

titanic_project(Seaborn)_exercise

May 23, 2019

1 Seaborn

```
In [ ]: import seaborn as sns
```

```
In [ ]: tips_df = sns.load_dataset('tips')
        tips_df.head()
```

```
In [ ]: tips_df['sex'].value_counts().plot(kind='bar')
```

1.0.1 factorplot

```
In [ ]: sns.factorplot('sex', data=tips_df, kind='count')
```

```
In [ ]: sns.factorplot('smoker', data=tips_df, kind='count')
```

```
In [ ]: sns.factorplot('day', data=tips_df, kind='count')
```

```
In [ ]: sns.factorplot('time', data=tips_df, kind='count')
```

```
In [ ]: sns.factorplot(x='sex', data=tips_df, kind='count', col='day')
```

```
In [ ]: sns.factorplot(x='sex', data=tips_df, kind='count', col='day', col_wrap=2)
```

```
In [ ]: sns.factorplot(x='smoker', data=tips_df, kind='count', col='time')
        sns.factorplot(x='smoker', data=tips_df, kind='count', row='time')
        sns.factorplot(x='smoker', data=tips_df, kind='count', hue='time')
```

```
In [ ]: sns.factorplot('size', data=tips_df, kind='count', hue='sex', order=[1,2,3,4,5,6])
```

```
In [ ]: sns.factorplot(x="sex", y="tip", data=tips_df)
```

```
In [ ]: sns.factorplot(x="time", y="tip", data=tips_df, col='sex')
```

```
In [ ]: sns.factorplot(x="smoker", y="tip", data=tips_df)
```

```
In [ ]: sns.jointplot(tips_df['total_bill'],tips_df['tip'])
```

2

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

In [ ]: pwd

In [ ]: !dir

In [ ]: import pandas as pd
        from pandas import Series, DataFrame

        titanic_df = pd.read_csv("titanic_data.csv")

In [ ]: titanic_df.head()

In [ ]: titanic_df.info()

In [ ]: titanic_df.describe()

In [ ]: titanic_df.head()

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

2.0.1 Who

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

In [ ]: titanic_data = df

In [ ]: sns.factorplot('sex', data=titanic_data, kind = 'count')

In [ ]: sns.factorplot('pclass', data=titanic_data, kind='count', order=[1,2,3])

In [ ]: sns.factorplot('pclass', data=titanic_data, kind='count', hue='sex', \
                        order=[1,2,3])

In [ ]: sns.factorplot('sex', data=titanic_data, kind='count', hue='pclass', \
                        hue_order=[1,2,3])

In [ ]: sns.factorplot('pclass', data=titanic_data, kind='count', hue='who', \
                        order=[3,2,1])

In [ ]: titanic_data['age'].hist(bins=70)
```

2.0.2 From Where

```
In [ ]: titanic_data.head()

In [ ]: cabin_df = DataFrame(titanic_data['deck'].dropna())

In [ ]: sns.factorplot('deck', data=cabin_df, kind='count')

In [ ]: sns.factorplot('deck', data=cabin_df, kind='count', \
                        palette='GnBu', order=list('ABCDEFGG'))

In [ ]: sns.factorplot('embark_town', data=titanic_data, hue='pclass', \
                        hue_order=[1,2,3], kind='count')
```

2.0.3 alone

```
In [ ]: titanic_data.head()

In [ ]: sns.factorplot('alone', data=titanic_data, palette='Blues', kind='count')
```

2.0.4 survived

```
In [ ]: # map column
        titanic_data['alive'] = titanic_data.survived.map({0: "no", 1: "yes"})

        sns.factorplot('alive', data=titanic_data, kind='count')

In [ ]: sns.factorplot('pclass', 'survived', data=titanic_data, order=[1,2,3])

In [ ]: sns.factorplot('pclass', 'survived', hue='who', \
                        data=titanic_data, order=[1,2,3], dodge=True)

In [ ]: sns.factorplot('who', 'survived', hue='pclass', \
                        data=titanic_data, dodge=True, hue_order=[1,2,3])

In [ ]: sns.lmplot('age', 'survived', data=titanic_data)

In [ ]: sns.lmplot('age', 'survived', hue='pclass', \
                    data=titanic_data, hue_order=[1,2,3])

In [ ]: # grouping..
        generations=[10,20,40,60,80]
        sns.lmplot('age', 'survived', hue='pclass', \
                    data=titanic_data, palette='winter', \
                    hue_order=[1,2,3])

In [ ]: sns.lmplot('age', 'survived', hue='sex', \
                    data=titanic_data, palette='winter')

In [ ]:
```