CS2134 Homework Assignment 4
Fall 2016
Due* 11:00 p.m. Tues. Oct 11, 2016

October 6, 2016

Assignment 4*B* include a programming portion and a written portion. The programming portion must compile and consist of a single file (hw04B.cpp), and the type written portion should consist of a single file (hw04Bwritten) in a *.pdf* format. Be sure to include your name at the beginning of each file! You must hand in both files via NYU Classes.

Programming Part:

1. **Enter data from the file** `MTA_train_stop_data.txt`**.** The data from this assignment is from http://www.mta.info/developers/download.html. (Please note that we will only be using some of the information in this file for this assignment.[1])
   In the batch phase you will read all the data from the file called `MTA_train_stop_data.txt` into a container of type `vector<trainStopData>`.
   Your program will define the class `trainStopData`. It has the following private member variables :

```
string stop_id;    // id of train stop (1st token)
string stop_name;  // name of station (4th token)
double stop_lat;   // latitude of train stop location
double stop_lon;   // longitude of train stop location
```

   Your class should also have a constructor and the following public member functions:

```
string get_id( ) const
string get_stop_name( ) const
double get_latitude( ) const
double get_longitude( ) const
```

*A bonus of %10 percent will be given if you turn in this homework assignment by Mon. Oct 10 at 11:00 p.m.

[1]The data is in a common format; please read https://developers.google.com/transit/gtfs/reference?csw=1 for more information.

2. In this question you will write the code to determine if a `vector` contains duplicate items. Your function will return `true` if there are duplicates and `false` otherwise. You will solve this problem in four ways (i.e. you will write four functions).

   (a) Use the `list` class to help you find the duplicates. (Use a "brute force" technique by comparing all items with each other.)

   (b) Use the STL `sort` algorithm, and the `vector` class to help you find the duplicates

   (c) Use the `set` class to help you find duplicates

   (d) Use the `unordered_set` class to help you find duplicates

Written Part

1. In programming question 2, what are the average case and the worst case running times for each of the different ways you used to find duplicates? Write your answers using big-Oh notation.