

Лекція № 1

ВСТУП ДО ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

План

- 1) Поняття інформаційних технологій
- 2) Використання інформаційних технологій в роботі психолога
- 3) Практичне використання інформаційних технологій при аналізі даних

Поняття інформаційних технологій

В широкому сенсі слово **технологія** - це спосіб освоєння людиною матеріального світу за допомогою соціально організованої діяльності, що включає три компоненти: інформаційну(наукові принципи та обґрунтування), матеріальну(знаряддя праці) та соціальну(фахівці, які мають професійні навички). Ця тріада становить сутність сучасного розуміння поняття технологія.

Інформація - це сукупність відомостей (даних), які сприймають із навколишнього середовища (вхідна інформація), видають у навколишнє середовище (вихідна інформація) або зберігають всередині певної системи.

Інформація існує у вигляді документів, креслень, рисунків, текстів, звуковиз в світлових сигналів, електричних та нервових імпульсів тощо. Саме слово 'інформатика' походить від латинського information, що означає виклад, роз'яснення факту, події.

Найбільш важливими властивостями інформації є:

- об'єктивність та суб'єктивність;
- повнота;
- достовірність;
- адекватність;
- доступність;
- актуальність.

Обробка даних містить в собі множину різних операцій. Основними операціями є:

- збір даних - накопичення інформації з метою забезпечення достатньої повноти для прийняття рішення;
- формалізація даних - приведення даних, що надходять із різних джерел до однакової форми;
- фільтрація даних - усунення зайвих даних, які не потрібні для прийняття рішень;
- сортування даних - впорядкування даних за заданою ознакою з метою зручності використання;
- архівація даних - збереження даних у зручній та доступній формі;
- захист даних - комплекс дій, що скеровані на запобігання втрат, відтворення та

модифікації даних;

– транспортування даних - прийом та передача даних між віддаленими користувачами інформаційного процесу. Джерело даних прийнято називати сервером, а споживача - клієнтом;

– перетворення даних - перетворення даних з однієї форми в іншу, або з однієї структури в іншу, або зміна типу носія.

Інформаційні технології (ІТ) - сукупність методів, виробничих процесів і програмно-технічних засобів, інтегрованих з метою збирання, обробки, зберігання, розповсюдження, відображення і використання [інформації](#) в інтересах її [користувачів](#).

Інформаційна система

Інформаційна система - взаємозв'язана сукупність засобів, методів і персоналу, використовувана для зберігання, оброблення та видачі інформації з метою вирішення конкретного завдання.

Сучасне розуміння інформаційної системи передбачає використання комп'ютера як основного технічного засобу обробки інформації. Комп'ютери, оснащені спеціалізованими програмними засобами, є технічною базою та інструментом інформаційної системи.

У роботі інформаційної системи можна виділити такі етапи:

Зародження даних - формування первинних повідомлень, що фіксують результати певних операцій, властивості об'єктів і суб'єктів управління, параметри процесів, зміст нормативних та юридичних актів тощо.

Накопичення і систематизація даних - організація такого їх розміщення, яке б забезпечувало б швидкий пошук і відбір потрібних відомостей, методичне оновлення даних, захист їх від спотворень, втрати, деформування цілісності та ін.

Обробка даних - процеси, внаслідок яких на підставі раніше накопичених даних формуються нові види даних: узагальюючі, аналітичні, рекомендаційні, прогнозні. Похідні дані також можуть зазнавати подальшого оброблення, даючи відомості глибшої узагальненості і т.д.

Відображення даних - подання їх у формі, придатній для сприйняття людиною. Передусім - це виведення на друк, тобто виготовлення документів на так званих твердих (паперових) носіях. Широко використовують побудову графічних ілюстративних матеріалів і формування звукових сигналів.

Розповсюдження ІТ

[Україна](#) за рівнем розвитку інформаційних технологій у світі посіла 75 місце.

Такі дані оприлюднила міжнародна громадська організація Всесвітній економічний форум у своїй шостій щорічній доповіді. У попередньому рейтингу Україна займала 76 позицію за розвитком ІТ-технологій.

У складеному рейтингу лідирує [Данія](#) - завдяки зразковій нормативно-правовій базі і чіткій політиці держави з поширення інформаційних технологій.

Друге місце зайняла [Швеція](#), яка піднялася на шість позицій, ставши однією з країн із найбільш зростаючим ІТ - сектором економіки. Також у першу трійку потрапив Сінгапур. У

першу десятку увійшли [Фінляндія](#), [Швейцарія](#), [Нідерланди](#), [США](#), [Ісландія](#), [Великобританія](#) і [Норвегія](#).

2. Використання інформаційних технологій в роботі психолога

Для розв'язання завдань, пов'язаних з аналізом даних випадкових та непередбачуваних впливів, за останні двісті років був напрацьований потужний і гнучкий арсенал методів, який у сукупності називається математичною статистикою, прикладною статистикою або аналізом даних.

Разом із тим поняття математичної статистики і аналізу даних дещо різняться. В статистичних дослідженнях виходять з апріорної моделі та намагаються оцінити її параметри. В аналізі даних намагаються виходити з даних як таких, маючи мінімум апріорних ідей стосовно їх структури. Далі намагаються зрозуміти, як організовані дані, які змінні або групи змінних пов'язані між собою і яким чином, тобто зрозуміти структуру даних виходячи з них самих.

Методи прикладної статистики і аналізу даних дозволяють:

- виявити закономірності на фоні випадковостей;
- робити обґрунтовані висновки та прогнози;
- давати оцінки ймовірностей виконання або невиконання цих прогнозів.

Основними завданнями обробки даних в інформаційно-аналітичній діяльності є:

- агрегування інформації;
- розрахунок узагальнюючих показників і проведення на їх основі комплексного порівняльного аналізу;
 - аналіз закономірностей розподілу;
 - аналіз взаємозв'язків;
 - аналіз закономірностей динаміки.

Агрегування інформації передбачає організацію її збирання, накопичення та сортування за класифікаційними та ієрархічними ознаками.

Розрахунок узагальнюючих статистичних показників має за мету оцінити результати спостережень з визначенням характеристик розподілу та можливим порівнянням рядів даних.

Аналіз закономірностей розподілу передбачає оцінку структури і структурних зрушень, варіації і концентрації явищ.

Для аналізу взаємозв'язків здійснюється оцінка щільності та перевірка істотності зв'язків.

При аналізі динаміки виконується індексний моніторинг та вивчається інтенсивність динаміки і тенденції розвитку.

Практичне використання інформаційних технологій при аналізі даних

Засоби аналізу даних на комп'ютерах

Статистика дозволяє:

- компактно описати дані;
- зрозуміти структуру даних;
- провести класифікацію;
- визначити закономірності у хаосі випадкових подій і явищ.

Зараз розроблена велика кількість програмного забезпечення, яке у тій чи іншій мірі може бути використано для статистичного аналізу і обробки даних. Таке програмне забезпечення можливо поділити на три категорії за ступенем складності і номенклатурою вирішуваних завдань:

1. Універсальні електронні таблиці.
2. Універсальні математичні пакети.
3. Спеціалізовані програми і пакети для статистичної обробки і аналізу даних.

Аналіз можливостей табличного процесора

Комп'ютер дозволяє представляти їх в електронному вигляді, а це надає можливість не лише відображати, але й обробляти дані. Клас програм, що використовується задля цієї мети, називається *електронними таблицями*.

Термін "електронна таблиця" використовується для позначення простої у використанні комп'ютерної програми, яка призначена для обробки даних. Особливість електронних таблиць полягає в можливості застосування формул для визначення взаємних зв'язків між значеннями різних комірок. Розрахунок за заданими формулами виконується автоматично. Зміна змісту будь-якої комірки призводить до перерахунку значень усіх комірок, які пов'язані формульними співвідношеннями і, таким чином, до оновлення всієї таблиці у відповідності зі зміною даних.

Найбільш широке застосування з високою ефективністю електронні таблиці знайшли при вирішенні таких типів задач:

- Проведення однотипних розрахунків над великими масивами даних;
- Автоматизація підсумкових обчислень;
- Розв'язання задач шляхом підбору значень параметрів, табулювання формул;
- Обробка результатів експериментів;
- Проведення пошуку оптимальних значень параметрів;
- Підготовка табличних документів;
- Побудова діаграм і графіків за існуючими даними.

В програмі обробки електронних таблиць використовується декілька десятків статистичних функцій, які призначені для вирішення наступних завдань:

- визначення чисельних характеристик;
- генерування рядів значень, що відповідають певному закону;
- статистичне оцінювання міри відповідності;
- елементарний аналіз;
- нескладне прогнозування;
- допоміжні функції.

Пакет аналізу має дозволяє виконувати наступні операції з обробки даних:

1. *Однофакторний дисперсійний аналіз* - використовується для перевірки гіпотези про подібність середніх значень двох чи більш вибірок, що належать однієї і тієї ж генеральної сукупності.

2. *Двохфакторний дисперсійний аналіз з повтореннями* - являє собою більш складний варіант однофакторного аналізу та включає більш ніж одну вибірку для кожної групи даних.

3. *Двохфакторний дисперсійний аналіз без повторення* - представляє собою двухфакторний аналіз дисперсії, що не включає більш однієї вибірки на групу.

4. *Кореляційний аналіз* - використовується для кількісної оцінки взаємозв'язку двох наборів даних, представлених у безрозмірному вигляді.

5. *Коваріаційний аналіз* - використовується для обчислення середнього добутку відхилень точок даних від відносних середніх. Коваріація є мірою зв'язку між двома діапазонами даних.

6. *Описова статистика* – застосовується для генерації одномірного статистичного звіту, що містить інформацію про центральну тенденцію і мінливість вхідних даних.

7. *Експонентне згладжування* - призначається для пророкування значення на основі прогнозу для попереднього періоду, скоректованого з урахуванням погрешностей у цьому прогнозі.

8. *Двохвибірковий F-тест для дисперсії* - застосовується для порівняння дисперсій двох генеральних сукупностей.

9. *Двохвибірковий t-тест з однаковими дисперсіями* - служить для перевірки гіпотези про рівність середніх для двох вибірок.

10. *Двохвибірковий t-тест Ст'юдента* - використовується для перевірки гіпотези про рівність середніх для двох вибірок даних з різних генеральних сукупностей.

11. *Парний двухвибірковий t-тест для середніх* - використовується для перевірки гіпотези про розходження середніх для двох вибірок даних.

12. *Двохвибірковий z-тест для середніх з відомими дисперсіями* - використовується для перевірки гіпотези про розходження між середніми двох генеральних сукупностей.

13. *Гістограма* - використовується для обчислення вибірових і інтегральних частот улучення даних у зазначені інтервали значень, при цьому, генеруються числа влучень для

заданого діапазону осередків.

14. *Ковзне середнє* - використовується для розрахунку значень у прогнозованому періоді на основі середнього значення перемінної для зазначеного числа попередніх періодів.

15. *Генерація випадкових чисел* - використовується для заповнення діапазону випадковими числами, витягнутими з одного чи декількох розподілів.

16. *Ранг і персентиль* - використовується для висновку таблиці, що містить порядковий і процентний ранги для кожного значення в наборі даних.

17. *Лінійний регресійний аналіз* - полягає в підборі графіка для набору спостережень за допомогою методу найменших квадратів.

18. *Аналіз Фур'є* - призначається для рішення задач у лінійних системах і аналізу періодичних даних, використовуючи метод швидкого перетворення Фур'є.

19. *Вибірка* - створює вибірку з генеральної сукупності, розглядаючи вхідний діапазон як генеральну сукупність.