Problem Description

Overview

I have an amazing opportunity from my company that will set up its primary IT services operations in New Castle, New South Wales. I have been asked to to lead this effort of setting up this operations and ensure we identify the right location for the office space and also a great neighborhood for the pool of talent we will be hiring.

As a part of this major project I will be taking you through my journey of exploration and give you an amazing glimpse of Newcastle and its surroundings. Are you ready to come along and find a great place to work and live in Newcastle?

Our company also intends to use this study to share the details with our prospective employees who we are currently recruiting. It is very critical to choose the right location for the office as it will serve as a hub for our operations in the New South Wales region. It is important to be cost effective but also be closer to bigger city like Sydney which is the capital of New South Wales to ensure we attract the right pool of talent. Our company has major offices in Adelaide, South Australia and Brisbane, Queensland and is now planning to expand in New Castle, New South Wales as part of a major planned growth and expansion opportunities due to the increased revenue projections of our service offerings.

Audience

Primary Audience will be our Corporate and Real Estate Strategy of our company who will be evaluating the final results and recommendations of this study to determine the site and office space for our New South Wales office in New Castle.

Data Description

For the data I found a JSON file of australian cities that I downloaded from the link below and processed the information into a pandas dataframe.

!wget -q -O 'australia_cities.json'
https://simplemaps.com/static/data/country-cities/au/au.json
print('Data downloaded!')

JSON Data Structure

▼ **root:** *[] 219 items*
  ▼ **0:**
      city: "Sydney"
      admin: "New South Wales"
      country: "Australia"
      population_proper: "3641422"
      iso2: "AU"
      capital: "admin"
      lat: "-33.861481"
      lng: "151.205475"
      population: "4630000"

JSON Data Structure processed into a Pandas Dataframe

	City	Admin	Country	Population Proper	ISO2	Capital	Lat	Lng	Population
0	Sydney	New South Wales	Australia	3641422	AU	admin	-33.861481	151.205475	4630000
1	South Melbourne	Victoria	Australia	93625	AU		-37.833333	144.966667	4170000
2	Brisbane	Queensland	Australia	926353	AU	admin	-27.47101	153.024292	1860000
3	Perth	Western Australia	Australia	880216	AU	admin	-31.95224	115.861397	1532000
4	Adelaide	South Australia	Australia	836354	AU	admin	-34.928661	138.598633	1145000
5	Newcastle	New South Wales	Australia	497955	AU		-32.927792	151.784485	1134616
6	Gold Coast	Queensland	Australia	332249	AU		-28.00029	153.430878	527660
7	Cranbourne	Victoria	Australia	39419	AU		-38.1	145.283333	460491
8	Canberra	Australian Capital Territory	Australia	234032	AU	primary	-35.27603	149.13435	327700
9	Wollongong	New South Wales	Australia	141725	AU		-34.424	150.893448	260914
10	Geelong	Victoria	Australia	137681	AU		-38.153461	144.358093	160991
11	Cairns	Queensland	Australia	109989	AU		-16.92304	145.766251	154225
12	Townsville	Queensland	Australia	119470	AU		-19.26639	146.805695	138954
13	Albury	New South Wales	Australia	32810	AU		-36.074823	146.924006	104258

Newcastle Neighborhoods

However I needed more granular information of Newcastle and its neighborhoods that I was not able to locate a CSV or JSON file. So I built my own dataset of Newcastle Neighborhoods by looking up Wikipedia and getting each of it neighborhoods coordinates-

https://en.wikipedia.org/wiki/List_of_suburbs_in_Greater_Newcastle,_New_South_Wales

City of Newcastle suburbs [edit]

- Adamstown
- Adamstown Heights[n 1]
- Bar Beach
- Beresfield
- Birmingham Gardens
- Black Hill[n 2]
- Broadmeadow
- Callaghan (University)
- Carrington
- Cooks Hill
- Elermore Vale
- Fletcher
- Georgetown

- Hamilton
- Hamilton East
- Hamilton North
- Hamilton South
- Hexham
- The Hill
- Islington
- Jesmond
- The Junction
- Kooragang
- Kotara
- Lambton
- Lenaghan

- Maryland
- Maryville
- Mayfield
- Mayfield East
- Mayfield North
- Mayfield West
- Merewether
- Merewether Heights
- Minmi
- New Lambton
- New Lambton Heights[n 1]
- Newcastle
- Newcastle East

- Newcastle West
- North Lambton
- Rankin Park[n 1]
- Sandgate
- Shortland
- Stockton
- Tarro
- Tighes Hill
- Wallsend
- Warabrook
- Waratah
- Waratah West
- Wickham

df_newcastle_neighborhood = pd.read_csv('newcastle_neighborhood.csv')

df_newcastle_neighborhood

	Neighborhood	City	Lat	Lng
0	Adamstown	Newcastle	-32.9344	151.7258
1	Adamstown Heights	Newcastle	-32.9477	151.7152
2	Bar Beach	Newcastle	-32.9400	151.7678
3	Beresfield	Newcastle	-32.7990	151.6350
4	Birmingham Gardens	Newcastle	-32.8937	32.8937
5	Black Hill	Newcastle	-32.8400	151.6120
6	Broadmeadow	Newcastle	-32.9252	151.7342
7	Callaghan	Newcastle	-32.8900	151.7010
8	Carrington	Newcastle	-32.9150	151.7680

Use of Foursquare data

Foursquare API for Newcastle, New South Wales and its neighborhoods will be used to analyze nearby venues, get the coordinates of the venues to get more information of the neighborhood areas as show below.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Adamstown	-32.9344	151.7258	Adams Ribs & Pizza	-32.934490	151.729330	Pizza Place
1	Adamstown	-32.9344	151.7258	Cupcake Espresso Adamstown	-32.935927	151.725734	Cupcake Shop
2	Adamstown	-32.9344	151.7258	Domino's Pizza	-32.934716	151.725785	Pizza Place
3	Adamstown	-32.9344	151.7258	7 Eleven	-32.934659	151.729345	Convenience Store
4	Adamstown	-32.9344	151.7258	The Nags Head Hotel	-32.935290	151.725370	Pub
5	Adamstown	-32.9344	151.7258	The Gates Hotel	-32.933370	151.721400	Beer Garden
6	Adamstown Heights	-32.9477	151.7152	Doors 4U Garage Doors	-32.948633	151.710709	Home Service
7	Adamstown Heights	-32.9477	151.7152	Mimco Kotara	-32.948681	151.719802	Jewelry Store
8	Bar Beach	-32.9400	151.7678	Bar Beach	-32.941137	151.768059	Beach

Clustering will then be used to determining the most common venues for each of the neighborhoods to then determine what will be the most suitable area for the New South Wales HQ that will then be most suitable from a business perspective and also be liked by the employees who will be hired and working in the location.

The recommendations will be submitted to the Senior Leadership Team in the company to review and make the final decision. Come along with me for this beautiful explorative journey and hope you enjoy it as much as I do!