Базы Данных: легче, чем кажется





Агенда

- Постановка задачи: БД для ОГО
- Свойства БД: порядок прежде всего
- Нарушение свойств: антипримеры
- Начинаем проектирование: вспоминаем грамматику
- Модели данных: выбор очевиден
- Реляционные СУБД: выбор не так уж и очевиден
- Нормализация отношений: долго, скучно, но необходимо
- Метод ER-диаграмм: нормальные отношения сразу
- Основы SQL: конфигурирование и CRUD
- Работа с БД в Python через DB-API
- Решение задачи: бета-версия и первые результаты
- SQL-инъекции и защита от них



Постановка задачи: БД для ОГО

Список утверждений об - «ОГО» - одной грандиозной организации:

- организация состоит из сотрудников, у которых есть ФИО, должность и уникальный табельный номер;
- сотрудники работают в различных проектах: один сотрудник может участвовать в нескольких проектах; в проекте, разумеется, может быть несколько сотрудников; бывают проекты без сотрудников (которые еще только планируются) и сотрудники без проектов (проект закрылся, но его сотрудников пока не сократили);
- у каждого проекта есть название и уникальный идентификатор;
- в каждом проекте есть не более одного менеджера;
- все сотрудники получают зарплату (как ни странно), которая складывается из должностного оклада и премии.



Постановка задачи

Задачи, которые должна решать База Данных и программа-клиент для работы с ней в 0-й версии:

- хотя бы запускаться и не падать в течение первого часа работы;
- хранить данных о сотрудниках, проектах и связях между ними;
- предоставлять каждому сотруднику возможность посмотреть свою и только свою зарплату;
- предоставлять каждому менеджеру возможность смотреть и редактировать премии всех сотрудников его проекта (но свои, разумеется, редактировать нельзя), а также исключать сотрудников из этого проекта (опять же, кроме себя).



Свойства баз данных: порядок прежде всего

База данных — это систематизированный набор данных, отображающий атрибуты и взаимосвязь объектов предметной области и предназначенный для удовлетворения информационных потребностей пользователей.





VS



Свойства баз данных

- Целостность (полнота, непротиворечивость, адекватность)
- Неизбыточность
- Безопасность

А что будет, если их нарушить?





Антипримеры (не повторять, опасно для кармы!)

ФИО	Должность	Оклад
Иванов И.И.	инженер	50000
Петров П.П.	старший инженер	51000
Сидоров С.С.	менеджер проекта	100000
Иванов И.И.	инженер	50000

Нарушение целостности (полноты) информация неполная, т.к. записи не могут быть однозначно идентифицированы

ФИО	Должность	Оклад
Иванов И.И.	инженер	50000
Петров П.П.	старший инженер	51000
Сидоров С.С.	менеджер проекта	100000
Иванов И.И.	инженер	60000

Нарушение неизбыточности информация избыточная, т.к. оклад должен однозначно определяться должностью

ФИО	Должность	Оклад
Иванов И.И.	инженер	150000
Петров П.П.	старший инженер	51000
Сидоров С.С.	менеджер проекта	10000
Иванов И.И.	инженер	50000

Нарушение безопасности отсутствуют уровни доступа к информации, защита от изменения информации произвольным пользователем



Начинаем проектирование: вспоминаем грамматику

Формулируем утверждения о предметной области, строго выделяя подлежащие и дополнения (это будут сущности), и сказуемые (это будут связи)

- сотрудник относится к проекту
- **сотрудник** <u>имеет</u> **оклад** и **премию** (из которых складывается зарплата, если кто не знает)
- сотрудник имеет должность (инженер, старший инженер, менеджер)
- оклад сотрудника однозначно определяется должностью



Модели данных: иерархическая — простая, как дерево

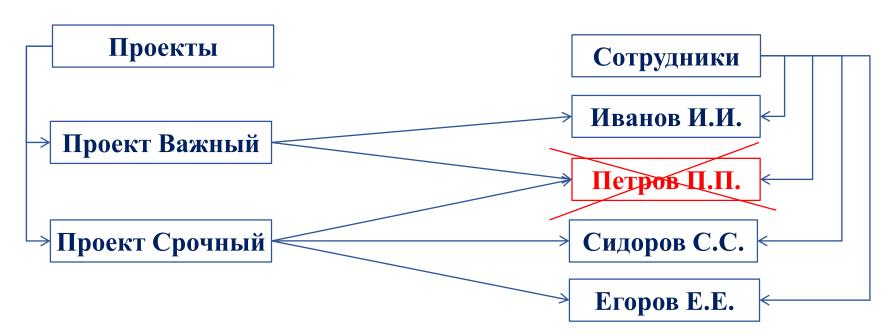
- Иерархическая (структура дерево)
- + удобная, если сущностей мало и связи простые
- отсутствие гибкости (как отразить возможность нахождения сотрудника сразу в нескольких проектах?)





Модели данных: сетевая — легко запутаться

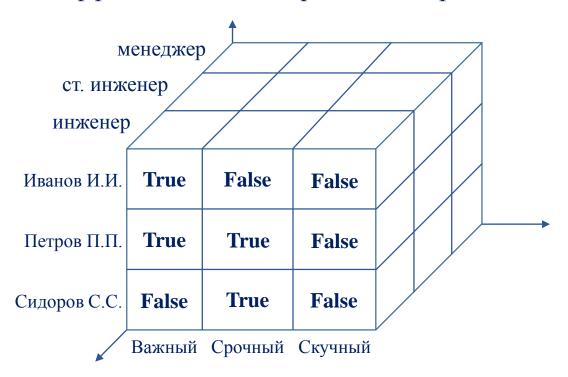
- Сетевая (структура граф)
- + более гибкая, чем иерархическая
- сложно контролировать целостность и неизбыточность





Модели данных: многомерная — для спецзадач

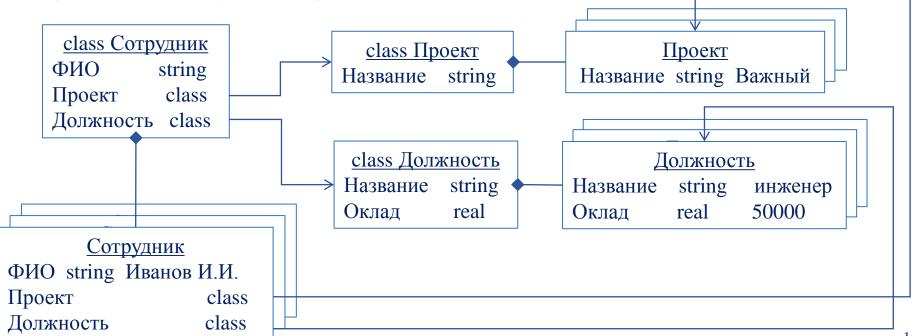
- Многомерная (структура гиперкуб)
- + удобная для аналитической обработки больших объемов данных (особенно, привязанных ко времени)
- громоздкая и неэффективная для оперативной обработки информации





Модели данных: объектно-ориентированная – для ORM

- Объектно-ориентированная (структура тоже дерево)
- + удобная для работы с отдельными объектами, полноценно представляющими соответствующие сущности со всеми их связями
- сложность алгоритмов и низкая скорость выполнения запросов для обработки совокупности разнотипных объектов





Модели данных: реляционная – то, что надо!

- Реляционная (структура таблицы)
- + удобная для понимания, физической реализации и оперативной обработки данных
- в общем случае необходимо анализировать совокупность таблиц даже если модифицируются атрибуты отдельной сущности

Сотрудники		
ID	ФИО	Должность
1	Иванов И.И.	инженер
2	Петров П.П.	старший инженер
3	Сидоров С.С.	менеджер проекта
4	Егоров Е.Е.	инженер

Проекты		
ID Название		
1	Важный	
2	Срочный	
3 Скучный		

Сотрудник-Проект		
ID сотрудника	ID проекта	
1	1	
2	1	
2	2	
3	2	
4	2	

Должность-Оклад		
Название	Оклад	
инженер	50000	
старший инженер	51000	
менеджер проекта	100000	



Модели данных: постреляционная — лучшее — враг хорошего

- Постреляционная (структура таблицы с возможностью вложенности)
- + можно заменить совокупность связанных реляционных таблиц одной постреляционной
- сложно контролировать целостность и неизбыточность

Проекты		
ID	Название	ID сотрудника
1	Важный	1
		1
2	Срочный	2
		2
		2
3	Скучный	



Выбор реляционной СУБД: прагматичность превыше всего





Нормализация отношений: долго, скучно, но важно

5НФ



3НФ

2НФ

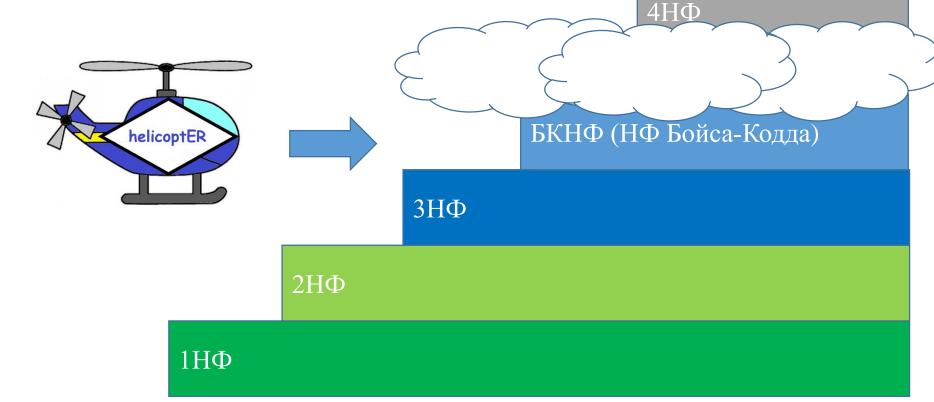
1НФ







5НФ



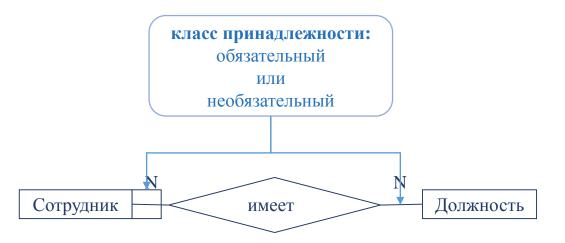


Метод ER-диаграмм











Метод ER-диаграмм





 Сотрудник
 Премия

 Иванов И.И.
 30000

 Петров П.П.
 50000

 Сидоров С.С.
 30000

 Егоров Е.Е.
 20000

 Новый Н.Н.
 20000









Метод ER-диаграмм: 6 правил

- [1]:[1] = одно отношение
- ➤ 1:[1] = два отношения
- ➤ 1:[N] = два отношения
- 1:1 = три отношения
- 1:N = три отношения
- ightharpoonup N:N = три отношения



Метод ER-диаграмм: результат проектирования

Сотрудники					
Таб. номер	ФИО	Должность	Премия	Логин	Пароль
1	Иванов И.И.	инженер	30000	ivanovi	ivanov123
2	Петров П.П.	старший инженер	50000	petrovp	p1e2t3
3	Сидоров С.С.	менеджер проекта	30000	sidorovs	zayka88
4	Егоров Е.Е.	инженер	20000	egorove	qwerty
5	Новый Н.Н.	инженер	20000	novyin	a1111

Проекты		
ID Название		
1	Важный	
2	Срочный	
3	Скучный	

Должность-Оклад		
Название Оклад		
инженер	50000	
старший инженер	51000	
ведущий инженер	70000	
менеджер проекта	100000	

Сотрудник-Проект		
Таб. номер	ID проекта	
1	1	
2	1	
2	2	
3	2	
4	2	



Первичный ключ уникальный идентификатор кортежа

Сотрудники							
Таб. номер		ФИО	Должность	Премия	Логин	Пароль	
1	1 Иванов И.И. инженер 2 Петров П.П. старший инженер 3 Сидоров С.С. менеджер проекта 4 Е оров Е.Е. инженер 5 Новый Н.Н. инженер		30000	ivanovi	ivanov123		
2			старший инженер	50000	petrovp	p1e2t3	
3			менеджер проекта	30000	sidorovs	zayka88	
4			инженер	20000	egorove	qwerty	
5			инженер	20000	novyin	a1111	

Проекты							
-ID	Название						
1	Зажный						
2	Срочный						
3	Скучный						

	Должность-Оклад				
	Название	Оклад			
	инженер	50000			
	старший инженер	51000			
	ведущий инженер	70000			
	менеджер проекта	100000			

Сотрудник-Проект					
Tah Homen ID BROOKTA					
1	1				
2	1				
2	2				
3	2				
4	2				



Внешний ключ: служит для связи двух отношений

Сотрудники								
Таб. номер	ФИО	Должность	Премия Логин		Пароль			
1	Иванов И.И.	инженер	30000	ivanovi	ivanov123			
2 Петров П.П.		старший инженер	50000	petrovp	p1e2t3			
3	Сидоров С.С.	менеджер проекта	30000	sidorovs	zayka88			
4 Егоров Е.Е.		инженер	20000	egorove	qwerty			
5	Новый Н.Н.	инженер	20000	novyin	a1111			

Проекты						
ID Название						
1	Важный					
2	Срочный					
3 Скучный						

Должность-Оклад					
Название Оклад					
инженер	50000				
старший инженер	51000				
ведущий инженер	70000				
менеджер проекта	100000				

	Сотрудник-Проект				
Таб. номер ID проек					
	1		1		
	2		1		
	2		2		
	3		2		
	4		2		



Основы SQL

SQL – Structured Query Language – язык структурированных запросов.

Условно подразделяется на:

- ▶ DDL Data Definition Language язык определения данных (операторы: CREATE, ALTER, DROP)
- ➤ DML Data Manipulation Language язык манипулирования данными (операторы: SELECT, INSERT, UPDATE, DELETE)
- ▶ DCL Data Control Language язык определения доступа к данным (операторы: GRANT, REVOKE, DENY)
- ➤ TCL Transaction Control Language язык управления транзакциями (операторы: COMMIT, ROLLBACK, SAVEPOINT)



Oсновы SQL: конфигурирование

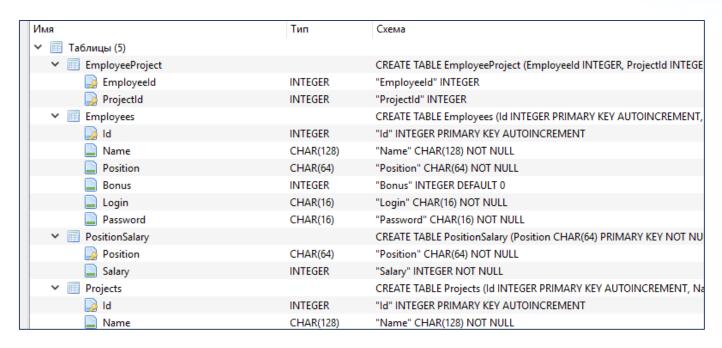
Конфигурируем Базу Данных: каждую таблицу создаем с помощью оператора CREATE.

```
CREATE TABLE Employees
(Id
          INTEGER
                    PRIMARY KEY AUTOINCREMENT,
Name CHAR (128) NOT NULL,
Position CHAR(64)
                   NOT NULL,
        INTEGER DEFAULT 0,
Bonus
Login CHAR(16) NOT NULL,
Password CHAR(16) NOT NULL);
CREATE TABLE Projects
(Id
                     PRIMARY KEY AUTOINCREMENT,
          TNTEGER
         CHAR (128)
                   NOT NULL);
Name
CREATE TABLE PositionSalary
(Position CHAR(64) PRIMARY KEY NOT NULL,
         INTEGER
Salary
                   NOT NULL);
CREATE TABLE EmployeeProject
(EmployeeId INTEGER,
ProjectId
           INTEGER,
PRIMARY KEY (EmployeeId, ProjectId));
```



Oсновы SQL: конфигурирование

Результат можно увидеть в наглядной форме с помощью SQLite браузера.



Удалить любую из созданных таблиц (при необходимости) можно с помощью оператора DROP.



Основы SQL: операторы CRUD

После создания необходимых таблиц, мы можем вносить в них данные. Для обозначения основных действий с данными существует специальная аббревиатура — CRUD (create, read, update, delete — «создать, прочесть, обновить, удалить») — акроним, обозначающий четыре базовые функции, используемые при работе с персистентными хранилищами данных.

В соответствии с CRUD в SQL имеются следующие операторы:

- ➤ INSERT оператор языка SQL, который позволяет добавить строку со значениями в таблицу.
- > SELECT оператор запроса в языке SQL, возвращающий набор данных (выборку) из базы данных. Имеет множество опций.
- ➤ UPDATE оператор языка SQL, позволяющий обновить значения в заданных столбцах таблицы.
- ➤ DELETE в языках, подобных SQL, операция удаления записей из таблицы. Критерий отбора записей для удаления определяется выражением WHERE. В случае, если критерий отбора не определён, выполняется удаление всех записей.



Основы SQL: операторы CRUD

INSERT

```
INSERT INTO Employees (Name, Position, Bonus, Login, Password) VALUES ('Иванов И.И.', 'инженер', 30000, 'ivanovi', 'ivanov123');
```

SELECT

```
SELECT * FROM Employees;
# или:
SELECT Id, Name, Bonus FROM Employees
WHERE Bonus > 20000 ORDER BY Bonus DESC LIMIT 3;
```

UPDATE

```
UPDATE Employees SET Bonus = 40000 WHERE Id = 1;
```

DELETE

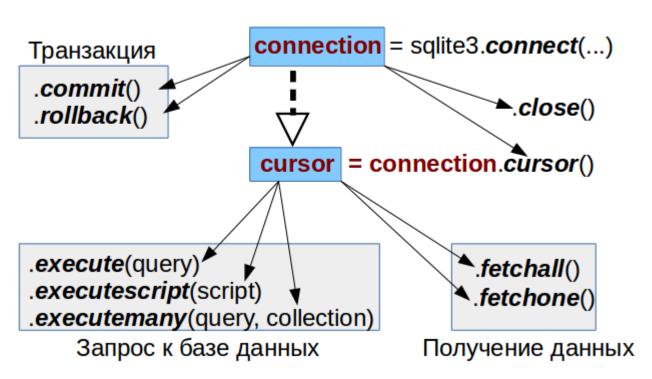
```
DELETE FROM Employees;
# или:
DELETE FROM Employees WHERE Id = 5;
```



Работа с БД в Python через DB-API

PEP 249 определяет DP-API - набор методов и интерфейсов для работы с базами данных

Python DB-API методы





БД ОГО через DB-API

Для работы с SQLite в Python используется библиотека sqlite3.

```
# Импортируем библиотеку, соответствующую типу нашей базы данных
import sqlite3
# Файл базы данных
# Если вместо файла указать : memory:, то база будет создана
# в оперативной памяти, а не в файле.
db name = "ogo.db"
# Создаем соединение с нашей базой данных
# Если файл базы данных еще не создан, он создастся автоматически.
conn = sqlite3.connect(db name)
# При необходимости меняем тип row factory, чтоб в ответах
# базы данных отображались названия атрибутов.
conn.row factory = sqlite3.Row
# РАБОТАЕМ С БАЗОЙ
# Не забываем закрыть соединение с базой данных после работы
conn.close()
```



Конфигурирование: CREATE

```
# Конфигурирование базы данных (если необходимо выполнить в скрипте)
def configure db(conn):
    cur = conn.cursor()
    # Создаем таблицу Employees
   cur.execute ("CREATE TABLE Employees"
                     (Id
                               INTEGER
                                          PRIMARY KEY AUTOINCREMENT,"
                                          NOT NULL,"
                     Name
                               CHAR (128)
                     Position CHAR(64)
                                          NOT NULL,"
                                          DEFAULT 0,"
                     Bonus
                               INTEGER
                                          NOT NULL,"
                     Login CHAR (16)
                                          NOT NULL)")
                     Password CHAR(16)
    # Создаем таблицу Projects
    cur.execute("CREATE TABLE Projects"
                     (Id
                               INTEGER
                                         PRIMARY KEY AUTOINCREMENT,"
                                         NOT NULL)")
                     Name
                               CHAR (128)
    # Создаем таблицу PositionSalary
    cur.execute("CREATE TABLE PositionSalary"
                     (Position CHAR(64) PRIMARY KEY NOT NULL,"
                     Salary
                               INTEGER
                                         NOT NULL)")
    # Создаем таблицу EmployeeProject
    cur.execute ("CREATE TABLE EmployeeProject"
                     (EmployeeId INTEGER,"
                     ProjectId INTEGER,"
                     PRIMARY KEY (EmployeeId, ProjectId))")
```



Добавление записей: INSERT



Добавление записей: INSERT

```
# Добавление записей в таблицу ДолжностьОклад
def insert position(conn, position, salary):
    cur = conn.cursor()
    cur.execute ("INSERT INTO PositionSalary (Position, Salary)"
                " VALUES (:position, :salary)",
                {'position': position, 'salary': salary})
    conn.commit()
# Добавление записей в таблицу Сотрудники
def insert employee (conn, name, position, bonus, login, pwd):
    cur = conn.cursor()
    cur.execute ("INSERT INTO Employees (Name, Position, Bonus, Login, Password)"
                " VALUES (:name, :position, :bonus, :login, :pwd)",
                {'name': name, 'position': position, 'bonus': bonus,
                 'login': login, 'pwd': pwd})
    conn.commit()
# Добавление записей в таблицу СотрудникиПроекты
def add employee to project(conn, employee id, project id):
    cur = conn.cursor()
    cur.execute("INSERT INTO EmployeeProject (EmployeeId, ProjectId)"
                " VALUES (:employeeId, :projectId)",
                {'employeeId': employee id, 'projectId': project id})
    conn.commit()
```



Создаем БД ОГО

```
db name = "ogo.db"
db exists = os.path.exists(db name)
conn = sqlite3.connect(db name)
conn.row factory = sqlite3.Row
if not db exists:
    configure db (conn)
    insert project (conn, "Важный")
    insert project (conn, "Срочный")
    insert position(conn, "инженер", 50000)
    insert position (conn, "старший инженер", 51000)
    insert position(conn, "менеджер проекта", 100000)
    insert employee (conn, "Иванов И.И.", "инженер", 30000,
                    "ivanovi", "ivanov123")
    insert employee (conn, "Петров П.П.", "старший инженер", 50000,
                    "petrovp", "p1e2t3")
    insert employee (conn, "Сидоров С.С.", "менеджер проекта", 30000,
                    "sidorovs", "zayka88")
    add employee to project (conn, 1, 1)
    add employee to project (conn, 2, 1)
    add employee to project (conn, 2, 2)
    add employee to project (conn, 3, 2)
```



Чтение данных: SELECT

```
# Проверка наличия пользователя в базе данных
# с указанным логином/паролем
def authentication(conn, login, pwd):
    cur = conn.cursor()
    # Делаем SELECT запрос к базе данных, используя обычный SQL-синтаксис
    cur.execute ("SELECT E.Id, E.Name, E.Position, EP.ProjectId"
                " FROM Employees AS E, EmployeeProject AS EP"
                " WHERE E.Id = EP.EmployeeId"
                " AND E.Login = :login AND E.Password = :pwd",
                {'login': login, 'pwd': pwd})
    # Получаем результат сделанного запроса
    return cur.fetchone()
# Проверка наличия указанного сотрудника в указанном проекте
def is employee in project (conn, employee id, project id):
    cur = conn.cursor()
    cur.execute("SELECT EP.ProjectId"
                " FROM EmployeeProject AS EP"
                " WHERE EP.EmployeeId = :employee id"
                " AND EP.ProjectId = :project id",
                {'employee id': employee id, 'project id': project id})
    return bool(cur.fetchone())
```



Чтение данных: SELECT

ld	Name	Position	Bonus		Login	Password
1	Иванов И.И.	инженер	30000	iν	anovi	ivanov123
2	Петров П.П.	старший инженер	50000	ре	trovp	p1e2t3
3	Сидоров С.С.	менеджер проекта	30000	si	dorovs	zayka88

PositionSalary					
Position	Salary				
инженер	50000				
старший инженер	51000				
менеджер проекта	100000				

Id	Name Position		Salary	Bonus	
1	Иванов И.И.	инженер	50000	30000	
2	Петров П.П.	старший инженер	51000	50000	
3	Сидоров С.С.	менеджер проекта	100000	30000	



Чтение данных: SELECT

		ee			
Id	Name	Bonus	Login	Password	
1	Иванов И.И.	инженер	30000	ivanovi	ivanov123
2	Петров П.П.	старший инженер	50000	petrovp	p1e2t3
3	Сидоров С.С.	менеджер проекта	30000	idorovs	zayka88

PositionSalary			
Position	Salary		
инженер	50000		
старший инженер	51000		
менеджер проекта	100000		

(Id	Name	Position	Salary	Bonus	ProjectID
	1	Иванов И.И.	инженер	50000	30000	1
	2	Петров П.П.	старший инженер	51000	50000	1
	2	Петров П.П.	старший инженер	51000	50000	2
	3	Сидоров С.С.	менеджер проекта	100000	30000	2

EmployeeProject			
EmployeeID	ProjectID		
1	1		
2	1		
2	2		
3	2		



Изменение данных: UPDATE и DELETE

```
# Изменение премии сотрудника
def update employee bonus(conn, employee id, new bonus):
    cur = conn.cursor()
    # Делаем UPDATE запрос к базе данных, используя обычный SQL-синтаксис
    cur.execute("UPDATE Employees"
                " SET Bonus = :new bonus"
                " WHERE Id = :employee id",
                {'employee id': employee id, 'new bonus': new bonus})
    conn.commit()
# Удаление сотрудника из проекта (но не из базы данных)
def delete employee from project (conn, employee id, project id):
    cur = conn.cursor()
    # Делаем DELETE запрос к базе данных, используя обычный SQL-синтаксис
    cur.execute("DELETE FROM EmployeeProject"
                " WHERE EmployeeId = :employee id"
                " AND ProjectId = :project id",
                {'employee id': employee id, 'project id': project id})
    conn.commit()
```



Решение задачи: бета-версия

```
login = input("Логин: ")
pwd = input ("Пароль: ")
res = authentication(conn, login, pwd)
if res:
    user = dict(res)
   print("Здравствуйте, {}".format(user['Name']))
    if user['Position'] == "менеджер проекта":
        show manager info(conn, user['ProjectId'])
        id upd = int(input("Изменение премии. ID сотрудника (0 - отмена): "))
        if id upd:
            if (id upd != user['Id'] and
                    is employee in project(conn, id upd, user['ProjectId'])):
                new bonus = input ("Новая премия: ")
                update employee bonus (conn, id upd, new bonus)
            else:
                print ("Невозможно изменить премию для данного сотрудника")
        id del = int(input("Удаление сотрудника. ID сотрудника (0 - отмена): "))
        if id del:
            if id del != user['Id']:
                delete employee from project(conn, id del, user['ProjectId'])
            else:
                print ("Невозможно удалить данного сотрудника из проекта")
    else:
        show employee info(conn, user['Id'])
else:
   print("Доступ запрещен")
```



Решение задачи: тестирование бета-версии

```
Логин: sidorovs
Пароль: zayka88
Здравствуйте, Сидоров С.С.
Информация для менеджера:
{'Id': 2, 'Name': 'Петров П.П.', 'Pay': 101000}
{'Id': 3, 'Name': 'Сидоров С.С.', 'Pay': 130000}
Изменение премии. ID сотрудника (0 - отмена): 2
Новая премия: 60000
Удаление сотрудника. ID сотрудника (0 - отмена): 0
```

```
Логин: sidorovs
Пароль: zayka88
Здравствуйте, Сидоров С.С.
Информация для менеджера:
{'Id': 2, 'Name': 'Петров П.П.', 'Pay': 111000}
{'Id': 3, 'Name': 'Сидоров С.С.', 'Pay': 130000}
Изменение премии. ID сотрудника (0 - отмена): 0
Удаление сотрудника. ID сотрудника (0 - отмена): 2
```

```
Логин: sidorovs
Пароль: 123
Доступ запрещен
```



SQL-инъекции: уязвимый код

```
Логин: ivanovi
Пароль: ivanov123
Здравствуйте, Иванов И.И.
Информация для сотрудника:
{'Id': 1, 'Name': 'Иванов И.И.', 'Pay': 80000}
```

Логин: ivanovi Пароль: 123 Доступ запрещен

```
Логин: ivanovi
Пароль: 123' OR 'a'='a
Здравствуйте, Иванов И.И.
Информация для сотрудника:
{'Id': 1, 'Name': 'Иванов И.И.', 'Pay': 80000}
```



SQL-инъекции: защищенный код

```
def authentication(conn, login, pwd):
    cur = conn.cursor()
    cur.execute ("SELECT E.Id, E.Name, E.Position, EP.ProjectId"
                " FROM Employees AS E, EmployeeProject AS EP"
                " WHERE E.Id = EP.EmployeeId"
                " AND E.Login = :login AND E.Password = :pwd",
                {'login': login, 'pwd': pwd})
    return cur.fetchone()
def authentication2(conn, login, pwd):
    cur = conn.cursor()
    cur.execute ("SELECT E.Id, E.Name, E.Position, EP.ProjectId"
                " FROM Employees AS E, EmployeeProject AS EP"
                " WHERE E.Id = EP.EmployeeId"
                " AND E.Login = ? AND E.Password = ?",
                (login, pwd))
    return cur.fetchone()
```

```
Логин: ivanovi
Пароль: 123' OR 'a'='a
Доступ запрещен
```



далее ничего не придумали, импровизируй