



INNOMATICS

RESEARCH LABS



Analysis on wedding halls



OBJECTIVES:

- *The primary objective of the study is to get to know about the cost sufficient halls in hyderabad ,to do the event of marriage .*
- *To know about the halls in hyderabad they are different sites available . I had gone through the website of WedMeGood.*
- *To know the best halls available :*
- *We need to know about the price,ratings,Location & hall details.*
- *We need to full-fill the requirements as per the client.*

What is Web Scraping ?

- ❖ Web scraping is an automatic method to obtain large amounts of data from websites. Most of this data is unstructured data in an HTML format which is then converted into structured data in a spreadsheet or a database so that it can be used in various applications.



This is how site looks like :

Locality ▾ No. of Guests ▾ Price per plate ▾ **Venue Type** ▾ Space ▾ Room Count ▾ Rating ▾ Review count ▾

4 Star & Above Hotels in Hyderabad
Showing 65 results as per your search criteria

Need Help? WMG Genie can help out! ▾ Search Wedding Venues... ▾ List Grid

4 Star & Above Hotels Clear all



Taj Deccan ★ 4.8 (6 reviews)
Hyderabad 4 Star & Above Hotels, Banquet Halls
Veg ₹ 2,950 per plate Non veg ₹ 3,304 per plate
Indoor Outdoor +3 more



Hotel Baseraa ★ 5.0 (1 review)
Secunderabad, Hyderabad 4 Star & Above Hotels, Banquet Halls, Hotels
Veg ₹ 500 per plate Non veg ₹ 600 per plate
50-1200 pax Indoor +2 more



Radisson Hyderabad ★ 5.0 (2 reviews)
Gachibowli, Hyderabad 4 Star & Above Hotels, Banquet Halls, Hotels
Veg ₹ 1,000 per plate Non veg ₹ 1,200 per plate
Indoor Inhouse catering only +1 more



Novotel Hyderabad Convention Centre ★ 5.0 (3 reviews)
Kothaguda, Hyderabad 4 Star & Above Hotels, Banquet Halls, Hotels
Veg ₹ 2,800 per plate Non veg ₹ 3,200 per plate



B Plus L ★ 5.0 (5 reviews)
Kothaguda, Hyderabad 4 Star & Above Hotels, Banquet Halls, Hotels
Veg ₹ 1,000 per plate Non veg ₹ 1,200 per plate



Taj Banjara ★ 4.8 (2 reviews)
Banjara Hills, Hyderabad 4 Star & Above Hotels, Banquet Halls, Hotels
Veg ₹ 1,300 per plate Non veg ₹ 1,500 per plate
Chat with us

Elements Console ▾ 4 6 ⚙

```
<div class="vendor-card extra-radius">
  <div class="flex">
    <div class="vendor-picture margin-l-10 margin-t-10" style="border-radius: 10px;">
      <a target="_blank" href="/wedding-venues/Radisson-42073">
        
      </a>
    </div>
    <div class="info-icon pointer" data-tip="true" data-for="10008" src="https://images.wedmegood.com/react-frontend-v4/statistics/media/NewInfoIconCropped.5709ead4.png" alt="currentitem="fa> ...
  </div>
</div>
 ...
... ).margin-l-10.margin-t-10 a img#vendor-card-2-object-fit-cover.w-100 ...
```

Styles Computed Layout Event Listeners DOM Breakpoints ▾

Filter :hover .cls +

Console What's New X

Highlights from the Chrome 105 update

Step-by-step replay in the Recorder panel

 new

Libraries used in the Analysis & scrapping :

```
In [1]: import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
In [ ]: !pip install selenium
```

```
In [1]: from selenium import webdriver
```

```
In [2]: path=r"C:\Users\LENOVO\Downloads\chromedriver_win32\chromedriver.exe"
```

```
In [3]: driver=webdriver.Chrome(path)
```

```
In [4]: website="https://www.wedmegood.com/vendors/hyderabad/wedding-venues/"  
driver.get(website)  
driver.maximize_window()
```



❖ This is how uncleaned data frame looks like:

```
In [78]: df.head()
```

```
Out[78]:
```

	Hall_name	Food_Price	rating_reviews	Seating	Location	hall_deatils
0	Dream Valley Resorts	Veg\n1,500\nNon veg\n1,800	4.0\n(9 reviews)	100-5000 pax\nIndoor\n+5 more	Moinabad, Hyderabad	Banquet Halls, Lawns / Farmhouses
1	Officers Mess Function Plaza	Veg\n400\nper plate\nNon veg\n500\nper plate	5.0\n(1 review)	600-900 pax\nIndoor\n+2 more	Hyderabad	Banquet Halls
2	Radon Convention Hall	Veg\n500\nper plate\nNon veg\n600\nper plate	5.0\n(1 review)	500-800 pax\nIndoor\n+2 more	Falaknuma, Hyderabad	Banquet Halls
3	Sitara Paradise Ameerpet	Veg\n300\nper plate\nNon veg\n500\nper plate	5.0\n(1 review)	100-300 pax\nIndoor\n+2 more	Punjagutta, Hyderabad	Banquet Halls, Lawns / Farmhouses
4	KLCC Conventions	On Request	5.0\n(1 review)	500-8000 pax\nIndoor\n+3 more	Shamshabad, Hyderabad	Banquet Halls, Lawns / Farmhouses

```
In [79]: df.shape
```

```
Out[79]: (600, 6)
```

```
In [80]: df.to_csv('wedding_hall_project_final.csv',index=False)
```

```
In [81]: driver.quit()
```

★ Challenges faced in web scraping & data cleaning :

- The website was not allowing us to extract the data from next page , so i did it by executing the program in different cell.
- In data-cleaning i faced the problems on :
 - 1) nan values
 - 2) duplicate values
 - 3) data-types
 - 4) outliers
 - 5) adding extra columns to clean the data.
- Removing the special characters
- Applying statistical methods to study given data and interpreting the results .

Raw data (Vs) clean Data.

In [78]: df.head()

Out[78]:

	Hall_name	Food_Price	rating_reviews	Seating	Location	hall_deatils
0	Dream Valley Resorts	Veg\n1,500\nNon veg\n1,800	4.0\n(9 reviews)	100-5000 pax\nIndoor\n+5 more	Moinabad, Hyderabad	Banquet Halls, Lawns / Farmhouses
1	Officers Mess Function Plaza	Veg\n400\nper plate\nNon veg\n500\nper plate	5.0\n(1 review)	600-900 pax\nIndoor\n+2 more	Hyderabad	Banquet Halls
2	Radon Convention Hall	Veg\n500\nper plate\nNon veg\n600\nper plate	5.0\n(1 review)	500-800 pax\nIndoor\n+2 more	Falaknuma, Hyderabad	Banquet Halls
3	Sitara Paradise Ameerpet	Veg\n300\nper plate\nNon veg\n500\nper plate	5.0\n(1 review)	100-300 pax\nIndoor\n+2 more	Punjagutta, Hyderabad	Banquet Halls, Lawns / Farmhouses
4	KLCC Conventions	On Request	5.0\n(1 review)	500-8000 pax\nIndoor\n+3 more	Shamshabad, Hyderabad	Banquet Halls, Lawns / Farmhouses

Data frame after Data Cleaning

In [78]: df.head()

Out[78]:

	Hall_name	Location	hall_deatils	Veg	Non-Veg	reviews	Rating	min_Seating	max_Seating
0	Dream Valley Resorts	Moinabad, Hyderabad	Banquet Halls, Lawns / Farmhouses	1500	1800	9	4.0	100	5000
1	Officers Mess Function Plaza	Hyderabad	Banquet Halls	400	500	1	5.0	600	900
2	Radon Convention Hall	Falaknuma, Hyderabad	Banquet Halls	500	600	1	5.0	500	800
3	Sitara Paradise Ameerpet	Punjagutta, Hyderabad	Banquet Halls, Lawns / Farmhouses	300	500	1	5.0	100	300
4	KLCC Conventions	Shamshabad, Hyderabad	Banquet Halls, Lawns / Farmhouses	400	700	1	5.0	500	8000

Data Analysis :-

In [6]: df.info()

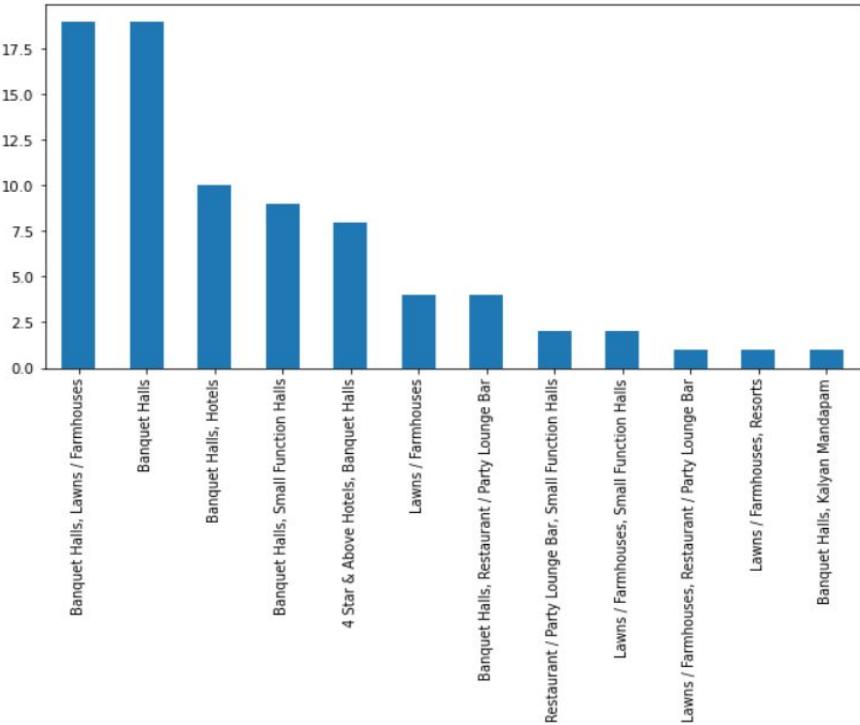
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 593 entries, 0 to 592
Data columns (total 9 columns):
 #   Column      Non-Null Count  Dtype  
 ---  --          --          --      
 0   Hall_name   593 non-null    object  
 1   Location    593 non-null    object  
 2   hall_deatails 593 non-null  object  
 3   Veg          593 non-null    int64  
 4   Non-Veg     593 non-null    int64  
 5   reviews      593 non-null    int64  
 6   Rating       593 non-null    float64 
 7   min_Seating 593 non-null    int64  
 8   max_Seating 593 non-null    int64  
dtypes: float64(1), int64(5), object(3)
memory usage: 41.8+ KB
```

In [7]: df.describe()

Out[7]:

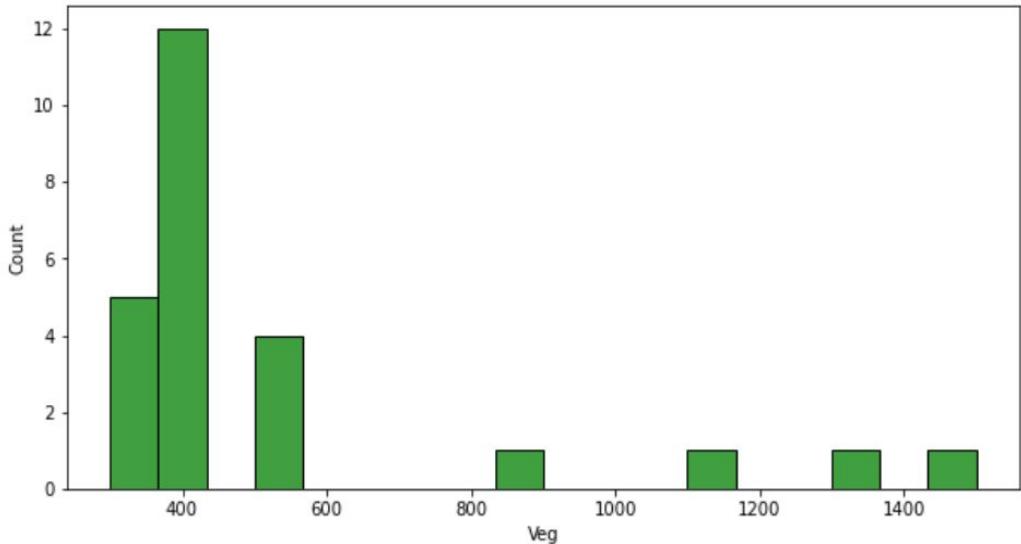
	Veg	Non-Veg	reviews	Rating	min_Seating	max_Seating
count	593.000000	593.000000	593.000000	593.000000	593.000000	593.000000
mean	467.516020	642.401349	0.281619	0.673187	313.212479	599.603710
std	284.825623	302.659681	0.956812	1.654327	264.960652	623.759691
min	80.000000	170.000000	0.000000	0.000000	15.000000	50.000000
25%	350.000000	450.000000	0.000000	0.000000	150.000000	250.000000
50%	400.000000	700.000000	0.000000	0.000000	250.000000	450.000000
75%	500.000000	700.000000	0.000000	0.000000	400.000000	700.000000
max	2950.000000	3304.000000	9.000000	5.000000	1575.000000	8000.000000

Hall Details Distribution.



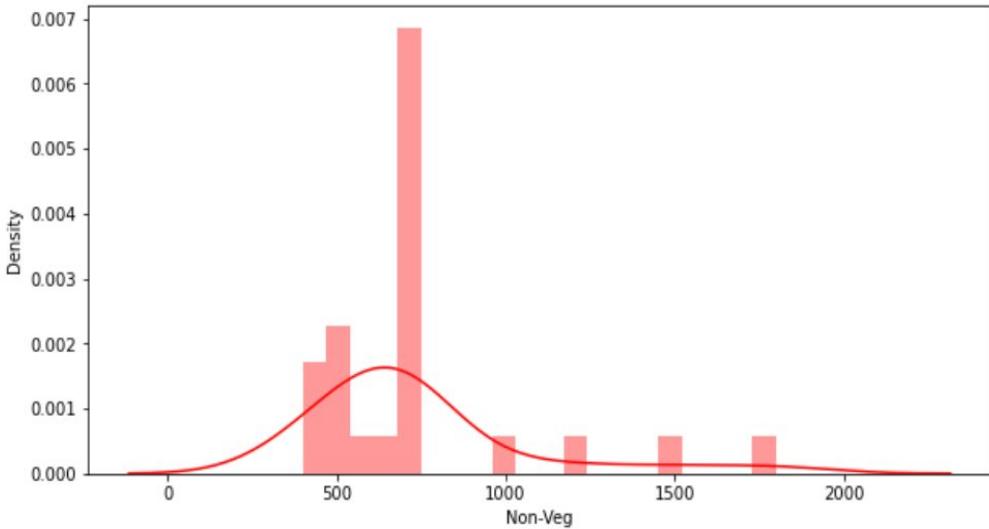
- Here we can see the distribution of the hall_details type.
- where maximum people are going through banquet halls,as compare to others.

Veg price Distribution.



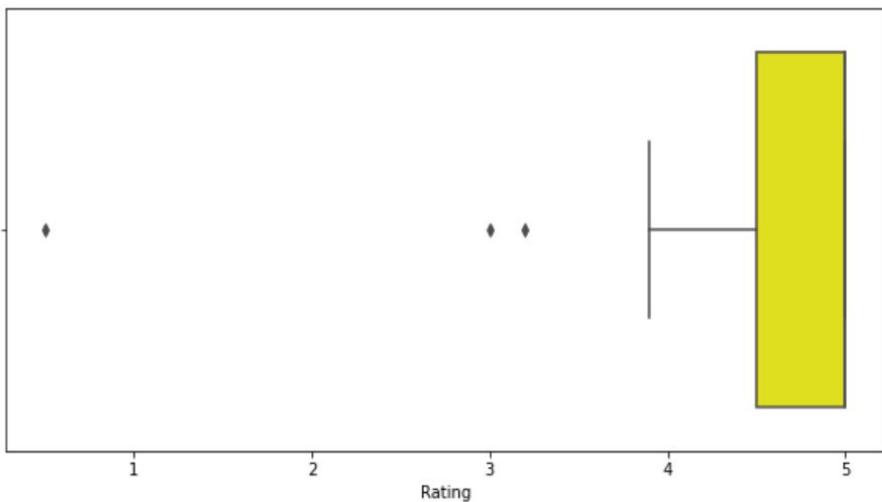
- ❖ here we can see the high no.of people are going through cost sufficient.
- ❖ as we know less no.of people prefer veg in compare to non-veg.
- ❖ approximate around(<30% of ppl will prefer Veg).

Non-Veg price Distribution.



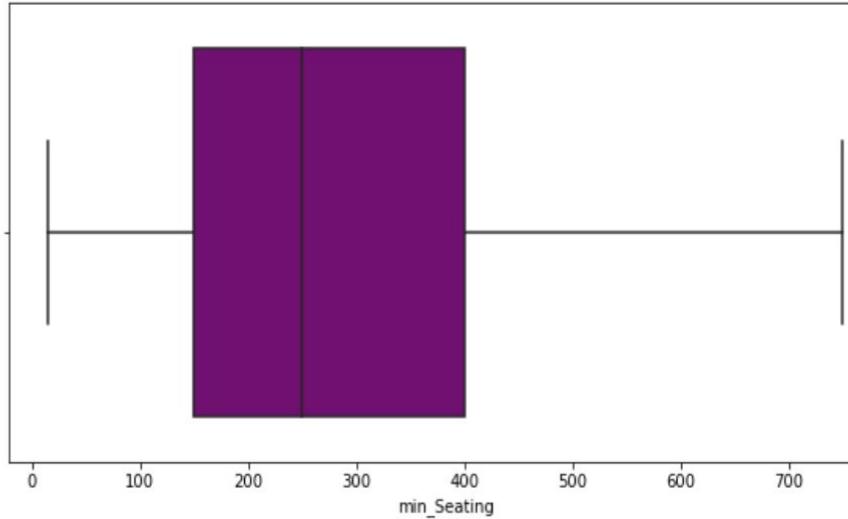
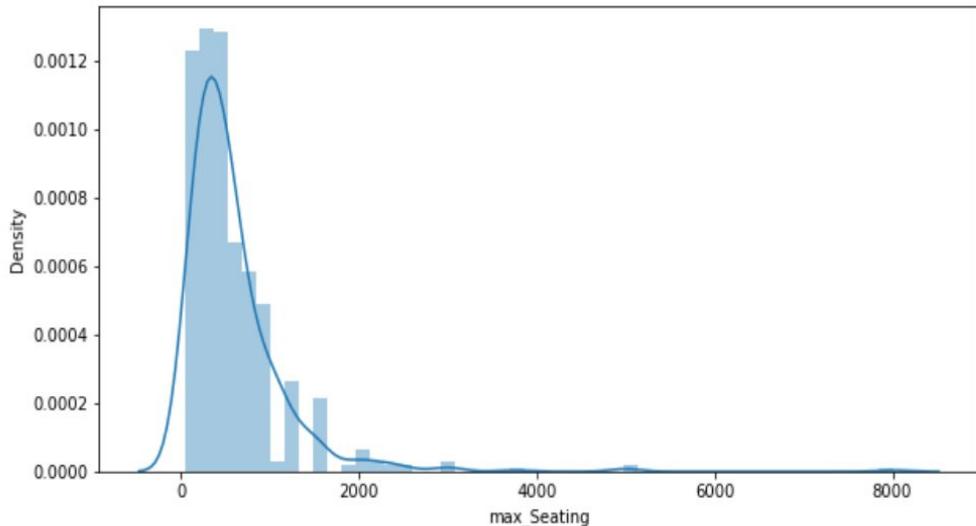
- ★ Here we can see that more cost is non-veg/plate .
- ★ but average price is among (500-800).
- ★ so in this case we should go with help of the reviews in taste.

Ratings.



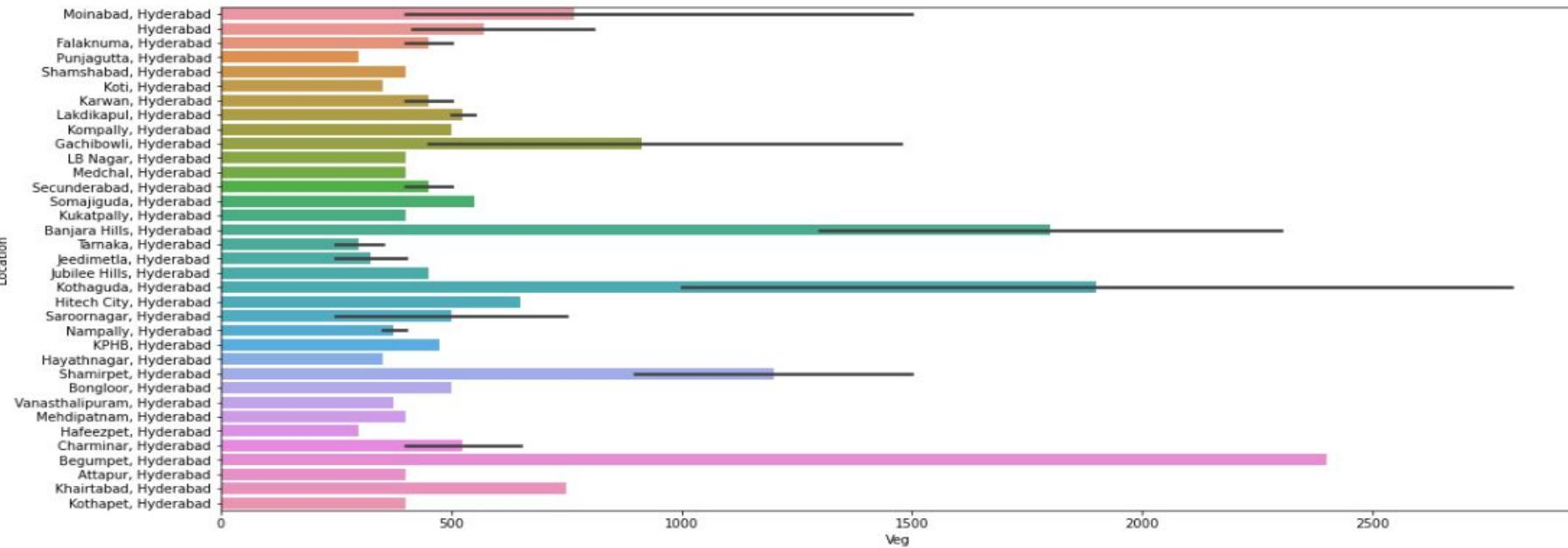
- Here we can see that ratings of all function halls
- majorly all are above 4-star rating. but some are having less ratings.
- we can see outliers in boxplot.

Min & Max Seating Distribution.



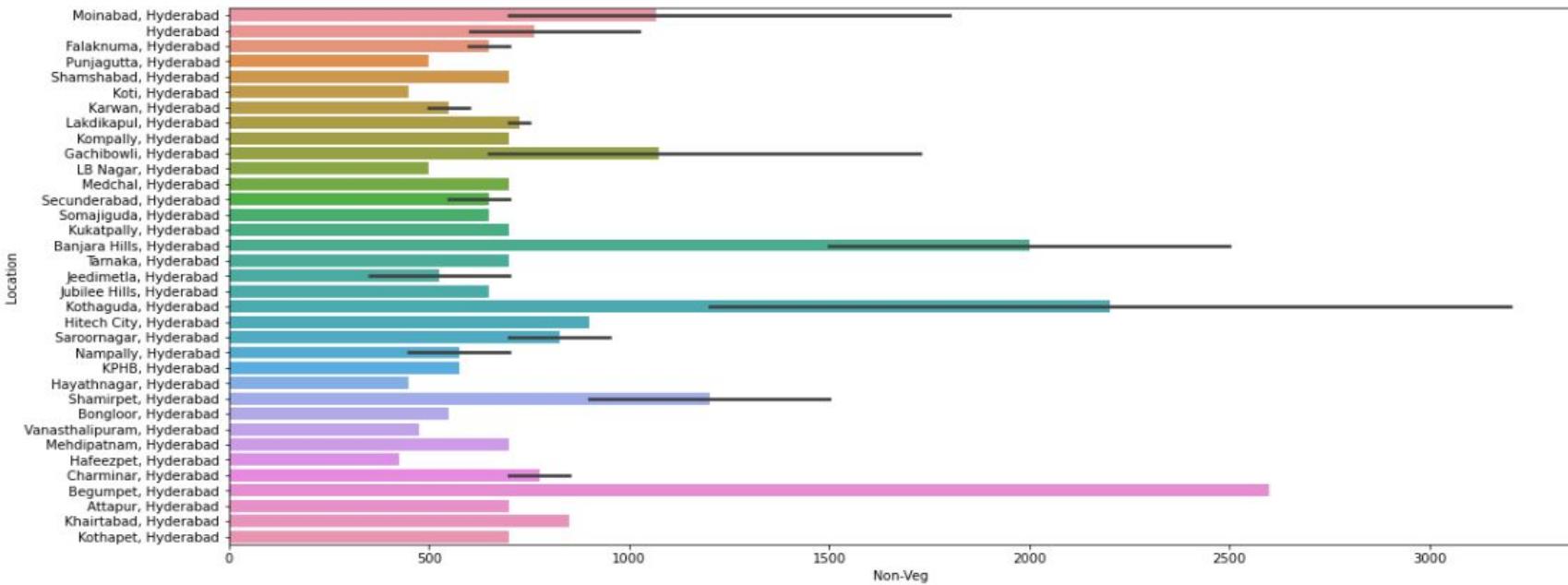
- Here we can see that max_Seatings is left skewed in nature.
- and min_seating is of range(100-700).but majority of them comes under 250-300 seatings.

Bivariate :- Location (vs) Veg price.



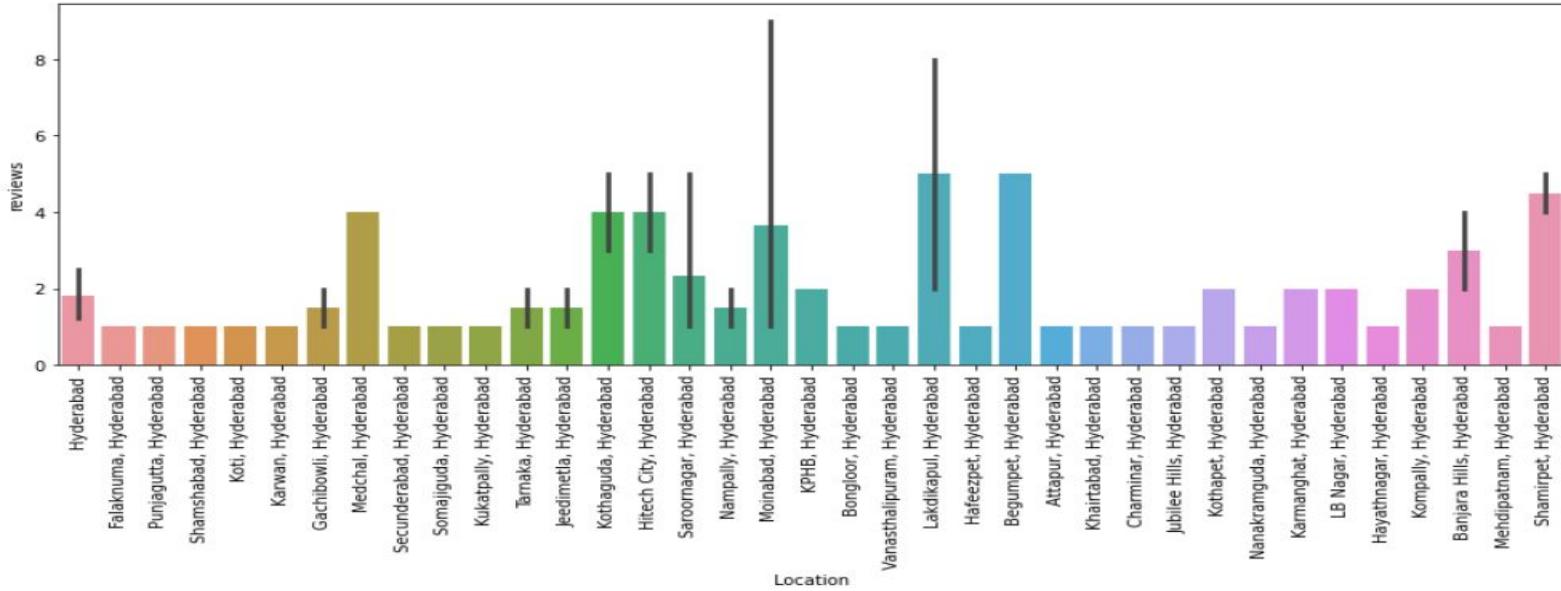
- here we can see that highest(veg-price/ plate)is kothaguda.and 2nd is banjara hills.
- we need to see for cost sufficient. (avg go for >500 is best)/plate.

Location (Vs) Non-Veg price.



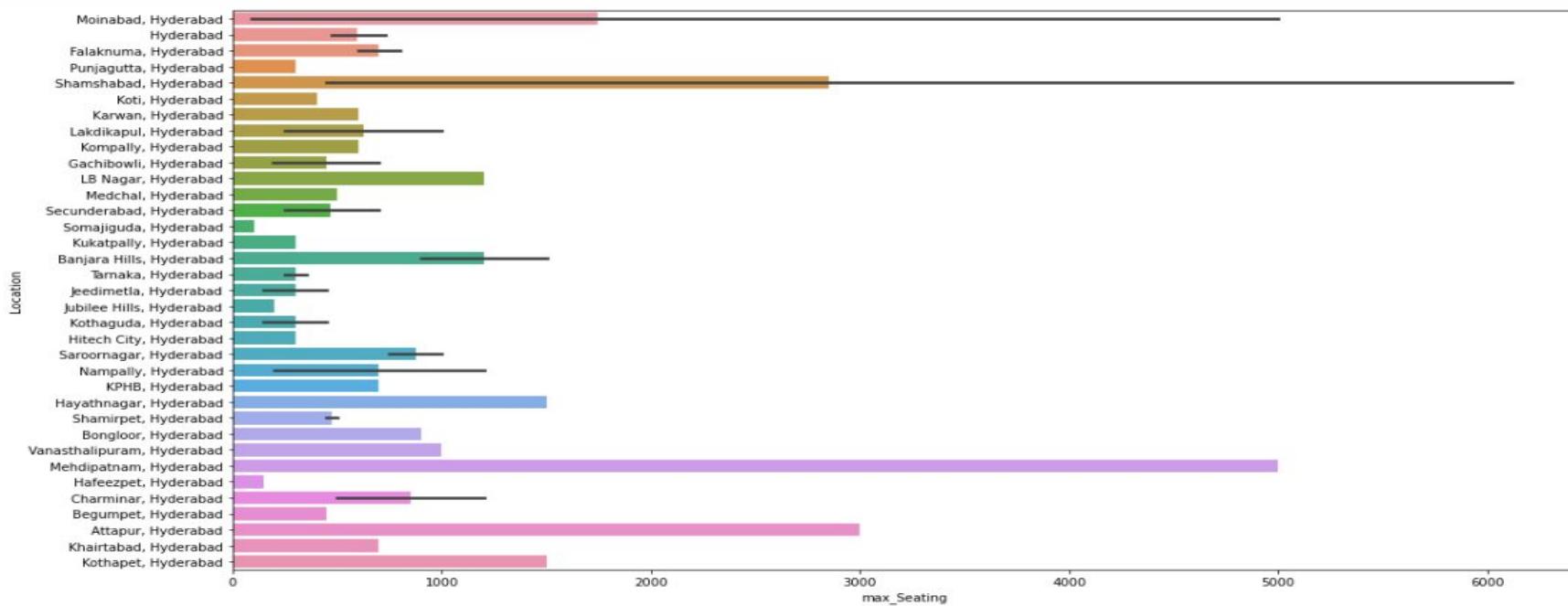
- As same as veg situation, we can see here.
- so most of people will prefer to go with non-veg. so, we have take it in Large quantity but for a reasonable price.
- so that, the price_range goes from (600-800)/plate.

Location (Vs) reviews



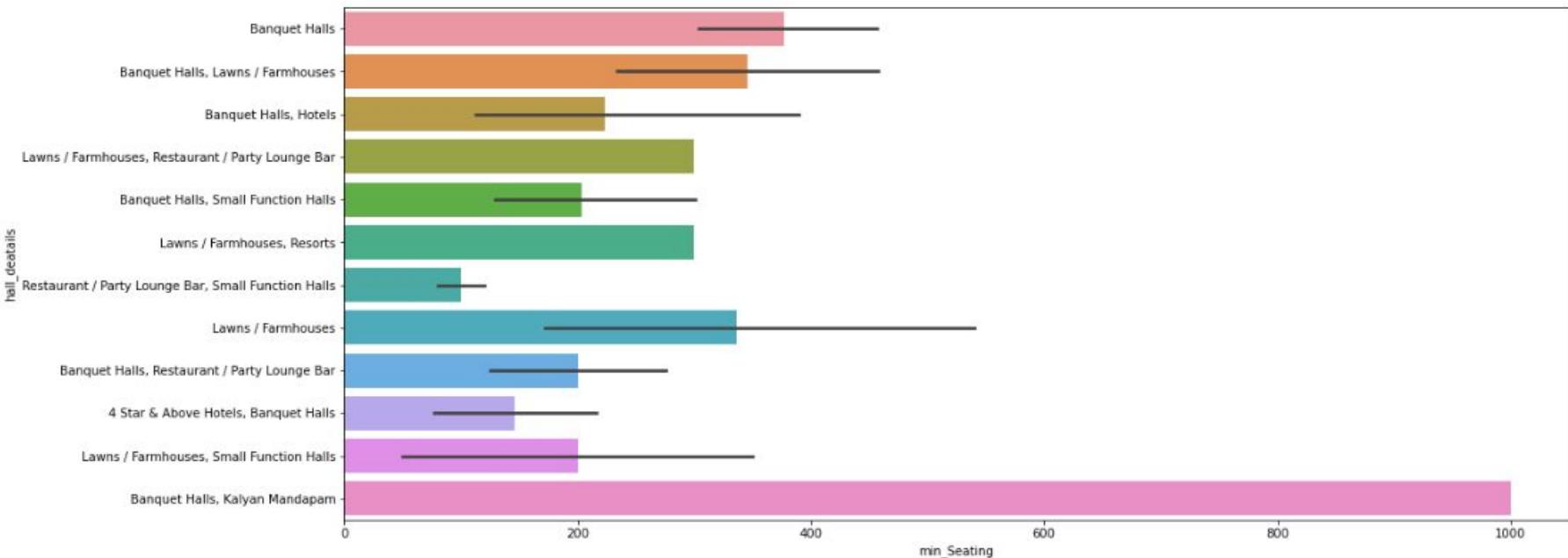
- here we can see that, most no.of people given review for moinabad area. and second is Lakdikapul area.
- were as kacheguda and banjara hills are given good reviews but not the best.

Location (Vs) max_Seating



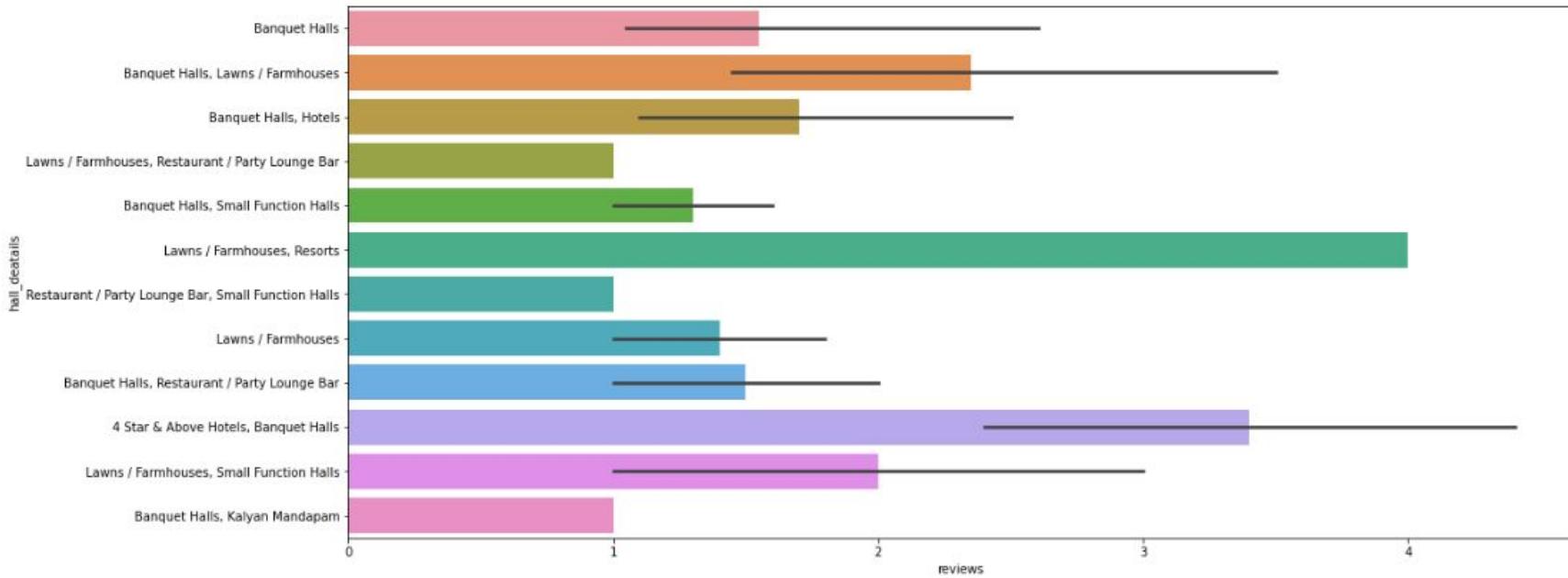
- ❖ As per the approximate people will come to hall for seating is 1000 because , the on going other events will maintain the out-flow & in-flow of the people .

Hall details (Vs) min_Seating.



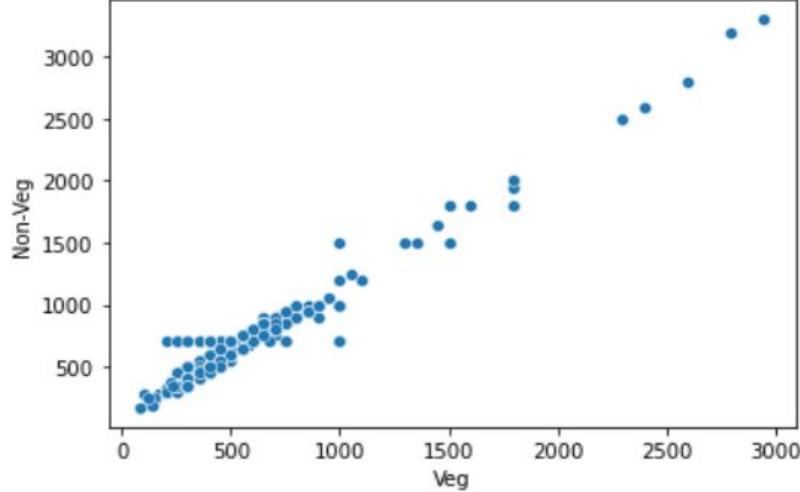
- We see that, most of the people are going through the (less than 300)Seatings in for small functions.

Hall details (Vs) reviews



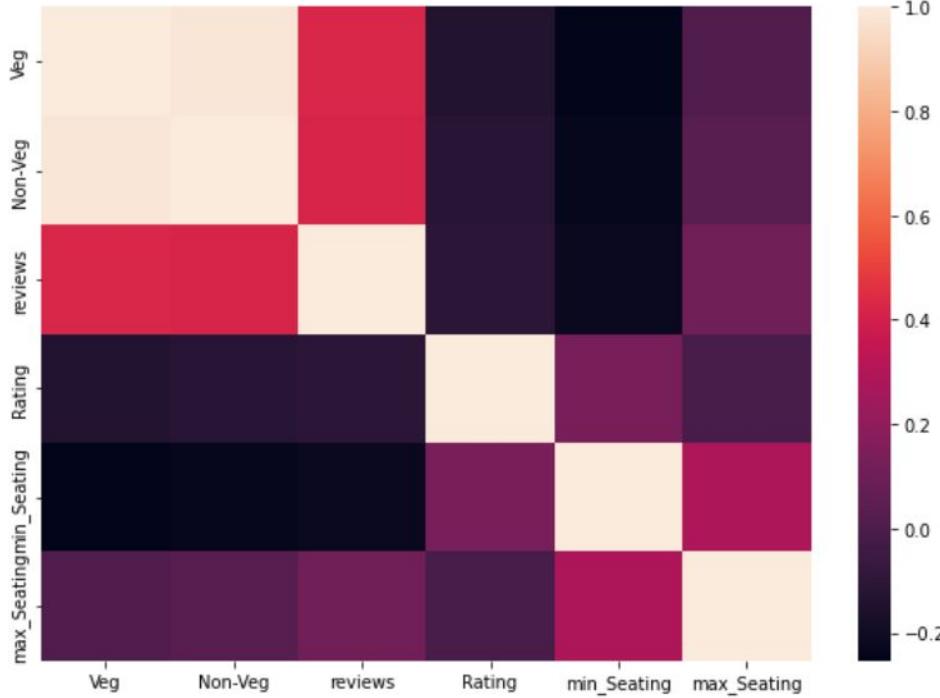
- hall_details which are 4 Star&above hotels,banquet halls. are the most preferred.among the people.
- and as well as lawns/Farmhouses,Resorts halls are preferred by people. (majorly they would be high class ppl)

Veg (Vs) Non-Veg.



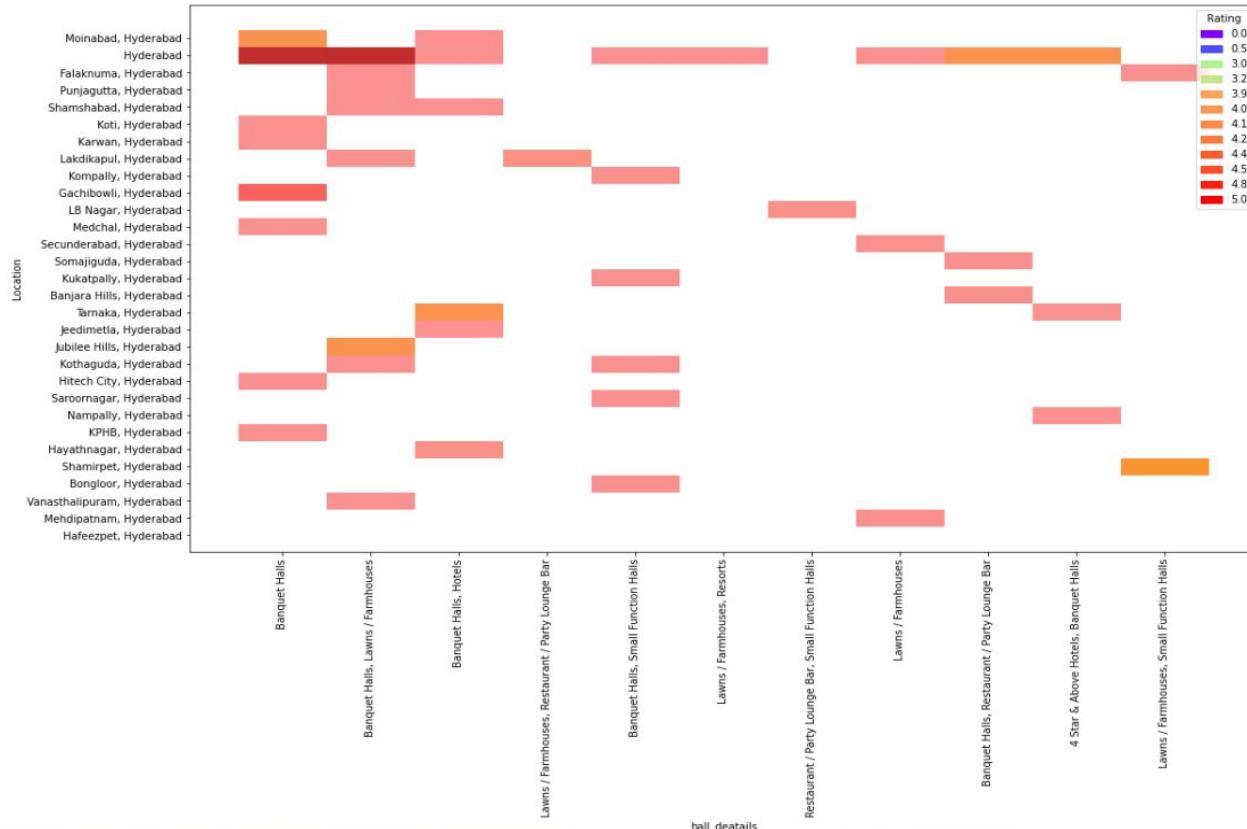
1. here we can see that , there is good correlation between veg price & non-veg price.
2. where we can clearly seen through the scatter plot.
3. it is gradually increasing in manner.

Correlation on Data-frame.



- ★ we can see a good correlation in veg & non-veg.
- ★ and moderate correlation in reviews&veg&non-veg.
- ★ remaining are majorly to +ve correlation side .

Multi-variate :-



❖ by seeing the plot.

we can say that hyderabad area is having highest ratings with hall_details of banquet halls among all others .

→ we explore more from this plot.

→ which area is best for doing events.

Client Requirement :-

→ The client requirement is :-

1. He needs to have good location as per the ratings and reviews.
2. And he wants the best hall_Type in banquet halls.
3. He said that 65% of the people are Non-Veg and 35% are Veg people.
4. He's Seating Capacity would be across 600.
5. This all requirement must be satisfied under the cost of 4 Lakhs.

Best output for client requirement.

```
best_halls=df_new2[
    (df_new2['Rating']>4)
    & (df_new2['hall_deatails'].str.contains('Banquet Halls','Banquet Halls, Lawns / Farmhouses'))
    & (df_new2['max_Seating'].between(600,800)
    & (df_new2['Client_Price']<400000.0))
]
```

```
best_halls
```

	Hall_name	Location	hall_deatails	Veg	Non-Veg	reviews	Rating	min_Seating	max_Seating	Client_Price
2	Radon Convention Hall	Falaknuma, Hyderabad	Banquet Halls	500	600	1	5.0	500	800	339000.0
6	Taj Convention	Karwan, Hyderabad	Banquet Halls	400	500	1	5.0	400	600	279000.0
8	Suvarna Kalyana Vedika	Hyderabad	Banquet Halls	400	700	1	5.0	550	800	357000.0
9	Hotel Kass	Kompally, Hyderabad	Banquet Halls, Hotels	500	700	2	4.8	400	600	378000.0
12	SR Convention	Gachibowli, Hyderabad	Banquet Halls, Small Function Halls	400	700	2	5.0	410	600	357000.0
15	Tivoli Gardens	Secunderabad, Hyderabad	Banquet Halls, Lawns / Farmhouses	400	700	1	5.0	500	700	357000.0
35	Vaaraahi Banquet Halls	KPHB, Hyderabad	Banquet Halls	475	575	2	5.0	100	700	324000.0
43	Noor Palace Function Hall	Falaknuma, Hyderabad	Banquet Halls, Small Function Halls	400	700	1	5.0	500	600	357000.0
61	Sri Laxminarayan Bhavan	Saroor Nagar, Hyderabad	Banquet Halls	250	700	1	5.0	500	750	325500.0
76	Ashoka Garden	Karwan, Hyderabad	Banquet Halls, Lawns / Farmhouses	500	600	1	5.0	500	600	339000.0

❖ Now:-
client can go with the choose he want.

Conclusion:-

- We can say that Location & Hall_type are Finalized by ratings and reviews & Seating's.
- Were as food is also finalized by reviews. so, here reviews& ratings are which plays major role in all the Accept's.
- If ratings&reviews are not available we can go with our own catering. and then hall price must based on hour-wise rule(or)day-wise rule. (apporox per day = 1 Lakh).
- Mostly halls are booked for Night time in compare to day time.(Price may various in Wedding season's)
- They are many more site for booking events. i felt WedMeGood is good see the surroundings near places in Hyderabad.

THANK
YOU

