

# Data Visualization Project

In this document I give a brief intro about the project, what I found and info about the dataset I used

## The Data Set:-

I used the Housing prices dataset from kaggle from this [link](#)

Attached is file that contains the description of the fields in the Data set and after cleaning it consists of 1199 rows and 62 columns

## Main findings from the Dataset:-

In the attached slides presentation I explained my findings :-

- The most correlated numerical features with the Sales price are[ the overall quality of a house, its basement, its garage capacity , the ground living area and the garage area]
- The most common neighborhoods are [NAMES, CollgCR,,Oldtown, Edwards and Somerst]
- The most expensive neighborhoods are [StoneBr,NindgHT,NoRedge,Tember,Vedkeer]
- The most common sale type is the [WD] sale type
- The most common sale condition is the[Normal] sale condition
- There is a positive linear relationship between the sale price and the area
- The relationship between the sale price, area and the zoning classification[ We can see that the ['Residential Low Density'] classification zone is the most popular, most expensive and has the largest areas
- The price of the houses related to the year the house was built in[there is no linear relationship between the building year and the price]