

Project Proposal

Sujal Bharathi Harindranath

2023-11-01

Group Info:

Sujal Bharathi Harindranath

Title:

Deciphering Housing Market Dynamics: An Analysis of the Ames Housing Data Set

Narrative:

Understanding the dynamics of residential property valuation is crucial for both homeowners and potential buyers in the real estate market. The Ames Housing Data Set, meticulously curated by Dean De Cock, offers a comprehensive view of the housing market in Ames, Iowa. This data set, which has been refined for academic use, provides a unique opportunity to delve into the factors that influence the sale price of homes in this region. Through detailed statistical analysis and interpretation of the Ames Housing Data Set, we aspire to provide valuable insights for various stakeholders in the real estate market, contributing to more informed decision-making processes in this vital sector.

A broader research question that encompasses multiple aspects of the Ames Housing Data Set could be: “What are the key determinants of residential property values in Ames, and how do various structural, spatial, and functional characteristics of homes influence their market price?” This question allows for an expansive analysis that includes factors such as the age and condition of the property, size of different areas (basement, first floor, second floor), presence of additional features like fireplaces and porches, and the size of the lot.

Hypotheses:

1. Property Size and Sale Price: There is a positive correlation between the size of a property (measured in square feet) and its sale price. Larger homes, including those with more bedrooms and larger living areas, tend to sell for higher prices.
2. Location within the City: Properties located in certain neighborhoods or areas within Ames command higher sale prices due to factors such as perceived desirability, proximity to amenities, and lower crime rates.
3. Property Age and Sale Price: Newer homes, or those that have been recently remodeled, sell for higher prices compared to older homes. This hypothesis is based on the assumption that newer properties require less maintenance and are more likely to have modern amenities.

Data Description:

The Ames Housing Data Set, curated by Dean De Cock from Truman State University, is a subset of a larger collection detailing residential property information in Ames, Iowa. This specific dataset includes 1,359 observations (houses) with 27 numeric variables. These variables cover aspects such as lot size, overall quality and condition, year built and remodeled, square footage, number of bathrooms and bedrooms, garage size, and porch areas.

The data was originally compiled from 2006 to 2010 and is tailored for educational use in statistical analysis, especially in regression modeling in real estate. The data, in its original and modified forms, can be accessed through the Journal of Statistics Education website (<https://jse.amstat.org/v19n3/decock>). The subset for this project was refined to include only houses with less than 1500 square feet of living area, sold under ‘Normal’ conditions, and without any missing values. This dataset provides a comprehensive view of the housing market in Ames, making it suitable for various real estate market analyses.

Data Dictionary:

The full data set contains 1,359 homes and the following 27 variables. SalePrice is considered as the response variable here.

- **LotArea:** Lot size in square feet.
- **OverallQual:** Overall quality of the house’s material and finish. The scale ranges from 1 (Very Poor) to 9 (Very Excellent).
- **OverallCond:** Overall condition rating. The scale ranges from 1 (Very Poor) to 9 (Very Excellent).
- **YearBuilt:** Original construction date.
- **YearRemodAdd:** Remodel date.
- **BsmtFinSf1:** Type 1 finished square feet.
- **BsmtFinType2:** Type 2 finished square feet.
- **TotalBsmtSf:** Total square feet of basement area.
- **FirstFlrSf:** First floor square feet.
- **SecondFlrSf:** Second floor square feet.
- **GrLivArea:** Above grade (ground) living area square feet.
- **BsmtFullBath:** Number of full bathrooms in the basement.
- **BsmtHalfBath:** Number of half baths in the basement.
- **FullBath:** Number of full bathrooms above ground.
- **HalfBath:** Number of half baths above ground.
- **BedroomAbvGr:** Number of Bedrooms above ground.
- **KitchenAbvGr:** Number of Kitchens above ground.
- **TotRmsAbvGrd:** Total rooms above ground (does not include bathrooms).
- **Fireplaces:** Number of fireplaces.
- **GarageCars:** Size of garage in car capacity.
- **GarageArea:** Size of garage in square feet.
- **WoodDeckSf:** Wood deck area in square feet.
- **OpenPorchSf:** Open porch area in square feet.
- **EnclosedPorch:** Enclosed porch area in square feet.
- **ThreeSsnPorch:** Three season porch area in square feet.
- **ScreenPorch:** Screen porch area in square feet.
- **SalePrice:** The property’s sale price in dollars. This is the response variable.

Evidence of Data:

##	LotArea	OverallQual	OverallCond	YearBuilt	YearRemodAdd	BsmtFinSf1	BsmtFinSf2
## 1	11622	5	6	1961	1961	468	144
## 2	14267	6	6	1958	1958	923	0
## 3	4920	8	5	2001	2001	616	0
## 4	5005	8	5	1992	1992	263	0
## 5	7980	6	7	1992	2007	935	0
## 6	8402	6	5	1998	1998	0	0
## 7	10176	7	5	1990	1990	637	0
## 8	19138	4	5	1951	1951	120	0
## 9	11241	6	7	1970	1970	578	0
## 10	12537	5	6	1971	2008	734	0
##	TotalBsmtSf	FirstFlrSf	SecondFlrSf	GrLivArea	BsmtFullBath	BsmtHalfBath	

## 1	882	896	0	896	0	0
## 2	1329	1329	0	1329	0	0
## 3	1338	1338	0	1338	1	0
## 4	1280	1280	0	1280	0	0
## 5	1168	1187	0	1187	1	0
## 6	789	789	676	1465	0	0
## 7	1300	1341	0	1341	1	0
## 8	864	864	0	864	0	0
## 9	1004	1004	0	1004	1	0
## 10	1078	1078	0	1078	1	0
##	FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr	TotRmsAbvGrd	Fireplaces
## 1	1	0	2	1	5	0
## 2	1	1	3	1	6	0
## 3	2	0	2	1	6	0
## 4	2	0	2	1	5	0
## 5	2	0	3	1	6	0
## 6	2	1	3	1	7	1
## 7	1	1	2	1	5	1
## 8	1	0	2	1	4	0
## 9	1	0	2	1	5	1
## 10	1	1	3	1	6	1
##	GarageCars	GarageArea	WoodDeckSf	OpenPorchSf	EnclosedPorch	ThreeSsnPorch
## 1	1	730	140	0	0	0
## 2	1	312	393	36	0	0
## 3	2	582	0	0	170	0
## 4	2	506	0	82	0	0
## 5	2	420	483	21	0	0
## 6	2	393	0	75	0	0
## 7	2	506	192	0	0	0
## 8	2	400	0	0	0	0
## 9	2	480	0	0	0	0
## 10	2	500	0	0	0	0
##	ScreenPorch	SalePrice				
## 1	120	105000				
## 2	0	172000				
## 3	0	213500				
## 4	144	191500				
## 5	0	185000				
## 6	0	180400				
## 7	0	171500				
## 8	0	141000				
## 9	0	149000				
## 10	0	149900				

Code Appendix

```
library(tidyverse, quietly = TRUE)

ames_housing <- read.csv("/Users/Suju/Downloads/ames_housing.csv")

data <- as_tibble(ames_housing)

head(ames_housing, n = 10)
```