**Assumptions used for Data Analysis:**

1. A given unit on the xy-plane is equal to 1 miles geographically.
2. The bus can house and charge 100 scooters at a given time.
3. The bus will travel between clusters at an average close to its maximum speed (50 mph).
4. Travel time between scooters within a given cluster is negligible.
5. The time required to charge a given scooter via the bus is equal to:
6. A cluster is defined as a group of scooters within 0.4 radial units of the central scooter location.

**Introduction:**

Scooters are a fast and energy efficient way of moving around a busy city landscape. They provide the consumers affordable and easily available transportation without adding to traffic congestion. The only problem with these scooters is the fact that they run on a finite energy supply stored within their batteries and must be charged in order to stay operational. In order to do so, the Xtern Xpress ridesharing service needs an efficient plan to charge scooters after consumers are done using them.

The scooter location and battery power data, given by the TechPointX team, was utilized in calculating the number of scooters within a given cluster and the subsequent battery power average of said clusters. The data of distance from each scooter to the mega charging bus was also calculated but later realized to be irrelevant to the findings of the evaluation. Therefore, the route by the mega bus and the overall strategy of the charging system was found to be connected to specific scooter distance from the bus parking location.

**Programming Used for Data Manipulation (C++):**

The data calculations and final results found throughout this analysis were calculated through programs written in C++ code for the evaluation of large data samples. These files include cluster\_mapping.cpp and distance\_eq.cpp.

In the distanc\_eq.cpp file, the class for point values was created as seen below. It was made up of two variables of type double and consisted of the proper get and set member functions as seen in Figure 1 on the next page. The point class was essential to evaluating the point coordinates of the data set given using a single vector for all scooter point values. The ID of each scooter was then understood to be the vector location of the scooter.

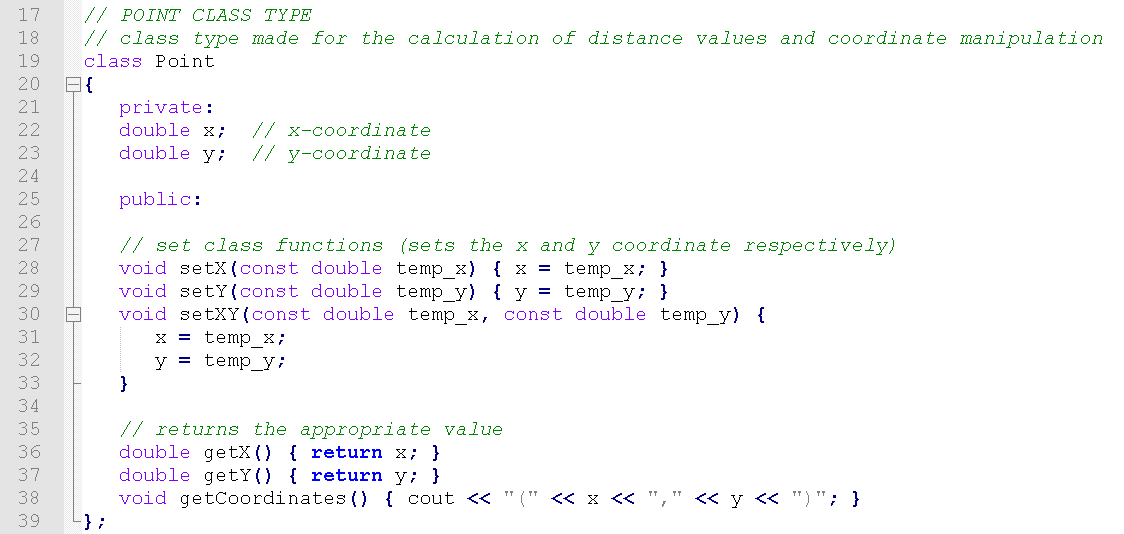


Figure 1: Point Class

To read in the scooter data values, a function was written that used the simplified data table (put into .txt file format) to add elements into proper vectors. This function disregarded the scooter ID because it was understood that the vector location given by the program would be equal to its ID number. This can be seen in Figure 2 below.

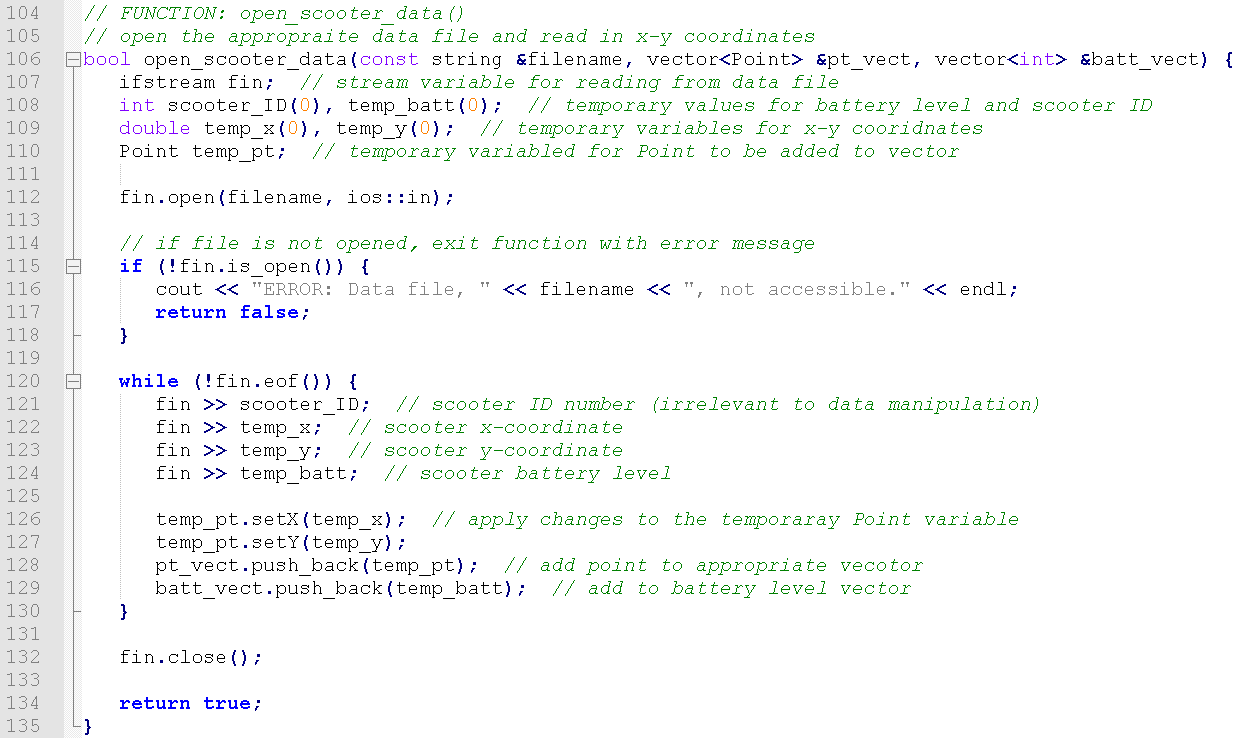


Figure 2: open\_scooter\_data() Function Definition

Once the data from the files was read into the program and assigned to the appropriate vectors. The program calculated the distance of each scooter from the bus and wrote those values to a .csv file so that the values could be graphed. Unfortunately, after calculation and subsequent transfer of values to the main data set excel file, it was observed these distance values played no role in what the route of the bus and ultimate charging method for the scooters would be. That being said, the point class, distance equation, and the reading and writing functions would be utilized to some extent in the cluster\_mapping.cpp file which yielded important results.

The file cluster\_mapping.cpp was written to evaluate the number of scooters and average battery value of each of the 19 clusters. This was done by approximating the center point for all of the clusters seen in Figure X (figure shown in the Data Analysis and Results section) and calculating all points within these clusters using a cluster radii of 0.04 coordinate units and the distance function shown in Figure 3 below. This was done after the values of the scooters was read in using the open\_scooter\_data() function shown in Figure 2.

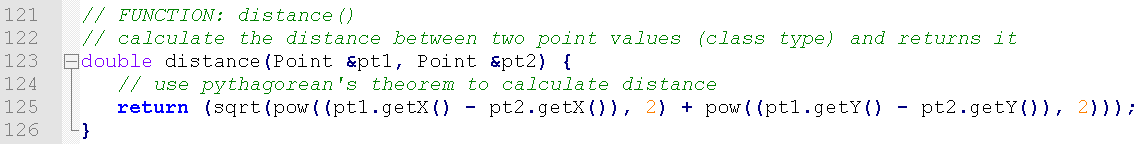


Figure 3: distance() Function Definition

The scooter locations were categorized into the appropriate their cluster by comparing their respective location to that of the central coordinate of all the clusters. If they were within 0.04 units of the central coordinate, they were understood to be located in that cluster. This process can be seen in the Figure 4 code below.

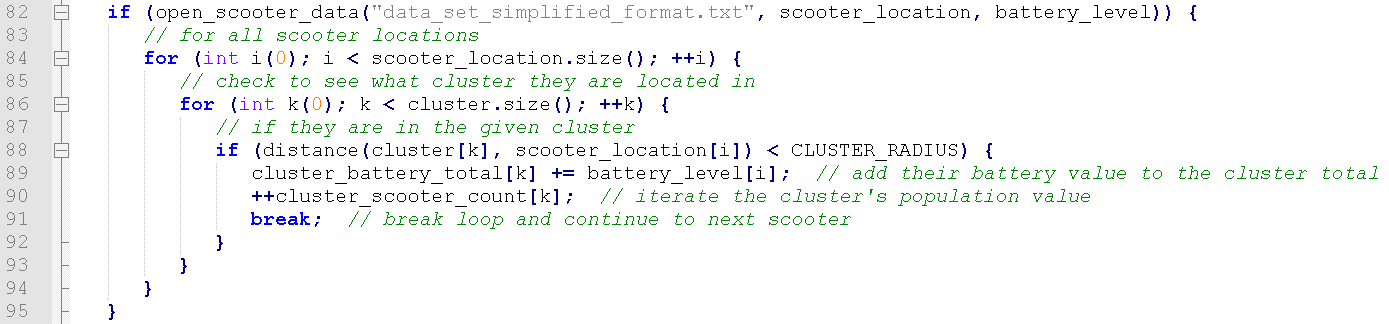


Figure 4: Cluster Mapping for Loops and if Statement

**Data Analysis and Results:**

It should be noted that the total scooter population calculated by finding all scooters within each cluster is roughly 99.96% the actual number of scooters owned by the Xtern Xpress ridesharing company. This is because of the error associated with the cluster radius size and its inability to account for all scooters within every cluster. That being said, the data and percentages calculated by this process should be considered close enough to the actual data of each cluster because of the extremely small error found in the exclusion of only 0.04% of scooters.

The scooter data given contained xy-coordinates for more than 25000 individual scooters along with their ID numbers and their respective battery charge. These scooters were found to be organized across a coordinate plane in 19 differing clusters, or regions of high scooter concentration. The plot below (Figure 5) displays this trend along with the cluster ID numbers used in the data manipulation and route planning.

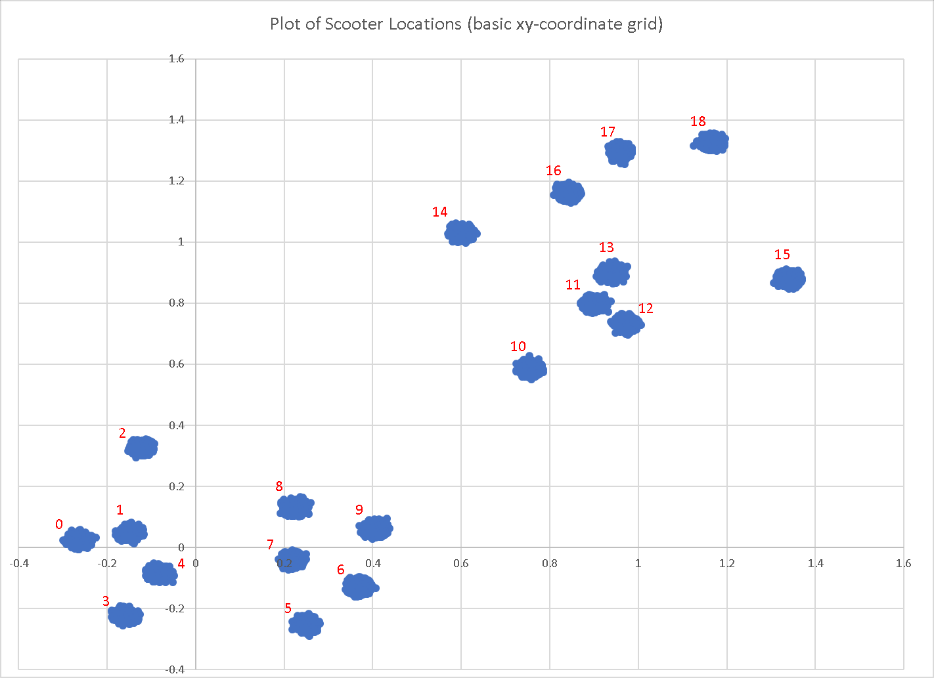


Figure 5: Scooter Locations and Cluster ID Numbers

The 19 clusters displayed in the plot have a relatively equal distribution of scooters apart from that of cluster 8, which contains 8.58% of the total scooter population, making it the most densely packed of the 19 clusters as seen in Figure 6 on the next page. All other clusters had a scooter population equal to roughly 5-6% with a few small deviations. This data can be found in Table 1, located on the next page.

Although cluster 8 has the highest specific cluster percentage of scooters, the clusters 11-13 are close enough to one another that travel between them would be relatively short and therefore negligible to the travel time of the mega charging bus. The combination of scooters within these three clusters is equal to 4070 scooters, or 15.86% the total population.

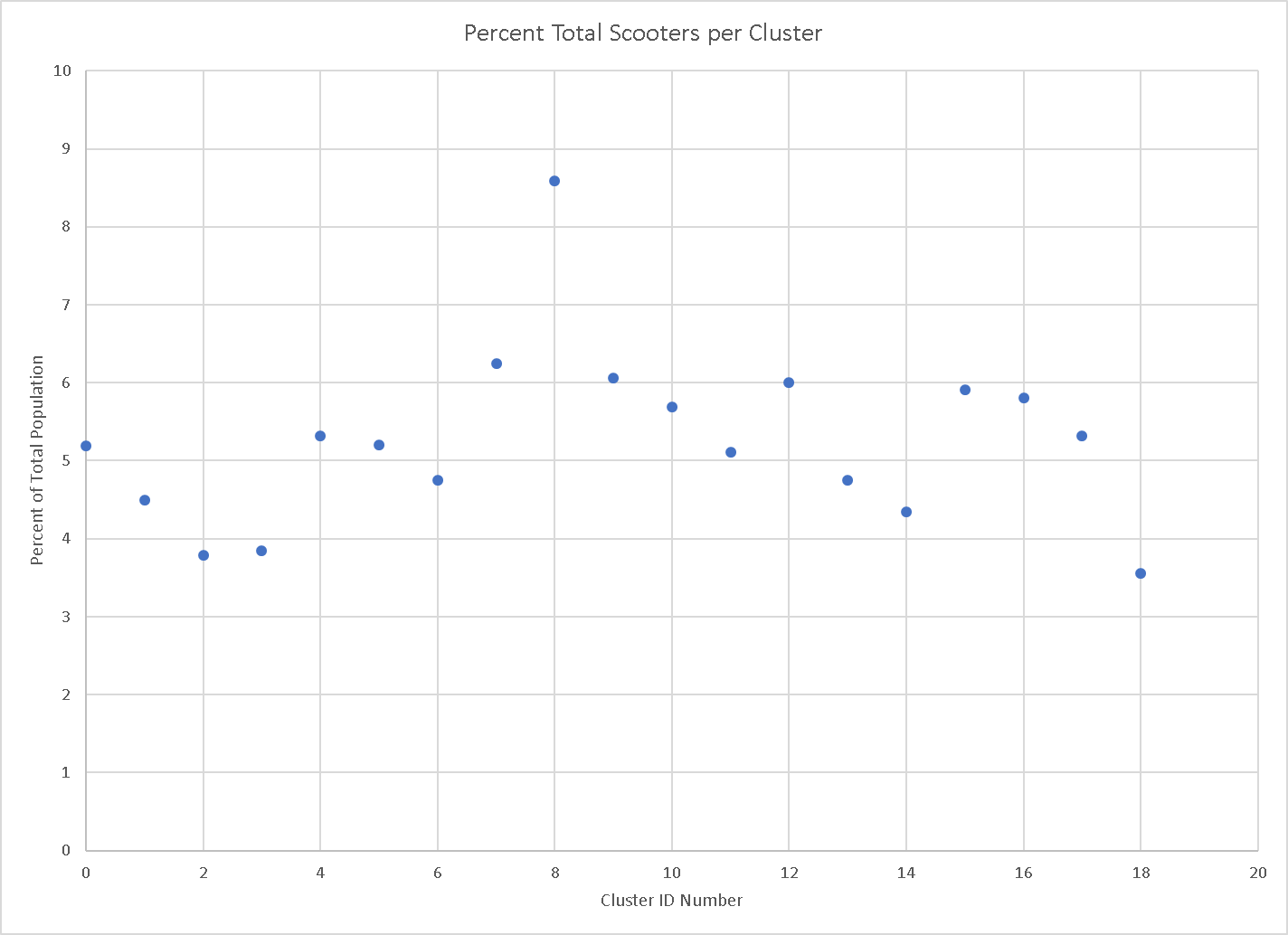


Figure 5: Scooter Percentage per Cluster



Table 1: Cluster Data

Within the cluster data table, the average battery power contained in each scooter was calculated in order to evaluate what clusters need the most initial attention by the Xtern Xpress charging team. With all clusters having a rough 2.5 charge rating, this data has no impact on the final findings of the charging system put in place by the company and the route taken by the bus. That being said, using this data, it is possible to calculate the average time needed to charge each bike which is proportional to (5 – average battery power). Therefore, the average bike needs approximately 2.5 hours of charge in order to reach full capacity.