

# Зачётное задание по дисциплине «Основы анализа данных»

---

## Оглавление

### 1. Теоритическая часть

- Минимальное, максимальное значения
- Сумма
- Среднее арифметическое
- Медиана
- Мода
- Среднеквадратичное отклонение

### 2. Техническая часть

- Начало работы
- Вычисления
- Результаты

## Теоритическая часть

### Минимальное, максимальное значения

- **Минимальное** - минимальное значение в заданном промежутке, массиве.
- **Максимальное** - максимальное значение в заданном промежутке, массиве.

---

### Сумма

**Сумма** - результат применения операции сложения величин, либо результат последовательного выполнения нескольких операций сложения.

---

### Среднее арифметическое

**Среднее арифметическое** - разновидность среднего значения. Определяется как число, равное сумме всех чисел множества, делённой на их количество.

---

### Медиана

**Медиана** - число, которое находится в середине этого набора, если его упорядочить по возрастанию, то есть такое число, что половина из элементов набора не меньше него, а другая половина не больше.

---

## Мода

**Мода** - значение во множестве наблюдений, которое встречается наиболее часто.

---

## Среднеквадратичное отклонение

**Среднеквадратичное отклонение** - наиболее распространённый показатель рассеивания значений случайной величины относительно её.

---

## Техническая часть

### Начало работы

Импортируем необходимые библиотеки и методы для статистических вычислений:

```
import csv
from statistics import mode
from numpy import median, std
```

Получим данные из файла:

```
with open('diamonds.csv') as file:
    reader = csv.reader(file, delimiter=',')

    bigData = (list(reader))

    del bigData[0] #удаление заголовков
```

### Вычисления

Подготовка. Определим словарь, в котором будут находиться записи следующего характера:

`parameter: value(column)`, где `value` - список чисел определенного столбца.

```
parameters = {'carat': [float(row[1]) for row in bigData],
              'price': [float(row[7]) for row in bigData],
              'x': [float(row[-3]) for row in bigData],
              'y': [float(row[-2]) for row in bigData],
              'z': [float(row[-1]) for row in bigData]}
```

Определим функции вычислений, в которые будут передоваться параметр и значение:

```
def getStdev(parameter, data):  
    return f'Parameter: {parameter}, Результат: {std(data)}'  
  
def getMode(parameter, data):  
    return f'Parameter: {parameter}, Результат: {mode(data)}'  
  
def getMean(parameter, data):  
    return f'Parameter: {parameter}, Результат: {sum(data) / len(bigData)}'  
  
def getMedian(parameter, data):  
    return f'Parameter: {parameter}, Результат: {median(data)}'  
  
def getSum(parameter, data):  
    return f'Parameter: {parameter}, Результат: {sum(data)}'  
  
def getMaxMin(parameter, data):  
    return f'Parameter: {parameter}, Результат:\n    Max: {max(data)}\n    Min: {min(data)}'
```

Запустим вычисления:

```
for key in parameters:  
    print(getStdev(key, parameters[key]))  
    print(getMode(key, parameters[key]))  
    print(getMean(key, parameters[key]))  
    print(getMedian(key, parameters[key]))  
    print(getSum(key, parameters[key]))  
    print(getMaxMin(key, parameters[key]))
```

---

## Результаты

```
Parameter: carat, Результат: 0.47400685050996644  
Parameter: carat, Результат: 0.3  
Parameter: carat, Результат: 0.7979397478679852  
Parameter: carat, Результат: 0.7  
Parameter: carat, Результат: 43040.869999999912  
Parameter: carat, Результат:  
    Max: 5.01  
    Min: 0.2  
Parameter: price, Результат: 3989.4027576288736  
Parameter: price, Результат: 605.0  
Parameter: price, Результат: 3932.799721913237  
Parameter: price, Результат: 2401.0  
Parameter: price, Результат: 212135217.0
```

```
Parameter: price, Результат:  
  Max: 18823.0  
  Min: 326.0  
Parameter: x, Результат: 1.1217503485171316  
Parameter: x, Результат: 4.37  
Parameter: x, Результат: 5.731157211716609  
Parameter: x, Результат: 5.7  
Parameter: x, Результат: 309138.6199999939  
Parameter: x, Результат:  
  Max: 10.74  
  Min: 0.0  
Parameter: y, Результат: 1.1421240869900022  
Parameter: y, Результат: 4.34  
Parameter: y, Результат: 5.734525954764462  
Parameter: y, Результат: 5.71  
Parameter: y, Результат: 309320.32999999507  
Parameter: y, Результат:  
  Max: 58.9  
  Min: 0.0  
Parameter: z, Результат: 0.7056923054027403  
Parameter: z, Результат: 2.7  
Parameter: z, Результат: 3.5387337782723316  
Parameter: z, Результат: 3.53  
Parameter: z, Результат: 190879.30000000956  
Parameter: z, Результат:  
  Max: 31.8  
  Min: 0.0
```