Project 8

New York Home Price Modeling

Description

This project uses Data Science to predict home prices using various features for the state of New York. The goal is to predict accurate prices so consumers can make well informed decisions when selling their homes. The project uses various visualizations to understand the data provided. The methods used in this project involve two models: Linear Regression and Random Forests Regression.

The goal of this project is to find a model that can accurately predict home prices to help consumers accurately price their homes when listing them on the market.

Data Dictionary

Data Dictionary

column	description	data_type
status	Current standing of the home (for sale or ready to build)	object
bed	Number of beds in the home	float64
bath	Number of baths in the home	float64
acre_lot	Size of the lot	float64
city	City where the home is located	object
state	State where the home is located	object
zip_code	Zip code of the home	float64
house_size	Square fottage of the home	float64
prev_sold_date	Date when the home was previously sold	object
price	Current sale price or previously sold price if the house is not for sale	float64

Data Transformations

Nulls droped for modeling

Latitude and Longitude added from nomi

NY Subset created for modeling

- Subset 1 includes zip code and prices

- Subset 2 includes all numeric fields in data

Project Status

The project has been submitted and waiting sign-off from stakeholders