

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 ринку транспортних послуг в Україні»

(назва теми)

здобувача Бабак Людмили Олександрівни
(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

Бабак Людмили Олександрівни

(прізвище, ім'я, по батькові)

 очної (денної) **форми навчання**

очної (денної), заочної, дистанційної

на підготовку кваліфікаційної бакалаврської роботи

на тему «Аналіз та моделювання ринку транспортних послуг 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/> та інших відкритих джерел.

План кваліфікаційної бакалаврської роботи

Розділ 1	Теоретичні аспекти та тенденції розвитку ринку транспортних послуг
Розділ 2	Аналіз структури та динаміки ринку транспортних послуг в Україні
Розділ 3	Моделювання показників ринку транспортних послуг

Об'єкт дослідження:	ринок транспортних послуг в Україні
Предмет дослідження:	математичні методи та моделі дослідження ринку транспортних послуг
Мета кваліфікаційної бакалаврської роботи:	аналіз ринку транспортних послуг з метою виявлення основних тенденцій та факторів впливу, а також моделювання показників ринку за допомогою регресійного аналізу

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

У розділі 1: 1) дослідити поняття та особливості ринку транспортних послуг; 2) охарактеризувати поточний стан розвитку логістичного ринку; 3) проаналізувати закономірності та тенденції розвитку ринку транспортних послуг.

У розділі 2: 1) проаналізувати структуру та структурні зрушення за видами транспорту; 2) оцінити збалансованість ринку транспортних послуг за регіонами країни; 3) провести кластерний аналіз показників ринку транспортних послуг в регіональному розрізі.

У розділі 3: 1) описати інформаційну базу для моделювання; 2) змоделювати показники ринку транспортних послуг за допомогою регресійного аналізу; 3) оцінити точність економетричної моделі; 4) визначити головні чинники впливу на ринок транспортних послуг; 5) здійснити порівняльний аналіз точності побудованих прогнозів показників ринку транспортних послуг.

Завдання підготував
науковий керівник

(підпис)

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

Завдання одержав
здобувач

(підпис)

Л.О. Бабак
(ініціали, прізвище)
«__» _____ 2024 р.

Реферат

Кваліфікаційна бакалаврська робота містить 60 сторінок, 9 таблиць, 55 рисунків, 13 формул та перелік використаних джерел з 35 найменувань, додатки.

«Аналіз та моделювання ринку транспортних послуг в Україні»

Об'єктом дослідження є ринок транспортних послуг в Україні.

Предметом дослідження є математичні методи та моделі дослідження ринку транспортних послуг.

Мета та завдання дослідження. Метою виконання кваліфікаційної бакалаврської роботи є аналіз ринку транспортних послуг з метою виявлення основних тенденцій та факторів впливу, а також моделювання показників ринку за допомогою регресійного аналізу.

Для досягнення поставленої мети визначено такі *завдання*:

- проаналізувати закономірності та тенденції розвитку ринку транспортних послуг;
- дослідити структуру та структурні зрушення за видами транспорту;
- оцінити збалансованість ринку транспортних послуг за регіонами країни;
- змодельовати показники ринку транспортних послуг;
- оцінити точність отриманих результатів моделювання та розрахованих прогнозів.

Теоретична, методична та практична значущість отриманих результатів. Під час дослідження було розглянуто теоретичні аспекти та сучасні методи аналізу ринку транспортних послуг. Досліджено структурні зрушення пасажиро- та вантажообігу у різних видах транспорту. Проведено регіональний аналіз ринку транспортних послуг, а також здійснено моделювання та прогнозування ключових показників його розвитку. Точність прогнозів перевірено на основі реальних даних, і визначено прогноз з найменшою похибкою.

Практичні результати дослідження полягають у можливості покращення процесу оптимізації ринку транспортних послуг на основі розроблених моделей. Це сприятиме підвищенню ефективності перевезень і конкурентоспроможності, покращенню управління логістичними процесами та забезпеченню збалансованості ринку за регіонами. Використання цих методів допоможе адаптуватися до макроекономічних змін та прогнозувати ключові показники ринку, зменшуючи вплив ризиків та покращуючи стратегічне планування.

Рік виконання кваліфікаційної бакалаврської роботи – 2024.

Рік захисту роботи – 2024.

Ключові слова: пасажирообіг, вантажообіг, ринок транспортних послуг, метод множинної регресії, економетрична модель, кластерний аналіз, прогнозування, ARIMA-модель.

В і д г у к
про кваліфікаційну бакалаврську роботу
здобувача освітньо-професійної програми «Економічна кібернетика»
Бабак Людмили Олександрівни
на тему: «Аналіз та моделювання ринку транспортних послуг в Україні»

1. Актуальність теми: У сучасних умовах глобалізації і швидкої зміни умов ринку аналіз та моделювання показників стратегічно важливої галузі транспортних послуг дозволяють оцінювати вплив факторів, прогнозувати економічні тенденції та розробляти стратегії для оптимізації використання транспортної інфраструктури, 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 та науковій роботі.
3. Наявність самостійних розробок автора: побудовані економетричні моделі показників пасажирообігу та вантажообігу, охарактеризовано вплив на них основних макроекономічних факторів, проведено кластерний аналіз регіонів країни за основними показниками ринку транспортних послуг, а також здійснено прогнозування ключових показників ринку.
4. Цінність теоретичних висновків та практичних рекомендацій полягає у можливості оптимізації ринку транспортних послуг за допомогою розроблених моделей, що сприятиме підвищенню ефективності перевезень, швидкої адаптації до макроекономічних змін, зниженню впливу ризиків та покращенню стратегічного планування в цій сфері.
5. Наявність недоліків: Доцільним в роботі є розширення системи показників в оцінюванні впливу на ключові показники ринку транспортних послуг. 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 Бабак Л.О. «Аналіз та моделювання ринку транспортних послуг в Україні» є самостійним завершеним дослідженням, відповідає вимогам, що висувуються 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 ринку транспортних послуг 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
Бабак Людмили

На сучасному етапі розвитку країни транспортна індустрія зазнає постійних змін через технологічний прогрес, зміни в економіці та зміни вимог споживачів. Під час воєнного стану оцінювання закономірностей та тенденцій розвитку ринку транспортних послуг України відіграє важливу роль, оскільки ефективний ринок транспортних послуг є однією з головних ознак високорозвиненої країни. Використання різних статистичних та економетричних методів аналізу ринку транспортних послуг стає актуальним при дослідженні розвитку транспортної сфери та визначає значимість і важливість бакалаврської роботи.

Дана кваліфікаційна робота присвячена узагальненню теоретичних основ вивчення ринку транспортних послуг, аналізу ринку в цілому та окремих його складових. Окрему увагу приділено дослідженню ринку в регіональному розрізі, побудові економетричних моделей та прогнозуванню. В роботі проведений теоретико-методологічний аналіз ринку транспортних послуг, розглянуто основні концепції та визначення, пов'язані з наданням транспортних послуг.

Особливе місце в роботі приділено побудові економетричних моделей впливу окремих факторів на показники пасажирообігу та вантажообігу, що дозволить покращити процес оптимізації ринку транспортних послуг на основі розроблених моделей. Заслугове на особливу увагу кластерний аналіз регіонів країни за основними показниками ринку транспортних послуг, що сприятиме покращенню управління логістичними процесами та забезпеченню збалансованості ринку за регіонами. Чільне місце в роботі займає прогнозування ключових показників ринку, в тому числі за допомогою ARIMA-моделей.

Всі розрахунки проведені в програмному середовищі R-studio, в якому автор адаптувала програмні коди із зовнішніх джерел для реалізації власних даних, що можна вважати підтвердженням самостійності виконання даної роботи. Інформаційною базою дослідження слугували статистичні дані Національного банку та Державної служби статистики України.

Висновки за результатами аналізу обґрунтовані і свідчать про високий рівень теоретичної підготовки і опанування практичних навичок аналізу та узагальнення статистичної інформації щодо складних закономірностей соціально-економічних явищ і процесів. Робота логічна та послідовна, результати дослідження відповідають поставленим меті та завданням.

До незначних недоліків можна віднести доцільність оцінювання розвитку ринку транспортних послуг не тільки в розрізі окремих регіонів країни, а й з урахуванням міжнародного досвіду та здійсненням міжнародних порівнянь.

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

Рецензент:

Канд.техн.наук, доцент
доцент кафедри системного аналізу
та інформаційних технологій
Маріупольського державного
університету

07 червня 2024 року



(підпис)

Г.В. МАРТИНЮК

Особистий підпис	<i>Г.В. Мартинюк</i>	доцента, зав. каф.
САТ МДУ,	<i>Мартинюк Г.В.</i>	засвідчую
факт бази		
Начальник загального відділу	<i>Анастасія БЕЗІМА</i>	

ЗМІСТ

ВСТУП	3
РОЗДІЛ 1. ТЕОРЕТИЧНІ АСПЕКТИ ТА ТЕНДЕНЦІЇ РОЗВИТКУ РИНКУ ТРАНСПОРТНИХ ПОСЛУГ	5
1.1 Сутність та структура ринку транспортних послуг.....	5
1.2 Особливості функціонування ринку транспортних послуг	7
1.3 Аналіз поточного стану ринку транспортних послуг.....	9
РОЗДІЛ 2. АНАЛІЗ СТРУКТУРИ ТА ДИНАМІКИ РИНКУ ТРАНСПОРТНИХ ПОСЛУГ В УКРАЇНІ	11
2.1 Структура та структурні зрушення за видами транспорту	11
2.3 Динаміка основних показників ринку транспортних послуг.....	14
2.3 Оцінювання збалансованості ринку транспортних послуг за регіонами країни	18
РОЗДІЛ 3. МОДЕЛЮВАННЯ ПОКАЗНИКІВ РИНКУ ТРАНСПОРТНИХ ПОСЛУГ	28
3.1 Інформаційна база дослідження	28
3.2 Регресійний аналіз для моделювання показників ринку транспортних послуг.....	29
3.3 Побудова моделей ARIMA показників ринку транспортних послуг	49
3.4 Прогнозування пасажирообігу та вантажообігу.....	57
ВИСНОВКИ	59
ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ	61
ДОДАТКИ	65

ВСТУП

Актуальність теми. Ринок транспортних послуг є не лише складовою частиною нашого повсякденного життя, але й важливим елементом соціально-економічного простору. Внутрішні та зовнішні переміщення пасажирів та вантажів є постійним явищем, яке визначає зручність та доступність для кожного. Зміни на ринку транспортних послуг несуть значні трансформації у соціальному та економічному житті. Наприклад, карантинні обмеження відкрили нові можливості для онлайн-роботи та навчання, що вплинуло на звички та спосіб життя багатьох людей.

Планування маршрутів, розвиток інфраструктури та розподіл ресурсів – усе це стає більш ефективним завдяки аналізу ринку транспортних послуг. Моделювання основних показників даного ринку та їх прогнозування є особливо актуальними через зростаючий попит, технологічні інновації, конкурентне середовище, екологічні вимоги та необхідність ефективного планування інфраструктури.

Аналіз останніх досліджень і публікацій. Дослідженням ринку транспортних послуг займалися такі науковці, як А. Брайковська, А. Матвєєва, І. Лукасевич-Крутник, В. Яворський, Ю. Стащенко, О. Гавриловський, А. Церковна, В. Макаренко та інші. Проте питання аналізу ключових показників ринку транспортних послуг, їх моделювання та прогнозування залишаються недостатньо дослідженими.

Метою виконання кваліфікаційної бакалаврської роботи є аналіз ринку транспортних послуг з метою виявлення основних тенденцій та факторів впливу, а також моделювання показників ринку за допомогою регресійного аналізу.

Для досягнення поставленої мети необхідно вирішити такі завдання:

- проаналізувати закономірності та тенденції розвитку ринку транспортних послуг;
- дослідити структуру та структурні зрушення за видами транспорту;

- оцінити збалансованість ринку транспортних послуг за регіонами країни;
- змоделювати показники ринку транспортних послуг;
- оцінити точність отриманих результатів моделювання та розрахованих прогнозів.

Об'єкт дослідження – ринок транспортних послуг в Україні.

Предмет дослідження – математичні методи та моделі дослідження ринку транспортних послуг.

Теоретичним та методологічним підґрунтям проведення дослідження стали положення сучасної економічної теорії та загальнонаукові й спеціальні методи пізнання. У процесі наукового дослідження було використано математичні моделі та методи машинного навчання, зокрема кореляційний, регресійний аналізи, методи прогнозування, в тому числі ARIMA моделі.

Інформаційна база для практичних досліджень створена на основі офіційної статистичної звітності Державної служби статистики та Національного банку України, а також матеріалах наукових публікацій та Інтернет-ресурсів, присвячених аналізу ринку транспортних послуг.

Робота складається зі вступу, трьох розділів, висновків та переліку використаних джерел. Обсяг кваліфікаційної роботи складає 60 сторінок. Робота містить 55 рисунків, 9 таблиць, 13 формул та 35 літературних джерел.

РОЗДІЛ 1

ТЕОРЕТИЧНІ АСПЕКТИ ТА ТЕНДЕНЦІЇ РОЗВИТКУ РИНКУ ТРАНСПОРТНИХ ПОСЛУГ

1.1 Сутність та структура ринку транспортних послуг

Транспортні перевезення є основою існування будь-якої держави. Саме завдяки розвитку транспортно-економічних зв'язків розширюються відносини між країнами, зростають обсяги виробництва, покращується рівень життя населення тощо. Сфера транспортних послуг розвивається та удосконалюється з кожним днем. Сучасні технології роблять перевезення більш екологічними, швидкими, безпечними, зручними, ніж вони були 20 років тому. Важливо розуміти сутність ринку транспортних послуг для покращення і вдосконалення його не лише в межах України, а й на міжнародному рівні.

Існує багато варіантів тлумачення поняття ринку транспортних послуг. Наприклад, А. Брайковська визначає ринок транспортних послуг як систему відносин між фізичними та юридичними особами – суб'єктами різних типів, форм власності та господарювання в світовому і внутрішньому економічному просторі щодо організації та купівлі-продажу транспортних послуг на основі дії об'єктивних економічних законів та правових норм [1, с. 51]. На думку А. Матвєєвої, ринок транспортних послуг – це складна система врегульованих нормами права та забезпечених засобами правового регулювання відносин, що складаються між виробниками та споживачами транспортних послуг, а також іншими учасниками ринку та державою в особі уповноважених органів державної влади з приводу перевезення пасажирів, багажу та вантажів як на національному, так і на міжнародному рівні [2, с. 142].

І. Лукасевич-Крутник в одному із своїх досліджень зазначає, що транспортна послуга – це дія або сукупність дій, спрямованих на здійснення

перевезення та (або) надання інших допоміжних для здійснення переміщення пасажирів, багажів, вантажів операцій [3, с. 116].

Узагальнюючи вищенаведені поняття ринок транспортних послуг можна визначити як систему економічних відносин між суб'єктами різних форм власності та господарювання, що надають та споживають послуги перевезення, регульовану економічними законами та правовими нормами.

Суб'єктами транспортного процесу виступають такі учасники як виробники, замовники, покупці, транспортні компанії та посередники, які вступають у складні економічні та комерційні відносини, що регулюються різними положеннями національного законодавства, міжнародного права та традицій [4, с. 127].

Транспортні послуги займають значну частку українського експорту послуг [5]. Тож дослідження ринку перевезень варто уваги. Аналізуючи структуру транспортних послуг дослідники виділяють такі види транспорту як автомобільний, залізничний, авіаційний, водний, міський електротранспорт [6]. Кожен з них займає значиме місце у композиції українських перевезень, оскільки всі вони є розвиненими в Україні [6].

Український залізничний транспорт є найважливішим сектором дорожньо-транспортного комплексу країни, на який припадає майже 82% вантажних перевезень і 36% пасажирських перевезень усіх видів транспорту разом узятих [7, 8].

Автомобільний транспорт є найпоширенішим видом транспорту в Україні. Але його основною проблемою є поганий стан доріг через кліматичні умови країни, недосконале дорожнє будівництво, пересування великогабаритних автомобілів на дорогах загального користування тощо. Для цього важливо суворо контролювати технології виробництва та ремонту дорожнього покриття, залучати іноземних інвесторів та забезпечувати державне фінансування в цілому [8; 9].

Авіаційний транспорт є менш поширеним порівнюючи з попередніми, але не менш важливим елементом транспортної системи України. Суб'єктами повітряного транспорту є всі види підприємств, організацій та установ, діяльність

яких спрямована на створення необхідних умов для використання повітряного простору особами, які використовують повітряні судна [8; 10].

Водний транспорт України поєднує в собі синтез підгалузей морського та річкового транспорту [8; 11]. Найважливіші порти – це Одеса, Ізмаїл, Іллічівськ, Херсон, Миколаїв, Севастополь, Ялта, Феодосія та Керч. Великі портові міста на узбережжі Азовського моря включають Маріуполь і Бердянськ [11]. Значна частка імпорту та експорту України здійснювалась саме через морські порти. Внутрішній водний транспорт є економічно ефективним способом транспортування товарів і матеріалів на великі відстані, знижуючи транспортні витрати та підвищуючи конкурентоспроможність української продукції [8; 12].

Міський електротранспорт є невід'ємною частиною інтегрованої транспортної системи, метою якої є перевезення населення трамваєм, тролейбусом і метро на маршрутах, що задовольняють життєво важливі потреби населення [13]. Трамвайний та тролейбусний транспорт доступний у багатьох містах країни, на відміну від метрополітену. Цей вид перевезень є відкритим лише у Києві, Харкові, Дніпрі та Кривому Розі [14].

Кожна складова транспортної системи України відіграє значну роль у загальній композиції ринку транспортних послуг. Дослідження структури дозволяє оцінити вплив соціально-економічних явищ на зміни динаміки кожного виду транспорту.

1.2 Особливості функціонування ринку транспортних послуг

Функціонування ринку транспортних послуг залежить від основних суб'єктів цього ринку, а саме транспортної системи та споживача транспортної продукції – вантажовідправник та вантажоодержувач. Внаслідок їх взаємодії має досягатись компроміс, баланс економічних інтересів, встановлення партнерських відносин тощо [15, с.17].

Ринок транспортних послуг невід’ємно пов’язаний з усіма галузями життя. Тож важливо аби послуги були якісними. Це зменшить ризики функціонування сфер соціально-економічного життя. Якість транспортних послуг впливає на конкурентоспроможність бізнесу та більшості галузей національної економіки [16].

Ефективне функціонування транспортного ринку залежить від якості надання послуг. Рівень обслуговування означає, що надана послуга відповідає очікуванням споживача та певним заздалегідь визначеним стандартам [17, с. 126]. Якість надання послуг залежить від організації транспортного процесу, експлуатаційних характеристик транспорту, стану дорожньої мережі, містобудування та інших факторів і безпосередньо впливає на час у дорозі, комфорт і зручність поїздки та рівень тарифів [18, с. 49].

Особливістю діяльності ринку транспортних послуг в Україні є взаємодія різних видів транспорту для вирішення питань перевезень пасажирів та вантажу якомога швидшим та якіснішим способом. Для налагодження ефективної взаємодії учасників ринку варто покращувати якість транспортної мережі доріг різного виду призначення, покращувати рівень обслуговування, створювати комфортні умови для поїздки тощо. Успішне функціонування транспортного ринку вимагає рівних умов конкуренції для всіх видів транспорту [19, с. 39].

Головними проблемами на ринку транспортних послуг є недосконалість системи регулювання та управління розвитком дорожньо-транспортного комплексу, брак інформації для прийняття управлінських рішень, недостатнє фінансування розвитку дорожньої мережі та транспортної інфраструктури, а також брак знань та застосування технологій [20, с. 29-30]. Подолання цих викликів є провідним вектором розвитку діяльності транспортного ринку України.

Основними показниками функціонування ринку транспортних послуг є пасажирообіг та вантажообіг. Пасажирообіг – це кількість пасажирів, перевезених на певну відстань за певний період. Обчислюється як добуток кількості пасажирів на відстань перевезення і вимірюється в пасажиро-кілометрах [21]. Вантажообіг визначає кількість вантажів, перевезених за певний період і на певну відстань.

Розраховується як добуток обсягу перевезених вантажів на відстань перевезення, одиницею виміру якого є тонно-кілометр [21; 22].

Аналізуючи та моделюючи ці показники, можна оцінити економічну ситуацію в країні чи регіоні, попит на транспортні послуги, якість транспортних послуг, оптимізацію маршрутів, розподіл транспортних засобів та розкладів руху тощо.

1.3 Аналіз поточного стану ринку транспортних послуг

Соціально-економічні явища привносять свої корективи у функціонування ринку. Так у 2020 році набула поширення коронавірусна хвороба (COVID-19) в Україні. Це мало суттєвий вплив на перевезення не тільки пасажирів, а й вантажів. Наступною кризовою точкою стало введення воєнного стану по всій підконтрольній території України. Кожне з цих явищ додало незворотних змін у соціально-економічне життя суспільства.

За словами міністра економіки Юлії Свириденко, відкрито альтернативний морський коридор для експорту товарів та послуг з України, запущено механізм страхування суден від воєнних ризиків, а також налагоджено річкову логістику на Дунаї. Очікується, що ці та інші позитивні зрушення сприятимуть зростанню експорту товарів та послуг з України на 9% у 2024 році. За прогнозами Міністерства економіки експорт зростатиме на 19,4% у 2025 році і на 20,6% у 2026 році [23; 24]. Важливо зазначити, що у січні 2022 року 80% експорту все ще здійснювалося морем, але до квітня 2022 року цей показник знизився до 13%. Більшість поставок наразі здійснюється залізничним та автомобільним транспортом, а також збільшилася кількість поставок до країн ЄС [25].

Україна зазнала значних економічних збитків – приблизно 486 мільярдів доларів. Основні потреби України до 2024 року включають 2,3 мільярда доларів США на реконструкцію транспорту та логістики [26]. Відновлення транспортної

інфраструктури має важливе значення для стабілізації економіки та забезпечення безперебійного функціонування транспортного ринку.

Основним постачальником залізничних транспортних послуг є АТ «Укрзалізниця». Вона працює над покращенням сервісу для пасажирів, розвиває власний додаток, програму лояльності [27], а також працює над оновленням парку локомотивів. У рамках проєкту Світового банку «Відновлення критично важливої логістичної інфраструктури та мережевого сполучення (RELINC)» та за допомогою фінансування Європейського банку реконструкції та розвитку (ЄБРР) планується придбати до 80 сучасних потужних електровозів [28]. Це сприяє зростанню перевезення більшої кількості пасажирів, заохочуючи їх користуватись більш комфортними та якісними послугами.

Автомобільні перевезення забезпечуються численною кількістю компаній, що забезпечують вантажні та пасажирські перевезення, як внутрішні, так і міжнародні.

Викликами для міського електротранспорту є ризики руйнування інфраструктури, що може змінити варіанти для переміщень по місту.

Загалом, транспортний сектор характеризується автоматизацією всіх процесів шляхом модернізації існуючих базових технологій. Ринок транспортних послуг активно розвивається, став невід'ємною частиною суспільної інфраструктури і, очевидно, розвиватиметься й надалі, незважаючи на різні перешкоди [5].

РОЗДІЛ 2

АНАЛІЗ СТРУКТУРИ ТА ДИНАМІКИ РИНКУ ТРАНСПОРТНИХ ПОСЛУГ В УКРАЇНІ

2.1 Структура та структурні зрушення за видами транспорту

Структурні зрушення за видами транспорту дозволяють оцінити популярність та визначити вплив соціально-економічних змін. Структуру пасажирообігу в Україні за видами транспорту за 2014-2022 роки представлено на рис. 2.1.

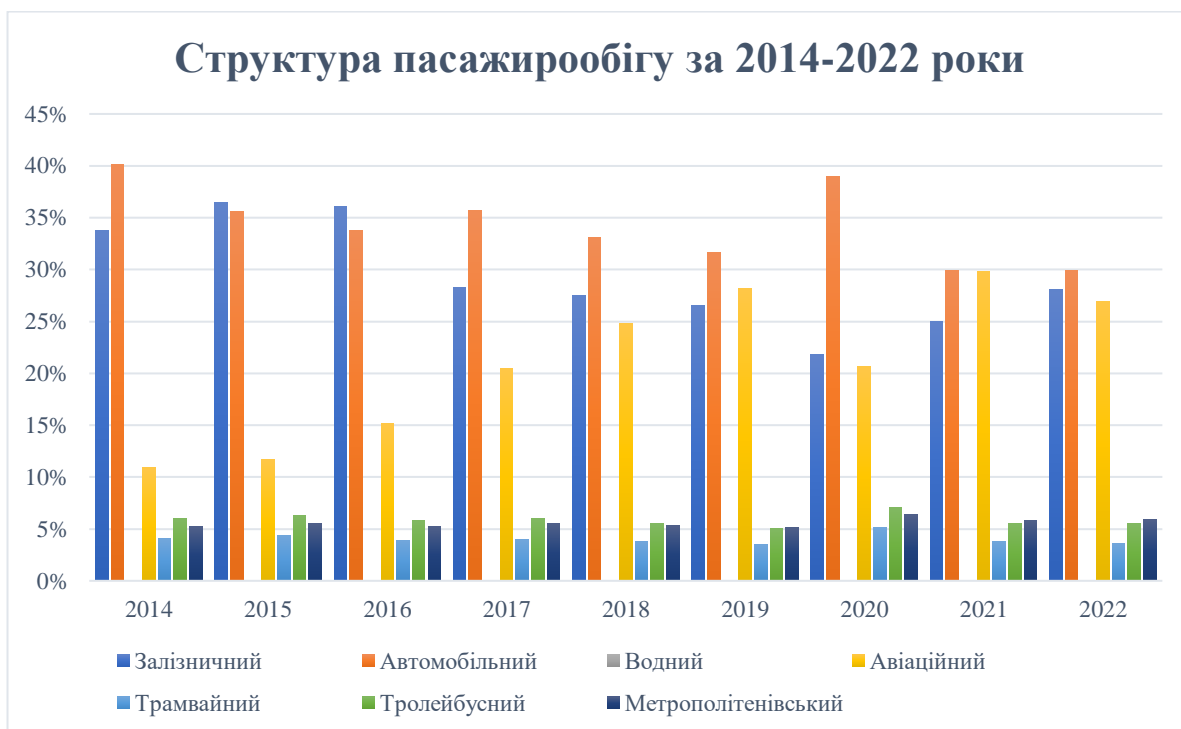


Рисунок 2.1 – Структура пасажирообігу в Україні за 2014-2022 роки

Джерело: розроблено автором на основі [33-34]

Структура пасажирообігу за різними видами транспорту постійно змінювалась. Щодо залізничного транспорту, то за період з 2015 року до 2020 частка пасажирообігу зменшилася на 14 в.п., з 36% до 22% відповідно. Частка авіаційного транспорту досягла максимуму у 2021 році, збільшивши свою присутність на ринку транспортних послуг на 19% порівняно з 2014 роком.

Важливо зазначити, що у 2020 році спостерігається різке зменшення кількості перевезених пасажирів залізничним та авіаційним транспортом, і збільшення автомобільних перевезень. Частка автомобільних перевезень пасажирів у 2020 році була наближена до рівня 2014 року, 39 і 40% відповідно. Це спричинено початком пандемії COVID-19 та відповідними обмеженнями у пересуванні осіб різними видами транспорту, заборона перетину кордону, що видно по зменшенню частки залізничних та авіаційних перевезень.

Оцінювання структури пасажироперевезень дозволяє більш детально проаналізувати зміни в соціально-економічному житті. На рис. 2.2 представлено структурні компоненти пасажирообігу в Україні за 2014 рік.

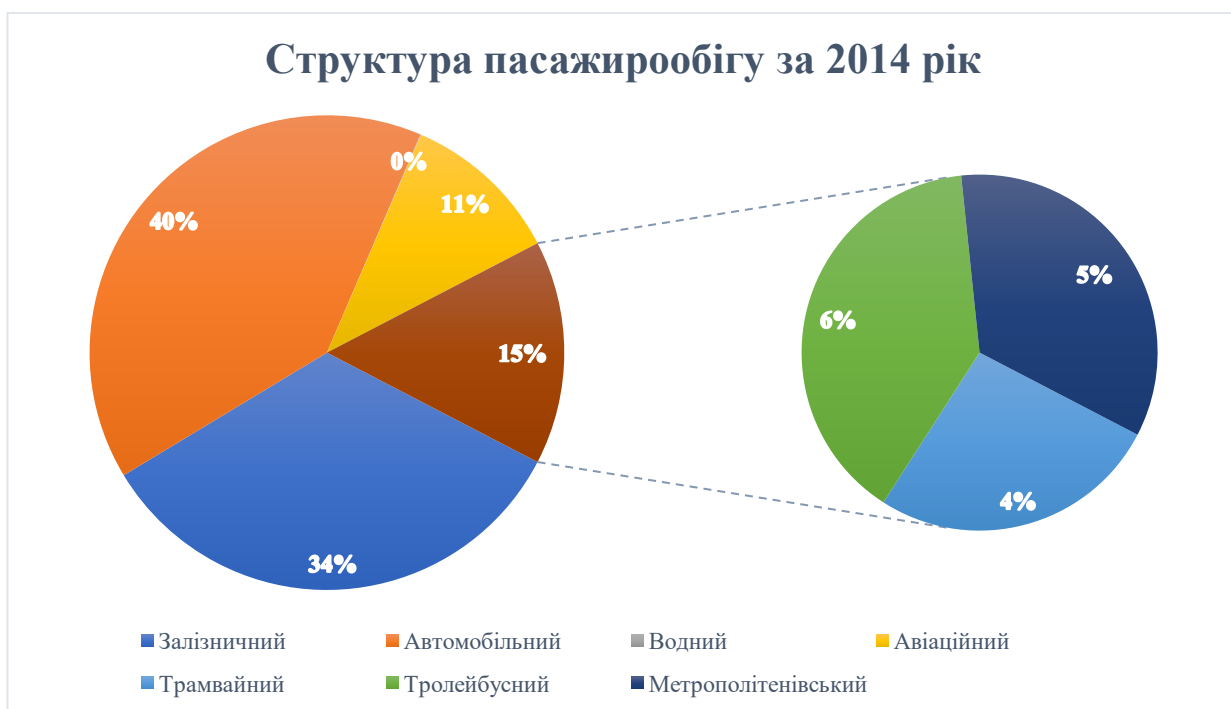


Рисунок 2.2 – Структура пасажироперевезень в Україні за 2014 рік

Джерело: розроблено автором на основі [33-34]

Значну частку займає автомобільний транспорт на який припадає 40% від усього пасажирообігу 2014 року. Залізничні перевезення є менш популярними і займають 34%. Лише 11% займає авіаційний транспорт. З 15% які припадають на міський електротранспорт найбільше для перевезень люди користувались тролейбусами (6%), трохи менше – послугами метрополітену (5%) і найменше – трамвайними (4%).

Вплив пандемії COVID-19 поширився й на транспортні вподобання пасажирів. Структура зазнала значних змін у 2020 році. Але вже у 2021 році та січні 2022 року структурні частки почали виходити на докризовий рівень. Основні складові кількості перевезених пасажирів зображено на рис. 2.3.



Рисунок 2.3 – Структура пасажирообігу в Україні за січень 2022 року

Джерело: розроблено автором на основі [33-34]

Значні структурні зрушення відбулись за період 2014-2022 років. Майже у рівних частках представлено залізничний, автомобільний та авіаційний транспорт, 28%, 30% і 27% відповідно. Автомобільна складова зменшилась на 10%, залізнична – на 6%. Саме на ці 16% зросла присутність авіаційного перевезень на транспортному ринку. Помінялись своїми позиціями тролейбусний та метрополітенівський транспорт, з 6% на 5% та з 5% на 6% відповідно. Трамвайні перевезення стабільно тримають постійне місце у структурі, незначно змінивши свої позиції лише у 2019 (3%) та 2020 роках (5%).

2.3 Динаміка основних показників ринку транспортних послуг

Аналізуючи ринок транспортних послуг важливо дослідити динаміку основних показників. Це дасть змогу визначити, які зрушення відбувались і з'ясувати, що ж стало причиною для змін. Динаміку обсягів пасажирообігу та вантажообігу представлено на рис. 2.4.



Рисунок 2.4 – Динаміка пасажирообігу та вантажообігу за 2014-2023 рр.

Джерело: розроблено автором на основі [33-34]

Для показника пасажирообігу характерною є сезонність з 2014 по 2020 роки. Найбільша кількість пасажироперевезень тоді припадала на 3 квартал. Але з настанням пандемії COVID-19 сезонність зникла. Тоді з'явилися і нововведення, без яких важко тепер уявити сучасне життя. За рахунок карантинних обмежень, в тому числі і переміщень як по країні, так і за її межі, різко зменшився потік бажаючих подорожувати. Але вже до 3 кварталу 2021 року пасажирообіг поступово зростав. Наступним критичним моментом став 2 квартал 2022 року. Саме тоді розпочалось повномасштабне вторгнення. Поступово потік пасажирів

збільшується, але навряд найближчим часом вийде на рівень допандемічного періоду, або хоча б до повномасштабного.

Для вантажообігу сезонність відсутня. Пандемія COVID-19 внесла свої корективи в динаміку вантажообігу, обсяг якого значно зменшився. Ще більший вплив має повномасштабне вторгнення, оскільки у 2 кварталі 2022 року обсяг вантажоперевезень зменшився приблизно вдвічі порівняно з попереднім періодом. Тож обидва фактори можуть впливати на моделювання показників, оскільки їх значення є незворотнім.

Для більш глибокого дослідження варто проаналізувати динаміку пасажирообігу в розрізі видів транспорту за 2014-2022 роки (рис. 2.5).

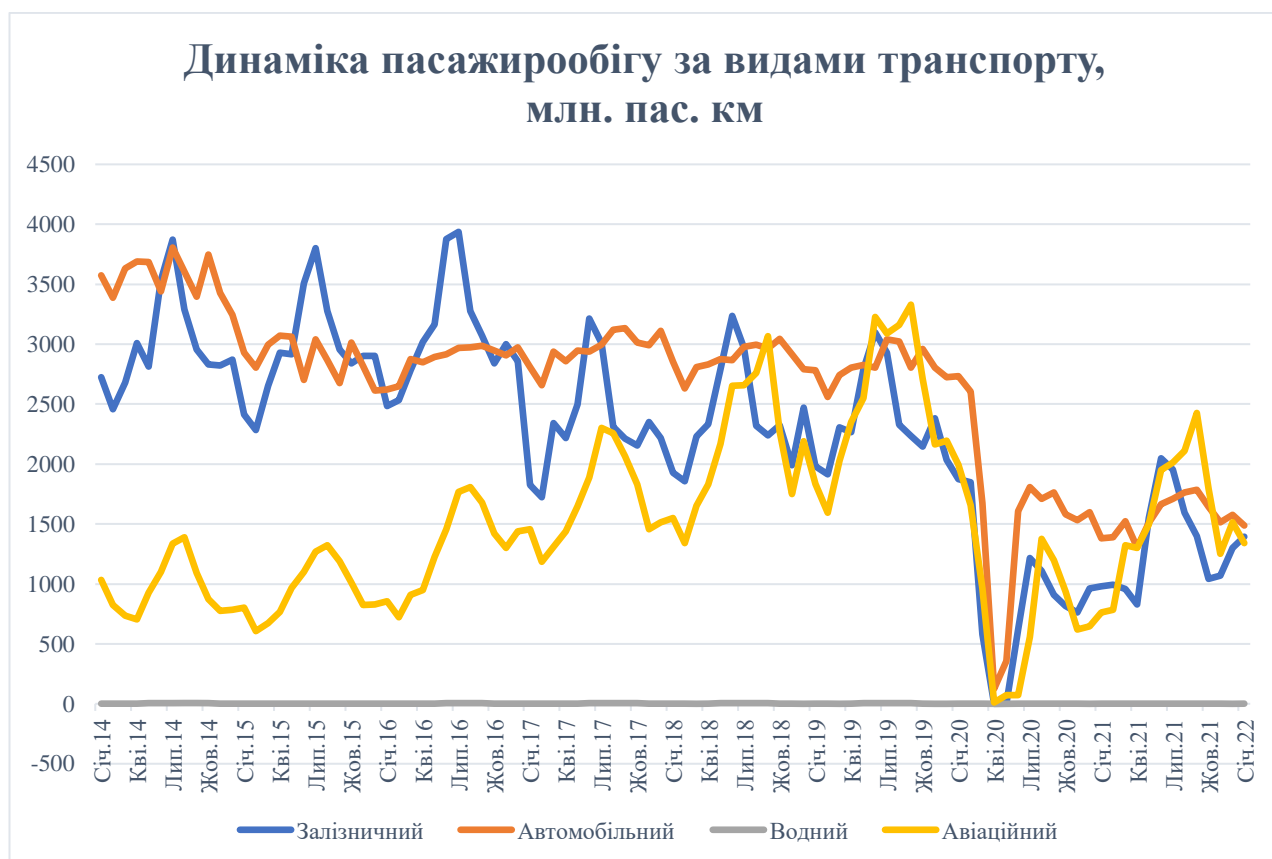


Рисунок 2.5 – Динаміка кількості перевезених пасажирів за видами транспорту в Україні за 2014-2022 роки

Джерело: розроблено автором на основі [33-34]

Залізничний транспорт майже відтворює загальну динаміку пасажирообігу. Варто зазначити, що за видами транспорту доступна статистика до січня 2022 року. Далі статистична інформація не оновлюється у зв'язку з воєнним станом. Звісно

спостерігається, що у квітні 2020 року відбулось значне скорочення обсягів пасажироперевезень з оголошенням пандемії COVID-19 та введенням карантинних обмежень. Залізничний транспорт почав втрачати свою популярність з січня 2017 року, його обсяг зменшився на 50%. Авіаційний транспорт нарощував свою популярність до жовтня 2019 року. Після послаблення обмежень він досить швидко відновив темпи зростання кількості перевезених пасажирів. Але з настанням повномасштабного вторгнення цей вид транспорту не використовується. Автомобільний транспорт теж зазнав незначних змін, але тримається майже на одному рівні. Щодо міського електротранспорту, то спостерігається незначна сезонність. Більш деталізована динаміка обсягу пасажироперевезень міським електротранспортом представлена на рис. 2.6.



Рисунок 2.6 – Динаміка пасажироперевезень за видами міського електротранспорту України за 2014-2022 роки

Джерело: розроблено автором на основі [33-34]

Динаміка пасажирообігу за видами міського електротранспорту відтворює поведінку людей, зважаючи переважно на вихідні та святкові дні. Наприклад, зменшується обсяг пасажироперевезень у січні, оскільки більше святкових днів.

Також цікаво проаналізувати динаміку пасажирообороту водного транспорту (рис. 2. 7), оскільки його обсяги значно менші за інші види транспорту.



Рисунок 2.7 – Динаміка пасажирообігу водного транспорту за 2014-2022 роки

Джерело: розроблено автором на основі [33-34]

Характерною ознакою водного транспорту є сезонність, оскільки в холодну пору року спостерігаються найменші обсяги перевезень, а влітку – найбільші. Також за результатами аналізу після введення карантинних обмежень спостерігається значне зменшення пасажирообігу в Україні за 2020-2022 роки. Наприклад, порівнюючи обсяг перевезених пасажирів цим видом транспорту у вересні 2021 року з аналогічним періодом 2019 року, спостерігається його зменшення на 62%. Це обумовлює необхідність оцінювання збалансованості ринку транспортних послуг за регіонами країни.

2.3 Оцінювання збалансованості ринку транспортних послуг за регіонами країни

В Україні ринок транспортних послуг має суттєві відмінності за регіонами. Тому наступним кроком проведено аналіз структури кількості перевезених пасажирів залізничним та автомобільним транспортом в регіональному розрізі. Кількість перевезених пасажирів залізничним транспортом у 2016 році представлено на рис. 2.8.



Рисунок 2.8 – Кількість перевезених пасажирів залізничним транспортом за регіонами України у 2016 році

Джерело: розроблено автором на основі [33-34]

Лідерами за обсягом пасажироперевезень у 2016 році були Київська, Харківська, Дніпропетровська та Львівська області. Для цих областей характерна розвинена транспортна інфраструктура та висока щільність населення. Найменше перевезено пасажирів в Миколаївській, Херсонській та Чернівецькій областях. Це

спричинено меншою кількістю населення в регіонах та частотою перевезень, а також економічними факторами.

Для порівняння та аналізу змін в динаміці на рис. 2.9 наведено кількість перевезених пасажирів за регіонами країни в 2020 році.

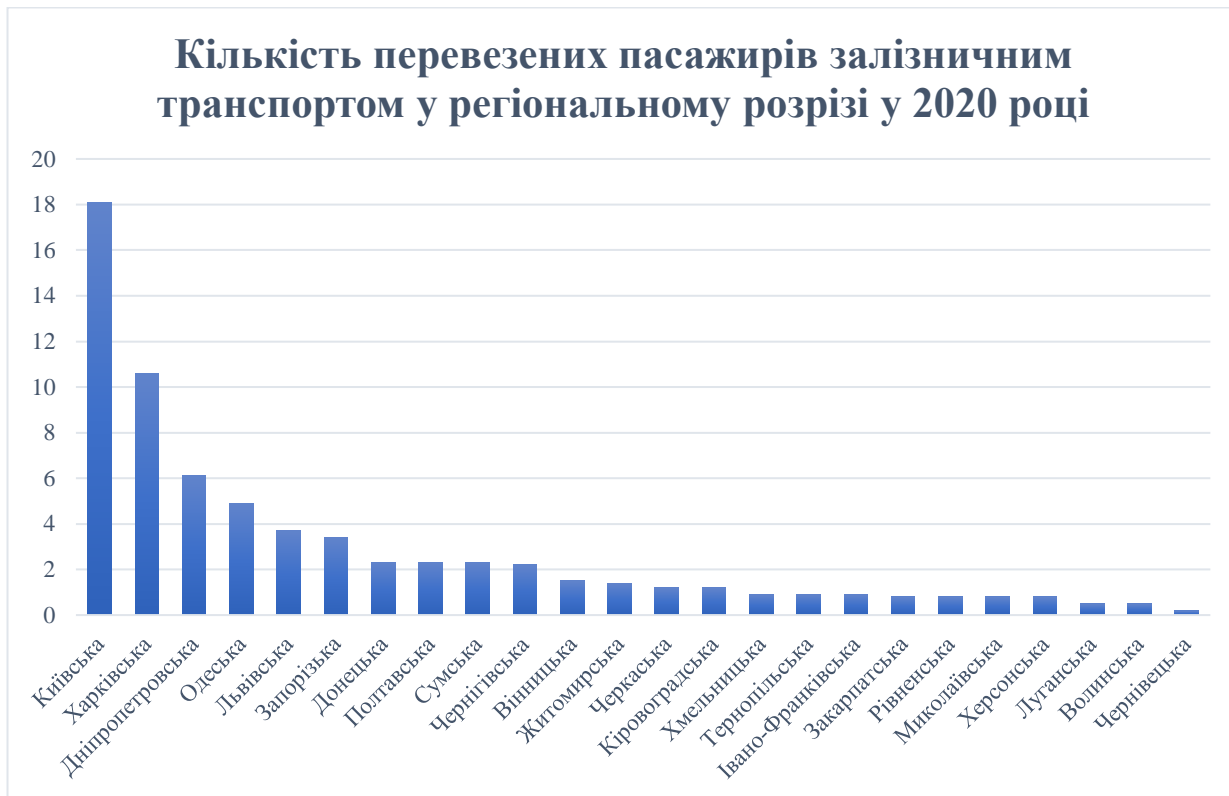


Рисунок 2.9 – Кількість перевезених пасажирів залізничним транспортом за регіонами України у 2020 році

Джерело: розроблено автором на основі [33-34]

У 2020 році лідером була Київська, Харківська і Дніпропетровська області. А Львівська знизила своє розташування в рейтингу. Також є зміни для аутсайдерів – Херсонська, Луганська, Волинська області, а також Чернівецька, яка не змінила свої позиції. Харківська область, яка була однією з провідних у галузі перевезень залізничним транспортом пасажирів, у 2020 році змінила свою позицію на друге місце в рейтингу. Варто зауважити, що пасажирообіг цієї області за 4 роки зменшився в 4 рази. Це спричинено настанням пандемії COVID-19, що потягнуло за собою зменшення пасажиропотоку у зв'язку з карантинними обмеженнями.

Автомобільний транспорт за аналізований період завжди мав перевагу над іншими видами. Він є більш комфортним, швидким та мобільним, що й виводить

його на лідерські позиції. На рис. 2.10 представлено кількість перевезених пасажирів у 2016 році за регіонами України.

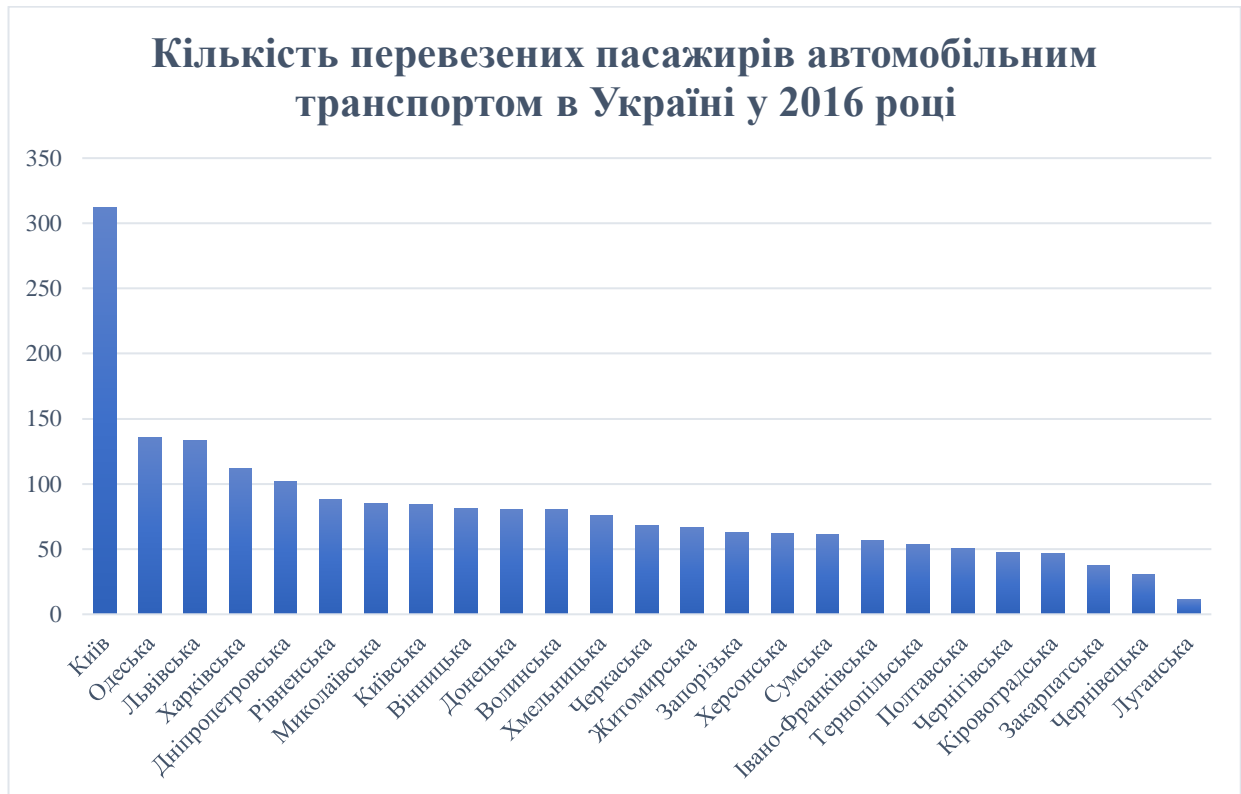


Рисунок 2.10 – Кількість перевезених пасажирів автомобільним транспортом за регіонами України у 2016 році

Джерело: розроблено автором на основі [33-34]

У 2016 році лідерські позиції займали м. Київ, а також Одеська, Львівська, Харківська області. Вони є туристичними, культурними, науковими та освітніми центрами України. Найменша кількість перевезених пасажирів була в Закарпатській, Чернівецькій та Луганських областях. Основними причинами є географічна віддаленість, невелика площа і населення, небезпека, що спричинила відтік жителів області відповідно.

За 5 років відбулось багато змін, а отже змінився рейтинг областей за обсягом пасажироперевезень. Кількість перевезених пасажирів автомобільним транспортом у 2021 році в Україні зображено на рис. 2.11.



Рисунок 2.11 – Рейтинг кількості перевезених пасажирів автомобільним транспортом за регіонами України у 2021 році

Джерело: розроблено автором на основі [33-34]

Абсолютним лідером в рейтингу у 2021 році є місто Київ, який є економічним центром та має високу густоту населення. Далі лідирують Донецька, Львівська та Дніпропетровська, які є промисловими регіонами, туристичними центрами та мають високий рівень економічної активності в регіонах. Рейтинг закривають Закарпатська, Чернівецька, Луганська області, які за 5 років не змінили свої позиції.

Варто проаналізувати зміну позицій регіонів у 2016 та 2021 роках за показником вантажообігу автомобільного транспорту. На рис. 2.12 представлено вантажоперевезення в Україні у 2016 році.



Рисунок 2.12 – Вантажообіг автомобільного транспорту за регіонами України у 2016 році

Джерело: розроблено автором на основі [33-34]

У 2016 році лідерські позиції займають місто Київ, Закарпатська, Львівська, Дніпропетровська та Київська області. Аутсайдерами є Житомирська, Луганська та Сумська області. Такий рейтинг спричинений рівнем економічної активності в регіонах, а також потребою у забезпеченні споживачів різними видами продукції зважаючи на щільність населення.

Для порівняльного аналізу на рис. 2.13 подано рейтинг вантажообігу за автомобільним транспортом у регіональному розрізі в 2021 році в Україні.



Рисунок 2.13 – Рейтинг вантажообігу за автомобільним транспортом за регіонами України у 2021 році

Джерело: розроблено автором на основі [33-34]

У 2021 році лідерами за обсягом вантажообігу є місто Київ, Львівська, Дніпропетровська, Закарпатська та Київська області. Для них характерним є високий попит населення на товари та послуги, значна частка промисловості та виробництва, значний обсяг міжнародних перевезень, а також розвинена транспортна інфраструктура. Порівнюючи з попереднім періодом Закарпатська область змінила свої позиції і перейшла з другого на четверте місце. Аутсайдерами є Чернівецька, Житомирська, Луганська області. Причиною таких низьких позицій є менша економічна активність, кількість промислових підприємств, віддаленість від транспортних коридорів та ринків збуту тощо.

Зміна лідерських позицій областей зумовлена різними факторами. Зосереджено увагу на областях, які в період пандемії змінили своє положення. Для кластерного аналізу було обрано дані вантажообігу по областях та по місту Києву за період 2016-2021 років. На рис. А.1 та А.2 (додаток А) представлено ієрархічні дендрограми за показником вантажообігу за період 2016-2019 та 2020-2021 років

відповідно. Побудовані дендрограми спричинили вибір кількості кластерів. Найбільш доцільним варіантом є розподіл на 4 кластери за ці періоди.

Графічний розподіл за методом к-середніх на кластери подано на рис. А.3 та А.4 (додаток А). Варто звернути увагу, що точки груп не перетинаються, а отже кількість кластерів було обрано вдало і така кластеризація є якісною.

В результаті кластерного аналізу за показником вантажообігу за 2016-2019 роки було утворено 4 групи регіонів, їх структуру подано у табл. 2.1.

Таблиця 2.1 – Структура кластерів вантажообігу в Україні за 2016-2019 роки

Кластер	Регіон
1	Вінницька, Івано-Франківська, Миколаївська, Запорізька, Тернопільська, Херсонська, Кіровоградська, Чернівецька, Чернігівська, Житомирська, Луганська, Сумська
2	Закарпатська, Львівська, Дніпропетровська, Київська, Харківська
3	Полтавська, Одеська, Хмельницька, Донецька, Волинська, Черкаська, Рівненська
4	м. Київ

Джерело: розроблено автором на основі [33-34]

Оцінюючи утворені кластери за період 2016-2019 років необхідно охарактеризувати їх значимість з економічної точки зору. У кластері 1 зосереджено області, що є важливими транспортними вузлами, а також регіони з розвиненими агропромисловими та промисловими комплексами. Другий кластер зібрав у собі області, що є великими центрами важкої та легкої промисловості, вони є важливими для логістичних потоків та мають розвинуту транспортну інфраструктуру. До кластеру 3 увійшли області, що можуть виступати як джерело сировини для інших регіонів. Окремим четвертим кластером є місто Київ, оскільки він має статус столиці та велике значення як адміністративного, економічного та культурного центру.

У табл. 2.2 показано розподіл областей України за показником вантажообігу за 2020-2021 роки.

Таблиця 2.2 – Структура кластерів вантажообігу в Україні за 2020-2021 роки

Кластер	Регіон
1	Хмельницька, Донецька, Івано-Франківська, Миколаївська, Запорізька, Тернопільська, Херсонська, Кіровоградська, Чернівецька, Чернігівська, Житомирська, Луганська, Сумська
2	Закарпатська, Львівська, Дніпропетровська, Київська
3	Харківська, Полтавська, Одеська, Волинська, Черкаська, Вінницька, Рівненська
4	м. Київ

Джерело: розроблено автором на основі [33-34]

Період 2020-2021 років характеризується значними змінами у соціально-економічному просторі держави та світу. Тож відбулись певні зрушення і в утворенні кластерів. Характеристика груп суттєво не змінилась, але є зміни у положеннях певних областей у секторах. Таким чином, Харківська та Вінницька області втратили свої позиції на ринку перевезень та зменшили обсяги вантажообігу, що і спричинило їх перехід у 3 кластер. Хмельницька ж область навпаки покращила свої результати роботи та перейшла до першої групи.

Також було проведено дослідження на основі даних лише по областях і не враховуючи місто Київ за аналогічні періоди. У табл. А.1 та А.2 (додаток А) подано структуру розподілу областей за показником вантажообігу за 2016-2019 та 2020-2021 роки відповідно. Тож при аналізі вантажообігу було виявлено, що місто Київ виступає окремим кластером, при відсутності його в аналізі структура кластерів не змінилась. Тож обидва варіанти є прийнятними для аналізу.

Для аналізу пасажирообігу були проведені аналогічні дослідження на основі даних кількості перевезених пасажирів за період 2016-2019 та 2020-2021 років. Тож для більш детального аналізу проаналізовано результати, що враховують всі області та місто Київ. Побудована дендрограма подана на рис. А.5 (додаток А) допомогла визначити, що варто поділити області на 5 кластерів. Період 2020-2021 років виніс свої корективи, але проведений ієрархічний аналіз, поданий на рис. А.6 (додаток А) допоміг визначити, що варто здійснити розподіл областей на 5

кластерів також. Позитивним результатом є те, що такий розподіл не спричинив перетин областей різних кластерів за методом к-середніх за період 2016-2019 та 2020-2021 років, що подано на рис. А.7 та А.8 (додаток А) відповідно.

Перелік областей розподілених на групи внаслідок проведеного кластерного аналізу за період 2016-2019 років подано у табл. 2.3.

Таблиця 2.3 – Структура кластерів кількості перевезених пасажирів в Україні за 2016-2019 роки

Кластер	Регіон
1	м. Київ
2	Дніпропетровська, Львівська, Одеська, Харківська
3	Вінницька, Донецька, Київська, Миколаївська, Рівненська
4	Волинська, Житомирська, Запорізька, Івано-Франківська, Кіровоградська, Полтавська, Сумська, Тернопільська, Херсонська, Хмельницька, Черкаська, Чернігівська
5	Закарпатська, Луганська, Чернівецька

Джерело: розроблено автором на основі [33-34]

Високий рівень пасажиропотоку характерний для першого кластеру, до якого увійшло лише місто Київ. Для регіонів другого кластеру притаманний високий рівень розвиненості транспортної інфраструктури. Ці області мають стратегічне географічне положення та є важливими транспортними вузлами України. Для третього кластеру властиві розвинені транспортні мережі та значний обсяг пасажиропотоку та вантажообігу. Області з четвертого сектору мають значно менший обсяг транспортного потоку, ніж попередні кластери, але мають велике значення для забезпечення регіональної та міжрегіональної доступності. До п'ятого кластеру увійшли регіони, які мають у порівнянні найменш розвинену транспортну інфраструктуру, але є значимими у забезпеченні комунікацій у віддалених та гірських районах.

У табл. 2.4 представлено кластеризацію кількості перевезених пасажирів за 2020-2021 роки в Україні.

Таблиця 2.4 – Структура кластерів кількості перевезених пасажирів в Україні за 2020-2021 роки

Кластер	Регіон
1	м. Київ
2	Донецька, Львівська
3	Вінницька, Дніпропетровська, Запорізька, Івано-Франківська, Київська, Миколаївська, Одеська, Рівненська, Харківська
4	Волинська, Житомирська, Кіровоградська, Полтавська, Сумська, Тернопільська, Херсонська, Хмельницька, Черкаська, Чернігівська
5	Закарпатська, Луганська, Чернівецька

Джерело: розроблено автором на основі [33-34]

Аналізуючи дані за 2020-2021 роки важливо звернути увагу на наповнення кластерів і зміни, що відбулись за період введення карантинних обмежень. Місто Київ продовжує залишатись лідером пасажироперевезень, що пов'язано з високою щільністю населення в регіоні та можливостей для реалізації їхнього потенціалу. Другому кластеру властивий високий пасажиропотік, оскільки Донецька область є промисловим центром, а Львівська – туристичним. У регіонах з третього кластеру середня кількість перевезень, він об'єднує області з різноманітними спрямуваннями економічного розвитку: туризм, важка промисловість та металургія, морські перевезення тощо. Четвертий кластер об'єднує менш розвинені економічно райони, переважає аграрний сектор, також менша щільність населення, що і спричиняє менший пасажиропотік. Жодних змін у структурі п'ятого сектору не відбулось. Як і раніше йому притаманні найнижча кількість перевезених пасажирів та менш розвинена інфраструктура транспортної мережі.

Аналогічні результати розподілу отримані і без врахування міста Києва при кластеризації за кількістю перевезених пасажирів за 2016-2019 та 2020-2021 роки, що подано у табл. А.3 та А.4 (додаток А). Отже суттєвого впливу на результати воно не має, бо структура груп областей залишилась без змін.

РОЗДІЛ 3

МОДЕЛЮВАННЯ ПОКАЗНИКІВ РИНКУ ТРАНСПОРТНИХ ПОСЛУГ

3.1 Інформаційна база дослідження

Для дослідження було використано методи кореляційного, регресійного аналізів, а також методи прогнозування для оцінки пасажирообігу та вантажообігу, а саме прогноз за моделями ARIMA та з використанням інструменту «Аркуш прогнозу» у середовищі MS Excel.

Кореляційний аналіз – метод дослідження взаємозалежності ознак у генеральній сукупності, які є випадковими величинами з нормальним законом розподілу [29]. У ході кореляційного аналізу ми визначаємо зв'язок між змінними та його силу. Слабкий кореляційний зв'язок каже про те, що змінні майже не залежать один від одного. Сильний зв'язок свідчить про те, що змінні є дуже залежними. Саме цю силу вимагає регресійний аналіз.

Наступним етапом використано модель множинної лінійної регресії. Це модель, яка пояснює залежність між змінними за допомогою лінійного взаємозв'язку. Для побудови якісної моделі, яку можна буде використати для прогнозу, треба визначати значущість параметрів моделі. Для регресії важливо аналізувати коефіцієнт детермінації, адже він вказує на те, на скільки відсотків рівняння регресії пояснює поведінку залежної змінної за розрахованою моделлю [30].

Моделі ARIMA використовуються для аналізу і прогнозування часового ряду. Вони спрямовані на опис автокореляцій в даних [31]. Вибір кращої серед побудованих моделей здійснюється за допомогою показника AICс – чим менше його значення, тим вища якість моделі [32]. Оцінка і порівняння таких різних варіантів моделей здійснюється за допомогою розрахунку похибок MAPE, MAE,

RMSE тощо. Важливо, щоб серед порівнюваних варіантів моделей кожен з цих показників якості був найменшим.

3.2 Регресійний аналіз для моделювання показників ринку транспортних послуг

Кореляційний аналіз

У багатьох економічних дослідженнях важливо знайти залежність між декількома ознаками та встановити їх взаємний зв'язок. Цю залежність можна виявити за допомогою коефіцієнта кореляції. Проведено кореляційний аналіз для підготовленої вибірки представленої на рис. Б.1 (додаток Б). Розраховані коефіцієнти кореляції для змінних подані у вигляді матриці на рис. Б.2 (додаток Б).

Великий вплив на пасажирообіг та вантажообіг мають курс долара, середня заробітна плата по країні, фактор настання пандемії COVID-19 та повномасштабного вторгнення. У свою чергу наявна сильна кореляція для вищезазначених факторів може бути першопричиною мультиколінеарності при моделюванні основних показників ринку транспортних послуг.

Моделювання пасажирообігу в Україні

Методика оцінювання пасажирообігу має бути придатною для практичного використання. Тому використано найбільш доцільний метод побудови і аналізу множинної лінійної регресії.

Множинна лінійна регресія – це модель, яка пояснює залежність між змінними за допомогою лінійного взаємозв'язку.

Загальна формула моделі множинної лінійної регресії має вигляд:

$$y_i = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + u_i \quad (3.1)$$

Для побудови моделей лінійної регресії було обрано квартальні дані за період 2014-2022 років по пасажирообігу, вантажообігу, середній заробітній платі по країні, курс долара тощо. Опис вхідних даних моделювання наведено у табл. 3.1.

Таблиця 3.1 – Вхідні дані для моделювання пасажирообігу та вантажообігу

Позначення	Опис	Одиниця вимірювання
pas_obig	Пасажирообіг	млн. пас. км
vantaj_obig	Вантажообіг	млн. тон км
dollar_rate	Курс долара	грн
bezrob	Кількість безробітного населення на кінець періоду	тис. осіб
average_salary	Середня заробітна плата по країні	грн
index	Індекс споживчих цін	%
invest	Прямі інвестиції в Україну	млн. дол. США
import_transport	Імпорт транспортних послуг	млн. дол. США

Джерело: розроблено автором на основі [33-35]

Для побудови моделі множинної лінійної регресії з кількісними змінними використано дані за період з 2014 по 2022 роки. Було обрано залежною змінною пасажирообіг (Y_i), а незалежними – курс долара (X_1), кількість безробітних на кінець періоду (X_2), середню заробітну плату по країні (X_3), індекс споживчих цін (X_4), прямі інвестиції в Україну (X_5), імпорт транспортних послуг (X_6).

Отримані у ході проведених розрахунків коефіцієнти, які будуть використані для побудови моделі, наведено на рис. 3.1.

Call:

```
lm(formula = pas_obig ~ dollar_rate + bezrob + average_salary +
    index + invest + import_transport, data = data_pas_van)
```

Coefficients:

(Intercept)	dollar_rate	bezrob	average_salary
83335.1925	-62.3486	-42.4445	-1.9949
index	invest	import_transport	
-326.2517	-0.6155	10.9427	

Рисунок 3.1 – Коефіцієнти для моделі пасажирообігу з кількісними змінними

Джерело: розроблено автором на основі [33-35]

Отже, модель має вигляд:

$$y_i = 83335,19 - 62,35 * x_1 - 42,44 * x_2 - 1,99 * x_3 - 326,25 * x_4 - 0,62 * x_5 + 10,94 * x_6 + u_i \quad (3.2)$$

де y_i – пасажирообіг,

x_1 – значення курсу долара,

x_2 – кількість безробітного населення,

x_3 – середня заробітна плата по країні,

x_4 – індекс споживчих цін,

x_5 – прямі інвестиції в Україну,

x_6 – імпорт транспортних послуг.

Інтерпретація отриманих показників:

b_0 - характеристика рівня автономності: значення пасажирообігу становить 83335,19 млн. пас. км., якщо значення незалежних змінних дорівнює 0;

b_1 – при збільшенні курсу долара на 1 грн. значення пасажирообігу зменшиться на 62,35 млн. пас. км.;

b_2 – зростання кількості безробітного населення на 1 тис. осіб спричинить зменшення пасажирообігу на 42,44 млн. пас. км.;

b_3 – збільшення середньої заробітної плати по країні на 1 грн. призведе до спадання пасажирообігу на 1,99 млн. пас. км.;

b_4 – при зростанні індексу споживчих цін на 1% значення пасажирообігу зменшиться на 326,25 млн. пас. км.;

b_5 – якщо значення прямих інвестицій в Україну збільшиться на 1 млн. дол. США, то пасажирообіг зменшиться на 0,62 млн. пас. км.;

b_6 – збільшення імпорту транспортних послуг на 1 млн. дол. США спричинить зростання пасажирообігу на 10,94 млн. пас. км.

Тільки при збільшенні імпорту транспортних послуг пасажирообіг зростатиме. Усі інші використані змінні за цією моделлю призведуть до зменшення.

Розраховані коефіцієнти еластичності для незалежних факторів моделі наведено у табл. 3.2.

Таблиця 3.2 – Коефіцієнти еластичності моделі пасажирообігу

E(X1)	-0,074477678
E(X2)	-0,749378661
E(X3)	-0,836479128
E(X4)	-1,606210547
E(X5)	0,020547255
E(X6)	0,298246376

Джерело: розроблено автором на основі [33-35]

Згідно з розрахованими коефіцієнтами еластичності для 6 незалежних факторів (X1, X2, X3, X4, X5, X6) варто зазначити, що найбільш впливовим фактором на пасажирообіг (Y) є індекс споживчих цін, менш впливовим – середня заробітна плата по країні, середнім за впливовістю є кількість безробітного населення на кінець періоду, найменш впливовими є імпорт транспортних послуг, курс долара і прямі інвестиції в Україну.

Результати здійсненої оцінки значимості параметрів моделі пасажирообороту наведено на рис. Б.3 (додаток Б). Статистично значимими показниками у цій моделі є кількість безробітного населення, середня заробітна плата по країні та індекс споживчих цін. Коефіцієнт детермінації дорівнює 0,8019, тобто це рівняння регресії на 80,19% пояснює поведінку залежної змінної. Модель непогано описує зв'язок між параметрами моделі, адже коефіцієнт детермінації та коефіцієнт детермінації скоригований ближчі до 1. Щільність зв'язку між параметрами моделі – сильна.

Результати здійсненої перевірки побудованої моделі на автокореляцію за критерієм Дарбіна-Уотсона наведено на рис. 3.2.

```

lag Autocorrelation D-W Statistic p-value
1      0.5255462      0.7870651      0
Alternative hypothesis: rho != 0

```

Рисунок 3.2 – Перевірка моделі пасажирообороту з кількісними змінними на автокореляцію

Джерело: розроблено автором на основі [33-35]

Тестова статистика становить 0,787, а відповідне значення p — 0. Оскільки p -value менше 0,05, можна відхилити нульову гіпотезу та зробити висновок, що залишки в цій моделі регресії є автокорельованими. Для вирішення проблеми можна додати до моделі фіктивні змінні, наприклад, фактор настання пандемії COVID-19 або ж фактор повномасштабного вторгнення.

Підсумки проведеного аналізу моделі на мультиколінеарність за допомогою функцій `omcdiag` і `imcdiag` представлено на рис. 3.3 та рис. 3.4, які забезпечують загальну та індивідуальну перевірку відповідно.

Overall Multicollinearity Diagnostics

	MC Results	detection
Determinant $ X'X $:	0.0589	0
Farrar Chi-Square:	91.0857	1
Red Indicator:	0.3733	0
Sum of Lambda Inverse:	20.1032	0
Theil's Method:	-0.6991	0
Condition Number:	99.9105	1

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

Рисунок 3.3 – Загальна перевірка моделі пасажирообороту з кількісними змінними на мультиколінеарність

Джерело: розроблено автором на основі [33-35]

Значення стандартизованого показника 0,0589 є дуже малим. Розрахункове значення статистики тесту Хі-квадрата становить 91,085, і це свідчить про наявність мультиколінеарності в моделі, оскільки розраховане значення більше табличного.

All Individual Multicollinearity Diagnostics Result

	VIF	TOL	Wi	Fi	Leamer	CVIF	Klein	IND1	IND2
dollar_rate	6.3859	0.1566	32.3152	41.7405	0.3957	6.7024	1	0.0261	1.5286
bezrob	3.4499	0.2899	14.6993	18.9866	0.5384	3.6209	0	0.0483	1.2870
average_salary	4.3803	0.2283	20.2817	26.1972	0.4778	4.5974	0	0.0380	1.3986
index	1.1825	0.8457	1.0947	1.4140	0.9196	1.2411	0	0.1410	0.2796
invest	1.1240	0.8897	0.7438	0.9608	0.9432	1.1797	0	0.1483	0.1999
import_transport	3.5808	0.2793	15.4847	20.0011	0.5285	3.7583	0	0.0465	1.3062

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

dollar_rate , invest , import_transport , coefficient(s) are non-significant may be due to multicollinearity

R-square of y on all x: 0.8019

* use method argument to check which regressors may be the reason of collinearity
 =====

Рисунок 3.4 – Індивідуальна перевірка моделі пасажирообороту з кількісними змінними на мультиколінеарність

Джерело: розроблено автором на основі [33-35]

F-статистика для змінних курсу долара (6,39), кількості безробітного населення (3,45), середньої заробітної плати по країні (4,38), імпорт транспортних послуг (3,58). Теоретичне значення F становить 2,36. Тобто ці змінні є першопричиною мультиколінеарності, оскільки розрахункове більше табличного.

Результат проведеної перевірки на гетероскедастичність залишків моделі за тестом Бройша-Пагана представлено на рис. 3.5.

Breusch-Pagan test

```
data: line_reg_multiple
BP = 4.7845, df = 6, p-value = 0.5717
```

Рисунок 3.5 – Перевірка моделі пасажирообігу з кількісними змінними на гетероскедастичність

Джерело: розроблено автором на основі [33-35]

Оскільки значення p-value 0,5717, то немає достатніх підстав для відкидання нульової гіпотези. Тобто немає причин вважати, що у моделі наявна гетероскедастичність.

Моделювання пасажирообігу (модель з якісними факторами)

Для побудови моделі множинної лінійної регресії з якісними змінними обрано залежною змінною пасажирообіг (Y_i), а незалежними – настання пандемії COVID-19 (d_1), повномасштабне вторгнення в Україну (d_2).

У ході проведених розрахунків отримано коефіцієнти, які будуть використані для побудови моделі подані на рис. 3.6.

```
Call:
lm(formula = pas_obig ~ covid19 + war, data = data_pas_van)

Coefficients:
(Intercept)      covid19          war
      25692         -11721         -6185
```

Рисунок 3.6 – Коефіцієнти моделі пасажирообігу з якісними змінними

Джерело: розроблено автором на основі [33-35]

Отже, модель має вигляд:

$$y_i = 25692 - 11721 * d_1 - 6185 * d_2 + u_i \quad (3.3)$$

де y_i – пасажирообіг,

d_1 – пандемія коронавірусу (0 – життя до настання пандемії, 1 – після),

d_2 – повномасштабне вторгнення (0 – життя до повномасштабного вторгнення, 1 – після).

Інтерпретація отриманих показників:

b_0 – характеристика рівня автономності: обсяг пасажирообороту становить 25692 млн. пас. км., якщо значення незалежних змінних дорівнює 0;

b_1 – якщо пандемія вже настала, то значення пасажирообігу зменшиться на 11721 млн. пас. км.;

b_2 – якщо вже відбулось повномасштабне вторгнення, то пасажирообіг зменшиться на 6185 млн. пас. км.

Результати оцінки значимості параметрів моделі з якісними факторами подано на рис. 3.7.

```

Call:
lm(formula = pas_obig ~ covid19 + war, data = data_pas_van)

Residuals:
    Min       1Q   Median       3Q      Max
-9947.2 -1742.5   145.9  1873.7  5374.0

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  25691.9     602.8  42.618 < 2e-16 ***
covid19     -11720.8    1205.7  -9.721 3.27e-11 ***
war         -17906.2    1595.0 -11.227 8.27e-13 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2953 on 33 degrees of freedom
Multiple R-squared:  0.8493,    Adjusted R-squared:  0.8401
F-statistic: 92.97 on 2 and 33 DF,  p-value: 2.757e-14

```

Рисунок 3.7 – Оцінка значимості моделі пасажирообороту з якісними змінними

Джерело: розроблено автором на основі [33-35]

Статистично значимими показниками у цій моделі є обидва фактори. Коефіцієнт детермінації дорівнює 0,8493, тобто це рівняння регресії на 84,93% пояснює поведінку залежної змінної. Модель гарно описує зв'язок між параметрами моделі, адже коефіцієнт детермінації та коефіцієнт детермінації скоригований ближчі до 1.

За критерієм Дарбіна-Уотсона було перевірено модель на наявність автокореляції. Підсумок наведено на рис. 3.8.

```

lag Autocorrelation D-W Statistic p-value
  1      -0.06668927     2.130777    0.958
Alternative hypothesis: rho != 0

```

Рисунок 3.8 – Перевірка за критерієм Дарбіна-Уотсона моделі пасажирообігу з якісними змінними

Джерело: розроблено автором на основі [33-35]

Тестова статистика становить 2,13, а відповідне значення p — 0,958. Оскільки p -value більше 0,05, можна прийняти нульову гіпотезу та зробити висновок, що залишки в цій моделі регресії не є автокорельованими.

Підсумки реалізованої загальної та індивідуальної перевірки аналізу моделі з якісними факторами на мультиколінеарність представлено на рис. 3.9 та 3.10 відповідно.

```
Overall Multicollinearity Diagnostics

MC Results detection
Determinant |X'X|:      0.9643      0
Farrar Chi-Square:    1.2183      0
Red Indicator:        0.1890      0
Sum of Lambda Inverse: 2.0741      0
Theil's Method:      -0.7778      0
Condition Number:     1.9319      0

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

Рисунок 3.9 – Загальна перевірка моделі пасажирообігу з якісними змінними на мультиколінеарність

Джерело: розроблено автором на основі [33-35]

Значення стандартизованого показника 0,964 є невеликим. Розрахункове значення статистики тесту Хі-квадрата становить 1,21, і це свідчить про відсутність мультиколінеарності в моделі, оскільки розраховане значення менше табличного.

```
All Individual Multicollinearity Diagnostics Result

      VIF   TOL   Wi Fi Leamer   CVIF Klein   IND1 IND2
covid19 1.037 0.9643 1.2593 Inf  0.982 0.5062   0 0.0284   1
war      1.037 0.9643 1.2593 Inf  0.982 0.5062   0 0.0284   1

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: 0.8493

* use method argument to check which regressors may be the reason of collinearity
=====
```

Рисунок 3.10 – Індивідуальна перевірка моделі пасажирообороту з якісними змінними на мультиколінеарність

Джерело: розроблено автором на основі [33-35]

F-статистика для обох факторів менше за теоретичне значення F (3,26). Тобто ці змінні вказують на відсутність мультиколінеарності в моделі.

Підсумок аналізу моделі на гетероскедастичність залишків моделі з якісними факторами за тестом Бройша-Пагана наведено на рис. 3.11.

```

studentized Breusch-Pagan test
data: line_reg_bin
BP = 5.9985, df = 2, p-value = 0.04982

```

Рисунок 3.11 – Аналіз моделі пасажирообігу з якісними змінними на гетероскедастичність

Джерело: розроблено автором на основі [33-35]

Оскільки значення p-value 0,049, то є підстави для відкидання нульової гіпотези. Тобто варто зауважити, що у моделі наявне явище гетероскедастичності.

Оптимальна модель і прогнозування пасажирообігу

Наступним кроком було визначено показники, які будуть значимими для моделі. Тож розраховані коефіцієнти для моделі зі значимими факторами зображено на рис. 3.12.

```

Call:
lm(formula = pas_obig ~ bezrob + import_transport + covid19 +
    war, data = data_pas_van)

Coefficients:
(Intercept)      bezrob  import_transport      covid19          war
  26574.29      -17.70       10.55      -11936.63      -8750.93

```

Рисунок 3.12 – Коефіцієнти моделі пасажирообігу зі значимими змінними

Джерело: розроблено автором на основі [33-35]

Модель має вигляд:

$$y_i = 26574,29 - 17,7 * x_1 + 10,55 * x_2 - 11936,63 * d_1 - 8750,93 * d_2 + u_i \quad (3.4)$$

де y_i – пасажирообіг,

x_1 – кількість безробітного населення,

x_2 – імпорт транспортних послуг,

d_1 – пандемія COVID-19 (0 – життя до настання пандемії, 1 – після),

d_2 – повномасштабне вторгнення (0 – життя до повномасштабного вторгнення, 1 – після).

Інтерпретація отриманих показників:

b_0 - значення пасажирообігу становить 26574,29 млн. пас. км., якщо значення незалежних змінних дорівнює 0;

b_1 – при збільшенні кількості безробітного населення на 1 тис. осіб пасажирообіг зменшиться на 17,7 млн. пас. км.;

b_2 – зростання імпорту транспортних послуг на 1 млн. дол. США спричинить збільшення пасажирообігу на 10,55 млн. пас. км.;

b_3 – врахування фактору настання пандемії призводить до зменшення пасажирообігу на 11936,63 млн. пас. км.;

b_4 – якщо враховувати, що вже відбулось повномасштабне вторгнення, то пасажирообіг зменшиться на 8750,93 млн. пас. км.

Перевірку значимості факторів наведено на рис. Б.3 (додаток Б). Значення p-value для кожного показника менше за 0,05, тож всі вони є значущими для оптимальної моделі. Тож можна використовувати цю модель для прогнозу. Коефіцієнт детермінації становить 0,9295, тобто це рівняння регресії на 92,95% пояснює поведінку залежної змінної. Модель добре відображає зв'язок між параметрами.

Найбільш впливовим фактором при моделюванні пасажирообігу є настання пандемії COVID-19. Середнім за впливовістю є фактор повномасштабного вторгнення. Найменший вплив мають кількість безробітного населення та імпорту транспортних послуг.

Здійснено перевірку оптимальної моделі на автокореляцію за критерієм Дарбіна-Уотсона, її результат наведено на рис. 3.13.

```

lag Autocorrelation D-W Statistic p-value
1 -0.3350142 2.624964 0.168
Alternative hypothesis: rho != 0

```

Рисунок 3.13 – Аналіз оптимальної моделі пасажирообороту на автокореляцію

Джерело: розроблено автором на основі [33-35]

Тестова статистика становить 2,63, а відповідне значення p-value - 0,168. Оскільки показник p-value більше 0,05, можна прийняти нульову гіпотезу та зробити висновок, що залишки в цій моделі регресії не є автокорельованими.

Підсумки проведеного аналізу моделі на мультиколінеарність за допомогою функцій `omcdiag` та `imcdiag` зображено на рис. 3.14 та 3.15 відповідно.

```
Call:
omcdiag(mod = line_reg_opt)

Overall Multicollinearity Diagnostics

Determinant |X'X|:          MC Results detection
Farrar Chi-Square:      0.5573           0
Red Indicator:          19.1962          1
Sum of Lambda Inverse: 0.2977           0
Theil's Method:        5.3991           0
Condition Number:      -1.8261           0
                    23.3314           0

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

Рисунок 3.14 – Загальна перевірка на мультиколінеарність оптимальної моделі пасажирообігу

Джерело: розроблено автором на основі [33-35]

Значення стандартизованого показника 0,56 є невеликим. Розрахункове значення статистики тесту Хі-квадрата становить 19,19, і це свідчить про відсутність мультиколінеарності в моделі, оскільки розраховане значення менше табличного.

```
Call:
imcdiag(mod = line_reg_opt)

All Individual Multicollinearity Diagnostics Result

          VIF   TOL   Wi   Fi Leamer   CVIF Klein   IND1   IND2
bezrob    1.6663 0.6001 7.1076 10.9946 0.7747 0.3899   0 0.0563 1.6622
import_transport 1.3377 0.7476 3.6021 5.5720 0.8646 0.3130   0 0.0701 1.0493
covid19   1.0674 0.9368 0.7192 1.1125 0.9679 0.2498   0 0.0878 0.2626
war       1.3277 0.7532 3.4954 5.4069 0.8679 0.3107   0 0.0706 1.0259

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: 0.9295

* use method argument to check which regressors may be the reason of collinearity
=====
```

Рисунок 3.15 – Індивідуальна перевірка мультиколінеарності моделі пасажирообігу зі значимими факторами

Джерело: розроблено автором на основі [33-35]

F-статистика для усіх факторів менше за теоретичне значення F (2,63). Тобто ці змінні вказують на відсутність мультиколінеарності в моделі.

Результати здійсненої перевірки на гетероскедастичність залишків моделі за тестом Бройша-Пагана для оптимальної моделі пасажирообороту представлено на рис. 3.16.

Breusch-Pagan test

```
data: line_reg_opt
BP = 12.772, df = 4, p-value = 0.01244
```

Рисунок 3.16 – Аналіз гетероскедастичності залишків моделі пасажирообігу зі значимими факторами

Джерело: розроблено автором на основі [33-35]

Оскільки значення p-value 0,012, то є підстави для відкидання нульової гіпотези. Тобто у моделі присутнє явище гетероскедастичності. На результати подальшого прогнозування наявність або відсутність гетероскедастичності не вплине.

Моделювання вантажообігу

Аналогічно до методики оцінювання пасажирообігу оцінено вантажообіг за допомогою лінійної регресії.

Для побудови моделі множинної лінійної регресії з кількісними змінними використано дані за період з 2014 по 2022 роки. Залежною змінною є вантажообіг (Y_i), а незалежними – курс долара (X_1), кількість безробітних на кінець періоду (X_2), середню заробітну плату по країні (X_3), індекс споживчих цін (X_4), прямі інвестиції в Україну (X_5), імпорт транспортних послуг (X_6).

Коефіцієнти розраховані для побудови моделі наведено на рис. 3.17.

```

Call:
lm(formula = vantaj_obig ~ dollar_rate + bezrob + average_salary +
    index + invest + import_transport, data = data_pas_van)

Coefficients:
(Intercept)      dollar_rate      bezrob      average_salary      index
199991.925         6.815         16.895         -2.624         -1209.767
invest import_transport
-2.150         29.509

```

Рисунок 3.17 – Коефіцієнти моделі вантажообігу з кількісними змінними

Джерело: розроблено автором на основі [33-35]

Отже, модель має вигляд:

$$y_i = 199991,93 + 6,82 * x_1 + 16,89 * x_2 - 2,62 * x_3 - 1209,77 * x_4 - 2,15 * x_5 + 29,51 * x_6 + u_i \quad (3.5)$$

де y_i – вантажообіг,

x_1 – значення курсу долара,

x_2 – кількість безробітного населення,

x_3 – середня заробітна плата по країні,

x_4 – індекс споживчих цін,

x_5 – прямі інвестиції в Україну,

x_6 – імпорт транспортних послуг.

Інтерпретація отриманих показників:

b_0 - характеристика рівня автономності: значення вантажообігу становить 199991,93 млн. тон км., якщо значення незалежних змінних дорівнює 0;

b_1 – при зростанні курсу долара на 1 грн. значення вантажообігу збільшиться на 6,82 млн. тон км;

b_2 – збільшення кількості безробітного населення на 1 тис. осіб спричинить зростання вантажообігу на 16,89 млн. тон км.;

b_3 – зростання середньої заробітної плати по країні на 1 грн. призводить до зменшення вантажообігу на 2,62 млн. тон км;

b_4 – якщо значення індексу споживчих цін збільшиться на 1%, то значення вантажообігу зменшиться на 1209,77 млн. тон км;

b_5 – збільшення обсягу прямих інвестицій в Україну на 1 млн. дол. США зумовить зменшення обсягів вантажообігу на 2,15 млн. тон км;

b_6 – при збільшенні імпорту транспортних послуг на 1 млн. дол. США значення вантажообігу збільшиться на 29,51 млн. тон км.

Тож при зростанні показників курсу долара, кількості безробітного населення та імпорту транспортних послуг вантажообіг зростатиме. Усі інші використані змінні за цією моделлю призведуть до його зменшення.

Розраховані коефіцієнти еластичності для незалежних факторів моделі вантажообігу представлено у табл. 3.3.

Таблиця 3.3 – Коефіцієнти еластичності моделі вантажообігу

E(X1)	0,002
E(X2)	0,083
E(X3)	-0,306
E(X4)	-1,653
E(X5)	0,020
E(X6)	0,223

Джерело: розроблено автором на основі [33-35]

Згідно до розрахованих коефіцієнтів еластичності для незалежних факторів (X1, X2, X3, X4, X5, X6) важливо зауважити, що найбільш впливовим фактором на вантажообіг (Y), є індекс споживчих цін, менш впливовим є середня заробітна плата по країні, середніми за впливовістю є імпорт транспортних послуг та кількість безробітного населення, найменш впливовими є прямі інвестиції в Україну і курс долара.

Підсумок здійсненого аналізу значимості параметрів моделі наведено на рис. 3.18.

```
Call:
lm(formula = vantaj_obig ~ dollar_rate + bezrob + average_salary +
    index + invest + import_transport, data = data_pas_van)

Residuals:
    Min       1Q   Median       3Q      Max
-21690  -6049   2461   6100  13295

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.000e+05  5.688e+04   3.516  0.00146 **
dollar_rate  6.815e+00  7.521e+02   0.009  0.99283
bezrob       1.689e+01  3.790e+01   0.446  0.65909
average_salary -2.624e+00  8.749e-01  -3.000  0.00550 **
index       -1.210e+03  4.258e+02  -2.841  0.00815 **
invest      -2.150e+00  1.441e+00  -1.492  0.14642
import_transport  2.951e+01  2.903e+01   1.017  0.31774
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10020 on 29 degrees of freedom
Multiple R-squared:  0.5837,    Adjusted R-squared:  0.4976
F-statistic: 6.776 on 6 and 29 DF,  p-value: 0.0001447
```

Рисунок 3.18 – Оцінка значущості параметрів моделі вантажообігу з кількісними змінними

Джерело: розроблено автором на основі [33-35]

Статистично значущими показниками моделі з кількісними змінними для показника вантажообігу є середня заробітна плата по країні та індекс споживчих цін. Коефіцієнт детермінації дорівнює 0,5837, тобто це рівняння регресії на 58,37% пояснює поведінку залежної змінної. Модель непогано описує зв'язок між факторами. Характерна середня щільність зв'язку між параметрами моделі.

Критерій Дарбіна-Уотсона є визначальним при перевірці моделі на автокореляцію. Результати аналізу наведено на рис. 3.19.

```
lag Autocorrelation D-w Statistic p-value
1 0.4327294 0.9830177 0
Alternative hypothesis: rho != 0
```

Рисунок 3.19 – Перевірка на автокореляцію залишків моделі вантажообігу з кількісними змінними

Джерело: розроблено автором на основі [33-35]

Тестова статистика становить 0,98, а відповідне значення p-value - 0. Оскільки p-value менше 0,05, відхилено нульову гіпотезу та визначено, що залишки в цій моделі регресії є автокорельованими. Для вирішення проблеми можна додати до моделі фіктивні змінні аналогічно до моделей пасажирообігу.

Підсумки індивідуальної перевірки моделі вантажообігу з кількісними факторами наведено на рис. Б.4 (додаток Б). Зважаючи, що результати є аналогічними до аналізу моделі пасажирообігу, то варто зазначити, що змінні курс долара, кількість безробітного населення, середня заробітна плата по Україні та імпорт транспортних послуг можуть бути першопричинами мультиколінеарності.

Гетероскедастичність залишків моделі перевірено за тестом Бройша-Пагана. Результати проведеного аналізу наведено на рис. 3.20.

```
studentized Breusch-Pagan test
data: line_reg_v_multiple
BP = 7.1546, df = 6, p-value = 0.3068
```

Рисунок 3.20 – Перевірка на гетероскедастичність залишків моделі вантажообігу з кількісними змінними

Джерело: розроблено автором на основі [33-35]

Оскільки значення р-value 0,31, то відсутні підстави для відкидання нульової гіпотези. Отже, немає причин вважати, що у моделі присутнє явище гетероскедастичності.

Моделювання вантажообігу (модель з якісними факторами)

При побудові моделі множинної лінійної регресії з якісними змінними залежною змінною обрано вантажообіг (Y_i), а незалежними – настання пандемії COVID-19 (d_1), повномасштабне вторгнення в Україну (d_2).

Отримані коефіцієнти для побудови моделі вантажообігу з якісними змінними представлено на рис. 3.21.

```
Call:
lm(formula = vantaj_obig ~ covid19 + war, data = data_pas_van)

Coefficients:
(Intercept)      covid19          war
      82900        -10436         -41217
```

Рисунок 3.21 – Коефіцієнти моделі вантажообігу з якісними змінними

Джерело: розроблено автором на основі [33-35]

Отже, модель має вигляд:

$$y_i = 82900 - 10436 * d_1 - 41217 * d_2 + u_i \quad (3.6)$$

де y_i – вантажообіг,

d_1 – пандемія коронавірусу (0 – життя до настання пандемії, 1 – після),

d_2 – повномасштабне вторгнення (0 – життя до повномасштабного вторгнення, 1 – після).

Інтерпретація отриманих показників:

b_0 – значення вантажообігу становить 82900 млн. тон. км., якщо значення незалежних змінних дорівнює 0;

b_1 – настання пандемії спричиняє зменшення значення вантажообігу на 10436 млн. тон. км.;

b_2 – врахування настання повномасштабного вторгнення зумовлює зменшення вантажообігу на 41217 млн. тон. км.

На відміну від моделі пасажирообігу з якісними змінними, фактор настання повномасштабного вторгнення має більший вплив на обсяги вантажообігу, ніж пандемія COVID-19.

Результати здійсненої перевірки значущості факторів моделі з якісними змінними наведено на рис. Б.5 (додаток Б). Статистично значимими є обидва фактори моделі. Коефіцієнт детермінації дорівнює 0,8506, тобто це рівняння регресії на 85,06% пояснює поведінку залежної змінної. Модель гарно відображає зв'язок між параметрами моделі. Щільність зв'язку параметрів моделі - сильна.

Результати перевірки моделі з якісними змінними на автокореляцію за критерієм Дарбіна-Уотсона подано на рис. 3.22.

```
lag Autocorrelation D-W Statistic p-value
1      0.1515124      1.655004    0.176
Alternative hypothesis: rho != 0
```

Рисунок 3.22 – Перевірка на автокореляцію моделі з якісними змінними

Джерело: розроблено автором на основі [33-35]

Тестова статистика становить 1,66, а відповідне значення p-value - 0,176. Оскільки p-value більше 0,05, нульова гіпотеза може бути прийнята та з'ясовано, що залишки в цій моделі регресії не є автокорельованими.

Результати проведеного аналізу моделі на мультиколінеарність, що забезпечує індивідуальну перевірку представлено на рис. Б.6 (додаток Б). Оскільки F-статистика для обох факторів менше за теоретичне значення, то ці змінні вказують на відсутність мультиколінеарності в моделі.

Гетероскедастичність залишків моделі перевірено за тестом Бройша-Пагана. Результати перевірки наведено на рис. 3.23.

```
studentized Breusch-Pagan test
data: line_reg_v_bin
BP = 2.9587, df = 2, p-value = 0.2278
```

Рисунок 3.23 – Перевірка на гетероскедастичність моделі вантажообігу з якісними змінними

Джерело: розроблено автором на основі [33-35]

Оскільки значення p-value 0,23, немає підстав для відкидати нульової гіпотези. Це означає, що у моделі відсутнє явище гетероскедастичності.

Оптимальна модель і прогнозування вантажообігу

Було проведено багато спроб і проаналізовано можливі варіанти покращення моделі. У підсумку було визначено показники, які будуть значимими для моделі. Розраховані коефіцієнти для моделі вантажообігу зі значимими факторами наведено на рис. 3.24.

```
Call:
lm(formula = vantaj_obig ~ import_transport + covid19 + war,
    data = data_pas_van)

Coefficients:
      (Intercept)  import_transport          covid19           war
      70466.99         22.15        -11166.53        -42529.53
```

Рисунок 3.24 – Коефіцієнти оптимальної моделі вантажообігу

Джерело: розроблено автором на основі [33-35]

Модель має вигляд:

$$y_i = 70466,99 + 22,15 * x_1 - 11166,53 * d_1 - 42529,53 * d_2 + u_i \quad (3.7)$$

де y_i – вантажообіг,

x_1 – імпорт транспортних послуг,

d_1 – пандемія COVID-19 (0 – життя до настання пандемії, 1 – після),

d_2 – повномасштабне вторгнення (0 – життя до повномасштабного вторгнення, 1 – після).

Інтерпретація отриманих показників:

b_0 – значення вантажообігу становить 70466,99 млн. тон км., якщо значення незалежних змінних дорівнює 0;

b_1 – збільшення обсягів імпорту транспортних послуг на 1 млн. дол. США зумовлює збільшення вантажообігу на 22,15 млн. тон км.;

b_2 – якщо пандемія вже настала, то значення вантажообігу зменшиться на 11166,53 млн. тон км.;

b_3 – якщо вже відбулось повномасштабне вторгнення, то вантажообіг зменшиться на 42529,53 млн. тон км.

Результатом перевірки значущості параметрів наведеної на рис. Б.7 (додаток Б). Підтверджено, що кожен фактор оптимальної моделі є значимим. Рівняння регресії на 87,95% пояснює поведінку залежної змінної. Модель гарно описує зв'язок між параметрами моделі. Отже цю модель допустимо використовувати для прогнозу.

Результат проведеної перевірки оптимальної моделі на автокореляцію за критерієм Дарбіна-Уотсона наведено на рис. 3.25.

```

lag Autocorrelation D-W Statistic p-value
1 0.1045496 1.723832 0.192
Alternative hypothesis: rho != 0

```

Рисунок 3.25 – Аналіз автокореляції за критерієм Дарбіна-Уотсона оптимальної моделі вантажообігу

Джерело: розроблено автором на основі [33-35]

Тестова статистика дорівнює 1,72, а відповідне значення p становить 0,192. Оскільки p -value більше 0,05, нульову гіпотезу можна прийняти, що дозволяє зробити висновок про відсутність автокореляції у цій регресійній моделі.

Підсумок індивідуального аналізу моделі на мультиколінеарність за допомогою функції `imcdiag` наведено на рис. Б.8 (додаток Б). Оскільки F -статистика кожного фактору менше за теоретичне значення $F(3,26)$, то варто зазначити, що мультиколінеарність у моделі відсутня.

Результати здійснення перевірки на гетероскедастичність залишків оптимальної моделі зображено на рис. 3. 26.

```

studentized Breusch-Pagan test

data: line_reg_v_opt
BP = 4.3743, df = 3, p-value = 0.2238

```

Рисунок 3.26 – Перевірка оптимальної моделі вантажообігу за тестом Бройша-Пагана

Джерело: розроблено автором на основі [33-35]

Оскільки значення p -value становить 0,22, відсутні підстави для відхилення нульової гіпотези. Тож гетероскедастичність залишків за тестом Бройша-Пагана є відсутньою у моделі.

3.3 Побудова моделей ARIMA показників ринку транспортних послуг

Моделі ARIMA для пасажиробігу

Визначення порядків для моделі грає важливу роль при побудові моделей ARIMA. Графіки автокореляції та часткової автокореляції зображено на рис. 3.27.

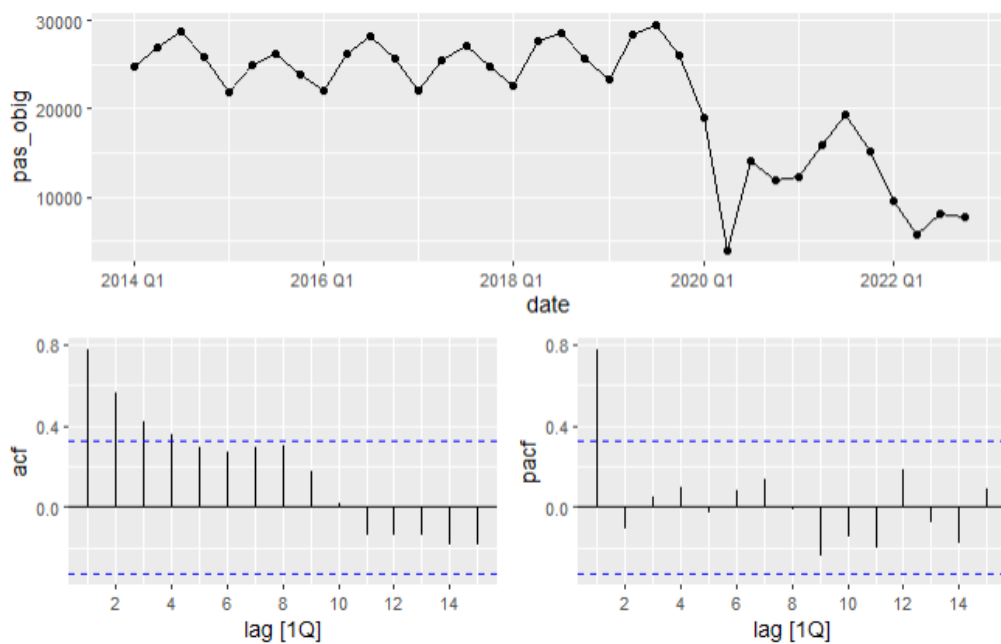


Рисунок 3.27 – Графік автокореляції та часткової автокореляції пасажиробігу

Джерело: розроблено автором на основі [33-35]

Порядок авторегресії (p) визначається за допомогою графіка PACF, а саме номер останнього значущого стовпчика. Кращим варіантом для моделі є p на рівні 1. Порядок середнього ковзання (q) відповідає останньому значущому стовпчику графіка ACF. Тобто, $q = 4$ є найбільш оптимальним варіантом. Порядок диференціювання (d) визначається шляхом підбору, якщо ряд не є стаціонарним.

Стаціонарність ряду перевірено за тестом KPSS (Квятковського-Філіпса-Шмідта-Шина), де нульова гіпотеза полягає в тому, що дані стаціонарні та несезонні [32]. Результати перевірки подано на рис. 3.28.

kpss_stat <dbl>	kpss_pvalue <dbl>	ndiffs <int>
0.7080074	0.01281751	1

Рисунок 3.28 – Тест KPSS для пасажирообігу

Джерело: розроблено автором на основі [33-35]

Статистика тесту kpss 0,71, p-value на рівні 0,01, що менше за 0,05. Тобто, є підстави для відхилення нульової гіпотези. Отже, дані не є стаціонарними і є сезонними.

Результати застосування ряду перших різниць зображено на рис. 3.29.

kpss_stat <dbl>	kpss_pvalue <dbl>	ndiffs <int>
0.1336677	0.1	0

Рисунок 3.29 – Тест KPSS з використанням ряду перших різниць для пасажирообігу

Джерело: розроблено автором на основі [33-35]

Оскільки p-value = 0,1, тобто більше за 0,05. Тобто є підстави, щоб прийняти нульову гіпотезу про те, що ряд є стаціонарним і несезонним. Також цей варіант є кращим, за попередній, оскільки значення kpss_stat = 0,13 є меншим. ndiffs показує, що більше не варто застосовувати різниць.

Серед різноманіття варіантів було обрано сезонні та несезонні моделі ARIMA для пасажирообороту і оцінено їх якість. Результати наведено на рис. 3.30.

.model <chr>	sigma2 <dbl>	log_lik <dbl>	AIC <dbl>	AICc <dbl>	BIC <dbl>	ar_roots <list>	ma_roots <list>
search	17070796	-340.5882	687.1764	687.9506	691.8424	<cplx [8]>	<cplx [0]>
arima410100	17110801	-338.9952	689.9904	692.9904	699.3225	<cplx [8]>	<cplx [0]>
arima211200	17328542	-339.2020	690.4040	693.4040	699.7361	<cplx [10]>	<cplx [1]>
arima113	19259863	-341.1939	692.3877	694.4567	700.1644	<cplx [1]>	<cplx [3]>
arima311	19336311	-341.2658	692.5316	694.6006	700.3084	<cplx [3]>	<cplx [1]>

Рисунок 3.30 – Оцінка якості моделей ARIMA для пасажирообігу

Джерело: розроблено автором на основі [33-35]

За критерієм AICс сезонна модель search є найкращою, оскільки показник найменший. У топ-3 входять також моделі ARIMA(4, 1, 0) (1, 0, 0) і ARIMA(2, 1, 1) (2, 0, 0).

Коефіцієнти моделі search_p наведено на рис. 3.31.

```
Series: pas_obig
Model: ARIMA(0,1,0)(2,0,0)[4]

Coefficients:
      sar1      sar2
      0.0830   0.3420
s.e.   0.1489   0.1526

sigma^2 estimated as 17070796:  log likelihood=-340.59
AIC=687.18  AICc=687.95  BIC=691.84
```

Рисунок 3.31 – Коефіцієнти для моделі пасажирообігу search_p

Джерело: розроблено автором на основі [33-35]

Загальний вигляд моделі search_p (ARIMA (0, 1, 0) (2, 0, 0)):

$$y_t = 0,083 * y_{t-4} + 0,342y_{t-8} + \varepsilon_t \quad (3.8)$$

де ε_t – білий шум зі стандартним відхиленням $4131,68 = \sqrt{17070796}$;

y_{t-4} , y_{t-8} – сезонні авторегресійні компоненти, тобто значення часового ряду у попередньому сезоні, 4 квартали назад або ж 8.

Розраховані коефіцієнти для моделі ARIMA (4, 1, 0) (1, 0, 0) представлено на рис. 3.32.

```
Series: pas_obig
Model: ARIMA(4,1,0)(1,0,0)[4]

Coefficients:
      ar1      ar2      ar3      ar4      sar1
      -0.1175  -0.2376  -0.1415  0.4952  -0.5356
s.e.   0.1471   0.1399   0.1447   0.1625   0.1594

sigma^2 estimated as 17110801:  log likelihood=-339
AIC=689.99  AICc=692.99  BIC=699.32
```

Рисунок 3.32 – Коефіцієнти для моделі ARIMA (4, 1, 0) (1, 0, 0)

Джерело: розроблено автором на основі [33-35]

Модель ARIMA (4, 1, 0) (1, 0, 0) має вигляд:

$$y_t = -0,118 * y_{t-1} - 0,238y_{t-2} - 0,142y_{t-3} - 0,495y_{t-4} - 0,537y_{t-4} + \varepsilon_t \quad (3.9)$$

де ε_t це білий шум зі стандартним відхиленням $4136,52 = \sqrt{171110801}$;

y_t – значення часового ряду в момент часу t ;

y_{t-1} – значення часового ряду в момент часу $t-1$;

y_{t-2} – значення часового ряду в момент часу $t-2$;

y_{t-3} – значення часового ряду в момент часу $t-3$;

y_{t-4} – значення часового ряду в момент часу $t-4$.

Результати здійсненого розрахунку коефіцієнтів моделі ARIMA (2, 1, 1) (2, 0, 0) представлено на рис. 3.33.

```
Series: pas_obig
Model: ARIMA(2,1,1)(2,0,0)[4]

Coefficients:
      ar1      ar2      ma1      sar1      sar2
      0.6500  -0.0970  -0.8038   0.1141   0.3323
s.e.    0.2959   0.2005   0.2535   0.1680   0.1681

sigma^2 estimated as 17328542:  log likelihood=-339.2
AIC=690.4  AICc=693.4  BIC=699.74
```

Рисунок 3.33 – Коефіцієнти для моделі пасажирообігу ARIMA (2, 1, 1) (2, 0, 0)

Джерело: розроблено автором на основі [33-35]

Модель ARIMA (2, 1, 1) (2, 0, 0) має вигляд:

$$y_t = 0,65 * y_{t-1} - 0,097y_{t-2} - 0,804\varepsilon_{t-1} - 0,114y_{t-4} - 0,332y_{t-8} + \varepsilon_t \quad (3.10)$$

де ε_t це білий шум зі стандартним відхиленням $4162,76 = \sqrt{17328542}$,

y_t – значення часового ряду в момент часу t ;

y_{t-1} – значення часового ряду в момент часу $t-1$;

y_{t-2} – значення часового ряду в момент часу $t-2$;

y_{t-4} – значення часового ряду в момент часу $t-4$ (попередній сезон);

y_{t-8} – значення часового ряду в момент часу $t-8$ (два попередні сезони);

ε_{t-1} – значення білого шуму в момент часу $t-1$.

Підсумки проведеної перевірки моделей за тестом Льюнга-Бокса представлено на рис. 3.34.

.model <chr>	lb_stat <dbl>	lb_pvalue <dbl>
arima113	0.5428132	0.4612692
arima211200	0.4178552	0.5180092
arima311	0.6765094	0.4107913
arima410100	0.7238048	0.3948987
search	1.9034912	0.1676881

Рисунок 3.34 – Тест Льюнга-Бокса для моделей пасажирообігу

Джерело: розроблено автором на основі [33-35]

За результатами тесту для всіх моделей p-value більше 0,05, тож нульова гіпотеза про те, що залишки є білим шумом приймається. Тобто залишки є схожими на білий шум.

Розраховані прогностні значення на 2023 рік за моделями ARIMA для пасажирообігу подано на рис. Б.9 (додаток Б). Порівняння прогнозів за топ-3 моделями з реальними значеннями представлено на рис. Б.10 – Б.12 (додаток Б). Отже, модель ARIMA (4, 1, 0) (1, 0, 0) має найбільш наближені до реальних значень прогностовані результати.

Моделі ARIMA для вантажообігу

Визначення порядків для моделі є ключовим аспектом при побудові моделей ARIMA. Для вантажообігу графіки автокореляції та часткової автокореляції подано на рис. 3.35.

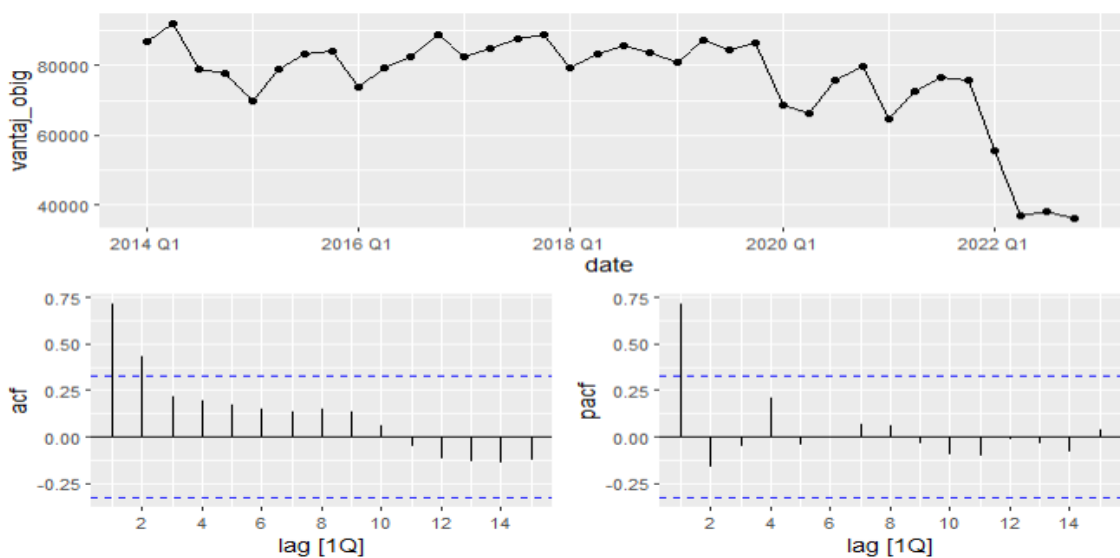


Рисунок 3.35 – Графік автокореляції та часткової автокореляції вантажообігу

Джерело: розроблено автором на основі [33-35]

Для моделювання вантажообігу порядок авторегресії дорівнює 1. Порядок середнього ковзання $q = 2$ виявився найкращим варіантом. Порядок диференціювання визначатиметься шляхом підбору, якщо ряд нестационарний.

Стаціонарність ряду перевіряється за тестом KPSS, нульова гіпотеза якого полягає в тому, що дані стаціонарні та несезонні [32]. Результати перевірки представлено на рис. 3.36.

kpss_stat <dbl>	kpss_pvalue <dbl>	ndiffs <int>
0.5826459	0.02421401	1

Рисунок 3.36 – Тест KPSS для вантажообігу

Джерело: розроблено автором на основі [33-35]

Наявні підстави для відхилення нульової гіпотези, оскільки статистика тесту kpss 0,58, p-value на рівні 0,02, а це менше за 0,05. Тобто дані не є стаціонарними і вони є сезонними.

Для вирішення питання стаціонарності застосовується ряд перших різниць, що подано на рис. 3.37.

kpss_stat <dbl>	kpss_pvalue <dbl>	ndiffs <int>
0.2800277	0.1	0

Рисунок 3.37 – Тест KPSS із використанням ряду перших різниць для вантажообігу

Джерело: розроблено автором на основі [33-35]

Застосування ряду перших різниць при диференціюванні вплинуло на те, що ряд став стаціонарним і несезонним. Це оцінено за основним показником p-value = 0,1, тобто більше за 0,05. Також зменшилось значення kpss_stat = 0,28. ndiffs показує, що більше не варто застосовувати різниць. Тож цей варіант є більш вдалим за попередній для створення прогнозів.

Для моделювання вантажообігу створено 5 різних варіантів моделей ARIMA. Оцінювання якості моделей представлено на рис. 3.38.

.model <chr>	sigma2 <dbl>	log_lik <dbl>	AIC <dbl>	AICc <dbl>	BIC <dbl>	ar_roots <list>	ma_roots <list>
search_v	44270457	-327.2607	662.5215	664.0030	668.3844	<cplx [5]>	<cplx [0]>
arima210	67206732	-364.0731	734.1463	734.9205	738.8123	<cplx [2]>	<cplx [0]>
arima310	64032355	-362.8130	733.6260	734.9593	739.8474	<cplx [3]>	<cplx [0]>
arima310001	61885187	-361.7251	733.4502	735.5191	741.2269	<cplx [3]>	<cplx [4]>
arima211	68189182	-363.8082	735.6164	736.9498	741.8378	<cplx [2]>	<cplx [1]>

Рисунок 3.38 – Оцінка якості моделей ARIMA для вантажообігу

Джерело: розроблено автором на основі [33-35]

За критерієм AICc найкращими є моделі search_v, ARIMA (2, 1, 0), ARIMA (3, 1, 0).

На рис. 3.39 подано коефіцієнти для моделі search_v (ARIMA (0,1,0)(1,1,0)).

```
Series: vantaj_obig
Model: ARIMA(1,0,0)(1,1,0)[4] w/ drift
```

Coefficients:

```
      ar1      sar1  constant
      0.9256 -0.6639 -1232.295
s.e.  0.0683  0.1662   954.182
```

```
sigma^2 estimated as 44270457: log likelihood=-327.26
AIC=662.52  AICc=664  BIC=668.38
```

Рисунок 3.39 – Коефіцієнти для моделі вантажообігу ARIMA (0, 1, 0) (1, 1, 0)

Джерело: розроблено автором на основі [33-35]

Модель ARIMA (0, 1, 0) (1, 1, 0) має вигляд:

$$y_t = -1232,295 + 0,926 * y_{t-1} - 0,664y_{t-4} + \varepsilon_t \quad (3.11)$$

де ε_t це білий шум зі стандартним відхиленням $6653,61 = \sqrt{44\,270\,457}$,

y_t – значення часового ряду в момент часу t ;

y_{t-1} – значення часового ряду в момент часу $t-1$;

y_{t-4} – значення часового ряду в момент часу $t-4$.

На рис. 3.40 представлено коефіцієнти для моделі ARIMA (2, 1, 0).

```
Series: vantaj_obig
Model: ARIMA(2,1,0)

Coefficients:
      ar1      ar2
    0.0065  -0.1790
s.e.  0.1657   0.1686

sigma^2 estimated as 67206732:  log likelihood=-364.07
AIC=734.15  AICc=734.92  BIC=738.81
```

Рисунок 3.40 – Коефіцієнти для моделі вантажообігу ARIMA (2, 1, 0)

Джерело: розроблено автором на основі [33-35]

Модель ARIMA (2, 1, 0) має вигляд:

$$y_t = 0,007 * y_{t-1} + 0,179y_{t-2} + \varepsilon_t \quad (3.12)$$

де ε_t це білий шум зі стандартним відхиленням $8197,97 = \sqrt{67\,206\,732}$;

y_{t-1} – значення часового ряду в момент часу $t-1$;

y_{t-2} – значення часового ряду в момент часу $t-2$.

На рис. 3.41 подано коефіцієнти для моделі ARIMA (3, 1, 0).

```
Series: vantaj_obig
Model: ARIMA(3,1,0)

Coefficients:
      ar1      ar2      ar3
    -0.0504  -0.1670  -0.2878
s.e.   0.1623   0.1616   0.1763

sigma^2 estimated as 64032355:  log likelihood=-362.81
AIC=733.63  AICc=734.96  BIC=739.85
```

Рисунок 3.41 – Коефіцієнти для моделі вантажообігу ARIMA (3, 1, 0)

Джерело: розроблено автором на основі [33-35]

Модель ARIMA (3, 1, 0) має вигляд:

$$y_t = -0,05 * y_{t-1} - 0,167y_{t-2} - 0,288y_{t-3} + \varepsilon_t \quad (3.13)$$

де ε_t це білий шум зі стандартним відхиленням $8002,02 = \sqrt{64\,032\,355}$;

y_t – значення часового ряду в момент часу t ;

y_{t-1} – значення часового ряду в момент часу $t-1$;

y_{t-2} – значення часового ряду в момент часу $t-2$;

y_{t-3} – значення часового ряду в момент часу $t-3$.

Результати аналізу моделей за тестом Льюнга-Бокса зображено на рис. 3.42.

.model <chr>	lb_stat <dbl>	lb_pvalue <dbl>
arima210	8.9350592	0.002797484
arima211	7.5441643	0.006020474
arima310	2.7445646	0.097585657
arima310001	0.5152258	0.472885359
search_v	2.8023099	0.094128612

Рисунок 3.42 – Перевірка моделі за тестом Льюнга-Бокса

Джерело: розроблено автором на основі [33-35]

Оскільки за результатами тесту Льюнга-Бокса для більшості моделей p-value більше 0,05, то нульова гіпотеза про те, що залишки є білим шумом приймається. Лише для моделей ARIMA (2, 1, 0) та ARIMA (2, 1, 1) залишки не є схожими на білий шум.

Розраховані значення прогнозів на 2023 рік за моделями ARIMA для вантажообігу представлено на рис. Б.13 (додаток Б). Порівняння прогнозних значень з реальними за топ-3 моделями зображено на рис. Б.14 – Б.16 (додаток Б). Таким чином, модель ARIMA (2, 1, 0) гарантує прогнозовані результати, які найбільше відповідають реальним значенням.

3.4 Прогнозування пасажирообігу та вантажообігу

На основі побудованих регресійних та ARIMA моделей розраховано прогнозні значення для залежних показників, а саме пасажирообігу та вантажообігу. Також для прогнозування застосовано «Аркуш прогнозу» у середовищі Excel та побудовано графіки прогнозів, що представлені на рис. Б.17 та Б.19, таблиці з прогнозними даними наведено на рис. Б.18 та Б.20 (додаток Б).

Розраховані похибки для кожного виду прогнозу пасажирообігу наведено у табл. 3. 4.

Таблиця 3.4 – Похибки прогнозів пасажирообігу на 2023 рік

Пасажирообіг				
Тип прогнозу	MAPE	MAE	MSE	RMSE
<i>Лінійна регресія</i>	22,18	2061,17	5085132,80	2255,02
<i>Аркуш прогнозу</i>	33,31	3433,48	14525760,20	3811,27
<i>arima113</i>	10,09	1085,51	1928445,02	1388,68
arima410100	3,56	326,36	137479,20	370,78
<i>arima311</i>	12,36	1319,92	2721941,24	1649,83
<i>arima211200</i>	7,29	738,42	699046,71	836,09
<i>search</i>	14,41	1450,36	2413059,49	1553,40

Джерело: розроблено автором на основі [33-35]

Розраховані похибки для 7 різних варіантів прогнозу пасажирообігу, як основного показника ринку транспортних послуг. Для прогнозів варто обрати лише дві моделі ARIMA (4, 1, 0) (1, 0, 0) та ARIMA (2, 1, 0) (0, 1, 0), оскільки за показником MAPE саме ці моделі входять до допустимого інтервалу 10%. Найбільш точним є прогноз за моделлю ARIMA (4, 1, 0) (1, 0, 0). Отже, саме цю побудовану модель найбільш доцільно використовувати для прогнозування пасажирообігу у часі.

Для кожного виду прогнозу вантажообігу на 2023 рік розраховано похибки, які наведено у табл. 3.5.

Таблиця 3.5 – Похибки прогнозів вантажообігу на 2023 рік

Вантажообіг				
Тип прогнозу	MAPE	MAE	MSE	RMSE
<i>Лінійна регресія</i>	17,68	7246,12	53410910,14	7308,28
<i>Аркуш прогнозу</i>	26,55	11148,77	195848155,74	13994,58
<i>search_v</i>	40,12	16348,98	270554918,02	16448,55
<i>arima210</i>	10,78	4467,28	23039039,49	4799,90
arima310	0,21	23,31	6783187,01	2604,46
<i>arima310001</i>	21,08	8729,98	90948179,12	9536,68
<i>arima211</i>	5,50	2316,62	8509990,39	2917,19

Джерело: розроблено автором на основі [33-35]

Розраховані похибки для 7 різних варіантів прогнозу вантажообігу, як основного показника ринку транспортних послуг. Для прогнозів варто обрати лише моделі ARIMA (3, 1, 0) та ARIMA (2, 1, 1), адже за показником MAPE вони входять до допустимого інтервалу 10%. Найбільшу точність має прогноз за моделлю ARIMA (3, 1, 0). Тож, найбільш доцільно здійснювати прогнозування саме за цією моделлю.

ВИСНОВКИ

Актуальність аналізу та моделювання показників ринку транспортних послуг зумовлена необхідністю оптимізації логістичних процесів, підвищення ефективності перевезень та забезпечення конкурентоспроможності у сучасних умовах глобалізації та цифрової трансформації економіки.

У даній роботі було узагальнено теоретичні аспекти та сучасні методи аналізу ринку транспортних послуг. Крім того, проведено аналіз ринку транспортних послуг у регіональному розрізі, а також здійснено моделювання та прогнозування ключових показників його розвитку.

На основі проведеного аналізу було виявлено, що за період з 2014 по січень 2022 років незначних змін зазнала структура обсягу перевезених пасажирів. Загалом темп приросту пасажирообігу за період 2014-2023 років зменшився на 59,3%, для вантажообігу – на 43,6%.

В роботі здійснено оцінювання збалансованості ринку транспортних послуг за регіонами. За результатами аналізу виявлено, що лідерські позиції в рейтингах перевезених пасажирів та вантажообігу за обрані періоди займали переважно Київська, Харківська, Дніпропетровська та Львівська області, оскільки саме вони у кожному рейтингу входили до топ-5. Спільними рисами для цих регіонів є розвинена промисловість різних спрямувань, стратегічне транспортне сполучення, висока щільність населення та інвестиційна привабливість.

Результатом кластерного аналізу для основних показників ринку транспортних послуг є розподіл регіонів за обсягом пасажирообігу на 4 кластери, а для вантажообігу на 5 груп. Спільною рисою таких розподілів є виокремлення міста Києва окремим кластером незважаючи на період та показник. Оскільки для столиці характерною ознакою є висока густина населення та розвинена транспортна інфраструктура. Також Київ є головним адміністративним, економічним та транспортним центром, що забезпечує великі обсяги пасажирообігу та вантажообігу порівняно з іншими регіонами. Розподіл інших областей також має

досить логічне обґрунтування структури отриманих кластерів. Зміни у наповненні груп у порівнювані періоди спричинені настанням пандемії COVID-19 та введеними карантинними обмеженнями, що мали значний вплив на економіку країни, зокрема на транспортну сферу.

У результаті проведеного кореляційного аналізу виявлено, що великий вплив на пасажирообіг та вантажообіг мають курс долара, середня заробітна плата по країні, фактор настання пандемії COVID-19 та повномасштабного вторгнення, але саме ці показники можуть бути причинами мультиколінеарності при моделюванні і спотворювати результати прогнозування. Тому у подальшому дослідженні ці показники не були додані одночасно в оптимальні моделі, що відобразилося на якості отриманих результатів.

В роботі побудовано регресійну модель для пасажирообігу і виявлено, що найбільш впливовими факторами на обсяг пасажирообігу є кількість безробітного населення, імпорт транспортних послуг та фіктивні змінні, що характеризують настання пандемії COVID-19 та повномасштабного вторгнення. Значення коефіцієнту детермінації 0,9295 свідчить про те, що згідно побудованої моделі варіація обсягу пасажирообігу на 92,95% визначається варіацією включених до моделі факторів, та на 7,05% – іншими недослідженими факторами.

Також було побудовано регресійну модель для вантажообігу і виявлено, що найбільший вплив має фактор імпорту транспортних послуг та бінарні змінні для пандемії COVID-19 й повномасштабного вторгнення. Коефіцієнт детермінації 0,8795 вказує на те, що в побудованій моделі 87,95% варіації обсягу вантажообігу пояснюється варіацією включених факторів, тоді як 12,05% зумовлено іншими, недослідженими факторами.

ARIMA-моделі дали більш точні результати прогнозування як пасажирообігу, так і вантажообігу, оскільки за показником MAPE результати не вийшли за межі рекомендованого діапазону 10%. Найбільш точний прогноз пасажирообігу дає модель ARIMA (4, 1, 0) (1, 0, 0), а для вантажообігу – ARIMA (3, 1, 0). Тому саме ці моделі рекомендовано використовувати для прогнозування основних показників ринку транспортних послуг.

ПЕРЕЛІК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Брайковська А. Дослідження особливостей формування ринку транспортних послуг як середовища функціонування підприємств транспорту. *Економіст*. 2012. № 9 (311). URL: http://nbuv.gov.ua/UJRN/econ_2012_9_14 (дата звернення: 30.04.2024).
2. Матвеева А. Ринок транспортних послуг: поняття, учасники та об'єкти. *Вісник Академії правових наук України*. 2011. №1 (64). С. 135-143.
3. Лукасевич-Крутник І. С. Поняття та ознаки транспортних послуг. *Юридичний вісник "Повітряне і космічне право"*. 2016. № 2 (39). URL: http://nbuv.gov.ua/UJRN/Npnau_2016_2_20 (дата звернення: 30.04.2024).
4. Яворський В. В., Лук'янчук К. П. Роль організаційно-економічних важелів публічного управління на ринок транспортних послуг в Україні. *Сучасні управлінські та соціально-економічні аспекти розвитку держави, регіонів та суб'єктів господарювання в умовах трансформації публічного управління: матеріали міжнар. наук.-практ. конф., Одеса, 2023*. URL: https://economics.net.ua/files/science/admin_men/2023/tezy.pdf#page=127 (дата звернення: 01.05.2024).
5. Стащенко Ю., Гавриловський О. Сучасні тенденції розвитку глобального ринку транспортно-логістичних послуг. *Економіка та суспільство*. 2022. № 39. URL: <https://doi.org/10.32782/2524-0072/2022-39-87> (дата звернення: 01.05.2024).
6. Україна: транспорт. *Електронна версія «Великої української енциклопедії»*: веб-сайт. URL: https://vue.gov.ua/Україна:_транспорт (дата звернення: 27.04.2024).
7. Інформація про Українські залізниці. *Міністерство розвитку громад, територій та інфраструктури України*: веб-сайт. URL: <https://mtu.gov.ua/content/informaciya-pro-ukrainski-zaliznici.html> (дата звернення: 29.04.2024).
8. Транспортна екологія: навч. посіб. / О. І. Запорожець та ін. ; за заг. ред. С. В. Бойченка. Київ : НАУ, 2017. 507 с. URL: <https://bit.ly/3UXTWK4> (дата звернення: 30.04.2024).

9. Проблеми та перспективи розвитку транспортної галузі України. *Наукове мислення*: веб-сайт. URL: <https://naukam.triada.in.ua/index.php/konferentsiji/50-dvadsyata-vseukrajinska-praktichno-piznavalna-internet-konferentsiya/461-problemi-ta-perspektivi-rozvitku-transportnoji-galuzi-ukrajini> (дата звернення: 03.05.2024).
10. Правові засади державного управління повітряним транспортом України. *Букліб*: веб-сайт URL: <https://buklib.net/books/25349/> (дата звернення: 03.05.2024).
11. Морський і річковий транспорт України. *Інформаційно-довідковий сайт «Україна»*: веб-сайт. URL: http://proukraine.net.ua/?page_id=469 (дата звернення: 29.04.2024).
12. Церковна А. В., Рахімі А. М. Внутрішній водний транспорт України: проблеми та перспективи. *Ринкова економіка: сучасна теорія і практика управління*. 2023. Т. 21 №3(52). URL: [https://doi.org/10.18524/2413-9998.2022.3\(52\).275828](https://doi.org/10.18524/2413-9998.2022.3(52).275828) (дата звернення: 03.05.2024).
13. Дивінець О. Л. Стратегічні пріоритети розвитку міських електротранспортних підприємств України. *Європейський вектор економічного розвитку*. 2015. № 1 (18). URL: http://nbuv.gov.ua/UJRN/ever_2015_1_11 (дата звернення: 05.05.2024).
14. Метрополітени України. *INSIDE-UA*: веб-сайт. URL: <https://inside-ua.com/metro> (дата звернення: 03.05.2024).
15. Іщенко Н.М. Маркетинг транспортних послуг : навч. посіб. Миколаїв : ЧДУ, 2013. 272 с. URL: <https://lib.chmnu.edu.ua/pdf/posibnuku/333/4.pdf> (дата звернення: 27.04.2024).
16. Ареф'єва, О., Годун, В. Стратегізація підвищення конкурентоспроможності підприємства на ринку транспортних послуг. *Економіка та суспільство*. 2023. № 55. URL: <https://doi.org/10.32782/2524-0072/2023-55-113> (дата звернення: 05.05.2024).
17. Васильєв О. Л., Богомаз С. М. Оцінка якості послуг залізничного транспорту. *Вісник економіки транспорту і промисловості*. 2019. № 67. URL: <https://doi.org/10.18664/338.47:338.45.v0i67.180884> (дата звернення: 07.05.2024).

18. Макаренко В. Д., Луб'яний П. В., Войтович О. А. Якість пасажирських транспортних послуг на основі нефінансових показників. *Вісник Херсонського національного технічного університету*. 2023. № 1 (84). URL: <https://doi.org/10.35546/kntu2078-4481.2023.1.6> (дата звернення: 07.05.2024).
19. Маркетинг транспортних послуг: навч. посіб. / О. І. Зоріна та ін. ; за ред. О. І. Зоріної. Харків : УкрДУЗТ. 2018. URL: http://lib.kart.edu.ua/bitstream/123456789/2443/1/Навчальний_посібник.pdf (дата звернення: 07.05.2024).
20. Давідіч Ю. О., Фалецька Г. І., Ольхова М. В. Конспект лекцій з дисципліни «Ефективність транспорту». Харків : ХНУМГ ім. О. М. Бекетова. 2019. 74 с. URL: http://eprints.kname.edu.ua/51058/1/2017_репоз_КЛ_146Л_ЕТ.pdf (дата звернення: 07.05.2024).
21. Терміни з теми "Транспорт України". *Школа сучасних знань*: веб-сайт. URL: http://www.zhu.edu.ua/mk_school/mod/glossary/view.php?id=4409 (дата звернення: 07.05.2024).
22. Техніко-економічні основи транспорту. *Сервер електронних курсів ТНПУ ім. В. Гнатюка*: веб-сайт. URL: https://elr.tnpu.edu.ua/pluginfile.php/84751/mod_resource/content/1/Lekcija_18.pdf (дата звернення: 05.05.2024).
23. Експортні можливості у 2024: морський коридор, страхування суден та річкова логістика. *Дія.Бізнес Експортний напрям*: веб-сайт. URL: https://export.gov.ua/news/5173-eksportni_mozhlivosti_u_2024_morskii_koridor_strakhuvannia_suden_ta_richkova_logistika (дата звернення: 05.05.2024).
24. Експорт може зрости на 9% у 2024 році – позиція Мінекономіки. *«Дебет-Кредит» Український фінансово-бухгалтерський портал*: веб-сайт. URL: <https://news.dtkr.ua/society/economics/88130-eksport-moze-zrosti-na-9-u-2024-roci-rozicijia-minekonomiki> (дата звернення: 07.05.2024).
25. Розвиток транспорту і логістики в Україні. *Блог UC.Market*: веб-сайт. URL: <https://blog.youcontrol.market/rozvitok-transportu-i-loghistiki-v-ukrayini/> (дата звернення: 05.05.2024).

26. Потреби України на відновлення сфери транспорту та логістики у 2024 році становлять 2,3 млрд доларів. *Центр транспортних стратегій*: веб-сайт. URL: https://cfts.org.ua/news/2024/02/16/potrebi_ukrani_z_vidnovlennya_sferi_transportu_ta_logistiki_u_2024_rotsi_stanovlyat_23_mlrd_dolariv_miu_78207 (дата звернення: 03.05.2024).

27. Укрзалізниця завершила заміну постільної білизни та переглянула її вартість. *Укрзалізниця*: веб-сайт. URL: https://www.uz.gov.ua/press_center/up_to_date_topic/628899/ (дата звернення: 22.05.2024).

28. Укрзалізниця планує поповнити парк сучасними електровозами. *Укрзалізниця*: веб-сайт. URL: https://www.uz.gov.ua/press_center/up_to_date_topic/629307/ (дата звернення: 24.05.2024).

29. Метод і методика економічного аналізу. *Луцький національний технічний університет*: веб-сайт. URL: https://elib.lntu.edu.ua/sites/default/files/elib_upload/Талах_Т_А_2/page5.html (дата звернення: 07.05.2024).

30. Загальна лінійна економетрична модель. *StudFiles*: веб-сайт. URL: <https://studfile.net/preview/7867367/page:6/> (дата звернення: 05.05.2024).

31. Hyndman R. J., Athanasopoulos G. Forecasting: Principles and Practice. OTexts. 2021. URL: <https://otexts.com/fpp3/> (дата звернення: 05.05.2024).

32. Regression Model Accuracy Metrics: R-square, AIC, BIC, Cp and more. *Statistical tools for high-throughput data analysis*: веб-сайт. URL: <http://www.sthda.com/english/articles/38-regression-model-validation/158-regression-model-accuracy-metrics-r-square-aic-bic-cp-and-more/> (дата звернення: 03.05.2024).

33. Показники ринку транспортних послуг. *Державна служба статистики України*: веб-сайт. URL: <https://www.ukrstat.gov.ua/> (дата звернення: 05.05.2024).

34. Транспорт. *Держстат*: веб-сайт. URL: <https://stat.gov.ua/uk> (дата звернення: 07.05.2024).

35. Макроекономічні показники. *Національний банк України*: веб-сайт. URL: <https://bank.gov.ua/> (дата звернення: 05.05.2024).

ДОДАТКИ

ДОДАТОК А

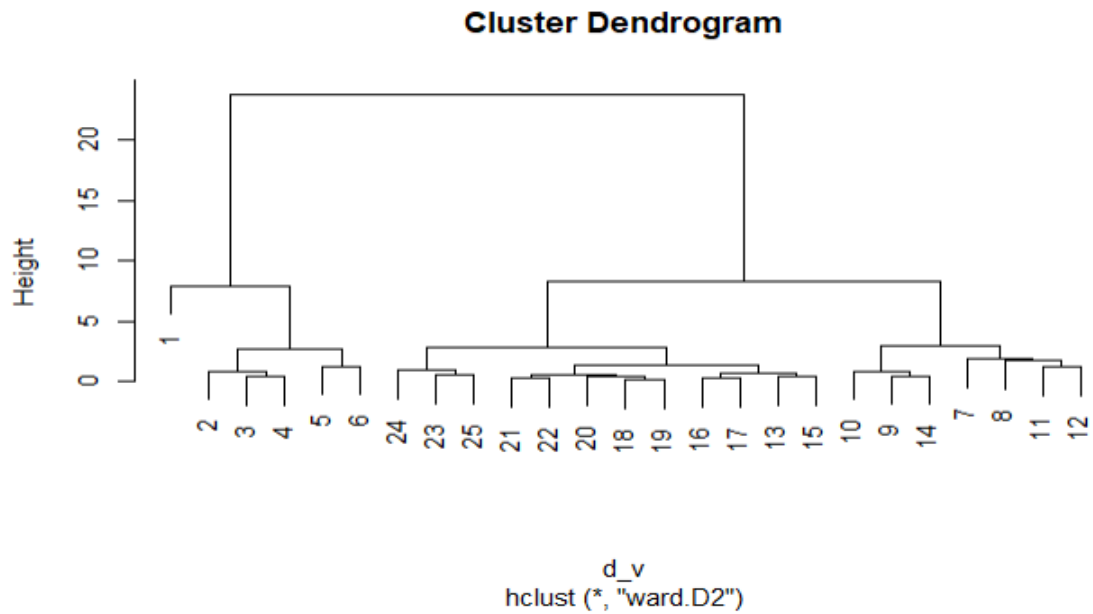


Рисунок А.1 – Дендрограма для показника вантажообігу за 2016-2019 роки
Джерело: розроблено автором на основі [33-35]

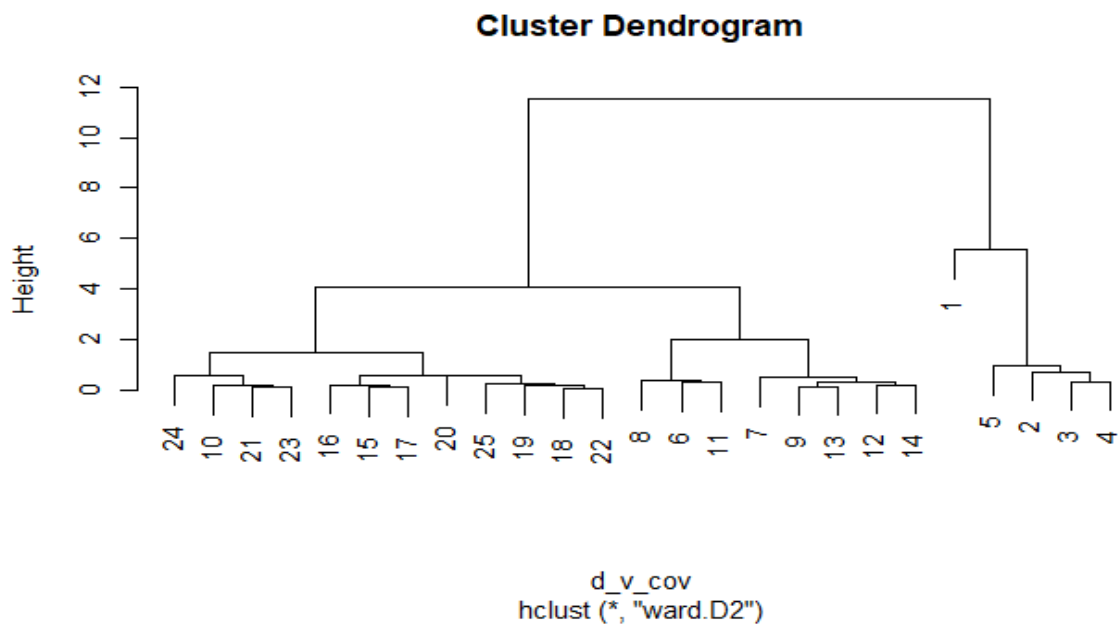


Рисунок А.2 – Дендрограма для показника вантажообігу в Україні за 2020-2021 роки

Джерело: розроблено автором на основі [33-35]

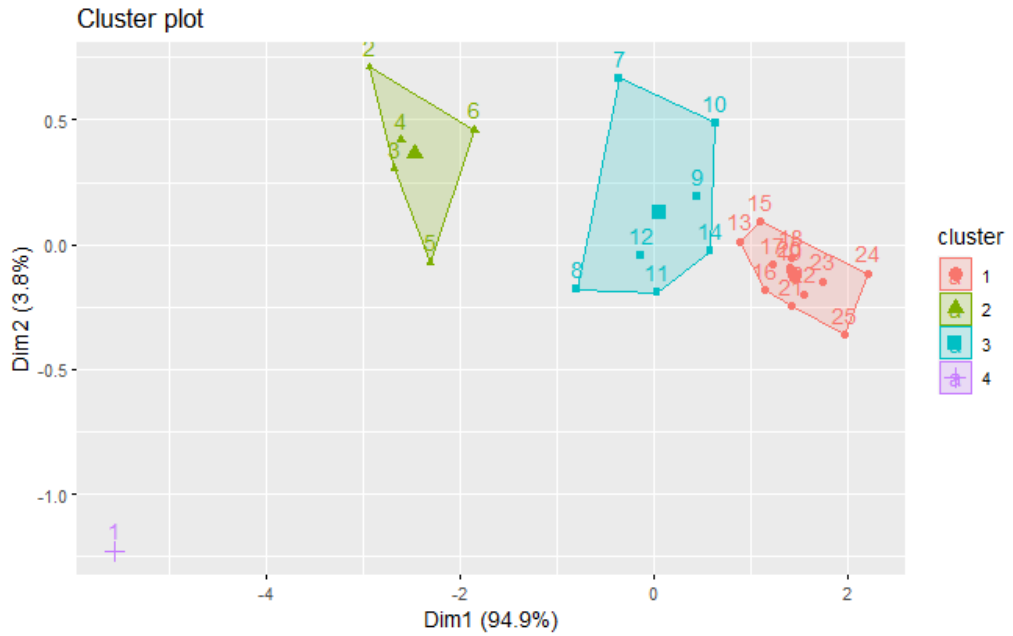


Рисунок А.3 – Розподіл показника вантажообігу в Україні на кластери методом к-середніх за 2016-2019 роки

Джерело: розроблено автором на основі [33-35]

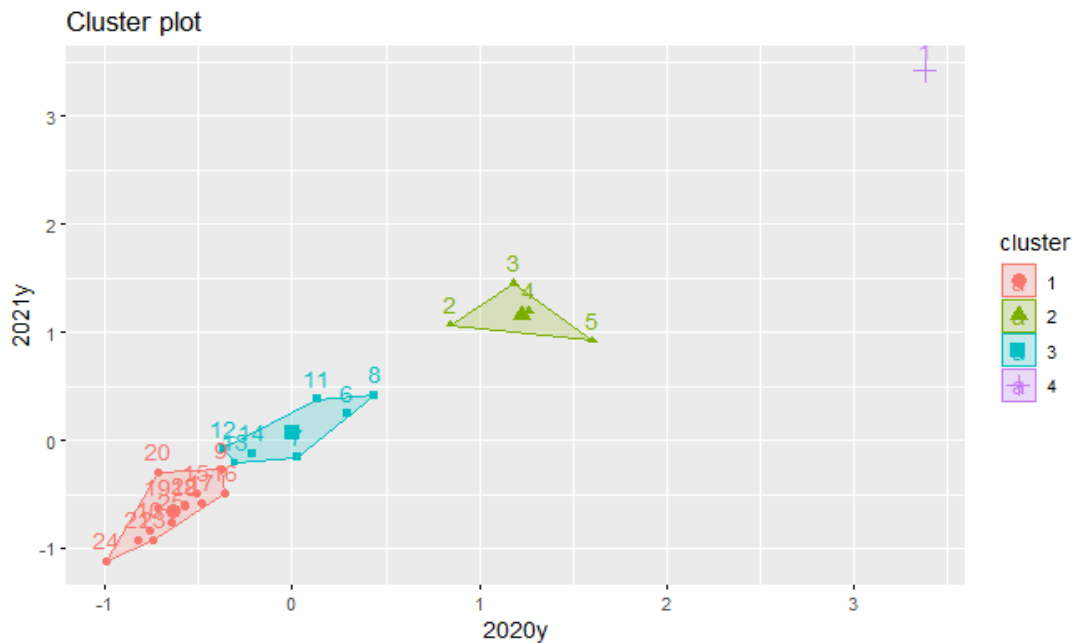


Рисунок А.4 – Розподіл показника вантажообігу в Україні на кластери методом к-середніх за 2020-2021 роки

Джерело: розроблено автором на основі [33-35]

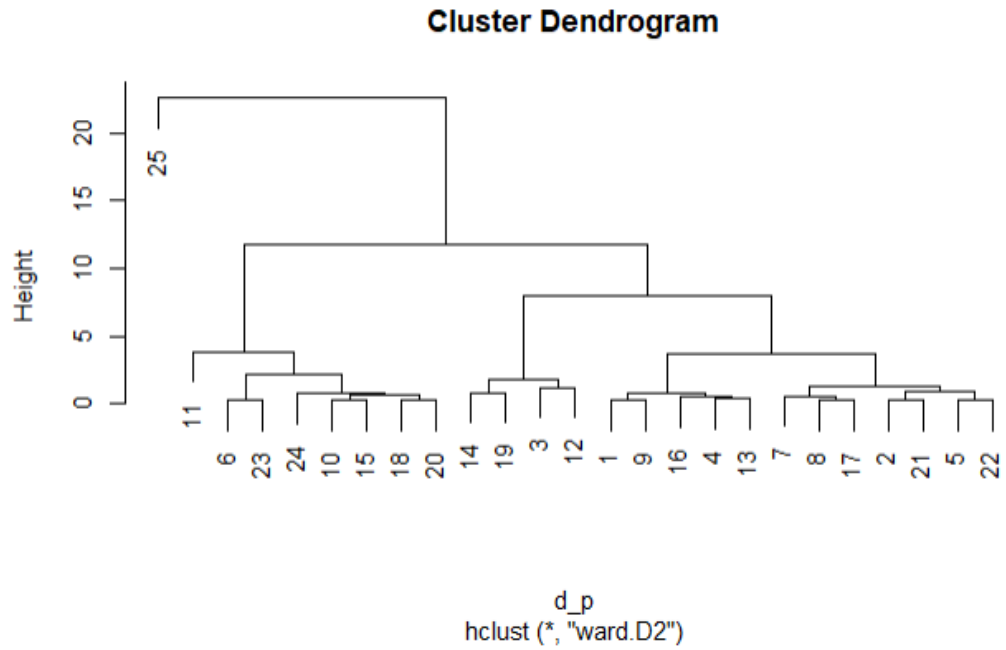


Рисунок А.5 – Дендрограма для кількості перевезених пасажирів в Україні за 2016-2019 роки

Джерело: розроблено автором на основі [33-35]

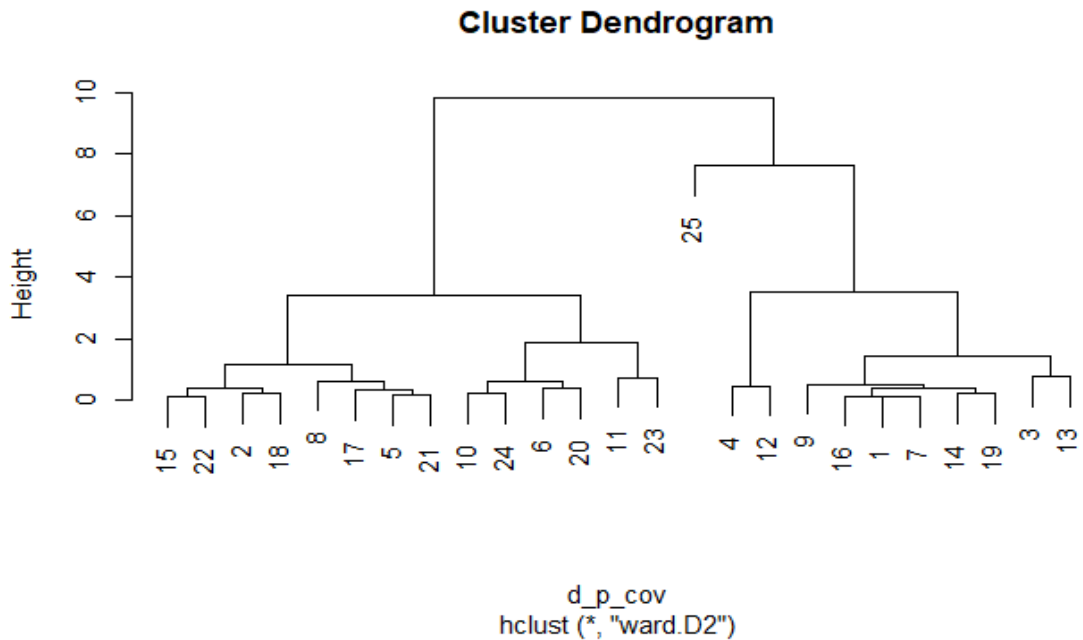


Рисунок А.6 – Дендрограма для кількості перевезених пасажирів в Україні за 2020-2021 роки

Джерело: розроблено автором на основі [33-35]

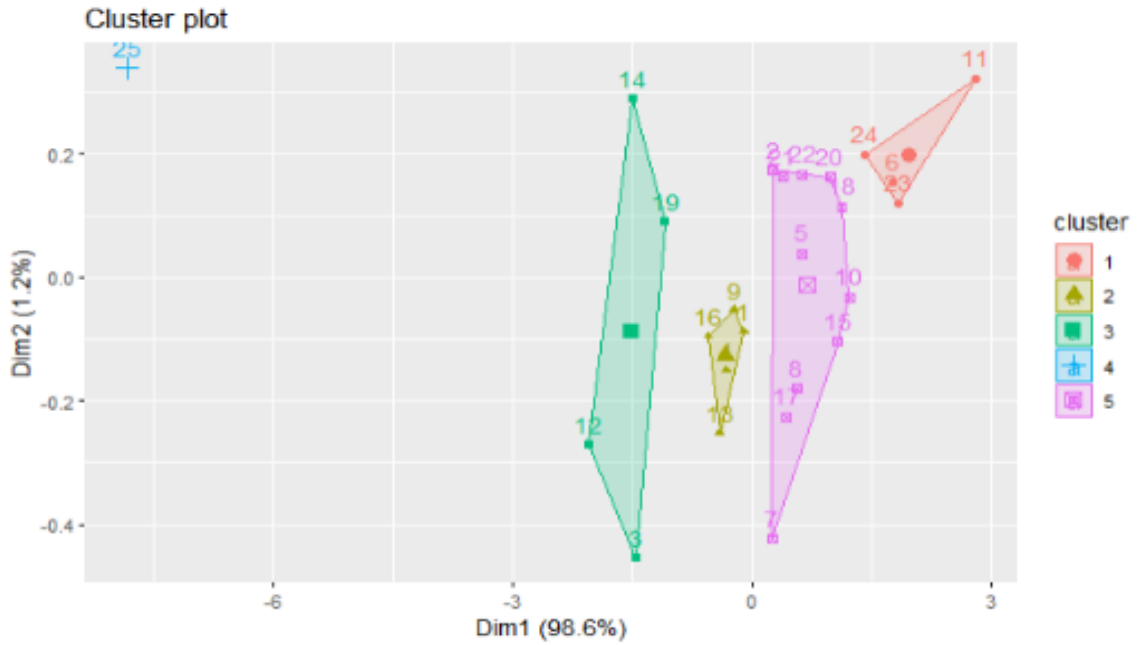


Рисунок А.7 – Розподіл кількості перевезених пасажирів в Україні на кластери методом к-середніх за 2016-2019 роки
 Джерело: розроблено автором на основі [33-35]

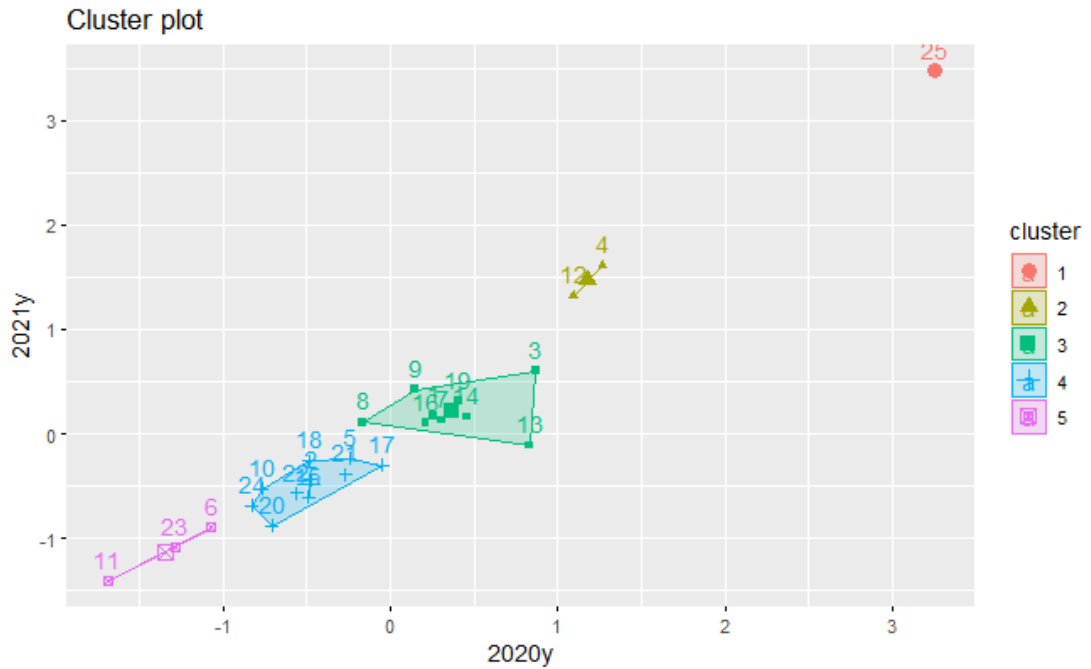


Рисунок А.8 – Розподіл кількості перевезених пасажирів в Україні на кластери методом к-середніх за 2020-2021 роки
 Джерело: розроблено автором на основі [33-35]

Таблиця А.1 – Структура кластерів вантажообігу за областями України за 2016-2019 роки

<i>Кластер</i>	<i>Область</i>
1	Полтавська, Одеська, Хмельницька, Донецька, Волинська, Черкаська, Рівненська
2	Вінницька, Івано-Франківська, Миколаївська, Запорізька, Тернопільська, Херсонська, Кіровоградська, Чернівецька, Чернігівська, Житомирська, Луганська, Сумська
3	Закарпатська, Львівська, Дніпропетровська, Київська, Харківська

Джерело: розроблено автором на основі [33-35]

Таблиця А.2 – Структура кластерів вантажообігу за областями України за 2020-2021 роки

<i>Кластер</i>	<i>Область</i>
1	Закарпатська, Львівська, Дніпропетровська, Київська
2	Хмельницька, Донецька, Івано-Франківська, Миколаївська, Запорізька, Тернопільська, Херсонська, Кіровоградська, Чернівецька, Чернігівська, Житомирська, Луганська, Сумська
3	Харківська, Полтавська, Одеська, Волинська, Черкаська, Вінницька, Рівненська

Джерело: розроблено автором на основі [33-35]

Таблиця А.3 – Структура кластерів кількості перевезених пасажирів за областями України за 2016-2019 роки

<i>Кластер</i>	<i>Область</i>
1	Дніпропетровська, Львівська, Одеська, Харківська
2	Вінницька, Донецька, Київська, Миколаївська, Рівненська
3	Волинська, Житомирська, Запорізька, Івано-Франківська, Кіровоградська, Полтавська, Сумська, Тернопільська, Херсонська, Хмельницька, Черкаська, Чернігівська
4	Закарпатська, Луганська, Чернівецька

Джерело: розроблено автором на основі [33-35]

Таблиця А.4 – Структура кластерів кількості перевезених пасажирів за областями України за 2020-2021 роки

<i>Кластер</i>	<i>Область</i>
1	Донецька, Львівська
2	Вінницька, Дніпропетровська, Запорізька, Івано-Франківська, Київська, Миколаївська, Одеська, Рівненська, Харківська
3	Волинська, Житомирська, Кіровоградська, Полтавська, Сумська, Тернопільська, Херсонська, Хмельницька, Черкаська, Чернігівська
4	Закарпатська, Луганська, Чернівецька

Джерело: розроблено автором на основі [33-35]

ДОДАТОК Б

Дата	Пасажиробіг (млн. пас. км)	Вантажобіг (млн. тонн км)	Курс долара (грн)	К-сть безроб на кінець періоду (тис. осіб)	Середня заробітна плата (грн)	Індекс споживчих цін (%)	Прямі інвестиції в Україну (млн. дол. США)	Імпорт транспортних послуг (млн. дол. США)
1 кв 2014	24829,6	86842	8,86	492,3	3245	103	665	696
2 кв 2014	26984,3	91864,6	11,69	437,5	3488	108,3	319	694
3 кв 2014	28691,8	78798,3	12,58	418,1	3463	104,1	-725	695
4 кв 2014	25835,7	77893	14,42	512,2	3685	107,5	-558	642
1 кв 2015	21866,9	69648,7	22,45	506,8	3650	120,3	3390	469
2 кв 2015	25015,8	78867,2	21,04	443,9	4113	117	-3159	480
3 кв 2015	26210,5	83294,9	21,44	407,4	4313	100,5	-1170	508
4 кв 2015	23943,2	84069,1	23,60	490,8	4753	101,4	1175	490
1 кв 2016	22101,4	73603,6	26,14	467,5	4622	101,5	-468	412
2 кв 2016	26248	79357,8	25,07	388,9	5072	103,4	-2200	435
3 кв 2016	28196,5	82319,8	25,45	341,5	5311	101,4	-1109	559
4 кв 2016	25653,4	88863,8	26,10	390,8	5744	105,6	-178	528
1 кв 2017	22037,5	82310,6	27,05	406,8	6323	103,9	-964	463
2 кв 2017	25542,6	84710,4	26,33	330,2	6953	103,8	-1374	518
3 кв 2017	27099,2	87523,7	26,01	303,0	7268	102,1	-391	565
4 кв 2017	24729,3	88810,1	27,30	354,4	7878	103,1	-717	563
1 кв 2018	22618,7	79293,6	27,17	366,9	7974	103,5	-1870	495
2 кв 2018	27641	83222,2	26,19	303,9	8782	100,8	-912	544
3 кв 2018	28569,3	85520,4	27,78	287,1	9063	101,2	-21	602
4 кв 2018	25617	83820	28,08	341,7	9651	103,9	-2056	596
1 кв 2019	23362,4	80862,9	27,33	340,7	9630	102,4	-864	561
2 кв 2019	28410,7	87195,4	26,55	287,1	10430	101,2	-1428	600
3 кв 2019	29428,5	84621,5	24,80	268,2	10732	99,8	-2101	684
4 кв 2019	25973,4	86282,7	24,24	338,2	11223	100,6	-782	671
1 кв 2020	18995,5	68464,1	25,85	349,4	11007	100,7	1630	534
2 кв 2020	4023,9	66307,4	26,86	517,7	10850	101,3	-1274	372
3 кв 2020	14119,2	75660,7	27,82	433,4	11749	99,7	7	467
4 кв 2020	11896,2	79647,1	28,39	459,2	12780	103,2	-305	547
1 кв 2021	12230,7	64522,3	28,00	449,7	12835	104,1	-1517	526
2 кв 2021	15982,4	72792,5	27,48	344,8	13784	102,2	-1193	617
3 кв 2021	19345,1	76582,3	26,77	285,9	14194	101,1	-2524	793
4 кв 2021	15176	75738,3	26,93	295,0	15254	102,3	-2284	898
1 кв 2022	9634,3	55474,2	29,10	286,9	14620	107,6	577	672
2 кв 2022	5691,3	36875,3	29,25	316,4	13723	109,2	-286	477
3 кв 2022	8102,5	38022,3	36,57	260,5	14619	103,7	-332	616
4 кв 2022	7714,7	36359,5	36,57	186,5	16477	103,9	-172	717
1 кв 2023	8144,1	39544,3	36,57	137,3	15436	103	-1145	753
2 кв 2023	9553,2	39637,5	36,57	113,4	17176	101,5	-1304	709
3 кв 2023	11168	40264,3	36,57	103,3	17937	98,5	-1640	745
4 кв 2023	10522,3	43913,8	36,91	96,1	19231	102	-116	831

Рисунок Б.1 – Вхідні дані для моделювання основних показників ринку транспортних послуг

Джерело: розроблено автором на основі [33-35]

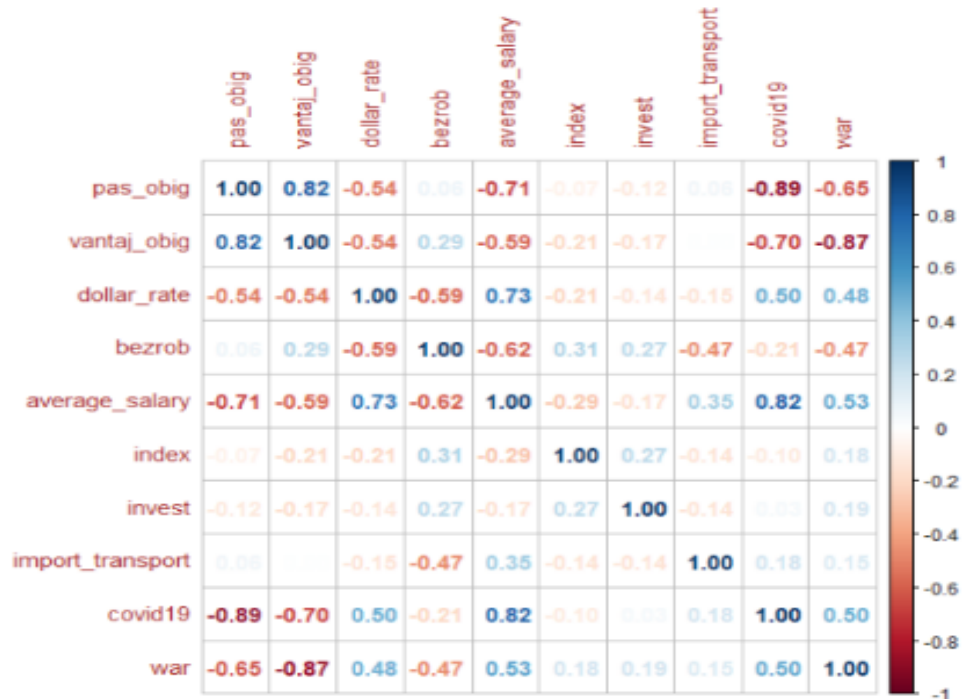


Рисунок Б.2 – Кореляційна матриця

Джерело: розроблено автором на основі [33-35]

```

Call:
lm(formula = pas_obig ~ dollar_rate + bezrob + average_salary +
    index + invest + import_transport, data = data_pas_van)

Residuals:
    Min       1Q   Median       3Q      Max
-6608.8 -2189.5   374.4  2413.2  5888.7

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  83335.1925  20492.5111   4.067 0.000334 ***
dollar_rate   -62.3486    270.9656  -0.230 0.819631
bezrob        -42.4445     13.6560  -3.108 0.004191 **
average_salary -1.9949     0.3152  -6.329 6.49e-07 ***
index        -326.2517    153.4254  -2.126 0.042100 *
invest        -0.6155     0.5190  -1.186 0.245313
import_transport 10.9427     10.4580   1.046 0.304041
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3611 on 29 degrees of freedom
Multiple R-squared:  0.8019,    Adjusted R-squared:  0.761
F-statistic: 19.57 on 6 and 29 DF,  p-value: 5.403e-09

```

Рисунок Б.2 – Оцінка значимості параметрів моделі пасажирообігу з кількісними змінними

Джерело: розроблено автором на основі [33-35]

```

Call:
lm(formula = pas_obig ~ bezrob + import_transport + covid19 +
    war, data = data_pas_van)

Residuals:
    Min       1Q   Median       3Q      Max
-5375.7  -813.9   378.8  1276.5  4908.5

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  26574.291   3609.756   7.362 2.74e-08 ***
bezrob       -17.698     5.478   -3.231 0.00292 **
import_transport 10.550     3.690   2.859 0.00753 **
covid19     -11936.627    863.420 -13.825 8.46e-15 ***
war         -8750.932    1440.809  -6.074 9.95e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2085 on 31 degrees of freedom
Multiple R-squared:  0.9295,    Adjusted R-squared:  0.9203
F-statistic: 102.1 on 4 and 31 DF,  p-value: < 2.2e-16

```

Рисунок Б.3 – Значимість параметрів оптимальної моделі пасажирообігу

Джерело: розроблено автором на основі [33-35]

```
Call:
lm(diag(mod = line_reg_v_multiple)
```

All Individual Multicollinearity Diagnostics Result

	VIF	TOL	Wi	Fi	Leamer	CVIF	Klein	IND1	IND2
dollar_rate	6.3859	0.1566	32.3152	41.7405	0.3957	13.0426	1	0.0261	1.5286
bezrob	3.4499	0.2899	14.6993	18.9866	0.5384	7.0461	1	0.0483	1.2870
average_salary	4.3803	0.2283	20.2817	26.1972	0.4778	8.9463	1	0.0380	1.3986
index	1.1825	0.8457	1.0947	1.4140	0.9196	2.4150	0	0.1410	0.2796
invest	1.1240	0.8897	0.7438	0.9608	0.9432	2.2956	0	0.1483	0.1999
import_transport	3.5808	0.2793	15.4847	20.0011	0.5285	7.3134	1	0.0465	1.3062

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

dollar_rate , bezrob , invest , import_transport , coefficient(s) are non-significant may be due to multicollinearity

R-square of y on all x: 0.5837

* use method argument to check which regressors may be the reason of collinearity

Рисунок Б.4 – Індивідуальна перевірка факторів моделі вантажообігу на мультиколінеарність

Джерело: розроблено автором на основі [33-35]

```
Call:
lm(formula = vantaj_obig ~ covid19 + war, data = data_pas_van)
```

Residuals:

Min	1Q	Median	3Q	Max
-13251.1	-4008.3	361.6	3522.7	13791.4

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	82900	1149	72.155	< 2e-16 ***
covid19	-10436	2298	-4.541	7.08e-05 ***
war	-41217	3040	-13.559	4.80e-15 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5629 on 33 degrees of freedom

Multiple R-squared: 0.8506, Adjusted R-squared: 0.8416

F-statistic: 93.96 on 2 and 33 DF, p-value: 2.375e-14

Рисунок Б.5 – Значимість параметрів моделі вантажообігу з якісними змінними

Джерело: розроблено автором на основі [33-35]

```

Call:
lmcdiag(mod = line_reg_v_bin)

All Individual Multicollinearity Diagnostics Result

      VIF   TOL   Wi Fi Leamer  CVIF Klein  IND1 IND2
covid19 1.037 0.9643 1.2593 Inf  0.982 0.6905   0 0.0284   1
war      1.037 0.9643 1.2593 Inf  0.982 0.6905   0 0.0284   1

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: 0.8506

* use method argument to check which regressors may be the reason of collinearity
=====

```

Рисунок Б.6 – Індивідуальна перевірка на мультиколінеарність моделі вантажообігу з якісними змінними

Джерело: розроблено автором на основі [33-35]

```

Call:
lm(formula = vantaj_obig ~ import_transport + covid19 + war,
    data = data_pas_van)

Residuals:
    Min       1Q   Median       3Q      Max
-11207.6  -2341.0   -230.3    2752.9   12650.5

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  70466.985   4614.557   15.271 3.00e-16 ***
import_transport  22.152     8.007    2.767 0.00933 **
covid19     -11166.527   2112.785   -5.285 8.67e-06 ***
war         -42529.532   2813.298  -15.117 3.99e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5135 on 32 degrees of freedom
Multiple R-squared:  0.8795,    Adjusted R-squared:  0.8682
F-statistic: 77.82 on 3 and 32 DF,  p-value: 8.643e-15

```

Рисунок Б.7 – Перевірка значимості параметрів оптимальної моделі вантажообігу

Джерело: розроблено автором на основі [33-35]

```
Call:
lmcdiag(mod = line_reg_v_opt)
```

All Individual Multicollinearity Diagnostics Result

	VIF	TOL	Wi	Fi	Leamer	CVIF	Klein	IND1	IND2
import_transport	1.0384	0.9630	0.6332	1.3049	0.9813	0.5579	0	0.0584	0.7348
covid19	1.0535	0.9492	0.8830	1.8195	0.9743	0.5661	0	0.0575	1.0099
war	1.0674	0.9369	1.1119	2.2913	0.9679	0.5735	0	0.0568	1.2552

```
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: 0.8795

* use method argument to check which regressors may be the reason of collinearity

=====

Рисунок Б.8 – Аналіз мультиколінеарності оптимальної моделі вантажообігу

Джерело: розроблено автором на основі [33-35]

.model <chr>	date <S3: yearquarter>	pas_obig <S3: distribution>	.mean <dbl>
arima113	2023 Q1	N(8345, 1.9e+07)	8344.700
arima113	2023 Q2	N(8697, 3.5e+07)	8697.152
arima113	2023 Q3	N(9055, 4.4e+07)	9055.193
arima113	2023 Q4	N(8949, 4.9e+07)	8948.532
arima311	2023 Q1	N(8237, 1.9e+07)	8237.467
arima311	2023 Q2	N(8613, 3.5e+07)	8613.073
arima311	2023 Q3	N(8660, 4.2e+07)	8660.122
arima311	2023 Q4	N(8597, 4.7e+07)	8597.269
arima410100	2023 Q1	N(7608, 1.7e+07)	7607.883
arima410100	2023 Q2	N(9150, 3e+07)	9150.187
arima410100	2023 Q3	N(11113, 3.8e+07)	11113.189
arima410100	2023 Q4	N(10211, 4.3e+07)	10210.905
arima211200	2023 Q1	N(7828, 1.7e+07)	7828.363
arima211200	2023 Q2	N(8958, 3e+07)	8957.653
arima211200	2023 Q3	N(10505, 3.7e+07)	10504.706
arima211200	2023 Q4	N(9143, 4.2e+07)	9143.221
search	2023 Q1	N(7369, 1.7e+07)	7369.090
search	2023 Q2	N(8325, 3.4e+07)	8324.890
search	2023 Q3	N(9675, 5.1e+07)	9675.107
search	2023 Q4	N(8217, 6.8e+07)	8217.057

Рисунок Б.9 – Прогноз за моделями ARIMA для пасажирообігу

Джерело: розроблено автором на основі [33-35]

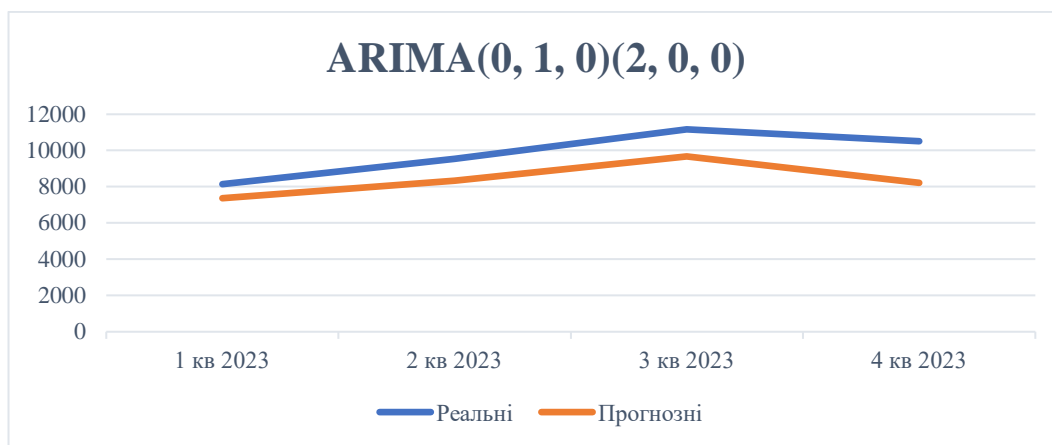


Рисунок Б.10 – Прогноз за моделлю ARIMA (0, 1, 0) (2, 0, 0) для пасажиробігу

Джерело: розроблено автором на основі [33-35]

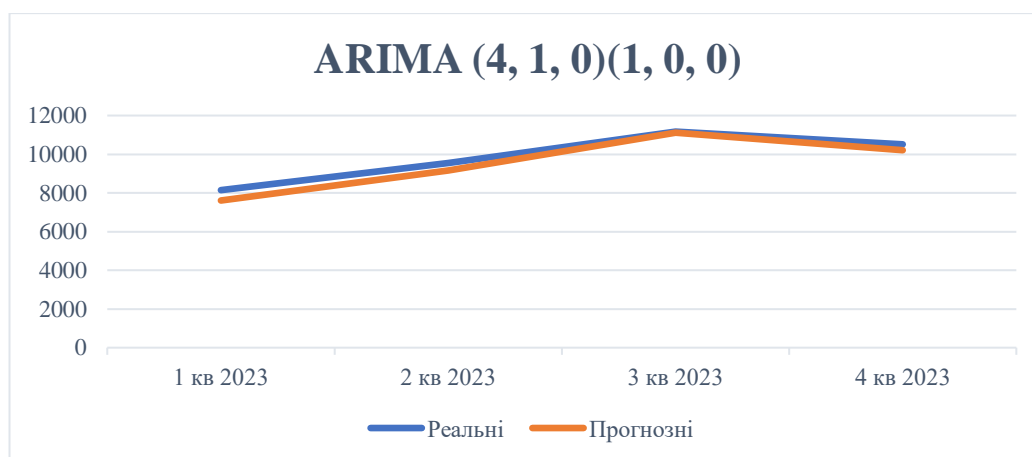


Рисунок Б.11 – Прогноз за моделлю ARIMA (4, 1, 0) (1, 0, 0) для пасажиробігу

Джерело: розроблено автором на основі [33-35]

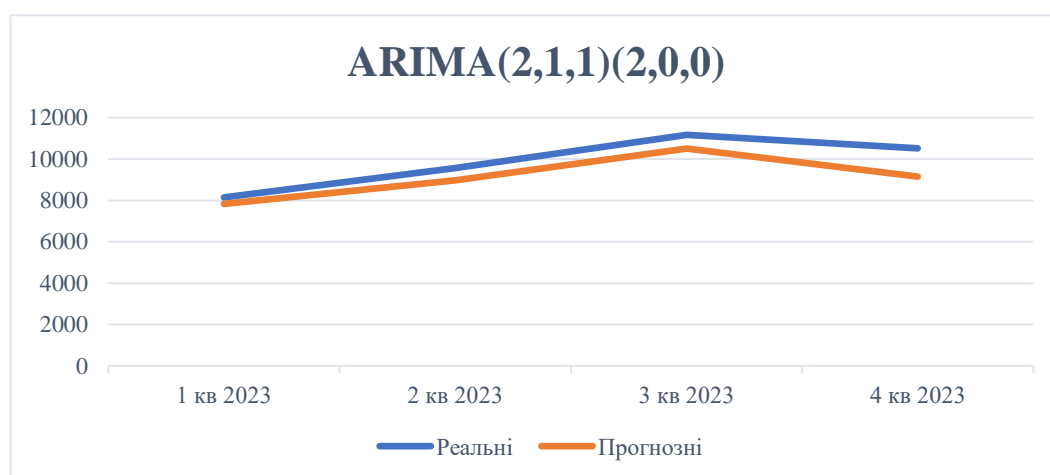


Рисунок Б.12 – Прогноз за моделлю ARIMA (2, 1, 1) (2, 0, 0) для пасажиробігу

Джерело: розроблено автором на основі [33-35]

.model <chr>	date <S3: yearquarter>	vantaj_obig <S3: distribution>	.mean <dbl>
arima210	2023 Q1	N(36143, 6.7e+07)	36143.38
arima210	2023 Q2	N(36440, 1.4e+08)	36439.66
arima210	2023 Q3	N(36480, 1.8e+08)	36480.27
arima210	2023 Q4	N(36427, 2.3e+08)	36427.49
arima310	2023 Q1	N(41604, 6.4e+07)	41603.98
arima310	2023 Q2	N(41287, 1.2e+08)	41287.47
arima310	2023 Q3	N(40906, 1.6e+08)	40906.20
arima310	2023 Q4	N(39469, 1.8e+08)	39469.02
arima211	2023 Q1	N(37946, 6.8e+07)	37945.78
arima211	2023 Q2	N(38812, 1.4e+08)	38812.15
arima211	2023 Q3	N(38766, 1.8e+08)	38765.63
arima211	2023 Q4	N(38570, 2.1e+08)	38569.88
arima310001	2023 Q1	N(36311, 6.2e+07)	36310.72
arima310001	2023 Q2	N(31148, 1.3e+08)	31148.36
arima310001	2023 Q3	N(31137, 1.8e+08)	31137.33
arima310001	2023 Q4	N(29844, 2.1e+08)	29843.66
search_v	2023 Q1	N(21396, 4.4e+07)	21396.25
search_v	2023 Q2	N(22383, 8.2e+07)	22383.30
search_v	2023 Q3	N(26903, 1.1e+08)	26903.31
search_v	2023 Q4	N(27282, 1.4e+08)	27282.26

Рисунок Б.13 – Прогнозовані значення вантажообігу за моделями ARIMA

Джерело: розроблено автором на основі [33-35]



Рисунок Б.14 – Прогноз вантажообігу за моделлю ARIMA (1, 0, 0) (1, 1, 0)

Джерело: розроблено автором на основі [33-35]

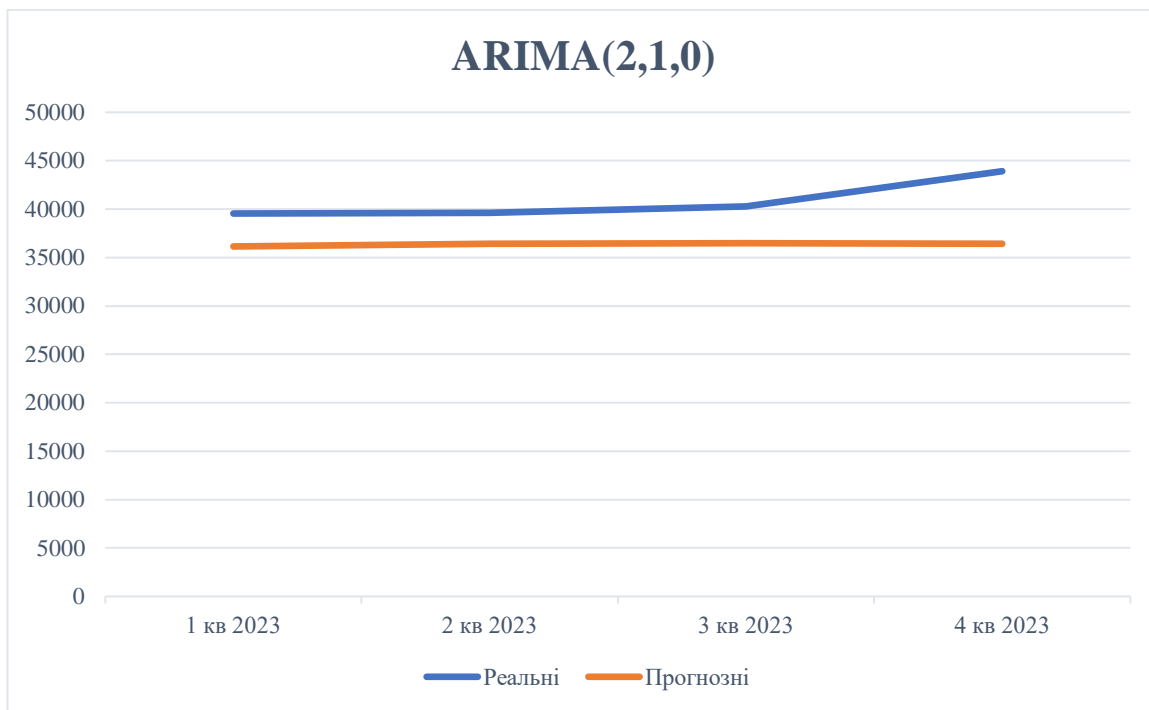


Рисунок Б.15 – Прогнозовані значення вантажообігу за моделлю ARIMA (2, 1, 0)

Джерело: розроблено автором на основі [33-35]

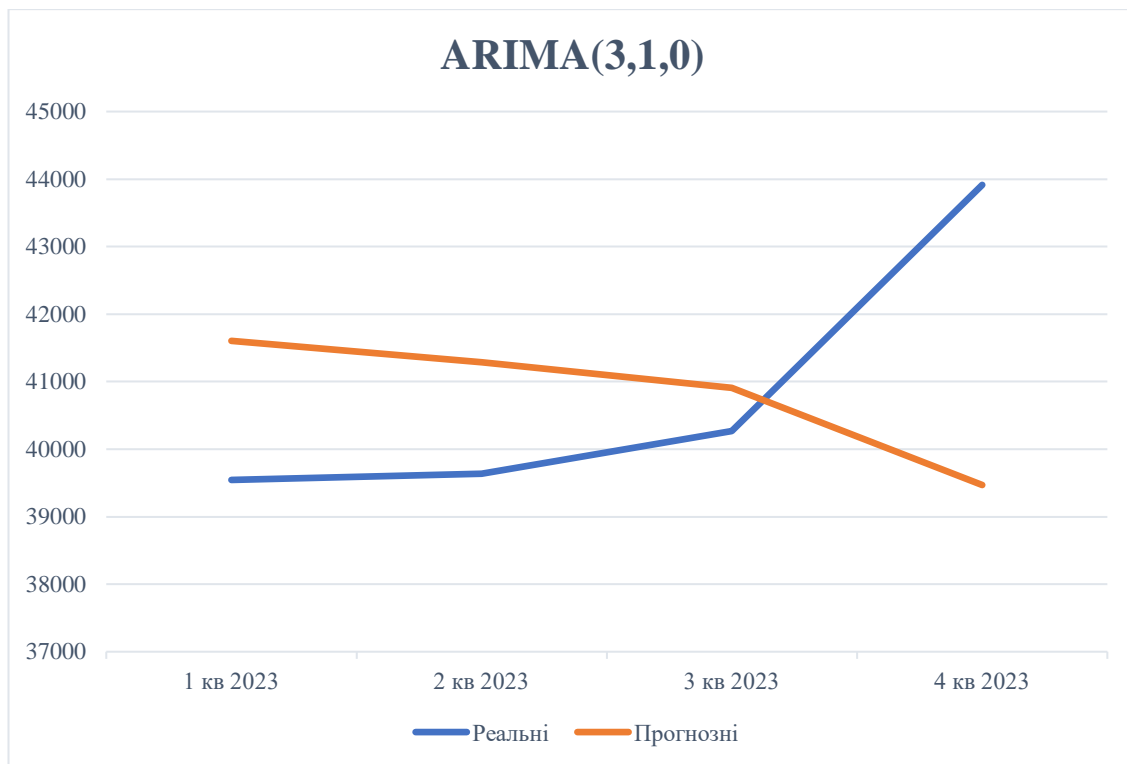


Рисунок Б.16 – Прогноз за моделлю ARIMA (3, 1, 0) для вантажообігу

Джерело: розроблено автором на основі [33-35]



Рисунок Б.17 – Графік прогнозу «Аркуш прогнозу» для вантажообігу

Джерело: розроблено автором на основі [33-35]

Дата	Вантажообіг (млн.)	Прогноз (Вантажо)	Низька ймовірність	Висока ймовірність
01.01.2014	86842			
01.04.2014	91864,6			
01.07.2014	78798,3			
01.10.2014	77893			
01.01.2015	69648,7			
01.04.2015	78867,2			
01.07.2015	83294,9			
01.10.2015	84069,1			
01.01.2016	73603,6			
01.04.2016	79357,8			
01.07.2016	82319,8			
01.10.2016	88863,8			
01.01.2017	82310,6			
01.04.2017	84710,4			
01.07.2017	87523,7			
01.10.2017	88810,1			
01.01.2018	79293,6			
01.04.2018	83222,2			
01.07.2018	85520,4			
01.10.2018	83820			
01.01.2019	80862,9			
01.04.2019	87195,4			
01.07.2019	84621,5			
01.10.2019	86282,7			
01.01.2020	68464,1			
01.04.2020	66307,4			
01.07.2020	75660,7			
01.10.2020	79647,1			
01.01.2021	64522,3			
01.04.2021	72792,5			
01.07.2021	76582,3			
01.10.2021	75738,3			
01.01.2022	55474,2			
01.04.2022	36875,3			
01.07.2022	38022,3			
01.10.2022	36359,5	36359,5	36359,50	36359,50
01.01.2023		38910,68809	20865,35	56956,03
01.04.2023		32764,3667	14163,66	51365,08
01.07.2023		26618,04531	7068,93	46167,16
01.10.2023		20471,72393	-473,49	41416,93
01.01.2024		14325,40254	-8480,53	37131,33
01.04.2024		8179,081154	-16938,97	33297,14
01.07.2024		2032,759767	-25817,61	29883,13
01.10.2024		-4113,56162	-35078,26	26851,13
01.01.2025		-10259,88301	-44682,95	24163,18

Рисунок Б.18 – Таблиця прогнозу «Аркуш прогнозу» для вантажообігу

Джерело: розроблено автором на основі [33-35]



Рисунок Б.19 – Графік «Аркуш прогнозу» для пасажирообігу

Джерело: розроблено автором на основі [33-35]

Дата	Пасажирообіг (млн. пас. км)	Прогноз(Пасажи)	Низька ймовірні	Висока ймовірні
01.01.2014	24829,6			
01.04.2014	26984,3			
01.07.2014	28691,8			
01.10.2014	25835,7			
01.01.2015	21866,9			
01.04.2015	25015,8			
01.07.2015	26210,5			
01.10.2015	23943,2			
01.01.2016	22101,4			
01.04.2016	26248			
01.07.2016	28196,5			
01.10.2016	25653,4			
01.01.2017	22037,5			
01.04.2017	25542,6			
01.07.2017	27099,2			
01.10.2017	24729,3			
01.01.2018	22618,7			
01.04.2018	27641			
01.07.2018	28569,3			
01.10.2018	25617			
01.01.2019	23362,4			
01.04.2019	28410,7			
01.07.2019	29428,5			
01.10.2019	25973,4			
01.01.2020	18995,5			
01.04.2020	4023,9			
01.07.2020	14119,2			
01.10.2020	11896,2			
01.01.2021	12230,7			
01.04.2021	15982,4			
01.07.2021	19345,1			
01.10.2021	15176			
01.01.2022	9634,3			
01.04.2022	5691,3			
01.07.2022	8102,5			
01.10.2022	7714,7	7714,7	7714,70	7714,70
01.01.2023		7178,205566	-1382,65	15739,06
01.04.2023		6668,349812	-2906,81	16243,51
01.07.2023		6158,494057	-4336,87	16653,86
01.10.2023		5648,638303	-5695,75	16993,03
01.01.2024		5138,782548	-6998,42	17275,98
01.04.2024		4628,926794	-8255,25	17513,10
01.07.2024		4119,071039	-9473,81	17711,95
01.10.2024		3609,215284	-10659,81	17878,24
01.01.2025		3099,35953	-11817,68	18016,40

Рисунок Б.20 – Прогнозні значення «Аркуш прогнозу» для пасажирообігу

Джерело: розроблено автором на основі [33-35]

ДОДАТОК В

```

library(dplyr) # попередня обробка даних
library(corrplot) # кореляційна матриця
library(car) # перевірка на автокореляцію
library(mctest) # перевірка на мультиколінеарність
library(ppcor) # кореляційна матриця
library(GGally) # ggpairs
library(lmtest) # тест на гетероскедастичність
library(broom)
library(factoextra)
library(cluster)
Завантажуємо датасети
library(readxl)
data_pas_van <- read_excel("D:/RStudio/data_pas_van.xlsx") # для регресійного аналізу
data_cluster <- read_excel("D:/RStudio/data_cluster.xlsx") # кластерний пасажирообіг з врахуванням Києва
amount_v_cl <- read_excel("D:/RStudio/amount_v_cl.xlsx") # кластерний вантажообіг з врахуванням Києва
amount_pas_cl <- read_excel("D:/RStudio/amount_pas_cl.xlsx") # кластерний пасажирообіг без
am_pas_cl <- read_excel("D:/RStudio/am_pas_cl.xlsx") # кластерний вантажообіг без
# Кластерний аналіз
cluster_d <- data_cluster %>%
  dplyr::select('2016y', '2017y', '2018y', '2019y', '2020y', '2021y')
d <- dist(scale(cluster_d), method = "manhattan")
hc <- hclust(d, method = "ward.D2")
plot(hc)
# Ієрархічний для вантажообігу 2016-2019
am_vantaj_cl <- data_cluster %>%
  dplyr::select(geo, '2016y', '2017y', '2018y', '2019y')
vantaj_cl <- data_cluster %>%
  dplyr::select('2016y', '2017y', '2018y', '2019y')
d_v <- dist(scale(vantaj_cl), method = "manhattan")
hc_v <- hclust(d_v, method = "ward.D2")
plot(hc_v)
clus_v <- scale(vantaj_cl)
fviz_nbclust(clus_v, kmeans, method = "wss")
gap_stat_v <- clusGap(vantaj_cl,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_v)
# метод k-середніх
set.seed(1)
km_v <- kmeans(vantaj_cl, centers = 4, nstart = 25)
km_v
fviz_cluster(km_v, data = vantaj_cl)
aggregate(clus_v, by=list(cluster=km_v$cluster), mean)
final_data_v <- cbind(cluster = km_v$cluster, am_vantaj_cl)
final_data_v
# Ієрархічний для вантажообігу 2020-2021
am_vantaj_cl_cov <- data_cluster %>%
  dplyr::select(geo, '2020y', '2021y')
vantaj_cl_cov <- data_cluster %>%
  dplyr::select('2020y', '2021y')
d_v_cov <- dist(scale(vantaj_cl_cov), method = "manhattan")
hc_v_cov <- hclust(d_v_cov, method = "ward.D2")

```

```

plot(hc_v_cov)
clus_v_cov <- scale(vantaj_cl_cov)
fviz_nbclust(clus_v_cov, kmeans, method = "wss")
gap_stat_v_cov <- clusGap(vantaj_cl_cov,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_v_cov)
# метод k-середніх
set.seed(1)
km_v_cov <- kmeans(vantaj_cl_cov, centers = 4, nstart = 25)
km_v_cov
fviz_cluster(km_v_cov, data = vantaj_cl_cov)
aggregate(clus_v_cov, by=list(cluster=km_v_cov$cluster), mean)
final_data_v_cov <- cbind(cluster = km_v_cov$cluster, am_vantaj_cl_cov)
final_data_v_cov
# Ієрархічний для вантажообігу 2016-2019 без Києва
am_v_cl <- amount_v_cl %>%
  dplyr::select(geo, '2016y', '2017y', '2018y', '2019y')
v_cl <- amount_v_cl %>%
  dplyr::select('2016y', '2017y', '2018y', '2019y')
d_vnk <- dist(scale(v_cl), method = "manhattan")
hc_vnk <- hclust(d_vnk, method = "ward.D2")
plot(hc_vnk)
clus_vnk <- scale(v_cl)
fviz_nbclust(clus_vnk, kmeans, method = "wss")
gap_stat_vnk <- clusGap(v_cl,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_vnk)
# метод k-середніх
set.seed(1)
km_vnk <- kmeans(v_cl, centers = 3, nstart = 25)
km_vnk
fviz_cluster(km_vnk, data = v_cl)
aggregate(clus_vnk, by=list(cluster=km_vnk$cluster), mean)
final_data_vnk <- cbind(cluster = km_vnk$cluster, am_v_cl)
final_data_vnk
# Ієрархічний для вантажообігу 2020-2021
am_v_cl_cov <- amount_v_cl %>%
  dplyr::select(geo, '2020y', '2021y')
v_cl_cov <- amount_v_cl %>%
  dplyr::select('2020y', '2021y')
d_vnk_cov <- dist(scale(v_cl_cov), method = "manhattan")
hc_vnk_cov <- hclust(d_vnk_cov, method = "ward.D2")
plot(hc_vnk_cov)
clus_vnk_cov <- scale(v_cl_cov)
fviz_nbclust(clus_vnk_cov, kmeans, method = "wss")
gap_stat_vnk_cov <- clusGap(v_cl_cov,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,

```

```

    B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_vnk_cov)
# метод k-середніх
set.seed(1)
km_vnk_cov <- kmeans(v_cl_cov, centers = 3, nstart = 25)
km_vnk_cov
fviz_cluster(km_vnk_cov, data = v_cl_cov)
aggregate(clus_vnk_cov, by=list(cluster=km_vnk_cov$cluster), mean)
final_data_vnk_cov <- cbind(cluster = km_vnk_cov$cluster, am_v_cl_cov)
final_data_vnk_cov
# Ієрархічний для кількості пасажирів 2016-2019
amount_pas_cl_1 <- amount_pas_cl %>%
  dplyr::select(geo, '2016y', '2017y', '2018y', '2019y')
pas_cl <- amount_pas_cl %>%
  dplyr::select('2016y', '2017y', '2018y', '2019y')
d_p <- dist(scale(pas_cl), method = "manhattan")
hc_p <- hclust(d_p, method = "ward.D2")
plot(hc_p)
clus_p <- scale(pas_cl)
fviz_nbclust(clus_p, kmeans, method = "wss")
gap_stat_p <- clusGap(pas_cl,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_p)
# метод k-середніх
set.seed(1)
km_p <- kmeans(pas_cl, centers =5, nstart = 25)
km_p
fviz_cluster(km_p, data = pas_cl)
aggregate(clus_p, by=list(cluster=km_p$cluster), mean)
final_data_p <- cbind(cluster = km_p$cluster, amount_pas_cl_1)
final_data_p
# Ієрархічний для кількості пасажирів 2020-2021
am_pas_cl_1 <- amount_pas_cl %>%
  dplyr::select(geo, '2020y', '2021y')
pas_cl_cov <- amount_pas_cl %>%
  dplyr::select('2020y', '2021y')
d_p_cov <- dist(scale(pas_cl_cov), method = "manhattan")
hc_p_cov <- hclust(d_p_cov, method = "ward.D2")
plot(hc_p_cov)
clus_p_cov <- scale(pas_cl_cov)
fviz_nbclust(clus_p_cov, kmeans, method = "wss")
gap_stat_p_cov <- clusGap(pas_cl_cov,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_p_cov)
# метод k-середніх
set.seed(1)
km_p_cov <- kmeans(pas_cl_cov, centers =5, nstart = 25)
km_p_cov

```

```

fviz_cluster(km_p_cov, data = pas_cl_cov)
aggregate(clus_p_cov, by=list(cluster=km_p_cov$cluster), mean)
final_data_p_cov <- cbind(cluster = km_p_cov$cluster, am_pas_cl_1)
final_data_p_cov
# Ієрархічний для кількості пасажирів 2016-2019
am_pas_cl_nk <- am_pas_cl %>%
  dplyr::select(geo, '2016y', '2017y', '2018y', '2019y')
pas_cl_nk <- am_pas_cl %>%
  dplyr::select('2016y', '2017y', '2018y', '2019y')
d_pnk <- dist(scale(pas_cl_nk), method = "manhattan")
hc_pnk <- hclust(d_pnk, method = "ward.D2")
plot(hc_pnk)
clus_pnk <- scale(pas_cl_nk)
fviz_nbclust(clus_pnk, kmeans, method = "wss")
gap_stat_pnk <- clusGap(pas_cl_nk,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_pnk)
# метод k-середніх
set.seed(1)
km_pnk <- kmeans(pas_cl_nk, centers =4, nstart = 25)
km_pnk
fviz_cluster(km_pnk, data = pas_cl_nk)
aggregate(clus_pnk, by=list(cluster=km_pnk$cluster), mean)
final_data_pnk <- cbind(cluster = km_pnk$cluster, am_pas_cl_nk)
final_data_pnk
# Ієрархічний для кількості пасажирів 2020-2021
am_pas_cl_nk_cov <- am_pas_cl %>%
  dplyr::select(geo, '2020y', '2021y')
pas_cl_nk_cov <- am_pas_cl %>%
  dplyr::select('2020y', '2021y')
d_pnk_cov <- dist(scale(pas_cl_nk_cov), method = "manhattan")
hc_pnk_cov <- hclust(d_pnk_cov, method = "ward.D2")
plot(hc_pnk_cov)
clus_pnk_cov <- scale(pas_cl_nk_cov)
fviz_nbclust(clus_pnk_cov, kmeans, method = "wss")
gap_stat_pnk_cov <- clusGap(pas_cl_nk_cov,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat_pnk_cov)
# метод k-середніх
set.seed(1)
km_pnk_cov <- kmeans(pas_cl_nk_cov, centers =5, nstart = 25)
km_pnk_cov
fviz_cluster(km_pnk_cov, data = pas_cl_nk_cov)
aggregate(clus_pnk_cov, by=list(cluster=km_pnk_cov$cluster), mean)
final_data_pnk_cov <- cbind(cluster = km_pnk_cov$cluster, am_pas_cl_nk_cov)
final_data_pnk_cov
# Метод k-середніх
clus <- scale(cluster_d)
fviz_nbclust(clus, kmeans, method = "wss")

```

```

# Метод k-середніх
gap_stat<-clusGap(cluster_d,
  FUN = kmeans,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat)
gap_stat1<-clusGap(cluster_d,
  FUN = pam,
  nstart = 20,
  K.max = 7,
  B = 20)
# графік оптимальної кількості кластерів
fviz_gap_stat(gap_stat1)
# метод k-середніх
set.seed(1)
km<-kmeans(cluster_d, centers = 5, nstart = 25)
km
fviz_cluster(km, data = cluster_d)
aggregate(clus, by=list(cluster=km$cluster), mean)
final_data <- cbind(data_cluster, cluster = km$cluster)
final_data
# Переглянемо його
glimpse(data_pas_van)
# Перевіримо кореляцію в даних
data_correl <- data_pas_van[,2:11]
corrplot(cor(data_correl), method = "number", rect.col = "", tl.col = "brown")
# Побудуємо модель, яка включає всі змінні
line_reg_multiple<-lm(pas_obig ~ dollar_rate+bezrob+average_salary+index+invest+import_transport,
  data=data_pas_van)
line_reg_multiple
summary(line_reg_multiple)
# Перевіримо на автокореляцію
durbinWatsonTest((line_reg_multiple))
# Перевіримо на мультиколінеарність
omcdiag(line_reg_multiple)
imcdiag(line_reg_multiple)
# Тест на гетероскедастичність
bptest(line_reg_multiple,studentize=FALSE)
# Створимо модель тільки з якісними факторами
line_reg_bin<-lm(pas_obig ~ covid19+war, data=data_pas_van)
line_reg_bin
# Перевірка значимості
summary(line_reg_bin)
durbinWatsonTest(line_reg_bin)
omcdiag(line_reg_bin)
imcdiag(line_reg_bin)
bptest(line_reg_bin)
# Переглянемо модель тільки зі значимими показниками
line_reg_opt<-lm(pas_obig ~ bezrob+import_transport+covid19+war, data=data_pas_van)
line_reg_opt
summary(line_reg_opt)
# Перевірка на мультиколінеарність
durbinWatsonTest((line_reg_opt))
omcdiag(line_reg_opt)
imcdiag(line_reg_opt)

```

```

# Тест на гетероскедастичність
bptest(line_reg_opt, studentize=FALSE)
# Побудовано моделі для вантажообігу
line_reg_v_multiple <- lm(vantaj_obig ~ dollar_rate+bezrob+average_salary+index+invest+import_transport,
data = data_pas_van)
line_reg_v_multiple
summary(line_reg_v_multiple)
durbinWatsonTest(line_reg_v_multiple)
omcdiag(line_reg_v_multiple)
imcdiag(line_reg_v_multiple)
bptest(line_reg_v_multiple)
# Модель з якісними
line_reg_v_bin<-lm(vantaj_obig ~ covid19+war, data=data_pas_van)
line_reg_v_bin
summary(line_reg_v_bin)
durbinWatsonTest(line_reg_v_bin)
omcdiag(line_reg_v_bin)
imcdiag(line_reg_v_bin)
bptest(line_reg_v_bin)
# Оптимальна модель для вантажообігу
line_reg_v_opt<-lm(vantaj_obig ~ import_transport+covid19+war, data=data_pas_van)
line_reg_v_opt
summary(line_reg_v_opt)
durbinWatsonTest(line_reg_v_opt)
omcdiag(line_reg_v_opt)
imcdiag(line_reg_v_opt)
bptest(line_reg_v_opt)
# Моделі ARIMA
library(fpp3)
# задаємо формат дати
data_pas_van$date <- as.Date(data_pas_van$date)
data_pas_van_ts <- as_tsibble(data_pas_van, index = date)
data_pas_van_ts
data_pas_van_ts <- data_pas_van_ts %>%
  mutate(date = yearquarter(date))
# переглянемо датасет
data_pas_van_ts
# ARIMA
data_pas_van_ts %>%
  gg_tsdisplay(pas_obig, plot_type='partial')
# Тест KPSS
data_pas_van_ts %>%
  features(pas_obig, unitroot_kpss)
data_pas_van_ts %>%
  features(pas_obig, unitroot_ndiffs)
# Диференціювання
data_pas_van_ts %>%
  features(difference(pas_obig), unitroot_kpss)
data_pas_van_ts %>%
  features(difference(pas_obig), unitroot_ndiffs)
pas_fit <- data_pas_van_ts %>%
  model(arima113 = ARIMA(pas_obig ~ pdq(1,1,3)+ PDQ(0,0,0)),
        arima311 = ARIMA(pas_obig ~ pdq(3,1,1)+ PDQ(0,0,0)),
        arima410100 = ARIMA(pas_obig ~ pdq(4,1,0)+ PDQ(1,0,0)),
        arima211200 = ARIMA(pas_obig ~ pdq(2,1,1) + PDQ(2,0,0)),
        search = ARIMA(pas_obig, stepwise=FALSE))
glance(pas_fit) %>% arrange(AICc)

```

```

fc_pas <- pas_fit %>%
  forecast(h=4)
fc_pas
search_p <- data_pas_van_ts %>%
  model(search = ARIMA(pas_obig, stepwise=FALSE))
arima410100 <- data_pas_van_ts %>%
  model(arima410100 = ARIMA(pas_obig ~ pdq(4,1,0)+ PDQ(1,0,0)))
arima211200 <- data_pas_van_ts %>%
  model(arima211200 = ARIMA(pas_obig ~ pdq(2,1,1) + PDQ(2,0,0)))
report(search_p)
report(arima410100)
report(arima211200)
pas_fit %>%
  forecast(h=4) %>%
  filter(.model=='arima410100') %>%
  autoplot(data_pas_van_ts)
pas_fit_f %>%
  forecast(h=4) %>%
  filter(.model=='search') %>%
  autoplot(data_pas_van_ts)
augment(pas_fit) %>%
  features(.innov, ljung_box, lag = 4, dof = 3)
# Вантажообіг
data_pas_van_ts %>%
  gg_tsdisplay(vantaj_obig, plot_type='partial')
# Тест KPSS
data_pas_van_ts %>%
  features(vantaj_obig, unitroot_kpss)
data_pas_van_ts %>%
  features(vantaj_obig, unitroot_ndiffs)
# Диференціювання
data_pas_van_ts %>%
  features(difference(vantaj_obig), unitroot_kpss)
data_pas_van_ts %>%
  features(difference(vantaj_obig), unitroot_ndiffs)
van_fit <- data_pas_van_ts %>%
  model(arima210 = ARIMA(vantaj_obig ~ pdq(2,1,0)+ PDQ(0,0,0)),
        arima310 = ARIMA(vantaj_obig ~ pdq(3,1,0)+ PDQ(0,0,0)),
        arima211 = ARIMA(vantaj_obig ~ pdq(2,1,1)+ PDQ(0,0,0)),
        arima310001 = ARIMA(vantaj_obig ~ pdq(3,1,0) + PDQ(0,0,1)),
        search_v = ARIMA(vantaj_obig, stepwise=FALSE))
glance(van_fit) %>% arrange(AICc)
fc_cv_v <- van_fit %>%
  forecast(h=4)
fc_cv_v
search_v <- data_pas_van_ts %>%
  model(search_v = ARIMA(vantaj_obig, stepwise=FALSE))
arima210 <- data_pas_van_ts %>%
  model(arima210 = ARIMA(vantaj_obig ~ pdq(2,1,0)+ PDQ(0,0,0)))
arima310 <- data_pas_van_ts %>%
  model(arima310 = ARIMA(vantaj_obig ~ pdq(3,1,0)+ PDQ(0,0,0)))
report(search_v)
report(arima210)
report(arima310)
augment(van_fit) %>%
  features(.innov, ljung_box, lag = 4, dof = 3)

```



метадані

Заголовок

Бабак_Мазур_плагіат

Автор Науковий керівник / Експерт

Бабак Мазур

підрозділ

кафедра математичного моделювання та статистики

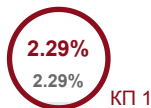
Тривога

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		1
Інтервали		0
Мікропробіли		0
Білі знаки		0
Парафрази (SmartMarks)		8

Обсяг знайдених подібностей

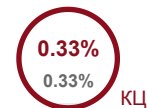
Коефіцієнт подібності визначає, який відсоток тексту по відношенню до загального обсягу тексту було знайдено в різних джерелах. Зверніть увагу, що високі значення коефіцієнта не автоматично означають плагіат. Звіт має аналізувати компетентна / уповноважена особа.

**25**

Довжина фрази для коефіцієнта подібності 2

**10267**

Кількість слів

**76791**

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)	Колір тексту
1	Дишкантиюк_стаття.docx 11/2/2021 Publishing House "Helvetica" (Видавничий дім "Гельветика")	40	0.39 %
2	http://www.visnyk.kh.ua/web/uploads/journals_pdf/1_2011.pdf	32	0.31 %
3	Моделювання та прогнозування показників ринку праці 6/4/2020 Kyiv National Economic University named after Vadym Hetman KNEU (кафедра економіко-математичного моделювання)	31	0.30 %
4	http://at.kdu.edu.ua/Files/Nauka/Konf_KrNU_2016-2.pdf	20	0.19 %

5	http://www.economyandsociety.in.ua/index.php/journal/article/download/1427/1374/	17	0.17 %
6	Моделювання задач логістики на підґрунті генетичних алгоритмів та аналізу мереж 6/3/2020 Kyiv National Economic University named after Vadym Hetman KNEU (кафедра економіко-математичного моделювання)	17	0.17 %
7	Substantiating the Choice of a Strategy for Development of Tourism Resources in the Region Pysareva Iryna V.;	16	0.16 %
8	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%AD%D0%BA%D0%BE%D0%BD%D0%BE%D0%BC%D0%B8%D0%BA%D0%B%D0%BE%D1%80%D0%B3%D0%B0%D0%BD%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D0%B8/%D0%A0%D0%B5%D0%B3%D1%83%D0%BB%D1%8F%D1%82%D0%BE%D1%80%D0%BD%D0%B0-%D0%BF%D0%BE%D0%BB%D1%96%D1%82%D0%B8%D0%BA%D0%B0-%D1%83-%D1%81%D1%84%D0%B5%D1%80%D1%96-345082.html	14	0.14 %
9	https://ndc-ipr.org/media/ndc_old/documents/Biznes_Inform_5_2015.pdf	12	0.12 %
10	Аналіз економічної діяльності малих підприємств за регіонами України 6/4/2020 Kyiv National Economic University named after Vadym Hetman KNEU (кафедра економіко-математичного моделювання)	10	0.10 %

з бази даних RefBooks (0.16 %)

ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)	
джерело: Paperity			
1	Substantiating the Choice of a Strategy for Development of Tourism Resources in the Region Pysareva Iryna V.;	16 (1)	0.16 %

з домашньої бази даних (0.56 %)

ПОРЯДКОВИЙ НОМЕР	ЗАГОЛОВОК	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)	
1	Моделювання та прогнозування показників ринку праці 6/4/2020 Kyiv National Economic University named after Vadym Hetman KNEU (кафедра економіко-математичного моделювання)	31 (1)	0.30 %
2	Моделювання задач логістики на підґрунті генетичних алгоритмів та аналізу мереж 6/3/2020 Kyiv National Economic University named after Vadym Hetman KNEU (кафедра економіко-математичного моделювання)	17 (1)	0.17 %
3	Аналіз економічної діяльності малих підприємств за регіонами України 6/4/2020 Kyiv National Economic University named after Vadym Hetman KNEU (кафедра економіко-математичного моделювання)	10 (1)	0.10 %

з програми обміну базами даних (0.39 %)

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

1	Дишкантюк_стаття.docx 11/2/2021 Publishing House "Helvetica" (Видавничий дім "Гельветика")	40 (1)	0.39 %
---	--------------------------------------------------------------------------------------------------	--------	--------

з Інтернету (1.18 %)

ПОРЯДКОВИЙ НОМЕР	ДЖЕРЕЛО URL	КІЛЬКІСТЬ ІДЕНТИЧНИХ СЛІВ (ФРАГМЕНТІВ)	
1	http://www.visnyk.kh.ua/web/uploads/journals_pdf/1_2011.pdf	32 (1)	0.31 %
2	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%AD%D0%BA%D0%BE%D0%BD%D0%BE%D0%BC%D0%B8%D0%BA%D0%B%D0%BE%D1%80%D0%B3%D0%B0%D0%BD%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D0%B8/%D0%A0%D0%B5%D0%B3%D1%83%D0%BB%D1%8F%D1%82%D0%BE%D1%80%D0%BD%D0%B0-%D0%BF%D0%BE%D0%BB%D1%96%D1%82%D0%B8%D0%BA%D0%B0-%D1%83-%D1%81%D1%84%D0%B5%D1%80%D1%96-345082.html	24 (2)	0.23 %
3	http://at.kdu.edu.ua/Files/Nauka/Konf_KrNU_2016-2.pdf	20 (1)	0.19 %
4	https://ndc-ipr.org/media/ndc_old/documents/Biznes_Inform_5_2015.pdf	18 (2)	0.18 %
5	http://www.economyandsociety.in.ua/index.php/journal/article/download/1427/1374/	17 (1)	0.17 %
6	https://www.stud24.ru/literature/specifika-modalnih-chastok-v-suchasnj/338474-1019784-page1.html	10 (1)	0.10 %

Список прийнятих фрагментів (немає прийнятих фрагментів)

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