In [ ]:

```
```

In [12]:

```
# IMPORT LIBRARIES
import numpy  as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [13]:

```
a=pd.read_csv(r"C:\Users\user\Downloads\4_drug200.csv")
a
```

Out[13]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY
...	...	...	...	...	...	...
195	56	F	LOW	HIGH	11.567	drugC
196	16	M	LOW	HIGH	12.006	drugC
197	52	M	NORMAL	HIGH	9.894	drugX
198	23	M	NORMAL	NORMAL	14.020	drugX
199	40	F	LOW	NORMAL	11.349	drugX

200 rows × 6 columns

In [14]:

```
a=a.head(10)
a
```

Out[14]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY
5	22	F	NORMAL	HIGH	8.607	drugX
6	49	F	NORMAL	HIGH	16.275	drugY
7	41	M	LOW	HIGH	11.037	drugC
8	60	M	NORMAL	HIGH	15.171	drugY
9	43	M	LOW	NORMAL	19.368	drugY

In [15]:

```
# to find
a.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 6 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   Age          10 non-null     int64  
 1   Sex          10 non-null     object 
 2   BP           10 non-null     object 
 3   Cholesterol  10 non-null     object 
 4   Na_to_K      10 non-null     float64
 5   Drug         10 non-null     object 
dtypes: float64(1), int64(1), object(4)
memory usage: 608.0+ bytes
```

In [16]:

```
# to display summary of statastic
a.describe()
```

Out[16]:

	Age	Na_to_K
count	10.000000	10.000000
mean	42.100000	14.486100
std	13.916018	5.482634
min	22.000000	7.798000
25%	31.250000	10.344750
50%	45.000000	14.132000
75%	48.500000	17.601000
max	61.000000	25.355000

In [17]:

```
# to display colum heading
a.columns
```

Out[17]:

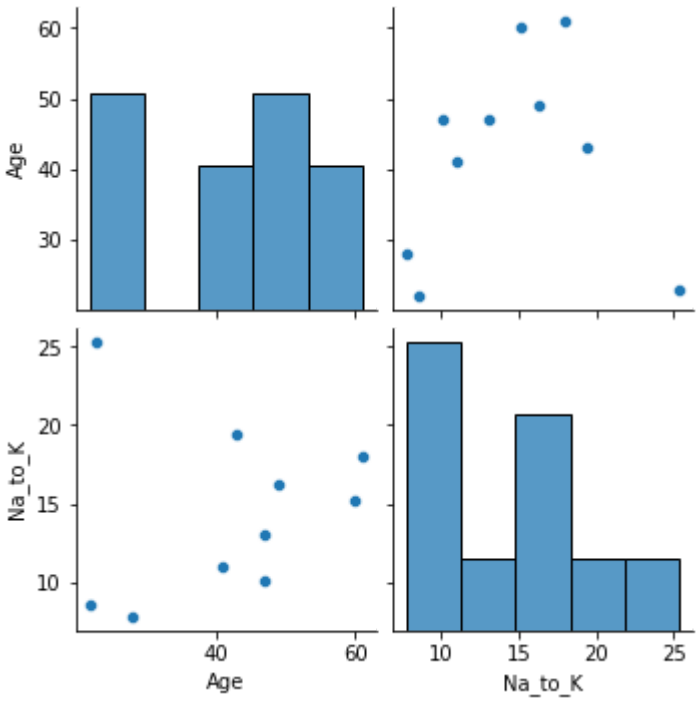
```
Index(['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug'], dtype='objec
t')
```

In [18]:

```
sns.pairplot(a)
```

Out[18]:

```
<seaborn.axisgrid.PairGrid at 0x203a96b1d60>
```

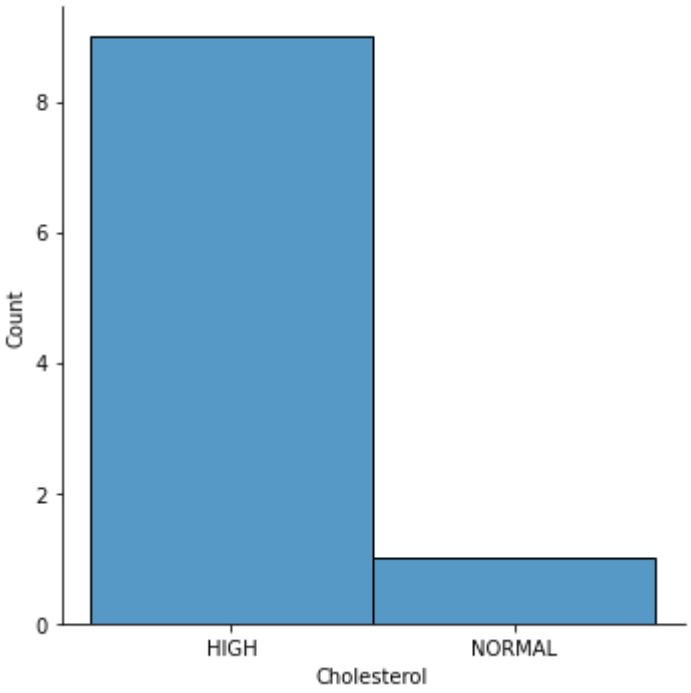


In [20]:

```
sns.displot(a["Cholesterol"])
```

Out[20]:

```
<seaborn.axisgrid.FacetGrid at 0x203aa3c8370>
```



In [21]:

```
b=a[['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug']]
b
```

Out[21]:

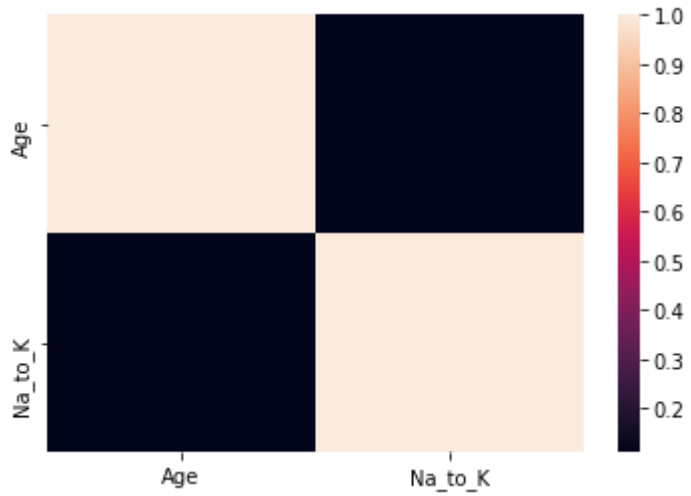
	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY
5	22	F	NORMAL	HIGH	8.607	drugX
6	49	F	NORMAL	HIGH	16.275	drugY
7	41	M	LOW	HIGH	11.037	drugC
8	60	M	NORMAL	HIGH	15.171	drugY
9	43	M	LOW	NORMAL	19.368	drugY

In [22]:

```
sns.heatmap(b.corr())
```

Out[22]:

<AxesSubplot:>



In [33]:

```
x=a[['Age','Na_to_K']]
y=a['Age']
```

In [34]:

```
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)
```

In [35]:

```
from sklearn.linear_model import LinearRegression
lr=LinearRegression()
lr.fit(x_train,y_train)
```

Out[35]:

```
LinearRegression()
```

In [36]:

```
lr.intercept_
```

Out[36]:

```
0.0
```

In [37]:

```
coeff=pd.DataFrame(lr.coef_,x.columns,columns=['Co-efficient'])
coeff
```

Out[37]:

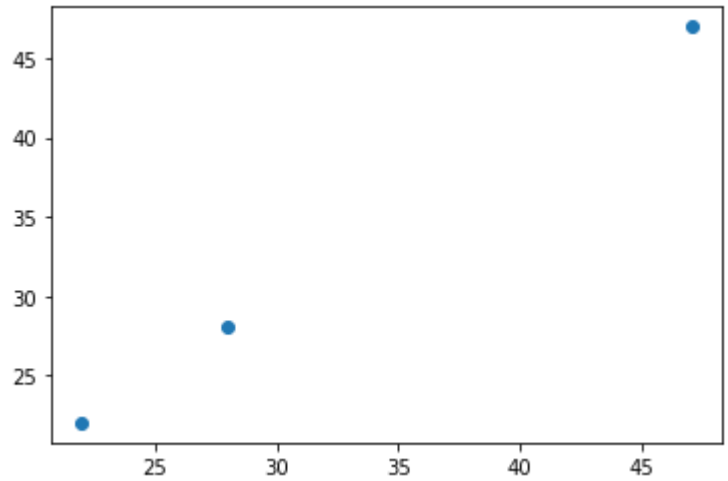
	Co-efficient
Age	1.000000e+00
Na_to_K	2.817941e-16

In [38]:

```
prediction = lr.predict(x_test)
plt.scatter(y_test,prediction)
```

Out[38]:

```
<matplotlib.collections.PathCollection at 0x203abee22b0>
```



In [39]:

```
lr.score(x_test,y_test)
```

Out[39]:

```
1.0
```

In [40]:

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

Out[40]:

```
1.0
```

In [41]:

```
from sklearn.linear_model import Ridge,Lasso
```

In [42]:

```
rr=Ridge(alpha=10)
rr.fit(x_test,y_test)
```

Out[42]:

```
Ridge(alpha=10)
```

In [43]:

```
rr.score(x_test,y_test)
```

Out[43]:

```
0.9991838929162196
```

In [44]:

```
la=Lasso(alpha=10)
la.fit(x_test,y_test)
```

Out[44]:

```
Lasso(alpha=10)
```

In [45]:

```
la.score(x_test,y_test)
```

Out[45]:

```
0.9922449745520275
```

In [ ]:

```

```