**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** **3**](#_Toc4967)

[1.1. Найменування програми 3](#_Toc4968)

[1.2. Призначення і галузь застосування 3](#_Toc4969)

[**2. Вимоги до програми** **4**](#_Toc4970)

[2.1. Вимоги до функціональних характеристик 4](#_Toc4971)

[2.2. Вимоги до надійності 5](#_Toc4972)

[2.2.1. Відмови через некоректні дії користувачів системи 5](#_Toc4973)

[**3. Умови експлуатації** **6**](#_Toc4974)

[3.1. Кліматичні умови експлуатації 6](#_Toc4975)

[3.2. Вимоги до складу і параметрів технічних засобів 6](#_Toc4976)

[**4. Вимоги до програмної документації** **6**](#_Toc4977)

[4.1. Попередній склад програмної документації 6](#_Toc4978)

[**5. Техніко-економічні показники** **6**](#_Toc4979)

[5.1. Економічні переваги розробки 6](#_Toc4980)

[**6. Стадії і етапи розробки** **7**](#_Toc4981)

[6.1. Стадії розробки 7](#_Toc4982)

[6.2. Етапи розробки 7](#_Toc4983)

[**7. Порядок контролю і приймання** **7**](#_Toc4984)

[7.1. Види випробувань 7](#_Toc4985)

[7.2. Загальні вимоги до приймання роботи 8](#_Toc4986)

# Найменування та галузь застосування розробки

## Найменування програми

Найменування програми: “Система аналізу цін споживчих товарів”.

## Галузь застосування

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

1. Продовоьчих 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.

# Дата початку та закінчення курсової роботи

Початок: 20 жовтня 2021р

Кінцевий термін: 14 грудня 2021р

# 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.

# 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
* Matpotlib – 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.

# 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.

# Вибір засобів розробки

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

# 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 |