

seabornって便利なやつを使います

- Seabornも実に様々な形式のプロット機能を備えている!
- 今回は主に前処理で使っていく描画形式を解説!
- その他の各種プロットについては <https://seaborn.pydata.org/examples/> を参照

```
In [1]: import seaborn as sns
import pandas as pd

%matplotlib inline
# Jupyter でグラフなどをinlineで表示
import matplotlib.pyplot as plt
```

```
In [2]: tips=pd.read_csv('data/tips.csv')
```

```
In [3]: tips # チップをもらった金額や曜日などのデータセット
```

```
Out[3]:
```

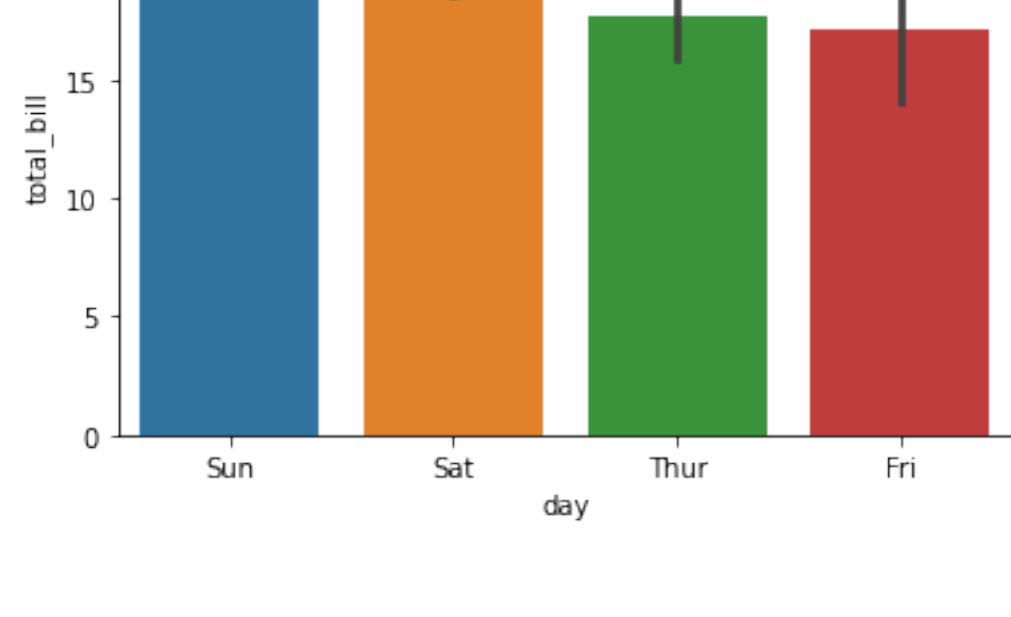
| | total_bill | tip | sex | smoker | day | time | size |
|-----|------------|------|--------|--------|------|--------|------|
| 0 | 16.99 | 1.01 | Female | No | Sun | Dinner | 2 |
| 1 | 10.34 | 1.66 | Male | No | Sun | Dinner | 3 |
| 2 | 21.01 | 3.50 | Male | No | Sun | Dinner | 3 |
| 3 | 23.68 | 3.31 | Male | No | Sun | Dinner | 2 |
| 4 | 24.59 | 3.61 | Female | No | Sun | Dinner | 4 |
| ... | ... | ... | ... | ... | ... | ... | ... |
| 239 | 29.03 | 5.92 | Male | No | Sat | Dinner | 3 |
| 240 | 27.18 | 2.00 | Female | Yes | Sat | Dinner | 2 |
| 241 | 22.67 | 2.00 | Male | Yes | Sat | Dinner | 2 |
| 242 | 17.82 | 1.75 | Male | No | Sat | Dinner | 2 |
| 243 | 18.78 | 3.00 | Female | No | Thur | Dinner | 2 |

244 rows × 7 columns

例題１：曜日と支払い総額の関係

```
In [4]: sns.barplot(x="day", y="total_bill", data=tips)
```

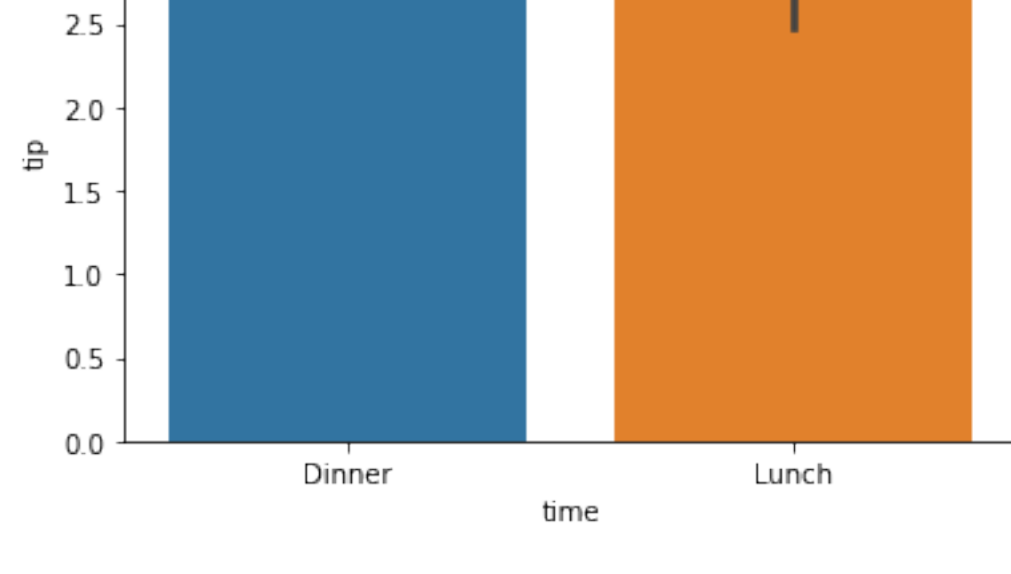
```
Out[4]: <AxesSubplot:xlabel='day', ylabel='total_bill'>
```



課題１：時間帯とチップの関係を図示しなさい

```
In [5]: sns.barplot(x="time", y="tip", data=tips)
```

```
Out[5]: <AxesSubplot:xlabel='time', ylabel='tip'>
```



```
In [ ]:
```

```
In [6]: titanic = sns.load_dataset("titanic") # データセットを読み込み
```

```
In [7]: titanic
```

```
Out[7]:
```

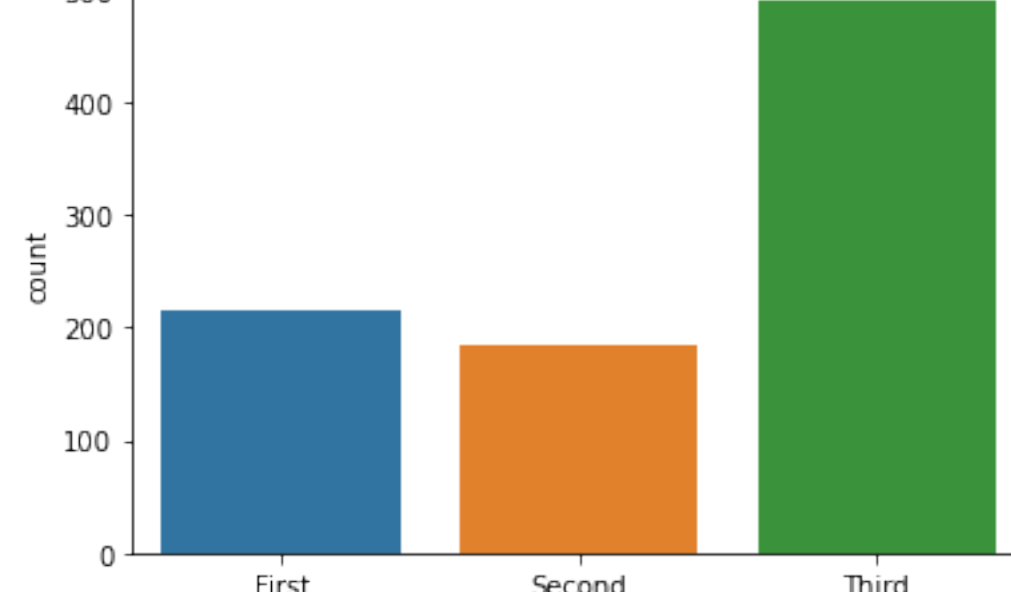
| | survived | pclass | sex | age | sibsp | parch | fare | embarked | class | who | adult_male | deck | embark_town | alive | alone |
|-----|----------|--------|--------|------|-------|-------|---------|----------|--------|-------|------------|------|-------------|-------|-------|
| 0 | 0 | 3 | male | 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 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 886 | 0 | 2 | male | 27.0 | 0 | 0 | 13.0000 | S | Second | man | True | NaN | Southampton | no | True |
| 887 | 1 | 1 | female | 19.0 | 0 | 0 | 30.0000 | S | First | woman | False | B | Southampton | yes | True |
| 888 | 0 | 3 | female | NaN | 1 | 2 | 23.4500 | S | Third | woman | False | NaN | Southampton | no | False |
| 889 | 1 | 1 | male | 26.0 | 0 | 0 | 30.0000 | C | First | man | True | C | Cherbourg | yes | True |
| 890 | 0 | 3 | male | 32.0 | 0 | 0 | 7.7500 | Q | Third | man | True | NaN | Queenstown | no | True |

891 rows × 15 columns

例題２－１：各船室等級の人数を図示

```
In [8]: sns.countplot(x="class", data=titanic)
```

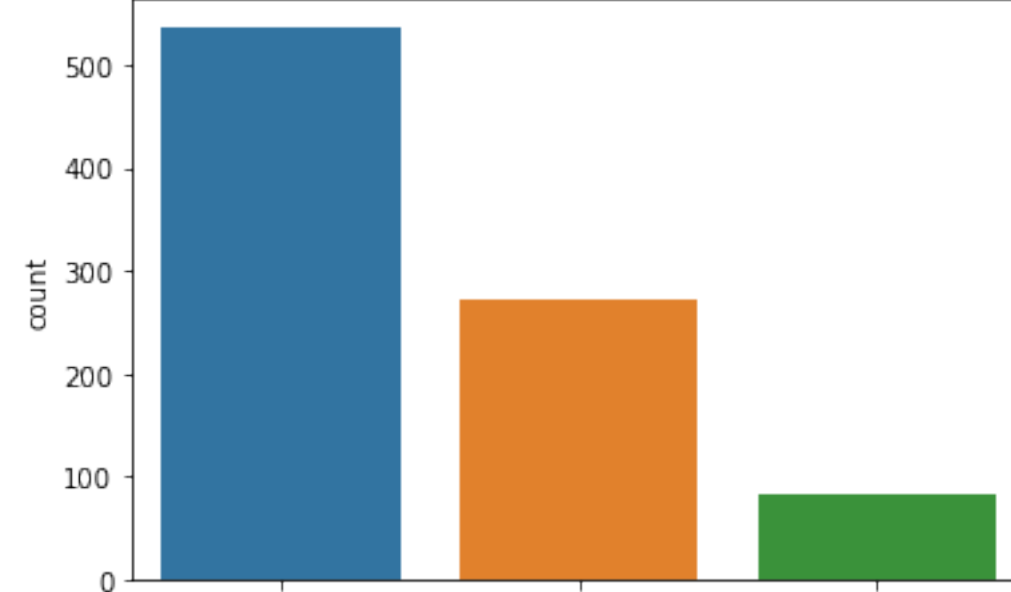
```
Out[8]: <AxesSubplot:xlabel='class', ylabel='count'>
```



課題２－１：男女子供の数を図示

```
In [9]: sns.countplot(x="who", data=titanic)
```

```
Out[9]: <AxesSubplot:xlabel='who', ylabel='count'>
```

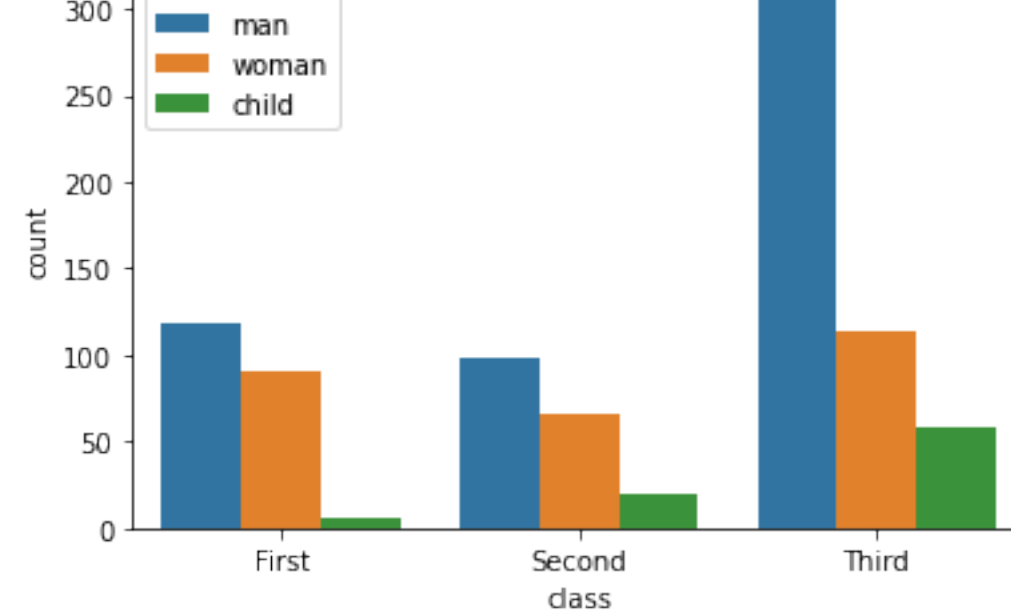


```
In [ ]:
```

例題２－２：２つのカテゴリ（class, whoで分類）

```
In [10]: sns.countplot(x="class", hue="who", data=titanic)
```

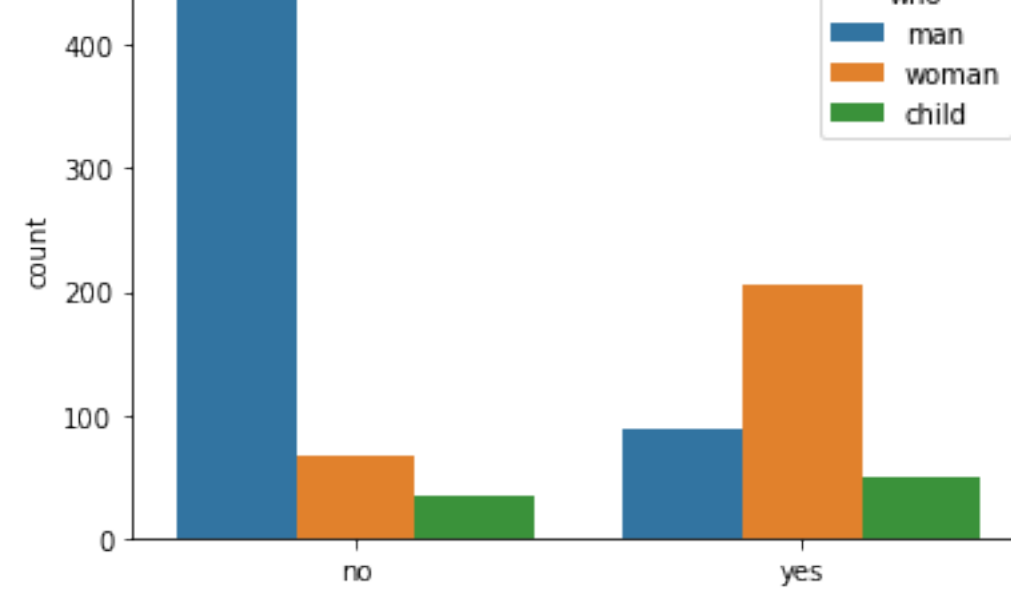
```
Out[10]: <AxesSubplot:xlabel='class', ylabel='count'>
```



課題２－２：任意の２カテゴリで分類

```
In [11]: sns.countplot(x="alive", hue="who", data=titanic)
```

```
Out[11]: <AxesSubplot:xlabel='alive', ylabel='count'>
```

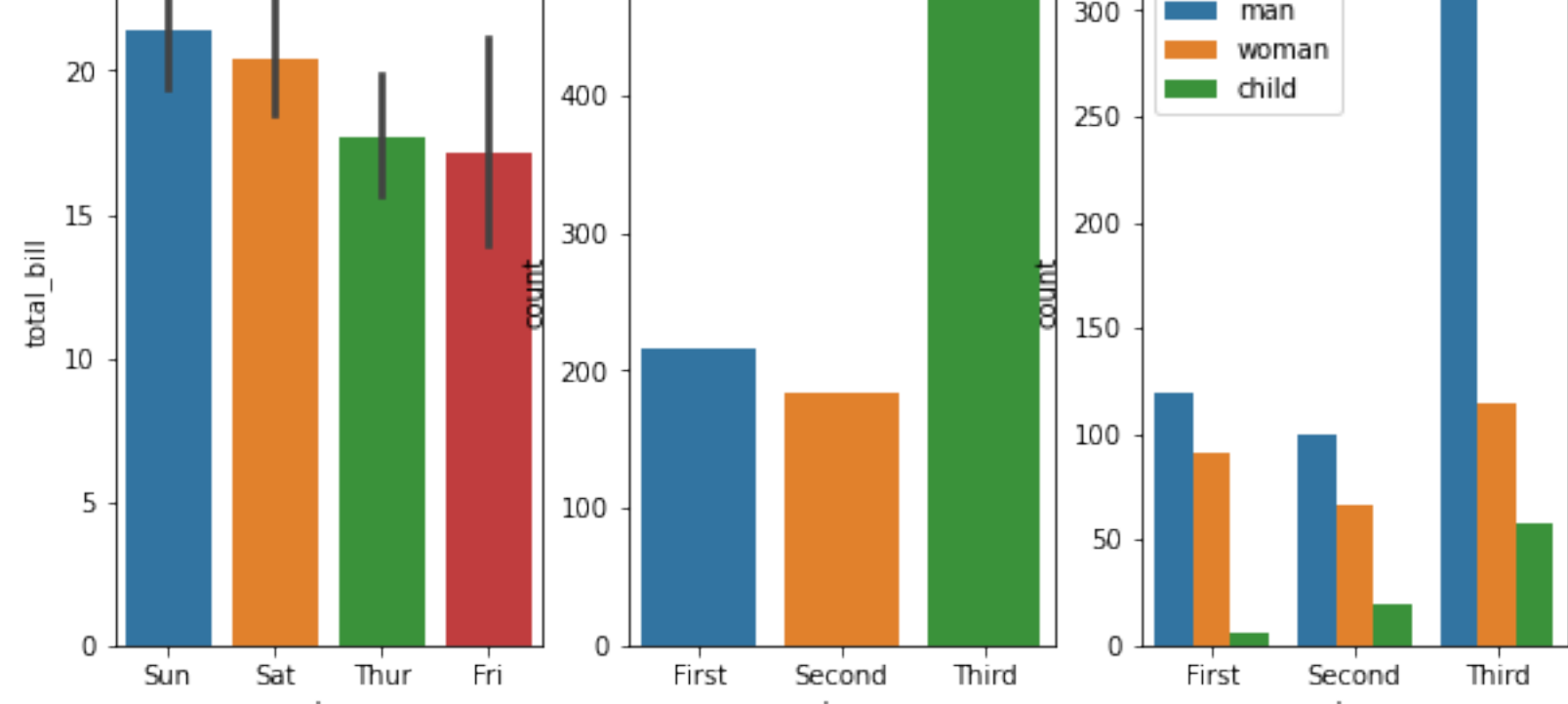


```
In [ ]:
```

例題３：axを用いて複数のsnsを図示

```
In [12]: fig, ax = plt.subplots(1,3, figsize=(10,5))
sns.barplot(x="day", y="total_bill", data=tips, ax =ax[0])
sns.countplot(x="class", data=titanic, ax =ax[1])
sns.countplot(x="class", hue="who", data=titanic, ax =ax[2])
```

```
Out[12]: <AxesSubplot:xlabel='class', ylabel='count'>
```

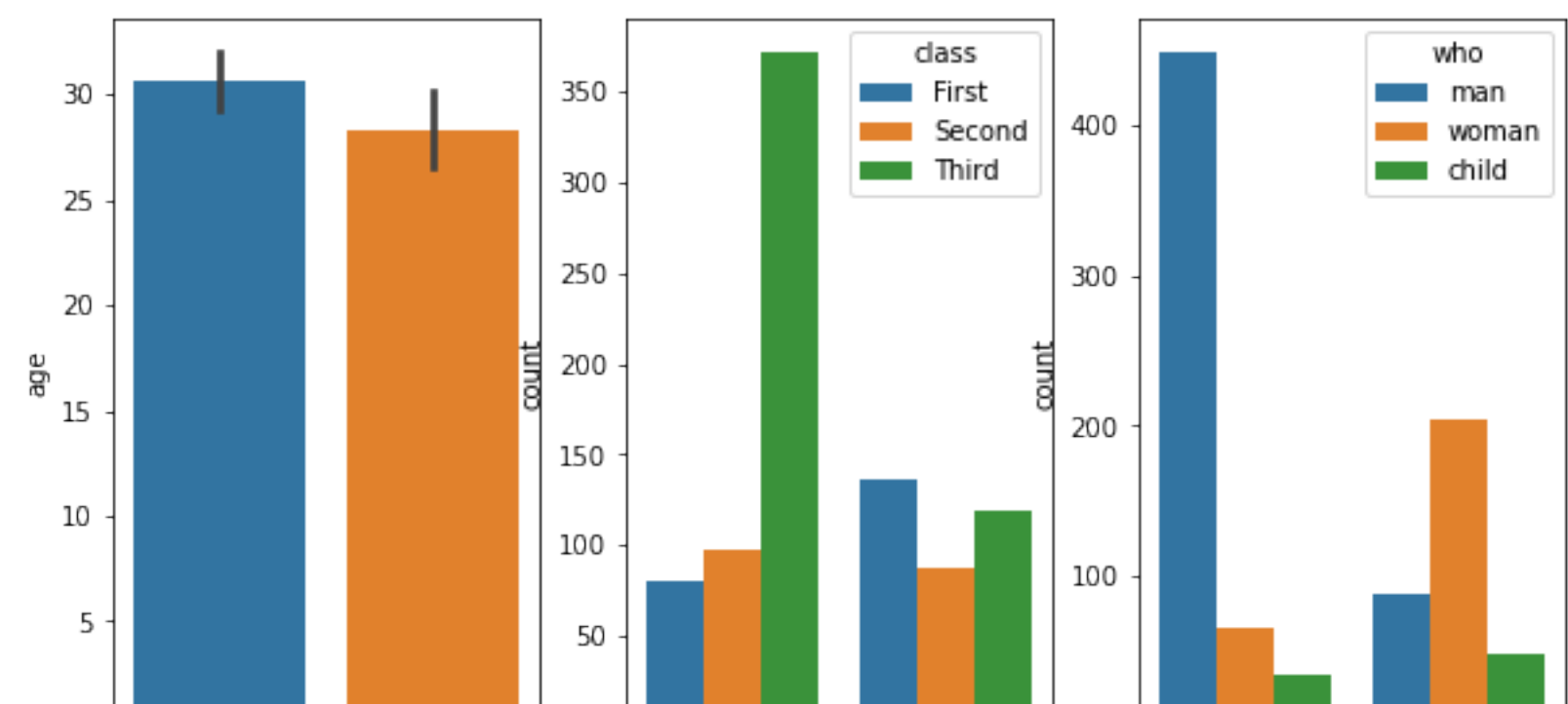


課題３：ax, snsを用いてタイタニックの死因を図示してください。

男性の方が死亡率が高い傾向にある。高齢者の方が生き残る傾向にある。などの関係を図示してください。

```
In [13]: fig, ax = plt.subplots(1,3, figsize=(10,5))
sns.barplot(x="alive", y="age", data=titanic, ax =ax[0])
sns.countplot(x="alive", hue="class", data=titanic, ax =ax[1])
sns.countplot(x="alive", hue="who", data=titanic, ax =ax[2])
```

```
Out[13]: <AxesSubplot:xlabel='alive', ylabel='count'>
```



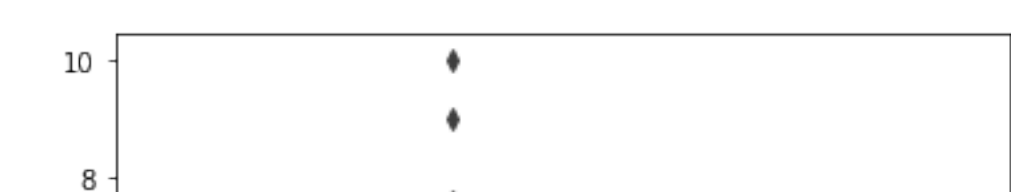
結論：生存者は若い。等級が高いほど生き残っている。女性の方が生き残っている。

```
In [ ]:
```

例題４：曜日とチップの関係の箱ヒゲ図をかけ

```
In [14]: sns.boxplot(x="day", y="tip", data=tips)
```

```
Out[14]: <AxesSubplot:xlabel='day', ylabel='tip'>
```



課題４：時間帯と支払い総額の関係の箱ヒゲ図をかけ

```
In [15]: sns.boxplot(x="time", y="total_bill", data=tips)
```

```
Out[15]: <AxesSubplot:xlabel='time', ylabel='total_bill'>
```



```
In [ ]:
```

その他便利機能

色々一気に見える

```
In [ ]:
```

```
In [16]: sns.pairplot(data=tips)
```

```
Out[16]: <seaborn.axisgrid.PairGrid at 0x7fe184bef4a8>
```



各データの相関関係が見れる

```
In [17]: tips.corr()
```

```
Out[17]:
```

| | total_bill | tip | size |
|------------|------------|----------|----------|
| total_bill | 1.000000 | 0.675734 | 0.598315 |
| tip | 0.675734 | 1.000000 | 0.489299 |
| size | 0.598315 | 0.489299 | 1.000000 |

表にカラーをつける

```
In [18]: tips.corr().style.background_gradient('summer_r')
```

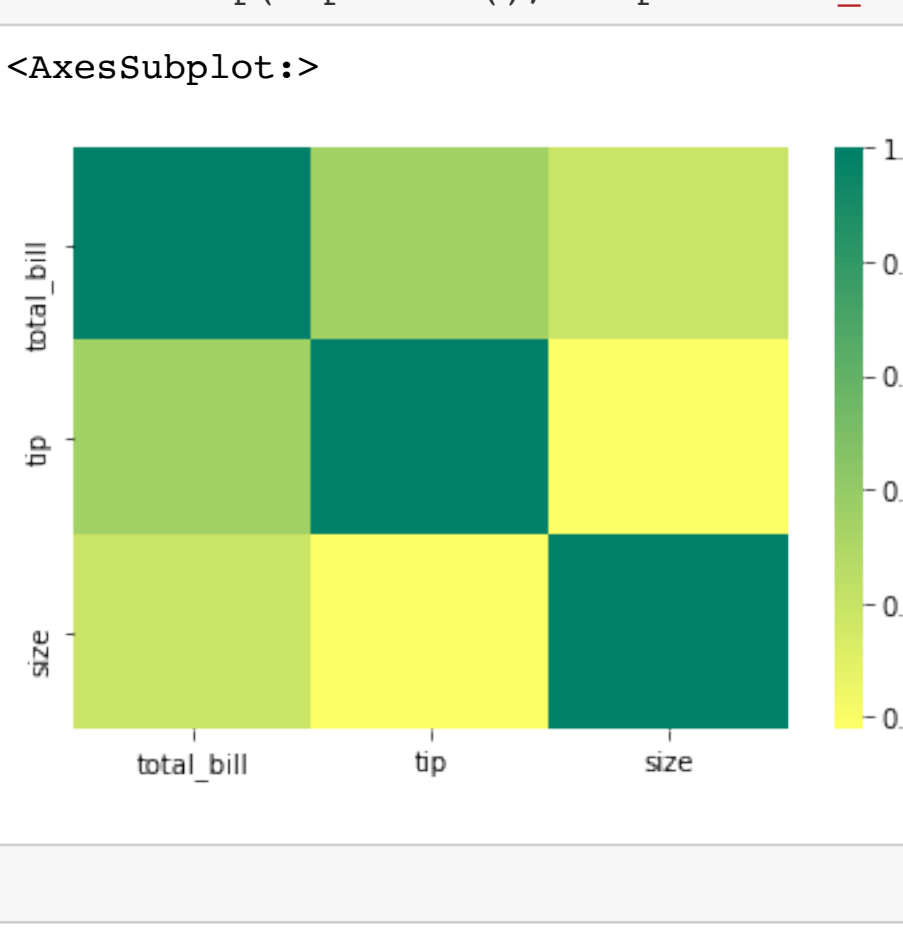
```
Out[18]:
```

| | total_bill | tip | size |
|------------|------------|----------|----------|
| total_bill | 1.000000 | 0.675734 | 0.598315 |
| tip | 0.675734 | 1.000000 | 0.489299 |
| size | 0.598315 | 0.489299 | 1.000000 |

スケールバーをつける

```
In [19]: sns.heatmap(tips.corr(), cmap='summer_r')
```

```
Out[19]: <AxesSubplot>
```



```
In [ ]:
```

課題５：タイタニックのデータセットをもとに自分の好きな用にデータ分析してみる。

なぜ死んだのか？という乗客が多いのか？など結論とその根拠となるデータ、グラフをわかりやすく見せる。 <https://seaborn.pydata.org/examples/>

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```

```
In [ ]:
```