

THE BATTLE OF NEIGHBOURHOODS

Capstone Final Project – Aiswariya VJ

Context

- Safety and accessibility is a top requirement for anyone moving to a new City/Country
- Melbourne, capital of Victoria Australia is the second most liveable city in the world
- Every year over 100000 migrants call Victoria their new home
- All this means is Victorian market is very competitive with a variety of cultures and people
- It is important for the migrants to find a safe, comfortable, accessible and affordable place to settle down
- People moving interstate would also want this information to move comfortably
- People living in Victoria love Melbourne for its diversity and access to sports, entertainment and many other factors



Problem Nature & Target Audience

- □ The goal of this project is to
 - To show insights on crime data and top venues across top 307 liveable suburbs in victoria enabling our target audience to make the right decision when moving into Victoria
- Target Audience
 - 1. Young couples who enjoy city buzz
 - 2. Migrants new to Victoria
 - 3. Families moving within Victoria and interstate



Capstone Project Presentation - By Ash VJ

Victoria Melbourne Facts

- Melbourne is officially the fox capital of the world
- Melbourne used to have a history if public bathing
- Melbourne's tramway system is the largest outside Europe and the fourth largest in the world
- World's first feature film, was filmed and made in Melbourne in 1906 (Ned Kelly Gang)
- Melbourne was originally going to be named BATMANIA
- 38% of Melbourne's population was born overseas
- Melbourne is the first host of Olympic Games outside of Europe and North America
- The world's largest stained glass ceiling is located in Melbourne
- Melbourne has the highest number of cafes and restaurants per number of people than any other city in the world.
- Melbourne has been named the world's most liveable city 7 times in a row to be beat by Vienna (Austria) last year

Data Acquisition

- □ For this project we will be acquiring data from 2 key sources and various other sources
- Victoria Crime Data the first data source is the Victorian crime/offence data for the year 2019
- 2. FourSquare Venue Data the second data set holds the top venues around various Victorian suburbs
- 3. Victorian Location Data the data consists of all latitude and longitudes for various postcodes/suburbs within Victoria
- 4. Suburb Statistics this data consists of the top 307 liveable suburbs based on various factors developed by Deloitte Analytics

Victorian Crime Data

- The Victorian crime data has been downloaded from
 https://www.crimestatistics.vic.gov.au/crime-statisticslatest-crime-data/download-data-0
- Our initial subset contained over 34K records across 8 columns

Y	ear	Area	Postcode	Suburb	Offence Division	Offence Subdivision	Offence Subgroup	Incidents Recorded
0 20	019	Alpine	3691	Dederang	A Crimes against the person	A20 Assault and related offences	A212 Non-FV Serious assault	1
1 20	019	Alpine	3691	Dederang	A Crimes against the person	Other crimes against the person	Other crimes against the person	1

 Various data cleansing techniques were used to group data, remove redundant columns, and come up a usable subset of records (~7K rows)

	Suburb	Offence	Incidents
0	Abbeyard	B Property and deception offences	1
1	Abbotsford	A Crimes against the person	137
2	Abbotsford	B Property and deception offences	891
3	Abbotsford	C Drug offences	68
4	Abbotsford	D Public order and security offences	80

Capstone Project Presentation - By Ash VJ

Latitude and Longitude Data

- There were multiple ways to extract the latitude and longitude information
- The first option easily was from Geopy and Nominatim, but because of the sheer number of records, we were not able to extract the latitude and longitude information for all suburbs
- 2. The second option was to use the latitude and longitude open source data available online. We downloaded this information from Matthew Proctor's website https://www.matthewproctor.com/full australian postcodes vic

	Locality	Latitude	Longitude
0	BENNISON	-39.140761	146.360412
1	BOOLARONG	-39.140761	146.360412

3. After cleansing, we were able to map the latitude and longitude to our Crime dataset

	Suburb	Offence	Incidents	Latitude	Longitude
0	ABBEYARD	B Property and deception offences	1	-36.554037	146.827590
1	ABBOTSFORD	A Crimes against the person	137	-37.803515	144.998203
2	ABBOTSFORD	B Property and deception offences	891	-37.803515	144.998203

Suburb Statistics

- Domain.com.au and realestate.com.au are the top 2 real estate sites used across
 Australia for buying/renting various property types
- Domain Liveable Melbourne study ranks 307 suburbs in Victoria on 17 indicators to give a list of the most liveable suburbs in the state
- This information is available on <u>Domain.com.au</u>
- This data, once downloaded and cleansed gave us the following rankings along with suburbs and was merged with Victorian Crime Data bringing our records to 1326 suburbs and offences

	Ranking	Suburb
0	1	South Yarra
1	2	East Melbourne
2	3	Carlton
3	4	Fitzroy North
4	5	Hawthorn

	Suburb	Offence	Incidents	Latitude	Longitude
0	SOUTH YARRA	A Crimes against the person	292.0	-37.840679	144.991264
1	SOUTH YARRA	B Property and deception offences	1545.0	-37.840679	144.991264
2	SOUTH YARRA	C Drug offences	91.0	-37.840679	144.991264
3	SOUTH YARRA	D Public order and security offences	108.0	-37.840679	144.991264
4	SOUTH YARRA	E Justice procedures offences	167.0	-37.840679	144.991264

FourSquare Venue Data

- The 307 Victorian suburb latitude and longitude will be used as the input for FourSquare API. Using FourSquare, we will extract venue information for various venues in the top 307 suburbs in Victoria
- Below is what we will receive from FourSquare (Over 2500 venues were downloaded)

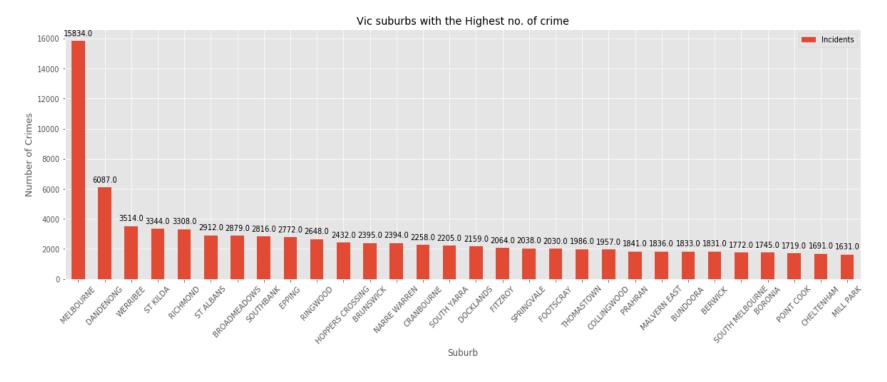
	Suburb	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	MEADOW HEIGHTS	-37.653917	144.923609	Morgan's SUPA IGA Meadow Heights	-37.650340	144.922540	Grocery Store
1	MEADOW HEIGHTS	-37.653917	144.923609	Centro Meadow Heights	-37.650631	144.921664	Shopping Mall
2	MEADOW HEIGHTS	-37.653917	144.923609	Bakers Boutique & Patisserie	-37.650754	144.922006	Bakery
3	MEADOW HEIGHTS	-37.653917	144.923609	My Mates Pizza	-37.650498	144.921988	Pizza Place
4	MEADOW HEIGHTS	-37.653917	144.923609	Kebab House	-37.650444	144.922156	Middle Eastern Restaurant
5	KILSYTH	-37.819364	145.313346	McDonald's	-37.819227	145.314506	Fast Food Restaurant
6	KILSYTH	-37.819364	145.313346	7 Eleven	-37.817650	145.317149	Convenience Store
7	KILSYTH	-37.819364	145.313346	Red Rooster	-37.818590	145.316260	Fast Food Restaurant
8	KILSYTH	-37.819364	145.313346	Nitro Gym	-37.818373	145.316046	Gym

Methodology

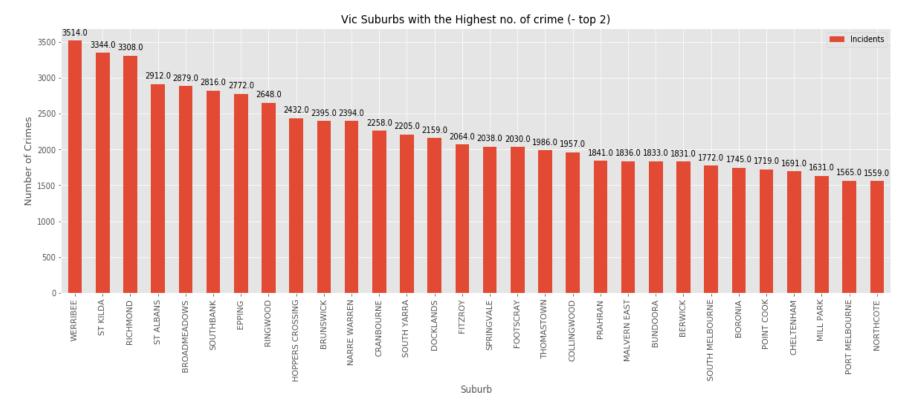
- We will use data visualisation techniques and various analytical techniques to come up with the facts for enabling our target audience to achieve a decision
- Data visualisation will provide us
 - Top crime offence recorded suburbs
 - Lowest crime offence recorded suburbs
 - Types of Crimes in Victoria
- Inferential Statistical Testing
 - Top venues latitude and score
 - Top venues longitude and score
- Modelling
 - K-Means Clustering to identify clusters and target audience
- Mapping
 - Various folium maps presenting suburbs and crime data



Victorian Suburbs with highest number of crime (includes the suburb Melbourne,
 but having Melbourne there is masking all the other data)

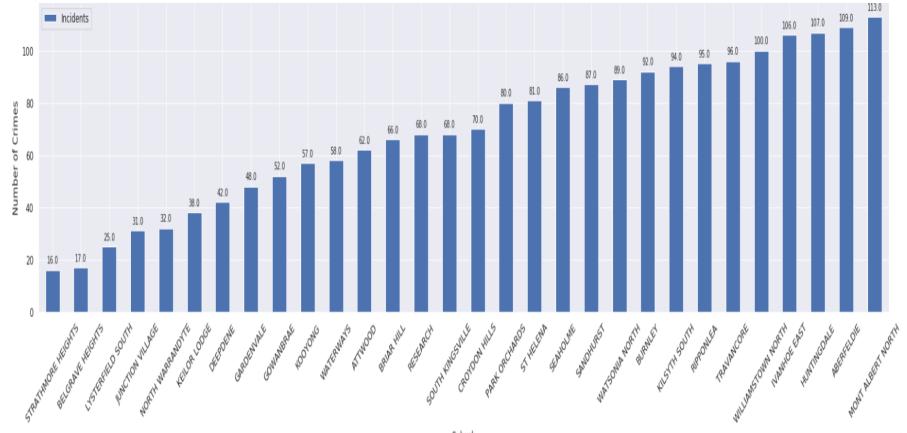


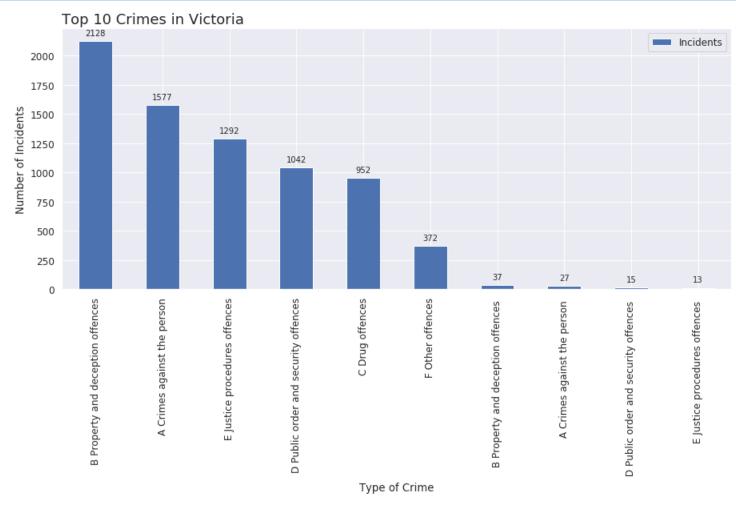
 Victorian Suburbs with highest number of crime (without Melbourne and Dandenong)



Capstone Project Presentation - By Ash VJ

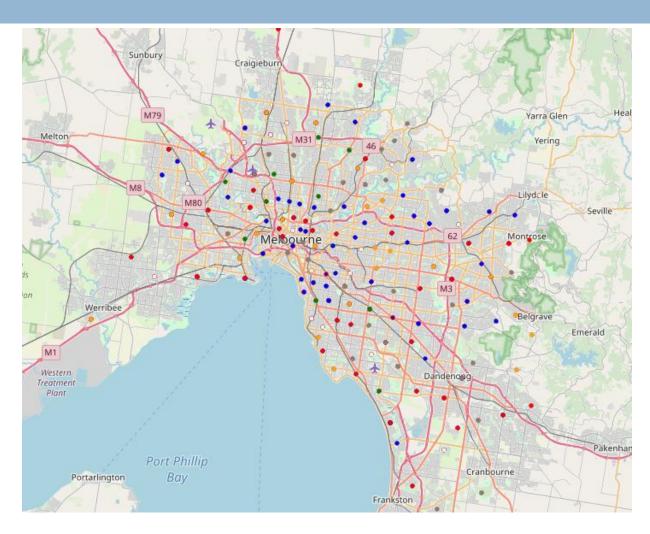
Victorian Suburbs with lowest number of crime





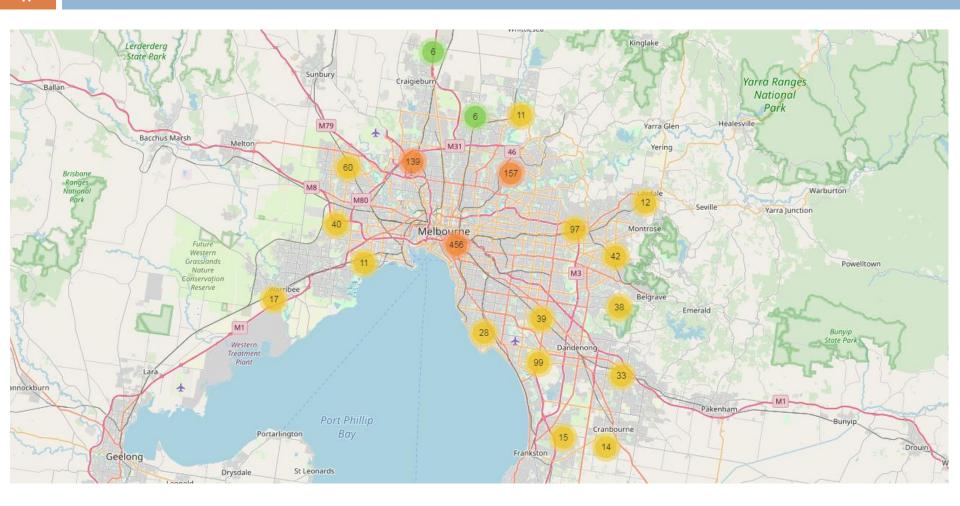
Capstone Project Presentation - By Ash VJ

Maps – Crime Types

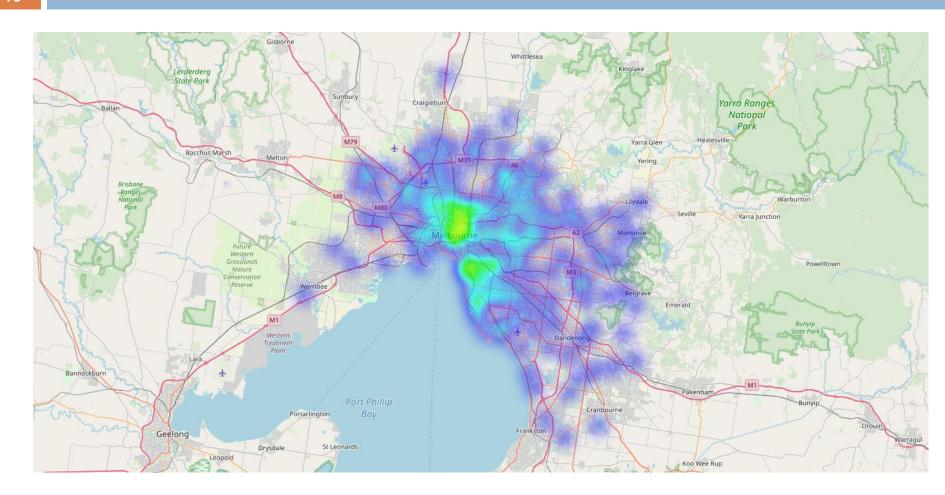


Capstone Project Presentation - By Ash VJ

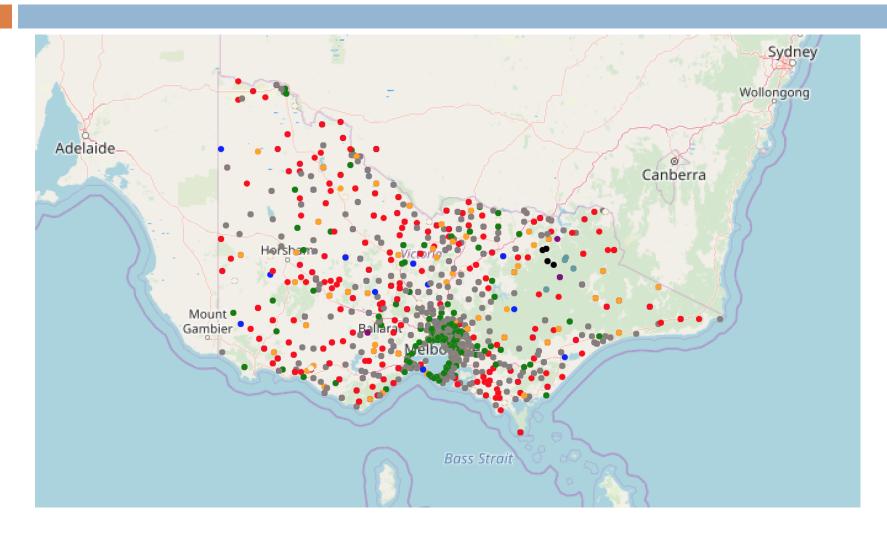
Maps – Crime Cluster



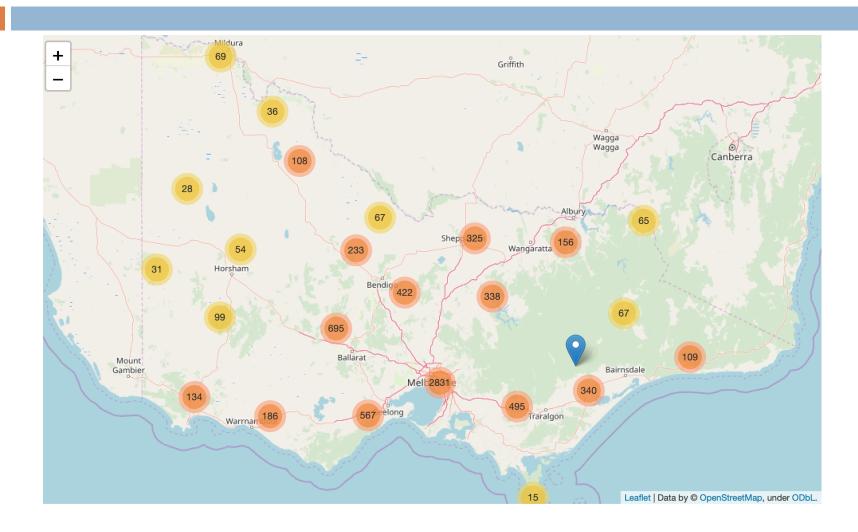
Maps – Crime Heat Map



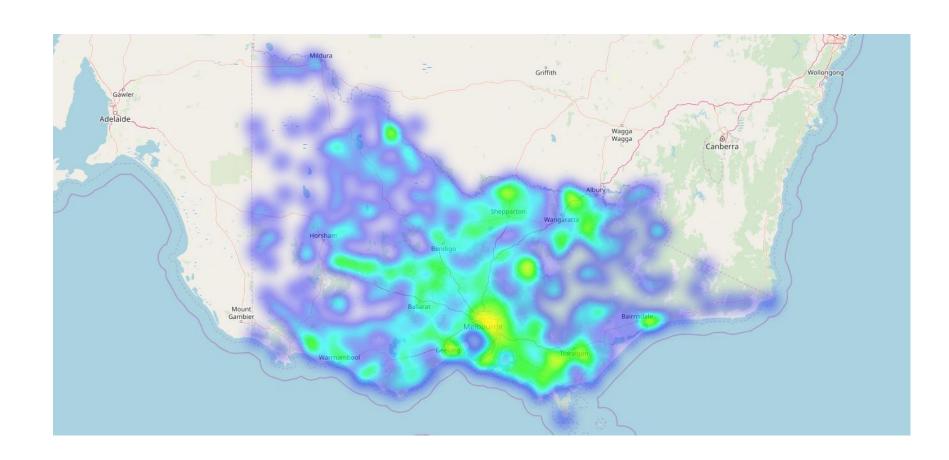
Maps - Crime Types (Vic)



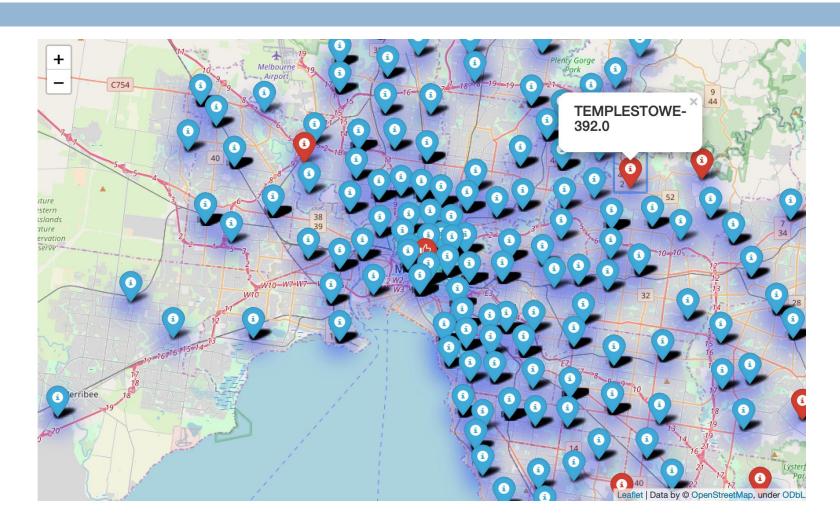
Maps - Crime Cluster (Vic)



Maps – Crime Heat Map (Vic)



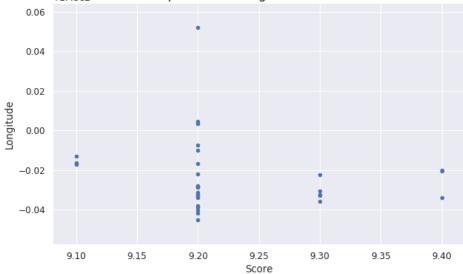
Maps – Crime & Venues (Vic)



Inferential Statistical Analysis

- Inferential statistics allows us to provide insight on a given topic. There are many types of statistical tests that allows one to make inferences. Some of the common statistical tests are:
 - Correlations Chi-square test Independent t-test (aka. Student's t-test) Paired sample t-test Welch's t-test Wilcoxon signed-rank test Linear regression Logistic regression One-way Analysis of Variance (ANOVA) Two-way/N-way ANOVA In this section we will investigate if there are any obvious inferential statistician methods that can help us when modelling the data.

 Scatter Plot of Top Venues Longitude and Score



Capstone Project Presentation - By Ash VJ

Model Creation and Valuation

- The first step in Model creation is to prepare the data for Modelling, we will be removing all unneeded columns and use numerical data only for mode creation
- Lets find the unique venue categories and the venue categories for each suburb
- Using one hot encoding technique we can map 256 unique venue categories to each suburb with the counts. This basically returns us a subset of 277 suburbs across 257 unique venues
- We can then pick up the top 10 venue categories for each suburb. E.g.

	ABBOTSFORD		SOUTH YARRA	
	venue	freq	venue fr	'eq
0	Café	0.19	0 Café 0.	19
1	Pub	0.14	1 Italian Restaurant 0.	10
2	Rock Climbing Spot	0.05	Japanese Restaurant 0.	96
3	Thrift / Vintage Store	0.05	3 Hotel 0.	05
4	Record Shop		4 Bar 0.	04
5	Farmers Market	0.05	5 Convenience Store 0.	03
6	Convenience Store	0.05	6 Thai Restaurant 0.	03
7	Coffee Shop	0.05	7 Grocery Store 0.	03
8	Gay Bar	0.05	8 Juice Bar 0.	03
9	Burger Joint	0.05	9 Coffee Shop 0.	03
	_			

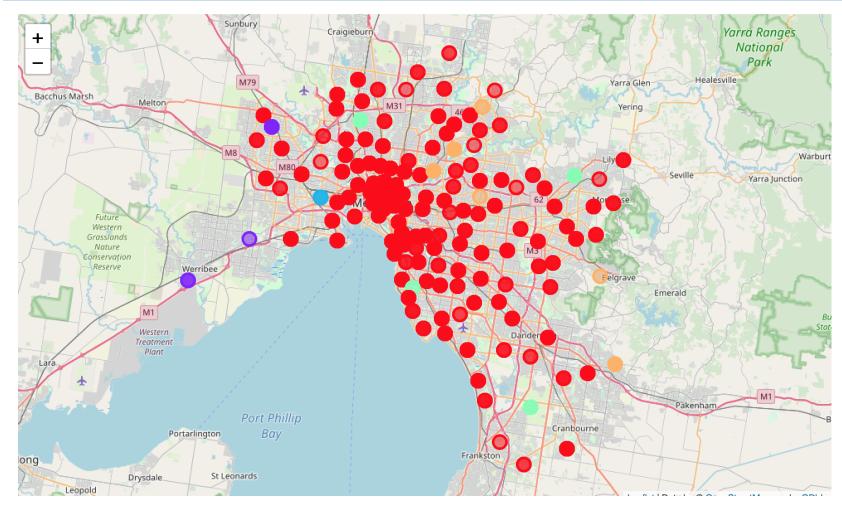
K-Means Clustering

We transpose the data to enable clustering

	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	ABBOTSFORD	Café	Pub	Thrift / Vintage Store	Rock Climbing Spot	Gay Bar	Farmers Market	Greek Restaurant	Grocery Store	Dive Bar	Coffee Shop
1	ABERFELDIE	Gym / Fitness Center	Coffee Shop	Café	Food & Drink Shop	Field	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flea Market	Zoo Exhibit
2	ALBION	Café	Pizza Place	Chinese Restaurant	Filipino Restaurant	Malay Restaurant	Skating Rink	Fast Food Restaurant	Grocery Store	Gym	Sandwich Place
3	ALPHINGTON	Train Station	Café	Farmers Market	Zoo Exhibit	Flea Market	Field	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flower Shop
4	AI TONA	Café	Convenience Store	Thai Restaurant	Train Station	7nn Exhibit	Flea Market	Filinino Restaurant	Fish & Chins Shon	Fish Market	Flower Shon

	Suburb	1st Most Common Venue	2nd Most Common Venue	3rd Mos
0	ABBOTSFORD	Café	Pub	Thrift / Vi
1	ABERFELDIE	Gym / Fitness Center	Coffee Shop	
2	ALBION	Café	Pizza Place	Chinese
3	ALPHINGTON	Train Station	Café	Farn
4	ALTONA	Café	Convenience Store	Thai
5	ARDEER	Portuguese Restaurant	Fast Food Restaurant	
6	ARMADALE	Café	Light Rail Station	Gr
7	ASCOT VALE	Park	Restaurant	
8	ASHBURTON	Pharmacy	Thai Restaurant	
9	ASPENDALE	Furniture / Home Store	Fish Market	
10	ASPENDALE GARDENS	Furniture / Home Store	Fish Market	
11	ATTWOOD	Bakerv	Dive Bar	Gn

K-Means Clustering - Map



Capstone Project Presentation - By Ash VJ

K-Means Clustering – 1st Cluster

	Suburb	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	SOUTH YARRA	-37.840679	144.991264	Master Roll Vietnam	-37.839016	144.992614	Vietnamese Restaurant	0	Café	Italian Restaurant	Japanese Restaurant	Hotel	Bar	Juice Bar	Coffee Shop	Convenience Store	Grocery Store	Thai Restaurant
1	SOUTH YARRA	-37.840679	144.991264	Dainty Sichuan Restaurant	-37.839282	144.992979	Szechuan Restaurant	0	Café	Italian Restaurant	Japanese Restaurant	Hotel	Bar	Juice Bar	Coffee Shop	Convenience Store	Grocery Store	Thai Restaurant
2	SOUTH YARRA	-37.840679	144.991264	Cucinetta	-37.838380	144.990080	Italian Restaurant	0	Café	Italian Restaurant	Japanese Restaurant	Hotel	Bar	Juice Bar	Coffee Shop	Convenience Store	Grocery Store	Thai Restaurant
3	SOUTH YARRA	-37.840679	144.991264	Darling Cafe	-37.838493	144.991432	Café	0	Café	Italian Restaurant	Japanese Restaurant	Hotel	Bar	Juice Bar	Coffee Shop	Convenience Store	Grocery Store	Thai Restaurant
4	SOUTH YARRA	-37.840679	144.991264	Squires Loft Steakhouse	-37.839104	144.992570	Steakhouse	0	Café	Italian Restaurant	Japanese Restaurant	Hotel	Bar	Juice Bar	Coffee Shop	Convenience Store	Grocery Store	Thai Restaurant
5	SOUTH YARRA	-37.840679	144.991264	Ned's Bake	-37.839060	144.990870	Bakery	0	Café	Italian Restaurant	Japanese Restaurant	Hotel	Bar	Juice Bar	Coffee Shop	Convenience Store	Grocery Store	Thai Restaurant

K-Means Clustering – 2nd Cluster

Subi	rb Latitude	Longitude	Venue					1st Most Common Venue				5th Most Common Venue				9th Most Common Venue	10th Most Common Venue
ELTH	M -37.713299	145.158579	Beauty Locker Pty Ltd	-37.717258	145.160782	Cosmetics Shop	1	Cosmetics Shop	Zoo Exhibit	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
ELTH. NOR	M TH -37.713299	145.158579	Beauty Locker Pty Ltd	-37.717258	145.160782	Cosmetics Shop	1	Cosmetics Shop	Zoo Exhibit	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
ESEAR	CH -37.713299	145.158579	Beauty Locker Pty Ltd	-37.717258	145.160782	Cosmetics Shop	1	Cosmetics Shop	Zoo Exhibit	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop

K-Means Clustering – 3rd Cluster

	Suburb	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	TSI MOSI Common Venue	znu most Common Venue	ora wost Common Venue	4th Most Common Venue	oui wost Common Venue	our wost Common Venue	rui most Common Venue	oun wost Common Venue	Sui Most Common Venue	Common Venue
MEL	.BOURNE	-38.365017	144.765920	Stringers Road Reserve	-38.362743	144.763987	Park	2	Park	Zoo Exhibit	Indian Restaurant	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Flower Shop
	IVANHOE	-37.772675	145.048573	Ivanhoe Park	-37.772876	145.052403	Park	2	Park	Ice Cream Shop	Pet Store	Zoo Exhibit	Flea Market	Field	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flower Shop
	IVANHOE	-37.772675	145.048573	Chelsworth Park	-37.771078	145.048118	Park	2	Park	Ice Cream Shop	Pet Store	Zoo Exhibit	Flea Market	Field	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flower Shop
	IVANHOE	-37.772675	145.048573	All 4 Pets	-37.771980	145.045830	Pet Store	2	Park	Ice Cream Shop	Pet Store	Zoo Exhibit	Flea Market	Field	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flower Shop
	IVANHOE	-37.772675	145.048573	Fairy Hills Gourmet Ice Cream	-37.771650	145.045720	Ice Cream Shop	2	Park	Ice Cream Shop	Pet Store	Zoo Exhibit	Flea Market	Field	Filipino Restaurant	Fish & Chips Shop	Fish Market	Flower Shop
	IVANH∩F									Ice Cream					Filinino	Fish & Chins		

K-Means Clustering – 4th Cluster

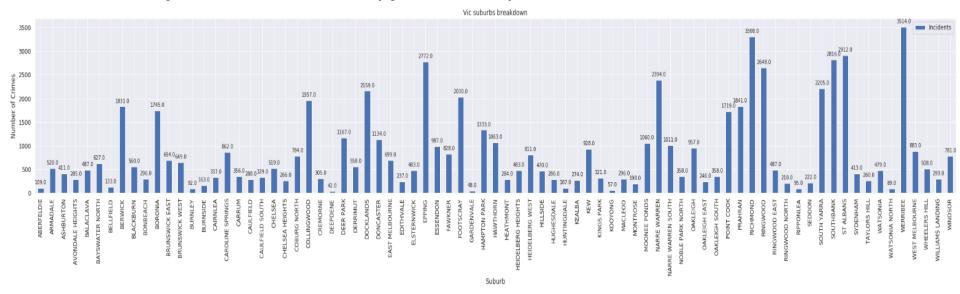
Suburb	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	ora wost Common Venue	4ui most Common Venue	our wost Common Venue	our wost Common Venue	7ui most Common Venue	Bull Most Common Venue	9th Most Common Venue	10th Most Common Venue
IGSVILLE	-37.807135	144.861162	Amazing Tiles & Stone	-37.809818	144.861839	Construction & Landscaping	3	Construction & Landscaping	Zoo Exhibit	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
WEST OTSCRAY	-37.807135	144.861162	Amazing Tiles & Stone	-37.809818	144.861839	Construction & Landscaping	3	Construction & Landscaping	Zoo Exhibit	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
IDSTONE	-37.807135	144.861162	Amazing Tiles & Stone	-37.809818	144.861839	Construction & Landscaping	3	Construction & Landscaping	Zoo Exhibit	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
SOOKTAN	-37.807135	144.861162	Amazing Tiles & Stone	-37.809818	144.861839	Construction & Landscaping	3	Construction & Landscaping	Zoo Exhibit	Fast Food Restaurant	Frozen Yogurt Shop	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop

K-Means Clustering – 5th Cluster

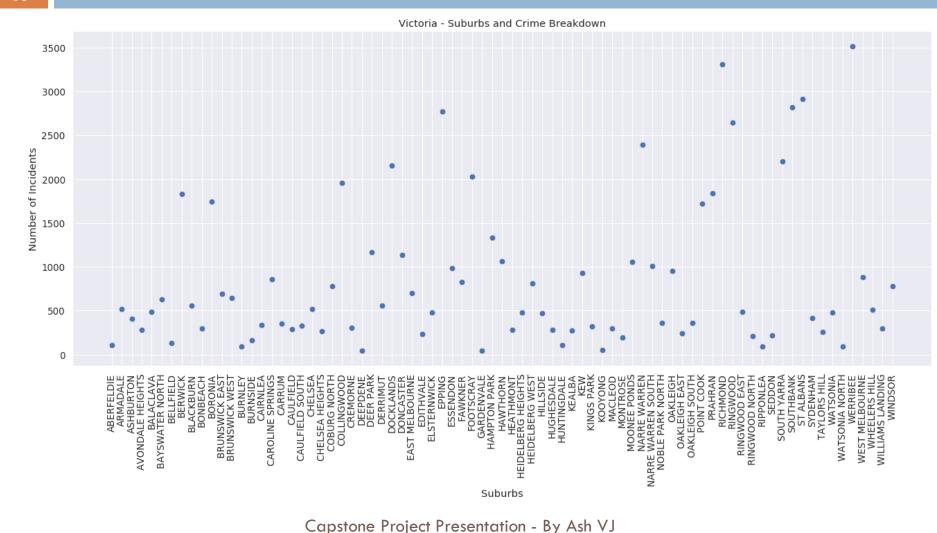
Suburb	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
HOPPERS ROSSING	-37.837165	144.705831	Alfred Park	-37.837863	144.700572	Playground	4	Playground	Zoo Exhibit	Farmers Market	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Flower Shop
INT COOK	-37.916240	144.642090	Riverwalk Village	-37.918152	144.643094	Playground	4	Playground	Zoo Exhibit	Farmers Market	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Flower Shop
VERRIBEE	-37.916240	144.642090	Riverwalk Village	-37.918152	144.643094	Playground	4	Playground	Zoo Exhibit	Farmers Market	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Flower Shop
)ERRIMUT	-37.916240	144.642090	Riverwalk Village	-37.918152	144.643094	Playground	4	Playground	Zoo Exhibit	Farmers Market	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Flower Shop
YDENHAM	-37.715579	144.780474	Proctor Cresent Reserve	-37.711862	144.783197	Playground	4	Music Store	Playground	Zoo Exhibit	Fast Food Restaurant	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
YDENHAM	-37.715579	144.780474	AKA Music	-37.711250	144.780020	Music Store	4	Music Store	Playground	Zoo Exhibit	Fast Food Restaurant	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
WILLIAMS LANDING	-37.861998	144.743016	Williams Landing Playground	-37.859709	144.743765	Playground	4	Playground	Zoo Exhibit	Farmers Market	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop	Flower Shop
.ORS HILL	-37.715579	144.780474	Proctor Cresent Reserve	-37.711862	144.783197	Playground	4	Music Store	Playground	Zoo Exhibit	Fast Food Restaurant	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
.ORS HILL	-37.715579	144.780474	AKA Music	-37.711250	144.780020	Music Store	4	Music Store	Playground	Zoo Exhibit	Fast Food Restaurant	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
HILLSIDE	-37.715579	144.780474	Proctor Cresent Reserve	-37.711862	144.783197	Playground	4	Music Store	Playground	Zoo Exhibit	Fast Food Restaurant	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop
HILLSIDE	-37.715579	144.780474	AKA Music	-37.711250	144.780020	Music Store	4	Music Store	Playground	Zoo Exhibit	Fast Food Restaurant	Fried Chicken Joint	French Restaurant	Football Stadium	Food Truck	Food Court	Food & Drink Shop

Analysis – For Target Audience

- Our target audience is young couples, migrants and new families. From the above, let us make an assumption on what our target audience will be looking for when settling down
 - Young Couples Pub, Pizza, Café
 - Families Grocery Store, Pharmacy, Train Station
 - Migrants Train Station, Playground, Bakery



Analysis – For Target Audience (cont'd)



Results & Discussion

- The aim of this project is to assist young couples, new families, and migrants to come to Victoria and stay in a safe and convenient location
- Say for example, as an young couple, they are looking for party life and what not, they will be focussing on cluster 1, we have also identified the possible suburbs they would be interested in using their first, second and third choice
- If a family is planning to move in an area which has playgrounds and heath fitness venues, they would probably consider cluster 2, 4 and 5. While someone who is super exited about Motor bikes would actually choose a suburb in cluster 3
- These are some of the findings with our specific target audience. We have also only considered suburbs which are within Melbourne's reach and have been listed in the top 300 suburbs to live in Victoria. Any one should be able to view these clusters and get the outcomes they need.

Conclusion and Further Development

- In the current state of affairs, the most important thing for a human is to make effective decisions considering all the factors which will impact the decision
- Doing this without the help of technology is not feasible in this day and age
- The above project actually enables a person to use this project and make a informed decision based on two key factors - 1. SAFETY and 2. INTEREST
- The future of this project will include other factors like Rent, public transport ratings, school ratings and most of all age group of people living in that suburb.