

**Final Assessment for Data Analytics**

HOUSE RENT PRICE   
ANALYSIS REPORT

**Prepared by**

|  |  |
| --- | --- |
| **Pralad Kadel**  Data Analyst Apprentice | Chakupat-10, Lalitpur, Nepal  [coderush.com.np](http://coderush.com.np/) |

**September 28, 2022**

# Introduction

To understand the general house rate prices in India, Code Rush was asked to conduct an evaluation of the rent prices of six major cities in the country. This report contains all information concerning said assessment. Various variables of the house are evaluated (like house size, furnishing status, no. of bathroom, BHK etc.) to see their consequences in the overall rent price of the house. Other factors like floor number, tenant preference, area type etc. are also evaluated to analyze their overall affect in the rent price. The assessment is also a mean to determine the general price range for customer with different needs and requirements.

# Key Aspects of the Assessment

## Data:

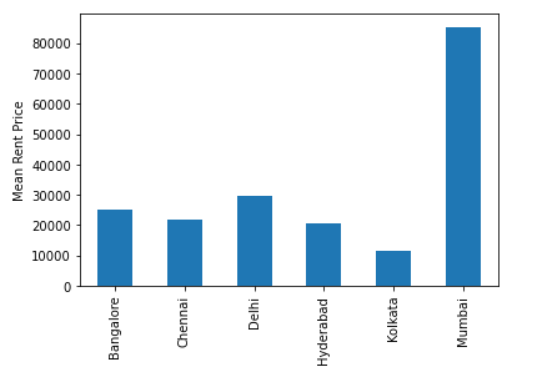
Source of the data is unknown as the sample data was provided by the CodeRush team directly as a part of the assessment.

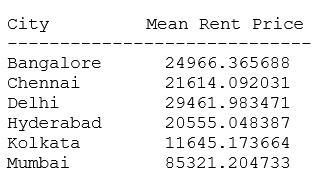
## Methods:

For the analysis of data, various types of descriptive statistics like measure of frequency, correlation, dispersion etc. are used so that we can present the data in a meaningful and understandable way allowing us a simplified interpretation and visualization of the data. We have used various univariate plots (like histogram, pie chart etc.) and multivariate plots (like scatter, linear model plots etc.) for the analysis of the data to understand their relationship with the house rent prices.

## Results and Analysis:

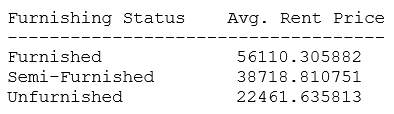
Our task is to analyze the given rent data and identify how the different variables affects the overall rent price of the houses. For that, we have filtered our dataset based on different variables like furnishing status, city, area, area type, no of BHK etc. and observed their relationship with the rent prices. We have also built a predictive model based on the data and some of the results of our analysis are shown and explained below

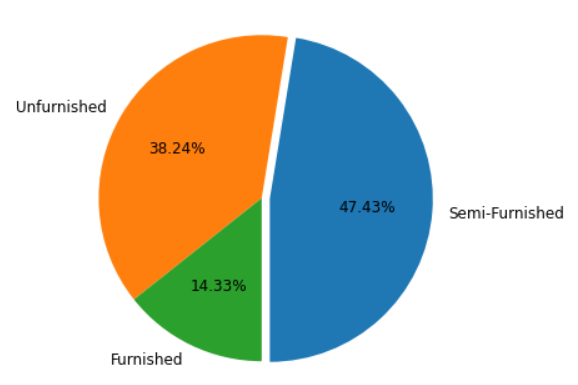


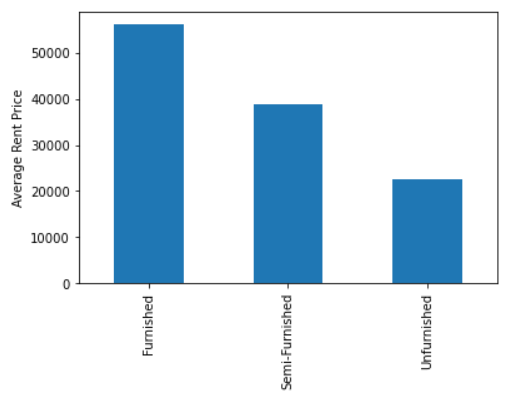


*Figure: Mean house rent price based on different city*

From the bar diagram, we observed that the rent price is highest in the Mumbai city with the average of INR 85,321. The cheapest city to live in is Kolkata with the average rent price of INR 11,645 and the average cost to rent a house in all other cities is between INR 20,000 to INR 30,000.

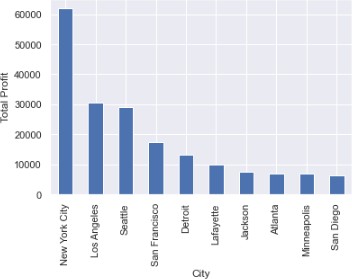






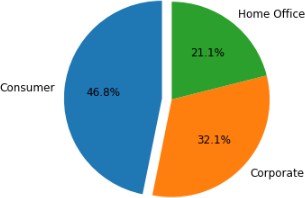
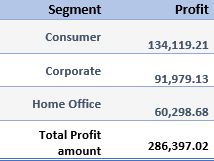
*Figure: Illustration of furnishing status and rent price relationship*

47.43% of the houses in rent are semi-furnished, 38.28% are unfurnished and remaining 14.33% of the houses are fully furnished. We also observed that the average price of rent for a furnished, semi-furnished and unfurnished houses are INR 56110, INR 38718 and INR 22461 respectively. We inferred that the furnished house cost more than double the unfurnished house.



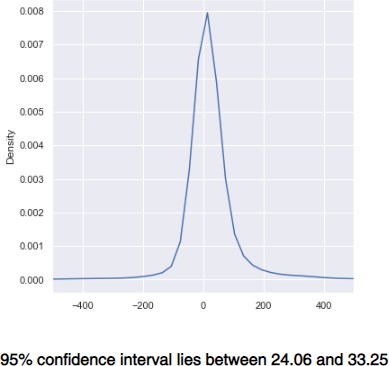
*Figure: Top ten most profitable cities*

In total, superstore sold around 531 different cities throughout the country. Among them, the city that made the most profit is New York City covering 28.66% of the total profit made by the store. The top ten most selling cities with their profit amount is shown in the figure above covering 68.38% of the total profit amount.



*Figure: Profit made by store based on different segments*

Based on the sample data, the superstore made most of its profit from the consumer segment with 46.8% of the total profit coming from the consumers. It made 32.1% and 21.1% of its profit from corporate and home office segment respectively.

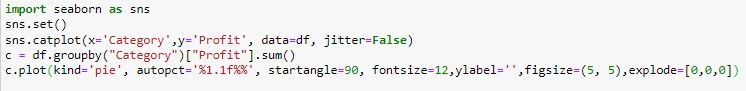


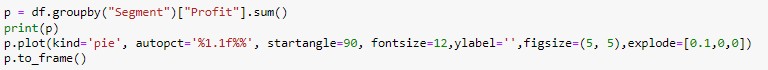
Using Z-score test, we evaluated that the range of profit will be between $24.06 and $33.25 with 95% confidence interval (i.e., 95 percent of our average profit data lies within this range). We also evaluated that there is 51.93% of probability that we make $40 or greater amount of profit and 54.87% of probability that the store makes no profit (or suffers loss).

# Conclusion and Recommendation

Considering that the data provided is accurate, the team can conclude that the superstore is considerably in profit. However, the store still has more than 50% of chance that it will make zero profit or goes into loss. Therefore, the store must focus on improving the identified weak selling categories and segment. Since furniture was the least profitable product, the store can focus on researching the most selling, modern furniture and add them on their store. The store is making the most amount of profit in New York, so the store can open new store fronts in this city. The store also needs to work on home office segment and improve in this field.

# Appendices





In [5): df['Country'] ..value\_counts()

Dut(5] : United State.s. 9994

Name: Country, cltype: int64

In [6): a df.grouphy("City")["Profit"J.s.um()

a pd. Series .. sort\_values (a, ascending=False)

a

Dut(6]: City

Ne1• York City Los P.ngeles. Seati:le

San Francis.co Detroit

62036.9837

30440.7579

29156.0967

17507.3854

1.3181.7908

Chicago -6654. 5688

Lancaster - 7239. 0634

San Antonio -7299.0502

Houston -10153. 5485

Philadelphia -1.3837. 7674

Name: Profit, Length: 531, dtype: float64

In [7): print( 'New York Profit Percenage = ', 62036.98/286397 .02\*100) a[:10).s.um()

print('Top 10 cities Profit Percenage = ', 190118.5284/286397.02\*100)

Ne," York Pr·ofit Percenage = 21. 661182089115314

Top 10 cities. Profit Percenage = 66.38285845292664 In [8): a [:10). plot (kind=' bar', ylabel = 'Total Profit')

**In** [11): profit= df['Profit']

profit. describe()

|  |  |
| --- | --- |
| **out [11]** : count | 9994.000000 |
| mean | 28. 656896, |
| s:td | 234. W0108 |
| m:i.n | -6599. 978,000 |
| 25% | 1. 728750 |
| 50% | 8 . 66650'0 |
| 75% | 29.364'00'0 |
| max | 8399.976000 |

Name:: Profit, dtype: float64

**In** [12] : irr o rt mat plot lib . pyplot as p1 t

**-from** scipy.stats import norm

prob.....:40 = norm...cdf( 40, 28. 656896, 234. 260108)

print('$40 and gr-eater Pr,obability is', pl'Cb\_40"'100,'\n') prob\_0 = nonm,. cdf(0, 28.656895 234. 260108)

print('No Pr,ofit Probability is', (1 - p,r,ob\_0)"'100) profit.p1ot(kind="kde",figsize=(6,6), xlim=(-500,500))

$40 and greater Pl'Cbability is 51..'93096303466416 No Pl'C.fit ProbahiJ.ity is 54.86809213427477