Christopher Kruger

4 October 2023

Homework 3 – Narrative of the File

By taking in the two csv files based off historical data of Citi Bike usage in NYC for the Spring season and Winter season of 2016, the program allows the user to understand the usage statistics without having to read the original file with difficulty. By giving an output that is easily readable, this program is far superior to having to try and understand the csv file.

The program begins by printing the last five lines in the Winter season as brief context for what the original data looked like. Following this, the program tells the user how many data points were recorded in the original file (excluding the header). From this total number, the user is also told how many of which were Customers or Subscribers. The percentage of customers compared to all users is also given.

Next, the program gives the first five lines in the Spring season data set as a reference. The program will again tell the user how many total data points were recorded (excluding the header), how many customers there were, and how many subscribers there were. The program then tells the user which file is smaller in total data point size, then says which file/season had a higher percentage of customers vs total entries. To summarize the information presented, the program then tells the user which file or season had more riders than the other. It also says which file or season had more customers/non-subscribers than the other.

In conclusion: the winter season had 21350 data points, 1057 customers, 20293 subscribers, and customers representing 4.95% of the total user base. The spring season had 44615 data points, 5841 customers, and 38774 subscribers. There were more riders and customers/non-subscribers in the Spring than in the Winter.