

Determining a Suitably Located Cafe in Sydney, Australia

1. Introduction

Sydney is the most populous city in the state of New South Wales, as well as in Australia. According to the 2016 Census, over 4.8 million people call Sydney home. Sydney is a global city of commerce, being the home location of ASX, one of the world's foremost stock exchanges, as well as being a regional hub for many companies in business, including investment banks, consultancies and accounting firms. In addition, Sydney is the most visited location in Australia, with no shortage of tourist attractions such as Sydney Opera House, Sydney Harbour Bridge and Bondi Beach.

As a bustling global city with a raft of economic activity, both white-collar and blue-collar, throughout the CBD (Central Business District), whether it is a multinational chain or a local 'mum-and-dad' business, it can be an attractive place to set up business to tap into the many thousands of workers who commute to the CBD every working day and the tourists who come along to experience life in Sydney. Conversely, it can often mean having extreme competition in competing for the many potential customers, and unsurprisingly, businesses must prepare well to survive in cutthroat competition – even popular eateries and cafes have closed down.

Coffee shops, and by extension, cafes, are big business across all Australian metropolitan areas. There appears to be a distinct 'café culture' around Sydney, with a distinctive range of coffee styles not found elsewhere around the world.

2. Business Problem

Given its population and its vibrant business culture, many cafes have opened up around Sydney. However, with strong competition and variety of different kinds of cuisines and offering arounds, opening up a new café in Sydney is no easy task. Location is a prime factor in the success or failure of an establishment. Thereby, for this project, the problem is defined to be "Where should a café be opened in the Sydney CBD and surrounds?"

The key stakeholders in this problem are potential café owners, as well as existing owners who could be reconsidering their business strategy, or are looking to relocate to their business to tap in on new customers. Landlords and developers of multi-use properties may also have interest.

Location is paramount to success in order to tap into the most possible number of customers, whether it during the morning rush to work, a coffee break in the morning, for lunch hour, or if opening up late to cater for the tourists. It also goes without saying that a café in the middle of the city is a very expensive investment. The stakes may be high but with tens, if not hundreds, of thousands of people living and working around Sydney, there is huge potential to attract customers should the business is executed well. It is therefore imperative for owners to carefully consider location and show care for this problem to maximise their potential in taking a slice of this billion-dollar industry.

Other factors that may affect suitability of location, such as ongoing rental costs and structural integrity of the building the potential café may be located, are not considered in this project.

3. Data

Data for this project was conveniently obtained from an open-source database on a website maintained by Matthew Proctor. Here, a .csv file was downloaded that includes a listing of postcodes across all of Australia – some 3200 unique postcodes. The format was conveniently organized, such that it was easily scraped and converted into a pandas dataframe. Clearly the required data was evident – ‘postcode’, ‘locality’ (equivalent to neighbourhood in North American English terms), ‘latitude’ and ‘longitude’.

Foursquare API and other relevant Python libraries such as ‘geocoder’ were also used, which tapped into latitude and longitude to enable Foursquare data to be leveraged.

Other columns of data, such as ‘region’ and ‘Lat_precise’, while being invaluable to an app developer, were a surplus to this project, and thereby were omitted from the project. Thereby, a dataframe consisting of ‘postcode’, ‘locality’, ‘latitude’ and ‘longitude’ suffices.

Obviously we only need to consider data around Sydney CBD. This involves restricting data based on postcode. Sydney CBD itself has postcode 2000, and neighbourhoods immediately surrounding the CBD have postcodes from 2000 to 2010. However, there are some neighbourhoods that are clearly much further away from the CBD than others, but have a postcode within the aforementioned range, whereas there are some neighbourhoods that are of closer range but with a postcode not within the range. For simplicity, the project will consider only postcodes 2000 to 2025. From these 12 postcodes alone, there are over 1200 venues to be analysed in this project.

No other datasets are used in this project.