# Notes:

* 2000 words
* A form of spatial analysis which needs a lot more attention, and more surveys to be carried out.

# Limitations:

## Data Limitations:

* 1 month
* Only ~3,000 respondents
* Bias (types of bias and uncertainty)

## Methodological Limitations and Uncertainty:

* Direction is 16
* MAUP -> consisting of the scale and zonation problem (﻿The scale problem is defined as the variation of results when the same areal data are aggregated into larger areal units for analysis.) (From Chen, 2019) (The zonation or aggregation problem Is when the analysis varies based on how the area is divided up even if using the same units)
* Problem with assigning rush hour and not (i.e. what about overlap) -> same problem with directional LISA as one trip may pass into rush hour
* Model is built on Montreal, may have a completely different

Uncertainty -> see Bruce’s lectures and notes

﻿Furthermore, in order to be able to predict transport demand or traffic, not only are real-time data required but also historic data. (Li *et al.*, 2016 -> find another ref, but basically trying to say that historical is needed as well).

## Key Points:

* Visualisation and Interactivity:
  + Can we make big geospatial data analysis and visualisation available to an end-user through interactivity? (maybe Li *et al.*, 2016) -> currently not
  + Videos of change over time may be needed for space-time investigation
* Fractal emergence of travel
  + (fractal) emergence in patterns of travel (Li *et al.*, 2016)
  + Self-similarity
  + See Batty refs

# Discussion

Clustering:

Requires post-hoc information, to be implemented we have to use a progressive clustering algorithm where the clustering is re-evaluated every so often (if we wanted it to be real time) Otherwise, can look at information backwards. I.e. KMeans can cluster new points in space, LDA new points in time, space-time can cluster emergence of time-space clusters.

Purpose classification:

\* Which clustering technique performs best

Model is built on Montreal, may have a completely different result for other place, although we can infer spatial-temporal trends from the results, this may ‘frozen’ in time and space

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). Indeed, it is erroneous to assume that what is examined in across this covered region Montreal at the time of the study period can at all be scaled up to Montreal at a different point in time (i.e. to Winter or 5 years in the future or past), let alone to another city. It is easier to assume instead that is useful information for studying a network of interconnected movement.

*On Neighbourhood effect ->* People often traverse neighbourhoods and boundaries throughout one day (Kwan, 2018)

On omitted-variable bias (OVB): “occurs when a statistical model leaves out one or more relevant variables” (i.e. purpose?)

Other models:

CANT DO LSTM or RNN as time is not regular

Can do CNN because of grid

CNN-LSTM using videos of each trip could prove important

Further Research:

[Better modelling] “there are uses of machine learning methods that have been extended to account for the limitations of working with spatio-temporal data (such as the integration of convolutional neural networks and LSTM methods (Shi *et al*., 2015; Yu *et al.*, 2017; Han *et al.*, 2019)“. Using a better model that accounts for space-time (CNN-LSTM) -> which you input a video of trips

# Conclusion:

Model is built on Montreal, may have a completely different result for other place, although we can infer spatial-temporal trends from the results, this may ‘frozen’ in time and space