

Adjacency Matrices and Lists

These instructions are pertinent to the **programming project** of [IS421: Enterprise Analytics for Decision Support](#), Semester 2, 2016.

This folder contains the adjacency matrices for the **Singapore road network**. Specifically, the following two CSV files that represent the adjacency matrices are contained:

- **edge_idx_matrix** -- adjacency of the nodes are represented by the **non-negative** edge indices (i.e., zero or positive integers) between all pairs of adjacent nodes.
- **travel_time_matrix** -- adjacent nodes (i.e., the edges) are represented by the **non-zero** travel time (in minutes) between them.

The matrices are sparse and of dimension 9,948 x 9,948 each. The first matrix represents the node adjacencies (i.e., the edges) via the edge indices, which are non-negative integers used to name the edges. The second matrix represents those edges via the (maximal) travel times (in minutes) through the road segments. Therefore, a negative or zero entry in each matrix means that the respective row and column nodes are *not* adjacent. In addition, all the matrices are symmetrical and have negative/zero diagonals. This means the road network is an **undirected graph**. The total number of edges is 11,153.

Each CSV file is a square matrix **with headers**. The rows and columns are exactly identical. The headers name the columns from 0 to 9947 (i.e., the number of columns - 1). This can be conveniently used to name or index the nodes. Otherwise, for most practical purposes, each header line (i.e., the first line of each CSV file) should be ignored when reading the file. Note that each of these CSV file, once expanded, can be **quite large** (~ 400 MB). It is thus essential that you write a program to read these files and construct an undirected graph out of them to represent the road network.

In addition to the adjacency matrices, the following two CSV files are included that represent the adjacency of the edges more compactly (and much easier to read) -- these are similar to (but not exactly) the adjacency lists:

- **edge_idx_list** -- for the edge index.
- **travel_time_list** -- for the travel time.

In these two files, only adjacent nodes are present (in the first two columns), where each pair of adjacent nodes forms an edge (road segment) of the graph. The third column represents the edge index and the travel time, respectively. You should, again, ignore the header (the first line) of each CSV file when reading in. You are free to choose whichever format you prefer for your own implementation.