Yoga and Pilates in Victoria: Finding a Market Need

By Amanjit Gill

The situation

2.2 million Australians practise yoga.

1.2 million Australians practise pilates.

Large increase in participation.

Large number of businesses offering yoga/pilates services. Risk of market saturation.

The problem

Where is there a market need for a new yoga/pilates studio?

Focus on Victoria, Australia.

Data

- Shapefile of local government areas (LGAs) in Victoria
- General community profiles from Australian Bureau of Statistics (ABS)
- Australian Statistical Geography Standard (ASGS) for Victoria

Data (cont...)

- Foursquare API data on yoga/pilates studios in Victoria
- Regional population growth from ABS

Data (cont...)

Final cleaned dataset:

Feature Name	Feature Description
lga_code	unique identifier for the LGA
lga_name	name of the LGA
total_women	total number of women in the LGA
median_fam_inc_weekly	median family weekly income in the LGA
pcent_women_employed	percent of women in the LGA who are employed
pcent_women_high_income	percent of women in the LGA on a high income
pcent_women_tertiary_educ	percent of women in the LGA who are tertiary educated
studios_per_10000_women	number of studios in the LGA per 10000 women

Methodology

1. Exploratory analysis
   - descriptive statistics
   - correlation coefficient

2. Machine learning
   - k-Means clustering to find LGAs that are under-serviced by yoga/pilates studios

Results - Descriptive Statistics

Feature	Middle 50% of Values	Range (Min to Max)
median_fam_inc_weekly	1321.0 to 1814.5	1001.0 to 2765.0
pcent_women_employed	46.6 to 53.9	36.8 to 63.9
pcent_women_high_income	25.5 to 32.2	19.7 to 50.5
pcent_women_tertiary_educ	20.6 to 31.9	14.5 to 51.8

Large variability implies that these features may have an impact on prevalence of studios in an LGA.

Results - Correlation

Feature	Pearson Correlation
median_fam_inc_weekly	0.566
pcent_women_employed	0.377
pcent_women_high_income	0.601
pcent_women_tertiary_educ	0.634

Three features have a <u>moderate</u> correlation with prevalence of studios in an LGA. One feature has a <u>weak</u> correlation.

Results - Machine learning

Cluster number	0	1	2	3
mean median_fam_inc_weekly	1535	2633	1935	1282
mean pcent_women_employed	52.27	58.64	54.36	45.02
mean pcent_women_high_income	28.23	45.50	34.19	24.81
mean pcent_women_tertiary_educ	23.57	48.08	36.13	20.85
mean studios_per_10000_women	0.430	4.928	2.634	0.622

Observations

Clusters 0 and 3: Very few studios present; unfavourable demographics for a new studio

Cluster 1: Many studios present; favourable demographics but already well-serviced
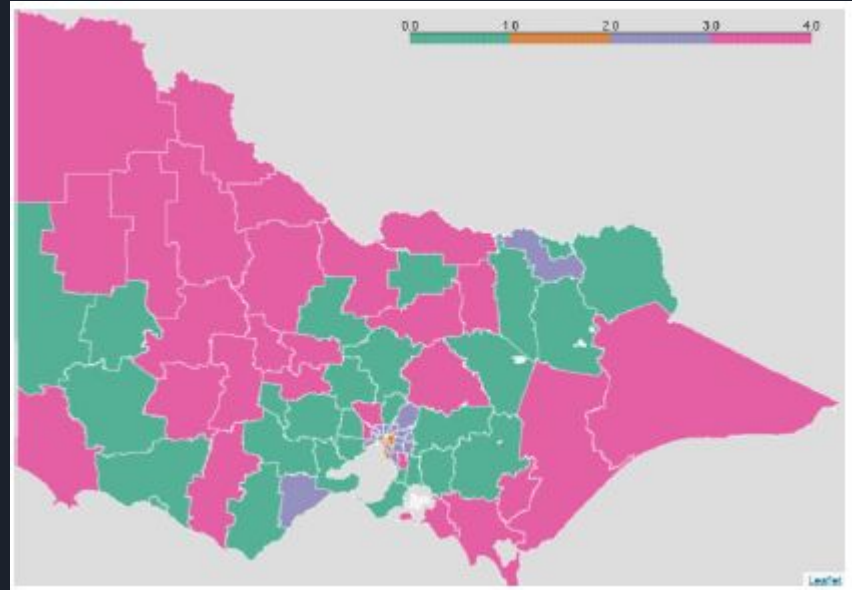
Cluster 2: Some studios present; favourable demographics for a new studio

Observations (cont…)

Green = Cluster 0

Pink = Cluster 3

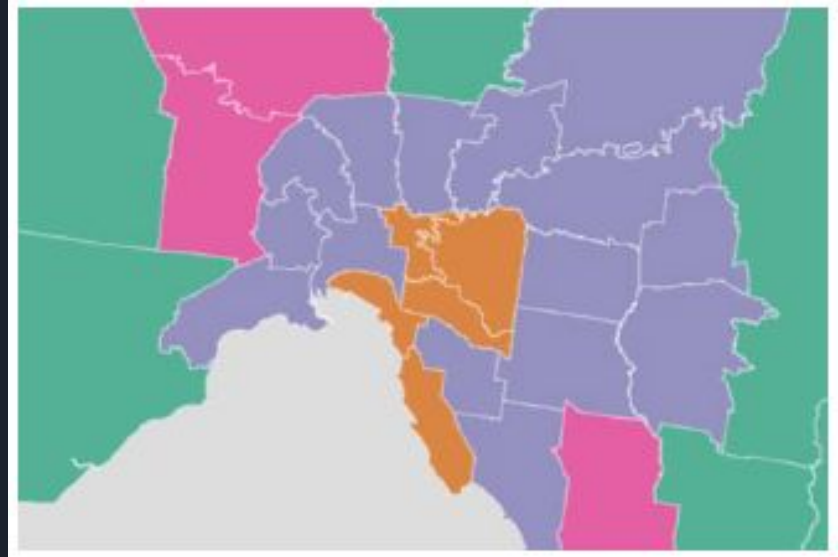Regional Victoria is not recommended for a new studio.

Observations

Brown = Cluster 1

Blue = Cluster 2

Well-serviced, wealthy, areas are all in central Melbourne.

Most of Cluster 2 surrounds Cluster 1.

Final recommendation

Cluster 2 has favourable demographics but is not well-serviced by yoga/pilates studios.

Therefore, recommend that any new studio is located in a Cluster 2 LGA.

Banyule	Knox	Moonee Valley
Darebin	Manningham	Moreland
Glen Eira	Maribyrnong	Nillumbik
Hobsons Bay	Maroondah	Surf Coast
Indigo	Melbourne	Whitehorse
Kingston	Monash	