

Основные понятия модуля

Математическая статистика — раздел математики, который занимается систематизацией и обработкой данных для их использования, для получения выводов.

Статистические данные — упорядоченные, классифицированные данные о каком-то явлении или процессе.

Описательная статистика применяется, чтобы извлечь из данных полезные идеи.

Меры центральной тенденции

| Мера | Описание |
|---------|---|
| Среднее | Сумма всех элементов, поделенная на количество элементов в числовом ряду. |
| Медиана | Средний элемент упорядоченного числового ряда. |
| Мода | Самый часто встречающийся элемент в числовом ряду. |

Корреляция

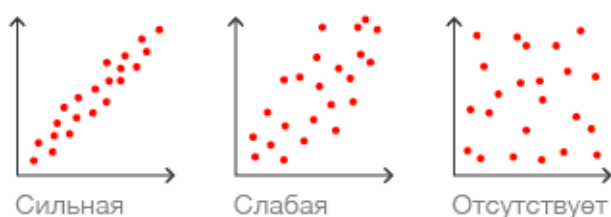
| Связь | Описание |
|---------------|--|
| Положительная | Коэффициент корреляции > 0 . Увеличение одной из величин ведёт к закономерному увеличению другой (других) величин. |
| Отрицательная | Коэффициент корреляции < 0 . Увеличение одной из величин ведёт к закономерному уменьшению другой (других) величин. |
| Нулевая | Переменные независимы друг от друга. |

Типы корреляций

| Тип | Метод | Описание |
|----------|---------------------|--|
| Пирсона | method = 'pearson' | Используется для непрерывных переменных с нормальным распределением |
| Спирмена | method = 'spearman' | Используется для вычисления взаимосвязей между категориальными переменными. |
| Кендалла | method = 'kendall' | Используется для вычисления взаимосвязей между категориальными переменными. Более устойчива к ошибкам и выбросам в данных. |

Типы связи

| Связь | Коэффициент корреляции |
|---|------------------------|
| Отсутствие связи или очень слабая связь | 0 - +/-0.3 |
| Слабая связь | +/-0.3 - +/-0.5 |
| Средняя связь | +/-0.5 - +/-0.7 |
| Сильная связь | +/-0.7 - +/-0.9 |
| Очень сильная или абсолютная связь | +/-0.9 - +/-1 |

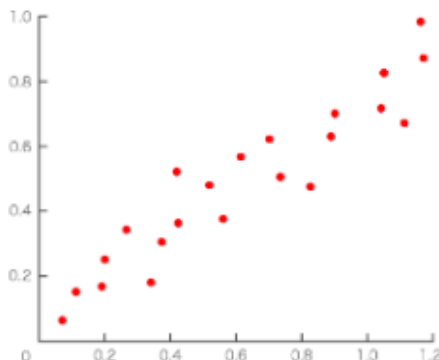


Визуализация корреляций (библиотека Seaborn)

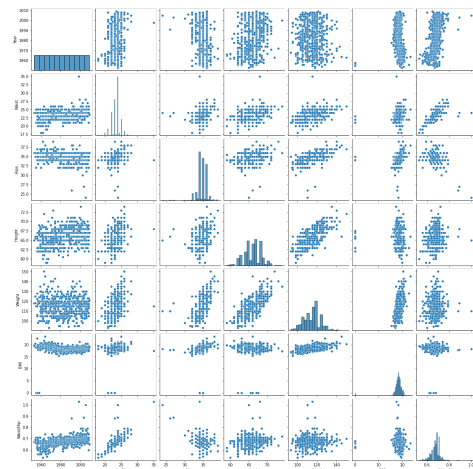
Связь

scatterplot()

Коэффициент корреляции



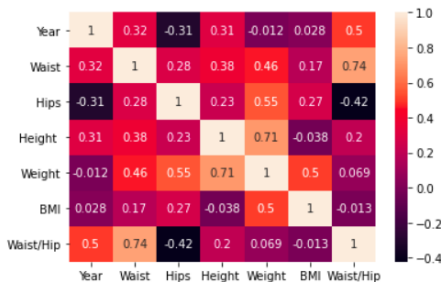
pairplot()



heatmap()

sns.heatmap(data.corr(), annot = True)

<AxesSubplot:>



| | Year | Waist | Hips | Height | Weight | BMI | Waist/Hip |
|-----------|--------|-------|-------|--------|--------|--------|-----------|
| Year | 1 | 0.32 | -0.31 | 0.31 | -0.012 | 0.028 | 0.5 |
| Waist | 0.32 | 1 | 0.28 | 0.38 | 0.46 | 0.17 | 0.74 |
| Hips | -0.31 | 0.28 | 1 | 0.23 | 0.55 | 0.27 | -0.42 |
| Height | 0.31 | 0.38 | 0.23 | 1 | 0.71 | -0.038 | 0.2 |
| Weight | -0.012 | 0.46 | 0.55 | 0.71 | 1 | 0.5 | 0.069 |
| BMI | 0.028 | 0.17 | 0.27 | -0.038 | 0.5 | 1 | -0.013 |
| Waist/Hip | 0.5 | 0.74 | -0.42 | 0.2 | 0.069 | -0.013 | 1 |