

НУБІП України

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ

Факультет інформаційних технологій

НУБІП України

УДК 004.94:37.091.212.2

«ПОГОДЖЕНО»

«ДОПУСКАЄТЬСЯ ДО ЗАХИСТУ»

Декан факультету

Завідувач кафедри комп'ютерних наук

НУБІП України

інформаційних технологій

Глазунова О.І., д.п.н., професор

Голуб Б.Д., к.т.н., доцент

_____ 2021 р.

«30» листопада 2021 р.

НУБІП України

МАГІСТЕРСЬКА КВАЛІФІКАЦІЙНА РОБОТА

на тему Інформаційно-аналітична система показників вступу на ІІ-спеціальності та метрик відкритих електронних ресурсів ЗВО

Спеціальність 122 Комп'ютерні науки

(код + назва)

Освітня програма Інформаційні управляючі системи і технології

(назва)

Орієнтація освітньої програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Гарант освітньої програми «Інформаційні управляючі системи і технології»

доктор техн. н., професор

Бондаренко В.Є.

(науковий ступінь та вчене звання)

(підпис)


(ПІБ)

НУБІП України

Керівник магістерської кваліфікаційної роботи

канд. техн. наук, доцент

(науковий ступінь та вчене звання)


(підпис)

Ткаченко О.М.

(ПІБ)

Виконав


(підпис)

Коломієць В.А.

(ПІБ студента)

НУБІП України

КИЇВ-2021

НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ
І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ
Факультет інформаційних технологій

ЗАТВЕРДЖУЮ

Завідувач кафедри комп'ютерних наук

К.Т.Н., доцент

Голуб Б.Л.

(науковий ступінь, вчене звання) (підпис) (ПДБ)

«29» жовтня 2020 року

ЗАВДАННЯ

ДО ВИКОНАННЯ МАГІСТЕРСЬКОЇ КВАЛІФІКАЦІЙНОЇ РОБОТИ СТУДЕНТУ

Коломієць Віктору Анатолійовичу

(прізвище, ім'я, по батькові)

Спеціальність 122 «Комп'ютерні науки»

Освітня програма Інформаційні управляючі системи і технології

Орієнтація освітньої програми освітньо-професійна

Тема магістерської кваліфікаційної роботи Інформаційно-аналітична система показників вступу на ІТ-спеціальності та метрик відкритих електронних ресурсів ЗВО

затверджена наказом ректора НУБіП України від «29» жовтня 2020 р. №1634 «С»

Термін подання завершеної роботи на кафедру 30 листопада 2021 р.

Вихідні дані до магістерської кваліфікаційної роботи рейтинг ЗВО на основі метрик статистика вступних кампаній за період 2018-2021 рр.

Перелік питань, що підлягають дослідженню:

№ з/п	Питання, що підлягає дослідженню	Строк виконання	Примітка
1.	Системний аналіз предметної області	01.09.2021	
2.	Моделювання системи	10.09.2021	
3.	Розробка системи	17.09.2021	
4.	Результати дослідження	29.10.2021	
5.	Попередній захист	30.11.2021	
6.	Захист	14.12.2021	

Дата видачі завдання «29» жовтня 2020 р.

Керівник магістерської кваліфікаційної роботи

ініціали)

Завдання прийняв до виконання

(підпис)

Коломієць В.А.

(прізвище та ініціали студента)

ЗМІСТ		
НУБІП України		
ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ		5
ВСТУП		6
1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ		10
1.1	10	
1.2	12	
1.3	18	
2 МОДЕЛЮВАННЯ СИСТЕМИ		20
2.1 Моделювання предметної області		20
2.1.1	Діаграма прецедентів.	21
2.1.2	Діаграма розгортання.	22
2.1.3	Діаграма послідовності.	23
2.1.4	Діаграма діяльності.	24
2.2 Опис вузлів, з яких збирається інформація		26
2.3 OLAP		29
2.4 SSIS, SSAS та SSRS		32
2.5 Використані технології та програмні засоби при обробці даних та управлінні сховищем даних		34
2.5.1	Мова програмування Python.	34
2.5.2	Microsoft SQL Server.	34
2.5.3	Microsoft SQL Server Management Studio.	34
2.5.4	Microsoft Excel.	35
2.5.5	Microsoft Visual Studio.	35
2.5.6	Microsoft Visual Studio Code.	35
2.6 Використані технології та програмні засоби при створенні інформаційно-аналітичної системи		35
2.6.1	Мова програмування JavaScript.	36
2.6.2	API.	36
2.6.3	Node.JS.	36
2.6.4	Front-end framework Vue.JS.	37
2.6.5	Flexmonster – JavaScript Pivot Table & Charts Component.	37
НУБІП України		

3 РОЗРОБКА СИСТЕМИ	38
3.1 Сховище даних	38
3.2 Обробка даних з ресурсів	39
3.3 Розробка інформаційно-аналітичної системи «клієнт-сервер»	43
3.3.1 Серверна частина.	43
3.3.2 Клієнтська частина.	45
4 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ	47
4.1 Аналіз отриманих КТЕ	47
4.2 Аналіз отриманих графічних звітів	47
4.3 Аналіз показників кореляції	51
ВИСНОВКИ	53
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	55
Додаток А	61
Додаток Б	62
Додаток В	64

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

API (Application Programming Interface) – прикладний програмний інтерфейс.

BI (Business Intelligence) – технологія обробки і аналізу даних.

CSV (Comma-separated values) – формат файлу з даними, які розділено комами.

IDE (Integrated Development Environment) – інтегроване середовище для комплексного рішення задач та його обробки.

MS SQL (Microsoft SQL Server) – сервер баз даних від Microsoft.

OLAP (Online Analysis Process) – методи аналізу в реальному часі.

OLTP (Online Transaction Process) – методи транзакцій в реальному часі.

SQL (Structured Query Language) – мова структурованих запитів.

SSAS (SQL Server Analysis Service) – модуль аналізу в BI SQL Server.

SSIS (SQL Server Integration Service) – модуль інтеграції в BI SQL Server.

SSMS (SQL Server Management Studio) – середовище для керування базами даних/сховищами даних.

SSRS (SQL Server Reporting Service) – модуль звітів в BI SQL Server.

БД – база даних.

ЄДЕБО – Єдина державна електронна база з питань освіти.

ЗВО – заклад вищої освіти.

ККЕ – ключовий показник ефективності.

СД – сховище даних.

СУБД – система управління базою даних.

НУБІП України

ВСТУП

У сучасному світі в еру розвитку інформаційних технологій якісний інтернет ресурс є надзвичайно важливим для закладів вищої освіти (ЗВО). У

ньому зібрано всю інформацію про ЗВО, його складові, освітні програми, навчальні матеріали, дані про співробітників та інше. Проте, наскільки це впливає на кількість заяв абітурієнтів? Завдяки розробленій системі можна

з'ясувати: чи є залежність між кількістю поданих заяв абітурієнтами на спеціальність 122 «Комп'ютерні науки» та якістю інтернет ресурсів ЗВО, що дозволить оптимізувати витрати маркетингових ресурсів ЗВО. Таким чином, обрана тема є актуальною.

Об'єкт дослідження – вступна кампанія ЗВО.

Предмет дослідження – інформаційно-аналітична система показників вступу.

Методи дослідження: монографічний, статистичний, аналітичний, вебметричний аналіз. Для формування та обробки набору даних використовується мова програмування Python. Для зберігання даних використовується СД в MS SQL Server. Для керування сховищем даних використовується СУБД SSMS. Для проведення аналізу даних та перевірки гіпотези, розрахунку КПЕ використовується технологія OLAP та засоби BI (такі як SASS, SSIS та SSRS), front-end framework Vue.JS та Node.JS.

Метою наукового дослідження є визначення кореляції між кількістю поданих заяв абітурієнтами та структурами відкритих електронних ресурсів ЗВО за допомогою створеної інформаційно-аналітичної системи та метрик. Задача аналізу досить складна, оскільки не існує єдиної структури вебконтенту для різних ЗВО, відповідно, практично неможливо визначити, які саме з індексованих документів ресурсом вебметрик можна вважати такими, які переглядаються потенційними абітурієнтами та їх батьками.

Для досягнення цієї мети поставлено такі завдання:

- проаналізувати існуючі дослідження на дану тематику;

НУВБІП України

- зібрати інформацію щодо кількості поданих заяв на обрану галузь ЗВО за допомогою ЄДЕБО;
- зібрати інформацію щодо рейтингів ЗВО та їхніх вебметрик з ресурсу World Ranking Web of Universities;

НУВБІП України

- об'єднати зібрану інформацію за назвою ЗВО;
- обробити отриману інформацію і перетворити на зручний для використання формат даних;
- створити та заповнити сховище даних;
- створити проекти з використанням засобів ВІ;

НУВБІП України

- розробити інформаційно-аналітичну систему;
- побудувати графіки, які відображатимуть різні аспекти об'єкту дослідження: загальну динаміку абітурієнтів, середній бал вступників, відображення кількості поданих заяв абітурієнтами та загального

НУВБІП України

- значення метрики для ЗВО;
- побудувати візуальне представлення кореляції між вебметриками та кількістю поданих заяв абітурієнтами, обрахувати загальне значення кореляції;

НУВБІП України

- провести аналіз отриманих даних;
- зробити висновки щодо залежності структур інтернет представництв ЗВО та кількістю поданих заяв абітурієнтами, надати рекомендації.

В якості даних, що підлягають аналізу, використано статистики з ресурсів ЄДЕБО [1] та Ranking web of universities [2]. Останній ресурс використовує 4 метрики для формування рейтингів ЗВО. Ці дані об'єднані в один файл за назвами ЗВО та оброблені для можливості імпортування. Створене сховище даних наповнене даними зі даного файлу.

Наукова новизна: вперше було розроблено інформаційно-аналітичну систему, з допомогою якої досліджено залежність між вебметриками й кількістю поданих заяв абітурієнтами на IT спеціальність 122 «Комп'ютерна наука» в українські ЗВО.

Апробація проведеної роботи відбулася в рамках XI Міжнародної

науково-практичної конференції молодих учених «Інформаційні технології: економіка, техніка, освіта» 10-11 листопада 2020 року, м. Київ, Україна [3]. В даних тезах було висвітлено важливість загальнодоступних ресурсів (вебсайтів)

для ЗВО, посилаючись на статтю 30 «Прозорість та інформаційна відкритість закладу освіти» Закону про освіту №2145-VIII від 5 вересня 2017 року [4] та на

декілька праць; було розглянуто один із відомих світових рейтингів університетів світу – Ranking web of universities. Крім цього, робота була

апробована в рамках IV всеукраїнської науково-практичної інтернет конференції студентів та аспірантів «Теоретичні та прикладні аспекти розробки

комп'ютерних систем» 29 квітня 2021 року, м. Київ, Україна [5]. В даних тезах описуються кроки збору інформації, їх обробки, створення та наповнення

сховища даних, проектування OLAP кубу та формування звітів за допомогою SSRS; розраховано загальний коефіцієнт кореляції відносно даних за період

2018-2020 рр.; сформувано попередній висновок щодо основної гіпотези дослідження. Також була апробація в рамках XII міжнародної науково-

практичної конференції молодих учених «Інформаційні технології: економіка, техніка, освіта» 11-12 листопада 2021 року, м. Київ, Україна [6]. В цих тезах

висвітлено розроблену систему «клієнт-сервер», аналіз побудованих нею графіків та таблиці кореляції за період 2018-2021 рр.; зроблено загальний висновок щодо результатів дослідження.

До складу пояснювальної записки входять: вступ, 4 розділи, висновки, список використаних джерел (57 найменувань) та 3 додатки. Робота містить 30

рисунків. Загальний обсяг роботи становить 64 сторінки, основний текст роботи викладено на 44 сторінках. У вступі описано актуальність роботи, проблематику,

предмет, об'єкт, мету, завдання, методи дослідження, а також зазначено апробації роботи.

В першому розділі пояснювальної записки міститься інформація про предметну область, розглянуті інформаційні джерела за тематикою дослідження,

проаналізовані існуючі рішення; поставлено основні завдання, які повинні бути виконані в результаті виконання; коротка характеристика інформаційно-

аналітичної системи. В другому розділі описано основні теоретичні відомості про моделювання системи; використані технології на програмні засоби; наведено різні діаграми, що описують розроблену інформаційно-аналітичну систему;

описано вузли, з яких збирається інформація. Третій розділ описує розробку

системи; створення сховища даних, підготування та обробка зібраної інформації,

заповнення сховища даних, створення проєктів та інформаційно-аналітичної системи типу «клієнт-сервер». В четвертому розділі аналізуються отримані КІЕ,

графічні звіти й таблиця кореляцій, сформовано короткі висновки по кожному з

них. У висновках сформульовано короткі підсумки кожного розділу, наводяться

результати виконання дослідження та виконані завдання, надаються рекомендації.

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

1 СИСТЕМНИЙ АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Опис предметної області

Вебунна кампанія є важливою для кожного ЗВО, оскільки потрібно заохотити якнайбільше талановитих абітурієнтів вступити саме в цей ЗВО. Як наслідок – посісти найкращі місця в рейтингах, мати більше фінансування з боку держави і випустити фахівців, які зможуть позитивно впливати на розвиток української науки. А для того, щоб мати більшу кількість абітурієнтів, необхідно мати поширеність і популярність серед публіки. Один із способів бути видимим для майбутніх студентів – мати якісний вебсайт в мережі Інтернет, який наповнений актуальною необхідною інформацією, має відомості про структурні підрозділи, склад співробітників, дані про освітні програми, різноманітні літературні матеріали, тощо. Ще можна зазначити наявність представництв у соціальних мережах, що дозволяє створювати таргетовані кампанії на вступників. Також не менш важливими факторами є поширеність публікацій співробітників, цитованість їхніх праць, оскільки це показує науковий рівень закладу і внесок її співробітників у науку.

В свою чергу, всі ці показники мають вплив і на місце в рейтингах. Наприклад, один із найбільш відомих світових рейтингів є Ranking web of universities, в якому оцінено більше 31 тисячі ЗВО зі всього світу. Станом на 2021 рік в це число включено 314 українських ЗВО. Цей проект є ініціативою Cybermetrics Lab, дослідницької групи Consejo Superior de Investigaciones Científicas (CSIC) – найбільшого державного дослідницького органу в Іспанії. CSIC підпорядкований Міністерству освіти і його головною метою є сприяння науковим дослідженням, формуванню нових дослідників для підвищення прогресу науково-технічного рівня країни, що в свою чергу сприятиме підвищенню добробуту громадян.

В цьому рейтингу ЗВО світу оцінюються за 4 показниками (метриками): Presense, Visibility, Transparency та Excellence. Перший показник – присутність –

це кількість проіндексованих пошуковою системою Google сторінок сайту ЗВО [7]. Другий показник – видимість або вплив – це комбінація отриманих даних від таких ресурсів як Majestic SEO та Ahrefs щодо кількості зовнішніх посилань на домен сайту ЗВО та кількості доменів, які посилалися, оскільки це може свідчити

про видимість університету і важливість опублікованих матеріалів. Третій показник – відкритість – це цитованість тої 210 авторів. Для цього використовується платформа Google Scholar. І четвертий показник – висока якість – кількість робіт серед 10% найбільш цитованих у кожній з усіх 27 дисциплін БД за даними Scimago.

Згідно попередньої методики Ranking web of universities у 2020 році дані показники мали вагу 5%, 50%, 10% та 35% відповідно. У 2021 дослідницька група вирішила припинити використання показника присутності, змінивши вагу останніх трьох параметрів на 50%, 10% та 40% відповідно. Спосіб оцінювання

ЗВО може змінюватися час від часу, бо він знаходиться в процесі постійного вдосконалення. Варто зазначити, що мета розробників методики полягає не в тому, щоб оцінити вебсайти, дизайн, зручність використання чи популярність їх вмісту за кількістю відвідувачів. Вебіндикатори розглядаються як правильна, комплексна, глибока оцінка глобальної діяльності університету з урахуванням його результатів, а також їх актуальності та впливу.

Даний ресурс було обрано через відомість у світі, відкриту методологію і широке оцінювання, наявність даних за великий період часу та для великої кількості українських ЗВО. Крім цього, дослідницька група відповідально ставиться до оцінювання і тому виключає або «штрафує» ЗВО, які намагаються підвищити свою рейтингову позицію шляхом штучного збільшення вебтрафіку, дублювання опублікованих робіт та використання неправильних практик. Даний факт має важливу роль у формуванні справедливих оцінок. Ще вагомим

фактором у виборі саме цього ресурсу є те, що він щороку фігурує в інформаційному полі українських ЗВО. Наприклад, такі заклади як: Харківський національний університет радіоелектроніки, Київський університет ім. Бориса Грінченка, Київський політехнічний інститут, Київський національний

економічний університет вже щорічно слідкують за позицією своїх закладів у даному рейтингу та публікують різноманітні новини [8-11]. А КПІ взагалі має окремий вебсайт [12], який було створено для допомоги своїм підрозділам, щоб ті збільшили свій вклад у підвищення рейтингу ЗВО в Ranking web of universities.

Інші системи рейтингів з використанням вебметрик ЗВО було вирішено не враховувати в даній роботі, оскільки вони можуть містити кардинально іншу методологію (яка, крім того, може бути не загальнодоступною), інший (менший) набір ЗВО, й тому комбінація різних джерел з різними підходами щодо оцінювання в рамках одного дослідження могла б призвести до хибних результатів.

Розроблена інформаційно-аналітична система використовує дані статистики вступних кампаній абітурієнтів та рейтингових списків ЗВО з метриками за період 2018-2021 рр., зібрані з ресурсів ЄДЕБО та Ranking web of universities відповідно.

1.2 Огляд інформаційних джерел та аналіз існуючих рішень

Українські вчені І. В. Гончарук та Н. П. Юрчук у своїй статті «Організація єдиного електронного науково-освітнього простору сучасного університету» обґрунтували необхідність створення єдиного електронного науково-освітнього простору (ЄНОП) для ЗВО [13]. Згідно з авторами, «єдиний електронний науково-освітній простір ЗВО» – це інтеграційна система. Вона включає інформаційні, матеріально-технічні та людські ресурси, забезпечує автоматизацію управлінських і науково-освітніх процесів, а також робить обробку, передачу й зберігання інформації узгодженим. На їх думку, діджиталізація освіти вимагає роботи в кількох напрямках. Так, ЄНОП може містити такі компоненти: інформаційні ресурси, організаційні структури, засоби інформаційної взаємодії студентів і ЗВО. У статті описано аналіз досвіду організації ЄНОП в Вінницькому національному аграрному університеті (ВНАУ), який було реалізовано шляхом впровадження електронної системи

управління університетом «Сократ». Це веборієнтована клієнт-серверна система. Як результат, було доведено, що якість освіти у ЗВО може бути підвищена завдяки інформаційно-комунікаційним технологіям і створенню єдиного електронного науково-освітнього простору.

У статті І. Шеробури «Використання інформаційних технологій в управлінні вищим навчальним закладом» описані напрями в управлінні ЗВО, в яких теоретично можуть використовуватися інформаційні технології [14]:

- надання майбутнім фахівцям можливостей здобуття освіти з використанням інформаційно-комунікаційних технологій (ІКТ) під час професійної підготовки;
- розвиток інформаційного менеджменту та ресурсів.

Крім того, авторка вказує загальні тенденції діджиталізації освіти, а саме створення та підтримка вебсайтів ЗВО, наповнення їх необхідним контентом.

Щоб досягнути мети діджиталізації освіти, потрібно особливо акцентувати увагу на створенні інформаційно-освітнього простору у ЗВО, а також використовувати засоби ІКТ та комп'ютерно-орієнтованих технологій. Зі статті очевидно, що діджиталізація суспільства не може відбутися без діджиталізації освіти.

Також у якості джерела було розглянуто автореферат Левченко Н. П. «Відкриті електронні ресурси у діяльності бібліотек закладів вищої освіти: специфіка формування, управління, доступ» [15]. Авторка вважає, що завдання створення єдиного електронного інформаційного простору (ЄЕІП) у ЗВО має бути делеговано бібліотекам. Успішна реалізація такого простору є важливим рушієм прискорення процесів діджиталізації в усіх напрямках діяльності ЗВО. Створення ЄЕІП значно вплине на досягнення цілей освітньої діяльності, а також на накопичення і надання навчально-наукової інформації.

Левченко Н.П. провела дослідження за критеріями Webometrics, яке показало, що розвиток та підтримка відкритих електронних ресурсів ЗВО позитивно впливають на підвищення позиції ЗВО у загальному рейтингу. Це дуже важливо для ЗВО, тому що бажання здобути якомога вищу позицію у світових рейтингах є мотиватором для пошуку ефективної стратегії розвитку

ЗВО. Також це стимулює зростання конкурентоспроможності та вдосконалення внутрішньої організації освітньої діяльності.

Чотири метрики Webometrics (присутність, видимість, прозорість та висока якість) показують, що наукові публікації відіграють важливу роль у формуванні позиції ЗВО в рейтингу. Більш того, висока публікаційна активність є показником авторитетності та впливовості вчених, що позитивно впливає на імідж ЗВО. Отже, науковці змотивовані писати та публікувати якісні роботи у

провідних наукових виданнях, таким чином сприяючи збільшенню обсягу та інформативності наукових публікацій. Тут на перший план виходять розвиток та

підтримка відкритих електронних ресурсів ЗВО. Такі ресурси роблять вагомий внесок у підвищення якості вищої освіти в українських ЗВО. Окрім того, науковці можуть підвищувати авторитетність у своїй сфері науки, створюючи та

наповнюючи релевантним контентом свої бібліометричні портрети на таких наукометричних платформах як Google Scholar, Scopus, Web of Science, тощо.

Такий вид наукової діяльності (за умови постійної активності) позитивно впливає і на науковий рейтинг науковця, і на позицію ЗВО у рейтингових списках світових університетів.

У статті «Webometrics: evolution of social media presence of universities» [16] було проведено дослідження кореляції між рейтингом вебметричних університетів та рейтингами, наданими відомими світовими рейтингами університетів за період 2005–2016 років. Серед останніх можна відмітити такий

як світовий рейтинг університетів QS [17]. Авторами було розроблено власний рейтинг для університетів за період 2005–2011, оскільки сам Ranking web of universities не мав даних за цей період. Їхній рейтинг має сильну кореляцію з методологією оцінювання Ranking web of universities. Було обчислено рейтинг

Webometrics для 30 найкращих університетів Північної Америки, Європи та Азії за період 2005–2016 років. Результати вказують на позитивну кореляцію для

університетів Північної Америки та Європи, але слабку кореляцію для університетів Азії. Це можна пояснити тим, що азіатські університети не приділяли особливої уваги своїм вебсайтам у порівнянні з університетами

Північної Америки та Європи. Загальні результати показують, що північноамериканські та європейські університети вищі за рейтингом порівняно з азіатськими.

Стаття «Statistical assesment of Webometric rating indicators of Ukrainian higher education institutions» [18] присвячена статистичному аналізу інструментів, що використовуються для оцінки міжнародної конкурентоспроможності ЗВО. Аналіз здійснено на основі рейтингу університетів Webometric. Метою дослідження авторів цієї статті є детальний розгляд одного з найважливіших і широко визнаних світових рейтингів – Ranking Web of World Universities, визначення місця в ньому українських вищих навчальних закладів, визначення перспектив подальшого розвитку та відповідного коригування їх вебполітики. На основі отриманої статистичної оцінки було виявлено причини недостатньо високих рейтингових позицій українських вишів у системі Webometrics, а саме: для світових університетів найбільш ваговими метриками є Вплив (Impact) та Відкритість (Openness), а для українських – Присутність (Presense), Перевага за якістю (Excellence), Вплив (Impact), середні показники яких нижчі за середнє значення Впливу. Розроблено пропозиції щодо покращення представницьких показників українських ЗВО в мережі Інтернет та підняття ЗВО у міжнародному рейтингу за критеріями Webometric, що в свою чергу також буде мати вплив на якісний розвиток вітчизняної науково-освітньої школи в контексті світових тенденцій. Серед пропозицій можна відмітити такі: збільшення кількості сторінок з різноманітним контентом; систематично розмішувати тези доповідей, виступи на семінарах та інші матеріали; розробляти електронні курси; збільшення присутності ЗВО у соціальних мережах; включення посилань на сайт у наукових публікаціях, тощо.

Проблематикою рейтингів українських ЗВО також займалися автори статті «Webometrics ranking analysis and possible ways to improve the position of the university» [19], в якій було сформульовано проблему низьких позицій українських університетів у списку світових рейтингів. Крім цього, було наведено загальні рішення щодо підвищення рейтингу університету.

Встановлено, що для підвищення рейтингу конкретного ЗВО необхідна розробка індивідуальної стратегії просування ЗВО у рейтингу з побудовою прогнозної моделі оцінки Webometrics та дослідженням своєї активності.

Важливість підвищення рейтингів ЗВО задля завоювання більшої частки висококонкурентного ринку освітніх послуг доведено у статті «Webometrics ranking of universities as a factor of gaining competitive advantage in the market for educational services» [20]. Представляючи на своїх офіційних вебсайтах всю

достовірну інформацію про зміст та сферу своєї діяльності, вищі навчальні заклади звертаються до цільової аудиторії. Проте, на превеликий жаль, деякі університети зловживають цією можливістю передачі важливої інформації через мережу Інтернет і вдаються до несправедливої практики підвищення свого рейтингу, за що в подальшому будуть «оштрафовані» дослідниками із

Webometrics. Для обґрунтування значущості вебметричного рейтингування університетів як інструменту маркетингу авторами було проаналізовано функціонування офіційного сайту «Вінницький державний педагогічний університет ім. Михайла Коцюбинського» (ВДПУ). Отримані результати однозначно свідчать про недостатній рівень присутності ВСПУ в мережі, тому було надано перелік пропозицій щодо покращення ситуації.

Залежність між соціальною мережею Facebook, вебсайтом ЗВО та рейтингом Ranking web of universities досліджувалось в статті «Online engagement in higher education institutions: the triple relationship among Webometrics ranking, Facebook and website» [21]. Як стало зрозуміло з дослідження, важлива не тільки кількість шанувальників або відвідувань вебсайту, але й їхня зацікавленість (активність). Тому було досліджено: чи рівень зацікавленості позитивно впливає на рейтинг вебметрії і чи більше зацікавленості на Facebook також викликає більшу відвідуваність вебсайту.

Також було надано теоретичні та управлінські наслідки щодо важливості залучення Facebook для ЗВО.

Крім цього, було знайдено дослідження з точки зору впливу доступності вебсайту ЗВО на позицію в рейтингу. У роботі «Exploring the Relationship between

Web Presence and Web Usability for Universities: A Case Study from Turkey» [22]

було досліджено зв'язок між вебприсутністю та зручністю використання університетів. З цією метою було відібрано п'ять турецьких університетів із

рейтингу університетів Webometrics у липні 2012 року з точки зору показника

вебприсутності. Результати оцінки та кореляційного аналізу показали низку

помітних висновків. По-перше, було виявлено значущу позитивну кореляцію

між присутністю в Інтернеті та зручністю використання університетських

вебсайтів. Це вказує на те, що якщо обсяг вебприсутності університету низький,

то і рівень доступності вебсайту цього університету слабкий. Можна зробити

висновок, що користувачі витрачають менше часу і стикаються з меншими

проблемами зі зручністю використання вебсайтів університетів, які мають

сильнішу присутність в Інтернеті. Іншими словами, університетські вебсайти, які

мають сильну присутність в Інтернеті, швидше за все, задовольняють потреби

користувачів. Крім того, оцінка зручності використання виявила кілька

поширених проблем дизайну на вибраних університетських вебсайтах, і ці

висновки здебільшого підтверджують ті, які були запропоновані відповідними

попередніми дослідженнями. Ця оцінка робить акцент на загальних проблемах

зручності використання, які слід враховувати при розробці або оцінці

університетських вебсайтів.

З практичної точки зору результати цього дослідження мають важливі

наслідки для університетів. По-перше, згідно з висновками цього дослідження,

університети можуть покращити зручність використання в Інтернеті,

приділяючи важливе значення обсягу своєї присутності в Інтернеті. По-друге,

університети можуть оцінити свій рівень зручності використання в Інтернеті,

досліджуючи свій рейтинг присутності в Інтернеті та використовуючи

різноманітні інструменти для аналізу доступності вебресурсів. Після такого

аналізу ЗВО можуть підвищити свій рейтинг у системі рейтингу Webometrics,

покращуючи зручність використання своїх вебсайтів. Крім того, це дослідження

також дає вказівки для подальших досліджень щодо дослідження зв'язку між

присутністю в Інтернеті та зручністю використання в Інтернеті.

Якщо підсумувати вище згадані джерела, розвиток і підтримка інтернет представництва ЗВО (вебсайти, відкриті електронні ресурси, вебпредставництва у соціальних мережах) є важливими задля розвитку, покращення якості та конкурентоспроможності закладів. Соціальні мережі мають позитивний вплив на

позиціонування ЗВО у рейтингах. Також не менш важливою є підтримка наукового статусу закладів: активна публікаційна діяльність, створення й наповнення бібліометричних профілів в таких системах Google Scholar, Scopus, Web of Science тощо. Деякі ЗВО, наприклад, КНУ, вже проводять семінари по

актуалізації профілів своїх співробітників у даних системах [23]. Проте, в даних джерелах не розповідається про те, який вплив мають показники метрик на кількість поданих заяв абітурієнтами, що в свою чергу підкреслює наукову новизну даного дослідження. Тому є необхідність у створенні інформаційно-аналітичної системи, яка зможе показати, чи є кореляція між кількістю поданих

заяв студентами і метриками інтернет представництв ЗВО.

На порталі спеціальної інформаційної системи УкрПатенту за предметом дослідження було знайдено такі патенти: «Інформаційно-аналітична система» [24], «Комп'ютерна програма «Моніторингова інформаційно-аналітична система»» [25], «Комп'ютерна програма «Інформаційно-аналітична система «Університет»» [26], «Опис комп'ютерної програми «Моніторингова інформаційно-аналітична система»» [27].

За запитом «інформаційно-аналітична система показників вступної кампанії» у пошуковій мережі не вдалося знайти прямих аналогів системі, що розроблялася. Із подібних у даній сфері можна відмітити інформаційно-аналітичну систему «Університет», яку використовує Сумський державний університет для підтримки своїх інформаційно-аналітичних процесів [28].

1.3 Постановка задачі

Основне завдання – перевірити наявність кореляції між кількістю поданих заяв абітурієнтами на спеціальність 122 «Комп'ютерні науки» та інтернет

ресурсами ЗВО для гіпотези «зі збільшенням кількості заяв покращується (зменшується число) загальне значення метрики» та надати рекомендації.

Для вирішення поставленої задачі необхідно виконати завдання:

- зібрати необхідні дані з ресурсів ЄДЕБО, а саме: кількість поданих заяв в українські ЗВО за період 2018-2021 рр., середній і мінімальний бал вступника, кількість допущених і рекомендованих абітурієнтів;

- зібрати дані з Ranking web of universities, а саме: рейтингові списки українських ЗВО зі значеннями 4 метрик за період 2018-2021 рр.;

- створити сховище даних та імпортувати ці дані в нього;
- розробити інформаційно-аналітичну систему, яка буде підключатися до даного сховища, виводити інформацію у вигляді графіків та таблиць, розраховувати коефіцієнти кореляції;

- створити проєкти типу SSRS та SASS, розгорнути куб, побудувати КПЕ та звіти;

- проаналізувати побудовані графіки й отримані дані щодо загальної кореляції, зробити відповідні висновки та надати рекомендації.

Інформаційно-аналітична система забезпечує:

- збір інформації про вступні кампанії;
- збір інформації про рейтингові списки ЗВО, значення метрик;
- вивід інформації для аналізу у вигляді графіків та таблиць;
- розрахунок коефіцієнтів кореляції.

НУБІП України

2. МОДЕЛЮВАННЯ СИСТЕМИ

2.1 Моделювання предметної області

Уніфікована мова моделювання (UML) — це мова моделювання загального призначення для розробки в області розробки програмного забезпечення, яка призначена забезпечити стандартний спосіб візуалізації дизайну системи [29].

Створення UML спочатку було мотивовано бажанням стандартизувати розрізнені системи позначення та підходи до проєктування програмного забезпечення. Вона була розроблена в Rational Software у 1994-1995 роках, а подальша розробка велась до 1996 року.

У 1997 році UML була прийнята як стандарт Групою управління об'єктами (OMG), і з тих пір керується цією організацією. У 2005 році UML також була опублікована Міжнародною організацією зі стандартизації (ISO) як затверджений стандарт ISO. З тих пір цей стандарт періодично переглядався, щоб охопити останню версію UML.

UML пропонує спосіб візуалізації архітектурних планів системи на діаграмі, включаючи такі елементи, як:

- будь-які види діяльності (роботи);
- окремі компоненти системи і як вони можуть взаємодіяти з іншими програмними компонентами;
- як буде працювати система;
- як сутності взаємодіють з іншими (компоненти та інтерфейси);
- зовнішній інтерфейс користувача.

Хоча спочатку UML була призначена для об'єктно-орієнтованої проєктної документації, потім вона була розширена до більшого набору і була визнана корисною у багатьох аспектах.

Далі інформаційно-аналітично систему буде описано за допомогою діаграми прецедентів, розгортання, послідовності та діяльності.

2.1.1 Діаграма прецедентів.

Діаграма варіантів використання (або прецедентів) це графічне зображення можливої взаємодії користувача з системою [30]. Діаграма

прецедентів показує різні варіанти використання та різні типи користувачів, яких система має і часто супроводжується діаграмами інших типів. Варіанти використання представлені колами або еліпсами. Діючі особи часто зображені у вигляді фігурок.

У той час, як сам варіант використання може детально розглянути кожен можливість, діаграма прецедентів може допомогти представити систему на більш високому рівні. Раніше було сказано, що «діаграми варіантів використання є кресленнями для вашої системи».

Завдяки своїй спрощеній природі діаграми прецедентів можуть бути хорошим інструментом комунікації для зацікавлених сторін. Малюнки намагаються імітувати реальний світ і забезпечують уявлення зацікавленої сторони, щоб зрозуміти, як буде розроблена система. Було проведено дослідження, щоб визначити, чи була дійсна ситуація для діаграм варіантів використання взагалі, чи вони непотрібні [31]. В результаті було виявлено, що діаграми прецедентів більш спрощено передають зацікавленим сторонам намір системи і що вони «інтерпретуються більш повно, ніж діаграми класів».

Діаграму прецедентів інформаційно-аналітичної системи наведено на рисунку 2.1. На ній зображено розподіл ключових сторін системи, їхню діяльність. Наприклад, аналітик формує КПЕ і звіти та надає їх керівнику інформаційної служби, а оператор Webometrics формує метрики, списки ЗВО та оцінює їх.

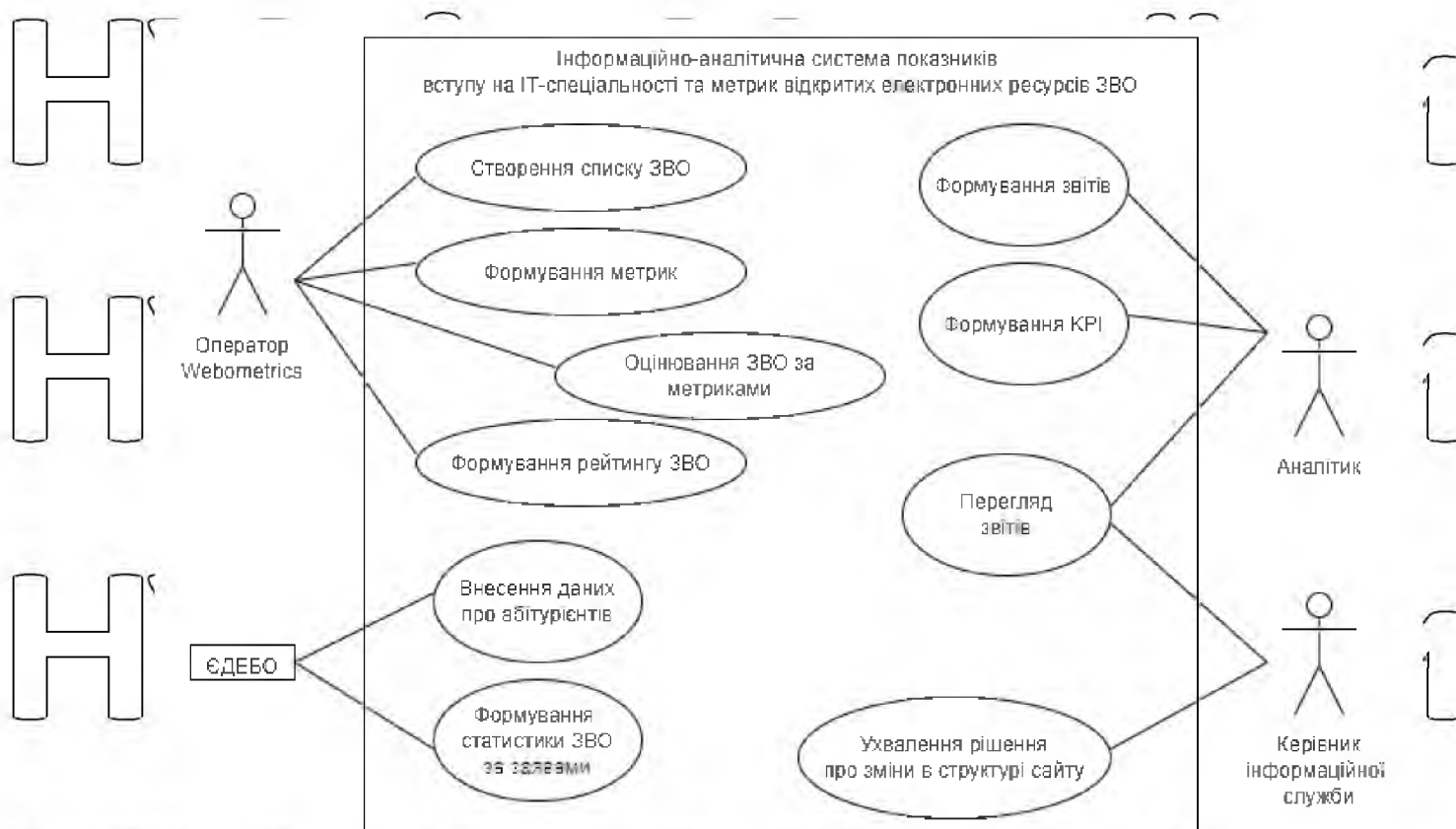


Рис. 2.1 Діаграма прецедентів

2.1.2 Діаграма розгортання.

Діаграма розгортання в UML моделює фізичне розгортання артефактів на вузлах [32]. Для опису вебсайту, наприклад, діаграма розгортання покаже, які апаратні компоненти («вузли») існують (наприклад, вебсервер, сервер додатків і сервер баз даних), які програмні компоненти («артефакти») працюють на кожному вузолі (наприклад, вебдодаток, база даних), а також спосіб з'єднання різних частин (наприклад, JDBC, REST, RMI).

Вузли відображаються у вигляді квадратів, а артефакти, які виділені кожному вузлу, відображаються у вигляді прямокутників у межах. Вузли можуть мати підвузли, які виглядають як вкладені блоки. Один вузол на схемі розгортання може концептуально представляти кілька фізичних вузлів, наприклад кластер серверів баз даних.

Існує два типи вузлів:

- вузол пристрою;
- вузол середовища виконання.

Вузли пристроїв — це фізичні обчислювальні ресурси з пам'яттю обробки та службами для виконання програмного забезпечення, наприклад типових комп'ютерів або мобільних телефонів. Вузол середовища виконання — це програмний обчислювальний ресурс, який працює всередині зовнішнього вузла і який сам надає послугу для розміщення та виконання інших програмних елементів.

Топологію системи наведено на рисунку 2.2:

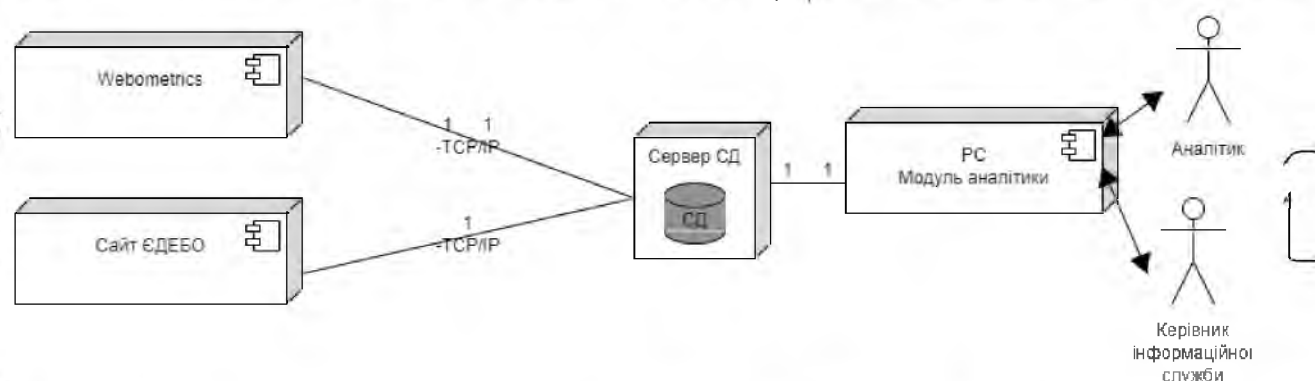


Рис. 2.2 Топологія системи

2.1.3 Діаграма послідовності

Діаграма послідовності або діаграма системної послідовності показує взаємодії об'єктів, упорядкованих у часовій послідовності в області програмної інженерії [33]. Вона зображує об'єкти, залучені до сценарію, і послідовність повідомлень, якими обмінюються об'єкти, необхідні для виконання функціональних можливостей сценарію. Діаграми послідовності зазвичай асоціюються з реалізацією варіантів використання в логічному уявленні системи, що розробляється. Діаграми послідовності іноді називають діаграмами подій або сценаріями подій.

Для конкретного сценарію варіанту використання діаграма показує події, які генерують зовнішні суб'єкти, їх порядок та можливі міжсистемні події.

Усі системи розглядаються як чорний ящик; діаграма акцентує увагу на подіях, які перетинають кордон системи від дійових осіб до систем. Для основного сценарію успіху варіанту використання та частих або складних альтернативних сценаріїв слід створити діаграму послідовності системи.

Діаграма послідовності показує у вигляді паралельних вертикальних ліній (ліній життя) різні процеси або об'єкти, які живуть одночасно, а також у вигляді горизонтальних стрілок повідомлення, якими вони обмінюються, у тому

порядку, в якому вони відбуваються. Це дозволяє специфікувати прості сценарії часу виконання в графічному вигляді.

Діаграма послідовності системи повинна вказувати та показувати наступне:

- зовнішніх дійових осіб;
- повідомлення (методи), викликані цими акторами;
- повертання значень (якщо такі є), пов'язані з попередніми повідомленнями;
- вказівку будь-яких циклів з областю ітерації.

Діаграму послідовності для системи наведено нижче на рисунку 2.3:

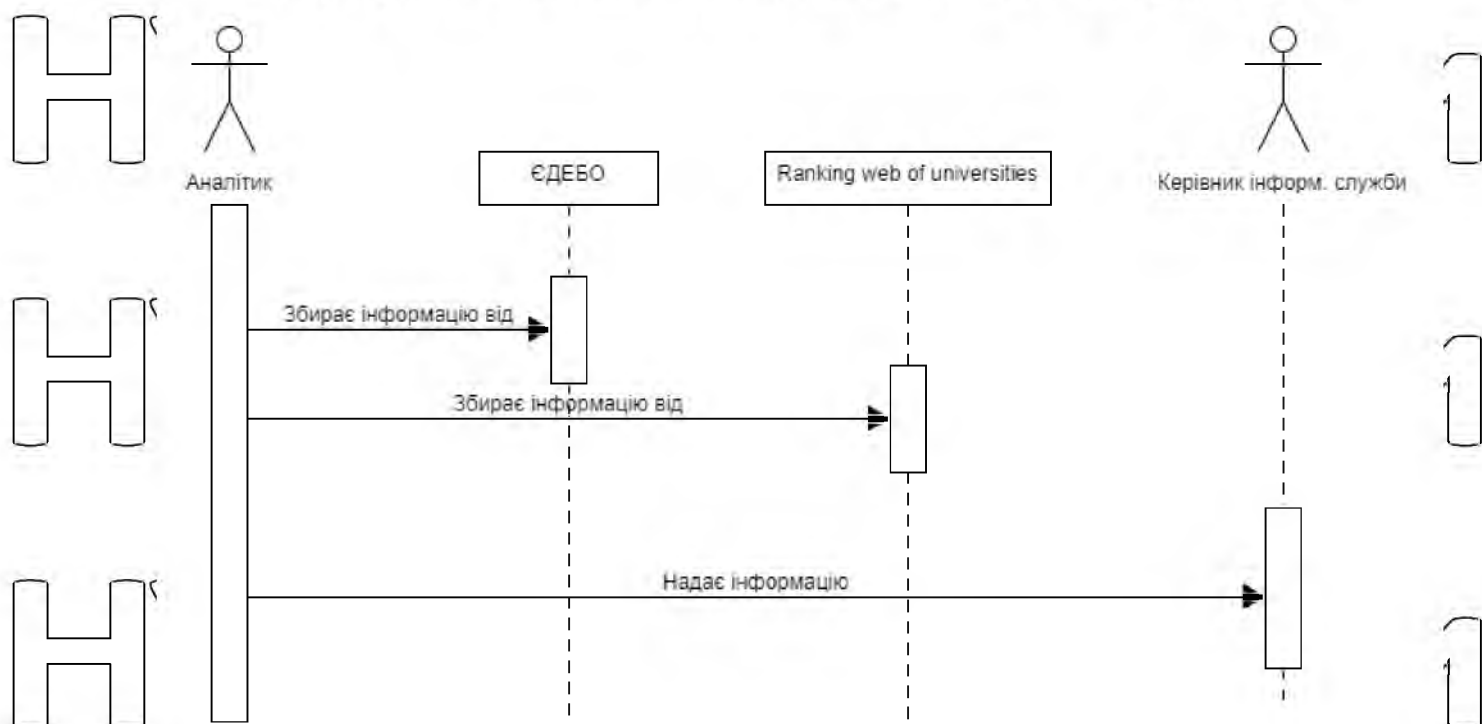


Рис. 2.3 Діаграма послідовності

2.1.4 Діаграма діяльності.

Діаграма діяльності – це графічне представлення робочих процесів, поетапних дій і дій з підтримкою вибору, ітерації та паралельності [34]. В UML

діаграми діяльності призначені для моделювання як обчислювальних, так і організаційних процесів (тобто робочих процесів), а також потсків даних, що перетинаються з пов'язаними видами діяльності. Хоча діаграми діяльності в першу чергу показують загальний потік контролю, вони також можуть включати

елементи, що показують потік даних між діями через одне або кілька сховищ даних.

Діаграми діяльності будуються з обмеженої кількості фігур, з'єднаних стрілками. Найважливіші типи фігур та їхня роль:

- крапки: представляють дії;
- діаманти: символізують рішення;
- смуги: представляють початок (розщеплення) або кінець (об'єднання) одночасних дій;
- чорне коло: представляє початок (початковий вузол) робочого процесу;
- обведене чорне коло: представляє кінець (кінцевий вузол).

Стрілки йдуть від початку до кінця і представляють порядок, у якому відбувається дія.

Діаграми діяльності можна розглядати як форму структурованої блок-схеми в поєднанні з традиційною схемою потоку даних. У типових методах блок-схем відсутні конструкції для вираження паралельності. Однак символи об'єднання та розбиття на діаграмах діяльності вирішують це лише для простих випадків; сенс моделі не зрозумілий, коли вони довільно поєднуються з рішеннями або циклами.

Діаграму активності системи наведено на рисунку 2.4:

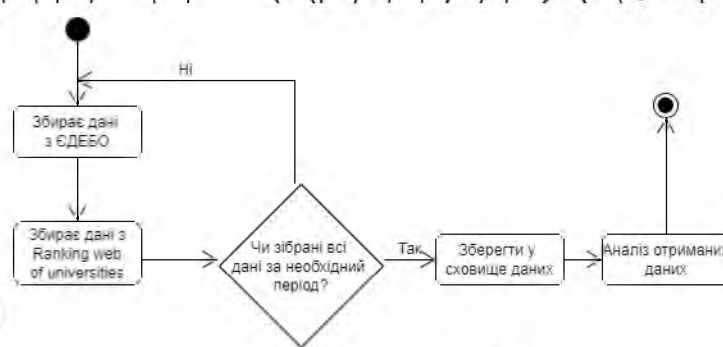


Рис. 2.4 Діаграма активності

2.2 Опіє вузлів, з яких збирається інформація

Як було раніше зображено на діаграмі розгортання (див. рис. 2), головними джерелами інформації для сховища даних є ресурси ЄДЕБО та Ranking web of universities.

На порталі ЄДЕБО в розділі «Статистика» «За спеціальностями (спеціалізаціями) і закладами освіти» (рис. 2.5) було зібрано дані за період 2018-2021 рр. (для попередніх років був використаний відповідний розділ на сайті «Попередні роки») з такими параметрами фільтрів:

- спеціальність: Т22 «Комп'ютерні науки»;
- освітній рівень: бакалавр;
- форма навчання: денна.

ОС	Назва спеціальності	Рік початку року	Відомості про спеціальність	Умови вступу	Інформація про спеціальність	Статус спеціальності	Розподіл за спеціальністю	Міс. Рівн. 2020	Заг. Рівн. 2020	Попередні роки	Міс. Рівн. 2021	Заг. Рівн. 2021
Т22	Комп'ютерні науки	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
Т22	Комп'ютерні науки	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
Т22	Комп'ютерні науки	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028
Т22	Комп'ютерні науки	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028

Рис. 2.5 Статистика вступної кампанії 2021 на порталі ЄДЕБО

Дані отримані дані було перенесено в Excel файл, зведено, очищено, додано колонку року та збережено у форматі CSV для подальшого використання.

Один із таких файлів обробленої статистики вступної кампанії з ЄДЕБО наведено на рисунку 2.6:

	A	B	C	D	E	F	G	
1	University	Усього по Допущен Усього реі Мін.бал (І Сер.бал (І Year						
2	Відокремлений підрозділ Національного університету Біоресурсів і природокористування України "Бережанський агротехнічний інститут" (Тернопільська)	19	10	4 186.466	189.438	2021		
3	Вінницький національний аграрний університет (Вінниця)	327	136	2 171.562	176.139	2021		
4	Вінницький національний технічний університет (Вінниця)	895	551	53 180.489	185.329	2021		
5	Волинський національний університет імені Лесі Українки (Волинська)	317	218	26 177.076	183.914	2021		
6	Державний вищий навчальний заклад "Донецький національний технічний університет" (Донецька)	68	56	13 158.080	168.328	2021		
7	Державний вищий навчальний заклад "Київський національний економічний університет імені Вадима Гетьмана" (Київ)	1 409	646	55 181.100	183.362	2021		
8	Державний вищий навчальний заклад "Національний лісотехнічний університет України" (Львівська)	616	379	26 177.684	180.944	2021		
9	Державний вищий навчальний заклад "Приазовський державний технічний університет" (Донецька)	304	247	30 174.512	184.718	2021		
10	Державний вищий навчальний заклад "Придніпровська державна академія будівництва та архітектури" (Дніпропетровська)	185	86	6 172.380	174.440	2021		
11	Державний вищий навчальний заклад "Прикарпатський національний університет імені Василя Стефаника" (Івано-Франківська)	475	282	30 172.967	179.524	2021		
12	Державний вищий навчальний заклад "Ужгородський національний університет" (Закарпатська)	252	169	25 172.788	184.403	2021		
13	Державний вищий навчальний заклад "Український державний хіміко-технологічний університет" (Дніпропетровська)	208	98	8 171.562	175.513	2021		
14	Державний заклад "Луганський національний університет імені Тараса Шевченка" (Луганська)	42	27	7 135.928	145.930	2021		
15	Державний університет "Житомирська політехніка" (Житомирська)	446	273	11 184.808	189.099	2021		
16	Державний університет "Одеська політехніка" (Одеська)	1 850	1 188	138 171.054	183.293	2021		
17	Державний університет економіки і технологій (Дніпропетровська)	158	90	4 175.848	180.387	2021		
18	Державний університет інтелектуальних технологій і зв'язу (Одеська)	375	227	19 171.911	176.504	2021		
19	Державний університет інфраструктури та технологій (Київ)	480	137	6 173.600	175.958	2021		
20	Державний університет телекомунікацій (Київ)	1 270	337	12 180.100	183.100	2021		
21	Дніпровський національний університет імені Олександра Гончара (Дніпропетровська)	593	342	32 176.460	181.051	2021		
22	Донецька державна машинобудівна академія (Донецька)	161	142	82 132.600	156.790	2021		
23	Донецький державний університет управління (Донецька)	75	59	20 151.005	166.221	2021		
24	Донецький національний університет імені Василя Стуса (Вінниця)	470	297	38 177.888	183.410	2021		
25	Дрогобицький державний педагогічний університет імені Івана Франка (Львівська)	119	76	6 174.624	180.819	2021		
26	Житомирський державний університет імені Івана Франка (Житомирська)	183	82	6 175.396	181.023	2021		
27	Закарпатський національний університет (Закарпатська)	276	134	7 176.358	181.781	2021		
28	Західноукраїнський національний університет (Тернопільська)	455	262	29 175.134	179.830	2021		
29	Кам'янець-Подільський національний університет імені Івана Огієнка (Хмельницька)	133	61	6 176.904	178.883	2021		
30	Київський національний торговельно-економічний університет (Київ)	1 150	372	25 181.200	183.443	2021		
31	Київський національний університет будівництва та архітектури (Київ)	699	303	24 175.400	179.930	2021		
32	Київський національний університет імені Тараса Шевченка (Київ)	3 840	1 547	205 186.200	192.404	2021		
33	Київський національний університет технологій та дизайну (Київ)	744	369	29 174.100	177.641	2021		
34	Київський університет імені Бориса Грінченка (Київ)	551	196	10 181.200	182.850	2021		

Рис. 2.6 CSV файл обробленої статистики з ресурсу ЄДБНО за 2021 р.

Дані з ресурсу Ranking web of universities за 2021 рік було зібрано на сторінці України (рис. 2.7) в Excel файл.

Рис. 2.7 Сайт Ranking web of universities

Проте, сайт не містить даних за попередні роки, тому довелося скористатися Інтернет архівом «WayBack machine» [35], який дозволяє переглянути «знімки» сайтів у минулому (рис. 2.8).

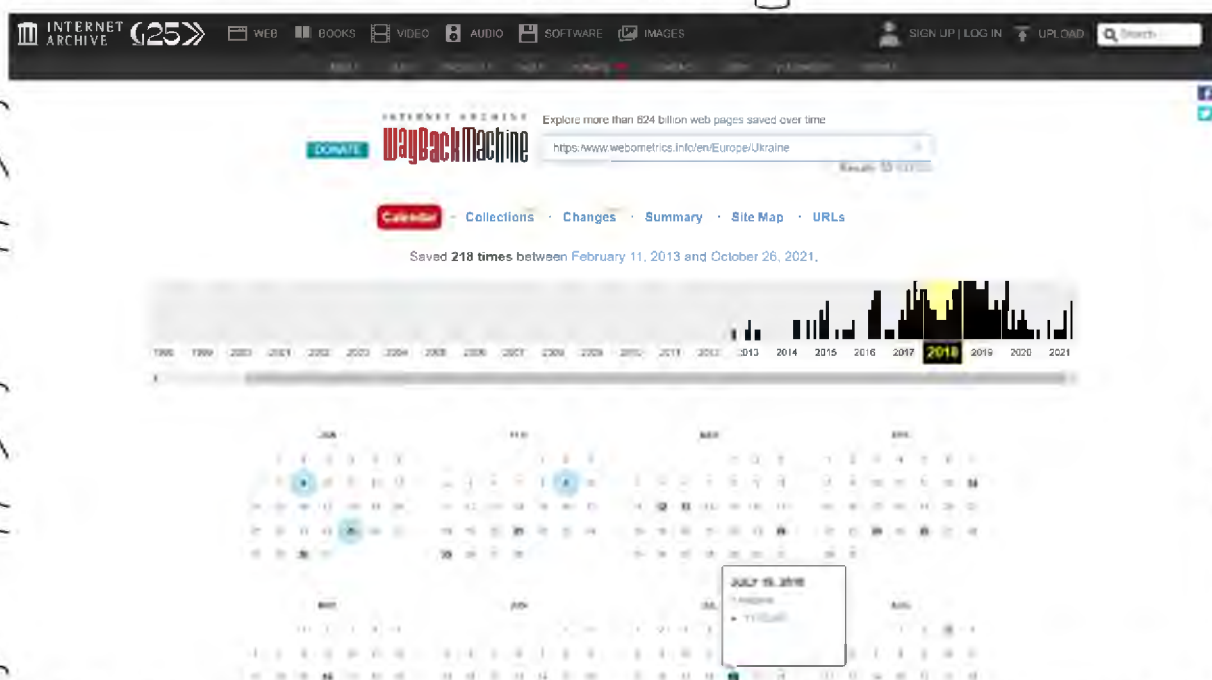


Рис. 2.8 Вебсайт «WayBack machine»

Завдяки ньому було зібрано решту даних за період 2018-2020 рр. Один із готових оброблених CSV файлів статистики з ресурсу Ranking web of universities подано на рисунку 2.9:

	A	B	C	D	E	F	G	H
1	University	ranking	World Ran	Presence	F Impact	Rai	Openness	Excellence Year
2	Київський національний університет Тараса Шевченка	1	1272	759	1727	1741	1590	2018
3	Національний технічний університет України Київський політехнічний інститут	2	1242	475	1479	1530	2615	2018
4	Львівський національний університет Івана Франка	3	1033	1322	1363	2515	2871	2018
5	Сумський державний університет	4	1067	894	1817	1594	2819	2018
6	Харківський національний університет В.Н. Каразіна	5	2258	1196	3339	3595	2365	2018
7	Тернопільський національний економічний університет	6	2300	1851	521	2180	4851	2018
8	Одеський національний університет І.І. Мечникова	7	2455	1552	3170	2129	3000	2018
9	Національний університет Львівська політехніка	8	2489	1420	4184	3648	2434	2018
10	Національний технічний університет Харківський політехнічний інститут	9	2522	963	3601	1660	3257	2018
11	Національний Університет біоресурсів і природокористування України	10	3114	1815	4440	2825	3728	2018
12	Чернівецький національний університет Ю.Федьковича	11	3180	2074	8440	2470	2776	2018
13	Запорізький національний університет	12	3223	1688	1571	2978	4752	2018
14	Національний університет Києво-Могилянська академія	13	3410	2657	4672	2270	4654	2018
15	Дніпропетровський національний університет О.Гончара	14	3518	1620	5458	4865	3958	2018
16	Львівський національний медичний університет	15	3770	4485	3018	2655	3885	2018
17	Харківський національний університет радіоелектроніки	16	3859	4024	10597	3153	3324	2018
18	Ужгородський національний університет	17	3946	1295	5650	2604	3912	2018
19	Вінницький національний технічний університет	18	3997	686	6646	2299	4650	2018
20	Національний медичний університет О.О. Богомольця	19	4084	6442	8644	2474	4164	2018
21	Національний гірничий університет	20	4287	2299	4046	3898	3276	2018
22	Харківський національний університет міського господарства О.М. Бекетова	21	4336	1856	1533	3456	4008	2018
23	Національний авіаційний університет	22	4344	1191	2090	3157	4208	2018
24	Національний педагогічний університет М.П. Драгоманова	23	4439	1635	2674	2150	5008	2018
25	Одеський національний політехнічний університет	24	4439	3498	8183	2379	4754	2018
26	Івано-Франківський національний технічний університет нафти і газу	25	4584	1219	3554	4527	3562	2018
27	Буковинський державний медичний університет	26	4584	3175	2182	4228	4008	2018
28	Київський національний економічний університет	27	4642	1993	7376	1707	5276	2018
29	Донецький національний університет	28	4859	4277	7426	4179	4895	2018
30	Київський національний університет будівництва і архітектури	29	4785	6009	3640	4381	4754	2018
31	Тернопільський державний медичний університет	30	4898	1231	3542	4096	4008	2018
32	Харківський національний економічний університет	31	4924	4543	3784	2994	4008	2018
33	Одеська національна академія харчових технологій	32	5305	4840	3739	3598	4008	2018
34	Донецький національний технічний університет	33	5379	2925	3823	4355	4008	2018

Рис. 2.9 Файл статистики з ресурсу Ranking web of universities за 2018 рік

2.3 OLAP

Аналітична обробка в режимі онлайн або OLAP — це підхід до швидкої відповіді на багатовимірні аналітичні запити у обчислювальній роботі [36].

OLAP є частиною ширшої категорії бізнес-аналітики, яка також охоплює реляційні бази даних, написання звітів і аналіз даних. Типові застосування OLAP включають бізнес-звітність для продажів, маркетинг, управлінську звітність, управління бізнес-процесами, бюджетування та прогнозування, фінансову звітність та інше.

Інструменти OLAP дозволяють користувачам інтерактивно аналізувати багатовимірні дані з різних точок зору. OLAP складається з трьох основних аналітичних операцій: консолідації (згортання), деталізації та нарізки.

Консолідація передбачає об'єднання даних, які можна накопичити та обчислити в одному або кількох вимірах. Наприклад, усі офіси збуту зведені до відділу продажів або відділу продажів, щоб передбачити тенденції продажів. Навпаки, деталізації — це техніка, яка дозволяє користувачам орієнтуватися в деталях.

Наприклад, користувачі можуть переглядати продажі за окремими продуктами, які складають продажі в регіоні. Нарізка — це функція, за допомогою якої користувачі можуть витягувати (нарізати) певний набір даних куба OLAP і переглядати (нарізати) фрагменти з різних точок зору. Ці точки зору іноді називають вимірами (наприклад, перегляд одних і тих самих продажів за продавцем, за датою, або за клієнтом, або за продуктом, або за регіоном тощо).

Бази даних, які налаштовані для OLAP, використовують багатовимірну модель даних, що дозволяє виконувати складні аналітичні та спеціальні запити зі швидким часом виконання. Вони запозичують аспекти навігаційних баз даних, ієрархічних баз даних і реляційних баз даних.

OLAP, як правило, відрізняється від OLTP (онлайн-обробки транзакцій), яка зазвичай характеризується набагато менш складними запитами у більшому обсязі для обробки транзакцій, і використовуються не для цілей бізнес-аналітики чи звітності. У той час як системи OLAP здебільшого оптимізовані для читання,

OLTP має обробляти всі види запитів (читання, вставка, оновлення та видалення).

В основі будь-якої системи OLAP лежить OLAP-куб (також званий «багатовимірним кубом») або гіперкубом). Він складається з числових фактів, які

називаються мірами і класифікуються за розмірами. Міри розміщуються на перетинах гіперкуба, який охоплюється розмірами як векторний простір.

Матричний інтерфейс – це класичний інтерфейс для маніпулювання кубом

OLAP, подібно до зведених таблиць у програмі електронних таблиць, який виконує операції проєкції вздовж вимірів, наприклад агрегування або усереднення.

Метадані куба, як правило, створюються на основі схеми-зірки або схеми сніжинки або сукупності таблиць у реляційній базі даних. Показники виводяться з записів у таблиці фактів, а розміри — з таблиць вимірювань.

Системи OLAP традиційно класифікуються за 3 видами:

1. Багатовимірний OLAP (MOLAP): є класичною формою OLAP, яку іноді називають просто OLAP. MOLAP зберігає дані в оптимізованому багатовимірному масиві, а не в реляційній базі даних. Деякі інструменти MOLAP

вимагають попереднього обчислення та зберігання отриманих даних, наприклад

консолідації – операції, відомої як обробка. Такі інструменти MOLAP зазвичай використовують попередньо розрахований набір даних, який називається кубом даних. Куб даних містить усі можливі відповіді на заданий діапазон запитань.

Завдяки цьому вони дуже швидко відповідають на запити. З іншого боку,

оновлення може зайняти багато часу залежно від ступеня попереднього обчислення. Інші інструменти MOLAP, зокрема ті, що реалізують

функціональну модель бази даних, не обчислюють попередньо отримані дані, а виконують усі обчислення за вимогою, крім тих, які раніше були запитані та

збережені в кеші

2. Реляційний OLAP (ROLAP): працює безпосередньо з реляційними базами даних і не вимагає попереднього обчислення. Базові дані та таблиці

вимірів зберігаються як реляційні таблиці, а нові таблиці створюються для

зберігання агрегованої інформації. Ця методологія покладається на маніпулювання даними, що зберігаються в реляційній базі даних, щоб створити вигляд традиційної функціональності OLAP. По суті, кожна дія нарізки

еквівалентна додаванню речення «WHERE» в оператор SQL. Інструменти

ROLAP не використовують попередньо обчислені куби даних, а натомість подають запит до стандартної реляційної бази даних та її таблиць, щоб повернути дані, необхідні для відповіді на запитання. Інструменти ROLAP мають

можливість задавати будь-які запитання, оскільки методологія не обмежується

вмістом куба. ROLAP також має можливість деталізації до найнижчого рівня

деталізації в базі даних. Хоча ROLAP використовує джерело реляційної бази даних, зазвичай база даних повинна бути ретельно розроблена для використання ROLAP.

3. Гібридний OLAP (HOLAP): через небажаний компроміс між додатковою

вартістю ETL та повільною продуктивністю запитів призвів до того, що більшість комерційних інструментів OLAP тепер використовують підхід «гібридний OLAP» (HOLAP), який дозволяє розробнику моделі вирішувати, яка

частина даних буде зберігатися в MOLAP та яка частина в ROLAP. У галузі

немає чіткої згоди щодо того, що таке «гібридний OLAP», за винятком того, що

база даних буде розділяти дані між реляційним і спеціалізованим сховищем.

Наприклад, для деяких постачальників база даних HOLAP використовуватиме реляційні таблиці для зберігання більшої кількості детальних даних і

використовуватиме спеціалізоване сховище принаймні для деяких аспектів

меншої кількості більш агрегованих або менш детальних даних. HOLAP усуває

недоліки MOLAP і ROLAP шляхом поєднання можливостей обох підходів.

Інструменти HOLAP можуть використовувати як попередньо обчислені куби,

так і реляційні джерела даних.

2.4 SSIS, SSAS та SSRS

Для того, щоб оперувати даними OLAP кубу та створювати різноманітні звіти за ними, можна використовувати такі служби, наприклад, як SSIS, SSAS та SSRS від компанії Microsoft.

SQL Server Служби Integration Services (SSIS) – це платформа для побудови рішень щодо інтеграції та перетворення даних рівня підприємства [37].

Її можна використовувати для вирішення складних бізнес-завдань шляхом копіювання та завантаження файлів, завантаження сховищ даних, очищення та інтелектуального аналізу даних, а також управління об'єктами та даними SQL Server.

SSIS можуть вилучати та перетворювати дані з ряду таких джерел, як файли XML-даних, неструктуровані файли та джерела реляційних даних, а потім завантажувати ці дані в один або кілька реляційних об'єктів.

Служби Integration Services містять широкий набір вбудованих завдань та перетворень, а також графічні засоби для створення пакетів. Крім того, вони включають базу даних каталогу для зберігання, виконання та адміністрування пакетів.

Analysis Services – це засіб аналітичних даних, що використовується у службі підтримки прийняття рішень та бізнес-аналітики [38]. Цей засіб надає можливості моделі семантичних даних корпоративного рівня для бізнес-аналітики (BI), аналізу даних та створення звітів, таких як Power BI, Excel, Reporting Services та інших засобів візуалізації даних.

Три продукти Microsoft, які використовують цей засіб:

1. Azure Analysis Services. Створені як ресурс Azure, ресурси сервера Azure Analysis Services підтримують табличні моделі, DirectQuery, розділи, безпека на рівні рядків, двосторонні зв'язки та переклади.

2. Power BI Premium. Механізм VertiPaq Analysis Services забезпечує програмування, підтримку клієнтських додатків та інструментів для наборів даних Power BI Premium і Premium для наборів користувачів з використанням

клієнтських бібліотек та API, які підтримують протокол XMLA відкритого стандарту. Набори даних Power BI Premium підтримують підключення через кінцеві точки XMLA для операцій лише читання та читання-запису від Microsoft і сторонніх клієнтських програм та інструментів.

3. Служби SQL Server Analysis Services (SSAS). Встановлені як локальний екземпляр або екземпляр сервера VM, SQL Server Analysis Services підтримує табличні моделі на всіх рівнях сумісності (залежно від версії), багатовимірні моделі, аналіз даних і Power Pivot для SharePoint.

Із цих 3 продуктів в даному дослідженні використовуються SQL Server Analysis Services.

Служби звітів SQL Server (SSRS) надають набір локальних інструментів і служб, які створюють, розгортають і керують мобільними та розбитими на сторінки звітами [39].

Рішення SSRS гнучко надає потрібну інформацію потрібним користувачам. Вони можуть переглядати звіти через веббраузер, на своєму мобільному пристрої або електронною поштою.

SQL Server Reporting Services пропонує такий набір продуктів:

1. «Традиційні» звіти з розбивкою на сторінки. Можна створювати різноманітні звіти різних типів та гнучко налаштовувати їхнє відображення: підписи, колір, розмір, вісь, тощо.

2. Нові мобільні звіти з адаптивним макетом, який адаптується до різних пристроїв і різних способів їх зберігання.

3. Сучасний вебпортал, який можна переглядати в будь-якому сучасному браузері. На новому порталі можна впорядковувати та відображати мобільні та розбиті на сторінки звіти SSRS та КІЕ. Також можна зберігати робочі книги Excel на порталі.

2.5 Використані технології та програмні засоби при обробці даних та управлінні сховищем даних

2.5.1 Мова програмування Python.

Python — це інтерпретована мова програмування загального призначення високого рівня [40]. Її мовні конструкції, а також її об'єктно-орієнтований підхід спрямовані на те, щоб допомогти програмістам писати зрозумілий, логічний код для проєктів різного масштабу.

Python динамічно типізується і збирає сміття. Вона підтримує декілька парадигм програмування, включаючи структуроване (зокрема, процедурне), об'єктно-орієнтоване та функціональне програмування. Її часто описують як мову з «батареями» через її повну стандартну бібліотеку. А за допомогою PIP (Package installer for Python) можна встановлювати сторонні бібліотеки з індексу пакунків Python або з інших джерел [41]. Python незмінно вважається однією з найпопулярніших мов програмування [42–45].

2.5.2 Microsoft SQL Server.

Microsoft SQL Server — це система управління реляційними БД (СУБД), яка була розроблена компанією Microsoft [46]. Як сервер баз даних, це програмний продукт, основною функцією якого є зберігання та отримання даних за запитом інших програмних додатків, які можуть працювати як на тому ж комп'ютері, так і на іншому комп'ютері в мережі (включаючи Інтернет).

2.5.3 Microsoft SQL Server Management Studio.

SQL Server Management Studio (SSMS) — це інтегроване середовище для керування будь-якою інфраструктурою SQL: від SQL Server до бази даних SQL Azure [47]. SSMS надає інструменти для налаштування, моніторингу та адміністрування екземплярів SQL Server і баз даних.

SSMS можна використовувати для запитів, проєктування та керування базами даних і сховищами даних, де б вони не були — на локальному комп'ютері чи в хмарі.

2.5.4 Microsoft Excel.

Microsoft Excel — це провідна програма для роботи з електронними таблицями, а також потужний інструмент візуалізації та аналізу даних [48]. Крім цього дане ПЗ може широко використовуватись як рішення для бізнес-аналізу: підключатися до БД чи СД, будувати графіки, прогнози, розрахунки, КПЕ, тощо.

2.5.5 Microsoft Visual Studio.

Інтегроване середовище розробки (IDE) — це багатофункціональна програма, яка підтримує багато аспектів розробки програмного забезпечення [49]. Visual Studio IDE — це засіб, який можна використовувати для редагування, налагодження та створення коду, а потім — і для публікації програми. Окрім стандартного редактора та налагоджувача, які надають більшість IDE, Visual Studio включає різноманітні компілятори, інструменти завершення коду, дизайну інтерфейсу та багато інших функцій для покращення процесу розробки програмного забезпечення.

2.5.6 Microsoft Visual Studio Code.

Visual Studio Code — це легкий і потужний редактор вихідного коду, який доступний для таких операційних систем як Windows, macOS та Linux [50]. Він поставляється з вбудованою підтримкою JavaScript, TypeScript і Node.js. Цей редактор має величезну екосистему розширень для інших мов (наприклад, C++, C#, Java, Python, PHP, Go), а також середовищ виконання (наприклад, .NET і Unity).

2.6 Використані технології та програмні засоби при створенні інформаційно-аналітичної системи

При розробці інформаційно-аналітичної системи було вирішено створити клієнт-серверну систему, яка буде написана мовою JavaScript з використанням front-end на Vue.JS. Вебсайт відправлятиме запити на отримання даних по API на back-end Node.JS і на основі отриманих даних відображати різноманітні графіки,

побудовані за допомогою компоненти Flexmonster – JavaScript Pivot Table & Charts Component та таблицю кореляцій для ЗВО з загальним значенням.

2.6.1 Мова програмування JavaScript.

JavaScript (JS) — це легка, інтерпретована або just-in-time (JIT) скомпільована мова програмування з функціями першого класу [51]. Хоча мова найбільш відома як мова сценаріїв для вебсторінок, багато середовищ, які не є браузерами, також використовують її, наприклад, Node.js, Apache CouchDB та Adobe Acrobat. JavaScript — це заснована на прототипах, багатопарадигмальна, однопоточна динамічна мова, яка підтримує об'єктно-орієнтований, імперативний та декларативний (наприклад, функціональне програмування) стилі.

2.6.2 API.

Інтерфейс програмування прикладних програм, або API, дозволяє компаніям відкривати дані та функціональні можливості своїх програм для зовнішніх сторонніх розробників, бізнес-партнерів і внутрішніх відділів у своїх компаніях [52]. Це дозволяє службам і продуктам спілкуватися один з одним і використовувати дані та функціональні можливості один одного через документований інтерфейс. Розробникам не потрібно знати, як реалізовано API; вони просто використовують інтерфейс для спілкування з іншими продуктами та послугами. За останнє десятиліття використання API зросло до такої міри, що багато з найпопулярніших вебдодатків сьогодні були б неможливими без API.

2.6.3 Node.JS.

Node.js — це середовище виконання JavaScript з відкритим вихідним кодом та кросплатформністю [53]. Це популярний інструмент практично для будь-яких проєктів. Node.js запускається на двигуні JavaScript V8, що є ядром Google Chrome, поза браузером. Це дозволяє Node.js бути дуже продуктивним. Програма Node.js виконується в одному процесі, не створюючи новий потік для кожного запиту. Node.js надає в стандартній бібліотеці набір асинхронних примітивів вводу-виводу, які запобігають блокуванню коду JavaScript, і загалом бібліотеки в Node.js написані з використанням неблокуючих парадигм, що

робить поведінку блокування скоріше винятком, ніж нормою. Крім цього дане середовище є зручним для full-stack розробників, оскільки для реалізації серверної та клієнтської частини використовується одна і та ж мова програмування.

2.6.4 Front-end framework Vue.JS.

Vue це прогресивна платформа для створення інтерфейсів користувача (UI) [54]. На відміну від інших монолітних фреймворків, Vue був розроблений з нуля так, щоб його можна було поступово розширювати. Основна бібліотека зосереджена лише на шарі перегляду, її легко підібрати та інтегрувати з іншими бібліотеками або існуючими проєктами. З іншого боку, Vue також ідеально здатний працювати над складними односторінковими додатками в поєднанні з сучасними інструментами та допоміжними бібліотеками.

2.6.5 Flexmonster – JavaScript Pivot Table & Charts Component.

Даний пачунок дозволяє створювати зведені таблиці і пропонує широкий спектр функцій для візуалізації даних: налаштування звіту, об'єднання, фільтрування, сортування та групування даних [55]. Є можливість переходу угору та вниз, щоб переглянути дані з різних точок зору. Присутнє збереження результатів, експорт в PDF, Excel, HTML, CSV або PNG одним кліком.

НУБІП України

3 РОЗРОБКА СИСТЕМИ

3.1 Сховище даних

Розроблене сховище даних має тип «зірка».

ER-модель сховища зображено на рисунку 3.1.

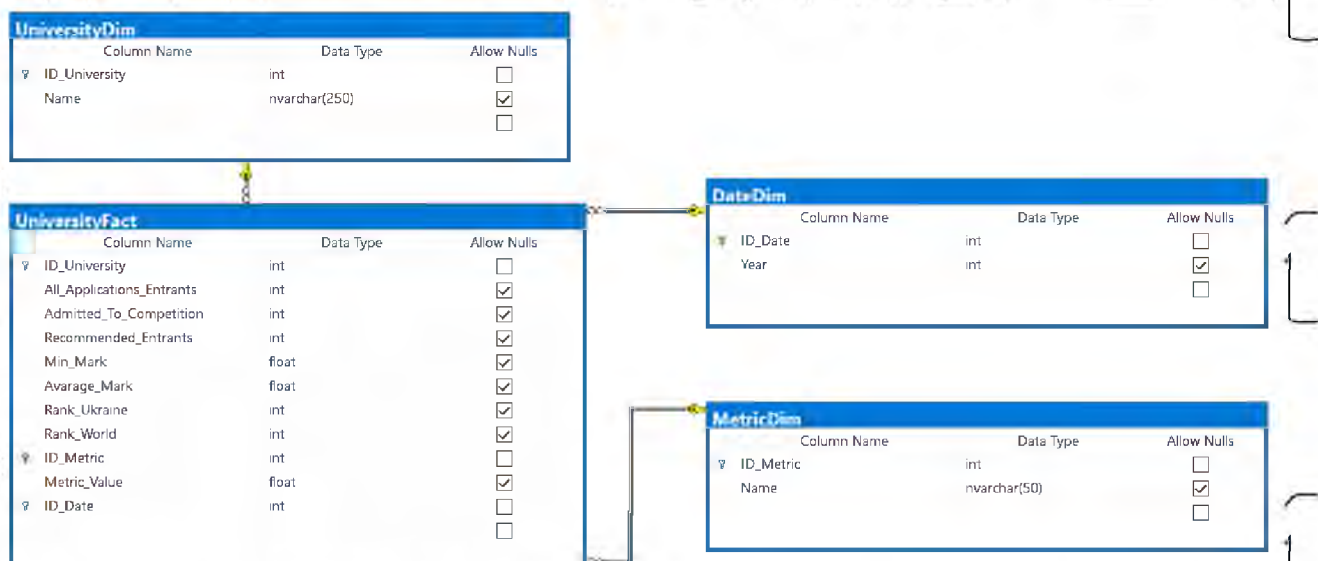


Рис. 3.1 ER-модель сховища даних

Діаграму сховища даних в середовищі SSMS наведено на рисунку 3.2:

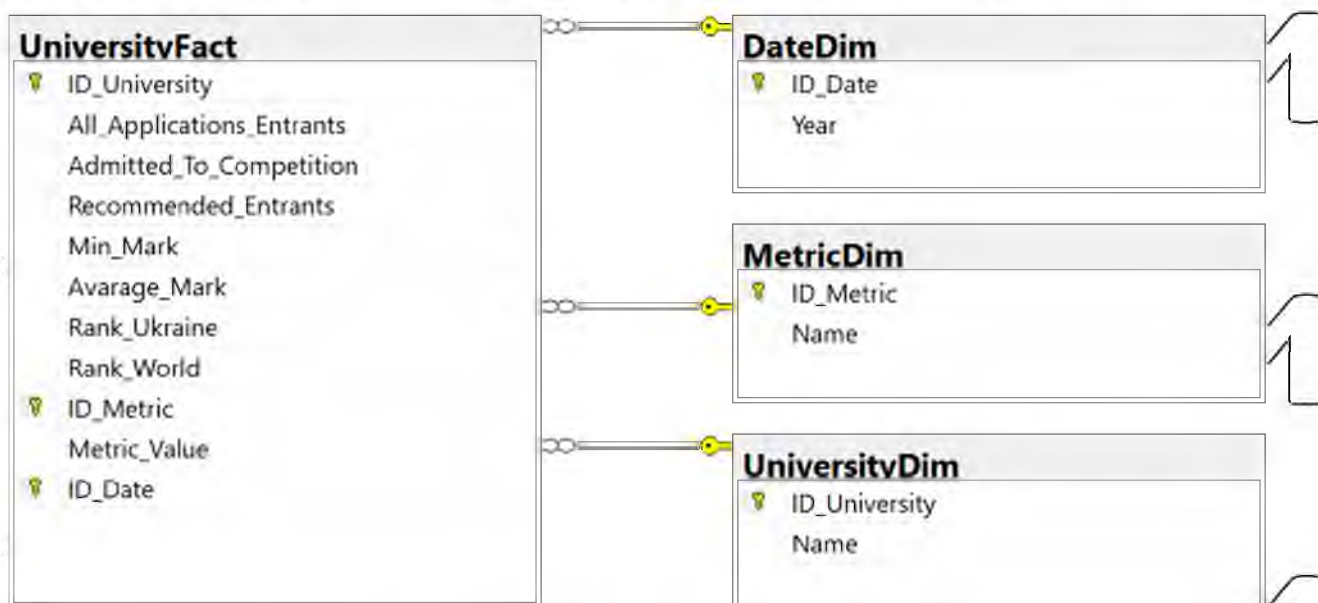


Рис. 3.2 Діаграма сховища даних

Опис таблиць та зв'язків у СД:

Таблиця «DateDim» – це таблиця виміру часу, в даному випадку, років. Зв'язок один до багатьох з таблицею «UniversityFact».

Таблиця «MetricDim» – це таблиця виміру метрик. В даному випадку, в ній знаходяться 4 метрики, які було отримано з ресурсу Ranking web of universities, а також 5 метрика – загальне значення з 4-х попередніх. Зв'язок один до багатьох з таблицею «UniversityFact».

Таблиця «UniversityDim» – це таблиця виміру ЗВО. В даній таблиці знаходяться всі ЗВО, для яких створено дані-факти. Зв'язок один до багатьох з таблицею «UniversityFact».

Таблиця «UniversityFact» – це таблиця фактів університетів. За комбінацією ключів з таблицями «MetricDim», «UniversityDim» та «DateDim» можна отримати такі дані про ЗВО: загальна кількість поданих заяв абітурієнтами, кількість допущених абітурієнтів, кількість рекомендованих абітурієнтів, мінімальний та середній бали абітурієнта, позиція ЗВО у рейтингу Ranking web of universities у світі та в Україні, значення метрики за певний рік.

3.2 Обробка даних з ресурсів

Як вже було згадано в розділі 2.2 «Опис вузлів, з яких збирається інформація», отримані дані з ЄДЕБО та Ranking web of universities потрібно було обробити для подальшого завантаження у СД. Для цієї мети було розроблено програму з використанням мови програмування Python, яка за один запуск аналізує пару CSV файлів з ЄДЕБО та Ranking web of universities за один рік, знаходить схожі ЗВО за назвою (за найбільшим коефіцієнтом подібності рядків), створює новий об'єднаний файл, в який заносить рядок ЗВО з фрагментами даних з вище згаданих двох файлів. Таким чином було оброблено 4 пари файлів. Потім по чергово ці пари були об'єднані в один файл за період 2018-2021 рр. Приклад коду такої програми наведено в додатку А. Було розраховано значення загальної метриками за формулами з використанням програмного засобу

Microsoft Excel!
 - для даних за період 2018-2020 рр.:

$$G = P * 0.5 + V * 0.5 + T * 0.1 + E * 0.35,$$

(1)

де G – загальне значення метрики;
 P – присутність;
 V – видимість;

T – прозорість;

E – висока якість.

- для даних за 2021 р.:

$$G = V * 0.5 + T * 0.1 + E * 0.4,$$

(2)

де G – загальне значення метрики;

V – видимість;

T – прозорість;

E – висока якість.

Проте, отримані об'єднані дані необхідно було перетворити в структуру, подібну до структури таблиць в СД, щоб можна було скористатися вбудованими інструментами імпорту даних. Тому було розроблену додаткову програму, яка створює окремі файли з ІД та назвою ЗВО для таблиці «UniversityDim» та файл з закодованими даними відносно ІД ЗВО, метрик та реків. Приклад коду програми наведено в додатку Б.

Таблиця-вимір метрик була заповнена за допомогою SQL запити, який наведений на рисунку 3.5:

```
USE [UniversitiesMetricsDWH]
GO
INSERT INTO [dbo].[MetricDim]([ID_Metric], [Name])
VALUES
(N'0',N'All'),
(N'1',N'Presence Rank*'),
(N'2',N'Impact Rank*'),
(N'3',N'Openness Rank*'),
(N'4',N'Excellence Rank*')
```


Рис. 3.3 SQL-запит заповнення таблиці «MetricDim»

Таблиця-вимір ЗВО та таблиця фактів були заповнені за допомогою вбудованої функції «Імпортувати» в SSMS. Всього в таблиці фактів налічується 920 записів (рис. 3.4) для 46 ЗВО.

ID_University	All_Applications_Entrants	Admitted_To_Competition	Recommended_Entrants	Min_Mark	Average_Mark	Rank_Ukraine	Rank_World	ID_Metric	Metric_Value	ID_Date
1	2882	2320	282	185,487	190,455	8	2469	0	3379,7	1
2	2882	2320	282	185,487	190,455	8	2469	1	1420	1
3	2882	2320	282	185,487	190,455	8	2469	2	4184	1
4	2882	2320	282	185,487	190,455	8	2469	3	3648	1
5	2882	2320	282	185,487	190,455	8	2469	4	2434	1
6	2400	1808	196	176,358	185,347	16	3859	0	6978,4	1
7	2400	1808	196	176,358	185,347	16	3859	1	4024	1
8	2400	1808	196	176,358	185,347	16	3859	2	10597	1
9	2400	1808	196	176,358	185,347	16	3859	3	3153	1
10	2400	1808	196	176,358	185,347	16	3859	4	3324	1
11	2336	1913	175	180,846	187,736	3	1933	0	2014,65	1
12	2336	1913	175	180,846	187,736	3	1933	1	1322	1
13	2336	1913	175	180,846	187,736	3	1933	2	1383	1
14	2336	1913	175	180,846	187,736	3	1933	3	2515	1
15	2336	1913	175	180,846	187,736	3	1933	4	2873	1
16	2169	1566	161	182,2	190,467	1	1272	0	1562,25	1
17	2169	1566	161	182,2	190,467	1	1272	1	759	1
18	2169	1566	161	182,2	190,467	1	1272	2	1727	1

Рис. 3.4 Фрагмент записів в таблиці фактів

Крім цього було створено проект з використанням інструментів BI: «Analysis Services Multidimensional and Data Mining» у IDE MS Visual Studio 2019. В ньому було підключено джерело даних (СД), створено види джерела (Data Source Views), підключено виміри та розгорнуто куб MOLAP (рис. 3.5).

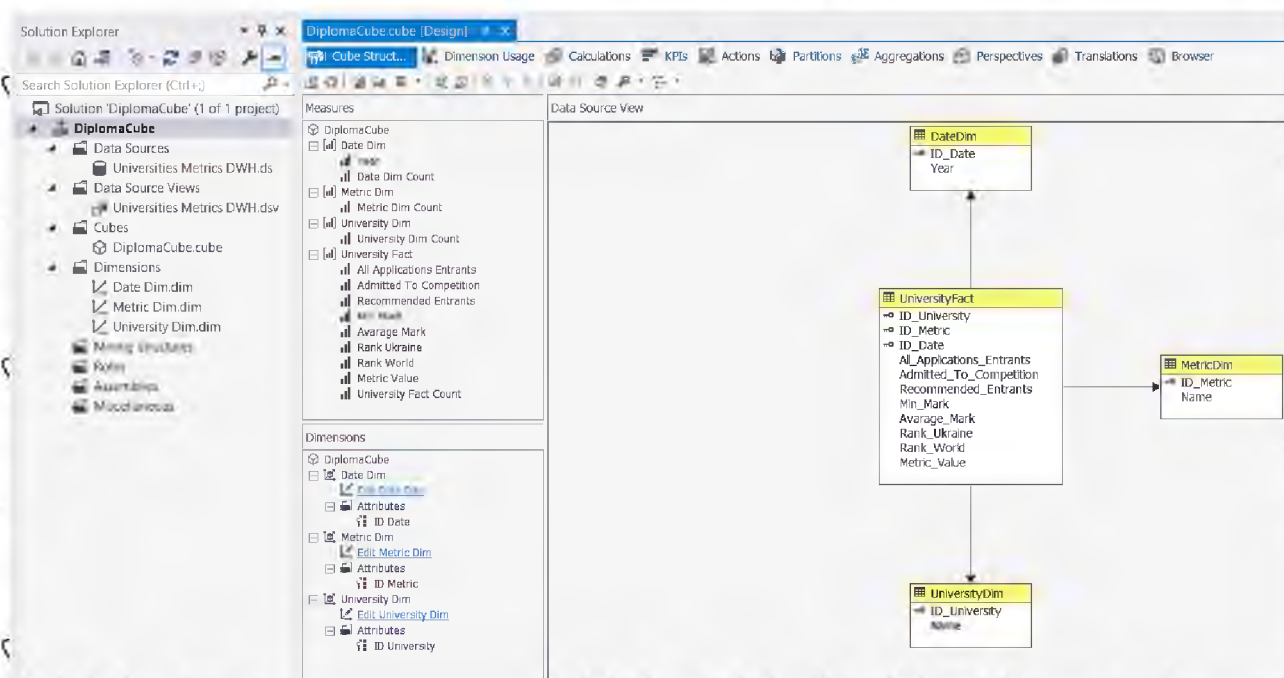


Рис. 3.5 Розгорнутий куб з вимірами

Також було створено такі КПЕ: середнього балу всіх ЗВО за 2018-2021 рр.

(рис. 3.6); загальної кількості поданих заяв абітурієнтами та їхнього середнього балу НУБІП за період 2018-2021 рр. та окремо для 2021 року.

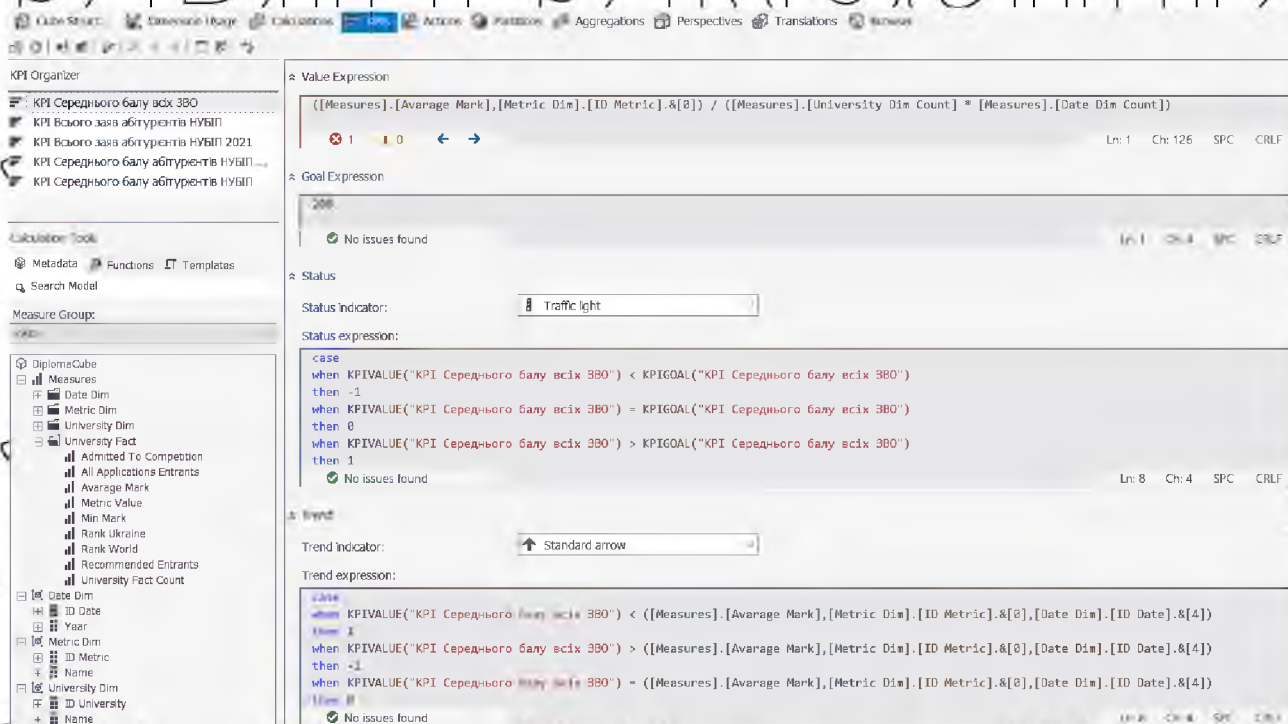


Рис. 3.6 Налаштування для КПІ Середнього балу абітурієнтів для 46 ЗВО

Ще було створено проєкт типу Report service project (рис. 3.7). Додано Data source, Dataset. Створено звіти, які демонструють співставлення кількості заяв абітурієнтів до загального значення метрики для різних наборів ЗВО.



Рис. 3.7 Проєкт сервісу звітів

3.3 Розробка інформаційно-аналітичної системи «клієнт-сервер»

3.3.1 Серверна частина.

Для розробки системи було створено проєкт у Visual Studio Code з окремими каталогами для серверної і клієнтської частини. Для керування пакунками використовувався Node Package Manager (NPM). Налаштовано синхронізацію з Github репозиторієм [56]. Структуру проєкту зображено на рисунку 3.8:

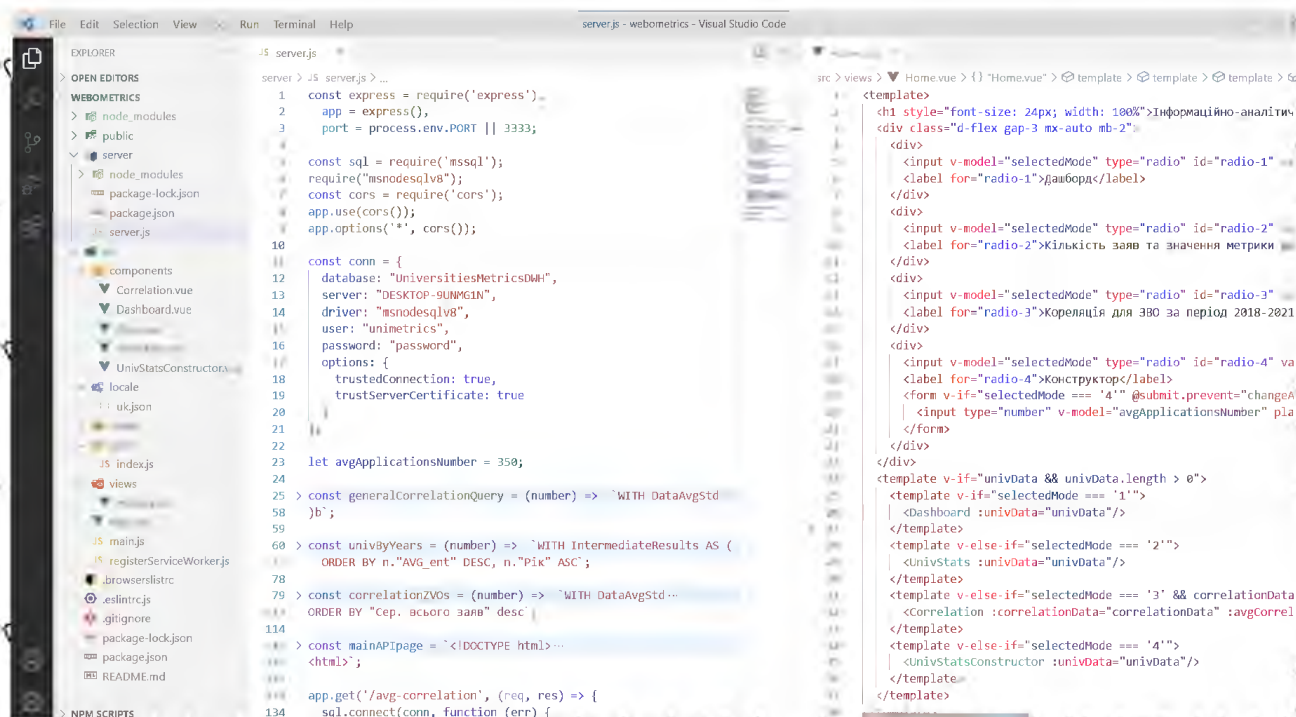


Рис. 3.8 Структура проєкту системи «клієнт-сервер»

Для реалізації серверної частини було використано Node.js. Програма підключається до сховища даних та за допомогою API запитів (для яких виконуються різні SQL запити) надає необхідні дані. Налаштовано Cross-Origin Resource Sharing (CORS), щоб можна було здійснювати запити з іншого вебсайту. Крім цього, можна змінювати параметр фільтра середньої кількості заяв ЗВО за допомогою окремого POST запиту. В ньому відбувається перевірка вхідного значення на правильність і у протилежному випадку видається відповідь з кодом 422. Код цього запиту, а також код запиту отримання всіх табличних даних статистики по ЗВО за період 2018-2021 рр. наведено на

Рис. 3.9:

```

app.get('/univ-stats', (req, res) => {
  sql.connect(conn, function (err) {
    if (err) console.log(err);
    var request = new sql.Request();
    request.query(univByYears(avgApplicationsNumber), function (err, recordset) {
      if (err) console.log(err);
      res.send(recordset.recordset);
    });
  });
});

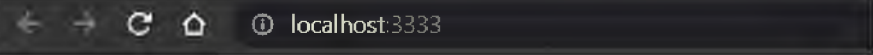
app.use(express.json());
app.post('/set-avg-applications', function (request, response) {
  if(!request.body) return response.sendStatus(400);
  if(request.body?.number && /^\d+$/ .test(request.body.number)) {
    avgApplicationsNumber = parseInt(request.body.number);
    console.log(`New avg number ${avgApplicationsNumber}`);
    return response.json({number: avgApplicationsNumber});
  }else{
    return response.sendStatus(422);
  }
});

app.get('/avg-application-number', (req, res) => {
  res.json({number: avgApplicationsNumber});
});

```

Рис. 3.9 Приклади функції API запитів

Початкова сторінка сервера містить в собі коротку документацію про наявні адреси та їхнє призначення (рис. 3.10).



Документація API:

- [GET] [/univ-correlations](#): отримати кореляцію по ЗВО.
- [GET] [/avg-correlation](#): отримати загальне число кореляції.
- [GET] [/univ-stats](#): отримати статистику по ЗВО.
- [GET] [/avg-application-number](#): отримати середнє число заяв абітурієнтів.
- [POST] [/set-avg-applications](#): встановити середнє число заяв абітурієнтів.

Рис. 3.10 Головна сторінка сервера

Як можна побачити, є адреси для отримання повної статистики по ЗВО, середнього числа заяв абітурієнтів (дане число використовується як фільтр для SQL запитів), кореляції між кількістю поданих заяв абітурієнтами та метриками ЗВО. Кореляція розраховується за допомогою формули коефіцієнта кореляції Пірсона [57] та обчислюється в SQL запиті (додаток В):

$$r_{xy} = \frac{\sum_{i=1}^m (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^m (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^m (y_i - \bar{y})^2}} \quad (3)$$

де r_{xy} – коефіцієнт кореляції Пірсона,

$x_1 \dots x_m, y_1 \dots y_m$ – вибірки даних;

\bar{x}, \bar{y} – вибіркові середні.

3.3.2 Клієнтська частина.

Клієнтська частина була розроблена з допомогою використання Front-end Framework Vue.js 3 версії. Для побудови та відображення графіків за отриманими з API даними використовується навунок Flexmonster – JavaScript Pivot Table & Charts Component. Кожен екземпляр Flexmonster має задані власні налаштування: джерело даних, вид звіту, сортування, агрегацію даних, тощо. При початковому відкритті вебсайту завантажуються необхідні дані за допомогою методу «loadData()» з використанням вбудованої JavaScript функції «fetch()» (рис. 3.11). В цьому ж методі отримується від сервера число-параметр кількості середніх поданих заяв абітурієнта в рамках одного ЗВО і встановлюється у якості змінної, яка використовується для фільтрування даних.

```
loadData(){
  //load all needed data
  fetch("http://localhost:3333/avg-application-number")
    .then(response => response.json())
    .then(jsonResponse => this.avgApplicationsNumber = jsonResponse && jsonResponse.number ? jsonResponse.number : 350);
  fetch("http://localhost:3333/univ-stats")
    .then(response => response.json())
    .then(jsonResponse => this.univData = jsonResponse || []);
  fetch("http://localhost:3333/univ-correlations")
    .then(response => response.json())
    .then(jsonResponse => this.correlationData = jsonResponse || []);
  fetch("http://localhost:3333/avg-correlation")
    .then(response => response.json())
    .then(jsonResponse => jsonResponse ? this.avgCorrelation = +((jsonResponse[0] * 100).toFixed(2)) : this.avgCorrelation = "");
}
```

Рис. 3.11 Завантаження даних при відкритті вебсайту

Всередині вебзастосунку використовуються радіо-кнопки, які дозволяють перемикається між різними панелями звітів. Доступно 4 опції: дашборд (загальна панель звітів), кількість заяв та значення метрики за період 2018-2021, кореляція для ЗВО за період 2018-2021 та конструктор звітів. При зміні вибраної опції сторінка не перезавантажується повністю, а оновлюється лише частина секції.

Приклад відкритої вкладки дашборд наведено на рисунку 3.12. На ній

НУБІП України

відображаються 4 звіти з використанням лінійних графіків, кругової діаграми та стовпчикової діаграми з лінією.

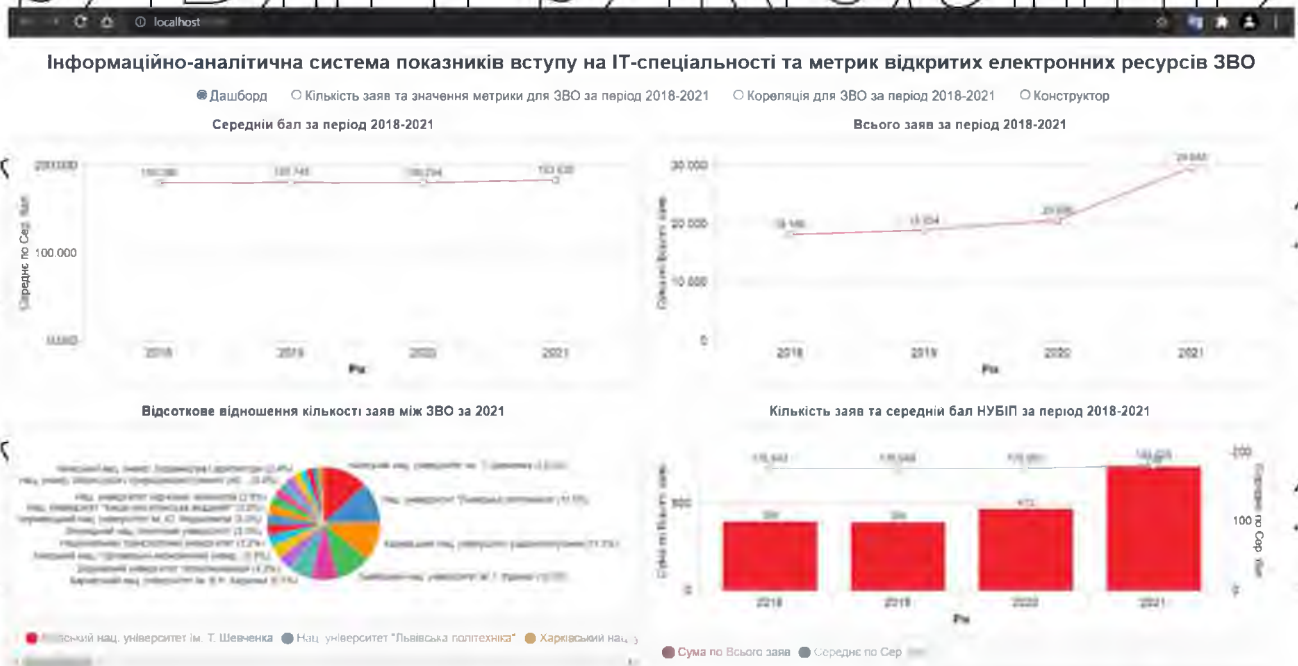


Рис. 3.12 Початкова сторінка вебсистеми

Приклад вкладки конструктора звітів наведено на рисунку 3.13. В даному розділі можна власноруч створювати звіт за потреби та модифікувати число-параметр для запитів у відповідному полі біля назви розділу.



Рис. 3.13 Розділ «Конструктор»

НУБІП України

4 РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

4.1 Аналіз отриманих КПЕ

В процесі дослідження було отримано КПЕ та різноманітні графічні звіти. На рисунку 4.1 наведено 5 отриманих КПЕ в проєкті SASS у Visual Studio

Display Structure	Value	Goal	Status	Trend
KPI Всього заяв абітурієнтів НУБІП	495.25	570.277173913043		↑
KPI Всього заяв абітурієнтів НУБІП 2021	716	775.173913043478		↑
KPI Середнього балу абітурієнтів НУБІП	177.22125	178.625836956522		↑
KPI Середнього балу абітурієнтів НУБІП 2021	181.026	181.128456521739		↑
KPI Середнього балу всіх ЗВО	178.625836956522	200		↑

Рис. 4.1 Отримані КПЕ

Перший КПЕ відображає кількість всіх заяв абітурієнтів НУБІП за період 2018-2021 рр. Як значення цілі було обрано середню кількість заяв абітурієнтів 46 ЗВО за цей період. Як можна побачити, індикатор вказує на те, що отримане значення знаходиться нижче цільового, проте має зростаючий тренд.

Другий КПЕ створено також для всіх заяв абітурієнтів НУБІП, але вже лише за 2021 рік. Значення цілі: середня кількість заяв абітурієнтів 46 ЗВО за 2021 рік. Отриманий результат аналогічний такому, який було отримано в попередньому КПЕ. Крім цього, можна помітити значне збільшення середньої кількості заяв.

Створені за аналогією наступні два КПЕ для середнього балу абітурієнтів НУБІП за 2018-2021 рр. та за 2021 р. мають такий же висновок.

КПЕ середнього балу всіх 46 ЗВО показує 179.63 бали, що є хорошим результатом враховуючи вибірку різноманітних у напрямках ЗВО з усієї країни.

Тренд має зростаючий показник.

4.2 Аналіз отриманих графічних звітів

За допомогою розробленої інформаційно-аналітичної системи було

НУВБІП УКРАЇНИ

Отримано такі графічні звіти

1. Значення загального середнього балу абітурієнтів за період 2018-2021 рр. для 46 ЗВО (рис. 4.2). Даний звіт демонструє загальну зростаючу динаміку цього показника.

Середній бал за період 2018-2021

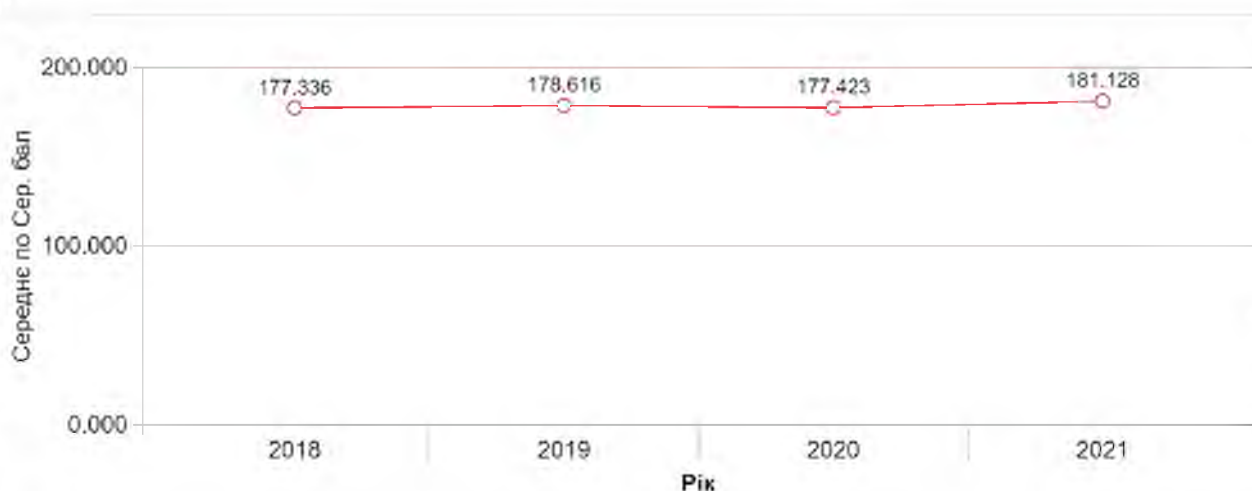


Рис. 4.2 Загальне значення середнього балу абітурієнтів за період 2018-2021 рр.

2. Всього заяв абітурієнтів за період 2018-2021 рр. (рис. 4.3) для 46 ЗВО. Звіт також демонструє загальну зростаючу тенденцію. Крім цього, можна відмітити значне збільшення кількості заяв у 2021 році в порівнянні з попереднім періодом.

Всього заяв за період 2018-2021

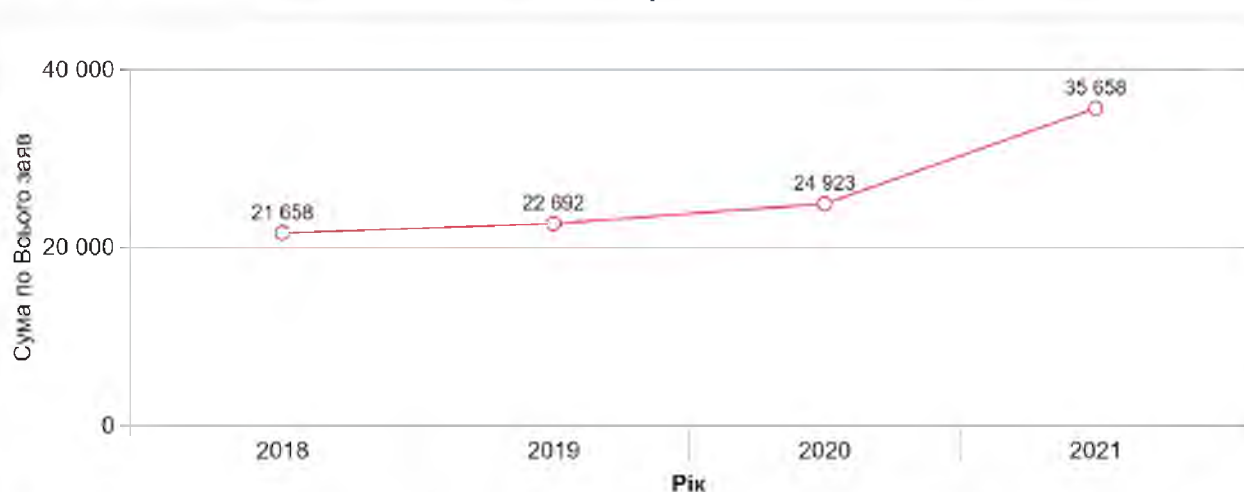


Рис. 4.3 Загальна кількість поданих заяв абітурієнтами за період 2018-2021 рр.

3. Відсоткове відношення кількості заяв між 46 ЗВО за 2021 рік (рис. 4.4).

Цей звіт показує, що сумарно більше третини заяв (38,2%) припадає на

Відсоткове відношення кількості заяв між ЗВО за 2021

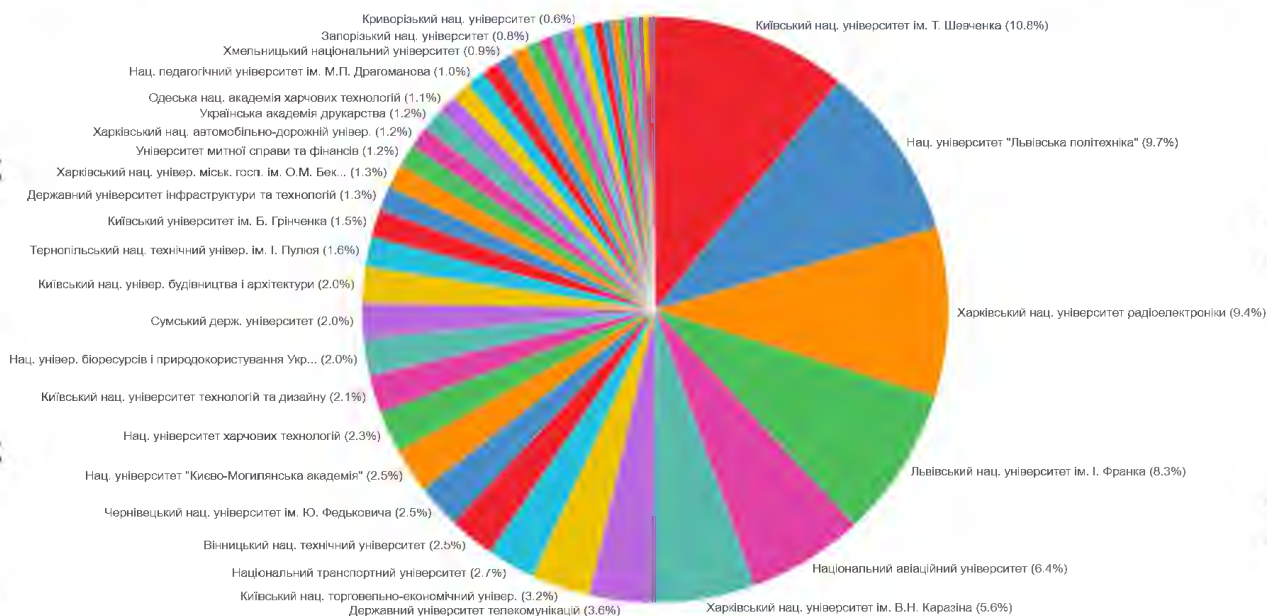


Рис. 4.4 Відсоткове відношення кількості заяв між ЗВО за 2021 рік

4.7 Комплексний звіт, який відображає значення середнього балу та кількості поданих заяв абітурієнтами НУБІП за період 2018-2021 рр. (рис. 4.5). Він показує аналогічну ситуацію загальних звітів (див. рис. 4.2-4.3) – кількість заяв і середній бал абітурієнтів зростають.

Кількість заяв та середній бал НУБІП за період 2018-2021



Рис. 4.5 Кількість заяв та середній бал абітурієнтів НУБІП за період 2018-2021 рр.

5. Комплексний звіт (рис. 4.6), що відображає загальну картину кількості поданих заяв абітурієнтами за спаданням середнього значення цього

показник та загального значення метрики (менше чітко – краще) за період 2018-2021 рр. для ЗВО України. Динаміка загального значення метрики зі зменшенням кількості поданих абітурієнтами заяв має нестабільний зростаючий характер.

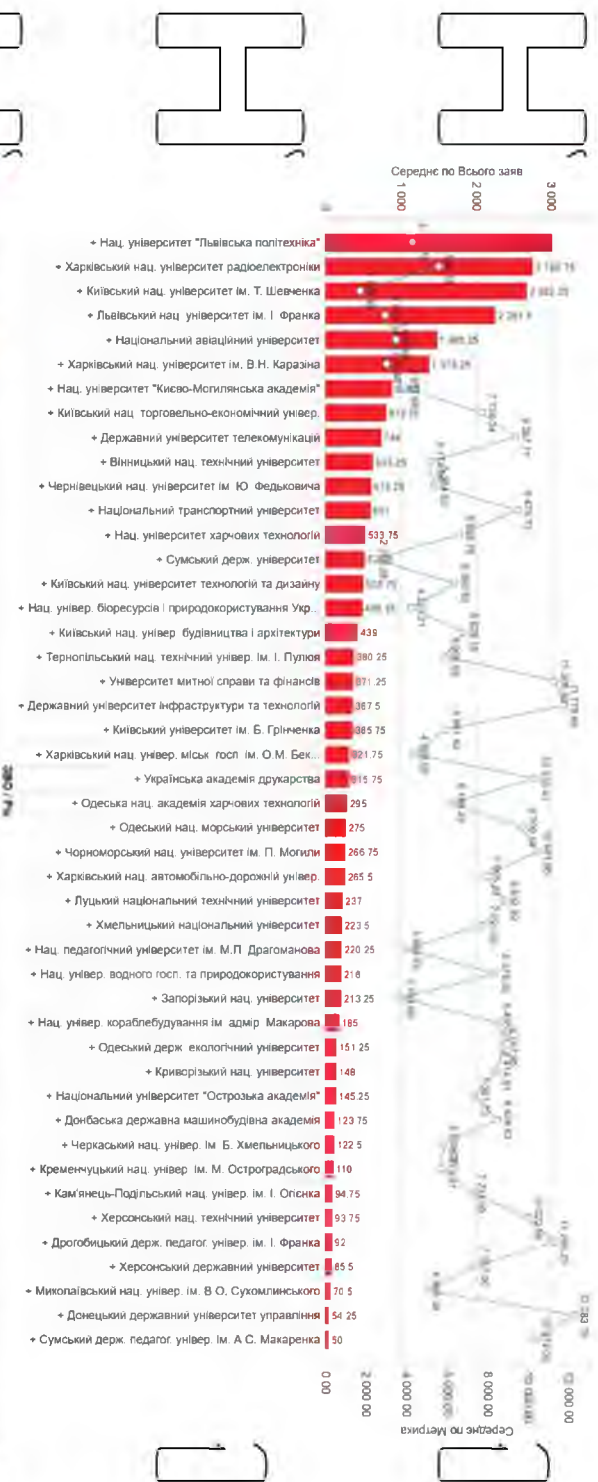


Рис. 4.6 Кількість поданих заяв абітурієнтами та загальне значення метрики за період 2018-2021 рр.

6. Звіт для перших 7 ЗВО зі списку вище, який покаже відношення кількості поданих заяв та загального значення метрики (менше число – краще) за період 2018-2021 рр. з деталізацією за кожним роком (рис. 4.7) у вигляді двох стовпців. Після аналізу даного звіту можна зробити

припущення про певну закономірність: зі збільшенням кількості заяв абітурієнтів значення метрики зменшується (показується). Проте спостерігається і виключення. Крім цього, можна побачити, що у Національній школі університету «Львівська політехніка» у 2020 році було значне погіршення показника загального значення метрики.

НУБІП України

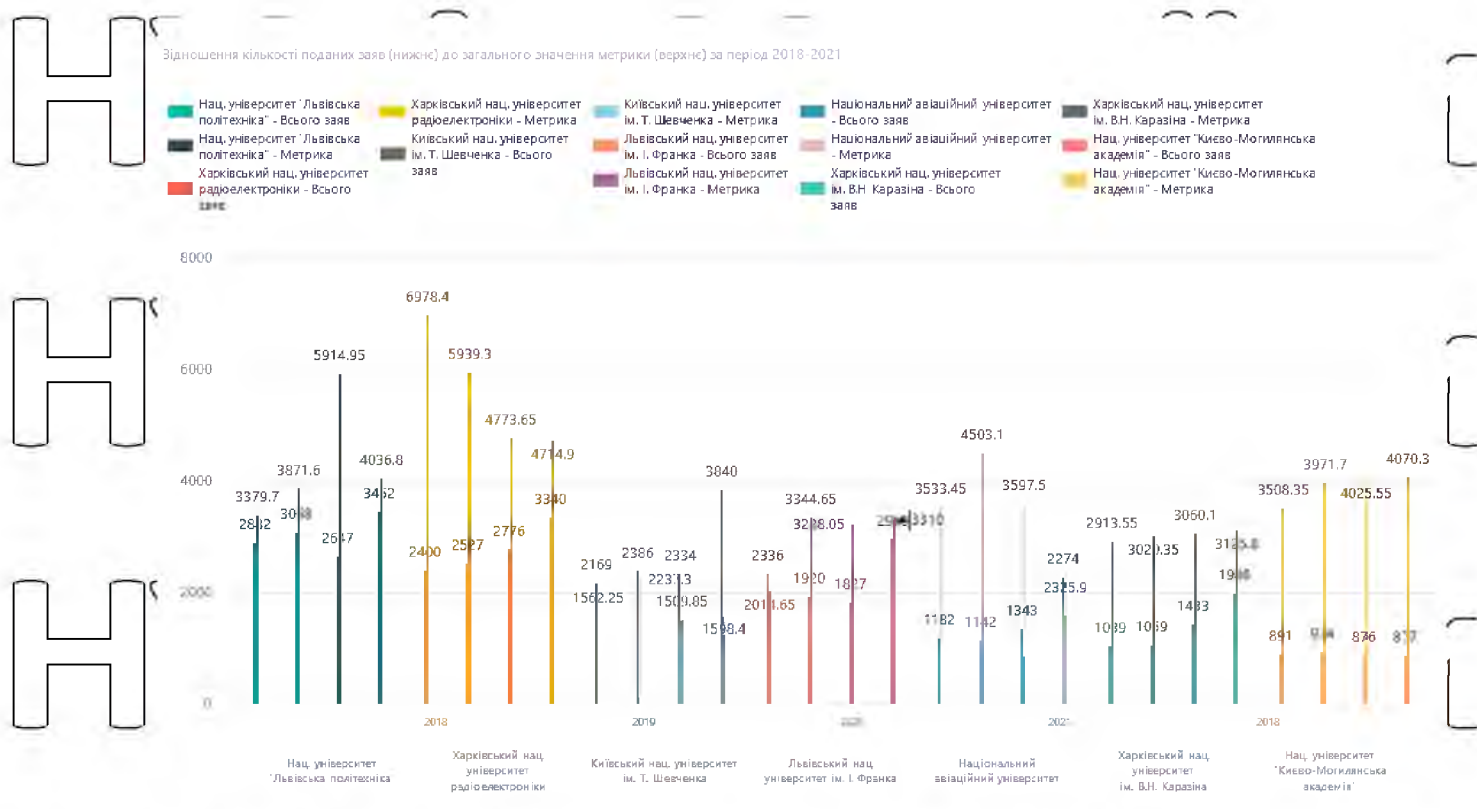


Рис. 4.7 Відношення кількості поданих заяв до загального значення метрики для 7 ЗВО за період 2018-2021 рр.

4.3 Аналіз показників кореляції

За допомогою отриманої інформації від сервера було виведено таблицю кореляції між кількістю поданих заяв абітурієнтами та загальним значенням метрики для ЗВО за період 2018-2021 рр. і розраховано загальний коефіцієнт кореляції з характеристикою (рис. 4.8). Проте було вирішено аналізувати вибірку з середнім значенням кількості заяв ЗВО більше 350, оскільки для ЗВО з малою кількістю заяв існує велика похибка, коли незначна зміна кількості заяв може суттєво змінити коефіцієнт кореляції в ту чи іншу сторону. Тому теловна гіпотеза дослідження перевіряється на вибірці з 21 ЗВО, які було відсортовано за спаданням середньої кількості поданих заяв.

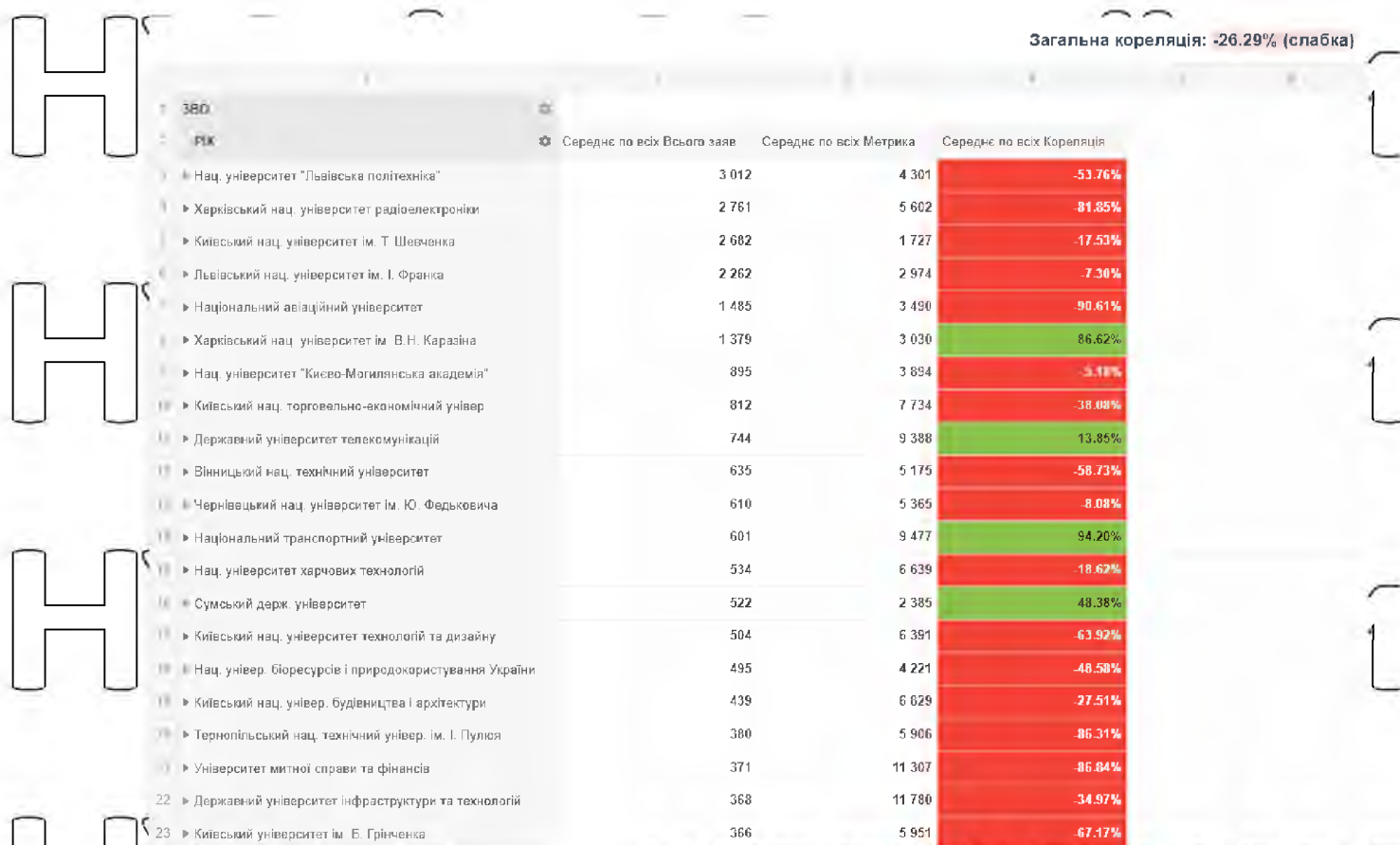


Рис. 4.8 Таблиця кореляції ЗВО

З отриманих даних видно, що для більшості ЗВО з вибірки справджується гіпотеза «зі збільшенням кількості заяв покращується (зменшується) загальне значення метрики» з різною часткою (від -5.18% до -90,61%). Загальне значення коефіцієнту кореляції для вибірки з 21/ЗВО становить -26.29% (слабка). Проте наявні і виключення з цього правила: ХНУК, ДУТ, НТУ та СумДУ. Крім цього, якщо розглядати СумДУ на фоні загальної статистики, то за середнім значенням загальної метрики за період 2018-2021 рр. він знаходиться на другому місці після КНУ і перед ЛНУ та ХНУК, проте має малу кількість заяв абітурієнтів в порівнянні з цією трійкою. Хоча це не означає, що не потрібно підтримувати і покращувати стан своїх інтернет представництв, як це було з'ясовано з проаналізованих раніше джерел.

ВИСНОВКИ

У ході виконання дослідження було описано предметну область, встановлено, що таке Ranking web of universities, за якими метриками даний

ресурс єдине ЗВО, обґрунтовано доцільність використання саме цього ресурсу як основного для дослідження гіпотези «зі збільшенням кількості заяв абітурієнтів на ІТ спеціальність 122 «Комп'ютерні науки» зменшується

(покращується) загальне значення метрики». Проаналізовано існуючі дослідження на тематику дослідження. За допомогою даних джерел було

встановлено, що важливо мати якісний і наповнений вебсайт ЗВО (встановлено навіть відповідний закон з вимогами українським законодавством), який відповідає стандартам доступності; мати представництва у соціальних мережах;

сприяти посиленню публікаційної активності своїх співробітників у

наукометричних базах даних, оскільки це сприяє покращенню рейтингової позиції організації. Крім цього, було з'ясовано, що українські ЗВО також слідкують за своїми позиціями в даному рейтингу і намагаються їх покращити.

Таку зацікавленість у рейтингу Ranking web of universities можна пояснити тим,

що високу оцінку відомим світовим ресурсом можна використати як один із факторів у формуванні іміджу ЗВО, що в свою чергу може бути використано під час вступних кампаній.

Було зібрано необхідні дані з ресурсу ЄДЕБО та World Ranking Web of

Universities, оброблено за допомогою створених програм в об'єднаний

результуючий файл. Описано основні теоретичні відомості про моделювання системи, використані технології та програмні засоби, наведено різні діаграми, що окреслюють розроблену систему, описано інформаційні вузли. Створено

сховище даних, яке було заповнене отриманим раніше файлом з обробленою

статистикою для 46 ЗВО з вище згаданих ресурсів. Створено проекти SSAS та SSRS, за допомогою яких було сформовано різні KPI та звіти. Розроблено інформаційно-аналітичну систему «клієнт-сервер», яка відображає різні графічні

звіти у вебінтерфейсі та розраховує кореляцію між кількістю поданих заяв

абітурієнтами та загальним значенням метрики для ЗВО.

Після аналізу отриманих КПЕ та звітів було встановлено, що

- загальне значення середнього балу абітурієнтів щорічно покращується;
- кількість заяв абітурієнтів щороку зростає, а у 2021 відбулося значне збільшення в порівнянні з попереднім періодом;

- кількість заяв абітурієнтів НУБІП та їхній середній бал є нижчими за загальні середні показники для 46 ЗВО;

- більше третини всіх заяв абітурієнтів (38,2%) у 2021 році припадає на 4 ЗВО: КНУ, НУ «Львівська політехніка», ХНУР та ЛНУ;

- розрахований загальний коефіцієнт кореляції для гіпотези для 21 ЗВО становить -26.29%.

Якщо підсумувати все вище сказане, вебметрика ЗВО не має суттєвого

впливу на кількість заяв абітурієнтів і тому можна зробити висновок, що існують

інші фактори, які впливають на вибір абітурієнтів в сторону того чи іншого ЗВО.

Можливо, аналіз цієї гіпотези в рамках лише однієї спеціальності 122 «Комп'ютерні науки» є недостатньо повним і тому для майбутніх досліджень

можна перевірити дану гіпотезу з загальною статистикою по всіх спеціальностях

(спеціалізаціях). Однак, як було раніше з'ясовано з проаналізованих джерел,

важливо підтримувати інтернет представництва ЗВО, збільшувати кількість публікацій, сприяти збільшенню їх цитованості, оскільки це має позитивний

вплив на позицію в рейтингу, за яким слідкують чимало ЗВО у нашій країні. Крім

цього, не потрібно забувати про студентів, яким необхідний якісно виконаний,

доступний, наповнений необхідною інформацією про освітні процеси, новини,

співробітників, матеріали вебресурс, і що не також вимагає законодавство України. Тобто потреба у якісному вебсайті ЗВО завжди була, є і буде надалі,

особливо в нинішню епоху пандемії коронавірусу й дистанційного навчання.

Тому відповідальним за вебпредставництва ЗВО можна порекомендувати й далі

підтримувати і покращувати свої ресурси, а співробітникам – збільшувати

кількість публікацій і свій вклад у розвиток української науки.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. ЄДЕБО Статистика за ЗВО. URL: <https://vstup.edbo.gov.ua/statistics/konkurs-universities/> (дата звернення:

04.11.2021).

2. Ranking web of universities Ukraine. URL: <https://www.webometrics.info/en/europe/ukraine> (дата звернення:

04.11.2021).

3. Коломієць В.А. Інформаційно-аналітична система показників вступу на ІТ-спеціальності та метрик відкритих електронних ресурсів ЗВО. *Інформаційні технології: економіка, техніка, освіта 2020*: збірник тез та доповідей XI Міжнародної науково-практичної конференції молодих учених (м. Київ, 10-11 листопада 2020 р.). Київ, 2020. С. 109-110.

4. Закон «Про вищу освіту». URL: <https://zakon.rada.gov.ua/laws/show/1556-18> (дата звернення: 04.11.2021).

5. Коломієць В.А. Інформаційно-аналітична система показників вступу на ІТ-спеціальності та метрик відкритих електронних ресурсів ЗВО. *Теоретичні та прикладні аспекти розробки комп'ютерних систем '2021*: збірник наукових праць за матеріалами IV Всеукраїнської науково-практичної інтернет конференції студентів і аспірантів, 29 квіт. 2021 р., НУБіП України, Київ. Київ, 2020. С. 30-31.

6. Коломієць В.А. Інформаційно-аналітична система показників вступу на ІТ-спеціальності та метрик відкритих електронних ресурсів ЗВО. *Інформаційні технології: економіка, техніка, освіта 2021*: збірник тез та доповідей XII Міжнародної науково-практичної конференції молодих учених (м. Київ, 11-12 листопада 2021 р.). Київ, 2020.

7. Вебметричний рейтинг університетів світу – Вікіпедія. URL: https://uk.wikipedia.org/wiki/Вебметричний_рейтинг_університетів_світу (дата звернення: 04.11.2021).

8. ХНУРЕ покращив позиції у рейтингу Ranking of world universities. URL:

<https://nure.ua/nure-pokrashtiv-pozicii-u-reitingu-ranking-of-world-universities> (дата звернення: 04.11.2021).

9. Університет Грінченка в Топ-50 найкращих українських ЗВО за даними міжнародного рейтингу Webometrics. URL:

<https://kubg.edu.ua/pro-universitet/news/podiji/7199-universytet-hrinchenka-v-top-50-naikashchykh-ukrainskykh-zvo-za-danymy-mizhnarodnoho-reytingu-webometrics.html> (дата звернення: 04.11.2021).

10. Липневий рейтинг Webometrics 2017 року – КПІ. URL:

<https://webometr.kpi.ua/2017-webometrics-2> (дата звернення: 04.11.2021).

11. Рейтинг Webometrics-2021 КНЕУ. URL:

https://ivo.kneu.edu.ua/ua/dosl_glot/prof_soitir_webom/webometrics2021/

(дата звернення: 04.11.2021).

12. Рейтинги сайтів КПІ ім. Ігоря Сікорського. URL: <https://webometr.kpi.ua/>

(дата звернення: 04.11.2021).

13. Гончарук І.В., Н.П. Юрчук Організація єдиного електронного науково-освітнього простору сучасного університету. Економіка. Фінанси.

Менеджмент: актуальні питання науки і практики. 2018. № 12. С. 75-87.

14. Шоробура І. Використання інформаційних технологій в управлінні вищим навчальним закладом. Педагогічний дискурс. 2016. № 21. С. 187-192.

15. Левченко Н. П. Відкриті електронні ресурси у діяльності бібліотек закладів вищої освіти: специфіка формування, управління, доступ: автореф. на здобуття наук. ступеня канд. наук із соціальних комунікацій за спец. 27.00.03. К., 2020. 20 с.

16. Sarwar R., Zia A., Nawaz R. Webometrics: evolution of social media presence of universities. Scientometrics. 2021. Vol. 126 P. 951-967.

<https://doi.org/10.1007/s11192-020-03804-y> (Last accessed: 04.11.2021).

17. QS World University Rankings URL:

<https://www.topuniversities.com/university-rankings/world-university-rankings/2021> (дата звернення: 04.11.2021)

18. Kvitka S., Rachynskii A., Borodin Y., Starushenko G., Lesina T., Kichuk A.

Statistical assessment of Webometric rating indicators of Ukrainian higher education institutions. *Management Theory and Studies for Rural Business and Infrastructure Development*. 2021. Vol. 43, №2. P. 177-194. URL:

<https://ejournals.vdu.lt/index.php/mtsrbid/article/view/2272> (Last accessed: 04.11.2021).

19. Yakymenko I., Kazymyr V., Lytyyn S. Webometrics ranking analysis and possible ways to improve the position of the university. 2020 IEEE 14th International Conference on Dependable Systems, Services and Technologies (DESSERT). 2020. P. 422-426. DOI: 10.1109/DESSERT50317.2020.9124999.

20. Umansky V., Shevchenko L., Bezugly A. Webometrics ranking of universities as a factor of gaining competitive advantage in the market for educational services. *Information Technologies and Learning Tools*. 2020. Vol. 77, №3. P. 324-336. URL: <https://doi.org/10.33407/itlt.v77i3.3261> (Last accessed: 04.11.2021).

21. Santos S., Brito C. Online engagement in higher education institutions: the triple relationship among Webometrics ranking, Facebook and website. *ICER 2017 Proceedings*. 2017. P. 6758-6765.

22. Peker S., Kueukozer-Caydar S., Cagiltay K. Exploring the Relationship between Web Presence and Web Usability for Universities: A Case Study from Turkey. *Program electronic library and information systems*. 2016. Vol. 50. P. 157-174. DOI: 10.1108/PROG-04-2014-0024.

23. Практичні заняття з актуалізації профілів авторів у Google Scholar – Рейтинги сайтів КНУ ім. Ігоря Сікорського. URL: <https://webometr.kpi.ua/node/463> (дата звернення: 04.11.2021).

24. Інформаційно-аналітична система : пат. 140372 Україна : МПК G05B23/00, G06F17/00. № и 2019 07240; заявл. 01.07.2019; опубл. 25.02.2020, Бюл. № 4)

25. Комп'ютерна програма «Моніторингова інформаційно-аналітична система»: пат. 35870 Україна. Заявл. 10.11.2010; опубл. 15.03.2011, Бюл. № 23.

26. Комп'ютерна програма «Інформаційно-аналітична система «Університет»» пат. 43459 Україна. Заявл. 22.02.2012; опубл. 02.07.2012, Бюл. № 27.

27. Опис комп'ютерної програми «Моніторингова інформаційно-аналітична система»; пат. 35871 Україна. Заявл. 10.11.2010; опубл. 15.03.2011, Бюл. № 23.

28. Інформаційно-аналітична система «Університет» ІІІ СумДУ. URL: <https://it.sumdu.edu.ua/golovna-ua/struktura/informatsiyno-analitichna-sistema-universitet> (дата звернення: 04.11.2021).

29. Unified Modeling Language - Wikipedia. URL: https://en.wikipedia.org/wiki/Unified_Modeling_Language (дата звернення: 04.11.2021).

30. Use case diagram - Wikipedia. URL: https://en.wikipedia.org/wiki/Use_case_diagram (дата звернення: 04.11.2021).

31. Siau Keng, Lee Lihynn. Are use case and class diagrams complementary in requirements analysis? An experimental study on use case and class diagrams in UML. *Requirements Engineering*. London, 2004. Vol. 9, №4. P. 229-237. DOI: 10.1007/s00766-004-0203-7

32. Deployment diagram - Wikipedia. URL: https://en.wikipedia.org/wiki/Deployment_diagram (дата звернення: 04.11.2021).

33. Sequence diagram - Wikipedia. URL: https://en.wikipedia.org/wiki/Sequence_diagram (дата звернення: 04.11.2021).

34. Activity diagram - Wikipedia. URL: https://en.wikipedia.org/wiki/Activity_diagram (дата звернення: 04.11.2021).

35. WayBack Machine. URL: https://web.archive.org/web/20180401000000*/https://www.webometrics.info/en/Europe/Ukraine (дата звернення: 04.11.2021).

36. OLAP - Wikipedia. URL:

https://en.wikipedia.org/wiki/Online_analytical_processing (дата звернення: 04.11.2021).

37. SQL Server Integration Services – Microsoft. URL:

<https://docs.microsoft.com/ru-ru/sql/integration-services/sql-server-integration-services?view=sql-server-ver15> (дата звернення: 04.11.2021).

38. Основные сведения об Analysis Services Microsoft. URL:

<https://docs.microsoft.com/ru-ru/analysis-services/analysis-services-overview?view=asallproducts-allversions> (дата звернення: 04.11.2021).

39. What is SQL Server Reporting Services (SSRS)? – Microsoft. URL:

<https://docs.microsoft.com/en-us/sql/reporting-services/create-deploy-and-manage-mobile-and-paginated-reports?view=sql-server-ver15> (дата звернення: 04.11.2021).

40. Python (programming language) – Wikipedia. URL:

[https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language)) (дата звернення: 04.11.2021).

41. Python package installer – PyPI. URL: <https://pypi.org/project/pip/> (дата

звернення: 04.11.2021).

42. The State of Developer Ecosystem 2020 – JetBrains. URL:

<https://www.jetbrains.com/lo/devecosystem/2020/> (дата звернення: 04.11.2021).

43. Stack Overflow Developer Survey 2020. URL:

<https://insights.stackoverflow.com/survey/2020/#most-popular-technologies> (дата звернення: 04.11.2021).

44. TIOBE Index for November 2021. URL: <https://www.tiobe.com/tiobe-index/>

(дата звернення: 04.11.2021).

45. PYPL Popularity of Programming Language. URL:

<https://pypi.github.io/PYPL.html> (дата звернення: 04.11.2021).

46. Microsoft SQL Server Wikipedia URL:

https://en.wikipedia.org/wiki/Microsoft_SQL_Server (дата звернення: 04.11.2021).

47. Download SQL Server Management Studio (SSMS) – Microsoft. URL: <https://docs.microsoft.com/en-us/sql/ssms/download-sql-server-management-studio-ssms?view=sql-server-ver15> (дата звернення: 04.11.2021).

48. Microsoft Excel Spreadsheet Software – Microsoft 365. URL: <https://www.microsoft.com/en-us/microsoft-365/excel> (дата звернення: 04.11.2021).

49. Microsoft Visual Studio URL: <https://docs.microsoft.com/en-us/visualstudio/get-started/visual-studio-ide?view=vs-2022> (дата звернення: 04.11.2021).

50. Microsoft Visual Studio Code. URL: <https://code.visualstudio.com/docs> (дата звернення: 04.11.2021).

51. JavaScript – MDN. URL: <https://developer.mozilla.org/en-US/docs/Web/javascript> (дата звернення: 04.11.2021).

52. What is an application programming interface (API)? IBM. URL: <https://www.ibm.com/cloud/learn/api> (дата звернення: 04.11.2021).

53. Introduction to Node.js. URL: <https://nodejs.dev/learn> (дата звернення: 04.11.2021).

54. Introduction – Vue.js URL: <https://vuejs.org/v2/guide/> (дата звернення: 04.11.2021).

55. JavaScript Pivot Table & Charts Component – Flexmonster. URL: <https://www.flexmonster.com/?r=fr5> (дата звернення: 04.11.2021).

56. Репозиторій «vikkolts/webometrics-diploma» – Github. URL: <https://github.com/vikkolts/webometrics-diploma> (дата звернення: 04.11.2021).

57. Коефіцієнт кореляції Пірсона – Вікіпедія. URL: https://uk.wikipedia.org/wiki/Коефіцієнт_кореляції_Пірсона (дата звернення: 04.11.2021).

НУБІП України

Додаток А
Програмний код об'єднання даних в один файл

```
from difflib import SequenceMatcher
```

```
import pandas as pd
```

```
import csv
```

```
dAbit = []
```

```
dWebometrics = []
```

```
listTotal = []
```

```
with open('abit_2018-2021.csv', 'r', encoding="utf8") as f1:
```

```
    csv_reader = list(csv.reader(f1, delimiter=';'))
```

```
    listTotal += csv_reader[0] #add columns to res file
```

```
    for row in csv_reader[1:]:
```

```
        dAbit.append([row[0], row[1:]]) #add elem to array as [name, [ZV0data, ...]]
```

```
with open('Webometrics_ukraine_2018-2021_with_totals.csv', 'r', encoding="utf8") as f2:
```

```
    csv_reader = list(csv.reader(f2, delimiter=';'))
```

```
    listTotal += csv_reader[0][1:] #add columns to res file
```

```
    for row in list(csv_reader)[1:]:
```

```
        dWebometrics.append([row[0], row[1:]])
```

```
resFile = []
```

```
for index1,nameOfList1 in enumerate(dAbit):
```

```
    for index2,nameOfList2 in enumerate(dWebometrics):
```

```
        if (nameOfList1[1][5] == '2021' and nameOfList1[1][5] == nameOfList2[1][7]):
```

```
            tempNum = SequenceMatcher(None, nameOfList1[0], nameOfList2[0]).ratio()
```

```
            if (tempNum >= 0.9): #check output for false merge
```

```
                print(str(nameOfList1[0]) + " - " + str(nameOfList2[0]) + ": " +
```

```
str(tempNum))
```

```
Н resFile.append([nameOfList1[0], *dAbit[index1 - 3][1],  
*dWebometrics[index2 - 3][1]]) # append data from 2018 for this ZVO  
resFile.append([nameOfList1[0], *dAbit[index1 - 2][1],  
*dWebometrics[index2 - 2][1]]) # append data from 2019 for this ZVO  
resFile.append([nameOfList1[0], *dAbit[index1 - 1][1],  
*dWebometrics[index2 - 1][1]]) # append data from 2020 for this ZVO  
Н resFile.append([nameOfList1[0], *dAbit[index1][1],  
*dWebometrics[index2][1]]) # append data from 2021 for this ZVO  
pd.DataFrame(data = resFile, columns = listTotal).to_csv('abit_webometrics_2018-  
2021.csv', encoding = 'utf-8-sig', index= False, sep = ';')
```

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

Додаток Б
Програмний код перетворення даних для подальшого імпортування у сховище даних

```

from numpy import int8
import pandas as pd
import csv

#form list of unique universities with its Id
dAbit = dict()

listUniTotal = ['ID', 'Name']

with open('abit_webometrics_2018-2021_cleared.csv', 'r', encoding="utf8") as f:
    csv_reader = list(csv.reader(f, delimiter=';'))
    index=1
    for row in csv_reader[1:]:
        dAbit[row[0]] =int8(index/4) #code uni by its unique ID
        index+=1

resFile = []
ind=1
for nameOfList1 in dAbit.keys():
    #print(nameOfList1)
    resFile.append([ind] + [nameOfList1])
    ind+=1

pd.DataFrame(data = resFile, columns = listUniTotal).to_csv('uniTableDWH.csv',
encoding = 'utf-8-sig', index= False, sep =';')

#code year, metrics to id. Split metrics by its ID to seperate rows
listUni =[]
listTotal = ['University',
'Усього подано заяв',
'Допущено до конкурсу',
'Усього рекомендовано',
'Мін.бал',
'Сер.бал',
'Ukraine Rank',

```

```

    'World Rank',
    'ID_Metric',
    'Metric_Value',
    'Year']

```

```

dmetricID = {"Presence": "1",
             "Impact": "2",
             "Openess": "3",
             "Excellence": "4",
             "All": "0"}

```

```

years = {
    "2018": "1",
    "2019": "2",
    "2020": "3",
    "2021": "4"
}

```

```

with open('abit_webometrics_2018-2021_cleared.csv', 'r', encoding="utf8") as f:
    csv_reader = list(csv.reader(f, delimiter=';'))
    for row in csv_reader[1:]:
        listUni.append(row[0:])

```

```

resFile = []

for i in range (0, len(listUni)):
    #form temp row part
    tempDic1, tempDic2, tempDic3, tempDic4, tempDic5 = listUni[i][1:8],
listUni[i][1:8], listUni[i][1:8], listUni[i][1:8], listUni[i][1:8]

```

```

    tempDic1.append(dmetricID["Presence"])
    tempDic1.append(listUni[i][8])
    tempDic1.append(years[listUni[i][13]])
    tempDic2.append(dmetricID["Impact"])
    tempDic2.append(listUni[i][9])
    tempDic2.append(years[listUni[i][13]])
    tempDic3.append(dmetricID["Openess"])
    tempDic3.append(listUni[i][10])
    tempDic3.append(years[listUni[i][13]])
    tempDic4.append(dmetricID["Excellence"])

```

```

H tempDic4.append(listUni[i][11])
tempDic4.append(years[listUni[i][13]])
tempDic5.append(dmetricID["All"])
tempDic5.append(listUni[i][12])
tempDic5.append(years[listUni[i][13]])

```

```

H #get uni ID by name
uniId = dAbit[listUni[i][0]]
resFile.append([uniId] + tempDic1)
resFile.append([uniId] + tempDic2)
resFile.append([uniId] + tempDic3)
resFile.append([uniId] + tempDic4)
resFile.append([uniId] + tempDic5)

```

```

H pd.DataFrame(data = resFile, columns =
listTotal).to_csv('devided_abit_webometrics_2018-2021_cleared.csv', encoding='utf-
8-sig', index= False, sep =';')

```

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

Додаток В
Програмний код з використанням SQL запитів для обчислення коефіцієнта кореляції Пірсона для ЗВО

```

const correlationZVOs = (number) => `WITH DataAvgStd
AS (SELECT ЗВО,
      STDEV([Всього заяв]) over(partition by ЗВО) AS XStdev,
      STDEV(Метрика) over(partition by ЗВО) AS YSTDev,
      COUNT(*) over(partition by ЗВО) AS SampleSize,
      AVG([Всього заяв]) OVER(PARTITION BY "ЗВО") AS "Сер. всього заяв",
      AVG("Метрика") OVER(PARTITION BY "ЗВО") AS "Сер. метрика",
      ( [Всього заяв] - AVG([Всього заяв]) over(partition by ЗВО)) * ( Метрика
- AVG(Метрика) over(partition by ЗВО)) AS ExpectedValue,
      "Всього заяв",
      "Метрика",
      "Рік"
FROM (
SELECT UniversityDim.Name AS "ЗВО"
      ,All_Applications_Entrants AS "Всього заяв"
      ,Admitted_To_Competition AS "Допущено"
      ,Recommended_Entrants AS "Рекомендовано"
      ,Min_Mark AS "Мін. бал"
      ,Avarage_Mark AS "Сер. бал"
      ,Rank_Ukraine AS "Рейтинг Україна"
      ,Rank_World AS "Рейтинг світ"
      ,Metric_Value AS "Метрика"
      ,DateDim.Year AS "Рік"
FROM UniversityFact
JOIN DateDim ON DateDim.ID_Date = UniversityFact.ID_Date
JOIN UniversityDim ON UniversityFact.ID_University = UniversityDim.ID_University
WHERE ID_Metric=0
)a)
SELECT ЗВО,
"Всього заяв",
"Метрика",
"Рік",
SUM(ExpectedValue) over(partition by ЗВО) / (SampleSize - 1 ) / ( XStdev * YSTDev
) AS "Кореляція"

```



```
Н FROM DataAvgStd  
WHERE "Сер. всього заяв" > ${number}  
ORDER BY "Сер. всього заяв" desc`;
```

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України

НУБІП України