

## QUESTION – 2

	Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	Alley	LotShape	LandContour	Utilities	...	MasVnrArea	ExterQual	ExterCond	Foundation	Bsr
0	1	60	RL	65.0	8450	Pave	1	Reg	Lvl	AllPub	...	196	Gd	TA	PConc	
1	2	20	RL	80.0	9600	Pave	1	Reg	Lvl	AllPub	...	0	TA	TA	CBlock	
2	3	60	RL	68.0	11250	Pave	1	IR1	Lvl	AllPub	...	162	Gd	TA	PConc	
3	4	70	RL	60.0	9550	Pave	1	IR1	Lvl	AllPub	...	0	TA	TA	BrkTil	
4	5	60	RL	84.0	14260	Pave	1	IR1	Lvl	AllPub	...	350	Gd	TA	PConc	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
94	95	60	RL	69.0	9337	Pave	1	IR1	Lvl	AllPub	...	0	TA	Gd	PConc	
95	96	60	RL	1.0	9765	Pave	1	IR2	Lvl	AllPub	...	68	Ex	Gd	PConc	
96	97	20	RL	78.0	10264	Pave	1	IR1	Lvl	AllPub	...	183	Gd	TA	PConc	
97	98	20	RL	73.0	10921	Pave	1	Reg	HLS	AllPub	...	48	TA	TA	CBlock	
98	99	30	RL	85.0	10625	Pave	1	Reg	Lvl	AllPub	...	0	TA	TA	BrkTil	

99 rows x 36 columns

```

In [13]: #checking how many null values are there
df2.isna().sum().sum()

Out[13]: 0

```

## QUESTION – 3

```

# Count each word in the file
for j in range(i+1, len(words)):
    if(words[i] == words[j]):
        count = count + 1;

# If the count value is more
# than highest frequency then
if(count > frequency):
    frequency = count;
    frequent_word = words[i];

print("Most frequently used word: " + frequent_word)
print("Frequency: " + str(frequency))
file.close();

```

Most frequently used word: python  
Frequency: 4

## QUESTION – 4

```
In [43]: df.sort_values(by='rating')
```

```
Out[43]:
```

	ProductID	price	rating
42	B005P0HHGK	170	0.10
96	B000JEHAHS	133	0.12
11	B0001PB9FY	183	0.19
18	B005DUM9UQ	138	0.24
15	B002GWHC0G	165	0.26
...	...	...	...
36	B002SRYRE8	189	4.67
41	B00374XSVY	120	4.70
60	B002TDK0VK	196	4.80
40	B0037ZFEW4	117	4.85
89	B003YDP5PA	169	4.98

99 rows × 3 columns