Report

In solving the Career Fair Programming challenge, in addition to the ***VehicleLogAnalysis*** class, I created two extra classes named ***InterestPoint.java*** and ***Store.java****.* The former class held data about an interest point such as the longitude, latitude, radius, and id. The latter class held data about the number of times the bus has passed through an interest point, total time, date, and id of a trip. I also created methods called ***calculateDistance****,* ***getMean****,* ***getVariance****,* and***getStandardDeviation***to get the distance between two GPS points, get the mean, get the variance and the standard deviation respectively. I researched about how to calculate the distance between two GPS points and how to compute the variance and standard deviation of a data set before I could implement the methods discussed above. My approach of solving each task is discussed below under each task heading.

Task 1

In solving this task, I read the vehicle log file and picked out the longitude and latitude where the bus parked each night. I did this by picking out the last GPS point of each day. I then found the average of the GPS points which became the point the bus parked each night.

Task 2

In solving this problem, I created an interest point object for each interest point on the "1.interest\_points.csv" file and a store object for each interest point. Next, I looped through each GPS point on the "2.vehicle\_log.csv" file and computed the distance between that point and each of the interest points and compared the distance with the radius of each interest point. If the distance was less than any interest point’s radius then the bus passed through that point, then I stored the id and date in the store object. Next, I compared the previous id and the current id, if they were different, I increased the count variable in the corresponding store object of that interest point. If the ids were the same, I computed the time and added it to the total time variable in the store object. I did this for each point. In the end, the count variable held the number of times the bus passed through an interest point and the total time variable held the total time in minutes each time the bus passed through the same point. Next, I found the average time spent in each entry by dividing the total time spent in each interest point by the number of times the bus passed through it.

Task 3

For this task, I computed the duration in minutes of each trip made by the bus. I Classified all trips with a duration of less than 13 minutes as unusual trips made by the bus. Next, I computed the standard deviation of all trips whose duration were equal or more than 13 minutes and I used 2 standard deviations above the mean to compute my upper bound. Any trip that fell above my upper bound was also classified as an unusual trip.

# References

Khanacademy. (n.d.). *Calculating Standard Deviation Step by Step*. Retrieved from KHANACADEMY: https://www.khanacademy.org/math/probability/data-distributions-a1/summarizing-spread-distributions/a/calculating-standard-deviation-step-by-step

Miracle, R. (2015, January 6). *Tutorial: Calculating the distance between Map points*. Retrieved from Corona Labs: https://coronalabs.com/blog/2015/01/06/tutorial-calculating-the-distance-between-map-points/

Scripts, M. T. (n.d.). *Calculate distance, bearing and more between Latitude/Longitude points*. Retrieved from Movable Type Scripts: http://www.movable-type.co.uk/scripts/latlong.html