

Обработка данных с помощью регрессии на Python

Регрессия — это метод статистического анализа, который используется для моделирования и анализа зависимостей между переменными. Основная цель регрессии — предсказать значение одной переменной (зависимой) на основе значений других переменных (независимых).

Существует несколько типов регрессии:

Линейная регрессия: Моделирует линейную зависимость между переменными.

Полиномиальная регрессия: Расширяет линейную регрессию, позволяя моделировать нелинейные зависимости.

Логистическая регрессия: Используется для бинарной классификации.

Регрессия с регуляризацией: Например, Lasso и Ridge, которые помогают избежать переобучения.

Обработка данных

Перед тем как применять регрессию, необходимо провести предварительную обработку данных. Это включает в себя несколько этапов:

Сбор данных: Данные могут быть собраны из различных источников (файлы CSV, базы данных, API и т.д.).

Очистка данных:

Удаление или заполнение пропущенных значений.

Устранение дубликатов.

Исправление ошибок в данных.

Преобразование данных:

Для обработки данных с помощью регрессии на Python, а также для определения коэффициента корреляции и коэффициента детерминации, можно использовать библиотеки pandas, numpy, и scikit-learn. Ниже приведен пример, который демонстрирует, как это сделать.

Пример: Регрессия на основе данных о росте и весе

В этом примере мы будем использовать случайные данные о росте и весе людей. Мы определим коэффициент корреляции и коэффициент детерминации, а также построим линейную регрессионную модель.

Шаг 1: Установка необходимых библиотек

Если у вас еще не установлены необходимые библиотеки, вы можете установить их с помощью pip:

```
bash
pip install pandas numpy scikit-learn matplotlib
```

Шаг 2: Импорт библиотек

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression
from sklearn.metrics import r2_score
```

Шаг 3: Создание данных

Создадим случайные данные о росте и весе:

```
# Генерация случайных данных
np.random.seed(0)
height = np.random.normal(170, 10, 100) # Рост в см
weight = height * 0.5 + np.random.normal(0, 5, 100) + 50 # Вес в кг с некоторым шумом
```

```
# Создание DataFrame
```

```
data = pd.DataFrame({'Height': height, 'Weight': weight})
```

Шаг 4: Определение коэффициента корреляции

Коэффициент корреляции можно вычислить с помощью метода `.corr()`:

```
correlation = data.corr().loc['Height', 'Weight']
```

```
print(f'Коэффициент корреляции между ростом и весом: {correlation:.2f}')
```

Шаг 5: Построение линейной регрессионной модели

Теперь мы создадим линейную регрессионную модель:

```
# Определение переменных X и y
```

```
X = data[['Height']] # Независимая переменная (рост)
```

```
y = data['Weight'] # Зависимая переменная (вес)
```

```
# Создание модели линейной регрессии
```

```
model = LinearRegression()
```

```
model.fit(X, y)
```

```
# Предсказание значений веса на основе роста
```

```
y_pred = model.predict(X)
```

```
# Коэффициент детерминации (R2)
```

```
r_squared = r2_score(y, y_pred)
```

```
print(f'Коэффициент детерминации (R2): {r_squared:.2f}')
```

Шаг 6: Визуализация результатов

визуализируем результаты:

```
# Построение графика
```

```
plt.scatter(data['Height'], data['Weight'], color='blue', label='Данные')
```

```
plt.plot(data['Height'], y_pred, color='red', label='Линейная регрессия')
```

```
plt.xlabel('Рост (см)')
```

```
plt.ylabel('Вес (кг)')
```

```
plt.title('Зависимость веса от роста')
```

```
plt.legend()
```

```
plt.show()
```

Код для выполнения:

```
import pandas as pd
```

```
import numpy as np
```

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

```
from sklearn.linear_model import LinearRegression
```

```
from sklearn.metrics import r2_score
```

```
# Генерация случайных данных
```

```
np.random.seed(0)
```

```
height = np.random.normal(170, 10, 100) # Рост в см
```

```
weight = height * 0.5 + np.random.normal(0, 5, 100) + 50 # Вес в кг с некоторым шумом
```

```
# Создание DataFrame
```

```
data = pd.DataFrame({'Height': height, 'Weight': weight})
```

```
# Определение коэффициента корреляции
```

```
correlation = data.corr().loc['Height', 'Weight']
```

```
print(f'Коэффициент корреляции между ростом и весом: {correlation:.2f}')
```

```
# Определение переменных X и y
```

```
X = data[['Height']]
```

```
y = data['Weight']
```

```
# Создание модели линейной регрессии
```

```

model = LinearRegression()
model.fit(X, y)
# Предсказание значений веса на основе роста
y_pred = model.predict(X)
# Коэффициент детерминации (R2)
r_squared = r2_score(y, y_pred)
print(f'Коэффициент детерминации (R2): {r_squared:.2f}')
# Построение графика
plt.scatter(data['Height'], data['Weight'], color='blue', label='Данные')
plt.plot(data['Height'], y_pred, color='red', label='Линейная регрессия')
plt.xlabel('Рост (см)')
plt.ylabel('Вес (кг)')
plt.title('Зависимость веса от роста')
plt.legend()
plt.show()

```

Результаты:

После выполнения кода вы получите:

Коэффициент корреляции между ростом и весом.

Коэффициент детерминации R^2 , который показывает долю вариации зависимой переменной (веса), объясняемую независимой переменной (ростом).

График зависимости веса от роста с наложенной линией регрессии.

В следующем примере мы создадим набор данных, содержащий информацию о площади жилья и его цене, а затем вычислим коэффициент корреляции и коэффициент детерминации.

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Создание набора данных
data = {
    'area': [1500, 2000, 2500, 3000, 3500],
    'price': [300000, 400000, 500000, 600000, 700000]
}
df = pd.DataFrame(data)

# Вычисление коэффициента корреляции
correlation_matrix = df.corr()
correlation = correlation_matrix.loc['area', 'price']
print("Коэффициент корреляции между площадью и ценой:", correlation)

# Построение линейной регрессии для определения R2
X = df[['area']]
y = df['price']

model = LinearRegression()
model.fit(X, y)

# Предсказание цен на основе площади
y_pred = model.predict(X)

# Вычисление R2 (коэффициент детерминации)

```

```
r_squared = model.score(X, y)
print("Коэффициент детерминации R^2:", r_squared)

# Визуализация данных и линии регрессии
plt.scatter(df['area'], df['price'], color='blue', label='Данные')
plt.plot(df['area'], y_pred, color='red', label='Линия регрессии')
plt.xlabel('Площадь (кв. футы)')
plt.ylabel('Цена ($)')
plt.title('Зависимость цены от площади')
plt.legend()
plt.show()
```

Создание набора данных: Мы создаем DataFrame с двумя переменными: area (площадь) и price (цена).

Коэффициент корреляции: Мы используем метод `.corr()` для вычисления матрицы корреляций. Затем извлекаем значение корреляции между площадью и ценой.

Линейная регрессия: Мы создаем модель линейной регрессии с помощью `LinearRegression` из библиотеки `sklearn`. Обучаем модель на наших данных.

Коэффициент детерминации (R^2): Мы используем метод `.score()` модели для получения значения R^2 .

Визуализация: Мы строим график рассеяния для визуализации зависимости между площадью и ценой, а также добавляем линию регрессии.

Результаты:

Коэффициент корреляции показывает степень линейной зависимости между двумя переменными.

Коэффициент детерминации (R^2) показывает долю вариации зависимой переменной (цены), объясняемую независимой переменной (площадью). Значение R^2 варьируется от 0 до 1; чем ближе к 1, тем лучше модель объясняет данные.

Этот пример демонстрирует основные шаги для вычисления коэффициента корреляции и коэффициента детерминации в Python.

Задания для выполнения

Каждая из этих задач может быть решена с использованием методов вычисления коэффициента корреляции (pearson, spearman или kendall) с помощью библиотеки pandas, а также определения коэффициента детерминации при помощи линейной регрессии из библиотеки scikit-learn.

1. Зависимость роста растений от количества удобрений.

Данные: высота растений (см) и количество удобрений (г).

Найти корреляцию между количеством удобрений и ростом растений.

Построить график зависимости.

| № | высота растений (см) | количество удобрений (г). |
|-----|----------------------|---------------------------|
| 1. | 11 | 5 |
| 2. | 12 | 1 |
| 3. | 5 | 12 |
| 4. | 7 | 3 |
| 5. | 9 | 5 |
| 6. | 23 | 8 |
| 7. | 2 | 9 |
| 8. | 9 | 4 |
| 9. | 5 | 3 |
| 10. | 9 | 6 |
| 11. | 7 | 5 |
| 12. | 12 | 1 |
| 13. | 11 | 23 |
| 14. | 13 | 2 |
| 15. | 9 | 5 |
| 16. | 7 | 4 |
| 17. | 2 | 6 |
| 18. | 1 | 9 |
| 19. | 6 | 8 |
| 20. | 18 | 2 |
| 21. | 23 | 8 |

2. Влияние температуры на продажи мороженого

Данные: температура (°C) и количество проданного мороженого.

Задача: Определить, как температура влияет на продажи.

Построить график зависимости.

| № | температура (°C) | количество проданного мороженого (ед). |
|----|------------------|--|
| 1. | 30 | 75 |
| 2. | 12 | 12 |

| | | |
|-----|----|-----|
| 3. | 5 | 1 |
| 4. | 35 | 140 |
| 5. | 9 | 5 |
| 6. | 23 | 34 |
| 7. | 2 | 0 |
| 8. | 40 | 340 |
| 9. | 5 | 3 |
| 10. | 10 | 14 |
| 11. | 8 | 11 |
| 12. | 12 | 1 |
| 13. | 23 | 18 |
| 14. | 13 | 2 |
| 15. | 40 | 520 |
| 16. | 12 | 16 |
| 17. | 12 | 5 |
| 18. | 0 | 2 |
| 19. | 18 | 6 |
| 20. | 12 | 20 |
| 21. | 37 | 120 |

3. Связь между временем обучения и оценками

Данные: время обучения (часы) и оценки студентов.

Задача: Найти корреляцию между временем обучения и оценками (баллы).

Построить график зависимости.

| № | время обучения (часы) | оценки студентов (баллы) |
|-----|-----------------------|--------------------------|
| 1. | 50 | 50 |
| 2. | 55 | 55 |
| 3. | 60 | 60 |
| 4. | 65 | 50 |
| 5. | 70 | 55 |
| 6. | 75 | 60 |
| 7. | 80 | 65 |
| 8. | 85 | 70 |
| 9. | 90 | 75 |
| 10. | 95 | 55 |
| 11. | 100 | 80 |
| 12. | 105 | 85 |
| 13. | 110 | 60 |
| 14. | 120 | 55 |
| 15. | 130 | 60 |
| 16. | 140 | 90 |

| | | |
|-----|-----|----|
| 17. | 145 | 95 |
| 18. | 150 | 95 |
| 19. | 155 | 95 |
| 20. | 160 | 90 |
| 21. | 165 | 10 |

4. Влияние рекламы на продажи

Данные: расходы на рекламу (тг) и объем продаж(тг).

Задача: Определить, как реклама влияет на продажи.

Построить график зависимости.

| № | расходы на рекламу (тг) | объем продаж (тг) |
|-----|-------------------------|-------------------|
| 1. | 500 | 1000 |
| 2. | 600 | 1100 |
| 3. | 700 | 1350 |
| 4. | 850 | 1450 |
| 5. | 700 | 1350 |
| 6. | 750 | 1400 |
| 7. | 850 | 1500 |
| 8. | 900 | 1600 |
| 9. | 1200 | 1800 |
| 10. | 1300 | 2600 |
| 11. | 1400 | 3700 |
| 12. | 1600 | 9800 |
| 13. | 2500 | 25000 |
| 14. | 4500 | 34000 |
| 15. | 5000 | 90000 |
| 16. | 1700 | 10700 |
| 17. | 1800 | 11000 |
| 18. | 6000 | 118000 |
| 19. | 7000 | 78000 |
| 20. | 9000 | 89000 |
| 21. | 9800 | 95000 |

5. Связь между количеством часов сна и продуктивностью

Данные: количество часов сна и уровень продуктивности (баллы).

Задача: Определить, как сон влияет на продуктивность.

Построить график зависимости.

| № | количеством часов сна | продуктивности |
|----|-----------------------|----------------|
| 1. | 1 | 50 |
| 2. | 2 | 100 |
| 3. | 3 | 150 |

| | | |
|-----|----|-----|
| 4. | 4 | 200 |
| 5. | 5 | 250 |
| 6. | 6 | 300 |
| 7. | 7 | 350 |
| 8. | 8 | 400 |
| 9. | 9 | 450 |
| 10. | 10 | 500 |
| 11. | 11 | 550 |
| 12. | 12 | 600 |
| 13. | 13 | 650 |

6. Влияние возраста на физическую активность

Данные: возраст (годы) и количество шагов в день.

Задача: Найти корреляцию между возрастом и физической активностью.

Построить график зависимости.

| № | возраст (годы) | количество шагов в день |
|-----|----------------|-------------------------|
| 1. | 25 | 500 |
| 2. | 30 | 1000 |
| 3. | 35 | 1100 |
| 4. | 40 | 1200 |
| 5. | 45 | 1300 |
| 6. | 50 | 1400 |
| 7. | 55 | 5000 |
| 8. | 60 | 1500 |
| 9. | 65 | 1600 |
| 10. | 70 | 1700 |
| 11. | 80 | 1900 |

7. Зависимость цен на нефть от курса доллара

Данные: цена нефти (₽) и курс доллара (₽/\$).

Задача: Определить, как курс доллара влияет на цены на нефть.

Построить график зависимости.

| № | цена нефти | курс доллара |
|----|------------|--------------|
| 1. | 500 | 100 |
| 2. | 550 | 150 |
| 3. | 600 | 200 |
| 4. | 650 | 250 |
| 5. | 700 | 300 |
| 6. | 750 | 350 |
| 7. | 800 | 400 |
| 8. | 850 | 450 |

| | | |
|-----|------|------|
| 9. | 900 | 500 |
| 10. | 950 | 550 |
| 11. | 1000 | 600 |
| 12. | 1050 | 650 |
| 13. | 1100 | 700 |
| 14. | 1200 | 750 |
| 15. | 1300 | 800 |
| 16. | 1400 | 850 |
| 17. | 1450 | 900 |
| 18. | 1500 | 950 |
| 19. | 1550 | 1000 |
| 20. | 1600 | 1000 |
| 21. | 1650 | 100 |

8. Связь между уровнем стресса и здоровьем

Данные: уровень стресса (баллы) и количество заболеваний.

Задача: Найти корреляцию между стрессом и здоровьем.

Построить график зависимости.

| № | уровень стресса (баллы) | количество заболеваний |
|----|-------------------------|------------------------|
| 1. | 1 | 3 |
| 2. | 2 | 5 |
| 3. | 3 | 9 |
| 4. | 4 | 15 |
| 5. | 5 | 17 |

9. Влияние времени в социальных сетях на настроение

Данные: время в социальных сетях (часы) и уровень счастья (баллы).

Задача: Определить, как время в социальных сетях влияет на настроение.

Построить график зависимости.

| № | время в социальных сетях (часы) | уровень счастья (баллы) |
|-----|---------------------------------|-------------------------|
| 1. | 1 | 1 |
| 2. | 2 | 1,5 |
| 3. | 3 | 2 |
| 4. | 4 | 0,5 |
| 5. | 5 | 3 |
| 6. | 6 | 3,7 |
| 7. | 7 | 2,5 |
| 8. | 8 | 0,7 |
| 9. | 9 | 2,3 |
| 10. | 10 | 6 |
| 11. | 11 | 7 |

10. Зависимость веса от роста

- Данные: рост (см) и вес (кг).

- Задача: Найти корреляцию между ростом и весом.

Построить график зависимости.

| № | рост (см) | вес (кг) |
|-----|-----------|----------|
| 1. | 50 | 15 |
| 2. | 60 | 20 |
| 3. | 70 | 30 |
| 4. | 80 | 35 |
| 5. | 90 | 40 |
| 6. | 100 | 45 |
| 7. | 110 | 50 |
| 8. | 120 | 55 |
| 9. | 130 | 60 |
| 10. | 140 | 65 |
| 11. | 150 | 70 |
| 12. | 160 | 75 |
| 13. | 170 | 100 |
| 14. | 180 | 110 |
| 15. | 190 | 120 |

11. Связь между количеством кофеина и уровнем энергии

- Данные: количество кофеина (мг) и уровень энергии (баллы).

- Задача: Определить, как кофеин влияет на уровень энергии.

Построить график зависимости.

| № | количество кофеина (мг) | уровень энергии (баллы) |
|-----|-------------------------|-------------------------|
| 1. | 3 | 15 |
| 2. | 4 | 20 |
| 3. | 5 | 30 |
| 4. | 6 | 35 |
| 5. | 7 | 40 |
| 6. | 8 | 45 |
| 7. | 9 | 50 |
| 8. | 10 | 55 |
| 9. | 11 | 60 |
| 10. | 12 | 65 |
| 11. | 13 | 70 |
| 12. | 14 | 75 |
| 13. | 15 | 100 |
| 14. | 16 | 110 |

| | | |
|-----|----|-----|
| 15. | 17 | 120 |
|-----|----|-----|

12. Связь между уровнем загрязнения воздуха и заболеваемостью

- Данные: уровень загрязнения (мкг/м³) и количество заболеваний.

- Задача: Найти корреляцию между загрязнением воздуха и здоровьем населения.

Построить график зависимости.

| № | уровень загрязнения (мкг/м ³) | количество заболеваний |
|-----|---|------------------------|
| 1. | 500 | 1 |
| 2. | 550 | 1 |
| 3. | 600 | 2 |
| 4. | 650 | 2 |
| 5. | 700 | 3 |
| 6. | 750 | 4 |
| 7. | 800 | 4 |
| 8. | 850 | 4 |
| 9. | 900 | 5 |
| 10. | 950 | 5 |
| 11. | 1000 | 6 |
| 12. | 1050 | 6 |
| 13. | 1100 | 7 |
| 14. | 1200 | 7 |
| 15. | 1300 | 8 |
| 16. | 1400 | 8 |
| 17. | 1450 | 9 |
| 18. | 1500 | 9 |
| 19. | 1550 | 10 |
| 20. | 1600 | 10 |
| 21. | 1650 | 10 |

13. Влияние цен на жилье на спрос

- Данные: цена жилья (\$) и количество проданных домов.

- Задача: Определить, как цена жилья влияет на спрос.

Построить график зависимости.

| № | цена жилья (\$) | количество проданных домов |
|----|-----------------|----------------------------|
| 1. | 5000 | 10 |
| 2. | 5500 | 10 |
| 3. | 6000 | 20 |
| 4. | 6500 | 20 |
| 5. | 7000 | 30 |

| | | |
|-----|-------|-----|
| 6. | 7500 | 40 |
| 7. | 8000 | 40 |
| 8. | 8500 | 40 |
| 9. | 9000 | 50 |
| 10. | 9500 | 50 |
| 11. | 10000 | 60 |
| 12. | 10500 | 60 |
| 13. | 11000 | 70 |
| 14. | 12000 | 70 |
| 15. | 13000 | 80 |
| 16. | 14000 | 80 |
| 17. | 14500 | 90 |
| 18. | 15000 | 90 |
| 19. | 15500 | 100 |
| 20. | 16000 | 100 |
| 21. | 16500 | 100 |

14. Зависимость времени работы от зарплаты

- Данные: часы работы в неделю и зарплата (\$).

- Задача: Найти корреляцию между часами работы и зарплатой.

Построить график зависимости.

| № | часы работы | зарплата (\$) |
|----|-------------|---------------|
| 1. | 24 | 1000 |
| 2. | 36 | 1200 |
| 3. | 48 | 1500 |
| 4. | 64 | 1800 |
| 5. | 72 | 2200 |

15. Связь между количеством тренировок в неделю и физической формой

Данные: количество тренировок в неделю и уровень физической формы (баллы).

Задача: Определить, как тренировки влияют на физическую форму.

Построить график зависимости.

| № | количество тренировок | уровень физической формы (баллы) |
|----|-----------------------|----------------------------------|
| 1. | 3 | 100 |
| 2. | 4 | 200 |
| 3. | 5 | 300 |
| 4. | 6 | 350 |
| 5. | 7 | 400 |
| 6. | 8 | 450 |

| | | |
|-----|----|------|
| 7. | 9 | 500 |
| 8. | 10 | 550 |
| 9. | 11 | 600 |
| 10. | 12 | 650 |
| 11. | 13 | 700 |
| 12. | 14 | 750 |
| 13. | 15 | 1000 |
| 14. | 16 | 1100 |
| 15. | 17 | 1200 |

16. Зависимость расходов семьи от дохода

Данные: доход семьи (\$) и расходы семьи (\$).

Задача: Определить, как доход влияет на расходы семьи.

Построить график зависимости.

| № | доход семьи (\$) | расходы семьи (\$) |
|-----|------------------|--------------------|
| 1. | 100 | 70 |
| 2. | 200 | 90 |
| 3. | 300 | 100 |
| 4. | 400 | 150 |
| 5. | 500 | 200 |
| 6. | 600 | 350 |
| 7. | 700 | 450 |
| 8. | 800 | 500 |
| 9. | 900 | 600 |
| 10. | 1000 | 700 |
| 11. | 1100 | 800 |
| 12. | 1200 | 900 |
| 13. | 1300 | 1000 |
| 14. | 1400 | 1100 |
| 15. | 1500 | 1200 |
| 16. | 1600 | 1225 |
| 17. | 1700 | 1350 |
| 18. | 1800 | 1500 |
| 19. | 1900 | 1600 |
| 20. | 2000 | 1900 |

17. Связь между количеством книг, прочитанных за год, и уровнем интеллекта

Данные: количество прочитанных книг за год и IQ.

Задача: Найти корреляцию между чтением книг и уровнем интеллекта.

Построить график зависимости.

| № | количество прочитанных книг | IQ |
|-----|-----------------------------|-----|
| 1. | 5 | 10 |
| 2. | 7 | 10 |
| 3. | 9 | 20 |
| 4. | 11 | 20 |
| 5. | 13 | 30 |
| 6. | 15 | 40 |
| 7. | 17 | 40 |
| 8. | 19 | 40 |
| 9. | 20 | 50 |
| 10. | 22 | 50 |
| 11. | 24 | 60 |
| 12. | 26 | 60 |
| 13. | 28 | 70 |
| 14. | 30 | 70 |
| 15. | 35 | 80 |
| 16. | 40 | 80 |
| 17. | 45 | 90 |
| 18. | 50 | 90 |
| 19. | 60 | 100 |
| 20. | 65 | 100 |
| 21. | 75 | 100 |

18. Зависимость уровня настроения от финансового состояния

Данные: финансовое состояние (\$) и уровень благосостояния (баллы).

Задача: Найти корреляцию между финансами и благосостоянием.

Построить график зависимости.

| № | финансовое состояние (\$) | уровень благосостояния |
|-----|---------------------------|------------------------|
| 1. | 100 | 5 |
| 2. | 200 | 10 |
| 3. | 300 | 15 |
| 4. | 400 | 20 |
| 5. | 500 | 25 |
| 6. | 550 | 30 |
| 7. | 600 | 35 |
| 8. | 700 | 40 |
| 9. | 80 | 45 |
| 10. | 900 | 50 |
| 11. | 1000 | 55 |
| 12. | 1100 | 60 |
| 13. | 1200 | 65 |

| | | |
|-----|------|-----|
| 14. | 1300 | 70 |
| 15. | 1400 | 75 |
| 16. | 1400 | 80 |
| 17. | 1500 | 85 |
| 18. | 1600 | 90 |
| 19. | 1700 | 95 |
| 20. | 1800 | 100 |
| 21. | 2000 | 110 |

19. Связь между количеством времени, проведенного за компьютером, с проблемами со зрением

Данные: время за компьютером (часы/день) и степень проблем со зрением.

Задача: Определить влияние времени за компьютером на здоровье глаз.

Построить график зависимости.

| № | время за компьютером | степень проблем со зрением |
|-----|----------------------|----------------------------|
| 1. | 1 | 1 |
| 2. | 2 | 1 |
| 3. | 3 | 1 |
| 4. | 4 | 2 |
| 5. | 5 | 2 |
| 6. | 6 | 2 |
| 7. | 7 | 2 |
| 8. | 8 | 2 |
| 9. | 9 | 3 |
| 10. | 10 | 3 |
| 11. | 11 | 3 |
| 12. | 12 | 3 |
| 13. | 13 | 3 |
| 14. | 14 | 4 |
| 15. | 15 | 4 |
| 16. | 16 | 4 |
| 17. | 17 | 5 |
| 18. | 18 | 5 |
| 19. | 19 | 5 |
| 20. | 20 | 5 |
| 21. | 21 | |

20. Влияние диеты на уровень холестерина

Данные: тип диеты (баллы по шкале здоровья) против уровня холестерина.

Задача: Найти корреляцию между диетой человека кал уровнем холестерина в крови.

Построить график зависимости.

| № | тип диеты | уровень холестерина |
|----|-----------|---------------------|
| 1. | 1 | 50 |
| 2. | 2 | 100 |
| 3. | 3 | 150 |
| 4. | 4 | 200 |
| 5. | 5 | 250 |