

|  |
| --- |
| **Analysis of Hospitality Trends in Melbourne 2002 - 2020** |
| Data Bootcamp Project 1 |
| **Authors:**  **Brandon Lee**  **Michael Dunne**  **Rosalina Indriani**  **Dhiren Patel** |

Contents

[Introduction 3](#_Toc99396079)

[Scope Of Work 3](#_Toc99396080)

[Limitation of Scope 3](#_Toc99396081)

[Record Of Work Done 4](#_Toc99396082)

[Obtain Data Set and Clean Data 4](#_Toc99396083)

[Methodology 4](#_Toc99396084)

[Results 4](#_Toc99396085)

[Industry Type Trends Over 2002 - 2020 5](#_Toc99396086)

[Methodology 5](#_Toc99396087)

[Results 5](#_Toc99396088)

[Figure 1 6](#_Toc99396089)

[Number of Businesses per Industry Per Suburb in 2020 7](#_Toc99396090)

[Methodology 7](#_Toc99396091)

[Results 7](#_Toc99396092)

[Table 1 8](#_Toc99396093)

[Figure 2 8](#_Toc99396094)

[Figure 3: Location of Businesses in Data Set 9](#_Toc99396095)

[Opening and Closing Trends per year 10](#_Toc99396096)

[Results 10](#_Toc99396097)

[Average Sitting per Suburbs & Industry over the timespan 14](#_Toc99396098)

[Methodology 14](#_Toc99396099)

[Results 14](#_Toc99396100)

[Seating Capacity over Industry 17](#_Toc99396101)

[Methodology 17](#_Toc99396102)

[Results 17](#_Toc99396103)

[Conclusion 22](#_Toc99396104)

# Introduction

The analysis looked at data obtained from data.melbourne.vic.gov.au of Melbourne hospitality industry’s seating capacity.

The objective of this analysis is to gain insights from the data to gain an understanding of Melbourne’s hospitality industry.

This analysis could be used by:

* Prospective hospitality owners/managers, eager to better understand the Melbourne Market.
* Investors in the hospitality industry to gain insights as to market saturations, and trends in the Melbourne market.
* Property buyers who would like to gain knowledge of the hospitality services available within the areas of Melbourne.

## Scope Of Work

The analysis leverages off the data provided through the following link:

https://data.melbourne.vic.gov.au/Business/Cafes-and-restaurants-with-seating-capacity/xt2y-tnn9

The scope includes:

* Review the data to gain an understanding of the data set;
* Clean the data to remove any records with missing fields;
* Conduct various analysis over the data to gain insights and trends into the hospitality industry within the Melbourne metro area.

### Limitation of Scope

The following limitations were noted:

* Data set accumulated over the years from 2002 to 2020
* Data only included suburbs within Melbourne’s CBD
* Of the 69 different industries, Cafés and Restaurants comprised 74.8% of the Data
* Furthermore, 4/69 categories comprised 96.49% of the data, the remaining 65/69 needing to be consolidated

# Record Of Work Done

## Obtain Data Set and Clean Data

The objective of this step is to obtain and clean the data set ready for analysis.

### Methodology

1. Download the data from https://data.melbourne.vic.gov.au/Business/Cafes-and-restaurants-with-seating-capacity/xt2y-tnn9
2. Load the data into Python
3. Gain an understanding of the data:
   1. Identify the variety of industry types within the dataset
   2. Identify the number of unique address (by the business address) within the dataset
   3. Identify the number of unique location within the dataset, which should match with the business address
   4. Identify the number of unique location by "Building address" within the dataset.
4. Consolidate the industry type into 5 categories:
   1. The top 4 most frequent "Industry (ANZSIC4) description" category; and
   2. All other categories will be re-categorised into "Others".
5. Drop any rows with missing data.

### Results

* There is a total of:
  + 69 different industry types in the dataset.
  + 6,917 unique addresses and locations
* The top 4 Industry types were identified as follows:



All other industries have been recategorized into “Others”.

* The raw data contained 54,052 records. Once the data have been cleaned of those records with missing fields, 53,990 records remained (i.e. 62 records contained missing fields and were excluded from the analysis.

## Industry Type Trends Over 2002 - 2020

The objective of this analysis is to review the number of businesses that existed within each of the 5 industry types. This will give insight as to:

* The industry that is growing/waning through time; and
* The most popular industry across time.

It should be noted that each business may have more than one record within the data set. This is due to the data recording 2 separate entries per business to account for indoor seating and outdoor seating. Consequently, the data must be further cleansed to remove the duplication of businesses.

### Methodology

1. Remove duplicated business:
   1. Create a key field based on the year and the business address (this will help remove duplication due to “outdoor”/”indoor” listing, but not due to “year”)
   2. Drop the columns “Number of Seats” and “Seating Type”. These are irrelevant to the analysis.
   3. Conduct a de-duplication of data frame to remove duplicate records.
2. Group data by the 5 industry groups.
3. Create a line graph that shows the number of business that exist in each industry group over time.
   1. From the previous data set, use the .getgroup() function to extract the data for each industry group.
   2. Use the plt.plot function to plot all industry data into a line graph.

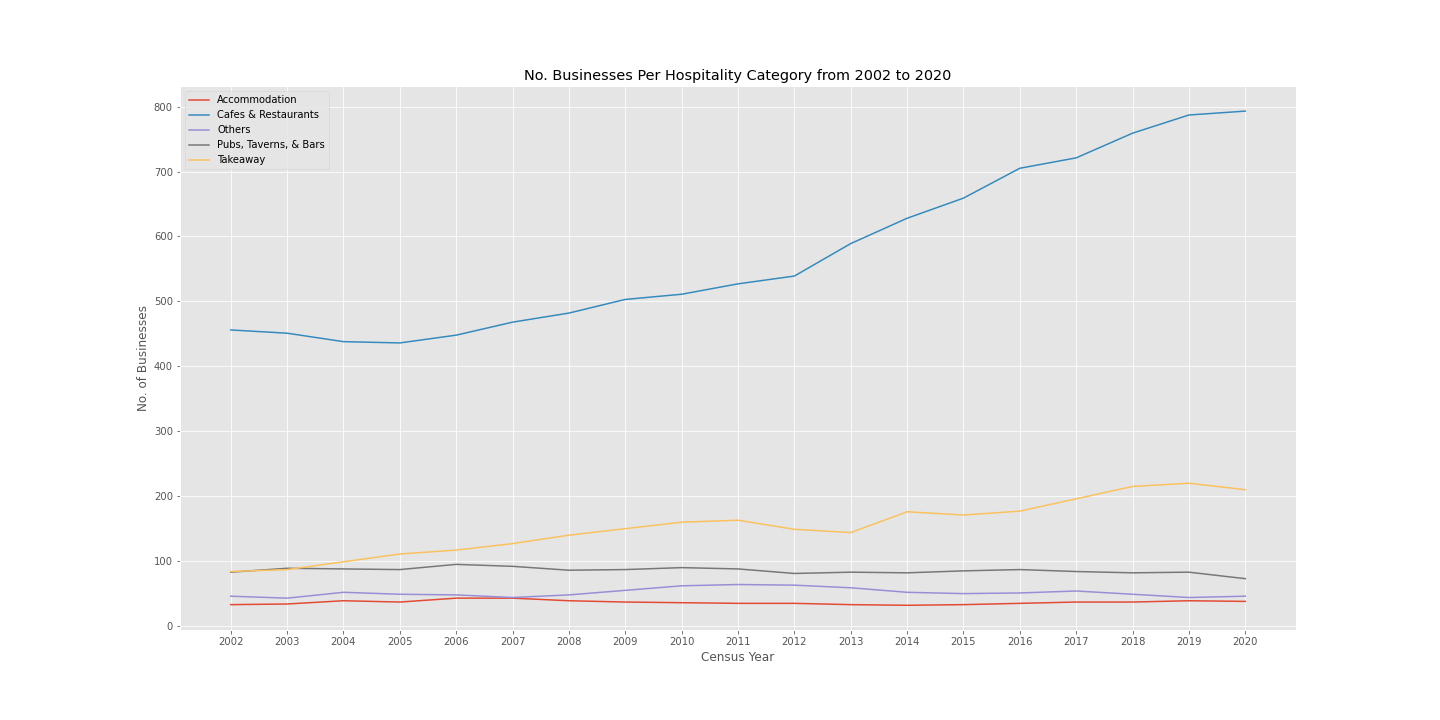
### Results

There were 17,095 records once the data has been cleansed from duplicated business.

Referring to Figure 1 (next page), the following observations were noted:

* Cafes & Restaurants has always been a dominant hospitality industry in Melbourne.
* There was a significant increase in the number of Cafes & Restaurants in Melbourne from 2012 onwards.
* Takeaway industry is the next dominant industry and has been steadily increasing over time.
* There appears to be a slight decrease in the number of Pubs, Taverns, and Bars over time. This may indicate that patrons who visits Melbourne do not have a focus on visiting Pubs, Taverns, and Bars establishments.
* It was surprising to see that there was a decline in Takeaway industry from 2019 to 2020. Whilst COVID would be the logical explanation for this dip, it was expected that Takeaway establishment would have flourished during this time as residents were in lockdown and required takeaway food services. That said, there was a large out-flux of residence from the Melbourne metropolitan area; thus the demand for such service would have decreased significantly.

#### Figure 1



## Number of Businesses per Industry Per Suburb in 2020

The objective of this analysis is to gain insight as to the number of businesses that existed in each industry in each suburb of inner Melbourne.

This analysis will leverage of the de-duplicated data set from the “Industry Type Trends Over 2002 – 2020” analysis (hence referred to as the “De-Duped Data”).

### Methodology

1. Using the De-Duped Data, group the data by “Census year” column.
2. Extract the 2020 Census year column into a new data frame.
3. Identify the number of unique suburbs in the data frame.
4. Create a pivot table that identifies the number of businesses that exists per category type in each suburb
5. Create a bar graph to illustrate the breakdown of industry type per suburb in the 2020 year
6. Using Google Map APIs, map out the businesses contained in the data set.

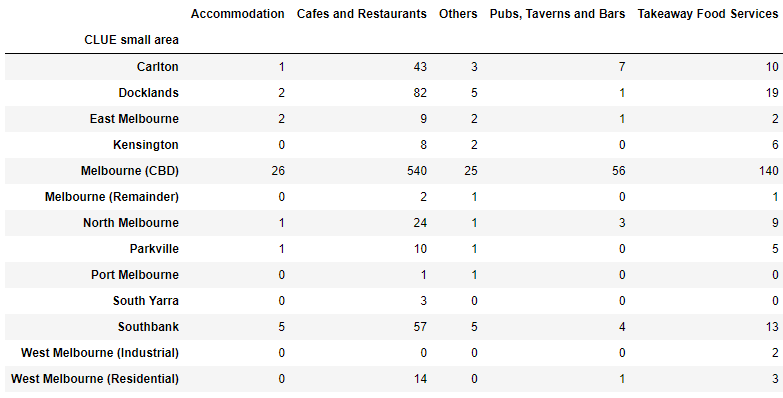
### Results

The breakdown of the number of businesses per industry type per suburb in 2020 is represented in table 1, figure 2, and figure 3 in the following page.

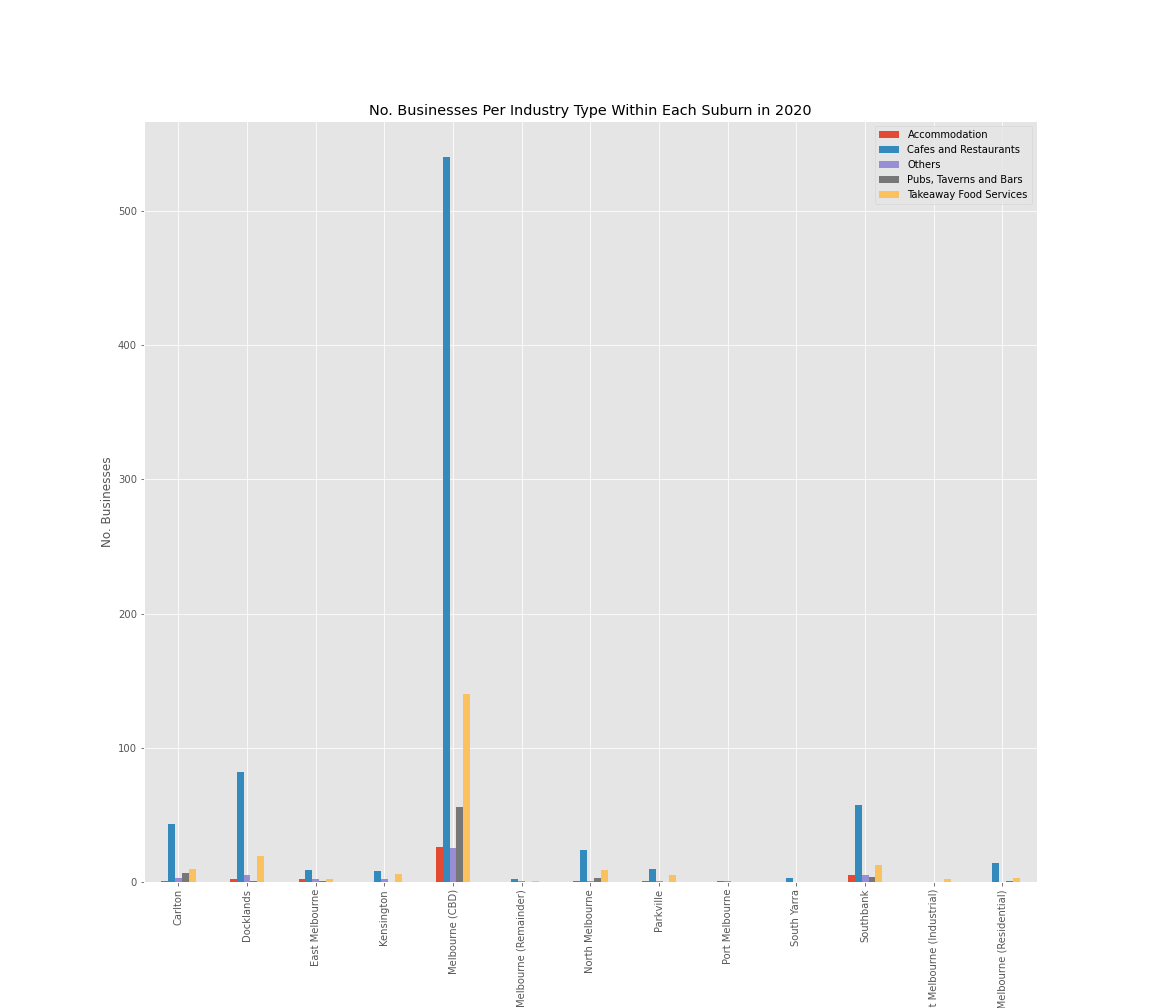
The following observations were made:

* Out of the 7 suburbs, Melbourne CBD holds the most hospitality businesses.
* Predominantly, Cafes and Restaurant businesses reside in Melbourne CBD. Those who would like to open a Café or Restaurant in Melbourne CBD should be aware that the market is already saturated and will face fierce competition.
* West Melbourne (Industrial) has the least number of hospitality businesses. This is to be expected as the area is an industrial area and would not expect a vast range of hospitality services.
  + Businesses who are intending to trade from this area should be aware that their employees will not have close access to many hospitality services.
  + Potentially a pub or tavern that concentrates on lunch service would do well in this area as there is not a lot of competition.
* Port Melbourne and South Yarra were 2 suburbs that had a surprisingly low number of hospitality services in the area.
  + This area is a heavily residential area; consequently, there may not be commercial infrastructure readily available to open any kind of hospitality industry.
  + It would be worth investigating opening a Café or takeaway business in these area as there is little competition that services the suburb.

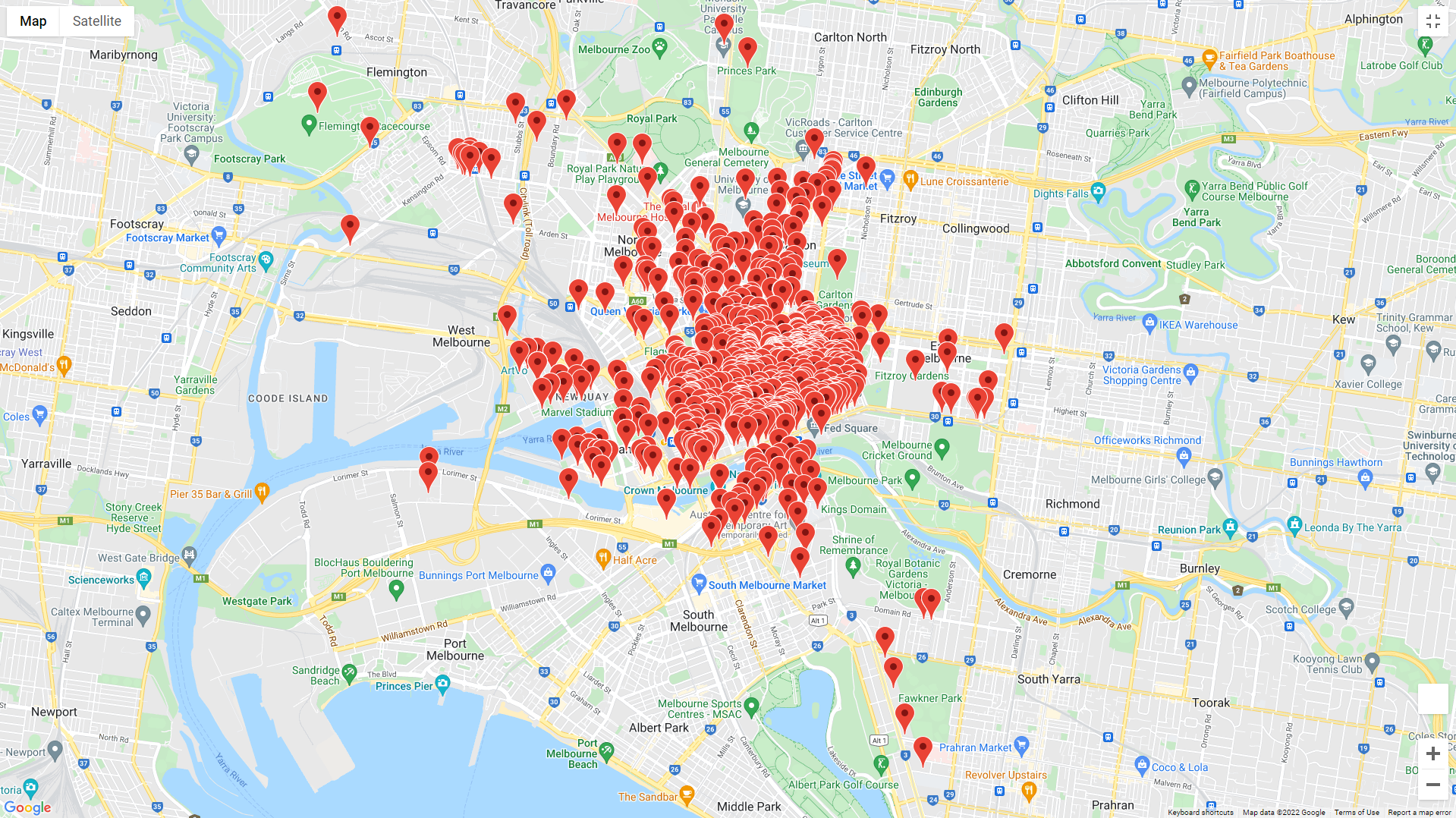
#### Table 1



#### Figure 2



#### Figure 3: Location of Businesses in Data Set



## Opening and Closing Trends per year

The purpose of this analysis is to gain a better understanding of opening and closing trends over the timespan of the dataset, with particular focus being cantered towards the last two years (i.e., 2019 to 2020).

It has been well documented Hospitality was hit harder than any other industry in the wake of the COVID-19 pandemic, which is why this time frame was highlighted during the data investigation process.

Methodology

1. Drop any rows with missing data, creating a cleaner dataframe
2. Using the clean dataframe, Index([ ]) the 14 columns from the dataset.
3. After extracting these columns, groupby() the particular columns you wish to use for your graph (stacked bar graph).   
   For this example *“Year”*, *“Area”* & *“Industry Count*”, were the three highlighted.
4. A pivot table of: 247 rows x 3 columns is created, listing the count of each of the  
   13 industries across each year (2002 to 2020), of the 19 year time frame   
   (13 x 19 = 247 rows)
5. Using this data, we are able to create a stacked bar graph, illustrating the count of each industry for each area, across every year of the dataset.

### Results

The breakdown of the overall count of industry (hospitality) per area, across each year is represented in Figure #1.

*Some of the key observations from this bar graph include the following:*

* The count of Industry increased by 89.8% from 2002 to 2020.

|  |  |
| --- | --- |
| **Year** | **Count of Industry** |
| 2002 | 1666 |
| 2020 | 3162 |

* Of all the 13 areas. only West Melbourne (Industrial) declined in count from 2002 to 2020.
* Docklands was the fastest growing area, increasing by 1514%

|  |  |
| --- | --- |
| **Year** | **Count of Industry** |
| 2002 | 22 |
| 2020 | 355 |

* A major component of this was the construction of Docklands Stadium in the year 2000 (clear example of gentrification), along with the redevelopment of Southern Cross station, prior to the 2006 Commonwealth Games.
* Port Melbourne had the second highest count variance, increasing by 1033%.

|  |  |
| --- | --- |
| **Year** | **Count of Industry** |
| 2002 | 3 |
| 2020 | 31 |

* Similar to Docklands, a refurbishment prior to the 2006 Commonwealth Games was a key factor for this increase. Additionally, the construction of sporting arenas, shopping centres, and private shipping ports has also played a significant role.
* South Yarra was not only surprisingly stagnant, but also illustrated very low counts throughout the dataset, only increasing by 50% through the time frame.

|  |  |
| --- | --- |
| **Year** | **Count of Industry** |
| 2002 | 10 |
| 2020 | 15 |

* ***Figure #1: Count of Industry from 2002 to 2020***   
  Chart, bar chart

  Description automatically generated*Effects of Covid: 2019 to 2020*
* As expected, 9/13 areas showed declines in industry counts, following 18 consecutive years of relatively steady growth. Best illustrated in the Melbourne CBD line plot below.

*Figure #2 Melbourne CBD line plot*

**Chart, line chart

Description automatically generated**- 3/13 areas (Kensington, Port Melbourne, & South Yara) remained stagnant

Count of Area

|  |  |  |  |
| --- | --- | --- | --- |
| **Year** | **Kensington** | **Port Melbourne** | **South Yarra** |
| 2019 | 61 | 31 | 15 |
| 2020 | 61 | 31 | 15 |

Docklands surprisingly continued to increase its industry count by +13 (3.8%), from 2019 to 2020.

The majority of this increase being attributed to the opening of +12 Cafes & Restaurants.

Chart, line chart

Description automatically generated

## Average Sitting per Suburbs & Industry over the timespan

The objective of this analysis is to see insight of the average number of seats that existed in per suburb or per industry of inner Melbourne.

### Methodology

1. Drop some rows with missing data, combine some industry categories into one to make it cleaner.
2. Using Clean Dataset, Groupby requires columns (Census Year, CLUE Small Area, Recategorized industry, Seating type) and finds average value of Number of seats column.
3. Create a table of census year, CLUE Small area and Average value of Number of seats to analyze Average Sitting per Suburb over the timespan!
4. Create a table of census year, Recategorized industry and Average value of Number of seats to analysis Average Sitting per industry over the timespan!
5. Create bar graphs using above analysed data.

### Results

The analysed data in graph 1 shows average number of seats as per suburb over the time span.

* It shows that Kensington is ahead of any other areas in average number of seats between 2002 – 2020 regardless of less number of cafes, restaurants, pubs and accommodation as because majority Venues are bigger and with larger seating capacity such as racecourse, community centre, fine dine restaurants.
* Whereas West Melbourne (Industrial) is very less sitting capacity as because it is industrial area so there is no such venues like hotel, restaurant or accommodation except few takeaway shops and cafes which are not with larger in seating capacity and they actually don’t need it really because people are going there just for breakfast and lunch between their break.
* After Kensington, Melbourne (Remainder) is 2nd highest in sitting capacity which has more fine dine restaurants, cafes and pubs than any other areas which shows that people actually like to go for coffee, lunch , dinner or for drink in larger groups.

The analysed data in graph 2 shows average number of seats as per Industry over the time span.

* It shows that ‘other’ has more seating capacity than any other industry because there are 65 different industries combined into this category to make data look cleaner. Which includes larger venues like functions room, racecourse etc.
* In one particular category, pubs and bars have more sitting capacity than any other categories as per analysis because people would like to go to pubs and bars in larger groups such as with friends or with a group of colleagues for a party.

Table

Description automatically generated

**Graph 1**

Chart, bar chart, histogram

Description automatically generated

Table

Description automatically generated

**Graph 2**

Chart, bar chart

Description automatically generated

## Seating Capacity over Industry

The objective of this analysis is to obtain an insight as to trends over industry in comparison with the seating preference.

### Methodology

1. Using the clean dataframe, group the data by “Census year” column.
2. change the data type of census year from integer to object
3. Breakdown the re-categorisation industry per year, per seating type.
4. Create a pivot table that identifies the number of seating per year: by suburb and by the industry
5. Create a bar graph to illustrate the breakdown
6. Using Google Map APIs, map out the top 10 businesses with largest seating capacity

### Results

The breakdown of the seating capacity over the years are represented in figures below.

based on the finding, below are the summary of observations:

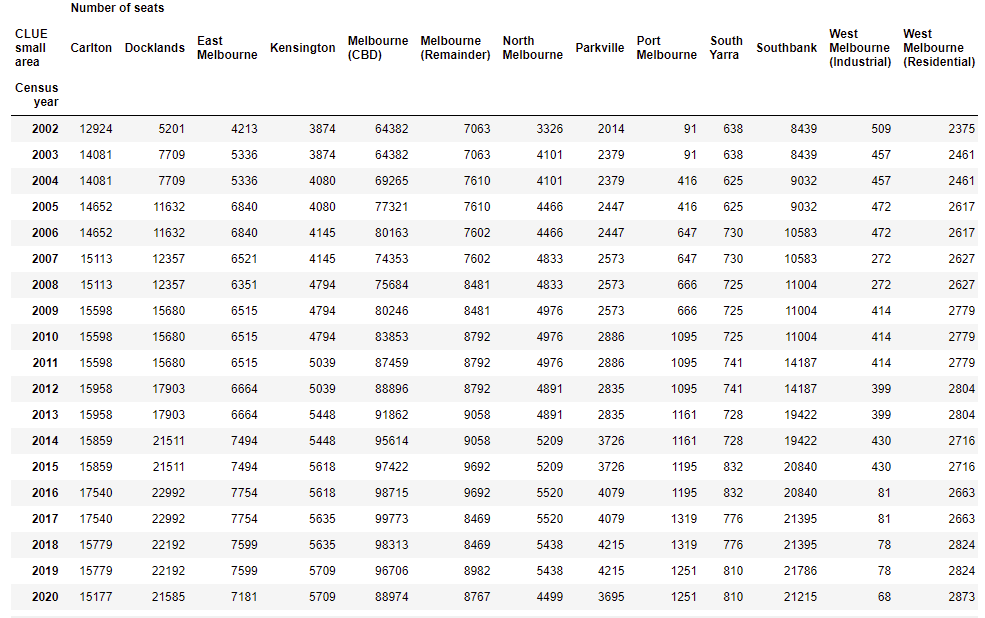
* Government regulation changes to accommodate indoor and outdoor settings. There's a certain guideline that business owners have to follow for outdoor sitting. This might hinder the business owner from putting more seating outside. In addition to unpredicted weather in Melbourne, indoor seating still preferred compared to the alfresco dining.
* Majority of the business still utilise the indoor seating, even though there's a slight increase of al-fresco or outdoor dining over the years. This trend most likely will continue growing in steady pace, especially after Covid period

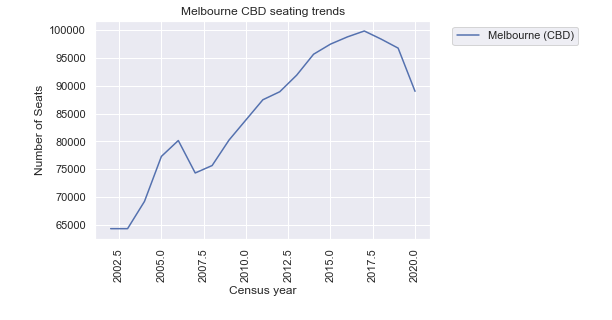
**table 1**

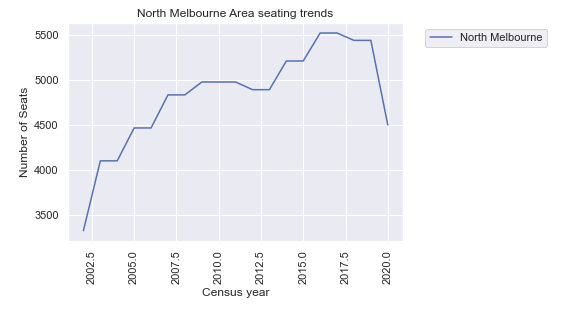


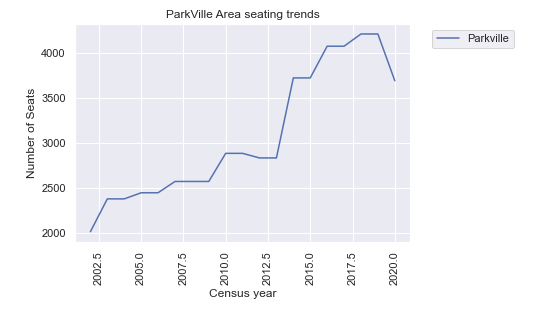
* For suburbs, there's been a steady increase over the years.
* Interestingly, between 2015-2016, there’s substantial increase for Carlton area, approximately 10% where total seat was close to 16k to 17.5K and interestingly, there’s massive decrease in West Melbourne (industrial) area in the same year around 80% from 400 total seating to 80 seating.
* This number returned back to 15K from 2017-2018 for Carlton, whilst the West Melbourne Industrial area kept decreasing in terms of the total seating up to 2020.
* For the CBD Melbourne area, the seating capacity reached its peak in 2016-2017 and slowly decreased till 2020. This could be because the CBD area is already overcrowded and has less room for growth.
* Overall growth for seating in other Metropolitan Melbourne areas is steady but not significant to a couple of areas such as: Docklands, Kensington, Port Melbourne and Southbank.
* Overall decrease, significantly in area of West Melbourne Industry, followed by North Melbourne and Parkville which showing a decrease approximately for 10% from 2019-2020.

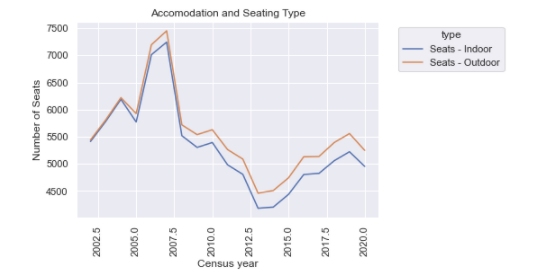
**table 2**

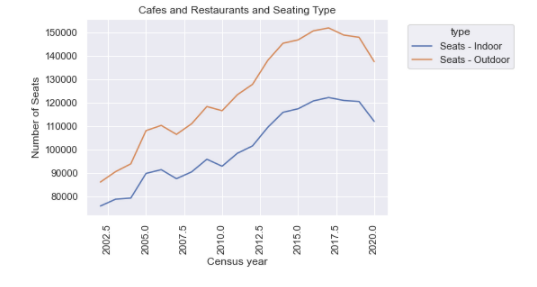


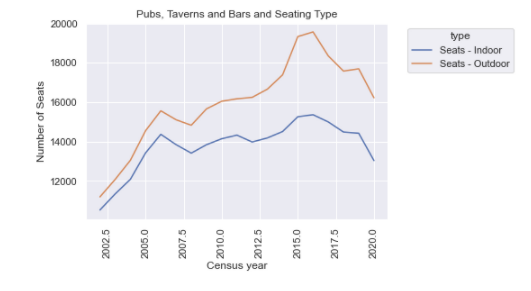




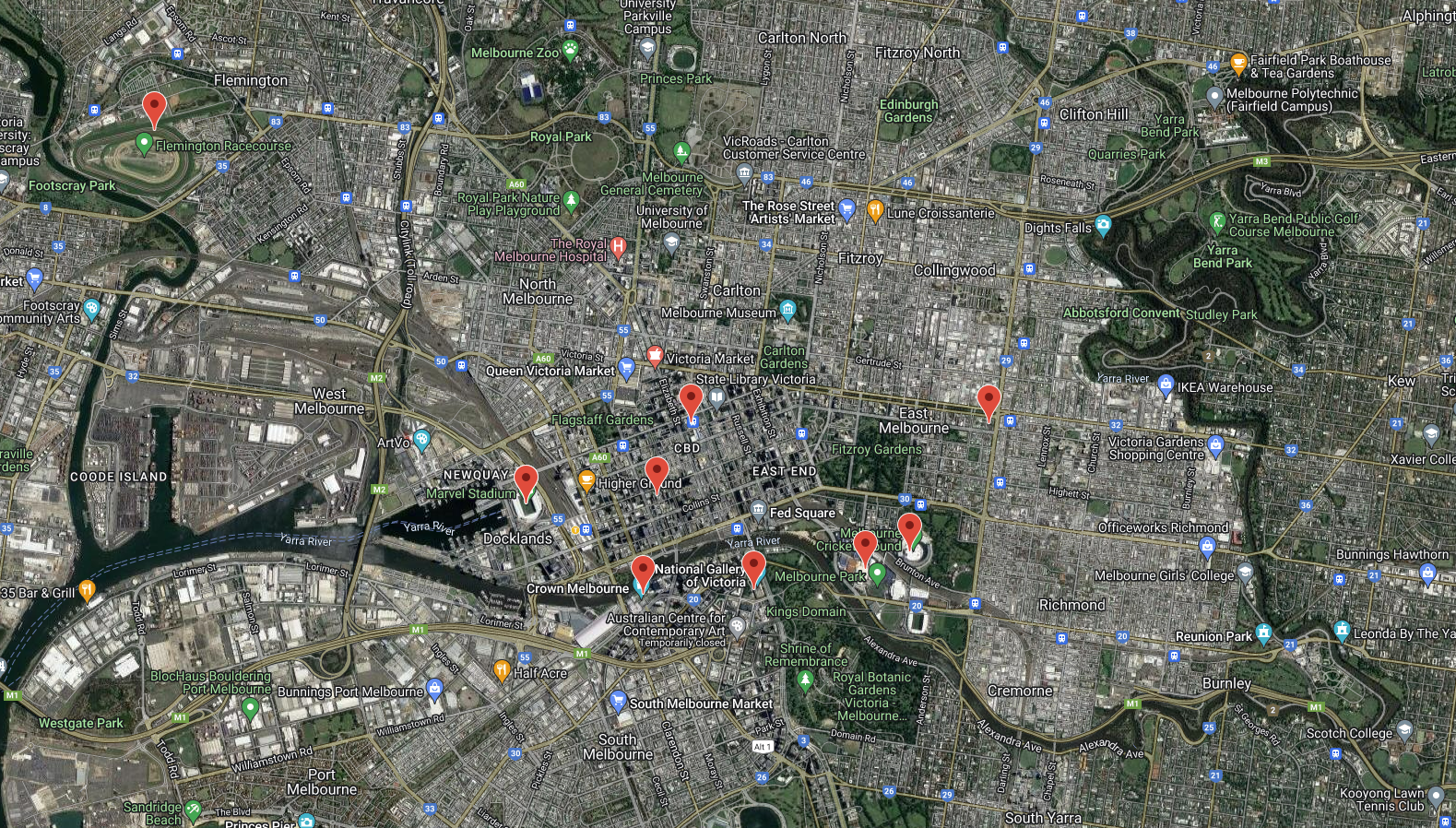


* For the industry and trends over seating type and its number, interestingly, there’s massive growth during 2005-2007 for accommodation industry type which when it’s quick to reach its peak, over the following years also showing quick dips. The number slowly increased steadily from 2012 to 2019. 
* Cafes and Restaurant, where the majority of the business sitting, showing steady increase, hand in hand for indoor seating as well as outdoor sitting even though from year 2017, it showed decrease in overall seating for both arrangement.





Top 9 areas in Metro Melbourne where this showing largest seating capacity.



These locations include: Rod Laver Arena Melbourne, Flemington Racecourse, MCG, Marvel Stadium, NGV, Crown Entertainment, RACV, Catholic Leadership Centre and Melbourne Central.

# Conclusion

Based on the analysis results, the following salient points were noted:

* There was a noticed decrease in the number of seating available and number of businesses opened from the 2019 to the 2020 year. This would largely be caused by the world wide pandemic as Melbourne enters into its various lockdown restrictions.
* The data reflects Melbourne’s Café culture with a large majority of the businesses falling into the Café’s and Restaurant industry and it resides mainly in the Melbourne CBD area.
* Changing over government regulations relating to seatings (specifically in CBD area) effecting the overall growth of the restaurant and café industry. This includes spacing and furniture issues.