

Házi feladatok megoldása 3. Sűrűsödéspontok és sűrű régiók

Smahajcsik-Szabó Tamás, M9IJYM

1. Készíts sűrűségváltozókat a 8 Diener-item segítségével. Korreláltasd egymással őket és válassz ki közülük egy olyat, amelyik az összes többivel minimum 0,80-as szinten korrelál!

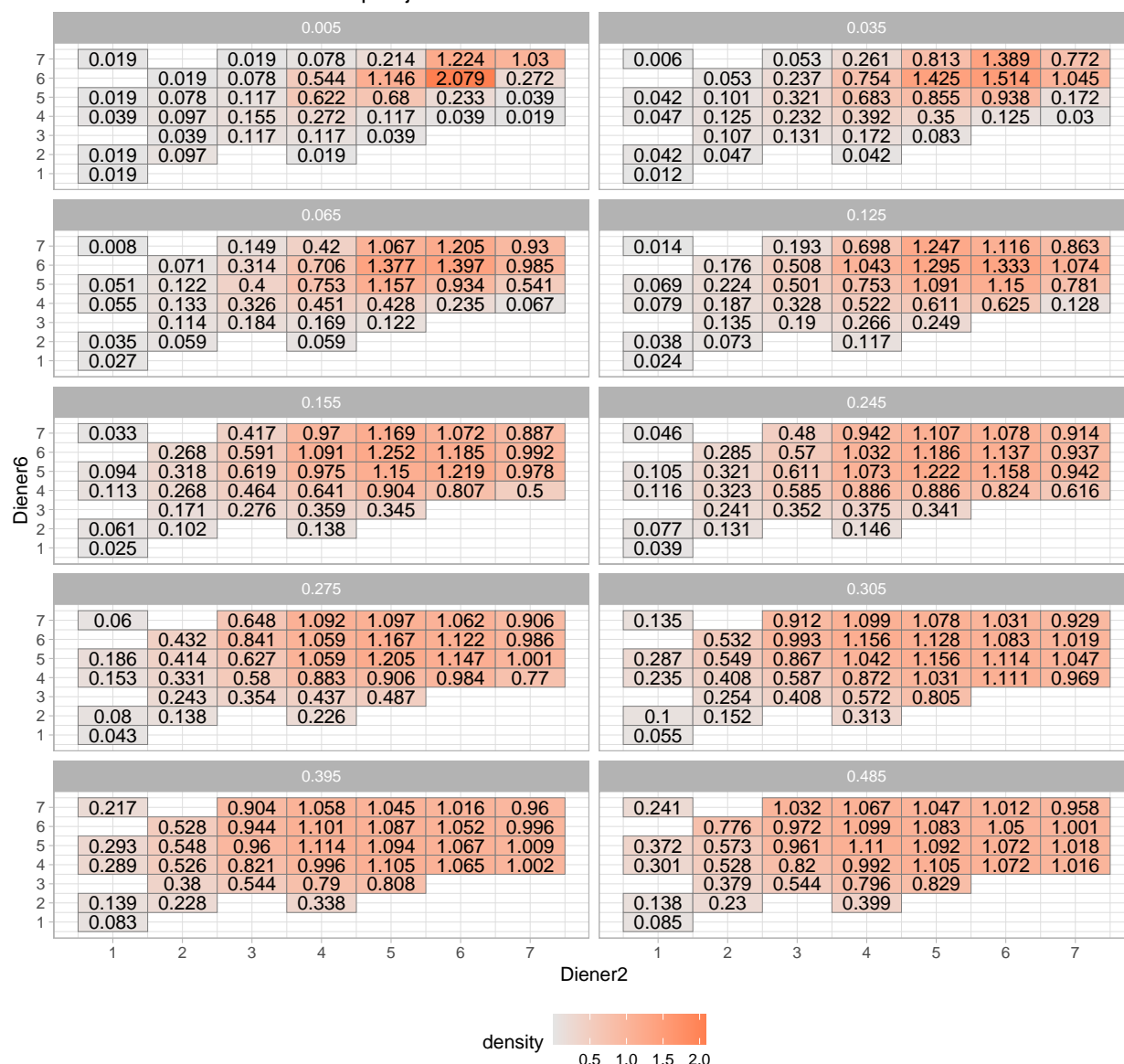
A Diener-tételekkel képzett density változók közül a Dense11 nevűt választottam, mely a leíró statisztikája (átlaga 1, szórása pedig 0.138) nyomán kevésbé szélsőséges eloszlású, mint a hasonlóan szoros, a többi változóval $r > 0.80$ korrelációt mutató Dense10 vagy a Dense9. Fontos, hogy az elemzéshez az *R* nyelvet használtam, így a változók neve eltér a ROPStattól. A másik eltérés, hogy az itt használt függvény opcionálisan egy előszűrést végez a density változókon: kiveszi azokat a sűrűsödésvaltozókat, melyek nem eredményeznek kellően nagy emelkedést a szomszédok átlagos számában.

2. A 2.1. alfejezetben leírt módszerrel keress sűrű régiókat a Diener2 és a Diener6 item kétdimenziós terében! Hogyan tudnád jellemezni a két legsűrűbb régiót?

A Diener2 és Diener6 itemek tekintetében az alábbi ábra tájékoztat a sűrű régiókról. Az ábrán különböző sugárértékek mentén került leképezésre a sűrűség.

Leginkább szembetűnő a Diener2 és Diener6 tételek 6-6 binjének magas sűrűsége a bal felső, legalacsonyabb sugár alapú feltérképezésen. Ezt úgy értelmezem, hogy a társas kapcsolatokat jónak, támogatónak megélők (6) jó eséllyel önmagukat és életüket is kellemesen élik meg, kielégítőnek minősítik. Viszont mint a 6-6 övezetet körülvevő, a nagyobb sugár mentén való leképezésnél szembetűnőbb, ám annal ugyan kevésbé sűrű, a többi övezeti metszésponthoz képest mégis nagyobb sűrűségű területből látszik (így a 6-7, 5-6 és 7-7 cellákból), a minta nagy része mindkét tétel tekintetében az 5-7 övezetekben tömörül.

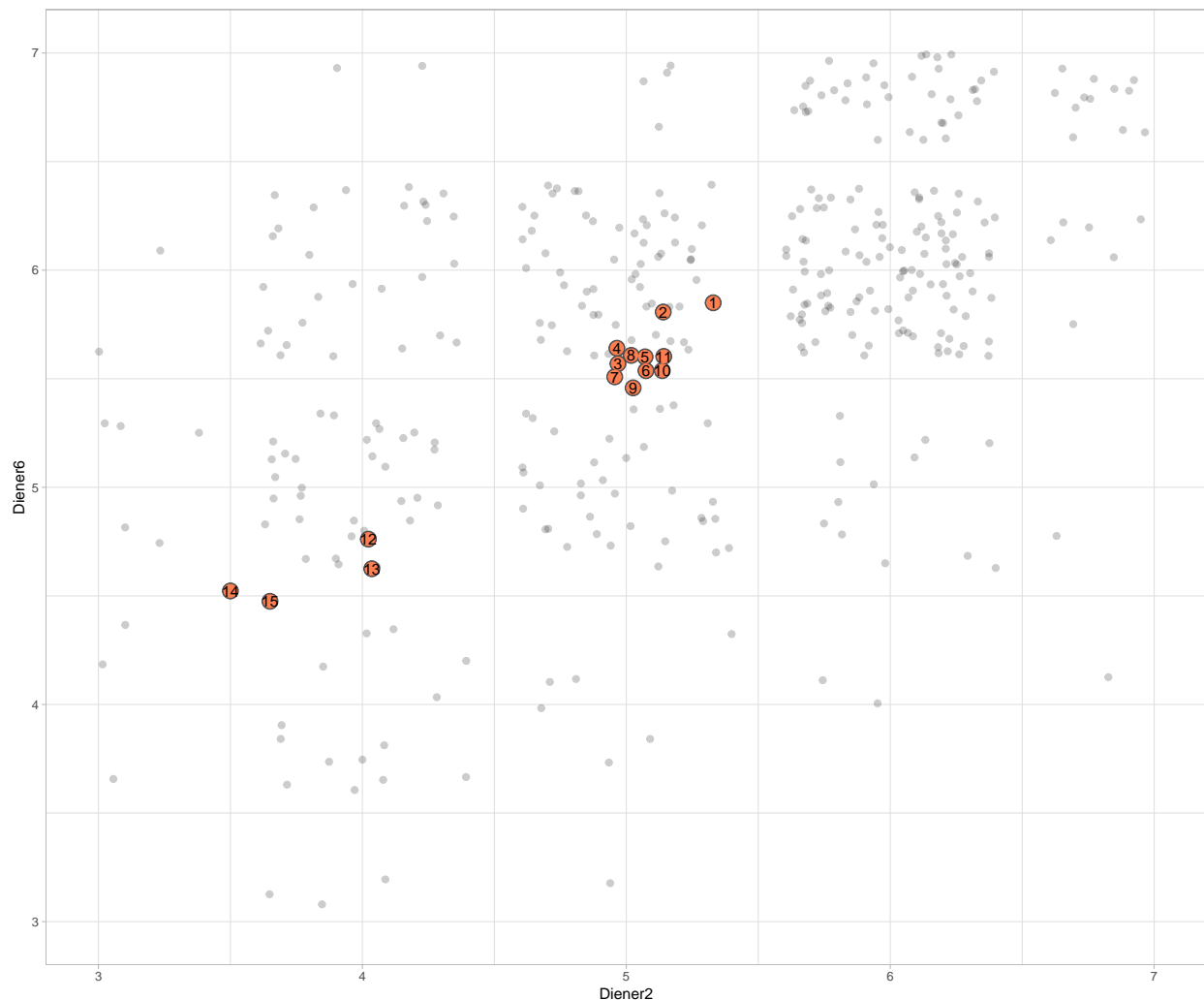
Diener2 es Diener6 s.r..södéspontjai különböz.. radius érték mentén



A számértékek átlag s.r..ség értékeket jelölnek cellánként

3. A 2.2.1. alpontban leírt módszerrel keress DP sűrűsödéspontokat a Diener2 és a Diener6 item kétdimenziós terében, 200 kezdeti sűrűsödéspontot beállítva és nem standardizálva a változókat! Hogyan tudnád jellemezni a kapott sűrűsödéspontokat? Van-e valamilyen közülük az előző feladat sűrű régióihoz?

2sűrűsödéspontok 00 kezdeti értékkel első körben 15 sűrűsödéspontot kerestem. Az alábbi ábra mutatja a sűrűsödéspontok elhelyezkedését az eredeti adateloszlásra exponálva. Fontos, hogy a tételek értékeinek diszkrét jellege miatt az *R ggplot2* csomagjának *geom_jitter()* függvényét használtam a pontdiagrammhoz, mely véletlenszerű értékekkel mesterseges szórást ad diszkrét változók ponteloszlásához a jobb láthatóság végett. Az eredményeket úgy tudom jellemezni, hogy a fentebb azonosított, 6-6 övezetmetszés mentén jelennek meg sűrűsödéspontok. Nagyobb célérték mellett az alsóbb területekben is találtam sűrűsödéspontokat.



Az alábbi táblázatban kiemeltem néhányat a dense point-okból:

Diener2	Diener6	id
5.329218	5.849794	1
4.964413	5.640569	4
3.500000	4.522727	14

A megadott Diener2 es Diener6 értékek összevont adatpontok átlagértékei.

4. Mentsd el a kapott DP-ket. A Dpcode változó oszlopában az üres cellákba írd 0-t (Szerkesztés/Keres, cserél/Aktuális oszlopban ...) segítségével, majd hasonlítsd össze a 0, 1, 2 kódú személyeket a MET skálái segítségével! Melyik MET-skálánál a legnagyobb az etanégyszet és mekkora ez az érték? Értelmezd szakmailag is a kapott eredményt. Önmagában a DPCode szerinti csoportosításnak a Diener2 es Diener6 tételek tekintetében nincs szignifikáns csoportképző hatása, azaz az elvártnak megfelelően, a sűrűsádpontok, noha mesterséges aggregátumok, e két változó alapján nem különböznek az eredeti mintától.

Az alábbi táblázat ezen ANOVA eredményét összegzi.

Df Sum Sq Mean Sq F value Pr(>F)

```
## Diener2      1  0.058 0.05839  2.063  0.152
## Diener6      1  0.013 0.01282  0.453  0.501
## Residuals   512 14.492 0.02830
```

A MET-tételek esetében a Jóllét és a Reziliencia tételek tekintetében $p < 0.01$ szintű eltérés mutatkozik az eredeti minta javára, mely az alkalmazott imputáció miatt alakulhatott így. A sűrűsödéspontok statikus, alacsony varianciájú MET pontokat kaptak kNN -alapú (iker) imputációval, értékük stabil és alacsonyabb a minta eredeti részénél tapasztalt értékeknél. Ezt összegzi az alábbi output:

```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Jóllét      1  0.211 0.21114   7.670 0.00582 **
## Savor       1  0.051 0.05098   1.852 0.17418
## AVhat       1  0.035 0.03518   1.278 0.25878
## Önreg       1  0.085 0.08458   3.073 0.08023 .
## Rezil       1  0.120 0.11960   4.345 0.03763 *
## M_Flow      1  0.056 0.05612   2.039 0.15396
## Diener2     1  0.047 0.04651   1.690 0.19424
## Diener6     1  0.030 0.03021   1.098 0.29531
## Residuals   506 13.929 0.02753
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```