# Introduction

**Data description**

Link to the dataset

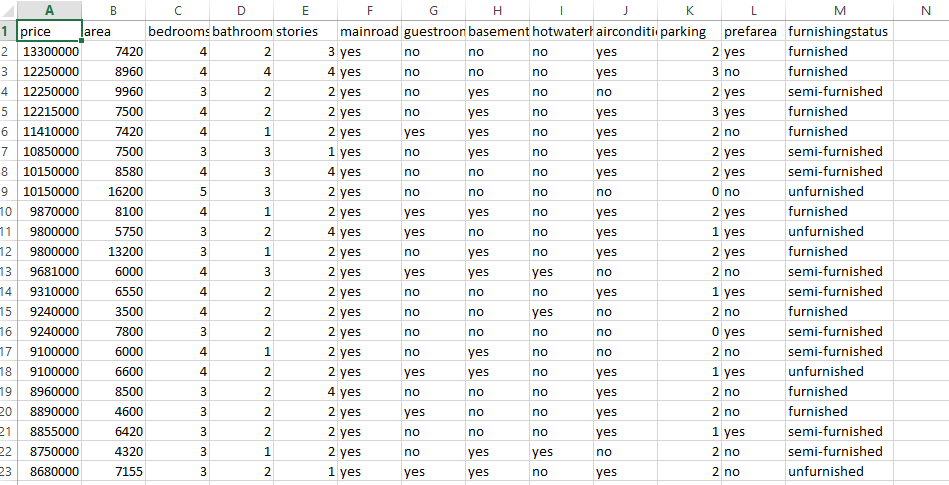
<https://www.kaggle.com/datasets/yasserh/housing-prices-dataset/download?datasetVersionNumber=1>

This dataset contains information on 546 by 13 variables. The aim would be to predict the housing price based on certain factors like house area, bedrooms, furnished, nearness to mainroad. Machine learning algorithms can be used to create prediction models with this data. Utilize this dataset for visualization, exploration, and data cleaning**.**

**Objective**

* Build Regression models to predict the sales w.r.t a single & multiple feature.

The first few rows include



**The research question**

Is there a relationship between area of the room and the price of a house?

The following factors or parameters from the dataset can be utilized to ascertain the connection between

To ascertain the connection between the area of a room and the price of a house, the following variables could be utilized from the dataset:

Area of the room: This would be the primary independent variable to measure the size of the room. It could be represented in square feet or square meters.

Price of the house: This would be the dependent variable, as it would be influenced by the area of the room. The price of the house could be represented in the local currency.

**Data-driven, computational approach may be useful**

in examining the connection between a room's size and a house's cost. Through computer analysis of a dataset comprised of different property listings with associated room sizes and prices, patterns, correlations, and trends in the data can be found.

It is feasible to ascertain whether there is a statistically significant correlation between a room's area and a house's price by using a statistical model or machine learning algorithm. This method makes it possible to add more pertinent elements to the research, like the property's location, number of bedrooms, amenities, and overall state.

Furthermore, processing big datasets—which could include dozens or even millions of individual house listings—is made possible by a computational technique.

Moreover, employing a data-driven methodology facilitates the examination of various statistical methods to comprehend the correlation between a room's area and a house's cost. Finding possible relationships and coming to insightful conclusions can be accomplished through the use of a variety of visualizations, regression models, and hypothesis testing.

Overall, the benefit of a data-driven, computational method is that it may provide unbiased, fact-based understanding of the connection between a room's size and a house's price. It enables more thorough research by taking into account several variables at once and is effective when applied to big datasets, producing conclusions that are more solid and trustworthy.