UNIVERSIDADE FEDERAL DA PARAÍBA

CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA

DEPARTAMENTO DE ESTATÍSTICA

Análise Multivariada II, 2012.2, Prof. Marcelo Ferreira

Aluno: Carlos Alberto Alves de Meneses, Matrícula: 20180003202

Atividade: Avaliação 4

Análise de Agrupamentos

- **Introdução**

Segundo Bruce e Bruce (2019), os métodos de agrupamentos podem ser usados para identificar grupos de dados significativos. Por exemplo, usando os cliques da web e dados demográficos de usuários em um site, podemos ser capaz de agrupar diferentes tipos de usuários. O site poderia, então, ser personalizado para esses diferentes tipos.

Nesse trabalho iremos propor e escolher o(s) melhor(es) modelo(s) de agrupamentos para o conjunto de dados de Water Qualiity (Drinking water portability, water_potability {datasets}) disponível em: https://www.kaggle.com/datasets/adityakadiwal/water-potability

- **Objetivo**

Reduzir a dimensão dos dados para um conjunto mais gerenciável de variáveis;

Obter uma melhor percepção interna do conjunto de dados e de como as diferentes variáveis se relacionam umas com as outras;

Filtrar e analisar as variáveis e descobrir relacionamentos.

ANÁLISE EXPLORATÓRIA

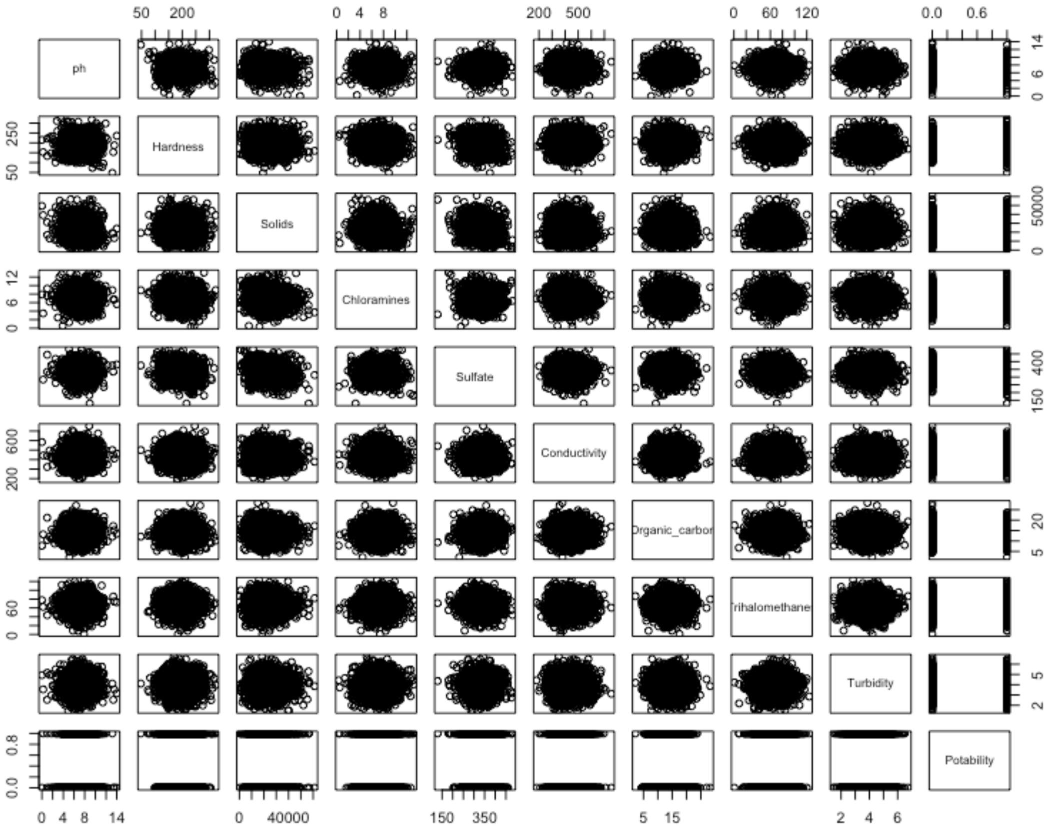
```
> head(df)
        ph Hardness   Solids Chloramines  Sulfate
1       NA 204.8905 20791.32    7.300212 368.5164
2 3.716080 129.4229 18630.06    6.635246       NA
3 8.099124 224.2363 19909.54    9.275884       NA
4 8.316766 214.3734 22018.42    8.059332 356.8861
5 9.092223 181.1015 17978.99    6.546600 310.1357
6 5.584087 188.3133 28748.69    7.544869 326.6784
  Conductivity Organic_carbon Trihalomethanes Turbidity
```

```
1     564.3087      10.379783        86.99097  2.963135
2     592.8854      15.180013        56.32908  4.500656
3     418.6062      16.868637        66.42009  3.055934
4     363.2665      18.436524       100.34167  4.628771
5     398.4108      11.558279        31.99799  4.075075
6     280.4679       8.399735       54.91786  2.559708
  Potability
1          0
2          0
3          0
4          0
5          0
6          0
> skim(df)
── Data Summary ────────────────────────
                           Values
Name                       df
Number of rows             3276
Number of columns          10
_______________________
Column type frequency:
  numeric                  10
________________________
Group variables            None

── Variable type: numeric ──────────────────────────────────────
   skim_variable  n_missing complete_rate      mean       sd
 1 ph                   491         0.850     7.08      1.59
 2 Hardness               0         1       196.       32.9
 3 Solids                 0         1     22014.     8769.
 4 Chloramines            0         1         7.12      1.58
 5 Sulfate              781         0.762   334.       41.4
```

6	Conductivity	0	1	426.	80.8
7	Organic_carbon	0	1	14.3	3.31
8	Trihalomethanes	162	0.951	66.4	16.2
9	Turbidity	0	1	3.97	0.780
10	Potability	0	1	0.390	0.488

	p0	p25	p50	p75	p100	hist
1	0	6.09	7.04	8.06	14	
2	47.4	177.	197.	217.	323.	
3	321.	15667.	20928.	27333.	61227.	
4	0.352	6.13	7.13	8.11	13.1	
5	129	308.	333.	360.	481.	
6	181.	366.	422.	482.	753.	
7	2.20	12.1	14.2	16.6	28.3	
8	0.738	55.8	66.6	77.3	124	
9	1.45	3.44	3.96	4.50	6.74	
10	0	0	0	1	1	

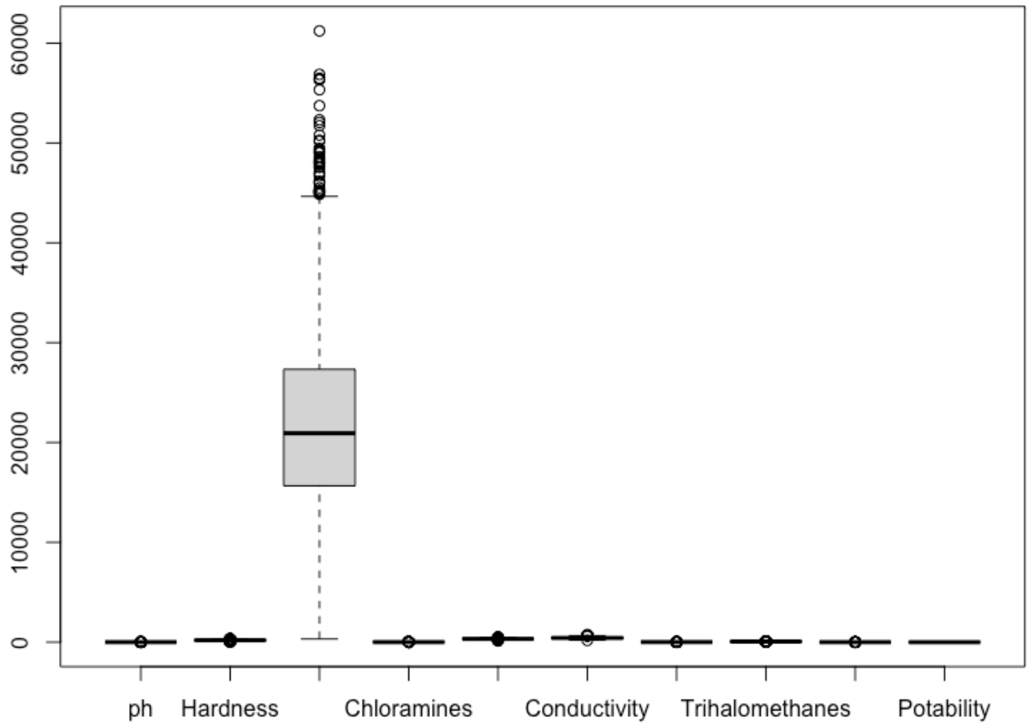
Gráfico 1



A análise inicial demonstra que o banco é composto por 3276 observações e 1º variáveis. Observamos também a presença de NaNs, que serão tratadas posteriormente.

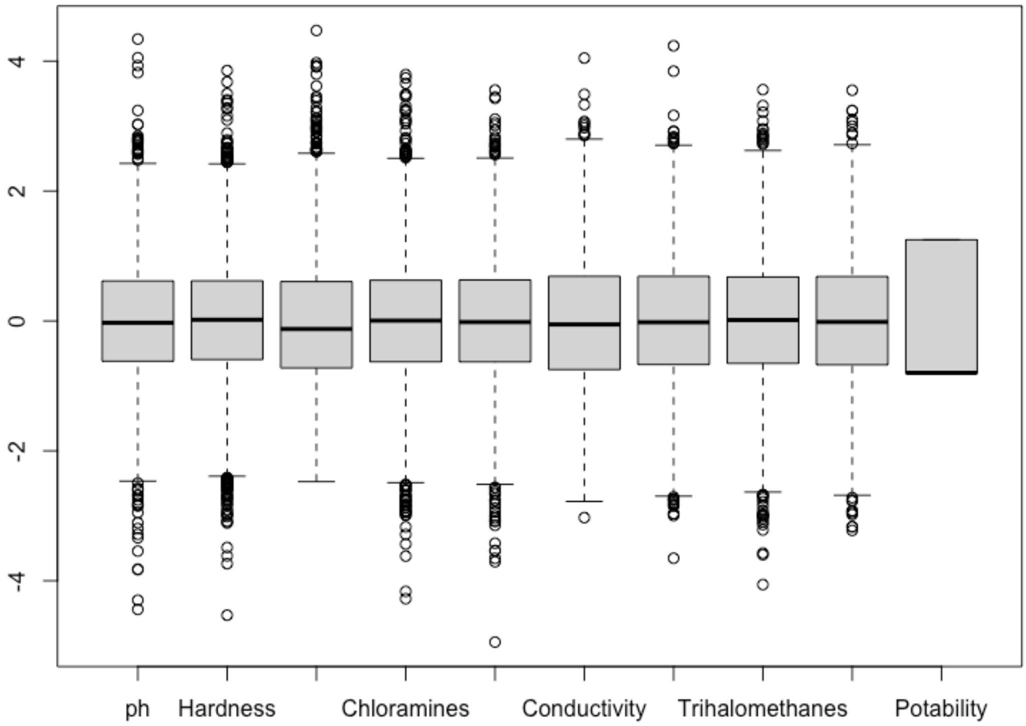
Verificando se os dados estão padronizados

Gráfico 2



O gráfico de boxplot (gráfico 2), mostra que os dados não estão padronizados, apresentando a variável Solids se destacando em relação as demais variáveis, portanto, iremos realizar a padronização do banco de dados utilizando a função scale().

Gráfico 3

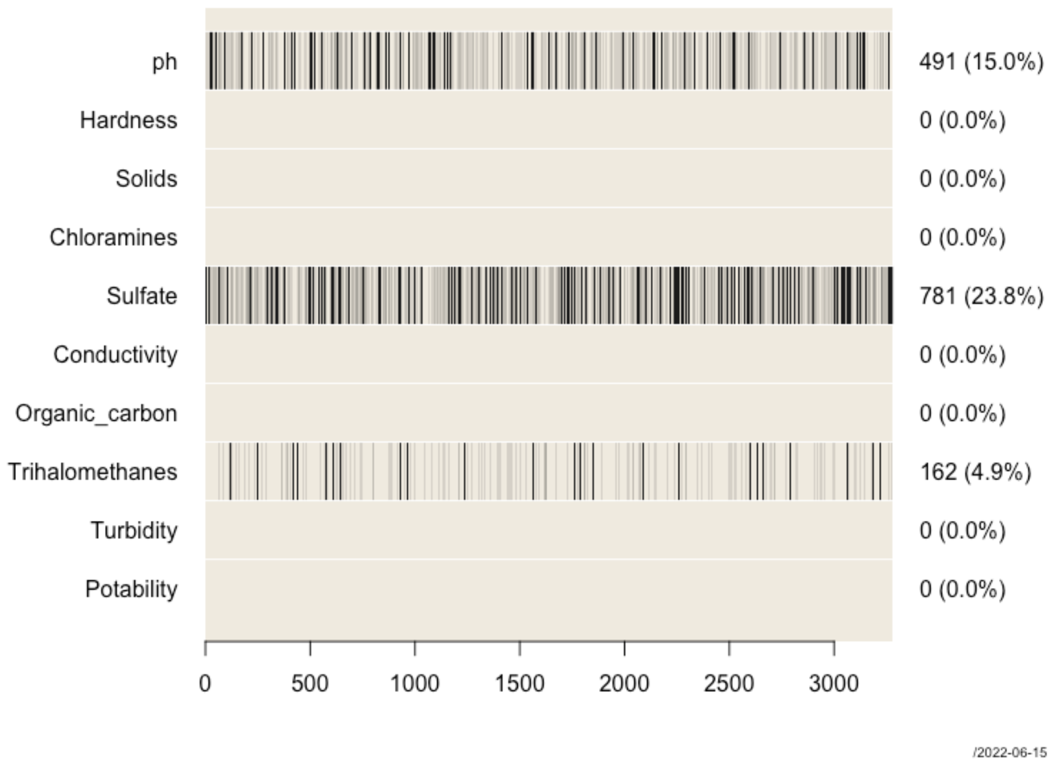


O gráfico de boxplot (gráfico 3) mostra que os dados foram padronizados após a aplicação da função scale**().**

Dados faltantes

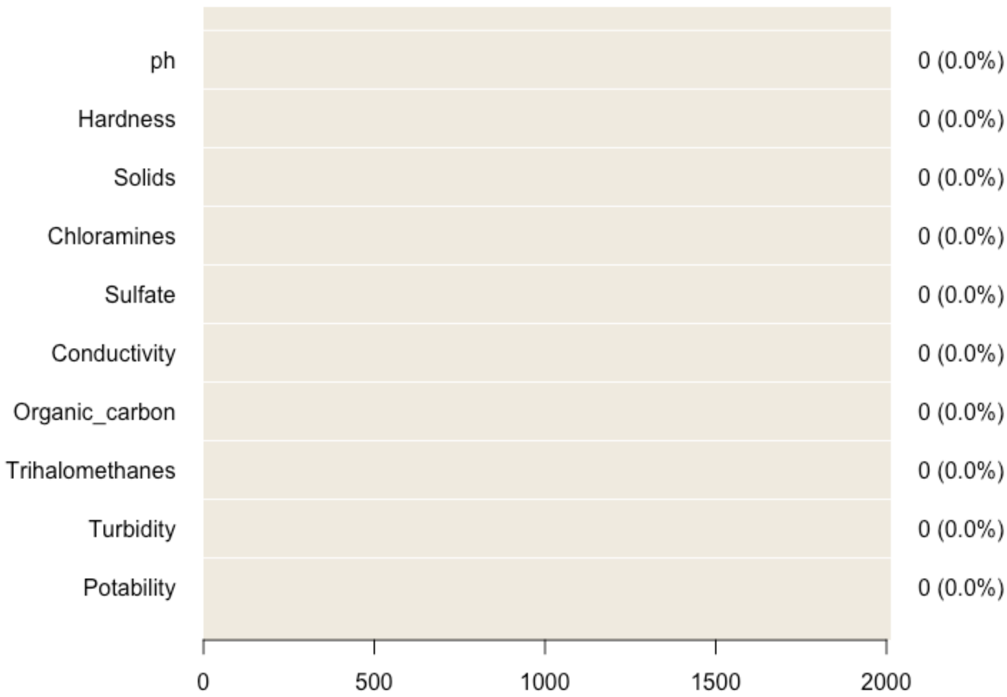
Vamos analisar nossa base de dados com observações faltantes. A primeira pergunta que devemos fazer é: como estão distribuídos os dados faltantes?

Gráfico 4



Observamos no (gráfico 4) a presença de NaNs e como eles estão distribuídos nas variáveis. Iremos proceder a limpeza dos dados faltantes utilizando a função omit.na().

Gráfico 5



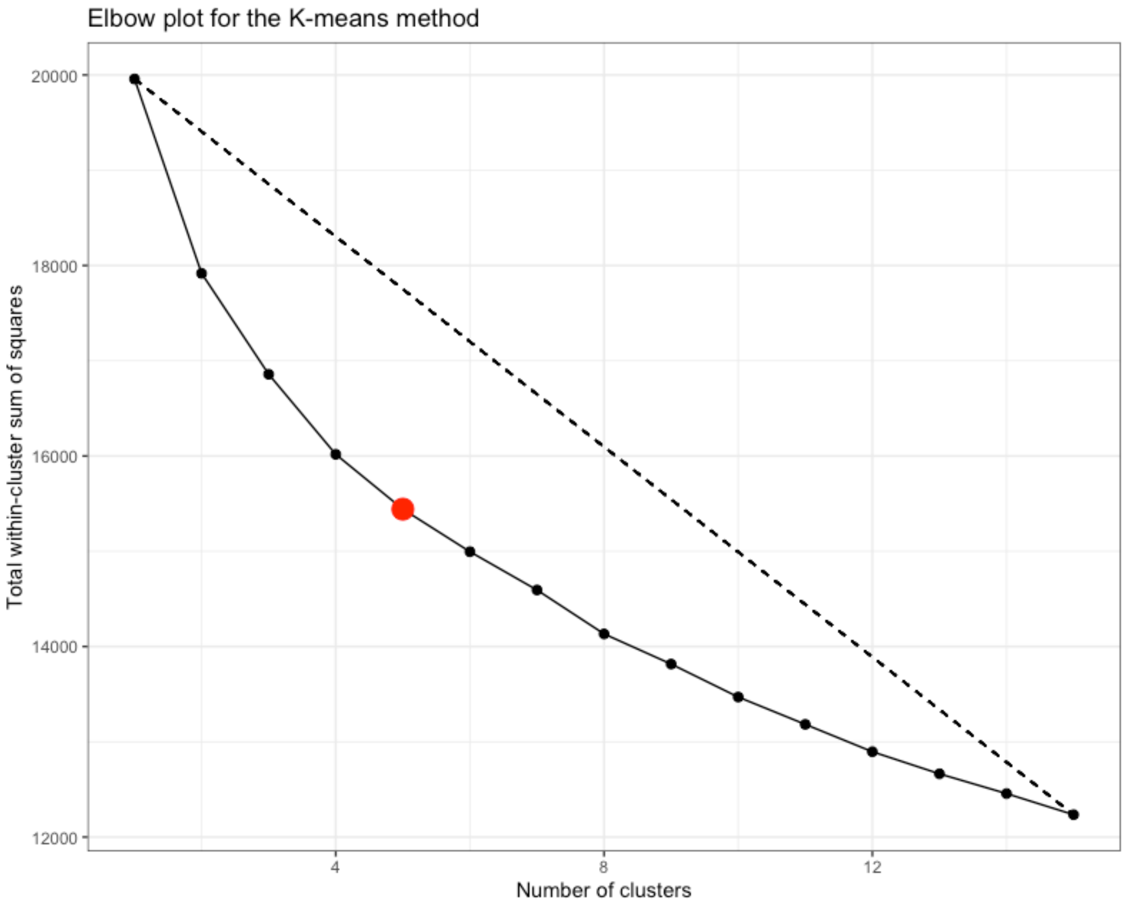
O (gráfico 5) mostra que os NaNs foram todos retirados do nosso banco de dados.

K-Means

O método de clusterização K-means classifica os objetos dentro de múltiplos grupos, de forma que a variação intra-cluster seja minimizada pela soma dos quadrados das distâncias Euclidianas entre os itens e seus centroides.

Encontrando o número ótimo de grupos para o K-means:

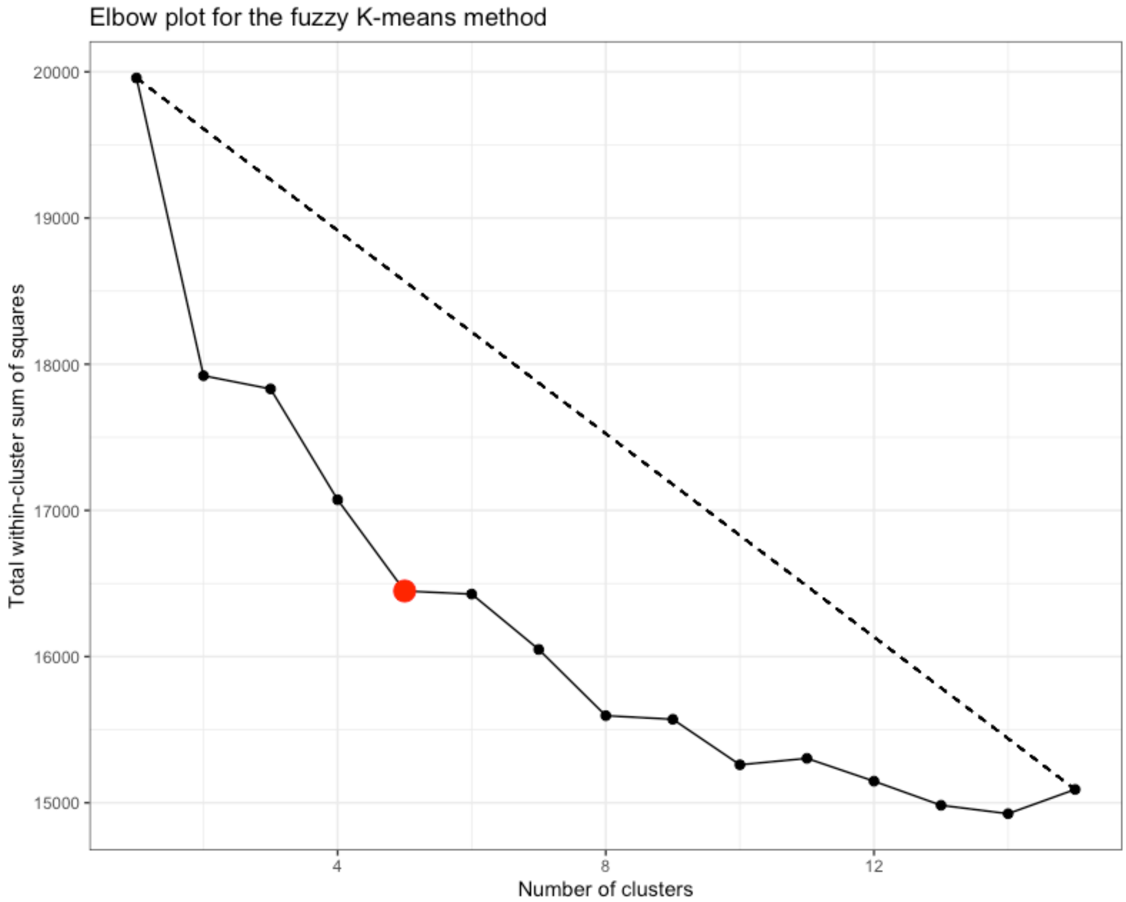
Gráfico 6



O (gráfico 6) mostra que o K "ótimo" é igual a 5 para o nosso banco de dados.

Encontrando agora o número ótimo de grupos para o Fuzzy K-means

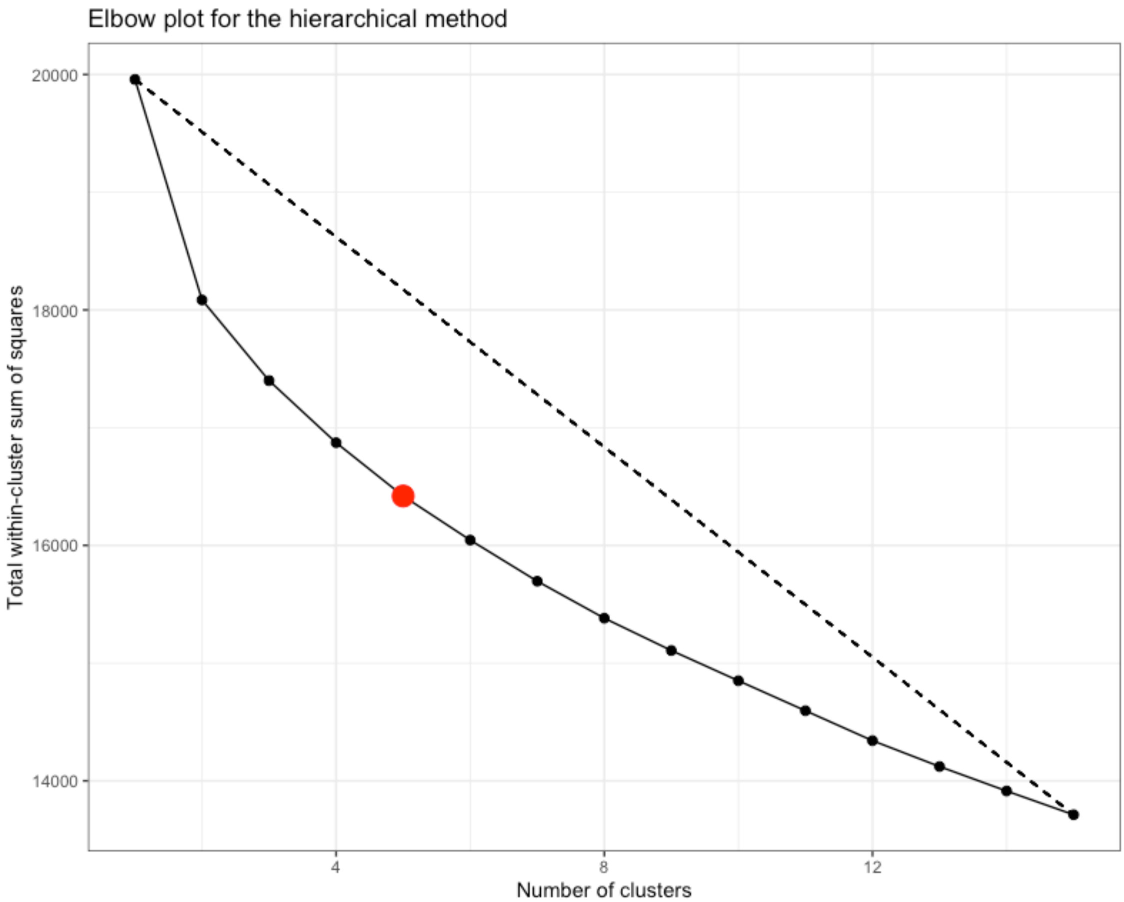
Gráfico 7



O (gráfico 7) nos mostra que o K "ótimo" também é igual a 5 através do método Fuzzy K-means.

Encontrando o número ótimo de grupos para o método hierárquico

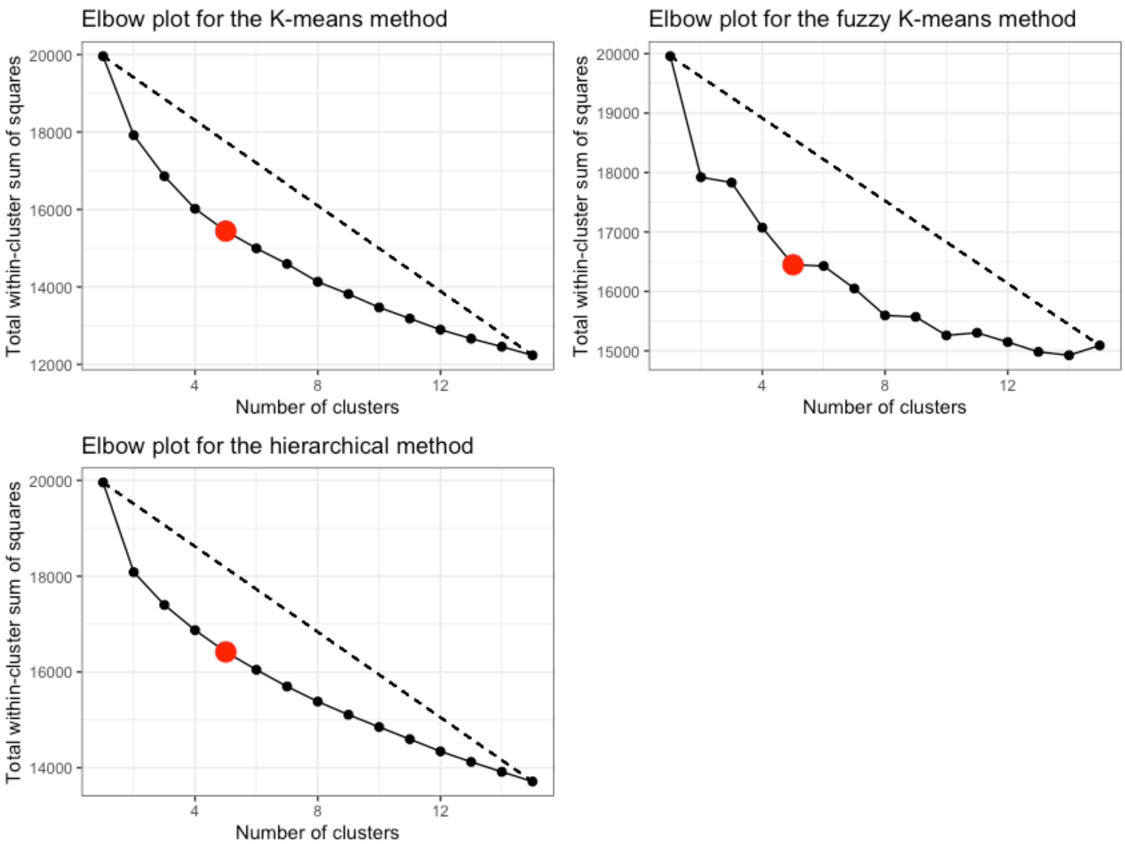
Gráfico 8



Assim como os métodos anteriores, o (gráfico 8) mostra que o K "ótimo" também é igual a 5 através do método hierárquico.

Portanto, nosso nclust = c(5,5,5) conforme é demonstrado pelo (gráfico 9)..

Gráfico 9



Agrupamento usando o K-means

K-means clustering with 5 clusters of sizes 342, 438, 449, 443, 339

Cluster means:

	ph	Hardness	Solids	Chloramines	Sulfate
1	0.4523146	-0.14368862	0.5857378	0.21014697	-0.95242877

```
2  0.6987339  0.18649100 -0.4859894 -0.19355669  0.20715838
3 -0.3055112  0.07540519 -0.3756107 -0.08447741  0.67564369
4 -0.8794434 -0.33853850  0.5279489  0.27616911 -0.16523622
5  0.2141124  0.17410636 -0.2208180 -0.16573173 -0.06468456
  Conductivity Organic_carbon Trihalomethanes   Turbidity
1  -0.11013143     -0.1672606      0.16456456  0.07016385
2   0.11405998     -0.1387859      0.55610610 -0.43436076
3   0.05630854      0.1190917     -0.07685825  0.03741296
4   0.02706308     -0.1798285      0.16445455  0.20518053
5  -0.12262689      0.5557541     -0.99596410  0.19511377
  Potability
1  1.2381740
2 -0.7622131
3  1.2455959
4 -0.7765171
5 -0.7391862

Clustering vector:
  [1] 2 5 4 2 2 2 4 2 4 2 2 4 4 4 4 2 2 5 4 4 5 4 4 5 2 2 5 2 4 5
 [31] 5 5 4 4 5 4 5 2 4 4 4 4 5 4 2 2 4 4 2 2 5 2 2 5 2 4 4 2 4 5
 [61] 5 2 4 5 2 5 4 2 2 4 2 2 4 5 5 4 5 5 5 4 2 2 5 4 2 5 2 5 2
 [91] 2 4 4 4 2 5 2 4 5 2 5 2 2 2 4 2 4 4 5 4 4 5 4 4 4 4 2 5 2 5
[121] 2 4 5 2 5 5 2 4 4 4 2 2 5 4 2 2 5 4 2 2 5 4 5 2 4 4 5 5 2 5
[151] 5 2 4 4 1 1 1 1 3 1 1 3 3 3 3 3 3 3 1 3 1 1 1 1 1 5 1 3 1 1
[181] 1 1 1 1 3 3 1 3 1 1 1 3 3 3 3 3 1 1 1 1 1 3 1 1 1 3 3 1 1 3
[211] 1 3 3 3 3 1 1 1 1 3 3 3 3 3 3 3 3 3 1 1 1 3 1 3 1 1 1 1 3 1
[241] 3 1 3 1 3 1 3 3 1 3 3 1 3 3 3 3 5 4 4 4 4 2 4 5 2 2 4 4 4 2
[271] 2 2 2 5 5 2 4 4 2 4 2 5 2 5 5 2 2 4 2 2 2 2 5 2 4 2 5 4 5 4
[301] 2 5 2 5 4 4 2 2 4 4 2 4 5 2 4 5 2 4 4 4 4 4 5 2 4 4 4 4 4 5
[331] 5 2 4 2 5 4 2 4 4 4 4 4 5 2 2 4 2 4 5 5 2 2 2 4 5 5 4 5 4 2
[361] 2 2 2 4 2 4 2 5 4 5 5 4 5 2 4 2 5 2 2 2 4 2 2 5 4 4 4 4 4 2
[391] 5 4 2 4 5 2 4 4 2 4 5 5 2 2 2 4 3 1 2 3 1 3 3 3 1 1 1 3 3 1
```

```
 [421] 1 1 1 1 1 4 3 1 3 3 1 1 1 3 1 1 3 1 3 3 1 4 3 1 4 3 1 1 1 3
 [451] 1 3 3 3 3 3 3 3 1 3 3 1 1 3 1 1 1 1 3 3 1 3 3 1 1 5 3 1 2 3
 [481] 1 3 3 3 3 1 2 1 1 1 1 3 1 1 1 2 3 1 3 3 1 1 3 3 3 4 4 5 2 5
 [511] 4 4 4 2 2 4 2 4 2 4 2 2 2 4 4 2 5 5 5 2 2 2 5 5 4 2 2 4 2 4
 [541] 4 4 2 4 2 2 4 2 5 4 2 4 4 4 5 2 5 5 4 5 5 5 2 4 5 2 5 4 5 2
 [571] 5 4 2 5 4 5 4 2 2 5 2 4 4 4 2 4 2 4 2 4 2 4 4 2 5 2 5 5 4 4
 [601] 2 4 4 2 2 5 5 5 5 5 4 4 4 5 4 2 5 2 4 5 5 5 2 4 2 5 2 4 2 2
 [631] 4 4 2 5 2 5 2 2 2 2 4 4 2 4 2 4 5 4 4 5 2 2 5 2 2 5 2 4 5 2
 [661] 2 2 5 2 5 2 5 4 3 1 3 1 1 3 1 3 1 3 3 3 3 3 3 1 1 3 3 1 1 3
 [691] 3 3 1 1 3 1 3 1 3 5 1 3 3 3 1 3 1 3 1 3 1 3 1 1 3 3 3 1 1 1
 [721] 1 1 3 3 3 1 1 3 1 3 3 1 1 3 3 3 3 1 3 1 1 1 3 3 3 3 1 1 3 1
 [751] 1 3 3 3 3 3 1 3 3 3 1 4 4 2 4 4 4 5 4 5 4 2 4 4 2 5 2 4 2 4
 [781] 2 4 5 2 5 4 4 4 2 2 2 5 2 2 4 4 2 2 4 4 4 4 2 2 5 2 4 2 2 5
 [811] 4 4 5 2 2 5 2 2 5 4 4 2 5 5 4 5 2 4 4 5 5 2 2 4 2 5 2 5 5 4
 [841] 4 2 2 4 5 5 2 5 5 5 2 4 5 4 2 2 4 5 4 5 2 2 4 2 4 2 5 4 4 2
 [871] 4 2 5 5 2 2 2 2 5 2 2 5 2 4 2 2 5 2 4 2 4 4 2 2 2 5 2 2 4 4
 [901] 4 4 4 4 4 5 4 4 4 3 3 3 1 1 3 2 3 3 3 3 3 1 1 3 3 1 1 3 1 3
 [931] 3 3 1 1 3 1 3 1 3 3 1 3 3 1 3 1 1 1 3 1 3 1 1 1 3 1 3 3 1 1
 [961] 3 1 3 3 3 3 1 3 3 3 3 1 3 1 1 1 3 3 3 1 1 1 3 1 3 1 3 3 3 3
 [991] 1 3 1 3 1 3 3 3 3 1
 [ reached getOption("max.print") -- omitted 1011 entries ]

Within cluster sum of squares by cluster:
[1] 3005.709 3228.914 3575.035 3289.782 2342.790
 (between_SS / total_SS =  22.6 %)

Available components:

[1] "cluster"      "centers"      "totss"        "withinss"
[5] "tot.withinss" "betweenss"    "size"         "iter"
[9] "ifault"
```

Agrupamento usando o Fuzzy K-means

Fuzzy c-means clustering with 5 clusters

Cluster centers:

	ph	Hardness	Solids	Chloramines	Sulfate
1	0.003309876	-0.01217055	-0.01101303	0.007647791	-0.01331860
2	0.003219246	-0.01223407	-0.01103082	0.007620648	-0.01330774
3	0.003297004	-0.01218615	-0.01102080	0.007621521	-0.01330940
4	0.003179042	-0.01228230	-0.01102310	0.007646414	-0.01327141
5	0.003288098	-0.01217124	-0.01102451	0.007558935	-0.01332433

	Conductivity	Organic_carbon	Trihalomethanes	Turbidity
1	0.003902742	0.02196240	0.0003139698	0.003792758
2	0.004011753	0.02196099	0.0002991695	0.003802998
3	0.003967437	0.02203686	0.0003188474	0.003757085
4	0.003979881	0.02196812	0.0002652721	0.003760499
5	0.004014568	0.02201057	0.0002143065	0.003740428

	Potability
1	0.02742239
2	0.02673830
3	0.02723825
4	0.02677277
5	0.02682970

Memberships:

	1	2	3	4	5
[1,]	0.1999929	0.2000073	0.1999990	0.2000021	0.1999987
[2,]	0.1999846	0.2000076	0.1999870	0.2000067	0.2000142
[3,]	0.1999913	0.2000053	0.1999901	0.2000108	0.2000026
[4,]	0.1999986	0.1999999	0.1999996	0.1999992	0.2000027
[5,]	0.1999788	0.2000136	0.1999878	0.2000022	0.2000176
[6,]	0.1999932	0.2000033	0.1999992	0.1999996	0.2000046

[7,] 0.1999651 0.2000214 0.1999843 0.2000228 0.2000064

[8,] 0.1999863 0.2000069 0.1999943 0.2000054 0.2000070

[9,] 0.1999874 0.2000111 0.1999910 0.2000113 0.1999992

[10,] 0.1999922 0.2000055 0.1999961 0.1999985 0.2000077

[11,] 0.1999783 0.2000080 0.1999903 0.2000061 0.2000172

[12,] 0.1999694 0.2000156 0.1999906 0.2000126 0.2000118

[13,] 0.1999845 0.2000137 0.1999871 0.2000167 0.1999981

[14,] 0.1999790 0.2000226 0.1999852 0.2000154 0.1999978

[15,] 0.1999871 0.2000070 0.1999904 0.2000117 0.2000038

[16,] 0.1999914 0.2000122 0.1999874 0.2000067 0.2000022

[17,] 0.1999926 0.2000072 0.1999939 0.2000054 0.2000009

[18,] 0.1999868 0.2000056 0.1999894 0.2000074 0.2000108

[19,] 0.1999634 0.2000226 0.1999731 0.2000246 0.2000163

[20,] 0.1999832 0.2000125 0.1999878 0.2000125 0.2000040

[21,] 0.1999845 0.2000047 0.1999911 0.2000032 0.2000165

[22,] 0.1999818 0.2000178 0.1999794 0.2000196 0.2000013

[23,] 0.1999929 0.2000088 0.1999902 0.2000054 0.2000027

[24,] 0.1999945 0.2000017 0.2000006 0.1999989 0.2000044

[25,] 0.1999844 0.2000094 0.1999931 0.2000045 0.2000086

[26,] 0.1999954 0.2000051 0.1999948 0.2000006 0.2000041

[27,] 0.1999747 0.2000096 0.1999856 0.2000051 0.2000250

[28,] 0.1999908 0.2000050 0.1999955 0.2000047 0.2000040

[29,] 0.1999783 0.2000142 0.1999881 0.2000112 0.2000082

[30,] 0.1999788 0.2000073 0.1999937 0.2000030 0.2000171

[31,] 0.1999856 0.2000041 0.1999934 0.2000029 0.2000140

[32,] 0.1999927 0.2000053 0.1999916 0.2000043 0.2000061

[33,] 0.1999898 0.2000138 0.1999897 0.2000111 0.1999957

[34,] 0.1999930 0.2000066 0.1999908 0.2000089 0.2000008

[35,] 0.1999858 0.2000049 0.1999894 0.2000068 0.2000131

[36,] 0.1999903 0.2000079 0.1999951 0.2000050 0.2000017

[37,] 0.1999552 0.2000174 0.1999827 0.2000185 0.2000262

[38,] 0.1999902 0.2000084 0.1999937 0.2000032 0.2000044

[39,]	0.1999857	0.2000101	0.1999887	0.2000125	0.2000030
[40,]	0.1999812	0.2000139	0.1999889	0.2000121	0.2000039
[41,]	0.1999899	0.2000098	0.1999901	0.2000125	0.1999978
[42,]	0.1999794	0.2000109	0.1999880	0.2000159	0.2000058
[43,]	0.1999728	0.2000088	0.1999902	0.2000125	0.2000156
[44,]	0.1999902	0.2000090	0.1999900	0.2000103	0.2000005
[45,]	0.1999742	0.2000194	0.1999863	0.2000092	0.2000109
[46,]	0.1999865	0.2000081	0.1999938	0.2000061	0.2000054
[47,]	0.1999884	0.2000099	0.1999936	0.2000064	0.2000017
[48,]	0.1999719	0.2000227	0.1999810	0.2000131	0.2000113
[49,]	0.1999750	0.2000083	0.1999911	0.2000029	0.2000226
[50,]	0.1999973	0.2000031	0.1999973	0.1999977	0.2000045
[51,]	0.1999823	0.2000046	0.1999922	0.2000061	0.2000148
[52,]	0.1999894	0.2000020	0.1999981	0.2000004	0.2000101
[53,]	0.1999784	0.2000133	0.1999865	0.2000080	0.2000138
[54,]	0.1999792	0.2000130	0.1999879	0.2000102	0.2000097
[55,]	0.1999916	0.2000128	0.1999900	0.2000056	0.1999999
[56,]	0.1999861	0.2000137	0.1999914	0.2000123	0.1999964
[57,]	0.1999775	0.2000134	0.1999868	0.2000160	0.2000064
[58,]	0.1999811	0.2000115	0.1999863	0.2000121	0.2000091
[59,]	0.1999768	0.2000147	0.1999824	0.2000218	0.2000043
[60,]	0.1999859	0.2000067	0.1999959	0.2000036	0.2000079
[61,]	0.1999601	0.2000186	0.1999844	0.2000163	0.2000206
[62,]	0.1999991	0.2000055	0.1999951	0.1999999	0.2000003
[63,]	0.1999873	0.2000092	0.1999893	0.2000131	0.2000011
[64,]	0.1999790	0.2000078	0.1999920	0.2000033	0.2000179
[65,]	0.1999894	0.2000112	0.1999875	0.2000061	0.2000057
[66,]	0.1999940	0.2000017	0.1999973	0.2000032	0.2000038
[67,]	0.1999875	0.2000064	0.1999952	0.2000109	0.2000000
[68,]	0.1999967	0.2000042	0.1999977	0.2000013	0.2000000
[69,]	0.1999886	0.2000025	0.1999969	0.1999971	0.2000149
[70,]	0.1999832	0.2000099	0.1999934	0.2000130	0.2000006

[71,]	0.1999782	0.2000109	0.1999925	0.2000052	0.2000132
[72,]	0.1999965	0.2000034	0.1999939	0.2000028	0.2000034
[73,]	0.1999862	0.2000153	0.1999842	0.2000158	0.1999985
[74,]	0.1999534	0.2000186	0.1999795	0.2000186	0.2000299
[75,]	0.1999859	0.2000075	0.1999934	0.2000050	0.2000081
[76,]	0.1999757	0.2000109	0.1999909	0.2000155	0.2000070
[77,]	0.1999811	0.2000057	0.1999939	0.2000073	0.2000119
[78,]	0.1999924	0.2000028	0.1999939	0.2000044	0.2000064
[79,]	0.1999864	0.2000047	0.1999934	0.2000097	0.2000058
[80,]	0.1999628	0.2000194	0.1999853	0.2000173	0.2000152
[81,]	0.1999776	0.2000161	0.1999825	0.2000218	0.2000020
[82,]	0.1999788	0.2000130	0.1999914	0.2000150	0.2000019
[83,]	0.1999751	0.2000122	0.1999885	0.2000025	0.2000217
[84,]	0.1999553	0.2000252	0.1999825	0.2000141	0.2000230
[85,]	0.1999879	0.2000052	0.1999920	0.2000086	0.2000063
[86,]	0.1999647	0.2000258	0.1999874	0.2000136	0.2000085
[87,]	0.1999722	0.2000108	0.1999877	0.2000130	0.2000164
[88,]	0.1999560	0.2000223	0.1999806	0.2000187	0.2000225
[89,]	0.1999961	0.2000025	0.1999966	0.1999983	0.2000064
[90,]	0.1999888	0.2000099	0.1999959	0.2000045	0.2000009
[91,]	0.1999928	0.2000075	0.1999954	0.2000073	0.1999970
[92,]	0.1999884	0.2000136	0.1999886	0.2000129	0.1999966
[93,]	0.1999939	0.2000074	0.1999916	0.2000043	0.2000028
[94,]	0.1999785	0.2000159	0.1999871	0.2000151	0.2000034
[95,]	0.1999878	0.2000054	0.1999905	0.2000088	0.2000074
[96,]	0.1999741	0.2000095	0.1999858	0.2000064	0.2000243
[97,]	0.1999942	0.2000090	0.1999920	0.2000024	0.2000024
[98,]	0.1999561	0.2000301	0.1999743	0.2000332	0.2000063
[99,]	0.1999815	0.2000117	0.1999889	0.2000130	0.2000049
[100,]	0.1999890	0.2000051	0.1999962	0.2000020	0.2000078
[101,]	0.1999912	0.2000073	0.1999943	0.2000033	0.2000040
[102,]	0.1999890	0.2000072	0.1999952	0.2000056	0.2000030

[103,] 0.1999791 0.2000121 0.1999920 0.2000062 0.2000107

[104,] 0.1999834 0.2000118 0.1999940 0.2000025 0.2000083

[105,] 0.1999814 0.2000124 0.1999896 0.2000169 0.1999998

[106,] 0.1999613 0.2000213 0.1999823 0.2000170 0.2000181

[107,] 0.1999594 0.2000261 0.1999752 0.2000200 0.2000193

[108,] 0.1999913 0.2000133 0.1999872 0.2000132 0.1999950

[109,] 0.1999953 0.2000018 0.1999968 0.2000021 0.2000040

[110,] 0.1999656 0.2000254 0.1999741 0.2000246 0.2000104

[111,] 0.1999868 0.2000065 0.1999931 0.2000081 0.2000055

[112,] 0.1999894 0.2000079 0.1999932 0.2000075 0.2000020

[113,] 0.1999776 0.2000151 0.1999875 0.2000130 0.2000068

[114,] 0.1999893 0.2000075 0.1999931 0.2000057 0.2000044

[115,] 0.1999925 0.2000095 0.1999928 0.2000089 0.1999963

[116,] 0.1999895 0.2000071 0.1999958 0.2000100 0.1999976

[117,] 0.1999814 0.2000205 0.1999800 0.2000179 0.2000001

[118,] 0.1999534 0.2000269 0.1999774 0.2000168 0.2000255

[119,] 0.1999887 0.2000080 0.1999949 0.2000032 0.2000052

[120,] 0.1999876 0.2000030 0.1999954 0.2000064 0.2000075

[121,] 0.1999900 0.2000022 0.1999991 0.1999968 0.2000120

[122,] 0.1999879 0.2000082 0.1999952 0.2000078 0.2000009

[123,] 0.1999903 0.2000072 0.1999902 0.2000043 0.2000080

[124,] 0.1999741 0.2000184 0.1999890 0.2000090 0.2000095

[125,] 0.1999840 0.2000101 0.1999919 0.2000074 0.2000067

[126,] 0.1999900 0.2000041 0.1999941 0.2000016 0.2000102

[127,] 0.1999845 0.2000065 0.1999974 0.1999983 0.2000132

[128,] 0.1999962 0.2000069 0.1999955 0.2000059 0.1999955

[129,] 0.1999537 0.2000350 0.1999756 0.2000249 0.2000108

[130,] 0.1999591 0.2000246 0.1999785 0.2000247 0.2000131

[131,] 0.1999882 0.2000062 0.1999936 0.2000029 0.2000092

[132,] 0.1999882 0.2000050 0.1999954 0.2000016 0.2000097

[133,] 0.1999767 0.2000094 0.1999892 0.2000016 0.2000231

[134,] 0.1999741 0.2000172 0.1999844 0.2000216 0.2000027

[135,] 0.1999859 0.2000049 0.1999967 0.2000060 0.2000065
[136,] 0.1999906 0.2000064 0.1999937 0.1999978 0.2000116
[137,] 0.1999810 0.2000129 0.1999867 0.2000053 0.2000141
[138,] 0.1999904 0.2000055 0.1999961 0.2000047 0.2000032
[139,] 0.1999908 0.2000072 0.1999968 0.1999993 0.2000060
[140,] 0.1999930 0.2000035 0.1999992 0.2000022 0.2000022
[141,] 0.1999640 0.2000177 0.1999850 0.2000108 0.2000225
[142,] 0.1999747 0.2000104 0.1999892 0.2000134 0.2000122
[143,] 0.1999396 0.2000396 0.1999699 0.2000275 0.2000233
[144,] 0.1999590 0.2000164 0.1999834 0.2000166 0.2000245
[145,] 0.1999868 0.2000066 0.1999934 0.2000103 0.2000029
[146,] 0.1999826 0.2000152 0.1999831 0.2000159 0.2000033
[147,] 0.1999534 0.2000210 0.1999847 0.2000136 0.2000273
[148,] 0.1999803 0.2000099 0.1999909 0.2000091 0.2000097
[149,] 0.1999945 0.2000049 0.1999981 0.1999966 0.2000059
[150,] 0.1999548 0.2000257 0.1999777 0.2000234 0.2000185
[151,] 0.1999893 0.2000004 0.1999946 0.2000021 0.2000136
[152,] 0.1999806 0.2000169 0.1999867 0.2000078 0.2000080
[153,] 0.1999955 0.2000036 0.1999970 0.2000026 0.2000013
[154,] 0.1999877 0.2000157 0.1999891 0.2000083 0.1999991
[155,] 0.2000113 0.1999954 0.2000056 0.1999947 0.1999930
[156,] 0.2000240 0.1999862 0.2000147 0.1999885 0.1999866
[157,] 0.2000127 0.1999940 0.2000054 0.1999925 0.1999955
[158,] 0.2000522 0.1999709 0.2000263 0.1999729 0.1999776
[159,] 0.2000196 0.1999873 0.2000115 0.1999880 0.1999936
[160,] 0.2000356 0.1999759 0.2000220 0.1999781 0.1999884
[161,] 0.2000087 0.1999930 0.2000061 0.1999970 0.1999952
[162,] 0.2000453 0.1999731 0.2000225 0.1999792 0.1999800
[163,] 0.2000169 0.1999884 0.2000094 0.1999903 0.1999951
[164,] 0.2000246 0.1999877 0.2000153 0.1999853 0.1999871
[165,] 0.2000190 0.1999867 0.2000145 0.1999841 0.1999957
[166,] 0.2000175 0.1999901 0.2000086 0.1999904 0.1999935

[167,] 0.2000443 0.1999719 0.2000269 0.1999755 0.1999814
[168,] 0.2000325 0.1999802 0.2000224 0.1999832 0.1999817
[169,] 0.2000201 0.1999897 0.2000100 0.1999852 0.1999950
[170,] 0.2000255 0.1999823 0.2000209 0.1999840 0.1999874
[171,] 0.2000118 0.1999944 0.2000061 0.1999956 0.1999921
[172,] 0.2000336 0.1999807 0.2000177 0.1999828 0.1999852
[173,] 0.2000094 0.1999967 0.2000033 0.1999955 0.1999951
[174,] 0.2000349 0.1999770 0.2000232 0.1999823 0.1999826
[175,] 0.2000075 0.1999969 0.2000023 0.1999964 0.1999969
[176,] 0.2000131 0.1999859 0.2000088 0.1999928 0.1999994
[177,] 0.2000104 0.1999943 0.2000044 0.1999967 0.1999943
[178,] 0.2000181 0.1999884 0.2000142 0.1999841 0.1999953
[179,] 0.2000165 0.1999905 0.2000091 0.1999931 0.1999908
[180,] 0.2000178 0.1999921 0.2000075 0.1999919 0.1999906
[181,] 0.2000121 0.1999925 0.2000075 0.1999943 0.1999936
[182,] 0.2000179 0.1999928 0.2000099 0.1999916 0.1999878
[183,] 0.2000271 0.1999877 0.2000145 0.1999854 0.1999852
[184,] 0.2000235 0.1999875 0.2000127 0.1999876 0.1999888
[185,] 0.2000066 0.1999957 0.2000050 0.1999940 0.1999987
[186,] 0.2000134 0.1999888 0.2000113 0.1999897 0.1999967
[187,] 0.2000354 0.1999788 0.2000196 0.1999769 0.1999893
[188,] 0.2000172 0.1999900 0.2000092 0.1999929 0.1999907
[189,] 0.2000482 0.1999673 0.2000282 0.1999779 0.1999784
[190,] 0.2000259 0.1999876 0.2000097 0.1999904 0.1999863
[191,] 0.2000160 0.1999938 0.2000087 0.1999923 0.1999892
[192,] 0.2000180 0.1999878 0.2000112 0.1999882 0.1999948
[193,] 0.2000193 0.1999872 0.2000144 0.1999866 0.1999926
[194,] 0.2000108 0.1999934 0.2000077 0.1999950 0.1999931
[195,] 0.2000215 0.1999864 0.2000147 0.1999850 0.1999924
[196,] 0.2000100 0.1999942 0.2000044 0.1999962 0.1999952
[197,] 0.2000218 0.1999894 0.2000127 0.1999892 0.1999870
[198,] 0.2000295 0.1999812 0.2000170 0.1999894 0.1999830

```
[199,] 0.2000076 0.1999942 0.2000043 0.1999974 0.1999964
[200,] 0.2000173 0.1999909 0.2000062 0.1999936 0.1999919
[ reached getOption("max.print") -- omitted 1811 rows ]
```

Closest hard clustering:

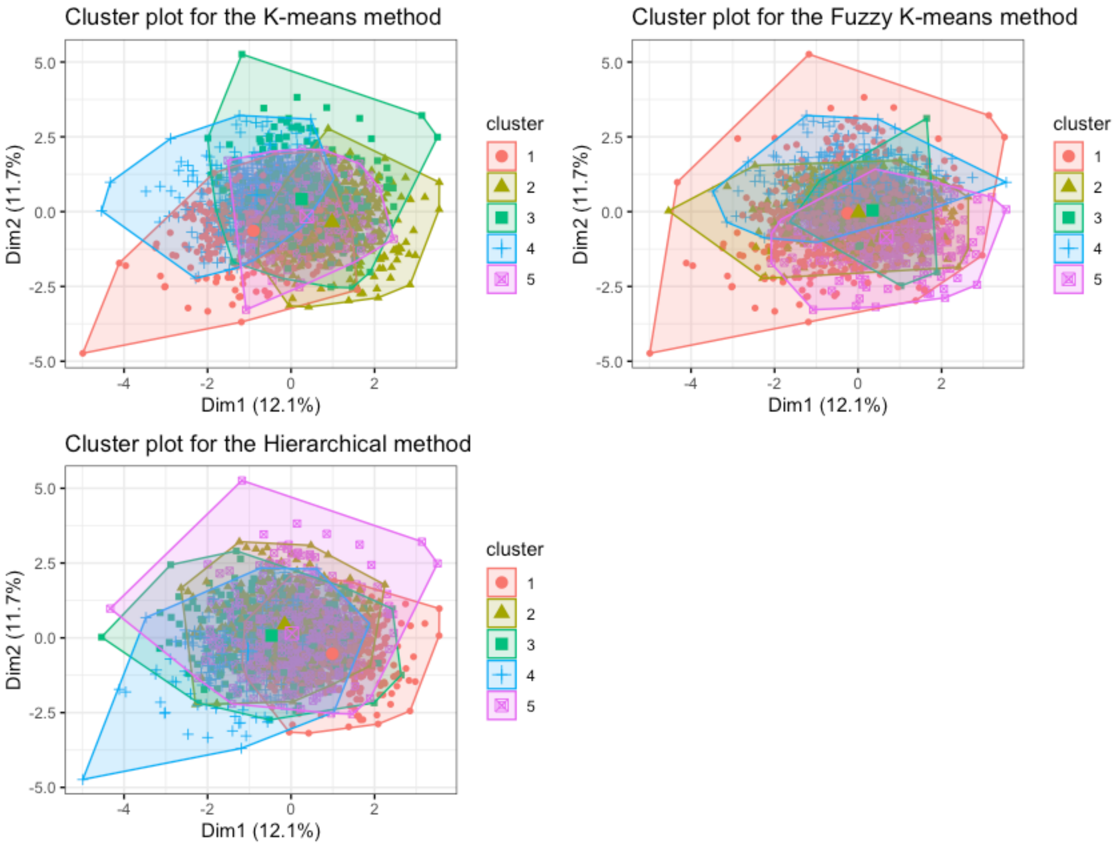
```
  [1] 2 5 4 5 5 5 4 5 4 5 5 2 4 2 4 2 2 5 4 4 5 4 2 5 2 2 5 2 2 5
 [31] 5 5 2 4 5 2 5 2 4 2 4 4 5 4 2 2 2 2 5 5 5 5 5 2 2 2 4 4 4 5
 [61] 5 2 4 5 2 5 4 2 5 4 5 2 4 5 5 4 5 5 4 2 4 4 5 2 4 2 5 5 5 2
 [91] 2 2 2 2 4 5 2 4 4 5 2 2 2 2 4 2 2 2 5 2 4 2 2 2 2 4 2 2 2 5
[121] 5 2 5 2 2 5 5 2 2 4 5 5 5 4 5 5 5 2 2 2 5 4 2 5 4 4 5 2 5 2
[151] 5 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[181] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[211] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1
[241] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 5 4 2 2 5 2 4 5 2 5 4 5 2 5
[271] 2 5 5 5 5 5 4 5 5 4 2 4 2 5 5 2 2 4 5 2 2 4 5 2 2 5 5 4 2 4
[301] 5 2 4 5 2 2 2 4 4 2 2 2 5 5 4 4 2 2 2 2 4 4 5 2 2 4 5 5 4 2
[331] 4 2 4 5 2 2 5 2 2 4 4 4 4 4 4 4 2 4 5 5 2 5 5 2 4 4 4 5 4 4
[361] 4 2 2 4 2 4 4 5 5 5 5 2 5 2 2 5 2 2 2 5 2 4 5 2 2 2 4 2 4 5
[391] 5 2 4 4 4 2 4 4 2 2 2 5 4 2 5 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[421] 1 1 1 1 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[451] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[481] 1 1 3 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 4 4 2 2 5
[511] 2 4 4 2 2 2 5 4 2 4 2 5 2 2 4 2 4 2 5 5 2 2 5 5 4 5 4 4 2 4
[541] 2 2 5 4 5 2 4 2 2 4 2 4 4 2 5 4 2 4 4 5 4 4 5 2 5 2 5 4 5 5
[571] 4 4 2 5 4 5 4 2 5 4 5 4 2 4 2 2 2 2 5 2 5 4 4 4 5 5 5 5 4 2
[601] 2 4 2 4 2 5 5 5 5 5 4 4 2 5 4 4 4 5 4 2 2 2 5 4 2 5 2 2 5 5
[631] 4 4 5 2 2 5 5 2 2 2 2 4 5 4 5 2 5 4 4 5 2 2 2 2 2 5 2 2 5 2
[661] 2 4 5 2 4 5 4 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[691] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[721] 1 1 1 1 3 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[751] 1 1 1 1 1 1 1 1 1 1 1 1 2 4 5 2 2 2 5 2 5 4 2 2 2 2 5 2 2 2 2
[781] 2 2 5 5 5 2 2 2 2 5 2 4 2 4 5 4 5 5 2 2 2 4 5 2 4 2 4 4 2 5
```

```
[811] 5 4 2 4 4 4 2 2 4 2 2 4 5 4 2 4 5 4 4 4 2 2 2 5 5 4 5 5 5 2
[841] 2 2 2 2 2 2 5 2 4 2 5 4 5 4 5 2 2 4 2 5 4 2 2 5 2 2 5 2 2 4
[871] 5 4 5 4 4 5 4 2 5 2 4 2 5 4 4 2 5 5 2 4 4 4 2 2 5 5 4 2 4 2
[901] 2 2 2 4 4 4 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[931] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[961] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
[991] 1 1 1 1 1 1 1 1 1 1
 [ reached getOption("max.print") -- omitted 1011 entries ]

Available components:
[1] "centers"     "size"        "cluster"     "membership"
[5] "iter"        "withinerror" "call"
```

Agrupamento usando o método hierárquico

```
Call:
stats::hclust(d = x, method = hc_method)

Cluster method   : ward.D2
Distance         : euclidean
Number of objects: 2011
```

Gráfico 10



O (gráfico 10) mostra os trés gráficos juntos (agrupamento usando o K-means, agrupamento usando o Fuzzy K-means e o agrupamento usando o método hierárquico).

Medidas de qualidade

Conectividade: varia de zero a infinito e deve ser minimizado:

c1

[1] 1228.178

c2

[1] 1005.103

> c3

[1] 1148.206

Dunn: varia de zero a infinito e deve ser maximizado:

d1

[1] 0.07689555

```
> d2
[1] 0.06866552
> d3
[1] 0.09785363
```

Silhueta: varia de -1 a 1. Quanto mais próximo de 1, melhor:

```
s.kmeans
      cluster neighbor     sil_width
  [1,]       2        4  1.831844e-01
  [2,]       5        2  2.896842e-01
  [3,]       4        5  1.687673e-01
  [4,]       2        5  2.604730e-01
  [5,]       2        5  3.784808e-02
  [6,]       2        5  1.035331e-01
  [7,]       4        2  1.626542e-01
  [8,]       2        5  7.739257e-02
  [9,]       4        2  3.263218e-01
 [10,]       2        5  1.317764e-01
 [11,]       2        5  1.447563e-01
 [12,]       4        2  1.435883e-01
 [13,]       4        5  2.354495e-01
 [14,]       4        5  1.942720e-01
 [15,]       4        5  9.349498e-02
 [16,]       2        4  5.998300e-02
 [17,]       2        5 -7.622350e-03
 [18,]       5        2  7.017586e-02
 [19,]       4        5  4.581699e-02
 [20,]       4        2  3.093371e-01
 [21,]       5        2  4.252714e-01
 [22,]       4        5  1.563988e-01
 [23,]       4        5 -1.563359e-03
 [24,]       5        2  1.109723e-01
 [25,]       2        5  3.095824e-01
```

[26,]	2	5	2.564220e-01
[27,]	5	2	1.815873e-01
[28,]	2	4	8.433189e-02
[29,]	4	2	3.702329e-02
[30,]	5	2	1.895317e-01
[31,]	5	2	2.836354e-01
[32,]	5	2	2.472065e-02
[33,]	4	2	2.350210e-01
[34,]	4	5	2.415453e-01
[35,]	5	4	3.746885e-01
[36,]	4	2	2.201505e-01
[37,]	5	2	2.529225e-01
[38,]	2	4	2.287303e-01
[39,]	4	5	2.556186e-01
[40,]	4	5	1.018729e-01
[41,]	4	2	3.509787e-01
[42,]	4	5	2.916135e-01
[43,]	5	2	3.237393e-02
[44,]	4	2	6.588044e-02
[45,]	2	4	3.200135e-01
[46,]	2	4	9.260463e-02
[47,]	4	2	1.012709e-01
[48,]	4	5	2.317469e-02
[49,]	2	5	2.476426e-02
[50,]	2	5	1.370838e-01
[51,]	5	2	3.685490e-02
[52,]	2	5	1.215756e-01
[53,]	2	5	7.068125e-02
[54,]	5	4	1.157344e-01
[55,]	2	4	4.201757e-02
[56,]	4	2	3.305081e-01
[57,]	4	5	1.124248e-01

[58,]	2	5	2.029458e-01
[59,]	4	5	2.378853e-01
[60,]	5	2	3.108515e-01
[61,]	5	2	8.935437e-02
[62,]	2	5	1.408178e-01
[63,]	4	5	1.203607e-01
[64,]	5	2	2.245238e-01
[65,]	2	5	1.825010e-01
[66,]	5	2	9.250533e-02
[67,]	4	2	1.936065e-01
[68,]	2	5	2.400548e-02
[69,]	2	5	9.301668e-02
[70,]	4	5	1.600393e-01
[71,]	2	5	1.114412e-01
[72,]	2	5	1.882876e-01
[73,]	4	2	1.046156e-01
[74,]	5	2	3.230883e-01
[75,]	5	2	6.896017e-02
[76,]	4	5	2.256980e-01
[77,]	5	2	1.244587e-01
[78,]	5	4	1.869642e-01
[79,]	5	4	1.466898e-01
[80,]	5	4	7.841933e-02
[81,]	4	2	1.015138e-01
[82,]	2	4	5.882869e-02
[83,]	2	5	1.407343e-01
[84,]	5	2	4.837604e-02
[85,]	4	5	1.551306e-01
[86,]	2	4	1.279042e-01
[87,]	5	4	1.644451e-01
[88,]	2	5	-3.417609e-02
[89,]	5	2	1.024634e-01

[90,]	2	5	1.635345e-01
[91,]	2	4	2.255233e-01
[92,]	4	2	2.883143e-01
[93,]	4	2	1.032701e-02
[94,]	4	5	1.948573e-01
[95,]	2	5	1.743417e-01
[96,]	5	2	3.293729e-01
[97,]	2	5	1.843015e-01
[98,]	4	5	1.378780e-01
[99,]	5	4	1.250863e-01
[100,]	2	5	4.628649e-02
[101,]	5	2	1.843438e-01
[102,]	2	4	2.103636e-01
[103,]	2	4	7.446782e-02
[104,]	2	5	1.695568e-01
[105,]	4	5	3.046943e-01
[106,]	2	4	4.959727e-02
[107,]	4	2	8.031656e-02
[108,]	4	2	2.923128e-01
[109,]	5	2	4.153338e-02
[110,]	4	5	1.168788e-01
[111,]	4	5	1.673807e-01
[112,]	5	4	2.001149e-01
[113,]	4	2	1.135556e-01
[114,]	4	2	1.683044e-01
[115,]	4	2	2.034093e-01
[116,]	4	5	1.787513e-01
[117,]	2	4	1.066148e-01
[118,]	5	2	1.572794e-01
[119,]	2	5	1.805401e-01
[120,]	5	4	1.485314e-01
[121,]	2	5	5.518096e-02

[122,]	4	5	1.621677e-01
[123,]	5	4	2.689460e-01
[124,]	2	4	3.822342e-01
[125,]	5	2	1.197814e-01
[126,]	5	2	5.105419e-02
[127,]	2	5	-8.376638e-04
[128,]	4	2	1.458188e-01
[129,]	4	2	1.580580e-01
[130,]	4	5	2.013258e-01
[131,]	2	5	3.283301e-02
[132,]	2	5	8.116167e-02
[133,]	5	2	1.410017e-01
[134,]	4	2	2.216038e-01
[135,]	2	5	4.018762e-02
[136,]	2	5	2.555946e-03
[137,]	5	2	2.701674e-01
[138,]	4	5	-8.506338e-04
[139,]	2	5	2.122364e-01
[140,]	2	5	2.279841e-02
[141,]	5	2	6.312962e-02
[142,]	4	5	1.636501e-02
[143,]	5	4	5.657555e-02
[144,]	2	5	7.327758e-03
[145,]	4	5	1.863949e-01
[146,]	4	5	5.936550e-02
[147,]	5	2	6.892592e-02
[148,]	5	4	3.946993e-02
[149,]	2	5	3.444881e-01
[150,]	5	2	1.042631e-01
[151,]	5	2	2.815286e-01
[152,]	2	4	8.518822e-02
[153,]	4	2	4.028414e-02

[154,]	4	2	5.985380e-02
[155,]	1	2	5.842928e-02
[156,]	1	3	1.989766e-01
[157,]	1	4	2.681399e-01
[158,]	1	3	1.752428e-01
[159,]	3	2	-1.779166e-02
[160,]	1	5	1.782404e-01
[161,]	1	3	1.483802e-01
[162,]	3	1	1.074368e-01
[163,]	3	5	2.160479e-01
[164,]	3	1	5.900625e-02
[165,]	3	2	1.860092e-01
[166,]	3	2	2.748629e-01
[167,]	3	1	7.021404e-02
[168,]	3	1	2.560815e-01
[169,]	1	3	8.667172e-02
[170,]	3	2	4.679697e-02
[171,]	1	4	1.846520e-01
[172,]	1	5	1.023449e-01
[173,]	1	4	1.745192e-01
[174,]	1	3	1.890743e-01
[175,]	1	4	1.245028e-01
[176,]	5	3	7.362302e-02
[177,]	1	5	1.785367e-01
[178,]	3	2	2.089281e-01
[179,]	1	2	1.824627e-01
[180,]	1	5	2.418021e-01
[181,]	1	4	2.412285e-01
[182,]	1	4	2.409267e-01
[183,]	1	4	2.169087e-01
[184,]	1	4	3.130290e-01
[185,]	3	2	1.376751e-01

[186,]	3	5	2.356902e-01
[187,]	1	3	2.491753e-01
[188,]	3	5	1.520628e-01
[189,]	1	3	9.534244e-02
[190,]	1	4	2.660292e-01
[191,]	1	2	6.814180e-02
[192,]	3	1	9.392700e-02
[193,]	3	5	4.741685e-02
[194,]	3	2	1.228699e-02
[195,]	3	1	1.276435e-01
[196,]	3	4	8.836078e-02
[197,]	1	4	2.730647e-01
[198,]	1	3	8.236711e-02
[199,]	1	5	1.929464e-01
[200,]	1	3	7.758022e-02
[201,]	1	4	2.353030e-01
[202,]	3	5	1.299993e-01
[203,]	1	4	2.338930e-01
[204,]	1	3	1.063594e-02
[205,]	1	3	1.230400e-01
[206,]	3	1	1.048622e-01
[207,]	3	2	2.783824e-01
[208,]	1	3	1.453771e-01
[209,]	1	3	-1.756803e-02
[210,]	3	5	2.641409e-01
[211,]	1	3	2.023083e-01
[212,]	3	2	2.445858e-01
[213,]	3	2	8.383845e-02
[214,]	3	5	1.350200e-01
[215,]	3	2	2.683484e-01
[216,]	1	3	9.365567e-02
[217,]	1	3	3.909741e-02

[218,]	1	2	-2.812185e-02
[219,]	1	4	2.530845e-01
[220,]	3	2	1.662162e-01
[221,]	3	5	1.059365e-01
[222,]	3	2	1.617319e-01
[223,]	3	2	4.979963e-02
[224,]	3	4	1.105973e-01
[225,]	3	5	2.369552e-01
[226,]	3	5	1.011413e-01
[227,]	3	1	2.114090e-01
[228,]	3	2	2.269510e-01
[229,]	1	4	1.370883e-01
[230,]	1	3	-1.821984e-02
[231,]	1	5	-5.052465e-02
[232,]	3	5	2.754825e-01
[233,]	1	4	1.518089e-01
[234,]	3	1	1.103200e-01
[235,]	1	2	1.977384e-01
[236,]	1	4	8.573649e-02
[237,]	1	3	1.093260e-01
[238,]	1	4	2.462425e-01
[239,]	3	1	2.378530e-01
[240,]	1	4	6.051549e-02
[241,]	3	2	-1.149106e-02
[242,]	1	5	1.420607e-01
[243,]	3	1	1.441381e-01
[244,]	1	2	5.911036e-02
[245,]	3	2	1.650691e-01
[246,]	1	3	9.934540e-02
[247,]	3	1	1.478189e-01
[248,]	3	5	1.059840e-01
[249,]	1	4	1.826691e-01

[250,]	3	5	1.578921e-01
[251,]	3	1	1.152664e-01
[252,]	1	5	8.392757e-02
[253,]	3	2	2.685246e-01
[254,]	3	2	5.670052e-02
[255,]	3	5	7.390700e-02
[256,]	3	4	2.411225e-01
[257,]	5	4	2.222750e-01
[258,]	4	5	1.688640e-01
[259,]	4	5	1.860308e-01
[260,]	4	5	4.725454e-02
[261,]	4	2	5.665839e-02
[262,]	2	5	3.835218e-01
[263,]	4	5	3.376025e-01
[264,]	5	2	1.876546e-01
[265,]	2	4	4.956960e-02
[266,]	2	4	4.499386e-02
[267,]	4	2	9.173724e-03
[268,]	4	2	6.915406e-02
[269,]	4	2	1.338653e-01
[270,]	2	5	-1.548491e-02
[271,]	2	4	1.598666e-01
[272,]	2	5	9.899549e-02
[273,]	2	5	2.562161e-01
[274,]	5	2	1.467064e-01
[275,]	5	4	3.514145e-01
[276,]	2	5	4.004941e-02
[277,]	4	1	2.470980e-01
[278,]	4	5	-1.776947e-02
[279,]	2	5	1.948848e-01
[280,]	4	5	2.211463e-01
[281,]	2	4	1.722388e-01

[282,]	5	4	1.792576e-01
[283,]	2	5	1.204390e-01
[284,]	5	2	1.735909e-01
[285,]	5	2	1.521683e-01
[286,]	2	5	3.224471e-01
[287,]	2	4	1.082601e-01
[288,]	4	5	1.714958e-01
[289,]	2	5	2.095243e-01
[290,]	2	5	1.693059e-01
[291,]	2	4	2.555228e-01
[292,]	2	5	2.564902e-02
[293,]	5	2	3.223737e-01
[294,]	2	4	1.838987e-01
[295,]	4	2	1.108433e-01
[296,]	2	5	3.893998e-02
[297,]	5	2	6.470776e-02
[298,]	4	5	2.892609e-01
[299,]	5	4	3.189754e-01
[300,]	4	5	1.405311e-01
[301,]	2	5	2.702115e-01
[302,]	5	4	6.330940e-02
[303,]	2	5	1.401120e-01
[304,]	5	4	4.025461e-01
[305,]	4	2	1.590486e-01
[306,]	4	5	-4.788160e-02
[307,]	2	4	8.126767e-02
[308,]	2	4	1.660821e-01
[309,]	4	5	5.027342e-02
[310,]	4	5	7.463574e-02
[311,]	2	5	-1.303413e-02
[312,]	4	2	5.533309e-06
[313,]	5	4	3.041489e-01

```
 [314,]       2        5  9.142482e-02
 [315,]       4        2 -3.355676e-03
 [316,]       5        4  5.839333e-02
 [317,]       2        5  6.263873e-02
 [318,]       4        5  1.326782e-01
 [319,]       4        5  2.498434e-01
 [320,]       4        2  1.600662e-01
 [321,]       4        2  1.667849e-01
 [322,]       4        5  1.950412e-01
 [323,]       5        2  2.457182e-01
 [324,]       2        4  2.747761e-01
 [325,]       4        2  2.219867e-02
 [326,]       4        5  1.101215e-01
 [327,]       4        5  2.796964e-02
 [328,]       4        5  2.500510e-02
 [329,]       4        5 -1.918989e-03
 [330,]       5        4  7.392956e-02
 [331,]       5        4  8.811792e-02
 [332,]       2        5  2.475462e-01
 [333,]       4        5  1.335788e-01
 [ reached getOption("max.print") -- omitted 1678 rows ]
attr(,"Ordered")
[1] FALSE
attr(,"call")
silhouette.default(x = fit.kmeans$cluster, dist = dist(df)^2)
attr(,"class")
[1] "silhouette"

> s.cmeans
       cluster neighbor    sil_width
  [1,]       2        4  0.2154876183
  [2,]       5        4  0.2177738708
```

```
 [3,]    4    2  0.2644617693
 [4,]    5    2  0.1183303865
 [5,]    5    2  0.0426623779
 [6,]    5    2  0.0384574634
 [7,]    4    2  0.0709831289
 [8,]    5    2 -0.0131594416
 [9,]    4    2  0.0233212216
[10,]    5    2  0.0256843469
[11,]    5    2  0.2344498660
[12,]    2    4  0.0725731086
[13,]    4    2  0.1210143652
[14,]    2    4  0.2393131800
[15,]    4    2  0.2602753331
[16,]    2    4  0.2420374063
[17,]    2    4  0.0886751298
[18,]    5    4  0.0731973478
[19,]    4    2  0.0737775680
[20,]    4    2 -0.0087545846
[21,]    5    4  0.3201486424
[22,]    4    2  0.0502712762
[23,]    2    4  0.1355956681
[24,]    5    2  0.1034797904
[25,]    2    5  0.0421625307
[26,]    2    5  0.0397278817
[27,]    5    2  0.3249708676
[28,]    2    4  0.0196793701
[29,]    2    4  0.1347502952
[30,]    5    2  0.2884020360
[31,]    5    4  0.3126692974
[32,]    5    2  0.0495104111
[33,]    2    4  0.1766108256
[34,]    4    2  0.1286283016
```

[35,]	5	4	0.1743445461
[36,]	2	4	0.1536182904
[37,]	5	4	0.1249018280
[38,]	2	5	0.1755589082
[39,]	4	2	0.1307408305
[40,]	2	4	0.0974130391
[41,]	4	2	0.1245415166
[42,]	4	2	0.2581622529
[43,]	5	4	0.0383597339
[44,]	4	2	0.0733305093
[45,]	2	5	0.2143710243
[46,]	2	4	0.0758516943
[47,]	2	4	0.2034794475
[48,]	2	4	0.2272877077
[49,]	5	2	0.3011204424
[50,]	5	2	0.0127887085
[51,]	5	4	0.2350235490
[52,]	5	3	0.2207786155
[53,]	5	2	-0.0010672798
[54,]	2	4	0.0864351388
[55,]	2	4	0.2752211318
[56,]	2	4	0.1048919205
[57,]	4	2	0.1123109535
[58,]	4	5	0.0231298812
[59,]	4	2	0.2742472384
[60,]	5	2	0.0423443111
[61,]	5	2	0.0817438541
[62,]	2	5	0.2112254171
[63,]	4	2	0.1573808161
[64,]	5	2	0.2759490290
[65,]	2	5	0.1506838497
[66,]	5	4	-0.0145447499

[67,]	4	2	0.2218743008
[68,]	2	5	0.1288017578
[69,]	5	2	0.3488362567
[70,]	4	2	0.1412114711
[71,]	5	2	0.0378721829
[72,]	2	5	-0.0391396097
[73,]	4	2	0.0189259933
[74,]	5	4	0.2029780061
[75,]	5	2	-0.0017674567
[76,]	4	2	0.2303669115
[77,]	5	4	0.1165935541
[78,]	5	4	0.0544737982
[79,]	4	5	0.1868807720
[80,]	2	4	0.0042551808
[81,]	4	2	0.1753835363
[82,]	4	2	0.0884244607
[83,]	5	2	0.2074062262
[84,]	2	5	0.0728311336
[85,]	4	5	0.1509634140
[86,]	2	4	0.2570510009
[87,]	5	4	0.0219181449
[88,]	5	2	0.0217629075
[89,]	5	2	0.1445272065
[90,]	2	4	0.2215711734
[91,]	2	4	0.0374514362
[92,]	2	4	0.0579316369
[93,]	2	4	0.1482530352
[94,]	2	4	0.0569110242
[95,]	4	5	0.0310335805
[96,]	5	2	0.3398237914
[97,]	2	5	0.1847020125
[98,]	4	2	0.0734502095

[99,]	4	2	0.0505231515
[100,]	5	2	0.0940906354
[101,]	2	5	0.0999764260
[102,]	2	4	0.1045208060
[103,]	2	5	0.0739809370
[104,]	2	5	0.1781564934
[105,]	4	2	0.2187105396
[106,]	2	5	0.0569644300
[107,]	2	4	0.1435245804
[108,]	2	4	0.0160224657
[109,]	5	4	0.0642984013
[110,]	2	4	0.0012706411
[111,]	4	2	0.1181305660
[112,]	2	4	0.0214273181
[113,]	2	4	0.1109670029
[114,]	2	4	0.0491243034
[115,]	2	4	0.0657204753
[116,]	4	2	0.1462312222
[117,]	2	4	0.1140584159
[118,]	2	5	0.0390122636
[119,]	2	5	0.1328422165
[120,]	5	4	-0.0213969358
[121,]	5	2	0.2897927799
[122,]	2	4	0.0449863100
[123,]	5	2	0.0048170822
[124,]	2	5	0.2073003032
[125,]	2	5	0.0745913076
[126,]	5	2	0.1665262789
[127,]	5	2	0.1717083909
[128,]	2	4	0.0577798887
[129,]	2	4	0.2398974230
[130,]	4	2	0.0241369616

[131,]	5	2	0.0912752081
[132,]	5	2	0.1599235134
[133,]	5	2	0.2758136943
[134,]	4	2	0.1588645148
[135,]	5	3	-0.0280425931
[136,]	5	2	0.1372264226
[137,]	5	2	0.0250491751
[138,]	2	4	0.0154330450
[139,]	2	5	0.1292405340
[140,]	2	5	0.0667951322
[141,]	5	2	0.1017326510
[142,]	4	5	0.0949518733
[143,]	2	4	0.2087149968
[144,]	5	4	0.1360209056
[145,]	4	2	0.1902511126
[146,]	4	2	0.0137101054
[147,]	5	2	0.0770211280
[148,]	2	5	0.0061373672
[149,]	5	2	-0.0645458715
[150,]	2	4	0.0093871437
[151,]	5	4	0.3120437192
[152,]	2	5	0.2718078293
[153,]	2	4	0.0214189844
[154,]	2	4	0.2600118436
[155,]	1	2	0.0201136610
[156,]	1	2	0.1235743221
[157,]	1	2	0.0567922774
[158,]	1	2	0.1669422812
[159,]	1	5	0.0954613527
[160,]	1	5	0.1691496611
[161,]	1	4	0.0172829361
[162,]	1	4	0.1725775832

[163,]	1	5	0.1002558380
[164,]	1	2	0.0291033217
[165,]	1	3	0.0547655404
[166,]	1	2	0.0829997332
[167,]	1	2	0.1883419768
[168,]	1	2	0.1097803166
[169,]	1	2	0.0539416066
[170,]	1	3	0.1138580182
[171,]	1	2	0.0712523210
[172,]	1	2	0.1541150796
[173,]	1	2	0.0076981694
[174,]	1	4	0.1479611852
[175,]	1	2	0.0370154427
[176,]	1	3	-0.1309062723
[177,]	1	4	0.0554272770
[178,]	1	3	0.0330377727
[179,]	1	4	0.0947073967
[180,]	1	2	0.0656120683
[181,]	1	4	0.0685980994
[182,]	1	2	-0.0051274989
[183,]	1	2	0.0518518450
[184,]	1	2	0.1069492280
[185,]	1	3	-0.0769548323
[186,]	1	3	-0.1990880097
[187,]	1	2	0.1461434920
[188,]	1	4	0.1123655086
[189,]	1	4	0.1215305922
[190,]	1	4	0.1112388689
[191,]	1	2	-0.0137622843
[192,]	1	3	0.1020401376
[193,]	1	3	0.0467272778
[194,]	1	2	0.0788423430

[195,]	1	2	0.0948619384
[196,]	1	4	0.0555981146
[197,]	1	2	0.0728635684
[198,]	1	4	0.0349220702
[199,]	1	4	0.0089280638
[200,]	1	4	0.1161585267
[201,]	1	2	0.0531631065
[202,]	1	5	-0.0129798663
[203,]	1	2	0.1003215104
[204,]	1	2	0.1631725915
[205,]	1	2	0.1910538974
[206,]	1	4	0.1718463947
[207,]	1	3	0.0127324691
[208,]	1	2	-0.0234490058
[209,]	1	4	0.0317232105
[210,]	1	2	0.1288765325
[211,]	1	2	0.1225262734
[212,]	1	3	0.0125926077
[213,]	1	3	-0.0642219993
[214,]	1	3	0.0500437905
[215,]	1	2	0.0756225760
[216,]	1	4	0.1578592051
[217,]	1	5	0.1743449509
[218,]	1	2	-0.0523806244
[219,]	1	2	0.0270274097
[220,]	1	2	0.1755084863
[221,]	1	5	0.1048424191
[222,]	1	2	0.0203192857
[223,]	1	2	-0.0260499299
[224,]	1	2	0.0995601726
[225,]	1	3	-0.2261168356
[226,]	1	4	-0.0133465230

[227,]	1	2	0.1256723682
[228,]	3	5	0.3779348618
[229,]	1	2	0.0305612835
[230,]	1	2	0.0659809080
[231,]	1	4	0.1446686527
[232,]	1	3	0.0930532449
[233,]	1	4	0.0064533346
[234,]	1	3	-0.0239519174
[235,]	1	2	0.0872101685
[236,]	1	4	0.1015697143
[237,]	1	4	0.0077003545
[238,]	1	2	0.0733558652
[239,]	1	4	0.0971236466
[240,]	1	2	-0.0362367952
[241,]	1	5	0.0975754407
[242,]	1	3	0.0762649267
[243,]	1	2	0.0761278673
[244,]	1	4	0.1010531783
[245,]	1	3	-0.0813817573
[246,]	1	4	0.1886037070
[247,]	1	2	0.1210270941
[248,]	1	3	-0.1031813665
[249,]	1	2	0.0648609418
[250,]	1	5	0.1017125736
[251,]	1	2	0.0481944503
[252,]	1	3	0.1469791563
[253,]	1	3	-0.1456050256
[254,]	1	5	0.0671957831
[255,]	1	3	-0.1254538397
[256,]	1	3	-0.0437291041
[257,]	5	2	0.1009825153
[258,]	4	2	0.0623147871

[259,]	2	4	0.1624401134
[260,]	2	4	-0.0166549472
[261,]	5	2	-0.0682331924
[262,]	2	5	0.1992901054
[263,]	4	2	0.1622347655
[264,]	5	3	-0.0922316408
[265,]	2	5	0.2715156452
[266,]	5	2	0.0658498043
[267,]	4	2	0.0160736495
[268,]	5	4	0.0296573467
[269,]	2	4	0.2965548847
[270,]	5	2	0.1964563852
[271,]	2	4	0.1943763453
[272,]	5	2	0.0879784839
[273,]	5	2	0.0716714939
[274,]	5	2	0.1025964311
[275,]	5	2	0.1111389626
[276,]	5	2	0.3092448578
[277,]	4	2	0.0700300751
[278,]	5	2	0.0549494494
[279,]	5	2	0.2755527215
[280,]	4	2	0.1151639599
[281,]	2	4	0.0848271578
[282,]	4	2	0.0595839996
[283,]	2	5	0.1381305311
[284,]	5	2	0.1659947809
[285,]	5	2	0.2071697111
[286,]	2	4	0.0259916126
[287,]	2	4	0.2037371779
[288,]	4	2	0.2484134165
[289,]	5	2	0.2294493669
[290,]	2	4	0.0025589901

[291,]	2	5	0.2249827480
[292,]	4	5	0.0144226437
[293,]	5	4	0.0635128498
[294,]	2	4	0.0510702672
[295,]	2	4	-0.0086235438
[296,]	5	4	0.1612931034
[297,]	5	4	0.0958724275
[298,]	4	2	0.1101593373
[299,]	2	5	-0.0066419411
[300,]	4	2	0.1238250581
[301,]	5	2	0.1848709714
[302,]	2	4	0.2032159416
[303,]	4	5	-0.0269890455
[304,]	5	4	0.0150294779
[305,]	2	4	0.2134295581
[306,]	2	4	0.0281460939
[307,]	2	4	0.0938988675
[308,]	4	2	0.1403855102
[309,]	4	2	0.1704102828
[310,]	2	4	0.0150446249
[311,]	2	5	0.1925452451
[312,]	2	5	0.2134771810
[313,]	5	4	0.0358304726
[314,]	5	4	0.0707870796
[315,]	4	5	0.0657158319
[316,]	4	2	0.0720877816
[317,]	2	5	0.1036133085
[318,]	2	4	0.2265866980
[319,]	2	4	0.0492519822
[320,]	2	4	0.1265009691
[321,]	4	2	0.0707736241
[322,]	4	2	0.1802229098

```
[323,]       5        4  0.2294555468
[324,]       2        5  0.2455465568
[325,]       2        4  0.0923489647
[326,]       4        2  0.0493233875
[327,]       5        3  0.0455609751
[328,]       5        4 -0.0247233723
[329,]       4        2  0.2189371528
[330,]       2        5  0.2431625917
[331,]       4        2  0.0061729324
[332,]       2        5  0.1072234016
[333,]       4        2  0.1109403866
 [ reached getOption("max.print") -- omitted 1678 rows ]
attr(,"Ordered")
[1] FALSE
attr(,"call")
silhouette.default(x = fit.cmeans$cluster, dist = dist(df)^2)
attr(,"class")
[1] "silhouette"

> s.hclust
       cluster neighbor     sil_width
  [1,]       1        2  0.1050899565
  [2,]       2        1  0.0533908487
  [3,]       3        2  0.2961344762
  [4,]       1        3  0.1476883081
  [5,]       2        1 -0.0228558836
  [6,]       1        2  0.2354852149
  [7,]       3        2 -0.0183307159
  [8,]       2        1  0.0729061190
  [9,]       3        2  0.0125521246
 [10,]       1        2  0.2243088195
 [11,]       1        3  0.0477483918
```

[12,]	1	2	-0.1385166982
[13,]	2	3	0.2058253022
[14,]	2	1	0.2537237133
[15,]	3	2	0.0856016555
[16,]	2	1	0.1914169243
[17,]	4	2	-0.1768972017
[18,]	1	3	-0.1204657370
[19,]	2	3	0.1036891936
[20,]	3	2	0.0584828165
[21,]	1	2	0.0306750141
[22,]	2	3	0.0748913269
[23,]	2	1	0.1236851897
[24,]	1	2	0.1893132641
[25,]	2	1	-0.1172985852
[26,]	1	2	0.0738960311
[27,]	1	2	0.0925431812
[28,]	2	1	0.0732470142
[29,]	2	3	0.2223600525
[30,]	1	2	0.2516364688
[31,]	1	2	0.1369910212
[32,]	1	3	-0.0177364904
[33,]	2	3	0.0724081853
[34,]	3	2	0.2032057890
[35,]	2	3	0.0464948306
[36,]	1	2	-0.1370421121
[37,]	2	1	0.0373710086
[38,]	2	1	0.0578977576
[39,]	2	3	0.1100621693
[40,]	2	1	0.3541488913
[41,]	3	2	0.0725046032
[42,]	3	2	0.0860005156
[43,]	1	3	-0.1012001946

[44,]	3	2	0.1783660315
[45,]	1	2	0.0496459557
[46,]	1	3	-0.0505028974
[47,]	2	3	0.1979655311
[48,]	1	2	-0.1501466227
[49,]	1	2	0.2875228214
[50,]	1	2	0.1868440292
[51,]	1	3	0.1112676480
[52,]	1	3	0.1977423657
[53,]	2	1	0.0544145933
[54,]	2	1	0.2673800347
[55,]	2	1	0.1232597630
[56,]	2	3	0.2536916489
[57,]	2	3	0.1707432211
[58,]	3	1	-0.0540681634
[59,]	3	2	0.0843396394
[60,]	2	1	0.0139348288
[61,]	1	2	0.1718134756
[62,]	1	2	0.0168995272
[63,]	3	2	0.0378307334
[64,]	1	2	0.0640159803
[65,]	1	2	-0.0278934914
[66,]	1	3	0.0088398952
[67,]	3	2	0.2277915100
[68,]	1	2	0.1238883406
[69,]	1	2	0.3679012025
[70,]	2	3	0.0888515791
[71,]	1	3	0.2269737617
[72,]	1	3	0.0345169826
[73,]	2	3	0.0182916807
[74,]	1	2	-0.0070836574
[75,]	2	1	0.1365332645

[76,]	3	2	0.0658401845
[77,]	1	2	0.0045611044
[78,]	1	3	-0.0552915707
[79,]	2	3	-0.0318901897
[80,]	2	3	0.1908565778
[81,]	2	3	0.0020810518
[82,]	3	1	0.0314805064
[83,]	1	2	0.2476234321
[84,]	1	2	0.0126698021
[85,]	3	2	0.2109470655
[86,]	1	2	0.0509649902
[87,]	1	3	-0.1647222849
[88,]	2	1	0.0529680563
[89,]	1	2	0.2331972177
[90,]	2	1	0.0594660916
[91,]	1	3	-0.0035277061
[92,]	2	3	0.1648438513
[93,]	2	3	0.1367548540
[94,]	2	3	0.1943732789
[95,]	3	1	0.1855117425
[96,]	2	1	-0.1155193215
[97,]	2	1	-0.1028620327
[98,]	2	3	0.2942869591
[99,]	2	1	0.2774159826
[100,]	1	3	0.2534057441
[101,]	2	1	0.0569083583
[102,]	2	1	0.0261155966
[103,]	1	2	0.0534798167
[104,]	2	1	-0.0012496155
[105,]	3	2	-0.0798804456
[106,]	2	1	-0.0064726965
[107,]	2	3	0.1851856389

[108,]	2	3	-0.0208349774
[109,]	1	3	0.0739873604
[110,]	2	3	0.2564943904
[111,]	1	3	-0.2273462697
[112,]	2	1	0.1960761973
[113,]	2	3	0.2645959837
[114,]	3	2	0.1257334020
[115,]	2	3	0.2384947891
[116,]	2	3	0.1374100646
[117,]	2	1	0.1710365823
[118,]	2	1	0.1232800898
[119,]	2	1	0.0855889253
[120,]	3	1	0.1714399239
[121,]	1	2	0.2622772674
[122,]	2	3	0.2200645277
[123,]	2	1	0.1953822977
[124,]	1	2	0.1006192402
[125,]	2	1	0.1042958562
[126,]	1	3	0.0666771492
[127,]	1	2	0.3072554173
[128,]	2	3	-0.0502005554
[129,]	2	1	0.2273973225
[130,]	2	3	0.2175522028
[131,]	2	1	-0.0059493407
[132,]	1	3	0.2693824855
[133,]	1	2	0.1556785637
[134,]	2	3	0.0260029117
[135,]	1	2	0.0645641832
[136,]	1	2	0.2201973636
[137,]	1	2	-0.0067481120
[138,]	1	3	-0.0332940744
[139,]	1	2	0.1520990703

[140,]	2	1	0.0489636242
[141,]	2	1	-0.0132142766
[142,]	1	3	-0.2087842761
[143,]	2	1	0.2644725620
[144,]	1	3	0.0397070535
[145,]	3	2	0.1613054752
[146,]	2	3	0.2112685633
[147,]	2	1	-0.0727427620
[148,]	3	2	0.0107724115
[149,]	1	2	0.2174347337
[150,]	2	1	0.0105958873
[151,]	1	3	0.1327782811
[152,]	2	1	0.1545198025
[153,]	1	3	-0.0516549574
[154,]	2	1	0.1138884107
[155,]	4	5	0.0296881839
[156,]	4	5	0.1157480070
[157,]	4	3	0.1321387628
[158,]	5	4	0.0529693860
[159,]	5	1	0.0776484216
[160,]	5	4	0.0206808951
[161,]	4	5	0.0970668076
[162,]	4	5	-0.2802617597
[163,]	5	3	0.1261539451
[164,]	5	2	0.1494028333
[165,]	5	1	0.1242009787
[166,]	5	1	0.2186289437
[167,]	5	1	0.1944857462
[168,]	5	2	0.2550809146
[169,]	5	4	0.1067119924
[170,]	4	5	-0.1750287844
[171,]	4	2	0.1615769756

[172,]	4	5	-0.0171170528
[173,]	4	3	0.1090364677
[174,]	5	4	-0.0016195044
[175,]	4	3	0.0021197230
[176,]	5	2	0.1560423040
[177,]	4	2	0.1567141599
[178,]	5	1	0.1305244962
[179,]	4	5	0.1529874722
[180,]	4	5	0.1418123438
[181,]	4	5	0.1901674234
[182,]	5	4	-0.1568819253
[183,]	5	4	-0.1310531679
[184,]	5	4	-0.2321297401
[185,]	5	1	0.1284892597
[186,]	5	1	0.1340304939
[187,]	1	4	-0.1174762401
[188,]	4	5	-0.1347107540
[189,]	5	4	0.0944095350
[190,]	4	3	0.0971719911
[191,]	4	5	0.0349324319
[192,]	5	3	0.1687386862
[193,]	4	5	-0.2381065243
[194,]	4	5	-0.0604343966
[195,]	5	1	0.1620870520
[196,]	5	2	0.0656087204
[197,]	5	4	-0.2206177216
[198,]	4	5	-0.0650379611
[199,]	4	5	0.1687061907
[200,]	4	2	0.0832968586
[201,]	5	4	-0.1384101564
[202,]	5	3	0.1057815260
[203,]	4	3	0.1370120358

[204,]	4	5	-0.0559059230
[205,]	5	4	0.1273774977
[206,]	5	2	0.1425172802
[207,]	5	1	0.2225751107
[208,]	4	5	0.0009916807
[209,]	4	5	0.0020182408
[210,]	5	3	0.2533488487
[211,]	4	5	0.0314256834
[212,]	5	2	0.1013774726
[213,]	5	1	0.0607634673
[214,]	2	5	-0.1250901817
[215,]	5	3	0.2039800515
[216,]	5	4	0.0170783725
[217,]	5	4	0.0719413307
[218,]	5	4	0.0552112229
[219,]	4	3	0.1918053496
[220,]	1	5	-0.1575213536
[221,]	5	1	0.1709393898
[222,]	5	1	0.2041127979
[223,]	5	1	0.1131454393
[224,]	5	3	0.1160061617
[225,]	5	1	0.1288113491
[226,]	5	2	0.1309411745
[227,]	5	2	0.1639196680
[228,]	5	1	0.1039967783
[229,]	4	3	-0.0291977359
[230,]	5	4	0.1613422001
[231,]	5	4	0.1005674581
[232,]	5	2	0.1888635943
[233,]	5	3	-0.1045970079
[234,]	5	2	0.1332608501
[235,]	4	5	0.1663061945

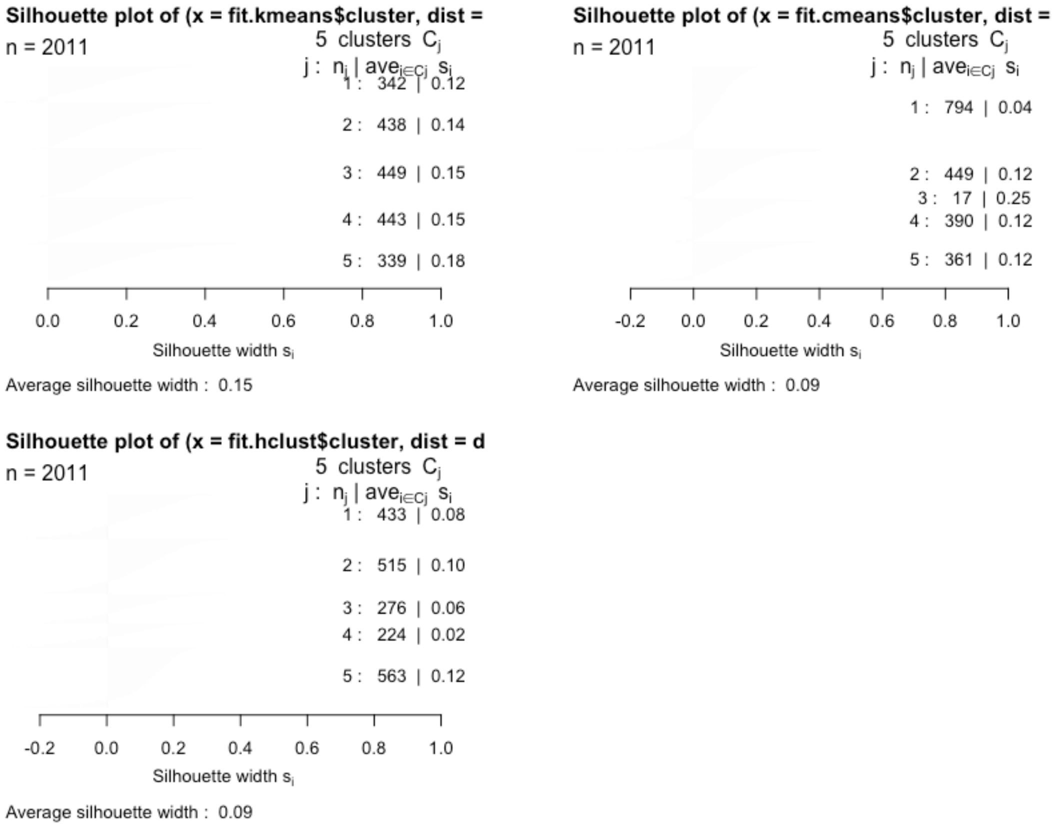
[236,]	5	3	-0.0757560863
[237,]	4	5	0.1096574436
[238,]	4	5	0.1306369468
[239,]	5	2	0.1524132226
[240,]	4	2	0.0184400598
[241,]	4	5	-0.2260214894
[242,]	5	4	-0.0606311656
[243,]	5	1	0.1365602191
[244,]	5	3	0.0385481579
[245,]	5	1	0.1433870655
[246,]	5	4	0.1123460474
[247,]	5	2	0.2075552403
[248,]	5	3	0.1645954095
[249,]	4	3	0.1233820174
[250,]	5	1	0.1926716299
[251,]	5	2	0.1160678048
[252,]	4	5	-0.0022547708
[253,]	5	1	0.1997382385
[254,]	5	1	0.0706140858
[255,]	5	2	0.1276569656
[256,]	5	2	0.1099494540
[257,]	2	1	0.0844556449
[258,]	2	3	-0.0520779958
[259,]	2	3	0.2824797256
[260,]	3	2	0.0251814384
[261,]	1	2	-0.0841106863
[262,]	1	2	0.1810469403
[263,]	3	2	0.0489319340
[264,]	1	3	0.0662790806
[265,]	1	2	-0.0124201480
[266,]	3	1	0.1494413690
[267,]	2	3	0.0553750815

[268,]	1	3	-0.2191232217
[269,]	2	1	0.3042345090
[270,]	1	2	0.2885795419
[271,]	2	1	0.1719383432
[272,]	1	2	0.0406401432
[273,]	1	2	0.2276799201
[274,]	1	2	-0.0215600894
[275,]	2	1	0.1703579012
[276,]	1	3	0.2178437094
[277,]	3	2	0.1558766090
[278,]	3	2	0.1073745971
[279,]	1	2	0.2582896141
[280,]	2	3	0.1929657144
[281,]	3	1	0.0781269241
[282,]	2	3	0.1010598893
[283,]	2	1	-0.1316808301
[284,]	2	1	-0.0225607763
[285,]	1	2	-0.0399575719
[286,]	1	2	0.0776323141
[287,]	2	1	0.0572007657
[288,]	2	3	0.0456760466
[289,]	1	3	0.2025620474
[290,]	1	2	0.1139724009
[291,]	1	2	0.0563852539
[292,]	3	1	-0.1419401508
[293,]	2	1	-0.0347602441
[294,]	2	1	0.0656321630
[295,]	3	1	0.1742154935
[296,]	1	3	0.0376707267
[297,]	3	2	0.2265316566
[298,]	3	2	0.0980389199
[299,]	2	1	0.1819498117

[300,]	3	2	-0.0773427087
[301,]	1	2	0.2586957822
[302,]	2	1	0.2889090798
[303,]	1	2	0.1338165281
[304,]	1	2	-0.0864805491
[305,]	2	3	0.2349853409
[306,]	2	3	0.1710793769
[307,]	2	1	0.1190019765
[308,]	3	1	0.0151321985
[309,]	3	2	-0.0588941011
[310,]	2	3	0.2247649730
[311,]	1	2	0.0647540958
[312,]	2	1	0.0220975967
[313,]	1	2	-0.1166346234
[314,]	1	3	-0.1438039737
[315,]	3	1	0.1940859787
[316,]	2	3	0.2598533437
[317,]	2	1	0.1785265808
[318,]	2	1	0.3299422623
[319,]	2	3	0.1243248425
[320,]	2	1	0.2688291533
[321,]	3	2	0.1759776787
[322,]	3	2	0.2479225273
[323,]	1	2	0.1720639309
[324,]	1	2	0.0932389927
[325,]	1	2	-0.1867724695
[326,]	2	3	0.2232767809
[327,]	3	1	0.0553978581
[328,]	3	2	0.1031224961
[329,]	3	2	0.0279776787
[330,]	2	1	0.2655880289
[331,]	2	3	0.1917995044

```
 [332,]       1        2  0.1390786351
 [333,]       3        2  0.1476567081
 [ reached getOption("max.print") -- omitted 1678 rows ]
attr(,"Ordered")
[1] FALSE
attr(,"call")
silhouette.default(x = fit.hclust$cluster, dist = dist(df)^2)
attr(,"class")
[1] "silhouette"
```

Tabela 1



A avaliação de modelos de agrupamentos não tem como levar em consideração uma informação de valores observados de uma variável alvo (ground truth).

Portanto, costuma-se usar métricas (tabela 1) que comparam os elementos dos grupos como os centroides (critério J) ou o quão similar os elementos de um grupo são entre si (coesão) em comparação aos outros clusters (separação), como é o caso da medida de silhueta (silhouette).

Resultados

> results

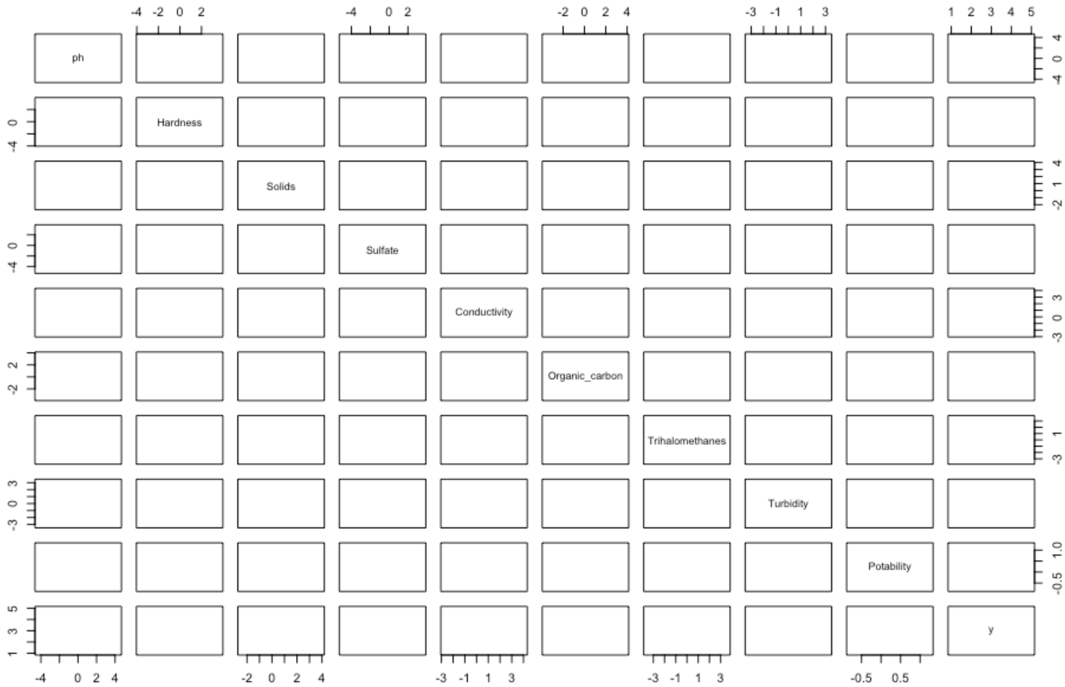
	Connectivity	Dunn	Silhouette
K-means (K = 5)	1228.178	0.07689555	0.14777435
Fuzzy K-means (K = 5)	1005.103	0.06866552	0.09128905
Hierarchical (K = 5)	1148.206	0.09785363	0.08706726

Conclusão: O método Fuzzy "vence" em 2 medidas (Connectivity e Dunn), dessa forma, utilizaremos o agrupamento fornecido por esse método.

Análise exploratória do agrupamento utilizando o método Fuzzy

```
> head(df)
         ph   Hardness       Solids Chloramines    Sulfate
1  0.7752344  0.5475678  0.0004932291   0.5919175  0.5579943
2  1.2616222 -0.4643582 -0.4601783194  -0.3636424 -0.5707833
3 -0.9387754 -0.2450192  0.7680379557   0.2669421 -0.1713654
4  1.9714164  1.5724639  0.7681552843   0.2470693  1.4459727
5  0.9753717  0.2126544 -0.9513523841  -1.6166335 -0.7355946
6  2.5713102  0.9386313  0.3957789739   1.2348821  1.6965528
  Conductivity Organic_carbon Trihalomethanes  Turbidity
1   -0.7787111      1.2549429       2.0986315  0.8482820
2   -0.3438864     -0.8242313      -2.1266326  0.1387643
3   -1.8031411     -1.7790047      -0.7096399 -1.8030621
4   -1.7637504     -0.1497130       1.1256417 -1.6579018
5    0.5988629     -0.5807314      -0.2224409  0.5569558
6    1.7034577      1.1011662       0.3449957  0.5174076
  Potability
1 -0.7996527
2 -0.7996527
3 -0.7996527
4 -0.7996527
5 -0.7996527
6 -0.7996527
```

Gráfico 11



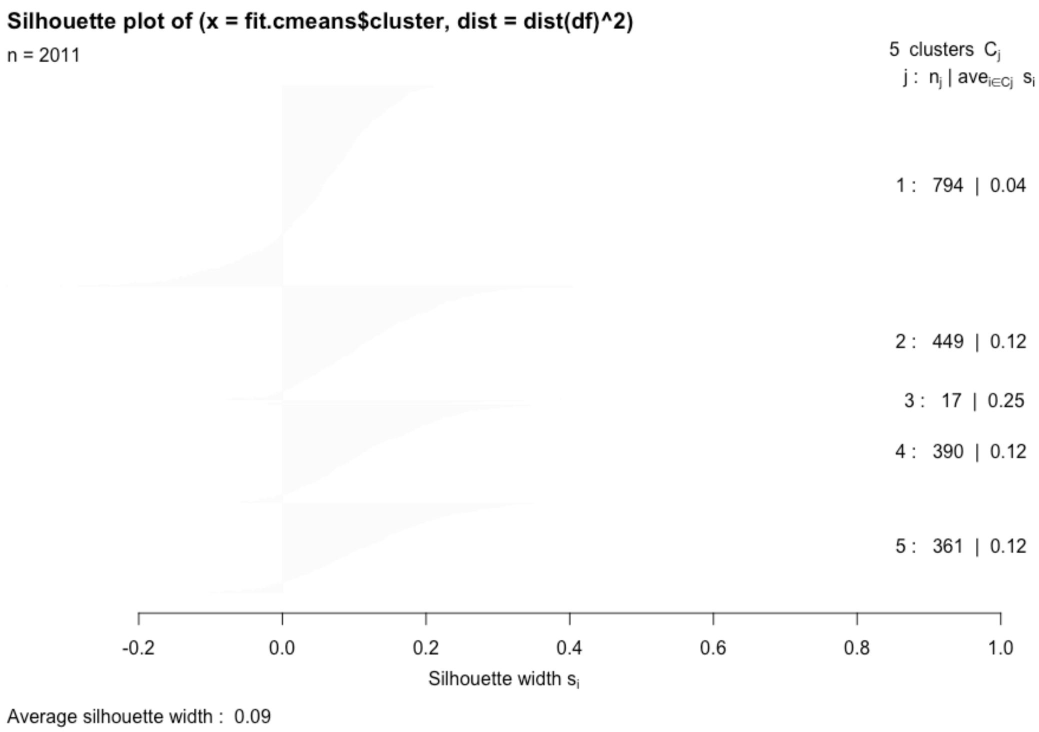
Desc(fit.cmeans$cluster, digits = 2, plotit = F)

```
------------------------------------------------------------------------
fit.cmeans$cluster (integer)

  length      n    NAs  unique    0s  mean  meanCI'
   2'011   2'011      0       5     0  2.54    2.47
          100.0%   0.0%          0.0%          2.61

     .05     .10    .25  median   .75   .90     .95
    1.00    1.00   1.00    2.00  4.00  5.00    5.00

   range     sd  vcoef     mad   IQR  skew    kurt
    4.00    1.58   0.62    1.48  3.00  0.45   -1.45

   level  freq   perc  cumfreq  cumperc
1      1   794  39.5%      794    39.5%
2      2   449  22.3%    1'243    61.8%
```

3	3	17	0.8%	1'260	62.7%
4	4	390	19.4%	1'650	82.0%
5	5	361	18.0%	2'011	100.0%

' 95%-CI (classic)

Gráfico 12



Referência

BRUCE, Peter; BRUCE, Andrew. **Estatística Prática para Cientistas de Dados**: 50 conceitos essenciais. Rio de Janeiro: Alta Books, 2019. 320 p. Luciana Ferraz

Anexo

Código R

```
library(tidyverse) # Gráficos, manipulação e transformação dos dados
library(cluster) # Avaliação dos grupos
library(clValid) # Avaliação dos grupos
library(e1071) # Fuzzy K-médias
library(factoextra) # Vizualização de grupos
library(skimr) # Análise exploratória de dados
library(gridExtra) # Ferramentas gráficas
library(ggforce) # Ferramentas gráficas
library(DescTools)
library(mice)

########################################
#                                      #
#----- Análise de Agrupamentos -----#
#                                      #
########################################

# Função que obtém o WSS para o método hierárquico
# NÃO MUDAR!
get_wss <- function(d, cluster){
  d <- stats::as.dist(d)
  cn <- max(cluster)
  clusterf <- as.factor(cluster)
  clusterl <- levels(clusterf)
```

```
  cnn <- length(clusterl)

  if (cn != cnn) {
    warning("cluster renumbered because maximum != number of clusters")
    for (i in 1:cnn) cluster[clusterf == clusterl[i]] <- i
    cn <- cnn
  }
  cwn <- cn
  # Compute total within sum of square
  dmat <- as.matrix(d)
  within.cluster.ss <- 0
  for (i in 1:cn) {
    cluster.size <- sum(cluster == i)
    di <- as.dist(dmat[cluster == i, cluster == i])
    within.cluster.ss <- within.cluster.ss + sum(di^2)/cluster.size
  }
  within.cluster.ss
}

# Função que contrói o "gráfico do cotovelo" e aponta o K ótimo
# NÃO MUDAR!
elbow.plot <- function(x, kmax = 15, alg = "kmeans") {
  # alg = c("kmeans", "cmeans", "hclust")
  wss <- c()
  if (alg == "kmeans") {
    for (i in 1:kmax) {
      set.seed(13)
      tmp <- kmeans(x, i)
      # wss[i] <- get_wss(dist(x), tmp$cluster)
      wss[i] <- tmp$tot.withinss
    }
    tmp <- data.frame(k = 1:kmax, wss)
```

```
  max_k <- max(tmp$k)
  max_k_wss <- tmp$wss[which.max(tmp$k)]
  max_wss <- max(tmp$wss)
  max_wss_k <- tmp$k[which.max(tmp$wss)]
  max_df <- data.frame(x = c(max_wss_k, max_k), y = c(max_wss, max_k_wss))
  tmp_lm <- lm(max_df$y ~ max_df$x)
  d <- c()
  for(i in 1:kmax) {
    d <- c(d, abs(coef(tmp_lm)[2]*i - tmp$wss[i] + coef(tmp_lm)[1]) /
           sqrt(coef(tmp_lm)[2]^2 + 1^2))
  }
  tmp$d <- d
  ggplot(data = tmp, aes(k, wss)) +
    geom_line() +
    geom_segment(aes(x = k[1], y = wss[1],
                  xend = max(k), yend = wss[which.max(k)]),
              linetype = "dashed") +
    geom_point(aes(size = (d == max(d)), color = (d == max(d))),
           show.legend = FALSE) +
    scale_size_manual(values = c(2,5)) +
    scale_color_manual(values = c("black", "red")) +
    labs(x = "Number of clusters",
       y = "Total within-cluster sum of squares",
       title = "Elbow plot for the K-means method") +
    theme_bw()
}
else if (alg == "cmeans") {
  for (i in 1:kmax) {
    if (i == 1) {
      wss[i] <- get_wss(dist(x), rep(1, nrow(x)))
    }
    else {
```

```
    set.seed(13)
    tmp <- cmeans(x, i)
    wss[i] <- get_wss(dist(x), tmp$cluster)
    # wss[i] <- tmp$sumsqrs$tot.within.ss
  }
}
tmp <- data.frame(k = 1:kmax, wss)
max_k <- max(tmp$k)
max_k_wss <- tmp$wss[which.max(tmp$k)]
max_wss <- max(tmp$wss)
max_wss_k <- tmp$k[which.max(tmp$wss)]
max_df <- data.frame(x = c(max_wss_k, max_k), y = c(max_wss, max_k_wss))
tmp_lm <- lm(max_df$y ~ max_df$x)
d <- c()
for(i in 1:kmax) {
  d <- c(d, abs(coef(tmp_lm)[2]*i - tmp$wss[i] + coef(tmp_lm)[1]) /
           sqrt(coef(tmp_lm)[2]^2 + 1^2))
}
tmp$d <- d
ggplot(data = tmp, aes(k, wss)) +
  geom_line() +
  geom_segment(aes(x = k[1], y = wss[1],
                   xend = max(k), yend = wss[which.max(k)]),
               linetype = "dashed") +
  geom_point(aes(size = (d == max(d)), color = (d == max(d))),
             show.legend = FALSE) +
  scale_size_manual(values = c(2,5)) +
  scale_color_manual(values = c("black", "red")) +
  labs(x = "Number of clusters",
       y = "Total within-cluster sum of squares",
       title = "Elbow plot for the fuzzy K-means method") +
  theme_bw()
```

```
}
else if (alg == "hclust") {
  for (i in 1:kmax) {
    set.seed(13)
    tmp <- hcut(x, i)
    wss[i] <- get_wss(dist(x), tmp$cluster)
  }
  tmp <- data.frame(k = 1:kmax, wss)
  max_k <- max(tmp$k)
  max_k_wss <- tmp$wss[which.max(tmp$k)]
  max_wss <- max(tmp$wss)
  max_wss_k <- tmp$k[which.max(tmp$wss)]
  max_df <- data.frame(x = c(max_wss_k, max_k), y = c(max_wss, max_k_wss))
  tmp_lm <- lm(max_df$y ~ max_df$x)
  d <- c()
  for(i in 1:kmax) {
    d <- c(d, abs(coef(tmp_lm)[2]*i - tmp$wss[i] + coef(tmp_lm)[1]) /
             sqrt(coef(tmp_lm)[2]^2 + 1^2))
  }
  tmp$d <- d
  ggplot(data = tmp, aes(k, wss)) +
    geom_line() +
    geom_segment(aes(x = k[1], y = wss[1],
                     xend = max(k), yend = wss[which.max(k)]),
                 linetype = "dashed") +
    geom_point(aes(size = (d == max(d)), color = (d == max(d))),
               show.legend = FALSE) +
    scale_size_manual(values = c(2,5)) +
    scale_color_manual(values = c("black", "red")) +
    labs(x = "Number of clusters",
         y = "Total within-cluster sum of squares",
         title = "Elbow plot for the hierarchical method") +
```

```
    theme_bw()
  }
}

# Função vizualização dos grupos
# NÃO MUDAR!
cluster_viz <- function(data, clusters,
                        axes = c(1, 2), geom = c("point", "text"), repel = TRUE,
                        show.clust.cent = TRUE, ellipse = TRUE, ellipse.type = "convex",
                        ellipse.level = 0.95, ellipse.alpha = 0.2, shape = NULL,
                        pointsize = 1.5, labelsize = 12, main = "Cluster plot",
                        ggtheme = theme_bw()) {
  require(factoextra)
  data <- scale(data)
  pca <- stats::prcomp(data, scale = FALSE, center = FALSE)
  ind <- facto_summarize(pca, element = "ind", result = "coord", axes = axes)
  eig <- get_eigenvalue(pca)[axes, 2]
  xlab <- paste0("Dim", axes[1], " (", round(eig[1], 1), "%)")
  ylab <- paste0("Dim", axes[2], " (", round(eig[2], 1), "%)")
  colnames(ind)[2:3] <- c("x", "y")
  label_coord <- ind
  lab <- NULL
  if ("text" %in% geom)
    lab <- "name"
  if (is.null(shape))
    shape <- "cluster"
  plot.data <- cbind.data.frame(ind, cluster = clusters, stringsAsFactors = TRUE)
  label_coord <- cbind.data.frame(label_coord, cluster = clusters, stringsAsFactors =
TRUE)
  p <- ggpubr::ggscatter(plot.data, "x", "y", color = "cluster",
                         shape = shape, size = pointsize, point = "point" %in% geom,
                         label = lab, font.label = labelsize, repel = repel,
```

```
                    mean.point = show.clust.cent, ellipse = ellipse, ellipse.type =
ellipse.type,
                    ellipse.alpha = ellipse.alpha, ellipse.level = ellipse.level,
                    main = main, xlab = xlab, ylab = ylab, ggtheme = ggtheme)
  p
}

# Mudar diretório de acordo com o endereço de onde o dataset foi salvo
# no seu compuador
# onde <- "C:\\Users\\marce\\Dropbox\\DE-UFPB\\disciplinas\\2021-1\\AM\\Aulas"
onde <- "/Users/user/Documents/ESTATÍSTICA/MULTIVARIADA 2/PROVA
4/PROVA_4"
setwd(onde)
# Mudar nome do arquivo de acordo com o dataset
df <- read.csv("water_potability.csv")
df

# Exploração inicial
head(df)
skim(df)
pairs(df)

# Padronização dos dados
boxplot(df)
df <- scale(df)

#Dados faltantes
PlotMiss(df, col = colorRampPalette(c( "gray10", "gray90"))(1))
#df <- complete(mice(df, printFlag=F))
df <- na.omit(df)
df <- as.data.frame(df)

# Encontrar número ótimo de grupos para o K-means
```

```
set.seed(13)
p1 <- elbow.plot(df)
p1
# K "ótimo" é igual a 5. Mudar de acordo com o seu resultado

# Encontrar número ótimo de grupos para o Fuzzy K-means
set.seed(13)
p2 <- elbow.plot(df, alg = "cmeans")
p2
# K "ótimo" é igual a 5. Mudar de acordo com o seu resultado

# Encontrar número ótimo de grupos para o método hierárquico
set.seed(13)
p3 <- elbow.plot(df, alg = "hclust")
p3
# K "ótimo" é igual a 5. Mudar de acordo com o seu resultado

grid.arrange(p1, p2, p3, nrow = 2) # Os trés gráficos juntos
nclust = c(5,5,5) # Número de grupos de acordo com os gráficos

# Agrupamento usando o K-means
set.seed(13)
fit.kmeans <- kmeans(df, nclust[1]) # Ajustar K de acordo com o gráfico
fit.kmeans
g1 <- cluster_viz(df, as.factor(fit.kmeans$cluster), geom = "point",
            main = "Cluster plot for the K-means method")

# Agrupamento usando o Fuzzy K-means
set.seed(13)
fit.cmeans <- cmeans(df, nclust[2]) # Ajustar K de acordo com o gráfico
fit.cmeans
g2 <- cluster_viz(df, as.factor(fit.cmeans$cluster), geom = "point",
```

```
                main = "Cluster plot for the Fuzzy K-means method")

# Agrupamento usando o método hierárquico
set.seed(13)
fit.hclust <- hcut(df, nclust[3]) # Ajustar K de acordo com o gráfico
fit.hclust
g3 <- cluster_viz(df, as.factor(fit.hclust$cluster), geom = "point",
                main = "Cluster plot for the Hierarchical method")

grid.arrange(g1, g2, g3, nrow = 2) # Os trés gráficos juntos

########################
# Medidas de qualidade #
########################

# Conectividade: varia de zero a infinito
# e deve ser minimizado
c1 <- connectivity(clusters = fit.kmeans$cluster, Data = df)
c2 <- connectivity(clusters = fit.cmeans$cluster, Data = df)
c3 <- connectivity(clusters = fit.hclust$cluster, Data = df)

# Dunn: varia de zero a infinito
# e deve ser maximizado
d1 <- dunn(clusters = fit.kmeans$cluster, Data = df)
d2 <- dunn(clusters = fit.cmeans$cluster, Data = df)
d3 <- dunn(clusters = fit.hclust$cluster, Data = df)

# Silhueta: varia de -1 a 1
# Quanto mais próximo de 1, melhor
s.kmeans <- silhouette(fit.kmeans$cluster, dist = dist(df)^2)
s.cmeans <- silhouette(fit.cmeans$cluster, dist = dist(df)^2)
s.hclust <- silhouette(fit.hclust$cluster, dist = dist(df)^2)
```

```
par(mfrow = c(2,2))
plot(s.kmeans)
plot(s.cmeans)
plot(s.hclust)
s1 <- summary(s.kmeans)$avg.width
s2 <- summary(s.cmeans)$avg.width
s3 <- summary(s.hclust)$avg.width

results <- matrix(c(c1, c2, c3, d1, d2, d3, s1, s2, s3), nrow = 3)
colnames(results) <- c("Connectivity", "Dunn", "Silhouette")
row.names(results) <- c(paste0("K-means (K = ",nclust[1],")"),
                    paste0("Fuzzy K-means (K = ",nclust[2],")"),
                    paste0("Hierarchical (K = ",nclust[3],")"))
results
# O método Fuzzy "vence" em 2 medidas (Connectivity e Dunn),
# dessa forma, utilizaremos o agrupamento fornecido por esse método
# MUDAR ESSA CONCLUSÃO DE ACORDO COM OS SEUS RESULTADOS!

head(df)
df$y <- factor(as.numeric(fit.cmeans$cluster))
pairs(df[,-4], col = df[,0])
#Fim
Desc(fit.cmeans, digits = 2, plotit = F)
```