

# ***Australian Population Characteristics & Venues Data Analysis for young professionals***

## **1. Description of the Capstone Project / Business Scope /**

Australia is one of the largest countries in Asia Pacific region with population over 25 million people and shy population density of 3 persons per square kilometer. The specific census organisation /aka stakeholder/ approached us to conduct a deeper investigation of how population of Australia is spread across the continent and which cities accommodate majority of its citizens. Once the biggest urban area is selected, the stakeholder wants to know which suburbs are the most popular with young professionals and what venues/categories are associated with those suburbs. This will allow local state departments to build smarter infrastructure planning strategies and create more venues if necessary.

The report will focus on statistical information from Worldometers website which presents live data information about population trending and major capital cities. The real estate data will indicate which suburbs are best for young professionals and how they are related to Foursquare API information.

## **2. Data Sources**

To provide the analysis that meet stakeholder's requirement, the following data sources have been taken into consideration:

2a. Live population statistics from the 'Worldometers' website:

<https://www.worldometers.info/world-population/australia-population/>

2b. Geo location information (latitude/longitude) from Google map

2c. Real estate statistics from 'OpenAgent' website that indicate urban suburbs popular with young professionals:

<https://www.openagent.com.au/blog/best-suburbs-to-live-in-sydney>

2d. **Foursquare API** to get the most common venues and its categories of given suburbs.

## **3. Data Injection and Preparation**

All data have been carefully analysed and assessed against the scope of the project.

3a. Live population statistics from 'Worldometers' have been scraped from the website and injected into multiple Pandas DataFrame within Jupyter notebook.

3b. Geo location information has been recovered from Goggle Map's 'Search' engine and injected into csv files accordingly. Those files then have been uploaded to Jupyter notebooks and modified into Pandas Data Frame.

3c. Real estate statistics from 'OpenAgent' site have been scraped from website and injected into manually created csv file with names of suburbs popular for young professionals and its geo locations. Then the csv file has been uploaded into Jupyter notebook and translated into Pandas Data Frame.

4c. Venus information has been extracted from the **Foursquare API** and injected into Pandas Data Frame.

#### 4. Methodology

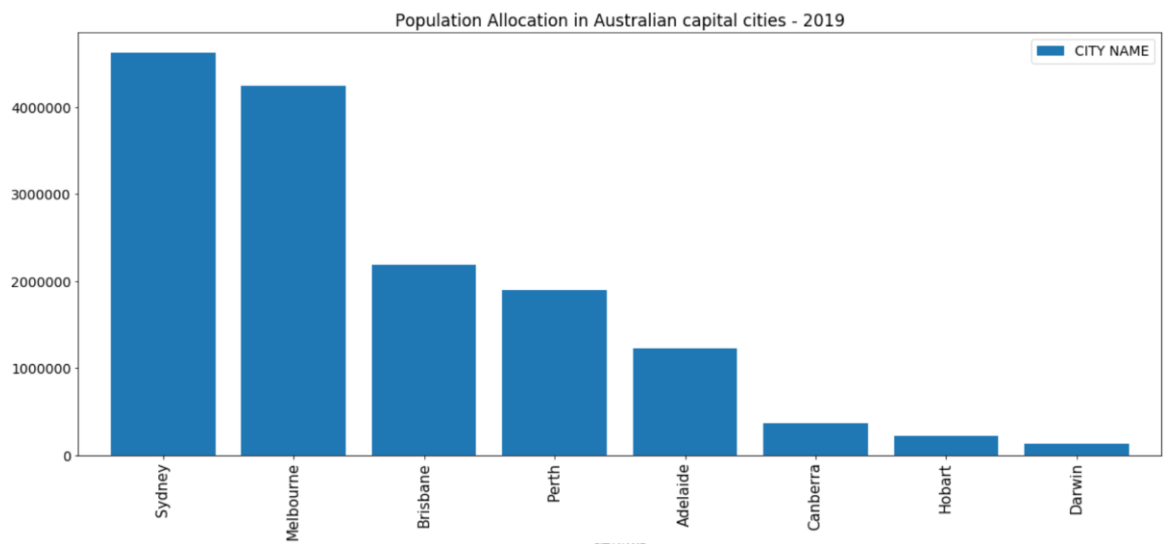
All collected data have been modified into Pandas-related Data Frames for further investigation. They have been correctly indexed and displayed in a form of tables to emphasise the business scope.

This table shows population data versus capital city name:

POPULATION

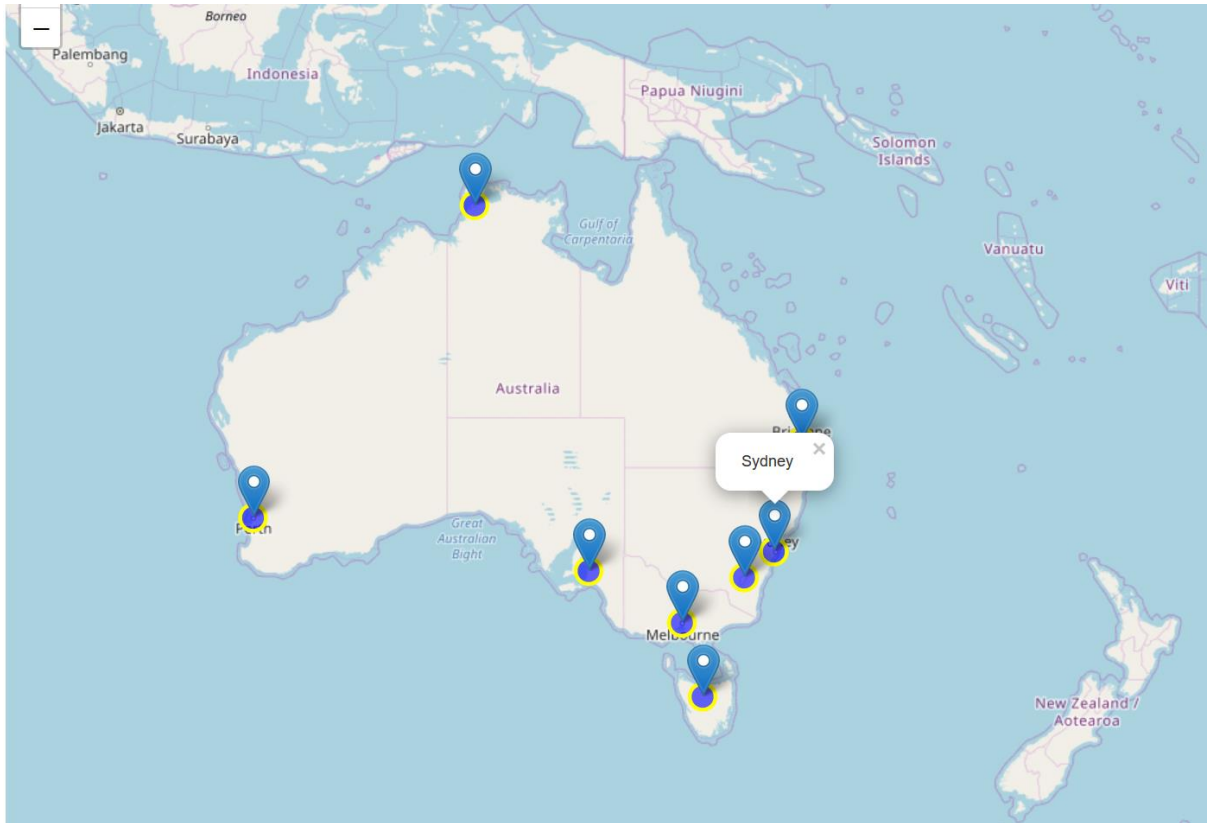
CITY NAME	
Sydney	4627345
Melbourne	4246375
Brisbane	2189878
Perth	1896548
Adelaide	1225235
Canberra	367752
Hobart	216656
Darwin	129062

This leads to the creation of the bar chart that displays the relation between capital cities:



Adding the latitude and longitude information to Pandas data frame resulted with the folium map of Australian capital cities:

	CITY	POPULATION	Latitude	Longitude
0	Sydney	4627345	-33.8675	151.2070
1	Melbourne	4246375	-37.8141	144.9633
2	Brisbane	2189878	-27.4710	153.0234
3	Perth	1896548	-31.9530	115.8575
4	Adelaide	1225235	-34.9286	138.6000
5	Canberra	367752	-35.3082	149.1242
6	Hobart	216656	-41.6400	146.3159
7	Darwin	129062	-12.4628	130.8418



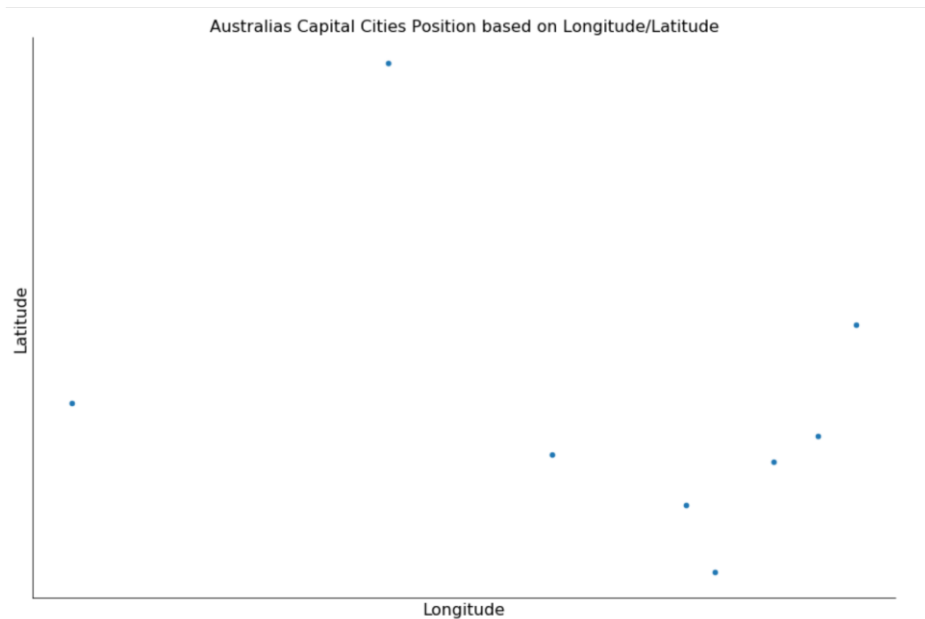
Additionally, the information about Sydney’s suburbs popular with young professionals has been added to the assessment.

	Latitude	Longitude
<b>Suburbs</b>		
<b>Redfern</b>	-33.89	151.200
<b>Waterloo</b>	-33.90	151.210
<b>Alexandria</b>	-33.90	151.200
<b>Glebe</b>	-33.88	151.180
<b>Newtown</b>	-33.90	151.170
<b>Ultimo</b>	-33.88	151.200
<b>Sydney_CBD</b>	-33.87	151.207

## 5. Results

The available population data indicated that population of Australia is mostly distributed around coastal cities which also are capital cities.

The scatter plot matches exactly the shape of the continent.



The most populated city of Australia is Sydney – 4.6 million - and is a home for many young professionals that select the CBD suburbs to live in.

The Foursquare API information clearly indicate that this CBD district attract young population because of the available venues and

3]:	Count	Gym	1	Thai Restaurant	6
categories		Gymnastics Gym	1	Theater	2
Art Gallery	1	History Museum	1	Trail	1
Asian Restaurant	1	Hostel	1	Vegetarian / Vegan Restaurant	2
Australian Restaurant	1	Hotel	1	Vietnamese Restaurant	2
BBQ Joint	1			Whisky Bar	1
Bakery	3	Ice Cream Shop	3	Wine Bar	1
Bar	6	Italian Restaurant	1	Yoga Studio	2
Bookstore	2	Japanese Restaurant	2		
Brewery	1	Lebanese Restaurant	1		
Café	16	Malay Restaurant	2		
Chinese Restaurant	1	Massage Studio	1		
Cocktail Bar	1	Museum	1		
Coffee Shop	8	Music Venue	1		
Comedy Club	1	Park	2		
Dessert Shop	1	Performing Arts Venue	1		
Dive Bar	1	Pizza Place	1		
Dumpling Restaurant	1	Pool	3		
Farmers Market	1	Pub	4		
Flea Market	1	Ramen Restaurant	1		
Fruit & Vegetable Store	1	Restaurant	1		
Garden	1	Science Museum	1		
		Spa	1		
		Speakeasy	1		

## 6. Discussion

As we can see from the data analysis, majority of the population of Australia lives in Urban areas (85.7%) which match the capital cities. The local population migration (including overseas migrants) also settle down in those urban metropolis.

```
Population of Australia - 2019: 25203198
Urban population in Australia - 2019: 85.7 %
Total Number of Migrants (Net) from 1955 till 2019 in Australia = 2056606
```

That poses much impact on capital cities resources and their infrastructures.

## 7. Conclusion

As a result, young professionals are heading to big cities, like Sydney to search for more opportunities and fulfilment of their career. Proximity to nightlife, good eating venues (confirmed by Foursquare APIs) and social aspects makes those big cities only viable option.