

## 17. tétel: Vizualizáció

Prósz Aurél

2019. június 9.

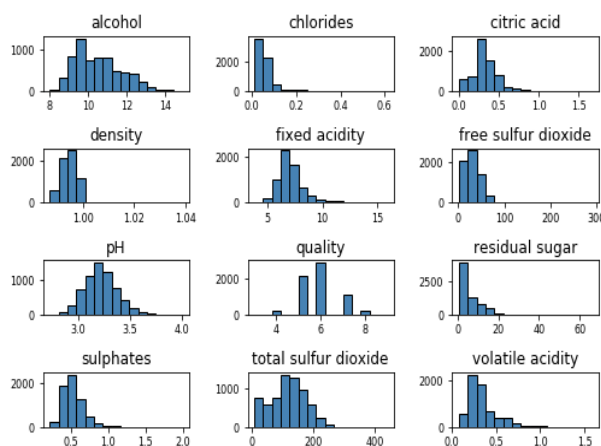
# 1. Tudományos adatok két és háromdimenziós megjelenítésének technikái

Többdimenziós tudományos adatok megjelenítésére számos módszer áll rendelkezésre, a továbbiakban ezeket szedtem össze kis magyarázat és példa keretében.

Tekintsünk egy példa adatkészletet, például az úgynevezett *Wine Quality Data Set*-et, amely letölthető innen: <https://archive.ics.uci.edu/ml/index.php>. Az adatkészlet a portugál *Vinho Verde* bor tulajdonságait írja le, és ezen mutatom be a vizualizációs technikákat.

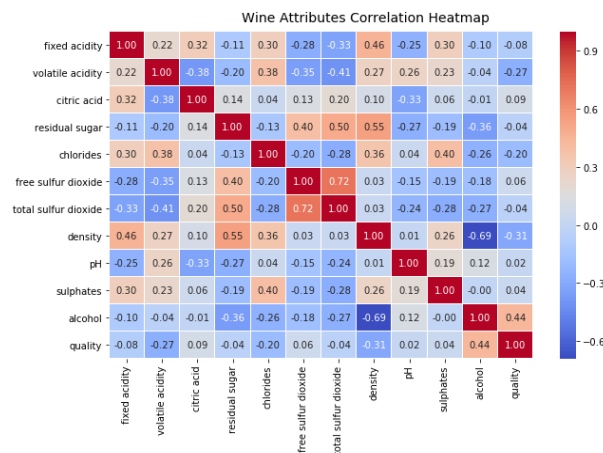
## 1.1. 1D és 2D vizualizáció

Egy dimenzióban, amennyiben egy változó eloszlására vagyunk kíváncsiak, ábrázolhatjuk az adatot **hisztogramként**. Ha egyszerre több változót kívánunk ábrázolni, úgy lehetőség van őket egyenként elrendezni egy rácson.



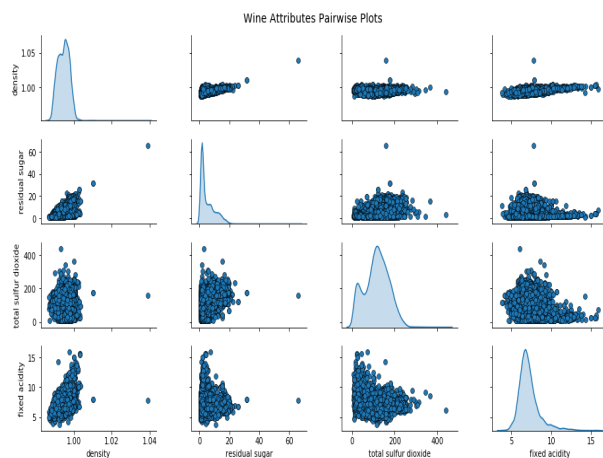
1. ábra. 1D adatok ábrázolása hisztogramként.

Egyszerű módszer az adatokon belüli mintázatok és korrelációk felfedezésére az úgynevezett korrelációs módszer kiszámítása, majd annak vizualizációja **heatmap** formájában. A heatmap esetében az adott párokhoz tartozó értékek jelzik a korreláció, vagy általános esetben bármilyen skalár értékű függvény nagyságát.



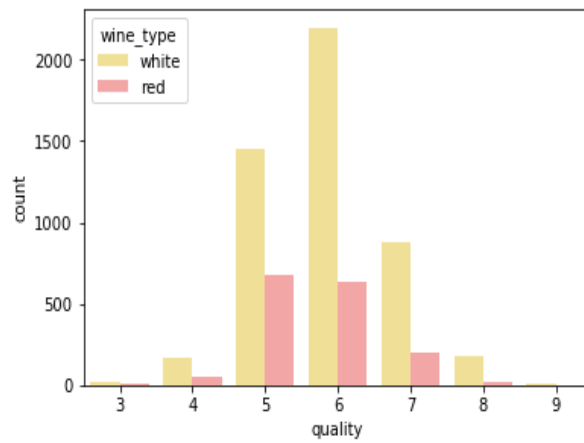
2. ábra. Korrelációs mátrix megjelenítése heatmap formájában.

Két változó közötti összefüggés megjelenítésére alkalmas a **scatterplot**, míg a hisztogramokhoz hasonlóan több változó esetén ezt is megjeleníthetjük egy rácson.



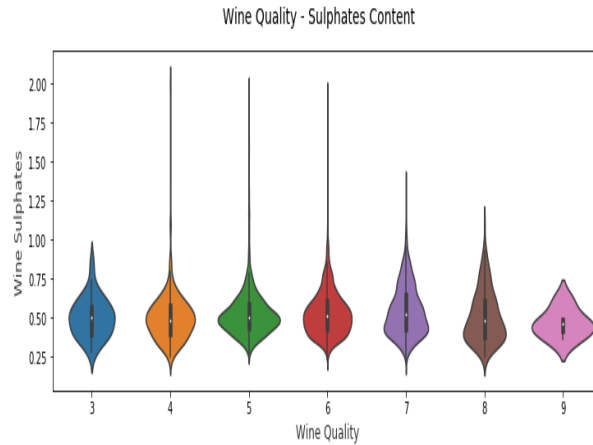
3. ábra. Scatterplotok megjelenítése rácson.

Több dimenziós, kategorikus adat megjelenítésére szolgál az úgynevezett **multiple bar plot**, ahol egy klasszikus bar ploton tudunk megjeleníteni több változót is.



4. ábra. Multiple bar plot két változóval.

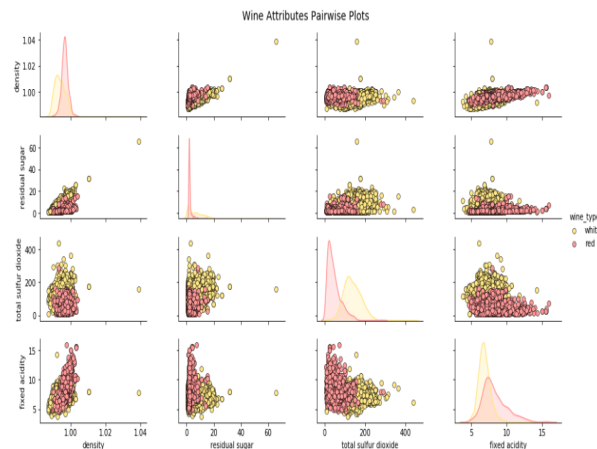
A **box plot** több információt tartalmazó változata a **violin plot**, amely többváltozós adatot vizualizál olyan módon, hogy az eloszlás alakja is leolvasható az outlier adatpontokkal együtt.



5. ábra. Violin plot több változóval és kategorikus színezéssel.

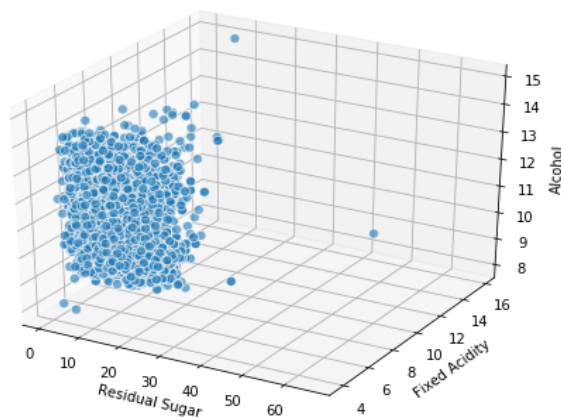
## 1.2. 3D adatvizualizáció

Amennyiben egyszerre 3 változót szeretnénk megjeleníteni, lehetőségünk van alkalmazni a már korábban látott **scatterplot**-ot egy rácson elhelyezve, de színkódolás segítségével újabb változó értékét megjeleníthetjük.



6. ábra. **2D scatterplot** rácson színkóddal.

Ha mind a három változónk folytonos, akkor érdemes az úgynevezett **3D scatterplot**-on ábrázolni, amelyben az adatpontok mélységében tároljuk a plusz információt a kétdimenziós megjelenítéshez képest.



7. ábra. **3D scatterplot**.

Magasabb dimenzióban az eddigiek analógiájára bővíthetjük a megjelenítést újabb színkódolás, forma, méret, vagy egyéb paraméterek bevezetésével.

Bővebb információkért, illetve egzotikusabb vizualizációs technikákért érdemes megtekinteni a következő linket: <https://towardsdatascience.com/the-art-of-effective-visualization-of-multi-dimensional-data-6c7202990c57>

## 2. Skalár-, vektor- és tenzorértékű adatok ábrázolása

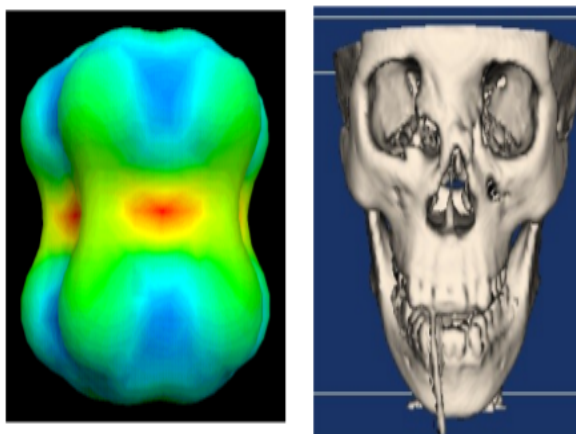
<https://booksite.elsevier.com/samplechapters/9780123875822/9780123875822.PDF>

### 2.1. Skalár értékű adatok ábrázolása

Skalár értékű adatokat az előző részekben tárgyaltakon kívül ábrázolhatunk még felületeken, vagy térfogatokban olyan módon, hogy színekkel jelöljük a konkrét értékeket, ezt nevezzük **col-**

or mapping-nak.

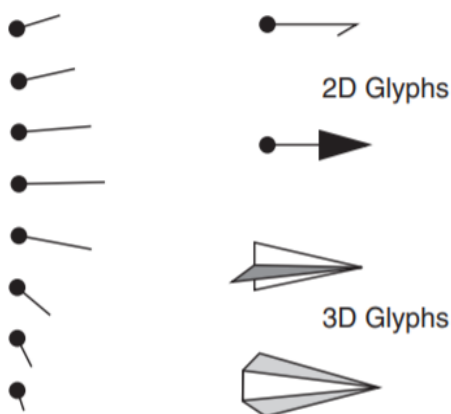
Másik módszer az úgynevezett **contouring**, amely egyfajta kibővítése a color mapping-nak olyan módon, hogy a szinte azonos színű (és értékű) régiókat kontúrvonalakkal (vagy kontúrfelületekkel) elkeríti, és ezeket a területeket (vagy térfogatokat) ugyanolyan színnel színezi. Tipikus példa erre a megjelenítésre az időjárást ábrázoló térképek, ahol a konstans hőmérsékletet jellemző kontúrvonalak választják el az egyes régiókat.



8. ábra. Példa a kontúrvonalakra és a kontúrfelületekre.

## 2.2. Vektor értékű adatok ábrázolása

Vektor értékű adatok ábrázolását a legegyszerűbb módon úgy valósíthatjuk meg, hogy minden egyes adatponthoz társítunk egy vonalat, amelynek meghatározott iránya és nagysága van. A technika neve **hedgehog**, a keletkező vizualizáció kinézete miatt. A megjelenítést tovább lehet finomítani a vonalak színezésével és méretük változtatásával. Ezen kívül az egyszerű vonalak helyett alkalmazhatunk 2D és 3D irányított objektumokat is, amelyek többdimenziós adatok megjelenítésénél lehet hasznos, mert könnyebb a jelölt irányt megállapítani az alkalmazással.



9. ábra. Példa az egyszerű vonalas vektorábrázolásra, illetve a bonyolultabb 2D és 3D objektumok használatára.

A vektor értékű adatok ábrázolásánál figyelni kell, hogy 3 dimenziós ábrázolásoknál sokszor nehéz pontosan megállapítani a vektorok irányítottságát, hiszen a 3D térből vetítjük le a megjelenítést a 2D térbe. Ezen kívül fontos, hogy ne jelenítsünk meg egyszerre túl sok adatpontot, mert a vektorok összemosódhatnak és értelmezhetetlenné válik az ábra.

## 2.3. Tenzor értékű adatok ábrázolása

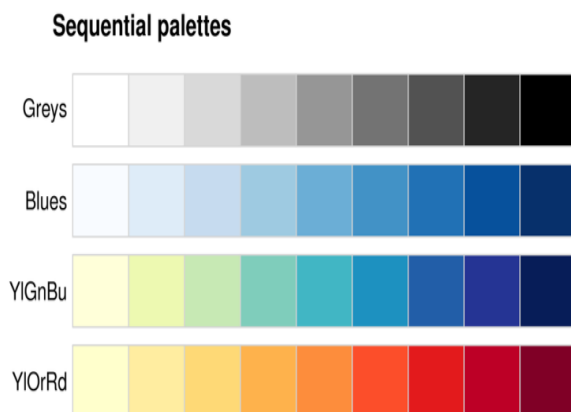
TODO

## 3. Színek

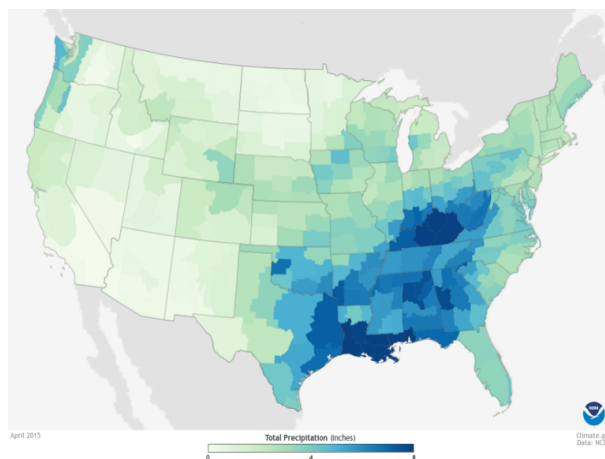
Tudományos publikációkban rendkívül fontos szerepe van a megfelelő színválasztásnak. Ehhez előre definiált színpaletták állnak rendelkezésre, a következőkben ezeket tekintem át. Általános alapelv a színválasztásnál, hogy a megjelenített adatokat egyértelművé tegye, ne torzítsa el, és az esetleges színtévesztő olvasók számára is problémamentes legyen a megjelenítés.

### 3.1. Szekvenciális adatok színezése

Amennyiben olyan adatot kívánunk megjeleníteni, amely értékei folytonosak és sorba rendezhetőek nagyság szerint, úgy érdemes az úgynevezett szekvenciális paletták közül válogatni. Ennek előnye, hogy a legtöbb ember a sötétebb színeket, például a feketét a "több", vagy "sűrűbb" adathoz társítja, míg a világosabb színek az ellenkező hatást érik el.



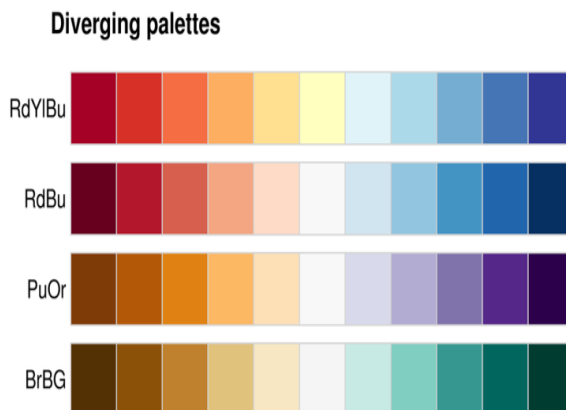
10. ábra. Szekvenciális paletták



11. ábra. Egy példa a szekvenciális adatok színezésére.

### 3.2. Divergens adatok színezése

Ha az adataink divergens struktúrájúak, tehát tartozik hozzá egy nullpont, amihez képest lehet az adatok értéke nagyobb, vagy kisebb, akkor célszerű a divergens színpalettákból válogatni. Ezek olyan paletták, amelyek két vége tendál a sötétebb színek felé, így mutatva, hogy azok a részei a "jobban" negatívak, vagy pozitívak.

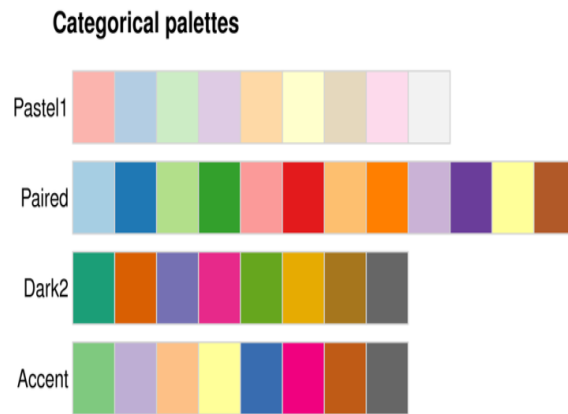


12. ábra. Divergens paletták

### 3.3. Kategorikus adatok színezése

Kategorikus adatok esetében, amelyek nem állíthatóak sorba, törekednünk kell arra, hogy egymástól jól megkülönböztethető színeket használjunk, illetve, hogy egy ábrázoláson csak megfelelő mennyiségű kategóriát szerepeltessünk, hogy ne legyenek összekeverhetőek az adatpontok.

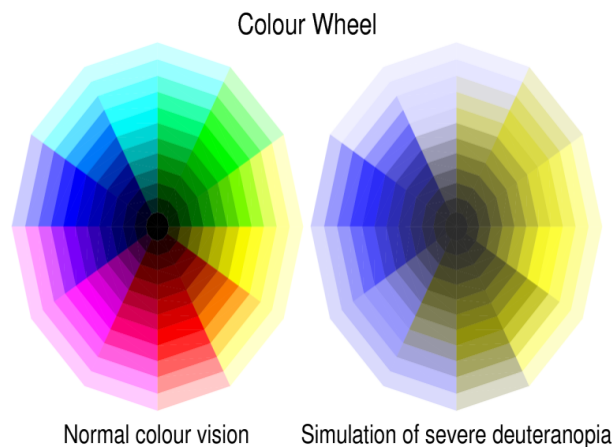




13. ábra. **Kategorikus paletták**

### 3.4. Egyéb szempontok

Fontos szempont, hogy az adatok megjelenítésénél gondoljunk színtévesztő olvasókra. Ehhez általános szabály, hogy pirosat és zöldet soha ne használjunk együtt, mert a színtévesztés leggyakoribb fajtája (8% férfiaknál, 0.5% nőknél) pont ezt a két színt teszi nehezen megkülönböztethetővé. Ezen szabály betartása mellett rendelkezésre állnak kifejezetten színtévesztőknek kifejlesztett színpaletták is, ezek használata is indokolt lehet.



14. ábra. **Színtévesztés szimulálása.**

	Original	Simulation				Hue	for Photoshop, Illustrator, Freehand, etc.		for Word, PowerPoint, Canvas, etc.	
		Protan	Deutan	Tritan			C,M,Y,K (%)	R,G,B (0-255)	R,G,B (%)	
1					Black	0°	(0,0,0,100)	(0,0,0)	(0,0,0)	
2					Orange	41°	(0,50,100,0)	(230,159,0)	(90,60,0)	
3					Sky Blue	202°	(80,0,0,0)	(86,180,233)	(35,70,90)	
4					bluish Green	164°	(97,0,75,0)	(0,158,115)	(0,60,50)	
5					Yellow	56°	(10,5,90,0)	(240,228,66)	(95,90,25)	
6					Blue	202°	(100,50,0,0)	(0,114,178)	(0,45,70)	
7					Vermillion	27°	(0,80,100,0)	(213,94,0)	(80,40,0)	
8					reddish Purple	326°	(10,70,0,0)	(204,121,167)	(80,60,70)	

15. ábra. Színtévesztő-barát színpaletták

## 4. Térbeli adatok megjelenítése, volometrikus és szintfelületes ábrázolás.

TODO

## 5. Háromdimenziós renderelési technikák, ray casting, ray tracing.

TODO

## 6. A sztereoszkópius megjelenítés technológiái

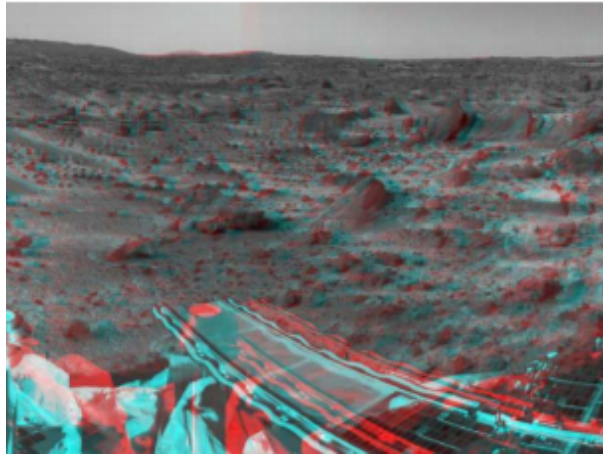
Sztereoszkópiának hívunk bármilyen olyan technológiát, amely lehetőséget ad egy 2D kép olyan módon való megjelenítésére, amely valódi 3D képnek tűnő hatást kelt. Fotók, vagy videók esetében ezt úgy érik el, hogy különböző képet mutatnak a különböző szemeknek.

A térlátás abból adódik, hogy az emberi szemek közelítőleg 5 cm-re helyezkednek el egymáshoz képest, így egy adott jelenséget a két szem kicsit más szögből látja. Lehetőség van ennek analógiájára olyan kamerát készíteni, amely két lencsével rendelkezik, reprodukálva az emberi látás sajátosságait. Alapvetően kétfajta technikát különböztetünk meg: a sztereoszkópikus és az autosztereoszkópikus megjelenítéseket. Az elsőnél a megfigyelőnek szüksége van valamilyen eszköz, például speciális szemüveg viselésére a 3D hatás megtapasztalására, míg a másodikonál nincs.

### 6.1. Anaglif megjelenítés

A sztereoszkópikus technológia felbontható több csoportra. Az első ilyen módszer az úgynevezett anaglif megjelenítés. Ennek során egy képet két különböző szögből vesznek fel, majd két különböző színnel látnak el, amelyet hasonló színezésű szemüveggel (egy szín az egyik, másik

szín a másik szemnek) vizsgálva előhozható a 3D megjelenítés. Általában a két használt szín a piros és a kék. Előnyei a technikának, hogy nagyon egyszerűen lehet előállítani a kompozit képet, és olcsó a hozzá tartozó szemüveg. Ezen kívül fizikailag is kinyomtatható a kép, a hatás úgy is megmarad. Hátrányai, hogy nehéz valódi színeket megjeleníteni rajta, a keletkező kép gyakran homályos.



16. ábra. Anaglif kép.

## 6.2. Polarizációs szemüvegek

Ez a technika hasonló az anaglif megjelenítéshez, de itt színkódolás helyett különböző polarizációs szűrőkön vezetik át a két szemhez tartozó képet. Előnye az előzőhöz képest a jobb színhűség, de hátránya, hogy a megjelenő kép homályos és fényszegény a szűrők fényelnyelése miatt.

## 6.3. Shutter szemüvegek

Ennek a technika során a speciális szemüveg periodikusan kitakarja egyszer az egyik, másszor a másik szem képét, így a kijelző videokártyájával megfelelően szinkronizálva, majd azon megfelelő időben megjelenítve a képeket elérhető a 3D hatás. Előnye, hogy sokkal meggyőzőbb képet ad az előző két technikával szemben, de hátránya, hogy a speciális szemüvegek gyakran nagyon drágák, és értelemszerűen a keletkező képet kinyomtatni se lehet.

## 6.4. Autosztereoszkópikus technikák

Az autosztereoszkópikus technikák is felbonthatóak több csoportra.

[https://www.cg.tuwien.ac.at/courses/Seminar/WS2006/stereo\\_techniques.pdf](https://www.cg.tuwien.ac.at/courses/Seminar/WS2006/stereo_techniques.pdf)

# 7. Számítógépes geometria

TODO