

TD: Analyse des correspondances multiples et révisions sur l'AFC

Exercice 1

Le tableau suivant représente la couleur des cheveux, celle des yeux, et le sexe dans une population de 12 individus.

	Cheveux	Yeux	Sexe
1	Noir	Bleu	Homme
2	Noir	Bleu	Femme
3	Noir	Brun	Homme
4	Noir	Brun	Homme
5	Brun	Bleu	Femme
6	Brun	Brun	Homme
7	Brun	Brun	Homme
8	Roux	Bleu	Homme
9	Roux	Bleu	Homme
10	Roux	Bleu	Femme
11	Roux	Brun	Femme
12	Roux	Brun	Femme

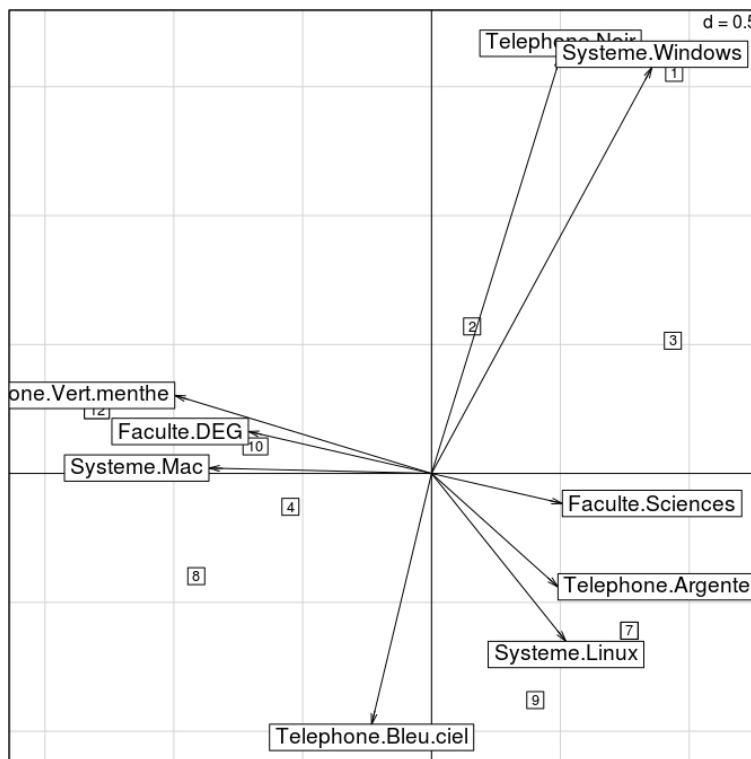
1. Déterminer les tableaux de contingence des variables croisées deux à deux.
2. Représenter les données sous forme de tableau disjonctif Z et calculer ses marges.
3. Calculer le tableau de Burt $B = Z^T Z$.

Exercice 2

Un fabricant de téléphones portables a effectué une étude de marché dans une population de 12 étudiants. Le tableau suivant représente la couleur du téléphone portable, le système d'exploitation de l'ordinateur de chaque étudiant, ainsi que la faculté où il est inscrit.

	Couleur	Système	Faculté
1	Noir	Windows	Sciences
2	Noir	Linux	DEG
3	Argenté	Windows	Sciences
4	Argenté	Mac	DEG
5	Argenté	Linux	Sciences
6	Argenté	Linux	Sciences
7	Argenté	Linux	Sciences
8	Bleu ciel	Mac	DEG
9	Bleu ciel	Linux	Sciences
10	Vert menthe	Mac	Sciences
11	Vert menthe	Mac	DEG
12	Vert menthe	Mac	DEG

1. Déterminer les tableaux de contingence des variables croisées deux à deux.
2. Représenter les données sous forme de tableau disjonctif Z et calculer ses marges.
3. Calculer le tableau de Burt B .
4. Au vu du résultat de l'ACM ci-dessous, quelle stratégie conseilleriez-vous au vendeur participant au forum en faculté DEG?



```

> acmin$TOT
      inertia      cum      ratio
1 0.72427724 0.7242772 0.3621386
2 0.43333543 1.1576127 0.5788063
3 0.35498491 1.5125976 0.7562988
4 0.26414381 1.7767414 0.8883707
5 0.19115518 1.9678966 0.9839483
6 0.03210343 2.0000000 1.0000000

> acmin$row.abs
      Axis1 Axis2 Axis3
1   1020   4643   267
2     28    626  2283
3   1009    510   969
4    345     32   174
5    678    715   258
6    678    715   258
7    678    715   258
8    959    305  2169
9    186   1485  1908
10   537     21   995
11  1941    116   230
12  1941    116   230

> acmin$row.rel
      Axis1 Axis2 Axis3 con.tra
1   2481   6761  -319   1488
2     95   1251 -3741   1083
3   3697   1117  1741    988
4  -2142   -119   528    583
5   5032  -3174   938    488
6   5032  -3174   938    488
7   5032  -3174   938    488
8  -3205   -610 -3554   1083
9     683 -3257 -3428    988
10 -2738     65  2486    710
11 -8724    313   507    806
12 -8724    313   507    806

> acmin$col.abs
      Comp1 Comp2 Comp3
Telephone.Argente      768   617  1607
Telephone.Bleu.ciel     69  1207  3825
Telephone.Noir          320  3338  1931
Telephone.Vert.menthe  1900   176  1199
Systeme.Linux           868  1352   352
Systeme.Mac            2380     1   152
Systeme.Windows        933  3165   103
Faculte.DEG            1611    85   485
Faculte.Sciences       1151    60   347

> acmin$col.rel
      Comp1 Comp2 Comp3 con.tra
Telephone.Argente      2859 -1374  2934    972
Telephone.Bleu.ciel    -180 -1882 -4888   1389
Telephone.Noir         833  5207 -2467   1389
Telephone.Vert.menthe -5503   304  1702   1250
Systeme.Linux          3234 -3013  -642    972
Systeme.Mac            -8867     3   278    972
Systeme.Windows        2434  4937   131   1389
Faculte.DEG            -6001   189  -886    972
Faculte.Sciences       6001  -189   886    694
    
```

Exercice 3

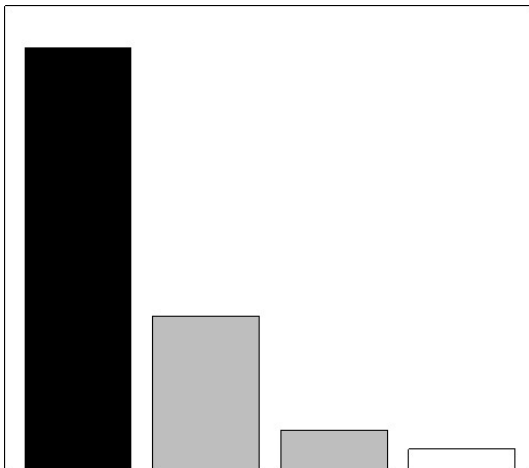
Le tableau donne la répartition d'une population par classe d'âge et loisir préféré:

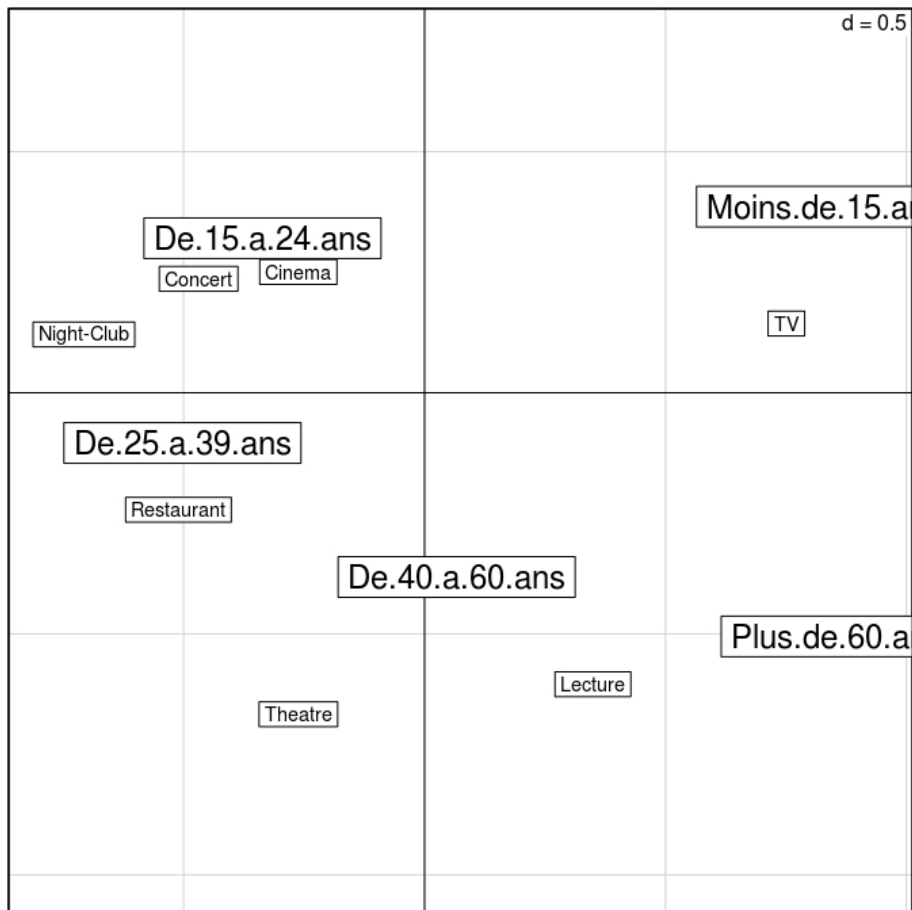
	Moins de 15 ans	15 à 24 ans	25 à 39 ans	40 à 60 ans	Plus de 60 ans
TV	322	114	72	135	130
Théâtre	1	17	85	92	14
Cinéma	90	220	192	87	7
Lecture	23	38	57	73	80
Restaurant	7	53	158	49	13
Night-Club	0	87	109	21	0
Concert	27	153	130	47	1

Voici les résultats d'une AFC. On demande de répondre aux questions suivantes:

1. Discuter les inerties.
2. Quels sont les modalités contribuant beaucoup au premier axe factoriel? Au second axe factoriel? Interpréter chaque axe en termes de type de loisir et de classe d'âge.
3. Quels sont les loisirs préférés de chaque classe d'âge? Comparer la projection sur le plan factoriel principal et le tableau de contingence.
4. Quelles sont les classes d'âge les plus concernées par chaque type de loisir?

```
> afc <- dudi.coa(tab, scan = FALSE)
> afcin <- inertia.dudi(afc, col.inertia=T, row.inertia=T)
> afcin$TOT
      inertia      cum      ratio
1 0.29004756 0.2900476 0.6641339
2 0.10542386 0.3954714 0.9055273
3 0.02729088 0.4227623 0.9680164
4 0.01396823 0.4367305 1.0000000
```





Contributions absolues et relatives:

```
> afcin$row.abs
      Axis1 Axis2
TV      5557   560
Theatre   183 3261
Cinema   521 1313
Lecture  422 3476
Restaurant 931  578
Night-Club 1383 111
Concert  1003  701
```

```
> afcin$col.abs
      Comp1 Comp2
Moins.de.15.ans 4115 2457
De.15.a.24.ans  983 2459
De.25.a.39.ans 2584  305
De.40.a.60.ans   29 2583
Plus.de.60.ans 2289 2195
```

```
> afcin$row.rel
      Axis1 Axis2 con.tra
TV      9563   350  3859
Theatre -1020 -6614  1190
Cinema  -5018  4592   690
Lecture  1970 -5903  1422
Restaurant -6424 -1450   963
Night-Club -9304   272   987
Concert  -7491  1903   890
```

```
> afcin$col.rel
      Comp1 Comp2 con.tra
Moins.de.15.ans 7991 1734  3420
De.15.a.24.ans -4521 4110  1444
De.25.a.39.ans -8882 -382  1932
De.40.a.60.ans   216 -7121  876
Plus.de.60.ans 6531 -2276 2328
```

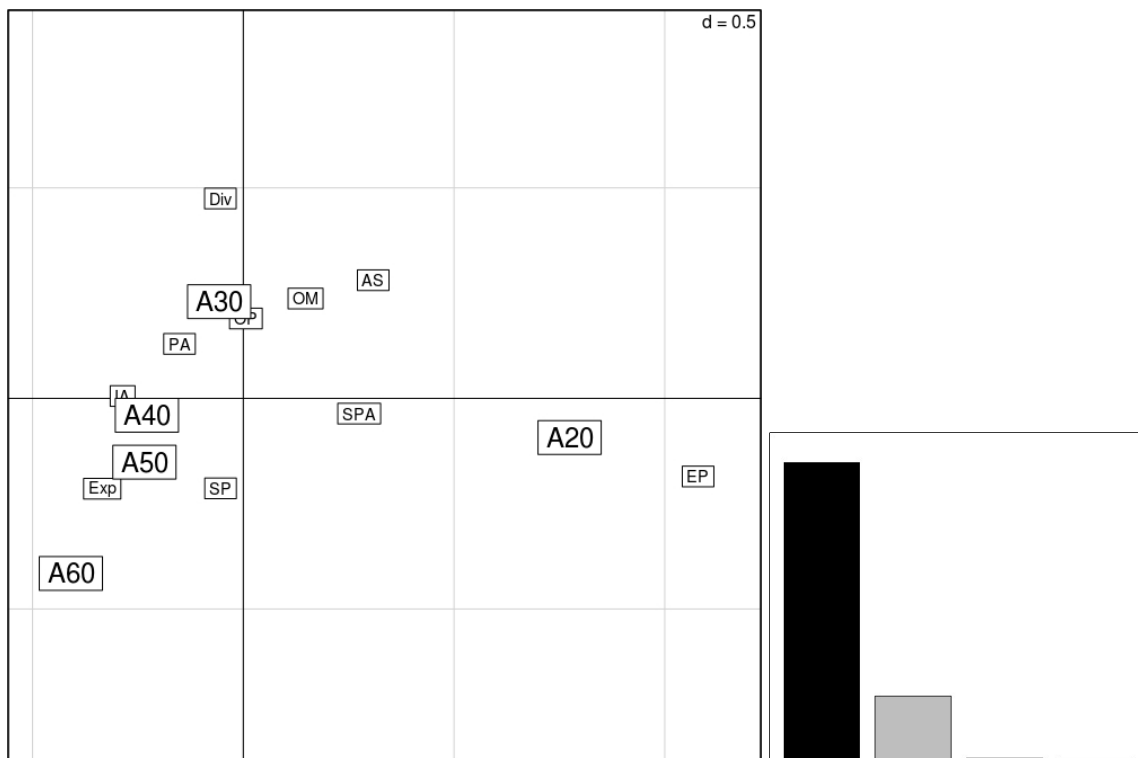
Exercice 4

On veut étudier le tableau croisé ci-dessous, issu d'une enquête de l'Association pour l'Emploi des Cadres de l'Agriculture, en 1975. Cette table donne la répartition par secteur et par classe d'âge des 76155 cadres de l'agriculture ayant répondu au questionnaire. Elle permet d'étudier la structure d'âge des différents secteurs.

SECTEURS/AGES	20–29	30–39	40–49	50–59	60–65
Secteur public (SP)	1480	1805	2603	1606	576
Secteur public autonome (SPA)	1409	1375	1251	650	164
Organ. professionnelles (OP)	2916	6875	4943	2099	489
Industries agricoles (IA)	1049	4870	4494	2380	1124
Exploitations (Exp)	1058	3385	4979	2341	1534
Enseignement privé (EP)	3864	1278	703	370	151
Para-agricole (PA)	700	2189	2153	700	256
Outre-mer (OM)	430	810	533	160	40
Divers (Div)	55	276	81	42	20
Agents spécialisé (AS)	1061	1676	816	309	27

Voici les résultats d'une AFC faite sur ces données.

```
> afcin$TOT
      inertia      cum      ratio
1 0.144238872 0.1442389 0.7961274
2 0.032225767 0.1764646 0.9739978
3 0.002759967 0.1792246 0.9892314
4 0.001951001 0.1811756 1.0000000
```



Contributions absolues:

```
> afcin$row.abs
  Axis1 Axis2
SP      21 1500
SPA    334   26
OP       1 2543
IA   1038    2
Exp  1356 2491
EP   6749   890
PA    125   410
OM     39   453
Div     1   436
AS    336 1249
```

```
> afcin$col.abs
  Comp1 Comp2
A20  7651   492
A30   73  5281
A40 1074   148
A50  534   996
A60  667  3084
```

Contributions relatives:

```
> afcin$row.rel
  Axis1 Axis2 con.tra
SP   -478 -7506   355
SPA  9322  -161   285
OP    11  9876   458
IA  -9600    4   861
Exp -6925 -2843 1559
EP  9695  -286 5542
PA -4461  3267   223
OM  2677  6903   117
Div  -95  7251   107
AS  5423  4499   494
```

```
> afcin$col.rel
  Comp1 Comp2 con.tra
A20  9858  -142  6179
A30  -574  9237  1017
A40 -9026  -277   947
A50 -6152 -2564   691
A60 -4557 -4706  1166
```

Interpréter ces résultats. On demande en particulier

1. de discuter les valeurs des inerties;
2. d'interpréter les axes.

Les questions typiques que l'on se pose à partir d'un tel tableau croisé sont:

1. la structure diffère-t-elle beaucoup d'un secteur à l'autre?
2. y-a-t-il des secteurs ayant des structures semblables?
3. quels sont les secteurs "jeunes", les secteurs "vieux"?
4. dans quels secteurs les jeunes sont-ils entrés durant les dix dernières années?
5. dans quels secteurs y aura-t-il beaucoup de départs à la retraite au cours des dix prochaines années?