Seaborn Exercises

January 16, 2020

1 Seaborn Exercises

Visualize data with Python Seaborn package

1.1 The Data

The following exercises are based on a famous titanic data set.

[1]:
```
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
```

[2]:
```
sns.set_style('darkgrid')
```

[3]:
```
titanic = sns.load_dataset('titanic')
```

[4]:
```
titanic.head()
```

[4]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	deck	embark_town	alive	alone
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	NaN	Southampton	no	False
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	C	Cherbourg	yes	False
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	NaN	Southampton	yes	True
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	C	Southampton	yes	False
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	NaN	Southampton	no	True

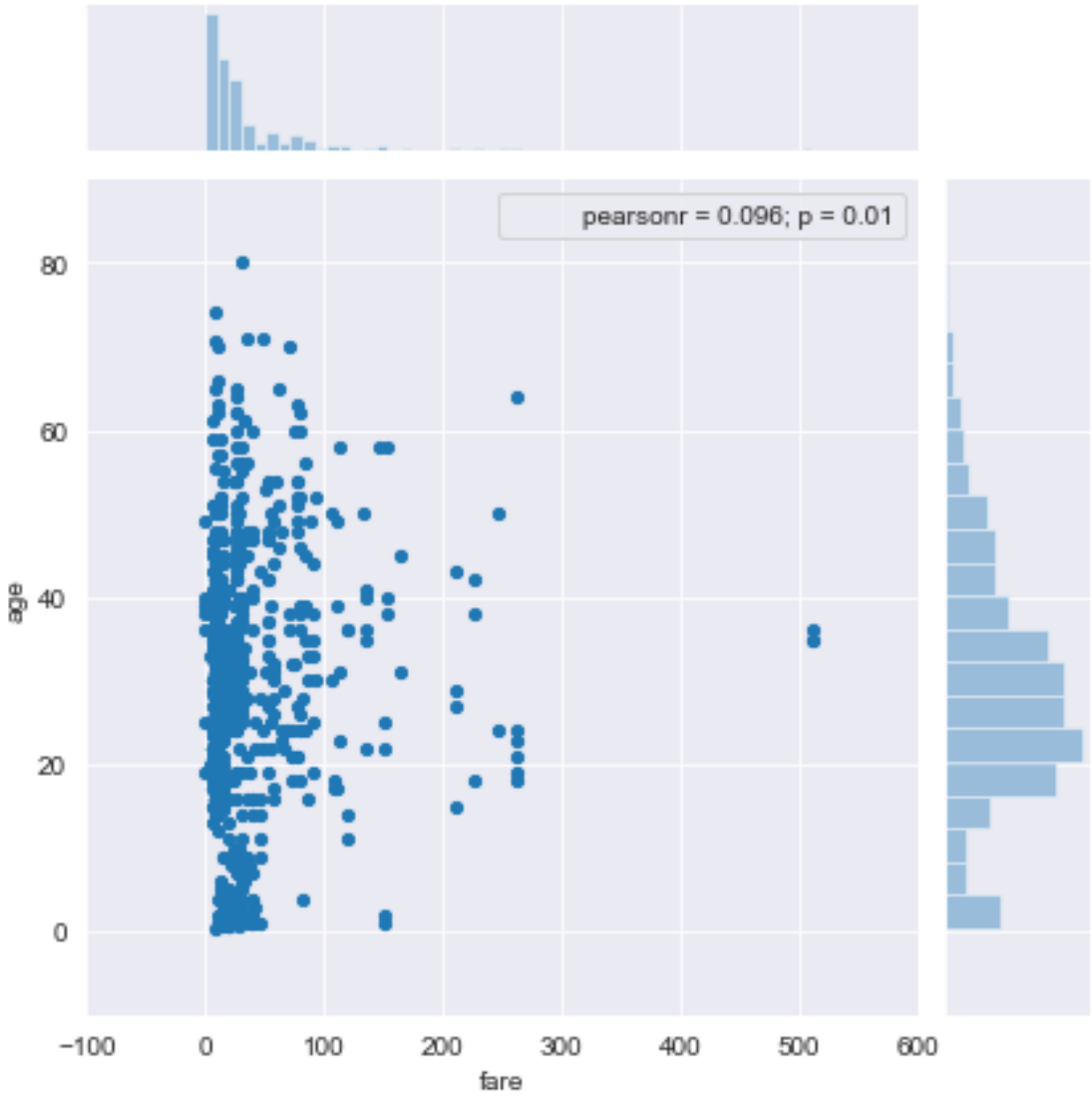
1.2 Exercises

1.2.1 Joint plot

[5]:
```
from scipy.stats import pearsonr
sns.jointplot(x='fare', y='age', data=titanic, xlim=(-100,600), ylim=(-10,90),␣
↪marker='.', s=70, stat_func=pearsonr)
```

```
/Users/stella/opt/anaconda3/lib/python3.7/site-
packages/seaborn/axisgrid.py:1847: UserWarning: JointGrid annotation is
deprecated and will be removed in a future release.
  warnings.warn(UserWarning(msg))
```

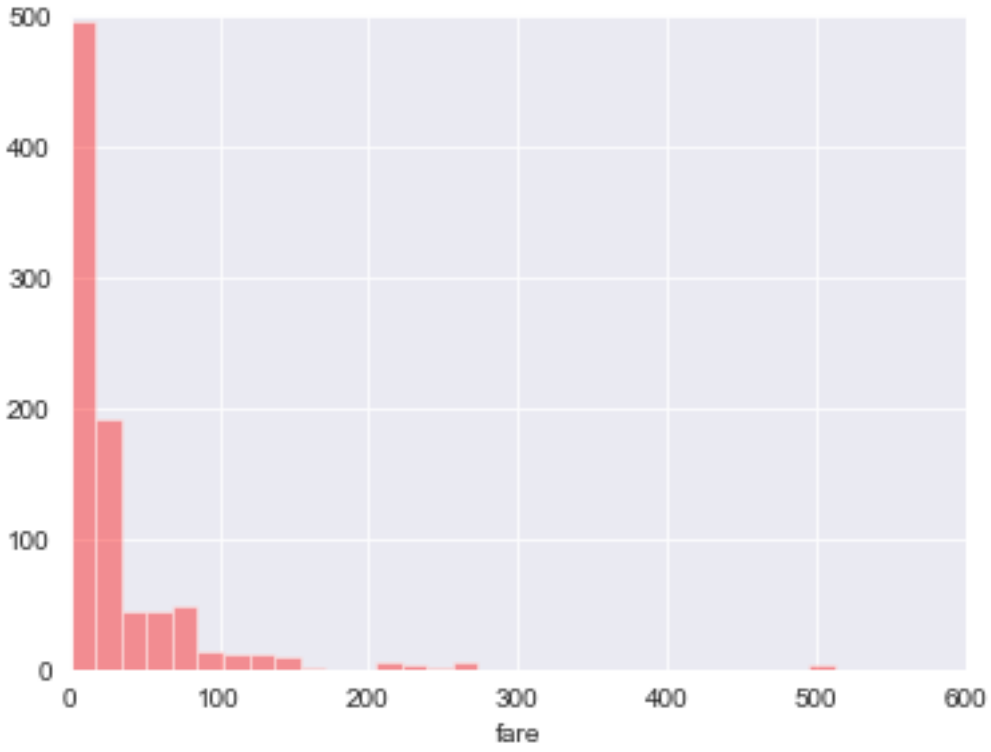
[5]: <seaborn.axisgrid.JointGrid at 0x1a254d9790>



1.2.2 Distribution plot

[6]:
```
plt.figure(figsize=(6,4.5))
sns.distplot(titanic['fare'], kde=False, bins=30, color='red')
plt.xlim(0,600)
plt.ylim(0,500)
```

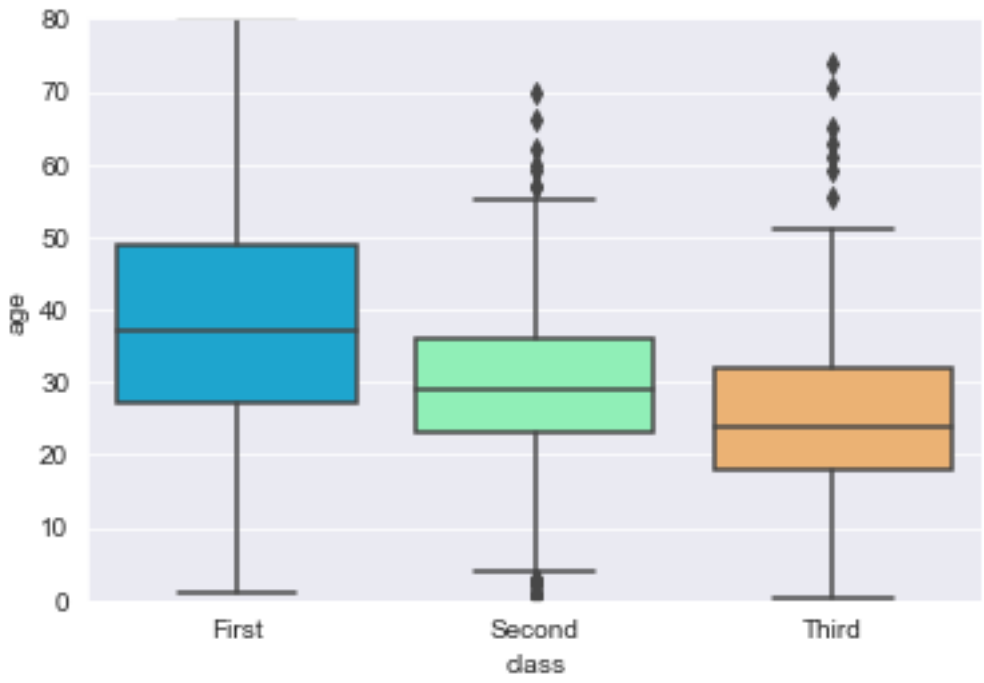
[6]: (0, 500)



1.2.3 Box plot

[7]:
```
sns.boxplot(x='class', y='age', data=titanic, palette='rainbow')
plt.ylim(0,80)
```

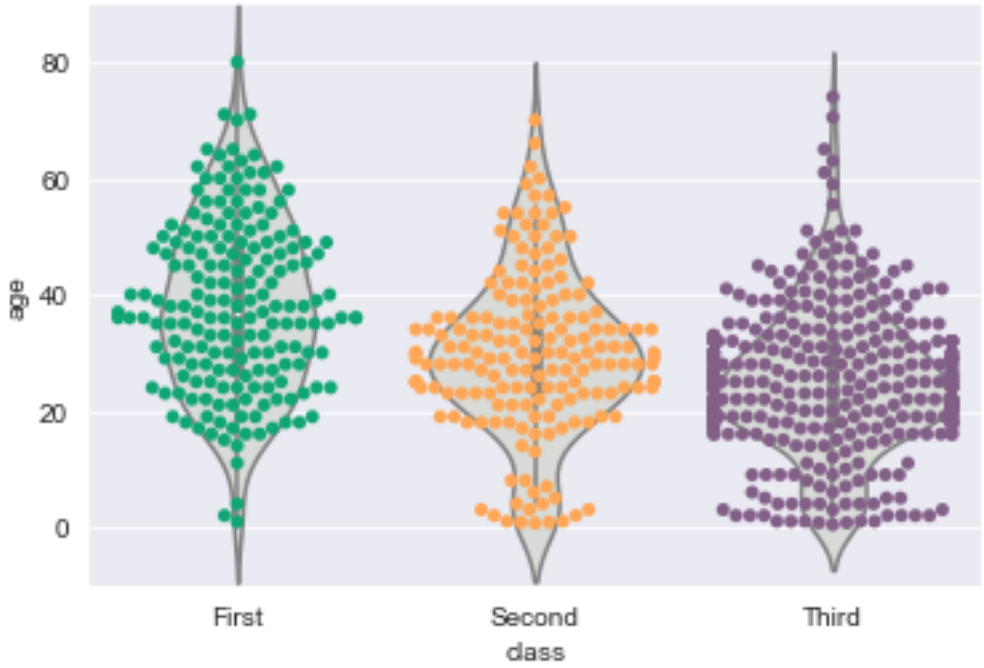
[7]: (0, 80)



1.2.4 Violin plot on top of Swarm plot

[8]:
```
mypal2 = {'First':sns.xkcd_rgb['bluish green'], 'Second':sns.xkcd_rgb['pale␣
↪orange'], 'Third':sns.xkcd_rgb['dusty purple']}
sns.violinplot(x='class', y='age', data=titanic, color=sns.xkcd_rgb['light␣
↪grey'])
sns.swarmplot(x='class', y='age', data=titanic, palette=mypal2)
plt.ylim(-10,90)
```

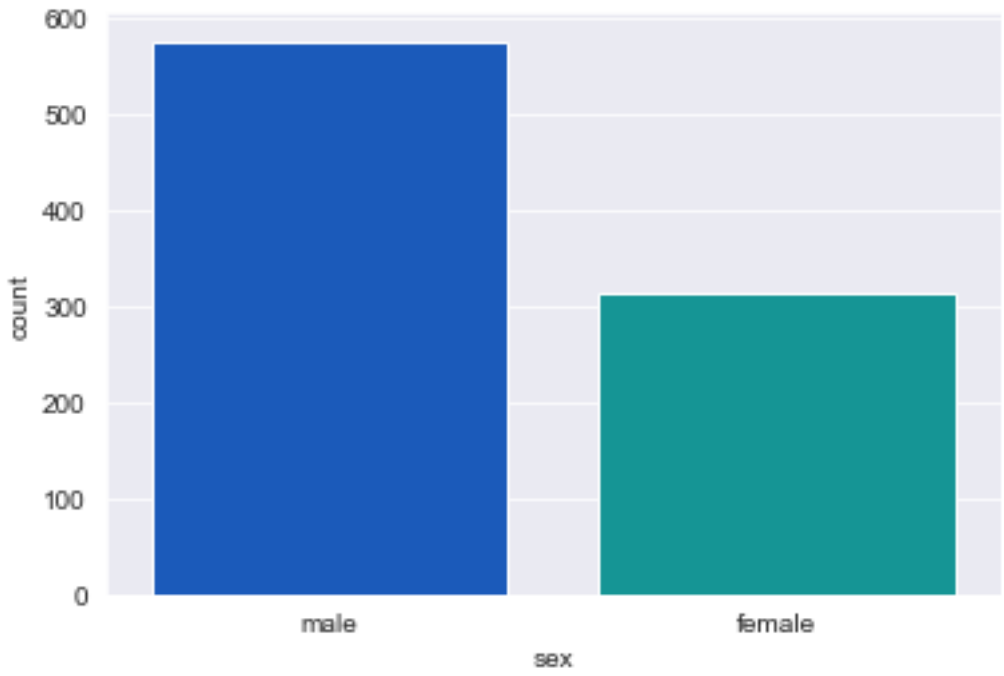
[8]: (-10, 90)



1.2.5 Count plot

[9]:
```
sns.countplot(titanic['sex'], palette='winter')
```

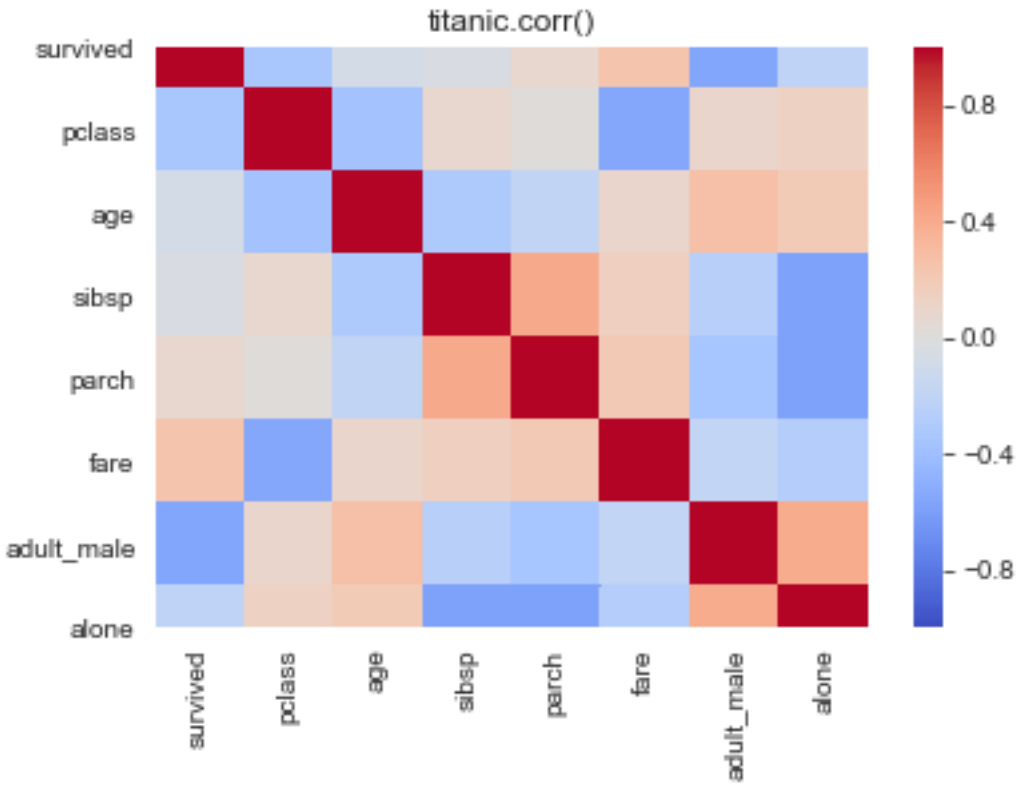
[9]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1a262d3890>
```



1.2.6 Heatmap

[10]:
```
tc = titanic.corr()
sns.heatmap(tc, vmin=-1, vmax=1, cmap='coolwarm')
plt.title('titanic.corr()')
```

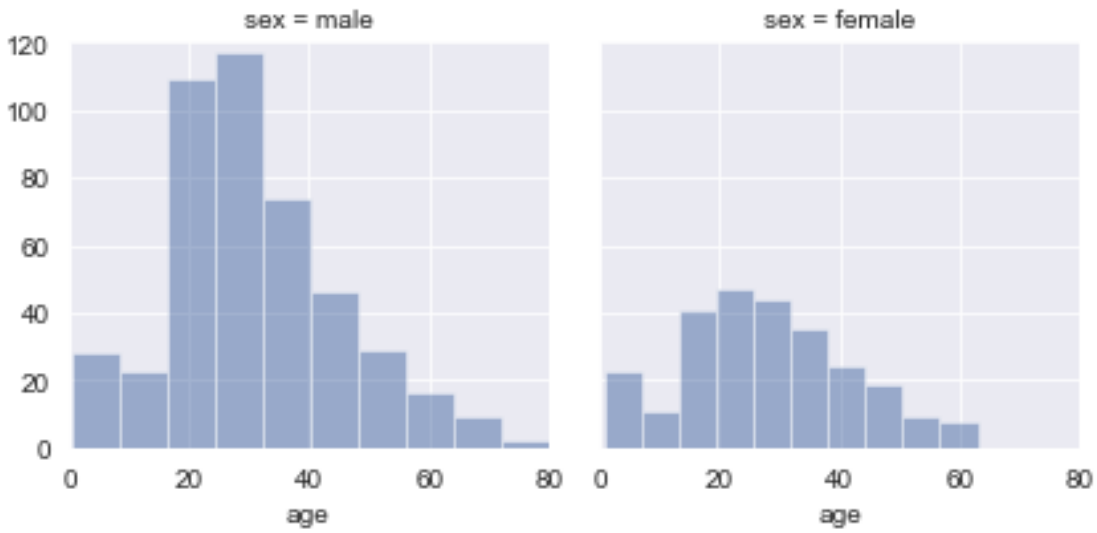
[10]: Text(0.5, 1, 'titanic.corr()')



1.2.7 Facet grid

[11]:
```
g = sns.FacetGrid(col='sex', data=titanic, xlim=(0,80), ylim=(0,120))
g.map(sns.distplot, 'age', kde=False, bins=10, color=sns.xkcd_rgb['cobalt'])
```

[11]:
```
<seaborn.axisgrid.FacetGrid at 0x1a26542890>
```


sex = male
sex = female
120
100
80
60
40
20
0
0
20
40
60
80
age
age