## Artificial Intelligence and

## Machine Learning

Project Report

Semester-IV (Batch-2022)

House Rent Prediction

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Ms. Shagun Sharma

Karan Kumar Garg, 2210990478 (G-10)

Karanpreet Singh, 2210990482 (G-10)

Karishma Dhawan, 2210990483 (G-10)

Lokesh Garg, 2210990544 (G-10)

**Department of Computer Science and Engineering**

**Chitkara University Institute of Engineering & Technology,**

**Chitkara University, Punjab**

**Table of Content**

|  |  |  |
| --- | --- | --- |
| **S.No.** | **Section** | **Page No.** |
| 1. | Introduction | 1-4 |
| 2. | Problem Definition and Requirements | 5-6 |
| 3. | Proposed Design/Methodology | 7-9 |
| 4. | Results | 10-17 |
| 5. | References | 18 |

1. **Introduction**

In recent years, the real estate market has experienced significant fluctuations, making it increasingly challenging for both tenants and landlords to determine fair and competitive rental prices. This unpredictability often leads to dissatisfaction and financial strain for both parties involved. To address this issue, we propose a project aimed at leveraging Artificial Intelligence (AI) and Machine Learning (ML) techniques to predict house rent prices accurately.

* 1. **Background**

The real estate rental market plays a crucial role in urban economies, providing housing options for a significant portion of the population. In many cities around the world, renting a home is often more feasible than purchasing one due to financial constraints, lifestyle preferences, and mobility requirements. However, the rental market is influenced by a myriad of factors, making it highly dynamic and sometimes unpredictable. Traditional methods of estimating rental prices often fall short due to their inability to process and analyze large volumes of data efficiently.

* 1. **Objectives**

The objective of the House Rent Prediction Project is to develop a robust machine learning model that accurately predicts rental prices for residential properties. This objective is broken down into several specific goals:**Develop Accurate Prediction Models:**Create and train machine learning models capable of predicting house rental prices with high accuracy. These models should be able to consider various factors such as property characteristics, location, and market trends.

1

**Identify Key Influencing Factors:**Analyze and identify the most significant factors that influence rental prices. This includes features like property size, number of rooms, proximity to amenities, neighbourhood quality, and historical price trends.

**Enhance Decision-Making for Stakeholders:**Provide landlords with a tool to set competitive and fair rental prices based on market data and trends.Help tenants make informed decisions by providing accurate rental price estimates for properties they are considering.

**Improve Market Transparency:**Increase transparency in the rental market by making data-driven price estimates accessible to all stakeholders. This will help reduce information asymmetry between landlords and tenants.

**Incorporate Real-Time Data and Trends:**Ensure that the predictive models can incorporate and adapt to real-time data and emerging market trends to maintain their accuracy and relevance over time.

**Develop a User-Friendly Application:**Create an intuitive, user-friendly application or web interface where users can input property details and receive instant rent price predictions. The interface should be easy to use for both landlords and tenants.

**Facilitate Continuous Improvement:**Implement a feedback loop where the model can be continuously improved based on new data and user feedback, ensuring that the predictions remain accurate and reliable over time.

2

**Promote Fairness and Equity:**Ensure that the predictive model is free from biases that could unfairly influence rental prices based on discriminatory factors. This involves careful consideration of the data used and the development of fair machine learning practices.

* 1. **Significance**

**Economic Significance**Market Efficiency: Aligns rental prices with market conditions, reducing supply-demand mismatches.Investment Decisions: Aids landlords and investors in making profitable decisions.Pricing Strategy: Helps set competitive and fair rental prices, optimizing occupancy rates and income.

**Social Significance**Tenant Empowerment: Provides transparent rental price information for informed decisions.Fair Housing Practices: Promotes fairness by using objective data to set prices.Housing Accessibility: Highlights affordable areas, aiding low and middle-income families.

**Technological Significance**AI and ML Advancement: Applies advanced technologies to real-world problems.Data Utilization: Demonstrates effective integration and analysis of large datasets.Model Development: Contributes to the creation of advanced predictive models.

**Strategic and Policy Significance**Urban Planning: Provides insights for data-driven urban development policies.Regulatory Oversight: Helps regulators monitor and ensure fair rental practices.

3

**Environmental and Community Significance**Sustainable Development: Supports initiatives by identifying demand patterns.Community Stability: Contributes to stability by promoting fair and accurate rental pricing.

4

1. **Problem Definition and Requirements**

The real estate rental market is characterized by its complexity and variability, influenced by numerous factors such as location, property features, amenities, and market conditions. Accurately predicting rental prices is a challenging task due to the dynamic nature of these factors. Traditional methods of rent price estimation often lack precision and fail to account for the multifaceted nature of the rental market.

* 1. **Software Requirements**

**Development Environment**Languages: PythonIDEs: Jupyter Notebook, PyCharm, VS Code

**Data Handling and Preprocessing**Storage: CSV (Comma separated File)Libraries: Pandas, NumPyData Cleaning: Feature-engine

**Machine Learning and Model Development**Frameworks: Scikit-learnVisualization: Matplotlib, Seaborn, PlotlyEvaluation: SciPy

* 1. **Hardware Requirements**

**Development Environment**Processor: Multi-core processor (e.g., Intel Core i4 or higher)RAM: At least 4GB for handling datasets and modelsStorage: SSD storage with sufficient space for datasets and development environments

5

**Model Training and Development**GPU: Optional but recommended for accelerating modelCUDA-enabled GPU: Required for GPU acceleration with TensorFlow or PyTorch

* 1. **Data Sets**

**Rental Listings Data:** Information about properties, amenities, and prices.**Property Characteristics:** Bedrooms, bathrooms, square footage, etc.

6

1. **Proposed Design / Methodology**

**Data Collection and Preprocessing:**Data Sources: Gather rental listings, historical rental data, geospatial information, demographic data, property characteristics, market indices, user-generated feedback, and optionally weather data.Data Cleaning: Handle missing values, outliers, and inconsistencies. Normalize and standardize numerical features, encode categorical variables, and address any data quality issues.**Feature Engineering:**Feature Selection: Identify relevant features that impact rental prices, such as property size, location, amenities, demographics, economic indicators, and user feedback.Feature Transformation: Create new features or transform existing ones to enhance model performance. For example, derive additional features from geospatial data (e.g., distance to amenities) or combine multiple features to capture interactions.**Model Selection and Training:**Algorithm Selection: Choose appropriate machine learning algorithms based on the dataset characteristics and problem complexity. Options may include linear regression, decision trees, random forests, gradient boosting machines, and neural networks.Model Training: Split the dataset into training and testing sets. Train the selected models using the training data and evaluate their performance using suitable metrics like Mean Absolute Error (MAE), Mean Squared Error (MSE), or Root Mean Squared Error (RMSE).

7

**Model Evaluation and Tuning:**Cross-Validation: Perform cross-validation to assess the generalization performance of the models and detect overfitting.Hyperparameter Tuning: Optimize model hyperparameters using techniques like grid search or random search to improve performance further.

* 1. **File Structure**

C:\Users\ \_\_\

│

├── House\_Rent\_Dataset.csv

│

├── AIML.ipynb

│

* 1. **Algorithm Used**

**Linear Regression:**Simple and interpretable algorithm suitable for predicting continuous variables like rental prices. It assumes a linear relationship between the input features and the target variable.

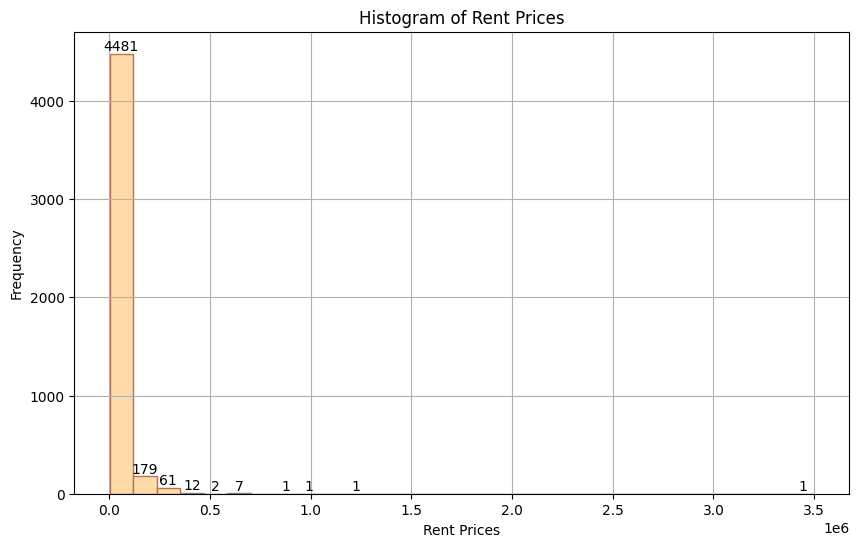
**Random Forests:**Ensemble learning technique that builds multiple decision trees and aggregates their predictions to improve accuracy and reduce overfitting. Random forests are robust against noise and outliers in the data.

8

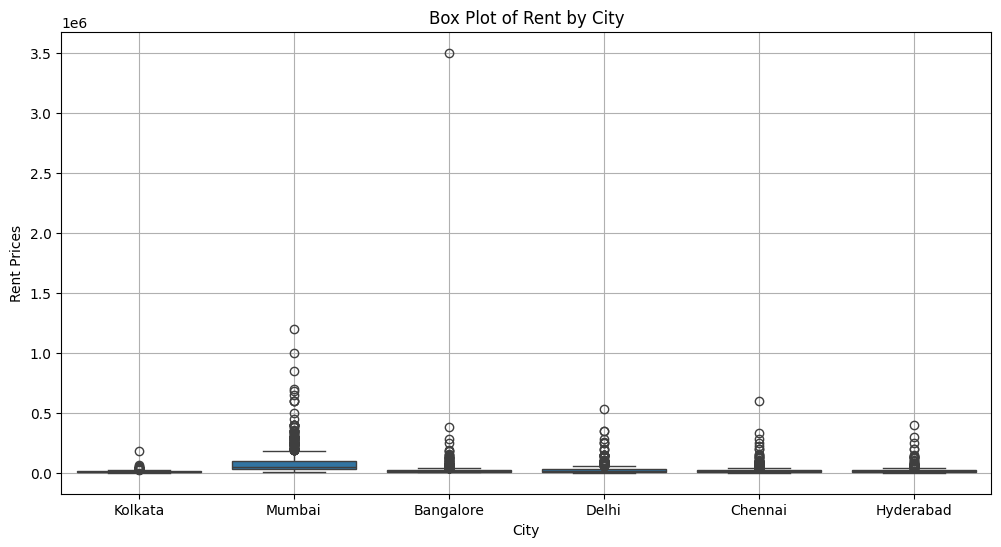
**Support Vector Machines (SVM):**Supervised learning algorithm that finds the hyperplane that best separates the classes in high-dimensional space. SVM can be applied to regression tasks by using the epsilon-insensitive loss function.

**4. Results**

9

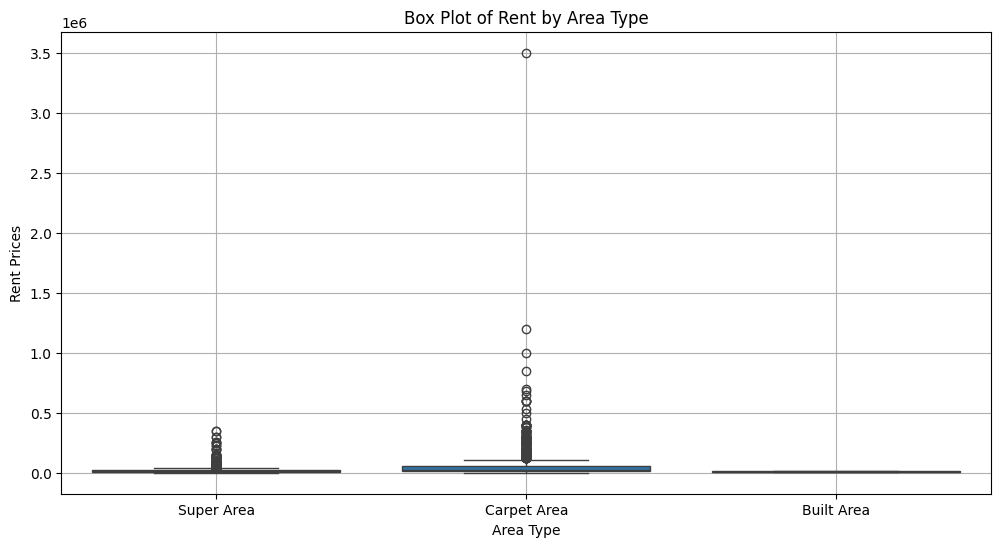
* 

It gives the Frequency of houses according to its Rent Price.

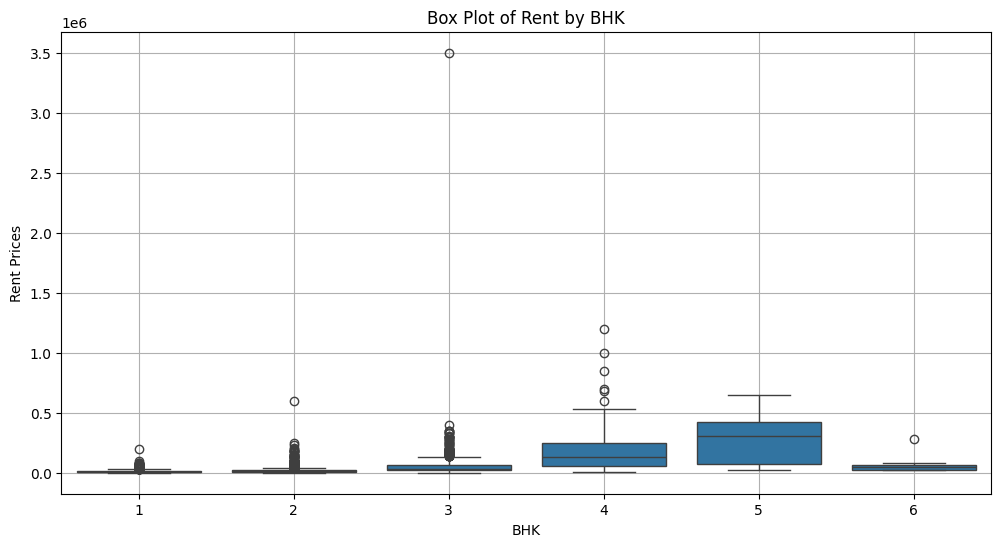
* 

It gives the Rent Prices of houses according to its city.

10

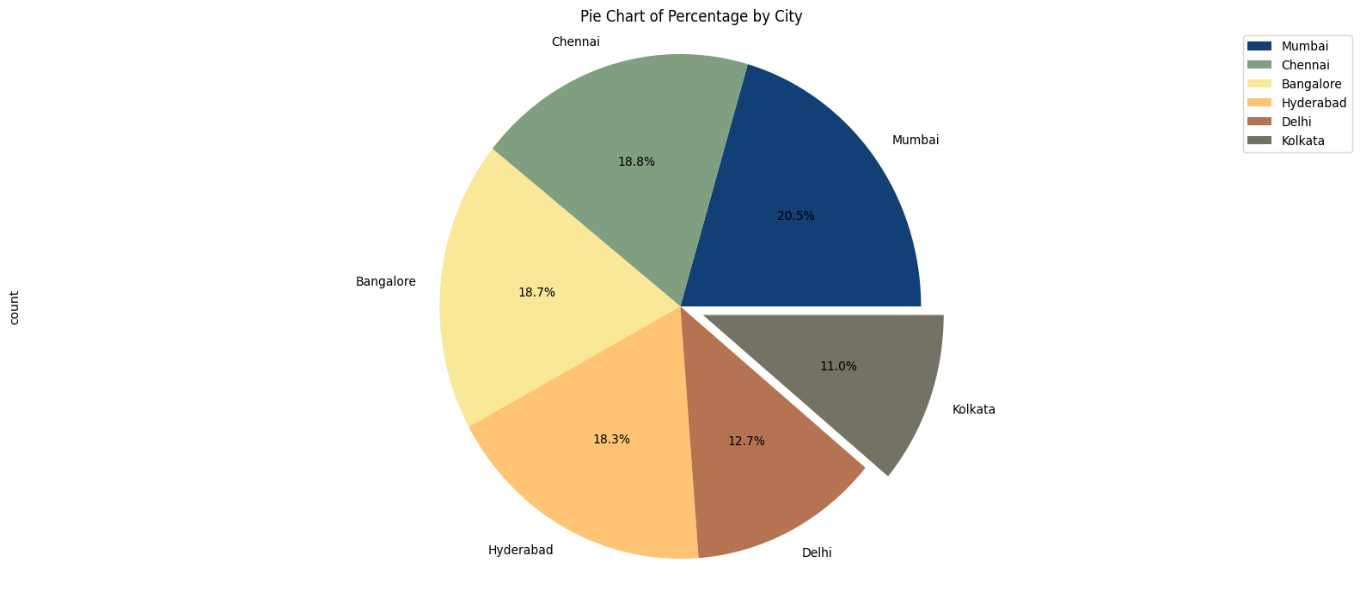
* ****

It gives the Rent Prices of houses according to its Area Type.

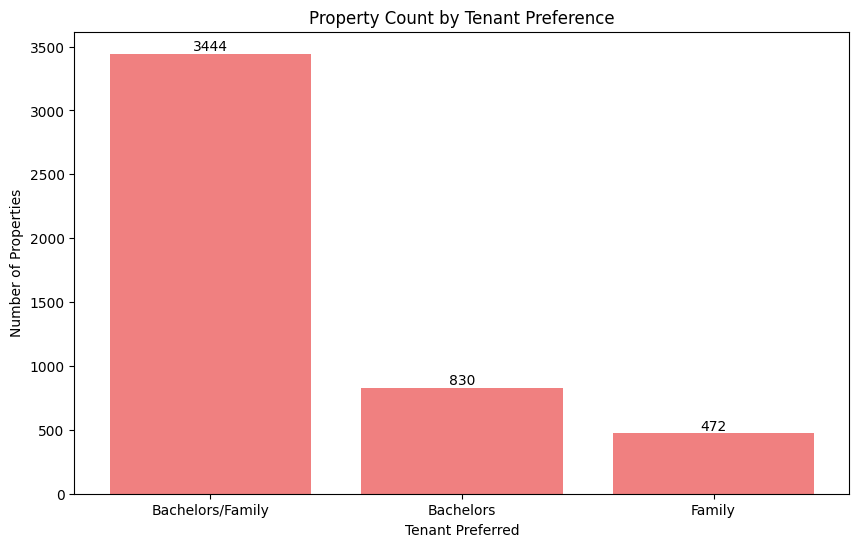
* ****

It gives the Rent Prices of houses according to its BHK.

11

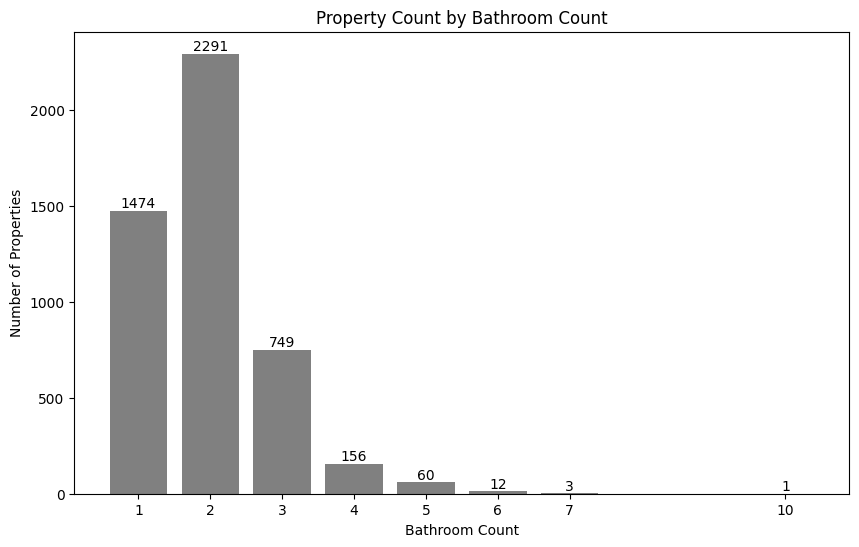
* ****

It gives the available no of rent houses in a particular city.

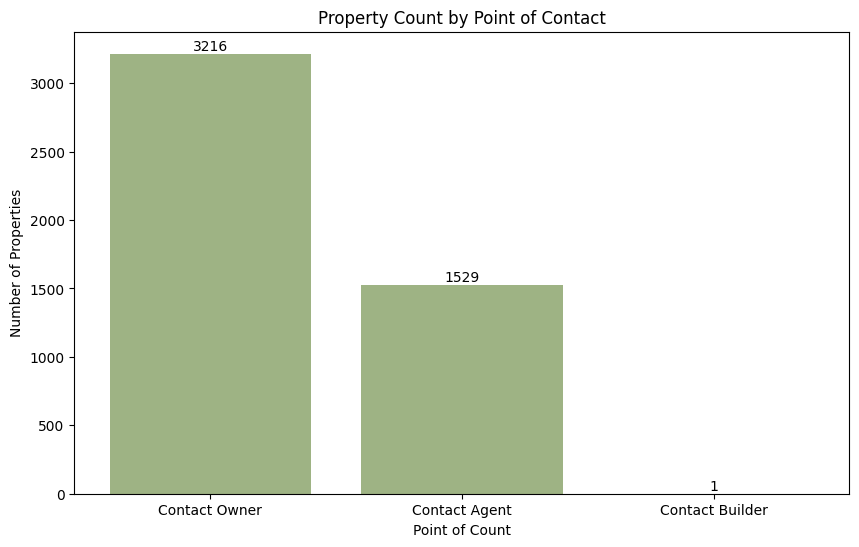
* ****

It gives the Number of houses according to Talent Preferred.

12

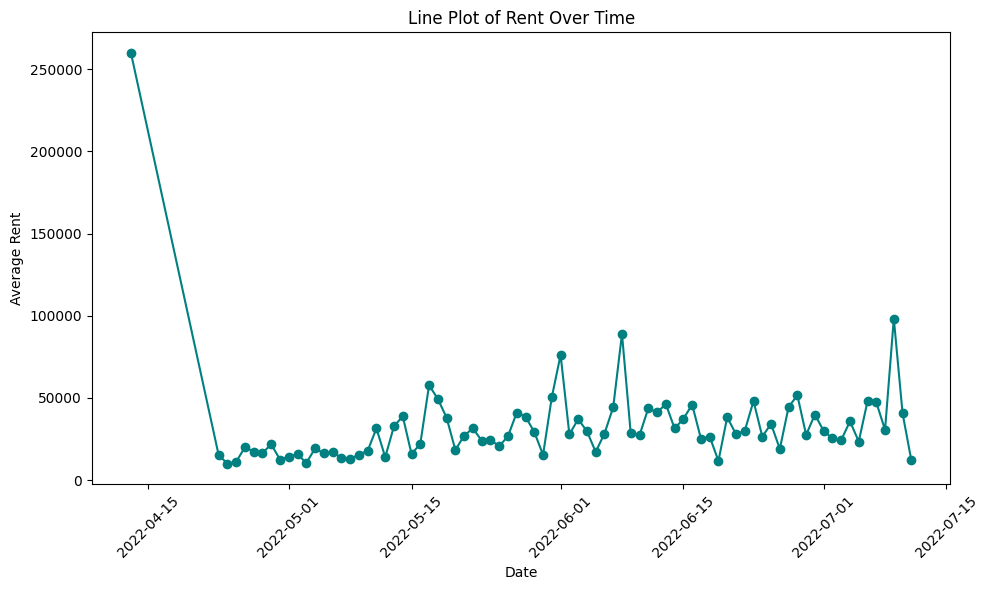
* ****

It gives the Number of houses according to Bathroom count.

****

It gives the Number of houses according to Contact.

13

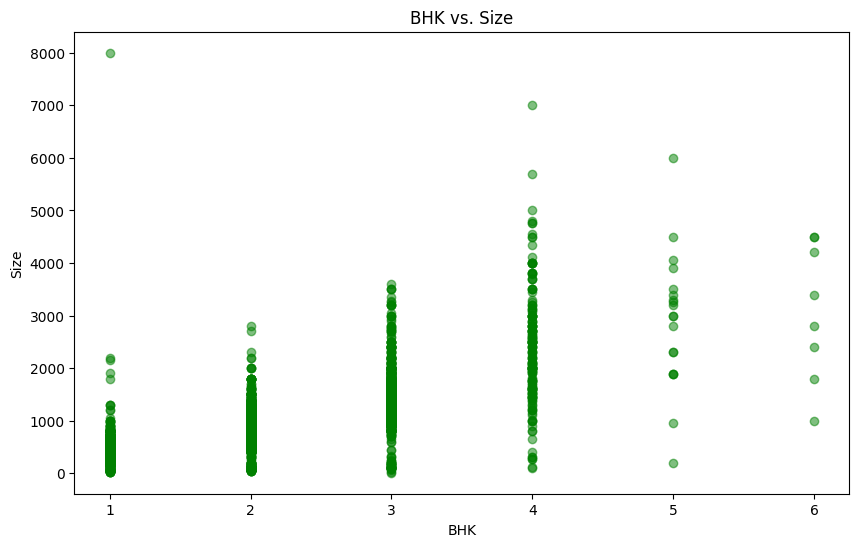
* ****

It gives the Average rent according to the Date.

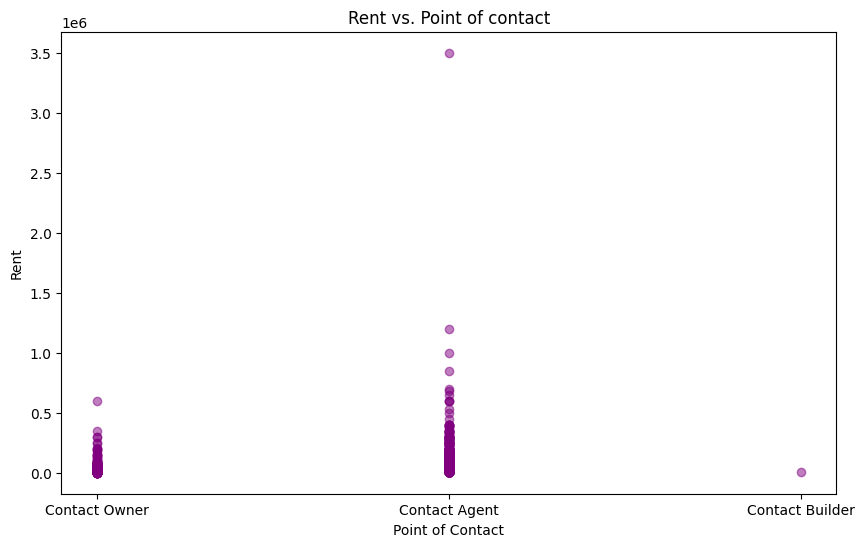


It gives the Rent of houses according to its size.

14

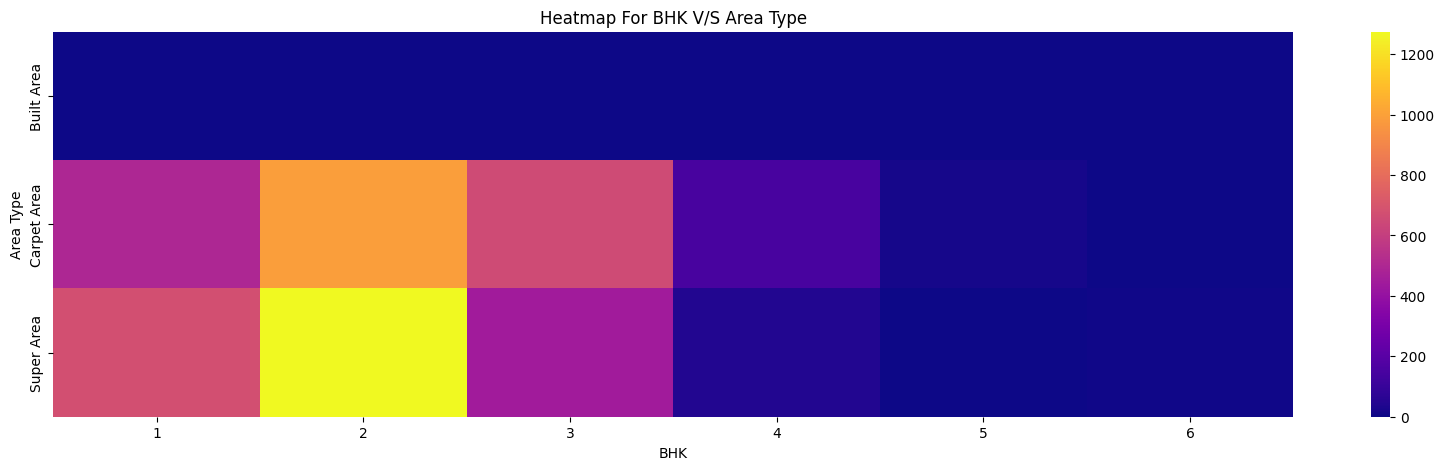
* ****

It gives the size of houses according to its BHK.

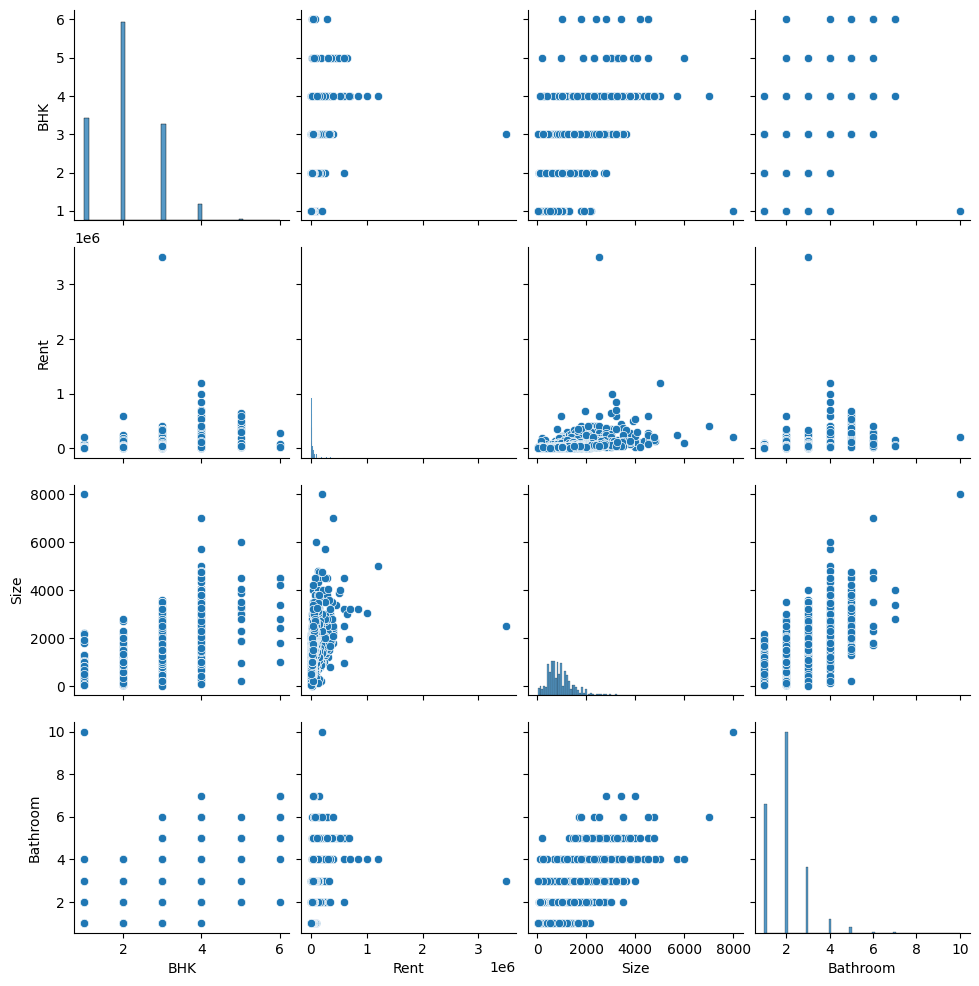
****

It gives the Rent of houses according to its Point of contact.

15

* ****

It gives the Area type of houses according to its BHK.

* ****

16

**Models Trained**

1. **Linear Regression:**

Library Used: sklearn

Time taken to train: 0.4 seconds

Model Accuracy: 0.7330051055316977 == 73.3%

1. **Random Forests:**

Library Used: sklearn

Time taken to train: 2.7 seconds

Model Accuracy: 0.8303840902289875 == 83%

1. **Support Vector Machines (SVM):**

Library Used: sklearn

Time taken to train: 0.6 seconds

Model Accuracy: 0.7922955360034856 == 79.2%

17

1. **References**

https://www.kaggle.com/

<https://pandas.pydata.org/docs/>

<https://scikit-learn.org/stable/>

<https://matplotlib.org/stable/index.html>

<https://seaborn.pydata.org/>

<https://github.com/>

18