**Conversion Script Usage Instructions**

These scripts can be used to aggregate dissemination area (DA) level count data to the community cluster level. Each community cluster comprises a number of DAs. The scripts reference a conversion file specifying which DAs belong to which community clusters. For a DA split between two or more community clusters, the script weights its contribution to each cluster based on the number of buildings in that DA which lie in respective cluster.

For instance, DA 12090793 is split between the Fall River and Cole Harbour / Eastern Passage community clusters. Of the 506 buildings in DA 12090793, 183 lie in the Cole Harbour / Eastern Passage cluster, and 323 lie in the Fall River cluster. Therefore, DA 12090793 contributes 183/506 = 36.2% of its data value to the Cole Harbour / Eastern Passage cluster, and 323/506 = 63.8% to the Fall River Cluster.

***Instructions for R:***

* Download the DA-CC Conversion Environment file (DA-CC Conversion Environment.RData) to your computer and open it in R
* In the R console, write convert(“*path to DA-level data”*)
	+ Example: convert(“C:/work/DA Level.csv”)
* Unless otherwise specified, the output community cluster data will be written to “Aggregated Community Count Data.csv” in the same directory as the DA data file. However, users can specify a different filename or a different directory:
	+ convert(“C:/work/DA Level.csv”, “CC-Level.csv”) will save the output to C:/work/CC-Level.csv
	+ convert(“C:/work/Living Alone – DA Level.csv”, “C:/data/CC-Level.csv) will save the output to C:/data/CC-Level.csv

***Instructions for SAS***

* Download both DA-CC-Conversion-File.csv and DA-CC-Conversion-Script.sas
* Open DA-CC-Conversion-Script.sas and on lines 5-8 enter the path to DA-CC-Conversion-File.csv, your data file, and where you would like the output to be placed, for example:
	+ %let conversion\_file\_path="C:/work/DA-CC-Conversion-File.csv";
	+ %let da\_data\_file\_path="C:/work/DA Level.csv";
	+ %let output\_file\_path = "C:/data/CC Level.csv";

***File Format***

* The original, DA level data must be stored in .csv format
* Numerical DA identifiers must be in the first column
* Any number of data columns can be included
* DAs which are not matched in the conversion file will be aggregated under “Unknown” community cluster
* The table below provides an example of the necessary formatting

Format of original data csv file:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Dissemination Area | *Data column 1* | *Data Column 2* |  . . .  |  *Data Column N* |
| 12010030 |  |  |  |  |
| 12010031 |  |  |  |  |
| … |  |  |  |  |