

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА**

Інститут інформаційних технологій в економіці

Кафедра штучного інтелекту, моделювання та статистики

Освітньо-професійна програма «Економічна кібернетика»

Галузь знань 05 «Соціальні та поведінкові науки»
Спеціальність 051 «Економіка»

Форма здобуття освіти: очна (денна)

КВАЛІФІКАЦІЙНА БАКАЛАВРСЬКА РОБОТА

на тему **«Оцінювання зв'язку між економічним та соціальним розвитком
країн світу»**

(назва теми)

здобувача Онуфрієнко Яни Володимирівни
(ПІБ)

(підпис)

Науковий керівник: доктор фізико-математичних наук,
професор Ольга ПРИТОМАНОВА

(підпис)

**Робота допущена до захисту перед екзаменаційною комісією
з атестації здобувачів вищої освіти (ЕК)**

Завідувач кафедри: кандидат фізико-математичних наук,
професор Галина ВЕЛИКОІВАНЕНКО

(підпис)

Київ 2026

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВАДИМА ГЕТЬМАНА**

**Навчально-науковий інститут
«Інститут інформаційних технологій в економіці»**

Кафедра штучного інтелекту, моделювання та статистики

Освітньо-професійна програма
Галузь знань
Спеціальність

Економічна кібернетика
05 Соціальні та поведінкові науки
051 Економіка

ПОГОДЖЕНО

Керівник проектної групи (гарант)
освітньо-професійної програми
Ольга ПРИТОМАНОВА

(підпис)

_____ 2026 р.

ЗАТВЕРДЖУЮ

Завідувач кафедри математичного
моделювання та статистики
Галина ВЕЛИКОІВАНЕНКО

(підпис)

_____ 2026 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

здобувача вищої освіти

Онуфрієнко Яни Володимирівни

(прізвище, ім'я, по батькові)

_____ очної (денної) _____ **форми здобуття освіти**
очної (денної), заочної, дистанційної

на підготовку кваліфікаційної бакалаврської роботи

на тему

«Оцінювання зв'язку між економічним та соціальним розвитком країн світу»

Тему затверджено наказом ректора Університету від «___» _____ 20__ р. № _____

Кваліфікаційна бакалаврська робота виконується на матеріалах сайту Групи світового банку <https://www.worldbank.org/ext/en/home>, сайту Transparency International <https://www.transparency.org/en/> та інших відкритих джерел.

План кваліфікаційної бакалаврської роботи

Розділ 1	Теоретичні та методичні основи оцінювання соціально-економічного розвитку країн світу
Розділ 2	Статистичний аналіз соціально-економічних показників країн світу
Розділ 3	Багатовимірний аналіз взаємозв'язку соціально-економічних показників країн світу

Об'єкт дослідження:	соціально-економічний розвиток країн світу
Предмет дослідження:	взаємозв'язок між рівнем економічного добробуту, який характеризується показником ВВП на душу населення, та соціально-економічними індикаторами: індексом сприйняття корупції, індексом демократії, індексом злочинності та індексом щастя
Мета виконання кваліфікаційної бакалаврської роботи:	виявлення та кількісне оцінювання взаємозв'язків між ВВП на душу населення та обраними соціально-економічними показниками на основі методів статистичного та багатовимірного аналізу, а також нейронечіткого моделювання

Конкретні завдання, які здобувач повинен виконати для досягнення поставленої мети:

У розділі 1: 1) обґрунтувати вибір показників для проведення дослідження; 2) проаналізувати теоретичні підходи до оцінювання соціально-економічного розвитку країн; 3) розглянути методи статистичного, багатовимірного та нейронечіткого аналізу.

У розділі 2: 1) сформувати інформаційну базу дослідження на основі міжнародних статистичних джерел; 2) провести аналіз таблиць спряженості та парний регресійний аналіз взаємозв'язків між показниками.

У розділі 3: 1) побудувати множинну регресійну модель та виконати кластеризацію країн; 2) застосувати нейронечіткі технології для виявлення можливих нелінійних залежностей між досліджуваними показниками .

Завдання підготував
науковий керівник

(підпис)

Ольга ПРИТОМАНОВА
(ім'я, прізвище)

«__» _____ 20__ р.

Завдання одержав
здобувач

(підпис)

Яна ОНУФРІЄНКО
(ім'я, прізвище)

«__» _____ 20__ р.

Реферат

Кваліфікаційна бакалаврська робота складається зі вступу, трьох розділів, висновків та переліку використаних джерел, що містить 41 найменування на 5 сторінках. Загальний обсяг роботи складає 67 сторінок друкованого тексту, з яких 62 сторінки основної частини. Робота містить 32 рисунки та 11 таблиць.

«Оцінювання зв'язку між економічним та соціальним розвитком країн світу»

Об'єктом дослідження є соціально-економічний розвиток країн світу.

Предметом дослідження є взаємозв'язок між рівнем економічного добробуту, який характеризується показником ВВП на душу населення, та соціально-економічними індикаторами: індексом сприйняття корупції, індексом демократії, індексом злочинності та індексом щастя.

Мета та завдання виконання дослідження. Метою роботи є виявлення та кількісне оцінювання взаємозв'язків між ВВП на душу населення та обраними соціально-економічними показниками на основі методів статистичного та багатовимірного аналізу, а також нейронечіткового моделювання.

Відповідно до поставленої мети було визначено такі завдання:

- обґрунтувати вибір показників для проведення дослідження;
- проаналізувати теоретичні підходи до оцінювання соціально-економічного розвитку країн;
- розглянути методи статистичного, багатовимірного та нейронечіткового аналізу;
- сформувати інформаційну базу дослідження на основі міжнародних статистичних джерел;
- провести аналіз таблиць спряженості та парний регресійний аналіз взаємозв'язків між показниками;
- побудувати множинну регресійну модель та виконати кластеризацію країн;
- застосувати нейронечіткі технології для виявлення можливих нелінійних залежностей між досліджуваними показниками.

Теоретична та методична значущість роботи полягає у комплексному поєднанні методів статистичного аналізу, кластеризації та нейронечіткового моделювання для дослідження соціально-економічного розвитку країн світу. У роботі апробовано різні підходи до аналізу взаємозв'язків між економічними, інституційними та соціальними характеристиками суспільства на єдиній інформаційній базі.

Практична значущість отриманих результатів полягає у можливості використання запропонованого підходу для виявлення факторів, що пов'язані з рівнем економічного добробуту країн, а також для формування аналітичних висновків і підтримки прийняття управлінських рішень у сфері соціально-економічної політики.

Рік виконання кваліфікаційної бакалаврської роботи – 2026.

Рік захисту роботи – 2026.

Ключові слова: ВВП на душу населення, Індекс сприйняття корупції, Індекс демократії, Індекс злочинності, Індекс щастя, таблиці спряженості ознак, регресійний аналіз, кластерний аналіз, нейронечіткі технології.

В і д г у к
про кваліфікаційну бакалаврську роботу
здобувачки навчально-наукового інституту
«Інститут інформаційних технологій в економіці»
освітньо-професійної програми «Економічна кібернетика»

Онуфрієнко Яни Володимирівни

на тему **«Оцінювання зв'язку між економічним та соціальним розвитком
країн світу»**

1. Актуальність теми пов'язана з тим, що в умовах глобальної нестабільності та структурних трансформацій світової економіки питання визначення факторів, які зумовлюють рівень економічного добробуту країн, набуває особливої практичної значущості.

2. Позитивні риси кваліфікаційної бакалаврської роботи: у роботі наведено результати сучасних досліджень щодо взаємозв'язку між економічним добробутом країн та інституційними й соціальними характеристиками суспільства. Для аналізу досліджуваних залежностей використано комплекс статистичних методів, зокрема аналіз таблиць спряженості, кореляційно-регресійний аналіз, методи кластеризації та нейронечітке моделювання, що дозволяє враховувати як лінійні, так і нелінійні зв'язки між показниками.

3. Наявність самостійних розробок автора: вибір теми та виконання кваліфікаційної роботи проведено автором повністю самостійно. Кваліфікаційна робота відповідає затвердженому індивідуальному завданню та повністю оформлена відповідно до вимог до кваліфікаційних робіт.

4. Цінність теоретичних висновків та практичних рекомендацій: полягає в апробації та порівнянні різних підходів до аналізу соціально-економічних процесів на єдиній інформаційній базі, що дасть можливість оцінити їхню ефективність. Практичні результати дослідження можуть бути використані для підтримки прийняття управлінських рішень у сфері державної соціально-економічної політики.

Робота Онуфрієнко Яни Володимирівни пройшла апробацію на 93-й студентській науковій конференції «Інноваційні проєктні рішення на шляху до економічного відродження України», секція «Кількісні методи в економіці» (24 квітня 2026 року, Київ).

5. Загальна оцінка кваліфікаційної бакалаврської роботи та її допущення до захисту перед ЕК: кваліфікаційна бакалаврська робота Онуфрієнко Я.В. виконана на *високому рівні* та свідчить про високий рівень сформованості компетентностей її автора, передбачених освітньою-професійною програмою «Економічна кібернетика» першого (бакалаврського) рівня вищої освіти за спеціальністю «Економіка». Вважаю, що робота може бути допущена до захисту перед ЕК.

Науковий керівник: професор кафедри
штучного інтелекту, моделювання та статистики,
доктор фізико-математичних наук, професор _____ Ольга ПРИТОМАНОВА

«12» червня 2026 р.

ЗМІСТ

ВСТУП.....	3
РОЗДІЛ 1. ТЕОРЕТИЧНІ ТА МЕТОДИЧНІ ОСНОВИ ОЦІНЮВАННЯ СОЦІАЛЬНО-ЕКОНОМІЧНОГО РОЗВИТКУ КРАЇН СВІТУ	5
1.1 Загальна характеристика соціально-економічних показників	5
1.2 Статистичні методи оцінювання зв'язку між показниками	10
1.3 Багатовимірні статистичні методи інтелектуального аналізу даних.	15
1.4 Огляд нейронечітких технологій для ідентифікації залежностей між соціально-економічними показниками	18
РОЗДІЛ 2. СТАТИСТИЧНИЙ АНАЛІЗ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ПОКАЗНИКІВ КРАЇН СВІТУ	22
2.1 Інформаційна база для моделювання.....	22
2.2 Аналіз динаміки досліджуваних показників.....	24
2.3 Аналіз таблиць спряженості соціально-економічних показників	30
2.4 Побудова парних регресійних моделей.....	38
РОЗДІЛ 3. БАГАТОВИМІРНИЙ АНАЛІЗ ВЗАЄМОЗВ'ЯЗКУ СОЦІАЛЬНО- ЕКОНОМІЧНИХ ПОКАЗНИКІВ КРАЇН СВІТУ	44
3.1 Кластерний аналіз	44
3.2 Модель множинної лінійної регресії	48
3.3 Застосування нейронечітких технологій для оцінювання зв'язку між економічним та соціальним розвитком країн.....	53
ВИСНОВКИ	60
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	63

ВСТУП

В умовах глобальної нестабільності та структурних трансформацій світової економіки питання визначення факторів, що зумовлюють рівень економічного добробуту країн, набуває особливої практичної значущості. Традиційний підхід, зосереджений виключно на економічних детермінантах зростання, виявляється недостатнім для повноцінного розуміння розвитку суспільства: дедалі більше досліджень засвідчує, що соціальні явища тісно пов'язані з економічним добробутом.

Об'єктом даного дослідження є соціально-економічний розвиток країн світу, предметом — взаємозв'язок між ВВП на душу населення як індикатором економічного добробуту та чотирма соціально-економічними індексами: Індексом щастя (World Happiness Report), Індексом сприйняття корупції (Transparency International), Індексом демократії (Economist Intelligence Unit) та Індексом злочинності (Numbeo). Дослідження проводиться на даних 113 країн світу за 2024 рік.

Метою роботи є виявлення та кількісне оцінювання характеру зазначених взаємозв'язків із застосуванням методів статистичного аналізу — зокрема аналізу таблиць спряженості з використанням критерію χ^2 -квадрат та мір асоціації — а також нейронечіткого моделювання, яке дозволяє врахувати нелінійний характер досліджуваних залежностей. Комплексне охоплення чотирьох соціальних вимірів одночасно на даних країн світу із застосуванням нейронечітких методів становить наукову новизну роботи.

Для досягнення поставленої мети необхідно виконати такі завдання:

- навести загальну характеристику соціально-економічних показників та обрати ті, що будуть застосовані у роботі;
- навести коротку характеристику методів статистичного аналізу, багатовимірного аналізу, нейронечітких технологій, які будуть

застосовані для виявлення зв'язку між показниками соціально-економічного розвитку країн світу;

- обґрунтувати інформаційну базу дослідження та провести графічний аналіз динаміки обраних показників;
- застосувати таблиці спряженості ознак та парний регресійний аналіз для виявлення можливих зв'язків між обраними показниками;
- провести багатовимірний аналіз можливих взаємозв'язків між обраними показниками;
- застосувати нейронечіткі технології для виявлення можливого нелінійного зв'язку між обраними показниками.

Теоретична, методична та практична значущість отриманих результатів полягає у комплексному дослідженні взаємозв'язків між економічними та соціальними характеристиками країн світу. У роботі наведено результати сучасних досліджень щодо взаємозв'язку між економічним добробутом країн та інституційними й соціальними характеристиками суспільства. Для аналізу досліджуваних залежностей використано комплекс статистичних методів, зокрема аналіз таблиць спряженості, кореляційно-регресійний аналіз, методи кластеризації та нейронечітке моделювання, що дозволяє враховувати як лінійні, так і нелінійні зв'язки між показниками.

Методична значущість роботи полягає в апробації та порівнянні різних підходів до аналізу соціально-економічних процесів на єдиній інформаційній базі, що дасть можливість оцінити їхню ефективність та інформативність при дослідженні факторів економічного добробуту. Практична цінність отриманих результатів полягає у можливості використання запропонованого підходу для виявлення ключових чинників соціально-економічного розвитку країн, формування аналітичних висновків щодо впливу інституційних та соціальних характеристик на рівень добробуту населення, а також для підтримки процесів прийняття управлінських рішень у сфері державної соціально-економічної політики.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ТА МЕТОДИЧНІ ОСНОВИ ОЦІНЮВАННЯ СОЦІАЛЬНО-ЕКОНОМІЧНОГО РОЗВИТКУ КРАЇН СВІТУ

1.1 Загальна характеристика соціально-економічних показників

Питання взаємозв'язку між рівнем економічного розвитку та показниками суспільного добробуту перебуває в центрі уваги сучасних наукових досліджень. В умовах посилення глобальної нестабільності та зростання ролі інституційних чинників актуалізується необхідність комплексного аналізу детермінант економічного зростання з урахуванням не лише суто економічних, але й соціальних вимірів розвитку. Вивчення факторів, що впливають на рівень ВВП на душу населення, дозволяє розробити ефективні стратегії економічної політики та підвищити добробут населення [1]. Водночас сучасна наукова думка дедалі більше наголошує на обмеженості ВВП як єдиного критерію розвитку суспільства і вказує на необхідність його доповнення соціальними індикаторами.

Об'єктом даного дослідження є соціально-економічний розвиток країн світу, предметом — взаємозв'язок між ВВП на душу населення та чотирма соціально-економічними індексами: Індексом щастя, Індексом сприйняття корупції, Індексом демократії та Індексом злочинності. Вибір саме цих індексів зумовлений їх взаємодоповнюваністю: кожен з них відображає окремий вимір соціально-економічного середовища країни — суб'єктивний добробут, інституційну якість, характер політичного управління та рівень соціальної безпеки. Метою роботи є виявлення та оцінювання характеру зазначених взаємозв'язків на вибірці 113 країн світу за 2024 рік із застосуванням методів статистичного аналізу та нейронечіткого моделювання — підходу, який залишається малодослідженим у контексті аналізу соціально-економічних індексів на даних країн світу, що й становить наукову новизну роботи.

Аналіз наявної літератури свідчить, що зв'язок між ВВП на душу населення та зазначеними індексами є підтвердженим емпірично, однак досліджується переважно для окремих пар змінних, в межах конкретних регіонів або з використанням методів, що не дозволяють врахувати нелінійний характер взаємодії між показниками. У праці Й.-Й. Лі та К.-Л. Го «The Happiness-Economic Well-Being Nexus: New Insights From Global Panel Data» [2] підтверджується двостороння залежність між рівнем щастя за Шкалою Кантріла та ВВП на душу населення на панелі 104 країн. Дослідження М. Ф. Абдулджабера та М. Ондера «When we can't see the wood for the trees: the lurking effect of sustainability on corruption» [3] встановлює, що збільшення ВВП на 10 000 доларів США супроводжується покращенням Індексу сприйняття корупції Transparency International на 3,3 бали. У праці А. І. Фаджрянї та М. Шафії «The Influence of GDP per Capita, Income Inequality, Population Density, and Unemployment on the Crime Index in 7 ASEAN Countries» [4] виявлено від'ємний довгостроковий вплив ВВП на душу населення на індекс злочинності Numbeo у країнах АСЕАН — зростання ВВП на 1% зумовлює зниження індексу злочинності на 1,11%. Нарешті, дослідження Качан О., Кмитюк Т. та Даценко Н. «Економетричне моделювання факторів економічного зростання на основі панельних даних» [1] підтверджує значущий позитивний вплив якості державного управління на ВВП на душу населення в різних групах країн, що обґрунтовує включення Індексу демократії до аналізу. Жодне з розглянутих досліджень не охоплює всі чотири зазначені виміри одночасно, що і визначає актуальність даної роботи.

ВВП на душу населення

ВВП на душу населення є одним із найбільш поширених показників для оцінювання та порівняння рівня економічного розвитку країн. ВВП являє собою загальну вартість усіх товарів і послуг, вироблених у країні за рік, і є ключовим економічним індикатором у системі національних рахунків. Хоч валовий внутрішній дохід і дає уявлення про розмір економіки в грошовому вираженні, через різниці в площі та населенні для порівняння країн між собою він не підходить. Замість цього використовується показник ВВП на душу населення, який

розраховується шляхом ділення загального ВВП країни на чисельність її населення і відображає середній обсяг економічного виробництва на одну особу. Даний показник є одним із найбільш поширених індикаторів рівня економічного розвитку та добробуту населення і широко використовується при порівняльних дослідженнях між країнами [1,5].

Використаний у роботі показник ВВП на душу населення сформований Групою світового банку. Група світового банку складається з п'яти організацій, які працюють разом та свою загальну мету визначають як зменшення рівня бідності та сприяння сталому розвитку за допомогою знань, фінансування та досвіду. Установи, що входять до складу Групи:

- Міжнародний банк реконструкції та розвитку;
- Міжнародна асоціація розвитку;
- Міжнародна фінансова корпорація;
- Багатостороння агенція з гарантій інвестицій;
- Міжнародний центр з урегулювання інвестиційних спорів [6].

Група почала своє існування у 1944 році, коли в результаті Бреттон-Вудської фінансової конференції було сформовано Міжнародний валютний фонд та Міжнародний банк реконструкції [7].

Валовий внутрішній продукт — це сукупний дохід, отриманий від виробництва товарів і послуг на економічній території протягом звітного періоду. Його можна виміряти трьома різними способами: за допомогою методів витрат, доходів або виробництва. Для отримання оцінки на душу населення, основний показник поділено на загальну чисельність населення. Цей показник виражено в доларах США у поточних цінах, тобто без урахування змін цін з часом.

Валовий внутрішній продукт (ВВП), валовий національний дохід (ВНД) та інші агреговані показники дають уявлення про розмір і стан економіки, вимірюючи загальну економічну активність у країні. Вони також дозволяють порівнювати економіки різних країн [8].

Індекс щастя

Щорічний Звіт рівня щастя у світі публікується з 2012 року Дослідницьким центром добробуту Оксфордського університету у партнерстві з Gallup, Мережею рішень для сталого розвитку ООН та спеціальною редакційною колегією. Глобальний рейтинг щастя базується на одному запитанні з опитування Gallup World Poll, сформованому на основі Шкали Кантріла: «Будь ласка, уявіть собі драбину зі сходинками, пронумерованими від 0 до 10. Верхівка драбини представляє найкраще можливе для вас життя, а низ драбини – найгірше. На якій сходинці драбини, на вашу думку, ви перебуваєте в цей момент?» [10].

У звіті за 2023 рік було розглянуто взаємозв'язок між рейтингом за шкалою щастя та рівнем ВВП на особу країн. Результати показали, що хоча країни з розвинутою економікою загалом мають вищі бали, присутні й винятки. Декілька країн з ринками, що розвиваються, країн із середнім рівнем доходу, і навіть деякі країни, що розвиваються з низьким рівнем доходу, випереджають своїх колег із розвинутою економікою. Це підкреслює, як фактори поза доходом відіграють вирішальну роль у визначенні рівню задоволеності життям [12].

Індекс сприйняття корупції

Transparency International це антикорупційна організація, заснована 1993 року Пітером Айгеном, колишнім працівником Світового банку. Свою місію компанія наразі визначає так: зупинити корупцію та сприяти прозорості, підзвітності та доброчесності на всіх рівнях і в усіх секторах суспільства. Transparency International є незалежною, неурядовою, некомерційною організацією, яка співпрацює з однодумцями по всьому світу, щоб покласти край несправедливості корупції [13, 14].

Індекс сприйняття корупції (ІСК) – це найпоширеніший у світі глобальний рейтинг корупції. Він вимірює, наскільки корумпованим сприймається державний сектор кожної країни, на думку експертів та бізнесменів. Рейтинг кожної країни є комбінацією щонайменше 3 джерел даних, взятих з 13 різних опитувань та оцінок щодо корупції. Ці дані збираються різноманітними авторитетними установами, включно зі Світовим банком та Всесвітнім економічним форумом. Оцінка країни

за цим індексом – це сприйнятий рівень корупції в державному секторі за шкалою від 0 до 100, де 0 означає високий рівень корумпованості, а 100 – низький [16].

Індекс демократії

Індекс створює заснований у 1946 році дослідницький та аналітичний підрозділ The Economist Group, сестринської компанії газети The Economist. Для задоволення потреб споживачів команда економістів, галузевих спеціалістів, політичних аналітиків та консультантів невпинно працює над створенням найякісніших даних, дослідженнями та аналізом усіх питань, від національних виборів та міжнародної торгівлі до продовольчої безпеки та сталого розвитку міст [18].

Індекс демократії EIU вимірює стан демократії у 167 країнах і територіях, оцінюючи виборчі процеси, управління, політичну участь, політичну культуру та громадянські свободи. Індекс базується на 60 показниках, згрупованих у п'ять категорій:

- виборчий процес та плюралізм;
- функціонування уряду;
- політична участь;
- політична культура;
- громадянські свободи.

Країни оцінюються від нуля до десяти у цих п'яти категоріях, а загальний рейтинг є середнім значенням індексів категорій. На основі загального балу кожену країну розподіляють на чотири типи режимів: «повні демократії», «недосконалі демократії», «гібридні режими» та «авторитарні режими» [19].

Індекс злочинності

Numbeo – це найбільша у світі база даних, що містить дані користувачів про міста та країни світу. Numbeo надає актуальну та своєчасну інформацію про умови життя у світі, в тому числі вартість життя, показники житла, охорону здоров'я, дорожній рух, злочинність та забруднення [22].

Місія Numbeo, відповідно до вебсайту, полягає в тому, щоб надати окремим особам та підприємствам точну та актуальну інформацію про вартість та якість життя, різні соціально-економічні фактори в містах та країнах світу. Заснований Младеном Адамовичем у 2009 році, сайт прагне бути провідним постачальником краудсорсингових даних, що дозволить користувачу приймати обґрунтовані рішення, сприятиме прозорості та глобальному розумінню [23].

Процес збору даних Numbeo включає поєднання даних, згенерованих користувачами, та інформації, зібраної вручну з авторитетних джерел, таких як вебсайти супермаркетів, веб-сайти служб таксі та державних установ. Зібрані дані обробляються двічі на рік і для підвищення надійності їм присвоюється вага, втричі вища, ніж даним, що були згенеровані користувачами. Для забезпечення точності даних використовуються алгоритми фільтрації спаму, які виключають дані, отримані від користувачів із підозрілою поведінкою [24].

Індекс злочинності оцінює загальний рівень злочинності в місті чи країні. Він класифікується таким чином:

- Дуже низький: 20 і нижче;
- Низький: 20.01-40;
- Помірний: 40.01-60;
- Високий: 60.01-80;
- Дуже високий: вище 80 [26].

1.2 Статистичні методи оцінювання зв'язку між показниками

Шкали вимірювання статистичних показників (якісні та кількісні)

Першим етапом будь-якого статистичного дослідження є вимірювання досліджуваних показників, тобто присвоєння їм певних значень. В залежності від того, які дані були сформовані та які математичні операції доречні з ними, можна

виділити чотири основних типи шкал вимірювання: номінальну, порядкову, інтервальну та шкалу відношень.

Номінальна шкала (або найменувань, категоріальна) є найнижчим рівнем вимірювання, оскільки не виконує функції ранжування елементів. Ця шкала розподіляє одиниці даних на групи на основі їх подібності між собою. Будь-якого особливого порядку чи визначення зв'язку між групами або елементами не відбувається. Ця шкала підходить для даних нечислового характеру, або таких, де числа не мають аналітичного значення, наприклад: стать, національність, вік, так/ні тощо.

Порядкова шкала (або ординальна, рангова) на відміну від номінальної виконує функцію ранжування, тобто її елементи розташовуються за зростанням чи спаданням в залежності від вираженості їх властивостей або ознак. Відстань між рангами не є чітко визначеною. Ця шкала утворює чітку ієрархію спостережень, а її прикладом є рівень задоволеності клієнта (від "дуже задоволений" до "дуже незадоволений"), шкала оцінювання знання учнів та інші.

Інтервальна шкала, на відміну від порядкової, має чітко визначені ранги, тобто інтервали між сусідніми значеннями тут однакові. Це означає, що дана шкала дозволяє визначити, наскільки одне значення більше чи менше за інше. Однак важливо зауважити, що нульове значення у цій шкалі є умовним - наприклад, як у температурній шкалі Цельсія чи Фаренгейта.

Шкали відношень є найбільш досконалими серед розглянутих та мають природний абсолютний нуль, який вказує на повну відсутність вимірювальної властивості у спостереження. Завдяки цій особливості, над такими даними можливо здійснювати будь-які математичні операції. Тобто, для двох різних величин, виміряних з допомогою шкали відношень, можна визначити різницю величини. [29].

Методи оцінювання зв'язку між якісними та кількісними показниками

Безперечно, першим кроком вибору методу статистичного аналізу є визначення шкали, у якій вимірюються обрані дані. Для оцінювання залежності між двома кількісними змінними класично використовують кореляційний аналіз

(коефіцієнти кореляції, наприклад Спірмена або Пірсона). Коли у дослідженні наявні якісні змінні, необхідні інші підходи.

У статистичному аналізі вибір методу для оцінювання зв'язку між змінними залежить насамперед від типів шкал, у яких вимірюються ці змінні. Для встановлення зв'язку між двома кількісними змінними традиційно застосовується кореляційний аналіз (наприклад, коефіцієнт кореляції Пірсона або Спірмена). Однак коли одна або обидві змінні є якісними (категоріальними), необхідні інші підходи.

Категоріальні дані можуть бути отримані безпосередньо (наприклад, поділ показника політичного устрою країни на типи: «повна демократія», «недосконала демократія» тощо), або перетворені з числових. Цей процес називається дискретизацією — означає поділ числової змінної на визначені класи. Прикладом може бути поділ країн світу на групи в залежності від їх рівня ВВП на душу населення

Одним із найпоширеніших методів дослідження взаємозв'язку між категоріальними ознаками є аналіз таблиць спряженості. Для дослідження залежності використовують критерій χ^2 та похідні від нього міри асоціації [30].

Аналіз таблиць спряженості ознак

Таблицями спряженості називаються спеціальні таблиці, які використовуються з метою дослідження взаємозв'язку між двома чи більшою кількістю змінних. Розрізняють двовходові (для розгляду зв'язку між двома змінними) та багатовходові (для більш ніж двох змінних) таблиці.

Двовходові таблиця спряженості ознак використовуються для аналізу зв'язку між категоріями двох змінних. У загальному вигляді вигляд таблиці такий: розмір $I \times J$, де I означає кількість категорій першої змінної та кількість рядків у таблиці, а J означає категорії другої змінної, тобто рядки. Кожна клітинка таблиці містить спостереження n_{ij} , відповідно до рядка та стовпця, у якому вона знаходиться. [31].

Багатовходові таблиці мають на меті те саме, що й двовходові, тобто дослідити взаємодію між змінними, але на відміну від останніх розглядають

зв'язок трьох чи більше змінних [29]. Для візуалізації результатів аналізу цих таблиць спряженості використовують декілька типів графіків:

- Гістограма для порівняння кількості спостережень між комбінаціями змінних. Якщо взяти графік з накопиченням, можна додатково визначити та порівняти частки у групах;
- Для візуалізації зв'язку числових змінних можна взяти діаграму розсіювання. Два показники будуть відображатися осями x та y , а колір точок (спостережень) позначатиме третю характеристику;
- Лінійна діаграма може використовуватись для порівняння кількості спостережень у групах;
- Теплові карти використовують для таблиць з більш ніж трьома вхідними змінними. Вони відображають значення кожної клітинки таблиці у вигляді кольорових прямокутників, де колір змінюється залежно від значення показника [29].

В кінці кожного рядка та стовпця таблиці обраховуються суми:

Значення поза основними клітинками таблиці називаються маргінальними (marginals): n_{i+} позначає суму по i -му рядку, а n_{+j} — суму по j -му стовпцю (символ $+$ вказує на те, елемент таблиці i або j сумується) [31]. На основі цих маргінальних сум розраховуються очікувані частоти — значення, які спостерігалися б у кожній клітинці за умови статистичної незалежності двох змінних. Очікувана частота для клітинки (i, j) визначається за формулою [31]:

$$E_{ij} = \frac{n_{i+} \times n_{+j}}{n_{++}}, \quad (1.1)$$

де n_{++} — загальний обсяг вибірки.

Порівняння спостережуваних частот O_{ij} з очікуваними E_{ij} є основою критерію χ^2 .

Критерій χ^2 (Chi-square) для оцінювання незалежності ознак

Критерій χ^2 (критерій хі-квадрат Пірсона) є непараметричним статистичним тестом, що використовується для перевірки нульової гіпотези H_0 про відсутність зв'язку між двома категоріальними змінними [30]. Він не вимагає нормального

розподілу даних і однорідності дисперсій, що робить його особливо придатним для аналізу індексів і рейтингових показників.

Статистика хі-квадрат обчислюється за формулою:

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \quad (1.2)$$

де O_{ij} — спостережувана частота в клітинці (i, j) , а E_{ij} — очікувана частота.

Розмір відхилення між спостережуваними та очікуваними частотами визначає розмір значення χ^2 . Для прийняття або відхилення нульової гіпотези розраховане значення χ^2 порівнюється з критичним значенням розподілу хі-квадрат. Для цього необхідні заданий рівні значущості α та розраховані за наступною формулою ступені свободи:

$$df = (\text{Кількість рядків} - 1)(\text{Кількість стовпців} - 1) \quad (1.3)$$

У випадку, коли розраховане значення перевищує критичне, нульова гіпотеза H_0 про відсутність залежності відхиляється на користь альтернативної гіпотези H_1 , яка свідчить про наявність зв'язку між змінними.

Не слід забувати про те, що критерій χ^2 вказує лише на значущість, а не на силу зв'язку між показниками. Великий хі-квадрат вказує лише на те, що залежність існує, а не на те, що вона є сильною. Крім того, значення χ^2 залежить від обсягу вибірки: при великій вибірці навіть слабкий зв'язок може дати статистично значущий результат [33]. Саме тому разом з цим критерієм необхідно застосовувати методи перевірки сили зв'язку.

Коефіцієнт спряженості C та міра Крамера V

Для вимірювання сили статистичного зв'язку між двома категоріальними змінними після отримання значущого результату χ^2 -тесту застосовуються похідні міри асоціації.

Коефіцієнт спряженості C (Pearson's contingency coefficient) — також запропонований Пірсоном показник, що може застосовуватися для таблиць будь-якого розміру $I \times J$ [31]. Він розраховується за формулою:

$$C = \sqrt{\frac{\chi^2}{\chi^2 + n}}, \quad (1.4)$$

де n — це загальний обсяг вибірки.

Оскільки значення n не може бути нульовим, розрахований результат C завжди залишається в діапазоні $0 \leq C < 1$ [33]. Важливо зауважити, що максимально досяжне значення коефіцієнта спряженості не є 1, а визначається розміром таблиці і обчислюється за формулою:

$$C_{max} = \sqrt{\frac{k-1}{k}}, \quad (1.5)$$

де k — кількість рядків або стовпців таблиці, залежно від того яке значення менше.

Отже, для таблиці 2×2 , наприклад, максимальне значення C буде лише близько 0,71, навіть за умови абсолютного зв'язку між розглянутими змінними. З цієї причини коефіцієнти спряженості розраховані для різних таблиць можна порівнювати тільки у випадку, якщо розмірність таблиць однакова [33].

Коефіцієнт Крамера V (Cramer's V) є придатним для використання з таблицями будь-якого розміру. Цей показник розраховується за формулою [30]:

$$V = \sqrt{\frac{\chi^2}{n(k-1)}}, \quad (1.6)$$

де k — менше значення з кількості рядків або стовпців таблиці.

Цей коефіцієнт вимірює силу зв'язку та є більш універсальним у застосуванні через відсутність обмежень на розмірність таблиць. Значення може змінюватися від 0 до 1 включно, де 0 означає відсутність залежності, а 1 — ідеальний зв'язок між змінними.

1.3 Багатовимірні статистичні методи інтелектуального аналізу даних

Кластерний аналіз

Кластерний аналіз є формальним дослідженням методів та алгоритмів групування об'єктів відповідно до їхніх виміряних або сприйнятих внутрішніх характеристик чи подібності. На відміну від класифікації, кластерний аналіз не

використовує категоріальні мітки, яка б заздалегідь визначала належність об'єктів до груп, — саме ця відсутність попередньо відомих категорій відрізняє кластеризацію від дискримінантного аналізу. Метою кластерного аналізу є виявлення структури в даних, тому за своєю природою він є дослідницьким інструментом [40].

Кластеризація має тривалу і широку міждисциплінарну історію: її методи розробляли та застосовували таксономісти, соціологи, психологи, біологи, статистики, математики, інженери та медичні дослідники. Методи кластеризації поділяються на дві основні групи: ієрархічні та ітеративні. Ієрархічні алгоритми рекурсивно об'єднують окремі об'єкти у кластери, тоді як ітеративні (неієрархічні) алгоритми визначають усі кластери одночасно через задавання їх числа заздалегідь. Найпростішим і найпопулярнішим серед ітеративних алгоритмів є метод k -середніх [40].

Метод k -середніх (k -means) вперше опубліковано у 1955 році, а його незалежне відкриття у різних наукових галузях здійснили Штайнгауз (1956), Ллойд (запропоновано у 1957 р., опубліковано у 1982 р.) та МакКвін (1967). Незважаючи давнину його відкриття, цей алгоритм залишається одним із найбільш широко вживаних — завдяки простоті реалізації, ефективності та підтвердженій практичній результативності [40].

Формально метод k -середніх вирішує задачу розбиття множини з n об'єктів на k кластерів таким чином, щоб мінімізувати суму квадратів відстаней між кожним об'єктом та центроїдом його кластера. Процес починається з вибору кількості k кластерів, після чого на кожній ітерації кожен об'єкт призначається до найближчого до нього центра кластера, а центри перераховуються як середні значення всіх об'єктів, що потрапили до відповідного кластера. Цей процес повторюється до стабілізації належності об'єктів до кластерів [40].

Алгоритм потребує від користувача задання трьох параметрів: кількості кластерів k , способу початкової ініціалізації та метрики відстані. Найбільш критичним є вибір k , оскільки жодного універсального математичного критерію для цього не існує — на практиці алгоритм запускається кілька разів з різними

значеннями k , а найбільш змістовне розбиття обирається з урахуванням предметної області [40].

Регресійний аналіз

Парний та множинний регресійний аналіз є одними з найбільш поширених методів економіко-статистичного моделювання, які використовуються для дослідження взаємозв'язків між соціально-економічними показниками та оцінювання впливу факторів на результативну ознаку. Регресійний аналіз дозволяє не лише встановити напрямок зв'язку між змінними, але й кількісно оцінити силу такого впливу [4].

Парна лінійна регресія використовується у випадках, коли необхідно дослідити залежність однієї результативної змінної від одного пояснювального фактора. Загальний вигляд моделі парної лінійної регресії [4]:

$$Y = \beta_0 + \beta_1 X + \varepsilon, \quad (1.7)$$

де Y — залежна змінна; X — незалежна змінна; β_0 — вільний член моделі;

β_1 — коефіцієнт регресії, який показує зміну результативної ознаки при зміні фактора на одну одиницю; ε — випадкова похибка моделі.

Основною перевагою парної регресії є простота інтерпретації результатів. Вона дозволяє оцінити характер зв'язку між двома показниками та визначити, наскільки зміни факторної ознаки пов'язані зі змінами результативної змінної. Для оцінювання якості моделі використовуються коефіцієнт детермінації (R^2), F -критерій та статистична значущість параметрів регресії [4].

Проте в реальних соціально-економічних процесах результативний показник зазвичай формується під впливом не одного, а кількох факторів одночасно. У таких випадках застосовується множинна лінійна регресія, яка дозволяє враховувати комплексний вплив декількох незалежних змінних на результативну ознаку [41].

Загальний вигляд множинної лінійної регресії [41]:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon, \quad (1.8)$$

де $X_1, X_2, \dots, X_k$ — факторні ознаки, а коефіцієнти $\beta_1, \beta_2, \dots, \beta_k$ характеризують вплив кожного фактора за умови фіксованих значень інших змінних.

Множинна регресія дозволяє оцінити відносний внесок окремих факторів у формування результативного показника та визначити їх статистичну значущість. Однією з ключових переваг такого підходу є можливість дослідження складних багатофакторних процесів, що особливо важливо під час аналізу економічного розвитку країн, оскільки рівень добробуту населення формується під впливом політичних, інституційних, соціальних та економічних чинників одночасно [41].

При побудові множинних регресійних моделей важливо враховувати можливу мультиколінеарність між незалежними змінними, оскільки висока взаємна кореляція факторів може спотворювати оцінки параметрів моделі та ускладнювати інтерпретацію результатів. Тому після побудови моделі доцільно додатково здійснювати перевірку її статистичних характеристик та якості апроксимації даних. Зокрема, оцінюється статистична значущість моделі за F -критерієм Фішера та окремих коефіцієнтів за t -критерієм Стюдента, аналізується коефіцієнт детермінації, перевіряється наявність мультиколінеарності за допомогою показника VIF, а також досліджуються залишки моделі на предмет виконання основних припущень лінійної регресії. Такі перевірки дозволяють оцінити адекватність побудованої моделі та надійність отриманих результатів [41].

1.4 Огляд нейронечітких технологій для ідентифікації залежностей між соціально-економічними показниками

Якщо між змінними утворюються складні та нелінійні взаємозв'язки, докладно дослідити їх такими традиційними статистичними методами як кореляційний аналіз чи критерій χ^2 стає неможливо. Можливості цих підходів дозволяють виявити зв'язок, проте подальший аналіз значно ускладнюється в умовах можливої невизначеності, неточності чи неповноти використовуваних даних. Аби впоратися з цією перешкодою можна звернутися до застосування нейронечітких технологій.

Нечітка логіка та нечіткі множини

Розділ нечіткої логіки, що стоїть в основі нейронечітких технологій, був уперше виокремлений Лотфі Заде у 1965 році. У цій системі кожен елемент множини має ступінь належності до неї, оцінений від 0 до 1 [34].

Такий підхід дозволяє дослідникам чітко окреслити та формалізувати абстрактні поняття, такі як, наприклад, «розвинена демократія», «низький рівень злочинності» та інші категорії, які можуть використовуватись при описі соціально-економічних показників розвитку. У фінансово-економічних системах, які найчастіше характеризуються високою невизначеністю, застосування нечіткої логіки є більш ніж доречним та логічним.

Основним поняттям нечіткого моделювання є функція належності, яка визначає ступінь, з яким досліджувані значення відповідають визначеним категоріям або лінгвістичним термам (наприклад таким як «низький», «середній», «високий»). З допомогою функцій належності формують нечітку базу знань, наповнену нечіткими правилами сформульованими так: «ЯКЩО ... ТО ...». Перетворення нечіткого значення на нечітке називають фазифікацією, а зворотне перетворення — дефазифікацією.

Ключовим поняттям нечіткого моделювання є функція належності — функція, що визначає ступінь, з яким кожне значення змінної відповідає певному лінгвістичному терму (наприклад, «низький», «середній», «високий»). На базі функцій належності формуються нечіткі правила виведення у форматі «ЯКЩО ... ТО ...», які разом утворюють нечітку базу знань. Процес перетворення чіткого числового значення на нечітке називається фазифікацією, а зворотне перетворення — дефазифікацією [35].

Моделі нечіткого виведення Мамдані та Сугено

Одним із найпоширеніших підходів до побудови нечітких систем є алгоритм Мамдані. У ньому як антецедент (вхід), так і консеквент (вихід) нечіткого правила є нечіткими множинами. Головна перевага цього типу системи — висока інтерпретабельність: правила набувають вигляду, зрозумілого для дослідника, а параметри легко пояснити змістовно. Оскільки бази правил у системах Мамдані

більш інтуїтивно зрозумілі, вони добре підходять для застосування в задачах, де правила було попередньо сформовано експертом [36]. Для задач аналізу соціально-економічних залежностей між країнами, де важливо не лише отримати результат, але й пояснити його зміст, модель Мамдані є доцільним вибором.

Модель Сугено відрізняється тим, що на виході кожного правила отримується не нечітка множина, а константа або лінійна функція виду:

$$z_i = a_i x + b_i y + c_i, \quad (1.9)$$

де x та y — вхідні значення даних, a_i , b_i та c_i — коефіцієнти виражені константами.

Кінцевий вихід системи обчислюється як зважене середнє виходів усіх правил, що робить дефазифікацію значно простішою й обчислювально ефективнішою порівняно з моделлю Мамдані [36]. Завдяки лінійній залежності кожного правила від вхідних змінних модель Сугено добре підходить для задач оптимізації та адаптивних технік, зокрема є основою для побудови системи ANFIS [36].

Нейронні мережі та нейронечіткі системи

Штучні нейронні мережі — це специфічні математичні моделі, збудовані з вузлів-нейронів здатних до взаємодії та навчання на вхідних даних. Ці моделі здатні автоматично виявляти закономірності та навчатися через процес тренування, проте їхнє функціонування відбувається за принципом "чорної скриньки", тобто при прийнятті рішень вони не надають чіткої логіки або алгоритму, який був використаний в цьому процесі. З цієї причини відбувається поєднання нейронних мереж та нечіткої логіки.

Нейронечіткий підхід дозволяє поєднати переваги обох його складових — механізми навчання нейронних мереж та зрозумілість нечітких правил. В результаті застосування такої системи виходить отримати структурований та інтерпретабельний результат, де параметри моделі (функції належності та коефіцієнти) були налаштовані нейронною мережею на основі вхідних даних. Таким чином можна будувати зрозумілі для дослідника моделі з функцією автоматичного навчання [37].

Система ANFIS

ANFIS (Adaptive Neuro-Fuzzy Inference System — адаптивна нейронечітка система виводу) — це архітектура, яка поєднує нечіткий висновок типу Сугено з навчанням нейронної мережі [38]. Модель Сугено лежить в основі ANFIS з тої причини, що її лінійна структура консеквентів правил добре поєднується з методами оптимізації, а обчислювальна ефективність дозволяє добре навчати параметри системи на даних [36].

В процесі навчання ANFIS модель автоматично змінює свої параметри для досягнення найбільш точного зв'язку між вхідними та вихідними даними. Один етап навчання відповідає за налаштування математичних параметрів нечітких правил, а інший — за коригування функцій належності. Це дозволяє системі адаптуватися до даних і водночас враховувати експертні знання, закладені у нечітких правилах [38].

Структурно ANFIS реалізується як п'ятишарова нейронна мережа, де кожен шар виконує певний крок нечіткого виведення типу Сугено:

- перший шар — фазифікація;
- другий — обчислення сили спрацювання кожного правила;
- третій — нормалізація сил спрацювання;
- четвертий — зважений вихід кожного правила;
- п'ятий — підсумовування у кінцевий результат [38].

Варто також відзначити, що нечітка логіка дозволяє будувати моделі нелінійних функцій довільної складності та спиратися на досвід експертів — на відміну від звичайних нейронних мереж, які генерують непрозорі, важко інтерпретовані моделі [35].

РОЗДІЛ 2

СТАТИСТИЧНИЙ АНАЛІЗ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ПОКАЗНИКІВ КРАЇН СВІТУ

2.1 Інформаційна база для моделювання

Для проведення статистичного аналізу було сформовано набір даних країн світу за 2024 рік. До датасету включено ряд соціально-економічних показників:

- ВВП на душу населення (GDP per capita) — дані Групи Світового банку;
- Індекс щастя (World Happiness Index) — дані World Happiness Report;
- Індекс сприйняття корупції (Corruption Perceptions Index) — дані Transparency International;
- Індекс демократії (Democracy Index) — дані Economist Intelligence Unit;
- Індекс злочинності (Crime Index) — дані Numbeo.

Обробка та статистичний аналіз даних виконувалися у середовищі RStudio з використанням широкого функціоналу мови програмування R.

На першому етапі до програмного середовища було імпортовано зібрані з відкритих джерел дані та виокремлено лише необхідні стовпці з первісних наборів [8, 9, 15, 21, 25]. Було проведено уніфікацію назв змінних з метою подальшого формування єдиної таблиці. Назви показників було подано у скороченому вигляді: GDPpc, WorldHapp, CorrPrecep, DemocInd та CrimeInd.

Через різні джерела отримання даних виникла проблема з різним написанням назв країн, наприклад «Turkey» та «Turkiye», «Congo» та «Republic of the Congo», «Korea, Rep.» та «South Korea» тощо. Було проведено ручну стандартизацію назв країн з використанням ISO кодів (вони були наведені не у всіх

первинних датасетах). Також із набору даних було виключено агреговані регіональні показники та території, які не були представлені в усіх джерелах даних.

Таблиці об'єднувалися поступово, за назвами та кодами країн, аби уникнути помилок та розбіжностей. В результаті було сформовано єдиний датафрейм, який містив такі змінні:

- назву країни;
- ISO-код країни;
- ВВП на душу населення;
- Індекс сприйняття корупції;
- Індекс демократії;
- Індекс злочинності;
- Індекс щастя.

На завершальному етапі підготовки даних було видалено спостереження з пропущеними значеннями за допомогою функції `na.omit()`. Остаточний набір даних містив 113 країн світу.

Описова статистика сформованого набору даних показала значну варіативність між країнами за рівнем соціально-економічного розвитку. Статистику відображено у табл. 2.1:

Таблиця 2.1 – Описова статистика використаних у дослідженні показників

	ВВП на душу населення	Індекс сприйняття корупції	Індекс демократії	Індекс злочинності	Індекс щастя
Мінімум	656,8	10	1	15,6	3,3
Максимум	137781,7	90	9,8	81,2	7,7
Середнє	23358,6	46,6	6	44,7	5,8

Джерело: розроблено авторкою на основі [8, 9, 15, 21, 25]

Зокрема, мінімальне значення ВВП на душу населення становило 656,8 дол. США, тоді як максимальне — 137781,7 дол. США. Середнє значення показника склало 23358,6 дол. США. Індекс сприйняття корупції змінювався в межах від 10 до 90 балів, індекс демократії — від 1 до 9,8, індекс злочинності — від 15,6 до 81,2, а індекс щастя — від 3,3 до 7,7 бала.

Сформований набір даних став основою для подальшого статистичного аналізу теми.

2.2 Аналіз динаміки досліджуваних показників

З метою оцінки тенденцій зміни досліджуваних показників проведено динамічний аналіз їх панельних даних. Через неповноту даних за різні роки кількість країн взята для панельних даних показників не однакова:

- у ВВП на душу та Індексу демократії залишилося 113 країн, як і в основній вибірці за 2024 рік;
- Індекс злочинності налічує всього 70 країн (за більш ранні роки кількість країн у вибірці значно менша, ніж пізніше), проте додані також спостереження за 2025 рік;
- Індекс сприйняття корупції має 112 країн;
- Індекс щастя налічує 110 країн, також дані у вільному доступі є тільки за 2020-2024 рр.

На зібраних та оброблених панельних даних було створено п'ять візуалізацій (рис. 2.1-2.5). Окремі графіки для кожного показника відображають три криві, які відповідають мінімальним, максимальним та середнім значенням спостережень за охоплені роки. Такий підхід дозволяє ознайомитися з загальним напрямом змін у даних.

Аналіз динаміки ВВП на душу населення свідчить про наявність загальної тенденції до зростання як максимальних, так і середніх значень показника протягом 2012–2024 років. Максимальні значення ВВП на душу населення у досліджуваній вибірці коливалися приблизно в межах 105–137 тис. дол. США, тоді як середні значення перебували на рівні близько 18–24 тис. дол. США. Такий значний розрив між максимальними та середніми значеннями вказує на суттєву нерівномірність

економічного розвитку між країнами вибірки та наявність групи високорозвинених країн із дуже високими доходами населення.

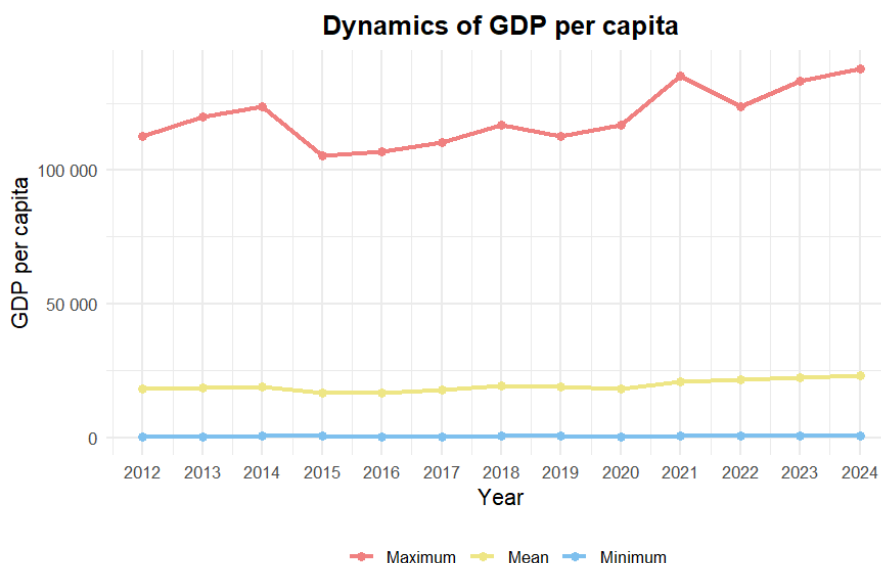


Рисунок 2.1 – Динаміка ВВП на душу населення в період 2012-2024 рр.

Джерело: розроблено авторкою

У 2015 році спостерігається помітне зниження максимального значення показника — приблизно зі 124 тис. до 105 тис. дол. США. Імовірною причиною такого спаду могли бути наслідки уповільнення світової економіки, падіння світових цін на нафту та інші сировинні ресурси та нестабільність фінансових ринків у цей період. Після цього динаміка поступово стабілізується та переходить до зростання.

Особливо помітне підвищення максимального значення ВВП на душу населення спостерігається у 2021 році, коли показник перевищив приблизно 130 тис. дол. США. Можна припустити, що тут проявляються економічне відновлення після пандемії COVID-19, відповідна активізація міжнародної торгівлі, збільшення обсягів інвестицій та масштабні стимулюючі заходи урядів і центральних банків. Після незначного зниження у 2022 році показник знову демонструє висхідну тенденцію у 2023–2024 роках.

Середнє значення ВВП на душу населення не має таких явних коливань, як максимальне, проте також поступово зростає майже протягом усього досліджуваного періоду. Мінімальні значення показника залишаються дуже

низькими і протягом усього періоду майже не змінюються. Видно, що для країн із низьким рівнем доходів населення характерні низькі темпи економічного зростання та наявність структурних економічних проблем.

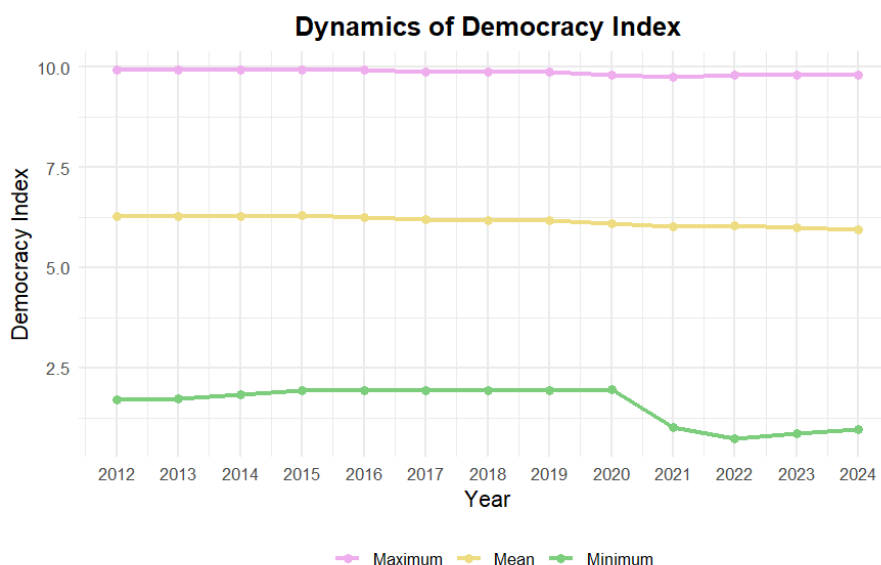


Рисунок 2.2 – Динаміка Індексу демократії в період 2012-2024 рр.

Джерело: розроблено авторкою

Максимальні значення індексу демократії протягом усього досліджуваного періоду залишалися близькими до 10 балів, тобто на рівні повноцінних демократій. Це вказує на те, що країни вибірки з усталеним демократичним політичним устроєм здебільшого зберігають високі стандарти демократичного розвитку, ефективність політичних інститутів та високий рівень громадянських свобод незалежно від факторів впливу.

Середні значення індексу демократії демонструють поступову слабку тенденцію до зниження: приблизно з 6,3 бала у 2012 році до близько 6 балів у 2024 році. Хоча зміни є незначними, така динаміка може свідчити про поступове послаблення демократичних процесів у частині країн вибірки, посилення авторитарних тенденцій та зниження рівня довіри до державних інститутів.

Найбільш помітні зміни спостерігаються у мінімальних значеннях індексу демократії. У 2021 році відбулося різке зниження показника — приблизно з 2 до 1 бала, після чого надалі він залишався на такому ж низькому рівні. Імовірно причиною такого погіршення могли стати наслідки пандемії COVID-19, під час

якої у багатьох країнах були запроваджені надзвичайні обмеження та відповідне посилення контролю з боку держави, обмеження громадянських свобод та погіршення політичної стабільності. Сильний негативний вплив скоріш за все також мали численні збройні конфлікти, політичні кризи у багатьох державах світу, як от початок повномасштабного вторгнення росії в Україну у 2022 році.

Загалом можна зробити висновок, що незважаючи на збереження стійкого демократичного устрою у частині держав, у світі спостерігається тенденція до певного погіршення середнього рівня демократії та посилення диференціації між країнами за рівнем демократичного розвитку.

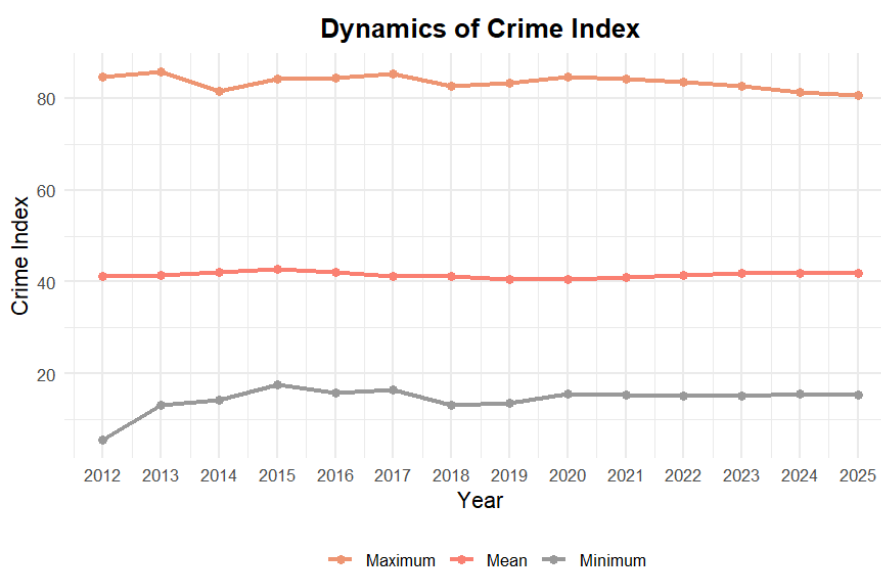


Рисунок 2.3 – Динаміка Індексу злочинності в період 2012-2025 рр.

Джерело: розроблено авторкою

Максимальні значення індексу злочинності у досліджуваній вибірці коливалися переважно в межах 81–86 балів, що вказує на збереження у вибірці частки країн із дуже високим рівнем злочинності протягом усього періоду дослідження.

У 2014 році спостерігається деяке зниження максимального значення індексу злочинності — приблизно з 86 до 81 бала. Причинами такого зниження могли бути зміни у методології оцінювання, посилення заходів безпеки у окремих країнах або тимчасове зниження рівня окремих видів злочинності, однак таке покращення все одно було не тривалим. Надалі показник знову зріс і у 2015–2020

роках перебував на рівні близько 84–85 балів. Після 2021 року спостерігається поступове слабке зниження максимальних значень.

Середні значення індексу злочинності залишалися відносно стабільними — на рівні приблизно 40–42 балів. Різких глобальних змін у середньому рівні злочинності серед країн вибірки не спостерігається. Мінімальні значення індексу злочинності змінювалися більш помітно. У 2012 році мінімальний показник становив близько 6 балів, після чого вже у 2013–2015 роках зріс до приблизно 13–17 балів. Причиною такого різкого зростання цілком може бути зміна у методології формування показника, адже 2012 рік — це перший рік оприлюднення Індексу злочинності. Однак не слід відкидати можливості, що після 2012 року країни з найнижчим рівнем злочинності почали стикатися зі зростанням окремих кримінальних ризиків, наприклад кіберзлочинності або фінансових злочинів. Після 2020 року мінімальні значення індексу стабілізувалися на рівні близько 15–16 балів.

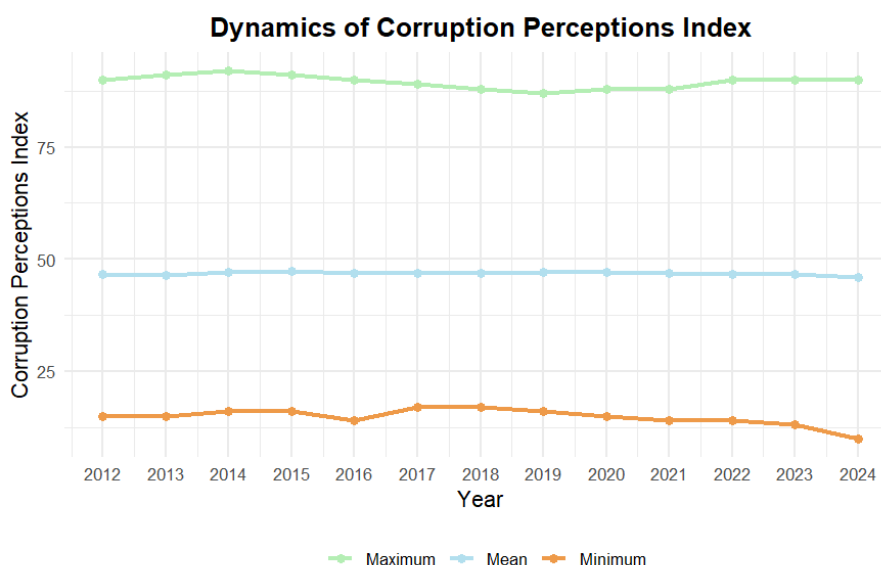


Рисунок 2.4 – Динаміка Індексу сприйняття корупції в період 2012-2024 рр.

Джерело: розроблено авторкою

Максимальні значення індексу коливаються в межах приблизно 88–91 балів, ця оцінка відповідає країнам із дуже низьким рівнем сприйняття корупції та високою ефективністю державних інституцій. Після незначного зниження у 2017–2021 роках максимальні значення знову зростають у 2022–2024 роках. Зростання

індексу вказує на можливе посилення антикорупційної політики, цифровізацію державних послуг та підвищення прозорості управління в окремих країнах.

Середнє значення індексу залишається майже незмінним — близько 46–47 балів протягом усього досліджуваного періоду. Отже, загалом у вибірці країн суттєвого прогресу у боротьбі з корупцією не відбулося. Країни продовжують демонструвати середній або помірно високий рівень сприйняття корупції в державному секторі.

Мінімальні значення індексу є значно нижчими, а у 2024 році знижуються ще більше приблизно до 10 балів, тобто рівень корупції зріс ще більше у деяких країнах вибірки (на це вказує також і невелике падіння в середньому рівні індексу цього року). Таке погіршення може бути викликане політичною нестабільністю, слабкістю державних інституцій, воєнними конфліктами, низьким рівнем верховенства права та недостатньою ефективністю чи кількістю антикорупційних реформ у частині країн вибірки.

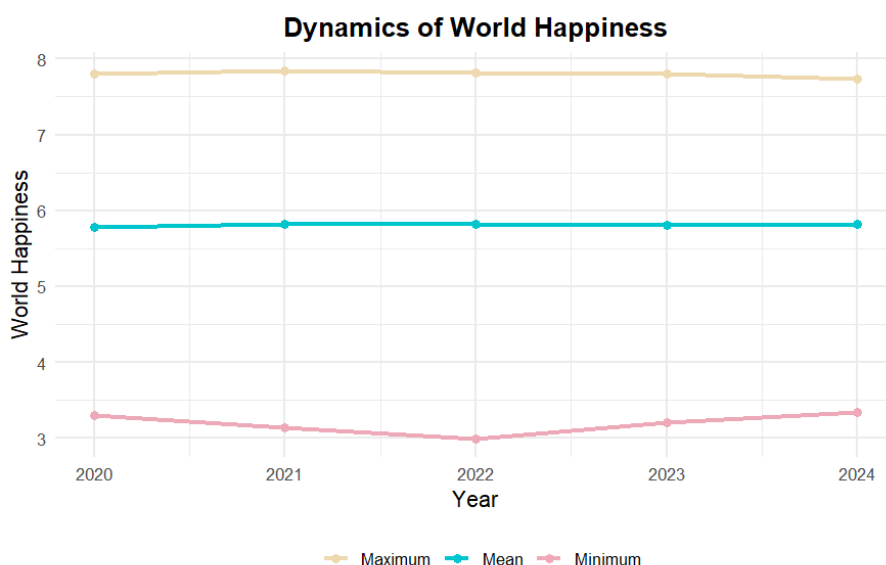


Рисунок 2.5 – Результат кореляційного аналізу у вигляді теплової мапи з вказаними коефіцієнтами

Джерело: розроблено авторкою

За 2020–2024 роки Індекс щастя демонструє загальну стабільність як середніх, так і максимальних значень показника. Середній рівень індексу щастя протягом досліджуваного періоду залишався близько 5,8 бала, а максимальні

значення індексу коливалися в межах 7,7–7,8. Це вказує на те, що країни з найвищим рівнем добробуту, соціальної підтримки, економічної стабільності та довіри до державних інституцій змогли зберегти високий рівень задоволеності життям навіть в умовах глобальних криз.

Мінімальні значення індексу у 2020–2022 роках поступово знижувалися — приблизно з 3,3 до 3 балів, після чого у 2023–2024 роках спостерігається відновлення. Таке зниження могло бути пов'язане з негативними соціально-економічними наслідками пандемії COVID-19: зростанням безробіття, погіршенням психоемоційного стану населення та посиленням економічної нестабільності у багатьох країнах. Подальше покращення показника в такому випадку вказує на поступову адаптацію економік і суспільств до нових умов та часткове відновлення рівня життя населення.

2.3 Аналіз таблиць спряженості соціально-економічних показників

Для первинного оцінювання взаємозв'язків між соціально-економічними показниками було здійснено кореляційний аналіз за допомогою функції `cor()`. Побудовано матрицю кореляцій та її графічну візуалізацію у вигляді теплової карти, де колір клітинок і числові значення відображають силу та напрям зв'язку між показниками.

У результаті розрахунків отримано значення коефіцієнтів кореляції, відображені на рис. 2.6:

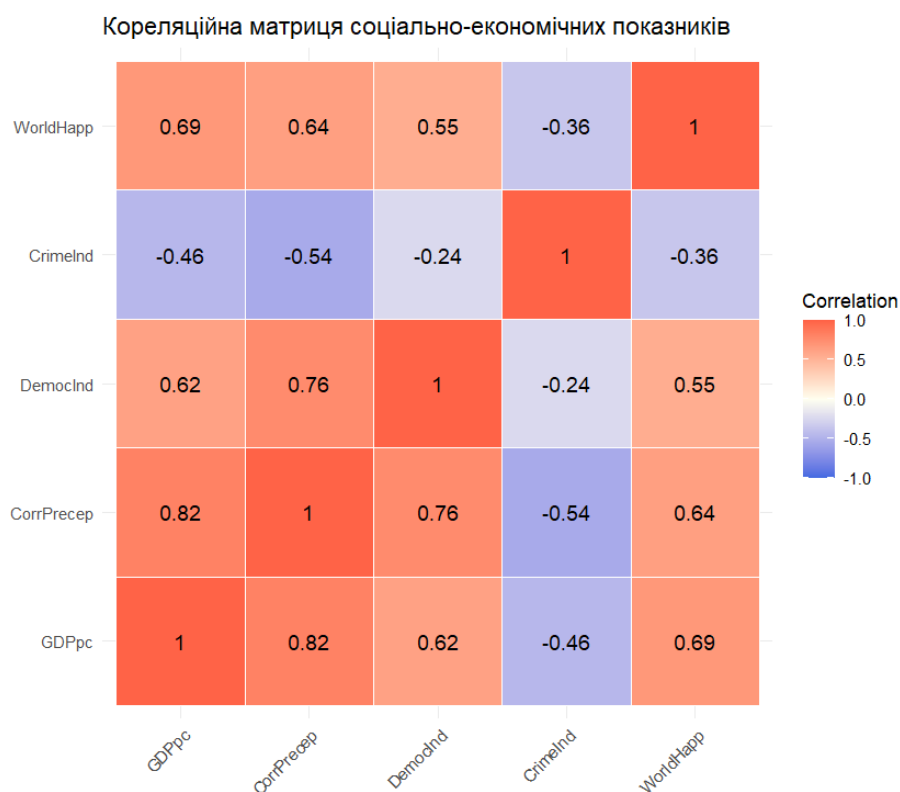


Рисунок 2.6 – Результат кореляційного аналізу у вигляді теплової мапи з вказаними коефіцієнтами

Джерело: розроблено авторкою на основі [8, 9, 15, 21, 25]

Найсильніший додатний зв'язок спостерігається між ВВП на душу населення та Індексом сприйняття корупції ($r = 0,82$). Це свідчить, що країни з вищим рівнем економічного розвитку частіше характеризуються нижчим рівнем корупції та більшою прозорістю державного сектору. Помітний додатний зв'язок також спостерігається між Індексом сприйняття корупції та Індексом демократії ($r = 0,76$), що може вказувати на те, що демократичні режими частіше мають ефективніші механізми контролю корупційної діяльності.

Кореляція між ВВП на душу населення та Індексом щастя становить $r = 0,69$ та вказує на помірно сильний додатний зв'язок між економічним добробутом країни та рівнем задоволеності життям населення.

Для Індексу злочинності спостерігається від'ємний зв'язок з більшістю показників. Зокрема, між ВВП на душу населення та Індексом злочинності отримано коефіцієнт $r = -0,46$. Він говорить про помірну тенденцію до зниження рівня злочинності у більш економічно розвинених країнах.

Отже, вже на початковому етапі аналізу можемо припустити наявність статистичних взаємозв'язків між обраними соціально-економічними характеристиками країн світу. Доцільним буде продовжити дослідження цих залежностей за допомогою інших методів аналізу.

Для дослідження структури взаємозв'язків між соціально-економічними показниками було застосовано таблиці спряженості.

Перед побудовою таблиць вихідні кількісні показники були перетворені у категоріальні змінні. Категоризація здійснювалася здебільшого відповідно до категорій, визначених укладачами показників, які описувалися у першому розділі цієї роботи.

Індекс сприйняття корупції було згруповано за рівнем корупції та прозорості, індекс демократії — за типами політичного режиму, а індекс злочинності — за рівнем криміногенної ситуації. Для індексу щастя було використано квартильний поділ вибірки, оскільки чіткої категоризації автори звіту цього індексу не надають. Для показника ВВП на душу населення було використано класифікаційні вказівки Світового Банку на FY25 (фінансовий рік 2025, що розпочався 1 Липня 2024 року):

- Низький дохід: $\leq 1,145$;
- Нижче середнього: 1,146-4,515;
- Вище середнього: 4,516-14,005;
- Високий дохід: $> 14,005$.

Далі для кожної пари ВВП та показників було побудовано двовходові таблиці спряженості. У таблицях відображено абсолютні частоти спостережень, тобто кількість країн, які одночасно належать до певних категорій двох показників. Для статистичної перевірки наявності взаємозв'язку між категоріальними змінними застосовано критерій хі-квадрат Пірсона (`chisq.test`). Оскільки критерій хі-квадрат лише підтверджує або спростовує статистичну значущість залежності, з метою оцінки сили взаємозв'язку між змінними та порівняння результатів між таблицями розраховано міру Крамера (Cramer's V) та коефіцієнт спряженості.

Результати розрахунків згруповано у табл. 2.2-2.5 (в дужках після кожного абсолютного значення записаний відсоток, який значення складає від загальної суми за рядком):

Таблиця 2.2 – Таблиця спряженості ознак ВВП на душу населення та Індексу сприйняття корупції

	Висока корупція (0-29)	Помірна корупція (30-49)	Низька корупція (50-69)	Дуже низька корупція (70- 100)
Низький дохід	3 (75%)	1 (25%)	0	0
Нижче середнього	11 (45%)	13 (54,2%)	0	0
Вище середнього	6 (17,1%)	25 (71,4%)	4 (11,4%)	0
Високий дохід	1 (2%)	11 (22%)	20 (40%)	18 (36%)

Джерело: розроблено авторкою

Таблиця 2.3 – Таблиця спряженості ознак ВВП на душу населення та Індексу демократії

	Авторитарний	Гібридний	Недосконала демократія	Повна демократія
Низький дохід	2 (59%)	2 (50%)	0	0
Нижче середнього	10 (41,7%)	10 (41,7%)	4 (16,7%)	0
Вище середнього	8 (22,9%)	9 (25,7%)	17 (48,6%)	1 (2,9%)
Високий дохід	5 (10%)	3 (6%)	19 (38%)	23 (46%)

Джерело: розроблено авторкою

Таблиця 2.4 – Таблиця спряженості ознак ВВП на душу населення та Індексу злочинності

	Дуже низький (0–20)	Низький (20– 40)	Помірний (40– 60)	Високий (60–80)	Дуже високий (80–100)
Низький дохід	0	0	2 (50%)	2 (50%)	0
Нижче середнього	0	2 (8,3%)	15 (62,5%)	6 (25%)	1 (4,2%)
Вище середнього	0	7 (20%)	18 (51,4%)	10 (28,6%)	0
Високий дохід	1 (2%)	30 (60%)	19 (38%)	0	0

Джерело: розроблено авторкою

Таблиця 2.5 – Таблиця спряженості ознак ВВП на душу населення та Індексу щастя

	Низький	Нижче середнього	Вище середнього	Високий
Низький дохід	3 (75%)	1 (25%)	0	0
Нижче середнього	17 (70,8%)	4 (16,7%)	3 (12,5%)	0
Вище середнього	8 (22,9%)	19 (54,3%)	8 (22,9%)	0
Високий дохід	1 (2%)	4 (9%)	17 (34%)	28 (56%)

Джерело: розроблено авторкою

Крім того, для оцінювання зв'язку було використано критерій χ^2 Пірсона, міру Крамера V та коефіцієнт спряженості C . Усі отримані результати можна побачити у табл. 2.6:

Таблиця 2.6 – Розраховані показники оцінки зв'язку ВВП та індексів

	ВВП на душу та Індекс сприйняття корупції	ВВП на душу та Індекс демократії	ВВП на душу та Індекс злочинності	ВВП на душу та Індекс щастя
χ^2	74,683	52,504	40,483	92,007
p-value	< 0,001	< 0,001	< 0,001	< 0,001
V	0,4694	0,3935	0,3456	0,5210
C	0,6308	0,5632	0,5136	0,6699

Джерело: розроблено авторкою

Для всіх побудованих таблиць спряженості значення критерію χ^2 -квадрат Пірсона виявилися статистично значущими ($p < 0,001$), отже можна відхилити гіпотезу про незалежність ознак: між рівнем ВВП на душу населення та досліджуваними показниками існує статистично значущий взаємозв'язок. Значення коефіцієнта C перебували в межах від приблизно 0,514 до 0,670, це вказує на наявність помірного зв'язку між ознаками. Найвище значення коефіцієнта спряженості було отримано для ВВП та індексу щастя населення ($C=0,670$), а найменше — для ВВП та індексу злочинності ($C=0,514$).

Значення міри Крамера для рівня доходу країни та індексу щастя населення становило $V = 0,521$, це помірно сильний зв'язок. Аналіз таблиці спряженості показує чітку тенденцію: у країнах із високим рівнем доходу переважають високі значення індексу щастя, відповідно в країнах із низьким та нижчим за середній рівнем доходу домінують низькі значення цього показника. Зокрема, серед країн із високим рівнем доходу 56% потрапили до групи з найвищими значеннями індексу

щастя, тоді як серед країн із низьким доходом таких країн не було зовсім. Натомість, 75% країн із низьким доходом віднесені до групи з найнижчим рівнем щастя. Отримані результати свідчать про суттєвий вплив економічного добробуту на суб'єктивну оцінку якості життя населення.

Досить сильний зв'язок також спостерігається між рівнем ВВП та індексом сприйняття корупції ($V = 0,469$). Із підвищенням рівня доходу країн простежується зменшення поширеності корупції та зростання частки країн із високими показниками прозорості. Серед країн із високим рівнем доходу 40% характеризуються помірною та ще 36% — високою некоруптованістю. Серед країн із низьким доходом цілком відсутні країни з великими оцінками прозорості, натомість 75% країн із цієї дохідної категорії мають високий рівень корупції. Можна зробити висновок про значущий взаємозв'язок між економічним розвитком та якістю інституційного середовища.

Аналіз взаємозв'язку між рівнем доходу та індексом демократії також підтвердив наявність статистично значущої залежності ($V = 0,394$). У країнах із високим рівнем доходу 46% країн належать до категорії повних демократій, а ще 38% — до категорії недосконалих демократій. Серед країн із низьким доходом повні демократії взагалі відсутні, домінують авторитарні та гібридні режими. Для країн із доходом нижче середнього ситуація схожа: частка авторитарних та гібридних режимів — по 41,7% відповідно.

Найменшу, хоча також статистично значущу силу зв'язку було виявлено між рівнем доходу та індексом злочинності ($V = 0,346$). На відміну від інших показників, залежність тут є менш однорідною, однак загальна тенденція все ж простежується. Для країн із високим рівнем доходу характерні переважно низькі або помірні рівні злочинності: 60% таких країн потрапили до категорії з низьким рівнем злочинності, а випадки високої злочинності були відсутні. У країнах із нижчим рівнем доходу значно частіше спостерігаються помірні та високі значення індексу — серед країн із низьким доходом половина країн належить до категорії з високим рівнем злочинності.

Результати, отримані в результаті формування таблиць спряженості ознак далі візуалізовано через декілька стовпчикових діаграм:

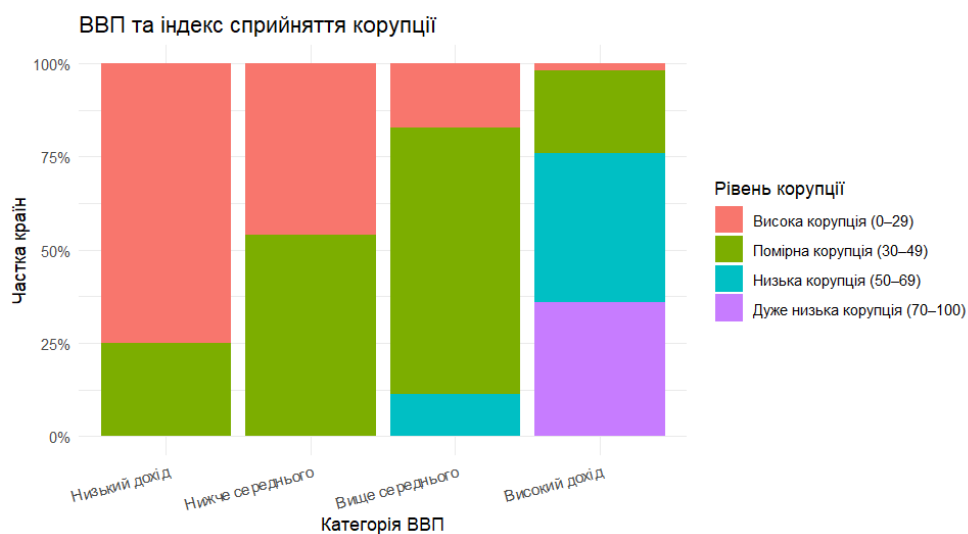


Рисунок 2.7 – Візуалізація, створена на основі спряженості ознак ВВП на душу населення та Індексу сприйняття корупції

Джерело: розроблено авторкою

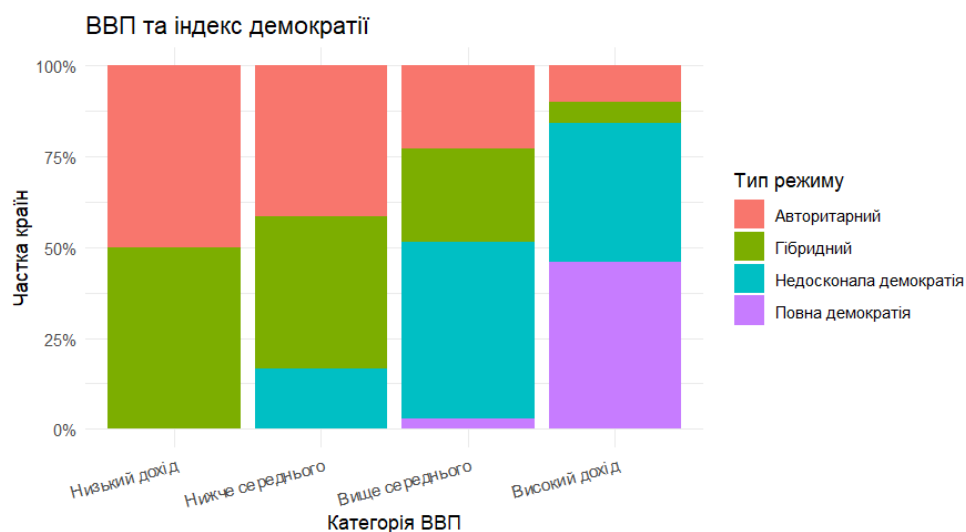


Рисунок 2.8 – Візуалізація, створена на основі спряженості ознак ВВП на душу населення та Індексу демократії

Джерело: розроблено авторкою

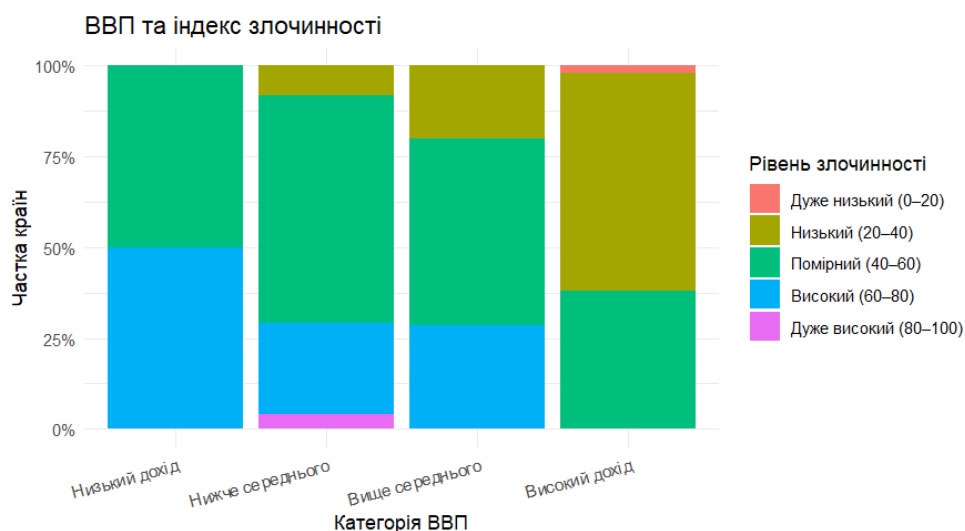


Рисунок 2.9 – Візуалізація, створена на основі спряженості ознак ВВП на душу населення та Індексу злочинності

Джерело: розроблено авторкою

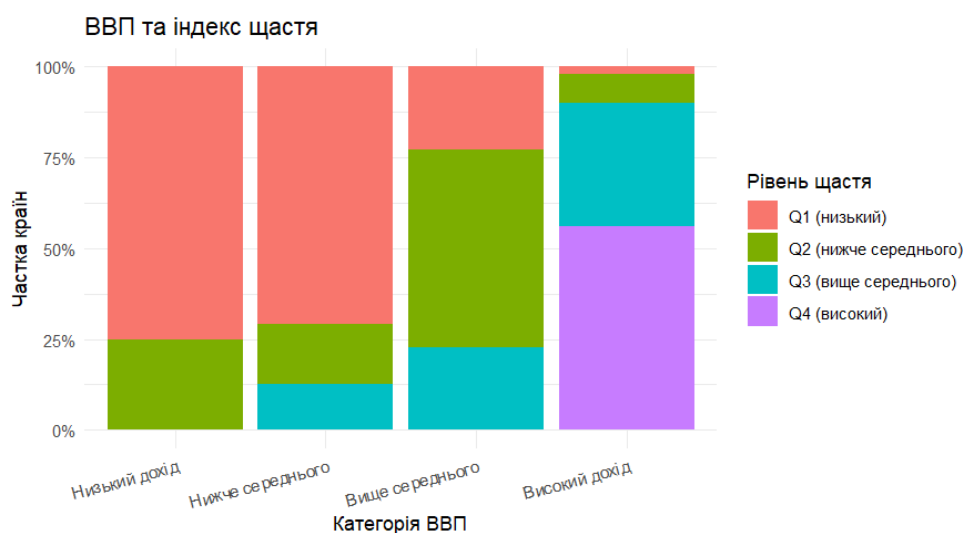


Рисунок 2.10 – Візуалізація, створена на основі спряженості ознак ВВП на душу населення та Індексу щастя

Джерело: розроблено авторкою

Чіткі структурні зміни, де зі зростанням рівня ВВП помітно збільшується частка країн із позитивними характеристиками, найяскравіше видно для показників сприйняття корупції та демократії. Загалом, графіки добре відображають динаміку зміни індексів залежно від рівня доходу.

2.4 Побудова парних регресійних моделей

Перед побудовою регресійних моделей було здійснено логарифмування показника ВВП на душу населення. Необхідність такого перетворення пов'язана з тим, що початковий показник GDPpc характеризується значною варіацією між країнами та асиметричним розподілом. Оскільки вибірка охоплює широкий обсяг держав (113), в ній одночасно присутні країни з низькими та дуже високими значеннями ВВП на душу населення. Додаткову обробку показника було здійснено для того, щоб уникнути небажаного впливу екстремальних значень на результати моделювання.

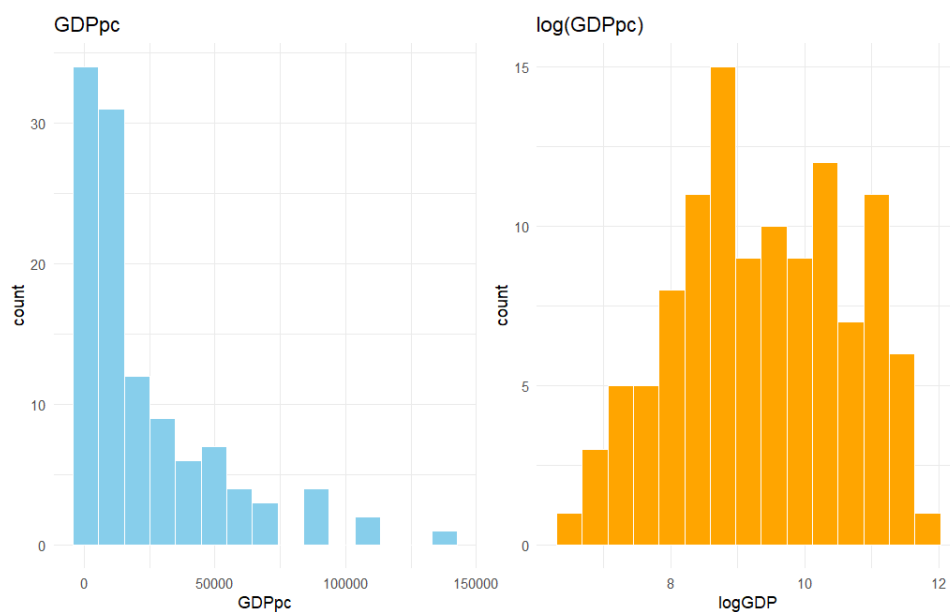


Рисунок 2.11 – Розподіл ВВП на душу населення до та після логарифмування

Джерело: розроблено авторкою

Логарифмування дозволяє зменшити масштаб варіації показника, послабити вплив викидів та наблизити розподіл змінної до нормального. Крім того, використання логарифмічної форми залежної змінної сприяє підвищенню стабільності оцінок параметрів регресійної моделі та покращенню виконання припущень класичної лінійної регресії. Логарифмований показник надалі використовувався як залежна змінна у побудованих регресійних моделях.

Для оцінки впливу рівня корупції на економічний добробут було побудовано парну лінійну регресійну модель:

$$\ln(GDPpc) = 6,806 + 0,0549 \times CorrPresep + \varepsilon$$

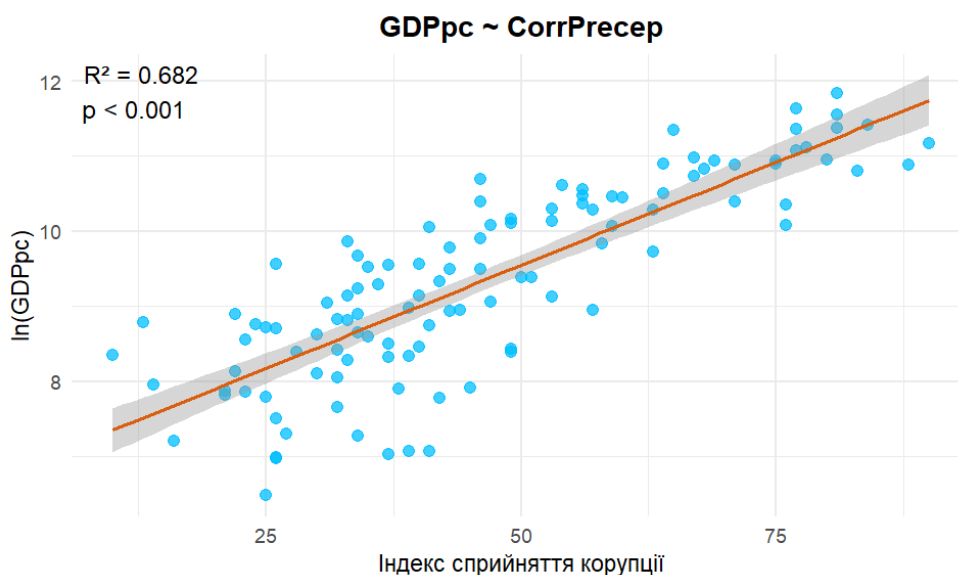


Рисунок 2.12 – Парна регресійна модель ВВП на душу населення та Індексу сприйняття корупції

Джерело: розроблено авторкою

Результати оцінювання показали, що коефіцієнт при змінній *CorrPresep* є статистично значущим ($p < 0,001$). Додатний знак коефіцієнта свідчить про наявність прямого зв'язку між індексом сприйняття корупції та рівнем ВВП на душу населення. Оскільки вищі значення індексу відповідають нижчому рівню корупції, зростання прозорості державних інститутів асоціюється зі збільшенням економічного добробуту населення.

Коефіцієнт детермінації R^2 становить 0,682, тобто близько 68,2 % варіації логарифма ВВП на душу населення пояснюється відмінностями в рівні сприйняття корупції між країнами. Серед побудованих парних моделей цей результат R^2 найвищий.

Парна регресійна модель ВВП на душу населення та індексу демократії має вигляд:

$$\ln(GDPpc) = 7,035 + 0,390 \times DemocInd + \varepsilon$$

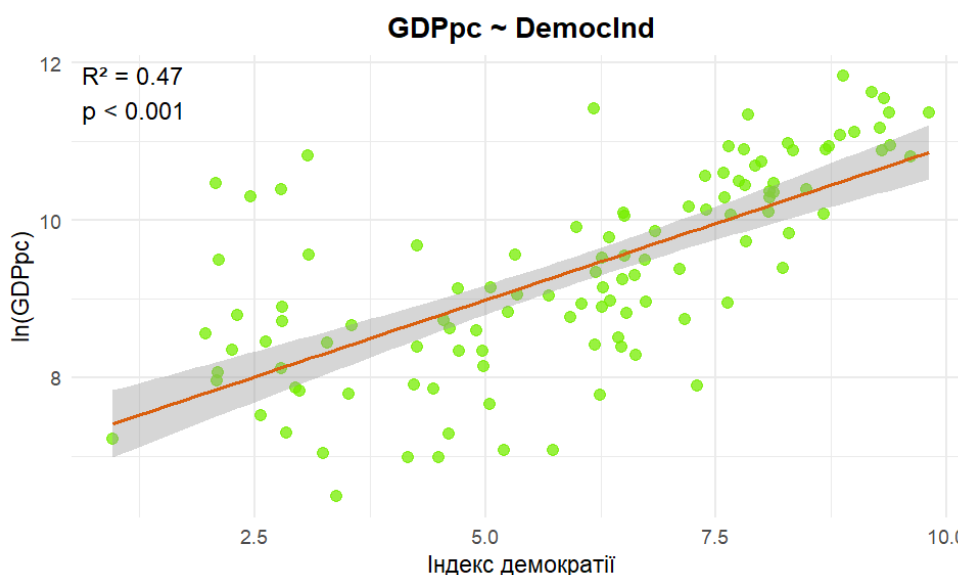


Рисунок 2.13 – Парна регресійна модель ВВП на душу населення та Індексу демократії

Джерело: розроблено авторкою

Коефіцієнт при індексі демократії є додатним та статистично значущим ($p < 0,001$). Отже, країни з вищим рівнем демократичного розвитку, як правило, характеризуються вищим рівнем ВВП на душу населення.

Коефіцієнт детермінації R^2 становить 0,470, тобто індекс демократії пояснює приблизно 47,0 % варіації залежного показника (ВВП на душу населення).

Отже, хоча вплив демократії є досить суттєвим, його пояснювальна здатність помітно нижча порівняно з попередньо розглянутим індексом сприйняття корупції. За результатом парної регресійної моделі рівень демократії важливий для економічного розвитку, однак не є його єдиним визначальним чинником.

Для дослідження взаємозв'язку між рівнем злочинності та економічним розвитком було побудовано модель:

$$\ln(GDPpc) = 11,605 + 0,0502 \times CrimeInd + \varepsilon$$

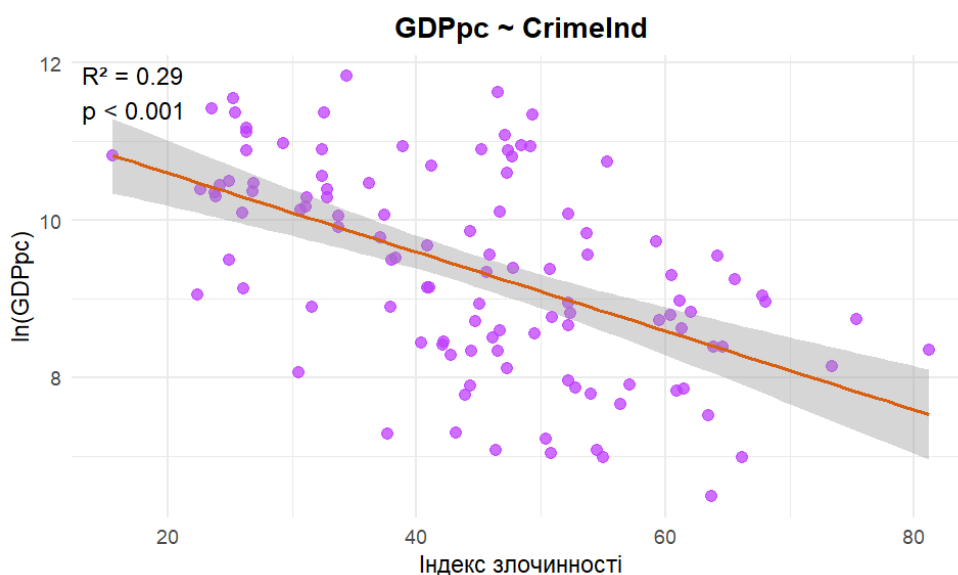


Рисунок 2.14 – Парна регресійна модель ВВП на душу населення та Індексу злочинності

Джерело: розроблено авторкою

Коефіцієнт при індексі злочинності є від’ємним і статистично значущим ($p < 0,001$), тому існує обернений зв’язок між досліджуваними показниками. Це означає, що зі зростанням рівня злочинності економічний добробут населення має тенденцію до зниження.

Коефіцієнт детермінації R^2 становить 0,290, модель пояснює лише 29,0 % варіації ВВП. Порівняно з іншими досліджуваними факторами індекс злочинності демонструє найменшу пояснювальну здатність. Це свідчить про те, що хоча злочинність пов’язана з рівнем економічного розвитку, серед розглянутих лінійних моделей зв’язок цього показника з ВВП на душу населення найменш тісний.

Парна регресійна модель має вигляд:

$$\ln(GDPpc) = 3,347 + 1,032 \times CorrPrecep + \varepsilon$$

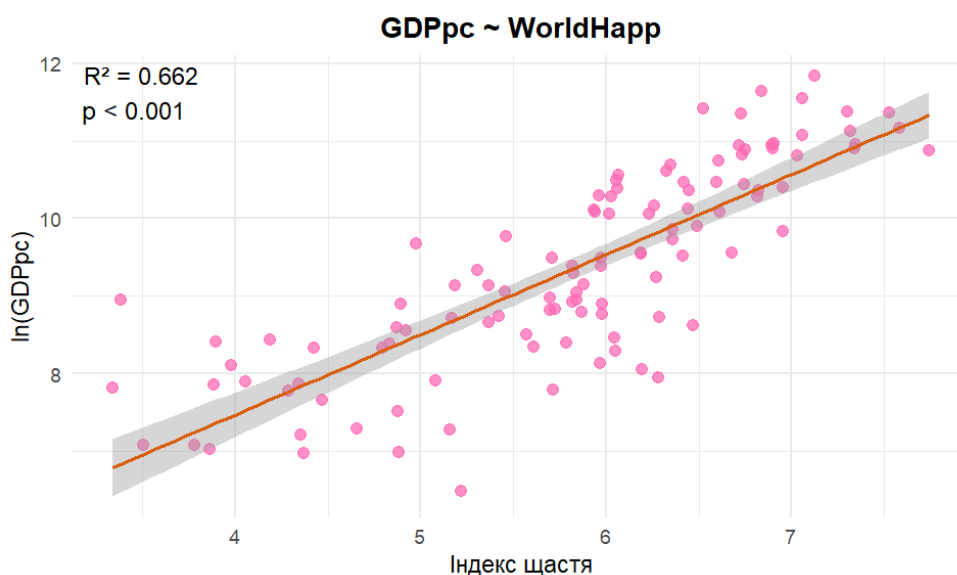


Рисунок 2.15 – Парна регресійна модель ВВП на душу населення та Індексу щастя

Джерело: розроблено авторкою

Коефіцієнт при індексі щастя є додатним і статистично значущим ($p < 0,001$). Це означає, що країни з вищими показниками суб'єктивного благополуччя населення зазвичай характеризуються вищим рівнем ВВП на душу населення.

Коефіцієнт детермінації R^2 становить 0,662, тобто близько 66,2 % варіації логарифма ВВП на душу населення пояснюється відмінностями в індексі щастя. За величиною коефіцієнта детермінації ця лінійна модель поступається лише моделі з індексом сприйняття корупції. Отримані результати свідчать про тісний взаємозв'язок між економічним добробутом та загальною задоволеністю життям населення.

Порівняння побудованих моделей показало, що всі досліджувані фактори мають статистично значущий зв'язок із ВВП на душу населення ($p < 0,001$). Показники всіх парних моделей можна побачити у табл. 2.7:

Таблиця 2.7 – Показники отриманих парних регресійних моделей

Модель	Коефіцієнт при X	R^2	p-value
GDPpc ~ CorrPrecep	0,0549	0,682	<0,001
GDPpc ~ DemocInd	0,3903	0,470	<0,001
GDPpc ~ CrimeInd	-0,0502	0,290	<0,001
GDPpc ~ WorldHapp	1,0318	0,662	<0,001

Джерело: розроблено авторкою

Отже, серед досліджуваних показників найбільш тісно з рівнем економічного розвитку пов'язані якість державного управління та суб'єктивний добробут населення, тоді як вплив злочинності виявився найслабшим. Це дає підстави очікувати, що в множинній регресійній моделі найбільший внесок у пояснення ВВП на душу населення забезпечуватимуть саме індекси сприйняття корупції та щастя населення.

РОЗДІЛ 3

БАГАТОВИМІРНИЙ АНАЛІЗ ВЗАЄМОЗВ'ЯЗКУ СОЦІАЛЬНО-ЕКОНОМІЧНИХ ПОКАЗНИКІВ КРАЇН СВІТУ

3.1 Кластерний аналіз

Перед проведенням кластерного аналізу було здійснено підготовку даних. Для усунення впливу масштабу показників дані було стандартизовано, в результаті чого кожна змінна була приведена до однакового масштабу з середнім значенням 0 та стандартним відхиленням 1.

Наступним кроком було побудовано дендрограму методом ієрархічної кластеризації. Її можна побачити на рис. 3.1:



Рисунок 3.1 – Дендрограма країн вибірки

Джерело: розроблено авторкою

Дендрограма використовувалася для того, щоб попередньо оцінити можливу кількість кластерів у сукупності. На основі візуального аналізу можна припустити поділ на 3 або 5 кластерів.

Для визначення оптимальної кількості кластерів далі було застосовано методи ліктя та силуету. Метод ліктя ґрунтується на оцінюванні внутрішньокластерної дисперсії залежно від кількості кластерів. Зі збільшенням кількості кластерів внутрішньокластерна варіація зменшується, однак після певного моменту темп покращення суттєво сповільнюється. Саме ця точка сповільнення і розглядається як потенційно оптимальна кількість кластерів.

Застосування методу ліктя можна побачити на рис. 3.2:

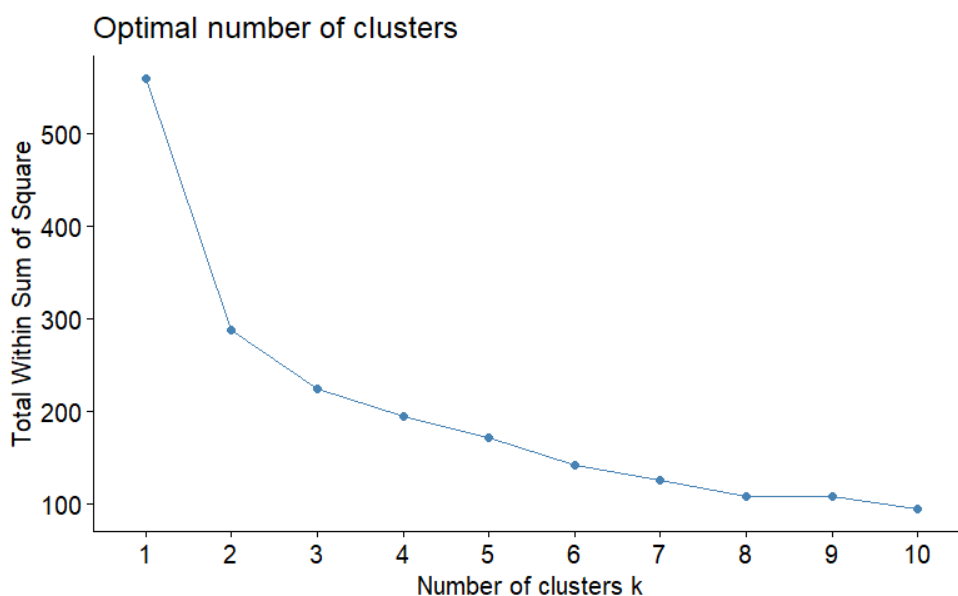


Рисунок 3.2 – Метод ліктя

Джерело: розроблено авторкою

Найбільш різке падіння кривої (зменшення дисперсії) відбувається при переході до двох кластерів, після чого крива стає більш плавною. Тобто, метод ліктя найбільш явно вказує на два кластери.

Коефіцієнт силуету також вказав на можливість поділу сукупності на дві великі групи. Однак слід зауважити, що поділ на два кластери є дуже загальним та обмежує можливості аналізу взаємодії показників соціально-економічної природи. З цієї причини скориставшись попередніми результатами дендрограми було виділено 3 кластери. Було створене графічне зображення цих груп, його показано на рис. 3.3:

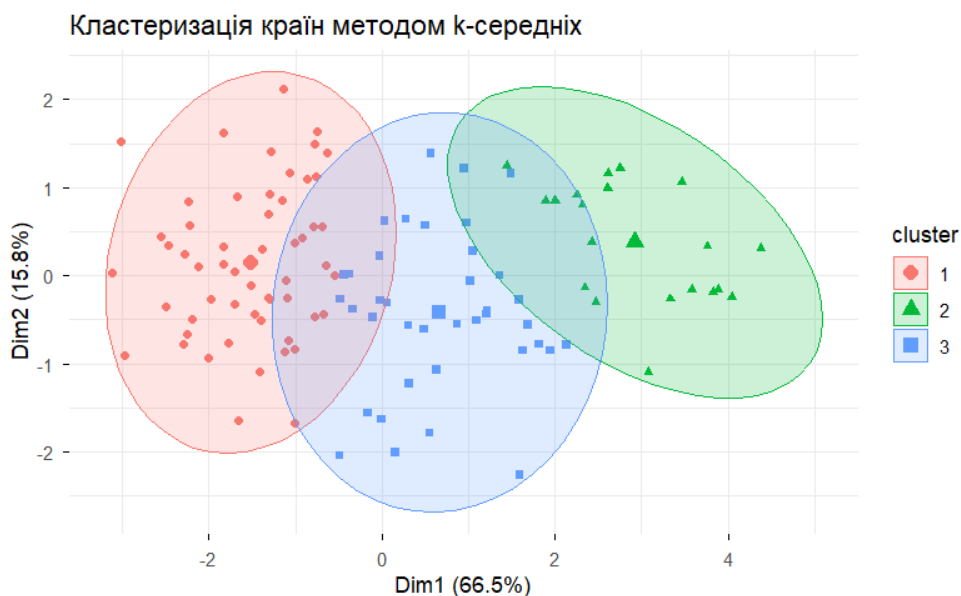


Рисунок 3.3 – Візуалізація трьох кластерів

Джерело: розроблено авторкою

Видно, що поділ на три кластери забезпечує чітке просторове розмежування груп та чітко розділяє досліджувані країни за рівнем соціально-економічного розвитку. З огляду на це для подальшого аналізу обрано саме таку кількість кластерів, а сама кластеризація проводилася методом k-середніх (k-means).

Отже, за результатами кластерного аналізу методом k-середніх усю сукупність із 113 країн було поділено на три кластери, які відрізняються рівнем економічного розвитку, сприйняттям корупції, розвитком демократичних інститутів, рівнем злочинності та суб'єктивним рівнем щастя населення. Візуалізація кластерів показала достатньо чітке розмежування груп у просторі головних компонент. Середні значення за кожним кластером подані у табл. 3.1:

Таблиця 3.1 – Середні значення трьох кластерів

Номер кластеру	Кількість країн	ВВП на душу	Індекс сприйняття корупції	Індекс демократії	Індекс злочинності	Індекс щастя
1	55	5309	31,9	4,53	53,6	5,13
2	20	72209	76,5	8,68	37,8	7,06
3	38	23780	52	6,59	35,4	6,19

Джерело: розроблено авторкою

Перший кластер є найбільшим і охоплює 55 країн переважно з нижчим рівнем соціально-економічного розвитку, серед яких Україна, Індія, Пакистан, Нігерія, Єгипет, Бразилія, ПАР, Індонезія, Мексика та інші. Для цього кластера

характерний найнижчий середній рівень ВВП на душу населення — близько 5304 дол. США. Середнє значення індексу сприйняття корупції становить лише 31,9 бала, що свідчить про доволі високий рівень корупції. Також країни цього кластера характеризуються нижчим рівнем демократії (4,53 бала), найвищим середнім рівнем злочинності (53,6 бала) та найнижчим рівнем щастя населення (5,13 бала). Отже, перший кластер можна охарактеризувати як групу країн із нижчим рівнем економічного розвитку та відносно несприятливими соціально-інституційними характеристиками.

Другий кластер налічує 20 країн і об'єднує найбільш економічно розвинені держави світу: США, Канаду, Німеччину, Швецію, Норвегію, Швейцарію, Данію, Австралію і т.д. Цей кластер демонструє найвищі значення майже за всіма досліджуваними показниками: середній ВВП на душу населення становить понад 72 тис. дол. США, а середній індекс сприйняття корупції дорівнює 76,5 бала і вказує на низький рівень корупції та високу прозорість інституцій. Країни другого кластера мають також найвищий рівень демократії (8,68 бала) та найвищі оцінки щастя населення (7,06 бала). Рівень злочинності у цій групі є помірним і нижчим, ніж у першому кластері. Отже, цей кластер можна охарактеризувати як групу високорозвинених демократичних країн із високою якістю життя та ефективними суспільними інститутами.

Третій кластер охоплює 38 країн із середніми характеристиками розвитку. До нього входять Польща, Чехія, Португалія, Південна Корея, Японія, Італія, Чилі, ОАЕ, Саудівська Аравія та інші країни. Середній ВВП на душу населення у цьому кластері становить приблизно 23,8 тис. дол. США та значно перевищує показники першого кластера, хоч і суттєво поступається другому. Рівень сприйняття корупції (52 бали), демократії (6,59 бала) та щастя населення (6,19 бала) також займає проміжне положення між першою та другою групами. Водночас саме цей кластер характеризується найнижчим середнім рівнем злочинності — 35,4 бала (хоч ця оцінка і не надто нижча за оцінку кластеру 2). Отже, третій кластер можна трактувати як групу країн із середнім або перехідним рівнем соціально-економічного розвитку.

3.2 Модель множинної лінійної регресії

Для оцінювання спільного впливу рівня корупції, демократії, злочинності та суб'єктивного добробуту населення на рівень економічного розвитку було побудовано множинну лінійну регресійну модель:

$$\ln(GDPpc) = 5,041 + 0,0229 \times CorrPrecep + 0,0738 \times DemocInd - 0,0144 \times CrimeInd + 0,5933 \times WorldHapp + \varepsilon$$

Основні результати оцінювання параметрів моделі наведено в табл. 3.2.:

Таблиця 3.2 – Оцінки параметрів множинної регресійної моделі

Змінна	Коефіцієнт	Std. Error	t-статистика	p-value
Константа	5,041	0,413	12,207	<0,001
CorrPrecep	0,0229	0,0051	4,457	<0,001
DemocInd	0,0738	0,0364	2,026	0,045
CrimeInd	-0,0144	0,0046	-3,161	0,002
WorldHapp	0,5933	0,0645	9,202	<0,001

Джерело: розроблено авторкою

Згідно з отриманими результатами всі незалежні змінні є статистично значущими на рівні значущості 5%. Найбільший позитивний вплив на залежну змінну має індекс щастя населення, а індекс злочинності очікувано демонструє негативний вплив. Позитивний зв'язок також спостерігається для індексу сприйняття корупції та індексу демократії.

Основні показники якості побудованої моделі наведено в табл. 3.3:

Таблиця 3.3 – Показники якості множинної регресійної моделі

Показник	Значення
Multiple R ²	0,8375
Adjusted R ²	0,8315
Residual Standard Error	0,523
F-statistic	139,2
p-value (F-test)	<0,001
Кількість спостережень	113

Джерело: розроблено авторкою

Коефіцієнт детермінації свідчить, що модель пояснює 83,8 % варіації логарифма ВВП на душу населення. Скоригований коефіцієнт детермінації має близьке до не скоригованого R² значення, що підтверджує відсутність надлишкових

змінних у моделі. Значення F -критерію та відповідний рівень значущості підтверджують статистичну значущість моделі в цілому.

Водночас високі значення статистичних показників самі по собі не гарантують адекватності моделі, тому для додаткової оцінки якості моделі було побудовано графік фактичних та розрахункових (модельних) значень логарифма ВВП на душу населення (рис. 3.1):

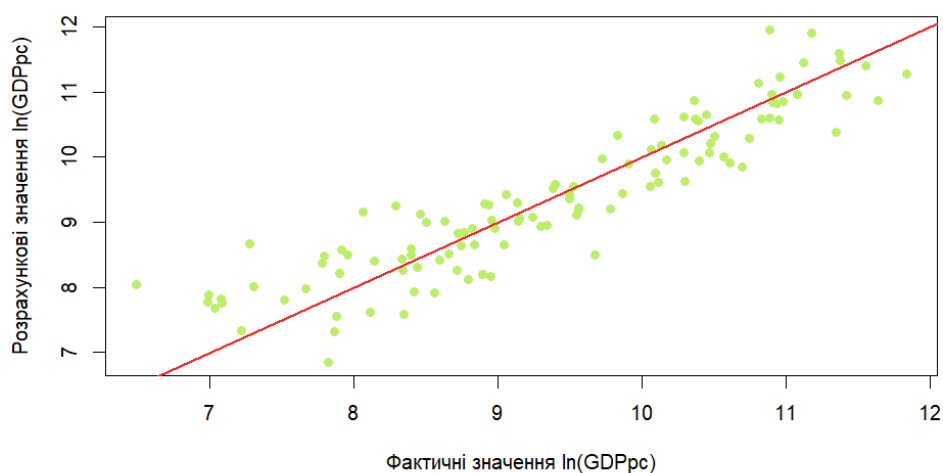


Рисунок 3.4 – Перевірка мультиколінеарності

Джерело: розроблено авторкою

Як видно з рисунка 3.4, більшість точок розташована поблизу прямої, що вказує на високу точність апроксимації даних побудованою множинною регресійною моделлю. Значних систематичних відхилень не спостерігається, тобто можна підтвердити адекватність моделі для опису взаємозв'язку між рівнем ВВП на душу населення та досліджуваними соціально-економічними показниками.

Після оцінювання якості моделі необхідно перевірити відсутність мультиколінеарності між незалежними змінними. Для оцінювання взаємозалежності між цими змінними були розраховані критерії інфляції дисперсії (VIF):

Таблиця 3.4 – Перевірка мультиколінеарності

Змінна	VIF
CorrPresep	3,967
DemocInd	2,725
CrimeInd	1,589
WorldHapp	1,719

Джерело: розроблено авторкою

Усі значення VIF є нижчими за критичний рівень 5, тому ознак суттєвої мультиколінеарності між незалежними змінними не виявлено.

Нормальність залишків оцінювалася за допомогою гістограми, Q-Q графіка та тесту Шапіро-Уїлка.

За результатами тесту Шапіро-Уїлка отримано значення $W = 0,988$, $p = 0,410$. Оскільки p -value перевищує 0,05, немає підстав відхиляти гіпотезу про нормальний розподіл залишків. Однак окрім статистичного тесту, доцільно також здійснити графічний аналіз розподілу залишків:

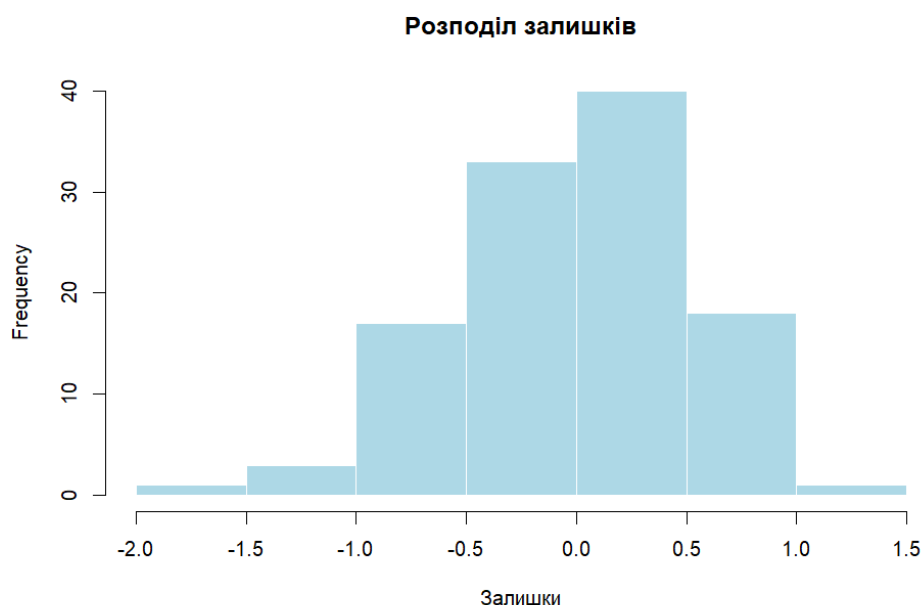


Рисунок 3.5 – Гістограма залишків моделі

Джерело: розроблено авторкою

Гістограма демонструє близький до симетричного розподіл залишків, без виражених ознак зміщення. Далі побудовано Q-Q графік, який дозволяє порівняти емпіричний розподіл залишків із теоретичним нормальним розподілом:

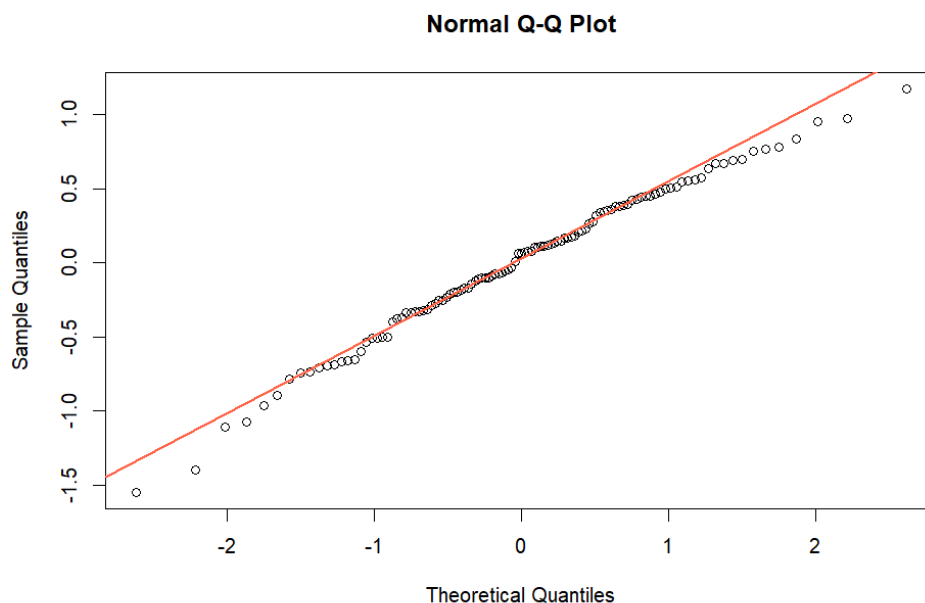


Рисунок 3.6 – Q-Q графік залишків

Джерело: розроблено авторкою

Хоч і присутні певні відхилення з країв, більшість точок розташована вздовж теоретичної прямої, що узгоджується з попередньо отриманими результатами та свідчить про виконання припущення нормальності залишків.

Для перевірки сталості дисперсії (гомоскедастичності) залишків було використано тест Бройша-Пагана. За результатами тесту $BP = 6,376$, а $p = 0,173$. Оскільки p -value перевищує $0,05$, немає підстав відхиляти гіпотезу про сталість дисперсії залишків. Однак крім статистичного тесту, припущення про сталість дисперсії залишків доцільно перевірити графічно шляхом аналізу залежності залишків від прогнозованих значень:

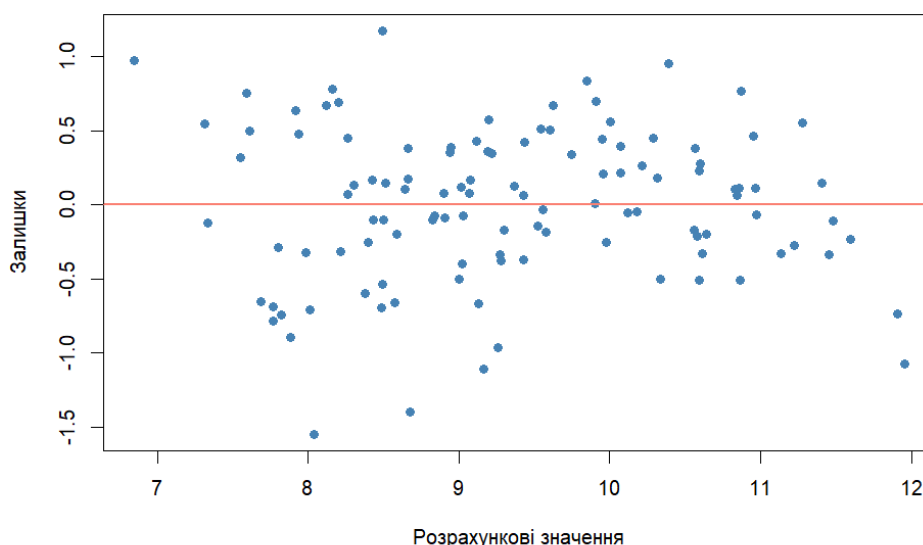


Рисунок 3.7 – Графік залежності залишків від розрахункових значень

Джерело: розроблено авторкою

На рисунку не спостерігається систематичного розширення або звуження розкиду залишків зі зміною рівня розрахункових значень. Точки розташовані відносно рівномірно навколо нульової лінії, що свідчить про відсутність виражених ознак гетероскедастичності. Отже, результати тесту Бройша-Пагана підтвердилися.

В результаті перевірки автокореляції залишків з допомогою тесту Дарбіна-Вотсона отримано результат $DW = 1,997$, а $p = 0,492$. Як видно, значення статистики Дарбіна-Вотсона практично дорівнює 2, отже свідчить про відсутність автокореляції залишків. Припущення про незалежність випадкових похибок виконується.

Отже, після перевірки основних припущень регресійного аналізу можна зробити висновок щодо загальної адекватності побудованої моделі та її здатності описувати взаємозв'язок між досліджуваними показниками.

3.3 Застосування нейронечітких технологій для оцінювання зв'язку між економічним та соціальним розвитком країн

Застосування у попередньому підрозділі методу множинної лінійної регресії до оцінювання зв'язку між економічним та соціальним розвитком країн показало, що зв'язок може бути більш складним, нелінійним. У першому розділі описані нейронечіткі технології, які дозволяють виявляти складну нелінійну залежність між показниками, отже побудуємо нечітку систему для оцінювання залежності між показником ВВП на душу населення та показниками соціального розвитку країни: індексами щастя, демократії, несприйняття корупції та злочинності.

Побудова нечіткої системи складається з двох етапів: структурна ідентифікація та параметрична ідентифікація. На етапі структурної ідентифікації визначається загальна структура нечіткої моделі на основі експериментальної та/або експертної інформації за допомогою нечітких продукційних правил «Якщо-То», які складають нечітку базу знань про наявність зв'язку між показниками моделі. Нечітку модель за результатами структурної ідентифікації можна покращити, тобто зменшити відхилення модельних значень результату від їх експериментальних (реальних, спостережених) значень. Оптимізація параметрів нечіткої моделі може виконуватися за допомогою різних методів оптимізації, як класичних (градієнтного типу), так і методів генетичної оптимізації.

Побудуємо нечітку систему для оцінювання залежності між обраними показниками соціально-економічного розвитку країн за допомогою системи MatLab.

Опишемо етапи побудови нечіткої системи у модулі Fuzzy Logic Toolbox системи MatLab.

1-й етап (структурна ідентифікація).

Фазифікація (уведення нечіткості). Функції належності вхідних факторів та вихідного показника представлені на рис. 3.8.

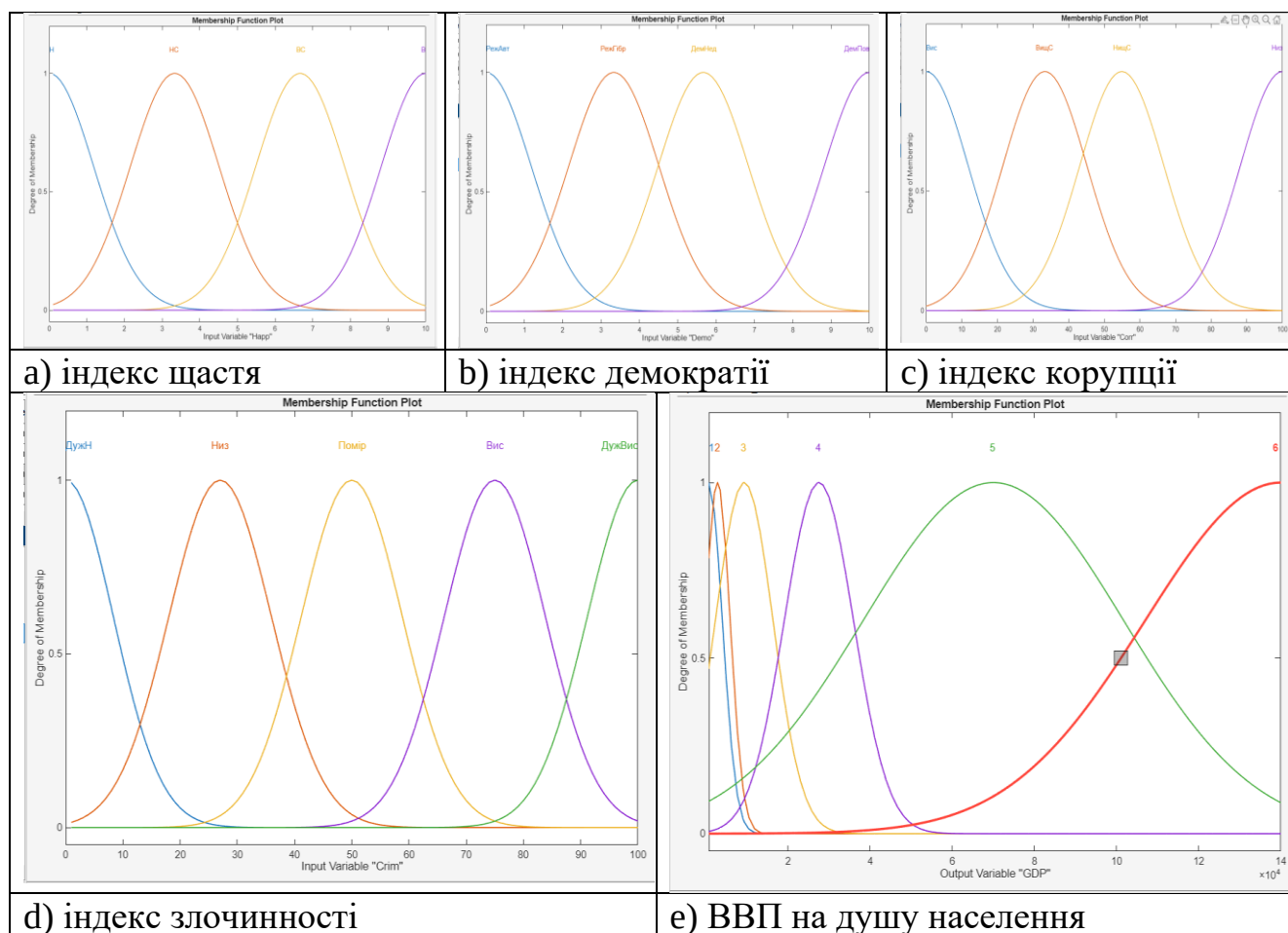


Рисунок 3.8 – Графіки функцій належності входних та вихідної змінних нечіткої моделі

Джерело: розроблено авторкою

Побудова нечіткої бази знань.

Загальна структура нечіткої моделі типу Мамдані представлена на рис. 3.9.

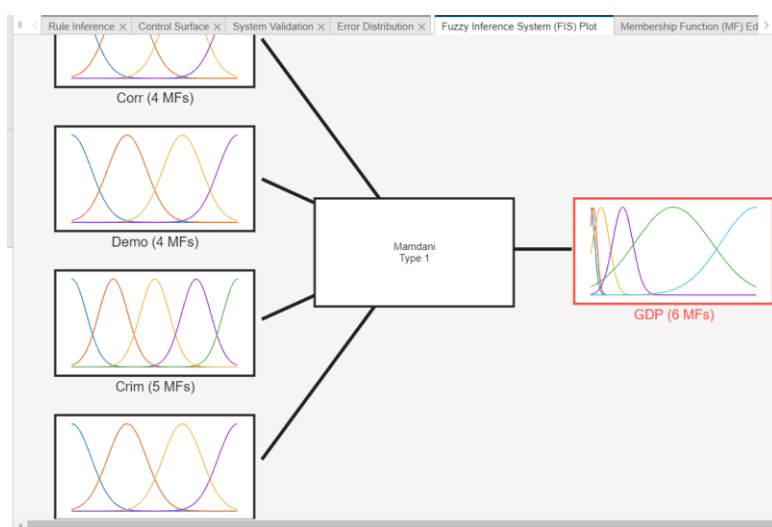


Рисунок 3.9 – Структура нечіткої системи типу Мамдані

Джерело: розроблено авторкою

Фрагменти нечіткої бази знань з 36 правил представлені на рис. 3.10.

	Rule	Weight	Name
1	If Corr is mf3 and Demo is mf3 and Crim is mf1 and Happ is mf3 then GDP	1	rule1
2	If Corr is mf3 and Demo is mf3 and Crim is mf2 and Happ is mf3 then GDP	1	rule2
3	If Corr is mf3 and Demo is mf3 and Crim is mf1 and Happ is mf3 then GDP	1	rule3
4	If Corr is mf3 and Demo is mf3 and Crim is mf1 and Happ is mf2 then GDP	1	rule4
5	If Corr is mf1 and Demo is mf1 and Crim is mf3 and Happ is mf1 then GDP	1	rule5
32	If Corr is mf2 and Demo is mf2 and Crim is mf2 and Happ is mf2 then GDP	1	rule32
33	If Corr is mf3 and Demo is mf3 and Crim is mf1 and Happ is mf3 then GDP	1	rule33
34	If Corr is mf3 and Demo is mf3 and Crim is mf2 and Happ is mf3 then GDP	1	rule34
35	If Corr is mf3 and Demo is mf1 and Crim is mf1 and Happ is mf3 then GDP	1	rule35
36	If Corr is mf3 and Demo is mf2 and Crim is mf1 and Happ is mf3 then GDP	1	rule36

Рисунок 3.10 – Фрагменти нечіткої бази знань

Джерело: розроблено авторкою

Нечітке логічне виведення. Процес нечіткого логічного виведення для правил з номерами 19-24 представлений на рис. 3.11.

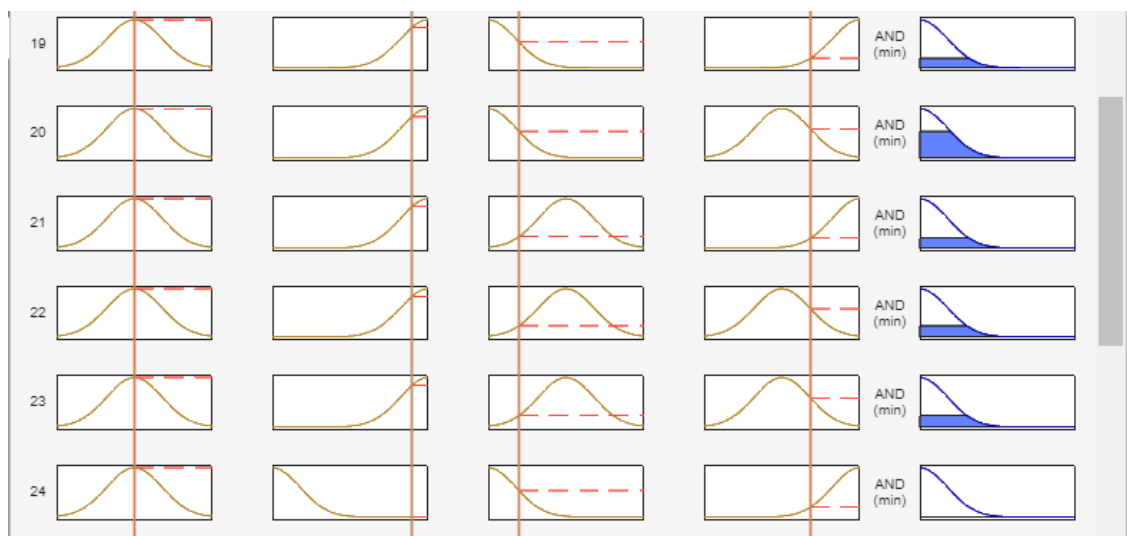


Рисунок 3.11 – Фрагменти процесу нечіткого логічного виведення

Джерело: розроблено авторкою

2-й етап (параметрична ідентифікація, настройка параметрів моделі)

На рис. 3.12, 3.13 представлені результати настройки параметрів нечіткої моделі (параметрів усіх функцій належності та ваг правил нечіткої бази знань).

Результати відновлених залежностей (після настройки нечіткої моделі) ВВП на душу населення від індексів соціального розвитку країни показані на рис. 3.14.

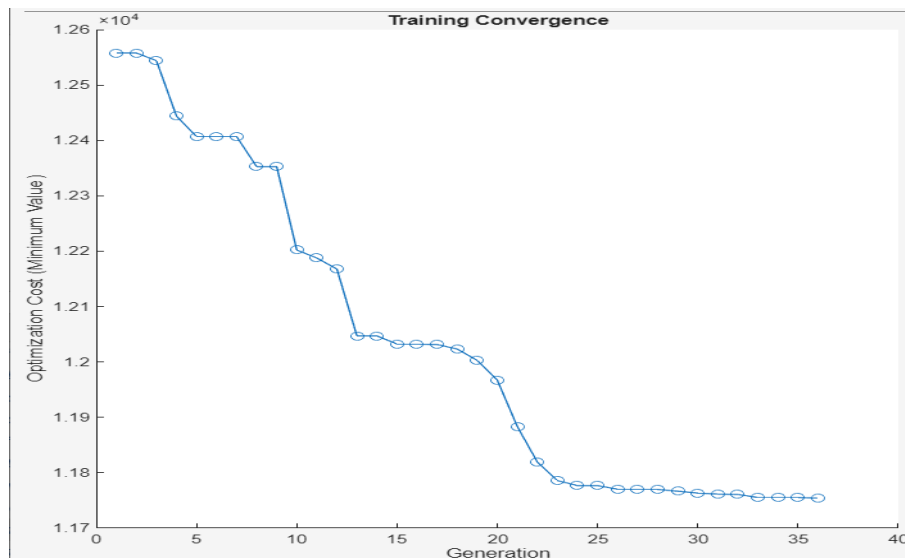


Рисунок 3.12 – Оптимізація параметрів нечіткої моделі за допомогою алгоритму зворотного поширення помилки нейронної мережі

Джерело: розроблено авторкою

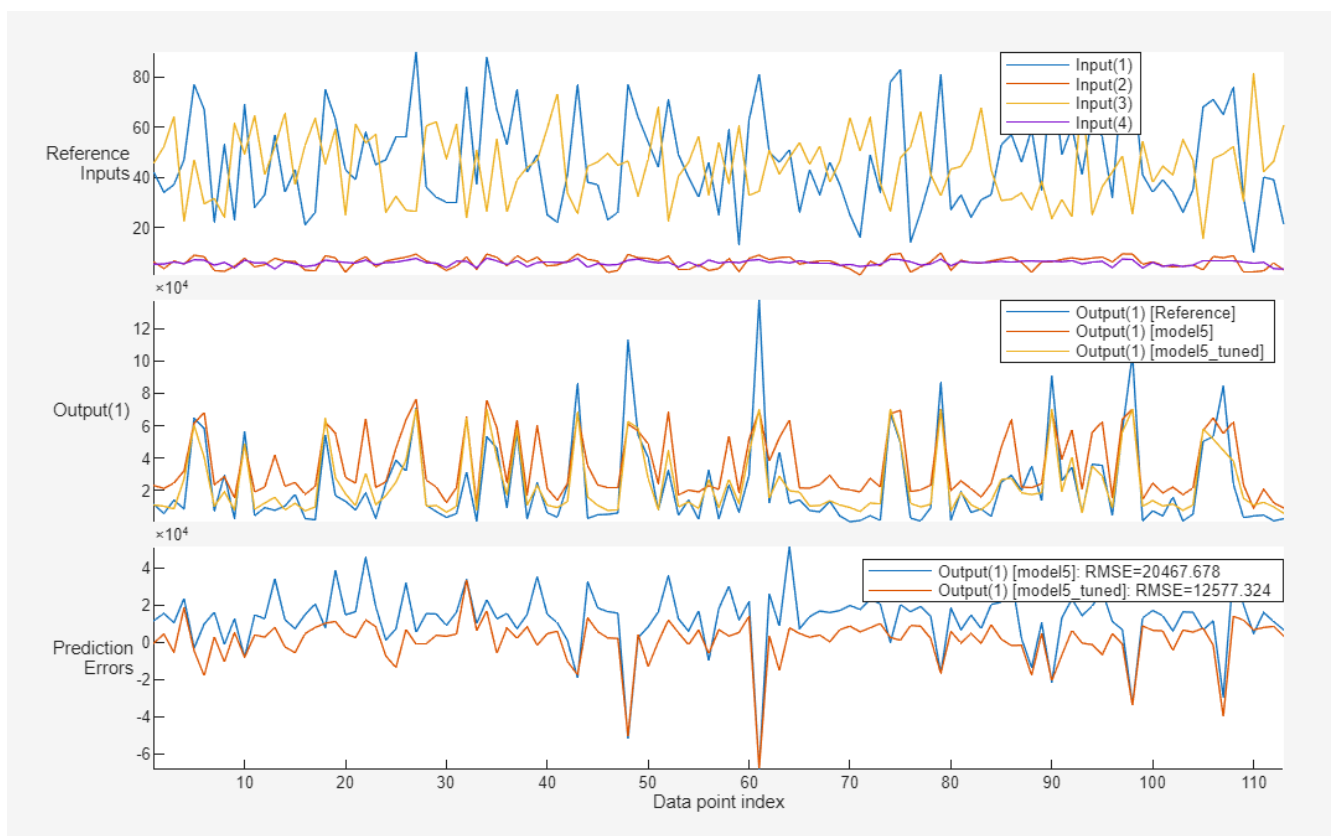


Рисунок 3.13 – Результати настройки параметрів нечіткої моделі

Джерело: розроблено авторкою

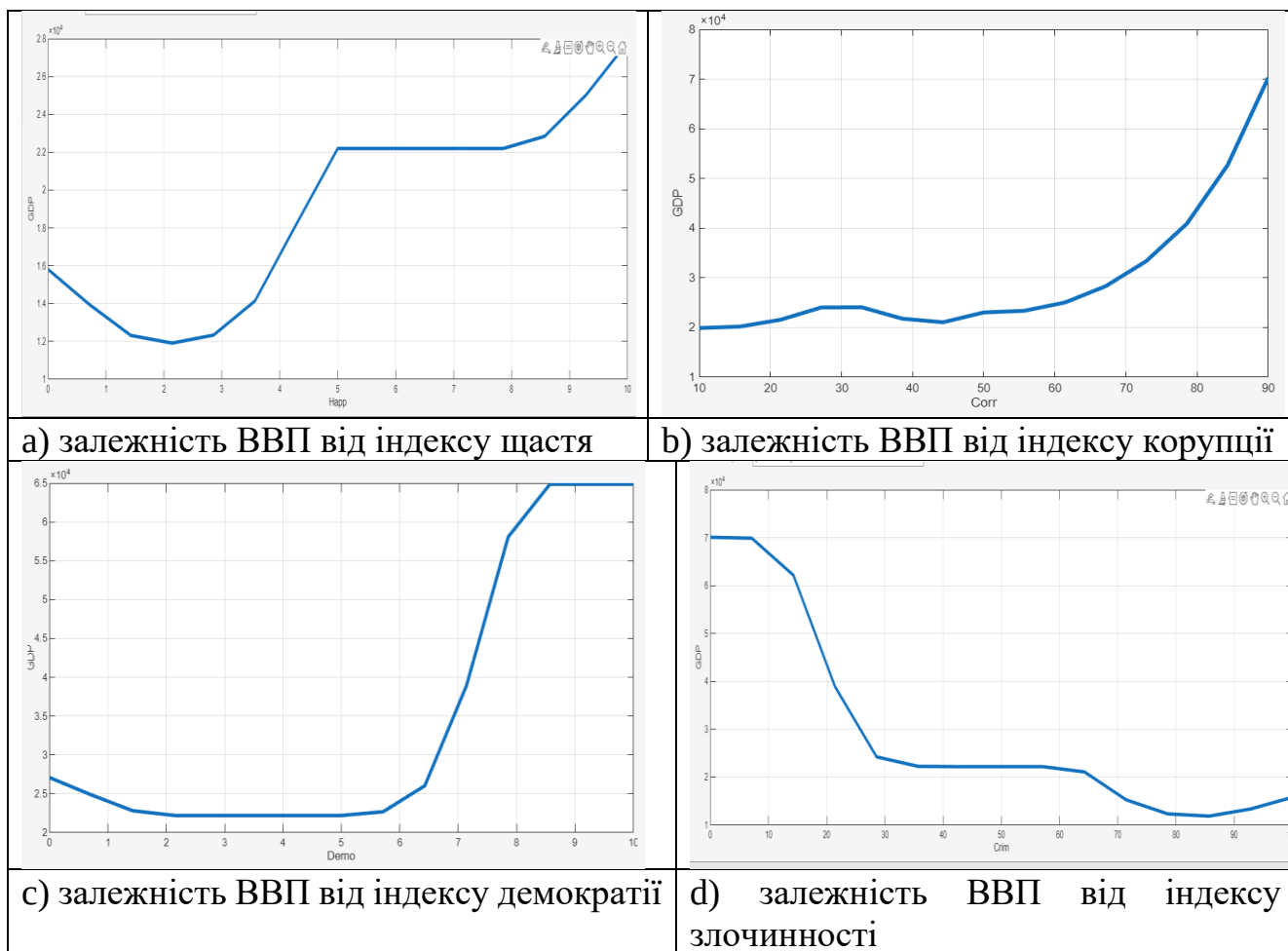


Рисунок 3.14 - Залежності показника ВВП на душу населення від індексів соціального розвитку країн світу

Джерело: розроблено авторкою

Як видно з рис. 3.14 виявлено нелінійну залежність між ВВП на душу населення та індексом щастя, індексом несприйняття корупції, індексом демократії. Залежності між ВВП на душу населення та індексом злочинності має зворотній характер, що відповідає очікуванням.

На рис. 3.15-3.17 представлені тривимірні графіки залежності ВВП на душу населення від попарних комбінацій індексів соціального розвитку.

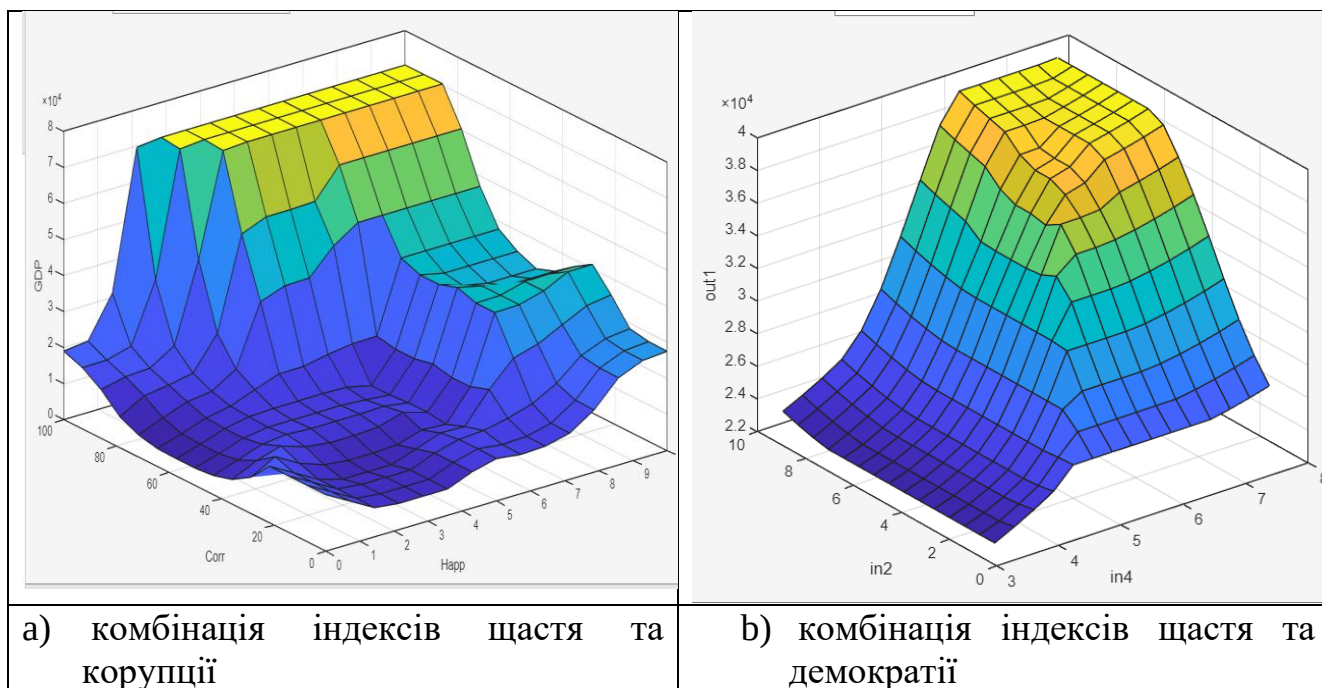


Рисунок 3.15 - Залежність ВВП на душу населення від попарних комбінацій індексів щастя та корупції, індексів щастя і демократії при значеннях інших індексів, фіксованих на середньому рівні

Джерело: розроблено авторкою

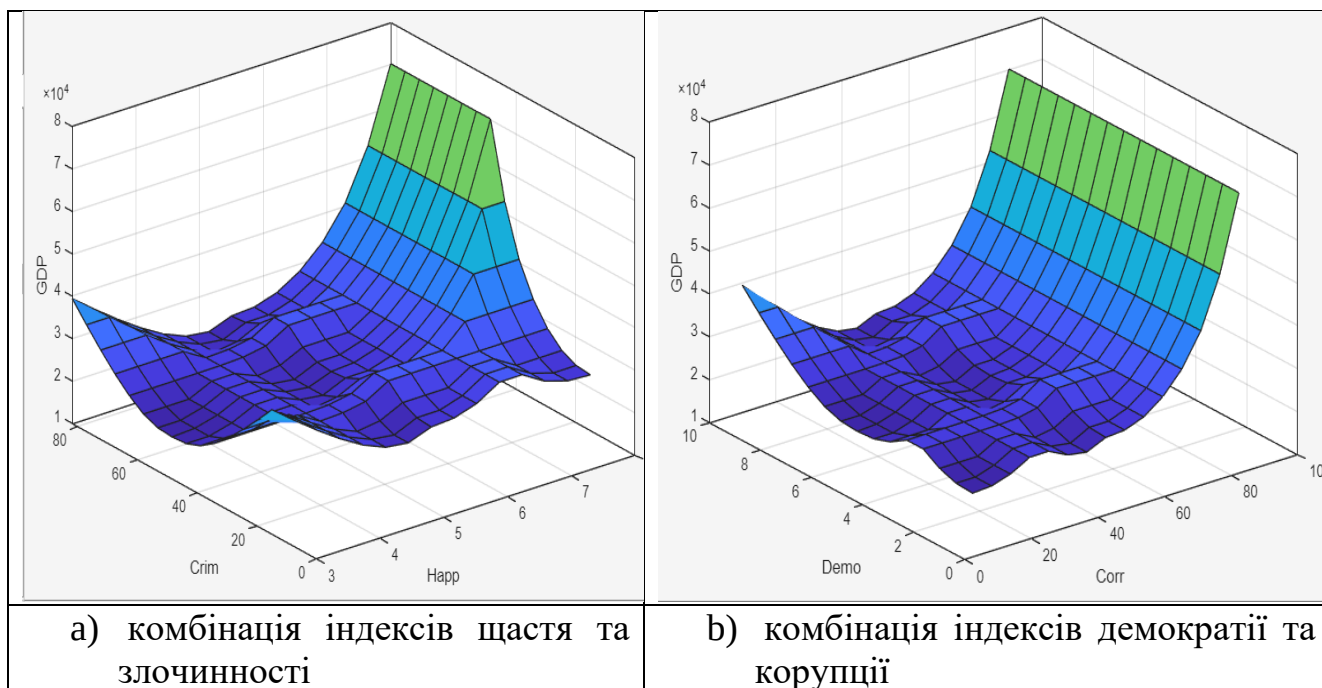


Рисунок 3.16 - Залежність ВВП на душу населення від попарних комбінацій індексів щастя та злочинності, індексів демократії та корупції при значеннях інших індексів, фіксованих на середньому рівні

Джерело: розроблено авторкою

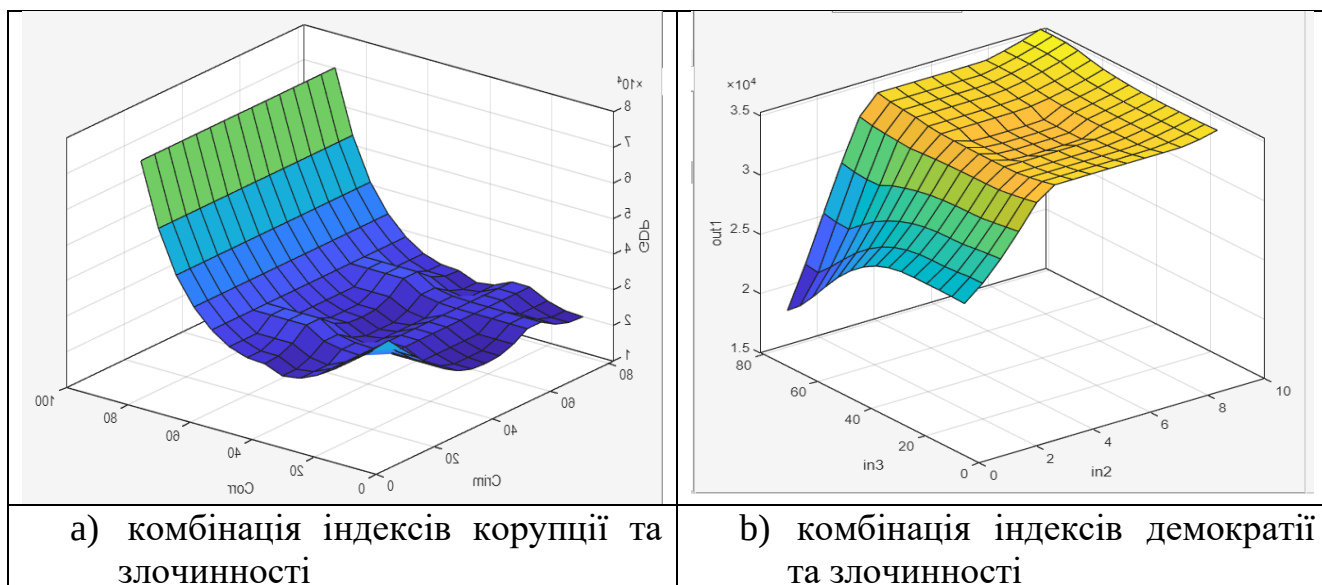


Рисунок 3.17 - Залежність ВВП на душу населення від попарних комбінацій індексів корупції та злочинності, індексів демократії та злочинності при значеннях інших індексів, фіксованих на середньому рівні

Джерело: розроблено авторкою

Оцінка точності нечіткої моделі типу Мамдані проводилась за критерієм RMSE: значення критерію до настройки 20 486, після настройки 12 577 (рис. 3.12).

ВИСНОВКИ

У роботі було досліджено взаємозв'язок між рівнем економічного добробуту країн світу та низкою соціально-економічних показників: індексом сприйняття корупції, індексом демократії, індексом злочинності та індексом щастя населення. Інформаційною базою дослідження стали дані 113 країн світу за 2024 рік.

У процесі виконання роботи було проаналізовано теоретичні підходи до оцінювання соціально-економічного розвитку країн та розглянуто показники, що характеризують економічні, інституційні та соціальні аспекти розвитку суспільства. Після огляду наукової літератури для подальшого аналізу було обрано ВВП на душу населення як узагальнюючий показник економічного добробуту, а також індекси сприйняття корупції, демократії, злочинності та щастя населення як потенційні фактори його формування.

Було розглянуто та обґрунтовано застосування методів статистичного аналізу, багатовимірного аналізу та нейронечіткого моделювання для дослідження взаємозв'язків між обраними показниками. Використання комплексу різних методів дозволило оцінити досліджувані залежності як з позицій класичного статистичного аналізу, так і з урахуванням можливих нелінійних ефектів.

На основі зібраних даних проведено попередній аналіз показників та досліджено особливості їх розподілу. Підтверджено значну диференціацію країн за рівнем ВВП на душу населення, у зв'язку з чим для подальшого регресійного аналізу було використано логарифмування залежної змінної.

За результатами аналізу таблиць спряженості було встановлено статистично значущий зв'язок між рівнем економічного добробуту та всіма досліджуваними соціально-економічними характеристиками ($p < 0,001$). Найтісніший зв'язок виявлено для індексу щастя (V Крамера = 0,521) та індексу сприйняття корупції (V Крамера = 0,469), а для індексів злочинності та демократії зв'язок виявився слабшим (V Крамера = 0,346 та 0,394).

Парний регресійний аналіз показав, що всі досліджувані фактори мають статистично значущий зв'язок із ВВП на душу населення, проте не скрізь він виявився сильний. Найбільшу пояснювальну здатність продемонстрував індекс сприйняття корупції ($R^2 = 0,682$), дещо нижчу — індекс щастя населення ($R^2 = 0,662$). В той же час, для індексу демократії коефіцієнт детермінації становив 0,470, а для індексу злочинності — всього 0,290. Результати вказують на те, що високий рівень прозорості державних інститутів, демократичного розвитку та задоволеності життям асоціюється з вищим рівнем економічного добробуту, тоді як зростання злочинності пов'язане зі зниженням ВВП на душу населення. Тим не менш, за результатами парної регресії пояснювальна здатність індексів демократії та злочинності виявилася нижчою.

У межах багатовимірного аналізу було проведено кластеризацію країн методом k-середніх. За результатами дослідження було виділено три кластери країн, що суттєво відрізняються за рівнем економічного розвитку та значеннями соціально-економічних індикаторів. Перший кластер об'єднав країни з відносно низьким рівнем доходів, високою корупцією та нижчими показниками добробуту населення. Другий кластер сформували найбільш економічно розвинені держави світу з високими значеннями індексів демократії, щастя та сприйняття корупції. Третій кластер представлений країнами із середнім рівнем розвитку, які займають проміжне положення між двома попередніми групами. Такі результати підтвердили існування груп країн зі схожими соціально-економічними характеристиками та засвідчили, що високий рівень економічного розвитку зазвичай супроводжується вищими значеннями індексів щастя, демократії та сприйняття корупції.

Побудована множинна регресійна модель підтвердила статистичну значущість усіх включених факторів та продемонструвала високий рівень пояснення залежної змінної ($R^2 = 0,838$). Серед включених факторів найбільш вагомими виявилися індекс щастя населення та індекс сприйняття корупції, які зберегли статистичну значущість і після врахування впливу інших пояснювальних змінних. Проведені діагностичні перевірки показали відсутність суттєвої

мультиколінеарності між факторами, нормальний розподіл залишків, гомоскедастичність та незалежність випадкових похибок, таким чином підтвердивши адекватність і статистичну надійність побудованої моделі.

Застосування нейронечіткого моделювання дозволило дослідити можливі нелінійні залежності між показниками. Отримані результати вказали на наявний нелінійний зв'язок між ВВП на душу населення та усіма розглянутими індексами. Таким чином, результати нейронечіткого моделювання загалом підтвердили висновки, отримані за допомогою регресійного аналізу.

Проведене дослідження підтвердило наявність тісного взаємозв'язку між економічним добробутом країн світу та їх інституційними й соціальними характеристиками. Результати аналізу таблиць спряженості, регресійного моделювання, кластерного аналізу та нейронечітких технологій виявили узгоджені закономірності, які вказали на важливу роль якості державного управління та суб'єктивного благополуччя населення у формуванні рівня економічного розвитку. Тобто, серед досліджуваних показників найбільш вагомими факторами виявилися індекс сприйняття корупції та індекс щастя населення, тоді як зв'язок із показниками демократії та злочинності виявився менш вираженим.

Слід зауважити, що тема дослідження зв'язку між економічними та соціальними показниками потребує подальших досліджень та аналізу. Подальші праці у цьому напрямку мають потенціал принести значну користь для економічного аналізу та стати у пригоді для практичного застосування в управлінських процесах.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Lee, Y.-Y., Goh, K.-L. The Happiness-Economic Well-Being Nexus: New Insights From Global Panel Data. *SAGE Open*, 2023, vol. 13, no. 4. DOI: 10.1177/21582440231199659. (date of access: 23.05.2026).
2. Abduljaber, M.F., Onder, M. When we can't see the wood for the trees: the lurking effect of sustainability on corruption. *Cogent Social Sciences*, 2024, vol. 10, no. 1. DOI: 10.1080/23311886.2024.2318859. (date of access: 23.05.2026).
3. Fajryani, A.I., Syafii, M. The Influence of GDP per Capita, Income Inequality, Population Density, and Unemployment on the Crime Index in 7 ASEAN Countries. *TALENTA Conference Series: LWSA*, 2026, vol. 9, no. 1. DOI: 10.32734/lwsa.v9i1.2731. (date of access: 23.05.2026).
4. Hogg, H.D.J., Saunders, K.E.A., Wormald, R. The clinician's guide to interpreting a regression analysis. *Eye*, 2022, vol. 36, no. 9, pp. 1714–1717. DOI: 10.1038/s41433-022-01949-z. (date of access: 23.05.2026).
5. Качан О.І., Кмитюк Т.Л., Даценко Н.С. Економетричне моделювання факторів економічного зростання на основі панельних даних. *Моделювання розвитку економічних систем*, 2025. DOI: 10.31891/mdes/2025-18-54. (дата звернення: 23.05.2026)
6. Who We Are. World Bank Group. URL: <https://www.worldbank.org/ext/en/who-we-are> (date of access: 11.05.2026).
7. Explore History. World Bank Group. URL: <https://www.worldbank.org/en/archive/history#> (date of access: 11.05.2026).
8. GDP per capita (current US\$). World Bank Group. URL: <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD> (date of access: 11.05.2026).

9. Data sharing. World Happiness Report. URL: <https://www.worldhappiness.report/data-sharing/> (date of access: 21.02.2026).
10. About Us. World Happiness Report. URL: <https://www.worldhappiness.report/about/> (date of access: 11.05.2026).
11. World Happiness Report: origins, purpose, contents, impact and future. Brockmann H., Fernandez-Urbano R. (eds). *Encyclopedia of Happiness, Quality of Life and Subjective Wellbeing*, 2024. P. 296–303. DOI: <https://doi.org/10.4337/9781800889675.00048> (date of access: 21.02.2026).
12. Stanley A. Looking Beyond GDP. International Monetary Fund. URL: <https://www.imf.org/en/publications/fandd/issues/2024/03/picture-this-looking-beyond-gdp> (date of access: 11.05.2026).
13. About. Transparency International. URL: <https://www.transparency.org/en/about> (date of access: 21.02.2026).
14. Our story. Transparency International. URL: <https://www.transparency.org/en/our-story> (date of access: 21.02.2026).
15. Corruption Perceptions Index 2024. Transparency International. URL: <https://www.transparency.org/en/cpi/2024> (date of access: 21.02.2026).
16. The ABCs of the CPI: How the Corruption Perceptions Index is calculated. Transparency International. URL: <https://www.transparency.org/en/news/how-cpi-scores-are-calculated> (date of access: 21.02.2026).
17. Christos P., Eleni K., Dimitrios K., Zacharias D., Athanasios A., Panagiotis L. Corruption Perception Index (CPI), as an Index of Economic Growth for European Countries. *Theoretical Economics Letters*, 2018, Vol. 8. P. 524–537. DOI: 10.4236/tel.2018.83037. (date of access: 21.02.2026).
18. About. *Economist Intelligence*. URL: <https://www.eiu.com/n/about/> (date of access: 23.02.2026).

19. The Democracy Index. *Economist Intelligence*. URL: <https://www.eiu.com/n/global-themes/democracy-index/> (date of access: 23.02.2026).
20. Giuseppe V. M. Democracy and economic growth: comparative perspectives. School of International Studies, University of Trento, 2015. DOI: 10.13140/RG.2.2.26240.48640 (date of access: 23.02.2026).
21. Democracy index Economist Intelligence Unit. Our world in data. URL: <https://ourworldindata.org/grapher/democracy-index-eiu> (date of access: 23.02.2026).
22. Common. Numbeo. URL: <https://www.numbeo.com/common/> (date of access: 23.02.2026).
23. About This Website. Numbeo. URL: <https://www.numbeo.com/common/about.jsp> (date of access: 23.02.2026).
24. Methodology and motivation. Numbeo. URL: https://www.numbeo.com/common/motivation_and_methodology.jsp (date of access: 23.02.2026).
25. Crime index by Country 2024. Numbeo. URL: https://www.numbeo.com/crime/rankings_by_country.jsp?title=2024 (date of access: 23.02.2026).
26. Understanding Crime Indexes. Numbeo. URL: https://www.numbeo.com/crime/indices_explained.jsp (date of access: 23.02.2026).
27. Karpavicius T., Stavytsky A., Giedraitis VR., Ulvidienė E., Kharlamova G., Kavaliauskaite B. What Determines the Crime Rate? A Macroeconomic Case Study. *Economies*, 2024, Vol. 12. P. 250. DOI: <https://doi.org/10.3390/economies12090250> (date of access: 23.02.2026).
28. Corruption Perceptions Index 2025. Transparency International. URL: <https://www.transparency.org/en/cpi/2025> (date of access: 02.03.2026).
29. Кількісні методи в економіці: навч. посіб. / Великоіваненко Г.І., Піскунова О.В., Ващаєв та ін. Київ: КНЕУ, 2024. 392 с.

30. McHugh, M.L. The Chi-square test of independence. *Biochemia Medica*, vol. 23, no. 2, 2013, pp. 143–149. DOI: 10.11613/BM.2013.018. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC3900058/> (date of access: 12.05.2026).
31. Newsom, J.T. Analysis of Contingency Tables: Course handout for Psy 525/625 Categorical Data Analysis. Portland State University, 2025. URL: https://web.pdx.edu/~newsomj/cdaclass/ho_chisq.pdf (date of access: 12.05.2026).
32. Newsom, J.T. Phi, Cramer's V, and the Contingency Coefficient: Course handout for Psy 525/625 Categorical Data Analysis. Portland State University, 2025. URL: https://web.pdx.edu/~newsomj/cdaclass/ho_phi.pdf (date of access: 12.05.2026).
33. Sheskin, D.J. Handbook of Parametric and Nonparametric Statistical Procedures. 2nd ed. Boca Raton: Chapman & Hall/CRC, 2000. URL: <https://archive.org/details/david-j.-sheskin-david-sheskin-handbook-of-para-book-fi.org> (date of access: 12.05.2026).
34. Zadeh L.A. Fuzzy Sets. *Information and Control*, vol. 8, 1965, pp. 338–353. DOI: 10.1016/S0019-9958(65)90241-X. URL: <https://www.sciencedirect.com/science/article/pii/S001999586590241X> (date of access: 13.05.2026).
35. Kozlovskiy S., Syniehub P., Kozlovskiy A., Lavrov R. Intellectual capital management of the business community based on the neuro-fuzzy hybrid system. *Neuro-Fuzzy Modeling Techniques in Economics*, no. 11, 2022, pp. 25–47. DOI: 10.33111/nfmte.2022.025. URL: <https://nfmte.kneu.ua/archive/2022/11.02> (date of access: 13.05.2026).
36. The MathWorks, Inc. Fuzzy Logic Toolbox™ User's Guide. 2025. 1154 p. URL: https://www.mathworks.com/help/pdf_doc/fuzzy/fuzzy Ug.pdf (дата звернення: 13.05.2026).
37. Castillo-Téllez, M. et al. Fuzzy Logic Concepts, Developments and Implementation. *Information*, vol. 15, no. 10, 2024, article 656. DOI:

- 10.3390/info15100656. URL: <https://www.mdpi.com/2078-2489/15/10/656> (date of access: 13.05.2026).
- 38.Jang J.S.R. ANFIS: Adaptive-Network-Based Fuzzy Inference System. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 23, no. 3, 1993, pp. 665–685. DOI: 10.1109/21.256541. URL: https://www.researchgate.net/publication/3113825_ANFIS_Adaptive-Network-based_Fuzzy_Inference_System (date of access: 13.05.2026).
- 39.Jain, A.K. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, vol. 31, no. 8, 2010, pp. 651–666. DOI: 10.1016/j.patrec.2009.09.011. URL: <https://cse.msu.edu/~cse802/Papers/DataClustering50yearsBeyondKmeans.pdf> (date of access: 25.05.2026).
- 40.Eurostat. Comparing GDP: growth rate and per capita. *Statistics Explained*. Luxembourg: Publications Office of the European Union, 2025. URL: <https://ec.europa.eu/eurostat/statistics-explained/SEPDF/cache/63213.pdf> (date of access: 25.05.2026).
- 41.Eydi, A., Mohtashami Borzadaran, G.R., Ahmadi, J. The variable selection methods and algorithms in the multiple linear model. *Communications in Statistics – Theory and Methods*, vol. 53, no. 4, 2024, pp. 1245–1265. DOI: 10.1080/03610926.2022.2027449. (date of access: 25.05.2026).

Метадані

ДОКУМЕНТ

Заголовок

Онуфрієнко_Притоманова_плагіат

Автор

Онуфрієнко

Науковий керівник / Експерт

Притоманова

ІД документа

334108013

ОРГАНІЗАЦІЯ

Назва організації

Kyiv National Economic University named after Vadym Hetman
KNEU

підрозділ

кафедра штучного інтелекту, моделювання та статистики

ЗВІТ

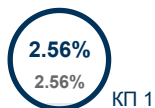
Дата звіту

5/31/2026

Дата редагування

Обсяг знайдених подібностей

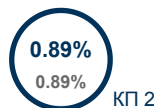
Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.



КП 1

25

Довжина фрази для коефіцієнта подібності 2



КП 2

10101

Кількість слів



КЦ

76397

Кількість символів

Тривога

У цьому розділі ви знайдете інформацію щодо текстових спотворень. Ці спотворення в тексті можуть говорити про МОЖЛИВІ маніпуляції в тексті. Спотворення в тексті можуть мати навмисний характер, але частіше характер технічних помилок при конвертації документа та його збереженні, тому ми рекомендуємо вам підходити до аналізу цього модуля відповідально. У разі виникнення запитань, просимо звертатися до нашої служби підтримки.

Заміна букв		1
Інтервали		0
Мікропробіли		0
Білі знаки		0
Парафрази (SmartMarks)		9

Джерела

Нижче наведений список джерел. В цьому списку є джерела із різних баз даних. Колір тексту означає в якому джерелі він був знайдений. Ці джерела і значення Коефіцієнту Подібності не відображають прямого плагіату. Необхідно відкрити кожне джерело і проаналізувати зміст і правильність оформлення джерела.

10 найдовших фраз

Колір тексту

#	НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
---	--	---