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

In [2]:
```
df=pd.read_csv(r"C:\Users\Admin\Downloads\4_drug200 - 4_drug200.csv")
df
```

Out[2]:

	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 [3]:
```
df.info()
```

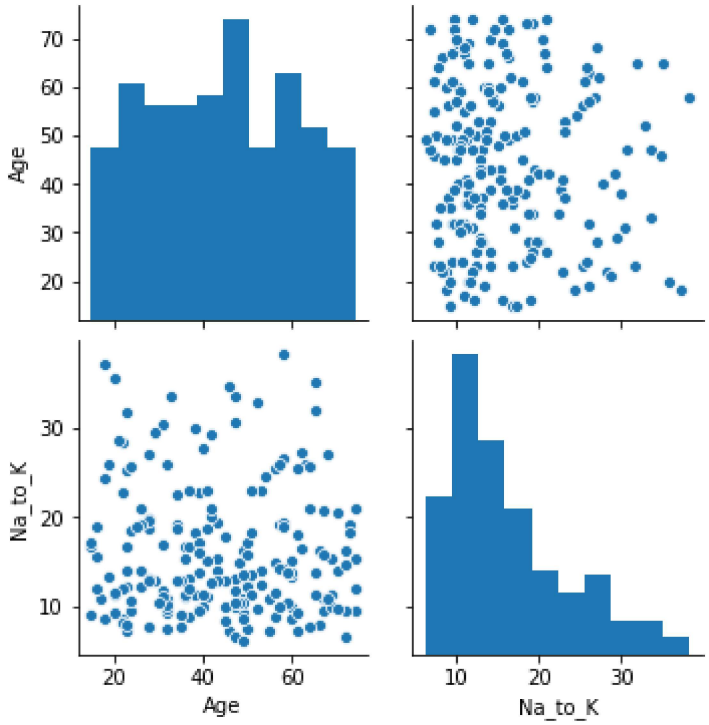
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 200 entries, 0 to 199
Data columns (total 6 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   Age          200 non-null    int64  
 1   Sex          200 non-null    object 
 2   BP           200 non-null    object 
 3   Cholesterol  200 non-null    object 
 4   Na_to_K      200 non-null    float64
 5   Drug         200 non-null    object 
dtypes: float64(1), int64(1), object(4)
memory usage: 9.5+ KB
```

In [4]:
```
df.columns
```

Out[4]:
```
Index(['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug'], dtype='object')
```

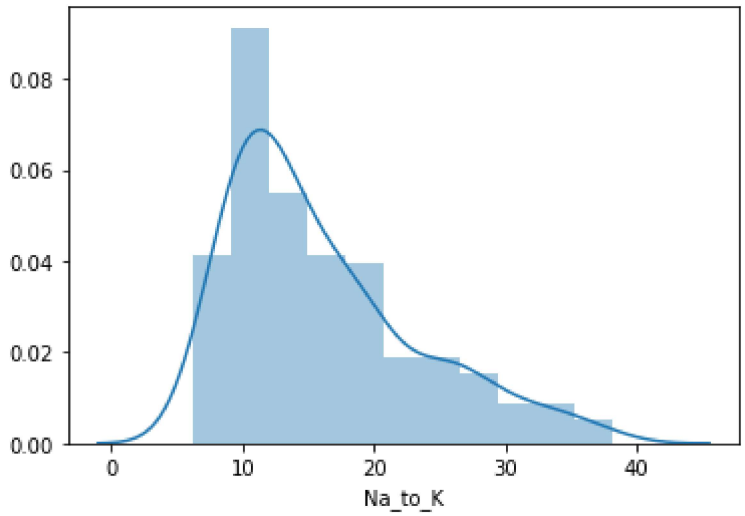
In [5]:
```
sns.pairplot(df)
```

Out[5]:
```
<seaborn.axisgrid.PairGrid at 0x1e21adf93d0>
```



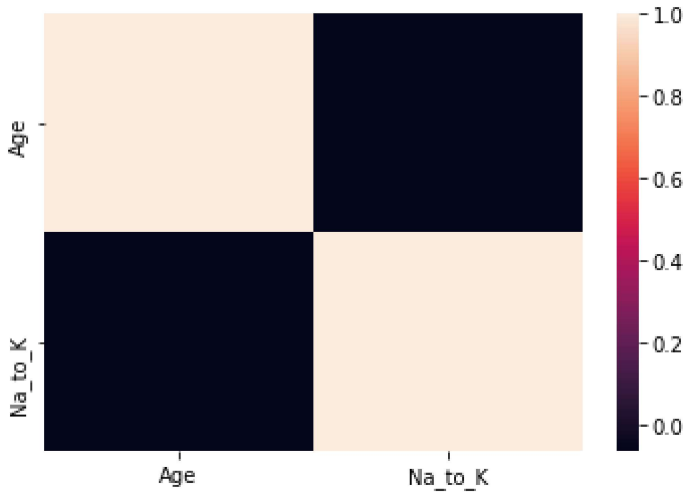
In [7]:
```
sns.distplot(df['Na_to_K'])
```

Out[7]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1e21fde3a90>
```



In [8]:
```
sns.heatmap(df.corr())
```

Out[8]: <matplotlib.axes._subplots.AxesSubplot at 0x1e21ffb84c0>



In [9]:
```
x=df[['Age']]
y=df['Na_to_K']
```

In [10]:
```
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 [11]:
```
from sklearn.linear_model import LinearRegression
lr= LinearRegression()
lr.fit(x_train,y_train)
```

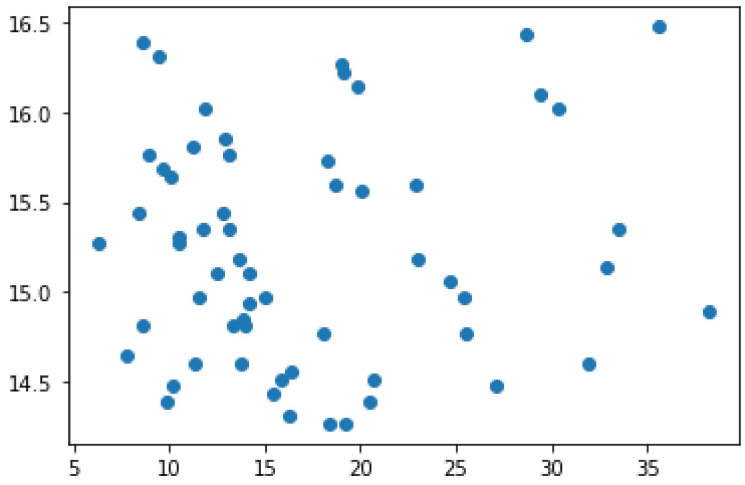
Out[11]: LinearRegression()

In [12]:
```
print(lr.intercept_)
```

```
17.31259028739985
```

In [13]:
```
prediction= lr.predict(x_test)
plt.scatter(y_test,prediction)
```

Out[13]: <matplotlib.collections.PathCollection at 0x1e220a9d760>



In [14]:
```
print(lr.score(x_test,y_test))
```

```
-0.0669626754972179
```

In [15]:
```
print(lr.score(x_train,y_train))
```

```
0.009988229995840037
```

In [16]:
```
from sklearn.linear_model import Ridge,Lasso
```

In [17]:
```
rr=Ridge(alpha=10)
rr.fit(x_train,y_train)
```

Out[17]: Ridge(alpha=10)

In [18]:
```
rr.score(x_test,y_test)
```

Out[18]: -0.0669561274313637

In [19]:
```
la=Lasso(alpha=10)
la.fit(x_train,y_train)
```

Out[19]: Lasso(alpha=10)

In [20]:
```
la.score(x_test,y_test)
```

Out[20]: -0.05148619823265066

In [ ]: