**A picture containing text

Description automatically generated**

**A01 – REPORT.**

# Description and Relevance

Bus routes are planned to take into consideration traffic patterns, patterns of movement etc. Over a period, these assumptions change. Traffic in one part of the city may increase while traffic in another part of the city may decrease. Also, the distribution of traffic throughout the day may change. All this implies that bus schedules need constant updating to reflect the current conditions.

One way to approach this problem is to consider every one-hour window and percentage the ratio of times buses has been late at every stop.

The routes and one-hour windows where we have the highest percentage of buses reporting late will be the prime candidates for replanning.

This dataset allows us to compute this information.

# Novelty

This problem is different from the four other problems in the assignment.

In the first problem, we take spatial rectangle based on the latitude and longitude. Then for every hour, within this rectangle, we calculate how often congestion is reported.

In the second problem, we want to calculate the timetable of a physical vehicle, as a pair of lineID and stationID, for a given day.

In the third problem, we compute the station that has the highest number of buses stopping at it so that we can use that station for best reach of advertisements.

In the fourth problem, we compute the distance travelled by individual vehicles so that we can find out when to send them for service.

This new problem is different from all others because we compute how many times a route has reported delays at stations, and then aggregate them by route and hour. This is not done for any of the given problems.