

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

Навчально-науковий інститут
«Інститут інформаційних технологій в економіці»

Кафедра математичного моделювання та статистики

Освітньо-професійна програма

«Економічна кібернетика»

Галузь знань

05 «Соціальні та поведінкові науки»

Спеціальність

051 «Економіка»

Форма навчання: очна (денна)

КВАЛІФІКАЦІЙНА БАКАЛАВРСЬКА РОБОТА

на тему **«Аналіз та моделювання індексу споживчих цін в Україні»**
(назва теми)

здобувача Музики Івонни Петрівни

(ПІБ)

(підпис)

Науковий керівник: кандидат економічних наук, доцент
Мазур Маргарита Вікторівна

(підпис)

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

В.о. завідувача кафедри: кандидат фізико-математичних наук,
професор Великоіваненко Г.І.

(підпис)

Київ 2024

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

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

Кафедра математичного моделювання та статистики

Освітньо-професійна програма

Галузь знань

Спеціальність

Економічна кібернетика

05 соціальні та поведінкові науки

051 економіка

ПОГОДЖЕНО

Керівник проектної групи (гарант)
освітньо-професійної програми

О.М. Притоманова

(підпис)

2024 р.

ЗАТВЕРДЖУЮ

В.о. завідувача кафедри
математичного моделювання та
статистики

Г.І. Великоіваненко

(підпис)

2024 р.

ІНДИВІДУАЛЬНЕ ЗАВДАННЯ

здобувача вищої освіти

Музики Іванни Петрівни

(прізвище, ім'я, по батькові)

 очної (денної) **форми навчання**

очної (денної), заочної, дистанційної

на підготовку кваліфікаційної бакалаврської роботи

на тему «Аналіз та моделювання індексу споживчих цін в Україні»

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

Кваліфікаційна бакалаврська робота виконується на матеріалах

сайтів Національного банку України <https://bank.gov.ua/>, Державної служби статистики України <https://www.ukrstat.gov.ua/>, Міністерства фінансів України <https://minfin.gov.ua/> та інших відкритих джерел.

План кваліфікаційної бакалаврської роботи

Розділ 1	Теоретичні та методичні основи дослідження індексу споживчих цін
Розділ 2	Регіональний аналіз індексу споживчих цін
Розділ 3	Моделювання індексу споживчих цін як основного показника інфляції в країні

Об'єкт дослідження:	індекс споживчих цін в Україні
Предмет дослідження:	методи та моделі машинного навчання, що застосовуються для дослідження індексу споживчих цін
Мета кваліфікаційної бакалаврської роботи:	проведення регіонального аналізу, моделювання індексу споживчих цін за допомогою регресійного аналізу, ARIMA та ETS моделей, здійснення на їх основі короткочасного прогнозу та порівняння точності моделей між собою

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

У розділі 1: 1) узагальнити теоретичні аспекти індексу споживчих цін як основного показника інфляції в країні; 2) охарактеризувати сучасні методи розрахунку індексу споживчих цін; 3) проаналізувати особливості розрахунку індексу споживчих цін за регіонами; 4) описати методичні підходи до аналізу індексу споживчих цін.

У розділі 2: 1) провести порівняльний аналіз регіонів за рівнем інфляції; 2) оцінити стан та тенденції розвитку індексу споживчих цін за регіонами; 3) здійснити кластеризацію регіонів за індексом споживчих цін за допомогою методів машинного навчання.

У розділі 3: 1) описати інформаційну базу для моделювання; 2) змоделювати індекс споживчих цін за допомогою регресійного аналізу, ARIMA та ETS моделей; 3) оцінити точність економетричних моделей; 4) здійснити короткостроковий прогноз.

Завдання підготував
науковий керівник

(підпис)

М.В. Мазур
(ініціали, прізвище)
«__» _____ 2024 р.

Завдання одержав
здобувач

(підпис)

І.П. Музика
(ініціали, прізвище)
«__» _____ 2024 р.

Реферат

Кваліфікаційна бакалаврська робота містить 60 сторінок, 4 таблиці, 51 рисунок, 5 формул, перелік використаних джерел з 38 найменувань, додатки.

«Аналіз та моделювання індексу споживчих цін в Україні»

Об'єктом дослідження є індекс споживчих цін в Україні.

Предметом дослідження є методи та моделі машинного навчання, що застосовуються для дослідження індексу споживчих цін.

Мета та завдання виконання дослідження. Метою виконання кваліфікаційної бакалаврської роботи є проведення регіонального аналізу, моделювання індексу споживчих цін за допомогою регресійного аналізу, ARIMA та ETS моделей, здійснення на їх основі короткочасного прогнозу та порівняння точності моделей між собою.

Відповідно до поставленої мети визначені такі *завдання*:

- узагальнити теоретичні аспекти та охарактеризувати сучасні методи розрахунку індексу споживчих цін;
- дослідити закономірності та тенденції розвитку індексу споживчих цін;
- провести регіональний аналіз індексу споживчих цін;
- побудувати моделі регресії, ARIMA та ETS моделі для прогнозування індексу споживчих цін;
- створити короткостроковий прогноз на основі трьох моделей;
- оцінити точність побудованих економетричних моделей.

Теоретична, методична та практична значущість отриманих результатів. Результати даної роботи узагальнюють інформацію щодо індексу споживчих цін, демонструючи його роль, як основного показника інфляції в Україні, а також розширюють наявні знання про методи та моделі машинного навчання, що застосовуються для дослідження індексу інфляції.

Застосування різних методів та моделей для аналізу індексу споживчих цін, включаючи регресійний аналіз, ARIMA та ETS моделі, а також карти Кохонена для регіонального аналізу, можуть бути використані в подальшому для аналізу інших економічних показників.

Практична цінність бакалаврської роботи полягає у аналізі індексу інфляції та перевірці кількох моделей з метою вибору найкращої для прогнозування ІСЦ, що дозволяє в подальшому використовувати її у реальних установах для планування монетарної політики, регулювання рівня видатків державного бюджету тощо.

Рік виконання кваліфікаційної бакалаврської роботи – 2024.

Рік захисту роботи – 2024.

Ключові слова: індекс споживчих цін, метод множинної регресії, економетрична модель, кластеризація, ARIMA модель, ETS модель, карти Кохонена, прогнозування.

В і д г у к
про кваліфікаційну бакалаврську роботу
здобувача освітньо-професійної програми «Економічна кібернетика»
Музики Іванни Петрівни
на тему: «Аналіз та моделювання індексу споживчих цін в Україні»

1. Актуальність теми: В сучасних нестабільних умовах розвитку економіки аналіз та моделювання індексу споживчих цін як основного показника інфляції в країні допомагають прогнозувати економічні тенденції, оперативно реагувати на швидкі зміни цін для мінімізації негативного впливу інфляції та підтримувати економічну стабільність, що обумовлює актуальність даної бакалаврської роботи.
2. Позитивні риси кваліфікаційної бакалаврської роботи: Індивідуальне завдання виконане здобувачем у повному обсязі. В роботі доцільно та грамотно використано математичні методи дослідження, що дозволило автору сформулювати обґрунтовані висновки і рекомендації. Окремі результати проведеного дослідження відображено у тезах конференції, науковій роботі, а також впроваджено у навчальний процес.
3. Наявність самостійних розробок автора: побудована економетрична модель індексу споживчих цін, охарактеризовано вплив на нього основних макроекономічних факторів, здійснено регіональний аналіз індексу інфляції з використанням методу кластеризації, а також побудовано прогноз інфляції.
4. Цінність теоретичних висновків та практичних рекомендацій полягає у можливості ефективніше передбачати зміни споживчих цін та адаптувати стратегії управління через проведений автором аналіз побудованих математичних моделей, що допомагає мінімізувати негативні наслідки інфляції та забезпечити фінансову стабільність для уряду, бізнесу та населення.
5. Наявність недоліків: Для кращого розуміння впливу інфляційних процесів на зміни в економіці доцільним є порівняння індексу споживчих цін з іншими показниками інфляції. Даний недолік залишає місце для майбутніх досліджень цієї теми.
6. Загальна оцінка кваліфікаційної бакалаврської роботи та її допущення до захисту перед ЕК: Кваліфікаційна бакалаврська робота Музики І.П. «Аналіз та моделювання індексу споживчих цін в Україні» є самостійним завершеним дослідженням, відповідає вимогам, що висуваються для даного типу робіт, може бути рекомендована до захисту перед ЕК. Рівень виконання кваліфікаційної бакалаврської роботи «високий».

Науковий керівник:

к.е.н., доцент, доцент кафедри

математичного моделювання та статистики

Мазур М.В.

“ ____ ” _____ 20__ року

РЕЦЕНЗІЯ
на кваліфікаційну роботу освітнього ступеня бакалавр
на тему: «Аналіз та моделювання індексу споживчих цін в Україні»
студентки 4-го курсу спеціальності 051 «Економіка»
освітньо-професійної програми «Економічна кібернетика»
ННІ «Інститут інформаційних технологій в економіці»
Київського національного економічного університету ім. Вадима Гетьмана
Музики Іванни

Індекс споживчих цін як ключовий показник інфляції є важливим для ефективного керування економікою, оскільки впливає на більшість макроекономічних показників, що визначають фінансове становище населення, рівень інвестицій та загальний економічний клімат країни. У ситуації воєнного стану аналіз та моделювання індексу споживчих цін стає надзвичайно актуальним, оскільки воєнні конфлікти часто призводять до швидкого зростання цін через збільшену нестабільність інфраструктури та зменшення обсягів виробництва.

Розкриття теоретико-методологічних та практичних засад аналізу та моделювання індексу споживчих цін дозволило автору дослідити масштаби інфляційних процесів в Україні. Проведено регіональний аналіз, моделювання індексу споживчих цін за допомогою регресійного аналізу, ARIMA та ETS моделей, здійснено короткочасний прогноз індексу інфляції та порівняння точності моделей між собою.

На особливу увагу заслуговує побудована в роботі економетрична модель індексу споживчих цін, що дозволило визначити взаємозв'язки між досліджуваними показниками і виявити фактори, які мають статистично значимий вплив на індекс споживчих цін. Для прогнозування індексу споживчих цін було використано регресійний аналіз, ARIMA та ETS моделі, порівняння точності яких дозволило обрати найкращу модель для прогнозу індексу інфляції. Кластерний аналіз регіонів України за допомогою карт Кохонена займає чільне місце. Отримані результати демонструють глибоке володіння Музикою І. методикою статистичного дослідження та моделювання.

Для виконання всіх розрахунків застосовувались програмне середовище R-studio, Deductor Studio Academic та програма MS Excel. Інформаційною базою слугували статистичні дані Національного банку, Державної служби статистики та Міністерства фінансів України.

Робота логічна та послідовна. Представлений матеріал, послідовність його викладання, глибина розкриття теми, аргументованість висновків дозволяють позитивно оцінити авторський внесок у розробку практичного інструментарію щодо аналізу та моделювання індексу інфляції в країні.

До незначних недоліків варто віднести недостатню обґрунтованість системи показників в оцінюванні впливу на індекс споживчих цін, однак це не знижує загальний рівень проведеного дослідження.

Загалом, за своїм спрямуванням та змістом бакалаврська робота відповідає вимогам, що висуваються до підготовки робіт ОС бакалавр, рекомендується до захисту і заслуговує на високу оцінку.

Рецензент:

Канд.техн.наук, доцент

доцент кафедри системного аналізу

та інформаційних технологій

Маріупольського державного

університету



(підпис)

Г.В. МАРТИНЮК

07 червня 2024 року

Особистий підпис Г.В. Мартинюк, доцент, зав. кафедр

САП МДУ, Мартинюк Г.В. засвідчую

факт без

Начальник загального відділу Анастасія БЕЗУЛА

ЗМІСТ

ВСТУП.....	3
РОЗДІЛ 1. ТЕОРЕТИЧНІ ТА МЕТОДИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ ІНДЕКСУ СПОЖИВЧИХ ЦІН.....	5
1.1 Теоретичні аспекти дослідження індексу споживчих цін як основного показника інфляції в країні.....	5
1.2 Характеристика сучасних методів розрахунку індексу інфляції.....	7
1.3 Опис методів та інструментів, що використовуються для дослідження індексу інфляції	10
РОЗДІЛ 2. РЕГІОНАЛЬНИЙ АНАЛІЗ ІНДЕКСУ СПОЖИВЧИХ ЦІН.....	15
2.1 Характеристика рівня інфляції за регіонами України	15
2.2 Кластерний аналіз регіонів за ІСЦ за допомогою карт Кохонена.....	17
РОЗДІЛ 3. МОДЕЛЮВАННЯ ІНДЕКСУ СПОЖИВЧИХ ЦІН ЯК ОСНОВНОГО ПОКАЗНИКА ІНФЛЯЦІЇ В КРАЇНІ	27
3.1 Інформаційна база моделювання та розвідувальний аналіз даних	27
3.2 Моделювання ІСЦ за допомогою регресійної, ARIMA та ETS моделей ..	30
3.3 Прогнозування індексу інфляції	49
3.4 Оцінка точності побудованих моделей	57
ВИСНОВКИ	59
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	61
ДОДАТКИ	66

ВСТУП

Актуальність теми. Індекс споживчих цін є важливим показником, який демонструє рівень цін на найпоширеніші товари та послуги, які купує населення в країні для невиробничого споживання. Цей показник є основним індикатором інфляції, тому його вивчення та прогнозування є важливим як для державних установ, банків, так і для населення. Для державних і фінансових установ це важливо для створення прогнозування та побудови виваженої політики на наступні періоди, а для населення – для розуміння очікуваного рівня цін.

Аналіз останніх досліджень і публікацій. У багатьох наукових роботах вітчизняних та зарубіжних вчених досліджено індекс споживчих цін. Зокрема моделювання індексу інфляції розглянуто у працях М. Когтева та В. Новик, де використовувались моделі регресії. У роботах О. Катуніної та Н. Шаповаленко застосовувались різні варіації та модифікації ARIMA моделей. Незважаючи на те, що є багато досліджень індексу споживчих цін, проблема прогнозування даного показника залишається не до кінця дослідженою, враховуючи постійні зміни та потрясіння в економіці України за останні роки. Регіональний аналіз індексу інфляції відображено в працях М. Мазур, а карти Кохонена для проведення кластерного аналізу використовуються у роботах А. Матвійчука та В. Смісова.

Метою виконання кваліфікаційної бакалаврської роботи є проведення регіонального аналізу, моделювання індексу споживчих цін за допомогою регресійного аналізу, ARIMA та ETS моделей, здійснення на їх основі короткочасного прогнозу та порівняння точності моделей між собою.

Для досягнення поставленої мети необхідно виконати наступні завдання:

- 1) узагальнити теоретичні аспекти та охарактеризувати сучасні методи розрахунку індексу споживчих цін;
- 2) дослідити закономірності та тенденції розвитку індексу споживчих цін;
- 3) провести регіональний аналіз індексу споживчих цін;

- 4) побудувати моделі регресії, ARIMA та ETS моделі для прогнозування індексу споживчих цін;
- 5) створити короткостроковий прогноз на основі трьох моделей;
- 6) оцінити точність побудованих економетричних моделей.

Об'єкт дослідження – індекс споживчих цін в Україні.

Предмет дослідження – методи та моделі машинного навчання, що застосовуються для дослідження індексу споживчих цін.

Основними методами дослідження, які використовуються в роботі, є карти Кохонена, кореляційний, розвідувальний, регресійний аналізи, а також ARIMA та ETS моделі. Практична цінність бакалаврської роботи полягає у аналізі індексу інфляції та перевірці кількох моделей з метою вибору найкращої для прогнозування даного показника, що дозволяє в подальшому використовувати її у реальних установах для планування монетарної політики, регулювання рівня видатків державного бюджету тощо.

Бакалаврська робота виконується на матеріалах сайтів Національного банку України, Державної служби статистики України, Міністерства фінансів України та інших відкритих джерел.

Робота складається зі вступу, трьох розділів, висновків, переліку використаних джерел та додатків. Обсяг кваліфікаційної роботи складає 60 сторінок. Робота містить 51 рисунок, 4 таблиці, 5 формул та 38 літературних джерел.

РОЗДІЛ 1

ТЕОРЕТИЧНІ ТА МЕТОДИЧНІ ОСНОВИ ДОСЛІДЖЕННЯ ІНДЕКСУ СПОЖИВЧИХ ЦІН

1.1 Теоретичні аспекти дослідження індексу споживчих цін як основного показника інфляції в країні

Інфляція є одним з ключових економічних показників, що відображає загальне зростання цін на товари та послуги протягом певного періоду часу. Інфляція з латині перекладається як роздуття та довготривале підвищення цін, що відображає купівельну спроможність грошей [1].

В цілому, підвищення рівня цін на деякі товари – ще не вважається ознакою інфляції, бо ціни на інші товари можуть залишатись стабільними або падати. Інфляція означає підвищення цін на більшість товарів та послуг, тобто збільшується середнє значення рівня цін, і такі зміни вже відчуває населення країни [2].

Розрізняють різні види інфляції. За темпами зростання виділяють помірну, галопуючу та гіперінфляцію.

Повзуча (помірна) інфляція спостерігається тоді, коли в країні рівень цін зростають поступово, не більше 10% на рік [1; 2]. Такий вид інфляції є відносно стабільним, таке зростання рівня цін спонукає людей більше заощаджувати кошти, бо їхня купівельна спроможність мало знецінюється [1].

Галопуюча інфляція – більш значне зростання цін у порівнянні з помірною, для такого виду інфляції характерно збільшення цін на рік в межах від 10% до 50% [1]. Така інфляція є досить важкою для населення, так як ціни на товари та послуги досить сильно зростають [2].

Гіперінфляція – темпи зростання цін обчислюються сотнями, а то й тисячами відсотків на рік [1]. Даний вид інфляції є найбільш небезпечним, бо ціни на товари та послуги можуть зростати на тисячі відсотків за рік [2].

Причини виникнення інфляції можуть бути різними, але виділено декілька основних видів:

- Інфляція попиту, така причина інфляції може виникати, коли сукупний попит перевищує обсяги виробництва [3]. Тобто населення має більше коштів, які бажає витратити, але не має достатньої кількості товарів та послуг на ринку [3].
- Інфляція пропозиції – виникає тоді, коли відбувається збільшення витрат на виробництво, це може бути збільшення цін на оплату праці, сировину, подорожчанням імпорту тощо [3].
- Фінансові (монетарні) та інші фактори інфляції – такі причини можуть виникати, коли збільшуються видатки державного бюджету, які не підкріплені зростання пропозиції товарів та послуг [3].
- Ще одним досить цікавим фактором є інфляційні очікування. Дана причина може виникати тоді, коли населення в очікування інфляції намагаються підвищити свої ціни на товари та послуги, для того, щоб запобігти втрат у подальшому [3]. І якщо так робить хоча б частина населення, то ціни будуть зростати, і владі буде необхідно докласти зусиль, щоб стабілізувати ціни [3].

Індекс споживчих цін (ІСЦ, або англійською CPI) є офіційним індексом інфляції в країні і є показником, що відображає зміни загального рівня цін на товари та послуги, які купує населення для невиробничого споживання [4].

ІСЦ є важливим показником, що характеризує інфляційний процес в економіці країни [5]. Його використовують, щоб вирішити багато питань держаної політики прогнозування та аналізу цінових процесів в країні, перегляду рівнів грошових доходів та встановлення мінімальних соціальних гарантій для населення [5]. Він також застосовується при вирішенні правових спорів, і є важливим інструментом для перерахунку показників системи національних рахунків у постійні ціни [5].

Національний банк України використовує ІСЦ для того, щоб визначити рівень інфляції та допомагає скорегувати процентні ставки, з метою контролю грошової маси та стримання інфляційного тиску [6].

ІСЦ використовують для перерахунку макроекономічних показників, наприклад реального ВВП. Такі розрахунки дозволяють провести порівняння економічної діяльності за різні періоди з врахування зміни рівня цін [7].

У рамках державного статистичного спостереження (ДСС) Державна служба статистики формує такі показники [8]:

1) ІСЦ, у % – показник, що відображає зміни загального рівня цін на товари та послуги, які споживає населення, протягом певного періоду [8]. Він є показником, який вимірює зміну вартості фіксованого набору споживчих товарів та послуг поточного періоду і порівнянні з базисним [8].

2) Базовий індекс споживчих цін (БІСЦ), у % – субіндекс ІСЦ, який відображає інфляцію, викликану попитом, і найточніше демонструє ефекти від зміни реальних монетарних умов [8]. Він характеризує динаміку цін, мінімізуючи короткострокові нерівномірні зміни, спричинених адміністративним регулюванням та шоками пропозиції [8].

3) Середні споживчі ціни на товари (послуги) (у гривнях) – розраховуються на основі фактичних цін, які сплачує споживач в моменті придбання продуктів [8].

Окрім вищенаведених індексів є багато індексів цін, які розраховують у різних галузях, наприклад: індекс цін виробників промислової продукції, індекс цін у будівництві, індекс цін на житло, індекс цін на вантажні перевезення залізничним транспортом тощо. Однак ІСЦ вважається основним показником інфляції в Україні.

1.2 Характеристика сучасних методів розрахунку індексу інфляції

ІСЦ є незамінним інструментом для управління економікою країни. В аналізі та передбаченні інфляційних тенденцій важливим аспектом є використання ефективних методів розрахунку індексу інфляції.

Збирання та обробка даних для розрахунку ІСЦ вимагає високої точності та є досить складним, але ІСЦ є важливим показником, який використовується у

різних сферах економіки. Дані індексу інфляції, також, використовуються для наукових досліджень та аналітичних робіт, для прикладу як дана кваліфікаційна робота.

Для розрахунку ІСЦ у кожній країні є свій споживчий кошик, в який входять різні товари та послуги, які відрізняються в залежності від країн [2]. В споживчому кошику (наборі) українського населення міститься 335 товарів та послуг, це вдвічі менше ніж споживчий набір населення Німеччини, Швеції і мешканців інших більш розвинених країн [2].

На основі споживчого кошика, який є єдиним для всіх регіонів країни, здійснюються моніторинг за змінами цін та тарифів, це дає можливість розраховувати ІСЦ [8]. Один раз на 5 років переглядається споживчий набір в Україні [8]. Якщо відбуваються певні значні зміни на регіональних споживчих ринках та у структурі витрат домогосподарств, період перегляду може бути скорочений [8].

Фахівці територіальних органів Держстату щомісячно збирають інформацію про споживчі ціни та тарифи, шляхом реєстрації цін (тарифів) на продукти зі споживчого кошика [8]. З 1 по 25 число кожного місяця відбувається реєстрація цін (тарифів) у всіх містах [8]. Але для деяких окремих товарів-представників реєстрація цін здійснюється протягом усього місяця [8]. До цього переліку входять товари та послуги, частка, яких становить понад 0,5% від споживчих витрат домогосподарств і ціни на які зазнають значних коливань протягом місяця [8].

Під товаром (послугою)-представником мається на увазі певна сукупність видів товарів та послуг, які мають схожі споживчі властивості та призначення, а також мають схожі тенденції у зміні цін й тарифів [7].

Формування вагової структури (вагових коефіцієнтів, ваг) створюється шляхом поєднання даних щодо структури фактичних споживчих грошових витрат домогосподарств з даними про споживні витрати на кінцеве споживання (за винятком певних витрат) [8]. Вагові коефіцієнти демонструють відносну важливість товарів (послуг)-представників, яка вимірюється їхньою часткою у споживчих грошових витратах домогосподарств, і відображає ступінь впливу зміни

цін кожного товару або послуги на загальний ІСЦ [8]. Агрегація індексів від найнижчого рівня до найвищого здійснюється за допомогою вагових коефіцієнтів [8].

Вагові коефіцієнти обчислюються як частка витрат домогосподарств на кожний представницький товар чи послугу в загальних споживчих витратах у фактичних цінах за цей рік [3]. Кожен товар або послуга має свою частку в розрахунках ІСЦ, яка розраховується з точністю до 0,00001 [8].

На національному рівні середня споживча ціна на товари або послуги в Україні обчислюються на основі середніх регіональних цін за формулою середньої арифметичної зваженої [8]. Ваги – це середньорічна кількість населення кожного з регіонів у відповідних даних по Україні в цілому, помножена на частку споживчих грошових витрат домогосподарств [8].

Об'єкти статистичних спостережень, тобто цільова група – це сукупність юридичних осіб та їх різних підрозділів, що здійснюють свою діяльність у сфері послуг, на ринках, із роздрібною торгівлю, де якраз таки поводиться реєстрація цін, іншими словами тарифів, на товари та послуги для повсякденного вжитку [8].

У межах проведення державного статистичного спостереження використовується комбінація статистичних методів, тобто застосовуються дані обстеження методом несутільного вивчення сукупностей, із залучанням методу відсікання, а також результати проведення інших ДСС [8].

Розрахунки ІСЦ за квартал, період із початку року та інших, здійснюється за допомогою «ланцюгового» методу, тобто шляхом множення місячних індексів споживчих цін між собою [8].

ІСЦ на регіональному рівні використовується для аналізу зміни цін та інфляційних процесів у різних куточках країни й дає можливість контролювати за розробляти заходи для стримування інфляції.

Узагальнений розрахунок ІСЦ відображено на рис. 1.1:



Рисунок 1.1 – Послідовність розрахунку ІСЦ в Україні

Джерело: розроблено автором на основі [9]

В даній роботі для побудови моделей використано квартальні індекси споживчих цін, які розраховувались добутком помісячних даних ІСЦ (% до попереднього місяця). База даних щодо ІСЦ в цілому по країні за 2013-2023 роки, а також в розрізі окремих її регіонів за 2019-2023 роки сформована на основі даних Держстату [10; 11].

1.3 Опис методів та інструментів, що використовуються для дослідження індексу інфляції

Для побудови карт Кохонена використано аналітичну платформу Deductor Studio Academic [12]. Це програмне забезпечення спрямоване на аналіз та обробку даних [12].

Більшість розрахунків та візуалізації були проведені у середовищі RStudio, за допомогою мови програмування R. R – це мова програмування, яку можна використовувати для статистичних досліджень, аналізу, візуалізації даних тощо [13]. R має потужні функції та пакети, які надають можливість для багатьох статистичних методів, наприклад різні види регресії, ARIMA та ETS моделі, різні статистичні тести, кластеризація і безліч іншого [13].

Також, для множинної лінійної регресійної моделі розраховані коефіцієнти еластичності та для деяких показників здійснено прогноз за допомогою ковзного середнього у середовищі Microsoft Excel.

Для аналізу та моделювання індексу споживчих цін використовуються наступні методи:

1. Карти Кохонена

Карти Кохонена, їх інша назва карти самоорганізації, вони являються методом нейронних мереж. Карти самоорганізації були розроблені Тейво Кохоненом, фінським ученим у 1980-х роках [14; 15]. Ці карти використовують метод навчання без учителя, звідки і взяли свою другу назву [14]. Ці карти складаються з нейронів, які пов'язані з нейронами вхідного шару [14]. Їхньою основною відмінністю від інших нейронних мереж є те, що вони не мають вихідного шару [14]. Алгоритм карт Кохонена складається з кількох наступних кроків, а саме: на нейрони надходять дані (вхідний вектор), далі обирається такий нейрон, який є найближчим до вхідного вектора, цей обраний нейрон називають «нейроном-переможцем» [14]. Далі здійснюється корекція «нейрона-переможця» та інших сусідніх до нього нейронів, які коректують свої ваги в сторону вхідних даних враховуючи коефіцієнт темпу навчання [14]. Такий процес повторюють для всіх даних вхідного вектора, такі кроки дозволяють адаптуватись нейронам карти до структури даних. У результаті отримано проекцію багатовимірного простору на двовимірну карту [14; 15]. Кластеризація та візуалізації великих масивів даних є основним призначенням карт Кохонена [14; 16; 17]. Кластеризація (кластерний аналіз) – це групування набору об'єктів на менші групи, під назвою – кластери,

таким чином, щоб об'єкти в кожному кластері були схожими між собою, а об'єкти різних груп відрізнялись [18].

2. Кореляційний аналіз – це метод вимірювання та оцінки того, як різні статистичні змінні пов'язані між собою [19], тобто він використовується для визначення взаємозв'язку між кількома змінними і показує наскільки сильним є зв'язок [20]. У даній роботі цей аналіз використаний для зменшення ймовірності виникнення мультиколінеарності у моделі регресії.

3. Розвідувальний аналіз – дозволяє провести попередній аналіз даних, шляхом представлення даних у графічному табличному вигляді, схемах, діаграмах тощо [21].

4. Регресійний аналіз – це метод допомагає зрозуміти взаємозв'язок між однією залежною змінною та однією або декількома незалежними змінними [22]. За допомогою такого аналізу можна дослідити як зміни у незалежних змінних пов'язані зі змінами в залежній змінній [22; 23; 24].

5. ARIMA модель складається з наступних частин:

- AR: авторегресія (лагові значення як вхідні дані) [25]
- I: інтегрування (диференціювання для стаціонарності) [25]
- MA: ковзна середня (лагові помилки як вхідні дані) [25]

Модель ARIMA рідко можна інтерпретувати з точки зору структур даних, таких як тренд і сезонність [25]. Це комплексна модель і може опрацьовувати багато видів, шаблонів часових рядів [26; 27]. Таку модель використовують до стаціонарних часових рядів. Стаціонарні часові ряди, це ті, для яких виконується 3 принципи:

- Приблизно горизонтальний ряд, тобто у ньому немає тренду, тобто він розвивається горизонтально [25].
- Постійно дисперсія [25].
- Відсутність передбачуваних патернів поведінки, наприклад сезонності чи чогось іншого [25].

Тобто, в цілому, ряд має виглядати приблизно як «шум», тобто без певного тренду або повторюваних коливань. Для того, щоб забезпечити стаціонарність часового ряду, далі будуть проведені тести на діагностику того чи варто проводити трансформації часового ряду. Для того щоб ідентифікувати, яким є часовий ряд стаціонарним чи ні, потрібно побудувати графік часового ряду, який варто аналізувати візуально, далі можна побудувати корелограму автокореляції (ACF), в цілому для стаціонарних часових рядів значення автокореляції швидко падають до 0, тобто декілька перших значень (лагів) є статистично значущими, а наступні значення вже не є значущими. А для нестационарних рядів – значення автокореляції дуже повільно наближаються до 0.

Наступним етапом дослідження є опис компонентів ARIMA моделі.

Авторегресійна модель (AR) (цій частині у моделі ARIMA відповідає параметр p): авторегресія (лагові значення використовуються як вхідні дані), тобто будується модель залежності від попередніх значень нашого часового ряду.

Модель ковзного середнього (MA) - модель авторегресії, яка шукає залежність від ε_t – білого шуму. Генерується, якусь випадкову величину білого шуму з якимось параметрами середнього і стандартного відхилення, і від них знаходиться ця модель авторегресії. Тобто, це множинна регресія залежності від помилок як предикторів. Параметром MA в ARIMA є q .

Авторегресійна модель ковзного середнього це – ARMA модель. Тобто об'єднується AR - модель авторегресії залежності від реальних даних та MA - залежність від білого шуму. Отже це є ARMA модель. Якщо до цієї моделі додати ще частинку I (інтегрування), то сформується модель ARIMA. Інтегрування (I) (у моделі ARIMA цій частині відповідає параметр d) - говорить про те, скільки різниць потрібно взяти до часового ряду, щоб він став стаціонарним.

ARIMA(p,d,q) моделі, складові [25]:

1. AR: p = порядок авторегресійної частини [25],
2. I: d = порядок різниці [25],
3. MA: q = порядок ковзної середньої [25].

У роботі використано автоматична модель ARIMA, тобто програма сама підбирає найбільш доцільні параметри p, d, q .

6. ETS модель

ETS (Exponential smoothing) модель або інша назва моделі – експоненціального згладжування [28]. Дана модель використовується для прогнозування часових рядів [28]. Особливістю цієї моделі є те, що даний метод в якому «давні», тобто більш старі значення змінної мають менший вплив на прогнозовані значення, а ніж ті, що є більш новими [29]. Тобто, це дозволяє моделі врахувати історичну поведінку даних, при цьому більше спираючись на нещодавні зміни поведінки часового ряду [29]. Компонентами моделі є:

- Error (Похибки) – дана компонента відповідає за певні випадкові зміни в даних, які не описуються трендом або сезонністю. Похибки можуть бути адитивними «A» або мультиплікативними «M» [30; 31].
- Trend (Тренд) – ця компонента вказує на певну тенденцію в часовому ряді. Тренд у моделі може бути відсутній «N», адитивний «A», мультиплікативний «M» та демпфований «Ad» [30; 31].
- Seasonal (Сезонність) – враховує певні зміни у даних що повторюються з певною періодичністю. Сезонна компонента може мати наступні значення: відсутній «N», адитивний «A» та мультиплікативний «M» [30; 31].

У рівняннях моделі ETS використовуються наступні параметри [32]: α – задає зміну рівнів часового ряду, β – описує зміну тренду, γ – демонструє зміну сезонності. Ці параметри змінюються у наступних межах: $\alpha \in [0,1)$, $\beta \in [0,\alpha)$, $\gamma \in [0,1-\alpha)$ [30]. Такі обмеження для гіперпараметрів потрібні для того, щоб «старіші» значення мали менший вплив на прогнози. Також, у ETS модель задає певні початкові значення кожного компонента, на основі всіх спостережень часового ряду, модель автоматично обраховує початкові значення: рівня (ℓ_t), тренду (b_t) та сезонності (s_t). Декілька з перелічених методів доцільно застосувати для проведення регіонального аналізу індексу споживчих цін.

РОЗДІЛ 2

РЕГІОНАЛЬНИЙ АНАЛІЗ ІНДЕКСУ СПОЖИВЧИХ ЦІН

2.1 Характеристика рівня інфляції за регіонами України

Україна є досить неоднорідною країною з точки зору соціально-економічного розвитку та географічних умов. Ця різноманітність свідчить про значний економічний потенціал регіонів, адже кожен з них має унікальні характеристики, що впливають на індекс споживчих цін. Тому актуальним завданням є аналіз регіональних особливостей на основі офіційної статистики України. Різні регіони мають свої пріоритетні напрями розвитку, кліматичні умови, технологічний рівень та інноваційні можливості. Структурні зрушення, розвиток та оновлення економіки країни залежать від змін в ІСЦ у різних регіонах [33].

Порівняння значень ІСЦ за регіонами за 2023 рік представлено на рис. 2.1.

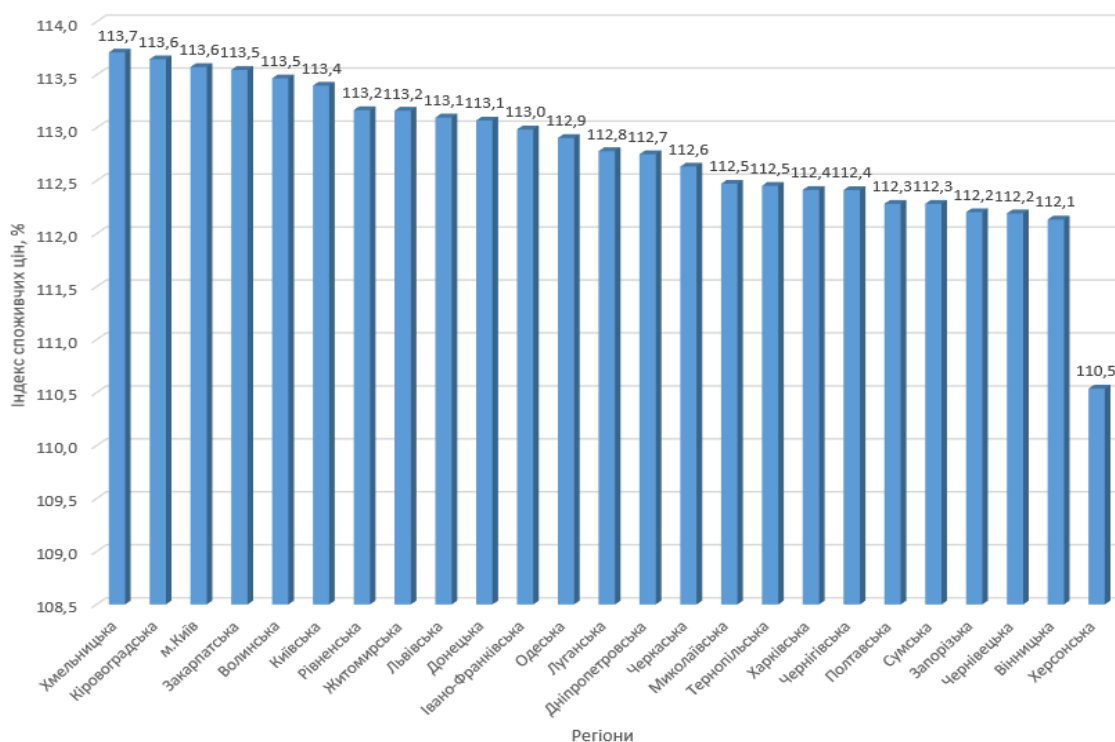


Рисунок 2.1 – Індекс споживчих цін за регіонами України за 2023р.

Джерело: розроблено автором за даними [8]

Рис. 2.1 відображає ІСЦ (у % до попереднього року) для різних регіонів України. На графіку помітні суттєві відмінності в значеннях індексу інфляції між регіонами. Таким чином, у 2023 році до трійки регіонів з найвищими значеннями ІСЦ входили Хмельницька, Кіровоградська області та м. Київ, до регіонів з найнижчими значеннями ІСЦ входили Чернівецька, Вінницька та Херсонська області. Такі відмінності можуть бути спричинені різним впливом військового стану на економіку регіонів, різні економічні умови та вплив місцевих факторів на інфляцію в кожному регіоні.

На рис. 2.2 представлено візуалізацію середніх значень ІСЦ по регіонах України за 2019-2023 роки. Аналіз показав, що середні значення ІСЦ за 5 років (рис. 2.1), значно відрізняються від результатів аналізу 2023 року (рис. 2.1)

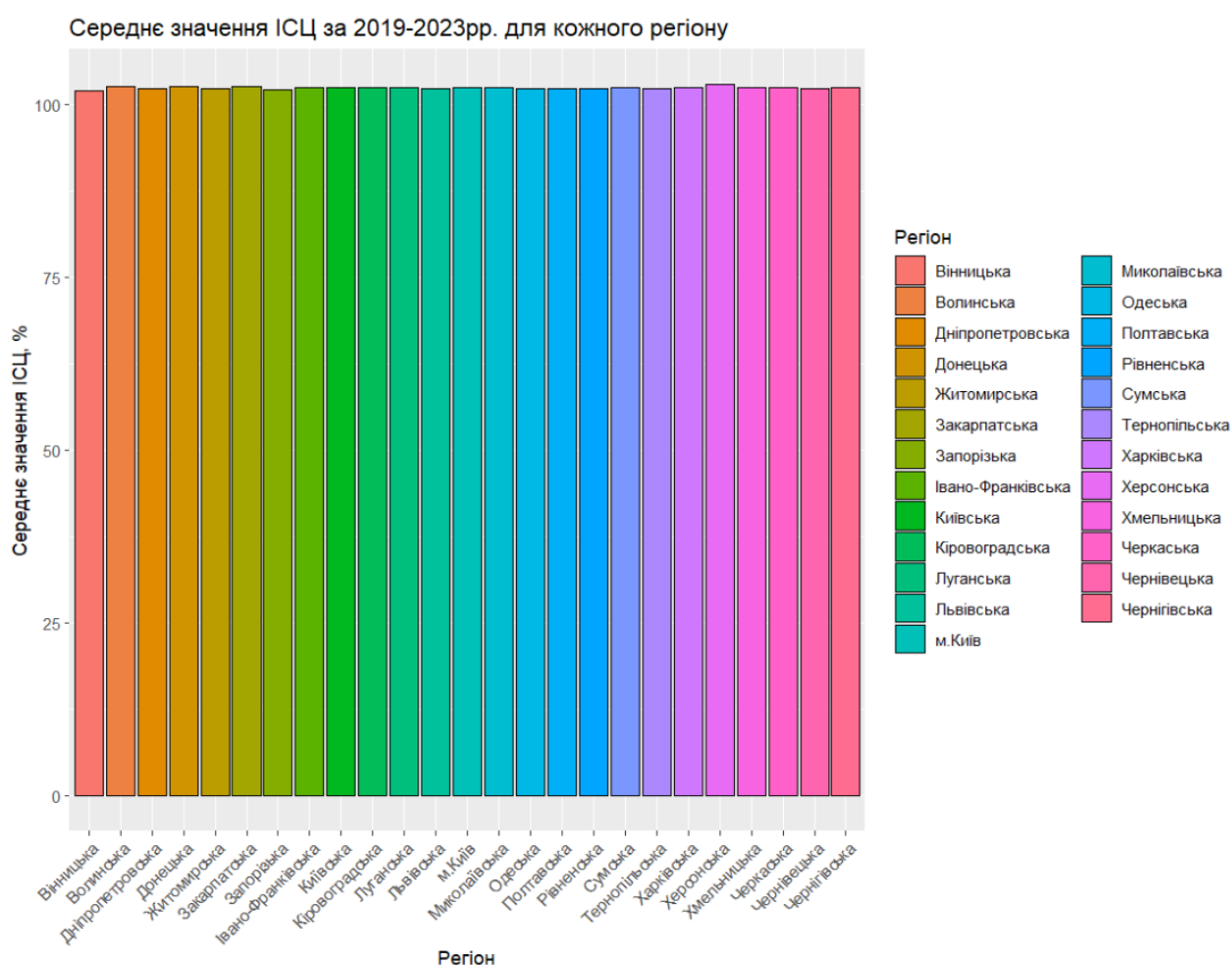


Рисунок 2.2 – Середнє значення ІСЦ регіонів України за 2019-2023рр.

Джерело: розроблено автором за даними [11]

За даними рис. 2.2 можна зробити висновок, що середнє значення ІСЦ по регіонах за 2019-2023 роки залишають приблизно на одному рівні у середньостроковій перспективі, тобто за 5 років. Можна сказати, що зовнішні умови та потрясіння економіки, такі як воєнний стан та пандемія COVID-19, мінімально впливають на відмінності ІСЦ за регіонами країни за останні п'ять років.

2.2 Кластерний аналіз регіонів за ІСЦ за допомогою карт Кохонена

Наступним кроком у даній роботі виконана кластеризація регіонів України за допомогою карт самоорганізації, а саме карти Кохонена. Даний крок допоможе визначити, у який регіонах присутній найбільший індекс споживчих цін, тобто у який регіонах більший рівень інфляції, а у яких зростання цін на товари кінцевого споживання є низьким. Цей метод також дозволить виявити, які регіони є подібними за поведінкою ІСЦ.

Варто зазначити, що для побудови самоорганізаційної карти використані поквартальні дані індексу споживчих цін у відсотках до попереднього місяця. Їх розраховано шляхом перемноження місячних значень ІСЦ. Дані були взяті з джерела [11]. Виключено дані по м. Севастополь та Автономної Республіки Крим. Значення за цими регіонами були виключені, оскільки містили пропущені значення.

Отже, у роботі побудовано три карти Кохонена за три роки. Одна карта за 2019 рік, який є допандемічним періодом, друга карта за 2021 рік – пандемічний період та третя карта за 2023 – період воєнного стану.

У Deductor Studio Academic використано функцію «Майстер обробки», далі необхідно обрати побудову Карти Кохонена, наступним кроком є визначення, які стопці є вхідними та інформаційними, так як на меті немає побудови класифікації, то і вихідних змінних немає. При розбитті на навчальну та тестову вибірки, першій віддано всі 100% даних, так як поставлена задача саме у класифікації, нічого тестуватись і перевірятись на точність не буде. Кількість епох навчання встановлено

на рівні 5000. Функцію сусідства (функція наближення) встановлено Гауссова. Також, варто зазначити, що кількість кластерів встановлена автоматично. Отже, наступним кроком побудовані карти Кохонена за 3 періоди.

Карта Кохонена допандемічного періоду зображена на рис. 2.3.

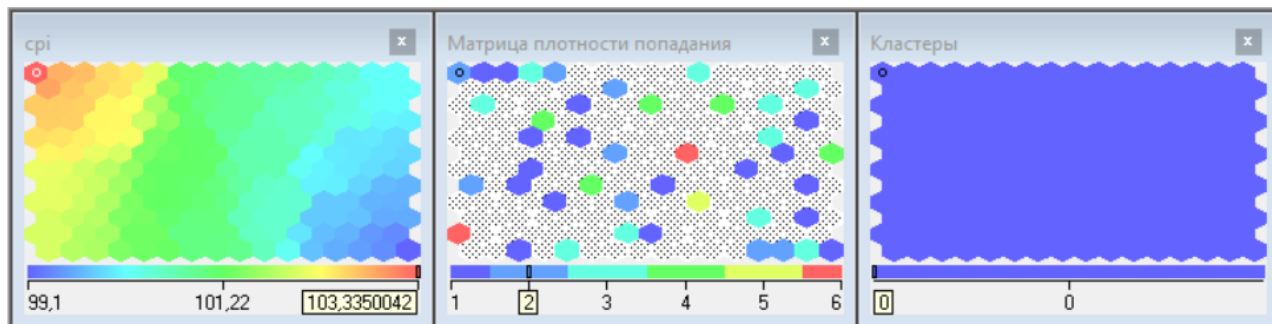


Рисунок 2.3 – Карта Кохонена для допандемічного періоду (2019 рік)
розподіл за регіонами

Джерело: розроблено автором за даними [11]

На рис. 2.3 спостерігається, що автоматично побудовано лише один кластер. Це означає, що значення ІСЦ по регіонах за допандемічний період значно не відрізняються один від одного. Аналізуючи карту Кохонена у комірці, яка знаходиться у лівому верхньому куті значення індексу інфляції становить 103,33% та 103,34% у м. Києві та Харківської області відповідно. Це максимальні значення, які присутні за аналізований період. У протилежному куті карти (нижньому правому) знаходяться відповідно мінімальне значення ІСЦ – 99,1% у Херсонській області. Таку тенденцію, щодо розподілу регіонів за індексом інфляції можна спостерігати на першій частині рис. 2.3. У другій частині рис. 2.3 матриця щільності попадання, де кольорами відображено кількість регіонів, які потрапили у ту чи іншу комірку.

Профіль кластеру для допандемічного періоду представлено на рис. 2.4.

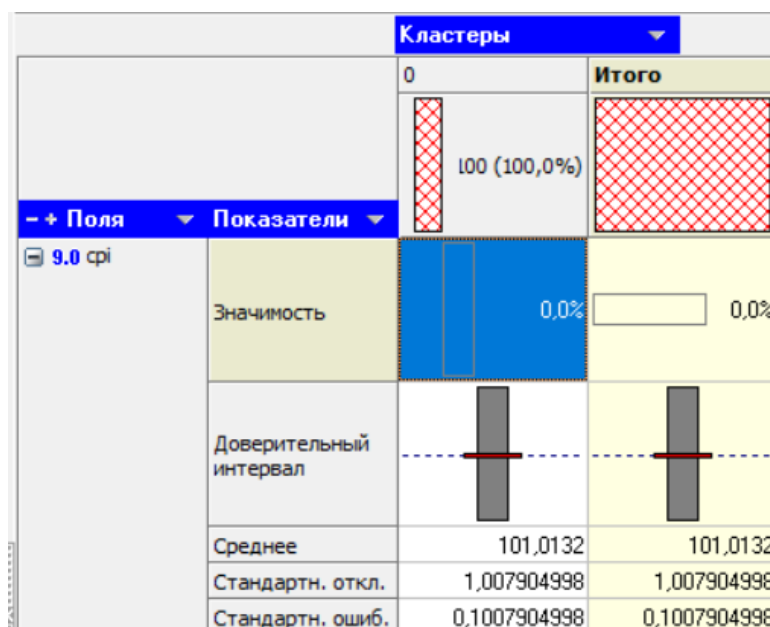


Рисунок 2.4 – Профіль кластера допандемічного періоду

Джерело: розроблено автором на основі [11]

На рис. 2.4 можна спостерігати, що середнє значення ІСЦ за регіонами, які містяться у кластері для допандемічного періоду, становить 101% і у нього входять всі 100% спостережень.

У таблиці 2.1 наведено назви регіонів та номери кварталів, які входять до побудованого кластеру. У таблиці 2.1 відображено регіони та квартали за допандемічний період, які входять в побудований кластер. Тобто, в цілому, спостерігається, що всі аналізовані регіони за чотири квартали 2019 року входять в один кластер.

Таблиця 2.1 – Опис кластерів за допандемічний період

Номер кластеру	Допандемічний період (2019 рік)
0	Вінницька (1-4 кв.), Волинська (1-4 кв.), Дніпропетровська (1-4 кв.), Донецька (1-4 кв.), Житомирська (1-4 кв.), Закарпатська (1-4 кв.), Запорізька (1-4 кв.), Івано-Франківська (1-4 кв.), Київська (1-4 кв.), Кіровоградська (1-4 кв.), Луганська (1-4 кв.), Львівська (1-4 кв.), м.Київ (1-4 кв.), Миколаївська (1-4 кв.), Одеська (1-4 кв.), Полтавська (1-4 кв.), Рівненська (1-4 кв.), Сумська (1-4 кв.), Тернопільська (1-4 кв.), Харківська (1-4 кв.), Херсонська (1-4 кв.), Хмельницька (1-4 кв.), Черкаська (1-4 кв.), Чернівецька (1-4 кв.), Чернігівська (1-4 кв.).

Джерело: розроблено автором на основі [11]

Карта Кохонена за 2021 рік, тобто пандемічний період, представлена на рис.2.5.

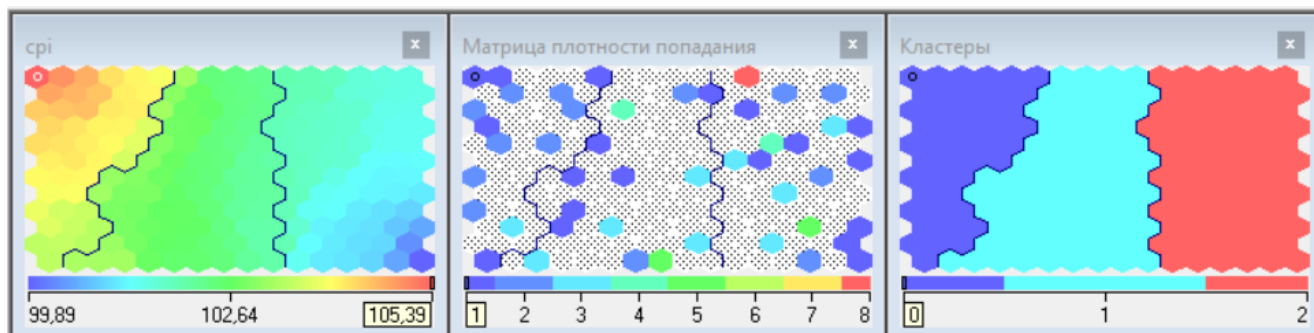


Рисунок 2.5 – Карта Кохонена за пандемічний період розподіл за регіонами

Джерело: розроблено автором на основі [11]

За пандемічний період отримано наступні результати (рис. 2.5). Спостерігається поділ регіонів на три кластери за ІСЦ. Такий розподіл демонструє, що регіони досить сильно відрізняються між собою за значенням індексу інфляції.

Для пандемічного періоду профілі кластерів представлено на рис. 2.6.

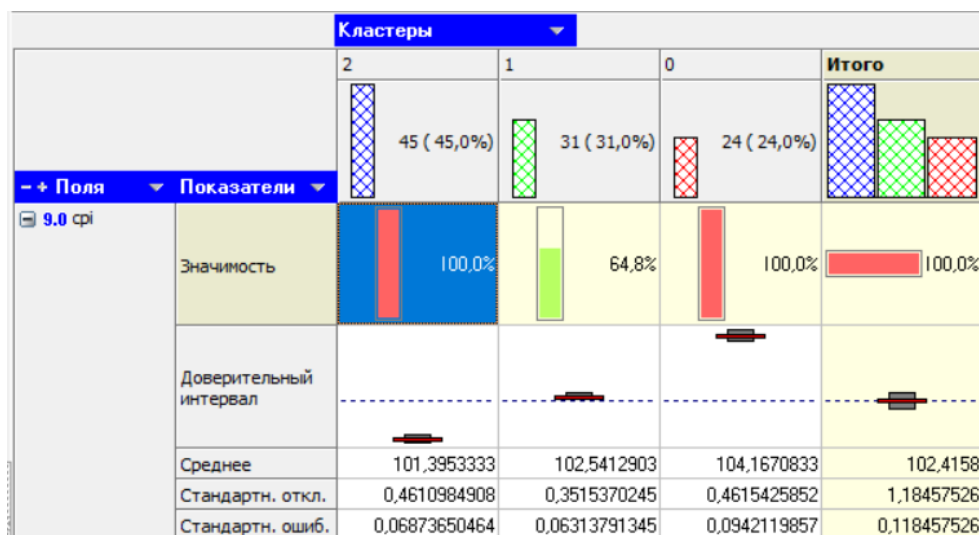


Рисунок 2.6 – Профілі кластерів пандемічного періоду

Джерело: розроблено автором на основі [11]

За даними рис. 2.6 отримано результати, що у нульовому кластері середнє значення ІСЦ по регіонах, які входять у цей кластер, становить 104,2%, у кластері під номером 1 середнє значення індексу інфляції дорівнює 102,5%, а у кластерів №2 – 101,4%. Такі середні значення ІСЦ по кластерах вказують на те, що за

пандемічний період регіони досить сильно відрізнялись за аналізованим показником. Варто зазначити, що у нульовий кластер входить 24% спостережень, у кластер під номером 1 входить 31% спостережень за аналізований пандемічний період, а кластер №2 складається з найбільшої кількості спостережень – 45%.

У таблиці 2.2 наведено розподіл регіонів за ІСЦ, які входять у відповідні кластери за 2021 рік, тобто за період пандемії COVID-19.

Таблиця 2.2 – Опис кластерів за регіонами у 2021 році (пандемічний період)

Номер кластеру	Пандемічний період (2021 рік)
0	Волинська (1 кв.), Дніпропетровська (1 кв.), Донецька (1 кв.), Житомирська (1 кв.), Закарпатська (1 кв.), Запорізька (1 кв.), Івано-Франківська (1 кв.), Київська (1 кв.), Кіровоградська (1 кв.), Луганська (1 кв.), Львівська (1 кв.), м.Київ (1 кв.), Миколаївська (1 кв.), Одеська (1 кв.), Полтавська (1 кв.), Рівненська (1 кв.), Сумська (1 кв.), Тернопільська (1 кв.), Харківська (1 кв.), Херсонська (1 кв.), Хмельницька (1 кв.), Черкаська (1 кв.), Чернівецька (1 кв.), Чернігівська (1 кв.).
1	Вінницька (1 кв.), Донецька (2 кв.), Запорізька (2 кв.), Івано-Франківська (2 кв.), Луганська (2 кв.), м.Київ (2 кв.), Одеська (2 кв.), Тернопільська (2 кв.), Харківська (2 кв.), Черкаська (2 кв.), Полтавська (3 кв.), Вінницька (4 кв.), Волинська (4 кв.), Дніпропетровська (4 кв.), Донецька (4 кв.), Житомирська (4 кв.), Закарпатська (4 кв.), Івано-Франківська (4 кв.), Київська (4 кв.), Кіровоградська (4 кв.), Луганська (4 кв.), Львівська (4 кв.), Миколаївська (4 кв.), Одеська (4 кв.), Рівненська (4 кв.), Сумська (4 кв.), Тернопільська (4 кв.), Харківська (4 кв.), Хмельницька (4 кв.), Чернівецька (4 кв.), Чернігівська (4 кв.).
2	Вінницька (2 кв.), Волинська (2 кв.), Дніпропетровська (2 кв.), Житомирська (2 кв.), Закарпатська (2 кв.), Київська (2 кв.), Кіровоградська (2 кв.), Львівська (2 кв.), Миколаївська (2 кв.), Полтавська (2 кв.), Рівненська (2 кв.), Сумська (2 кв.), Херсонська (2 кв.), Хмельницька (2 кв.), Чернівецька (2 кв.), Чернігівська (2 кв.), Вінницька (3 кв.), Волинська (3 кв.), Дніпропетровська (3 кв.), Донецька (3 кв.), Житомирська (3 кв.), Закарпатська (3 кв.), Запорізька (3 кв.), Івано-Франківська (3 кв.), Київська (3 кв.), Кіровоградська (3 кв.), Луганська (3 кв.), Львівська (3 кв.), м.Київ (3 кв.), Миколаївська (3 кв.), Одеська (3 кв.), Рівненська (3 кв.), Сумська (3 кв.), Тернопільська (3 кв.), Харківська (3 кв.), Херсонська (3 кв.), Хмельницька (3 кв.), Черкаська (3 кв.), Чернівецька (3 кв.), Чернігівська (3 кв.), Запорізька (4 кв.), м.Київ (4 кв.), Полтавська (4 кв.), Херсонська (4 кв.), Черкаська (4 кв.).

Джерело: розроблено автором на основі [11]

Для 2023 року, тобто періоду воєнного стану, карти Кохонена зображено на рис. 2.7.

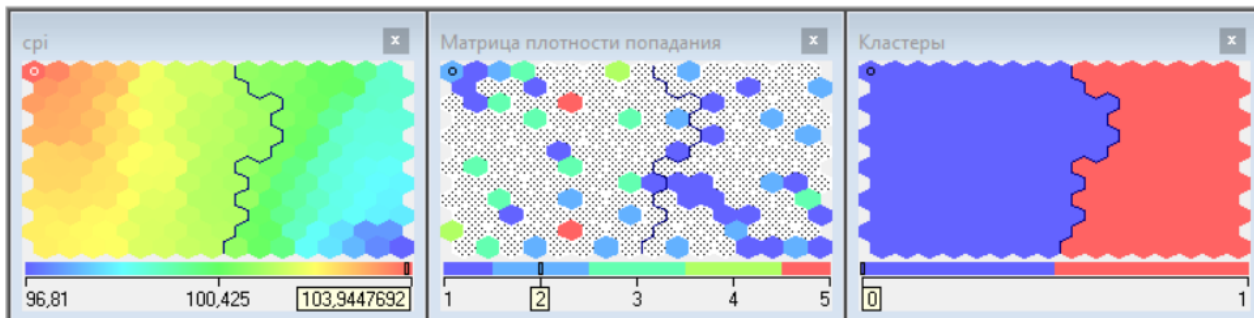


Рисунок 2.7 – Карти Кохонена розподілу ІСЦ за регіонами 2023 рік

Джерело: розроблено автором за даними [11]

За даними рис. 2.7 можна зробити висновок, що за аналізований період воєнного стану регіони були розподілені на два кластери. Суттєва різниця між значеннями ІСЦ за регіонами зменшилась з плином часу, тому спостерігається вже не 3 кластери, як за пандемічний період. Відповідно розрив між різними значеннями індексу інфляції зменшується, вказуючи на те, що воєнний стан більш однорідно впливає на економічну ситуацію у кожному регіоні.

Для воєнного стану профілі побудованих кластерів відтворено на рис. 2.8.

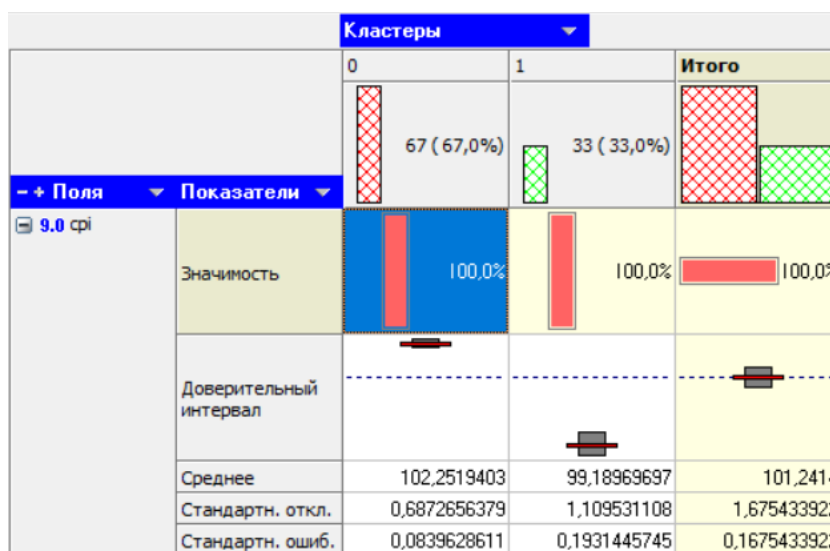


Рисунок 2.8 – Профілі кластерів за період воєнного стану

Джерело: розроблено автором за даними [11]

З рис. 2.8 спостерігається, що кластер під номер 0 містить 67% спостережень аналізованого періоду, а кластер №1 містить 33% спостережень за період воєнного

стану. Середнє значення у нульовому кластері становить 102,3%, а у 1-му кластері – 99,2%. Підсумовуючи стало відомо, що у нульовому кластері містяться регіони, які мають вищі значення ІСЦ, а у першому кластері – нижчі.

У таблиці 2.3 наведено назви та номери кварталів, які входять в побудовані кластери за воєнний період.

Таблиця 2.3 – Опис кластерів за період воєнного стану

Номер кластеру	Воєнний період (2023 рік)
0	Вінницька (1 кв.), Волинська (1 кв.), Дніпропетровська (1 кв.), Донецька (1 кв.), Закарпатська (1 кв.), Запорізька (1 кв.), Івано-Франківська (1 кв.), Київська (1 кв.), Кіровоградська (1 кв.), Луганська (1 кв.), Львівська (1 кв.), м.Київ (1 кв.), Миколаївська (1 кв.), Одеська (1 кв.), Полтавська (1 кв.), Рівненська (1 кв.), Сумська (1 кв.), Тернопільська (1 кв.), Харківська (1 кв.), Хмельницька (1 кв.), Черкаська (1 кв.), Чернівецька (1 кв.), Чернігівська (1 кв.), Волинська (2 кв.), Дніпропетровська (2 кв.), Донецька (2 кв.), Житомирська (2 кв.), Запорізька (2 кв.), Івано-Франківська (2 кв.), Київська (2 кв.), Кіровоградська (2 кв.), Луганська (2 кв.), Львівська (2 кв.), м.Київ (2 кв.), Миколаївська (2 кв.), Одеська (2 кв.), Рівненська (2 кв.), Харківська (2 кв.), Херсонська (2 кв.), Хмельницька (2 кв.), Черкаська (2 кв.), Вінницька (4 кв.), Волинська (4 кв.), Дніпропетровська (4 кв.), Донецька (4 кв.), Житомирська (4 кв.), Закарпатська (4 кв.), Запорізька (4 кв.), Івано-Франківська (4 кв.), Київська (4 кв.), Кіровоградська (4 кв.), Луганська (4 кв.), Львівська (4 кв.), м.Київ (4 кв.), Миколаївська (4 кв.), Одеська (4 кв.), Полтавська (4 кв.), Рівненська (4 кв.), Сумська (4 кв.), Тернопільська (4 кв.), Харківська (4 кв.), Херсонська (4 кв.), Хмельницька (4 кв.), Черкаська (4 кв.), Чернівецька (4 кв.), Чернігівська (4 кв.).
1	Вінницька (2 кв.), Вінницька (3 кв.), Волинська (3 кв.), Дніпропетровська (3 кв.), Донецька (3 кв.), Житомирська (3 кв.), Закарпатська (2 кв.), Закарпатська (3 кв.), Запорізька (3 кв.), Івано-Франківська (3 кв.), Київська (3 кв.), Кіровоградська (3 кв.), Луганська (3 кв.), Львівська (3 кв.), м.Київ (3 кв.), Миколаївська (3 кв.), Одеська (3 кв.), Полтавська (2 кв.), Полтавська (3 кв.), Рівненська (3 кв.), Сумська (2 кв.), Сумська (3 кв.), Тернопільська (2 кв.), Тернопільська (3 кв.), Харківська (3 кв.), Черкаська (2 кв.), Полтавська (3 кв.), Вінницька (4 кв.), Волинська (4 кв.), Херсонська (1 кв.), Херсонська (3 кв.), Хмельницька (3 кв.), Черкаська (3 кв.), Чернівецька (2 кв.), Чернівецька (3 кв.), Чернігівська (2 кв.), Чернігівська (3 кв.).

Джерело: розроблено автором на основі [11]

Наступним кроком проаналізовано матрицю щільності попадання за 2019 рік, тобто допандемічний період, яка наведена на рис. 2.9.

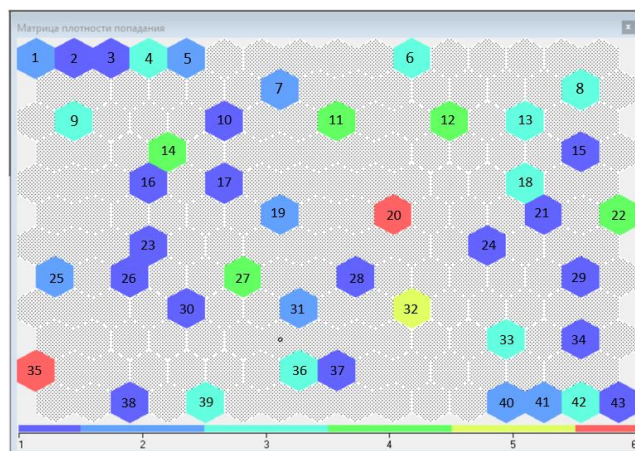


Рисунок 2.9 – Матриця щільності попадання карти Кохонена за 2019 рік
(допандемічний період)

Джерело: розроблено автором за даними [11]

На рис. 2.9 наведено матрицю щільності та номери кожної комірки за 2019 рік (табл. А.1 додатка А) будуть наведено відповідні номери комірок, а також регіони, які входять до кожної комірки. Матриця щільності попадання дає можливість візуально оцінити розподіл регіонів між комірками карти, різними кольорами на карті відображено кількість спостережень, які потрапили у конкретну комірку з зазначеним номером.

На рис. 2.10 відображено матрицю щільності попадання за 2021 рік.

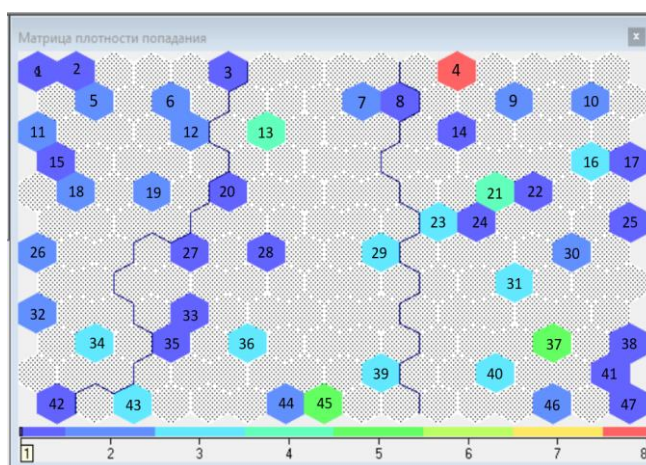


Рисунок 2.10 – Матриця щільності попадання карти Кохонена за 2021 рік
(пандемічний період)

Джерело: розроблено автором за даними [11]

У таблиці А.2 (додаток А) будуть розглянуті комірки з матриці щільності з рис. 2.10.

Наступною варто проаналізувати матрицю щільності для воєнного періоду, яка наведена на рис. 2.11.

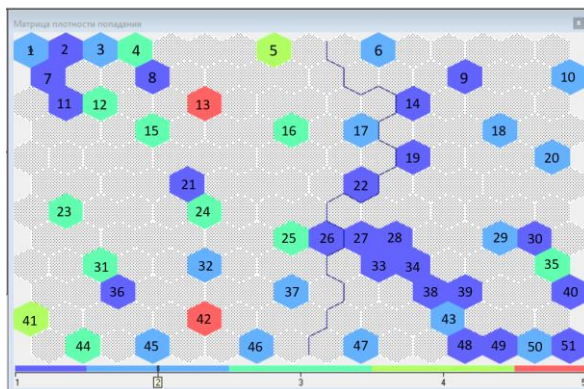


Рисунок 2.11 – Матриця щільності попадання карти Кохонена за 2023 рік

Джерело: розроблено автором за даними [11]

Опис комірок з рис. 2.11 наведено у таблиці А.3 (додаток А). Після здійснення детального аналізу регіонів за 2019, 2021 та 2023 роки, можна зробити висновок, що кластерний аналіз проведений з використання карт Кохонена допоміг виявити регіональні відмінності та схожості в зміні індексу споживчих цін у різні періоди: допандемічний, пандемічний та воєнний. Такий підхід є корисним інструментом для аналізу регіональних економічних умов та виявлення тенденцій, що впливають на інфляцію. Карти Кохонена продемонстрували свою ефективність у виявленні груп регіонів з подібними економічними поведінками, це дозволило глибше розглянути як зміна зовнішніх умов впливає на індекс споживчих цін у різних регіонах.

Допандемічний період, тобто аналіз за 2019 рік, показав, що регіони країни були відносно стабільними в значеннях інфляції. Найвищі значення ІСЦ зафіксовано в Київській та Харківській областях – 103,33% та 103,34% відповідно. І навпаки, найнижче значення ІСЦ спостерігалось в Херсонській області – 99,1%. Такі результати показують, що до початку пандемії COVID-19 ситуація в країні була відносно стабільною. В загальному, можна зробити висновок, що у допандемічний період значення ІСЦ у регіонах України значно не відрізнялись один від одного. У

пандемічний період індекс інфляції мав найбільші розбіжності значень по регіонах, відповідно до цього і побудовано 3 кластери. У воєнний період регіони за ІСЦ поділено всього на 2 кластери, тобто зберігаються розбіжності між регіонами, але все ж вони не є такими великими, як у пандемічний період.

Пандемія 2021 року суттєво змінила ситуацію. Карта Кохонена показує, що ІСЦ групує регіони в три окремі кластери, це свідчить про те, що рівень інфляції в різних регіонах сильно відрізняється. Перший кластер включав регіони з найбільшим середнім ІСЦ (104,2%), другий кластер з середнім ІСЦ 102.5%, і третій кластер з найнижчим середнім ІСЦ (101,4%). Відмінності в інфляції між регіонами в Україні показують, що пандемія вплинула на кожен регіон по-різному.

Аналіз періоду воєнного стану, тобто 2023 рік, продемонстрував, що регіони України розподілились на два кластери. В нульовому кластері середнє значення становить 102,3%, а у першому – 99,2%. Це вказує на те, що розрив у значеннях ІСЦ між регіонами зменшився після пандемії, але ще існують помітні відмінності.

Між комірками матриць щільностей попадання найбільше переміщувались: Київська, Херсонська, Харківська та Полтавська області. Варто зазначити, що високі значення ІСЦ в основному були виявлені у м. Київ, Київській, Харківській та Одеській областях, вони часто демонстрували високі значення у трьох періодах. Відносно низькі значення показували такі регіони, а саме, Херсонська, Закарпатська та Сумська області, такі, значення індексу інфляції можуть бути спричинені невеликим впливом зовнішніх факторів на стабільність цін. Також, можна припустити, що у пандемічний та воєнний періоди існує тенденція до зменшення відмінностей між регіонами за значеннями ІСЦ та поверненням їх до значень, які будуть мало відрізнятись один від одного, звідси слідує, що значення індексу інфляції будуть подібні на більш «спокійні» роки, тобто на допандемічний період. Отже, в цілому можна стверджувати, що кластерний аналіз з використанням карт Кохонена наглядно продемонстрував чутливість індексу споживчих цін, тобто інфляційних процесів в регіонах України, до зміни зовнішніх умов та потрясінь.

РОЗДІЛ 3

МОДЕЛЮВАННЯ ІНДЕКСУ СПОЖИВЧИХ ЦІН ЯК ОСНОВНОГО ПОКАЗНИКА ІНФЛЯЦІЇ В КРАЇНІ

3.1 Інформаційна база моделювання та розвідувальний аналіз даних

Для моделювання індексу споживчих цін сформовано базу даних [10; 32; 34]. У додатку Б наведено інтерпретація, тобто пояснення змінних, які було зібрано для аналізу. Код R, який був використаний у даній роботі, міститься у додатку В.

Частина змінних спрогнозована за допомогою ковзного середнього. Через воєнний стан в Україні деякі показники опубліковані лише до 2021 року. Відповідно, для проведення розрахунків прийнято рішення їх спрогнозувати, хоча б до 2022 року, щоб оцінити вплив змінних на індекс споживчих цін.

Для коректної побудови множинної лінійної регресії спочатку проведено кореляційний аналіз змінних, на основі яких може бути побудована модель. Такий аналіз проводиться для того, щоб виявити зв'язок між незалежними змінними і, у подальшому, не обрати в модель корельовані між собою змінні, з метою виключення проблем з мультиколінеарністю моделі.

Розглянуто кореляційну матрицю незалежних змінних (додаток Д). У матриці представлено, що між багатьма змінними присутня висока кореляція. Наприклад, між витратами населення та грошовою базою кореляція на рівні 0,94, між витратами населення та готівковими коштами випущеними в обіг кореляція також, 0,94, між витратами населення та номінальним ВВП кореляція - 0,95. Також, висока кореляція між грошовою базою та номінальним ВВП, між готівковими коштами випущеними в обіг і номінальним ВВП, між видатками Державного бюджету та грошовою базою. Ще висока кореляція між експортом та імпортом тощо.

Підсумовуючи варто зазначити, що у даному наборі даних багато змінних, які мають високу кореляцію тому, при побудові модель приділено особливу увагу при виборі змінних, щоб потім не стикнутись з проблемою, коли фактори у моделі корелюватимуть один з одним.

Проаналізовано динаміку залежної змінної для подальшої побудови лінійної регресії – індекс споживчих цін на рис. 3.1:



Рисунок 3.1 – Динаміка ІСЦ в Україні за 2013-2022 роки

Джерело: розроблено автором на основі [10]

На рис. 3.1 спостерігається, що з 2014 року ІСЦ почав різко зростати і досяг свого максимуму за весь досліджуваний період у 1 кварталі 2015 року. Скоріш за все таке швидке зростання та пікове значення на рівні близьким до 120%, через важкий перебіг військового конфлікту на сході України. Після такого великого пікового значення, індекс інфляції падав впродовж інших кварталів 2015 року, і вже з 2016 значення ІСЦ частково стабілізувалось. Але у 1-2 кварталах 2022 року розпочався воєнний стан в Україні й індекс споживчих цін знову дестабілізувався, хоча впродовж 2022 року прослідковується тенденція на спадання індексу інфляції.

Ці військові конфлікти потягли за собою й інші зміни та дестабілізаційні процеси в економіці України, а саме різке зростання курсу гривні до долара та євро, зниження рівня іноземних інвестицій в Україну, що могло вплинути на кількість сплачених податків резидентами і нерезидентами, а відповідно це могло призвести до зменшення наповненості державного бюджету, а в подальшому до скорочення видатків і падіння ВВП тощо.

Далі варто розглянути візуалізації змінних, які були вибрані для побудови множинної лінійної регресії (рис. 3.2):

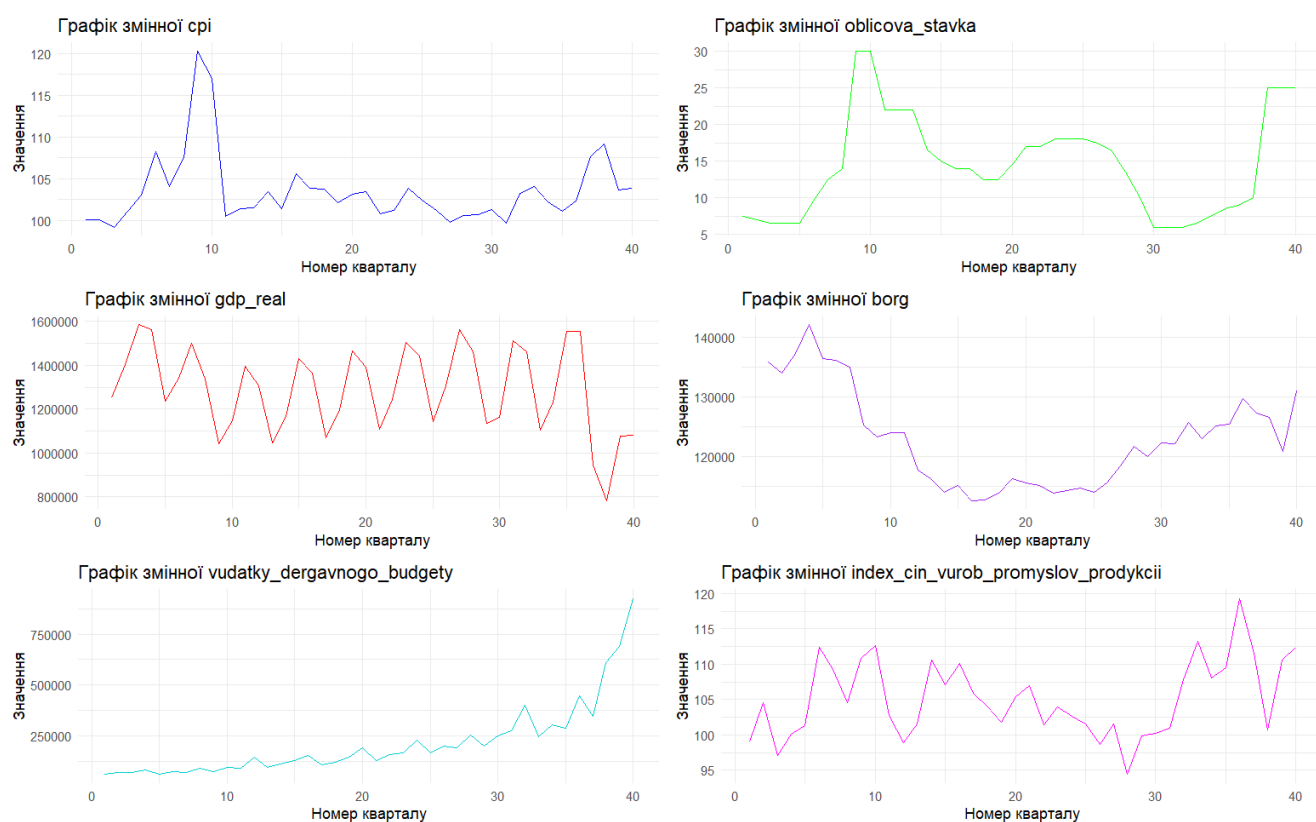


Рисунок 3.2 – Впливові фактори на ІСЦ

Джерело: розроблено автором на основі [10; 32; 34]

Опис змін індексу споживчих цін наведений вище (рис. 3.1).

На графіку динаміки облікової ставки НБУ (рис. 3.2), спостерігається пікове значення у 30%, таке значення ставка набуває у той самий період, що і індекс інфляції, а саме у 2015 році, мабуть, це через потрясіння економіки внаслідок військових дій, далі з плином часу ставка спадає і коливається в середньому навколо

15%, але з початком військового стану в Україні, облікова ставка знову почала зростати до 25%, але не досягла рівня 2015 року.

Реальний ВВП також відображений на загальному графіку (рис. 3.2), можна стверджувати, що в цілому цей показник є сезонних, тобто приблизно на першому кварталі кожного року він має низьке значення, впродовж року збільшується і до першого кварталу наступного року знову зменшується і т.д.

Валовий зовнішній борг візуалізовано на рис. 3.2, на ньому представлено, що під час великих потрясінь для економіки, які вже були описані вище, цей показник зростає, а у більш спокійні роки, або у час коли економіка «адаптувалась» до змін, значення боргу зменшувались.

Графік видатків Державного бюджету (рис. 3.2), на ньому спостерігається постійний ріст видатків впродовж років. У 2022 році цей показник досяг свого пікового значення тому, що відбулось зростання фінансування сектору оборони та безпеки [35].

Візуалізація індексу цін виробників промислової продукції (рис. 3.2), демонструє, що до 30 кварталу, у вибірці це 2020 рік, цей індекс має тенденцію до зниження, але після цього періоду, він починає доволі різко зростати, це може бути спричинене пандемією COVID-19, коли багато підприємств та промислових об'єктів припиняли свою роботу через поширення захворювання.

3.2 Моделювання ІСЦ за допомогою регресійної, ARIMA та ETS моделей

Наступним кроком є застосування до ІСЦ регресійного аналізу. З усіх наведених змінних шляхом перебору їх різноманітних можливих комбінацій, знайдено найбільш оптимальну модель множинної лінійної регресії, яка складається з наступних змінних: `cpi`, `oblicova_stavka`, `gdp_real`, `borg`, `vudatky_derzhavnogo_budgety`, `index_cin_vurob_promyslov_prodykcii`.

Для цієї моделі були розраховані наступні коефіцієнти (рис. 3.3):

```
Call:
lm(formula = cpi ~ oblicova_stavka + gdp_real + borg + vudatky_derzhavnogo_budgety +
    index_cin_vurob_promyslov_prodykcii, data = data)

Coefficients:
            (Intercept)                oblicova_stavka                gdp_real
                5.434e+01                   3.253e-01                -7.024e-06
                   borg            vudatky_derzhavnogo_budgety index_cin_vurob_promyslov_prodykcii
                1.452e-04                   -8.151e-06                   3.562e-01
```

Рисунок 3.3 – Коефіцієнти моделі лінійної регресії

Джерело: розроблено автором на основі [10; 32; 34]

Отже, рівняння моделі матиме такий вигляд:

$$y_i = 54.34 + 0.3253 \cdot x_1 - 0.000007024 \cdot x_2 + 0.0001452 \cdot x_3 - 0.000008151 \cdot x_4 + 0.3562 \cdot x_5 + u_i, \quad (3.1)$$

де y_i – Індекс споживчих цін, %,

x_1 – Облікова ставка НБУ, на кінець кварталу, %,

x_2 – Реальний ВВП, млн грн.

x_3 – Валовий зовнішній борг, млн дол. США,

x_4 – Видатки Державного бюджету, млн грн,

x_5 – Індеси цін виробників промислової продукції (до попереднього кварталу).

Опис інтерпретації коефіцієнтів моделі з формули 3.1:

b_0 – індекс споживчих цін за цією моделлю становитиме 54,34%, за умови, що всі інші незалежні змінні дорівнюватимуть 0;

b_1 – якщо значення облікової ставки НБУ збільшиться на 1%, то індекс споживчих цін збільшиться на 0,3253%;

b_2 – якщо значення реального ВВП збільшиться на 1 млн грн, то індекс споживчих цін зменшиться на 0.000007024 %;

b_3 – якщо значення валового зовнішнього боргу України збільшиться на 1 млн дол. США, то значення індексу споживчих цін збільшиться на 0,0001452%;

b_4 – якщо значення видатків Державного бюджету збільшиться на 1 млн грн, то значення індексу споживчих цін зменшиться на 0,000008151%;

b_5 – якщо індекс цін виробників промислової продукції збільшиться на 1%, то індекс споживчих цін в Україні збільшиться на 0,3562%.

Отже, на підставі цієї моделі можна зробити висновок, що при збільшенні облікової ставки НБУ, валового зовнішнього боргу України та індексу цін виробників промислової продукції зростатиме індекс споживчих цін в Україні. А за умови збільшення значення реального ВВП та видатків Державного бюджету ІСЦ зменшуватиметься.

Далі розглянуто статистичну значущість змінних у моделі (рис. 3.4):

```
Call:
lm(formula = cpi ~ oblicova_stavka + gdp_real + borg + vudatky_dergavnogo_budgety +
    index_cin_vurob_promyslov_prodykcii, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-5.0884 -1.7087 -0.1409  1.8235  6.7450

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      5.434e+01  1.140e+01   4.766 3.45e-05 ***
oblicova_stavka   3.253e-01  7.492e-02   4.342 0.000120 ***
gdp_real         -7.024e-06  2.477e-06  -2.836 0.007639 **
borg              1.452e-04  5.750e-05   2.526 0.016370 *
vudatky_dergavnogo_budgety -8.151e-06  2.564e-06  -3.179 0.003144 **
index_cin_vurob_promyslov_prodykcii  3.562e-01  8.284e-02   4.300 0.000136 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.666 on 34 degrees of freedom
Multiple R-squared:  0.6585,    Adjusted R-squared:  0.6083
F-statistic: 13.11 on 5 and 34 DF,  p-value: 3.818e-07
```

Рисунок 3.4 – Результати моделі лінійної регресії

Джерело: розроблено автором на основі [10; 32; 34]

Спостерігається на рис. 3.4, що всі змінні моделі є статистично значущими, тобто значення p-value менше за 0,05, до того ж в R це відображається зірочками (*) після значень p-value, біля усіх змінних присутні зірочки, тому всі ці фактори моделі є статистично значущими, іншими словами змінні є впливовими на індекс споживчих цін.

R-квадрат (R-squared) для цієї моделі становить 65,85%, це означає, що на 65.85% рівняння регресії пояснює поведінку індексу споживчих цін. Це досить високий R-квадрат, бо його значення більше 50%, це говорить про те, що модель добре пояснює дані.

Також, для покращення точності моделі варто спробувати ввести фіктивні змінні кварталів, у такому випадку модель матиме вигляд, як зображено на рис. 3.5.

Є можливість, що введення таких змінних допоможе врахувати сезонні коливання, які матимуть вплив на індекс інфляції. Якщо такі фіктивні незалежні змінні будуть значимими, то це означатиме, що є якийсь певний специфічний вплив квартальних факторів. Якщо такий вплив буде наявним, тобто p -value меншим за 0,05, то є ймовірність, що покращиться прогностична здатність моделі.

```
Call:
lm(formula = cpi ~ oblicova_stavka + gdp_real + borg + vudatky_dergavnogo_budgety +
    index_cin_vurob_promyslov_prodykcii + q2 + q3 + q4, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-4.6033 -1.6934  0.2788  1.2829  7.2514

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      5.798e+01  1.257e+01   4.612 6.51e-05 ***
oblicova_stavka   2.950e-01  8.963e-02   3.291  0.00250 **
gdp_real         -1.055e-05  5.836e-06  -1.808  0.08029 .
borg              1.471e-04  5.769e-05   2.550  0.01592 *
vudatky_dergavnogo_budgety
-1.059e-05  3.682e-06  -2.876  0.00722 **
index_cin_vurob_promyslov_prodykcii
 3.596e-01  8.329e-02   4.318  0.00015 ***
q2              1.229e+00  1.386e+00   0.887  0.38192
q3              1.099e+00  2.525e+00   0.435  0.66643
q4              2.774e+00  2.401e+00   1.155  0.25683
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.666 on 31 degrees of freedom
Multiple R-squared:  0.6885,    Adjusted R-squared:  0.6082
F-statistic: 8.566 on 8 and 31 DF,  p-value: 4.439e-06
```

Рисунок 3.5 – Результати моделі лінійної регресії з фіктивними
квартальними змінними

Джерело: розроблено автором на основі [10; 32; 34]

На рис. 3.5 наведена модель з фіктивними змінними кварталів. Вона є гіршою як модель без цих факторів, бо введені змінні є статистично незначущими, тобто p -value > 0,05. Хоча R-squared становить 68,85%, що є більшим, ніж у моделі без фіктивних змінних, це значення вказує на стільки відсотків рівняння моделі з незалежними змінними пояснюють поведінку ІСЦ.

Можна спробувати замінити квартальні фіктивні змінні на змінну часу, коли в країні розпочалась пандемія COVID-19, тобто яка містить значення 1 та 0, всі значення до 1 кварталу 2020 року становлять 0, а з 1 кварталу 2020 року, всі наступні квартали – 1. Модель отримала такий вигляд (рис. 3.6):

```

Call:
lm(formula = cpi ~ oblicova_stavka + gdp_real + borg + vudatky_derzavnogo_budgety +
    index_cin_vurob_promyslov_prodykcii + covid, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-5.7233 -1.5645 -0.3536  1.6623  6.0642

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      5.407e+01  1.134e+01   4.770 3.62e-05 ***
oblicova_stavka   4.029e-01  9.902e-02   4.068 0.000277 ***
gdp_real        -5.811e-06  2.665e-06  -2.180 0.036496 *
borg             1.493e-04  5.725e-05   2.607 0.013598 *
vudatky_derzavnogo_budgety
-1.228e-05  4.307e-06  -2.850 0.007469 **
index_cin_vurob_promyslov_prodykcii
 3.306e-01  8.513e-02   3.883 0.000468 ***
covid            2.201e+00  1.853e+00   1.188 0.243213
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.65 on 33 degrees of freedom
Multiple R-squared:  0.6725,    Adjusted R-squared:  0.613
F-statistic: 11.3 on 6 and 33 DF,  p-value: 7.746e-07

```

Рисунок 3.6 – Результати моделі лінійної регресії з фіктивною змінною часу Covid-19

Джерело: розроблено автором на основі [10; 32; 34]

З рис. 3.6 очевидно, що R-squared збільшився у порівнянні зі звичайною лінійною регресією, але фіктивна змінна, що відповідає за COVID-19 є незначущою, тобто $p\text{-value} > 0,05$.

Включення змінної воєнного стану до моделі дозволяє оцінити його вплив на ІСЦ. Тому було прийнято рішення, спробувати ввести фіктивну змінну воєнного стану в Україну, у якій значення з 1 кварталу 2022 року становитимуть 1, а до цього періоду 0. Наявна така модель (рис. 3.7):

```

Call:
lm(formula = cpi ~ oblicova_stavka + gdp_real + borg + vudatky_derzavnogo_budgety +
    index_cin_vurob_promyslov_prodykcii + war, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-5.0568 -1.7435 -0.2762  1.9493  6.5659

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      5.403e+01  1.158e+01   4.665 4.93e-05 ***
oblicova_stavka   3.323e-01  7.828e-02   4.245 0.000167 ***
gdp_real        -7.620e-06  3.000e-06  -2.540 0.015985 *
borg             1.551e-04  6.426e-05   2.413 0.021551 *
vudatky_derzavnogo_budgety
-6.914e-06  4.294e-06  -1.610 0.116890
index_cin_vurob_promyslov_prodykcii
 3.526e-01  8.452e-02   4.172 0.000206 ***
war            -1.116e+00  3.083e+00  -0.362 0.719754
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.7 on 33 degrees of freedom
Multiple R-squared:  0.6599,    Adjusted R-squared:  0.598
F-statistic: 10.67 on 6 and 33 DF,  p-value: 1.399e-06

```

Рисунок 3.7 – Результати моделі лінійної регресії з фіктивною змінною часу воєнного стану в Україні

Джерело: розроблено автором на основі [10; 32; 34]

Хоча значення R-squared моделі рис. 3.7 є вищим (65,99%) в порівнянні з моделлю на рис. 3.4, введена фіктивна змінна war, тобто періоду воєнного стану є незначущою у моделі – $p\text{-value} > 0,05$. Ще, варто спробувати застосувати дві фіктивні змінні до моделі, на період COVID-19 та воєнного стану. З включенням таких змінних отримано наступну модель (рис. 3.8):

```
Call:
lm(formula = cpi ~ oblicova_stavka + gdp_real + borg + vudatky_derzhavnogo_budgety +
  index_cin_vurob_promyslov_prodykcii + covid + war, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-5.704 -1.596 -0.383  1.659  6.032

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      5.400e+01  1.154e+01   4.680 5.02e-05 ***
oblicova_stavka   4.032e-01  1.006e-01   4.007 0.000343 ***
gdp_real        -5.980e-06  3.331e-06  -1.796 0.082023 .
borg             1.516e-04  6.410e-05   2.366 0.024228 *
vudatky_derzhavnogo_budgety
-1.190e-05  6.183e-06  -1.924 0.063304 .
index_cin_vurob_promyslov_prodykcii
 3.302e-01  8.658e-02   3.813 0.000590 ***
covid           2.161e+00  1.937e+00   1.116 0.272780
war            -2.762e-01  3.162e+00  -0.087 0.930947
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.69 on 32 degrees of freedom
Multiple R-squared:  0.6726,    Adjusted R-squared:  0.601
F-statistic: 9.392 on 7 and 32 DF,  p-value: 2.769e-06
```

Рисунок 3.8 – Результати моделі лінійної регресії з двома фіктивними змінними часу Covid-19 та воєнного стану в Україні

Джерело: розроблено автором на основі [10; 32; 34]

На рис. 3.8 зображено, що фіктивні змінні є незначущими, тобто $p\text{-value} > 0,05$, хоча і R-squared є вищим, ніж у моделі без фіктивних змінних.

Отже, найкращою для подальшого аналізу та прогнозування є найперша побудована лінійна регресійна модель (див. рис. 3.4), де не включені будь-які фіктивні змінні, її і використано для подальшого аналізу.

Наступним показником для аналізу є коефіцієнт еластичності для незалежних змінних моделі, тобто x_i , значення наведені у табл. 3.1.

Таблиця 3.1 – Коефіцієнти еластичності для лінійної регресійної моделі

E(X1)	0,0452
E(X2)	-0,0876
E(X3)	0,1723
E(X4)	-0,0168
E(X5)	0,3618

Джерело: розроблено автором на основі [10; 32; 34]

На основі коефіцієнта еластичності можна зробити висновок, що найбільш впливовим фактором на індекс споживчих цін є індекс цін виробників промислової продукції, наступним за впливовістю є валовий зовнішній борг України. Менш впливовими факторами є значення реального ВВП та облікова ставка. А найменш впливовим фактором є видатки Державного бюджету.

Наступним кроком є перевірка змінних на автокореляцію за допомогою критерію Дарбіна-Уотсона (рис. 3.9):

```
lag Autocorrelation D-W Statistic p-value
1      0.3073245      1.335819      0.004
Alternative hypothesis: rho != 0
```

Рисунок 3.9 – Результати критерію Дарбіна-Уотсона

Джерело: розроблено автором на основі [10; 32; 34]

З рис. 3.9 відомо, що значення $p\text{-value}=0,004<0,05$, отже варто відхилити нульову гіпотезу про те, що фактори є незалежними один від одного, і стверджувати, що залишки лінійної регресії є автокорельованими.

Далі необхідно перевірити модель на наявність мультиколінеарності за допомогою функцій в R: `omcdiag` і `imcdiag` (див. рис. 3.10).

```
Call:
omcdiag(mod = line_reg)

Overall Multicollinearity Diagnostics
Determinant |X'X|:      0.5473      0
Farrar Chi-Square: 22.0007      1
Red Indicator:  0.2525      0
Sum of Lambda Inverse: 6.2794      0
Theil's Method: -1.6445      0
Condition Number: 75.8341      1

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test

Call:
imcdiag(mod = line_reg)

All Individual Multicollinearity Diagnostics Result
              VIF    TOL    Wi    F1 Leamer    CVIF    Klein    IND1    IND2
oblicova_stavka 1.4366 0.6961 3.8201 5.2390 0.8343 1.8859 0 0.0796 1.5355
gdp_real        1.3092 0.7639 2.7051 3.7099 0.8740 1.7187 0 0.0873 1.1932
borg            1.2101 0.8264 1.8380 2.5206 0.9091 1.5886 0 0.0944 0.8771
vudatky_derzavnogo_budgety 1.2068 0.8286 1.8099 2.4821 0.9103 1.5843 0 0.0947 0.8660
index_cin_vurob_promyslov_prodykcii 1.1168 0.8954 1.0217 1.4012 0.9463 1.4661 0 0.1023 0.5283

1 --> COLLINEARITY is detected by the test
0 --> COLLINEARITY is not detected by the test
* all coefficients have significant t-ratios

R-square of y on all x: 0.6585
* use method argument to check which regressors may be the reason of collinearity
```

Рисунок 3.10 – Результати перевірки моделі на мультиколінеарність

Джерело: розроблено автором на основі [10; 32; 34]

Мультиколінеарність виникає тоді, коли дві або більше змінних у моделі корелюють між собою. Детермінант матриці $X'X$, має значення досить віддалене від 0 тому, цей тест показує, що у даних немає мультиколінеарності. Тест Фаррара Хі-квадрата дорівнює 22, а це вказує на наявність мультиколінеарності. Наступні

тести дають розрізнені значення щодо наявності або відсутності мультиколінеарності, якщо дивитись в цілому, то 4 тести вказують на те, що мультиколінеарність відсутня, лише 2 показують її наявність. Тому, можна зробити висновок, що мультиколінеарності у даних немає.

Якщо розглядати другу частину рис. 3.10, то значення VIF (Variance Inflation Factor) для всіх змінних моделі є низькими, вони не перевищують 5, а є близькими до 1, тому у висновку варто зазначити про відсутність мультиколінеарності між змінними моделі.

Наступним кроком є перевірка моделі на гетероскедастичність з допомогою тесту Брейша-Пагана (рис. 3.11):

```
studentized Breusch-Pagan test

data: line_reg
BP = 15.878, df = 5, p-value = 0.007201
```

Рисунок 3.11 – Результати тесту Брейша-Пагана

Джерело: розроблено автором на основі [10; 32; 34]

Даний тест показує чи присутня у залишках моделі не однакова для всіх спостережень дисперсія. Статистика тесту дорівнює 15.878, а відповідне значення p дорівнює 0.007201. Оскільки p -значення менше 0,05, варто відкинути нульову гіпотезу. Тому, можна сказати, що гетероскедастичність присутня в моделі регресії.

Далі варто перейти до побудови моделі ARIMA. Даний тим моделі є ефективним та потужним інструментом, яка має високі прогностичні властивості часових рядів. Тому, наступним кроком побудовано ARIMA модель для ІСЦ, бо даний вид моделі допомагає врахувати особливості часового ряду індексу інфляції, а саме тренд та сезонність, шляхом включення інтегрування, авторегресії та ковзного середнього. Такий вид моделі є досить гнучкою та підлаштовується під різні дані, використовуючи різні значення гіперпараметрів, які описані в пп.1.3 розділу 1. Перед тим, як будувати таку модель, потрібно переконатися, що дані є стаціонарними тому, що ARIMA модель застосовується до стаціонарних даних. Для цього варто проаналізувати динаміку індексу споживчих цін, рис. 3.12:

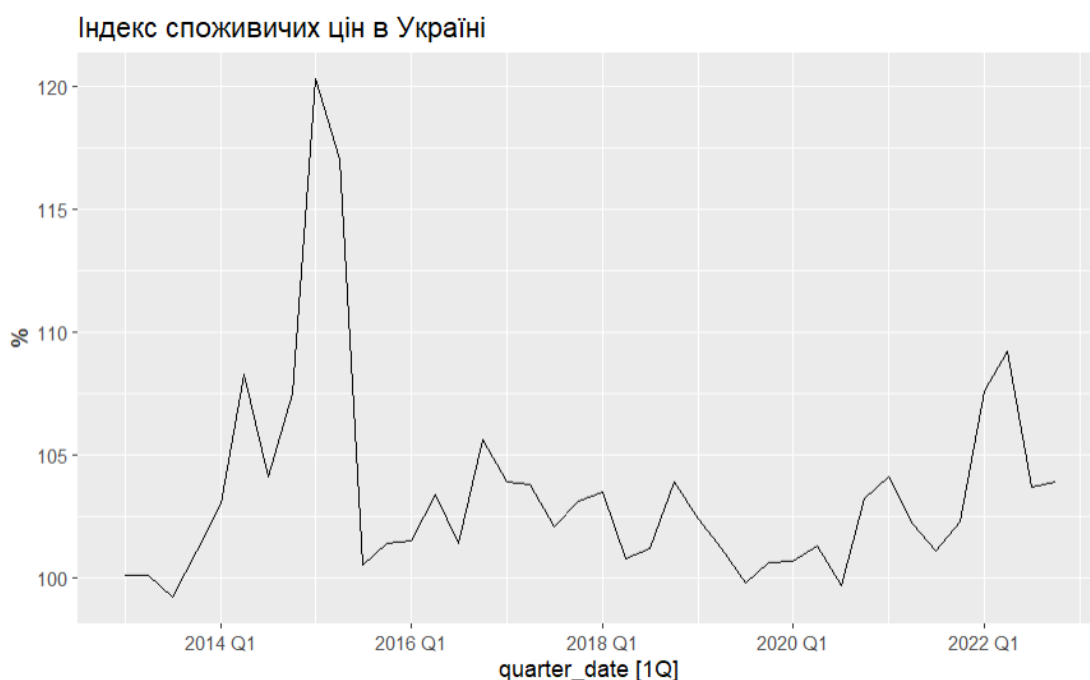


Рисунок 3.12 – Візуалізація ІСЦ в Україні за 2013-2022 роки

Джерело: розроблено автором на основі [10]

З даного графіку (рис. 3.12) важко зробити висновки, щодо того чи часовий ряд є стаціонарним, але варто зазначити, що сезонність не надто прослідковується, тренду на перши погляд також немає, ряд розвивається приблизно горизонтально.

Наступним кроком є розрахунок коефіцієнт Херста (рис. 3.13):

coef_hurst
<dbl>
0.8691932

Рисунок 3.13 – Результати розрахунку коефіцієнта Херста

Джерело: розроблено автором на основі [10]

Показник Херста вказує на наявність довготривалої пам'яті у часового ряду. Цей показник може коливатись від 0 до 1. Значення показника Херста для часового ряду, є досить близькими до 1, це говорить про те, що часовий ряд поводитьсь доволі стабільно, має певну довготривалу пам'ять та схильний до наслідування тренду. Тобто, можна припустити, що значення часового ряду при прогнозуванні не будуть суттєво змінювати свій напрямок, тобто як рухався тренд, то скоріш за все так і рухатиметься далі.

Наступною варто розглянути корелограму автокореляції для усіх спостережень на рисунку 3.14, даний аналіз є одним зі способів для визначення стаціонарності часового ряду.

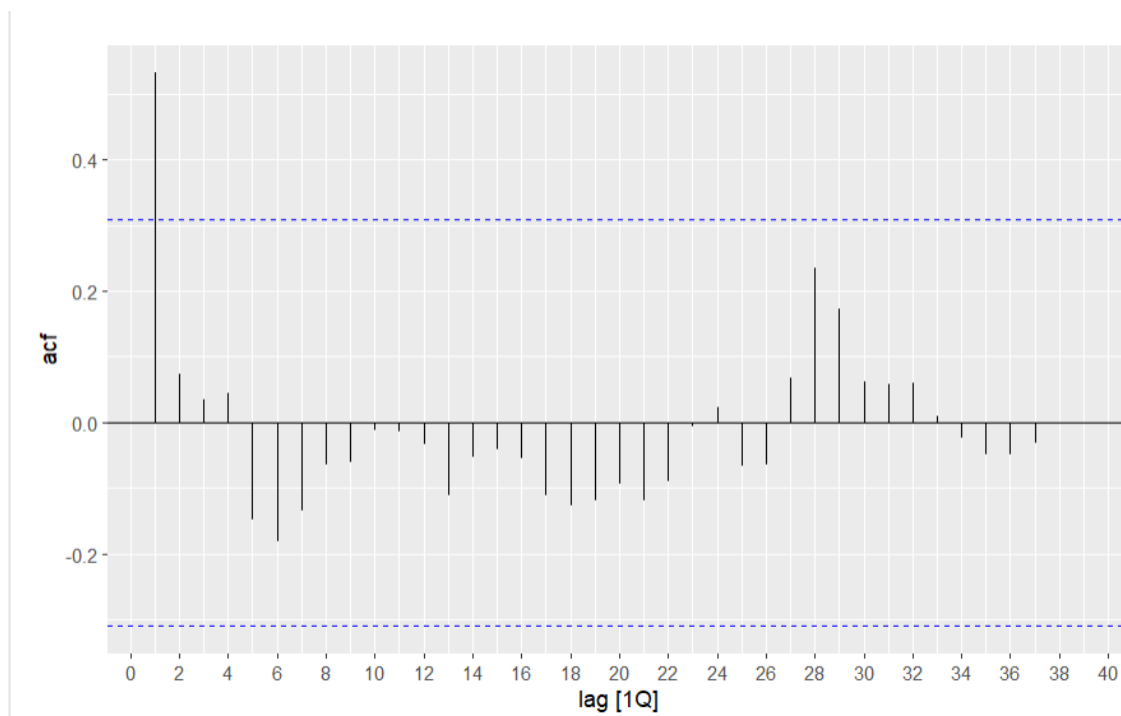


Рисунок 3.14 – Результати розрахунку коефіцієнта Херста

Джерело: розроблено автором на основі [10]

На даній корелограмі видно, що є тільки один (перший) статистично значущий лаг, а далі всі значення спадають дуже різко. Це наводить на думку, що можливо даний часовий ряд є стаціонарним.

Далі використано KPSS тест, він зображений на рис. 3.15.

kpss_stat <dbl>	kpss_pvalue <dbl>
0.1247917	0.1

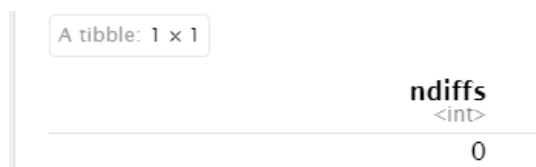
Рисунок 3.15 – Результати розрахунку KPSS тесту

Джерело: розроблено автором на основі [10]

KPSS тест або його інша назва – тест Квятковського-Філіпса-Шмідта-Шина, нульова гіпотеза цього тесту полягає в тому, що дані є стаціонарні та несезонні.

За даним тестом очевидно, що $p\text{-value} > 0.05$, це дає підстави прийняти нульову гіпотезу про те, що даний часовий ряд є стаціонарним і несезонним. Тобто, до нього не варто застосовувати перетворення.

Далі розглянуто функцію `unitroot_ndiffs` (рис. 3.16), яка покаже, які ряди різниць потрібно застосовувати, для трансформації часового ряду на стаціонарний:



ndiffs <int>
0

Рисунок 3.16 – Результати функції `unitroot_ndiffs` (необхідна к-сть рядів різниць)

Джерело: розроблено автором на основі [10]

Таким чином, за даними рис. 3.16 до часового ряду не варто використовувати ніякі перетворення, ні ряд перших різниць, ні ряд сезонних різниць тощо.

Отже, на основі всіх вищенаведених проведених тестів, можна зробити висновок, що часовий ряд ІСЦ є стаціонарним. Результати тестів вказують на те, що його не потрібно трансформувати, тобто модель ARIMA варто будувати на даному часовому ряді без перетворень.

Далі побудовано модель ARIMA, але для початку варто розглянути наступний рис. 3.17, на якому зображено часовий ряд індексу споживчих цін, корелограма автокореляція та часткової автокореляції, для того, щоб визначити можливі значення гіперпараметрів моделі - p , d , q .

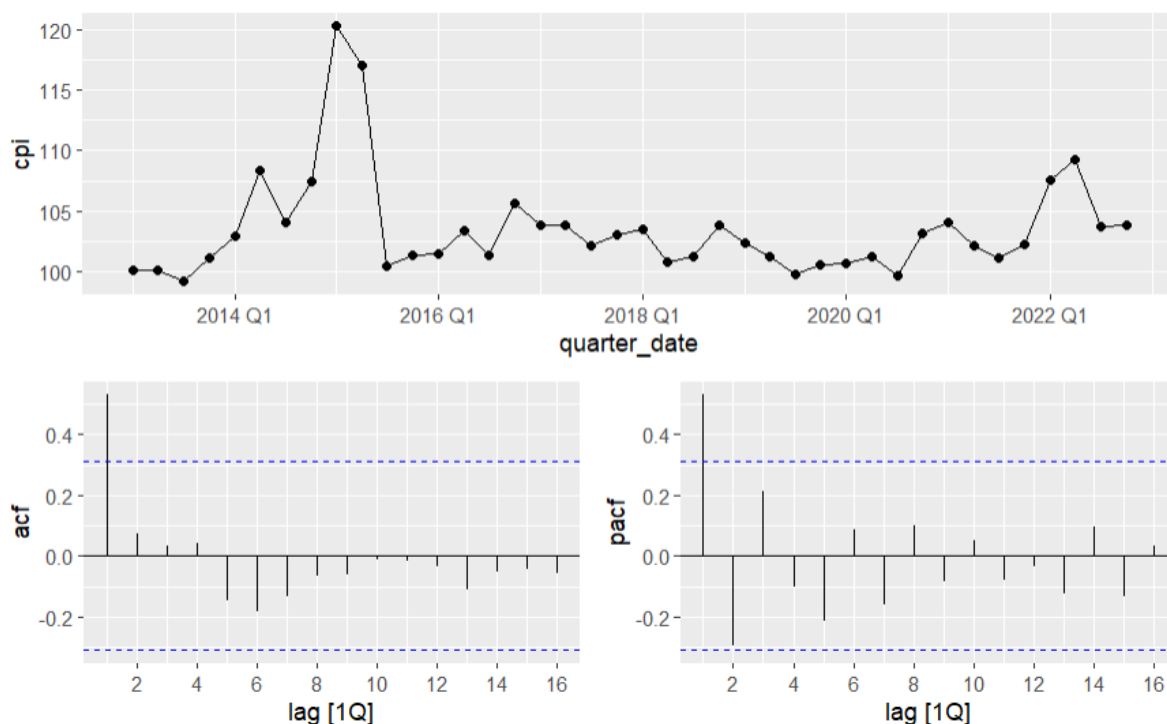


Рисунок 3.17 – Аналіз часового ряду ІСЦ (візуалізація, корелограми автокореляції та часткової автокореляції)

Джерело: розроблено автором на основі [10]

На даній корелограмі автокореляції спостерігається, що є тільки один (перший) статистично значущий лаг, а далі всі значення спадають дуже різко. Це підтверджує гіпотезу про стаціонарність даного часового ряду.

Часткова автокореляція (PACF) - це та ж сама автокореляція, співвідношення у поточного, до його попередніх значень, коли вплив проміжків не враховується, наприклад якщо взяти y_t – це поточне значення, і y_{t-10} , то от цей проміжок між 1 і 10 значення у частковій автокореляції не враховується. За допомогою цього графіку (рис. 3.17) можна підібрати гіперпараметри для АРІМА. В залежності від кількості значущих часткових автокореляцій встановлюється значення AR, тобто параметр p на рівні кількості значущих часткових автокореляцій. Якщо на графіку зображено велика кількість значущих значень PACF, це вказує на те, що присутня значна залежність поточних значень від попередні спостережень часового ряду, у такому випадку варто брати більше значення параметра p .

Для підбору МА, тобто ковзного середнього варто дивитись на графік автокореляції (ACF), тобто, якщо на такому графіку є тільки один статистично значущий лаг, то взято значення МА, тобто параметра q на рівні 1, якщо два статистично значущих лаги то $q=2$, і так далі.

За допомогою рис. 3.17 можна спробувати підібрати самостійно «вручну» гіперпараметри для ARIMA. Для аналізованого у роботі часового ряду зображено, що часткова автокореляція має лише одне значуще значення, тобто можна припустити, що AR, параметр моделі $p=1$. На основі графіка автокореляції можна підібрати МА – параметр q , спостерігається, що на графіку лише один значущий лаг, тому можна припустити, що $q=1$, а частина моделі I, якому відповідає параметр d , скоріш за все $d=0$, бо за попередніми тестами, визначено, що ніяких трансформацій часовий ряд не потребує.

Далі потрібно підтвердити або спростувати попередні твердження щодо найоптимальніших параметрів для ARIMA моделі, побудувавши авто ARIMA модель (рис. 3.18):

```
Series: cpi
Model: ARIMA(1,0,2) w/ mean

Coefficients:
          ar1      ma1      ma2  constant
      -0.9685   1.7507   0.7941   203.7604
s.e.      0.0629   0.1269   0.1121    1.7380

sigma^2 estimated as 10.91:  log likelihood=-103.35
AIC=216.69   AICc=218.46   BIC=225.14
```

Рисунок 3.18 – Результати ARIMA моделі для часового ряду ІСЦ за 2013-2022 роки

Джерело: розроблено автором на основі [10]

На рис. 3.18 зазначено, що оптимальними критеріями для моделі є $p=1$, $d=0$, $q=2$, тобто при підборі параметрів «вручну» здійснено помилку щодо частини моделі МА, тобто параметра q , з корелограми автокореляції здавалось, що він має дорівнювати 1, а не 2, всі інші параметри були визначені правильно.

Отже, модель ARIMA має рівняння наступного вигляду (формула 3.2):

$$y_t = 203.76 - 0.96 \cdot y_{t-1} + 1.75 \cdot \varepsilon_{t-1} + 0.79 \cdot \varepsilon_{t-2} + \varepsilon_t, \quad (3.2)$$

де y_t – поточне значення часового ряду,

y_{t-1} – попереднє значення часового ряду,

ε_{t-1} – попереднє значення білого шуму,

ε_{t-2} – два значення назад від поточного значення білого шуму,

ε_t – білий шум.

Далі варто розглянути інформаційні критерії для даної моделі на рис. 3.19:

.model <chr>	sigma2 <dbl>	log_lik <dbl>	AIC <dbl>	AICc <dbl>	BIC <dbl>
auto	10.91274	-103.3459	216.6919	218.4566	225.1363

1 row

Рисунок 3.19 – Інформаційні критерії для ARIMA моделі індексу інфляції

Джерело: розроблено автором на основі [10]

Ці всі показники (див. рис. 3.19) застосовуються для оцінки якості моделі і порівняння різних моделей між собою, далі у ході роботи будуть порівнюватись роботи за інформаційним критерієм AIC. В цілому, зараз можна сказати, що sigma2 – оцінка дисперсії помилок даної моделі, і чим меншим є дане значення, тим краще. log_lik – логарифмічна правдоподібність моделі, і чим більше значення даного критерію тим краще. Для індексу споживчих цін воно становить -103,3, це погано, оскільки вказує на те що модель не дуже добре відповідає даним. AIC, AICc (скоригований критерій AIC) та BIC (Байєсівський інформаційний критерій) використовуються для порівняння моделей між собою, чим нижчими є значення цих критеріїв у моделі, тим кращою вважається дана модель.

Далі здійснено аналіз візуалізації залишків ARIMA моделі (рис. 3.20):

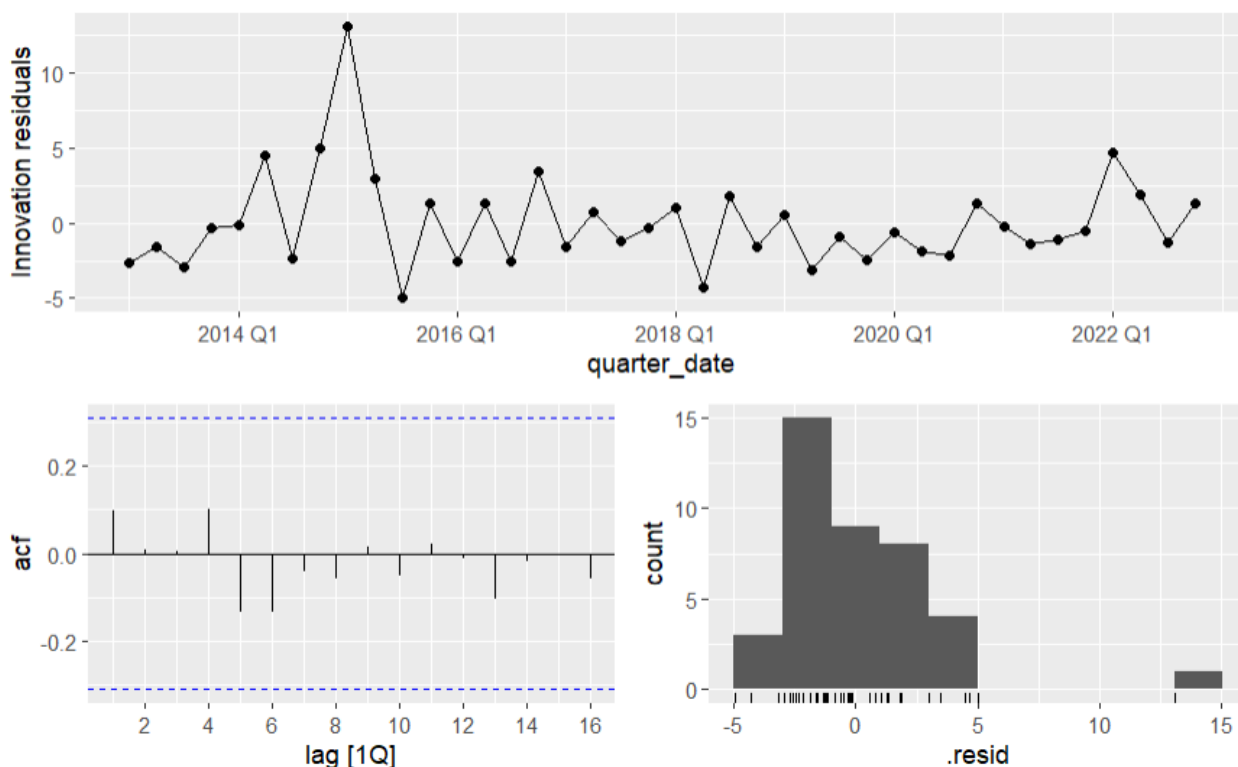


Рисунок 3.20 – Аналіз залишків ARIMA моделі

Джерело: розроблено автором на основі [10]

За рис. 3.20 можна оцінити автокорельованість залишків. Спостерігається, що всі значення автокореляції по залишках є в межах довірчих інтервалів, тобто коефіцієнти автокореляції є статистично не значущими. Варто сказати, що модель добре може прогнозувати майбутні значення, бо витягла всю необхідну інформацію з початкових даних і ніякої важливої інформації не залишилось у залишках. На гістограмі (рис. 3.20) прослідковується, що скоріше за все нормального розподілу залишків немає. Також, на рис. 3.20 спостерігається, що залишки схожі на білий шум.

Наступним розглянуто тест Лjung-Бокса, який перевіряє залишки моделі на автокорельованість (рис. 3.21):

.model <chr>	lb_stat <dbl>	lb_pvalue <dbl>
auto	0.8977679	0.3433809

Рисунок 3.21 – Результати тесту Лjung-Бокса

Джерело: розроблено автором на основі [10]

За тестом Лjunga-Бокса отримано, що $p\text{-value}=0,34$, це є значно більшим за 0,05, відповідно немає підстав відхилити нульову гіпотезу, яка говорить про те, що залишки моделі є білим шумом. Тобто, цей тест показує, що залишки моделі є білим шумом, отже між залишками немає автокореляції.

У даній роботі, не будуть перебиратись різні специфікації ETS моделей, а використовуватиметься автоматична модель, яка сама обирає оптимальну модель, шляхом побудови всіх можливих специфікацій моделі. Оптимальна модель даним способом обирається на основі значення скорегованого інформаційного критерію AIC, тобто AICc (дані критерії, це по факту просто числа, вони не мають певної розмірності виміру, наприклад можна сказати, що значення цього показника може вимірюватись у відсотках або атомах). Тобто, іншими словами найоптимальнішою вважається та модель, яка має найнижче значення критерію AICc.

Далі побудовано ETS модель. Отримано результат, який відображено на рисунку 3.22:

```
Series: cpi
Model: ETS(M,N,A)
Smoothing parameters:
  alpha = 0.8862782
  gamma = 0.0001000163

Initial states:
  l[0]      s[0]      s[-1]      s[-2]      s[-3]
99.76393 -0.6145282 -2.040098 1.271512 1.383114

sigma^2: 0.0012

      AIC      AICc      BIC
255.7172 259.2172 267.5393
```

Рисунок 3.21 – Автоматична ETS модель для ІСЦ від 2013 до 2022 рр.

Джерело: розроблено автором на основі [10]

Для часового ряду ІСЦ отримано, що $\alpha = 0.886$, $\gamma = 0.0001$. Тобто в цілому можна сказати, що значення альфа вказує на те, що значення рівня часового ряду змінюються певним чином між середнім значення загального часового ряду та постійними змінами, тобто оновленнями нових даних. А γ

досить близька до 0, а це вказує на те, що сезонність має якесь певне фіксоване значення, наприклад середнє значення.

Отже, отримано такі компоненти оптимальної моделі ETS(M,N,A):

- Мультиплікативні помилки (Error) – значення M у дужках моделі;
- Відсутній тренд (Trend) – друге значення N у моделі;
- Адитивна сезонність (Seasonality) – третє значення, тобто A у моделі.

У отриманій моделі наявні α та γ гіперпараметри згладжування, а також початкові значення ℓ_0 та $s_0, s_{-1}, s_{-2}, s_{-3}$. Коефіцієнта β немає, бо у моделі не присутня змінна, яка прослідковує тренд.

ℓ_0 – перше значення, яке задає згладжений рівень часового ряду. Це значення обирається за рахунок оптимізації, тобто якимось певним методом перебору за допомогою використаних пакетів R відбувається вибір найбільш оптимального значення. За замовчуванням критерії оптимізації – це значення лог-правдоподібності "lik", яке використано у даній роботі.

$\sigma^2 = 0.0012$ – сума квадратів залишків, це мінімальне значення, яке можна досягти за такої специфікації моделі.

$s_0, s_{-1}, s_{-2}, s_{-3}$ – значення згладжування сезонності, так як у роботі використано квартальні дані, то модель бере їх 4 значення назад, якби були помісячні, то 12 значень назад і так далі. Індокси внизу – це значення змінної m , тобто скільки значень назад береться в залежності від сезонності часового ряду.

Дану модель не можна описати одним рівнянням як це зроблено для лінійної регресії або ARIMA моделі. ETS описується трьома рівняннями, а саме формули (3.3-3.5).

Модель ETS(M,N,A) описується наступними рівняннями (3.3-3.5) [30]:

Рівняння (формула) прогнозу:

$$y_t = (l_{t-1} + s_{t-m})(1 + \varepsilon_t), \quad (3.3)$$

Рівняння згладжування:

$$l_t = l_{t-1} + 0,886 \cdot (l_{t-1} + s_{t-m}) \cdot \varepsilon_t, \quad (3.4)$$

Рівняння сезонності:

$$s_t = s_{t-m} + 0,0001 \cdot (l_{t-1} + s_{t-m}) \cdot \varepsilon_t, \quad (3.5)$$

де ε_t – білий шум, тобто має середнє значення рівним нулю, нормальним розподілом та дисперсією σ^2 [31].

На рисунку 3.22 візуалізовано декомпозицію часового ряду:

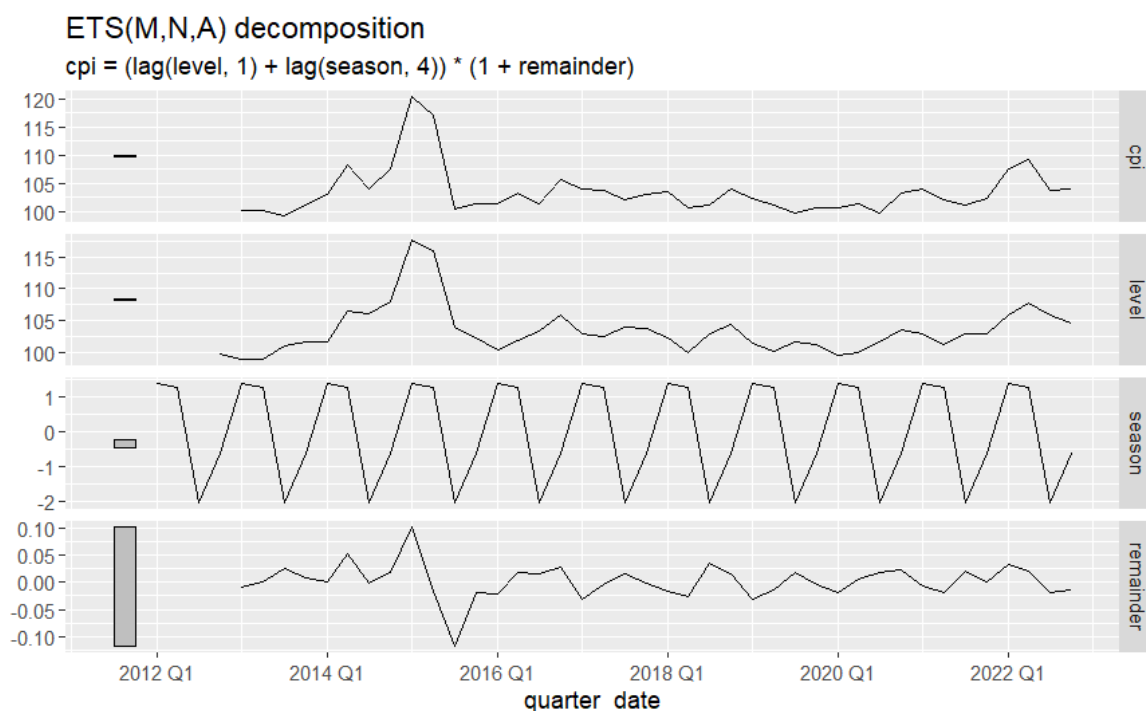


Рисунок 3.22 – Декомпозиція часового ряду ІСЦ

Джерело: розроблено автором на основі [10]

На вищенаведеному рис. 3.22 зображено декомпозицію, тобто компоненти моделі. Перший графік - просто значення ІСЦ. Другий - зміна рівнів часового ряду, третій сезонність даного часового ряду і четвертий - залишки.

Далі варто розглянути компоненти моделі (див. рис. 3.23):

0: 44 x 7 [1Q] 1: .model [1] 2: $cpi = (\text{lag}(\text{level}, 1) + \text{lag}(\text{season}, 4)) * (1 + \text{remainder})$

.model <chr>	quarter_date <S3: yearquarter>	cpi <dbl>	level <dbl>	season <dbl>	remainder <dbl>	.fitted <dbl>
auto	2012 Q1	NA	NA	1.3831141	NA	NA
auto	2012 Q2	NA	NA	1.2715120	NA	NA
auto	2012 Q3	NA	NA	-2.0400979	NA	NA
auto	2012 Q4	NA	99.76393	-0.6145282	NA	NA
auto	2013 Q1	100.1	98.83596	1.3830094	-1.035170e-02	101.14704
auto	2013 Q2	100.1	98.82934	1.2715112	-7.461519e-05	100.10747
auto	2013 Q3	99.2	100.96594	-2.0398568	2.490732e-02	96.78924
auto	2013 Q4	101.1	101.62940	-0.6144533	7.459650e-03	100.35141
auto	2014 Q1	103.0	101.61840	1.3830082	-1.204415e-04	103.01241
auto	2014 Q2	108.3	106.41324	1.2720523	5.258132e-02	102.88991

Рисунок 3.23 – Компоненти побудованої ETS моделі

Джерело: розроблено автором на основі [10]

Перші 4 значення таблиці були створені самою моделлю, бо у аналізованих даних ІСЦ присутні реальні дані лише з 2013 року.

Стовпець `level`, тобто зміна рівня зміни часового ряду. Перше його значення - ℓ_0 , тобто 3-ій рядок був також визначений на основі даної моделі.

Стовпець `season` – сезонна складова моделі, перші 4 значення, також були обраховані моделлю це і є початкові значення згладження сезонності - $s_0, s_{-1}, s_{-2}, s_{-3}$.

`Remainder` – залишки моделі.

`.fitted` – це стовпець, який містить спрогнозовані значення даною моделлю, на основі всіх складових.

Далі розглянуто аналіз залишків моделі на рисунку 3.24:

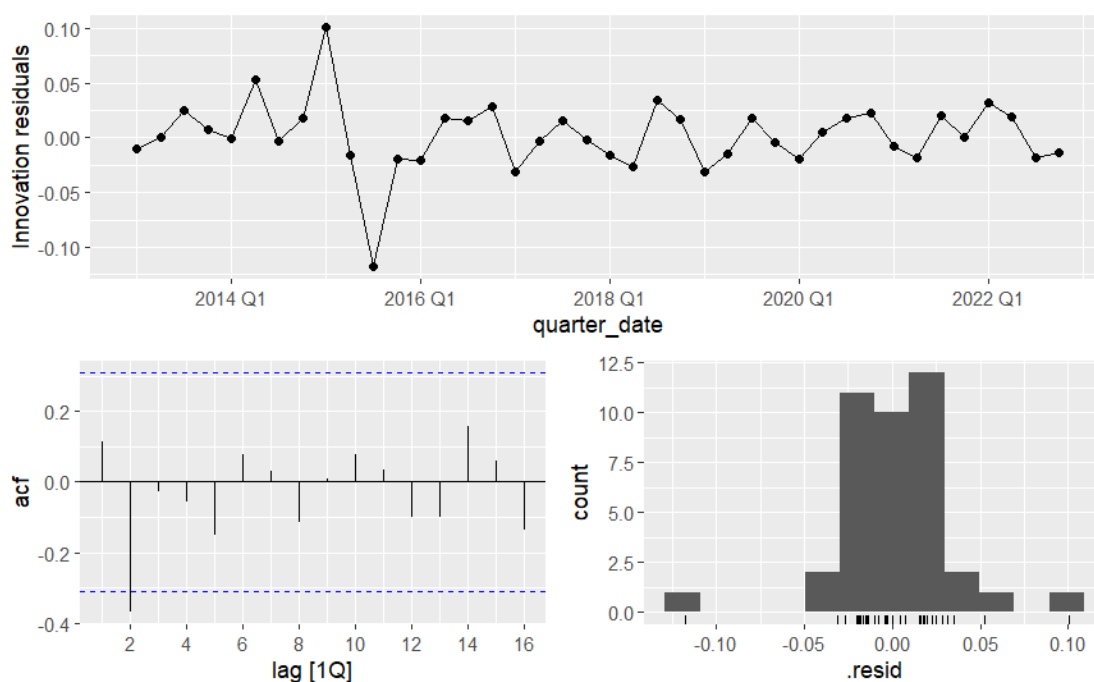


Рисунок 3.24 – Візуалізація залишків ETS моделі

Джерело: розроблено автором на основі [10]

За даним графіком (див. рис. 3.24) наявна можливість оцінити автокорельованість залишків. Всі значення автокореляції, окрім одного, по залишках не виходять за межі довірчих інтервалів, що вказує на те, що ці коефіцієнти автокореляції є статистично незначущими, це демонструє відсутність постійного зв'язку між значеннями.

Тобто, виходить, що модель доволі добре може прогнозувати майбутні значення, бо у залишках не лишилось якоїсь важливої інформації, яку модель ще б можна використати.

Отож, з рис. 3.24 видно, що залишки схожі на білий шум та середнє значення дорівнює 0. Також, прослідковується, що залишки не розподілені за нормальним законом, а це може негативно вплинути на можливість моделі прогнозувати.

Далі перевірено залишки моделі на автокореляцію за допомогою тесту Лjung-Бокса (див. рис. 3.25):

.model <chr>	lb_stat <dbl>	lb_pvalue <dbl>
auto	6.733957	0.009459455

Рисунок 3.25 – Тест Лjung-Бокса для ETS моделі

Джерело: розроблено автором на основі [10]

За тестом Лjung-Бокса отримано, що $p\text{-value}=0,009$, це значення є меншим за 0,05, відповідно є підстави відхилити нульову гіпотезу, яка говорить про те, що залишки моделі є білим шумом. Тобто, цей тест показує, що залишки моделі не є білим шумом, тобто між залишками присутня автокореляція. Але як раніше зображено на рис. 3.24, тобто на корелограмі автокореляції залишків, існує лише один викид, який виходить за межі довірчих інтервалів, тому таку автокореляцію цілком можна вважати несуттєвою.

3.3 Прогнозування індексу інфляції

Далі розглядатиметься прогнозування ІСЦ на основі побудованих моделей.

Потрібно перевірити як модель лінійної регресії, на основі раніше отриманого рівняння (формула 3.1) може прогнозувати значення індексу споживчих цін в Україні. Тобто, при зборі даних для даної роботи, були знайдені значення

деяких показників за 2023 рік, зокрема й індексу споживчих цін, їх і використано для перевірки прогнозних значень моделі. Всі інші значення показників, яких немає у вільному доступі були спрогнозовані за допомогою методу ковзного середнього.

Отож, у рівняння (формула 3.1) моделі підставляються нові (за 2023 рік) дані за незалежними змінними, для прогнозування майбутніх значень ІСЦ і отримуються обраховані значення індексу споживчих цін (рис. 3.26):

```
## {r}

# Розраховуємо значення індексу споживчих цін за допомогою рівняння
data2023$cpi_calculated <- with(data2023, {
  54.34 + 0.3253 * oblicova_stavka - 0.000007024 * gdp_real + 0.0001452 * borg - 0.000008151 * vudatky_dergavnogo_budgety +
  0.3562 * index_cin_vurob_promyslov_prodykcii
})

# Перевірка результату
print(data2023$cpi_calculated)

##
```

[1] 121.01261 108.76593 94.07444 106.43094

Рисунок 3.26 – Результати прогнозування ІСЦ за допомогою лінійної регресійної моделі на 2023 рік

Джерело: розроблено автором на основі [10]

На вищенаведеному рис. 3.26 представлено, що для першого кварталу 2023 року значення індексу споживчих цін становить 121,01%, для 2 кварталу 2023 року – 108,77%, для третього 94,07% і для четвертого кварталу 2023 року – 106,43%.

На рис. 3.27 зображено розраховані значення за допомогою рівняння лінійної регресійної моделі (стовпець 1) та реальні значення (стовпець 2) індексу споживчих цін:

cpi_calculated <dbl>	cpi <dbl>
121.01261	103.0
108.76593	101.5
94.07444	98.5
106.43094	102.0

Рисунок 3.27 – Порівняння реальних значень ІСЦ та прогнозованих значень за допомогою лінійної регресії на 4 квартали 2023 року

Джерело: розроблено автором на основі [10]

Варто зазначити, що реальні та розраховані значення досить сильно відрізняються. Такі результати скоріш за все вказують на певні недоліки моделі або вплив на модель неврахованих факторів.

Так як результати за регресійною моделлю є не надто якісними, у роботі використано ще 2 інші моделі, які містять більш складні розрахунки, але можуть дати кращі результати.

На наступному рис. 3.28 графічно відображено різницю між реальними і розрахованими значеннями індексу споживчих цін:

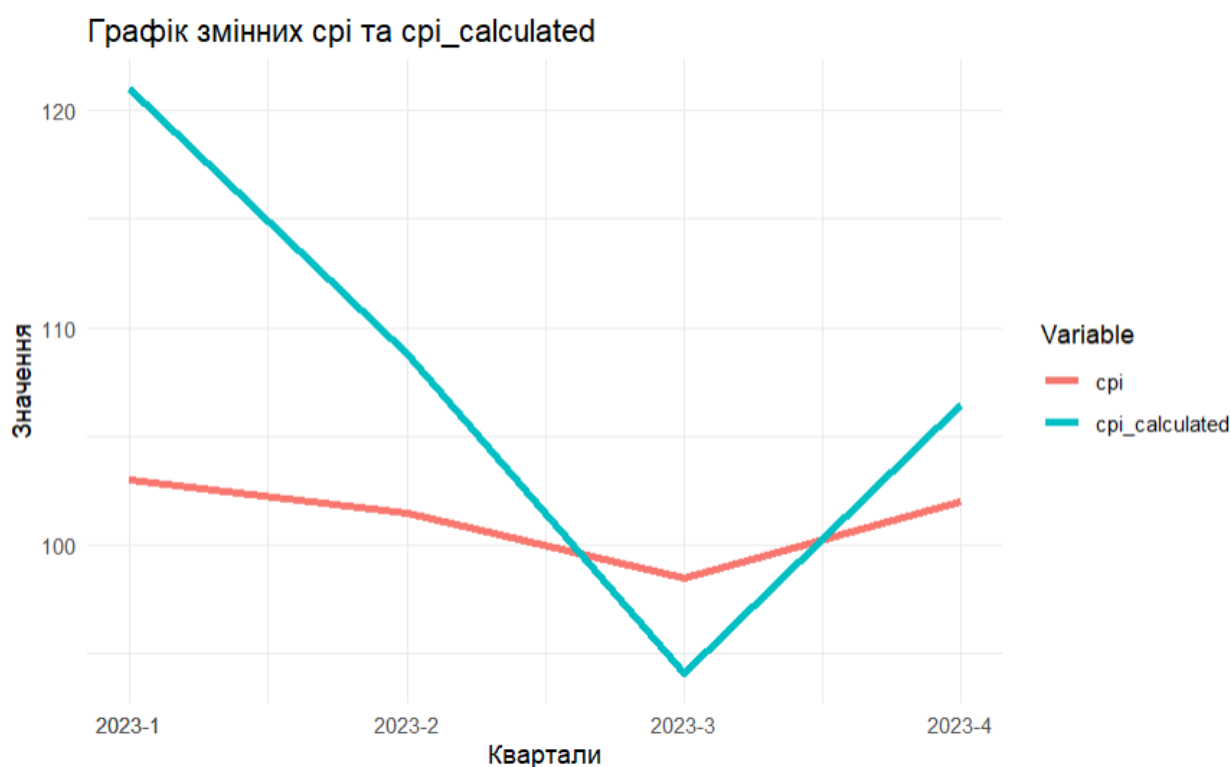


Рисунок 3.28 – Візуалізація реальних значень ІСЦ та прогнозованих значень за допомогою лінійної регресії

Джерело: розроблено автором на основі [10]

З візуальної оцінки стає відомо, що розраховані значення не дуже близькі до реальних, особливо у першому кварталі 2023 року.

Далі варто проаналізувати відсоткове співвідношення між реальними та розрахованими значеннями (рис. 3.29):

```

```{r}
Додаємо новий стовпець для збереження відсоткового співвідношення
data2023$percentage_relation <- (data2023$cpi_calculated / data2023$cpi) -1

Виведемо результати
print(data2023$percentage_relation)

```

```

```
[1] 0.17487974 0.07158548 -0.04492959 0.04344063
```

Рисунок 3.29 – Відсоткове співвідношення між реальними та розрахованими на основі лінійної регресії значеннями ІСЦ на 2023 рік

Джерело: розроблено автором на основі [10]

Тобто відомо, що для 1-го кварталу 2023 р. розраховане значення індексу споживчих цін на 17,49% вище за реальні значення, для 2-го кварталу розраховане значення за допомогою лінійної регресії більше за реальне на 7,16%. Для 3-го кварталу розраховане значення менше за реальне на 4,49%, а для 4-го кварталу 2023 року розрахований індекс споживчих цін вищий за реальний на 4,34% .

Побудовано прогноз на 4 періоди вперед на основі моделі ARIMA, тобто на 4 квартали 2023 року отримано такий результат (рис. 3.30):

| .model
<chr> | quarter_date
<S3: yearquarter> | cpi
<S3: distribution> | .mean
<dbl> |
|-----------------|-----------------------------------|---------------------------|----------------|
| auto | 2023 Q1 | N(104, 11) | 104.4189 |
| auto | 2023 Q2 | N(104, 18) | 103.6688 |
| auto | 2023 Q3 | N(103, 18) | 103.3597 |
| auto | 2023 Q4 | N(104, 18) | 103.6591 |

Рисунок 3.30 – Прогнозні значення ІСЦ на 4 квартали 2023 року на основі моделі ARIMA

Джерело: розроблено автором на основі [10]

Рисунок 3.30 не дає корисної інформації чи наочного представлення тому, на наступному рис. 3.31 візуалізовано прогнозовані значення:



Рисунок 3.31 – Прогнозні значення на основі моделі ARIMA за 2023 рік та реальні значення ІСЦ за 2013-2022 роки

Джерело: розроблено автором на основі [10]

Проаналізовано графік (рис. 3.31), і зроблено висновок, що, на перший погляд, значення спрогнозовані досить добре. Тобто отримані значення адекватно продовжують тенденцію даних минулих років.

Розглянуто ці прогнозовані значення (стовпець 1) разом із реальними (стовпець 2), як це зроблено для лінійної регресійної моделі (рис. 3.32):

| cpi_calculat_arima
<dbl> | cpi
<dbl> |
|------------------------------------|---------------------|
| 104.4189 | 103.0 |
| 103.6688 | 101.5 |
| 103.3597 | 98.5 |
| 103.6591 | 102.0 |

Рисунок 3.32 – Порівняння реальних значень ІСЦ та прогнозованих значень за допомогою ARIMA моделі на 4 квартали 2023 року

Джерело: розроблено автором на основі [10]

На рис. 3.32 зображено, що прогнозовані значення досить близькі до реальних, принаймні набагато ближчі, ніж ті, які були розраховані за допомогою попередньої моделі багатфакторної лінійної регресії.

Візуалізовано значення, для кращого сприйняття (рис. 3.33):

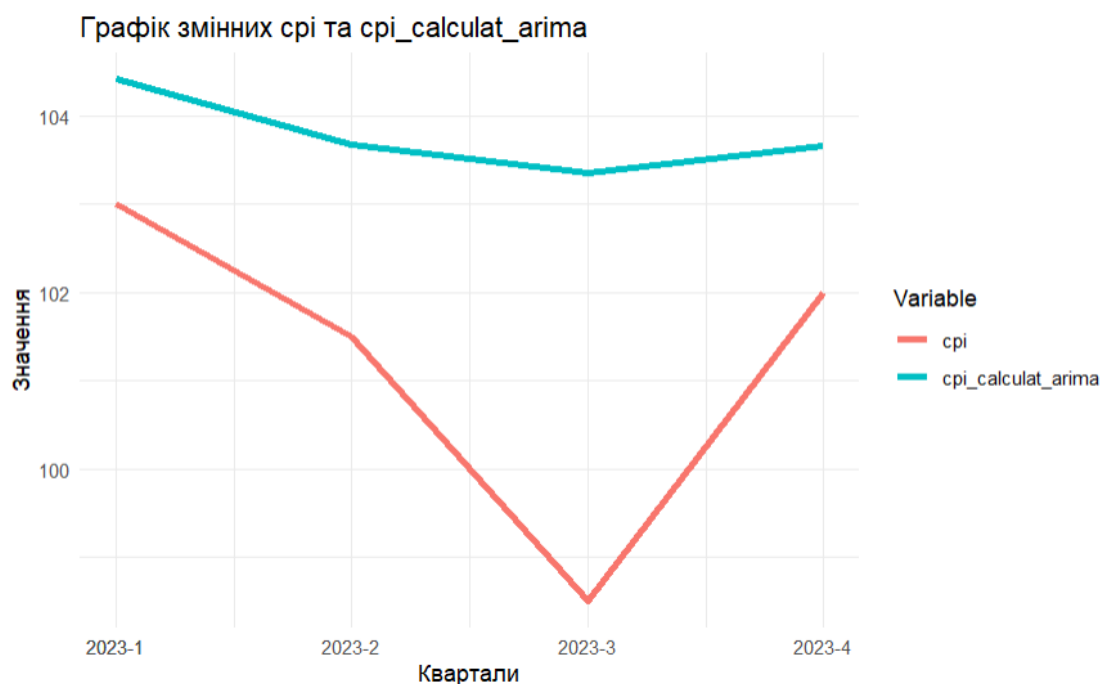


Рисунок 3.33 – Візуалізація реальних значень ІСЦ та прогнозованих значень за допомогою ARIMA моделі на 2023 рік

Джерело: розроблено автором на основі [10]

Отже, на графіку (рис. 3.33) відображено, що прогнозовані значення не ідентичні реальним, але модель досить добре спрогнозувала індекс споживчих цін.

Наступним, розглянуто відсоткове співвідношення між спрогнозованими і реальними значеннями за 4 квартали 2023 року (рис. 3.34):

```

```{r}
Додаємо новий стовпець для збереження відсоткового співвідношення
data2023$percentage_relation_arima <- (data2023$cpi_calculat_arima / data2023$cpi) -1

Виведемо результати
print(data2023$percentage_relation_arima)
```

```

[1] 0.01377613 0.02136770 0.04933749 0.01626547

Рисунок 3.34 – Відсоткове співвідношення між реальними та прогнозованими на основі ARIMA моделі значеннями ІСЦ на 2023 рік

Джерело: розроблено автором на основі [10]

Всі значення позитивні, це означає, що прогнозовані значення є вищими, а ніж реальні. Для першого кварталу 2023 року прогнозоване значення на 1,38% більше за реальне. Для другого кварталу – спрогнозоване більше за реальне на 2,14%, для 3-го – прогнозоване більше за реальне на 4,9%, і у 4-му кварталі 2023 року прогнозне значення більше за реальне на 1,63%. Отже, в цілому, прогнозовані значення досить близькі до реальних.

За допомогою побудованої моделі ETS створено прогноз на 4 періоди вперед, так як у роботі використані квартальні дані, то прогноз отримано на 2023 рік. На рис. 3.35 зображено прогнозовані значення за допомогою ETS моделі та реальні значення ІСЦ за даний період.

| cpi_calculat_ets
<dbl> | cpi
<dbl> |
|----------------------------------|---------------------|
| 106.0616 | 103.0 |
| 105.9502 | 101.5 |
| 102.6383 | 98.5 |
| 104.0645 | 102.0 |

Рисунок 3.35 – Відсоткове співвідношення між реальними та прогнозованими на основі ETS моделі значеннями ІСЦ на 2023 рік

Джерело: розроблено автором на основі [10]

Видно з рис. 3.35, що прогнозовані значення досить сильно відрізняються від реальних. В усіх періодах спрогнозовані значення є більшими за реальні. Тобто, можливо для подальших досліджень варто використати метод демпфування тренду. Наступним розглянуто візуалізацію реальних та спрогнозованих за допомогою ETS моделі значень, вона зображена на рис. 3.36.



Рисунок 3.36 – Прогнозні значення на основі ETS моделі за 2023 рік та реальні значення ІСЦ за 2013-2022 роки

Джерело: розроблено автором на основі [10]

Прогнозні значення зображені на рис. 3.36 досить непогано відповідають загальній тенденції ІСЦ за попередні роки. Далі порівняно отримані значення з реальними (див. рис. 3.37).

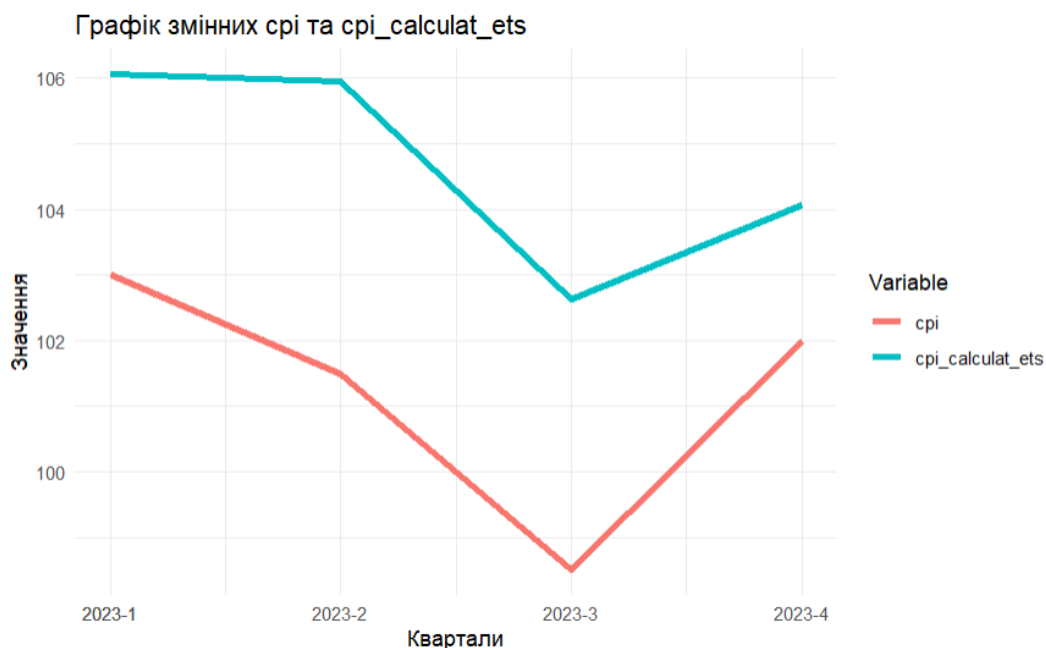


Рисунок 3.37 – Порівняння реальних значень ІСЦ та прогнозованих значень за допомогою ETS моделі на квартали 2023 рік

Джерело: розроблено автором на основі [10]

На рис. 3.37 видно, що модель ETS досить непогано спрогнозувала поведінку ІСЦ у часі, але показала дещо завищені значення, тобто спрогнозувала рівень інфляції вищим, ніж він був насправді.

Далі варто розглянути відсоткове співвідношення між спрогнозованими і реальними значеннями ІСЦ, це відображено на рис. 3.38.

```

```{r}
Додаємо новий стовпець для збереження відсоткового співвідношення
data2023$percentage_relation_ets <- (data2023$cpi_calculat_ets / data2023$cpi) -1

Виведемо результати
print(data2023$percentage_relation_ets)
```

```

[1] 0.02972467 0.04384399 0.04201320 0.02024052

Рисунок 3.38 – Відсоткове співвідношення між реальними та прогнозованими на основі ETS моделі значеннями ІСЦ на кварталам 2023 рік

Джерело: розроблено автором на основі [10]

Для 1 кварталу 2023 року отримане з моделі значення 2,97% більше за реальне. Для 2 кв. – більше за реальне на 4,38%, для 3-го – отримане значення більше за реальне на 4,2%, а у четвертому кварталі 2023 року спрогнозоване більше за реальне значення на 2,02%. Отже, значення, які були отримані за допомогою ETS моделі відрізняються від реальних.

3.4 Оцінка точності побудованих моделей

Останнім кроком у даній роботі є порівняння моделей створених у роботі за допомогою наступних критеріїв (рис. 3.39):

| model
<chr> | MAE
<dbl> | MSE
<dbl> | RMSE
<dbl> | MAPE
<dbl> | AIC
<dbl> | AICc
<dbl> |
|------------------|--------------|--------------|---------------|---------------|--------------|---------------|
| Лінійна регресія | 8.533762 | 104.116703 | 10.203759 | 8.370886 | 206.2073 | 210.8525 |
| ARIMA модель | 2.526646 | 8.271706 | 2.876057 | 2.518670 | 216.6919 | 218.4566 |
| ETS модель | 3.428660 | 12.641362 | 3.555469 | 3.395560 | 255.7172 | 259.2172 |

Рисунок 3.39 – Оцінка точності моделей: ARIMA, ETS та лінійної регресії

Джерело: розроблено автором на основі [10]

На рис. 3.39 наведено 6 критеріїв, а саме: Середня абсолютна помилка (Mean Absolute Error, MAE), визначає величину помилки прогнозу в цілому [36]. Середня квадратична помилка (Mean Squared Error, MSE), дана метрика враховує наявність рідкісних екстремальних помилок [36]. Корінь з середньої квадратичної помилки (Root Mean Squared Error, RMSE) [36]. Середня абсолютна відсоткова помилка (Mean Absolute Percentage Error, MAPE) [37 с. 46]. Показник MAPE має бути менше 10-15%, це свідчить про те, що прогностична якість моделі з досить високою, якщо ж значення моделі за цим показником перевищують рекомендований діапазон, це означає, що модель видаватиме досить погані прогнози. Інформаційний критерій Акаїке (AIC) - це математичний метод оцінки того, наскільки добре модель відповідає даним, на основі яких вона була згенерована [38]. Та скорегований інформаційний критерій AIC – AICс, він сильніше штрафує модель за складність. AICс є більш надійним показником моделей, а ніж AIC, особливо для моделей з невеликими обсягами спостережень, запобігаючи перенаванчання моделі. Варто зазначити, що, чим нижчі значення цих описаних похибок і інформаційного критерію, тим кращою є модель. Низькі значення вищенаведених похибок та помилок вказують на більш точні можливості моделі прогнозувати якісні значення на майбутні періоди.

Отже, можна спостерігати, що за показниками MAE, MSE, RMSE, MAPE модель APIMA має нижчі значення ніж лінійна регресія та ETS модель. Показник MAPE для всіх моделей має значення нижчі за 10%, це вказує, на те, що в цілому всі проаналізовані моделі можуть досить добре прогнозувати ІСЦ. Тобто, за похибками кращою є APIMA модель, але за інформаційним критерієм AIC, кращою є модель лінійної регресії, це дивно, але скоріш за все це пояснюється тим, що регресія має більшу ефективність використання інформації, а ніж APIMA та ETS моделі або ймовірно – регресія є більш простою моделлю і відповідно була менше «штрафована» за складність, хоча результати прогнозування регресії не можна вважати надто близькими до реальних значень.

ВИСНОВКИ

На початку роботи описано та узагальнено теоретичні знання про інфляцію, індекс споживчих цін та його розрахунку. За допомогою Deductor Studio Academic проаналізовано індекс споживчих цін по регіонах. Використовуючи R та MS Excel, побудовано модель множинної лінійної регресії, ARIMA та ETS моделі для аналізу та прогнозування індексу споживчих цін в Україні на чотири квартали 2023 року.

Для проведення регіонального аналізу використано карти Кохонена, які добре себе показали, у виявленні регіонів, які є схожими за своєю економічною поведінкою за показником ІСЦ. За допомогою регіонального аналізу виявлені значні відмінності між значеннями ІСЦ за 3 періоди, а саме допандемічний період (2019 рік), пандемічний (2021 рік) та період воєнного стану (2023 рік). Такий аналіз показав, що зміна значень індексу інфляції та переміщення регіонів між кластерами досить сильно залежить від зміни зовнішніх умов.

В цілому, за допандемічний період значення ІСЦ по регіонах не мали сильних відмінностей один від одного, тому карти Кохонена, виділили лише 1 загальний кластер. З початком пандемії COVID-19 значення ІСЦ в розрізі регіонів мали суттєві відмінності, тому даний період поділено на 3 кластери. Відмінності в значеннях індексу інфляції в різних регіонах, вказує на те, що пандемія вплинула на кожен регіон з різною силою. У період воєнного стану, регіон України були поділені на 2 кластери, тобто у порівнянні з пандемічним періодом, розрив між значеннями ІСЦ зменшився, але відмінності все ще присутні.

За результатами регіонального аналізу стало відомо, що найбільше між комірками карти Кохонена переміщувались Київська, Херсонська, Харківська та Полтавська області. За всі 3 аналізовані періоди, в середньому найбільші значення ІСЦ мали такі регіони: м. Київ, Київській, Харківській та Одеській областях, а найменші – Закарпатська, Сумська та Херсонська області.

Під час проведення кореляційного аналізу виявлено, що між багатьма зібраними змінними існує висока кореляція. При виборі змінних у регресійну

модель, де залежною змінною виступав ІСЦ, приділено особливо увагу вибору показників, які б не мали між собою високої кореляції, щоб запобігти появі мультиколінеарності у побудованих моделях. За допомогою розвідувального аналізу даних, розглянуто ІСЦ та інші впливові на нього фактори і описано ймовірні причини певних патернів поведінки даних показників. Регресійний аналіз використовувався для вивчення зв'язків між ІСЦ та такими показниками, як облікова ставка НБУ, реальний ВВП, валовий зовнішній борг, видатки Державного бюджету, індекс цін виробників промислової продукції та різними фіктивними змінними. Визначено найкращу регресійну модель. Також, побудовано ETS модель та оцінено дані ІСЦ на стаціонарність і на основі стаціонарного часового ряду побудовано ARIMA модель.

За допомогою моделі лінійної множинної регресії, ARIMA та ETS моделі створено короткостроковий прогноз на 4 квартали 2023 року, а потім результати кожної моделі порівнювались з реальними значеннями ІСЦ за цей період. Також, оцінено моделі за точністю та порівняно їх між собою. За показниками MAE, MAPE, RMSE, MSE кращою була визнана модель ARIMA. За інформаційними критеріями AIC та AICс кращою виходить множинна лінійна регресія, такі результати пояснюються тим, що лінійна модель має більшу ефективність використання інформації через те, що складається з інших змінних, окрім індексу споживчих цін. Також, такі значення інформаційних критеріїв можуть бути спричинені тим, що регресія є досить простою моделлю, а ARIMA та ETS моделі більш складними. Показники AIC та AICс, по суті, «штрафують» моделі за складність, відповідно регресія є менш «штрафованою», а інші моделі сильніше «штрафовані», хоча, якщо розглядати прогноз, то модель регресії не надто добре прогнозує значення у порівнянні з реальними значеннями ІСЦ.

Отже, при порівнянні прогнозованих значень з реальними за MAE, MSE, RMSE, MAPE, можна зробити висновок, що ARIMA модель показала кращі значення, тобто ближчі до реальних, ніж ETS та регресійна моделі. Тому підсумовуючи, з цією роботи стало відомо, що для подальших прогнозів варто використовувати модель ARIMA, бо вона показала найкращі результати.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Тінус Д. Що таке інфляція: поняття, причини, наслідки. *Daliz Finance* : веб-сайт. URL: <https://daliz.finance/blog/shcho-take-inflyatsiya-ponyattya-prychyny-naslidky> (дата звернення: 24.05.2024).
2. Гордійчук А. Ю. Моделі оцінювання впливу інфляційних процесів на економіку України : магістерська дисертація : 051. Київ. 2021. 117 с. URL: <https://ela.kpi.ua/server/api/core/bitstreams/130c13b1-bce5-420b-bda6c8b57c626b83/content> (дата звернення: 24.05.2024).
3. Саприкіна Т., Хайєк Ш. Інфляція, її причини та наслідки в сучасному світі. *Україна фінансова - інформаційно-аналітичний портал агентства фінансового розвитку* : веб-сайт. URL: https://ufin.com.ua/analit_mat/drn/166.htm (дата звернення: 23.05.2024).
4. Індекс інфляції в Україні 2024. *Мінфін* : веб-сайт. URL: <https://index.minfin.com.ua/ua/economy/index/inflation/> (дата звернення: 21.04.2024).
5. Індекс інфляції (роз'яснення). *Верховна Рада України* : веб-сайт. URL: <https://zakon.rada.gov.ua/rada/show/n0079697-06#Text> (дата звернення: 28.05.2024).
6. Інфляційний звіт. *Національний банк України*. 2023. URL: https://bank.gov.ua/admin_uploads/article/IR_2023-Q2.pdf (дата звернення: 23.05.2024).
7. Індекс споживчих цін: сприйняття та реальність : посіб. для користувачів. Київ : *Держкомстат України*. 2006. 56 с. URL: <https://www.lv.ukrstat.gov.ua/ukr/themes/13/posibnyk.pdf> (дата звернення: 23.05.2024).
8. Статистичний збірник «Індекси споживчих цін» 2023. *Державна служба статистики України*. URL: https://www.ukrstat.gov.ua/druk/publicat/kat_u/2024/03/zb%20ICC_2023.pdf (дата звернення: 21.04.2024).

9. Манцуров І. Г. та ін. Статистика ринків : практикум / за заг. ред. д-ра екон. наук, проф. І. Г. Манцура. Київ: КНЕУ, 2010. 498 с.
10. Індеси споживчих цін (інфляція). *Державна служба статистики України* : веб-сайт. URL: https://www.ukrstat.gov.ua/operativ/menu/menu_u/cit.htm (дата звернення: 08.05.2024).
11. Індеси споживчих цін за регіонами - Набори даних. *Єдиний державний веб-портал відкритих даних* : веб-сайт. URL: <https://data.gov.ua/dataset/c8511be5-a8f6-40ce-b4af-ba9ffeeae11c/resource/be6438f9-8d41-488f-a4e8-deddd0217a03f> (дата звернення: 25.05.2024).
12. Новожилова М.В., Петрова О.О., Чуб О.І. Інтелектуальний аналіз даних в аналітичній платформі Deductor : практикум. Харків. 2017. 160 с. URL: http://econ2016.krasalex.com/wp-content/uploads/2021/01/Navch.-pos_Laborator.-prakykum_Deductor_KNIT_122_bakalavry.pdf (дата звернення: 23.05.2024).
13. Мова програмування для обробки даних R-Studio. *RBC group* : веб-сайт. URL: <https://www.rbcgrp.com/ua/it-platformy/r-studio/> (дата звернення: 26.05.2024).
14. Матвійчук А.В. Штучний інтелект в економіці: нейронні мережі, нечітка логіка : монографія. К.: КНЕУ. 2011. 439 с. URL: https://library.kneu.edu.ua/cgi-bin/ecgi64/cgiirbis_64.exe?LNG=&C21COM=2&I21DBN=ELIB&P21DBN=ELIB&Z21ID=1884171100582165011&Image_file_name=MONOG_PUBL%5Cmatvijchuk%5Cartific_intelligence_econ.pdf (дата звернення: 23.05.2024).
15. Максимчук Т. М., Блохін О. Застосування самоорганізованих карт в умовах кластеризації. *Мехатронні системи і комп'ютерні технології. Вища математика*. 2018. URL: https://er.knutd.edu.ua/bitstream/123456789/11825/1/NRMSE2018_V2_P485-486.pdf (дата звернення: 24.05.2024).
16. Матвійчук А. В. Моделювання фінансової стійкості підприємств із застосуванням теорій нечіткої логіки, нейронних мереж і дискримінантного аналізу. *Вісник Національної академії наук України*. 2010. № 9. URL: <http://dspace.nbu.gov.ua/bitstream/handle/123456789/27566/03-Matviichuk.pdf?sequence=1> (дата звернення: 22.05.2024).

17. Смесова В. Л. Нейромережевий підхід до кластеризації країн за показниками, що характеризують процеси формування та реалізації економічних інтересів. *Науковий вісник Ужгородського національного університету*. 2018. Вип. 19(3). URL: http://www.visnyk-econom.uzhnu.uz.ua/archive/19_3_2018ua/12.pdf (дата звернення: 22.05.2024).
18. Кластерний аналіз. *Wiki THTU* : веб-сайт. URL: https://wiki.tntu.edu.ua/Кластерний_аналіз (дата звернення: 25.05.2024).
19. Статистика (МН). *Центр дистанційного навчання СНАУ*. URL: <https://cdn.snau.edu.ua/moodle/mod/page/view.php?id=10970&lang=uk> (дата звернення: 21.04.2024).
20. Чеманова А. О. Інтелектуальна система підтримки прийняття рішень (ІСППР) для дослідження актуарних фінансових ризиків : магістерська дисертація. Київ. 2024. URL: <https://ela.kpi.ua/server/api/core/bitstreams/9f43c383-21d1-401a-b371-8df20690412d/content> (дата звернення: 24.05.2024).
21. Методичні вказівки до виконання практичних робіт з дисципліни «Інтелектуальний аналіз даних» для студентів спеціальності 122 «Комп'ютерні науки» / уклад. Т.В. Федорончак. Запоріжжя: ЗНТУ. 2017. URL: http://eir.zntu.edu.ua/bitstream/123456789/2898/1/Fedoronchak_Methodical_instructions.pdf (дата звернення: 21.04.2024).
22. Аббадія Дж. Використання регресійного аналізу для розуміння складних взаємозв'язків. *Mind the Graph* : веб-сайт. URL: <https://mindthegraph.com/blog/uk/регресійний-аналіз/> (дата звернення: 21.04.2024).
23. Когтєв М. Дослідження залежності індексу споживчих цін. *ЛЮГОС. Мистецтво наукової думки*. 2019. Т. 3. URL: <https://ojs.ukrlogos.in.ua/index.php/2617-7064/article/download/143/127> (дата звернення: 08.05.2024).
24. Новик В.О. Моделювання процесу інфляції в Україні та вплив на нього економічних чинників. *«Інформаційні технології та моделювання систем»* : матеріали всеукр. студент. науково-практ. конф., м. Житомир, 25.04.2019 р. Житомир, 2019. С. 98–101. URL: http://ir.polissiauniver.edu.ua/bitstream/123456789/12041/1/Zbirnyk_KTMS_2019.pdf#page=98 (дата звернення: 08.05.2024).

25. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. OTexts, 2021. URL: <https://otexts.com/fpp3/> (дата звернення: 27.05.2024).

26. Катуніна О.С, Кваша Н.С. Прогнозне моделювання рівня інфляції в Україні на базі інструментів машинного навчання. *Моделювання та інформаційні технології в економіці*. К. КНЕУ. 2021. Вип.101. С. 93–116. URL: https://kneu.edu.ua/userfiles/zb_mise/101/9.pdf (дата звернення: 08.05.2024).

27. Shapovalenko N. A Suite of Models for CPI Forecasting. *Visnyk of the National Bank of Ukraine*. 2021. No. 252. P. 4–36. URL: <https://doi.org/10.26531/vnbu2021.252.01> (дата звернення: 08.05.2024).

28. Козачок Л. Аналіз та прогнозування часових рядів динаміки значень досліджуваної величини за допомогою експоненціального згладжування методом Брауна. *Автомобіль і електроніка. Сучасні технології*. 2017. Т. 12. URL : https://rcf.khadi.kharkov.ua/fileadmin/F-HIGHWAY/Інформатики_і_прикладної_математики/veit_2017_12_28.pdf (дата звернення: 24.05.2024).

29. Підхід до прогнозування дієвості державного управління з використанням технологій OSINT / Додонов О.Г. та ін. *Інформаційні технології та безпека* : матеріали XIX міжнар. наук.-практ. конф. 2019. С. 230-233. URL: http://e.ieu.edu.ua/bitstream/123456789/719/1/%d0%94%d0%be%d0%b4%d0%be%d0%bd%d0%be%d0%b2_%d0%9f%d1%96%d0%b4%d1%85%d1%96%d0%b4_2019.pdf (дата звернення: 24.05.2024).

30. Svetunkov I. Forecasting and Analytics with the Augmented Dynamic Adaptive Model (ADAM). 2024. URL: <https://openforecast.org/adam/> (дата звернення: 24.05.2024).

31. Воробець І.В. Моделі та методи прогнозування часових рядів із властивістю циклічності : кваліфікаційна робота магістра за спеціальністю «122 — комп'ютерні науки». Тернопіль : ТНТУ, 2023. 76 с. URL: https://elartu.tntu.edu.ua/bitstream/lib/41161/1/2023_KRM_SNm-61_Vorobets_I_V.pdf (дата звернення: 24.05.2024).

32. Міністерство фінансів України : веб-сайт. URL: <http://www.minfin.com.ua> (дата звернення: 21.04.2024).

33. Мазур М. В., Сіренко К. В. Статистичний аналіз регіональних відмінностей інфляції в Україні. *Вісник Хмельницького національного університету*. 2017. № 5. URL: http://journals.khnu.km.ua/vestnik/pdf/ekon/pdfbase/2017/2017_5/jrn/pdf/8.pdf (дата звернення: 23.05.2024).

34. Національний банк України: веб-сайт. URL: <https://bank.gov.ua/> (дата звернення: 21.04.2024).

35. Інформація Міністерства фінансів України про виконання Державного бюджету України за 2022 рік. *Міністерство фінансів України*. URL: https://mof.gov.ua/storage/files/Інформація_щодо_викон_ДБУ_за_2022_рік_01_03_2023.doc.pdf (дата звернення: 21.04.2024).

36. Критерії вибору перцептронної моделі для прогнозування: аналіз і практичні рекомендації щодо їх використання / Руденко О.Г та ін. *Біоніка інтелекту*. 2018. Т. 2, № 91. С. 31–40. URL: <https://journals.indexcopernicus.com/api/file/viewByFileId/645659.pdf> (дата звернення: 21.04.2024).

37. Красножон К. Г. Інформаційно-аналітична система для аналізу якості прогнозування часових рядів різної природи : дипломний проєкт. Київ. 2021. 118 с. URL: <https://ela.kpi.ua/server/api/core/bitstreams/90bfc703-f8b4-488f-9d4f-6715c38c0f02/content> (дата звернення: 21.04.2024).

38. Bevans R. Akaike Information Criterion. When & How to Use It (Example). *Scribbr* : веб-сайт. URL: <https://www.scribbr.com/statistics/akaike-information-criterion/> (дата звернення: 21.04.2024).

ДОДАТКИ

Додаток А

Таблиця А.1 – Опис комірок кластеру матриці щільності за допандемічний період

| № комірки | Назва регіону та номери кварталів 2019 року |
|-----------|--|
| 1 | м.Київ (1 кв.), Харківська (1 кв.) |
| 2 | Херсонська (1 кв.) |
| 3 | Полтавська (1 кв.) |
| 4 | Дніпропетровська (1 кв.), Запорізька (1 кв.), Київська (1 кв.) |
| 5 | Кіровоградська (1 кв.), Чернігівська (1 кв.) |
| 6 | Київська (4 кв.), Сумська (4 кв.), Харківська (4 кв.) |
| 7 | Черкаська (2 кв.), Хмельницька (4 кв.) |
| 8 | Волинська (3 кв.), Житомирська (4 кв.), Черкаська (4 кв.) |
| 9 | Донецька (1 кв.), Івано-Франківська (1 кв.), Одеська (1 кв.) |
| 10 | Івано-Франківська (2 кв.) |
| 11 | Житомирська (2 кв.), Донецька (4 кв.), Луганська (4 кв.), Чернігівська (4 кв.) |
| 12 | Харківська (2 кв.), Донецька (3 кв.), Рівненська (3 кв.), Кіровоградська (4 кв.) |
| 13 | Житомирська (3 кв.), Рівненська (4 кв.), Тернопільська (4 кв.) |
| 14 | Миколаївська (1 кв.), Рівненська (1 кв.), Сумська (1 кв.), Чернівецька (1 кв.) |
| 15 | Луганська (3 кв.) |
| 16 | Закарпатська (4 кв.) |
| 17 | Волинська (2 кв.) |
| 18 | Миколаївська (3 кв.), Харківська (3 кв.), Херсонська (4 кв.) |
| 19 | Запорізька (2 кв.), м.Київ (2 кв.) |
| 20 | Закарпатська (2 кв.), Хмельницька (2 кв.), Івано-Франківська (4 кв.), Львівська (4 кв.), Одеська (4 кв.), Полтавська (4 кв.) |
| 21 | Львівська (3 кв.) |
| 22 | Запорізька (3 кв.), Київська (3 кв.), Черкаська (3 кв.), Чернігівська (3 кв.) |
| 23 | Донецька (2 кв.) |
| 24 | Вінницька (2 кв.) |
| 25 | Житомирська (1 кв.), Львівська (1 кв.) |
| 26 | Луганська (1 кв.) |
| 27 | Волинська (4 кв.), Одеська (2 кв.), Полтавська (2 кв.), Тернопільська (2 кв.) |
| 28 | Чернівецька (2 кв.) |

| | |
|----|--|
| 29 | Чернівецька (3 кв.) |
| 30 | Херсонська (2 кв.) |
| 31 | Київська (2 кв.), Кіровоградська (2 кв.) |
| 32 | Вінницька (4 кв.), Дніпропетровська (4 кв.), Миколаївська (4 кв.), Тернопільська (3 кв.), Хмельницька (3 кв.) |
| 33 | Вінницька (3 кв.), м.Київ (3 кв.), Полтавська (3 кв.) |
| 34 | Сумська (3 кв.) |
| 35 | Вінницька (1 кв.), Волинська (1 кв.), Сумська (2 кв.), Тернопільська (1 кв.), Хмельницька (1 кв.), Черкаська (1 кв.) |
| 36 | Рівненська (2 кв.), Чернівецька (4 кв.), Чернігівська (2 кв.) |
| 37 | Миколаївська (2 кв.) |
| 38 | Закарпатська (1 кв.) |
| 39 | Дніпропетровська (2 кв.), Луганська (2 кв.), Львівська (2 кв.) |
| 40 | Запорізька (4 кв.), м.Київ (4 кв.) |
| 41 | Закарпатська (3 кв.), Кіровоградська (3 кв.) |
| 42 | Дніпропетровська (3 кв.), Івано-Франківська (3 кв.), Одеська (3 кв.) |
| 43 | Херсонська (3 кв.) |

Таблиця А.2 – Опис комірок кластеру матриці щільності за пандемічний період

| № комірки | Назва регіону та номери кварталів 2021 року |
|-----------|---|
| 1 | Кіровоградська (1 кв.) |
| 2 | Херсонська (1 кв.) |
| 3 | Рівненська (1 кв.) |
| 4 | Запорізька (4 кв.), м.Київ (4 кв.), Миколаївська (2 кв.), Рівненська (2 кв.), Хмельницька (2 кв.), Черкаська (3 кв.), Черкаська (4 кв.), Чернігівська (2 кв.) |
| 5 | Київська (1 кв.), Полтавська (1 кв.) |
| 6 | Волинська (1 кв.), Чернігівська (1 кв.) |
| 7 | Одеська (4 кв.), Хмельницька (4 кв.) |
| 8 | Житомирська (4 кв.) |
| 9 | Волинська (2 кв.), Полтавська (4 кв.) |
| 10 | Львівська (2 кв.), Чернівецька (2 кв.) |
| 11 | Донецька (1 кв.), Луганська (1 кв.) |
| 12 | Харківська (1 кв.), Черкаська (1 кв.) |
| 13 | Вінницька (1 кв.), Волинська (4 кв.), Одеська (2 кв.), Харківська (4 кв.) |

| | |
|----|--|
| 14 | Херсонська (3 кв.) |
| 15 | м.Київ (1 кв.) |
| 16 | Вінницька (3 кв.), Закарпатська (3 кв.), Кіровоградська (3 кв.) |
| 17 | Миколаївська (3 кв.) |
| 18 | Запорізька (1 кв.), Миколаївська (1 кв.) |
| 19 | Закарпатська (1 кв.), Одеська (1 кв.) |
| 20 | Луганська (4 кв.) |
| 21 | Житомирська (2 кв.), Київська (2 кв.), Сумська (2 кв.), Хмельницька (3 кв.) |
| 22 | Сумська (3 кв.) |
| 23 | Вінницька (2 кв.), Дніпропетровська (2 кв.), Херсонська (2 кв.) |
| 24 | Закарпатська (2 кв.) |
| 25 | Житомирська (3 кв.) |
| 26 | Львівська (1 кв.), Сумська (1 кв.) |
| 27 | Полтавська (3 кв.) |
| 28 | Запорізька (2 кв.) |
| 29 | Донецька (4 кв.), Івано-Франківська (2 кв.), Чернівецька (4 кв.) |
| 30 | Дніпропетровська (3 кв.), Тернопільська (3 кв.) |
| 31 | Івано-Франківська (3 кв.), Київська (3 кв.), Чернігівська (3 кв.) |
| 32 | Дніпропетровська (1 кв.), Івано-Франківська (1 кв.) |
| 33 | м.Київ (2 кв.) |
| 34 | Житомирська (1 кв.), Хмельницька (1 кв.), Чернівецька (1 кв.) |
| 35 | Чернігівська (4 кв.) |
| 36 | Закарпатська (4 кв.), Івано-Франківська (4 кв.), Львівська (4 кв.) |
| 37 | Кіровоградська (2 кв.), Донецька (3 кв.), Одеська (3 кв.), Рівненська (3 кв.), Чернівецька (3 кв.) |
| 38 | Львівська (3 кв.) |
| 39 | Тернопільська (2 кв.), Луганська (2 кв.), Харківська (2 кв.) |
| 40 | Волинська (3 кв.), Луганська (3 кв.), Херсонська (4 кв.) |
| 41 | м.Київ (3 кв.) |
| 42 | Тернопільська (1 кв.) |
| 43 | Донецька (2 кв.), Черкаська (2 кв.), Сумська (4 кв.) |
| 44 | Вінницька (4 кв.), Тернопільська (4 кв.) |
| 45 | Дніпропетровська (4 кв.), Київська (4 кв.), Кіровоградська (4 кв.), Миколаївська (4 кв.), Рівненська (4 кв.) |
| 46 | Полтавська (2 кв.), Харківська (3 кв.) |
| 47 | Запорізька (3 кв.) |

Таблиця А.3 – Опис комірок кластеру матриці щільності за період воєнного стану

| № комірки | Назва регіону та номери кварталів 2023 року |
|-----------|---|
| 1 | Закарпатська (1 кв.), м.Київ (1 кв.) |
| 2 | Одеська (1 кв.) |
| 3 | Запорізька (1 кв.), Львівська (1 кв.) |
| 4 | Донецька (1 кв.), Луганська (1 кв.), Чернівецька (1 кв.) |
| 5 | Івано-Франківська (2 кв.), Миколаївська (2 кв.), Житомирська (4 кв.), Херсонська (4 кв.) |
| 6 | Вінницька (2 кв.), Чернівецька (2 кв.) |
| 7 | Миколаївська (1 кв.) |
| 8 | Івано-Франківська (1 кв.) |
| 9 | Херсонська (1 кв.) |
| 10 | Волинська (3 кв.), Рівненська (3 кв.) |
| 11 | Харківська (1 кв.) |
| 12 | Дніпропетровська (1 кв.), Кіровоградська (1 кв.), м.Київ (4 кв.) |
| 13 | Житомирська (2 кв.), Кіровоградська (4 кв.), Луганська (4 кв.), Львівська (4 кв.), Чернігівська (4 кв.) |
| 14 | Чернігівська (2 кв.) |
| 15 | Волинська (1 кв.), Житомирська (1 кв.), Рівненська (1 кв.) |
| 16 | Луганська (2 кв.), Черкаська (2 кв.), Кіровоградська (2 кв.) |
| 17 | Львівська (2 кв.), Рівненська (2 кв.) |
| 18 | Закарпатська (3 кв.), Полтавська (3 кв.) |
| 19 | Сумська (2 кв.) |
| 20 | Сумська (3 кв.), Харківська (3 кв.) |
| 21 | Херсонська (2 кв.) |
| 22 | Черкаська (4 кв.) |
| 23 | Сумська (1 кв.), Хмельницька (1 кв.), Чернівецька (4 кв.) |
| 24 | Полтавська (4 кв.), Сумська (4 кв.), Хмельницька (4 кв.) |
| 25 | Волинська (2 кв.), Дніпропетровська (2 кв.), Харківська (2 кв.) |
| 26 | Київська (4 кв.) |
| 27 | Тернопільська (2 кв.) |
| 28 | Київська (3 кв.) |
| 29 | Вінницька (3 кв.), Тернопільська (3 кв.) |
| 30 | Івано-Франківська (3 кв.) |

| | |
|----|--|
| 31 | Волинська (4 кв.), Полтавська (1 кв.), Чернігівська (1 кв.) |
| 32 | Донецька (4 кв.), Запорізька (4 кв.) |
| 33 | Черкаська (3 кв.) |
| 34 | Хмельницька (3 кв.) |
| 35 | Донецька (3 кв.), Запорізька (3 кв.), Львівська (3 кв.) |
| 36 | м.Київ (2 кв.) |
| 37 | Київська (2 кв.), Одеська (4 кв.) |
| 38 | Чернігівська (3 кв.) |
| 39 | Житомирська (3 кв.) |
| 40 | Чернівецька (3 кв.) |
| 41 | Вінницька (4 кв.), Київська (1 кв.), Миколаївська (4 кв.), Тернопільська (1 кв.) |
| 42 | Дніпропетровська (4 кв.), Івано-Франківська (4 кв.), Одеська (2 кв.), Рівненська (4 кв.), Харківська (4 кв.) |
| 43 | Кіровоградська (3 кв.), Луганська (3 кв.) |
| 44 | Закарпатська (4 кв.), Тернопільська (4 кв.), Черкаська (1 кв.) |
| 45 | Вінницька (1 кв.), Запорізька (2 кв.) |
| 46 | Донецька (2 кв.), Хмельницька (2 кв.) |
| 47 | Закарпатська (2 кв.), Полтавська (2 кв.) |
| 48 | Дніпропетровська (3 кв.) |
| 49 | Миколаївська (3 кв.) |
| 50 | Одеська (3 кв.), Херсонська (3 кв.) |
| 51 | м.Київ (3 кв.) |

Таблиця Б.1 – Таблиця назв змінних, які були зібрані для аналізу

| № | Назва змінної | Опис |
|----|--------------------------------------|--|
| 1 | quarter | У цій змінній вказані квартали, за які було зібрано усі дані |
| 2 | oblicova_stavka | Облікова ставка НБУ, % |
| 3 | groshova_baza | Грошова база, млн грн |
| 4 | kurs_hryvni | Офіційний курс гривні до долара США (за 1 од.) |
| 5 | vutraty_naselen | Витрати населення, млн. грн |
| 6 | cash | Готівкові кошти, випущені в обіг, млн грн |
| 7 | index_cin_vurob_promyslo_v_prodykcii | Індекс цін виробників промислової продукції, % |
| 8 | kilkist_bezrobit | Кількість безробітних, тис. осіб |
| 9 | zoltovalutn_reservey | Золотовалютні резерви України, млн. дол. США |
| 10 | borg | Валовий зовнішній борг, млн дол. США |
| 11 | riven_bezrobit | Рівень безробіття в Україні, % |
| 12 | invest_v_ukraine | Прямі інвестиції в Україну, млн. дол. США |
| 13 | vudatky_derzhavnogo_budgety | Видатки Державного бюджету, млн. грн |
| 14 | deficit_derzhavnogo_budgety | Дефіцит Державного бюджету, млн. грн |
| 15 | import | Імпорт товарів та послуг, млн. дол. США |
| 16 | export | Експорт товарів та послуг, млн. дол. США |
| 17 | gdp_nom | Номінальний ВВП, млн. грн |
| 18 | gdp_real | Реальний ВВП, млн. грн |
| 19 | cpi | Індекс споживчих цін, % |

Код R, який був використаний у роботі

```

library(readxl)
library(dplyr)
library(broom)
library(combinat)
library(purrr)
library(fpp3)
library(tidyr)
library(GGally)
library(repr)
library(corrplot)
library(car)
library(mctest)
library(ppcor)
library(lmtest)
library(tsibble)
library(feasts)
library(ggplot2)
library(forecast)
library(gridExtra)
library(ggrepel)
library(Metrics)
data_kart_koh <- read_excel("C:/Users/User/Desktop/region_cpi_2019_2023.xlsx")
head(data_kart_koh)
data_kart_koh$quarter_date <- as.yearqtr(paste(data_kart_koh$year, data_kart_koh$quarter), format =
"%Y %q")
print(data_kart_koh)
data_kart_koh_avg <- aggregate(data_kart_koh$cpi, by=list(Category=data_kart_koh$region),
FUN=mean)
ggplot(data_kart_koh_avg, aes(x=Category, y=x, fill=Category)) +
  geom_bar(stat="identity", color="black") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  labs(x="Region", y="Середнє значення ІЦІ", fill="Region") +
  ggtitle("Середнє значення ІЦІ за 2019-2023рр. для кожного регіону")
data <- read_excel("C:/Users/User/Desktop/skorocheni_dani.xlsx")
head(data)
quarters <- substr(data$quarter, 1, 2)
years <- substr(data$quarter, 4, 7)
data$quarter_date <- as.yearqtr(paste(years, quarters))
print(data)
cor_matrix <- cor(data[,2:17])
corrplot(cor_matrix, method = "number")
ggplot(data, aes(x = quarter_date, y = cpi)) +
  geom_line(color = "blue", size = 1.2) +
  theme_minimal() +
  ggtitle("Графік динаміки індексу споживчих цін в Україні,%") +
  xlab("Квартали") +
  ylab("Значення індексу споживчих цін, %")
p1 <- ggplot(data, aes(x = 1:nrow(data), y = cpi)) +
  geom_line(color = "blue") +
  theme_minimal() +
  ggtitle("Графік змінної cpi") +
  xlab("Номер кварталу") +

```

```

  ylab("Значення")
p2 <- ggplot(data, aes(x = 1:nrow(data), y = oblicova_stavka)) +
  geom_line(color = "green") +
  theme_minimal() +
  ggtitle("Графік змінної oblicova_stavka") +
  xlab("Номер кварталу") +
  ylab("Значення")
p3 <- ggplot(data, aes(x = 1:nrow(data), y = gdp_real)) +
  geom_line(color = "red") +
  theme_minimal() +
  ggtitle("Графік змінної gdp_real") +
  xlab("Номер кварталу") +
  ylab("Значення")
p4 <- ggplot(data, aes(x = 1:nrow(data), y = borg)) +
  geom_line(color = "purple") +
  theme_minimal() +
  ggtitle("Графік змінної borg") +
  xlab("Номер кварталу") +
  ylab("Значення")
p5 <- ggplot(data, aes(x = 1:nrow(data), y = vudatky_derzavnogo_budgety)) +
  geom_line(color = "#00CED1") +
  theme_minimal() +
  ggtitle("Графік змінної vudatky_derzavnogo_budgety") +
  xlab("Номер кварталу") +
  ylab("Значення")
p6 <- ggplot(data, aes(x = 1:nrow(data), y = index_cin_vurob_promyslov_prodykcii)) +
  geom_line(color = "magenta") +
  theme_minimal() +
  ggtitle("Графік змінної index_cin_vurob_promyslov_prodykcii") +
  xlab("Номер кварталу") +
  ylab("Значення")
grid.arrange(p1, p2, p3, p4, p5, p6, ncol = 2)
line_reg <-
lm(cpi~oblicova_stavka+gdp_real+borg+vudatky_derzavnogo_budgety+index_cin_vurob_promyslov_prodykci
i, data=data)
line_reg
summary (line_reg)
line_reg1 <-
lm(cpi~oblicova_stavka+gdp_real+borg+vudatky_derzavnogo_budgety+index_cin_vurob_promyslov_prodykci
i+q2+q3+q4, data=data)
summary (line_reg1)
line_reg2 <-
lm(cpi~oblicova_stavka+gdp_real+borg+vudatky_derzavnogo_budgety+index_cin_vurob_promyslov_prodykci
i+covid, data=data)
summary (line_reg2)
line_reg3 <-
lm(cpi~oblicova_stavka+gdp_real+borg+vudatky_derzavnogo_budgety+index_cin_vurob_promyslov_prodykci
i+war, data=data)
summary (line_reg3)
line_reg4 <-
lm(cpi~oblicova_stavka+gdp_real+borg+vudatky_derzavnogo_budgety+index_cin_vurob_promyslov_prodykci
i+covid+war, data=data)
summary (line_reg4)
durbinWatsonTest((line_reg))
omcdiag (line_reg)
imcdiag (line_reg)

```

```

bptest(line_reg)
data2023 <- read_excel("C:/Users/User/Desktop/data_tilku_2023.xlsx")
head(data2023)
quarters <- substr(data2023$quarter, 1, 2)
years <- substr(data2023$quarter, 4, 7)
data2023$quarter_date <- as.yearqtr(paste(years, quarters))
print(data2023)
quarters <- substr(data1$quarter, 1, 2)
years <- substr(data1$quarter, 4, 7)
data1$quarter_date <- as.yearqtr(paste(years, quarters))
print(data1)
data_cpi <- data1 %>%
  dplyr::select(quarter_date, cpi)%>%
  mutate(quarter_date = yearquarter(quarter_date)) %>%
  as_tsibble(index = quarter_date)
data_cpi
data_cpi %>%
  autoplot(cpi) +
  labs(title = "Індекс споживчих цін в Україні", y = "%")
data_cpi %>%
  features(cpi, list(coef_hurst))
data_cpi %>% ACF(cpi, lag_max = 40) %>% autoplot()
data_cpi %>%
  features(cpi, unitroot_kpss)
data_cpi %>%
  features(cpi, unitroot_ndiffs)
data_cpi %>%
  gg_tsdisplay(cpi, plot_type='partial')
fit <- data_cpi %>%
  model(auto = ARIMA(cpi))
report(fit)
glance(fit) %>% dplyr::select(.model:BIC)
fit %>% gg_tsresiduals()
augment(fit) %>%
  features(.innov, ljung_box, lag = 4, dof = 3)
fit_ets <- data_cpi %>%
  model(auto = ETS(cpi))
report(fit_ets)
components(fit_ets) %>% autoplot()
components(fit_ets) %>%
  left_join(fitted(fit_ets), by = c(".model", "quarter_date"))
fit_ets %>% gg_tsresiduals()
augment(fit_ets) %>%
  features(.innov, ljung_box, lag = 4, dof = 3)
data2023$cpi_calculated <- with(data2023, {
  54.34 + 0.3253 * oblicova_stavka - 0.000007024 * gdp_real + 0.0001452 * borg - 0.000008151 *
vudatky_dergavnogo_budgety + 0.3562 * index_cin_vurob_promyslov_prodykcii})
print(data2023$cpi_calculated)
select_cpi <- dplyr::select(data2023, cpi_calculated, cpi)
select_cpi
ggplot(data2023, aes(x = quarter_date)) +
  geom_line(aes(y = cpi, color = "cpi"), size = 1.5) +
  geom_line(aes(y = cpi_calculated, color = "cpi_calculated"), size = 1.5) +
  labs(color = "Variable") +
  theme_minimal() +
  ggtitle("Графік змінних cpi та cpi_calculated") +

```

```

xlab("Квартали") +
ylab("Значення")
data2023$percentage_relation <- (data2023$cpi_calculated / data2023$cpi) -1
print(data2023$percentage_relation)
fc <- fit %>%
  forecast(h=4)
fc
fc <- fc %>%
  mutate( dm80 = hilo(cpi, 80), dm95 = hilo(cpi, 95))%>%
  mutate(
    lower80 = dm80$lower,
    upper80 = dm80$upper,
    lower95 = dm95$lower,
    upper95 = dm95$upper,)
combined_data <- dplyr::bind_rows(data_cpi, fc)
combined_data$cpi <- as.numeric(mean(combined_data$cpi))
combined_data$.mean <- as.numeric(mean(combined_data$.mean))
ggplot() +
  geom_line(data = combined_data, aes(x = quarter_date, y = cpi), color = "black") +
  geom_line(data = combined_data, aes(x = quarter_date, y = .mean), color = "blue") +
  geom_ribbon(data = combined_data, aes(x = quarter_date, ymin = lower80, ymax = upper80), fill =
"lightblue", alpha = 0.5) +
  geom_ribbon(data = combined_data, aes(x = quarter_date, ymin = lower95, ymax = upper95), fill =
"lightgreen", alpha = 0.5) +
  labs(title = "Прогнозовані значення індексу споживчих цін з довірчими інтервалами", x =
"Квартальна дата", y = "Індекс споживчих цін") +
  theme_minimal()
data2023$cpi_calculat_arma <- fc$.mean
select_cpi <- dplyr::select(data2023, cpi_calculat_arma, cpi)
select_cpi
ggplot(data2023, aes(x = quarter_date)) +
  geom_line(aes(y = cpi, color = "cpi"), size = 1.5) +
  geom_line(aes(y = cpi_calculat_arma, color = "cpi_calculat_arma"), size = 1.5) +
  labs(color = "Variable") +
  theme_minimal() +
  ggtitle("Графік змінних cpi та cpi_calculat_arma") +
  xlab("Квартали") +
  ylab("Значення")
data2023$percentage_relation_arma <- (data2023$cpi_calculat_arma / data2023$cpi) -1
print(data2023$percentage_relation_arma)
fc_ets <- fit_ets %>%
  forecast(h=4)
fc_ets
fc_ets <- fc_ets %>%
  mutate( dm80 = hilo(cpi, 80), dm95 = hilo(cpi, 95))%>%
  mutate(
    lower80 = dm80$lower,
    upper80 = dm80$upper,
    lower95 = dm95$lower,
    upper95 = dm95$upper,)
data2023$cpi_calculat_ets <- fc_ets$.mean
select_cpi <- dplyr::select(data2023, cpi_calculat_ets, cpi)
select_cpi
combined_data <- dplyr::bind_rows(data_cpi, fc_ets)
combined_data$cpi <- as.numeric(mean(combined_data$cpi))
combined_data$.mean <- as.numeric(mean(combined_data$.mean))

```

```

ggplot() +
  geom_line(data = combined_data, aes(x = quarter_date, y = cpi), color = "black") +
  geom_line(data = combined_data, aes(x = quarter_date, y = .mean), color = "red") +
  geom_ribbon(data = combined_data, aes(x = quarter_date, ymin = lower80, ymax = upper80), fill =
"orange", alpha = 0.5) +
  geom_ribbon(data = combined_data, aes(x = quarter_date, ymin = lower95, ymax = upper95), fill =
"#F9C96A", alpha = 0.5) +
  labs(title = "Прогнозовані значення індексу споживчих цін за ETS моделлю", x = "Квартальна
дата", y = "Індекс споживчих цін") +
  theme_minimal()
ggplot(data2023, aes(x = quarter_date)) +
  geom_line(aes(y = cpi, color = "cpi"), size = 1.5) +
  geom_line(aes(y = cpi_calculat_ets, color = "cpi_calculat_ets"), size = 1.5) +
  labs(color = "Variable") +
  theme_minimal() +
  ggtitle("Графік змінних cpi та cpi_calculat_ets") +
  xlab("Квартали") +
  ylab("Значення")
data2023$percentage_relation_ets <- (data2023$cpi_calculat_ets / data2023$cpi) -1
print(data2023$percentage_relation_ets)
actual_values <- data2023$cpi
predicted_values <- data2023$cpi_calculated
library("AICcmodavg")
mae_lr <- mae(actual_values, predicted_values)
mse_lr <- mse(actual_values, predicted_values)
rmse_lr <- rmse(actual_values, predicted_values)
mape_value_lr <- mape(actual_values, predicted_values)
aic_value_lr <- AIC(line_reg)
aic_c_value_lr <- AICc(line_reg)
predicted_val_ar <- data2023$cpi_calculat_arima
mae_ar <- mae(actual_values, predicted_val_ar)
mse_ar <- mse(actual_values, predicted_val_ar)
rmse_ar <- rmse(actual_values, predicted_val_ar)
mape_value_ar <- mape(actual_values, predicted_val_ar)
aic_value_ar <- glance(fit) %>% dplyr::pull(AIC)
aic_c_value_ar <- glance(fit) %>% dplyr::pull(AICc)
predicted_val_ets <- data2023$cpi_calculat_ets
mae_ets <- mae(actual_values, predicted_val_ets)
mse_ets <- mse(actual_values, predicted_val_ets)
rmse_ets <- rmse(actual_values, predicted_val_ets)
mape_value_ets <- mape(actual_values, predicted_val_ets)
aic_value_ets <- glance(fit_ets) %>% dplyr::pull(AIC)
aic_c_value_ets <- glance(fit_ets) %>% dplyr::pull(AICc)
results_all <- tibble(
  model = c("Лінійна регресія", "ARIMA модель", "ETS модель"),
  MAE = c(mae_lr, mae_ar, mae_ets),
  MSE = c(mse_lr, mse_ar, mse_ets),
  RMSE = c(rmse_lr, rmse_ar, rmse_ets),
  MAPE = c(mape_value_lr, mape_value_ar, mape_value_ets),
  AIC = c(aic_value_lr, aic_value_ar, aic_value_ets),
  AICc = c(aic_c_value_lr, aic_c_value_ar, aic_c_value_ets))
print(results_all)

```

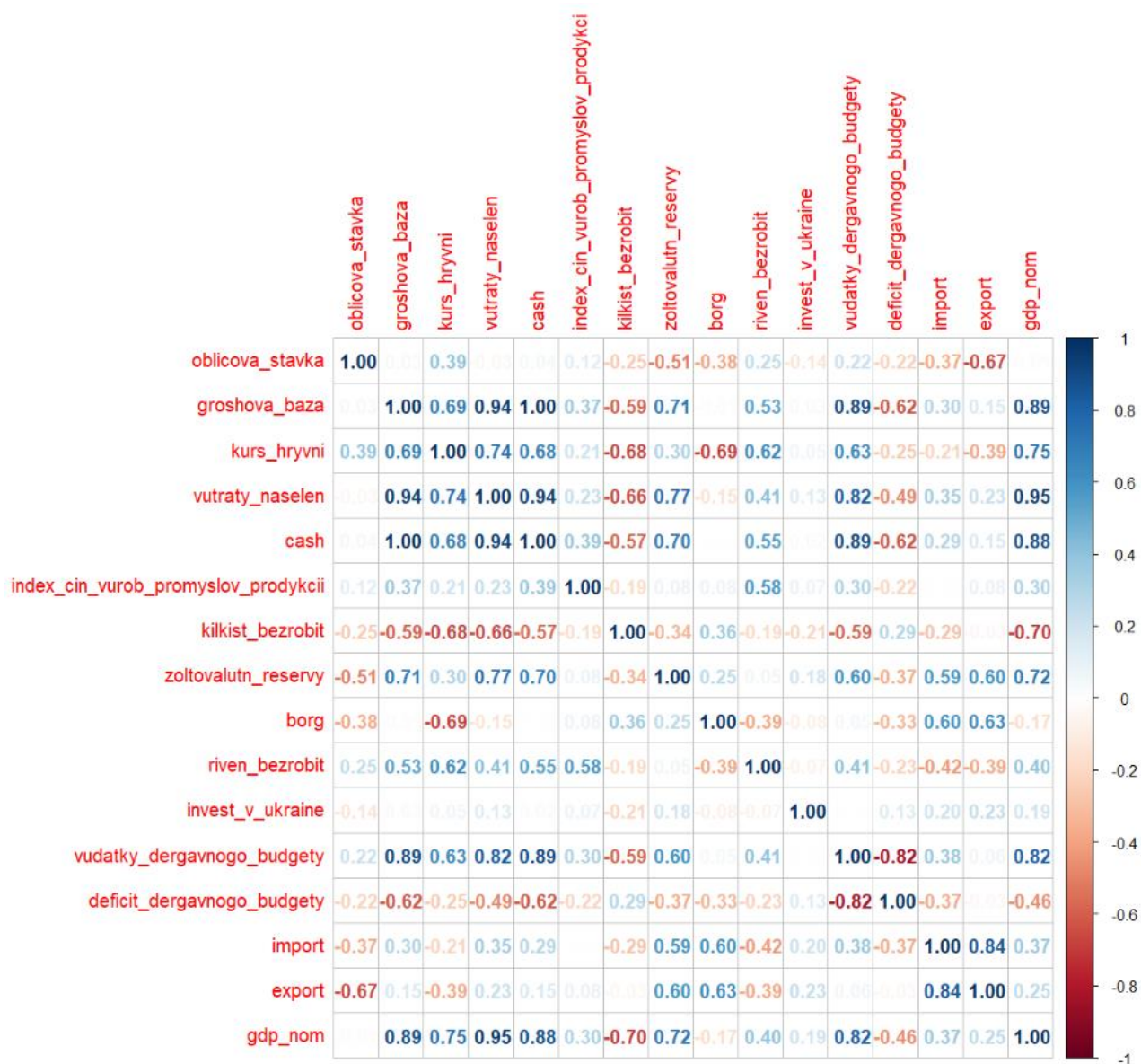


Рисунок Д.1 – Кореляційна матриця зібраних змінних



метадані

Заголовок

Музика_Мазур_плагіат

Автор

Науковий керівник / Експерт

Музика**Мазур**

підрозділ

кафедра математичного моделювання та статистики

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 |  | 13 |
| Інтервали |  | 0 |
| Мікропробіли |  | 13 |
| Білі знаки |  | 0 |
| Парафрази (SmartMarks) |  | 5 |

Обсяг знайдених подібностей

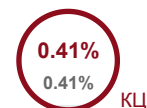
Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.



25
Довжина фрази для коефіцієнта подібності 2



11689
Кількість слів



82467
Кількість символів

Подібності за списком джерел

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 | https://ir.kneu.edu.ua/server/api/core/bitstreams/f1825f68-0bfb-4a6c-9ef4-894c7a3875b9/content | 41 | 0.35 % |
| 2 | https://library.if.ua/book/77/5550.html | 19 | 0.16 % |

| | | | |
|----|---|----|--------|
| 3 | https://ru.essays.club/%D0%AD%D0%BA%D0%BE%D0%BD%D0%BE%D0%BC%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B5-%D0%B4%D0%B8%D1%81%D1%86%D0%B8%D0%BF%D0%BB%D0%B8%D0%BD%D1%8B/%D0%91%D1%83%D1%85%D0%B3%D0%B0%D0%BB%D1%82%D0%B5%D1%80%D1%81%D0%BA%D0%B8%D0%B9-%D1%83%D1%87%D1%91%D1%82-%D0%B8-%D0%B0%D1%83%D0%B4%D0%B8%D1%82/%D0%A2%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D1%96%D1%97-%D0%BE%D0%B1%D0%BB%D1%96%D0%BA%D1%83-%D1%82%D0%B0-%D0%B0%D1%83%D0%B4%D0%B8%D1%82%D1%83-449617.html | 14 | 0.12 % |
| 4 | https://ir.kneu.edu.ua/server/api/core/bitstreams/01411c28-e954-480c-9741-533e5822b9ac/content | 12 | 0.10 % |
| 5 | https://ir.kneu.edu.ua/server/api/core/bitstreams/43414356-20b6-4998-a15f-f537153222d7/content | 12 | 0.10 % |
| 6 | https://ir.kneu.edu.ua/server/api/core/bitstreams/f1825f68-0bfb-4a6c-9ef4-894c7a3875b9/content | 11 | 0.09 % |
| 7 | https://knowledge.allbest.ru/marketing/3c0a65625a2bc68a4d53a89421206c36_0.html | 10 | 0.09 % |
| 8 | https://ir.kneu.edu.ua/server/api/core/bitstreams/b932a2e6-a370-4e13-ad35-e0090533cd70/content | 10 | 0.09 % |
| 9 | https://knowledge.allbest.ru/marketing/3c0a65625a2bc68a4d53a89421206c36_0.html | 8 | 0.07 % |
| 10 | http://www.repository.hneu.edu.ua/bitstream/123456789/11691/1/2015-8~1.PDF | 8 | 0.07 % |

з бази даних RefBooks (0.11 %)

| ПОРЯДКОВИЙ
НОМЕР | ЗАГОЛОВОК | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ
(ФРАГМЕНТІВ) | |
|---------------------|---|---|--------|
| джерело: Paperity | | | |
| 1 | Control of compressor station units using a fuzzy inverse model
О.О. Безсонов, О. Г. Лебедєв, О.Г. Руденко, О. С. Романюк; | 13 (2) | 0.11 % |

з домашньої бази даних (0.00 %)

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

з програми обміну базами даних (0.17 %)

| ПОРЯДКОВИЙ
НОМЕР | ЗАГОЛОВОК | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ
(ФРАГМЕНТІВ) | |
|---------------------|--|---|--------|
| 1 | SUMDU/out2013/Valiutno_kursova_polityka.pdf
7/22/2019
Summy State University (SUMDU) | 15 (3) | 0.13 % |
| 2 | Методи оцінювання та прогнозування в дослідженні процесів на фондовому ринку
5/14/2024
V. N. Karazin Kharkiv National University (KGNU) (Факультет математики і інформатики - кафедра теоретичної та прикладної інформатики) | 5 (1) | 0.04 % |

з Інтернету (1.45 %)

| ПОРЯДКОВИЙ
НОМЕР | ДЖЕРЕЛО URL | КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ
(ФРАГМЕНТІВ) | |
|---------------------|---|---|--------|
| 1 | https://ir.kneu.edu.ua/server/api/core/bitstreams/f1825f68-0bfb-4a6c-9ef4-894c7a3875b9/content | 52 (2) | 0.44 % |
| 2 | http://www.repository.hneu.edu.ua/bitstream/123456789/11691/1/2015-8~1.PDF | 20 (3) | 0.17 % |

| | | | |
|---|---|--------|--------|
| 3 | https://library.if.ua/book/77/5550.html | 19 (1) | 0.16 % |
| 4 | https://knowledge.allbest.ru/marketing/3c0a65625a2bc68a4d53a89421206c36_0.html | 18 (2) | 0.15 % |
| 5 | https://ru.essays.club/%D0%AD%D0%BA%D0%BE%D0%BD%D0%BE%D0%BC%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B8%D0%B5-%D0%B4%D0%B8%D1%81%D1%86%D0%B8%D0%BF%D0%BB%D0%B8%D0%BD%D1%8B/%D0%91%D1%83%D1%85%D0%B3%D0%B0%D0%BB%D1%82%D0%B5%D1%80%D1%81%D0%BA%D0%B8%D0%B9-%D1%83%D1%87%D1%91%D1%82-%D0%B8-%D0%B0%D1%83%D0%B4%D0%B8%D1%82/%D0%A2%D0%B5%D1%85%D0%BD%D0%BE%D0%BB%D0%BE%D0%B3%D1%96%D1%97-%D0%BE%D0%B1%D0%BB%D1%96%D0%BA%D1%83-%D1%82%D0%B0-%D0%B0%D1%83%D0%B4%D0%B8%D1%82%D1%83-449617.html | 14 (1) | 0.12 % |
| 6 | https://ir.kneu.edu.ua/server/api/core/bitstreams/01411c28-e954-480c-9741-533e5822b9ac/content | 12 (1) | 0.10 % |
| 7 | http://irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21COM=2&I21DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/nvmdue_2015_2(1)_41.pdf | 12 (2) | 0.10 % |
| 8 | https://ir.kneu.edu.ua/server/api/core/bitstreams/43414356-20b6-4998-a15f-f537153222d7/content | 12 (1) | 0.10 % |
| 9 | https://ir.kneu.edu.ua/server/api/core/bitstreams/b932a2e6-a370-4e13-ad35-e0090533cd70/content | 10 (1) | 0.09 % |

Список прийнятих фрагментів (немає прийнятих фрагментів)

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