**Міністерство освіти і науки України**

**Національний університет харчових технологій**

Кафедра інформаційних технологій, штучного

інтелекту та кібербезпеки

**Звіт**

із лабораторної роботи № **1**

з дисципліни «Технології обробки великих даних»

на тему: «Розгортання робочого середовища та завантаження даних. »

Виконав:

Студент групи КН-1-3М

Кучерявий М. В.

Перевірив:

асистент Струзік В.А.

**Київ — 2025**

**Завдання:**

Завантажити обраний набір даних в ClickHouse, а саме в базу даних default у вигляді нової таблиці з механізмом MergeTree.

CREATE TABLE default.heart\_attack\_prediction

(

`Country` String,

`CountryUUIDHash` UInt64 MATERIALIZED reinterpretAsUInt64(generateUUIDv4()),

`Age` Int64,

`Gender` FixedString(50),

`Cholesterol\_Level` Float64,

`Blood\_Pressure` Float64,

`Smoking\_History` FixedString(50),

`Alcohol\_Consumption` Float64,

`Physical\_Activity` FixedString(50),

`Obesity` FixedString(50),

`Diabetes` FixedString(50),

`Family\_History` FixedString(50),

`Stress\_Levels` Float64,

`Dietary\_Habits` FixedString(50),

`Heart\_Attack\_History` FixedString(50),

`Chest\_Pain` FixedString(50),

`Exercise\_Induced\_Angina` FixedString(50),

`Resting\_ECG` FixedString(50),

`Max\_Heart\_Rate\_Achieved` Int64,

`Thalassemia` FixedString(50),

`HDL\_Cholesterol` Float64,

`LDL\_Cholesterol` Float64,

`Triglycerides` Float64,

`Heart\_Disease\_Risk` FixedString(50),

`Medication\_Adherence` FixedString(50),

`Urbanization\_Level` FixedString(50),

`Air\_Pollution\_Exposure` Float64,

`Access\_To\_Healthcare` FixedString(50),

`Education\_Level` FixedString(50),

`Income\_Level` FixedString(50),

`Heart\_Attack\_Outcome` FixedString(50)

)

ENGINE = MergeTree

ORDER BY CountryUUIDHash

SAMPLE BY CountryUUIDHash

SETTINGS index\_granularity = 8192

Завантажимо данні за допомогою clickhouse-client прокинувши 9000 порт в віртуальній машині.   
clickhouse-client --host=127.0.0.1 --port=9000 --user=default --query="INSERT INTO heart\_attack\_prediction FORMAT CSVWithNames" < ./heart\_attack\_predictions.csv

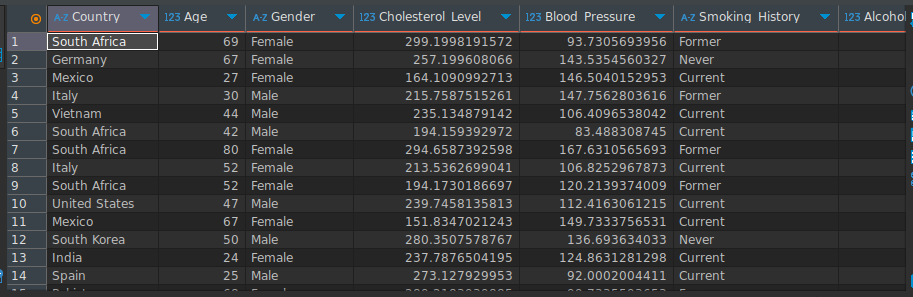
Виконаємо декілька тестових запитів:

Перші 20 рядків

**SELECT** \*

**FROM** heart\_attack\_prediction

**LIMIT** 20;



Кількісь записів

**SELECT** **count**()

**FROM** heart\_attack\_prediction

