IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Price Prediction and Business Analysis Web-Application for Entrepreneurs

[1]Soham Desai, [2]Shlok Asher, [3]Het Patel, [4]Stevina Correia

[1,2,3]Student, [4]Assistant Professor, Department of Information Technology

Department of Information Technology, Dwarkadas Jamnadas Sanghvi College Of Engineering, Mumbai, India

Abstract: As of 2020, the Indian food processing market reached a value of almost 31 billion rupees. Although this value depreciated in the following months due to the COVID pandemic, the restaurant industry is rapidly recovering its customer base and profits. In the coming years, we can only see further growth, wherein the industry has a (Compound Annual Growth Rate) CAGR of about 12% for every forthcoming year until 2026.[1]

Before one opens a business, a major step in the process is to put in the research for sustainability and profitability. This research is often tedious, and requires inputs from various sources which may or may not require specialized opinions. For example, if a person wishes to open a new restaurant, they need to analyze the target audience, the location, the potential expenditures and the competition. In this paper, we put forward a web application which aggregates all this information and provides it on a single platform. Through this platform, the client can make an educated estimate for their finances and expansion plans.

Index Terms **- restaurant location business, business analysis, property price predictor**

I. INTRODUCTION

In lieu of the sizeable increase in restaurant businesses over the last decade as well as the predicted uptick in the coming months, it is no surprise that ambitious entrepreneurs want to dip their fingers in this booming industry. However, with this, the need for easily accessible information has also grown leaps and bounds.

Given how a copious quantity of data is still on-ground and needs to be compiled, most entrepreneurs need to put in extra hours to process this data and make business decisions. This in turn also deters those who are not familiar with the city or surroundings.

To tackle this problem, we aim to develop web-app that acts as a one stop shop which provides free commercial properties with the requisite filters, along with quick and ready analysis. The analysis will cover:

1. The economic demographic of the target area
2. Restaurants with similar price ranges and/or cuisines in the area
3. A price prediction for the preferred property

II. REVIEW OF LITERATURE

As it stands, MagicBricks offers some analysis pertaining to the locality selected. They show the locality rating, which is based on cleanliness, ease of commuting and surrounding amenities/businesses. They also show the price trends in said localities in the past, and give an average of expected price in the current scenario. Lastly, they show the demand and supply of properties in said locality.

All these displayed factors can be compared among localities on MagicBricks to understand the pros and cons of areas, as well as their respective prices.

Similarly, 99acres also shows the surrounding amenities for a particular property but doesn't have the depth of information that MagicBricks has.

Zomato on the other hand, shows information about restaurants. It gives us details on the cuisine, price for two and the facilities in the restaurants. By changing the target location and manipulating the filters (eg: cuisine, rating, price), we can also see the restaurants in the target area that meet this criteria. When it comes to existing work, there is undoubtedly a significant amount of information available. However, it is scattered. While MagicBricks shows us details regarding the locality, it doesn't give us a precise idea of the businesses in the area aside from an arbitrary rating.

Meanwhile, Zomato shows us restaurants but they are all only accessible when we apply the location and filters manually. They cannot be directly accessed or compared with the available properties in the area, and their respective prices.

Currently, there is currently no website that offers both facilities together in a concise and consumable form.

III. PROPOSED SOLUTION

A. Obtaining the data

To kick off our project, we had to first get two datasets – one, commercial properties in Mumbai that are up for sale or rent, and two, restaurants in Mumbai. Both datasets would contain extensive information regarding the specifications. For example, in properties, our dataset includes features such as – the exact address, total area, carpet area, images, number of amenities, proximity to facilities such as malls, Leadership in Energy and Environmental Design (LEED) certification, furnishing status, etc.

The restaurants dataset includes price, cuisine type, category of restaurant, timing, and rating.

The first dataset was obtained through scraping MagicBricks website. We used Python on Jupyter Notebook with Pandas, BeautifulSoup and Selenium libraries. First, each property link was collected manually from the site, and was put into an Excel file. The file was then run through the scraping code which traversed through each link and pulled the necessary information. A new Excel file was made with the updated information. The data was then pre-processed in MS Excel and Python to make it fit for the ML algorithms.

B. Implementing Machine learning algorithms on the data

Machine Learning is widely used in the real estate industry. The most common application of it, is price prediction wherein multiple quantifiable and tangible characteristics of a property are considered and used to predict a price.

There are various machine learning algorithms that have been used for this, and the best performing ones generally depend case to case. This is because each dataset differs from one another.

As a baseline however, the algorithms look to quantify different factors associated to the property, known as ‘features’. For example, the square foot area, number of lifts, level of furnishing etc.

The algorithm used i.e. decision tree algorithm, has yielded poor results for cases like price prediction of housing. [2] However, in our case for commercial flats, it has given us results with 94% accuracy.

As for price prediction, we tested three different algorithms and picked the optimal one out of them. The algorithms were decision tree algorithm, multilinear regression and random forest regression.

C. Building the front-end

For our front-end, we used HTML (HyperText Markup Language), CSS (Cascading Style Sheets), and JS (JavaScript).

Since our web application is dynamic and offers the users quite a few choices, we had to use JavaScript for coding these areas. Meanwhile, the beautification and the overall look-and-feel of the site was done through HTML and CSS

As part of the analysis, the users are asked the details as shown in figure 1, where they must enter necessary details in order to obtain a predicted price. This estimated price is obtained from our dataset of properties and can help the users get an idea about the possible property price during negotiations.

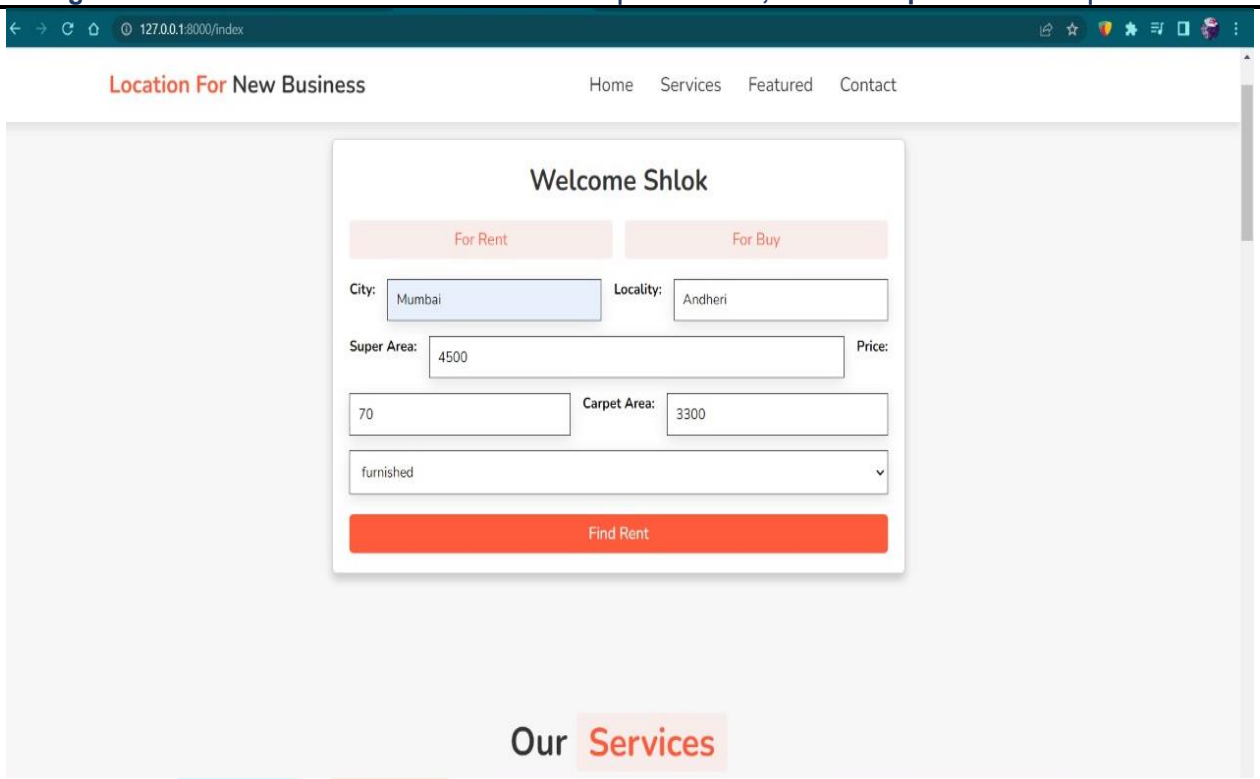


Figure 1: User input for rent prediction

Aside from this, we also offer two more services as shown in figure 2. One being the property search service, which is split into properties available for rent, and properties that are available for buying. The second service is restaurant analysis.

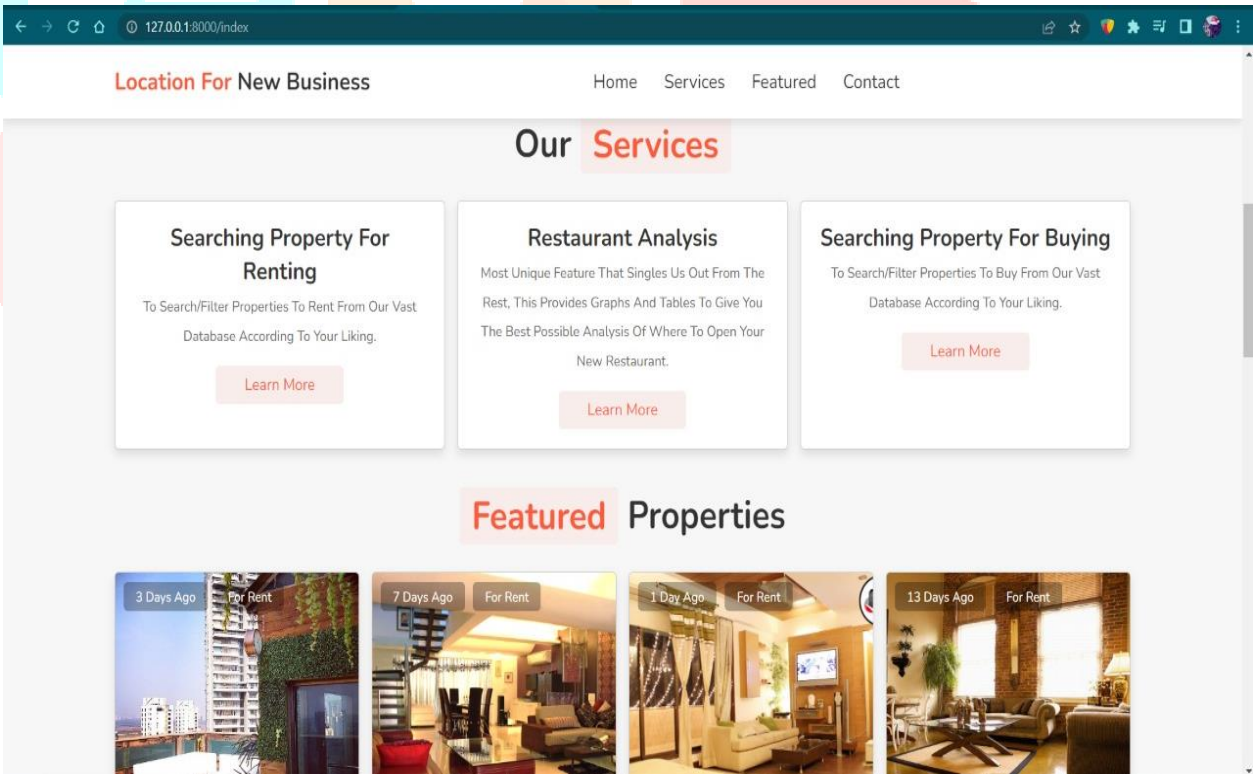


Figure 2: Home page, displaying all our services

In figure 3, we have the user input certain variables so that they can narrow down their search. The variables are location, minimum and maximum price, and minimum and maximum area.

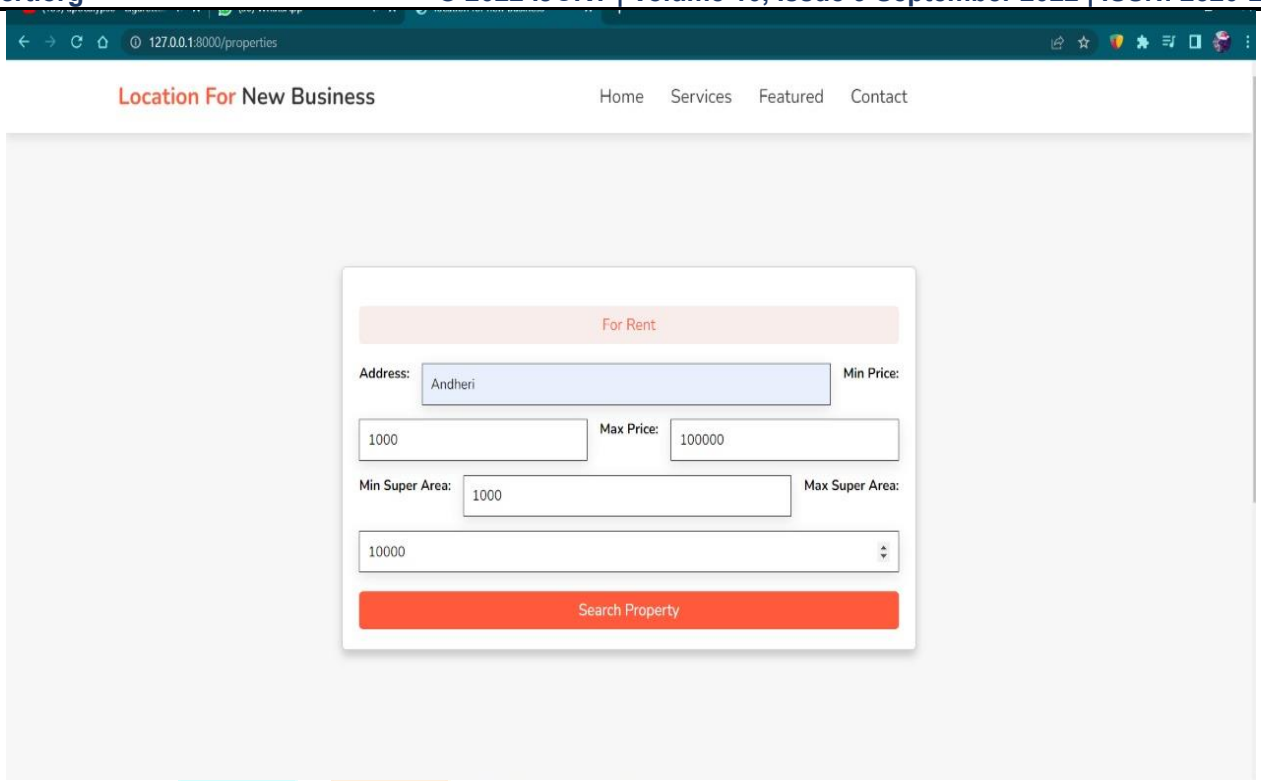


Figure 3: Input variables for finding properties available for rent

Based on the input, we generate a list of properties that the users can view and target, as seen in figure 4.

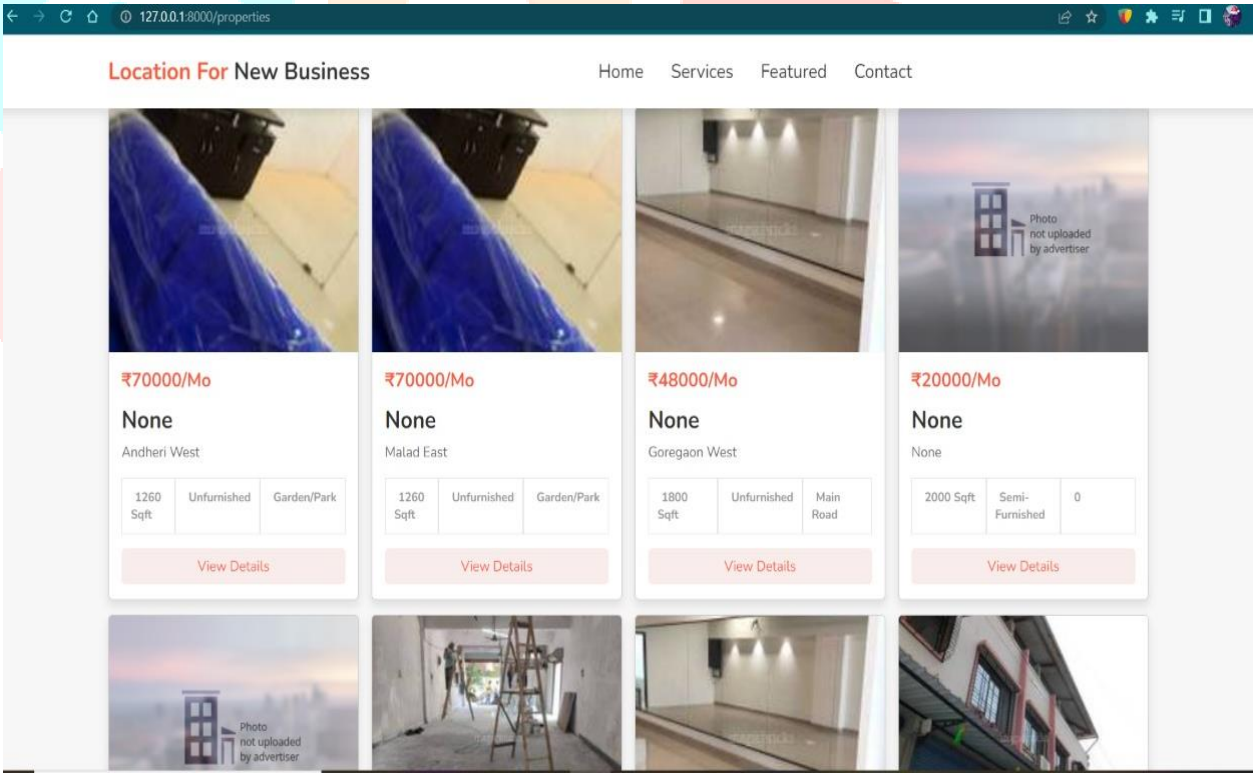


Figure 4: Displaying properties that match criteria

Lastly, the restaurant analysis service first asks the user to input a target cuisine. Based on this, the analysis is given, which we will discuss under the results and discussion section. This can be seen in figure 5.

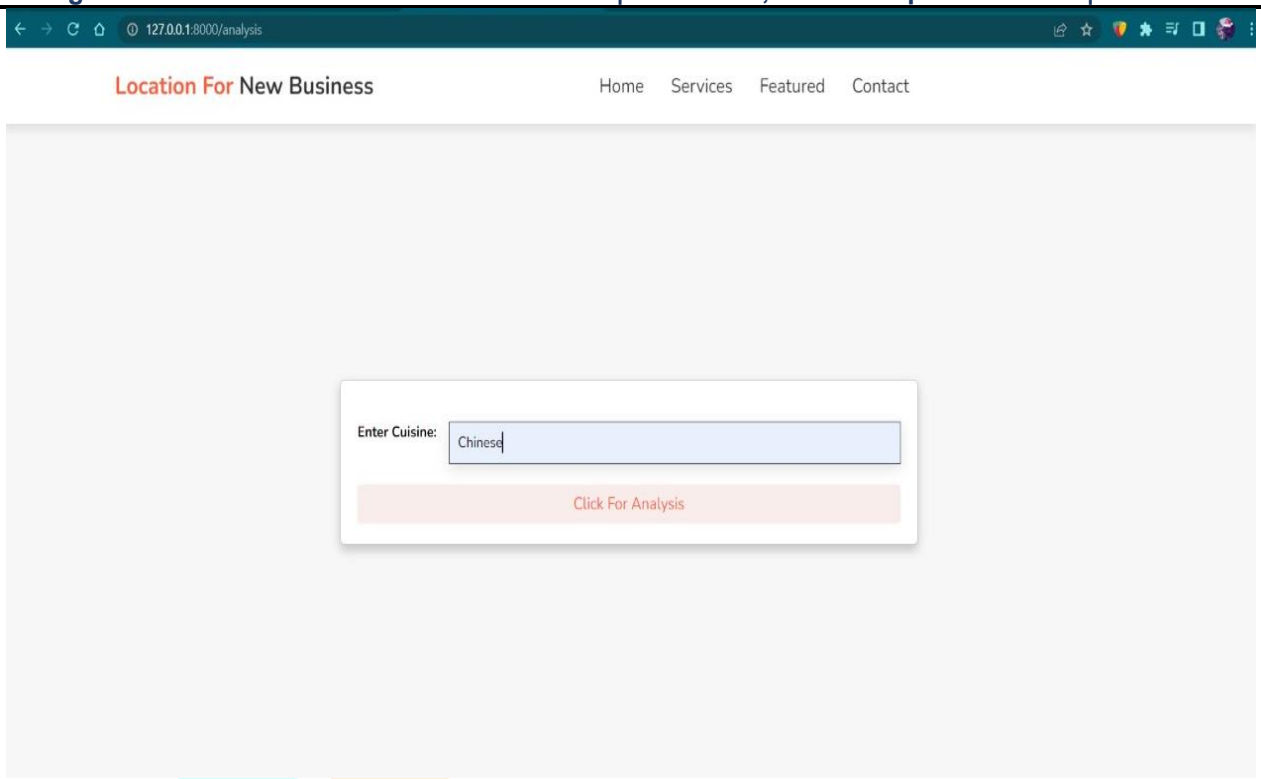


Figure 5: Entry Page for Restaurant Analysis

For the forms on the site, we used the CSRF token. CSRF stands for Cross Site Request Forgeries. CSRF is essentially Django middleware which protects a user from malicious third-party sites that aim to request credentials, and then save them for misuse at a later time.

IV. RESULTS AND DISCUSSION

Out of all the algorithms implemented and tested, we found that decision tree algorithm was the best fit for us, with an accuracy score of 94.157%

Table 1: The types of regression implemented, and their scores

Type of regression	Accuracy score
Random Forest	89.491%
Multilinear	92.108%
Decision Tree	94.157%

Figure 6 shows us a sample predicted value on our site using the decision tree algorithm.

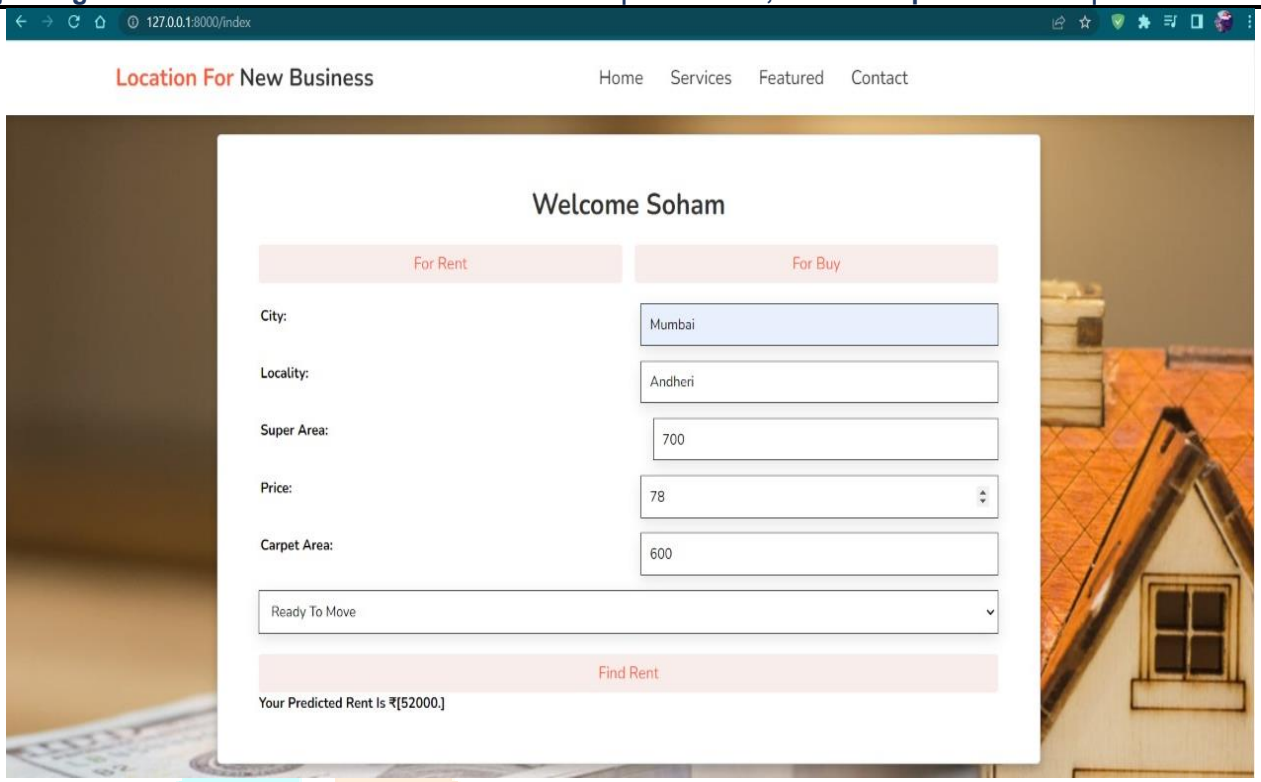


Figure 6: Output of rent prediction based on user input

Once the user enters the cuisine type on the restaurant analysis page, the following tables are shown:

1. The first half of figure 7 simply shows us the number of restaurants which correspond to each rating type, to understand the average quality for each cuisine and the density for the same. Eg: for Chinese cuisine, the bulk of restaurants lie in the 'average' and 'good' rating, and the second part shows us the top restaurants for the chosen cuisine.

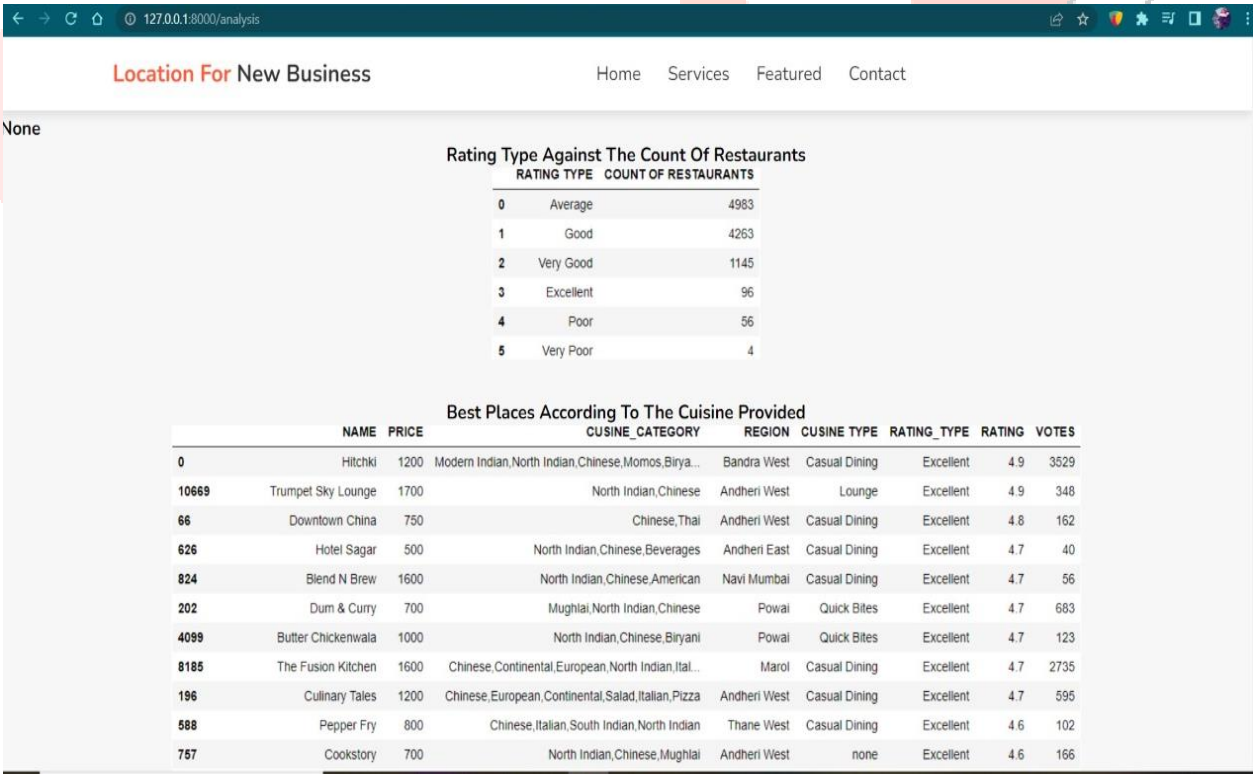
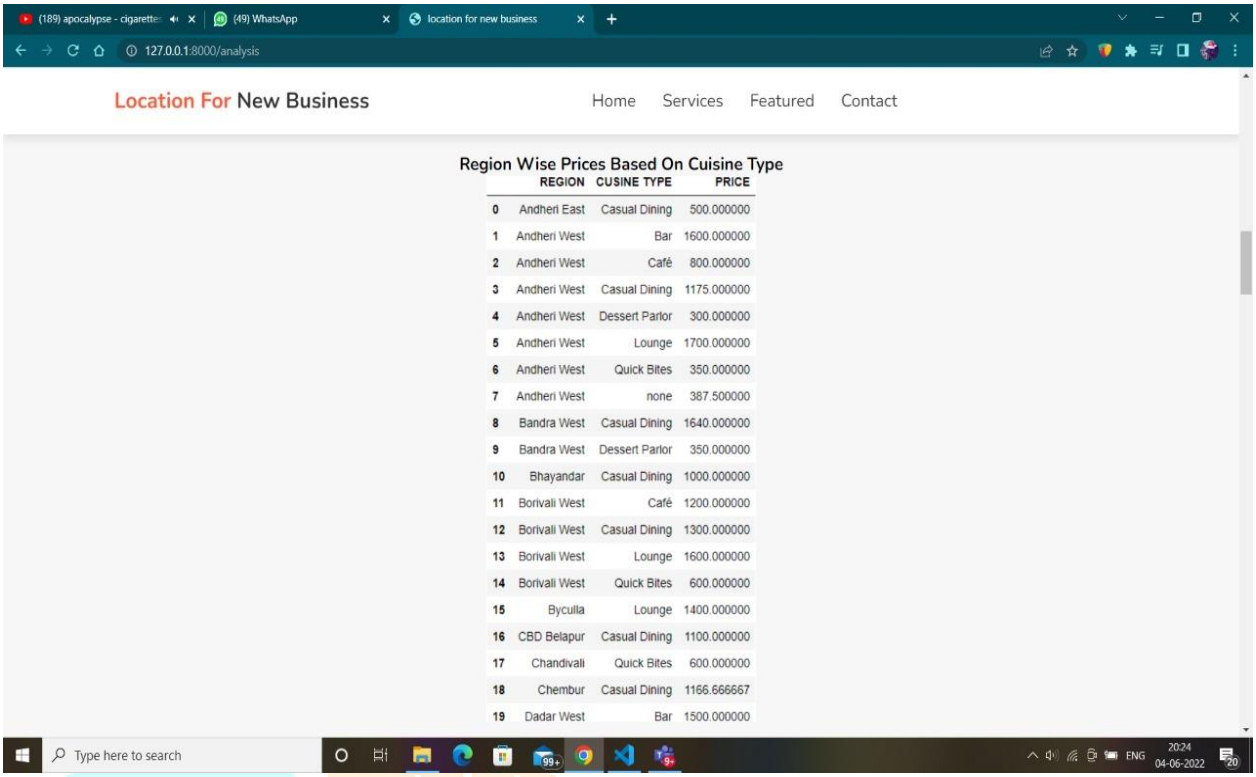


Figure 7: Rating Distribution of restaurants and the top places for input cuisine

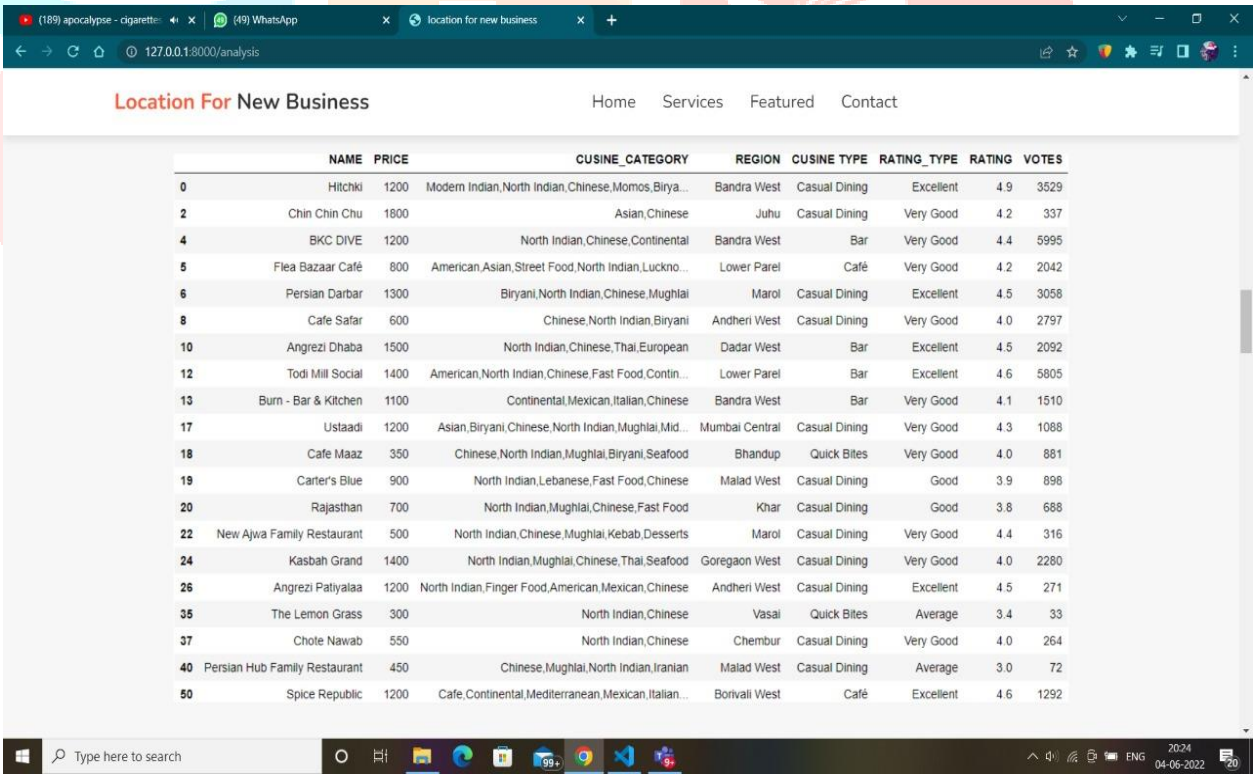
2. As shown in figure 8, we can see the average price range for each cuisine type sorted by region.



	REGION	CUSINE TYPE	PRICE
0	Andheri East	Casual Dining	500.000000
1	Andheri West	Bar	1600.000000
2	Andheri West	Café	800.000000
3	Andheri West	Casual Dining	1175.000000
4	Andheri West	Dessert Parlor	300.000000
5	Andheri West	Lounge	1700.000000
6	Andheri West	Quick Bites	350.000000
7	Andheri West	none	387.500000
8	Bandra West	Casual Dining	1640.000000
9	Bandra West	Dessert Parlor	350.000000
10	Bhayandar	Casual Dining	1000.000000
11	Borivali West	Café	1200.000000
12	Borivali West	Casual Dining	1300.000000
13	Borivali West	Lounge	1600.000000
14	Borivali West	Quick Bites	600.000000
15	Byculla	Lounge	1400.000000
16	CBD Belapur	Casual Dining	1100.000000
17	Chandivali	Quick Bites	600.000000
18	Chembur	Casual Dining	1166.666667
19	Dadar West	Bar	1500.000000

Figure 8: Region wise price based on cuisine type

3. Figure 9 is all-inclusive, and shows us exactly which restaurants/food outlets serve the target cuisine and their relevant details. This puts a microscope on the already filtered competition.



	NAME	PRICE	CUSINE_CATEGORY	REGION	CUSINE TYPE	RATING_TYPE	RATING	VOTES
0	Hitchki	1200	Modern Indian,North Indian,Chinese,Momos,Birya...	Bandra West	Casual Dining	Excellent	4.9	3529
2	Chin Chin Chu	1800	Asian,Chinese	Juhu	Casual Dining	Very Good	4.2	337
4	BKC DIVE	1200	North Indian,Chinese,Continental	Bandra West	Bar	Very Good	4.4	5995
5	Flea Bazaar Café	800	American,Asian,Street Food,North Indian,Luckno...	Lower Parel	Café	Very Good	4.2	2042
6	Persian Darbar	1300	Biryani,North Indian,Chinese,Mughlai	Marol	Casual Dining	Excellent	4.5	3058
8	Cafe Safar	600	Chinese,North Indian,Biryani	Andheri West	Casual Dining	Very Good	4.0	2797
10	Angrezi Dhaba	1500	North Indian,Chinese,Thai,European	Dadar West	Bar	Excellent	4.5	2092
12	Todi Mill Social	1400	American,North Indian,Chinese,Fast Food,Contin...	Lower Parel	Bar	Excellent	4.6	5805
13	Burn - Bar & Kitchen	1100	Continental,Mexican,Italian,Chinese	Bandra West	Bar	Very Good	4.1	1510
17	Ustaadi	1200	Asian,Biryani,Chinese,North Indian,Mughlai,Mid...	Mumbai Central	Casual Dining	Very Good	4.3	1088
18	Cafe Maaz	350	Chinese,North Indian,Mughlai,Biryani,Seafood	Bhandup	Quick Bites	Very Good	4.0	881
19	Carter's Blue	900	North Indian,Lebanese,Fast Food,Chinese	Malad West	Casual Dining	Good	3.9	898
20	Rajasthan	700	North Indian,Mughlai,Chinese,Fast Food	Khar	Casual Dining	Good	3.8	688
22	New Ajwa Family Restaurant	500	North Indian,Chinese,Mughlai,Kebab,Desserts	Marol	Casual Dining	Very Good	4.4	316
24	Kasbah Grand	1400	North Indian,Mughlai,Chinese,Thai,Seafood	Goregaon West	Casual Dining	Very Good	4.0	2280
26	Angrezi Patiyalaa	1200	North Indian,Finger Food,American,Mexican,Chinese	Andheri West	Casual Dining	Excellent	4.5	271
35	The Lemon Grass	300	North Indian,Chinese	Vasai	Quick Bites	Average	3.4	33
37	Chote Nawab	550	North Indian,Chinese	Chembur	Casual Dining	Very Good	4.0	264
40	Persian Hub Family Restaurant	450	Chinese,Mughlai,North Indian,Iranian	Malad West	Casual Dining	Average	3.0	72
50	Spice Republic	1200	Cafe,Continental,Mediterranean,Mexican,Italian...	Borivali West	Café	Excellent	4.6	1292

Figure 9: Restaurant mastersheet

4. In figure 10, we can see how many restaurants in the same cuisine are present in the different areas of Mumbai, and their average price. This is a great way to show where there could be scarcity of a certain cuisine, or where there could be inflated prices.

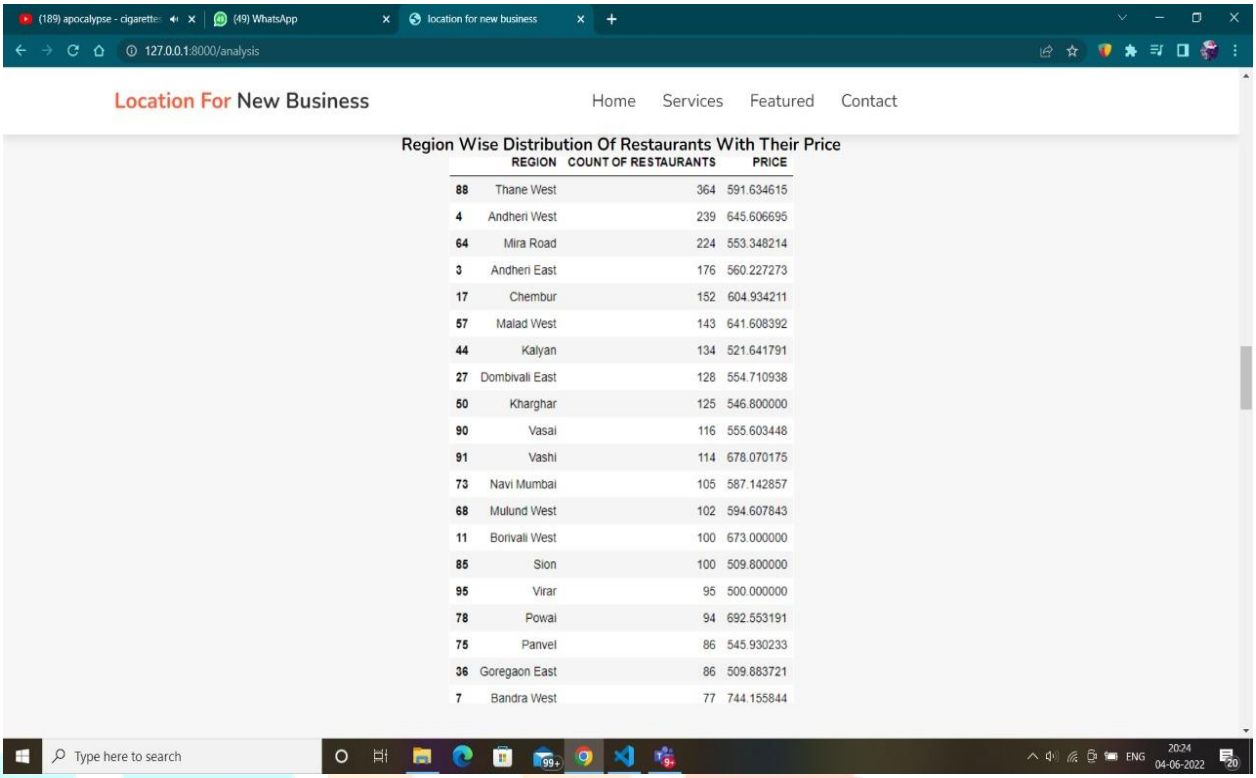


Figure 10: restaurant density with respect to cuisine and their average price

5. Figure 11 simply shows us the cuisine types and their corresponding average prices.

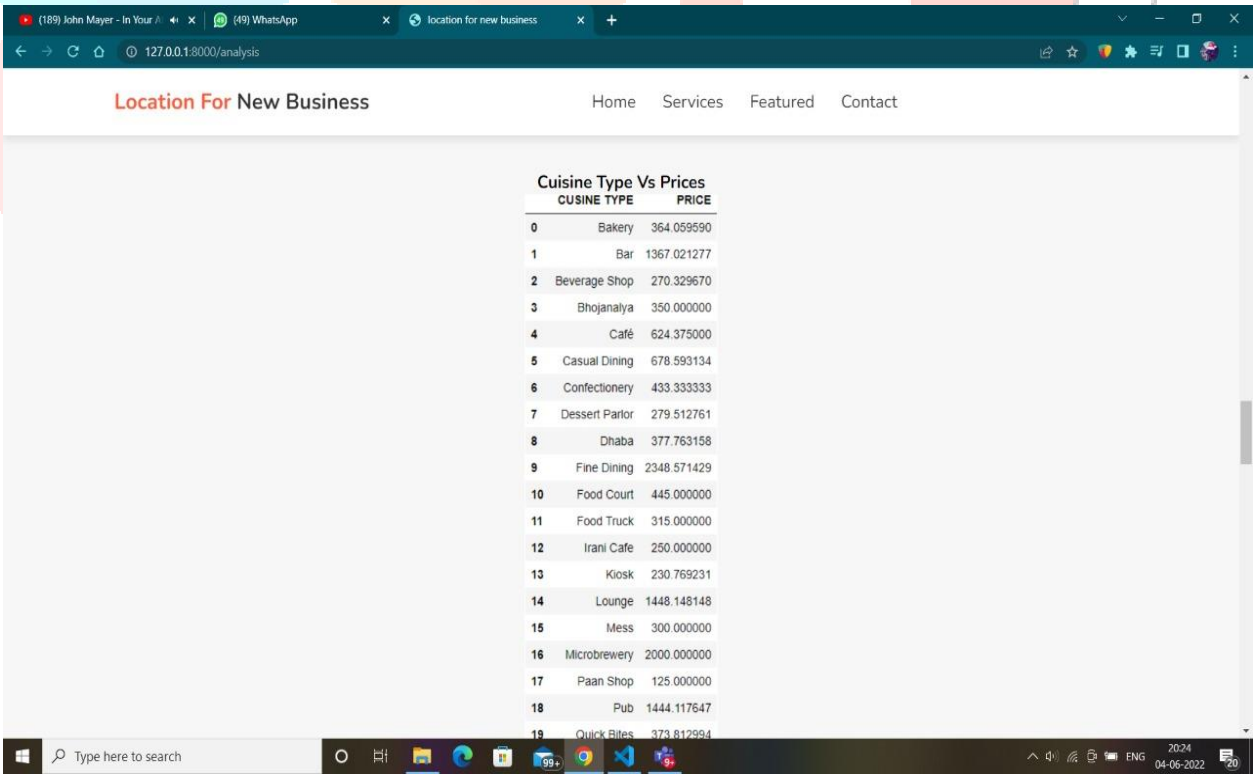


Figure 11: Cuisine types and their average prices

6. Figure 12 compares region and their average prices. Generally, it is presumed that as we go south in Mumbai, the price of outlets increases. Our analysis proves the same.

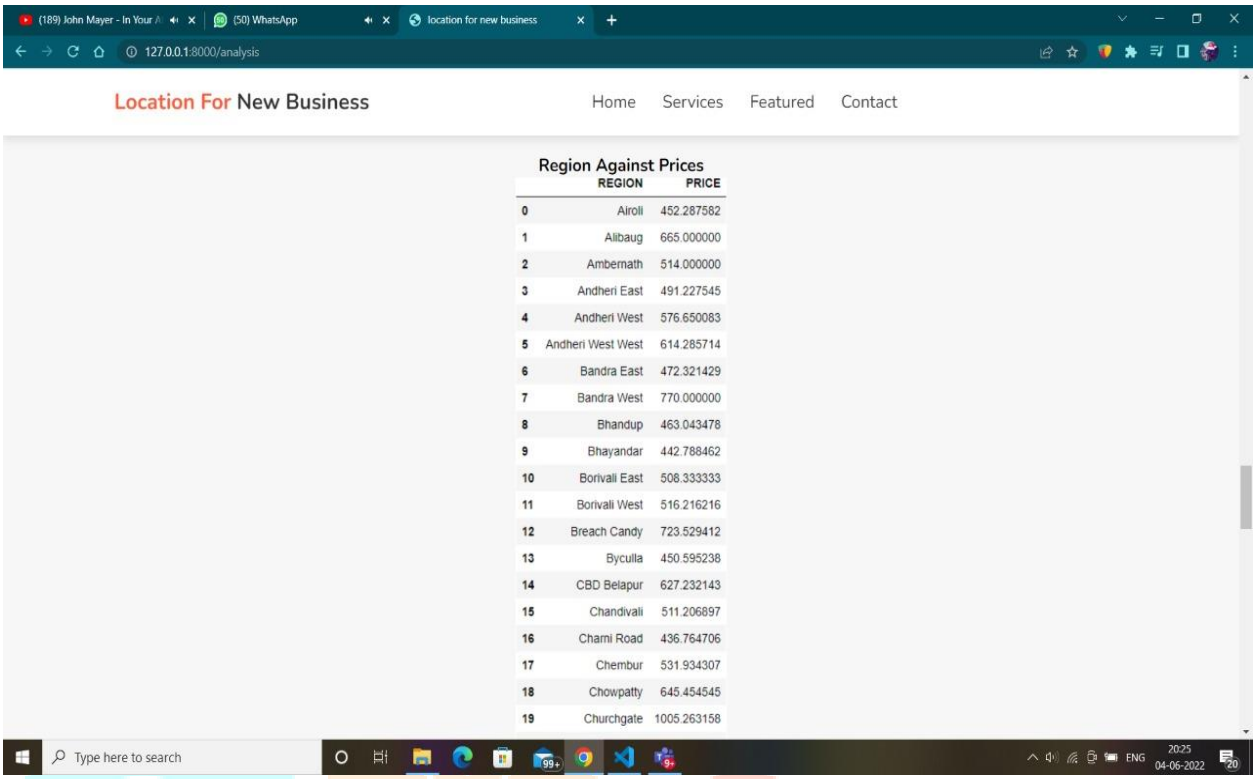


Figure 12: Price per region

7. Lastly, figure 13 adds detail to the previous one by also adding other cuisines to the equation. Thus showing whether there is any anomalous inflation of a particular cuisine in a certain area.

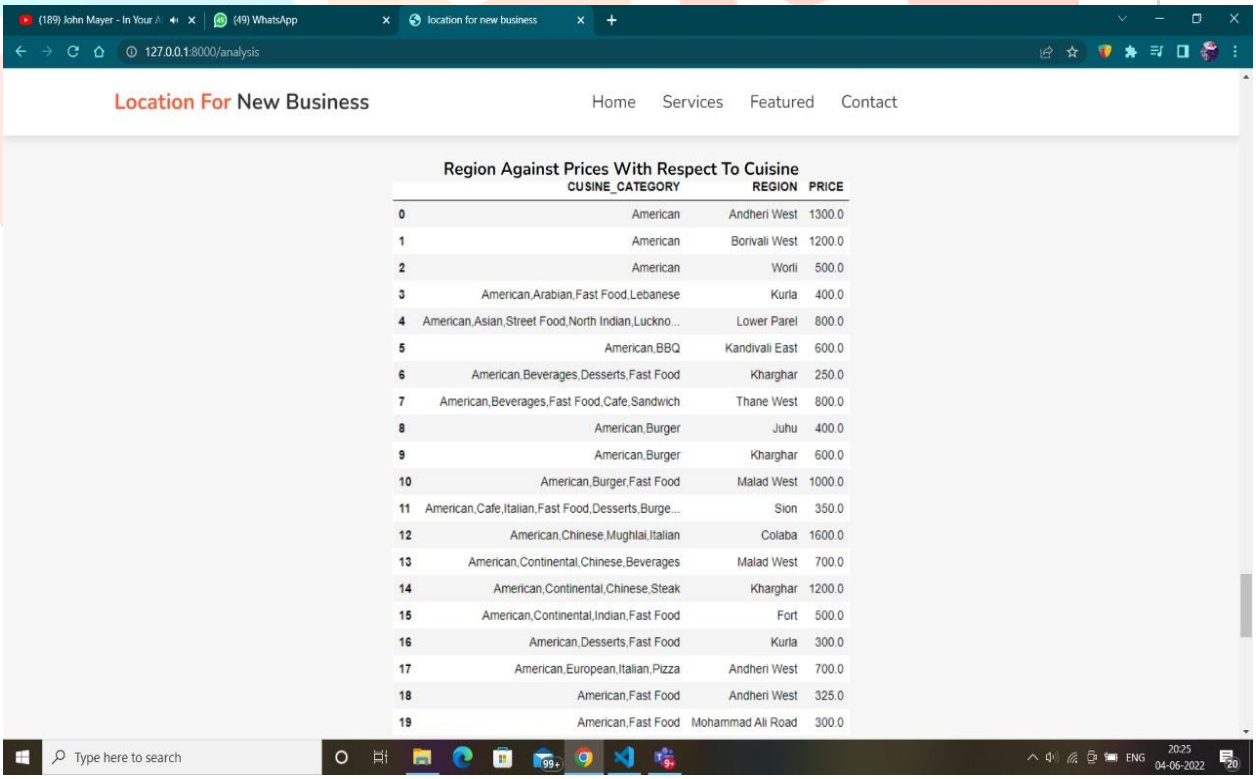
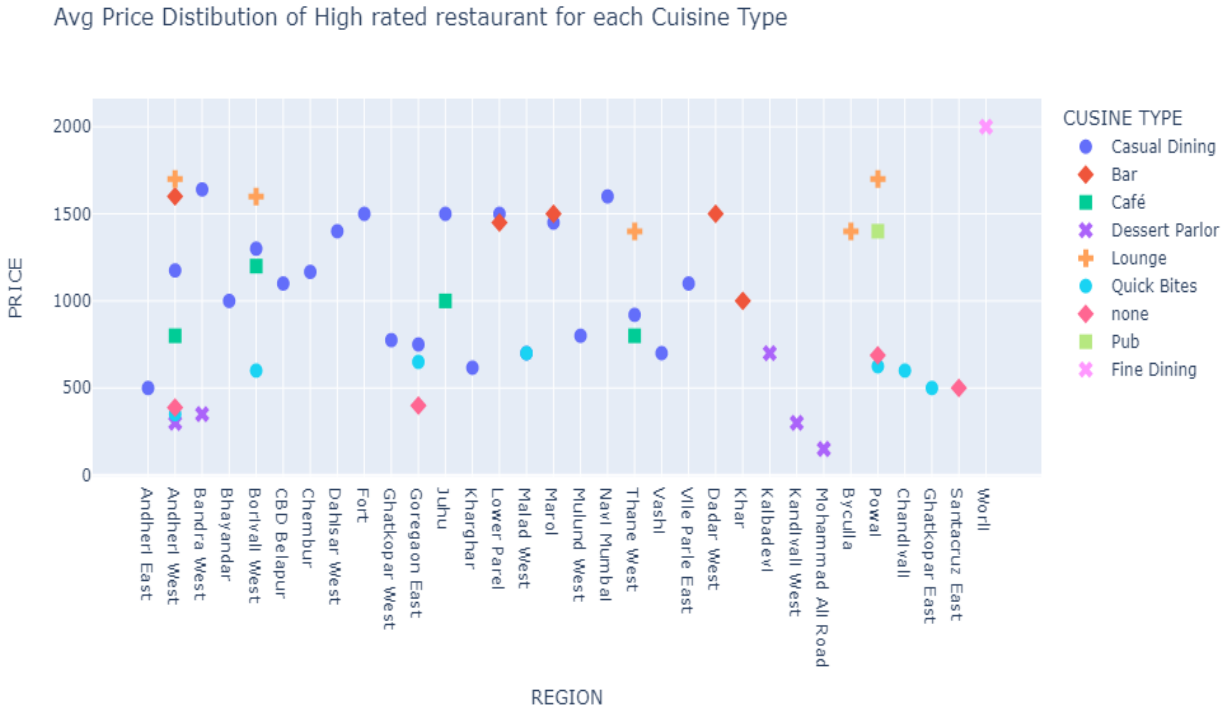


Figure 13: Regions and their prices with respect to cuisine

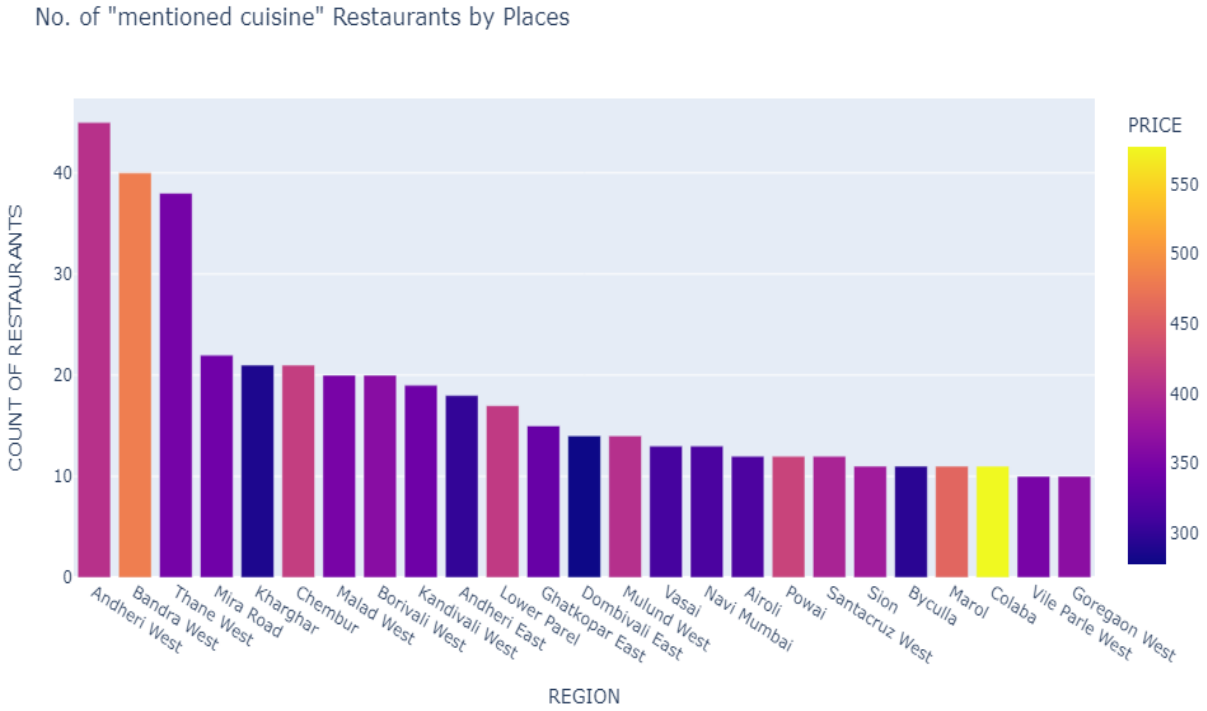
On top of the tabular analysis, we also provide a variety of graphs to better visualize the data

As shown in graph 1, we can see the average price of every high rated restaurant with respect to cuisine types.



Graph 1: Plot of price distribution of the best rated restaurants in each cuisine type

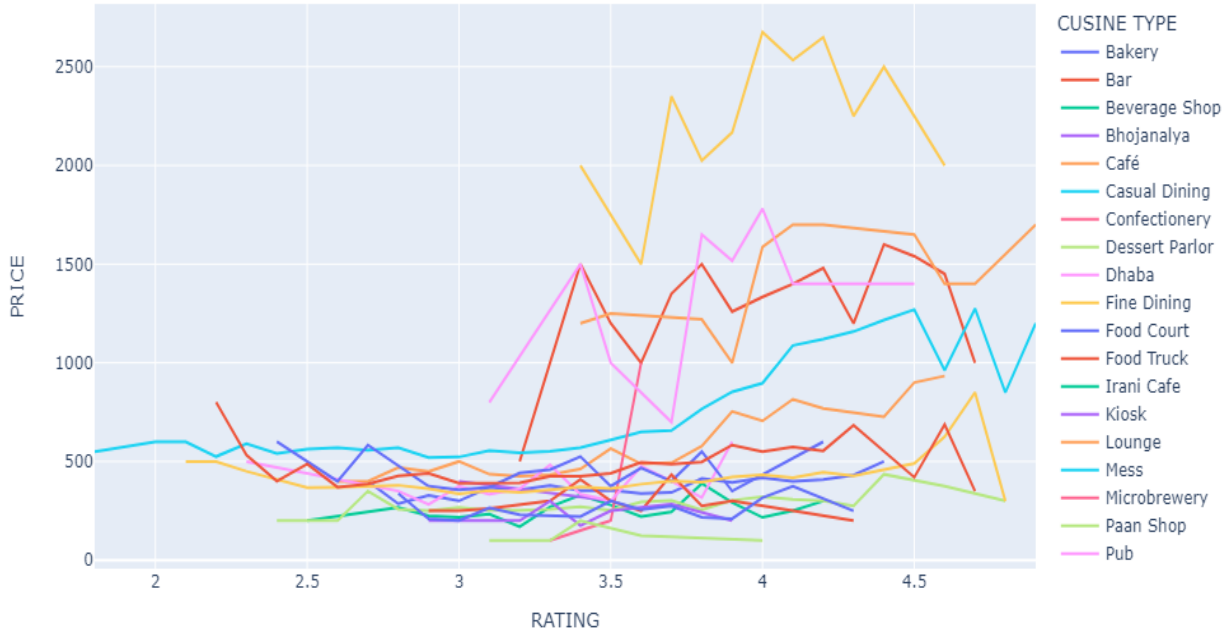
In graph 2, we can see the average price and the count of restaurants of the target cuisine, sorted by area.



Graph 2: Bar graph of frequency of restaurants for selected cuisine based on regions

Graph 3 shows the average price of each cuisine type, corresponding to our cuisine clustered on the basis of ratings.

Here we can see that Fine Dining, irrespective of the rating has a very inconsistent average price whereas Casual Dining has a smooth linear growth



Graph 3: Graph of average price of each cuisine type, with respect to cuisines clustered on basis of ratings

V. CONCLUSION

While web applications such as MagicBricks and 99acres provide a lot of detail regarding available properties, and Zomato contains information about restaurants, there is no system that shows properties as well as analysis of nearby businesses. Hence, the proposed solution combines data from both avenues and acts as a one-stop shop for entrepreneurs. The analysis given covers:

1. The prevailing customer "cost-for-two" for the selected cuisine, in the target area
2. Competition in the target area for the selected cuisine
3. An estimated price for the property with user-input variables

The scope can be expanded to multiple businesses, and multiple cities. The analysis can also be more detailed should more information be available.

VI. ACKNOWLEDGEMENT

We would like to use this occasion to express our appreciation to everyone who helped us, directly or indirectly, to complete our work. Thanks in particular to Dr. Vinaya Sawant, HOD of Information Technology, for her insightful counsel and support. Many thanks to Ms. Stevina Correia, an assistant professor of information technology and our guide, for her insightful remarks and ideas on the writing of this report. We would like to express our profound gratitude to all of the Department of Information Technology faculty for their encouragement and advice, which have allowed us to refine our work and become what we are today. We are grateful for the emotional support our parents and friends provided us during times of work when our energy were at their lowest.

REFERENCES

[1] "Indian Food Processing Market: Industry Trends, Share, Size, Growth, Opportunity and Forecast 2021-2026", imarcgroup.com, Market Overview. Available:HTTPS://WWW.IMARCGROUP.COM/INDIAN-FOOD-PROCESSING-MARKET [ACCESSED MAY 29TH ,2022

[2] ZULKIFLEY, NOR & RAHMAN, SHUZLINA & NOR HASBIAH, UBAIDULLAH & IBRAHIM, ISMAIL. (2020). HOUSE PRICE PREDICTION USING A MACHINE LEARNING MODEL: A SURVEY OF LITERATURE. INTERNATIONAL JOURNAL OF MODERN EDUCATION AND COMPUTER SCIENCE. 12. 46-54. 10.5815/IJMECS.2020.06.04.