Company: FlipRoboTechnologies

Location: Banglore

.....Project Report : Housing Project.....

.....Prediction of Sale Price.....

.....SUBMITTED BY .....Jyoti Chaudhary

Under the guidance of DataTrained.....

CERTIFICATE

This is to certify that the this project report entitled "**Housing SalePrice Project ANALYSIS & PREDICTION**" is a bonafide work carried out by Jyoti Chaudhary in Project fulfillment for **DataScience** from

FlipRoboTechnologies- Internship during the year 2022-2023. This Project Report has been approved as it satisfies the company requirements.

.....Housing Project .....

Problem statement

Houses are one of the necessary need of each and every person around the globe and therefore housing and real estate market is

one of the markets which is one of the major contributors in the world's economy. It is a very large market and there are

various companies working in the domain. Data science comes as a very important tool to solve problems in the domain to help

the companies increase their overall revenue, profits, improving their marketing strategies and focusing on changing trends in

house sales and purchases. Predictive modelling, Market mix modelling, recommendation systems are some of the machine learning

techniques used for achieving the business goals for housing companies. Our problem is related to one such housing company.

A US-based housing company named Surprise Housing has decided to enter the Australian market. The company uses data analytics

to purchase houses at a price below their actual values and flip them at a higher price. For the same purpose, the company has

collected a data set from the sale of houses in Australia. The data is provided in the CSV file below. The company is looking

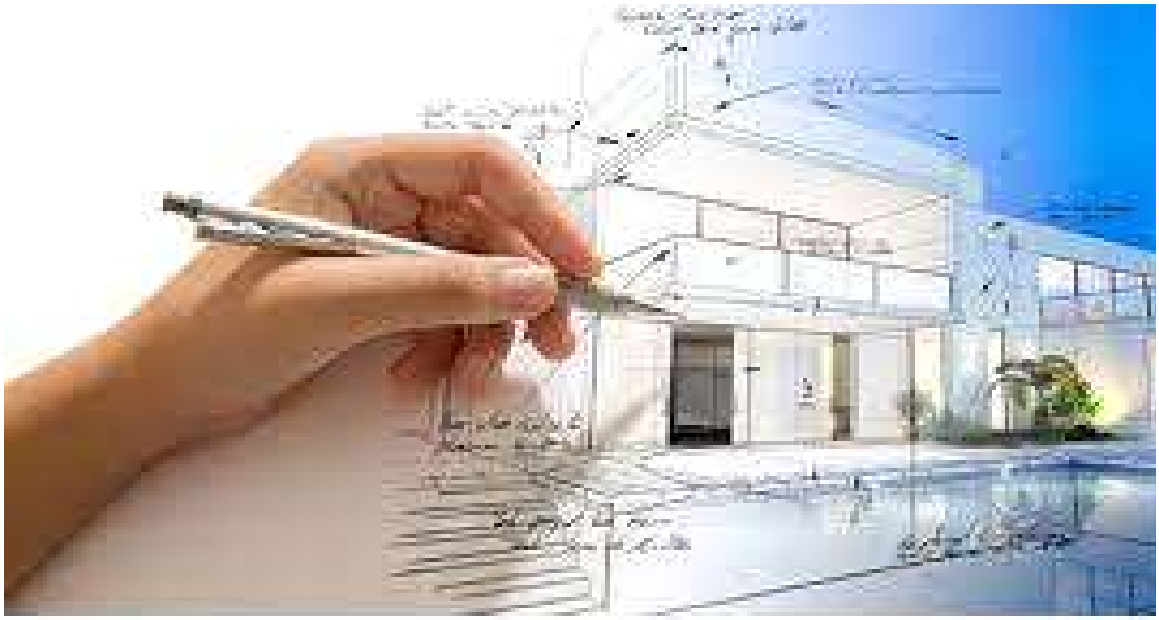
at prospective properties to buy houses to enter the market. You are required to build a model using Machine Learning in order

to predict the actual value of the prospective properties and decide

whether to invest in them or not. For this company wants

to know: • Which variables are important to predict the price of variable? • How do these variables describe the price of the

house?



Attributes Information : .....

MSSubClass: Identifies the type of dwelling involved in the sale.

20	**1-STORY 1946 & NEWER ALL STYLES**
30	**1-STORY 1945 & OLDER**
40	**1-STORY W/FINISHED ATTIC ALL AGES**
45	**1-1/2 STORY - UNFINISHED ALL AGES**
50	**1-1/2 STORY FINISHED ALL AGES**

60 2-STORY 1946 & NEWER

70 2-STORY 1945 & OLDER

75 2-1/2 STORY ALL AGES

80 SPLIT OR MULTI-LEVEL

85 SPLIT FOYER

90 DUPLEX - ALL STYLES AND AGES

120 1-STORY PUD (Planned Unit Development) - 1946 & NEWER

150 1-1/2 STORY PUD - ALL AGES

160 2-STORY PUD - 1946 & NEWER

180 PUD - MULTILEVEL - INCL SPLIT LEV/FOYER

190 2 FAMILY CONVERSION - ALL STYLES AND AGES

MSZoning: Identifies the general zoning classification of the sale.

A Agriculture

C Commercial

FVFloating Village Residential

I Industrial

RH Residential High Density

RLResidential Low Density

RP Residential Low Density Park

RM Residential Medium Density

LotFrontage: Linear feet of street connected to property

LotArea: Lot size in square feet

Street: Type of road access to property

Grvl	**Gravel**
Pave	**Paved**

Alley: Type of alley access to property

Grvl	**Gravel**
Pave	**Paved**
NA	**No alley access**

LotShape: General shape of property

Reg	**Regular**
IR1	**Slightly irregular**
IR2	**Moderately Irregular**
IR3	**Irregular**

LandContour: Flatness of the property

Lvl **Near Flat/Level**

Bnk **Banked - Quick and significant rise from street grade to building**

HLS **Hillside - Significant slope from side to side**

Low **Depression**

Utilities: Type of utilities available

AllPub **All public Utilities (E,G,W,& S)**

NoSewr **Electricity, Gas, and Water (Septic Tank)**

NoSeWa **Electricity and Gas Only**

ELO **Electricity only**

LotConfig: Lot configuration

Inside **Inside lot**

Corner **Corner lot**

CulDSac **Cul-de-sac**

FR2 **Frontage on 2 sides of property**

FR3 **Frontage on 3 sides of property**

LandSlope: Slope of property

Gtl Gentle slope

Mod Moderate Slope

Sev Severe Slope

Neighborhood: Physical locations within Ames city limits

Blmngtn Bloomington Heights

Blueste Bluestem

BrDale Briardale

BrkSide Brookside

ClearCr Clear Creek

CollgCr College Creek

Crawfor Crawford

Edwards Edwards

Gilbert Gilbert

IDOTRR Iowa DOT and Rail Road

MeadowV Meadow Village

Mitchel Mitchell

Names North Ames

NoRidge Northridge

NPkVill Northpark Villa

NridgHt Northridge Heights

NWAmes Northwest Ames

OldTown Old Town

SWISU South & West of Iowa State University

Sawyer Sawyer

SawyerW Sawyer West

Somerst Somerset

StoneBr Stone Brook

Timber Timberland

Veenker Veenker

Condition1: Proximity to various conditions

Artery Adjacent to arterial street

Feedr Adjacent to feeder street

Norm Normal

RRNn Within 200' of North-South Railroad

RRAn Adjacent to North-South Railroad

PosN Near positive off-site feature--park, greenbelt, etc.

PosA Adjacent to postive off-site feature

RRNe Within 200' of East-West Railroad

RRAe Adjacent to East-West Railroad

Condition2: Proximity to various conditions (if more than one is present)

Artery	**Adjacent to arterial street**
Feedr	**Adjacent to feeder street**
Norm	**Normal**
RRNn	**Within 200' of North-South Railroad**
RRAn	**Adjacent to North-South Railroad**
PosN	**Near positive off-site feature--park, greenbelt, etc.**
PosA	**Adjacent to postive off-site feature**
RRNe	**Within 200' of East-West Railroad**
RRAe	**Adjacent to East-West Railroad**

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

OverallQual: Rates the overall material and finish of the house

10 Very Excellent

9 Excellent

8 Very Good

7 Good

6 Above Average

5 Average

4 Below Average

3 Fair

2 Poor

1 Very Poor

OverallCond: Rates the overall condition of the house

10 Very Excellent

9 Excellent

8 Very Good

7 Good

6 Above Average

5 Average

4 Below Average

3 Fair

2 Poor

1 Very Poor

YearBuilt: Original construction date

YearRemodAdd: Remodel date (same as construction date if no remodeling or additions)

RoofStyle: Type of roof

Flat Flat

Gable Gable

Gambrel Gabrel (Barn)

Hip Hip

Mansard Mansard

Shed Shed

RoofMatl: Roof material

ClyTile Clay or Tile

CompShg Standard (Composite) Shingle

Membran Membrane

Metal Metal

Roll Roll

Tar&Grv Gravel & Tar

WdShake Wood Shakes

WdShngl Wood Shingles

Exterior1st: Exterior covering on house

AsbShng Asbestos Shingles

AsphShn Asphalt Shingles

BrkComm Brick Common

BrkFace Brick Face

CBlock Cinder Block

CemntBd Cement Board

HdBoard Hard Board

ImStucc Imitation Stucco

MetalSd Metal Siding

Other Other

Plywood Plywood

PreCast PreCast

Stone Stone

Stucco Stucco

VinylSd Vinyl Siding

Wd Sdng Wood Siding

WdShing Wood Shingles

Exterior2nd: Exterior covering on house (if more than one material)

AsbShng Asbestos Shingles

AsphShn Asphalt Shingles

BrkComm Brick Common

BrkFace Brick Face

CBlock Cinder Block

CemntBd Cement Board

HdBoard Hard Board

ImStucc Imitation Stucco

MetalSd Metal Siding

Other Other

Plywood Plywood

PreCast PreCast

Stone Stone

Stucco Stucco

VinylSd Vinyl Siding

Wd Sdng Wood Siding

WdShing Wood Shingles

MasVnrType: Masonry veneer type

BrkCmn Brick Common

BrkFace Brick Face

CBlock Cinder Block

None None

Stone Stone

MasVnrArea: Masonry veneer area in square feet

ExterQual: Evaluates the quality of the material on the exterior

Ex Excellent

Gd Good

TA Average/Typical

Fa Fair

Po Poor

ExterCond: Evaluates the present condition of the material on the exterior

Ex Excellent

Gd Good

TA Average/Typical

Fa Fair

Po Poor

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

BsmtCond: Evaluates the general condition of the basement

Ex Excellent

Gd Good

TA Typical - slight dampness allowed

Fa Fair - dampness or some cracking or settling

Po Poor - Severe cracking, settling, or wetness

NA No Basement

BsmtExposure: Refers to walkout or garden level walls

Gd Good Exposure

Av Average Exposure (split levels or foyers typically score average or above)

Mn Mimimum Exposure

No No Exposure

NA No Basement

BsmtFinType1: Rating of basement finished area

GLQ Good Living Quarters

ALQ Average Living Quarters

BLQ Below Average Living Quarters

Rec Average Rec Room

LwQ Low Quality

Unf Unfinshed

NA No Basement

BsmtFinSF1: Type 1 finished square feet

BsmtFinType2: Rating of basement finished area (if multiple types)

GLQ Good Living Quarters

ALQ Average Living Quarters

BLQ Below Average Living Quarters

Rec Average Rec Room

LwQ Low Quality

Unf Unfinshed

NA No Basement

BsmtFinSF2: Type 2 finished square feet

BsmtUnfSF: Unfinished square feet of basement area

TotalBsmtSF: Total square feet of basement area

Heating: Type of heating

Floor	**Floor Furnace**
GasA	**Gas forced warm air furnace**
GasW	**Gas hot water or steam heat**
Grav	**Gravity furnace**
OthW	**Hot water or steam heat other than gas**
Wall	**Wall furnace**

HeatingQC: Heating quality and condition

Ex	**Excellent**
Gd	**Good**
TA	**Average/Typical**
Fa	**Fair**
Po	**Poor**

CentralAir: Central air conditioning

N	**No**

Y Yes

Electrical: Electrical system

SBrkr	**Standard Circuit Breakers & Romex**
FuseA	**Fuse Box over 60 AMP and all Romex wiring (Average)**
FuseF	**60 AMP Fuse Box and mostly Romex wiring (Fair)**
FuseP	**60 AMP Fuse Box and mostly knob & tube wiring (poor)**
Mix	**Mixed**

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: Bedrooms above grade (does NOT include basement bedrooms)

Kitchen: Kitchens above grade

KitchenQual: Kitchen quality

Ex	**Excellent**
Gd	**Good**
TA	**Typical/Average**
Fa	**Fair**
Po	**Poor**

TotRmsAbvGrd: Total rooms above grade (does not include bathrooms)

Functional: Home functionality (Assume typical unless deductions are warranted)

Typ	**Typical Functionality**
Min1	**Minor Deductions 1**
Min2	**Minor Deductions 2**

Mod Moderate Deductions

Maj1 Major Deductions 1

Maj2 Major Deductions 2

Sev Severely Damaged

Sal Salvage only

Fireplaces: Number of fireplaces

FireplaceQu: Fireplace quality

Ex Excellent - Exceptional Masonry Fireplace

Gd Good - Masonry Fireplace in main level

TA Average - Prefabricated Fireplace in main living area or Masonry Fireplace in basement

Fa Fair - Prefabricated Fireplace in basement

Po Poor - Ben Franklin Stove

NA No Fireplace

GarageType: Garage location

2Types More than one type of garage

Attchd Attached to home

Basment Basement Garage

BuiltIn Built-In (Garage part of house - typically has room above garage)

CarPort Car Port

Detchd Detached from home

NA No Garage

GarageYrBlt: Year garage was built

GarageFinish: Interior finish of the garage

Fin Finished

RFn Rough Finished

Unf Unfinished

NA No Garage

GarageCars: Size of garage in car capacity

GarageArea: Size of garage in square feet

GarageQual: Garage quality

Ex Excellent

Gd Good

TA Typical/Average

Fa Fair

Po Poor

NA No Garage

GarageCond: Garage condition

Ex Excellent

Gd Good

TA Typical/Average

Fa Fair

Po Poor

NA No Garage

PavedDrive: Paved driveway

Y Paved

P Partial Pavement

N Dirt/Gravel

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

Ex	**Excellent**
Gd	**Good**
TA	**Average/Typical**
Fa	**Fair**
NA	**No Pool**

Fence: Fence quality

GdPrv	**Good Privacy**
MnPrv	**Minimum Privacy**
GdWo	**Good Wood**
MnWw	**Minimum Wood/Wire**
NA	**No Fence**

MiscFeature: Miscellaneous feature not covered in other categories

Elev	**Elevator**
Gar2	**2nd Garage (if not described in garage section)**
Othr	**Other**
Shed	**Shed (over 100 SF)**
TenC	**Tennis Court**
NA	**None**

MiscVal: $Value of miscellaneous feature

MoSold: Month Sold (MM)

YrSold: Year Sold (YYYY)

SaleType: Type of sale

WD	**Warranty Deed - Conventional**
CWD	**Warranty Deed - Cash**
VWD	**Warranty Deed - VA Loan**
New	**Home just constructed and sold**
COD	**Court Officer Deed/Estate**

Con Contract 15% Down payment regular terms

ConLw Contract Low Down payment and low interest

ConLI Contract Low Interest

ConLD Contract Low Down

Oth Other

SaleCondition: Condition of sale

Normal Normal Sale

Abnorml Abnormal Sale - trade, foreclosure, short sale

AdjLand Adjoining Land Purchase

Alloca Allocation - two linked properties with separate deeds, typically condo with a garage unit

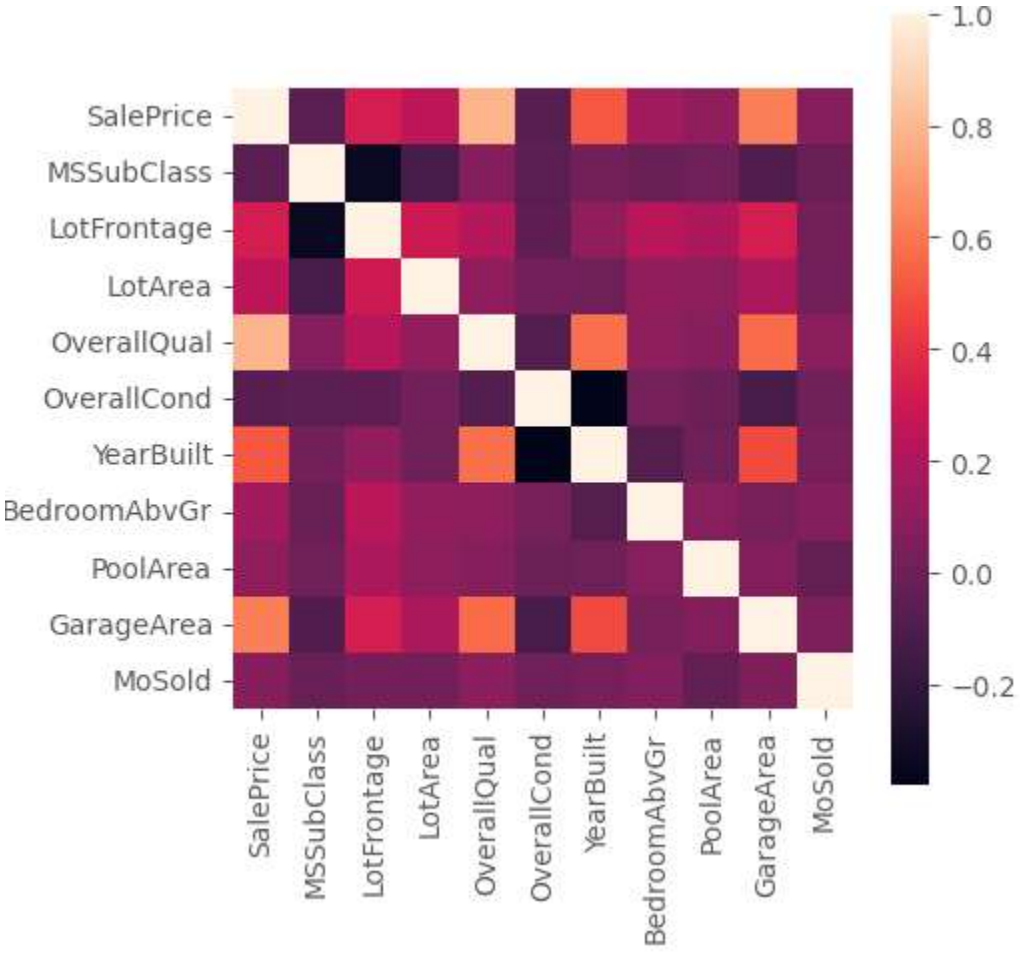
Family Sale between family members

Partial Home was not completed when last assessed (associated with New Homes)00000000000000000000

Experimental Results :

Analysis.....

Correlation:



CONCLUSION:

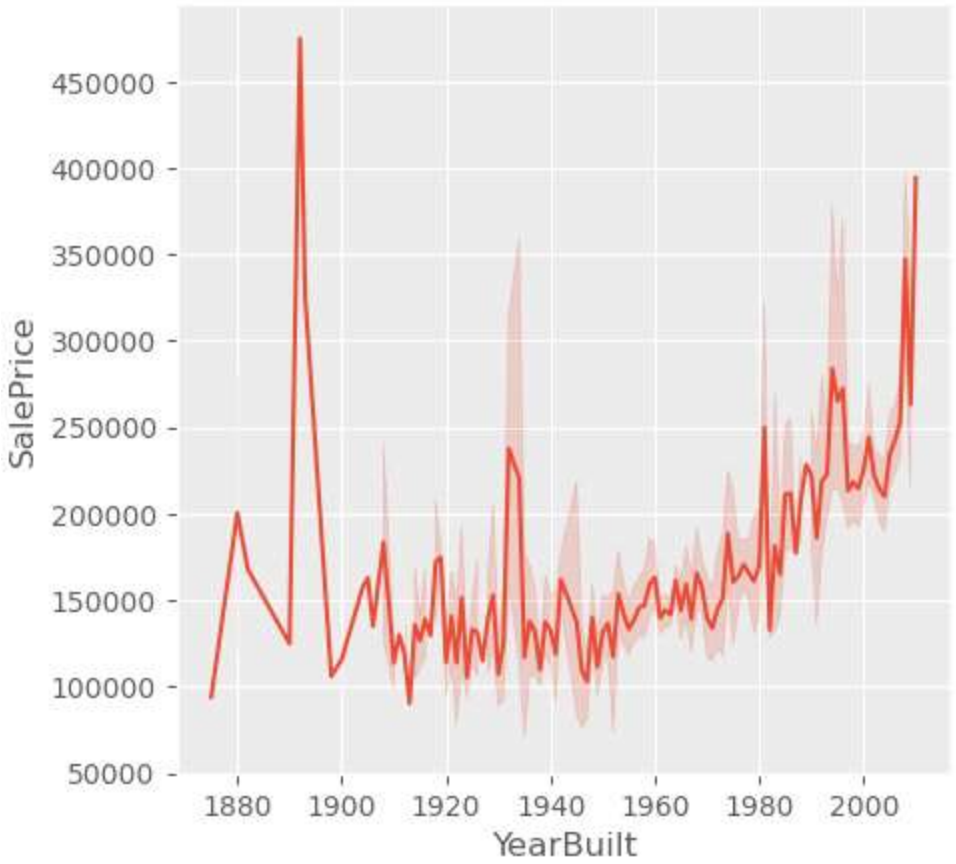
We can see that the features OverallQual, GarageArea and YearBuilt are closely correlated with the sale price.

That means these features play an important role in determining the SalePrice of a house. Similarly a lot of

inferences can be made just by looking at the heatmap but we're not going to go through each one of them.

If you don't know what a heatmap is and wish to learn.

LINEPLOT



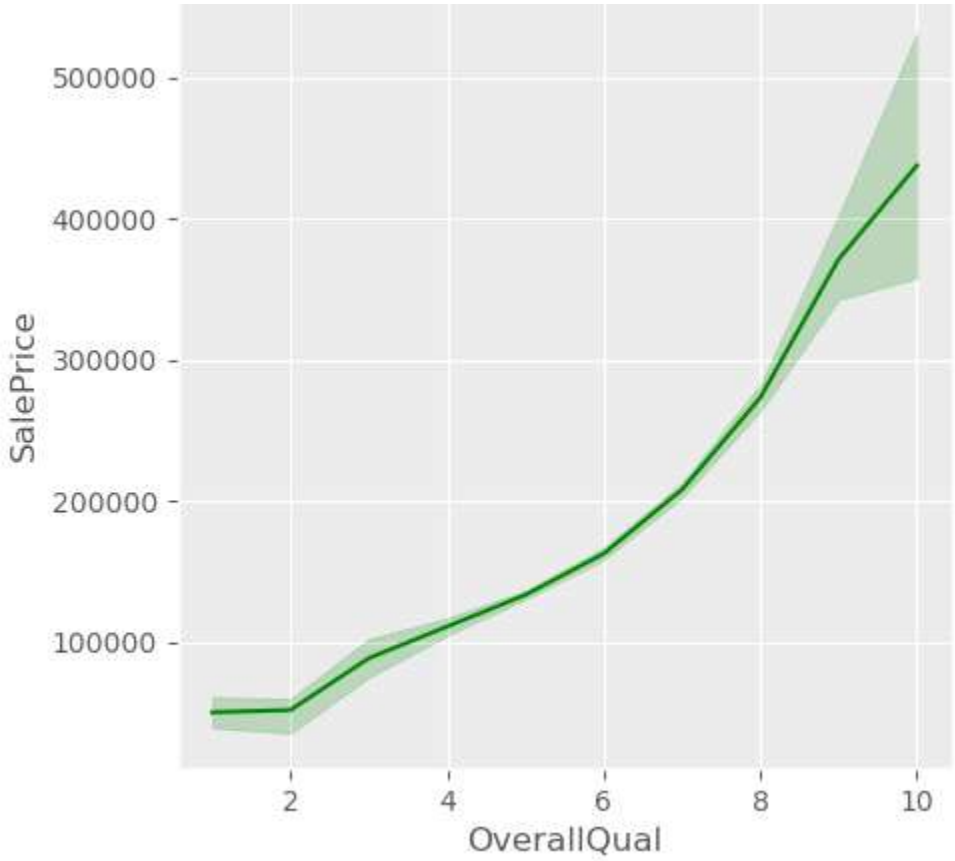
CONCLUSION:

You might have noticed a significant increase in SalePrices just after the start of the 21st century,

which is pretty interesting. What's even more surprising is the late 1800s saw phenomenal increase in

SalePrices but dropping way below even before the end of that century.

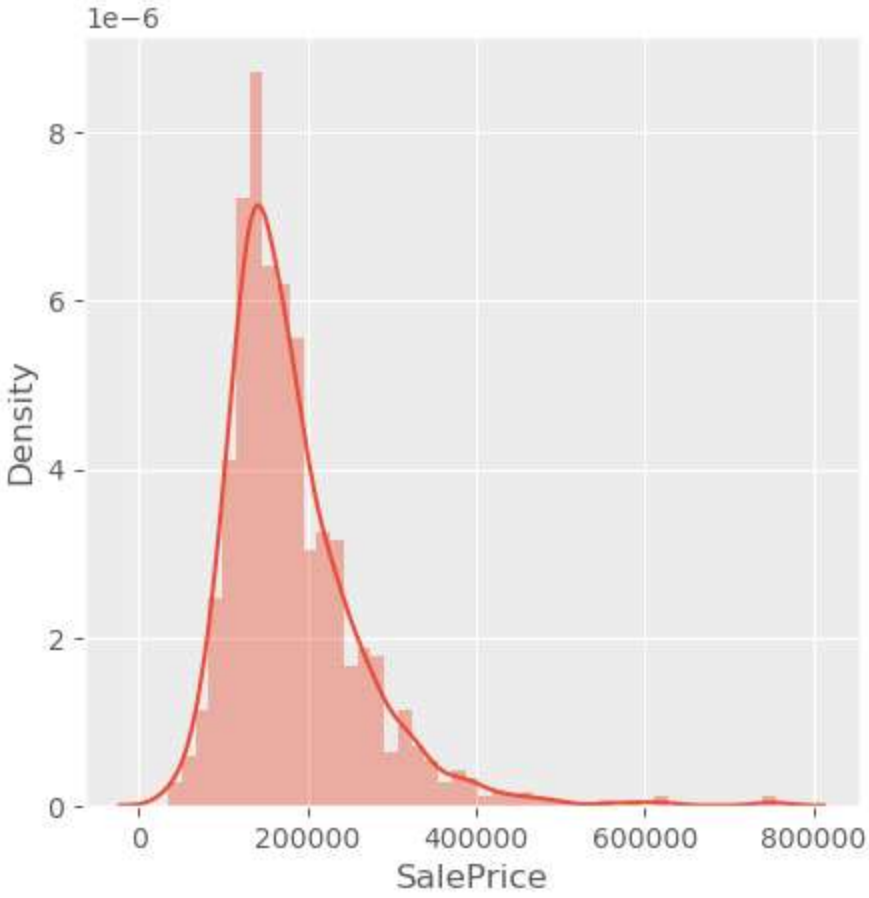
Lineplot:



CONCLUSION:

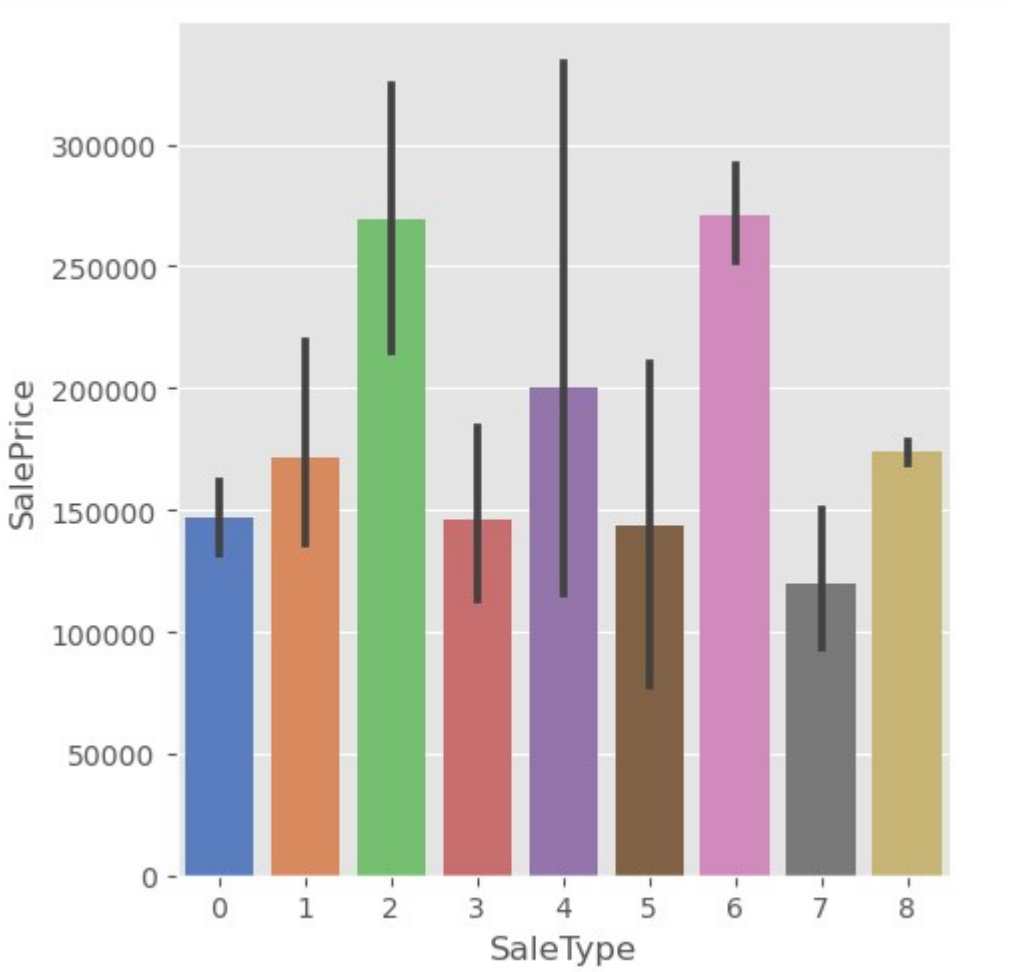
We can see that the SalePrices increase rapidly with houses with better overall quality which is pretty reasonable.

Distplot:



CONCLUSION:

It's clear that most of the houses, about 80% are sold for a price in the region on 100,000 and 200,000 dollars. Looks like there are a lot of houses sold in the mid 100,000s which makes up most of the data.



CONCLUSION:

It is clear that the selling prices for Sale types 2 and 6 are significantly higher than the others.

.....Housing project.....