




Визуализация данных. Matplotlib, Seaborn, Plotly

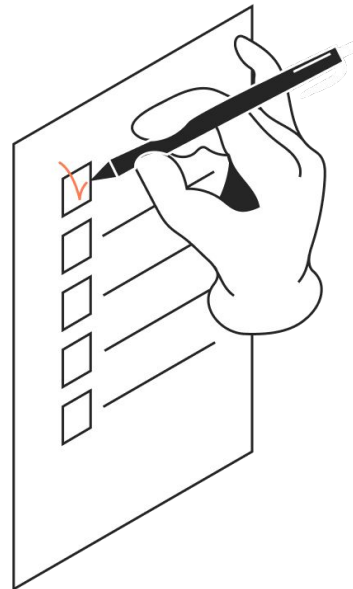
Визуализация данных — это сила!





Что будет на уроке сегодня

-  Зачем визуализировать данные?
-  Как одни и те же графики строить в разных библиотеках. Matplotlib/Seaborn
-  Plotly и почему о ней всё время забывают





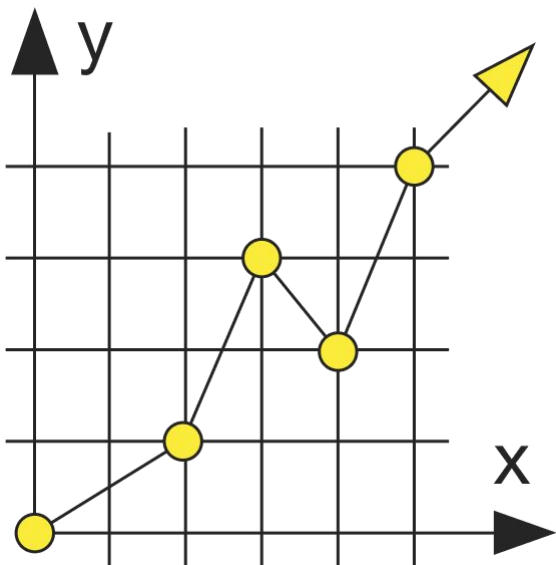
Вопрос

Зачем визуализировать данные?





Зачем визуализировать данные?



- EDA, для лучшего понимания данных
- Наглядно представить результаты



Matplotlib VS Seaborn





Что будем визуализировать?

```
1 data = sns.load_dataset("tips")
2 data.head()
```

	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

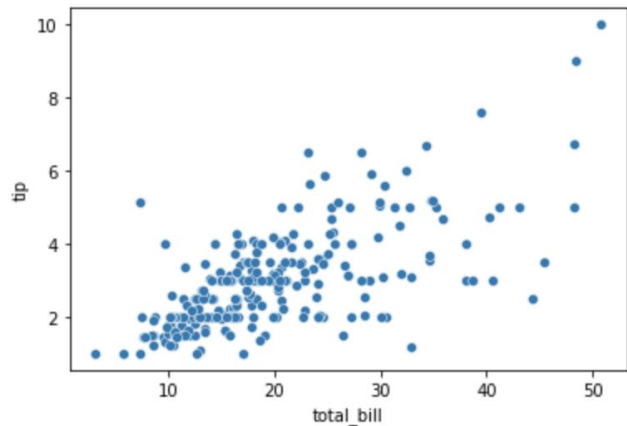
- **total_bill** — счёт в ресторане
- **tip** — сумма оставленных чаевых
- **sex** — пол клиента
- **smoker** — курил ли клиент
- **day** — день недели
- **time** — тип приёма пищи/время приёма пищи



Диаграмма рассеяния — Scatter plot

```
1 sns.scatterplot(data=data, x="total_bill", y="tip")
```

<AxesSubplot:xlabel='total_bill', ylabel='tip'>



```
1 plt.scatter(data.total_bill, data.tip);
```

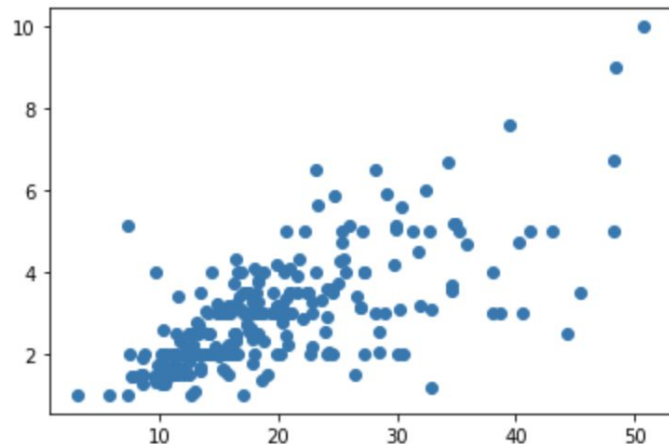
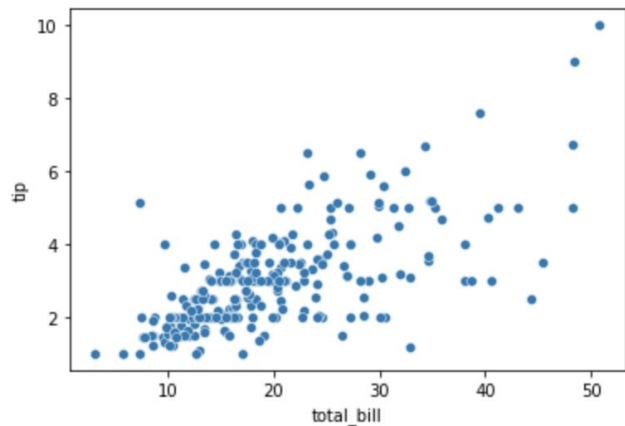




Диаграмма рассеяния — Scatter plot

```
1 sns.scatterplot(data=data, x="total_bill", y="tip")
```

<AxesSubplot:xlabel='total_bill', ylabel='tip'>



```
1 plt.scatter(data.total_bill, data.tip)
2 plt.xlabel('total_bill')
3 plt.ylabel('tip');
```

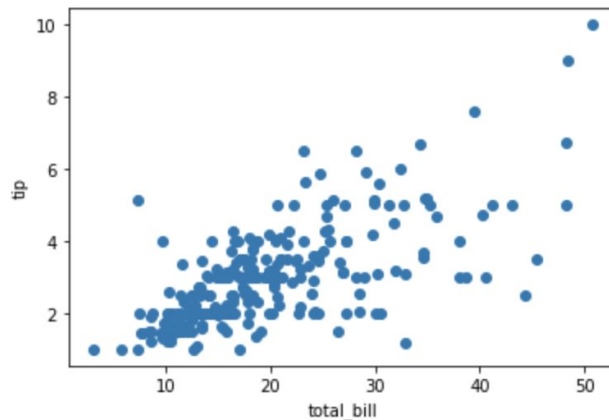
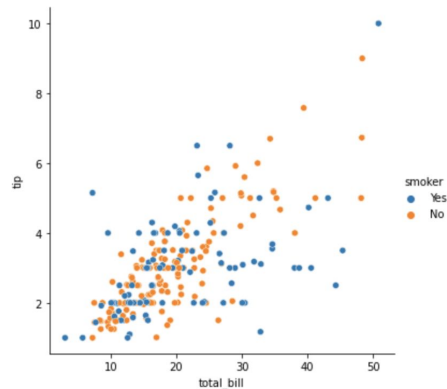


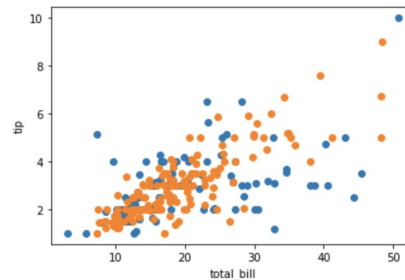


Диаграмма рассеяния — Scatter plot

```
1 sns.relplot(x = "total_bill", y = "tip", hue= "smoker", data=data);
```



```
1 plt.scatter(data[data.smoker == 'Yes'].total_bill, data[data.smoker == 'Yes'].tip)  
2 plt.scatter(data[data.smoker == 'No'].total_bill, data[data.smoker == 'No'].tip)  
3 plt.xlabel('total_bill')  
4 plt.ylabel('tip');
```

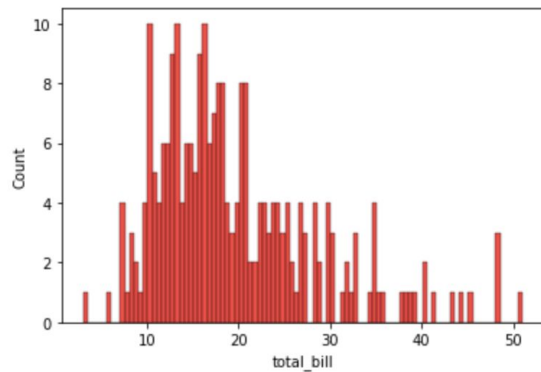




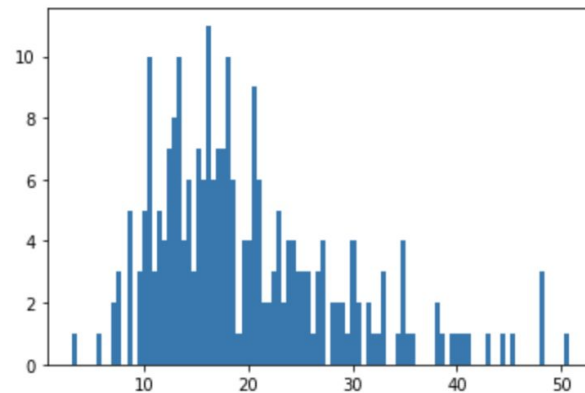
Гистограмма — Hist plot

```
: 1 sns.histplot(data["total_bill"], binwidth=0.5, color='red')
```

```
: <AxesSubplot:xlabel='total_bill', ylabel='Count'>
```



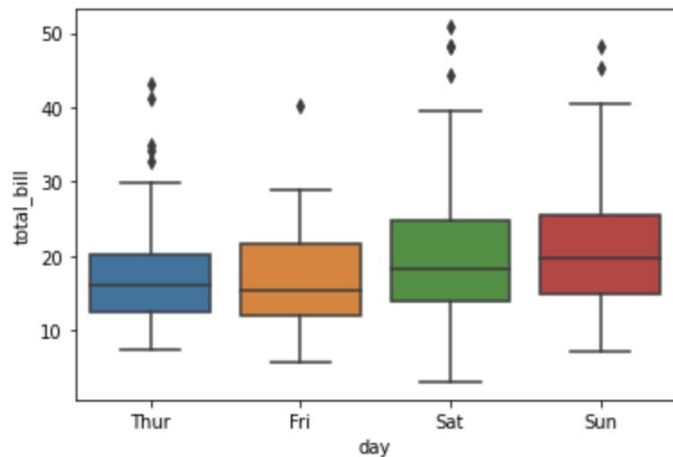
```
1 plt.hist(data.total_bill, bins=100);
```



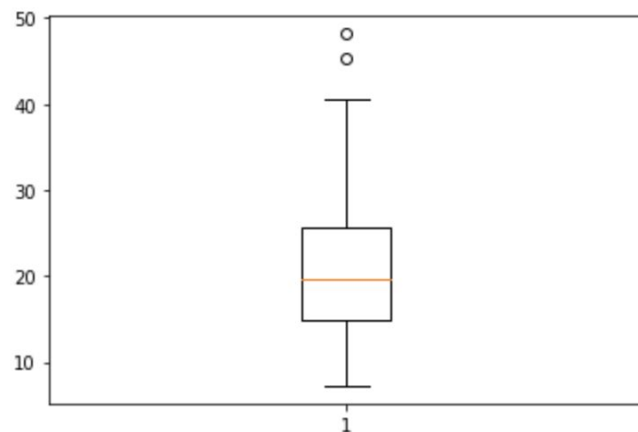


Ящик с усами — Hist plot

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



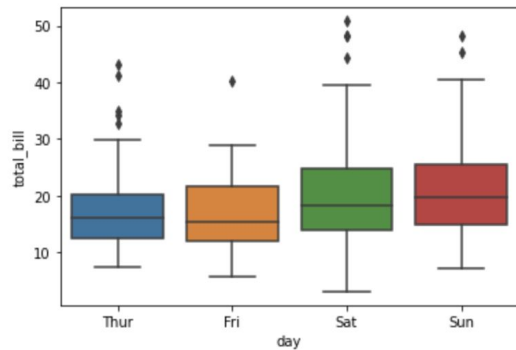
```
1 plt.boxplot(data[data.day=='Sun'].total_bill);
```



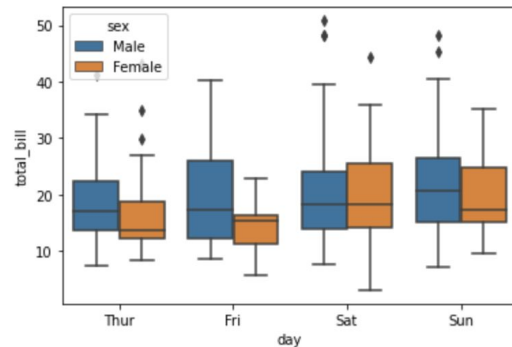


Ящик с усами — Hist plot

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



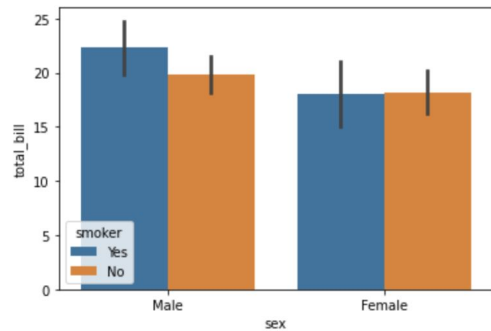
```
1 sns.boxplot(x="day", y = "total_bill", hue = "sex", data = data);
```



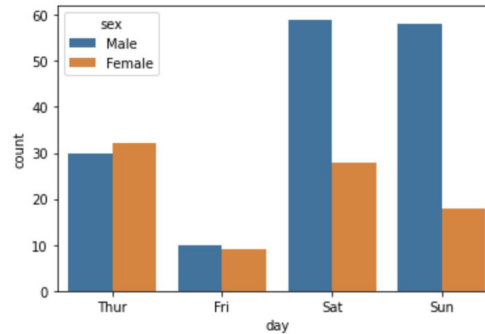


Гистограмма — Bar Plot

```
1 sns.barplot(x = "sex", y = "total_bill", hue = "smoker", data = data);
```



```
1 sns.countplot(x = "day", hue = "sex", data = data);
```





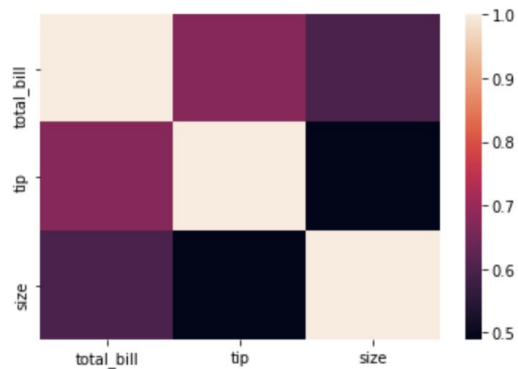
Тепловая карта — Heatmap

```
1 data.corr()
```

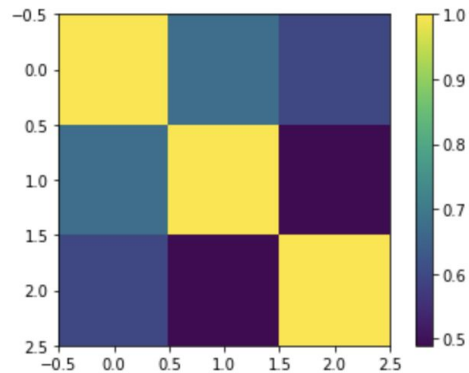
	total_bill	tip	size
total_bill	1.000000	0.675734	0.598315
tip	0.675734	1.000000	0.489299
size	0.598315	0.489299	1.000000

```
1 sns.heatmap(data.corr())
```

<AxesSubplot:>



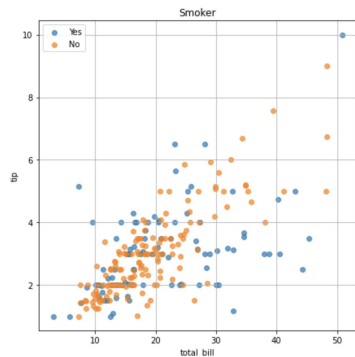
```
1 plt.imshow(data.corr())  
2 plt.colorbar();
```





А что ещё есть?

```
1 alpha = 0.7
2 plt.figure(figsize=(7, 7))
3
4 plt.scatter(data[data.smoker == 'Yes'].total_bill, data[data.smoker == 'Yes'].tip, alpha=alpha, label='Yes')
5 plt.scatter(data[data.smoker == 'No'].total_bill, data[data.smoker == 'No'].tip, alpha=alpha, label='No')
6
7 plt.title('Smoker')
8 plt.xlabel('total_bill')
9 plt.ylabel('tip')
10
11 plt.grid()
12 plt.legend();
```



- **alpha** — прозрачность
- **label** — имя графика
- **.legend()** — легенда
- **.title()** — название графика
- **.xlabel(), .ylabel()** — названия осей
- **.grid()** — сетка



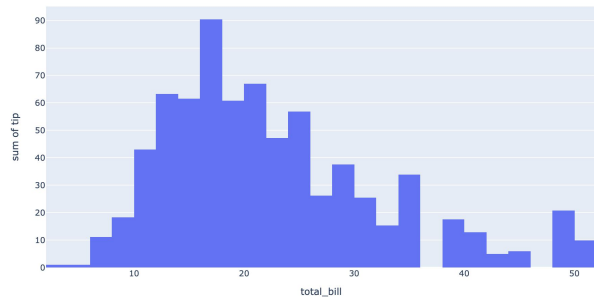
Plotly. О ней всегда забывают



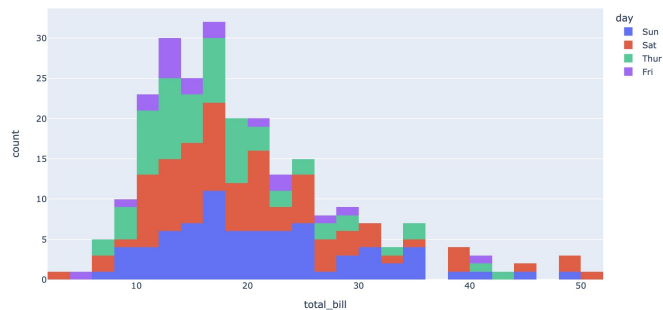


Гистограмма — Hist plot

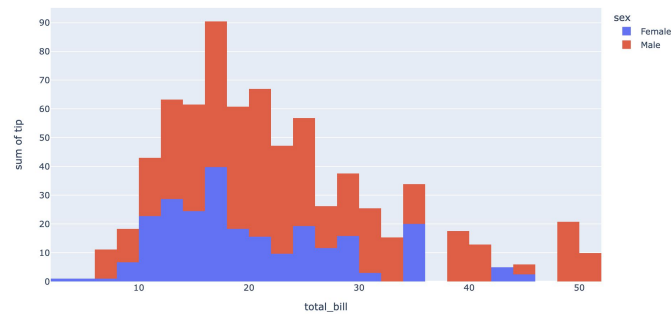
```
1 px.histogram(data, x="total_bill", y="tip")
```



```
1 px.histogram(data, x="total_bill", color="day")
```



```
1 px.histogram(data, x="total_bill", y="tip", color="sex")
```

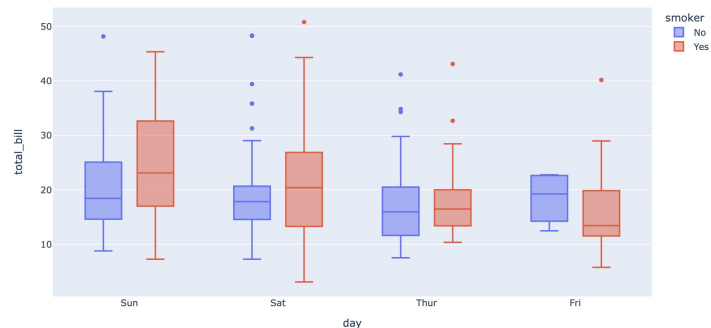
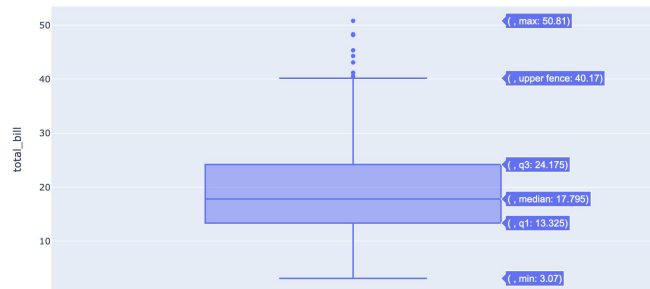


Для иллюстрации интерактивности графиков
перейдём в jupyter notebook →

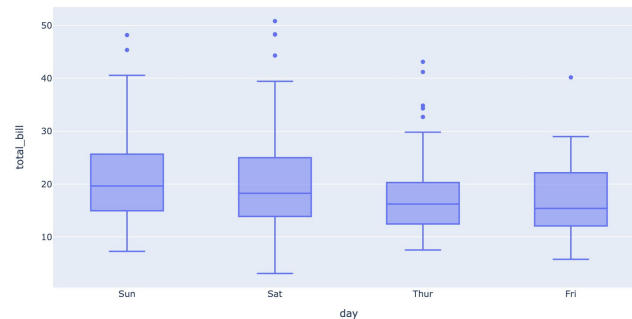


Гистограмма — Hist plot

```
1 px.box(data, y="total_bill")
```



```
1 px.box(data, x="day", y="total_bill")
```

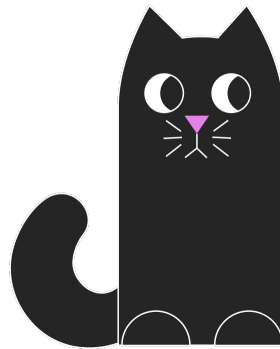


Для иллюстрации интерактивности графиков
перейдём в jupyter notebook →



Практическое задание

Ищите практическое задание в notebook с уроком.





Что мы узнали сегодня

- 📌 Обсудили, зачем визуализировать данные
- 📌 Повторили, какие бывают типы графиков
- 📌 Научились строить графики с использованием разных библиотек
- 📌 Обсудили плюсы и минусы библиотек и когда какую стоит использовать





Вопросы?

Вопросы?



Вопросы?

