is available in “Input for Assignment01” folder]

|  |  |
| --- | --- |
| **input file** | **output file** |
|  |  |

1. **WAP a python program that reads the output file (of 1.1) and generates**

* Multiple plots such as Histogram, Line chart, Pie chart, etc for count/total for P1, P2, P3, P4, P5, Total-of-all}. All the plots must be saved into multiple .png files (filenames must be <RollNo>-<Plot-type>.png | Example 101903001-histrogram-count.png, 101903001-histrogram-total.png 101903001-line-count.png, 101903001-line-total.png, etc).
* Generate the different statistics and save to a txt file: such as {Min, Max, Mean, Median, SD, distribution, Count number of missing values, Count Non numeric Values, etc for P1, P2, P3, P4, P5, Total-of-all.

**Please note: I/O for each program**

|  |  |  |
| --- | --- | --- |
| **Program Name** | **Input File** | **Output File** |
| **10155-1.py** | input.csv | 1019033001-output.csv |
| **10155-2.py** | output.csv | 1019033001-histogram.png  1019033001-line.png  1019033001-pie.png  1019033001-graph1.png  1019033001-graph2.png  1019033001-graph3.png  1019033001-statistics.txt  **1019033001-log.txt** |