

Seaborn 기본

디자인 12기 송민수

- Seaborn 의 모든 내용을 다루는 것이 아니고 기본적으로 그래프 그리는 방법만 다룹니다
- 이 자료에서는 그래프의 생김새만 파악하고 실습자료에서 구체적으로 확인해보겠습니다
- 추가적으로 공부하고 싶으신 분들은 구글링하시면 양질의 자료가 많습니다

목차

1. 다양한 시각화 라이브러리 소개
2. Seaborn plot 소개

1. 다양한 시각화 라이브러리 소개

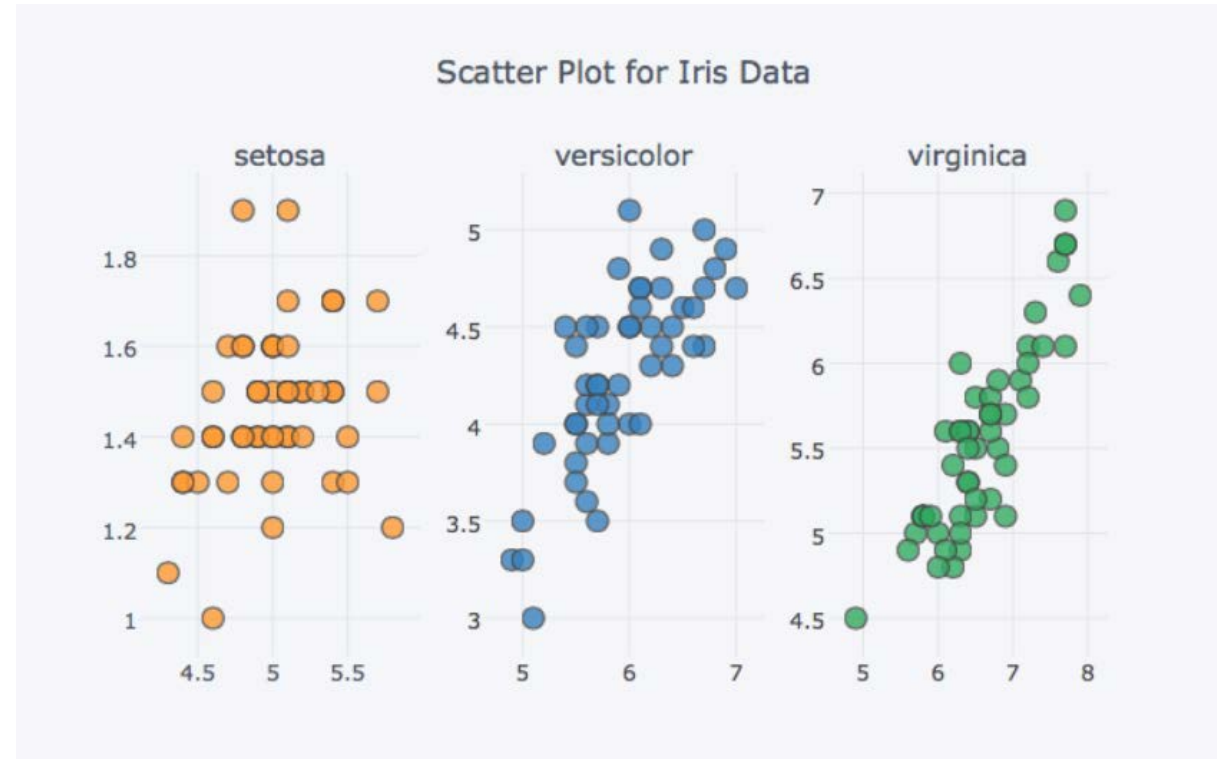
Cufflinks

특징

- 인터랙티브한 시각화
- 사용하기 매우 간단(12기 방학세션 때 이거로 배움)
- 메모리 소모가 큼

참고사이트:

- <https://plot.ly/ipython-notebooks/cufflinks/>
- http://lumiamitie.github.io/python/cufflinks_basic/



1. 다양한 시각화 라이브러리 소개

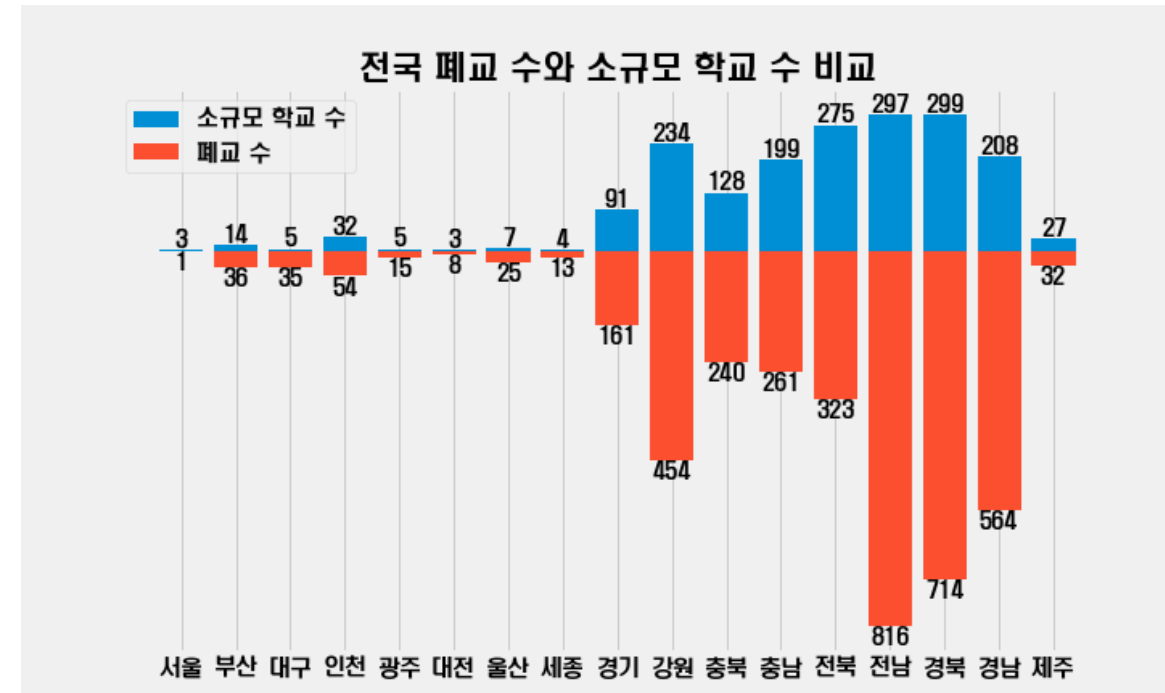
Matplotlib

특징

- 인터랙티브한 시각화도 가능함(하지만 웹이 아니면 굳이 할 이유는 없음)
- 자유도가 높음 = customizing 하여서 다양한 그래프 그리기 가능
- 퀄리티가 좋음
- 공부하기 어려움;;

참고사이트:

<https://matplotlib.org/>



1. 다양한 시각화 라이브러리 소개

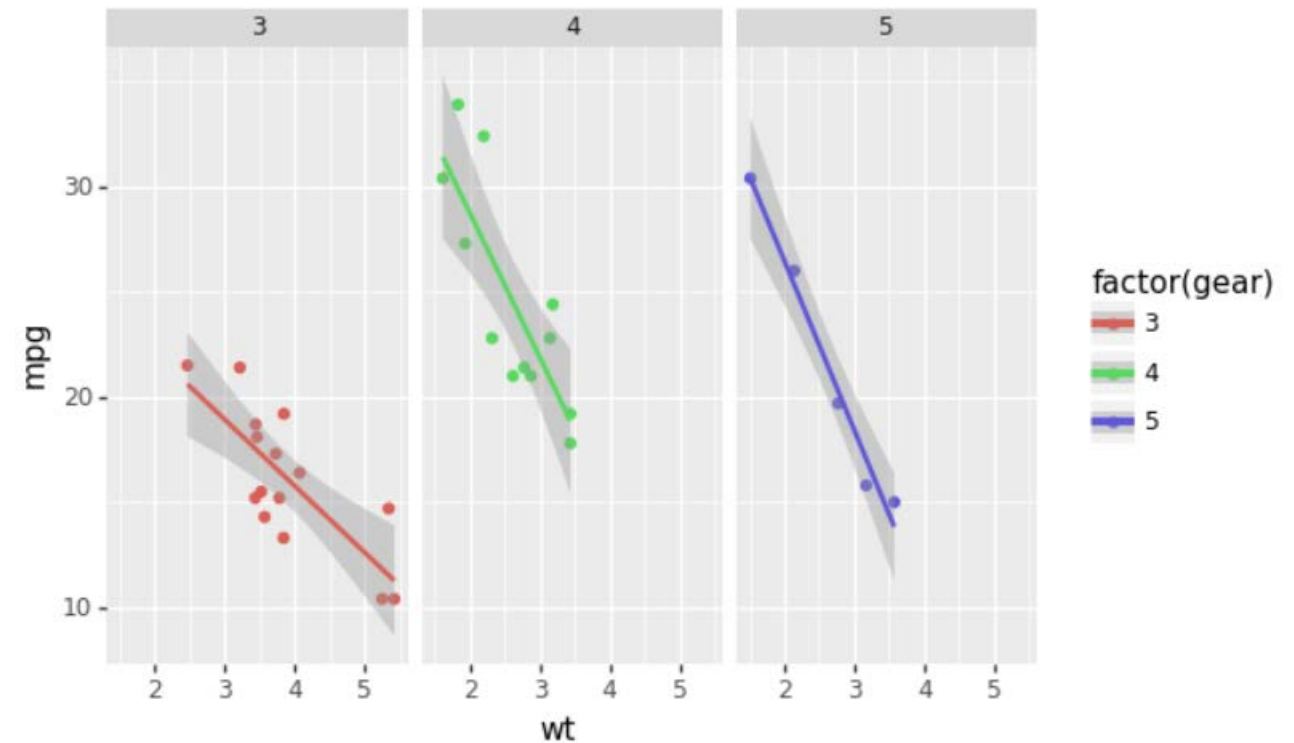
Plotnine

특징

- R + python
- 김철웅 교수님의 '탐색적 자료 분석' 을 수강한 응용통계학생들은 이게 편할 수도 있음
- 하지만 python 유저들이 굳이 사용할 필요는 없어보임 물론 R 유저들은 그냥 R 쓰겠지

참고사이트:

<https://plotnine.readthedocs.io/en/stable/>



1. 다양한 시각화 라이브러리 소개

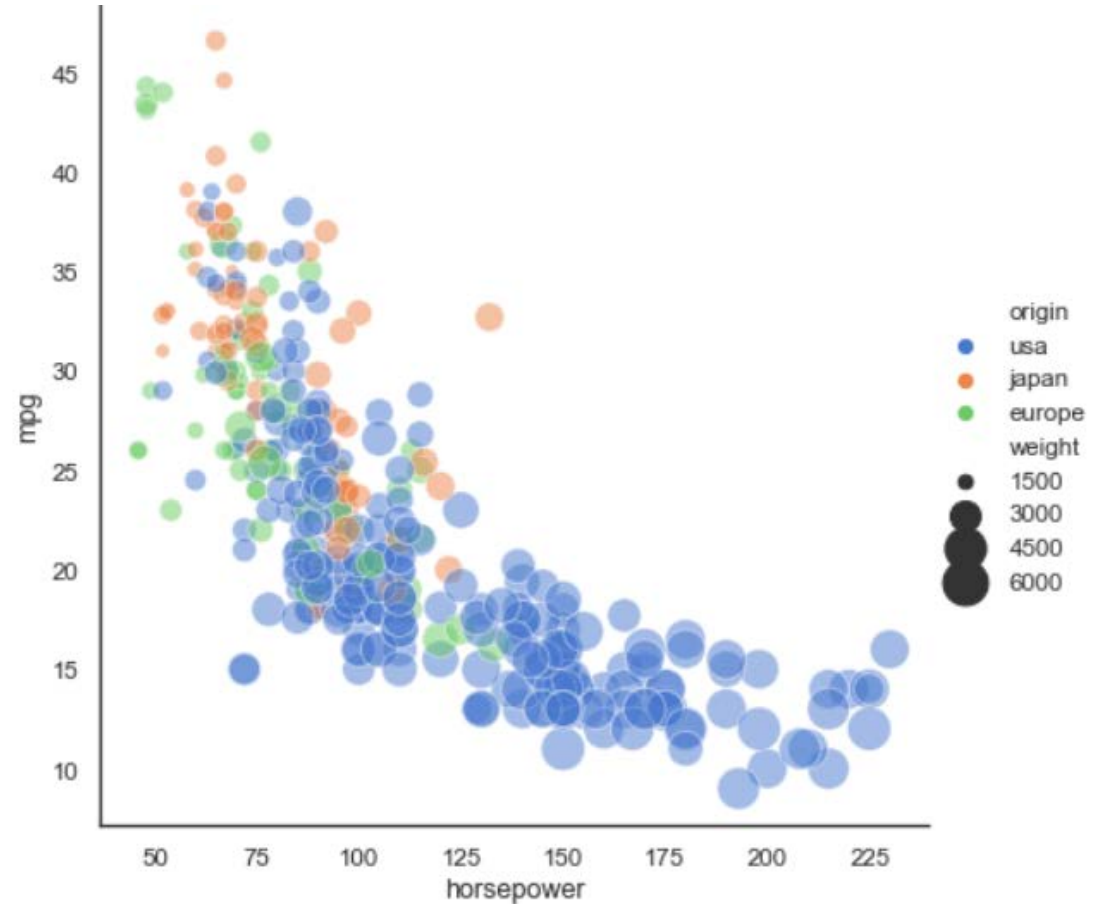
Seaborn

특징

- 예쁨
- Matplotlib 보다는 쉬움
- 하지만 다소 기본적인 그래프
- EDA 하기에는 매우 용이하다고 판단

참고사이트:

<https://seaborn.pydata.org/>

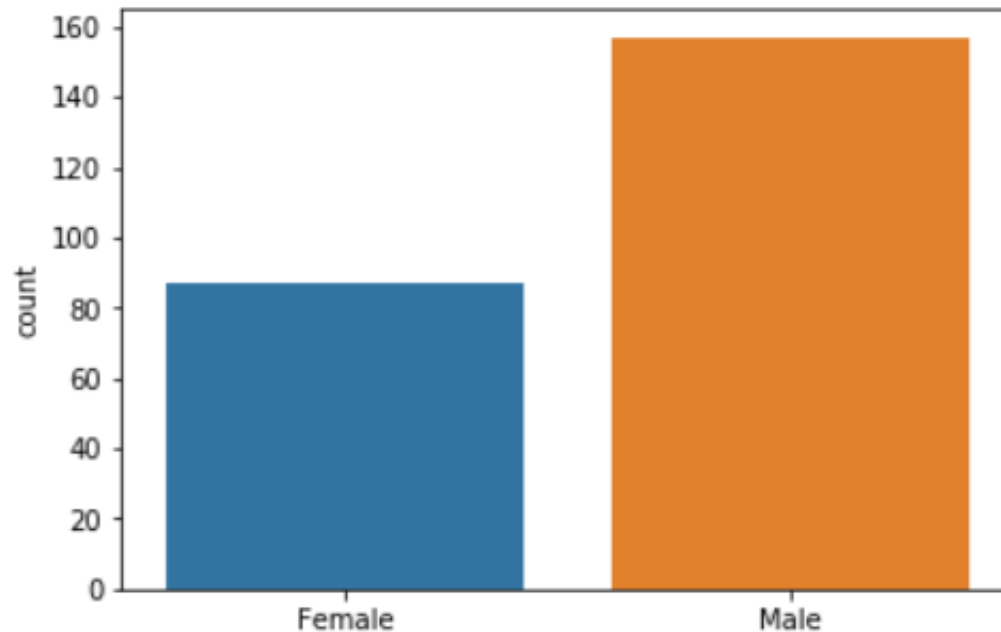


sns.countplot

*(x=None, y=None, hue=None, data=None, order=None, hue_order=None, orient=None, color=None, palette=None, saturation=0.75, dodge=True, ax=None, **kwargs)*

```
sns.countplot(x = 'sex', data = tips)
```

<matplotlib.axes._subplots.AxesSubplot at 0x297fcfbecf8>

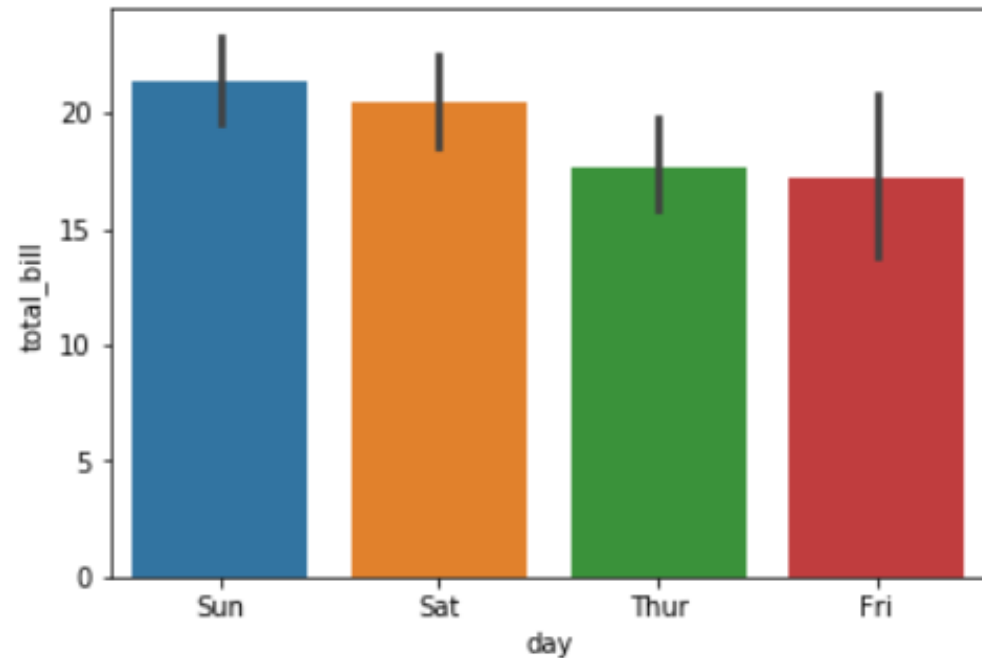


sns.barplot

*(x=None, y=None, hue=None, data=None, order=None, hue_order=None, estimator=<function mean>, ci=95, n_boot=1000, units=None, orient=None, color=None, palette=None, saturation=0.75, errcolor='.26', errwidth=None, capsize=None, dodge=True, ax=None, **kwargs)*

```
sns.barplot(x = 'day', y = 'total_bill', data = tips)
```

<matplotlib.axes._subplots.AxesSubplot at 0x297fc37d668>

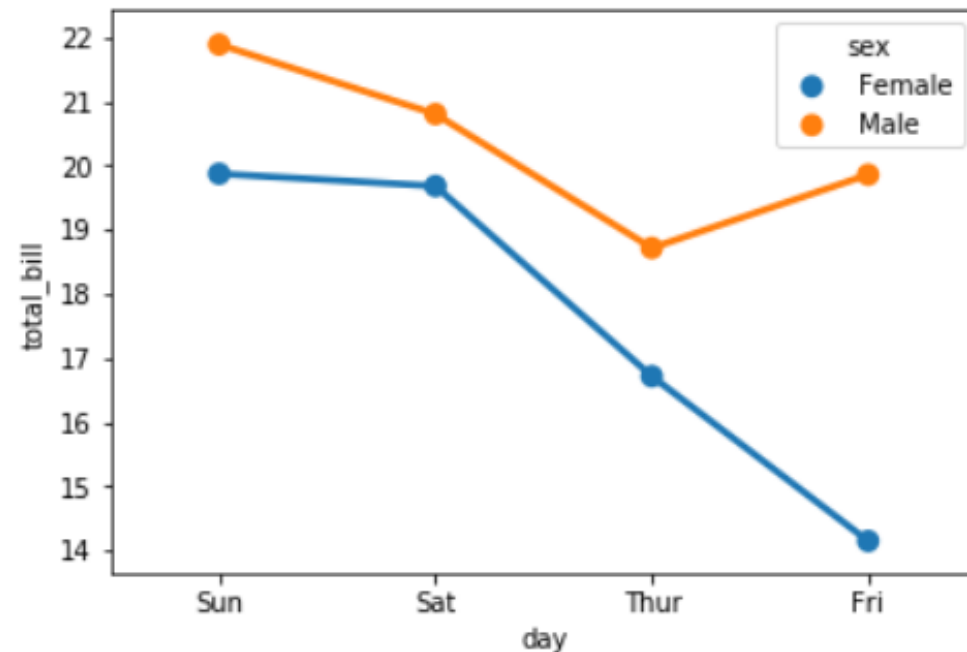


sns.pointplot

*(x=None, y=None, hue=None, data=None, order=None, hue_order=None, estimator=<function mean>, ci=95, n_boot=1000, units=None, markers='o', linestyle='-', dodge=False, join=True, scale=1, orient=None, color=None, palette=None, errwidth=None, capsize=None, ax=None, **kwargs)*

```
sns.pointplot(x = 'day', y = 'total_bill',  
              hue = 'sex', data = tips, ci = None)  
# ci : size of confidence interval
```

<matplotlib.axes._subplots.AxesSubplot at 0x2212a51bfd0>

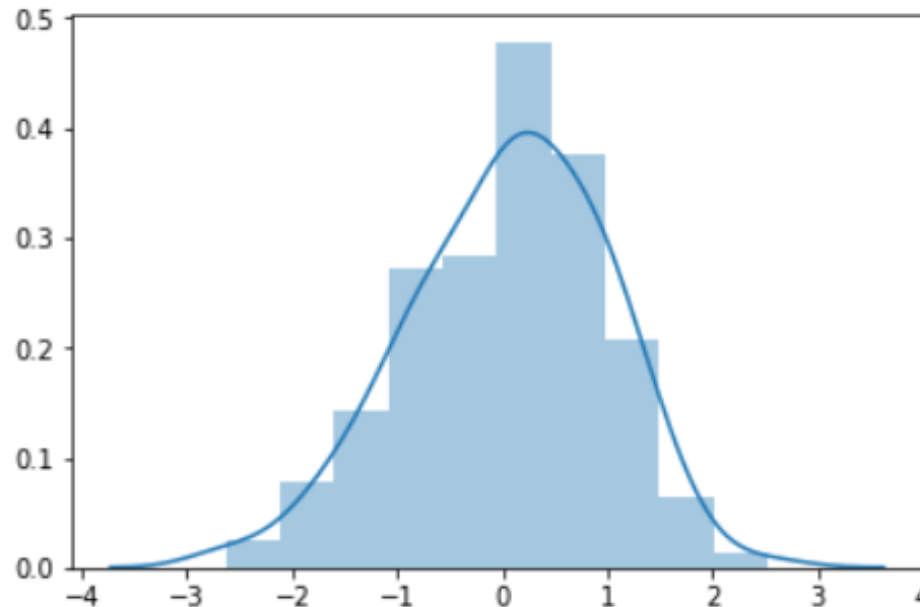


sns.distplot

(a, bins=None, hist=True, kde=True, rug=False, fit=None, hist_kws=None, kde_kws=None, rug_kws=None, fit_kws=None, color=None, vertical=False, norm_hist=False, axlabel=None, label=None, ax=None)

```
num = np.random.randn(150)
sns.distplot(num)
# default 모양
```

<matplotlib.axes._subplots.AxesSubplot at 0x297fcb2f4e0>

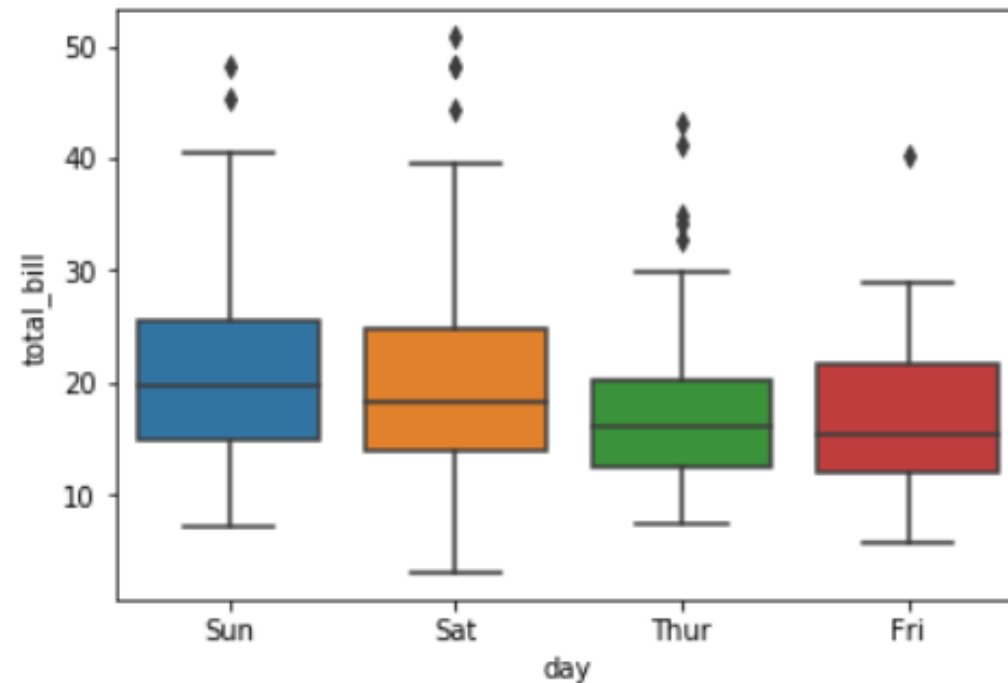


sns.boxplot

*(x=None, y=None, hue=None, data=None, order=None, hue_order=None, orient=None, color=None, palette=None, saturation=0.75, width=0.8, dodge=True, fliersize=5, linewidth=None, whis=1.5, notch=False, ax=None, **kwargs)*

```
sns.boxplot(x="day", y="total_bill", data=tips)
```

<matplotlib.axes._subplots.AxesSubplot at 0x297fc16b4e0>

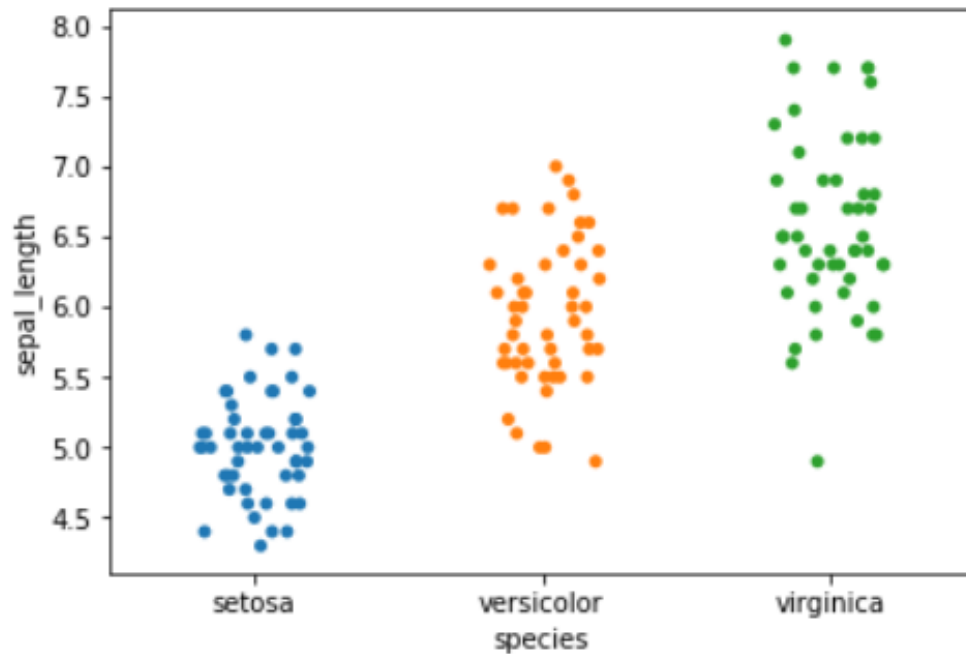


sns.stripplot

*(x=None, y=None, hue=None, data=None, order=None, hue_order=None, jitter=True, dodge=False, orient=None, color=None, palette=None, size=5, edgecolor='gray', linewidth=0, ax=None, **kwargs)*

```
sns.stripplot(x = 'species', y = 'sepal_length', data = iris, jitter = 0.2 )  
#jitter 를 통해 복잡한 분포일 때 보기 편하게 퍼트려줌
```

<matplotlib.axes._subplots.AxesSubplot at 0x297fc804748>

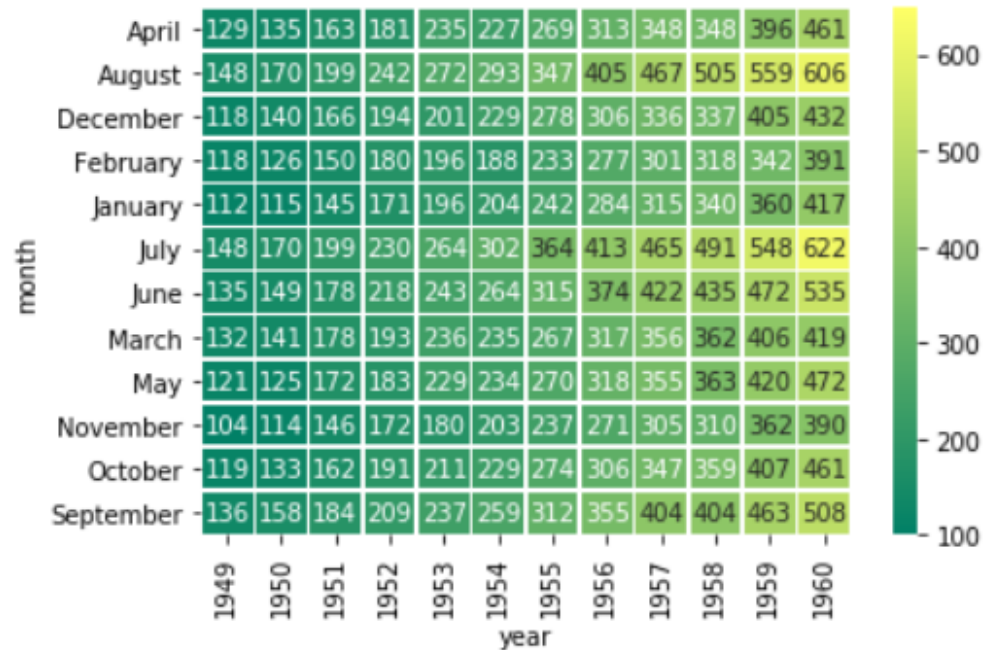


sns.heatmap

*(data, vmin=None, vmax=None, cmap=None, center=None, robust=False, annot=None, fmt='.2g', annot_kws=None, linewidths=0, linecolor='white', cbar=True, cbar_kws=None, cbar_ax=None, square=False, xticklabels='auto', yticklabels='auto', mask=None, ax=None, **kwargs)*

```
sns.heatmap(flights, annot = True,  
            fmt = 'd', # fmt : 숫자의 형태(여기서는 정수)  
            linewidths = 0.5, vmin = 100, vmax = 650,  
            cmap = 'summer' ) #cmap으로 컬러맵조정
```

<matplotlib.axes._subplots.AxesSubplot at 0x297fd05d358>

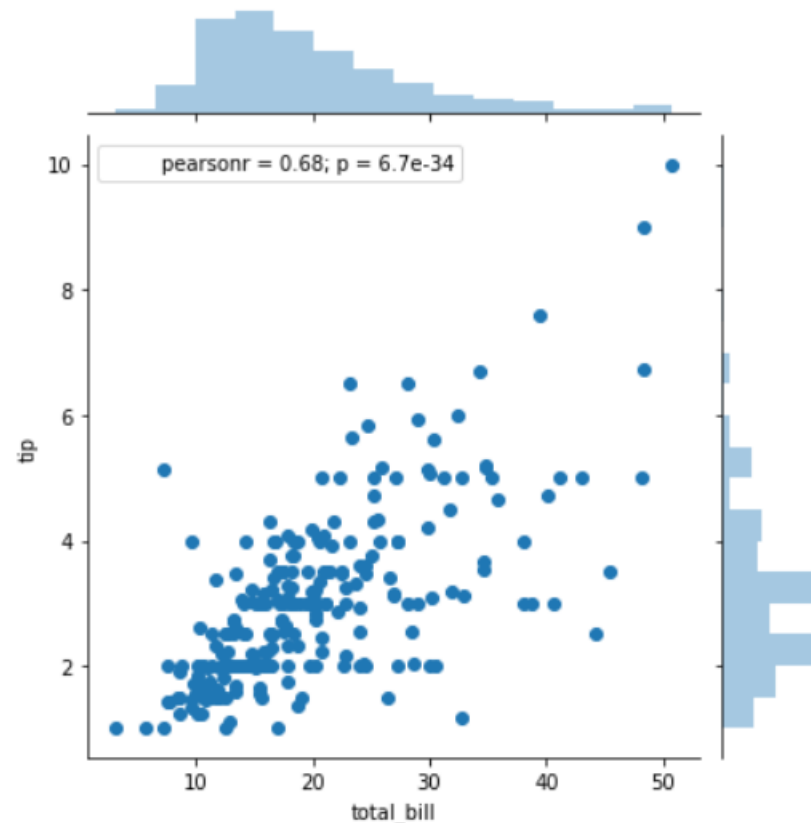


sns.jointplot

*(x, y, data=None, kind='scatter', stat_func=None, color=None, height=6, ratio=5, space=0.2, dropna=True, xlim=None, ylim=None, joint_kws=None, marginal_kws=None, annot_kws=None, **kwargs)*

```
sns.jointplot(x = 'total_bill', y = 'tip', data = tips)
```

```
<seaborn.axisgrid.JointGrid at 0x297fcf127f0>
```

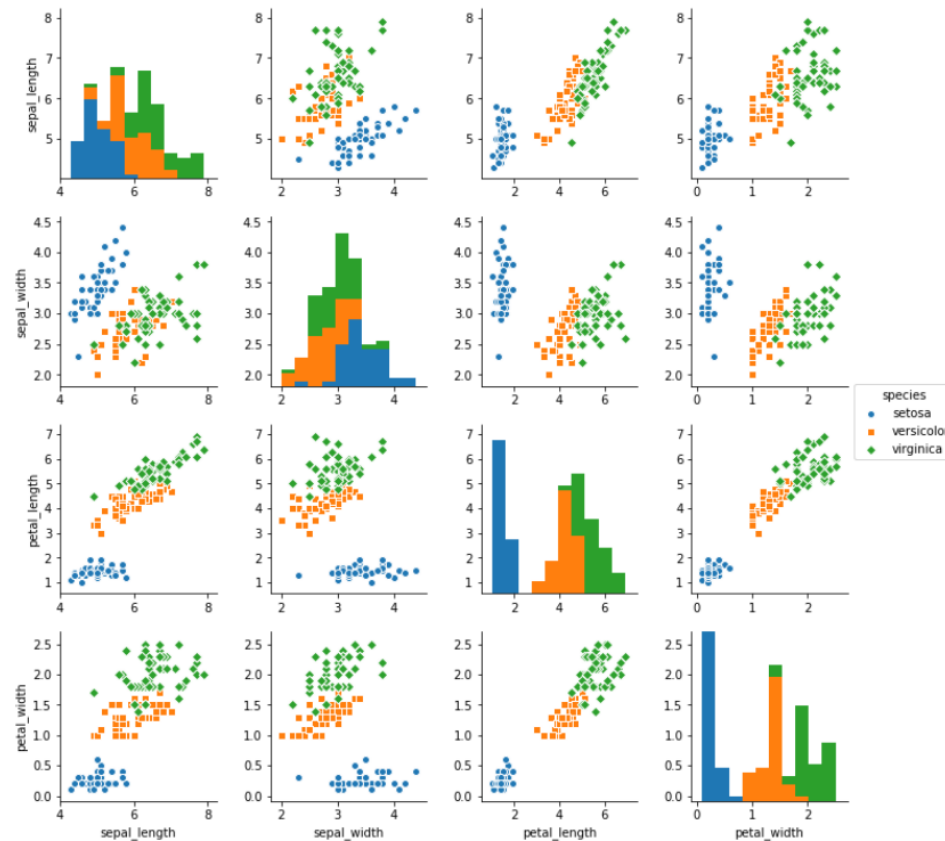


sns.pairplot

(data, hue=None, hue_order=None, palette=None, vars=None, x_vars=None, y_vars=None, kind='scatter', diag_kind='auto', markers=None, height=2.5, aspect=1, dropna=True, plot_kws=None, diag_kws=None, grid_kws=None, size=None)

```
sns.pairplot(iris, hue="species", markers=["o", "s", "D"])
```

<seaborn.axisgrid.PairGrid at 0x297fc92b5c0>



sns.regplot

(x, y, data=None, x_estimator=None, x_bins=None, x_ci='ci', scatter=True, fit_reg=True, ci=95, n_boot=1000, units=None, order=1, logistic=False, lowess=False, robust=False, logx=False, x_partial=None, y_partial=None, truncate=False, dropna=True, x_jitter=None, y_jitter=None, label=None, color=None, marker='o', scatter_kws=None, line_kws=None, ax=None)

```
sns.regplot(x = 'total_bill', y = 'tip', data = tips)  
#데이터를 가장 잘 설명하는 reg 선을 그려줌
```

<matplotlib.axes._subplots.AxesSubplot at 0x29781c30518>

