INNOMATICS
RESEARCH LABS

WEB SCRAPING AND EXPLORATORY DATA ANALYSIS
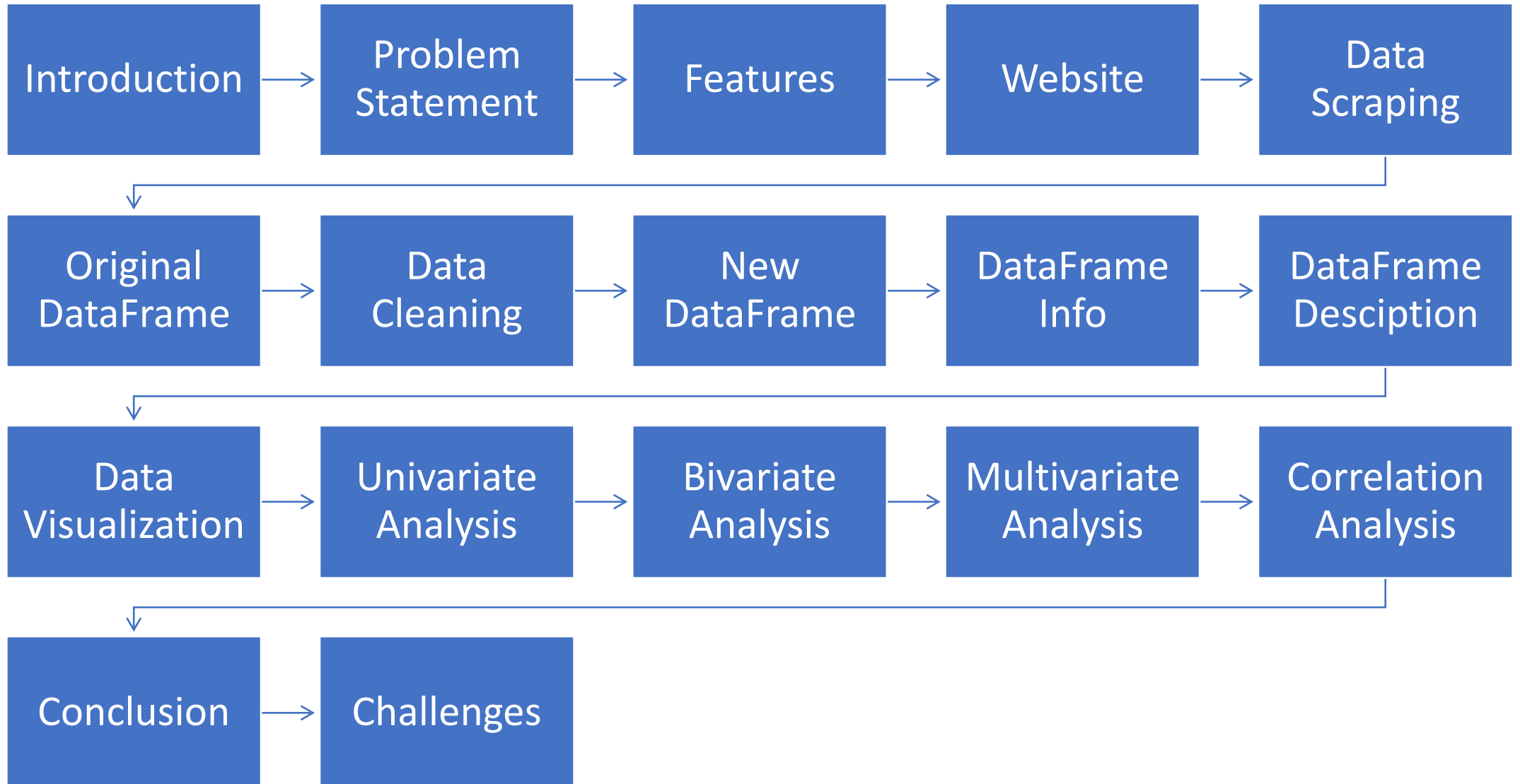
OYO Rooms

Presented By:

Tanvi Bakshi

Batch : 143

Coming Up Next..

Introduction →	Problem Statement →	Features →	Website →	Data Scraping
Original DataFrame →	Data Cleaning →	New DataFrame →	DataFrame Info →	DataFrame Desciption
Data Visualization →	Univariate Analysis →	Bivariate Analysis →	Multivariate Analysis →	Correlation Analysis
Conclusion →	Challenges			

Problem Statement

Comparative Analysis of different features between two major Cities in terms of OYO Rooms

- Hyderabad
- Delhi

OYO Website

OYO

Delhi ⌖ Near me | Tue, 17 May – Wed, 18 May | 1 Room, 1 Guest | **Search** | 🌐 EN 👤 Login / Signup

⚠ Please check the travel advisory issued by the concerned state government/local authorities before booking, as some places may have COVID-19 related travel/lodging restrictions.

Filters
Clear All

Price

₹450   ₹2855

Collections

☐ sanitised b4 ur eyes

☐ Vaccinated Staff

☐ Family OYOs

☐ For Group Travellers

☐ Local IDs accepted

+ View More

Categories

☐ **OYO Rooms** - Super affordable stays with essential amenities

☐ **Townhouse** - Your friendly, premium neighbourhood hotel- Serviced by OYO

☐ **Flagship** - Affordable hotels at Prime locations-

Hotels in Delhi (145 OYOs)

Map View 🔵   Sort By [Popularity ▾]

Earn ₹500
Create an account to get ₹500 as OYO Money

OYO 61891 Hotel Kabila
Karol Bagh, Delhi · 📍 5.0 km

🔥 13 people booked this hotel in last 6 hours

4.5 ★ (2414 Ratings) · Excellent

⊘ Reception   📶 Free Wifi   ⚡ Power backup   + 12 more

₹799 ₹2070 **61% off**
per room per night

[View Details] [Book Now]

OYO 70893 KI Khari Inn
Near Ashok Park Main, Rampura Marg, Delhi · 📍 6.7 km

🔥 8 people booked this hotel today

4.8 ★ (1141 Ratings) · Excellent

🚗 Parking Facility   ⊘ Reception   ⊘ Daily Housekeeping   + 7 more

Major Features for Analysis

- Offer Price

- Discounts

- Categories

- Original Price

- Ratings

- Features of the Hotels

Data Scraping

WEB SCRAPING

DataFrame

```
df.head()
```

	Hotel Name	Cost Per Night	Rating	Features	Location
0	OYO 28715 Acclaim	₹500₹196374% offper room per nightView Details...	4.5 (6 Ratings)·Excellent	Hair Dryer Reception Free Wifi + 15 more	Near Abdul Gaffar Khan Marg, Vasant Kunj, Delhi
1	Capital O 88089 Hotel Elite	₹549₹254778% offper room per nightView Details...	4.5 (74 Ratings)·Excellent	Parking Facility Reception Free Wifi + 14 ...	India, Delhi
2	OYO 87715 Comfort In	₹541₹202573% offper room per nightView Details...	4.6 (136 Ratings)·Excellent	Elevator Free Wifi Geyser + 7 more	India, Delhi
3	OYO 87875 Krazy Home	₹500₹192574% offper room per nightView Details...	4.5 (15 Ratings)·Excellent	Elevator Free Wifi Geyser + 6 more	India, Delhi
4	OYO 22355 Hotel Istay	₹699₹237170% offper room per nightView Details...	4.6 (176 Ratings)·Excellent	Reception Free Wifi Power backup + 13 more	Near New Delhi Railway Station, Delhi

```
df.head()
```

	Hotel Name	Cost Per Night	Rating	Features	Location
0	SPOT ON 37214 Classic Lodge	₹406₹119065% offper room per nightView Details...	4.5 (870 Ratings)·Excellent	AC Parking Facility Elevator + 8 more	Near Board Of Intermediate Office, Hyderabad· ...
1	OYO 24204 Nav Bharath Residency	₹978₹301767% offper room per nightView Details...	4.4 (988 Ratings)·Very Good	Reception Free Wifi Power backup + 17 more	Near Central Gurudwara Saheb, Gowliguda, Hyder...
2	OYO 35467 Shiridi Residency	₹1014₹222654% offper room per nightView Detail...	4.7 (937 Ratings)·Excellent	Reception Free Wifi Security + 10 more	Near Chermas, Abids, Hyderabad· 1.3 km
3	SPOT ON 75815 Hotel Kanha Residency	₹338₹123372% offper room per nightView Details...	3.8 (394 Ratings)·Good	TV	India, Hyderabad· 3.1 km
4	OYO 74978 Hotel Aadaab	₹646₹231772% offper room per nightView Details...	3.3 (240 Ratings)·Fair	Reception Free Wifi 24/7 Checkin + 12 more	Near Nampally railway station, Mahesh Nagar, H...

Data Cleaning

- Identifying Null Values
- Identifying Missing Values
- Treating rows and columns with null values and missing values using Mean, Median & Mode
- Data Mugging, using regular expressions
- Treating columns and converting required columns from object to int or float datatype

Real Python
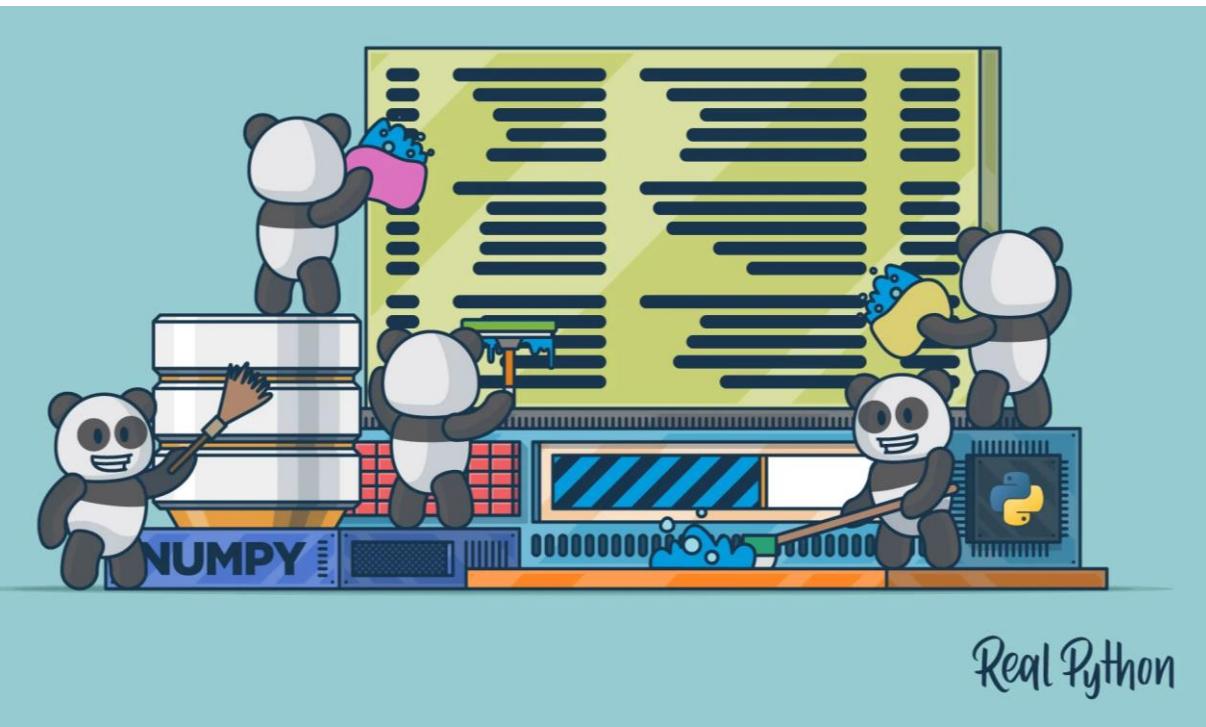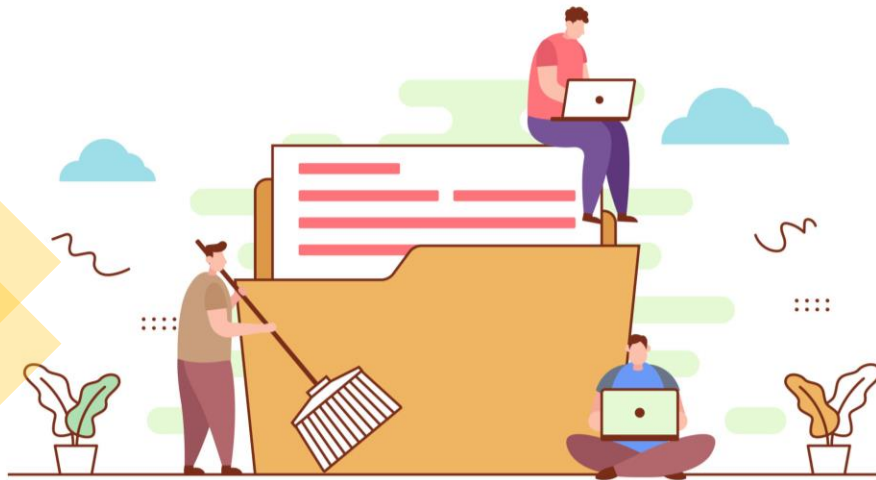
DataFrame

```
In [46]: d3.head()
```

	City	Category	Hotel_Name	Locality	Original Price	Discount	Offer Price	Star Rating	Rating	Remarks	Wi-Fi	CCTV Cameras	Doctor on Call	Parking Facility	Power backup	Elevator	Reception	AC
0	Hyderabad	SPOT ON	ON 37214 Classic Lodge	Near Board Of Intermediate Office	1190.0	65.0	406.0	4.5	870	Excellent	no	no	no	no	no	no	no	no
1	Hyderabad	OYO	24204 Nav Bharath Residency	Near Central Gurudwara Saheb	3017.0	67.0	978.0	4.4	988	Very	yes	yes	yes	yes	yes	yes	yes	yes
2	Hyderabad	OYO	35467 Shiridi Residency	Near Chermas	2226.0	54.0	1014.0	4.7	937	Excellent	yes	yes	yes	yes	yes	yes	yes	yes
3	Hyderabad	SPOT ON	ON 75815 Hotel Kanha Residency	India	1233.0	72.0	338.0	3.8	394	Good	no	no	no	no	no	no	no	no

```
In [97]: d3.shape
```

```
Out[97]: (326, 18)
```

```
d3.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 326 entries, 0 to 325
Data columns (total 18 columns):
 #   Column            Non-Null Count   Dtype
---  ------            --------------   -----
 0   City              326 non-null     object
 1   Category          326 non-null     object
 2   Hotel_Name        326 non-null     object
 3   Locality          326 non-null     object
 4   Original Price    326 non-null     float64
 5   Discount          326 non-null     float64
 6   Offer Price       326 non-null     float64
 7   Star Rating       326 non-null     float64
 8   Rating            326 non-null     int64
 9   Remarks           326 non-null     object
 10  Wi-Fi             326 non-null     object
 11  CCTV Cameras      326 non-null     object
 12  Doctor on Call    326 non-null     object
 13  Parking Facility  326 non-null     object
 14  Power backup      326 non-null     object
 15  Elevator          326 non-null     object
 16  Reception         326 non-null     object
 17  AC                326 non-null     object
dtypes: float64(4), int64(1), object(13)
memory usage: 46.0+ KB
```

DataFrame Info

# DataFrame Description Numerical Columns

```
d3.describe()
```

	Original Price	Discount	Offer Price	Star Rating	Rating
count	326.000000	326.000000	326.000000	326.000000	326.000000
mean	2051.950704	61.226481	838.000000	3.846497	420.892638
std	529.206196	9.497575	337.963293	0.597485	811.978930
min	1000.000000	47.000000	266.000000	1.600000	0.000000
25%	1797.000000	54.000000	625.000000	3.400000	4.000000
50%	2051.950704	61.226481	775.000000	3.900000	8.000000
75%	2260.250000	69.000000	947.500000	4.300000	433.750000
max	4441.000000	83.000000	2332.000000	4.900000	5131.000000

Data
Visualization

OYO are in Top Category

Category	
OYO	230
Collection O	29
SPOT ON	28
Capital O	28
Townhouse	11
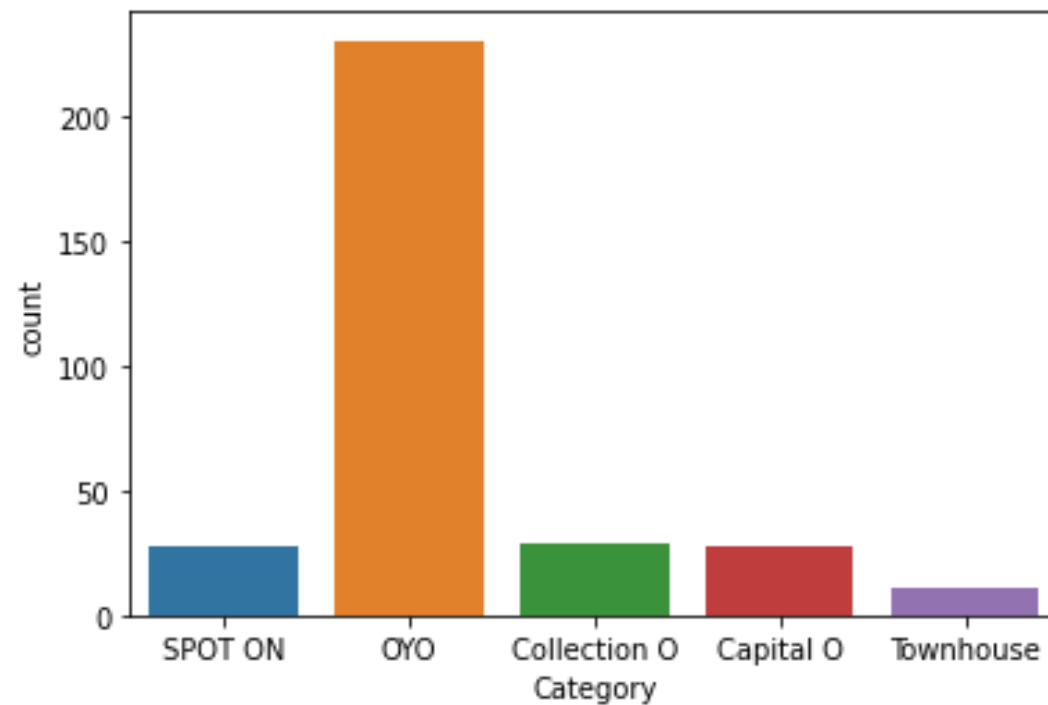
Delhi has a greater number of OYO's

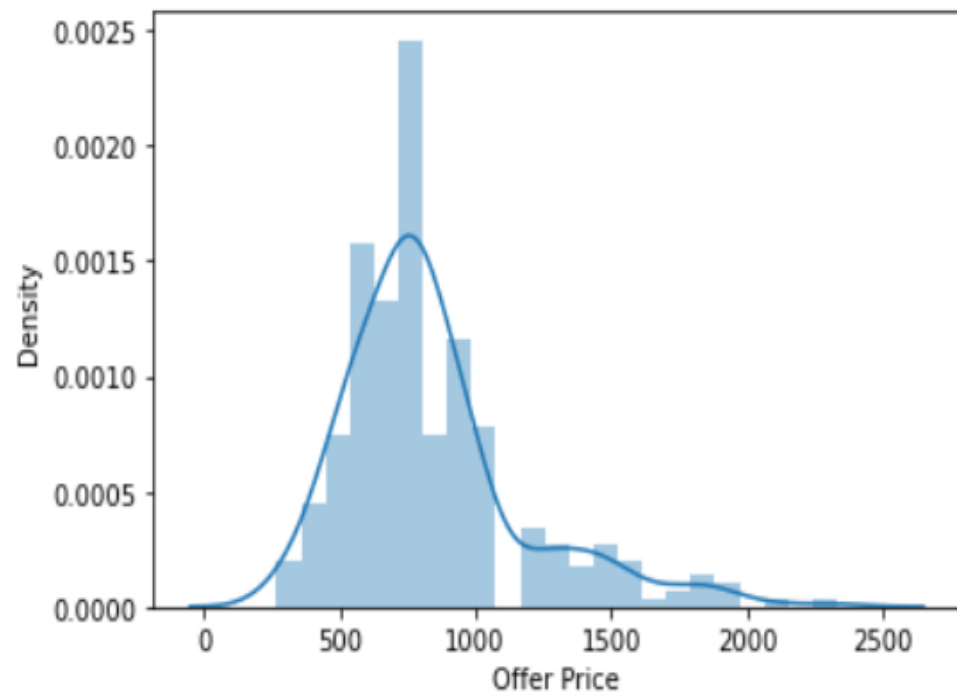| City	
Delhi	180
Hyderabad	146
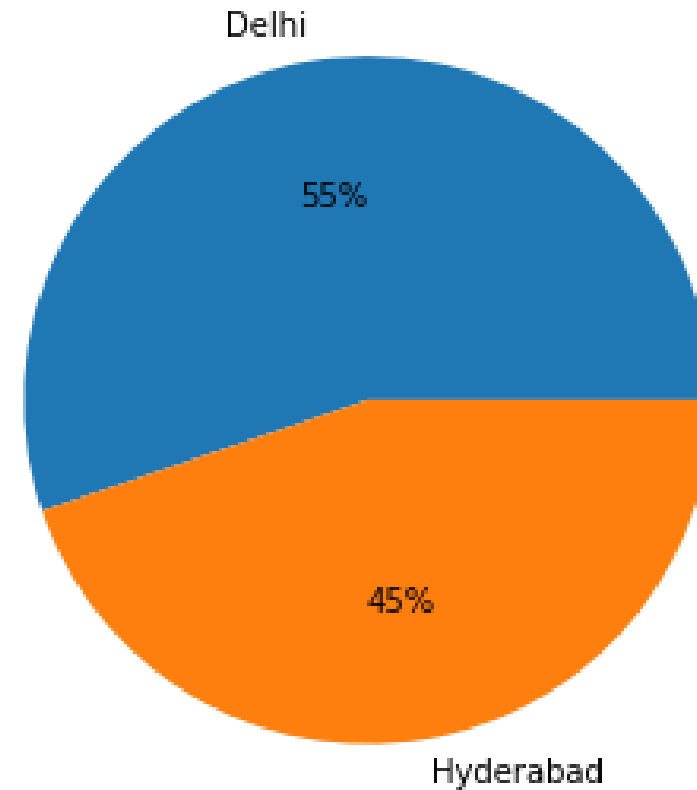
Original Price

```
count        326.000000
mean        2051.950704
std          529.206196
min         1000.000000
25%         1797.000000
50%         2051.950704
75%         2260.250000
max         4441.000000
Name: Original Price, dtype: float64
```
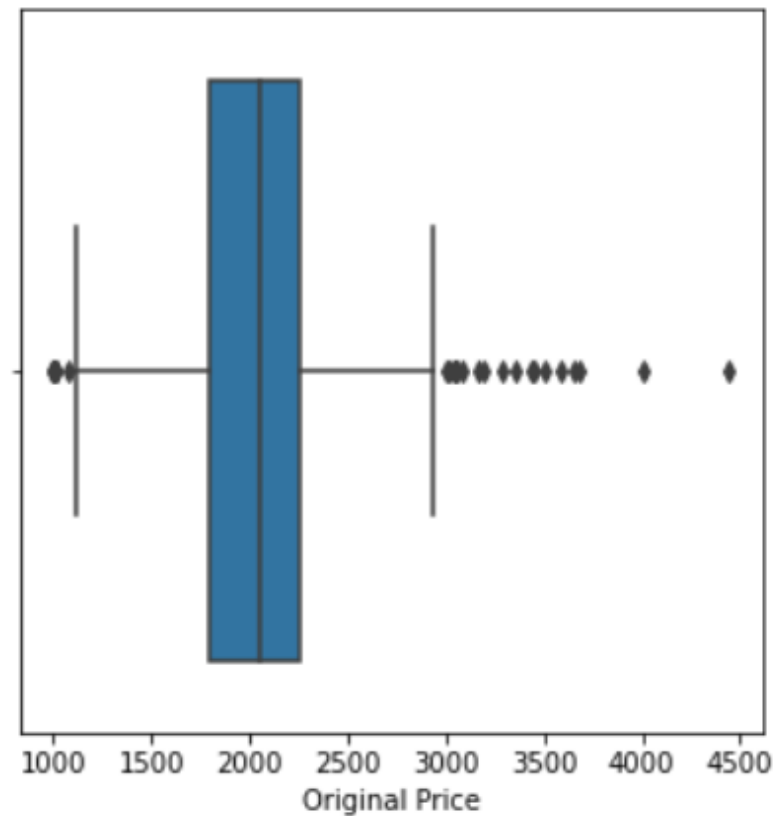
Offer Pricing


```
:   count        326.000000
    mean         838.000000
    std          337.963293
    min          266.000000
    25%          625.000000
    50%          775.000000
    75%          947.500000
    max         2332.000000
    Name: Offer Price, dtype: float64
```
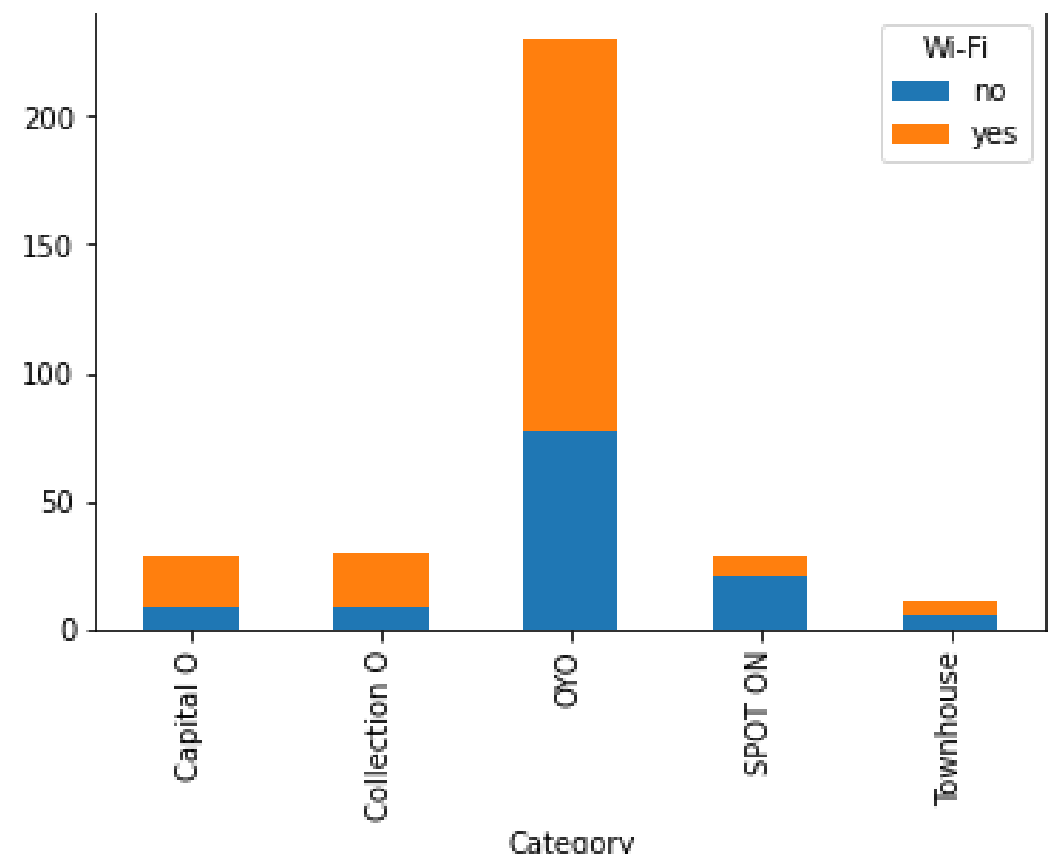
47% is the highest discount given
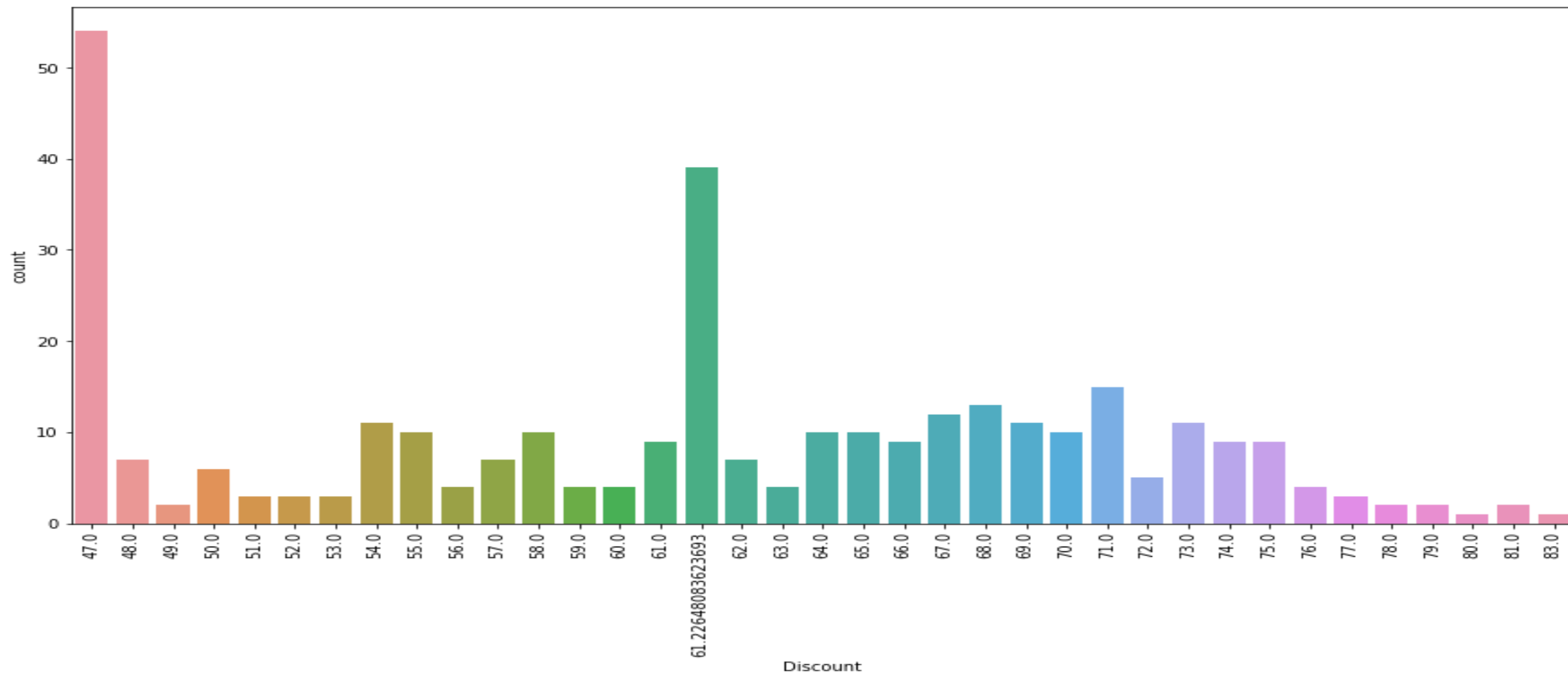
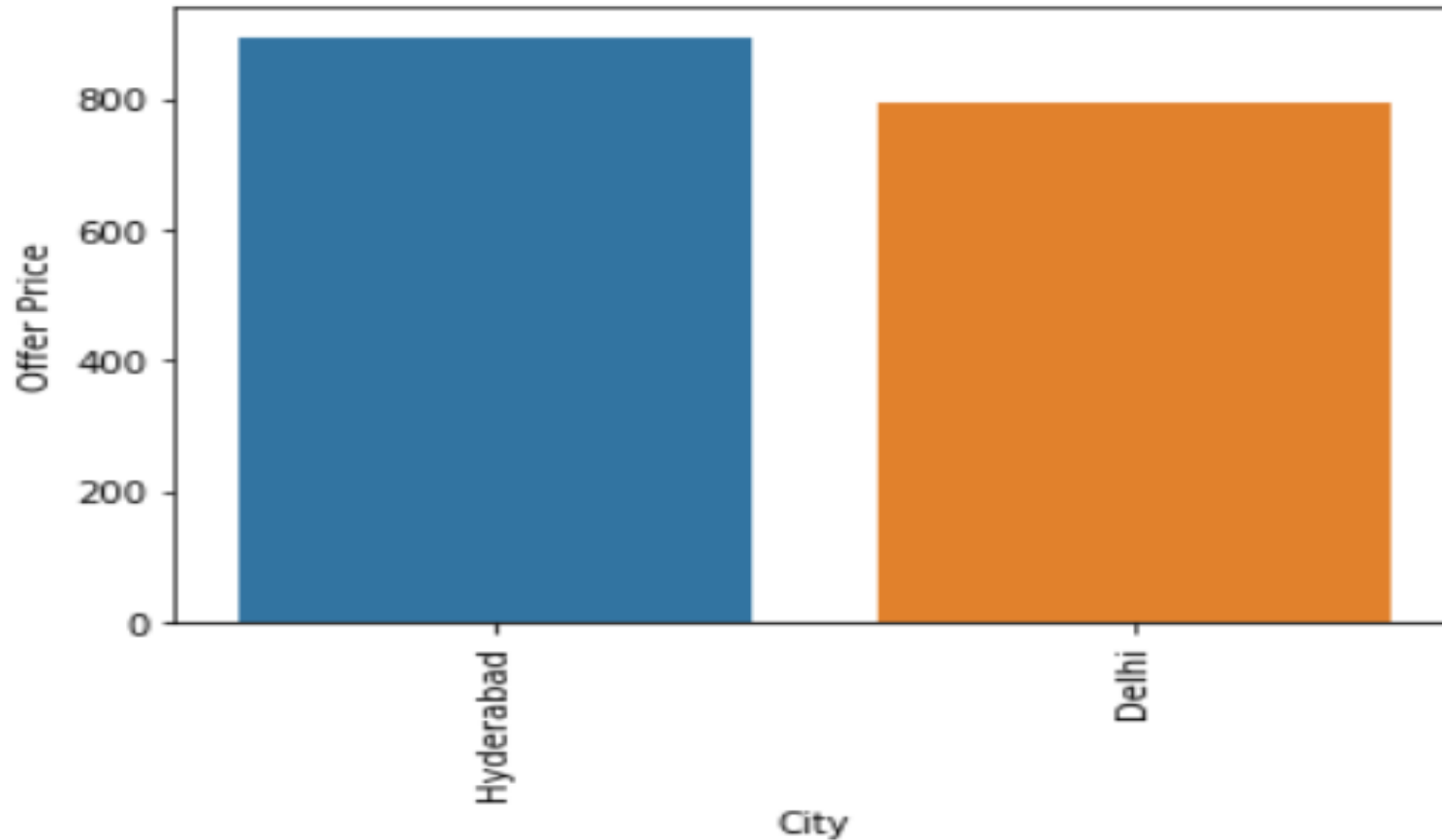# Collection O and Capital O category has the highest Original Price
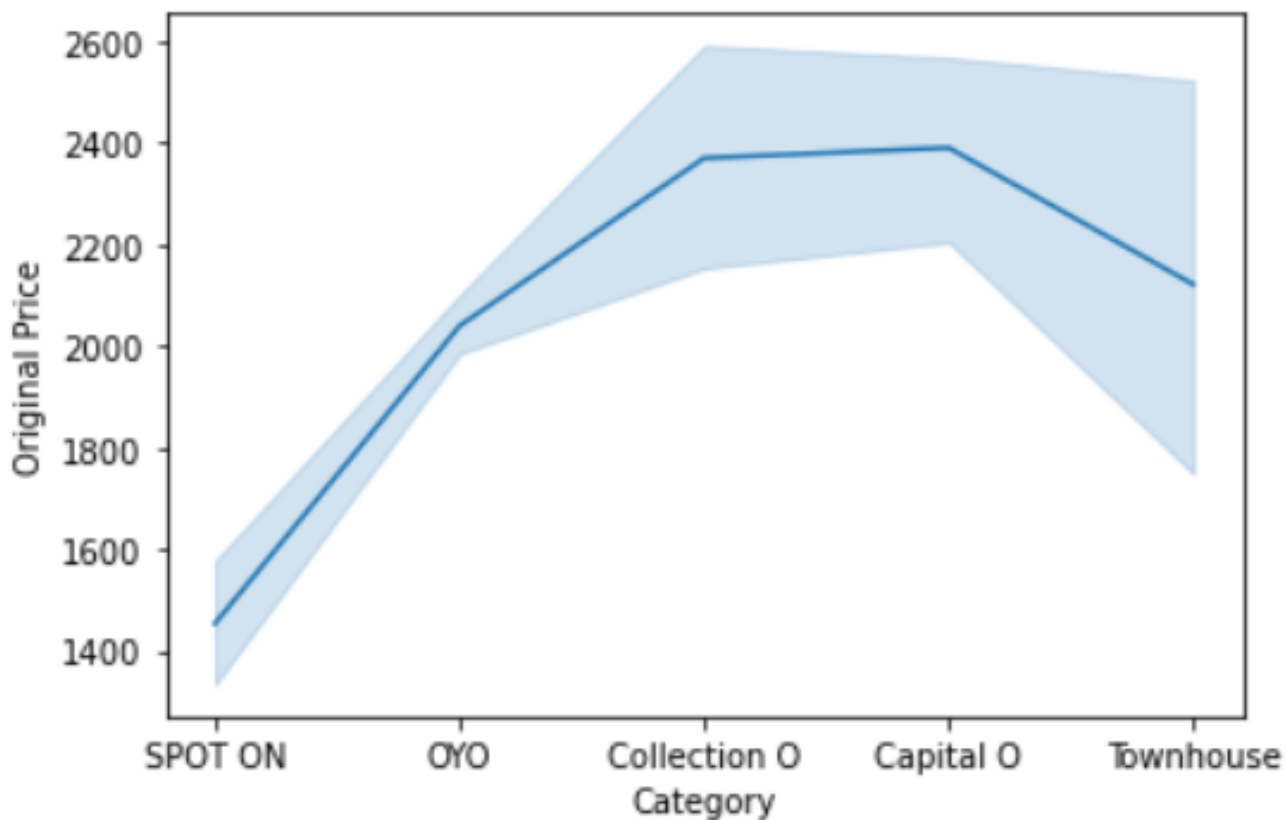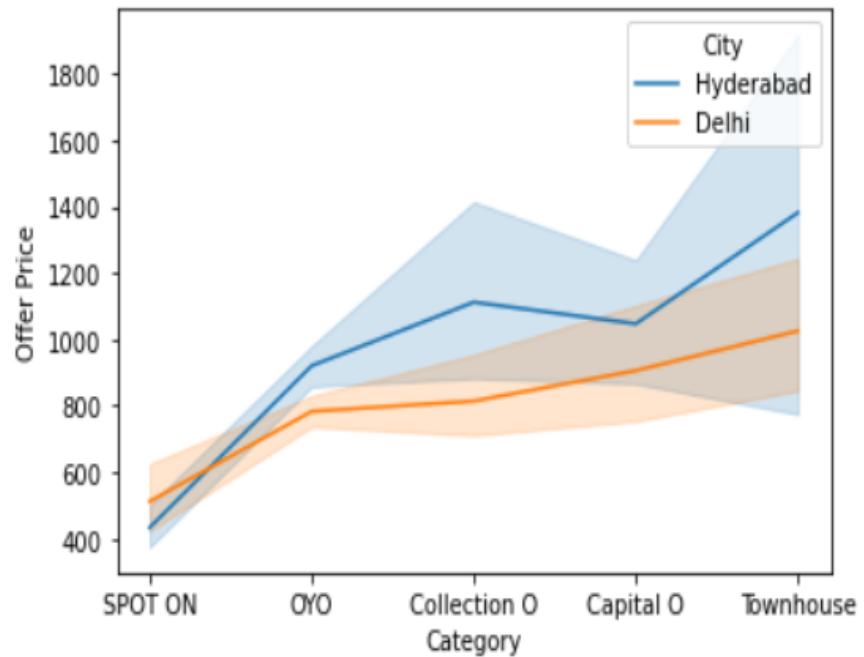
OYO category has WI-FI

Wi-Fi	no	yes
Category		
Capital O	8	20
Collection O	8	21
OYO	77	153
SPOT ON	21	7
Townhouse	5	6

Offer Price of Hyderabad is high as compared to Delhi

# Hyderabad City has more avg offer price OYOs than Delhi

Category	City	max	median	mean	min
Capital O	Delhi	1902.0	801.5	905.642857	549.0
	Hyderabad	1726.0	942.5	1046.642857	549.0
Collection O	Delhi	1684.0	740.0	813.705882	559.0
	Hyderabad	2332.0	902.0	1111.750000	664.0
OYO	Delhi	1807.0	699.0	782.787879	489.0
	Hyderabad	2107.0	775.0	920.244898	330.0
SPOT ON	Delhi	799.0	454.0	511.666667	364.0
	Hyderabad	778.0	406.0	432.631579	266.0
Townhouse	Delhi	1612.0	930.0	1025.125000	715.0
	Hyderabad	1917.0	1452.0	1381.333333	775.0

OYO has maximum ratings in comparison
Collection O
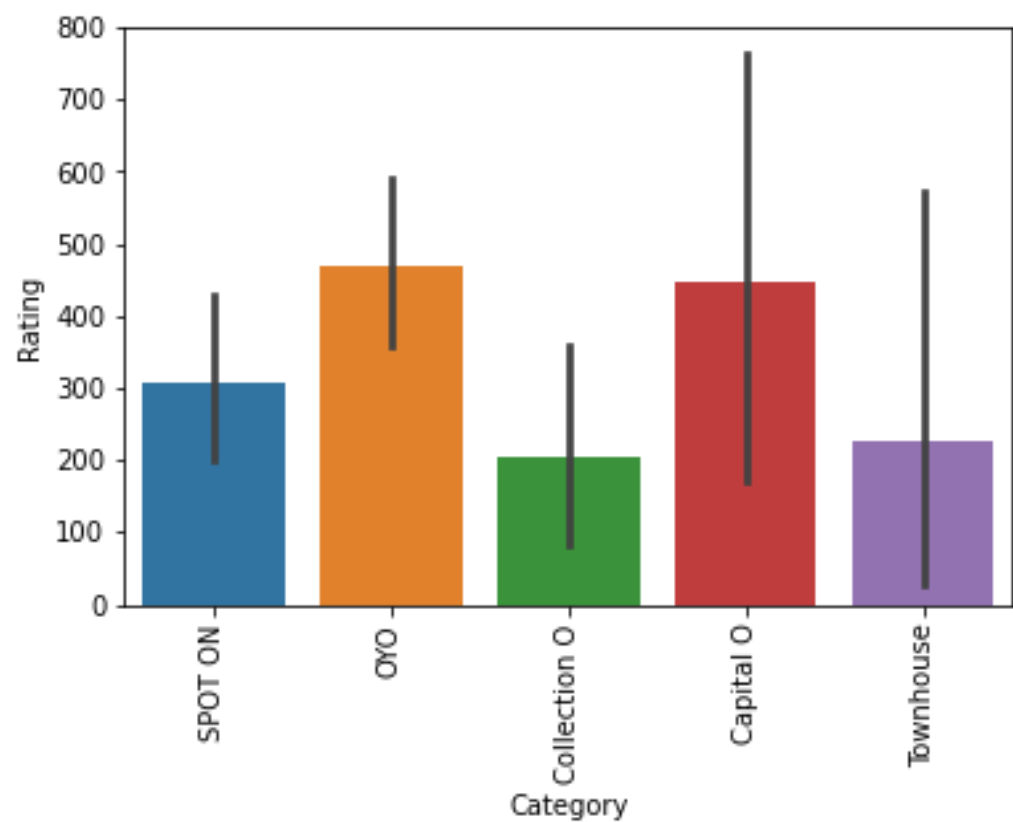
Category	City	median	mean	max	min
Capital O	Delhi	4.0	4.642857	9	0
	Hyderabad	516.0	887.071429	3194	8
Collection O	Delhi	2.0	3.882353	9	0
	Hyderabad	439.5	492.000000	1331	19
OYO	Delhi	5.0	5.030303	9	0
	Hyderabad	746.0	1091.153061	5131	8
SPOT ON	Delhi	4.0	4.000000	8	0
	Hyderabad	370.0	453.736842	1172	56
Townhouse	Delhi	6.5	5.875000	9	1
	Hyderabad	308.0	818.666667	1889	259

Hyderabad is positively related, as the offer price is constantly increasing as compared to original price
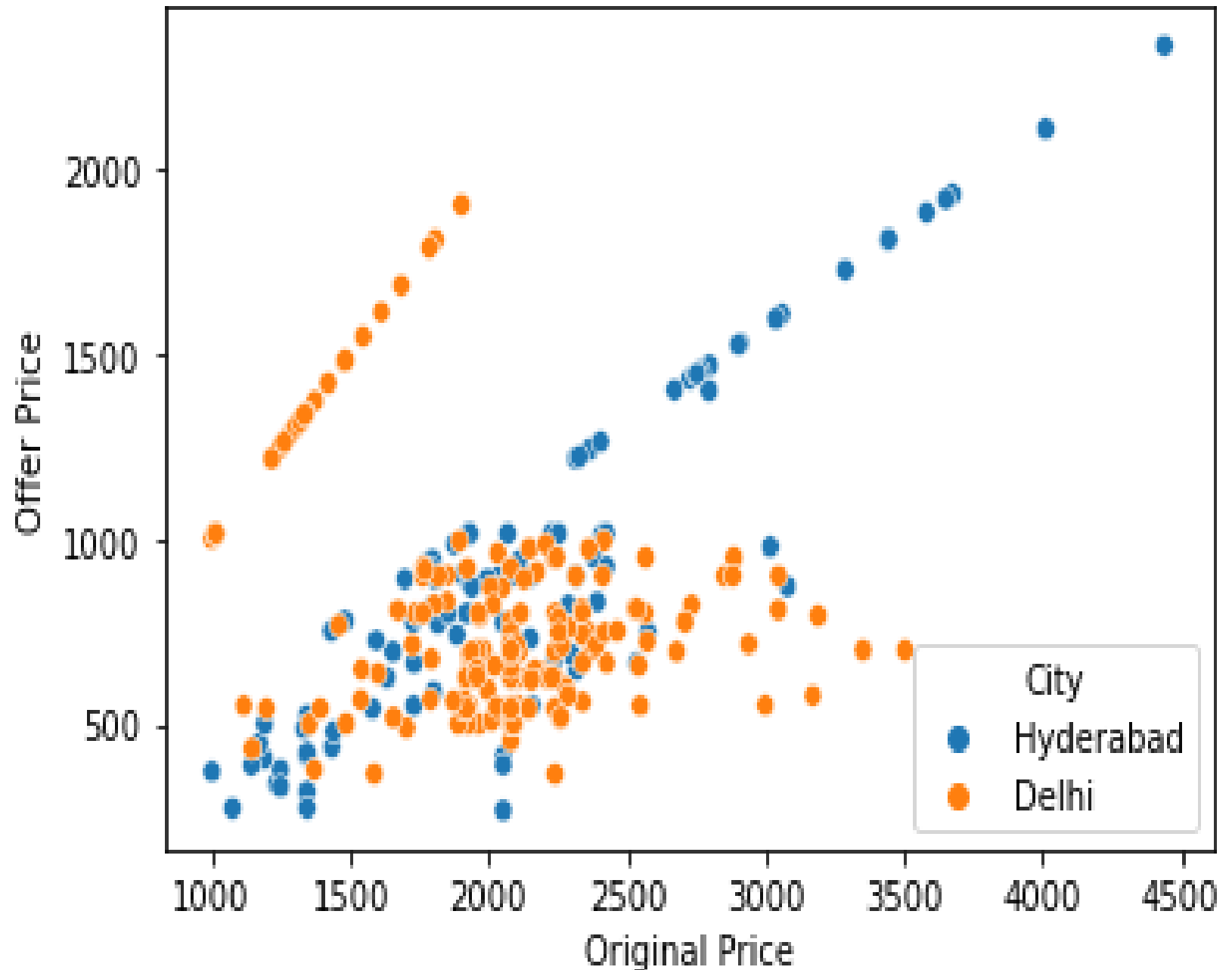
Correlation
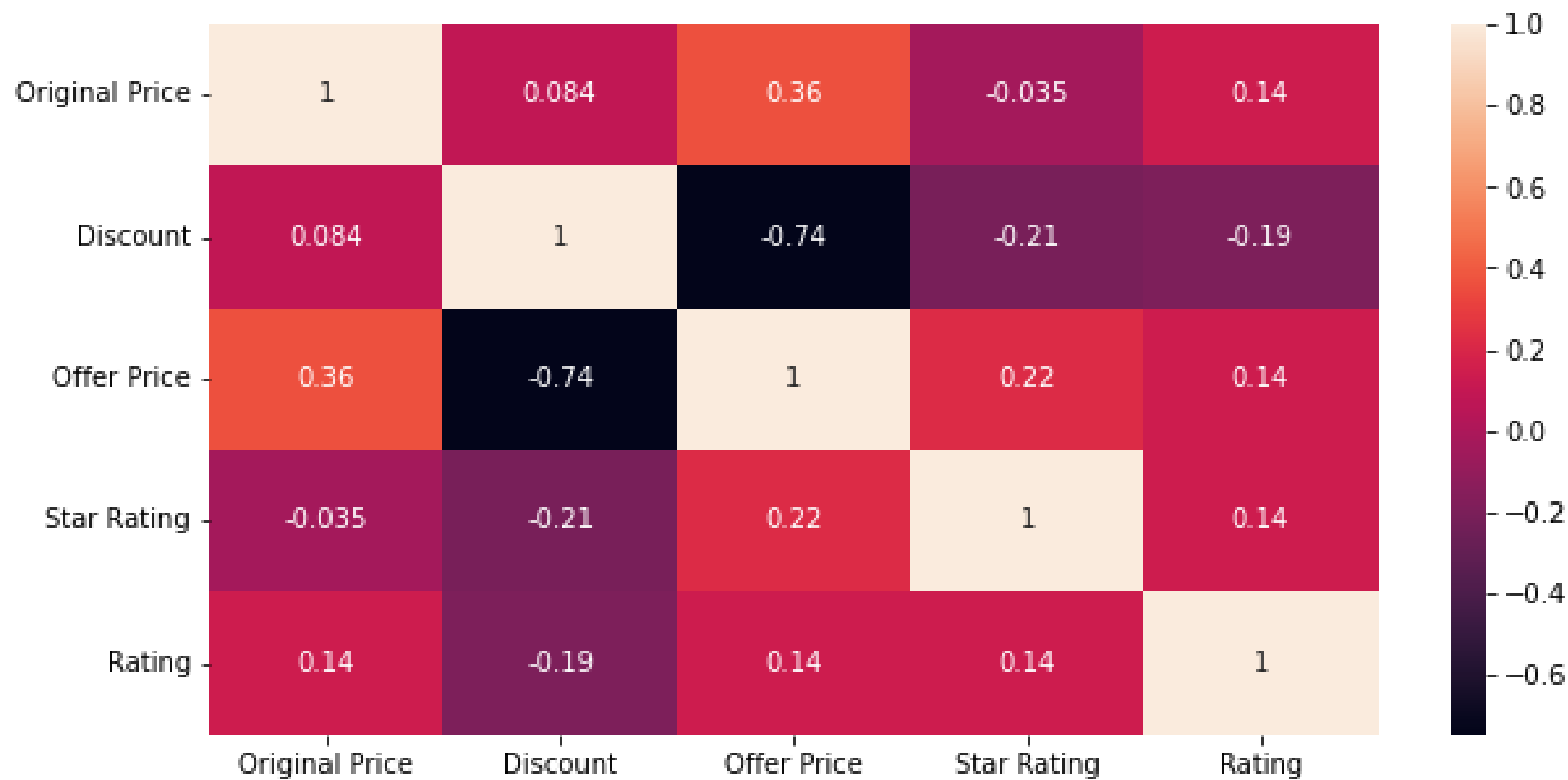
Conclusions

Offered Pricing

Category wise

Discount

Rating

Original Price

Features

Hyderabad, exceeds Delhi in the above data analysis.

If we talk about other individual Variable Comparison, then Category wise, Oyo Rooms out of all its segments are in the highest profitable line.

Number of Oyo Rooms is higher in Delhi as Hyderbad and it has got maximum Ratings

Challenges

Extraction of data was difficult, used Selenium

To fill null values in the given data set

Thank You!!!

- Mentor
- Friends