MSc Data Analytics Dissertation: Visualisation of Traffic Flow Data for Glasgow City Council

Project Context/Scene Setting Document

**Elijah Reid**

**202093659**

**2021**

# Organisational Context

The Glasgow City Council is responsible for running the majority of public services in and around the Glasgow area, and in order to better apply itself to so many different simultaneous projects it is split up into many different departments and teams, two of which are the instigators of this project.

The Strategic Innovation and Technology (SIT) Data team are responsible for finding better ways to use data, technology, and user engagement to create new insights that can help to inform and shape new developments across both the Council and the City of Glasgow.

Traffcom are responsible for the controlling and management of traffic in the Glasgow City area, with their responsibilities including the maintenance and operation of over 1000 sets of traffic lights across the city, which make use of various adaptive control techniques. The signals generate and collect a large amount of real time data, some of which is made available to other users and departments in the council.

The SIT Data Team and Traffcom have begun to recently collaborate on several projects to consider how the Council might open up new city data to the public, how they can analyse it to validate its accuracy and utility for further use, as well as how they might be able to visualise the data to generate new insights and further understanding around what the data tells them about Glasgow.

# Background to Project

Traffcom wants to explore whether the traffic flow data generated from the pre-existing Urban Traffic Control, or SCOOT system, for Glasgow can be augmented with directional information, and following this whether it can be visualised for the city traffic network.

A proof-of-concept project featuring a small subset of the data was released to the MSc in Data Analytics programme, of which I am a part of, at the University of Strathclyde in October 2020 in order to judge the feasibility of the project by bringing multiple outside perspectives in to analyse the data. After receiving several promising ideas back, the Traffcom and SIT data teams proceeded to set up a more ambitious project to be undertaken over the summer by a student from one of these groups.

The project will make use of data already collected from the Glasgow SCOOT system, which collects data from road-side sensors and adjusts traffic signals in response to real-time traffic flows. There are currently over 1500 existing links in the network, with traffic flow data being sent to a central database every 15 minutes, with data points going back as far as May 2018. Currently live and historical traffic flow data can be requested via APIs; however, they require specific queries to provide relevant data as well as missing a directional component, limiting their use for fully understanding the data and creating visualisations.

# Project Value and Expected Outcomes

The project involves coming up with a method to establish traffic direction for each of the nodes in the network of traffic sensors across the city and then use this data to create visualisation, ideally using software tools that are already in use or easily available to the Glasgow City Council.

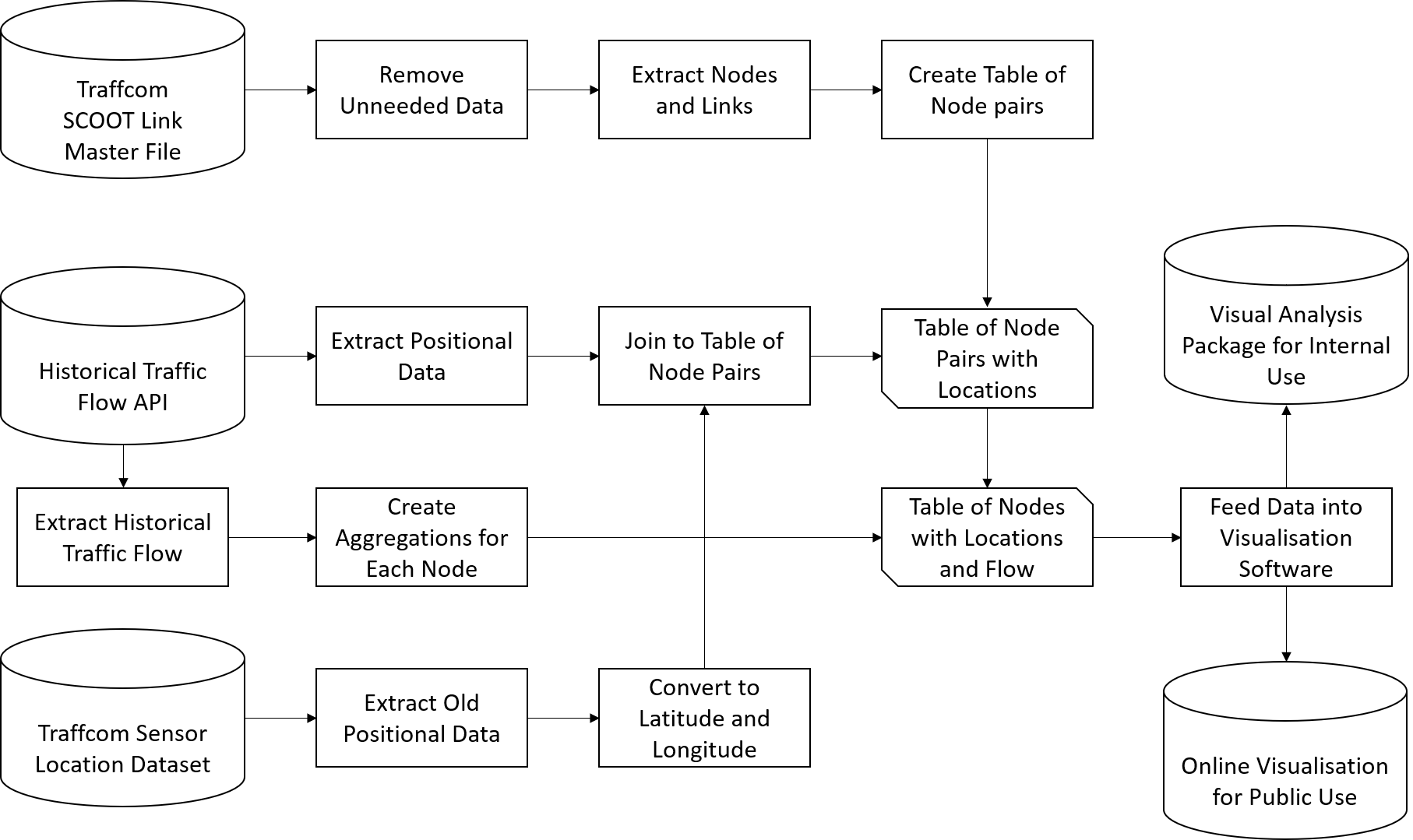
The visualisation will be based off the 3+ years of historic data collected by the Glasgow City Council, as this will allow for the end users to analyse and predict traffic trends based off this. It is expected to be at least somewhat configurable to allow different groups to view their areas of interests as well as allowing different time periods and different routes/areas to be selectable for more in-depth analysis.

The plan for the project is to set up an ETL pipeline which will extract each SCOOT link from the master text file maintained by Traffcom, and then transform this extracted data into a table of pairs of linked nodes. As new links are very rarely added and only manually, the script to update the links will only run on an irregular basis and might even be only activated manually by Traffcom management.

Data will also be pulled in from the Historical Traffic Flow API to provide the location co-ordinates for each link as well as the traffic flow there over the past few years. This will be scheduled, either with a simple Cron script or a more advanced solution like Apache Airflow to update the traffic flow every 15 minutes. An additional script for undocumented or links with old positional data will calculate the bearings for entry links, as well as converting the old British National Grid co-ordinates into Latitude and Longitude to ensure all positional data is in the same format.

Once all the links and related data are processed, it will then be fed into a Geographic Information System (GIS) Mapping software (either Arc GIS which is already in use at the council, or another free software such as Leaflet, Plotly or Gephi). The map will highlight all the links, with an indication of the volume and direction of traffic flow, with additional information being potentially added if time allows. This map will be included in a visual analytics software for internal use (either a more in-depth ArcGIS setup or a simpler product such as Microsoft Power BI or Tableau) as well as being hosted on the web for public access.

# Initial Data Pipeline Mock-up



# Client Use Case

The visualisation produced as an end-product of the project will be of use to a range of stakeholders, including traffic engineers and managers who can use it to monitor live traffic and predict future trends based on the historical data.

City stakeholders will be able to use it to anticipate the building or removal of connecting roads as well as the effect on traffic from closing roads for works at different times of the day and year.

Finally, residents of Glasgow will also be able to make use of the data and visualisation as it will be hosted as an app as part of the Glasgow Open Data initiative, allowing them to keep track of traffic themselves whilst also allowing local entrepreneurs to make use of the data for their own future projects.

# Project Plan

## Initial Plan

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Week 1 | Week 2 | Week 3 | Week 4 | Week 5 | Week 6 | Week 7 | Week 8 | Week 9 | Week 10 | Week 11 | Week 12 |
| 31-May | 07-Jun | 14-Jun | 21-Jun | 28-Jun | 05-Jul | 12-Jul | 19-Jul | 26-Jul | 02-Aug | 09-Aug | 16-Aug |
| Intro to project, datasets and currently used tools  Meeting with traffic, data and visualisation departments  Extract relevant link data from Network file  Ethics Form completed | Gained access to ArcGIS and FME  Complete Overview Document  Begin research into traffic visualisation techniques  Extract all co-ordinate data for each link from API  Team Meeting Wednesday | Get each SCOOT link plotted on a map (without directions)  Create a table containing every junction connection pair with latitude and longitude for each  Continue visualisation research  Team Meeting Wednesday | Write a function to calculate the bearings for each of these pairs and add to table  Begin Literature Review and Methodology section  Further visualisation research if necessary  Team Meeting Wednesday | Begin trying to visualise connections between the junctions as lines  Team Meeting Wednesday  Check in with SLP (Wed) and Pierro and Liz (Fri) | Continue visualisation process as per week 5  Redraft Literature Review and Methodology  Team Meeting Wednesday | If visualisation has worked attempt to automate process from weeks 1 to 7  Gain access to data for traffic flow and other information to be overlayed  Team Meeting Wednesday | Add the flow information to the visualisation, colour coding by set criteria (to be agreed upon)  Begin writing client report for GCC internal use  Discuss system for internal use- using Python and R, creating a dashboard (PowerBI etc) or online  Team Meeting Wednesday | Begin to implement the internal system  Begin research in ways to host visualisation online for public access  Create aggregates of a few different time periods to speed up user requests  Team Meeting Wednesday  Check in with SLP (Wed) and Pierro and Liz (Fri) | Continue working on the internal system  Begin creation of web version of the visualisation  Potentially start adding other data sources if time allows  Team Meeting Wednesday | Continue working on the public version  Review and redraft client report and technical manual  Team Meeting Wednesday | Handover of project and mop-up any remaining issues  Write conclusions and recommendations for dissertation and spend next few weeks reviewing  Team Meeting Wednesday |

## Updated Plan

# References

* Glasgow City Council (2018) *Digital Glasgow Strategy*, Glasgow: Scottish Government.
* Director of Governance and Solicitor to the Council (March 2018) *CORPORATE ASSET MANAGEMENT PLAN - UPDATE*, Glasgow: Glasgow City Council.