Using spatial and temporal predictors to classify why people travel across a city: A study analysing the spatio-temporal characteristics of trip purpose within Montréal, Canada between September – October 2017.

*This dissertation is submitted in part requirement for the MSc in the Centre for Advanced Spatial Analysis, Bartlett Faculty of the Build Environment, UCL.*

*A colour version of this dissertation is available online.*

Word Count:11,986

Thomas J. Keel – 18110348

TMSSDSAVIS01: MSc Spatial Data Science & Visualisation University College LondonSupervisor: Huanfa Chen

*30th August 2019*

# Abstract

The quantification of trip purpose when people move across cities, remains an area within broader mobility studies without an extensive investigation. In the past, this investigation has been hindered by two factors: (1) the absence of relevant data which details the purpose of people’s travel in a city and (2) the difficulty in representing the space-time dynamics of mobility within models/metrics. Regarding (1), in recent years, smartphones travel surveys have provided researchers a platform to study travel within a city on a large scale. In turn, this has fuelled an eruption of volunteered geographic information (VGI), which provides qualitative information about travel at increasingly finer temporal and spatial scales in cities. This study makes uses of one such study: the *2017 MTL Trajet* project – a survey examining travel behaviour and intent patterns across Montreal between 18th September 2017 – 18th October 2017. Results from this survey provide insight into how and why people travel and could be used to inform the better planning of essential and non-essential services at the city-level. This project builds on a small body of research into methods that uncover spatial and temporal interdependencies of data before assessing the ability of three distinct classification models (Support Vector Machines, Random Forest and Artificial Neural Networks) used to predict the trip purpose.

**Key Words:** Trip Purpose, Mobility, Volunteered Geographic Information, Spatio-Temporal Investigation, Classification Modelling.

# Declaration

I, Thomas Keel, hereby declare that this dissertation is all my own original work and that all sources have been acknowledged. It is 12,000 words in length

Signed:

Date: 28th August 2019

Acknowledgements

Introduction

List of Figures and Tables

List of Acronyms and Tables

Chapter 1. Introduction

Chapter 2. Literature Review

Chapter 3. Methodology

Chapter 4. Results

Chapter 5. Discussion

Chapter 6. Conclusion

Bibliography

Appendix 1. Notification of intent not to apply for ethical approval

Appendix 2. Link to scripts used for the analysis

Table of Contents

[Abstract 1](#_Toc17543675)

[Declaration 2](#_Toc17543676)

[List of Figures 4](#_Toc17543677)

[List of Tables 4](#_Toc17543678)

[List of Acronyms and Abbreviations 5](#_Toc17543679)

[Acknowledgments 5](#_Toc17543680)

[Chapter 1. Introduction 6](#_Toc17543681)

[1.1. Research Overview and Questions 6](#_Toc17543682)

[1.2 Motivation 7](#_Toc17543683)

[1.3 Approach 7](#_Toc17543684)

[1.4 Outline 9](#_Toc17543685)

# List of Figures

**Figure 1.1** Screenshot from the MTL Trajet App (left; source: Patterson, 2017) and example of prompt within Itinerum apps (right; source: Patterson *et al.*, 2019)………………………..……….1

**Figure 2.1** Framework of the Itinerum Mobile and Desktop app (Source: Patterson & Fitzsimmons, 2017)…………………………………………………………………………………………………...…………2

**Figure 2.2** Example of GPS trace with location collection priorities within an Itinerum platform app (A); Example of the on screen prompt after an Itinerum platform app stops recording movement (B) (Source: Patterson *et al.*, 2019)……………………………………………………..4

**Figure 2.3** Example of *Map Matching* done by the Open Source Routing Machine when processing the raw GPS trace from user devices (Source: Hamouni, 2018)…………………………..4

**Figure 3.1** Montreal within Quebec, Canada………………………...………………………..……………………5

**Figure** **3.2** GPS routes from the MTL Trajet plotted within the study area…………...………...……5

**Figure 3.3** Location of Montreal within the study area. ……………………………………………………….5

[**Figure 3.4** Bar chart showing the count of each category of land use]

**Figure 3.5** Map showing land use changes within the City of Montreal………..…………...…………5

**Figure 3.6.** Circular contour plot (windrose; left) and circular histogram (right) showing the direction of trips (circle bands indicate count of trips)

**Figure 3.7** Example of the spatial join between a route and the underlying dissemination areas (route in **blue**; overlapping dissemination areas in **red**)

**Figure 4.1** Line plot showing the amount of recorded trips taken from the MTL Trajet app between 18th September 2019– 18th October 2019 (weekends indicated in **purple** ; data from PDO, 2017).

**Figure 4.2** Boxplots (top), Kernel Density Estimation (middle) and Quantile-Quantile (bottom) plots showing the distribution of trip distance and duration.

**Figure 4.3** 1-Dimensional Kernel Density Estimation plot of trip distances.

**Figure 4.4** 1-Dimensional Kernel Density Estimation plot of trip durations.

**Figure 4.5** Average trip distance and duration as proportion of the mean.

# List of Tables

**Table 3.1** Description of the key variables from data from the MTL Trajet survey before pre-processing.

**Table 3.2** Categories of mode and purpose allowed for trips in the MTL Trajet survey.

**Table 3.3** Description and cover of Land Use categories within the City of Montreal.

**Table 3.4** Definition of Rush hourand Off-peak hours used in this study.

**Table 3.5** Description of the key variables from the MTL Trajet survey after pre-processing.

**Table 4.1** Number of trips removed from the analysis

**Table 4.2** Summary statistics for Distance and Duration of trips (converted to km and minutes; N=177,938)

# List of Acronyms and Abbreviations

**ANN** – Artificial Neural Networks

**CNN** – Convolutional Neural Networks

**DA** – Dissemination Areas

**GPS** – Global Positioning System

**LDA** – Latent Dirichlet Allocation

**MAUP** – Modifiable Areal Unit Problem

**MLP** – Multi-Layer Perceptron

**MTUP** – Modifiable Temporal Unit Problem

**RF** – Random Forest

**SVM** – Support Vector Machine

**VGI** – Volunteered Geographic Information

# Acknowledgments

I would like to thank my family for their thorough support

Secondly, I would like to thank my supervisor Huanfa for their comments and advice and the staff at CASA for the support this year.

Finally, I am extremely grateful for those people I have met throughout this year at CASA, who have

Especially my flat mates James and George.

All dissertations are submitted electronically, but we also require ***two printed and bound copies to be submitted as well***. Double spacing must be used, except for indented quotations, tables, bibliographies and footnotes, which should be single-spaced. The left- side margin should be not less than 40mm (1.5 inches) and other margins not less than 20mm (0.75 inch).

# Chapter 1. Introduction

## 1.1 Research Overview and Questions

The purposes by which people make trips along transport networks on a large scale still remains an area with a distinct lack of investigation within broader literature (Yazdizadeh *et al.,* 2019). This somewhat owes itself to the absence of data available to study trip purpose.

In recent years, smartphones have provided researchers an opportunity to study and record large scale geospatial movement of people as, smartphones generally exhibit the same mobility behaviour as their carriers (Jahromi *et al.*, 2016). Further, mobile apps provide researchers a platform to ask users questions about their movement so we can infer transport behaviour patterns (Li et al., 2016). Through mobile app surveys, we can thus gain insight into population movement behaviour throughout a city and throughout time. Indeed, we can infer preferences into travel-mode choice they take when travelling for certain activities.

Arguably, by studying these behaviours on such a scale, they can use them to inform city-level decision-making to help plan a variety essential and non-essential services (Attard *et al.*, 2016). For example, if we knew that people have a higher propensity to use public transport when travelling to leisure activities in certain parts of a city, transport authorities could use this information couple bus/metro routes/schedules to locations and opening hours of leisure services [may need better example].

Despite this, there is still a gap in knowledge of understanding of travel intent in most cities, owing to the fact that no research has been initiated there. One exception is in Montreal, where a number of mobile applications have been created in recent years (since 2016) to study how people move across the city based on their smartphones. This report makes use of the most recent available data from one of these studies: The *2017 MTL Trajet* study carried out between 18th September 2017 – 18th October 2017. We use data from this study to assess an over-arching research question along with two of derived questions:

**Main Research Question:**

Can we effectively classify trip purpose of the respondent to the MTL Trajet survey using spatial and temporal indicators within Montreal?

**Sub-Questions:**

1. Which spatial and temporal indicators are most important for the classification of trip purpose?
2. Which type of classification algorithm performs are most effective in the classification of trip purpose?

## 1.2 Motivation

Movement can be thought of as an interaction between an origin and destination, whereby people move across space and through time to go from the they are in to the they to where they want to be (Murray *et al.*, 2012). Transport, being a by-product of movement, is as such a derived demand for a particular destination (Golledge & Gärling, 2001). Studying the purpose of movement that individuals take between an origin and destination within a transport network underpins our comprehension of the human behaviour within a city (Kwan & Neutens, 2012). Indeed, if we can discern the activities for which people travel for and at which temporal and spatial scales, we can better manage and plan essential services within a city.

The improvement in smartphone and geolocation technology mean we currently have a unique opportunity to study

with smartphones enormous potential to collect location data for many purposes, hence ﻿Itinerum platform created (Patterson *et al.*, 2019)

Smartphones capable of environmental geospatial information at an unprecedented level (Li et al., 2016; etc. (i.e. Patterson)).

Big data measurement and influx has extended to the extent that it is real-time, which gives us a unique opportunity to study geographical phenomena in real time (Goodchild, 2013)

Big geographic data allows for us to not only study the spatial and temporal interactions but also interactions of socio-economic factors [this is what this research aims to do] (Cheng *et al.*, 2017).

The motivation of this study is thus to deconstruct the spatial, temporal and spatial-temporal profiles of trip purposes and try to model them.

*Why Montreal*

Smart cities [i.e.] the instrumentation of cities which are now providing vast amount of real-time data (from Li *et al.*, 2016). Increase in the amount of geo-tagged/geo-referenced data

ESSENTIAL: Smart City with interacting networks and GPS signals (Jahromi et al., 2016)

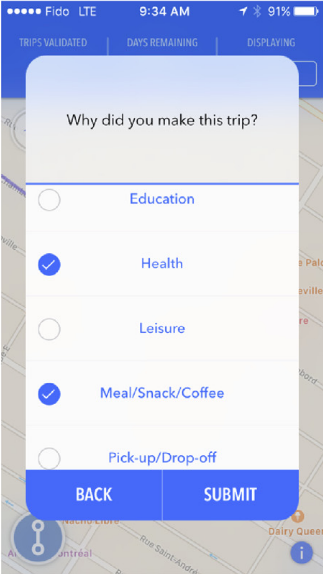
## 1.3 Approach

*MTL Trajet*

This study makes use of data from the *2017 MTL Trajet* survey originally collected by researchers at the Transportation Research for Integrated Planning (TRIP) lab, Concordia University (Patterson & Fitzsimmons, 2017). This survey was created to study travel behaviour and is currently being used to inform city and transport planning decisions in the city of Montreal (MTL Trajet, 2017).

The methodology of this project differs from previous research as it concerns itself with the modelling (through classification) of the movement of people in Montreal based on given purposes and mode. It is hoped that this modelling, in combination with an overarching spatial-temporal investigation, can be used infer something […] about movement within a city.

**Figure 1.1**

** **

**Figure 1.1** Screenshot from the MTL Trajet App (left; source: Patterson, 2017) and example of prompt within Itinerum apps (right; source: Patterson *et al.*, 2019).

Although, this study primarily focusses in on Montréal and this may not be transferred to other cities.

* Understanding mobility through mobile phone has kicked off (Zhao *et al.*, 2019)

Primarily the movement of people is of concern to time-space analysis.

The dataset forming the backbone of this report provides a look into the way that people move.

This study also concerns itself with the modelling through classification on purpose of the movement. It is hoped, in combination with a spatio-temporal investigation, this analysis presented can infer something about movement at a higher scale within a city. Although, it must not be forgotten that this study primarily focusses in on Montréal and this may not be transferred to other cities (Ergodoic and Ecological Fallacy).

This study attempts to break away from its data-driven approach to provide more context

* \*Trip sentiment

## 1.4 Outline

This following examine the related literature and reviews the philosophy of the metrics introduced in 2. After this, in 3 detail is provided on the data . Results in 4 first create a ESTDA before modelling.