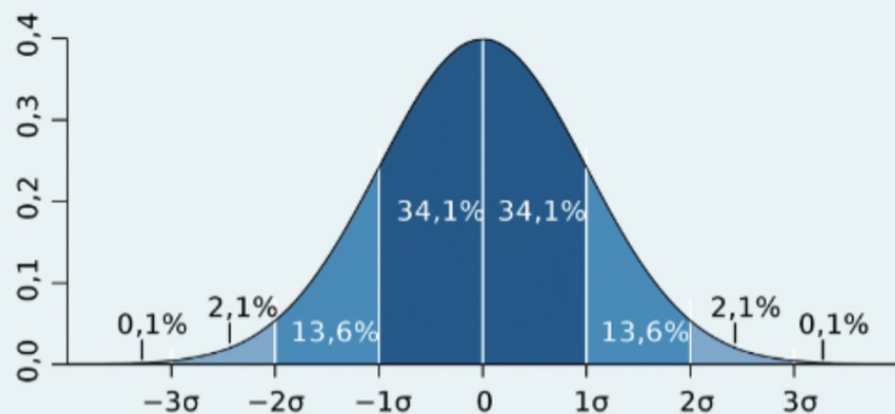


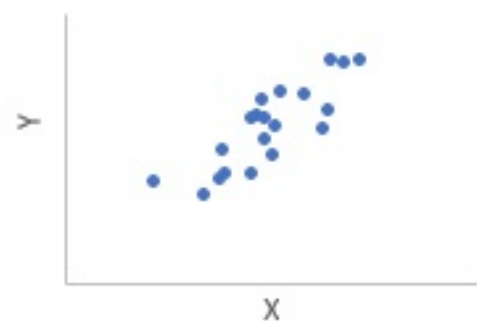
Предположим, что распределение веса всех взрослых женщин на планете нормальное. Если среднее — 62 кг, а среднеквадратичное отклонение — 8 кг, то какие выводы можно сделать?



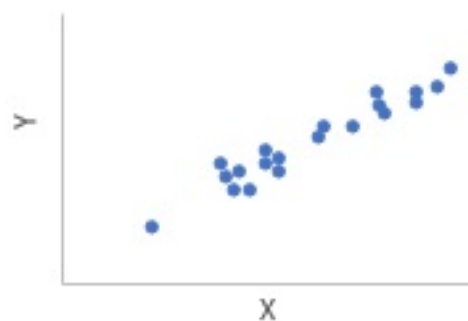
- ☐ 47.7 % взрослых женщин весит от 62 до 86 кг.
- ☐ 13.6 % взрослых женщин весит от 46 до 62 кг
- ☐ 0.1 % взрослых женщин весит больше 86 кг.
- ☐ Вес 68.2 % взрослых женщин находится между 58 и 68 кг.

Мы хотим узнать, какой средний уровень образования у совершеннолетних женщин в России. Из вариантов ниже выберите тот, в котором не происходит гарантированного смещения выборки:

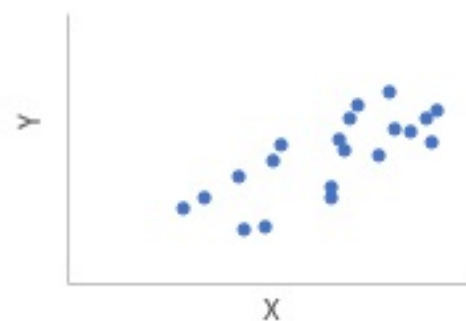
- ☐ Опрос был проведен посредством социальных сетей для их пользователей.
- ☐ Были опрошены только работающие женщины.
- ☐ Было опрошено такое количество женщин из каждого региона, которое соответствует доле женщин, проживающих в этом регионе согласно переписи населения.
- ☐ Были опрошены жительницы всех городов с населением более 300 000 человек.



Прямая



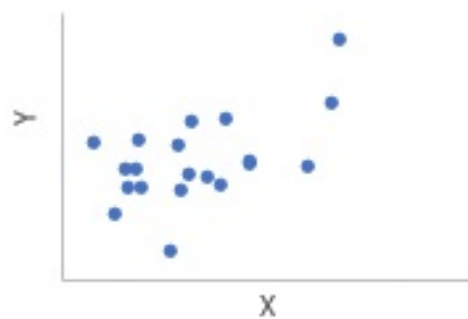
Сильная



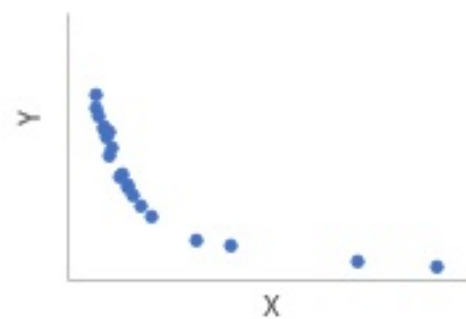
Линейная



Обратная

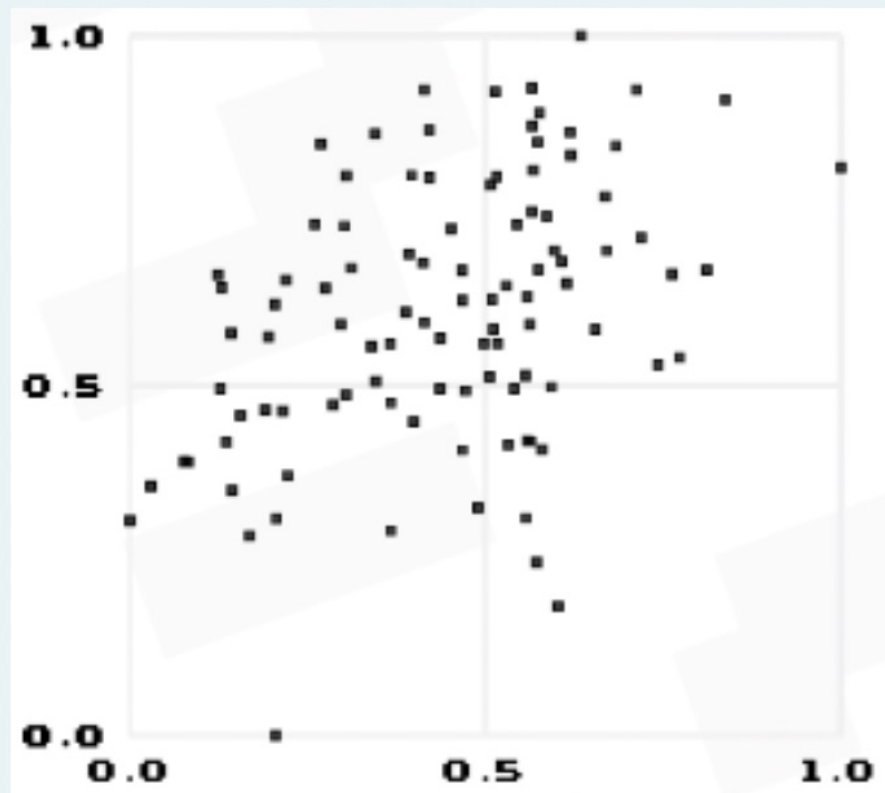


Слабая



Нелинейная

Оцените коэффициент корреляции данных на графике.



- ☐ - 0.21
- ☐ 0.82
- ☐ 0.41
- ☐ -0.54

Мы провели исследование и выявили, что у сотрудников компании уровень удовлетворенности трудом коррелирует с их продуктивностью на работе, коэффициент равен 0.81. Какие выводы можно точно сделать из этого наблюдения?

- ☐ Между удовлетворенностью работой и продуктивностью слабая прямая взаимосвязь.
- ☐ Если у сотрудника повышается продуктивность, то, скорее всего, повысится и удовлетворенность работой.
- ☐ Продуктивность на работе влияет на удовлетворенность собственным трудом.
- ☐ Удовлетворенность собственным трудом напрямую влияет на продуктивность.

Кот Матроскин ежедневно в течение недели ходил на рыбалку, чтобы поймать больше рыбок для своего любимого аквариума. Оказалось, что пойманные им рыбки обладают следующими характеристиками:

- Цвет: [серая, серая, золотая, синяя, золотая]
- Длина (см): [15, 12.1, 20, 24.2, 18.7]
- Масса (г): [800, 495, 302, 1001, 256]

Выберите две характеристики, между которыми можно корректно рассчитать выборочный коэффициент корреляции Пирсона, и вычислите этот коэффициент по приведённым данным. Ответ округлите до сотых.

Пример ответа: 0.97

Ответ:

Выбросы Outliers

это наблюдение в анализируемых данных, значение которого сильно отличается от других.

Как опознать выброс? Эвристики.
Например, расстояние три стандартных отклонения от среднего.

С точки зрения исследователя:
принять решение, как их
“нейтрализовать”.

С точки зрения журналиста:
что-то, что нужно расследовать!

Сила выбросов

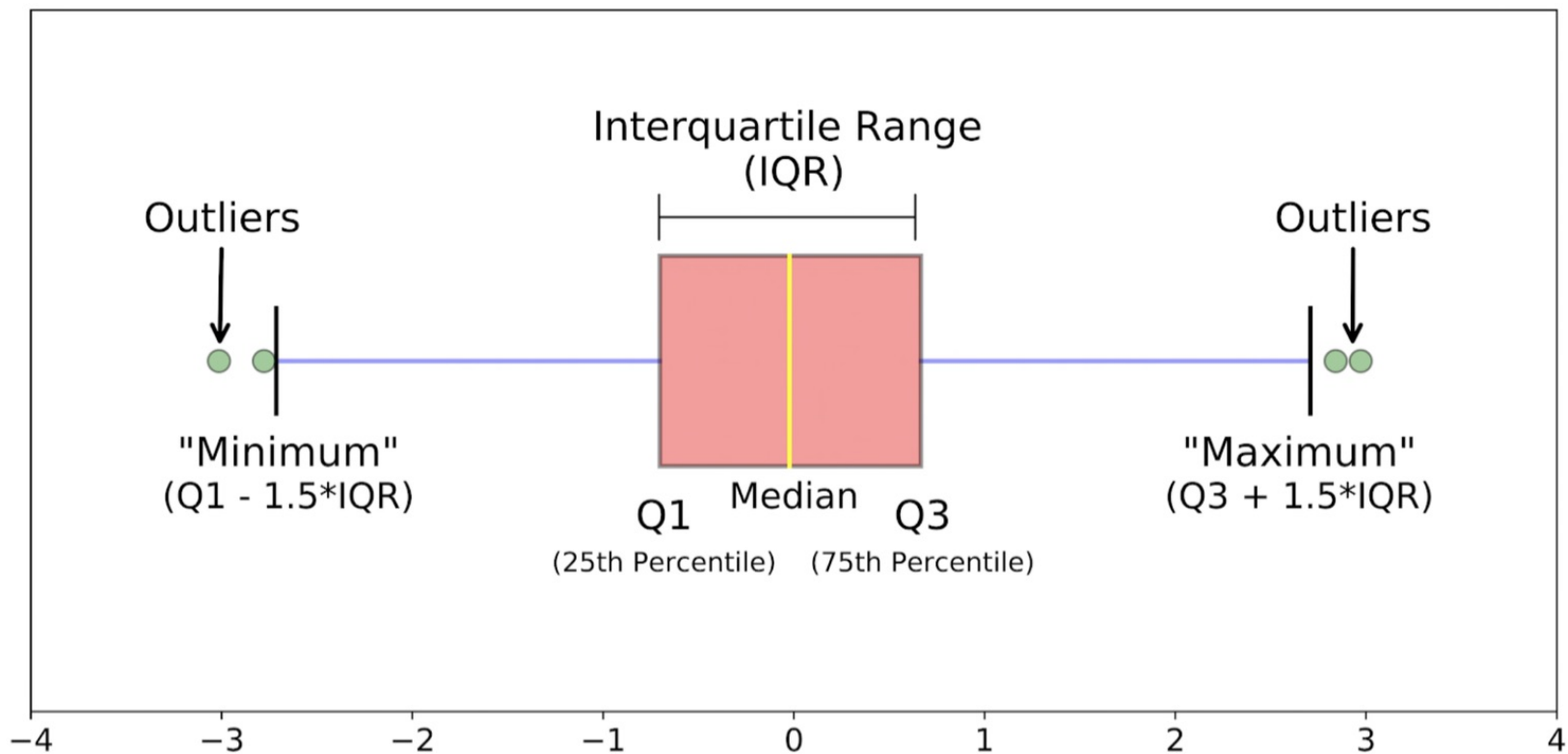
Слабые выбросы

- Те значения, которые меньше 25% перцентиля минус $1,5 \cdot \text{ИКР}$ (интерквартильный размах размах) или больше 75% персентиля плюс $1,5 \cdot \text{ИКР}$.
- Или находятся от среднего на расстоянии от двух до трех стандартных отклонений.

Сильные выбросы

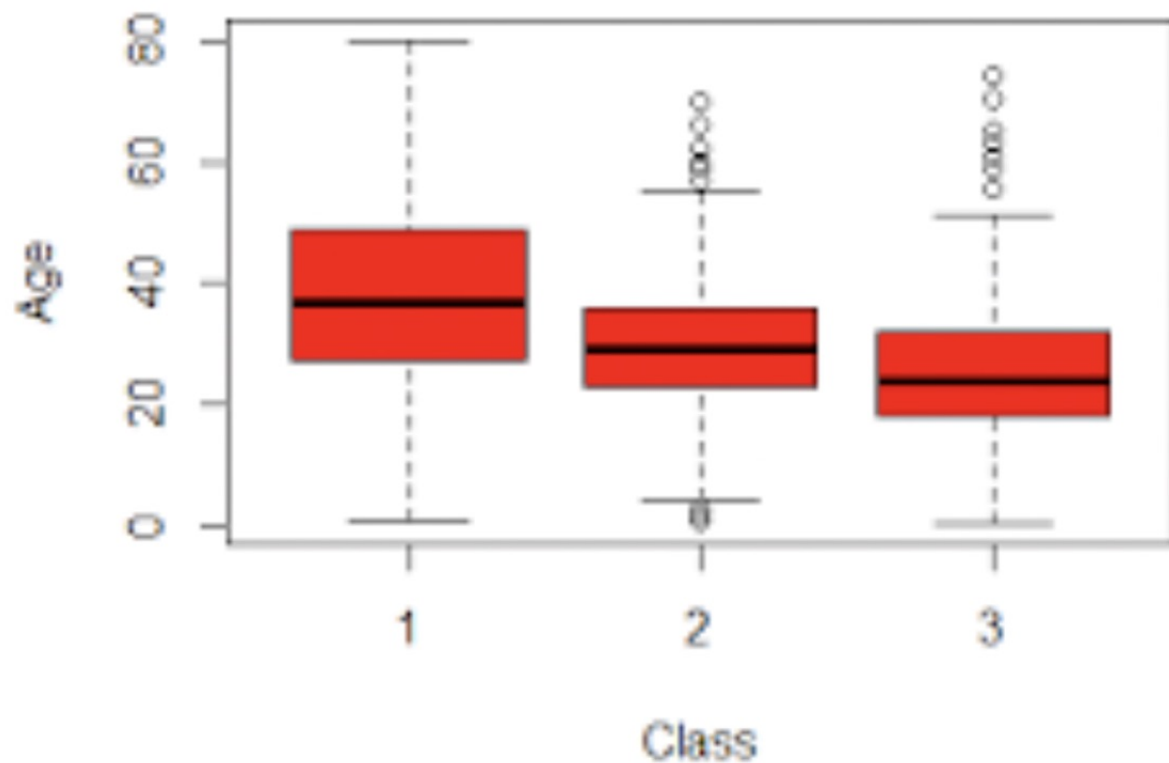
- Те значения, которые меньше 25% перцентиля минус $3 \cdot \text{ИКР}$ (межквартильный размах) или больше 75% персентиля плюс $3 \cdot \text{ИКР}$.
- Или дальше от среднего чем три стандартных отклонения.

Выбросы



В ящике-с-усами, усы как раз отграничивают выборку от выбросов.

Выбросы



Найдите выбросы
на графике.

