**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ “КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО”**

Факультет прикладної математики

Кафедра програмного забезпечення комп’ютерних систем

**КУРСОВИЙ ПРОЕКТ**

з дисципліни “Бази даних”

спеціальність 121 – Програмна інженерія

на тему: **Моніторингова система кліматичних показників довкілля**

**Студент**

**групи** ​КП-71 Пилипенко Іван Олександрович \_\_\_\_\_\_\_\_\_ (ПІБ) (підпис)

**Викладач**

**к.т.н, доцент кафедри**

**СПіСКС** **Петрашенко А.В.**  \_\_\_\_\_\_\_\_\_

(підпис)

Захищено з оцінкою \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Київ – 2020

## **Анотація**

Метою розробки даного курсового проекту є набуття практичних навичок розробки сучасного програмного забезпечення, що взаємодіє з пост реляційними базами даних, а також здобуття навичок оформлення відповідного текстового, програмного та ілюстративного матеріалу у формі проектної документації. У результаті виконання курсового проекту було опановано навик розробляти програмне забезпечення для пост реляційних баз даних, володіння основами використання СУБД, а також інструментальними засобами аналізу великих обсягів даних.

Темою даного курсового проекту є створення моніторингової системи кліматичних показників довкілля. У документі викладена актуальність та проблематика аналізу великого обсягу даних, аналіз використаного інструментарію (опис мови програмування, використаних бібліотек та СУБД), описана структура бази даних, опис розробленого програмного забезпечення (загальний, опис модулів та основних алгоритмів роботи).

Результатами даного проекту стали діаграми та графіки, що зображають результати дослідження кліматичних показників.

## **Зміст**

[**Анотація**](#_9jyhzqr5uf10) **1**

[**Зміст**](#_9fz3c6l1rrjs) **2**

[**Вступ**](#_2q4b2t7er08g) **3**

[**Аналіз інструментарію для виконання курсового проект**](#_bbv7sji0a10g) **4**

[Аналіз СКБД](#_mgeezdf7zi82) 4

[Обґрунтування вибору мови програмування](#_6jkiqf6aq67n) 6

[Обґрунтування вибору бібліотек і фреймворків](#_9vu3crz2ezw8) 6

[**Структура бази даних**](#_1cxrfnlbyfoi) **7**

[**Аналіз функціонування засобів масштабування**](#_oizgttm5cizj) **8**

[Загальний алгоритм додавання нового кластеру шардингу, який складається з кількох серверів реплікації](#_yhfxaezfflf1) 9

[Загальний алгоритм додавання нового вузла реплікації серверів конфігурації](#_slv5ncz55s9) 10

[**Опис результатів аналізу предметної галузі**](#_30c5e5f98af0) **10**

[**Висновки**](#_7cwogsamug9e) **11**

[**Література**](#_ymmlg1w3cuxn) **11**

[**Додаток А**](#_jl5on67udgk) **12**

[**Додаток Б**](#_psx5t0c5f5s5) **16**

## **Вступ**

Під час виконання проекту було створено систему моніторингу погоди. Всі данні було отримано з сайту Open Weather, що пропонує можливість безкоштовного доступу до API з погодою.

Актуальність

На сьогоднішній день кожен користується сайтами з погодою, для того щоб знати що сьогодні вдягти, яка погода буде завтра або на вихідних. Сайти з погодою є досить корисними на даний час, хоча погода часто змінюється і прогнози погоди часто є хибними, нам все одно хочеться знати хочаб приблизну погоду.

Програма збирає дані з 15 різних міст України і зберігає данні про температуру, вологість, тиск та швидкість вітру. Також проект вираховує моду і медіану показників, та будує графік залежності температури від вологості.

Мета розробки:

Створення програмного забезпечення, що забезпечує роботу наведених далі пунктів:

1. Попередня обробка даних

• Засоби генерації даних. Підключення до API сайту з погодою для отримання даних для різних міст.

• Засоби фільтрації і валідації даних. Розроблено ПЗ, що фільтрує данні, отримані з API та залишає лише потрібні, також переводить різні системи вимірювань даних.

2. Основний модуль процесу. Ззібрання даних клімату різних міст України, побудова деяких графіків.

Дані для аналізу були запозичені із відкритого API за посиланням <https://openweathermap.org/current>.

## 

## **Аналіз інструментарію для виконання курсового проект**

### **Аналіз СКБД**

В процесі виконання цього курсового проекту перед нами встала потреба кешувати та зберігати дані про вакансії між запусками аналізатора. Кожного разу читати дані із CSV - файлів є дуже дорогою операцією, тож було прийняте рішення використати СКБД. В якості СКБД були розглянуті варіанти: PostgreSQL, MongoDB. З порівняльною характеристикою цих СУБД можна ознайомитися в таблиці 1.

*таблиця 1. Порівняльна характеристика СКБД*

|  |  |  |  |
| --- | --- | --- | --- |
| Критерій порівняння | Назва СКБД | | |
| MongoDB | PostgreSQL | CassandraDB |
| Реляційні дані | ні | так | так |
| Схема даних | динамічна | статична і динамічна | статична і динамічна |
| Підтримка ієрархічних даних | так | так | ні |
| Має відкритий вихідний код | так | так | так |
| Транзакції | ні | так | так |
| Атомарність операцій | всередині документа | по всій БД | всередині партиції |
| Мова запитів | JSON/JavaScript | SQL | CQL |
| Найлегший спосіб масштабування | горизонтальний | вертикальний | горизонтальний |
| Підтримка шардингів | так | так ( важка конфігурація) | так (може зберігати партиції на різних машинах) |
| Приклад використання | Великі дані (мільярди записів) з великою кількістю паралельних оновлень, де цілісність і узгодженість даних не потрібно. | Транзакційні і операційні програми, вигода яких в нормалізованому формі, об'єднаннях, обмеження даних і підтримки транзакцій. | Багато запитів на запис/читання у одиницю часу, до даних можна задіяти партиціювання за ключем, дані мають лише первинні індекси |
| Наявність бібліотек для мови програмування Node 12 | так | так | так |
| Підтримка реплікації | так, автоматичне переобрання головного процесу | За принципом master-slave | так, через партиціювання |
| Засіб збереження та відновлення даних | mongodump | pg\_dump | не має окремого доданка, виконується засобами CQL |
| Форма збереження даних | документы JSON | таблиця | таблиця |

За результатами порівнянні цих СУБД було прийнято рішення зупинитися на NoSQL рішеннях. Оскільки вона чудово поєднує в собі переваги неструктурованих баз даних. Крім цього класичним прикладом використання NoSQL СУБД є системи збору та аналізу даних, до яких можна застосувати індексування за первинними та вторинними ключами.

Ця база даних є об’єктно орієнтованою та дозволяє зберігати великі масиви неструктурованих даних. На відміну від SQL баз даних ми можемо зберігати дані у “сирому” об’єктному вигляді, який використовується програмою та є більш близьким за структурою до моделі даних, яку буде використовувати ПЗ написане з використанням мови програмування JavaScript. Це пришвидшить збір, збереження та отримання даних програмним забезпеченням. Оскільки MongoDB є представником NoSQL баз даних, вона не потребує жорсткої схеми даних, що дозволяє пришвидшити процес розробки та зробити його більш гнучким. Окрім цього дана СУБД підтримує горизонтальне масштабування за допомогою шардингу з метою зменшення навантаження на кожен окремий вузол шляхом розподілення навантаження між ними всіма.

Нижче наведено перелік основних переваг:

* Динамічна схема
* Швидкість запису у колекцію
* Швидкість читання із колекції

### 

### 

### **Обґрунтування вибору мови програмування**

Мовою програмування для ПЗ було обрано JavaScript, а саме Node js.Оскільки це динамічна, [об'єктно-орієнтована](https://uk.wikipedia.org/wiki/%D0%9E%D0%B1%27%D1%94%D0%BA%D1%82%D0%BD%D0%BE-%D0%BE%D1%80%D1%96%D1%94%D0%BD%D1%82%D0%BE%D0%B2%D0%B0%D0%BD%D0%B5_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F) [прототипна](https://uk.wikipedia.org/wiki/%D0%9F%D1%80%D0%BE%D1%82%D0%BE%D1%82%D0%B8%D0%BF%D0%BD%D0%B5_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F" \o "Прототипне програмування) [мова програмування](https://uk.wikipedia.org/wiki/%D0%9C%D0%BE%D0%B2%D0%B0_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F). Вона дозволяє створити гнучку систему, що спрошує розробку та подальше розширення програми.

### 

### 

### **Обґрунтування вибору бібліотек і фреймворків**

Використані бібліотеки:

Google Charts - це багатофункціональний набір інструментів для візуалізації даних. За допомогою нього можна відносно легко будувати графіки та діаграми на сайті.

mongoose — це бібліотека JavaScript, яка створена для взаємодії з базами даних включаючи MongoDB

simple-statistics — це бібліотека JavaScript, яка реалізує статистичні методи. Вона допомагає програмістам використовувати силу статистики, а статистики розуміють код. Бібліотека вичерпно задокументована, написана у простому та доброзичливому стилі та ретельно перевірена

Express — програмний каркас розробки серверної частини веб-застосунків для Node.js, реалізований як вільне і відкрите програмне забезпечення під ліцензією MIT. Він спроектований для створення веб-застосунків і API. Де-факто є стандартним каркасом для Node.js. Автор фреймворка, TJ Holowaychuk, описує його як створений на основі написаного на мові Ruby каркаса Sinatra, маючи на увазі, що він мінімалістичний, але має велику кількість плагінів, що підключаються.

## 

## **Структура бази даних**

База даних складається з однієї колекції, в якій зберігаються відомості міста з кліматичними показниками. Загальна структура документа в базі даних приведена у таблиці 2.

*таблиця 2. Опис властивостей документа у базі даних*

|  |  |  |
| --- | --- | --- |
| Назва властивості | Тип | Опис |
| \_id | ObjectId | Ідентифікатор запису |
| name | String | Назва міста |
| date | String | Дата виміру |
| temperature | String | Температура |
| pressure | String | Тиск |
| humidity | String | Вологість |
| windSpeed | String | Швидкість вітру |

**Опис результатів аналізу предметної галузі**

В результаті виконання курсового проекту було проаналізовано кліматичні дані в різних містах України. Було виконано наступне:

* Проаналізована залежність температури повітря від вологості;
* Складена таблиця міст України з показниками атмосферних параметрів;
* Знайдені мода та медіана атмосферних параметрів;
* Розроблений графічний інтерфейс для візуалізації отриманих даних;
* Додана функція експорту та імпорту бази даних.

## 

## **Висновки**

В Ході роботи було отримано практичні навички обробки даних та роботи з данними за допомогою мови програмування JavaScript та MongoDB.

Також було отримано нові навички роботи з певними бібліотеками, які спрощують та покращують роботу з JsavaScript. Було складено таблиці порівняння баз данних. На основі цих даних було обрано в якості СУБД NoSQL рішення MongoDB.

На основі зібраних даних було побудовано таблиці та графіки. Дані аналізу приведені у Додатку А.

В ході виконання даного курсового проекту було набуто практичних навичок розробки програмного забезпечення, що взаємодіє з NoSQL базою даних, а також були покращені навички оформлення відповідного текстового, програмного та ілюстративного матеріалу у формі проектної документації. Також було набуто навиків в роботі з бібліотеками для побудови графічної частини роботи.

## **Література**

1. PostgreSQL vs MongoDB<https://habr.com/post/272735/>;
2. Визуализация с Google Chart Tools API <https://habr.com/ru/post/304296/>
3. Mongodb <https://docs.mongodb.com/>
4. JavaScript <https://www.jetbrains.com/help/phpstorm/viewing-javascript-reference.html?gclid=Cj0KCQjwirz3BRD_ARIsAImf7LOHP8rqF-M7XZyl8hxRoDFd0Ta74D6cXVM3tT1jlHxOe3NAIa39JRwaAvfoEALw_wcB>

## 

## **Додаток А**

**

Рис.1 – таблиця з кліматичними даними

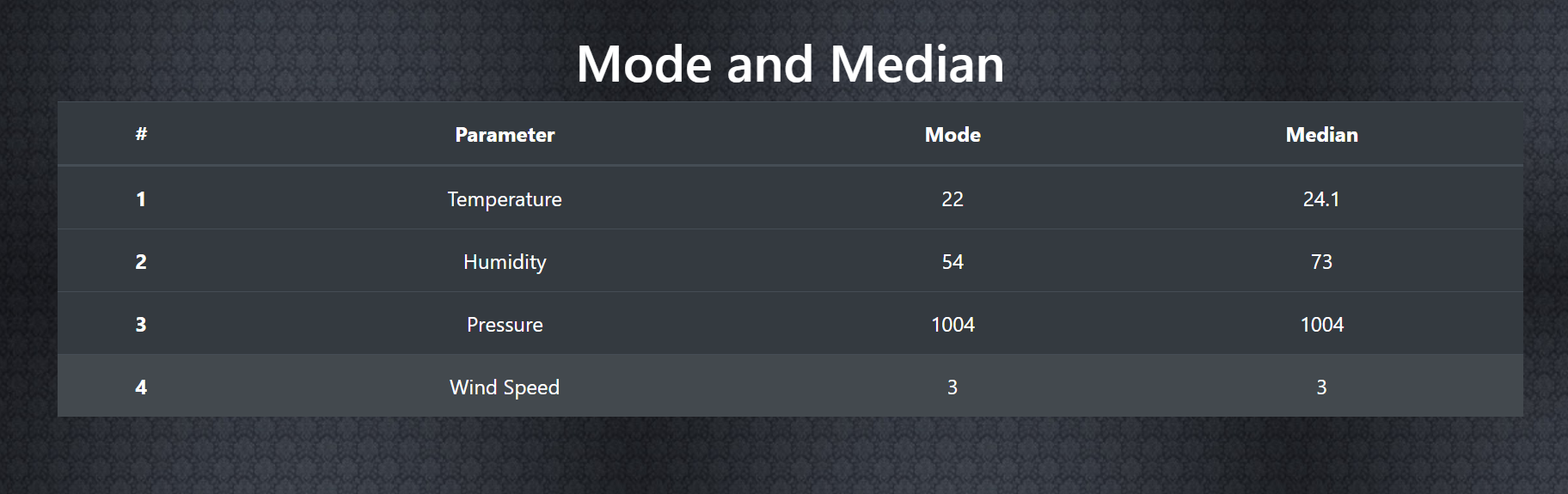
\

Рис.2 – мода та медіана показників

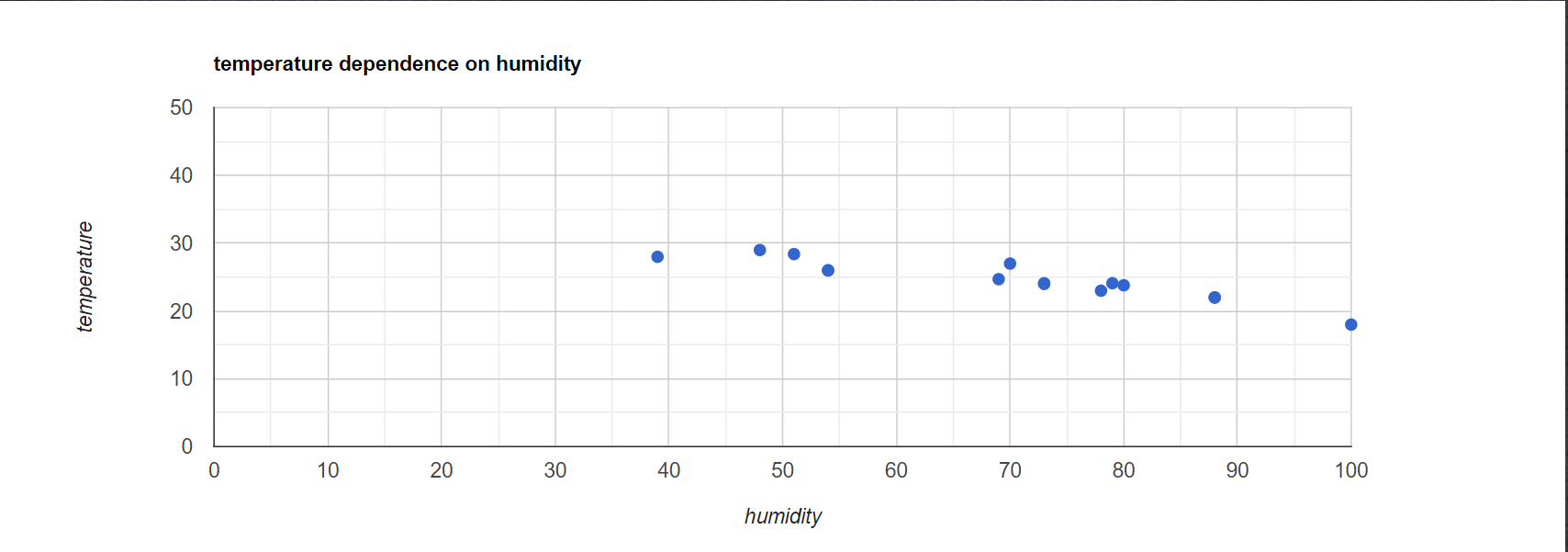


Рис.3 – залежність вологості від температури