

Rapport ACP etat

M1 Maths IMSD

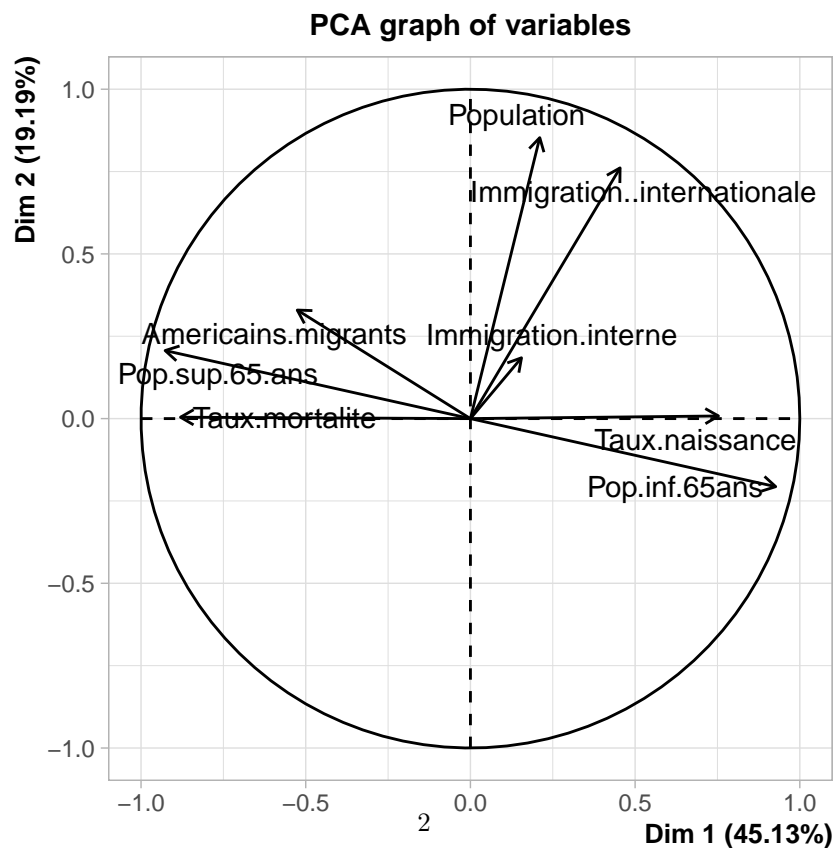
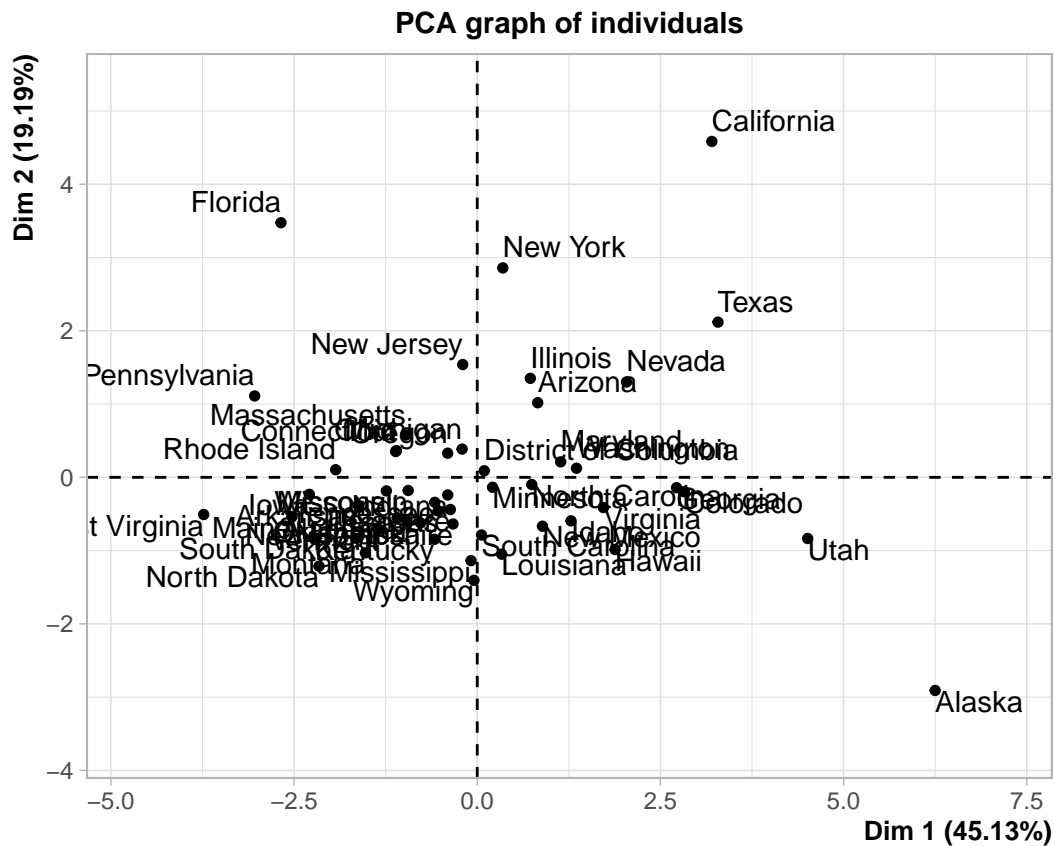
Ghilen TAGNIT HAMMOU

2024-05-08

Contents

Analyse et interprétation	2
ACP avec les variables actives	2
La qualité globale de représentation du nuage par le premier plan factoriel	3
Nombres d'axes principaux	3
La qualité de représentation des individus	4
La qualité de représentation des individus trié dans le premier plan factoriel	6
La qualité de représentation des individus graphique dans le premier plan factoriel	7
Les individus ayant la plus forte contribution à la formation du premier axe principal	7
La qualité de représentation des variables	8
Interprétation des premiers plans factoriels	10
Les caractéristiques de certains pays suivant leur position dans le plan factoriel	13

ACP avec les variables actives



Les valeurs propres

```
##          eigenvalue percentage of variance cumulative percentage of variance
## comp 1 3.610540e+00      4.513175e+01      45.13175
## comp 2 1.534894e+00      1.918617e+01      64.31793
## comp 3 1.089758e+00      1.362197e+01      77.93990
## comp 4 7.787879e-01      9.734848e+00      87.67474
## comp 5 4.528872e-01      5.661090e+00      93.33583
## comp 6 3.366277e-01      4.207846e+00      97.54368
## comp 7 1.965056e-01      2.456321e+00     100.00000
## comp 8 3.081460e-28      3.851825e-27     100.00000
```

La somme des valeurs propres

```
## [1] 8
```

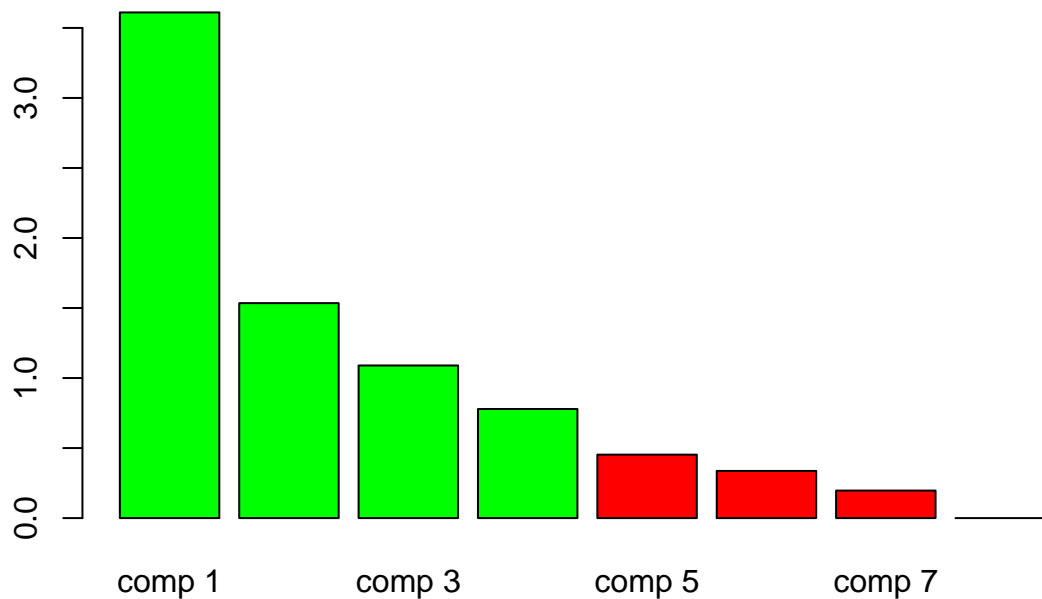
- On remarque que c'est bien égal au nombre de colonnes. la nullité de la dernière valeur propre s'explique du fait que la Pop inf à 65 ans est liée à la Pop sup à 65 ans.

La qualité globale de représentation du nuage par le premier plan factoriel

- Sur le premier plan factoriel on a 64.32% de l'information.

Nombres d'axes principaux

```
##          eigenvalue percentage of variance cumulative percentage of variance
## comp 1 3.610540e+00      4.513175e+01      45.13175
## comp 2 1.534894e+00      1.918617e+01      64.31793
## comp 3 1.089758e+00      1.362197e+01      77.93990
## comp 4 7.787879e-01      9.734848e+00      87.67474
## comp 5 4.528872e-01      5.661090e+00      93.33583
## comp 6 3.366277e-01      4.207846e+00      97.54368
## comp 7 1.965056e-01      2.456321e+00     100.00000
## comp 8 3.081460e-28      3.851825e-27     100.00000
```



- Avec la règle de Kaiser on peut se limiter à 3 dimensions compte tenu des 3 premières valeurs propres qui sont supérieures à 1. On obtient aussi plus de 87.67% de l'information avec les 4 premiers axes. Donc on choisit raisonnablement 4 dimensions.

La qualité de représentation des individus

La matrice Ψ des coordonnées des individus

##	Dim.1	Dim.2	Dim.3	Dim.4
## Alabama	-1.09225808	-0.55137507	-0.528624629	0.565975126
## Alaska	6.24968765	-2.90866872	-0.162754337	-1.015825283
## Arizona	0.82530920	1.01781349	1.694269759	-0.055178230
## Arkansas	-1.62537430	-0.37984009	-0.166027068	0.408541065
## California	3.20200720	4.58388132	-1.386157209	0.110078769
## Colorado	2.82550060	-0.19665091	1.438917992	0.413111688
## Connecticut	-1.11364363	0.35202148	0.005487548	-0.298949029
## Delaware	-0.36721557	-0.43903290	0.841082463	-0.064498965
## District of Columbia	0.09567505	0.08979763	-1.121144202	-0.766536404
## Florida	-2.67842619	3.47491636	1.302958048	-1.664099964
## Georgia	2.71989446	-0.14164840	0.472963744	0.636059999
## Hawaii	1.88160672	-0.98553059	-0.485885571	-4.351310099
## Idaho	1.28029703	-0.59405738	0.983158773	0.643014032
## Illinois	0.72508586	1.35113225	-1.358909958	0.282809871
## Indiana	-0.40233706	-0.24098398	-0.389350808	0.968601936
## Iowa	-2.29161242	-0.23223585	-0.585149278	0.105737803
## Kansas	-0.53174051	-0.45595473	-0.949172140	-0.237781759
## Kentucky	-0.58121564	-0.84523241	-0.129162642	0.044965080

## Louisiana	0.33312390	-1.04913989	-1.470747113	0.754592693
## Maine	-2.52913191	-0.53914589	1.468746305	-0.159785781
## Maryland	1.13867371	0.21131287	0.285021126	-0.145814668
## Massachusetts	-0.97574037	0.57673008	-0.267674784	0.023221201
## Michigan	-0.20677920	0.38581442	-0.513095859	0.871659323
## Minnesota	0.20919291	-0.13575451	0.339994646	0.837436392
## Mississippi	-0.08474246	-1.13726725	-0.939173233	0.532505193
## Missouri	-1.23993389	-0.18558743	-0.130291822	0.330831828
## Montana	-1.52692287	-1.03047750	0.202486736	0.060329302
## Nebraska	-0.77461793	-0.62236235	-0.815371153	-0.153559875
## Nevada	2.03701182	1.30100407	3.966412030	-0.005856354
## New Hampshire	-0.33058638	-0.63675960	2.133452301	0.825022194
## New Jersey	-0.19781333	1.53908372	-0.342501268	-0.238405537
## New Mexico	0.89009454	-0.66706970	-0.595938174	0.058301158
## New York	0.34937280	2.85809670	-1.788866722	-0.198872519
## North Carolina	0.74456677	-0.09874819	0.280503288	-0.534786997
## North Dakota	-2.15897960	-1.21226810	-1.447943788	-1.174439111
## Ohio	-1.10322408	0.36460173	-0.884762510	0.804029219
## Oklahoma	-0.96109997	-0.59370738	-0.611747788	-0.288958309
## Oregon	-0.40395147	0.32931642	0.974896372	0.413375360
## Pennsylvania	-3.03811159	1.11074819	-0.446550418	-0.095004119
## Rhode Island	-1.93140581	0.10360963	0.936287183	-0.430494439
## South Carolina	0.05842768	-0.78718372	0.362671688	-0.216480503
## South Dakota	-1.64922318	-0.83256956	-0.358354199	-0.068133114
## Tennessee	-0.57932519	-0.34156819	0.008524383	0.666936553
## Texas	3.28547051	2.11807539	-0.565252600	0.632424692
## Utah	4.50979967	-0.83523017	-1.295354039	1.784539872
## Vermont	-1.34086102	-0.70941731	1.197384515	0.467491699
## Virginia	1.72196002	-0.41335666	0.159609241	-1.685676324
## Washington	1.35447861	0.12390925	0.664519809	-0.280064109
## West Virginia	-3.73381527	-0.50778790	-0.171064747	0.050280710
## Wisconsin	-0.94221404	-0.17942790	0.263568577	0.654334044
## Wyoming	-0.04493374	-1.40582480	-0.075888470	0.184304690

La qualité de représentation des individus trié dans le sous-espace engendré par 4 composantes principales

##	Iowa	West Virginia	Indiana
##	0.9999883	0.9992809	0.9987422
##	Ohio	Pennsylvania	Texas
##	0.9985601	0.9973437	0.9971995
##	Illinois	Rhode Island	Maryland
##	0.9960613	0.9931686	0.9924138
##	New Mexico	Virginia	Hawaii
##	0.9889290	0.9857091	0.9845705
##	Oregon	Colorado	Nebraska
##	0.9818842	0.9813567	0.9747744
##	Nevada	Idaho	Georgia
##	0.9736629	0.9718674	0.9717572
##	California	Alaska	Kentucky
##	0.9702395	0.9678130	0.9646359
##	Florida	Missouri	South Dakota
##	0.9616824	0.9607768	0.9606516
##	North Dakota	Maine	Louisiana

##	0.9488643	0.9466471	0.9365716
##	New York	Utah	Tennessee
##	0.9333430	0.9314635	0.9175757
##	Connecticut	District of Columbia	New Jersey
##	0.9058968	0.9058544	0.9001064
##	Kansas	Oklahoma	Michigan
##	0.8955213	0.8895390	0.8825639
##	Montana	Wisconsin	Massachusetts
##	0.8820128	0.8670971	0.8629411
##	Washington	New Hampshire	South Carolina
##	0.8480718	0.8421542	0.8410299
##	Alabama	Delaware	Arkansas
##	0.8142017	0.7799037	0.7388907
##	North Carolina	Minnesota	Arizona
##	0.7131414	0.7079686	0.6839181
##	Vermont	Wyoming	Mississippi
##	0.6360112	0.6277004	0.6218791

- Les 5 individus les mieux représentés sont Iowa, West Virginia, Indiana, Ochio et Pennsylvania avec des valeurs respectives (0.9999883 ; 0.9992809 ; 0.9987422 ; 0.9985601 ; 0.9973437).
- Les 5 individus les moins bien représentés sont Mississippi, Wyoming, Vermont, Arizona et Minnesota avec des valeurs respectives (0.6218791 ; 0.6277004 ; 0.6360112 ; 0.6839181 ; 0.7079686).
- Ici on peut pas parler d'individus très mal représentés mais moins bien car la plus faible valeur vaut 0.6218791.

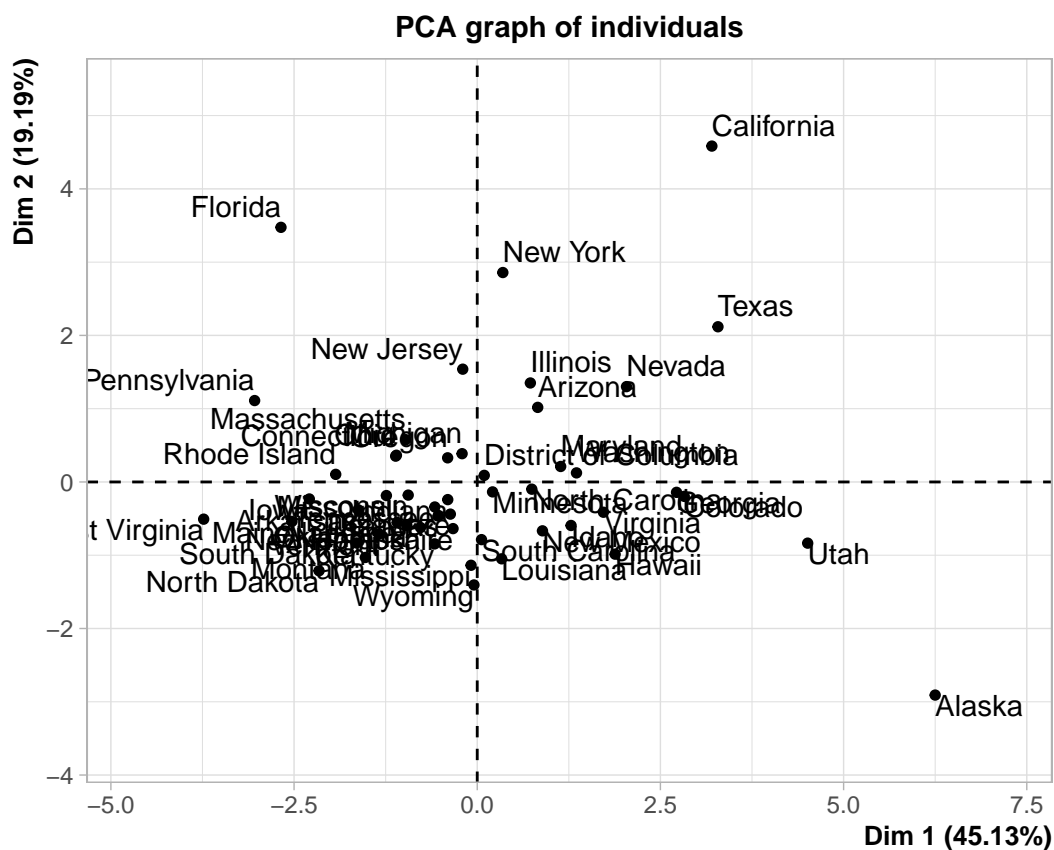
La qualité de représentation des individus trié dans le premier plan factoriel

##	West Virginia	Pennsylvania	Texas
##	0.997048531	0.977865341	0.952358566
##	Kentucky	Alaska	Iowa
##	0.947787613	0.946726563	0.937507998
##	South Dakota	Maryland	California
##	0.924605880	0.921956482	0.913730510
##	Georgia	Missouri	Montana
##	0.895879607	0.889254346	0.870560375
##	Connecticut	New Jersey	Massachusetts
##	0.850178524	0.839399522	0.817030632
##	Florida	Rhode Island	Colorado
##	0.780548230	0.773576745	0.767062064
##	New Mexico	Utah	Maine
##	0.766737167	0.756577772	0.713692553
##	Arkansas	New York	Washington
##	0.690681274	0.671109284	0.661988009
##	Oklahoma	South Carolina	Wyoming
##	0.654714191	0.653829166	0.615343788
##	North Dakota	Alabama	Nebraska
##	0.605548634	0.581307781	0.574336565
##	Idaho	Wisconsin	Illinois
##	0.574124064	0.562715598	0.547472560
##	Virginia	Ohio	Tennessee
##	0.514942363	0.485047052	0.462576882
##	North Carolina	Vermont	Mississippi
##	0.433140308	0.370201618	0.327954329
##	Kansas	Louisiana	Nevada

##	0.303416740	0.287718606	0.263652461
##	Arizona	Delaware	Oregon
##	0.255810294	0.245862104	0.191471013
##	Hawaii	Indiana	Michigan
##	0.187575271	0.167722643	0.139221971
##	New Hampshire	Minnesota	District of Columbia
##	0.075430084	0.050085201	0.008377232

- Les 5 individus les mieux représentés sont West Virginia, Pennsylvania, Texas, Kentucky et Alaska avec des valeurs respectives (0.997048531 ; 0.977865341 ; 0.952358566 ; 0.947787613 ; 0.946726563).
- Les 5 individus les moins bien représentés sont District of columbia, Minnesota, New Hampshire, Michigan et Indiana avec des valeurs respectives (0.008377232 ; 0.050085201 ; 0.075430084 ; 0.139221971 ; 0.167722643).

La qualité de représentation des individus graphique dans le premier plan factoriel



- On voit bien que l'Alaska et la Pennsylvanie sont bien représentées.
- On voit bien que District of columbia et Minnesota sont mal représentées, elles sont près du centre.

Les individus ayant la plus forte contribution à la formation du premier axe principal

##	Alaska	Utah	West Virginia
##	21.211640340	11.045163025	7.571174994
##	Texas	California	Pennsylvania
##	5.862093971	5.568038617	5.012623334
##	Colorado	Georgia	Florida

##	4.335592326	4.017554113	3.895982482
##	Maine	Iowa	North Dakota
##	3.473766265	2.851937237	2.531364625
##	Nevada	Rhode Island	Hawaii
##	2.253433394	2.025837930	1.922716878
##	Virginia	South Dakota	Arkansas
##	1.610288750	1.477122426	1.434710926
##	Montana	Washington	Vermont
##	1.266169452	0.996327068	0.976394137
##	Idaho	Missouri	Maryland
##	0.890182610	0.834938890	0.704135456
##	Connecticut	Ohio	Alabama
##	0.673519391	0.660975090	0.647900266
##	Massachusetts	Oklahoma	Wisconsin
##	0.517042424	0.501642997	0.482121799
##	New Mexico	Arizona	Nebraska
##	0.430258937	0.369905685	0.325861249
##	North Carolina	Illinois	Kentucky
##	0.301068227	0.285519976	0.183456127
##	Tennessee	Kansas	Oregon
##	0.182264656	0.153552583	0.088616790
##	Indiana	Delaware	New York
##	0.087909884	0.073231815	0.066288136
##	Louisiana	New Hampshire	Minnesota
##	0.060265563	0.059350932	0.023765752
##	Michigan	New Jersey	District of Columbia
##	0.023220488	0.021250479	0.004971129
##	Mississippi	South Carolina	Wyoming
##	0.003899956	0.001853937	0.001096485

- On remarque que les individus Alaska, Utah et West Virginia ont une forte contribution à la formation du premier axe principal suivie de manière moins forte du Texas, California, et Pennsylvania.

La qualité de représentation des variables

La matrice Φ des coordonnées des variables

##	Dim.1	Dim.2	Dim.3	Dim.4
## Population	0.2104598	0.852902089	-0.278780255	0.06110702
## Immigration.interne	0.1550099	0.184622984	0.935512192	0.01984345
## Americains.migrants	-0.5256081	0.329546513	0.078511285	0.75563997
## Immigration..internationale	0.4543315	0.761116451	0.055043153	-0.27408314
## Taux.naissance	0.7529892	0.007929195	-0.305925782	0.12873385
## Taux.mortalite	-0.8801867	0.004193804	-0.184577741	0.03158157
## Pop.inf.65ans	0.9267730	-0.206623468	-0.001180088	0.23555988
## Pop.sup.65.ans	-0.9267730	0.206623468	0.001180088	-0.23555988

Sur le premier axe

##	Dim.1	Dim.2	Dim.3	Dim.4
## Population	0.04429331	7.274420e-01	7.771843e-02	0.0037340681
## Immigration.interne	0.02402806	3.408565e-02	8.751831e-01	0.0003937625
## Americains.migrants	0.27626388	1.086009e-01	6.164022e-03	0.5709917628
## Immigration..internationale	0.20641714	5.792983e-01	3.029749e-03	0.0751215660
## Taux.naissance	0.56699270	6.287213e-05	9.359058e-02	0.0165724049


```

## Taux.mortalite      0.77472870 1.758799e-05 3.406894e-02 0.0009973955
## Pop.inf.65ans       0.85890824 4.269326e-02 1.392607e-06 0.0554884562
## Pop.sup.65.ans     0.85890824 4.269326e-02 1.392607e-06 0.0554884562
##                    Dim.5
## Population          1.239005e-02
## Immigration.interne  3.875420e-02
## Americains.migrants 3.428350e-04
## Immigration..internationale 6.928901e-05
## Taux.naissance      3.141547e-01
## Taux.mortalite      5.435783e-02
## Pop.inf.65ans       1.640918e-02
## Pop.sup.65.ans      1.640918e-02

```

- Les variables Pop sup 65 ans, Pop inf 65 ans et Taux de mortalité sont bien représentées (entre 0.77 et 0.86).
- La variable Taux de naissance est moyennement bien représentée (0.57).
- Les variables Population, Immigration interne, Americains migrants et Immigration internationale sont très mal représentées (entre 0.04 et 0.20).

Dans le premier plan factoriel

```

##              Population      Immigration.interne
##              0.77173529      0.05811371
##      Americains.migrants Immigration..internationale
##              0.38486478      0.78571539
##              Taux.naissance      Taux.mortalite
##              0.56705558      0.77474629
##              Pop.inf.65ans      Pop.sup.65.ans
##              0.90160150      0.90160150

```

- Les variables Pop sup 65 ans, Pop inf 65 ans, Immigration internationale, Taux de mortalité, et Population sont bien représentées (entre 0.77 et 0.90).
- La variable Taux de naissance est moyennement bien représentée (0.57).
- Les variables Immigration interne et Americains migrants sont très mal représentées (entre 0.06 et 0.38).

Dans le sous-espace à 4 dimensions

```

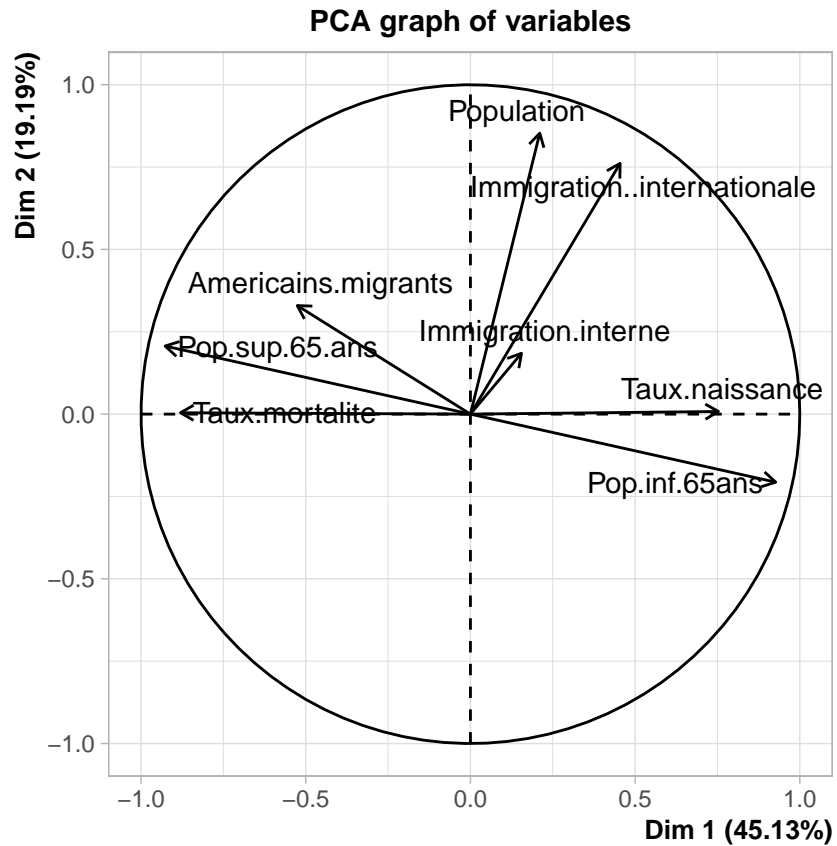
##              Population      Immigration.interne
##              0.8531878      0.9336905
##      Americains.migrants Immigration..internationale
##              0.9620206      0.8638667
##              Taux.naissance      Taux.mortalite
##              0.6772186      0.8098126
##              Pop.inf.65ans      Pop.sup.65.ans
##              0.9570913      0.9570913

```

- La qualité de représentation est supérieure à 0.93 pour Immigration interne, Pop sup 65 ans, Pop inf 65 ans et Americains migrants.
- La qualité de représentation est entre 0.81 et 0.86 pour Taux de mortalité, Population et Immigration internationale.
- La qualité de représentation de Taux de naissance est à 0.68.

Interprétation des premiers plans factoriels

Plan (1,2)



* La dim 1 apporte 45% de l'information.

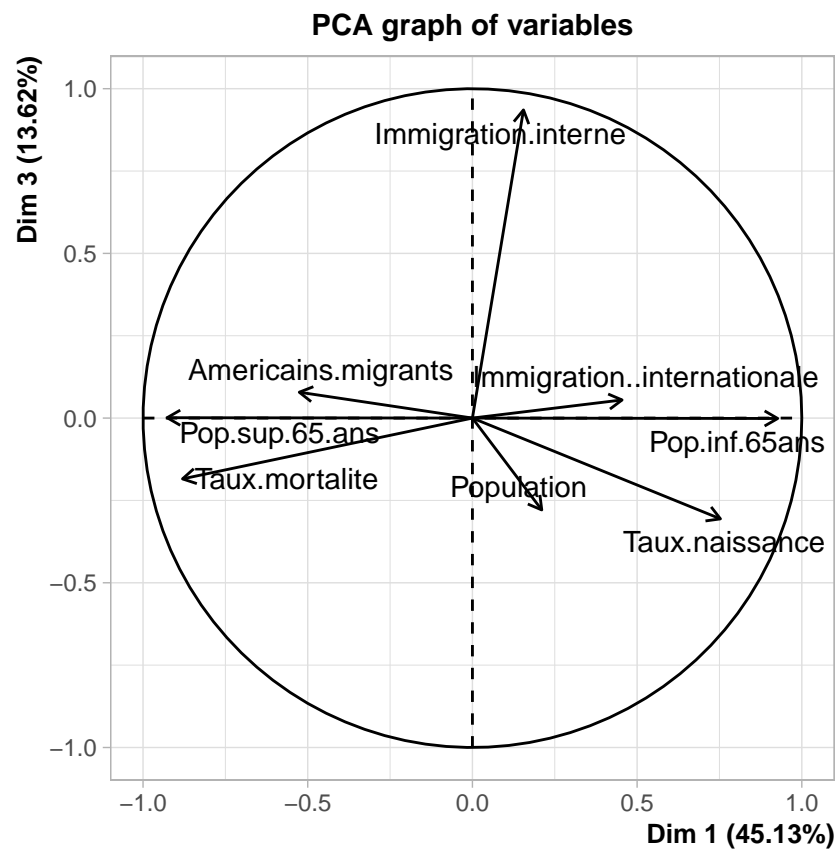
+ Elle est corrélée positivement avec Taux de naissance et Pop inf à 65 ans.

+ Elle est corrélée négativement avec Taux de mortalité et Pop sup à 65 ans.

* La dim 2 apporte 19% de l'information.

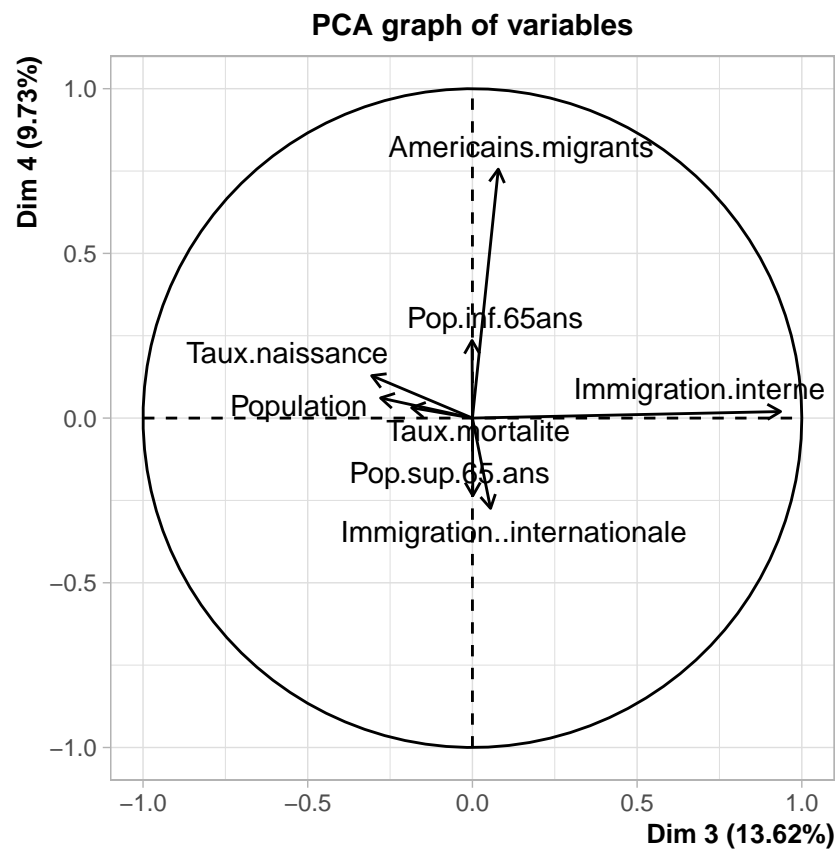
+ Elle est corrélée positivement avec Immigration internationale et Population.

Plan (1,3)



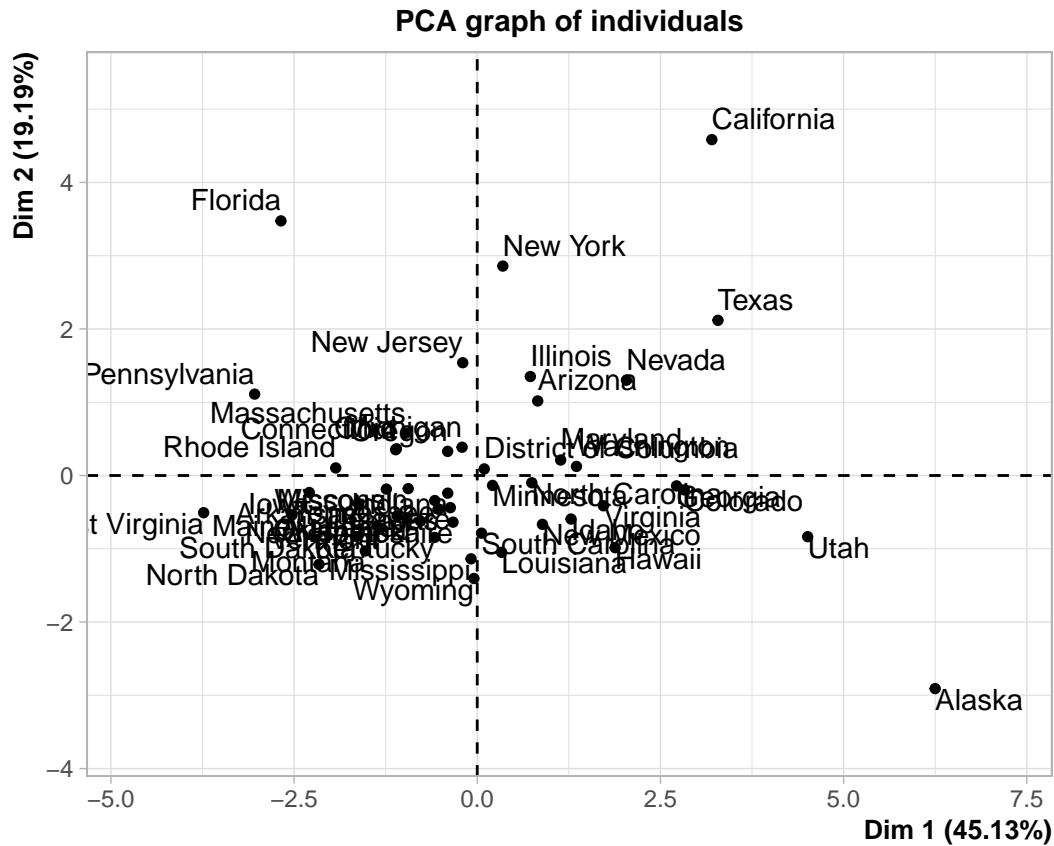
- La dim 3 apporte 14% de l'information.
- + Elle est corrélée positivement avec Immigration interne.

Plan (3,4)



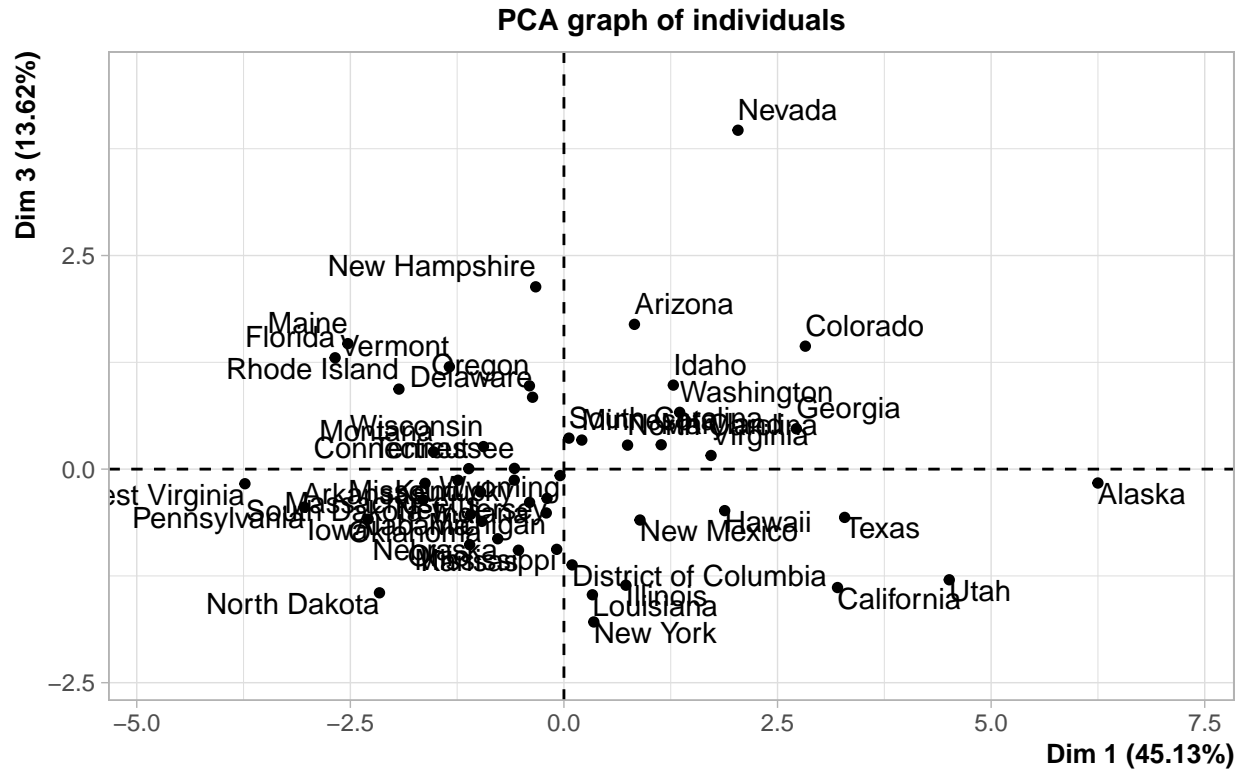
- La dim 4 apporte 10% de l'information.
- + Elle est corrélée positivement avec Americains migrants.

Les caractéristiques de certains pays suivant leur position dans le plan factoriel Plan (1,2)



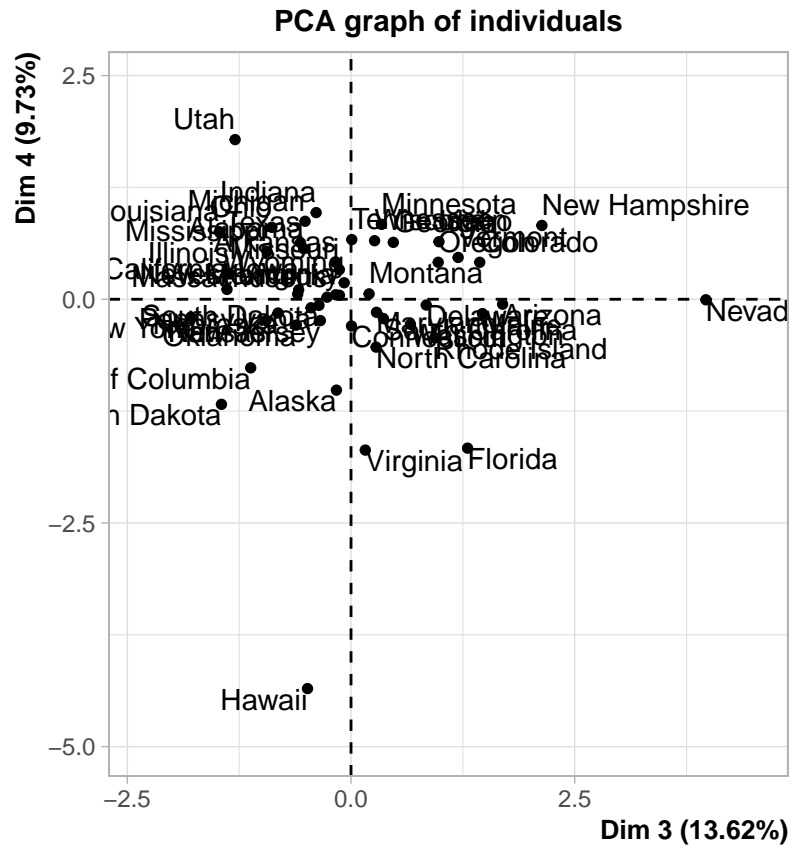
- Les individus placés à droite ont un Taux de naissance élevé et une Pop inf à 65 ans (Alaska, California, Texas et Utah).
- Les individus placés à gauche ont un Taux de mortalité élevé et une Pop sup à 65 ans (West Virginia, pennsylvania, Florida et South Dakota).
- Les individus placés en haut ont une forte immigration internaionale et Population (California, Florida, New York, Texas et New Jersey).
- Les individus placés en bas ont une faible immigration internaionale et Population (Alaska et Wyoming).

Plan (1,3)



- Les individus placés en haut ont une forte immigration interne (Nevada, New Hampshire, Arizona et Florida).
- Les individus placés en bas ont une faible immigration interne (New York et North Dakota).

Plan (3,4)



- Les individus placés en haut ont reçu une forte migration américaine (Utah).
- Les individus placés en bas ont reçu une faible migration américaine (Hawaii, Virginia, Florida et Alaska).