

СТАТИСТИЧНІ ГРАФІКИ

Статистичний графік – це спосіб наочного зображення й узагальнення статистичних даних про соціально - економічні явища і процеси за допомогою геометричних образів, малюнків або схематичних географічних карт.

- **Основні елементи графіка:**
- поле графіка;
- графічні образи;
- масштабні орієнтири;
- експлікація графіка.
- **Поле графіка** – це простір, на якому розташовані геометричні та інші знаки, тобто графічне зображення.
- **Головна задача графіків** - наочне представлення фактів соціально-економічного життя. Вони знаходять широке вживання для популяризації різних статистичних даних і виконують важливу контрольну функцію, виявляючи різного роду помилки і неточності в цифровому матеріалі. Разом з тим, їх вживання має деякі обмеження:
- Графік не може включити стільки даних, скільки їх може містити таблиця.
- Графік часто використовується для зображення загальної ситуації, а не деталей, тому представляє, як правило, округлені, не зовсім точні дані.
- Побудова графіка більш трудомістка ніж побудова таблиці.

- Розмір поля залежить від призначення графіка. Найчастіше застосовують графіки у вигляді прямокутників і квадратів. Співвідношення сторін прямокутників беруть від 1: 1,3 до 1: 1,5 , якщо вертикальну сторону прийняти за 1.
- **Графічний образ** – це сукупність різноманітних геометричних та графічних знаків, за допомогою яких відображають статистичні величини. У статистичних графіках використовують такі геометричні знаки, як крапки, відрізки прямих ліній, квадрати, прямокутники, кола, півкола, сектори, а також негеометричні знаки-символи у вигляді силуетів або малюнків. Це і є основою графіка, його мовою.

- **Масштабні орієнтири** – це масштаб, масштабні шкали і масштабні знаки, які використовуються для визначення розмірів геометричних та інших графічних знаків.
- **Масштаб** – умовна міра переведення числового значення статистичного явища в графічне і навпаки. тобто це довжина відрізка шкали, прийнята за числову одиницю.
- **Масштабна шкала** – це лінія, поділена на відрізки точками відповідно до прийнятого масштабу. Довжину відрізків між сусідніми поділками шкали називають **графічним інтервалом**.
- **Експлікація графіка** – це пояснення, що розкривають його зміст і основні елементи: заголовок (назва) графіка, одиниці виміру, умовні позначення.
- Назва графіка має відповідати на три запитання – “що?”, “коли?”, “де?”.

Класифікація статистичних графіків

- **За загальним призначенням:**

- аналітичні;
- ілюстративні;
- інформаційні.

- **За функціонально-цільовим призначенням:**

- групувань;
- рядів розподілу;
- динаміки;
- взаємозв'язку;
- порівняння.

- **За формою графічних образів:**

- крапкові;
- лінійні;
- площинні;
- просторові ;
- фігурні.

- **За типом системи координат:**

- у прямокутній системі координат;
- у полярній системі координат.

- **За виглядом поля графіка:**

- діаграми;
- картограми;
- картодіаграми;
- центрограми.

- **З огляду на розв'язувані завдання:**

- порівняння статистичних величин;
- структури і структурних зрушень;
- зображення динаміки статистичних показників;
- контролю виконання плану;
- розташування і поширення в просторі;
- варіаційних рядів (тема 1.3.);

- **взаємозв'язку і взаємозалежності**

- **За масштабними шкалами:**

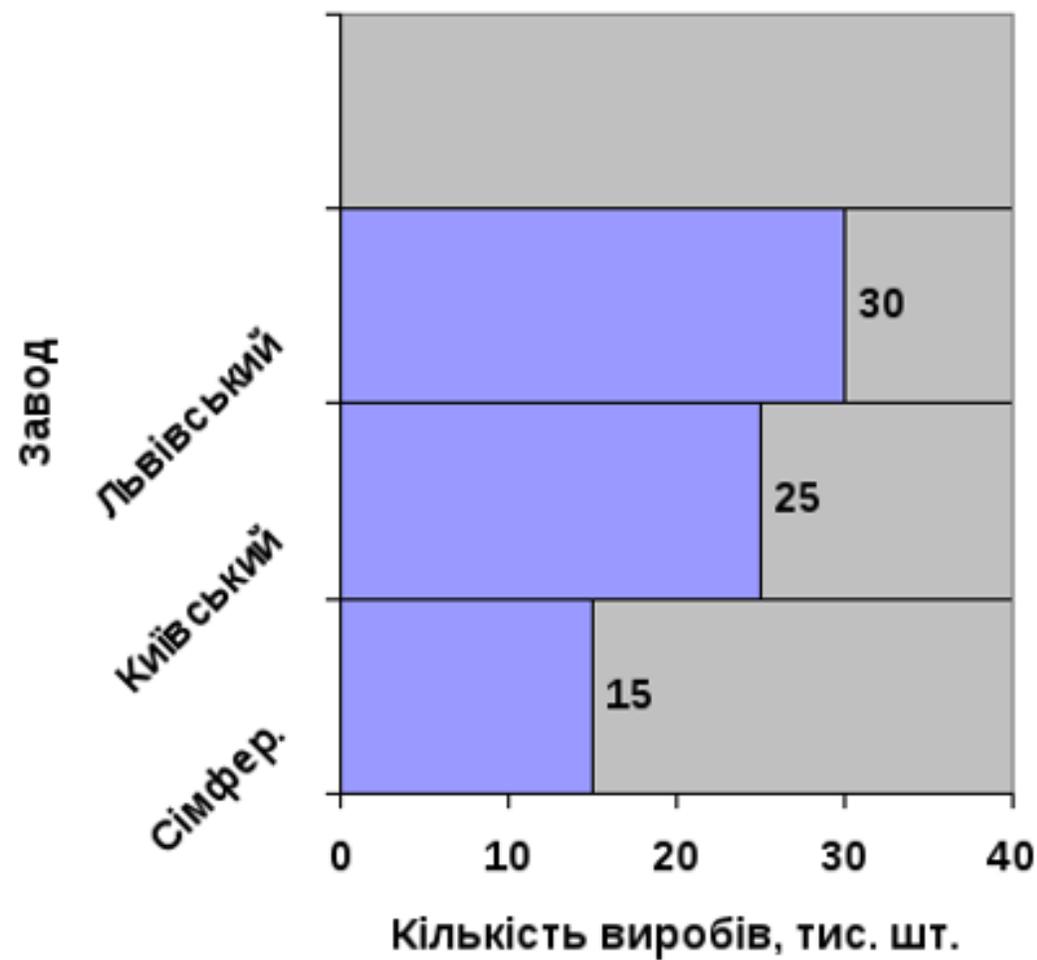
- з рівномірними шкалами;
- з функціональними шкалами;
- з мішаними шкалами.

Графіки порівняння статистичних величин

- Здебільшого для графічного порівняння величин статистичного показника застосовують діаграми.
- **Діаграма** –це вид графіка, в якому цифровим (кількісним) даним відповідають різні геометричні фігури і лінії.
- Залежно від способу зображення статистичних даних вони можуть бути в одному вимірі, коли ці дані зображують у вигляді прямих ліній або смуг однакової ширини, і в двох вимірах (площинні), на яких дані зображують за допомогою площ геометричних фігур.
- До першого виду належать: лінійні, стовпчикові, стрічкові та ін.; до другого – прямокутні (квадратні, “Знак Варзара”), колові, секторні, радіальні, фігурні.

Таблиця. Виробництво телевізорів кольорового зображення заводами України за певний період часу.

Завод	Загальний випуск телевізорів , тис. шт	Середня ціна одного телевізора, гр. од.	Загальна вартість вироблених телевізорів, тис. гр. од.
Львівський	30	650	19500
Київський	25	450	11250
Ужгородський	15	670	10050



Виробництво
телевізорів
кольорового
зображення
заводами України за
певний проміжок
часу.

Завод

Фігурні діаграми - є деякою зміною стрічкової діаграми. В цих діаграмах смуга ділиться на рівні прямокутники, в яких показники зображаються схематичними фігурами, що відображають зміст показника.

Якщо треба графічно представити два показники, один з яких наполовину менше іншого, то менший показник зображається половиною символу (половина автомобіля). На фігурних діаграмах порівнюється не тільки кількість фігур, але і довжина смуг (стрічок).

НАПРИКЛАД, виробництво телевізорів в країні можна зобразити таким чином (1995 рік, тис. шт.).

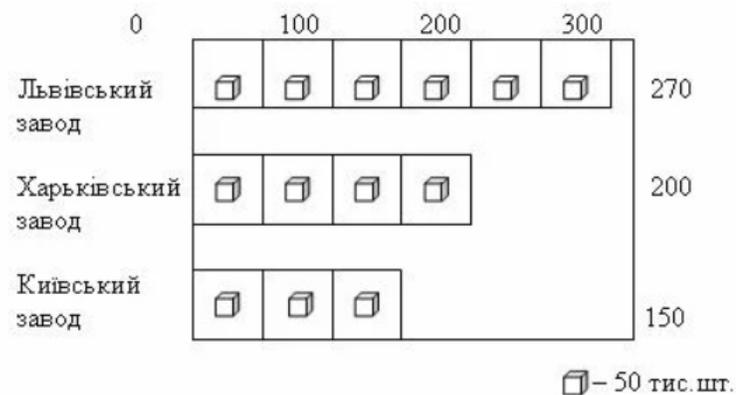


Рис. 4 Приклад фігурної діаграми

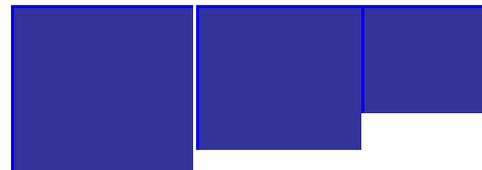
Або таким чином:



Рис. 5 Приклад фігурної діаграми

• **Квадратні діаграми** використовують для порівняння абсолютних величин. Щоб визначити сторони квадрата, потрібно добути корінь квадратний із абсолютного значення, що характеризує явище, у нашому випадку з обсягу виробництва телевізорів.

Для Львівського заводу $\sqrt{30} \approx 5,5$; Київського - $\sqrt{25} = 5$; Ужгородського- $\sqrt{45} \approx 3,9$.
Добираючи масштаб, орієнтуються на найбільше значення. Приймавши масштаб $2=1$ см, визначимо сторони квадратів для заводу Львова – $5,5:2=2,75$ см; Києва – $5,0:2=2,5$ см; Ужгород – $3,9:2=1,95$ см. За цими даними побудуємо квадратну діаграму



Завод: Львівський Київський Ужгородський

Виробництво телевізорів кольорового зображення заводами України за певний проміжок часу.

- **Колові діаграми** своєю площею відображують величину досліджуваного явища. Радіуси кіл пропорційні кореню квадратному значень порівнюваних показників. Отже, щоб знайти радіус, потрібно добути корінь квадратний із абсолютних значень.

• **Стрічкові діаграми.** На відміну від стовпчикових, при побудові стрічкових діаграм, якими зображують розмір явища, розташовують не по вертикалі, а по горизонталі. Вимоги, що ставляться до побудови цього виду діаграм, аналогічні вимогам до стовпчикових діаграм.

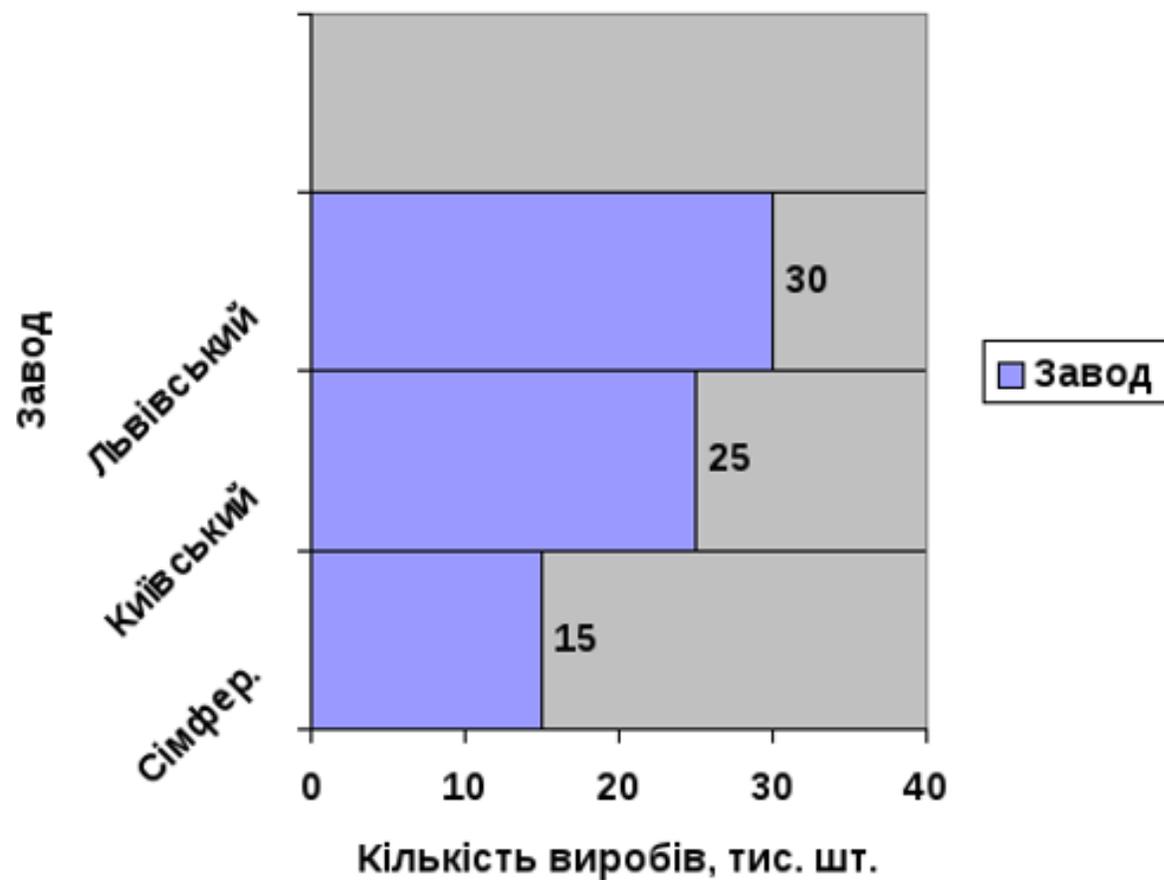
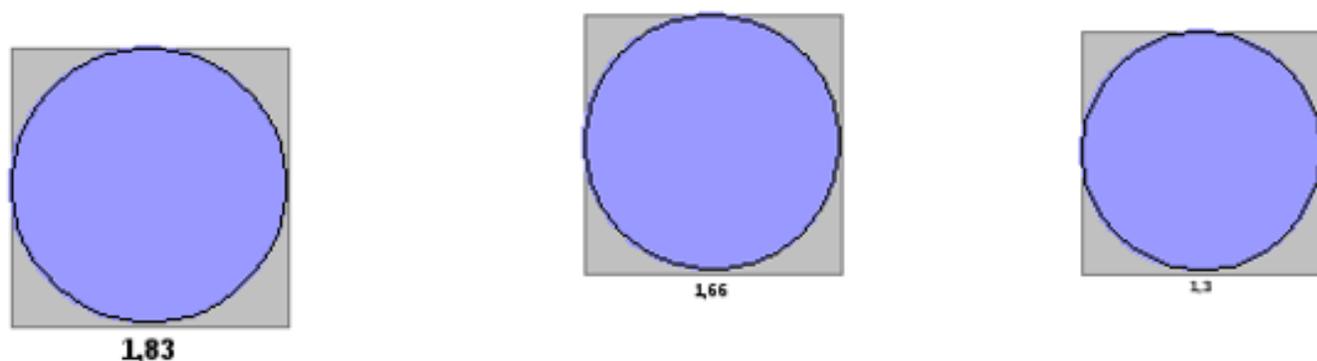


Рис. 1.8.2. Виробництво телевізорів кольорового зображення заводами України за певний проміжок часу.

- **Колові діаграми** своєю площею відображують величину досліджуваного явища. Радіуси кіл пропорційні кореню квадратному значень порівнюваних показників. Отже, щоб знайти радіус, потрібно добути корінь квадратний із абсолютних значень.

Довжина радіуса при масштабі $3=1$ см становить: для заводу Львова : $5,5:3=1,83$ см; Києва – $5,0:3=1,66$ см; Ужгорода – $3,9:3=1,30$ см. Після того, як визначили довжини радіусів, описують кола і дістають відповідну діаграму (рис. 1.8.4.).



Завод: Львівський Київський Ужгородський

Рис. 1.8.4. Виробництво телевізорів кольорового зображення заводами України за певний проміжок часу.

- **Знаки Варзара** застосовуються для характеристики трьох взаємозв'язаних величин. Названі на честь відомого статистика професори В.Є. Варзара, який їх запропонував. Вони засновані також на порівнянні площ прямокутників, в яких основою є один показник, а висотою прямокутника інший.
- *НАПРИКЛАД*, валовий збір зерна характеризується добутком врожайності на площу. Тоді на осі абсцис відкладають площу, а на осі ординат - значення врожайності.
- Їх добуток і дасть площу прямокутника, що характеризує валовий збір. Це і буде діаграма Варзара.

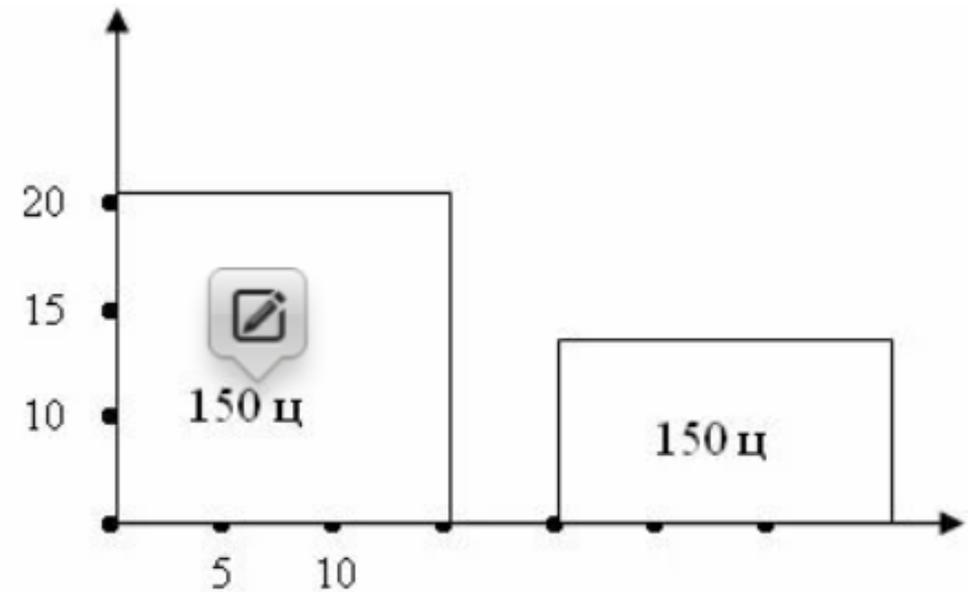


Рис. 7. Приклад знаків Варзара

Зображення структурних явищ і структурних зрушень

Для статистичного дослідження складу сукупності використовують **структурні діаграми** – діаграми співвідношення питомої ваги, які характеризують відношення окремих частин сукупності до загального їх обсягу.

Такі діаграми поділяють на стовпчикові, стрічкові і секторні.

Стовпчикові і стрічкові діаграми застосовують не лише для порівняння величин між собою, а й водночас для порівняння їхніх частин.

Таблиця 1.8.2. Обсяги виробництва м'яса за деякі роки.

Рік	Забійна вага м'яса, тис. тонн				
	Загальна	Яловичина й телятина	Свинина	Баранина й козлятина	М'ясо птиці та інше
2000	4,7	1,9	1,7	0,7	0,4
2010	8,7	3,3	3,3	1,0	1,1
2020	15,0	6,7	5,1	0,8	2,4

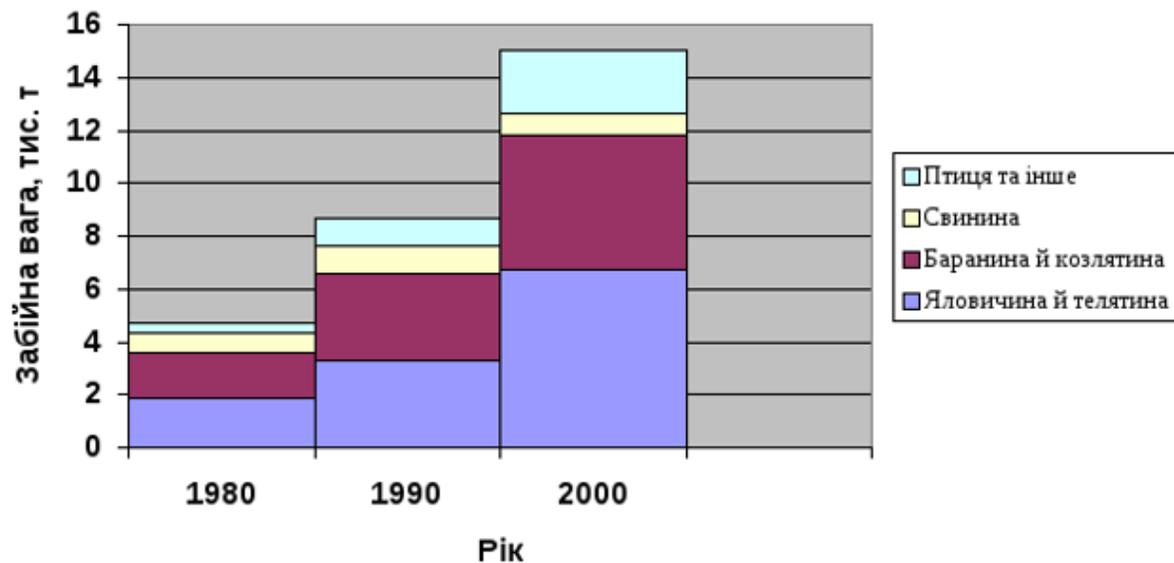
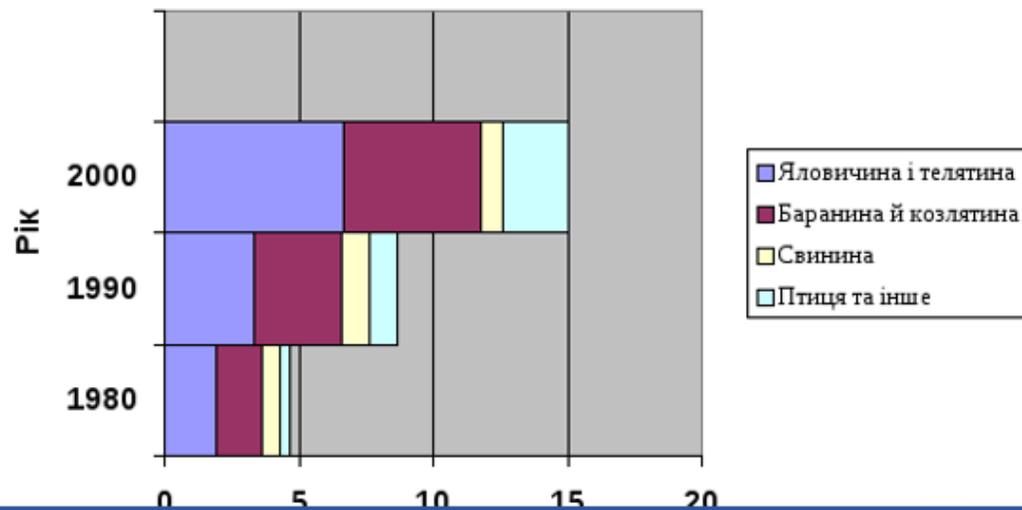


Рис. 1.8.6 (а) Обсяги і структура виробництва м'яса за деякі роки.

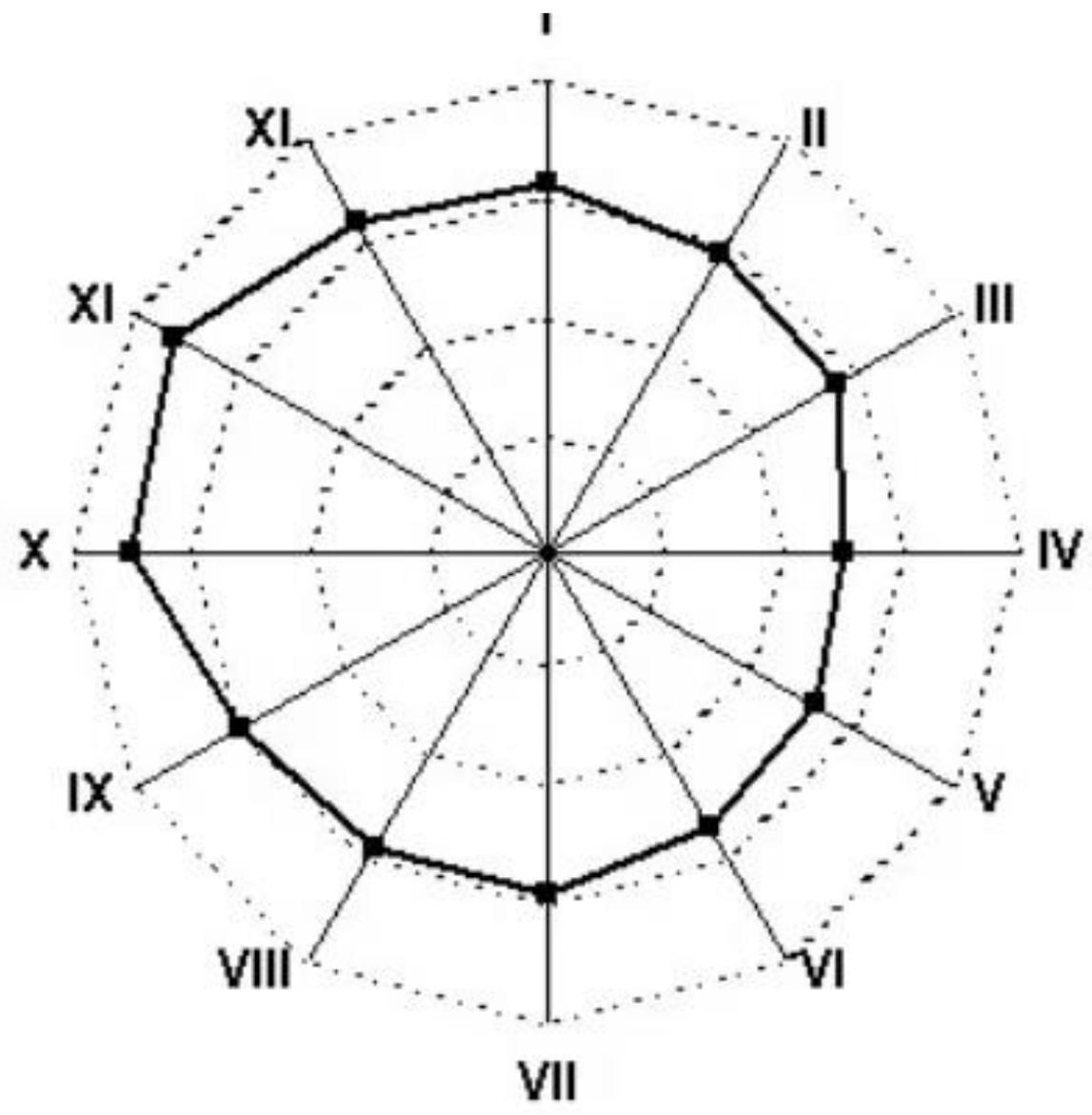
Аналогічно будемо і стрічкову діаграму, тільки в цьому разі масштабна шкала відкладається на осі абсцис, а перпендикулярно до осі ординат малюють смужки (стрічки), які відображають статистичне явище (рис. 1.8.6. б).



- **Секторні діаграми** – це графічні зображення на площі круга, розділеного радіусами на окремі сектори за кількістю різновидів номінальних ознак. Такі діаграми ілюструють структуру і структурні зрушення явищ. На секторних діаграмах можна зображувати частини абсолютних значень явищ або їх процентне вираження.
- Якщо секторна діаграма враховує лише питому вагу частин явища, круги креслять однакових діаметрів. Весь розмір явища приймають за 100 %, розраховують частки окремих його частин у відсотках. Круг поділяють на сектори пропорційно до частин зображуваного цілого.



- **Замкнуті** відображають внутрішньорічні сезонні коливання одного року, **спіральні** - внутрішньорічні коливання за ряд років. Радіальні діаграми можуть бути побудовані двома способами залежно від того, що прийнято як пункт відліку - центр круга (полюс), або коло. Звичайно приймається пунктом відліку центр круга.
- Замкнуті радіальні діаграми будуються таким чином:
 - викреслюється круг, радіус якого в масштабі рівний середньомісячному показнику;
 - потім круг ділиться на 12 частин (місяців) за прикладом годинного циферблата;
 - на кожному з одержаних 12ти радіусів згідно масштабу відкладаються відповідні місячні дані (якщо дані перевищують середньомісячні значення вони відкладаються на продовженні радіусів за межами круга);
 - з'єднавши кінці відрізків на радіусах одержимо замкнуту радіальну діаграму, що характеризує сезонні коливання.
- **НАПРИКЛАД**, виробництво м'яса в одному з регіонів склало в січні 3 100 т, лютому - 2 900 ., вересні - 3 000 т, жовтні - 3 500 т, листопаді - 3 600 т, грудні 3 200 т. Разом за рік 36 000 т, а в середньому за місяць - 3 000 т. Тоді:



—■— M=B 1 CM=1000
T

Контрольно-планові графіки

- Графічний метод широко використовують для поточного контролю за виконанням плану. Форми графічного зображення для порівняння планових і фактичних показників досить різноманітні. Основні види графіків:
- лінійні графіки виконання плану;
- обліково-планові графіки.
- **Лінійні графіки виконання плану** будуються аналогічно лінійним графікам динаміки, при цьому на одному графіку доцільно показати не тільки планові і фактичні показники за звітний період, а й фактичні за минулий рік.
- **Обліково-планові графіки** застосовують для наочного контролю виконання плану одночасно на кількох об'єктах.
- Їх будують на спеціально розграфленій сітці, яка має вигляд таблиці й на якій по горизонталі відкладають одиниці часу (день, п'ятиденку, декаду, місяць, квартал), а по вертикалі - об'єкти дослідження.
- Кожний відрізок по горизонталі відповідає 100 % виконання планового завдання, який у свою чергу, ділиться на п'ять різних клітин по 20 %.
- Ступінь виконання плану для кожного об'єкта зображують двома лініями : тонкою штриховою, яка вказує на ступінь виконання плану за одиницю часу, і жирною суцільною, яка характеризує виконання плану за звітний період у цілому.

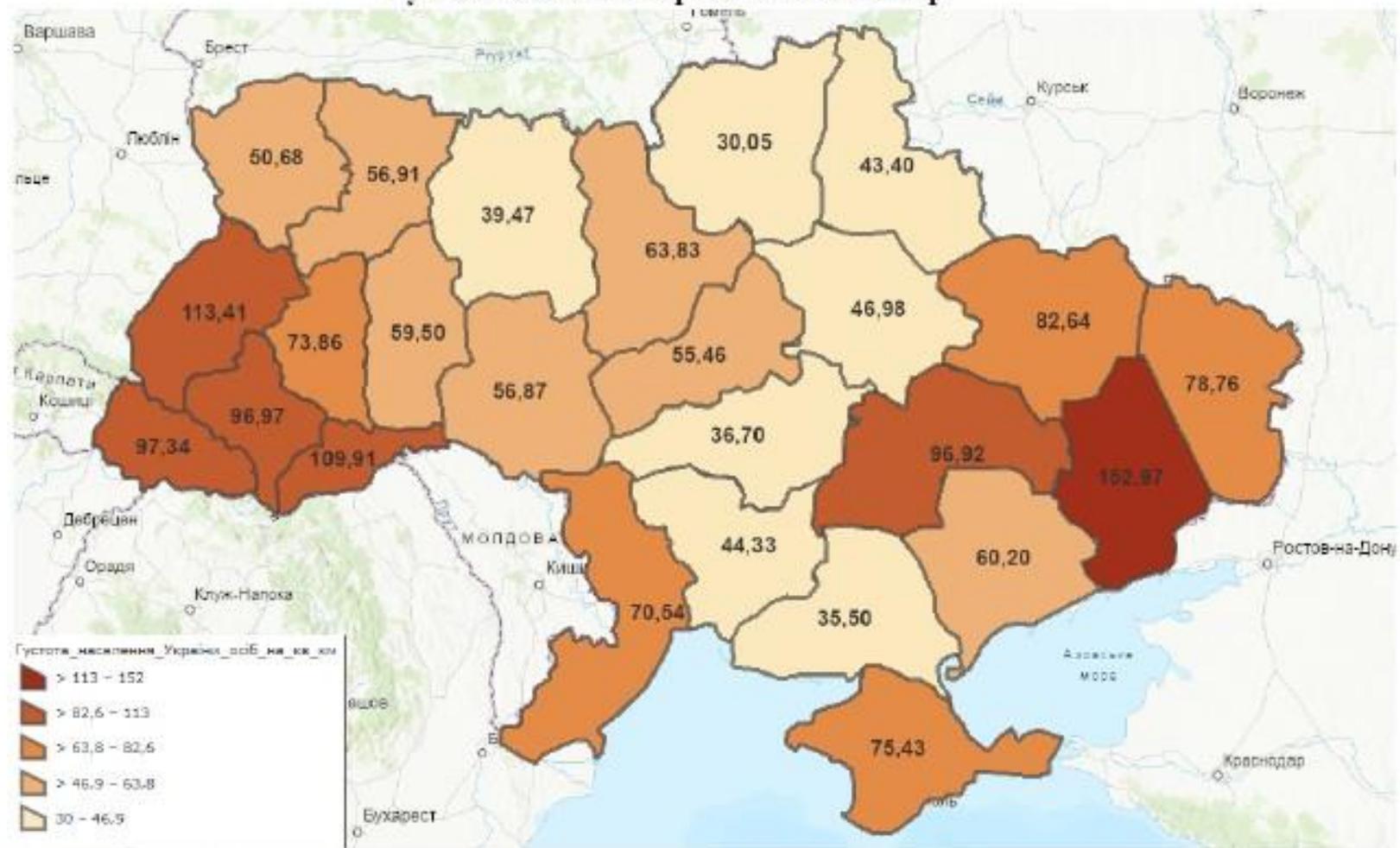
Графічне зображення динаміки статистичних показників

- Зміну статистичних явищ у часі ілюструють динамічні графіки. Динаміку явищ часто описують стовпчикові, стрічкові, квадратні, кругові та фігурні діаграми, в яких кожний стовпчик, стрічка, квадрат зображують обсяг статистичного явища на певну дату або за відповідний період часу.
- Крім зазначених, часто застосовують і лінійні графіки.

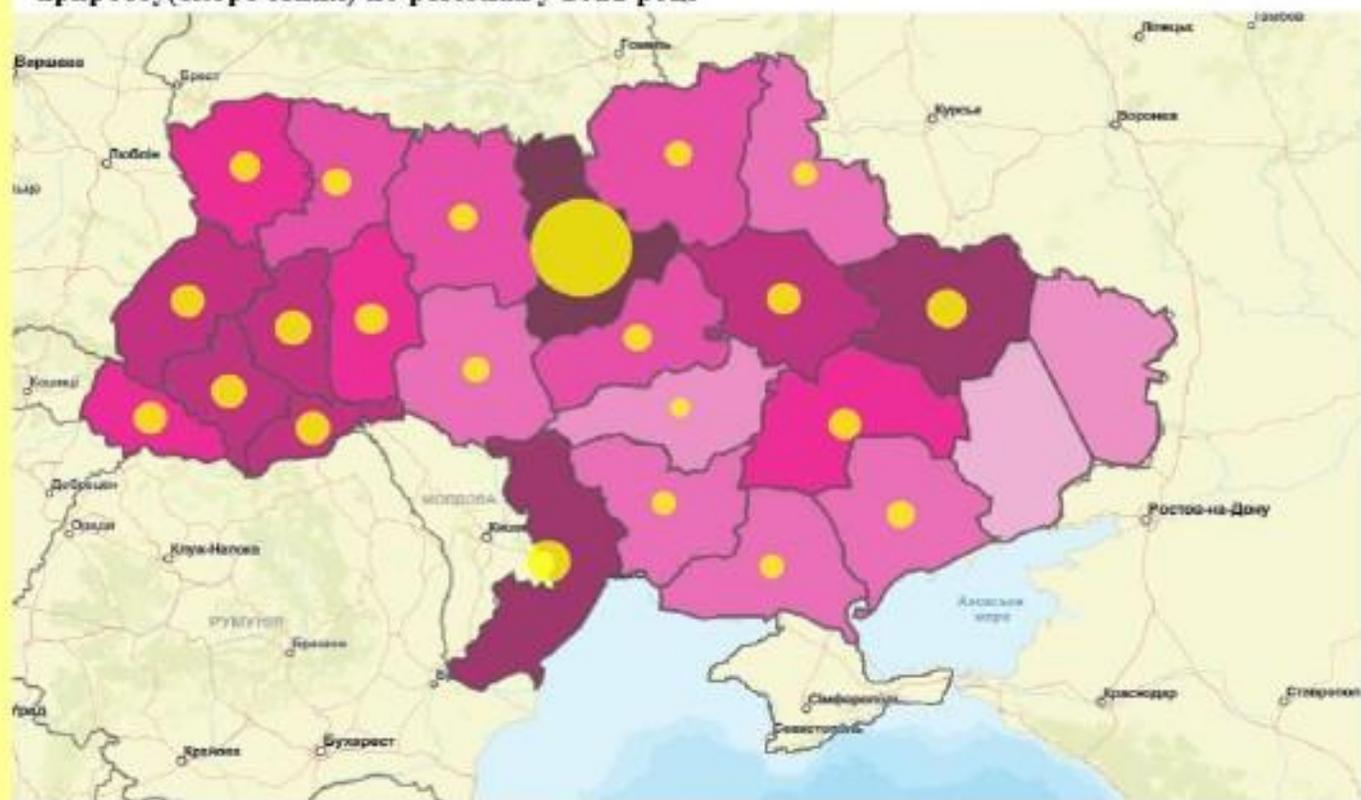
- **Лінійні графіки** характеризують зміну явищ у часі, виявляють залежність між двома показниками тощо. Їх будують за допомогою прямокутної системи координат, на осі абсцис якої розташовані характеристики часу, а на осі ординат – рівні динаміки явища.
- До побудови лінійних графіків ставлять такі вимоги:
- графік повинен читатися по горизонталі зліва направо, по вертикалі – знизу вверху;
- на осі ординат обов'язково позначається нульова величина. У випадках, коли дотримання цього правила пов'язане зі значним зменшенням масштабу та погіршенням наочності, слід зробити розрив по всіх ординатах (при цьому нульова лінія зберігається);
- відрізки на осі абсцис повинні відповідати інтервалам (для рядів динаміки – періоду часу);
- нульова лінія повинна різко відрізнятися від інших паралельних ліній;
- при побудові графіку із застосуванням процентної шкали треба чітко виділити лінію, яка означає 100%;
- крива лінія діаграми повинна різко відрізнятися від лінії сітки;
- цифрові показники розміщують таким чином, щоб їх легко можна було прочитати;
- площа графіка повинна бути квадратною або прямокутною.
- Лінійні графіки дають можливість наочно визначити періоди часу, коли явища зростали (зменшувались) більш чи менш інтенсивно, або лишалися без змін.

- **Картограма** – це схематична географічна карта, на якій розподіл зображуваних явищ по території подається за допомогою відповідних графічних і тонових символів (штриховки, крапки тощо). Залежно від використовуваних символів розрізняють фонові, ізолінійні і крапкові.
- **Картограма** – це поєднання схематичної географічної карти із діаграмою. Головна їх особливість полягає в тому, що ці явища відображають на контурній географічній карті за допомогою спеціальних знаків – символів у вигляді стовпчиків, квадратів, кругів тощо. Розмір геометричного знаку відповідає обсягу даного явища у певному регіоні.
- **Центрограма** – це контурна карта, на якій розміщують короткі цифрові таблиці з інформацією про історико-географічний розвиток і розташування досліджуваного явища чи процесу. Центрограми ще називають історико-географічними картами, бо вони дають змогу наносити цілі статистико-географічні списки для різних територій у вигляді цифрових (кількісних) рядів на карті і в такий спосіб наочно простежити окремі аспекти динаміки досліджуваного процесу.

Густота населення України на 01.01.2022 р.



Міграційний приріст(скорочення) населення та загальний коефіцієнт міграційного приросту(скорочення) по регіонах у 2021 році



Області України

Міграційний_приріст_скорочення_на_1000_осіб_напрямого_населення



Області України

Міграційний_приріст_скорочення_осіб



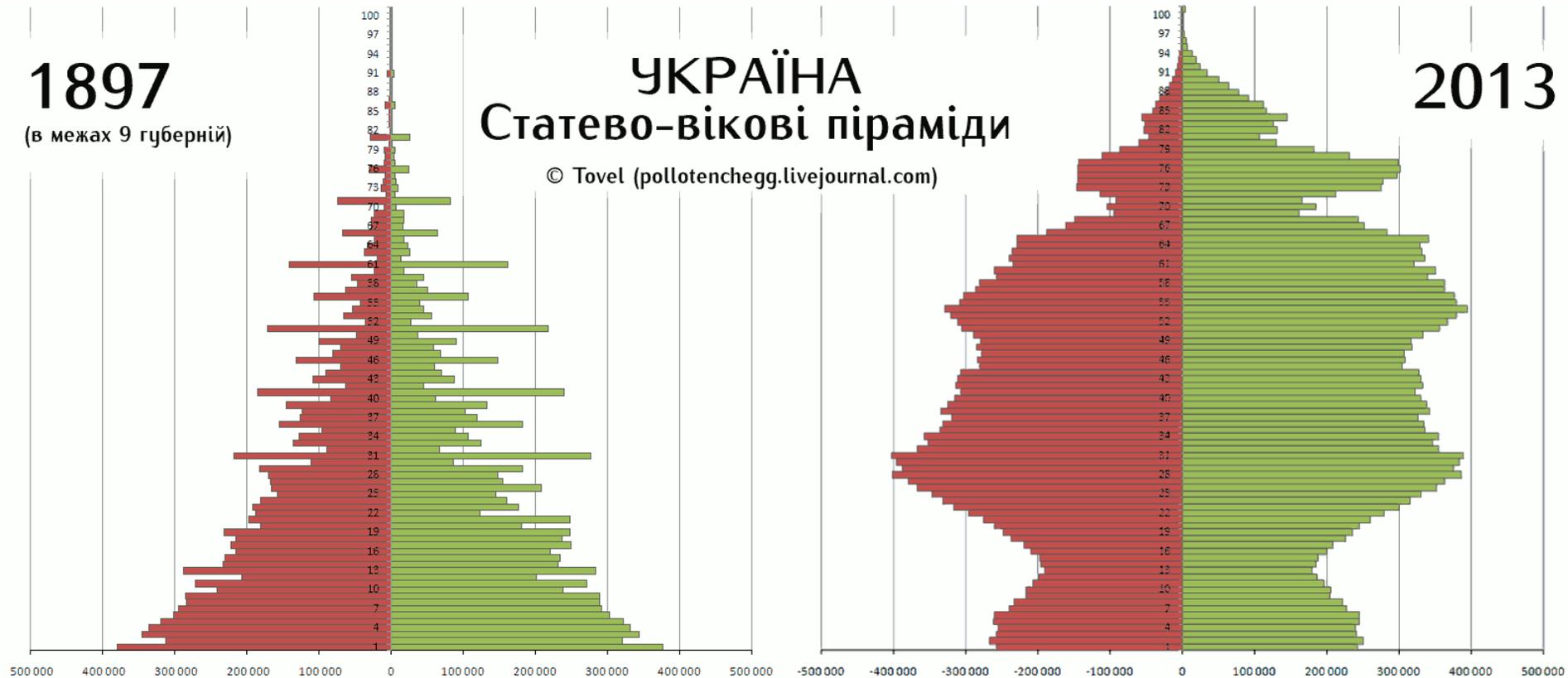
Візуалізація

- Таблична форма представлення даних годиться для їх зберігання і обробки, але вже на фазі аналізу ми використовуємо інші, графічні представлення даних, наприклад діаграми розсіяння, графіки, гістограми та стовпчасті діаграми. Графічна форма полегшує сприйняття інформації як цілого, вочевиднює її особливості, тенденції та аномалії — що цікаво не лише аналітикові, а й остаточному споживачеві.
- Але кінцеві візуалізації від робочих з процесу аналізу відрізняє не лише більша увага до охайності і естетичності. Якщо в процесі аналізу ми *шукаємо* закономірності й цікаві виники, в процесі кінцевої візуалізації ми *(як правило) наперед знаємо, що хочемо показати*. Ба навіть більше — ми знаємо, *які емоції ми хочемо викликати у споживача, що у наявних даних має його вразити*.
- Крім того, якщо робочі візуалізації зрозумілі для людини, що підготована до сприйняття даних, перебуває в певному контексті роботи з ними, то кінцеві візуалізації мають *занурювати у контекст* непідготовану людину.
- Методи розв'язку цих задач різноманітні, і частина з них виходить за межі візуалізації даних. Для занурення непідготовленої людини в контекст, виділення в даних «героя», тощо, найчастіше використовуються засоби **інфографіки**, яка, в свою чергу, може як включати елементи візуалізації даних, так і обходитися цілковито без них, *ілюструючи* певний набір даних, а не *відображаючи* його.

Задачі візуалізації

- В першу чергу, візуалізація дозволяє охоплювати великі обсяги інформації, наче стискаючи її, роблячи компактною. Так само, вона дозволяє зробити доступнішим сприйняття складної інформації, пошвидшуючи порівняння величин і полегшуючи виявлення паттернів в даних.
- Важливою рисою візуалізованих даних є їхня переконливість, тому дуже важливо уникати викривлення інформації в процесі візуалізації. Зокрема, в хорошій візуалізації наочність загальної картини не заважає сприйняттю деталей.
- Але головне, що вирізняє візуалізацію як інфопродукт від використання візуалізації як техніки аналізу — *чітке повідомлення*. Якщо в даних нема нічого такого, що хотілося би в них показати сприймачеві, жодна візуалізація не зможе зробити їх цікавими.
- Відповідно, побудові візуалізації має передувати момент осяяння, усвідомлення в даних чогось такого, що вони промовляють, чим варто поділитися. Це може бути просто масштаб чогось, якась залежність, закономірність, співвідношення, аномалія, тощо. Іншими словами, в даних має з'явитися *герой* і вони мають стати його *пригодю*.
- Далі, треба обрати метод візуалізації, годящий для донесення цього повідомлення. Слід пам'ятати, що увага людини, опрацьовуючи патерни в полі уваги, завжди розділяє патерн на дві сутності — *фігуру* і *тло*, важливе і *те, серед чого воно важливе*. За тривалого споглядання багатого на деталі патерна, можливий розпад його на субпатерни, в кожному з яких так само відбувається знайдення фігури і тла.
- В хорошій візуалізації «головний герой» одразу помітний і сприймається фігурою, а його стосунки з тлом передають чи принаймні відтіняють головне повідомлення. Тому різні повідомлення вимагають різних технік візуалізації.
- Але якою б не була техніка візуалізації, говорити мають самі дані.

Наприклад, віко-статева (популяційна, демографічна) піраміда, стандартний інструмент візуалізації даних про демографічні тенденції популяції та порівняння популяцій, являє собою дві гістограми, поставлені вертикально нулем до нуля.



Неважко помітити, що ця візуалізація дозволяє швидко бачити вікові і статеві дисбаланси в популяції а також миттєво оцінювати співвідношення дітей, людей репродуктивного віку і старих. Для більшості повідомлень на цю тему ця діаграма достатня.

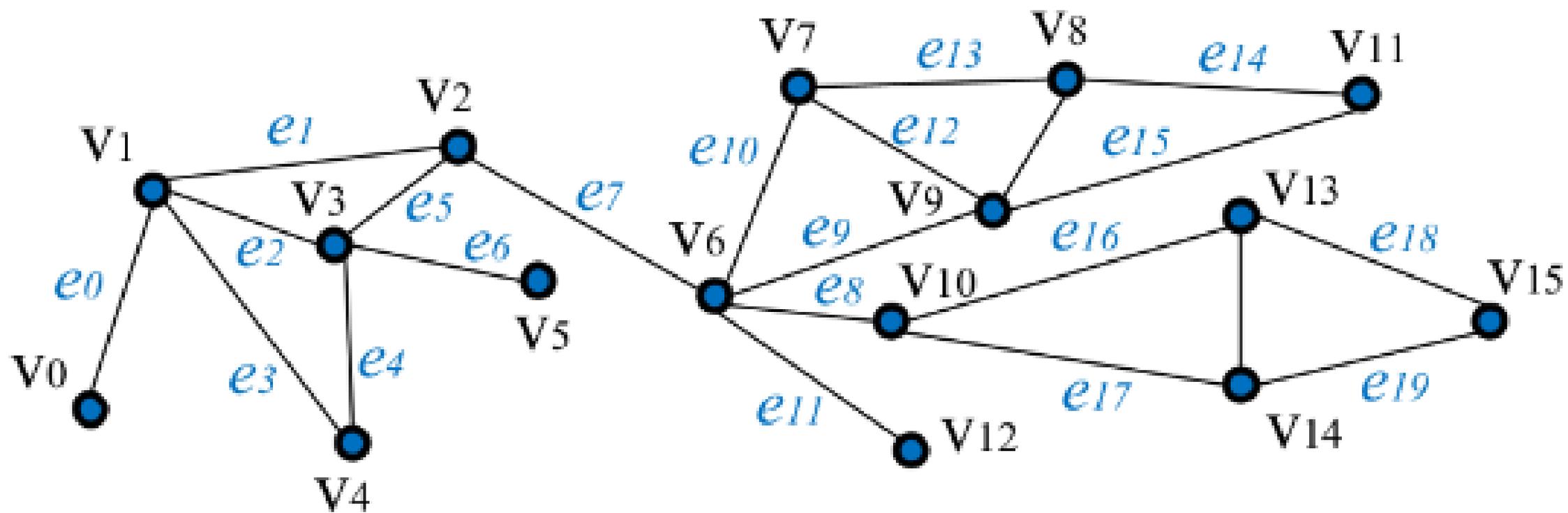
Таблиця як техніка візуалізації

- Говорячи раніше про оптимізацію структури даних, ми вже розглядали таблиці як засіб *візуалізації даних*. Справді, належним чином побудована таблиця може дуже ефективно унаочнювати певні властивості даних. Важко давати загальні поради щодо організації таких таблиць, кожен окремий випадок може потребувати спеціального рішення, але деякі принципи досить-таки прості, хоча не абсолютні.
- **Заголовок містить згадку основної змінної**, іншими словами, якщо у таблиці є «головний герой», він має бути присутній в заголовку.
- **Одиниці виміру в заголовку чи заголовкових комірках у шапці чи боковику**, іншими словами, в комірках мають бути саме значення.
- **Об'єднані комірки в шапці чи боковику** мають відбивати структуру даних.
- **Таблиця не має бути переускладненою та зовеликою**, краще подати декілька простих табличок, а якщо неможливо уникнути значного розміру таблиці, слід дублювати шапку на кожній сторінці і вводити елементи візуального структурування, наприклад, підзаголовки в об'єднаних рядках.
- **Комірки не зобов'язані містити цифри чи слова** — таблиця може бути носієм простих діаграм, графіків і гістограм, крім того, використання піктограм в таблиці може пришвидшити її сприйняття.
- Оформлення таблиці як засобу візуалізації може бути дуже різноманітним, але слід уникати зайвих ліній — достатня відбивка білим простором і лінії, що відокремлюють шапку і боковик від даних. Вирівнювання тексту в таблиці варто робити по лівому краю, але якщо ми маємо багато числових значень, що відрізняються на кілька розрядів, їх варто вирівнювати по правому краю, або, якщо вони є десятковими дробами, за десятковим символом.
- Грамотно і лаконічно побудовані таблиці легко сприймаються і прикрашають сторінку. Якщо око «намагається оминати» вашу таблицю, перевірте, чи все з нею гаразд, чи не переобтяжена вона зайвими графічними елементами, чи достатні відбивки, тощо.

Інші техніки візуалізації

- Існує ще багато технік візуалізації — від досить строгих, як Ейлерові кола та споріднені з ними діаграми Венна, що використовуються для унаочнення стосунків множин і їх перетинів.
- Для демонстрації структур і залежностей використовуються *графи* і *організаційні діаграми*. Для деяких візуалізацій доречно використати не декартові координати, а, наприклад, полярні.





Стовпчаста діаграма

Column Chart

Embedded Charts with Subcomponents

Що ми хочемо показати?

Розподіл

Одна змінна

Небагато точок даних / категорій / інтервалів

Стовпчикова гістограма

Column Histogram

Розподіл

Одна змінна

Багато точок даних

Лінійна гістограма

Line Histogram

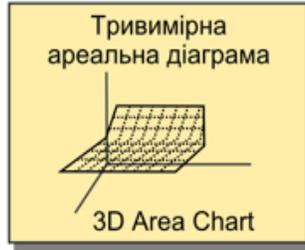
Розподіл

Дві змінні

Діаграма розкиду

Scatter Chart

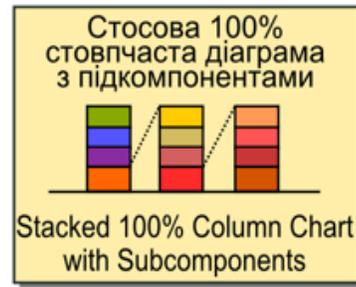
Розподіл Три змінні



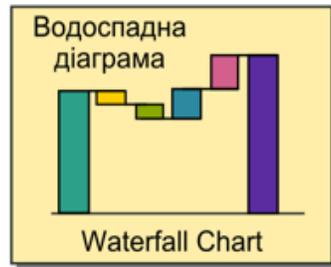
Композицію (будову) Статичну Проста частина в цілому



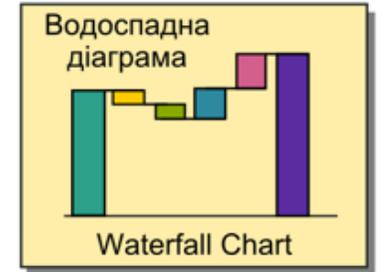
Композицію (будову) Статичну Компоненти компонентів



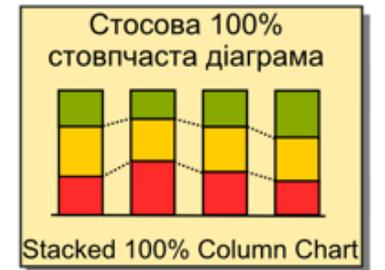
Композицію (будову) Статичну Накопичення додавань і віднімань в цілому



Композицію (будову) Статичну Накопичення додавань і віднімань в цілому

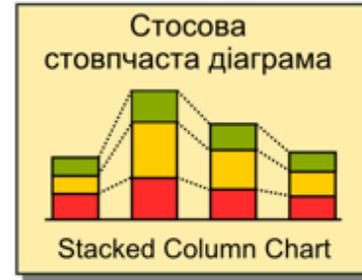


Зміну протягом часу періодів Небагато Важливі тільки відносні відмінності.



Зміну Небагато Важливі
протягом часу періодів абсолютні,
відносні,
відмінності

і
і



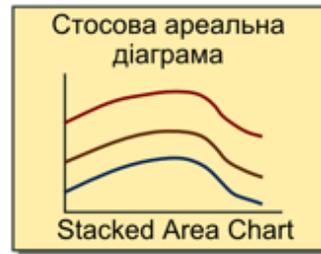
Зміну Багато Важливі
протягом часу періодів відносні
відмінності

тільки

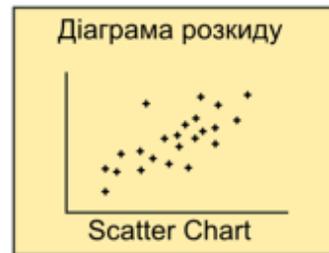


Зміну Багато Важливі
протягом часу періодів абсолютні,
відносні
відмінності

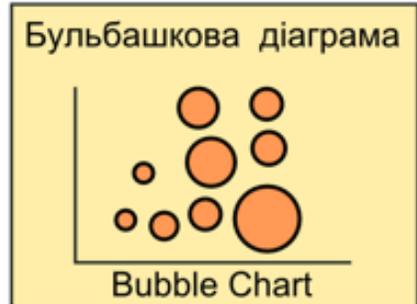
і
і



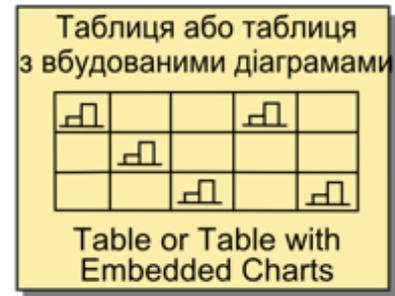
Зв'язок чи Двох
залежність змінних



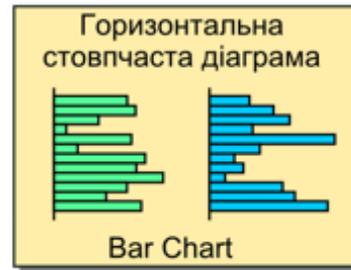
Зв'язок чи ір'юх
залежність змінних



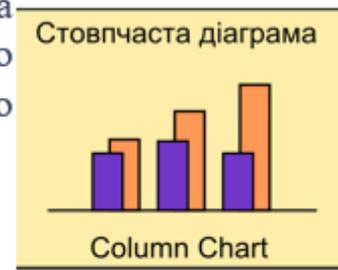
Порівняння Між Одна змінна на
об'єктами об'єкт, багато
категорій



Порівняння Між Одна змінна на
об'єктами об'єкт, мало
категорій, багато
об'єктів



Порівняння Між Одна змінна на
об'єктами об'єкт, мало
категорій, мало
об'єктів



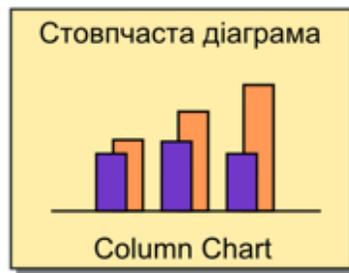
Порівняння В часі Багато періодів —
циклічні дані



Порівняння В часі Багато періодів —
ациклічні дані



Порівняння В часі Мало періодів,
одна чи декілька
категорій



Порівняння В часі Мало періодів,
багато категорій

