In [6]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import tensorflow
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import layer,Dense,Dropout
```

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_7728\2926707417.py in <module>
      2 import numpy as np
      3 import matplotlib.pyplot as plt
----> 4 import tensorflow
      5 from tensorflow.keras.models import Sequential
      6 from tensorflow.keras.layers import layer,Dense,Dropout

ModuleNotFoundError: No module named 'tensorflow'
```

In [9]:
```
data=pd.read_csv("D:\\NMDS\\drug200.csv")
data.head()
```

Out[9]:

	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

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

Out[10]:
```
Index(['Age', 'Sex', 'BP', 'Cholesterol', 'Na_to_K', 'Drug'], dtype='object')
```

In [12]:
```
data.Drug
```

Out[12]:
```
0      DrugY
1      drugC
2      drugC
3      drugX
4      DrugY
       ...
195    drugC
196    drugC
197    drugX
198    drugX
199    drugX
Name: Drug, Length: 200, dtype: object
```

In [13]:
```
data.isnull()
```

Out[13]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
...	...	...	...	...	...	...
195	False	False	False	False	False	False
196	False	False	False	False	False	False
197	False	False	False	False	False	False
198	False	False	False	False	False	False
199	False	False	False	False	False	False

200 rows × 6 columns

In [14]:
```
data.isnull().sum()
```

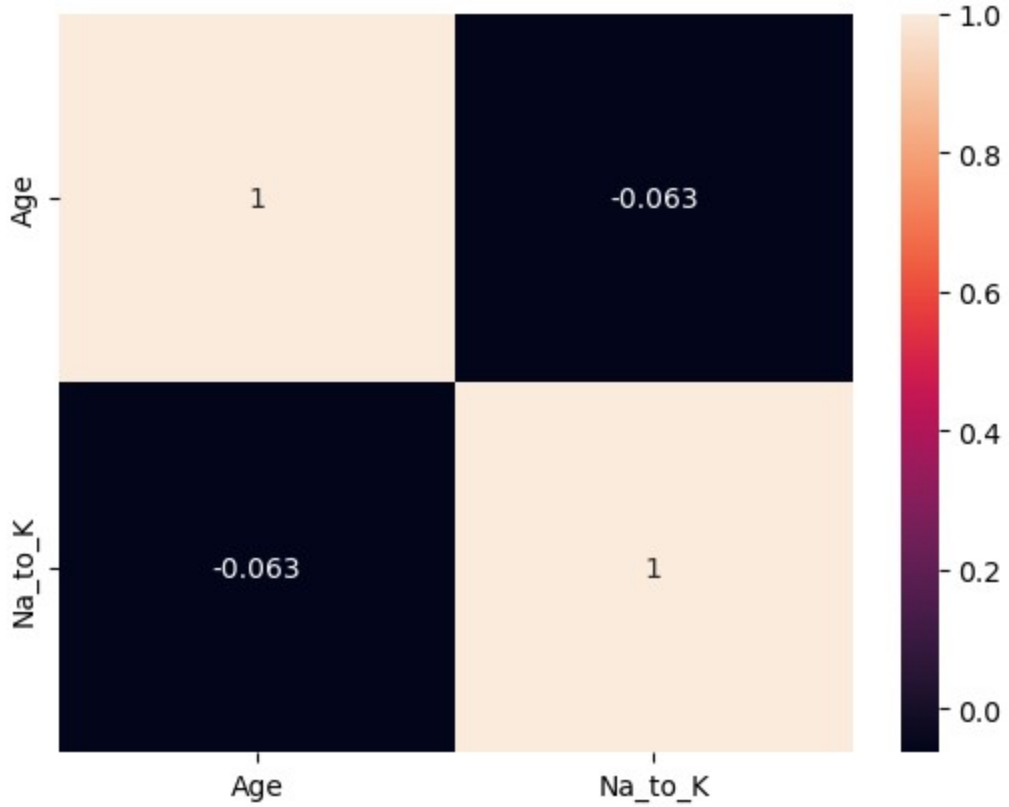
Out[14]:
```
Age            0
Sex            0
BP             0
Cholesterol    0
Na_to_K        0
Drug           0
dtype: int64
```

In [16]:
```
data.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 [18]:
```
import seaborn as sns
sns.heatmap(data.corr(),annot=True)
```

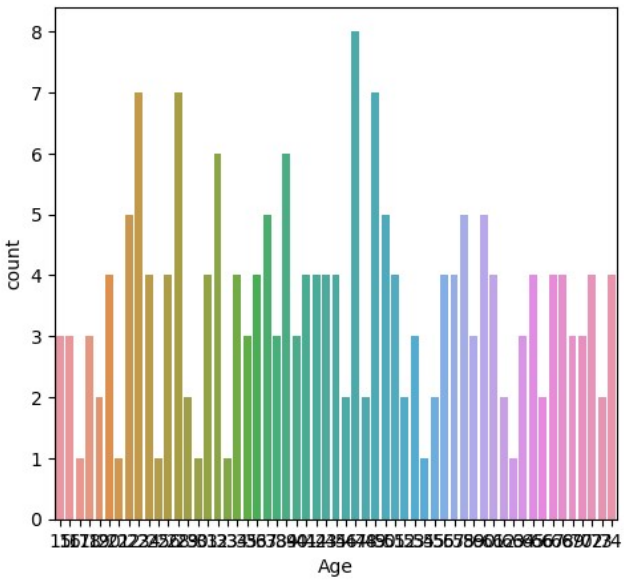
Out[18]:
```
<AxesSubplot:>
```

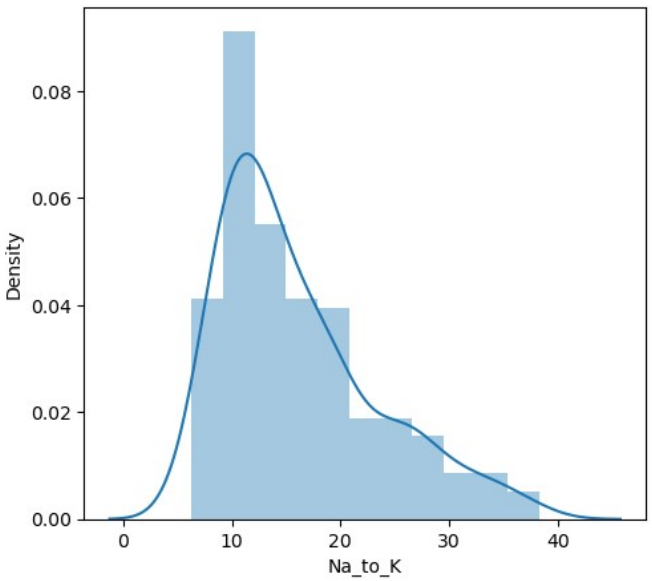


In [23]:
```
plt.figure(figsize=(12,5))
plt.subplot(1,2,1)
sns.countplot(data['Age'])
plt.subplot(1,2,2)
sns.distplot(data['Na_to_K'])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\_decorators.py:36: FutureWarning:
Pass the following variable as a keyword arg: x. From version 0.12, the only valid po
sitional argument will be `data`, and passing other arguments without an explicit key
word will result in an error or misinterpretation.
  warnings.warn(
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2619: FutureWarni
ng: `distplot` is a deprecated function and will be removed in a future version. Plea
se adapt your code to use either `displot` (a figure-level function with similar flex
ibility) or `histplot` (an axes-level function for histograms).
  warnings.warn(msg, FutureWarning)
```

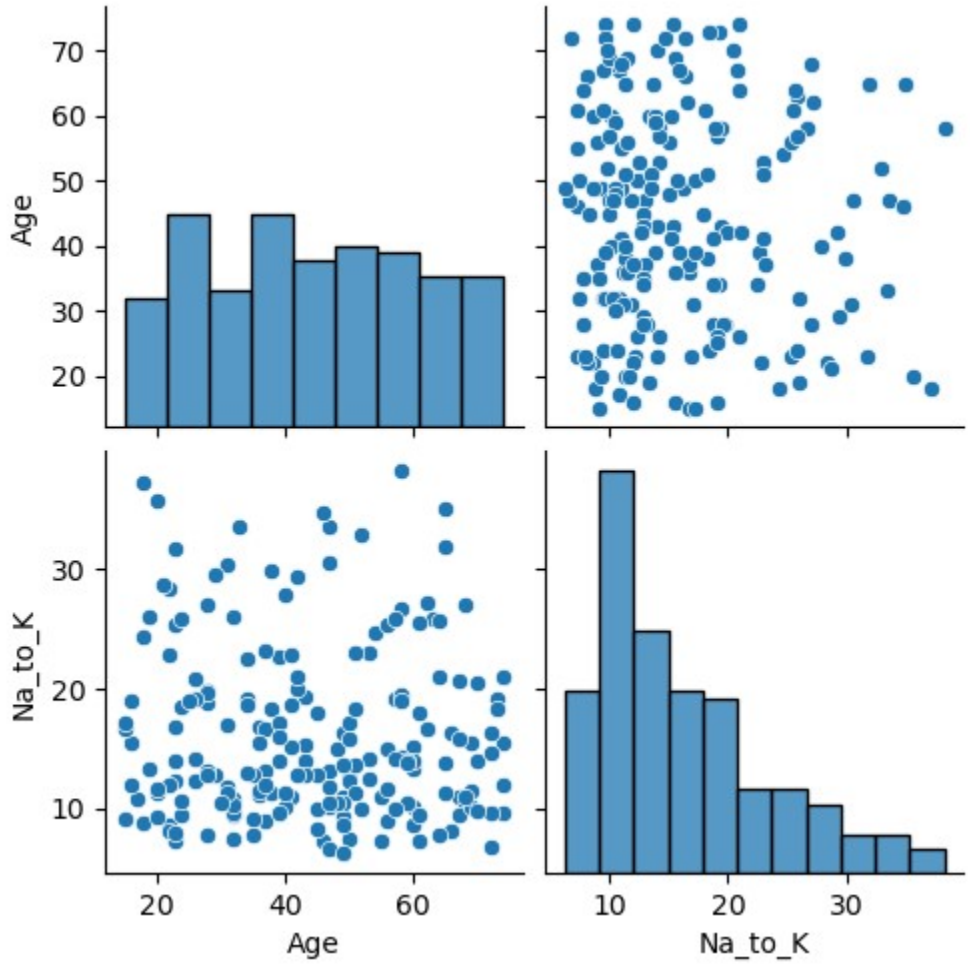
Out[23]:<AxesSubplot:xlabel='Na_to_K', ylabel='Density'>





In [26]:
```
sns.pairplot(data=data,markers=["^","v"],palette="Inferno")
```

Out[26]:<seaborn.axisgrid.PairGrid at 0x2a8ef2df280>



In [...
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test=train_test_split(x_resamble,y_resamble,test_size=0
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_7728\2491902902.py in <module>
      1 from sklearn.model_selection import train_test_split
----> 2 x_train,x_test,y_train,y_test=train_test_split(x_resamble,y_resamble,test_siz
e=0.2,random_state=0)

NameError: name 'x_resamble' is not defined
```

In [33]:
```
import sklearn
from sklearn.preprocessing import LabelEncoder, OneHotEncoder
from sklearn.ensample import RandomForestClassifier
rfr1=RandomForestClassifier().fit(x_os,y_os.values.ravel())
y_pred=rfr1.pred(x_test_os)
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_7728\3161771132.py in <module>
      2 from sklearn.preprocessing import LabelEncoder, OneHotEncoder
      3 #from sklearn.ensample import RandomForestClassifier
----> 4 rfr1=RandomForestClassifier().fit(x_os,y_os.values.ravel())
      5 y_pred=rfr1.pred(x_test_os)

NameError: name 'RandomForestClassifier' is not defined
```

In [34]:
```
from xgboost import XGBClassifier
xgb1=XGBClassifier()
xgb1.fit(x_os,y_os)
```

```
---------------------------------------------------------------------------
ModuleNotFoundError                       Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_7728\2150144178.py in <module>
----> 1 from xgboost import XGBClassifier
      2 xgb1=XGBClassifier()
      3 xgb1.fit(x_os,y_os)

ModuleNotFoundError: No module named 'xgboost'
```

In [35]:
```
from sklearn.svm import SVC
from sklearn.metrics import accuracy_score,classification_report
sv=SVC
sv.fit(x_bal,y_bal)
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call last)
~\AppData\Local\Temp\ipykernel_7728\1912348629.py in <module>
      2 from sklearn.metrics import accuracy_score,classification_report
      3 sv=SVC
----> 4 sv.fit(x_bal,y_bal)

NameError: name 'x_bal' is not defined
```

In [ ]: