 **teetangh** Completed Pandas Data Visualisation

865d005 yesterday

1 contributor

Download History

5.4 MB

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

from numpy.random import randn, randint, uniform, sample
```

```
In [2]: df = pd.DataFrame(randn(1000), index=pd.date_range('2019-06-07', periods=1000), columns=['value'])
ts = pd.Series(randn(1000), index=pd.date_range('2019-06-07', periods=1000))
```

```
In [3]: df['value'] = df['value'].cumsum()
df.head()
```

```
Out[3]:
```

	value
2019-06-07	-0.069767
2019-06-08	0.855426
2019-06-09	2.292657
2019-06-10	2.361732
2019-06-11	2.128053

```
In [4]: ts = ts.cumsum()
ts.head()
```

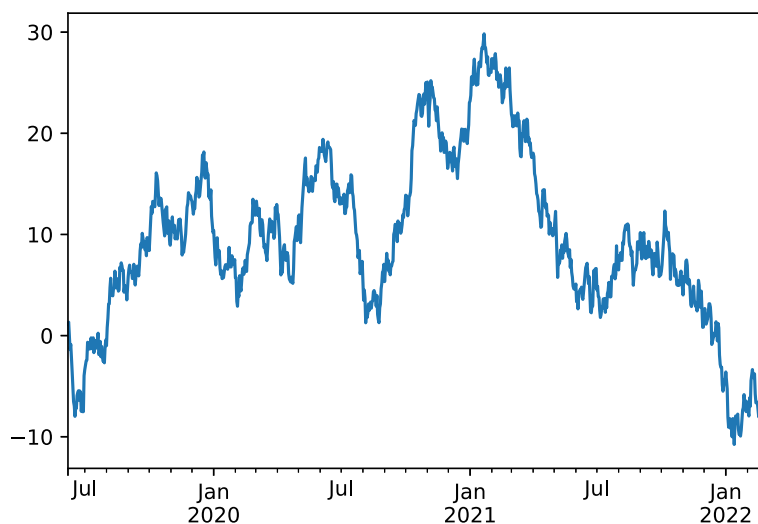
```
Out[4]: 2019-06-07    1.157698
2019-06-08    1.330188
2019-06-09    0.046533
2019-06-10   -1.386425
2019-06-11   -0.871049
Freq: D, dtype: float64
```

```
In [5]: type(df), type(ts)
```

```
Out[5]: (pandas.core.frame.DataFrame, pandas.core.series.Series)
```

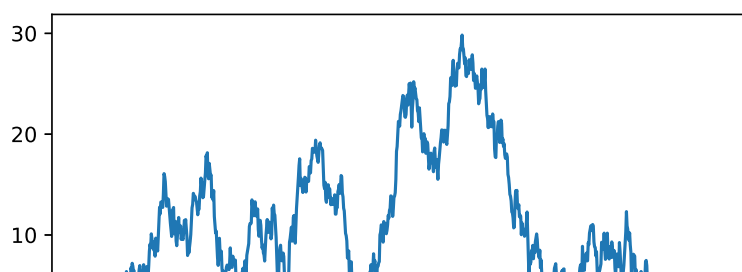
```
In [6]: ts.plot()
```

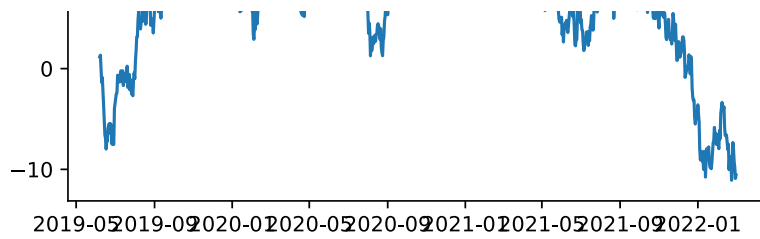
```
Out[6]: <matplotlib.axes._subplots.AxesSubplot at 0x2427a850>
```



```
In [7]: plt.plot(ts)
```

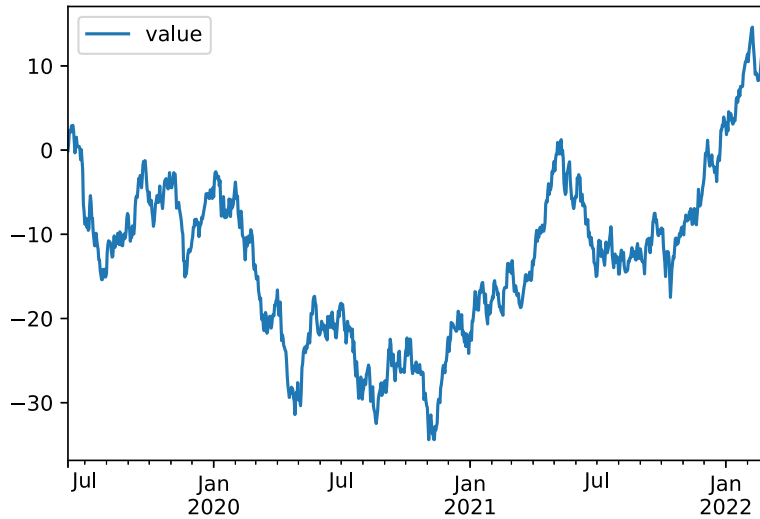
```
Out[7]: [<matplotlib.lines.Line2D at 0x2538ece8>]
```





In [8]: `df.plot()`

Out[8]: `<matplotlib.axes._subplots.AxesSubplot at 0x25a54f70>`



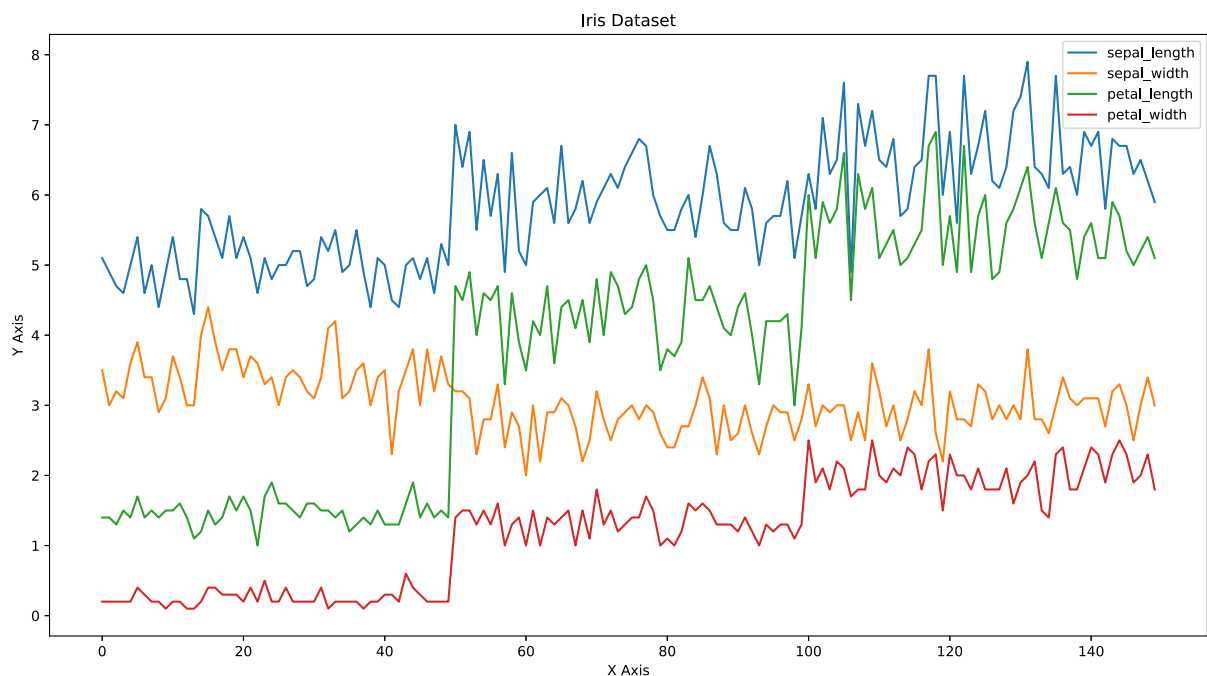
In [9]: `iris = sns.load_dataset('iris')`
`iris.head()`

Out[9]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

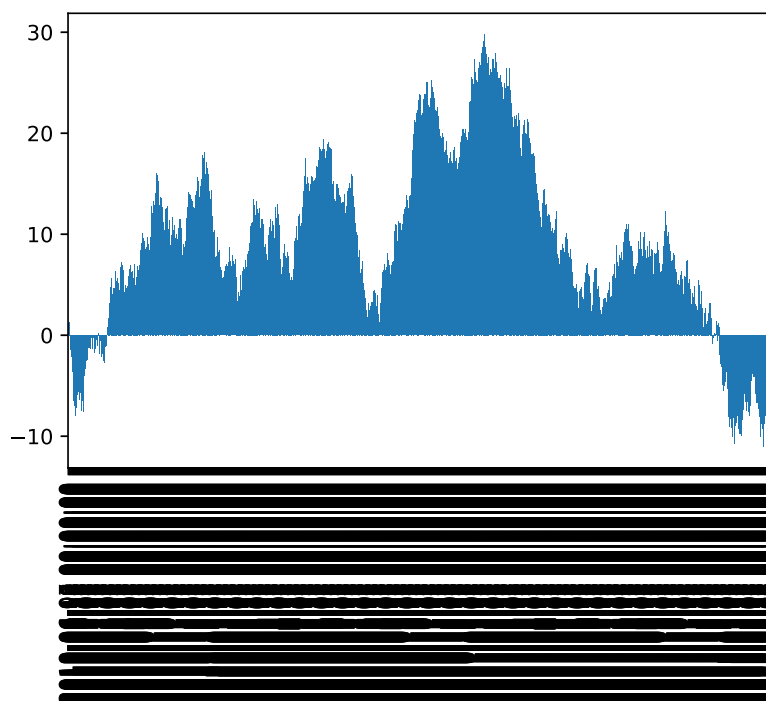
In [10]: `ax = iris.plot(figsize=(15,8),title="Iris Dataset")`
`ax.set_xlabel('X Axis')`
`ax.set_ylabel('Y Axis')`

Out[10]: `Text(0, 0.5, 'Y Axis')`



```
In [11]: ts.plot(kind = 'bar')
```

```
Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0x25b114f0>
```



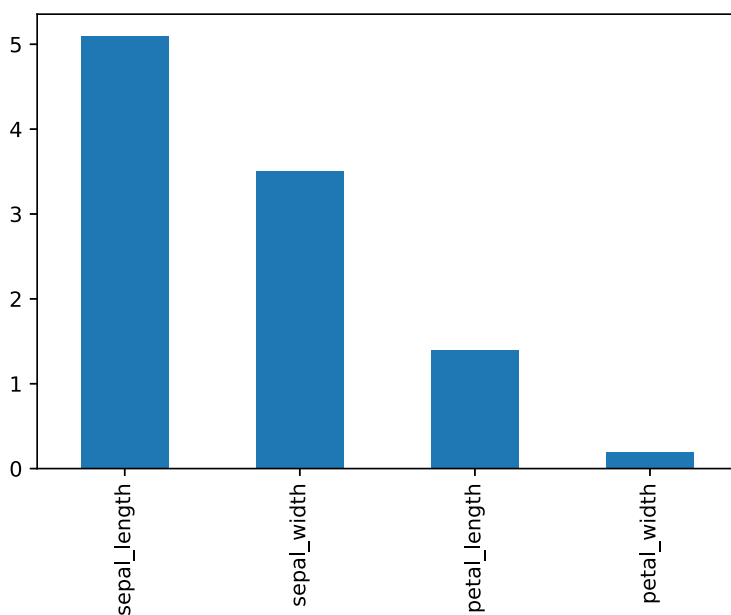
```
In [12]: df = iris.drop(['species'],axis=1)
df.head()
```

```
Out[12]:
```

	sepal_length	sepal_width	petal_length	petal_width
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

```
In [13]: ## Same Result
# df.iloc[0].plot(kind='bar')
df.iloc[0].plot.bar()
```

```
Out[13]: <matplotlib.axes._subplots.AxesSubplot at 0x25b1f730>
```



```
In [14]: titanic = sns.load_dataset('titanic')
titanic.head()
```

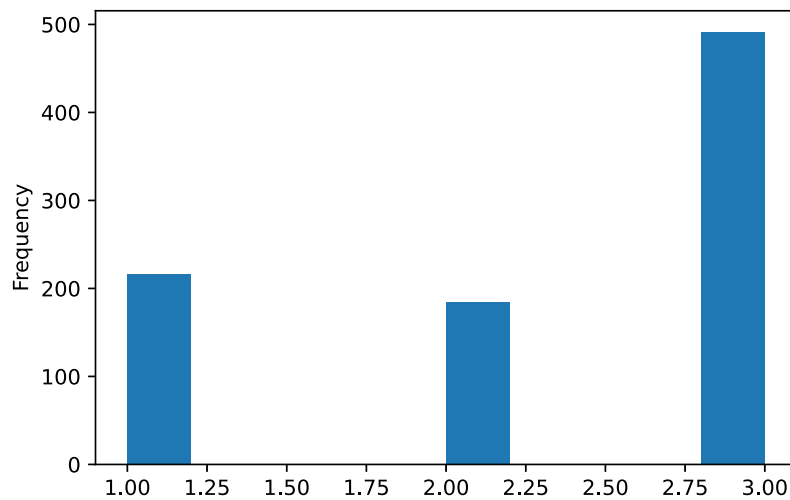
```
Out[14]:
```

	survived	passenger_id	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town
--	----------	--------------	-----	-----	-------	-------	------	----------	-------	-----	------------	------	-------------

	survived	pclass	sex	age	sibsp	parth	fare	embarked	class	who	adult_male	deck	embark_town
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton

```
In [15]: # titanic['pclass'].plot(kind='bar') # Too Dense
titanic['pclass'].plot(kind='hist')
```

```
Out[15]: <matplotlib.axes._subplots.AxesSubplot at 0x26886880>
```



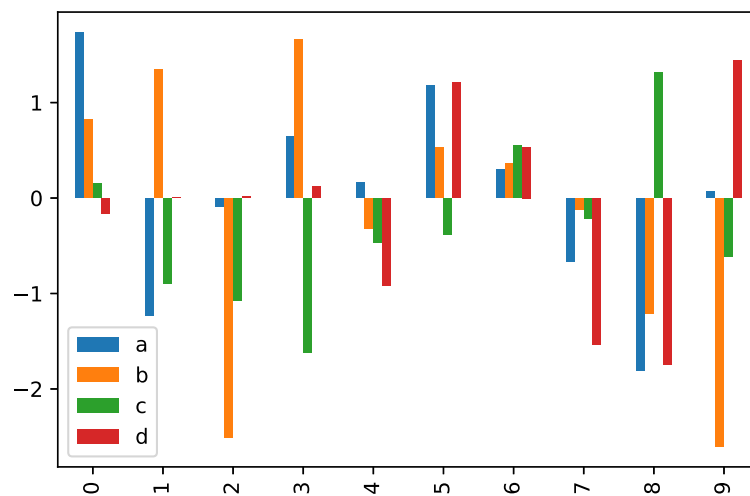
```
In [16]: df = pd.DataFrame(randn(10,4),columns=['a','b','c','d'])
df.head()
```

```
Out[16]:
```

	a	b	c	d
0	1.731810	0.820223	0.148647	-0.163526
1	-1.235545	1.345377	-0.895509	0.001187
2	-0.092366	-2.511320	-1.078872	0.015677
3	0.644796	1.657930	-1.621722	0.124866
4	0.159099	-0.314360	-0.470257	-0.912141

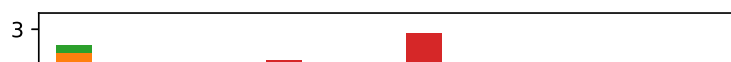
```
In [17]: df.plot.bar()
```

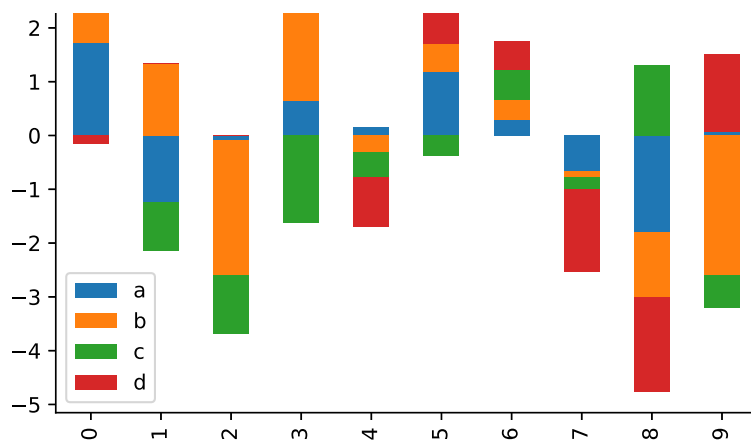
```
Out[17]: <matplotlib.axes._subplots.AxesSubplot at 0x2674e8c8>
```



```
In [18]: df.plot(kind='bar',stacked=True)
```

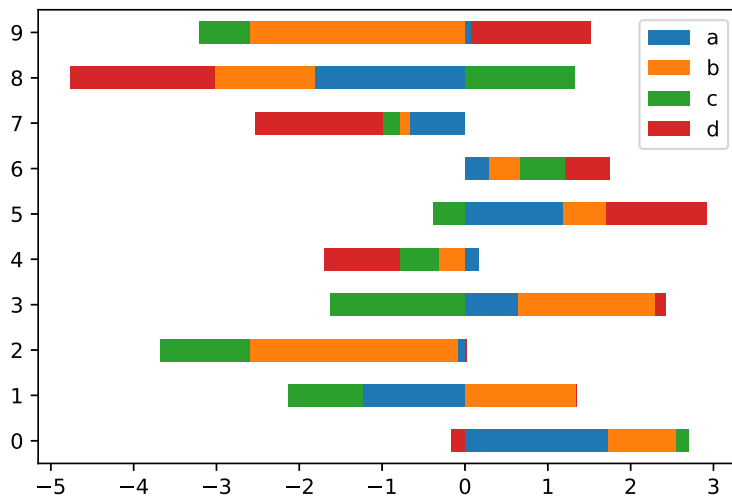
```
Out[18]: <matplotlib.axes._subplots.AxesSubplot at 0x279646b8>
```





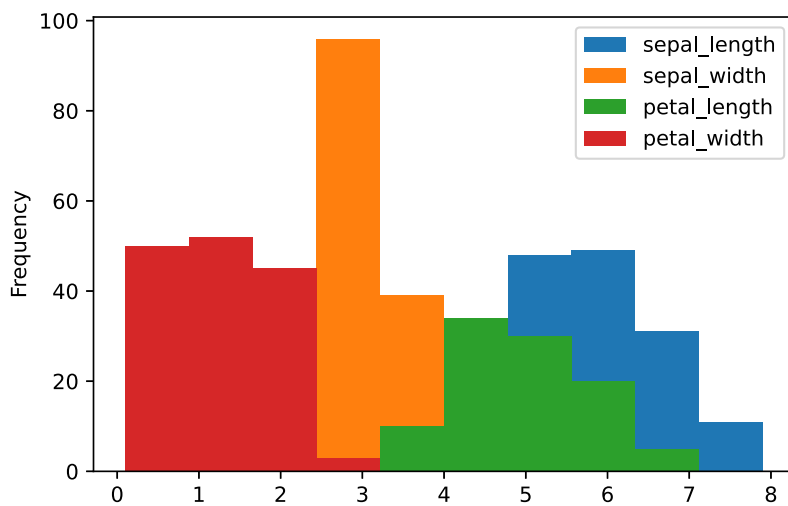
```
In [19]: df.plot.barh(stacked=True)
```

```
Out[19]: <matplotlib.axes._subplots.AxesSubplot at 0x27987fa0>
```



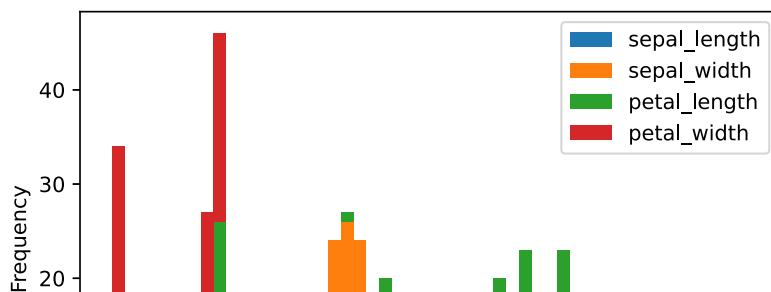
```
In [20]: # iris.plot.hist()
iris.plot(kind='hist')
```

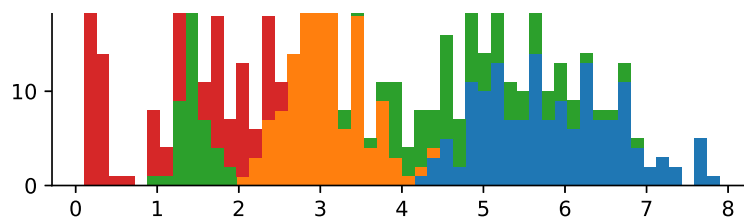
```
Out[20]: <matplotlib.axes._subplots.AxesSubplot at 0x27a4e4c0>
```



```
In [21]: iris.plot(kind='hist', stacked = True, bins= 50)
```

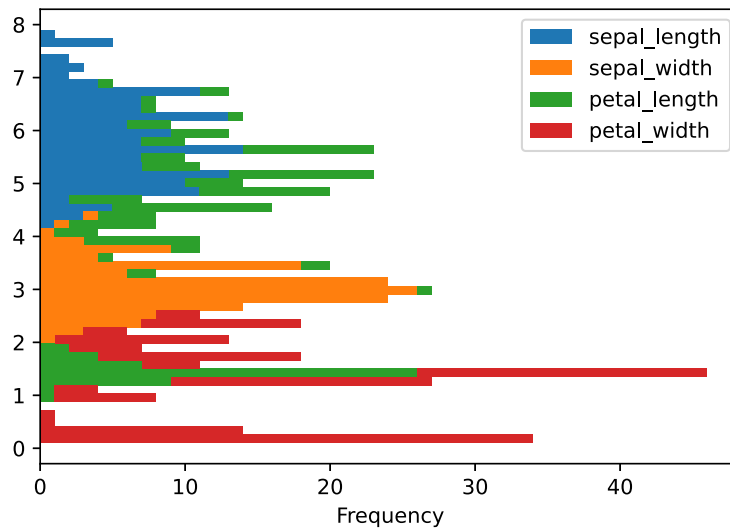
```
Out[21]: <matplotlib.axes._subplots.AxesSubplot at 0x28b8bac0>
```





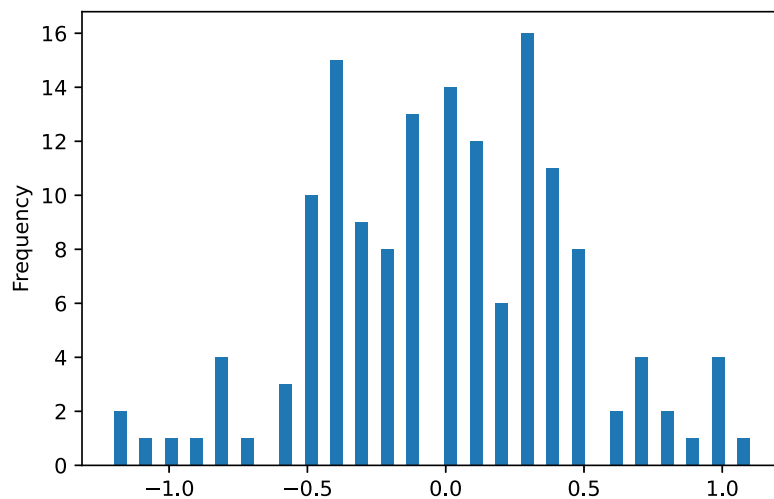
```
In [22]: iris.plot(kind='hist',stacked = True,bins= 50,orientation='horizontal')
```

```
Out[22]: <matplotlib.axes._subplots.AxesSubplot at 0x2864fcb8>
```



```
In [23]: iris['sepal_width'].diff().plot(kind='hist',stacked = True,bins= 50)
```

```
Out[23]: <matplotlib.axes._subplots.AxesSubplot at 0x2873b388>
```



```
In [24]: df = iris.drop(['species'],axis=1)
```

```
In [25]: df.diff()
```

```
Out[25]:
```

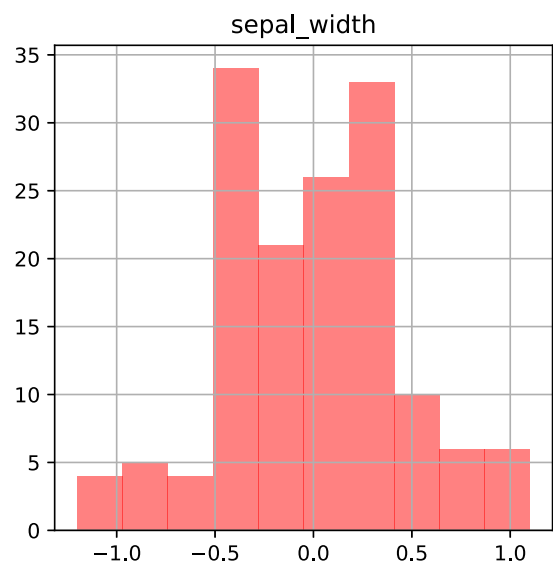
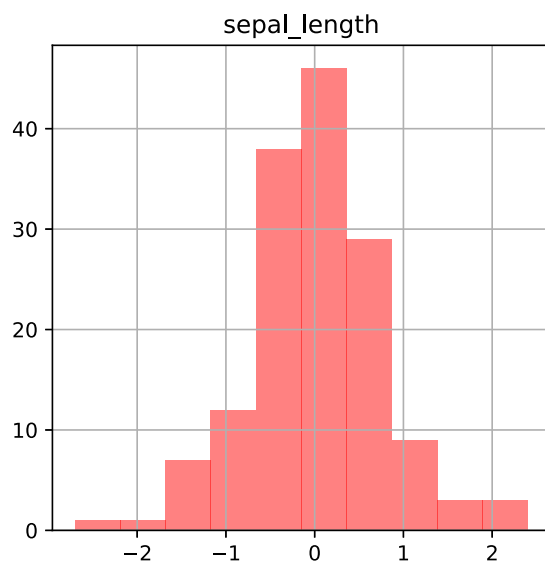
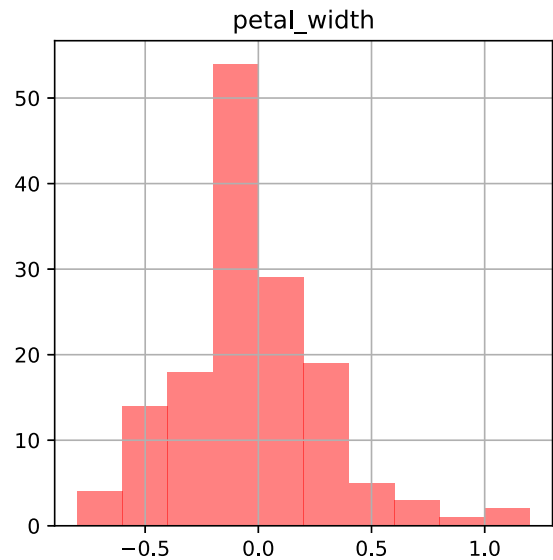
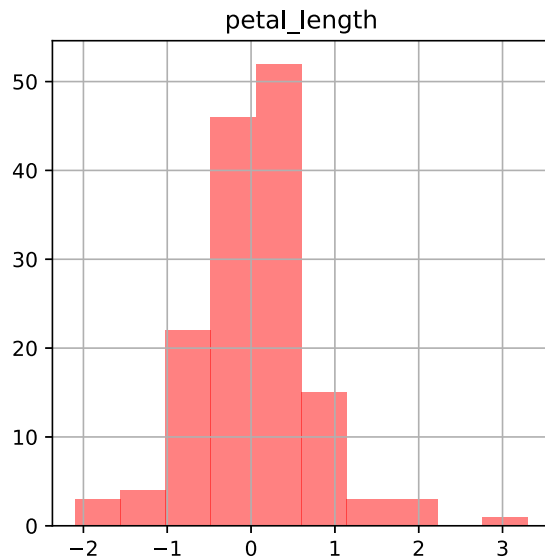
	sepal_length	sepal_width	petal_length	petal_width
0	NaN	NaN	NaN	NaN
1	-0.2	-0.5	0.0	0.0
2	-0.2	0.2	-0.1	0.0
3	-0.1	-0.1	0.2	0.0
4	0.4	0.5	-0.1	0.0
...
145	0.0	-0.3	-0.5	-0.2
146	-0.4	-0.5	-0.2	-0.4
147	0.2	0.5	0.2	0.1
148	-0.3	0.4	0.2	0.3

149	-0.3	-0.4	-0.3	-0.5
-----	------	------	------	------

150 rows × 4 columns

```
In [26]: df.diff().hist(color='r',alpha=0.5,figsize=(10,10))
```

```
Out[26]: array([[<matplotlib.axes._subplots.AxesSubplot object at 0x289D38F8>,
  <matplotlib.axes._subplots.AxesSubplot object at 0x285AD6B8>],
  [<matplotlib.axes._subplots.AxesSubplot object at 0x2861F958>,
  <matplotlib.axes._subplots.AxesSubplot object at 0x2875EC58>]],
  dtype=object)
```

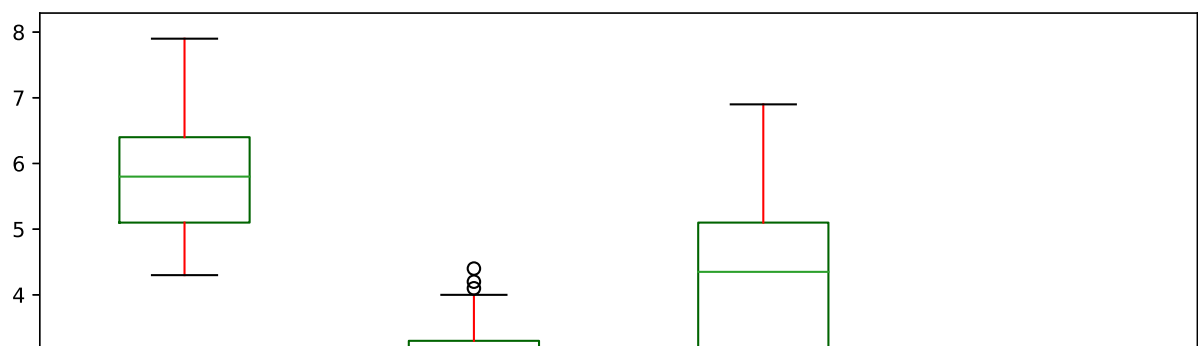


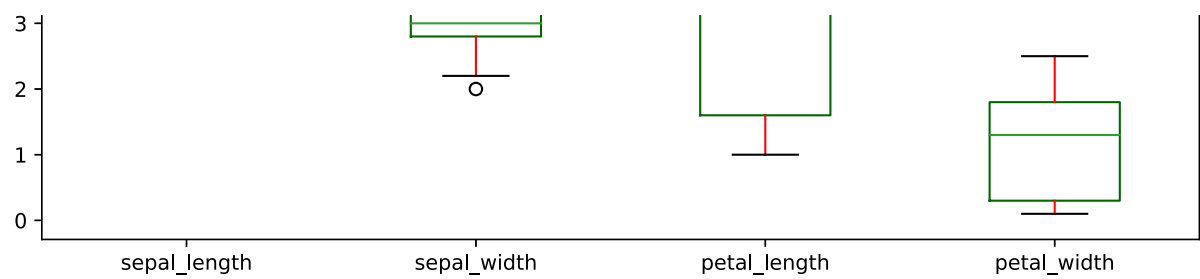
```
In [27]: color = {'boxes': 'DarkGreen', 'whiskers':'r'}
  color
```

```
Out[27]: {'boxes': 'DarkGreen', 'whiskers': 'r'}
```

```
In [28]: df.plot(kind='box',figsize=(10,5),color=color)
```

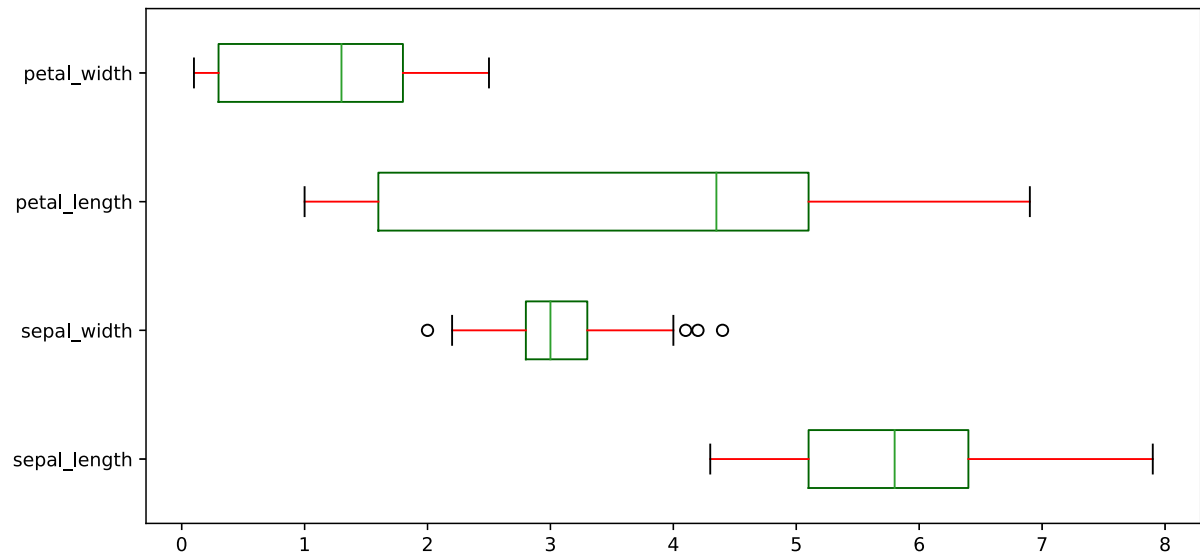
```
Out[28]: <matplotlib.axes._subplots.AxesSubplot at 0x28505370>
```





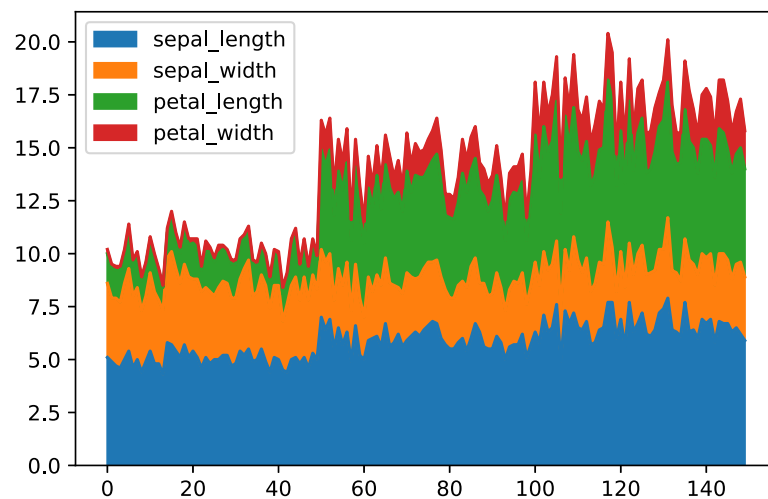
```
In [29]: df.plot(kind='box',figsize=(10,5),color='color',vert=False)
```

```
Out[29]: <matplotlib.axes._subplots.AxesSubplot at 0x28930808>
```



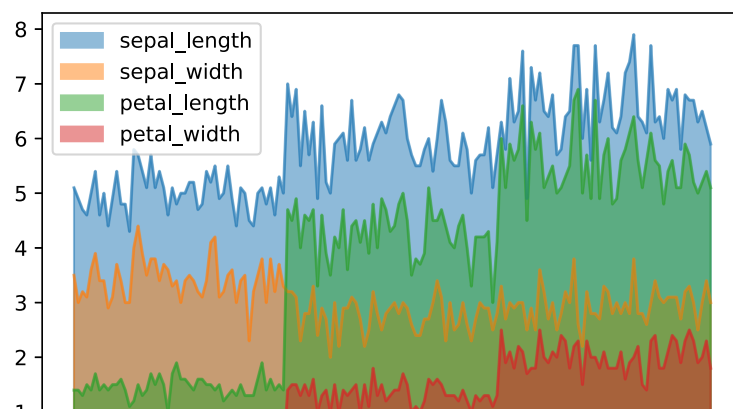
```
In [30]: # df.plot(kind='area')
df.plot.area()
```

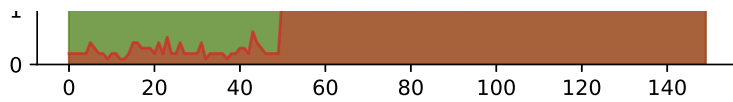
```
Out[30]: <matplotlib.axes._subplots.AxesSubplot at 0x28c0fe68>
```



```
In [31]: df.plot.area(stacked=False)
```

```
Out[31]: <matplotlib.axes._subplots.AxesSubplot at 0x289389e8>
```

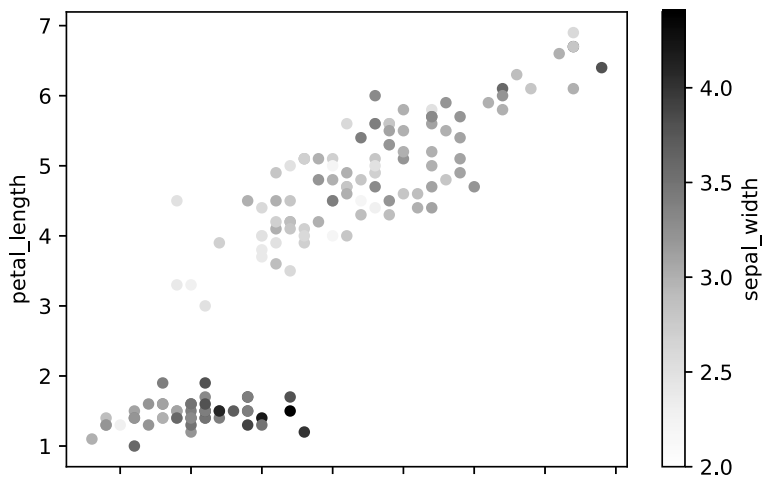




```
In [ ]: df.plot.scatter(x='sepal_length',y='petal_length')
```

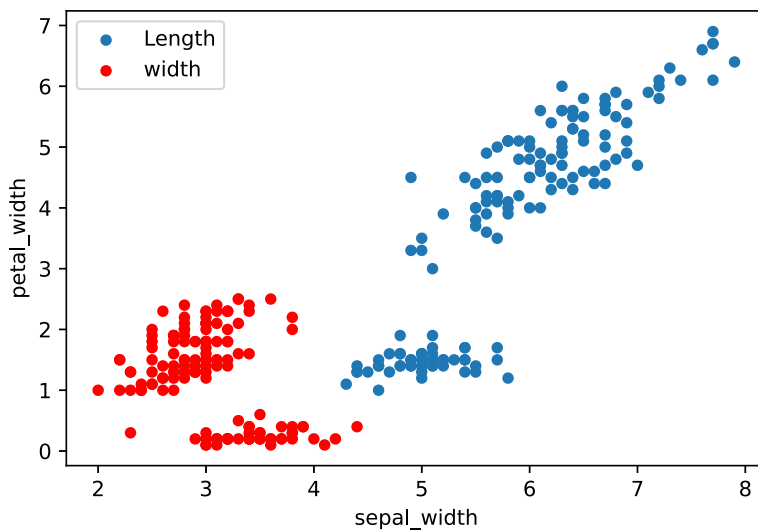
```
In [34]: df.plot.scatter(x='sepal_length',y='petal_length',c='sepal_width')
```

```
Out[34]: <matplotlib.axes._subplots.AxesSubplot at 0x290a9df0>
```



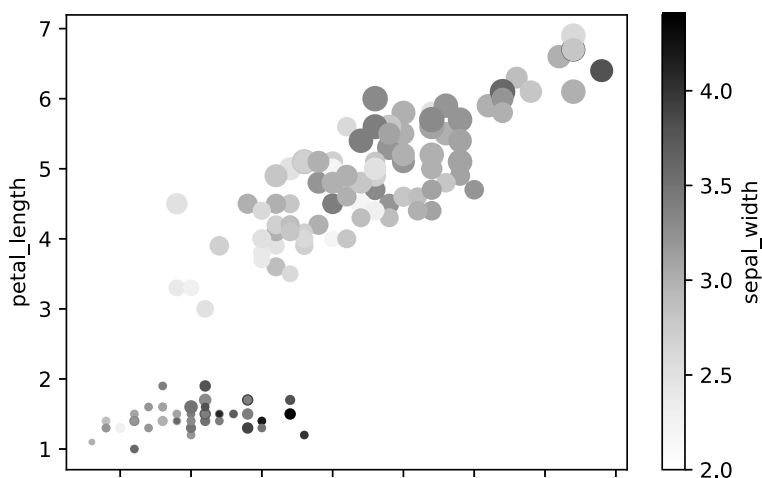
```
In [35]: ax = df.plot.scatter(x='sepal_length',y='petal_length',label='Length')
df.plot.scatter(x='sepal_width',y='petal_width',label='width',ax=ax,color='r')
```

```
Out[35]: <matplotlib.axes._subplots.AxesSubplot at 0x29115040>
```



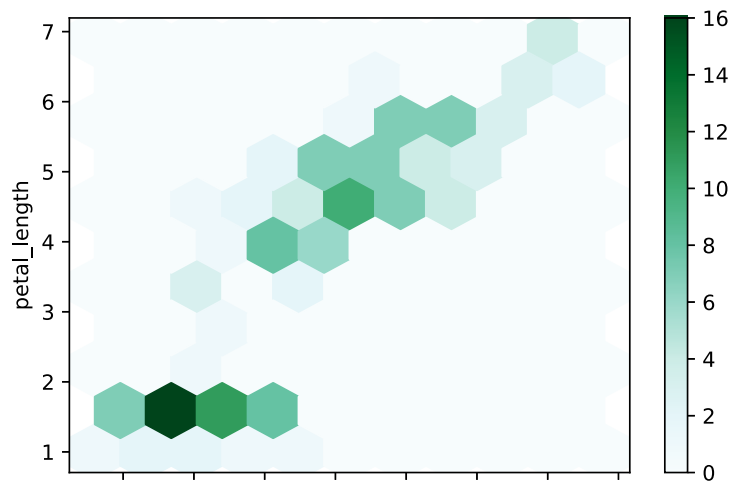
```
In [36]: df.plot.scatter(x='sepal_length',y='petal_length',c='sepal_width',s=df['petal_width']*50)
```

```
Out[36]: <matplotlib.axes._subplots.AxesSubplot at 0x29147b98>
```



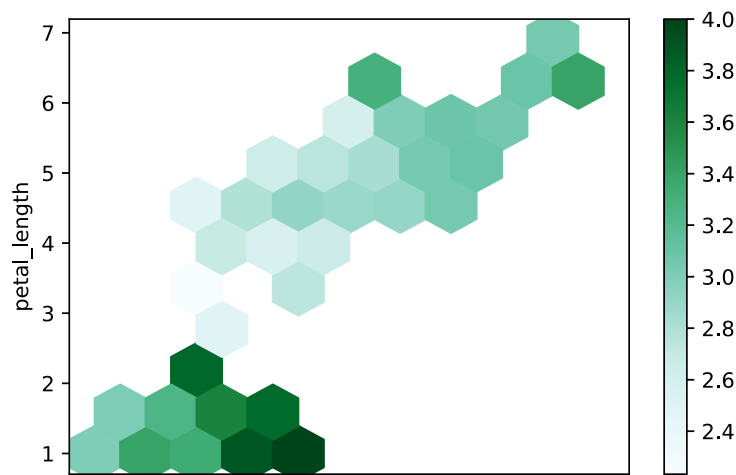
```
In [37]: df.plot.hexbin(x='sepal_length',y='petal_length',gridsize=10)
```

Out[37]: <matplotlib.axes._subplots.AxesSubplot at 0x286cd610>



```
In [38]: df.plot.hexbin(x='sepal_length',y='petal_length',gridsize=10,C='sepal_width')
```

Out[38]: <matplotlib.axes._subplots.AxesSubplot at 0x2918cac0>

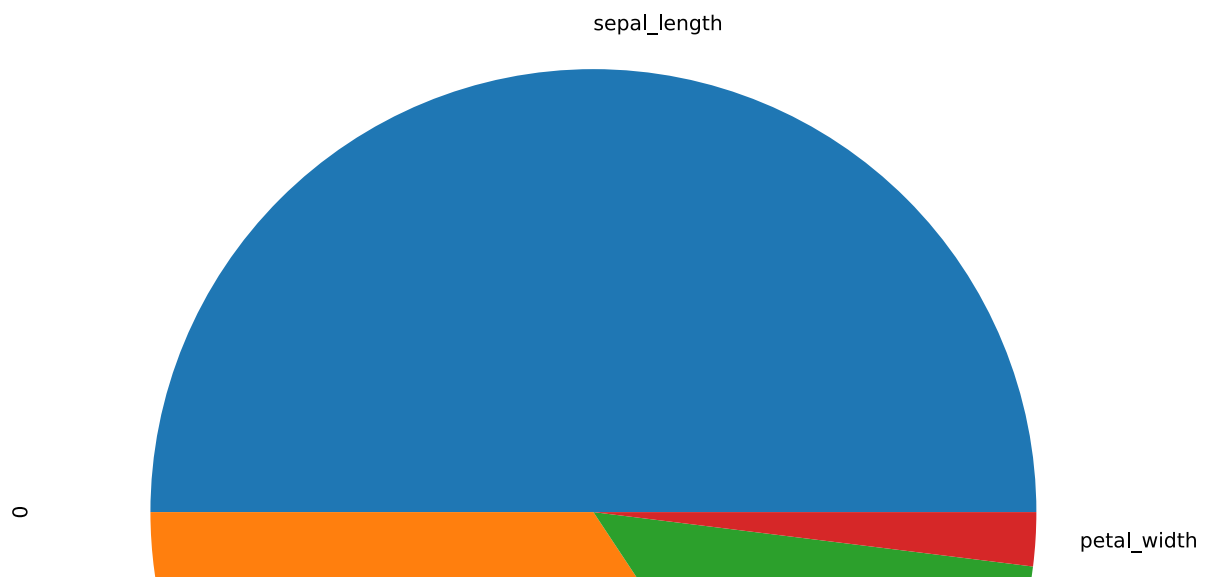


```
In [39]: d = df.iloc[0]
d
```

```
Out[39]: sepal_length    5.1
sepal_width      3.5
petal_length      1.4
petal_width      0.2
Name: 0, dtype: float64
```

```
In [41]: d.plot.pie(figsize=(10,10))
```

Out[41]: <matplotlib.axes._subplots.AxesSubplot at 0x29c8faa8>





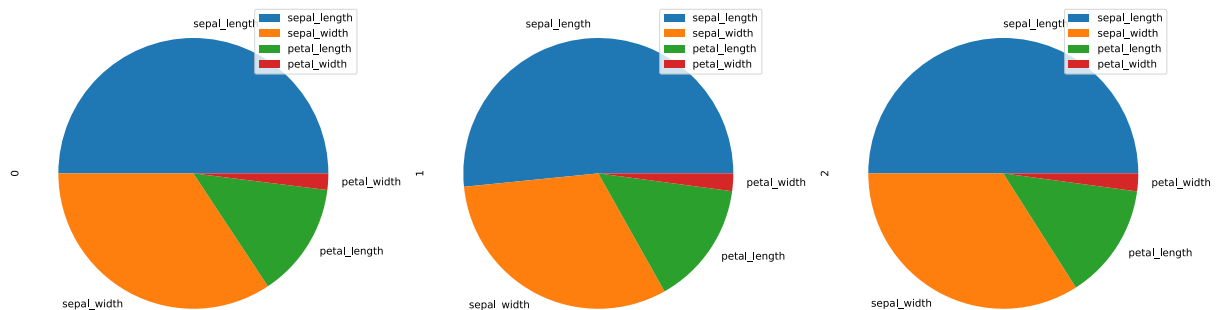
```
In [51]: d = df.head(3).T
d
```

```
Out[51]:
```

	0	1	2
sepal_length	5.1	4.9	4.7
sepal_width	3.5	3.0	3.2
petal_length	1.4	1.4	1.3
petal_width	0.2	0.2	0.2

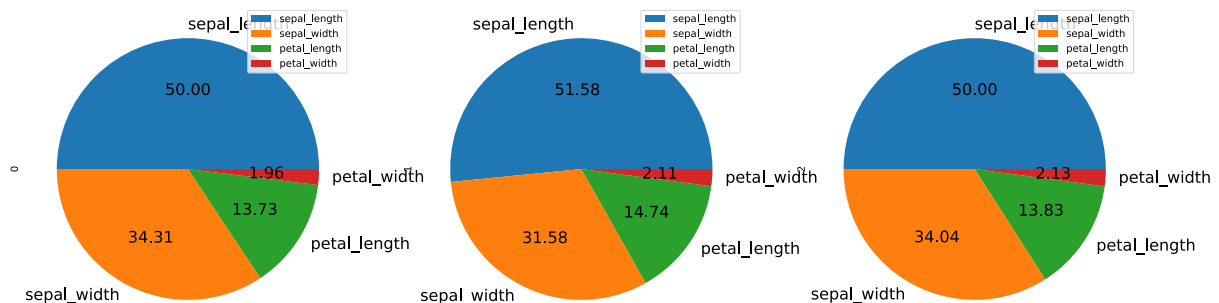
```
In [52]: d.plot.pie(subplots=True,figsize=(20,20))
```

```
Out[52]: array([<matplotlib.axes._subplots.AxesSubplot object at 0x2AB8E448>,
<matplotlib.axes._subplots.AxesSubplot object at 0x2ABA2940>,
<matplotlib.axes._subplots.AxesSubplot object at 0x2ABA28E0>],
dtype=object)
```



```
In [53]: d.plot.pie(subplots=True,figsize=(20,20),fontsize=16,autopct='%0.2f')
```

```
Out[53]: array([<matplotlib.axes._subplots.AxesSubplot object at 0x2C1F7EF8>,
<matplotlib.axes._subplots.AxesSubplot object at 0x2C1C25B0>,
<matplotlib.axes._subplots.AxesSubplot object at 0x2C1E89D0>],
dtype=object)
```

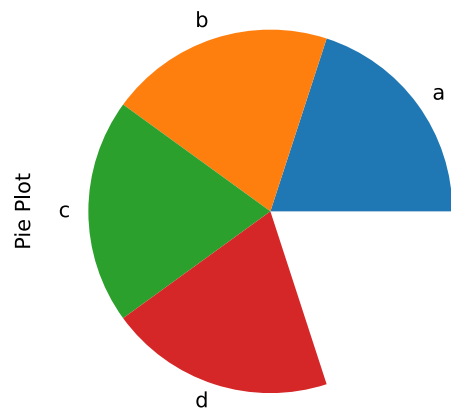


```
In [54]: [0.1]*4
```

```
Out[54]: [0.1, 0.1, 0.1, 0.1]
```

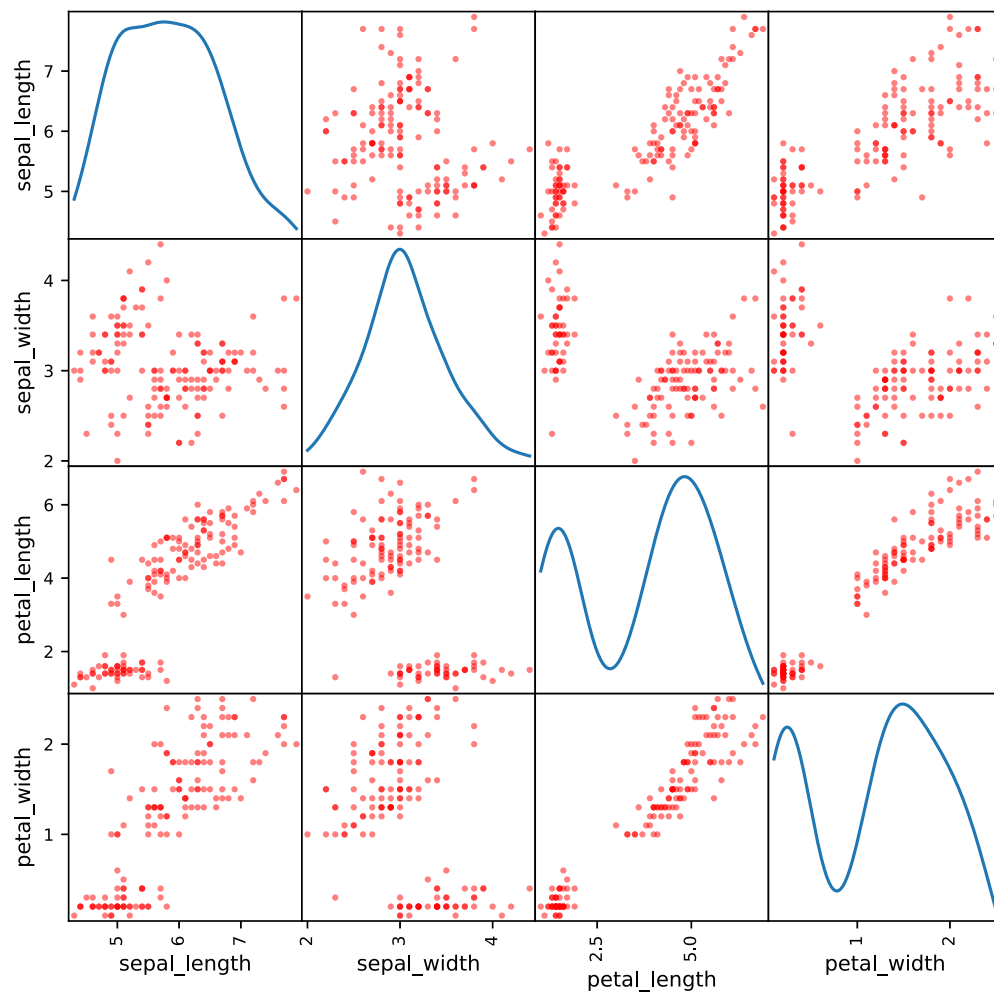
```
In [56]: series = pd.Series([0.2]*4,index=['a','b','c','d'],name='Pie Plot')
series.plot.pie()
```

```
Out[56]: <matplotlib.axes._subplots.AxesSubplot at 0x29758c70>
```



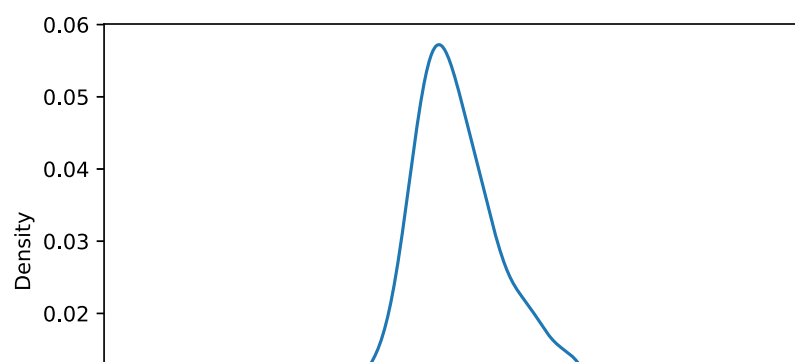
In [57]: `from pandas.plotting import scatter_matrix`

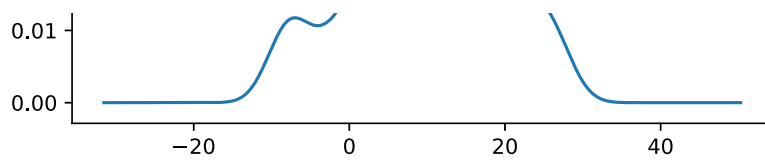
In [61]: `scatter_matrix(df,figsize=(8,8),diagonal='kde',color='r')`
`plt.show()`



In [62]: `ts.plot.kde()`

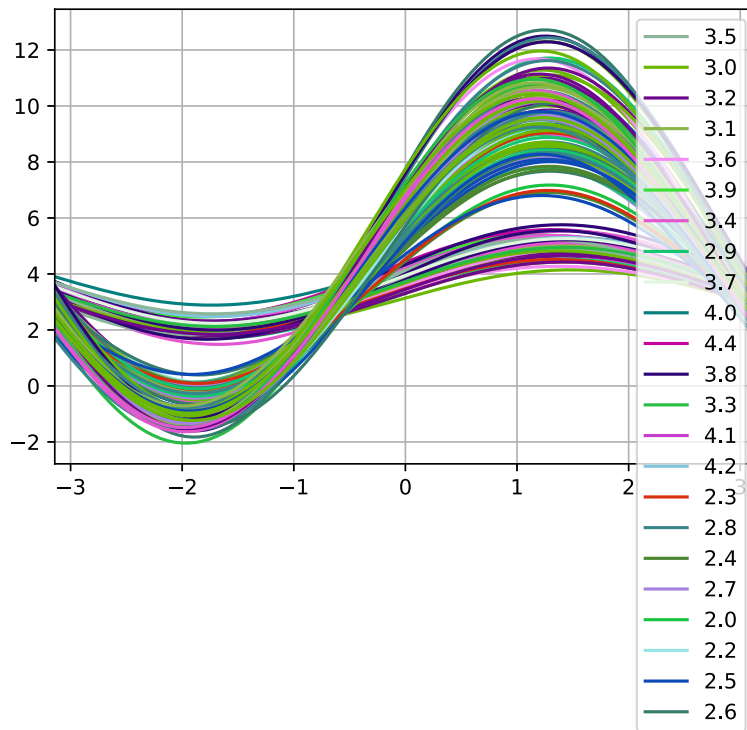
Out[62]: `<matplotlib.axes._subplots.AxesSubplot at 0x2e3ebbf8>`





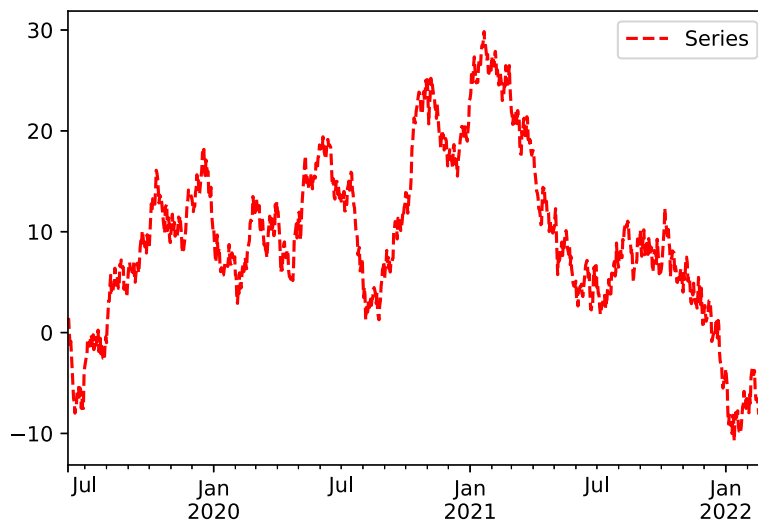
```
In [63]: from pandas.plotting import andrews_curves
andrews_curves(df, 'sepal_width')
```

```
Out[63]: <matplotlib.axes._subplots.AxesSubplot at 0x2e468e38>
```



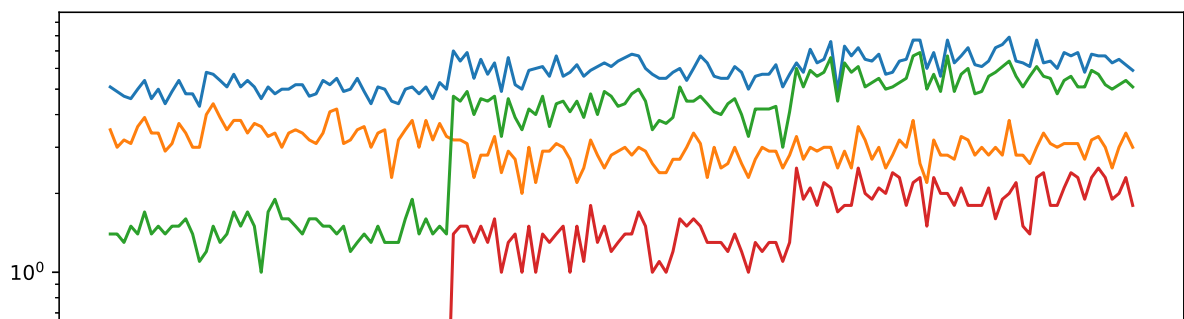
```
In [64]: ts.plot(style='r--', label='Series', legend=True)
```

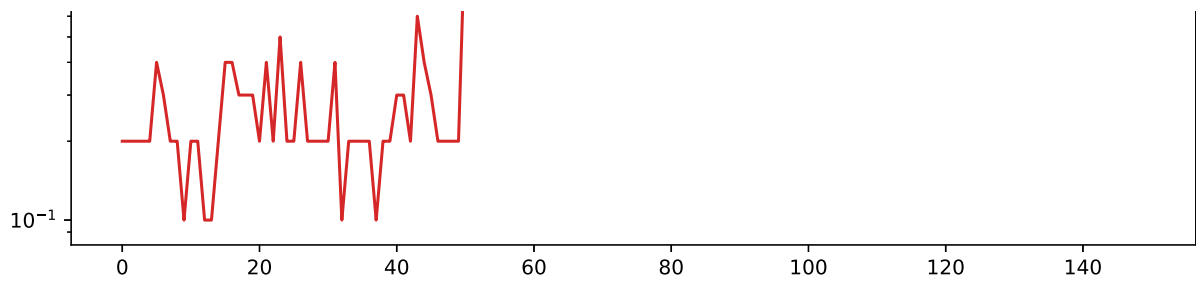
```
Out[64]: <matplotlib.axes._subplots.AxesSubplot at 0x2e750538>
```



```
In [66]: df.plot(legend=False, figsize=(10,5), logy=True)
```

```
Out[66]: <matplotlib.axes._subplots.AxesSubplot at 0x2f8ff640>
```





In [67]: `df.head()`

Out[67]:

	sepal_length	sepal_width	petal_length	petal_width
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

In [68]: `x = df.drop(['sepal_width', 'petal_width'], axis=1)`
`x.head()`

Out[68]:

	sepal_length	petal_length
0	5.1	1.4
1	4.9	1.4
2	4.7	1.3
3	4.6	1.5
4	5.0	1.4

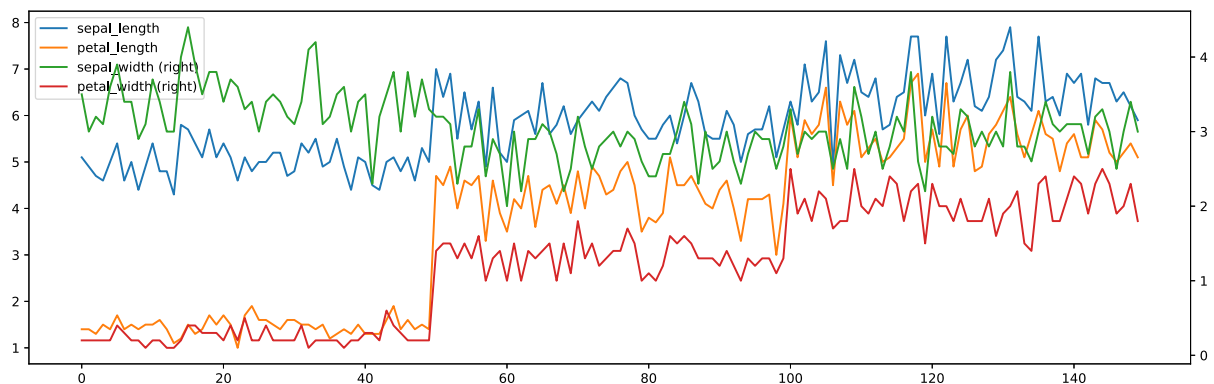
In [69]: `y = df.drop(['sepal_length', 'petal_length'], axis=1)`
`y.head()`

Out[69]:

	sepal_width	petal_width
0	3.5	0.2
1	3.0	0.2
2	3.2	0.2
3	3.1	0.2
4	3.6	0.2

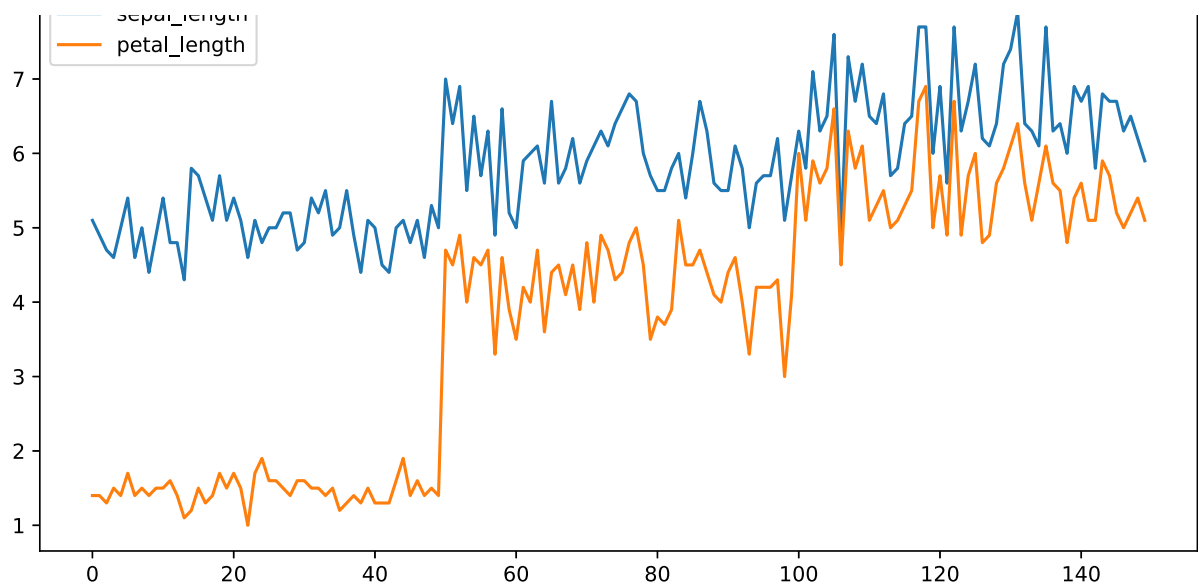
In [73]: `ax = x.plot()`
`y.plot(figsize=(16,5), secondary_y=True, ax=ax)`

Out[73]: `<matplotlib.axes._subplots.AxesSubplot at 0x2e764370>`

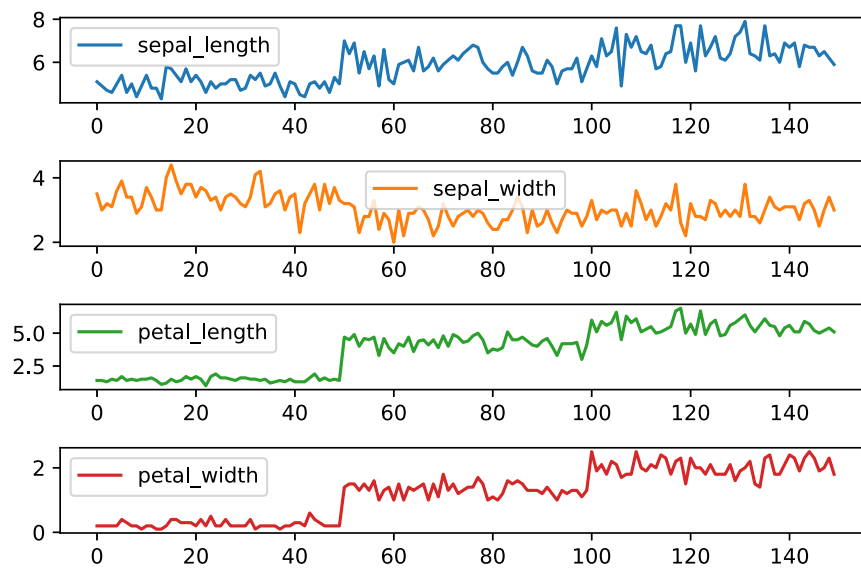


In [74]: `x.plot(figsize=(10,5), x_compat=True)`

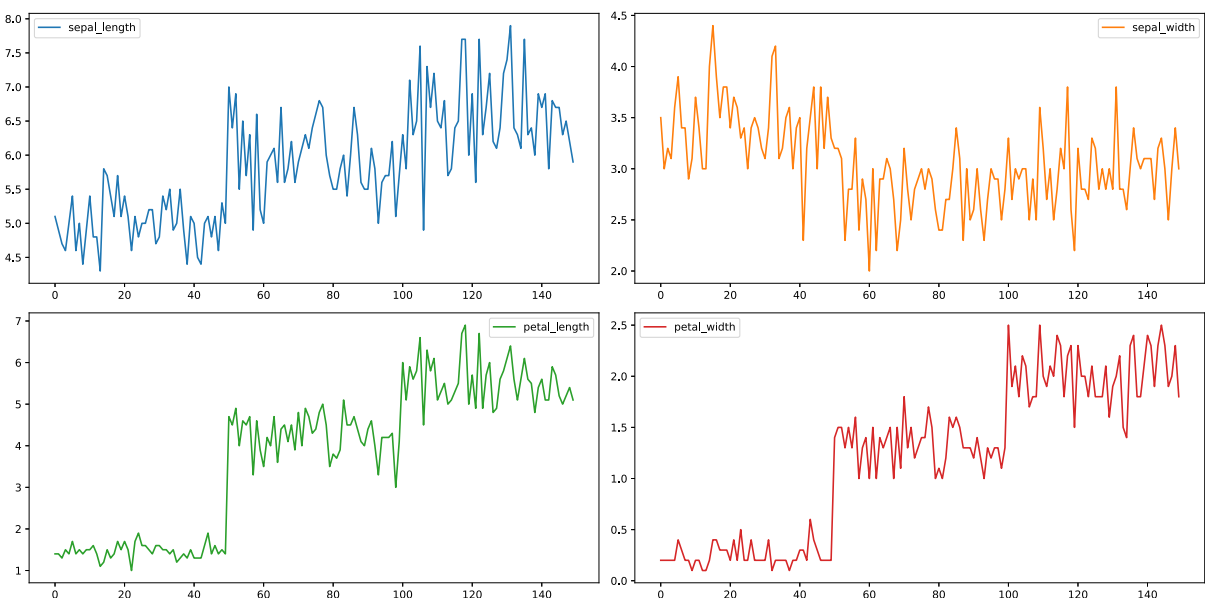
Out[74]: `<matplotlib.axes._subplots.AxesSubplot at 0x2e3e3bc8>`



```
In [79]: # df.plot(subplots=True,sharex = True)
df.plot(subplots=True,sharex = False)
plt.tight_layout()
```



```
In [81]: df.plot(subplots=True,sharex = False,layout=(2,2),figsize=(16,8))
plt.tight_layout()
```



In []:

