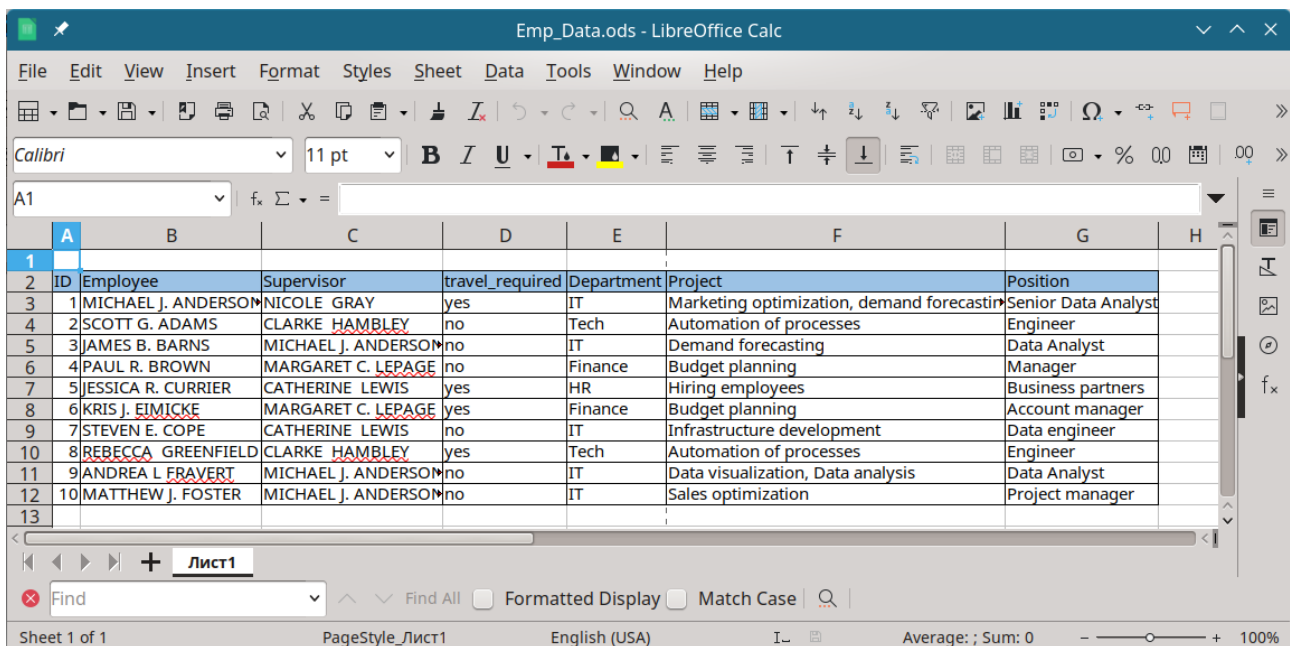


Домашнее задание 1

Матесов Данил

Часть 1

1. Создайте базу данных Emp_Data.



ID	Employee	Supervisor	travel_required	Department	Project	Position
1	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Marketing optimization, demand forecasting	Senior Data Analyst
2	SCOTT G. ADAMS	CLARKE HAMBLEY	no	Tech	Automation of processes	Engineer
3	JAMES B. BARNES	MICHAEL J. ANDERSON	no	IT	Demand forecasting	Data Analyst
4	PAUL R. BROWN	MARGARET C. LEPAGE	no	Finance	Budget planning	Manager
5	JESSICA R. CURRIER	CATHERINE LEWIS	yes	HR	Hiring employees	Business partners
6	KRIS J. EIMICKE	MARGARET C. LEPAGE	yes	Finance	Budget planning	Account manager
7	STEVEN E. COPE	CATHERINE LEWIS	no	IT	Infrastructure development	Data engineer
8	REBECCA GREENFIELD	CLARKE HAMBLEY	yes	Tech	Automation of processes	Engineer
9	ANDREA L. FRAVERT	MICHAEL J. ANDERSON	no	IT	Data visualization, Data analysis	Data Analyst
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON	no	IT	Sales optimization	Project manager

2. Проведите нормализацию таблицы, по ссылке ниже. На каждом этапе нормализации опишите к какой нормальной форме вы приводите таблицу и почему таблица в этой нормальной форме не находится.

чтобы база данных находилась в 1 нормальной форме, необходимо чтобы ее таблицы соблюдали следующие реляционные принципы:

- В таблице не должно быть дублирующих строк
- В каждой ячейке таблицы хранится атомарное значение (одно не составное значение)
- В столбце хранятся данные одного типа
- Отсутствуют массивы и списки в любом виде

В Emp_Data в столбце Project хранится список проектов, избавляемся от этого

ID	Employee	Supervisor	travel_required	Department	Project	Position
1	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Marketing optimization	Senior Data Analyst
1	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Demand forecasting	Senior Data Analyst
2	SCOTT G. ADAMS	CLARKE HAMBLEY	no	Tech	Automation of processes	Engineer
3	JAMES B. BARNES	MICHAEL J. ANDERSON	no	IT	Demand forecasting	Data Analyst
4	PAUL R. BROWN	MARGARET C. LEPAGE	no	Finance	Budget planning	Manager
5	JESSICA R. CURRIER	CATHERINE LEWIS	yes	HR	Hiring employees	Business partners
6	KRIS J. EIMICKE	MARGARET C. LEPAGE	yes	Finance	Budget planning	Account manager
7	STEVEN E. COPE	CATHERINE LEWIS	no	IT	Infrastructure development	Data engineer
8	REBECCA GREENFIELD	CLARKE HAMBLEY	yes	Tech	Automation of processes	Engineer
9	ANDREA L. FRAVERT	MICHAEL J. ANDERSON	no	IT	Data visualization	Data Analyst
9	ANDREA L. FRAVERT	MICHAEL J. ANDERSON	no	IT	Data analysis	Data Analyst
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON	no	IT	Sales optimization	Project manager

Чтобы база данных находилась во второй нормальной форме (2NF), необходимо чтобы ее таблицы удовлетворяли следующим требованиям:

- Таблица должна находиться в первой нормальной форме
- Таблица должна иметь ключ
- Все неключевые столбцы таблицы должны зависеть от полного ключа (в случае если он составной)

После предыдущего преобразования мы потеряли уникальность поля ID, и теперь имеем составной первичный ключ из полей Employee + Project. Т.е. второе требование выполнено,

9	ANDREA L FRAVERT	MICHAEL J. ANDERSON	no	IT	Data visualization	Data Analyst
9	ANDREA L FRAVERT	MICHAEL J. ANDERSON	no	IT	Data analysis	Data Analyst
10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON	no	IT	Sales optimization	Project manager

Первичный ключ

но не выполнено третье, т. к. для разных значений ПК остальные поля имеют одинаковые значения. Поэтому применяем декомпозицию и выделяем в отдельные таблицы сущности «Project», «Department», «Position» и «Suprtvisor».

ID	Employee	travel_required
1	MICHAEL J. ANDERSON	yes
2	SCOTT G. ADAMS	no
3	JAMES B. BARNES	no
4	PAUL R. BROWN	no
5	JESSICA R. CURRIER	yes
6	KRIS J. EIMICKE	yes
7	STEVEN E. COPE	no
8	REBECCA GREENFIELD	yes
9	ANDREA L FRAVERT	no
10	MATTHEW J. FOSTER	no

Supervisor
CATHERINE LEWIS
CLARKE HAMBLEY
MARGARET C. LEPAGE
MICHAEL J. ANDERSON
NICOLE GRAY

ID	Project
1	Automation of processes
2	Budget planning
3	Data analysis
4	Data visualization
5	Demand forecasting
6	Hiring employees
7	Infrastructure development
8	Marketing optimization
9	Sales optimization

Department
Finance
HR
IT
Tech

Position
Account manager
Business partners
Data Analyst
Data engineer
Engineer
Manager
Project manager
Senior Data Analyst

Связь между сотрудниками и проектами имеет тип «многие ко многим», поэтому для ее реализации понадобится еще одна таблица

Employee	Project
1	8
1	5
2	1
3	5
4	2
5	6
6	2
7	7
8	1
9	3
9	4
10	9

После этих преобразования мы снова получили несоставной первичный ключ в таблице сотрудников

Требование третьей нормальной формы (3NF) заключается в том, чтобы в таблицах отсутствовала транзитивная зависимость.

Транзитивная зависимость – это когда неключевые столбцы зависят от значений других неключевых столбцов.

Транзитивных зависимостей в таблице нет. Итоговая схема в приложенном файле Emp_Data.db

Запрос по нормализованным данным выдает тот же результат что и оригинал

```
1 SELECT
2     employee.id,
3     employee.employee,
4     supervisor.supervisor,
5     employee.travel_required,
6     department.department,
7     position.position
8 FROM employee as employee
9     join department as department on employee.department = department.id
10    join supervisor as supervisor on employee.supervisor = supervisor.id
11    join position as position on employee.position = position.id
12
```

Grid view Form view

Total rows loaded: 10

	id	employee	supervisor	travel_re	departm	position
1	1	MICHAEL J. ANDERSON	NICOLE GRAY	yes	IT	Senior Data Analyst
2	2	SCOTT G. ADAMS	CLARKE HAMBLEY	no	Tech	Engineer
3	3	JAMES B. BARNS	MICHAEL J. ANDERSON	no	IT	Data Analyst
4	4	PAUL R. BROWN	MARGARET C. LEPAGE	no	Finance	Manager
5	5	JESSICA R. CURRIER	CATHERINE LEWIS	yes	HR	Business partners
6	6	KRIS J. EIMICKE	MARGARET C. LEPAGE	yes	Finance	Account manager
7	7	STEVEN E. COPE	CATHERINE LEWIS	no	IT	Data engineer
8	8	REBECCA GREENFIELD	CLARKE HAMBLEY	yes	Tech	Engineer
9	9	ANDREA L FRAVERT	MICHAEL J. ANDERSON	no	IT	Data Analyst
10	10	MATTHEW J. FOSTER	MICHAEL J. ANDERSON	no	IT	Project manager