```
In [58]:   import numpy as np
           import pandas as pd
           import seaborn as sns
           import warnings
           warnings.filterwarnings("ignore")
```

```
In [3]:    df=pd.read_csv("data2.csv")
```

```
In [4]:    df
```

Out[4]:

	Name	Python	Machine Learning	Age	AI	NLP	Cv
0	jay	90	?	23	90	?	23
1	raj	missing	89	NaN	missing	89	NaN
2	kumar	78	78	?	78	78	?
3	suraj	NaN	NaN	missing	NaN	NaN	missing
4	pawan	missing	NaN	21	missing	NaN	21
5	viraj	65	56	22	65	56	22
6	vijay	45	67	?	45	67	?
7	sumit	80	45	24	80	45	24
8	akash	70	67	missing	70	67	missing
9	sujit	NaN	NaN	22	NaN	NaN	22

```
In [5]:    df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 7 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Name              10 non-null     object
 1   Python            8 non-null      object
 2   Machine Learning  7 non-null      object
 3   Age               9 non-null      object
 4   AI                8 non-null      object
 5   NLP               7 non-null      object
 6   Cv                9 non-null      object
dtypes: object(7)
memory usage: 688.0+ bytes
```

```
In [6]:    df[ "Python"].value_counts()
```

```
Out[6]:    missing    2
           90         1
           78         1
           65         1
           45         1
           80         1
           70         1
           Name: Python, dtype: int64
```

```
In [10]:   for i in df:
                   print(df[i].value_counts())
```

```
           jay     1
           raj     1
```

```
         kumar    1
         suraj    1
         pawan    1
         viraj    1
         vijay    1
         sumit    1
         akash    1
         sujit    1
         Name: Name, dtype: int64
         missing    2
         90         1
         78         1
         65         1
         45         1
         80         1
         70         1
         Name: Python, dtype: int64
         67    2
         ?     1
         89    1
         78    1
         56    1
         45    1
         Name: Machine Learning, dtype: int64
         ?          2
         missing    2
         22         2
         23         1
         21         1
         24         1
         Name: Age, dtype: int64
         missing    2
         90         1
         78         1
         65         1
         45         1
         80         1
         70         1
         Name: AI, dtype: int64
         67    2
         ?     1
         89    1
         78    1
         56    1
         45    1
         Name: NLP, dtype: int64
         ?          2
         missing    2
         22         2
         23         1
         21         1
         24         1
         Name: Cv, dtype: int64
```

In [11]:
```python
for i in df:
    df[i].replace("?",np.nan,inplace=True)
```

In [12]:
```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 7 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
```

```
        0   Name                10 non-null     object
        1   Python              8 non-null      object
        2   Machine Learning    6 non-null      object
        3   Age                 7 non-null      object
        4   AI                  8 non-null      object
        5   NLP                 6 non-null      object
        6   Cv                  7 non-null      object
       dtypes: object(7)
       memory usage: 688.0+ bytes
```

In [14]:
```python
for i in df:
    df[i].replace("missing",np.nan,inplace=True)
```

In [56]:
```python
for i in df1:
    df1[i]=df1[i].astype("float64")
```

In [30]:
```python
df1=df.iloc[:,1:8]
```

In [31]:
```python
df1
```

Out[31]:

	Python	Machine Learning	Age	AI	NLP	Cv
0	90	NaN	23	90	NaN	23
1	NaN	89	NaN	NaN	89	NaN
2	78	78	NaN	78	78	NaN
3	NaN	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	21	NaN	NaN	21
5	65	56	22	65	56	22
6	45	67	NaN	45	67	NaN
7	80	45	24	80	45	24
8	70	67	NaN	70	67	NaN
9	NaN	NaN	22	NaN	NaN	22

In [34]:
```python
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 6 columns):
 #   Column            Non-Null Count  Dtype
---  ------            --------------  -----
 0   Python            6 non-null      float64
 1   Machine Learning  6 non-null      float64
 2   Age               5 non-null      float64
 3   AI                6 non-null      float64
 4   NLP               6 non-null      float64
 5   Cv                5 non-null      float64
dtypes: float64(6)
memory usage: 608.0 bytes
```

In [45]:
```python
mean=[]
for i in df1:
    allmean=df1[i].mean()
    mean.append(allmean)
print(mean)
```

```
[71.33333333333333, 67.0, 22.4, 71.33333333333333, 67.0, 22.4]
```

```
In [57]:   for i in df1:
               for j in mean:
                   df1[i].fillna(j,inplace=True)
```

```
In [51]:   df2=df.iloc[:,0:1]
```

```
In [52]:   df3=df2.join(df1)
```
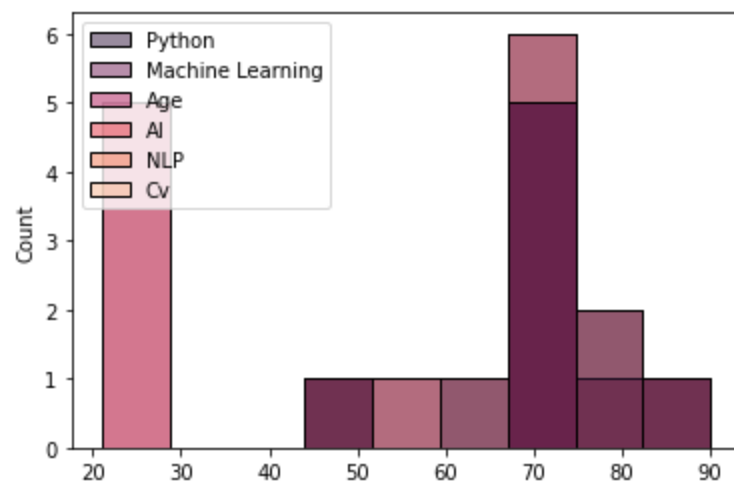
```
In [53]:   df3
```

Out[53]:

	Name	Python	Machine Learning	Age	AI	NLP	Cv
0	jay	90.000000	71.333333	23.000000	90.000000	71.333333	23.000000
1	raj	71.333333	89.000000	71.333333	71.333333	89.000000	71.333333
2	kumar	78.000000	78.000000	71.333333	78.000000	78.000000	71.333333
3	suraj	71.333333	71.333333	71.333333	71.333333	71.333333	71.333333
4	pawan	71.333333	71.333333	21.000000	71.333333	71.333333	21.000000
5	viraj	65.000000	56.000000	22.000000	65.000000	56.000000	22.000000
6	vijay	45.000000	67.000000	71.333333	45.000000	67.000000	71.333333
7	sumit	80.000000	45.000000	24.000000	80.000000	45.000000	24.000000
8	akash	70.000000	67.000000	71.333333	70.000000	67.000000	71.333333
9	sujit	71.333333	71.333333	22.000000	71.333333	71.333333	22.000000

```
In [54]:   df3.describe()
```

Out[54]:

	Python	Machine Learning	Age	AI	NLP	Cv
count	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000
mean	71.333333	68.733333	46.866667	71.333333	68.733333	46.866667
std	11.563032	11.808973	25.801330	11.563032	11.808973	25.801330
min	45.000000	45.000000	21.000000	45.000000	45.000000	21.000000
25%	70.333333	67.000000	22.250000	70.333333	67.000000	22.250000
50%	71.333333	71.333333	47.666667	71.333333	71.333333	47.666667
75%	76.333333	71.333333	71.333333	76.333333	71.333333	71.333333
max	90.000000	89.000000	71.333333	90.000000	89.000000	71.333333

```
In [63]:   sns.histplot(data=df3,palette="rocket");
```

In [65]: `sns.boxplot(data=df3);`