Factors Influencing House prices in Ames, IOwa

Name

11/27/2022

## Introduction

The goal of this project is to analyze and visualize data from the kaggle competition project’s housing prices data collection. When asked to describe their ideal home, most purchasers would not start with the basement ceiling height or the closeness to an east-west railroad. This experiment demonstrates that price negotiations are influenced by factors other than the number of bedrooms or a white picket fence.

## Packages required

library(readxl)  
library(ggplot2)  
library(ggthemes)

## Data

When various criteria defining (almost) every characteristic of residential houses in Ames, Iowa are evaluated, home buyers may make selections on where to look for their dream home at a fair price. The Ames Housing dataset was created by Dean De Cock for use in data science courses. It’s a great choice for data scientists looking for an updated and improved version of the frequently discussed Boston Housing data collection. The data was previously segregated into training and testing sets because the competition was envisioned as a machine learning challenge, but we only consider the training set in our research.

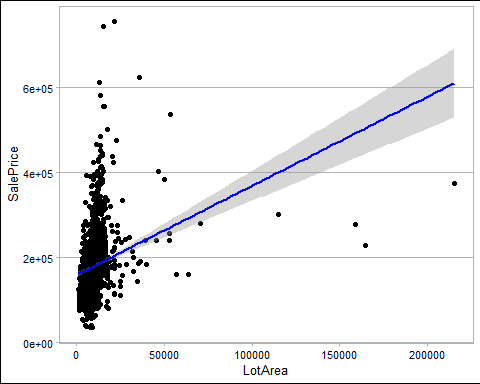
## Analysis and Results

Loading the data set in R as in the chunk below

train <- read\_excel("C:/Users/aland/OneDrive/Documents/house-prices-advanced-regression-techniques.xlsx")

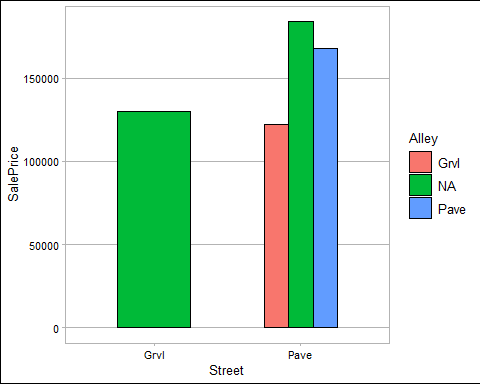
The first stage in data analysis will be to display the training set data using ggplot2 in R in what is known as exploratory data analysis (EDA). One of the elements that may influence the price of a property is lot area in square feet, which we can depict using ggplot2.

ggplot(train, aes(x=LotArea, y=SalePrice))+  
 geom\_point()+ theme\_calc() +  
geom\_smooth(method = 'lm', formula = y ~ x, se = TRUE, color = "blue")

 According to the scatter plot above, lot size does not appear to have a significant impact on home values. That instance, a large lot area does not always imply that the house will be pricey.

Aside from lot size, another aspect that may impact the price of a house is the type of road and alley access to the property. A ggplot can show us how pricing for paved and graveled roads, as well as ally type, compare.

ggplot(train, aes(x=Street, y=SalePrice,fill=Alley))+  
 geom\_bar(position=position\_dodge(), stat="summary", colour='black', width = 0.5)+theme\_calc()



There was no lane providing access to any of the residences on the graveled roadways. Houses with paved roads were more expensive than those with gravel roads. Surprisingly, residences without alleys were more expensive than those with graveled or paved alleyways.

Another element that might have an impact on house prices is the kind and design of home. The following types of dwellings are possible:

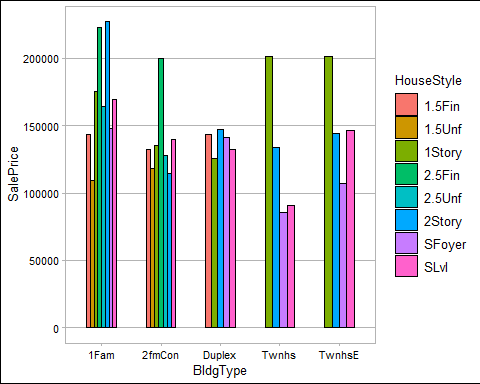
#### BldgType: Type of dwelling

1Fam Single-family Detached   
 2FmCon Two-family Conversion; originally built as one-family dwelling  
 Duplx Duplex  
 TwnhsE Townhouse End Unit  
 TwnhsI Townhouse Inside Unit

#### HouseStyle: Style of dwelling

1Story One story  
 1.5Fin One and one-half story: 2nd level finished  
 1.5Unf One and one-half story: 2nd level unfinished  
 2Story Two story  
 2.5Fin Two and one-half story: 2nd level finished  
 2.5Unf Two and one-half story: 2nd level unfinished  
 SFoyer Split Foyer  
 SLvl Split Level

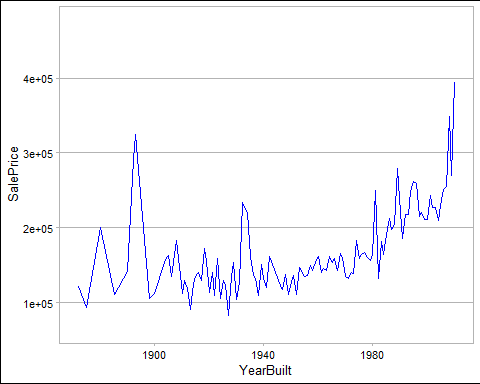
ggplot(train, aes(x=BldgType, y=SalePrice,fill=HouseStyle))+  
 geom\_bar(position=position\_dodge(), stat="summary", colour='black', width = 0.5)+theme\_calc()



The graph above clearly shows that properties in towns with one storey have moderately higher costs. It is also worth noting that houses for a single family that are two stories tend to be more expensive than any other categories.

Though we would anticipate house prices to rise over time owing to inflation, prices might also fluctuate over time due to new innovations, therefore we investigate the year that a house was built and their pricing.

ggplot(train, aes(x=YearBuilt, y=SalePrice))+  
 geom\_line(position=position\_dodge(), stat="summary", colour='blue', width = 0.5)+theme\_calc()

 Generally it is clear that the prices of the house were going up in time.

Another consideration that many buyers examine when purchasing a home is the kind of foundation and the assessed height of the basement. In the graph below, we group the foundation type and basement height to show how their prices compare.

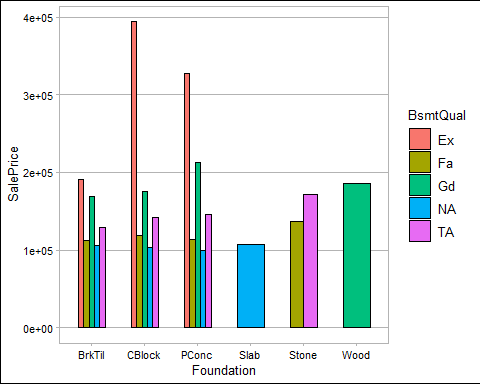
#### Foundation: Type of foundation

BrkTil Brick & Tile  
 CBlock Cinder Block  
 PConc Poured Contrete   
 Slab Slab  
 Stone Stone  
 Wood Wood

#### BsmtQual: Evaluates the height of the basement

Ex Excellent (100+ inches)   
 Gd Good (90-99 inches)  
 TA Typical (80-89 inches)  
 Fa Fair (70-79 inches)  
 Po Poor (<70 inches  
 NA No Basement

ggplot(train, aes(x=Foundation, y=SalePrice, fill=BsmtQual))+  
 geom\_bar(position=position\_dodge(), stat="summary", colour='black', width = 0.5)+theme\_calc()



Houses made of cinder block and poured concrete with great basement heights command very high rates.

## Conclusion/Summary

We infer that more than one factor influences property prices. It is commonly assumed that residences in prime locations would always be expensive, but this investigation proves otherwise. Many elements, such as lot size, street and alley access to the property, foundation and its appraisal, and the year the property was built, among others, constantly influence the price of the house.