

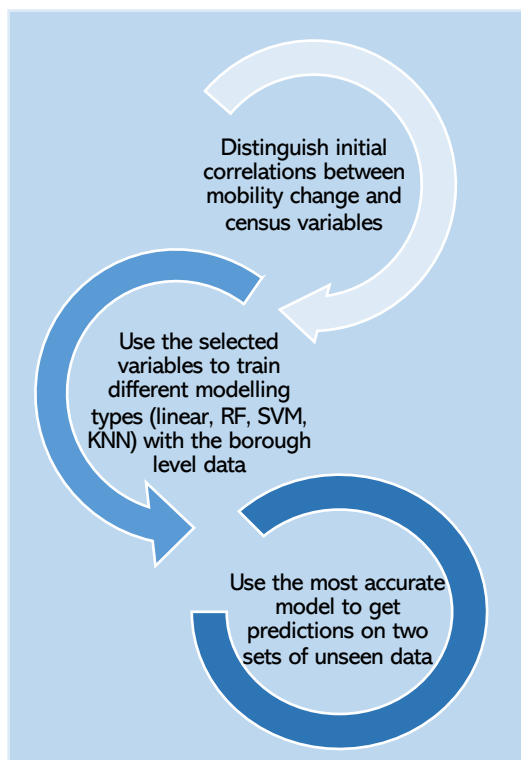
A machine learning approach to understanding factors influencing public transport usage during a pandemic

Ellie Marfleet

Research Summary

- The project aims to investigate factors affecting mobility during a pandemic, using the UK's lockdown period in November 2020 and London as a case study. To do this, the project investigates the relationship between census variables (such as occupation, industry, qualifications, ethnicity and car ownership) and mobility in London boroughs to identify what socio-demographic variables are predictors of increased mobility during a period of government-imposed lockdown. The project therefore aims to assess whether individuals in practical industries are more mobile and thus may have a higher exposure rate to the virus due to their inability to switch to remote working, and if mobility in London during lockdown can be successfully modelled using machine learning approaches.

Methods



Results

- The final model identified those in financial, real estate and professional industries to be the most important variable
- Car ownership and roles in practical industries (manufacturing, construction, skilled trades) were also highly important features
- The model was considerably better at predicting mobility in the remaining London boroughs but should not be used as a prototype model for any city due to its predictive performance when tested with unseen data referring to locations outside of London

Figure 1: Mobility change in London boroughs during November's lockdown

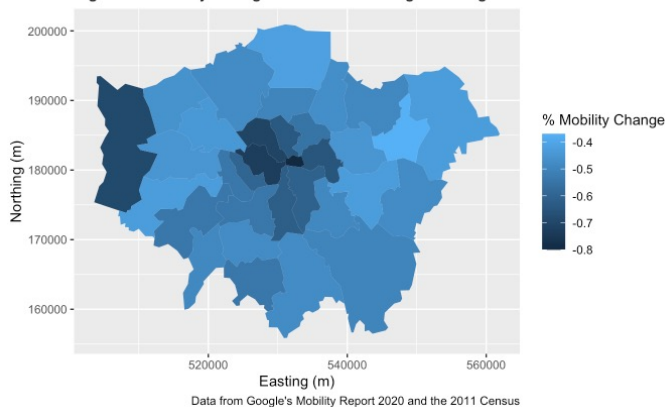
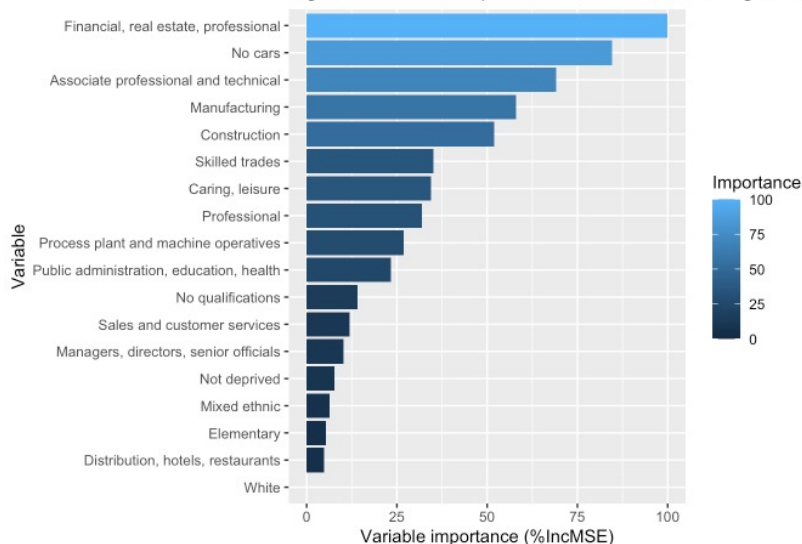


Figure 6: Ranked importance variables according to RF



Issues



Borough level analysis only – thus ecological fallacy issues and small sized dataset



Unclear what modes of transport the mobility indicator reflects as Google merely states "transit stations"



Mobility indicator derived from Google's location data - biased towards android users