

Generic Real Estate Consulting Project

GROUP 38



Group Member

1

Haoran Guo



Data collection

2

Jia man Xu



3

Muhan Chu



Data collection

4

Xinxiu Dai



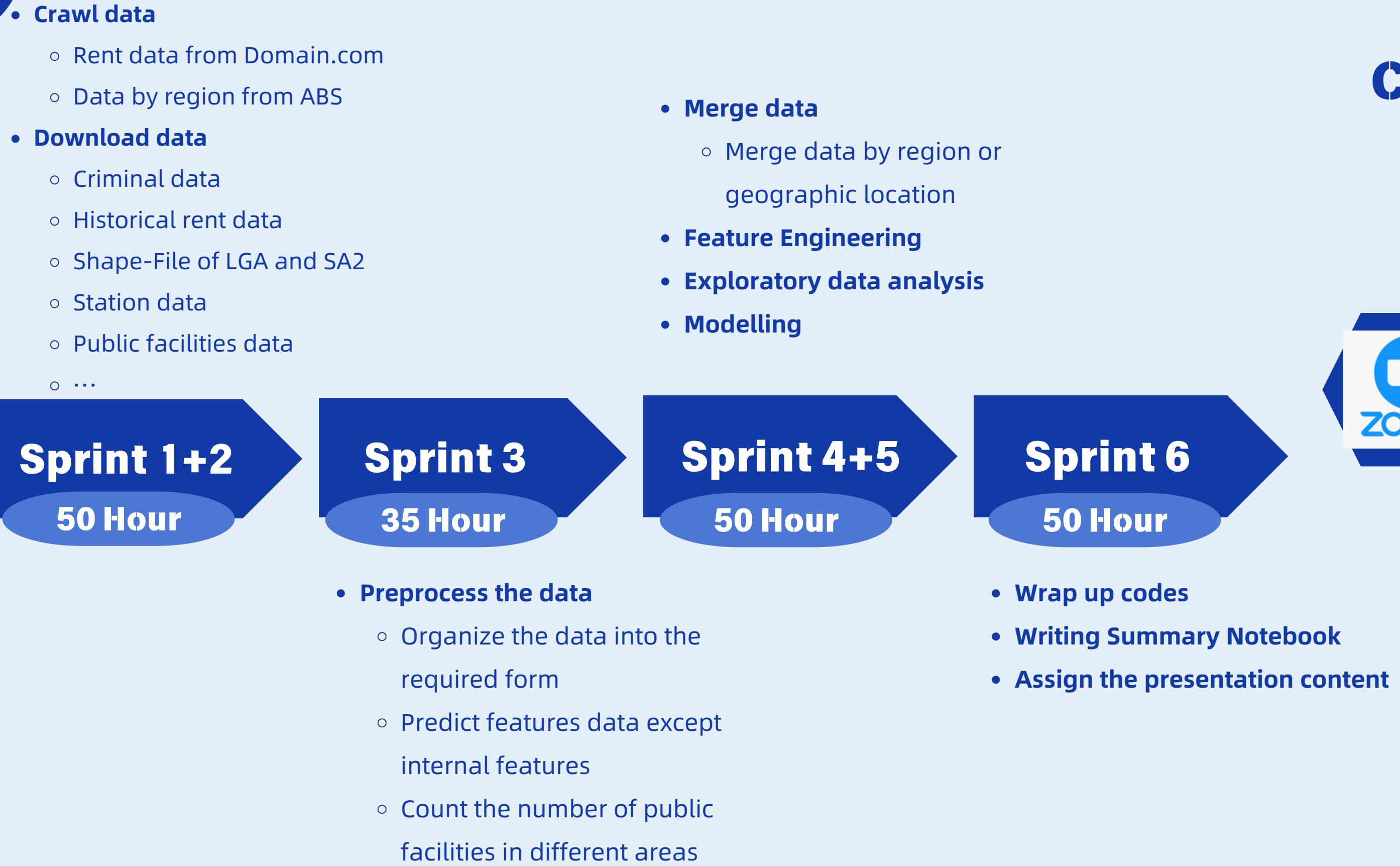
5

zhihan Wang

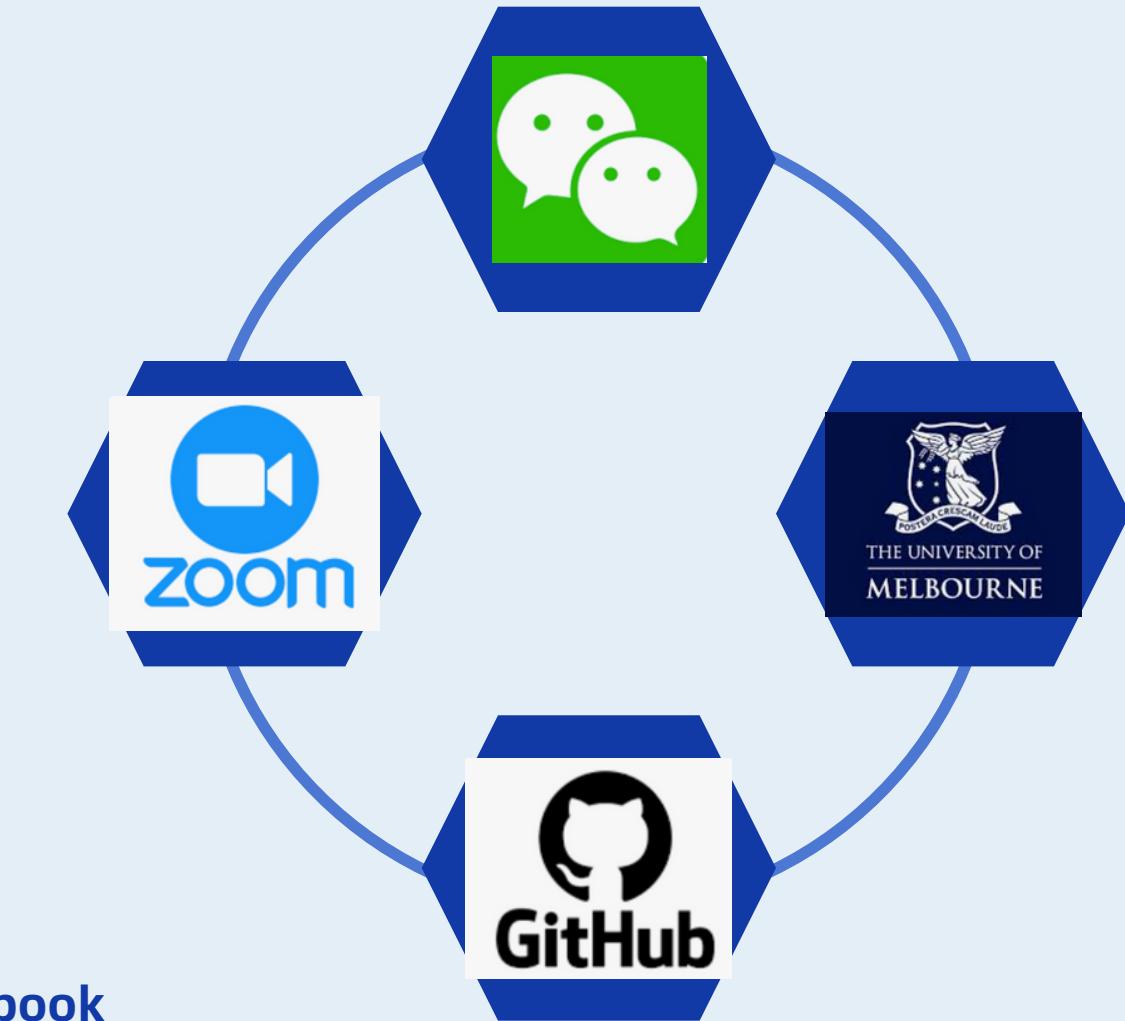


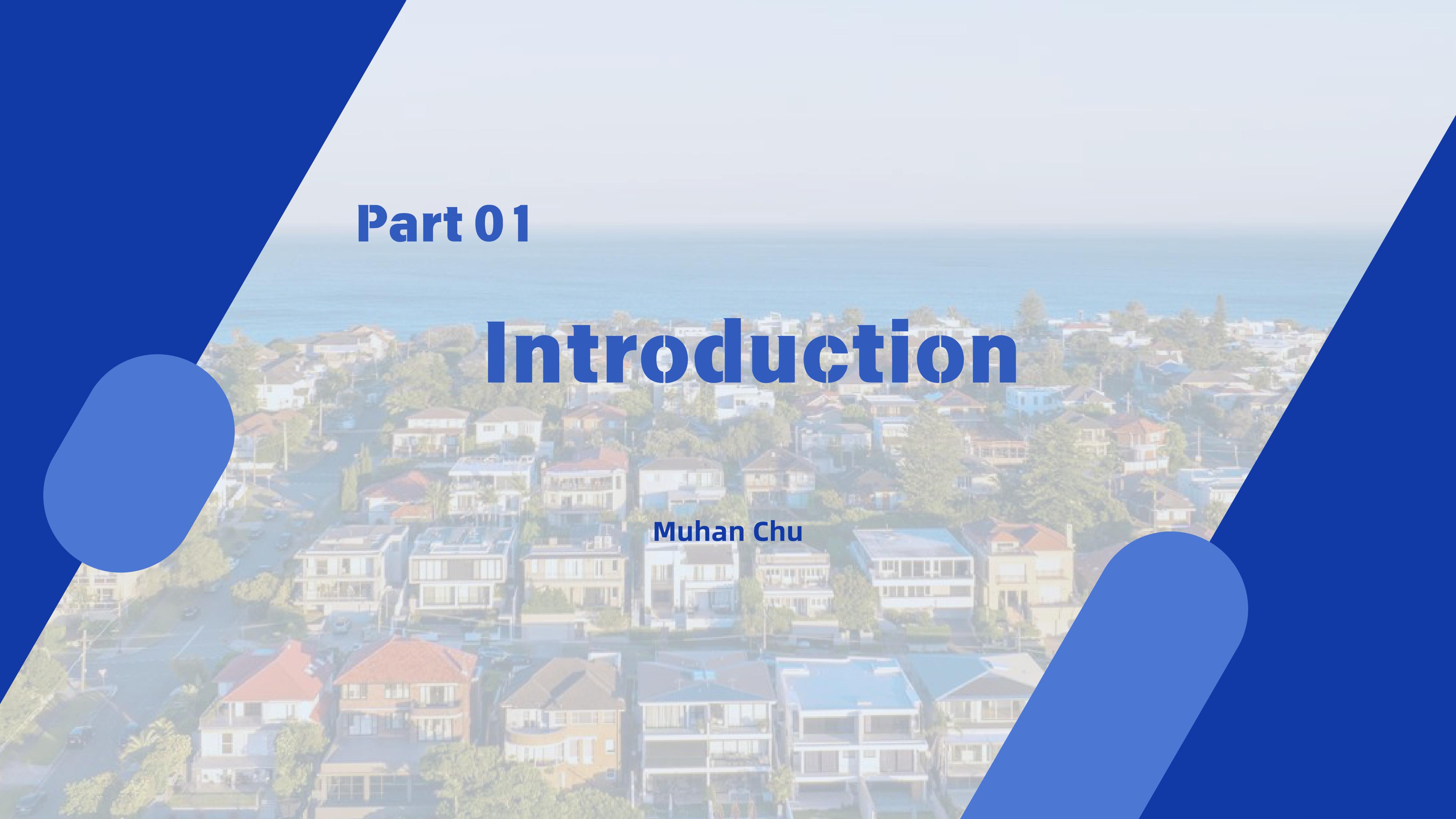
Machine Learning Model Developer Machine Learning Model Developer

TimeLine of the Project



Communication



The background image shows a coastal town built on a hillside. The town is densely packed with houses of various sizes and colors, mostly with red roofs. The hillside is covered with green trees and bushes. In the foreground, there are some streets with parked cars. The town extends towards the horizon where it meets a clear blue ocean under a bright sky.

Part 01

Introduction

Muhan Chu

Project Overview

1

What are the key internal and external factors that impact rental prices?

2

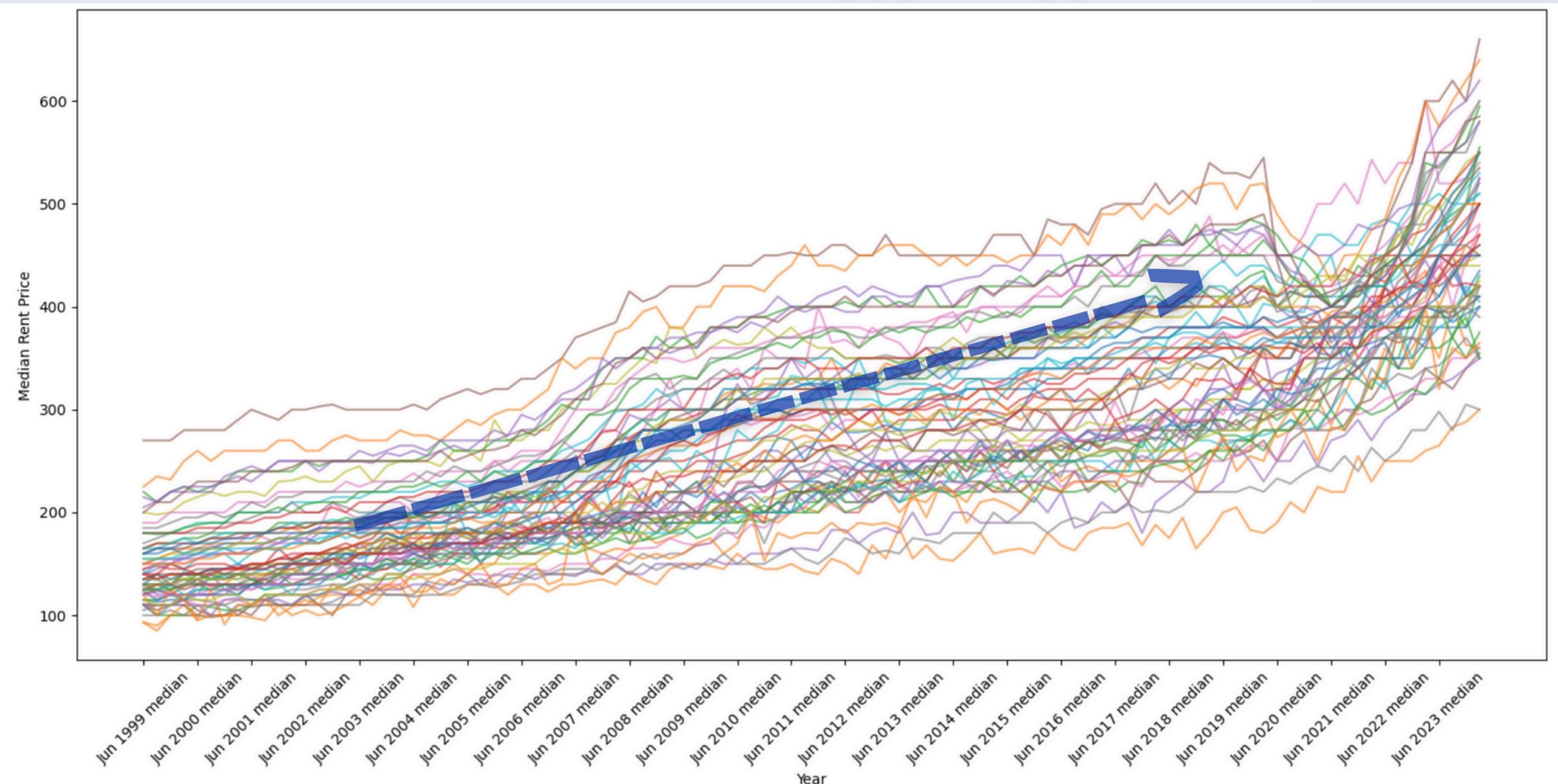
Which regions are projected to have the highest rent growth rates over the next three years?

3

Which region are the most livable and affordable?

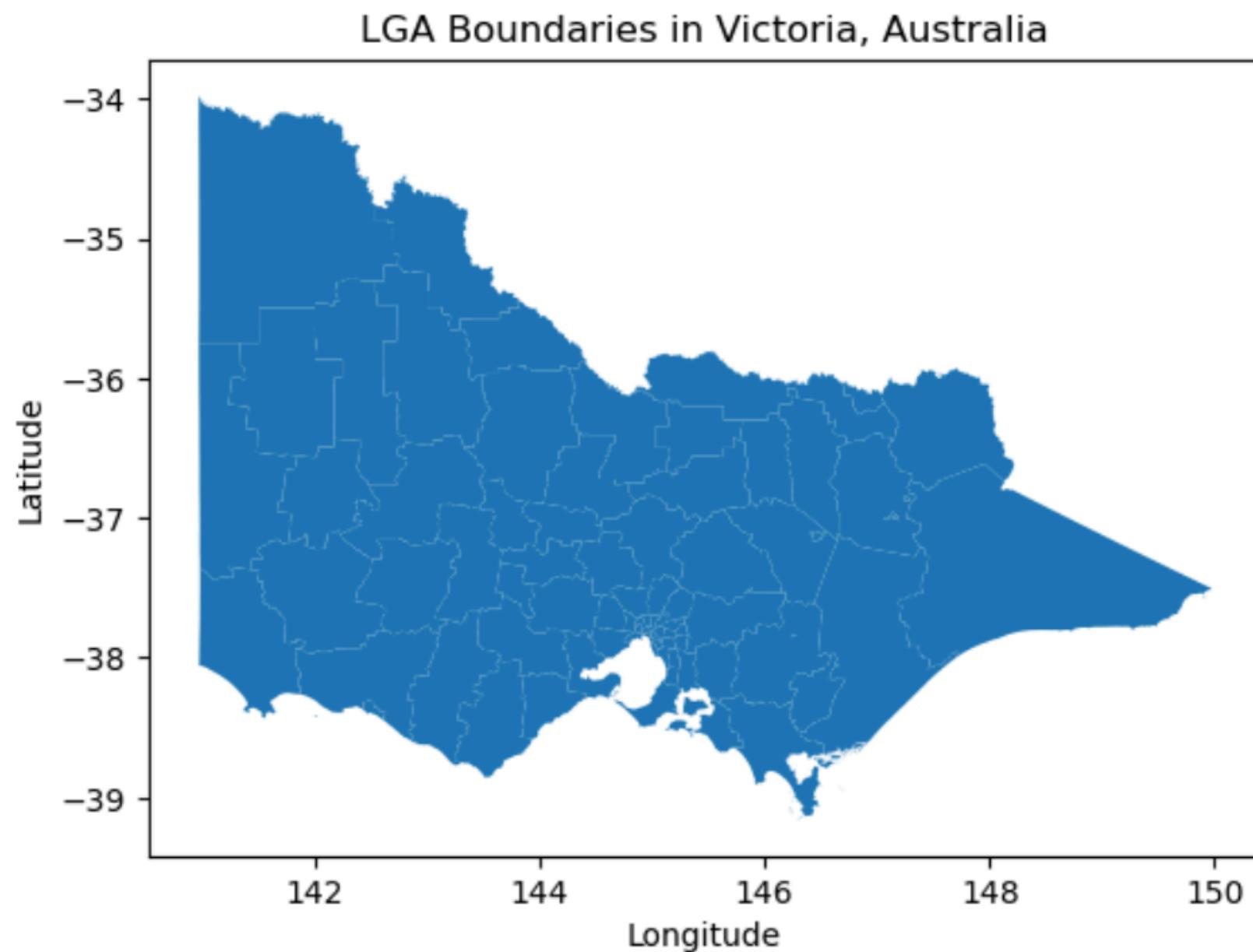


Rent background in Victoria

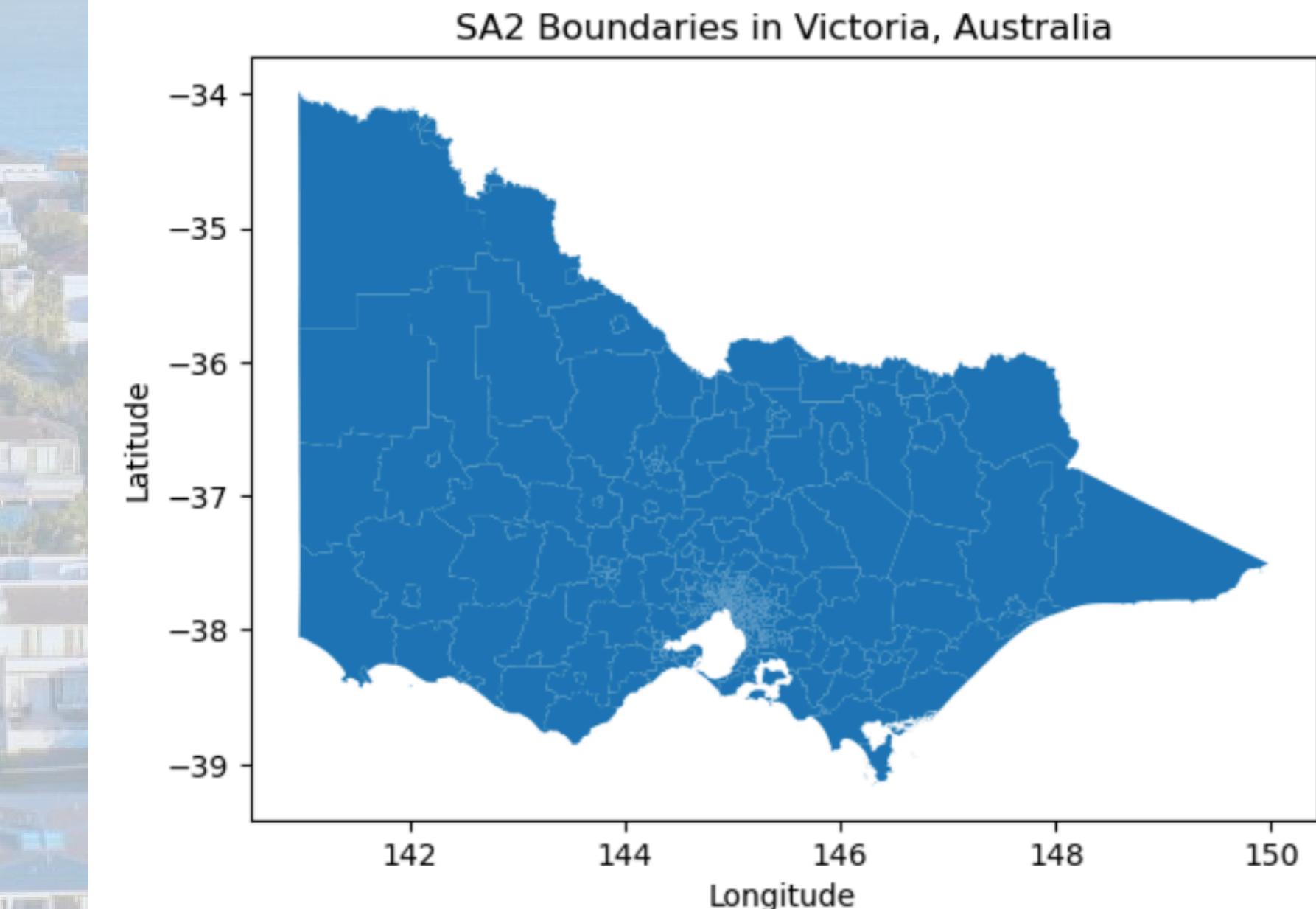


Geographical information of Victoria

Local Government Area



Statistical Area Level 2



Part 02

Data and strategy

Zhihan Wang, Haoran Guo, Xinxiu Dai

Security and Privacy

Data Security:

Only authorized members of the team have access to our code, data and API

Data Privacy:

We follow the principle of collecting only the minimum amount of data necessary for our business and avoid excessive collection of information.



What data do we have?

Domain Rental Properties Data

Criminal data

Victoria state stops data

Nearby facilities data

schools, parks, fire stations, police stations,
shopping centers, supermarkets, restaurants,
cinemas, distance to university, distance to CBD

Map visualization data

LGA historical rent data

ABS region data

resident population, income, Number of jobs,
Land area, etc...

Correlation Analysis for Numerical Features : Spearman Correlation

Spearman Feature	Spearman Value
Cost text	1.000
Bath info	0.531
Bed info	0.524
Median monthly household mortgage payment (\$)	0.397
Median weekly household rental payment (\$)	0.388
Median price of established house transfers (\$)	0.348
Parking	0.326
Average crime count	-0.292
Median total income (excl. Government pensions) (\$)	0.216
Land area (ha)	-0.188

Spearman Feature	Spearman Value
Cost text	1.000
Distance to Monash University, Clayton Campus	-0.306
Distance to Deakin University - Burwood Campus	-0.298
Distance to La Trobe University (Bundoora Campus)	-0.243
Distance to Melbourne Central	-0.228
Distance to the University of Melbourne, Parkville Campus	-0.220
Secondary/Primary School Count Within 3000m	0.183
Cinema/Theatre Count Within 3000m	0.134
Distance to Hospital	-0.130
Restaurant/Bar Count Within 1000m	0.106

Correlation and Statistical Analysis for Feature Importance

Categorical Features: ANOVA

Feature	p-value
Property Type: House	<0.05
Property Type: New Apartments / Off the Plan	0.84
Property Type: New House & Land	0.81
Property Type: Studio	<0.05
Property Type: Terrace	0.25
Property Type: Townhouse	0.04
Property Type: Villa	0.95
Post Code	<0.05
Region Encoded	<0.05

Model Analysis for Feature Importance

Tree-based model

- REF-CF

Random forest

By combining multiple decision trees, the model improve the accuracy and robustness.

Extreme Gradient Boosting

Combine multiple weak learners to improve the prediction accuracy

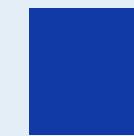
Internal Features

- Number of Bathroom
- Number of Bedroom
- Number of Parking Spaces

External Features

- Median price of established house transfers
- Distance to universities
- Number of Cinema within 3km

Prediction Model Analysis



Time series

Prophet

Easy to use, with automatic parameter optimization.



Regression

Linear regression

suitable for linearly separable data

SARIMA

Effectively captures long-term trends and seasonality.

Random forest regression

Predicting continuous outcomes through the multiple decision trees

Part 03

Result Analysis

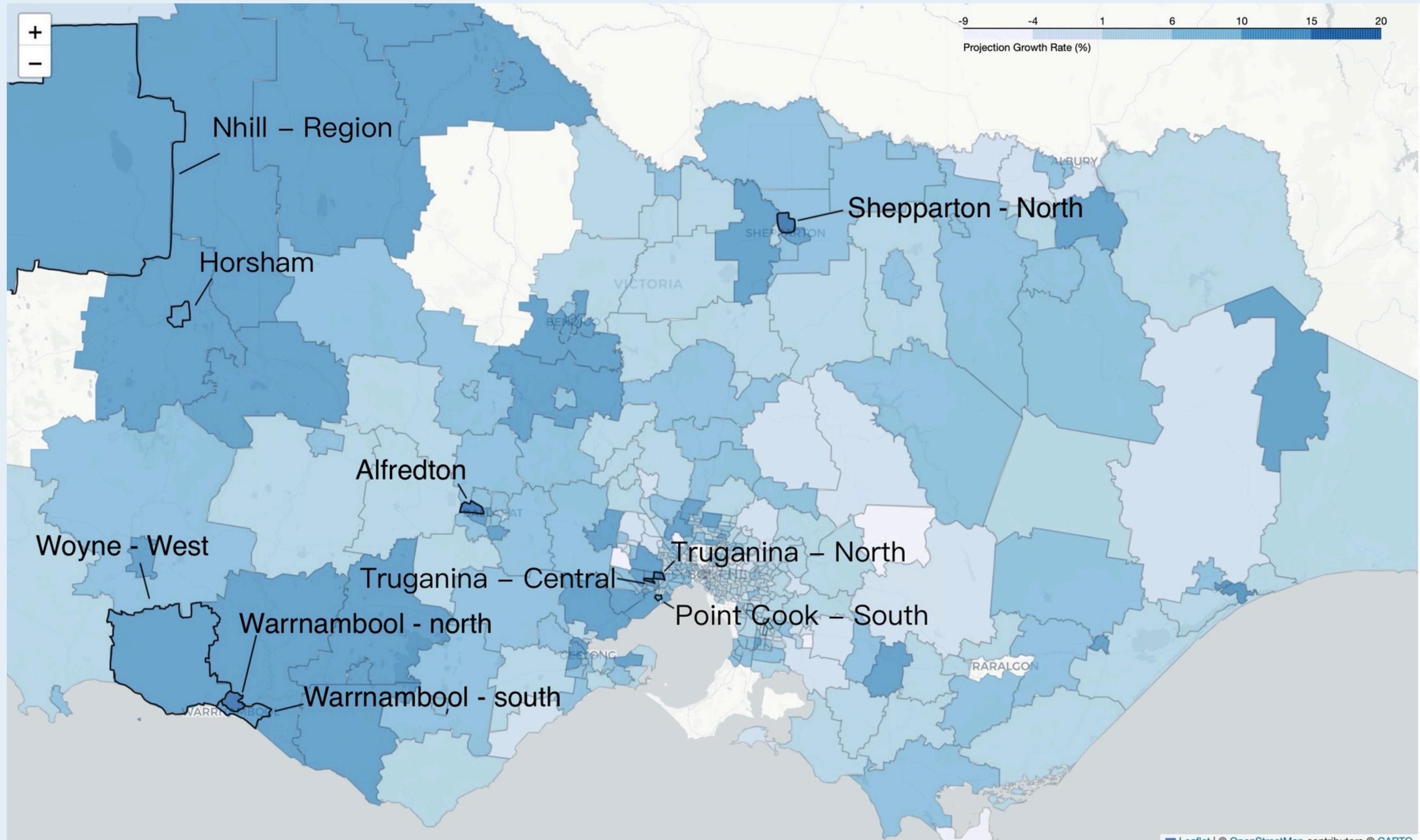
Xinxiu Dai, Jiaman(Mandy) Xu

What are the top 10 suburbs with the highest predicted growth rate?

Result Analysis - Top 10 Regions with the Highest Rental Growth Rates

SA2 NAME	Mean Growth Rate	SA2 NAME	Mean Growth Rate
Shepparton - North	20.90%	Nhill Region	15.49%
Truganina - North	17.64%	Warrnambool - South	15.44%
Warrnambool - North	17.48%	Horsham	15.25%
Tarneit - Central	17.22%	Moyne - West	15.13%
Alfredton	16.23%	Point Cook - South	15.08%

What are the top 10 suburbs with the highest predicted growth rate?



What are the most liveable and affordable suburbs?

Result Analysis - Key Factors in Liveability Index

Affordability

Rent-to-income ratio

Compares median rent to household income

Lower ratios = better affordability



Safety

Evaluated through **crime rates** and
homelessness data

Lower crime and homelessness
= higher safety scores



Amenities

Availability of **parks, schools, healthcare, and shops**

Higher amenity availability
= better quality of life

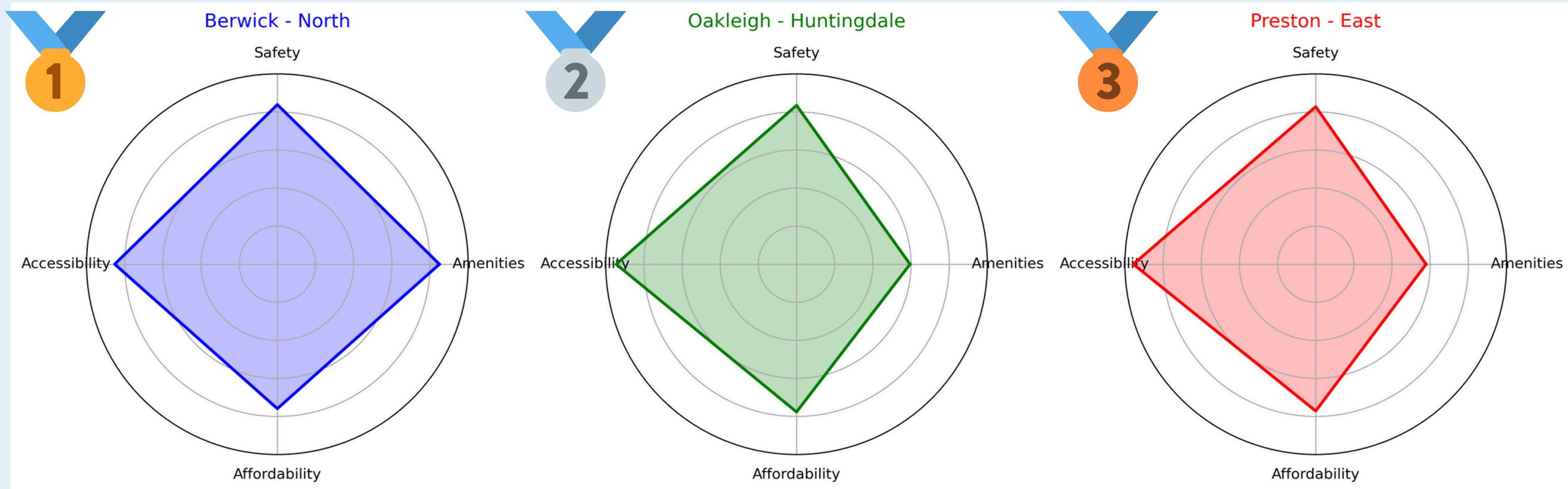
Accessibility

Public transport options and proximity
to **Melbourne CBD**

Better connectivity
= higher accessibility scores

What are the most liveable and affordable suburbs?

Result Analysis - Top 3 Suburbs Based on Liveability Index



Berwick - North

Excels in amenities

Oakleigh - Huntingdale

Strong in accessibility

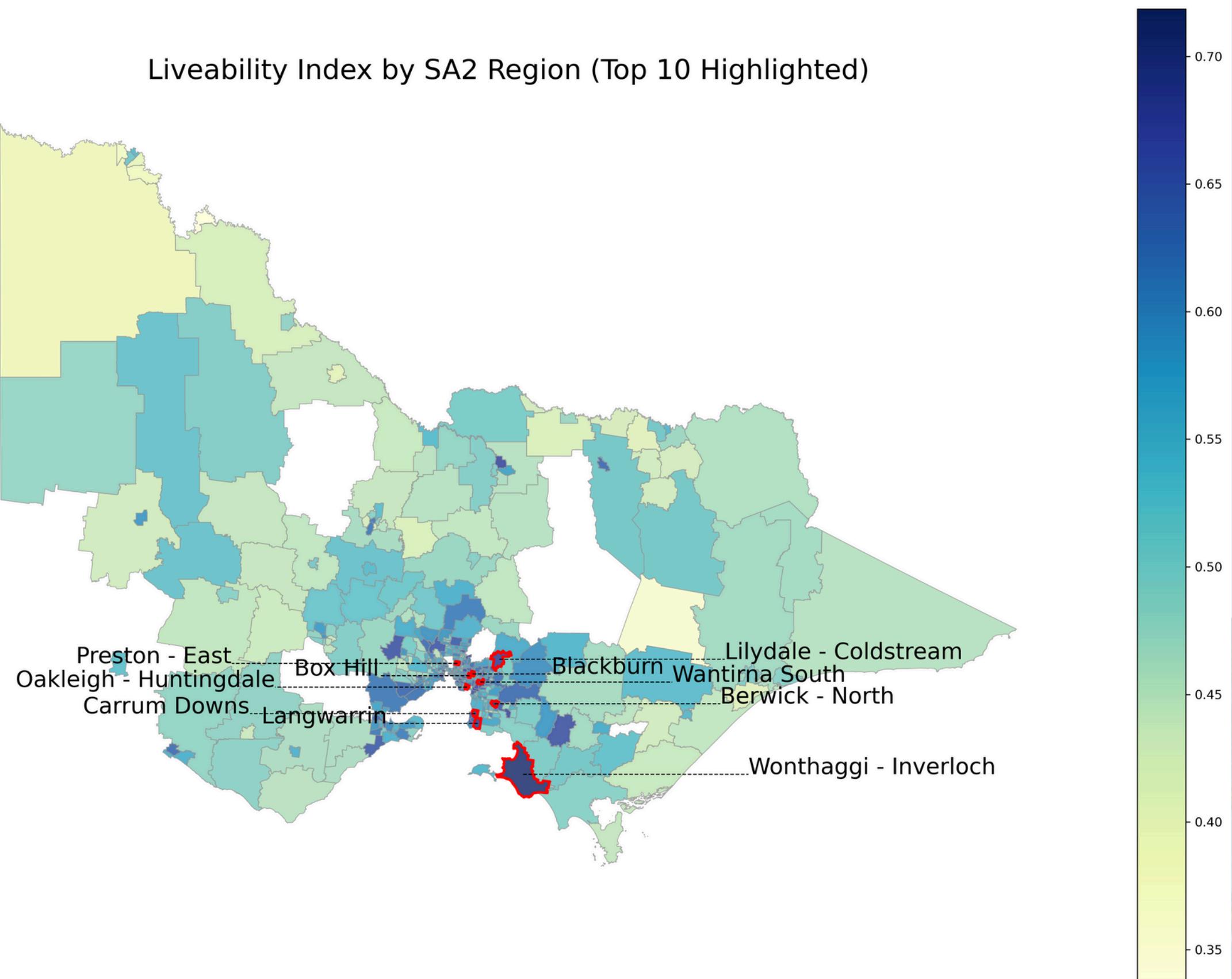
Preston - East

Leads in accessibility

What are the most liveable and affordable suburbs?

Result Analysis - Top 10 Liveable Suburbs

Liveability Index by SA2 Region (Top 10 Highlighted)



Rank	SA2	Liveability Index
1	Berwick - North	0.7214
2	Oakleigh - Huntingdale	0.7140
3	Preston - East	0.7110
4	Wonthaggi - Inverloch	0.7084
5	Carrum Downs	0.6970
6	Blackburn	0.6960
7	Langwarrin	0.6869
8	Wantirna South	0.6848
9	Lilydale - Coldstream	0.6818
10	Box Hill	0.6806

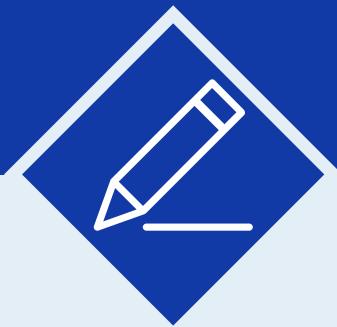
Part 04

Discussion

Muhan Chu

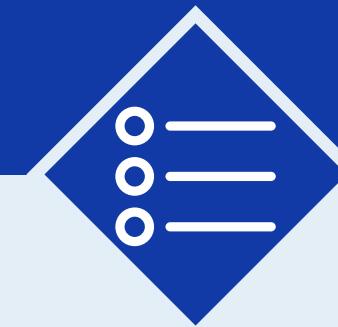
Limitations

1. Not every region has enough rental data
2. Absence of External data
3. Data consistency



Difficulties

1. When combining the data, the SA2 names are different for different datasets
2. The dimensionality of the data is large
3. dynamic data scraping (Time consuming)



Assumptions

1. The predicted external features of previous years and the future are close to the true value
2. Median historical rent data according to LGA is also suitable for SA2
3. The number or location of public facilities remains the same



Improvement

1. Crawl more comprehensive rental information from more websites
2. Use more granular historical rental data



Recommendation



Predict rental price by:

- Internal feature
 - Number of Bathroom
 - Number of Bedroom
 - Parking

External feature

- Median price of established house transfers
- distance to monash university
- cinema theatre count within 3000m



Invest region with the highest growth rate over the next three years:

- South West Victoria
- South Melbourne



Advertise properties in affordable, liveable areas:

- South-East Melbourne
- Eastern Melbourne

Business Development



Turned into money

- **Real estate investor**
 - Use the model to identify regions of high reward
- **Real estate agency or rental platform**
 - Rent pricing tool
 - Tools for matching customer needs and listings



Current progression and approach

- Feature selection
- Modelling
- Result analysis

TimeFrame 3 – 6 Month

- User interface
- Deploying models
- Performance optimisation
- Conducting market promotion

The background of the slide features a photograph of a coastal town, likely Bondi Beach in Sydney, Australia. The town is built on a hillside, with numerous houses and apartment buildings of various colors (white, blue, yellow) and architectural styles. The town is surrounded by lush green trees. In the distance, the calm blue ocean meets a clear sky. The overall atmosphere is bright and airy.

**THANK YOU
GROUP 38**