

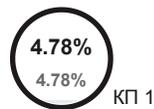
## Звіт подібності

## Метадані

Назва організації		Назва організації		
National University of Water and Environmental Engineering		National University of Water and Environmental Engineering		
Заголовок				
2025_Mishchuk_126_master.docx				
Автор		Науковий керівник / Експерт		
Міщук Вікторія Андріївна		Міщук Вікторія Андріївна		
Кількість слів	Кількість символів	Дата звіту	Дата редагування	ІД документу
13811	100322	12/22/2025	---	332948992

## 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.



13811

Кількість слів



100322

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		378
Білі знаки		0
Парафрази (SmartMarks)		33

## 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

ПОРЯДКОВИЙ НОМЕР	НАЗВА ТА АДРЕСА ДЖЕРЕЛА URL (НАЗВА БАЗИ)	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
1	<a href="https://ep3.nuwm.edu.ua/30214/1/04-05-84%D0%9C.pdf">https://ep3.nuwm.edu.ua/30214/1/04-05-84%D0%9C.pdf</a>	67 0.49 %
2	Особливості впливу факторів життя на захворюваність карієсом серед населення віком від 18 до 24 років у м. Ужгород і Ужгородському районі 2/23/2018 Kharkiv National Medical University (KNMU) (Науковий відділ ХНМУ)	20 0.14 %
3	<a href="https://echas.vnu.edu.ua/index.php/echas/article/download/509/419/">https://echas.vnu.edu.ua/index.php/echas/article/download/509/419/</a>	20 0.14 %

4	<a href="http://elar.khnu.km.ua/bitstream/123456789/9762/1/%D0%91%D0%BE%D0%BD%D0%B4%D0%B0%D1%80%D0%B5%D0%BD%D0%BA%D0%BE%20%D0%9E%D0%86%20%D0%A1%D1%82%D0%B0%D1%82%D1%82%D1%8F.pdf">http://elar.khnu.km.ua/bitstream/123456789/9762/1/%D0%91%D0%BE%D0%BD%D0%B4%D0%B0%D1%80%D0%B5%D0%BD%D0%BA%D0%BE%20%D0%9E%D0%86%20%D0%A1%D1%82%D0%B0%D1%82%D1%82%D1%8F.pdf</a>	20 0.14 %
5	<a href="https://sdc-journal.com/index.php/journal/article/download/392/315/">https://sdc-journal.com/index.php/journal/article/download/392/315/</a>	19 0.14 %
6	2025-1-Олійник_Самолук-ПП.docx 5/9/2025 National University of Water and Environmental Engineering (National University of Water and Environmental Engineering)	19 0.14 %
7	<a href="http://repository.hneu.edu.ua/bitstream/123456789/25844/1/%D0%9A%D1%80%D1%83%D1%82%D0%BE%D0%B2%D0%B0_%D0%A1%D1%82%D0%B0%D0%B2%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0_k_23_04_2021.pdf">http://repository.hneu.edu.ua/bitstream/123456789/25844/1/%D0%9A%D1%80%D1%83%D1%82%D0%BE%D0%B2%D0%B0_%D0%A1%D1%82%D0%B0%D0%B2%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0_k_23_04_2021.pdf</a>	18 0.13 %
8	<a href="https://otherreferats.allbest.ru/management/01230310_0.html">https://otherreferats.allbest.ru/management/01230310_0.html</a>	17 0.12 %
9	<a href="https://archive.mcn.org.ua/index.php/conference-proceeding/article/download/1012/1032/1045">https://archive.mcn.org.ua/index.php/conference-proceeding/article/download/1012/1032/1045</a>	17 0.12 %
10	<a href="https://sdc-journal.com/index.php/journal/article/download/392/315/">https://sdc-journal.com/index.php/journal/article/download/392/315/</a>	17 0.12 %

### з бази даних RefBooks (0.00 %)



ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
------------------	-----------	--

### з домашньої бази даних (0.24 %)



ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
------------------	-----------	--

1	2025-1-Олійник_Самолук-ПП.docx 5/9/2025 National University of Water and Environmental Engineering (National University of Water and Environmental Engineering)	19 (1) 0.14 %
2	Більчук магістерська робота 161225.docx 12/16/2025 National University of Water and Environmental Engineering (National University of Water and Environmental Engineering)	14 (2) 0.10 %

### з програми обміну базами даних (0.26 %)



ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
------------------	-----------	--

3	Особливості впливу факторів життя на захворюваність карієсом серед населення віком від 18 до 24 років у м. Ужгород і Ужгородському районі 2/23/2018 Kharkiv National Medical University (KNMU) (Науковий відділ ХНМУ)	20 (1) 0.14 %
4	Каркач-Семигіна_БЕСПЛ № 3.docx 6/26/2023 Publishing House "Helvetica" (Видавничий дім "Гельветика")	10 (1) 0.07 %
5	ФОРМУВАННЯ ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНОЇ КОМПЕТЕНТНОСТІ В УЧНІВ ПОЧАТКОВОЇ ШКОЛИ..docx 12/10/2020 Rivne State Humanities University (Півненський державний гуманітарний університет)	6 (1) 0.04 %

### з Інтернету (4.28 %)



ПОРЯДКОВИЙ НОМЕР	ДЖЕРЕЛО URL	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)
------------------	-------------	--

6	<a href="https://sdc-journal.com/index.php/journal/article/download/392/315/">https://sdc-journal.com/index.php/journal/article/download/392/315/</a>	148 (10) 1.07 %
---	---	-----------------

7	<a href="https://otherreferats.allbest.ru/management/01230310_0.html">https://otherreferats.allbest.ru/management/01230310_0.html</a>	95 (8) 0.69 %
8	<a href="https://ep3.nuwm.edu.ua/30214/1/04-05-84%D0%9C.pdf">https://ep3.nuwm.edu.ua/30214/1/04-05-84%D0%9C.pdf</a>	76 (2) 0.55 %
9	<a href="https://archive.mcmd.org.ua/index.php/conference-proceeding/article/download/1012/1032/1045">https://archive.mcmd.org.ua/index.php/conference-proceeding/article/download/1012/1032/1045</a>	58 (4) 0.42 %
10	<a href="https://revolution.allbest.ru/pedagogics/00996226_0.html">https://revolution.allbest.ru/pedagogics/00996226_0.html</a>	51 (5) 0.37 %
11	<a href="https://echas.vnu.edu.ua/index.php/echas/article/download/509/419/">https://echas.vnu.edu.ua/index.php/echas/article/download/509/419/</a>	34 (2) 0.25 %
12	<a href="http://repository.hneu.edu.ua/bitstream/123456789/25844/1/%D0%9A%D1%80%D1%83%D1%82%D0%BE%D0%B2%D0%B0_%D0%A1%D1%82%D0%B0%D0%B2%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0_k_23_04_2021.pdf">http://repository.hneu.edu.ua/bitstream/123456789/25844/1/%D0%9A%D1%80%D1%83%D1%82%D0%BE%D0%B2%D0%B0_%D0%A1%D1%82%D0%B0%D0%B2%D0%B5%D1%80%D1%81%D1%8C%D0%BA%D0%B0_k_23_04_2021.pdf</a>	30 (2) 0.22 %
13	<a href="https://repository.sspu.edu.ua/bitstreams/017bff31-37d0-47eb-9f6e-5ca74a9e0272/download">https://repository.sspu.edu.ua/bitstreams/017bff31-37d0-47eb-9f6e-5ca74a9e0272/download</a>	23 (3) 0.17 %
14	<a href="http://elar.khnu.km.ua/bitstream/123456789/9762/1/%D0%91%D0%BE%D0%BD%D0%B4%D0%B0%D1%80%D0%B5%D0%BD%D0%BA%D0%BE%20%D0%9E%D0%86%20%D0%A1%D1%82%D0%B0%D1%82%D1%82%D1%8F.pdf">http://elar.khnu.km.ua/bitstream/123456789/9762/1/%D0%91%D0%BE%D0%BD%D0%B4%D0%B0%D1%80%D0%B5%D0%BD%D0%BA%D0%BE%20%D0%9E%D0%86%20%D0%A1%D1%82%D0%B0%D1%82%D1%82%D1%8F.pdf</a>	20 (1) 0.14 %
15	<a href="http://binpo.com.ua/wp-content/uploads/2021/04/DigComp-2.0-%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0-%D1%86%D0%B8%D1%84%D1%80%D0%BE%D0%B2%D0%BE%D1%97-%D0%BA%D0%BE%D0%BC%D0%BF%D0%B5%D1%82%D0%B5%D0%BD%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%96-%D0%B3%D1%80%D0%BE%D0%BC%D0%B0%D0%B4%D1%8F%D0%BD.pdf">http://binpo.com.ua/wp-content/uploads/2021/04/DigComp-2.0-%D0%A1%D0%B8%D1%81%D1%82%D0%B5%D0%BC%D0%B0-%D1%86%D0%B8%D1%84%D1%80%D0%BE%D0%B2%D0%BE%D1%97-%D0%BA%D0%BE%D0%BC%D0%BF%D0%B5%D1%82%D0%B5%D0%BD%D1%82%D0%BD%D0%BE%D1%81%D1%82%D1%96-%D0%B3%D1%80%D0%BE%D0%BC%D0%B0%D0%B4%D1%8F%D0%BD.pdf</a>	14 (2) 0.10 %
16	<a href="https://economyandsociety.in.ua/index.php/journal/article/download/1476/1422/">https://economyandsociety.in.ua/index.php/journal/article/download/1476/1422/</a>	13 (2) 0.09 %
17	<a href="https://1library.net/document/ydlnl1gz-digital-competence-framework-citizens-european-proficiency-definition-technologies.html">https://1library.net/document/ydlnl1gz-digital-competence-framework-citizens-european-proficiency-definition-technologies.html</a>	12 (2) 0.09 %
18	<a href="https://kursovi.com.ua/files/21000677.docx">https://kursovi.com.ua/files/21000677.docx</a>	9 (1) 0.07 %
19	<a href="https://revolution.allbest.ru/economy/01239173_0.html">https://revolution.allbest.ru/economy/01239173_0.html</a>	8 (1) 0.06 %

## Список прийнятих фрагментів

ПОРЯДКОВИЙ НОМЕР	ЗМІСТ	КІЛЬКІСТЬ ОДНАКОВИХ СЛІВ (ФРАГМЕНТІВ)
------------------	-------	---------------------------------------

2  
8 **МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ** Національний університет водного господарства та природокористування Навчально-науковий інститут кібернетики, інформаційних технологій та інженерії Кафедра комп'ютерних технологій та економічної кібернетики  
Допущено до захисту: Завідувач кафедри комп'ютерних технологій та економічної кібернетики д. е. н., проф. П. М. Грицюк

« \_\_\_\_\_ » 2025\_р. **КВАЛІФІКАЦІЙНА РОБОТА на здобуття ступеня «магістр» за освітньо-професійною програмою «Інформаційні технології в бізнесі» спеціальності 126 «Інформаційні системи та технології» на тему: «Інноваційні технології для формування цифрових навичок. Аналітичне обґрунтування»**

Виконала:

8 **здобувач вищої освіти 2 курсу, групи ІТБ-61м**

Міщук Вікторія Андріївна

Керівник:

2 **к. т. н., доцент Джоші О.І.**

Рецензент:

2 **д.е.н., професор Грицюк П.М.**

Рівне - 2025

РЕФЕРАТ

Кваліфікаційна робота магістра: 61 с., 25 рис., 4 табл., 21 літературне джерело.

Актуальність теми: Умови сучасної цифрової трансформації диктують нові вимоги до людського капіталу, де цифрові навички стають не просто

додатковою перевагою, а базовою умовою професійної реалізації та економічного добробуту. Використання інноваційних технологій аналізу даних дозволяє виявити приховані закономірності у формуванні цифрового профілю громадянина, що є необхідним для розробки ефективних інформаційних систем підтримки навчання.

Об'єкт дослідження - інноваційні технології для формування цифрових навичок.

Предметом дослідження - аналітичне обґрунтування доцільності впровадження інноваційних технологій для формування цифрових навичок населення.

Метою магістерської роботи є проведення комплексного аналізу доцільності впровадження інноваційних технологій як ключового інструменту для формування цифрових навичок у сучасних умовах.

У магістерській роботі було: проаналізовано теоретичні засади цифрової трансформації та сучасні моделі цифрових компетенцій; обґрунтовано вибір інструментарію інтелектуального аналізу даних для дослідження рівня цифрової грамотності; проведено детальний аналіз цифрових навичок населення та впливу факторів на них; розроблено практичні рекомендації щодо персоналізації процесу формування цифрових навичок.

Ключові слова: Інноваційні технології, цифрові навички, аналітика, інтелектуальний аналіз даних, факторний аналіз, інформаційні системи.

## 18 МІСТ

### ТОС lo "1-3" lh lz lu ВСТУП 4

#### РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ АНАЛІТИЧНОГО ОБҐРУНТУВАННЯ 6

1.1 Теоретичний огляд та класифікація сучасних цифрових навичок 6

1.2 Методологія аналітичного обґрунтування та візуалізації даних 10

#### РОЗДІЛ 2. ІНФОРМАЦІЙНО-АНАЛІТИЧНЕ ЗАБЕЗПЕЧЕННЯ ДОСЛІДЖЕННЯ ЦИФРОВИХ НАВИЧОК 16

2.1 Дослідження програмних продуктів для аналізу даних 16

2.2 Дослідження структури даних, введення умовних позначень та попередній аналіз 21

#### РОЗДІЛ 3. РЕАЛІЗАЦІЯ АНАЛІЗУ ДАНИХ ТА РОЗРОБКА РЕКОМЕНДАЦІЙ ЩОДО ФОРМУВАННЯ ЦИФРОВИХ НАВИЧОК 34

3.1 Аналіз категорій цифрових навичок 34

3.2 Аналіз рівнів цифрових навичок 36

3.3 Дослідження впливу факторів на рівень цифрових компетентностей населення 43

3.4 Рекомендації для розробки ІС щодо формування цифрових навичок 55

ВИСНОВКИ 57

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ 59

ДОДАТКИ 62

### ВСТУП

Умови сучасної цифрової трансформації диктують нові вимоги до людського капіталу, де цифрові навички стають не просто додатковою перевагою, а базовою умовою професійної реалізації та економічного добробуту. Динамічний розвиток технологій, перехід до економіки знань та зміна форматів праці (дистанційна робота, фріланс) потребують постійного моніторингу та вдосконалення цифрових компетенцій. Попри значну кількість освітніх ресурсів, залишається гострою проблема персоналізації навчання та об'єктивного обґрунтування того, які саме навички критично впливають на дохід та кар'єрний успіх. Використання інноваційних технологій аналізу даних (Data Science, інтелектуальний аналіз) дозволяє виявити приховані закономірності у формуванні цифрового профілю громадянина, що є необхідним для розробки ефективних інформаційних систем підтримки навчання.

Метою роботи є проведення комплексного аналізу доцільності впровадження інноваційних технологій як ключового інструменту для формування цифрових навичок у сучасних умовах.

Об'єкт дослідження - інноваційні технології для формування цифрових навичок.

Предмет дослідження - аналітичне обґрунтування доцільності впровадження інноваційних технологій для формування цифрових навичок населення.

Методи дослідження включають: системний аналіз (при вивченні предметної області), методи математичної статистики та кореляційного аналізу (в середовищі Python), методи інтелектуального аналізу даних та дерева класифікації (в середовищі RStudio), а також методи візуалізації даних (Tableau та Python).

Для досягнення поставленої мети у роботі необхідно вирішити наступні основні завдання дослідження:

1. Аналіз предметної області
2. Аналіз програмних продуктів для аналітичного обґрунтування інноваційних технологій
3. Аналіз категорій цифрових навичок
4. Аналіз цифрових навичок на момент вступу в університет, випуску з університету та на сьогодні
5. Дослідження впливу факторів ( $i =$ ) на цифрові навички
6. Розробка рекомендацій для розробки ІС щодо формування цифрових навичок

Практичне значення одержаних результатів полягає у використанні отриманих результатів та сформульованих рекомендацій при розробці інформаційної системи формування цифрових навичок, що дозволяє автоматизувати процес діагностики прогалів та побудови індивідуальних освітніх траєкторій.

Логіка дослідження зумовила структуру кваліфікаційної роботи на здобуття ступеня «магістр»: вступ, три розділи, висновки, список використаних джерел та додатки.

У першому розділі аналізуються теоретичні основи цифрових навичок та сучасні моделі цифрових компетенцій, зокрема DigComp 2.2, а також методологія аналітики та візуалізації.

У другому розділі проведено огляд програмних продуктів для збору та аналізу даних, обґрунтовано вибір інструментарію (Python, R, Tableau) для проведення подальшого дослідження та проведено попередній аналіз даних.

У третьому розділі описується процес комплексного статистичного та інтелектуального аналізу цифрових навичок респондентів. Виявлено ключові фактори впливу на їхній рівень та доведено нелінійний характер зв'язку між цифровими компетенціями та факторами.

Програмні продукти та технології, що використовувались під час написання кваліфікаційної роботи: Google сервіси, Jupyter (Python), RStudio (R),

## РОЗДІЛ 1. ТЕОРЕТИЧНІ ЗАСАДИ АНАЛІТИЧНОГО ОБҐРУНТУВАННЯ

## 1.1 Теоретичний огляд та класифікація сучасних цифрових навичок

У сучасних умовах глобальної цифровізації та стрімкого технологічного прогресу, формування цифрових навичок стає не лише додатковою перевагою, але й фундаментальною передумовою життєдіяльності та розвитку людського капіталу в цілому. Динамічне впровадження інноваційних рішень у всі сфери суспільного життя - від державного управління до приватного сектору - вимагає від громадян не лише базової комп'ютерної грамотності, а й комплексного володіння цифровими компетентностями, що дозволяють ефективно адаптуватися до сучасних вимог. Тому, вважаємо важливим глибоко оцінити сутність поняття «цифрові компетентності», систематизувати підходи до їх визначення та проаналізувати існуючі світові стандарти, серед яких ключове місце посідає Європейська рамка цифрових компетентностей для громадян (DigComp 2.2) [1].

Аналіз сучасної наукової літератури дозволяє виділити декілька підходів до визначення цифрових компетентностей, що відображають різні аспекти цифровізації економіки та суспільства. Зокрема, А. Магдіч та О. Задоя розглядають цифрові навички як ключовий індикатор людського капіталу, що визначає здатність країни до інноваційного розвитку. У їхньому розумінні цифрова компетентність є інтегральною характеристикою, яка включає не лише технічне володіння ІКТ, а й здатність використовувати онлайн-сервіси для задоволення соціальних та економічних потреб [4].

Інший ракурс пропонує А. Мостова, яка пов'язує цифрові навички безпосередньо з процесами цифрової трансформації бізнесу. Автор наголошує, що цифрова компетентність - це набір критичних знань та вмінь, які дозволяють фахівцям адаптуватися до нових бізнес-моделей та ефективно взаємодіяти у межах цифрових екосистем [3]. Водночас Г. Жекало акцентує увагу на інституційному аспекті, визначаючи цифрові навички як складову «цифрового табло» держави. Згідно з цим підходом, компетентність розглядається через призму доступу до технологій та спроможності громадян використовувати цифрові публічні послуги, що є основою для побудови сучасних взаємовідносин між державою та суспільством [2].

Спільним для цих підходів є визнання того, що цифрова компетентність вийшла за межі суто технічної грамотності та стала стратегічним ресурсом, що забезпечує конкурентоспроможність особистості в умовах цифрової економіки.

Різноманітність наукових підходів підкреслює складність явища цифрових навичок. Для прикладних досліджень та розробки освітніх стратегій необхідна єдина уніфікована система показників. На європейському просторі такою еталонною моделлю стала «**Рамка цифрової компетентності для громадян Digital Competence Framework for Citizens**», відома як DigComp. Актуальна версія стандарту - DigComp 2.2 [1]. Вона інтегрує в собі теоретичні напрацювання щодо знань, умінь та ставлень особистості в цифровому середовищі та пропонує чітку дворівневу структуру, що включає 5 концептуальних сфер та 21 дескриптор конкретних компетенцій. Саме ця модель дозволяє трансформувати абстрактне поняття «цифрова грамотність» у вимірювані аналітичні показники, що становить методологічну основу нашого подальшого дослідження.

Методологія DigComp 2.2 виділяє такі ключові категорії з відповідними компетенціями [1]:

## 1. Інформаційна та дата-грамотність

15 **Перегляд, пошук і фільтрація даних, інформації та цифрового контенту:** здатність визначити потреби в інформації та здійснювати ефективний пошук у цифрових середовищах.

- Оцінювання 15 **даних, інформації та цифрового контенту:** вміння 10 **аналізувати, порівнювати та критично оцінювати достовірність і надійність джерел** інформації.

13 **Управління даними, інформацією та цифровим контентом:** навички організації, зберігання, отримання та опрацювання даних в цифрових середовищах.

13 **Комунікація та співпраця - Взаємодія за допомогою цифрових технологій:** розуміння та використання різноманітних комунікаційних засобів.

- Поширення 13 **та обмін даними за допомогою цифрових технологій:** вміння ділитися 7 **даними, інформацією та цифровим вмістом з іншими за допомогою відповідних цифрових технологій.**

- Реалізація громадянської позиції за допомогою 9 **цифрових технологій:** вміння **контактувати із суспільством, користуватися державними та приватними послугами завдяки використанню цифрових технологій.**

7 **Співпраця за допомогою цифрових технологій:** вміння **використовувати цифрові інструменти та технології для** процесів співпраці, а також для спільного створення та **спільного створення даних, ресурсів і знань.** - **Мережевий етикет:** знання «нетикету», 6 **тобто володіння правилами поведінки та етикету в цифровому середовищі.**

- **Управління цифровою ідентичністю:** вміння **створювати та управляти аккаунтами,** 7 **мати можливість захищати власну репутацію, мати справу з даними,** що створюються 10 **за допомогою кількох цифрових інструментів, середовищ і служб.** 3. **Створення цифрового контенту**

- Створення контенту в різних форматах: розробка текстових, аудіо- та відеоматеріалів, а також програм.

- Редагування та інтеграція 12 **цифрового контенту:** вміння **змінювати, покращувати, використовувати наявний цифровий контент задля створення нового контенту.**

- Авторське право і ліцензії: 12 **обізнаність щодо авторських прав та політики ліцензування відносно даних, інформації та цифрового контенту.**

- **Програмування:** вміння **писати програмний код,** розробляти 7 **послідовність зрозумілих інструкцій для обчислювальної системи для вирішення заданої проблеми або виконання конкретного завдання.**

4. **Безпека - Захист пристроїв** та безпечне підключення до мережі Інтернет. 6 **вміння захистити пристрої та контент, знання заходів безпеки, розуміння ризиків та загроз.** - **Захист персональних даних: збереження приватності** у цифровому просторі.

- Захист здоров'я і благополуччя: 6 **знання та навички для збереження свого здоров'я та інших з точки зору використання цифрових технологій.**

6 **Захист навколишнього середовища:** розуміння впливу цифрових технологій на екологію, навколишнє середовище, з точки зору їх використання, утилізації.

## 5. Вирішення проблем

- Вирішення технічних 19 **проблем:** вміння **вирішувати технічні проблеми, що виникають із** комп'ютерною технікою, програмним забезпеченням, та мережами.

- Визначення потреб та їх технологічного вирішення: вміння визначати потреби та знаходити відповідні технічні рішення, або налаштувати <sup>6</sup>

### цифрові технології до власних потреб.

- Творче використання цифрових технологій: вміння <sup>6</sup> завдяки цифровим технологіям створювати знання, процеси та продукти, індивідуально або колективно, з метою вирішення повсякденних життєвих та професійних проблем.

- Виявлення прогалів цифрової компетентності: самооцінка рівня власної цифрової компетентності, виявлення та усунення прогалів в знаннях та навиках, можливість підтримувати інших у розвитку їхньої цифрової компетентності.

Отже, деталізація сфер та дескрипторів моделі DigComp 2.2 підтверджує, що сучасна цифрова компетентність є багаторівневим утворенням, яке охоплює не лише технічну вправність, а й когнітивні, соціальні та етичні аспекти взаємодії у цифровому середовищі. Актуальність такого підходу зумовлена трансформацією ринку праці, де здатність фахівця до постійного оновлення знань, критичного оцінювання інформації та безпечної роботи з даними стає ключовим фактором його професійної стійкості. Саме уніфікована структура європейської рамки дозволяє вийти за межі теоретичних описів і перейти до об'єктивної оцінки реального стану сформованості навичок у різних професійних групах.

Таким чином, систематизація цифрових компетентностей за методологією DigComp 2.2 закладає базис для подальшого аналітичного дослідження. Вона забезпечує необхідну систему метрик для вимірювання динаміки розвитку людського капіталу в умовах прискореної цифровізації. Подальший аналіз того, як ці теоретично визначені компетенції реалізуються та трансформуються у професійній діяльності, дозволить виявити критичні прогалини та сформулювати рекомендації для їх усунення, що і становить прикладну цінність нашої роботи.

## 1.2. Методологія аналітичного обґрунтування та візуалізації даних

Для дослідження процесу формування цифрових навичок, насамперед варто визначитись з методологією на якій буде побудована майбутня робота. В роботі використано дослідницький підхід, який включає сучасні статистичні інструменти обробки інформації та поєднує у собі графічну інтерпретацію отриманих результатів. Фундаментом методичної основи виступають принципи доказовості, цілісності підходу. Таке поєднання забезпечує достовірність отриманих висновків і можливість їх практичного застосування.

Алгоритм дослідження реалізується через чітку послідовність взаємопов'язаних етапів. Початковим етапом виступає підготовка даних до аналізу, виявлення, пропущених значень, уніфікація форматів, перевірка узгодженості інформації. Для більш зручної роботи з великим масивом даних, було здійснено введення умовних позначень для назв змінних. Даний етап є особливо важливим, оскільки якість вхідних даних і вибір методів аналізу впливають на достовірність подальших результатів.

Наступним етапом виступає безпосередня робота з даними, відбувається первинне групування, використовуються методи математичної статистики (розрахунок середніх значень, медіан, відхилень) для опису поточного стану рівня цифрових навичок. Однією із складових аналітичного обґрунтування є кореляційний аналіз, що дає змогу визначити існування, характер спрямованості та інтенсивність взаємозв'язку між факторами та рівнем сформованості цифрових навичок. Використано два основні підходи кореляційного аналізу, які краще підходять до кожного елементу аналізу та дозволяють отримати цілісну картину взаємозв'язків між змінними.

Коефіцієнт кореляції Пірсона є параметричною мірою лінійного зв'язку між двома безперервними змінними і приймає значення в інтервалі від -1 до +1 [8]. Даний статистичний показник застосовується для дослідження взаємозв'язку двох змінних, виміряних у метричних шкалах на одній і тій же вибірці, та дозволяє визначити, наскільки пропорційна мінливість обох параметрів. Значення +1 означає ідеальну пряму лінійну залежність, -1 - ідеальну обернену лінійну залежність, а 0 свідчить про відсутність лінійного зв'язку. Застосування цього інструменту є оптимальним за умови нормального або близького до нормального розподілу даних та відсутності аномальних значень.

<sup>3</sup> Коефіцієнт рангової кореляції Спірмена є непараметричною мірою статистичної залежності між двома змінними та оцінює, наскільки добре можна описати відношення між ними за допомогою монотонної функції [6]. Його перевагою є те, що він базується на рангах спостережень замість їх абсолютних значень, що робить його стійким до викидів та придатним для аналізу порядкових даних. Цей інструмент є ефективним для опрацювання неметричних змінних або кількісних показників, які були попередньо ранжовані за певною ознакою.

Візуалізація даних є невід'ємною складовою аналітичного обґрунтування, оскільки забезпечує наочне представлення складних залежностей та сприяє ефективному донесенню результатів дослідження різним аудиторіям. Вона дозволяє охоплювати великі обсяги інформації, роблячи її стислою та зрозумілою. У роботі використані різні типи візуалізацій, кожен з яких має своє призначення та певну область застосування.

Стовпчикова діаграма є одним з найпоширеніших та найбільш інтуїтивно зрозумілих способів подання інформації [9]. Вона використовується для відображення кількісного порівняння значень між різними категоріями даних. У цій візуалізації кожна категорія представлена прямокутником, висота якого пропорційна значенню відповідного показника. Така структура наглядно демонструє мінімуми та максимуми у межах вибірки.

Найчастіше використовується при проведенні порівняльного аналізу та групуванні результатів за певними ознаками. Використання відповідних кольорових схем, підписів та легенд забезпечує чітке донесення інформації та уникнення неоднозначності інтерпретації. Також її використовують для візуалізації часових рядів, де кожен стовпець відповідає певному етапу розвитку явища. Це дозволяє відстежувати темпи змін та порівнювати інтенсивність процесів у різні періоди, що є важливим для оцінки зростання рівня цифрових компетентностей від моменту початку навчання до етапу професійної реалізації.

Радарна діаграма є способом візуалізації даних, який відображає значення кількох змінних у вигляді поля на системі координат, де кожна вісь виходить з центральної точки під рівним кутом. Вона дозволяє одночасно відображати від трьох і більше показників та забезпечує швидке візуальне порівняння профілів за множиною критеріїв, що особливо ефективно для виявлення сильних та слабких сторін досліджуваних об'єктів.

Методологічно важливою особливістю є необхідність стандартизації шкал вимірювання через нормалізацію до єдиної шкали (0-100 або 0-1), що забезпечує коректність порівняння та запобігає викривленню інтерпретації. Віддаленість точки від центру визначає рівень розвитку відповідної компетентності, виміряний за шкалою від початкового до просунутого рівня.

Box-plot застосовується для візуалізації статистичного розподілу даних та оцінки їхньої варіативності. Графічна модель відображає п'ять основних характеристик вибірки: мінімальне та максимальне значення, медіану, а також перший і третій квартилі. Центральний «ящик» охоплює міжквартильний діапазон, що містить 50% спостережень, а вертикальні лінії («вуса») вказують на розсіювання даних поза його межами. Нижній і верхній квартилі описують по 25% відповідно. Дана модель займає менше місця порівняно з іншими формами візуалізації і тому особливо корисна при порівнянні розподілів між багатьма групами одночасно.

Особливу увагу варто приділити побудові дерева рішення, як одному з методів машинного навчання, що використовується для класифікації даних та виявлення ієрархії факторів впливу.

Дерево рішень - це графічна концепція, що представляє послідовність рішень, заснованих на вхідних значеннях, має гілкоподібну структуру, де кожен вузол представляє певний тест або умову, а гілки вказують на можливі варіанти відповідей або наступні дії [10]. Її метою є створення прогнозуючої моделі на основі змінних і параметрів на вході, що дозволяє виявити ієрархію факторів впливу, визначаючи, які характеристики мають найбільший вплив на параметри. Візуалізація є зрозумілою, може бути використана для подальших висновків, розробки рекомендацій,

пошуку закономірностей.

Використання технік Data Mining, зокрема виявлення асоціативних зв'язків (Association Rule Mining, ARM), у соціологічних дослідженнях розкриває багатовимірні взаємозалежності, які залишаються недоступними для класичних статистичних методів. Робота з масивом даних, що містить 397 учасників опитування та 21 тематичне питання, оцінене за 8-бальною шкалою, потребує інструментів, здатних перетворювати числові показники у логічні конструкції. Класичний кореляційний підхід показує лише парні лінійні зв'язки між змінними. ARM натомість будує комплексні моделі, демонструючи, як сукупність оцінок за певними питаннями впливає на відповіді щодо інших [20].

Методологія пошуку асоціацій ґрунтується на концепції транзакційних структур даних. Кожна заповнена анкета інтерпретується як окрема транзакція, а відповіді - як набір елементів. Для опитування з 21 питанням транзакція  $T$  становить підмножину загального простору можливих відповідей  $I$ . Основною формою представлення результату є правило  $X \Rightarrow Y$ , де  $X$  (антецедент) та  $Y$  (консеквент) є наборами елементів, що не мають спільних членів ( $X \cap Y = \emptyset$ ). Антецедент ( $X$ ) - передумова або комбінація відповідей, що формує умову, і може включати одну чи декілька оцінок. Консеквент ( $Y$ ) - наслідок, що виникає за наявності передумови.

Ключовими показниками ефективності асоціативного правила є підтримка (Support), достовірність (Confidence) та ліфт (Lift). Підтримка визначає частку учасників, які одночасно задовольняють умовам антецедента та консеквента за формулою  $\text{Supp}(X \Rightarrow Y) = P(X \cup Y)$ . Достовірність (Confidence) відображає умовну ймовірність, тобто частоту, з якою учасники з ознакою  $X$  також мають ознаку  $Y$ , обчислюється як  $\text{Conf}(X \Rightarrow Y) = P(Y|X) = P(X \cap Y) / P(X)$ . Ліфт (Lift) є критичним для розмежування реальних асоціацій від випадковостей і обчислюється як  $\text{Lift}(X \Rightarrow Y) = P(Y|X) / P(Y)$ . Інтерпретація залежить від значення:  $\text{Lift} = 1$  означає незалежність елементів,  $\text{Lift} > 1$  вказує на позитивний зв'язок (наприклад,  $\text{Lift} = 2.5$  означає, що ймовірність  $Y$  за наявності  $X$  у 2.5 рази вища за загальну), а  $\text{Lift} < 1$  сигналізує про негативний зв'язок, що може вказувати на взаємовиключні установки.

Оскільки вихідні дані представлені оцінками від 1 до 8, виникає потреба в застосуванні методів числового асоціативного аналізу (Numerical Association Rule Mining, NARM) [21]. Пряме застосування класичного алгоритму Apriori до вісьми значень для кожного з 21 питань призведе до надмірної фрагментації даних, де кожне правило матиме мізерну підтримку. Для вирішення цієї проблеми застосовується дискретизація - процес розбиття безперервних або багатозначних числових шкал на інтервали (біни).

Виявлення асоціативних зв'язків у масиві дослідження відкриває можливості для типологізації учасників, прогнозування відповідей, виявлення нетипових патернів та глибшого розуміння взаємозв'язків між різними аспектами досліджуваного явища.

Застосування обраної методології забезпечує системність у дослідженні цифрових навичок. Поєднання класичного статистичного апарату з алгоритмами інтелектуального аналізу дозволяє зафіксувати поточні показники компетентностей та розкрити ієрархію факторів, що зумовлюють їхній розвиток. Використання радарних діаграм та моделей класифікації створює базу для інтерпретації результатів, наочно відображаючи профілі навичок у різних часових вимірах. Сформована методична основа є необхідною передумовою для переходу до практичної частини роботи, де на основі визначеного інструментарію буде проведено аналіз структури емпіричних даних.

## РОЗДІЛ 2. ІНФОРМАЦІЙНО-АНАЛІТИЧНЕ ЗАБЕЗПЕЧЕННЯ ДОСЛІДЖЕННЯ ЦИФРОВИХ НАВИЧОК

### 2.1 Дослідження програмних продуктів для аналізу даних

Проведення комплексного дослідження цифрових компетентностей вимагає застосування надійного інформаційно-аналітичного забезпечення, здатного ефективно обробляти великі масиви багатовимірних даних. Оскільки аналіз охоплює повний цикл - від первинної підготовки даних до виявлення прихованих закономірностей та побудови змістовних графічних результатів - критично важливим є обґрунтований вибір програмного інструментарію. Сучасна практика інтелектуального аналізу даних передбачає використання спеціалізованих мов програмування та аналітичних платформ, кожна з яких виконує специфічну роль у досягненні поставленої мети. Дослідження програмних продуктів у межах цієї роботи дозволяє виділити найбільш ефективні рішення для статистичного оцінювання та візуалізації результатів.

Ключовим інструментом для проведення статистичного аналізу в межах даного дослідження було обрано мову програмування Python. Згідно з дослідженням Дж. Рогель-Салазара, Python на сьогодні є провідною екосистемою для науки про дані (Data Science), що зумовлено його універсальністю та наявністю потужних бібліотек, таких як Pandas, NumPy та Scikit-learn, які дозволяють виконувати повний цикл обробки інформації: від підготовки даних до реалізації складних алгоритмів інтелектуального аналізу [5]. Використання середовища Jupyter Notebook забезпечує відтворюваність аналізу та можливість поєднання програмного коду з науковими візуалізаціями, що підвищує прозорість та достовірність отриманих результатів аналітичного обґрунтування.

Вибір Python як базового інструменту дослідження зумовлений низкою стратегічних переваг, проте його застосування має певні обмеження, які враховувалися під час аналізу.

До ключових переваг мови відносимо багату екосистему спеціалізованих бібліотек: наявність таких пакетів, як Pandas для маніпуляцій з даними та Matplotlib або Seaborn для побудови візуалізацій, значно спрощує процес аналітичного обґрунтування [7]. Висока читабельність та синтаксис мови дозволяє досліднику зосередитися на логіці аналізу, а не на технічних складнощах програмування. Також Python легко поєднується з іншими інструментами та форматами даних (наприклад, .xlsx чи .csv), що забезпечує гнучкість при роботі з різними джерелами.

Водночас, використання Python має певні недоліки, наприклад при роботі з дуже об'ємними датасетами структури даних Python можуть вимагати значних ресурсів оперативної пам'яті. Інша імовірна проблема це глобальне блокування інтерпретатора. Це обмеження може ускладнювати розпаралелювання обчислень на багатоядерних процесорах, хоча для завдань описової статистики та побудови статистичних моделей цей фактор не є критичним.

Незважаючи на зазначені недоліки, Python залишається найкращим вибором для сучасних аналітичних досліджень завдяки своїй модульній архітектурі. Основна потужність цієї мови реалізується за допомогою розгалуженої системи спеціалізованих бібліотек, які оптимізовані для виконання конкретних завдань обробки та візуалізації великих масивів даних. Саме наявність розвиненого інструментарію дозволяє перетворити Python на універсальну платформу для інтелектуального аналізу, де кожен етап - від структурування сирих даних до статистичної перевірки гіпотез - забезпечується окремим високоєфективним програмним модулем.

Важливою складовою аналітичного забезпечення є використання спеціалізованих бібліотек мови Python, які у сукупності формують цілісну екосистему для обробки та дослідження даних. Зокрема, бібліотека Pandas розглядається як фундаментальний засіб для маніпуляцій зі структурами даних, пропонуючи високорівневі об'єкти, такі як DataFrame. Її наукова цінність полягає у можливості інтеграції різнорідних масивів інформації, забезпеченні методів для реструктуризації даних, а також у наявності потужного апарату для попередньої обробки, що включає фільтрацію, агрегування та очищення масивів від статистичного шуму.

Обчислювальним ядром системи виступає бібліотека NumPy, яка забезпечує підтримку векторних операцій над багатовимірними масивами та матрицями. Використання NumPy дозволяє оптимізувати математичні обчислення завдяки реалізації алгоритмів, що працюють з пам'яттю на

низькому рівні, що є критично важливим для виконання високорівневих статистичних операцій [11]. Саме ця бібліотека слугує базою для більшості методів інтелектуального аналізу, забезпечуючи швидкість та точність матричних перетворень.

Візуалізація результатів дослідження реалізується за допомогою бібліотеки Matplotlib, а саме її модуля pyplot [11]. Вона є стандартом у науковому середовищі для побудови статичних графіків високої якості. Гнучкість бібліотеки дозволяє здійснювати налаштування кожного графічного елемента, що забезпечує відповідність візуалізацій академічним вимогам до подання результатів.

Для поглибленого статистичного аналізу застосовується модуль Scipy.stats, який містить широкий інструментарій для аналізу ймовірнісних розподілів та проведення параметричних і непараметричних тестів [11]. Його функціонал дозволяє проводити сувору перевірку статистичних гіпотез, що мінімізує суб'єктивізм при інтерпретації виявлених залежностей.

Доповнює графічний апарат бібліотека Seaborn, яка базується на Matplotlib, проте пропонує вищий рівень абстракції для створення складних статистичних моделей візуалізації. Seaborn оптимізована для роботи зі структурами Pandas і дозволяє відображати складні взаємозв'язки між декількома змінними, забезпечуючи естетичну чіткість та високу інформативність графічних моделей.

Отже, сформований перелік бібліотек Python дає можливість до повного циклу роботи з даними, проте, для більш глибокого та інтелектуального аналізу, вважаємо необхідним використання декількох інструментів.

Другим важливим компонентом аналітичного забезпечення є мова програмування R, яка традиційно вважається стандартом для наукових статистичних обчислень [12]. На відміну від універсального Python, R було створено статистиками для статистиків, що зумовлює наявність у цій екосистемі найбільш передових методів аналізу даних та інтелектуального моделювання. Її застосування дозволяє розширити межі дослідження через використання спеціалізованих пакетів, орієнтованих на візуалізацію ієрархічних залежностей та точну обробку даних.

Бібліотека graph.plot призначена для професійної візуалізації результатів роботи алгоритмів побудови дерев рішень [13]. Вона трансформуює складні математичні структури у зрозумілі деревоподібні схеми, де кожен вузол та гілка відображають конкретні умови та статистичні розподіли.

Використання цієї бібліотеки забезпечує високу наочність процесу прийняття рішень моделлю, дозволяючи досліднику та читачу простежити шлях формування висновку для кожної групи респондентів.

Завершальним етапом формування аналітичного забезпечення є використання платформи Tableau, яка слугує для підсумкової візуалізації результатів. На відміну від засобів програмування, цей інструмент дозволяє перетворити масиви цифр та результати складних моделей у зрозумілі графіки, що значно полегшує аналіз отриманих даних.

Основні можливості Tableau зосереджені на створенні динамічних порівняльних моделей. Інструментарій платформи дозволяє наочно зіставляти показники за різними категоріями та часовими періодами, що допомагає виявляти тенденції та закономірності у великих масивах інформації. Завдяки вбудованим механізмам обробки даних, платформа забезпечує високу точність відображення статистичних показників, що дозволяє уникнути помилок при їхньому візуальному представленні.

Крім того, Tableau пропонує широкі можливості для багатофакторного аналізу. Система дозволяє об'єднувати різні типи даних в одному графічному просторі, що допомагає бачити зв'язки між окремими показниками. Такий підхід забезпечує перетворення результатів математичних розрахунків у чітку та зрозумілу базу, яка є необхідною для формування якісних висновків дослідження.

Цілі застосування обраних для дослідження програмних продуктів наведено в табл.2.1.

Таблиця 2.1

Цілі застосування програмних продуктів

No	Назва ПП	Ціль застосування
1	Google сервіси	Використано як інструмент для проведення онлайн-опитування та збору первинних даних від 397 респондентів за допомогою Google Forms; Google Sheets застосовано для автоматичної акумуляції відповідей, їхнього попереднього перегляду та експорту у формат .xlsx для подальшої обробки.
2	Python (Jupyter Notebook)	Мова програмування Python використана як основний інструмент для повного циклу аналітики: від первинної обробки та очищення масивів даних до перевірки статистичних гіпотез. Також застосовується для побудови складних візуалізацій.
3	R (RStudio)	Використано алгоритми інтелектуального аналізу даних для виявлення прихованих закономірностей та визначення пріоритетності факторів, що впливають на професійну успішність.
4	Tableau	Інструмент використано для побудови інтерактивних візуалізацій.

Таким чином, проведений аналіз програмних продуктів дозволив сформувати набір технологій для аналітичного дослідження формування цифрових навичок.

## 2.2 Дослідження структури даних, введення умовних позначень та попередній аналіз

Ефективність аналітичного обґрунтування безпосередньо залежить від якості підготовки вхідної інформації та розуміння її внутрішньої структури. У межах цього підрозділу здійснюється перехід від збору первинних даних до формування впорядкованого масиву, придатного для статистичної обробки та інтелектуального аналізу.

Першочерговим завданням при дослідженні структури даних є класифікація змінних за їхньою природою, що визначає подальший математичний апарат роботи (рис.2.1).

Якісні (категоріальні) дані використовуються для опису ознак об'єктів, які не мають прямого числового вираження, таких як стать, рівень освіти чи сфера професійної діяльності. У межах нашого аналізу такі дані дозволяють групувати респондентів та виявляти специфічні особливості цифрового розвитку окремих соціальних категорій. Оскільки якісні змінні зазвичай представлені у текстовому форматі, вони потребують попередньої трансформації для коректного сприйняття алгоритмами інтелектуального аналізу.

Кількісні дані відображають міру або величину досліджуваних явищ і дозволяють проводити безпосередні арифметичні та статистичні операції. Робота з кількісними показниками дає можливість розраховувати середні значення, аналізувати варіативність відповідей та будувати математичні моделі динаміки навичок. Саме поєднання цих двох типів даних у межах одного масиву дозволяє отримати комплексну картину впливу зовнішніх чинників на рівень цифрової грамотності.

Рис. 2.1. Типи даних [18]

Емпіричну основу дослідження складає масив первинних даних, отриманий шляхом опитування респондентів, що представляють різні сфери економічної діяльності (охорона здоров'я, будівництво, транспорт, освіта тощо). Загальний обсяг вибірки становить 397 спостережень, що дозволяє забезпечити репрезентативність результатів та високу точність при застосуванні методів інтелектуального аналізу.

Дані для дослідження зібрані в межах науково-дослідної роботи на замовлення МОН України **«Вища освіта в умовах війни та повоєнного відновлення: детермінанти розвитку задля подолання загроз відтворенню людського капіталу»** (реєстраційний номер: 0124U000351) в червні - липні 2024 р.

Структурно сформований масив містить 76 змінних, які для зручності подальшої обробки були розподілені на два функціональні блоки:

- 1) Блок результативних показників (Блок X): охоплює 63 змінні, що відображають самооцінку рівня володіння 21 цифровою компетенцією згідно з моделлю DigComp 2.2 [1]. Унікальність цього блоку полягає у фіксації даних у трьох референтних часових точках: рівень навичок на момент вступу до закладу вищої освіти, рівень при випуску та актуальний стан на робочому місці. Кожна компетенція оцінювалася за 8-бальною шкалою, що дає змогу детально простежити динаміку професійного зростання.
- 2) Блок контекстуальних характеристик (Блок Y): включає 14 змінних, які описують соціально-економічний та професійний профіль респондента. Сюди належать такі параметри, як стать, вік, рівень доходу, сфера діяльності, розмір підприємства, показники кар'єрного успіху та інші. Поєднання цих характеристик із даними блоку X дозволяє виявити фактори, що найбільше корелюють із рівнем цифрової грамотності фахівця. Наступним етапом підготовки масиву стало введення умовних позначень змінних, що є важливим для автоматизації розрахунків та запобігання системним помилкам при написанні програмного коду. Оскільки первинні дані містили розгорнуті текстові запитання з анкети, їх було замінено на лаконічні ідентифікатори, що відображають структуру дослідження (Таблиця 2.2).

Таблиця 2.2

Умовні позначення змінних блоку X

Позначення Вихідна назва

- | Позначення | Вихідна назва  |
|------------|--|
| x10        | <b>Перегляд, пошук і фільтрація даних, інформації та цифрового контенту</b>  |
| x2         | <b>Оцінювання даних, інформації та цифрового контенту</b> (вміння <b>аналізувати, порівнювати та критично оцінювати достовірність і надійність джерел даних, інформації та цифрового контенту</b> )  |
| x3         | <b>Управління даними, інформацією та цифровим контентом</b> (навички організації, зберігання, отримання та опрацювання даних в цифрових середовищах)   |
| x4         | <b>Засвідчення за допомогою цифрових технологій</b> (вміння використовувати та розуміти цифрові комунікаційні засоби)  |
| x5         | <b>Поширення та обмін даними за допомогою цифрових технологій</b> (вміння ділитися <b>даними, інформацією та цифровим вмістом з іншими за допомогою відповідних цифрових технологій</b> )  |
| x6         | Реалізація громадянської позиції <b>за допомогою цифрових технологій</b> (вміння <b>контактувати із суспільством, користуватися державними та приватними послугами завдяки використанню цифрових технологій</b> )                          |
| x7         | <b>Співпраця за допомогою цифрових технологій</b> (вміння <b>використовувати цифрові інструменти та технології для</b> процесів співпраці, а також для спільного створення та спільного створення даних, ресурсів і знань)                 |
| x8         | Мережевий етикет (знання «нетикету»), <b>якщо володіння правилами поведінки та етикету в цифровому середовищі</b>  |
| x9         | <b>Управління цифровою ідентичністю</b> (вміння <b>створювати та управляти аккаунтами, мати можливість захищати власну репутацію, мати справу з даними, що створюються за допомогою кількох цифрових інструментів, середовищ і служб</b> ) |
| x10        | <b>Створення цифрового контенту</b> в різних форматах (наприклад, аудіо, зображення, текст, відео, програми)   |
| x11        | Редагування <b>та інтеграція цифрового контенту</b> (вміння <b>змінювати, покращувати, використовувати наявний цифровий контент задля створення нового контенту</b> )  |
| x12        | Авторське право і ліцензії <b>обізнаність щодо авторських прав та політики ліцензування відносно даних, інформації та цифрового контенту</b> )   |
| x13        | Програмування (вміння писати програмний код, розробляти <b>послідовність зрозумілих інструкцій для обчислювальної системи для вирішення заданої проблеми або виконання конкретного завдання</b> )  |
| x14        | <b>Захист пристроїв та безпечне підключення до мережі Інтернет</b> (вміння <b>захистити пристрої та контент, знання заходів безпеки, розуміння ризиків та загроз</b> )   |
| x15        | <b>Захист персональних даних та приватності</b>  |
| x16        | Захист здоров'я і благополуччя <b>знання та навички для збереження свого здоров'я та інших з точки зору використання цифрових технологій</b>   |
| x17        | <b>Захист навколишнього середовища</b> (розуміння впливу цифрових технологій на екологію, навколишнє середовище, з точки зору їх використання, утилізації)   |
| x18        | Вирішення технічних <b>проблем</b> (вміння <b>вирішувати технічні проблеми, що виникають із комп'ютерною технікою, програмним забезпеченням, та мережами</b> )   |
| x19        | Визначення потреб та їх технологічного вирішення (вміння <b>визначати потреби та знаходити відповідні технічні рішення, або налаштувати цифрові технології до власних потреб</b> )   |
| x20        | Творче використання цифрових технологій (вміння <b>завдяки цифровим технологіям створювати знання, процеси та продукти, індивідуально або колективно, з метою вирішення повсякденних життєвих та професійних проблем</b> )                 |
| x21        | Виявлення прогалів цифрової компетентності (самооцінка рівня власної цифрової компетентності, виявлення та усунення прогалів в знаннях та навиках, можливість підтримувати інших у розвитку їхньої цифрової компетентності)                |

Для блоку цифрових компетентностей було впроваджено систему індексів xp\_t, де p відповідає номеру компетенції за моделлю DigComp 2.2 (від 1 до 21), а t вказує на часовий період (0 - вступ до ЗВО, 1 - випуск, 2 - поточний рівень). Наприклад, змінна x1\_0 ідентифікує навичку "Перегляд, пошук і фільтрація даних" на момент початку навчання. Такий підхід дозволяє легко оперувати масивами даних при побудові порівняльних графіків та розрахунку динаміки зростання показників.

Паралельно було проведено введення умовних позначень для контекстуальних характеристик за допомогою індексів up (Таблиця 2.3). Це дозволило уніфікувати назви таких параметрів, як соціально-демографічні ознаки респондентів, умови їхньої професійної діяльності та суб'єктивні показники кар'єрного успіху. Сформована система умовних позначень забезпечує прозорість структури даних та спрощує процес звернення до конкретних стовпців масиву під час інтелектуального аналізу.

Таблиця 2.3

Умовні позначення змінних блоку Y

Позначення Вихідна назва

- | Позначення | Вихідна назва  |
|------------|--|
| u6         | Оберіть з наведеного переліку навички, які, на Ваш погляд, є найважливішими у Вашій діяльності сьогодні чи дають змогу відчувати |

впевненість у подальшому професійному просуванні (оберіть не більше 10 варіантів відповідей)

y7 Якщо в наведеному переліку є цифрові навички, які Ви вважали дуже важливими раніше (на етапі вибору професії), а пізніше переконались, що насправді вони не визначальні для вашої професії, або за потреби їх легко здобути з невеликими витратами часу та коштів, відмітьте їх, будь ласка (виберіть будь-яку кількість варіантів відповіді):

y8 Самооцінка Вашої успішності у кар'єрі, враховуючи, що 1 - професійні амбіції на даний час реалізовані мінімально або не реалізовані зовсім, 5 - реалізовані повністю:

y9 Як би Ви оцінили посаду, яку зараз обіймаєте:

y10 Орієнтовний рівень Вашого щомісячного доходу:

y11 Галузь знань, за якою Ви навчались в університеті:

y12 Чи працюєте у сфері діяльності відповідно до освіти, здобутої в університеті:

y13 У якій сфері Ви працюєте зараз:

y14 Розмір підприємства, де Ви працюєте

y15 Скільки років минуло від завершення навчання:

y16 Вкажіть область в якій Ви проживаєте

y17 Місце Вашого проживання

y18 Ваша стать

Для проведення глибокого аналізу та забезпечення сумісності масиву з алгоритмами машинного навчання було здійснено трансформацію якісних ознак у числові за допомогою методу Label Encoding. Цей етап є обов'язковим, оскільки більшість сучасних аналітичних модулів орієнтовані на роботу з числовими векторами, а не текстовими категоріями.

У межах нашого дослідження така заміна дозволила не лише уніфікувати формат даних, а й зберегти ієрархічну логіку в межах окремих категорій. Зокрема, для змінної y10 (розмір доходу) було впроваджено ранжування від 1 до 6, де найменше значення відповідає мінімальному доходу респондентів, а найбільше - рівню понад 30 000 грн. Аналогічний підхід застосовано до параметрів y14 (масштаб підприємства) та y17 (тип населеного пункту), де кожній категорії присвоєно відповідну вагу залежно від її обсягу чи значущості.

Забезпечення достовірності та обґрунтованості результатів починається ще на етапі попереднього аналізу та підготовки даних. При попаданні неякісних (аномальних даних) в масив, може відбутись викривлення отриманих результатів що, як наслідок, може призвести до хибних висновків.

У використаному масиві даних, респонденти дали відповідь, на питання, обмежені 8-бальною шкалою, через це значення не виходитимуть за цю межу. У даному випадку аномальні значення проявляються через нетипові комбінації відповідей. Значення, які мають значне відхилення від норми це не про погані дані, перш за все це говорить про те, що їм потрібно приділити більше уваги. Це може бути як нетиповий випадок, так і свідчити про певні проблеми зі збором інформації [15].

Insufficient Effort Responding (IER) - це явище в опитуваннях та дослідженнях, коли респонденти відповідають на запитання неухважно, поспіхом або недбало, замість того, щоб ретельно читати та обмірковувати відповіді, що знижує якість та достовірність даних. Це може проявлятися у вигляді випадкових відповідей, нехтуванням інструкцій, що спотворює результати. Причинами цьому може виступати неуважність, відсутність зацікавленості, втома, навмисне викривлення відповідей або величина опитування [16].

Виявлення сумнівних даних отриманих шляхом проведення анкетування, реалізується через використання одновимірної та багатовимірної аналізу. Одновимірний підхід орієнтований на вивчення кожної окремої змінної, а багатовимірний - враховує взаємозв'язки між відповідями. В одновимірному аналізі найпоширенішими є статистичні критерії відхилення, метод міжквартильного розмаху (IQR) та робастні (стійкі до збурень та помилок у даних) метрики, зокрема медіанне абсолютне відхилення (MAD). Представлені методи дозволяють ідентифікувати екстремальні значення, що значно відхиляються від типового розподілу відповідей [19]. Однак, у контексті опитувань такі методи мають обмежену ефективність, оскільки аномалія часто криється не в самій оцінці, а в патернах відповідей респондента. Наприклад, якщо дані мають сильний від'ємний скос (більшість відповідей - 6, 7, 8), то оцінка "1" буде ідентифікована як аномальна.

Відстань Махаланобіса (2.1), - багатовимірний метод, - враховує не лише відхилення від середніх значень, але й кореляційну структуру між усіма питаннями [14]. Це дозволяє ідентифікувати випадки, коли респондент порушує типові зв'язки між конструктами, навіть якщо окремі відповіді знаходяться в межах норми.

, (2.1)

де  $x$  - вектор відповідей респондента,

- вектор вибірових середніх,

- обернена коваріаційна матриця.

Особливе значення для виявлення хибних відповідей має аналіз внутрішньоіндивідуальної варіативності (IRV). Надто низький IRV вказує на прямолінійність (straightlining), коли респондент дає однакові або майже однакові відповіді на всі питання. Натомість, надто високий IRV може свідчити про випадкові відповіді, коли хаотично чергуються оцінки без вдумливого читання запитань [17].

У дослідженні застосовано метод виявлення straightlining, оскільки він дає можливість відкинути найочевидніші недбаді відповіді, у яких повна відсутність варіативності в оцінках. Для виявлення таких аномалій була використана функція apply в RStudio (цифра 1 означає, що операція проводиться над рядками, а цифра 2 - над стовпчиками). Критерієм виключення було наявність абсолютно однакових значень у всіх відповідях (рис. 2.2).

Рис. 2.2. Результат виконання функції apply в RStudio

У результаті перевірки первинного масиву даних (n=397) було виявлено 34 респонденти (8,56% від загальної вибірки), які давали однакову відповідь на всі питання. Детальний аналіз структури виключених відповідей представлено в таблиці 2.4.

Таблиця 2.4

Структура респондентів, які надали однакові відповіді

Усі відповіді оцінені однаково      Кількість респондентів, осіб

"1" 1

"2"	1
"3"	9
"4"	2
"8"	21
Разом	34

Попри виявлення певної частки респондентів із нульовою внутрішньоіндивідуальною варіативністю (straightlining), у подальшому дослідженні було прийнято рішення використовувати повний масив даних (n=397). Такий підхід зумовлений прагненням забезпечити цілісність вибірки та врахувати специфіку самооцінювання в умовах професійної діяльності. Виключення цих відповідей могло б призвести до втрати інформації про крайні групи користувачів: тих, хто об'єктивно оцінює свої навички як початкові (відповіді "1") або, навпаки, впевнений у своєму досконалому володінні технологіями (відповіді "8"). Збереження повної вибірки дозволяє зберегти реальну варіативність думок та забезпечує більш високу точність при побудові дерев рішень, де кожна модель поведінки, навіть прямолінійна, може відображати специфічну групу фахівців у межах загального ринку праці.

Окрім виявлення аномальних патернів у відповідях, критично важливим аспектом підготовки масиву є перевірка його повноти. Наявність пропущених значень може призвести до системних помилок при розрахунку статистичних метрик або некоректної роботи алгоритмів машинного навчання, що вимагає проведення процедури верифікації кожної змінної перед початком аналізу.

Для реалізації цього завдання було використано інструментарій мови програмування Python, зокрема бібліотеку Pandas, яка дозволяє автоматизувати пошук пустих клітинок у великих масивах даних. За допомогою програмного коду було проведено суцільну перевірку всіх 76 змінних для кожного з 397 респондентів. Результати виконання скрипту підтвердили високу якість зібраної інформації: пропущені значення у масиві відсутні (рис. 2.3), що свідчить про ретельність заповнення анкет та дозволяє використовувати повну вибірку без додаткових маніпуляцій щодо заповнення чи видалення даних.

Рис. 2.3. Пошук пропущених значень

Після завершення етапів введення умовних позначень та верифікації цілісності даних, було проведено первинний дескриптивний аналіз за допомогою інструментарію Python. Це дозволило отримати узагальнену статистичну картину за всіма 21 компетенціями для поточного часового зрізу (станом "на сьогодні"). Використання функції data.describe() забезпечило розрахунок ключових метрик: середніх значень (mean), стандартного відхилення (std), мінімальних та максимальних оцінок, а також кватильного розподілу (25%, 50%, 75%). Аналіз отриманої таблиці (рис. 2.4) свідчить про високу щільність даних. Зокрема, показник count для всіх змінних становить 397.00, що ще раз підтверджує відсутність пропусків. Середні значення (mean) за більшістю компетенцій коливаються в межах від 5.38 до 6.49 балів за 8-бальною шкалою, що вказує на відносно високий рівень впевненості респондентів у власних цифрових навичках. Водночас значення стандартного відхилення близько 2 балів сигналізує про помірну варіативність відповідей, що створює необхідні умови для подальшого виявлення закономірностей за допомогою методів кореляції та класифікації.

Рис. 2.4. Статистичний опис даних для поточного часового зрізу

Для глибокого розуміння чинників, що впливають на рівень цифрової грамотності, необхідно проаналізувати склад респондентів за ключовими ознаками. Вибірка у 397 осіб є неоднорідною, що дозволяє дослідити цифрові навички крізь призму різних соціальних та економічних груп. Для оптимізації об'єму та забезпечення наочності характеристики респондентів згруповано у комбіновані рисунки за логічними ознаками.

Першим блоком аналізу є демографічний та інфраструктурний профіль учасників дослідження (рис. 2.5). Розподіл за статтю демонструє переважну частку жінок у вибірці (57,4%), що дозволяє враховувати гендерні особливості сприйняття технологій. Водночас структура за місцем проживання вказує на переважання мешканців середніх та великих міст (36,3% та 25,4% відповідно), що безпосередньо має зв'язок з рівнем доступу до цифрової інфраструктури та якісної освіти (Програмні коди для побудови діаграм представлені в додатку А).

Рис. 2.5. Розподіл респондентів за статтю та місцем проживання

Економічний контекст життєдіяльності респондентів охарактеризовано через параметри рівня доходу та масштабу підприємства (рис. 2.6). Зіставлення цих показників дозволяє оцінити ресурсні можливості фахівців. Більшість опитаних зосереджена у категорії доходу більше 30 тис. грн., при цьому значна частина вибірки репрезентує середні та великі підприємства. Така структура підтверджує, що результати дослідження відображають стан навичок найбільш активної та затребуваної частини трудових ресурсів.

Рис. 2.6. Розподіл респондентів за рівнем доходу та розміром підприємства

Заключним етапом опису вибірки є аналіз професійного статусу та кар'єрної самоідентифікації (рис. 2.7). Дані щодо займаних посад свідчать про те, що вибірка охоплює як фахівців початкового рівня, так і керівну ланку. Важливим доповненням є самооцінка успішності в кар'єрі, де більшість респондентів схиляється до середніх/високих значень. Це створює аналітичне підґрунтя для перевірки гіпотези про те, чи є високий рівень цифрових навичок ключовим чинником суб'єктивного відчуття професійної реалізації.

Рис. 2.7. Розподіл респондентів за займаною посадою та самооцінкою успішності в кар'єрі

Таким чином, проведений комплекс підготовчих заходів дозволив трансформувати первинну інформацію у впорядкований та верифікований масив даних. Перевірка цілісності в середовищі Python підтвердила відсутність пропусків, а аналіз внутрішньоіндивідуальної варіативності в RStudio дозволив оцінити якість відповідей респондентів. Сформований соціально-демографічний та професійний портрет вибірки свідчить про її репрезентативність для дослідження цифрових компетентностей у динаміці. Завдяки впровадженню системи умовних позначень та числових рангів, масив повністю адаптований для застосування статистичних методів і алгоритмів машинного навчання, що створює необхідне підґрунтя для переходу до інтерпретації результатів та виявлення факторів впливу на рівень цифрової грамотності фахівців.

### 3.1. Аналіз категорій цифрових навичок

Першим етапом аналітичного дослідження є оцінка рівня сформованості цифрових компетентностей респондентів у межах п'яти концептуальних сфер моделі DigComp 2.2. Оскільки кожна сфера включає декілька дескрипторів, для отримання узагальнених показників було проведено агрегацію даних за групами: g1 (Інформаційна грамотність), g2 (Комунікація та співпраця), g3 (Створення контенту), g4 (Безпека) та g5 (Вирішення проблем).

Технічна реалізація аналізу базується на використанні бібліотек NumPy та Matplotlib мови програмування Python. Алгоритм передбачав розрахунок ієрархічних середніх значень: на першому рівні обчислювався середній бал для кожної з 21 змінної, на другому - визначався підсумковий інтегральний показник для всієї категорії. Для візуалізації «цифрового профілю» фахівців було обрано радарну діаграму (рис. 3.1), побудовану в полярній системі координат. Це дозволило наочно зіставити п'ять вимірів компетентностей одночасно у трьох часових точках: при вступі до ЗВО ( $t=0$ ), при випуску ( $t=1$ ) та на поточному етапі професійної діяльності ( $t=2$ ).

Аналіз отриманої візуалізації дозволяє констатувати явище «експансії цифрового профілю». На початковому етапі (вступ до університету) контур діаграми є найбільш стиснутим, що відображає базовий рівень підготовки. Найнижчі показники зафіксовані у сферах створення контенту (g3) та вирішення проблем (g5), що свідчить про те, що до початку фахової підготовки цифрові навички молоді мають переважно стихійний та споживчий характер.

Етап випуску з університету характеризується помітним розширенням площі діаграми за всіма векторами. Це підтверджує гіпотезу про критичну роль вищої освіти у систематизації цифрових знань. Особливо виражене зростання спостерігається у групі g1 (Інформаційна грамотність), що пояснюється інтенсивною науково-дослідною роботою студентів, необхідністю здійснювати критичний аналіз великих обсягів інформації при написанні кваліфікаційних робіт, а також формуванням навичок академічної доброчесності та управління цифровим контентом. Таким чином, академічне середовище трансформує стихійне споживання інформації у професійну здатність оперувати даними як стратегічним ресурсом.

Рис. 3.1. Динаміка середніх значень цифрових навичок за категоріями (радарна діаграма)

Найбільший інтерес представляє зовнішній контур діаграми («Сьогодні»), який демонструє стан компетентностей у реальних ринкових умовах. Подальше зміщення ліній від центру вказує на безперервний розвиток навичок вже після завершення формальної освіти. Тобто, професійна діяльність змушує фахівців не просто використовувати технології, а критично підходити до вибору інструментів та творчо вирішувати технічні завдання.

Таким чином, радарна діаграма підтверджує, що формування цифрового капіталу є циклічним процесом, де кожен наступний етап життєдіяльності значно розширює функціональні можливості особистості у цифровому просторі.

### 3.2. Аналіз рівнів цифрових навичок

Після проведення інтегральної оцінки за основними сферами моделі DigComp 2.2 у попередньому підрозділі, виникає наукова необхідність у здійсненні більш детального, покомпонентного аналізу. Такий підхід дозволяє ідентифікувати конкретні сильні сторони та критичні прогалини у цифровій підготовці фахівців, які можуть залишатися непомітними при аналізі узагальнених середніх показників груп. Аналіз рівнів цифрових навичок присвячений дослідженню кожної з 21 компетентності окремо, що забезпечує високу точність та доказовість отриманих висновків. Для реалізації цього завдання та формування бази для поглибленої візуалізації було розроблено алгоритм обробки даних з наступною його реалізацією у середовищі Python (рис. 3.2). Метою цього етапу стала трансформація первинного масиву спостережень у компактну аналітичну таблицю середніх значень. Процес підготовки включав такі кроки:

1. Сегментація за часовими ознаками: за допомогою програмних фільтрів було сформовано три групи змінних, що відповідають рівням навичок на момент вступу ( $t=0$ ), випуску ( $t=1$ ) та поточному стану ( $t=2$ ).
2. Статистична агрегація: для кожної з 21 компетентності було розраховано середнє арифметичне значення у межах кожного часового зрізу. Це дозволило перевести аналіз із рівня окремих анкет на рівень середньостатистичного профілю вибірки.
3. Структурування результатів: отримані показники інтегровано у зведену таблицю (DataFrame), де кожен рядок відповідає конкретному індикатору моделі DigComp 2.2, а стовпці відображають динаміку його змін.

Кінцевим продуктом цього етапу став файл `summary_means.xlsx`, який виступив джерелом даних для платформи Tableau. Такий методологічний перехід дозволив поєднати обчислювальну потужність Python із можливостями високорівневої візуалізації Tableau, забезпечуючи високу якість графічного представлення результатів аналізу.

Рис. 3.2. Формування датафрейму в середовищі Python для роботи з Tableau

Для проведення детального аналізу актуального стану цифрової грамотності респондентів у середовищі Tableau було побудовано стовпчикову діаграму середніх значень за кожним із 21 індикатора моделі DigComp 2.2 (рис. 3.3). Ця візуалізація дозволяє ідентифікувати ієрархію цифрових компетентностей та виокремити ті навички, які є найбільш та найменш розвиненими у фахівців на поточному етапі.

Рис. 3.3 Середнє значення показників на сьогодні

Отже, щодо середнього значення показників станом на сьогодні, більшість демонструють високі значення, які перевищують загальний середній рівень (5.86). Це свідчить про загалом сильну цифрову компетентність за більшістю напрямків.

Наприклад, такі важливі навички, як x1 (Перегляд, пошук, фільтрація даних) та x5 (Поширення та обмін даними), отримали одні з найвищих оцінок (6.49 та 6.45 відповідно). Це підтверджує, що в умовах сучасного інформаційного суспільства базові навички роботи з даними та мережевої комунікації є найбільш відпрацьованими та затребуваними.

Також варто звернути увагу на показник x13 (Програмування), який має найнижче середнє значення (3.72) і значно відхиляється від середньої лінії вниз. Це чітко вказує на найбільшу прогалину в поточній цифровій компетентності та необхідність посилення навчання та практики в цій сфері.

Такі показники, як x14 (Захист пристроїв), x15 (Захист персональних даних), x16 (Захист здоров'я) та інші індикатори сфери безпеки, знаходяться ближче до середньої лінії. Це також області, де можна було б посилити знання для досягнення найвищих результатів. Отримані дані

підкреслюють пріоритетність подальшого розвитку критичних навичок кібергігієни та програмування як стратегічних елементів професійної конкурентоспроможності.

Для об'єктивного вимірювання внеску системи вищої освіти у розвиток людського капіталу було проведено порівняльний аналіз стану цифрових компетентностей у двох критичних точках: на момент вступу ( $t=0$ ) та на момент завершення навчання в університеті ( $t=1$ ). Така постановка задачі дозволяє ізолювати академічний вплив від подальшого професійного досвіду та оцінити, наскільки успішно освітні програми адаптують студентів до вимог цифрової економіки. У середовищі Tableau для цього було побудовано діаграму з подвійними стовпцями, що наочно демонструє амплітуду змін за кожним дескриптором.

Аналіз візуалізації, представленої на рисунку 3.4, підтверджує стійке зростання рівня компетентності за абсолютно всіма показниками. Зокрема, наочно видно суттєвий прогрес за такими індикаторами, як  $x_1$  (Перегляд, пошук, фільтрація даних), де рівень зріс із 4.42 до 5.86, та  $x_2$  (Оцінювання даних) - з 4.05 до 5.66. Важливо зазначити, що навіть складні технічні навички, як-от  $x_{13}$  (Програмування), продемонстрували позитивну динаміку, піднявшись із 2.37 до 3.24 бала.

Рис. 3.4 Середнє значення показників на момент вступу та випуску з університету

Суттєвий приріст також зафіксовано за показниками  $x_5$  (Поширення та обмін даними) та  $x_{10}$  (Створення цифрового контенту), що вказує на активне використання цифрових інструментів в академічному середовищі. Отже, університет успішно формує цифрові навички, переводячи студентів із базового рівня на високий рівень компетентності до моменту їхнього випуску. Отже, освітній процес виступає ключовим етапом систематизації знань, створюючи необхідний фундамент для професійної самореалізації.

Наступним логічним етапом дослідження стало порівняння середніх показників цифрової грамотності, зафіксованих на момент завершення навчання у закладі вищої освіти, із їхнім актуальним рівнем у теперішній час. Отримані результати, відображені на рисунку 3.5, демонструють позитивну динаміку за всіма без винятку індикаторами, що свідчить про безперервність процесу вдосконалення людського капіталу в умовах реального сектору економіки. Така тенденція підтверджує спроможність випускників до самостійного засвоєння нових знань та успішного пристосування до вимог сучасного ринку праці.

Рис. 3.5 Середнє значення показників на момент випуску з університету та на сьогодні

Детальний аналіз візуалізованих даних дозволяє виділити кілька важливих закономірностей. Зокрема, найбільш виражене зростання спостерігається у сфері пошуку та управління даними, а також у питаннях інформаційної безпеки, що пояснюється щоденною практичною необхідністю захисту пристроїв та персональних відомостей у професійній діяльності. Водночас навички, пов'язані з розробкою програмного коду, хоча і демонструють приріст, залишаються на відносно низькому рівні, що вказує на вузьку спеціалізацію цієї компетенції. Отже, якщо університетське середовище формує цілісний теоретичний фундамент, то подальша професійна практика стає майданчиком для відшліфування прикладних умінь. Стабільне підвищення оцінок за всіма дескрипторами є свідченням того, що респонденти не лише зберігають здобутий рівень підготовки, а й активно нарощують свій інтелектуальний потенціал під впливом технологічних змін.

Для отримання комплексної оцінки прогресу цифрових навичок було проведено узагальнюючий аналіз за всіма часовими інтервалами. Зведені дані, представлені на рисунку 3.6 у вигляді лінійного графіка, наочно демонструють ієрархію зростання компетентностей: від базового рівня при вступі до університету до фахового рівня у поточній професійній діяльності. Така візуалізація дозволяє простежити «життєвий цикл» кожної окремої навички та оцінити стійкість позитивної динаміки.

Рис. 3.6 Середнє значення показників станом на всі періоди

Доповнюючи загальну картину, деталізовані радарні графіки (рис. 3.7) дозволяють оцінити збалансованість цифрового профілю в кожній точці. На момент вступу ( $t=0$ ) спостерігається значна нерівномірність: відносно високі показники з пошуку інформації ( $x_1$ ) супроводжуються критично низькими балами з програмування ( $x_{13}$ ) та цифрової безпеки ( $x_{14}$ ). На момент випуску ( $t=1$ ) та у поточному стані ( $t=2$ ) площа «цифрового дзеркала» суттєво розширюється, набуваючи більш правильної форми. Це свідчить про гармонізацію навичок та перетворення окремих умінь на цілісну систему компетентностей. Особливо показовим є те, що на сьогодні більшість показників перевищили позначку у 6 балів, що за шкалою DigComp 2.2 відповідає високому рівню володіння технологіями.

Рис. 3.7. Деталізовані радарні графіки

Завершальним етапом аналізу часової динаміки стало зіставлення інтенсивності розвитку цифрових компетентностей у межах академічного середовища та в процесі подальшої професійної практики. Для цього було побудовано діаграму порівняння приростів (рис. 3.8), яка дозволяє чітко розмежувати внесок системи вищої освіти та досвіду, здобутого на робочих місцях. Результати візуалізації неспростовно доводять, що саме період навчання в університеті забезпечує найбільш вагомий розвиток цифрового потенціалу особистості.

Рис. 3.8. Діаграма порівняння приростів

Особливо виражена перевага освітнього процесу спостерігається у формуванні навичок першої групи, що стосуються інформаційної та дата-грамотності. Зокрема, найбільша амплітуда зростання зафіксована за показниками перегляду, пошуку та фільтрації даних ( $x_1$ ), критичного оцінювання інформації ( $x_2$ ) та системного управління цифровим контентом ( $x_3$ ). Суттєвий розрив між «академічним» та «професійним» приростами за цими дескрипторами вказує на те, що університетська підготовка закладає складні когнітивні механізми роботи з даними, які надалі стають фундаментом для виконання професійних обов'язків.

Разом з тим, професійний досвід забезпечує хоч і менший за обсягом, проте стабільний приріст за всіма показниками, що свідчить про доцільність концепції безперервного навчання.

### 3.3. Дослідження впливу факторів на рівень цифрових компетентностей населення

Продовженням аналізу часової динаміки стало виявлення ключових детермінант, які зумовлюють нерівномірність формування цифрових навичок серед різних категорій населення. На даному етапі дослідження постає завдання не просто зафіксувати рівень компетентностей, а обґрунтувати вплив низки соціально-економічних та демографічних факторів на кінцевий рівень навичок. Для досягнення високої точності

висновків було застосовано інструментарій інтелектуального аналізу даних, зокрема метод побудови дерев рішень у середовищі RStudio, а також розрахунок коефіцієнтів кореляції в середовищі Jupyter. Такий підхід дозволяє ідентифікувати найбільш вагомі змінні - від гендерних характеристик до рівня доходу та розміру підприємства, де працюють респонденти. Виявлення цих взаємозв'язків є необхідною умовою для розробки цільових рекомендацій щодо впровадження інноваційних технологій, які б враховували специфіку кожної групи користувачів та сприяли максимально ефективному розвитку їхнього цифрового потенціалу.

Першочерговим завданням у межах факторного аналізу стало з'ясування ролі гендерної приналежності у процесі формування цифрових навичок. Для забезпечення об'єктивності висновків було використано комплексний підхід, що поєднує методи класичної статистики та алгоритми машинного навчання. Розрахунки проводилися мовою програмування Python (повний текст скрипту наведено у Додатку В), а статистична значущість результатів перевірялася за допомогою t-критерію Стьюдента для незалежних вибірок.

Згідно з отриманими даними, значення t-статистики становить  $-0,161$ , а показник p-value дорівнює  $0,8718$ . Оскільки отримане значення ймовірності значно перевищує критичний поріг у  $0,05$ , ми маємо всі підстави стверджувати, що різниця між групами не є статистично значущою. Отже, у досліджуваній вибірці стать респондентів не виступає фактором впливу на рівень їхньої цифрової компетентності.

Візуальним підтвердженням цього висновку є побудована діаграма «ящик з вусями» (рис. 3.9), яка демонструє розподіл середніх значень навичок для чоловіків та жінок. Медіанні значення обох груп знаходяться на ідентичному рівні (близько 6 балів), а практичний збіг інтерквартильних розмахів свідчить про високу однорідність розподілів. Це вказує на те, що доступ до технологій та інтенсивність їх освоєння в сучасних умовах є рівномірними для обох статей.

Для більш детальної перевірки прихованих залежностей у середовищі RStudio було побудовано дерево рішень (код наведено у Додатку В), мета якого полягала у спробі класифікувати респондентів за статтю на основі їхніх цифрових дескрипторів. Результати побудови дерева (рис. 3.10) підтверджують складність такої сепарації. Корінь дерева охоплює всю вибірку ( $n = 397$ ), де жінки складають  $57,4\%$ , а чоловіки -  $42,6\%$ . Алгоритм намагається виділити групи через порогові значення навичок, як-от програмування (x13) чи оцінювання даних (x2), проте структура дерева виявляється надто розгалуженою та не демонструє домінуючих «гендерних» компетентностей.

Рис. 3.9. Діаграма «box plot»

Рис. 3.10. Дерево рішень за статтю

Таким чином, інтегрований аналіз дозволяє зробити впевнений висновок: у досліджуваному середовищі фактор статі є нейтральним. Цифрова грамотність визначається не біологічними характеристиками, а, ймовірно, іншими параметрами, що ми розглянемо у наступних дослідженнях. Наступним проведеним аналізом взаємозв'язку між рівнем сформованості цифрових компетентностей та економічним добробутом респондентів, вираженим через розмір їхнього доходу (y10). Для вирішення цього завдання було використано дворівневу методологію. На першому рівні за допомогою мови програмування Python (бібліотеки Pandas та NumPy) було проведено класичний кореляційний аналіз для пошуку лінійних залежностей (повний текст коду представлено у Додатку Г).

Отримані результати лінійної кореляції (рис. 3.13) демонструють дуже слабку силу зв'язку, де найбільш значущими виявилися навички виявлення прогалин цифрової компетентності (x21, коефіцієнт  $0,157$ ), вирішення технічних проблем (x18,  $0,115$ ) та визначення технологічних потреб (x19,  $0,112$ ). Це дозволяє зробити попередній висновок, що сам по собі високий рівень окремих навичок не гарантує автоматичного пропорційного зростання доходів у лінійній прогресії.

Рис. 3.11. Результати лінійної кореляції з рівнем доходу

Для виявлення більш глибоких, нелінійних закономірностей на другому рівні було застосовано інтелектуальний аналіз даних у середовищі RStudio за допомогою алгоритму побудови дерев класифікації (бібліотека граф). Унікальність цього методу полягає у здатності виявляти приховані комбінації факторів та специфічні порогові значення, які критично впливають на цільову змінну доходу. Побудоване дерево рішень (рис. 3.12) підтверджує, що дохід визначається не сумарним рівнем грамотності, а структурованим поєднанням компетенцій.

Рис. 3.12. Дерево рішень за рівнем доходу

Глибокий аналіз структури побудованого дерева рішень дозволяє ідентифікувати специфічні «точки зростання», які визначають економічну ефективність використання цифрового капіталу. Фундаментальним фільтром у цій ієрархії виступає навичка програмування (x13), дескриптор якої є кореневим вузлом дерева з пороговим значенням  $5,5$  балів. Респонденти, чий рівень володіння цією компетенцією дорівнює або перевищує вказану позначку, демонструють значно вищу ймовірність потрапляння до групи з доходом понад  $30,001$  грн, що складає  $0,55$  у відповідній гілці.

Подальший аналіз виявляє ефект синергії навичок, де для фахівців із високим рівнем програмування критично важливим стає поєднання екологічної відповідальності та креативності. Зокрема, досягнення порогу в  $7,5$  балів за показниками захисту навколишнього середовища (x17) та творчого використання технологій (x20) фактично маркує перехід спеціаліста до категорії високооплачуваних професіоналів, інтегрованих у складні інноваційні системи.

Водночас дослідження виявило існування певних компенсаторних механізмів для осіб із нижчим рівнем підготовки в галузі програмування (x13 &lt;  $5,5$ ). Для цієї категорії респондентів можливість збереження відносно високого рівня доходу напряму залежить від їхньої спроможності працювати в безпековому та правовому полях, а саме підтримки високої компетенції у захисті персональних даних ( $x15 \geq 3,5$ ) та ґрунтовної обізнаності з авторським правом ( $x12 \geq 7,5$ ). Отже, дерево рішень наочно ілюструє, що цифрова компетентність конвертується в економічний успіх не автоматично, а через формування специфічних профілів, де дефіцит одних складних навичок має обов'язково перекриватися високим рівнем базових регуляторних та етичних умінь.

Таким чином, використання комбінованого інструментарію Python та R дозволило довести, що вплив цифрових навичок на розмір доходу має ступінчастий та системний характер. Висока економічна віддача досягається через досягнення конкретних порогових значень у ключових компетенціях та їх гармонійне поєднання, що є вагомим аргументом на користь впровадження спеціалізованих інноваційних програм навчання для дорослого населення.

Важливим вектором дослідження став аналіз впливу цифрових компетентностей на успішність кар'єри респондентів (y9). Для коректної математичної обробки категоріальні дані щодо посад були трансформовані у числові ранги: від некерівних позицій (1) до топменеджменту (4). 3

огляду на рангову природу змінних, для пошуку лінійних залежностей у середовищі Python було застосовано метод кореляції Спірмена (текст програми наведено у Додатку Д).

Отримані результати (рис. 3.13) демонструють слабкий прямий зв'язок між кар'єрним статусом та рівнем окремих навичок.

Рис. 3.13. Результати лінійної кореляції з успішністю кар'єри

Проте, аналіз за допомогою дерева рішень у середовищі RStudio (бібліотека `gpart`) показав, що успішність кар'єри залежить від цифрових навичок нелінійно. Алгоритм класифікації дозволив ідентифікувати складні комбінації та порогові рівні компетенцій, які розділяють керівні та виконавські позиції. Згідно з побудованою моделлю (рис. 3.14), керівні посади середньої та вищої ланки частіше пов'язані з усвідомленим використанням технологій у питаннях безпеки та вирішення технічних проблем.

Зокрема, коренем дерева виступає дескриптор авторського права ( $x_{12}$ ), де низькі значення ( $< 2.5$ ) однозначно маркують некерівні позиції. Для тих, хто долає цей поріг, ключову роль починають відігравати безпекові аспекти ( $x_{17}$  - захист навколишнього середовища та  $x_{16}$  - захист здоров'я). Цікавим є спостереження, що високий рівень самооцінки власних навичок ( $x_{21} \geq 5.5$ ) у поєднанні з помірним рівнем програмування ( $x_{13} < 6.5$ ) найчастіше веде до позицій керівників середньої ланки. Водночас високі, але вузькоспеціалізовані технічні навички без належного балансу з навичками безпеки та комунікації частіше корелюють із самозайнятістю або некерівними посадами.

Таким чином, кар'єрний успіх формується не через максимальний рівень окремих умінь, а через їхній структурований баланс, де управлінська відповідальність підкріплюється цифровою етикою та безпековою грамотністю.

Рис. 3.14. Дерево рішень за успішністю кар'єри

Дослідження самооцінки успішності кар'єри ( $y_8$ ) виявило цікаву специфіку його залежності від рівня цифрової грамотності. За результатами кореляційного аналізу, проведеного мовою Python (скрипт у Додатку Е), встановлено прямий, хоча і помірний за силою зв'язок між розвиненими навичками та впевненістю у власному успіху (рис. 3.15). Найбільш виражений вплив мають навички програмування ( $x_{13}$ , коефіцієнт  $0.242$ ), захисту пристроїв ( $x_{14}$ ,  $0.205$ ) та знання авторського права ( $x_{12}$ ,  $0.193$ ).

Рис. 3.15. Результати лінійної кореляції з самооцінкою успішності кар'єри

Водночас аналіз дерева рішень у середовищі RStudio підтверджує, що самооцінка пов'язана з цифровими навичками нелінійно. Вищу оцінку власного кар'єрного успіху частіше мають ті, хто досягає критичних порогових рівнів у програмуванні ( $x_{13}$ ), вирішенні технічних проблем ( $x_{18}$ ) та цифровій безпеці ( $x_{14}$ ). Наприклад, при рівні володіння правовими аспектами  $x_{12} \geq 6.5$ , ймовірність високої самооцінки складає  $0.66$ . Натомість окремі високі результати без балансу з іншими компетенціями не завжди підвищують впевненість у собі. Отже, суб'єктивне відчуття успішності формується не просто через максимальний рівень окремих умінь, а через їхню гармонійну комбінацію та практичне усвідомлення власної спроможності у цифровому просторі.

Завершальним етапом дослідження факторів впливу став аналіз взаємозв'язку між розміром підприємства ( $y_{14}$ ) та рівнем цифрових компетентностей працівників. На початковому етапі розрахунки, проведені мовою Python для всієї вибірки респондентів (рис. 3.17), не виявили лінійних залежностей (код скрипту представлено у Додатку Є).

Рис. 3.16. Дерево рішень за самооцінкою успішності кар'єри

Рис. 3.17. Результати лінійної кореляції з розміром підприємства

Однак глибша дескриптивна аналітика із застосуванням фільтрації даних показала специфічні закономірності для представників малого бізнесу. При відборі лише тих осіб, які працевлаштовані на мікропідприємствах або є самозайнятими (значення  $y_{14}$  в межах  $\{1, 2\}$ ), було виявлено слабкий обернений зв'язок. Це свідчить про те, що такі категорії працівників мають нижчий рівень знань та навичок у критичних сферах, зокрема у захисті здоров'я та благополуччя при використанні технологій ( $x_{16}$ , коефіцієнт  $-0.269$ ).

Рис. 3.18. Результати лінійної кореляції відфільтрованих даних з розміром підприємства

Для ідентифікації прихованої структури цифрових навичок залежно від типу організації було побудовано дерево рішень у середовищі RStudio (рис. 3.19).

Отримані результати підтверджують, що тип підприємства пов'язаний не стільки із середнім рівнем навичок, скільки з їхньою специфічною архітектурою. Великі компанії (понад 250 працівників) частіше залучають фахівців із сильнішими навичками цифрової безпеки та мережевої взаємодії. Зокрема, коренем дерева виступає дескриптор захисту пристроїв ( $x_{14}$ ): за високих значень цього показника ( $\geq 7.5$ ) та високої компетенції у захисті персональних даних ( $x_{15} \geq 7.5$ ) ймовірність працевлаштування на середніх та великих підприємствах значно зростає. Натомість малі підприємства та самозайняті є характернішими для респондентів із нижчими або неузгодженими цифровими компетенціями, особливо у сферах програмного кодування та системної безпеки.

Отже, масштаб підприємства виступає потужним зовнішнім регулятором, який формує запит на певний цифровий профіль працівника: від універсальної, але часто фрагментарної грамотності у малому бізнесі до високоспеціалізованої та безпеково-орієнтованої моделі у великих корпоративних структурах.

Рис. 3.19. Дерево рішень за розміром підприємства

Загалом, проведений аналіз підтвердив, що рівень цифрових компетентностей населення не залежить від статі, проте корелює з економічними та професійними показниками. Виявлено, що вплив навичок на дохід та кар'єрний успіх має нелінійний характер: вирішальним є не загальний рівень грамотності, а досягнення порогових значень у таких «**сильних**» компетенціях, як програмування, цифрова безпека та правова обізнаність. При цьому великі підприємства виступають стимулом для формування цілісного безпекового профілю працівника, тоді як у малій самозайнятості спостерігається дефіцит навичок технічного самозахисту. Таким чином, для підвищення конкурентоспроможності на ринку праці

необхідне не просто нарощування балів, а збалансоване поєднання технічних умінь із високим рівнем відповідальності та безпеки.

3.4. Рекомендації для розробки ІС щодо формування цифрових навичок

Результати проведеного аналітичного обґрунтування стали підґрунтям для розробки функціональних вимог до інформаційної системи, спрямованої на персоналізований розвиток цифрових компетентностей. Основна ідея полягала у створенні гнучкого інструменту, який не просто пропонує навчальний контент, а адаптує його під індивідуальний профіль користувача, виявлений під час аналізу.

Ключовою рекомендацією стала дворівнева структура системи, що поєднує модуль діагностики (тестування) та модуль навчання (курси).

Оскільки дослідження виявило нелінійний характер зв'язку між навичками та успішністю, система повинна насамперед ідентифікувати «критичні прогалини». Наприклад, враховуючи низький середній бал за дескриптором x13 (програмування) та його високий вплив на дохід, ІС має пропонувати курси-прогалини саме з тих напрямків, які є «пусковими гачками» для кар'єрного зростання.

Для забезпечення високої точності навчання було запропоновано систему фільтрації контенту за трьома основними осями. По-перше, це рівні складності від 1 до 8 згідно з методологією DigComp 2.2, що дозволяє користувачеві поступово нарощувати складність матеріалу. По-друге, контент структурувати за п'ятьма категоріями цифрових навичок (від інформаційної грамотності до вирішення проблем), що забезпечує збалансований розвиток профілю. По-третє, додати фільтр за тривалістю курсу, оскільки платформа розрахована на створення максимально зручного та комфортного простору для користування.

Таким чином, розроблена ІС базується на принципах предиктивної аналітики: вона використовує виявлені порогові значення як орієнтири для формування індивідуальних освітніх траєкторій. Такий підхід дозволяє перетворити інформаційну систему з простого сховища курсів на інтелектуальний помічник, який допомагає користувачеві досягати саме тих рівнів компетентності, що мають найбільшу вартість на сучасному ринку праці.

## ВИСНОВКИ

У ході виконання кваліфікаційної магістерської роботи було проведено комплексне аналітичне обґрунтування доцільності впровадження інноваційних технологій для формування цифрових навичок.

Відповідно до поставлених завдань отримано наступні результати:

1. Проведено аналіз предметної області, який підтвердив критичну роль цифрової трансформації для розвитку людського капіталу. На основі моделі DigComp 2.2 виділено 21 дескриптор цифрових компетенцій, що охоплюють п'ять ключових категорій: від інформаційної грамотності до вирішення технічних проблем.
2. Здійснено аналіз програмних продуктів, у результаті якого обґрунтовано використання мов програмування Python (середовище Jupyter Notebook) та R (RStudio) як основних інструментів для повного циклу аналітики та виявлення прихованих закономірностей. Для візуалізації використано платформу Tableau.
3. Проаналізовано 5 категорій цифрових навичок, що дозволило структурувати дані за 21 компетенцією у динаміці їх розвитку. Визначено, що найбільш розвиненими на сьогодні є навички інформаційної грамотності та комунікації, тоді як категорія програмування потребує особливої уваги через найнижчі показники самооцінки.
4. Виконано порівняльний аналіз цифрових навичок на трьох етапах: вступу до ЗВО, випуску та поточної діяльності. Доведено, що найбільший приріст компетентностей забезпечує саме університетська освіта, яка формує необхідний теоретичний та прикладний фундамент для подальшої професійної адаптації.
5. Досліджено вплив факторів (уп) на цифрові навички за допомогою методів кореляційного аналізу та дерев рішень. Встановлено, що стать не є визначальним чинником, натомість виявлено нелінійну залежність доходу та кар'єрного успіху від порогових значень ключових компетенцій, таких як програмування, безпека та авторське право.
6. Розроблено рекомендації для розробки ІС, які передбачають створення адаптивної системи діагностики та навчання. Рекомендовано впровадження персоналізованих траєкторій на основі виявлених прогалин, систему фільтрації за рівнями DigComp 2.2 та категоріями цифрових навичок, що дозволить ефективно нарощувати цифровий потенціал користувачів.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Vuorikari, R., Kluzer, S. and Punie, Y., DigComp 2.2: The Digital Competence Framework for Citizens - With new examples of knowledge, skills and attitudes, EUR 31006 EN, Publications Office of the European Union, Luxembourg, 2022, ISBN 978-92-76-48882-8, doi:10.2760/115376, JRC128415.
2. Жекало, Г. І. Цифрова економіка України: проблеми та перспективи розвитку [Текст] / Г. І. Жекало // Науковий вісник Ужгородського національного університету : серія: Міжнародні економічні відносини та світове господарство / голов. ред. М. М. Палінчак. - Ужгород : Гельветика, 2019. - Вип. 26. No Ч. 1. - С. 56-60
3. Мостова А. Д. Цифрова трансформація бізнесу і суспільства: роль цифрових навичок / А. Д. Мостова, І. В. Тараненко, Г. В. Щолокова // Східна Європа: економіка, бізнес та управління. - 2023. - No 41 - С. 51-57
4. Магдич А.С., Задова О.А. Цифрові навички та компетенції в Україні та Європі: виклики та перспективи. European Vector of Economic Development. 2025. No1 (38). С.58-75. DOI 10.32342/3041-2153-2025-1-38-5
5. Rogel-Salazar, J. (2025). Data Science and Analytics with Python (2nd ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781003482062>
6. Статистична обробка даних: навч. посіб. / О.В. Перегуда, О.А. Капустян, О.Б. Курилко.–Електронне видання, 2022.–103 с
7. Python 3.14.2 documentation: веб сайт. URL: <https://docs.python.org/3/>
8. СТАТИСТИЧНИЙ АНАЛІЗ ДАНИХ / Пашко А.О. : Електронне видання, -2019.-55 с.
9. Радкевич, О. (2021). ДІАГРАМИ, ГРАФІКИ ТА СХЕМИ ЯК ІНСТРУМЕНТАРІЙ ПРЕДСТАВЛЕННЯ ПРОЄКТНОЇ ІНФОРМАЦІЇ. Професійна педагогіка, 1(22), 197-212. <https://doi.org/10.32835/2707-3092.2021.22.197-212>
10. Mienye, I. D., & Sun, Y. (2024). A Survey of Decision Trees: Concepts, Algorithms, and Applications. IEEE Access. DOI: 10.1109/ACCESS.2024.3416838
11. McKinney, Wes. Python for data analysis: Data wrangling with Pandas, NumPy, and IPython. " O'Reilly Media, Inc.", 2012.
12. Chan, B.K.C. (2018). Data Analysis Using R Programming. In: Biostatistics for Human Genetic Epidemiology. Advances in Experimental Medicine and Biology, vol 1082. Springer, Cham. [https://doi.org/10.1007/978-3-319-93791-5\\_2](https://doi.org/10.1007/978-3-319-93791-5_2)
13. Rpart.plot: Plot an rpart model. A simplified interface to the prp function. : веб сайт. URL: <https://www.rdocumentation.org/packages/rpart.plot/versions/3.1.3/topics/rpart.plot>
14. Ghorbani H. MAHALANOBIS DISTANCE AND ITS APPLICATION FOR DETECTING MULTIVARIATE OUTLIERS. Facta Universitatis, Series:

Mathematics and Informatics. 2019. P. 583. URL: <https://doi.org/10.22190/fumi1903583g> (date of access: 21.12.2025).

15. Schroeders U., Schmidt C., Gnamb T. Detecting Careless Responding in Survey Data Using Stochastic Gradient Boosting. Educational and Psychological Measurement. 2021. P. 001316442110047. URL: <https://doi.org/10.1177/00131644211004708> (date of access: 21.12.2025).

16. Using Markov Chains to Detect Careless Responding in Survey Research / T. Biemann et al. Organizational Research Methods. 2025. URL: <https://doi.org/10.1177/10944281251334778> (date of access: 21.12.2025).

17. Straightlining: Overview of Measurement, Comparison of Indicators, and Effects in Mail-Web Mixed-Mode Surveys / Y. Kim et al. Social Science Computer Review. 2018. Vol. 37, no. 2. P. 214-233. URL: <https://doi.org/10.1177/0894439317752406> (date of access: 21.12.2025).

18. 4 Scales Of Measurement In Research - Knowledge Base. Knowledge Base. URL: <https://knowledgebase.centreforelites.com/4-scales-of-measurement-in-research/> (date of access: 21.12.2025).

19. Abu-Shawiesh, M. O. (2008). A simple robust control chart based on MAD. Journal of Mathematics and Statistics, 4(2), 102.

20. Kaur J., Madan N. Association Rule Mining: A Survey. International Journal of Hybrid Information Technology. 2015. Vol. 8, no. 7. P. 239-242. URL: <https://doi.org/10.14257/ijhit.2015.8.7.22> (date of access: 21.12.2025).

21. Kaushik, Minakshi & Sharma, Rahul & Fister jr, Iztok & Draheim, Dirk. (2023). Numerical Association Rule Mining: A Systematic Literature Review. <https://doi.org/10.48550/arXiv.2307.00662>

## ДОДАТКИ

### Додаток А

Програмні коди для побудови кругових діаграм в середовищі Jupyter

### Додаток Б

Програмний код для побудови радарного графіка за категоріями в середовищі Jupyter

### Додаток В

Програмний код для побудови boxplot (Python) та дерева рішень (R) за статтю

```
1) root 397 169 woman (0.42569270 0.57430730) x13_t2&gt;=7.5 57 17 man (0.70175439 0.29824561) x10_t2&lt; 7.5 21 1 man (0.95238095 0.04761905) x10_t2&gt;=7.5 36 16 man (0.55555556 0.44444444) x17_t2&gt;=7.5 28 8 man (0.71428571 0.28571429) x17_t2&lt; 7.5 8 0 woman (0.00000000 1.00000000) x13_t2&lt; 7.5 340 129 woman (0.37941176 0.62058824) x2_t2&lt; 6.5 131 65 man (0.50381679 0.49618321) x19_t2&gt;=6.5 18 4 man (0.77777778 0.22222222) x19_t2&lt; 6.5 113 52 woman (0.46017699 0.53982301) x17_t2&lt; 3.5 60 25 man (0.58333333 0.41666667) x18_t2&gt;=3.5 19 1 man (0.94736842 0.05263158) x18_t2&lt; 3.5 41 17 woman (0.41463415 0.58536585) x7_t2&lt; 3.5 30 14 man (0.53333333 0.46666667) x7_t2&lt; 2.5 11 3 man (0.72727273 0.27272727) x7_t2&gt;=2.5 19 8 woman (0.42105263 0.57894737) x7_t2&gt;=3.5 11 1 woman (0.09090909 0.90909091) x17_t2&gt;=3.5 53 17 woman (0.32075472 0.67924528) x21_t2&gt;=3.5 41 17 woman (0.41463415 0.58536585) x4_t2&lt; 3.5 8 2 man (0.75000000 0.25000000) x4_t2&gt;=3.5 33 11 woman (0.33333333 0.66666667) x16_t2&lt; 3.5 8 2 man (0.75000000 0.25000000) x16_t2&gt;=3.5 25 5 woman (0.20000000 0.80000000) x21_t2&lt; 3.5 12 0 woman (0.00000000 1.00000000) x2_t2&gt;=6.5 209 63 woman (0.30143541 0.69856459) x19_t2&gt;=7.5 49 24 man (0.51020408 0.48979592) x13_t2&lt; 3.5 17 3 man (0.82352941 0.17647059) x13_t2&gt;=3.5 32 11 woman (0.34375000 0.65625000) x19_t2&lt; 7.5 160 38 woman (0.23750000 0.76250000) x10_t2&lt; 6.5 73 29 woman (0.39726027 0.60273973) x19_t2&gt;=4.5 60 28 woman (0.46666667 0.53333333) x3_t2&gt;=7.5 12 2 man (0.83333333 0.16666667) x3_t2&lt; 7.5 48 18 woman (0.37500000 0.62500000) x14_t2&gt;=4.5 35 16 woman (0.45714286 0.54285714) x5_t2&lt; 7.5 27 12 man (0.55555556 0.44444444) x6_t2&lt; 6.5 20 7 man (0.65000000 0.35000000) x6_t2&gt;=6.5 7 2 woman (0.28571429 0.71428571) x5_t2&gt;=7.5 8 1 woman (0.12500000 0.87500000) x14_t2&lt; 4.5 13 2 woman (0.15384615 0.84615385) x19_t2&lt; 4.5 13 1 woman (0.07692308 0.92307692) x10_t2&gt;=6.5 87 9 woman (0.10344828 0.89655172) *
```

### Додаток Г

Програмний код для побудови лінійної кореляції (Python) та дерева рішень (R) за розміром доходу

```
1) root 397 230 &gt; 30001 грн. (0.42 0.16 0.14 0.12 0.1 0.063) x13_t2&gt;=5.5 106 48 &gt; 30001 грн. (0.55 0.2 0.094 0.038 0.11 0.0094) x17_t2&lt; 7.5 56 17 &gt; 30001 грн. (0.7 0.14 0.11 0.054 0 0) x13_t2&gt;=7.5 23 2 &gt; 30001 грн. (0.91 0.087 0 0 0) x13_t2&lt; 7.5 33 15 &gt; 30001 грн. (0.55 0.18 0.18 0.091 0 0) x3_t2&lt; 7.5 26 8 &gt; 30001 грн. (0.69 0.038 0.19 0.077 0 0) x3_t2&gt;=7.5 7 2 10001 - 15001 грн (0 0.71 0.14 0.14 0 0) x17_t2&gt;=7.5 50 31 &gt; 30001 грн. (0.38 0.26 0.08 0.02 0.24 0.02) x20_t2&lt; 7.5 13 3 &gt; 30001 грн. (0.77 0.15 0 0.077 0 0) x20_t2&gt;=7.5 37 25 25001 - 30000 грн. (0.24 0.3 0.11 0 0.32 0.027) x10_t2&lt; 7.5 7 2 10001 - 15001 грн (0.29 0.71 0 0 0) x10_t2&gt;=7.5 30 18 25001 - 30000 грн. (0.23 0.2 0.13 0 0.4 0.033) x13_t2&lt; 5.5 291 182 &gt; 30001 грн. (0.37 0.14 0.15 0.15 0.096 0.082) x15_t2&lt; 3.5 66 30 &gt; 30001 грн. (0.55 0.12 0.17 0.015 0.015 0.14) x19_t2&gt;=2.5 52 17 &gt; 30001 грн. (0.67 0.077 0.12 0 0.019 0.12) x19_t2&lt; 2.5 14 9 15001 - 20000 грн. (0.071 0.29 0.36 0.071 0 0.21) x15_t2&gt;=3.5 225 152 &gt; 30001 грн. (0.32 0.15 0.15 0.19 0.12 0.067) x12_t2&gt;=7.5 46 19 &gt; 30001 грн. (0.59 0.11 0.087 0.17 0 0.043) x18_t2&lt; 6.5 20 1 &gt; 30001 грн. (0.95 0 0.05 0 0) x18_t2&gt;=6.5 26 18 &gt; 30001 грн. (0.31 0.19 0.12 0.31 0 0.077) x17_t2&lt; 6.5 10 2 &gt; 30001 грн. (0.8 0 0.2 0 0) x17_t2&gt;=6.5 16 8 20001 - 25000 грн. (0 0.31 0.062 0.5 0 0.12) x12_t2&lt; 7.5 179 133 &gt; 30001 грн. (0.26 0.16 0.17 0.19 0.15 0.073) x16_t2&gt;=3.5 157 113 &gt; 30001 грн. (0.28 0.17 0.18 0.14 0.16 0.076) x11_t2&lt; 6.5 111 71 &gt; 30001 грн. (0.36 0.17 0.21 0.11 0.099 0.054) x15_t2&lt; 5.5 56 28 &gt; 30001 грн. (0.5 0.11 0.14 0.12 0.11 0.018) x17_t2&gt;=5.5 16 1 &gt; 30001 грн. (0.94 0.062 0 0 0) x17_t2&lt; 5.5 40 27 &gt; 30001 грн. (0.32 0.12 0.2 0.18 0.15 0.025) x3_t2&lt; 6.5 31 18 &gt; 30001 грн. (0.42 0.13 0.097 0.13 0.19 0.032) x10_t2&gt;=4.5 8 0 &gt; 30001 грн. (1 0 0 0 0) x10_t2&lt; 4.5 23 17 25001 - 30000 грн. (0.22 0.17 0.13 0.17 0.26 0.043) x13_t2&gt;=2.5 12 7 &gt; 30001 грн. (0.42 0.25 0.17 0.17 0 0) x13_t2&lt; 2.5 11 5 25001 - 30000 грн. (0 0.091 0.091 0.18 0.55 0.091) x3_t2&gt;=6.5 9 4 15001 - 20000 грн. (0 0.11 0.56 0.33 0 0) x15_t2&gt;=5.5 55 40 15001 - 20000 грн. (0.22 0.24 0.27 0.091 0.091 0.091) x4_t2&gt;=7.5 10 4 &gt; 30001 грн. (0.6 0.1 0.1 0.1 0.1 0) x4_t2&lt; 7.5 45 31 15001 - 20000 грн. (0.13 0.27 0.31 0.089 0.089 0.11) x17_t2&lt; 6.5 30 20 10001 - 15001 грн (0.17 0.33 0.17 0.1 0.13 0.1) x16_t2&lt; 5.5 12 8 &gt; 30001 грн. (0.33 0.083 0.25 0.17 0 0.17) x16_t2&gt;=5.5 18 9 10001 - 15001 грн (0.056 0.5 0.11 0.056 0.22 0.056) x17_t2&gt;=6.5 15 6 15001 - 20000 грн. (0.067 0.13 0.6 0.067 0 0.13) x11_t2&gt;=6.5 46 32 25001 - 30000 грн. (0.087 0.15 0.11 0.22 0.3 0.13) x11_t2&gt;=7.5 17 11 20001 - 25000 грн. (0.24 0.24 0.059 0.35 0 0.12) x11_t2&lt; 7.5 29 15 25001 - 30000 грн. (0 0.1 0.14 0.14 0.48 0.14) x16_t2&lt; 3.5 22 10
```

## Додаток Д

Програмний код для побудови лінійної кореляції (Python) та дерева рішень (R) за успішністю кар'єри(1) root 397 236 не керівна посада за наймом (0.013 0.37 0.41 0.15 0.068(2) x12\_t2&lt; 2.5 36(9) не керівна посада за наймом (0.056 0.14 0.75 0.056 0(3) x12\_t2&gt;=2.5 361 220 керівник середньої ланки (0.0083 0.39 0.37 0.16 0.075(6) x17\_t2&gt;=2.5 346 207 керівник середньої ланки (0.0087 0.4 0.38 0.16 0.052(12) x16\_t2&gt;=2.5 338 199 керівник середньої ланки (0.0089 0.41 0.39 0.14 0.053(24) x17\_t2&gt;=3.5 305 173 керівник середньої ланки (0.0098 0.43 0.36 0.15 0.052(48) x13\_t2&lt; 6.5 233 120 керівник середньої ланки (0.0086 0.48 0.35 0.12 0.034(96) x21\_t2&gt;=5.5 143(62) керівник середньої ланки (0.007 0.57 0.3 0.098 0.028(192) x16\_t2&lt; 6.5 60(16) керівник середньої ланки (0.017 0.73 0.22 0.033 0(193) x16\_t2&gt;=6.5 83(46) керівник середньої ланки (0 0.45 0.36 0.14 0.048(386) x7\_t2&lt; 7.5 57(23) керівник середньої ланки (0 0.6 0.26 0.088 0.053(772) x18\_t2&gt;=5.5 49(16) керівник середньої ланки (0 0.67 0.22 0.1 0(1544) x6\_t2&gt;=7.5 25(3) керівник середньої ланки (0 0.88 0.04 0.08 0(1545) x6\_t2&lt; 7.5 24(13) керівник середньої ланки (0 0.46 0.42 0.12 0(3090) x5\_t2&gt;=7.5 9(1) керівник середньої ланки (0 0.89 0.11 0 0(3091) x5\_t2&lt; 7.5 15(6) не керівна посада за наймом (0 0.2 0.6 0.2 0(773) x18\_t2&lt; 5.5 8(4) не керівна посада за наймом (0 0.12 0.5 0 0.38(387) x7\_t2&gt;=7.5 26(11) не керівна посада за наймом (0 0.12 0.58 0.27 0.038(774) x20\_t2&gt;=6.5 18(4) не керівна посада за наймом (0 0.17 0.78 0 0.056(775) x20\_t2&lt; 6.5 8(1) самозайнятий, працюю на себе (0 0 0.12 0.88 0(97) x21\_t2&lt; 5.5 90(52) не керівна посада за наймом (0.011 0.36 0.42 0.17 0.044(194) x11\_t2&lt; 4.5 45(21) керівник середньої ланки (0 0.53 0.22 0.2 0.044(388) x18\_t2&gt;=5.5 12(0) керівник середньої ланки (0 1 0 0 0(389) x18\_t2&lt; 5.5 33(21) керівник середньої ланки (0 0.36 0.3 0.27 0.061(778) x12\_t2&lt; 3.5 13(5) керівник середньої ланки (0 0.62 0.31 0.077 0(779) x12\_t2&gt;=3.5 20(12) самозайнятий, працюю на себе (0 0.2 0.3 0.4 0.1(1558) x3\_t2&lt; 3.5 7(3) керівник середньої ланки (0 0.57 0.29 0 0.14(1559) x3\_t2&gt;=3.5 13(5) самозайнятий, працюю на себе (0 0 0.31 0.62 0.077(195) x11\_t2&gt;=4.5 45(17) не керівна посада за наймом (0.022 0.18 0.62 0.13 0.044(49) x13\_t2&gt;=6.5 72(44) не керівна посада за наймом (0.014 0.26 0.39 0.22 0.11(98) x3\_t2&gt;=6.5 63(35) не керівна посада за наймом (0.016 0.3 0.44 0.13 0.11(196) x20\_t2&gt;=7.5 47(25) не керівна посада за наймом (0.021 0.36 0.47 0.021 0.13(392) x14\_t2&gt;=7.5 36(19) керівник середньої ланки (0.028 0.47 0.36 0.028 0.11(393) x14\_t2&lt; 7.5 11(2) не керівна посада за наймом (0 0 0.82 0 0.18(197) x20\_t2&lt; 7.5 16(9) самозайнятий, працюю на себе (0 0.12 0.38 0.44 0.063(99) x3\_t2&lt; 6.5 9(1) самозайнятий, працюю на себе (0 0 0 0.89 0.11(25) x17\_t2&lt; 3.5 33(11) не керівна посада за наймом (0 0.21 0.67 0.061 0.061(13) x16\_t2&lt; 2.5 8(1) самозайнятий, працюю на себе (0 0 0.12 0.88 0(7) x17\_t2&lt; 2.5 15(6) топ-керівник (0 0.13 0.13 0.13 0.6) \*

## Додаток Е

Програмний код для побудови лінійної кореляції (Python) та дерева рішень (R) за самооцінкою успішності кар'єри(1) root 397 217 4 (0.025 0.055 0.35 0.45 0.12(2) x12\_t2&lt; 6.5 245 142 3 (0.041 0.069 0.42 0.33 0.14(4) x13\_t2&lt; 7.5 231 128 3 (0.043 0.065 0.45 0.33 0.11(8) x18\_t2&gt;=5.5 91(37) 3 (0.033 0.033 0.59 0.31 0.033(16) x7\_t2&lt; 5.5 21(3) 3 (0 0 0.86 0.095 0.048(17) x7\_t2&gt;=5.5 70(34) 3 (0.043 0.043 0.51 0.37 0.029(34) x19\_t2&gt;=6.5 47(19) 3 (0.043 0.064 0.6 0.26 0.043(68) x9\_t2&gt;=5.5 36(10) 3 (0.028 0.083 0.72 0.14 0.028(69) x9\_t2&lt; 5.5 11(4) 4 (0.091 0 0.18 0.64 0.091(35) x19\_t2&lt; 6.5 23(9) 4 (0.043 0 0.35 0.61 0(70) x21\_t2&lt; 5.5 11(4) 3 (0 0 0.64 0.36 0(71) x21\_t2&gt;=5.5 12(2) 4 (0.083 0 0.083 0.83 0(9) x18\_t2&lt; 5.5 140(91) 3 (0.05 0.086 0.35 0.35 0.16(17) x11\_t2&lt; 2.5 15(3) 3 (0.067 0.067 0.8 0.067 0(19) x11\_t2&gt;=2.5 125(77) 4 (0.048 0.088 0.3 0.38 0.18(38) x16\_t2&gt;=3.5 75(47) 3 (0.027 0.13 0.37 0.28 0.19(76) x15\_t2&lt; 3.5 12(6) 2 (0 0.5 0.083 0.083 0.33(77) x15\_t2&gt;=3.5 63(36) 3 (0.032 0.063 0.43 0.32 0.16(154) x11\_t2&lt; 4.5 31(17) 3 (0.032 0.13 0.45 0.13 0.26(155) x11\_t2&gt;=4.5 32(16) 4 (0.031 0 0.41 0.5 0.062(310) x10\_t2&gt;=6.5 8(1) 3 (0.12 0 0.88 0 0(311) x10\_t2&lt; 6.5 24(8) 4 (0 0 0.25 0.67 0.083(39) x16\_t2&lt; 3.5 50(23) 4 (0.08 0.02 0.18 0.54 0.18(5) x13\_t2&gt;=7.5 14(5) 5 (0 0.14 0 0.21 0.64(3) x12\_t2&gt;=6.5 152(52) 4 (0 0.033 0.22 0.66 0.086(6) x11\_t2&lt; 7.5 74(33) 4 (0 0.054 0.36 0.55 0.027(12) x10\_t2&gt;=7.5 13(7) 3 (0 0.23 0.46 0.23 0.077(13) x10\_t2&lt; 7.5 61(23) 4 (0 0.016 0.34 0.62 0.016(26) x15\_t2&lt; 6.5 13(5) 3 (0 0 0.62 0.38 0(27) x15\_t2&gt;=6.5 48(15) 4 (0 0.021 0.27 0.69 0.021(54) x10\_t2&lt; 5.5 9(3) 3 (0 0 0.67 0.33 0(55) x10\_t2&gt;=5.5 39(9) 4 (0 0.026 0.18 0.77 0.026(7) x11\_t2&gt;=7.5 78(19) 4 (0 0.013 0.09 0.76 0.14) \*

## Додаток С

Програмний код для побудови лінійної кореляції (Python) та дерева рішень (R) за розміром підприємства(1) root 397 268 середнє підприємство: від 50 до 249 працівників; (0.27 0.21 0.091 0.11 0.32(2) x14\_t2&lt; 7.5 310 212 велике підприємство: 250 і більше працівників. (0.32 0.22 0.094 0.087 0.29(4) x6\_t2&gt;=6.5 135(74) велике підприємство: 250 і більше працівників. (0.45 0.18 0.11 0.081 0.18(8) x10\_t2&lt; 7.5 94(42) велике підприємство: 250 і більше працівників. (0.55 0.14 0.12 0.064 0.13(16) x2\_t2&gt;=6.5 84(33) велике підприємство: 250 і більше працівників. (0.61 0.15 0.083 0.071 0.083(17) x2\_t2&lt; 6.5 10(5) середнє підприємство: від 50 до 249 працівників; (0.1 0 0.4 0 0.5(9) x10\_t2&gt;=7.5 41(29) середнє підприємство: від 50 до 249 працівників; (0.22 0.27 0.098 0.12 0.29(18) x14\_t2&lt; 5.5 17(7) мале підприємство: від 10 до 49 працівників; (0.12 0.59 0.24 0 0.059(19) x14\_t2&gt;=5.5 24(13) середнє підприємство: від 50 до 249 працівників; (0.29 0.042 0 0.21 0.46(38) x6\_t2&gt;=7.5 11(5) велике підприємство: 250 і більше працівників. (0.55 0.091 0 0.36 0(39) x6\_t2&lt; 7.5 13(2) середнє підприємство: від 50 до 249 працівників; (0.077 0 0 0.077 0.85(5) x6\_t2&lt; 6.5 175 110 середнє підприємство: від 50 до 249 працівників; (0.21 0.25 0.08 0.091 0.37(10) x8\_t2&lt; 2.5 18(8) мале підприємство: від 10 до 49 працівників; (0.056 0.56 0.22 0.056 0.11(11) x8\_t2&gt;=2.5 157(94) середнє підприємство: від 50 до 249 працівників; (0.23 0.21 0.064 0.096 0.4(22) x13\_t2&gt;=3.5 41(28) велике підприємство: 250 і більше працівників. (0.32 0.24 0.15 0.12 0.17(23) x13\_t2&lt; 3.5 116(60) середнє підприємство: від 50 до 249 працівників; (0.2 0.2 0.034 0.086 0.48(46) x14\_t2&lt; 4.5 85(51) середнє підприємство: від 50 до 249 працівників; (0.21 0.26 0.035 0.094 0.4(92) x21\_t2&gt;=3.5 34(18) мале підприємство: від 10 до 49 працівників; (0.24 0.47 0 0.088 0.21(184) x21\_t2&gt;=5.5 8(3) велике підприємство: 250 і більше працівників. (0.62 0.12 0 0 0.25(185) x21\_t2&lt; 5.5 26(11) мале підприємство: від 10 до 49 працівників; (0.12 0.58 0 0.12 0.19(93) x21\_t2&lt; 3.5 51(24) середнє підприємство: від 50 до 249 працівників; (0.2 0.12 0.059 0.098 0.53(47) x14\_t2&gt;=4.5 31(9) середнє підприємство: від 50 до 249 працівників; (0.16 0.032 0.032 0.065 0.71(3) x14\_t2&gt;=7.5 87(47) середнє підприємство: від 50 до 249 працівників; (0.1 0.17 0.08 0.18 0.46(6) x15\_t2&lt; 7.5 15(6) самозайнятий (0.27 0 0.067 0.6 0.067(7) x15\_t2&gt;=7.5 72(33) середнє підприємство: від 50 до 249 працівників; (0.069 0.21 0.083 0.097 0.54(14) x12\_t2&lt; 6.5 8(3) мале підприємство: від 10 до 49 працівників; (0.12 0.62 0.25 0 0(15) x12\_t2&gt;=6.5 64(25) середнє підприємство: від 50 до 249 працівників; (0.062 0.16 0.062 0.11 0.61(30) x21\_t2&lt; 7.5 26(15) середнє підприємство: від 50 до 249 працівників; (0.038 0.23 0.038 0.27 0.42(60) x17\_t2&gt;=7.5 15(8) самозайнятий (0.067 0.4 0 0.47 0.067(61) x17\_t2&lt; 7.5 11(1) середнє підприємство: від 50 до 249 працівників; (0 0 0.091 0 0.91(31) x21\_t2&gt;=7.5 38(10) середнє підприємство: від 50 до 249 працівників; (0.079 0.11 0.079 0 0.74) \*