Quiz Notebook-Finished

January 18, 2022

0.1 Data Analysis Exercise

The objective of this notebook 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]:
               Ιd
                   MSSubClass MSZoning
                                           LotFrontage
                                                           LotArea Street Alley LotShape
     0
                1
                                                    65.0
                             60
                                       RL
                                                              8450
                                                                              NaN
                                                                      Pave
                                                                                         Reg
                2
     1
                             20
                                       RL
                                                    80.0
                                                              9600
                                                                      Pave
                                                                              NaN
                                                                                         Reg
                3
     2
                             60
                                       RL
                                                    68.0
                                                             11250
                                                                      Pave
                                                                              NaN
                                                                                         IR1
     3
                4
                             70
                                                    60.0
                                                                              NaN
                                       RL
                                                              9550
                                                                      Pave
                                                                                         IR1
                5
     4
                             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		ıve		Reg
1458	1459	20	RL	68.0		ıve		Reg
1459	1460	20	RL	75.0	9937 Pa	ıve	NaN	Reg
	LandContour	Utilities L	otConfig	LandSlope	Neighborhoo	d Co	ondition1	\
0	Lvl	AllPub	Inside	Gt1	CollgC	r	Norm	
1	Lvl	AllPub	FR2	Gt1	Veenke	er	Feedr	
2	Lvl	AllPub	Inside	Gt1	CollgC	r	Norm	
3	Lvl	AllPub	Corner	Gt1	Crawfo	r	Norm	
4	Lvl	AllPub	FR2	Gtl	NoRidg	ge	Norm	
 1455	 Lvl	 AllPub	 Inside	 Gtl	 Gilber	:t	Norm	
1456	Lvl	AllPub	Inside	Gtl	NWAme		Norm	
1457	Lvl	AllPub	Inside	Gtl	Crawfo		Norm	
1458	Lvl	AllPub	Inside	Gtl	NAme		Norm	
1459	Lvl	AllPub	Inside	Gtl	Edward		Norm	
	Condition2 B	ldgTyne Hou	seStvle	OverallOua	l OverallO	ond.	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	
			25 CO1 y			J	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	
1409	NOTIII	11. qiii	15001 y		5	O	1900	
	YearRemodAd	d RoofStyle	RoofMat	l Exterior1	st Exterior	2nd	MasVnrType	\
0	200	3 Gable	CompSh	g Vinyl	Sd Viny	rlSd	${\tt BrkFace}$	
1	197	6 Gable	CompSh	g Metal	Sd Meta	ılSd	None	
2	200	2 Gable	CompSh	g Vinyl	Sd Viny	rlSd	${ t BrkFace}$	
3	197	O Gable	CompSh	g Wd Sd	ng Wd S	hng	None	
4	200	O Gable	CompSh	g Vinyl	Sd Viny	rlSd	${ t BrkFace}$	
•••	•••	•••	•••	•••				
1455	200	0 Gable	CompSh	g Vinyl	Sd Viny	rlSd	None	
1456	198	8 Gable	CompSh	g Plywo	od Plyw	rood	Stone	
1457	200	6 Gable	CompSh	g Cemnt	Bd Cmen	ıtBd	None	
1458	199	6 Hip	CompSh	g Metal	Sd Meta	ılSd	None	
1459	196	5 Gable	CompShg	g HdBoa	rd HdBc	ard	None	
	MasVnrArea	ExterQual E	ExterCond	Foundation	BsmtQual B	SsmtC	Cond \	
0	196.0	Gd	TA	PConc			TA	
1	0.0	TA	TA	CBlock			TA	
2	162.0	Gd	TA	PConc			TA	

3	0.	0 T	Ά	TA	Ві	ckTil	TA	Go	d			
4	350.	0 G	d	TA		Conc	Gd	T	A			
	•••	•••	•••	•••		•••						
1455	0.	0 T	Ά	TA	F	Conc	Gd	T	A			
1456	119.	0 T	Ά	TA	CI	Block	Gd	T	A			
1457	0.	0 E	x	Gd	S	Stone	TA	Go	d			
1458	0.	0 T	Ά	TA	CE	Block	TA	T	A			
1459	0.	0 G	d	TA	CE	Block	TA	T	A			
	BsmtExposu	re BamtFin	Tune 1	RamtFir	nSF1	RemtFir	Tune?	BsmtFins	SF2	\		
0	_	No	GLQ	Domor 11	706	Domor II.	Unf	Domor 111	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			
- 	···	·	4									
1455		No	Unf		0		Unf		0			
1456		No	ALQ		790		Rec	:	163			
1457		No	GLQ		275		Unf		0			
1458		Mn	GLQ		49		Rec	10	029			
1459		No	BLQ		830		LwQ		290			
	${\tt BsmtUnfSF}$			ating He	eatir	ngQC Cer	ntralAi			\		
0	150		856	${\tt GasA}$		Ex	7	Y SI	Brkr			
1	284	1	262	${\tt GasA}$		Ex	7	Y SI	Brkr			
2	434		920	${\tt GasA}$		Ex	7	Y SI	Brkr			
3	540		756	${\tt GasA}$		Gd	7	Y SI	Brkr			
4	490	1	145	${\tt GasA}$		Ex	7	Y SI	3rkr			
•••	•••	•••	•••	•••		•••	•••					
1455	953		953	${\tt GasA}$		Ex			Brkr			
1456	589		542	${\tt GasA}$		TA			Brkr			
1457	877		152	${\tt GasA}$		Ex			Brkr			
1458	0		078	${\tt GasA}$		Gd			ıseA			
1459	136	1	256	GasA		Gd	`	Y SI	Brkr			
	1stFlrSF	2ndFlrSF	ΙοωΟιι	alFinSF	Grī	LivArea	RemtF1	ullBath	Remt	·HalfRa	a+h	١
0	856	854	шожща	0	411	1710	Domor	1	Dome	JIIGIIDO	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	
<u></u>						2100		<u></u>			v	
 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	
- 100	1200	J		v		_200		-			•	

	FullBath H	IalfBath Bed	roomAbvGr Ki	tchenAbvGr Ki	tchenQual	\	
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		
	${\tt TotRmsAbvGr}$	d Functional	Fireplaces	FireplaceQu G	arageType	${\tt GarageYrBlt}$	\
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	
	Compactiniah	ComomoComa	CamamaAmaa	CamamaDual Car	mamaCand D	orrodDmirro \	
0	GarageFinish RFn	_	GarageArea 548	GarageQual Ga: TA	ragecond Pa	avedDrive \ Y	
0	RFr		460	TA	TA	Y	
1 2	RFr		608	TA	TA	Y	
3	Unf		642	TA	TA	Y	
4	RFr		836	TA	TA	Y	
-	1011				ın	1	
 1455	 RFr	 1 2	460	 TA	 TA	Y	
1456	Unf		500	TA	TA	Y	
1457	RFr		252	TA	TA	Y	
1458	Unf		240	TA	TA	Y	
1459	Fir		276	TA	TA	Y	
1100	1 11	<u>. </u>	210	111	111	•	
	WoodDeckSF	OpenPorchSF	EnclosedPor	ch 3SsnPorch	ScreenPo	rch \	
0	0	61		0 0		0	
1	298	0		0 0		0	
2	0	42		0 0		0	
3	0	35	2	72 0		0	
4	192	84		0 0		0	
	•••	•••					

1455	0		40		0	0		0	
1456	349	9	()	0	0		0	
1457	(0	60)	0	0		0	
1458	366	3)	112	0		0	
1459	736	6	68	3	0	0		0	
	PoolArea H	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	${\tt MnPrv}$	NaN	0	2	2010	WD	
1457	0	NaN	${\tt GdPrv}$	Shed	2500	5	2010	WD	
1458	0	NaN	NaN	NaN	0	4	2010	WD	
1459	0	NaN	NaN	NaN	0	6	2008	WD	
	SaleCondit	ion Sa	lePrice	e					
0	Norn		208500)					
1	Norn	mal	181500)					
2	Norr	mal	223500)					
3	Abno	rml	140000)					
4	Nor	mal	250000)					

[1460 rows x 81 columns]

Normal

Normal

Normal

Normal

Normal

1455

1456

1457

1458

1459

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

0.1.3 Inspect missing values and deal with missing values

175000

210000

266500

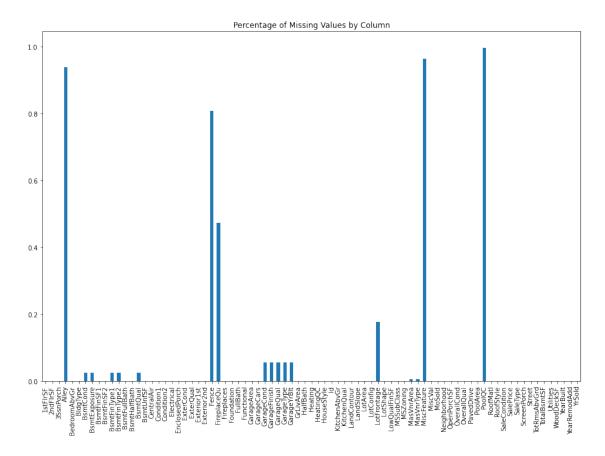
142125

147500

- 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 = 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['GarageYrBlt_2'] = boston['GarageYrBlt'].isnull()
  boston['LotFrontage_2'] = boston['LotFrontage'].isnull()
```

```
boston['MiscFeature_2'] = boston['MiscFeature'].isnull()
     boston['PoolQC_2'] = boston['PoolQC'].isnull()
     boston
                  MSSubClass MSZoning LotFrontage LotArea Street Alley LotShape \
[6]:
     0
               1
                           60
                                     RL
                                                 65.0
                                                           8450
                                                                   Pave
                                                                          NaN
                                                                                    Reg
     1
               2
                           20
                                     RL
                                                 80.0
                                                           9600
                                                                          NaN
                                                                   Pave
                                                                                    Reg
               3
     2
                           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
                                                                          NaN
                                                                                    IR1
                                                                   Pave
     1455 1456
                                     RL
                                                 62.0
                           60
                                                           7917
                                                                   Pave
                                                                          NaN
                                                                                    Reg
                                                 85.0
     1456
           1457
                           20
                                     RL
                                                          13175
                                                                   Pave
                                                                          NaN
                                                                                    Reg
     1457
           1458
                           70
                                     R.T.
                                                 66.0
                                                           9042
                                                                          NaN
                                                                   Pave
                                                                                    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
                                      Inside
                                                    Gtl
     0
                   Lvl
                           AllPub
                                                              CollgCr
                                                                             Norm
     1
                   Lvl
                           AllPub
                                         FR2
                                                    Gtl
                                                              Veenker
                                                                            Feedr
     2
                   Lvl
                           AllPub
                                      Inside
                                                    Gtl
                                                              CollgCr
                                                                              Norm
                                      Corner
     3
                   Lvl
                           AllPub
                                                    Gtl
                                                              Crawfor
                                                                              Norm
     4
                   Lvl
                           AllPub
                                         FR2
                                                    Gtl
                                                              NoRidge
                                                                              Norm
     1455
                           AllPub
                                      Inside
                                                    Gtl
                   Lvl
                                                              Gilbert
                                                                              Norm
                                      Inside
                                                    Gtl
     1456
                   Lvl
                           AllPub
                                                               NWAmes
                                                                              Norm
     1457
                   Lvl
                           AllPub
                                      Inside
                                                    Gtl
                                                              Crawfor
                                                                              Norm
     1458
                   Lvl
                           AllPub
                                      Inside
                                                    Gtl
                                                                 NAmes
                                                                              Norm
     1459
                   T.v.T
                           AllPub
                                      Inside
                                                    Gtl
                                                              Edwards
                                                                             Norm
                                                            OverallCond
          Condition2 BldgType HouseStyle OverallQual
                                                                          YearBuilt
     0
                 Norm
                           1Fam
                                     2Story
                                                         7
                                                                       5
                                                                                2003
     1
                 Norm
                           1Fam
                                     1Story
                                                         6
                                                                       8
                                                                                1976
     2
                 Norm
                                                         7
                                                                       5
                           1Fam
                                     2Story
                                                                                2001
     3
                 Norm
                           1Fam
                                     2Story
                                                         7
                                                                       5
                                                                                1915
     4
                 Norm
                           1Fam
                                     2Story
                                                                       5
                                                                                2000
     1455
                 Norm
                           1Fam
                                     2Story
                                                                       5
                                                                                1999
                                                         6
     1456
                 Norm
                           1Fam
                                                                       6
                                                                                1978
                                     1Story
                                                         6
                                                         7
                                                                       9
     1457
                 Norm
                           1Fam
                                     2Story
                                                                                1941
     1458
                 Norm
                           1Fam
                                     1Story
                                                         5
                                                                       6
                                                                                1950
                                                         5
     1459
                 Norm
                           1Fam
                                     1Story
                                                                       6
                                                                                1965
           YearRemodAdd RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType
     0
                    2003
                              Gable CompShg
                                                   VinylSd
                                                                VinylSd
                                                                            BrkFace
```

boston['MasVnrArea_2'] = boston['MasVnrArea'].isnull()
boston['MasVnrType_2'] = boston['MasVnrType'].isnull()

1	1976	Gable (CompShg	MetalSd	MetalSo	d None
2	2002		CompShg	VinylSd	VinylS	
3	1970	Gable (CompShg	Wd Sdng	Wd Shn	
4	2000	Gable (CompShg	VinylSd	VinylS	d BrkFace
•••	•••				•••	
1455	2000	Gable (CompShg	VinylSd	VinylS	d None
1456	1988	Gable (CompShg	Plywood	Plywoo	d Stone
1457	2006	Gable (CompShg	CemntBd	CmentBo	d None
1458	1996	Hip (CompShg	MetalSd	MetalSo	d None
1459	1965	Gable (CompShg	HdBoard	HdBoar	d None
	MasVnrArea Ex					
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
0	BsmtExposure B		BsmtFi			tFinSF2 \
0	No	GLQ	BsmtFi	706	Unf	0
1	No Gd	GLQ ALQ	BsmtFi	706 978	Unf Unf	0 0
1 2	No Gd Mn	GLQ ALQ GLQ	BsmtFi	706 978 486	Unf Unf Unf	0 0 0
1 2 3	No Gd Mn No	GLQ ALQ GLQ ALQ	BsmtFi	706 978 486 216	Unf Unf Unf Unf	0 0 0 0
1 2	No Gd Mn	GLQ ALQ GLQ		706 978 486 216 655	Unf Unf Unf Unf Unf	0 0 0
1 2 3 4 	No Gd Mn No Av	GLQ ALQ GLQ ALQ GLQ	BsmtFi 	706 978 486 216 655	Unf Unf Unf Unf Unf Unf	0 0 0 0
1 2 3 4 1455	No Gd Mn No Av 	GLQ ALQ GLQ ALQ GLQ Unf		706 978 486 216 655 	Unf Unf Unf Unf Unf Unf Unf	0 0 0 0 0
1 2 3 4 1455 1456	No Gd Mn No Av No	GLQ ALQ GLQ ALQ GLQ Unf ALQ		706 978 486 216 655 0	Unf Unf Unf Unf Unf Unf Unf Rec	0 0 0 0 0
1 2 3 4 1455 1456 1457	No Gd Mn No Av No No	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ		706 978 486 216 655 0 790 275	Unf Unf Unf Unf Unf Unf Unf Unf Rec Unf	0 0 0 0 0 0 163 0
1 2 3 4 1455 1456 1457 1458	No Gd Mn No Av No No No Mn	GLQ ALQ GLQ GLQ Unf ALQ GLQ GLQ		706 978 486 216 655 0 790 275 49	Unf Unf Unf Unf Unf Unf Unf Unf Rec Unf Rec	0 0 0 0 0 0 163 0 1029
1 2 3 4 1455 1456 1457	No Gd Mn No Av No No	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ		706 978 486 216 655 0 790 275	Unf Unf Unf Unf Unf Unf Unf Unf Rec Unf	0 0 0 0 0 0 163 0
1 2 3 4 1455 1456 1457 1458	No Gd Mn No Av No No No No No	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ GLQ		706 978 486 216 655 0 790 275 49 830	Unf Unf Unf Unf Unf Unf Unf Lnf Rec Unf Rec LwQ	0 0 0 0 0 163 0 1029 290
1 2 3 4 1455 1456 1457 1458 1459	No Gd Mn No Av No No No No No Mn No BsmtUnfSF To	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ	 eating H	706 978 486 216 655 0 790 275 49 830 eatingQC Cen	Unf Unf Unf Unf Unf Unf Cunf Rec Unf Rec LwQ	0 0 0 0 0 163 0 1029 290
1 2 3 4 1455 1456 1457 1458 1459	No Gd Mn No Av No No No No Mn No BsmtUnfSF To	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ BLQ	 eating H GasA	706 978 486 216 655 0 790 275 49 830 eatingQC Cen	Unf Unf Unf Unf Unf Unf Cunf Rec Unf Rec LwQ CtralAir Ele	0 0 0 0 0 163 0 1029 290 ectrical \
1 2 3 4 1455 1456 1457 1458 1459	No Gd Mn No Av No No No No Mn No BsmtUnfSF 150 284	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ StalBsmtSF He 856 1262	 eating H GasA GasA	706 978 486 216 655 0 790 275 49 830 eatingQC Cen Ex Ex	Unf Unf Unf Unf Unf Unf Cunf Rec Unf Rec LwQ CtralAir Ele Y	0 0 0 0 0 0 163 0 1029 290 ectrical \ SBrkr SBrkr
1 2 3 4 1455 1456 1457 1458 1459	No Gd Mn No Av No No No No Mn No BsmtUnfSF 150 284 434	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ btalBsmtSF He 856 1262 920	 GasA GasA GasA	706 978 486 216 655 0 790 275 49 830 eatingQC Cen Ex Ex Ex	Unf Unf Unf Unf Unf Unf Unf Rec Unf Rec LwQ tralAir Ele Y Y	0 0 0 0 0 163 0 1029 290 ectrical \ SBrkr SBrkr
1 2 3 4 1455 1456 1457 1458 1459	No Gd Mn No Av No No No No Mn No SsmtUnfSF To 150 284 434 540	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ BLQ DtalBsmtSF He 856 1262 920 756	 GasA GasA GasA GasA	706 978 486 216 655 0 790 275 49 830 eatingQC Cen Ex Ex Ex Cd	Unf Unf Unf Unf Unf Unf Cunf Rec Unf Rec LwQ CtralAir Ele Y Y Y	0 0 0 0 0 163 0 1029 290 ectrical \ SBrkr SBrkr SBrkr
1 2 3 4 1455 1456 1457 1458 1459	No Gd Mn No Av No No No Mn No SsmtUnfSF To 150 284 434 540 490	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ btalBsmtSF He 856 1262 920	 GasA GasA GasA GasA GasA	706 978 486 216 655 0 790 275 49 830 eatingQC Cen Ex Ex Ex Gd Ex	Unf Unf Unf Unf Unf Unf Unf Rec Unf Rec LwQ tralAir Ele Y Y	0 0 0 0 0 163 0 1029 290 ectrical \ SBrkr SBrkr
1 2 3 4 1455 1456 1457 1458 1459 0 1 2 3 4 	No Gd Mn No Av No No No Mn No EsmtUnfSF To 150 284 434 540 490	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ CtalBsmtSF He 856 1262 920 756 1145	 GasA GasA GasA GasA GasA	706 978 486 216 655 0 790 275 49 830 eatingQC Cen Ex Ex Ex Gd Ex 	Unf Unf Unf Unf Unf Unf Unf Rec Unf Rec LwQ AtralAir Ele Y Y Y Y Y Y	0 0 0 0 0 163 0 1029 290 ectrical \ SBrkr SBrkr SBrkr SBrkr
1 2 3 4 1455 1456 1457 1458 1459 0 1 2 3 4 1455	Solve See See See See See See See See See S	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ BLQ OtalBsmtSF He 856 1262 920 756 1145 953	 GasA GasA GasA GasA GasA GasA	706 978 486 216 655 0 790 275 49 830 TeatingQC Cen Ex Ex Ex Gd Ex Ex	Unf Unf Unf Unf Unf Unf Unf Rec Unf Rec LwQ UrralAir Ele Y Y Y Y Y Y Y Y	0 0 0 0 0 163 0 1029 290 ectrical \ SBrkr SBrkr SBrkr SBrkr SBrkr
1 2 3 4 1455 1456 1457 1458 1459 0 1 2 3 4 	No Gd Mn No Av No No No Mn No EsmtUnfSF To 150 284 434 540 490	GLQ ALQ GLQ ALQ GLQ Unf ALQ GLQ GLQ BLQ CtalBsmtSF He 856 1262 920 756 1145	 GasA GasA GasA GasA GasA	706 978 486 216 655 0 790 275 49 830 eatingQC Cen Ex Ex Ex Gd Ex 	Unf Unf Unf Unf Unf Unf Unf Rec Unf Rec LwQ AtralAir Ele Y Y Y Y Y Y	0 0 0 0 0 163 0 1029 290 ectrical \ SBrkr SBrkr SBrkr SBrkr

1458	0	1	.078	GasA	Gd		Y I	FuseA	
1459	136	1	.256	GasA	Gd		Y S	SBrkr	
	1stFlrSF	2ndFlrSF	I own	ualFinSF	GrLivArea	Ran	ntFullBath	BsmtHalfBath	\
0	856	854	томф	0	1710	וופת	1	DSIMCHAIIDACH 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	
-				. <u></u>	2100			· ·	
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	
1100	1200	V		V	1200		•	· ·	
	FullBath	HalfBath	Bedr	oomAbvGr	KitchenAby	vGr k	KitchenQua]	L \	
0	2	1		3		1	Go		
1	2	0		3		1	T		
2	2	1		3		1	Go		
3	1	0		3		1	Go		
4	2	1		4		1	Go		
-					•••		a.	•	
1455	2	1		3		1	TA	1	
1456	2	0		3		1	T.		
1457	2	0		4		1	Go		
1458	1	0		2		1	Go		
1459	1	1		3		1	TA		
	TotRmsAbv	Grd Functi	onal	Fireplace	s Fireplac	ceQu	GarageType	e GarageYrBlt	\
0		8	Тур		0	NaN	Attch	d 2003.0	
1		6	Тур		1	TA	Attch	d 1976.0	
2		6	Тур		1	TA	Attch	d 2001.0	
3		7	Тур		1	Gd	Detch	d 1998.0	
4		9	Тур		1	TA	Attch	2000.0	
•••	•••	•••		•••	•••	•••	•••		
1455		7	Тур		1	TA	Attch	i 1999.0	
1456		7	Min1		2	TA	Attch	i 1978.0	
1457		9	Тур		2	Gd	Attch	i 1941.0	
1458		5	Тур		0	NaN	Attch	1950.0	
1459		6	Тур		0	NaN	Attch	i 1965.0	
	Cama w - E	-h (1	. O =	Comp A		7 (Nama mc (1 1	Daniel d Daniel \	
	GarageFini	sh Garage Fn	Cars 2	GarageAre 54	_	ıaı (TA	arageCond TA	PavedDrive \ Y	
0									
1		Fn En	2	46		TA TA	TA	Y	
2		Fn nf	2	60		TA	TA	Y	
3	U	nf	3	64	: ∠	TA	TA	Y	

4	RFn		3	83	6		TA	TA	Y	
 1455	 DE				 ^		 T 1		V	
1455 1456	RFn Unf		2 2	46 50			TA TA	TA TA	Y Y	
1457	RFn		1	25:			TA	TA	Y	
1458	Unf						TA	TA	Y	
1459	Fin		1 1	24 27			TA	TA	Y	
1405	1.111		1	21	O		IA	IA	1	
	WoodDeckSF	Open	PorchSF	EnclosedP	orch	3Ssn		ScreenPo		
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 Po	olQC	Fence M	fiscFeature	Mis	cVal	MoSold	YrSold	SaleType	\
0	0	NaN	NaN	NaN		0	2		WD	
1	0	NaN	NaN	NaN		0	5		WD	
2	0	NaN	NaN	NaN		0	9		WD	
3	0	NaN	NaN	NaN		0	2		WD	
4	0	NaN	NaN	NaN		0	12	2008	WD	
	•••	•••		• •••	•••	•••	•••			
1455	0	NaN	NaN	NaN		0	8	2007	WD	
1456	0	NaN	${\tt MnPrv}$	NaN		0	2	2010	WD	
1457	0	NaN	${\tt GdPrv}$	Shed		2500	5	2010	WD	
1458	0	NaN	NaN	NaN		0	4	2010	WD	
1459	0	NaN	NaN	NaN		0	6	2008	WD	
	SaleCondition	n Sa	lePrice	Alley_2	BsmtC	ond_2	BsmtE	[xposure_2	2 \	
0	Norma	1	208500	True		False		False	Э	
1	Norma	1	181500	True		False		False	Э	
2	Norma	1	223500	True		False		False	e	
3	Abnorm	1	140000	True		False		False	Э	
4	Norma	1	250000	True		False		False	е	
	••• NT	,					•••			
1455	Norma		175000	True		False		False		
1456	Norma		210000	True		False		False		
1457	Norma		266500	True		False		False		
1458	Norma		142125	True		False		False		
1459	Norma	Τ	147500	True		False		False	9	

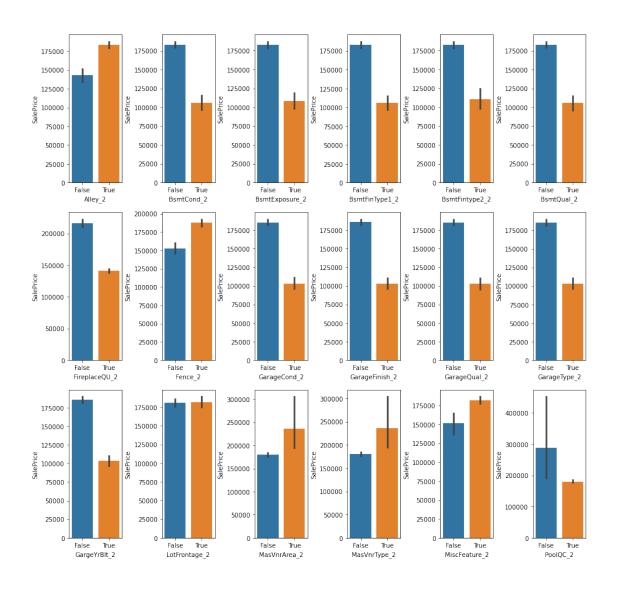
```
BsmtFinType1_2
                       BsmtFintype2_2
                                         BsmtQual_2 FireplaceQU_2
                                                                       Fence_2 \
0
                False
                                 False
                                 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

[1460 rows x 99 columns]

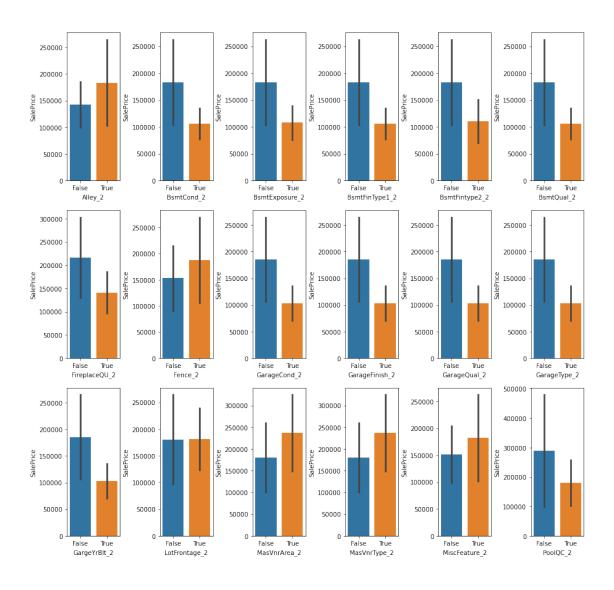
```
[7]: fig, axes = plt.subplots(3, 6, figsize=(15, 15))
fig.suptitle('Mean SalePrice of true and false values')

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)
```



```
sns.barplot(x =boston['BsmtFinType1_2'], y = boston['SalePrice'], ci='sd', u
 \rightarrowax=axes[0,3])
sns.barplot(x =boston['BsmtFintype2_2'], y = boston['SalePrice'], ci='sd', u
\rightarrowax=axes[0,4])
sns.barplot(x =boston['BsmtQual_2'], y = boston['SalePrice'], ci='sd',__
\rightarrowax=axes[0,5])
sns.barplot(x =boston['FireplaceQU_2'], y = boston['SalePrice'], ci='sd', u
\rightarrowax=axes[1,0])
sns.barplot(x =boston['Fence 2'], y = boston['SalePrice'], ci='sd', 
\rightarrowax=axes[1,1])
sns.barplot(x =boston['GarageCond_2'], y = boston['SalePrice'], ci='sd',__
\rightarrowax=axes[1,2])
sns.barplot(x =boston['GarageFinish_2'], y = boston['SalePrice'], ci='sd', u
\rightarrowax=axes[1,3])
sns.barplot(x =boston['GarageQual_2'], y = boston['SalePrice'], ci='sd', u
\rightarrowax=axes[1,4])
sns.barplot(x =boston['GarageType_2'], y = boston['SalePrice'], ci='sd',_
\rightarrowax=axes[1,5])
sns.barplot(x =boston['GargeYrBlt_2'], y = boston['SalePrice'], ci='sd', u
 \rightarrowax=axes[2,0])
sns.barplot(x =boston['LotFrontage_2'], y = boston['SalePrice'], ci='sd', u
\rightarrowax=axes[2,1])
sns.barplot(x =boston['MasVnrArea_2'], y = boston['SalePrice'], ci='sd',__
\rightarrowax=axes[2,2])
sns.barplot(x =boston['MasVnrType_2'], y = boston['SalePrice'], ci='sd', u
\rightarrowax=axes[2,3])
sns.barplot(x =boston['MiscFeature 2'], y = boston['SalePrice'], ci='sd', __
\rightarrowax=axes[2,4])
sns.barplot(x =boston['PoolQC 2'], y = boston['SalePrice'], ci='sd', ...
\rightarrowax=axes[2,5])
fig.subplots_adjust(wspace=0.75)
```



```
[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	${ t LotFrontage}$	${\tt LotArea}$	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	
	I andContour	Utilities L	otConfig	LandSlope	Neighbo	rhood Co	ondition1	\
0	Lvl	AllPub	Inside	Gt1	_	llgCr	Norm	`
1	Lvl	AllPub	FR2	Gt1		enker	Feedr	
2	Lvl	AllPub	Inside	Gtl		llgCr	Norm	
3	Lvl	AllPub	Corner	Gtl		awfor	Norm	
4	Lvl	AllPub	FR2	Gt1		Ridge	Norm	
-			1102		NO	ittage	WOIM	
 1455	 Lvl	AllPub	Inside	 Gtl	Gi.	 lbert	Norm	
1456	Lvl	AllPub	Inside	Gtl		WAmes	Norm	
1457	Lvl	AllPub	Inside	Gtl		awfor	Norm	
1458	Lvl	AllPub	Inside	Gtl		NAmes	Norm	
1459	Lvl	AllPub	Inside	Gtl		wards	Norm	
1100	211	11211 42	1110140	401	24	. ar ab	1101111	
	Condition2	BldgType Hou	seStyle	OverallQua	l Over	allCond	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	3
1457	Norm	1Fam	2Story		7	9	1941	
1458	Norm	1Fam	1Story		5	6	1950)
1459	Norm	1Fam	1Story		5	6	1965	•
	YearRemodA	dd RoofStyle	RoofMat1	Exterior1	st Exte	rior2nd	MasVnrTvpe	. \
0	20	•	CompShg			VinylSd	BrkFace	
1	19			•		MetalSd	None	
2	20		1 0			VinylSd		
3	19			•		Wd Shng	None	
4	20				J	VinylSd	BrkFace	
-				, · <i>y</i>	•••			
1455	20	00 Gable	 CompShg	 Vinyl	Sd '	WinylSd	None	<u>:</u>
1456	19		1 0	•		Plywood	Stone	
1457	20		1 0	•		CmentBd	None	
1458	19			,		MetalSd	None	
1459	19	-				HdBoard	None	
				,				

	MasVnrArea	ExterQual	${\tt ExterCond}$	${\tt Foundation}$	${\tt BsmtQual}$	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	${\tt BrkTil}$	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		
	BsmtExposur	•	_			BsmtFinSF2	\	
0	N		GLQ	706	Unf	0		
1	G		ALQ	978	Unf	0		
2	M:		GLQ	486	Unf	0		
3	N		ALQ	216	Unf	0		
4	A·	V	GLQ	655	Unf	0		
 1455	 N	···	Unf	0	 Unf	0		
1456	N		ALQ	790	Rec	163		
1457	N		GLQ	275	Unf	0		
1458	M:		GLQ	49	Rec	1029		
1459	N		BLQ	830	LwQ	290		
1400	141	O	סבע	000	тwQ	250		
	BsmtUnfSF	TotalBsmtS	SF Heating	HeatingQC (CentralAi:	r Electrical	L \	
0	150	85	_	Ex		Y SBrkı		
1	284	126		Ex		Y SBrkı		
2	434		20 GasA	Ex		Y SBrkı		
3	540		56 GasA	Gd		Y SBrkı		
4	490	114		Ex		Y SBrkı		
	•••			•••	•••			
1455	953	95	GasA	Ex	,	Y SBrkı	<u>-</u>	
1456	589	154	12 GasA	TA	,	Y SBrkı	2	
1457	877	115	52 GasA	Ex	,	Y SBrkı	2	
1458	0	107	78 GasA	Gd	,	Y Fuse	I	
1459	136	125	56 GasA	Gd	,	Y SBrkı	2	
			LowQualFinS			ullBath Bsr	ntHalfBath	\
0	856	854		0 17:		1	0	
1	1262	0		0 126		0	1	
2	920	866		0 178		1	0	
3	961	756		0 17:		1	0	
4	1145	1053		0 219	98	1	0	
•••	•••	•••	•••	•••	•••	•••		
1455	953	694		0 164	17	0	0	

```
1456
           2073
                                          0
                                                   2073
                                                                                      0
                          0
                                                                       1
1457
           1188
                      1152
                                          0
                                                   2340
                                                                       0
                                                                                       0
                                                                                       0
1458
           1078
                          0
                                          0
                                                   1078
                                                                       1
                          0
1459
           1256
                                                   1256
      FullBath
                 HalfBath
                             {\tt 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
                                                                       Gd
                          1
                                          4
1455
              2
                          1
                                          3
                                                          1
                                                                       TA
1456
              2
                          0
                                          3
                                                                       TA
                                                          1
1457
              2
                          0
                                          4
                                                          1
                                                                       Gd
1458
               1
                          0
                                          2
                                                                       Gd
                                                          1
                                          3
1459
               1
                          1
                                                          1
                                                                       TA
      TotRmsAbvGrd Functional Fireplaces GarageType
                                                             GarageYrBlt
                   8
                                                                   2003.0
0
                                             0
                                                    Attchd
                             Тур
1
                   6
                             Тур
                                             1
                                                    Attchd
                                                                   1976.0
2
                   6
                             Тур
                                             1
                                                    Attchd
                                                                   2001.0
3
                   7
                             Тур
                                             1
                                                    Detchd
                                                                   1998.0
4
                   9
                                                    Attchd
                                                                   2000.0
                             Тур
                                             1
                   7
1455
                             Тур
                                             1
                                                    Attchd
                                                                   1999.0
1456
                   7
                            Min1
                                                    Attchd
                                                                   1978.0
                                             2
1457
                   9
                                             2
                                                    Attchd
                                                                   1941.0
                             Тур
1458
                   5
                                             0
                             Тур
                                                    Attchd
                                                                   1950.0
1459
                   6
                                             0
                                                    Attchd
                                                                   1965.0
                             Тур
                                   GarageArea GarageQual GarageCond PavedDrive
     GarageFinish GarageCars
               RFn
                                           548
0
                               2
                                                         TA
                                                                     TA
                                                                                   Y
                RFn
                                2
                                                         TA
                                                                     TA
                                                                                   Y
1
                                           460
                                2
2
               RFn
                                           608
                                                         TA
                                                                     TA
                                                                                   Y
3
                Unf
                               3
                                           642
                                                         TA
                                                                     TΑ
                                                                                   Y
4
                               3
                                                                                   γ
               RFn
                                           836
                                                         TA
                                                                     TA
1455
                               2
                                                         TA
                                                                                   Y
                RFn
                                           460
                                                                     TA
1456
               Unf
                               2
                                           500
                                                         TA
                                                                     TA
                                                                                   Y
1457
               RFn
                                1
                                           252
                                                         TA
                                                                     TA
                                                                                   Y
1458
                                                         TA
                                                                                   Y
               Unf
                                1
                                           240
                                                                     TA
1459
               Fin
                                           276
                                                         TA
                                                                     TΑ
                                   EnclosedPorch
      WoodDeckSF
                    OpenPorchSF
                                                    3SsnPorch ScreenPorch
0
                 0
                                                 0
                                                             0
                              61
                                                                            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	${ t MiscVal}$	MoSold	YrSold	SaleType	${\tt SaleCondition}$	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 int64SaleType 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)

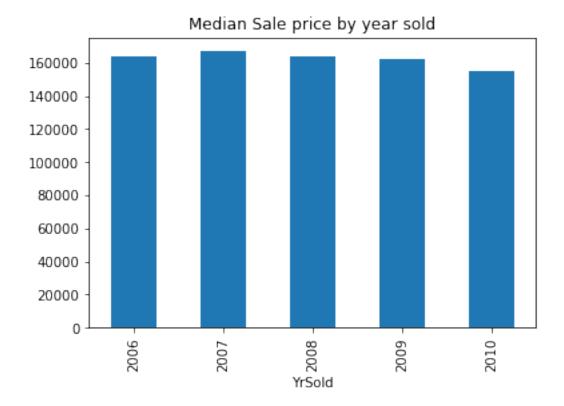
[12]: Median_Saleprice_by_YrSold = Dropped_na_boston.groupby('YrSold')['SalePrice'].

→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
- [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
```

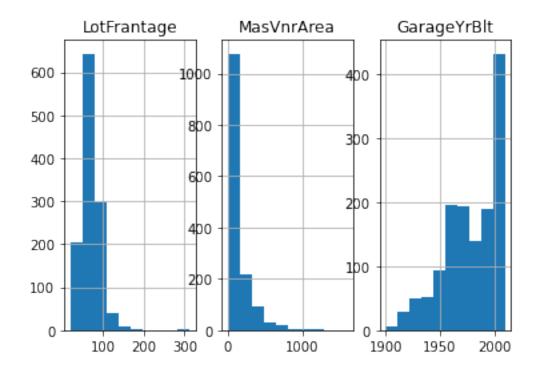
[14]: # There are 3 continuous variables in the dataset (LotFrontage, MasVnrArea, \Box \hookrightarrow GarageYrBlt)

```
[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()
```

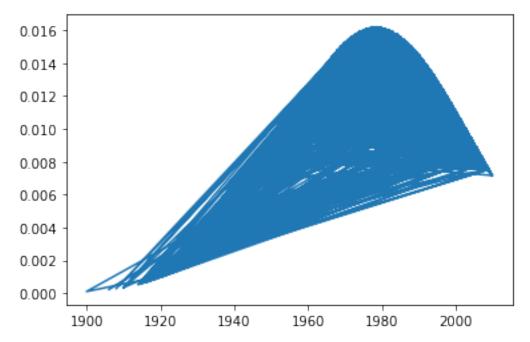
[15]: <AxesSubplot:title={'center':'GarageYrBlt'}>



```
[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

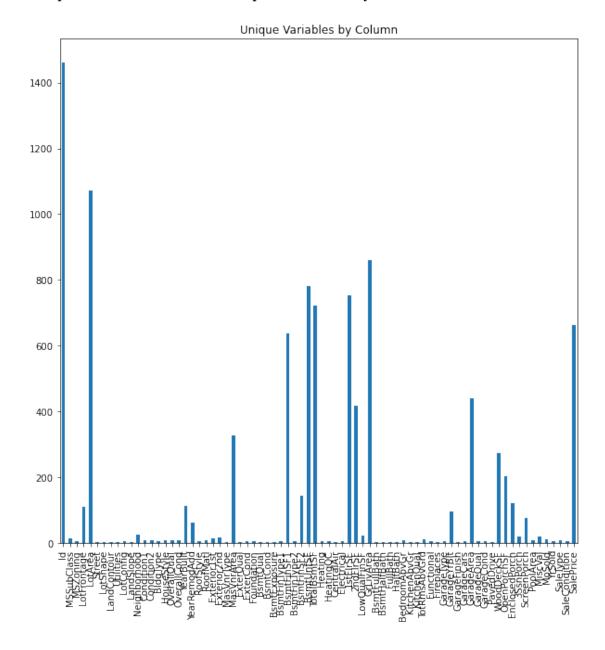
LotArea 1073
```

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'}>



```
→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, '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'] = Dropped na boston['HeatingQC'].map({'Po':
       →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]:
                  MSSubClass MSZoning LotFrontage LotArea Street LotShape \
              Ιd
      0
               1
                           60
                                    RL
                                               65.0
                                                         8450
                                                                Pave
                                                                          Reg
      1
               2
                                                         9600
                           20
                                    RL
                                               80.0
                                                                Pave
                                                                          Reg
      2
               3
                           60
                                    RL
                                               68.0
                                                        11250
                                                                Pave
                                                                          IR1
      3
               4
                           70
                                    RL
                                               60.0
                                                         9550
                                                                Pave
                                                                          IR1
      4
               5
                                    RL
                                               84.0
                           60
                                                        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
                           20
                                    RL
            1459
                                               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
                           AllPub
                                     Inside
                                                  Gtl
                                                            Crawfor
                   Lvl
                                                                          Norm
      1458
                   Lvl
                           AllPub
                                     Inside
                                                  Gtl
                                                              NAmes
                                                                          Norm
```

[21]: Dropped_na_boston['ExterQual_map'] = Dropped_na_boston['ExterQual'].map({'Po':__

1459	Lvl	AllPub	Inside	Gtl	Edwards	Norm	
	Condition2 Blo	dgType Hous	eStyle	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		2Story	7	5	1915	
4	Norm		2Story	8	5		
-							
1455	Norm	1Fam	2Story	6		1999	
1456		1Fam	1Story	6	6		
1457			•	7	g		
			2Story	5			
1458		1Fam	1Story		6		
1459	Norm	1Fam	1Story	5	6	1965	
	YearRemodAdd	RoofStyle	RoofMatl	Exterior1s	t Exterior2nd	MasVnrType	\
0	2003	Gable	CompShg	VinylS	d VinylSd	BrkFace	
1	1976	Gable	CompShg	•			
2	2002	Gable	CompShg				
3	1970	Gable	CompShg	•	•		
4	2000	Gable	CompShg				
···	2000	dabie	Company	, VIIIYID	vinyibo	DINIACE	
 1455	2000	Gable	CompShg	WinylS		None	
1456		Gable	CompShg	-	· ·		
1457		Gable	CompShg	•	•		
1458		Hip	CompShg				
1459	1965	Gable	CompShg	HdBoar	d HdBoard	None	
	MasVnrArea E	vtorOusl Ev	terCond	Foundation 1	Ram+Ousl Ram+	Cond \	
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	
 1/FE			т.		 	т л	
1455		TA	TA	PConc	Gd	TA	
1456		TA —	TA	CBlock	Gd 	TA	
1457		Ex	Gd	Stone	TA	Gd	
1458		TA	TA	CBlock	TA	TA	
1459	0.0	Gd	TA	CBlock	TA	TA	
	BsmtExposure I	BsmtFinTvpe	1 BsmtF	inSF1 BsmtF	inTvpe2 Bsmt	FinSF2 \	
0	No	GL		706	Unf	0	
1	Gd	AL	-	978	Unf	0	
2	Mn	GL		486	Unf	0	
3							
	No A	AL		216	Unf	0	
4	Av	GL	·Ų	655	Unf	0	

```
0
1455
                                 Unf
                                                  0
                                                                Unf
                  No
1456
                  No
                                 ALQ
                                               790
                                                                Rec
                                                                               163
                                                275
1457
                  No
                                 GLQ
                                                                Unf
                                                                                 0
1458
                                 GLQ
                                                 49
                                                                             1029
                  Mn
                                                                Rec
1459
                                 BLQ
                                               830
                                                                               290
                  No
                                                               LwQ
                    TotalBsmtSF Heating HeatingQC CentralAir Electrical
       {\tt BsmtUnfSF}
                                                                    Y
                                                                             SBrkr
0
                              856
                                       GasA
                                                     Ex
              150
1
              284
                             1262
                                       GasA
                                                      Ex
                                                                    Y
                                                                             SBrkr
2
                                                                    Y
              434
                              920
                                       GasA
                                                      Ex
                                                                             SBrkr
3
              540
                              756
                                       GasA
                                                      Gd
                                                                    Y
                                                                             SBrkr
4
              490
                             1145
                                       GasA
                                                      Ex
                                                                    Y
                                                                             SBrkr
                                                      •••
1455
              953
                              953
                                                                    Y
                                                                             SBrkr
                                       GasA
                                                      Ex
                                                                    Y
1456
              589
                             1542
                                       GasA
                                                      TΑ
                                                                             SBrkr
1457
                                                                    Y
              877
                             1152
                                       GasA
                                                      Ex
                                                                             SBrkr
1458
                             1078
                                                      Gd
                                                                     Y
                                                                             FuseA
                 0
                                       GasA
1459
                             1256
              136
                                       GasA
                                                      Gd
                                                                             SBrkr
       1stFlrSF
                   2ndFlrSF
                               {\tt LowQualFinSF}
                                                 {\tt GrLivArea}
                                                              {\tt BsmtFullBath}
                                                                                BsmtHalfBath
0
             856
                         854
                                             0
                                                       1710
                                                                            1
                                                                                             0
                                                       1262
1
            1262
                            0
                                             0
                                                                            0
                                                                                             1
2
             920
                         866
                                             0
                                                       1786
                                                                                             0
3
             961
                         756
                                                       1717
                                                                                             0
4
            1145
                        1053
                                             0
                                                       2198
1455
             953
                         694
                                             0
                                                       1647
                                                                            0
                                                                                             0
1456
            2073
                            0
                                             0
                                                       2073
                                                                            1
                                                                                             0
1457
            1188
                                             0
                                                       2340
                                                                            0
                                                                                             0
                        1152
1458
            1078
                            0
                                             0
                                                       1078
                                                                                             0
                            0
1459
            1256
                                                       1256
                   HalfBath
       FullBath
                               {\tt BedroomAbvGr}
                                                 KitchenAbvGr KitchenQual
0
               2
                            1
                                             3
                                                              1
                                                                            Gd
1
               2
                            0
                                             3
                                                              1
                                                                            TA
2
               2
                            1
                                             3
                                                                            Gd
                                                              1
3
               1
                            0
                                             3
                                                                            Gd
4
               2
                            1
                                             4
                                                                            Gd
               2
                                             3
                                                                            TA
1455
                            1
               2
                            0
                                                                            TA
1456
                                             3
                2
1457
                            0
                                             4
                                                                            Gd
                                                              1
1458
                1
                            0
                                             2
                                                              1
                                                                            Gd
1459
                1
                            1
                                             3
                                                                            TA
                                                              1
```

TotRmsAbvGrd Functional Fireplaces GarageType GarageYrBlt \

0		8		Тур		0	Attchd	2003	.0	
1		6		Тур		1	Attchd			
2		6		Тур		1	Attchd			
3		7		Тур		1	Detchd			
4		9		Тур		1	Attchd	2000	.0	
•••	•••		•••		•••	•••		•••		
145	5	7		Тур		1	Attchd	1999	. 0	
145		7	,	1)P Min1		2	Attchd			
			1							
145		9		Тур		2	Attchd			
145		5		Тур		0	Attchd			
145	9	6		Тур		0	Attchd	1965	.0	
	GarageFin	ish	Garage	Cars G	arageAre	a Gara	ageQual	GarageCond	PavedDrive	e \
0	_	RFn		2	54		TA	_		
1		RFn		2	46		TA			
2		RFn		2	60		TA			
3		Unf		3	64	2	TA	TA	Y	<i>[</i>
4		RFn		3	83	6	TA	TA	Y	7
145	5	RFn		2	46	in	TA	TA	Y	7
				2						
145		Unf			50		TA			
145		RFn		1	25		TA			
145	8	Unf		1	24	:0	TA	TA	Y	<i>[</i>
145	9	Fin		1	27	6	TA	TA	Y	7
	WoodDeck	SF O	penPor	hSF F	nclosedP	orch	3SsnPo:	rch Screen	Porch \	
0	WOO GEDOOL	0	Pom or	61		0	ODDIII O	0	0	
1	2	98		0		0		0	0	
2		0		42		0		0	0	
3		0		35		272		0	0	
4	1	.92		84		0		0	0	
			•••							
145		0	•••	40	•••	0		0	0	
145		349		0		0		0	0	
145	7	0		60		0		0	0	
145	8 3	866		0		112		0	0	
145	9 7	'36		68		0		0	0	
	PoolArea	Mic	cVal 1	MoSold	VrSold	TaleP	ma Sal	eCondition	SalePrice	\
^						Sare!	_			`
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	
								•••		
 145	5 0	····	0	 8	 2007		wD	 Normal	175000	
145	6 0)	0	2	2010		WD	Normal	210000	

```
1457
                      0
                             2500
                                          5
                                               2010
                                                                        Normal
                                                                                     266500
                                                             WD
       1458
                      0
                                0
                                          4
                                                2010
                                                             WD
                                                                        Normal
                                                                                     142125
       1459
                      0
                                0
                                          6
                                                2008
                                                                        Normal
                                                             WD
                                                                                     147500
                               ExterCond_map
             ExterQual_map
                                                BsmtQual_map
                                                                 BsmtCond_map
       0
                                             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
                                             3
       1456
                           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
                                   HeatingQC_map
             BsmtExposure_map
                                                    KitchenQual_map
                                                                        GarageQual_map
       0
                             {\tt NaN}
                                                                                     3.0
                                                 5
                                                                     3
       1
                             4.0
                                                                                     3.0
                                                 5
       2
                             {\tt NaN}
                                                                     4
                                                                                     3.0
       3
                                                 4
                                                                     4
                             NaN
                                                                                     3.0
       4
                             {\tt NaN}
                                                 5
                                                                     4
                                                                                     3.0
                                                                     3
                                                                                     3.0
       1455
                             NaN
                                                 5
                                                 3
                                                                     3
       1456
                                                                                     3.0
                             NaN
                                                 5
       1457
                                                                     4
                                                                                     3.0
                             NaN
       1458
                             NaN
                                                 4
                                                                     4
                                                                                     3.0
       1459
                             {\tt NaN}
                                                                     3
                                                                                     3.0
             GarageCond_map
       0
                          3.0
       1
                          3.0
       2
                           3.0
       3
                           3.0
       4
                           3.0
                          3.0
       1455
       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)
```

