FINAL PROJECT
DS-660 BUSINESS ANALYTICS

PROBLEM STATEMENT

Zomato is a multinational restaurant aggretor and food delivery company founded by deepinder goyal and Pankaj chaddah in 2008. Zomato provides information, menus and user reviews of restaurant as well as food delivery options from partner restaurant n select cities. As of 2019 service is available in 24 countries and in more than 1000 cities.

Lets consider you are a product manager at a company like Zomato. To release any feature, lets say you have four or five steps like ideation, design, implementation, testing and deployment. And you have 3-4 features like rate/review restaurant, filters for search, online ordering, etc. in line for next release. Write a medium post on how you would go about the development of these features using the Agile methodology, incremental and iterative way of development.

The ability to understand the problem, frame problem statements, generate hypotheses, define key metrics and provide tangible solutions is integral to any analyst job. Below are the question which are part of my problem statement.

1.How to increase AOV(average order value)?
2.How to reduce delivery costs?
3.Is it feasible to deliver multiple orders during single delivery and which which cuisines or restaurants can be targeted for a POC(proof of concept)?
4.For a particular month, say December the ice and desers orders reduced. Whats the root cause?
5.For a fixed budget of $10M to provide 10X growthon hyperpure should tier 1 or 2 or 3 cities can be targeted?
6.For a city, Pune, the overall delivery ratings are low. What could be the reasons?
7.Which restaurants should be used for customer acquisitions?
8.Yearly about 25%of restaurants in india shut down. Hoe can zomato prevent such a situation on its platform?
9.What metrics can be measure customer retention and customer satisfaction apart from rating and reviews?
10.What factors need to be considered while looking for an apt location to open a cloud kitchen in a Tier 1 city vs Tier 2?

INTRODUCTION

Founded in 2008 Zomato is a major food delivery aggregator with a markdown cap of 1 trillion INR. It started as Foodiebay, a restaurant recommendation product, at its peak, it has 35000 menus and Rs 60 Lakh monthly revenue. Foodiebay.com reroutes to zomato.com now.

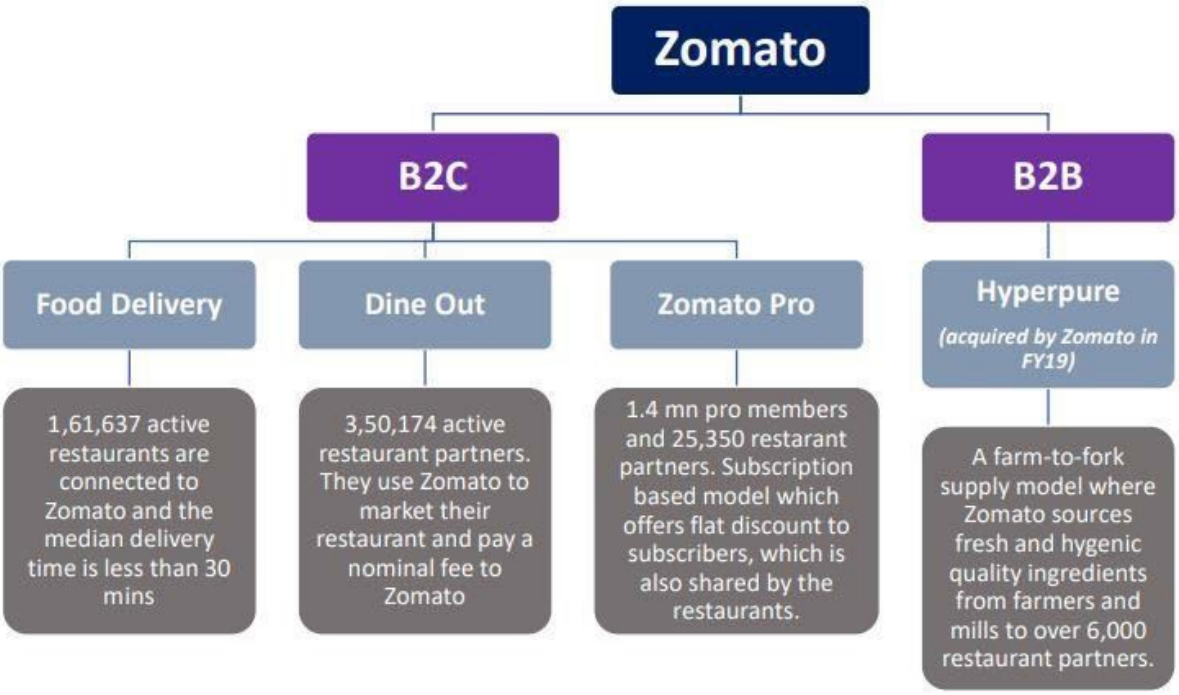
Swiggy is the major competitors in india, offerinf the same services as zomato. But zomato is ahead of swiggy with respect to its AOV by RS 35 according to a 2020 Goldman sachs report. This gives Zomato a shot in the arm to break even and be profitable as the higher the AOV higher the margins. The importance of high AOV will be explained in the later section.

Zomato has either invested or acquired in quite a few startups internationally to help grow its business, the recent one being grofers – Gurugram based online grocery firm. This helps zomato focus mainly and food delivery and lets grofers handle daily usag products. Swiggy its counterpart has swiggy mart which handles grocery delivery. Zomato believes it can provide its shareholder better value by investing in another firm than building its own grocery product.

METHODOLOGY.

With the present dataset, interesting insights pertaining to geography, restaurant type and rating van be generated. So I have used tableau for exploratory data analysis and data visualization using tableau. Tableau is easy drag and drop tool and is quite helpful in EDA's.

Even I worked on predictive analysis on Zomato data so that customer experience can be enriched, also to reduce cost or increase revenue. i.e machine learning.



ZOMATO RESTAURANTS DATA FOR PREDICTIVE ANALYTICS.

The kaggle dataset was downloaded using the zomato API. It contains dimensional information of restaurants listed on the platform. The columns available are listed below.

1.Restaurant id: Unique id of every restaurant across various cities of the world.
2.Restaurant name: Name of thr restaurant
3.Country code: Country in which restaurant is located
4.City: City in which restaurant is located.
5.Address: Address of the restaurant
6.Locality: Locality in the city.
7.Locality verbose: detailed description of the locality
8.Longitude: Longitude coordinate of the restaurant
9.Latitude: Latitude coordinated of the restaurant
10.Cuisines: cuisines offered by the restaurant
11.Average cost of two: cost for two people in currencies.
12.Currency: currency of the country
13.Has table booking : Yes/No
14.Has online delivery: yes/no
15.Is delivering: yes/no
16.Switch to order menu: yes/no
17.Price range: range of price of food
18.Aggregate rating: Average rating out of 5
19.Rating color: depending upon the average rating color.
20.Rating text: text on the basis of rating of rating
21.Votes: number of ratings given.

Switch to order menu is NO for all observations, hence it's of no value.

P*rice range* ranges from one to four, with four being premium-priced restaurants. Local currency

to dollar conversion as of 20211022. This helps to normalize prices across territories.

Currency	to_Dollars_20211022
Botswana Pula(P)	0.09
Brazilian Real(R$)	0.18
Dollar($)	1
Emirati Diram(AED)	0.27
Indian Rupees(Rs.)	0.013
Indonesian Rupiah(IDR)	0.000071
NewZealand($)	0.72
Pounds(Œ£)	1.38
Qatari Rial(QR)	0.27
Rand(R)	0.068
Sri Lankan Rupee(LKR)	0.005
Turkish Lira(TL)	0.11

Based on whether the restaurant has online delivery, and reservation a new tag is used to help segregate restaurants.

Has Table booking	Has Online delivery	New Tag
No	No	No_Reservation_No_Delievery
No	Yes	No_Reservation_Only_Delievery
Yes	No	Reservation_No_Delievery
Yes	Yes	Both_Available

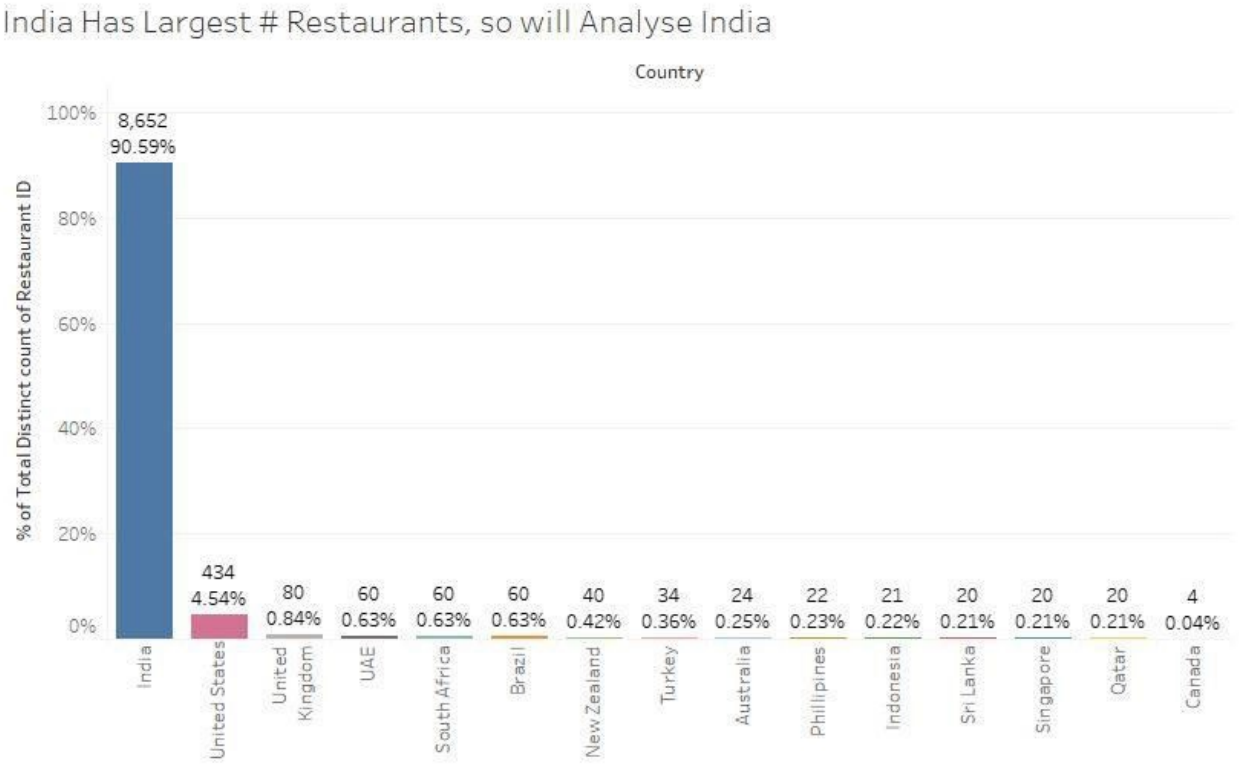
EDA and Data visualization using Tableau

With the present dataset, interesting insights pertaining to geography, restaurant type, and rating can be generated. Tableau is an easy drag and drop tool and is quite helpful in EDA's. We will be using the same to answer some basic questions.

1.Distribution of restaurants across countries?
2.Distribution of ratings, price range, the average cost for two, voters
3.Which cuisine has a higher rating?
4.Which country has the cheapest food prices?
5.How do votes and ratings vary across services(delivery/reservation) provided?

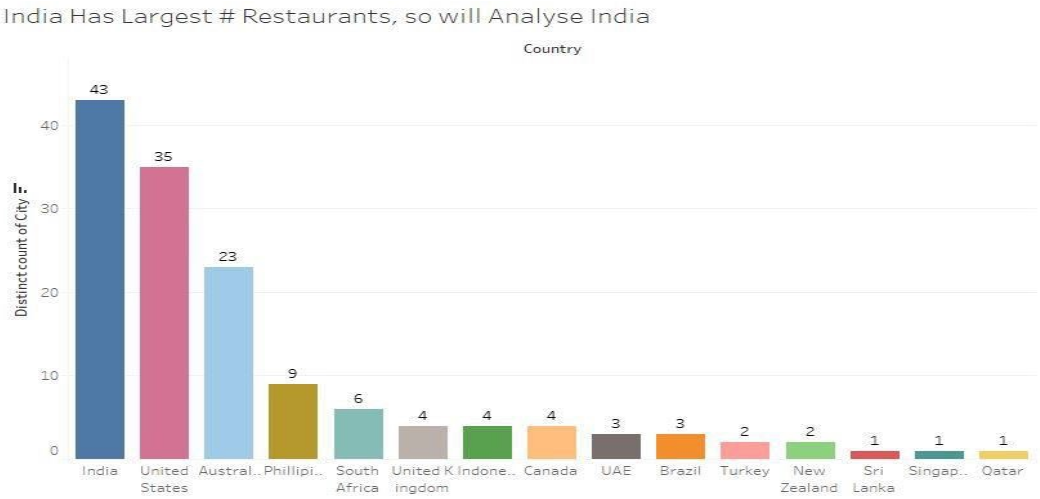
I have attached screenshot of the tableau visualization.

Distribution of # restaurants across geographies



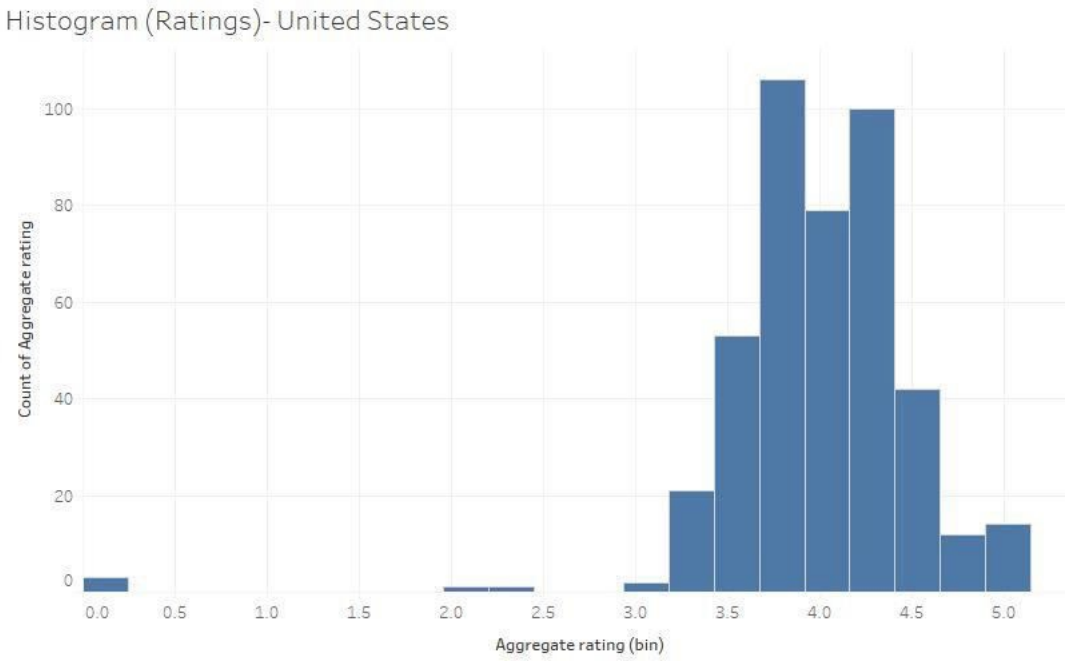
The current dataset has restaurant across 15 countires with about 90% and 5% observation for indian and 5% USA respectively. So that comparing the indian subcontinent with the USA could be an ideal analysis.

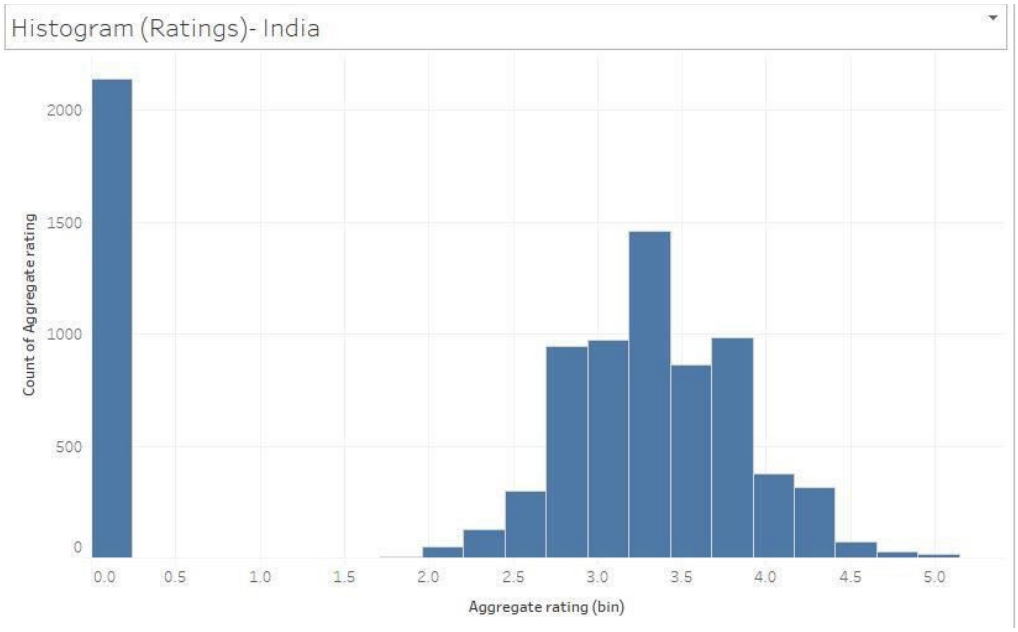
Penetration across cities



Zomato is present in 43 cities in idnia and 35 in the usa and 23 in Australia and the rest 12 cities are in single digit. So it validates the previous assumption that inida and USA can be compared.

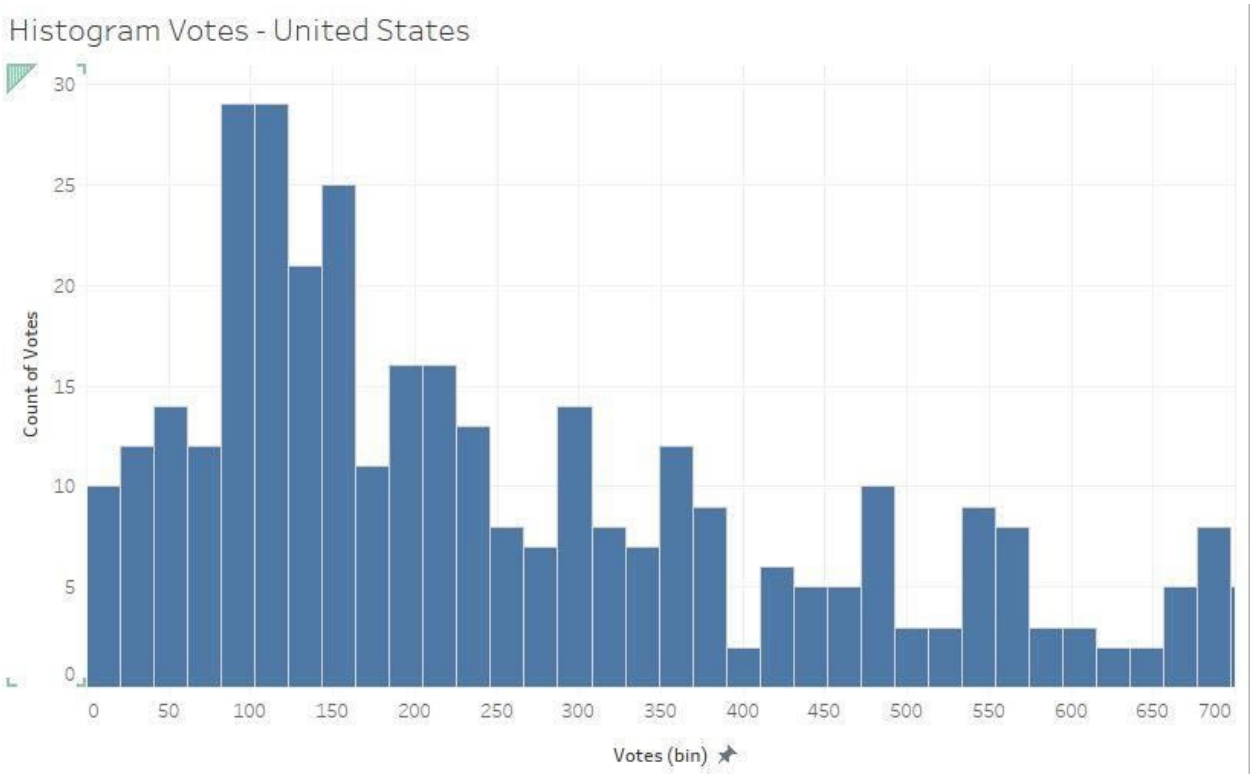
Distribution of Ratings

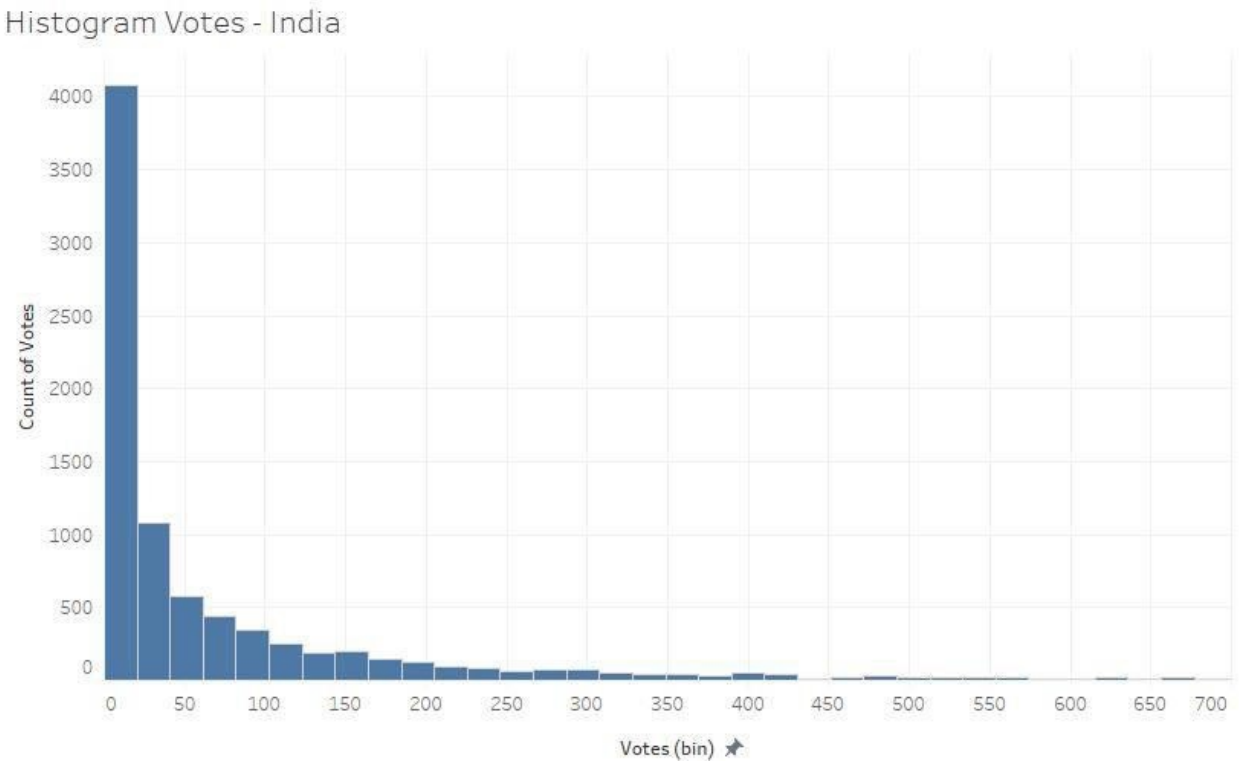




USA restaurants have higher ratings in compared with india restaurants, this could indicate a better service in the USA. A major indian restaurant ratings are zero.

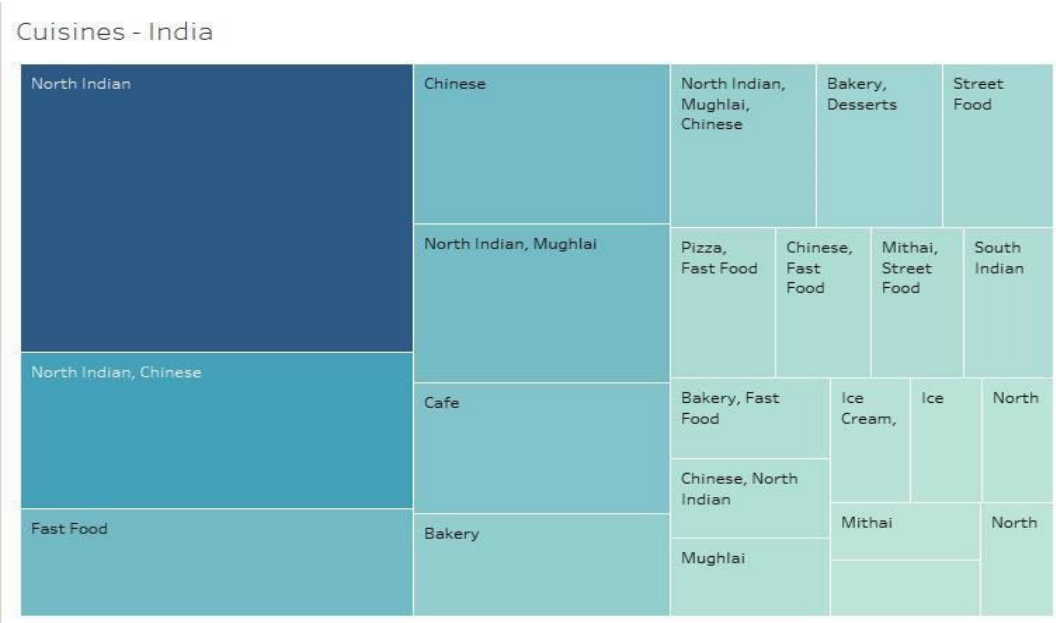
Question: could the 0 rating be due to a high number of low priced restaurant in india?



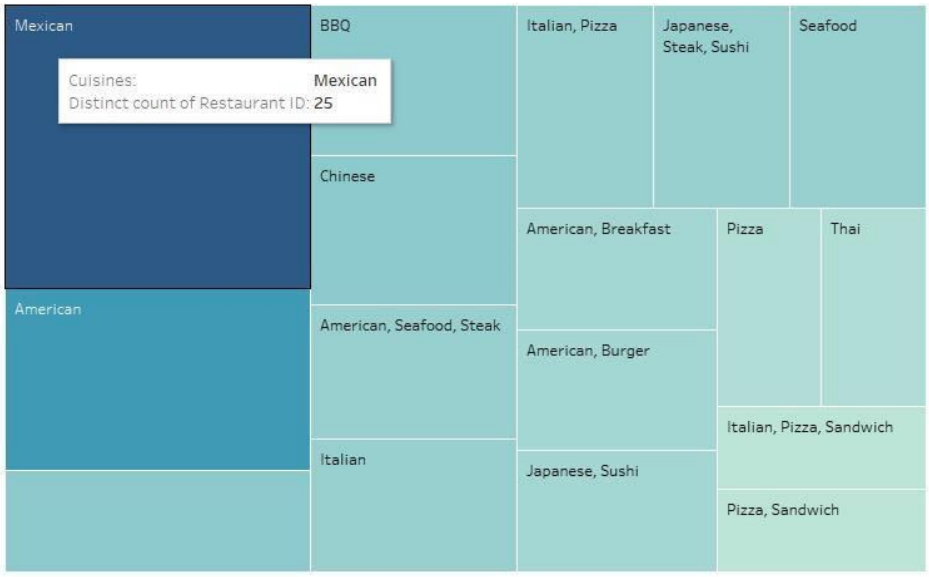


We check the rating count in the indian subcontinent vs the US. US customers have a higher vote count than Indians, reiterating the previous findings of ratings.

Cuisines



Cuisines - United States

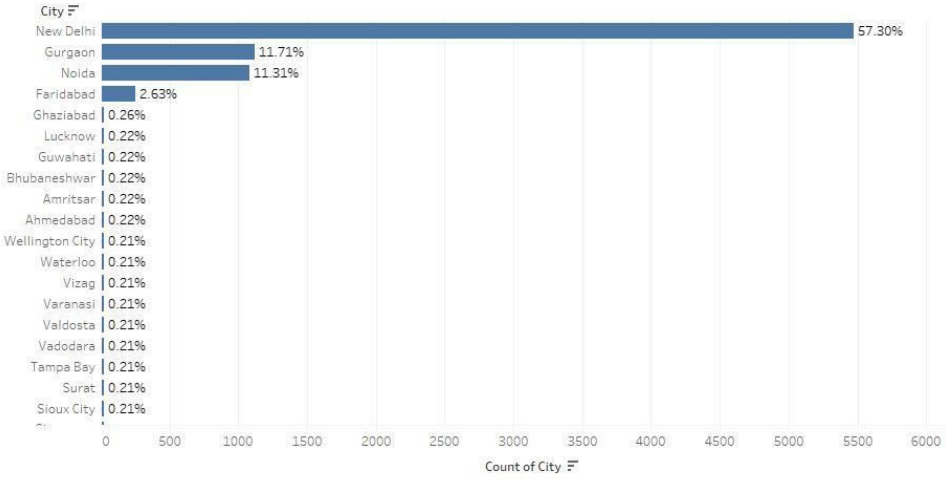


North Indian, Chinese, Fast Food, Mughlai are a few popular cuisines in india, so new restaurants opening up in this space can face stiff competition from establishes restaurants. The potential for growth in south indian cuisines seems good due to lack of competition and variety of vegetarian dishes.

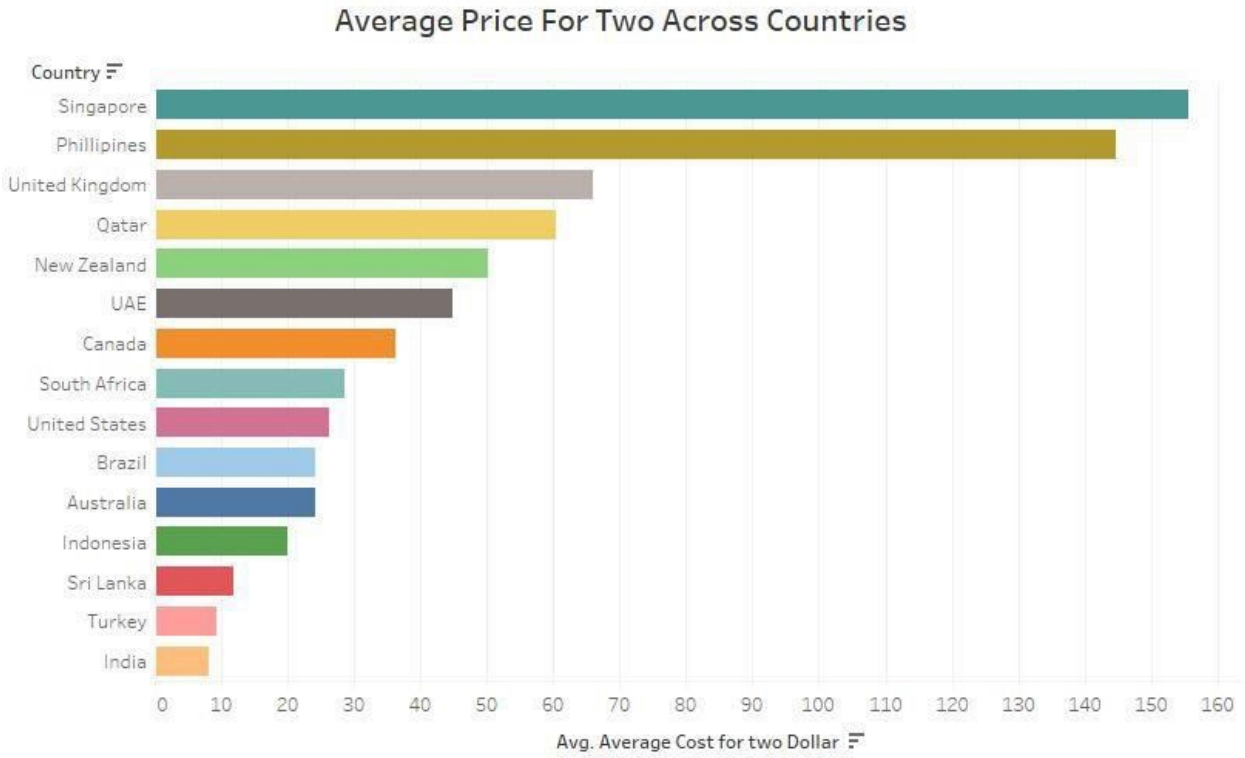
American love everything America – steak, seafood, burgers, BBQ apart from Mexican and Chines food.

Questions: North Indian food dominates the Indian taste buds, could this be because the major of the

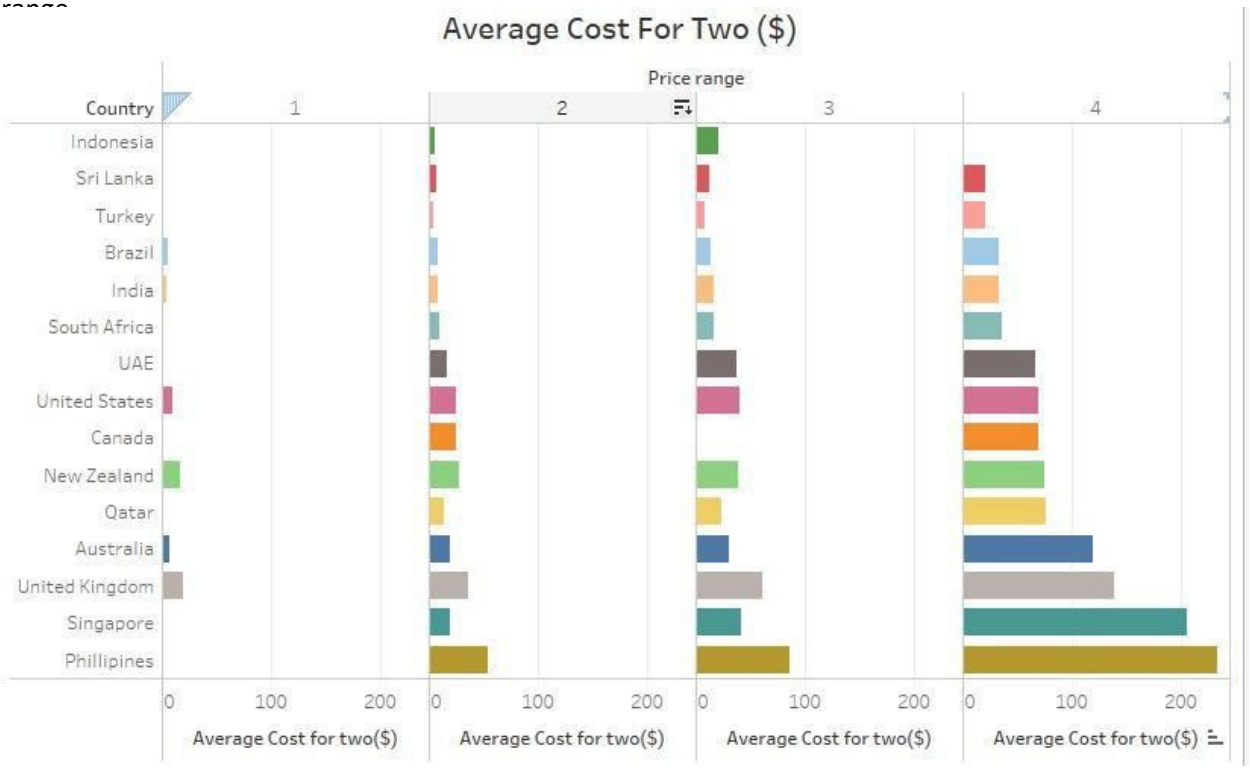
Restaurant/City



Which country has the cheaper food?



It seems that india has the lowest price for two amongst the 15 restaurants. Could it be due to the higher number of price range 1 restaurants in india? Lets break it down based on the price range



40% of retaurants across geographies fall into the cheaper price rang of 1, with 32%, 14%, 6% of restaurants falling into2,3,4 respectively, this is because indian price range 1 restaurants are dominating the dataset. Similar distribution can be found for indian restaurants as well.

For the US about 40% belong to price 2, with 31,25 and 5 belonging to 1,3,4 respectively

Overall Philippines($233), the UK($138), and Singapore($206) have higher food prices(price range 4) and Indonesia($5.3), turkey ($4.8) and sri lanka($6.3 have lower prices.

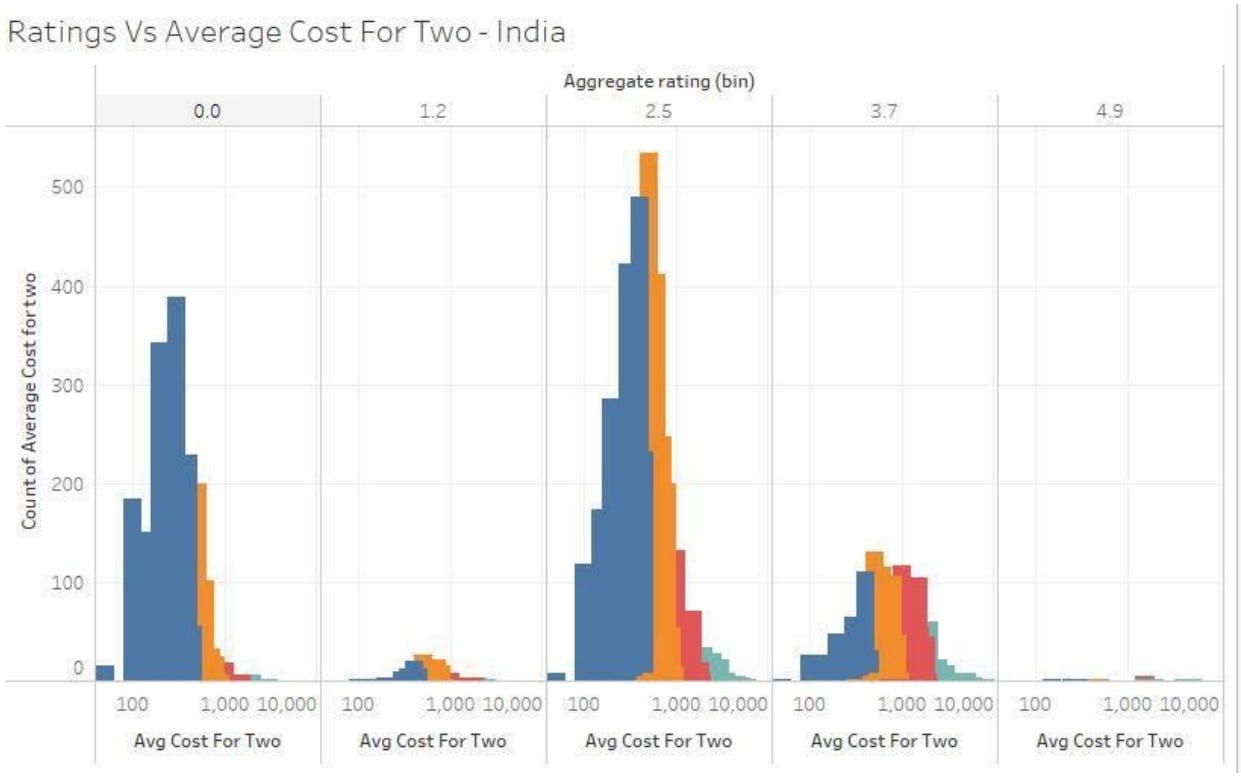
Distribution of average cost for in india based on price range.

1.Blue – price range 1
2.Orange- price range 2
3.Red- price range 3.
4.Teal- price range 4

In india, for the price ranges 1,2,3,4 the average price is RS 284, 620, 1258, 2583 respectively. The distribution is shifting from 200 on th left about 2500 on the right.

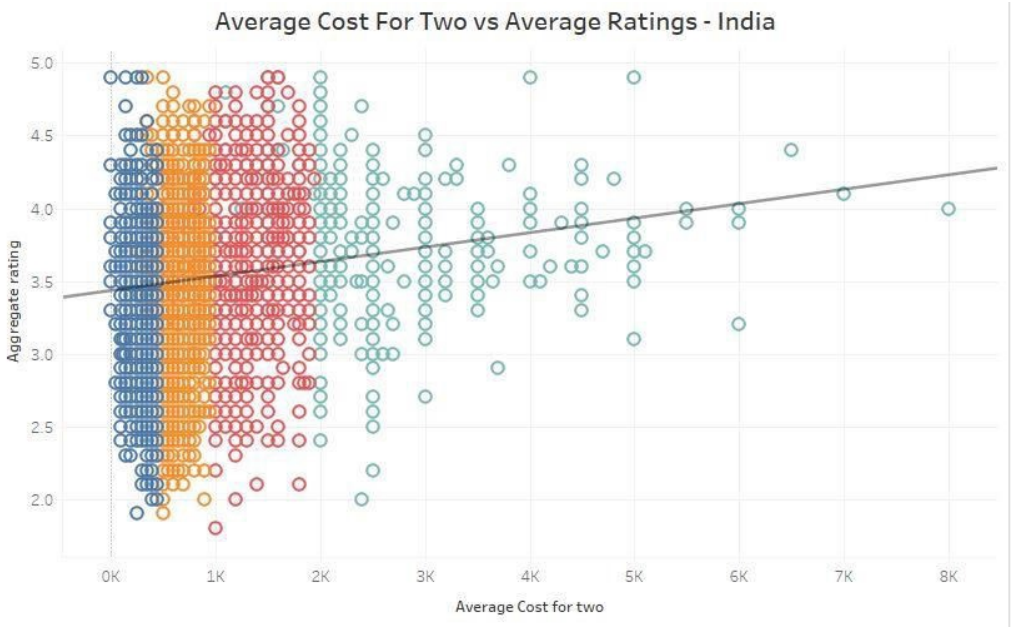
Question: Are price Range 1 restaurants providing value to zomato? Justify, if no what could be dome to increase the aov of the price range 1 restaurants?

Distribution of ratings across average cost for two in India

1. Blue – Price Range 1
2. Orange – Price Range 2
3. Red – Price Range 3
4. Teal – Price Range 4
5. 

The majority of ratings are between 2.5 to 3.7, and the rests are 0. As the rating increases, there is a slight shift in average cost for two as well. For ratings 0, the average cost is at about Rs 300-400 range but for rating 3.7 the average cost moves up to Rs 1000. There is a positive relationship between the

two. This can be explored using a scatter plot as well.



Question: Devise an A/B to analyze the impact of segregating ratings such that there are separate ratings for online delivery and dine-in. As these two are very different experiences, the rating system needs to be robust to provide holistic information yet be able to differentiate the two experiences.

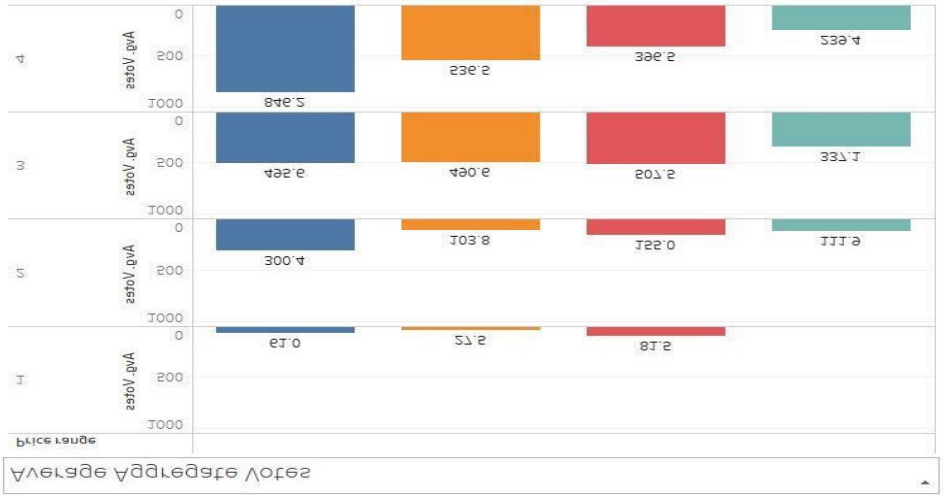
Analysis of Ratings, Votes and Costs for two across price range and new tag in India

1.Blue – Reservation and online ordering available
2.Orange – No reservation and no delivery (Zomato listing, advertisements)
3.Red – No reservation and only delivery

As the price range increases from 1 to 4 the rating gradually increases as well, in line with the previous scatter plot. Higher rating are for restaurants with online delivery (red) and delivery + reservations(blue). As people provide ratings on the app right after ordering food or before the next order this could explain the higher ratings.

Ratings of dine in restaurants without reservations or online delivery increases drastically with a price range from 1.7 (local dine-in) to 3.8 (premium dine-in chains), the highest increases across the 4 categories.

One intriguing observation – price range 1 with both facilities has an average rating of 3.7 and this fails to follow the overall trend observed previously. More data could help explain this outlier.



It’s a known fact that premium restaurants with delivery get higher share ratings and this is true for vote count as well.

For price range 4 with reservation and no delivery, the average cost for two is Rs 2900 this is an exception, where as for the rest of the price ranges, the cost is more or else the same. Where as for the price range 4 dine in costlier, this could happen because customers want to enjoy higher priced cuisines at the restaurants and only deliver lower prices items on the menu.

SUMMARY/RECOMMENDATIONS.

1.90% of observations belong to india.
2.USA has higher ratings than india and a majority of indian restaurants have 0 ratings. Either zomato should nudge customers to rate or these are newly onboarded restaurants.
3.US customers provide ratings more frequently and consistently than Indians.
4.Popular indian cuisines is north indian and foe USA its all Americana food
5.India and US have moderate food prices as compared to the rest of the countries.
6.Better affordability has led to restaurants offering in price range 1 but this could affects zomato's unit economics. The average price range is 1 is RS284, except this, all the rest price ranges provide positive returns based on previous analysis on unit economics.
7.Positive relations exist between ratings and the average cost for two, this could be because of the assumed notion thatpremium restaurants have good ambiance or better service, whereas regular restaurants cannot provide such an experience which could lead to lower ratings.
8.Provide reservations and online delivery to receive valuable customer feedback and ratings
9.Nudge price range 1 and 2 restaurants customers to vote more number of time.
10.Dine-in without reservation is cheaper for price range 1,2,3 restaurants as nocommision is charged.

END NOTES.

Data science has a direct impact on unit economics and in zomato case, its more true than ever.
The not-so-hyped aspects of data science are the ones that drive business growth and provide stability.

THANKYOU,
VIRENDRASINH MAFAJI RAJPUT