London Underground –
A Simulation Study: Step 1

Starting point is a data set obtained by freedom of information request by Rayleigh Collins, from Transport for London (August 2008). This dataset describes the topography of the underground network, the lines, the distance between stations and the average run time of trains between stations.

Each line has its particular rolling stock which for the purpose of this analysis can be described by two parameters:

- The maximal acceleration of the train (this is related to the weight of the train, the rating of the electrical engines and the power supply to the network),
- The maximal speed of the train (this is related to the design of the train and the layout of the tracks)

The run times provided are the sum of the drive time and dwell time for empty train runs, peak-time train runs and off-peak train runs in normal operation. The dwell time describes the average waiting time due to signalling constraints and the time required for passengers getting on and off the train.

The first step is to split-up the run times into drive time and dwell times and to determine distributions for the parameters. The drive model assumes that a train sets out with maximal acceleration until it reaches maximal velocity, then cruises at maximal velocity and finally decelerates with the same maximal acceleration to come to a stop in the next station. When the distance between two stations is not long enough, we assume that the train accelerates to mid-way point (with maximum positive acceleration) and decelerates thereafter (with the same negative acceleration). The minimum drive time ($\min_t$) between two stations can then be computed using the function `timeTo(a, v, d)`. The parameters $\max_{acc}$ and $\max_v$ are estimated for each line and stored in a dataframe with the line as index. For each line segment the running time is split into drive time and dwell time.

In step 1 of the simulation model, we read the dataset and estimate the parameters and per line and compute the minimal drive time and dwell times for each line segment.

We define a class Train that encapsulates the process of a train running along a track, i.e. one direction of a line. A process is implemented in SimPy as a Python generator method. We will follow the convention of calling this method **`process`**.