

1. Чисельність населення в м. Дніпро. *[/Електронний ресурс]*
Режим доступу: <https://index.minfin.com.ua/ua/reference/people/town/dnepr/>
2. Офіційна кількість переселенців. *[/Електронний ресурс]*
Режим доступу: <https://suspilne.media/309840-na-dnipropetrovsini-zbilsilas-kilkist-pereselenciv-oficijno-zareestruvalis-ponad-360-tisac/>

УДК 004.6

*Джога М. О., студент 3 курсу
спеціальності 122 «Комп'ютерні науки»
Ніколюк П. К., професор,
доктор фізико-математичних наук.*

МЕТОДИ ОБРОБКИ І АНАЛІЗУ ДАНИХ

Донецький національний університет імені Василя Стуса

Сучасні методи обробки та аналізу даних — це міждисциплінарні галузі, створені та розроблені на науковій основі математики та статистики, розпізнавання образів та штучного інтелекту.

Аналіз даних — це основний вид соціологічної дослідницької роботи, яка спрямована на виявлення стабільності, суттєвих ознак і тенденцій розвитку об'єкта дослідження, включаючи відбір і розрахунок показників, демонстрацію та доведення гіпотез, побудову висновків дослідження. На цій основі підтримується логічна узгодженість, послідовність і обґрунтованість усіх дослідницьких процедур.

Основні цілі аналізу даних: зафіксувати інформацію про об'єкт дослідження в символічній формі, визначити її достовірність, розробити характеристики та показники для об'єктивної та суб'єктивної оцінки процесу дослідження, довести та перевірити гіпотези, узагальнити результати, визначити напрям і форму його практичного застосування.

Основні вимоги до нагляду включають теоретичні аспекти, а також керівна роль методологічних принципів. Концептуальний зв'язок між усіма етапами аналізу та планами дослідження, процедури для забезпечення цілісності, достовірності інформації та отримання результатів дослідження.

Завдяки використанню логіки, математики та статистичних і інформаційні методи, ефективні процедури та сучасні технічні засоби, систематизація, стиснення та більш повне представлення інформації, ітераційний процес аналізу, підвищення рівня достовірності інформації на кожному наступному етапі дослідження, комплексне використання експертної компетенції [1].

Програма аналізу даних є невід'ємною частиною програми соціологічного дослідження. Його провідні завдання: визначити вид і склад необхідної інформації, визначити методи і засоби її реєстрації, вимірювання, обробки і перетворення, забезпечити достовірність даних, визначити форму інтерпретації,

узагальнити дані, встановити методи практичного застосування та оприлюднити результати досліджень.

Вимоги до даних: Потреба в даних як вхідних даних для аналізу визначається на основі вимог тих, хто керує аналізом. Загальний тип об'єкта, для якого збираються дані, називається експериментальною одиницею. Специфічні для населення змінні, такі як вік і дохід, можна ідентифікувати та охопити. Дані можуть являтися категоріальними або числовими.

Збір даних: інформація отримується з різноманітних ресурсів, фахівці можуть пересилати запити зберігачам даних, до прикладу, персоналу, що відповідає за інформаційні технології в організації. Дані також збираються з пристроїв, що знаходяться у навколо нас, до прикладу, камери, супутники та інші пристрої. Теж їх можна одержати різними шляхами, а саме перегляд інформаційних ресурсів з інтернету або опрацювання документів.

Обробка даних: Застосовується для трансформування непроаналізовану інформацію в корисні дані або нові практичні застосунки. Для початку отримані дані необхідно обробити чи організувати для розгляду. Наприклад, це може включати впорядкування інформації у таблиці для аналізування з використанням програмного продукту, що виконує статистичні функції та електронних таблиць [2].

Очищення даних: після аналізу та очищення є шанс, що інформація може бути неповною, дублюватися або містити помилки. Очищення даних виконується через спосіб введення та збереження коректної інформації. Очищення даних - це процес коригування інформації, який унеможливорює виникнення помилок і неточностей. Його завдання - це порівняння записів, знаходження неточностей інформації та покращення якості наявних даних. Проблемні дані також виявляють на основі певних методів аналізу. Наприклад, для фінансової інформації суму певних змінних порівнюють з загальнодоступними цифрами, які є надійними. Є декілька основних способів очищення інформації, що покладаються на те, якого типу дані у колекції. До цього можна віднести адреси електронної пошти, номери телефонів та інші подібні значення.

Методи кількісних даних використовуються для фільтрації даних, які, ймовірно, були введені некоректно. Щоб виявити та зменшити орфографічні помилками, можна застосувати перевірку на орфографію текстової інформації. Проте встановити правильність самого слова складніше [3].

Дослідницький аналіз даних: коли набір інформації очищено, його можна проаналізувати. Фахівці можуть використовувати різні методи, які застосовують пошукове аналізування даних для того, щоб краще зрозуміти повідомлення, що знаходяться в отриманій інформації. Процес аналізу та подальшого дослідження інформації іноді призводить до виникнення додаткового очищення даних чи додаткових запитів на дані. Для того, щоб краще опрацювати дані, застосовують описову статистику, наприклад середнє значення або медіану.

Візуалізація даних - це метод, за допомогою якої аналітики досліджують дані у графічному форматі, для отримання більшої інформації про зміст певних даних.

Алгоритми та моделювання. Математичні формули або моделі можна застосовувати до даних для відображення зв'язків між перемінними, наприклад, використання кореляції або причинно-наслідкового зв'язку.

Інференційна статистика використовує певні процедури, які обчислюють взаємовідношення між деякими показниками. Наприклад, регресійний аналіз можна застосовувати для моделювання того, як зміни в рекламному просторі впливають на динаміку в обсягу продажу.

$$Y = aX + b + \text{помилка} \quad (1)$$

Де, незалежна змінна X – реклама та залежна змінна Y – змінна продажів. З математичної точки зору Y (продажі) є функцією X (реклама), модель змодельована таким чином, що (a) і (b) мінімізують помилку, коли модель передбачає Y для заданого значення X .

Аналітики також можуть спробувати побудувати моделі, які описують дані, з метою спрощення аналізу та передачі результатів [4].

Завдяки використанню сучасних комп'ютерних систем процес обробки та аналізу даних кардинально змінилися, дані стали доступнішими та наочнішими. Вже не потрібно вручну виконувати трудомісткі обчислення, будувати таблиці та графіки - всі ці завдання бере на себе комп'ютер, а людині залишається тільки ставити завдання, вибирати методи їх вирішення та презентація отриманих результатів.

Список літератури:

- [1] Методи аналізу даних : навчальний посібник для студентів В.Є. Бахрушин. – Запоріжжя (дата звернення 29.11.2022)
- [2] Аналіз даних: Навчальний посібник. – Рівне: НУБГП URL: (дата звернення 29.11.2022)
- [3] Stud URL: https://stud.com.ua/88907/pedagogika/metodi_obrobki_danih_interpretatsiyni_metodi (дата звернення 29.11.2022)
- [4]. Wikipedia URL: https://uk.wikipedia.org/wiki/Аналіз_даних (дата звернення 29.11.2022)

УДК 004.6+51.0

*Оліх В.І., студентка 2 курсу
спеціальності 113 «Прикладна
математика»*

*Потапова Н.А., к.е.н., доцент,
доцент кафедри інформаційних
технологій*

ІСТОРІЯ РОЗВИТКУ МЕТОДІВ ОБЧИСЛЕНЬ

Донецький національний університет імені Василя Стуса, м. Вінниця