

# Первые графики и выводы

## Чтение данных из файла с использованием разделителей

```
In data = read_csv('file.csv', sep=';', decimal=',')

# sep – разделитель столбцов
# decimal – разделитель десятичных знаков
```

## Импорт библиотеки matplotlib

```
In import matplotlib.pyplot as plt
```

## Отображение диаграммы ящик с усами

```
In data['column'].boxplot()
plt.show()
```

## Числовое описание данных для колонки

```
In data['column'].describe()
```

## Изменение осей графика

```
In plt.xlim(x_min, x_max)
# мин и макс для оси X
```

```
In plt.ylim(y_min, y_max)
# мин и макс для оси Y
```

## Отображение гистограммы с n\_bins, min\_value, max\_value

```
In data['column'].hist(bins=n_bins, range=(min_value, max_value))
plt.show()

# отображение гистограммы с числом корзин n_bins и отображаемыми
# минимальным и максимальным значением min_value и max_value
```

# Словарь

### Гистограмма

это график, который показывает, как часто в наборе данных встречается то или иное значение

### Распределение

это все возможные значения переменной с частотой их появления

### Нормальное распределение

чаще всего встречается среднее значение и близкие к нему, а крайние значения встречаются довольно редко

### Распределение Пуассона

число событий в единицу времени, если они в среднем происходят с измеренной частотой

### Характерный разброс

то, какие значения оказались вдали от среднего, и насколько их много

### Числовое описание данных

среднее, медиана, стандартное отклонение, количество наблюдений в выборке и разброс их значений

### Квартиль

число, разбивающее выборку на две части: значения в одной части меньше квартиля, а в другой - больше

### Первый квартиль Q1

25% элементов меньше, а 75% – больше него

### Второй квартиль Q2 (медиана)

половина элементов больше и половина меньше неё

### Третий квартиль Q3

75% элементов меньше и 25% элементов больше него

### Межквартильный размах

это расстояние между первым Q1 и третьим Q3 квартилями

### Стандартное отклонение

числовая характеристика данных, которая показывает, насколько значения в выборке отличаются от среднего арифметического