

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙ-
СКОЙ ФЕДЕРАЦИИ**

**Федеральное государственное автономное
образовательное учреждение высшего образования
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»**

Кафедра инфокоммуникаций

Анализ данных

Отчет по лабораторной работе №2.20(C Duckdb)

Тема: «Изучение Duckdb»

Выполнил студент группы

ИВТ-б-о-21-1

Горшков В.И. « »_____20__г.

Подпись студента_____

Работа защищена « »_____20__г.

Проверил доцент

Кафедры инфокоммуникаций, старший
преподаватель

Воронкин Р.А.

(подпись)

Ставрополь 2023

Цель работы: исследовать возможности Duckdb.

Ход работы:

1. Создал репозиторий в GitHub, дополнил правила в .gitignore для работы с IDE PyCharm с ЯП Python, выбрал лицензию MIT, клонировал его на компьютер и организовал в соответствии с моделью ветвления git-flow.

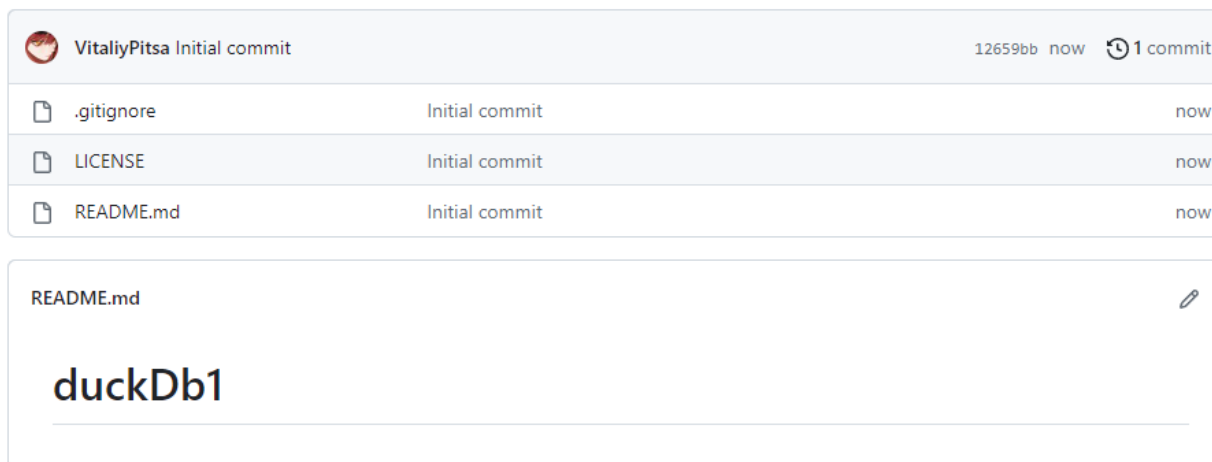


Рисунок 1.1 – Созданный репозиторий

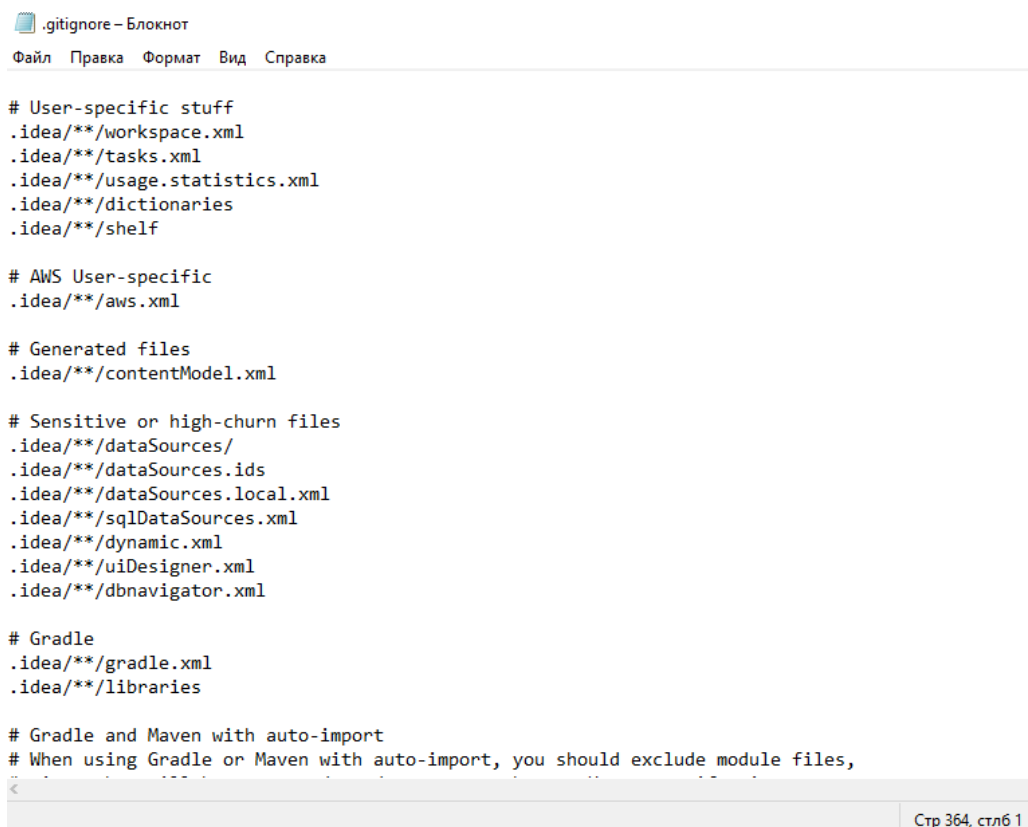


Рисунок 1.2 – Изменения в .gitignore

```

c:\Users\Admin\Desktop\git>git clone https://github.com/ItsMyLife1337/DuckDB1.git
Cloning into 'DuckDB1'...
remote: Enumerating objects: 4, done.
remote: Counting objects: 100% (4/4), done.
remote: Compressing objects: 100% (4/4), done.
remote: Total 4 (delta 0), reused 0 (delta 0), pack-reused 0
Receiving objects: 100% (4/4), done.

c:\Users\Admin\Desktop\git>cd /d c:\users\admin\desktop\git\Duckdb1

c:\Users\Admin\Desktop\git\DuckDB1>git flow init

Which branch should be used for bringing forth production releases?
- main
Branch name for production releases: [main]
Branch name for "next release" development: [develop]

How to name your supporting branch prefixes?
Feature branches? [feature/]
Bugfix branches? [bugfix/]
Release branches? [release/]
Hotfix branches? [hotfix/]
Support branches? [support/]
Version tag prefix? []
Hooks and filters directory? [C:/Users/Admin/Desktop/git/DuckDB1/.git/hooks]

c:\Users\Admin\Desktop\git\DuckDB1>git status

```

Рисунок 1.3 – Организация репозитория в соответствии с моделью ветвления git-flow

№2. Работа с duckdb.

Первое, что делаем необходимо сделать, это открываем консоль powershell и устанавливаем scoop. После установки scoop ставим пакет duckdb.

```

PS C:\Users\Admin> Set-ExecutionPolicy RemoteSigned -scope CurrentUser

Изменение политики выполнения
Политика выполнения защищает компьютер от ненадежных сценариев. Изменение политики выполнения может поставить под угрозу безопасность системы, как описано в разделе справки, вызываемом командой about_Execution_Policies и расположенном по адресу https://go.microsoft.com/fwlink/?LinkID=135170 . Вы хотите изменить политику выполнения?
[Y] Да - Y [A] Да для всех - A [N] Нет - N [L] Нет для всех - L [S] Приостановить - S [?] Справка
(значением по умолчанию является "N"):Y
PS C:\Users\Admin> iwr -useb get.scoop.sh | iex
Initializing...
Downloading ...
Creating shim...
Adding ~\scoop\shims to your path.
scoop was installed successfully!
Type 'scoop help' for instructions.
PS C:\Users\Admin> install duck.db
install : Имя 'install' не распознано как имя командлета, функции, файла сценария или выполняемой программы. Проверьте
правильность написания имени, а также наличие и правильность пути, после чего повторите попытку.
строка:1 знак:1
+ install duck.db
~ ~~~~~
+ CategoryInfo          : ObjectNotFound: (install:String) [], ParentContainsErrorRecordException
+ FullyQualifiedErrorId : CommandNotFoundException

PS C:\Users\Admin> scoop install duck.db
couldn't find manifest for 'duck.db'
PS C:\Users\Admin> scoop install duckdb
Installing 'duckdb' (0.8.0) [64bit] from main bucket
duckdb_cli-windows-amd64.zip (8,5 MB) [=====] 100%
Checking hash of duckdb_cli-windows-amd64.zip ... ok.
Extracting duckdb_cli-windows-amd64.zip ... done.
Linking ~\scoop\apps\duckdb\current => ~\scoop\apps\duckdb\0.8.0
Creating shim for 'duckdb'.
'duckdb' (0.8.0) was installed successfully!
PS C:\Users\Admin>

```

Рисунок 2 – Установка duckdb

```

D CREATE TABLE table1 AS SELECT * FROM read_csv_auto('c:\users\admin\downloads\covid19.csv');
D SELECT Country, CasesLast7d FROM table1 where CasesLast7d between 2500 and 3000;

```

Country varchar	CasesLast7d int64
Colombia	2512

```

D

```

Рисунок 3 – Запрос №1

```

D COPY (SELECT Country, CasesLast7d FROM table1 where CasesLast7d between 2500 and 3000) TO 'zapros1.csv' (HEADER, DELIMITER ',',');
D COPY (SELECT Country, CasesLast7d FROM table1 where CasesLast7d between 2500 and 3000) TO 'zapros1.json';
D

```

Рисунок 4 – Сохранение запроса в формате CSV и JSON

```

D SELECT Country, DeathsLast7d from table1 where DeathsLast7d == 0 limit 2;

```

Country varchar	DeathsLast7d int64
Falkland Islands	0
French Guiana	0

```

D COPY (SELECT Country, DeathsLast7d from table1 where DeathsLast7d == 0 limit 2) TO 'zapros2.csv' (HEADER, DELIMITER ',',');
D COPY (SELECT Country, DeathsLast7d from table1 where DeathsLast7d == 0 limit 2) TO 'zapros2.json';
D

```

Рисунок 5 – Запрос №2 и его сохранение

```

D SELECT Country, Population from table1 order by Population asc limit 10;

```

Country varchar	Population int64
Falkland Islands	3655
French Guiana	312076
Suriname	595733
Guyana	793124
Uruguay	3494576
Paraguay	7284217
Bolivia	11948668
Ecuador	18106806
Chile	19400316
Venezuela	28296430
10 rows 2 columns	

```

D COPY (SELECT Country, Population from table1 order by Population asc limit 10) TO 'zapros3.csv' (HEADER, DELIMITER ',',');
D COPY (SELECT Country, Population from table1 order by Population asc limit 10) TO 'zapros3.json';
D

```

Рисунок 6 – Запрос №3 и его сохранение

```

D SELECT Country, CasesPreceding7d, Population from table1 where CasesPreceding7d between 4000 and 4999 order by Population desc limit 3;

```

Country varchar	CasesPreceding7d int64	Population int64
Colombia	4312	51821865
Peru	4901	33766958
Ecuador	4629	18106806

```

D COPY (SELECT Country, CasesPreceding7d, Population from table1 where CasesPreceding7d between 4000 and 4999 order by Population desc limit 3) TO 'zapros4.csv' (HEADER, DELIMITER ',',');
Error: Parser Error: syntax error at or near "asc"
LINE 1: ...ion from table1 where CasesPreceding7d asc between 4000 and 4999 order by Popu...
                                     ^
D COPY (SELECT Country, CasesPreceding7d, Population from table1 where CasesPreceding7d between 4000 and 4999 order by Population desc limit 3) TO 'zapros4.csv' (HEADER, DELIMITER ',',');
Error: Binder Error: Referenced column "Cases" not found in FROM clause!
Candidate bindings: "table1.CasesLast7d"
LINE 1: COPY (SELECT Country, CasesPreceding7d, Population from tabl...
                                     ^
D COPY (SELECT Country, CasesPreceding7d, Population from table1 where CasesPreceding7d between 4000 and 4999 order by Population desc limit 3) TO 'zapros4.csv' (HEADER, DELIMITER ',',');
D COPY (SELECT Country, CasesPreceding7d, Population from table1 where CasesPreceding7d between 4000 and 4999 order by Population desc limit 3) TO 'zapros4.json';
D

```

Рисунок 7 – Запрос №4 и его сохранение

Вывод: в результате выполнения лабораторной работы были исследованы на практике базовые возможности системы управления базами данных Duckdb.