[**TODO TABLE** 2](#_Toc100684048)

[**Python** 3](#_Toc100684049)

[**Иммутабельность типов данных** 3](#_Toc100684050)

[**Списковое включение/представление списков (*list comprehension*)** 3](#_Toc100684051)

[**Контекстный менеджер** 3](#_Toc100684052)

[**Отличия кортежа от списка** 3](#_Toc100684053)

[**Функция *map*** 3](#_Toc100684054)

[**Функция *filter*** 4](#_Toc100684055)

[**Различия методов *type()* и *isinstance()*** 4](#_Toc100684056)

[**Протокол итерации (итерируемый объект)** 5](#_Toc100684057)

[**Итераторы** 5](#_Toc100684058)

[**Генераторы** 6](#_Toc100684059)

[**Отличия генераторов от итераторов** 7](#_Toc100684060)

[**\*args и \*\*kwargs** 8](#_Toc100684061)

[**Виртуальное окружение** 8](#_Toc100684062)

[**Слоты** 9](#_Toc100684063)

[**GIL** 9](#_Toc100684064)

[**is, id, ==** 9](#_Toc100684065)

[**Создание своего контекстного менеджера** 10](#_Toc100684066)

[**Индексатор класса** 10](#_Toc100684067)

[**Поверхностное и глубокое виды копирования объектов** 11](#_Toc100684068)

[**Пакеты и модули** 11](#_Toc100684069)

[**Чистая функция** 12](#_Toc100684070)

[**Именованные кортежи (*namedtuple*)** 12](#_Toc100684071)

[**Базы данных** 13](#_Toc100684072)

[**Схемы (зачем нужны)** 13](#_Toc100684073)

[**Виды соединений таблиц** 13](#_Toc100684074)

[**Индексы** 13](#_Toc100684075)

[***ACID* (свойства транзакционной СУБД)** 15](#_Toc100684076)

[**Уровни изоляции транзакций** 15](#_Toc100684077)

[**REST** 17](#_Toc100684078)

[**Идемпотентные методы** 17](#_Toc100684079)

**TODO TABLE:**

|  |  |  |  |
| --- | --- | --- | --- |
| Docker | Poetry | Django+debug toolbar | Celery |
| Redis | Магические методы | GIL | Asyncio |
| SOLID | Тестирование (unittest, mock-объекты) | Метаклассы (поверхностно описать) | REST/restful |
|  | Расписать важные вопросы на собесах по видосам на ютубе |  |  |

**Python**

## **Иммутабельность типов данных**

Тип переменной *a* **иммутабелен**, если после перезаписи значения переменной изменился ее идентификатор (*id(a)*). Изменение идентификатора указывает на то, что после перезаписи значения создался новый объект, следовательно, старый объект нельзя было изменить.

## **Списковое включение/представление списков (*list comprehension*)**

***List comprehension*** – короткий способ генерирования списков.

**Пример**

squares = [i \* i for i in range(10) if i % 2 == 0]

Представление списков состоит из трех элементов:

1. *expression* – выражение, возвращающее значение. В приведенном выше примере выражение *i \* i* выражением.
2. *member* – является объектом или значением в списке или итерируемым объектом (*iterable*). В приведенном выше примере значением элемента является *i*.
3. *iterable* – объект, возвращающий элементы. В примере выше этим объектом является *range*.

## **Контекстный менеджер**

<https://devpractice.ru/python-lesson-21-context-manager/>

## **Отличия кортежа от списка**

Элементы списка можно редактировать, а кортежа – нет.

## **Функция *map***

**Функция** *map* применяет функцию к каждому элементу последовательности и **возвращает итератор** с результатами.

**Пример**

list\_of\_str = ['1', '2', '5', '10']

list = list(map(int, list\_of\_str))

print(list)

> [1, 2, 5, 10]

Вместо *map* можно использовать *list comprehension*, но следует учитывать, что *map* **возвращает итератор**, а *list comprehension* – **список**. Если нужно сгенерировать большое количество элементов, то использование *map* выглядит уместнее в целях экономии памяти.

## **Функция *filter***

Функция *filter* применяет функцию ко всем элементам последовательности и **возвращает итератор** с теми объектами, для которых функция вернула *True*. Например, вернуть только те строки, в которых находятся числа:

list\_of\_strings = ['one', 'two', 'list', '', 'dict', '100', '1', '50']

list(filter(str.isdigit, list\_of\_strings))

> ['100', '1', '50']

Из списка чисел оставить только нечетные:

list(filter(lambda x: x % 2 == 1, [10, 111, 102, 213, 314, 515]))

> [111, 213, 515]

## **Различия методов *type()* и *isinstance()***

В отличие от *type*, функция *isinstance* **возвращает** не тип данных аргумента, а **булево значение**, говорящее о том, принадлежит объект к определенному классу или нет:

num = 4.44

print(isinstance(num, float))

> True

Также отличие от *type* состоит в том, что *isinstance* **"знает" о наследовании**. Функция воспринимает объект производного класса, как объект базового. Поэтому нельзя использовать *type* для проверок наследующихся типов.

**Пример**

class BaseExample:

pass

class DerivedExample(BaseExample):

pass

test = DerivedExample()

print(isinstance(test, BaseExample))

> True

## **Протокол итерации (итерируемый объект)**

**Итерируемый объект (*iterable*)** – это объект, возвращающий элементы по одному, а именно объект, **реализующий** **один из** методов: *\_\_iter\_\_* **или** *\_\_getitem\_\_* и метод *\_\_next\_\_.*

## **Итераторы**

**Итератор** – это объект, который может возвращать элементы последовательности по одному.

Технически же это любой объект, **реализующий метод** *\_\_next\_\_*, который должен вернуть **следующий элемент** или исключение *StopIteration,* если перечислены все элементы.

Также итератор **реализует метод** *\_\_iter\_\_*, **возвращающий итератор**. Если этого метода нет, функция *iter()* проверяет, нет ли метода *\_\_getitem\_\_* – метода, который позволяет получать элементы по индексу. Если метод *\_\_getitem\_\_* есть, **возвращается итератор**, который проходится по элементам, используя индекс (начиная с 0).  Если не реализован ни один из этих методов, тогда будет **вызвано исключение** *TypeError*.

**Пример**

numbers = [1, 2]

i = iter(numbers)

next(i)

> 1

next(i)

> 2

next(i)

> StopIteration Traceback (most recent call last) in () ----> 1 next(i) StopIteration:

Для того, чтобы итератор снова начал возвращать элементы, его надо **заново создать**.

**Итераторы полезны** тем, что они **отдают элементы по одному**. Например, при работе с файлом или любой последовательностью это полезно тем, что **в памяти будет находиться** не вся последовательность, а **только одно текущее значение** итератора.

## **Генераторы**

В языках программирования есть такие понятия, как ленивые/отложенные вычисления (*lazy evaluation*). Генераторы можно считать реализацией механизма отложенного вычисления.

Генератор генерирует значения, возвращающиеся по запросу, и после возврата одного значения выполнение функции-генератора приостанавливается до запроса следующего значения. Между запросами **генератор сохраняет свое состояние**.

Python позволяет **создавать** генераторы **двумя способами**:

1. **генераторное выражение**;
2. **функция-генератор**.

**Генераторное выражение** использует такой же синтаксис, как *list comprehension* (но используются скобки круглые), но **возвращает итератор**, а не список:

genexpr = (x\*\*2 for x in range(1,5))

next(genexpr)

> 1

next(genexpr)

> 4

next(genexpr)

> 9

**Генераторные функции** – это функции, где есть хотя бы одно выражение *yield*. При запуске генератора, функция выполняется до первого выражения *yield*. Генератор при этом встанет «на паузу» до следующей итерации. При следующей итерации выполнение генератора продолжится до очередного *yield*:

def f\_gen(m):

s = 1

for n in range(1,m):

yield n\*\*2 + s

s += 1

a = f\_gen(5)

for i in a:

print(i)

> 2

> 6

> 12

> 20

## **Отличия генераторов от итераторов**

**Итератор** – более общая концепция, чем генератор. Можно сказать, что генератор является итератором, но не наоборот.

Главной целью **итератора** является предоставление интерфейса для поочередного доступа к элементам коллекций и потоков данных.

Главная задача **генератора** – хранить закономерность, по которой генерируется последовательность.

## **\*args и \*\*kwargs**

***\*args*** – кортеж параметров (является типом *tuple*), позволяющий передавать в функцию неопределенное количество аргументов:

def getsum(\*numbers):

sum = 0

for num in numbers:

sum += num

return sum

print(add(2, 3)) # 5

print(add(2, 3, 5)) # 10

***\*\*kwargs*** – то же самое, что и \*args, но обязывает проставлять имена параметров и является типом *dictionary*. Позволяет передавать параметры в произвольном порядке:

def total\_fruits(\*\*fruits):

total = 0

for amount in fruits.values():

total += amount

return total

print(total\_fruits(banana=5, mango=7, apple=8))

print(total\_fruits(banana=5, mango=7, apple=8, oranges=10))

## **Виртуальное окружение**

Виртуальное окружение помогает изолировать зависимости проекта в отдельном пространстве (вместо глобального окружения ОС), привязанном к конкретному проекту. Это позволяет не конфликтовать зависимостям разных версий при наличии множества больших проектов на одной машине.

Также виртуальное окружение можно использовать для хранения приватных данных, необходимых для работы проекта (*API*-ключи, различные переменные, пароли и т.д.). Для этого можно создавать **переменные окружения** и получать их значения в конфигах приложения.

Поддержка виртуальных окружений есть в Python по умолчанию (начиная с версии 3.3) в виде *venv*.

**Примеры** использования на видео:

<https://www.youtube.com/watch?v=Y9MRCxq4DIc&ab_channel=%D0%94%D0%B8%D0%B4%D0%B6%D0%B8%D1%82%D0%B0%D0%BB%D0%B8%D0%B7%D0%B8%D1%80%D1%83%D0%B9%21>

## **Слоты**

1. <https://ru.stackoverflow.com/questions/1206730/%D0%9A%D0%B0%D0%BA%D0%BE%D0%B2%D0%B0-%D1%86%D0%B5%D0%BB%D1%8C-slots-%D0%B2-python>
2. <https://otus.ru/nest/post/664/>

## **GIL**

<https://tproger.ru/translations/global-interpreter-lock-guide/>

## **is, id, ==**

Ниже приведен наглядный **пример**

# определим три одинаковых списка

list\_1 = ['a', 'b', 'c']

list\_2 = list\_1

list\_3 = list(list\_1)

print(list\_1) # >> ['a', 'b', 'c']

print(list\_2) # >> ['a', 'b', 'c']

print(list\_3) # >> ['a', 'b', 'c']

# list\_1 равен list\_2 и по значению и как объект, т.к. оба указывают на один и тот же адрес в памяти

# (это можно проверить оператором is или сравнить id (эквивалентные способы))

print(list\_1 == list\_2, id(list\_1) == id(list\_2), list\_1 is list\_2)

>> True True True

# list\_2 равен list\_3 только ПО ЗНАЧЕНИЮ (==). А как объекты они разные

print(list\_2 == list\_3, id(list\_2) == id(list\_3), list\_2 is list\_3)

>> True False False

# добавим элемент в list\_1 и посмотрим, что станет с list\_2

list\_1.append('n')

print(list\_1)

>> ['a', 'b', 'c', 'n']

# list\_2 также изменится, потому что является ссылкой на list\_1

print(list\_2)

>> ['a', 'b', 'c', 'n']

# а list\_3 не изменится

print(list\_3)

>> ['a', 'b', 'c']

## **Создание своего контекстного менеджера**

Нужно реализовать протокол менеджера контекста одним из способов:

1. реализация методов \_\_*enter*\_\_ и \_\_*exit*\_\_.
2. использование модуля *contextlib* (через декоратор *@contextmanager*).

Пример есть в учебном проекте.

## **Индексатор класса**

Нужно реализовать метод *\_\_getitem\_\_* (см. пример в учебном проекте).

## **Поверхностное и глубокое виды копирования объектов**

*copy.copy(x)* – возвращает поверхностную копию *x*.

*copy.deepcopy(x)* – возвращает глубокую копию *x*.

Разница между **поверхностным** и **глубоким** копированием существенна только для **составных** объектов, содержащих изменяемые объекты (например, список списков, или словарь, в качестве значений которого – списки или словари).

**Пример**

import copy

test\_1 = [0, [1, 2]]

test\_copy = copy.copy(test\_1) # поверхностная копия

print(test\_1, test\_copy)

>>> [0, [1, 2]] [0, [1, 2]]

test\_copy[1].append(3)

print(test\_1, test\_copy)

>>> [0, [1, 2, 3]] [0, [1, 2, 3]]

# тройка добавилась и туда и туда, т.к. поверхностное копирование

не проводит копирование вложенных объектов

test\_1 = [0, [1, 2]]

test\_deepcopy = copy.deepcopy(test\_1) # глубокая копия

test\_deepcopy[1].append(3)

print(test\_1, test\_deepcopy)

>>> [0, [1, 2]] [0, [1, 2, 3]]

## **Пакеты и модули**

**Модуль** представляет собой **один файл** *Python*, который можно импортировать.

**Пакет** состоит из нескольких файлов/модулей *Python*. По сути это директория, которая может содержать поддиректории и модули. Особенностью является также то, что **пакет** **должен** **содержать пустой файл** с названием *\_\_init\_\_.py.*

## **Чистая функция**

Чистая функция – функция, удовлетворяющая следующим критериям:

1. **всегда возвращает одно** и то же **значение** **для одних и тех же** значений **входных** **параметров**;
2. **не имеет побочных эффектов** (не влияет на код, содержащийся вне функции. Например, не меняет значения глобальных переменных, не использует потоки ввода-вывода).

В общих чертах это функция, возвращаемое значение которой зависит исключительно от входных данных.

**Преимущества**:

* читаемость;
* портативность;
* легкая тестируемость.

## **Именованные кортежи (*namedtuple*)**

Именованный кортеж – это то же самое, что и обычный кортеж, но с той разницей, что теперь к элементу кортежа можно будет обращаться по имени, что улучшает читабельность кода (фактически это будет аналогом перечисления *Enum* в *C#*).

**Пример**:

from collections import namedtuple

# 1 параметр – имя типа, далее – список имен элементов кортежа

States = namedtuple('States', 'In\_Progress Open')

states = States(1, 4)

print(states)

>> States(In\_Progress=1, Open=4)

**Базы данных**

## **Схемы (зачем нужны)**

Есть несколько возможных объяснений, для чего стоит применять схемы:

* чтобы объединить объекты БД в логически связанные группы для облегчения управления ими.
* чтобы в одной базе сосуществовали разные приложения, и при этом не возникало конфликтов имён.

## **Виды соединений таблиц**

1. *INNER JOIN* – **внутреннее** соединение. Объединяет из соединяемых таблиц только те строки, для которых выполняется условие *ON*.
2. *LEFT JOIN* – **левостороннее** соединение. Строки из **левой** таблицы в любом случае попадут в результирующую выборку.
3. *RIGHT JOIN* – **правостороннее** соединение. Строки из **правой** таблицы в любом случае попадут в результирующую выборку.
4. *CROSS JOIN* – **перекрестное** соединение. Каждая запись левой таблицы объединяется с каждой записью правой таблицы (декартово произведение).

## **Индексы**

Индекс БД – объект, создаваемый с целью повышения производительности поиска данных. Индекс хранит ссылки на страницы, в которых лежат записи.

Индексы могут наоборот уменьшить производительность, если применяются в таблицах, где часто происходят операции *INSERT*, *DELETE* и *UPDATE*, т.к. нужно актуализировать состояние самих индексов.

Также нужно помнить о том, что **индексы занимают дополнительный объем памяти**, т.к. хранятся на диске.

|  |  |
| --- | --- |
| **Индекс** | **Кейсы для применения** |
| ***btree*** – **сбалансированное дерево**. Является индексом по умолчанию*.* Наиболее часто применяемый вид индекса | Любые данные, с которыми можно проводить операции сравнения |
| ***hash*-индексы** – хеширование значений индексированного поля таблицы | Работает только с простыми условиями равенства  Используется *PostgreSQL* для соединения таблиц, если у таблиц нет индексов. СУБД временно создает хеш-индексы для таблиц, соединяет их и затем удаляет хеш-индексы. |
| ***GiST (R-Tree)*** | Позволяет индексировать значения, перекрывающиеся между собой. Это могут быть значения многомерных структур хранения данных типа координат в многомерном пространстве  Индексация географических типов данных |
| ***GIN* (инвертированный индекс)** | Полнотекстовый поиск, поиск по *JSON*-полям |
| **Частичный индекс** | Применяется, когда нет цели индексировать данные целиком.  **Пример**: *CREATE INDEX idx\_table\_field ON table (field) WHERE field IS NOT NULL*  **Недостаток**: чтобы индекс применялся, необходимо в запросы «таскать» выражение, указанное при создании индекса (в примере это выражение *IS NOT NULL*). |
| **Функциональный индекс** | Применяется, когда есть необходимость проиндексировать не сами значения, а результаты функций от этих значений.  **Пример**: *CREATE INDEX idx\_table\_field ON table (LOWER(field))*  **Недостаток**: чтобы индекс применялся, необходимо в запросы «таскать» выражение, указанное при создании индекса (в примере это выражение *LOWER(field)*). |
| **Кластерный индекс** | Кластеризация сортирует данные по указанному индексу. То есть она применяется к таблице и ранее созданному индексу  Кластеризация является одноразовой операцией: последующие изменения в таблице нарушают порядок кластеризации |

## ***ACID* (свойства транзакционной СУБД)**

***Atomicity*** – **атомарность** операций. Гарантирует, что никакая [транзакция](https://ru.wikipedia.org/wiki/%D0%A2%D1%80%D0%B0%D0%BD%D0%B7%D0%B0%D0%BA%D1%86%D0%B8%D1%8F_(%D0%B8%D0%BD%D1%84%D0%BE%D1%80%D0%BC%D0%B0%D1%82%D0%B8%D0%BA%D0%B0)) не будет зафиксирована в системе частично. Будут выполнены все её операции, либо не выполнено ни одной.

***Consisitency*** – **консистентность, согласованность**. Гарантия целостности данных, выполнения определенных правил и ограничений.

***Isolation*** – **изолированность**. Выполнение параллельных транзакций не должно влиять на итоговый результат другой транзакции. На самом деле это очень дорогое требование, поэтому оно частично и по-разному выполняется в соответствии с уровнями изоляции транзакций.

***Durability*** – **долговечность**. Это свойство гарантирует устойчивость к ошибками. Изменения успешно завершенной транзакции не должны пропасть в случае сбоев.

## **Уровни изолированности транзакций**

Стандарт *SQL* определяет **4 уровня изоляции** транзакций.

Пусть *T1* – **первая транзакция**, а *T2* – **вторая** **транзакция**.

***L1. READ UNCOMMITTED*** («грязное» чтение) – *T1* **видит изменения**, выполненные **незавершенной** *T2*. Этот уровень не реализован в *PostgreSQL*, т.к. не представляет практической ценности.

***L2****\*****. READ COMMITTED*** (чтение зафиксированных данных) – *T1* увидит изменения, выполненные *T2*, после коммита (фиксации) *T2*.

Попытка изменения одних и тех же данных в *T2* до коммита изменения в *T1* будет заблокирована.

\***Является уровнем по умолчанию в *PostgreSQL***.

***L3. REPEATABLE READ*** (повторяемое чтение) – *T1* увидит изменения, выполненные *T2*, только после коммита *T2* и *T1*.

Попытка изменения одних и тех же данных в *T2* до коммита изменения в *T1* будет заблокирована.

***L4. SERIALIZABLE*** (сериализуемость) – уровень, на котором моделируется последовательное выполнение всех зафиксированных транзакций. Если транзакции обращаются к одной и той же таблице, зафиксируются изменения только той транзакции, которая зафиксируется быстрее всех.

Шаблон для наглядного представления и работы с уровнями изоляции (нужна заранее созданная БД с таблицей *test* и еще одним пользователем помимо *postgres*):

*T1* (сессия пользователя *postgres*):

-- READ COMMITTED

-- REPEATABLE READ

-- SERIALIZABLE

DELETE FROM test;

INSERT INTO test VALUES (1), (2);

SELECT \* FROM test

BEGIN TRANSACTION ISOLATION LEVEL REPEATABLE READ;

UPDATE test SET id = id + 10 WHERE id = 1;

COMMIT;

*T2* (сессия пользователя *test*):

SELECT \* FROM test

BEGIN TRANSACTION ISOLATION LEVEL REPEATABLE READ;

--- можно поменять id на 1, чтобы увидеть блокировки

UPDATE test SET id = id + 10 WHERE id = 2;

COMMIT;

**REST**

## **Идемпотентные методы**

Метод HTTP является **идемпотентным**, если один и тот же неоднократно выполненный запрос оказывает один и тот же эффект, не изменяющий состояние сервера. Другими словами, идемпотентный метод не должен иметь никаких побочных эффектов (*side-effects*), кроме сбора статистики или подобных операций.

Методы [***GET***](https://developer.mozilla.org/ru/docs/Web/HTTP/Methods/GET)***,***[***HEAD***](https://developer.mozilla.org/ru/docs/Web/HTTP/Methods/HEAD)***,***[***PUT***](https://developer.mozilla.org/ru/docs/Web/HTTP/Methods/PUT), [***DELETE***](https://developer.mozilla.org/ru/docs/Web/HTTP/Methods/DELETE) и ***OPTIONS*** **идемпотентны** в отличие от методов [***POST***](https://developer.mozilla.org/ru/docs/Web/HTTP/Methods/POST)и ***PATCH***.

Для идемпотентности нужно рассматривать только изменение фактического внутреннего состояния сервера, а возвращаемые запросами коды статуса могут отличаться: первый вызов [***DELETE***](https://developer.mozilla.org/ru/docs/Web/HTTP/Methods/DELETE) вернёт код [**200**](https://developer.mozilla.org/ru/docs/Web/HTTP/Status/200), в то время как последующие вызовы вернут код [404](https://developer.mozilla.org/ru/docs/Web/HTTP/Status/404).