In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings('ignore')
```

In [2]:
```
import os
os.chdir("C:\\Users\\jayes\Downloads")
```

In [3]:
```
data = pd.read_csv('dataset.csv')
data
```

Out[3]:

	brand	model	processor_brand	processor_name	processor_gnrtn	ram_gb	ram_type	ssd	hdd	os	...	display_size	warranty	Touchscreen	msoffice	latest_price	old_price	discount	star_rating	ratings	re
0	Lenovo	Ali-5225	AMD	Ali-5225 Processor	10th Gen	4 GB	DDR4	0 GB	1024 GB	Windows	...	Missing	0	No	No	24990	32790	23	3.7	63	
1	Lenovo	Ryzen 5	AMD	APU Dual Core	10th Gen	4 GB	DDR4	0 GB	512 GB	Windows	...	Missing	0	No	No	19590	21325	8	3.6	1894	
2	Avita	PURA	AMD	APU Dual Core	10th Gen	4 GB	DDR4	128 GB	0 GB	Windows	...	Missing	0	No	No	19990	27990	28	3.7	1153	
3	Avita	PURA	AMD	APU Dual Core	10th Gen	8 GB	DDR4	128 GB	0 GB	Windows	...	Missing	0	No	No	21490	27990	23	3.7	1153	
4	Avita	PURA	AMD	APU Dual Core	10th Gen	8 GB	DDR4	256 GB	0 GB	Windows	...	Missing	0	No	No	24490	33490	25	3.7	1857	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
891	ASUS	ROG	AMD	Ryzen 9	Missing	8 GB	DDR4	1024 GB	0 GB	Windows	...	16	1	No	No	144990	194990	25	0.0	0	
892	ASUS	Ryzen	AMD	Ryzen 9	Missing	8 GB	DDR4	1024 GB	0 GB	Windows	...	Missing	1	No	No	149990	0	0	0.0	0	
893	ASUS	Ryzen	AMD	Ryzen 9	Missing	8 GB	DDR4	1024 GB	0 GB	Windows	...	16	1	No	No	142990	0	0	0.0	0	
894	SAMSUNG	Galaxy	Qualcomm	Snapdragon 7c	Missing	4 GB	LPDDR4X	0 GB	512 GB	Windows	...	Missing	0	No	No	38990	47990	18	0.0	0	
895	Lenovo	Thinkpad	AMD	Ryzen 5	10th Gen	8 GB	DDR4	512 GB	0 GB	DOS	...	14	0	No	No	57490	78490	26	4.2	18	

896 rows × 23 columns

In [4]:
```
data.head()
```

Out[4]:

	brand	model	processor_brand	processor_name	processor_gnrtn	ram_gb	ram_type	ssd	hdd	os	...	display_size	warranty	Touchscreen	msoffice	latest_price	old_price	discount	star_rating	ratings	reviews
0	Lenovo	Ali-5225	AMD	Ali-5225 Processor	10th Gen	4 GB	DDR4	0 GB	1024 GB	Windows	...	Missing	0	No	No	24990	32790	23	3.7	63	12
1	Lenovo	Ideapad	AMD	APU Dual Core	10th Gen	4 GB	DDR4	0 GB	512 GB	Windows	...	Missing	0	No	No	19590	21325	8	3.6	1894	256
2	Avita	PURA	AMD	APU Dual Core	10th Gen	4 GB	DDR4	128 GB	0 GB	Windows	...	Missing	0	No	No	19990	27990	28	3.7	1153	159
3	Avita	PURA	AMD	APU Dual Core	10th Gen	8 GB	DDR4	128 GB	0 GB	Windows	...	Missing	0	No	No	21490	27990	23	3.7	1153	159
4	Avita	PURA	AMD	APU Dual Core	10th Gen	8 GB	DDR4	256 GB	0 GB	Windows	...	Missing	0	No	No	24490	33490	25	3.7	1857	254

5 rows × 23 columns

In [5]:
```
data.tail()
```

Out[5]:

	brand	model	processor_brand	processor_name	processor_gnrtn	ram_gb	ram_type	ssd	hdd	os	...	display_size	warranty	Touchscreen	msoffice	latest_price	old_price	discount	star_rating	ratings	reviews
891	ASUS	ROG	AMD	Ryzen 9	Missing	8 GB	DDR4	1024 GB	0 GB	Windows	...	16	1	No	No	144990	194990	25	0.0	0	0
892	ASUS	Ryzen	AMD	Ryzen 9	Missing	8 GB	DDR4	1024 GB	0 GB	Windows	...	Missing	1	No	No	149990	0	0	0.0	0	0
893	ASUS	Ryzen	AMD	Ryzen 9	Missing	8 GB	DDR4	1024 GB	0 GB	Windows	...	16	1	No	No	142990	0	0	0.0	0	0
894	SAMSUNG	Galaxy	Qualcomm	Snapdragon 7c	Missing	4 GB	LPDDR4X	0 GB	512 GB	Windows	...	Missing	0	No	No	38990	47990	18	0.0	0	0
895	Lenovo	Thinkpad	AMD	Ryzen 5	10th Gen	8 GB	DDR4	512 GB	0 GB	DOS	...	14	0	No	No	57490	78490	26	4.2	18	0

5 rows × 23 columns

In [6]:
```
data.describe()
```

Out[6]:

	graphic_card_gb	warranty	latest_price	old_price	discount	star_rating	ratings	reviews
count	896.000000	896.000000	896.000000	896.000000	896.000000	896.000000	896.000000	896.000000
mean	1.198661	0.693084	76309.865401	89334.154018	16.527902	2.980469	367.391741	48.152902
std	2.307454	0.936292	49013.354368	50719.645554	10.504046	1.965254	1166.309305	136.070906
min	0.000000	0.000000	13990.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	40490.000000	54940.500000	11.000000	0.000000	0.000000	0.000000
50%	0.000000	1.000000	61490.000000	78052.500000	18.000000	4.100000	15.000000	3.000000
75%	2.000000	1.000000	99090.000000	111019.500000	26.000000	4.400000	178.500000	23.250000
max	8.000000	3.000000	441990.000000	377799.000000	57.000000	5.000000	15279.000000	1947.000000

In [7]:
```
data.dtypes
```

Out[7]:
```
brand              object
model              object
processor_brand    object
processor_name     object
processor_gnrtn    object
ram_gb             object
ram_type           object
ssd                object
hdd                object
os                 object
os_bit             object
graphic_card_gb    int64
weight             object
display_size       object
warranty           int64
Touchscreen        object
msoffice           object
latest_price       int64
old_price          int64
discount           int64
star_rating        float64
ratings            int64
reviews            int64
dtype: object
```

In [8]:
```
data.shape
```

Out[8]:
```
(896, 23)
```

In [9]:
```
data.head(1)
```

Out[9]:

	brand	model	processor_brand	processor_name	processor_gnrtn	ram_gb	ram_type	ssd	hdd	os	...	display_size	warranty	Touchscreen	msoffice	latest_price	old_price	discount	star_rating	ratings	reviews
0	Lenovo	Ali-5225	AMD	Ali-5225 Processor	10th Gen	4 GB	DDR4	0 GB	1024 GB	Windows	...	Missing	0	No	No	24990	32790	23	3.7	63	12

1 rows × 23 columns

In [10]:
```
data.columns
```

Out[10]:
```
Index(['brand', 'model', 'processor_brand', 'processor_name',
       'processor_gnrtn', 'ram_gb', 'ram_type', 'ssd', 'hdd', 'os', 'os_bit',
       'graphic_card_gb', 'weight', 'display_size', 'warranty', 'Touchscreen',
       'msoffice', 'latest_price', 'old_price', 'discount', 'star_rating',
       'ratings', 'reviews'],
      dtype='object')
```

In [11]:
```
data.index
```

Out[11]:
```
RangeIndex(start=0, stop=896, step=1)
```

In [12]:
```
data.describe(include='all')
```

Out[12]:

	brand	model	processor_brand	processor_name	processor_gnrtn	ram_gb	ram_type	ssd	hdd	os	...	display_size	warranty	Touchscreen	msoffice	latest_price	old_price	discount	star_rating	ratings	reviews
count	896	896	896	896	896	896	896	896	896	896	...	896	896.000000	896	896	896.000000	896.000000	896.000000	896.000000	896.000000	896.000000
unique	21	117	5	28	8	4	6	8	4	3	...	16	NaN	2	2	NaN	NaN	NaN	NaN	NaN	NaN
top	ASUS	Missing	AMD	Ryzen 5	10th Gen	8 GB	DDR4	512 GB	0 GB	Windows	...	Missing	NaN	No	No	NaN	NaN	NaN	NaN	NaN	NaN
freq	254	85	680	312	346	454	760	417	896	832	...	332	NaN	793	836	NaN	NaN	NaN	NaN	NaN	NaN
mean	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	0.693084	NaN	NaN	76309.865401	89334.154018	16.527902	2.980469	367.391741	48.152902
std	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	0.936292	NaN	NaN	49013.354368	50719.645554	10.504046	1.965254	1166.309305	136.070906
min	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	0.000000	NaN	NaN	13990.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	0.000000	NaN	NaN	40490.000000	54940.500000	11.000000	0.000000	0.000000	0.000000
50%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	1.000000	NaN	NaN	61490.000000	78052.500000	18.000000	4.100000	15.000000	3.000000
75%	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	1.000000	NaN	NaN	99090.000000	111019.500000	26.000000	4.400000	178.500000	23.250000
max	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	3.000000	NaN	NaN	441990.000000	377799.000000	57.000000	5.000000	15279.000000	1947.000000

11 rows × 23 columns

In [13]:
```
data['ratings']
```

Out[13]:
```
0      63
1    1894
2    1153
3    1153
4    1857
     ...
891     0
892     0
893     0
894     0
895    18
Name: ratings, Length: 896, dtype: int64
```

In [14]:
```
data['brand']
```

Out[14]:
```
0      Lenovo
1      Lenovo
2       Avita
3       Avita
4       Avita
        ...
891      ASUS
892      ASUS
893      ASUS
894    SAMSUNG
895     Lenovo
Name: brand, Length: 896, dtype: object
```

In [15]:
```
data['ram_gb'].max()
```

Out[15]:
```
'8 GB'
```

In [16]:
```
data['graphic_card_gb'].max()
```

Out[16]:
```
8
```

In [19]:
```
data.isnull().sum()
```

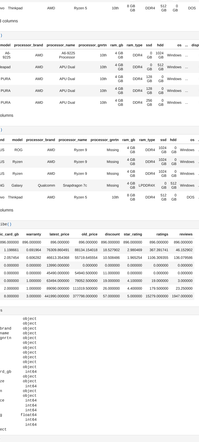
Out[19]:
```
brand             0
model             0
processor_brand   0
processor_name    0
processor_gnrtn   0
ram_gb            0
ram_type          0
ssd               0
hdd               0
os                0
os_bit            0
graphic_card_gb   0
weight            0
display_size      0
warranty          0
Touchscreen       0
msoffice          0
latest_price      0
old_price         0
discount          0
star_rating       0
ratings           0
reviews           0
dtype: int64
```

In [20]:
```
data.info()
```

Out[20]:
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 896 entries, 0 to 895
Data columns (total 23 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   brand            896 non-null    object
 1   model            896 non-null    object
 2   processor_brand  896 non-null    object
 3   processor_name   896 non-null    object
 4   processor_gnrtn  896 non-null    object
 5   ram_gb           896 non-null    object
 6   ram_type         896 non-null    object
 7   ssd              896 non-null    object
 8   hdd              896 non-null    object
 9   os               896 non-null    object
 10  os_bit           896 non-null    object
 11  graphic_card_gb  896 non-null    int64
 12  weight           896 non-null    object
 13  display_size     896 non-null    object
 14  warranty         896 non-null    int64
 15  Touchscreen      896 non-null    object
 16  msoffice         896 non-null    object
 17  latest_price     896 non-null    int64
 18  old_price        896 non-null    int64
 19  discount         896 non-null    int64
 20  star_rating      896 non-null    float64
 21  ratings          896 non-null    int64
 22  reviews          896 non-null    int64
dtypes: float64(1), int64(7), object(15)
memory usage: 161.1+ KB
```

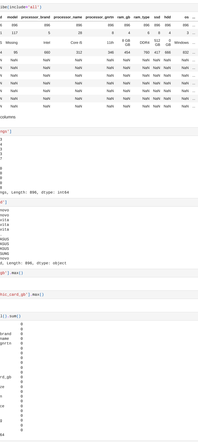
In [21]:
```
data.plot(x='old_price',y='graphic_card_gb',kind='scatter')
```

Out[21]:
```
<Axes: xlabel='old_price', ylabel='graphic_card_gb'>
```

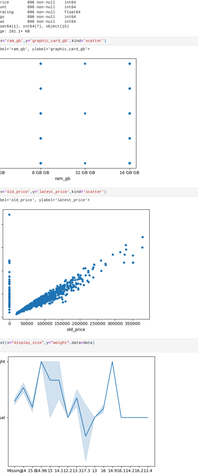


In [22]:
```
data.plot(x='old_price',y='latest_price',kind='scatter')
```

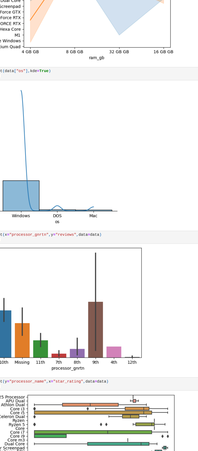
Out[22]:
```
<Axes: xlabel='old_price', ylabel='latest_price'>
```



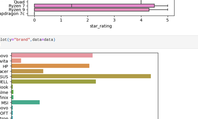
In [23]:
```
sns.lineplot(x="display_size",y="weight",data=data)
plt.show()
```



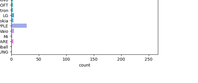
In [24]:
```
sns.lineplot(x="ram_gb",y="processor_name",data=data,hue="Touchscreen")
plt.show()
```



In [25]:
```
sns.displot(data['os'],kde=True)
plt.show()
```



In [26]:
```
sns.barplot(x='processor_gnrtn',y='reviews',data=data)
plt.show()
```



In [27]:
```
sns.boxplot(y="processor_name",x="star_rating",data=data)
plt.show()
```


In [28]:
```
sns.countplot(y="brand",data=data)
plt.show()
```