Case Studies in Scientific Computing [MA4306]

Project 3 – The Bikesharing Rebalancing Problem

Input data: NYC Citibike

For this dataset, I reconstructed the actual vehicle-based rebalancing that happened in September 2020 (shortly after that, the operator changed the format of the open source data and no longer included bike IDs. September is by far the busiest month, so I chose to use September). In order to do so, I checked which bikes did not have coherent End/Start stations, which means they must have been moved by the operator. I calculate 39560 rebalancing rides, which is realistic. The monthly operating report states there have been 51179 rebalancing rides, but this number includes user-based rebalancing. I chose the 11th as our concrete instance, because this day yields the most rebalancing actions (3180 to/from 615 stations) and can therefore be seen as a *worst case*.

The file '*nyc_instance.csv'* contains data of 548 stations and the depot. The following information is given:

- Id: the unique ID of the station/depot
- Delta: the amount of deficit/surplus at each station. If the number is negative (positive) we need to bring (take) bikes.
- y/x_coord: coordinates of the station/depot
- Note: Over the course of the month the sum of surplus and deficit is zero, but on single days it isn't. on the 11th it is -112, meaning we need to bring 112 bikes when leaving the depot (this happens because bicycles need repairs). That is why I prepared two instances, one with a vertice called 'dummy' holding a surplus of 112 bikes and one without that. Same for the adjacency matrices.

The adjacency matrix *'adjacency_matrix_nyc.csv'* is also present in the moodle folder.