Data_PreProcessing1

August 19, 2022

[1]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.datasets import load_iris
```

[2]:
```
iris_data = load_iris()

data = pd.DataFrame(iris_data.data, columns = iris_data.feature_names)
data
```

[2]:

	sepal length (cm)	sepal width (cm)	petal length (cm)	petal width (cm)
0	5.1	3.5	1.4	0.2
1	4.9	3.0	1.4	0.2
2	4.7	3.2	1.3	0.2
3	4.6	3.1	1.5	0.2
4	5.0	3.6	1.4	0.2
..	…	…	…	…
145	6.7	3.0	5.2	2.3
146	6.3	2.5	5.0	1.9
147	6.5	3.0	5.2	2.0
148	6.2	3.4	5.4	2.3
149	5.9	3.0	5.1	1.8

[150 rows x 4 columns]

[3]:
```
data.columns
```

[3]:
```
Index(['sepal length (cm)', 'sepal width (cm)', 'petal length (cm)',
       'petal width (cm)'],
      dtype='object')
```

[4]:
```
data.shape
```

[4]:
```
(150, 4)
```

[6]:
```
data['sepal length (cm)'].describe()
```

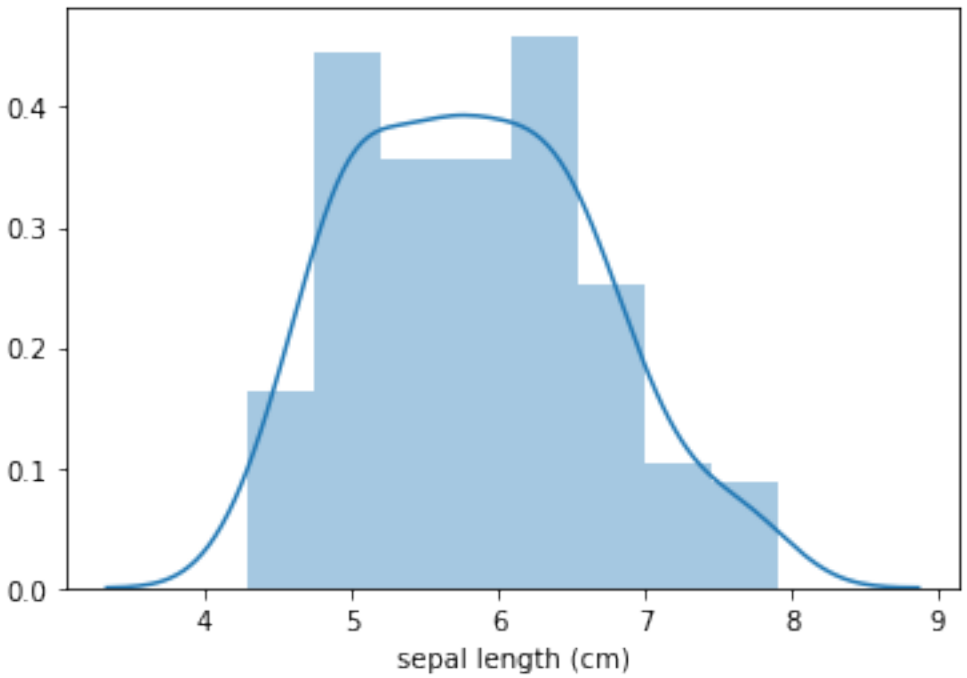
[6]:
```
count    150.000000
mean       5.843333
std        0.828066
min        4.300000
25%        5.100000
50%        5.800000
75%        6.400000
max        7.900000
Name: sepal length (cm), dtype: float64
```

[10]:
```
print('Skewness : ', data['sepal length (cm)'].skew())
print('Kurtosis : ', data['sepal length (cm)'].kurt())
```

```
Skewness :  0.3149109566369728
Kurtosis :  -0.5520640413156395
```

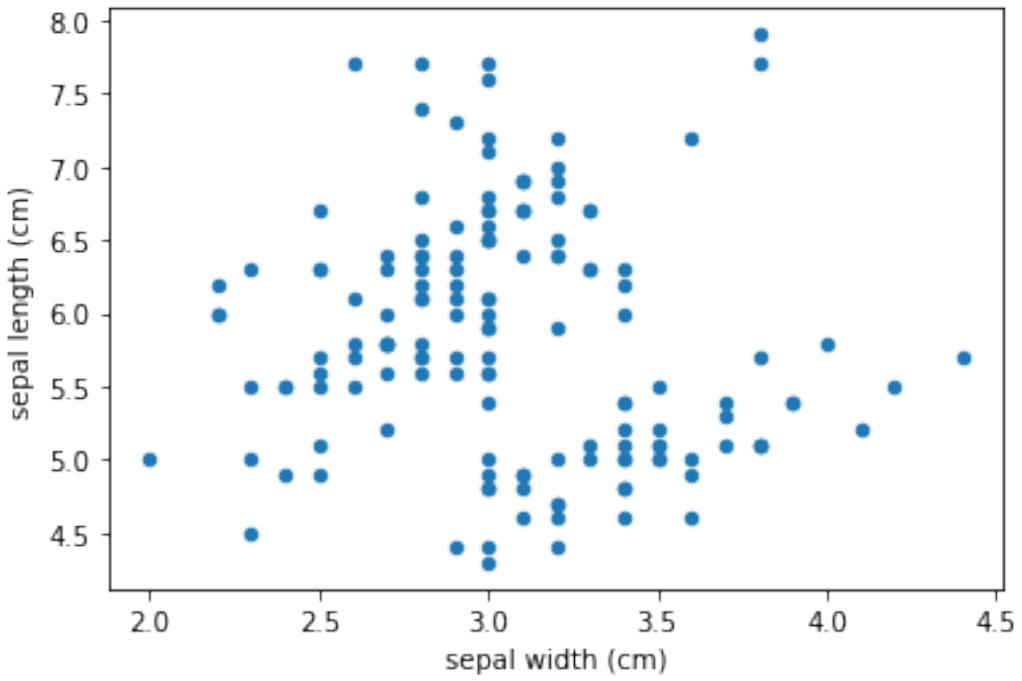
[7]:
```
sns.distplot(data['sepal length (cm)'])
```

[7]:
```
<AxesSubplot:xlabel='sepal length (cm)'>
```



[8]:
```
var='sepal width (cm)'
data.plot.scatter(var,'sepal length (cm)')
```

[8]:
```
<AxesSubplot:xlabel='sepal width (cm)', ylabel='sepal length (cm)'>
```



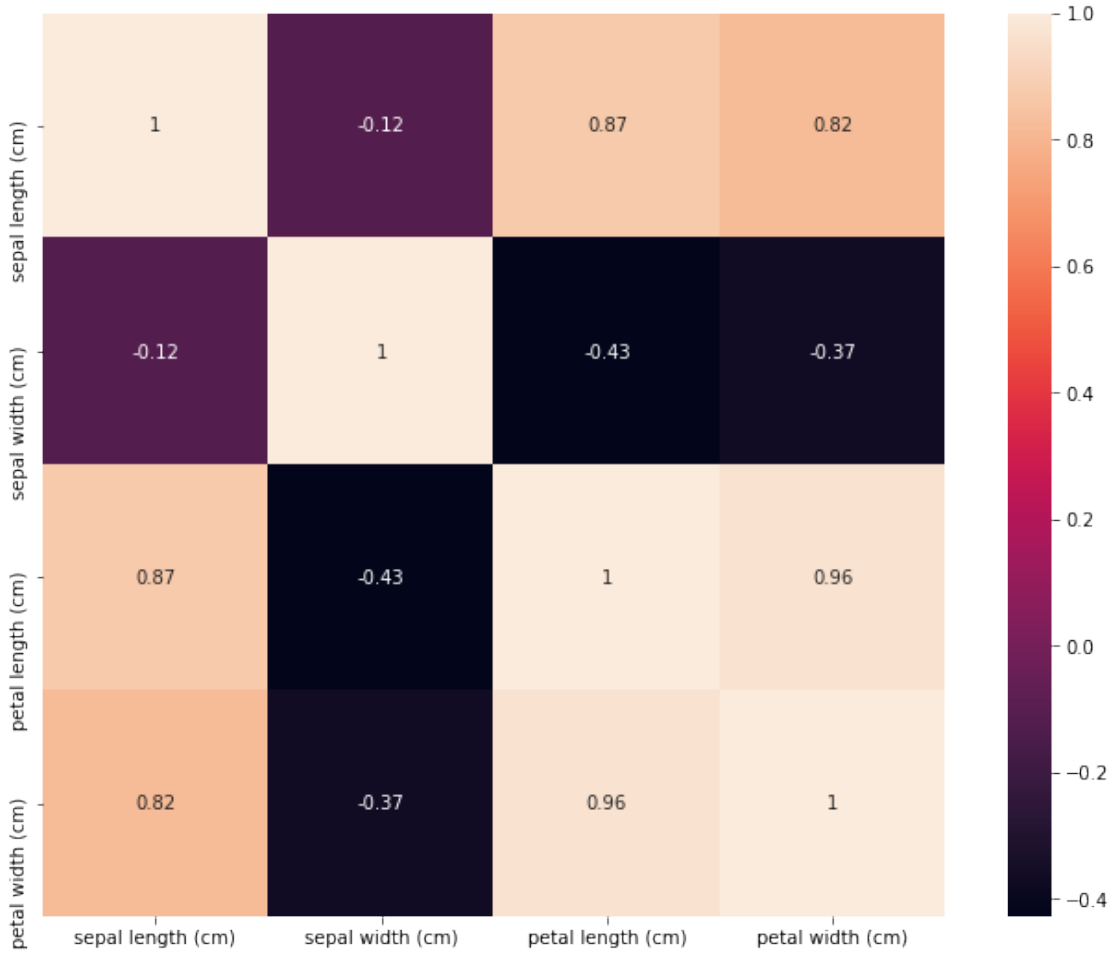
[14]:
```
data[['sepal length (cm)', 'sepal width (cm)']].corr()
```

[14]:

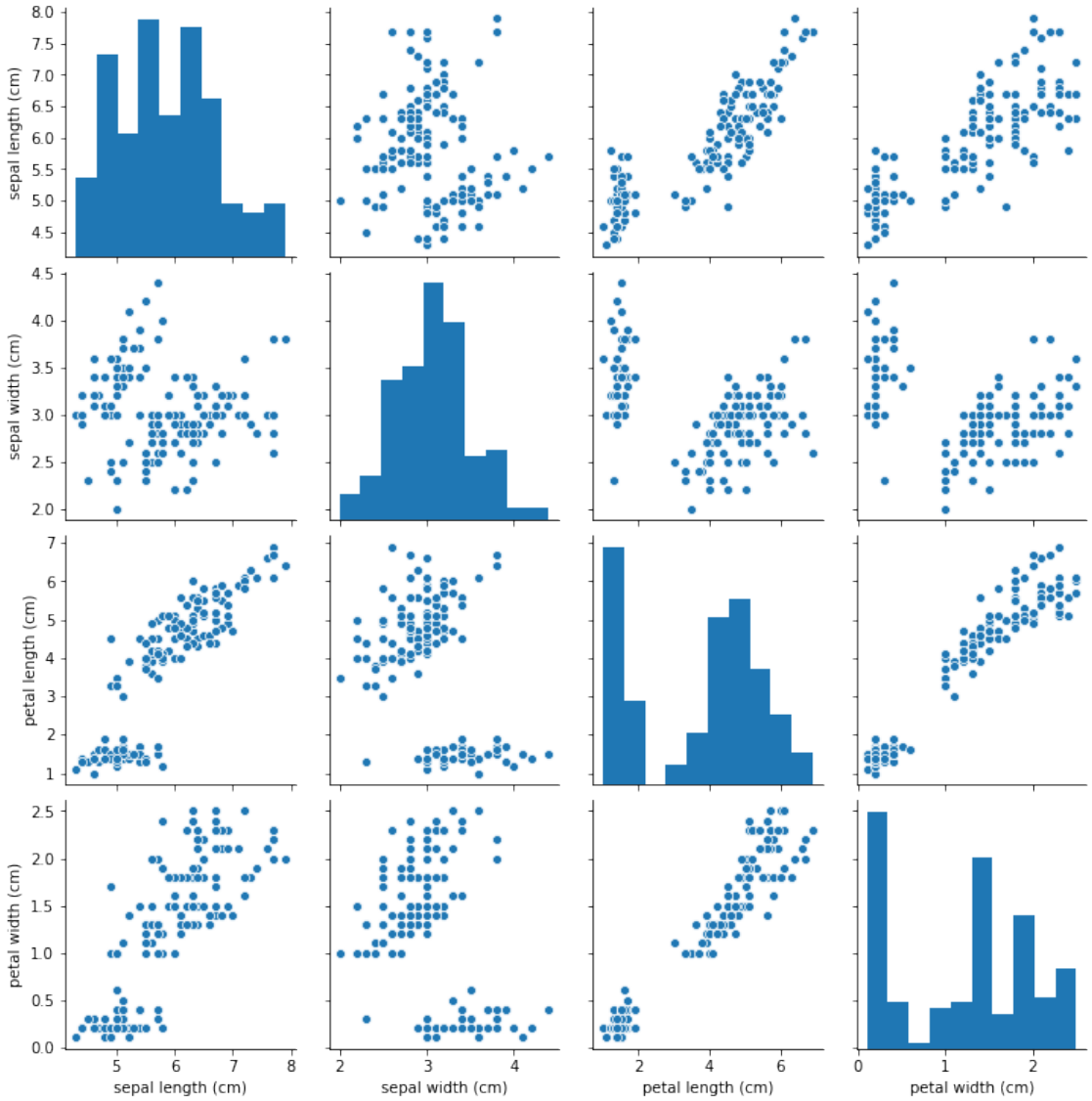
	sepal length (cm)	sepal width (cm)
sepal length (cm)	1.00000	-0.11757
sepal width (cm)	-0.11757	1.00000

[9]:
```
f, ax = plt.subplots(figsize = (12,9))
sns.heatmap(data.corr(), square=True, annot=True)
```

[9]: <AxesSubplot:>



[12]:
```
#scatterplot
cols= ['sepal length (cm)', 'sepal width (cm)', 'petal length (cm)',
       'petal width (cm)']
sns.pairplot(data[cols]);
```


sepal length (cm)
sepal width (cm)
petal length (cm)
petal width (cm)

[ ]: