



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ЯДЕРНЫЙ УНИВЕРСИТЕТ
«МИФИ»

Институт интеллектуальных кибернетических систем

Кафедра № 22

**Курсовая работа по моделированию предметных сред
и проектированию баз данных**

«Приложение по заказу кофе»

Выполнил	Шамаев С.Д.
Группа	Б21-514
Вариант	Приложение по заказу кофе
Преподаватель	Дадтеев К.М.
Проверяющий	Дадтеев К.М.
Оценка	

Москва 2024

Содержание

1.	ОПИСАНИЕ ПРЕДМЕТНОЙ ОБЛАСТИ	4
1.1.	Формулировка задания	4
1.2.	Конкретизация предметной области	4
1.3.	Пользователи системы	4
1.4.	Сроки хранения информации	4
1.5.	События, изменяющие состояние Базы Данных	4
1.6.	Основные запросы к Базе Данных (на естественном языке)	4
2.	КОНЦЕПТУАЛЬНО-ИНФОРМАЦИОННАЯ МОДЕЛЬ ПРЕДМЕТНОЙ ОБЛАСТИ	5
2.1.	ER-диаграмма модели (SilverRun)	5
2.2.	Оценка мощностных характеристик сущностей и связей	5
3.	КОНЦЕПТУАЛЬНОЕ ПРОЕКТИРОВАНИЕ	6
3.1.	Принятые проектные соглашения	6
3.2.	Обоснование выбора модели базы данных	6
3.3.	Используемые в системе кодификаторы	6
3.4.	Концептуальная модель Базы Данных (SilverRun)	6
4.	ЛОГИЧЕСКОЕ ПРОЕКТИРОВАНИЕ	7
4.1.	ER-диаграмма Базы Данных (ERWin Logical)	7
4.2.	Схемы отношений Базы Данных (ERWin Phisycal)	7
4.3.	Схема реляционной Базы Данных	7
4.4.	Схемы основных запросов на реляционной алгебре	7
5.	ФИЗИЧЕСКОЕ ПРОЕКТИРОВАНИЕ	8
5.1.	Обоснование выбора конкретной СУБД	8
5.2.	Создание Базы Данных	8
5.3.	Создание Таблиц	8
5.4.	Заполнение таблиц. ETL-процессы загрузки базы данных	8
5.5.	Запросы в терминах SQL	8
5.6.	Оценка размеров Базы Данных и каждого из файлов	8

1. Описание предметной области

1.1. Формулировка задания

Спроектировать базу данных сервиса по заказу кофе. База данных должна соединить все точки кофеен разных брендов с возможностью пользователя удаленно сделать заказ.

1.2. Конкретизация предметной области

Приложение для заказа кофе - это возможность сэкономить время и усилия. Оно создано для того, чтобы сделать заказ в ближайшей для клиента точке и к определенному времени. Нередко бывает, особенно в час пик, что слишком длинные очереди в кофейне не дают возможности выпить кофе. Также, как и в незнакомом районе найти кофейню бывает проблематично. Данное приложение позволяет решить подобные проблемы.

Приложение может предоставлять информацию о расположении ближайших кафе, рейтинге заведений, отзывах других пользователей. Пользователи могут оплачивать заказы онлайн или наличными при получении.

Дополнительные функции приложения могут включать возможность создания личного профиля, сохранение предпочтений и истории заказов, получение уведомлений о специальных предложениях и акциях.

Таким образом, оно должно отвечать следующим запросам:

- Находить ближайшие к клиенту кофейни (возможно, только определенного бренда)
- Оценивать время дороги до него
- Среднее время ожидания в этой точке в текущий момент времени
- Возможность сделать заказ дистанционно

Для базы данных, на которой будет работать приложение, можно выделить следующие сущности атрибуты:

1) Клиент

- a) ФИО
- b) Никнейм
- c) Телефон
- d) Домашний адрес
- e) Адрес текущего местоположения

2) Сеть кофеен

- a) Название
- b) Ассортимент

3) Товар

- a) Название
- b) Описание
- c) Цена

4) Филиал

- a) Адрес
- b) Среднее время ожидание в текущий момент

5) Заказ

- a) Время создания
- b) Сумма
- c) Платеж
- d) Филиал
- e) Сеть кофеен

6) Платеж

- a) Сумма
- b) Время создания
- c) Статус

1.3. Пользователи системы

Клиент - получение данных о расположении кофеен, об ассортименте сети кофеен, об среднем времени ожидания в каждой точке; возможность сделать заказ и оплатить его (или выбрать оплату на месте по получении);

Администратор сети кофеен - возможность предоставить ассортимент, указать филиалы, время ожидания в них, принять заказ к текущему филиалу;

1.4. Сроки хранения информации

Данные о заказах и платежах могут храниться в течение года для анализа и рейтинга клиентов, анализа трафика в сети кофеен, выбора самых популярных.

Среднее время ожидания в текущий момент изменяется каждый час (опционально), и хранить можно в течение недели, чтобы после была возможность установить значение по умолчанию в каждый час суток;

Расположение клиентов может сохраняться в течение недели, а после необходимость в них отпадает (тк поступает всё время новая актуальная информация), домашний адрес не удаляется, но может изменяться.

После истечения срока хранения данных их можно архивировать на отдельный носитель или удалять в соответствии с внутренними правилами компании и законодательством.

1.5. События, изменяющие состояние Базы Данных

Регистрация клиента,

указание клиентом текущего местоположения,

оформление заказа, оплата или отмена заказа, добавление или изменение ассортимента сети кофеен, добавление или удаление сети кофеен.

1.6. Основные запросы к Базе Данных (на естественном языке)

- Получение данных о местоположении филиалов определенной сети кофеен
- Узнать, какие филиалы определенной сети кофеен самые популярные
- Узнать, сколько филиал получает заказов
- Узнать, сколько филиал получает выручки за день
- Узнать, сколько сеть кофеен получает выручки за день
- Выбрать любимые сети кофеен для конкретного клиента (больше всего заказов)

2. Концептуально-информационная модель предметной области

2.1. ER-диаграмма модели

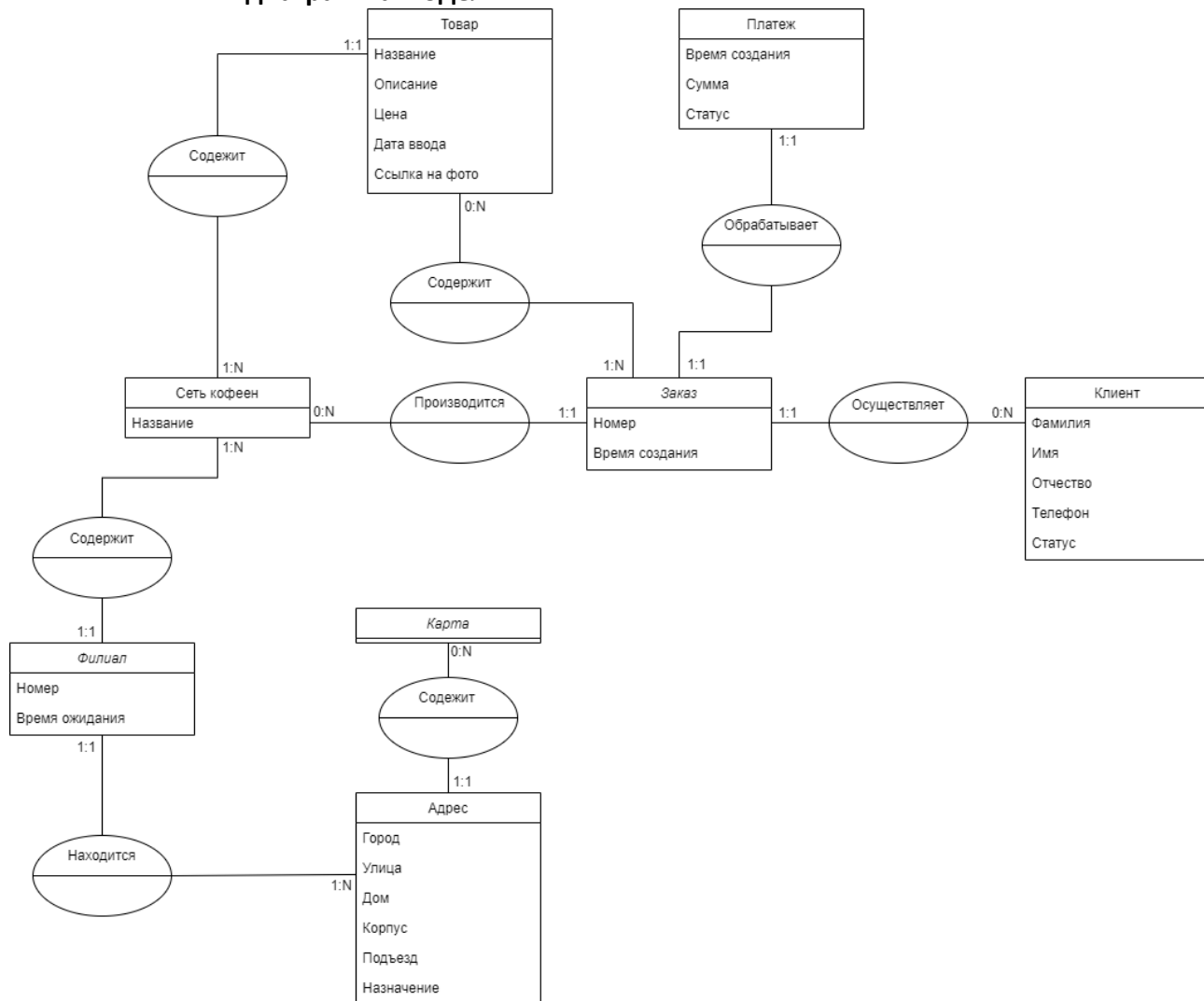


Рис. 2.1: КМПО ER-диаграмма базы данных для сервиса по заказу кофе

2.2. Оценка мощностных характеристик сущностей и связей

Сущность/связь	Мощность		
	минимальная	средняя	максимальная
Сеть кофеен	1	10	100
Товар	10	1000	30000
Филиал	1	500	40000
Клиент	0	1000	10000
Заказ	0	10000	1000000
Адрес	1	500	40000
Карта	1	1	1
Платёж	0	10000	1000000
Сеть кофеен содержит Товар	10	1000	30000
Сеть кофеен содержит Филиал	1	500	40000

Сеть кофеен производит Заказ	0	10000	1000000
Клиент осуществляет Заказ	0	10000	1000000
Филиал находится по Адресу	1	500	40000
Карта содержит Адрес	1	500	40000
Заказ содержит Товар	0	20000	200000000
Платёж обрабатывает Заказ	0	10000	1000000

По данным сети Cofix насчитывает 200 филиалов. Всего в их меню 27 позиций, причем для каждого существуют 3 размера. Итого в районе 100 позиций. В сети One Price Coffee всего 380 филиалов и всего 200 позиций.

Всего в кофейнях (небольших) обычно не более 50-100 позиций, у сети кофеен - не более 100 филиалов.

Всего сетей кофеен, которые согласятся работать с приложением по Москве не более 100. В среднем 20 крупных кофеен дают широкий охват мест (с учетом филиалов).

Товаров в одной сети для крупных сетей до 200 (максимум - 300). Минимум - 10, в среднем - 50. Тогда общее количество товаров = количество сетей кофеен * количество товаров в одной сети. Сеть кофеен содержит товар - по числу товаров.

Филиалов - для крупных сетей - до 400 (пример выше), для малых - 1, в среднем - 5. Тогда общее количество филиалов = количество сетей кофеен * количество филиалов в одной сети. Сеть кофеен содержит филиал - по числу филиалов. Филиал находится по адресу - по числу филиалов.

Количество клиентов может быть различным, в среднем 1000 (для окупаемости), максимум - 10000.

Количество заказов у одного клиента может быть от нуля (только зарегистрировался) до 100, в среднем - 10. Тогда общее количество заказов = количество клиентов * количество заказов для одного клиента. Сеть кофеен производит заказ - по числу заказов. Клиент осуществляет заказ - по числу заказов.

Адреса есть у каждого филиала, по запросу клиента находятся ближайшие к нему (адрес клиента не хранится). Тогда общее количество адресов = количество филиалов. Карта содержит все данные о всех филиалах, содержится в единственном экземпляре (к ней обращается клиент). Карта содержит адрес - по числу адресов.

Платёж есть у каждого заказа и ровно 1. Платёж обрабатывает заказ - по количеству заказов.

В одном заказе от 1 до 20 позиций товара, но в среднем - 2 (кофе, пончик). Тогда общее количество связи заказ содержит товар = количество заказов * количество товаров в одном заказе.

3. Концептуальное проектирование

3.1. Принятые проектные соглашения

Для заказа добавляется сущность «Список товаров», содержащий в себе список актуальных товаров заказа, а также конкретную сеть кофейни.

Для каждой сущности добавляется атрибут идентификатор, для удобной связи с другими сущностями.

Вводятся различные кодификаторы (см. далее).

3.2. Обоснование выбора модели базы данных

Для данной работы выбрана реляционная модель базы данных, в связи с тем, что она наиболее удобна для обработки транзакций. Данная база данных обладает высокой структурированностью, поэтому выбор реляционной модели является наилучшим решением.

3.3. Используемые в системе кодификаторы

Вводятся кодификаторы, связанные со способом оплаты (сразу через приложение онлайн, или в самой точке кофейни, наличным или безналичным расчетом). Данный кодификатор храним в ПО, т.к. не может быть изменений в составе данных кодификатора.

Способ оплаты		ПО
Код	Значение	
О	Онлайн	
N	При получении заказа наличными	
С	При получении заказа картой	

Вводятся кодификаторы, связанные со статусом клиента (новый клиент, надёжный клиент, золотой клиент и др.). Т.к. могут добавиться новые статусы (для большего разнообразия привилегий и ограничений), то для данного кодификатора разумно выделить отдельную сущность.

Статус		БД
Код	Значение	
0	новый клиент	
1	надёжный клиент	
2	золотой клиент	

3.4. Концептуальная модель Базы Данных (SilverRun)

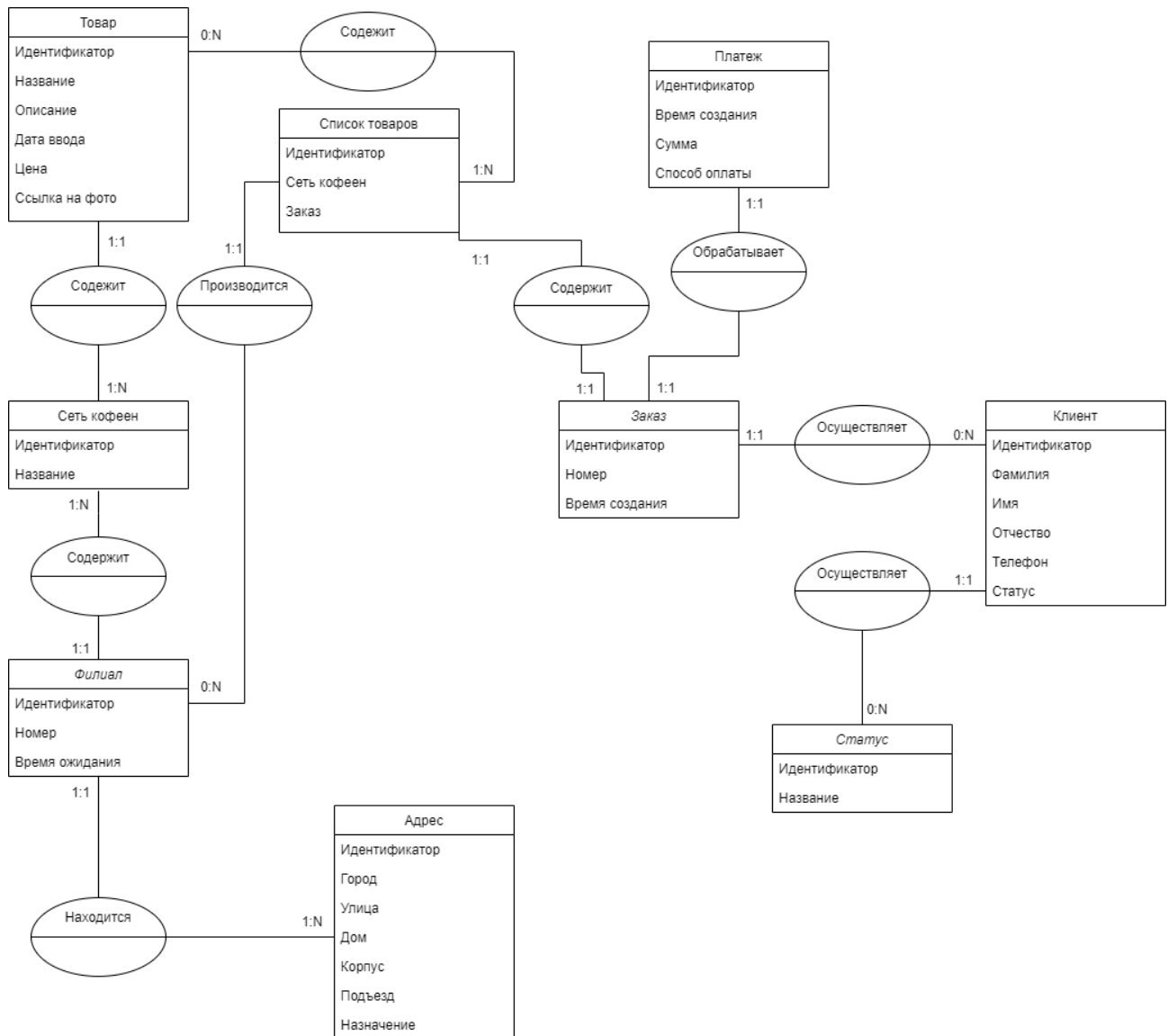


Рис. 3.1. КМБД на русском

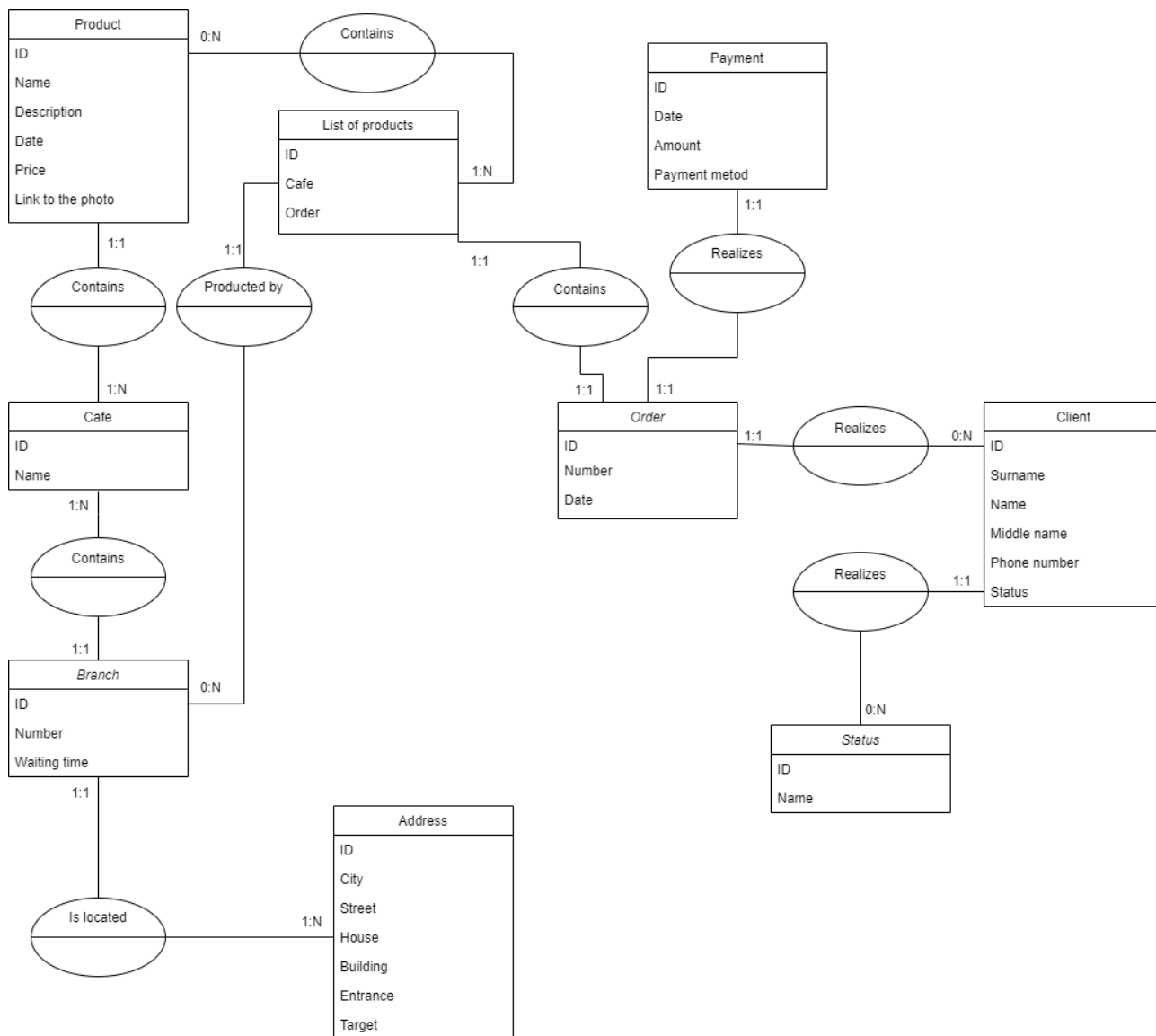


Рис. 3.2. КМБД на английском

4. Логическое проектирование

4.1. ER-диаграмма Базы Данных

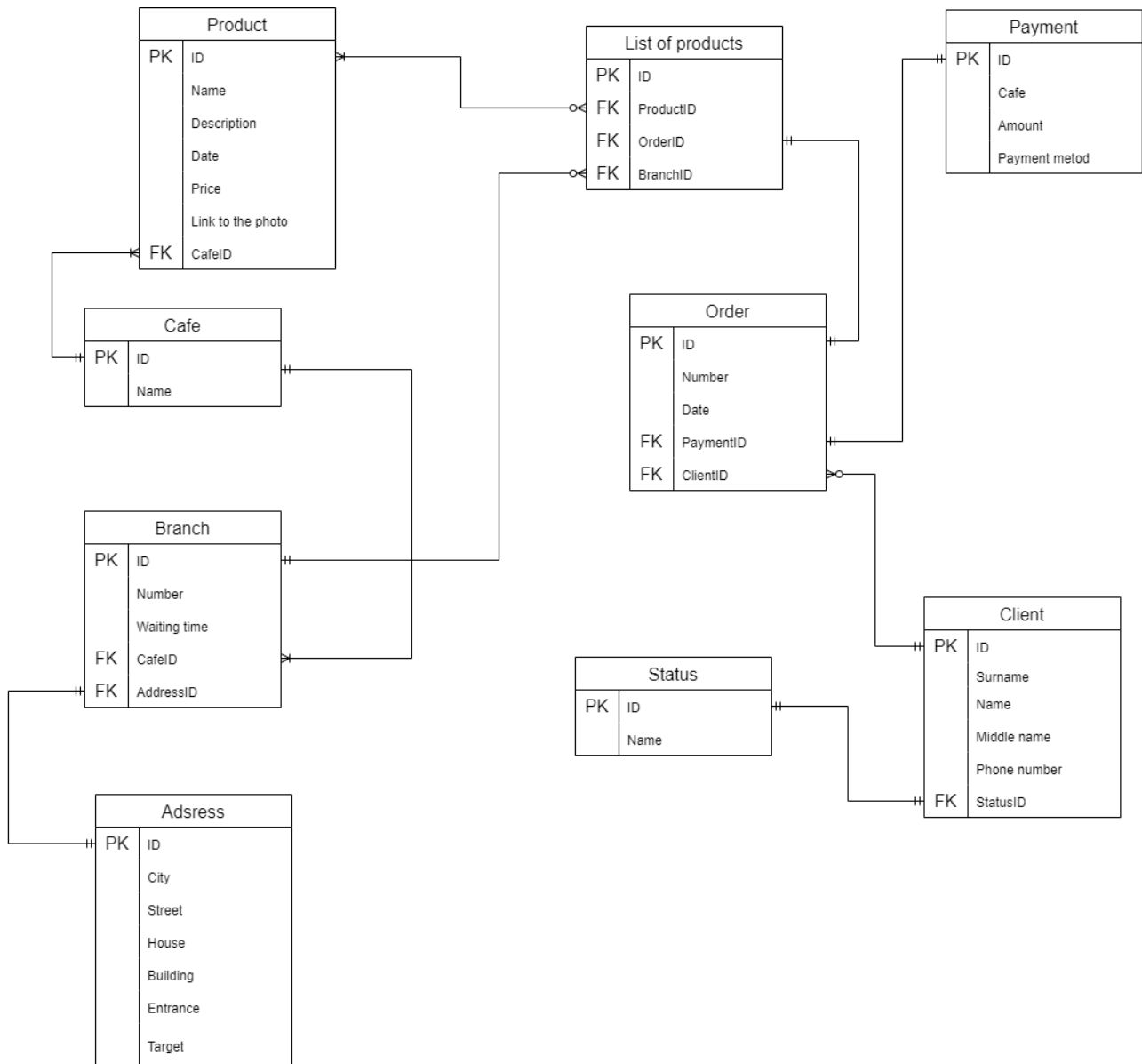


Рис. 4.1. ER-диаграмма баз данных

4.2. Схемы отношений Базы Данных (ERWin Phisycal)

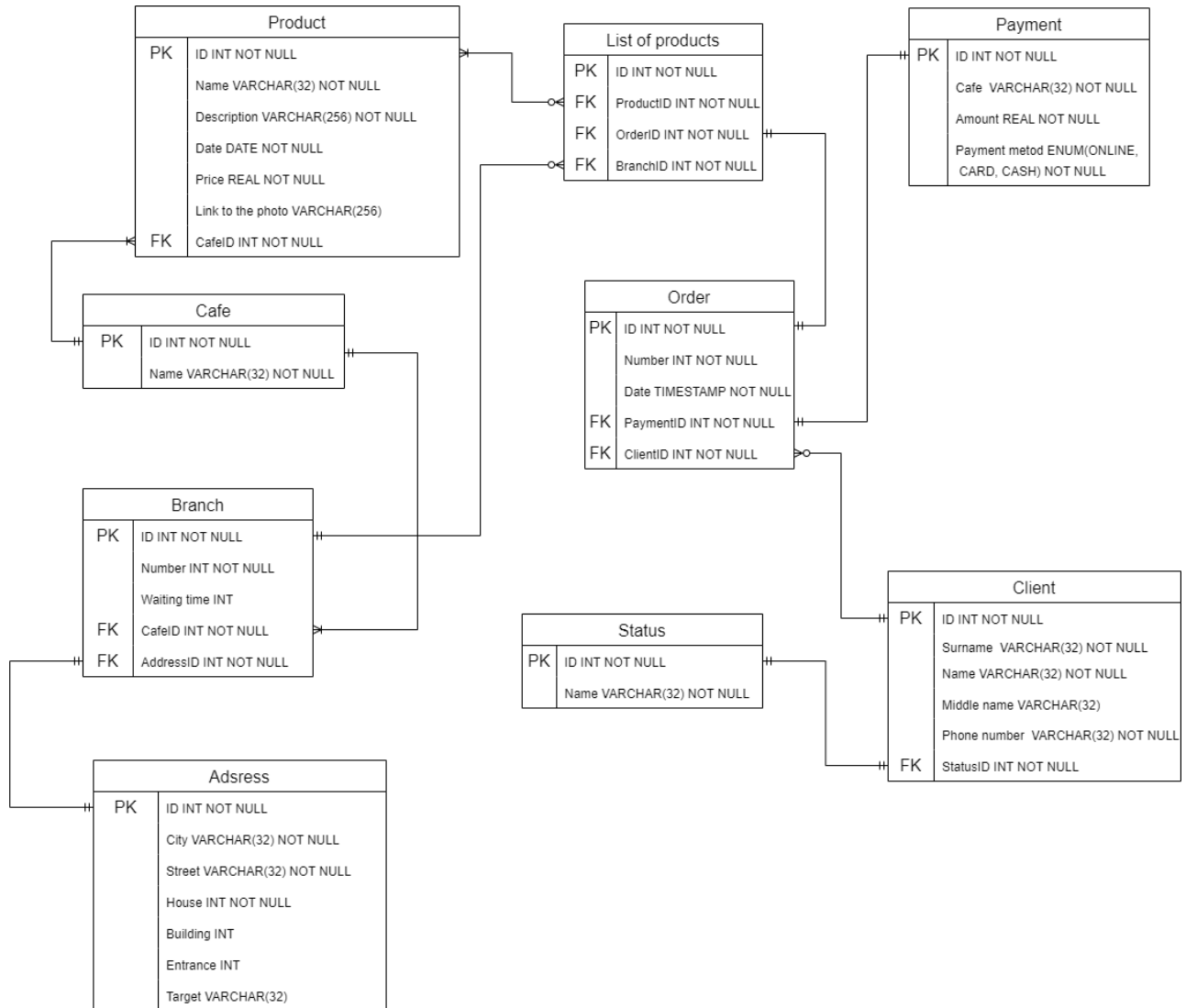


Рис. 4.2. Схема отношений базы данных

4.3. Схема реляционной Базы Данных

R1	Cafe(#ID, Name)
R2	Branch(#ID, Number, WaitingTime, CafeID, AddressID)
R3	Address(#ID, City, Street, House, Building, Entrance, Target)
R4	Product(#ID, Name, Description, CreationDate, Price, LinkToPhoto, CafeID)
R5	ListProduct(#ID, ProductID, OrdersID, BranchID)
R6	Orders(#ID, Number, CreationDate, PaymentID, ClientID)
R7	Payment(#ID, Café, Amount, PaymentMetod)
R8	Client(#ID, Surame, Name, MiddleName, PhoneNumber, StatusID)
R9	Status(#ID, Name)

4.4. Схемы основных запросов на реляционной алгебре

4.4.1. Получение данных о местоположении филиалов определенной сети кофеен (с именем name)

$T1 = R1[R1.Name = name]$

$T2 = R2[R2.CafeID = T1.ID]T1$

$T3 = T2[R2.AddressID = R3.ID]R3$

$T4 = T3[T3.City, T3.Street, T3.House, T3.Building, T3.Entrance]$

4.4.2. Узнать, какие филиалы определенной сети кофеен (с именем name) самые популярные

$T1 = R1[R1.Name = name]$

$T2 = R2[R2.CafeID = T1.ID]T1$

$T3 = R5[R5.BranchID = R2.ID]T2$

Далее производится группировка и агрегация над этой таблицей

4.4.3. Узнать, сколько определенной филиал (с номером number) определенной сети кофеен получает заказов

$T1 = R1[R1.Name = name]$

$T2 = R2[R2.CafeID = T1.ID]T1$

$T3 = T2[R2.ID = R5.BranchID]R5$

$T4 = T3[R2.Number = number]$

Далее производится группировка и агрегация над этой таблицей

4.4.4. Узнать, сколько филиал получает выручки за месяц

$T1 = R1[R1.Name = name]$

$T2 = R2[R2.CafeID = T1.ID]T1$

$T3 = T2[R2.ID = R5.BranchID]R5$

$T4 = T3[R2.Number = number]$

$T5 = T4[R5.OrderID = R6.ID, date1 < R6.Date < date2]R6$

$T6 = T5[R6.PaymentID = R7.hID]R7$

Далее производится группировка и агрегация над этой таблицей

4.4.5. Узнать, сколько сеть кофеен получает выручки за год

$T1 = R1[R1.Name = name]$

$T2 = R2[R2.CafeID = T1.ID]T1$

$T3 = T2[R2.ID = R5.BranchID]R5$

$T4 = T3[R5.OrderID = R6.ID, date1 < R6.Date < date2]]R6$

$T5 = T4[R6.PaymentID = R7.hID]R7$

Далее производится группировка и агрегация над этой таблицей

4.4.6. Выбрать любимые сети кофеен для конкретного клиента (по номеру телефона phone)

T1 = R8[R8.PhoneNumber = phone]

T2 = R6[R6.OrderID = R8.ID]T1

T3 = R5[R5.OrderID = R8.ID]T2

T4 = R2[R2.ID = T3.BranchID]T3

T5 = T4[R8.ID, R1.ID, R1.Name]

Далее производится группировка и агрегация над этой таблицей

5. Физическое проектирование

5.1. Обоснование выбора конкретной СУБД

В качестве СУБД была выбрана PostgreSQL, которая является одной из самых популярных на данный момент. Эта СУБД была выбрана в связи с следующими характеристиками:

1. PostgreSQL бесплатна и открыта для разработки и использования без ограничений лицензии.
2. Поддержка пользовательских типов данных, функций, индексов и других расширений позволяет адаптировать PostgreSQL под конкретный проект.
3. Оптимизации запросов и индексации обеспечивают высокую производительность при работе с большими объемами данных.
4. PostgreSQL гарантирует транзакционную целостность данных и обладает механизмами резервного копирования для защиты от потерь.
5. PostgreSQL можно установить на любую из популярных: Windows, MacOS, все популярные дистрибутивы Linux. Это добавляет удобство при разработке и развертывании системы.

5.2. Создание Базы Данных

Работа с БД приведена с использованием pgAdmin4, с помощью которого была создана база данных Cofe_db

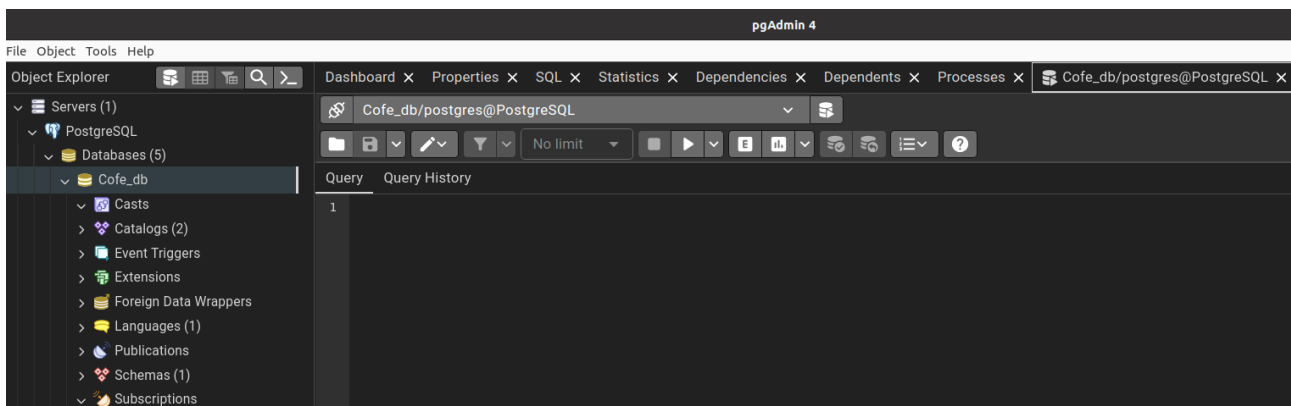


Рис. 5.1.Создание баз данных

5.3. Создание Таблиц

```
CREATE TYPE PAYMETOD AS ENUM ('Online', 'Card', 'Cash');

CREATE TABLE IF NOT EXISTS Cafe (
    ID INT PRIMARY KEY,
    Name VARCHAR(32) NOT NULL
);

CREATE TABLE IF NOT EXISTS Address (
    ID INT PRIMARY KEY,
    City VARCHAR(32) NOT NULL,
    Street VARCHAR(32) NOT NULL,
    House INT NOT NULL,
    Building INT,
    Entrance INT,
    Target VARCHAR(32)
);

CREATE TABLE IF NOT EXISTS Branch (
    ID INT PRIMARY KEY,
    Number INT NOT NULL,
    WaitingTime INT,
    CafeID INT NOT NULL REFERENCES Cafe(ID),
    AddressID INT NOT NULL REFERENCES Address(ID)
);

CREATE TABLE IF NOT EXISTS Product (
    ID INT PRIMARY KEY,
    Name VARCHAR(32) NOT NULL,
    Description VARCHAR(256) NOT NULL,
    CreationDate DATE NOT NULL,
    Price REAL NOT NULL,
    LinkToPhoto VARCHAR(256),
    CafeID INT NOT NULL REFERENCES Cafe(ID)
```

```

);

CREATE TABLE IF NOT EXISTS Payment(
    ID INT PRIMARY KEY,
    Cafe VARCHAR(32) NOT NULL,
    Amount REAL NOT NULL,
    PaymentMetod PAYMETOD NOT NULL
);

CREATE TABLE IF NOT EXISTS Status (
    ID INT PRIMARY KEY,
    Name VARCHAR(32) NOT NULL
);

CREATE TABLE IF NOT EXISTS Client (
    ID INT PRIMARY KEY,
    Surname VARCHAR(32) NOT NULL,
    Name VARCHAR(32) NOT NULL,
    MiddleName VARCHAR(32),
    PhoneNumber VARCHAR(32) NOT NULL,
    StatusID INT NOT NULL REFERENCES Status(ID)
);

CREATE TABLE IF NOT EXISTS Orders (
    ID INT PRIMARY KEY,
    Number INT NOT NULL,
    CreationDate TIMESTAMP NOT NULL,
    PaymentID INT NOT NULL REFERENCES Payment(ID),
    ClientID INT NOT NULL REFERENCES Client(ID)
);

CREATE TABLE IF NOT EXISTS ListProduct (
    ID INT PRIMARY KEY,
    ProductID INT NOT NULL REFERENCES Product(ID),
    OrdersID INT NOT NULL REFERENCES Orders(ID),
    BranchID INT NOT NULL REFERENCES Branch(ID)
);

```

Рис. 5.2. Создание таблиц

5.4. Заполнение таблиц. ETL-процессы загрузки базы данных

Данные о филиалах различных сети кофеен были получены с их сайтов или Яндекс карт (К примеру, ниже приведены части адресов филиалов для OnePriceCoffee, Cofix и Кофемании). Далее были получены Адреса в удобоваримом формате по средствам Python.

1	КОФЕМАНИЯ	Кутузовский проспект, 17	4.5 70 оценок
2	КОФЕМАНИЯ	Улица Покровка , 18 ст3	4.3 105 оценок
3	КОФЕМАНИЯ	Малый Черкасский переулок, 2	4.6 58 оценок
4	КОФЕМАНИЯ	Newa Towers-й Красногвардейский проезд, 22 ст2	4.6 29 оценок
5	КОФЕМАНИЯ	Ленинградский проспект, 37 к14 ст6	4.2 6 оценок
6	КОФЕМАНИЯ	Неглинная ГалереяТрубная площадь, 2	4.6 61 оценка
7	КОФЕМАНИЯ	БЦ Белая площадьУлица Лесная, 5	4.5 72 оценки
8	КОФЕМАНИЯ	МГК им. П.И. ЧайковскогоУлица Большая Никитская , 13	4.5 91 оценка
9	КОФЕМАНИЯ	Улица Новый Арбат, 19	4.6 63 оценки
10	КОФЕМАНИЯ	Кудринская площадь, 46 ст1	4.8 29 оценок
11	КОФЕМАНИЯ	БЦ SUMMITТверская улица, 22	4.7 48 оценок
12	КОФЕМАНИЯ	Комсомольский проспект, 21 ст2	4.8 59 оценок
13	КОФЕМАНИЯ	Мосфильмовская улица, 1а	4.7 18 оценок
14	КОФЕМАНИЯ	БЦ АврораСадовническая улица, 82 ст2	4.7 46 оценок
15	КОФЕМАНИЯ	Улица Большая Полянка, 2 ст2	4.5 21 оценка
16	КОФЕМАНИЯ	Four Winds Plaza1-я Тверская-Ямская улица, 21	4.6 29 оценок
17	КОФЕМАНИЯ	ГУМКрасная площадь, 3	4.6 26 оценок
18	КОФЕМАНИЯ	Улица Бавилова, 64 ст2	4.5 14 оценок
19	КОФЕМАНИЯ	БЦ Wall StreetВаловая улица, 35	4.8 38 оценок
20	КОФЕМАНИЯ	Даниловский рынокМятная улица, 74	5.0 9 оценок
21	КОФЕМАНИЯ	Спирidonьевский переулок, 17	5.0 1 оценка
22	КОФЕМАНИЯ	Усачевский рынокУлица Усачёва, 26	4.6 30 оценок
23	КОФЕМАНИЯ	ТРЦ ХорошоХорошёвское шоссе, 27	4.7 41 оценка
24	КОФЕМАНИЯ	МетрополисЛенинградское шоссе, 16а ст4	4.7 28 оценок
25	КОФЕМАНИЯ	Садовническая улица, 73 ст17	5.0 17 оценок
26	КОФЕМАНИЯ	Рынок на ЛенинскомЛенинский проспект, 108 ст1	5.0 2 оценки
27	КОФЕМАНИЯ	Крылатская улица, 2 ст2	5.0 23 оценки
28	КОФЕМАНИЯ	БЦ КрылатскийОсенняя улица, 11	4.8 28 оценок
29	КОФЕМАНИЯ	ТЦ Vnukovo Outlet Villageдеревня Лапшинка, вл8а к4	4.9 29 оценок
30	КОФЕМАНИЯ	Шереметьево аэропорт, терминал С	5.0 2 оценки
31	КОФЕМАНИЯ	Шереметьево аэропорт, терминал В	3.7 38 оценок
32	КОФЕМАНИЯ	Рублёво-Успенское шоссе, 22в	4.8 10 оценок
33	КОФЕМАНИЯ	Деревня Новинки, 116	5.0 5 оценок
34	КОФЕМАНИЯ	Центральная улица, 33	5.0 12 оценок
35	КОФЕМАНИЯ	Домодедовотерритория аэропорт Домодедово, 1	4.7 6 оценок
36	КОФЕМАНИЯ	Домодедовотерритория аэропорт Домодедово, 1	4.0 13 оценок

ONEPRICECOFFEE	Москва, 2-я Владимирская улица, 38	пн-пт 7:00-22:00 сб, вс: 8:00-22:00
ONEPRICECOFFEE	Москва, Измайловское шоссе, 69Д	Б: 7:00-22:00, В: 8:00-22:00
ONEPRICECOFFEE	Москва, Сходненская улица, 37	Б: 7:00 - 22:00, В: 8:00-22:00
ONEPRICECOFFEE	Москва, микрорайон Ikea к.2	пн-пт 10:00-23:00
ONEPRICECOFFEE	Москва, Комсомольская пл д.2	Будни 07:00-23:00, вых 07:00-22:00
ONEPRICECOFFEE	Москва, бульвар Адмирала Ушакова, 5	Б - 8:20-21:40, В - 9:00-21:40
ONEPRICECOFFEE	Москва, Комсомольская пл д.2	Круглосуточно
ONEPRICECOFFEE	Москва, ул. 10 лет Октября, 13	Б: 07:00-23:00, В: 08:00-22:00
ONEPRICECOFFEE	Москва, Дмитровское шоссе, 163А,	10:00-22:00
ONEPRICECOFFEE	Москва, ул. Красного маяка, д. 2	10:00-22:00
ONEPRICECOFFEE	Москва, Звенигородское ш., 1	07:30-22:30
ONEPRICECOFFEE	Москва, ул. Грузинский Вал, д. 45	7*23 будни, 8*23вых
ONEPRICECOFFEE	Москва, ул. Маршала Бирюзова, д. 16	07:00-21:00
ONEPRICECOFFEE	Москва, ул. Народного Ополчения, д.	07:00-21:00
ONEPRICECOFFEE	Москва, Ленинградский пр-кт, 74к1	Пн-Пт 7:00-23:00 Сб-Вс 8:00-22:00
ONEPRICECOFFEE	Москва, Чистопрудный б-р, д. 1А	пн-пт 8:00-22:00, пт 8:00-23:00, сб 9:00-23:00, вс 9:00-22:00
ONEPRICECOFFEE	Москва, ул. Корнея Чуковского, дом 2	08:00-22:00
ONEPRICECOFFEE	Москва, ул. Грузинский Вал, д. 28	7-22
ONEPRICECOFFEE	Москва, ул. Совхозная, 41	пн-пт 7.30-21.00
ONEPRICECOFFEE	Москва, Большой бульвар, дом 40	пн-пт 8:00-20:00
ONEPRICECOFFEE	Москва, ул. Зеленый проспект 79	7:00-22:00
ONEPRICECOFFEE	Москва, Ленинградский проспект 62	09:00-22:00
ONEPRICECOFFEE	Москва, Комсомольская пл д.2	Круглосуточно
ONEPRICECOFFEE	Москва, ул. Маршала Бирюзова, 16	07:00-22:00
ONEPRICECOFFEE	Москва, г. Московский, ул. Хабарова, д.2	10:00-22:00
ONEPRICECOFFEE	Москва, Ленинградское шоссе, 16А, стр. 4	09:00-23:00
ONEPRICECOFFEE	Москва, ул. Ткацкая, 36	8:00-18:00
ONEPRICECOFFEE	Москва, ул. Профсоюзная, 104	8:00-22:00
ONEPRICECOFFEE	Москва, Краснопресненская наб., 16 стр. 1	пн-пт 8:00-22:00 Сб-вскр 9:00-22:00

COFIX	ТЦ «Галерея Аэропорт». Ленинградский просп., 62А	Телефон: +7(925)417-71-20
COFIX	Пресненская наб., 4 стр. 1	Телефон: +7(925) 959-67-08
COFIX	Электродная улица, 2, стр. 1	Телефон: +7(977)942-16-47
COFIX	2-й Войковский проезд, 2/11	Телефон: +7(977)442-49-60
COFIX	9-я Парковая ул., 27	Телефон: +7 (999) 889-92-60
COFIX	Садовая-Спасская ул., 6, стр. 1	Телефон: 8(965)222-04-41
COFIX	Профсоюзная ул., 19	Телефон: +7 (977) 521-42-47
COFIX	Южнопортовая ул., 18	Телефон: +7 (999) 811-81-94
COFIX	Авиамоторная ул., 14	Телефон: +7 (968) 784-18-71
COFIX	Автозаводская ул., 13/1	Телефон: +7 (991) 332-67-17
COFIX	Андропова просп., 8	Телефон: +79773801454
COFIX	Арбат ул., 9 стр. 1	Телефон: +7(909)908-49-68
COFIX	Арена Плаза, ВТБ, Ленинградский просп., 36	Телефон: +7(977)904-01-48
COFIX	Барклай ул., 10	Телефон: +7(926)221-23-15
COFIX	Бауманская ул., 33/2 стр. 1	Телефон: +7(985)710-72-70
COFIX	Бауманская ул., 54	Телефон: 8(903) 101-92-19
COFIX	Тверская Застава площадь, 7 стр. 1	Телефон: +7 (926) 186-74-79
COFIX	Большая Бронная ул., 27/4	Телефон: +7(910)481-81-72
COFIX	Большая Полянка ул., д. 26, стр.2	Телефон: +7 (919) 765-48-84
COFIX	Большая Полянка ул., д. 26, стр.2	Телефон: +7(925)208-20-26

Рис. 5.3..Примеры адресов

Данные о клиентах получить невозможно, поэтому они были сгенерированы с помощью онлайн сайтов (имена) Python. Пример кода приведен ниже.


```

97 def client_creating():
98     N = 1000
99     client = ''
100     f = open('name.txt', 'r')
101     full_name = f.readlines()
102     f.close()
103
104     name = [full_name[i][:-1].split() for i in range(N)]
105     num = list(range(0, 10))
106     num = list(map(str, num))
107     phone_number = ['+7({}){}-{}-{}'.format(''.join(random.choices(num, k=3)),
108     ''.join(random.choices(num, k=3)),
109     ''.join(random.choices(num, k=2)),
110     ''.join(random.choices(num, k=2))) for i in range(N)]
111
112     status = random.choices([1, 2, 3], k=N)
113
114     for i in range(N):
115         client += 'INSERT INTO CLIENT (ID, SURNAME, NAME, MIDDLENAME, PHONENUMBER, STATUSID) VALUES' + \
116             ' ({}, \'\{}\', \'\{}\', \'\{}\', \'\{}\', {});\n'.format(i + 1, name[i][0], name[i][1], name[i][2],
117             phone_number[i], status[i])
118     return client
119
120 try:
121     # Подключение к существующей базе данных
122     connection = psycopg2.connect(dbname="Cofe_db",
123     user="sidg",
124     # пароль, который указали при установке PostgreSQL
125     password="1111",
126     host="127.0.0.1",
127     port="5432")
128     connection.set_isolation_level(ISOLATION_LEVEL_AUTOCOMMIT)
129     # Курсор для выполнения операций с базой данных
130     cursor = connection.cursor()
131     sql_creating_client = client_creating()
132     print(sql_creating_client)
133     cursor.execute(sql_creating_client)
134     connection.commit()
135
136
137 except (Exception, Error) as error:
138     print("Ошибка при работе с PostgreSQL", error)
139 finally:
140     if connection:
141         cursor.close()
142         connection.close()
143     print("Соединение с PostgreSQL закрыто")

```

Рис. 5.4. Генерация и заполнение Client

	id [PK] integer	city character varying (32)	street character varying (32)	house integer	building integer	entrance integer
52	52	Москва	Товарищеская улица	79	1	1
53	53	Москва	Ильинские Ворота, площадь	38	[null]	1
54	54	Москва	Микрорайон 10-й, микрорайон	57	[null]	[null]
55	55	Москва	Берников переулок	111	[null]	[null]
56	56	Москва	Микрорайон 12-й, микрорайон	13	[null]	[null]

	id [PK] integer	number integer	waitingtime integer	cafeid integer	addressid integer
1	1	43	[null]	6	104
2	2	251	[null]	4	215
3	3	63	4	4	490
4	4	44	4	4	114
5	5	299	5	2	460
6	6	329	4	9	160
7	7	355	3	8	175
8	8	301	1	8	485

	id [PK] integer	name character varying (32)
1	1	Cofix
2	2	OnePriceCoffee
3	3	Stars Coffee
4	4	Кофемания
5	5	Новатор
6	6	Finch
7	7	Starbucks

	id [PK] integer	surname character varying (32)	name character varying (32)	middlename character varying (32)	phonenummer character varying (32)	statusid integer
1	1	Березницкий	Геннадий	Русланович	+7(499)053-13-17	2
2	2	Валько	Руслан	Станиславович	+7(856)553-19-29	3
3	3	Капиносова	Жанна	Аркадьевна	+7(068)176-20-89	1
4	4	Толпыгин	Филипп	Ильдарович	+7(377)215-84-58	2
5	5	Приемова	Раиса	Радиковна	+7(777)572-70-47	2
6	6	Малышко	Диана	Николаевна	+7(541)920-01-20	3
7	7	Козлитина	Ксения	Романовна	+7(563)103-35-02	3
8	8	Городникова	Раиса	Ринатовна	+7(025)695-17-75	1

	id [PK] integer	productid integer	ordersid integer	branchid integer
1	1	88	1026	362
2	2	405	7189	56
3	3	421	7485	9
4	4	536	859	127
5	5	510	8761	382
6	6	532	7808	15
7	7	822	4127	48
8	8	47	2877	449

	id [PK] integer	cafe character varying (32)	amount real	paymentmetod paymetod
1	1	Success	250	Card
2	2	Success	310	Card
3	3	Success	310	Online
4	4	Success	160	Card
5	5	Success	220	Online
6	6	Success	220	Card
7	7	Success	130	Online
8	8	Success	220	Online

	id [PK] integer	number integer	creationdate timestamp without time zone	paymentid integer	clientid integer
1	1	7574	2023-04-06 15:48:15	7566	516
2	2	7411	2023-04-03 13:22:38	5004	520
3	3	9887	2023-10-19 19:21:25	7101	904
4	4	2810	2023-12-02 06:44:58	3418	882
5	5	8223	2023-12-04 03:53:48	4186	471
6	6	1980	2023-04-12 19:14:42	4466	221
7	7	7531	2023-04-20 20:35:28	4368	361
8	8	5895	2023-03-19 05:58:02	6624	661

	id [PK] integer	name character varying (32)
1	1	Новый клиент
2	2	Надежный клиент
3	3	Золотой клиент

Рис. 5.5. Примеры заполнения таблиц Address, Branch, Café, Client, ListProduct, Payment, Orders, Status (сверху вниз)

5.5. Запросы в терминах SQL

5.5.1. Получение данных о местоположении филиалов определенной сети кофеен (с именем name)

```
SELECT
a.city, a.street, a.house
FROM
address a
JOIN branch b ON a.ID = b.addressID
JOIN cafe c ON b.cafeID = c.ID
WHERE c.name = 'McCafe'
```

	city character varying (32) 🔒	street character varying (32) 🔒	house integer 🔒
1	Москва	Авиационный переулок	80
2	Москва	Красковская улица	119
3	Москва	Мещанская улица	84
4	Москва	Средний Каретный переулок	57
5	Москва	Ильменский проезд	31
6	Москва	Вельяминовская улица	61
7	Москва	Солдатская улица	53
8	Москва	Веткин проезд	67

Рис. 5.5. Выполнение запроса 1

5.5.2. Узнать, какие филиалы определенной сети кофеен (с именем name) самые популярные

```
SELECT
c.name name_cafe, b.number number_branch, COUNT(l.ID) number_orders_in_branch
FROM
listproduct l
JOIN branch b ON b.ID = l.branchID
JOIN cafe c ON b.cafeID = c.ID
WHERE c.name = 'McCafe'
GROUP BY c.name, b.number ORDER BY COUNT(l.ID) DESC
LIMIT 3
```

	name_cafe character varying (32) 🔒	number_branch integer 🔒	number_orders_in_branch bigint 🔒
1	McCafe	281	47
2	McCafe	56	41
3	McCafe	20	39

Рис. 5.6. Выполнение запроса 2

5.5.3. Узнать, сколько определенной филиал (с номером number) определенной сети кофеен получает заказов

```
SELECT
c.name name_cafe, b.number number_branch, COUNT(l.ID) number_orders_in_branch
FROM
listproduct l
JOIN branch b ON b.ID = l.branchID
JOIN cafe c ON b.cafeID = c.ID
WHERE c.name = 'McCafe' AND b.number = 99
GROUP BY c.name, b.number
```

	name_cafe character varying (32)	number_branch integer	number_orders_in_branch bigint
1	McCafe	99	25

Рис. 5.7. Выполнение запроса 3

5.5.4. Узнать, сколько филиал получает выручки за год

```
SELECT
c.name name_cafe, b.number number_branch, SUM(p.amount) revenue, COUNT(p.amount)
FROM
payment p
JOIN orders o ON p.ID = o.paymentID
JOIN listproduct l ON o.ID = l.ordersID
JOIN branch b ON b.ID = l.branchID
JOIN cafe c ON b.cafeID = c.ID
WHERE c.name = 'McCafe' AND b.number = 99 AND o.creationdate BETWEEN '2023-01-01' AND '2023-12-31'
GROUP BY c.name, b.number
```

	name_cafe character varying (32)	revenue real	number_of_order bigint
1	McCafe	26420	137

Рис. 5.8. Выполнение запроса 4

5.5.5. Узнать, сколько сеть кофеен получает выручки за месяц

```
SELECT
c.name name_cafe, SUM(p.amount) revenue, COUNT(p.amount) number_of_order
FROM
payment p
JOIN orders o ON p.ID = o.paymentID
JOIN listproduct l ON o.ID = l.ordersID
JOIN branch b ON b.ID = l.branchID
JOIN cafe c ON b.cafeID = c.ID
WHERE c.name = 'McCafe' AND o.creationdate BETWEEN '2023-10-01' AND '2023-10-31'
GROUP BY c.name
```

	name_cafe character varying (32) 🔒	number_branch integer 🔒	revenue real 🔒	count bigint 🔒
1	McCafe	99	4750	25

Рис. 5.9. Выполнение запроса 5

5.5.6. Выбрать любимые сети кофеен для конкретного клиента (по номеру телефона phone)

```
SELECT
c.name name_cafe, cl.surname, cl.name, cl.middlename, COUNT(cl.phonenumber) number_orders
FROM
cafe c
JOIN branch b ON c.ID = b.cafeID
JOIN listproduct l ON b.ID = l.branchID
JOIN orders o ON o.ID = l.ordersID
JOIN client cl ON cl.ID = o.clientID
WHERE cl.phonenumber = '+7(141)210-86-71'
GROUP BY c.name, cl.surname, cl.name, cl.middlename, cl.phonenumber ORDER BY COUNT(cl.phonenumber) DESC
LIMIT 5
```

	name_cafe character varying (32) 🔒	surname character varying (32) 🔒	name character varying (32) 🔒	middlename character varying (32) 🔒	number_orders bigint 🔒
1	OnePriceCoffee	Багданова	Лидия	Леонидовна	3
2	Starbucks	Багданова	Лидия	Леонидовна	3
3	McCafe	Багданова	Лидия	Леонидовна	3
4	Новатор	Багданова	Лидия	Леонидовна	2
5	Coffee Like	Багданова	Лидия	Леонидовна	2

Рис. 5.10. Выполнение запроса 6

5.6. Оценка размеров Базы Данных и каждого из файлов

Отношение	Атрибут	Тип данных	Размер, байт	Среднее количество	Объем, байт
Cafe	ID	INTEGER	4	10	400
	Name	VARCHAR(32)	36		
Address	ID	INTEGER	4	500	62000
	City	VARCHAR(32)	36		
	Street	VARCHAR(32)	36		
	House	INTEGER	4		
	Building	INTEGER	4		
	Entrance	INTEGER	4		
	Target	VARCHAR(32)	36		
Branch	ID	INTEGER	4	500	10000
	Number	INTEGER	4		
	WaitigTime	INTEGER	4		
	CafeID	INTEGER	4		
	AddressID	INTEGER	4		
Product	ID	INTEGER	4	1000	570000
	Name	VARCHAR(32)	36		
	Description	VARCHAR(256)	259		

	CreationDate	DATE	4		
	Price	REAL	4		
	LinkToPhoto	VARCHAR(256)	259		
	CafeID	INTEGER	4		
Payment	ID	INTEGER	4	10000	480000
	Cafe	VARCHAR(32)	36		
	Amount	REAL	4		
	PaymentMetod	ENUM	4		
Status	ID	INTEGER	4	3	120
	Name	VARCHAR(32)	36		
Client	ID	INTEGER	4	1000	152000
	Surname	VARCHAR(32)	36		
	Name	VARCHAR(32)	36		
	MiddleName	VARCHAR(32)	36		
	PhoneNumber	VARCHAR(32)	36		
	StatusID	INTEGER	4		
Orders	ID	INTEGER	4	10000	240000
	Number	INTEGER	4		
	CreationDate	TIMESTAMP	8		
	PaymentID	INTEGER	4		
	ClientID	INTEGER	4		
ListProduct	ID	INTEGER	4	100000	1600000
	ProductID	INTEGER	4		
	OrderID	INTEGER	4		
	BranchID	INTEGER	4		

Средний размер базы данных 3114520 байт, т.е. примерно 380 КБ.

6. Применение для аналитики

При помощи SQL-запросов выбираются данные из таблицы. Для удобства дальнейшей обработки они складываются в csv-файл. В дальнейшем этот файл обрабатывается в Python для построения графиков.

Выведем, как менялось количество заказов филиала 99 кофейни „McCafe“ с течением времени по месяцам.

```
SELECT
DATE_TRUNC('month', o.creationdate) months, c.name name_cafe, b.number number_branch, COUNT(l.ID) number_orders_in_branch
FROM
listproduct l
JOIN branch b ON b.ID = l.branchID
JOIN cafe c ON b.cafeID = c.ID
JOIN orders o ON o.ID = l.ordersID
WHERE c.name = 'McCafe' AND b.number = 99
GROUP BY months, c.name, b.number
```

	months timestamp without time zone	name_cafe character varying (32)	number_branch integer	number_orders_in_branch bigint
1	2023-01-01 00:00:00	McCafe	99	3
2	2023-02-01 00:00:00	McCafe	99	2
3	2023-03-01 00:00:00	McCafe	99	3
4	2023-05-01 00:00:00	McCafe	99	4
5	2023-07-01 00:00:00	McCafe	99	1
6	2023-08-01 00:00:00	McCafe	99	3
7	2023-09-01 00:00:00	McCafe	99	2
8	2023-10-01 00:00:00	McCafe	99	3

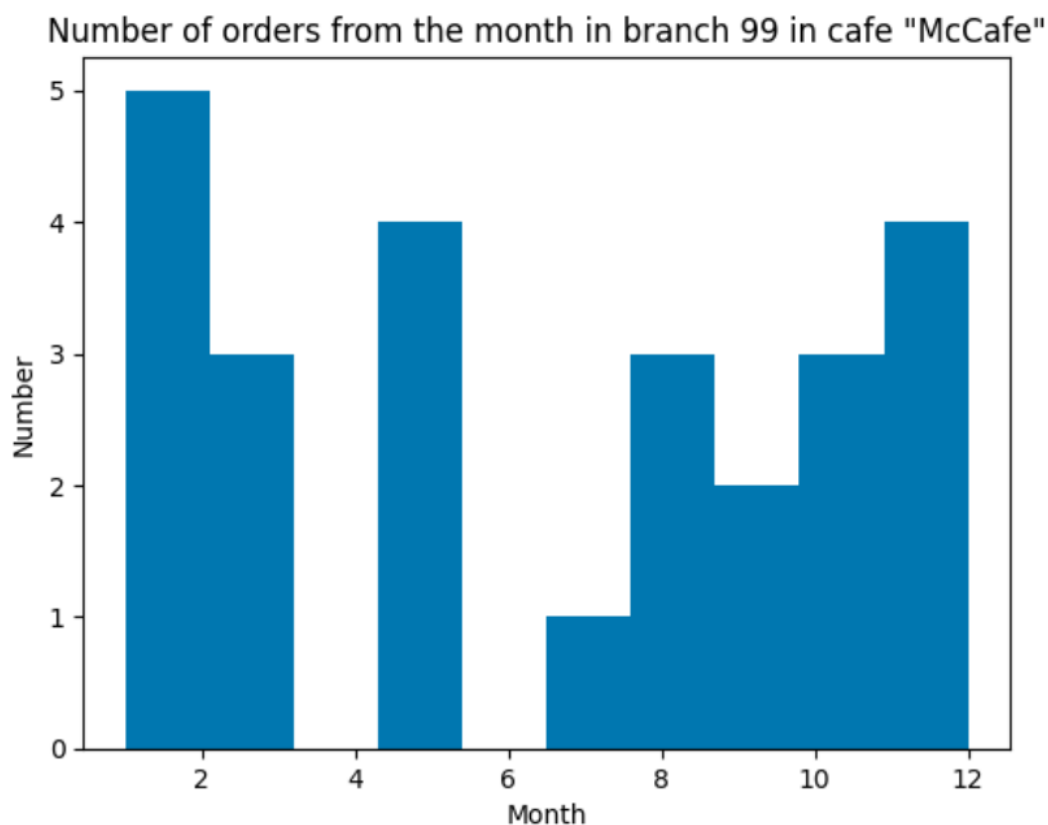


Рис. 6.1. Выполнение анализа 1

Выведем, как менялось количество заказов кофейни „McCafe“ с течением времени по месяцам.

```
SELECT
DATE_TRUNC('month', o.creationdate) months, c.name name_cafe, COUNT(l.ID) number_orders_in_branch
FROM
listproduct l
JOIN branch b ON b.ID = l.branchID
JOIN cafe c ON b.cafeID = c.ID
JOIN orders o ON o.ID = l.ordersID
WHERE c.name = 'McCafe'
GROUP BY months, c.name
```

	months timestamp without time zone 🔒	name_cafe character varying (32) 🔒	number_orders_in_branch bigint 🔒
1	2023-01-01 00:00:00	McCafe	152
2	2023-02-01 00:00:00	McCafe	119
3	2023-03-01 00:00:00	McCafe	127
4	2023-04-01 00:00:00	McCafe	120
5	2023-05-01 00:00:00	McCafe	142
6	2023-06-01 00:00:00	McCafe	124
7	2023-07-01 00:00:00	McCafe	125
8	2023-08-01 00:00:00	McCafe	150

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

gr1 = pd.read_csv('6_2.csv')
gr1['number_orders_in_branch'] = gr1['number_orders_in_branch'].astype('int')
for i in range(gr1.shape[0]):
    gr1['months'][i] = int(gr1['months'][i][5:7])

plt.hist(gr1['months'], weights=gr1['number_orders_in_branch'])
plt.xlabel('Month')
plt.ylabel('Number')
plt.title('Number of orders from the month in cafe "McCafe"')
plt.show()
```

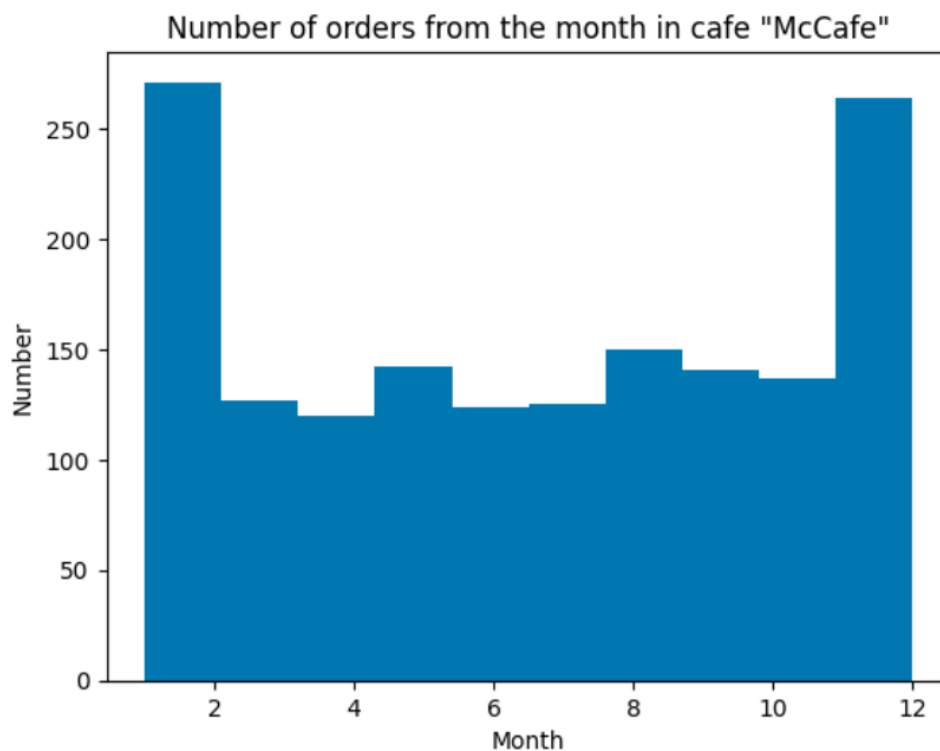



Рис. 6.2. Выполнение анализа 2

Выведем, в какие сети кофеен ходили новые клиенты.

```
SELECT
c.name name_cafe, s.name status, COUNT(cl.phonenumber) number_orders
FROM
cafe c
JOIN branch b ON c.ID = b.cafeID
JOIN listproduct l ON b.ID = l.branchID
JOIN orders o ON o.ID = l.ordersID
JOIN client cl ON cl.ID = o.clientID
JOIN status s ON s.ID = cl.statusID
WHERE s.name = 'Новый клиент'
GROUP BY c.name, s.name ORDER BY COUNT(cl.phonenumber) DESC
```

	name_cafe character varying (32) 🔒	status character varying (32) 🔒	number_orders bigint 🔒
1	Кофемания	Новый клиент	588
2	Coffee Like	Новый клиент	586
3	Stars Coffee	Новый клиент	534
4	McCafe	Новый клиент	526
5	OnePriceCoffee	Новый клиент	514
6	Starbucks	Новый клиент	497
7	Новатор	Новый клиент	482
8	Love Coffee	Новый клиент	473

```
gr1 = pd.read_csv('6_3.csv')
gr1['number_orders'] = gr1['number_orders'].astype('int')

plt.figure(figsize=(14,6))
plt.hist(gr1['name_cafe'], weights=gr1['number_orders'])
plt.xlabel('Name cafe')
plt.ylabel('Number')
plt.title('Number of orders by new client')
plt.show()
```

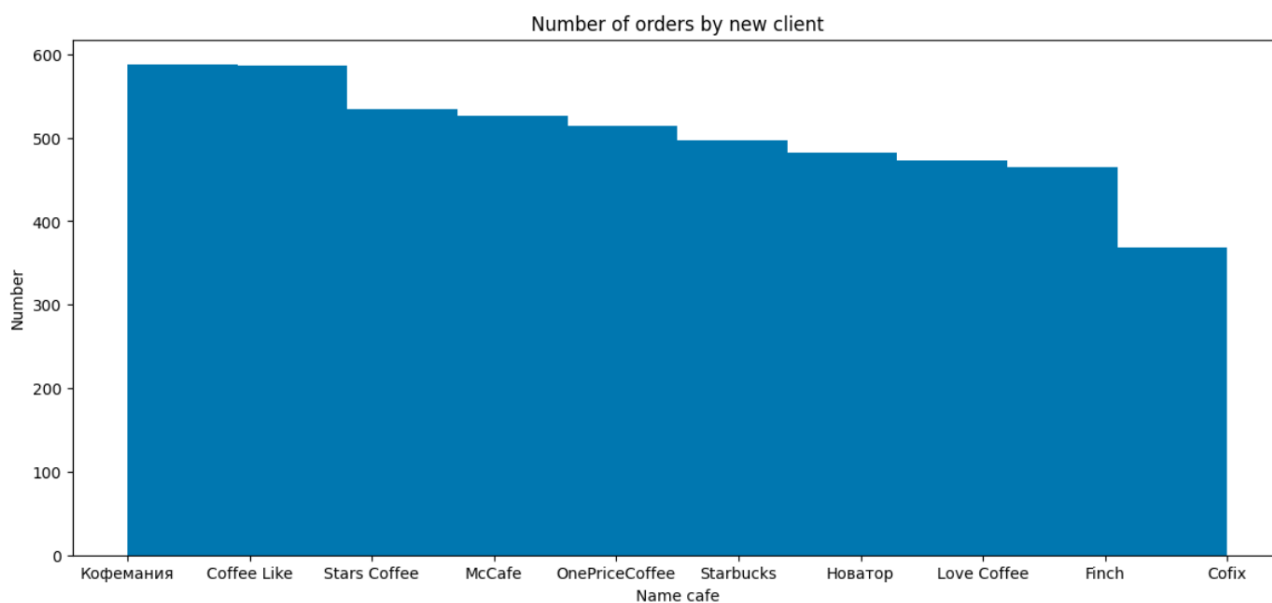


Рис. 6.3. Выполнение анализа 3

График посещаемости клиента с определенным номером телефона определенной сети кофеен.

```
SELECT
DATE_TRUNC('month', o.creationdate) months, c.name name_cafe, COUNT(l.ID) number_orders_in_branch
FROM
cafe c
JOIN branch b ON c.ID = b.cafeID
JOIN listproduct l ON b.ID = l.branchID
JOIN orders o ON o.ID = l.ordersID
JOIN client cl ON cl.ID = o.clientID
JOIN status s ON s.ID = cl.statusID
WHERE c.name = 'McCafe' AND s.name = 'Новый клиент'
GROUP BY months, name_cafe
```

	months timestamp without time zone	name_cafe character varying (32)	number_orders_in_branch bigint
1	2023-01-01 00:00:00	McCafe	45
2	2023-02-01 00:00:00	McCafe	47
3	2023-03-01 00:00:00	McCafe	39
4	2023-04-01 00:00:00	McCafe	37
5	2023-05-01 00:00:00	McCafe	43
6	2023-06-01 00:00:00	McCafe	34
7	2023-07-01 00:00:00	McCafe	33
8	2023-08-01 00:00:00	McCafe	46
9	2023-09-01 00:00:00	McCafe	59

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

gr1 = pd.read_csv('6_4.csv')
gr1['number_orders_in_branch'] = gr1['number_orders_in_branch'].astype('int')
for i in range(gr1.shape[0]):
    gr1['months'][i] = int(gr1['months'][i][5:7])

plt.plot(gr1['months'], gr1['number_orders_in_branch'], c = 'r')
plt.xlabel('Month')
plt.ylabel('Number')
plt.title('Number of orders in cafe "McCafe" by new client')
plt.show()
```

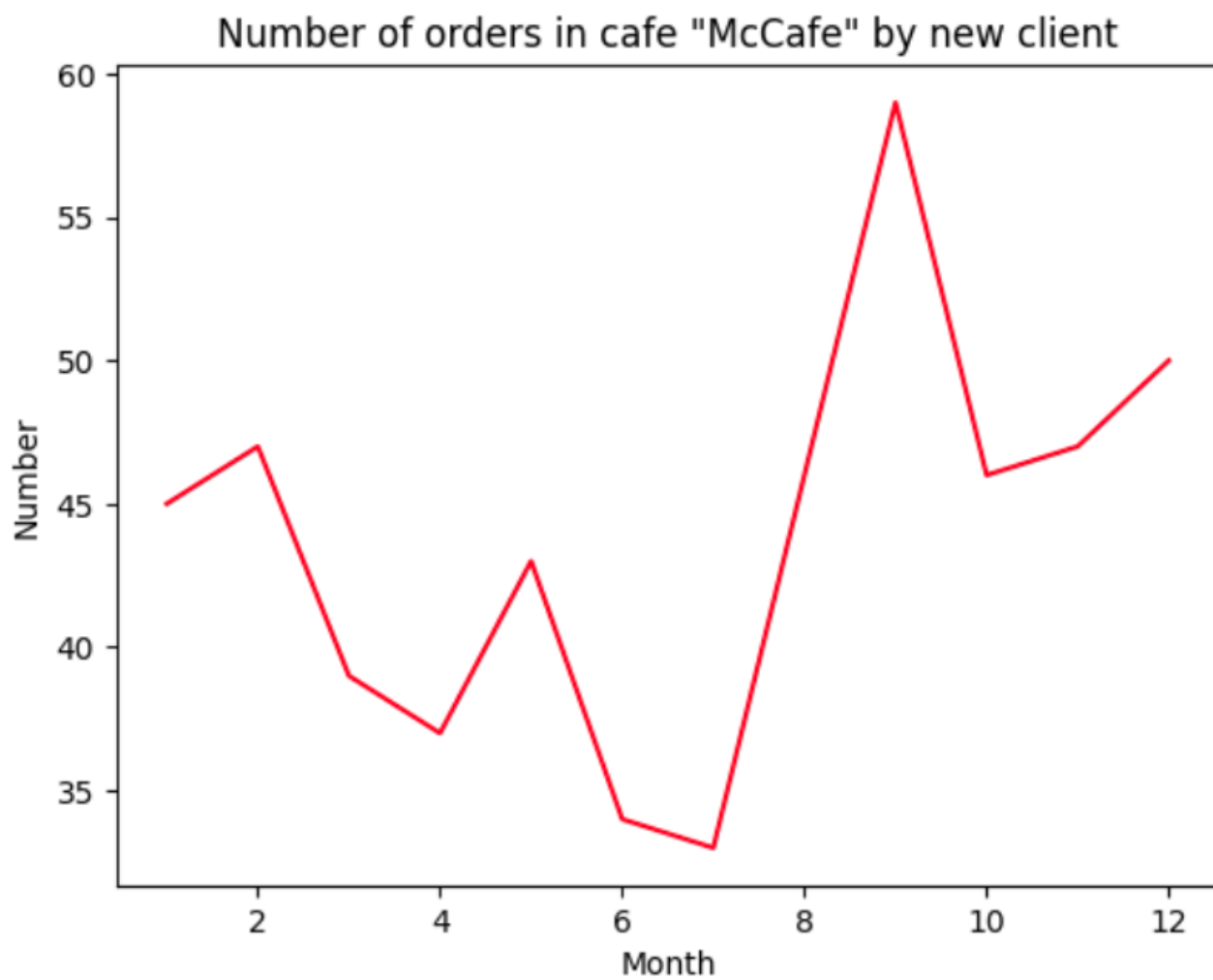


Рис. 6.4. Выполнение анализа 4