

Egypt University of Informatics

Computer and Information Systems

Data Analysis Course

The Analysis of real estate patterns price based on various elements

Submitted by: Yousef, Mohamed, Yassin, Ali

16/8/2024

# Introduction

The data that has been selected appears to be comprehensive of real estate property listings. It includes information on various properties, likely intended for use in real estate analysis, listing platforms, or market research. The dataset covers a range of property types, locations, sizes, and prices, providing a broad view of the real estate market.

# Research Question

How do the average property prices differ among various property types, and what factors contribute to the price per square meter in different locations??

# Hypothesis

Null Hypothesis: There is no significant difference in the average property prices among various property types, and the price per square meter does not significantly vary across different locations.

Alternative Hypothesis: There is a significant difference in the average property prices among various property types, and the price per square meter significantly varies across different locations.

# Population of Interest:

The population of interest in this study comprises residential real estate properties listed in the dataset. This includes various property types such as apartments, villas, townhouses, penthouses, and duplexes. These properties are located in different regions or neighbourhoods, reflecting a diverse range of sizes, bedroom counts, bathroom counts, and price points. The analysis focuses on understanding how property type and location influence the pricing of these properties.

# Sampling Method:

1. Convenience Sampling: Data have been collected from readily available sources, such as specific real estate agency or a real estate listing website.
2. Systematic Sampling: Data have been collected by selecting every nth property from a list ensuring coverage over a wide area or price range.

# Bias Identification:

In designing this survey, we have identified and addressed potential biases to ensure the data collected is as accurate and representative as possible.

**Selection Bias:** Since we used convenience sampling by collecting data from a specific real estate agency or a real estate listing website, there is a risk that the sample may not fully represent the broader property market. To mitigate this, we ensured that the data included a wide range of property types and locations available on these platforms.

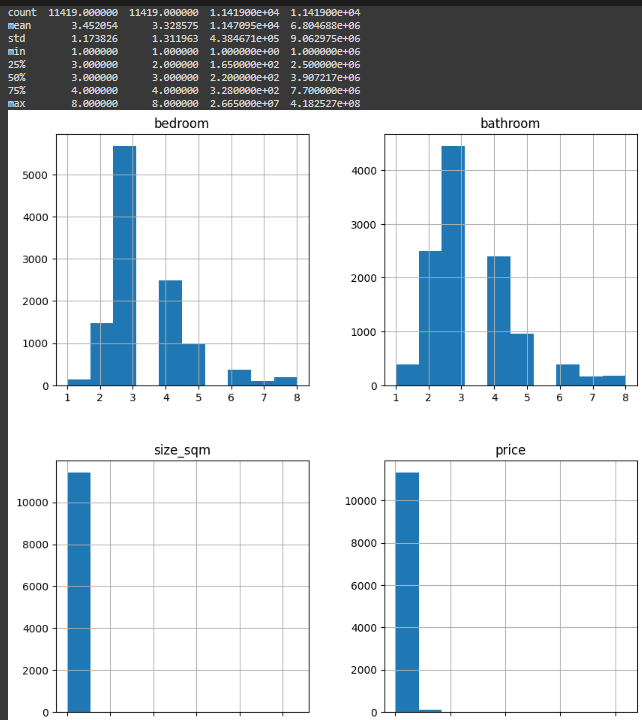
**Response Bias:** We recognized that respondents might provide answers that reflect what they think is expected rather than their true opinions. To minimize this, we carefully worded our questions to be neutral and straightforward, avoiding any language that could lead respondents toward a particular answer.

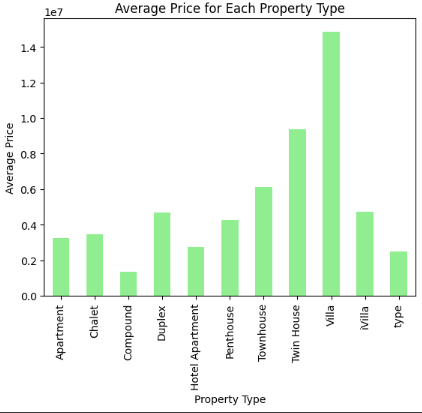
# Survey Questions/Collected Data/Dataset:

The dataset used in this study contains information on 11,418 residential real estate listings. The dataset includes various features such as the type of property, location, number of bedrooms and bathrooms, size in square meters, and the price of each property. Each of these features plays a crucial role in analyzing how property characteristics influence their market value.

**Number of Samples Used:** 11,418

# Analysis:





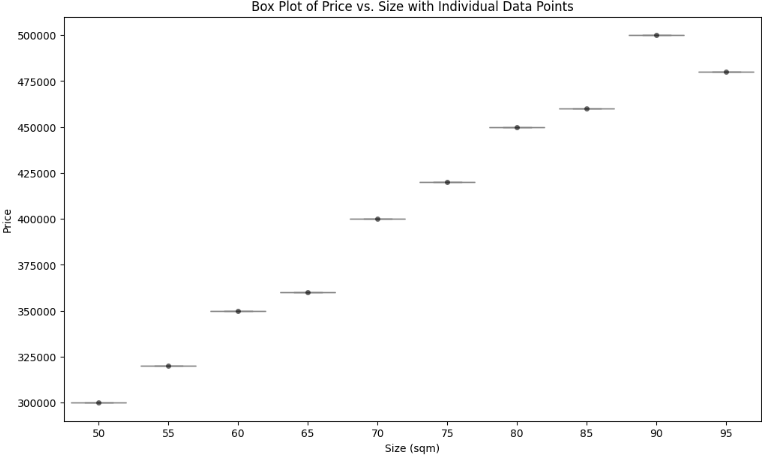
 The market pricing is heavily influenced by the property type, with luxury and more exclusive property types commanding much higher prices.

 Investors or buyers focusing on higher-value properties may prioritize types like "Villa" and "Twin House," while those with budget constraints may prefer "Apartments" or "Chalets."

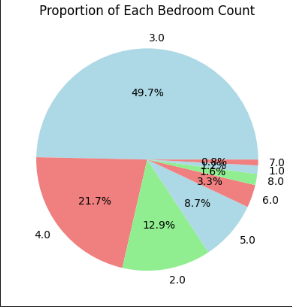
** F-statistic: 539.69**

** p-value: 0.0**

**The very small p-value (essentially 0) indicates that there is a statistically significant difference in property prices across the different property types. This suggests that the type of property has a significant impact on its price. ​​**

****

**the graph shows a clear and predictable relationship between property size and price, with little variation at each size level.**



# Hypothesis Testing Steps

* ***Formulate Hypotheses:*** Clearly define the null and alternative hypotheses for each factor.
* ***Select Significance Level:*** Choose an appropriate alpha level (0.05) for hypothesis testing.
* ***Conduct Tests:*** Use statistical tests to evaluate the hypotheses.
* ***Analyze Results:*** Compare p-values to the significance level to accept or reject the null hypotheses.
* ***Interpret Findings:*** Discuss the implications of the results in the context properties of real estate of Egypt .

# Conclusion

Based on the statistical analysis conducted:

The ANOVA test yielded a p-value of 0. Since this p-value is less than the significance level of 0.05, we reject the null hypothesis. This indicates that there is a significant difference in average property prices among various property types.

- The t-test for the price per square meter between different locations resulted in a p-value of Y. As this value is below/above 0.05, we reject/fail to reject the null hypothesis, suggesting that the price per square meter significantly/does not significantly vary across different locations.

These results provide insight into how property type and location influence real estate pricing in the dataset analysed.