

Coursera IBM Capstone Project

*Detecting the best waterfall spots in Tamilnadu to set up
a restaurant*

Gokulakannan S

11th June 2020



Introduction

Tamilnadu is one of the 2 states in India and has 28 beautiful waterfalls across the state. The waterfalls being very beautiful have resulted in sufficient increase of the revenue of the state. Lots of tourists visit the waterfalls during the weekends and vacations in order to relax themselves. Waterfalls are the best spots for vacations. Tamilnadu being a state with 28 waterfalls welcomes tourists from all over the world to spend their leisure time here. The travel agencies try their best to find the best amenities and attract the tourists with offers. New startups and hotel business caterers also find opportunities to start their business or expand their business at these tourist spots.

Situation

The tourists find it difficult to find a proper outlet for having meals. Hence, the project aims in finding the best waterfalls spot to set up a restaurant to meet the needs of tourists.

Interest

This might help the tourists in spending a comfortable vacation or weekend without any hustle or challenge in finding a decent meal outlet nearby. This also helps new startups to open its outlet in the vicinity of these waterfalls or renowned restaurants to open their branches here.

Data Sources

The data regarding the names of the waterfalls has been taken from the Wikipedia website listed [here](#). The data lacks the exact location of the waterfalls as well as the venues nearby. Hence these datas are acquired from the [Geocoder Package](#) in Python. The nearby venues are found using the [Foursquare API](#).

Data Wrangling

The data from the wikipedia site is first extracted and then the null values

were cleaned and then a perfect dataframe was created with a data size of (28,1). The data consists of waterfalls ranging from Agasthiyar Falls to Vattaparai Falls. The data is cleaned using Pandas Library and the data is parsed and used using the Beautiful Soup Library.

Methodology

The geographical coordinates of the individual waterfalls are obtained using the Geocoder Package. The coordinates are then merged with the original Dataframe Waterfall. A csv file is created storing the new dataset.

Map of Tamilnadu is created using Folium. The map indicates all the waterfalls present in Tamilnadu. We then search for top 50 facilities within the range of 1 kilometer. This results in creation of a new dataframe with (26,7) as data size consisting of the venues near the waterfalls. The number of the venues are counted and then analyzed for the presence of the facilities. Later the column names related to restaurant chains are collected and analysed. This data is then clustered.

Clustering

Three clusters are formed and then analysed based on the venues. On Visualization these led to 4, 4 and 2 waterfalls in each cluster. The clustering is done based on the locality of the restaurants, hotels nearby.

Results

From further analysis on clusters we found that 6 waterfalls were deficient in any restaurant or decent stalls nearby. The Waterfalls were-

- Courtallam(Cluster 1)
- Catherine Falls(Cluster 2)
- Thirparappu Falls(Cluster 2)
- Thirparappu(Cluster 2)
- Vaideki(Cluster 3)
- Siruvani Falls(Cluster 3)

Discussion

From the above data analysis, we can predict that it would be a good decision if the restaurants are constructed in Cluster 3. Cluster 1 seems to be equipped with good hotels and restaurants nearby. The district where these falls fall is Coimbatore, which is indeed a place with a huge crowd that can be attracted. The other districts include Kanyakumari, Nilgiris and Tenkasi.

Conclusion

After observing the data it is advisable for the startups or food chain restaurants to open up their restaurant or branch at **Coimbatore**. **Coimbatore** is an industrial city hence it is the best option to attract customers as well.

Bibliography

- <https://en.wikipedia.org/>
- <https://www.trawell.in/tamilnadu/waterfalls> (Picture inserted in the report)