Polytechnique Montréal

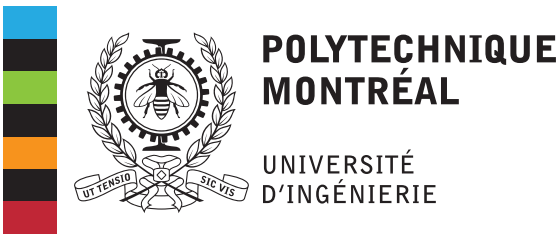
Lab Work 3

Spatial and statistic analyses

CIV8760E : Transport Data Management

Teaching Assistant :
Guillaume Neven

Fall 2024



1 Introduction

In this practical work, you will study traffic dynamics on Montreal's road network through a spatial and statistical analysis of travel times on certain road segments and their travel conditions using the travel time data and road segments information from the City of Montreal. Warning, the road segment LAT and LONG are using , rather than .. You need to change that before using Spatialite.

2 Spatial Analysis [7.5]

Before performing statistical analyses, the city asks you to match the various segments where travel times are collected with the city's boroughs (shapefile or GeoJSON file). To do so, find out which borough each link is located in. Create a map in QGIS, showing the different boroughs with their own unique color. Color the links in each region with the same color. Develop a strategy for links that overlap two boroughs and provide commentary.

You are free to use the program of your choice, except for the final map, which must be created in QGIS. However, you are encouraged to use Spatialite.

3 Statistical Analysis [10]

For the two subsections that follow, you must select ONE road segment to study that travels in both directions (e.g., from A to B AND from B to A, such as "Sherbrooke N06: Curatteau to SaintDonat" and "Sherbrooke S07: SaintDonat to Curatteau") and is longer than 1000 m. Clearly indicate the selected segment. It is important to select a segment that contains data for each of the years (2016 to 2019).

3.1 Travel Times and Speeds [5]

For the selected segment, describe the travel times and speeds by year. The data may vary significantly from year to year and may contain outliers. Therefore, it is important to implement some filtering to mitigate their effects and consider them in the analysis. Describe your filtering method and your assumptions.

Using the filtered speed data, perform a goodness-of-fit test to the theoretical probability distribution that you deem most relevant. Explain your choice, your assumptions, and the results. Carry out the analyses for both directions separately.

3.2 Time-Based Analysis [5]

By selecting one travel direction, describe the temporal distribution of the number of observations and travel times, according to months, days of the week, and hours of the

day. Justify the statistic used to represent travel times. Create graphs showing the following information per year (2016 to 2019) and provide commentary.

Finally, you should study the difference in travel times between different years: include a figure showing the distributions and use appropriate tests to determine if there has been a significant reduction in speed from one year to another.

4 Statistical Models [10]

For this final part, you must select one year (2016 or 2017) and five road segments in different neighborhoods of Montreal. You must then create a variable describing the travel conditions at each hour of each day. For example, the ratio of the average speed on the segment to the speed limit for each segment (speed limits are available on OpenStreetMap). You could also consider another way to derive the density (number of vehicles per unit of length) or congestion on the segments. This variable will be the dependent variable y, the variable to be explained by the model. Finally, you must enrich the hourly data with weather conditions (Historical Data - Government of Canada) and relevant characteristics of the road segments, such as the number of lanes, orientation (cardinal points), number of intersections, distance to downtown, etc. Clearly explain the steps to add these new characteristics to the selected segments for analysis. In addition to the variables already available in the data (travel time and road segments), you must add at least one variable related to temperature, one related to travel, and one related to the road segment. Comment on your choice of variables (in particular, select variables that are not highly correlated with each other).

Study the temporal, meteorological, and road segment characteristics associated with travel conditions on the chosen segments using a linear regression model. Describe your model development and verification process. Finally, comment on your final model and carry out a (graphical) residual analysis.

5 Submission of the Report

This practical work is to be done individually. Please do not exceed a report length of 10 pages and keep in mind that this is an exercise, so be concise.

Submit the report in PDF format by November 21st at 11:59 PM on Moodle. Make sure the report is free of grammatical errors and includes clear graphs and explanations. Points will be deducted for writing errors and inaccuracies in data analysis. If a language model (e.g., ChatGPT) is used, you must disclose its precise use.