

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ДОНЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ВАСИЛЯ СТУСА
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ І ПРИКЛАДНИХ ТЕХНОЛОГІЙ
КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

МИШКІВСЬКА ЯНА ВІКТОРІВНА

Допускається до захисту:
в.о. завідувача кафедри
інформаційних технологій,
д-р. техн. наук, професор
_____ Наталія ВЕСЕЛОВСЬКА
« ____ » _____ 2026 р.

**РОЗРОБКА ВІ-СИСТЕМИ ДЛЯ АНАЛІЗУ МІГРАЦІЙНИХ ПОТОКІВ
УКРАЇНЦІВ В УМОВАХ ВІЙНИ**

Спеціальність 122 Комп'ютерні науки

Кваліфікаційна (бакалаврська) робота

Керівник:

Тетяна БАЦЕНКО, доцент кафедри
інформаційних технологій,
к. т. н., доцент

Оцінка: _____ / _____ / _____
(бали/за шкалою ЄКТС/за національною шкалою)

Голова ЕК: _____

АНОТАЦІЯ

Мишківська Я. В. Розробка ВІ-системи для аналізу міграційних потоків українців в умовах війни. Спеціальність 122 «Комп'ютерні науки», освітня програма «Комп'ютерні науки». Донецький національний університет імені Василя Стуса, Вінниця, 2026.

У роботі розроблено ВІ-систему на базі Power BI для аналізу міграційних потоків українців у країнах ЄС. Реалізовано аналітичний дашборд для візуалізації географічного розподілу, часової динаміки та ключових міграційних показників. Для аналізу використано відкриті статистичні дані Eurostat, виконано кореляційний аналіз показників та побудовано інтерактивні візуалізації міграційних процесів.

Ключові слова: ВІ-система, Power BI, міграційні процеси, українські біженці, Eurostat, аналіз даних.

60 с., 6 табл., 17 рис., 2 дод., 42 джерел.

ABSTRACT

Myshkivska Y. V. Development of a BI System for Analysis of Migration Flows of Ukrainians under Wartime Conditions. Specialty 122 "Computer Science", educational program "Computer Science". Vasyl Stus Donetsk National University, Vinnytsia, 2026.

The qualification work presents the development of a Power BI-based BI system for analyzing migration flows of Ukrainians in EU countries. An analytical dashboard was developed to visualize geographical distribution, temporal dynamics, and key migration indicators. The analysis is based on open Eurostat datasets and includes correlation analysis and interactive visualization of migration processes.

Keywords: BI system, Power BI, migration processes, Ukrainian refugees, Eurostat, data analysis.

ЗМІСТ

ВСТУП.....	4
РОЗДІЛ 1. АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ВІ-ТЕХНОЛОГІЙ	6
1.1 Сутність та особливості міграційних процесів	6
1.2 Вплив війни на міграційні потоки українців	10
1.3 Використання ВІ-систем для аналізу даних.....	14
РОЗДІЛ 2. АНАЛІЗ ТА ПІДГОТОВКА ДАНИХ ДЛЯ ДОСЛІДЖЕННЯ МІГРАЦІЙНИХ ПОТОКІВ.....	21
2.1 Джерела даних про міграцію українців	21
2.2 Методи обробки та аналізу даних	25
2.3 Формування набору даних для ВІ-системи	31
РОЗДІЛ 3. Розробка ВІ-системи для аналізу міграційних потоків.....	36
3.1 Архітектура ВІ-системи.....	36
3.2 Реалізація аналітичного дашборду	38
3.3 Аналіз отриманих результатів	48
ВИСНОВКИ.....	55
СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ	57
ДОДАТОК А	61
ДОДАТОК А	63

ВСТУП

Міграційні процеси є одним із ключових факторів, що впливають на демографічну структуру населення, ринок праці, соціальну політику та економічний розвиток держав. Масове переміщення населення змінює навантаження на соціальні системи, формує нові економічні виклики та впливає на розподіл людських ресурсів між країнами.

Після початку повномасштабної війни в Україні у 2022 році міграційні процеси набули чималого масштабу. Мільйони громадян були змушені залишити територію країни та отримати тимчасовий захист або інші форми підтримки за кордоном. Це призвело до формування масштабних міграційних потоків, концентрації українського населення в окремих країнах Європейського Союзу та зміни динаміки соціально-економічних процесів у приймаючих державах.

Зростання обсягів даних формує потребу у сучасних методах їх аналізу та обробки [37]. Обробка великих наборів даних, інтеграція інформації з різних джерел та виявлення прихованих закономірностей потребують застосування сучасних інструментів Business Intelligence (BI). BI-системи забезпечують автоматизацію аналізу, інтерактивну візуалізацію, прогнозування тенденцій і підтримку прийняття рішень на основі даних[1].

Особливого значення набуває використання BI-технологій у поєднанні з методами аналітики даних та елементами машинного навчання [38]. Такі підходи дозволяють не лише досліджувати поточний стан міграційних процесів, але й оцінювати їхню динаміку, визначати взаємозв'язки між показниками.

Незважаючи на значну кількість досліджень у сфері міграції, питання комплексного аналізу міграційних процесів українців із використанням BI-систем, інтерактивних дашбордів та сучасних методів обробки даних залишається недостатньо дослідженим. Актуальним є створення аналітичних інструментів, що дозволяють поєднувати статистичні дані, візуалізацію, часовий аналіз, кореляційні залежності.

Метою дослідження є розробка ВІ-системи для аналізу міграційних потоків українців в умовах війни із використанням сучасних інструментів обробки даних, інтерактивної візуалізації та аналітичних методів.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- дослідити особливості міграції громадян України в умовах повномасштабної війни;
- розглянути сучасні ВІ-технології та інструменти аналізу даних;
- дослідити методи збору, очищення та інтеграції статистичних даних;
- сформувати набір даних для аналізу міграційних процесів;
- реалізувати аналітичний дашборд із використанням ВІ-систем;
- здійснити аналіз часових рядів, кореляційних залежностей та просторового розподілу міграційних потоків [32].

Об'єктом дослідження є міграційні процеси українців в умовах війни.

Предметом дослідження є методи аналізу, обробки та візуалізації міграційних даних із використанням ВІ-технологій та аналітичних інструментів.

Практичне значення роботи полягає у створенні ВІ-системи, що дозволяє здійснювати інтерактивний аналіз міграційних процесів, визначати ключові тенденції, оцінювати навантаження на країни ЄС, виявляти закономірності у даних та підтримувати прийняття управлінських рішень на основі статистичних показників.

Апробацію результатів дослідження здійснено шляхом публікації матеріалів конференції та статті у фаховому науковому виданні (категорії Б).

1. Мишківська Я. В., Баценко Т. В. Аналіз міграційних даних із використанням ВІ-систем. *Наука і техніка сьогодні*. 2026 № 4(58). С 2757–2765.
2. Myshkivska, Y., Lozynska L, T. Application of Business Intelligence Systems for Migration Data Analysis. In: *Topical Issues of Humanities, Technical, and Natural Sciences: Proceedings of the 10th International Scientific Conference for Bachelor, Master, Graduate Students, and Young Researchers*. Vinnytsia, 2026.

РОЗДІЛ 1

АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ ТА ВІ-ТЕХНОЛОГІЙ

1.1 Сутність та особливості міграційних процесів

Згідно з визначенням Міжнародної організації з міграції (МОМ), «міграція – це переміщення осіб, як поодинокі, так і в групах, з однієї країни чи місця до іншого, незалежно від того, чи вони перетнули міжнародні кордони або перемістилися всередині своєї країни, незалежно від причини, тривалості перебування там чи їхнього правового статусу»[28]. Це широке визначення охоплює добровільну та вимушену міграцію і може бути тимчасовою або постійною.

На міграційні процеси впливає низка різних факторів. Економічні, політичні, соціальні та демографічні фактори відіграють важливу роль у створенні міграційного процесу. Основною причиною міграції є економічна нерівність між країнами та регіонами. Чимало людей переїжджають в інше місце, оскільки шукають роботу або покращення умов життя. Як зазначає Світовий банк, міжнародна міграція є важливим гравцем у світовій економіці, оскільки грошові перекази мігрантів забезпечують значну частину доходів для багатьох країн, що розвиваються. Окрім грошових переказів, мігранти також допомагають підтримувати різні галузі промисловості, забезпечуючи робочу силу; таким чином, вони допомагають як країнам походження, так і країнам перебування зростати економічно.

Існує багато способів класифікації різних типів міграційних процесів. Ці класифікації ґрунтуються на географічному розташуванні, мотивації, тривалості перебування на новому місці та способі прийняття рішення про переїзд. Різні типи міграційних процесів наведено в таблиці 1.1:

Таблиця 1.1 – Класифікація міграційних процесів

Критерій	Вид міграції	Характеристика
За територією переміщення	Внутрішня	Переміщення населення в межах однієї держави
	Міжнародна	Переміщення населення через державні кордони
За причинами	Економічна	Пов'язана з пошуком роботи та кращих умов життя
	Освітня	Пов'язана з отриманням освіти
	Політична	Виникає через політичну нестабільність або переслідування
	Вимушена	Пов'язана з війнами, конфліктами або катастрофами
За тривалістю	Короткострокова	Переміщення на період до одного року
	Довгострокова	Перебування понад один рік
	Постійна	Остаточна зміна місця проживання
За характером рішення	Добровільна	Міграція за власним бажанням
	Вимушена	Міграція через зовнішні обставини

З табл 1.1 відображає, що міграційні процеси мають різноманітні форми та можуть класифікуватися за територією переміщення, причинами, тривалістю перебування й характером прийняття рішення, що дозволяє комплексно аналізувати особливості сучасної міграції.

За останні кілька десятиліть міграційні моделі різко змінилися, що зараз призводить до надзвичайно високого рівня міжнародної мобільності, яка, згідно з останніми оцінками Організації Об'єднаних Націй, становить понад 280 мільйонів мігрантів у світі. Ця тенденція підкреслює зростаючу важливість міграції у визначенні того, як глобальні економічні сили формуватимуть майбутні тенденції.

Військові конфлікти та політичні кризи зробили міграцію дедалі важливішою проблемою. Як наслідок, кількість внутрішньо переміщених осіб та біженців також зросла. За даними Управління Верховного комісара ООН у справах біженців, понад 110 мільйонів людей у всьому світі покинули свої домівки через насильство та переслідування.

Міграція є складним, різноманітним процесом, який може відбуватися через численні стимули, що впливають на міграційні моделі населення. Розуміння міграційних моделей може призвести до кращої оцінки того, як міграційні групи впливають на економічний розвиток та соціальний розвиток держав. Впровадження методів обробки великих даних та аналітичних платформ, а також використання систем бізнес-аналітики (BI-систем) дозволяє нам проводити як описовий аналіз міграційних потоків, так і знаходити приховані закономірності, виявляти тенденції та приймати рішення на основі даних, що підтримують управління як на державному, так і на міжнародному рівнях. Це особливо важливо при дослідженні міграційних процесів, оскільки аналіз значних обсягів статистичних даних дає можливість виявляти закономірності зміни показників та оцінювати їх динаміку [25]. Крім того, поєднання різних типів даних (таких як демографічні дані, статистика та демографічна інформація) допоможе створити загальне уявлення про міграцію, на яку значною мірою вплинула глобальна нестабільність та швидкі зміни в ній.

Ключовим елементом аналізу міграційних тенденцій є розуміння їх кількісних показників та пошук способів створення аналітичних моделей. У сучасній галузі дослідження розглядають міграцію не лише як соціальну подію, а й як складну, динамічну систему, яку можна математично представити за допомогою різноманітних статистичних та числових методів вимірювання. Деякі приклади включають: використання методології часових рядів та регресійного аналізу; а також методів кластерного та прогнозного аналізу — ми можемо знайти тенденції в різні роки, пов'язані з переміщенням людей з географічних районів [3].

Крім того, міграційні процеси є дуже невизначеними через багато зовнішніх факторів, включаючи політику, економіку, збройні конфлікти та зміни в законодавстві. Ці фактори ускладнюють вивчення міграції, тому необхідний комплексний підхід, який поєднує якісний та кількісний аналізи. Тому сучасні інформаційно-технологічні інструменти, що дозволяють працювати з великими

обсягами даних, включаючи всі різні змінні, що впливають на певний регіон, будуть дуже важливими в контексті досліджень міграції.

Міграція суттєво впливає на демографічну структуру населення країни. Коли відбувається відтік людей працездатного віку з країни, це призводить до старіння населення, зниження трудового потенціалу та нижчого економічного зростання цієї країни. Водночас країни, що приймають мігрантів, отримують додаткових працівників з інших країн, що може принести користь економіці країни, але може бути навантаженням на їхні соціальні системи. Вплив міграції проявляється не лише у зміні чисельності населення, але й у трансформації структури ринку праці, соціального забезпечення та економічної активності держав. Значний відтік працездатного населення може спричинити дефіцит трудових ресурсів, скорочення податкових надходжень та посилення демографічного навантаження на окремі групи населення.

Такі країни, отримують додаткову робочу силу та можливість частково компенсувати демографічне старіння населення. Проте це також може створювати додаткове навантаження на соціальну інфраструктуру, систему освіти та медичного забезпечення.

Таким чином, міграційні процеси мають комплексний характер та впливають не лише на окремих осіб, а й на соціально-економічний розвиток держав загалом. Міграція є важливим фактором у глобальному перерозподілі людських ресурсів

Оперативний моніторинг міграційних процесів є важливим сьогодні. Цей моніторинг здійснюється за допомогою цифрових платформ, статистичних ресурсів та аналітичних систем, які надають доступ до актуальних даних у режимі реального часу. Використання таких методів для оцінки міграційних потоків робить оцінку набагато точнішою, зміни можна легко виявити та створювати обґрунтовані прогнози.

Дослідження природи та характеристик міграційних процесів є важливим для проведення додаткового аналізу цих процесів з використанням сучасних технологій бізнес-аналітики (BI). Аналітичні інструменти, методи обробки даних

та інформаційні системи можуть бути використані для сприяння переходу від методології, що ґрунтується на описі, до підходу, що ґрунтується на аналізі та прогнозуванні; тим самим забезпечуючи розуміння сучасних міграційних тенденцій. Інформаційні системи забезпечують інтеграцію даних та підтримку аналітичних процесів [14].

1.2 Вплив війни на міграційні потоки українців

Повномасштабна війна в Україні з 2022 року різко змінила демографічні процеси для українців та моделі вимушеної міграції. Війна призвела до мільйонів переміщених осіб, як внутрішньо переміщених, так і через транскордонну міжнародну міграцію. Як правило, люди не мали іншого вибору, окрім як покинути свої домівки через небезпечні для життя ситуації: руйнування житла та інфраструктури, втрату роботи та загальне погіршення соціально-економічних умов; тому цей конфлікт спричинив один з найбільших міграційних потоків через Європу за останні кілька десятиліть.

Більшість міграційних процесів, спричинених війною, є вимушеними. Більшість людей мігрують не з економічних причин, а для того, щоб гарантувати свою безпеку та безпеку своїх сімей. Деякі люди в Україні мігрували всередині країни та зараз є внутрішньо переміщеними особами, а більшість тих, хто покинув країну, зараз отримують захист або статус біженця.

Перша хвиля міграції була зафіксована в перші місяці війни, і найбільша кількість українців втекла до сусідніх країн, таких як Польща, Румунія, Словаччина та Угорщина. Ці країни були географічно близькі одна до одної, що дозволяло багатьом людям швидко перетинати кордон, щоб знайти тимчасове житло. Деякі біженці пізніше переїхали до інших європейських країн, таких як Німеччина, Чехія, Італія та Іспанія.

Громадянам України було надано можливість в'їхати до Європейського Союзу через тимчасову форму захисту, яка дозволяла їм легально проживати та працювати в країнах Європейського Союзу, а також мати доступ до соціальної

допомоги. Це створило набагато простіший спосіб, який дозволив би членам Європейського Союзу підтримувати інтеграцію українських біженців та сприяти задоволенню їхніх основних потреб після того, як вони були змушені покинути Україну.

Статистичні дані, надані міжнародними організаціями, підтверджують масштаби вимушеної міграції, враховуючи кількість українців, які втекли з кінця 2023 року; за даними Управління Верховного комісара ООН у справах біженців, на той час у різних країнах світу перебувало приблизно 6 мільйонів українських біженців [39]. Деякі з них знаходилися в сусідніх країнах, а деякі були розкидані по різних частинах Європи та інших регіонах світу. Дані про загальну кількість українських біженців можна знайти в таблиці 1.2. наведено нижче[39]:

Таблиця 1.2 – Біженці з України у 2023 році

Показник	Кількість
Загальна кількість біженців з України	6 млн
Українці у сусідніх країнах	2.6 млн (44%)
Українці в інших країнах Європи та світу	3.4 млн
Зміна порівняно з 2022 роком	+5%

Наведені дані свідчать про значні масштаби вимушеної міграції українців. Майже половина біженців перебуває у сусідніх країнах, що пояснюється географічною близькістю та можливістю швидкого перетину кордону. Інша частина українців знайшла притулок у країнах Європи та інших регіонах світу. Особливістю сучасної вимушеної міграції українців є її неоднорідний характер. Частина населення переміщується до сусідніх держав із наміром тимчасового перебування, тоді як інші мігранти інтегруються в економічне та соціальне середовище країн перебування на довгостроковій основі.

На рішення щодо вибору країни перебування впливають різні фактори: географічна близькість, мовне середовище, наявність соціальної підтримки, можливість працевлаштування та доступ до освіти. Саме тому міграційні потоки українців мають нерівномірний характер розподілу між країнами.

Для більш детального аналізу географічного розподілу українських біженців можна використати статистичні дані Європейського статистичного агентства [24]. Візуалізація цих даних у вигляді карти дозволяє оцінити, у яких країнах Європи перебуває найбільша кількість українських громадян, які отримали тимчасовий захист. Приклад візуалізації через карти представлено на рис 1.1:

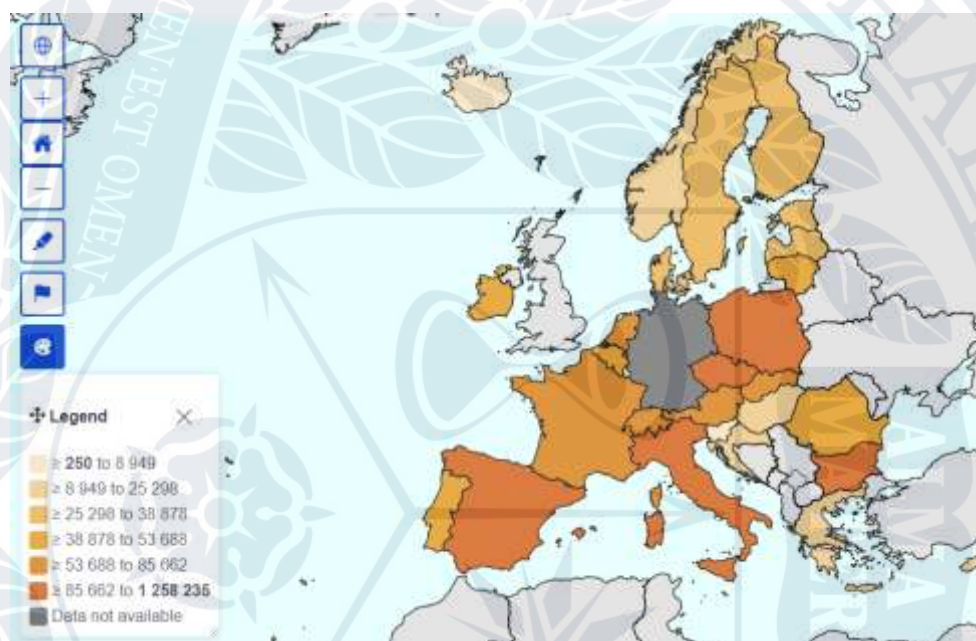


Рисунок 1.1 – Розподіл українських біженців у країнах Європи[22]

Наведені дані показують, що більшість українських біженців знаходяться в країнах Центральної та Західної Європи, зокрема в Польщі, Німеччині та Чеській Республіці. Причинами цього є близькість цих країн до України, а також їхні унікальні системи підтримки, доступні спеціально для громадян України.

Окрім зовнішньої міграції, спостерігається швидке зростання внутрішньої міграції, спричиненої переміщенням населення. Більшість цих людей були змушені покинути свої домівки та переїхати до іншого регіону України, де вважається безпечніше. За оцінками міжнародних організацій (МОП, УВКБ ООН тощо), станом на кінець 2023 року в Україні налічувалося приблизно 3,7 мільйона ВПО.

Економіка та соціальна структура країни будуть стикатися зі значним обсягом вимушеної міграції. Одним із наслідків відтоку населення є зменшення загальної кількості працездатного населення в країні. Це безпосередньо впливає на здатність країни генерувати економічну активність. Країни, які приймають біженців, також несуть відповідальність за забезпечення житлом, доступ до освіти, медичного обслуговування та пошук роботи.

З цих причин війна була одним з основних факторів, що вплинули на міграційні процеси в Україні. Масові міграції призвели до багатьох нинішніх внутрішніх та зовнішніх міграційних потоків за розмірами. Ці потоки є довгостроковими соціальними та економічними наслідками. Аналіз цих міграційних процесів є важливим для розуміння демографічної ситуації сьогодні та формування обґрунтованої державної політики щодо відбудови країни.

Міграція, спричинена війною, призведе не лише до короткострокових, а й до довгострокових наслідків для країни. Одним із довгострокових потенційних наслідків міграції, спричиненої війною, є можливість часткової втрати людського капіталу, оскільки значна частина мігрантів інтегрувалася в економічну та соціальну інфраструктуру країн, що їх приймають. Неможливість повернення мігрантів після завершення бойових дій може ускладнити повернення переміщених осіб, а також затримати відновлення національної економіки України.

Водночас, сучасні міграційні моделі українців також характеризуються швидким рухом та нерівномірним розподілом, тому існує потреба застосовувати сучасні аналітичні інструменти для їх вивчення. Бізнес-аналітика (BI), аналітичні та прогнозні моделі можуть забезпечити точніше вимірювання руху населення, визначити основні закономірності та створити ефективні політичні рішення на національному рівні.

Отже, необхідна ретельна оцінка впливу війни на міграційні системи, щоб розробити методи зміцнення демографічних умов, створити шлях для репатріації громадян та забезпечити подальше зростання України після закінчення війни.

1.3. Використання ВІ-систем для аналізу даних

ВІ-системи являють собою сукупність методів, інструментів і технологій, що дозволяють здійснювати збір, інтеграцію, обробку, аналіз і візуалізацію даних з метою отримання корисної інформації для прийняття обґрунтованих рішень [8]. Основною метою використання ВІ є перетворення великих обсягів сирих даних у структуровану інформацію, придатну для аналітики, прогнозування та підтримки управлінських рішень [12]. Ефективне використання аналітичних систем сприяє підвищенню якості підтримки прийняття рішень та оптимізації процесів аналізу даних [2].

У дослідженні міграційних процесів ВІ-системи відіграють особливо важливу роль, оскільки дозволяють працювати з великими масивами статистичних даних, що надходять із різних джерел: міжнародних організацій, державної статистики, відкритих баз даних та інших інформаційних ресурсів. Використання ВІ-підходу дає змогу не лише аналізувати поточний стан міграційних потоків, але й виявляти закономірності, визначати фактори впливу та прогнозувати тенденції. Методи інтелектуального аналізу даних дозволяють знаходити приховані залежності та підвищувати точність прогнозування [9].

Функціонування ВІ-систем базується на багаторівневій архітектурі наведено на рис 1.2. Першим рівнем є джерела даних, які можуть включати як структуровані (таблиці), так і неструктуровані дані. Наступним рівнем є ETL-процеси, що забезпечують очищення, трансформацію та інтеграцію даних. Після цього інформація потрапляє у сховище даних (Data Warehouse), де вона структурується для подальшого аналізу [41]. На аналітичному рівні використовуються інструменти Python, R і SQL.

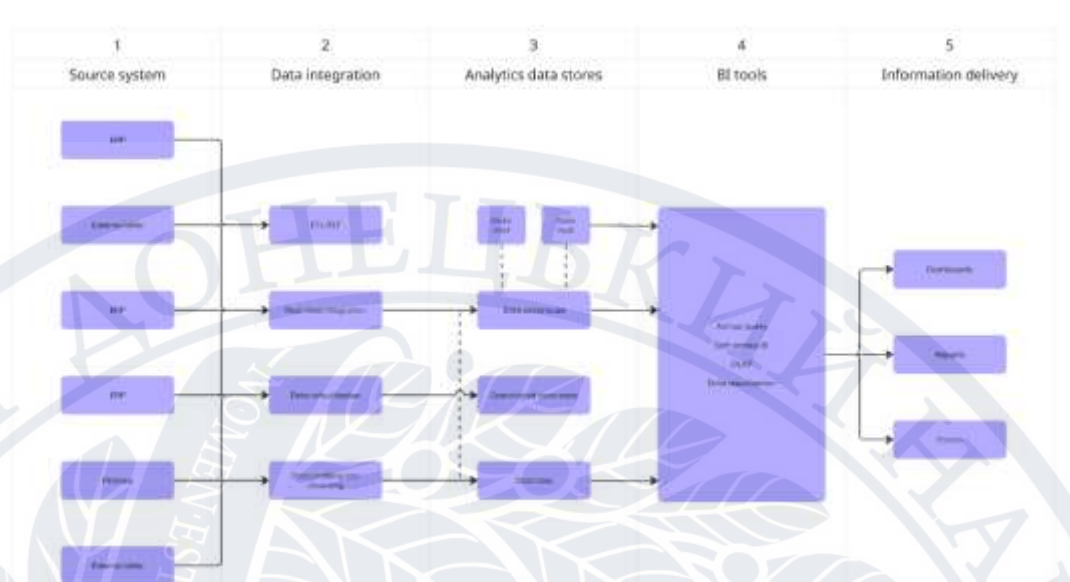


Рисунок 1.2 – Архітектура ВІ-системи[35]

Python також дозволяє застосовувати методи машинного навчання, що значно розширює можливості ВІ-систем. Він є одним із ключових інструментів сучасної аналітики даних. Завдяки бібліотекам *pandas* і *numpy* він дозволяє ефективно обробляти великі обсяги інформації, виконувати фільтрацію, групування та агрегацію даних. Використання бібліотек *matplotlib* і *seaborn* забезпечує побудову графіків і візуалізацій. Окрім базових операцій обробки даних, Python активно використовується для автоматизації процесів підготовки інформації, інтеграції даних із різних джерел та реалізації складних аналітичних алгоритмів. Його застосування дозволяє значно скоротити час виконання повторюваних операцій та підвищити точність аналізу за рахунок зменшення впливу людського фактора.

У сфері бізнес-аналітики Python часто використовується для реалізації ETL-процесів, які забезпечують отримання даних із різних джерел, їх очищення, трансформацію та підготовку до подальшого аналізу. Використання автоматизованих сценаріїв дозволяє підтримувати актуальність даних та спрощує оновлення аналітичних моделей.

Важливою перевагою Python є підтримка бібліотек машинного навчання, таких як *scikit-learn*, *TensorFlow* та інших інструментів, що дозволяють реалізовувати класифікацію, кластеризацію, прогнозування та інші підходи

інтелектуального аналізу даних. У контексті дослідження міграційних процесів це створює можливість виявлення прихованих закономірностей, оцінки тенденцій та формування прогнозів зміни показників у майбутньому.

Крім того, Python забезпечує інтеграцію з сучасними ВІ-платформами, зокрема Power BI, що дозволяє поєднувати аналітичні можливості програмування із засобами інтерактивної візуалізації. Такий підхід сприяє створенню комплексних аналітичних систем, здатних працювати з великими обсягами даних та підтримувати процес прийняття рішень.

Мова R використовується для глибокого статистичного аналізу, включаючи регресійні моделі, аналіз часових рядів та перевірку гіпотез. Вона широко застосовується у наукових дослідженнях.

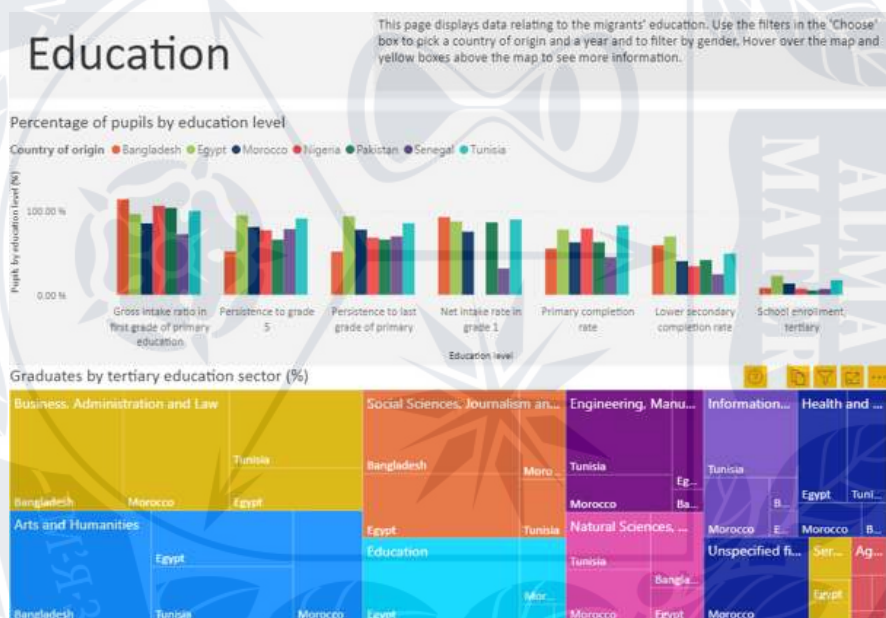


Рисунок 1.3 – Приклад дашборду в Power BI [35]

Power BI дозволяє створювати інтерактивні аналітичні панелі, які забезпечують зручне представлення даних див. рис. 1.3. Однією з ключових переваг Power BI є підтримка інтерактивної взаємодії між візуалізаціями, що дозволяє користувачам здійснювати багатовимірний аналіз даних. Використання фільтрів, сегментів та взаємопов'язаних графіків забезпечує можливість швидко змінювати параметри аналізу та адаптувати результати відповідно до поставлених завдань [10, 35].

Power BI підтримує створення KPI-показників, геопросторових карт, аналіз часових рядів, порівняльний аналіз та інтеграцію зовнішніх сценаріїв Python. Це робить платформу універсальним інструментом для реалізації BI-рішень у різних сферах, включаючи економіку, медицину, маркетинг та демографічні дослідження.

Особливо актуальним використання Power BI є у випадках роботи зі статистичними даними, де необхідно забезпечити одночасно зручність візуалізації та можливість глибокого аналітичного дослідження. Використання інтерактивних дашбордів сприяє більш ефективному сприйняттю інформації та пришвидшує виявлення тенденцій.

Використання візуальної аналітики сприяє швидшому виявленню закономірностей та інтерпретації результатів [17].

Порівняльну характеристику основних інструментів BI-аналітики наведено в таблиці 1.3.

Таблиця 1.3 – Інструменти BI-аналітики

Інструмент	Призначення	Недоліки
Python	Обробка та інтеграція даних, аналітика, ML	Вимагає навичок програмування
R	Статистичний аналіз і моделювання	Обмежена універсальність
SQL	Управління та вибірка даних із БД	Обмежені аналітичні можливості
Power BI	Візуалізація та бізнес-аналітика	Залежність від якості підготовки даних

Як видно з таблиці 1.3, різні інструменти BI-аналітики мають власні переваги та обмеження, тому їх вибір залежить від поставлених завдань і специфіки аналізу даних. Після аналізу інструментів BI-аналітики доцільно також розглянути підходи до зберігання та організації даних, зокрема концепції Data Warehouse та Data Lake [15].

Data Warehouse (сховище даних) призначене для зберігання структурованих, попередньо оброблених даних, які використовуються для

аналітики та формування звітності [27]. Дані у сховищі проходять етапи очищення, трансформації та інтеграції, що забезпечує їх узгодженість, високу якість та швидкий доступ при виконанні аналітичних запитів.

Натомість Data Lake використовується для зберігання сирих даних у різних форматах, включаючи структуровані, напівструктуровані та неструктуровані дані. Такий підхід дозволяє накопичувати великі обсяги інформації без попередньої обробки, що є особливо актуальним у випадках, коли дані надходять з різнорідних джерел і можуть бути використані для подальшого аналізу або машинного навчання.

У контексті даної роботи Data Warehouse доцільно використовувати для формування аналітичного набору міграційних даних, який застосовується у Power BI для побудови звітів та дашбордів. Водночас Data Lake може розглядатися як середовище для зберігання первинних даних з різних джерел (Eurostat, UNHCR тощо), які надалі проходять обробку за допомогою Python або R.

Таким чином, поєднання підходів Data Warehouse та Data Lake дозволяє створити гнучку та масштабовану архітектуру BI-системи, що забезпечує ефективну роботу з міграційними даними різної структури та рівня обробки.

Поряд із інструментами та архітектурою BI-систем важливим є розуміння типів аналітики, які використовуються для обробки даних та підтримки прийняття рішень. У сучасних BI-рішеннях виділяють декілька основних типів аналітики, що відрізняються рівнем складності та глибиною аналізу. Основні типи аналітики наведено в таблиці 1.4.

Таблиця 1.4 – Типи аналітики

Тип аналітики	Характеристика
Descriptive	Аналіз історичних даних з метою визначення того, що відбулося
Diagnostic	Визначення причин і факторів, що вплинули на події
Predictive	Прогнозування майбутніх тенденцій на основі наявних даних
Prescriptive	Формування рекомендацій щодо прийняття управлінських рішень

Як видно з таблиці 1.4, різні типи аналітики застосовуються залежно від поставлених завдань дослідження. У межах даної роботи основна увага приділяється описовій аналітиці, яка дає змогу оцінити особливості та динаміку міграційних процесів українців на основі наявних статистичних даних.

Таким чином, у результаті проведеного аналізу предметної області встановлено, що міграційні процеси є складним багатофакторним явищем, яке формується під впливом економічних, політичних, соціальних та демографічних чинників. Особливої актуальності вони набули в умовах повномасштабної війни в Україні, що спричинила масштабні внутрішні та зовнішні переміщення населення та суттєво вплинула на демографічну структуру країни.

Проведений аналіз показав, що сучасні міграційні потоки характеризуються високою динамічністю, географічною нерівномірністю та значною залежністю від зовнішніх факторів, що ускладнює їх дослідження традиційними методами. Це обумовлює необхідність застосування сучасних інформаційно-аналітичних підходів, зокрема ВІ-технологій.

У ході дослідження було розглянуто основні інструменти ВІ-аналітики, зокрема Python, R, SQL та Power BI, які у сукупності забезпечують повний цикл роботи з даними: від їх збору та обробки до побудови аналітичних моделей і візуалізації результатів. Встановлено, що найбільш ефективним є комплексний підхід, який поєднує можливості різних технологій у єдиній аналітичній системі.

Також було проаналізовано сучасні підходи до організації зберігання даних, зокрема концепції Data Warehouse та Data Lake, які забезпечують ефективне управління структурованими та неструктурованими даними. Це є важливим елементом побудови ВІ-систем, оскільки якість аналітики безпосередньо залежить від організації даних.

Окрему увагу приділено типам аналітики, які використовуються в ВІ-системах. Визначено, що описова, діагностична, прогностична та прескриптивна аналітика дозволяють не лише оцінювати поточний стан міграційних процесів,

але й виявляти причинно-наслідкові зв'язки, прогнозувати подальший розвиток подій та формувати рекомендації для прийняття рішень.

Таким чином, використання ВІ-систем у поєднанні з мовами програмування Python та R створює основу для побудови ефективної аналітичної платформи дослідження міграційних процесів. Це дозволяє перейти від описового аналізу до глибокого аналітичного та прогнозного підходу, що є необхідним для розуміння сучасних тенденцій та розробки обґрунтованих управлінських рішень.

Отримані результати аналізу предметної області та ВІ-технологій є базою для подальшої розробки та реалізації аналітичної системи, що буде представлена у наступних розділах роботи.

РОЗДІЛ 2

АНАЛІЗ ТА ПІДГОТОВКА ДАНИХ ДЛЯ ДОСЛІДЖЕННЯ МІГРАЦІЙНИХ ПОТОКІВ

2.1. Джерела даних про міграцію українців

Для забезпечення об'єктивності та достовірності аналізу необхідно використовувати комплексний підхід до вибору джерел даних, що включає міжнародні статистичні бази, гуманітарні платформи, державні ресурси, а також відкриті аналітичні датасети, представлено на табл. 2.1.

Таблиця 2.1 – Основні джерела даних

Джерело даних	Тип даних	Аналітичні можливості	Роль у ВІ-системі
Eurostat	Офіційна статистика ЄС	Аналіз кількості осіб із тимчасовим захистом, географічний розподіл, динаміка у часі	Базове джерело для аналізу міграції в країнах ЄС
UNHCR	Дані про біженців	Оцінка масштабів вимушеної міграції, порівняння між країнами, гуманітарні показники	Доповнює статистику оцінками міжнародних організацій
Держстат України	Національна статистика	Аналіз демографічних змін, внутрішньої міграції, регіональних показників	Забезпечує аналіз внутрішніх процесів
ДПСУ	Дані про перетин кордону	Оцінка інтенсивності виїзду/повернення, сезонні коливання	Дозволяє аналізувати динаміку потоків
HDX	Гуманітарні дані	Географічний аналіз, просторовий розподіл, кризові індикатори	Використовується для картографічної візуалізації
ACAPS	Аналітичні кризові дані	Аналіз факторів впливу (безпека, економіка, гуманітарна ситуація)	Дає змогу інтерпретувати причини міграції
Kaggle / Open Data	Підготовлені датасети	Тестування моделей, агрегований аналіз, навчальні сценарії	Допоміжне джерело для експериментів і моделювання

Як видно з таблиці 2.1, для дослідження міграційних потоків українців доцільно використовувати не одне джерело, а їх комбінацію. Eurostat і UNHCR забезпечують міжнародну статистику, державні джерела дозволяють уточнювати внутрішній контекст, а HDX, ACAPS та відкриті дата-платформи дають змогу доповнювати дослідження гуманітарними, географічними та кризовими показниками. Такий підхід є особливо важливим для ВІ-системи, оскільки вона має не лише відображати числові показники, а й допомагати пояснювати причини та наслідки міграційних змін.

Дані є ключовим елементом будь-якого аналітичного дослідження. У випадку міграційних процесів вони дозволяють:

- оцінити масштаби переміщення населення;
- визначити географічні напрямки міграції;
- проаналізувати демографічну структуру мігрантів;
- виявити фактори, що впливають на міграцію;
- побудувати прогностичні моделі.

У рамках ВІ-систем дані виступають основою для побудови дашбордів, аналітичних звітів та систем підтримки прийняття рішень.

На основі представлених джерел доцільно детальніше розглянути ключові міжнародні статистичні ресурси, які є основою для аналізу міграційних процесів.

Одним із ключових джерел даних є Eurostat — офіційне статистичне агентство Європейського Союзу [21]. Воно забезпечує доступ до великої кількості структурованих даних щодо міграції, зокрема:

- кількість українців, які отримали тимчасовий захист;
- щомісячну динаміку змін;
- розподіл за країнами ЄС;
- соціально-демографічні характеристики.

Перевагою Eurostat є використання стандартизованих форматів даних, що дозволяє легко інтегрувати їх у аналітичні системи. Крім того, дані доступні через API, що є важливим для автоматизації збору інформації.

Іншим важливим джерелом є United Nations High Commissioner for Refugees (UNHCR). Ця організація надає:

- глобальну статистику щодо біженців;
- оперативні дані про переміщення населення;
- інтерактивні карти та візуалізації [40].

UNHCR активно використовує сучасні інструменти аналізу даних, що дозволяє отримувати актуальну інформацію у реальному часі.

Таким чином, міжнародні статистичні джерела забезпечують високий рівень достовірності, стандартизації та порівнюваності даних, що є критично важливим для побудови аналітичних моделей та ВІ-рішень.

Серед спеціалізованих джерел варто виділити ACAPS — організацію, що займається аналізом гуманітарних криз [20]. Її датасети містять:

- індикатори безпеки;
- економічні показники
- рівень гуманітарної допомоги;
- дані про конфлікти.

Ці дані дозволяють дослідити причинно-наслідкові зв'язки між подіями та міграційними процесами.

Ще одним важливим ресурсом є Humanitarian Data Exchange — платформа, що об'єднує відкриті гуманітарні дані [26]. Вона надає:

- геопросторові набори даних;
- статистику переміщення населення;
- інтегровані аналітичні інструменти.

Національні джерела даних включають:

- Державну службу статистики України;
- Міністерство соціальної політики;
- Державну прикордонну службу.

Вони надають інформацію про кількість внутрішньо переміщених осіб, обсяги перетину державного кордону, а також соціально-економічні

характеристики населення, що дозволяє комплексно досліджувати міграційні процеси та їх вплив на суспільство. Проте ці дані мають певні обмеження, зокрема:

- затримки у публікації;
- обмежений доступ;
- відсутність уніфікованих форматів.

Державна служба статистики України забезпечує доступ до офіційних демографічних та соціально-економічних показників, які можуть використовуватися для аналізу внутрішніх міграційних процесів населення [4].

Важливу роль у сучасних дослідженнях відіграють відкриті платформи, такі як Kaggle, які надають користувачам доступ до готових датасетів, очищених і структурованих даних, а також забезпечують можливість інтеграції з мовами програмування Python та R для подальшого аналізу й моделювання даних [29]. Такі платформи значно спрощують процес аналізу, особливо на етапі підготовки даних.

Дані про міграцію можуть бути представлені у різних форматах:

- CSV — для табличних даних;
- JSON — для складних структур;
- XML — для обміну між системами;
- API — для автоматичного отримання даних.

Однією з ключових задач є об'єднання даних з різних джерел. Це включає:

- приведення до єдиного формату;
- очищення від дублювання;
- нормалізацію значень;
- синхронізацію часових періодів.

Цей процес є основою для побудови якісної аналітичної системи.

Основними проблемами є:

- різна методологія збору даних;
- неповнота інформації;

- відсутність стандартів;
- складність інтеграції;
- можливі помилки у даних.

Для вирішення цих проблем використовуються методи попередньої обробки, які будуть розглянуті у наступному підпункті.

У контексті ВІ-систем джерела даних відіграють ключову роль, оскільки саме вони формують основу для:

- побудови дашбордів;
- аналізу трендів;
- прийняття управлінських рішень;
- прогнозування.

Використання якісних даних дозволяє підвищити точність аналітики та ефективність системи.

Таким чином, аналіз міграційних процесів українців базується на використанні широкого спектра джерел даних, серед яких ключову роль відіграють міжнародні організації (Eurostat, UNHCR), гуманітарні платформи (ACAPS, HDX), державні установи та відкриті дата-ресурси (Kaggle). Комплексне використання цих джерел дозволяє отримати повну, достовірну та актуальну інформацію, необхідну для побудови ефективної ВІ-системи та проведення глибокого аналітичного дослідження.

2.2. Методи обробки та аналізу даних

Методи обробки та аналізу даних, пов'язаних з міграційними потоками, проходять шість основних етапів, починаючи зі збору даних, їх очищення, інтеграції з даними з інших джерел, перетворення даних для використання в аналізі, проведення власне аналізу та, нарешті, візуалізації результатів. Усі ці етапи відображають деякі сучасні принципи побудови аналітичних систем; вони зазвичай використовуються в середовищах бізнес-аналітики.

Обробка даних відбувається поетапно, причому кожен етап пов'язаний з усіма попередніми етапами. На першому етапі визначаються джерела даних, такі як міжнародна статистика, урядові дані та дані з загальнодоступних джерел. Потім дані очищаються від помилок, відсутніх даних та дублікатів записів перед обробкою.

Інтеграція даних – це об'єднання даних з різних джерел в одну структуру даних. Наступним кроком є перетворення даних (наприклад, нормалізація, агрегація та розробка нових показників). Останні два кроки – це аналіз та візуалізація даних для отримання аналітичних результатів та їх чіткого відображення.

Такий підхід забезпечує систематичну обробку даних та підвищує достовірність результатів дослідження.

Процес збору інформації є початковим етапом проведення дослідження, і якість усіх майбутніх результатів залежить від того, наскільки добре була зібрана інформація. Дослідження міграційних моделей використовують дані, надані міжнародними організаціями, такими як Євростат або УВКБ ООН, а також інформацію з урядових баз даних та загальнодоступних джерел.

Ви можете отримувати дані в багатьох різних форматах (електронні таблиці, звіти або автоматизовані інтерфейси доступу). Одним з ключових питань є забезпечення способу оновлення даних достатньо часто, щоб забезпечити динамічний аналіз, який відображає поточний стан змін. (Це включає як постійні дані, так і тимчасові зміни.)

Організація процесу збору даних повинна забезпечувати їх повноту, актуальність та відповідність цілям дослідження.

Попередня обробка даних включає очищення, інтеграцію та стандартизацію інформації [16]. Очищення даних є критично важливим етапом, оскільки первинні дані часто містять помилки, пропущені значення або некоректні записи. Основними проблемами є:

- відсутність окремих значень;
- дублювання записів;

- різні формати представлення даних;
- логічні помилки у значеннях.

Для вирішення цих проблем застосовуються методи видалення або заповнення пропущених значень, усунення дублікатів, а також приведення даних до єдиного формату. Крім того, важливим є проведення перевірки достовірності даних та їх узгодженості.

Якісне очищення даних дозволяє значно підвищити точність подальшого аналізу та уникнути викривлення результатів.

Помилки у первинних даних можуть призводити до викривлення результатів аналізу, формування некоректних висновків та зниження достовірності прогнозів. Саме тому процес очищення даних розглядається як один із найбільш важливих етапів побудови аналітичних систем

У сучасній практиці для очищення даних часто використовуються автоматизовані сценарії обробки, які дозволяють мінімізувати ручне втручання та забезпечити повторюваність процесів підготовки інформації.

Інтеграція даних передбачає об'єднання інформації з різних джерел у єдину систему. Це є складним процесом, оскільки дані можуть мати різну структуру, формат та рівень деталізації.

Основними завданнями інтеграції є:

- узгодження назв змінних;
- об'єднання таблиць;
- синхронізація часових періодів;
- усунення суперечностей між джерелами.

У результаті інтеграції формується єдиний набір даних, який є основою для подальшого аналізу. Цей етап є ключовим для забезпечення цілісності та узгодженості інформації.

Після інтеграції дані потребують додаткової обробки для підготовки до аналізу. Трансформація включає:

- нормалізацію значень;

- агрегування даних;
- створення нових показників;
- зміну структури даних.

Наприклад, можуть бути розраховані такі показники, як темпи зростання міграції або середні значення за певний період. Це дозволяє більш глибоко аналізувати тенденції та виявляти закономірності. На цьому етапі здійснюється безпосередній аналіз підготовлених даних. Основними методами є:

- описова статистика;
- аналіз динаміки;
- кореляційний аналіз;

Описова статистика дозволяє отримати загальне уявлення про структуру даних, тоді як кореляційний аналіз допомагає виявити взаємозв'язки між показниками. Регресійний аналіз може використовуватися для оцінки впливу різних факторів на міграційні процеси. Крім базових методів аналізу, у сучасних дослідженнях активно застосовуються методи аналізу часових рядів, які дозволяють оцінювати зміни показників у довгостроковій перспективі. Для міграційних процесів часові ряди є особливо важливими, оскільки дають можливість визначати періоди різкого зростання або зниження кількості мігрантів, а також досліджувати вплив зовнішніх факторів на динаміку переміщення населення.

Важливим напрямом аналізу є також використання кластерного підходу, який дозволяє групувати об'єкти за схожими характеристиками. У контексті міграційних досліджень кластеризація може застосовуватися для сегментації країн за рівнем міграційного навантаження, виявлення подібних тенденцій або пошуку груп держав із близькими характеристиками міграційних процесів.

Методи прогнозного аналізу створюють можливість оцінювати потенційні зміни показників у майбутньому. Такі підходи використовуються для прогнозування чисельності мігрантів, оцінки тенденцій та формування сценаріїв

розвитку подій. Прогнозування особливо актуальне в умовах нестабільності, коли швидкість зміни показників суттєво зростає.

Таблиця 2.2 – Методи аналізу міграційних даних

Метод аналізу	Які задачі вирішує	Доцільність використання у дослідженні
Описова статистика	Узагальнення даних, базовий аналіз	Початковий аналіз структури міграції
Кореляційний аналіз	Виявлення взаємозв'язків між показниками	Аналіз впливу факторів (економіка, війна)
Регресійний аналіз	Оцінка впливу факторів на міграцію	Прогнозування міграційних процесів
Аналіз часових рядів	Дослідження змін у часі	Аналіз динаміки міграції
Кластерний аналіз	Групування об'єктів	Сегментація країн за міграційними показниками
Візуальний аналіз	Інтерпретація даних через графіки	Представлення результатів
ВІ-аналітика	Комплексний аналіз і інтеграція	Побудова дашбордів і прийняття рішень

Як видно з таблиці 2.2, різні методи аналізу даних мають свої особливості, переваги та обмеження. Їх комплексне використання дозволяє отримати більш повне уявлення про міграційні процеси, виявити закономірності та побудувати обґрунтовані аналітичні висновки. У ВІ-системах важливу роль відіграє інтеграція статистичних методів із візуальними інструментами аналізу. Це забезпечує не лише точність розрахунків, а й покращує сприйняття результатів користувачами.

Оскільки міграційні процеси мають часову природу, важливим є аналіз часових рядів [5]. Він дозволяє:

- виявити тренди;
- визначити сезонні коливання;
- прогнозувати майбутні значення.

Методи прогнозування дозволяють оцінювати потенційні зміни показників та майбутні тенденції [13]. Такий аналіз є особливо актуальним для дослідження змін у кількості мігрантів протягом часу.

Візуалізація є невід'ємною частиною аналізу, оскільки дозволяє представити складні дані у зрозумілій формі [11]. Основними інструментами є:

- графіки;
- діаграми;
- карти;
- інтерактивні панелі.

Використання візуалізації дозволяє швидко виявляти закономірності та тенденції, що значно спрощує інтерпретацію результатів.

ВІ-системи відіграють ключову роль у сучасному аналізі даних. Вони забезпечують:

- інтеграцію даних із різних джерел;
- автоматизацію процесів обробки;
- створення інтерактивних дашбордів;
- аналіз даних у реальному часі.

Використання ВІ-систем дозволяє значно підвищити ефективність аналізу та забезпечити підтримку прийняття управлінських рішень.

Застосування сучасних методів обробки та аналізу даних має низку переваг:

- підвищення точності результатів;
- автоматизація процесів;
- скорочення часу обробки;
- можливість роботи з великими обсягами даних;
- виявлення прихованих закономірностей [19].

Це робить їх незамінними у дослідженні складних соціально-економічних процесів, таких як міграція.

Таким чином, методи обробки та аналізу даних є ключовим елементом дослідження міграційних процесів українців. Використання комплексного підходу, що включає очищення, інтеграцію, трансформацію та аналіз даних, дозволяє отримати достовірні результати та забезпечити ефективне функціонування ВІ-систем. Це, у свою чергу, створює основу для прийняття обґрунтованих управлінських рішень у сфері міграційної політики.

3.3 Формування набору даних для ВІ-системи

Формування набору даних для ВІ-системи є стратегічним етапом дослідження, який визначає ефективність подальшого аналізу, якість візуалізації та можливості аналітичної інтерпретації результатів. На відміну від процесів обробки даних, які були розглянуті у попередньому підпункті, даний етап зосереджується на побудові цілісної інформаційної моделі, що забезпечує логічну організацію даних для їх використання у ВІ-середовищі.

У межах дослідження міграційних процесів українців формування набору даних передбачає створення структурованої аналітичної системи, яка інтегрує інформацію з різних джерел, зокрема Eurostat та United Nations High Commissioner for Refugees, і забезпечує можливість багатовимірного аналізу.

Формування набору даних базується на ряді принципів, які забезпечують його ефективність:

- орієнтація на аналітику — структура даних повинна відповідати задачам дослідження;
- мінімізація надлишковості — уникнення дублювання інформації;
- гнучкість структури — можливість розширення набору даних;
- масштабованість — здатність працювати з великими обсягами інформації;
- узгодженість — використання єдиних стандартів даних.

Дотримання цих принципів дозволяє створити набір даних, який відповідає вимогам сучасних ВІ-систем.

Правильна організація структури даних визначає не лише швидкість виконання аналітичних запитів, але й точність отриманих результатів. Наявність надлишкових зв'язків, дублювання інформації або неузгодженості показників може негативно впливати на ефективність роботи ВІ-систем.

Крім того, сучасні аналітичні системи повинні забезпечувати можливість масштабування, оскільки обсяги статистичних даних постійно зростають. Тому під час формування наборів даних важливо враховувати можливість подальшого розширення структури без втрати продуктивності систем.

Набір даних для ВІ-системи формується у вигляді багаторівневої структури, яка включає:

- рівень джерел даних;
- рівень інтеграції;
- аналітичний рівень.

На першому рівні знаходяться сирі дані з різних джерел. Другий рівень забезпечує їх об'єднання та узгодження. Третій рівень — це вже підготовлена модель, яка використовується у ВІ-системі.

Такий підхід дозволяє розділити процеси збору та аналізу даних, що підвищує ефективність системи.

Для реалізації аналітичної системи найбільш ефективною є зіркова схема (Star Schema), яка широко застосовується у ВІ-аналітиці [30]. Основні елементи моделі:

- фактова таблиця — містить числові показники;
- вимірні таблиці — описують контекст даних.

Фактова таблиця акумулює основні метрики, тоді як вимірні таблиці забезпечують можливість деталізації та фільтрації. Застосування зіркової схеми дозволяє:

- оптимізувати швидкість запитів;
- спростити структуру даних;
- забезпечити зручність використання для кінцевого користувача.

Такий підхід дозволяє реалізувати багатовимірний аналіз даних.

На основі сформованого набору даних визначаються ключові показники, які використовуються у Ві-системі:

- загальна кількість мігрантів;
- зміна кількості у часі;
- частка мігрантів за країнами;
- показники динаміки;
- індекси змін.

Розробка таких показників дозволяє перейти від сирих даних до аналітичної інформації.

Для забезпечення коректності аналізу важливо забезпечити узгодженість даних. Це досягається шляхом:

- використання стандартних кодів;
- уніфікації назв;
- синхронізації часових періодів;
- усунення суперечностей.

Цей етап є критично важливим для запобігання помилок у Ві-системі. У Ві-системі важливим є не лише набір даних, а й зв'язки між ними. Основними типами зв'язків є:

- один-до-багатьох;
- ієрархічні зв'язки;
- часові залежності.

Правильне налаштування зв'язків дозволяє виконувати складні аналітичні запити. Для підвищення продуктивності Ві-системи необхідно:

- зменшити обсяг даних;
- уникати дублювання;
- використовувати ключі;
- оптимізувати структуру таблиць.

Це дозволяє значно підвищити швидкість обробки запитів.

Формування набору даних повинно враховувати майбутню візуалізацію. Структура набору даних визначає не лише можливість подальшої візуалізації інформації, але й продуктивність аналітичної системи загалом. Неправильно сформовані взаємозв'язки між таблицями можуть призводити до збільшення часу обробки запитів та зниження ефективності роботи системи.

Під час формування наборів даних важливо враховувати принципи масштабованості та гнучкості. Це означає, що структура повинна підтримувати можливість додавання нових джерел інформації без необхідності повної перебудови моделі.

Важливим аспектом є також мінімізація дублювання інформації, оскільки надлишковість даних може негативно впливати на швидкість виконання розрахунків та збільшувати складність підтримки моделі.

Крім того, організація структури наборів даних повинна забезпечувати підтримку інтерактивної аналітики, включаючи роботу фільтрів, часових атрибутів та похідних показників.

Сучасні ВІ-платформи орієнтовані на використання багатовимірних моделей даних, які забезпечують одночасний аналіз інформації за різними параметрами. Це створює умови для реалізації більш складних сценаріїв дослідження та підтримує багаторівневий аналіз показників.

Таким чином, формування наборів даних є не лише технічним етапом підготовки інформації, але й важливим компонентом забезпечення ефективності подальшого аналізу. Структура даних повинна підтримувати:

- побудову часових графіків;
- створення карт;
- порівняльний аналіз;
- інтерактивні фільтри.

Таким чином, дані формуються з урахуванням кінцевого використання. Після формування набір даних інтегрується у ВІ-систему, що дозволяє:

- створювати дашборди;

- аналізувати тенденції;
- будувати прогнози;
- підтримувати прийняття рішень.

Це є кінцевою метою формування аналітичної моделі.

Отже, формування набору даних для ВІ-системи є складним процесом, що передбачає побудову логічної структури даних, їх узгодження та оптимізацію. Використання сучасних підходів до моделювання даних, зокрема зіркової схеми, дозволяє забезпечити ефективний аналіз міграційних процесів українців. Це створює основу для побудови аналітичних систем та прийняття обґрунтованих управлінських рішень.

РОЗДІЛ 3

РОЗРОБКА ВІ-СИСТЕМИ ДЛЯ АНАЛІЗУ МІГРАЦІЙНИХ ПОТОКІВ

3.1 Архітектура ВІ-системи

Архітектура ВІ-системи визначає спосіб збору, підготовки, інтеграції, аналізу та візуалізації даних. У межах даної роботи розроблено ВІ-систему для аналізу міграційних потоків українців у країнах Європейського Союзу в умовах повномасштабної війни. Основною метою системи є забезпечення інтерактивного аналізу статистичних показників міграції, виявлення тенденцій змін та візуального представлення закономірностей міграційних процесів.

Запропонована архітектура ВІ-системи рис. 3.1 демонструє основні складові ВІ-системи, що використовуються для аналізу даних.

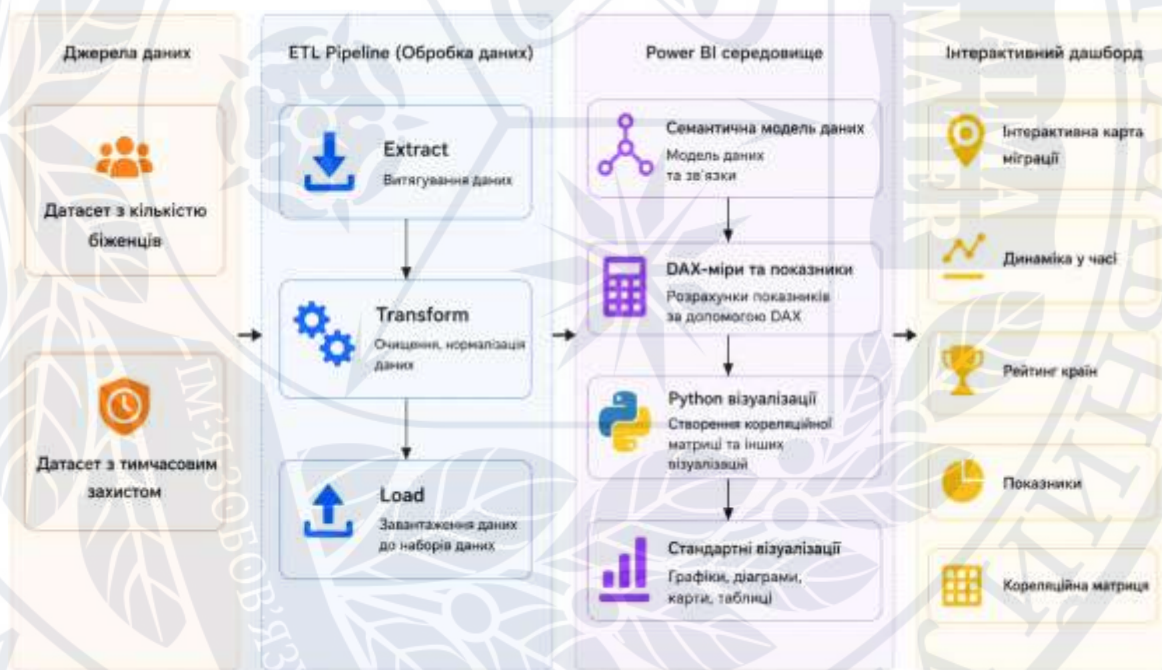


Рисунок 3.1 – Архітектура ВІ-системи

Основним джерелом інформації стали відкриті статистичні дані Eurostat, які містять інформацію щодо міграційних процесів українців у країнах Європейського Союзу.

Для формування комплексної аналітичної системи використовувались декілька наборів даних, які містять показники. Використання декількох датасетів

дозволило здійснити інтеграцію показників та сформувати ширший аналітичний контекст.

Підготовка даних здійснювалася із використанням Python, що дозволило автоматизувати процес обробки даних перед їх завантаженням у Power BI [36].

Основні етапи ETL включали:

- очищення пропущених значень;
- перетворення форматів;
- стандартизацію назв колонок;
- агрегування показників;
- створення нових змінних;
- підготовку часових атрибутів [18].

Використання Python дозволило мінімізувати помилки ручної обробки даних та забезпечити повторне використання сценаріїв. Приклад одного зі сценаріїв наведено в додатку. Частина операцій з підготовки даних, зокрема видалення дублікатів та очищення наборів даних, виконувалася окремо в Python, тоді як Power BI використовувався для подальшої візуалізації та аналізу даних.

Після підготовки дані були завантажені до Power BI, де сформовано модель даних.

Для забезпечення часового аналізу використовувались календарні атрибути, зокрема дата, рік, місяць та день. Це дозволило реалізувати аналіз динаміки міграції.

Аналітичний рівень реалізовано за допомогою DAX-формул, Python-сценаріїв та вбудованих інструментів Power BI. DAX забезпечує створення обчислюваних показників та реалізацію складних аналітичних розрахунків у Power BI [31]. У роботі реалізовано KPI-показники, кореляційний аналіз, аналіз часової динаміки міграційних процесів, геопросторову візуалізацію та сегментацію даних за країнами. Для розрахунків використовувались показники загальної кількості мігрантів, середнього приросту, середнього міграційного потоку та розмаху міграції. Частина попередньої обробки даних, зокрема

очищення наборів даних і видалення дублікатів, виконувалася за допомогою Python, тоді як Power BI використовувався для аналізу та візуалізації результатів. Рівень візуалізації

Фінальний рівень архітектури представлений інтерактивним дашбордом у Power BI. Інтерактивність системи дозволяє аналізувати показники залежно від вибраних країн та часових інтервалів

3.2 Реалізація аналітичного дашборду

Одним із ключових етапів розробки BI-системи є створення аналітичного дашборду, який забезпечує інтерактивну візуалізацію та аналіз міграційних процесів українців у країнах Європейського Союзу. У межах роботи дашборд реалізовано у середовищі Power BI Desktop із використанням відкритих статистичних даних Eurostat, підготовлених та оброблених за допомогою Python.

Основною метою створеного дашборду є забезпечення інструментів для аналізу динаміки міграційних потоків, дослідження географічного розподілу українців у країнах ЄС, оцінки міграційного навантаження, проведення кореляційного аналізу показників та інтерактивної роботи з даними для виявлення особливостей міграційних процесів. Побудова ефективних дашбордів передбачає дотримання принципів зрозумілої візуалізації та зручності інтерпретації показників [42].

На відміну від традиційних статичних звітів, розроблений дашборд дозволяє користувачу змінювати параметри аналізу, використовувати фільтри та досліджувати окремі аспекти міграційних процесів.

Для забезпечення коректного функціонування аналітичного дашборду та формування достовірних результатів аналізу було реалізовано послідовний процес підготовки, інтеграції та обробки даних.

Перед створенням візуалізацій було виконано обробку даних засобами Python. Вихідні датасети Eurostat містили різні формати дат, пропущені значення, неоднорідні структури таблиць та потребували інтеграції між собою.

Для підготовки використовувались такі етапи:

- очищення пропущених значень;
- перетворення форматів даних;
- агрегування показників;
- створення часових атрибутів;
- формування підготовлених таблиць для аналізу.

У роботі використовувались датасети:

- датасет з даними з кількістю біженців – головний;
- датасет з даними про осіб під тимчасовим захистом – додатковий [22].

Статистика щодо дозволів на проживання є важливим джерелом для дослідження довгострокових міграційних тенденцій та оцінки інтеграції мігрантів у країнах перебування [23]. Структуру використаних наборів даних наведено на рис. 3.2 та рис. 3.3

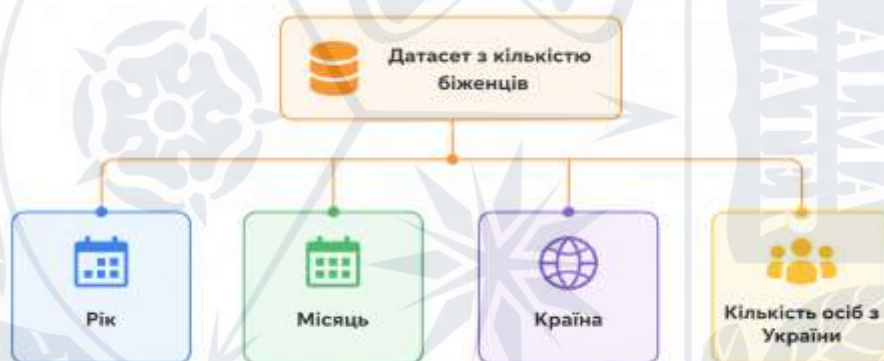


Рисунок 3.2 – Структура головного датасету



Рисунок 3.3 – Структура додаткового датасету

Для автоматизації процесу підготовки даних було розроблено Python-код, який виконує очищення, перетворення форматів, агрегування показників та формування підготовлених таблиць для подальшого імпорту в Power BI. Повний код обробки даних наведено у Додатку А, а фрагмент його реалізації подано на рис. 3.4.

```
for df in [migrants, residence, protection]:
    df.drop_duplicates(inplace=True)
    df.dropna(subset=["Country", "Year", "Value"], inplace=True)
    df["Year"] = pd.to_numeric(df["Year"], errors="coerce")
    df["Value"] = pd.to_numeric(df["Value"], errors="coerce")
```

Рисунок 3.4 – Фрагмент Python-коду підготовки даних

Основними етапами підготовки даних стали:

- очищення від пропущених значень;
- приведення показників до єдиного формату;
- агрегування статистичних показників;
- створення нових аналітичних показників.

Особливе значення мало узгодження часових інтервалів між різними datasets, оскільки частина джерел використовувала щомісячні значення, тоді як інші містили агреговані показники за більші часові проміжки. Це потребувало додаткового перетворення структури даних для забезпечення коректності аналізу.

Процес інтеграції даних дозволив сформувати єдину інформаційну модель, яка надалі використовувалася для побудови взаємозв'язків між таблицями у Power BI. Для забезпечення ефективності роботи системи використовувалися довідникові таблиці часу та країн, а також відповідні зв'язки між фактологічними таблицями. Це дозволило реалізувати централізовану систему фільтрації, покращити швидкість виконання запитів та забезпечити узгодженість результатів аналізу.

Після завершення етапу інтеграції було сформовано набір ключових показників, необхідних для дослідження міграційних процесів. До них належать загальна кількість українців, середні значення, показники зміни у часі, частка

окремих країн та інші похідні показники. Формування таких метрик дозволило перейти від первинних статистичних даних до більш глибокого аналітичного аналізу.

Завдяки комплексному підходу до підготовки та інтеграції даних було створено основу для подальшого використання DAX-розрахунків, побудови інтерактивних візуалізацій та реалізації багатовимірного аналізу у Power BI. Для забезпечення коректної роботи візуалізацій використовувались:

- DAX-міри;
- часові атрибути;
- обчислювані показники;
- інтегровані зв'язки між datasets.

Це забезпечило можливість централізованої фільтрації та узгодженого аналізу.

Для забезпечення коректної роботи візуалізацій та реалізації інтерактивного аналізу у Power BI використовувались різні елементи аналітичної моделі, які забезпечують узгодженість показників, автоматичне оновлення розрахунків та підтримку багатовимірного аналізу. До основних компонентів належать:

- DAX-міри — використовувались для створення обчислюваних показників, автоматичного перерахунку значень залежно від вибраних фільтрів та реалізації KPI-показників. Застосування DAX дозволяє формувати складні аналітичні розрахунки та забезпечує динамічне оновлення результатів аналізу;
- часові атрибути — використовувались для дослідження динаміки міграційних процесів у часовому розрізі. Формування полів року, місяця та дати дозволило реалізувати аналіз часових рядів, дослідження тенденцій та оцінку змін показників у різні періоди;
- обчислювані показники — застосовувались для визначення середніх значень, показників приросту, часток країн у загальних показниках, середнього міграційного навантаження та інших аналітичних характеристик;

- інтегровані зв'язки між datasets — забезпечували коректну взаємодію між таблицями, узгодженість результатів фільтрації та можливість комплексного аналізу показників із різних джерел даних.

Використання зазначених компонентів дозволило реалізувати централізовану систему фільтрації, забезпечити коректність розрахунків та підвищити ефективність аналізу даних. Завдяки цьому користувач може досліджувати показники залежно від вибраних часових інтервалів, країн або окремих категорій даних без втрати узгодженості результатів.

Розроблений аналітичний дашборд має багатокомпонентну структуру та складається з декількох логічних блоків, кожен із яких виконує окрему функцію в процесі аналізу міграційних процесів. Подібний підхід забезпечує структурованість представлення інформації та дозволяє користувачу швидко отримувати ключові аналітичні висновки.

У верхній частині дашборду реалізовано KPI-картки, які використовуються для швидкої оцінки поточного стану міграційних показників та дозволяють отримати узагальнену інформацію без необхідності детального аналізу графіків. Мова DAX є одним із ключових інструментів аналітики в середовищі Power BI. Використання DAX дозволяє створювати складні обчислення, реалізовувати похідні показники та автоматизувати розрахунки.

Однією з переваг DAX є можливість виконання контекстних обчислень, коли результати змінюються залежно від вибраних параметрів фільтрації. Це створює умови для реалізації багатовимірного аналізу та забезпечує гнучкість дослідження.

У межах роботи DAX використовувався для формування KPI, часових показників та похідних характеристик. До основних KPI належать:

- загальна кількість українців;
- середній приріст показників;
- частка країни у загальному показнику (%);
- середнє міграційне навантаження;
- похідні часові показники;

- додаткові розрахункові характеристики, сформовані за допомогою DAX.

Використання KPI-карток дозволяє швидко оцінювати зміни показників, порівнювати результати між різними періодами та виявляти основні тенденції міграційних процесів. Подібний формат подання інформації значно спрощує сприйняття великих обсягів статистичних даних та сприяє оперативному прийняттю аналітичних рішень.

Для реалізації KPI використовувалися DAX-міри, які забезпечують автоматичне оновлення показників залежно від вибраних параметрів фільтрації та дозволяють адаптувати результати аналізу до різних сценаріїв використання. Основні розрахункові показники та відповідні DAX-вирази наведено на рисунках 3.5–3.8.

```
Total Ukrainians = SUM('MigrationData'[Value])
```

Рисунок 3.5 – DAX-міра для обчислення загальної кількості українців

```
Country Share % =  
DIVIDE(  
    SUM('MigrationData'[Value]),  
    CALCULATE(  
        SUM('MigrationData'[Value]),  
        ALL('MigrationData'[Country])  
    )  
    ) * 100
```

Рисунок 3.6 – DAX-міра для розрахунку частки країни

```
Share of Ukrainians (%) =  
DIVIDE(  
    SUM('migr_asytpsm'[Value]),  
    CALCULATE(  
        SUM('migr_asytpsm'[Value]),  
        ALL('migr_asytpsm')  
    ),  
    0  
    ) * 100
```

Рисунок 3.7 – DAX-міра для визначення частки українців (%)

```

Max Migration Flow =
MAX(
    'migr_asytpfm'[Value]
)

```

Рисунок 3.8 – DAX-міра для визначення максимального міграційного потоку

Одним із центральних елементів є карта Європи, що відображає розподіл українців між країнами ЄС. Для покращення сприйняття використано темний стиль оформлення та синю кольорову градацію.

Карта дозволяє:

- визначати країни з найбільшою кількістю українців;
- оцінювати концентрацію міграційних потоків;
- аналізувати просторовий розподіл.

Для аналізу змін у часі було створено лінійний графік динаміки показників, який дозволяє відстежувати тенденції зміни кількості українців протягом досліджуваного періоду. Лінійна візуалізація забезпечує можливість:

- виявлення різких змін показників;
- аналізу періодів стабілізації;
- визначення загальних тенденцій.

Лінійні графіки є одним із найбільш ефективних інструментів для аналізу часових рядів, оскільки дозволяють візуально відображати зміну показників у динаміці та швидко визначати періоди різких змін або стабілізації. У контексті дослідження міграційних процесів використання часових графіків забезпечує можливість оцінювати довгострокові тенденції та аналізувати вплив зовнішніх факторів на зміну кількості українців у країнах Європейського Союзу.

Крім визначення загальних тенденцій, часові графіки дозволяють виявляти аномальні зміни показників, які можуть бути пов'язані зі змінами політики тимчасового захисту, економічною ситуацією або іншими факторами. Це підвищує аналітичну цінність дашборду та забезпечує більш глибоке розуміння досліджуваних процесів.

Таким чином, аналіз часових рядів дозволяє переходити від простого спостереження за статистичними показниками до інтерпретації закономірностей розвитку міграційних процесів.

Для проведення поглибленого статистичного аналізу використано Python Visual, за допомогою якого реалізовано кореляційну матрицю.

Кореляційний аналіз є важливим інструментом статистичного дослідження, який дозволяє оцінювати ступінь взаємозв'язку між окремими показниками. Використання кореляційної матриці забезпечує можливість швидкого виявлення позитивних, негативних або слабких залежностей між змінними.

У межах дослідження кореляційна матриця використовувалась для оцінки взаємозв'язків між часовими параметрами та міграційними показниками. Аналіз отриманих результатів показує, що зміни міграційних процесів формуються під впливом комплексу факторів, а не визначаються окремими змінними.

Варто зазначити, що слабкі значення кореляції не означають відсутності впливу, оскільки соціально-економічні процеси мають багатфакторний характер. На зміну показників можуть одночасно впливати безпекова ситуація, економічні умови, державна політика та інші зовнішні чинники.

Отже, використання кореляційного аналізу дозволяє не лише оцінювати взаємозв'язки між показниками, але й формувати гіпотези щодо причин зміни міграційних тенденцій.

Побудована матриця дозволяє оцінити взаємозв'язки між часовими параметрами та міграційними показниками.

Крім того, реалізовано порівняльний аналіз країн за показниками середнього навантаження, частки (%) та загальної кількості українців. Така візуалізація дозволяє визначати країни з найбільшим міграційним навантаженням, оцінювати відмінності між державами та досліджувати нерівномірність розподілу міграційних процесів у країнах Європейського Союзу. Використання порівняльного аналізу забезпечує можливість виявлення країн, які

прийняли найбільшу кількість українців, а також оцінки потенційного впливу міграційних процесів на соціально-економічні системи приймаючих держав.

Для підвищення функціональності дашборду та забезпечення гнучкої взаємодії з даними реалізовано інтерактивні елементи керування, які дозволяють адаптувати результати аналізу відповідно до потреб користувача. Зокрема, реалізовано:

- фільтрацію за країнами, що забезпечує можливість аналізу показників окремих держав та проведення міжкраїнного порівняння;
- вибір часових періодів, який дозволяє досліджувати динаміку змін показників за різні роки та виявляти тенденції розвитку міграційних процесів;
- аналіз окремих наборів даних (datasets), що забезпечує можливість порівняння різних міграційних показників та їх взаємозв'язків;
- зміну параметрів відображення, яка дозволяє деталізувати результати аналізу залежно від поставлених дослідницьких завдань.

Реалізовані інтерактивні механізми забезпечують адаптацію дашборду до різних сценаріїв використання та дозволяють виконувати як загальний огляд міграційних процесів, так і поглиблений аналіз окремих показників. Завдяки цьому користувач отримує можливість оперативно досліджувати тенденції, виявляти закономірності та формувати висновки на основі статистичних даних.

Особлива увага приділялась оформленню дашборду. Під час розробки інтерфейсу дашборду враховувались принципи сучасного інформаційного дизайну та UX-аналітики. Візуальна структура аналітичної системи повинна забезпечувати швидке сприйняття інформації, мінімізувати когнітивне навантаження та підтримувати логічну послідовність аналізу.

Для оформлення дашборду використовувалась світла кольорова тема, яка забезпечує покращену читабельність візуалізацій та спрощує сприйняття аналітичної інформації. Для виділення ключових показників застосовувалась синя кольорова палітра з акцентними відтінками, що дозволяє підвищити контрастність графічних елементів та забезпечити цілісність дизайну. Використання єдиного стилю оформлення сприяє покращенню візуального

сприйняття даних та підвищує зручність роботи з аналітичним дашбордом. Приклад використаних кольорів рис 3.9:

```

48
49 # світла палітра
50 colors = [
51     "#FFFFFF",
52     "#DBEAFE",
53     "#60A5FA",
54     "#1D4ED8"
55 ]
56
57 ]
58
59 custom_cmap = LinearSegmentedColormap.from_list(
60     "blue_theme",
61     colors
62 )
63
64

```

Рисунок 3.9 – Задання дизайну для графіка

Розташування КРІ-карток у верхній частині інтерфейсу забезпечує можливість швидкого отримання узагальненої інформації, тоді як основні графіки дозволяють переходити до більш деталізованого аналізу. Подібний підхід відповідає принципу ієрархії інформації, який широко використовується при створенні сучасних ВІ-рішень.

Таким чином, оформлення дашборду є не лише елементом дизайну, але й важливим компонентом підвищення ефективності аналітичного процесу. Подібний підхід дозволяє:

- покращити сприйняття інформації;
- підвищити зручність аналізу;
- забезпечити цілісність дизайну.

Таким чином, у межах роботи реалізовано інтерактивний аналітичний дашборд на основі Power BI, Python та даних Eurostat. Розроблена система включає КРІ-показники, часовий аналіз, картографічні візуалізації, кореляційний аналіз, прогнозування та інтерактивні елементи, що дозволяють виконувати комплексне дослідження міграційних процесів українців у країнах Європейського Союзу.

3.3 Аналіз отриманих результатів

У результаті реалізації ВІ-системи було сформовано інтерактивний аналітичний дашборд, який дозволяє досліджувати міграційні процеси українців у країнах Європейського Союзу в умовах повномасштабної війни. Отримані результати відображають просторові, часові та статистичні закономірності змін кількості українців під тимчасовим захистом. Повну інтерактивну версію розробленого дашборду з можливістю фільтрації, детального перегляду показників та аналізу міграційних процесів представлено у Додатку А.

На основі КРІ-карток встановлено, що загальна кількість українців під тимчасовим захистом у країнах Європейського Союзу становить приблизно 9,69 млн осіб. Аналіз також показав, що дослідження охоплює 27 країн, а середній приріст показників за останній період становить близько 1,2 %.

Отримані значення свідчать про збереження міграційної активності, хоча темпи приросту поступово стабілізуються.

Стабілізація темпів приросту кількості українців під тимчасовим захистом може свідчити про зміну характеру міграційних процесів. Якщо на початковому етапі повномасштабної війни міграція мала переважно екстрений характер, то згодом почали формуватися більш довгострокові тенденції перебування українців у країнах Європейського Союзу.

На початку військових дій значна частина громадян залишала територію України через безпекові ризики, невизначеність та необхідність забезпечення базових потреб. У результаті спостерігалось різке збільшення кількості осіб, які отримували тимчасовий захист. Надалі темпи приросту почали поступово знижуватися, що може пояснюватися декількома причинами.

По-перше, частина громадян адаптувалася до умов перебування за кордоном та інтегрувалася у приймаючі суспільства. Інтеграційні процеси включають працевлаштування, отримання освіти, доступ до медичних послуг та формування соціальних зв'язків. У результаті тимчасове перебування поступово може переходити у більш довгострокові форми проживання.

По-друге, стабілізація показників може бути пов'язана зі змінами міграційної політики країн Європейського Союзу. Механізми тимчасового захисту, програми підтримки та законодавчі особливості можуть впливати на швидкість зміни кількості мігрантів.

По-третє, частина населення поступово повертається в Україну, що також впливає на зміну темпів приросту показників. Рішення щодо повернення можуть залежати від безпекової ситуації, економічних факторів та особистих обставин.

Таким чином, отримані результати демонструють складний характер міграційних процесів, які формуються під впливом комплексу взаємопов'язаних факторів та потребують постійного моніторингу.

Картографічна візуалізація (рис.3.10) дозволила визначити країни з найбільшою концентрацією українців. Найбільше навантаження спостерігається у:

1. Німеччині;
2. Польщі;
3. Чехії;
4. Іспанії;
5. Румунії.

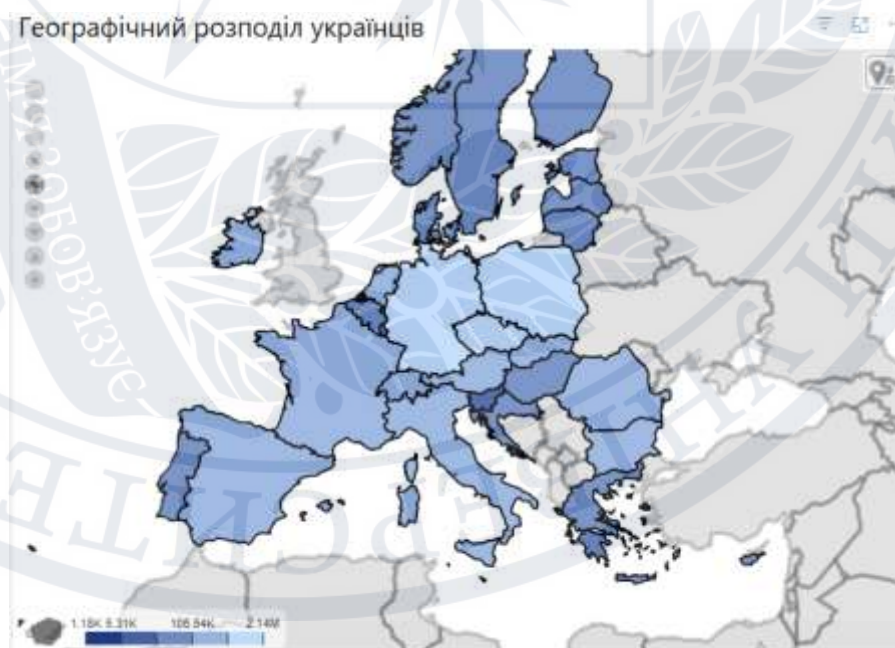


Рисунок 3.10 – Географічний розподіл

Просторовий аналіз показав нерівномірний розподіл міграційних потоків між країнами Європейського Союзу. Найбільша концентрація українців спостерігається у країнах, які характеризуються поєднанням географічної близькості, розвиненої економіки та наявності програм підтримки.

Польща стала одним із ключових напрямків через спільний кордон з Україною та спрощені механізми в'їзду після початку повномасштабної війни. Крім географічного чинника, важливу роль відіграють культурна близькість та схожість мовного середовища.

Німеччина демонструє високі показники завдяки розвинутій системі соціального забезпечення, можливостям працевлаштування та високому рівню економічного розвитку. Це робить країну одним із найбільш привабливих напрямків довгострокового перебування.

Подібна концентрація населення може створювати додаткове навантаження на соціальну інфраструктуру приймаючих держав. До основних сфер впливу належать:

- система освіти;
- охорона здоров'я;
- ринок праці;
- житловий сектор;
- соціальне забезпечення.

Водночас присутність великої кількості мігрантів може позитивно впливати на економіку країн через збільшення трудових ресурсів та споживчої активності.

Збільшення концентрації українців у цих країнах може впливати на соціальну інфраструктуру, ринок праці та системи тимчасового захисту.

Результати аналізу рис. 3.11 показують значне домінування окремих країн над іншими за кількістю українців. Найбільші значення спостерігаються для:

1. Німеччини;
2. Польщі;
3. Чехії.

При цьому різниця між країнами-лідерами та іншими країнами є суттєвою, що вказує на нерівномірний розподіл міграційних потоків.



Рисунок 3.11 – Топ 10 країн за кількістю біженців

Лінійний графік (рис. 3.12) відображає динаміку кількості українців у країнах Європейського Союзу протягом досліджуваного періоду. Найбільш різке зростання показників спостерігається у 2022 році, що пов'язано з початком повномасштабного вторгнення та масовим переміщенням населення. Внаслідок цього країни ЄС активували механізми підтримки, які забезпечили українцям можливість легального перебування, доступ до працевлаштування, освіти та соціальних послуг [6].

Після періоду стрімкого зростання простежується поступове зниження темпів приросту та перехід до більш стабільної динаміки.

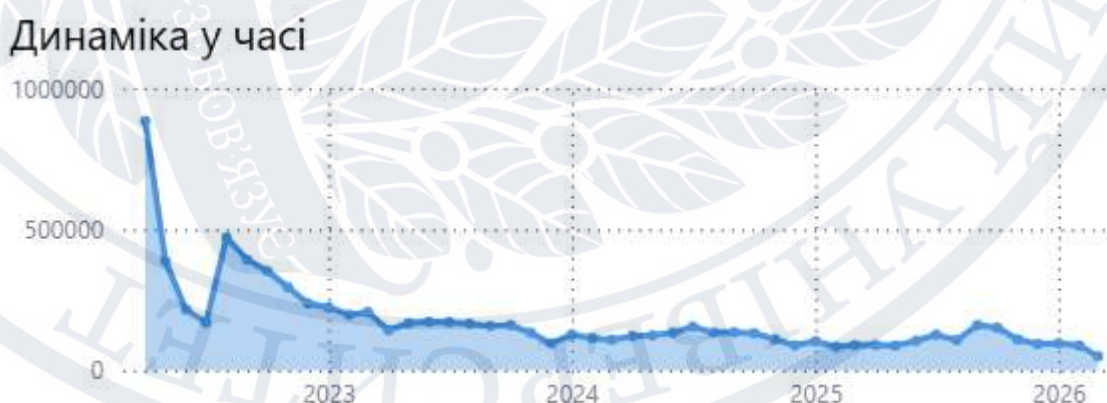


Рисунок 3.12 – Лінійний графік розподілу по часу

Як видно з рис. 3.12, після початкового пікового періоду відбувається стабілізація показників, що може свідчити про завершення фази різкого

збільшення міграційних потоків і формування більш сталих тенденцій перебування українців за кордоном. Отримані результати дозволяють припустити:

- стабілізацію міграційних процесів у країнах Європейського Союзу;
- адаптацію частини українців до соціально-економічних умов країн перебування;
- збільшення частки довгострокового перебування порівняно з короткостроковими переміщеннями;
- поступове формування нових моделей міграційної поведінки населення.

Таким чином, результати, представлені на рис. 3.12, демонструють поступовий перехід від різкого зростання міграційних потоків до більш стабільної динаміки.

Для дослідження взаємозв'язків між часовими та міграційними показниками було побудовано кореляційну матрицю (рис. 3.13). Аналіз отриманих результатів показав наявність сильної позитивної залежності між поточними та попередніми значеннями показників, що може свідчити про збереження загальних тенденцій зміни міграційних процесів у часі.

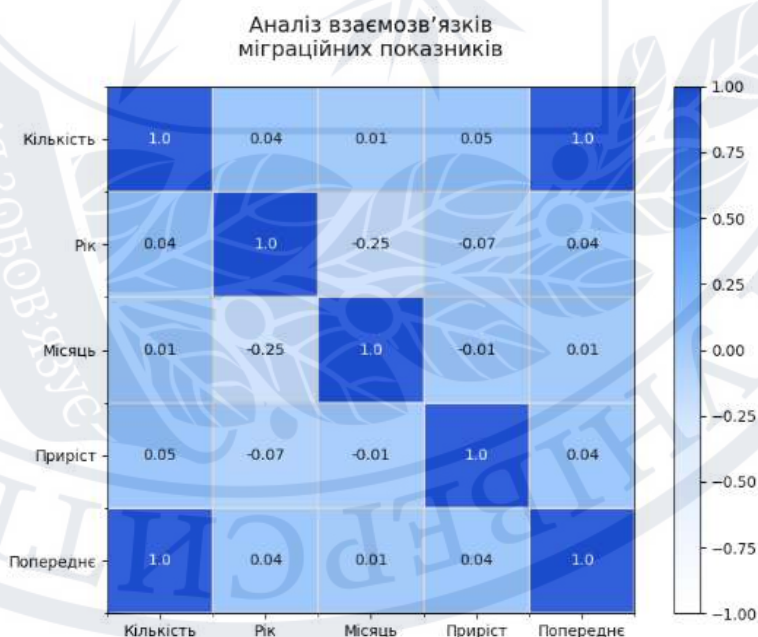


Рис 3.13– Кореляційна матриця

Як видно з рис. 3.13, між більшістю часових характеристик (рік, місяць) та міграційними показниками спостерігається низький рівень кореляції, що вказує на відсутність сильних лінійних взаємозв'язків між окремими змінними. Крім того, суттєвих негативних кореляцій також не виявлено, оскільки більшість значень перебувають поблизу нульової позначки.

Отримані результати дозволяють зробити висновок, що зміни міграційних показників мають відносно стабільний характер і можуть визначатися сукупністю зовнішніх факторів, а не впливом окремих часових параметрів. Це підтверджує складну багатофакторну природу міграційних процесів українців у країнах Європейського Союзу.

Для поглибленого аналізу міграційних процесів було сформовано рейтинг країн за середнім навантаженням, часткою та загальною кількістю українців у країнах Європейського Союзу (рис. 3.14). Отримані результати показують значну нерівномірність розподілу міграційних потоків між країнами.



Країни	Середнє навантаження	Частка (%)	Загальна кількість
Germany	746,769.24	13	98573540
Poland	664,454.66	13	97674835
Czech Republic	250,418.40	5	36811505
Spain	128,722.11	3	18922150
Romania	94,228.88	2	13851645
Italy	83,180.37	2	12227515
Slovakia	76,753.23	1	11052465
Netherlands	69,427.28	1	10205810
Bulgaria	68,210.58	1	10026955
Ireland	60,515.65	1	8895800
Austria	54,382.37	1	7341620
Belgium	49,061.84	1	7212090
Switzerland	41,939.59	1	6165120
Lithuania	39,473.74	1	5802640
Norway	39,174.25	1	5758615
France	38,828.71	1	5707820
Finland	38,004.66	1	5586685

Рисунок 3.14 – Порівняльна таблиця країн

Як видно з рис. 3.14, найбільше навантаження та загальна кількість українців спостерігаються у Німеччині, яка займає першу позицію за всіма основними показниками. Результати аналізу підтвердили концентрацію

основних міграційних потоків у Німеччині, Польщі та Чеській Республіці. Проведений аналіз дозволив виявити нерівномірність розподілу громадян України між країнами ЄС та визначити держави з найбільшим міграційним навантаженням.

Країни Західної та Північної Європи, зокрема Ірландія, Австрія, Бельгія та Швейцарія, демонструють нижчі абсолютні показники, однак залишаються важливими напрямками міграції українського населення. Водночас частка окремих країн у загальному розподілі міграційних потоків є відносно невеликою порівняно з лідерами рейтингу.

Отримані результати свідчать про концентрацію основних міграційних потоків у декількох країнах ЄС, тоді як інші держави мають значно менше навантаження. Це може вказувати на нерівномірний розподіл. Таким чином, розроблений дашборд забезпечує комплексне представлення міграційних показників та дозволяє досліджувати тенденції зміни кількості українців у країнах ЄС.

ВИСНОВКИ

У результаті виконання кваліфікаційної роботи розроблено програмно-аналітичну ВІ-систему для аналізу міграційних процесів українців на основі відкритих статистичних даних Eurostat. У роботі досліджено сучасні підходи до побудови систем Business Intelligence, визначено функціональні та інформаційні вимоги до системи аналізу міграційних процесів та сформовано концептуальну модель дослідження.

Для підготовки даних реалізовано ETL-процес із використанням мови програмування Python. Виконано очищення даних від пропусків і дублікатів, перетворення форматів атрибутів та підготовку наборів даних до подальшого аналізу. На основі підготовлених даних побудовано модель даних у Power BI та реалізовано систему аналітичних показників і KPI із використанням мови DAX.

Створено інтерактивний аналітичний дашборд, який містить KPI-картки, часові графіки, геопросторові візуалізації, рейтингові таблиці та засоби фільтрації даних. Розроблена система забезпечує аналіз динаміки міграційних процесів, міжкраїнні порівняння та дослідження просторового розподілу міграційних потоків.

У результаті проведеного дослідження визначено основні країни концентрації громадян України під тимчасовим захистом, проаналізовано часову динаміку міграційних процесів та оцінено міграційне навантаження між країнами Європейського Союзу. Отримані результати дозволили виявити ключові тенденції та закономірності міграційних процесів і підтвердили ефективність використання ВІ-технологій для аналізу великих масивів статистичних даних.

Практична цінність роботи полягає у можливості використання розробленої ВІ-системи як інструменту підтримки прийняття управлінських рішень державними органами України, міжнародними гуманітарними організаціями та громадськими об'єднаннями. Зокрема, геопросторова візуалізація розподілу громадян України за країнами перебування (рис. 3.4) та рейтинг країн за кількістю осіб під тимчасовим захистом (рис. 3.5) дозволяють визначати держави з найбільшою концентрацією українців. На основі цих даних

можуть прийматися стратегічні рішення щодо відкриття нових консульських пунктів, оптимізації роботи вже існуючих консульських установ, а також посилення дипломатичної підтримки громадян України в країнах, де зосереджена найбільша кількість українців, зокрема в Німеччині, Польщі та Чехії.

KPI-картки та графіки часової динаміки (рис. 3.6) можуть використовуватися для прийняття рішень щодо гуманітарної, соціальної та фінансової підтримки. Вони дозволяють визначати країни з найбільшою потребою в допомозі, планувати розподіл ресурсів та оперативно реагувати на зміни міграційних показників.

Аналіз часових рядів і тенденцій зміни міграційних показників (рис. 3.6) може використовуватися для планування програм повернення та реінтеграції громадян України, а також оцінювання потенційного повернення трудових ресурсів і підтримки процесів відновлення економіки України.

Порівняльний аналіз показників між країнами та оцінка міграційного навантаження (рис. 3.8) можуть застосовуватися для адаптації комунікаційної політики, підготовки інформаційних матеріалів та поширення успішних практик підтримки громадян України.

Результати кореляційного аналізу та узагальнені показники дашборду (рис. 3.7) можуть бути використані для перевірки аналітичних гіпотез, підвищення достовірності висновків та формування інформаційних повідомлень на основі об'єктивних даних.

Отже, поставлену мету роботи досягнуто шляхом проєктування та реалізації BI-системи аналізу міграційних процесів із використанням Python, Power BI та сучасних методів аналітичної обробки даних. Усі поставлені завдання виконано, а отримані результати підтверджують доцільність використання BI-технологій для створення інформаційно-аналітичних систем підтримки прийняття рішень.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Биков В. Ю. Цифрова трансформація суспільства і розвиток інформаційно-аналітичних систем. Інформаційні технології і засоби навчання. 2022. № 4. С. 15–28.
2. Васильєва Т. А. Аналіз даних та системи підтримки прийняття рішень. Суми : СумДУ, 2021. 356 с.
3. Глибовець М. М. Аналіз даних : навч. посіб. Київ : НаУКМА, 2020. 287 с.
4. Державна служба статистики України. Демографічна та соціальна статистика. URL: <https://data.gov.ua/dataset/f373010c-0dc1-4060-bd13-79cea54ee5d5>(дата звернення: 13.05.2026).
5. Дуброва О. В. Методи аналізу часових рядів у соціально-економічних дослідженнях. Економіка та суспільство. 2021. № 30. С. 45–53.
6. Європейська комісія. Тимчасовий захист для громадян України в ЄС. URL: https://commission.europa.eu/topics/eu-solidarity-ukraine/eu-assistance-ukraine/information-people-fleeing-war-ukraine/fleeing-ukraine-your-rights-eu_uk(дата звернення: 13.05.2026).
7. Захарова К. В., Січко Т. В. Використання інструменту POWER BI для аналізу фармацевтичного ринку. Прикладні аспекти сучасних міждисциплінарних досліджень. 2021. С. 97–99.
8. Коваленко В. В. Бізнес-аналітика та ВІ-системи : навч. посіб. Харків : ХНЕУ, 2022. 320 с.
9. Литвин В. В. Інтелектуальний аналіз даних : підручник. Львів : Новий Світ-2000, 2021. 406 с.
10. Матвійчук О. В. Аналітика даних у Power BI : навч. посіб. Київ : КПІ ім. Ігоря Сікорського, 2023. 210 с.
11. Морзе Н. В. Візуалізація даних та інформаційні панелі. Інформаційні технології в освіті. 2022. № 2. С. 33–47.
12. Олійник А. О. Сучасні ВІ-системи для аналізу великих даних. Наукові праці ЧДУ. 2023. № 1. С. 66–74.

- 13.Петренко І. В. Методи прогнозування в аналітичних системах. Комп'ютерно-інтегровані технології. 2021. № 5. С. 90–101.
- 14.Пономаренко В. С. Інформаційні системи і технології в економіці. Харків : ХНЕУ, 2020. 512 с.
- 15.Приймак В. І., Рибак Ю. М., Голубник О. Р. Сучасні підходи до зберігання даних на розподіленій інфраструктурі. Економіка і регіон. – 2023. – №4(91). – С. 236–242.
- 16.Семенюк О. А., Кирилащук Т. Г., Січко Т. В. Прикладні аспекти обробки даних в інформаційних системах. Комп'ютерні технології обробки даних. 2021. С. 212–213.
- 17.Синиця К. М. Візуальна аналітика даних у Power BI. Вісник ХНУ. 2023. № 6. С. 102–110.
- 18.Тарасенко О. О. Технології ETL у системах бізнес-аналітики. Науковий вісник НУБіП. 2021. № 320. С. 77–85.
- 19.Ткачук Н. О., Січко Т. В. Застосування Big Data у бізнесі. Комп'ютерні технології обробки даних. 2023. С. 224–226.
- 20.ACAPS. Ukraine Crisis Analysis. URL: <https://www.acaps.org/en/countries/ukraine> (дата зверненн: 13.05.2026).
- 21.Eurostat. Asylum applicants by citizenship. URL: https://ec.europa.eu/eurostat/databrowser/view/MIGR_ASYTPSM (дата звернення: 13.05.2026).
- 22.Eurostat. Migration and migrant population statistics. URL: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Migration_and_migrant_population_statistics (дата звернення: 13.05.2026).
- 23.Eurostat. Residence permits statistics. URL: <https://ec.europa.eu/eurostat/web/migration-asylum/asylum/database> (дата звернення: 13.05.2026).

- 24.Eurostat. Temporary protection for persons fleeing Ukraine. URL: https://ec.europa.eu/eurostat/databrowser/view/migr_asyappctza/default/table(дата звернення: 13.05.2026).
- 25.Han J., Kamber M., Pei J. Data Mining: Concepts and Techniques. 4th ed. Burlington : Morgan Kaufmann, 2022. 740 p.
- 26.Humanitarian Data Exchange. Ukraine Data Explorer. URL: <https://data.humdata.org/group/ukr>(дата звернення: 13.05.2026).
- 27.Inmon W. H. Building the Data Warehouse. 5th ed. Indianapolis : Wiley, 2021. 576 p.
- 28.International Organization for Migration. Key Migration Terms. URL: <https://www.iom.int/key-migration-terms>(дата звернення: 13.05.2026).
- 29.Kaggle. Ukraine Refugee Data. URL: <https://www.kaggle.com> (дата звернення: 13.05.2026).
- 30.Kimball R., Ross M. The Data Warehouse Toolkit. 3rd ed. Indianapolis : Wiley, 2020. 608 p.
- 31.Microsoft. Data Analysis Expressions (DAX). URL: <https://learn.microsoft.com/en-us/dax/>(дата звернення: 13.05.2026).
- 32.Microsoft. Forecasting in Power BI. URL: <https://community.fabric.microsoft.com/t5/Desktop/Forecasting-in-Power-BI-The-Theory/m-p/365736> (дата звернення: 13.05.2026).
- 33.Microsoft. What is Power BI? URL: <https://learn.microsoft.com/en-us/power-bi/fundamentals/power-bi-overview>(дата звернення: 13.05.2026).
- 34.Migration Partnership Facility. MPF Labour Mobility Partner Country Dashboard. URL: <https://www.migrationpartnershipfacility.eu/labour-mobility/labour-mobility-partner-country-dashboard/about> (дата звернення: 13.05.2026).
- 35.Miro. Business Intelligence Architecture Template. URL: <https://miro.com/templates/business-intelligence-architecture> (дата звернення: 13.05.2026).

36. Python Software Foundation. Python Documentation. URL: <https://docs.python.org/3> (дата звернення: 13.05.2026).
37. Reinsel D., Gantz J. The Digitization of the World: From Edge to Core. Framingham : IDC, 2021. 28 p.
38. Russell S., Norvig P. Artificial Intelligence: A Modern Approach. 4th ed. Pearson, 2021. 1136 p.
39. UNHCR. Global Trends: Forced Displacement in 2023. URL: <https://www.unhcr.org/global-trends-report-2023> (дата звернення: 13.05.2026).
40. UNHCR. Ukraine Refugee Situation. URL: <https://data.unhcr.org/en/situations/ukraine> (дата звернення: 13.05.2026).
41. Січко Т. В. Методи моделювання бізнес-процесів підприємства засобами системного аналізу. Галицький економічний вісник. 2016. № 2(51). С. 190–201. URL: <https://elartu.tntu.edu.ua/handle/123456789/20640>.
42. Нескородева Т.В., Федоров Є.Є., Січко Т.В., Нескородева А.Р. Експертні та рекомендаційні системи: навч. посібник. Вінниця: ДонНУ імені Василя Стуса, 2022, 208 с.

ДОДАТОК А

Код підготовки даних

У додатку наведено фрагменти програмного коду мовою Python, які використовувалися для очищення, трансформації, інтеграції та підготовки статистичних наборів даних до подальшого використання в середовищі Power BI.

```
import pandas as pd

migrants = pd.read_csv("migr_asytpsm__custom_21275923_spreadsheet_1_cleaned.csv")
residence = pd.read_csv("migr_resvalid__custom_3325683_spreadsheet_cleaned.csv")
protection = pd.read_csv("migr_asytpfm__custom_15841088_spreadsheet_cleaned.csv")

for df in [migrants, residence, protection]:
    df.drop_duplicates(inplace=True)
    df.dropna(subset=["Country", "Year", "Value"], inplace=True)
    df["Year"] = pd.to_numeric(df["Year"], errors="coerce")
    df["Value"] = pd.to_numeric(df["Value"], errors="coerce")

migrants.rename(columns={"Value": "Migrants"}, inplace=True)
residence.rename(columns={"Value": "Residence"}, inplace=True)
protection.rename(columns={"Value": "Protection"}, inplace=True)

migrants = migrants.groupby(["Country", "Year", as_index=False])["Migrants"].sum()
residence = residence.groupby(["Country", "Year", as_index=False])["Residence"].sum()
protection = protection.groupby(["Country", "Year", as_index=False])["Protection"].sum()

df = migrants.merge(
    residence,
    on=["Country", "Year"],
    how="left"
).merge(
    protection,
```

```
on=["Country","Year"],  
how="left"  
)
```

```
df.fillna(0, inplace=True)
```

```
df.to_csv(  
    "final_for_powerbi_correlation.csv",  
    index=False,  
    encoding="utf-8-sig"  
)
```



ДОДАТОК Б

Візуалізація розробленої ВІ-системи

У додатку наведено фрагменти візуальної частини дашборду рис. Б.1 та рис Б.2

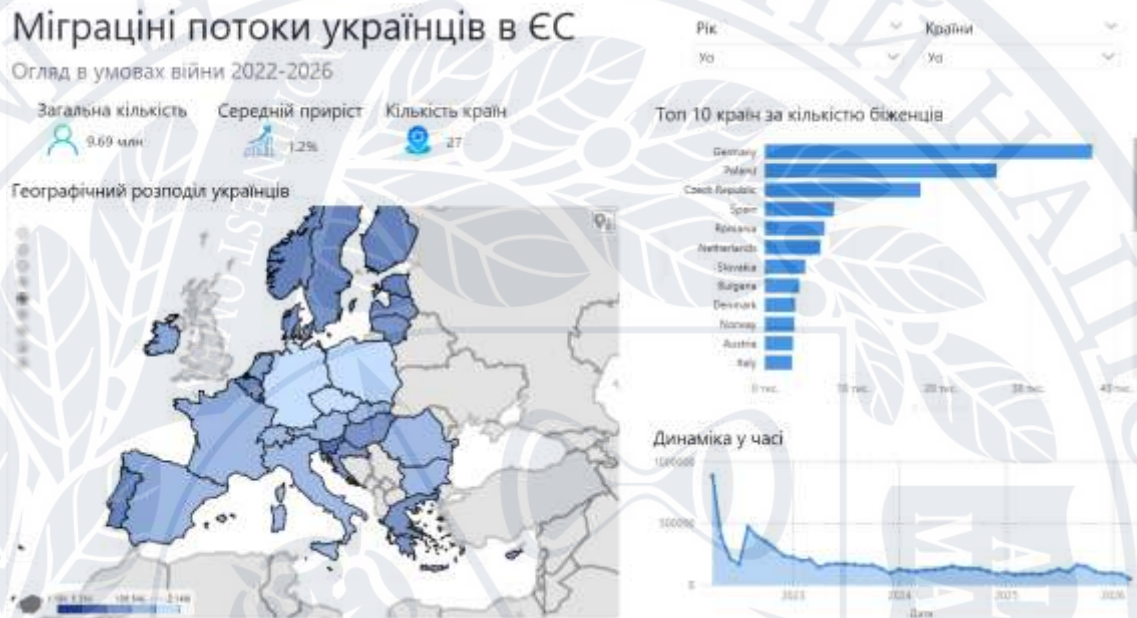


Рисунок Б.1 – Загальний вигляд дашборду частина 1

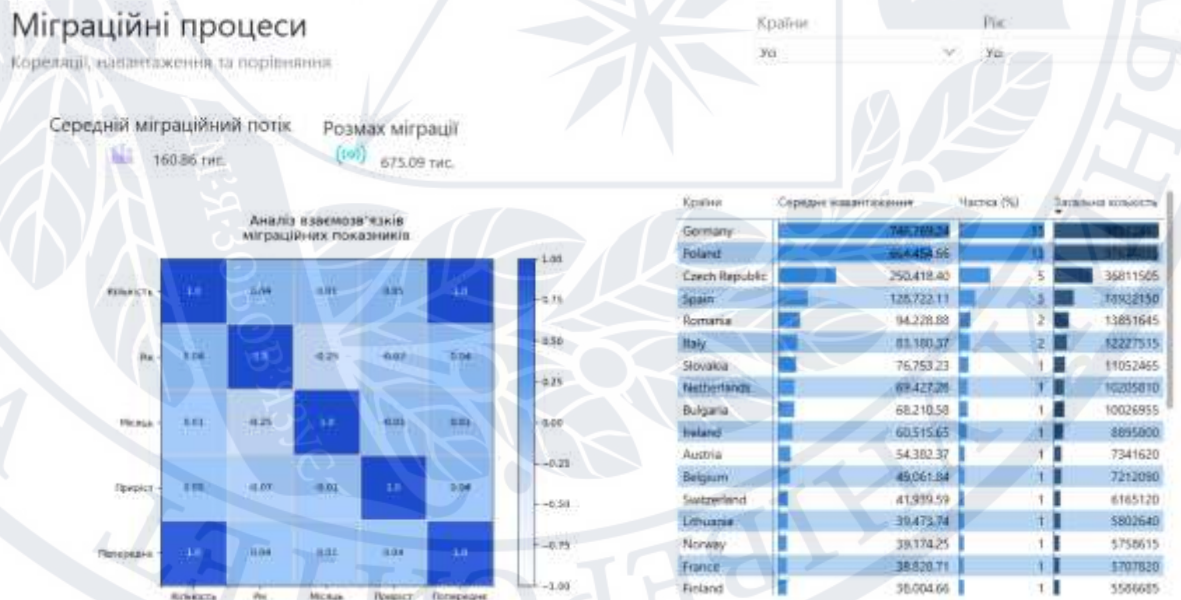


Рисунок Б.2 – Загальний вигляд дашборду частина 2