Лекция 3

Язык программирования Python.

План занятия

- Коллекция
- Индексация
- tuple
- list
 - **Срезы**
 - Списковые включения
- Операции над списками

--

- Hash function
- Словарь Dict
 - definition
- Множество Set
- Генераторы

План занятия

• Общее представление

Инициализация аргументов функции

В случае инициализации аргументов функции мутабельными объектами, важно учитывать что. Инициализация будет произведена один раз а не каждый раз при вызове функции

Инициализация аргументов функции

```
[1]: def foo(l=[]):
           l.append(1)
   [2]: def foo(l=[]):
   ...: l.append(1)
  ...: print(l)
   [3]: foo()
[1]
   [4]: foo()
   [5]: foo()
[1, 1, 1]
  [6]: foo([1,2,3])
```

Массив

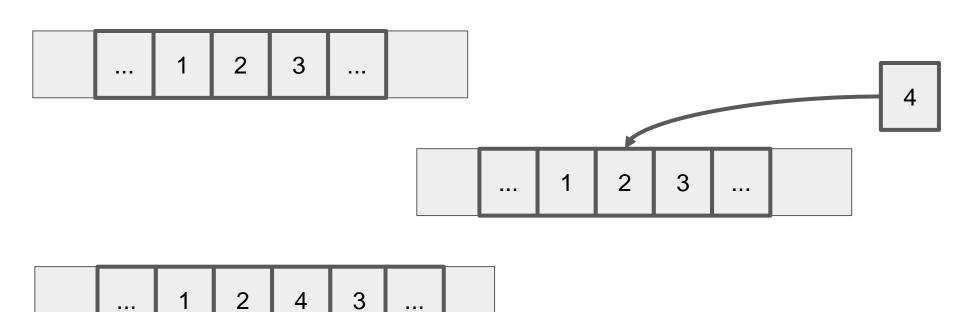
Фундаментальная структура данных, позволяющая хранить некоторый объем непересекающихся данных в непрерывном куске оперативной памяти. Обычно массивы представляют к каждому отдельной сущности массива через порядковый индексэто операция называется индексация. Поддерживается большинством языков программирования. Руthon не исключение.

Динамический массив

Обычно при создании массива предполагается, что массив будет иметь фиксированную длину, которая не меняется. Соответственно стандартный массив не поддерживает операции вставки и удаления элемента. Поэтому были введены специализированный тип данных как динамический массив.

Массив

RAM



Массив

RAM





Коллекции

- **Коллекция** программный объект, содержащий в себе, тем или иным образом, набор значений одного или различных типов, и позволяющий обращаться к этим значениям.
- Коллекция позволяет записывать в себя значения и извлекать их.
- **Назначение коллекции** служить хранилищем объектов и обеспечивать доступ к ним.

Индексация

Некоторые коллекции поддерживают операцию индексацию

В python операция обозначается [...] - аргументом операции является индекс, который указывает какой элемент коллекции нужно взять

Синтаксис:

collection_name[index]

где collection_name - переменная содержащая ссылку на коллекцию

index - указатель на элемент коллекции

Индексация

Для коллекций, сохраняющих порядок элементов, таких как list, tuple, range - индексом является порядковый номер элемента в коллекции.

ИНДЕКСАЦИЯ В РҮТНОМ ДЛЯ ТАКИХ КОЛЛЕКЦИЙ НАЧИНАЕТСЯ С 0! Также в python поддерживаются отрицательные индексы (см. примеры).

Для **dict**, **set**, ... - индексом является значение ключа элемента (об этом на др лекции).

Индексация

```
test.py X
  d = {'a':'data from dict'}
  3 print([0])
  4 print(d['a'])
      TERMINAL
[16:14:41] serg :: serg-pc → ~/tmp»
python3 test.py
data from list
data from dict
[16:14:52] serg :: serg-pc → ~/tmp»
```

Python - collection package

collections — Container datatypes

Source code: Lib/collections/__init__.py

This module implements specialized container datatypes providing alternatives to Python's general purpose builtin containers, dict, list, set, and tuple.

namedtuple()	factory function for creating tuple subclasses with named fields
deque	list-like container with fast appends and pops on either end
ChainMap	dict-like class for creating a single view of multiple mappings
Counter	dict subclass for counting hashable objects
OrderedDict	dict subclass that remembers the order entries were added
defaultdict	dict subclass that calls a factory function to supply missing values
UserDict	wrapper around dictionary objects for easier dict subclassing
UserList	wrapper around list objects for easier list subclassing
UserString	wrapper around string objects for easier string subclassing

Deprecated since version 3.3, will be removed in version 3.10: Moved Collections Abstract Base Classes to the collections.abc module. For backwards compatibility, they continue to be visible in this module through Python 3.9.

https://docs.python.org/3.8/library/collections.html

Python built-in collection

Sequence Types — list, tuple, range

There are three basic sequence types: lists, tuples, and range objects. Additional sequence types tailored for processing of binary data and text strings are described in dedicated sections.

tuple

```
test.py X
test.py > ...
      11 = 1,2,3
      12 = (1,2,3)
      print(l1)
      print(l2)
       TERMINAL DEBUG CONSOLE PROBLEMS
[15:10:10] serg :: serg-pc → ~/tmp»
 cd /home/serg/tmp ; /usr/bin/env /usr/bin/python3
g/tmp/test.py
(1, 2, 3)
(1, 2, 3)
[15:10:14] serg :: serg-pc → ~/tmp»
```

```
test.py x
test.py > ...
1 l1 = 1,2,3
D 2 l1[0] = 123

Exception has occurred: TypeError
'tuple' object does not support item assignment
File "/home/serg/tmp/test.py", line 2, in <module>
l1[0] = 123
```

list

```
test.py
🕏 test.py > ...
  1 l = [1,2,3,4]
      print(l)
       TERMINAL
[15:12:37] serg :: serg-pc → ~/tmp»
 cd /home/serg/tmp ; /usr/bin/env /usr/bin/python3
g/tmp/test.py
[1, 2, 3, 4]
[15:12:43] serg :: serg-pc → ~/tmp»
```

```
test.py X
🕏 test.py > ...
      l = [1,2,3,4]
  2 print(l)
   3 \quad l[2] = 3.14
       print(l)
       TERMINAL
[15:15:11] serg :: serg-pc → ~/tmp»
 cd /home/serg/tmp ; /usr/bin/env /usr/bin/python3
[1, 2, 3, 4]
[1, 2, 3.14, 4]
 [15:15:16] serg :: serg-pc → ~/tmp»
```

range - объект

Позволяет создавать генератор последовательности с заданными стартовым конечным (не включается) и шагом значениями. При этом последовательность не генерируется моментально.

range(begin, end, step)

```
In [16]: [1,2,3,4]
Out[16]: [1, 2, 3, 4]

In [17]: range(10)
Out[17]: range(0, 10)

In [18]: range(1, 10)
Out[18]: range(1, 10)

In [19]: range(1, 10, 2)
Out[19]: range(1, 10, 2)
```

```
In [20]: list(range(10))
Out[20]: [0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
In [21]: list(range(1, 10))
Out[21]: [1, 2, 3, 4, 5, 6, 7, 8, 9]
In [22]: list(range(1, 10, 2))
Out[22]: [1, 3, 5, 7, 9]
```

range - объект

```
In [23]: range(-100)
Out[23]: range(0, -100)
```

```
In [27]: list(range(10, 1, -1))
Out[27]: [10, 9, 8, 7, 6, 5, 4, 3, 2]
```

list

Operation	Result
s[i] = x	item i of s is replaced by x
s[i:j] = t	slice of s from i to j is replaced by the contents of the iterable t
del s[i:j]	same as s[i:j] = []
s[i:j:k] = t	the elements of s[i:j:k] are replaced by those of t (1)
del s[i:j:k]	removes the elements of s[i:j:k] from the list
s.append(x)	appends x to the end of the sequence (same as $s[len(s):len(s)] = [x]$)
s.clear()	removes all items from s (same as del s[:]) (5)
s.copy()	creates a shallow copy of s (same as s[:]) (5)
s.extend(t) or s += t	extends s with the contents of t (for the most part the same as s[len(s):len(s)] = t
s *= n	updates s with its contents repeated n times (6)
s.insert(i, x)	inserts x into s at the index given by i (same as $s[i:i] = [x]$)
s.pop([i])	retrieves the item at i and also removes it from s (2)
s.remove(x)	remove the first item from s where $s[i] == x (3)$
s.reverse()	reverses the items of s in place (4)

len

Встроенная функция <u>len</u> позволяет выяснить текущую длину коллекции тип коллекции не имеет значения.

```
In [12]: l = [1,2,3]
In [13]: m = {'key1':'value1', 'key2':'value2'}
In [14]: len(l)
Out[14]: 3
In [15]: len(m)
Out[15]: 2
```

append

Добавляет новый элемент в конец списка. Довольно эффективная процедура, так как при этом список **почти всегда** не реаллоцирует память.

insert

В случае если необходимо вставить элемент в середину или в начало списка, можно воспользоваться этой процедурой, но в этом случае список память будет реалоцированна, что может привести к потере производительности. Реалокация - процесс перераспределения памяти.

list. insert(i, x)

Insert an item at a given position. The first argument is the index of the element before which to insert, so a.insert(0, x) inserts at the front of the list, and a.insert(len(a), x) is equivalent to a.append(x).

index

Поиск элемента по значению в последовательности - вернет индекс элемента в списке. Если элемента нет то будет порождено исключение

list. **index**(x[, start[, end]])

Return zero-based index in the list of the first item whose value is equal to x. Raises a ValueError if there is no such item.

The optional arguments *start* and *end* are interpreted as in the slice notation and are used to limit the search to a particular subsequence of the list. The returned index is computed relative to the beginning of the full sequence rather than the *start* argument.

remove

Удаление элемента из списка по значению. Также в случае отсутствия такого элемента в списке будет порождено исключение

list. remove(x)

Remove the first item from the list whose value is equal to x. It raises a ValueError if there is no such item.

pop

Извлечь элемент из списка и вернуть его. В случае если мы извлекаем не последний элемент процесс также приведет к релокации памяти.

list. pop([i])

Remove the item at the given position in the list, and return it. If no index is specified, a.pop() removes and returns the last item in the list. (The square brackets around the *i* in the method signature denote that the parameter is optional, not that you should type square brackets at that position. You will see this notation frequently in the Python Library Reference.)

Списки арифметика (конкатенация)

Списки поддерживают операцию сложения

```
[n [1]: l = [1,2,3,4]
n [2]: l1 = [5,6,7,8]
  [3]: 12 = 1 + 11
ut[4]: [1, 2, 3, 4, 5, 6, 7, 8]
ut[6]: [1, 2, 3, 4, 5, 6, 7, 8]
```

Индексация списков

Для списков применима индексация с отрицательными значениями индекса при этом значению -1 будет соответствовать последний элемент списка, -2 предпоследний то есть

```
In [7]: l = [1,2,3]
In [8]: l[-1]
Out[8]: 3
In [9]: l[-1] == l[len(l) - 1]
Out[9]: True
```

Индексация списков

При обращение по индексу превосходящему длину списка будет сгенерировано исключение

```
In [10]: l[100]

IndexError Traceback (most recent call last)
<ipython-input-10-e2a0c2623844> in <module>
----> 1 l[100]

IndexError: list index out of range
```

Создание списков

Список можно создать несколькими способами

- 1. Явно
- 2. Списковые включения
- 3. Ключевое слово list

Явно

Через использование конструкции

[element1, element2, ..., elementN]

```
In [16]: [1,2,3,4]
Out[16]: [1, 2, 3, 4]
```

Более короткий синтаксис, который позволяет создать новый список на основе значений существующего списка.

[item for item in collection if condition]

Условный оператор может отсутствовать.

```
test.py
     # How to create list
     l = [] # empty list
      for i in range(10):
          l.append(i)
      print(l)
      l = [i for i in range(10)]
      print(l)
```

```
test.py
     # How to create list
     l = [] # empty list
      for i in range(10):
          l.append(i)
      print(l)
      l = [i for i in range(10)]
      print(l)
```

```
l = [[j for j in range(i)] for i in range(10)]
  3 for row in l:
          print(row)
      TERMINAL
[17:07:56] serg :: serg-pc → ~/tmp»
python3 test.py
```

```
test.py > ...
  2 l = [i for i in range(10) if i % 2 == 0]
     print(l)
      TERMINAL
[17:09:16] serg :: serg-pc → ~/tmp»
python3 test.py
[0, 2, 4, 6, 8]
[17:09:18] serg :: serg-pc → ~/tmp»
```

Ключевое слово list

Позволяет создавать списки в том числе десериализованные из другого типа последовательностей

```
In [34]: list('abcd')
Out[34]: ['a', 'b', 'c', 'd']
In [35]: list({'1':'2'})
Out[35]: ['1']
In [36]: list((1,2,3))
Out[36]: [1, 2, 3]
```

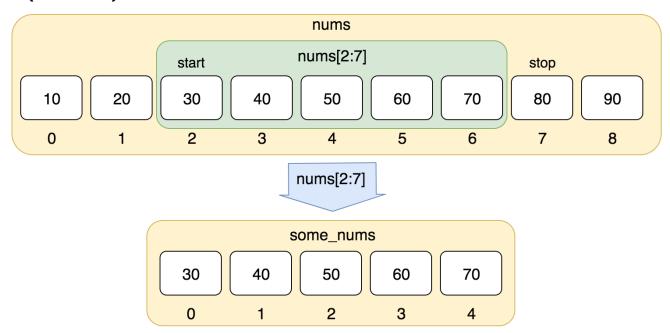
Срезы (slice)

Расширенная операция индексации, над сохраняющими порядок коллекциями, которая позволяет делать выборки из них. Параметры задаются значениями индексов правым (который включается в выборку) и левым (не включается в выборку). Синтаксис

collection_name[start:stop:step] -> вернет список в котором первым элементов будет элемент collection_name[start]

Значение индексов могут быть < 0. Значение step - шаг может отсутствовать по умолчанию он равен 1.

Срезы (slice)



Срезы (slice)

```
l = [i \text{ for } i \text{ in range}(10)]
        print(l)
        print(l[1:7])
        print(1[4:])
        print(l[:5])
OUTPUT TERMINAL DEBUG CONSOLE PROBLEMS
[16:43:49] serg :: serg-pc → ~/tmp»
python3 test.py
[0, 1, 2, 3, 4, 5, 6, 7, 8, 9]
[1, 2, 3, 4, 5, 6]
[4, 5, 6, 7, 8, 9]
[16:43:50] serg :: serg-pc → ~/tmp»
```

Tricks

Самый простой способ перевернуть список

```
In [7]: l
Gut[7]: [1, 2, 3, 4, 5, 6, 7, 8]
In [8]: l[::-1]
Gut[8]: [8, 7, 6, 5, 4, 3, 2, 1]
```

Взять каждый второй элемент

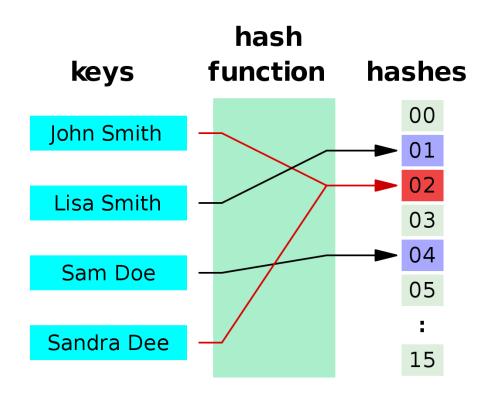
```
In [9]: l[::2]
Out[9]: [1, 3, 5, 7]
```

Hash function

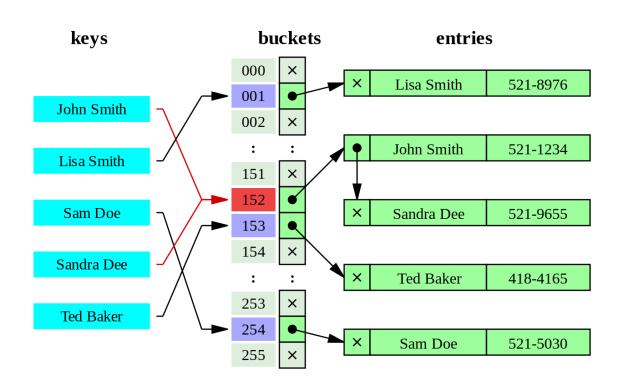
Неоднозначная функция принимающая на вход некоторые данные и возвращающая некоторое целое число ∈ [0, ..., N]. Хорошая хеш-функция должна удовлетворять двум свойствам:

- быстрое вычисление
- минимально количество коллизий

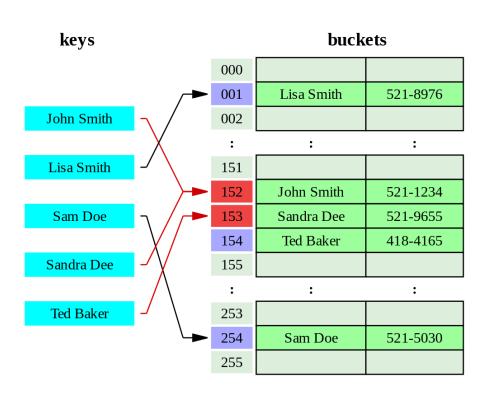
Hash function



Hash function (коллизия)



Hash function (коллизия)



Hash table

Еще называют хеш таблицей(дословный перевод), ассоциативный массив, словарь... Структура, которая не сохраняет порядок элементов при вставки, НО обеспечивает при этом быстрый доступ, быстрое удаление и вставку. В худшем и среднем случае работы этих функций производится за константное время в худшем за линейное.

Hash table

Python, как и почти все языки программирования дает возможность программисту создавать такие структуры. Ключами при этом могут быть только те типы данных, которые поддерживают операцию хеширования - значениями могут быть любые типы данных.

Hash table (python)

Словарь можно определить несколькими способами

- Явно
- Dict comprehension
- через цикл
- ключевое слово dict

Hash table (python)

```
dct.py > ...
       hash_table = {}
       hash_table["key1"] = "value"
       print(hash_table)
TERMINAL PROBLEMS 2 OUTPUT DEBUG CONSOLE
 rserg@MBP-serg ~/temp
$ python dct.py
{'key1': 'value'}
_serg@MBP-serg ~/temp
```

```
dct = {
              1: 1,
              2: 2 ** 2,
  10
  11
            3: 3 ** 2,
  12
        print(dct)
  13
TERMINAL
                               DEBUG CONSOLE
serg@MBP-serg ~/temp
$ python dct.py
{1: 1, 2: 4, 3: 9}
```

```
dct = {}
 22
      for i in range(1, 4):
 23
          key = i
 24 value = i ** 2
          dct[key] = value
 25
      print(dct)
 26
TERMINAL
rserg@MBP-serg ~/temp
-$ python dct.py
```

```
dest.pv
                           dct.py

    ■ Settings

ant.py
        dct: dict
dct.py
        dct = dict(
            key1="value1",
            key2="value2",
            key3="value3",
       print(dct)
TERMINAL
 rserg@MBP-serg ~/temp
 -$ python dct.py
{'key1': 'value1', 'key2': 'value2', 'key3': 'value3'}
```

```
>>> a = dict(one=1, two=2, three=3)
>>> b = {'one': 1, 'two': 2, 'three': 3}
>>> c = dict(zip(['one', 'two', 'three'], [1, 2, 3]))
>>> d = dict([('two', 2), ('one', 1), ('three', 3)])
>>> e = dict({'three': 3, 'one': 1, 'two': 2})
>>> f = dict({'one': 1, 'three': 3}, two=2)
>>> a == b == c == d == e == f
True
```

Dict - операции

dict.clear() - очищает словарь.

dict.copy() - возвращает копию словаря.

dict.get(key[, default]) - возвращает значение ключа, но если его нет, не бросает исключение, а возвращает default (по умолчанию None).

dict.items() - возвращает пары (ключ, значение).

dict.keys() - возвращает ключи в словаре.

dict.pop(key[, default]) - удаляет ключ и возвращает значение. Если ключа нет, возвращает default (по умолчанию бросает исключение).

dict.popitem() - удаляет и возвращает пару (ключ, значение). Если словарь пуст, бросает исключение KeyError. Помните, что словари неупорядочены.

Dict - операции

dict.setdefault(key[, default]) - возвращает значение ключа, но если его нет, не бросает исключение, а создает ключ со значением default (по умолчанию None).

dict.update([other]) - обновляет словарь, добавляя пары (ключ, значение) из other. Существующие ключи перезаписываются. Возвращает None (не новый словарь!).

dict.values() - возвращает значения в словаре.

Dict (проход)

```
dct = {i: i**2 for i in range(1,4)}
     print("----")
     for k in dct:
         print(k, dct[k])
     print("----")
     for k in dct.keys():
         print(k, dct[k])
     print("----")
     for v in dct.values():
         print(v)
     print("----")
     for k, v in dct.items():
         print(k, v)
TERMINAL PROBLEMS OUTPUT DEBUG CONSOLE
rserg@MBP-serg ~/temp
-$ python dct.py
3 9
3 9
2 4
```

Dict - ключи

- <u>immutable</u> objects
- <u>hashable</u>
- comparable \Rightarrow obj1 == obj2

Dict - ключи

Class	Description	Immutable?
bool	Boolean value	✓
int	integer (arbitrary magnitude)	✓
float	floating-point number	✓
list	mutable sequence of objects	
tuple	immutable sequence of objects	✓
str	character string	✓
set	unordered set of distinct objects	
frozenset	immutable form of set class	✓
dict	associative mapping (aka dictionary)	

Dict - ключи

```
In [2]: d = {[1,2,3]: 1}
TypeError
                                           Traceback (most recent call last)
<ipython-input-2-7e7847a84a44> in <module>()
---> 1 d = \{[1,2,3]: 1\}
TypeError: unhashable type: 'list'
In [3]: d = \{(1,2,3): 1\}
In [4]: d
  rt[4]: {(1, 2, 3): 1}
```

Set - множество

Структура данных очень похожая на словарь, но в отличие от словаря хранит только ключи. Ключи уникальны. Сам тип данных реализует теоретико-множественные операции.

Множество (Set)

```
In [13]: s1 = set((1,2,3))
In [14]: s2 = \{1, 2, 3\}
In [15]: s1
Out[15]: {1, 2, 3}
In [16]: s2
Out [16]: {1, 2, 3}
In [17]: s1 == s2
  t[17]: True
```

https://docs.python.org/3/tutorial/datastructures.html#sets

set операции

len(s) - вернет длину множества

х in s - проверить наличее элемента в множестве

x not in s - проверить отсутствие элемента в множестве

isdisjoint(other) - Верните **True**, если множество не имеет общих элементов с другими. Множества не пересекаются тогда и только тогда, когда их пересечение - пустое множество.

issubset(other) - наоборот

set операции

set <= other - Является ли каждый ли элемент в наборе состоит из другого.

set < other - Тоже самое но строго

set >= other

set > other

union(*others) ⇒ set | other | ... - Объединение множеств в одно результат множество

intersection(*others) \Rightarrow set & other & ... Пересечение всех множеств, то есть результат это множество из элементов общих для всех

difference(*others) ⇒ set - other - ...

set ^ other - Вернет новое множество элементов, которые содержаться только в одном из множеств, но не в обоих

сору() Верните неглубокую копию множетва.

Множество (Set) операции

```
In [31]: 1 in s1
Out [31]: True
In [32]: s2.add(4)
   [33]: s2
Out[33]: {1, 2, 3, 4}
   [34]: s1 ^ s2
Out [34]: {4}
  [35]: s1 | s2
Out [35]: {1, 2, 3, 4}
   [36]: s1 & s2
  t[36]: {1, 2, 3}
```

Определение

Generator functions allow you to declare a function that behaves like an iterator, i.e. it can be used in a for loop (https://wiki.python.org/moin/Generators).

При создании генератора нет необходимости выделять такое количество памяти, которое необходимо для хранения всего массива данных.

Создание

```
In [1]: gen = (i for i in range(10))
  [2]: type(gen)
 ıt[2]: generator
in [3]: for i in gen:
  ...: print(i)
```

Создание

```
class Arifmetic:
    def __init__(self, end, start=0, step=1):
        self.start = start
        self.n = end
        self.step = step
        self.cnt = 1
    def __next__(self):
        if self.cnt < self.n:</pre>
            self.cnt += 1
            self.start += self.step
            return self.start
        else:
            raise StopIteration()
    def __iter__(self):
        return self
a = Arifmetic(100, 2, 2)
print(sum(a))
```

Ключевое слово yield

```
def gen(start, end, step):
         while start < end:
             yield start
             start += step
     g = gen(1, 10, 1)
     for i in g:
         print(i)
      TERMINAL DEBUG CONSOLE
[13:10:56] serg :: serg-pc → ~/tmp»
python test.py
 13:10:57] serg :: serg-pc → ~/tmp»
```

Домашнее задание

- 1. Отсортируйте список случайной длины в зависимости от мода по возрастанию и по убыванию
- 2. Найдите сумму всех чисел меньше 1000, кратных 3 или 5 с помощью функции генератора
- 3. Запишите в словарь по ключам от 1 до 10, список чисел (подается на входе), которые делятся на соответствующие ключи