

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ

ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

Факультет прикладної математики

Кафедра прикладної математики

Звіт

із лабораторної роботи №2

із дисципліни «Введення до Баз Даних»

Виконав:

студент групи КМ-91

*Хом'як І. В.*

Перевірила:

*ст. викл.*

Бай Юлія

Київ-2021

## ЗАВДАННЯ

1. Визначитися з тематикою власної БД, обрати датасет на kaggle або data.gov.ua. Перевірити, що обраної тематики ще немає у таблиці. Записати тематику та посилання на датасет у стовпчики C, D.

Мій вибір впав на <https://www.kaggle.com/lasaljaywardena/global-food-prices-dataset>. Тематикою цього датасету є ціни на продукти на різних ринках світу з 1992 до 2017 року. В на кожному регіоні для кожного товару ціни вимірювали один раз.

2. Визначити сутності, атрибути та зв'язки інфологічної моделі БД. (Кількість сутностей та атрибутів повинна бути мінімально необхідною). Записати сутності та їх атрибути у стовпчик E.

Сутності:

- локація (location)
- ринок (market)
- ціни на товари (comdity\_price).

Атрибути сутностей:

- location:
  - ідентифікатор локації (location\_id) — Первинний Ключ
  - назва регіону (locality\_name)
  - назва країни (country\_name)
- market :
  - ідентифікатор ринку (market\_id) – Первинний Ключ
  - назва ринку (market\_name)
  - ідентифікатор локації де знаходиться ринок (location\_id)
  - тип ринку (market\_type), наприклад retail, wholesale, ...
- comodity\_price:
  - Тип товари (comodity) — Первинний Ключ

- ідентифікатор ринку Ринок (market\_id) — Первинний Ключ
- рік (year) — Первинний Ключ
- місяць (month) — Первинний Ключ
- ціна (price)
- валюта в якій ціна (currency)
- одиниця виміру товару (units)

Зв'язок: ціна за товар commodity\_price на ринку з ідентифікатором market\_id що знаходиться в локації location. Зв'язок сутності commodity\_price з сутністю market є зв'язком n:1, оскільки ми маємо багато товарів в різний час що продається на одному ринку. Зв'язок сутності market з сутністю location n:1, оскільки багато ринків можуть знаходитися в одному регіоні, але один ринок не може знаходитися в багатьох регіонах.

3. Побудувати ERD реляційної бази даних, використовуючи Draw.io (нотація для зв'язків між сутностями – «Crow's Foot»). Нормалізувати таблиці до 4NF включно. Оформити ERD бази даних у вигляді файлу erd.drawio. Назви таблиць БД після нормалізації записати у стовпчик F. (Увага! Кількість таблиць повинна бути не менше 3-х і не більше 10).

Нормалізація:

Оскільки

1) Таблиці location і market мають простий ключ. Ціна за товар відрізняється в залежності від типу товару, ринку, року і місяця. Одиниця виміру, ціна і валюта може мінятися з часом і бути різною на різних ринках може бути різною на різних ринках. Отже немає часткових залежностей.

2) Атомарність: кожен атрибут має лише одне значення, а не множину значень

3) Відсутні Транзитивні залежності

одиниця виміру, ціна чи валюта не визначають одна одну.

Оскільки різні ринки можуть мати однакову назву, назва ринку не визначає його тип чи місце знаходження.

Оскільки в різних країнах регіони можуть мати однакові назви. Ми не можемо визначи назву країни з назви регіону. Наприклад в цій базі даних в трьох країнах є “Western” (англ Західний) регіон.

4) Таблиці не містять нетривіальних багатозначних залежностей

5. Підготувати дані для заповнення таблиць, 5-10 рядків в кожну таблицю. Написати код для заповнення таблиць даними – populate.sql.

Файл populate.sql розміщено в проєкті [github.com/ihor7748/db\\_lab2\\_khomyak](https://github.com/ihor7748/db_lab2_khomyak) на Github.

6. Сформулювати 3 запити (файл query.sql), що відповідають заданій структурі:

- a) візуалізація – стовпчикова діаграма;
- b) візуалізація – кругова діаграма;
- c) візуалізація – графік залежності (в разі можна замінити іншою стовпчиковою діаграмою).

Запити:

- 1) ціни продуктів на ринку Львова.
- 2) кількість ринків по країнам в нашій базі даних.
- 3) середня ціна картоплі в Україні по рокам.

Короткий опис запитів записати у стовпчики G, H, I.

7. Написати код на Python, що реалізує підключення до PostgreSQL, виконує запити 6a, 6b, 6c та виводить результати їх виконання у консоль. Оформити даний код як файл main.py.

8. Реалізувати візуалізацію запитів 6a, 6b, 6c, побудувати dashboard та завантажити його на хостинг plotly. Зберегти файл як visualization.py. За бажанням можна виконати дане завдання за допомогою бібліотеки matplotlib без використання plotly.

9. Створити проєкт db\_lab2\_student01 на Github, в якому розташувати файли:

- erd.drawio
- create.sql
- populate.sql
- query.sql
- main.py
- visualization.py

та додати посилання на проект у стовпчик J. (student01 = Прізвище студента латиницею).