### Both Scatter and Boxplots

#### BedroomAbGr (Bedrooms above grade - does NOT include basement bedrooms)

Homes with 3 and 4 bedrooms add the most value and at 5 it starts diminishing. There are a lot of outliers with 2, 3, and 4 bedrooms but it is above the upper whisker which means that the outliers indicate a higher SalePrice.

#### BsmtFullBath (Basement full bathrooms)

Homes with 1 and 2 basement full bathrooms add value and at 3 it starts to diminish. There are a lot of outliers with 0 and 1 basement full bathrooms but it is above the upper whisker which means that the outliers indicate a higher SalePrice and the fact that the outliers are also present on 0 basement full baths indicates that bathrooms in the basement don’t add a lot to the SalePrice

#### BsmtHalfBath (Basement half bathrooms)

Homes with a basement half bathrooms do not appear to be a good predictor of SalePrice

#### Fireplaces (Number of fireplaces)

Homes increase with the more fireplaces they have. There are a lot of outliers with 0, 1, and 2 fireplaces but it is above the upper whisker which means that the outliers indicate a higher SalePrice.

#### FullBath (Full bathrooms above grade)

Homes increase with the more full baths they have. There are a lot of outliers with 1 and 2 full baths but it is above the upper whisker which means that the outliers indicate a higher SalePrice.

#### GarageCars (Size of garage in car capacity)

Homes increase with the more cars that can fit in the garage. There are a lot of outliers with 1, 2, and 3 cars but it is above the upper whisker which means that the outliers indicate a higher SalePrice.

#### HalfBath (Half baths above grade)

Homes increase with the with 1 half bath as opposed to 0, but there is a diminishing return when there are 2 half baths . There are a lot of outliers with 0 and 1 half baths but it is above the upper whisker which means that the outliers indicate a higher SalePrice.

#### KitchenAbvGr (Kitchens above grade)

Homes increase when there is at least 1 kitchen and there is not a lot of value added when additional kitchens exist within a house. If there are no kitchens the SalePrice is very low. There are a lot of outliers with homes with 1 kitchen which indicates that linking this to an additional attribute such as KitchenQual is needed to quantify the impact on the SalePrice and the number of kitchens is a poor predictor of SalePrice.

#### OverallCond (Rates the overall condition of the house)

The overall condition is a good predictor of SalePrice from the ratio of 1 to 5 but then it starts diminishing and leveling off from 6 to 8. When the overall condition is 9 it adds a lot of value to the SalePrice. Since there are a lot of outliers, specially at the OverallCond of 5, it indicates this need to be linked to additional features or we can only consider OveralllConds from 1 to 5 as good predictors

#### OverallQual (Rates the overall material and finish of the house)

The overall quality is the strongest predictor for the SalePrice of a home and as the ratio increases from 1 to 10, we can see the SalePrice increase positively. This will be the main feature we want to consider in our model

#### TotalRmsAbvGrd

The total rooms above grade do appear to add value as it increases but the whiskers get larget the higher it gets and there are a lot of outliers at 4, 5, 6, 7, and 8. This indicates it is not a good predictor for SalePrice

Features that appear to be good predictors of SalePrice:

- BedroomAbvGr - no

- BsmtFullBath - no

- Fireplaces

- Full Bath

- GarageCars

- OverallQual

- TotRmsAbvGr

- 1stFlrSF (First Floor square feet)

- GrLivArea (Above grade (ground) living area square feet)

- TotalBsmtSF (Total square feet of basement area)

- BsmtQual

- CentralAir

- Electrical

- ExterQual

- FirepaceQu

- GarageFinish

- HeatingQC

- KitchenQual

- PoolQC

### Scatter Plots

The following features have a positive correlation which indicates they are good predictors for determining the SalePrice

- 1stFlrSF (First Floor square feet)

- 2ndFlrSF (Second floor square feet) – some homes have no second floor so we may have to deal with this ater

- GarageArea (Size of garage in square feet)

- GarageYrBlt (Year garage was built)

- GrLivArea (Above grade (ground) living area square feet)

- LotArea (Lot size in square feet) – need to remove outliers to get a better look

- TotalBsmtSF (Total square feet of basement area)

The following features have no correlation and do not appear to be good predictors of SalePrice

- 3SsnPorch (Three season porch area in square feet)

- BsmtFinSF1 (Type 1 finished square feet)

- BsmtFinSF2 (Type 2 finished square feet)

- BsmtUnfSF (Unfinished square feet of basement area)

- EnclosedPorch (Enclosed porch area in square feet)

- LotFrontage (Linear feet of street connected to property)

- LowQualFinSF (Low quality finished square feet (all floors))

- MasVnrArea (Masonry veneer area in square feet)

- MiscVal ($Value of miscellaneous feature)

- OpenPorchSF (Open porch area in square feet)

- PoolArea (Pool area in square feet)

- SaleDate (combined SaleMo and SaleYr)

- ScreenPorch (Screen porch area in square feet)

- WoodDeckSF (Wood deck area in square feet)

- YearBuilt (Original construction date)

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

### Boxplots

The following features appear to be good predictors of SalePrice and should be considered to add a numeric value to each categorical value as a method to quantify SalePrice:

#### BsmtQual: Evaluates the height of the basement

- NA (No basement) = 0

- Po (Poor) = 1

- Fa (Fair) = 2

- TA (Typical) = 3

- Gd (Good) = 4

- Ex (Excellent) = 5

#### CentralAir: Central air conditioning

- N (No) = 1

- Y (Yes) = 2

#### Electrical: Electrical system

- NA (Missing) = 0

- Mix (Mixed) = 1

- FuseP (60 AMP Fuse Box and mostly knob & tube wiring (poor)) = 2

- FuseF (60 AMP Fuse Box and mostly Romex wiring (Fair)) = 3

- FuseA (Fuse Box over 60 AMP and all Romex wiring (Average)) = 4

- SBrkr (Standard Circuit Breakers & Romex) = 5

#### ExterQual: Evaluates the quality of the material on the exterior

- Po (Poor) = 1

- Fa (Fair) = 2

- TA (Typical) = 3

- Gd (Good) = 4

- Ex (Excellent) = 5

#### FireplaceQu: Fireplace quality

- Po (Poor) = 1

- NA (No Fireplace) = 2

- Fa (Fair) = 2

- Ta (Typical) = 3

- Gd (Good) = 4

- Ex (Excellent) = 5

#### GarageFinish: Interior finish of the garage

- NA (No Garage) = 1

- Unf (Unfinished) = 2

- RFn (Rough Finished) = 3

- Fin (Finished) = 4

#### HeatingQC: Heating quality and condition

- Po (Poor) = 1

- Fa (Fair) = 2

- TA (Typical) = 3

- Gd (Good) = 4

- Ex (Excellent) = 5

#### KitchenQual: Kitchen quality

- Po (Poor) = 1

- Fa (Fair) = 2

- TA (Typical) = 3

- Gd (Good) = 4

- Ex (Excellent) = 5

#### PoolQC: Pool quality

- Fa (Fair) = 0

- TA (Typical) = 0

- Gd (Good) = 0

- NA (No Pool) = 0

- Ex (Excellent) = 1

The following features do not appear to be good predictors of SalePrice:

- Alley: to many outliers on NA

- BsmtCondition

- BsmtExposure

- BsmtFinType1: to many outliers

- BsmtFinType2: to many outliers

- Condition1: no real value seen in plot

- Condition2: no real value seen in plot

- ExterCond

- Fence

- Foundation

- Functional

- GarageCond

- GarageQual

- GarageType

- Heating (HeatingQC is better)

- HouseStyle

- LandContour

- LandSlope

- LotConfig

- LotShape

- MSSubClass

- MSZoning

- MasVnrType

- MiscFeature

- PavedDrive

- PoolQC

- RoofMat

- RoofStyle

- SaleCondition

- SaleType

- Street

- Utilities

#### Heating

- Floor (Floor Furnace) = 1

- Wall (Wall furnace) = 1

- OthW (Hot water or steam heat other than gas) = 1

- Grav (Gravity furnace) = 2

- GasW (Gas hot water or steam heat) = 3

- GasA (Gas forced warm air furnace) = 4

#### BsmtExposure: Refers to walkout or garden level walls

- NA (No basement) = 0

- No (No Exposure) = 1

- Mn (Minimum Exposure) = 2

- Av (Average Exposure) = 3

- Gd (Good Exposure) = 4

\*\* Note there are some

#### BsmtCondition: Evaluates the general condition of the basement

- Po (Poor) = 1

- NA = 2

- Fa (Fair) = 3

- TA (Typical) = 4

- Gd (Good) = 5

- Ex (Excellet) = 6

\*\* Note there are some outliers we may have to consider on TA (Typical)