**Міністерство освіти і науки України**

**Національний університет харчових технологій**

Кафедра інформаційних технологій, штучного

інтелекту та кібербезпеки

**Звіт**

із лабораторної роботи № **2**

з дисципліни «Технології обробки великих даних»

на тему: «Створення запитів мовою ClickHouse SQL.»

Виконав:

Студент групи КН-1-3М

Кучерявий М. В.

Перевірив:

асистент Струзік В.А.

**Київ — 2025**

**Завдання:**

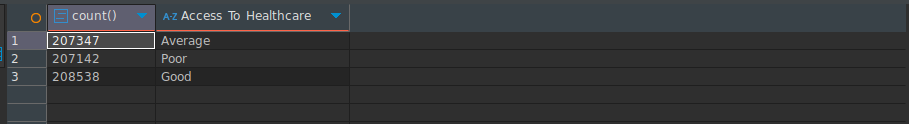
2. Створити 5 запитів на вибірку даних **SELECT** з використанням **GROUP BY**.

Кількісь записів згрупованих по яккості догляду за здоров’ям

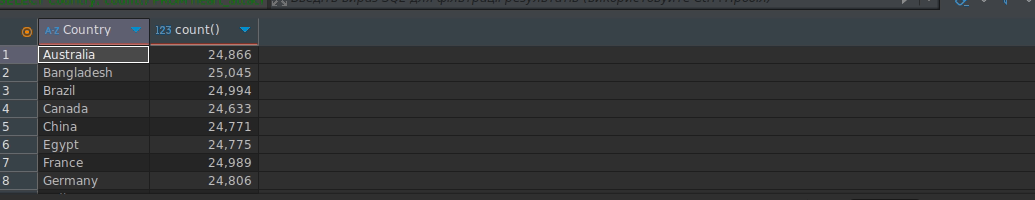
**SELECT** **count**(), Access\_To\_Healthcare

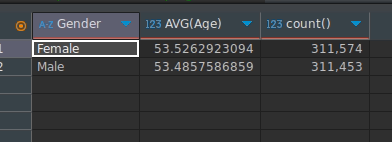
**FROM** heart\_attack\_prediction *h*

**GROUP** **BY** Access\_To\_Healthcare



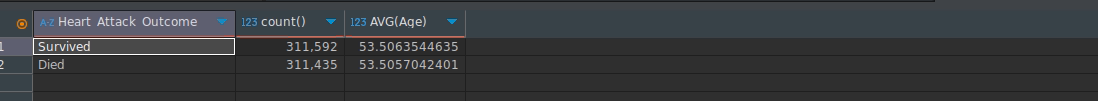
Кількість записів по країнам  
SELECT Country, count()   
FROM heart\_attack\_prediction hap   
GROUP BY hap.Country  
ORDER BY hap.Country ASC

  
Кількість записів відповідно до статі на середній вік кожної статі  
SELECT hap.Gender, AVG(hap.Age), count()  
FROM heart\_attack\_prediction hap   
GROUP BY hap.Gender

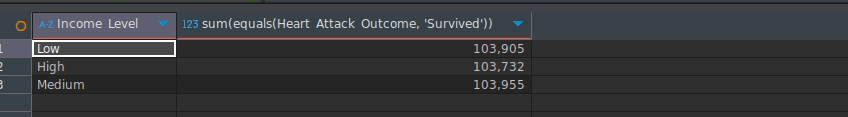


Кількість записів відповідно до результату серцевого нападу до середній вік кожного з результатів

SELECT hap.Heart\_Attack\_Outcome, count(), AVG(hap.Age)  
FROM heart\_attack\_prediction hap   
GROUP BY hap.Heart\_Attack\_Outcome



SELECT hap.Income\_Level, sum(hap.Heart\_Attack\_Outcome = 'Survived')  
FROM heart\_attack\_prediction hap   
GROUP BY hap.Income\_Level



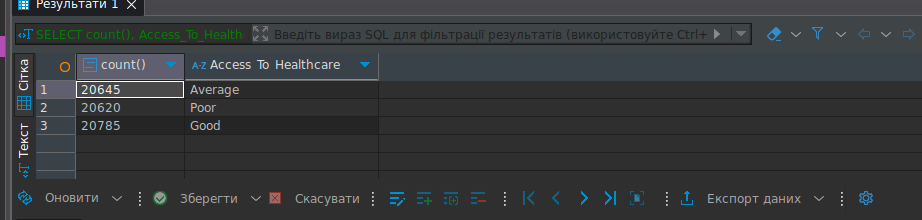
3. Запити, створені у попередньому пункті, доповнити використоювуючи секцю **SAMPLE**.

**SELECT** **count**(), Access\_To\_Healthcare

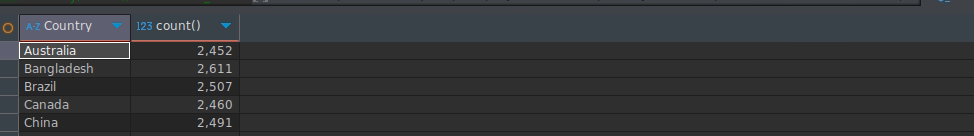
**FROM** heart\_attack\_prediction h

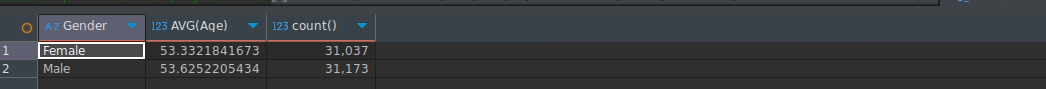
**SAMPLE** 0.1

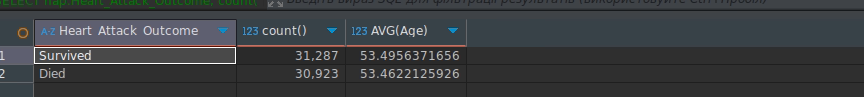
**GROUP** **BY** Access\_To\_Healthcare

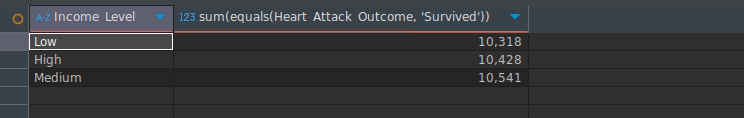


SELECT Country, count()   
FROM heart\_attack\_prediction hap   
SAMPLE 0.1  
GROUP BY hap.Country  
ORDER BY hap.Country ASC

  
  
  
SELECT hap.Gender, AVG(hap.Age), count()  
FROM heart\_attack\_prediction hap   
SAMPLE 0.1  
GROUP BY hap.Gender

  
  
SELECT hap.Heart\_Attack\_Outcome, count(), AVG(hap.Age)  
FROM heart\_attack\_prediction hap   
SAMPLE 0.1  
GROUP BY hap.Heart\_Attack\_Outcome

  
  
SELECT hap.Income\_Level, sum(hap.Heart\_Attack\_Outcome = 'Survived')  
FROM heart\_attack\_prediction hap   
SAMPLE 0.1  
GROUP BY hap.Income\_Level

  
Порівняти результати, що були отримані у п.2 та п.3.

У всіх запитах було використано SAMPLE 0.1 тобто 10% від усіх записів. Відповідно всі кількості в обрахунках змінились на x/10.

Також можна зазначити, що зменьшення кількость вибірки в конкретному датасеті не суттєво впливає на якісні результати. Наприклад: при повнів вибірці середній вік чоловіків / жінок – 53.52 / 53.48. При SAMPLE 0.1 чоловіків / жінок – 53.33 / 53.62. Средній вік летальних / нелетальних випадків 53.50 / 53.50. При SAMPLE 0.1 53.49 / 53.46. Різниця не є суттєвою. Також можна зазначити, що настільки близькі середні значення віку в різних вибірках свідчить про можливу штучність походження данних.