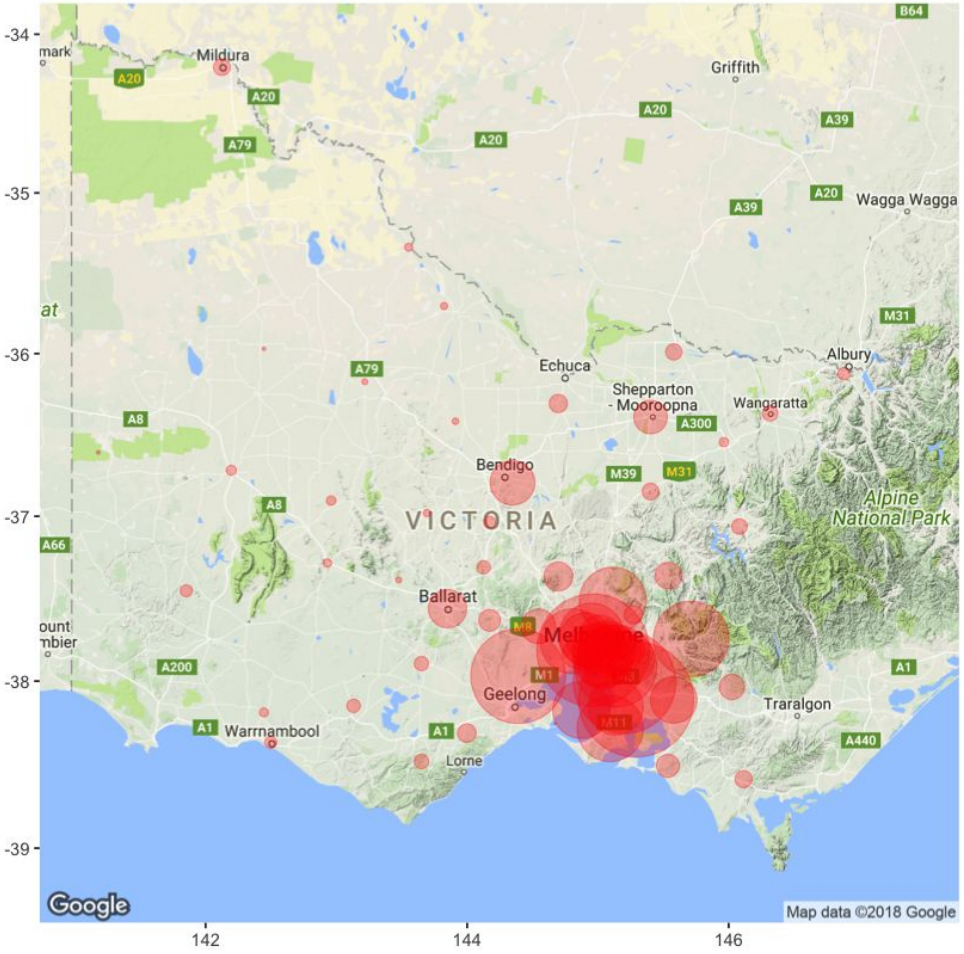
Solution 49^x Technical Interview

Dr. Jennifer Piscionere

What Data am I Using?

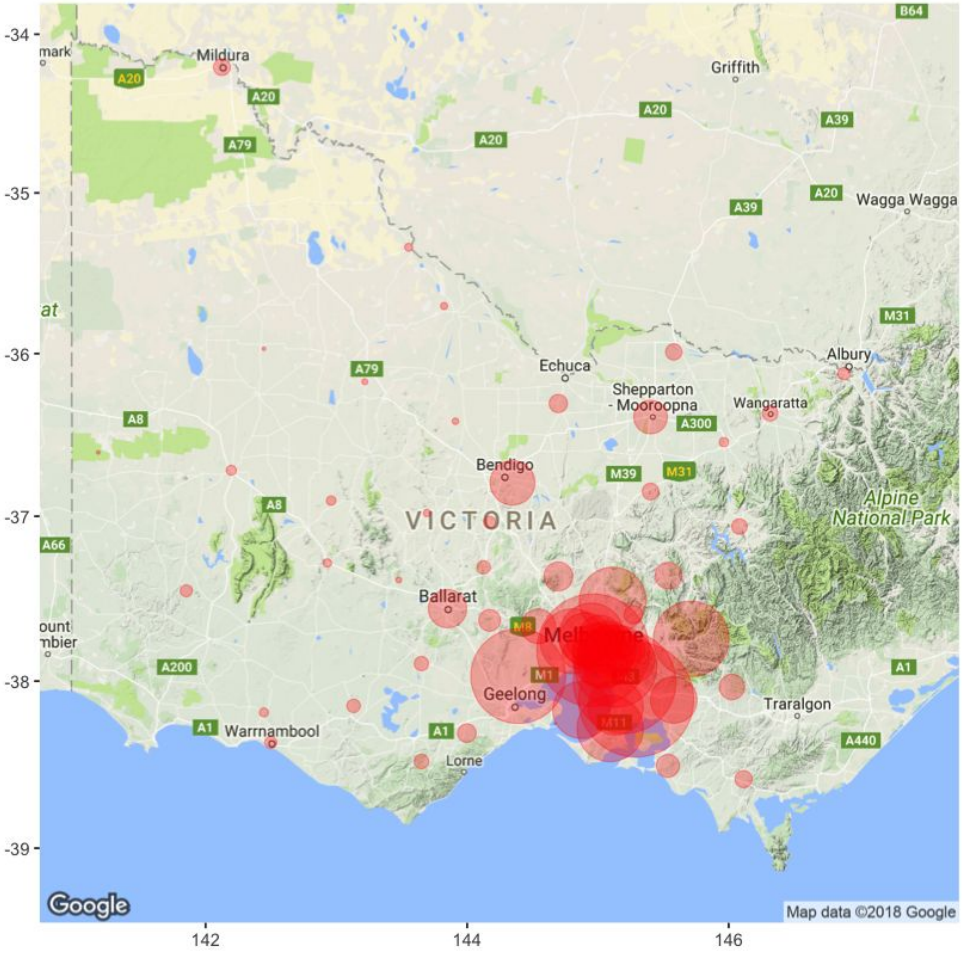
The number and type of all non-fatal* road accidents across Victoria, based on Local Government Areas.

Where are the accidents happening?

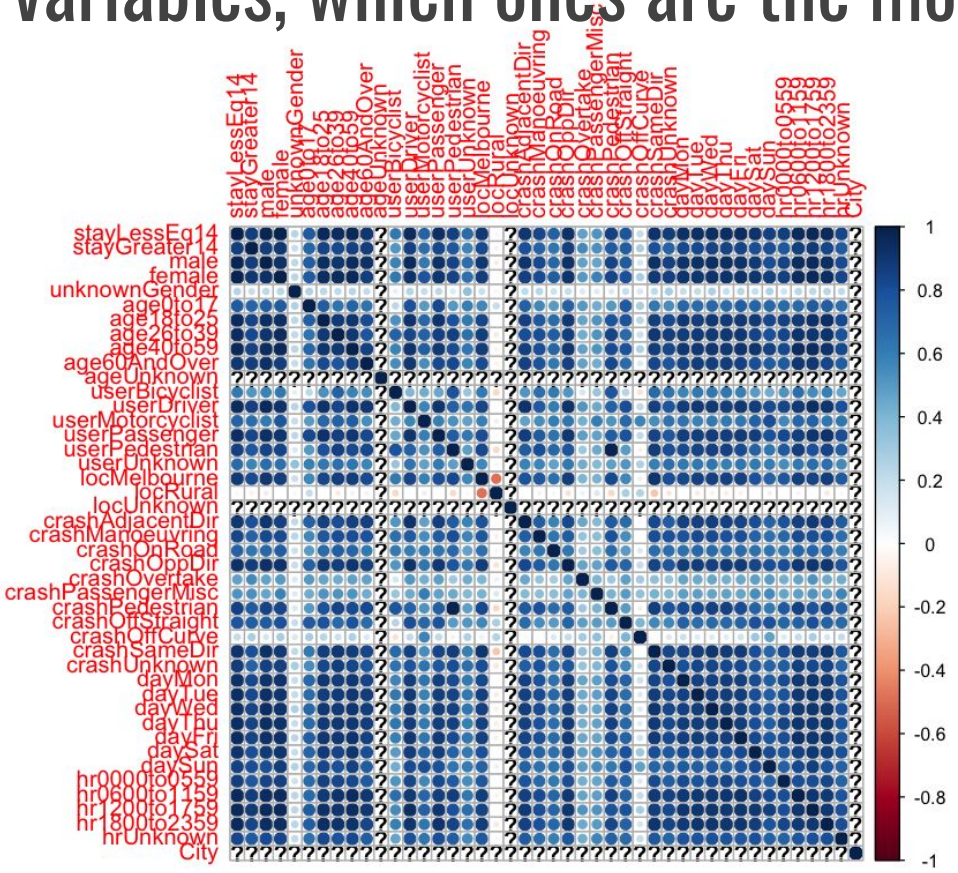


Where are the accidents happening?

*Made using R, ggplot and the Google Map API
To get lat/lon for each local area



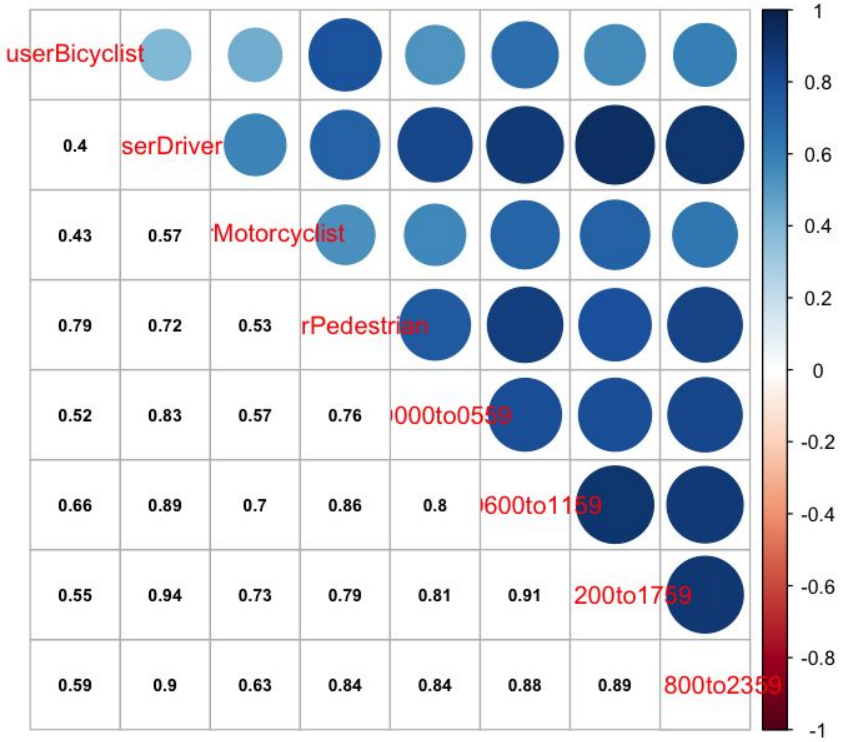
We have 47 variables, which ones are the most useful?



Which Variables are the Most Correlated?

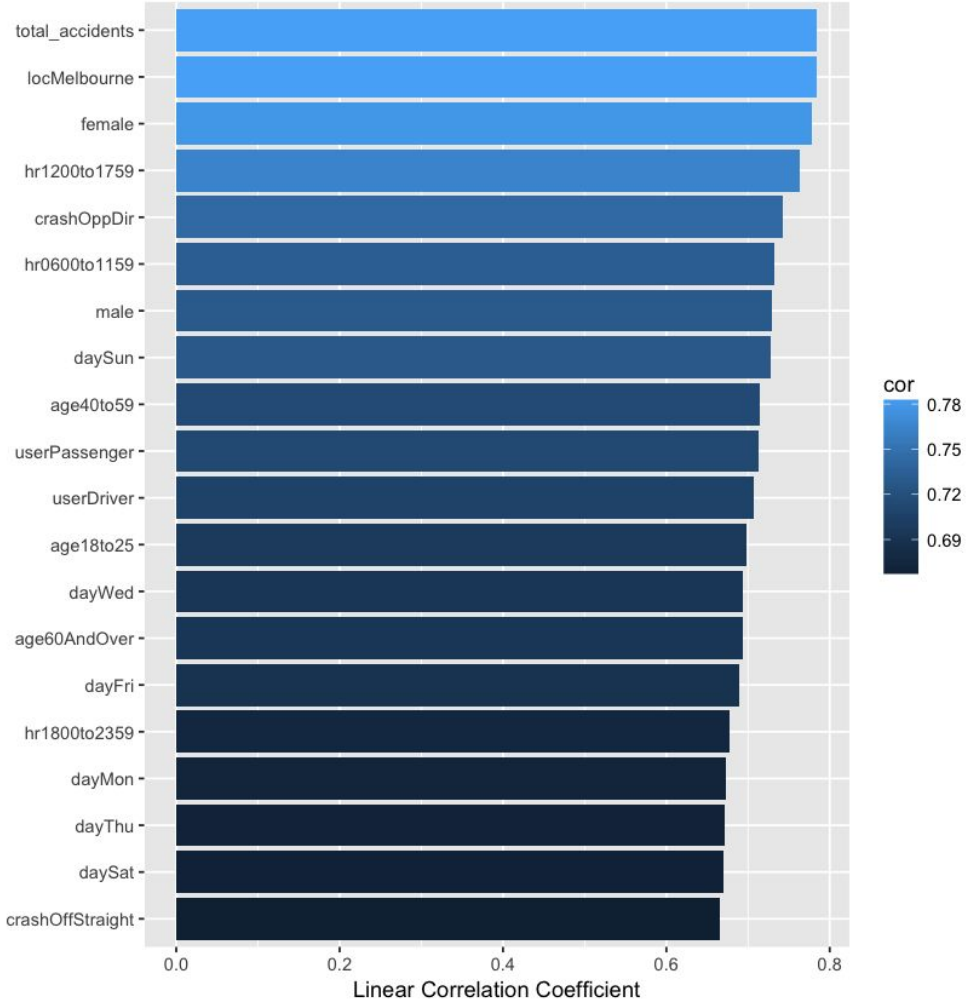
	row	column	cor	p
783	male	hr1200to1759	0.9505827	0
69	male	userDriver	0.9522033	0
16	stayLessEq14	age18to25	0.9537018	0
821	stayLessEq14	hr1800to2359	0.9544301	0
742	stayLessEq14	hr0600to1159	0.9566319	0
784	female	hr1200to1759	0.9582089	0
32	female	age40to59	0.9626181	0
781	stayLessEq14	hr1200to1759	0.9680465	0
29	stayLessEq14	age40to59	0.9696351	0
2	stayLessEq14	male	0.9781661	0
4	stayLessEq14	female	0.9805275	0
341	userPedestrian	crashPedestrian	0.9879493	0

Try Again: When do accidents happen?



What’s the point of all this?
Actionable Solutions

What Correlates With ‘Bad’ Accidents?


total_accidents
locMelbourne
female
hr1200to1759
crashOppDir
hr0600to1159
male
daySun
age40to59
userPassenger
userDriver
age18to25
dayWed
age60AndOver
dayFri
hr1800to2359
dayMon
dayThu
daySat
crashOffStraight
0.0
0.2
0.4
0.6
0.8
Linear Correlation Coefficient
cor
0.78
0.75
0.72
0.69

Based on these correlation, one recommendation for decreasing bad accidents is to increase signage about one way roads.