```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
df=pd.read_csv('train.csv')
```

```
df.head()
```

	battery_power	blue	clock_speed	dual_sim	fc	four_g	int_memory	m_dep	mobile_
0	842	0	2.2	0	1	0	7	0.6	1
1	1021	1	0.5	1	0	1	53	0.7	1
2	563	1	0.5	1	2	1	41	0.9	1
3	615	1	2.5	0	0	0	10	0.8	1
4	1821	1	1.2	0	13	1	44	0.6	1

```
df.shape
```

```
(2000, 21)
```

```
df.describe()
```

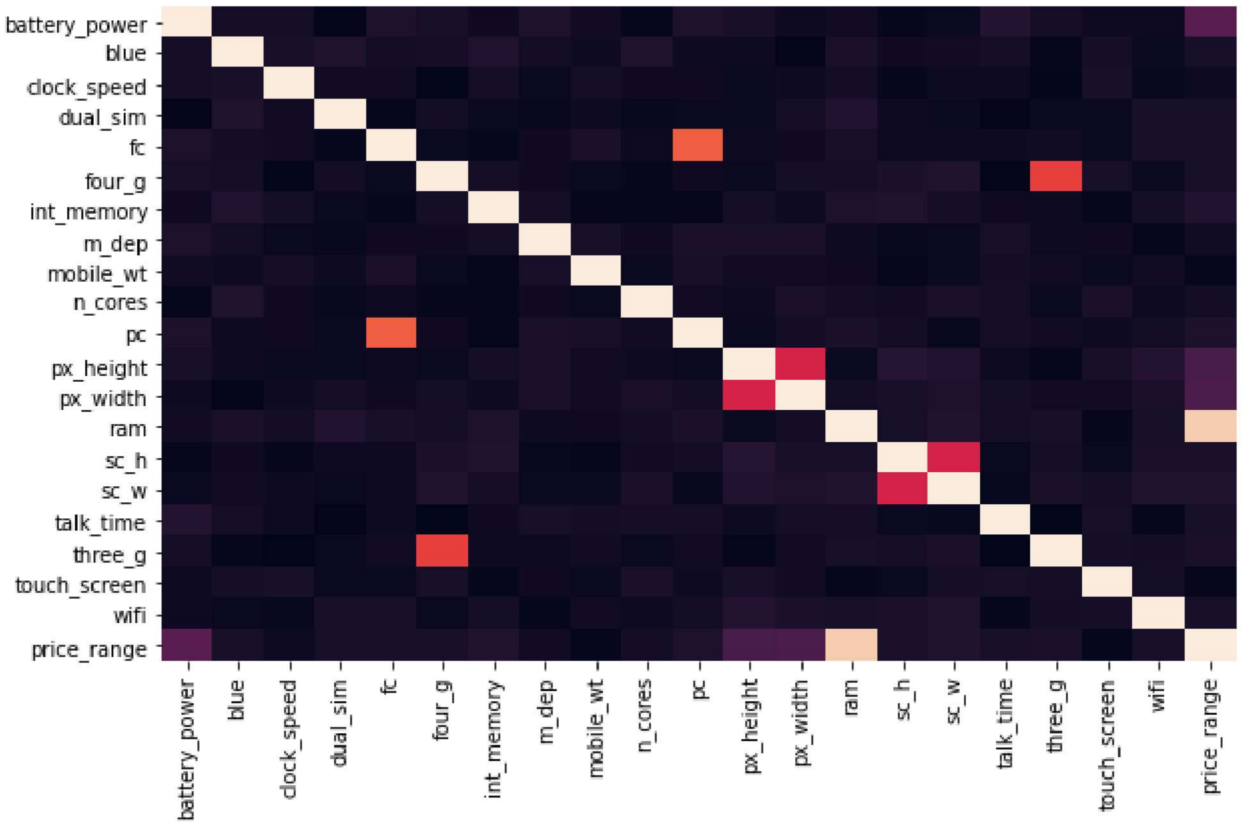
	battery_power	blue	clock_speed	dual_sim	fc	four_g	
count	2000.000000	2000.0000	2000.000000	2000.000000	2000.000000	2000.000000	2
mean	1238.518500	0.4950	1.522250	0.509500	4.309500	0.521500	
std	439.418206	0.5001	0.816004	0.500035	4.341444	0.499662	
min	501.000000	0.0000	0.500000	0.000000	0.000000	0.000000	
25%	851.750000	0.0000	0.700000	0.000000	1.000000	0.000000	
50%	1226.000000	0.0000	1.500000	1.000000	3.000000	1.000000	
75%	1615.250000	1.0000	2.200000	1.000000	7.000000	1.000000	
max	1998.000000	1.0000	3.000000	1.000000	19.000000	1.000000	

```
df.info()
```

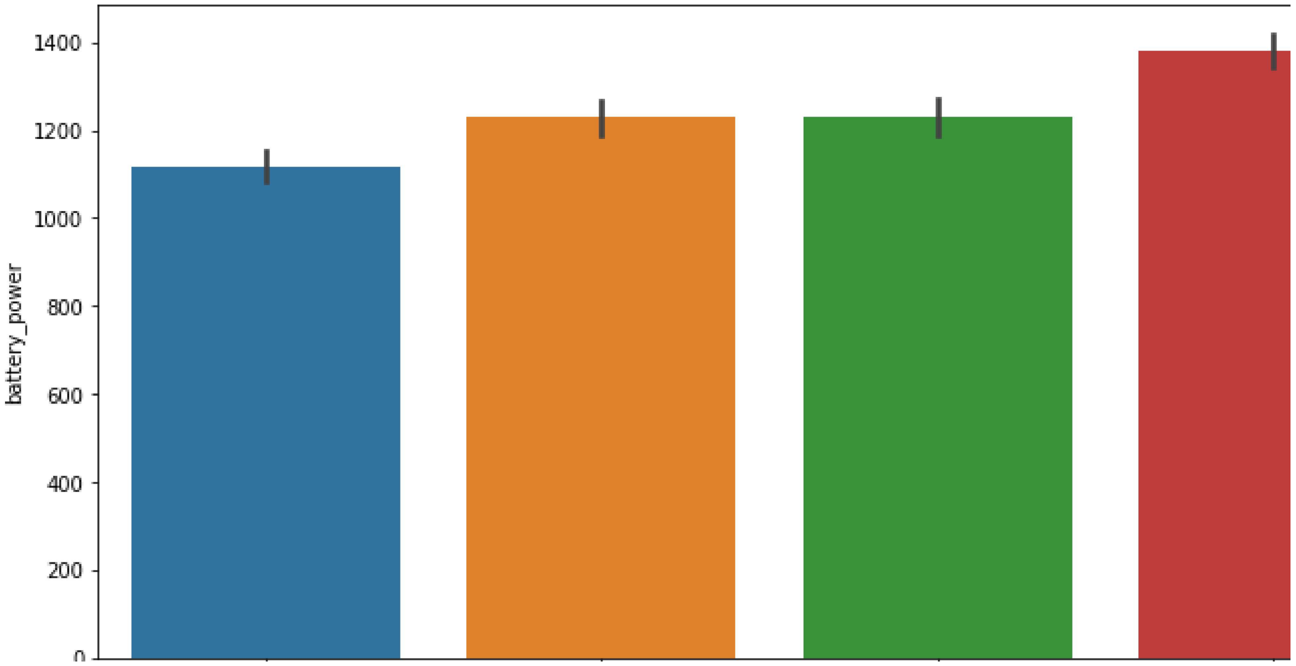
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2000 entries, 0 to 1999
Data columns (total 21 columns):
 #   Column         Non-Null Count  Dtype
---  ------         --------------  -----
 0   battery_power  2000 non-null   int64
 1   blue           2000 non-null   int64
```

```
     2   clock_speed    2000 non-null   float64
     3   dual_sim       2000 non-null   int64
     4   fc             2000 non-null   int64
     5   four_g         2000 non-null   int64
     6   int_memory     2000 non-null   int64
     7   m_dep          2000 non-null   float64
     8   mobile_wt      2000 non-null   int64
     9   n_cores        2000 non-null   int64
     10  pc             2000 non-null   int64
     11  px_height      2000 non-null   int64
     12  px_width       2000 non-null   int64
     13  ram            2000 non-null   int64
     14  sc_h           2000 non-null   int64
     15  sc_w           2000 non-null   int64
     16  talk_time      2000 non-null   int64
     17  three_g        2000 non-null   int64
     18  touch_screen   2000 non-null   int64
     19  wifi           2000 non-null   int64
     20  price_range    2000 non-null   int64
    dtypes: float64(2), int64(19)
    memory usage: 328.2 KB
```

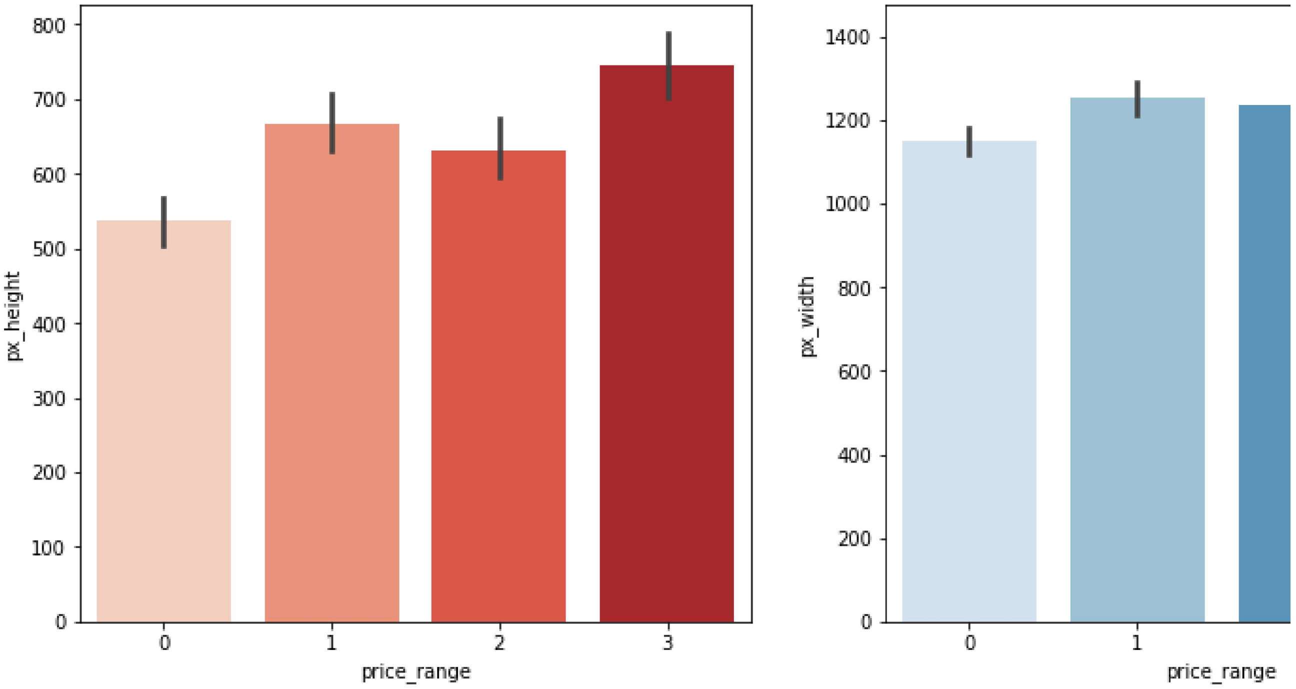
```
plt.figure(figsize=(12,6))
sns.heatmap(df.corr())
plt.show()
```



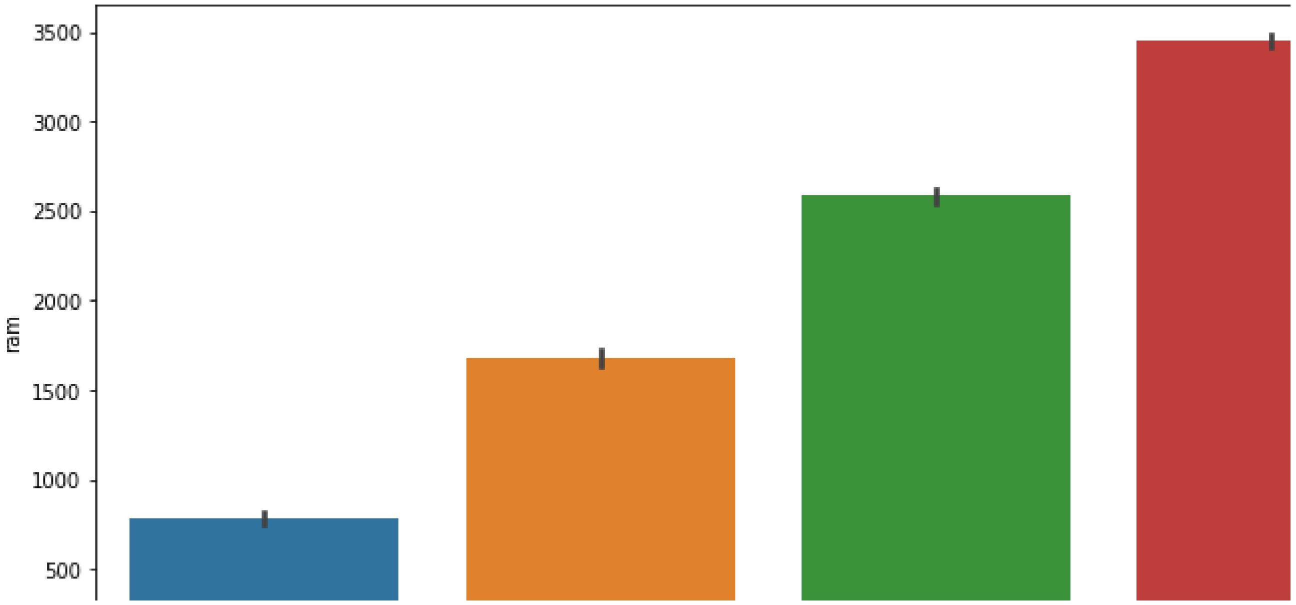
```
plt.figure(figsize=(12,6))
sns.barplot(x='price_range' , y ='battery_power' , data=df)
plt.show()
```



```
plt.figure(figsize=(14,6))
plt.subplot(1,2,1)
sns.barplot(x='price_range',y='px_height',data=df,palette='Reds')
plt.subplot(1,2,2)
sns.barplot(x='price_range',y='px_width',data=df,palette='Blues')
plt.show()
```

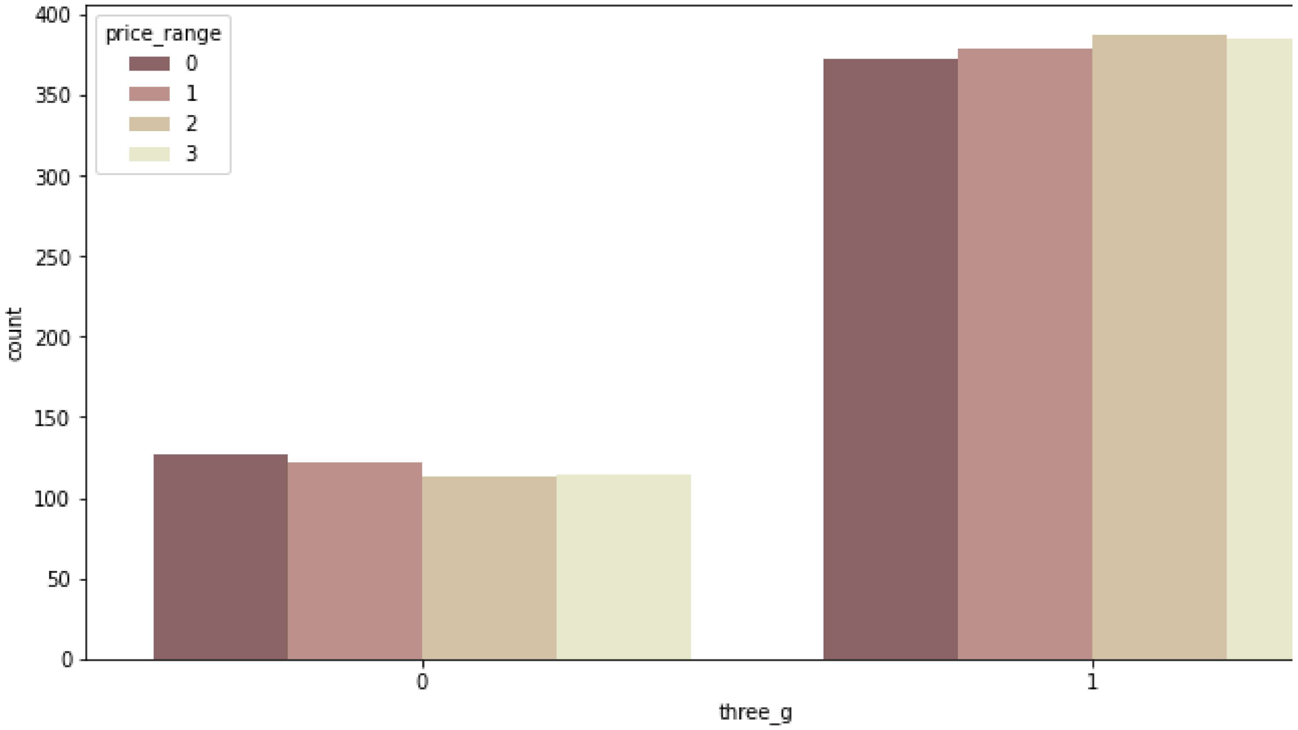


```
plt.figure(figsize=(12,6))
sns.barplot(x='price_range' , y='ram' ,data=df)
plt.show()
```



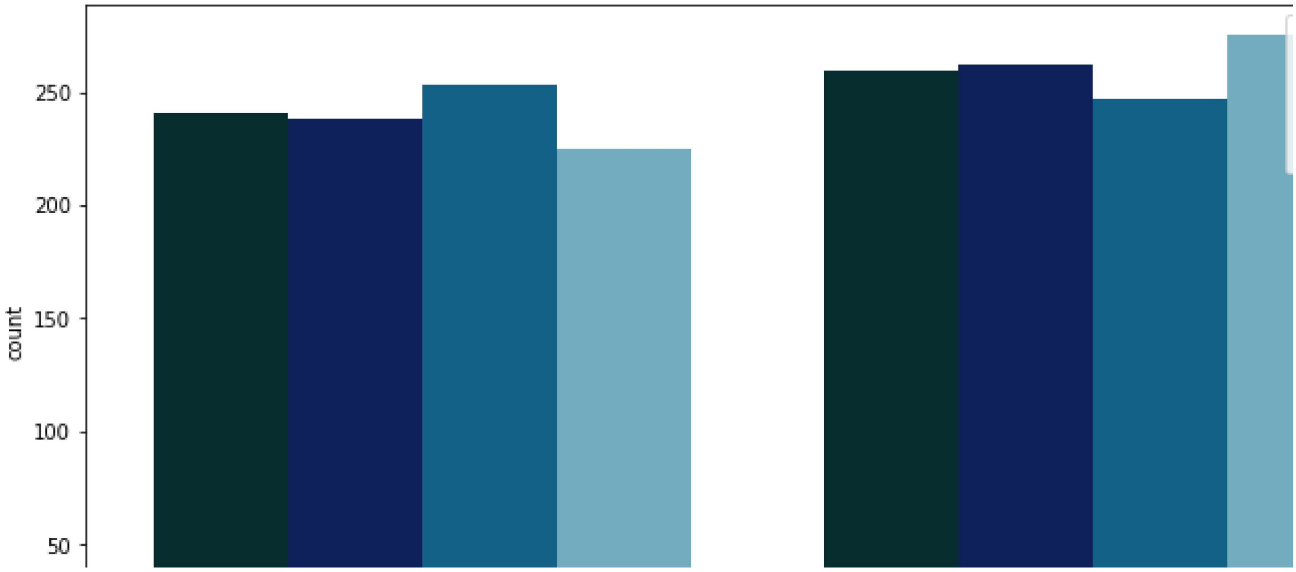
```
plt.figure(figsize=(12,6))
sns.countplot(df['three_g'], hue=df['price_range'],palette='pink')
plt.show()
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pass
  FutureWarning
```

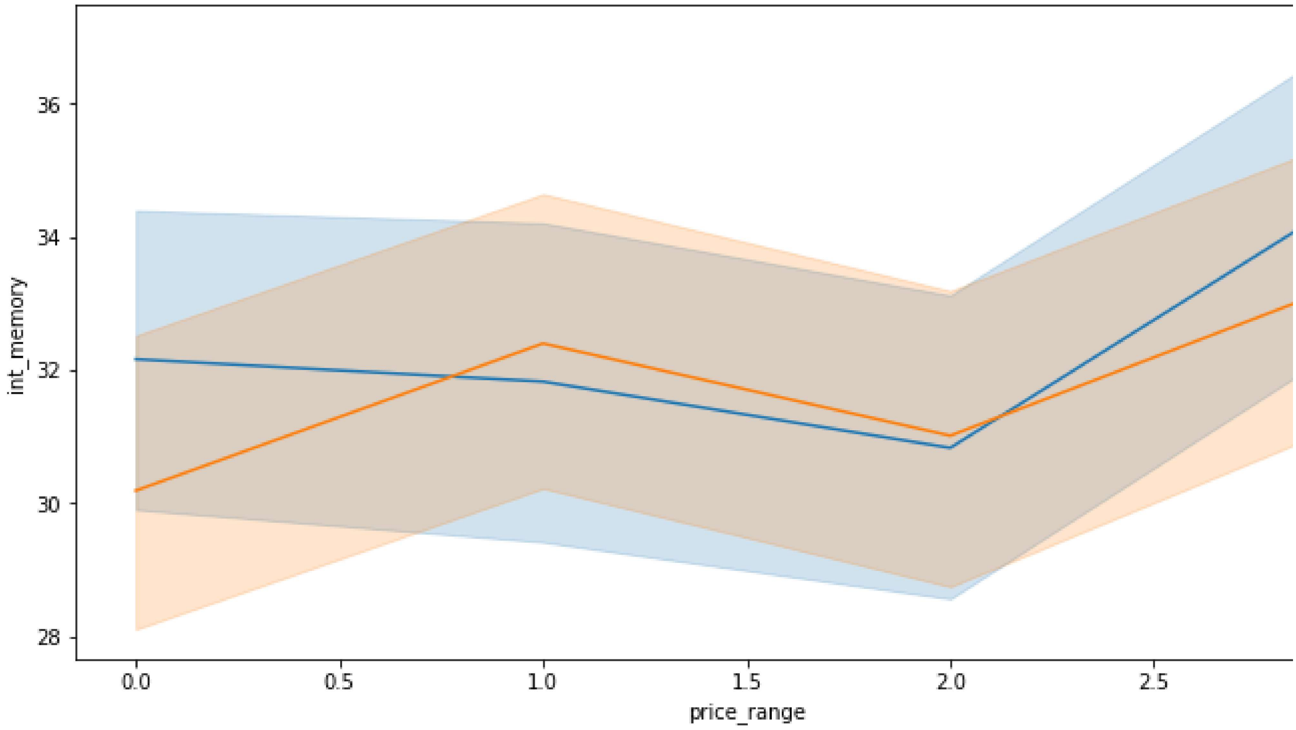


```
plt.figure(figsize=(12,6))
sns.countplot(df['four_g'], hue=df['price_range'],palette='ocean')
plt.show()
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pass
  FutureWarning
```



```
plt.figure(figsize=(12,6))
sns.lineplot(x='price_range', y='int_memory', data=df,hue='dual_sim')
plt.show()
```



```
x=df.drop(['price_range'],axis=1)
y=df['price_range']
```

```
from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y, test_size=0.3, random_state=0)
```

```
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors=10)
knn.fit(x_train,y_train)
```

```
KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='minkowski',
                     metric_params=None, n_jobs=None, n_neighbors=10, p=2,
                     weights='uniform')
```

```
knn.score(x_train,y_train)
```

```
0.9457142857142857
```