**Comparison of two large cities in South Africa**

1. **Introduction**

1.1 Background

The City of Cape Town is the legislative capital of South Africa. It is oftenly, called the Mother city and is located in the Western Cape province of South Africa. The is known for its natural setting as well as its landmarks such as Table Mountain and Cape Point. Cape Town is said to be the home to 64% of the Western Cape's population making it one of the most multicultural cities in the world. Cape town is a popular destination for immigrants, expatriates as well as tourists alike. The city boosts a population of 3 740 026 people, according to the South African National Census of 2011. Although, the City of Cape Town enjoys the support of many immigrants, expatriates and tourists it has to compete with two other cities one of which is Johannesburg.

Johannesburg often called Jozi or Joburg is the largest city in South Africa and one of the 50 largest urban areas in the world. It is the provincial capital of the wealthiest province in South Africa called Gauteng. The city is located in the mineral-rich Witwatersrand range of hills and is the centre of large-scale gold and diamond trade [3]. Johannesburg has a population of 4 434 827 people of which 73% are blacks, 18% are whites, 6% coloureds and 4% of the population are asians. This city is not traditional known as a tourist destination, but has seen many tourists due to its role as a transit point for connecting flights to other parts of South Africa.

1.2 Problem

These two cities have similarities as well as dissimilarities that set them apart from each other. This sometimes leads to tourists being unsure of which of the cities to choose as their tourist destination. This project thus aims to alleviate the stress and make the choice of immigrants, expatriates and tourists considering South African cities as their destination effortless. In addition, this project will assist the South African government as well as the South African business sector to able to tailor and improve their services in order to deliver the best experience for immigrants, expatriates and tourists.

* 1. Interested stakeholders

The people of interest in this project are immigrants, expatriates, tourists, the South African government as well as the South African business sector.

1. **Data acquisition and preparation**

2.1 Data sources

The data about the three cities will be scraped from multiple websites for the cities rental expenses, property prices, basic utilities, expected expenses for food items, expected costs for entertainment, transportation costs, quality of life scores as well as crime statistics. In addition, the Foursquare API will also be used to get specific type of venues and landmarks, explore a given venues or landmarks, and search for trending venues or landmarks for the two cities.

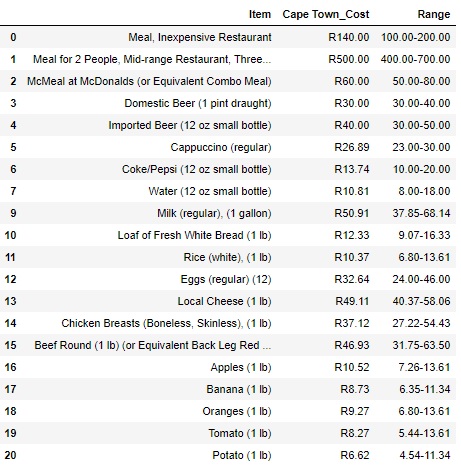
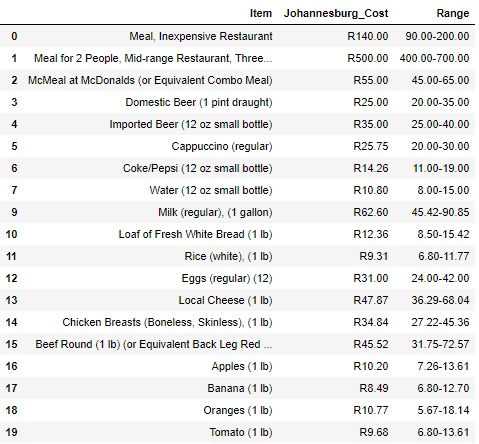
2.2  Data cleaning

The data acquired through web scrapping required some cleaning before it could be utilised. The cities rental expenses, property prices, basic utilities, expected expenses for food items, expected costs for entertainment, transportation costs, quality of life scores as well as crime statistics data for each city was represented in a pandas dataframe. In each dataframe the costs of each item had the currency sign located after the amount, thus that was corrected and all the amount reflected accurately. The, second thing we had to do was to combine the dataframes into a single dataframe where the costs for each city were given as columns and the items considered were given as the rows of the dataframe. When the new dataframe consisting of the costs of items in all cites was created it had rows that were duplicated, thus these needed to be removed in order to avoid redundant information.

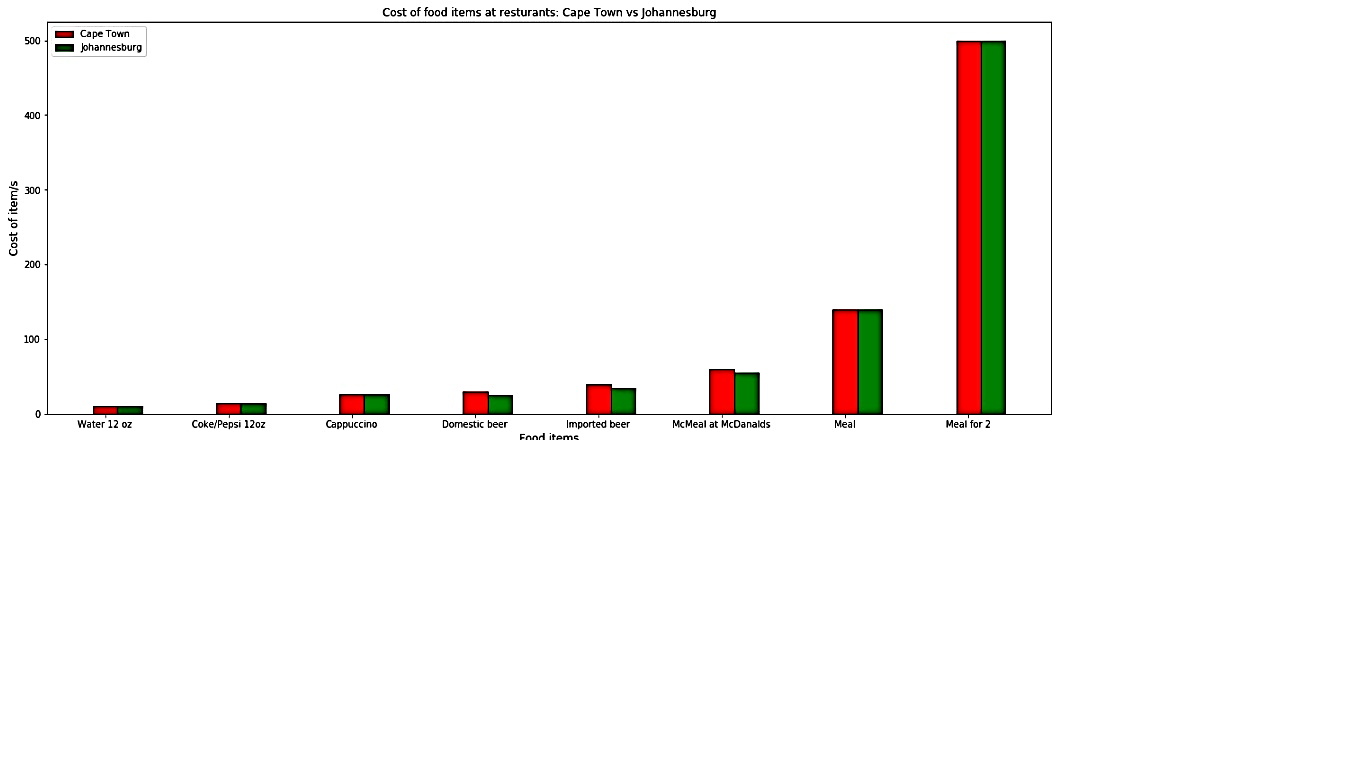
* 1. Feature selection

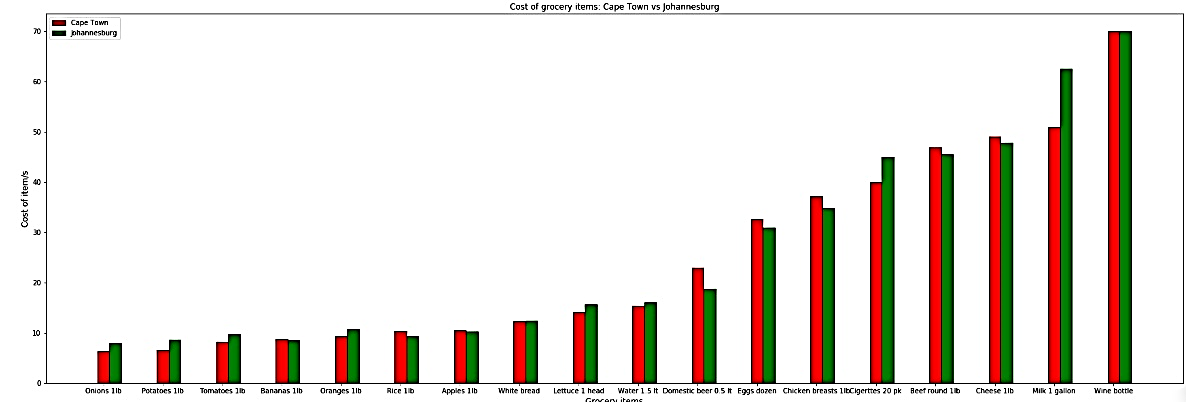
The data was collected from several sources, there were data that we did not consider, i.e. the range of prices for items sold in Cape town and Johannesburg. In addition, in the data received from the Four square application, only the most popular venues around the specific hotels in the Protea hotels company were considered in this project and these were mapped.

1. **Methodology**

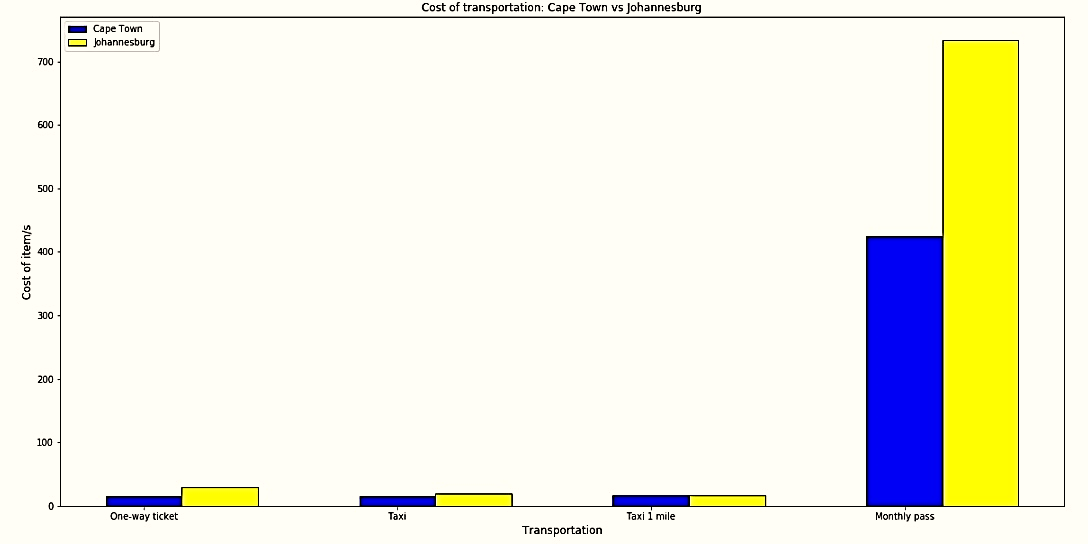
The github respiratory was used to store the resources of this project. The data scrapped from the Numbeo website that contained the cost of items in the two cities under consideration was seen to have multiple rows containing items ranging from food items, transportation to accommodation. In the columns the cost of the items as well as their typical price range

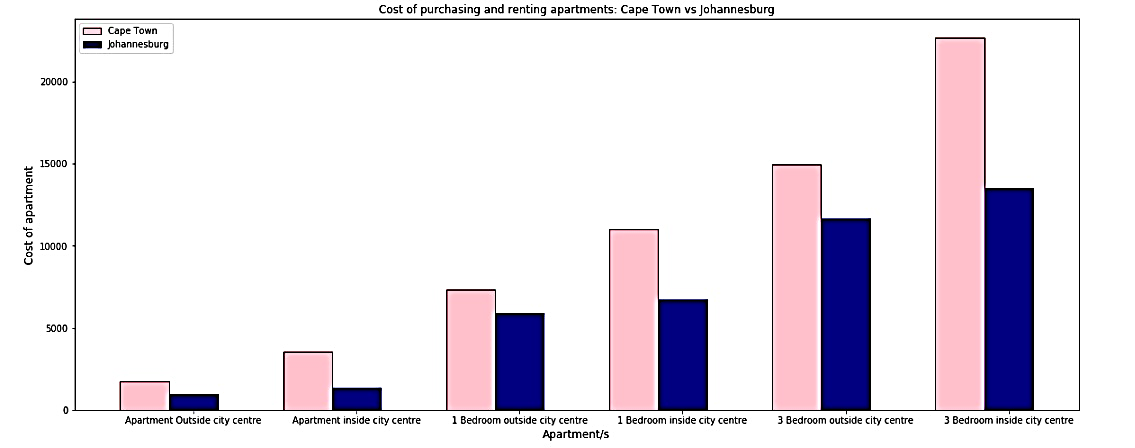
The dataframes were combined and duplicated data was removed. The dataframe contained prices for 54 items. The data was then split into categories which includes cost of items in restaurants, groceries, transportation, utilities, clothes and shoes, sports and leisure amongst other categories. The following are bar graphs that represent the costs under the identified categories for each of the two cities. It can be observed from the bar graph below that the cost for food items in resturants in the city of Cape town tend to be higher than in Johannesburg. Although this is the case some items such a cappacino, coke, etc. cost the same in the two cities.



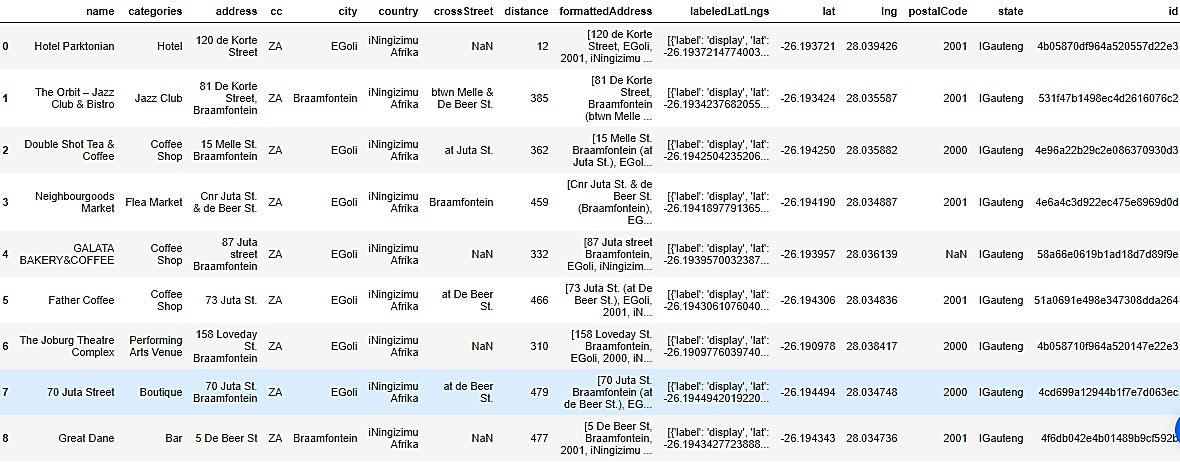
It can be observed that the cost of grocery items tends to be higher in the city of Johannesburg. In addition, some items are found to be costing the same in the two cities. Contrary to this, items such as rice, 1.5 litres domestic beer, etc. are cheaper in Johannesburg.

The issue around getting around these two cities is a crucial part if one is considering migrating or visiting the two cities. Transportation costs are much cheaper in the city of Cape town than in Johannesburg.

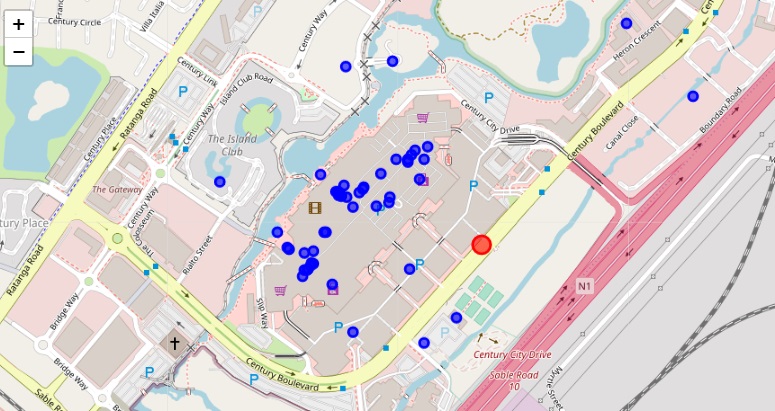
The high cost of accommodation in the city of Cape town is a concern that tourists and immigrants alike need to consider. This is proven by the observed higher apartment prices in Cape town than in Johannesburg.



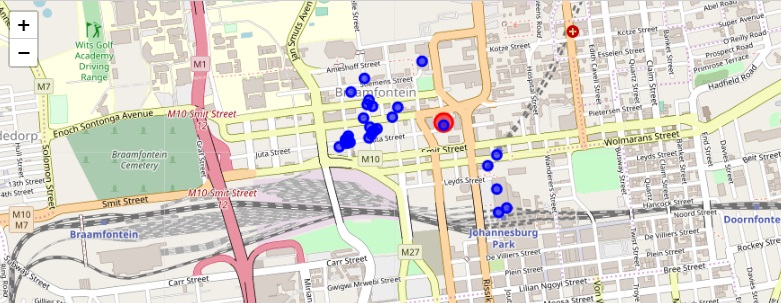
The Four square API was then used to find the coordinates of two highly rated hotels in the two cities under consideration. After this, 100 most popular venues 600 meters around these hotels were identified. Some of the popular venues around the Cape Town Marriott Hotel Crystal Towers can be seen below.

The following data shows some of the popular venues around the Protea Hotel Johannesburg Wanderers.

The folium library was then used to visualise the hotels as well as the most popular venues around them. Below is a map for the popular venues around the Cape Town Marriott Hotel Crystal Towers. The Four square API returned 55 venues which included restaurants, parks and other tourist attraction venues. The hotel is represented by the red circle while the venues are shown by the blue circles.



The Four square API returned 29 popular venues around the Protea Hotel Johannesburg Wanderers. Below is a map of these venues and the hotel where the hotel is represented by the red circle while the venues are shown by the blue circles.



1. **Results and discussions**

The cost of items in the city of Cape town and Johannesburg varied under the category of items considered. It was observed that Cape town sells food items at restaurants at a higher price than its counterpart. In addition, the price for renting or purchasing an apartment in this city is higher than the price of renting or purchasing a similar apartment in Johannesburg. Stores in Johannesburg were seen to be selling grocery items at a higher price than in Cape town. Also, getting around town in Johannesburg is observed to be costing the people a larger amount in this city. Although this was the case, some items were sold for the same amount in both these cities.

The data received from the Four square API showed that there were 55 popular venues 600 metres around the hotel in Cape Town, while 29 venues were obtained around the hotel in Johannesburg. This result can differ from another result where several hotels from each city are obtained.

1. **Conclusion**

In this study two of the largest cities in South Africa were compared. The cost of several items in the two cities was obtained and compared. The costs for items ranged from food items sold at restaurants to rent or purchasing prices for apartments. In addition, popular venues 600 meters around one of the highly rated hotels in these cities was obtained using the Four square API for each city. The prices for buying certain items in the two cities vary, as it was observed that the city of Cape town had higher costs for items than its counterpart when the cost of food items at restaurants and accommodation were concerned. Although this was the case, grocery and transportation costs were much cheaper at this city. The data obtained from the Four square API showed the most popular venues around the chosen highly rated hotel in each city. There were 55 popular venues 600 metres around the Cape Town Marriott Hotel Crystal Towers, while there 29 venues were obtained around the Protea Hotel Johannesburg Wanderers. The maps for each hotel and its surrounding popular venues was plotted using the folium library in python. The map with the hotel in Cape town was observed to show venues that are farther from the hotel when compared to that of the Johannesburg hotel. In addition, two outlier venues are observed in the Cape town hotel map. Therefore, one can choose which city to migrate to or visit based on which of the costs they considered to be significant to them. I would recommend that individual or family planning to migrate or visit South Africa to choose Cape Town as their destination. This is due to the fact that Cape town the cost of grocery items is much cheaper in this city. In addition, there are more popular venues under different categories, i.e. entertainment, sports and leisure, etc. Cape town which is further supported by the lower transportation costs in that city.

1. **Future work**

In this study only two cites were considered which give rise to a question of whether these two cities are the only ones in South Africa being considered by tourists, migrants and expatriates. Thus other large cities in South Africa can be considered for this comparison, for instance, Durban. In addition, there was only one hotel per city being considered for this study which can give rise to a possibility of a change in conclusion if several hotels and their popular surroundings were obtained from each city.

1. **References**

[1] Numbeo: Cost of Living. Retrieved from <https://www.numbeo.com/cost-of-living/in/Cape-Town>.

[2] City of Cape Town. Retrieved from <https://en.wikipedia.org/wiki/Cape_Town>

[3] City of Johannesburg. Retrieved from <https://en.wikipedia.org/wiki/Johannesburg>

[4] Four square API