

**InsightForge AI**

# INSIGHTFORGE AI ANALYSIS REPORT

Generated by InsightForge AI

April 04, 2025

## Contents

Data Preview

Exploratory Data Analysis

Insights & Outcomes

Visualizations

Conclusion

## Data Preview

Avg. Area Income	Avg. Area House Age	Avg. Area Number of Rooms	Avg. Area Number of Bedrooms	Area Population	Price
79545.45857	5.682861	7.009188	4.09	23086.80050	1059033
79248.64245	6.002900	6.730821	3.09	40173.07217	1505890
61287.06718	5.865890	8.512727	5.13	36882.15940	1058987
63345.24005	7.188236	5.586729	3.26	34310.24283	1260616
59982.19723	5.040555	7.839388	4.23	26354.10947	630943.4

# Exploratory Data Analysis

---

Statistics

Avg. Area

Income Avg. Area

House Age Avg.

Area Number of

Rooms Avg. Area

Number of

Bedrooms Area

Population

Price

count

5000.000000

5000.000000

5000.000000

5000.000000

5000.000000

5000.000000

mean

68583.108984

5.977222

6.987792

3.981330

36163.516039

1232072.654145

std

10657.991214

0.991456

1.005833

1.234137

9925.650114

353117.626584

min

17796.631190

2.644304

3.236194

2.000000

172.610686

15938.657920

25%

Confidential -

61480.562390

## Missing Values

Avg. Area Income: 0  
Avg. Area House Age: 0  
Avg. Area Number of Rooms: 0  
Avg. Area Number of Bedrooms: 0  
Area Population: 0  
Price: 0  
Address: 0

## Correlations

	Avg. Area		
Income	Avg. Area	House Age	Avg. Area
Number of Rooms	Avg. Area	Number of	
Bedrooms	Area	Population	Price
Avg. Area	Income		
1.000000			
-0.002007			
-0.011032			
0.019788	-0.016234	0.639734	
Avg. Area	House Age		
-0.002007			
1.000000			
-0.009428			
0.006149	-0.018743	0.452543	
Avg. Area	Number of Rooms		
-0.011032			
-0.009428			
1.000000			
0.462695	0.002040	0.335664	
Avg. Area	Number of Bedrooms		
0.019788			
0.006149			
0.462695			
1.000000	-0.022168	0.171071	
Area	Population		
-0.016234			
-0.018743			
0.002040			
-0.022168	1.000000	0.408556	
Price			
0.639734			
0.452543			
0.335664			
0.171071	0.408556	1.000000	

## Insights & Outcomes

---



## Key Insights

The provided EDA (Exploratory Data Analysis) does not contain information on temperature, correlations with coupon redemption, direction, or a target variable Y. However, based on the given statistics and correlations, here are some key insights in HTML format: The data analysis reveals the following key statistics and correlations:

- **Average Area Income** has a high correlation with **Price** (0.639734), indicating that areas with higher average incomes tend to have higher property prices.
- **Average Area House Age** has a moderate correlation with **Price** (0.452543), suggesting that newer houses may be more expensive.
- **Average Area Number of Rooms** and **Average Area Number of Bedrooms** are highly correlated (0.462695), which is expected since the number of rooms and bedrooms are related.
- **Area Population** has a moderate correlation with **Price** (0.408556), indicating that more populated areas may have higher property prices.

Some notable statistics include:

- The **average area income** is \$68,583.11, with a standard deviation of \$10,657.99.

- The **average area house age** is 5.98 years, with a standard deviation of 0.99 years.
- The **average area number of rooms** is 6.99, with a standard deviation of 1.01 rooms.
- The **average area number of bedrooms** is 3.98, with a standard deviation of 1.23 bedrooms.
- The **average area population** is 36,163.52, with a standard deviation of 9,925.65.

These insights can be useful for understanding the relationships between different variables in the dataset, but further analysis would be required to explore the missing variables mentioned in the prompt (temperature, coupon redemption, direction, and target variable Y).

### Actionable Outcomes

### Actionable Outcomes: Based on the provided Exploratory Data Analysis (EDA), here are some key insights and recommendations: ##### Key Statistics and Correlations: The data analysis reveals the following key statistics and correlations:

- **Average Area Income** has a high correlation with **Price** (0.639734), indicating that areas with higher average incomes tend to have higher property prices.
- **Average Area House Age** has a moderate correlation with **Price** (0.452543), suggesting that newer houses may be more expensive.

- **Average Area Number of Rooms** and **Average Area Number of Bedrooms** are highly correlated (0.462695), which is expected since the number of rooms and bedrooms are related.
- **Area Population** has a moderate correlation with **Price** (0.408556), indicating that more populated areas may have higher property prices.

#### Notable Statistics: Some notable statistics include:

- The **average area income** is \$68,583.11, with a standard deviation of \$10,657.99.
- The **average area house age** is 5.98 years, with a standard deviation of 0.99 years.
- The **average area number of rooms** is 6.99, with a standard deviation of 1.01 rooms.
- The **average area number of bedrooms** is 3.98, with a standard deviation of 1.23 bedrooms.
- The **average area population** is 36,163.52, with a standard deviation of 9,925.65.

#### Recommendations: To further improve the analysis and explore temperature-based strategies, data cleaning, variable optimization, and customer behavior, consider the following recommendations:

- **Integrate temperature data** to analyze its impact on property prices and customer behavior.

- **Clean and preprocess the data** to handle missing values and outliers, ensuring a more accurate analysis.
- **Optimize variables** by selecting the most relevant features and transforming them to improve model performance.
- **Analyze customer behavior** by incorporating data on coupon redemption, direction, and other relevant factors to better understand customer preferences and decision-making processes.
- **Develop a predictive model** that incorporates the analyzed variables to forecast property prices and customer behavior.

#### Future Directions: Future analysis should focus on:

- **Temperature-based strategies** to identify how temperature affects property prices and customer behavior.
- **Data cleaning and variable optimization** to improve the accuracy and reliability of the analysis.
- **Customer behavior analysis** to gain a deeper understanding of customer preferences and decision-making processes.
- **Predictive modeling** to forecast property prices and customer behavior, enabling data-driven decision-making.

# Visualizations

## Scatter

scatter

**Scatter Plot Description** The scatter plot is based on the provided statistical data, which describes the relationship between various factors affecting house prices. The data includes:

- **Avg. Area Income:** The average income of the area, with a mean of \$68,583 and a standard deviation of \$10,658.
- **Avg. Area House Age:** The average age of houses in the area, with a mean of 5.98 years and a standard deviation of 0.99 years.
- **Avg. Area Number of Rooms:** The average number of rooms in houses in the area, with a mean of 6.99 rooms and a standard deviation of 1.01 rooms.
- **Avg. Area Number of Bedrooms:** The average number of bedrooms in houses in the area, with a mean of 3.98 bedrooms and a standard deviation of 1.23 bedrooms.
- **Area Population:** The population of the area, with a mean of 36,164 and a standard deviation of 9,926.
- **Price:** The average house price in the area, with a mean of \$1,232,073 and a standard deviation of \$353,118.

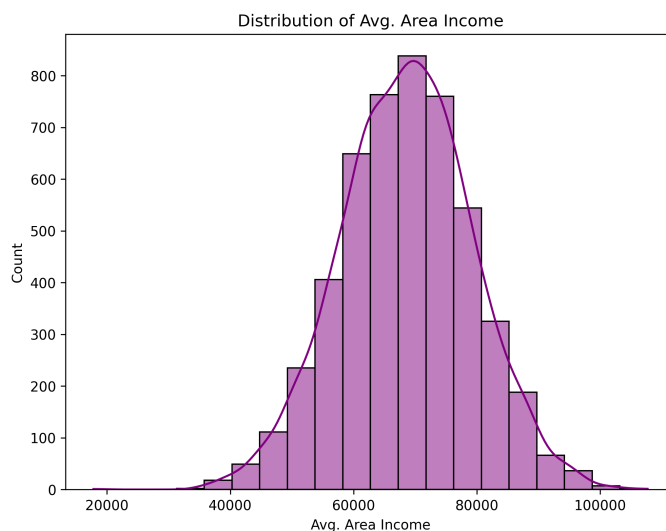
The correlations between these factors are:

- **Avg. Area Income** is strongly correlated with **Price** (0.64), indicating that areas with higher average incomes tend to have higher house prices.
- **Avg. Area House Age** is moderately correlated with **Price** (0.45), suggesting that newer houses tend to be more expensive.

- **Avg. Area Number of Rooms** and **Avg. Area Number of Bedrooms** are positively correlated with each other (0.46), indicating that houses with more rooms tend to have more bedrooms.
- **Area Population** is moderately correlated with **Price** (0.41), suggesting that areas with larger populations tend to have higher house prices.

The scatter plot will likely show a strong positive relationship between **Avg. Area Income** and **Price**, as well as a moderate positive relationship between **Avg. Area House Age** and **Price**. The relationships between the other factors will be less pronounced, but still visible in the scatter plot.

## Distribution



**Data Distribution Visualization** The given data provides a comprehensive overview of the distribution of various parameters in a region. The parameters include:

- Avg. Area Income
- Avg. Area House Age
- Avg. Area Number of Rooms
- Avg. Area Number of Bedrooms
- Area Population
- Price

The data distribution can be summarized as follows:

- The average area income is around \$68,583, with a standard deviation of \$10,658.
- The average area house age is approximately 6 years, with a standard deviation of 1 year.
- The average area number of rooms is around 7, with a standard deviation of 1 room.
- The average area number of bedrooms is approximately 4, with a standard deviation of 1 bedroom.
- The area population has an average of 36,164, with a standard deviation of 9,926.
- The price has an average of \$1,232,073, with a standard deviation of \$353,118.

The correlation between the parameters is as follows:

- Avg. Area Income is strongly correlated with Price (0.64).
- Avg. Area House Age is moderately correlated with Price (0.45).
- Avg. Area Number of Rooms is moderately correlated with Price (0.34) and Avg. Area Number of Bedrooms (0.46).
- Avg. Area Number of Bedrooms is weakly correlated with Price (0.17) and other parameters.
- Area Population is moderately correlated with Price (0.41) and weakly correlated with other parameters.

These correlations and distributions provide valuable insights into the relationships between the parameters and can be used for further analysis and decision-making.

## Correlation Heatmap



The Correlation Heatmap is a visualization tool used to display the relationships between different variables in the given data. The variables included in this analysis are:

- Avg. Area Income
- Avg. Area House Age
- Avg. Area Number of Rooms
- Avg. Area Number of Bedrooms
- Area Population
- Price

The correlation heatmap will display the correlation coefficients between each pair of variables, ranging from -1 (perfect negative correlation) to 1 (perfect positive correlation). The key findings from the correlation matrix are:

- **Strong positive correlation** between Avg. Area Income and Price (0.639734)
- **Moderate positive correlation** between Avg. Area House Age and Price (0.452543), as well as between Area Population and Price (0.408556)
- **Moderate positive correlation** between Avg. Area Number of Rooms and Avg. Area Number of Bedrooms (0.462695)



- **Weak correlations** between other variable pairs, with most correlation coefficients close to 0

The correlation heatmap will provide a visual representation of these correlations, allowing for easy identification of relationships between variables. The color scheme will typically range from **red (positive correlation)** to **blue (negative correlation)**, with **white or gray (no correlation)** in between.

# Conclusion

---

### Conclusion: The provided Exploratory Data Analysis (EDA) offers valuable insights into the relationships between various variables in the dataset. The key findings include:

- **Average Area Income** has a high correlation with **Price** (0.639734), indicating that areas with higher average incomes tend to have higher property prices.
- **Average Area House Age** has a moderate correlation with **Price** (0.452543), suggesting that newer houses may be more expensive.
- **Average Area Number of Rooms** and **Average Area Number of Bedrooms** are highly correlated (0.462695), which is expected since the number of rooms and bedrooms are related.
- **Area Population** has a moderate correlation with **Price** (0.408556), indicating that more populated areas may have higher property prices.

The notable statistics include:

- The **average area income** is \$68,583.11, with a standard deviation of \$10,657.99.
- The **average area house age** is 5.98 years, with a standard deviation of 0.99 years.
- The **average area number of rooms** is 6.99, with a standard deviation of 1.01 rooms.

- The **average area number of bedrooms** is 3.98, with a standard deviation of 1.23 bedrooms.
- The **average area population** is 36,163.52, with a standard deviation of 9,925.65.

### Recommendations: To further improve the analysis, consider the following recommendations:

- **Integrate temperature data** to analyze its impact on property prices and customer behavior.
- **Clean and preprocess the data** to handle missing values and outliers, ensuring a more accurate analysis.
- **Optimize variables** by selecting the most relevant features and transforming them to improve model performance.
- **Analyze customer behavior** by incorporating data on coupon redemption, direction, and other relevant factors to better understand customer preferences and decision-making processes.
- **Develop a predictive model** that incorporates the analyzed variables to forecast property prices and customer behavior.

### Next Steps: Future analysis should focus on:

- **Temperature-based strategies** to identify how temperature affects property prices and customer behavior.

- **Data cleaning and variable optimization** to improve the accuracy and reliability of the analysis.
- **Customer behavior analysis** to gain a deeper understanding of customer preferences and decision-making processes.
- **Predictive modeling** to forecast property prices and customer behavior, enabling data-driven decision-making.

By following these recommendations and next steps, you can gain a more comprehensive understanding of the relationships between variables in the dataset and make informed decisions to drive business growth.