In [1]:

```
pip install pandas
```

```
Requirement already satisfied: pandas in c:\users\sakshi potdarvc\anaconda3\lib\site-
packages (1.3.4)
Requirement already satisfied: pytz>=2017.3 in c:\users\sakshi potdarvc\anaconda3\lib
\site-packages (from pandas) (2021.3)
Requirement already satisfied: numpy>=1.17.3 in c:\users\sakshi potdarvc\anaconda3\li
b\site-packages (from pandas) (1.20.3)
Requirement already satisfied: python-dateutil>=2.7.3 in c:\users\sakshi potdarvc\ana
conda3\lib\site-packages (from pandas) (2.8.2)
Requirement already satisfied: six>=1.5 in c:\users\sakshi potdarvc\anaconda3\lib\sit
e-packages (from python-dateutil>=2.7.3->pandas) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

In [5]:

```
import pandas as pd
import numpy as np
```

In [7]:

```
a = pd.read_csv(r'Downloads\StudentsPerformance1 - StudentsPerformance1.csv')
```

In [9]:

```
a.describe()
```

Out[9]:

	math score	reading score	writing score
count	19.000000	19.000000	17.000000
mean	61.421053	67.736842	65.294118
std	19.712703	18.902798	20.250817
min	18.000000	32.000000	28.000000
25%	46.500000	53.500000	46.000000
50%	65.000000	72.000000	70.000000
75%	74.000000	82.000000	78.000000
max	90.000000	95.000000	93.000000

In [10]:

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

Out[10]:

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score
0	female	group B	bachelor's degree	standard	none	72	72	NaN
1	female	group C	some college	standard	completed	69	90	88.0
2	female	group B	master's degree	standard	none	90	95	93.0
3	male	group A	associate's degree	free/reduced	none	47	57	44.0
4	male	group C	some college	standard	none	76	78	75.0
5	female	group B	associate's degree	standard	none	71	83	78.0

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score
6	female	group B	some college	standard	completed	88	95	92.0
7	male	group B	some college	free/reduced	none	40	43	39.0
8	male	group D	high school	free/reduced	completed	64	64	67.0
9	female	group B	high school	free/reduced	none	38	60	NaN

In [11]:

```
a.isnull().sum()
```

Out[11]:

```
gender                         0
race/ethnicity                 0
parental level of education    1
lunch                          0
test preparation course        0
math score                     0
reading score                  0
writing score                  2
dtype: int64
```

In [13]:

```
a.dtypes
```

Out[13]:

```
gender                          object
race/ethnicity                  object
parental level of education     object
lunch                           object
test preparation course         object
math score                       int64
reading score                    int64
writing score                  float64
dtype: object
```

In [14]:

```
a.dropna
```

Out[14]:

```
<bound method DataFrame.dropna of     gender race/ethnicity parental level of educati
on         lunch  \
0   female        group B           bachelor's degree      standard
1   female        group C                some college      standard
2   female        group B             master's degree      standard
3     male        group A          associate's degree  free/reduced
4     male        group C                some college      standard
5   female        group B          associate's degree      standard
6   female        group B                some college      standard
7     male        group B                some college  free/reduced
8     male        group D                 high school  free/reduced
9   female        group B                 high school  free/reduced
10    male        group C          associate's degree      standard
11    male        group D          associate's degree      standard
12  female        group B                 high school      standard
13    male        group A                some college      standard
14  female        group A             master's degree      standard
15  female        group C                         NaN      standard
16    male        group C                 high school      standard
17  female        group B            some high school  free/reduced
18    male        group C             master's degree  free/reduced

   test preparation course  math score  reading score  writing score
0                     none          72             72            NaN
```

```
1            completed         69          90          88.0
2                 none         90          95          93.0
3                 none         47          57          44.0
4                 none         76          78          75.0
5                 none         71          83          78.0
6            completed         88          95          92.0
7                 none         40          43          39.0
8            completed         64          64          67.0
9                 none         38          60           NaN
10                none         58          54          52.0
11                none         40          52          43.0
12                none         65          81          73.0
13           completed         78          72          70.0
14                none         50          53          58.0
15                none         69          75          78.0
16                none         88          89          86.0
17                none         18          32          28.0
18           completed         46          42          46.0  >
```

In [15]:

```
a.describe()
```

Out[15]:

	math score	reading score	writing score
count	19.000000	19.000000	17.000000
mean	61.421053	67.736842	65.294118
std	19.712703	18.902798	20.250817
min	18.000000	32.000000	28.000000
25%	46.500000	53.500000	46.000000
50%	65.000000	72.000000	70.000000
75%	74.000000	82.000000	78.000000
max	90.000000	95.000000	93.000000

In [18]:

```
print(a['gender'].value_counts())
```

```
female    10
male       9
Name: gender, dtype: int64
```

In [21]:

```
a_Lunch = pd.get_dummies(a['lunch'])
a_new = pd.concat([a, a_Lunch], axis=1)
print(a_new)
```

```
    gender race/ethnicity parental level of education         lunch  \
0   female        group B           bachelor's degree      standard
1   female        group C                some college      standard
2   female        group B             master's degree      standard
3     male        group A          associate's degree  free/reduced
4     male        group C                some college      standard
5   female        group B          associate's degree      standard
6   female        group B                some college      standard
7     male        group B                some college  free/reduced
8     male        group D                 high school  free/reduced
9   female        group B                 high school  free/reduced
10    male        group C          associate's degree      standard
11    male        group D          associate's degree      standard
12  female        group B                 high school      standard
```

```
13    male        group A           some college      standard
14  female        group A        master's degree      standard
15  female        group C                    NaN      standard
16    male        group C            high school      standard
17  female        group B       some high school  free/reduced
18    male        group C        master's degree  free/reduced

   test preparation course  math score  reading score  writing score  \
0                     none          72             72            NaN   
1                completed          69             90           88.0   
2                     none          90             95           93.0   
3                     none          47             57           44.0   
4                     none          76             78           75.0   
5                     none          71             83           78.0   
6                completed          88             95           92.0   
7                     none          40             43           39.0   
8                completed          64             64           67.0   
9                     none          38             60            NaN   
10                    none          58             54           52.0   
11                    none          40             52           43.0   
12                    none          65             81           73.0   
13               completed          78             72           70.0   
14                    none          50             53           58.0   
15                    none          69             75           78.0   
16                    none          88             89           86.0   
17                    none          18             32           28.0   
18               completed          46             42           46.0   

    free/reduced  standard  
0              0         1  
1              0         1  
2              0         1  
3              1         0  
4              0         1  
5              0         1  
6              0         1  
7              1         0  
8              1         0  
9              1         0  
10             0         1  
11             0         1  
12             0         1  
13             0         1  
14             0         1  
15             0         1  
16             0         1  
17             1         0  
18             1         0  
```

MATPLOTLIB

In [22]:

```
import matplotlib.pyplot as plt
```

In [23]:

```
b = pd.read_csv(r'Downloads\Iris - Iris.csv')
```

In [24]:

```
b.describe()
```

Out[24]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
count	150.000000	150.000000	150.000000	150.000000	150.000000
mean	75.500000	5.843333	3.054000	3.758667	1.198667
std	43.445368	0.828066	0.433594	1.764420	0.763161
min	1.000000	4.300000	2.000000	1.000000	0.100000
25%	38.250000	5.100000	2.800000	1.600000	0.300000
50%	75.500000	5.800000	3.000000	4.350000	1.300000
75%	112.750000	6.400000	3.300000	5.100000	1.800000
max	150.000000	7.900000	4.400000	6.900000	2.500000

In [25]:

```
b.head(20)
```

Out[25]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species
0	1	5.1	3.5	1.4	0.2	Iris-setosa
1	2	4.9	3.0	1.4	0.2	Iris-setosa
2	3	4.7	3.2	1.3	0.2	Iris-setosa
3	4	4.6	3.1	1.5	0.2	Iris-setosa
4	5	5.0	3.6	1.4	0.2	Iris-setosa
5	6	5.4	3.9	1.7	0.4	Iris-setosa
6	7	4.6	3.4	1.4	0.3	Iris-setosa
7	8	5.0	3.4	1.5	0.2	Iris-setosa
8	9	4.4	2.9	1.4	0.2	Iris-setosa
9	10	4.9	3.1	1.5	0.1	Iris-setosa
10	11	5.4	3.7	1.5	0.2	Iris-setosa
11	12	4.8	3.4	1.6	0.2	Iris-setosa
12	13	4.8	3.0	1.4	0.1	Iris-setosa
13	14	4.3	3.0	1.1	0.1	Iris-setosa
14	15	5.8	4.0	1.2	0.2	Iris-setosa
15	16	5.7	4.4	1.5	0.4	Iris-setosa
16	17	5.4	3.9	1.3	0.4	Iris-setosa
17	18	5.1	3.5	1.4	0.3	Iris-setosa
18	19	5.7	3.8	1.7	0.3	Iris-setosa
19	20	5.1	3.8	1.5	0.3	Iris-setosa

In [27]:

```
b['Species'].unique()
```

Out[27]:

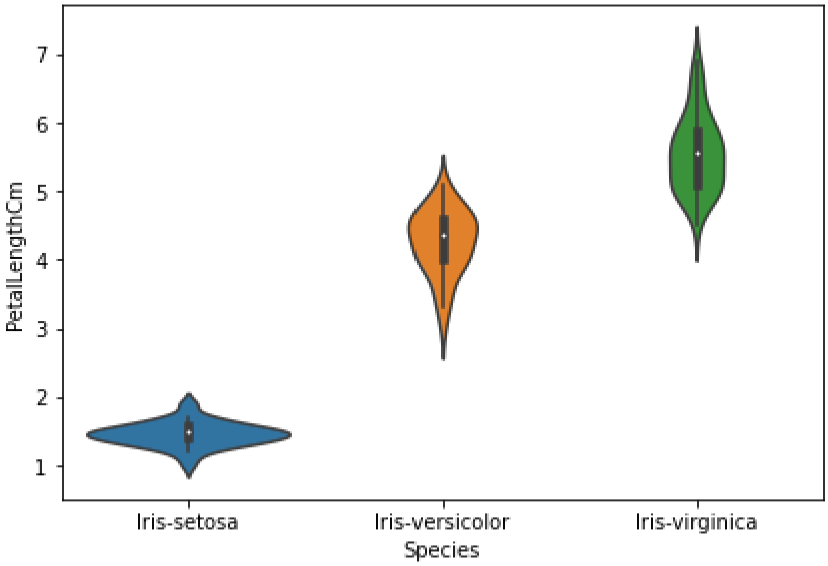
```
array(['Iris-setosa', 'Iris-versicolor', 'Iris-virginica'], dtype=object)
```

In [28]:

```
import seaborn as sns
```

In [36]:

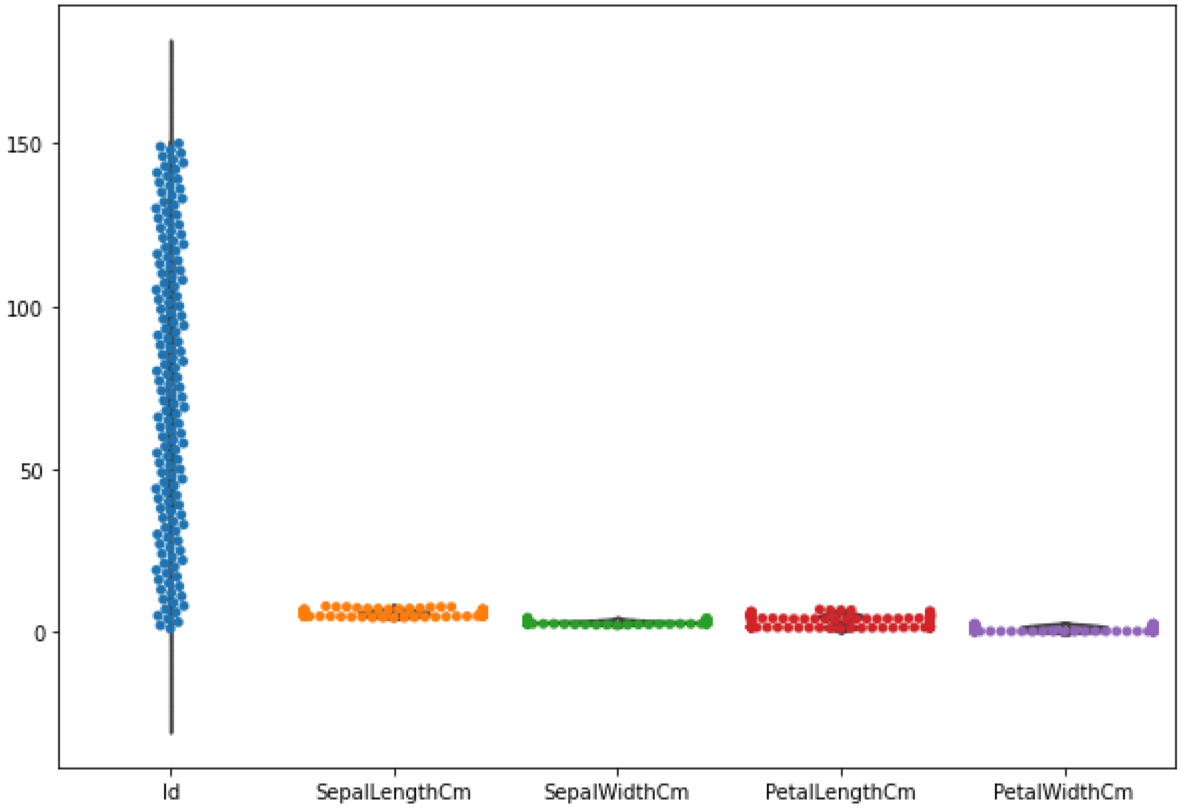
```
plt.figure(figsize=(15,10))
plt.subplot(2,2,1)
sns.violinplot(x='Species',y='PetalLengthCm',data=b)
# plt.subplot(2,2,2)
# sns.violinplot(x='Species',y='PetalLengthCm',data=b)
# plt.subplot(2,2,3)
# sns.violinplot(x='Species',y='PetalLengthCm',data=b)
# plt.subplot(2,2,4)
# sns.violinplot(x='Species',y='PetalLengthCm',data=b)
plt.show()
```



In [43]:

```
plt.subplots(figsize=(10,7))
sns.violinplot(data=b)
sns.swarmplot(data=b)
plt.show()
```

```
C:\Users\Sakshi Potdarvc\anaconda3\lib\site-packages\seaborn\categorical.py:1296: Use
rWarning: 80.0% of the points cannot be placed; you may want to decrease the size of
the markers or use stripplot.
  warnings.warn(msg, UserWarning)
C:\Users\Sakshi Potdarvc\anaconda3\lib\site-packages\seaborn\categorical.py:1296: Use
rWarning: 88.7% of the points cannot be placed; you may want to decrease the size of
the markers or use stripplot.
  warnings.warn(msg, UserWarning)
C:\Users\Sakshi Potdarvc\anaconda3\lib\site-packages\seaborn\categorical.py:1296: Use
rWarning: 74.0% of the points cannot be placed; you may want to decrease the size of
the markers or use stripplot.
  warnings.warn(msg, UserWarning)
```

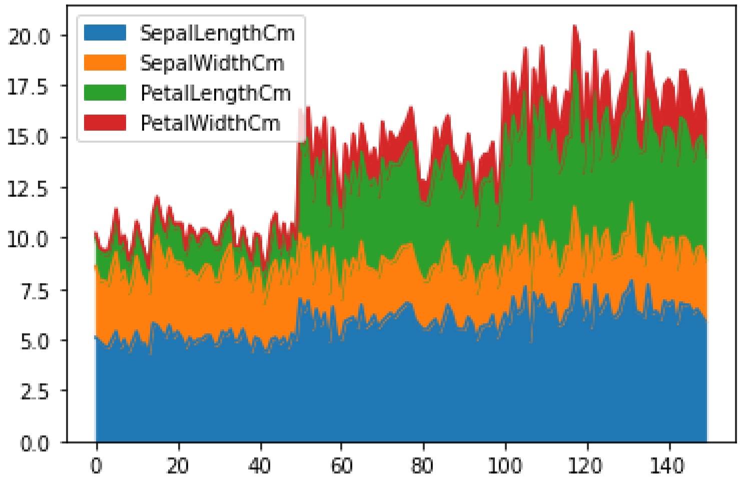


In [44]:

```
b.plot.area(y=['SepalLengthCm','SepalWidthCm', 'PetalLengthCm', 'PetalWidthCm'])
```

Out[44]:

```
<AxesSubplot:>
```



In [45]:

```
b.corr()
```

Out[45]:

	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
Id	1.000000	0.716676	-0.397729	0.882747	0.899759
SepalLengthCm	0.716676	1.000000	-0.109369	0.871754	0.817954
SepalWidthCm	-0.397729	-0.109369	1.000000	-0.420516	-0.356544
PetalLengthCm	0.882747	0.871754	-0.420516	1.000000	0.962757
PetalWidthCm	0.899759	0.817954	-0.356544	0.962757	1.000000

In [ ]:

```
In [ ]:
```