

Quiz Notebook-Finished

January 18, 2022

0.1 Data Analysis Exercise

The objective of this notebook is to gain an understanding of what your skill set looks like. You are free to use whichever libraries you need for this exercise. Please be sure to make comments where necessary and use the quickest and cleanest methods you know.

0.1.1 Task:

You will have to download the train.csv dataset onto your local machine from the following link <https://www.kaggle.com/c/house-prices-advanced-regression-techniques/data> and make an account on kaggle if you do not already. You will be using this dataset to answer all of the following questions. Make sure your graphs are clean and readable with titles.

```
[1]: import pandas as pd
      boston = pd.read_csv('train.csv')
      import numpy as np
      import matplotlib.pyplot as plt
      %matplotlib inline
      import seaborn as sns
      from scipy.stats import norm
      import statistics
```

0.1.2 Familiarize yourself with the data

- how many rows
- how many columns

```
[2]: pd.set_option('display.max_columns', None)
      boston
```

```
[2]:
```

	Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	Alley	LotShape	\
0	1	60	RL	65.0	8450	Pave	NaN	Reg	
1	2	20	RL	80.0	9600	Pave	NaN	Reg	
2	3	60	RL	68.0	11250	Pave	NaN	IR1	
3	4	70	RL	60.0	9550	Pave	NaN	IR1	
4	5	60	RL	84.0	14260	Pave	NaN	IR1	
...	
1455	1456	60	RL	62.0	7917	Pave	NaN	Reg	
1456	1457	20	RL	85.0	13175	Pave	NaN	Reg	

1457	1458	70	RL	66.0	9042	Pave	NaN	Reg
1458	1459	20	RL	68.0	9717	Pave	NaN	Reg
1459	1460	20	RL	75.0	9937	Pave	NaN	Reg

	LandContour	Utilities	LotConfig	LandSlope	Neighborhood	Condition1	\
0	Lvl	AllPub	Inside	Gtl	CollgCr	Norm	
1	Lvl	AllPub	FR2	Gtl	Veenker	Feedr	
2	Lvl	AllPub	Inside	Gtl	CollgCr	Norm	
3	Lvl	AllPub	Corner	Gtl	Crawfor	Norm	
4	Lvl	AllPub	FR2	Gtl	NoRidge	Norm	
...	
1455	Lvl	AllPub	Inside	Gtl	Gilbert	Norm	
1456	Lvl	AllPub	Inside	Gtl	NWAmes	Norm	
1457	Lvl	AllPub	Inside	Gtl	Crawfor	Norm	
1458	Lvl	AllPub	Inside	Gtl	NAmes	Norm	
1459	Lvl	AllPub	Inside	Gtl	Edwards	Norm	

	Condition2	BldgType	HouseStyle	OverallQual	OverallCond	YearBuilt	\
0	Norm	1Fam	2Story	7	5	2003	
1	Norm	1Fam	1Story	6	8	1976	
2	Norm	1Fam	2Story	7	5	2001	
3	Norm	1Fam	2Story	7	5	1915	
4	Norm	1Fam	2Story	8	5	2000	
...	
1455	Norm	1Fam	2Story	6	5	1999	
1456	Norm	1Fam	1Story	6	6	1978	
1457	Norm	1Fam	2Story	7	9	1941	
1458	Norm	1Fam	1Story	5	6	1950	
1459	Norm	1Fam	1Story	5	6	1965	

	YearRemodAdd	RoofStyle	RoofMatl	Exterior1st	Exterior2nd	MasVnrType	\
0	2003	Gable	CompShg	VinylSd	VinylSd	BrkFace	
1	1976	Gable	CompShg	MetalSd	MetalSd	None	
2	2002	Gable	CompShg	VinylSd	VinylSd	BrkFace	
3	1970	Gable	CompShg	Wd Sdng	Wd Shng	None	
4	2000	Gable	CompShg	VinylSd	VinylSd	BrkFace	
...	
1455	2000	Gable	CompShg	VinylSd	VinylSd	None	
1456	1988	Gable	CompShg	Plywood	Plywood	Stone	
1457	2006	Gable	CompShg	CemntBd	CmentBd	None	
1458	1996	Hip	CompShg	MetalSd	MetalSd	None	
1459	1965	Gable	CompShg	HdBoard	HdBoard	None	

	MasVnrArea	ExterQual	ExterCond	Foundation	BsmtQual	BsmtCond	\
0	196.0	Gd	TA	PConc	Gd	TA	
1	0.0	TA	TA	CBlock	Gd	TA	
2	162.0	Gd	TA	PConc	Gd	TA	

3	0.0	TA	TA	BrkTil	TA	Gd
4	350.0	Gd	TA	PConc	Gd	TA
...
1455	0.0	TA	TA	PConc	Gd	TA
1456	119.0	TA	TA	CBlock	Gd	TA
1457	0.0	Ex	Gd	Stone	TA	Gd
1458	0.0	TA	TA	CBlock	TA	TA
1459	0.0	Gd	TA	CBlock	TA	TA

	BsmtExposure	BsmtFinType1	BsmtFinSF1	BsmtFinType2	BsmtFinSF2	\
0	No	GLQ	706	Unf	0	
1	Gd	ALQ	978	Unf	0	
2	Mn	GLQ	486	Unf	0	
3	No	ALQ	216	Unf	0	
4	Av	GLQ	655	Unf	0	
...
1455	No	Unf	0	Unf	0	
1456	No	ALQ	790	Rec	163	
1457	No	GLQ	275	Unf	0	
1458	Mn	GLQ	49	Rec	1029	
1459	No	BLQ	830	LwQ	290	

	BsmtUnfSF	TotalBsmtSF	Heating	HeatingQC	CentralAir	Electrical	\
0	150	856	GasA	Ex	Y	SBrkr	
1	284	1262	GasA	Ex	Y	SBrkr	
2	434	920	GasA	Ex	Y	SBrkr	
3	540	756	GasA	Gd	Y	SBrkr	
4	490	1145	GasA	Ex	Y	SBrkr	
...
1455	953	953	GasA	Ex	Y	SBrkr	
1456	589	1542	GasA	TA	Y	SBrkr	
1457	877	1152	GasA	Ex	Y	SBrkr	
1458	0	1078	GasA	Gd	Y	FuseA	
1459	136	1256	GasA	Gd	Y	SBrkr	

	1stFlrSF	2ndFlrSF	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath	\
0	856	854	0	1710	1	0	
1	1262	0	0	1262	0	1	
2	920	866	0	1786	1	0	
3	961	756	0	1717	1	0	
4	1145	1053	0	2198	1	0	
...
1455	953	694	0	1647	0	0	
1456	2073	0	0	2073	1	0	
1457	1188	1152	0	2340	0	0	
1458	1078	0	0	1078	1	0	
1459	1256	0	0	1256	1	0	

	FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr	KitchenQual	\
0	2	1	3	1	Gd	
1	2	0	3	1	TA	
2	2	1	3	1	Gd	
3	1	0	3	1	Gd	
4	2	1	4	1	Gd	
...	
1455	2	1	3	1	TA	
1456	2	0	3	1	TA	
1457	2	0	4	1	Gd	
1458	1	0	2	1	Gd	
1459	1	1	3	1	TA	

	TotRmsAbvGrd	Functional	Fireplaces	FireplaceQu	GarageType	GarageYrBlt	\
0		8	Typ	0	NaN	Attchd	2003.0
1		6	Typ	1	TA	Attchd	1976.0
2		6	Typ	1	TA	Attchd	2001.0
3		7	Typ	1	Gd	Detchd	1998.0
4		9	Typ	1	TA	Attchd	2000.0
...	
1455		7	Typ	1	TA	Attchd	1999.0
1456		7	Min1	2	TA	Attchd	1978.0
1457		9	Typ	2	Gd	Attchd	1941.0
1458		5	Typ	0	NaN	Attchd	1950.0
1459		6	Typ	0	NaN	Attchd	1965.0

	GarageFinish	GarageCars	GarageArea	GarageQual	GarageCond	PavedDrive	\
0	RFn	2	548	TA	TA	Y	
1	RFn	2	460	TA	TA	Y	
2	RFn	2	608	TA	TA	Y	
3	Unf	3	642	TA	TA	Y	
4	RFn	3	836	TA	TA	Y	
...	
1455	RFn	2	460	TA	TA	Y	
1456	Unf	2	500	TA	TA	Y	
1457	RFn	1	252	TA	TA	Y	
1458	Unf	1	240	TA	TA	Y	
1459	Fin	1	276	TA	TA	Y	

	WoodDeckSF	OpenPorchSF	EnclosedPorch	3SsnPorch	ScreenPorch	\
0	0	61	0	0	0	
1	298	0	0	0	0	
2	0	42	0	0	0	
3	0	35	272	0	0	
4	192	84	0	0	0	
...	

1455	0	40	0	0	0
1456	349	0	0	0	0
1457	0	60	0	0	0
1458	366	0	112	0	0
1459	736	68	0	0	0

	PoolArea	PoolQC	Fence	MiscFeature	MiscVal	MoSold	YrSold	SaleType	\
0	0	NaN	NaN	NaN	0	2	2008	WD	
1	0	NaN	NaN	NaN	0	5	2007	WD	
2	0	NaN	NaN	NaN	0	9	2008	WD	
3	0	NaN	NaN	NaN	0	2	2006	WD	
4	0	NaN	NaN	NaN	0	12	2008	WD	
...	
1455	0	NaN	NaN	NaN	0	8	2007	WD	
1456	0	NaN	MnPrv	NaN	0	2	2010	WD	
1457	0	NaN	GdPrv	Shed	2500	5	2010	WD	
1458	0	NaN	NaN	NaN	0	4	2010	WD	
1459	0	NaN	NaN	NaN	0	6	2008	WD	

	SaleCondition	SalePrice
0	Normal	208500
1	Normal	181500
2	Normal	223500
3	Abnorml	140000
4	Normal	250000
...
1455	Normal	175000
1456	Normal	210000
1457	Normal	266500
1458	Normal	142125
1459	Normal	147500

[1460 rows x 81 columns]

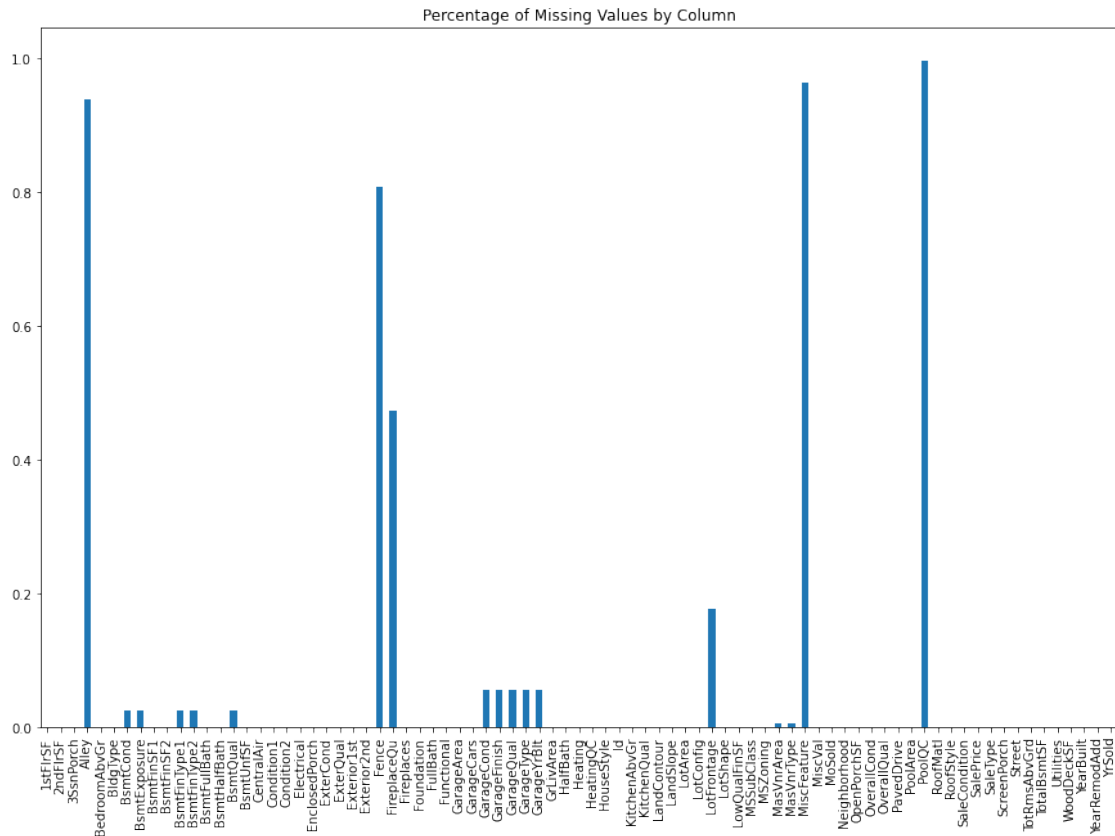
```
[3]: # The data set contains 1460 rows and 81 columns.
```

0.1.3 Inspect missing values and deal with missing values

- Graph the percentage of missing values
- Come up with 2 reasons why data might be missing for MiscFeature: Miscellaneous feature
- For every field with missing values, create a duplicate boolean field and graph the mean and
- Drop columns with greater than 30% of missing values

```
[4]: boston_missing_percentage = boston.isnull().sum()/boston.isnull().count().
      ↪sort_values(ascending = False)
      boston_missing_percentage.plot(kind = 'bar', figsize = (15,10), title = '
      ↪'Percentage of Missing Values by Column')
```

```
[4]: <AxesSubplot:title={'center':'Percentage of Missing Values by Column'}>
```



```
[5]: # There could simply be no homes with any Miscellaneous Features.
# Miscellaneous Features may be a combination of unique values in which no
→homes fall under.
```

```
[6]: boston['Alley_2'] = boston['Alley'].isnull()
boston['BsmtCond_2'] = boston['BsmtCond'].isnull()
boston['BsmtExposure_2'] = boston['BsmtExposure'].isnull()
boston['BsmtFinType1_2'] = boston['BsmtFinType1'].isnull()
boston['BsmtFinType2_2'] = boston['BsmtFinType2'].isnull()
boston['BsmtQual_2'] = boston['BsmtQual'].isnull()
boston['FireplaceQU_2'] = boston['FireplaceQu'].isnull()
boston['Fence_2'] = boston['Fence'].isnull()
boston['GarageCond_2'] = boston['GarageCond'].isnull()
boston['GarageFinish_2'] = boston['GarageFinish'].isnull()
boston['GarageQual_2'] = boston['GarageQual'].isnull()
boston['GarageType_2'] = boston['GarageType'].isnull()
boston['GargeYrBlt_2'] = boston['GarageYrBlt'].isnull()
boston['LotFrontage_2'] = boston['LotFrontage'].isnull()
```

```

boston['MasVnrArea_2'] = boston['MasVnrArea'].isnull()
boston['MasVnrType_2'] = boston['MasVnrType'].isnull()
boston['MiscFeature_2'] = boston['MiscFeature'].isnull()
boston['PoolQC_2'] = boston['PoolQC'].isnull()
boston

```

```

[6]:
      Id  MSSubClass  MSZoning  LotFrontage  LotArea  Street  Alley  LotShape  \
0      1          60        RL          65.0     8450   Pave   NaN      Reg
1      2          20        RL          80.0     9600   Pave   NaN      Reg
2      3          60        RL          68.0    11250   Pave   NaN     IR1
3      4          70        RL          60.0     9550   Pave   NaN     IR1
4      5          60        RL          84.0    14260   Pave   NaN     IR1
...  ...
1455  1456          60        RL          62.0     7917   Pave   NaN      Reg
1456  1457          20        RL          85.0    13175   Pave   NaN      Reg
1457  1458          70        RL          66.0     9042   Pave   NaN      Reg
1458  1459          20        RL          68.0     9717   Pave   NaN      Reg
1459  1460          20        RL          75.0     9937   Pave   NaN      Reg

      LandContour  Utilities  LotConfig  LandSlope  Neighborhood  Condition1  \
0          Lvl1    AllPub    Inside    Gtl    CollgCr    Norm
1          Lvl1    AllPub    FR2    Gtl    Veenker    Feedr
2          Lvl1    AllPub    Inside    Gtl    CollgCr    Norm
3          Lvl1    AllPub    Corner    Gtl    Crawfor    Norm
4          Lvl1    AllPub    FR2    Gtl    NoRidge    Norm
...  ...
1455          Lvl1    AllPub    Inside    Gtl    Gilbert    Norm
1456          Lvl1    AllPub    Inside    Gtl    NWAmes    Norm
1457          Lvl1    AllPub    Inside    Gtl    Crawfor    Norm
1458          Lvl1    AllPub    Inside    Gtl    NAmes    Norm
1459          Lvl1    AllPub    Inside    Gtl    Edwards    Norm

      Condition2  BldgType  HouseStyle  OverallQual  OverallCond  YearBuilt  \
0          Norm    1Fam    2Story          7          5        2003
1          Norm    1Fam    1Story          6          8        1976
2          Norm    1Fam    2Story          7          5        2001
3          Norm    1Fam    2Story          7          5        1915
4          Norm    1Fam    2Story          8          5        2000
...  ...
1455          Norm    1Fam    2Story          6          5        1999
1456          Norm    1Fam    1Story          6          6        1978
1457          Norm    1Fam    2Story          7          9        1941
1458          Norm    1Fam    1Story          5          6        1950
1459          Norm    1Fam    1Story          5          6        1965

      YearRemodAdd  RoofStyle  RoofMatl  Exterior1st  Exterior2nd  MasVnrType  \
0          2003    Gable  CompShg    VinylSd    VinylSd    BrkFace

```

1	1976	Gable	CompShg	MetalSd	MetalSd	None
2	2002	Gable	CompShg	VinylSd	VinylSd	BrkFace
3	1970	Gable	CompShg	Wd Sdng	Wd Shng	None
4	2000	Gable	CompShg	VinylSd	VinylSd	BrkFace
...
1455	2000	Gable	CompShg	VinylSd	VinylSd	None
1456	1988	Gable	CompShg	Plywood	Plywood	Stone
1457	2006	Gable	CompShg	CemntBd	CmentBd	None
1458	1996	Hip	CompShg	MetalSd	MetalSd	None
1459	1965	Gable	CompShg	HdBoard	HdBoard	None

	MasVnrArea	ExterQual	ExterCond	Foundation	BsmtQual	BsmtCond	\
0	196.0	Gd	TA	PConc	Gd	TA	
1	0.0	TA	TA	CBlock	Gd	TA	
2	162.0	Gd	TA	PConc	Gd	TA	
3	0.0	TA	TA	BrkTil	TA	Gd	
4	350.0	Gd	TA	PConc	Gd	TA	
...	
1455	0.0	TA	TA	PConc	Gd	TA	
1456	119.0	TA	TA	CBlock	Gd	TA	
1457	0.0	Ex	Gd	Stone	TA	Gd	
1458	0.0	TA	TA	CBlock	TA	TA	
1459	0.0	Gd	TA	CBlock	TA	TA	

	BsmtExposure	BsmtFinType1	BsmtFinSF1	BsmtFinType2	BsmtFinSF2	\
0	No	GLQ	706	Unf	0	
1	Gd	ALQ	978	Unf	0	
2	Mn	GLQ	486	Unf	0	
3	No	ALQ	216	Unf	0	
4	Av	GLQ	655	Unf	0	
...	
1455	No	Unf	0	Unf	0	
1456	No	ALQ	790	Rec	163	
1457	No	GLQ	275	Unf	0	
1458	Mn	GLQ	49	Rec	1029	
1459	No	BLQ	830	LwQ	290	

	BsmtUnfSF	TotalBsmtSF	Heating	HeatingQC	CentralAir	Electrical	\
0	150	856	GasA	Ex	Y	SBrkr	
1	284	1262	GasA	Ex	Y	SBrkr	
2	434	920	GasA	Ex	Y	SBrkr	
3	540	756	GasA	Gd	Y	SBrkr	
4	490	1145	GasA	Ex	Y	SBrkr	
...	
1455	953	953	GasA	Ex	Y	SBrkr	
1456	589	1542	GasA	TA	Y	SBrkr	
1457	877	1152	GasA	Ex	Y	SBrkr	

1458	0	1078	GasA	Gd	Y	FuseA
1459	136	1256	GasA	Gd	Y	SBrkr

	1stFlrSF	2ndFlrSF	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath	\
0	856	854	0	1710	1	0	
1	1262	0	0	1262	0	1	
2	920	866	0	1786	1	0	
3	961	756	0	1717	1	0	
4	1145	1053	0	2198	1	0	
...	
1455	953	694	0	1647	0	0	
1456	2073	0	0	2073	1	0	
1457	1188	1152	0	2340	0	0	
1458	1078	0	0	1078	1	0	
1459	1256	0	0	1256	1	0	

	FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr	KitchenQual	\
0	2	1	3	1	Gd	
1	2	0	3	1	TA	
2	2	1	3	1	Gd	
3	1	0	3	1	Gd	
4	2	1	4	1	Gd	
...	
1455	2	1	3	1	TA	
1456	2	0	3	1	TA	
1457	2	0	4	1	Gd	
1458	1	0	2	1	Gd	
1459	1	1	3	1	TA	

	TotRmsAbvGrd	Functional	Fireplaces	FireplaceQu	GarageType	GarageYrBlt	\
0		8	Typ	0	NaN	Attchd	2003.0
1		6	Typ	1	TA	Attchd	1976.0
2		6	Typ	1	TA	Attchd	2001.0
3		7	Typ	1	Gd	Detchd	1998.0
4		9	Typ	1	TA	Attchd	2000.0
...	
1455		7	Typ	1	TA	Attchd	1999.0
1456		7	Min1	2	TA	Attchd	1978.0
1457		9	Typ	2	Gd	Attchd	1941.0
1458		5	Typ	0	NaN	Attchd	1950.0
1459		6	Typ	0	NaN	Attchd	1965.0

	GarageFinish	GarageCars	GarageArea	GarageQual	GarageCond	PavedDrive	\
0	RFn	2	548	TA	TA	Y	
1	RFn	2	460	TA	TA	Y	
2	RFn	2	608	TA	TA	Y	
3	Unf	3	642	TA	TA	Y	

4	RFn	3	836	TA	TA	Y
...
1455	RFn	2	460	TA	TA	Y
1456	Unf	2	500	TA	TA	Y
1457	RFn	1	252	TA	TA	Y
1458	Unf	1	240	TA	TA	Y
1459	Fin	1	276	TA	TA	Y

	WoodDeckSF	OpenPorchSF	EnclosedPorch	3SsnPorch	ScreenPorch	\
0	0	61	0	0	0	
1	298	0	0	0	0	
2	0	42	0	0	0	
3	0	35	272	0	0	
4	192	84	0	0	0	
...	
1455	0	40	0	0	0	
1456	349	0	0	0	0	
1457	0	60	0	0	0	
1458	366	0	112	0	0	
1459	736	68	0	0	0	

	PoolArea	PoolQC	Fence	MiscFeature	MiscVal	MoSold	YrSold	SaleType	\
0	0	NaN	NaN	NaN	0	2	2008	WD	
1	0	NaN	NaN	NaN	0	5	2007	WD	
2	0	NaN	NaN	NaN	0	9	2008	WD	
3	0	NaN	NaN	NaN	0	2	2006	WD	
4	0	NaN	NaN	NaN	0	12	2008	WD	
...	
1455	0	NaN	NaN	NaN	0	8	2007	WD	
1456	0	NaN	MnPrv	NaN	0	2	2010	WD	
1457	0	NaN	GdPrv	Shed	2500	5	2010	WD	
1458	0	NaN	NaN	NaN	0	4	2010	WD	
1459	0	NaN	NaN	NaN	0	6	2008	WD	

	SaleCondition	SalePrice	Alley_2	BsmtCond_2	BsmtExposure_2	\
0	Normal	208500	True	False	False	
1	Normal	181500	True	False	False	
2	Normal	223500	True	False	False	
3	Abnorml	140000	True	False	False	
4	Normal	250000	True	False	False	
...	
1455	Normal	175000	True	False	False	
1456	Normal	210000	True	False	False	
1457	Normal	266500	True	False	False	
1458	Normal	142125	True	False	False	
1459	Normal	147500	True	False	False	

	BsmtFinType1_2	BsmtFintype2_2	BsmtQual_2	FireplaceQU_2	Fence_2	\
0	False	False	False	True	True	
1	False	False	False	False	True	
2	False	False	False	False	True	
3	False	False	False	False	True	
4	False	False	False	False	True	
...	
1455	False	False	False	False	True	
1456	False	False	False	False	False	
1457	False	False	False	False	False	
1458	False	False	False	True	True	
1459	False	False	False	True	True	

	GarageCond_2	GarageFinish_2	GarageQual_2	GarageType_2	GargeYrBlt_2	\
0	False	False	False	False	False	
1	False	False	False	False	False	
2	False	False	False	False	False	
3	False	False	False	False	False	
4	False	False	False	False	False	
...	
1455	False	False	False	False	False	
1456	False	False	False	False	False	
1457	False	False	False	False	False	
1458	False	False	False	False	False	
1459	False	False	False	False	False	

	LotFrontage_2	MasVnrArea_2	MasVnrType_2	MiscFeature_2	PoolQC_2
0	False	False	False	True	True
1	False	False	False	True	True
2	False	False	False	True	True
3	False	False	False	True	True
4	False	False	False	True	True
...
1455	False	False	False	True	True
1456	False	False	False	True	True
1457	False	False	False	False	True
1458	False	False	False	True	True
1459	False	False	False	True	True

[1460 rows x 99 columns]

```
[7]: fig, axes = plt.subplots(3, 6, figsize=(15, 15))
fig.suptitle('Mean SalePrice of true and false values')

sns.barplot(x =boston['Alley_2'], y = boston['SalePrice'], ax=axes[0,0])
sns.barplot(x =boston['BsmtCond_2'], y = boston['SalePrice'], ax=axes[0,1])
sns.barplot(x =boston['BsmtExposure_2'], y = boston['SalePrice'], ax=axes[0,2])
```

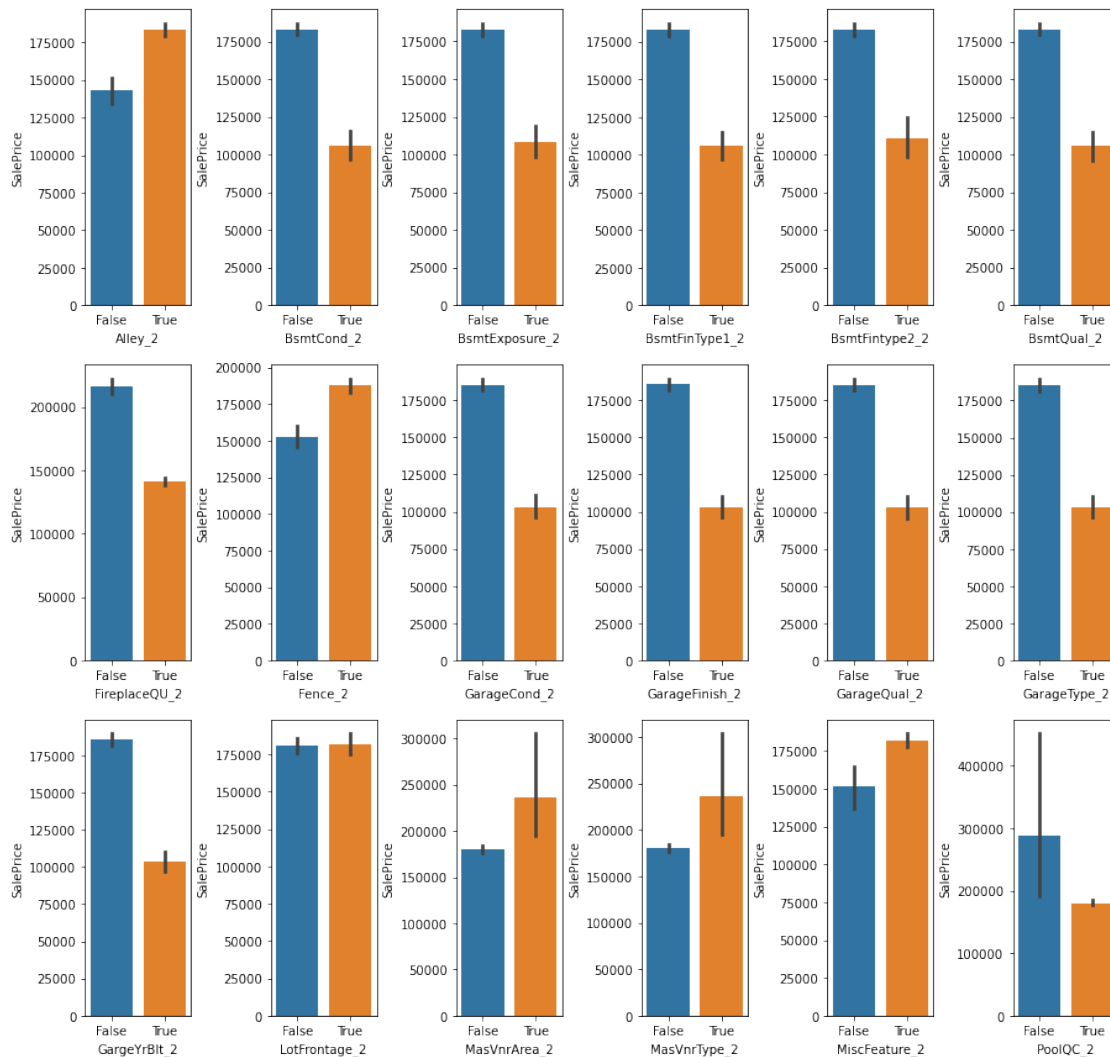
```

sns.barplot(x =boston['BsmtFinType1_2'], y = boston['SalePrice'], ax=axes[0,3])
sns.barplot(x =boston['BsmtFintype2_2'], y = boston['SalePrice'], ax=axes[0,4])
sns.barplot(x =boston['BsmtQual_2'], y = boston['SalePrice'], ax=axes[0,5])
sns.barplot(x =boston['FireplaceQU_2'], y = boston['SalePrice'], ax=axes[1,0])
sns.barplot(x =boston['Fence_2'], y = boston['SalePrice'], ax=axes[1,1])
sns.barplot(x =boston['GarageCond_2'], y = boston['SalePrice'], ax=axes[1,2])
sns.barplot(x =boston['GarageFinish_2'], y = boston['SalePrice'], ax=axes[1,3])
sns.barplot(x =boston['GarageQual_2'], y = boston['SalePrice'], ax=axes[1,4])
sns.barplot(x =boston['GarageType_2'], y = boston['SalePrice'], ax=axes[1,5])
sns.barplot(x =boston['GargeYrBlt_2'], y = boston['SalePrice'], ax=axes[2,0])
sns.barplot(x =boston['LotFrontage_2'], y = boston['SalePrice'], ax=axes[2,1])
sns.barplot(x =boston['MasVnrArea_2'], y = boston['SalePrice'], ax=axes[2,2])
sns.barplot(x =boston['MasVnrType_2'], y = boston['SalePrice'], ax=axes[2,3])
sns.barplot(x =boston['MiscFeature_2'], y = boston['SalePrice'], ax=axes[2,4])
sns.barplot(x =boston['PoolQC_2'], y = boston['SalePrice'], ax=axes[2,5])

fig.subplots_adjust(wspace=0.75)

```

Mean SalePrice of true and false values



```
[8]: fig, axes = plt.subplots(3, 6, figsize=(15, 15))
fig.suptitle('Std SalePrice of true and false values')

sns.barplot(x =boston['Alley_2'], y = boston['SalePrice'], ci='sd',
    ↳ax=axes[0,0])
sns.barplot(x =boston['BsmtCond_2'], y = boston['SalePrice'], ci='sd',
    ↳ax=axes[0,1])
sns.barplot(x =boston['BsmtExposure_2'], y = boston['SalePrice'], ci='sd',
    ↳ax=axes[0,2])
```

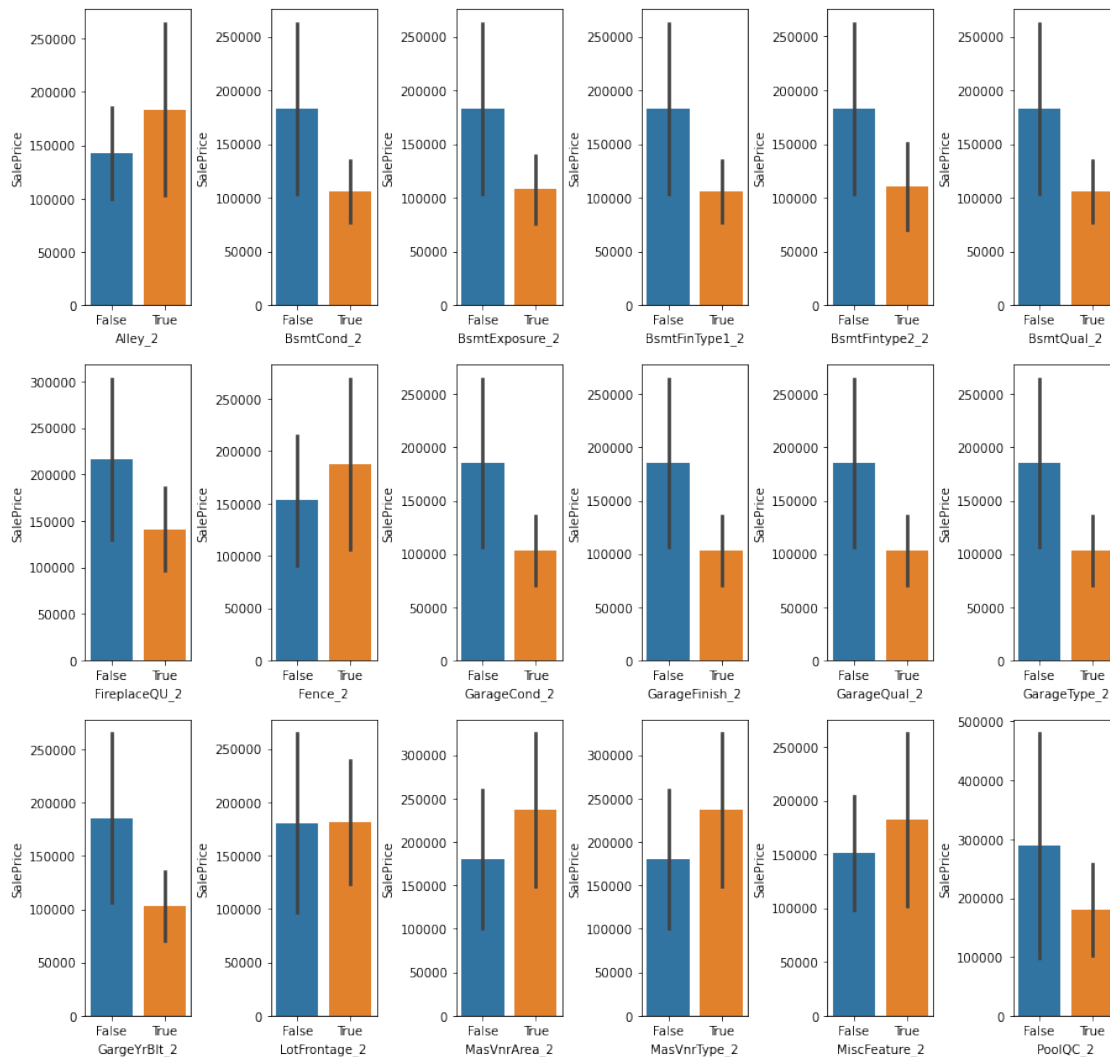
```

sns.barplot(x =boston['BsmtFinType1_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[0,3])
sns.barplot(x =boston['BsmtFintype2_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[0,4])
sns.barplot(x =boston['BsmtQual_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[0,5])
sns.barplot(x =boston['FireplaceQU_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[1,0])
sns.barplot(x =boston['Fence_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[1,1])
sns.barplot(x =boston['GarageCond_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[1,2])
sns.barplot(x =boston['GarageFinish_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[1,3])
sns.barplot(x =boston['GarageQual_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[1,4])
sns.barplot(x =boston['GarageType_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[1,5])
sns.barplot(x =boston['GargeYrBlt_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[2,0])
sns.barplot(x =boston['LotFrontage_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[2,1])
sns.barplot(x =boston['MasVnrArea_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[2,2])
sns.barplot(x =boston['MasVnrType_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[2,3])
sns.barplot(x =boston['MiscFeature_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[2,4])
sns.barplot(x =boston['PoolQC_2'], y = boston['SalePrice'], ci='sd',□
    ↳ax=axes[2,5])

fig.subplots_adjust(wspace=0.75)

```

Std SalePrice of true and false values



```
[9]: # Dropping columns with more than 30% of missing Vales
Dropped_na_boston = pd.read_csv('train.csv')
Dropped_na_boston = Dropped_na_boston.drop(columns=['Alley', 'Fence', '
↳ 'FireplaceQu', 'MiscFeature', 'PoolQC'])
Dropped_na_boston
```

[9]:

	Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	LotShape	\
0	1	60	RL	65.0	8450	Pave	Reg	
1	2	20	RL	80.0	9600	Pave	Reg	
2	3	60	RL	68.0	11250	Pave	IR1	
3	4	70	RL	60.0	9550	Pave	IR1	

4	5	60	RL	84.0	14260	Pave	IR1
...
1455	1456	60	RL	62.0	7917	Pave	Reg
1456	1457	20	RL	85.0	13175	Pave	Reg
1457	1458	70	RL	66.0	9042	Pave	Reg
1458	1459	20	RL	68.0	9717	Pave	Reg
1459	1460	20	RL	75.0	9937	Pave	Reg

	LandContour	Utilities	LotConfig	LandSlope	Neighborhood	Condition1	\
0	Lvl	AllPub	Inside	Gtl	CollgCr	Norm	
1	Lvl	AllPub	FR2	Gtl	Veenker	Feedr	
2	Lvl	AllPub	Inside	Gtl	CollgCr	Norm	
3	Lvl	AllPub	Corner	Gtl	Crawfor	Norm	
4	Lvl	AllPub	FR2	Gtl	NoRidge	Norm	
...	
1455	Lvl	AllPub	Inside	Gtl	Gilbert	Norm	
1456	Lvl	AllPub	Inside	Gtl	NWAmes	Norm	
1457	Lvl	AllPub	Inside	Gtl	Crawfor	Norm	
1458	Lvl	AllPub	Inside	Gtl	NAmes	Norm	
1459	Lvl	AllPub	Inside	Gtl	Edwards	Norm	

	Condition2	BldgType	HouseStyle	OverallQual	OverallCond	YearBuilt	\
0	Norm	1Fam	2Story	7	5	2003	
1	Norm	1Fam	1Story	6	8	1976	
2	Norm	1Fam	2Story	7	5	2001	
3	Norm	1Fam	2Story	7	5	1915	
4	Norm	1Fam	2Story	8	5	2000	
...	
1455	Norm	1Fam	2Story	6	5	1999	
1456	Norm	1Fam	1Story	6	6	1978	
1457	Norm	1Fam	2Story	7	9	1941	
1458	Norm	1Fam	1Story	5	6	1950	
1459	Norm	1Fam	1Story	5	6	1965	

	YearRemodAdd	RoofStyle	RoofMatl	Exterior1st	Exterior2nd	MasVnrType	\
0	2003	Gable	CompShg	VinylSd	VinylSd	BrkFace	
1	1976	Gable	CompShg	MetalSd	MetalSd	None	
2	2002	Gable	CompShg	VinylSd	VinylSd	BrkFace	
3	1970	Gable	CompShg	Wd Sdng	Wd Shng	None	
4	2000	Gable	CompShg	VinylSd	VinylSd	BrkFace	
...	
1455	2000	Gable	CompShg	VinylSd	VinylSd	None	
1456	1988	Gable	CompShg	Plywood	Plywood	Stone	
1457	2006	Gable	CompShg	CemntBd	CmentBd	None	
1458	1996	Hip	CompShg	MetalSd	MetalSd	None	
1459	1965	Gable	CompShg	HdBoard	HdBoard	None	

	MasVnrArea	ExterQual	ExterCond	Foundation	BsmtQual	BsmtCond	\
0	196.0	Gd	TA	PConc	Gd	TA	
1	0.0	TA	TA	CBlock	Gd	TA	
2	162.0	Gd	TA	PConc	Gd	TA	
3	0.0	TA	TA	BrkTil	TA	Gd	
4	350.0	Gd	TA	PConc	Gd	TA	
...	
1455	0.0	TA	TA	PConc	Gd	TA	
1456	119.0	TA	TA	CBlock	Gd	TA	
1457	0.0	Ex	Gd	Stone	TA	Gd	
1458	0.0	TA	TA	CBlock	TA	TA	
1459	0.0	Gd	TA	CBlock	TA	TA	

	BsmtExposure	BsmtFinType1	BsmtFinSF1	BsmtFinType2	BsmtFinSF2	\
0	No	GLQ	706	Unf	0	
1	Gd	ALQ	978	Unf	0	
2	Mn	GLQ	486	Unf	0	
3	No	ALQ	216	Unf	0	
4	Av	GLQ	655	Unf	0	
...	
1455	No	Unf	0	Unf	0	
1456	No	ALQ	790	Rec	163	
1457	No	GLQ	275	Unf	0	
1458	Mn	GLQ	49	Rec	1029	
1459	No	BLQ	830	LwQ	290	

	BsmtUnfSF	TotalBsmtSF	Heating	HeatingQC	CentralAir	Electrical	\
0	150	856	GasA	Ex	Y	SBrkr	
1	284	1262	GasA	Ex	Y	SBrkr	
2	434	920	GasA	Ex	Y	SBrkr	
3	540	756	GasA	Gd	Y	SBrkr	
4	490	1145	GasA	Ex	Y	SBrkr	
...	
1455	953	953	GasA	Ex	Y	SBrkr	
1456	589	1542	GasA	TA	Y	SBrkr	
1457	877	1152	GasA	Ex	Y	SBrkr	
1458	0	1078	GasA	Gd	Y	FuseA	
1459	136	1256	GasA	Gd	Y	SBrkr	

	1stFlrSF	2ndFlrSF	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath	\
0	856	854	0	1710	1	0	
1	1262	0	0	1262	0	1	
2	920	866	0	1786	1	0	
3	961	756	0	1717	1	0	
4	1145	1053	0	2198	1	0	
...	
1455	953	694	0	1647	0	0	

1456	2073	0	0	2073	1	0
1457	1188	1152	0	2340	0	0
1458	1078	0	0	1078	1	0
1459	1256	0	0	1256	1	0

	FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr	KitchenQual	\
0	2	1	3	1	Gd	
1	2	0	3	1	TA	
2	2	1	3	1	Gd	
3	1	0	3	1	Gd	
4	2	1	4	1	Gd	
...	
1455	2	1	3	1	TA	
1456	2	0	3	1	TA	
1457	2	0	4	1	Gd	
1458	1	0	2	1	Gd	
1459	1	1	3	1	TA	

	TotRmsAbvGrd	Functional	Fireplaces	GarageType	GarageYrBlt	\
0	8	Typ	0	Attchd	2003.0	
1	6	Typ	1	Attchd	1976.0	
2	6	Typ	1	Attchd	2001.0	
3	7	Typ	1	Detchd	1998.0	
4	9	Typ	1	Attchd	2000.0	
...	
1455	7	Typ	1	Attchd	1999.0	
1456	7	Min1	2	Attchd	1978.0	
1457	9	Typ	2	Attchd	1941.0	
1458	5	Typ	0	Attchd	1950.0	
1459	6	Typ	0	Attchd	1965.0	

	GarageFinish	GarageCars	GarageArea	GarageQual	GarageCond	PavedDrive	\
0	RFn	2	548	TA	TA	Y	
1	RFn	2	460	TA	TA	Y	
2	RFn	2	608	TA	TA	Y	
3	Unf	3	642	TA	TA	Y	
4	RFn	3	836	TA	TA	Y	
...	
1455	RFn	2	460	TA	TA	Y	
1456	Unf	2	500	TA	TA	Y	
1457	RFn	1	252	TA	TA	Y	
1458	Unf	1	240	TA	TA	Y	
1459	Fin	1	276	TA	TA	Y	

	WoodDeckSF	OpenPorchSF	EnclosedPorch	3SsnPorch	ScreenPorch	\
0	0	61	0	0	0	
1	298	0	0	0	0	

2	0	42	0	0	0
3	0	35	272	0	0
4	192	84	0	0	0
...
1455	0	40	0	0	0
1456	349	0	0	0	0
1457	0	60	0	0	0
1458	366	0	112	0	0
1459	736	68	0	0	0

	PoolArea	MiscVal	MoSold	YrSold	SaleType	SaleCondition	SalePrice
0	0	0	2	2008	WD	Normal	208500
1	0	0	5	2007	WD	Normal	181500
2	0	0	9	2008	WD	Normal	223500
3	0	0	2	2006	WD	Abnorml	140000
4	0	0	12	2008	WD	Normal	250000
...
1455	0	0	8	2007	WD	Normal	175000
1456	0	0	2	2010	WD	Normal	210000
1457	0	2500	5	2010	WD	Normal	266500
1458	0	0	4	2010	WD	Normal	142125
1459	0	0	6	2008	WD	Normal	147500

[1460 rows x 76 columns]

Work with Numerical variables

Discrete

- Determine how many discrete variables and how many temporal variables are in the dataset
- Graph the median SalePrice grouped by year sold (year sold x-axis, SalePrice y-axis)

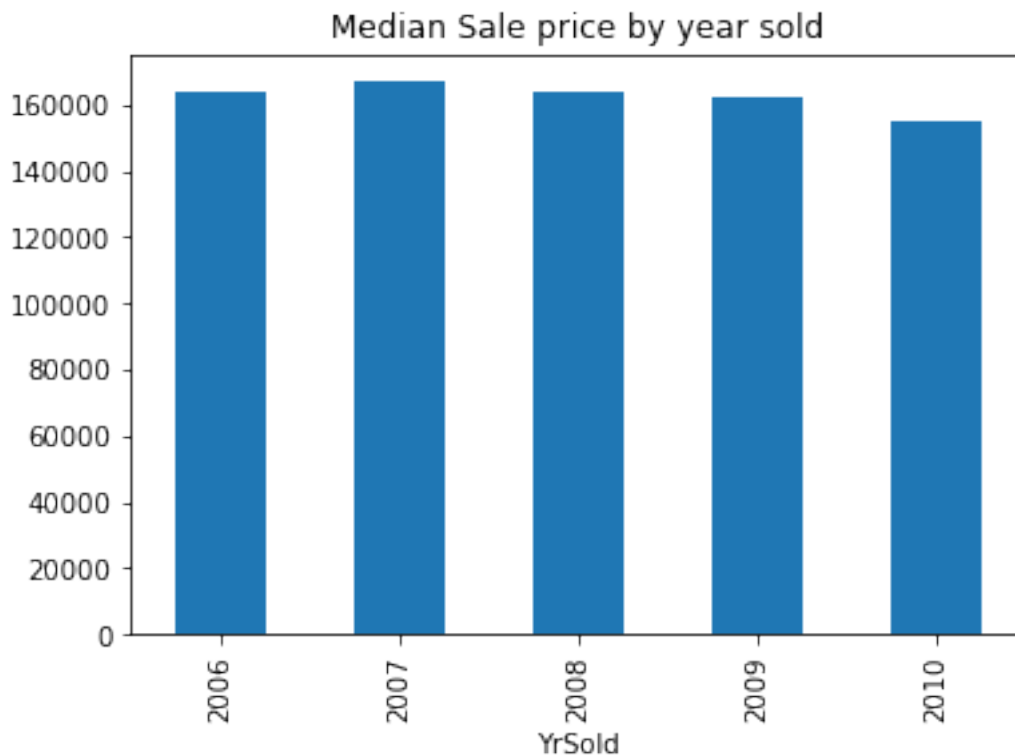
```
[10]: Dropped_na_boston.dtypes
```

```
[10]: Id                int64
      MSSubClass        int64
      MSZoning          object
      LotFrontage       float64
      LotArea           int64
      ...
      MoSold            int64
      YrSold            int64
      SaleType          object
      SaleCondition     object
      SalePrice         int64
      Length: 76, dtype: object
```

```
[11]: # There are 35 discrete variables in the dataset.
      # There are 4 temporal variables in the data set. (YrSold, GarageYrBlt,
      ↪YearRemodAdd, YearBuilt)

[12]: Median_Saleprice_by_YrSold = Dropped_na_boston.groupby('YrSold')['SalePrice'].
      ↪median()
      Median_Saleprice_by_YrSold.plot(kind = 'bar', title = 'Median Sale price by
      ↪year sold')

[12]: <AxesSubplot:title={'center':'Median Sale price by year sold'}, xlabel='YrSold'>
```



Continuous

- Determine the number of continuous variables
- Plot a histogram of each
- Use one normalization method on one column to create a normal-like distribution (does not need to be on all columns)

```
[13]: Dropped_na_boston.dtypes
```

```
[13]: Id                int64
      MSSubClass         int64
      MSZoning           object
      LotFrontage        float64
```

```

LotArea          int64
...
MoSold           int64
YrSold           int64
SaleType         object
SaleCondition    object
SalePrice        int64
Length: 76, dtype: object

```

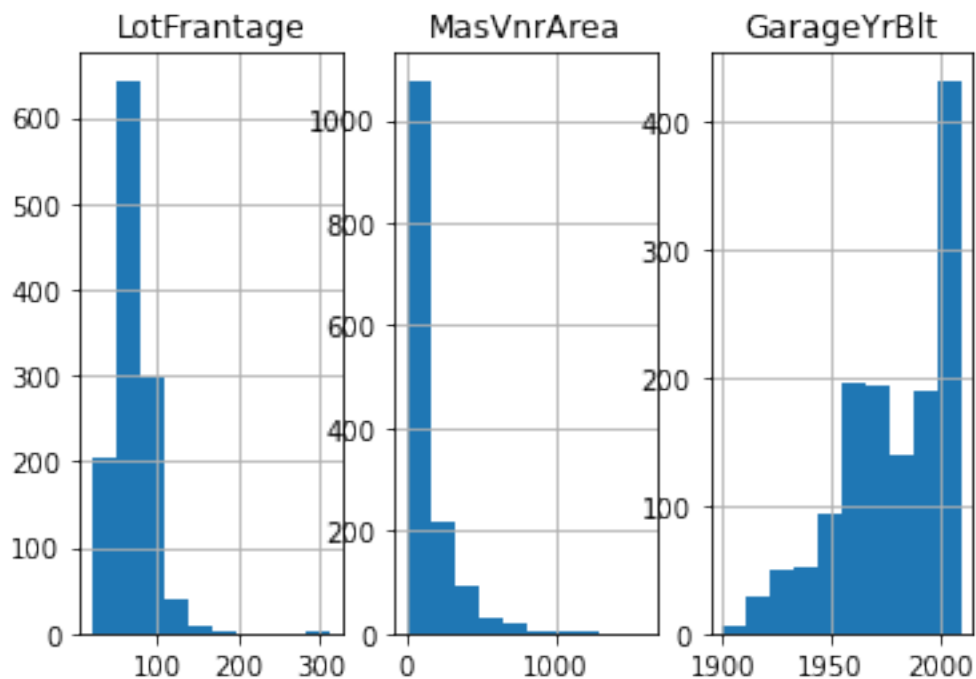
```
[14]: # There are 3 continuous variables in the dataset (LotFrontage, MasVnrArea,
      ↪ GarageYrBlt)
```

```
[15]: plt.subplot(1,3,1)
      plt.title('LotFrantage')
      Dropped_na_boston['LotFrontage'].hist()

      plt.subplot(1, 3, 2)
      plt.title('MasVnrArea')
      Dropped_na_boston['MasVnrArea'].hist()

      plt.subplot(1, 3, 3)
      plt.title('GarageYrBlt')
      Dropped_na_boston['GarageYrBlt'].hist()
```

```
[15]: <AxesSubplot:title={'center':'GarageYrBlt'}>
```

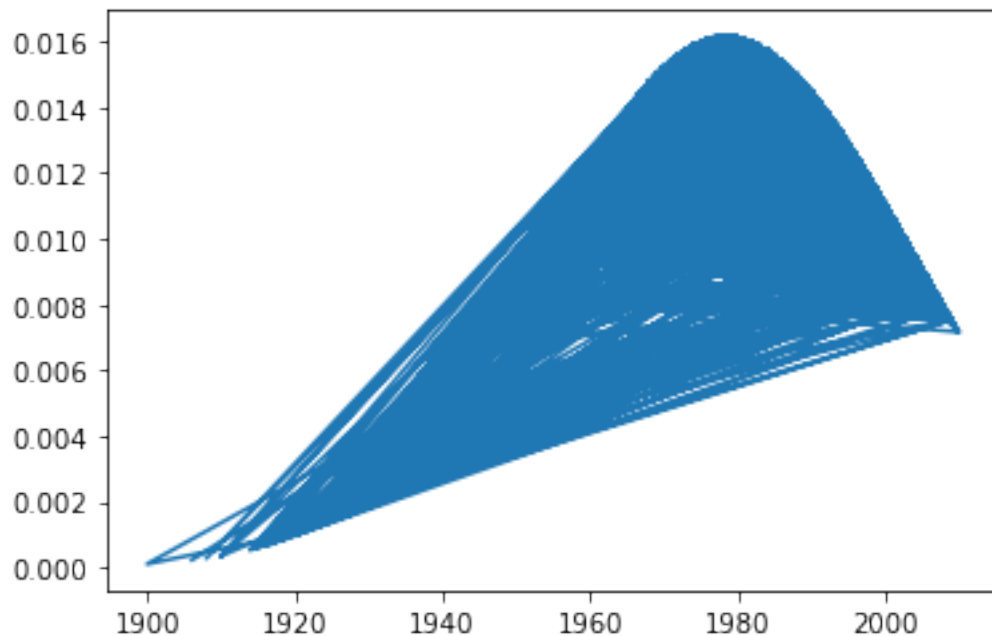


```
[16]: Dropped = Dropped_na_boston['GarageYrBlt']
```

```
[17]: Dropped = Dropped.dropna()
```

```
[18]: x_axis = Dropped
mean = statistics.mean(x_axis)
sd = statistics.stdev(x_axis)

plt.plot(x_axis, norm.pdf(x_axis, mean, sd))
plt.show()
```



0.1.4 Work with categorical variables

- Graph the number of unique variables per column
- Map the following on columns that apply:
qual_mappings = {'Po': 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0}
- Create a box and whisker plot on SalePrice for 3 columns that utilize this mapped metric

```
[19]: Unique_Variables_per_column = Dropped_na_boston.nunique()
Unique_Variables_per_column
```

```
[19]: Id                1460
MSSubClass             15
MSZoning                5
LotFrontage           110
LotArea               1073
```

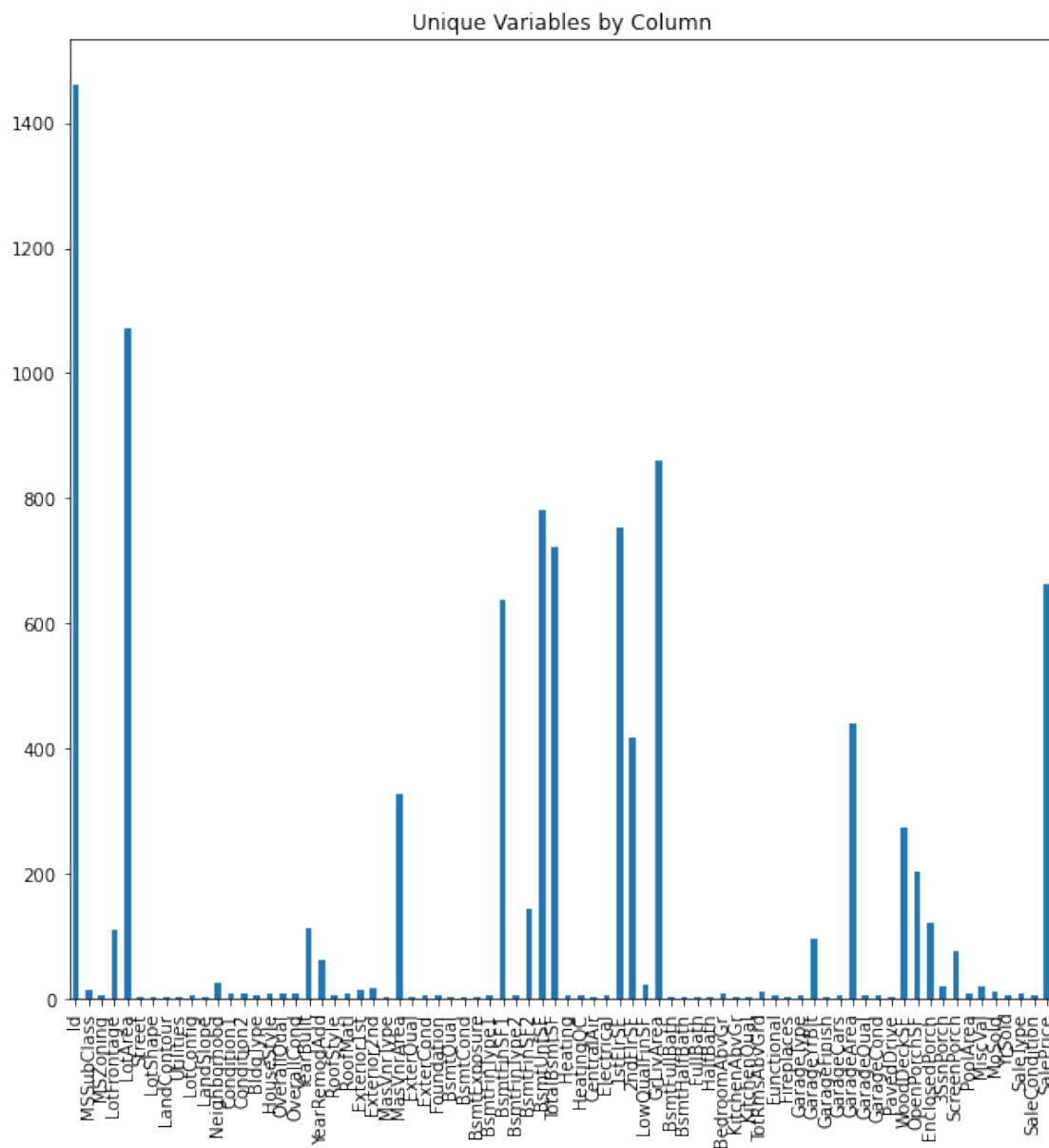
```

...
MoSold          12
YrSold          5
SaleType        9
SaleCondition    6
SalePrice      663
Length: 76, dtype: int64

```

```
[20]: Unique_Variables_per_column.plot(kind='bar', figsize=(10,10), title='Unique_
↳Variables by Column')
```

```
[20]: <AxesSubplot:title={ 'center': 'Unique Variables by Column' }>
```



```
[21]: Dropped_na_boston['ExterQual_map'] = Dropped_na_boston['ExterQual'].map({'Po': 1,
↪ 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['ExterCond_map'] = Dropped_na_boston['ExterCond'].map({'Po': 1,
↪ 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['BsmtQual_map'] = Dropped_na_boston['BsmtQual'].map({'Po': 1,
↪ 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['BsmtCond_map'] = Dropped_na_boston['BsmtCond'].map({'Po': 1,
↪ 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['BsmtExposure_map'] = Dropped_na_boston['BsmtExposure'].
↪ map({'Po': 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['HeatingQC_map'] = Dropped_na_boston['HeatingQC'].map({'Po': 1,
↪ 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['KitchenQual_map'] = Dropped_na_boston['KitchenQual'].
↪ map({'Po': 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['GarageQual_map'] = Dropped_na_boston['GarageQual'].map({'Po':
↪ 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston['GarageCond_map'] = Dropped_na_boston['GarageCond'].map({'Po':
↪ 1, 'Fa': 2, 'TA': 3, 'Gd': 4, 'Ex': 5, 'Missing': 0, 'NA': 0})
Dropped_na_boston
```

```
[21]:
```

	Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	LotShape	\
0	1	60	RL	65.0	8450	Pave	Reg	
1	2	20	RL	80.0	9600	Pave	Reg	
2	3	60	RL	68.0	11250	Pave	IR1	
3	4	70	RL	60.0	9550	Pave	IR1	
4	5	60	RL	84.0	14260	Pave	IR1	
...	
1455	1456	60	RL	62.0	7917	Pave	Reg	
1456	1457	20	RL	85.0	13175	Pave	Reg	
1457	1458	70	RL	66.0	9042	Pave	Reg	
1458	1459	20	RL	68.0	9717	Pave	Reg	
1459	1460	20	RL	75.0	9937	Pave	Reg	
	LandContour	Utilities	LotConfig	LandSlope	Neighborhood	Condition1	\	
0	Lvl	AllPub	Inside	Gtl	CollgCr	Norm		
1	Lvl	AllPub	FR2	Gtl	Veenker	Feedr		
2	Lvl	AllPub	Inside	Gtl	CollgCr	Norm		
3	Lvl	AllPub	Corner	Gtl	Crawfor	Norm		
4	Lvl	AllPub	FR2	Gtl	NoRidge	Norm		
...		
1455	Lvl	AllPub	Inside	Gtl	Gilbert	Norm		
1456	Lvl	AllPub	Inside	Gtl	NWAmes	Norm		
1457	Lvl	AllPub	Inside	Gtl	Crawfor	Norm		
1458	Lvl	AllPub	Inside	Gtl	NAmes	Norm		

1459	Lvl	AllPub	Inside	Gtl	Edwards	Norm
	Condition2	BldgType	HouseStyle	OverallQual	OverallCond	YearBuilt
0	Norm	1Fam	2Story	7	5	2003
1	Norm	1Fam	1Story	6	8	1976
2	Norm	1Fam	2Story	7	5	2001
3	Norm	1Fam	2Story	7	5	1915
4	Norm	1Fam	2Story	8	5	2000
...
1455	Norm	1Fam	2Story	6	5	1999
1456	Norm	1Fam	1Story	6	6	1978
1457	Norm	1Fam	2Story	7	9	1941
1458	Norm	1Fam	1Story	5	6	1950
1459	Norm	1Fam	1Story	5	6	1965

	YearRemodAdd	RoofStyle	RoofMatl	Exterior1st	Exterior2nd	MasVnrType
0	2003	Gable	CompShg	VinylSd	VinylSd	BrkFace
1	1976	Gable	CompShg	MetalSd	MetalSd	None
2	2002	Gable	CompShg	VinylSd	VinylSd	BrkFace
3	1970	Gable	CompShg	Wd Sdng	Wd Shng	None
4	2000	Gable	CompShg	VinylSd	VinylSd	BrkFace
...
1455	2000	Gable	CompShg	VinylSd	VinylSd	None
1456	1988	Gable	CompShg	Plywood	Plywood	Stone
1457	2006	Gable	CompShg	CemntBd	CmentBd	None
1458	1996	Hip	CompShg	MetalSd	MetalSd	None
1459	1965	Gable	CompShg	HdBoard	HdBoard	None

	MasVnrArea	ExterQual	ExterCond	Foundation	BsmtQual	BsmtCond
0	196.0	Gd	TA	PConc	Gd	TA
1	0.0	TA	TA	CBlock	Gd	TA
2	162.0	Gd	TA	PConc	Gd	TA
3	0.0	TA	TA	BrkTil	TA	Gd
4	350.0	Gd	TA	PConc	Gd	TA
...
1455	0.0	TA	TA	PConc	Gd	TA
1456	119.0	TA	TA	CBlock	Gd	TA
1457	0.0	Ex	Gd	Stone	TA	Gd
1458	0.0	TA	TA	CBlock	TA	TA
1459	0.0	Gd	TA	CBlock	TA	TA

	BsmtExposure	BsmtFinType1	BsmtFinSF1	BsmtFinType2	BsmtFinSF2
0	No	GLQ	706	Unf	0
1	Gd	ALQ	978	Unf	0
2	Mn	GLQ	486	Unf	0
3	No	ALQ	216	Unf	0
4	Av	GLQ	655	Unf	0

...
1455	No	Unf	0	Unf	0
1456	No	ALQ	790	Rec	163
1457	No	GLQ	275	Unf	0
1458	Mn	GLQ	49	Rec	1029
1459	No	BLQ	830	LwQ	290

	BsmtUnfSF	TotalBsmtSF	Heating	HeatingQC	CentralAir	Electrical	\
0	150	856	GasA	Ex	Y	SBrkr	
1	284	1262	GasA	Ex	Y	SBrkr	
2	434	920	GasA	Ex	Y	SBrkr	
3	540	756	GasA	Gd	Y	SBrkr	
4	490	1145	GasA	Ex	Y	SBrkr	
...	
1455	953	953	GasA	Ex	Y	SBrkr	
1456	589	1542	GasA	TA	Y	SBrkr	
1457	877	1152	GasA	Ex	Y	SBrkr	
1458	0	1078	GasA	Gd	Y	FuseA	
1459	136	1256	GasA	Gd	Y	SBrkr	

	1stFlrSF	2ndFlrSF	LowQualFinSF	GrLivArea	BsmtFullBath	BsmtHalfBath	\
0	856	854	0	1710	1	0	
1	1262	0	0	1262	0	1	
2	920	866	0	1786	1	0	
3	961	756	0	1717	1	0	
4	1145	1053	0	2198	1	0	
...	
1455	953	694	0	1647	0	0	
1456	2073	0	0	2073	1	0	
1457	1188	1152	0	2340	0	0	
1458	1078	0	0	1078	1	0	
1459	1256	0	0	1256	1	0	

	FullBath	HalfBath	BedroomAbvGr	KitchenAbvGr	KitchenQual	\
0	2	1	3	1	Gd	
1	2	0	3	1	TA	
2	2	1	3	1	Gd	
3	1	0	3	1	Gd	
4	2	1	4	1	Gd	
...	
1455	2	1	3	1	TA	
1456	2	0	3	1	TA	
1457	2	0	4	1	Gd	
1458	1	0	2	1	Gd	
1459	1	1	3	1	TA	

TotRmsAbvGrd	Functional	Fireplaces	GarageType	GarageYrBlt	\
--------------	------------	------------	------------	-------------	---

0	8	Typ	0	Attchd	2003.0
1	6	Typ	1	Attchd	1976.0
2	6	Typ	1	Attchd	2001.0
3	7	Typ	1	Detchd	1998.0
4	9	Typ	1	Attchd	2000.0
...
1455	7	Typ	1	Attchd	1999.0
1456	7	Min1	2	Attchd	1978.0
1457	9	Typ	2	Attchd	1941.0
1458	5	Typ	0	Attchd	1950.0
1459	6	Typ	0	Attchd	1965.0

	GarageFinish	GarageCars	GarageArea	GarageQual	GarageCond	PavedDrive	\
0	RFn	2	548	TA	TA	Y	
1	RFn	2	460	TA	TA	Y	
2	RFn	2	608	TA	TA	Y	
3	Unf	3	642	TA	TA	Y	
4	RFn	3	836	TA	TA	Y	
...	
1455	RFn	2	460	TA	TA	Y	
1456	Unf	2	500	TA	TA	Y	
1457	RFn	1	252	TA	TA	Y	
1458	Unf	1	240	TA	TA	Y	
1459	Fin	1	276	TA	TA	Y	

	WoodDeckSF	OpenPorchSF	EnclosedPorch	3SsnPorch	ScreenPorch	\
0	0	61	0	0	0	
1	298	0	0	0	0	
2	0	42	0	0	0	
3	0	35	272	0	0	
4	192	84	0	0	0	
...	
1455	0	40	0	0	0	
1456	349	0	0	0	0	
1457	0	60	0	0	0	
1458	366	0	112	0	0	
1459	736	68	0	0	0	

	PoolArea	MiscVal	MoSold	YrSold	SaleType	SaleCondition	SalePrice	\
0	0	0	2	2008	WD	Normal	208500	
1	0	0	5	2007	WD	Normal	181500	
2	0	0	9	2008	WD	Normal	223500	
3	0	0	2	2006	WD	Abnorml	140000	
4	0	0	12	2008	WD	Normal	250000	
...	
1455	0	0	8	2007	WD	Normal	175000	
1456	0	0	2	2010	WD	Normal	210000	

1457	0	2500	5	2010	WD	Normal	266500
1458	0	0	4	2010	WD	Normal	142125
1459	0	0	6	2008	WD	Normal	147500

	ExterQual_map	ExterCond_map	BsmtQual_map	BsmtCond_map	\
0	4	3	4.0	3.0	
1	3	3	4.0	3.0	
2	4	3	4.0	3.0	
3	3	3	3.0	4.0	
4	4	3	4.0	3.0	
...	
1455	3	3	4.0	3.0	
1456	3	3	4.0	3.0	
1457	5	4	3.0	4.0	
1458	3	3	3.0	3.0	
1459	4	3	3.0	3.0	

	BsmtExposure_map	HeatingQC_map	KitchenQual_map	GarageQual_map	\
0	NaN	5	4	3.0	
1	4.0	5	3	3.0	
2	NaN	5	4	3.0	
3	NaN	4	4	3.0	
4	NaN	5	4	3.0	
...	
1455	NaN	5	3	3.0	
1456	NaN	3	3	3.0	
1457	NaN	5	4	3.0	
1458	NaN	4	4	3.0	
1459	NaN	4	3	3.0	

	GarageCond_map
0	3.0
1	3.0
2	3.0
3	3.0
4	3.0
...	...
1455	3.0
1456	3.0
1457	3.0
1458	3.0
1459	3.0

[1460 rows x 85 columns]

```
[22]: plt.subplot(1,3,1)
plt.title('ExternalQual_map')
```

```
plt.boxplot(Dropped_na_boston['ExterQual_map'])

plt.subplot(1, 3, 2)
plt.title('HeatingQC_map')
plt.boxplot(Dropped_na_boston['HeatingQC_map'])

plt.subplot(1, 3, 3)
plt.title('KitchenQual_map')
plt.boxplot(Dropped_na_boston['KitchenQual_map'])

plt.subplots_adjust(wspace=0.5)
```

