The City College of New York

**Department of
Computer Science**

CSC 59927-B: Big Data Management & Analysis
SPRING 2017
Homework 4 – Setup Project Workspace
Due: **April 24nd 2017**

This is a group homework. The main objective is for you to setup a shared HDFS space on the CUSP cluster for your final project. You are asked to create an HDFS folder (e.g. under a user home folder) such that:
- All of input data required for your project are stored in this folder
- All of your group members can read and write any files under this folder
- Your instructor (username: **hvo**) also has read permissions for all files under this folder

Your submission: the full path to your shared HDFS space on the CUSP Cluster.

Useful resources and Example:
https://hadoop.apache.org/docs/r2.6.1/hadoop-project-dist/hadoop-common/FileSystemShell.html#setfacl
https://hortonworks.com/blog/hdfs-acls-fine-grained-permissions-hdfs-files-hadoop/

For example, if I'd like to share with user hk1953 a folder named "project" under my home folder on HDFS that contains our usual citibike.csv file and the subway folder in the previous lab, I can do the following:

```
hadoop fs –mkdir /user/hvo/project
hadoop fs –put citibike.csv subway /user/hvo/project/.
hadoop fs -setfacl –R –m default:user:hk1953:r–x /user/hvo/project
```

With the above command Himanshu would be able to read all files under my project folder. The Access Control specifier "default:user:hk1953:r-x" means that allowing the "user" "hk1953" the permission to "read" and "execute" (or list) all files and subfolders under this project folder. However, Himanshu cannot write new contents to this folder. Also, the "default:" prefix indicates that any additional contents written to this folder after this command will have the same permission as above. (Note: the "-R" switch above is for marking the permission recursively, aka. including all sub folders).

Now if I'd like to create a "scratch" space so that Himanshu can share data with me, I can do the following:

```
hadoop fs –mkdir /user/hvo/scratch
hadoop fs –setfacl –R –m default:user:hk1953:rwx /user/hvo/scratch
```