Pandas Data Visualization Exercise

January 16, 2020

1 Pandas Data Visualization Exercise

[20]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
df3 = pd.read_csv('df3')
%matplotlib inline
```

[2]:
```
df3.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 500 entries, 0 to 499
Data columns (total 4 columns):
a    500 non-null float64
b    500 non-null float64
c    500 non-null float64
d    500 non-null float64
dtypes: float64(4)
memory usage: 15.8 KB
```

[3]:
```
df3.head()
```

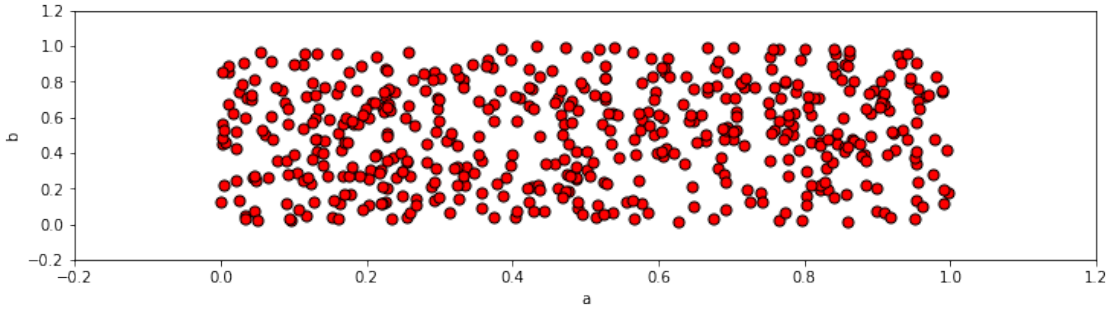
[3]:

	a	b	c	d
0	0.336272	0.325011	0.001020	0.401402
1	0.980265	0.831835	0.772288	0.076485
2	0.480387	0.686839	0.000575	0.746758
3	0.502106	0.305142	0.768608	0.654685
4	0.856602	0.171448	0.157971	0.321231

1.1 Scatter plot

[21]:
```
df3.plot.scatter(x='a', y='b', xlim=(-.2,1.2), ylim=(-.2,1.2), c='red',␣
↪figsize=(12,3),
                 marker='o', s=50, edgecolor='black')
```

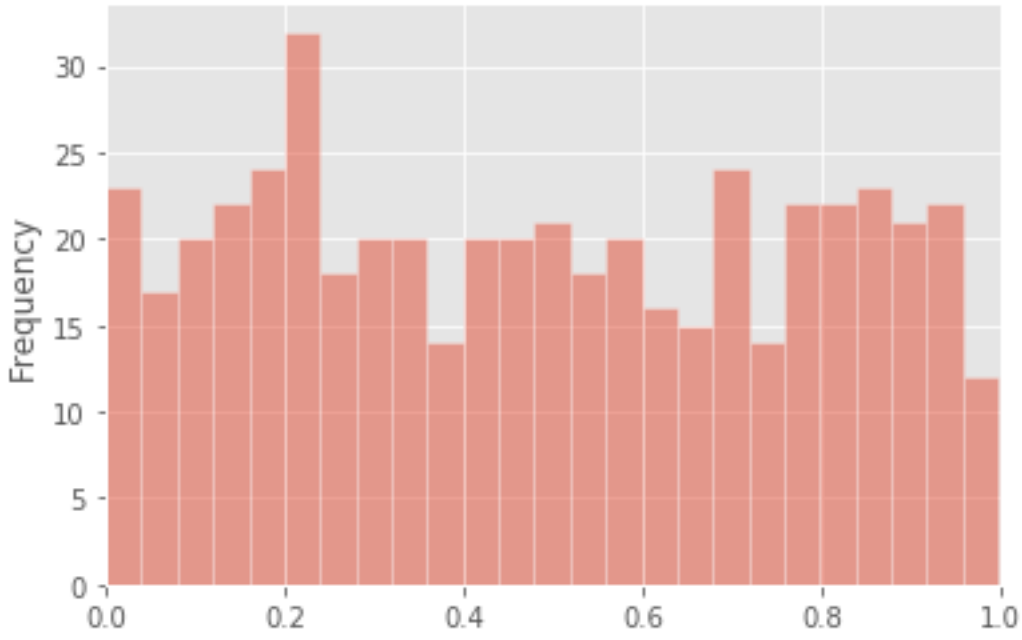
[21]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1a296b1850>
```



1.2 Histogram

[30]:
```
plt.style.use('ggplot')
df3['a'].plot.hist(xlim=(0,1), bins=25, alpha=0.5, edgecolor='white')
```

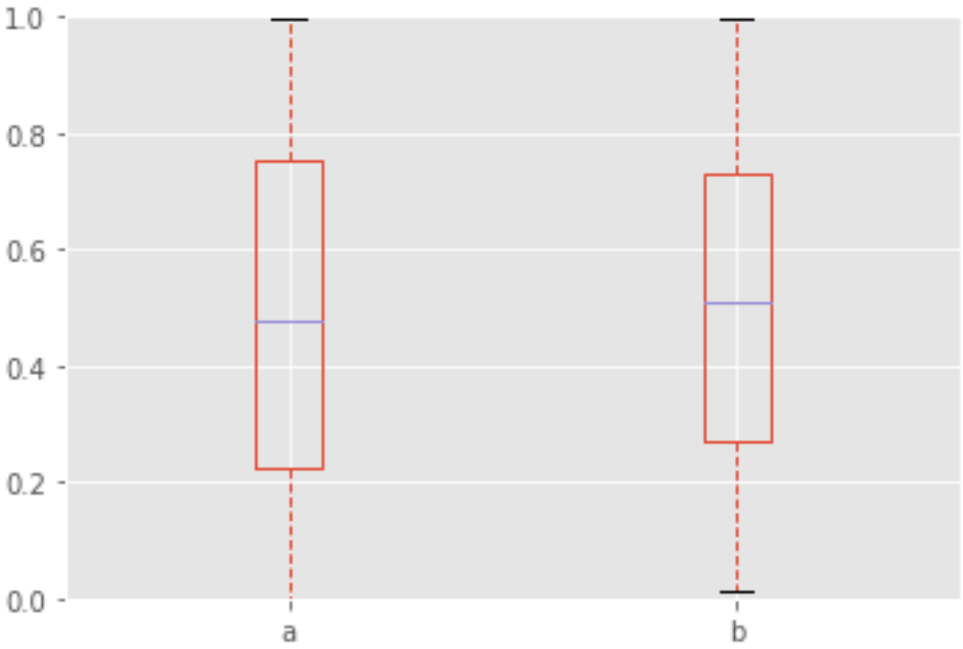
[30]: <matplotlib.axes._subplots.AxesSubplot at 0x1a29faeb90>



1.3 Box plot

[46]:
```
df3.plot.box(y=['a','b'], ylim=(0,1), whiskerprops=dict(linestyle='--'))
```

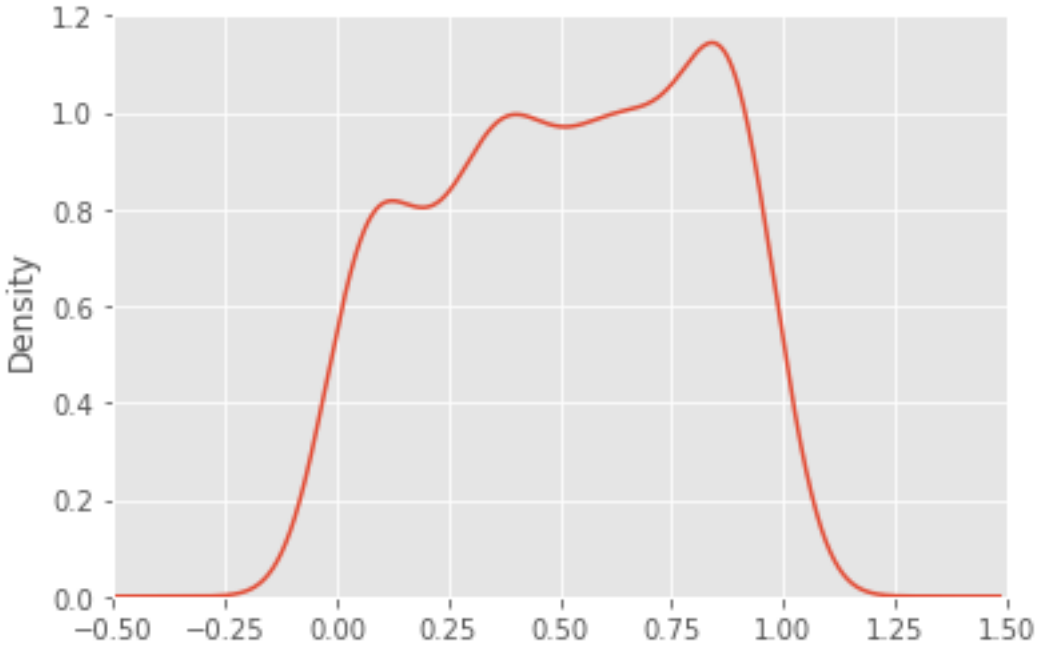
[46]: <matplotlib.axes._subplots.AxesSubplot at 0x1a2aaa3b10>



1.4 Kde plot

[49]:
```
df3['d'].plot.density(xlim=(-.5,1.5), ylim=(0,1.2))
```

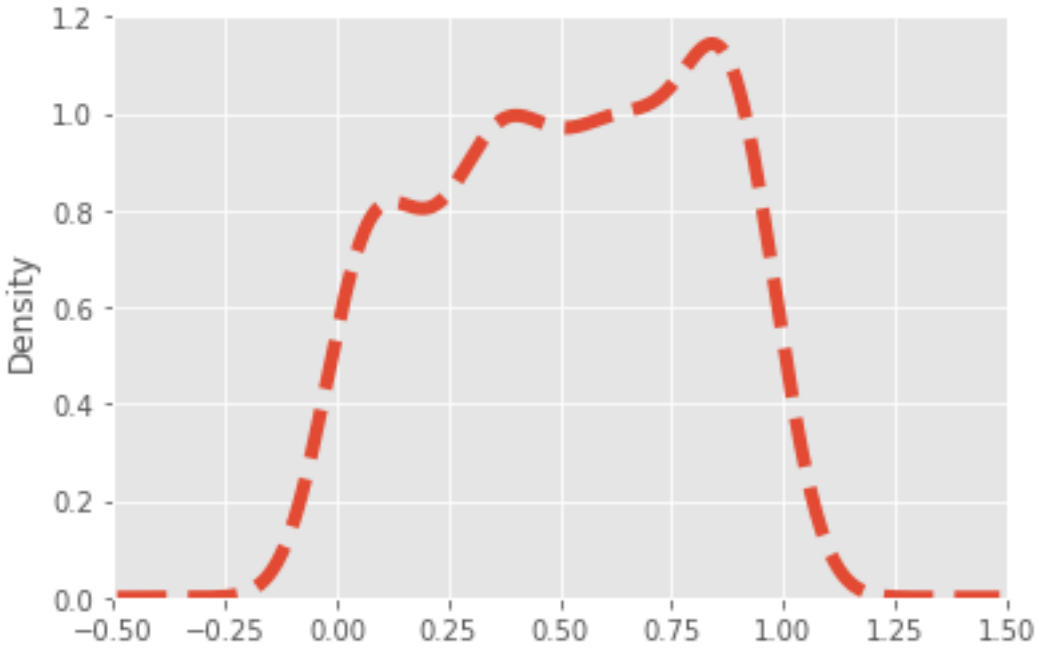
[49]: <matplotlib.axes._subplots.AxesSubplot at 0x1a2ad5d9d0>



Change line style and width

[58]:
```
df3['d'].plot.density(xlim=(-.5,1.5), ylim=(0,1.2), ls='--', lw=5)
```

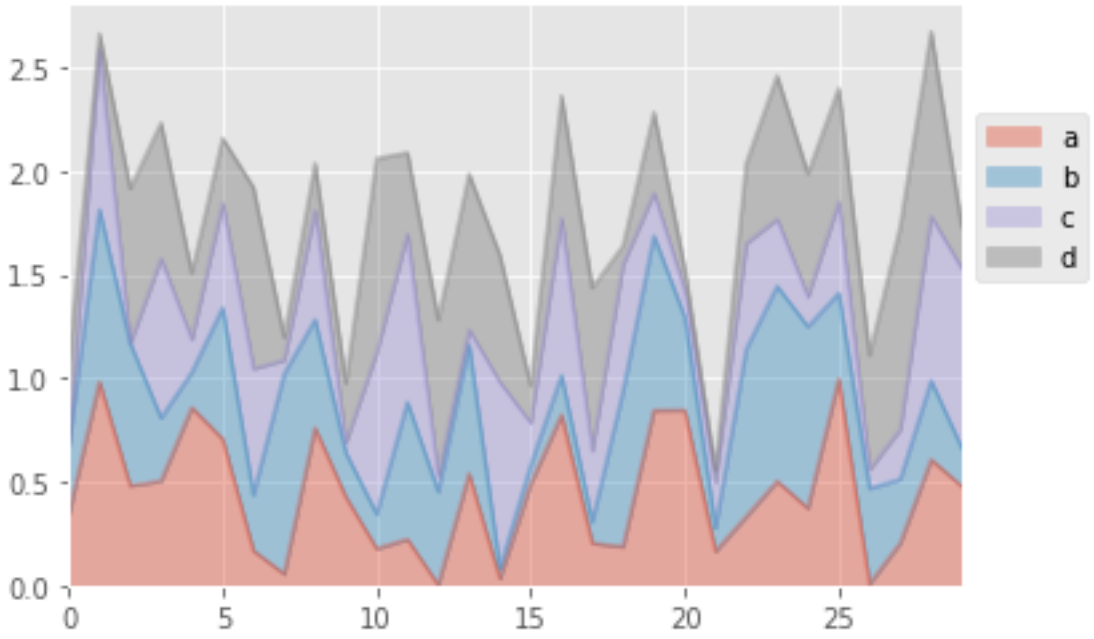
[58]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1a2b3f94d0>
```



1.5 Area plot

[99]:
```
df3.iloc[:30].plot.area(alpha=.4)
plt.legend(loc='lower left', bbox_to_anchor=(1,.5))
```

[99]:
```
<matplotlib.legend.Legend at 0x1a33c863d0>
```


2.5
2.0
1.5
1.0
0.5
0.0
0
5
10
15
20
25
a
b
c
d