

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ
ІМЕНІ ІГОРЯ СІКОРСЬКОГО”**

Факультет прикладної математики
Кафедра програмного забезпечення комп’ютерних систем

КУРСОВА РОБОТА

з дисципліни “Бази даних”

спеціальність 121 – Програмна інженерія

на тему: Система аналізу цін споживчих товарів
(назва теми)

Студент групи КП-01

Северин Юрій Юрійович
(ПІБ)

(підпис)

Викладач
к.т.н, доцент кафедри СПіСКС

Радченко К.О.

(підпис)

Захищено з оцінкою _____

Київ – 2021

Зміст

| | |
|--|--|
| 1. Вступ | Ошибка! Закладка не определена. |
| 1.1. Найменування програми | 2 |
| 1.2. Призначення і галузь застосування | 3 |
| 2. Вимоги до програми | Ошибка! Закладка не определена. |
| 2.1. Вимоги до функціональних характеристик | Ошибка! Закладка не определена. |
| 2.2. Вимоги до надійності | Ошибка! Закладка не определена. |
| 2.2.1. Відмови через некоректні дії користувачів системи | Ошибка! Закладка не определена. |
| 3. Умови експлуатації | Ошибка! Закладка не определена. |
| 3.1. Кліматичні умови експлуатації | Ошибка! Закладка не определена. |
| 3.2. Вимоги до складу і параметрів технічних засобів | Ошибка! Закладка не определена. |
| 4. Вимоги до програмної документації | Ошибка! Закладка не определена. |
| 4.1. Попередній склад програмної документації | Ошибка! Закладка не определена. |
| 5. Техніко-економічні показники | Ошибка! Закладка не определена. |
| 5.1. Економічні переваги розробки | Ошибка! Закладка не определена. |
| 6. Стадії і етапи розробки | Ошибка! Закладка не определена. |
| 6.1. Стадії розробки | Ошибка! Закладка не определена. |
| 6.2. Етапи розробки | Ошибка! Закладка не определена. |
| 7. Порядок контролю і приймання | Ошибка! Закладка не определена. |
| 7.1. Види випробувань | Ошибка! Закладка не определена. |
| 7.2. Загальні вимоги до приймання роботи | Ошибка! Закладка не определена. |

1. Найменування та галузь застосування розробки

1.1. Найменування програми

Найменування програми: “Система аналізу цін споживчих товарів”.

1.2. Галузь застосування

Програма призначена для створення, керування вмістом бази даних, що містить такі дані:

1. Продовольчих товарів
2. Наявності у магазинах
3. Статистики цін

Система призначена для аналізу цін споживчих товарів: фільтрації, зберігання та аналізу поточних цін на товари за певною категорією з метою визначення певних цінових трендів, передбачення популярності товарів тощо.

2. Дата початку та закінчення курсової роботи

Початок: 20 жовтня 2021р

Кінцевий термін: 14 грудня 2021р

3. Мета розробки

Створити систему аналізу цін споживчих товарів, що буде фільтрувати, зберігати та аналізувати поточні ціни на товари за певною категорією з метою визначення певних цінових трендів, передбачення популярності товарів тощо.

4. Вимоги до програмного забезпечення;

Підсистема попередньої обробки даних містить у собі:

- **Засіб генерації даних:** розроблення утиліти для збору інформації про школи через відкриті ресурси у Інтернеті з використанням Scrapy.
- **Засоби фільтрації та валідації даних:** розроблення додаткового функціоналу у вищезазначеній утиліті задля корегування отриманих даних та переходу до їх подальшої обробки та структуризації.

База даних: PostgreSQL

Засоби аналізу даних:

- NumPy – бібліотека для роботи із великими масивами даних
- Matplotlib – бібліотека для візуалізації даних у вигляді 2D і 3D графіків
- Pandas – бібліотека, яка використовуватиметься як надбудова до NumPy для структуризації роботи із масивами даних.
- XGBoost – прогнозування даних

Задачі аналізу даних:

- Структурувати усі отримані дані із різних датасетів про школи, оцінки учнів та демографічні особливості
- Здійснити валідацію даних та прибрати зайву інформацію із усіх датасетів
- Об'єднати інформацію та заповнити пусті місця у даних
- Провести кореляцію по усіх даних
- У місцях, де найбільша кореляція, здійснити більш детальний аналіз
- Сформувати діаграми розсіювання, виокремити кластери та зробити висновки щодо отриманих результатів
- Створити прогноз на основі оцінок учнів різних шкіл за минулі роки

Передбачається створення механізму **реплікації** даних та створення звітної інформації.

1. Вимоги можуть **динамічно** змінюватися під час розробки проекту.

5. Обґрунтування вибору СУБД

PostgreSQL. У Постгреса є безліч можливостей. Створений з використанням об'єктно-реляційної моделі, він підтримує складні структури і широкий спектр вбудованих і обумовлених користувачем типів даних. Він забезпечує розширену ємність даних і заслужив довіру дбайливим ставленням до цілісності даних.

6. Вибір засобів розробки

Мова Python 3.6-3.8 та відповідні бібліотеки напрямку Data Science.

7. Етапи розробки

| | |
|------------|--|
| 19.10.2021 | Створення ТЗ |
| 28.10.2021 | Засоби генерації даних |
| 02.11.2021 | Засоби фільтрації та валідації даних |
| 04.11.2021 | Створення і підключення БД |
| 14.11.2021 | Засоби реплікації |
| 21.11.2021 | Засоби аналізу |
| 02.12.2021 | Засоби резервування |
| 10.12.2021 | Структурування проекту та документації |
| 14.12.2021 | Захист |