

Крвавич Н.Б.

студент,

Національний університет «Львівська політехніка»

РОЛЬ ВІЗУАЛІЗАЦІЇ ДАНИХ У СИСТЕМНОМУ АНАЛІЗІ

Візуалізація даних – це представлення інформації у графічній формі, яка є кращою для сприйняття та опрацювання. Людям важко опрацьовувати й аналізувати інформацію, якщо вона є у вигляді числових даних. В такому випадку найкращим виходом для спрощення роботи та підвищення її ефективності є інфографіка.

Інформаційна графіка або інфографіка (англ. Information graphics; infographics) – це графічне візуальне подання інформації, даних або знань, призначених для швидкого та чіткого відображення комплексної інформації. Вона може покращити сприйняття інформації, використовуючи графічні матеріали для того, щоб підвищити можливості зорової системи людини бачити моделі і тенденції. Процес створення інфографіки можна розглядати як візуалізацію даних, створення інформаційних схем та моделей подання інформації [1].

Візуалізація вже протягом довгого часу набуває щораз більшої популярності. Великим поштовхом для того є її корисність та велика кількість безкоштовних засобів реалізації. Вона використовується у всіх галузях нашого життя. Дуже популярною є в політичній сфері. Зокрема журналісти, часто використовують візуалізацію інформації для кращої ілюстрації даних. На рисунку 1 зображено приклад такої візуалізації.

Є багато типів зорових образів, які можуть бути використані для відображення одного набору даних. Тому вкрай важливо визначити відповідний спосіб візуалізації для набору даних і інфографіки, беручи до уваги графічні функції, такі як позицію, розмір, форму і колір. Початково існує 5 типів візуалізаційних категорій – дані часового ряду, розподіл імовірностей, карти, ієрархії та мережі.

Є багато типів зорових образів, які можуть бути використані для відображення одного набору даних. Тому вкрай важливо визначити відповідний спосіб візуалізації для набору даних і інфографіки, беручи до уваги графічні функції, такі як позицію, розмір, форму і колір. Початково існує 5 типів візуалізаційних категорій – дані часового ряду, розподіл імовірностей, карти, ієрархії та мережі.

Для того, щоб після проведення візуалізації отримати очікуваний засіб для зручного опрацювання даних, потрібно правильно виконати її. Перш за все, треба чітко розуміти для чого потрібно провести представлення даних у візуальній формі, для яких людей це призначено та що потрібно донести до них.

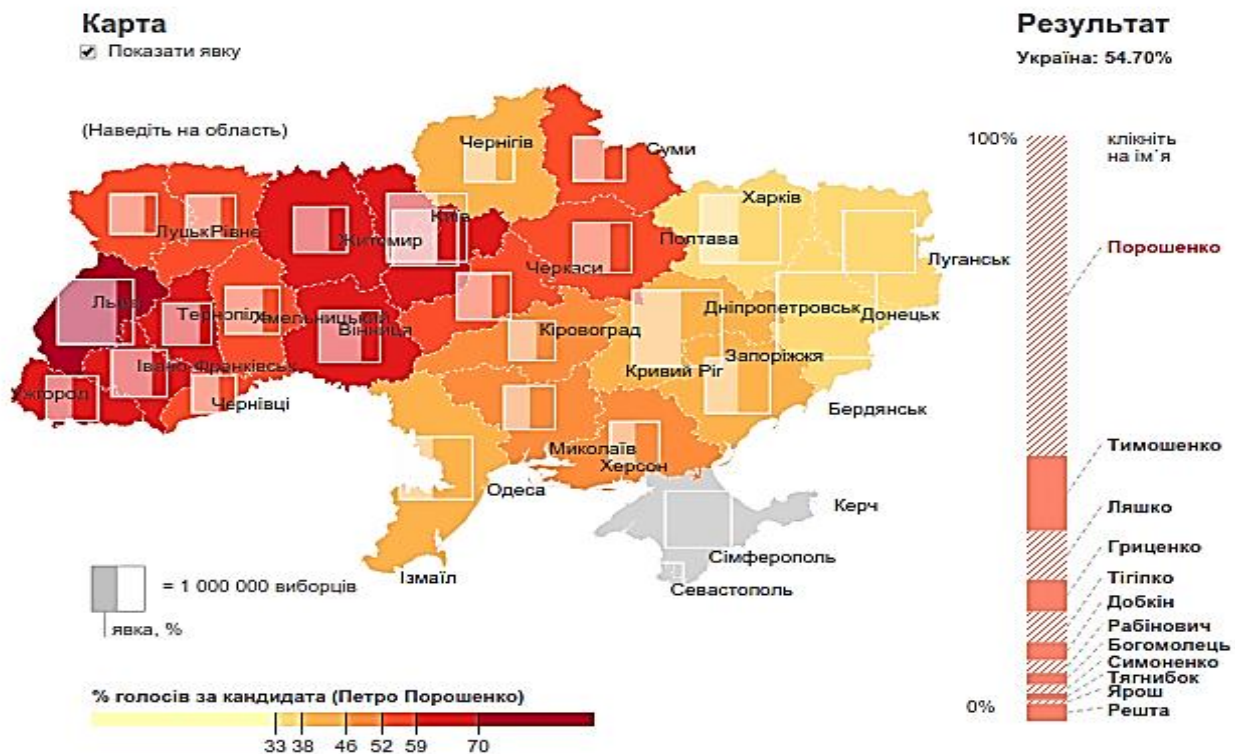


Рис. 1. Результат президентських виборів (травень 2014)

«Купа цифр з ілюстраціями – це не інфографіка. Це просто таблиця, бо ці дані не мають жодного контексту. Не зрозуміло, ці цифри великі чи малі? Візуалізація – це не питання створення красивих ілюстрацій або графіків. Корисне ніколи не відрізняється від красивого – це і є вся суть візуалізації. Вона красива не лише тому, що гарно виглядає, а й тому, що дає важливу інформацію. Крім того, візуалізація повинна бути правдивою», – Альберто Каїро [2].

Наступним кроком після візуалізації даних є аналіз отриманих результатів. Потрібно проаналізувати те в якій мірі було досягнуто очікуваного результату, чи можна побачити якісь цікаві закономірності та що це значить в контексті початкових даних. Іноді, попри гарне зовнішнє відображення, візуальне представлення не може показати нічого корисного, тоді потрібно виконати візуалізацію повторно, обдумавши причини попередньої невдачі [3].

Протягом усього процесу підготовки візуалізації, виконання і аналізу результатів дуже корисним є вести документацію, яка у будь-який час зможе допомогти відслідкувати проходження процесу.

Після проведеного аналізу часто відкриваються нові ідеї розвитку візуалізації. Для подальшої роботи в такому випадку необхідно провести трансформацію даних.

Однією з можливих дій є зміна масштабу. В цьому випадку для того щоб подивитися на певну рису в сукупності візуалізацій можна з'єднати багато елементів в одну групу.

Також можна провести фільтрацію для того щоб відсіяти ці елементи які в даний час не є важливими, тоді аналіз елементів які залишаться буде ефективнішим.

Можна зробити висновок, що візуалізація даних є дуже важливим інструментом для проведення системного аналізу інформації якою ми володіємо. Але для максимальної зручності виконання роботи необхідно правильно підійти до постановки завдання, обдумати мотиви візуалізації та звернути увагу на те, для якої аудиторії людей проводиться дане представлення.

На даний момент існує багато безкоштовних інструментів, які будуть корисними для якісного проведення візуалізації та подальшого аналізу роботи. До таких засобів належать:

- Електронні таблиці – LibreOffice, Excel чи Google Docs.
- Середовище для статистичного програмування – R (r-project.org) чи Pandas (pandas.pydata.org)
- Геоінформаційні системи (GIS) – Quantum GIS, ArcGIS, GRASS
- Бібліотеки візуалізацій – d3.js (mbostock.github.com/d3), Prefuse (prefuse.org), Flare (flare.prefuse.org)
- Інструменти для обробки даних – Google Refine, Datawrangler
- Середовище для візуалізацій, де не використовується програмування – ManyEyes, Tableau Public (tableausoftware.com/products/public) [4].

Важливим є подальше дослідження візуалізації даних, з метою розроблення кращих методів її проведення, які будуть забезпечувати безпомилковість проведення даної роботи з інформацією. Завдяки ві Отже, візуалізація даних – актуальне питання, яке є важливим в галузі системного аналізу.

Список використаних джерел:

1. Вікіпедія «Інфографіка» [електронне джерело]. – Режим доступу: <http://uk.wikipedia.org/wiki/Інфографіка>
2. Школа журналістики УКУ «Мистецтво візуалізації даних, – майстер-клас Альберто Каїро» [електронне джерело], – Режим доступу: <http://journalism.ucu.edu.ua/video/3790/>
3. Colin Ware Information Visualization: Perception for Design / Ware Colin – San Francisco, Morgan Kaufmann Publishers, 2004. – 400 p.
4. Scott Murray Interactive Data Visualization for the Web / Murray Scott – Sebastopol, O'Reilly Media, 2013. – 268 p.