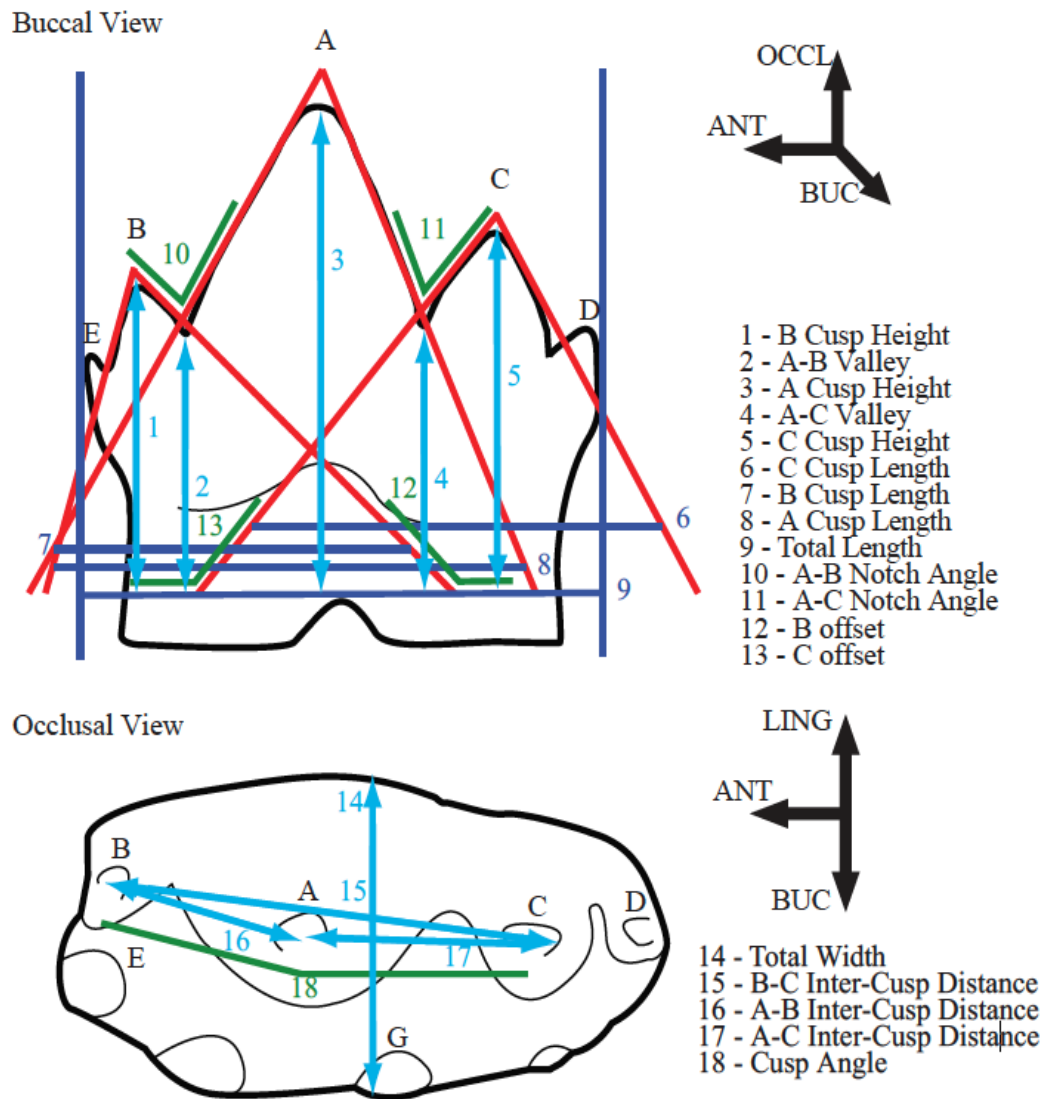


Sujet d'examen - Janvier 2020 - Durée 1h30

1 ACP

Les deux schémas ci-dessous détaillent les 18 mesures (i.e. variables) définissant la morphologie d'une molaire des premiers mammifères vivant entre -200 millions et -150 millions d'années :



Ces 18 mesures sont observées sur 22 animaux (i.e. individus) dont les restes ont été mis à jour. On dispose d'une variable supplémentaire qualitative fournissant l'espèce d'appartenance pour chaque mammifère et dont les modalités sont *Kuehneotherium* d'une part et *Morganucodon* d'autre part.

Questions. À partir des résultats de l'ACP fournis en Annexe 1, répondez aux questions suivantes en justifiant vos réponses :

1. Combien de dimensions retiendriez-vous pour analyser ces données ? Quel pourcentage d'inertie est associé au plan factoriel (1, 2) ?
2. Quelle est la variable la mieux représentée sur l'axe 1 ? Quel est l'individu le moins bien représenté sur l'axe 2 ?
3. Quelle est la corrélation entre la variable **HeightB** et l'axe factoriel 1 ? Quelle est la variable la plus corrélée avec l'axe factoriel 2 en valeur absolue ?
4. Un commentaire indique que les molaires des 11 individus représentés à droite sur l'axe 1 possèdent en moyenne un angle AB (numéroté 10 sur le schéma présentant les 18 mesures) plus petit que celui des 11 individus représentés à gauche de l'axe 1. Êtes-vous d'accord avec cette affirmation ?
5. Préciser la méthode permettant de représenter les deux modalités *Kuehneotherium* et *Morganucodon*.
6. Que représente la figure 2 ?
7. Quelle est la variable qui contribue le plus à la construction de l'axe 1 ? Quelle est la variable qui contribue le plus à la construction de l'axe 2 ?
8. Quel axe sépare le mieux les deux espèces de mammifères ? Quelle propriété doivent vérifier les variables qui distinguent le mieux les deux espèces de mammifères ?
9. Écrire un bref commentaire résumant les principales différences entre ces deux espèces de mammifères.
10. Quel axe sépare le mieux les deux mammifères 4 et 12 ? Écrire un bref commentaire résumant les principales différences entre ces deux mammifères.

2 AFC et AFCM

On s'intéresse à l'opinion public des européens vis à vis du **revenu universel**. Selon la définition sur Wikipédia, le **revenu universel** (encore appelé revenu de base) est une somme d'argent versée par une communauté politique à tous ses membres, sur une base individuelle, sans conditions de ressources ni obligation ou absence de

travail. Les données auxquelles on s'intéresse ont été extraites d'une enquête menée en 2016 ; elles concernent 9649 européens pour lesquels on a observé les variables suivantes :

- **country** = pays d'appartenance ; les modalités correspondent aux abbréviations officielles suivantes :

Union européenne (UE)

Belgique	(BE)	Grèce	(EL)	Lituanie	(LT)	Portugal	(PT)
Bulgarie	(BG)	Espagne	(ES)	Luxembourg	(LU)	Roumanie	(RO)
Tchéquie	(CZ)	France	(FR)	Hongrie	(HU)	Slovénie	(SI)
Danemark	(DK)	Croatie	(HR)	Malte	(MT)	Slovaquie	(SK)
Allemagne	(DE)	Italie	(IT)	Pays-Bas	(NL)	Finlande	(FI)
Estonie	(EE)	Chypre	(CY)	Autriche	(AT)	Suède	(SE)
Irlande	(IE)	Lettonie	(LV)	Pologne	(PL)	Royaume-Uni	(UK)

- **gender** = sexe (2 modalités : **female** / **male**),
- **rural_or_urban** = environnement du lieu d'habitation (2 modalités : **rural** / **urban**),
- **education** = niveau d'étude (4 modalités : **educ_high** / **educ_medium** / **educ_low** / **educ_no**),
- **job** = l'enquêté est-il en activité professionnelle à plein temps ? (2 modalités : **job_no** / **job_yes**),
- **children** = l'enquêté a-t-il à sa charge un ou plusieurs enfants ? (2 modalités : **child_no** / **child_yes**),
- **awareness** = niveau de connaissance du concept de revenu universel (4 modalités : **know_nothing** / **know_little** / **know_some** / **know_well**),
- **vote** = vote vis à vis du revenu universel si un référendum était organisé (3 modalités : **ABS** / **CONTRE** / **POUR**),
- **age_group** = class d'âge (3 modalités : **14_25** / **26_39** / **40_65**).

2.1 Analyse Factorielle des Correspondances

On étudie ici la relation entre la variable **awareness** "niveau de connaissance du concept de revenu universel" et la variable **vote** "positionnement vis à vis du revenu universel". En croisant ces deux variables, on obtient la table de contingence suivante :

Les résultats de l'AFC de cette table de contingence sont fournis dans l'Annexe 2.

	ABS	CONTRE	POUR
know_little	250	651	1256
know_nothing	460	321	601
know_some	198	840	2153
know_well	80	458	1728

TABLE 1 – Table de contingence

1. Combien d'axes sont nécessaires pour représenter la totalité de la relation entre ces deux variables ?
2. Sans faire de calculs, ordonner de la plus petite à la plus grande, les 3 quantités suivantes : $d_{\chi^2}(\text{know_well}, \text{POUR})$, $d_{\chi^2}(\text{know_well}, \text{CONTRE})$, $d_{\chi^2}(\text{know_well}, \text{ABS})$. Que pouvez-vous en déduire sur les européens favorables au revenu universel ?
3. À partir des résultats de l'AFC :
 - l'axe 1 est-il le lieu d'une association exclusive ?
 - quelle est l'intensité de la liaison entre ces 2 variables ?
 - quelle est la modalité de la variable **vote** qui contribue le plus à l'axe 1 ?
 - quelle est la modalité de la variable **awareness** qui contribue le plus à l'axe 1 ?
4. À partir de la représentation simultanée des modalités, commentez brièvement la relation entre ces 2 variables catégorielles.

2.2 Analyse Factorielle des Correspondances Multiples

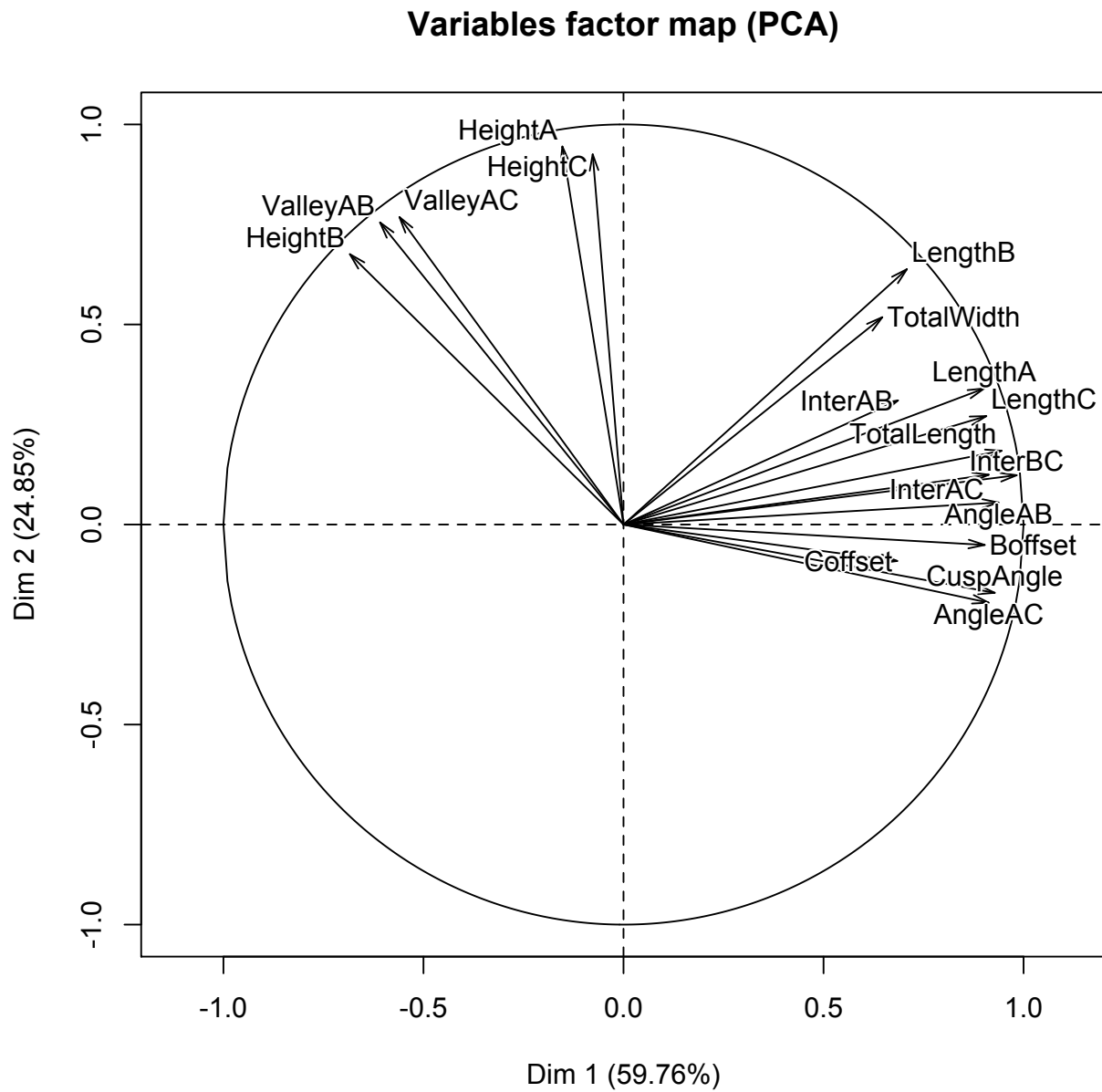
On enlève les individus ayant répondu **ABS** pour se concentrer uniquement sur ceux qui sont soit **POUR**, soit **CONTRE**. Vous trouverez dans l'Annexe 3 des résultats partiels de l'AFCM.

1. Est-ce que toutes les variables ont participé à la construction des axes factoriels ?
2. Que vaut l'inertie moyenne ?
3. Dans la représentation des variables catégorielles, quelle est l'abscisse de la variable **job** ? Que représente cette quantité ?

4. Donner deux pays pour lesquels on observe une différence marquée relativement à la variable `job`.
5. Donner deux pays pour lesquels on observe une différence marquée relativement à la variable `vote`.
6. À partir des représentations graphiques, qu'est-ce qui caractérise principalement les individus qui voteraient en faveur (i.e. `POUR`) du revenu universel de ceux qui y sont opposés (i.e vote `CONTRE`) ?

Annexes

Annexe 1 : ACP



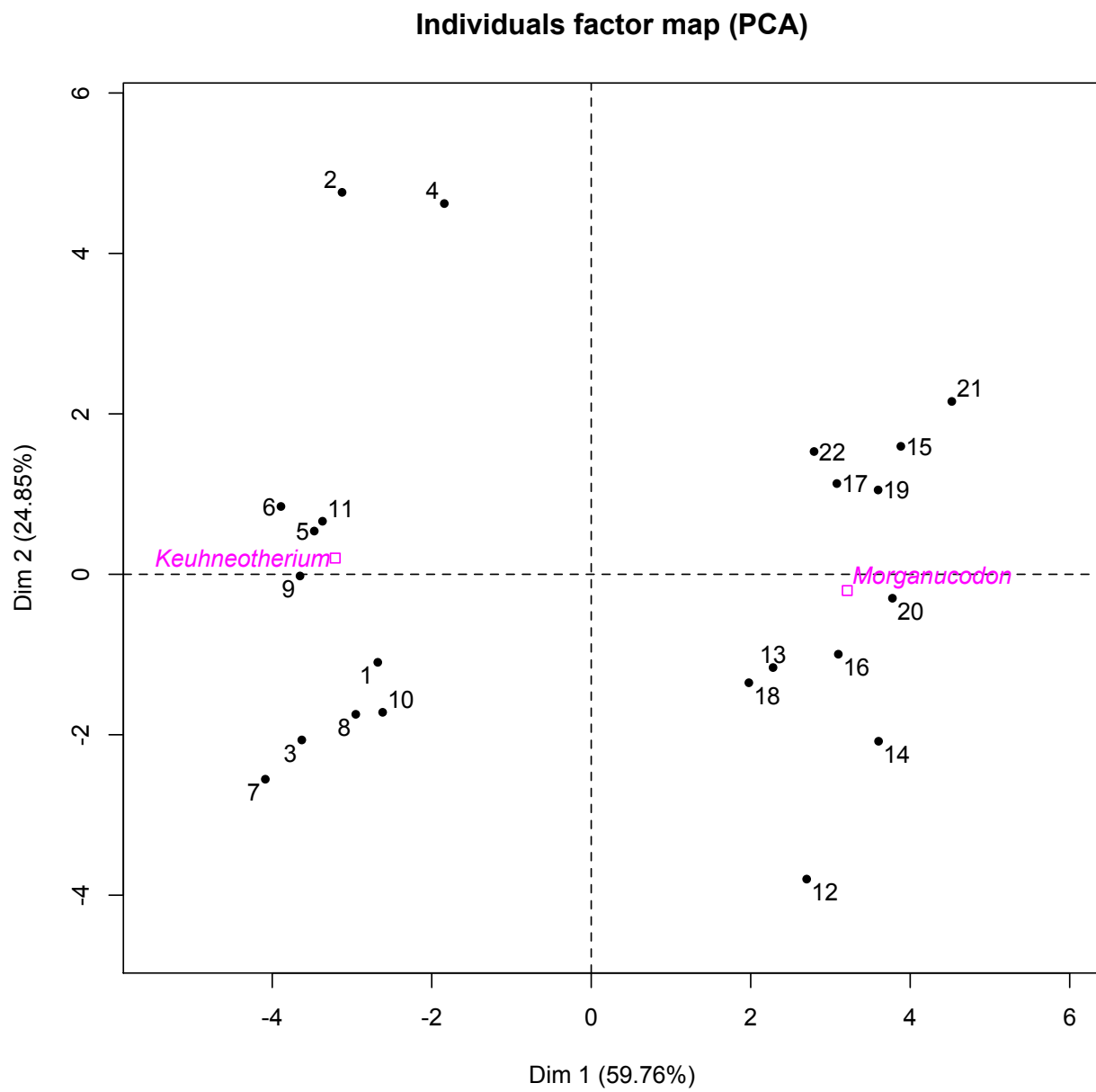


FIGURE 1 –

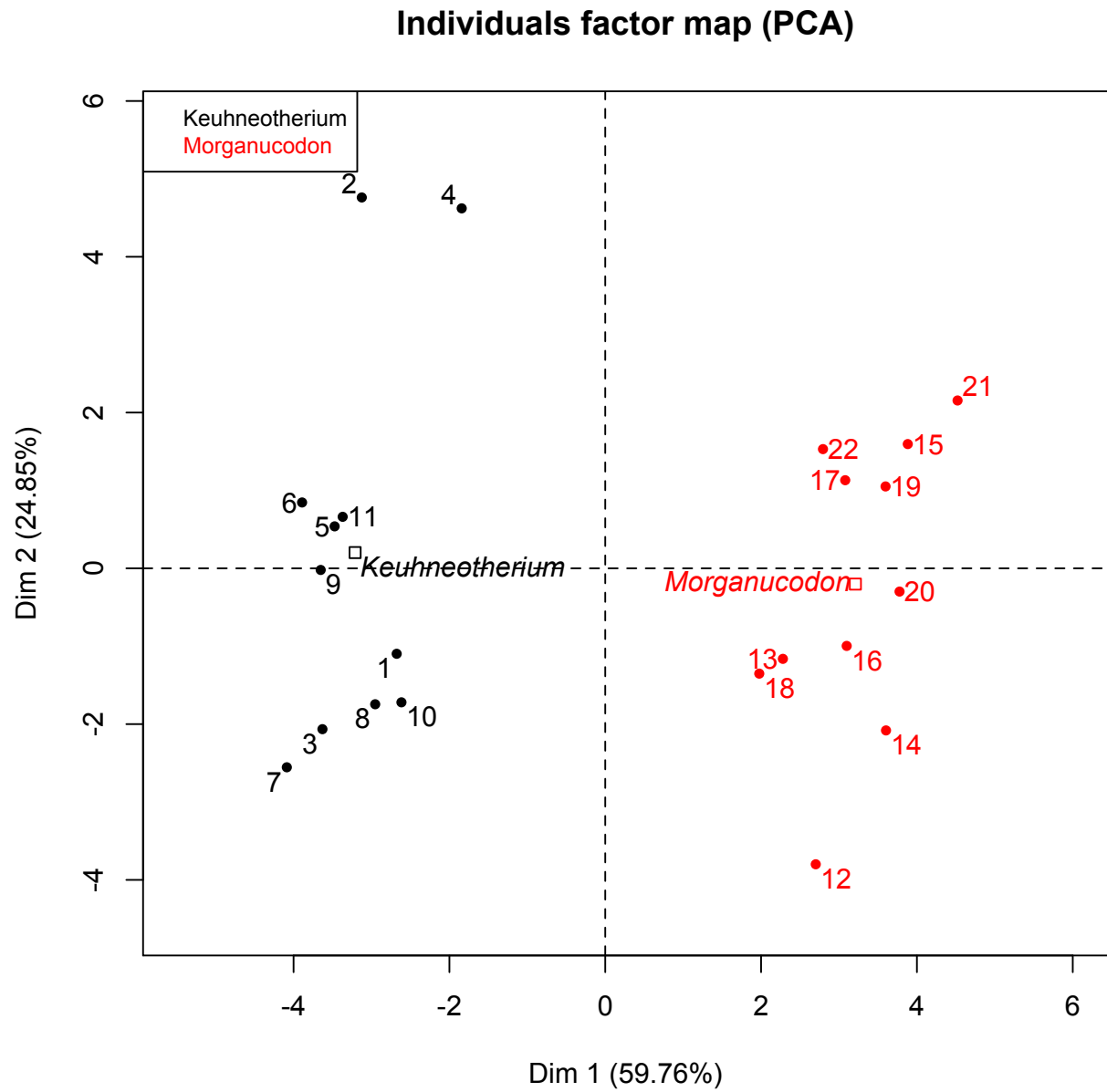
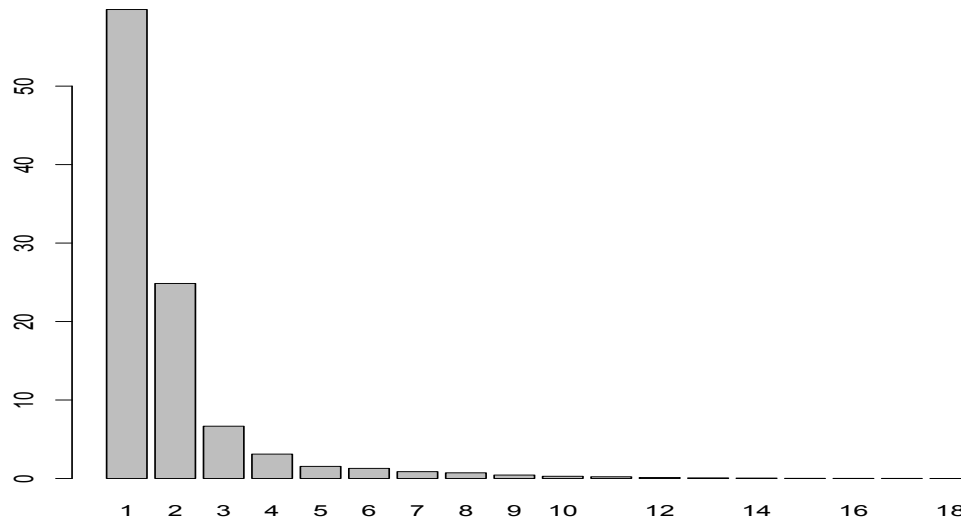


FIGURE 2 –

Décroissance des inerties (%)



Species	HeightA	HeightB	HeightC	ValleyAB
Keuhnetherium:11	Min. :0.990	Min. :0.4990	Min. :0.6150	Min. :0.3290
Morganucodon :11	1st Qu.:1.179	1st Qu.:0.6462	1st Qu.:0.7565	1st Qu.:0.4863
	Median :1.302	Median :0.7665	Median :0.7950	Median :0.5790
	Mean :1.276	Mean :0.7699	Mean :0.8145	Mean :0.5897
	3rd Qu.:1.361	3rd Qu.:0.8862	3rd Qu.:0.8770	3rd Qu.:0.6817
	Max. :1.622	Max. :1.1390	Max. :1.0290	Max. :0.9110
ValleyAC	LengthA	LengthB	LengthC	TotalLength
Min. :0.4210	Min. :0.7750	Min. :0.5040	Min. :0.5730	Min. :0.8430
1st Qu.:0.5483	1st Qu.:0.9235	1st Qu.:0.6122	1st Qu.:0.6940	1st Qu.:0.9615
Median :0.5825	Median :1.2055	Median :0.6955	Median :0.8785	Median :1.3440
Mean :0.6189	Mean :1.1785	Mean :0.7095	Mean :0.8604	Mean :1.3399
3rd Qu.:0.6597	3rd Qu.:1.4132	3rd Qu.:0.8180	3rd Qu.:1.0128	3rd Qu.:1.6805
Max. :0.9190	Max. :1.6300	Max. :0.9030	Max. :1.1880	Max. :1.9290
AngleAB	AngleAC	Boffset	Coffset	TotalWidth
Min. :40.23	Min. :45.53	Min. :110.7	Min. :113.0	Min. :0.4410
1st Qu.:48.33	1st Qu.:49.47	1st Qu.:119.1	1st Qu.:121.4	1st Qu.:0.5633
Median :60.58	Median :56.59	Median :128.1	Median :125.8	Median :0.6835
Mean :59.60	Mean :60.93	Mean :126.8	Mean :126.2	Mean :0.6523
3rd Qu.:67.85	3rd Qu.:73.55	3rd Qu.:134.9	3rd Qu.:132.4	3rd Qu.:0.7280
Max. :78.74	Max. :80.69	Max. :146.6	Max. :139.3	Max. :0.8570
InterAB	InterAC	InterBC	CuspAngle	
Min. :0.2790	Min. :0.2750	Min. :0.4990	Min. :105.0	
1st Qu.:0.3678	1st Qu.:0.3417	1st Qu.:0.6175	1st Qu.:116.1	
Median :0.4230	Median :0.3995	Median :0.8685	Median :157.9	
Mean :0.4338	Mean :0.4873	Mean :0.8686	Mean :147.4	
3rd Qu.:0.4865	3rd Qu.:0.6615	3rd Qu.:1.1390	3rd Qu.:176.0	
Max. :0.6310	Max. :0.8250	Max. :1.2620	Max. :179.4	

Eigenvalues										
	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6	Dim.7	Dim.8	Dim.9	Dim.10
Variance	10.757	4.473	1.199	0.559	0.277	0.231	0.156	0.130	0.078	0.050
% of var.	59.763	24.852	6.662	3.106	1.541	1.286	0.869	0.720	0.436	0.278
Cumulative % of var.	59.763	84.615	91.277	94.383	95.924	97.209	98.078	98.798	99.234	99.512
	Dim.11	Dim.12	Dim.13	Dim.14	Dim.15	Dim.16	Dim.17	Dim.18		
Variance	0.039	0.021	0.013	0.009	0.003	0.002	0.001	0.000		
% of var.	0.216	0.119	0.071	0.048	0.016	0.010	0.007	0.000		
Cumulative % of var.	99.729	99.848	99.919	99.967	99.982	99.992	100.000	100.000		
Individuals										
	Dist	Dim.1	ctr	cos2	Dim.2	ctr	cos2	Dim.3	ctr	cos2
1	3.557	-2.677	3.027	0.566	-1.097	1.223	0.095	1.235	5.784	0.121
2	5.761	-3.125	4.125	0.294	4.761	23.035	0.683	-0.250	0.237	0.002
3	4.250	-3.630	5.566	0.729	-2.065	4.333	0.236	0.197	0.147	0.002
4	5.201	-1.842	1.434	0.125	4.622	21.703	0.790	1.252	5.942	0.058
5	3.848	-3.473	5.097	0.815	0.539	0.295	0.020	-0.729	2.017	0.036
6	4.367	-3.891	6.397	0.794	0.846	0.726	0.037	-0.705	1.882	0.026
7	4.890	-4.087	7.057	0.698	-2.554	6.630	0.273	-0.354	0.475	0.005
8	3.959	-2.953	3.684	0.556	-1.745	3.095	0.194	-1.085	4.463	0.075
9	4.004	-3.652	5.635	0.832	-0.021	0.000	0.000	1.092	4.521	0.074
10	3.306	-2.616	2.891	0.626	-1.720	3.007	0.271	-0.303	0.349	0.008
11	3.790	-3.370	4.798	0.791	0.661	0.445	0.030	0.819	2.545	0.047
12	4.973	2.702	3.084	0.295	-3.800	14.670	0.584	1.296	6.362	0.068
13	2.857	2.280	2.196	0.637	-1.162	1.373	0.166	-0.753	2.147	0.069
14	4.464	3.603	5.486	0.652	-2.081	4.401	0.217	0.659	1.644	0.022
15	4.838	3.883	6.370	0.644	1.595	2.586	0.109	1.997	15.121	0.170
16	3.522	3.099	4.057	0.774	-0.996	1.008	0.080	-0.265	0.267	0.006
17	3.548	3.080	4.008	0.754	1.131	1.301	0.102	-0.666	1.681	0.035
18	2.918	1.977	1.651	0.459	-1.352	1.857	0.215	-1.314	6.543	0.203
19	4.222	3.597	5.468	0.726	1.051	1.123	0.062	-0.354	0.475	0.007
20	4.117	3.777	6.028	0.842	-0.298	0.090	0.005	0.984	3.669	0.057
21	5.079	4.522	8.639	0.793	2.155	4.718	0.180	0.222	0.187	0.002
22	4.468	2.794	3.299	0.391	1.531	2.380	0.117	-2.975	33.543	0.443

Variables									
	Dim.1	ctr	cos2	Dim.2	ctr	cos2	Dim.3	ctr	cos2
HeightA	-0.153	0.217	0.023	0.945	19.956	0.893	0.129	1.380	0.017
HeightB	-0.684	4.349	0.468	0.676	10.208	0.457	0.202	3.400	0.041
HeightC	-0.077	0.055	0.006	0.925	19.141	0.856	-0.331	9.159	0.110
ValleyAB	-0.609	3.443	0.370	0.755	12.744	0.570	0.100	0.841	0.010
ValleyAC	-0.560	2.916	0.314	0.769	13.203	0.591	-0.013	0.014	0.000
LengthA	0.902	7.556	0.813	0.340	2.585	0.116	-0.109	0.990	0.012
LengthB	0.708	4.666	0.502	0.638	9.106	0.407	-0.105	0.915	0.011
LengthC	0.907	7.646	0.823	0.271	1.642	0.073	0.197	3.226	0.039
TotalLength	0.945	8.305	0.893	0.184	0.760	0.034	-0.136	1.539	0.018
AngleAB	0.937	8.170	0.879	0.055	0.067	0.003	-0.160	2.129	0.026
AngleAC	0.913	7.744	0.833	-0.195	0.853	0.038	0.253	5.350	0.064
Boffset	0.903	7.578	0.815	-0.051	0.058	0.003	-0.137	1.571	0.019
Coffset	0.684	4.353	0.468	-0.091	0.186	0.008	0.661	36.477	0.437
TotalWidth	0.647	3.891	0.419	0.518	6.003	0.269	0.031	0.080	0.001
InterAB	0.686	4.377	0.471	0.310	2.151	0.096	0.527	23.127	0.277
InterAC	0.913	7.749	0.834	0.124	0.346	0.015	-0.281	6.608	0.079
InterBC	0.983	8.979	0.966	0.123	0.339	0.015	-0.052	0.222	0.003
CuspAngle	0.928	8.006	0.861	-0.171	0.651	0.029	-0.189	2.972	0.036

Supplementary categories										
	Dist	Dim.1	cos2	v.test	Dim.2	cos2	v.test	Dim.3	cos2	v.test
Keuhneotherium	3.219	-3.210	0.995	-4.485	0.202	0.004	0.438	0.106	0.001	0.445
Morganucodon	3.219	3.210	0.995	4.485	-0.202	0.004	-0.438	-0.106	0.001	-0.445

Annexe 2 : AFC

Eigenvalues

	Dim.1	Dim.2
Variance	0.106	0.007
% of var.	93.928	6.072
Cumulative % of var.	93.928	100.000

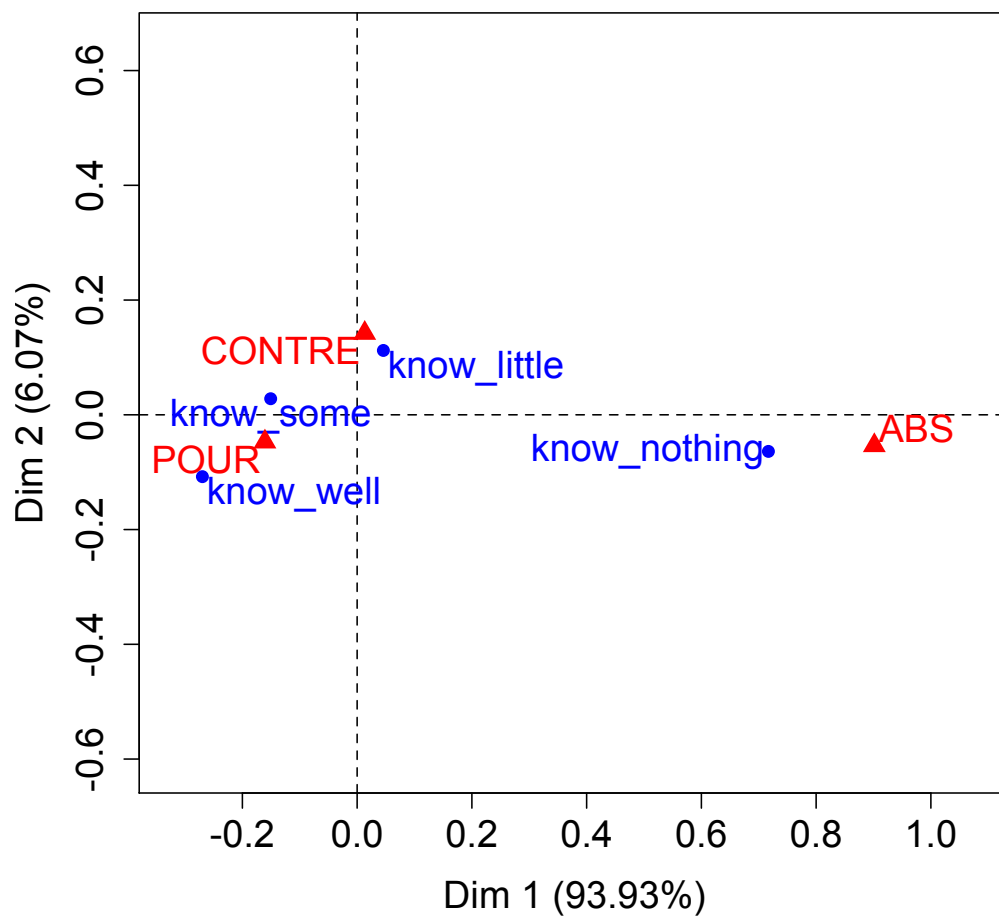
Rows

	Iner*1000	Dim.1	ctr	cos2	Dim.2	ctr	cos2
know_little	3.522	0.046	0.474	0.143	0.112	44.136	0.857
know_nothing	79.650	0.717	74.656	0.992	-0.064	9.144	0.008
know_some	8.343	-0.151	7.618	0.967	0.028	4.076	0.033
know_well	21.179	-0.270	17.251	0.862	-0.108	42.645	0.138

Columns

	Iner*1000	Dim.1	ctr	cos2	Dim.2	ctr	cos2
ABS	89.656	0.901	84.402	0.996	-0.054	4.603	0.004
CONTRE	5.155	0.013	0.042	0.009	0.142	74.697	0.991
POUR	17.883	-0.161	15.556	0.921	-0.047	20.700	0.079

CA factor map



Annexe 3 : AFCM

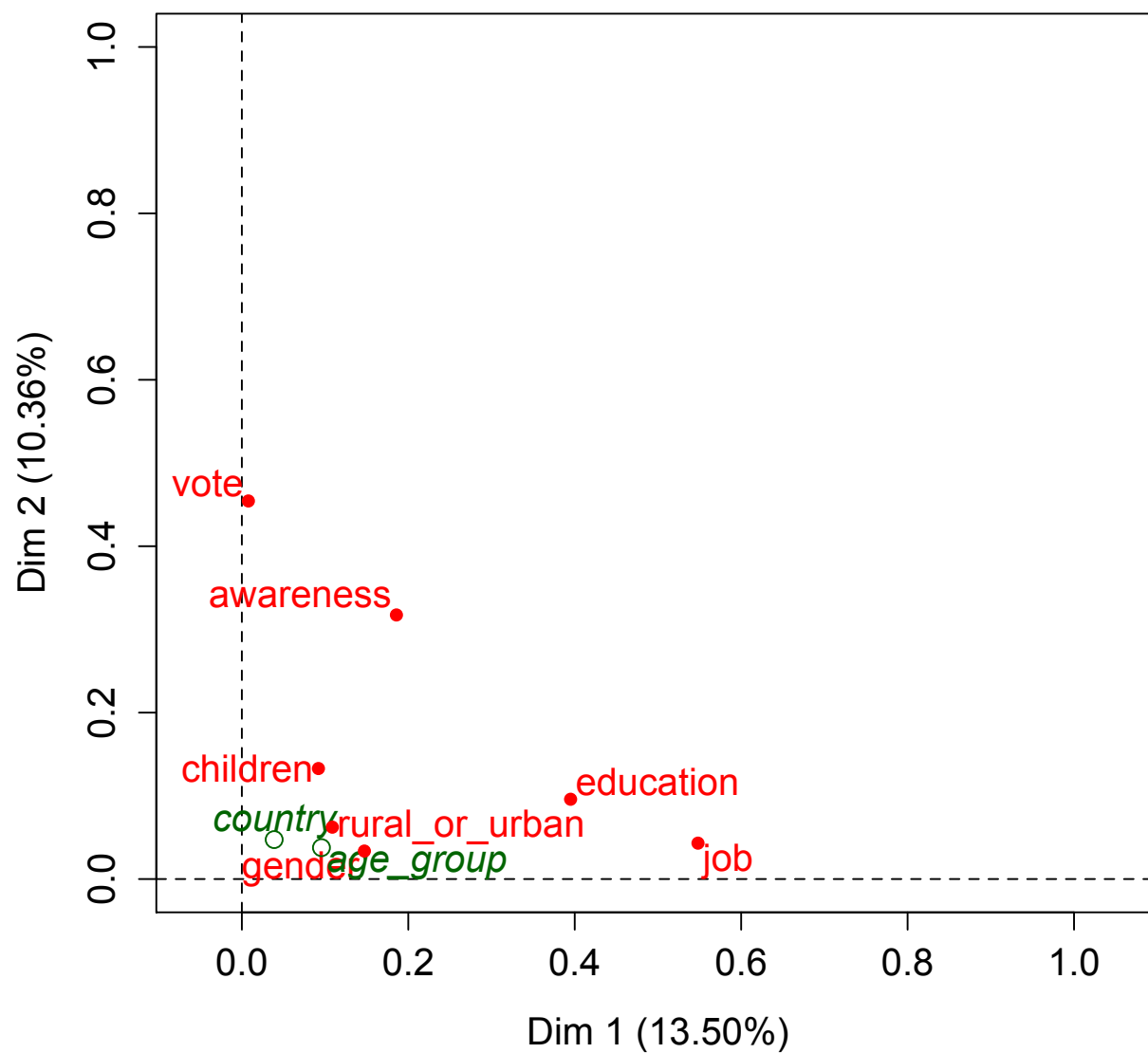


FIGURE 3 – Représentation des variables

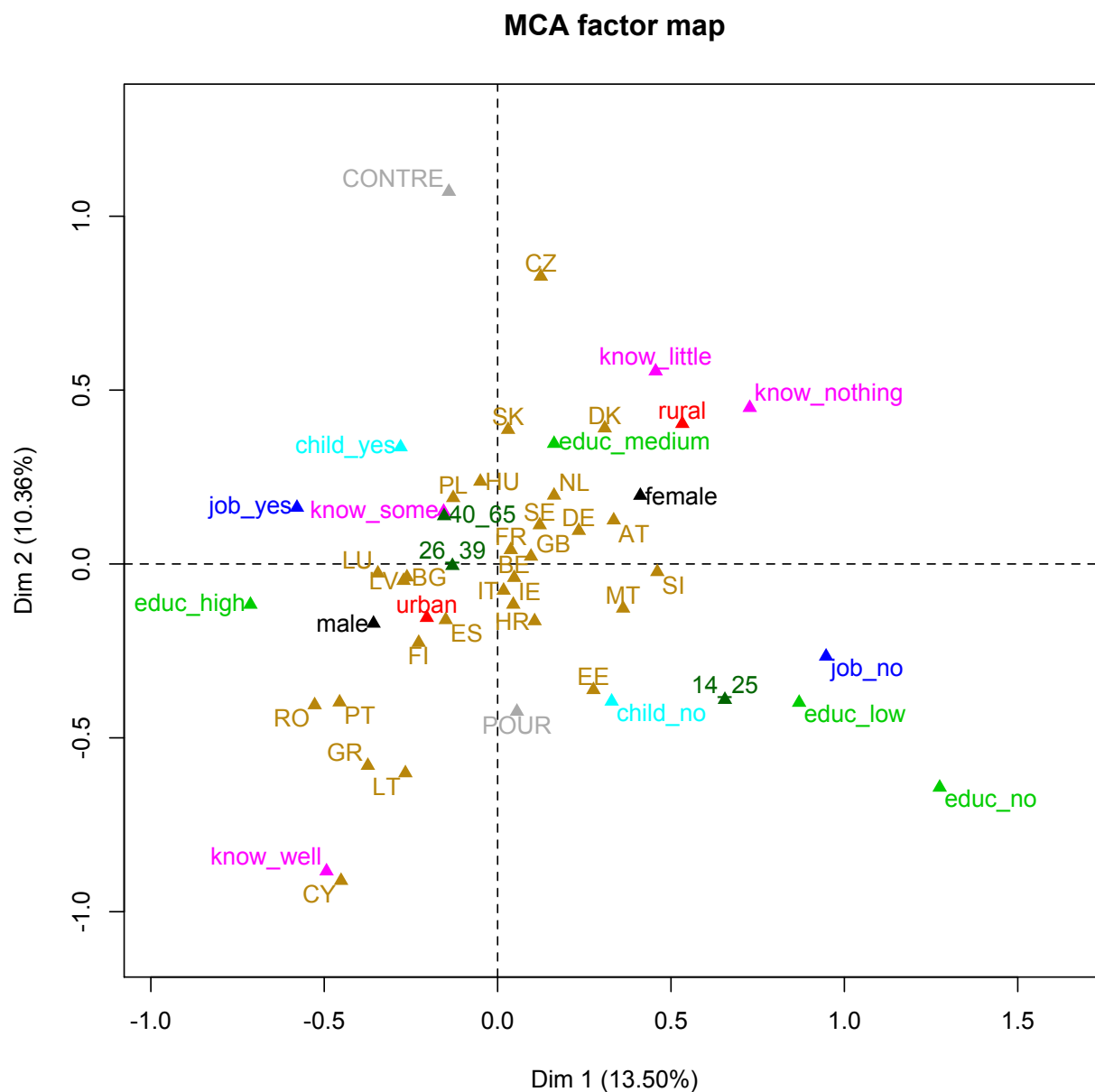


FIGURE 4 – Représentation des modalités

country	gender	rural_or_urban	education	job
DE :1206	female:3720	rural:2216	educ_high :3020	job_no :3033
IT :1009	male :4278	urban:5782	educ_low :1523	job_yes:4965
FR : 982			educ_medium:3216	
GB : 951			educ_no : 239	
ES : 890				
PL : 681				
(Other):2279				
children	awareness	vote	age_group	
child_no :3676	know_little :1907	CONTRE:2270	14_25:1455	
child_yes:4322	know_nothing: 922	POUR :5728	26_39:2342	
	know_some :2993		40_65:4201	
	know_well :2176			

Eigenvalues

	Dim.1	Dim.2	Dim.3	Dim.4	Dim.5	Dim.6
Variance	0.212	0.163	0.154	0.149	0.145	0.144
% of var.	13.497	10.360	9.780	9.484	9.243	9.174
Cumulative % of var.	13.497	23.857	33.637	43.121	52.364	61.538
	Dim.7	Dim.8	Dim.9	Dim.10	Dim.11	
Variance	0.140	0.136	0.126	0.112	0.091	
% of var.	8.933	8.659	8.001	7.106	5.763	
Cumulative % of var.	70.471	79.130	87.130	94.237	100.000	

Categorical variables (eta2)

	Dim.1	Dim.2	Dim.3
gender	0.147	0.034	0.385
rural_or_urban	0.109	0.062	0.002
education	0.395	0.096	0.178
job	0.548	0.043	0.013
children	0.092	0.133	0.268
awareness	0.186	0.317	0.105
vote	0.008	0.454	0.123

Supplementary categorical variables (eta2)

	Dim.1	Dim.2	Dim.3
country	0.039	0.047	0.029
age_group	0.096	0.038	0.031

