

AZ ADATVIZUALIZÁCIÓ ELVEINEK ÉS FEJLŐDÉSÉNEK ÚJRAÉRTELMEZÉSE A 21. SZÁZADI ADATTUDOMÁNYBAN

REINTERPRETING THE PRINCIPLES AND EVOLUTION OF DATA VISUALIZATION IN TWENTY-FIRST-CENTURY DATA SCIENCE

Sramo Eszter^{0009-0008-9335-5260 1*}, Dr. Oparin-Salamon Melinda^{0009-0006-9082-2922 1†}

Csendes István^{0009-0001-8119-768X 1}

¹ Menedzsment és Jogi Tanszék, Gazdaságtudományi Kar, Neumann János Egyetem, Magyarország

<https://doi.org/10.47833/2026.2.ECO.009>

Kulcsszavak:

adatvizualizáció
data science
adattudomány
statisztika

Keywords:

data visualization
data science
statistics

Cikk történet:

Beérkezett 2025. október 22.
Átdolgozva 2026. április 24.
Elfogadva 2026. május 10.

Összefoglalás

Az adatvizualizáció az adattudomány egyik jelentős pillére, amely hidat képez az adatok feldolgozása és azok megértése között. A tanulmány célja, hogy feltárja az adatvizualizáció történeti gyökereit és elméleti alapelveit, valamint bemutassa, miként alakították ezek a 21. századi adattudomány vizuális kultúráját. A tanulmány áttekinti az adatábrázolás fejlődését, majd vizsgálja a 21. századi digitális fordulat hatásait. A tanulmány kiemeli, hogy az olyan alapelvek, mint az arányosság, az egyszerűség és a torzításmentesség, ma is meghatározzák az feltáró adatelemzés, a modellértelmezés gyakorlatát. A modern data science eszközei ezeket a klasszikus elveket algoritmizált formában alkalmazzák, megteremtve a vizuális gondolkodás és a gépi adatfeldolgozás közös nyelvét.

Abstract

Data visualization is one of the key pillars of data science, serving as a bridge between data processing and understanding. The aim of this study is to explore the historical roots and main principles of data visualization, and to demonstrate how these have shaped the visual practices of twenty-first-century data science. The paper reviews the development of graphical data representation and examines the impact of the digital transformation of the twenty-first century. It highlights that fundamental principles such as proportionality, simplicity, and freedom from distortion continue to define the practice of exploratory data analysis and model interpretation. Modern data science tools apply these classical principles in algorithmic form, creating a common language between visual thinking and machine-based data processing.

† Kapcsolattartó szerző.
E-mail cím: salamon.melinda@nje.hu

1. Bevezetés

Az adatok korában a vizualizáció az emberi megértés egyik legfontosabb eszköze. A statisztikai ábrázolás több száz éves története során az adatvizualizáció folyamatosan alkalmazkodott a társadalmi, technológiai és tudományos környezet változásaihoz. A 18–19. század úttörői (pl. William Playfair, Florence Nightingale) a grafikus ábrázolás segítségével tették láthatóvá az addig nem ismert összefüggéseket. Ezek közös jellemzője, hogy az adatok vizuális formát öltve váltak értelmezhetővé és a meggyőzés eszközeivé. A grafikon az érvelés és a megértés támogatója lett.

A 21. században az adatvizualizáció új dimenziót nyert az adattudomány, vagy más néven data science térhódításával. Az adatok robbanásszerű növekedése, az interaktív technológiák megjelenése és a mesterséges intelligencia elterjedése jelentősen megváltoztatták a vizuális kultúrát. Míg korábban a grafikon nagyrészt a statisztikusok gyorsjelentéseit, hivatalos kiadványait illusztrálta, ma már ezen felül számos foglalkozás eszköztárába bekerült. Közülük kiemeljük a data scientistákat, akik a vizuális eszközök segítségével fedezik fel a mintázatokat, azonosítják az anomáliákat, és kommunikálják a modellek eredményeit. A vizualizáció így nem kiegészítő elem, hanem a data science gondolkodásmódjának alapvető eleme lett.

E folyamatban a klasszikus elvek – arányosság, egyszerűség, kontextus és pontosság – továbbra is érvényesek, ám új formát öltöttek. A vizualizáció ma algoritmizált: a modern szoftverek automatikusan alkalmazzák azokat a szabályokat, amelyeket a 20. század gondolkodói fogalmaztak meg.

Jelen tanulmány célja, hogy bemutassa, miként élnek tovább az adatvizualizáció történeti és elméleti alapjai a data science gyakorlatában. A vizsgálat kiterjed arra, hogyan válik a vizualizáció az feltárázó adatelemzés (EDA) és a modellértelmezés központi elemévé. A cikk arra a kérdésre keresi a választ, hogy a klasszikus adatvizualizációs elvek hogyan alkalmazkodtak a digitális korszak új kihívásaihoz.

2. Az adatvizualizáció fogalma

Friendly [1] szerint az adatvizualizáció a mennyiségi információk grafikus megjelenítése annak érdekében, hogy elősegítse az információk közlését és következtetések levonását. Értelmezésében az adatvizualizáció kettős funkciót tölt be. Egyrészt kommunikációs eszköz, melynek célja az információk hatékony közvetítése. Másrészt kognitív eszköz is, itt a megértés és a következtetések levonása a cél. Összefoglalva az adatvizualizáció nem csupán az adatok ábrázolása, hanem az adatokról való gondolkodás vizuális folyamata. Bertin [2] megfogalmazásában a grafikus ábrázolás az információ olyan vizuális megjelenítése, amely az összehasonlítást, a mintázatok felismerését és a megértést szolgálja. Szerinte az adatvizualizáció információs nyelv. Ezzel szemben Ware [3] az emberi észlelés felől közelít: a vizualizáció a megismerés kiterjesztése. Az adatvizualizáció modernbb definícióját fogalmazta meg: az vizualizáció számítógéppel támogatott, interaktív vizuális reprezentációk alkalmazása az absztrakt adatok megjelenítésére, a kognitív folyamatok erősítése céljából.

3. Az adatvizualizáció története

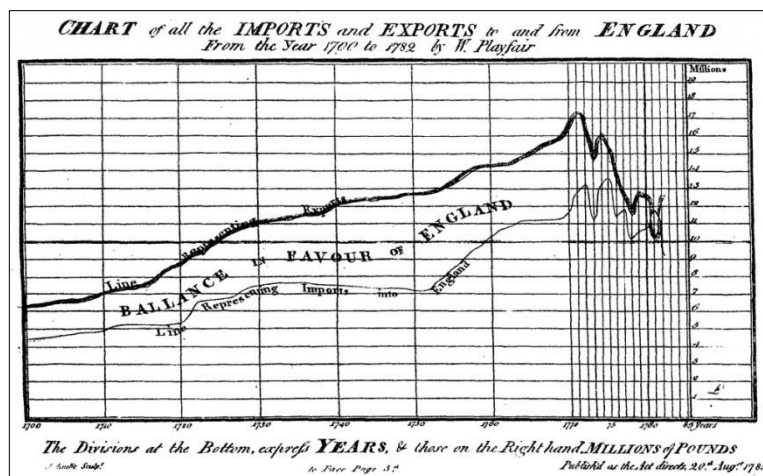
Az adatvizualizáció kezdetei a statisztika és a grafikus kommunikáció közös fejlődéséhez köthetők. A korai ábrák célja az volt, hogy az adatok mögött rejlő mintázatokat érzékletesen mutassák be, elősegítve a tényeken alapuló érvelést.

A kvantitatív információk legkorábbi grafikus ábrázolásai között említhetők meg a csillagok és más égitestek mozgását leíró diagramok, valamint a hajózási térképek. Ezt követően a kezdeti oszlopdiagramok megjelenése emelhető ki, azonban ezek még „a levegőben lógtak”. Ennek oka, hogy a koordináta geometria kialakulása csak jó 300 év múlva kezdődött el. A Descartes-féle koordináta rendszer használatával különböző grafikonok ábrázolása is lehetővé vált. Amellett, hogy a grafikai ábrázolásmód fejlődésnek indult (majd a fejlődés az 1800-as években fokozódott), arra is szükség volt, hogy a statisztika, mint a matematika egyik ága kifejlődjön, amely a diagramokhoz szükséges adatokat rendszerezi és biztosítja. A statisztikai adatok gyűjtése és rendszerezése a 18. századtól a népelességnyilvántartás, hadügy, valamint pénzügy területén voltak jellemzők. [1]

A XVIII. században született William Playfair nevéhez fűződik a statisztikai adatokon alapuló diagramok publikálásának megjelenése [4]. Ő már tisztában volt azzal, hogy miért és hogyan válnak a diagramok hatásos kommunikációs eszközzé. Az 1786-ban megjelent "The Commercial and Political Atlas" című könyvében 44 diagramon mutatta be a brit kereskedelem, az import, az export, a bevétel és a kiadás statisztikai adatait. Közöttük található a Skócia kivitelét és behozatalát összehasonlító ábra, mint a sávdiaagram első alkalmazása. Későbbi munkáiban már kör- és osztott kördiagramok segítségével hasonlított össze adatokat.

A 17-18. században számos jelentős szerző statisztikai grafikái a mai adatvizualizációs és elemzési technikák közvetlen elődjeinek tekinthetők. Munkáik alapját térképészeti, matematikai és statisztikai ismeretek jelentik, és ezt a tudásukat számszerűsíthető adatok vizuális formába öntésére használták fel. Grafikus ábrázolásaik közös tulajdonsága, hogy a hatékony információ-közvetítés érdekében a grafikonok, térképek szerkesztésében akár apróbb torzításokat is alkalmaztak, amennyiben ezek a változtatások hozzájárultak a közvetíteni kívánt tartalom, a lényeg kiemeléséhez.

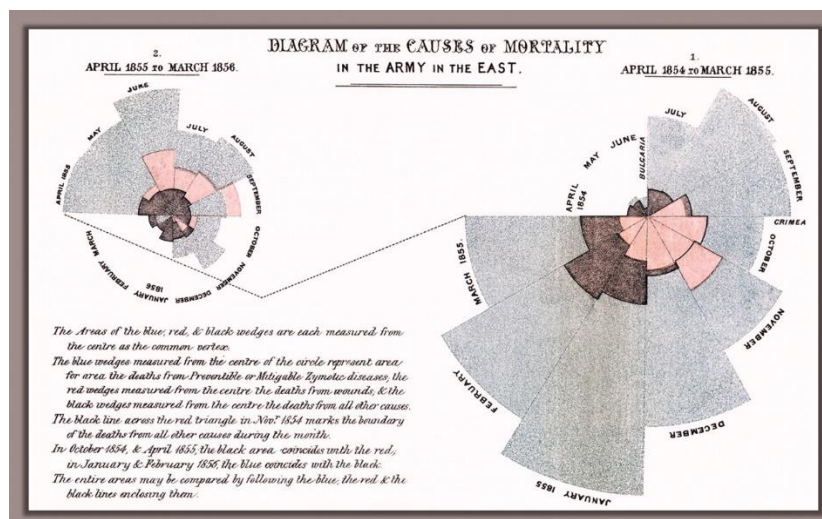
William Playfair szerint a kartográfia és a statisztikai diagramok a vizuális információ megjelenítés szempontjából alapvető hasonlóságokat mutatnak, ezzel az érveléssel támasztja alá, hogy vizuális elemzési eszközökkel bemutatott statisztikáinak gyűjteményét atlasznak nevezi. Ahogy a földrajzban a távolságok, úgy más mértékegységek, például a pénzmennyiség is ábrázolhatók vizuális formában, amennyiben adott pénzmennyiséghez mértani egységet rendelünk. Állítása szerint bármely számmal kifejezhető érték vonalként is felrajzolható [5]. A fentiekre kiváló példa az angol gazdaság exportjának, illetve importjának ábrázolása ugyanazon a diagramon, melyen az egyes évekre vonatkozó értékek vonalként összekötve olyan diagramot eredményeznek, amelyen jól látható az export és az import egymáshoz viszonyított változása az 1700-as években.



1. ábra. Anglia importjának és exportjának diagramja 1700-tól 1782-ig

Az eddig felsoroltakon túl William Playfair nevéhez fűződik a kördiagram, mint vizuális elemzési eszköz alkalmazása is. Playfair továbbá kiemeli, hogy a vizuális eszközök használatával az adatfeldolgozás, illetve az adatok információvá történő átalakításának ideje is jelentősen csökkenthető, ez pedig a későbbi adatvizualizációs eszközök esetében is kulcsfontosságú szempont. Elképzelése szerint a diagramok vizuális nyelvezete univerzális, ezért nyelvi korlátok áthidalására is alkalmas, különböző országok adatainak összehasonlítása így jelentősen egyszerűsíthető.

A kördiagram egyik speciális típusát, az úgynevezett rózsadiagramot, Florence Nightingale vezette be a XIX. század derekán azért, hogy meggyőzze Viktória királynőt a katonai kórházak feltételeinek javításáról. A krími háborúban szolgáló ápolónő, a sebesültek halálozási arányáról készített statisztikai grafikákat. Az általa készített ábrák a kördiagrammal mutatnak hasonlóságot, bár itt a kör 12 cikkelye az egyes hónapokat jelöli az óramutató járásának megfelelően. A körcikkelyek nagysága a halálozások számával arányos. [6]



2. ábra. Nightingale halálozási okokról szóló diagramja

Az adatvizualizáció XX. századi története egy átmeneti korszakot jelent, amikor is a kézzel készített grafikonoktól eljutunk a számítógéppel támogatott vizualizációig. Az adatvizualizáció szempontjából jelentős személyiség volt Bertin, aki elsőként rendszerezte a vizuális változókat (mint a pozíció, méret, szín, orientáció stb.). Ez a munkája a mai adatvizualizációs eszközök alapjait fektette le. [2] A feltáró adatelemzés atyja, John W. Tukey, amerikai statisztikus (akinek a nevéhez fűződik a dobozdiagram bevezetése is) a grafikus módszerek gondolkodástámogató szerepét támogatta. Az ő nevéhez fűződik az adatok hipotézisalkotás előtti vizuális feltárása. A munkássága hozzájárult ahhoz, hogy a data science jelentős eleme a vizualizáció.

A 21. század elejére az adatvizualizáció a statisztika, a számítógépes grafika és az ember-gép interakció határterületein helyezkedik el. Míg a 19–20. században a vizualizáció a statisztikai közlés eszköze volt, addig a digitális korszakban a hangsúly az interaktív felfedezésen, a valós idejű döntéstámogatáson és a kommunikatív narratívákon van. Az adatvizualizáció a korábbi statikus ábrákhoz képest ma dinamikus, többdimenziós és interaktív. Az olyan eszközök, mint a Tableau, Power BI, vagy D3.js, lehetővé teszik, hogy a felhasználó maga fedezze fel az adatstruktúrákat – a vizualizáció nem az elemzés végeredménye, hanem annak a folyamata lett. [3], [7]

4. A vizuális adatelemzés alapelvei

Az adatvizualizáció klasszikus alapelveit Tufte [8] fogalmazta meg. Ezek az egyszerűség, pontosság, arányosság, torzításmentesség és információs hatékonyság. Amíg a 20. században ezek elsősorban grafikai döntések voltak, addig a 21. században kódolási és elemzési döntésekké váltak.

Spence [4] az alábbi vizuális elemzési irányelveket fogalmazza meg:

- Helyes méretarányok, léptékek megválasztása kulcsfontosságú, annak érdekében, hogy a kívánt üzenetet megfelelően közvetítsük.
- Felismeri a kontrasztok szerepét a vizuális érzékelésben: erős, egymástól jól elkülöníthető színeket alkalmaz, de sosem használ 3-4 színnél többet.
- A diagramok vizuális tagolását rácsvonalak alkalmazásával tehetjük meg.
- Rövid magyarázatokat fűzött a diagramokhoz azok számára, akik nem járatosak a földrajzban és matematikai tudományokban. A feliratok és címkék elrendezése is fontos.

A vizuális adatelemzés feladatai közé tartozik a mintázatok keresése a feldolgozandó adatokban, valamint ezen mintázatok vizuális formába öntése a megfelelő ábrázolási mód, illetve diagramtípus kiválasztásával. [9] A vizuális elemzések készítőinek ismerniük és érteniük kell az elemezni kívánt adatokat, azok háttérét, az üzleti oldal elemzési szempontjait és elvárásait ahhoz, az adatokat ennek megfelelően tudják a döntési helyzet kontextusába helyezni. Fontos kérdés, hogy egy vizualizációs felületen milyen mennyiségű információ jeleníthető meg. Ez egyrészt jelenti az egyes diagramokon ábrázolt adat-dimenziók számát, valamint a műszerfal felületen egyszerre

elhelyezett diagramok és elemzési táblák számát és elrendezését. A felhasználói céloknak megfelelő vizuális elemzési táblák elkészítésének és bevezetésének elengedhetetlen feltétele a leendő felhasználók általi tesztelés, a visszajelzések begyűjtése, valamint értékelése, mivel a műszerfal elemek kiválasztása és elrendezése egészen egyedi igényeket követhet. A vizuális adatelemzésben tehát nem egyértelmű szabályok, hanem irányelvek vannak.

A jól strukturált műszerfal felület Eckerson [10] szerint a végletekig letisztult, nélkülöz minden öncélú esztétikai elemet, beleértve a rácsvonalakat és a változatos színhasználatot. Jóllehet, a jelenlegi adat-vizualizációs szoftvereken belül változatos színpaletta áll a felhasználók rendelkezésére, az üzleti elemzési irányelvek ezek óvatos, használatára intenek. Lehetőség szerint egyetlen színcsalád különböző árnyalatit érdemes konzekvens módon használni, egyetlen műszerfal felületen maximum hét színt.

A diagramszerkesztés során kiemelten fontos a léptékek, tengelyek arányainak helyes megválasztása – például a vonal- és a területdiagram esetében – az X illetve Y tengely rosszul megválasztott léptékei jelentősen torzíthatják az adatokból lesűrűzhető következtetéseket. Nem szabad megfeledkeznünk arról sem, hogy az adat-vizualizációnak a szöveges elemek illetve a számok ugyanúgy fontos részei lehetnek, mint a grafikus ábrázolások (Tufte, 2006). Bizonyos esetekben pedig grafikonok helyett kereszt táblák jelenthetik a megfelelő adat-vizualizációs megoldást.

5. A klasszikus elvek továbbélése a data science gyakorlatában

A data science szakemberek számára az adatvizualizáció nem csupán prezentációs eszköz, hanem a gondolkodás vizuális terepe. Az elvek ezáltal új, funkcionális szerepet kaptak. Az adattisztítás és adatfeltárás fázisában a pontosság és arányosság biztosítja, hogy a mintázatok valósak, ne pedig skálázási hibák eredményei legyenek. Ebben a fázisban a kutatók az alakészlelés mentén a csoportosulásokat, trendeket és kiugró értékeket vizuálisan mérik fel. A modellépítés és interpretálás fázisában viszont a modern data science modellek gyakran „fekete dobozként” működnek, és csak a vizualizáció tudja visszaadni az értelmezhetőséget. Itt a pontosság elve a modellek magyarázhatóságában újraértelmeződik: vizuális eszközök is segítenek megmutatni, hogy mely változók vannak hatással az output változóra és milyen irányban.

Az eredmények kommunikációjában a Tufte-féle tisztaság és adat–tinta arány játszik szerepet és a dashboardok és riportok minőségét határozza meg. A jól megválasztott vizualizáció befolyásolja a döntéshozatalt, illetve annak hitelességét.

A XXI. században megfigyelhető az elvek algoritmizálódása is. Az adattudomány egyik sajátossága, hogy a klasszikus vizualizációs elvek beépültek a szoftverekbe. A Python, R és a BI-eszközök fejlesztői automatizálták a Tufte- és Bertin-féle elvek egy részét:

1. Táblázat. Az adatvizualizáció elvei az adattudományban

<i>Klasszikus elv</i>	<i>Adattudományi megfelelő</i>
Pontosság és torzításmentesség	Model interpretability, bias detection
Egyszerűség és információs tisztaság	Minimalista, jól skálázott vizualizációk
Adat–tinta arány	Adat és vizualizációk egyensúlya
Kognitív olvashatóság	Felhasználó központú vizualizációk
Arányosság és kontextus	Megfelelő skálázás, normalizálás

Így az adatvizualizációs szoftver nemcsak eszköz, hanem az emberi vizuális tudás kódolt formája. A felhasználó – akár tudatosan, akár nem – Tufte és Bertin alapelveivel dolgozik minden alkalommal, amikor grafikont hoz létre. [11], [3], [2],[8]

A 2010-es évektől az adatvizualizációt a mesterséges intelligencia és az automatikus grafikon-generálás új szintre emelte. Ezek a rendszerek képesek a nyers adatstruktúrák automatikus vizuális

leképezésére. Így Tufte és Bertin elveit implicit módon kódolják (pl. perceptuálisan helyes színkálák, arányos tengelyek), tehát a klasszikus vizualizációs elvek itt is algoritmizálódnak.

6. Összegzés

Az adatvizualizáció története a statisztikai gondolkodás és a vizuális kommunikáció fejlődésének lenyomata, amely a XXI. században új szerepet kapott az adattudomány gyakorlatában. A tanulmány bemutatta, hogy az adatvizualizáció elvei – mint a pontosság, az arányosság, az információ tisztaság és a torzításmentesség – nem csupán történeti örökségként maradtak fenn, hanem beépültek a modern adatfeldolgozó rendszerek algoritmikus logikájába. A vizualizáció ma már nem a tudományos közlés végpontja, hanem a felfedezés, a modellértelmezés és a döntéstámogatás aktív folyamata.

Irodalomjegyzék

- [1] Friendly, Michael. 2008. „A Brief History of Data Visualization”. In *Handbook of Data Visualization*, by Chun-houh Chen, Wolfgang Härdle, és Antony Unwin. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-33037-0_2.
- [2] Bertin, Jacques. 1983. *Semiology of Graphics: Diagrams, Networks, Maps*. University of Wisconsin.
- [3] Ware, Colin. 2013. *Information Visualization: Perception for Design*. 3d edition. Interactive Technologies. Morgan Kaufmann.
- [4] Spence, Ian. 2006. „William Playfair and the psychology of graphs”. *Proceeding of the American Statistical Association, Section of Statistical graphics*, 2426–36
- [5] Sachs, Jonathan. 2012. „1786/1801: William Playfair, Statistical Graphics, and the Meaning of an Event”. *BRANCH: Britain, Representation, and Nineteenth-Century History*
- [6] Bradshaw, Noel-Ann. 2017. „Florence Nightingale (1820–1910): A Pioneer of Data Visualisation”. In *Women in Mathematics*, szerkesztette Janet L. Beery, Sarah J. Greenwald, Jacqueline A. Jensen-Vallin, és Maura B. Mast, köt. 10. *Association for Women in Mathematics Series*. Springer International Publishing. https://doi.org/10.1007/978-3-319-66694-5_11.
- [7] Heer, Jeffrey, és Ben Shneiderman. 2012. „Interactive dynamics for visual analysis: A taxonomy of tools that support the fluent and flexible use of visualizations”. *Queue* 10 (2): 30–55.
- [8] Tufte, Edward. 2016. „The Visual Display of Quantitative Information (1983)”. In *Diagrammatik-Reader*, szerkesztette Birgit Schneider, Christoph Ernst, és Jan Wöpping. De Gruyter. <https://doi.org/10.1515/9783050093833-035>.
- [9] Otten, Sjoors, Marco Spruit, és Remko Helms. 2015. „Towards Decision Analytics in Product Portfolio Management”. *Decision Analytics* 2 (1): 4. <https://doi.org/10.1186/s40165-015-0013-7>.
- [10] Eckerson, Wayne W. 2011. *Performance Dashboards: Measuring, Monitoring, and Managing Your Business*. 2nd ed. Finance Professional Collection. Wiley. <https://doi.org/10.1002/9781119199984>.
- [11] Shneiderman, Ben. 2020. „Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy”. *International Journal of Human–Computer Interaction* 36 (6): 495–504. <https://doi.org/10.1080/10447318.2020.1741118>.