

Final Project Proposal

Data Set

- **Link to Data Set:** [Kaggle House Prices Dataset](#)
- **Source and Collection Details:**
 - **Collected By:** Ames Housing Dataset
 - **Collection Method:** Aggregated from public records
 - **Collection Period:** 2006-2010
- **Number of Observations:** 1460
- **Column Names and Descriptions:**
 - **SalePrice:** The property's sale price in dollars (target variable).
 - **MSSubClass:** The building class.
 - **MSZoning:** The general zoning classification.
 - **LotFrontage:** Linear feet of street connected to property.
 - **LotArea:** Lot size in square feet.
 - **Street:** Type of road access.
 - **Alley:** Type of alley access.
 - **LotShape:** General shape of property.
 - **LandContour:** Flatness of the property.
 - **Utilities:** Type of utilities available.
 - **LotConfig:** Lot configuration.
 - **LandSlope:** Slope of property.
 - **Neighborhood:** Physical locations within Ames city limits.
 - **Condition1:** Proximity to main road or railroad.
 - **Condition2:** Proximity to main road or railroad (if a second is present).
 - **BldgType:** Type of dwelling.
 - **HouseStyle:** Style of dwelling.
 - **OverallQual:** Overall material and finish quality.
 - **OverallCond:** Overall condition rating.
 - **YearBuilt:** Original construction date.
 - **YearRemodAdd:** Remodel date.
 - **RoofStyle:** Type of roof.
 - **RoofMatl:** Roof material.
 - **Exterior1st:** Exterior covering on house.
 - **Exterior2nd:** Exterior covering on house (if more than one material).
 - **MasVnrType:** Masonry veneer type.
 - **MasVnrArea:** Masonry veneer area in square feet.
 - **ExterQual:** Exterior material quality.
 - **ExterCond:** Present condition of the material on the exterior.
 - **Foundation:** Type of foundation.
 - **BsmtQual:** Height of the basement.
 - **BsmtCond:** General condition of the basement.
 - **BsmtExposure:** Walkout or garden level basement walls.
 - **BsmtFinType1:** Quality of basement finished area.
 - **BsmtFinSF1:** Type 1 finished square feet.
 - **BsmtFinType2:** Quality of second finished area (if present).
 - **BsmtFinSF2:** Type 2 finished square feet.
 - **BsmtUnfSF:** Unfinished square feet of basement area.

- **TotalBsmtSF:** Total square feet of basement area.
 - **Heating:** Type of heating.
 - **HeatingQC:** Heating quality and condition.
 - **CentralAir:** Central air conditioning.
 - **Electrical:** Electrical system.
 - **1stFlrSF:** First Floor square feet.
 - **2ndFlrSF:** Second floor square feet.
 - **LowQualFinSF:** Low quality finished square feet (all floors).
 - **GrLivArea:** Above grade (ground) living area square feet.
 - **BsmtFullBath:** Basement full bathrooms.
 - **BsmtHalfBath:** Basement half bathrooms.
 - **FullBath:** Full bathrooms above grade.
 - **HalfBath:** Half baths above grade.
 - **Bedroom:** Number of bedrooms above basement level.
 - **Kitchen:** Number of kitchens.
 - **KitchenQual:** Kitchen quality.
 - **TotRmsAbvGrd:** Total rooms above grade (does not include bathrooms).
 - **Functional:** Home functionality rating.
 - **Fireplaces:** Number of fireplaces.
 - **FireplaceQu:** Fireplace quality.
 - **GarageType:** Garage location.
 - **GarageYrBlt:** Year garage was built.
 - **GarageFinish:** Interior finish of the garage.
 - **GarageCars:** Size of garage in car capacity.
 - **GarageArea:** Size of garage in square feet.
 - **GarageQual:** Garage quality.
 - **GarageCond:** Garage condition.
 - **PavedDrive:** Paved driveway.
 - **WoodDeckSF:** Wood deck area in square feet.
 - **OpenPorchSF:** Open porch area in square feet.
 - **EnclosedPorch:** Enclosed porch area in square feet.
 - **3SsnPorch:** Three season porch area in square feet.
 - **ScreenPorch:** Screen porch area in square feet.
 - **PoolArea:** Pool area in square feet.
 - **PoolQC:** Pool quality.
 - **Fence:** Fence quality.
 - **MiscFeature:** Miscellaneous feature not covered in other categories.
 - **MiscVal:** \$ Value of miscellaneous feature.
 - **MoSold:** Month Sold.
 - **YrSold:** Year Sold.
 - **SaleType:** Type of sale.
 - **SaleCondition:** Condition of sale.
- **Privacy/Confidentiality Notes:** data is public.

Context and Motivation

- **Context:** The housing market is a critical component of the economy. Understanding the factors that influence house prices can provide insights for buyers, sellers, and policymakers.
- **Motivation:** This analysis aims to identify key determinants of house prices, which can help in making informed decisions about buying or selling properties and policy formulation.

Research Questions

- **Subject Matter Research Questions:**
 - What are the main factors affecting house prices in the Ames data set?
 - How does the age of a house influence its price?
- **Translation to Statistical Research Questions:**
 - Which variables significantly predict house prices in a regression model?
 - Is there a statistically significant relationship between the year a house was built and its sale price?

Methods

- **Multiple Linear Regression:**
 - **How it will be used** - To model the relationship between sale price and various predictors (e.g., lot area, overall quality, year built).
 - **Why it's appropriate** - It allows for the assessment of multiple predictors simultaneously and helps in understanding their individual contributions.
- **t-Test:**
 - **How it will be used** - To compare the mean prices of houses built before and after a certain year (e.g., 2006).
 - **Why it's appropriate** - It helps to determine if there is a significant difference in prices between two groups of houses.
- **Chi-Squared Test:**
 - **How it will be used** - To examine the association between categorical variables such as house style and neighborhood.
 - **Why it's appropriate** - It's suitable for testing relationships between categorical variables.

END