

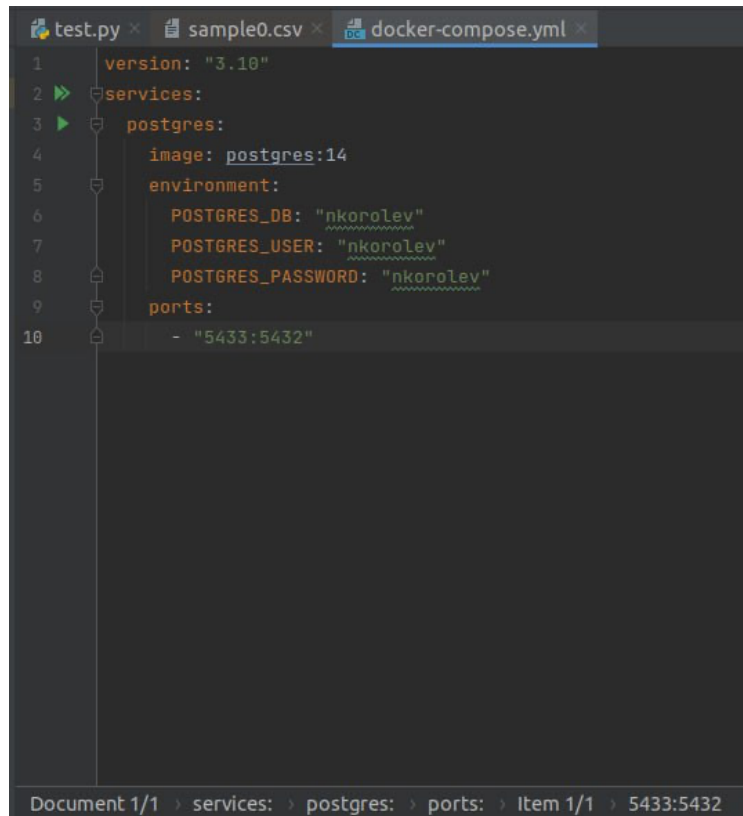
Отчет по лабораторной работе № 1

Рисунок 1 – Конфигурация файла docker-compose.yml

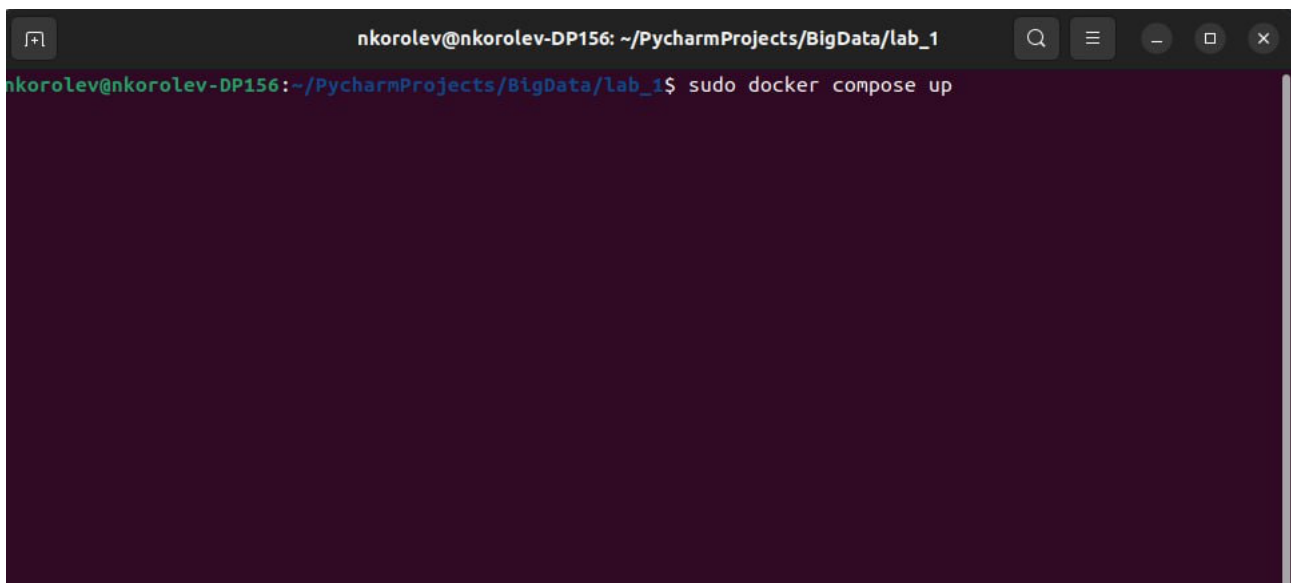


Рисунок 2 – «Поднятие» docker-compose с образом базы данных PostgreSQL

```
nkorolev@nkorolev-DP156: ~/PycharmProjects/BigData/lab_1
lab_1-postgres-1 | You can change this by editing pg_hba.conf or using the option -A, or
lab_1-postgres-1 | --auth-local and --auth-host, the next time you run initdb.
lab_1-postgres-1 | waiting for server to start....2023-03-09 19:57:16.649 UTC [48] LOG:  starting Postgre
SQL 14.7 (Debian 14.7-1.pgdg110+1) on x86_64-pc-linux-gnu, compiled by gcc (Debian 10.2.1-6) 10.2.1 202101
10, 64-bit
lab_1-postgres-1 | 2023-03-09 19:57:16.650 UTC [48] LOG:  listening on Unix socket "/var/run/postgresql/.
s.PGSQL.5432"
lab_1-postgres-1 | 2023-03-09 19:57:16.656 UTC [49] LOG:  database system was shut down at 2023-03-09 19:
57:16 UTC
lab_1-postgres-1 | 2023-03-09 19:57:16.660 UTC [48] LOG:  database system is ready to accept connections
lab_1-postgres-1 | done
lab_1-postgres-1 | server started
lab_1-postgres-1 | CREATE DATABASE
lab_1-postgres-1 |
lab_1-postgres-1 | /usr/local/bin/docker-entrypoint.sh: ignoring /docker-entrypoint-initdb.d/*
lab_1-postgres-1 | 2023-03-09 19:57:16.949 UTC [48] LOG:  received fast shutdown request
lab_1-postgres-1 | waiting for server to shut down....2023-03-09 19:57:16.950 UTC [48] LOG:  aborting any
active transactions
lab_1-postgres-1 | 2023-03-09 19:57:16.952 UTC [48] LOG:  background worker "logical replication launcher
" (PID 55) exited with exit code 1
lab_1-postgres-1 | 2023-03-09 19:57:16.952 UTC [50] LOG:  shutting down
lab_1-postgres-1 | 2023-03-09 19:57:16.964 UTC [48] LOG:  database system is shut down
lab_1-postgres-1 | done
lab_1-postgres-1 | server stopped
lab_1-postgres-1 |
lab_1-postgres-1 | PostgreSQL init process complete; ready for start up.
lab_1-postgres-1 |
lab_1-postgres-1 | 2023-03-09 19:57:17.070 UTC [1] LOG:  starting PostgreSQL 14.7 (Debian 14.7-1.pgdg110+
1) on x86_64-pc-linux-gnu, compiled by gcc (Debian 10.2.1-6) 10.2.1 20210110, 64-bit
lab_1-postgres-1 | 2023-03-09 19:57:17.070 UTC [1] LOG:  listening on IPv4 address "0.0.0.0", port 5432
lab_1-postgres-1 | 2023-03-09 19:57:17.070 UTC [1] LOG:  listening on IPv6 address ":::", port 5432
lab_1-postgres-1 | 2023-03-09 19:57:17.072 UTC [1] LOG:  listening on Unix socket "/var/run/postgresql/.s
.PGSQL.5432"
lab_1-postgres-1 | 2023-03-09 19:57:17.077 UTC [63] LOG:  database system was shut down at 2023-03-09 19:
57:16 UTC
lab_1-postgres-1 | 2023-03-09 19:57:17.083 UTC [1] LOG:  database system is ready to accept connections
```

Рисунок 3 – docker-compose с образом PostgreSQL запущен

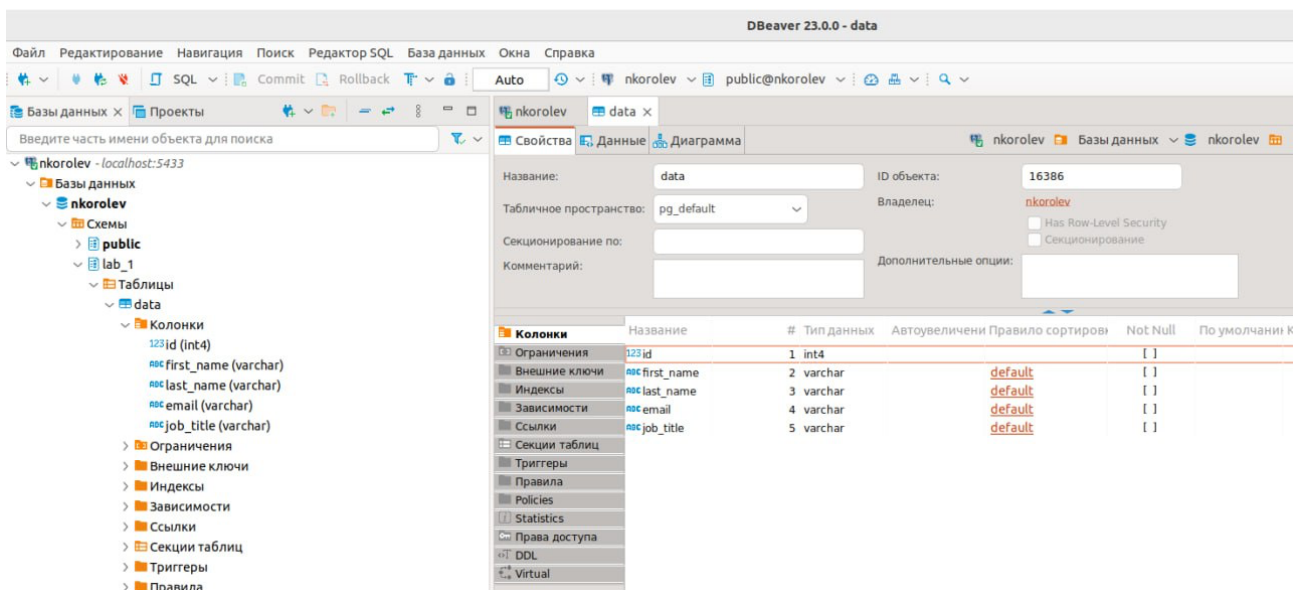


Рисунок 4 – В ПО для управления базами данных DBeaver в базе данных nkorolev создана схема lab_1 и таблица data с колонками id, first_name, last_name, email, job_title

```
test.py x sample0.csv x docker-compose.yml x
1 index,first_name,last_name,email,job_title
2 1,Shelby,Terrell,elijah57@example.net,Games developer
3 2,Phillip,Summers,bethany14@example.com,Phytotherapist
4 3,Kristine,Travis,bthompson@example.com,Homeopath
5 4,Yesenia,Martinez,kaitlinkaiser@example.com,Market researcher
6 5,Lori,Todd,buchananmanuel@example.net,Veterinary surgeon
7 6,Erin,Day,tconner@example.org,Waste management officer
8 7,Katherine,Buck,conniecowan@example.com,Intelligence analyst
9 8,Ricardo,Hinton,wyattbishop@example.com,Hydrogeologist
10 9,Dave,Farrell,nmccann@example.net,Lawyer
11 10,Isaiah,Downs,virginiaterrell@example.org,Engineer
12 11,Sheila,Ross,huangcathy@example.com,Advertising account executive
13 12,Stacy,Newton,rayloroy@example.org,Warden/ranger
14 13,Mandy,Blake,jefferynoble@example.org,"Scientist, clinical (histocompatibility and immunogenetics)"
15 14,Bridget,Nash,mercedes44@example.com,Social worker
16 15,Crystal,Farmer,pmiranda@example.org,Agricultural consultant
17 16,Thomas,Knight,braunpriscilla@example.net,Sport and exercise psychologist
18 17,Maurice,Rangel,sheenabanks@example.com,Secretary/administrator
19 18,Frank,Meadows,gbrewer@example.org,Audiological scientist
20 19,Alvin,Paul,gilbertdonaldson@example.com,Teacher
21 20,Jared,Mitchell,jcortez@example.com,Paediatric nurse
22 21,Jacqueline,Norton,carias@example.net,"Scientist, marine"
```

Рисунок 5 – Файл в формате .csv, данные из которого необходимо перенести в таблицу в базе данных

```
test.py x sample0.csv x docker-compose.yml x
1 import psycopg2
2
3
4 def input_values():
5     id = input('Write user_id')
6     first_name = input('Write first name')
7     last_name = input('Write last name')
8     email = input('Write email')
9     job_title = input('Write job title')
10    return id, first_name, last_name, email, job_title
11
12
13 if __name__ == "__main__":
14     conn = psycopg2.connect(host='localhost', port=5433, database='nkorolev', user='nkorolev', password='nkorolev')
15     cursor = conn.cursor()
16     while True:
17         print('1 - add row /n 2 - exit')
18         status_code = input()
19         if status_code == '1':
20             id, first_name, last_name, email, job_title = input_values()
21             sql = "INSERT INTO lab_1.data (id, first_name, last_name, email, job_title) VALUES ({}, " \
22                 "'{}', '{}', '{}', '{}');".format(id, first_name, last_name, email, job_title)
23             cursor.execute(sql)
24             elif status_code == '2':
25                 break
26             else:
27                 print('Error')
28     conn.commit()
29     conn.close()
```

Рисунок 6 – Python-скрипт для ручного занесения записей в таблицу в базе данных по одной строке за раз

```
Run: test x
/home/nkorolev/PycharmProjects/BigData/lab_1/venv/bin/python /home/nkorolev/PycharmProjects/BigData/lab_1/test.py
1 - add row /n 2 - exit
1
Write user_id1
Write first nameShelby
Write last nameTerrell
Write emailelijah57@example.net
Write job titleGames developer
1 - add row /n 2 - exit
2

Process finished with exit code 0
```

Рисунок 7 – Внесение данных в таблицу в ручном режиме (работает скрипт, представленный на рисунке 6)

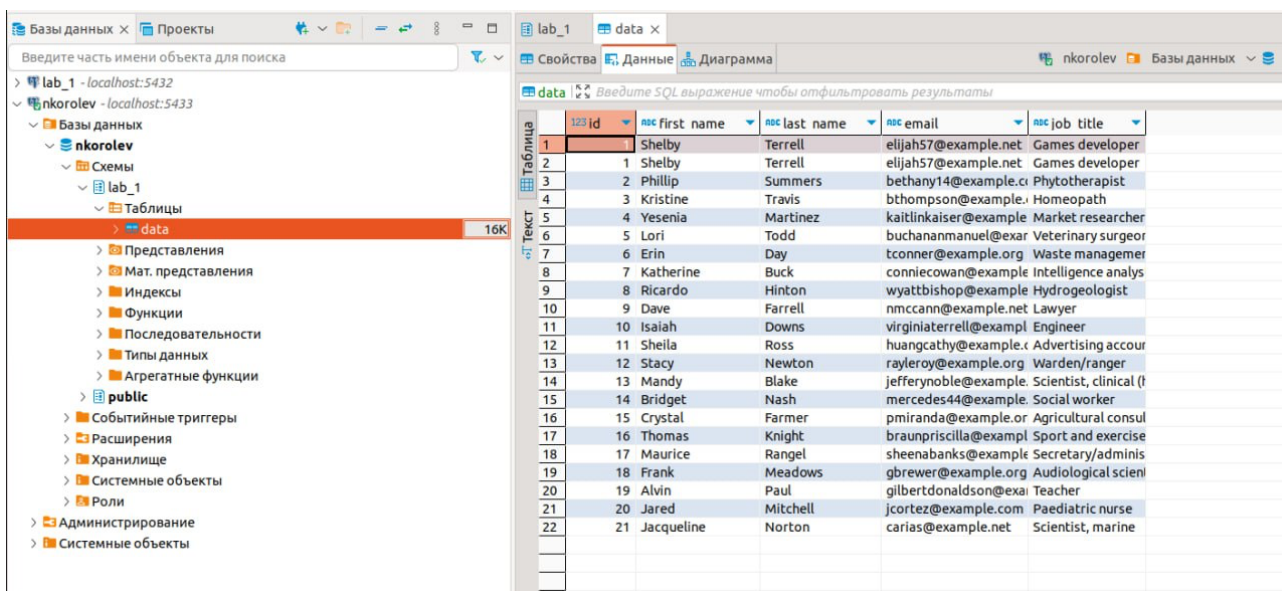
The screenshot shows a database management interface. On the left, a tree view displays the database structure: 'lab_1' (localhost:5432) contains a schema 'nkorolev', which has a table 'data'. The 'data' table has columns: 'id' (int4), 'first_name' (varchar), 'last_name' (varchar), 'email' (varchar), and 'job_title' (varchar). The main panel shows the 'data' table with one row of data.

id	first_name	last_name	email	job_title
1	Shelby	Terrell	elijah57@example.net	Games developer

Рисунок 8 – После внесения записи вручную с помощью Python-скрипта в таблице в базе данных при обновлении появляются внесенные нами данные


```
test.py x test_pd.py x sample0.csv x docker-compose.yml x
1 import psycopg2
2 import pandas as pd
3
4 conn = psycopg2.connect(host='localhost', port=5433, database='nkorolev', user='nkorolev', password='nkorolev')
5 cursor = conn.cursor()
6
7 if __name__ == "__main__":
8     data = pd.read_csv("sample0.csv")
9
10    for i in data.values:
11        cursor.execute("INSERT INTO lab_1.data (id, first_name, last_name, email, job_title) VALUES (%s, %s, %s, %s, \
12            %s);", [i[0], i[1], i[2], i[3], i[4]])
13    conn.commit()
14    conn.close()
```

Рисунок 9 – Для ускорения внесения данных из csv файла применим библиотеку для работы с данными Pandas и напишем новый скрипт, позволяющий автоматически переносить все данные в таблицу



	id	first name	last name	email	job title
1	1	Shelby	Terrell	elijah57@example.net	Games developer
2	1	Shelby	Terrell	elijah57@example.net	Games developer
3	2	Phillip	Summers	bethany14@example.c	Phytotherapist
4	3	Kristine	Travis	bthompson@example.	Homeopath
5	4	Yesenia	Martinez	kaitlinkaiser@example	Market researcher
6	5	Lori	Todd	buchananmanuel@exar	Veterinary surgeon
7	6	Erin	Day	tconner@example.org	Waste manager
8	7	Katherine	Buck	conniecowan@example	Intelligence analys
9	8	Ricardo	Hinton	wyattbishop@example	Hydrogeologist
10	9	Dave	Farrell	nmccann@example.net	Lawyer
11	10	Isaiah	Downs	virginiaterrell@exampl	Engineer
12	11	Sheila	Ross	huangcathy@example.c	Advertising account
13	12	Stacy	Newton	rayleroy@example.org	Warden/ranger
14	13	Mandy	Blake	jefferynoble@example	Scientist, clinical (
15	14	Bridget	Nash	mercedes44@example	Social worker
16	15	Crystal	Farmer	pmiranda@example.or	Agricultural consul
17	16	Thomas	Knight	braunpriscilla@exampl	Sport and exercise
18	17	Maurice	Rangel	sheenabanks@example	Secretary/adminis
19	18	Frank	Meadows	gbrewer@example.org	Audiological scienti
20	19	Alvin	Paul	gilbertdonaldson@exa	Teacher
21	20	Jared	Mitchell	jcortez@example.com	Paediatric nurse
22	21	Jacqueline	Norton	carias@example.net	Scientist, marine

Рисунок 10 – После успешной отработки скрипта, представленного на рисунке 10 видим, что все данные, содержащиеся в csv файле (рисунок 5) были перенесены в таблицу в базе данных