

Toronto vs Manhattan

Introduction

This project aims to help investors to obtain an insight on venues from Toronto, and Manhattan. The purpose of this project is to help these potential investors select a business to invest in based on the rating of that business.

Data Description

For the city of Toronto I will be scraping the data through Wikipedia link, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. The obtained data will be raw, so I am going to do some data manipulation and cleansing before being able to use it in the program. As for Manhattan, I have decided to export the processed data from "Neighborhoods-New-York-py-v1.0" to a csv file and then read to obtain the data needed to gain helpful insights, and perform comparison between the two cities.

Methodology

I have used pandas and matplotlib to explore the data and give an insight of the main idea of the project. I also used Foursight API to be able to retrieve venues and their categories of both cities. I used pandas to explore the data and manipulate it to be able to obtain the top 10 venues of each city and their percentage of each venue from the top 10. Using the artistic layer of matplotlib I was able to plot the processed data to give a clear visual representation of the data using bar chart.

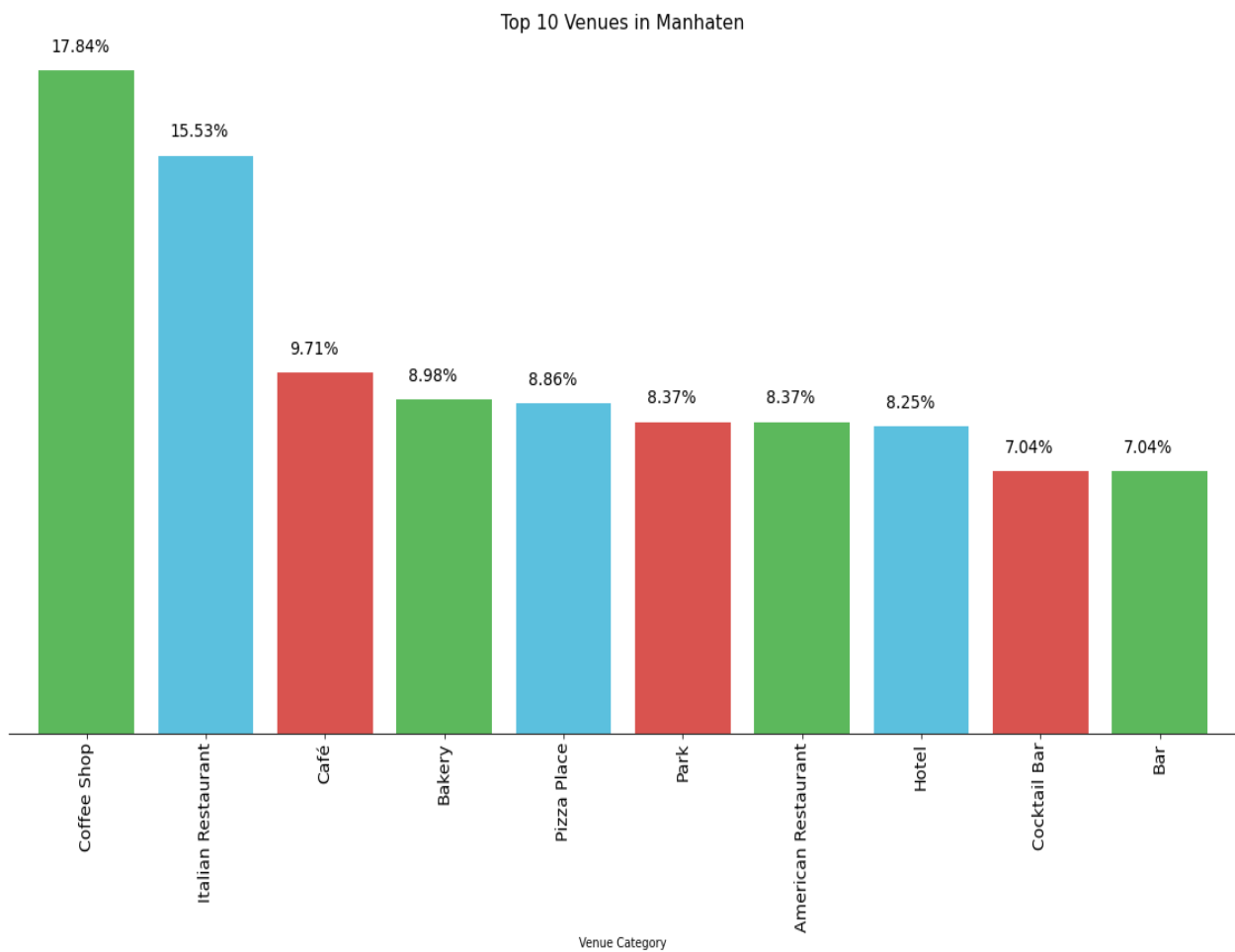
Results

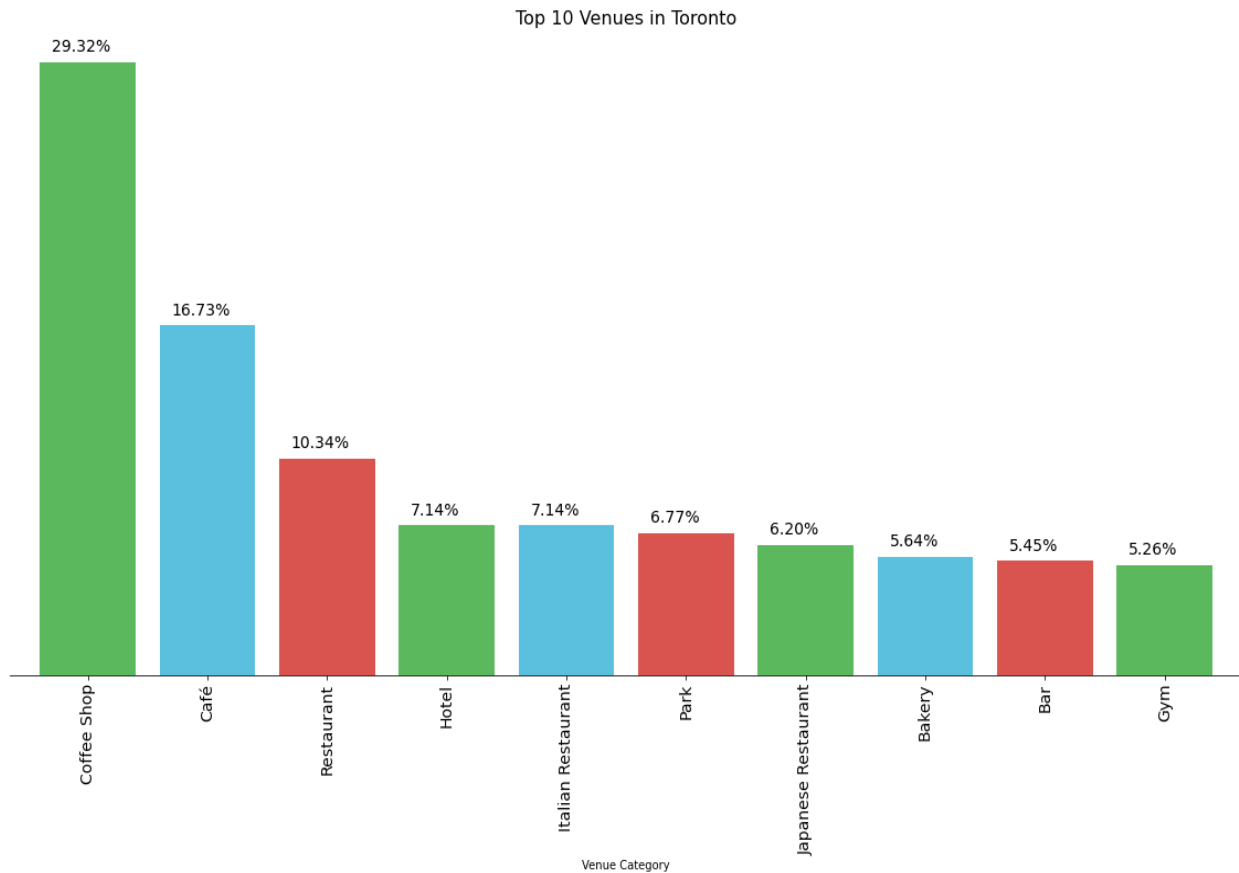
After manipulating the data and reaching and visualizing the top 10 venues of each city I have obtained the following:

In the city of Toronto the number 1 venue is Coffee shops with a percentage of 29.32%

In the city of Manhattan the number 1 venue is Coffee shops with a percentage of 17.84%

Based on the observation we can determine that the best investment in both cities would be a Coffee shop. Below is a graph that shows the data in both cities in case the investors are interested in other venues.





Discussion

If you examine the above graphs you clearly see that the difference between the number 1 venue and number 2 venue in Toronto is high with almost 13% gap. Whereas in Manhattan it is almost 2%. So when it comes to investing in Toronto I highly recommend the number 1 venue, but in the case of Manhattan since there is little to no difference between number 1 and 2 I would recommend either.

Conclusion

In this program we were able to examine and analyze venue using data analysis in a very simple and straightforward way. That only shows how implementing data analysis and data science in the business decision making can give more insights and suggest the best solution to solve a problem.