

## **Міжнародний благодійний фонд «Альянс громадського здоров'я»**

Сегментація респондентів операційного дослідження проєкту Drugstore 2025 року:  
результати кластерного аналізу

Поведінкові профілі, ризики та потреби людей, які вживають психоактивні речовини

**Київ, Україна**  
**Травень 2026 року**

### **Автори:**

**Вячеслав Кушаков**, директор департаменту поведінкового здоров'я та кризового реагування Альянсу громадського здоров'я.  
kushakov@aph.org.ua

**Галина Сергієнко**, керівниця відділу поведінкового здоров'я та телездоров'я Альянсу громадського здоров'я.  
sergienko@aph.org.ua

**Євген Большов**, старший дослідник-аналітик відділу поведінкового здоров'я та телездоров'я Альянсу громадського здоров'я.  
bolshov@aph.org.ua

**Контактна особа:** Євген Большов, [bolshov@aph.org.ua](mailto:bolshov@aph.org.ua)

© 2026 Міжнародний благодійний фонд «Альянс громадського здоров'я». Всі права захищені. Жодна частина цієї публікації не може бути відтворена без попереднього дозволу власника авторських прав.

Погляди, висловлювання та думки, що містяться у цьому звіті, належать авторам і не обов'язково відображають офіційну позицію МБФ «Альянс громадського здоров'я», його партнерів або донорів. Автори доклали всіх зусиль для забезпечення точності інформації, однак організація не несе відповідальності за можливі помилки та неточності.

Глосарій та список скорочень.....	5
Список таблиць .....	8
Список діаграм .....	9
1. Короткий виклад результатів.....	10
1.1. Загальна логіка і цілі сегментації.....	10
1.2. Основні результати .....	10
1.3. Найбільш пріоритетні сегменти для програмної роботи.....	11
1.4. Практичне значення сегментації для проєкту Drugstore.....	12
2. Методологія дослідження .....	13
2.1. Джерело даних.....	13
2.2. Дизайн дослідження .....	13
2.3. Цільова аудиторія дослідження .....	13
2.4. Особливості вибірки .....	13
2.5. Інтерпретація результатів у межах операційного дослідження.....	14
2.6. Статистичний аналіз та програмне забезпечення .....	14
3. Дані, змінні та підготовка до кластерного аналізу .....	16
3.1. Аналітична база кластеризації.....	16
3.2. Змінні, використані для кластеризації .....	16
3.2.1. Демографічні змінні.....	16
3.2.2. Змінні щодо ПАР, які вживаються неін'єкційно .....	16
3.2.3. Змінні щодо ін'єкційного вживання ПАР.....	17
3.2.4. Змінні щодо вживання алкоголю та поєднання алкоголю з ПАР .....	17
3.2.5. Змінні щодо сексуальної поведінки .....	17
3.3. Перекодування змінних та обробка пропущених значень .....	18
3.4. Робота із змінними різного типу та стандартизація змінних.....	19
4. Метод кластеризації.....	21
4.1. Методи кластерного аналізу, що використовувалися.....	21
4.2. Двоетапний кластерний аналіз .....	21
4.3. Ієрархічний кластерний аналіз.....	21
4.4. РАМ-кластеризація .....	22
5. Порівняння підходів до кластеризації та вибір робочого рішення.....	23
5.1. Вибір робочої моделі сегментації .....	23
5.2. Порівняння типологій, отриманих методом ієрархічного кластерного аналізу та двоетапним кластерним аналізом .....	23
6. Розподіл респондентів за кластерами .....	24
7. Профілі кластерів.....	25
7.1. Кластер 1 — жіночий відносно низько ризиковий сегмент із запитом на психологічну підтримку.....	25

7.2. Кластер 2 — жіночий сегмент підвищеного ризику, акцент на вживання стимуляторів та емпатогенів .....	26
7.3. Кластер 3 — чоловічий переважно гетеросексуальний сегмент помірного ризику...26	26
7.4. Кластер 4 — чоловічий ін'єкційно-поліспоживчий сегмент .....	27
7.5. Кластер 5 — ядро полівживання, групового сексу та соціальної вразливості.....	28
7.6. Узагальнення .....	28
8. Порівняльний профіль кластерів.....	29
8.1. Порівняння за демографічним профілем .....	29
8.2. Порівняння за профілем вживання ПАР.....	31
8.2.1. Ін'єкційне вживання ПАР.....	31
8.2.2. Види речовин, що вживаються неін'єкційно.....	32
8.2.3. Полівживання ПАР .....	34
8.2.4. Вживання алкоголю, одночасне вживання ПАР та алкоголю .....	35
8.3. Порівняння за сексуальною поведінкою.....	38
8.3.1. Груповий секс .....	38
8.3.2. Вживання ПАР під час останнього сексу.....	40
8.3.3. Гомосексуальні контакти.....	41
8.3.4. Транзакційний секс.....	42
8.4. Тестування на ВІЛ та ВІЛ-статус, тестування на інші ІПСШ.....	44
8.5. Потреби в психологічній, психіатричній, медичній допомозі та звернення за послугами.....	46
8.5.1. Психологічна та психіатрична допомога .....	46
8.5.2. Медична допомога .....	46
8.6. Вживання ПАР соціальним оточенням респондента, спайкінг, небажані прояви і проблеми, пов'язані із вживання ПАР .....	48
8.6.1. Соціальне оточення .....	48
8.6.2. Спайкінг .....	49
8.6.3. Небажані прояви та проблеми, пов'язані із вживанням ПАР .....	49
9. Профілювальний алгоритм для попереднього віднесення клієнта до сегмента .....	51
9.1. Метод побудови алгоритму.....	51
9.2. Вхідні змінні алгоритму.....	51
9.3. Принцип роботи алгоритму .....	51
9.4. Результативність алгоритму .....	52
9.5. Методичні застереження та перевірка альтернативних статистичних алгоритмів профілювання.....	53
9.6. Підсумок.....	54
9.7. Етичні застереження щодо використання алгоритму .....	55
10. Обмеження сегментації .....	56

Додатки.....	58
Додаток 1. Повний перелік змінних, використаних у кластерному аналізі .....	58
Додаток 2. Схема перекодування змінних і створення похідних змінних .....	62
Додаток 3. Методичні матеріали щодо вибору кількості кластерів.....	65
Додаток 4. Матриця класифікації профілюючого алгоритму.....	66
Додаток 5. Технічний опис профілюючого алгоритму (дискримінантного аналізу) .....	68
Додаток 6. Ілюстративні таблиці розподілу ключових змінних за кластерами .....	71
Додаток 7. Діагностика пропущених значень у змінних кластерного аналізу.....	73

# Глосарій та список скорочень

## I. Ключові групи та поведінкові практики

**ЧСЧ** – Чоловіки, які мають секс із чоловіками.

**Транзакційний секс** – Сексуальні контакти, у межах яких відбувається обмін сексу на гроші, матеріальні блага, психоактивні речовини або інші ресурси.

**Сексуалізоване вживання наркотиків / ПАР (SDU, Sexualised Drug Use)** – Умисне вживання психоактивних речовин перед або під час сексуальних контактів з метою зміни або інтенсифікації сексуального досвіду.

**Спайкінг** – Навмисне додавання психоактивних або сильнодіючих речовин у напій чи їжу людини без її відома та згоди.

## II. Профілактика, інфекції та послуги

**ВІЛ** – Вірус імунодефіциту людини.

**СНІД** – Синдром набутого імунодефіциту.

**ІПСШ** – Інфекції, що передаються статевим шляхом.

**Вірусні гепатити** – Група вірусних інфекцій, що уражають печінку; у контексті цього звіту передусім ідеться про гепатити В і С.

**ДКП / PrEP** – Доконтактна профілактика ВІЛ: прийом антиретровірусних препаратів для зниження ризику інфікування ВІЛ.

**Самотестування** – Самостійне проведення тесту, наприклад на ВІЛ, без безпосередньої участі медичного працівника.

**Зменшення шкоди** – Підхід до роботи з людьми, які вживають ПАР, спрямований на зниження ризиків для здоров'я та життя без обов'язкової вимоги негайної відмови від вживання.

**PartyBox** – Комплексний профілактичний набір, який отримують клієнти проекту Drugstore, що містить матеріали для зменшення шкоди, швидкі діагностичні тести на інфекції та реагенти для перевірки ПАР, та інформаційні ресурси.

## III. Алкоголь та пов'язані індикатори

**Стандартна порція алкоголю (standard drink)** – Умовна одиниця споживання алкоголю, що відповідає фіксованій кількості чистого етанолу; у цьому звіті визначається через конкретні об'єми напоїв із типовою міцністю.

**Епізодичне надмірне вживання (binge drinking)** – Епізод вживання великої кількості алкоголю за один раз; поріг визначається відповідно до використаної методики.

**HED (heavy episodic drinking)** – Епізодичне надмірне вживання алкоголю у визначенні ВООЗ: споживання щонайменше 60 г чистого алкоголю принаймні один раз протягом останніх 30 днів.

**Інтенсивне вживання алкоголю (heavy drinking)** – Часте епізодичне надмірне вживання алкоголю, визначення якого залежить від використаної методики (SAMHSA, CDC тощо).

#### **IV. Психоактивні речовини**

**ПАР** – Психоактивні речовини.

**2С-В** – Синтетична психоактивна речовина з групи фенетиламінів із психоделічними та стимулювальними ефектами.

**Альфа-PVP ( $\alpha$ -PVP)** – Синтетичний психостимулятор із групи катінонів із високим потенціалом токсичності та залежності.

**GHB/GBL** –  $\gamma$ -гідроксимасляна кислота та  $\gamma$ -бутиролактон; речовини з депресивною дією на центральну нервову систему.

**LSD** (лізергінова кислота діетиламід) – Напівсинтетичний психоделічний галюциноген.

**MDMA** (екстазі) – 3,4-метилендіоксиметамфетамін; синтетична психоактивна речовина з емпатогенними та стимулювальними властивостями.

**Прегабалін** (торгова назва «Лірика») – Габапентиноїд, який при немедичному використанні може викликати ейфорію та залежність.

**Синтетичні катіони / «солі»** – Група синтетичних стимулювальних речовин, до якої можуть належати мефедрон, альфа-PVP та інші подібні речовини.

#### **V. Статистичні та методичні терміни**

**Сегментація** – Поділ респондентів на групи зі схожими поведінковими профілями, ризиками та потребами.

**Кластер** – Група респондентів, які за набором використаних ознак є більш подібними між собою, ніж до респондентів з інших груп.

**Кластерний аналіз** – Статистичний підхід, який використовується для пошуку груп схожих випадків у масиві даних без наперед заданої належності до груп.

**Двоетапний кластерний аналіз (TwoStep Cluster Analysis)** – Метод кластеризації, придатний для великих масивів даних і змінних різних типів; у цьому звіті використовувався як один із підходів для перевірки сегментації.

**Ієрархічний кластерний аналіз** – Метод кластеризації, у якому респонденти поступово об'єднуються в більші групи на основі подібності між ними.

**Метод Варда** – Ієрархічний метод кластеризації, який на кожному кроці об'єднує ті кластери, злиття яких мінімально збільшує внутрішньокластерну дисперсію.

**Псевдометричні змінні** – Порядкові або бінарні змінні, які в окремих аналітичних процедурах технічно розглядаються як числові шкали.

**Відстань Говера** – Міра несхожості, що дозволяє одночасно працювати з числовими, порядковими та номінальними змінними.

**PAM (Partitioning Around Medoids)** – Метод кластеризації, який формує групи навколо медоїдів — реальних спостережень, найбільш центральних (типових) для своїх кластерів; може використовувати різні матриці відстаней або несхожості.

**Silhouette / коефіцієнт силуету** – Показник, який оцінює, наскільки добре випадки відповідають своєму кластеру порівняно з іншими кластерами.

**BIC (Bayesian Information Criterion)** – Статистичний критерій, який може використовуватися для порівняння моделей і вибору кількості кластерів.

**Дискримінантний аналіз** – Статистичний метод класифікації, який будує правила розрізнення наперед визначених груп на основі набору ознак; у цьому звіті використаний для попереднього віднесення нових респондентів до одного з кластерів.

**Негативно-біноміальна GLM** – Узагальнена лінійна модель, яка використовується для аналізу лічильних показників, наприклад кількості подій або кількості типів проблем.

## **VI. Інформаційні, аналітичні та облікові системи**

**KoboToolbox** – Онлайн-платформа для створення та адміністрування опитувань і збору даних.

**SPSS (Statistical Package for the Social Sciences)** – Програмне забезпечення для статистичного аналізу даних.

**R** – Мова програмування та програмне середовище для статистичного аналізу даних, їх обробки, візуалізації та моделювання.

**Google Sheets** – Онлайн-сервіс для роботи з електронними таблицями; у межах цього проекту використовується для технічної реалізації профілювального алгоритму.

## Список таблиць

Таблиця 1. Основні показники профілювального алгоритму на основі дискримінантного аналізу .....	52
Таблиця 2. Якість відтворення окремих кластерів у дискримінантному аналізі .....	53
Таблиця 3. Порівняння протестованих профілювальних підходів .....	54
Таблиця 4. Змінні, включені до кластерного аналізу.....	58
Таблиця 5. Перекодування змінних для кластерного аналізу .....	62
Таблиця 6. Похідні індикатори, створені для подальшого опису кластерів .....	63
Таблиця 7. Порівняння протестованих підходів до вибору кількості кластерів .....	65
Таблиця 8. Матриця класифікації профілювального алгоритму, крос-валідація (кількість випадків) .....	66
Таблиця 9. Частка правильно класифікованих випадків за кластерами, % .....	67
Таблиця 10. Основні технічні параметри дискримінантного аналізу .....	68
Таблиця 11. Канонічні дискримінантні функції.....	68
Таблиця 12. Перевірка статистичної значущості дискримінантних функцій .....	69
Таблиця 13. Якість класифікації .....	69
Таблиця 14. Частка правильно класифікованих випадків за кластерами, % .....	69
Таблиця 15. Перевірка передумови рівності коваріаційних матриць .....	70
Таблиця 16. Змінні з найбільш вираженим внеском у дискримінацію кластерів .....	70
Таблиця 17. Демографічний, ін'єкційний та алкогольний профіль кластерів .....	71
Таблиця 18. Сексуальна поведінка та сексуальний ризик у кластерах.....	72
Таблиця 19. Частка пропущених або невизначених відповідей у ключових блоках кластероутворюючих змінних .....	73
Таблиця 20. Розподіл респондентів за кількістю пропущених або невизначених значень серед кластероутворюючих змінних .....	73
Таблиця 21. Середня кількість пропущених або невизначених значень за кластерами.....	74

## Список діаграм

Діаграма 1. Розподіл респондентів за кластерами та інтегральна оцінка проблемності. ...	24
Діаграма 2. Статевий склад кластерів. ....	29
Діаграма 3. Середній вік за кластерами. ....	30
Діаграма 4. Матеріальне становище респондентів за кластерами. ....	31
Діаграма 5. Досвід ін'єкційного вживання ПАР за кластерами. ....	32
Діаграма 6. Вживання ПАР протягом 30 днів до опитування за кластерами. ....	33
Діаграма 7. Вживання ПАР протягом життя за кластерами. ....	34
Діаграма 8. Показники полівживання ПАР за кластерами. ....	35
Діаграма 9. Практики одночасного вживання алкоголю та ПАР за кластерами. ....	36
Діаграма 10. Вплив вживання ПАР на кількість алкоголю, що споживається. ....	37
Діаграма 11. Частка респондентів, які підпадають під поріг епізодичного надмірного вживання алкоголю за кластерами. ....	38
Діаграма 12. Показники досвіду групового сексу за кластерами. ....	39
Діаграма 13. Кількість епізодів групового сексу за 6 місяців за кластерами. ....	40
Діаграма 14. Практика сексуалізованого вживання ПАР за кластерами. ....	40
Діаграма 15. Досвід гомосексуальних стосунків серед жінок за кластерами. ....	41
Діаграма 16. Досвід гомосексуальних стосунків серед чоловіків за кластерами. ....	42
Діаграма 17. Досвід транзакційного сексу протягом останніх 12 місяців серед жінок за кластерами. ....	43
Діаграма 18. Досвід транзакційного сексу протягом останніх 12 місяців серед чоловіків за кластерами. ....	43
Діаграма 19. Результати тестування на ВІЛ за кластерами. ....	45
Діаграма 20. Вживання ПАР друзями респондента, за кластерами. ....	48
Діаграма 21. Кількість типів проблем, викликаних вживанням ПАР за кластерами. ....	50

## 1. Короткий виклад результатів

### 1.1. Загальна логіка і цілі сегментації

Сегментацію респондентів було проведено для того, щоб краще зрозуміти неоднорідність аудиторії проекту Drugstore<sup>1</sup>. Люди, які вживають психоактивні речовини, можуть суттєво відрізнятися за частотою та типами вживання ПАР, ін'єкційними практиками, поєднанням ПАР з алкоголем, сексуальною поведінкою, досвідом тестування на ІПСШ, потребами у психологічній, медичній та профілактичній допомозі тощо.

Загальний опис вибірки показує середню картину, але для роботи з клієнтами цього замало. Сегментація допомагає побачити, які групи мають подібні ризики й потреби та яку підтримку їм варто пропонувати насамперед. Відповідно, сегментація дає змогу побудувати систему визначення пріоритетів в діяльності проекту за умови обмежених ресурсів і витратити ці ресурси (час працівників, профілактичні матеріали, кошти) максимально ефективно.

Важливо, що ця сегментація не має на меті оцінити поширеність таких груп серед усіх людей, які вживають ПАР в Україні. Її завдання — описати респондентів дослідження та потенційних клієнтів проекту Drugstore у спосіб, який є корисним для сервісної роботи, маршрутизації клієнтів і формування подальших програмних рішень.

У результаті аналізу було виділено п'ять робочих сегментів. Вони не є жорсткими «типами особистості» і не означають, що всі люди всередині сегмента однакові. Це радше зручні профілі, які показують, де концентруються різні ризики та потреби.

### 1.2. Основні результати

Перший сегмент — це переважно молоді жінки з відносно нижчим рівнем поведінкових ризиків. У цій групі майже не має ін'єкційного вживання, більшість речовин вживаються рідко, рідше трапляється груповий секс і рідше поєднуються ПАР з алкоголем. Але цей сегмент не можна вважати «безпроблемним». Для нього особливо помітний запит на психологічну підтримку. Очевидно, низький рівень наркотичного або сексуального ризику не означає відсутності потреб у допомозі.

Другий сегмент також переважно жіночий, але вже з помітно активнішим профілем вживання ПАР. Для нього характерне вживання таких класів речовин як стимулятори та емпатоцени, частіше поєднання ПАР з алкоголем і вища залученість до сексуалізованого вживання. Це не найбільш проблемний сегмент, але він важливий для раннього втручання: саме тут профілактика може запобігти переходу до більш ризикових практик.

Третій сегмент — переважно чоловічий і гетеросексуальний. Він виглядає більш благополучним за більшістю показників ризику і водночас більш матеріально забезпеченим. Проте, це не означає, що його можна не враховувати. У цьому сегменті є вживання різних ПАР, є частина респондентів з ін'єкційним досвідом, а

---

<sup>1</sup> <https://drugstore.org.ua/uk/>

також помітні прогалини в регулярному тестуванні на ІПСШ. Для цієї групи важливі прості, ненав'язливі профілактичні пропозиції та зручна медична навігація.

Четвертий сегмент — один із ключових високоризикових кластерів. Це переважно чоловіки з виразним ін'єкційним компонентом, часто із полівживанням, частішим поєднанням ПАР з алкоголем. Тут ризики пов'язані не з одним фактором, а з їхнім накладанням: ін'єкційне вживання, кілька речовин, алкоголь, сексуальні ризики, потреба в медичній допомозі. Для Drugstore це один із головних сегментів для активної сервісної роботи.

П'ятий сегмент — найменший, але найбільш концентрований за ризиками. У ньому сходяться майже всі ключові проблемні лінії: полівживання, ін'єкційний досвід, груповий секс, сексуалізоване вживання ПАР, транзакційний секс, спайкінг і матеріальна вразливість. Саме цей сегмент потребує найбільш комплексної відповіді. Тут недостатньо інформаційних повідомлень або стандартного профілактичного набору, потрібна зв'язка послуг таких як зменшення шкоди, тестування на ІПСШ, сексуальне здоров'я, психологічна підтримка, соціальна та медична навігація.

Очевидно, ризики в аудиторії Drugstore розподілені дуже нерівномірно. В одних сегментах вони нижчі й більше вимагають профілактики, інформування і психологічної підтримки. В інших — ризики більш комплексні і потребують більш активного супроводу від профілактичних програм. Найбільш рельєфними є відмінності за ін'єкційним вживанням, полівживанням ПАР, поєднанням ПАР з алкоголем і сексуальним контекстом вживання.

### 1.3. Найбільш пріоритетні сегменти для програмної роботи

Найвищий пріоритет для програмної роботи мають кластери 5 і 4.

Кластер 5 є найбільш вразливим. Для цього сегмента потрібна не окрема послуга, а комплексний маршрут: тестування на ВІЛ, гепатити та інші ІПСШ, консультування щодо безпечнішого сексу, матеріали для зменшення шкоди, підтримка після ризикових ситуацій, психологічна допомога та перенаправлення до дружніх медичних і соціальних сервісів.

Кластер 4 є другим ключовим пріоритетом. Для нього особливо важливі послуги зменшення шкоди, доступ до тестування, медична навігація, консультації щодо поєднання речовин та алкоголю, а також робота з ризиками, пов'язаними із сексуальною поведінкою.

Кластер 2 можна розглядати як сегмент раннього втручання. Для цього сегмента важлива профілактика для сексуального здоров'я, безпечного дозвілля і регулярне тестування.

Кластери 1 і 3 не є головними високо ризиковими сегментами, але їх не варто залишати поза увагою. Для кластера 1 важлива психологічна підтримка як окрема точка входу до сервісів. Для кластера 3 — зручні пропозиції тестування, профілактичні повідомлення без надмірного тиску та маршрутизація до медичної допомоги. У цих сегментах корисніше працювати через легкий доступ, довіру та ненав'язливе інформування.

#### 1.4. Практичне значення сегментації для проекту Drugstore

Сегментація дає Drugstore практичну основу для більш адресної роботи з клієнтами в онлайн-середовищі. Оскільки взаємодія з потенційним клієнтом часто починається через цифрову анкету, важливо не лише зібрати відповіді, а й одразу перетворити їх на корисну рекомендацію та контекст для сервісної роботи. Профілювальний алгоритм дозволяє автоматично оцінити, до якого сегмента найближчий новий клієнт, і на цій основі запропонувати менеджеру або консультанту перший набір пріоритетних тем: тестування, сексуальне здоров'я, ін'єкційні ризики, психологічна підтримка, зменшення шкоди або повторний контакт.

Практична цінність сегментації полягає в тому, що для різних кластерів, які потребують різних сценаріїв взаємодії, можна одразу обирати найбільш доцільний набір послуг. Це також може допомогти краще налаштувати повідомлення і контент проєкт. Замість однакової комунікації для всіх Drugstore може формувати різні цифрові підказки: для одних — короткі матеріали про безпечніше вживання і самотестування; для інших — нагадування про регулярне тестування, PrEP/ДКП, презервативи й лубриканти; для третіх — інформацію про ін'єкційні ризики, реакенти, небажані реакції, змішування ПАР з алкоголем і дружні сервіси. Такий підхід робить онлайн-взаємодію менш загальною і більш наближеною до реальних ризиків конкретного клієнта.

Профілювальний алгоритм не має ухвалювати рішення замість фахівця, але може суттєво зменшити ризик пропустити важливу тему під час першого контакту. Його роль — швидко структурувати інформацію з анкети і показати, який сервісний маршрут варто розглянути першим. Для проєкту, що значною мірою працює онлайн, це особливо важливо: сегментація стає не лише аналітичним результатом звіту, а частиною цифрової інфраструктури Drugstore — від анкети до персоналізованої пропозиції послуг і подальшого супроводу клієнта.

## 2. Методологія дослідження

### 2.1. Джерело даних

Сегментацію проведено на основі даних операційного дослідження проєкту Drugstore серед людей, які вживають психоактивні речовини. Польовий етап дослідження тривав з 12 лютого по 26 грудня 2025 року. До фінального очищеного масиву даних увійшло 2 592 анкети.

Детальна інформація про методологію, дизайн, вибірку, інструмент збору даних, чистку даних, етичні аспекти та обмеження дослідження наведена в основному звіті, з яким можна ознайомитися за посиланням «Операційне дослідження проєкту Drugstore серед людей, які вживають психоактивні речовини: результати 2025 року» <https://aph.org.ua/wp-content/uploads/2026/04/Operatsijne-doslidzhennya-proyektu-Drugstore.pdf>.

### 2.2. Дизайн дослідження

Дослідження мало дизайн безперервного крос-секційного онлайн-опитування. Збір даних здійснювався на платформі KoboToolbox<sup>2</sup> методом самозаповнення структурованої анкети. Протягом року в анкеті використовувалася ротація окремих тематичних модулів: частина блоків запитань запускала на обмежений період і вимикалася після досягнення необхідної кількості відповідей або отримання даних для програмних потреб. Тому обсяг підвбірок для певних питань в цьому звіті може відрізнятись.

Дослідження було інтегроване в сервісну модель проєкту Drugstore: після проходження скринінгу та заповнення анкети респондент отримував унікальні коди, проходив верифікацію співробітником проєкту та міг отримати профілактичний набір PartyBox або інші релевантні послуги.

### 2.3. Цільова аудиторія дослідження

Цільовою аудиторією були люди віком 14 років і старше, які проживали в Україні, надали інформовану згоду на участь у дослідженні та мали досвід вживання психоактивних речовин, крім алкоголю та канабісу, протягом останніх 6 місяців.

Дослідження було орієнтоване насамперед на людей, які потенційно зацікавлені у взаємодії з проєктом Drugstore та отриманні послуг зменшення шкоди, тестування на ВІЛ, психологічного консультування тощо.

### 2.4. Особливості вибірки

Вибірка дослідження була неімовірнісною. Рекрутування здійснювалося через цифрові канали комунікації, соціальні мережі, онлайн-спільноти, ресурси

---

<sup>2</sup> <https://www.kobotoolbox.org/>

партнерських організацій, особисті рекомендації та мережеве поширення інформації про проект.

Такий підхід передбачає самовідбір учасників: до опитування долучалися ті, хто дізнався про дослідження, відповідав критеріям включення та був готовий взаємодіяти з проектом. Відповідно, вибірка не є репрезентативною для всієї популяції людей, які вживають ПАР в Україні.

## 2.5. Інтерпретація результатів у межах операційного дослідження

Результати сегментації слід інтерпретувати як опис поведінкових профілів, ризиків і потреб респондентів дослідження та потенційних клієнтів проекту Drugstore. Вони не є популяційною оцінкою поширеності окремих типів споживачів ПАР в Україні.

Аналітична цінність сегментації полягає в тому, що вона допомагає краще зрозуміти різні підгрупи залученої аудиторії, адаптувати сервісні пропозиції до їхніх потреб, уточнити логіку маршрутизації клієнтів і сформулювати практичні гіпотези для подальшої програмної роботи та досліджень.

Важливо зазначити, що під «ризиками» або «ризикованою поведінкою» у цьому дослідженні ми розуміємо не лише практики, які безпосередньо підвищують імовірність інфікування ВІЛ, а й ті, що можуть мати ширші негативні наслідки для людини. Йдеться, зокрема, про проблеми сексуального та ментального здоров'я, погіршення фізичного самопочуття, порушення соціальних зв'язків, погіршення матеріального становища тощо. Усі ці наслідки можуть бути пов'язані з практиками, які розглядаються в дослідженні, зокрема з певними видами сексуальної поведінки та вживанням психоактивних речовин.

## 2.6. Статистичний аналіз та програмне забезпечення

Аналіз даних у цьому звіті поєднував кластеризацію, описову статистику, порівняльний аналіз кластерів, побудову профілювального алгоритму для попереднього віднесення нових клієнтів до визначених сегментів.

Для побудови сегментації було протестовано три підходи до кластерного аналізу. Перший — двоетапний кластерний аналіз у SPSS<sup>3</sup>. Другий — ієрархічна кластеризація методом Варда у SPSS<sup>4</sup>. Третій — PAM-кластеризація на відстанях Говера в R<sup>5</sup>.

Для порівняння кластерів використовувалися частотні таблиці, перехресні таблиці та стандартні статистичні тести для категоріальних змінних та метричних змінних. Основним тестом для перевірки зв'язку між кластерною належністю та категоріальними показниками був критерій  $\chi^2$  Пірсона. Для оцінки сили зв'язку використовувався Cramer's V, а для уточнення конкретних міжкластерних відмінностей — порівняння колонкових пропорцій. Для метричних або лічильних

<sup>3</sup> <https://www.ibm.com/docs/en/spss-statistics/32.0.0?topic=features-twostep-cluster-analysis>

<sup>4</sup> <https://www.ibm.com/docs/en/spss-statistics/32.0.0?topic=analysis-hierarchical-cluster-method>

<sup>5</sup> <https://www.rdocumentation.org/packages/cluster/versions/2.1.8.2/topics/pam>

показників застосовувалися робастні тести (наприклад, Welch), непараметричні тести, зокрема Kruskal–Wallis і тест медіан. Для деяких лічильних показників, що мали вигляд кількості подій або кількості типів проблем, використовувалася узагальнена лінійна модель.

Після побудови кластерів було створено окремий профілювальний алгоритм для попереднього віднесення нового респондента або клієнта до найбільш імовірного сегмента. Для цього використовувався лінійний дискримінантний аналіз у SPSS. Додатково перевірялися альтернативні підходи — мультиноміальна логістична регресія та дерева рішень.

Основним програмним забезпеченням для статистичного аналізу був IBM SPSS Statistics. За допомогою цього програмного забезпечення будувалися нові аналітичні змінні, перехресні таблиці, проводилися  $\chi^2$ -тести, непараметричні тести, будувалися узагальненні лінійні моделі, проводився двоетапний кластерний аналіз, ієрархічна кластеризація методом Варда, дискримінантний аналіз і відбувалося тестування альтернативних профілюючих моделей. R використовувався як додаткове середовище для методичної перевірки кластеризації на змішаних даних, насамперед через розрахунок відстаней Говера та PAM-кластеризацію.

В R використовувалися такі пакети: `cluster`<sup>6</sup> — для функцій `daisy()`/відстань Говера,  `pam()`/Partitioning Around Medoids і силуетної діагностики.

---

<sup>6</sup> <https://www.rdocumentation.org/packages/cluster/versions/2.1.8.2>

### 3. Дані, змінні та підготовка до кластерного аналізу

#### 3.1. Аналітична база кластеризації

Сегментацію було виконано не на всьому фінальному масиві операційного дослідження 2025 року, а на аналітичній підвибірці, сформованій для кластерного аналізу. Після чистки до загального масиву дослідження увійшло 2 592 опитувальники. Через дуже малу чисельність окремих гендерних груп у вибірці до технічної моделі кластеризації було включено лише респондентів, які вказали жіночу або чоловічу стать. Респонденти, які вказали варіанти «трансгендерна жінка», «трансгендерний чоловік» або «небінарна особа», не були включені до побудови кластерного рішення через недостатню чисельність для стабільної кластеризації та ризик некоректної інтерпретації результатів для малих груп. Відповідно, до масиву кластерного аналізу увійшло 2 543 респонденти.

Одним із методичних аргументів на користь такого рішення є те, що змінна статі в цій моделі використовувалася як одна з ознак для розрахунку відстаней між респондентами. За дуже малих важко надійно оцінити, чи справді їхнє розміщення в певних кластерах відображає стійкий поведінковий профіль, а не випадкові особливості окремих респондентів.

Це рішення також важливе для подальшого використання кластерів у профілюючому алгоритмі: дуже малі групи не дають достатньої кількості респондентів для стабільного навчання або перевірки алгоритму, який попередньо відносить нового клієнта до одного з сегментів. Водночас ці респонденти залишаються частиною загальної аудиторії дослідження і враховуються в загальному описовому аналізі там, де це методично доцільно.

Це рішення є технічним і не означає виключення цих груп із цільової аудиторії проєкту Drugstore – вони отримують послуги проєкту на рівних з іншими.

#### 3.2. Змінні, використані для кластеризації

##### 3.2.1. Демографічні змінні

До кластеризації було включено дві демографічні змінні: вік респондента та стать. Вік був представлений як метрична змінна у повних роках, стать використовувалася як бінарна ознака.

##### 3.2.2. Змінні щодо ПАР, які вживаються неін'єкційно

Основний масив показників стосувався неін'єкційного вживання ПАР. До нього увійшли 14 окремих змінних, кожна з яких фіксувала давність останнього неін'єкційного вживання певної речовини. У модель було включено такі речовини: MDMA/екстазі, амфетамін, LSD, кокаїн, метамфетамін, кетамін, GHB/GBL, мефедрон, альфа-PVP, інші синтетичні катіони, гриби/псилоцибін, 2C-B,

опіати/опіоїди та прегабалін («Лірика»). Відкрита категорія «інше» до кластеризації не входила.

Ці показники відображали не інтенсивність споживання у дозах і не кількість епізодів, а саме давність останнього вживання кожної речовини. Таким чином, модель фіксувала радше профіль актуальності різних ПАР у поведінці респондента, ніж обсяг споживання.

### 3.2.3. Змінні щодо ін'єкційного вживання ПАР

Ін'єкційний компонент був представлений змінною, яка вимірювала давність ін'єкційного вживання психоактивних речовин: від ін'єкцій протягом останніх 30 днів до відсутності такого досвіду взагалі. Саме ця змінна задавала окремий вимір ризику, який був особливо важливий для відокремлення респондентів з історією або актуальною практикою ін'єкційного вживання від тих, хто ніколи не вживав ПАР ін'єкційним способом.

### 3.2.4. Змінні щодо вживання алкоголю та поєднання алкоголю з ПАР

Алкогольний блок у кластеризації було представлено трьома змінними. Одна з них відображала, чи траплялося респонденту протягом останніх 6 місяців одночасно або в межах однієї сесії вживання ПАР/вечірки вживати алкоголь та інші психоактивні речовини. Друга фіксувала суб'єктивне відчуття, чи збільшується кількість алкоголю при вживанні ПАР. Третя описувала частоту епізодів вживання 5 і більше стандартних порцій алкоголю за один раз протягом останнього року.

### 3.2.5. Змінні щодо сексуальної поведінки

Сексуальний блок у кластеризації складався з семи показників. П'ять змінних фіксували давність сексуальних контактів із різними типами партнерів: з чоловіками, з жінками, з двома або більше чоловіками одночасно, з двома або більше жінками одночасно, а також із двома або більше партнерами різної статі одночасно. У такий спосіб до кластеризації було включено структуру сексуальних партнерів респондента та наявність або відсутність практик групового сексу.

Також до сексуального блоку входили змінні, які, по-перше, відображали факт вживання ПАР перед або під час останнього сексуального контакту, тобто індикатор сексуалізованого вживання ПАР, по-друге, використання презерватива під час останнього сексу. Разом ці змінні дозволяли врахувати у сегментації сексуальні контексти ризиків.

### 3.3. Перекодування змінних та обробка пропущених значень

При підготовці до кластерного аналізу було перевірено рівень пропущених і невизначених відповідей у кластероутворюючих змінних. На рівні окремих змінних частка таких відповідей загалом була невисокою: вік і стать не мали пропусків, більшість поведінкових показників мали менше 5% пропущених значень. Найвищі частки були зафіксовані для змінної щодо збільшення вживання алкоголю при вживанні ПАР — 8,3%, вживання синтетичних катіонів — 7,3%, частоти епізодів вживання 5 і більше стандартних порцій алкоголю — 5,4%, ін'єкційного вживання ПАР — 5,0% та вживання 2С-В — 5,0%. Сексуальний блок мав нижчий рівень пропусків: для більшості відповідних змінних він становив приблизно 1–4%.

Водночас через велику кількість змінних у моделі повністю заповнені дані за всіма 27 кластероутворюючими показниками мали 62,2% респондентів, тоді як 37,8% мали хоча б одне пропущене або невизначене значення.

Для здійснення сегментації було підготовлено дві аналітичні версії даних. Перша використовувалася для кластерних процедур у SPSS — ієрархічної кластеризації та двоетапного кластерного аналізу. Друга використовувалася для альтернативної перевірки в R із застосуванням відстані Г'ювера та алгоритму PAM.

У версії даних для SPSS невизначені відповіді в змінних для кластеризації було попередньо перекодовано у валідні категорії. Це було потрібне для того, щоб зменшити втрату респондентів через пропущені значення та забезпечити порівнюваність кластерних рішень, оскільки кластерні процедури в SPSS включають в аналіз тільки тих респондентів, які не мають пропущених змінних в жодній змінній. Зокрема, відповіді типу «не знаю / не пам'ятаю», «відмова від відповіді» або «важко відповісти» не залишалися окремими пропущеними значеннями, а включалися до змістовно найближчих категорій.

У змінній щодо ін'єкційного вживання ПАР невизначені відповіді було прирівняно до категорії відсутності ін'єкційного досвіду. У блоці неін'єкційного вживання ПАР відповіді «не знаю / не пам'ятаю» для окремих речовин було перекодовано у категорію «ніколи». У блоці алкоголю невизначені відповіді було віднесено до варіантів, що відповідають відсутності або мінімальному прояву відповідної практики: відсутність поєднання алкоголю з ПАР, відсутність збільшення кількості алкоголю під впливом ПАР або найнижча частота епізодів вживання великої кількості алкоголю.

У сексуальному блоці невизначені відповіді щодо досвіду контактів із різними типами партнерів було перекодовано як відсутність відповідного досвіду. Невизначену відповідь щодо вживання ПАР перед або під час останнього сексу було прирівняно до відповіді «ні». Невизначену відповідь щодо використання презерватива під час останнього сексу було прирівняно до відповіді «так».

Таке перекодування мало прагматичний характер і було пов'язане з технічними вимогами та логікою використаних SPSS-процедур. Водночас, очевидно, воно не є повністю нейтральним методичним рішенням, оскільки невизначені відповіді здебільшого включалися в модель як менш ризикові стани. Тому це рішення слід розглядати як один із методичних компромісів прикладної сегментації.

Для альтернативної перевірки в R було використано іншу логіку роботи з даними. У цій версії невизначені або пропущені значення не перекодовувалися за SPSS-логікою у менш ризикові категорії. Відстань Говера, розрахована за допомогою пакету *daisy*, дозволяє працювати зі змішаними типами змінних і обчислювати подібність між респондентами на основі тих показників, за якими для конкретної пари респондентів є доступні значення. Після цього кластеризація виконувалася алгоритмом PAM на матриці відстаней.

Порівняння цих підходів було потрібне для того, щоб оцінити, наскільки отримане кластерне рішення залежить від способу підготовки даних та вибору методів кластеризації.

#### 3.4. Робота із змінними різного типу та стандартизація змінних

Стандартизація змінних застосовувалася насамперед в ієрархічній кластеризації SPSS. У цьому підході всі кластероутворюючі змінні розглядалися як метричні або псевдометричні, а відстані між респондентами обчислювалися за допомогою квадрату евклідової відстані. Перед розрахунком матриці відстаней змінні було стандартизовано до z-оцінок, тобто приведено до шкали із середнім значенням 0 та стандартним відхиленням 1.

Це було необхідно, оскільки до кластеризації входили показники з різним числовим діапазоном: вік, бінарна змінна статі, змінні давності вживання ПАР, ін'єкційного досвіду, поєднання ПАР з алкоголем, сексуальної поведінки та використання презерватива. Без стандартизації змінні з ширшим діапазоном значень могли б механічно сильніше впливати на розрахунок відстаней і, відповідно, на формування кластерів.

Така логіка спиралася на припущення, що частота або нещодавність певного досвіду може розглядатися як ступінь вираженості відповідної ознаки або практики у респондента. Наприклад, вживання певної ПАР протягом останніх 30 днів інтерпретувалося як прояв більшої частоти вживання цієї ПАР респондентом, ніж вживання понад 6 місяців тому або відсутність такого досвіду. Аналогічно, частіше поєднання алкоголю з ПАР, актуальніший сексуальний досвід певного типу або відсутність презерватива під час останнього сексу розглядалися як індикатори більшої інтенсивності відповідної поведінки.

Деякі публікації вказують, що такий підхід може нести ризики для валідності статистичного аналізу<sup>7</sup>, в той же час багато джерел вказують, що використання

---

<sup>7</sup> Torrin M. Liddell, John K. Kruschke. Analyzing ordinal data with metric models: What could possibly go wrong? *Journal of Experimental Social Psychology*, Volume 79, 2018, Pages 328-348, ISSN 0022-1031, <https://doi.org/10.1016/j.jesp.2018.08.009>.  
(<https://www.sciencedirect.com/science/article/pii/S0022103117307746>)

порядкових змінних в параметричних статистичних методах в соціальних науках припустимо і не зміщує результати аналітичних процедур<sup>8 9 10</sup>.

У двоетапному кластерному аналізі SPSS використовувалася інша логіка: змінні розглядалися як категоріальні (крім віку). Тому стандартизація, як в ієрархічній процедурі, тут не застосовувалася. Цей підхід був корисний як альтернативна перевірка, оскільки дозволяв оцінити, чи зберігається змістовна структура сегментів тоді, коли змінні не трактуються як псевдометричні шкали і застосовується інший алгоритм кластеризації.

В аналізі в R також використовувалася інша логіка вимірювання подібності. Відстань Говера спеціально пристосована для роботи зі змішаними типами даних<sup>11</sup>. Вона не потребує такої стандартизації як ієрархічна кластеризація, оскільки внесок змінних нормується автоматично в межах розрахунку відстані. Для порядкових і категоріальних показників це є методично більш природним варіантом, ніж пряме трактування всіх змінних як метричних.

Отже, три використані підходи спиралися на різну логіку підготовки даних: в ієрархічній кластеризації SPSS змінні були стандартизовані та розглядалися як метричні або псевдометричні; у двоетапному кластерному аналізі SPSS вони трактувалися як категоріальні; в R-підході подібність між респондентами оцінювалася через відстань Говера, яка краще відповідає змішаній природі використаних показників.

---

<sup>8</sup> Norman, G. Likert scales, levels of measurement and the “laws” of statistics. *Adv in Health Sci Educ* 15, 625–632 (2010). <https://doi.org/10.1007/s10459-010-9222-y>

<sup>9</sup> Huh, I., Gim, J. Exploration of Likert scale in terms of continuous variable with parametric statistical methods. *BMC Med Res Methodol* 25, 218 (2025). <https://doi.org/10.1186/s12874-025-02668-1>

<sup>10</sup> Sullivan, G. M., & Artino, A. R., Jr (2013). Analyzing and interpreting data from likert-type scales. *Journal of graduate medical education*, 5(4), 541–542. <https://doi.org/10.4300/JGME-5-4-18>

<sup>11</sup> Gower, J. C. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics*, 27(4), 857–871. <https://doi.org/10.2307/2528823>

## 4. Метод кластеризації

### 4.1. Методи кластерного аналізу, що використовувалися

Для пошуку найбільш інтерпретованого та практично корисного кластерного рішення ми протестували три підходи: двоетапний кластерний аналіз у SPSS, ієрархічну кластеризацію методом Варда у SPSS та кластеризацію PAM на відстанях Говера в R.

Ці методи відрізняються за алгоритмами побудови кластерів і мають різні вимоги до типів даних. Двоетапний алгоритм може працювати з категоріальними та метричними змінними одночасно, однак у нашому випадку має важливе методичне обмеження: фактично він розглядає категоріальні змінні як номінальні, через що в порядкових ознаках втрачається градієнт упорядкованості, а це цінна інформація для нашого аналізу.

Ієрархічний метод Варда є класичним підходом до кластеризації, який формує групи шляхом послідовного об'єднання об'єктів за критерієм мінімізації внутрішньо кластерної неоднорідності. Водночас для наших даних його застосування також потребує обережності, оскільки набір змінних поєднує бінарні та умовно псевдометричні ознаки, а також включає робочі рішення щодо обробки пропусків.

Окремо ми протестували PAM-кластеризацію на відстанях Говера в R — підхід, який методично краще пристосований до змішаних даних.

### 4.2. Двоетапний кластерний аналіз

Двоетапний кластерний аналіз у SPSS було виконано для 2543 спостережень із використанням likelihood distance, автоматичного вибору кількості кластерів за критерієм BIC та верхньої межі в 5 кластерів. В межах заданого діапазону 1–5 кластерів алгоритм підтримав 5-кластерне рішення: значення BIC послідовно зменшувалося від 134718,315 для одного кластера до 121046,554 для п'яти кластерів, а розміри виділених кластерів становили 22,7%, 25,5%, 21,2%, 11,2% та 19,4%. Водночас силуетна міра зв'язності та розділення кластерів становила лише 0,093, що свідчить про низьку статистичну якість кластеризації та слабку відокремленість сегментів.

### 4.3. Ієрархічний кластерний аналіз

Ієрархічний кластерний аналіз методом Варда було виконано для 2543 спостережень із подальшим розглядом 5- та 4-кластерного варіантів розбиття. Для 5-кластерного рішення розміри кластерів становили 18,0%, 25,1%, 31,9%, 18,8% та 6,1%. Силуетний аналіз на основі міри несхожості Говера показав низьку статистичну якість цього рішення: загальний середній коефіцієнт силуету становив 0,054.

Для 4-кластерного рішення розміри кластерів становили 43,1%, 31,9%, 18,8% та 6,1%, а загальний середній коефіцієнт силуету дорівнював 0,112. Порівняно з 5-

кластерним варіантом це вказує на дещо кращу, але все одно низьку якість кластеризації.

4-кластерне рішення фактично є вкладеним у 5-кластерне: три кластери залишаються незмінними, тоді як один великий кластер 4-кластерного варіанта розщеплюється на два окремі кластери у 5-кластерному рішенні. Важливо, що це розщеплення не призводить до утворення дрібних груп, оскільки нові кластери охоплюють 18,0% та 25,1% вибірки. Отже, 5-кластерне рішення можна розглядати як більш деталізований варіант сегментації, який уточнює внутрішню неоднорідність одного з великих кластерів, а не як повністю іншу або нестійку кластерну структуру.

Водночас перевага 5-кластерного рішення в цьому випадку пов'язана насамперед із більшою аналітичною деталізацією і практичною корисністю, а не з кращими формальними показниками якості кластеризації.

#### 4.4. PAM-кластеризація

Кластеризацію в R було виконано методом PAM (Partitioning Around Medoids) на основі матриці несхожості Говера<sup>12</sup>. Для зіставлення з іншими підходами спочатку було розглянуто 5-кластерне рішення. Розміри кластерів у цьому варіанті становили 19,0%, 21,9%, 10,4%, 29,2% та 19,6% вибірки. середній коефіцієнт силуету для 5-кластерного рішення становив лише 0,040, що свідчить про дуже низьку статистичну якість кластеризації та слабку відокремленість кластерів. Отже, це рішення не дає підстав трактувати 5-кластерну структуру як добре виражену в межах PAM на відстанях Говера.

Для додаткової технічної перевірки було побудовано PAM-рішення для 2–8 кластерів. Найвище середнє значення силуету було отримано для 2-кластерного варіанта (0,127), після чого зі збільшенням кількості кластерів цей показник послідовно зменшувався. Це вказує на те, що в межах цієї процедури дані підтримують дуже грубе розбиття на два кластери, ніж детальнішу сегментацію. Водночас 2-кластерне рішення не становить значного аналітичного чи практичного інтересу для цілей цього дослідження, оскільки є надто узагальненим і не дозволяє виділити змістовно диференційовані профілі клієнтів проєкту.

---

<sup>12</sup> <https://www.rdocumentation.org/packages/cluster/versions/2.1.8.2/topics/daisy>

## 5. Порівняння підходів до кластеризації та вибір робочого рішення

### 5.1. Вибір робочої моделі сегментації

Для перевірки стійкості сегментації ми протестували три підходи до кластеризації, що спираються на різну статистичну логіку. Формальні показники якості не дають підстав вважати будь-яке з отриманих рішень беззаперечно сильним: в усіх підходах показники силуету залишалися низькими, а отже кластерну структуру не слід трактувати як чітко відмежовану «природну» типологію. Водночас ці показники коректно інтерпретувати насамперед у межах тієї самої процедури та тієї самої метрики, тому вони використовувалися тут як внутрішній діагностичний орієнтир, а не як універсальна підстава для прямого ранжування всіх методів між собою<sup>13</sup>.

З урахуванням цього, при виборі фінального рішення ми орієнтувалися не стільки на формальні статистичні критерії, скільки на інтерпретованість та практичну корисність сегментації.

Важливо, що двоетапний алгоритм у межах автоматичного добору за критерієм BIC визначив 5-кластерне рішення як найкраще в заданому діапазоні (від 1 до 5 кластерів), що підтримує доцільність подальшого розгляду саме п'ятикластерної структури.

Серед протестованих варіантів найбільш рельєфним, змістовно послідовним і придатним для аналітичного опису виявилось 5-кластерне рішення, отримане в ієрархічному кластерному аналізі методом Варда. Саме його ми використовуємо далі як основну робочу сегментацію. Водночас результати двоетапної кластеризації розглядаються як важливий матеріал для порівняння й обговорення, оскільки вони не заперечують загальну структуру сегментації, а радше дозволяють додатково перевірити її стійкість і межі інтерпретації.

### 5.2. Порівняння типологій, отриманих методом ієрархічного кластерного аналізу та двоетапним кластерним аналізом

Перехресна таблиця двох класифікацій показала високий рівень відповідності між 5-кластерними рішеннями, отриманими методом Варда та двоетапним кластерним аналізом<sup>14</sup>. Найближче між підходами збігаються три базові сегменти: у двоетапному рішенні 65,8% кластера 1 становлять випадки з кластера 1 за методом Варда, 70,8% кластера 2 — випадки з кластера 2, а кластер 3 майже повністю відтворює відповідний кластер в методі Варда, оскільки 94,3% його складу припадає саме на нього. Найбільші розбіжності стосуються більш ризикованих профілів: двоетапний кластер 4 формується переважно за рахунок кластерів 5 і 4 за методом Варда, а двоетапний кластер 5 — передусім за рахунок кластера 4 та помітної частини кластера 3. Очевидно, двоетапний алгоритм досить

---

<sup>13</sup> Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics*, Volume 20, 1987, Pages 53-65, ISSN 0377-0427, [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).  
(<https://www.sciencedirect.com/science/article/pii/0377042787901257>)

<sup>14</sup> Cramer's  $V = 0,636$


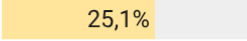


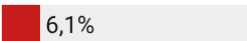
близько відтворює три базові сегменти, але інакше розмежовує більш ризиковані та змішані групи.

## 6. Розподіл респондентів за кластерами

У межах обраного 5-кластерного рішення, отриманого методом Варда, усі 2543 респонденти аналітичної вибірки були розподілені між п'ятьма кластерами. Кількісно вони представлені так: кластер 1 — 459 осіб, кластер 2 — 638 осіб, кластер 3 — 811 осіб, кластер 4 — 479 осіб, кластер 5 — 156 осіб. Найбільшим за чисельністю є кластер 3, тоді як найменшим — кластер 5. Водночас навіть найменший кластер має достатній обсяг для окремого змістовного аналізу та подальшого профілювання. У відносних показниках кластери становлять відповідно 18,0%, 25,1%, 31,9%, 18,8% та 6,1% аналітичної вибірки. Отже, майже третина респондентів належить до кластера 3, близько чверті — до кластера 2, а кластери 1 і 4 мають близьку вагу в загальній структурі вибірки.

*Діаграма 1. Розподіл респондентів за кластерами та інтегральна оцінка проблемності.*

## Розмір кластерів та інтегральна оцінка проблемності

Кластер	Частка вибірки, %	Інтегральний рівень проблемності кластеру
Кластер 1 — жіночий відносно низькоризиковий сегмент із запитом на психологічну підтримку	 18,0%	Низький за поведінковими ризиками; підвищений за потребою в психологічній підтримці
Кластер 2 — жіночий сегмент підвищеного ризику з акцентом на вживання стимуляторів та емпатогенів	 25,1%	Підвищений
Кластер 3 — чоловічий переважно гетеросексуальний сегмент помірного ризику	 31,9%	Помірний
Кластер 4 — чоловічий ін'єкційно-поліспоживчий сегмент	 18,8%	Високий
Кластер 5 — ядро полівживання, групового сексу та соціальної вразливості	 6,1%	Дуже високий

Створено за допомогою Datawrapper

## 7. Профілі кластерів

5-кластерне рішення методом Варда формує п'ять змістовно різних сегментів, які відрізняються за статевим складом, віком, матеріальним становищем, моделями вживання психоактивних речовин, ін'єкційним компонентом, сексуальною поведінкою та пов'язаними ризиками. У структурі сегментації простежуються дві відносно мало ризикові групи, один жіночий сегмент із підвищеною залученістю до вживання стимуляторів та емпатогенів, один переважно чоловічий сегмент із виразним ін'єкційним і поліспоживчим профілем та один найбільш концентрований з точки зору ризиків сегмент, у якому поєднуються полівживання, груповий секс, практики сексуалізованого вживання наркотиків та соціально-економічна вразливість. За сукупністю показників найбільш ризиковими виглядають кластери 4 і 5, тоді як кластери 1 і 3 описують порівняно благополучні профілі з різними гендерними характеристиками.

Робочі назви кластерів узагальнюють їхні найбільш виразні риси та використовуються як аналітичні позначення для компактного представлення профілів.

### 7.1. Кластер 1 — жіночий відносно низько ризиковий сегмент із запитом на психологічну підтримку

Молодий, найбільш «жіночий» сегмент відносно низького поведінкового ризику. Медіана віку тут становить 23 роки, середній вік – 24,3 роки; це разом з кластером 5 молодший полюс вибірки, і, ймовірно, найменш ризиковий. Ін'єкційний досвід в кластері майже відсутній: 0,7% мали його за останні 30 днів, 90,0% ніколи не вживали ПАР ін'єкціями. За спектром уживаних речовин кластер 1 характеризується відносно стриманим профілем. Для амфетаміну, мефедрону, кетаміну, альфа-PVP, інших солей та опіоїдів у ньому переважають відповіді «ніколи», а частки недавнього вживання нижчі, ніж у кластерах 2, 4 і 5. 19,4% учасників повідомили про вживання MDMA в останні 30 днів, амфетаміну — 10,5%, прегабаліну — 12,0%, тобто показники доволі низькі.

В цьому сегменті алкоголь із ПАР поєднується найрідше серед усіх кластерів: лише 7,2% респондентів роблять це «майже кожного разу» коли вживають ПАР. Сексуальний ризик також мінімальний: груповий секс протягом життя мали 19,4%, за останні 30 днів — ніхто, а ПАР перед або під час останнього сексу вживали 13,5%.

Водночас цей кластер не особливо включений у профілактику: 72,3% не тестувалися на жодну ІПСШ чи вірусний гепатит упродовж останніх шести місяців. Щодо медичної допомоги, то серед тих, хто звертався по неї, переважає використання приватного сектору: 69,0% проти 42,9% державного<sup>15</sup>.

Водночас для цього кластера виразно характерний запит на психологічну підтримку: 73,6% повідомили що в них останнім часом була потреба в допомозі психолога або психотерапевта, і це найвищий показник серед усіх сегментів.

---

<sup>15</sup> Респонденти могли користуватися і приватною, і державною медициною протягом року.

Матеріальне становище цього сегменту виглядає відносно стабільним: серед тих, хто відповів на запитання, 41,9% зазначили, що їм вистачає грошей на їжу й одяг і вони можуть дещо відкладати.

## 7.2. Кластер 2 — жіночий сегмент підвищеного ризику, акцент на вживання стимуляторів та емпатогенів

Цей сегмент також молодий і переважно жіночий, але вже з виразно підвищеним ризиком за рахунок більш частого вживання стимуляторів, поєднання алкоголю з ПАР та сексуальної поведінки. Медіана віку — 24 роки, середній вік — 24,9 років. Ін'єкційне вживання ПАР в кластері малопоширене: 2,0% за останні 30 днів; 84,0% ніколи не мали ін'єкційного досвіду. В той же час, інші поведінкові ризики помітно вищі, ніж у кластера 1: 41,1% майже щоразу вживають алкоголь при вживанні ПАР, 28,5% повідомляють, що під час вживання ПАР майже завжди п'ють більше алкоголю, 50,2% мали досвід групового сексу протягом життя, 10,5% — упродовж останніх 30 днів, а 41,8% вживали ПАР перед або під час останнього сексу.

Основна риса цього сегменту — активне вживання стимуляторів і емпатогенів ПАР. Про вживання амфетаміну в останні 30 днів повідомили 41,1% учасників кластеру, MDMA — 31,5%, мефедрону — 28,8%, прегабаліну — 30,4%, кокаїну — 20,1%, кетаміну — 15,4%. Для низки речовин частка відповіді «ніколи» тут суттєво нижча, ніж у кластерах 1 і 3.

Досвід транзакційного сексу протягом останніх 12 місяців зазначили 23,3% тих, хто відповідав на відповідне запитання.

71,9% респондентів цього кластера коли-небудь проходили тестування на ВІЛ, що є одним із вищих показників серед сегментів. У сукупності цей кластер доцільно розглядати як жіночу групу з активним вживанням стимуляторів і підвищеною залученістю до супутніх ризиків.

## 7.3. Кластер 3 — чоловічий переважно гетеросексуальний сегмент помірного ризику

Старший, переважно чоловічий сегмент помірного ризику. Медіана віку — 25 років; саме кластер 3 разом із кластером 4 формує «старший» полюс вибірки. Ін'єкційне вживання наркотиків тут радше рідкісне, хоча й не зовсім відсутнє: 3,6% за останні 30 днів, 4,7% — за останні шість місяців, 81,3% ніколи не вживали ін'єкційно. Регулярне поєднання алкоголю з ПАР трапляється значно рідше, ніж у кластерах 2, 4 і 5 (16,8%), досвід групового сексу протягом життя мають 28,5%, за останні 30 днів — 3%, а ПАР перед/під час останнього сексу вживали 20,6%.

За матеріальним становищем сегмент виглядає найбільш благополучним: 26,0% учасників, які відповіли на це запитання, зазначили, що можуть дозволити собі купувати деякі коштовні речі; це найвищий показник серед усіх кластерів.

За спектром речовин, що вживаються, кластер 3 займає проміжну позицію. Частка вживання амфетаміну за останні 30 днів становить 21,8%, MDMA — 22,9%, мефедрону — 9,0%, прегабаліну — 13,7%. Для багатьох позицій, зокрема

мефедрону, інших «солей», альфа-PVP та опіатів/опіоїдів, у цьому сегменті частіше фіксується відповідь «ніколи», ніж у кластерах 4 і 5.

Сексуальний профіль кластера 3 є переважно гетеросексуальним. 62,5% учасників мали секс із жінкою протягом останніх 30 днів, а 69,3% ніколи не мали сексу з чоловіком. Групові сексуальні практики тут поширені менше, ніж у кластерах 4 і 5: 93,5% ніколи не мали сексу з двома або більше чоловіками одночасно, 74,8% — з двома або більше жінками одночасно, 86,6% — з партнерами різної статі одночасно. ПАР перед або під час останнього сексу вживали 20,6%, про транзакційний секс протягом останніх 12 місяців повідомили 5,2% серед тих, хто відповідав

За сукупністю ознак це переважно чоловічий, більш ресурсний і відносно низько ризиковий сегмент.

#### 7.4. Кластер 4 — чоловічий ін'єкційно-поліспоживчий сегмент

Старший чоловічий ін'єкційно-поліспоживчий сегмент і один із двох головних високоризикових вузлів типології. Медіана віку — 25 років. В цьому кластері спостерігається виразний ін'єкційний компонент: 25,1% вживали ПАР ін'єкційно за останні 30 днів, ще 8,8% — за останні шість місяців, а частка тих, хто ніколи не мав такого досвіду, склала всього 56,6%. Алкогольний супровід вживання ПАР також інтенсивний: 48,0% майже щоразу поєднують алкоголь із ПАР, 31,1% майже завжди п'ють більше алкоголю під час вживання ПАР.

За матеріальним профілем сегмент є менш благополучним, ніж кластер 3: 13,6% серед тих, хто відповідав на запитання, зазначили, що їм не вистачає грошей навіть на їжу, а 39,0% — що грошей вистачає лише на їжу, але купувати одяг уже важко

Кластер 4 демонструє широкий спектр актуального вживання ПАР: амфетамін в останні 30 днів — 48,0%, MDMA — 37,4%, мефедрон — 33,6%, прегабалін — 49,7%, кокаїн — 25,3%, метамфетамін — 24,8%, кетамін — 19,0%, альфа-PVP — 17,5%, інші солі — 15,7%, опіати/опіоїди — 19,6%. За кількома позиціями цей сегмент поступається лише кластеру 5.

Сексуальна поведінка кластера 4 також пов'язана з помітним рівнем ризику. ПАР перед або під час останнього сексу вживали 54,3% учасників, а 20,8% серед релевантної підвибірки повідомили про досвід транзакційного сексу<sup>16</sup>. Серед тих, хто мав досвід групового сексу, 74,6% зазначили, що під час останнього такого контакту вживали ПАР.

У сукупності кластер 4 описує один із двох головних високоризикових сегментів моделі, у якому поєднуються ін'єкційне вживання, поліспоживання ПАР та суттєві сексуальні ризики.

---

<sup>16</sup> Блок про транзакційний секс ставився респондентом не весь час протягом дослідження в 2025 році, тому деякі учасники дослідження не відповідали на ці питання.

## 7.5. Кластер 5 — ядро полівживання, групового сексу та соціальної вразливості

Наймолодший за віком (медіана 22 роки, середній вік – 23,7 років), найменший за чисельністю, але такий, що демонструє концентрацію поведінкових ризиків. Саме тут видно найповніше накладання ін'єкційного, алкогольного та сексуального профілю ризиків: 31,4% вживали ПАР ін'єкційно за останні 30 днів, 19,2% — за останні шість місяців, а лише 34,0% ніколи не вживали ін'єкційно. Алкоголь супроводжує вживання ПАР найчастіше з усіх кластерів: 52,6% поєднують його з ПАР майже кожного разу, 35,3% під час вживання ПАР майже завжди випивають більше.

Найрізкіше кластер 5 вирізняється за сексуальною поведінкою: 100% мають досвід групового сексу протягом життя, 69,9% — за останні 30 днів, а 80,8% вживали ПАР перед або під час останнього сексу

За матеріальним становищем це найуразливіший сегмент: 24,0% респондентів, які відповіли на відповідне запитання, зазначили, що їм не вистачає грошей навіть на їжу, ще 40,0% — що грошей вистачає лише на їжу, але купівля одягу вже викликає труднощі.

У підсумку кластер 5 — це ядро ризиків.

## 7.6. Узагальнення

Отримана сегментація описує п'ять різних профілів. Кластер 1 концентрує переважно жінок із порівняно нижчим рівнем наркотичного та сексуального ризику, але з виразним запитом на психологічну підтримку. Кластер 2 також є переважно жіночим, проте має значно активніший профіль вживання стимуляторів і вищу залученість до супутніх ризикових практик. Кластер 3 формує переважно чоловічий і більш гетеросексуальний сегмент із помірним рівнем ризику та відносно кращим матеріальним становищем. Кластер 4 об'єднує переважно чоловіків із виразним ін'єкційним компонентом і широким поліспоживанням. Кластер 5 є найменшим за розміром, але найбільш концентрованим за поєднанням полівживання, групового сексу, сексуалізованого вживання наркотиків, транзакційного сексу і соціальної вразливості.

Така структура є достатньо виразною для подальшого аналізу потреб сегментів і підготовки практичних рекомендацій для таргетованих інтервенцій.

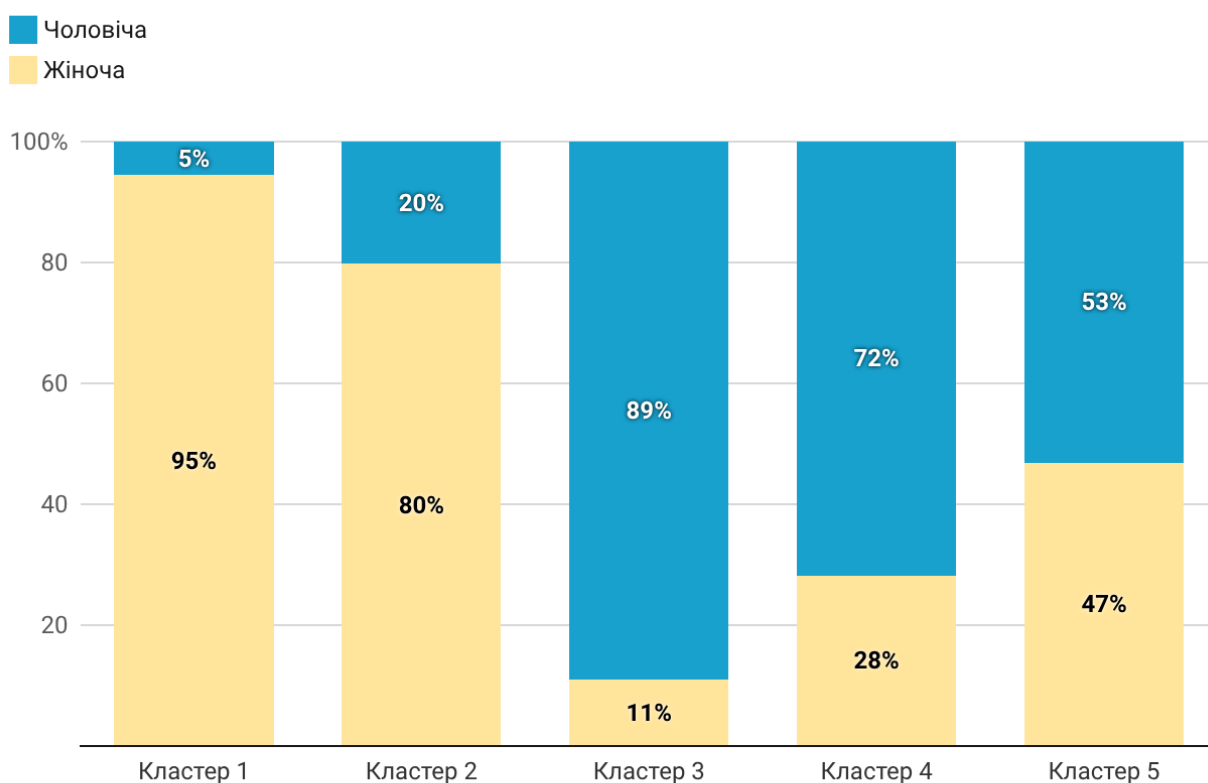
## 8. Порівняльний профіль кластерів

### 8.1. Порівняння за демографічним профілем

Кластери суттєво відрізняються за статевим складом респондентів. Кластер 1 є найбільш «жіночим»: 94,6% його учасників становлять жінки. Кластер 2 також має переважно жіночий склад — 79,9% жінок. Кластер 3, навпаки, є найбільш «чоловічим» сегментом: 88,9% його учасників — чоловіки. Кластер 4 теж має виразний чоловічий профіль, хоча менш однорідний, ніж кластер 3: 71,8% його учасників становлять чоловіки. Кластер 5 має найбільш збалансований статевий склад серед усіх сегментів: 53,2% чоловіків і 46,8% жінок. Загалом відмінності між кластерами за статтю є статистично значущими і мають високу силу зв'язку<sup>17</sup>.

Діаграма 2. Статевий склад кластерів.

### Статевий склад кластерів



Створено за допомогою Datawrapper

За віком кластери також розрізняються доволі чітко<sup>18</sup>. Наймолодшим є кластер 5, де середній вік становить 23,7 років. До молодших сегментів належать також кластер 1 із середнім віком 24,3 року та кластер 2 із середнім віком 24,9 року. За результатами робастних тестів загальна різниця між кластерами є статистично

<sup>17</sup> Cramer's V = 0,680; p < 0,001

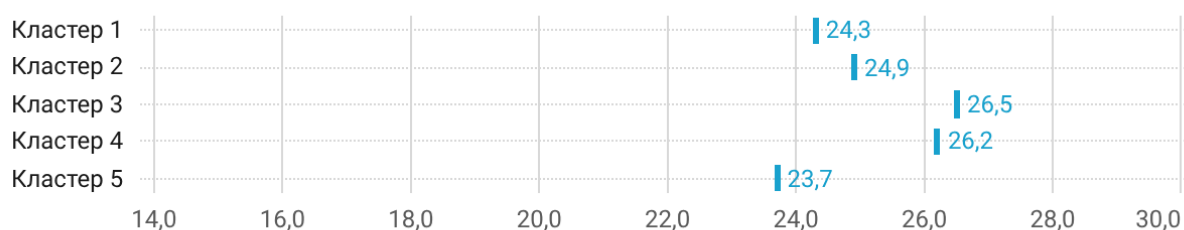
<sup>18</sup> Нагадаємо, що це дослідження групи людей, які практикують вживання ПАР, і ця група в принципі більш гомогенна за віком, ніж, наприклад, загальне населення. Тому хоча відмінності між кластерами за віком можуть здаватися не дуже принциповими, потрібно брати до уваги специфіку вибірки.

значущою<sup>19</sup>. Попарні порівняння показують, що кластери 3 і 4 є статистично значуще старшими за кластери 1, 2 і 5, тоді як між собою кластери 3 і 4 не відрізняються, так само як між собою не відрізняються кластери 1, 2 і 5. Кластери 1, 2 і 5 можна віднести до молодшої частини вибірки, а кластери 3 і 4 — до дещо старшої.

Непараметричний аналіз підтвердив, що вік статистично значуще відрізняється між кластерами на рівні медіан, і на рівні загального розподілу. Медіани віку становлять 22 роки у кластері 5, 23 роки у кластері 1, 24 роки у кластері 2 та по 25 років у кластерах 3 і 4. Кластери 3 і 4 є старшими за кластер 1, а кластер 5 — молодшим за кластери 2, 3 і 4. Водночас кластери 3 і 4 між собою не відрізняються, так само як немає надійної різниці між кластерами 1 і 5.

Діаграма 3. Середній вік за кластерами.

### Середній вік респондентів різних кластерів (в роках)



Створено за допомогою Datawrapper

Матеріальне становище також пов'язане з кластерною належністю<sup>20</sup>. Найбільш ресурсним виглядає кластер 3: 26,0% його учасників зазначили, що можуть дозволити собі купувати деякі коштовні речі; це найвищий показник серед усіх сегментів. Найбільш вразливим є кластер 5: 24,0% його представників повідомили, що їм не вистачає грошей навіть на їжу, а ще 40,0% зазначили, що грошей вистачає лише на їжу, але купувати одяг уже важко. Кластер 4 також демонструє ознаки матеріальних проблем: 13,6% його учасників не мають достатньо коштів навіть на їжу, а 39,0% мають труднощі з купівлею одягу. Кластери 1 і 2 займають проміжну позицію: в обох по 41,9% респондентів зазначили, що їм вистачає грошей на їжу та одяг і вони можуть дещо відкладати. Загальний зв'язок між кластерами та матеріальним становищем є статистично значущим, хоча його сила помірною<sup>21</sup>.

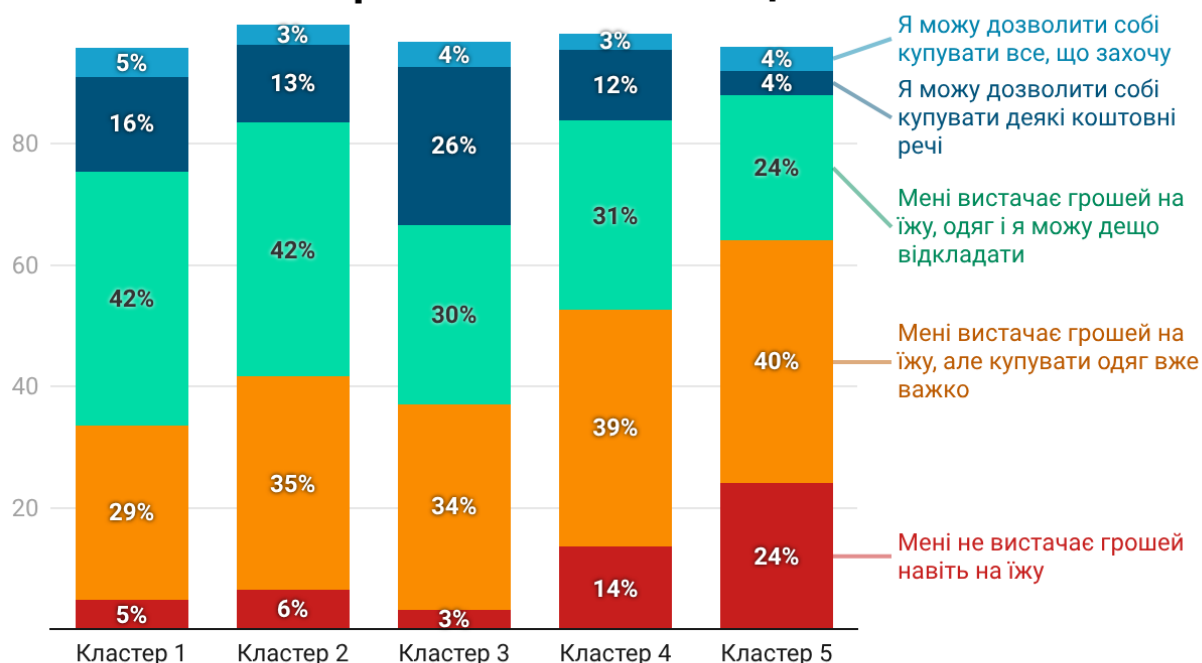
<sup>19</sup> Welch = 16,682;  $p < 0,001$ ; Brown–Forsythe = 17,199;  $p < 0,001$

<sup>20</sup> Цей блок слід інтерпретувати обережно, оскільки запитання про фінансове становище ставили 895 респондентам, або 35,2% аналітичної вибірки

<sup>21</sup> Cramer's V = 0,152;  $p < 0,001$

Діаграма 4. Матеріальне становище респондентів за кластерами.

## Показники матеріального становища



Створено за допомогою Datawrapper

## 8.2. Порівняння за профілем вживання ПАР

Кластери виразно відрізняються за профілем вживання психоактивних речовин, показниками поліспоживання та практиками поєднання ПАР з алкоголем. Найбільш контрастно сегменти розходяться за показником ін'єкційного вживання ПАР і за частотою недавнього вживання стимуляторів, прегабаліну та синтетичних катіонів. Загалом у структурі кластерів простежуються різні моделі: від відносно стриманого профілю з мінімальним ін'єкційним досвідом до сегмента з максимальною концентрацією ін'єкційного вживання, широкого спектра речовин і супутніх ризиків.

### 8.2.1. Ін'єкційне вживання ПАР

Найвищий рівень ін'єкційного вживання ПАР спостерігається в кластерах 5 і 4. У кластері 5 31,4% респондентів повідомили про ін'єкційне вживання протягом останніх 30 днів, ще 19,2% — протягом останніх 6 місяців, але не в останні 30 днів; лише 34,0% ніколи не вживали ін'єкційно. У кластері 4 відповідні частки становлять 25,1% і 8,8%, а частка тих, хто ніколи не мав ін'єкційного досвіду, дорівнює 56,6%. Кластер 3 має помітний, але значно слабший ін'єкційний компонент: 3,6% — за останні 30 днів, 4,7% — за останні 6 місяців, 81,3% ніколи не вживали ін'єкційно. В кластерах 1 і 2 ін'єкційне вживання ПАР є малопоширеним: 90,0% і 84,0% відповідно ніколи не мали такого досвіду. Отже, саме кластери 4 і 5 формують основну

концентрацію ін'єкційного ризику. Зв'язок між кластерною належністю та ін'єкційним досвідом є статистично значущим<sup>22</sup>.

Діаграма 5. Досвід ін'єкційного вживання ПАР за кластерами.

## Досвід ін'єкційного вживання ПАР



Створено за допомогою Datawrapper

### 8.2.2. Види речовин, що вживаються неін'єкційно

За спектром уживаних речовин кластер 5 має найширший профіль. Саме тут зафіксовано найвищі частки вживання в останні 30 днів для більшості позицій: MDMA — 53,2%, амфетамін — 58,3%, LSD — 39,7%, кокаїн — 33,3%, метамфетамін — 28,2%, кетамін — 21,2%, бутират — 19,9%, мефедрон — 46,8%, альфа-PVP — 25,6%, інші солі — 23,7%, гриби/псилоцибін — 35,9%, 2C-B — 15,4%, опіати/опіоїди — 19,2%, прегабалін — 50,6%. Кластер 4 посідає наступну позицію за інтенсивністю споживання: амфетамін в останні 30 днів — 48,0%, MDMA — 37,4%, мефедрон — 33,6%, прегабалін — 49,7%, кокаїн — 25,3%, метамфетамін — 24,8%, кетамін — 19,0%, альфа-PVP — 17,5%, інші солі — 15,7%, опіати/опіоїди — 19,6%. Таким чином, кластери 4 і 5 концентрують найбільш виражене поліспоживання та найбільшу частку недавнього вживання високоризикових речовин.

Кластер 2 формує окремий профіль активного вживання речовин, що мають стимулюючий та емпатогенний профіль дії, за відсутності вираженого ін'єкційного компонента. Для цього сегменту характерні високі частки вживання амфетаміну (41,1% за останні 30 днів), MDMA (31,5%), мефедрону (28,8%), прегабаліну (30,4%), кокаїну (20,1%) та кетаміну (15,4%). За низкою позицій цей сегмент поступається кластерам 4 і 5, проте помітно перевищує кластери 1 і 3. Кластер 3 займає проміжне

<sup>22</sup>  $\chi^2(24)=546,368$ ;  $p<0,001$ , Cramer's  $V=0,232$

місце: у ньому також присутнє вживання стимуляторів і прегабаліну, але менш інтенсивне, ніж у кластерах 4 і 5, і без настільки вираженого ін'єкційного компонента. Частки вживання за останні 30 днів становлять тут 22,9% для MDMA, 21,8% для амфетаміну, 13,7% для прегабаліну та 9,0% для мефедрону. Кластер 1 є найбільш стриманим за більшістю показників: для амфетаміну частка вживання за останні 30 днів становить 10,5%, для MDMA — 19,4%, для мефедрону — 4,1%, для прегабаліну — 12,0%, а для багатьох інших речовин переважає відповідь «ніколи».

Діаграма 6. Вживання ПАР протягом 30 днів до опитування за кластерами.

## Вживання різних ПАР протягом 30 днів до опитування

Речовина	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5
Екстазі (MDMA)	19,4%	31,5%	22,9%	37,4%	53,2%
Амфетамін	10,5%	41,1%	21,8%	48,0%	58,3%
LSD	5,0%	8,9%	6,5%	16,1%	39,7%
Кокаїн	3,7%	20,1%	10,1%	25,3%	33,3%
Метамфетамін	2,2%	11,0%	6,0%	24,8%	28,2%
Кетамін	0,9%	15,4%	4,8%	19,0%	21,2%
Бутират (ГНВ/GBL)	0,0%	2,5%	0,0%	4,4%	19,9%
Мефедрон	4,1%	28,8%	9,0%	33,6%	46,8%
Альфа-PVP	0,2%	6,0%	2,3%	17,5%	25,6%
Інші солі (катіони)	0,4%	2,0%	0,2%	15,7%	23,7%
Гриби/ Псилоцибін	7,6%	10,0%	8,5%	18,2%	35,9%
2С-В	0,0%	1,7%	1,6%	7,3%	15,4%
Опіати/опіоїди	0,0%	2,7%	3,8%	19,6%	19,2%
Лірика (прегабалін)	12,0%	30,4%	13,7%	49,7%	50,6%

Створено за допомогою Datawrapper

Діаграма 7. Вживання ПАР протягом життя за кластерами.

## Досвід вживання різних ПАР протягом життя

Речовина	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5
Екстазі (MDMA)	70,6%	91,7%	78,4%	94,6%	95,5%
Амфетамін	44,7%	85,1%	62,1%	93,8%	95,5%
LSD	36,9%	57,9%	54,3%	85,4%	92,9%
Кокаїн	27,9%	64,3%	43,8%	73,7%	75,0%
Метамфетамін	14,6%	46,1%	27,7%	77,7%	80,8%
Кетамін	12,0%	45,0%	24,3%	69,1%	63,5%
Бутират (GHB/GBL)	1,1%	18,3%	6,4%	40,3%	51,9%
Мефедрон	20,2%	67,1%	30,3%	80,6%	84,0%
Альфа-PVP	7,4%	30,6%	12,0%	58,5%	66,0%
Інші солі (катіони)	8,1%	15,5%	7,5%	49,5%	66,0%
Гриби/ Псилоцибін	43,7%	55,1%	53,7%	77,5%	80,1%
2С-В	4,1%	21,6%	17,5%	58,2%	57,7%
Опіати/опіоїди	4,2%	16,3%	14,2%	59,3%	60,2%
Лірика (прегабалін)	36,4%	61,0%	35,9%	85,8%	84,0%

Створено за допомогою Datawrapper

### 8.2.3. Полівживання ПАР

Для порівняння кластерів за кількістю різних психоактивних речовин, які респонденти вживали протягом останніх 30 днів – індикатор полівживання, було застосовано узагальнену лінійну модель<sup>23</sup>. Загалом відмінності між кластерами були статистично значущими<sup>24</sup>.

Оцінені середні значення показника кількості речовин, що вживалися протягом останніх 30 днів, зростали від кластера 1 до кластерів з більш вираженим ризиковим профілем. Найменшу середню кількість уживаних речовин мав кластер 1 — 0,66, далі йшли кластер 3 — 1,11 та кластер 2 — 2,12. Значно вищі значення спостерігалися у кластері 4 — 3,37, а максимальне — в кластері 5, де середнє становило 4,71 речовини за останні 30 днів. Усі парні відмінності між кластерами

<sup>23</sup> З негативно-біноміальним розподілом і log-link, оскільки змінна лічильникова з вираженою асиметрією розподілу.

<sup>24</sup> Omnibus likelihood ratio  $\chi^2(4)=593,302$ ,  $p<0,001$ ; ефект кластерної належності за Wald  $\chi^2(4)=556,282$ ,  $p<0,001$

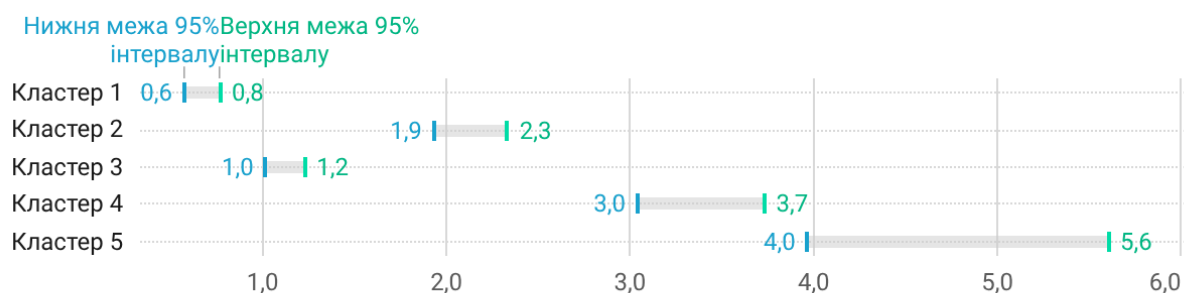
були статистично значущими, включно з різницею між кластерами 4 і 5<sup>25</sup>. Це свідчить, що сегменти розрізняються не лише за типом речовин, що вживаються а й за загальною інтенсивністю поточного полівживання.

У змістовному сенсі цей результат добре доповнює кластерну типологію: кластер 1 відповідає найнижчому рівню поточного полівживання, кластери 2 і 3 займають проміжне положення, тоді як кластери 4 і особливо 5 формують ядро найбільш інтенсивного та багатокомпонентного вживання.

Діаграма 8. Показники полівживання ПАР за кластерами.

## Кількість різних ПАР, які респонденти вживали протягом останніх 30 днів

95% довірчі інтервали оціненого за допомогою GLM-моделі середнього значення



Створено за допомогою Datawrapper

### 8.2.4. Вживання алкоголю, одночасне вживання ПАР та алкоголю

Відмінності між кластерами добре видно і в питаннях про поєднання ПАР з алкоголем та вживання алкоголю загалом. Регулярне одночасне вживання алкоголю з іншими психоактивними речовинами найчастіше трапляється у кластерах 5 і 4: 52,6% та 48,0% відповідно відповіли, що це відбувається майже кожного разу, коли вживаються ПАР. У кластері 2 така практика теж поширена — 41,1%, тоді як у кластері 3 цей показник становить 16,8%, а в кластері 1 ще менше — 7,2%. Загалом зв'язок між кластерною належністю та частотою одночасного вживання алкоголю з іншими ПАР є статистично значущим<sup>26</sup>.

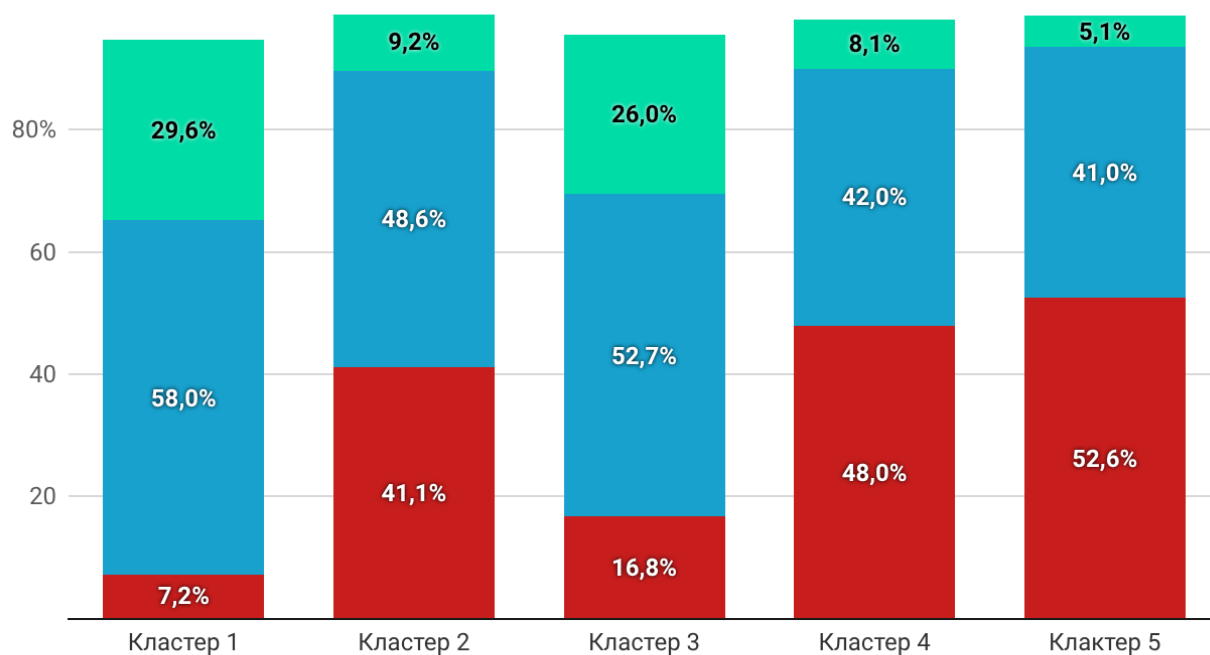
<sup>25</sup>  $p=0,003$

<sup>26</sup>  $\chi^2(12)=409,484$ ;  $p<0,001$ , Cramer's  $V=0,232$

Діаграма 9. Практики одночасного вживання алкоголю та ПАР за кластерами.

### Протягом останніх 6 місяців чи траплялося тобі одночасно (або впродовж однієї сесії/вечірки) вживати алкоголь та інші психоактивні речовини?

- Ні, ніколи
- Так, іноді / рідко
- Так, регулярно (майже кожен раз, коли вживаю психоактивні речовини)



Створено за допомогою Datawrapper

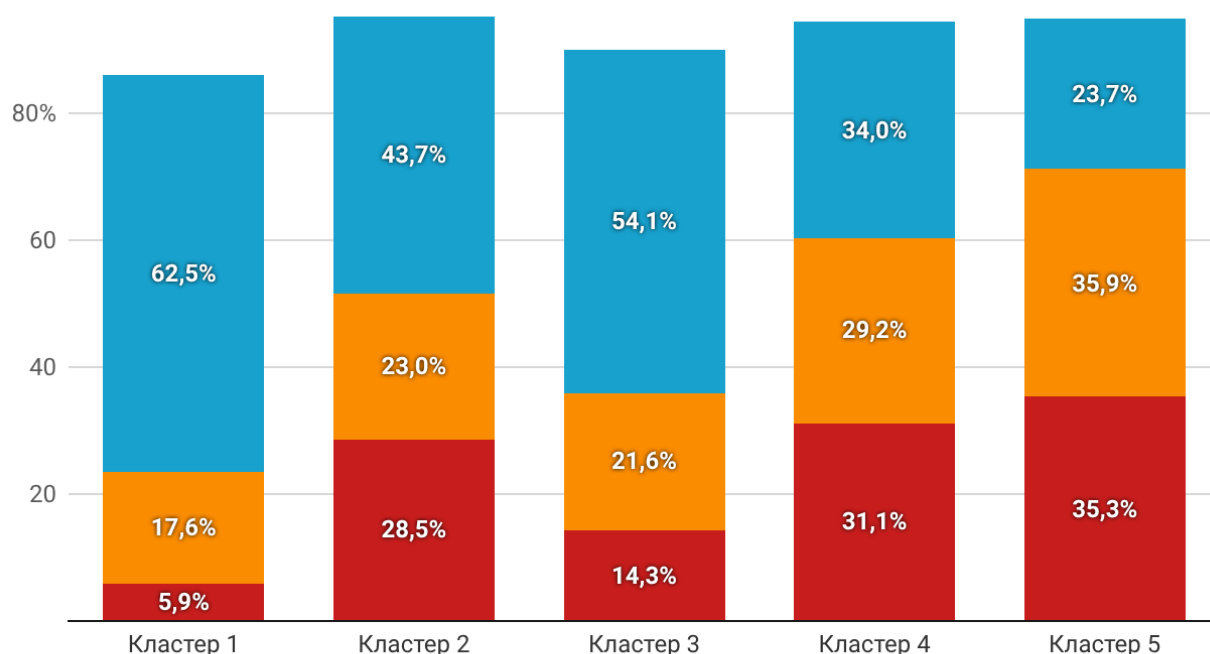
Також представники кластерів 4 і 5 частіше повідомляють, що під час вживання ПАР вони вживають більше алкоголю, ніж зазвичай: варіант «так, майже завжди» обрали 31,1% і 35,3% відповідно<sup>27</sup>.

<sup>27</sup>  $\chi^2(12)=253,628$ ;  $p<0,001$ , Cramer's  $V=0,183$

Діаграма 10. Вплив вживання ПАР на кількість алкоголю, що споживається.

## Чи помічав/ла ти, що при вживанні психоактивних речовин ти вживаєш більшу кількість алкоголю, ніж зазвичай?

■ Ні, кількість алкоголю не змінюється  
■ Іноді  
■ Так, майже завжди



Створено за допомогою Datawrapper

ВООЗ вважає, що епізодичне надмірне вживання алкоголю («heavy episodic drinking», HED) - визначене як споживання щонайменше 60 грам чистого алкоголю принаймні один раз протягом попередніх 30 днів - є найкращим показником споживання алкоголю, пов'язаного із несприятливими наслідками для здоров'я<sup>28</sup>. Відповідно, респонденти, які в нашому дослідженні відповіли, що випивають 5 стандартних порцій алкоголю (приблизно 70 грам) 2-4 рази на місяць, 2-3 рази на тиждень і 4 та більше разів на тиждень підпадають під це визначення ВООЗ. Це кожен другий респондент нашої вибірки (50,7%)<sup>29</sup>.

Частка респондентів, які підпадають під поріг епізодичного надмірного вживання алкоголю за визначенням ВООЗ, статистично значуще відрізняється між кластерами<sup>30</sup>. Це вказує на наявність зв'язку між кластерною належністю та

<sup>28</sup> World Health Organization. Global Status Report on Alcohol and Health 2018.

<https://iris.who.int/server/api/core/bitstreams/9530de1c-1fd2-4c20-a167-ec6ba7cb00c3/content>

<sup>29</sup> Звісно, наше питання напряду не вимірює вживання алкоголю за місяць, а використовує формулювання про рік, тому це наблизений індикатор HED.

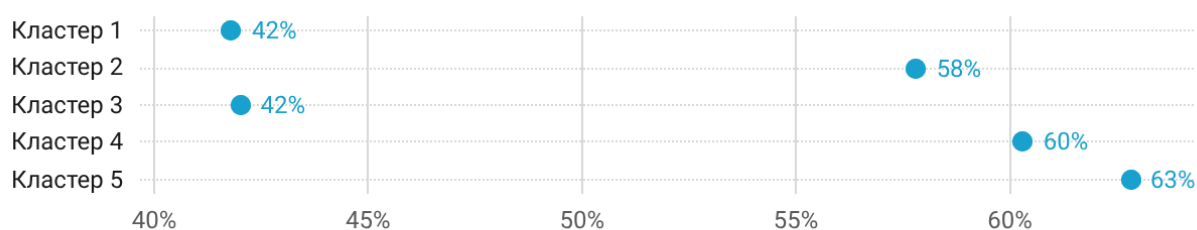
<sup>30</sup>  $\chi^2(4)=78,699$ ;  $p<0,001$ ; Cramer's  $V=0,176$

ризиковим патерном вживання алкоголю, хоча сила цього зв'язку є радше невеликою або помірною.

Найнижча частка респондентів, які підпадають під поріг епізодичного надмірного вживання спостерігається у кластерах 1 і 3 — 41,8% та 42,0% відповідно. Натомість у кластерах 2, 4 і 5 вона є суттєво вищою: 57,8%, 60,3% і 62,8%. За тестами порівняння колонкових пропорцій кластери 1 і 3 не відрізняються між собою, так само як між собою не відрізняються кластери 2, 4 і 5; водночас перша пара кластерів має статистично нижчі частки таких респондентів, ніж друга група.

Діаграма 11. Частка респондентів, які підпадають під поріг епізодичного надмірного вживання алкоголю за кластерами.

## Частка респонденти, які підпадають під поріг епізодичного надмірного вживання алкоголю



Створено за допомогою Datawrapper

Це добре узгоджується з загальною логікою сегментації, де саме кластери 4 і 5, а частково й кластер 2, демонструють більш виражений ризиковий поведінковий профіль.

### 8.3. Порівняння за сексуальною поведінкою

Кластери суттєво відрізняються за профілем сексуальної поведінки, однак цей блок слід інтерпретувати з урахуванням статевого складу сегментів. Частина відмінностей закономірно пов'язана з тим, що кластери 1 і 2 мають переважно жіночий склад, кластери 3 і 4 — переважно чоловічий, а кластер 5 є змішаним. Тому відмінності за статтю сексуальних партнерів, одностатевими контактами та груповими практиками відображають водночас і різний статевий склад кластерів, і реальні поведінкові розбіжності всередині жіночої та чоловічої підвибірок.

#### 8.3.1. Груповий секс

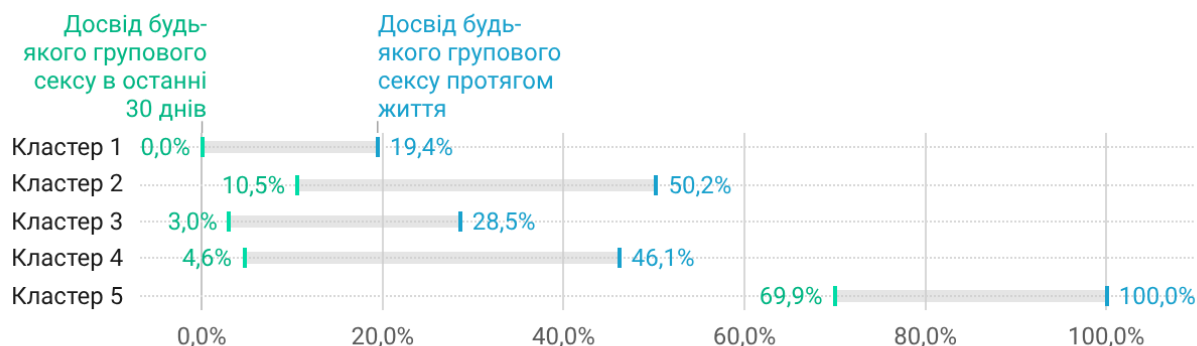
Поширеність досвіду групового сексу суттєво відрізняється між кластерами. Найнижчі показники має кластер 1: лише 19,2% респондентів повідомили про будь-який досвід групового сексу протягом життя, а про такий досвід в останні 30 днів не повідомив ніхто. Кластер 3 також характеризується відносно нижчим рівнем залученості до групового сексу, ніж кластери 2 і 4: відповідні частки становлять 28,5% протягом життя і 3% протягом останніх 30 днів. Натомість у кластерах 2 і 4 досвід групового сексу є значно поширенішим: протягом життя його мали 50,2% і

46,1% відповідно, а в останні 30 днів — 10,5% і 4,6%. Найбільш різко вирізняється кластер 5: всі його представники мали досвід групового сексу протягом життя (100%), а 69,9% — протягом останніх 30 днів. Це вказує на те, що саме кластер 5 є найбільш концентрованим сегментом за показниками поточного групового сексуального ризику.

Загалом зв'язок між кластерною належністю та досвідом групового сексу є статистично значущим і досить сильним як для показника протягом життя<sup>31</sup>, так і особливо для досвіду в останні 30 днів<sup>32</sup>.

Діаграма 12. Показники досвіду групового сексу за кластерами.

## Досвід будь якого групового сексу протягом життя та в останні 30 днів



Створено за допомогою Datawrapper

Якщо аналізувати тільки респондентів, які мали груповий секс за останні 6 місяців і відповіли на запитання про його частоту, то різні кластери також статистично значуще відрізняються за кількістю таких епізодів. Узагальнена лінійна модель<sup>33</sup> показала значущий ефект кластерної належності на кількість епізодів групового сексу за останні пів року<sup>34</sup>. Найвищу оцінену частоту має кластер 5 — у середньому 7,01 епізодів за останні 6 місяців, що статистично значуще перевищує показники всіх інших кластерів. Найнижчий показник має кластер 1 — 1,57 епізодів. Кластери 2, 3 і 4 займають проміжне положення і між собою статистично не відрізняються. Це свідчить, що кластер 5 вирізняється не лише більшою поширеністю групового сексу, а й значно вищою повторюваністю такого досвіду<sup>35</sup>.

<sup>31</sup>  $\chi^2(4)=454,316$ ;  $p<0,001$ ; Cramer's  $V=0,423$

<sup>32</sup>  $\chi^2(4)=752,365$ ;  $p<0,001$ ; Cramer's  $V=0,544$

<sup>33</sup> З негативно-біноміальним розподілом і log-link, оскільки змінна лічильникова з вираженою асиметрією розподілу.

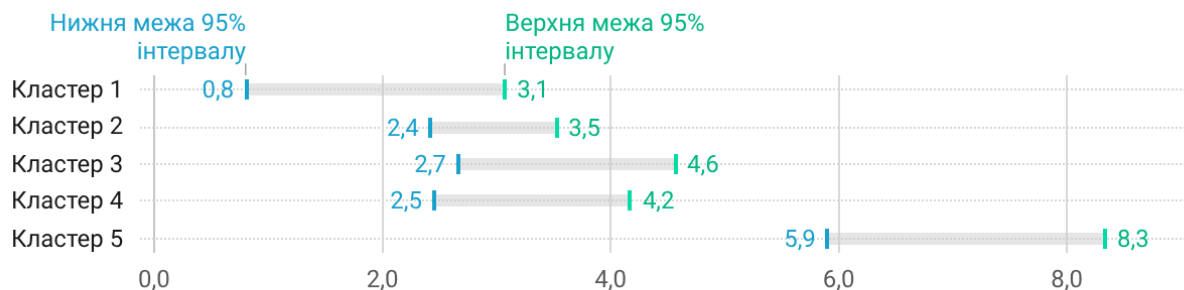
<sup>34</sup> Likelihood ratio  $\chi^2(4)=64,835$ ;  $p<0,001$

<sup>35</sup> Водночас результати слід інтерпретувати обережно з огляду на технічне попередження моделі про неповну збіжність.

Діаграма 13. Кількість епізодів групового сексу за 6 місяців за кластерами.

## Кількість епізодів групового сексу за 6 місяців

Серед респондентів, які практикували груповий секс. 95% довірчі інтервали оціненого за допомогою GLM-моделі середнього значення.



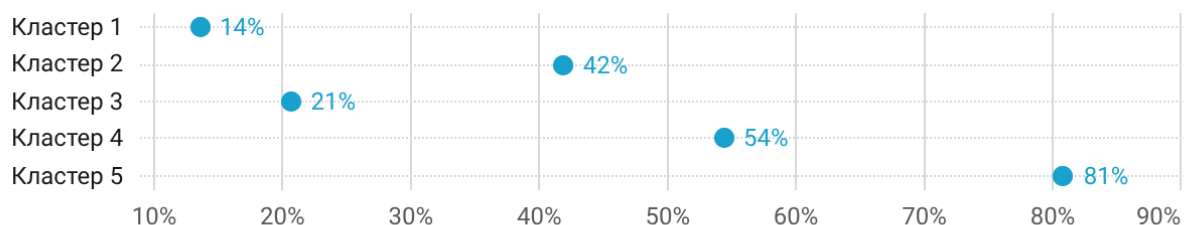
Створено за допомогою Datawrapper

### 8.3.2. Вживання ПАР під час останнього сексу

Поширеність практики вживання психоактивних речовин (крім алкоголю та канабісу) перед або під час останнього сексу суттєво відрізняється між кластерами<sup>36</sup>. Найнижча частка респондентів із таким досвідом зафіксована у кластері 1 — 13,5%. Порівняно низький показник має і кластер 3 — 20,6%. У кластері 2 частка тих, хто вживав ПАР перед або під час останнього сексу, є значно вищою і становить 41,8%. Ще вищий рівень спостерігається у кластері 4 — 54,3%. Найбільше вирізняється кластер 5: у ньому про вживання ПАР перед або під час останнього сексу повідомили 80,8% респондентів. Це вказує на те, що сексуалізоване вживання ПАР є особливо характерним саме для кластерів 4 і 5, а найбільш концентрованим — для кластера 5.

Діаграма 14. Практика сексуалізованого вживання ПАР за кластерами.

## Вживання ПАР під час останнього сексу (частка відповідей "так")



Створено за допомогою Datawrapper

Поширеність використання альфа-PVP, мефедрону або інших «солей» у сексуальному контексті суттєво відрізняється між кластерами. Найнижчі показники

<sup>36</sup>  $\chi^2(4)=81,616$ ;  $p<0,001$ , Cramer's  $V=0,179$

мають кластери 1 і 3 — 1,5% та 3,9% відповідно. У кластерах 2 і 4 така практика трапляється помітно частіше — 14,3% і 20,3%. Найбільше вирізняється кластер 5, де про використання цих речовин в сексуальному контексті повідомили 36,5% респондентів. Загалом зв'язок між кластерною належністю та використанням альфа-PVP, мефедрону або інших «солей» є статистично значущим і помірним за силою<sup>37</sup>. Кластер 5 є основним сегментом концентрації сексуалізованого вживання «солей», тоді як кластери 2 і 4 займають проміжне положення.

### 8.3.3. Гомосексуальні контакти

Поширеність одностатевих сексуальних контактів суттєво відрізняється між кластерами як серед жінок, так і серед чоловіків. Серед жінок найнижчі показники зафіксовані у кластері 1: 30,4% мали такий досвід протягом життя і 4,1% — в останні 30 днів. У кластерах 2, 3 і 4 частка жінок з одностатевим досвідом протягом життя є помітно вищою і становить відповідно 52,4%, 58,9% та 56,3%, а частка недавнього досвіду — 8,2%, 24,4% і 14,1%. Найбільш чітко вирізняється кластер 5: 95,9% жінок у ньому мали одностатеві контакти протягом життя, а 56,2% — в останні 30 днів. Відмінності між жіночими підгрупами кластерів є статистично значущими як для досвіду протягом життя<sup>38</sup>, так і для досвіду в останні 30 днів<sup>39</sup>.

Діаграма 15. Досвід гомосексуальних стосунків серед жінок за кластерами.

## Досвід гомосексуальних стосунків серед жінок ("останні 30 днів" - "впродовж життя")



Створено за допомогою Datawrapper

Серед чоловіків картина ще контрастніша. У кластерах 1 і 2 одностатеві контакти серед чоловіків фактично є нормою: про такий досвід протягом життя повідомили 96,0% і 96,1% відповідно, а в останні 30 днів — 72,0% і 82,8%. Натомість кластер 3 різко відрізняється: лише 23,4% чоловіків у ньому мали такий досвід протягом життя, а 11,4% — в останні 30 днів. Кластер 4 займає проміжне положення (31,1% і 14,5%), тоді як кластер 5 знову характеризується високими показниками: 66,3% протягом життя і 57,8% в останні 30 днів. Відмінності між чоловічими

<sup>37</sup>  $\chi^2(4)=232,868$ ;  $p<0,001$ ; Cramer's  $V=0,303$

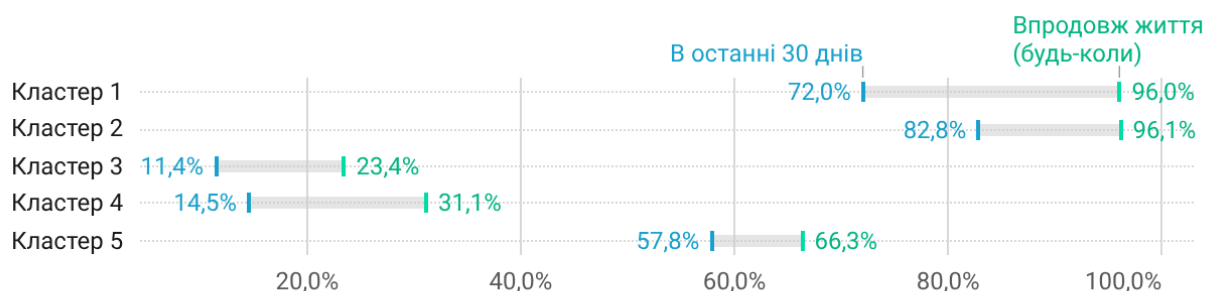
<sup>38</sup>  $\chi^2=132,676$ ;  $p<0,001$ ;  $V=0,327$

<sup>39</sup>  $\chi^2=188,125$ ;  $p<0,001$ ;  $V=0,389$

підгрупами кластерів також є статистично значущими і сильнішими, ніж серед жінок<sup>40</sup>.

Діаграма 16. Досвід гомосексуальних стосунків серед чоловіків за кластерами.

## Досвід гомосексуальних стосунків серед чоловіків ("останні 30 днів" - "впродовж життя")



Створено за допомогою Datawrapper

Зв'язок між кластерною належністю та одностатевими сексуальними контактами суттєво залежить від статі респондента. Для чоловіків кластери 1 і 2 концентрують гомосексуальний профіль (але частка чоловіків в цих кластерах дуже мала), тоді як кластер 3 є переважно гетеросексуальним. Серед жінок розподіл більш рівномірний між кластерами 2–4, але кластер 5 так само формує найбільш виражене ядро одностатевого досвіду, особливо недавнього. Водночас інтерпретацію окремих статево-кластерних підгруп слід робити обережно там, де кількість респондентів відповідної статі є невеликою, зокрема для чоловіків у кластері 1 та жінок у кластері 3 і 5.

### 8.3.4. Транзакційний секс

Показник поширеності досвіду транзакційного сексу аналізувався серед підвибірки респондентів, яким ставилося це запитання<sup>41</sup>. Загалом відмінності між кластерами є статистично значущими<sup>42</sup>. Найнижча частка респондентів, які повідомили про транзакційний секс (коли вони отримували винагороду за сексуальний контакт) протягом останніх 12 місяців, зафіксована в кластері 3 — 5,2% (один із найменш ризикових і найбільш чоловічих кластерів). Дещо вищий показник має кластер 1 — 9,6% (переважно жіночий і низько ризиковий кластер). У кластерах 2 і 4 частка таких респондентів є помітно більшою — 23,3% і 20,8% відповідно. Найбільш сильно вирізняється кластер 5: у ньому про участь в транзакційному сексі протягом останнього року повідомили 52,0% респондентів.

Якщо розглядати цей показник окремо за статтю, закономірності зберігаються, але стає видно певні відмінності в інтенсивності показника. Серед

<sup>40</sup> Для досвіду протягом життя  $\chi^2=322,495$ ;  $p<0,001$ ;  $V=0,498$ , для досвіду в останні 30 днів  $\chi^2=413,584$ ;  $p<0,001$ ;  $V=0,564$

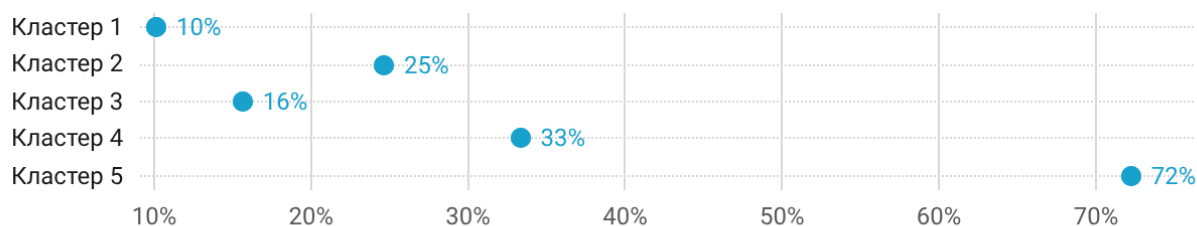
<sup>41</sup>  $n=895$ , або 35,2% усієї аналітичної вибірки

<sup>42</sup>  $\chi^2(8)=91,960$ ,  $p<0,001$ ; Cramer's  $V=0,227$

жінок транзакційний секс найрідше трапляється в кластері 1 (10,1%) і найчастіше — в кластері 5 (72,2%); у кластерах 2, 3 і 4 відповідні частки становлять 24,6%, 15,6% і 33,3%. Відмінності між жіночими підгрупами кластерів є статистично значущими<sup>43</sup>.

Діаграма 17. Досвід транзакційного сексу протягом останніх 12 місяців серед жінок за кластерами.

## Досвід транзакційного сексу за останні 12 місяців серед жінок

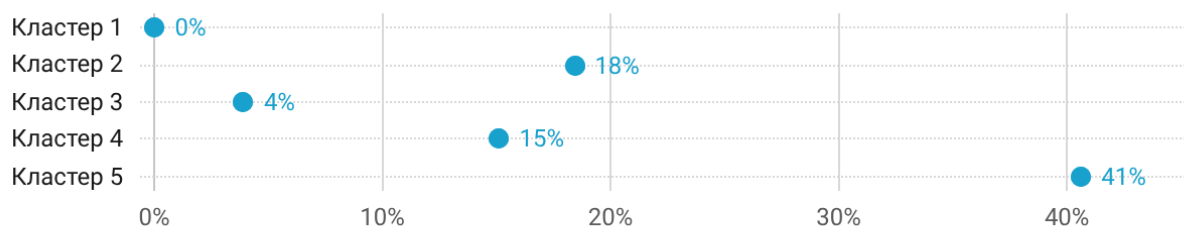


Створено за допомогою Datawrapper

Серед чоловіків показник також значуще відрізняється між кластерами<sup>44</sup>. Найнижчі значення спостерігаються в кластерах 1 і 3 (0,0% та 3,9%), проміжні — в кластерах 2 і 4 (18,4% та 15,1%), а найвищий показник знову має кластер 5 — 40,6%.

Діаграма 18. Досвід транзакційного сексу протягом останніх 12 місяців серед чоловіків за кластерами.

## Досвід транзакційного сексу за останні 12 місяців серед чоловіків



Створено за допомогою Datawrapper

Як видно, поширеність практики транзакційного сексу не є рівномірно розподіленою між сегментами. Такі респонденти концентруються в кластері 5, а також помітно частіше трапляються в кластерах 2 і 4, ніж у кластерах 1 і 3. Водночас інтерпретацію окремих статево-кластерних підгруп варто робити обережно там, де підвибірki відповідних груп є невеликими, зокрема для чоловіків у кластері 1 та жінок у кластері 5.

<sup>43</sup>  $\chi^2(8)=47,399$ ,  $p<0,001$ ; Cramer's  $V=0,231$

<sup>44</sup>  $\chi^2(8)=55,679$ ,  $p<0,001$ ; Cramer's  $V=0,248$

#### 8.4. Тестування на ВІЛ та ВІЛ-статус, тестування на інші ІПСШ

За показниками досвіду тестування на ВІЛ кластери відрізняються статистично значуще, але сила цих відмінностей є невеликою<sup>45</sup>. Частка респондентів, які коли-небудь проходили тестування на ВІЛ, коливається від 59,5% у кластері 1 та 59,8% у кластері 3 до 71,9% у кластері 2 і 69,9% у кластері 4; у кластері 5 цей показник становить 62,2%. Як видно, більш ризикові з точки зору поведінки кластери демонструють вищу поширеність досвіду тестування на ВІЛ.

Схожа картина спостерігається і щодо давності останнього тестування на ВІЛ. У кластерах 2, 4 і 5 останнє тестування загалом є дещо більш «свіжим»: сукупна частка тих, хто тестувався протягом останніх 6 місяців, становить 42,1%, 37,1% і 48,5% відповідно. У кластерах 1 і 3 вона нижча — 30,7% і 33,4%, а частка тих, хто тестувався більше 12 місяців тому, навпаки, вища: 47,3% у кластері 1 та 40,8% у кластері 3. Водночас і тут сила зв'язку залишається низькою<sup>46</sup>, тому доречніше говорити про помірні відмінності в регулярності тестування, ніж про різко відокремлені профілі.

Серед респондентів, які повідомили результат свого останнього тестування на ВІЛ, статистично значущі відмінності між кластерами є, але їхня сила є дуже невеликою<sup>47</sup>. Додатково слід враховувати, що кількість респондентів, які повідомили ВІЛ-позитивний результат, є малою (n=19).

Найвищі частки респондентів, які повідомили ВІЛ-позитивний результат останнього тесту, спостерігаються в кластерах 5 і 2. Якщо рахувати від усього складу кластерів, це 1,9% у кластері 5 та 1,6% у кластері 2; серед протестованих — 3,1% і 2,2% відповідно. У кластерах 3 і 4 ці частки нижчі, а в кластері 1 відповідей про позитивний результат тестування на ВІЛ не зафіксовано. За тестами порівняння колонкових пропорцій статистично підтверджується насамперед різниця між кластером 1 і кластером 5; для інших пар кластерів окремі відмінності надійно не підтверджені.

---

<sup>45</sup> Cramer's V = 0,084; p < 0,001

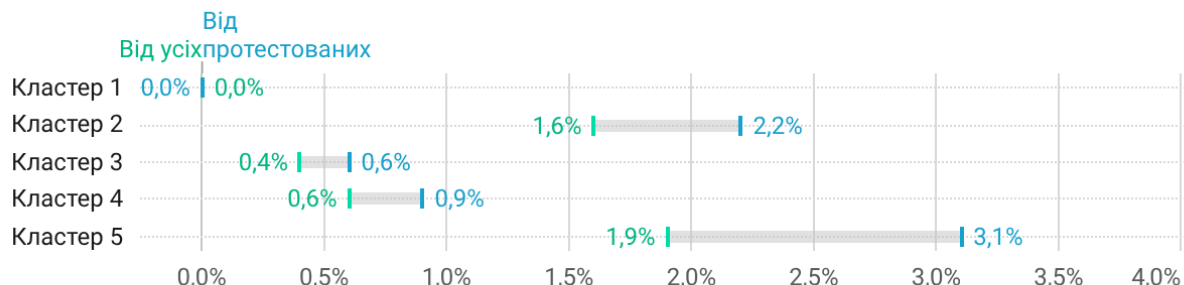
<sup>46</sup> Cramer's V = 0,069; p = 0,01

<sup>47</sup>  $\chi^2(16)=33,526$ ; p=0,006; Cramer's V=0,071

Діаграма 19. Результати тестування на ВІЛ за кластерами.

## Частка тих, хто повідомив про позитивний результат ВІЛ-тестування

(від усіх, та від тих, хто проходив тестування протягом життя)



Створено за допомогою Datawrapper

Тестування на інші ІПСШ та вірусні гепатити показує дещо більш строкату, але також не надто контрастну картину. Показник «не тестувався/лася на жодне з цих захворювань протягом останніх 6 місяців» статистично відрізняється між кластерами: така відповідь частіше трапляється в кластерах 1 (72,3%) і 3 (71,2%), рідше — у кластерах 2 (64,9%), 4 (62,7%) і 5 (66,4%)<sup>48</sup>. Це дає підстави припускати, що респонденти з кластерів 2 і 4 дещо частіше включені в практики регулярного тестування на ІПСШ та гепатити, як і випадку тестування на ВІЛ. За окремими інфекціями найбільш виразні відмінності спостерігаються для тестування на сифіліс: 28,0% у кластері 5, 22,9% у кластері 2, 19,6% у кластері 4 проти 16,6–16,7% у кластерах 1 і 3; зв'язок є статистично значущим, хоча й слабким<sup>49</sup>. Для гепатиту В (17,9–24,2%) і гепатиту С (16,9–24,5%) відмінності між кластерами не досягають статистичної значущості. Тестування на інші ІПСШ також розподіляється без різких контрастів: частки становлять 9,1% у кластері 1, 12,7% у кластері 2, 8,2% у кластері 3, 11,6% у кластері 4 та 14,0% у кластері 5. Загалом цей блок показує, що кластери 2, 4 і 5 дещо більше залучені до тестування на ІПСШ, але масштаб цих відмінностей залишається помірним.

<sup>48</sup>  $\chi^2(4)=11,129$ ;  $p=0,025$ ; Cramer's  $V=0,082$

<sup>49</sup> Cramer's  $V = 0,086$ ;  $p = 0,015$

## 8.5. Потреби в психологічній, психіатричній, медичній допомозі та звернення за послугами

### 8.5.1. Психологічна та психіатрична допомога

Наявність потреб у психологічній та психіатричній допомозі відрізняється між кластерами, хоча масштаб цих відмінностей є не дуже значним<sup>50</sup>. Потребу в допомозі психолога або психотерапевта протягом останніх 6 місяців найчастіше зазначали респонденти з кластерів 1 і 2 — 73,6% та 69,3% серед тих, хто відповів на це запитання (нагадаємо, що це найбільш «жіночі» кластери). У кластерах 3 (найбільш «чоловічий») і 4 цей показник нижчий — 52,1% і 56,0%, а в кластері 5 становить 46,7%. Потреба в допомозі психіатра також є найвищою в кластері 2 (52,7%), високою в кластерах 1, 4 і 5 (47,6%, 46,5% і 47,7%) і найнижчою в кластері 3 (32,6%). Як видно, кластери 1 і 2 мають найбільш виражений запит на психоемоційну підтримку, тоді як кластер 3 виглядає стриманішим і за цим показником.

Фактичне звернення по психологічну або психіатричну допомогу розподіляється менш контрастно, ніж сама потреба. Якщо об'єднати всі форми звернення, то найвищі частки спостерігаються в кластерах 1, 2 і 5: 42,6%, 40,3% і 40,2% відповідно. У кластері 4 будь-який досвід звернення має 30,0% респондентів, у кластері 3 — 25,7%. Очевидно, кластер 3 вирізняється не лише меншим запитом на психологічну й психіатричну допомогу, а й нижчою фактичною залученістю до таких послуг. Зв'язок тут теж статистично значущий, але невеликий<sup>51</sup>.

Серед тих, хто відчував потребу в психологічній або психіатричній допомозі, але не звертався по неї, найпоширенішим бар'єром у всіх кластерах є брак грошей або висока вартість послуг: від 54,1% у кластері 4 до 71,2% у кластері 2. Частина потенційних бар'єрів майже не відрізняється між сегментами — зокрема труднощі з транспортом, нестача часу для очного звернення чи невизначеність, куди саме звертатися. Натомість недовіра до психологів або психіатрів частіше фіксується в кластері 4 (27,9%), ніж в інших сегментах, а скепсис щодо ефективності такої допомоги найвиразніший у кластерах 4 і 5 (36,1% і 36,0%). Обидва ці бар'єри мають статистично значущі, хоча й помірні, відмінності між кластерами<sup>52</sup>. У підсумку найбільш напружений профіль бар'єрів до психологічної допомоги спостерігається в кластерах 4 і 5.

### 8.5.2. Медична допомога

Потреба в медичній допомозі також розподілена між сегментами нерівномірно. Найчастіше про неї повідомляли респонденти з кластерів 2 і 1 — 52,9% і 49,3%; дещо нижчі показники мають кластери 4 і 5 — 47,4% і 45,8%; найнижчий показник знову спостерігається в кластері 3 — 35,6%. Фактичне звернення по медичну допомогу найчастіше траплялося в кластерах 1 і 2 (56,8% і

<sup>50</sup> Для психологічних послуг: Cramer's V = 0,140; p < 0,001; для психіатричних Cramer's V = 0,131; p < 0,001.

<sup>51</sup> Cramer's V = 0,101; p < 0,001

<sup>52</sup> Для недовіри Cramer's V = 0,175; p = 0,002; для відповіді «не відчуваю, що це реально допоможе» Cramer's V = 0,165; p = 0,004.

54,9%), рідше — в кластерах 4 і 5 (45,9% і 47,7%), а найнижче — в кластері 3 (41,4%). Кластер 3 і в медичному блоці виглядає як сегмент із найнижчим рівнем як потреби, так і звернення по допомогу. Загальні зв'язки між кластерною належністю та потребою/зверненням по медичну допомогу є статистично значущими, але слабкими<sup>53</sup>.

Кластери відрізняються і за типом надавачів послуг, яким користувалися респонденти. Кластери 1 і 2 частіше користувалися приватною медициною: 69,0% і 71,1% відповідно зверталися до приватної клініки або приватного лікаря. У кластерах 3, 4 і 5 ця частка нижча — 57,9%, 52,0% і 47,1%<sup>54</sup>. Натомість державні клініки частіше використовують кластери 3, 4 і 5: 57,0%, 58,7% і 64,7% відповідно, тоді як у кластерах 1 і 2 цей показник становить 42,9% і 44,4%<sup>55</sup>. Звернення до неурядових організацій або мобільних сервісів трапляється рідко в усіх сегментах і не формує виразної осі диференціації.

Серед тих, хто потребував медичної допомоги, але не звертався по неї, відмінності між кластерами загалом менш виражені. Брак грошей або висока вартість послуг залишається найчастішою перешкодою майже в усіх сегментах, статистично значущих розбіжностей тут немає. Більшість інших бар'єрів — незнання, куди звертатися, недовіра до державних закладів, сумніви щодо ефективності допомоги, транспортні труднощі чи нестача часу — теж не показують чіткої кластерної диференціації. Єдиний бар'єр, який відрізняється статистично значуще, — страх осуду або стигми: він частіше трапляється в кластерах 4 і 5 (26,8% і 38,5%) і рідше — в кластерах 1 і 2 (11,6% і 10,9%). Це вказує на те, що для більш уразливих сегментів, особливо кластерів 4 і 5, доступ до медичних послуг додатково ускладнюється стигмою.

Узагальнено блок потреб і послуг показує таку конфігурацію: кластери 1 і 2 мають найбільш виражений запит на психологічну підтримку; кластер 3 вирізняється найнижчим рівнем і потреби, і звернення по допомогу; кластери 4 і 5 не завжди мають найвищий декларований попит, але демонструють більш напружений профіль бар'єрів, зокрема недовіру, сумнів у користі допомоги та стигму. У медичному блоці кластери 1–2 частіше користуються приватними послугами, тоді як кластери 3–5 частіше звертаються до державних закладів. Як видно, відмінності між сегментами проявляються не лише в рівні потреб, а й у способах взаємодії з системою допомоги та в типових бар'єрах доступу до неї.

---

<sup>53</sup> Для потреби Cramer's V = 0,115; p < 0,001; для звернень Cramer's V = 0,113; p < 0,001

<sup>54</sup>  $\chi^2(4)=24,250$ ; p<0,001; Cramer's V=0,172

<sup>55</sup>  $\chi^2(4)=18,839$ ; p=0,001; Cramer's V=0,152

## 8.6. Вживання ПАР соціальним оточенням респондента, спайкінг, небажані прояви і проблеми, пов'язані із вживання ПАР

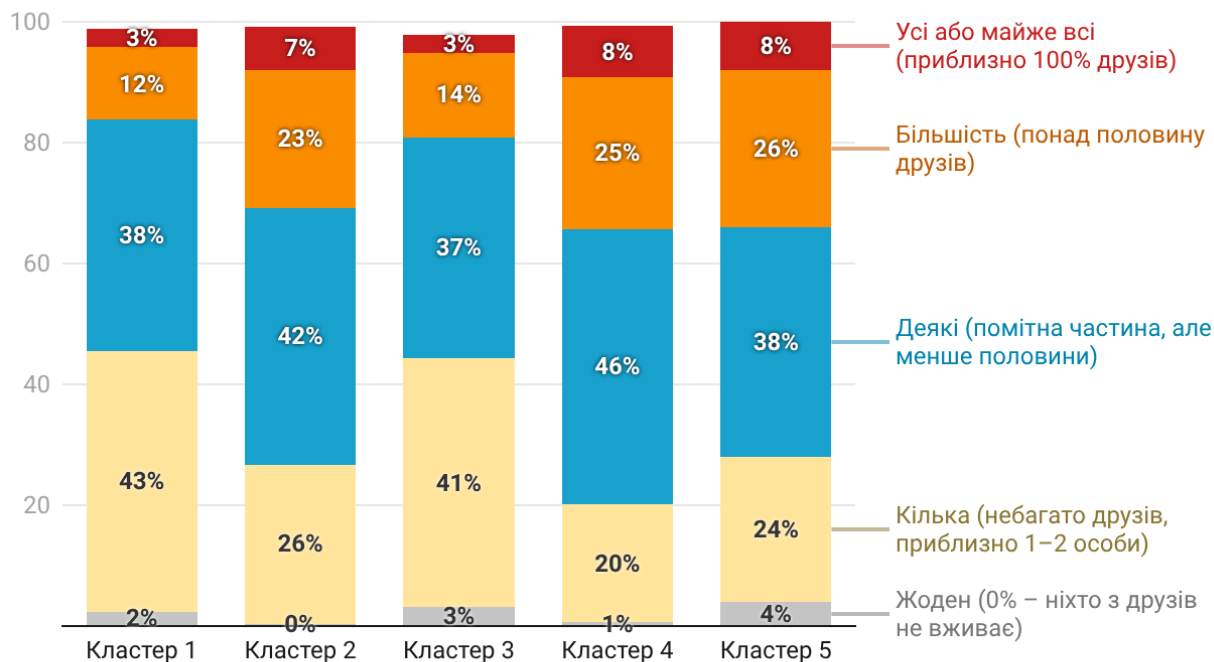
### 8.6.1. Соціальне оточення

Кластери помітно відрізняються за наявністю соціального оточення, яке практикує вживанням психоактивних речовин. Частка друзів, які теж вживають ПАР, статистично значуще пов'язана з кластерною належністю, сила цього зв'язку є помірною<sup>56</sup>.

Найбільш «насиченими» споживачами ПАР виглядають соціальні мережі в кластерах 4 і 5. У кластері 4 25,3% респондентів зазначили, що психоактивні речовини вживає більшість їхніх друзів, а ще 8,4% — що майже всі; у кластері 5 відповідні частки становлять 26,0% і 8,0%. У кластері 2 цей показник також є підвищеним: 22,9% повідомили, що вживає більшість друзів, ще 7,2% — що майже всі. Натомість у кластерах 1 і 3 частіше трапляється відповідь, що таких друзів лише кілька: 43,1% і 41,3% відповідно. Отже, кластери 4 і 5 характеризуються більш щільним включенням у соціальне середовище, де вживання ПАР є звичною практикою, тоді як у кластерах 1 і 3 такі мережі контактів зустрічаються рідше.

Діаграма 20. Вживання ПАР друзями респондента, за кластерами.

### Яка частка друзів респондента вживає ПАР



Створено за допомогою Datawrapper

<sup>56</sup> Cramer's V = 0,220; p < 0,001; цей показник доступний лише для 895 респондентів, або 35,2% аналітичної вибірки, тому його варто інтерпретувати обережно.

### 8.6.2. Спайкінг

Досвід спайкінгу нерівномірно розподілений між кластерами<sup>57</sup>. Найвища частка респондентів, які повідомили, що з ними траплялися випадки спайкінгу протягом останніх 6 місяців або раніше, зафіксована в кластері 5 — 46,2%. Далі йдуть кластер 4 — 22,8% і кластер 2 — 21,0%. У кластері 1 цей показник становить 14,6%, а в кластері 3 — 9,9%, що є найнижчим значенням серед усіх сегментів. Таким чином, кластер 5 різко виділяється як найбільш уразливий до спайкінгу, тоді як кластери 3 і 1 виглядають менш зачепленими цим досвідом.

### 8.6.3. Небажані прояви та проблеми, пов'язані із вживанням ПАР

Наявність досвіду небажаних проявів після вживання ПАР також демонструє чітку кластерну диференціацію<sup>58</sup>. Найчастіше про небажані прояви протягом останніх 30 днів повідомляли респонденти з кластерів 5 і 4 — 39,7% та 34,2% відповідно. У кластері 2 цей показник становить 27,7%, тоді як у кластерах 3 і 1 він нижчий — 13,7% і 10,0%. За сукупністю недавніх небажаних проявів — упродовж останніх 30 днів або останніх 6 місяців — найбільш навантаженими знову виглядають кластери 5, 4 і 2.

Частка тих, хто ніколи не мав небажаних проявів після вживання ПАР, є найвищою в кластерах 3 і 1 — 37,9% і 31,8%. У кластері 2 вона нижча і становить 19,0%, а в кластерах 4 і 5 — лише 13,2% і 11,5%. Це означає, що кластери 4 і 5 вирізняються не лише інтенсивнішим і ризикованішим профілем вживання, а й більшою накопиченою частотою негативних фізичних або психоемоційних реакцій після нього. Для кластера 2 характерний проміжний, але теж досить напружений профіль, тоді як кластери 1 і 3 виглядають менш проблемними за цим показником.

Для оцінки відмінностей між кластерами за загальною кількістю типів проблем, що виникли через вживання психоактивних речовин, було використано два підходи тест Kruskal–Wallis та тест різниці медіан. Обидва непараметричні тести показали статистично значущі відмінності між кластерами<sup>59</sup>.

Для наочнішої інтерпретації було також було побудовано узагальнену лінійну модель, вона підтвердила асоціацію кластерної належності та кількості типів проблем, викликаних вживання ПАР<sup>60</sup>. Найнижче оцінене середнє число типів проблем має кластер 1 (0,97), далі йде кластер 3 (1,20), потім кластер 2 (1,81). Найвищі значення спостерігаються у кластерах 4 і 5 — 2,53 та 2,46 відповідно.

Парні порівняння показали, що всі кластери статистично значуще відрізняються один від одного за цим показником, окрім пари кластерів 4 і 5: між ними різниці не виявлено<sup>61</sup>. Кластери 4 і 5 формують групу з найбільш вираженим проблемним профілем, кластер 2 займає проміжне положення, а кластери 1 і 3 є відносно менш обтяженими проблемами, причому кластер 1 має найнижчий рівень.

<sup>57</sup> Cramer's V = 0,230; p < 0,001

<sup>58</sup> Cramer's V = 0,150; p < 0,001

<sup>59</sup> Kruskal–Wallis H=305,301; p<0,001, median test  $\chi^2=221,250$ ; p<0,001.

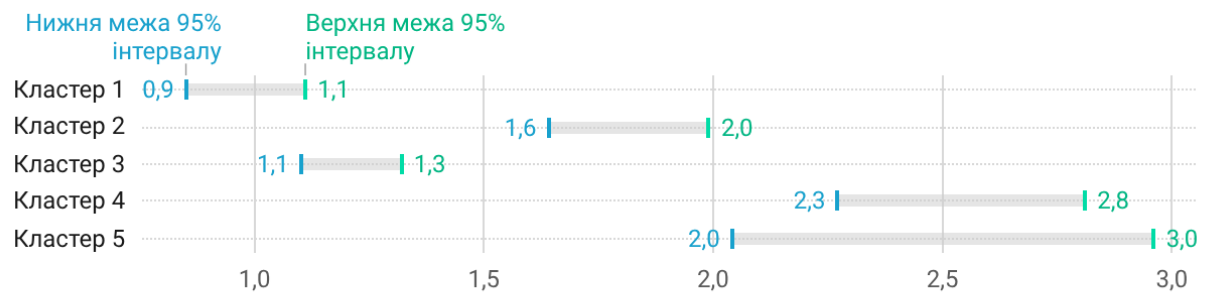
<sup>60</sup> Likelihood ratio  $\chi^2(4)=193,649$ ; p<0,001; Wald  $\chi^2(4)=189,744$ ; p<0,001

<sup>61</sup> Ані в Kruskal–Wallis pairwise comparisons, ані в NB-GLM (p=0,811)

Діаграма 21. Кількість типів проблем, викликаних вживанням ПАР за кластерами.

## Кількість типів проблем, викликаних вживанням ПАР, про які повідомили респонденти

95% довірчі інтервали оціненого за допомогою GLM-моделі середнього значення



Створено за допомогою Datawrapper

Цей блок добре доповнює загальну типологію. Кластери 4 і 5 мають не лише інтенсивніше вживання, а й складніше соціальне середовище та більший тягар негативних наслідків. Кластер 2 займає проміжну позицію, а кластери 1 і 3 виглядають менш навантаженими.

## 9. Профілювальний алгоритм для попереднього віднесення клієнта до сегмента

### 9.1. Метод побудови алгоритму

Після побудови кластерної моделі перед нами постало практичне завдання: як використати отриману сегментацію не лише для аналітичного опису, а й у реальній роботі з новими клієнтами. Кластерний аналіз добре підходить для виявлення структури в даних, однак сам по собі не є зручним інструментом для оперативного віднесення кожного нового респондента до вже описаного сегмента. Для прикладного використання сегментації було побудовано окремий класифікаційний механізм, який на основі відповідей респондента дозволяє оцінити, до якого з виділених кластерів найбільш імовірно належить новий потенційний клієнт. Саме цю функцію в нашому дослідженні виконує дискримінантний аналіз<sup>62</sup>.

Дискримінантний аналіз має лінійну й прозору логіку. Для кожного респондента на основі його відповідей обчислюються значення дискримінантних функцій, після чого алгоритм визначає, до якого сегмента він є найближчим. У нашому випадку цей підхід уже реалізовано в тестовій Google-таблиці: на основі лінійних формул система автоматично розраховує потрібні значення та визначає найбільш імовірний сегмент без додаткового ручного введення з боку менеджера. Це дозволяє використовувати результати статистичного моделювання в прикладному робочому інструменті.

Практична цінність такого рішення пов'язана з його простотою, прозорістю та відтворюваністю. Формули легко оновити в разі зміни моделі. Водночас алгоритм дає змогу автоматично профілювати нові випадки майже одразу після завершення анкети, що є важливим для повсякденної проектної роботи, де потрібно швидко перейти від відповідей респондента до підбору релевантних форматів підтримки, консультування або маршрутизації послуг.

### 9.2. Вхідні змінні алгоритму

До профілювального алгоритму було включено той самий набір змінних, який використовувався для побудови кластерної моделі. Це забезпечує змістовну й технічну узгодженість між кластеризацією та подальшим профілюванням нових випадків: алгоритм класифікує респондентів за тією самою логікою, за якою раніше було сформовано самі сегменти. Усі змінні використовувалися в тих самих кодуваннях і в тій самій аналітичній логіці, що й у кластерному аналізі<sup>63</sup>.

### 9.3. Принцип роботи алгоритму

Алгоритм працює на рівні окремого респондента. На основі його відповідей обчислюються значення дискримінантних функцій; їхня кількість на одиницю менша

<sup>62</sup> <https://www.ibm.com/docs/en/spss-statistics/32.0.0?topic=analysis-discriminant-statistics>

<sup>63</sup> Включно з обробкою пропущених значень.

за кількість кластерів, тобто, 4 в нашому випадку. Кожен респондент отримує власний набір таких значень функцій, які відображають його положення в просторі, що розмежовує сегменти. Після цього алгоритм порівнює отримані значення і визначає, до якого кластера респондент є найближчим за результатами моделі.

#### 9.4. Результативність алгоритму

Дискримінантний аналіз показав добру здатність відтворювати раніше побудовану 5-кластерну типологію. У модель увійшли всі 2543 спостереження без втрат через пропущені значення. Частка правильно класифікованих випадків становила 80,7%, а за крос-валідацією — 79,7%<sup>64</sup>. Найкраще відтворювалися кластери 1, 3 і 5, де частка правильної класифікації перевищувала 80%, тоді як найскладнішим для моделі виявився кластер 4. Перші дві канонічні дискримінантні функції пояснювали 86,0% розділювальної варіації, всі чотири функції були статистично значущими. Це свідчить, що побудований профілювальний алгоритм є достатньо ефективним для прикладного автоматичного віднесення нових респондентів до найбільш імовірного сегмента.

*Таблиця 1. Основні показники профілювального алгоритму на основі дискримінантного аналізу*

<b>Показник</b>	<b>Значення</b>
Кількість спостережень у моделі	2543
Кількість кластерів	5
Кількість дискримінантних функцій	4
Частка правильно класифікованих випадків	80,7%
Частка правильно класифікованих випадків за крос-валідацією	79,7%
Частка варіації, пояснена 1-ю функцією	52,5%
Частка варіації, пояснена 2-ю функцією	33,5%
Сукупно 1-ша і 2-га функції	86,0%
Значущість функцій	усі 4 функції статистично значущі

<sup>64</sup> Крос-валідація в дискримінантному аналізі показує, наскільки стійкою є класифікаційна модель: кожен випадок класифікується на основі функцій, побудованих без урахування цього самого випадку. Близькість загальної частки правильно класифікованих випадків до результату за крос-валідацією свідчить про відносну стабільність моделі.

Таблиця 2. Якість відтворення окремих кластерів у дискримінантному аналізі

Кластер	Частка правильно класифікованих випадків
Кластер 1	83,0%
Кластер 2	79,8%
Кластер 3	85,1%
Кластер 4	66,8%
Кластер 5	80,8%
Загалом	80,7%
Загалом, крос-валідація	79,7%

#### 9.5. Методичні застереження та перевірка альтернативних статистичних алгоритмів профілювання

Під час інтерпретації результатів слід враховувати, що лінійний дискримінантний аналіз має статистичні передумови, які в наших даних виконуються не повністю. Зокрема, тест Vox's M вказував на нерівність коваріаційних матриць груп, а значна частина предикторів були порядковими. Тому результати дискримінантного аналізу доцільно трактувати як прикладний інструмент класифікації, а не як модель із повністю дотриманими класичними передумовами застосування методу.

Слід також враховувати, що профілювальний алгоритм було побудовано на тій самій аналітичній вибірці, на якій попередньо формувалася сама кластерна типологія. Використання крос-валідації частково зменшує ризик надто оптимістичної оцінки якості, однак не замінює повноцінної зовнішньої перевірки на незалежних даних. Тому отримані показники точності доцільно трактувати насамперед як оцінку внутрішньої відтворюваності вже побудованої сегментації в межах наявного масиву, а не як остаточне підтвердження того, що алгоритм з такою самою якістю працюватиме на нових вибірках. Водночас для поточного етапу дослідження використання всієї аналітичної вибірки було практично виправданим, з огляду на потребу стабільно відтворити п'ятикластерну структуру, включно з найменшим кластером, що охоплював 156 респондентів. Крім того, кластерний аналіз – це «навчання без вчителя», тобто, ми ніколи не будемо знати «реальну» належність певного кейса до кластеру, оскільки самі будемо цю типологію. Відповідно, принципово відсутня інформація, яка б дала змогу провести повноцінну зовнішню валідацію алгоритму профілювання.

Для перевірки більш стійких альтернатив, які із статистичної точки зору більше відповідають типам наявних в нас даних, було протестовано кілька інших підходів до класифікації. Мультиноміальна логістична регресія, яка коректно працює із порядковими предикторами, на повному наборі предикторів показала дуже високу частку правильно класифікованих випадків, однак супроводжувалася числовою нестабільністю, зокрема сингулярностями матриці Гессе та надмірно великими оцінками окремих параметрів, через що її валідність залишалася непевною. Скорочена мультиноміальна модель, в яку увійшло менше предикторів, була більш інтерпретованою, але проблеми числової нестабільності збереглися, а загальна точність класифікації знизилася до 77,1%. За таких умов мультиноміальні моделі не дали достатньо надійної статистичної основи, щоб розглядати їх як кращий

фінальний профілювальний алгоритм виходячи з характеристик даних, з якими ми працюємо.

Для додаткової перевірки прикладної придатності також було протестовано деревоподібні класифікаційні моделі. Початкова модель на основі Exhaustive CHAID виявилася надто грубою: загальна частка правильно класифікованих випадків становила 62,3%, а за крос-валідацією помилка сягала 40,8%; особливо слабо модель відтворювала кластер 4, для якого правильна класифікація становила лише 16,7%. Після переходу до CRT-дерева з більш гнучкими параметрами якість покращилася: загальна точність зросла до 69,4%, а точність для кластера 4 — до 42,2%. Водночас навіть у цій конфігурації дерево залишилося помітно слабшим за дискримінантний аналіз, який показав вищу загальну точність і кращу відтворюваність сегментації. Крім того, зі зростанням якості дерево суттєво ускладнилося за структурою, частково втративши перевагу простоти й наочності. З огляду на це деревоподібні моделі доцільно розглядати як допоміжний інструмент перевірки та ілюстрації логіки сегментації, але не як основний профілювальний алгоритм.

Таблиця 3. Порівняння протестованих профілювальних підходів

Підхід	Загальна точність	Основне обмеження	Висновок
Дискримінантний аналіз	80,7%	Часткове порушення передумов методу	Основний алгоритм
Дискримінантний аналіз, крос-валідація	79,7%	Внутрішня, а не зовнішня валідація	Прийнятна відтворюваність
Мультиноміальна логістична регресія, повна	85,9%	Числова нестабільність, сингулярності	Не обрана
Мультиноміальна логістична регресія, скорочена	77,1%	Нестабільність зберіглася, нижча точність	Не обрана
Дерево Exhaustive CHAID	62,3%	Надто груба класифікація	Не обране
Дерево CRT	69,4%	Нижча точність і складніша структура	Допоміжний інструмент

## 9.6. Підсумок

З урахуванням поєднання точності, технічної прозорості, відтворюваності та легкості автоматизації саме дискримінантний аналіз було обрано як основний профілювальний алгоритм у межах цього дослідження. Його доцільно розглядати як прикладний інструмент автоматичного віднесення нових респондентів до найбільш імовірного сегмента в межах уже побудованої кластерної типології. Методичні обмеження цього підходу на доступних нам даних відомі, однак у поточній конфігурації саме він забезпечив найкращий баланс між статистичною якістю, інтерпретованістю та практичною придатністю.

## 9.7. Етичні застереження щодо використання алгоритму

Профілювальний алгоритм слід розглядати як допоміжний аналітичний інструмент для попереднього віднесення нового клієнта до найбільш імовірного сегмента. Його результат має ймовірнісний характер і відображає близькість профілю респондента до одного з раніше виділених кластерів, а не встановлює «справжню» належність до певної групи. Відповідно, автоматично визначений сегмент доцільно використовувати для орієнтовного профілювання, планування комунікації та підбору релевантних послуг, але не як жорстке або остаточне рішення щодо потреб конкретної людини.

Особливої обережності потребує те, що алгоритм спирається на чутливі дані, зокрема інформацію про вживання психоактивних речовин, сексуальну поведінку та інші поведінкові характеристики. Це вимагає обмеженого доступу до результатів профілювання, використання їх виключно в межах професійних завдань проекту та недопущення стигматизуючої інтерпретації сегментів. Назва або номер кластера не повинні сприйматися як характеристика «цінності» чи «проблемності» клієнта; йдеться про аналітичний профіль, корисний для кращого розуміння потреб і маршрутів підтримки.

Крім того, результати алгоритму не слід використовувати як єдину підставу для визначення необхідності в певних послугах, обмеження доступу до сервісів або пріоритезації допомоги без додаткового професійного аналізу. Остаточне рішення в кожному конкретному випадку має залишатися за фахівцем або командою проекту, які враховують повний контекст взаємодії з клієнтом. Практично коректне використання алгоритму передбачає регулярний перегляд його якості, перевірку того, чи не погіршується точність на нових даних, а також оновлення моделі в разі зміни структури клієнтської групи або логіки надання послуг.

## 10. Обмеження сегментації

Отриману сегментацію слід інтерпретувати з урахуванням низки методичних обмежень. Насамперед вони пов'язані з характером аналітичної вибірки. Дослідження спирається на неімовірнісну вибірку клієнтів і потенційних клієнтів проекту, які самостійно заповнювали анкету в межах онлайн-рекрутингу для подальшого отримання послуг. Відповідно, виділені кластери описують структуру саме цієї сукупності, а не репрезентують усю ширшу популяцію людей, які вживають психоактивні речовини в Україні. Сегментацію доцільно розглядати як інструмент прикладного розуміння внутрішньої структури клієнтської групи проекту, а не як підставу для широких популяційних узагальнень.

Друге обмеження стосується характеристик самих даних. Значна частина змінних, на яких будувалася сегментація, базується на самозвіті респондентів щодо чутливих тем: вживання психоактивних речовин, ін'єкційного досвіду, сексуальної поведінки, поєднання ПАР з алкоголем, тестування та звернення по послуги. У таких питаннях завжди існує ризик похибок пам'яті, неточного відтворення подій, різного розуміння формулювань, а також соціально бажаних відповідей. Навіть за умов анонімного самозаповнення це обмеження повністю не зникає. Відповідно, кластери відображають не «об'єктивно виміряну» поведінку, а зафіксовані в анкеті патерни самозвіту.

Окреме обмеження пов'язане з пропущеними значеннями та логікою підготовки даних до аналізу. Для побудови кластерних рішень використовувалися аналітичні змінні після перекодування пропущених відповідей. Це може впливати на фінальну структуру сегментів, особливо якщо пропуски розподілені нерівномірно або пов'язані зі змістом самих запитань.

Суттєвим є також обмеження, пов'язане з типами змінних і метрикою відстані. У сегментації одночасно використовувалися порядкові, номінальні та метричні змінні, а різні алгоритми по-різному працюють із таким набором даних. Зокрема, двоетапний кластерний аналіз у SPSS у практичному сенсі трактує порядкові змінні як номінальні й не повною мірою зберігає градієнт між упорядкованими категоріями. Ієрархічний метод Варда дозволяє зберігати градієнти ознак, але вимагає розгляду всіх змінних як метричних. Додаткова перевірка через відстань Говера та РАМ була методично важливою саме тому, що краще узгоджується зі змішаним типом даних, однак у нашому випадку не дала достатньо рельєфного й корисного на практиці рішення. Отже, жоден із протестованих підходів не можна вважати статистично беззаперечним; ідеться радше про вибір найбільш придатної робочої моделі за умов реальних обмежень даних.

Ще одне обмеження стосується вибору кількості кластерів. Формальні статистичні критерії не дали підстав говорити про дуже чітко відмежовану «природну» кластерну структуру. Силуетні показники для протестованих рішень залишалися низькими, а різні методи частково по-різному проводили межі між сегментами. Водночас двоетапний алгоритм вказав саме на п'ятикластерне рішення як оптимальне із статистичної точки зору, а порівняння рішень методом Варда і двоетапного алгоритму показало значний перетин їхньої структури. З точки зору інтерпретації та практичної корисності також п'ятикластерне рішення виявилось оптимальним. Тому вибір рішення із п'яти кластерів спирається на поєднання статистичних індикаторів, інтерпретованості, практичної корисності

отриманих профілів. Це слід враховувати: такий варіант сегментації має ряд переваг, але не є єдино можливим.

Нарешті, окремих застережень потребує практичне використання профілювального алгоритму, побудованого на основі дискримінантного аналізу. Його призначення полягає у попередньому автоматичному віднесенні нового респондента до найбільш імовірного сегмента на основі вже наявних відповідей анкети. Такий інструмент є корисним для оперативного профілювання, однак його результат має ймовірнісний характер і не повинен трактуватися як остаточний «діагноз» або жорстка характеристика конкретної людини.

## Додатки

### Додаток 1. Повний перелік змінних, використаних у кластерному аналізі

У кластерному аналізі використовувався набір змінних, що охоплював демографічні характеристики, досвід ін'єкційного та неін'єкційного вживання психоактивних речовин, поєднання ПАР з алкоголем, окремі показники сексуальної поведінки та ризику.

Таблиця 4. Змінні, включені до кластерного аналізу

Код змінної	Блок анкети	Формулювання / зміст змінної	Приклад шкали відповіді в анкеті	Тип змінної у масиві	Роль у кластерному аналізі
Q9	Скринінг	Вкажи свою стать	Жіноча; чоловіча; трансгендерна жінка; трансгендерний чоловік; небінарна особа	Nominal	Показник статті, для аналізу були відібрані тільки чоловіки та жінки
age	Скринінг	Скільки тобі повних років? / числова версія віку	Відкрита числова відповідь: повний вік у роках	Scale	Показник віку
Q16clust_rec	Досвід ін'єкційного вживання	Коли останній раз ти вживав/ла будь-які психоактивні речовини ін'єкційним шляхом?	Протягом останніх 30 днів; протягом останніх 6 місяців, але не останніх 30 днів; більше 6 місяців тому; більше 12 місяців тому; ніколи; не знаю / не пам'ятаю; відмова від відповіді	Ordinal	Відображає наявність і давність ін'єкційного досвіду
Q18.1clust_rec	Неін'єкційне вживання ПАР	Екстазі (MDMA): коли востаннє вживав/ла?	Останні 30 днів; останні 6 місяців, але не останні 30 днів; більше 6 місяців тому; ніколи; не знаю / не пам'ятаю	Ordinal	Давність вживання MDMA

Код змінної	Блок анкети	Формулювання / зміст змінної	Приклад шкали відповіді в анкеті	Тип змінної у масиві	Роль у кластерному аналізі
Q18.2clust_rec	Неін'єкційне вживання ПАР	Амфетамін: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання амфетаміну
Q18.3clust_rec	Неін'єкційне вживання ПАР	LSD: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання LSD
Q18.4clust_rec	Неін'єкційне вживання ПАР	Кокаїн: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання кокаїну
Q18.5clust_rec	Неін'єкційне вживання ПАР	Метамфетамін: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання метамфетаміну
Q18.6clust_rec	Неін'єкційне вживання ПАР	Кетамін: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання кетаміну
Q18.7clust_rec	Неін'єкційне вживання ПАР	Бутират (GHB, GBL): коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання GHB/GBL
Q18.8clust_rec	Неін'єкційне вживання ПАР	Мефедрон: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання мефедрону
Q18.9clust_rec	Неін'єкційне вживання ПАР	Альфа-PVP: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання alpha-PVP
Q18.10clust_rec	Неін'єкційне вживання ПАР	Інші солі (синтетичні катіони): коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання інших солей

Код змінної	Блок анкети	Формулювання / зміст змінної	Приклад шкали відповіді в анкеті	Тип змінної у масиві	Роль у кластерному аналізі
Q18.11clust_rec	Неін'єкційне вживання ПАР	Гриби / псилоцибін: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання псилоцибіну
Q18.12clust_rec	Неін'єкційне вживання ПАР	2С-В: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання 2С-В
Q18.13clust_rec	Неін'єкційне вживання ПАР	Опіати / опіоїди: коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність неін'єкційного вживання опіатів / опіоїдів
Q18.14clust_rec	Неін'єкційне вживання ПАР	Лірика (прегабалін): коли востаннє вживав/ла?	Та сама шкала давності неін'єкційного вживання	Ordinal	Давність вживання прегабаліну
Q19clust_rec	Поєднання ПАР та алкоголю	Чи траплялося одночасно вживати алкоголь та інші ПАР?	Так, регулярно (майже кожен раз, коли вживаю психоактивні речовини); так, іноді / рідко; ні, ніколи; важко відповісти	Ordinal	Частота змішаного вживання алкоголю і ПАР
Q21clust_rec	Поєднання ПАР та алкоголю	Чи збільшується кількість алкоголю під час вживання ПАР?	Так, майже завжди; іноді; ні, кількість алкоголю не змінюється; важко відповісти	Ordinal	Показник посилення алкогольного компоненту
Q22clust_rec	Поєднання ПАР та алкоголю	Як часто за останній рік ти випивав/ла 5+ стандартних порцій алкоголю за раз?	Ніколи; раз на місяць чи рідше; 2–4 рази на місяць; 2–3 рази на тиждень; 4 та більше разів на тиждень; важко відповісти	Ordinal	Індикатор heavy drinking
Q32_1clust_rec	Сексуальні стосунки	Коли ти востаннє займався/лася сексом з чоловіком?	Останні 30 днів; останні 6 місяців, але не останні 30 днів; більше 6	Ordinal	Давність сексуальних контактів з чоловіками

Код змінної	Блок анкети	Формулювання / зміст змінної	Приклад шкали відповіді в анкеті	Тип змінної у масиві	Роль у кластерному аналізі
			місяців тому; ніколи; не знаю / не пам'ятаю		
Q32_2clust_rec	Сексуальні стосунки	Коли ти востаннє займався/лася сексом з жінкою?	Та сама шкала давності сексуального досвіду	Ordinal	Давність сексуальних контактів з жінками
Q32_3clust_rec	Сексуальні стосунки	Коли ти востаннє займався/лася сексом з двома або більше чоловіками одночасно?	Та сама шкала давності сексуального досвіду	Ordinal	Давність досвіду групового сексу з чоловіками
Q32_4clust_rec	Сексуальні стосунки	Коли ти востаннє займався/лася сексом з двома або більше жінками одночасно?	Та сама шкала давності сексуального досвіду	Ordinal	Давність досвіду групового сексу з жінками
Q32_5clust_rec	Сексуальні стосунки	Коли ти востаннє займався/лася сексом з двома або більше партнерами різної статі одночасно?	Та сама шкала давності сексуального досвіду	Ordinal	Давність досвіду групового сексу зі змішаним складом партнерів
Q34clust_rec	Сексуальні стосунки	Перед або під час останнього сексу ти вживав/ла ПАР (крім марихуани та алкоголю)?	Так; ні; не знаю / не пам'ятаю	Nominal	Індикатор сексу під впливом ПАР
Q36clust_rec	Сексуальні стосунки	Чи використовував/ла ти або твій партнер(-и) презерватив під час останнього сексу?	Так; ні; не знаю / не пам'ятаю	Nominal	Індикатор бар'єрного захисту під час останнього сексу

## Додаток 2. Схема перекодування змінних і створення похідних змінних

Для проведення кластерного набору даних було створено аналітичні версії частини змінних. Перекодування мало на меті позбутися пропущених значень. У більшості випадків логіка перекодування полягала в тому, щоб об'єднати пропущені або невизначені відповіді на кшталт «не знаю / не пам'ятаю», «відмова від відповіді» з однією зі змістовних категорій, аби не втрачати спостереження в кластерному аналізі.

Таблиця 5. Перекодування змінних для кластерного аналізу

Змінна	Вихідні категорії	Правило перекодування	Підсумкова аналітична логіка
Q16clust_rec	1 — ін'єкційно протягом останніх 30 днів; 2 — протягом останніх 6 місяців, але не 30 днів; 3 — більше 6 місяців тому, але не більше 12; 4 — більше 12 місяців тому; 5 — ніколи; 98/99 — не знаю / відмова	98 thru 99 = 5	Невизначені відповіді об'єднані з категорією «ніколи»
Q18.1clust_rec – Q18.14clust_rec	1 — останні 30 днів; 2 — останні 6 місяців, але не останні 30 днів; 3 — більше 6 місяців тому; 5 — ніколи; 98 — не знаю / не пам'ятаю	98 = 5	Невизначені відповіді об'єднані з категорією «ніколи»
Q19clust_rec	1 — так, регулярно; 2 — так, іноді / рідко; 3 — ні, ніколи; 4 — важко відповісти	4 = 3	«Важко відповісти» об'єднано з категорією відсутності змішаного вживання
Q21clust_rec	1 — так, майже завжди; 2 — іноді; 3 — ні, кількість алкоголю не змінюється; 4 — важко відповісти	4 = 3	Невизначені відповіді об'єднані з категорією «ні, кількість алкоголю не змінюється»
Q22clust_rec	1 — ніколи; 2 — раз на місяць чи рідше; 3 — 2–4 рази на місяць; 4 — 2–3 рази на тиждень; 5 — 4 та більше разів на тиждень; 6 — важко відповісти	6 = 1	«Важко відповісти» об'єднано з категорією «ніколи»
Q32_1clust_rec – Q32_5clust_rec	1 — останні 30 днів; 2 — останні 6 місяців, але не останні 30 днів; 3 — більше 6 місяців тому; 4 — ніколи; 98 — не знаю / не пам'ятаю	98 = 4	Невизначені відповіді об'єднані з категорією «ніколи»

<b>Змінна</b>	<b>Вихідні категорії</b>	<b>Правило перекодування</b>	<b>Підсумкова аналітична логіка</b>
Q34clust_rec	1 — так; 2 — ні; 98 — не знаю / не пам'ятаю	98 = 2	Невизначені відповіді об'єднано з відповіддю «ні»
Q36clust_rec	1 — так; 2 — ні; 98 — не знаю / не пам'ятаю	98 = 1	Невизначені відповіді об'єднано з відповіддю «так»

*Таблиця 6. Похідні індикатори, створені для подальшого опису кластерів*

Окремо від кластерного набору в масиві було створено низку похідних індикаторів, які використовувалися не для побудови самих кластерів, а для їх подальшої інтерпретації та статистичного опису. Найважливіші з них наведено нижче.

<b>Змінна</b>	<b>Логіка побудови</b>	<b>Аналітичний зміст</b>
any_group_sex	1, якщо Q32_3 або Q32_4 або Q32_5 мають коди 1–3	Будь-який досвід групового сексу протягом життя
any_group_sex_6_month	1, якщо Q32_3 або Q32_4 або Q32_5 мають коди 1–2	Досвід групового сексу протягом останніх 6 місяців
any_group_sex_30_days	1, якщо Q32_3 або Q32_4 або Q32_5 = 1	Досвід групового сексу протягом останніх 30 днів
last_sex_was_in_group	На основі Q33: 0 — секс з одним партнером, 1 — груповий секс	Груповий характер останнього сексуального контакту
alpha_mefed_salts	1, якщо під час останнього сексу вживали мефедрон, alpha-PVP або інші солі (Q35_8, Q35_9, Q35_10)	Маркер вживання «солей» у сексуальному контексті
used30	Кількість змінних Q18.1clust_rec – Q18.14clust_rec з кодом 1	Число різних речовин, які респондент вживав упродовж останніх 30 днів
ev_male, ev_female	Дихотомізація питань про досвід сексу з чоловіками / жінками: 1–3 = так, 4 = ні	Будь-який сексуальний досвід з партнером відповідної статі

<b>Змінна</b>	<b>Логіка побудови</b>	<b>Аналітичний зміст</b>
ev_grp_male, ev_grp_female	Дихотомізація питань про груповий секс з чоловіками / жінками: 1–3 = так, 4 = ні	Будь-який досвід групового сексу з партнерами відповідної статі
d30_male, d30_female	Дихотомізація: 1 = так, 2–4 = ні	Секс з чоловіками / жінками протягом останніх 30 днів
d30_grp_male, d30_grp_female	Дихотомізація: 1 = так, 2–4 = ні	Груповий секс з чоловіками / жінками протягом останніх 30 днів
homo_ever	Побудовано окремо для чоловіків і жінок як маркер досвіду одностатевих стосунків	Досвід одностатевого сексуального контакту протягом життя
homo_30d	Аналогічно, але для останніх 30 днів	Одностатевий секс протягом останніх 30 днів

Додаток 3. Методичні матеріали щодо вибору кількості кластерів

Таблиця 7. Порівняння протестованих підходів до вибору кількості кластерів

Підхід	Розглянуті рішення	Ключовий статистичний результат	Інтерпретаційний висновок
TwoStep (SPSS)	1–5 кластерів	Автоматично обрано 5 кластерів за BIC; середній silhouette = 0,093	Підтримує 5-кластерну структуру, але з низькою відокремленістю
Ward (SPSS)	4 і 5 кластерів	4 кластери: silhouette = 0,112; 5 кластерів: silhouette = 0,054	5 кластерів слабші за силуетом, але сильніші за змістовною інтерпретованістю
PAM / Gower (R)	2–8 кластерів	Найкраще рішення за silhouette — 2 кластери (0,127); 5 кластерів = 0,040	2 кластери статистично кращі, але аналітично надто грубі
Порівняння Ward 5 vs TwoStep 5	5 і 5 кластерів	Cramer's V = 0,636	Спільне ядро сегментації відтворюється, розбіжності стосуються складніших ризикових профілів

Таблиця 8. Показники силуету для основних протестованих кластерних рішень

Рішення	Середній силует
TwoStep, 5 кластерів	0,093
Ward, 4 кластери	0,112
Ward, 5 кластерів	0,054
PAM/Gower, 2 кластери	0,127
PAM/Gower, 3 кластери	0,061
PAM/Gower, 4 кластери	0,044
PAM/Gower, 5 кластери	0,040

Додаток 4. Матриця класифікації профілюючого алгоритму

Таблиця 9. Матриця класифікації профілювального алгоритму, початкова класифікація (кількість випадків)

<b>Фактичний кластер</b>	<b>Прогноз 1</b>	<b>Прогноз 2</b>	<b>Прогноз 3</b>	<b>Прогноз 4</b>	<b>Прогноз 5</b>	<b>Усього</b>
1	385	59	14	1	0	459
2	62	516	29	22	9	638
3	47	31	694	32	7	811
4	3	54	76	328	18	479
5	0	6	4	17	129	156

Таблиця 8. Матриця класифікації профілювального алгоритму, крос-валідація (кількість випадків)

<b>Фактичний кластер</b>	<b>Прогноз 1</b>	<b>Прогноз 2</b>	<b>Прогноз 3</b>	<b>Прогноз 4</b>	<b>Прогноз 5</b>	<b>Усього</b>
1	381	63	14	1	0	459
2	66	509	30	22	11	638
3	49	33	690	32	7	811
4	3	57	79	320	20	479
5	0	7	4	19	126	156

Таблиця 9. Частка правильно класифікованих випадків за кластерами, %

Кластер	Початкова класифікація	Крос-валідація
1	83,9	83,0
2	80,9	79,8
3	85,6	85,1
4	68,5	66,8
5	82,7	80,8
<b>Загалом</b>	<b>80,7</b>	<b>79,7</b>

Додаток 5. Технічний опис профілюючого алгоритму (дискримінантного аналізу)

Таблиця 10. Основні технічні параметри дискримінантного аналізу

Параметр	Значення
Кількість кластерів	5
Кількість дискримінантних функцій	4
Кількість предикторів	27
Кількість спостережень у моделі	2543
Виключено через пропущені значення	0
Апріорні ймовірності	Пропорційно розмірам груп
Збережені результати	Прогнозована група, значення функцій, ймовірності належності до кластеру

Таблиця 11. Канонічні дискримінантні функції

Функція	Eigenvalue	% варіації	Кумулятивно, %	Канонічна кореляція
1	2,445	52,5	52,5	0,842
2	1,559	33,5	86,0	0,781
3	0,476	10,2	96,3	0,568
4	0,173	3,7	100,0	0,384

Таблиця 12. Перевірка статистичної значущості дискримінантних функцій

Перевірка функцій	Wilks' Lambda	Chi-square	df	p
1-4	0,065	6886,152	108	<0,001
2-4	0,226	3761,815	78	<0,001
3-4	0,577	1388,091	50	<0,001
4	0,852	403,977	24	<0,001

Таблиця 13. Якість класифікації

Показник	Значення
Частка правильно класифікованих випадків, початкова класифікація	80,7%
Частка правильно класифікованих випадків, крос-валідація	79,7%

Таблиця 14. Частка правильно класифікованих випадків за кластерами, %

Кластер	Початкова класифікація	Крос-валідація
1	83,9	83,0
2	80,9	79,8
3	85,6	85,1
4	68,5	66,8
5	82,7	80,8

Таблиця 15. Перевірка передумови рівності коваріаційних матриць

Показник	Значення
Box's M	12413,843
Approx. F	7,957
df1	1512
df2	1866125,663
p	<0,001

Перші дві дискримінантні функції пояснювали 86,0% розділювальної варіації, тому саме вони мають основне змістовне значення для інтерпретації моделі. За набором коефіцієнтів і зв'язків із функціями найбільш виражене навантаження мали стать, ін'єкційний досвід, окремі показники сексуальної поведінки, а також частина профілю вживання ПАР та одночасного вживання ПАР з алкоголем.

Таблиця 16. Змінні з найбільш вираженим внеском у дискримінацію кластерів

Змінна	Ознака впливу в моделі	Короткий зміст
Q9	Один із перших і найсильніших розділювачів	Стать респондента
Q16clust_rec	Помітне навантаження у функціях; виражений внесок у розмежування кластерів	Давність ін'єкційного вживання
Q32_1clust_rec	Відчутний внесок у функції, пов'язані з сексуальним профілем	Секс з чоловіком
Q32_2clust_rec	Відчутний внесок у функції, пов'язані з сексуальним профілем	Секс з жінкою
Q32_3clust_rec	Один із найвиразніших коефіцієнтів серед сексуальних змінних	Груповий секс з чоловіками
Q32_4clust_rec	Дуже сильне навантаження в одній із функцій	Груповий секс з жінками
Q32_5clust_rec	Помітний внесок у дискримінацію ризикових профілів	Груповий секс з партнерами різної статі
Q18.13clust_rec	Один із найсильніших зв'язків у матриці структур	Неін'єкційне вживання опіатів / опіодів
Q19clust_rec	Стабільний внесок у функції	Поєднання ПАР з алкоголем
Q21clust_rec	Додатково підсилює алкогольний профіль	Збільшення кількості алкоголю під час вживання ПАР

## Додаток 6. Ілюстративні таблиці розподілу ключових змінних за кластерами

У таблицях нижче наведено окремі ключові показники, які найкраще ілюструють відмінності між кластерами. Для категоріальних змінних подано частки в межах кластера, для віку — середнє значення.

Таблиця 17. Демографічний, ін'єкційний та алкогольний профіль кластерів

Показник	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5	р / тест
Розмір кластера, n	459	638	811	479	156	—
Частка у вибірці, %	18,0	25,1	31,9	18,8	6,1	—
Жінки, %	94,6	79,9	11,1	28,2	46,8	p<0,001; V=0,680
Середній вік, років	24,3	24,9	26,5	26,2	23,7	Welch p<0,001
Ін'єкційне вживання протягом останніх 30 днів, %	0,7	2,0	3,6	25,1	31,4	p<0,001; V=0,232
Ніколи не вживали ін'єкційно, %	90,0	84,0	81,3	56,6	34,0	p<0,001; V=0,232
Регулярно поєднували алкоголь та ПАР, %	7,2	41,1	16,8	48,0	52,6	p<0,001; V=0,232
Вживання 5 стандартних порцій алкоголю 2–3 рази на тиждень або частіше, %	16,1	25,6	18,4	36,1	35,8	p<0,001; V=0,115

Таблиця 18. Сексуальна поведінка та сексуальний ризик у кластерах

Показник	Кластер 1	Кластер 2	Кластер 3	Кластер 4	Кластер 5	p / тест
Будь-який груповий секс протягом останніх 6 місяців, %	13,7	41,5	17,1	29,0	99,4	p<0,001; V=0,452
Будь-який груповий секс протягом останніх 30 днів, %	0,0	10,5	3,0	4,6	69,9	p<0,001; V=0,569
Вживання ПАР перед або під час останнього сексу, %	13,5	41,8	20,6	54,3	80,8	p<0,001; V=0,287
Вживання мефедрону / alpha-PVP / інших солей у сексуальному контексті, %	1,5	14,3	3,9	20,3	36,5	p<0,001; V=0,303

Додаток 7. Діагностика пропущених значень у змінних кластерного аналізу

Таблиця 19. Частка пропущених або невизначених відповідей у ключових блоках кластероутворюючих змінних

<b>Змінна / блок</b>	<b>Частка пропущених або невизначених відповідей</b>
Вік	0,0%
Стать	0,0%
Сексуальні контакти з різними типами партнерів	1,4–2,0%
Вживання ПАР перед або під час останнього сексу	3,9%
Використання презерватива під час останнього сексу	3,6%
Одночасне вживання алкоголю та ПАР	3,1%
Ін'єкційне вживання ПАР	5,0%
Вживання 5+ стандартних порцій алкоголю	5,4%
Інші синтетичні катіони	7,3%
Збільшення кількості алкоголю при вживанні ПАР	8,3%

Таблиця 20. Розподіл респондентів за кількістю пропущених або невизначених значень серед кластероутворюючих змінних

<b>Кількість пропусків серед 27 змінних</b>	<b>Частка респондентів</b>
0	62,2%
1	19,2%
2	7,4%
3	4,4%
4	2,5%
5	1,4%
6	1,0%

**Кількість пропусків серед 27 змінних Частка респондентів**

7 і більше 1,9%

*Таблиця 21. Середня кількість пропущених або невизначених значень за кластерами*

**Кластер Середня кількість пропущених / невизначених значень**

1	1,14
2	0,82
3	1,03
4	0,70
5	1,02
Загалом	0,93