

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ ЕКОНОМІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ СЕМЕНА КУЗНЕЦЯ**

ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

КАФЕДРА ЕКОНОМІЧНОЇ КІБЕРНЕТИКИ І СИСТЕМНОГО АНАЛІЗУ

Рівень вищої освіти	Перший (бакалаврський)
Спеціальність	Економіка
Освітня програма	Економічна кібернетика
Група	6.04.051.020.18.1

ДИПЛОМНА РОБОТА

на тему: «Моделі оцінки впливу цифрових технологій
на економіку країн»

Виконав: студент Владислав ПАВЛЕНКО

Керівник: к.е.н., доцент Світлана ПРОКОПОВИЧ

Рецензент: д.е.н., професор, перший проректор, Заклад вищої освіти "Міжнародний науково-технічний університет імені академіка Юрія Бугая"
Оксана КАРПЕНКО

Харків – 2022 рік

РЕФЕРАТ

Звіт про дипломну роботу: 137 сторінок, 3 розділи, 81 малюнок, 2 таблиці, 32 джерела, 3 додатки.

Об'єктом дослідження є процеси впливу цифрових технологій на економіку країн.

Метою роботи є побудова комплексу економіко-математичних моделей, які дозволяють оцінити, проаналізувати та спрогнозувати вплив цифрових технологій на економіку країн.

Розглянуто сучасну концепцію та стратегію розвитку цифрових технологій та їх впливу на економіку країн.

Результати дослідження можуть бути використані у розробці теоретичних положень та методичного інструментарію оцінки впливу цифрових технологій на економічну діяльність країн.

КЛЮЧОВІ СЛОВА: ЦИФРОВІ ТЕХНОЛОГІЇ, БЛОКЧЕЙН, ЦИФРОВИЙ АКТИВ, ЕКОНОМІЧНА ДІЯЛЬНІСТЬ, МОДЕЛІ ПАНЕЛЬНИХ ДАНИХ, КЛАСТЕРНИЙ АНАЛІЗ, МОДЕЛІ БАГАТОФАКТОРНОЇ РЕГРЕСІЇ.

РЕФЕРАТ

Отчет о дипломной работе: 137 страниц, 3 раздела, 81 рисунок, 2 таблицы, 32 источника, 3 приложения.

Объектом исследования являются процессы влияния цифровых технологий на экономику стран.

Целью работы является построение комплекса экономико-математических моделей, позволяющих оценить, проанализировать и спрогнозировать влияние цифровых технологий на экономику стран.

Рассмотрена современная концепция и стратегия развития цифровых технологий и их влияние в экономику стран.

Результаты исследования могут использоваться в разработке теоретических положений и методического инструментария оценки влияния цифровых технологий на экономическую деятельность стран.

КЛЮЧЕВЫЕ СЛОВА: ЦИФРОВЫЕ ТЕХНОЛОГИИ, БЛОКЧЕЙН, ЦИФРОВЫЙ АКТИВ, ЭКОНОМИЧЕСКАЯ ДЕЯТЕЛЬНОСТЬ, МОДЕЛИ ПАНЕЛЬНЫХ ДАННЫХ, КЛАСТЕРНЫЙ АНАЛИЗ, МОДЕЛИ МНОГОФАКТОРНОЙ.

ABSTRACT

Thesis report: 137 pages, 3 sections, 81 figures, 2 tables, 32 sources, 3 appendices.

The object of research is the processes of impact of digital technologies on the economy of countries.

The aim of the work is to build a set of economic and mathematical models that allow to assess, analyze and predict the impact of digital technologies on the economy.

The modern concept and strategy of digital technologies development and their impact on the economy of the countries are considered.

The results of the study can be used in the development of theoretical provisions and methodological tools for assessing the impact of digital technologies on economic activity.

KEY WORDS: DIGITAL TECHNOLOGIES, BLOCKCHAIN, DIGITAL ASSETS, ECONOMIC ACTIVITY, PANEL DATA MODELS, CLUSTER ANALYTICS, CLUSTER ANALYSIS, MODELS.

ЗМІСТ

ВСТУП	7
РОЗДІЛ 1. ЦИФРОВІ ТЕХНОЛОГІЇ. БЛОКЧЕЙН ТА ЦИФРОВІ АКТИВИ	10
1.1. Поняття «цифрові технології», «блокчейн» та «цифрові активи»	10
1.2. Аспекти цифрового активу.....	15
1.3. Властивості та завдання цифрового активу.....	20
РОЗДІЛ 2. МЕТОДИ ТА МОДЕЛІ ОЦІНКИ ВПЛИВУ ЦИФРОВИХ ТЕХНОЛОГІЙ НА ЕКОНОМІКУ КРАЇН	25
2.1. Концептуальна схема побудови моделей оцінки впливу цифрових технологій на економіку країн.....	25
2.2. Моделі панельних даних та методи їх оцінювання	29
2.3. Методи кластерного аналізу.....	40
2.4. Методи побудови багатофакторної лінійної моделі.....	48
РОЗДІЛ 3. РЕАЛІЗАЦІЯ МОДЕЛЕЙ ОЦІНКИ ТА АНАЛІЗУ ВПЛИВУ ЦИФРОВИХ ТЕХНОЛОГІЙ НА ЕКОНОМІКУ КРАЇН	58
3.1. Аналіз показників країн за допомогою моделей панельних даних.....	58
3.2. Реалізація кластерного аналізу рівня цифрового та економічного розвитку країн	72
3.3. Розробка регресійної моделі оцінки впливу цифрових технологій на економіку України та побудова прогнозу	86
ВИСНОВОК.....	107
СПИСОК ВИКОРСТАНИХ ДЖЕРЕЛ.....	109
ДОДАТОК А.....	112
ДОДАТОК Б	125
ДОДАТОК В.....	132

ВСТУП

Бурхливий розвиток науково-технічного прогресу, особливо в останні десятиліття, змінив мислення сучасної людини: перехід на цифрову техніку та поширення інформаційних технологій зробило наше життя комфортнішим.

Людство живе під час захоплюючих технологічних інновацій. Цифрові технології сприяють трансформаційним змінам. Змінюються економічні парадигми. Нові технології змінюють ринки продуктів, фактори виробництва, бізнес та роботу. Останні досягнення в галузі штучного інтелекту та пов'язаних з ним інновацій розширюють межі цифрової революції. Цифрова трансформація прискорюється після пандемії COVID-19. Майбутнє настає швидше, ніж очікувалося.

Визначення цифрових технологій включає всі електронні засоби, автоматичні системи, технологічні пристрої та ресурси, які генерують, обробляють та зберігають інформацію. Новітні технології відкрили шлях до багатofункціональних інструментів, таких як розумний годинник і смартфон. Комп'ютери та ноутбуки поступово стають швидшими, зручнішими та потужнішими, ніж раніше. Завдяки всім цим інноваціям, технології також зробили наше життя простішим, швидшим, зручним, комфортним, точним і приємнішим.

Завдяки технологіям наш світ може стати більш рівноправним, миролюбним та справедливим. Розвиток цифрових технологій може сприяти підтримці та прискоренню досягнення цілей у сфері сталого розвитку – починаючи з ліквідації бідності, зниження коефіцієнтів смертності та закінчуючи цілями заохочення стійкого фермерського виробництва й забезпечення гідної роботи, а також досягнення загальної грамотності. Водночас технології можуть ставити під загрозу недоторканність приватного життя, підривати безпеку та посилювати нерівність. Використання технологій позначається як у здійсненні правами людини, так і на забезпеченні свободи її дій. Як і попереднім поколінням, людству належить зробити вибір щодо того, як використовувати нові технології та контролювати їх розвиток.

Нові технології мають великі перспективи. Вони створюють нові шляхи та можливості для більш процвітаючого майбутнього. Але вони також ставлять нові виклики економіці світу. Хоча цифрові технології вразили блиском і майстерністю своїх застосувань, вони поки що не повною мірою принесли очікувані дивіденди у вищому зростанні продуктивності. Дійсно, зростання сукупної продуктивності сповільнилося за останні кілька десятиліть у багатьох країнах. Відповідно, економічне зростання має тенденцію до зниження.

Водночас нерівність доходів і пов'язані з цим диспропорції зросли, особливо в країнах з розвинутою економікою, що викликало соціальне та політичне невдоволення. У різних сферах економіки спостерігається нерівномірна участь у нових можливостях, створених цифровою трансформацією. Багато з них залишаються позаду в галузях, робочій силі та різних сегментах суспільства.

Однією з важливих причин цих результатів є те, що органи влади країн та інституції повільно пристосовувалися до змін, що розгортаються. Щоб реалізувати перспективи сучасних розумних машин, держава також має бути готовою до змін. Країна має бути більш чутливішою до змін, щоб повністю охопити потенційний приріст продуктивності, економічного зростання та подолати зростання нерівності, оскільки технологічні збої створюють переможців і переможе-них.

Оскільки технології змінюють ринки, динаміку зростання та розподілу, політика має забезпечити, щоб ринки залишалися інклюзивними та підтримували широкий доступ до нових можливостей для компаній і працівників. Цифрову економіку необхідно розширити, щоб поширювати нові технології та можливості серед менших компаній і широких сегментів робочої сили.

Отже, у той час як цифрові технології дають значну віддачу від продуктивності, вони також створюють нові проблеми, оскільки змінюються виробничі процеси, джерела конкурентних переваг і ринкові структури.

Метою роботи є побудова комплексу економіко-математичних моделей, які дозволяють оцінити, проаналізувати та спрогнозувати вплив цифрових технологій на економіку країн.

Об'єктом дослідження є процеси впливу цифрових технологій на економіку країн.

Предметом дослідження є економіко-математичні методи та моделі оцінювання, аналізу та прогнозування.

Розроблені моделі дозволять надати комплексну оцінку впливу цифрових технологій, оцінити перспективи розвитку та визначити напрямки розвитку цифрових технологій для пришвидшення процесів економічного розвитку країн.

РОЗДІЛ 1. ЦИФРОВІ ТЕХНОЛОГІЇ. БЛОКЧЕЙН ТА ЦИФРОВІ АКТИВИ

1.1. Поняття «цифрові технології», «блокчейн» та «цифрові активи»

Цифрові технології проникають у всі аспекти нашого життя, такий процес отримав назву цифровізація та стає визначальною тенденцією найближчих десятиліть. Цифровізація у глобальному плані є концепцією економічної діяльності, заснованої на цифрових технологіях, що впроваджуються в різні сфери життя та виробництва. І ця концепція широко впроваджується у всіх країнах.

Для початку розділимо життя людини на складові. Людина залежить від ефективності роботи уряду, купує товари та послуги в Інтернеті та в магазинах, спілкується та використовує інформацію. Цифровізація робить всі ці процеси простішими.

Цифровізація прямо зараз застосовується в побуті, на виробництві, на роботі, у державних структурах, в бізнесі й т.д. Побутова цифровізація для більшості давно стала зрозумілою та звичною. І не йдеться про якісь нові та дорогі системи на кшталт «розумного» будинку – навіть звичайна мультиварка з управлінням по Wi-Fi, яка є у багатьох, являється яскравим прикладом побутової цифровізації. Багато людей часто використовують ще одну цифрову систему – сигналізацію. Спеціальний датчик стежить за тим, чи є в квартирі сторонні, й у разі виявлення повідомляє на пульт черговому. Виробнича цифровізація займається скороченням монотонної фізичної праці для людини, організовує й контролює трудові та виробничі процеси, забезпечує безпеку співробітників компанії. Допустимо, простий робітник йде на своє місце до верстата, але чомусь включається сирена, а самому співробітнику повідомляють про порушення техніки безпеки. Це «розумна» система відеоаналітики зауважила, що працівник увійшов до цеху без каски, подала сигнал і врятувала людині життя. "Розумні" системи за допомогою машинного зору виявляють браковані деталі, а система відеоспостереження слідкує за дотриманням техніки безпеки. Такі ж системи використову-

ються в невиробничих компаніях і навіть у невеликих офісах. Держава впроваджує цифровізацію у всі свої вертикалі. Система аналізу даних давно вже ефективно використовується у виконавчій владі. Яскравий приклад – розшук злочинців з використанням онлайн-камер на вулицях або можливість відправити звернення до потрібних інстанцій по Інтернету. Однією з головних заслуг цифровізації держави є зниження кількості паперової тяганини та бюрократії при оформленні документів. Довідки та паспорти можна замовляти через програму, там же зберігати та оновлювати всі дані.

Можна виділити кілька ключових напрямів цифрової трансформації: розробка нової цифрової бізнес-моделі; створення цифрових товарів та послуг; управління життєвим циклом продукту; автоматизований збір, зберігання та обробка інформації; використання цифрового проектування; управління виробничими процесами та мережами постачання; виконання адміністративних функцій; автоматизація ручної праці за допомогою використання роботів та електронного документообігу.

Вибір технологій для конкретного підприємства чи бізнесу залежить від того, навіщо виконується цифровізація та яких цілей необхідно досягти. На рис. 1.1 представлена взаємозалежність між інноваціями, навичками та цифровими технологіями.



Рис. 1.1. Взаємозв'язки між інноваціями, навичками та цифровими технологіями

Три вершини показують відносини між трьома динаміками. Системна взаємодія між цими трьома динаміками вимагає і зумовлює реорганізацією виробничих та інноваційних процесів як всередині компаній, так і між ними.

Єдина небезпека цифровізації – зниження важливості людини у багатьох процесах та ймовірне зникнення у майбутньому цілих професій, з якими краще справлятимуться програми [20].

Динамічний розвиток та запровадження цифрових технологій зумовили закономірну трансформацію суспільних відносин. Так, стрімкий розвиток технології блокчейн як концептуально нового виду обліку ресурсів та забезпечення найвищого ступеня прозорості, безпеки та довіри, а також активна інтеграція даної технології у різні сфери діяльності послужили відправною точкою формування нового виду соціально-економічних відносин, регулювання яких має здійснюватися у правовій площині.

Становлення та регулювання даних відносин відбувається на основі суспільних договорів, що має призвести до кардинального перетворення фундаментального принципу. Усе це свідчить про необхідності створення законодавчої бази з метою правового забезпечення нового виду соціально-економічних відносин. У зв'язку з цим виникла потреба у розгляді механізму роботи та принципів, на яких базується технологія блокчейн, для глибокого та коректного розуміння суті та практичної корисності даної технології.

Незважаючи на активне впровадження технології блокчейн та токенів блокчейну в різних сферах суспільного життя, їх можливості та переваги ще недостатньо вивчені. Серед теоретиків та практиків точаться дискусії щодо перспектив застосування технології блокчейн у фінансовій, банківській та соціальній сферах, а також у державному секторі.

Блокчейн є багатофункціональною та багаторівневою інформаційною технологією, яка призначена для надійного зберігання, обліку та передачі різної інформації. Потенціал використання даної технології охоплює всі сфери життя суспільства та має безліч сфер застосування. Блокчейн створює нові можливості з пошуку, організації, оцінки та передачі будь-яких дискретних одиниць. По суті

це нова організаційна парадигма для координації будь-якого виду діяльності [18].

Блокчейн відноситься до ключових технологій і є глобальною технологією XXI століття. Вона повинна сприяти вирішенню питань надійності та прозорості суспільних відносин, формування якісно нового довірчого рівня відносин та модифікації різноманітних систем обліку, зберігання та передачі інформації, зумовлених новими потребами та економічними ресурсами суспільства.

Властивості технології дозволяють реалізувати перераховані вище завдання за допомогою функціонування таких інструментів, як токени блокчейна. Якісно новим видом токена блокчейна є цифровий актив, який має унікальні властивості та атрибути, а також має ряд цільових завдань. При цьому властивості цифрового активу значно ширші за властивості інших токенів (блокчейна), що сприяє більш ефективній реалізації завдань самої технології.

Цифровий актив – інформаційний ресурс, похідний від права на цінність і звертається у розподіленому реєстрі у вигляді унікального ідентифікатора [2]. Запропоноване поняття «цифровий актив» представлено складовими (економічною, правовою, інформаційною, ціннісною), які взаємопов'язані та взаємозалежні.

Проаналізувавши функціональну природу блокчейна та цифрових активів, можна зробити висновок про те, що їхні властивості та функції органічно співвідносні. Так, функція обліку інформації, доступна в блокчейні, неможлива без наявності одиниці обліку, якою може бути цифровий актив. При цьому така властивість цифрового активу, як управління інформацією, нереалізована без відповідного простору – розподіленого реєстру або блокчейну. Таким чином, цифровий актив – це новий підхід до поняття токена розподіленого реєстру або блокчейна, і оскільки цифровий актив у певному сенсі можна віднести до інструменту блокчейна, його концептуальне призначення полягає в абсолютній реалізації властивостей цієї технології [10].

На сьогоднішній день поняття «цифровий актив» не має єдиного комплексного визначення, яке повністю розкривало б суть терміна, що значно ускладнює

розуміння багатьох процесів, пов'язаних з використанням цифрових активів, і досить часто впливає на спотворення та помилкове трактування терміну. Подібна термінологічна плутанина створює стійкі умови для подальшого вкорінення суперечливості та неоднозначності не лише самого поняття «цифровий актив», а й перспектив його використання.

Аналіз сучасних наукових публікацій та досліджень з питань цифрового активу свідчить, що сьогодні не існує чіткого визначення та розуміння цього поняття [17]. Найчастіше поняття «цифровий актив» вживається у широкому сенсі, має неточне і неоднозначне формулювання і ототожнюється з криптовалютой. Так, Лаптев В.А. наполягає на широкому трактуванні цифрового активу і стверджує, що вони включають будь-які активи, що мають економічну цінність [3]. Закон SF0125 штату Вайомінг від 01.07.19 року дає таке визначення: «Цифровий актив означає відображення економічних, майнових прав чи прав доступу, які зберігаються в машиночитаному форматі й включають цифрові споживчі активи, цифрові цінні папери і віртуальну валюту» [4]. Цифровий актив, за закладеною в нього суті, надає інформацію про цінність – є інформаційним ресурсом, який можна використовувати. Значимість цифрового активу полягає у його достовірності та унікальності, що забезпечуються внаслідок обігу цифрового активу в розподіленому реєстрі. На основі цих властивостей можна говорити про високе практичне значення цифрового активу в економіко-правовому аспекті.

На підставі встановлених сутнісно-сміслових особливостей поняття «цифровий актив» у статті «Обґрунтування поняття «Цифровий актив»: економіко-правовий аспект» сформульовано таке визначення: «Цифровий актив – інформаційний ресурс, похідний від права на цінність та звертається до розподіленого реєстру у вигляді унікального ідентифікатора» [2]. Таким чином, в рамках даної роботи під цифровим активом пропонується розуміти поняття запропоноване в даній статті.

1.2. Аспекти цифрового активу

Розглянемо різні аспекти цифрового активу.

Першим є економічний аспект. Невід'ємною складовою визначення цифрового активу є економічний аспект, виражений в інформаційному ресурсі, який має властивості практичної корисності, а також можливість багаторазового використання. При цьому в цифровому середовищі такий інформаційний ресурс представлений у вигляді унікального ідентифікатора. Таким чином, розглядаючи економічний аспект визначення цифрового активу, необхідно говорити про одноприродність та функціональну ідентичність цифрового активу як інформаційного ресурсу та унікального ідентифікатора.

Розглядаючи поняття «унікальний ідентифікатор», слід зазначити, що це поняття найчастіше використовують у суспільних відносинах. Так, при наданні документів до державного реєстру унікальний ідентифікатор є логічною послідовністю символів, за допомогою яких заявник передає на офіційному вебсайті спеціального уповноваженого органу з питань державної реєстрації інформацію про стан розгляду направлених ним електронних документів для реєстрації юридичної особи або фізичної особи – підприємця, а також отримує електронні документи, надіслані йому державним реєстратором [5].

В інформаційних системах унікальний ідентифікатор призначений для ідентифікації певного об'єкта в мережі й дозволяє виключити будь-яку можливість дублювання цього об'єкта, тим самим підтверджуючи його справжність [10]. Окремо слід зазначити той факт, що унікальний ідентифікатор вказує на форму представлення цифрового активу та свідчить про наявність унікального імені: середовище, в якому створюється та звертається цифровий актив, привласнює буквено-цифровий код, який дозволяє відрізнити конкретний цифровий актив у розподіленому реєстрі даних або блокчейні та отримати необхідну інформацію про нього. Основне призначення унікального ідентифікатора як загальноприйнятого критерію ідентифікації полягає у можливості дозволити уповноваженому суб'єкту ідентифікувати інформацію щодо її унікальності без центру координації

[10]. Таким чином, наявність унікального ідентифікатора дозволяє використовувати його для ідентифікації об'єкта або суб'єкта з достатньою мірою впевненості в тому, що даний ідентифікатор ніколи не буде навмисно використаний для цілей, які не закладені в нього спочатку, що є дуже важливим у сфері економіко-правових відносин.

Необхідно зауважити, що ідентифікація суб'єкта або об'єкта є не новим процесом, оскільки вже давно застосовується у різних галузях сфери суспільства. У електронному просторі унікальний ідентифікатор відображає: електронну пошту, номер телефону, хеш.

Так, ідентифікація суб'єктів та об'єктів давно й широко використовується, але останнім часом спостерігається яскраво виражена тенденція використання унікального ідентифікатора у середовищі виконання на базі технології блокчейн. За своєю природою блокчейн гарантує відсутність підробки та копіювання, що створює новий рівень надійності та довіри. Цінність унікального ідентифікатора полягає в тому, що за його допомогою забезпечується закріплення цифрового активу за інформацією про цінність [10]. Тому закріплення цифрового активу за інформацією про цінність за допомогою унікального ідентифікатора дозволяє безпечно оперувати цією інформацією. Таким чином, унікальний ідентифікатор, захищаючи та наділяючи інформацію невразливістю, формує економічний аспект цифрового активу.

Розглянемо правовий аспект. Якщо розглядати правовий аспект поняття цифрового активу як інформаційного ресурсу, необхідно говорити про формування такої його правової складової, як «похідність від права». За своєю сутністю цифровий актив є способом використання інформації про цінність, яка представлена у вигляді інформаційного ресурсу в інформаційній системі на базі технології блокчейн. Насамперед слід зазначити, що нормативне визначення поняття «похідність від права» у правовій сфері не встановлено. Похідність права як правового феномена використовують у всіх галузях права, тобто носить міжгалузевий характер.

Термін «право» можна розглядати як: формалізована сукупність правил поведінки; наука, що вивчає відповідні формалізовані правила поведінки; міра можливої поведінки, свобода дій, що надається державою. Останнє визначення права як міри можливої поведінки, свободи дій, що надається державою, є найближчим до справжнього розуміння в контексті визначення поняття «цифровий актив», представленого у статті «Обґрунтування поняття «Цифровий актив»: економіко-правовий аспект» [2].

Похідність від права виникає внаслідок трансформації базового явища права зі збереженням базових властивостей. Можна виділити такі властивості похідності від права на цінність: вторинність стосовно первинного (основного) права; прямий зв'язок похідності з таким правом; збереження базових властивостей первинних основних прав. Так, цифровий актив не є правом на цінність, але він має властивості похідного від права на цінність. Це означає, що власник цифрового активу має право на доступ до інформаційного ресурсу, який містить інформацію про комплекс прав на цінність та відомості про зобов'язання власника цінності, виконання яких він має право вимагати [10].

Таким чином, розглядаючи правовий аспект, можна стверджувати, що з передачею цифрового активу новому власнику також переходять і всі права попереднього власника [10]. Іншими словами, цифровий актив виступає цифровим відображенням цінності, тобто по суті цифровий актив не є правом на цінність, але має властивості похідного від права на цінність. Закладений у цифровий актив принцип дозволяє оперувати правом на інформаційний ресурс, похідний від права на цінність, тоді як сама цінність перебуває у власника.

Розглянемо інформаційний аспект. Сутність цифрового активу, що полягає в ототожненні його з інформаційним ресурсом, формує інформаційну складову в контексті визначення поняття «цифровий актив». Незважаючи на те, що в сучасній науці застосовується велика кількість термінів, пов'язаних з питаннями інформатизації суспільства, тільки деякі з них мають науково-ужитковий характер.

Термін «інформаційний ресурс» немає єдиного загальноприйнятого визначення, що пояснюється використанням даного терміна в різних сферах, при цьому підходи до розуміння інформаційного ресурсу також різняться. У законодавчих актах поняття «інформаційний ресурс» визначається як: сукупність документів в інформаційних системах: бібліотеках, архівах, банках даних тощо; організована сукупність документованої інформації, що включає бази даних, інші сукупності взаємозалежної інформації в інформаційних системах [2].

Інформаційні ресурси є продуктом інтелектуальної діяльності [6]. Таким чином, якщо розглядати інформаційний ресурс як структуровану та організовану певним чином інформацію, зафіксовану на матеріальному носії, яку можна зберігати, передавати, використовувати, поповнювати, то очевидно, що такий ресурс має властивості практичної значущості та корисності, а також можливість багаторазового використання. При цьому важливим є той факт, що інформація, акумульована конкретним чином і предметно оброблена, обумовлює кількість і якість інформаційних ресурсів, з чого слідує висновок: інформаційний ресурс має рідкісну властивість збільшення та розвитку інформаційного капіталу [10].

Розглядаючи інформаційний ресурс під кутом інформації про цінність, закладену в цифровий актив, необхідно відзначити такі властивості такої інформації: структурованість за певними параметрами та категоріями; фіксація на цифровому носії; зберігання, передача, обмін, використання тощо.

Дані властивості зумовлюють природу цифрового активу як інформаційного ресурсу, оскільки інформація, закладена в цифровий актив, має упорядкований і системний характер. Окремо слід зазначити, що інформацію про цінність формує її власник.

З вищевикладеного можна дійти висновку в тому, що концептуальний аналіз поняття інформаційного ресурсу зумовлює коректне розуміння компонента «інформаційний ресурс» цифрового активу. Насамперед, слід зазначити, що поняття «інформаційний ресурс» є органічною частиною структури поняття «цифровий актив» через те, що цифровий актив є ототожненням інформаційного ресурсу [10]. З огляду на те, що інформація про цінність включає інформацію про

комплекс прав і обмежень на цінність, об'єктом взаємовідносин власників цифрових активів є інформація про цінність, виражена в інформаційному ресурсі.

Наступним розглянемо ціннісний аспект. Однією зі складових цифрового активу є ціннісний аспект, на розгляд якого необхідно проаналізувати природу цінності. Так, виділяють інструментально-індикативний аспект цінності, який полягає в тому, що цінність за допомогою нормативного та казуального регулювання вказує на ті феномени, які покликані задовольняти потреби людей, і тому виступають цінностями, та інструментально-сумірний аспект цінності, що виражається у функції ранжирування соціальних цінностей за допомогою використання спеціальних інструментів [7].

Як зазначає Абішева О.К., «адаптація, формування, інтеграція та дезінтеграція цінностей у будь-яку зі сфер, процеси абсолютизації одних компонентів та знецінення інших породжують так званий фон дії цінностей» [8].

На становлення цінностей впливають такі фактори: попит у суспільстві; динаміка, темпи та напрямки розвитку цінностей; специфіка та характер розвитку громадських інститутів; особливості конкретної правової ситуації; існування різних матриць переваг, ієрархій ціннісних структур.

Слід зазначити, що суспільство прагне закріплення цінностей у нормативній площині, тобто їх редукування, у зв'язку з тим, що користування будь-якою цінністю ефективніше у разі, якщо вона виражена і захищена певною нормою.

Поняття «цінність» має ряд визначень та тлумачень, які залежать від предметної сфери їх застосування. Найбільш слушним у контексті цифрового активу є таке визначення: цінність – вартість конкретного об'єкта матеріальних або нематеріальних благ. Для детермінації вартості застосовують два підходи:

з точки зору класичної політичної економіки вартість слід розглядати як об'єктивну реальність, яка вимірюється за допомогою витраченої праці на виробництво тих чи інших благ;

представники неокласичної школи безпосередньо пов'язують вартість з рідкістю запитуваного блага.

Найпоширенішим є саме неокласичний підхід. У статті «Обґрунтування поняття «Цифровий актив»: економіко-правовий аспект» дано таке визначення: «Цінність – це значення вартості, важливості, корисності чогось, яка також може виражатися різними одиницями виміру в залежності від галузі практичного використання такої цінності та екосистеми, у якій така цінність існує» [2]. У визначенні поняття "цифровий актив" використання формулювання "похідний від права на цінність" означає, що, цифровий актив надає право на доступ до інформаційного ресурсу, який містить інформацію про цінність, за допомогою якої можливий доступ до самої цінності.

З вищесказаного можна зробити висновок, що у загальному розумінні цінності формуються як потреби існування або розвитку суспільства. Що стосується розгляду поняття цифрового активу «цінність» відбиває сутність, закладену в цифровий актив, цим формуючи ключовий елемент цифрового активу – його ціннісну складову.

1.3. Властивості та завдання цифрового активу

Виходячи з визначення цифрового активу як інформаційного ресурсу, похідного від права на цінність і представленому в розподіленому реєстрі у вигляді унікального ідентифікатора, необхідно позначити властивість цифрового активу, що зумовлює його співвідношення з інформаційним ресурсом [10]. Цифровий актив надає новий спосіб використання інформації про цінність, яка представлена у вигляді інформаційного ресурсу в інформаційній системі на базі технології розподіленого реєстру або блокчейну, що відкриває нові можливості для оперування комплексом прав на цінність [2].

Використовуючи цифровий актив, власник цінності використовує лише право на доступ до інформаційного ресурсу, що містить комплекс відомостей (інформації) про цінність, тоді як сама цінність перебуває у нього до пред'яв-

лення власником цифрового активу права вимоги до власника цінності щодо виконання зобов'язань, зафіксованих у такому інформаційному ресурсі [10]. За своєю природою цифровий актив є одиницею обліку обсягу допуску до інформаційного ресурсу, похідного від права на цінність. При передачі цифрового активу передається певний обсяг допуску до інформаційного ресурсу, де міститься інформація, закладена власником цінності. Володіння 100% обсягу допуску до певного інформаційного ресурсу надає власнику особливе право на використання такого інформаційного ресурсу: право вимоги у власника цінності виконати закладені на інформаційний ресурс зобов'язання; право обміняти наявні цифрові активи на інші цифрові активи. Слід зазначити, що з обміном цифровими активами також відбувається обмін певним обсягом допуску до інформаційних ресурсів, відповідним даним цифровим активам.

Цифровий актив дозволяє оперувати інформацією, що є не просто новим способом використання інформації, а включає процес організації взаємодії суб'єктів (власника цифрового активу і власника цінності) та об'єкта управління (цифрового активу), спрямованого на забезпечення: реалізації закладених у такий цифровий актив цілей; здійснення правовідносин, об'єктом яких є право на інформацію, похідну від права власності на цінність.

Управління інформацією за допомогою цифрового активу являє собою складне явище, що динамічно розвивається. Як соціальне явище таке управління виникає у разі виникнення потреби. Діяльність управління інформацією містить у собі постійний інноваційний компонент, який сприяє розвитку можливостей інформаційного ресурсу в ході використання інформаційної системи. При цьому така діяльність сприяє забезпеченню збереження та функціонування таких систем.

Розгляд основної властивості цифрового активу та встановлений зв'язок між цифровим активом як одиницею обсягу допуску (права на доступ) до інформаційного ресурсу та правом власності на цінність дозволяє говорити про новий об'єкт цивільних правовідносин – право на користування інформацією, похідною від права на цінність [16].

Відповідно до Міжнародного стандарту бухгалтерського обліку (International Accounting Standards) цифровий актив як інформаційний ресурс можна віднести до категорії «нематеріальних активів – немонетарних активів, що ідентифікуються, що не мають матеріально-речової форми» [9]. Така категоризація робить цифровий актив зрозумілим для оподаткування та незамінним для організації нового виду обліку та контролю цінностей суб'єктами господарювання. Завдання цифрових активів можна поділити на три групи: індивідуальні, галузеві та глобальні.

Індивідуальні завдання реалізуються у процесі індивідуального використання цифрового активу власником у розподіленому реєстрі та залежать від цільового призначення: облік та зберігання інформації, закладеної у цифровий актив; подальше розповсюдження інформації; тимчасове користування цифровим активом; фіксація прав інтелектуальної власності за допомогою цифрового активу тощо [10]. Так, власник цифрового активу самостійно визначає його індивідуальні завдання.

Галузеві задачі формуються в залежності від сфери застосування цифрових активів. Наприклад, завданням цифрового активу, що використовується як інформаційний ресурс у державному реєстрі, є систематичний облік даних [2]. Великі державні замовлення, наприклад, постачання сировини та обладнання, можуть фіксуватися та здійснюватися за допомогою відповідних цифрових активів, що забезпечить раціональний розподіл ресурсів та дозволить уникнути витоків коштів не за цільовим призначенням. У бухгалтерії цифровий актив може бути інструментом розрахунку числових показників, зведення до балансу активів і пасивів, а також іншої бухгалтерської інформації. Таким чином, завданням цифрових активів у бухгалтерській сфері є сприяння функціонуванню реєстру, в якому можна знаходити, відстежувати та перевіряти необхідні показники. Великі компанії можуть відображати товари та послуги у вигляді цифрових активів, розподіляти їх як заробітну плату та вести будь-яку внутрішню діяльність за допомогою цифрових активів та технології блокчейн, вирішуючи виробничі завдання та завдання обліку інформації.

Глобальні завдання цифрового активу включають: викорінення тіньової економіки та «подвійної» бухгалтерії; раціональний розподіл ресурсів за цільовим призначенням; усунення корупції та шахрайських операцій у сфері відносин, пов'язаних з обігом цінностей. Перелік завдань цифрового активу, виконання яких зумовлено його функціональними можливостями, дозволяє зробити висновок, що використання цифрового активу можливе у всіх сферах життя суспільства, зокрема соціальної, економічної, політичної та інших. Вирішення індивідуальних, галузевих та глобальних завдань цифрового активу сприяє створенню на базі технології блокчейн нових суспільних відносин у будь-якій із вищевказаних сфер, а також появі та оновленню в режимі реального часу надійних реєстрів та баз даних з обліку, зберігання та обігу інформації, закладеної в цифровий актив.

Трансформаційні процеси, зумовлені появою та поширенням технології блокчейн, докорінно змінюють підходи до роботи з інформацією. Принципи роботи блокчейна не просто сприяють її систематизації, вони перетворюють ціннісні характеристики такої інформації [2]. Так, інформація набуває значущості, яка може бути відображена у збільшенні попиту на її споживання та формування цінності. Крім цього, технологія блокчейн зміщує парадигму сприйняття активів, матеріальних цінностей та майнових прав, оскільки інформація, незважаючи на свою нематеріальну природу, стає в один ряд цінностей із матеріальними об'єктами. Процедура цифрування активів сприяє тому, що цифровий актив є цифровим відображенням структурованої та впорядкованої інформації про об'єкт, що набуває цінності серед свого звернення [10]. Так, відбувається своєрідна еволюція звичного обороту матеріальних цінностей, адже об'єктом обороту з урахуванням технології блокчейн є інформаційний ресурс як унікального ідентифікатора.

Подібні зміни у принципі звернення об'єктів потребують переосмислення майнових прав та правового регулювання. Право, як один із найважливіших соціальних регуляторів, має реагувати на потреби суспільства та перетворення соціальних відносин. При цьому важливим є виявлення можливостей інноваційних технологій, таких як блокчейн.

Використання об'єктів матеріального світу в цифровому просторі як одна з можливостей блокчейна сприяє спрощенню товарообігу, дозволяє суб'єктам обмінюватися інформацією про такий об'єкт без його фактичної передачі, а також керувати інформацією про такий об'єкт третім особам, тоді як сам об'єкт перебуває у власника [10]. Власне, з'являється новий вид правовідносин – «право на право».

Технологічні переваги блокчейна дозволяють виключити посередників (фінансових посередників, банки або третіх осіб) із відносин між суб'єктами, що, як наслідок, дозволяє оптимізувати фінансові витрати та тимчасові витрати. Слід зазначити, що на основі технології блокчейн з'явився новий спосіб управління інформацією у вигляді цифрового активу.

Технологія блокчейн спрямована не просто на покращення щоденного життя громадян, а й на підвищення якості товарів та послуг, ліквідацію адміністративних бар'єрів у роботі суб'єктів господарської діяльності, надання державних послуг, підвищення конкурентоспроможності економіки, відкритості та прозорості діяльності державних органів [2].

Нові засоби використання інформації призводять до появи нового ринку – ринку інформаційних ресурсів та послуг на основі технології блокчейн. Слід зазначити, що ці зміни не обов'язково спричинять кадрові скорочення, швидше, вони зумовляють необхідність перекваліфікації у напрямі інформаційних технологій та права.

Так, технологія блокчейн несе у собі кардинальні зміни у соціально-економічній, політико-правовій та інших сферах життя суспільства. При цьому правові зміни повинні бути такими, що дозволять дотримуватися балансу між збереженням фундаментальних суспільних та правових цінностей та безперешкодним впровадженням технології блокчейн. Призупинення або неприйняття змін на державному рівні, очевидно, спричинить природне відставання в індустрії і пов'язану з цим економічну стагнацію.

РОЗДІЛ 2. МЕТОДИ ТА МОДЕЛІ ОЦІНКИ ВПЛИВУ ЦИФРОВИХ ТЕХНОЛОГІЙ НА ЕКОНОМІКУ КРАЇН

2.1. Концептуальна схема побудови моделей оцінки впливу цифрових технологій на економіку країн

На основі розглянутого матеріалу та проаналізованого стану сучасної економіки країн, розроблено концептуальну схему побудови моделей оцінки впливу цифрових технологій на економіку країн, яка представлена на рис. 2.1.



Рис. 2.1. Концептуальна схема побудови моделей оцінки впливу цифрових технологій на економіку країн

Перший етап «Аналіз вхідних даних» включає три пункти:

збір та аналіз статистичних даних;
виділення ключових показників;
оцінка економічної ситуації в досліджуваних країнах.

Збір та аналіз статистичних даних проводився за допомогою сайту «Мін-фін» [11] та сайту «Гуманітарний портал ∞ ISSN 2310-1792 [24]». В ході пошуку статистичної інформації розглянуто річні показники економічної діяльності країн. Особливу увагу приділено показникам розвитку економіки країн та розвитку цифрових технологій. Дані розглянуто з 2012 року по 2020 рік.

Таким чином, виділено наступні ключові показники, що варто використати при побудові моделей:

ВВП на душу населення;
ВНД на душу населення;
рівень безробіття;
індекс електронного уряду;
індекс електронної участі;
індекс онлайн-сервісів;
індекс людського капіталу;
індекс телекомунікаційної інфраструктури;
відсоток користувачів Інтернетом.

Другий етап «Аналіз економіко математичних методів та моделей» включає: аналіз моделей панельних даних, аналіз та виділення методів побудови кластерного аналізу, аналіз та відбір методів побудови лінійної багатофакторної моделі.

У сучасних умовах, що характеризуються високим рівнем нестабільності та невизначеності, стає необхідним всебічний аналіз функціонування та розвитку соціально-економічних об'єктів за загальною інформаційною базою. Зазвичай вихідні дані представляють у вигляді "об'єкт-ознаки", де в рядках знаходяться об'єкти, а в стовпцях – ознаки. Для панельних даних додають ще один вимір – час [12]. У роботі аналіз буде здійснюватися по країнах та роках. Тому для аналізу показників використано моделі панельних даних.

Кластерний аналіз – це метод класифікаційного аналізу; його основне призначення – розбиття безлічі досліджуваних об'єктів і ознак на однорідні в деякому сенсі групи, або кластери. Це багатовимірний статистичний метод, тому передбачається, що вихідні дані можуть бути значного обсягу, тобто суттєво великими можуть бути як кількість об'єктів дослідження (спостережень), так і ознаки, що характеризують ці об'єкти [21]. У даній роботі доцільніше використати кластерний аналіз, щоб розбити країни на групи по показникам та проаналізувати ці кластери.

Аналіз впливу показників доцільніше робити за допомогою побудови багатофакторної лінійної моделі, допомагає більш детально та всебічно проаналізувати фактори впливу на досліджуваний об'єкт, оцінити вплив кожного фактору та їх взаємозв'язок, на основі отриманих результатів побудувати прогнозні значення досліджуваного показника.

Третій етап «Побудова моделей» включає безпосереднє практичне застосування відібраних економіко-математичних методів в побудові моделей.

У побудові моделей панельних даних та кластерного аналізу використано та розглянуто показники, що згадано вище, по сімдесяти країнам в період з 2012 року по 2020 рік.

У побудові багатофакторної лінійної моделі використано та розглянуто ключові показники, що було згадано, по Україні в період з 2012 року по 2020 рік.

Результатом моделювання на панельних даних є оцінка впливу незалежних змінних, що відображають рієнь цифрового розвитку країн, на залежні змінні, що характеризують економічний розвиток країн.

Результатом реалізації кластерного аналізу є розподілення країн на групи, однорідні за рівнем цифрового та економічного розвитку.

Результатом побудови багатофакторної лінійної моделі є оцінка впливу факторів на економіку країн, побудова найбільш адекватної моделі для прогнозування, та безпосередня побудова прогнозу впливу цифрових технологій на економіку України.

Показники, які будуть характеризувати економічну діяльність країн: ВВП на душу населення, ВНД на душу населення та рівень безробіття. Ці показники будуть результуючими, так як описують рівень економічного розвитку країн.

Показники які будуть характеризувати розвиток цифрових технологій країн: індекс розвитку електронного уряду, індекс електронної участі, індекс розвитку онлайн-сервісів, індекс людського капіталу, індекс телекомунікаційної інфраструктури, відсоток користувачів Інтернетом. Ці показники будуть факторними.

Індекс розвитку електронного уряду в країнах світу (Global E-Government Index) – це комплексний показник, який оцінює готовність та можливості національних державних структур у використанні інформаційно-комунікаційних технологій (ІКТ) для надання громадянам державних послуг [23].

Індекс електронної участі (E-participation index) – відображає розвиток сервісів активної комунікації між громадянами та державою.

Індекс розвитку онлайн-сервісів (Online Service Index) служить оцінкою розвитку та ефективності використання урядових порталів, веб-сайтів та надання електронних послуг. Індекс розвитку електронних послуг відображає всі стадії розвитку електронних послуг у країні.

Індекс людського капіталу (Human Capital Index) – це індекс вимірює досягнення країни щодо рівня освіти її населення за двома основними показниками: рівень грамотності дорослого населення та індекс сукупної частки учнів, які отримують початкову, середню та вищу освіту [22].

Індекс телекомунікаційної інфраструктури (Telecommunication Infrastructure Index) – складний, зважений індекс, що стосуються можливості інфраструктури постачати електронні послуги.

Відсоток користувачів Інтернетом (Internet users) – відображає відсоток користувачів Інтернетом на 1000 мешканців країни.

Всі ці показники відображають рівень впровадження цифрових технологій в країну та на скільки країна розвинута в області цифрових технологій.

Таким чином, в ході проведення дослідження розроблено аналітичні моделі оцінки впливу цифрових технологій на економіку країн, розбито країни на кластери по показникам, побудовано багатofакторну модель для прогнозу показників.

2.2. Моделі панельних даних та методи їх оцінювання

Особливістю моделей, побудованих на базі панельних даних, є можливість аналізу та вивчення як просторових даних, так і даних тимчасового типу. Переваги таких моделей полягають у можливості обліку та аналізу індивідуальних особливостей між економічними одиницями, ефективного усунення гетерогенності об'єктів на підставі коротких рядів, що неможливо виконати при побудові стандартних регресійних моделей [12]. У моделях ознаки розташовують по стовпцях, рядкам дані про i -му об'єкті за T періодів часу (рис. 2.2).

№ об'єкта	t	X _{1t}	X _{2t}	...	X _{mt}	Y _{it}
i=1	1	X ₁₁	X ₂₁	...	X _{m1}	Y ₁₁
	2	X ₁₂	X ₂₂	...	X _{m2}	Y ₁₂

	T	X _{1T}	X _{2T}	...	X _{mT}	Y _{1T}
i=2	1	X ₁₁	X ₂₁	...	X _{m1}	Y ₂₁
	2	X ₁₂	X ₂₂	...	X _{m2}	Y ₂₂

	T	X _{1T}	X _{2T}	...	X _{mT}	Y _{2T}
...	
i=n	1	X ₁₁	X ₂₁	...	X _{m1}	Y _{n1}
	2	X ₁₂	X ₂₂	...	X _{m2}	Y _{n2}

	T	X _{1T}	X _{2T}	...	X _{mT}	Y _{nT}

Рис. 2.2. Матриця панельних даних для побудови моделі

Панельні (просторові) дані – набір даних, що складаються зі спостережень за однотипними статистичними об'єктами протягом кількох часових періодів [12].

Досить часто індивідуальні фактори є суттєвими і корельовані з іншими пояснювальними змінними. У межах моделей регресії це означає, що виняток такого чинника призводить до зміщених оцінок інших параметрів. У такому випадку застосування моделей з панельними даними дозволяє отримати точніші оцінки [12].

У тому випадку, якщо кількість періодів часу, що спостерігаються, більша за кількість досліджуваних об'єктів, панельні дані називаються об'єднаним часовим рядом. Насправді безліч панельних даних включає спостереження протягом небагатьох періодів часу. У такому випадку важливим є моделювання відмінностей між досліджуваними об'єктами, їх гетерогенності (неоднорідності), ніж аналіз тимчасових ефектів. Але незважаючи на те, що тимчасові ефекти явно не моделюються, панельні дані мають інформацію про розвиток однотипних об'єктів у часі [12].

Якщо є дані про кожен об'єкт спостереження в кожний момент часу, тобто якщо немає пропущених спостережень, про панельні дані говорять, що вони збалансовані (або повні) (рис. 2.3. А). Збалансована панель є ідеальним випадком на практиці. Крім цього, для панельних даних притаманна загальна проблема обстежень — проблема самовідбору, пов'язана з недоступністю значень за певними позиціями, що призводить до зміщення даних. Якщо відсутність будь-якого спостереження має природну причину (наприклад, переїзд людини або домогосподарства з одного міста в інше) і можливо корелює з величинами, що спостерігаються, такий ефект називають панельною смертністю, або виснаженням (рис. 2.3. Б). Оголошення такого спостереження відсутнім може призвести до зміщених та неефективних оцінок. Для усунення ефекту виснаження (неконтрольованого) можна скористатися складанням вибірки з використанням ротаційної панелі. Для створення такої вибірки кожен об'єкт спостереження вибуває з обстеження на протязі певної кількості періодів часу, після чого до обстежуваної вибірки додають об'єкт, який спостерігається таку ж кількість періодів (рис. 2.3. В) [12].

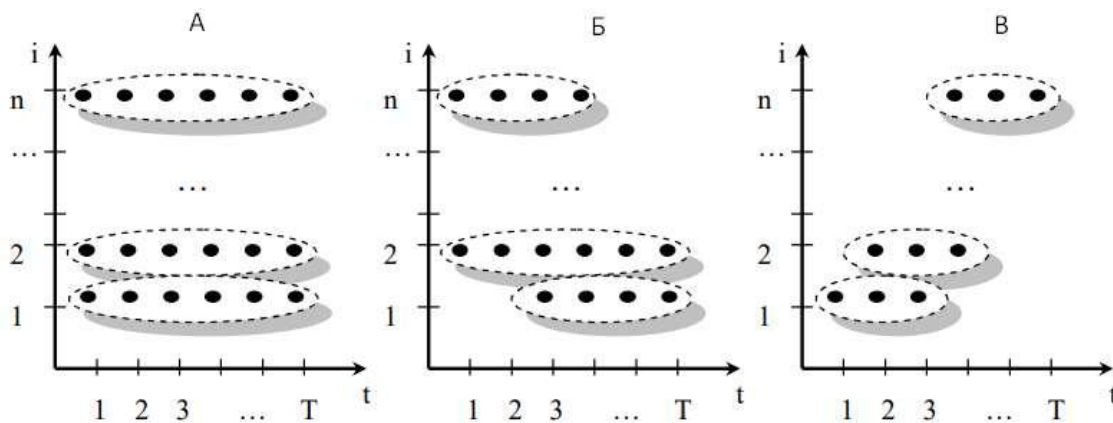


Рис. 2.3. Види панелей

Введемо наступні позначення для моделі панельних даних:

$$y_i = \begin{bmatrix} y_{i1} \\ \cdot \\ \cdot \\ y_{iT} \end{bmatrix}; y = \begin{bmatrix} y_1 \\ \cdot \\ \cdot \\ y_n \end{bmatrix}; X_i = \begin{bmatrix} x_{i1} \\ \cdot \\ \cdot \\ x_{iT} \end{bmatrix}; X = \begin{bmatrix} X_1 \\ \cdot \\ \cdot \\ X_n \end{bmatrix}; \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \cdot \\ \cdot \\ \varepsilon_{iT} \end{bmatrix}; \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \cdot \\ \cdot \\ \varepsilon_n \end{bmatrix}; \mu_{0i} = \begin{bmatrix} \mu_{01} \\ \cdot \\ \cdot \\ \mu_{0n} \end{bmatrix}; i = \overline{1, n}; t = \overline{1, T},$$

де y_{it} – залежна змінна для i -го об'єкта в момент часу t ;

y_i – об'єднана залежна змінна;

x_{it} – набір пояснюючих змінних для i -го об'єкта на момент часу;

x_i – набір пояснюючих змінних;

ε_{it} – помилка для i -го об'єкта на момент часу t ;

ε_i – набір індивідуальних помилок [12].

Традиційно лінійна залежність панельних даних для i -го об'єкта обраної генеральної сукупності має вигляд:

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it}.$$

Параметри α_i є адитивними константами, які підсумовують ефекти, характерні для конкретного об'єкта спостереження i та періоду часу t , і, отже, визначають середнє розташування y_{it} , якщо всі регресори зафіксовані на рівні $x_{it} = 0$. Параметри розташування прийнято називати неспостережуваними ефектами [12].

Найбільш прості припущення можуть мати такий вигляд:

$$y_{it} = x_{it}\beta + \varepsilon_{it}.$$

Представлена модель передбачає ніяких ефектів, притаманних окремих об'єктів спостереження чи моментів часу. У разі мають на увазі, що дані об'єднані. Параметри моделі оцінюються за допомогою МНК з усіх спостережень, не враховуючи специфіки панельних даних [12].

У моделях панельних даних обурення поділяються на кілька компонентів. Виділяють моделі з одно- та двокомпонентною помилкою. Найбільш поширені моделі з однокомпонентною складовою помилкою (одномірними шоками). У свою чергу одно- та двовимірні шоки можуть представляти фіксовані та випадкові ефекти. У першому випадку досліджують вплив специфічних для кожного об'єкта факторів, виражених у значеннях констант α_i за відсутності загального параметра розташування [12].

У другому випадку ефекти розбіжності у об'єктах спостереження моделюються не за допомогою параметрів, а через причини щодо виду розподілу [12]. Ефекти стають компонентом залишків, тобто параметри розташування u_i приймаються як випадкові величини з нульовим математичним очікуванням, а обурення ε_{it} некорельованих для різних періодів часу:

$$\alpha_{it} = \alpha, \beta_{it}, v_{it} = u_{it} + \varepsilon_{it},$$

$$Z_i = (y_i, X_i) \sim i. i. d.,$$

де $\alpha_{it}\varepsilon_{it}$ – незалежні для будь-яких i, j, t ;

$$\varepsilon_{it} \sim i. i. d., \text{ де } E[\varepsilon_{it}|X_i, u_i] = 0, E[\varepsilon_{it}^2|X_i, u_i] = \sigma_V^2;$$

$$u_i \sim i. i. d., \text{ де } E[u_i|X_i] = 0, E[u_i^2|X_i] = \sigma_M^2 [12].$$

У такому випадку моделі оцінюють $(k + 1)$ коефіцієнт і дисперсії σ_V^2 та σ_M^2 . Це можливо із застосуванням двоступінчастої процедури, коли спочатку оцінюються дисперсії, а потім коефіцієнти узагальненим методом найменших квадратів [12].

У моделях з двокомпонентною складовою помилки як ще одного компонента можуть бути включені часові ефекти: $v_{it} = u_i + \varepsilon_{it} + l_t$ з необхідними передумовами щодо розподілу l_t [12].

Залежно від припущень щодо характеру величини α_t розглядаються три моделі: модель з фіксованими ефектами, модель з випадковими ефектами та об'єднана модель регресії. Модель із фіксованими ефектами має вигляд:

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it}.$$

При цьому необхідно, щоб виконувались такі умови: помилки некоррелювали між собою як по i , так і по t , тобто: $E(\varepsilon_{it}) = 0, V(\varepsilon_{it}) = \sigma_\varepsilon^2$; помилки некоррелювали з регресорами [12].

Модель також можна розглядати як модель з індивідуальними фіктивними змінними, тобто для кожного об'єкта спостереження вводиться змінна, що має індивідуальний характер. Припускаючи наявність тих самих параметрів для всіх об'єктів спостереження в усі моменти часу, можна досліджувати наявність гетерогенності між об'єктами спостереження з інваріантним по відношенню до часу, але специфічним для кожного об'єкта спостереження параметром розташування [12].

Якщо ввести фіктивні змінні $d_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$ для кожної економічної одиниці, модель можна представити у вигляді стандартної моделі:

$$y_{it} = \sum_{j=1}^n \alpha_j d_{ij} + x_{ij}\beta + \varepsilon_{it}.$$

Отриману модель можна оцінити звичайним МНК, проте виникає складність в оцінці з великою кількістю параметрів [12]. Тому вводять середнє за часом для кожного об'єкта спостереження та при переході з отриманої формули до середніх за часом, отримуємо:

$$\bar{y}_i = \alpha_j + \bar{x}_i \beta + \bar{\varepsilon}_i, \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \bar{x}_i = \frac{1}{T} \sum_{t=1}^T x_{it}, \bar{\varepsilon}_i = \frac{1}{T} \sum_{t=1}^T \varepsilon_{it}.$$

Отримуємо рівняння генеральної сукупності з поправкою на середнє:

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i) \beta + \varepsilon_{it} - \bar{\varepsilon}_i.$$

Це перетворення називається внутрішньогруповим перетворенням [12]. Застосовуючи метод найменших квадратів до регресії через початок координат, отримуємо оцінку:

$$\hat{\beta}_{FE} = \left(\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1} * \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i).$$

Ця оцінка називається оцінкою моделі з фіксованими ефектами. Оскільки з даних віднімається середнє за часом, цю оцінку називають оцінкою з урахуванням варіації в межах об'єкта спостереження. В якості оцінок індивідуальних ефектів можна приймати $\hat{a}_i = \bar{y}_i - \bar{x}_i \hat{\beta}_{FE}$. Вони є незміщеними і устатковані для фіксованого n при $t \rightarrow \infty$. Вираз для матриці коваріацій оцінки $\hat{\beta}_{FE}$ має вид: $V(\hat{\beta}_{FE}) = \sigma_\varepsilon^2 (\sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)')^{-1}$ [12]. В якості оцінки дисперсії σ_ε^2 можна прийняти:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{nT - n - k} \sum_{i=1}^n \sum_{t=1}^T (y_{it} - \bar{y}_i - (x_{it} - \bar{x}_i) \hat{\beta}_{FE})^2.$$

Оцінки параметрів моделі з фіктивними змінними безпосередньо не обчислюють, оскільки виникає потреба звертати велику матрицю при введенні великої кількості додаткових фіктивних змінних. Натомість застосовують двоступінчасту процедуру, при якій спочатку обчислюють оцінку з урахуванням варіації між об'єктами спостереження, а потім індивідуальні ефекти [12].

Модель з випадковими ефектами досить добре адаптована до структур панельних даних, що сприяє усуненню деяких недоліків моделі з фіксованими ефектами, особливо для проблеми числа параметрів, що оцінюються. Свою сутність випадкові ефекти виражають у тому, що ефекти u_i , що описують гетерогенність, є випадковими змінними у сенсі випадковості вибірки з генеральної сукупності, тому що кожен об'єкт спостереження має специфічний, не залежний від часу, ефект. Тобто вибірка, що включає досліджувані об'єкти, розглядається як випадкова з деякої генеральної сукупності. Як і для моделі з фіксованими ефектами, випадкові ефекти відбивають наявність деяких індивідуальних характеристик об'єктів, інваріантних у часі, які складно спостерігати. Однак значення цих параметрів вводяться до складу помилки [12].

Рівняння моделі з випадковими ефектами має вигляд:

$$y_{it} = \mu + x_{it}\beta + u_i + \varepsilon_{it},$$

де u_i – випадкова помилка, інваріантна за часом для кожного об'єкта;
 μ – константа [12].

У моделі передбачається, що виконуються такі умови:

помилки ε_{it} некорельовані між собою $E(\varepsilon_{it}) = 0, V(\varepsilon_{it}) = \sigma_\varepsilon^2$;

помилки ε_{it} некорельовані з регресорами x_{js} при всіх i, t, j, s ;

помилки u_i некорельовані $E(u_{it}) = 0, V(u_{it}) = \sigma_u^2$;

помилки u_i некорельовані з регресорами x_{js} при всіх i, t, j, s ;

помилки u_i та ε_{it} некорельовані при всіх i, t, j, s [12].

Таким чином, у моделі передбачається наявність одних і тих же параметрів для всіх об'єктів спостереження у всі моменти часу, проте досліджується ефект гетерогенності об'єктів спостереження за допомогою введення постійного за часом, але індивідуального для кожного об'єкта спостереження доданку помилки u_i , який є незалежним від решти помилок [12].

Модель з випадковими ефектами не може бути оцінена з допомогою МНК ефективно, оскільки помилки за припущеннями моделі корелюють між собою

внаслідок наявності специфічного доданку для кожного об'єкта спостереження. Тому на першому кроці застосовують узагальнений метод найменших квадратів (УМНК), що зважає залишки відповідно до структури матриці коваріацій, що відповідає умовам моделі з випадковими ефектами. На практиці дисперсії залишків замінюють на встановленні оцінки $\hat{\sigma}_M^2$ та $\hat{\sigma}_V^2$. Можливі різні результати в залежності від застосовуваних методів оцінювання цих дисперсій. Найчастіше дисперсію незалежної компоненти залишків пропонують оцінювати, використовуючи квадрати різниць між оціненими залишками u_{it} та їх середніми по кожному об'єкту спостереження [12].

На другому кроці виконується УМНК-оцінювання з оцінками дисперсій в якості параметрів. Ця двокрокова процедура називається здійсненим узагальненим методом найменших квадратів (ЗУМНК) [12]. ЗУМНК-оцінка для рівняння моделі з випадковими ефектами є оцінкою цієї моделі:

$$\hat{\beta}_{RE} = W\hat{\beta}_B + (I_k - W)\hat{\beta}_{FE},$$

$$\hat{\beta}_B = \left(\sum_{i=1}^n (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1} \sum_{i=1}^n (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i); \bar{y} = \frac{1}{n} \sum_{y=1}^n \bar{y}_i, \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

$$W = (S_{xx}^w + (1 - \theta)^2 T S_{xx}^w)^{-1} S_{xx}^w; S_{xx}^w = \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)',$$

де I_k – одинична матриця, розмірності k (що пояснюють змінні) [12].

Параметр θ для методу вибирають так, щоб помилки цієї моделі не були взаємопов'язані в часі для різних значень t . Тоді параметр θ набуває вигляду:

$$\theta = 1 - \frac{\hat{\sigma}_V}{\sqrt{\hat{\sigma}_V^2 + T\hat{\sigma}_M^2}}; \hat{\sigma}_M^2 = \frac{\sum_{i=1}^N \bar{\hat{u}}_i^2}{n - k - 1} - \frac{1}{T} \hat{\sigma}_V^2; \hat{\sigma}_V^2 = \frac{\sum_{i=1}^n \sum_{t=1}^T (\hat{u}_{it} - \bar{\hat{u}}_i)^2}{nT - n - k}; u_{it} = \alpha_i + \varepsilon_{it},$$

де u_i – залишки, отримані за МНК у регресії з урахуванням варіації між об'єктами спостереження:

$$\bar{y}_i = \mu + \bar{x}_i' \beta + u_i + \bar{\varepsilon}_i; i = \overline{1, N},$$

де n – кількість об'єктів;
 k – кількість пояснюючих змінних [12].

Об'єднана модель регресії ігнорує панельну природу даних, не враховує індивідуальні особливості спостережень та тимчасові ефекти, і розпоряджає однаковою поведінку всіх об'єктів вибірки у всі періоди часу. Якщо ці припущення виконуються, то параметри моделі можуть бути оцінено за допомогою методу найменших квадратів (МНК) [12]. Формально модель може бути описана рівнянням:

$$y_{it} = \alpha_i + x_{it}\beta + \varepsilon_{it},$$

де x_{it} – вектор-рядок значень детермінованих регресорів;
 α_i і вектор-стовпець β – коефіцієнти регресії, однакові для всіх спостережень;
 ε_{it} – помилка для i -го об'єкта на момент часу t , нормальні та задовольняють умови класичної лінійної регресійної моделі, у тому числі умовою некорельованості з x_{it} [12].

Відповідна оцінка в матричній формі запису матиме вигляд:

$$\beta = (X'X)^{-1}X'Y.$$

Отже, розглянуто три основні регресії: об'єднану, з фіксованими ефектами і з випадковими ефектами. Оскільки в моделях панельних даних можуть бути індивідуальні розбіжності по кожному з об'єктів у досліджуваній вибірці, виявлення таких розбіжностей і відповідний вид специфікації моделі здійснюють за допомогою низки статистичних тестів [12]. Щоб з'ясувати оцінки якої з цих моделей найбільш адекватні для даних, необхідно попарне порівняння оцінених моделей.

Для перевірки статистичної значущості параметрів перетину (місце розташування) у моделі формують нульову гіпотезу H_0 таку, що $H_0: \mu_i = \mu_j$ для

будь-яких i, j , що відповідає моделі з одним і тим самим параметром μ для всіх об'єктів вибірки, тобто об'єднаної моделі регресії. Альтернативна гіпотеза полягає у тому, що $H_1: \mu_i \neq \mu_j$ хоча б для однієї пари i, j , що відповідає моделі із фіксованими ефектами [12]. Таку пару гіпотез можна перевірити за допомогою F-тесту:

$$F = \frac{R_{FE}^2 - R_{pool}^2}{1 - R_{FE}^2} * \frac{nT - n - d}{n - 1} = \frac{Q_{pool} - Q_{FE}}{Q_{FE}} * \frac{nT - n - d^{H_0}}{n - 1} \sim F(n - 1, nT - n - d),$$

де R_{FE}^2 – коефіцієнт множинної кореляції моделі з фіксованими ефектами;

R_{pool}^2 – коефіцієнт об'єднаної моделі регресії;

Q_{pool} – сума квадратів залишків об'єднаної моделі регресії;

Q_{FE} – сума квадратів залишків моделі із фіксованими ефектами;

n – кількість досліджуваних об'єктів;

T – кількість часових моментів;

d – кількість незалежних змінних моделі [12].

Якщо справедлива гіпотеза H_0 і виконується передумова нормального розподілу помилок (або у разі досить великих n), тестова статистика має (наближено) F-розподіл з $(n-1)$ та $(nT-n-k)$ ступенями свободи. Аналогічну гіпотезу можна перевірити за допомогою тесту Вальда. Однак його використання можливе у разі малої кількості об'єктів [12].

Перевірку на значущість випадкових ефектів моделі здійснюють за допомогою тесту множників Лагранжа, запропонованого Бреушем та Паганом, що базується на відповідній статистиці:

$$LM = \frac{nT}{2(T - 1)} \left(\frac{\sum_{i=1}^n (\sum_{t=1}^T e_{it})^2}{\sum_{i=1}^n \sum_{t=1}^T e_{it}^2} - 1 \right)^2,$$

де e_{it} – залишки об'єднаної моделі регресії [12].

За цим тестом висувають таку нульову гіпотезу H_0 : об'єднана модель регресії є окремим випадком моделі з випадковими ефектами, де відсутні помилки

u_i або $\sigma_u^2 = 0$. Відповідно в гіпотезі H_1 приймаємо $\sigma_u^2 > 0$. Якщо гіпотеза H_0 є вірною і виконується передумова відносно нормального розподілу помилок, статистика LM має асимптотичний χ^2 розподіл з одним ступенем свободи [12].

Оскільки найважливішою відмінністю підходів до моделювання гетерогенності об'єктів спостереження є співвідношення ефектів, що вводяться з регресорами: випадкові ефекти передбачаються некорельованими з регресорами, у той час як фіксовані ефекти можуть з ними корелювати, вибір моделі з фіксованими або випадковими ефектами залежить від того, корелює ефект чи ні [12].

При справедливості H_1 оцінки моделі з фіксованими ефектами спроможні, а оцінки моделі з випадковими ефектами неспроможні. У цьому випадку очікується суттєва різниця між оцінками даних двох моделей [12]. Виявлення такої різниці досліджують за допомогою відповідного тесту, що базується на статистиці Хаусмана:

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' \hat{\Phi}^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE}),$$

де $\hat{\Phi}$ – оцінка матриці коваріацій $(\hat{\beta}_{FE} - \hat{\beta}_{RE})$, що має асимптотичний розподіл χ^2 з d ступенями свободи;

$\hat{\beta}_{FE}$ – вектор оцінок моделі з фіксованими ефектами;

$\hat{\beta}_{RE}$ – вектор оцінок моделі з випадковими ефектами [12].

За цим тестом на підставі отриманих оцінок моделей з фіксованими та випадковими ефектами висувають гіпотезу H_0 таку, що оцінки моделі з випадковими ефектами є ефективними та обґрунтованими і не повинні відрізнятись від оцінок моделі з фіксованими ефектами. Відповідно, якщо справедлива альтернативна гіпотеза H_1 , різниця між оцінками моделі з випадковими та фіксованими ефектами є суттєвою, але оцінки моделі з фіксованими ефектами є обґрунтованими [12].

2.3. Методи кластерного аналізу

Кластерний аналіз – це розбиття всієї вихідної сукупності елементів на групи однорідних об'єктів, подібних між собою, але ці групи мають чіткі відмінності одна від одної [14]. На практиці кластерний аналіз має переваги перед іншими методами аналізу. Зокрема, мінімальні вимоги до даних. Проте, на кластерний аналіз не накладаються обмеження інших статистичних методів, такі як однорідність та обсяг вибірки. Можна аналізувати будь-які дані про будь-які об'єкти, і якщо зв'язок є – його буде видно.

Кластерний аналіз можна розглядати як інструмент дослідницького аналізу даних, спрямований на сортування різних об'єктів за значущими групами таким чином, щоб ступінь зв'язаності цих об'єктів був максимальним, якщо вони належать до однієї групи, і мінімальним, якщо не належать [13]. Кластерний аналіз використовується для виявлення прихованих структур або взаємозв'язків даних без необхідності пояснювати або інтерпретувати, що це за взаємозв'язок.

Класифікація спостережень на групи вимагає деяких методів для обчислення відстані або подібності між кожною парою спостережень. Результат цього обчислення відомий як матриця несхожості або відстань. Існує багато методів для обчислення відстані. Вибір мір відстаней є важливим кроком у кластеризації. На цьому кроці обчислюється подібність двох елементів (x, y) , і як це вплине на форму кластерів. Класичними методами вимірювання відстані є евклідова та манхеттенська відстані [13]. Евклідова відстань розраховується за формулою:

$$d_{euc}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Манхеттенська відстань розраховується за формулою:

$$d_{man}(x, y) = \sum_{i=1}^n |(x_i - y_i)|,$$

де x, y – два вектори довжини n [13].

Існують й інші методи, де відстань розраховується на основі кореляції. Відстань на основі кореляції визначається шляхом віднімання коефіцієнта кореляції від одиниці. Можна використовувати різні методи на основі кореляції. Метод Пірсона розраховується за формулою:

$$d_{cor}(x, y) = 1 - \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}.$$

Метод кореляції Спірмена обчислює кореляцію між рангом x і рангом y змінних. Формула розрахунку за методом Спірмена:

$$d_{spear}(x, y) = 1 - \frac{\sum_{i=1}^n (\hat{x}_i - \bar{\hat{x}})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (\hat{x}_i - \bar{\hat{x}})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}},$$

де $\hat{x}_i = rank(x_i)$ і $\hat{y}_i = rank(y_i)$ [13].

Вибір мір відстаней дуже важливий, оскільки має сильний вплив на результати кластеризації. Для найбільш поширеного програмного забезпечення для кластеризації мірою відстані за замовчуванням є евклідова відстань. Однак, залежно від типу даних і цілей дослідження, перевага може віддаватися іншим параметрам.

Алгоритми неієрахічного поділу здійснюють декомпозицію набору даних, що складається з n спостережень, на k груп (кластерів) із заздалегідь невідомими параметрами. При цьому виконується пошук центроїдів – максимально віддалених один від одного центрів згущень точок C_k з мінімальним розкидом усередині кожного кластера [19]. До алгоритмів поділу даних на кластери відносяться:

метод k -середніх Мак-Кіна, у якому кожен із k кластерів представлений центроїдом;

поділ навколо k медоїдів або PAM (Partitioning Around Medoids), де медоїд – це центроїд, координати якого зміщені до найближчого з вихідних об'єктів даних;

алгоритм CLARA (Clustering Large Applications) – метод, дуже схожий на PAM і використовують для аналізу великих наборів даних [19].

Об'єднання в кластери методом k -середніх – дуже простий та ефективний алгоритм, що має, однак, дві суттєві проблеми. По-перше, підсумкові результати чутливі до початкового випадкового вибору центрів груп. Можливе вирішення цієї проблеми полягає у багаторазовому виконанні алгоритму з різним випадковим призначенням початкових центроїдів. Ітерація з мінімальним значенням W_{total} відбирається як кінцевий варіант кластеризації. Друга проблема – необхідність задавати фіксоване число кластерів для розбиття, яке не завжди вибирається оптимальним [14]. Тому одним із завдань кластерного аналізу є підбір оптимального значення k , для якої існує кілька версій рішення. Основна ідея кластеризації k -середніх полягає у визначенні кластерів таким чином, щоб загальна внутрішньокластерна варіація була мінімізована [13]. Існує кілька доступних алгоритмів k -середніх. Стандартним алгоритмом є алгоритм Хартігана-Вонга, який визначає загальну варіацію всередині кластера як суму квадратів евклідових відстаней між елементами та відповідним центром за формулою:

$$W(C_k) = \sum_{x_i \in C_k} (x_i - \mu_k)^2,$$

де x_i – це точка даних, що належить кластеру C_k ;

μ_k – середнє значення балів, присвоєних кластеру C_k [13].

Кожне спостереження x_i призначається даному кластеру так, що сума квадратів відстані спостереження до призначених їм центрів скупчення μ_k зводилось до мінімуму [13]. Загальну варіацію всередині кластера можна розрахувати за формулою:

$$tot. \text{ withiness} = \sum_{k=1}^k W(C_k) = \sum_{k=1}^k \sum_{x_i \in C_k} (x_i - \mu_k)^2.$$

Загальна внутрішньокластерна сума квадратів вимірює компактність кластеризації, яка повинна бути якомога меншою.

Метод k середніх виконує кластеризацію таким чином:

перше – призначається число груп (k), куди мають бути розбиті дані. Випадково вибирається k об'єктів вихідного набору як початкові центри кластерів;

друге – кожному спостереженню присвоюється номер групи з найближчого центроїду, тобто на підставі найменшої евклідової відстані між об'єктом і точкою C_k ;

третє – перераховуються координати центроїдів μ_k всіх k кластерів та обчислюються внутрішньогрупові розкиди за формулою:

$$W(C_k) = \sum_{x_i \in C_k} (x_i - \mu_k)^2.$$

Якщо набір даних включає p змінних, μ_k являє собою вектор середніх з p елементами;

четверте – мінімізується загальний внутрішньогруповий розкид за формулою:

$$W_{total} = \sum_k W(C_k) \rightarrow \min,$$

де кроки 2 і 3 повторюються багаторазово, доки призначення груп не припиняють змінюватись або не досягнуто заданої кількості ітерацій [19].

Метод "ліктя" розглядає характер зміни розкиду W_{total} зі збільшенням числа груп k . Об'єднавши всі n спостережень в одну групу, маємо найбільшу внутрішньокластерну дисперсію, яка знижуватиметься до 0 при $k \rightarrow n$. На якомусь етапі можна побачити, що зниження цієї дисперсії сповільнюється – на графіці це відбувається в точці, що називається "ліктем" [14].

Альтернативою метода "ліктя" є використання статистики розриву або GAP-статистики, яка генерується на основі ресемплінгу та імітаційних процедур Монте-Карло. Нехай $E_n^*\{\log(W_k^*)\}$ означає оцінку середньої дисперсії W_k^* , отриманої бутстреп-методом, коли k кластерів утворені випадковими наборами об'єктів з вихідної вибірки розміром n [19]. Тоді статистика розраховується за формулою:

$$Gap_n(k) = E_n^*\{\log(W_k^*)\} - \log(W_k).$$

Статистика визначає відхилення дисперсії W_k від її очікуваної величини при справедливості нульової гіпотези про те, що вихідні дані утворюють тільки один кластер [19].

При порівняльному аналізі послідовності значень $Gap_n(k)$, $k = 2, \dots, K_{max}$ найбільше значення статистики відповідає найбільш корисному угрупованню, дисперсія якого максимально менша за внутрішньогрупову дисперсію кластерів, зібраних з випадкових об'єктів вихідної вибірки досліджуваних даних [19].

Метод РАМ багато в чому ідентичний алгоритму k -середніх за винятком того, що замість обчислення центроїдів здійснюється пошук k найбільш представницьких об'єктів (або медоїдів) серед аналізованих спостережень. Внутрішньокластерний розкид оцінюється по манхеттенській відстані [14].

Оскільки параметром функції РАМ також є число кластерів k , перевіримо оптимальність його значення третім найбільш популярним методом – по середній ширині силуету. Для кожного знайденого кластера розраховується "ширина силуету":

$$s_i = \frac{b(i) - a(i)}{\max[b(i), a(i)]},$$

де $a(i)$ – середня відстань між об'єктами i -го кластера;

$b(i)$ - середня відстань від об'єктів i -го кластера до іншого кластера, найближчого до i -го [13].

Для відбору оптимальної кількості кластерів будується графік залежності S_{mean} від k .

Метод CLARA є алгоритмом для кластеризації великих даних на основі PAM. Ідея алгоритму полягає в тому, щоб виконати випадкову вибірку у великих даних, потім використовувати алгоритм PAM для кожної вибіркової групи i , на решті, знайти центр кластера з найменшою вартістю в якості поточного серед кращих центральних точок, кластеризованих для кожної вибірки. Найкраща кластеризація для великих вибірок даних [14].

Методи ієрархічної кластеризації ґрунтуються на двох ідеях: агломерації (AGNES, Agglomerative Nesting), тобто послідовного об'єднання індивідуальних об'єктів або їх груп у більші підмножини, або зворотному за змістом процесі розбиття (DIANA, Divise Analysis), який починається з кореня і на кожному кроці ділить утворюючі групи за ступенем їхньої гетерогенності [14]. В обох випадках результатом роботи алгоритму є деревоподібна структура, або дендрограма.

Окремі версії агломеративної ієрархічної процедури відрізняються правилами обчислення відстані між кластерами. Наприклад, алгоритм середнього зв'язку на кожному наступному кроці поєднує два найближчих кластери, розраховуючи середню арифметичну дистанцію між усіма парами об'єктів. До алгоритму середнього зв'язку природно одразу додати ще два із взаємно протилежними тенденціями:

алгоритм поодинокого зв'язку, або "найближчого сусіда", коли відстань між кластерами оцінюється як мінімальна з дистанцій між парами об'єктів, один з яких входить до першого кластеру, а інший – до другого;

та алгоритм повного зв'язку або "далекого сусіда", коли обчислюється відстань між найбільш віддаленими об'єктами [14].

Глибший статистичний сенс у поняття відстані між кластерами вкладається при використанні центроїдного методу та методу мінімуму дисперсії Варда.

Стверджувати, який з методів краще, досить просто – це той, який найчастіше призводить до природної кластеризації. Проте наявні великі розбіжності у розумінні настільки тонкого предмета, як “природність”, змушують визнати цей термін нечітким.

Після створення кластерного рішення зазвичай виникає питання, наскільки воно стійке та статистично значуще. Тут існує емпіричне правило – стійке угруповання має зберігатися при зміні методів кластеризації: наприклад, якщо результати ієрархічного кластерного аналізу мають частку збігів більше 70% з групуванням за методом k -середніх, то припущення про стійкість приймається [19].

У теоретичному плані проблема перевірки адекватності кластеризації не вирішена, принаймні, без використання іншого виду аналізу чи знання належності об'єктів до відповідної групи. Можна виділити п'ять основних методів перевірки адекватності кластеризації:

кофенетична кореляція – не рекомендується та обмежена у використанні;
тести на значущість розбиття даних на кластери (багатомірний дисперсійний аналіз) завжди дають значний результат;

методика повторних (випадкових) вибірок – не доводить обґрунтованості рішення;

тести значущості для ознак, не використаних під час кластеризації – придатні лише за наявності повторних вимірів;

методи Монте-Карло дуже складні та доступні лише досвідченим математикам [14].

Взагалі запропоновано безліч методів та критеріїв оцінки якості результатів кластеризації. Можна виділити кілька підходів до валідації кластерів:

зовнішня валідація, яка полягає у порівнянні підсумків кластерного аналізу із заздалегідь відомим результатом;

відносна валідація, яка оцінює структуру кластерів, змінюючи різні параметри того самого алгоритму (наприклад, число груп k);

внутрішня валідація, яка використовує внутрішню інформацію процесу об'єднання у кластери (якщо зовнішня інформація відсутня);

оцінка стабільності об'єднання у кластери (або спеціальна версія внутрішньої валідації), що використовує методи ресемплінгу [14].

Одна з проблем кластеризації полягає в тому, що методи будуть формувати групи, навіть якщо аналізований набір даних є цілком випадковою структурою. Тому першим завданням валідації, яку рекомендується виконати перед початком кластерного аналізу, є оцінка загальної схильності даних до об'єднання в кластери.

Статистика Хопкінса є однією з індикаторів тенденції до групування. Для її розрахунку створюється псевдо-набір даних, згенерованих випадковим чином на основі розподілу з тим самим стандартним відхиленням, що і оригінальний набір даних. Для кожного спостереження i з n розраховується середня відстань до k найближчих сусідів: w_i між реальними об'єктами та q_i між штучними об'єктами та їх найближчими реальними сусідами [19]. Статистика Хопкінса розраховується за формулою:

$$H_{ind} = \frac{\sum_n w_i}{\sum_n q_i + \sum_n w_i}.$$

Тоді статистика Хопкінса, що перевищує 0.5, буде відповідати нульовій гіпотезі про те, що q_i і w_i подібні, а об'єкти, що групуються, розподілені випадково і однорідно. Розмір $H_{ind} > 0.25$ на 90% впевненості свідчить про наявну тенденцію до групування даних [19].

В даному розділі розглянуто методи кластеризації. Існує велика кількість методів та алгоритмів поділу даних на кластери. Але потрібно зазначити, що єдиного правильного рішення для побудови кластерного аналізу не існує. Для досягнення найкращого результату необхідно експериментувати з вибором методів розрахування відстаней, а іноді змінювати алгоритми для аналізу.

2.4. Методи побудови багатофакторної лінійної моделі

Регресія має ключове значення для формування та втілення основних процесів моделювання в дослідженні даної теми. Парна регресія може дати хороший результат при моделюванні, якщо впливом інших факторів, що впливають на об'єкт дослідження, можна знехтувати. Якщо ж цим впливом знехтувати не можна, то в цьому випадку слід спробувати виявити вплив інших факторів, ввівши їх в модель, тобто побудувати рівняння множинної регресії. Множинна регресія – це модель, де теоретичне (середнє) значення залежної змінної Y розглядається як функція кількох незалежних змінних X_1, X_2, \dots, X_m .

$$Y = f(X_1, X_2, \dots, X_m),$$

де Y – залежна змінна (результативна ознака);

X_m – незалежні або пояснювальні змінні (ознаки-фактори).

В моделюванні регресія – одна із ключових моделей для знаходження закономірностей розвитку. Метою регресії є створення моделей із великим числом вхідних факторів, котрі створюють окремий або загальний вплив на формування кінцевого показника.

Побудова рівняння множинної регресії починається з вирішення питання про специфікацію моделі. Включає в себе два питання: відбір факторів та вибір виду рівняння регресії. При цьому самі фактори X_1, X_2, \dots, X_m , що включаються в модель, повинні відповідати таким вимогам: бути кількісно вимірними; бути

тісно пов'язаними з результативною ознакою; не повинні бути корельованими між собою.

При порушенні третьої вимоги, неможливо визначити індивідуальний вплив окремих регресорів на результат Y , що є досить актуальним для здійснення прогнозів та прийняття управлінських рішень.

Існує ряд підходів подолання сильної міжфакторної кореляції. Найпростіший шлях усунення мультиколінеарності полягає у виключенні з моделі одного або декількох факторів. Інший підхід пов'язаний з перетворенням факторів, при якому зменшується кореляція між ними [26].

Можливі різні види рівнянь множинної регресії: лінійні і нелінійні. Найбільш широко використовується лінійна функція. Модель множинної лінійної регресії описується рівнянням:

$$y = a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m,$$

де коефіцієнти $a_j, j = 1, 2, \dots, m$ – характеризують середню зміну результату зі зміною фактора x_j на одиницю при незмінному значенні інших факторів.

Однак реальне співвідношення між соціально-економічними явищами і процесами далеко не завжди можна виразити лінійними функціями.

У такому випадку вдаються до нелінійних моделей, серед яких найбільш часто застосовуються:

степенева:

$$y = a_0 * x_1^{a_1} * x_2^{a_2} * \dots * x_m^{a_m},$$

де $a_j, j = 1, 2, \dots, m$ – коефіцієнти еластичності, які показують, на скільки відсотків зміниться в середньому результат y зі зміною фактора x_j на 1% при незмінності дії інших факторів:

експоненціальна:

$$y = e^{a_0 + a_1x_1 + a_2x_2 + \dots + a_mx_m};$$

гіперболічна:

$$y = a_0 + \frac{a_1}{x_1} + \frac{a_2}{x_2} + \dots + \frac{a_m}{x_m}.$$

Використовуються й інші нелінійні функції. Правильність вибору форми моделі визначається на етапі верифікації.

Розглянемо лінійну модель множинної регресії. Класичний підхід до оцінювання параметрів лінійної моделі множинної регресії заснований на методі найменших квадратів [27]. При цьому лінійна економетрична модель записується у векторно матричній формі такого вигляду:

$$\hat{Y} = X * \bar{a},$$

де \hat{Y} і \bar{a} – вектори,

X – матриця значень вихідних незалежних ознак.

Тобто вказані елементи із формули подаються у такому вигляді:

$$\hat{Y} = (Y_1, Y_2, \dots, Y_n),$$

$$X = \begin{pmatrix} 1 & x_{11} & x_{21} & \dots & x_{1m} \\ 1 & x_{12} & x_{22} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{nm} \end{pmatrix},$$

$$\bar{a} = (a_0, a_1, a_2, \dots, a_k).$$

Від такого уявлення моделі варто відразу перейти до можливості провести оцінку параметрів. Таку можливість забезпечує метод найменших квадратів [27]. Суть методу найменших квадратів (МНК) – оцінки параметрів такі, що сума квадратів відхилень фактичних значень залежної змінної y_i від розрахункових (теоретичних) y_{x_i} мінімальна:

$$\sum (y_i - \widehat{y}_{x_i})^2 \rightarrow \min$$

$$\varepsilon_i = y_i - \widehat{y}_{x_i}.$$

Тобто, з усієї безлічі побудованих ліній, лінія регресії на графіку вибирається таким чином, щоб сума квадратів відстаней по вертикалі між точками і цією лінією була б мінімальною, виконуючи тим самим умову формули.

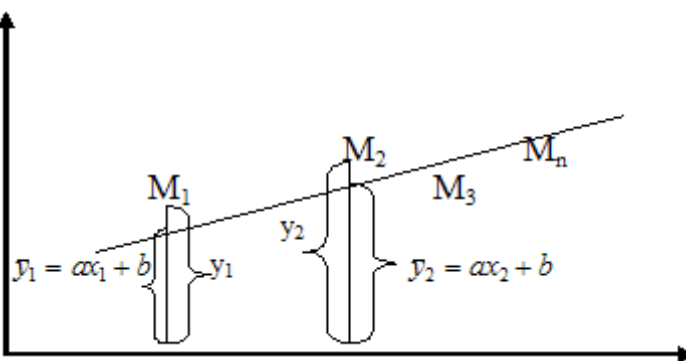


Рис. 2.4. Графічне зображення МНК

Вектор оцінок параметрів \bar{a} лінійної економетричної моделі в матричній формі визначається на основі наступного виразу:

$$\bar{a} = B^{-1} * \overline{YX},$$

$$B = X^T * X,$$

$$\overline{YX} = X^T * Y.$$

Наведені вище вирази ще називають як оператор оцінювання. При цьому значення вектора \bar{a} є рішенням системи нормальних рівнянь МНК.

Потрібно провести аналіз статистичної значимості моделі. Потрібно здійснити перевірку того чи впливає кожен із отриманих параметрів на зміну кінцевого показника. Використаємо для цього критерій Стюдента, тобто t-статистика:

$$t_{a_i} = \frac{a_i}{\sigma_{a_i}},$$

$$\sigma_{a_i} = \sqrt{\sigma_e^2 * B_{jj}^{-1}},$$

$$\sigma_e^2 = \frac{\sum_{i=1}^n e_i^2}{n - m - 1},$$

де σ_e^2 – оцінка дисперсії випадкової величини;
 B_{jj}^{-1} – діагональний елемент матриці B^{-1} .

Далі потрібно порівняти отриманні значення із критичним t_p для числа ступені свободи $k = n - m - 1$ і рівня значущості α .

За виконанням умови, коли емпіричне значення менше табличного, то можна стверджувати про невагомий вплив даного конкретного фактору на вихідну змінну. В протилежному випадку, при більшому емпіричному значенні можна сказати про наявність вагомий впливу на залежну змінну.

Коефіцієнт детермінації – це показник, який визначає частку розкиду залежною змінною Y , пояснюється регресією Y на X . Тому для перевірки адекватності отриманої моделі необхідно обчислити коефіцієнт детермінації за формулою:

$$R^2 = 1 - \frac{\sigma_e^2}{\sigma_y^2},$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}.$$

При наближенні до 1, даний показник відобразить покращену якість підібраної моделі.

Мультиколінеарності – це лінійна залежність між двома або кількома факторними змінними в рівнянні множинної регресії. Якщо така залежність є функціональною, то говорять про повну мультиколінеарність. Якщо ж вона є кореляційною, то має місце часткова мультиколінеарність. Відсутність мультиколінеарності є однією з бажаних передумов класичної лінійної множинної моделі. Це пов'язано з наступними міркуваннями:

у разі повної мультиколінеарності взагалі неможливо побудувати оцінки параметрів лінійної регресії за допомогою МНК;

у разі часткової мультиколінеарності оцінки параметрів регресії можуть бути ненадійними.

Головною причиною виникнення мультиколінеарності є наявність в досліджуваному об'єкті процесів, які одночасно впливають на деякі вхідні змінні, але не враховані в моделі. Це може бути результатом неякісного дослідження предметної області або складності взаємозв'язків параметрів досліджуваного об'єкта. При мультиколінеарності навіть використання деяких факторів уже не має певного розуміння та потреби. При цьому усі оцінки у моделі стають надто чутливими до певних змін у їх структурі.

Для вимірювання мультиколінеарності використовують такі методи:

Знаходження величини визначника матриці $X^T * X$:

$$|X^T X| = \sum_{i=1}^k \lambda_i \cdot \begin{vmatrix} b_{11} - \lambda & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} - \lambda & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n1} & \dots & \dots & b_{nn} - \lambda \end{vmatrix} = 0,$$

Ступінь мультиколінеарності між включеними в модель факторами залежить від того, як близько визначник цієї матриці наближається до нуля.

Мінімальність власного числа матриці $X^T * X$:

$$\lambda_{\min}(X^T * X) = \min_i(\lambda_i) = \lambda'_i.$$

У даному методі сила мультиколінеарності залежить від того, на скільки менше показник λ'_1 .

Міра обумовленості матриці $X^T * X$ по Нейману-Голдштейну:

$$\frac{\lambda_{\max}(X^T * X)}{\lambda_{\min}(X^T * X)}; \lambda_{\max} = \max_i(\lambda_i) = \lambda'_k.$$

У цьому методі сила мультиколінеарності залежить від того, як близько показник $\frac{\lambda'_k}{\lambda'_1}$ до нескінченності. Для даного відношення існує шкала, що дозволяє дати якісну оцінку ступеня мультиколінеарності за величиною міри Неймана-Голдштейна. За цією шкалою, якщо відношення більше 30, то мультиколінеарність середнього ступеня, а якщо більше 100, то мультиколінеарність велика.

Метод Феррара-Глобера. При його застосуванні реалізується пошук парних кореляції між всіма факторами. При цьому вдається виокремити оцінку загальної мультиколінеарності (між усіма ознаками) та часткової (лише для факторних змінних).

Способи усунення мультиколінеарності:

можна збільшити обсяг вибірки (чим більше даних, тим менше дисперсії оцінок МНК);

замінити мультиколінеарні фактори або перетворити їх – скласти, розділити, замінити їх середнім, але при цьому отримані факторні змінні повинні мати економічний сенс;

часто виключають змінні, високореліровані з іншими. Однак якщо будь-які чинники були включені в модель на основі рекомендацій економічної теорії, їх не слід виключати.

В разі, коли слід позбутись від мультиколінеарності, застосовують такі правила:

виключення одного з двох залежних факторних змінних;

застосування методу головних компонент, котрий перетворює вхідні фактори на некорельовані величини;

застосування методу рідж-регресії;

пошук додаткової інформації для моделі.

Автокореляція (послідовна кореляція) визначається як кореляція між спостерігаючими показниками, впорядкованими в часі (часові ряди) або в просторі (перехресні ряди).

При застосуванні МНК у моделях логічно отримувати такі наслідки автокореляції: оцінки параметрів втрачають характеристику ефективності; дисперсії оцінок стають зміщеними; дисперсія регресії знижується; розрахунки t-статистики та F-статистики можуть бути знайдені невірно.

Існують такі методи виявлення автокореляції: графічний аналіз залишків; критерій Дарбіна-Уотсона; метод рядів. Найбільш відомим критерієм для оцінки автокореляції залишків є показник Дарбіна-Уотсона:

$$DW = \frac{\sum_{t=2}^T (e_t - e_{t-1})^2}{\sum_{t=1}^T e_t^2}.$$

Алгоритм розрахунку даного критерію:

Крок 1. У побудованій регресії шукаються відхилення для кожного із результуючих спостережень e_t .

Крок 2. Обчислюється значення DW.

Крок 3. По таблиці критичних точок Дарбіна -Уотсона визначаються два числа d_l та d_u і здійснюються висновки за правилами:

$0 \leq DW \leq d_l$ – позитивна автокореляція;

$d_l \leq DW \leq d_u$ – невизначеність автокореляції;

$d_u \leq DW \leq 4 - d_u$ – відсутність автокореляції;

$4 - d_u \leq DW \leq 4 - d_l$ – невизначеність автокореляції;

$4 - d_l \leq DW \leq 4$ – від'ємна автокореляція.

Одне з важливих припущень при побудові регресійної моделі полягає в тому, що випадкові помилки моделі некоррельовані між собою й мають постійну дисперсію. Розглянемо випадок порушення передумов класичної регресійної моделі, що стосується сталості дисперсії кожної випадкової величини [1]. Якщо дисперсія залишків стала для кожного спостереження, то це явище називається гомоскедастичністю:

$$M\varepsilon_i^2 = D\varepsilon_i = \sigma_\varepsilon^2 = \text{const}, i = 1, n.$$

Якщо це припущення не задовольняється в якомусь окремому випадку, то наявна неоднорідність спостережень, тобто гетероскедастичність (помилки некорельовані, але мають несталу дисперсію) [1]. Якщо дисперсія залишків змінюється для кожного спостереження або групи спостережень, то це явище називається гетероскедастичністю:

$$M\varepsilon_i^2 = D\varepsilon_i = \sigma_\varepsilon^2 = \text{const}, i = 1, n.$$

Варто розуміти, що якщо незважаючи на гетероскедастичність будуть використовуватися звичайні процедури перевірки гіпотез (t-критерій, F -критерій), то висновки можуть бути неправильними. Гетероскедастичність є суттєвою проблемою, а тому потрібно вміти з'ясувати її наявність [1].

Існують різноманітні тести на виявлення гетероскедастичності: тест Глейзера, тест Парка, Тест Уайта та інші. У практичній частині буде застосований тест Глейзера на наявності гетероскедастичності [1].

Тест Глейзера (Glejser test) – це статистичний тест, що дозволяє оцінити наявність певного виду гетероскедастичності випадкових помилок регресійній моделі [1]. Тест заснований на такій моделі можливої залежності стандартного відхилення випадкової помилки моделі σ_t від деякого фактора x_{jt} :

$$\sigma_t = a_0 + a_1 x_{jt}^v + u_t.$$

Нульова гіпотеза полягає в рівності коефіцієнта a_1 нулю (відсутність гетероскедастичності даного виду). Якщо у тесті відкидається нульова гіпотеза, то гетероскедастичність даного виду визнається статистично значущою. Якщо нульова гіпотеза не відкидається, то швидше за все гетероскедастичності даного виду немає у моделі (проте, це не виключає можливості гетероскедастичності іншого виду) [1].

Розглянемо алгоритм тесту Глейзера.

1 крок. За допомогою звичайного МНК оцінюється вихідна регресійна модель і визначаються залишки регресії $e_t = Y_t - \hat{Y}_t$:

$$Y_t = a_i x_{it} + e_t.$$

2 крок. Побудова допоміжних регресійних функцій, що характеризують залежність величини залишків за модулем від пояснюючої змінної x_{jt} , яка може зумовити зміну дисперсії залишків, виду:

$$e_t = a_0 + a_1 x_{jt}^v + u_t.$$

де v – величина, яка може набувати значення (-1; -0.5; 0.5; 1; 2).

3 крок. Для кожного значення v перевіряється статистична значущість коефіцієнта a_1 за допомогою критерію Стьюдента або еквівалентного йому в даному випадку F-тесту на значущість допоміжної регресії в цілому. Якщо для деяких v коефіцієнт a_1 визнається значущим (тестова статистика більше критичного значення), то гетероскедастичність даного виду визнається значущою і вибирається модель із тим значенням v , для якого коефіцієнт a_1 найбільш значущий (з найбільшим значенням тестової статистики). Перевага цього методу полягає в можливості розрізнити випадок чистої і змішаної гетероскедастичності [1]. Можливі чотири випадки, які наведено у табл.2.1.

Таблиця 2.1

Висновки за тестом Глейзера

Оцінка	a_1 - статистично значуща	a_1 - статистично незначуща
a_0 - статистично значуща	Чиста і змішана гетероскедастичність	Змішана гетероскедастичність
a_0 - статистично незначуща	Чиста гетероскедастичність	Гетероскедастичність відсутня

Нульова гіпотеза про відсутність гетероскедастичності полягає в рівності коефіцієнта a_1 нулю. Відхилення цієї гіпотези означає наявність гетероскедастичності зазначеного виду, прийняття нульової гіпотези означає, що гетероскедастичність даного виду відсутня [1].

РОЗДІЛ 3. РЕАЛІЗАЦІЯ МОДЕЛЕЙ ОЦІНКИ ТА АНАЛІЗУ ВПЛИВУ ЦИФРОВИХ ТЕХНОЛОГІЙ НА ЕКОНОМІКУ КРАЇН

3.1. Аналіз показників країн за допомогою моделей панельних даних

Для аналіз впливу факторних змінних на результуючі побудуємо моделі панельних даних. Взято дані для 70 країн світу з 2012 року по 2020 рік. Частину даних зображено в таблиці 3.1. Повний об'єм даних можна переглянути в додатку А.

Таблиця 3.1

Вихідні дані (фрагмент)

Рік	Країна	ВВП на душу населення	Рівень безробіття	ВНД на душу населення	E-Government Index	E-Participation Index	Online Service Index	Human Capital Index	Telecommunication Infrastructure Index	Internet users (%)
		Y1	Y2	Y3	X1	X2	X3	X4	X5	X6
2012	VietNam	1443.50	2.00	4840.00	0.52	0.11	0.42	0.74	0.40	36.80
2012	Uzbekistan	1814.47	4.90	5780.00	0.51	0.24	0.50	0.83	0.21	23.60
2012	Uruguay	12981.14	6.30	16830.00	0.63	0.18	0.55	0.90	0.44	54.45
2012	Ukraine	3153.74	7.50	9190.00	0.57	0.16	0.42	0.92	0.35	35.27
2012	Turkey	12128.30	8.40	20620.00	0.53	0.05	0.46	0.77	0.35	45.13
2012	Tunisia	4146.10	16.70	9790.00	0.48	0.37	0.48	0.68	0.29	41.44
2012	Thailand	5437.88	0.70	14190.00	0.51	0.32	0.51	0.78	0.24	26.46
2012	Switzerland	77844.19	2.90	61140.00	0.81	0.34	0.67	0.89	0.88	85.20
2012	Sweden	53432.64	8.10	46470.00	0.86	0.68	0.84	0.91	0.82	93.18
2012	Sudan	1825.58	14.80	3590.00	0.26	0.08	0.25	0.46	0.07	21.00
2012	Spain	29235.83	24.80	31470.00	0.78	0.50	0.76	0.94	0.63	69.81
2012	Slovenia	22989.93	8.90	28590.00	0.75	0.21	0.67	0.93	0.65	68.35
2012	Singapore	50143.78	2.00	78420.00	0.85	0.95	1.00	0.85	0.69	72.00
2012	Serbia	5886.89	24.60	13480.00	0.63	0.24	0.58	0.85	0.47	48.10

R – це мова програмування та обчислювальне середовище для обробки інформації сучасними статистичними методами. Працювати з R зручніше в

RStudio. Подальші розрахунки та побудова моделей буде виконуватись у RStudio.

Завантажимо дані по країнам за роками в RStudio та пакет даних для маніпуляцій з даними.

```
1 library("knitr")
2 library("texreg")
3 library("tidyverse")
4 library("plm")
```

Рис. 3.1. Пакет даних в RStudio

Позначимо набір даних як панельні дані на рис. 3.2. Країни позначимо як індивідуальний індекс (i) в формулі, а роки – часовий індекс (t).

```
h <- pdata.frame(Panel,
                 index = c("Country", "Year"),
                 row.names = TRUE)
```

Рис. 3.2. Формула перетворення даних в панельні дані

Для аналізу впливу цифрових технологій на економіку країн потрібно побудувати та проаналізувати моделі з фіксованими ефектами (Fixed Effects), випадковими ефектами (Random Effects) та об'єднаної моделі регресії (Pooling Regression) для.

Побудуємо та проаналізуємо вплив факторних змінних (індекс розвитку електронного уряду, індекс електронної участі, індекс розвитку онлайн-сервісів, індекс людського капіталу, індекс телекомунікаційної інфраструктури, відсоток користувачів Інтернетом) на ВВП на душу населення країн. Результати аналізу побудованих моделей зображені на рис. 3.3 – 3.5.

```

> m.re <- plm(Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "random")
> summary(m.re)
Oneway (individual) effect Random Effect Model
(Swamy-Arora's transformation)

Call:
plm(formula = Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "random")

Balanced Panel: n = 70, T = 9, N = 630

Effects:
              var      std.dev share
idiosyncratic 9025072      3004 0.046
individual    187409365     13690 0.954
theta: 0.927

Residuals:
      Min.   1st Qu.   Median     3rd Qu.    Max.
-40755.27 -1306.74  -542.23   841.26  17860.49

Coefficients:
              Estimate Std. Error z-value Pr(>|z|)
(Intercept)  8.2943e+03  4.3398e+03  1.9112 0.055977 .
X1           -2.7844e+06  9.2052e+06 -0.3025 0.762283
X2            5.2690e+03  1.6591e+03  3.1758 0.001494 **
X3            9.2702e+05  3.0684e+06  0.3021 0.762563
X4            9.4286e+05  3.0687e+06  0.3073 0.758650
X5            9.2327e+05  3.0681e+06  0.3009 0.763469
X6            3.5862e+01  2.9060e+01  1.2341 0.217168
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    7114500000
Residual Sum of Squares: 6711200000
R-Squared:                0.056693
Adj. R-Squared:          0.047609
Chisq: 37.4428 on 6 DF, p-value: 1.4433e-06

```

Рис. 3.3. Результат аналізу моделі з випадковими ефектами для Y1

```

> m.fe <- plm(Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "within")
> summary(m.fe)
Oneway (individual) effect within Model

Call:
plm(formula = Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "within")

Balanced Panel: n = 70, T = 9, N = 630

Residuals:
      Min.   1st Qu.   Median     3rd Qu.    Max.
-41885.077 -514.539   -40.973   583.570  15147.042

Coefficients:
              Estimate Std. Error t-value Pr(>|t|)
X1  -2.5724e+06  8.4335e+06 -0.3050 0.7604597
X2   5.2906e+03  1.5291e+03  3.4600 0.0005818 ***
X3   8.5715e+05  2.8112e+06  0.3049 0.7605508
X4   8.6228e+05  2.8114e+06  0.3067 0.7591808
X5   8.5272e+05  2.8109e+06  0.3034 0.7617245
X6   4.7862e-01  2.7037e+01  0.0177 0.9858826
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    5320200000
Residual Sum of Squares: 4999900000
R-Squared:                0.060205
Adj. R-Squared:          -0.067024
F-statistic: 5.91504 on 6 and 554 DF, p-value: 5.3301e-06

```

Рис. 3.4. Результат аналізу моделі з фіксованими ефектами для Y1

```

> m.pooled <- plm(Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "pooling")
> summary(m.pooled)
Pooling Model

Call:
plm(formula = Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "pooling")

Balanced Panel: n = 70, T = 9, N = 630

Residuals:
    Min.    1st Qu.    Median     3rd Qu.     Max.
-50362.03 -10342.01  -822.41   7378.23   70776.76

Coefficients:
              Estimate Std. Error t-value Pr(>|t|)
(Intercept) -2.2057e+04  4.3732e+03  -5.0436 6.001e-07 ***
X1           8.3767e+07  3.8536e+07   2.1737 0.030104 *
X2          -2.8245e+04  5.8094e+03  -4.8621 1.472e-06 ***
X3          -2.7897e+07  1.2846e+07  -2.1717 0.030256 *
X4          -2.7931e+07  1.2846e+07  -2.1744 0.030053 *
X5          -2.7854e+07  1.2845e+07  -2.1685 0.030496 *
X6           2.1336e+02  7.4363e+01   2.8692 0.004255 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 3.4245e+11
Residual Sum of Squares: 1.5174e+11
R-Squared: 0.55691
Adj. R-Squared: 0.55264
F-statistic: 130.504 on 6 and 623 DF, p-value: < 2.22e-16

```

Рис. 3.5. Результат аналізу об'єднаної моделі регресії для Y1

Проаналізуємо коефіцієнт детермінації в моделях. У моделях з випадковими ефектами і фіксованими ефектами коефіцієнт менше 0.1, це вказує на те, що всі вхідні елементи дуже слабо (за шкалою Чеддока) пояснюють долю варіації результуючого показника. У об'єднаній моделі регресії коефіцієнт детермінації більше 0.5, це вказує на те, що всі вхідні елементи помітно (за шкалою Чеддока) пояснюють долю варіації результуючого показника. Критерій Фішера вказує, що моделі є адекватними та значущими. Порівнявши критерій Стюдена можна зробити висновок, що в моделях з випадковими ефектами та фіксованими ефектами на Y1 (ВВП на душу населення) істотно впливає X2 (індекс електронної участі), у об'єднаній моделі регресії істотно впливають всі показники, особливо великий вплив має X2 (індекс електронної участі) та X6 (відсоток користувачів Інтернетом).

Проведемо три теста для порівняння моделей. Порівняємо модель з фіксованими ефектами та об'єднану модель регресії за допомогою F-теста. Результат на рис. 3.6.

```
> pFtest(m.fe, m.pooled)

      F test for individual effects

data:  Y1 ~ X1 + X2 + X3 + X4 + X5 + X6
F = 235.64, df1 = 69, df2 = 554, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Рис. 3.6. Результат порівняння моделі з фіксованими ефектами та об'єднаної моделі регресії за допомогою F-теста для Y1

Так як р-значення менше 0.05, то можна зробити висновок, що модель з фіксованими ефектами є найкращим вибором.

Порівняємо модель з фіксованими ефектами та модель з випадковими ефектами за допомогою теста Хаусмана. Результати на рис. 3.7.

```
> phtest(m.fe, m.re)

      Hausman Test

data:  Y1 ~ X1 + X2 + X3 + X4 + X5 + X6
chisq = 192.61, df = 6, p-value < 2.2e-16
alternative hypothesis: one model is inconsistent
```

Рис. 3.7. Результат порівняння моделі з фіксованими ефектами та моделі з випадковими ефектами за допомогою теста Хаусмана для Y1

Так як р-значення менше 0.05, то можна зробити висновок, що потрібно обрати модель з фіксованими ефектами.

Порівняємо модель з випадковими ефектами та об'єднану модель регресії за допомогою теста Бройша-Пагана, який є одним з тестів множників Лагранжа. Результати на рис. 3.8.

```
> plmtest(m.re, type = "bp")  
  
Lagrange Multiplier Test - (Breusch-Pagan) for balanced panels  
  
data: Y1 ~ X1 + X2 + X3 + X4 + X5 + X6  
chisq = 1454.4, df = 1, p-value < 2.2e-16  
alternative hypothesis: significant effects
```

Рис. 3.8. Результат порівняння моделі з випадковими ефектами та об'єднаної моделі регресії теста Бройша-Пагана для Y1

Так як р-значення менше 0.05, то можна зробити висновок, що гіпотеза про вірність об'єднаної моделі регресії відкидається. Потрібно обрати модель з випадковими ефектами.

Побудувавши моделі з фіксованими ефектами, випадковими ефектами та об'єднану модель регресії виявлено які факторні показники істотно впливаю на ВВП на душу населення. Також виявлено, що моделі є адекватними та значимими, об'єднана модель регресії краще використовувати для аналізу, так як всі вхідні елементи помітно пояснюють долю варіації результуючого показника. За результатами порівняльних тестів можна зробити висновок, що для аналізу даних краще обрати модель з фіксованими ефектами.

Побудуємо та проаналізуємо вплив факторних змінних на рівень безробіття країн. Результати аналізу побудованих моделей зображені на рис. 3.9 – 3.11.

```

> m.re <- plm(Y2 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "random")
> summary(m.re)
Oneway (individual) effect Random Effect Model
(Swamy-Arora's transformation)

Call:
plm(formula = Y2 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "random")

Balanced Panel: n = 70, T = 9, N = 630

Effects:
              var std.dev share
idiosyncratic 4.518  2.125 0.179
individual    20.717  4.552 0.821
theta: 0.8462

Residuals:
      Min.  1st Qu.  Median  3rd Qu.  Max.
-6.95507 -1.23142 -0.35844  0.93533 10.57586

Coefficients:
              Estimate Std. Error z-value Pr(>|z|)
(Intercept) 1.3196e+01  2.2203e+00  5.9434 2.792e-09 ***
X1           5.7234e+03  5.9435e+03  0.9630  0.3356
X2          -6.1545e+00  1.0537e+00 -5.8406 5.201e-09 ***
X3          -1.9058e+03  1.9812e+03 -0.9619  0.3361
X4          -1.9149e+03  1.9813e+03 -0.9665  0.3338
X5          -1.9038e+03  1.9810e+03 -0.9611  0.3365
X6           1.4153e-02  1.8044e-02  0.7844  0.4328
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total sum of squares: 3116.3
Residual sum of squares: 2814.5
R-squared: 0.096834
Adj. R-squared: 0.088136
Chisq: 66.7959 on 6 DF, p-value: 1.8529e-12

```

Рис. 3.9. Результат аналізу моделі з випадковими ефектами для Y2

```

> m.fe <- plm(Y2 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "within")
> summary(m.fe)
Oneway (individual) effect within Model

Call:
plm(formula = Y2 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "within")

Balanced Panel: n = 70, T = 9, N = 630

Residuals:
    Min.    1st Qu.    Median    3rd Qu.    Max.
-7.746649 -0.898540 -0.016671  0.778820 10.699208

Coefficients:
            Estimate Std. Error t-value Pr(>|t|)
x1  5.4921e+03  5.9667e+03  0.9205  0.3577
x2 -6.4774e+00  1.0818e+00 -5.9875 3.833e-09 ***
x3 -1.8284e+03  1.9889e+03 -0.9193  0.3583
x4 -1.8373e+03  1.9891e+03 -0.9237  0.3560
x5 -1.8266e+03  1.9887e+03 -0.9185  0.3588
x6  2.1187e-02  1.9129e-02  1.1076  0.2685
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    2809.7
Residual Sum of Squares: 2502.7
R-Squared:              0.10927
Adj. R-Squared:        -0.011319
F-statistic: 11.3266 on 6 and 554 DF, p-value: 5.9031e-12

```

Рис. 3.10. Результат аналізу моделі з фіксованими ефектами для Y2

```

> m.pooled <- plm(Y2 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "pooling")
> summary(m.pooled)
Pooling Model

Call:
plm(formula = Y2 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "pooling")

Balanced Panel: n = 70, T = 9, N = 630

Residuals:
    Min.    1st Qu.    Median    3rd Qu.    Max.
-9.0571 -3.3408 -1.1040  2.0139 18.8124

Coefficients:
            Estimate Std. Error t-value Pr(>|t|)
(Intercept) 1.2103e+01  1.3688e+00  8.8426 <2e-16 ***
x1          1.0599e+04  1.2061e+04  0.8787  0.3799
x2          -1.5724e+00  1.8182e+00 -0.8648  0.3875
x3          -3.5363e+03  4.0205e+03 -0.8796  0.3794
x4          -3.5326e+03  4.0205e+03 -0.8786  0.3799
x5          -3.5335e+03  4.0203e+03 -0.8789  0.3798
x6          -7.9992e-03  2.3275e-02 -0.3437  0.7312
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    15768
Residual Sum of Squares: 14864
R-Squared:              0.057309
Adj. R-Squared:        0.04823
F-statistic: 6.31228 on 6 and 623 DF, p-value: 1.8546e-06

```

Рис. 3.11. Результат аналізу об'єднаної моделі регресії для Y2

Проаналізувавши коефіцієнт детермінації трьох моделей можна побачити, що показник менше 0.1, це вказує на те, що всі вхідні елементи слабо пояснюють долю варіації результуючого показника. Критерій Фішера вказує, що моделі є адекватними та значимими. Порівнявши критерій Стюдена можна зробити висновок, що в моделях з випадковими ефектами та фіксованими ефектами на Y2 (рівень безробіття) істотно впливає X2 (індекс електронної участі), у об'єднаній моделі регресії всі показники не мають істотного впливу на результуючий показник.

Проведемо три теста для порівняння моделей. Порівняємо модель з фіксованими ефектами та об'єднану модель регресії за допомогою F-теста. Результат на рис. 3.12.

```
> pFtest(m.fe, m.pooled)

          F test for individual effects

data:  Y2 ~ X1 + X2 + X3 + X4 + X5 + X6
F = 39.657, df1 = 69, df2 = 554, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Рис. 3.12. Результат порівняння моделі з фіксованими ефектами та об'єднаній моделі регресії для Y2

Так як р-значення менше 0.05, то можна зробити висновок, що модель з фіксованими ефектами є найкращим вибором.

Порівняємо модель з фіксованими ефектами та модель з випадковими ефектами за допомогою теста Хаусмана. Результати на рис. 3.13.

```
> phtest(m.fe, m.re)

          Hausman Test

data:  Y2 ~ X1 + X2 + X3 + X4 + X5 + X6
chisq = 6.0189, df = 6, p-value = 0.4211
alternative hypothesis: one model is inconsistent
```

Рис. 3.13. Результат порівняння моделі з фіксованими ефектами та моделі з випадковими ефектами для Y2

Так як р-значення більше 0.05, то можна зробити висновок, що потрібно обрати модель з випадковими ефектами.

Порівняємо модель з випадковими ефектами та об'єднану модель регресії за допомогою теста Бройша-Пагана, який є одним з тестів множників Лагранжа. Результати на рис. 3.14.

```
> plmtest(m.re, type = "bp")

Lagrange Multiplier Test - (Breusch-Pagan) for balanced panels

data:  Y2 ~ X1 + X2 + X3 + X4 + X5 + X6
chisq = 1589.3, df = 1, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Рис. 3.14. Результат порівняння моделі з випадковими ефектами та об'єднаної моделі регресії для Y2

Так як р-значення менше 0.05, то можна зробити висновок, що гіпотеза про вірність об'єднаної моделі регресії відкидається. Потрібно обрати модель з випадковими ефектами.

Побудувавши моделі з фіксованими ефектами, випадковими ефектами та об'єднаної моделі регресії виявлено які факторні показники істотно впливають на рівень безробіття. Також виявлено, що моделі є адекватними та значимими, але в побудованих моделях всі вхідні елементи слабо пояснюють долю варіації рівня безробіття. За результатами порівняльних тестів можна зробити висновок, що для аналізу даних краще обрати модель з випадковими ефектами.

Побудуємо та проаналізуємо вплив факторних змінних на ВНД на душу населення країн. Результати аналізу побудованих моделей зображені на рис. 3.15 – 3.17.

```

> m.re <- plm(Y3 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "random")
> summary(m.re)
Oneway (individual) effect Random Effect Model
  (Swamy-Arora's transformation)

Call:
plm(formula = Y3 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "random")

Balanced Panel: n = 70, T = 9, N = 630

Effects:
              var  std.dev share
idiosyncratic 10905023    3302 0.128
individual    74461934    8629 0.872
theta: 0.8735

Residuals:
   Min.  1st Qu.  Median    3rd Qu.    Max.
-9847.57 -2064.79  -498.66  1343.42 29530.46

Coefficients:
              Estimate  Std. Error  z-value  Pr(>|z|)
(Intercept) -1.2124e+04  4.1526e+03  -2.9196  0.003505 **
X1           -6.6871e+06  1.0338e+07  -0.6468  0.517749
X2            1.1555e+04  1.8446e+03   6.2646  3.738e-10 ***
X3            2.2253e+06  3.4462e+06   0.6457  0.518455
X4            2.2665e+06  3.4464e+06   0.6576  0.510765
X5            2.2243e+06  3.4458e+06   0.6455  0.518592
X6            1.2536e+02  3.1841e+01   3.9370  8.252e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    1.1975e+10
Residual Sum of Squares: 8495300000
R-Squared:               0.29058
Adj. R-Squared:          0.28375
ChiSq: 255.182 on 6 DF, p-value: < 2.22e-16

```

Рис. 3.15. Результат аналізу моделі з випадковими ефектами для Y3

```

> m.fe <- plm(Y3 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "within")
> summary(m.fe)
Oneway (individual) effect within Model

Call:
plm(formula = Y3 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "within")

Balanced Panel: n = 70, T = 9, N = 630

Residuals:
    Min.    1st Qu.    Median    3rd Qu.    Max.
-15105.364  -1479.348   -30.955   1328.895   24758.374

Coefficients:
      Estimate Std. Error t-value Pr(>|t|)
x1 -6.4544e+06  9.2703e+06 -0.6962  0.48657
x2  1.2472e+04  1.6808e+03  7.4200 4.431e-13 ***
x3  2.1481e+06  3.0901e+06  0.6951  0.48726
x4  2.1756e+06  3.0904e+06  0.7040  0.48173
x5  2.1458e+06  3.0898e+06  0.6945  0.48766
x6  6.7547e+01  2.9720e+01  2.2728  0.02342 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total sum of Squares:    8334800000
Residual sum of Squares: 6041400000
R-Squared:              0.27516
Adj. R-Squared:         0.17703
F-statistic: 35.0515 on 6 and 554 DF, p-value: < 2.22e-16

```

Рис. 3.16. Результат аналізу моделі з фіксованими ефектами для Y3

```

> m.pooled <- plm(Y3 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "pooling")
> summary(m.pooled)
Pooling Model

Call:
plm(formula = Y3 ~ x1 + x2 + x3 + x4 + x5 + x6, data = h, model = "pooling")

Balanced Panel: n = 70, T = 9, N = 630

Residuals:
    Min.    1st Qu.    Median    3rd Qu.    Max.
-24877.51  -6545.01   -992.19   5617.18   59391.05

Coefficients:
      Estimate Std. Error t-value Pr(>|t|)
(Intercept) -1.7207e+04  3.0663e+03 -5.6115 3.021e-08 ***
x1           3.3336e+07  2.7020e+07  1.2338 0.2177597
x2          -2.0960e+04  4.0733e+03 -5.1457 3.577e-07 ***
x3          -1.1085e+07  9.0068e+06 -1.2307 0.2188863
x4          -1.1117e+07  9.0069e+06 -1.2343 0.2175687
x5          -1.1056e+07  9.0062e+06 -1.2276 0.2200688
x6           1.8974e+02  5.2140e+01  3.6390 0.0002964 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total sum of Squares:    2.3567e+11
Residual sum of Squares: 7.4597e+10
R-Squared:              0.68347
Adj. R-Squared:         0.68043
F-statistic: 224.207 on 6 and 623 DF, p-value: < 2.22e-16

```

Рис. 3.17. Результат аналізу об'єднаної моделі регресії для Y3

Проаналізуємо коефіцієнт детермінації в моделях. У моделях з випадковими ефектами і фіксованими ефектами коефіцієнт менше 0.3. Це вказує на те, що всі вхідні елементи слабо (за шкалою Чеддока) пояснюють долю варіації результуючого показника. У об'єднаній моделі регресії коефіцієнт детермінації менше 0.7, це вказує на те, що всі вхідні елементи помітно (за шкалою Чеддока) пояснюють долю варіації результуючого показника. Критерій Фішера вказує, що моделі є адекватними та значущими. Порівнявши критерій Стюдена можна зробити висновок, що в усіх моделях на Y3 (ВНД на душу населення) істотно впливають X2 (індекс електронної участі) та X6 (відсоток користувачів Інтернетом).

Проведемо три теста для порівняння моделей. Порівняємо модель з фіксованими ефектами та об'єднану модель регресії за допомогою F-тесту. Результат на рис. 3.18.

```
> pFtest(m.fe, m.pooled)

      F test for individual effects

data:  Y3 ~ X1 + X2 + X3 + X4 + X5 + X6
F = 91.111, df1 = 69, df2 = 554, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Рис. 3.18. Результат порівняння моделі з фіксованими ефектами та об'єднаної моделі регресії для Y3

Так як р-значення менше 0.05, то можна зробити висновок, що модель з фіксованими ефектами є найкращим вибором.

Порівняємо модель з фіксованими ефектами та модель з випадковими ефектами за допомогою тесту Хаусмана. Результати на рис. 3.19.

```
> phtest(m.fe, m.re)

      Hausman Test

data:  Y3 ~ X1 + X2 + X3 + X4 + X5 + X6
chisq = 84.824, df = 6, p-value = 3.59e-16
alternative hypothesis: one model is inconsistent
```

Рис. 3.19. Результат порівняння моделі з фіксованими ефектами та моделі з випадковими ефектами для Y3

Так як р-значення менше 0.05, то можна зробити висновок, що потрібно обрати модель з фіксованими ефектами.

Порівняємо модель з випадковими ефектами та об'єднану модель регресії за допомогою теста Бройша-Пагана, який є одним з тестів множників Лагранжа. Результати на рис. 3.20.

```
> plmtest(m.re, type = "bp")

Lagrange Multiplier Test - (Breusch-Pagan) for balanced panels

data:  Y3 ~ x1 + x2 + x3 + x4 + x5 + x6
chisq = 1030.4, df = 1, p-value < 2.2e-16
alternative hypothesis: significant effects
```

Рис. 3.20. Результат порівняння моделі з випадковими ефектами та об'єднаної моделі регресії для Y3

Так як р-значення менше 0.05, то можна зробити висновок, що гіпотеза про вірність об'єднаної моделі регресії відкидається. Потрібно обрати модель з випадковими ефектами.

Побудувавши моделі з фіксованими ефектами, випадковими ефектами та об'єднані моделі регресії виявлено які факторні показники істотно впливають на ВНД на душу населення. Також виявлено, що моделі є адекватними та значимими, об'єднану модель регресії краще використовувати для аналізу, так як всі вхідні елементи помітно пояснюють долю варіації результуючого показника. За результатами порівняльних тестів можна зробити висновок, що для аналізу даних краще обрати модель з фіксованими ефектами.

Отже, за допомогою побудованих моделей панельних даних та F-теста, теста Бройша-Пагана та теста Хаусмана можна зробити висновок, що вплив цифрових технологій на економіку країн наявний, це помітно на таких показниках як ВВП на душу населення й ВНД на душу населення, також на результуючі показники, які відображають економічну діяльність країн, істотний вплив мають індекс електронної участі та відсоток користувачів Інтернетом. Аналіз моделей та

тестів показали які краще моделі використовувати для подальшої діагностики результируючих показників.

3.2. Реалізація кластерного аналізу рівня цифрового та економічного розвитку країн

Тепер відберемо дані, які потрібні нам для кластерного аналізу. Побудуємо кластери країн за значеннями ВВП на душу населення (Y1), ВНД на душу населення (Y3), рівень безробіття (Y2), індекс електронного уряду (X1), індекс електронної участі (X2), індекс онлайн-сервісів (X3), індекс людського капіталу (X4), індекс телекомунікаційної інфраструктури (X5), відсоток користувачів Інтернетом (X6). Дані візьмемо за 2020 рік. Звернемося до бібліотеки `dplyr` та виберемо відповідні стовпці. А потім збережемо їх у базу `d`. Назвемо рядки у базі даних за назвами країн. Результат на рис. 3.21.

```
library(dplyr)
d <- Pane1111 %>% select(Y1:X6)
rownames(d) <- Pane1111$Country
```

Рис. 3.21. Побудова бази даних `d` для кластерного аналізу

Тепер потрібно об'єднати дані в кластери методом `k`-середніх. Так як підсумкові результати чутливі до початкового випадкового вибору центрів груп потрібно багаторазово виконати алгоритм з різним випадковим призначенням початкових центроїдів. Число таких ітерацій становить 25. Результат на рис. 3.22.

```
> library(cluster)
> df.stand <- as.data.frame(scale(d))
> set.seed(5)
> c(kmeans(df.stand, centers = 5, nstart = 1)$tot.withinss,
+   kmeans(df.stand, centers = 5, nstart = 25)$tot.withinss)
[1] 161.8131 157.4573
```

Рис. 3.22. Побудова бази даних `d` для кластерного аналізу

Вдалося за 25 повторів алгоритму дещо зменшити значення критерію оптимальності.

Тепер потрібно визначити кількість кластерів для розбиття країн на групи. За допомогою метода «ліктя» визначимо кількість кластерів. Результат наведено на рис. 3.23.

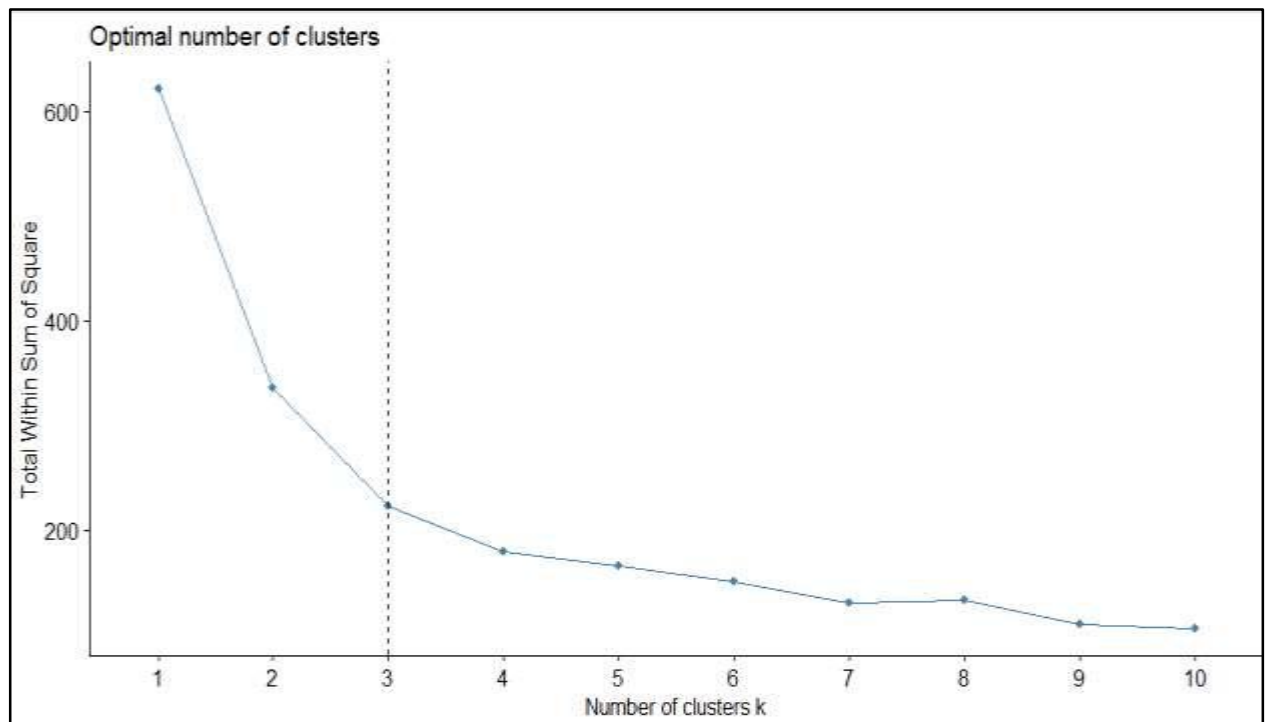


Рис. 3.23. Графік вибору оптимальної кількості кластерів за методом “ліктя”

Даний метод показав, що оптимальна кількість кластерів 3.

Спиратися на результат одного методу не є правильним, тому використаєм інші методи аналізу кількості кластерів.

Спробуємо метод GAP-статистики.

Результати на рис. 3.24 – 3.25.


```

> set.seed(123)
> gap_stat <- clusGap(d, FUN = kmeans, nstart = 10, K.max = 10, B = 70)
Clustering k = 1,2,..., K.max (= 10): .. done
Bootstrapping, b = 1,2,..., B (= 70) [one "." per sample]:
..... 50
..... 70
> print(gap_stat, method = "firstmax")
Clustering Gap statistic ["clusGap"] from call:
clusGap(x = d, FUNcluster = kmeans, K.max = 10, B = 70, nstart = 10)
B=70 simulated reference sets, k = 1..10; spaceHO="scaledPCA"
--> Number of clusters (method 'firstmax'): 5
      logw      E.logw      gap      SE.sim
[1,] 13.30150 13.65244 0.3509451 0.04037236
[2,] 12.74615 13.18195 0.4358049 0.04300740
[3,] 12.37025 12.96627 0.5960168 0.04290192
[4,] 12.16565 12.80666 0.6410105 0.04038472
[5,] 11.95367 12.66136 0.7076910 0.04421326
[6,] 11.83176 12.53537 0.7036150 0.04390278
[7,] 11.70981 12.42328 0.7134698 0.04619447
[8,] 11.59859 12.32695 0.7283645 0.04613511
[9,] 11.49832 12.24167 0.7433509 0.04806962
[10,] 11.39663 12.16181 0.7651772 0.04928608
> fviz_gap_stat(gap_stat)

```

Рис. 3.24. Результат GAP-статистики з вибором кількості кластерів

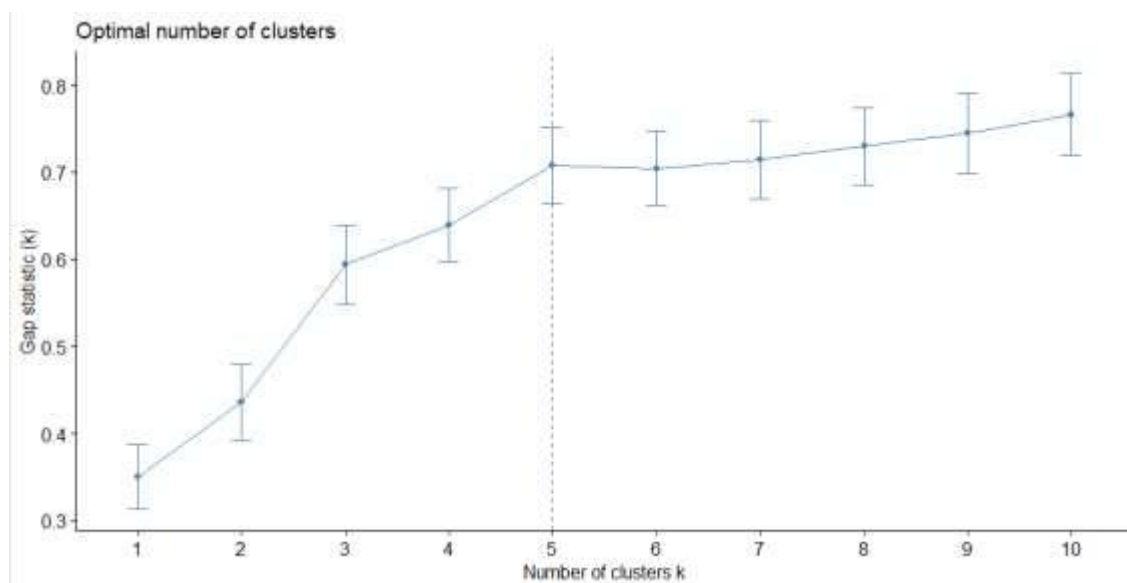


Рис. 3.25. Графік GAP-статистики вибору оптимальної кількості кластерів

Даний метод показав, що оптимальна кількість кластерів 5. Тепер використаємо метод РАМ-статистики для вибору оптимальної кількості кластерів. Результати на рис. 3.26 – 3.27.

```

> set.seed(123)
> gap_stat <- clusGap(d, FUN = pam, K.max = 7, B = 100)
Clustering k = 1,2,..., K.max (= 7): .. done
Bootstrapping, b = 1,2,..., B (= 100) [one "." per sample]:
..... 50
..... 100
> print(gap_stat, method = "firstmax")
Clustering Gap statistic ["clusGap"] from call:
clusGap(x = d, FUNcluster = pam, K.max = 7, B = 100)
B=100 simulated reference sets, k = 1.7; space0="scaledPCA"
--> Number of clusters (method 'firstmax'): 5
      logw   E.logw      gap   SE.sim
[1,] 13.30150 13.65806 0.3565612 0.04494791
[2,] 12.73612 13.19248 0.4563690 0.04717927
[3,] 12.36755 12.98321 0.6156556 0.04780220
[4,] 12.16801 12.82418 0.6561713 0.04739049
[5,] 11.95367 12.68081 0.7271432 0.04910890
[6,] 11.83919 12.55871 0.7195187 0.04610742
[7,] 11.73291 12.44980 0.7168928 0.04675739

```

Рис. 3.26. Результат РАМ-статистики з вибором кількості кластерів

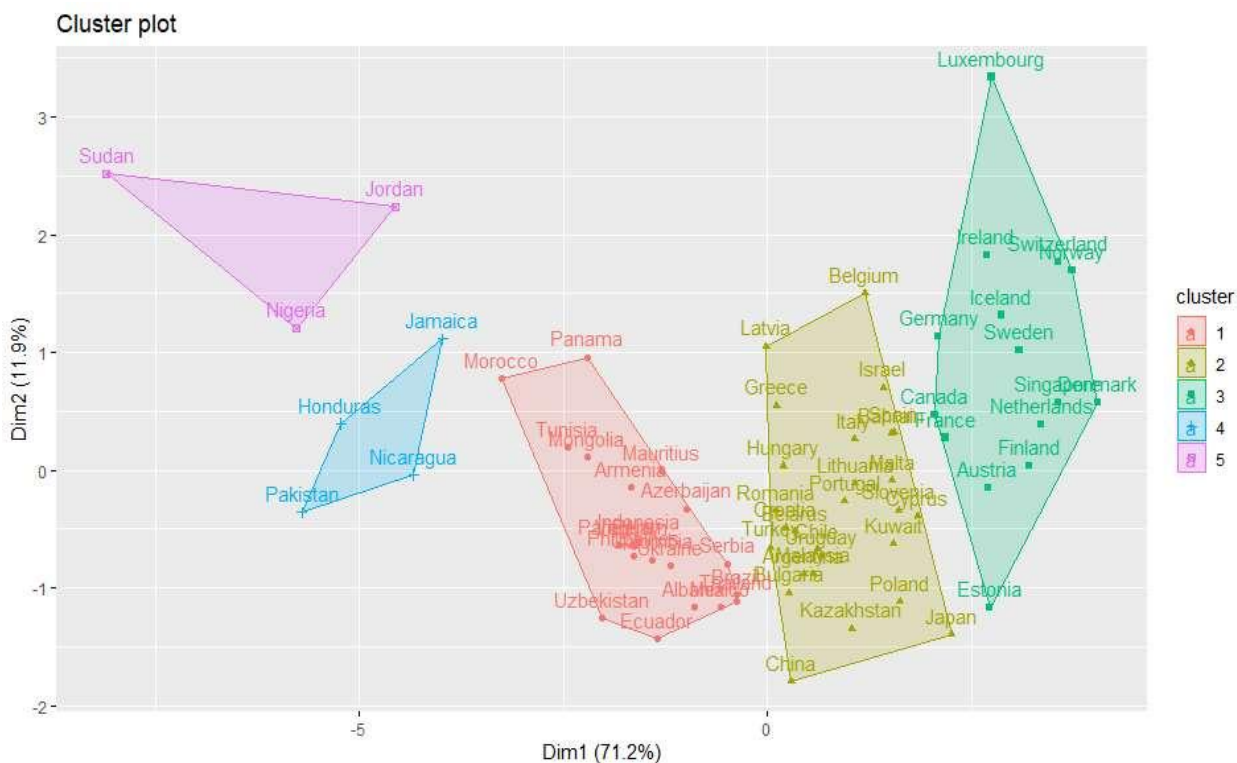


Рис. 3.27. Діаграма розподілу країн за кластерами, отриманими методом РАМ

Даний метод показав, що найкраще розбити дані на 5 кластерів. Тому потрібно використати процедуру Nbclust, яка включає 30 різних індексів знаходження оптимальної кількості кластерів, і видає результат, за який проголосували більшість методів. Результати на рис. 3.28 – 3.30.

```
> res$Best.nc
      KL      CH Hartigan  CCC      Scott Marriot  TrCovw Tracew Friedman  Rubin Cindex
Number_clusters 4.0000 3.0000 3.0000 4.000 3.0000 3.00 3.000 3.0000 3 4.0000 8.0000
Value_Index     5.4668 60.2319 18.8769 0.095 104.4617 50110.92 1019.458 71.1749 8381394 -0.3839 0.2836
      DB Silhouette  Duda PseudoT2  Beale Ratkowsky  Ball PtBiserial Frey McClain  Dunn
Number_clusters 3.0000 2.0000 2.0000 2.0000 3.0000 2.0000 3.000 2.0000 3.0000 1 2.0000 3.0000
Value_Index     0.9971 0.3866 0.7359 12.2035 1.6225 0.4734 94.092 0.573 NA 0.4809 0.1776
      Hubert SDbw  Dindex  SDbw
Number_clusters 0 2.000 0 8.0000
Value_Index     0 1.794 0 0.3234
```

Рис. 3.28. Результати розрахунку за 30 індексами

```
> fviz_nbclust(res)
Among all indices:
=====
* 2 proposed 0 as the best number of clusters
* 1 proposed 1 as the best number of clusters
* 6 proposed 2 as the best number of clusters
* 12 proposed 3 as the best number of clusters
* 3 proposed 4 as the best number of clusters
* 2 proposed 8 as the best number of clusters

Conclusion
=====
* According to the majority rule, the best number of clusters is 3 .
```

Рис. 3.29. Число кластерів обрані більшістю методами

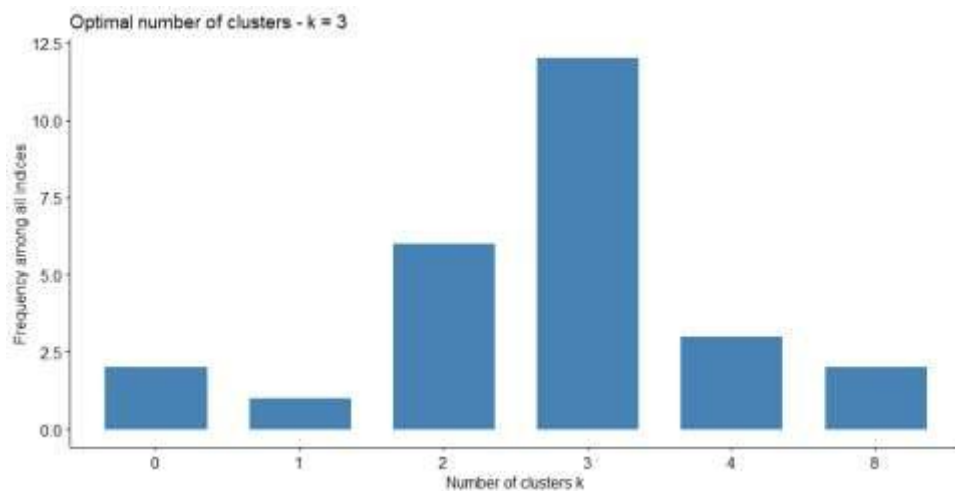


Рис. 3.30. Діаграма вибору оптимальної кількості кластерів за 30 індексами

Отже, за умов невизначеності різні алгоритми можуть породжувати конкуруючі рішення. Використані методи показали різні оптимальні вибори кількості кластерів. За результатами методів кластеризації країни можна розбити на 3 та 5 кластерів.

Тепер реалізуємо ієрархічний кластерний аналіз, використовуючи метод Варда. Побудуємо дендограму та подивимося нам всі можливі кластери, які можна знайти в наших даних. Результат на рис. 3.31.

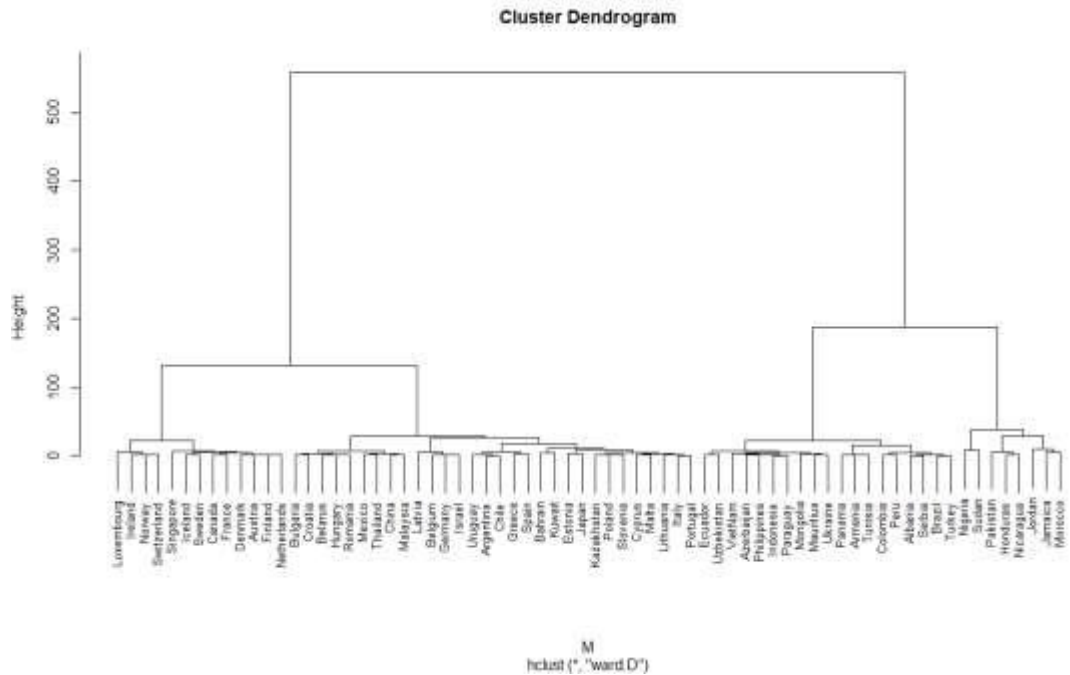


Рис. 3.31. Дендограма

Тепер потрібно виділити кількість кластерів. На основі попередніх аналізів методів вибору оптимальної кількості кластерів оберемо 3 та 5 кластерів. Результати на рис. 3.32 – 3.33.

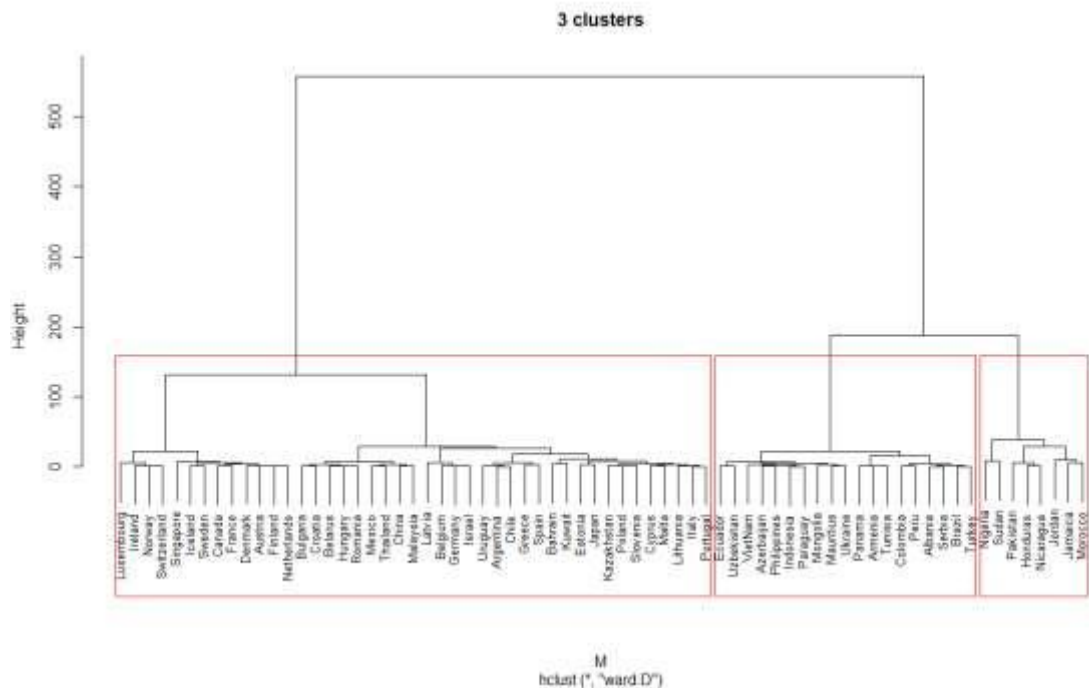


Рис. 3.32. Дендограма з вибором 3 кластерів

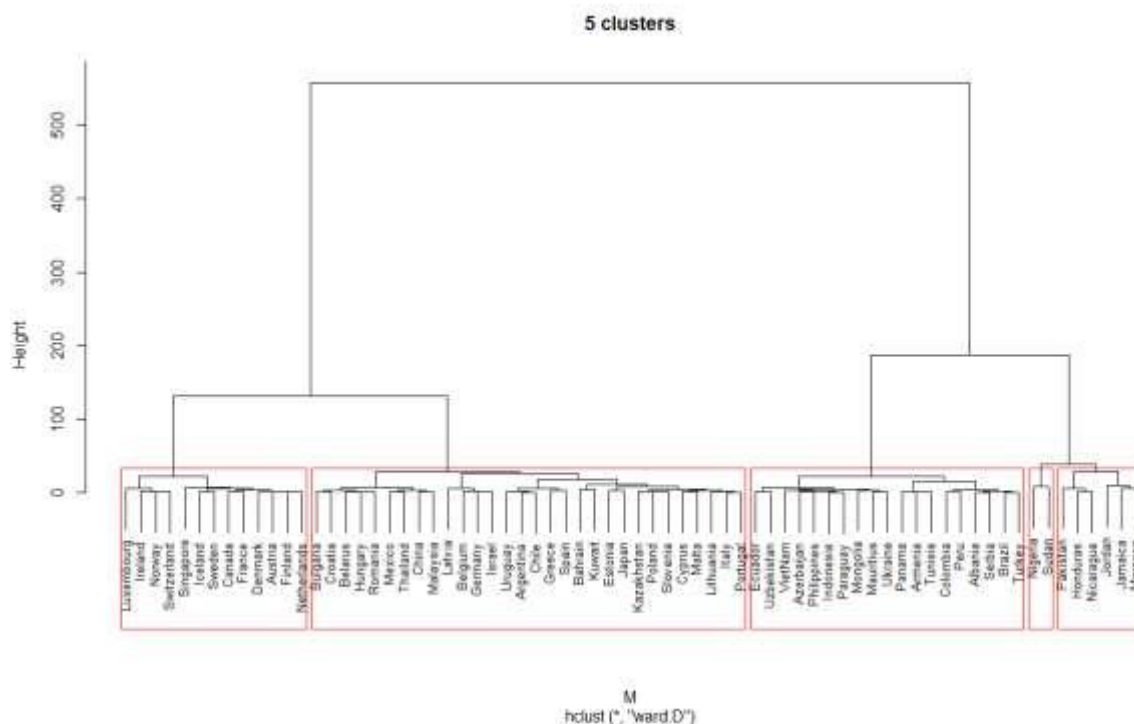


Рис. 3.33. Дендрограма з вибором 5 кластерів

Тепер потрібно здійснити оцінку якості кластеризації. Перевіримо, чи кластери, які ми бачили на дендрограмі, вийшли обґрунтовано, виходячи з даних, а не випадково, через кілька викидів або якогось шуму в даних. На дендрограмі над кожним кластером (над кожною “гілкою” дендрограми) підписано p -value – ймовірність того, що кластер утворений не випадково (наприклад, через наявність у даних якихось дивних значень чи шуму), а обґрунтовано, на основі наявних показників. Значення підписані двома кольорами. Зелене число – це майже незміщене p -value, пораховане на основі вибірок, отриманих в результаті багатоточкового бутстрапування, червоне число – те ж p -value, але пораховане за підсумками звичайного, нормального бутстрапування. Крім того, можна вибрати бажану ймовірність – за якої гіпотеза про те, що кластер утворений не випадково, не відкидається, і виділити такі кластери на дендрограмі. Оберемо ймовірність 90%. Побудуємо графік для перевірки. Результат на рис. 3.34.

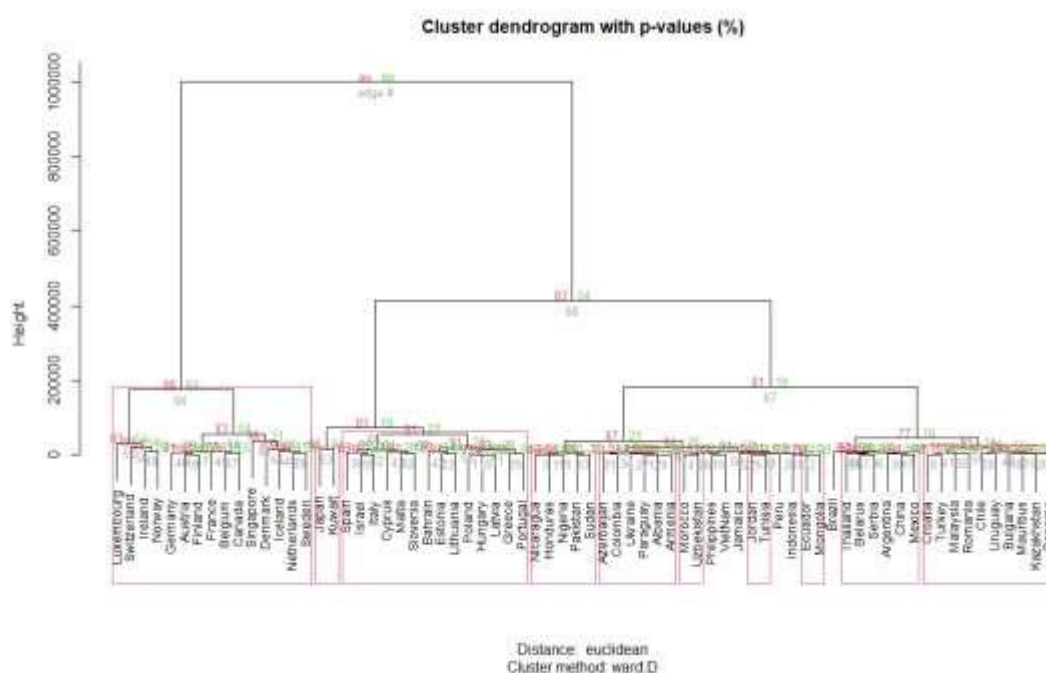


Рис. 3.34. Графік ймовірності випадкових кластерів

Бачимо, що з ймовірністю 90% скільки кластерів є не випадковими.

Випадок з 3 кластерами доволі цікавий, тому використаємо його. Так країни можна поділити на 3 групи:

кластер № 2 – з високим рівнем розвитком цифрових технологій та економіки (Люксембург, Норвегія, Ірландія, Швейцарія, Данія, Сінгапур, Швеція, Нідерланди, Канада, Фінляндія, Ісландія, Австрія, Німеччина, Бельгія, Франція, Ізраїль, Італія, Кіпр, Іспанія, Мальта, Словенія, Португалія, Греція, Естонія, Бахрейн, Литва, Польща, Угорщина, Латвія, Хорватія, Чилі, Уругвай, Румунія, Малайзія, Казахстан, Мексика, Болгарія, Аргентина, Китай, Білорусь, Таїланд, Японія, Кувейт);

кластер № 1 – з середнім рівнем розвитком цифрових технологій та економіки (Туреччина, Бразилія, Панама, Маврикій, Колумбія, Сербія, Перу, Азербайджан, Парагвай, Албанія, Еквадор, Вірменія, Індонезія, Монголія, Туніс, Україна, Філіппіни, Узбекистан, В'єтнам);

кластер № 3 – з низьким рівнем розвитком цифрових технологій та економіки (Ямайка, Йорданія, Марокко, Нігерія, Гондурас, Судан, Нікарагуа, Пакистан).

Кластери вийшли доволі логічними. Витягнемо з отриманого розбиття на кластери мітки для спостережень, щоб зрозуміти, які країни в одному кластері, а які – у різних. Приналежність країн до кластера зображено на рис. 3.35.

```
> groups3 <- cutree(hc, k = 3)
> groups3
```

Albania	Argentina	Armenia	Austria	Azerbaijan	Bahrain
1	2	1	2	1	2
Belarus	Belgium	Brazil	Bulgaria	Canada	Chile
2	2	1	2	2	2
China	Colombia	Croatia	Cyprus	Denmark	Ecuador
2	1	2	2	2	1
Estonia	Finland	France	Germany	Greece	Honduras
2	2	2	2	2	3
Hungary	Iceland	Indonesia	Ireland	Israel	Italy
2	2	1	2	2	2
Jamaica	Japan	Jordan	Kazakhstan	Kuwait	Latvia
3	2	3	2	2	2
Lithuania	Luxembourg	Malaysia	Malta	Mauritius	Mexico
2	2	2	2	1	2
Mongolia	Morocco	Netherlands	Nicaragua	Nigeria	Norway
1	3	2	3	3	2
Pakistan	Panama	Paraguay	Peru	Philippines	Poland
3	1	1	1	1	2
Portugal	Romania	Serbia	Singapore	Slovenia	Spain
2	2	1	2	2	2
Sudan	Sweden	Switzerland	Thailand	Tunisia	Turkey
3	2	2	2	1	1
Ukraine	Uruguay	Uzbekistan	Vietnam		
1	2	1	1		

Рис. 3.35. Країна та її кластер

Побудуємо графік з кластерами та країнами, які відносяться до певного кластера.

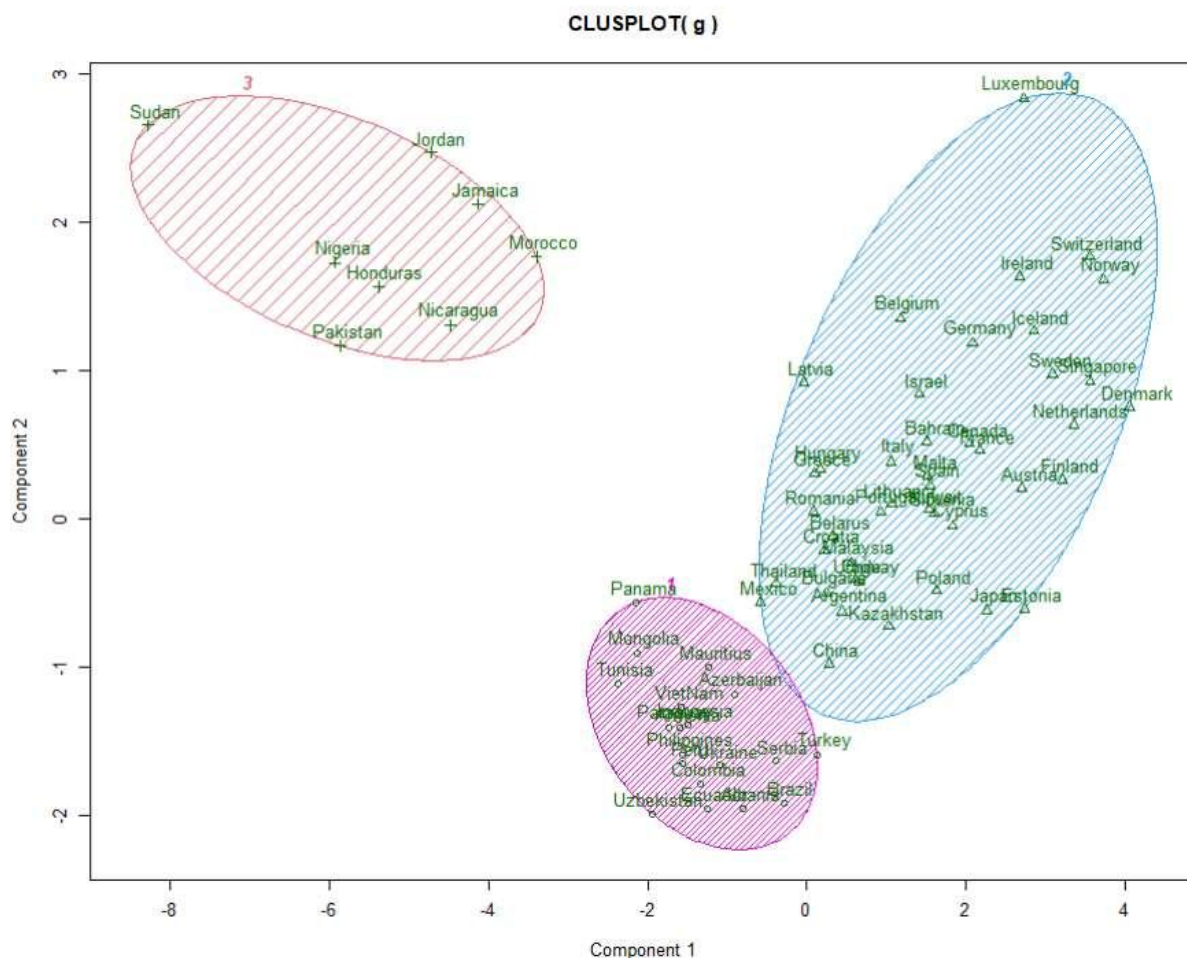


Рис. 3.36. Графік розбиття країн на кластери

Можна прослідкувати, що кластер з країнами з середнім рівнем розвитку економіки та цифрових технологій (рожевий) та кластер з країнами з високим рівнем розвитку економіки та цифрових технологій (блакитний) дуже сильно наближені один до одного.

Побудуємо діаграми розсіювання для пар показників і виділятимемо точки, що відносяться до різних кластерів, різним кольором. Результати деяких пар зображено на рис. 3.37 – 3.41. Результати всіх пар зображено в додатку Б.

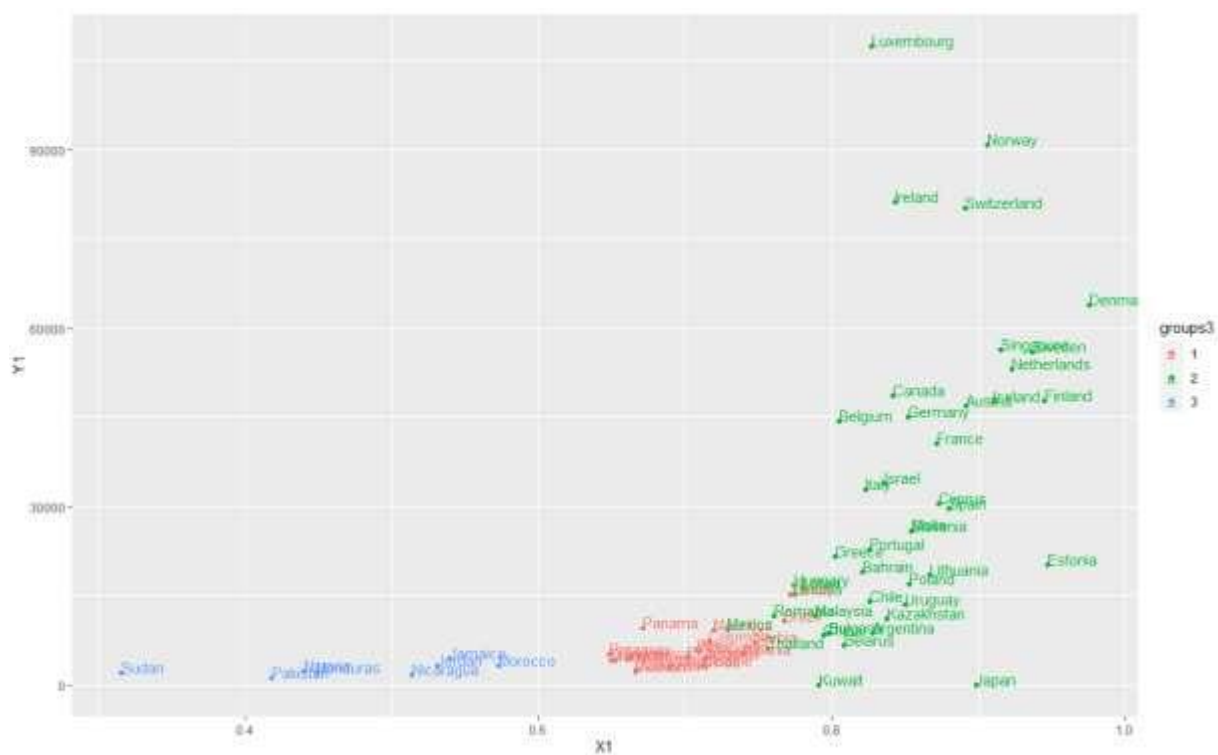


Рис. 3.37. Діаграма розсіювання для пари Y1 та X1

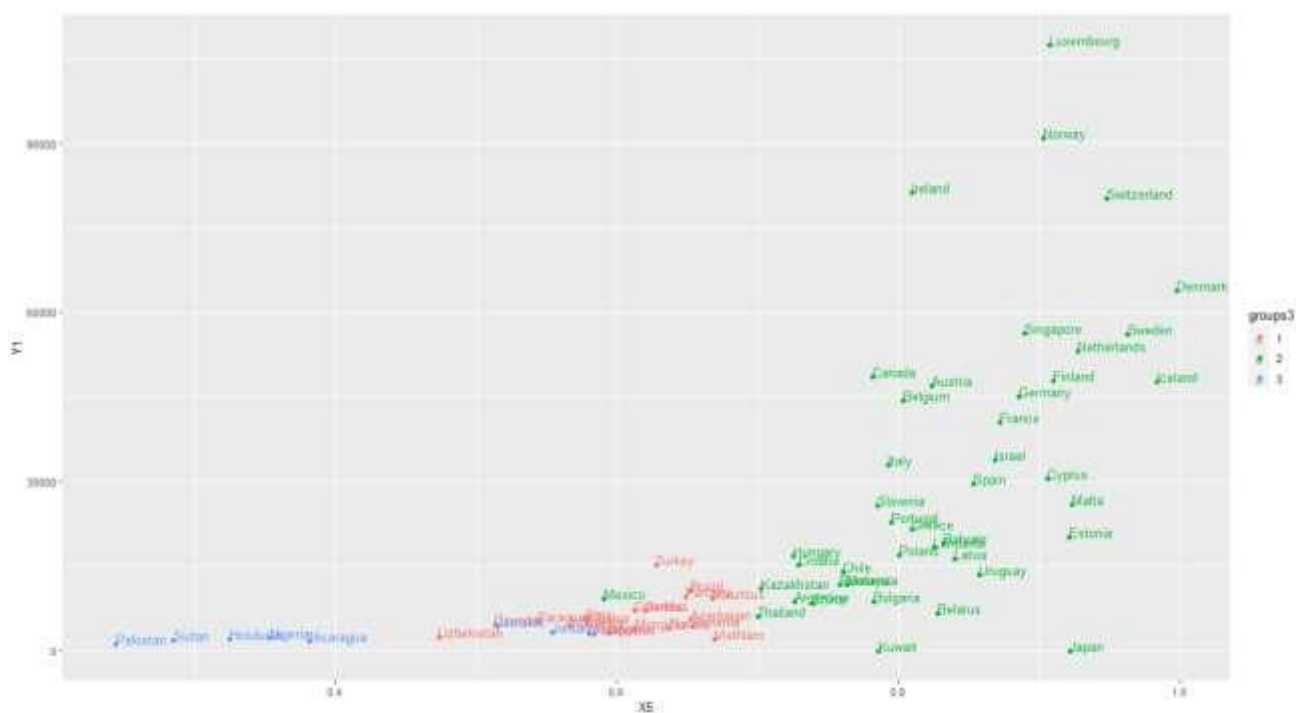


Рис. 3.38. Діаграма розсіювання для пари Y1 та X5

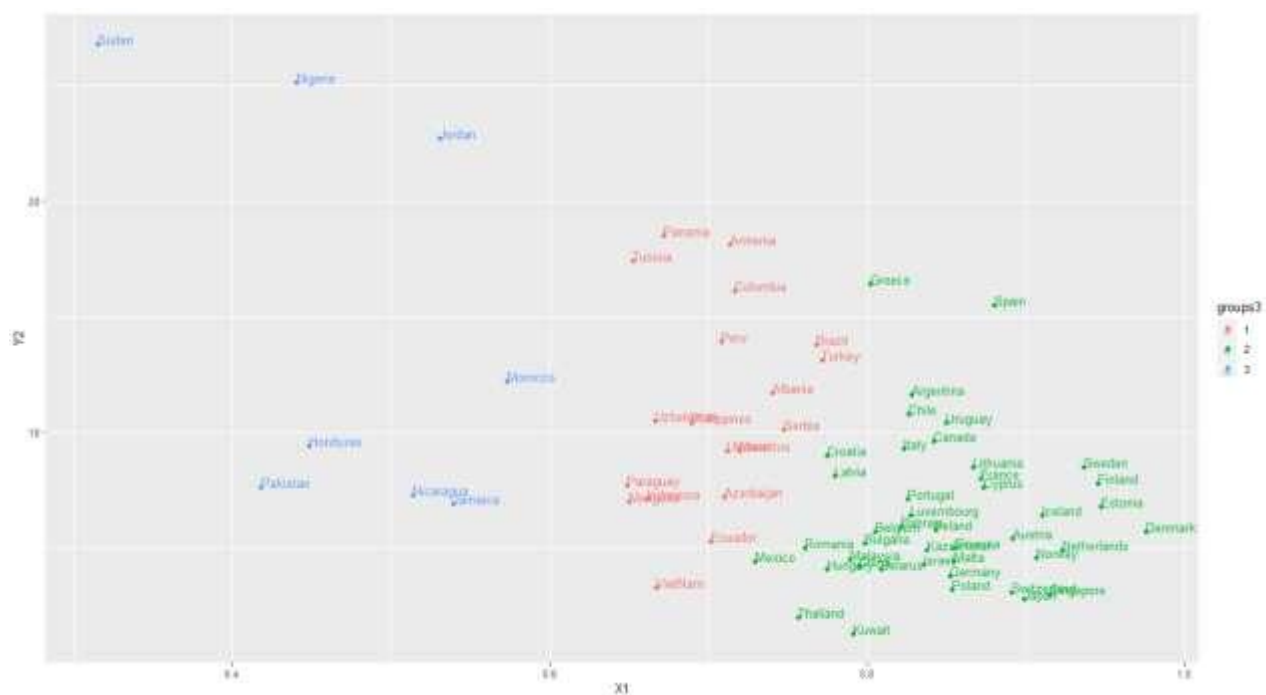


Рис. 3.39. Діаграма розсіювання для пари Y2 та X1

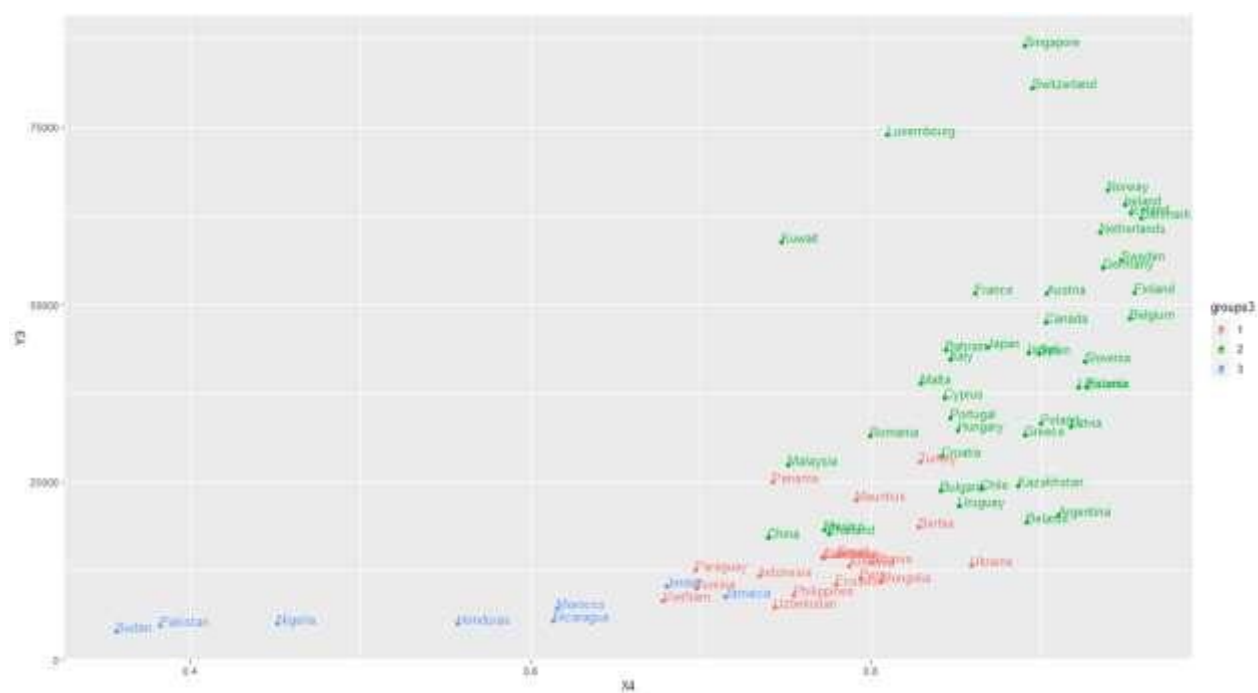


Рис. 3.40. Діаграма розсіювання для пари Y3 та X4

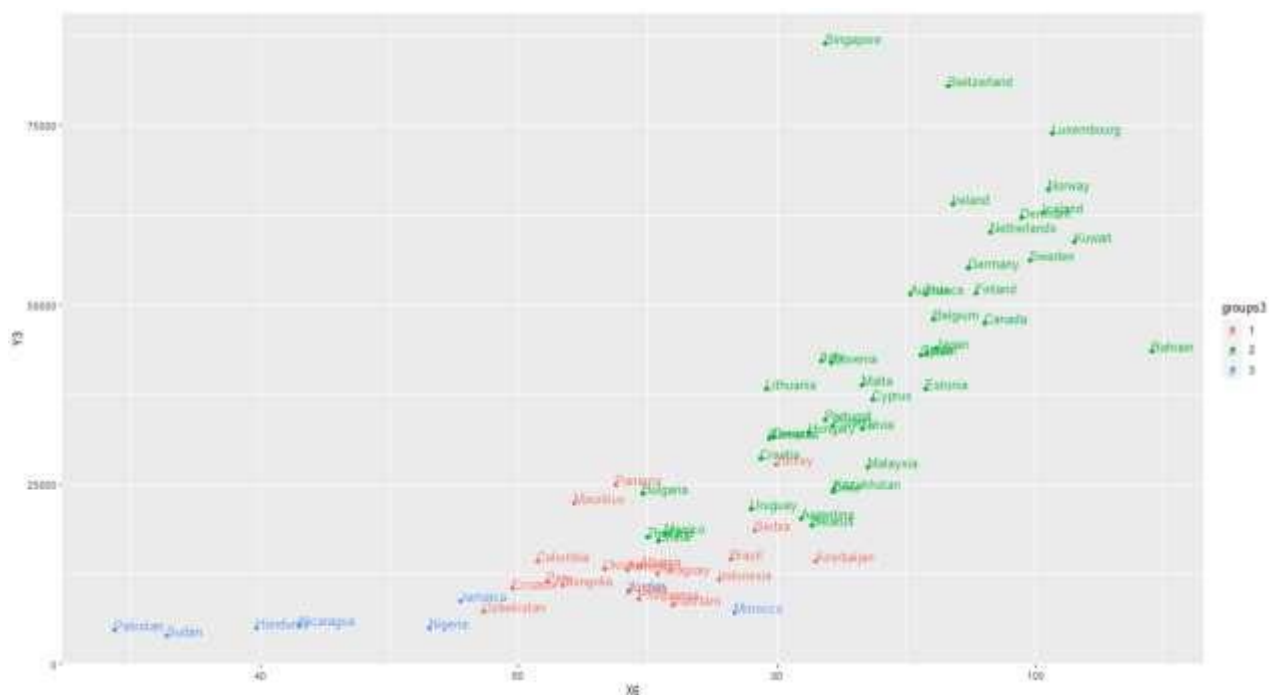


Рис. 3.41. Діаграма розсіювання для пари Y3 та X6

Дивлячись на ці діаграми розсіювання можна сказати, що розподіл на три кластери цілком вдався: точки різного кольору утворюють сформовані, досить "щільні" групи. Але є винятки, це пов'язано з тим, що країни мають досить нетипові показники, тому навряд чи вони рівно ляжуть у будь-який поділ на групи.

Проведемо аналіз по діаграмам розсіювання. На рис. 3.37 можна спостерігати, що Японія, Кувейт, Білорусь та Естонія мають низький рівень ВВП на душу населення, але високий індекс цифрового уряду. Люксембург має високий рівень ВВП на душу населення і високий індекс цифрового уряду. Також можна помітити, що Мексика, Тайланд та Румунія мають зміщення до кластера країн з середнім рівнем розвитку цифрових технологій та економіки, тому за цією парою показників ці країни можна віднести до кластера № 1. Бразилія та Туреччина мають зміщення до кластера країн з високим розвитком цифрових технологій та економіки, тому за цією парою показників ці країни можна віднести до кластера № 2.

На рис. 3.38 можна спостерігати, що Японія, Білорусь, Кувейт, Естонія і Мальта мають низький рівень ВВП на душу населення, але високий індекс телекомунікаційної інфраструктури. Люксембург та Норвегія мають високий рівень

ВВП на душу населення та достатньо високий індекс телекомунікаційної інфраструктури. Такі країни як Казахстан, Аргентина та Тайланд мають низький рівень ВВП на душу населення, але вище середнього індекс телекомунікаційної інфраструктури. Йорданія, Ямайка, Марокко та Мексика мають зміщення до кластера краї з середнім розвитком цифрових технологій та економіки, тому за цією парою показників ці країни можна віднести до кластера № 1.

На рис. 3.39 можна спостерігати, що Греція та Іспанія, мають високий рівень безробіття, але високий індекс цифрового уряду. Кувейт та Тайланд мають низький рівень безробіття, але високий індекс цифрового уряду. Данія має дуже високий індекс цифрового уряду і середній рівень безробіття. Мексика мають зміщення до кластера краї з середнім розвитком цифрових технологій та економіки, тому за цією парою показників ці країни можна віднести до кластера № 1.

На рис. 3.40 можна спостерігати, що Кувейт має високий рівень ВНД на душу населення, але середній індекс людського капіталу. Швейцарія та Сінгапур мають високий рівень ВНД на душу населення і достатньо високий індекс людського капіталу. Малайзія, Китай, Тайланд, Мексика, Йорданія та Ямайка мають зміщення до кластера краї з середнім розвитком цифрових технологій та економіки, тому за цією парою показників ці країни можна віднести до кластера № 1. Туреччину має зміщення до кластера країн з високим розвитком цифрових технологій та економіки, тому за цією парою показників цю країну можна віднести до кластера № 2.

На рис. 3.41 можна спостерігати, що Сінгапур має високий рівень ВНД на душу населення та достатньо високий відсоток користувачів Інтернетом. Бахрейн має середній рівень ВНД на душу населення та дуже високий відсоток користувачів Інтернетом. Ямайка, Нігерія, Йорданія, Марокко, Болгарія, Мексика, Тайланд, Китай, Уругвай, Аргентина, Білорусь мають зміщення до кластера краї з середнім розвитком цифрових технологій та економіки, тому за цією парою показників ці країни можна віднести до кластера № 1. Туреччина має зміщення до кластера країн з високим розвитком цифрових технологій та економіки, тому за цією парою показників цю країну можна віднести до кластера № 2.

Отже, за допомогою кластерного аналізу вдалося: визначити оптимальну кількість кластерів для країн та їх показників за 2020 рік, побудувати дендограму та здійснити ієрархічний кластерний аналіз, зробити оцінку якості кластерного аналізу, розділити країни на 3 групи та побудувати діаграми розсіювання для пар показників. Кластерний аналіз показав як краще розбити країни на 3 групи за тими показниками, які відображають цифровий та економічний розвиток краї.

3.3. Розробка регресійної моделі оцінки впливу цифрових технологій на економіку України та побудова прогнозу

Побудуємо модель множинної регресії для України по досліджуваним показникам з 2012 року по 2020 рік. Дана модель дозволить зробити прогноз значень для результуючих показників. Побудуємо гістограму розподілу значень Y1 (ВВП на душу населення), Y2 (рівень безробіття) та Y3 (ВНД на душу населення). Результати на рис. 3.42 – 3.44.

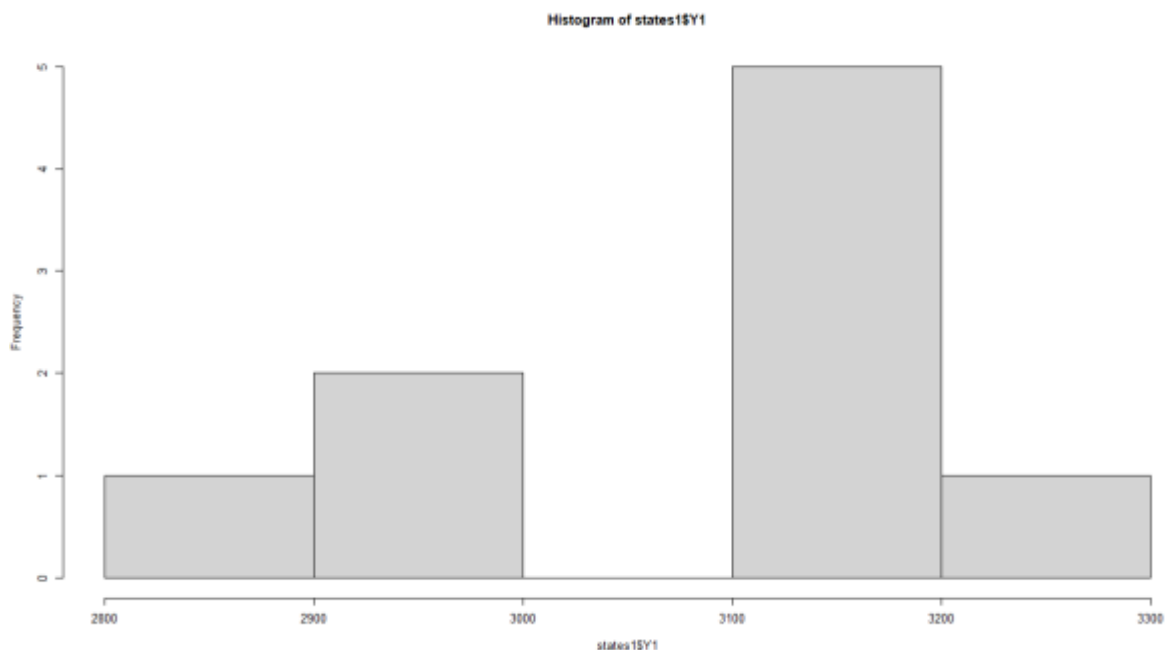


Рис. 3.42. Гістограма розподілу значень Y1

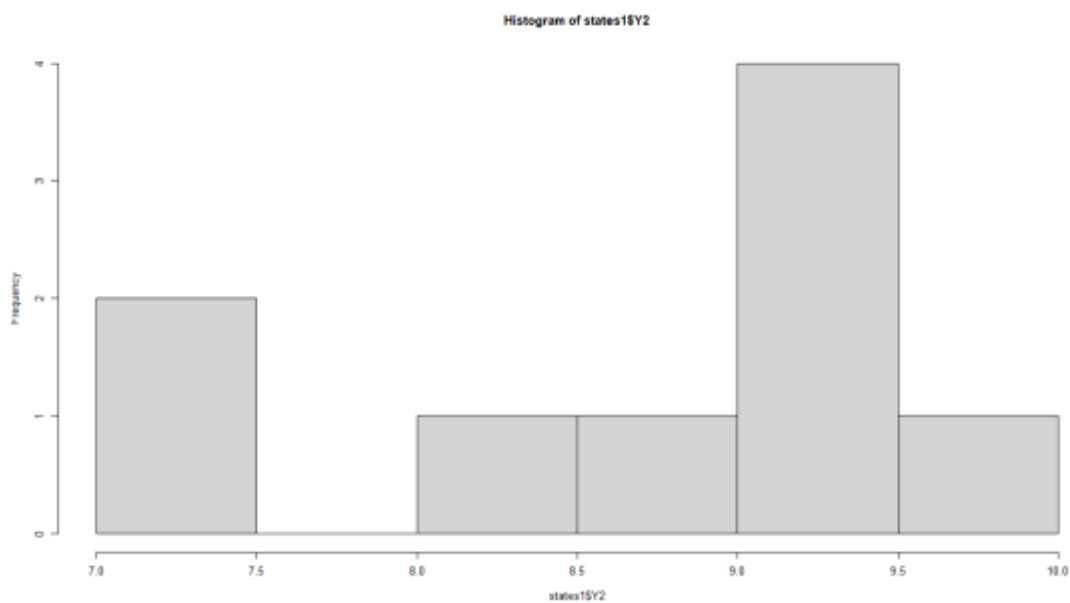


Рис. 3.43. Гістограма розподілу значень Y2

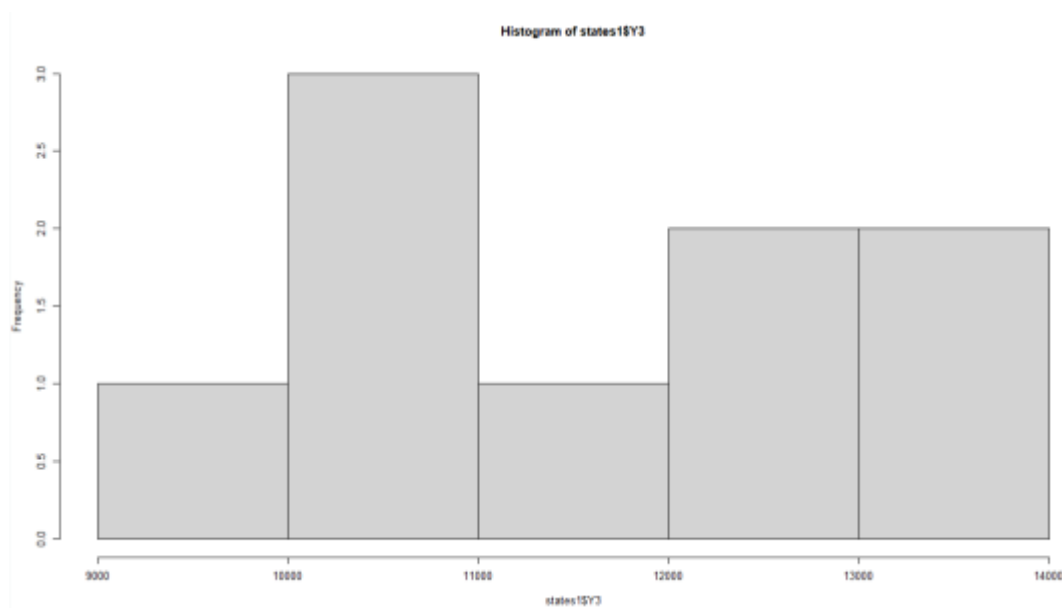


Рис. 3.44. Гістограма розподілу значень Y3

Подивимося як розподілені значення змінних між собою попарно на рис. 3.45.

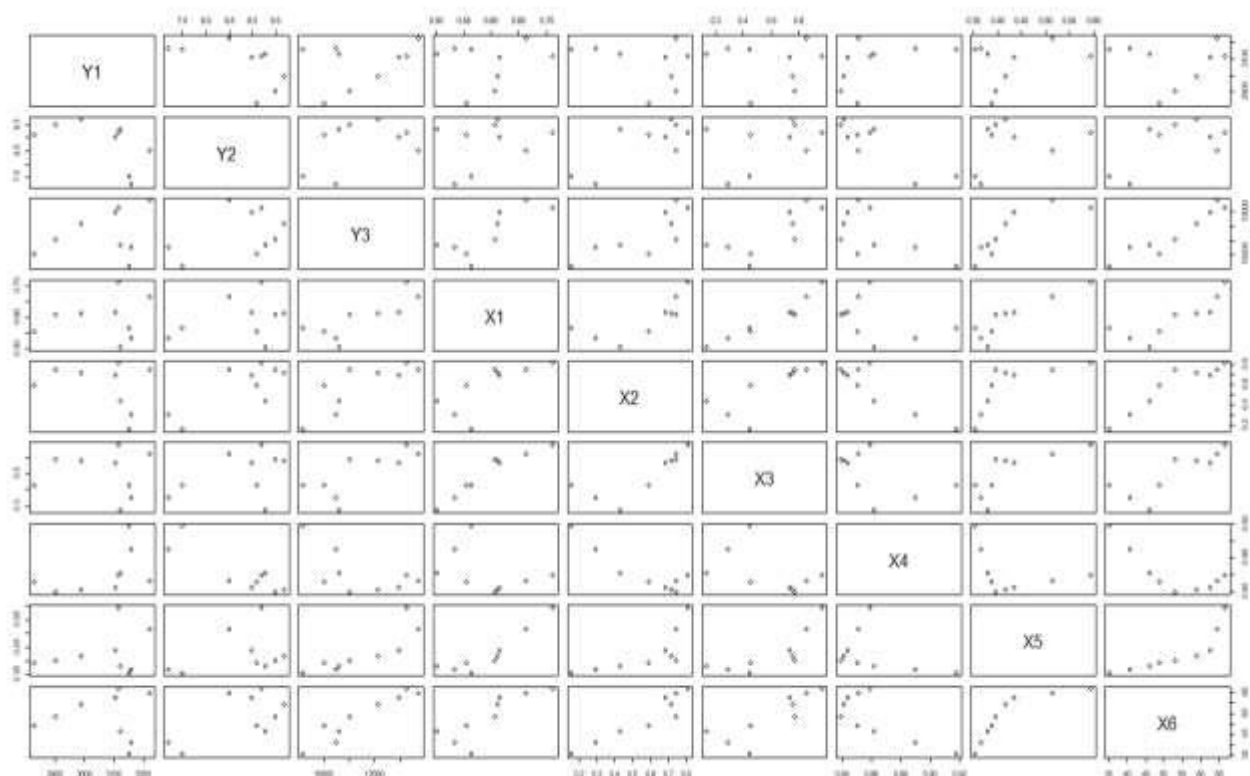


Рис. 3.45. Графік розподілення значень змінних між собою попарно

Порахуємо матрицю взаємних кореляцій Пірсона всіх змінних між собою та ранговий коефіцієнт кореляції Спірмена. Результат на рис. 3.46 – 3.48.

	Y1	Y2	Y3	X1	X2	X3	X4	X5	X6
Y1	1.00	-0.55	0.32	0.15	-0.29	-0.04	0.48	0.28	0.03
Y2	-0.55	1.00	0.38	0.29	0.77	0.39	-0.86	0.30	0.61
Y3	0.32	0.38	1.00	0.81	0.80	0.76	-0.60	0.84	0.95
X1	0.15	0.29	0.81	1.00	0.74	0.96	-0.36	0.90	0.83
X2	-0.29	0.77	0.80	0.74	1.00	0.80	-0.89	0.71	0.93
X3	-0.04	0.39	0.76	0.96	0.80	1.00	-0.49	0.76	0.82
X4	0.48	-0.86	-0.60	-0.36	-0.89	-0.49	1.00	-0.36	-0.76
X5	0.28	0.30	0.84	0.90	0.71	0.76	-0.36	1.00	0.84
X6	0.03	0.61	0.95	0.83	0.93	0.82	-0.76	0.84	1.00

Рис. 3.46. Матриця взаємних кореляцій всіх змінних між собою

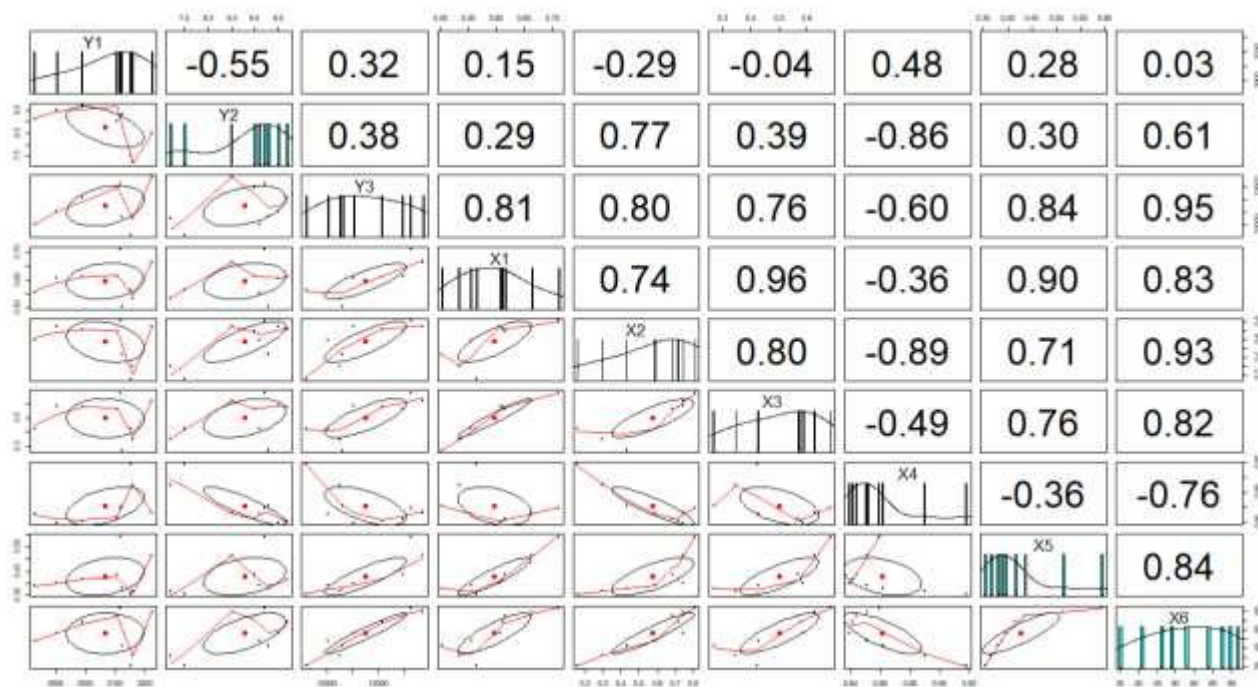


Рис. 3.47. Діаграми розподілу даних та значення коефіцієнтів кореляції Пірсона

	Y1	Y2	Y3	X1	X2	X3	X4	X5	X6
Y1	1.00	-0.67	0.15	0.02	-0.18	-0.15	0.72	-0.12	-0.12
Y2	-0.67	1.00	0.25	0.12	0.47	0.30	-0.68	0.33	0.33
Y3	0.15	0.25	1.00	0.82	0.87	0.77	-0.42	0.93	0.93
X1	0.02	0.12	0.82	1.00	0.82	0.92	-0.37	0.88	0.88
X2	-0.18	0.47	0.87	0.82	1.00	0.93	-0.58	0.93	0.93
X3	-0.15	0.30	0.77	0.92	0.93	1.00	-0.52	0.87	0.87
X4	0.72	-0.68	-0.42	-0.37	-0.58	-0.52	1.00	-0.52	-0.52
X5	-0.12	0.33	0.93	0.88	0.93	0.87	-0.52	1.00	1.00
X6	-0.12	0.33	0.93	0.88	0.93	0.87	-0.52	1.00	1.00

Рис. 3.48. Ранговий коефіцієнт кореляції Спірмена

За допомогою побудованих графіків та розрахованих значень коефіцієнтів кореляції Пірсона та рангових коефіцієнтів кореляції Спірмена можна прослідкувати тісноту зв'язків між факторними та результуючими змінними. Також можна зробити висновок про наявність мультиколінеарності між факторними змінними.

Побудуємо лінійну множинну регресію для Y1. Також перевіримо статистичну значимість параметрів моделі та її адекватність. Результати побудови моделі на рис. 3.49.


```

Call:
lm(formula = Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, data = states1[train.ids,
])

Residuals:
    1     2     3     4     5     6     7     8     9
-9.779 19.558 32.601 -84.759 72.104 -59.449 -13.100 85.648 -42.824

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -38539.56   71844.74  -0.536   0.629
X1           -22541.87   42415.14  -0.531   0.632
X2              5278.99   11494.58   0.459   0.677
X3              3820.00    7765.58   0.492   0.657
X4            54153.00   95929.14   0.565   0.612
X5                NA         NA         NA     NA
X6              65.21     76.56   0.852   0.457

Residual standard error: 94.52 on 3 degrees of freedom
(91 пропущенное наблюдение удалено)
Multiple R-squared:  0.8053,    Adjusted R-squared:  0.4808
F-statistic: 2.482 on 5 and 3 DF,  p-value: 0.2424

```

Рис. 3.49. Результати багатofакторного регресійного аналізу для Y1

Проаналізувавши результати, можна сказати, що модель не є статистично значимою та адекватною. У даній моделі велика кількість результуючих змінних на короткий період дослідження, також між факторними змінними існує тісний лінійний взаємозв'язок. Спробуємо застосувати метод покрокового включення факторних змінних. Подальші розрахунки та побудова моделі буде виконуватись у ППП STATISTICA. STATISTICA дозволяє проводити різні процедури (модулі) обробки статистичних даних (в термінології програми – аналізи): розрахунок описових статистик; аналіз динамічних рядів й прогнозування; множинна регресія; дискримінантний аналіз; аналіз відповідностей; кластерний аналіз; факторний аналіз, дисперсійний аналіз і та інші [15]. Результат зображено на рис 3.51.

Summary of Stepwise Regression; DV: Y1 (Spreadsheet1_(Recovered))							
Variable	Step +in/-out	Multiple R	Multiple R-square	R-square change	F - to entr/rem	p-value	Variables included
X4	1	0,476876	0,227411	0,227411	2,060446	0,194302	1
X6	2	0,761850	0,580415	0,353004	5,047903	0,065734	2
X1	3	0,887713	0,788034	0,207619	4,897474	0,077808	3

Рис. 3.50. Результат метода покрокового включення для Y1

Регресійна модель, після застосування метода покрокового включення, зображена на рис. 3.51.

Regression Summary for Dependent Variable: Y1 (Spreadsheet1)						
R= ,88771299 R ² = ,78803436 Adjusted R ² = ,66085497						
F(3,5)=6,1962 p<,03881 Std.Error of estimate: 76,392						
N=9	b*	Std.Err. of b*	b	Std.Err. of b	t(5)	p-value
Intercept			-5152,80	1921,508	-2,68164	0,043734
X4	1,86206	0,445919	9395,57	2250,014	4,17578	0,008690
X6	2,38884	0,739755	28,50	8,826	3,22923	0,023226
X1	-1,15174	0,520438	-2320,22	1048,438	-2,21302	0,077808

Рис. 3.51. Регресійна модель методом покрокового включення для Y1.

Застосувавши метод покрокового включення, бачимо, що факторні змінні X2, X3 та X5 не були включені в побудову моделі. Проаналізуємо отримані результати моделі. Коефіцієнт множинної кореляції рівний 88,77%, що показує величину тісноти лінійного зв'язку між Y1 та факторними змінними. Коефіцієнт детермінації рівний 78,80%, який показує, що всі вхідні елементи практично повністю пояснюють долю варіації результуючого показника. $F_R = 6.2$ – розраховане значення критерію Фішера. $F_{0,05} = 5.41$ – табличне значення критерію Фішера. Оскільки $F_R > F_{0,05}$, то дана модель є значимою та адекватною. Далі проаналізуємо критерій Стьюдента. $t_{0,05}(5) = 2.571$ – табличне значення критерія Стьюдента. Порівнявши табличне з $|t_{a_n}|$, можемо зробити висновок, що отримані значення для коефіцієнтів a_n статистично значимі для факторних змінних X4 та X6, тобто вони істотно впливають на величину Y1. Проте фактор X1 не має статистичної значущості. Отже, можна зробити висновок, що отримана модель є адекватною і може використовуватися для прогнозування. Таким чином, маємо варіант лінійної множинної економетричної моделі:

$$Y1 = -5152.80 + 9395.57 * X4 + 28.50 * X6 - 2320.22 * X1.$$

Результат розрахунку критерію Дарбіна-Уотсона представлено на рис. 3.52.

	Durbin-Watson d (Spreadsheet1) and serial correlation of residuals		
	Durbin-Watson d	Serial Corr.	
Estimate	3,124859	-0,660912	

Рис. 3.52. Розрахунок критерію Дарбіна-Уотсона для Y1

Розраховане значення критерію дорівнює $DW = 3.12$, табличні значення критерію $dl = 0.63$; $du = 1.70$. Отже, $4 - du \leq d \leq 4 - dl$, приймається гіпотеза про невизначеність автокореляції.

Наступним етапом дослідження є тестування гетероскедастичності. Проведемо аналіз моделі за допомогою тесту Глейзера за допомогою проміжних розрахунків у табличному процесорі MS Excel. Даний тест припускає, що дисперсія випадкових залишків являє собою певну функцію залежності від якогось фактору (або факторів). Знайдені залишки та проміжні розрахунки наведені рис. 3.53.

Рік	Y1	X4	X6	Залишки, e	e
2012	3153.735	0.91758	35.27	-50.65261392	50.65261
2013	3160.081	0.88959	40.95413	58.29556224	58.29556
2014	3123.9	0.8616	46.23598	129.0704414	129.0704
2015	2828.885	0.850275	48.88464	-128.2042514	128.2043
2016	2903.795	0.83895	53.00097	-31.43180927	31.43181
2017	2988.504	0.841275	58.88948	-24.0602215	24.06022
2018	3106.039	0.8436	62.55316	40.20721945	40.20722
2019	3224.56	0.85135	64.57658	91.40102655	91.40103
2020	3115.861	0.8591	66.6	-84.62535362	84.62535

Рис. 3.53. Залишки та проміжні розрахунки для Y1

Проведемо перевірку залишків на гіпотезу про наявність гетероскедастичності за тестом Глейзера. Знайдемо параметри регресійного рівняння для X4 та X6 за допомогою надбудови «Аналіз даних», інструменту «Регресія». Результати регресійного аналізу для змінних наведені на рис. 3.54 – 3.55.

Тест Глейзера X4								
Регрессионная статистика								
Множественный R	0.002278							
R-квадрат	5.19E-06							
Нормированный R-квадрат	-0.14285							
Стандартная ошибка	42.32319							
Наблюдения	9							
Дисперсионный анализ								
	df	SS	MS	F	значимость F			
Регрессия	1	0.065081	0.0650813	3.63329E-05	0.995359			
Остаток	7	12538.77	1791.25221					
Итого	8	12538.83						
Коэффициент								
	коэффициент	стандартная ошибка	t-статистика	P-Значение	нижние 95%	верхние 95%	нижние 95.0%	верхние 95.0%
Y-пересечение	67.89431	496.0557	0.136868329	0.89498826	-1105.09	1240.88	-1105.09	1240.88
Переменная X 1	3.469439	575.5851	0.006027674	0.995358826	-1357.57	1364.512	-1357.57	1364.512

Рис. 3.54. Регресійний аналіз змінної X4 по залишкам для змінної Y1

Тест Глейзера X6								
Регрессионная статистика								
Множественный R	0.075697							
R-квадрат	0.00573							
Нормированный R-квадрат	-0.13631							
Стандартная ошибка	42.20187							
Наблюдения	9							
Дисперсионный анализ								
	df	SS	MS	F	значимость F			
Регрессия	1	71.84735	71.84734898	0.040341	0.846526			
Остаток	7	12466.98	1780.9976					
Итого	8	12538.83						
Коэффициент								
	коэффициент	стандартная ошибка	t-статистика	P-Значение	нижние 95%	верхние 95%	нижние 95.0%	верхние 95.0%
Y-пересечение	85.32826	73.28236	1.164376614	0.282416	-87.957	258.6135	-87.957	258.6135
Переменная X 1	-0.27257	1.357072	-0.200850865	0.846526	-3.48153	2.936395	-3.48153	2.936395

Рис. 3.55. Регресійний аналіз змінної X6 по залишкам для змінної Y1

Задав $k=1$, одержуємо наступні рівняння регресії:

$$|e| = 67.89431 + 3.469439 * x_4 + v; t_b = 0.006;$$

$$|e| = 85.32826 - 0.27257 * x_6 + v; t_b = -0.2009.$$

Для визначення наявності гетероскедастичності знайдемо табличне значення критерію Стюдента та порівняємо його зі знайденими. Критерій Стюдента $t_{\text{табл.}}(0.05; 9) = 2.262$. У результаті порівняння табличного та розрахованих значень, можемо зробити висновок про відсутність гетероскедастичності у факторів X4 та X6, адже їх t-критерії є меншими за табличне значення. Отже, гіпотеза про наявність гетероскедастичності залишків моделі відсунена,

адже визначено, що рівняння регресії в тесті Глейзера не є значимим, тож можна затверджувати, що залишки моделі гомоскедастичні. Побудовану модель для змінної Y1 можна використовувати для прогнозування.

Побудуємо лінійну множинну регресію для Y2. Також перевіримо статистичну значимість параметрів моделі та її адекватність. Результати побудови моделі на рис. 3.56.

```
Call:
lm(formula = Y2 ~ X1 + X2 + X3 + X4 + X5 + X6, data = states1[train.ids,
])

Residuals:
    1     2     3     4     5     6     7     8     9
0.42016 -0.84033  0.40216  0.03601 -0.10585  0.17570  0.12679 -0.42928  0.21464

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -253.4009    506.9298  -0.500   0.652
X1           -162.0232    299.2772  -0.541   0.626
X2             47.7455     81.1047   0.589   0.597
X3             27.7873     54.7932   0.507   0.647
X4            352.7754    676.8670   0.521   0.638
X5              NA         NA         NA     NA
X6              0.2560     0.5402   0.474   0.668

Residual standard error: 0.6669 on 3 degrees of freedom
(91 пропущенное наблюдение удалено)
Multiple R-squared:  0.784,    Adjusted R-squared:  0.4239
F-statistic: 2.177 on 5 and 3 DF,  p-value: 0.2773
```

Рис. 3.56. Результати багатofакторного регресійного аналізу для Y2

Проаналізувавши результати, можна сказати, що модель не є статистично значимою та адекватною. У даній моделі аналогічна ситуація, що й в побудованій моделі для Y1 з кількістю результуючих змінних на короткий період дослідження та тісний зв'язок між факторними змінними. Спробуємо застосувати метод покрокового включення факторних змінних. Результат зображено на рис 3.57.

Summary of Stepwise Regression; DV: Y2 (Spreadsheet1_(Recovered))							
Variable	Step +in/-out	Multiple R	Multiple R-square	R-square change	F - to entr/rem	p-value	Variables included
X4	1	0,857293	0,734951	0,734951	19,41018	0,003135	1

Рис. 3.57. Результат метода покрокового включення для Y2

Регресійна модель, після застосування метода покрокового включення, зображена на рис. 3.58.

Regression Summary for Dependent Variable: Y2 (Spreadsheet1_(Recovered))						
R= ,85729264 R ² = ,73495067 Adjusted R ² = ,69708648						
F(1,7)=19,410 p<,00314 Std.Error of estimate: ,48356						
N=9	b*	Std.Err. of b*	b	Std.Err. of b	t(7)	p-value
Intercept			33,7377	5,667668	5,95266	0,000569
X4	-0,857293	0,194587	-28,9733	6,576329	-4,40570	0,003135

Рис. 3.58. Регресійна модель методом покрокового включення для Y2

Застосувавши метод покрокового включення, бачимо, що факторні змінні X1, X2, X3, X5 та X6 не були включені в побудову моделі. Проаналізуємо отримані результати моделі. Коефіцієнт множинної кореляції рівний 85,73%, що показує величину тісноти лінійного зв'язку між Y2 та факторними змінними. Коефіцієнт детермінації рівний 73,5%, який показує, що всі вхідні елементи практично повністю пояснюють долю варіації результуючого показника. $F_R = 19.41$ – розраховане значення критерію Фішера. $F_{0.05} = 5.59$ – табличне значення критерію Фішера. Оскільки $F_R > F_{0.05}$, то дана модель є значимою та адекватною. Далі проаналізуємо критерій Стьюдента. $t_{0.05}(7) = 2.365$ – табличне значення критерія Стьюдента. Порівнявши табличне з $|t_{a_n}|$, можемо зробити висновок, що отримані значення для коефіцієнтів a_n статистично значимі для факторних змінних X4, тобто змінна істотно впливає на величину Y2. Отже, можна зробити висновок, що отримана модель є адекватною і може використовуватися для прогнозування. Таким чином, маємо варіант лінійної множинної економетричної моделі:

$$Y2 = 33.7377 - 28.9733 * X4.$$

Результат розрахунку критерію Дарбіна-Уотсона представлено на рис. 3.59.

Durbin-Watson d (Spreadsheet1_(Recovered)) and serial correlation of residuals				
	Durbin-Watson d	Serial Corr.		
Estimate	2,799504	-0,513975		

Рис. 3.59. Розрахунок критерію Дарбіна-Уотсона для Y2

Розраховане значення критерію дорівнює $DW = 2.7995$, табличні значення критерію $dl = 0.82$; $du = 1.32$. Отже, $4 - du \leq d \leq 4 - dl$, приймається гіпотеза про невизначеність автокореляції.

Наступним етапом дослідження є тестування гетероскедастичності. Проведемо аналіз моделі за допомогою тесту Глейзера за допомогою проміжних розрахунків у табличному процесорі MS Excel. Знайдені залишки та проміжні розрахунки наведені рис. 3.60.

Рік	Y	X4	Залишки, e	e
2012	7.5	0.91758	0.347625528	0.347626
2013	7.2	0.88959	-0.763337726	0.763338
2014	9.3	0.8616	0.525699021	0.525699
2015	9.1	0.850275	-0.002423839	0.002424
2016	9.5	0.83895	0.069453301	0.069453
2017	9.7	0.841275	0.336816272	0.336816
2018	9	0.8436	-0.295820757	0.295821
2019	8.5	0.85135	-0.571277519	0.571278
2020	9.2	0.8591	0.353265718	0.353266

Рис. 3.60. Залишки та проміжні розрахунки для Y2

Проведемо перевірку залишків на гіпотезу про наявність гетероскедастичності за тестом Глейзера. Знайдемо параметри регресійного рівняння для X4 за допомогою надбудови «Аналіз даних», інструменту «Регресія». Результати регресійного аналізу для змінних наведені на рис. 3.61.

Тест Глейзера X4								
<i>Регрессионная статистика</i>								
Множественный R	0.418964							
R-квадрат	0.175531							
Нормированный R-квадрат	0.05775							
Стандартная ошибка	0.23069							
Наблюдения	9							
<i>Дисперсионный анализ</i>								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>значимость F</i>			
Регрессия	1	0.079311	0.079311156	1.490314748	0.261689			
Остаток	7	0.372524	0.053217722					
Итого	8	0.451835						
	<i>Коэффициент</i>	<i>стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>нижние 95%</i>	<i>верхние 95%</i>	<i>нижние 95.0%</i>	<i>верхние 95.0%</i>
Y-пересечение	-2.93661	2.703835	-1.086089549	0.313418868	-9.33016	3.456947	-9.33016	3.456947
Переменная X 1	3.829996	3.137324	1.220784481	0.261688522	-3.5886	11.24859	-3.5886	11.24859

Рис. 3.61. Регресійний аналіз змінної X4 по залишкам для змінної Y2

Задав $k=1$, одержуємо наступні рівняння регресії:

$$|e| = -2.9366 + 3.83 * x_4 + v; t_b = 0.2617.$$

Для визначення наявності гетероскедастичності знайдемо табличне значення критерію Стьюдента та порівняємо його зі знайденими. Критерій Стьюдента $t_{\text{табл.}}(0.05; 9) = 2.262$. У результаті порівняння табличного та розрахованих значень, можемо зробити висновок про відсутність гетероскедастичності у фактора X4, адже їх t-критерії є меншими за табличне значення. Отже, гіпотеза про наявність гетероскедастичності залишків моделі відсунена, адже визначено, що рівняння регресії в тесті Глейзера не є значимим, тож можна затверджувати, що залишки моделі гомоскедастичні. Побудовану модель для змінної Y2 можна використовувати для прогнозування.

Побудуємо лінійну множинну регресію для Y3. Також перевіримо статистичну значимість параметрів моделі та її адекватність. Результати побудови моделі на рис. 3.62.


```

Call:
lm(formula = Y3 ~ X1 + X2 + X3 + X4 + X5 + X6, data = states1[train.ids,
])

Residuals:
    1     2     3     4     5     6     7     8     9
-196.66  393.32 -38.38 -316.56  298.25 -279.94 -96.69  473.32 -236.66

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -225645.3   380520.3  -0.593   0.595
X1           -135803.5   224648.6  -0.605   0.588
X2             31799.4    60880.2   0.522   0.638
X3             25057.3    41129.8   0.609   0.585
X4            304742.4   508081.6   0.600   0.591
X5              NA         NA         NA     NA
X6              468.1     405.5     1.154   0.332

Residual standard error: 500.6 on 3 degrees of freedom
(91 пропущенное наблюдение удалено)
Multiple R-squared:  0.9627,    Adjusted R-squared:  0.9004
F-statistic: 15.47 on 5 and 3 DF,  p-value: 0.02368

```

Рис. 3.62. Результати багатofакторного регресійного аналізу для Y3

Проаналізувавши результати, можна сказати, що модель не є статистично значимою та адекватною. У даній моделі аналогічна ситуація, що й в побудованих моделях для Y1 та Y2 з кількістю результуючих змінних на короткий період дослідження та тісний зв'язок між факторними змінними. Спробуємо застосувати метод покрокового включення факторних змінних. Результат зображено на рис 3.63.

Summary of Stepwise Regression; DV: Y3 (Spreadsheet1_(Recovered))							
Variable	Step +in/-out	Multiple R	Multiple R-square	R-square change	F - to entr/rem	p-value	Variables included
X6	1	0,949836	0,902188	0,902188	64,56612	0,000089	1
X2	2	0,978792	0,958034	0,055846	7,98440	0,030129	2

Рис. 3.63. Результат метода покрокового включення для Y3

Регресійна модель, після застосування метода покрокового включення, зображена на рис. 3.64.

Regression Summary for Dependent Variable: Y3 (Spreadsheet1_(Recovered))						
R= ,97879210 R ² = ,95803397 Adjusted R ² = ,94404530						
F(2,6)=68,486 p<,00007 Std.Error of estimate: 375,28						
N=9	b*	Std.Err. of b*	b	Std.Err. of b	t(6)	p-value
Intercept			2198,83	972,263	2,26156	0,064410
X6	1,555159	0,229969	224,40	33,183	6,76247	0,000510
X2	-0,649817	0,229969	-4503,41	1593,752	-2,82567	0,030129

Рис. 3.64. Регресійна модель методом покрокового включення для Y3

Застосувавши метод покрокового включення, бачимо, що факторні змінні X1, X3, X4 та X5 не були включені в побудову моделі. Проаналізуємо отримані результати моделі. Коефіцієнт множинної кореляції рівний 97,88%, що показує величину тісноти лінійного зв'язку між Y3 та факторними змінними. Коефіцієнт детермінації рівний 95,80%, який показує, що всі вхідні елементи практично повністю пояснюють долю варіації результуючого показника. $F_R = 68,486$ – розраховане значення критерію Фішера. $F_{0,05} = 5.14$ – табличне значення критерію Фішера. Оскільки $F_R > F_{0,05}$, то дана модель є значимою та адекватною. Далі проаналізуємо критерій Стьюдента. $t_{0,05}(6) = 2.447$ – табличне значення критерія Стьюдента. Порівнявши табличне з $|t_{a_n}|$, можемо зробити висновок, що отримані значення для коефіцієнтів a_n статистично значимі для факторних змінних X2 та X6, тобто вони істотно впливають на величину Y3. Отже, можна зробити висновок, що отримана модель є адекватною і може використовуватися для прогнозування. Таким чином, маємо варіант лінійної множинної економетричної моделі:

$$Y3 = 2198.83 + 224.40 * X6 - 4503.41 * X2.$$

Результат розрахунку критерію Дарбіна-Уотсона представлено на рис. 3.65.

Durbin-Watson d (Spreadsheet1_(Recovered)) and serial correlation of residuals			
	Durbin-Watson d	Serial Corr.	
Estimate	2,895584	-0,544710	

Рис. 3.65. Розрахунок критерію Дарбіна-Уотсона для Y3

Розраховане значення критерію дорівнює $DW = 2.896$, табличні значення критерію $dl = 0.63$; $du = 1.70$. Отже, $4 - du \leq d \leq 4 - dl$, приймається гіпотеза про невизначеність автокореляції.

Наступним етапом дослідження є тестування гетероскедастичності. Проведемо аналіз моделі за допомогою тесту Глейзера за допомогою проміжних розрахунків у табличному процесорі MS Excel. Знайдені залишки та проміжні розрахунки наведені рис. 3.66.

Рік	Y3	X2	X6	Залишки, e	e
2012	9190	0.1579	35.27	-212.3288456	212.3288
2013	10510	0.294635	40.95413	447.9272119	447.9272
2014	10620	0.43137	46.23598	-11.54454899	11.54455
2015	10040	0.588565	48.88464	-477.9913957	477.9914
2016	11040	0.74576	53.00097	306.2194443	306.2194
2017	12140	0.71558	58.88948	-51.07490218	51.0749
2018	12960	0.6854	62.55316	-189.1166322	189.1166
2019	13750	0.74745	64.57658	426.2643454	426.2643
2020	13260	0.8095	66.6	-238.3546769	238.3547

Рис. 3.66. Залишки та проміжні розрахунки для Y3

Проведемо перевірку залишків на гіпотезу про наявність гетероскедастичності за тестом Глейзера. Знайдемо параметри регресійного рівняння для X2 та X6 за допомогою надбудови «Аналіз даних», інструменту «Регресія». Результати регресійного аналізу для змінних наведені на рис. 3.67 – 3.68.

Тест Глейзера X2								
<i>Регрессионная статистика</i>								
Множественный R	0.000416523							
R-квадрат	1.73491E-07							
Нормированный R-квадрат	-0.142856945							
Стандартная ошибка	179.568555							
Наблюдения	9							
Дисперсионный анализ								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
Регрессия	1	0.03915946	0.039159464	1.21444E-06	0.999151466			
Остаток	7	225714.062	32244.86594					
Итого	8	225714.101						
	<i>Коэффициенты</i>	<i>стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>нижние 95%</i>	<i>верхние 95%</i>	<i>нижние 95.0%</i>	<i>верхние 95.0%</i>
Y-пересечение	262.1377807	170.364466	1.538688127	0.167770875	-140.7101667	664.9857	-140.71	664.9857
Переменная X 1	0.305627513	277.334799	0.001102016	0.999151466	-655.4869644	656.0982	-655.487	656.0982

Рис. 3.67. Регресійний аналіз змінної X2 по залишкам для змінної Y3

Тест Глейзера X6								
<i>Регрессионная статистика</i>								
Множественный R	0.06174142							
R-квадрат	0.003812003							
Нормированный R-квадрат	-0.138500568							
Стандартная ошибка	179.2259858							
Наблюдения	9							
Дисперсионный анализ								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Значимость F</i>			
Регрессия	1	860.4228262	860.4228262	0.02678613	0.87462			
Остаток	7	224853.6779	32121.95399					
Итого	8	225714.1007						
	<i>Коэффициенты</i>	<i>стандартная ошибка</i>	<i>t-статистика</i>	<i>P-Значение</i>	<i>нижние 95%</i>	<i>верхние 95%</i>	<i>нижние 95.0%</i>	<i>верхние 95.0%</i>
Y-пересечение	312.3021563	311.2208841	1.003474292	0.349051265	-423.618	1048.223	-423.618	1048.223
Переменная X 1	-0.943250488	5.763310989	-0.163664687	0.874620283	-14.5713	12.68481	-14.5713	12.68481

Рис. 3.68. Регресійний аналіз змінної X6 по залишкам для змінної Y3

Задав $k=1$, одержуємо наступні рівняння регресії:

$$|e| = 262.1378 + 0.3056 * x_2 + v; t_b = 0.0011;$$

$$|e| = 312.3022 - 0.94325 * x_6 + v; t_b = -0.1637.$$

Для визначення наявності гетероскедастичності знайдемо табличне значення критерію Стюдента та порівняємо його зі знайденими. Критерій Стюдента $t_{\text{табл.}}(0.05; 9) = 2.262$. У результаті порівняння табличного та розрахованих значень, можемо зробити висновок про відсутність гетероскедастичності у факторів X2 та X6, адже їх t-критерії є меншими за табличне значення. Отже, гіпотеза про наявність гетероскедастичності залишків моделі відсунена,

адже визначено, що рівняння регресії в тесті Глейзера не є значимим, тож можна затверджувати, що залишки моделі гомоскедастичні. Побудовану модель для змінної Y3 можна використовувати для прогнозування.

Тепер використаємо побудовані моделі для прогнозування результуючих змінних на 2021 рік. Дослідимо вплив факторних змінних на результуючі при їх збільшенні та зменшенні.

Спочатку проведемо дослідження для Y1. Збільшимо та зменшимо X4 (індекс людського капіталу) на 5% при незмінних інших факторів, по відношенню до 2020 року. Результат на рис. 3.69 – 3.70.

Predicting Values for (Spreadshe variable: Y1			
Variable	b-Weight	Value	b-Weight * Value
X4	9395,57	0,90206	8475,32
X6	28,50	66,60000	1898,15
X1	-2320,22	0,71190	-1651,76
Intercept			-5152,80
Predicted			3568,90
-95,0%CL			3252,80
+95,0%CL			3885,01

Рис. 3.69. Збільшення X4 на 5% для результуючої змінної Y1

Predicting Values for (Spreadshe variable: Y1			
Variable	b-Weight	Value	b-Weight * Value
X4	9395,57	0,81615	7668,15
X6	28,50	66,60000	1898,15
X1	-2320,22	0,71190	-1651,76
Intercept			-5152,80
Predicted			2761,73
-95,0%CL			2504,52
+95,0%CL			3018,94

Рис. 3.70. Зменшення X4 на 5% для результуючої змінної Y1

Отже, при збільшенні X4 на 5% бачимо, що $Y1_{\text{пр}} = 3568.90$, і значення може коливатися: $3252.80 \leq \widehat{Y1} \leq 3885.01$. Прогнозне значення на 2021 рік

більше на 14.54% за значення 2020 року, при умові, якщо індекс людського капіталу збільшити на 5%. При зменшенні X4 на 5% бачимо, що $Y1_{\text{пр}} = 2504.52$, і значення може коливатися: $2761.73 \leq \widehat{Y1} \leq 3018.94$. Прогнозне значення на 2021 рік менше на 19.62% за значення 2020 року, при умові, якщо індекс людського капіталу зменшити на 5%.

Збільшимо та зменшимо X4 (індекс людського капіталу), X6 (відсоток користувачів Інтернетом) та X1 (індекс електронного уряду) на 5%, по відношенню до 2020 року. Результат на рис. 3.71 – 3.72.

Predicting Values for (Spreadshe variable: Y1			
Variable	b-Weight	Value	b-Weight * Value
X4	9395,57	0,90206	8475,32
X6	28,50	69,93000	1993,06
X1	-2320,22	0,74750	-1734,35
Intercept			-5152,80
Predicted			3581,22
-95,0%CL			3248,53
+95,0%CL			3913,92

Рис. 3.71. Збільшення X4, X6 та X1 на 5% для результуючої змінної Y1

Predicting Values for (Spreadshe variable: Y1			
Variable	b-Weight	Value	b-Weight * Value
X4	9395,57	0,81615	7668,15
X6	28,50	63,27000	1803,24
X1	-2320,22	0,67630	-1569,18
Intercept			-5152,80
Predicted			2749,41
-95,0%CL			2511,65
+95,0%CL			2987,18

Рис. 3.72. Зменшення X4, X6 та X1 на 5% для результуючої змінної Y1

Отже, при збільшенні X1, X4, X6 на 5% бачимо, що $Y1_{\text{пр}} = 3581.22$, і значення може коливатися: $3248.53 \leq \widehat{Y1} \leq 3913.92$. Прогнозне значення на

2021 рік більше на 14.94% за значення 2020 року, при умові, якщо індекс людського капіталу, індекс електронного уряду та відсоток користувачів Інтернетом збільшити на 5%. При зменшенні X_1 , X_4 , X_6 на 5% бачимо, що $Y_{1\text{пр}} = 2749.41$, і значення може коливатися: $2511.65 \leq \widehat{Y}_1 \leq 2987.18$. Прогнозне значення на 2021 рік менше на 11.76% за значення 2020 року, при умові, якщо індекс людського капіталу, індекс електронного уряду та відсоток користувачів Інтернетом зменшити на 5%.

Проведемо дослідження для Y_2 . Збільшимо та зменшимо X_4 (індекс людського капіталу) на 5%, по відношенню до 2020 року. Результат на рис. 3.73 – 3.74.

Predicting Values for (Spreadshe variable: Y2)			
Variable	b-Weight	Value	b-Weight * Value
X4	-28,9733	0,902055	-26,1355
Intercept			33,7377
Predicted			7,6022
-95,0%CL			6,8650
+95,0%CL			8,3393

Рис. 3.73. Збільшення X_4 на 5% для результуючої змінної Y_2

Predicting Values for (Spreadshe variable: Y2)			
Variable	b-Weight	Value	b-Weight * Value
X4	-28,9733	0,816145	-23,6464
Intercept			33,7377
Predicted			10,0913
-95,0%CL			9,2899
+95,0%CL			10,8927

Рис. 3.74. Зменшення X_4 на 5% для результуючої змінної Y_2

Отже, при збільшенні X_4 на 5% бачимо, що $Y_{2\text{пр}} = 7.6022$, і значення може коливатися: $6.8650 \leq \widehat{Y}_2 \leq 8.3393$. Прогнозне значення на 2021 рік

менше на 17,37% за значення 2020 року, при умові, якщо індекс людського капіталу збільшити на 5%. При зменшенні X_4 на 5% бачимо, що $Y_{2PP} = 10.89$, і значення може коливатися: $9.2899 \leq \widehat{Y}_2 \leq 10.8927$. Прогнозне значення на 2021 рік більше на 18.37% за значення 2020 року, при умові, якщо індекс людського капіталу зменшити на 5%.

Проведемо дослідження для Y_3 . Збільшимо та зменшимо X_2 (індекс електронної участі) та X_6 (відсоток користувачів Інтернетом) на 5%, по відношенню до 2020 року. Результат на рис. 3.75 – 3.76.

Predicting Values for (Spreadsheet variable: Y3)			
Variable	b-Weight	Value	b-Weight * Value
X6	224,40	69,93000	15692,29
X2	-4503,41	0,84998	-3827,79
Intercept			2198,83
Predicted			14063,33
-95,0%CL			13440,92
+95,0%CL			14685,74

Рис. 3.75. Збільшення X_2 та X_6 на 5% для результуючої змінної Y_3

Predicting Values for (Spreadsheet variable: Y3)			
Variable	b-Weight	Value	b-Weight * Value
X6	224,40	63,27000	14197,78
X2	-4503,41	0,76903	-3463,24
Intercept			2198,83
Predicted			12933,38
-95,0%CL			12501,91
+95,0%CL			13364,84

Рис. 3.76. Зменшення X_2 та X_6 на 5% для результуючої змінної Y_3

Отже, при збільшенні X_2 та X_6 на 5% бачимо, що $Y_{3PP} = 14063.33$, і значення може коливатися: $13440.92 \leq \widehat{Y}_3 \leq 14685.74$. Прогнозне значення на 2021 рік більше на 6.06% за значення 2020 року, при умові, якщо відсоток користувачів Інтернетом та індекс електронної участі збільшити на 5%. При

зменшенні X_2 та X_6 на 5% бачимо, що $Y_{3\text{пр}} = 12933.38$, і значення може коливатися: $12501.91 \leq \widehat{Y}_3 \leq 13364.84$. Прогнозне значення на 2021 рік менше на 2.46% за значення 2020 року, при умові, якщо відсоток користувачів Інтернетом та індекс електронної участі зменшити на 5%.

Отже, побудовано рівняння багатофакторної регресії для прогнозування результуючих змінних, оцінено вплив факторних змінних на модель та побудовано прогноз результуючих змінних при зростанні та зменшенні факторних змінних. Побудовані моделі показали, що на ВВП на душу населення істотно впливають індекс електронного уряду, індекс людського капіталу та відсоток користувачів Інтернетом, на рівень безробіття істотно впливає індекс людського капіталу, на ВНД на душу населення істотно впливає індекс електронної участі та відсоток користувачів Інтернетом. Побудовані прогнози показали, як вплине зростання та зменшення показників розвитку цифрових технологій на економічну діяльність України.

Отже, в даному розділі за допомогою моделей панельних даних та проведених тестів оцінено вплив цифрових технологій на економічну діяльність країн, аналіз моделей та побудовані тести показали які краще моделі використовувати для діагностики результуючих показників. Також в даному розділі побудовано кластерні аналізи та поділено досліджуванні країни на групи по обраних показниках, побудовано регресійні моделі для України по показникам Y_1 (ВВП на душу населення), Y_2 (рівень безробіття) та Y_3 (ВНД на душу населення), оцінено вплив факторних змінних на модель, побудовано прогноз для результуючих показників при збільшенні та зменшенні змінних, які істотно впливають на модель та були включенні в побудову моделі.

ВИСНОВОК

В даній роботі було розглянуто моделі оцінки впливу цифрових технологій на економіку країн. Поняття «цифрові технології», «блокчейн» та «цифрові активи» дали зрозуміти що собою представляють цифрові технології та як вони можуть бути пов'язані з економікою країн та світом.

В роботі для проведення дослідження взято показники, які характеризують економічну діяльність країн та розвиток цифрових технологій. Показники які характеризують економічну діяльність країн: ВВП на душу населення, ВНД на душу населення та рівень безробіття. Ці показники результуючі, так як описують рівень економічного розвитку країн. Показники які характеризують розвиток цифрових технологій країн: індекс розвитку електронного уряду, індекс електронної участі, індекс розвитку онлайн-сервісів, індекс людського капіталу, індекс телекомунікаційної інфраструктури, відсоток користувачів Інтернетом. Ці показники факторні.

За допомогою побудованих моделей панельних даних та F-теста, теста Бройша-Пагана та теста Хаусмана можна досліджено, що вплив цифрових технологій на економіку країн наявний, це помітно на таких показниках як ВВП на душу населення й ВНД на душу населення. Також на результуючі показники, які відображають економічну діяльність країн, істотний вплив мають індекс електронної участі та відсоток користувачів Інтернетом. Аналіз моделей та тестів показали які краще моделі використовувати для подальшої діагностики результуючих показників.

Завдяки кластерному аналізу вдалося: визначити оптимальну кількість кластерів для країн та їх показників за 2020 рік, побудувати дендограму та здійснити ієрархічний кластерний аналіз, зробити оцінку якості кластерного аналізу, розділити країни на 3 групи та побудувати діаграми розсіювання для пар показників. Кластерний аналіз показав як краще розбити країни на 3 групи за тими показниками, які відображають цифровий та економічний розвиток краї.

Побудовані рівняння багатофакторної регресії дали змогу зробити прогноз результуючих змінних. Завдяки моделям багатофакторної регресії оцінено вплив показників, що відображають розвиток цифрових технологій, на показники, що відображають економічну діяльність краї. Побудовано прогноз при зростанні та зменшенні факторних змінних. Побудовані моделі показали, що на ВВП на душу населення істотно впливають індекс електронного уряду, індекс людського капіталу та відсоток користувачів Інтернетом, на рівень безробіття істотно впливає індекс людського капіталу, на ВНД на душу населення істотно впливає індекс електронної участі та відсоток користувачів Інтернетом. Побудовані прогнози показали, як вплине зростання та зменшення показників розвитку цифрових технологій на економічну діяльність України.

СПИСОК ВИКОРСТАНИХ ДЖЕРЕЛ

1. Прикладна економетрика : навч. посіб. : у двох частинах. Частина 1 : [Електронне видання] / Л. С. Гур'янова, Т. С. Клебанова, С. В. Прокопович та ін. – Харків : ХНЕУ ім. С. Кузнеця, 2016. – 235 с.
2. Kud A. Substantiation of the term «digital asset»: economic and legal aspects. *International Journal of Education and Science*. 2019. Vol. 2. No.1. P. 41–52. doi: 10.26697/ijes.2019.1.06. URL: [http://culturehealth.org/ijes_archive/IJES,Vol.2,No1,2019_\(6\).pdf](http://culturehealth.org/ijes_archive/IJES,Vol.2,No1,2019_(6).pdf).
3. Лаптев В.А. Цифрові активи як громадянські права. Юридична наука та практика: Вісник Нижегородської академії МВС Росії. URL: <https://cyberleninka.ru/article/v/tsifrovye-aktivy-kak-obekty-grazhdanskih-prav>
4. Правове життя сучасної України : у 2 т. : матеріали Міжнар. наук.-практ. конф. (м. Одеса, 17 трав. 2019 р.) / відп. ред. Г. О. Ульянова. – Одеса : Видавничий дім «Гельветика», 2019. – Т. 2. – 788 с.
5. Про затвердження Порядку подання та обігу електронних документів державному реєстратору: Наказ Міністерства Юстиції України від 23 жовтня 2011 року №997/19735. URL: <https://zakon.rada.gov.ua/laws/show/z0200-16>.
6. Поляков В.П., Косарев В.П. Информатика економістів: підручник. М: Вид. "Юрайт", 2014. 524 с.
7. Бобньова М.І. Соціальні норми та регуляція поведінки. М: Наука, 1978. 312 с.
8. Абішева О.К. Про поняття «цінність». Питання філософії. Науково-теоретичний журнал. 2002. №3. С. 139-146.
9. Міжнародний стандарт фінансової звітності (IAS) 38 "Нематеріальні активи". – [Електронний ресурс]. – Режим доступу: https://www.minfin.ru/common/img/uploaded/library/no_date/2012/IAS_38.pdf.
10. Цифрові активи та їх економіко-правове регулювання у світлі розвитку технології блокчейн: монографія / Олександр Кудь, Микола Кучерявенко, Євген Смичок. – Харків: Право, 2019. – 384 с.

11. Дані по країнам – [Електронний ресурс]. – Режим доступу: <https://minfin.com.ua/>
12. Прикладна економетрика : навч. посіб. : у двох частинах. Частина 2 : [Електронне видання] / Л. С. Гур'янова, Т. С. Клебанова, С. В. Прокопович та ін. – Харків : ХНЕУ ім. С. Кузнеця, 2016. – 252 с.
13. K-means Cluster Analysis – [Електронний ресурс]. – Режим доступу: https://uc-r.github.io/kmeans_clustering
14. Шітіков В. К., Мастицький С. Е. (2017) Класифікація, регресія, алгоритми Data Mining з використанням R. – [Електронний ресурс]. – Режим доступу: <https://ranalytics.github.io/data-mining/index.html>
15. Фетісов В. С. Пакет статистичного аналізу даних STATISTICA : навч. посіб. / В. С. Фетісов. – Ніжин : НДУ ім. М. Гоголя, 2018. – 114 с.
16. Кудь О.О., Кучерявенко М.П., Смичок Є.М. Цифрові активи та їх правове регулювання у світі розвитку технології блокчейн : монографія. Харків : Право, 2019. 216 с.
17. Професійний розвиток фахівців в умовах цифровізації суспільства: сучасні тренди : наук.-практ. конф. з міжнар. участю : Матеріали конф. / наук. ред. О. М. Спірін, О. А. Острианська. Житомир: ЖДУ ім. І. Франка, 2021. 279 с.
18. The 9th International scientific and practical conference “World science: problems, prospects and innovations” (May 19-21, 2021) Perfect Publishing, Toronto, Canada. 2021. 678 p.
19. Cluster Policy of Innovative Development of the National Economy: Integration and Infrastructure Aspects : monograph / under the editorship of professor Svitlana Smerichevska. Poznań: Wydawnictwo naukowe WSPIA, 2020. 382 pages.
20. Цифровізація науки та сучасні тренди її розвитку: матеріали I міжнародної студентської наукової конференції (Т. 2), 26 березня, 2021 рік. Дніпро, Україна: Молодіжна наукова ліга.
21. Огляд методів статистичного аналізу даних – [Електронний ресурс]. – Режим доступу: chrome-extension://efaidnbmninnibpcjpcglclefindmkaj/http://statlab.kubsu.ru/sites/project_bank/claster.pdf

22. Рейтинг країн світу з індексу людського капіталу – [Електронний ресурс]. – Режим доступу: <https://gtmarket.ru/ratings/human-capital-index>

23. Рейтинг країн світу по індексу розвитку електронного уряду – [Електронний ресурс]. – Режим доступу: <https://gtmarket.ru/ratings/e-government-development-index>

24. Дані по країнам по показникам – [Електронний ресурс]. – Режим доступу: <https://gtmarket.ru/>

25. Дані по країнам по показникам – [Електронний ресурс]. – Режим доступу: <https://gtmarket.ru/>

26. Лекції – Економетрика – [Електронний ресурс]. – Режим доступу: <https://uadoc.zavantag.com/text/21113/index-1.html>

27. Ачкасов А. Є. Конспект лекцій з курсу «Економіко-математичне моделювання» (для студентів 3 курсу заочної форми навчання бакалаврів за галуззю знань 0305 «Економіка і підприємництво», напрями підготовки 6.030504 «Економіка підприємства», 6.030509 «Облік і аудит») / А. Є. Ачкасов, О. О. Воронков; Харк. нац. акад. міськ. госп-ва. – Х.: ХНАМГ, 2011.– 204 с.

28. Ponomarenko V., Rayevnyeva O., Yermachenko V., Aksonova I. and Brovko O. (2021). Digitalization as a development factor of innovative-active university // *Problems and Perspectives in Management*, 19(4), pp. 213-231.

29. Manucharyan M.G. The role of digitalization in the development of the agro-industrial complex of Armenia // *Большая Евразия: развитие, безопасность, сотрудничество*. 2022. №5-1. – pp. 885-887.

30. Aptikasheva P.R., Mishura L.G. Impact of digitalization on business environment // *Экономика и экологический менеджмент*. 2021. №4. – pp. 66-73.

31. Akberdina V.V. Digitalization of industrial markets: regional characteristics // *Управленец*. 2018. №6. - pp. 78-87.

32. Olkhovik V. V., Lyutova O. I., Juchnevicius E. ECONOMIC GROWTH MODELS AND FDI IN THE CIS COUNTRIES DURING THE PERIOD OF DIGITALIZATION // *Финансовый журнал*. 2022. №2. – pp. 73-90.

ДОДАТОК А

Дані по показника країн з 2012 року по 2020 рік.

Survey Year	Country Name	E-Government Index	E-Participation Index	Online Service Index	Human Capital Index	Telecommunication Infrastructure Index	ВВП на душу населения (\$)	Internet users (%)	Уровень безработицы (%)	ВНД на душу населения (\$)
2012	Albania	0.52	0.11	0.42	0.79	0.34	4 276.61	54.66	13.40	10450
2012	Argentina	0.62	0.29	0.53	0.90	0.44	10 649.84	55.80	7.20	19150
2012	Armenia	0.50	0.00	0.33	0.85	0.32	3 603.58	37.50	17.30	9740
2012	Austria	0.78	0.37	0.75	0.91	0.70	48 172.24	80.03	5.20	46500
2012	Azerbaijan	0.50	0.13	0.37	0.83	0.30	5 725.63	54.20	5.20	14980
2012	Bahrain	0.69	0.66	0.86	0.80	0.42	20 924.73	88.00	3.70	49020
2012	Belarus	0.61	0.08	0.41	0.91	0.50	6 478.70	46.91	5.50	17660
2012	Belgium	0.77	0.13	0.65	0.93	0.74	44 361.21	80.72	7.60	43270
2012	Brazil	0.62	0.50	0.67	0.82	0.36	11 745.78	48.56	7.40	14670
2012	Bulgaria	0.61	0.03	0.49	0.85	0.50	7 083.60	51.90	12.40	16030
2012	Canada	0.84	0.68	0.89	0.92	0.72	48 903.22	83.00	7.40	41540
2012	Chile	0.68	0.66	0.75	0.88	0.40	14 035.67	55.05	6.60	20490
2012	China	0.54	0.21	0.53	0.77	0.30	5 325.36	42.30	4.10	11170
2012	Colombia	0.66	0.74	0.84	0.84	0.29	6 911.76	48.98	10.40	11560
2012	Croatia	0.73	0.29	0.64	0.86	0.70	13 677.42	61.94	18.60	20450
2012	Cyprus	0.65	0.08	0.56	0.88	0.52	28 873.62	60.69	11.80	31610
2012	Denmark	0.89	0.55	0.86	0.95	0.86	58 487.79	92.26	7.80	45820
2012	Ecuador	0.49	0.24	0.46	0.75	0.25	5 122.18	35.14	4.10	10160
2012	Estonia	0.80	0.76	0.82	0.91	0.66	16 485.05	78.39	10.00	25110
2012	Finland	0.85	0.74	0.88	0.95	0.72	46 538.55	89.88	7.80	41060
2012	France	0.86	0.58	0.88	0.92	0.79	41 258.27	81.44	9.80	38500
2012	Germany	0.81	0.76	0.75	0.90	0.77	44 070.91	82.35	5.40	44400
2012	Greece	0.69	0.34	0.58	0.93	0.55	22 437.22	55.07	24.40	25120
2012	Honduras	0.43	0.13	0.38	0.71	0.22	1 982.01	18.12	3.80	3750
2012	Hungary	0.72	0.45	0.69	0.91	0.57	13 367.99	70.58	10.70	22230
2012	Iceland	0.78	0.16	0.54	0.93	0.88	44 132.32	96.21	6.60	36690
2012	Indonesia	0.49	0.21	0.50	0.80	0.19	3 421.28	14.52	6.10	9440
2012	Ireland	0.71	0.13	0.54	0.95	0.66	48 633.54	76.92	15.50	37280
2012	Israel	0.81	0.89	0.85	0.89	0.69	31 701.84	70.80	6.90	30880
2012	Italy	0.72	0.26	0.58	0.91	0.67	35 019.47	55.83	10.90	36440
2012	Jamaica	0.46	0.00	0.31	0.79	0.27	4 703.09	33.79	13.90	8080
2012	Japan	0.80	0.74	0.86	0.90	0.65	45 276.87	79.50	4.30	38190
2012	Jordan	0.49	0.11	0.39	0.80	0.27	3 529.52	37.00	12.20	9080
2012	Kazakhstan	0.68	0.95	0.78	0.91	0.36	9 923.42	61.91	5.30	19050
2012	Kuwait	0.60	0.18	0.58	0.79	0.42	40 287.32	70.45	2.00	87030
2012	Latvia	0.66	0.21	0.59	0.89	0.51	13 023.71	73.12	15.00	21260
2012	Lithuania	0.73	0.53	0.70	0.92	0.58	13 683.86	67.23	13.40	23930
2012	Luxembourg	0.80	0.39	0.70	0.84	0.86	102 404.61	91.95	6.10	63710
2012	Malaysia	0.67	0.50	0.79	0.77	0.45	9 743.09	65.80	3.00	22140

2012	Malta	0.71	0.26	0.61	0.81	0.72	22 503.13	68.20	6.20	28820
2012	Mauritius	0.51	0.08	0.43	0.76	0.33	8 580.09	35.42	8.10	17570
2012	Mexico	0.62	0.58	0.73	0.83	0.31	9 690.87	39.75	4.90	16830
2012	Mongolia	0.54	0.61	0.59	0.87	0.18	3 353.09	16.40	8.20	9440
2012	Morocco	0.42	0.39	0.54	0.44	0.28	2 995.45	55.42	8.70	6570
2012	Netherlands	0.91	1.00	0.96	0.94	0.83	50 780.70	92.86	6.80	48080
2012	Nicaragua	0.36	0.13	0.31	0.65	0.12	1 657.65	13.50	5.90	4250
2012	Nigeria	0.27	0.18	0.22	0.45	0.13	2 372.46	16.10	10.60	4740
2012	Norway	0.86	0.68	0.86	0.93	0.79	88 604.57	94.65	3.30	65910
2012	Pakistan	0.28	0.13	0.37	0.36	0.12	1 006.08	9.96	6.00	3960
2012	Panama	0.57	0.32	0.46	0.82	0.44	9 541.00	40.30	4.10	17410
2012	Paraguay	0.48	0.16	0.46	0.79	0.20	4 392.14	29.34	4.60	9580
2012	Peru	0.52	0.39	0.52	0.79	0.26	5 642.58	38.20	6.80	10090
2012	Philippines	0.51	0.21	0.50	0.83	0.21	2 379.67	36.24	7.00	6990
2012	Poland	0.64	0.18	0.54	0.90	0.49	13 380.99	62.31	10.40	22860
2012	Portugal	0.72	0.37	0.65	0.89	0.60	21 337.29	60.34	16.50	25810
2012	Romania	0.61	0.08	0.52	0.88	0.42	8 621.87	45.88	6.80	18540
2012	Serbia	0.63	0.24	0.58	0.85	0.47	5 886.89	48.10	24.60	13480
2012	Singapore	0.85	0.95	1.00	0.85	0.69	50 143.78	72.00	2.00	78420
2012	Slovenia	0.75	0.21	0.67	0.93	0.65	22 989.93	68.35	8.90	28590
2012	Spain	0.78	0.50	0.76	0.94	0.63	29 235.83	69.81	24.80	31470
2012	Sudan	0.26	0.08	0.25	0.46	0.07	1 825.58	21.00	14.80	3590
2012	Sweden	0.86	0.68	0.84	0.91	0.82	53 432.64	93.18	8.10	46470
2012	Switzerland	0.81	0.34	0.67	0.89	0.88	77 844.19	85.20	2.90	61140
2012	Thailand	0.51	0.32	0.51	0.78	0.24	5 437.88	26.46	0.70	14190
2012	Tunisia	0.48	0.37	0.48	0.68	0.29	4 146.10	41.44	16.70	9790
2012	Turkey	0.53	0.05	0.46	0.77	0.35	12 128.30	45.13	8.40	20620
2012	Ukraine	0.57	0.16	0.42	0.92	0.35	3 153.74	35.27	7.50	9190
2012	Uruguay	0.63	0.18	0.55	0.90	0.44	12 981.14	54.45	6.30	16830
2012	Uzbekistan	0.51	0.24	0.50	0.83	0.21	1 814.47	23.60	4.90	5780
2012	Vietnam	0.52	0.11	0.42	0.74	0.40	1 443.50	36.80	2.00	4840
2013	Albania	0.51	0.32	0.44	0.75	0.35	4 327.38	57.20	15.90	10750
2013	Argentina	0.63	0.42	0.54	0.88	0.46	10 784.63	59.90	7.10	19650
2013	Armenia	0.54	0.26	0.47	0.81	0.36	3 705.35	41.90	16.20	10440
2013	Austria	0.79	0.50	0.75	0.89	0.73	47 901.37	80.62	5.40	48050
2013	Azerbaijan	0.52	0.28	0.40	0.79	0.38	5 982.33	73.00	5.00	16230
2013	Bahrain	0.75	0.74	0.90	0.79	0.56	21 805.12	90.00	4.40	49600
2013	Belarus	0.61	0.22	0.37	0.90	0.56	6 542.68	54.17	5.40	18260
2013	Belgium	0.76	0.38	0.66	0.91	0.72	44 355.38	82.17	8.60	44750
2013	Brazil	0.61	0.60	0.64	0.78	0.41	11 993.48	51.04	7.20	15370
2013	Bulgaria	0.58	0.14	0.36	0.82	0.55	7 146.56	53.06	13.00	16220
2013	Canada	0.84	0.75	0.90	0.91	0.72	49 516.28	85.80	7.10	43620
2013	Chile	0.69	0.80	0.79	0.85	0.45	14 461.17	58.00	6.10	21300
2013	China	0.54	0.43	0.57	0.72	0.33	5 710.67	45.80	4.10	11830
2013	Colombia	0.64	0.81	0.82	0.79	0.31	7 201.01	51.70	9.70	12280
2013	Croatia	0.68	0.31	0.55	0.83	0.66	13 654.25	66.75	19.80	21380
2013	Cyprus	0.62	0.20	0.52	0.83	0.52	27 044.30	65.45	15.90	30020
2013	Denmark	0.85	0.55	0.76	0.93	0.87	58 788.08	94.63	7.40	48210
2013	Ecuador	0.50	0.36	0.47	0.73	0.29	5 295.68	40.28	4.20	10990
2013	Estonia	0.81	0.76	0.80	0.90	0.73	16 766.53	80.00	8.60	26980

2013	Finland	0.85	0.72	0.83	0.93	0.79	45 906.92	91.51	8.30	41660
2013	France	0.88	0.77	0.94	0.90	0.80	41 282.99	81.92	10.30	40410
2013	Germany	0.80	0.73	0.71	0.89	0.79	44 143.13	84.17	5.20	46040
2013	Greece	0.70	0.57	0.59	0.90	0.60	21 980.85	59.87	27.50	25940
2013	Honduras	0.42	0.23	0.39	0.67	0.21	2 000.79	17.80	4.10	3850
2013	Hungary	0.69	0.45	0.62	0.89	0.57	13 654.29	72.64	9.80	23810
2013	Iceland	0.79	0.32	0.58	0.92	0.87	45 707.04	96.55	5.80	42100
2013	Indonesia	0.47	0.25	0.43	0.74	0.25	3 563.32	14.94	6.30	9770
2013	Ireland	0.75	0.39	0.61	0.96	0.68	48 971.66	78.25	13.80	40400
2013	Israel	0.81	0.88	0.86	0.87	0.70	32 422.55	70.25	6.20	33460
2013	Italy	0.74	0.52	0.66	0.88	0.67	33 978.55	58.46	12.40	36250
2013	Jamaica	0.45	0.10	0.31	0.76	0.27	4 700.02	37.10	15.30	8240
2013	Japan	0.84	0.85	0.90	0.88	0.75	46 249.21	88.22	4.00	40280
2013	Jordan	0.50	0.29	0.46	0.76	0.29	3 439.25	41.40	12.60	9230
2013	Kazakhstan	0.71	0.86	0.77	0.89	0.47	10 368.50	63.30	5.20	21900
2013	Kuwait	0.61	0.31	0.58	0.75	0.50	38 698.83	75.46	1.90	84000
2013	Latvia	0.69	0.46	0.64	0.86	0.56	13 468.06	75.23	11.90	22760
2013	Lithuania	0.73	0.59	0.73	0.89	0.57	14 313.77	68.45	11.80	26130
2013	Luxembourg	0.78	0.47	0.66	0.81	0.87	103 721.75	93.78	6.80	63840
2013	Malaysia	0.64	0.51	0.73	0.74	0.45	10 061.70	57.06	3.10	22690
2013	Malta	0.68	0.37	0.51	0.80	0.74	23 403.89	68.91	6.10	30690
2013	Mauritius	0.52	0.30	0.45	0.72	0.39	8 848.89	40.12	8.00	19240
2013	Mexico	0.60	0.59	0.70	0.79	0.31	9 693.72	43.46	4.90	16870
2013	Mongolia	0.55	0.65	0.60	0.83	0.22	3 669.54	17.70	7.90	9950
2013	Morocco	0.46	0.60	0.62	0.47	0.31	3 087.12	56.00	9.50	6820
2013	Netherlands	0.90	1.00	0.94	0.93	0.83	50 565.30	93.96	8.20	49910
2013	Nicaragua	0.32	0.11	0.20	0.61	0.14	1 716.39	15.50	5.70	4390
2013	Nigeria	0.28	0.26	0.26	0.42	0.16	2 463.89	19.10	10.00	4970
2013	Norway	0.85	0.69	0.81	0.94	0.80	88 444.89	95.05	3.80	67670
2013	Pakistan	0.27	0.23	0.34	0.35	0.12	1 028.45	10.90	6.00	4040
2013	Panama	0.55	0.40	0.42	0.78	0.45	10 027.35	44.03	4.10	19480
2013	Paraguay	0.43	0.21	0.34	0.73	0.21	4 691.53	36.90	5.00	10620
2013	Peru	0.53	0.55	0.57	0.76	0.27	5 919.21	39.20	5.90	10710
2013	Philippines	0.49	0.39	0.49	0.77	0.23	2 497.68	48.10	7.10	7390
2013	Poland	0.65	0.34	0.54	0.87	0.53	13 539.80	62.85	10.60	23740
2013	Portugal	0.70	0.51	0.65	0.86	0.61	21 256.76	62.10	17.10	27560
2013	Romania	0.58	0.27	0.48	0.84	0.43	8 980.28	49.76	7.10	19490
2013	Serbia	0.59	0.32	0.48	0.81	0.47	6 086.73	53.45	23.00	14060
2013	Singapore	0.88	0.92	1.00	0.85	0.79	51 724.98	80.90	1.90	78530
2013	Slovenia	0.70	0.30	0.55	0.92	0.64	22 722.44	72.68	10.20	29570
2013	Spain	0.81	0.64	0.85	0.93	0.65	28 910.76	71.64	26.10	32220
2013	Sudan	0.26	0.18	0.27	0.38	0.13	1 903.45	22.70	15.20	3510
2013	Sweden	0.84	0.65	0.77	0.90	0.85	53 611.10	94.78	8.20	47350
2013	Switzerland	0.77	0.36	0.59	0.87	0.85	78 356.05	86.34	3.20	63430
2013	Thailand	0.49	0.43	0.48	0.72	0.26	5 558.74	28.94	0.70	14420
2013	Tunisia	0.51	0.51	0.56	0.68	0.30	4 221.84	43.80	15.30	9780
2013	Turkey	0.54	0.27	0.51	0.74	0.35	12 936.63	46.25	9.00	22240
2013	Ukraine	0.53	0.29	0.35	0.89	0.37	3 160.08	40.95	7.20	10510
2013	Uruguay	0.69	0.58	0.70	0.86	0.50	13 541.22	57.69	6.50	18110

2013	Uzbekistan	0.49	0.35	0.47	0.78	0.22	1 921.83	26.80	4.90	6050
2013	VietNam	0.50	0.30	0.42	0.68	0.39	1 505.83	38.50	2.20	5130
2014	Albania	0.50	0.53	0.45	0.71	0.35	4 413.30	60.10	17.50	11360
2014	Argentina	0.63	0.55	0.55	0.86	0.48	10 398.69	64.70	7.30	19250
2014	Armenia	0.59	0.53	0.61	0.77	0.39	3 819.22	54.62	17.60	10490
2014	Austria	0.79	0.63	0.75	0.87	0.76	47 842.75	81.00	6.10	48860
2014	Azerbaijan	0.55	0.43	0.43	0.75	0.46	6 073.40	75.00	4.90	16850
2014	Bahrain	0.81	0.82	0.94	0.78	0.71	22 395.34	90.50	3.80	48600
2014	Belarus	0.61	0.35	0.32	0.89	0.61	6 649.65	59.02	5.40	18360
2014	Belgium	0.76	0.63	0.68	0.89	0.70	44 855.97	85.00	8.70	45870
2014	Brazil	0.60	0.71	0.60	0.74	0.47	11 951.21	54.55	6.90	15430
2014	Bulgaria	0.54	0.25	0.24	0.80	0.59	7 322.81	55.49	11.50	17410
2014	Canada	0.84	0.82	0.91	0.90	0.72	50 427.89	87.12	7.00	45010
2014	Chile	0.71	0.94	0.82	0.82	0.49	14 561.32	61.11	6.50	21910
2014	China	0.55	0.65	0.61	0.67	0.36	6 103.75	47.90	4.10	12570
2014	Colombia	0.62	0.88	0.79	0.73	0.33	7 449.33	52.57	9.10	12910
2014	Croatia	0.63	0.33	0.46	0.79	0.63	13 663.17	68.57	19.30	21600
2014	Cyprus	0.60	0.31	0.47	0.78	0.53	26 843.42	69.33	16.10	29280
2014	Denmark	0.82	0.55	0.66	0.91	0.87	59 437.95	95.99	6.90	49640
2014	Ecuador	0.51	0.49	0.48	0.70	0.33	5 412.13	45.59	3.80	11540
2014	Estonia	0.82	0.76	0.77	0.89	0.79	17 312.74	84.24	7.40	28310
2014	Finland	0.84	0.71	0.77	0.90	0.86	45 550.63	86.53	8.80	42180
2014	France	0.89	0.96	1.00	0.88	0.80	41 480.77	83.75	10.30	41010
2014	Germany	0.79	0.71	0.67	0.89	0.80	44 930.79	86.19	5.00	47950
2014	Greece	0.71	0.80	0.61	0.87	0.65	22 282.24	63.21	26.50	26810
2014	Honduras	0.41	0.33	0.40	0.63	0.20	2 025.82	19.08	5.50	4090
2014	Hungary	0.66	0.45	0.56	0.87	0.57	14 270.11	75.65	7.50	24540
2014	Iceland	0.80	0.49	0.61	0.92	0.86	45 964.00	98.16	5.40	43500
2014	Indonesia	0.45	0.29	0.36	0.68	0.31	3 692.99	17.14	5.90	9930
2014	Ireland	0.78	0.65	0.68	0.96	0.70	52 815.53	83.49	11.90	43200
2014	Israel	0.82	0.86	0.87	0.85	0.72	33 053.19	75.02	5.90	34060
2014	Italy	0.76	0.78	0.75	0.86	0.67	33 666.69	55.64	12.80	36200
2014	Jamaica	0.44	0.20	0.31	0.73	0.28	4 705.40	40.40	13.80	8360
2014	Japan	0.89	0.96	0.94	0.86	0.86	46 484.16	89.11	3.60	40590
2014	Jordan	0.52	0.47	0.52	0.72	0.31	3 396.24	46.20	11.90	9080
2014	Kazakhstan	0.73	0.76	0.75	0.86	0.57	10 646.03	66.00	5.10	22190
2014	Kuwait	0.63	0.43	0.57	0.72	0.59	37 158.67	78.70	1.70	76820
2014	Latvia	0.72	0.71	0.70	0.83	0.62	13 741.46	75.83	10.80	23830
2014	Lithuania	0.73	0.65	0.76	0.86	0.57	14 948.03	72.13	10.70	27850
2014	Luxembourg	0.76	0.55	0.62	0.78	0.87	105 658.52	94.67	7.10	67290
2014	Malaysia	0.61	0.53	0.68	0.71	0.45	10 524.05	63.67	2.90	23790
2014	Malta	0.65	0.47	0.40	0.79	0.77	24 692.43	73.17	5.70	32560
2014	Mauritius	0.53	0.53	0.47	0.69	0.44	9 163.63	44.80	7.80	19920
2014	Mexico	0.57	0.61	0.66	0.74	0.31	9 843.40	44.39	4.80	17610
2014	Mongolia	0.56	0.69	0.61	0.79	0.27	3 880.35	19.94	7.90	10160
2014	Morocco	0.51	0.80	0.69	0.49	0.33	3 125.08	56.80	9.70	6450
2014	Netherlands	0.89	1.00	0.93	0.92	0.82	51 100.84	91.67	8.30	49110
2014	Nicaragua	0.28	0.10	0.09	0.56	0.17	1 775.02	17.60	6.60	4760
2014	Nigeria	0.29	0.33	0.31	0.38	0.19	2 550.47	21.00	7.80	5320
2014	Norway	0.84	0.69	0.76	0.94	0.81	89 175.50	96.30	3.60	67980

2014	Pakistan	0.26	0.33	0.32	0.33	0.12	1 054.24	12.00	6.00	4170
2014	Panama	0.52	0.49	0.37	0.75	0.46	10 357.51	44.92	4.80	20900
2014	Paraguay	0.37	0.25	0.23	0.67	0.22	4 873.43	43.00	6.00	11040
2014	Peru	0.54	0.71	0.63	0.73	0.27	5 996.49	40.20	5.90	10990
2014	Philippines	0.48	0.57	0.48	0.71	0.25	2 612.85	49.60	6.80	7760
2014	Poland	0.65	0.49	0.54	0.84	0.56	14 007.73	66.60	9.20	24520
2014	Portugal	0.69	0.65	0.64	0.82	0.61	21 540.99	64.59	14.50	28150
2014	Romania	0.56	0.47	0.44	0.81	0.44	9 339.28	54.08	6.80	20560
2014	Serbia	0.55	0.41	0.39	0.78	0.47	6 018.17	62.08	21.00	14100
2014	Singapore	0.91	0.90	0.99	0.85	0.88	53 068.36	79.03	2.00	81490
2014	Slovenia	0.65	0.39	0.43	0.91	0.62	23 328.47	71.59	9.80	30590
2014	Spain	0.84	0.78	0.94	0.92	0.66	29 398.61	76.19	24.40	33400
2014	Sudan	0.26	0.27	0.29	0.31	0.18	1 988.95	24.64	19.80	4450
2014	Sweden	0.82	0.61	0.70	0.88	0.89	54 492.60	92.52	8.10	48180
2014	Switzerland	0.73	0.37	0.50	0.86	0.82	79 299.83	87.40	3.00	64210
2014	Thailand	0.46	0.55	0.44	0.66	0.28	5 589.33	34.89	0.80	14680
2014	Tunisia	0.54	0.65	0.64	0.67	0.31	4 299.85	46.16	15.00	10070
2014	Turkey	0.54	0.49	0.56	0.71	0.36	13 346.48	51.04	9.90	23880
2014	Ukraine	0.50	0.43	0.27	0.86	0.38	3 123.90	46.24	9.30	10620
2014	Uruguay	0.74	0.98	0.85	0.81	0.56	13 934.56	61.46	6.60	18780
2014	Uzbekistan	0.47	0.47	0.45	0.73	0.23	2 025.35	35.50	5.10	6250
2014	VietNam	0.47	0.49	0.42	0.61	0.38	1 579.20	41.00	2.10	5470
2015	Albania	0.52	0.59	0.52	0.68	0.35	4 524.37	63.25	17.10	11800
2015	Argentina	0.66	0.59	0.63	0.87	0.49	10 568.16	68.04	6.50	19700
2015	Armenia	0.55	0.53	0.52	0.75	0.39	3 923.71	59.10	18.50	10370
2015	Austria	0.81	0.75	0.83	0.85	0.73	47 789.39	83.94	6.20	49370
2015	Azerbaijan	0.59	0.55	0.56	0.73	0.47	6 064.47	77.00	5.00	14370
2015	Bahrain	0.79	0.78	0.88	0.75	0.74	22 353.33	93.48	3.40	43030
2015	Belarus	0.63	0.46	0.40	0.88	0.62	6 384.82	62.23	5.80	17240
2015	Belgium	0.77	0.64	0.69	0.93	0.69	45 507.23	85.05	8.70	46850
2015	Brazil	0.62	0.72	0.67	0.71	0.48	11 431.16	58.33	8.60	14470
2015	Bulgaria	0.59	0.47	0.40	0.79	0.58	7 663.72	56.66	9.20	17770
2015	Canada	0.84	0.87	0.93	0.88	0.69	50 382.86	90.00	6.90	43990
2015	Chile	0.70	0.84	0.80	0.82	0.50	14 722.36	76.63	6.30	22010
2015	China	0.58	0.73	0.69	0.68	0.36	6 500.42	50.30	4.10	12930
2015	Colombia	0.62	0.82	0.79	0.72	0.36	7 580.28	55.90	8.90	13080
2015	Croatia	0.67	0.56	0.61	0.80	0.61	14 111.25	69.85	17.10	22860
2015	Cyprus	0.60	0.42	0.50	0.78	0.51	27 865.59	71.72	14.90	31610
2015	Denmark	0.83	0.68	0.72	0.93	0.85	60 402.13	96.33	6.30	50550
2015	Ecuador	0.53	0.53	0.56	0.71	0.34	5 330.54	48.94	4.80	10870
2015	Estonia	0.83	0.79	0.83	0.88	0.76	17 620.59	88.41	6.20	28750
2015	Finland	0.86	0.81	0.86	0.92	0.81	45 647.67	86.42	9.60	42920
2015	France	0.87	0.93	0.97	0.86	0.78	41 793.54	78.01	10.40	41730
2015	Germany	0.80	0.73	0.75	0.89	0.77	45 208.06	87.59	4.60	48690
2015	Greece	0.70	0.71	0.59	0.88	0.63	22 337.15	66.83	24.90	26690
2015	Honduras	0.38	0.36	0.36	0.60	0.20	2 067.29	27.10	6.10	4490
2015	Hungary	0.67	0.47	0.59	0.85	0.56	14 850.43	72.83	6.60	25570
2015	Iceland	0.78	0.58	0.62	0.91	0.82	47 505.70	98.20	4.50	46470
2015	Indonesia	0.45	0.33	0.36	0.68	0.30	3 824.27	22.06	6.20	9910
2015	Ireland	0.77	0.68	0.70	0.94	0.68	65 490.86	83.49	9.90	53070
2015	Israel	0.80	0.85	0.87	0.86	0.67	33 123.53	77.35	5.20	35170

2015	Italy	0.77	0.85	0.81	0.83	0.66	33 961.44	58.14	12.00	36640
2015	Jamaica	0.45	0.23	0.34	0.71	0.30	4 722.67	42.22	13.50	8520
2015	Japan	0.87	0.97	0.91	0.84	0.84	47 102.58	91.06	3.40	41940
2015	Jordan	0.51	0.46	0.49	0.73	0.33	3 350.39	60.11	13.10	9360
2015	Kazakhstan	0.73	0.68	0.76	0.85	0.57	10 617.47	70.83	5.00	21760
2015	Kuwait	0.67	0.54	0.61	0.72	0.66	35 969.38	72.00	1.30	52460
2015	Latvia	0.70	0.62	0.65	0.84	0.60	14 409.54	79.20	9.90	24840
2015	Lithuania	0.75	0.74	0.79	0.86	0.60	15 394.81	71.38	9.10	27790
2015	Luxembourg	0.76	0.62	0.67	0.78	0.85	107 638.21	96.38	6.60	67570
2015	Malaysia	0.61	0.60	0.70	0.70	0.44	10 912.15	71.06	3.20	24130
2015	Malta	0.70	0.63	0.60	0.76	0.73	26 426.68	75.96	5.40	35380
2015	Mauritius	0.58	0.60	0.59	0.70	0.45	9 476.67	50.14	7.90	20860
2015	Mexico	0.60	0.74	0.75	0.72	0.31	10 042.14	57.43	4.30	17830
2015	Mongolia	0.54	0.70	0.56	0.77	0.28	3 895.42	22.50	7.50	9880
2015	Morocco	0.51	0.82	0.72	0.48	0.34	3 222.05	57.08	9.20	6890
2015	Netherlands	0.88	0.97	0.93	0.92	0.78	51 871.58	91.72	7.90	50330
2015	Nicaragua	0.33	0.25	0.24	0.55	0.19	1 836.04	19.70	5.90	5090
2015	Nigeria	0.31	0.34	0.36	0.38	0.19	2 549.72	36.00	9.00	5280
2015	Norway	0.82	0.72	0.78	0.92	0.77	90 029.36	96.81	4.50	62880
2015	Pakistan	0.26	0.35	0.32	0.33	0.12	1 081.29	14.00	5.90	4300
2015	Panama	0.51	0.43	0.35	0.73	0.44	10 765.93	51.21	5.10	23700
2015	Paraguay	0.44	0.42	0.41	0.66	0.24	4 950.62	49.72	5.40	11050
2015	Peru	0.54	0.62	0.63	0.72	0.27	6 114.23	40.85	6.50	11180
2015	Philippines	0.53	0.58	0.57	0.69	0.31	2 735.19	36.00	6.30	8010
2015	Poland	0.68	0.69	0.62	0.86	0.57	14 610.88	68.00	7.70	25880
2015	Portugal	0.70	0.65	0.69	0.82	0.60	22 018.01	68.63	12.90	28870
2015	Romania	0.56	0.55	0.45	0.79	0.45	9 660.43	55.76	6.80	21380
2015	Serbia	0.63	0.62	0.61	0.78	0.51	6 157.25	65.32	19.30	14240
2015	Singapore	0.90	0.91	0.98	0.84	0.86	54 009.74	79.01	1.90	81040
2015	Slovenia	0.71	0.56	0.64	0.90	0.60	23 826.13	73.10	9.00	30650
2015	Spain	0.83	0.86	0.93	0.90	0.66	30 549.79	78.69	22.10	34900
2015	Sudan	0.26	0.26	0.25	0.33	0.19	2 019.50	26.61	21.60	4320
2015	Sweden	0.85	0.69	0.79	0.90	0.85	56 339.99	90.61	7.60	49500
2015	Switzerland	0.74	0.47	0.55	0.86	0.81	79 701.98	87.48	3.20	67430
2015	Thailand	0.51	0.57	0.50	0.68	0.35	5 741.36	39.32	0.90	15010
2015	Tunisia	0.55	0.67	0.68	0.66	0.33	4 304.55	46.50	15.40	9890
2015	Turkey	0.57	0.56	0.58	0.75	0.37	13 924.13	53.74	10.30	25460
2015	Ukraine	0.56	0.59	0.43	0.85	0.39	2 828.89	48.88	9.10	10040
2015	Uruguay	0.73	0.85	0.81	0.80	0.59	13 938.77	64.57	7.50	19290
2015	Uzbekistan	0.51	0.57	0.57	0.71	0.24	2 138.57	42.80	5.20	6450
2015	VietNam	0.49	0.59	0.49	0.61	0.38	1 667.17	45.00	2.30	5790
2016	Albania	0.53	0.64	0.59	0.65	0.35	4 681.83	66.36	15.20	12280
2016	Argentina	0.70	0.63	0.71	0.88	0.50	10 239.48	70.97	8.50	19860
2016	Armenia	0.52	0.53	0.43	0.73	0.39	3 917.38	64.35	18.00	10950
2016	Austria	0.82	0.88	0.91	0.84	0.71	48 215.89	84.32	6.50	52680
2016	Azerbaijan	0.63	0.68	0.68	0.72	0.49	5 813.33	78.20	5.00	13430
2016	Bahrain	0.77	0.75	0.83	0.72	0.78	22 272.94	98.00	3.70	42280
2016	Belarus	0.66	0.56	0.49	0.87	0.63	6 215.70	71.11	5.90	16900

2016	Belgium	0.79	0.64	0.71	0.97	0.68	45 850.93	86.52	7.90	49060
2016	Brazil	0.64	0.73	0.73	0.68	0.50	10 965.97	60.87	11.70	13950
2016	Bulgaria	0.64	0.69	0.57	0.79	0.56	8 012.51	59.83	7.70	19400
2016	Canada	0.83	0.92	0.96	0.86	0.67	50 314.42	91.16	7.10	45900
2016	Chile	0.69	0.75	0.78	0.81	0.50	14 777.15	83.56	6.70	22710
2016	China	0.61	0.81	0.77	0.69	0.37	6 908.11	53.20	4.00	13520
2016	Colombia	0.62	0.76	0.79	0.70	0.38	7 633.39	58.14	9.20	13660
2016	Croatia	0.72	0.78	0.75	0.80	0.60	14 707.24	72.70	15.00	24130
2016	Cyprus	0.60	0.53	0.54	0.78	0.49	29 523.87	75.90	13.00	34300
2016	Denmark	0.85	0.81	0.78	0.95	0.82	61 877.98	96.97	6.00	53250
2016	Ecuador	0.56	0.58	0.63	0.71	0.34	5 176.06	54.06	5.20	10830
2016	Estonia	0.83	0.81	0.89	0.88	0.73	18 177.01	87.24	6.80	30870
2016	Finland	0.88	0.92	0.94	0.94	0.76	46 796.34	87.70	9.00	45140
2016	France	0.85	0.90	0.94	0.84	0.75	42 140.04	79.27	10.00	43830
2016	Germany	0.82	0.76	0.84	0.89	0.73	45 844.64	84.17	4.10	51830
2016	Greece	0.69	0.61	0.58	0.89	0.60	22 320.37	69.09	23.60	27330
2016	Honduras	0.36	0.39	0.31	0.57	0.20	2 111.19	29.50	6.70	4790
2016	Hungary	0.67	0.49	0.63	0.83	0.56	15 213.05	79.26	5.00	27200
2016	Iceland	0.77	0.66	0.62	0.89	0.78	49 804.17	98.24	3.30	50990
2016	Indonesia	0.45	0.37	0.36	0.68	0.30	3 968.02	25.45	5.60	10160
2016	Ireland	0.77	0.71	0.72	0.92	0.66	66 046.81	83.80	8.40	58440
2016	Israel	0.78	0.83	0.86	0.86	0.62	33 723.41	79.65	4.80	37450
2016	Italy	0.78	0.92	0.87	0.81	0.65	34 459.21	61.32	11.70	40040
2016	Jamaica	0.45	0.27	0.36	0.69	0.32	4 762.54	44.37	13.20	8800
2016	Japan	0.84	0.98	0.88	0.83	0.83	47 403.05	93.18	3.10	41310
2016	Jordan	0.51	0.46	0.46	0.73	0.35	3 314.30	62.30	15.30	9370
2016	Kazakhstan	0.72	0.59	0.77	0.84	0.57	10 582.70	74.59	4.90	21490
2016	Kuwait	0.71	0.64	0.65	0.73	0.74	35 887.11	78.37	1.20	49900
2016	Latvia	0.68	0.53	0.61	0.85	0.58	14 886.89	79.84	9.60	26670
2016	Lithuania	0.77	0.83	0.83	0.87	0.63	15 984.41	74.38	7.90	29820
2016	Luxembourg	0.77	0.69	0.72	0.78	0.82	110 162.12	98.14	6.30	70450
2016	Malaysia	0.62	0.68	0.72	0.70	0.44	11 244.06	78.79	3.50	24840
2016	Malta	0.74	0.78	0.80	0.73	0.70	26 885.38	78.08	4.70	36000
2016	Mauritius	0.62	0.66	0.70	0.71	0.46	9 833.61	52.19	7.30	22390
2016	Mexico	0.62	0.88	0.85	0.70	0.31	10 183.03	59.54	3.90	18810
2016	Mongolia	0.52	0.71	0.51	0.76	0.28	3 866.25	22.27	10.00	9860
2016	Morocco	0.52	0.83	0.74	0.47	0.34	3 212.82	58.27	9.00	6980
2016	Netherlands	0.87	0.95	0.93	0.92	0.75	52 727.10	90.41	7.00	51520
2016	Nicaragua	0.38	0.41	0.38	0.55	0.21	1 895.22	24.57	4.50	5420
2016	Nigeria	0.33	0.36	0.41	0.38	0.20	2 443.44	25.67	13.40	5120
2016	Norway	0.81	0.76	0.80	0.90	0.73	90 195.96	97.30	4.70	61520
2016	Pakistan	0.26	0.37	0.33	0.32	0.13	1 117.50	12.39	6.00	4330
2016	Panama	0.49	0.37	0.33	0.72	0.42	11 107.24	54.00	5.50	25840
2016	Paraguay	0.50	0.58	0.60	0.64	0.25	5 094.03	53.40	6.00	11560
2016	Peru	0.54	0.54	0.63	0.72	0.27	6 262.37	45.46	6.70	11520
2016	Philippines	0.58	0.59	0.67	0.68	0.38	2 886.90	55.50	5.50	8560
2016	Poland	0.72	0.88	0.70	0.87	0.59	15 076.39	73.30	6.30	27210
2016	Portugal	0.71	0.66	0.75	0.81	0.58	22 533.63	70.42	11.50	30840
2016	Romania	0.56	0.63	0.46	0.77	0.45	10 172.95	59.50	5.90	23950
2016	Serbia	0.71	0.83	0.82	0.78	0.54	6 396.22	67.06	16.90	14990
2016	Singapore	0.88	0.92	0.97	0.84	0.84	55 088.35	84.45	2.10	83760

2016	Slovenia	0.78	0.73	0.85	0.90	0.59	24 568.64	75.50	8.00	33080
2016	Spain	0.81	0.93	0.91	0.88	0.65	31 449.28	80.56	19.60	37380
2016	Sudan	0.25	0.25	0.22	0.36	0.19	2 042.11	14.10	20.60	4460
2016	Sweden	0.87	0.76	0.88	0.92	0.81	56 788.54	89.65	7.10	50720
2016	Switzerland	0.75	0.58	0.60	0.86	0.80	80 448.70	89.13	3.30	68440
2016	Thailand	0.55	0.59	0.55	0.69	0.41	5 916.48	47.50	1.00	15840
2016	Tunisia	0.57	0.69	0.72	0.64	0.35	4 310.53	49.60	15.50	10070
2016	Turkey	0.59	0.63	0.60	0.79	0.38	14 152.83	58.35	10.90	26230
2016	Ukraine	0.61	0.75	0.59	0.84	0.40	2 903.80	53.00	9.50	11040
2016	Uruguay	0.72	0.71	0.78	0.78	0.61	14 124.11	66.40	7.90	21410
2016	Uzbekistan	0.54	0.68	0.69	0.70	0.25	2 229.78	46.79	5.20	6520
2016	VietNam	0.51	0.69	0.57	0.60	0.37	1 752.50	53.00	2.30	6300
2017	Albania	0.59	0.70	0.67	0.72	0.39	4 864.31	71.85	13.70	12800
2017	Argentina	0.72	0.63	0.73	0.87	0.55	10 419.46	74.29	8.40	23000
2017	Armenia	0.56	0.55	0.50	0.74	0.43	4 198.82	64.74	17.80	12540
2017	Austria	0.83	0.85	0.89	0.85	0.74	49 031.11	87.94	5.90	53660
2017	Azerbaijan	0.64	0.68	0.71	0.73	0.50	5 765.41	79.00	5.00	13510
2017	Bahrain	0.79	0.77	0.81	0.75	0.81	22 167.03	95.88	3.60	45030
2017	Belarus	0.71	0.72	0.61	0.87	0.66	6 375.29	74.44	5.70	17580
2017	Belgium	0.80	0.70	0.73	0.97	0.69	46 415.28	87.68	7.20	50900
2017	Brazil	0.69	0.85	0.83	0.72	0.51	11 021.72	67.47	12.90	14250
2017	Bulgaria	0.68	0.78	0.66	0.80	0.57	8 353.97	63.41	6.20	20870
2017	Canada	0.83	0.91	0.94	0.87	0.67	51 225.67	92.70	6.40	47700
2017	Chile	0.71	0.78	0.80	0.82	0.52	14 740.61	82.33	7.00	23460
2017	China	0.64	0.86	0.81	0.70	0.42	7 346.84	54.30	3.90	14330
2017	Colombia	0.66	0.84	0.84	0.72	0.41	7 620.92	62.26	9.40	13910
2017	Croatia	0.71	0.77	0.71	0.81	0.60	15 396.80	67.10	12.40	26410
2017	Cyprus	0.69	0.67	0.66	0.79	0.61	30 758.31	80.74	11.10	36810
2017	Denmark	0.88	0.91	0.89	0.95	0.81	63 216.00	97.10	5.80	56570
2017	Ecuador	0.59	0.63	0.68	0.73	0.36	5 205.76	54.74	4.60	11360
2017	Estonia	0.84	0.86	0.90	0.88	0.75	19 153.39	88.10	5.80	33040
2017	Finland	0.88	0.96	0.95	0.95	0.74	48 177.08	87.47	8.80	47590
2017	France	0.86	0.93	0.96	0.85	0.77	42 980.73	80.50	9.40	45580
2017	Germany	0.85	0.84	0.89	0.90	0.76	46 862.04	84.39	3.80	54210
2017	Greece	0.74	0.74	0.70	0.89	0.62	22 651.01	69.89	21.50	28490
2017	Honduras	0.40	0.47	0.41	0.59	0.21	2 176.29	32.14	5.50	5220
2017	Hungary	0.70	0.60	0.68	0.83	0.58	15 912.01	76.75	4.00	28310
2017	Iceland	0.80	0.67	0.68	0.92	0.81	50 690.39	98.26	3.30	53920
2017	Indonesia	0.49	0.50	0.47	0.68	0.31	4 120.44	32.34	5.50	10590
2017	Ireland	0.80	0.82	0.78	0.94	0.68	71 296.09	84.11	6.70	62640
2017	Israel	0.79	0.83	0.84	0.86	0.66	34 251.95	81.58	4.20	38640
2017	Italy	0.80	0.94	0.91	0.82	0.66	35 086.48	63.08	11.30	41810
2017	Jamaica	0.46	0.29	0.34	0.69	0.36	4 785.98	55.07	11.70	9330
2017	Japan	0.86	0.98	0.91	0.84	0.83	48 510.61	91.73	2.80	42430
2017	Jordan	0.53	0.47	0.47	0.74	0.39	3 303.42	66.79	18.30	9950
2017	Kazakhstan	0.74	0.72	0.82	0.84	0.57	10 867.74	76.43	4.90	22160
2017	Kuwait	0.72	0.67	0.72	0.71	0.74	33 359.40	98.00	1.30	58770
2017	Latvia	0.69	0.61	0.64	0.83	0.60	15 507.65	80.11	8.70	28580
2017	Lithuania	0.76	0.82	0.81	0.85	0.63	16 903.68	77.62	7.10	32540
2017	Luxembourg	0.80	0.82	0.82	0.78	0.81	109 452.96	97.36	5.80	72410

2017	Malaysia	0.67	0.78	0.80	0.70	0.50	11 736.99	80.14	3.40	25910
2017	Malta	0.77	0.81	0.82	0.76	0.73	28 283.03	81.01	4.00	38300
2017	Mauritius	0.65	0.68	0.72	0.72	0.50	10 199.48	55.40	7.10	23540
2017	Mexico	0.65	0.91	0.89	0.70	0.36	10 277.88	63.85	3.40	19210
2017	Mongolia	0.55	0.72	0.56	0.77	0.32	3 997.47	23.71	8.80	9720
2017	Morocco	0.52	0.80	0.70	0.50	0.36	3 305.87	61.76	10.60	7170
2017	Netherlands	0.87	0.97	0.93	0.92	0.76	53 942.09	93.20	5.90	55500
2017	Nicaragua	0.40	0.40	0.39	0.57	0.25	1 957.87	27.86	3.70	5690
2017	Nigeria	0.35	0.42	0.47	0.40	0.19	2 399.73	7.47	17.50	5030
2017	Norway	0.83	0.87	0.88	0.90	0.72	91 549.04	96.36	4.20	66400
2017	Pakistan	0.31	0.44	0.44	0.34	0.14	1 155.31	17.11	6.00	4500
2017	Panama	0.55	0.55	0.50	0.72	0.44	11 529.23	59.95	6.10	28610
2017	Paraguay	0.51	0.57	0.58	0.66	0.30	5 269.72	61.08	6.10	12170
2017	Peru	0.59	0.70	0.72	0.72	0.33	6 314.29	50.45	6.90	11930
2017	Philippines	0.61	0.77	0.77	0.70	0.37	3 042.69	60.05	5.70	9020
2017	Poland	0.76	0.89	0.82	0.87	0.58	15 802.70	75.99	5.00	28820
2017	Portugal	0.76	0.78	0.84	0.81	0.62	23 380.69	73.79	9.20	32270
2017	Romania	0.61	0.67	0.56	0.78	0.50	10 980.84	63.75	4.90	26760
2017	Serbia	0.71	0.82	0.78	0.78	0.58	6 565.46	70.33	14.90	15540
2017	Singapore	0.88	0.94	0.98	0.85	0.82	57 527.56	84.45	2.20	88090
2017	Slovenia	0.77	0.77	0.82	0.89	0.61	25 729.58	78.89	6.60	35850
2017	Spain	0.83	0.96	0.93	0.88	0.67	32 308.59	84.60	17.20	39540
2017	Sudan	0.25	0.20	0.19	0.37	0.18	2 086.97	30.87	19.60	4270
2017	Sweden	0.88	0.85	0.91	0.93	0.80	57 467.25	93.01	6.90	52850
2017	Switzerland	0.80	0.71	0.72	0.86	0.82	80 964.60	89.69	3.10	67870
2017	Thailand	0.60	0.62	0.59	0.74	0.47	6 142.41	52.89	1.20	16640
2017	Tunisia	0.60	0.75	0.76	0.65	0.38	4 343.45	55.50	15.50	10280
2017	Turkey	0.65	0.74	0.75	0.80	0.40	14 972.88	64.68	10.90	27550
2017	Ukraine	0.61	0.72	0.58	0.84	0.42	2 988.50	58.89	9.70	12140
2017	Uruguay	0.75	0.81	0.83	0.78	0.66	14 301.79	70.32	7.90	21730
2017	Uzbekistan	0.58	0.72	0.74	0.72	0.29	2 290.38	48.70	5.80	6650
2017	VietNam	0.55	0.69	0.65	0.63	0.38	1 852.88	58.14	2.20	6650
2018	Albania	0.65	0.76	0.74	0.79	0.43	5 074.86	70.74	12.30	13530
2018	Argentina	0.73	0.62	0.75	0.86	0.59	10 049.56	76.20	9.20	22470
2018	Armenia	0.59	0.57	0.56	0.75	0.47	4 406.76	68.25	19.00	13250
2018	Austria	0.83	0.83	0.87	0.85	0.77	50 051.79	87.48	5.20	56530
2018	Azerbaijan	0.66	0.68	0.73	0.74	0.51	5 801.44	79.80	4.90	13790
2018	Bahrain	0.81	0.80	0.80	0.79	0.85	21 465.76	98.64	3.90	44700
2018	Belarus	0.76	0.88	0.74	0.87	0.69	6 586.30	79.13	4.80	18650
2018	Belgium	0.81	0.76	0.76	0.97	0.69	47 030.58	88.66	6.00	53180
2018	Brazil	0.73	0.97	0.92	0.75	0.52	11 130.72	70.43	12.40	14600
2018	Bulgaria	0.72	0.87	0.76	0.81	0.58	8 674.72	64.78	5.30	22190
2018	Canada	0.83	0.91	0.93	0.87	0.67	51 734.27	93.55	5.90	49390
2018	Chile	0.74	0.82	0.83	0.83	0.54	15 076.86	82.82	7.40	23620
2018	China	0.68	0.90	0.86	0.71	0.47	7 807.06	58.43	3.80	15520
2018	Colombia	0.69	0.92	0.88	0.74	0.44	7 698.11	64.13	9.70	14480
2018	Croatia	0.70	0.77	0.68	0.82	0.61	15 971.15	75.29	9.90	28140
2018	Cyprus	0.77	0.82	0.78	0.81	0.73	31 977.81	84.43	8.40	38930
2018	Denmark	0.92	1.00	1.00	0.95	0.80	64 271.88	97.32	5.10	59060
2018	Ecuador	0.61	0.67	0.73	0.74	0.37	5 180.60	55.42	3.70	11530

2018	Estonia	0.85	0.91	0.90	0.88	0.76	19 918.16	89.36	5.40	35620
2018	Finland	0.88	1.00	0.97	0.95	0.73	48 744.38	88.89	7.40	49950
2018	France	0.88	0.97	0.98	0.86	0.80	43 631.40	82.04	9.00	47600
2018	Germany	0.88	0.92	0.93	0.90	0.80	47 313.85	87.04	3.40	56440
2018	Greece	0.78	0.88	0.82	0.89	0.64	23 050.85	72.24	19.30	29330
2018	Honduras	0.45	0.54	0.51	0.60	0.23	2 222.61	34.03	5.60	5360
2018	Hungary	0.73	0.71	0.74	0.84	0.61	16 792.49	76.07	3.60	30660
2018	Iceland	0.83	0.69	0.73	0.94	0.83	51 677.65	99.01	3.10	55340
2018	Indonesia	0.53	0.62	0.57	0.69	0.32	4 284.76	39.90	5.20	11300
2018	Ireland	0.83	0.93	0.83	0.96	0.70	76 415.36	84.52	5.80	66860
2018	Israel	0.80	0.83	0.83	0.86	0.71	34 767.38	83.73	4.00	40310
2018	Italy	0.82	0.96	0.95	0.83	0.68	35 484.28	74.39	10.60	43590
2018	Jamaica	0.47	0.31	0.32	0.70	0.39	4 853.16	55.18	9.10	9590
2018	Japan	0.88	0.98	0.95	0.84	0.84	48 766.13	91.28	2.40	43220
2018	Jordan	0.56	0.48	0.49	0.74	0.44	3 306.66	67.24	18.60	10210
2018	Kazakhstan	0.76	0.84	0.87	0.84	0.57	11 165.57	78.90	4.80	22940
2018	Kuwait	0.74	0.69	0.79	0.69	0.74	33 112.13	99.60	1.10	58550
2018	Latvia	0.70	0.69	0.67	0.81	0.62	16 257.91	83.58	7.40	30290
2018	Lithuania	0.75	0.80	0.80	0.83	0.63	17 737.60	79.72	6.10	35220
2018	Luxembourg	0.83	0.94	0.92	0.78	0.80	110 701.88	97.06	5.10	74310
2018	Malaysia	0.72	0.89	0.89	0.70	0.56	12 131.71	81.20	3.30	27340
2018	Malta	0.80	0.85	0.84	0.80	0.77	28 727.02	81.66	3.70	41160
2018	Mauritius	0.67	0.69	0.73	0.73	0.54	10 577.18	58.60	6.90	25460
2018	Mexico	0.68	0.94	0.92	0.70	0.42	10 385.83	65.77	3.30	19710
2018	Mongolia	0.58	0.74	0.60	0.79	0.36	4 210.85	47.13	7.80	10950
2018	Morocco	0.52	0.78	0.67	0.53	0.37	3 366.85	64.80	9.40	7480
2018	Netherlands	0.88	0.99	0.93	0.92	0.78	54 894.13	92.57	4.90	58560
2018	Nicaragua	0.42	0.39	0.40	0.58	0.28	1 868.41	31.65	5.50	5570
2018	Nigeria	0.38	0.48	0.53	0.43	0.19	2 383.42	18.88	22.60	5040
2018	Norway	0.86	0.98	0.95	0.90	0.71	91 964.26	96.49	3.90	72330
2018	Pakistan	0.36	0.50	0.55	0.37	0.15	1 197.84	17.09	5.80	4770
2018	Panama	0.61	0.72	0.66	0.71	0.45	11 743.60	61.81	6.00	29530
2018	Paraguay	0.53	0.57	0.56	0.67	0.35	5 368.98	64.99	6.20	12700
2018	Peru	0.65	0.87	0.82	0.73	0.39	6 453.08	55.05	6.70	12430
2018	Philippines	0.65	0.94	0.88	0.72	0.35	3 190.79	51.54	5.30	9650
2018	Poland	0.79	0.89	0.93	0.87	0.58	16 648.77	77.54	3.90	30650
2018	Portugal	0.80	0.90	0.93	0.82	0.66	24 085.42	74.66	7.20	34080
2018	Romania	0.67	0.71	0.66	0.79	0.55	11 539.78	70.68	4.20	28710
2018	Serbia	0.72	0.81	0.74	0.79	0.62	6 898.18	73.36	14.10	16830
2018	Singapore	0.88	0.97	0.99	0.86	0.80	59 260.57	88.17	2.10	88930
2018	Slovenia	0.77	0.81	0.80	0.89	0.62	26 760.48	79.75	5.20	38290
2018	Spain	0.84	0.98	0.94	0.89	0.70	32 949.08	86.11	15.30	40790
2018	Sudan	0.24	0.14	0.15	0.39	0.18	2 095.05	31.35	19.50	4360
2018	Sweden	0.89	0.94	0.94	0.94	0.78	57 911.23	93.75	6.50	54570
2018	Switzerland	0.85	0.84	0.85	0.87	0.84	82 792.26	91.42	2.50	70230
2018	Thailand	0.65	0.65	0.64	0.79	0.53	6 379.60	56.82	1.10	17630
2018	Tunisia	0.63	0.80	0.81	0.66	0.41	4 408.33	64.19	15.50	10640
2018	Turkey	0.71	0.86	0.89	0.81	0.43	15 186.83	71.04	10.90	27510
2018	Ukraine	0.62	0.69	0.57	0.84	0.44	3 106.04	62.55	9.00	12960
2018	Uruguay	0.79	0.92	0.89	0.77	0.70	14 317.46	74.38	8.40	22250

2018	Uzbekistan	0.62	0.76	0.79	0.74	0.33	2 373.54	55.20	9.30	7120
2018	VietNam	0.59	0.69	0.74	0.65	0.39	1 964.36	69.85	2.20	7290
2019	Albania	0.70	0.80	0.79	0.79	0.51	5 207.31	69.64	11.50	14040
2019	Argentina	0.78	0.74	0.80	0.88	0.66	9 742.46	78.10	9.80	22080
2019	Armenia	0.65	0.66	0.63	0.77	0.56	4 732.07	68.32	18.30	14470
2019	Austria	0.86	0.90	0.91	0.88	0.80	50 536.66	87.75	4.80	58940
2019	Azerbaijan	0.68	0.69	0.72	0.75	0.58	5 895.20	81.40	4.80	14410
2019	Bahrain	0.82	0.79	0.79	0.82	0.84	20 936.08	99.70	4.70	44330
2019	Belarus	0.79	0.82	0.72	0.88	0.76	6 725.06	82.79	4.20	19520
2019	Belgium	0.81	0.71	0.71	0.96	0.75	47 611.91	90.37	5.40	55370
2019	Brazil	0.75	0.94	0.90	0.77	0.59	11 203.27	73.42	12.00	14980
2019	Bulgaria	0.76	0.88	0.77	0.83	0.68	9 058.74	67.95	4.30	24180
2019	Canada	0.83	0.93	0.89	0.89	0.73	51 956.84	94.40	5.80	50010
2019	Chile	0.78	0.84	0.84	0.85	0.65	15 039.92	83.31	7.20	25040
2019	China	0.74	0.93	0.88	0.72	0.61	8 242.05	62.55	3.60	16730
2019	Colombia	0.70	0.90	0.82	0.76	0.53	7 843.54	65.01	10.50	15280
2019	Croatia	0.74	0.83	0.72	0.83	0.67	16 519.04	79.08	7.80	29790
2019	Cyprus	0.82	0.89	0.83	0.83	0.82	32 517.24	86.06	7.10	39850
2019	Denmark	0.95	0.98	0.99	0.95	0.90	65 867.00	98.05	5.00	62120
2019	Ecuador	0.66	0.74	0.77	0.76	0.44	5 094.95	56.78	3.80	11520
2019	Estonia	0.90	0.96	0.95	0.90	0.84	20 835.80	89.53	4.40	37940
2019	Finland	0.91	0.98	0.97	0.95	0.82	49 310.46	89.61	6.70	51800
2019	France	0.88	0.94	0.93	0.86	0.83	44 192.81	83.34	8.40	50400
2019	Germany	0.86	0.84	0.83	0.92	0.84	47 469.48	88.13	3.20	57410
2019	Greece	0.79	0.83	0.76	0.89	0.73	23 503.58	75.67	17.30	30620
2019	Honduras	0.45	0.52	0.49	0.58	0.28	2 244.45	35.92	5.40	5520
2019	Hungary	0.75	0.69	0.74	0.84	0.67	17 579.84	80.37	3.30	33070
2019	Iceland	0.87	0.73	0.76	0.94	0.91	51 852.43	99.81	3.90	59590
2019	Indonesia	0.59	0.68	0.63	0.71	0.44	4 450.64	47.69	5.20	11940
2019	Ireland	0.84	0.89	0.80	0.96	0.75	79 572.31	89.06	5.00	69190
2019	Israel	0.82	0.77	0.79	0.88	0.79	35 286.86	86.79	3.80	41750
2019	Italy	0.82	0.89	0.89	0.84	0.74	35 998.97	78.89	9.90	45240
2019	Jamaica	0.50	0.34	0.35	0.70	0.45	4 874.17	55.29	7.70	9930
2019	Japan	0.89	0.99	0.93	0.86	0.88	49 000.26	91.79	2.40	43760
2019	Jordan	0.54	0.41	0.43	0.71	0.50	3 325.81	67.70	19.10	10500
2019	Kazakhstan	0.80	0.86	0.90	0.86	0.64	11 518.52	81.88	4.80	24040
2019	Kuwait	0.77	0.80	0.82	0.72	0.76	32 702.25	99.54	1.20	58930
2019	Latvia	0.74	0.63	0.62	0.87	0.73	16 703.20	86.14	6.30	31590
2019	Lithuania	0.81	0.77	0.83	0.88	0.73	18 556.26	81.58	6.30	37420
2019	Luxembourg	0.83	0.82	0.84	0.80	0.85	111 043.53	99.18	5.40	74110
2019	Malaysia	0.75	0.87	0.87	0.73	0.66	12 486.67	84.21	3.30	28780
2019	Malta	0.83	0.84	0.83	0.81	0.84	29 149.58	85.78	3.60	43240
2019	Mauritius	0.69	0.67	0.71	0.76	0.61	10 892.54	64.00	6.70	26790
2019	Mexico	0.71	0.88	0.87	0.74	0.50	10 267.50	70.07	3.50	19860
2019	Mongolia	0.62	0.67	0.56	0.80	0.49	4 352.64	51.08	10.00	11400
2019	Morocco	0.55	0.64	0.60	0.57	0.47	3 407.79	74.38	10.20	7700
2019	Netherlands	0.90	0.98	0.92	0.93	0.85	55 450.50	93.29	4.40	59700
2019	Nicaragua	0.47	0.46	0.47	0.60	0.33	1 777.77	35.43	6.10	5470
2019	Nigeria	0.41	0.49	0.52	0.44	0.27	2 374.37	30.28	23.20	5180
2019	Norway	0.88	0.94	0.91	0.92	0.81	92 123.71	98.00	3.70	70330

2019	Pakistan	0.39	0.51	0.59	0.38	0.20	1 185.46	17.07	6.90	4790
2019	Panama	0.64	0.65	0.64	0.73	0.55	11 902.33	63.63	7.10	30550
2019	Paraguay	0.59	0.66	0.63	0.68	0.45	5 280.18	68.52	6.60	12760
2019	Peru	0.68	0.81	0.79	0.76	0.48	6 489.57	59.95	6.60	12820
2019	Philippines	0.67	0.84	0.81	0.74	0.47	3 340.13	43.03	5.10	10220
2019	Poland	0.82	0.93	0.89	0.88	0.69	17 409.03	84.52	3.30	32790
2019	Portugal	0.81	0.86	0.88	0.83	0.73	24 679.03	75.35	6.70	35940
2019	Romania	0.71	0.76	0.69	0.80	0.65	12 079.55	73.66	3.90	31840
2019	Serbia	0.73	0.82	0.77	0.81	0.62	7 229.94	77.42	11.60	17910
2019	Singapore	0.90	0.97	0.98	0.87	0.85	59 374.44	88.95	2.30	90320
2019	Slovenia	0.81	0.84	0.83	0.91	0.70	27 421.03	83.11	4.50	40530
2019	Spain	0.86	0.91	0.91	0.89	0.78	33 352.33	90.72	14.10	42250
2019	Sudan	0.28	0.18	0.23	0.37	0.23	2 018.41	31.84	22.10	4140
2019	Sweden	0.91	0.88	0.92	0.94	0.87	58 112.44	94.49	7.00	56670
2019	Switzerland	0.87	0.87	0.84	0.88	0.90	83 093.19	93.15	2.30	73620
2019	Thailand	0.71	0.71	0.72	0.78	0.62	6 505.72	66.65	1.00	18530
2019	Tunisia	0.64	0.74	0.71	0.68	0.52	4 404.99	66.70	14.90	10720
2019	Turkey	0.74	0.88	0.87	0.82	0.53	15 125.88	73.98	13.70	26860
2019	Ukraine	0.66	0.75	0.63	0.85	0.52	3 224.56	64.58	8.50	13750
2019	Uruguay	0.82	0.89	0.87	0.81	0.78	14 315.99	76.95	8.90	22810
2019	Uzbekistan	0.64	0.78	0.79	0.74	0.40	2 464.48	56.25	9.00	7400
2019	VietNam	0.63	0.70	0.69	0.67	0.53	2 082.24	68.70	2.20	7890
2020	Albania	0.74	0.85	0.84	0.80	0.58	5 064.06	69.40	11.70	13580
2020	Argentina	0.83	0.86	0.85	0.91	0.73	8 692.71	81.90	11.60	20210
2020	Armenia	0.71	0.75	0.70	0.79	0.65	4 364.31	68.40	18.10	13080
2020	Austria	0.89	0.98	0.95	0.90	0.82	47 008.71	90.30	5.40	51460
2020	Azerbaijan	0.71	0.69	0.71	0.77	0.65	5 593.22	83.00	7.20	14300
2020	Bahrain	0.82	0.77	0.79	0.84	0.83	19 018.55	109.00	5.90	43560
2020	Belarus	0.81	0.75	0.71	0.89	0.83	6 678.00	82.70	4.10	19320
2020	Belgium	0.80	0.65	0.66	0.95	0.80	44 361.25	92.10	5.70	48030
2020	Brazil	0.77	0.90	0.87	0.78	0.65	10 672.18	76.40	13.80	14550
2020	Bulgaria	0.80	0.89	0.77	0.84	0.78	8 743.22	69.60	5.20	23780
2020	Canada	0.84	0.94	0.84	0.90	0.78	48 617.09	96.10	9.60	47500
2020	Chile	0.83	0.86	0.85	0.86	0.76	14 050.18	84.30	10.80	23980
2020	China	0.79	0.96	0.91	0.74	0.74	8 405.18	70.80	4.20	17200
2020	Colombia	0.72	0.87	0.76	0.77	0.61	7 228.44	61.50	16.10	14280
2020	Croatia	0.77	0.89	0.75	0.84	0.73	15 204.35	78.70	9.00	28630
2020	Cyprus	0.87	0.95	0.87	0.84	0.91	30 438.47	87.40	7.60	36840
2020	Denmark	0.98	0.96	0.97	0.96	1.00	63 880.09	98.90	5.70	62180
2020	Ecuador	0.70	0.80	0.81	0.78	0.51	4 628.32	59.50	5.30	10580
2020	Estonia	0.95	1.00	0.99	0.93	0.92	20 161.65	91.50	6.80	38320
2020	Finland	0.95	0.95	0.97	0.95	0.91	47 863.80	95.40	7.80	51650
2020	France	0.87	0.90	0.88	0.86	0.87	40 520.65	91.50	8.00	51560
2020	Germany	0.85	0.75	0.74	0.94	0.89	45 065.08	94.80	3.80	55220
2020	Greece	0.80	0.79	0.71	0.89	0.81	21 577.21	79.70	16.40	31580
2020	Honduras	0.45	0.49	0.46	0.56	0.32	2 010.54	39.70	9.40	5050
2020	Hungary	0.77	0.68	0.75	0.85	0.73	16 745.09	82.50	4.10	32260
2020	Iceland	0.91	0.77	0.79	0.95	0.98	47 630.33	100.60	6.40	62850
2020	Indonesia	0.66	0.75	0.68	0.73	0.57	4 312.35	75.50	7.10	11750
2020	Ireland	0.84	0.86	0.77	0.95	0.81	81 297.07	93.60	5.80	64000
2020	Israel	0.84	0.71	0.75	0.89	0.87	33 817.42	91.60	4.30	43110

2020	Italy	0.82	0.82	0.83	0.85	0.79	32 901.88	83.40	9.30	42270
2020	Jamaica	0.54	0.37	0.39	0.71	0.52	4 357.96	55.50	6.90	8850
2020	Japan	0.90	0.99	0.91	0.87	0.92	0.00	92.30	2.80	43960
2020	Jordan	0.53	0.33	0.36	0.68	0.55	3 241.66	68.60	22.70	10320
2020	Kazakhstan	0.84	0.88	0.92	0.89	0.70	11 075.01	84.40	4.90	24380
2020	Kuwait	0.79	0.90	0.84	0.75	0.79	0.00	103.00	1.30	58800
2020	Latvia	0.78	0.58	0.58	0.92	0.84	16 201.83	86.60	8.10	32780
2020	Lithuania	0.87	0.74	0.85	0.92	0.82	18 390.76	79.20	8.50	38360
2020	Luxembourg	0.83	0.70	0.76	0.81	0.91	107 457.98	101.30	6.40	73910
2020	Malaysia	0.79	0.86	0.85	0.75	0.76	11 637.35	87.00	4.50	27370
2020	Malta	0.85	0.83	0.81	0.83	0.92	26 013.51	86.60	4.40	38800
2020	Mauritius	0.72	0.64	0.70	0.79	0.67	9 272.84	64.30	9.20	22390
2020	Mexico	0.73	0.82	0.82	0.77	0.59	9 322.34	71.30	4.40	18170
2020	Mongolia	0.65	0.61	0.53	0.81	0.61	4 053.63	63.40	7.00	10930
2020	Morocco	0.57	0.51	0.52	0.62	0.58	3 126.96	76.70	12.20	7190
2020	Netherlands	0.92	0.96	0.91	0.93	0.93	53 081.31	96.50	4.90	60120
2020	Nicaragua	0.51	0.52	0.55	0.61	0.38	1 721.82	43.00	7.30	5410
2020	Nigeria	0.44	0.49	0.52	0.45	0.35	2 273.22	53.10	25.10	5000
2020	Norway	0.91	0.90	0.88	0.94	0.90	90 884.61	101.00	4.60	66020
2020	Pakistan	0.42	0.52	0.63	0.38	0.24	1 168.34	28.70	7.60	4770
2020	Panama	0.67	0.58	0.62	0.74	0.65	9 611.29	67.50	18.50	25010
2020	Paraguay	0.65	0.75	0.71	0.70	0.54	5 162.96	70.70	7.70	12590
2020	Peru	0.71	0.76	0.75	0.79	0.58	5 685.37	62.20	13.90	11490
2020	Philippines	0.69	0.75	0.73	0.75	0.58	2 980.02	69.30	10.40	9040
2020	Poland	0.85	0.96	0.86	0.90	0.80	16 945.24	84.30	3.20	33220
2020	Portugal	0.83	0.82	0.84	0.85	0.79	22 769.68	83.70	7.10	33980
2020	Romania	0.76	0.81	0.72	0.80	0.76	11 664.85	79.40	5.00	31410
2020	Serbia	0.75	0.82	0.79	0.83	0.62	7 197.46	78.30	10.10	18650
2020	Singapore	0.92	0.98	0.96	0.89	0.89	56 349.03	83.70	3.00	86480
2020	Slovenia	0.85	0.86	0.85	0.93	0.79	25 759.22	84.20	5.00	41980
2020	Spain	0.88	0.85	0.89	0.90	0.85	29 600.35	91.10	15.50	43000
2020	Sudan	0.32	0.21	0.31	0.36	0.28	1 939.93	32.80	26.80	3970
2020	Sweden	0.94	0.82	0.90	0.95	0.96	56 068.26	99.60	8.50	56270
2020	Switzerland	0.89	0.90	0.83	0.89	0.95	80 132.43	93.20	3.10	80500
2020	Thailand	0.76	0.77	0.79	0.78	0.70	6 094.43	70.00	2.00	17730
2020	Tunisia	0.65	0.69	0.62	0.70	0.64	3 983.94	68.50	17.40	9990
2020	Turkey	0.77	0.89	0.86	0.83	0.63	15 225.61	79.90	13.10	27780
2020	Ukraine	0.71	0.81	0.68	0.86	0.59	3 115.86	66.60	9.20	13260
2020	Uruguay	0.85	0.86	0.84	0.85	0.86	13 430.72	78.00	10.40	21630
2020	Uzbekistan	0.67	0.81	0.78	0.74	0.47	2 457.40	57.30	10.50	7350
2020	VietNam	0.67	0.70	0.65	0.68	0.67	2 123.46	71.90	3.30	8200

ДОДАТОК Б

Діаграми розсіювання для пар показників по країнам.

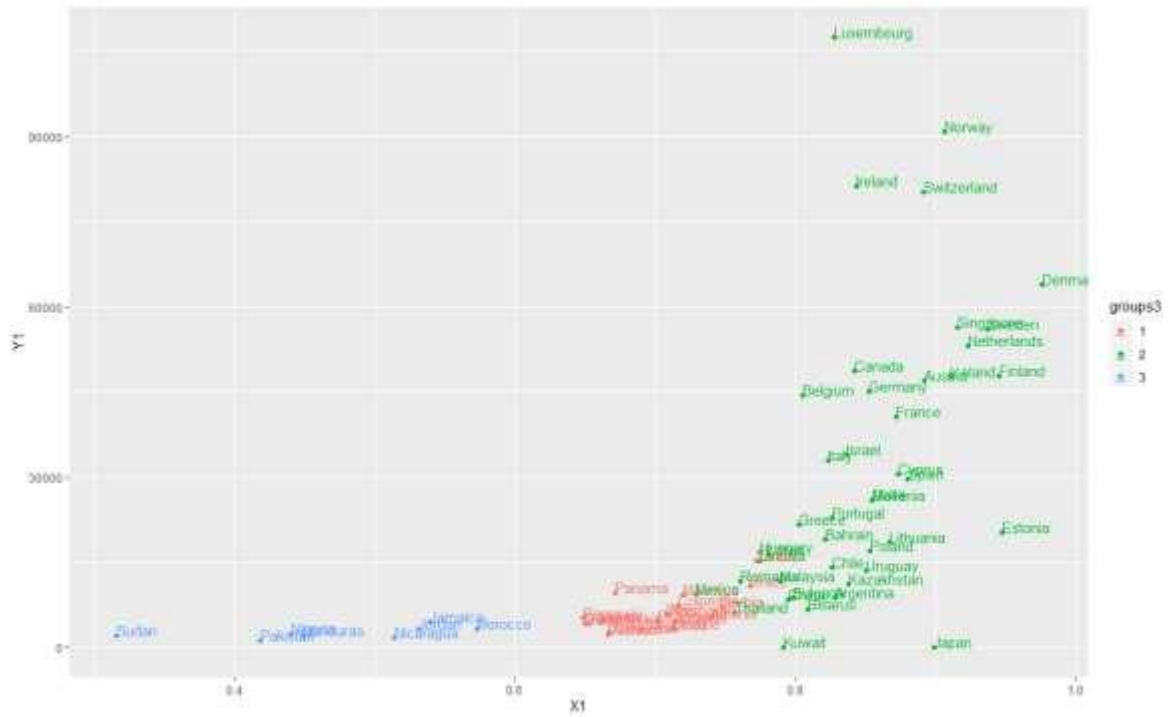


Рис. Б.1. Діаграма розсіювання для пари Y1 та X1

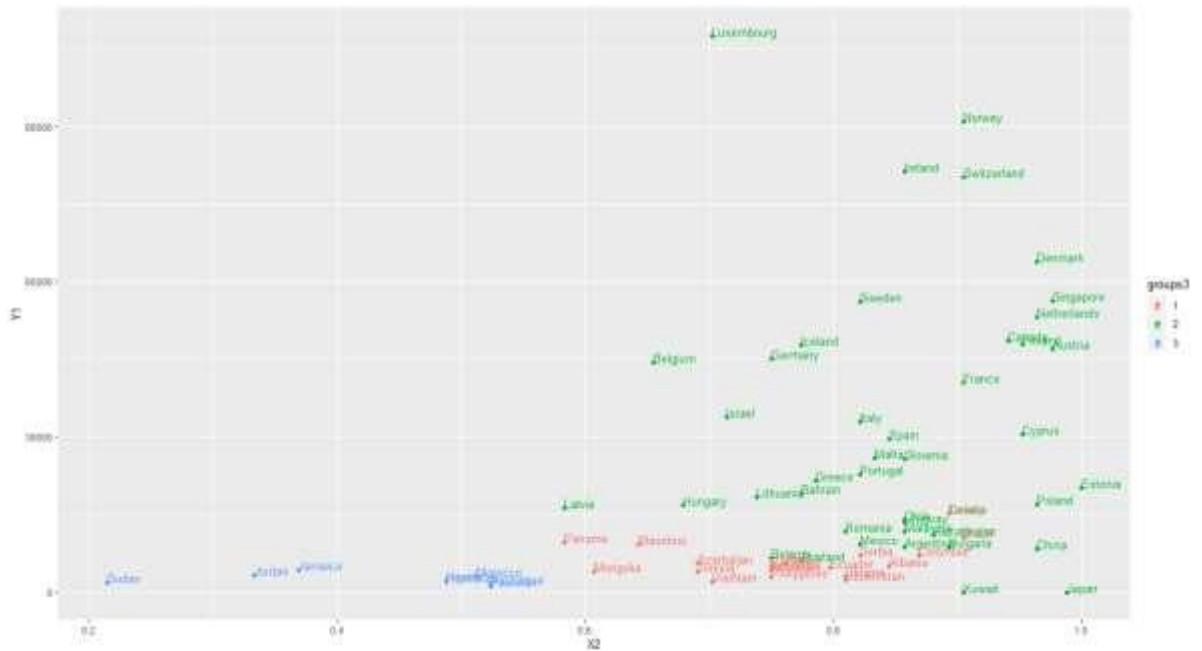


Рис. Б.2. Діаграма розсіювання для пари Y1 та X2

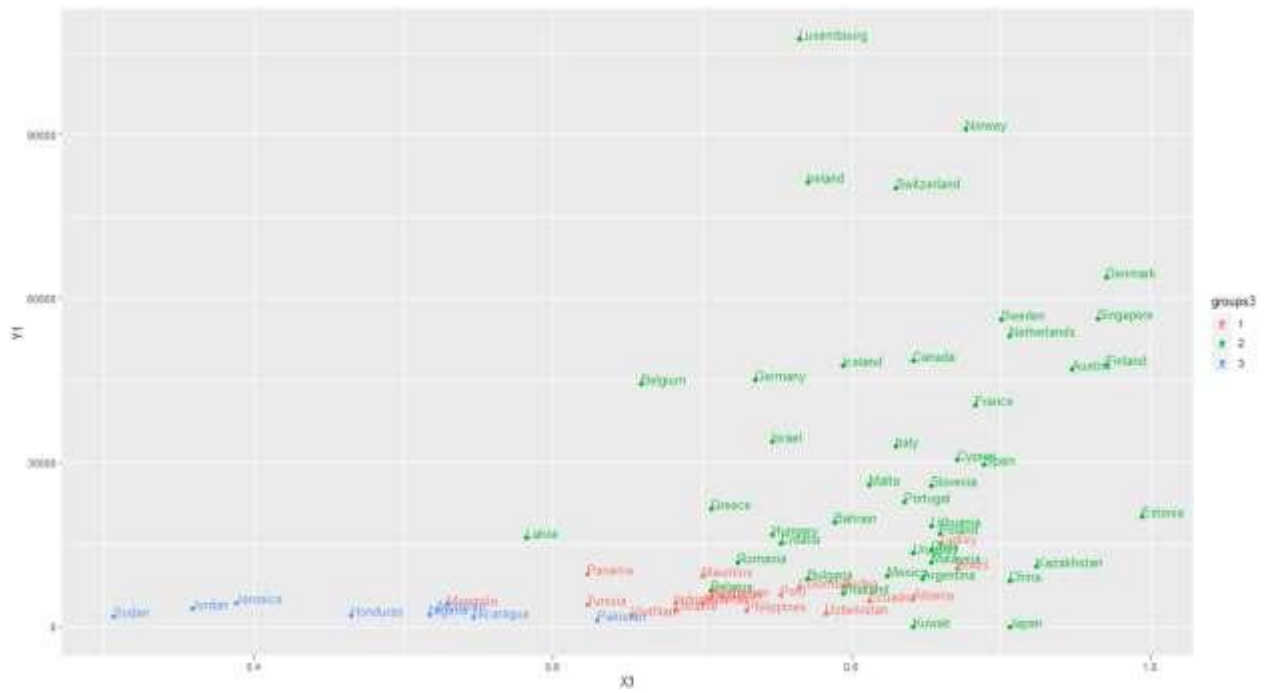


Рис. Б.3. Діаграма розсіювання для пари Y1 та X3

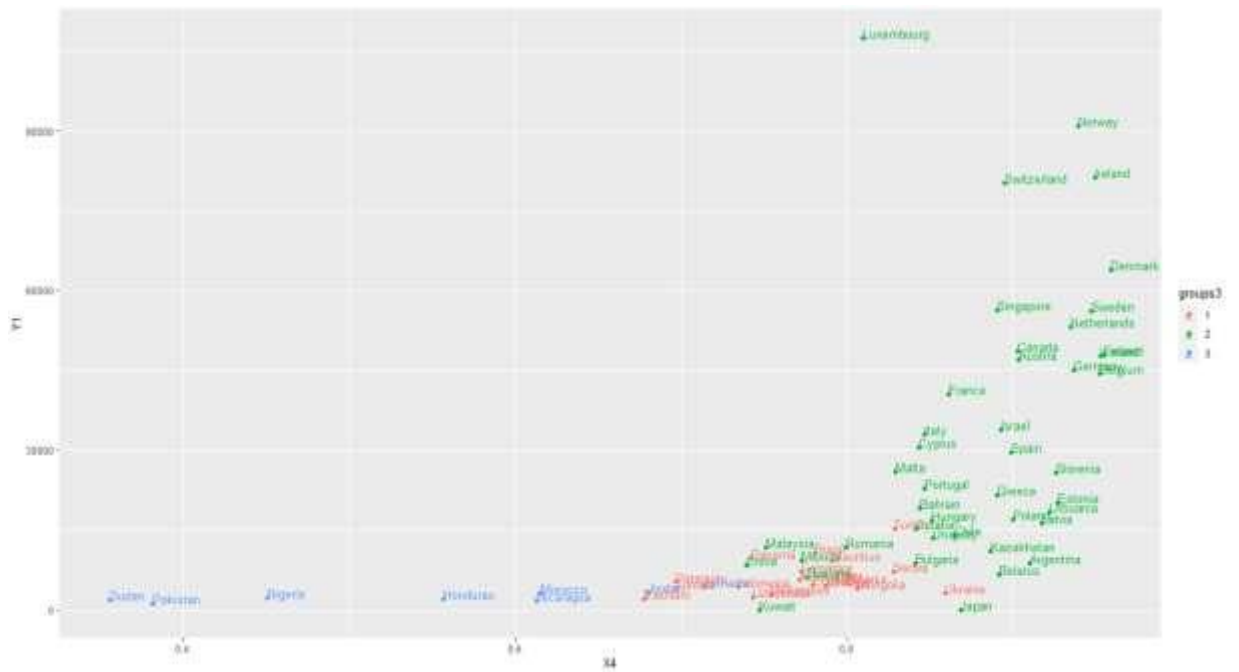


Рис. Б.4. Діаграма розсіювання для пари Y1 та X4

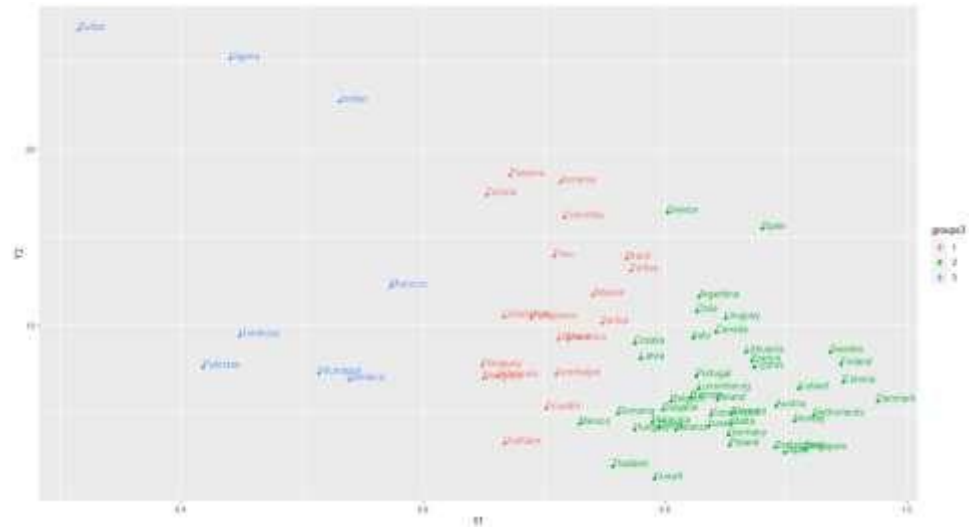


Рис. Б.7. Діаграма розсіювання для пари Y2 та X1

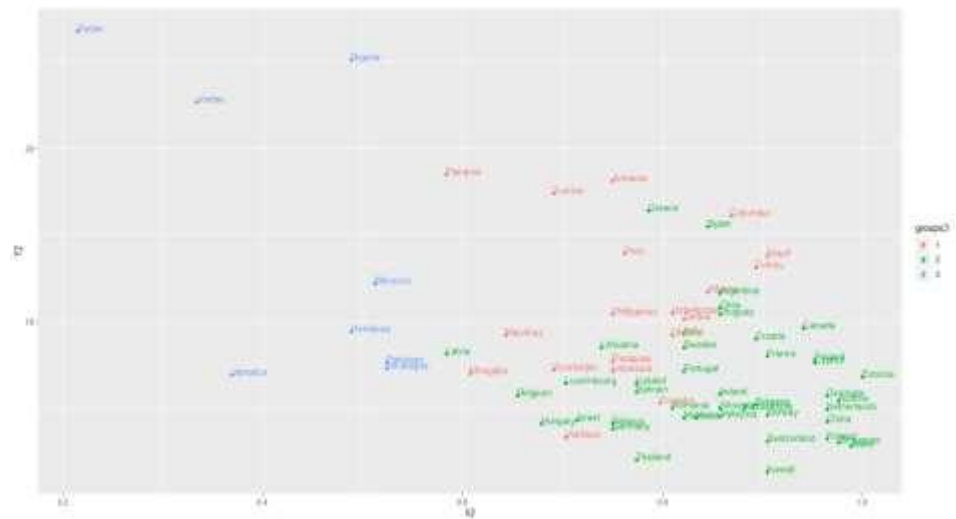


Рис. Б.8. Діаграма розсіювання для пари Y2 та X2

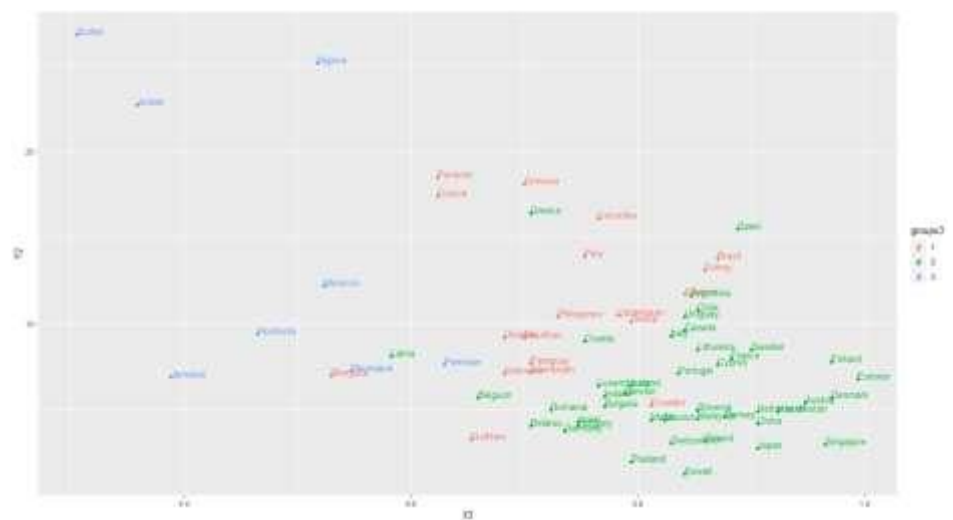


Рис. Б.9. Діаграма розсіювання для пари Y2 та X3

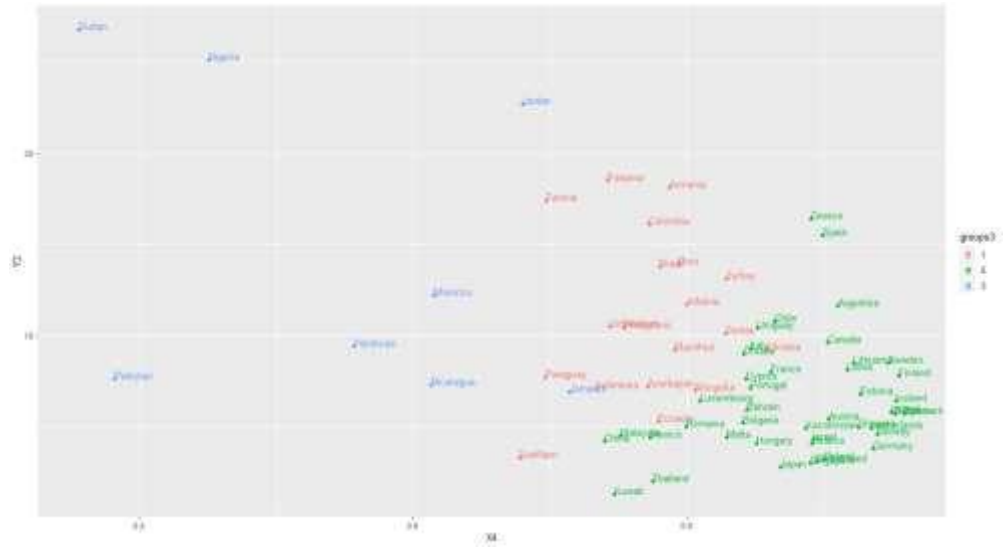


Рис. Б.10. Діаграма розсіювання для пари Y2 та X4

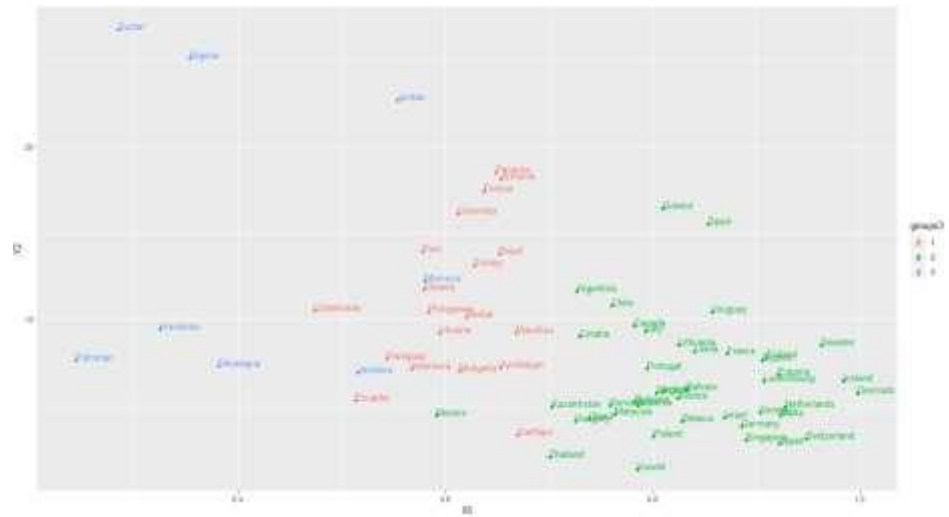


Рис. Б.11. Діаграма розсіювання для пари Y2 та X5

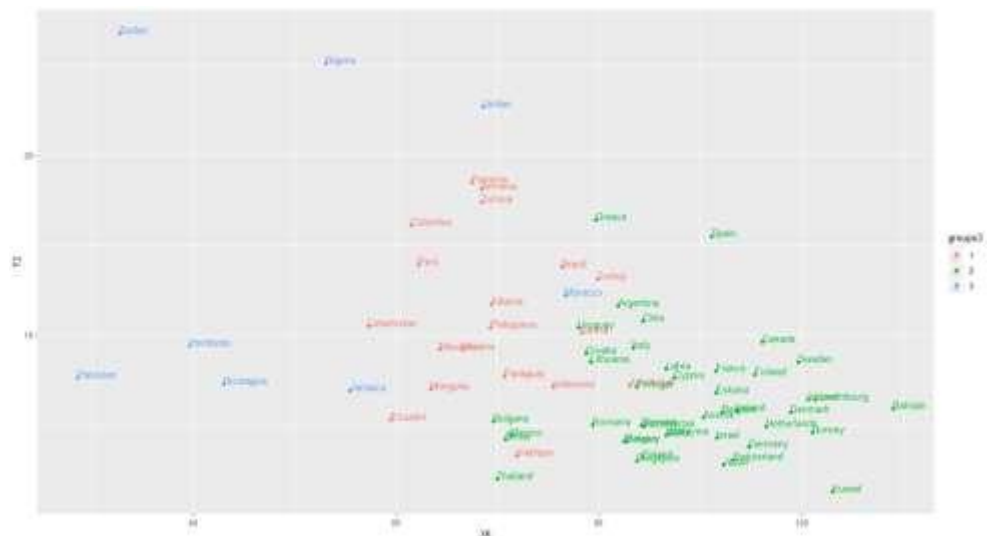


Рис. Б.12. Діаграма розсіювання для пари Y2 та X6

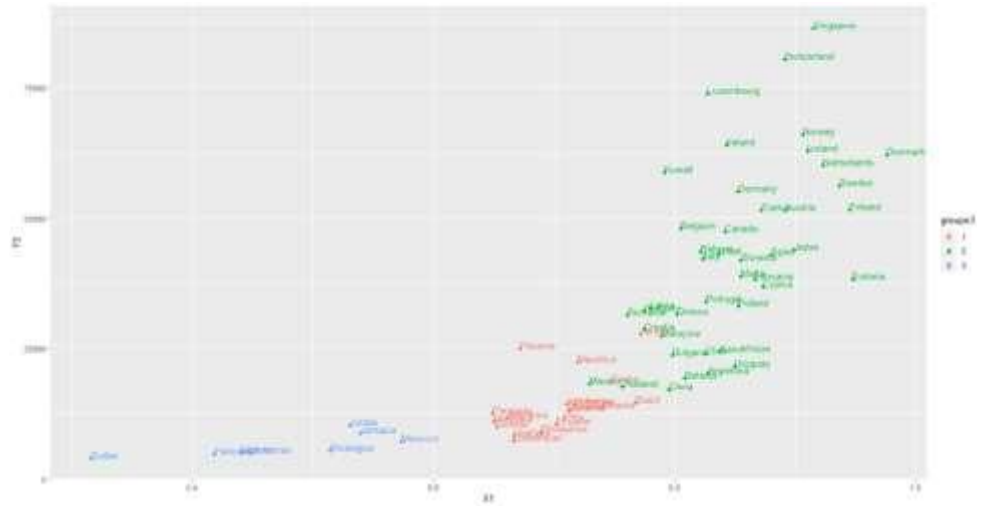


Рис. Б.13. Діаграма розсіювання для пари Y3 та X1

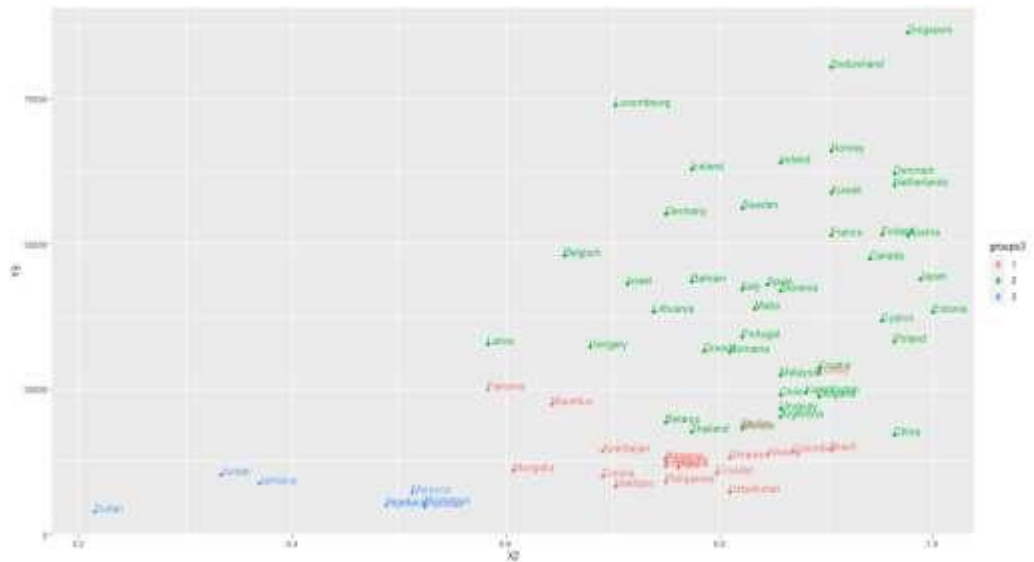


Рис. Б.14. Діаграма розсіювання для пари Y3 та X2

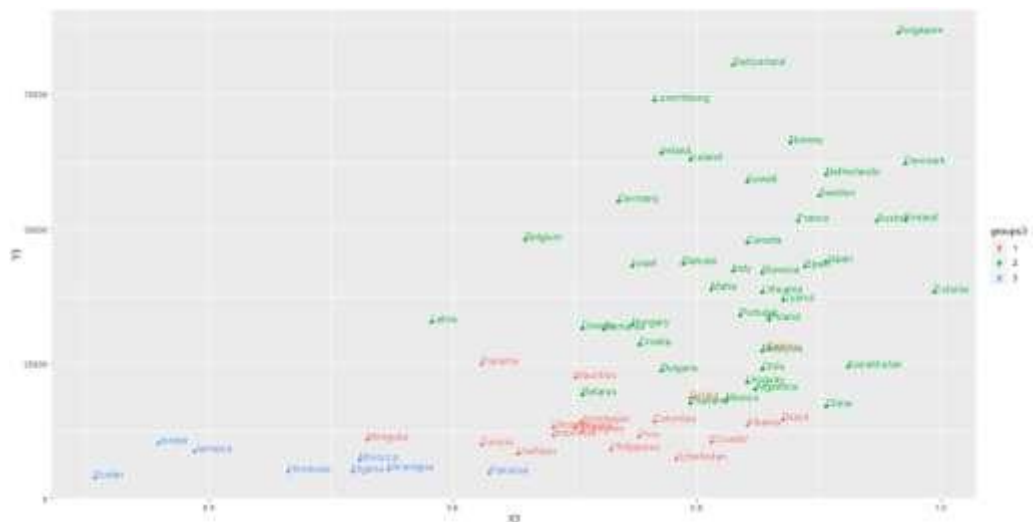


Рис. Б.15. Діаграма розсіювання для пари Y3 та X3

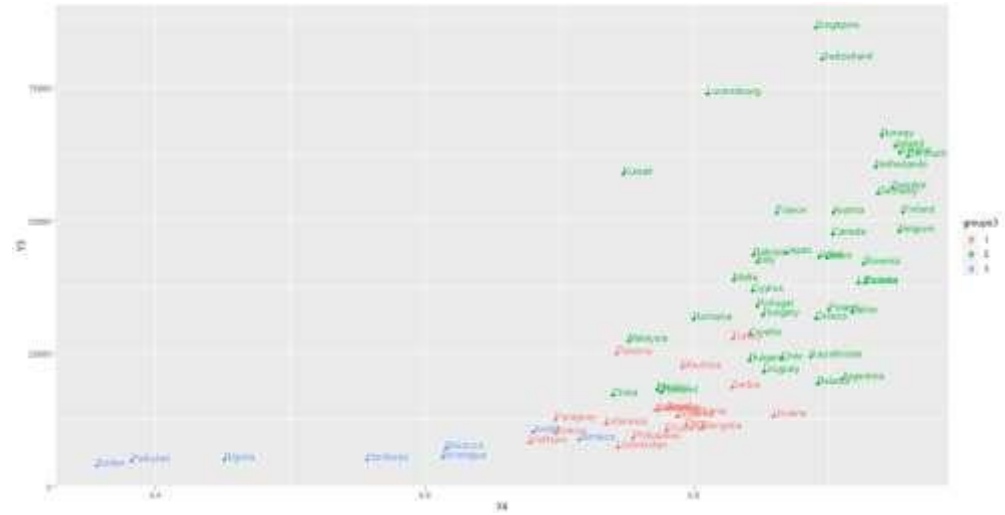


Рис. Б.16. Діаграма розсіювання для пари Y3 та X4

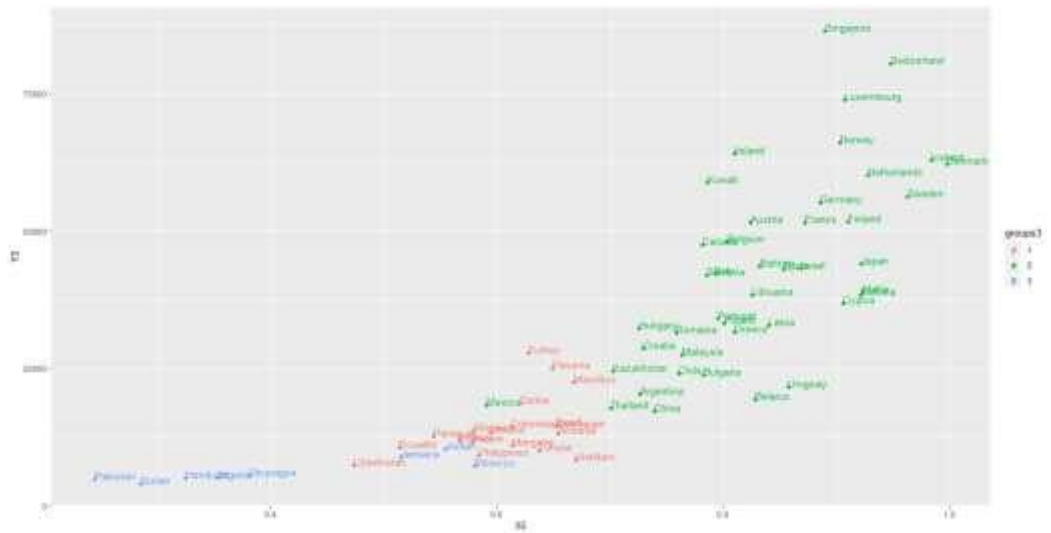


Рис. Б.17. Діаграма розсіювання для пари Y3 та X5

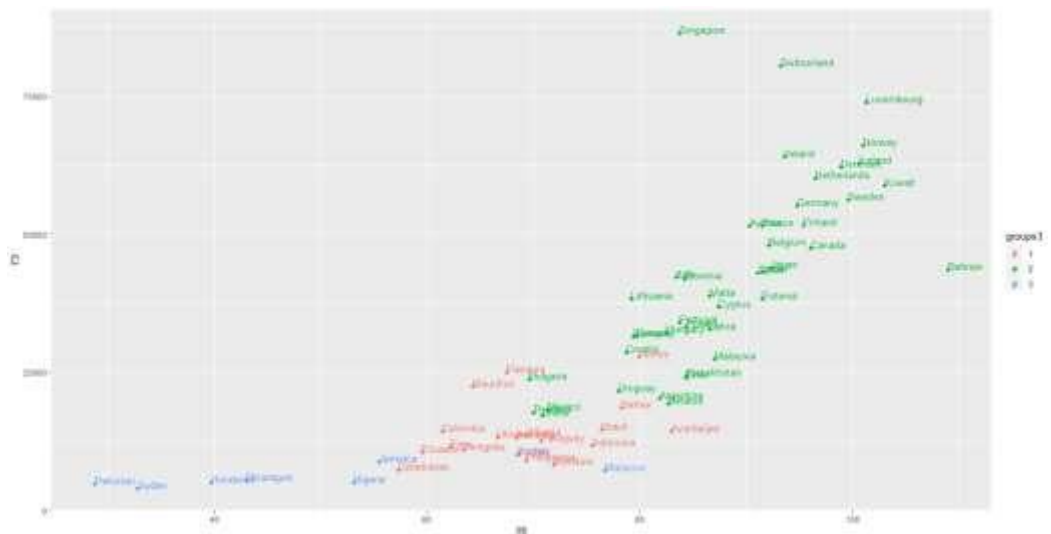


Рис. Б.18. Діаграма розсіювання для пари Y3 та X6

ДОДАТОК В

Код RStudio для побудови моделей та аналізу.

Дані для аналізу представлені в додатку А.

```
#Побудова моделей панельних даних та їх аналіз
h <- pdata.frame(Panel,
  index = c("Country", "Year"),
  row.names = TRUE)
m.re <- plm(Y1 ~ X1 + X2 + X3 + X4 +X5 + X6, data = h, model = "random")
summary(m.re)
m.fe <- plm(Y1 ~ X1 + X2 + X3 + X4 +X5 + X6, data = h, model = "within")
summary(m.fe)
m.pooled <- plm(Y1 ~ X1 + X2 + X3 + X4 +X5 + X6, data = h, model =
"pooling")
summary(m.pooled)
pFtest(m.fe, m.pooled)
phtest(m.fe, m.re)
plmtest(m.re, type = "bp")
m.re <- plm(Y2 ~ X1 + X2 + X3 + X4 +X5 + X6, data = h, model = "random")
summary(m.re)
m.fe <- plm(Y2 ~ X1 + X2 + X3 + X4 +X5 + X6, data = h, model = "within")
summary(m.fe)
m.pooled <- plm(Y2 ~ X1 + X2 + X3 + X4 +X5 + X6, data = h, model =
"pooling")
summary(m.pooled)
pFtest(m.fe, m.pooled)
phtest(m.fe, m.re)
plmtest(m.re, type = "bp")
m.re <- plm(Y3 ~ X1 + X2 + X3 + X4 +X5 + X6, data = h, model = "random")
```

```

summary(m.re)
m.fe <- plm(Y3 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model = "within")
summary(m.fe)
m.pooled <- plm(Y3 ~ X1 + X2 + X3 + X4 + X5 + X6, data = h, model =
"pooling")
summary(m.pooled)
pFtest(m.fe, m.pooled)
phtest(m.fe, m.re)
plmtest(m.re, type = "bp")

#Кластерний аналіз
library(dplyr)
d <- Panellll %>% select(Y1:X6)
rownames(d) <- Panellll$Country

library(cluster)
df.stand <- as.data.frame(scale(d))
set.seed(5)
c(kmeans(df.stand, centers = 5, nstart = 1)$tot.withinss,
  kmeans(df.stand, centers = 5, nstart = 25)$tot.withinss)

k.max <- 15 # максимальна кількість кластерів
wss <- sapply(1:k.max, function(k){
  kmeans(df.stand, k, nstart = 10)$tot.withinss
})
library(factoextra)
fviz_nbclust(df.stand, kmeans, method = "wss") +
  geom_vline(xintercept = 3, linetype = 2)

set.seed(123)

```

```
gap_stat <- clusGap(d, FUN = kmeans, nstart = 10, K.max = 10, B = 70)
print(gap_stat, method = "firstmax")
fviz_gap_stat(gap_stat)
```

```
set.seed(123)
gap_stat <- clusGap(d, FUN = pam, K.max = 7, B = 100)
print(gap_stat, method = "firstmax")
fviz_cluster(pam(df.stand, 5), stand = FALSE)
```

```
install.packages("NbClust")
library(NbClust)
res <- NbClust(df.stand, min.nc = 2, max.nc = 8, method = "kmeans")
res$Best.nc
library(factoextra)
fviz_nbclust(res)
```

```
M <- dist(scale(d))^2
hc <- hclust(M, method = "ward.D")
plot(hc, cex = 0.8)
plot(hc, cex = 0.8, main = "3 clusters")
rect.hclust(hc, k = 3, border="red")
plot(hc, cex = 0.8, main = "5 clusters")
rect.hclust(hc, k = 5, border="red")
```

```
install.packages("pvclust")
library(pvclust)
fit <- pvclust(t(d), method.hclust = "ward", method.dist = "euclidean")
plot(fit, cex = 0.9)
plot(fit, cex = 0.9)
pvrect(fit, alpha = 0.90)
```

```

groups3 <- cutree(hc, k = 3)
groups3
d <- d %>% mutate(groups3 = factor(groups3), Country = Panellll$Country)
View(d)
library(ggplot2)
ggplot(data = d, aes(x = X1, y = Y1, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X2, y = Y1, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X3, y = Y1, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X4, y = Y1, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X5, y = Y1, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X6, y = Y1, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X1, y = Y2, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X2, y = Y2, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X3, y = Y2, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X4, y = Y2, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X5, y = Y2, color = groups3)) + geom_point() +
  geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X6, y = Y2, color = groups3)) + geom_point() +

```

```

geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X1, y = Y3, color = groups3)) + geom_point() +
geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X2, y = Y3, color = groups3)) + geom_point() +
geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X3, y = Y3, color = groups3)) + geom_point() +
geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X4, y = Y3, color = groups3)) + geom_point() +
geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X5, y = Y3, color = groups3)) + geom_point() +
geom_text(aes(label = Country, vjust = 0, hjust = 0))
ggplot(data = d, aes(x = X6, y = Y3, color = groups3)) + geom_point() +
geom_text(aes(label = Country, vjust = 0, hjust = 0))
aggregate(df.stand,by=list(d$groups3),FUN=mean)
g <- data.frame(df.stand, d$groups3)
clusplot(g, d$groups3, color=TRUE, shade=TRUE, labels=2, lines=0)

#Побудова багатofакторної моделі регресії
states1 <- as.data.frame(model[,c("Y1", "Y2", "Y3", "X1", "X2", "X3", "X4",
"X5", "X6")])
hist(states1$Y1)
hist(states1$Y2)
hist(states1$Y3)
pairs(states1)
round(cor(states1), 2)
psych::pairs.panels(states1)
round(cor(states1, method="spearman"), 2)
train.ids <- 1:100
test.ids <- 101:200

```

```
my1_1 <- lm(Y1 ~ X1 + X2 + X3 + X4 + X5 + X6, states1[train.ids,])
```

```
summary(my1_1)
```

```
my1_2 <- lm(Y1 ~ X1 + X2 + X3 + X5 + X6, states1[train.ids,])
```

```
summary(my1_2)
```

```
my1_3 <- lm(Y1 ~ X1 + X2 + X3 + X4 + X5 + X6 + 0, states1[train.ids,])
```

```
summary(my1_3)
```

```
my1_4 <- lm(Y1 ~ X2 + X4 + X5 + 0, states1[train.ids,])
```

```
summary(my1_4)
```

```
anova(my1_2,my1_4)
```

```
summary(aov(Y1 ~ X2 + X4 + X5 + 0, data = states1))
```