

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ

федеральное государственное автономное образовательное учреждение высшего
образования Санкт-Петербургский национальный
исследовательский университет информационных технологий механики и оптики

Мега факультет трансляционных информационных технологий

Факультет информационных технологий и программирования

Лабораторная работа № 4
По дисциплине «Проектирование баз данных»
Маскировка и анонимизация данных

Выполнил студент группы М34091

Джахан Исрат

Санкт-Петербург

2024

Задачи:

- 1. Замаскировать поля с конфиденциальными данными.
- 2. Провести анонимизацию данных.

Порядок выполнения работы:

- 1. Установите расширение PostgreSQL Anonymizer.
- 2. Выберите поля, которые необходимо замаскировать и модифицируйте уже существующие таблицы или представления. Динамическая маскировка данных.
- 3. Выберите данные, которые можно анализировать, скрыв, обобщив конфиденциальные данные. Создайте три Materialized Views используя:
 - 1) Generalization – заменяет данные более широкими и менее точными значениями, диапазонами.
 - 2) Используйте 2 стратегии анонимизации из списка:
 - Destruction
 - Adding Noise
 - Randomization
 - Faking
 - Advanced Faking
 - Pseudonymization
 - Generic Hashing
 - Partial scrambling
- 4. Предоставить отчёт.

Выбранная таблица для маскировки данных:

Таблица client

	client_id [PK] integer	client_name character varying (100)	email character varying (100)	phone character varying (15)	address character varying (255)	date_of_birth date
1	1	Ivan Petrov	ivan.petrov@example.ru	19812345678	Lenina St, 5	1985-02-15
2	2	Olga Sidorova	olga.sidorova@example.ru	89098765432	Kirova St, 12	1990-06-25
3	3	Dmitry Ivanov	dmitry.ivanov@example.ru	89034567890	Gagarina St, 25	1988-08-10
4	4	Anna Kuznetsova	anna.kuznetsova@example.ru	89101223344	Pushkina St, 8	1983-07-01
5	5	Sergey Popov	sergey.popov@example.ru	89111223344	Mira St, 17	1983-07-01
6	6	Elena Smirnova	elena.smirnova@example.ru	89091223345	Zhukova St, 21	1991-09-19
7	7	Roman Kozlov	roman.kozlov@example.ru	89098765431	Victory St, 42	1995-04-29
8	8	Irina Fedorova	irina.fedorova@example.ru	89055557788	Volgogradsky Prospect, 25	1995-04-29
9	9	Mikhail Orlov	mikhail.orlov@example.ru	89012345678	Tverskaya St, 15	1987-01-10
10	10	Nina Voronova	nina.voronova@example.ru	89091223346	Nevsky Prospect, 19	1993-11-15
11	11	Yuri Zaitsev	yuri.zaitsev@example.ru	89071223347	Moscow St, 45	1984-10-25
12	12	Natalya Lebedeva	natalya.lebedeva@example.ru	89012345678	Liteyny Prospect, 20	1995-05-05
13	13	Andrey Grigoriev	andrey.grigoriev@example.ru	89023456789	Kutuzovsky Prospect, 14	1986-02-18
14	14	Ekaterina Sorokina	ekaterina.sorokina@example.ru	89044556677	Arbat St, 37	1986-02-18
15	15	Vladimir Sokolov	vladimir.sokolov@example.ru	89012345678	Bolshaya Morskaya St, 9	1990-06-17
16	16	Marina Pavlova	marina.pavlova@example.ru	89087654321	Frunzenskaya Embankment, 10	1988-09-25
17	17	Alexey Gusev	alexey.gusev@example.ru	89098765432	Rublevskoye Shosse, 58	1988-09-25
18	18	Oksana Krylova	oksana.krylova@example.ru	89023456789	Taganskaya St, 29	1991-09-15
19	19	Viktor Nikitin	viktor.nikitin@example.ru	89012345679	Sadovaya St, 70	1985-05-15
20	20	Anastasia Belova	anastasia.belova@example.ru	89077889900	Profsoyuznaya St, 8	1993-08-01

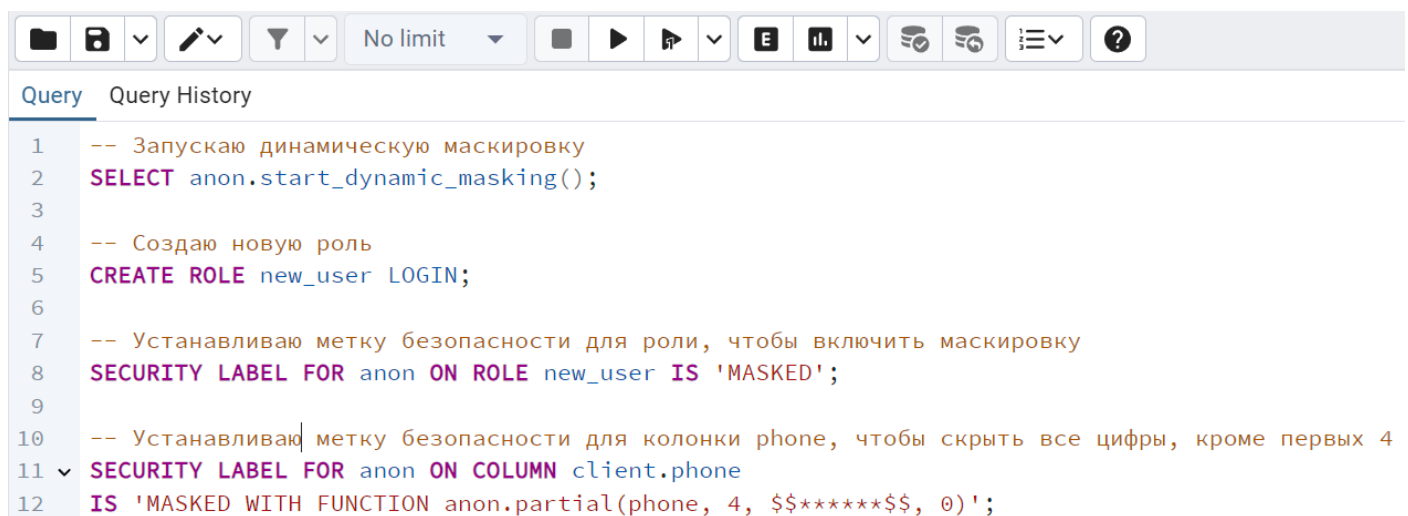
1. Установка расширения PostgreSQL Anonymizer:

- 1) `docker pull registry.gitlab.com/dalibo/postgresql_anonymizer:stable`
- 2) `docker run -d -e POSTGRES_PASSWORD=x -p 6543:5432 registry.gitlab.com/dalibo/postgresql_anonymizer`
- 3) `docker ps`

CONTAINER ID	IMAGE	COMMAND	CREATED	STATUS	PORTS	NAMES
942a1aadfc7c	registry.gitlab.com/dalibo/postgresql_anonymizer	"docker-entrypoint.s..."	12 seconds ago	Up 11 seconds	0.0.0.0:6543->5432/tcp	angry_greider

2. Динамическая маскировка:

Для начала необходимо создать новую роль в системе, установить метку безопасности и включить маскировку:



```
1  -- Запускаю динамическую маскировку
2  SELECT anon.start_dynamic_masking();
3
4  -- Создаю новую роль
5  CREATE ROLE new_user LOGIN;
6
7  -- Устанавливаю метку безопасности для роли, чтобы включить маскировку
8  SECURITY LABEL FOR anon ON ROLE new_user IS 'MASKED';
9
10 -- Устанавливаю метку безопасности для колонки phone, чтобы скрыть все цифры, кроме первых 4
11 SECURITY LABEL FOR anon ON COLUMN client.phone
12 IS 'MASKED WITH FUNCTION anon.partial(phone, 4, '$$*****$$', 0)';
```

В результате:

	client_name character varying (100)	masked_phone text	email character varying (100)	address character varying (255)	date_of_birth date
1	Ivan Petrov	1981*****	ivan.petrov@example.ru	Lenina St, 5	1985-02-15
2	Olga Sidorova	8909*****	olga.sidorova@example.ru	Kirova St, 12	1990-06-25
3	Dmitry Ivanov	8903*****	dmitry.ivanov@example.ru	Gagarina St, 25	1988-08-10
4	Anna Kuznetsova	8910*****	anna.kuznetsova@example.ru	Pushkina St, 8	1983-07-01
5	Sergey Popov	8911*****	sergey.popov@example.ru	Mira St, 17	1983-07-01
6	Elena Smirnova	8909*****	elena.smirnova@example.ru	Zhukova St, 21	1991-09-19
7	Roman Kozlov	8909*****	roman.kozlov@example.ru	Victory St, 42	1995-04-29
8	Irina Fedorova	8905*****	irina.fedorova@example.ru	Volgogradsky Prospect, 25	1995-04-29
9	Mikhail Orlov	8901*****	mikhail.orlov@example.ru	Tverskaya St, 15	1987-01-10
10	Nina Voronova	8909*****	nina.voronova@example.ru	Nevsky Prospect, 19	1993-11-15
11	Yuri Zaitsev	8907*****	yuri.zaitsev@example.ru	Moscow St, 45	1984-10-25
12	Natalya Lebedeva	8901*****	natalya.lebedeva@example.ru	Liteyny Prospect, 20	1995-05-05
13	Andrey Grigoriev	8902*****	andrey.grigoriev@example.ru	Kutuzovsky Prospect, 14	1986-02-18
14	Ekaterina Sorokina	8904*****	ekaterina.sorokina@example.ru	Arbat St, 37	1986-02-18
15	Vladimir Sokolov	8901*****	vladimir.sokolov@example.ru	Bolshaya Morskaya St, 9	1990-06-17
16	Marina Pavlova	8908*****	marina.pavlova@example.ru	Frunzenskaya Embankment, 10	1988-09-25
17	Alexey Gusev	8909*****	alexey.gusev@example.ru	Rublevskoye Shosse, 58	1988-09-25
18	Oksana Krylova	8902*****	oksana.krylova@example.ru	Taganskaya St, 29	1991-09-15
19	Viktor Nikitin	8901*****	viktor.nikitin@example.ru	Sadovaya St, 70	1985-05-15
20	Anastasia Belova	8907*****	anastasia.belova@example.ru	Profsoyuznaya St, 8	1993-08-01

3. Materialized View с использованием Generalization:

1) При использовании anon:

Generalization в данном случае заменяет данные более широкими и менее точными значениями, **диапазонами**. **date_of_birth** – диапазон 10 лет



Query Query History

```

1  CREATE MATERIALIZED VIEW Client_Generalized_Date AS
2  SELECT
3      client_name,
4      email,
5      phone,
6      address,
7      anon.generalize_daterange(date_of_birth, 'decade') AS date_of_birth
8  FROM Client;
```

Итог:

	client_name character varying (100) 🔒	email character varying (100) 🔒	phone character varying (15) 🔒	address character varying (255) 🔒	date_of_birth daterange 🔒
1	Ivan Petrov	ivan.petrov@example.ru	19812345678	Lenina St, 5	[1980-01-01,1990-01-01)
2	Olga Sidorova	olga.sidorova@example.ru	89098765432	Kirova St, 12	[1990-01-01,2000-01-01)
3	Dmitry Ivanov	dmitry.ivanov@example.ru	89034567890	Gagarina St, 25	[1980-01-01,1990-01-01)
4	Anna Kuznetsova	anna.kuznetsova@example.ru	89101223344	Pushkina St, 8	[1980-01-01,1990-01-01)
5	Sergey Popov	sergey.popov@example.ru	89111223344	Mira St, 17	[1980-01-01,1990-01-01)
6	Elena Smirnova	elena.smirnova@example.ru	89091223345	Zhukova St, 21	[1990-01-01,2000-01-01)
7	Roman Kozlov	roman.kozlov@example.ru	89098765431	Victory St, 42	[1990-01-01,2000-01-01)
8	Irina Fedorova	irina.fedorova@example.ru	89055557788	Volgogradsky Prospect, 25	[1990-01-01,2000-01-01)
9	Mikhail Orlov	mikhail.orlov@example.ru	89012345678	Tverskaya St, 15	[1980-01-01,1990-01-01)
10	Nina Voronova	nina.voronova@example.ru	89091223346	Nevsky Prospect, 19	[1990-01-01,2000-01-01)
11	Yuri Zaitsev	yuri.zaitsev@example.ru	89071223347	Moscow St, 45	[1980-01-01,1990-01-01)
12	Natalya Lebedeva	natalya.lebedeva@example.ru	89012345678	Liteyny Prospect, 20	[1990-01-01,2000-01-01)
13	Andrey Grigoriev	andrey.grigoriev@example.ru	89023456789	Kutuzovsky Prospect, 14	[1980-01-01,1990-01-01)
14	Ekaterina Sorokina	ekaterina.sorokina@example.ru	89044556677	Arbat St, 37	[1980-01-01,1990-01-01)
15	Vladimir Sokolov	vladimir.sokolov@example.ru	89012345678	Bolshaya Morskaya St, 9	[1990-01-01,2000-01-01)
16	Marina Pavlova	marina.pavlova@example.ru	89087654321	Frunzenskaya Embankment, 10	[1980-01-01,1990-01-01)
17	Alexey Gusev	alexey.gusev@example.ru	89098765432	Rublevskoye Shosse, 58	[1980-01-01,1990-01-01)
18	Oksana Krylova	oksana.krylova@example.ru	89023456789	Taganskaya St, 29	[1990-01-01,2000-01-01)
19	Viktor Nikitin	viktor.nikitin@example.ru	89012345679	Sadovaya St, 70	[1980-01-01,1990-01-01)
20	Anastasia Belova	anastasia.belova@example.ru	89077889900	Profsoyuznaya St, 8	[1990-01-01,2000-01-01)

2) В данном случае хочу просто привести пример частичной маскировки к столбцам **email** и **phone**, которая скрывает часть настоящих данных.

No limit

[Query](#)
[Query History](#)

```

1  CREATE MATERIALIZED VIEW Client_Generalizeds AS
2  SELECT
3      client_name,
4      CONCAT(SUBSTRING(email, 1, 3), '****@****.***') AS email,
5      LEFT(phone, 6) || 'XXXXX' AS phone,
6      address,
7      date_of_birth
8  FROM Client;

```

Итог:

	client_name character varying (100) 🔒	email text 🔒	phone text 🔒	address character varying (255) 🔒	date_of_birth date 🔒
1	Ivan Petrov	iva****@****.***	198123XXXXXX	Lenina St, 5	1985-02-15
2	Olga Sidorova	olg****@****.***	890987XXXXXX	Kirova St, 12	1990-06-25
3	Dmitry Ivanov	dmi****@****.***	890345XXXXXX	Gagarina St, 25	1988-08-10
4	Anna Kuznetsova	ann****@****.***	891012XXXXXX	Pushkina St, 8	1983-07-01
5	Sergey Popov	ser****@****.***	891112XXXXXX	Mira St, 17	1983-07-01
6	Elena Smirnova	ele****@****.***	890912XXXXXX	Zhukova St, 21	1991-09-19
7	Roman Kozlov	rom****@****.***	890987XXXXXX	Victory St, 42	1995-04-29
8	Irina Fedorova	iri****@****.***	890555XXXXXX	Volgogradsky Prospect, 25	1995-04-29
9	Mikhail Orlov	mik****@****.***	890123XXXXXX	Tverskaya St, 15	1987-01-10
10	Nina Voronova	nin****@****.***	890912XXXXXX	Nevsky Prospect, 19	1993-11-15
11	Yuri Zaitsev	yur****@****.***	890712XXXXXX	Moscow St, 45	1984-10-25
12	Natalya Lebedeva	nat****@****.***	890123XXXXXX	Liteyny Prospect, 20	1995-05-05
13	Andrey Grigoriev	and****@****.***	890234XXXXXX	Kutuzovsky Prospect, 14	1986-02-18
14	Ekaterina Sorokina	eka****@****.***	890445XXXXXX	Arbat St, 37	1986-02-18
15	Vladimir Sokolov	vla****@****.***	890123XXXXXX	Bolshaya Morskaya St, 9	1990-06-17
16	Marina Pavlova	mar****@****.***	890876XXXXXX	Frunzenskaya Embankment, 10	1988-09-25
17	Alexey Gusev	ale****@****.***	890987XXXXXX	Rublevskoye Shosse, 58	1988-09-25
18	Oksana Krylova	oks****@****.***	890234XXXXXX	Taganskaya St, 29	1991-09-15
19	Viktor Nikitin	vik****@****.***	890123XXXXXX	Sadovaya St, 70	1985-05-15
20	Anastasia Belova	ana****@****.***	890778XXXXXX	Profsoyuznaya St, 8	1993-08-01

4. Materialized View с использованием Faking:

1) При использовании расширения anon:

Очень важное уточнение:

Чтобы использовать функции имитации, сначала необходимо `init()` расширение в вашей базе данных:

```
SELECT anon.init();
```

Функция `init()` импортирует набор случайных данных по умолчанию (идентификаторы, имена, города и т. д.).

▼

▼

No limit ▼

▼

E

Query

Query History

1 ▼

CREATE MATERIALIZED VIEW Client_Generalized_Email AS

SELECT

client_name,

anon.fake_email() as email,

phone,

address,

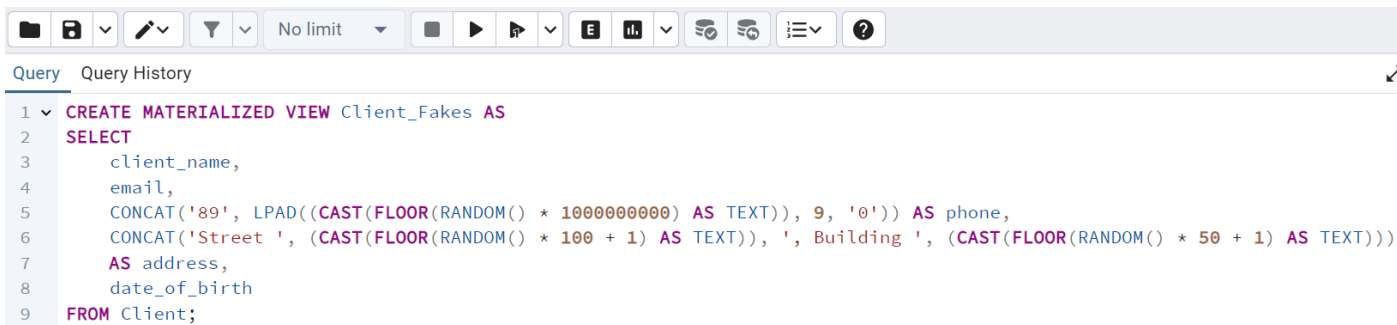
date_of_birth

FROM Client;






Итог:

	client_name character varying (100) 🔒	email text 🔒	phone character varying (15) 🔒	address character varying (255) 🔒	date_of_birth date 🔒
1	Ivan Petrov	glennjason@example.com	19812345678	Lenina St, 5	1985-02-15
2	Olga Sidorova	coltonmorgan@example.com	89098765432	Kirova St, 12	1990-06-25
3	Dmitry Ivanov	heidihernandez@example.net	89034567890	Gagarina St, 25	1988-08-10
4	Anna Kuznetsova	zsoto@example.com	89101223344	Pushkina St, 8	1983-07-01
5	Sergey Popov	castillotimothy@example.com	89111223344	Mira St, 17	1983-07-01
6	Elena Smirnova	anthonyreed@example.org	89091223345	Zhukova St, 21	1991-09-19
7	Roman Kozlov	edwardscarlos@example.com	89098765431	Victory St, 42	1995-04-29
8	Irina Fedorova	wesley47@example.org	89055557788	Volgogradsky Prospect, 25	1995-04-29
9	Mikhail Orlov	laura11@example.org	89012345678	Tverskaya St, 15	1987-01-10
10	Nina Voronova	wardtyler@example.org	89091223346	Nevsky Prospect, 19	1993-11-15
11	Yuri Zaitsev	gonzaleznicholas@example.net	89071223347	Moscow St, 45	1984-10-25
12	Natalya Lebedeva	gregory00@example.org	89012345678	Liteyny Prospect, 20	1995-05-05
13	Andrey Grigoriev	ricardo67@example.net	89023456789	Kutuzovsky Prospect, 14	1986-02-18
14	Ekaterina Sorokina	taylorgill@example.org	89044556677	Arbat St, 37	1986-02-18
15	Vladimir Sokolov	walterkristi@example.org	89012345678	Bolshaya Morskaya St, 9	1990-06-17
16	Marina Pavlova	boyerkrystal@example.com	89087654321	Frunzenskaya Embankment, 10	1988-09-25
17	Alexey Gusev	ramoscynthia@example.org	89098765432	Rublevskoye Shosse, 58	1988-09-25
18	Oksana Krylova	cweber@example.com	89023456789	Taganskaya St, 29	1991-09-15
19	Viktor Nikitin	hannahharrison@example.com	89012345679	Sadovaya St, 70	1985-05-15
20	Anastasia Belova	erin49@example.org	89077889900	Profsoyuznaya St, 8	1993-08-01

2) При использовании этого метода происходит следующее: **phone**: генерирую новый номер, начинающийся с '89', а остальные цифры случайным образом генерируются для реалистичного вида. **address**: подделанный адрес, состоящий из случайно сгенерированного названия улицы и номера здания.



Итог:

	client_name character varying (100) 	email character varying (100) 	phone text 	address text 	date_of_birth date 
1	Ivan Petrov	ivan.petrov@example.ru	89299670424	Street 6, Building 25	1985-02-15
2	Olga Sidorova	olga.sidorova@example.ru	89355871448	Street 68, Building 2	1990-06-25
3	Dmitry Ivanov	dmitry.ivanov@example.ru	89046711337	Street 85, Building 30	1988-08-10
4	Anna Kuznetsova	anna.kuznetsova@example.ru	89977759709	Street 28, Building 30	1983-07-01
5	Sergey Popov	sergey.popov@example.ru	89250491593	Street 45, Building 18	1983-07-01
6	Elena Smirnova	elena.smirnova@example.ru	89952470882	Street 96, Building 11	1991-09-19
7	Roman Kozlov	roman.kozlov@example.ru	89390298282	Street 37, Building 22	1995-04-29
8	Irina Fedorova	irina.fedorova@example.ru	89383603376	Street 19, Building 27	1995-04-29
9	Mikhail Orlov	mikhail.orlov@example.ru	89640549897	Street 2, Building 47	1987-01-10
10	Nina Voronova	nina.voronova@example.ru	89457676341	Street 79, Building 16	1993-11-15
11	Yuri Zaitsev	yuri.zaitsev@example.ru	89922379813	Street 6, Building 23	1984-10-25
12	Natalya Lebedeva	natalya.lebedeva@example.ru	89173752236	Street 10, Building 11	1995-05-05
13	Andrey Grigoriev	andrey.grigoriev@example.ru	89816632453	Street 83, Building 11	1986-02-18
14	Ekaterina Sorokina	ekaterina.sorokina@example.ru	89142881284	Street 14, Building 28	1986-02-18
15	Vladimir Sokolov	vladimir.sokolov@example.ru	89897166211	Street 81, Building 45	1990-06-17
16	Marina Pavlova	marina.pavlova@example.ru	89400173079	Street 38, Building 5	1988-09-25
17	Alexey Gusev	alexey.gusev@example.ru	89725038018	Street 18, Building 20	1988-09-25
18	Oksana Krylova	oksana.krylova@example.ru	89132276202	Street 59, Building 39	1991-09-15
19	Viktor Nikitin	viktor.nikitin@example.ru	89947556209	Street 76, Building 35	1985-05-15
20	Anastasia Belova	anastasia.belova@example.ru	89223982984	Street 38, Building 37	1993-08-01

5. Materialized View с использованием Noise:

При использовании anon:

Эта функция добавляет случайный шум к дате рождения в пределах ± 2 дней.


```
1 CREATE MATERIALIZED VIEW Client_With_Noise AS
2 SELECT
3     client_name,
4     email,
5     phone,
6     address,
7     anon.dnoise(date_of_birth, interval '2 days') AS date_of_birth
8 FROM Client;
```

Itor:

	client_name character varying (100)	email character varying (100)	phone character varying (15)	address character varying (255)	date_of_birth date
1	Ivan Petrov	ivan.petrov@example.ru	19812345678	Lenina St, 5	1985-02-13
2	Olga Sidorova	olga.sidorova@example.ru	89098765432	Kirova St, 12	1990-06-25
3	Dmitry Ivanov	dmitry.ivanov@example.ru	89034567890	Gagarina St, 25	1988-08-09
4	Anna Kuznetsova	anna.kuznetsova@example.ru	89101223344	Pushkina St, 8	1983-07-02
5	Sergey Popov	sergey.popov@example.ru	89111223344	Mira St, 17	1983-06-29
6	Elena Smirnova	elena.smirnova@example.ru	89091223345	Zhukova St, 21	1991-09-18
7	Roman Kozlov	roman.kozlov@example.ru	89098765431	Victory St, 42	1995-04-30
8	Irina Fedorova	irina.fedorova@example.ru	89055557788	Volgogradsky Prospect, 25	1995-04-27
9	Mikhail Orlov	mikhail.orlov@example.ru	89012345678	Tverskaya St, 15	1987-01-11
10	Nina Voronova	nina.voronova@example.ru	89091223346	Nevsky Prospect, 19	1993-11-13
11	Yuri Zaitsev	yuri.zaitsev@example.ru	89071223347	Moscow St, 45	1984-10-23
12	Natalya Lebedeva	natalya.lebedeva@example.ru	89012345678	Liteyny Prospect, 20	1995-05-03
13	Andrey Grigoriev	andrey.grigoriev@example.ru	89023456789	Kutuzovsky Prospect, 14	1986-02-18
14	Ekaterina Sorokina	ekaterina.sorokina@example.ru	89044556677	Arbat St, 37	1986-02-18
15	Vladimir Sokolov	vladimir.sokolov@example.ru	89012345678	Bolshaya Morskaya St, 9	1990-06-17
16	Marina Pavlova	marina.pavlova@example.ru	89087654321	Frunzenskaya Embankment, 10	1988-09-26
17	Alexey Gusev	alexey.gusev@example.ru	89098765432	Rublevskoye Shosse, 58	1988-09-25
18	Oksana Krylova	oksana.krylova@example.ru	89023456789	Taganskaya St, 29	1991-09-13
19	Viktor Nikitin	viktor.nikitin@example.ru	89012345679	Sadovaya St, 70	1985-05-13
20	Anastasia Belova	anastasia.belova@example.ru	89077889900	Profsoyuznaya St, 8	1993-08-02