

АНАЛІЗ НАБОРУ ДАНИХ ЗВЕРНЕНЬ ГРОМАДЯН У МІСТІ ВІННИЦІ ДЛЯ ПОБУДОВИ ІНТЕРАКТИВНОГО ДАШБОРДУ ДЛЯ ПРОГНОЗУВАННЯ

Вінницький національний технічний університет

Анотація

Дану роботу присвячено аналізу відкритого набору даних звернень громадян у місті Вінниці з метою подальшого використання для побудови інтерактивного дашборду для прогнозування. Проведено дослідження структури даних, визначено основні ознаки, виконано попередній аналіз та оцінено потенціал даних для моделювання часових залежностей. Набір даних дозволяє виявляти тенденції звернень та створити основу для їх прогнозування.

Ключові слова: відкриті дані, звернення громадян, аналіз даних, дашборд, прогнозування, часові ряди.

Abstract

This work is dedicated to analyzing an open dataset of citizen inquiries in the city of Vinnytsia with the aim of using it to build an interactive forecasting dashboard. We examined the data structure, identified key features, conducted a preliminary analysis, and assessed the data's potential for modeling temporal dependencies. The dataset allows for the identification of trends in citizen inquiries and provides a basis for forecasting them

Keywords: open data, citizen inquiries, data analysis, dashboard, forecasting, time series.

Вступ

З розвитком концепції відкритих даних органи місцевого самоврядування надають доступ до інформації про звернення громадян, що відкриває нові можливості для аналітики та прогнозування.

Звернення громадян є важливим індикатором стану міської інфраструктури та рівня задоволеності населення, оскільки відображають актуальні проблеми у різних сферах життєдіяльності міста.

Аналіз таких даних є важливим для підвищення ефективності роботи міських служб, оскільки дозволяє виявляти проблемні сфери, сезонні коливання та навантаження на інфраструктуру.

Метою даної роботи є дослідження набору даних звернень громадян для подальшого використання у системі прогнозування та візуалізації у вигляді інтерактивного дашборду.

Формування та опис набору даних

У даному дослідженні використано відкритий набір даних «Дані про надходження звернень на телефонні гарячі лінії, в аварійно-диспетчерські служби, телефонні центри тощо. Відділ оперативного реагування "Цілодобова варта" Вінницької міської ради»[1], розміщений на державному порталі відкритих даних Вінницької міської ради[2], відкритість даного набору даних дозволила працювати з ним на платформі Kaggle[3].

Набір містить інформацію про звернення громадян, що надійшли у відділ оперативного реагування "Цілодобова варта" Вінницької міської ради та включає такі основні характеристики:

- дата реєстрації звернення;
- тип звернення;
- тематика (сфера проблеми);
- спосіб надходження;
- геодані (поштовий індекс, країна, регіон, район, територіальна громада, населений пункт, вулиця, номер об'єкта);
- статус розгляду звернення;
- виконавчий орган або відповідальна структура;
- результат розгляду заявки;
- інші супровідні атрибути (ідентифікатори, опис тощо).

Дані мають табличну структуру та представлені у форматі, придатному для подальшої обробки. Для зручності аналізу та візуалізації проведено попередній перегляд набору даних, що дозволяє оцінити обсяг та структуру інформації, наявність пропущених значень та типи даних. Приклад структури набору даних наведено на рис. 1.

	uid	receivedDateTime	type	kind	content	accrualMethod	CATUTTC	addressPostCode	addressAdminUnitL1	addressAdminUnitL2	...	addressAdminUnitL4	addressPostName	addressThoroughfare
0	289873	2025-01-01T01:23:14	Заява	Холодне водопостачання	Течія зовнішньої мережі	Засоби телефонного зв'язку	UA05020030010063857	21050.0	Україна	Вінницька область	—	Вінницька	Вінниця	вулиця Червонохрестівська
1	289874	2025-01-01T06:26:58	Заява	Запобігання і реагування на надзвичайні ситуації	Пожежа (інше)	Засоби телефонного зв'язку	UA05020030010063857	21050.0	Україна	Вінницька область	—	Вінницька	Вінниця	вулиця Соборна
2	289875	2025-01-01T07:00:15	Скарга	Освітлення	Вимкнути зовнішнє освітлення	Засоби телефонного зв'язку	UA05020030010063857	21023.0	Україна	Вінницька область	—	Вінницька	Вінниця	провулок 2-й Омеляна Грабця
3	289876	2025-01-01T08:18:39	Заява	Газопостачання	Запах газу в під'їзді	Засоби телефонного зв'язку	UA05020030010063857	21009.0	Україна	Вінницька область	—	Вінницька	Вінниця	вулиця Академіка Янгеля
4	289877	2025-01-01T08:28:26	Заява	Гаряче водопостачання	Течія мережі по стояку	Засоби телефонного зв'язку	UA05020030010063857	21029.0	Україна	Вінницька область	—	Вінницька	Вінниця	вулиця Миколи Ващука

Рис. 1. Приклад структури набору даних

Даний рисунок демонструє перші кілька рядків набору даних після завантаження та відображає структуру колонок. Зокрема, ми можемо побачити наявність дати звернення, категорії type, опису проблеми та інших супровідних ознак. Важливо відзначити, що дані мають різноманітні типи: числові, текстові та категоріальні. Таке представлення дозволяє швидко оцінити кількість записів та типи даних у кожній колонці. За допомогою цього перегляду можна визначити, які колонки будуть використані для візуалізації та прогнозування, а які – для побудови категоріальних статистик.

Аналіз даних

Для аналізу набору даних було проведено базовий огляд числових та категоріальних ознак, а також часових характеристик[4].

Аналіз показав, що кількість звернень не рівномірна протягом року. Спостерігаються періоди пікового навантаження, що дозволяє визначити найбільш активні місяці для органів місцевого самоврядування. Тому, було побудовано розподіл звернень по місяцях для оцінки сезонних тенденцій, що представлено на рис. 2[5].



Рис. 2. Розподіл звернень по місяцях

На даному графіку видно, що деякі місяці мають значно більше звернень, ніж інші. Це може бути пов'язано з сезонними особливостями, роботою певних служб або специфічними міськими подіями. Наприклад, пік звернень може збігатися з періодами закінчення податкових або адміністративних процедур. Даний графік допомагає виявити тренди та дозволяє надати дані для підготовки органів влади до пікових навантажень. Крім того, він слугує наочною ілюстрацією того, як дані про звернення змінюються протягом року, що є важливою інформацією для подальшого прогнозування.

Для категоріальних ознак, зокрема type, було визначено найпопулярніші значення, що дозволяє оцінити найбільш поширені типи звернень.

Даний графік демонструє найбільш часті категорії звернень громадян. Ми можемо побачити, які типи звернень найчастіше виникають у місті Вінниці, що зображено на рис. 3. Це дозволяє ефективніше планувати ресурси, виділяти пріоритети та розподіляти навантаження між відповідними підрозділами. Також на рис. 4 зображено, що певні категорії зустрічаються набагато частіше за інші, що важливо враховувати при побудові прогнозової моделі та дашборду.

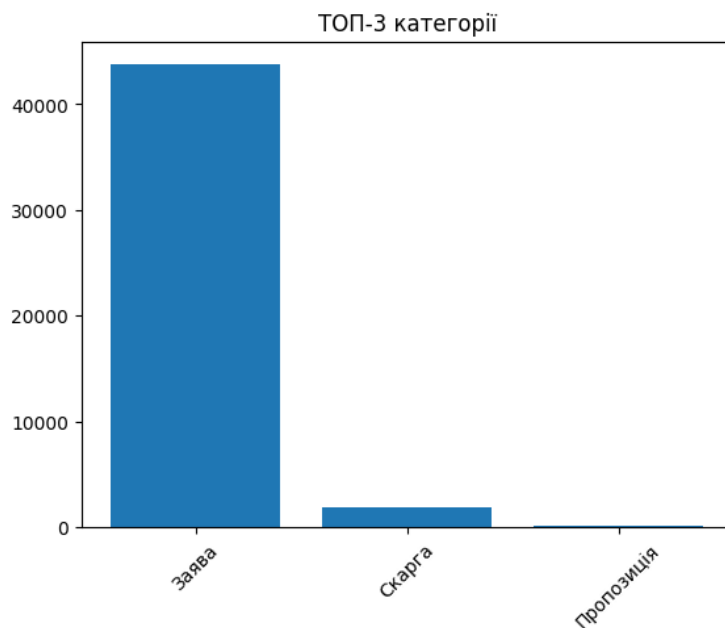


Рис. 3. Топ 3 категорії звернень

	type	count
0	Заява	43738
1	Скарга	1921
2	Пропозиція	156

Рис. 4. Кількість звернень відповідно до певної категорії

Для оцінки часової активності звернень також була побудована таблиця звернень громадян по днях тижня, яка показує, у які дні та місяці спостерігається найбільше звернень. Таблиця демонструє розподіл звернень за двома вимірами: місяць і день тижня. Вона дозволяє визначити закономірності у поведінці громадян, показати пікові дні тижня та місяці, коли навантаження на органи влади є максимальним. Це корисно для планування ресурсів та оптимізації роботи підрозділів. Таке представлення є своєрідною

матрицею активності звернень і допомагає візуально оцінити взаємозв'язки між часом та кількістю звернень. Таблиця активності звернень представлена на рис. 5, також до даної таблиці була побудована теплова карта, яка зображена на рис. 6.

weekday	Friday	Monday	Saturday	Sunday	Thursday	Tuesday	Wednesday
month							
1	522	524	302	246	547	463	512
2	417	456	320	266	455	426	424
3	478	648	395	424	477	418	450
4	499	518	314	311	505	665	652
5	607	542	402	395	587	639	598
6	634	838	400	473	570	660	584
7	615	805	350	320	768	808	818
8	701	552	396	406	619	533	662
9	545	751	308	261	546	674	589
10	797	638	379	348	808	559	704
11	673	830	623	630	748	804	746
12	486	782	350	338	487	766	729

Рис. 5. Таблиця активності звернень по днях тижня

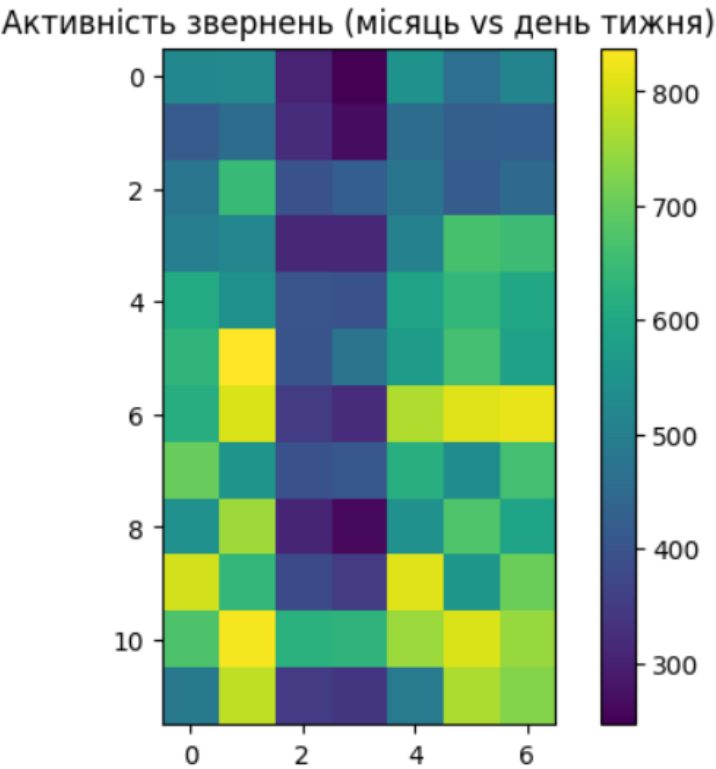


Рис. 6. Теплова карта активності звернень по днях тижня

Даний графік являє собою теплову карту, де по осі X представлені дні тижня, а по осі Y – місяці року. Інтенсивність кольору відображає кількість звернень у кожному комірниці. Це дозволяє швидко виявити пікові комбінації дня і місяця, які відповідають максимальному навантаженню на міські служби. Використання такого графіку є важливим для інтерактивного дашборду, оскільки він дозволяє користувачеві одразу побачити критичні періоди та зрозуміти закономірності у поведінці громадян.

Результати дослідження

На основі проведеного аналізу можна зробити висновок, що набір даних містить детальну інформацію про звернення громадян до органів місцевого самоврядування у місті Вінниці, що дозволяє здійснювати як часовий, так і категоріальний аналіз. Дослідження показало, що кількість звернень є нерівномірною у часі, а саме спостерігаються піки в окремі місяці та дні тижня, що свідчить про наявність певних закономірностей у поведінці громадян. Аналіз популярних категорій звернень дав змогу визначити основні проблеми, з якими стикається населення, а також виділити напрями роботи, що потребують підвищеної уваги з боку органів влади.

Окрему увагу було приділено аналізу залежності між місяцями та днями тижня, який показав, що максимальна активність громадян припадає на конкретні дні, що є важливим для оптимізації роботи відповідних підрозділів. Візуалізація розподілу звернень за місяцями дозволила виявити сезонні тенденції та сформулювати передумови для прогнозування майбутніх періодів. Побудовані графіки та таблиці дали змогу наочно оцінити навантаження на служби та підвищити ефективність планування їх діяльності.

Отримані дані також відкривають можливості для створення інтерактивних дашбордів, які дозволяють відслідковувати тенденції, аналізувати звернення за різними параметрами та формувати обґрунтовані рішення на основі отриманих графіків. Аналіз кореляції між днем тижня та місяцем підтвердив наявність стабільних закономірностей, а також дозволив виявити проблемні часові інтервали, що потребують додаткової уваги.

Крім того, дані є достатньо повними для побудови моделей прогнозування на основі часових рядів і категоріальних ознак, що дає змогу оцінювати майбутнє навантаження на міські служби. Інтерактивні фільтри та візуалізації значно спрощують доступ до необхідної інформації та покращують сприйняття результатів аналізу. У цілому результати дослідження демонструють високий потенціал для автоматизації процесів прогнозування звернень громадян, виявлення часових і тематичних трендів, а також формування рекомендацій щодо оптимізації роботи міських служб і підвищення якості обслуговування населення.

Висновки

Результати проведеного дослідження підтверджують високу аналітичну цінність використаного набору даних щодо звернень громадян до відділу «Цілодобова варта» Вінницької міської ради. Дані є достатньо повними та структурованими, що дозволило здійснити ефективний часовий і категоріальний аналіз. Встановлено, що звернення мають нерівномірний розподіл у часі з вираженими піковими періодами за місяцями та днями тижня, що свідчить про наявність стабільних закономірностей у поведінці громадян.

Аналіз категорій звернень дозволив визначити основні проблемні напрями, які потребують найбільшої уваги з боку міських служб, а також створює основу для більш ефективного розподілу ресурсів. Використання графіків, таблиць і теплових карт продемонструвало свою ефективність у виявленні закономірностей та спрощенні сприйняття даних, що є важливим для прийняття рішень.

Крім того, результати показали, що даний набір даних є придатним для побудови прогнозних моделей і створення інтерактивних дашбордів, що дозволяє автоматизувати процеси аналізу та прогнозування навантаження на міські служби. Це сприяє підвищенню ефективності планування, оптимізації роботи органів місцевого самоврядування та покращенню якості обслуговування громадян.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Дані про надходження звернень на телефонні гарячі лінії, в аварійно-диспетчерські служби, телефонні центри тощо. Відділ оперативного реагування "Цілодобова варта" Вінницької міської ради. [Електронний ресурс]. – Режим доступу: <https://opendata.gov.ua/dataset/informaciya-zi-zvernen-gromadyan/resource/7470a248-1b89-4132-8c97-4db09f85ae74>
2. Портал відкритих даних Вінницької МР, [Електронний ресурс] – Режим доступу: <https://opendata.gov.ua/>
3. Платформа Kaggle. [Електронний ресурс] – Режим доступу: <https://www.kaggle.com/>
4. Upton G., Brawn D. Data Analysis: A Gentle Introduction for Future Data Scientists. – Oxford University Press, 2023. – 161 p.
5. McKinney W. Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter. – 3rd ed. – Sebastopol: O'Reilly Media, 2022. – 544 p.

Липко Максим Віталійович – студент групи СА-22Б, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: makjim18@gmail.com

Варчук Ілона Вячеславівна – к.т.н., доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, e-mail: ivvarchuk@vntu.edu.ua

Lypko Maksym V. – student group SA-22B of Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail makjim18@gmail.com

Varchuk Iлона V. - Ph.D., Assistant Professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: ivvarchuk@vntu.edu.ua