Report on the House Price Data Set

Dimensions of the data set

In total, there are ***1460** observations* and ***81** variables.*

Occurrence of missing values

There are missing values in the following variables:

-	LotFrontage:	259	-	BsmtFinType1:	37	-	GarageQual:	81
-	Alley:	1369	-	BsmtFinType2:	38	-	GarageCond:	81
-	MasVnrType:	872	-	Electrical:	1	-	PoolQC:	1453
-	MasVnrArea:	8	-	FireplaceQu:	690	-	Fence:	1179
-	BsmtQual:	37	-	GarageType:	81	-	MiscFeature:	1406
-	BsmtCond:	37	-	GarageYrBlt:	81			
-	BsmtExposure:	38	-	GarageFinish:	81			

Distribution patterns

Numerical variables:
- **LotArea:** Right-skewed
- **TotalBsmtSF**: Right-skewed
- **1stFlrSF**: Right-skewed
- **2ndFlrSF**: Right-skewed
- **GrLivArea**: More or less normally distributed
- **GarageArea**: Normally distributed
- **WoodDeckSF**: Right skewed
- **OpenPorchSF:** Right skewed
- **EnclosedPorch**: Right skewed
- **PoolArea**: Right skewed

Categorical variables:
- **FullBath**: Most houses have *1-2* full bathrooms
- **HalfBath**: Most houses have *0-1* half bathrooms
- **LandSlope**: Most houses are built on *gentle* slopes
- **BldgType**: Most houses are *single-family detached*
- **PavedDrive**: Most houses have *paved* driveways
- **GarageType**: Most houses have *attached/detached* garages
- **Fireplaces**: Most houses have *0-1* fireplaces
- **GarageCars**: Most garages have capacity for *1-2* cars
- **Utilities**: Almost all houses run on all *public* utilities
- **Foundation**: Most houses have *poured concrete/cinder* block foundation

Bivariate analysis

- SalePrice Vs (i) TotalBsmtSF (ii) GrLivArea
    - TotalBsmtSF and GrLivArea have a **strong positive** relationship with Sale Price.

- SalePrice Vs (i) GarageArea (ii) OpenPorchSF (iii) WoodDeckSF
    - GarageArea, OpenPorchSF and WoodDeckSF have a **weak positive** relationship with Sale Price.

- SalePrice Vs FullBath
    - Generally, higher the no. of full baths in a house, higher the sale price

- SalePrice Vs Foundation
    - Houses with **poured concrete** foundations are **more expensive** than the rest.

- SalePrice Vs GarageType
    - Houses with **built-in/attached** garages are **more expensive** than the rest.

- SalePrice Vs Fireplaces
    - Generally, higher the no. of fireplaces in a house, higher the sale price.

- SalePrice Vs PavedDrive
    - Houses that have **paved driveways** are **more expensive** than the rest.

Figures

Lot Area Histogram

TotalBsmtSF Histogram

1stFlrSF Histogram

2ndFlrSF Histogram

GrLivArea Histogram

GarageArea Histogram

Utilities

Foundation

SalePrice vs TotalBsmtSF

SalePrice vs GarageArea

SalePrice vs GrLivArea

SalePrice vs OpenPorchSF