SELMA BA-SULOUH

INFLUENCING FACTORS OF URBAN HOUSING SIZE IN SAUDI ARABIA

REGRESSION ANALYSIS

AGENDA

INTRO

- Housing design plays a crucial role in shaping the quality of life for residents (Norberg-Schulz, 1984).

- Housing size could affect the social relations of complex family structures in Saudi Arabia.

- Providing developers with appropriate housing size will increase stability of housing development, thus reducing home modifications and relocation.

BACKGROUND

LITERATE REVIEW

• Size matters. How does the number of dwellings affect housing co-production? (Gérald, 2022)

• Housing Price, Family Structures and Household Consumption: Empirical Evidence of Taiwan. (Hung et al., 2023)

• Influence Factors and Regression Model of Urban Housing Prices Based on Internet Open Access Data. (Wu et al., 2018)
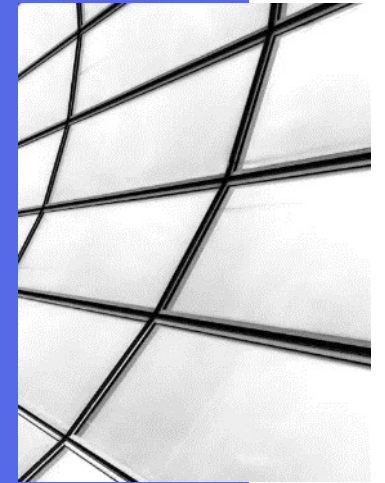
BACKGROUND

RESEARCH QUESTION

How does the urban housing size differs depending on number of rooms and household size?
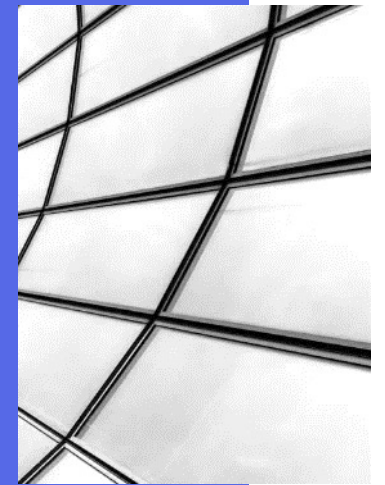
BACKGROUND

LIMITATIONS

• Data source reliability: The data used in this analysis is pulled from private property listing website affecting its accuracy.

• Missing data: the dependency ratio (children, elderly).

• Identifying outliers: Limiting the study size to eliminate them and making sure the irrelevancy of outliers to the model.

• Data coverage: the data is limited to the main cities in Saudi Arabia Riyadh, Ad Dammam, Jeddah, and Al Khobar, which is not generalizable to all Saudi Arabia.
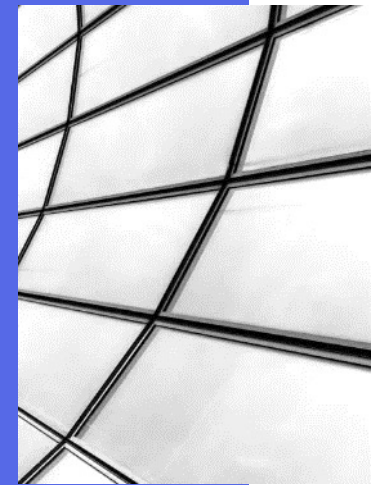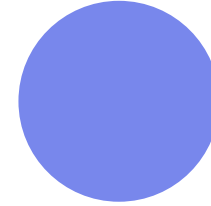
PRIMARY GOAL

To help private housing developers predict the accurate housing size for sustainable development.

H1: Housing size does not directly affect the number of rooms.

H2: Household size has the main effect on number of rooms.

METHODOLOGY
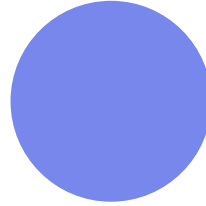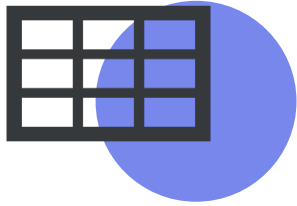
DATA

- Characteristics of Property Listing (Aqar.com)
  - 3,713 observations.
  - Variabls:
    - city
    - unit number
    - **size**
    - property_age
    - **bedrooms**

    - **bathrooms**
    - **livingrooms**
    - kitchen
    - garage
    - driver_room

    - maid_room
    - furnished
    - ac
    - roof
    - pool

    - frontyard
    - basement
    - duplex
    - stairs
    - elevator

    - fireplace
    - **price**
    - **household size**

METHODOLOGY
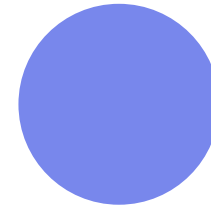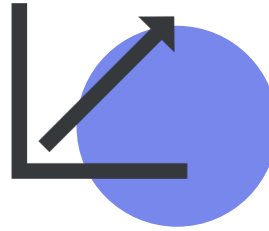
MODEL

- The analysis used housing size ('size') as the dependent variable y and the influence factors ('bedrooms', 'bathrooms', 'livingrooms', 'houshold size') as independent variables $x1$, $x2$, …, xn, the **linear regression model** is shown in equation, where $\beta0$ is the regression constant, $\beta1$ to βn are respectively the regression coefficient of each factor, and ε the random error.

$$y = \beta0 + \beta1\ x1 + \beta2\ x2 + … + \beta n\ xn + \varepsilon$$

METHODOLOGY

RESULTS
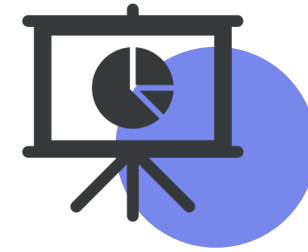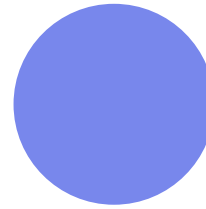
	size	bedrooms	bathrooms	kitchen	livingrooms	houshold size
mean	362.930245	5.085376	4.607326	0.909238	2.244008	3.973337
max	2944.000000	7.000000	5.000000	1.000000	5.000000	7.000000
min	100.000000	1.000000	1.000000	0.000000	1.000000	1.000000
median	330.000000	5.000000	5.000000	1.000000	2.000000	4.000000
std	176.249091	1.229787	0.702804	0.287309	0.915464	0.923875

- Calculation of mean, maximum, minimum, median, and standard deviation.

- To crate a statistical understanding of the data and simplifying it.

- The average housing size is over 362m2. the average number of bedrooms is 5 rooms. for living rooms is 2 rooms and for a household size is 4 people

METHODOLOGY
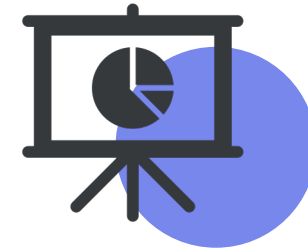
RESULTS

	size	bedrooms	bathrooms	livingrooms	houshold size
size	1.000000	0.241273	0.164381	0.166831	0.023070
bedrooms	0.241273	1.000000	0.428416	0.118842	0.015519
bathrooms	0.164381	0.428416	1.000000	0.406890	0.018723
livingrooms	0.166831	0.118842	0.406890	1.000000	0.019798
houshold size	0.023070	0.015519	0.018723	0.019798	1.000000

- Calculation of correlation coefficients of all related variables.

- To measure the change in one variable caused by the change in other variables

METHODOLOGY

RESULTS

Correlation Matrix
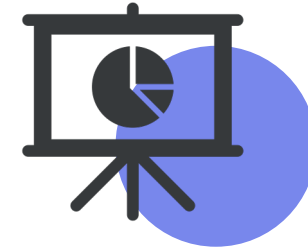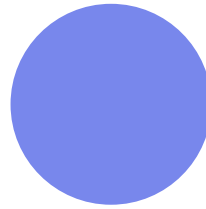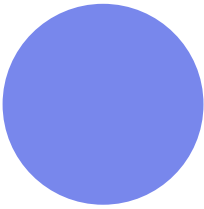
Combined Scatter Plot: Size vs Bedrooms, Bathrooms, Kitchen, Livingrooms
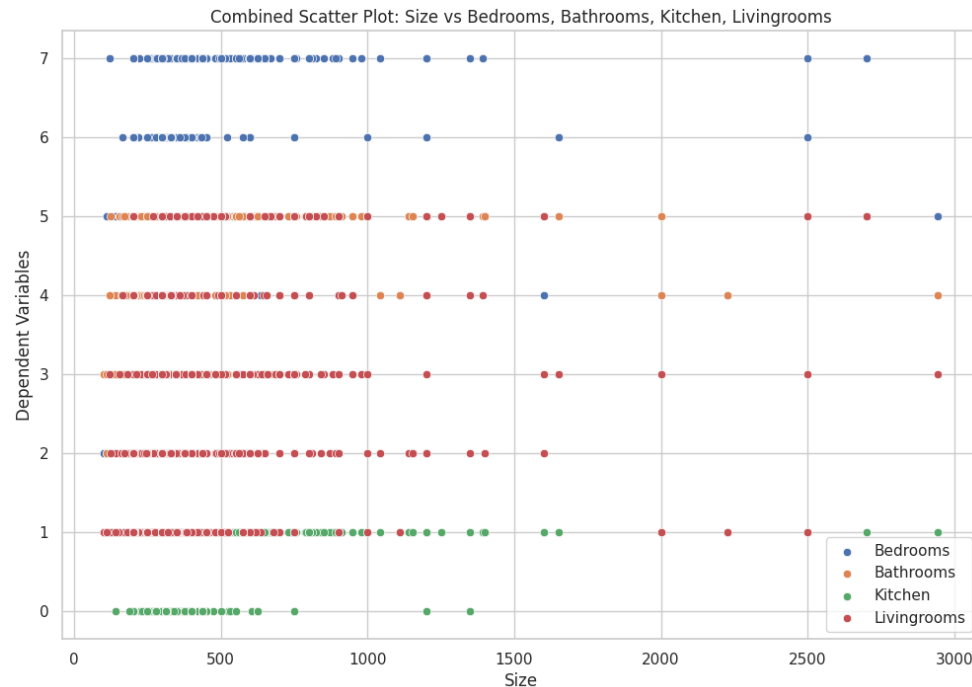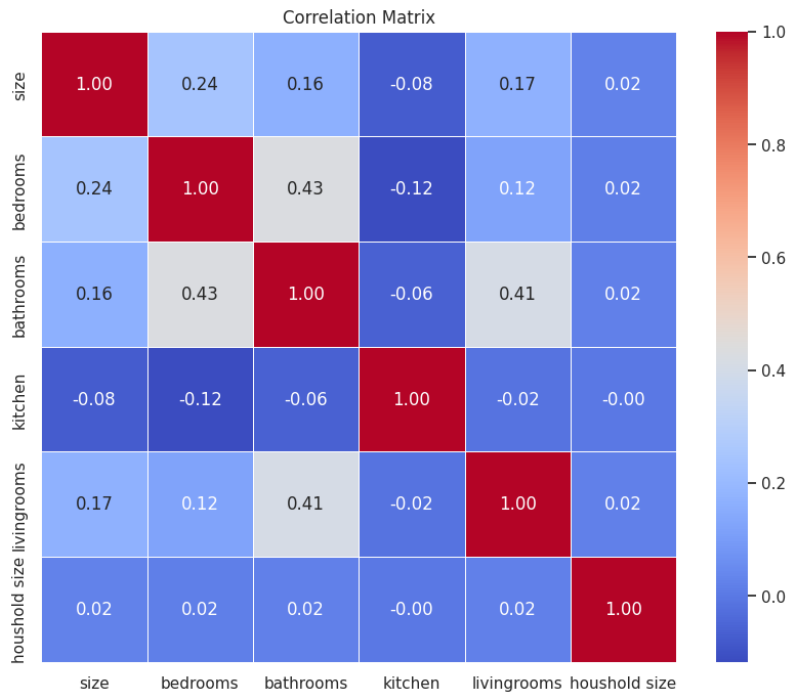
- No relation found possibly due to high number of variables studied

METHODOLOGY
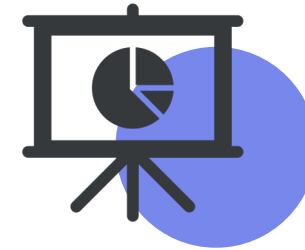
RESULTS

```
                        OLS Regression Results
==============================================================================
Dep. Variable:                   size   R-squared:                       0.078
Model:                            OLS   Adj. R-squared:                  0.077
Method:                 Least Squares   F-statistic:                     78.46
Date:                Mon, 04 Dec 2023   Prob (F-statistic):           5.56e-64
Time:                        18:29:06   Log-Likelihood:                -24320.
No. Observations:                3713   AIC:                         4.865e+04
Df Residuals:                    3708   BIC:                         4.868e+04
Df Model:                           4
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          114.7681     22.152      5.181      0.000      71.337     158.199
bedrooms        31.2732      2.507     12.475      0.000      26.358      36.188
bathrooms        4.0234      4.768      0.844      0.399      -5.324      13.371
livingrooms     25.8059      3.331      7.747      0.000      19.275      32.337
houshold size    3.1914      3.009      1.061      0.289      -2.708       9.091
==============================================================================
Omnibus:                     4096.256   Durbin-Watson:                   1.865
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           493423.253
Skew:                           5.464   Prob(JB):                         0.00
Kurtosis:                      58.407   Cond. No.                         67.6
==============================================================================
```

- The R-square shows that the model explains %7 of the variation in variables which is low.

- However, the P-value shows significance between size and the collective variables (bedrooms, and livingrooms).

- An increase in bedrooms by 30m2 has an effect on size.

- An increase in livingrooms by 25m2 has an effect on size.

METHODOLOGY

RESULTS

Linear Regression: Actual vs Predicted Size with Regression Line
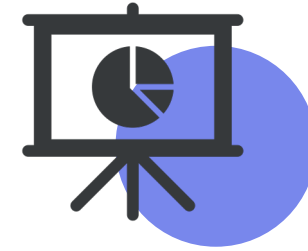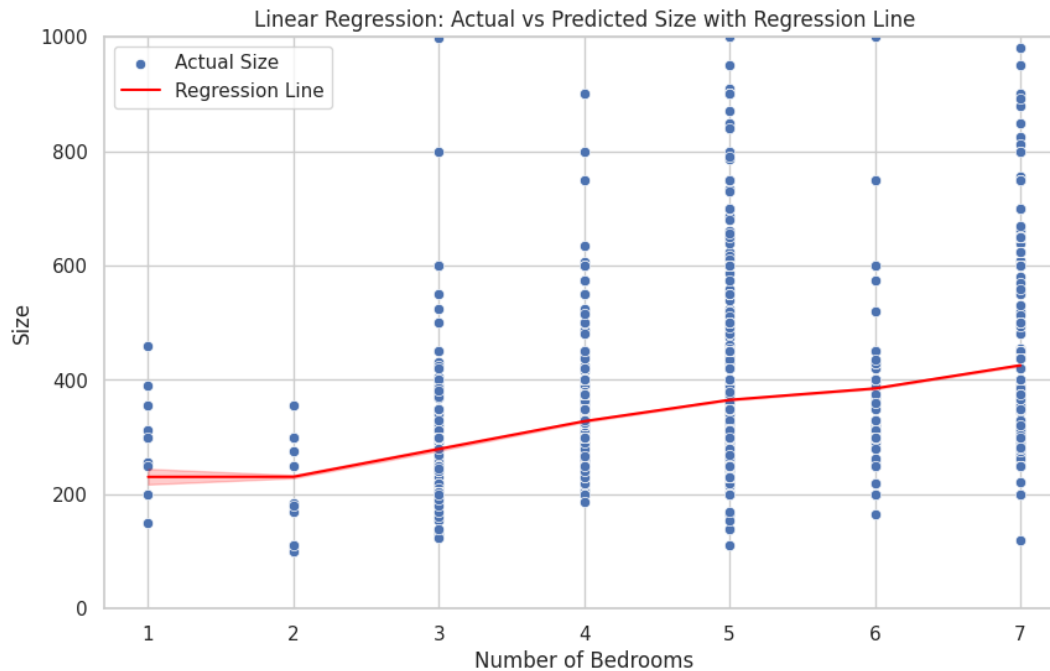
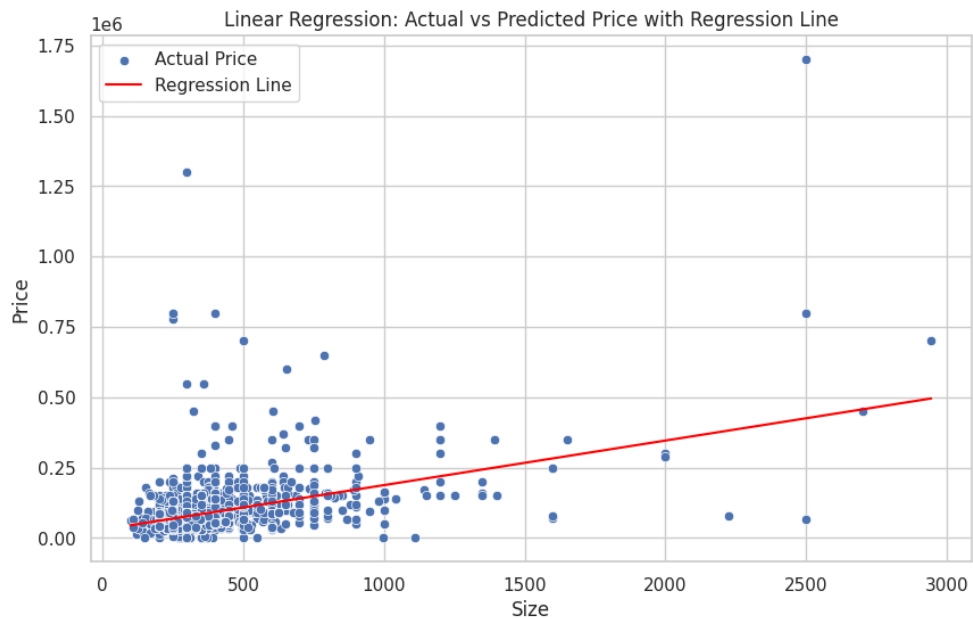- A linear regression study was preformed to predict the increase in size with the variation in number of bedrooms.

DISCUSSION

- The analysis is significant as an increase in number of bedrooms increases the size. Consequently, increasing the price on the property listing.

Linear Regression: Actual vs Predicted Price with Regression Line

```
                        OLS Regression Results
==============================================================================
Dep. Variable:                  price   R-squared:                       0.161
Model:                            OLS   Adj. R-squared:                  0.161
Method:                 Least Squares   F-statistic:                     710.7
Date:                Mon, 04 Dec 2023   Prob (F-statistic):          2.05e-143
Time:                        18:32:36   Log-Likelihood:                -46347.
No. Observations:                3713   AIC:                         9.270e+04
Df Residuals:                    3711   BIC:                         9.271e+04
Df Model:                           1
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const         2.964e+04   2397.666     12.361      0.000    2.49e+04    3.43e+04
size           158.4320      5.943     26.659      0.000     146.780     170.084
==============================================================================
Omnibus:                     4798.139   Durbin-Watson:                   1.878
Prob(Omnibus):                  0.000   Jarque-Bera (JB):          1650991.979
Skew:                           6.852   Prob(JB):                         0.00
Kurtosis:                     105.391   Cond. No.                         924.
==============================================================================
```
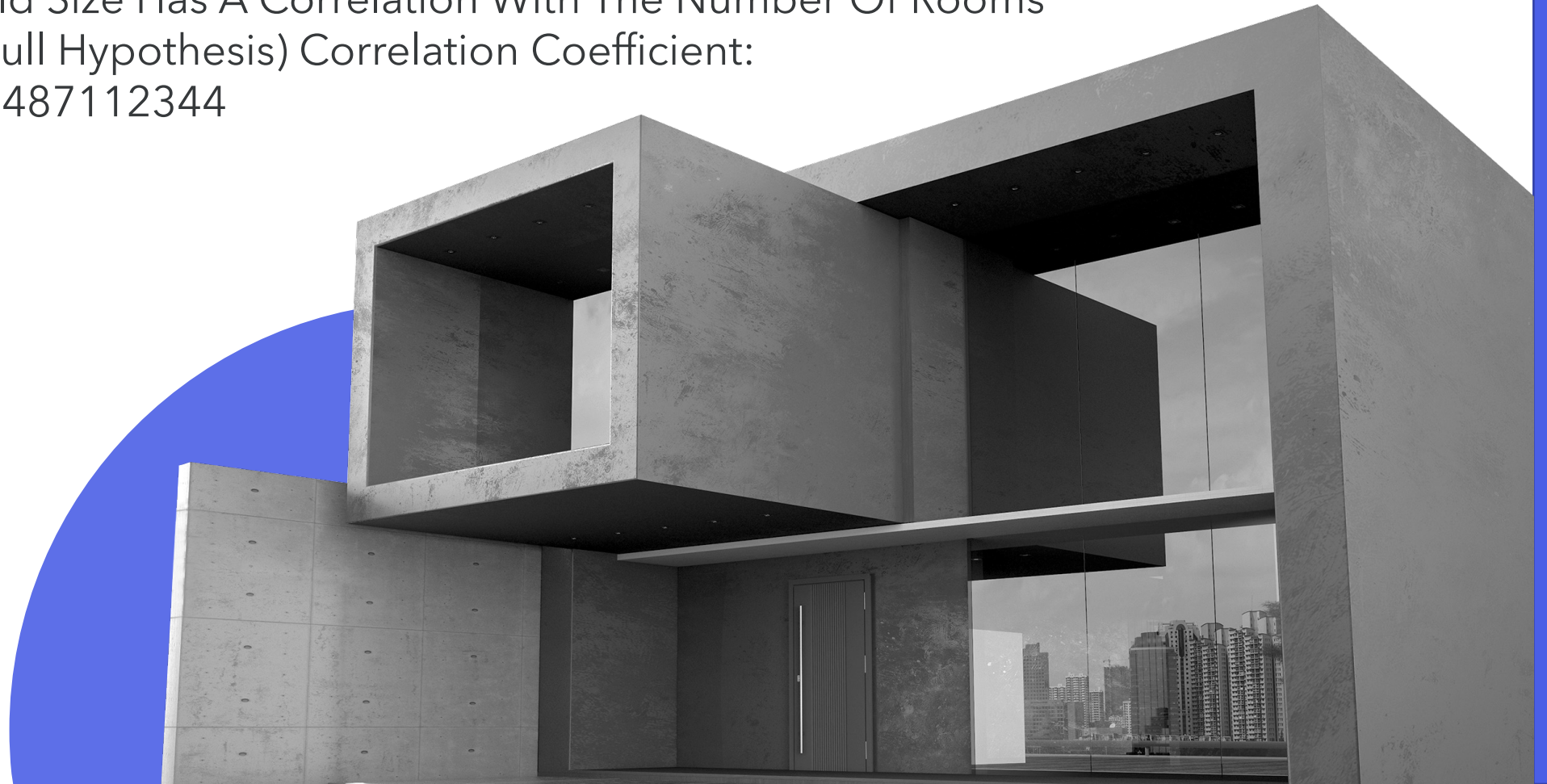
H1: Housing Size Has A Correlation With The Number Of Rooms (Reject The Null Hypothesis) Correlation Coefficient: 0.24127269546672064

H2: Household Size Has A Correlation With The Number Of Rooms (Reject The Null Hypothesis) Correlation Coefficient: 0.015519338487112344

THANK YOU