

Housing Price Analysis

By: Charles Dillaway

Agenda

Some Topics For This Presentation:

- Client Background
- Project Purpose
- Data Processing
- Exploratory Data Analysis
- Machine Learning Predictions
- Recommendations/Conclusion

Background

For this slide I want to dive into background info on the client, I haven't thought of what would make a good client yet though

Goals

1. Uncover patterns within the dataset to understand the market demand of houses and their features in Ames, Iowa
2. Prepare the dataset for modeling through data-processing
3. Develop the machine learning model to drive accurate predictions of housing prices within Ames, Iowa

Purpose

My Analysis will show:

1. How the housing market of Ames, Iowa differs from that of other regions within the United States.
2. Why understanding the market differences within the town of Ames will help us drive more accurate predictions.
3. How the developed machine learning model will help all sorts of industries understand the housing market within Ames

Data Processing

Data Overview

Shape of the dataframe: 2930 Rows, 82 Columns

Total Missing Values: 15749 NaN values

Number of Object Columns: 43 object Columns

Key Columns for the dataset:

- SalePrice (target Variable)
- Overall Qual (Strong Correlation with Target Variable)
- Garage Area (highlights an interesting fact about the community)

Handling Object Values

I will talk about why I had to handle them, and will mention atleast one of the methods that I used in order to change the object columns into a integer object

Missing Values

Will discussed why i dropped NaN values after I converted them to object values

Outlier Handling

Will discuss how I handled the outliers, so I could create two reliable regression plots

Exploratory Data Analysis

Visualization 1

Is a histogram that shows the
distributions of housing prices

Visualization 2

Is a heatmap that shows the correlation coefficients of each feature by the target variable

Visualization 3

Is a bar plot that shows the sale price
by the type of garage



Visualization 4

Is a bar plot that shows the sale price
by the overall quality of the house



Visualization 5

Is a regression plot that shows the correlation between the sale price and the Lot area

Visualization 6

Is a regression plot that shows the correlation between the sale price by the area of the garage

Predictions

Model Performance

This will explain the type of model that I chose (linear regression) and it will also show the performance metrics of the model to show that it would be reliable at making predictions on the cost of a house

Exploratory Data Analysis

Recommendations

This will discuss the recommendations that I have for the client

In Conclusion

Will explain everything that we did throughout the presentation (restating the agenda)

Questions?