

## MAE5776 - 1º Semestre/2022 – Análise Multivariada - Lista 3

### Alunos:

Fernando F. Paulos Vieira - nº USP: 13492870

Leandro Alves da Silva - nº USP: 11868023

Thiago Ferreira Miranda - nº USP: 11925711

1 - A partir de uma matriz de dados normalizados  $Y_{n \times p}^*$ , considere a matriz de covariâncias  $nS_{p \times p} = Y^{*'}Y^* = V\Lambda V'$ , tal que  $V_{p \times p} = (V_1, \dots, V_p)$  e  $\Lambda = \text{diag}(\lambda_j)$  são matrizes de autovetores (das colunas de  $Y_{n \times p}^*$ ) e autovalores, respectivamente, e a matriz de distâncias  $D_{n \times n}$ , tal que seus elementos são função dos elementos de  $B_{n \times n} = Y^*Y^{*'} = U\Lambda U'$ , com  $U_{n \times n} = (U_1, \dots, U_n)$  matriz de autovetores (das linhas de  $Y_{n \times p}^*$ ). Três pesquisadores realizaram análises estatísticas e chegaram à seguinte redução de dimensionalidade de  $Y^*$ .

Pesquisador 1:  $Y_{n \times p}^* \rightarrow \tilde{Y}_{n \times 2} = Y^*(V_1 \ V_2)$

Pesquisador 2:  $Y_{n \times p}^* \rightarrow \tilde{Y}_{n \times 2} = Y^*\left(\frac{V_1}{\sqrt{\lambda_1}} \ \frac{V_2}{\sqrt{\lambda_2}}\right)$

Pesquisador 3:  $Y_{n \times p}^* \rightarrow \tilde{Y}_{n \times 2} = Y^*(V_1\sqrt{\lambda_1} \ V_2\sqrt{\lambda_2})$

1.1 - Qual análise estatística cada pesquisador realizou? Que propriedades dos dados estão preservadas em cada caso? Eles partiram do mesmo objetivo? Faça suposições necessárias.

R:

O pesquisador 1 - Realizou uma Análise de Componentes Principais. Buscou preservar a variância total dos dados (ou a maior proporção da variância total que possa ser preservada) em 2 componentes.

O pesquisador 2 - Realizou uma Análise Fatorial Exploratória. Buscou aproximar a matriz de covariâncias em termos de fatores latentes comuns e específicos, descrevendo as variáveis em função de 2 fatores latentes.

O pesquisador 3 - Realizou uma Análise de Coordenadas Principais ou Escalonamento Multidimensional. Analogamente ao Pesquisador 1, buscou preservar a variância total dos dados (ou a maior proporção da variância total que possa ser preservada).

Em um primeiro momento, acredita-se que estes pesquisadores partiram de um mesmo objetivo de redução de dimensionalidade das variáveis presentes nos dados originais disponíveis, buscando representar as  $p$  variáveis em  $m$  variáveis (neste caso,  $m = 2$ ).

1.2 - Simule dados e realize as análises dos três pesquisadores. Interprete os resultados.

Parâmetros da Simulação:

Amostra	Vetor de Médias	Matriz de Covariâncias
$n_1 = 100$	$\mu_1 = (5, 7, 10, 11, 6)$	$\Sigma_1 = \begin{pmatrix} 1 & 0.6 & 0.5 & 0.7 & 1 \\ 0.6 & 1.5 & 0.2 & 0.6 & 1.2 \\ 0.5 & 0.2 & 3 & 1.5 & 0.9 \\ 0.7 & 0.6 & 1.5 & 3.5 & 1.3 \\ 1 & 1.2 & 0.9 & 1.3 & 2 \end{pmatrix}$

Pesquisador 1 - Análise de Componentes Principais:

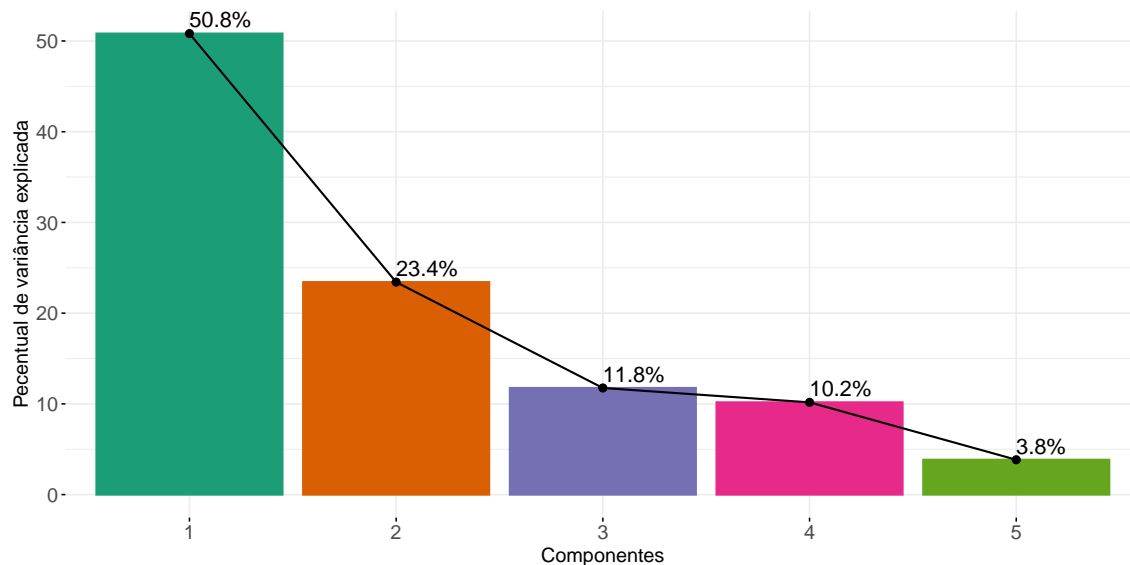
O pesquisador 1 buscou reduzir o banco de dados para dois componentes. Na sua análise de componentes principais com dados padronizados, os dois componentes encontrados explicam 74,2% da variabilidade total presentes nos dados. O primeiro componente encontrado é formado, principalmente, pela variáveis V1, V2, V4 e V5, com carga maior na variável V5(0,5828), este componente representa 50,8% da variabilidade total dos dados. No segundo componente destacam-se as variáveis V2, V3 e V4, sendo a V3 a variável com maior carga, este componente explicou 23,4% da variabilidade presente nos dados.

Autovalores dos componentes:

PC1	PC2	PC3	PC4	PC5
2.5407	1.1711	0.5877	0.5084	0.192

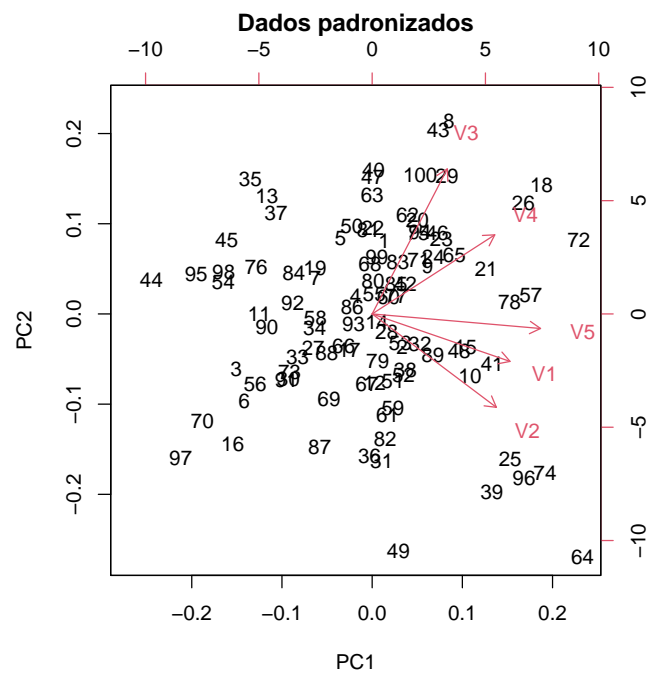
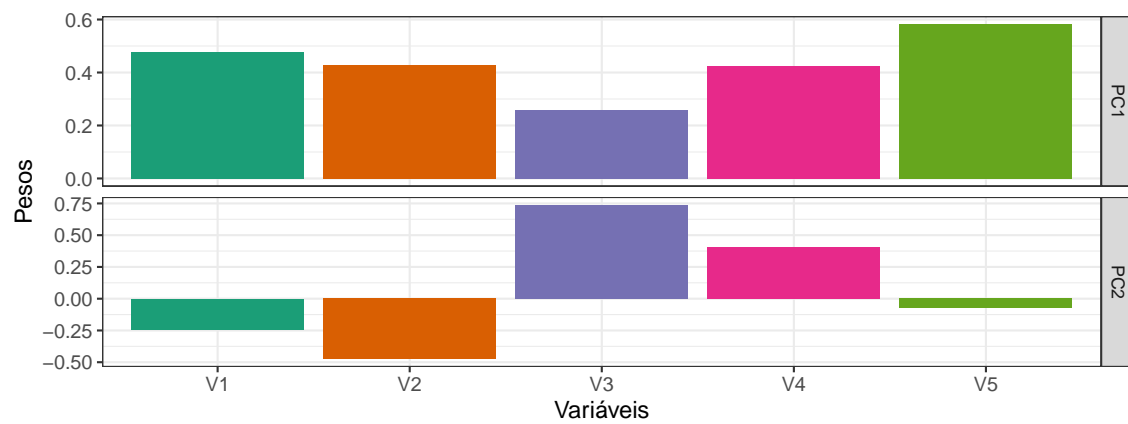
Importância dos componentes:

	PC1	PC2	PC3	PC4	PC5
Standard deviation	1.5940	1.0822	0.7666	0.7130	0.4382
Proportion of Variance	0.5082	0.2342	0.1175	0.1017	0.0384
Cumulative Proportion	0.5082	0.7424	0.8599	0.9616	1.0000



Autovetores:

	PC1	PC2	PC3	PC4	PC5
V1	0.4773	-0.2420	0.6577	0.3750	0.3748
V2	0.4299	-0.4760	-0.4156	-0.5143	0.3891
V3	0.2600	0.7394	0.2632	-0.5228	0.2075
V4	0.4244	0.4032	-0.5669	0.5580	0.1563
V5	0.5828	-0.0742	0.0635	-0.1009	-0.8004



### Pesquisador 2 - Análise Fatorial Exploratória via Máxima Verossimilhança - Sem Rotação:

O pesquisador 2 buscou reduzir a dimensionalidade dos dados padronizados para dois fatores por meio da análise fatorial exploratória, onde buscou preservar a variância comum dos dados padronizados. Com dois fatores estimados, o pesquisador 2 conseguiu explicar 62,6% da variância total a partir da communalidade. O fator 1 é majoritariamente composto por todas as variáveis, com grande influência da variável V5, enquanto o fator 2 é caracterizado pela variáveis V2, V3 e V4, com ênfase para a variável V3.

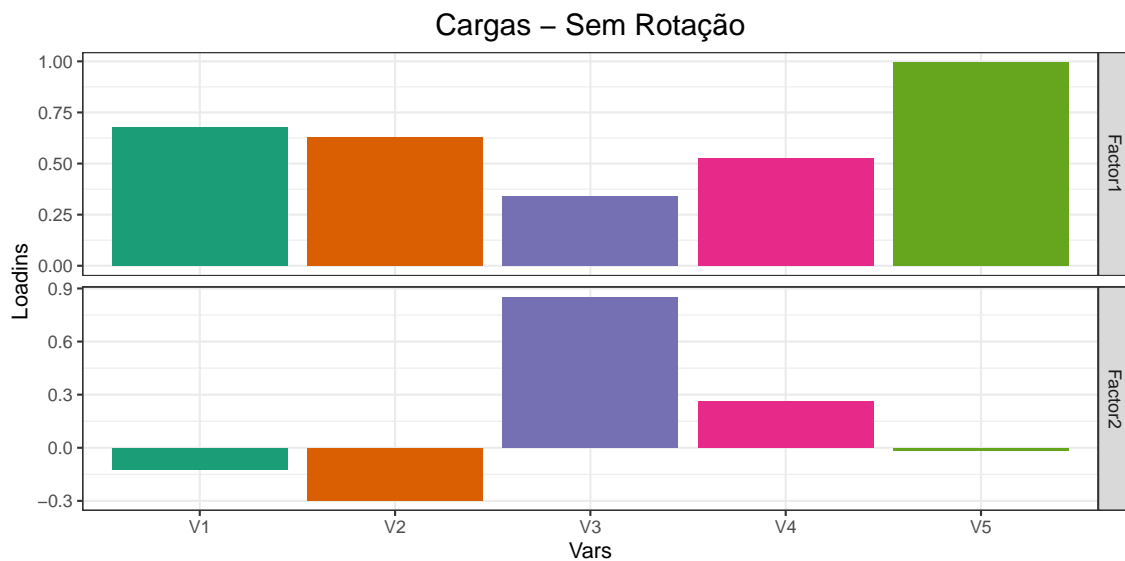
Cargas Fatoriais:

	Factor1	Factor2
V1	0.6763300	-0.1239107
V2	0.6271729	-0.2968888
V3	0.3401737	0.8520002
V4	0.5258317	0.2614647
V5	0.9948799	-0.0152961

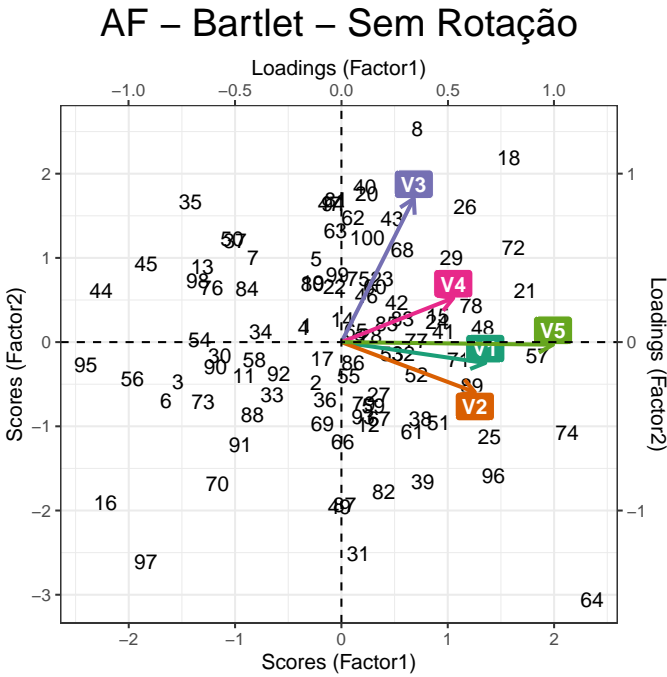
Variância Explicada:

	Factor1	Factor2
SS loadings	2.233	0.898
Proportion Var	0.447	0.180
Cumulative Var	0.447	0.626

Cargas ou coeficientes dos fatores comuns:



Biplot:



Especificidades:

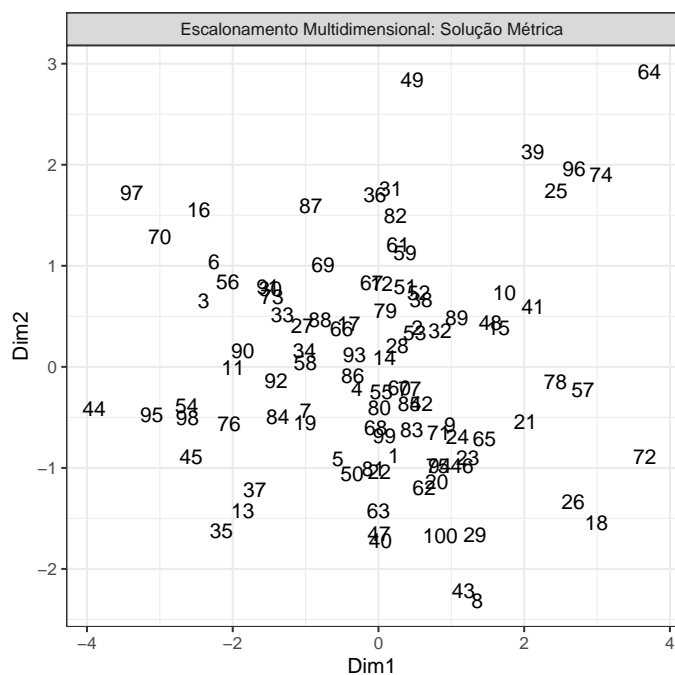
V1	V2	V3	V4	V5
0.527	0.518	0.158	0.655	0.01

Matrizes de correlação para a qual a aproximação é feita:

	V1	V2	V3	V4	V5
V1	1.0000	0.4255	0.1227	0.2989	0.6754
V2	0.4255	1.0000	-0.0403	0.2431	0.6290
V3	0.1227	-0.0403	1.0000	0.3997	0.3254
V4	0.2989	0.2431	0.3997	1.0000	0.5196
V5	0.6754	0.6290	0.3254	0.5196	1.0000

### Pesquisador 3 - Escalonamento Multidimensional:

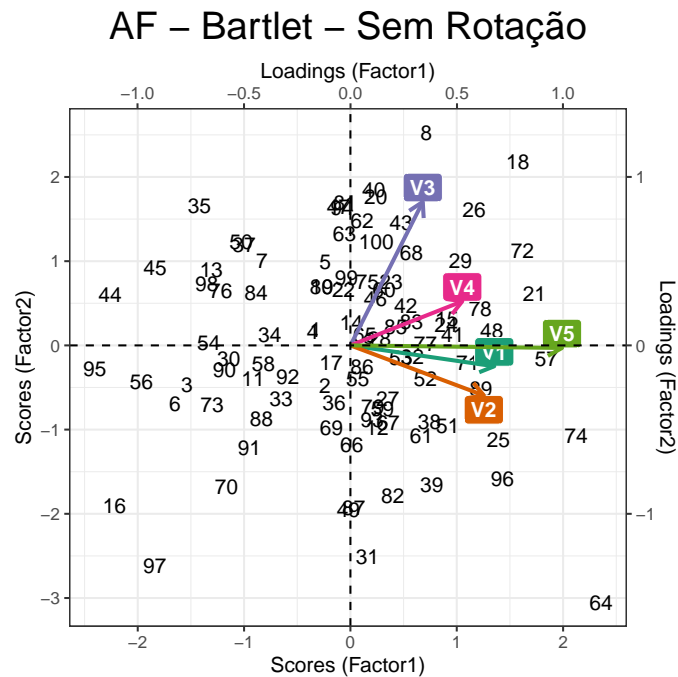
O pesquisador pesquisador 3 buscou mapear as distâncias entre os dados padronizados através da técnica de interdependência Escalonamento Multidimensional, técnica que facilitou a representação gráfica dos dados em duas dimensões.



1.3 - Para os dados simulados, obtenha uma representação Biplot. Como esse gráfico é construído?

R: O gráfico biplot representa cada observação dos dados, sendo construído pela dispersão de pares de escores gerados por alguma análise de redução de dimensionalidade e pelo sentido das cargas que geram tais escores. O que proporciona uma representação gráfica simultânea das  $n$  observações e  $p$  variáveis em  $R^2$ .

Biplot entre os dois fatores da análise fatorial:



2 - Considere os dados “bodyfat” disponíveis na biblioteca TH.data do R. Neste caso, a matriz de trabalho contém 71 observações avaliadas em 10 variáveis. Gere 5 observações (para tanto, adote um critério) e considere seu novo conjunto de dados “bodyfat\_new”. Com base na matriz de trabalho resultante realize as seguintes análises:

### 2.1 - Componentes Principais.

R: Para o conjunto de dados “bodyfat\_new”, com 76 observações, foram realizadas duas análises de componentes principais, a primeira com os dados originais e a segunda com os dados padronizados.

A análise considerando dos dados originais conseguiu reunir em dois componentes cerca de 95% da variabilidade total dos dados, enquanto que a análise com os dados padronizados conseguiu explicar cerca de 80%.

Ao analisármos para os componentes gerados pela primeira análise, percebemos que apenas as variáveis *age*, *DEXfat*, *hipcirc* e *waistcirc* receberam cargas relevantes. Podemos observar também que o primeiro componente é formado por cargas negativas nessas variáveis (*age*, *DEXfat*, *hipcirc* e *waistcirc*). Enquanto que no segundo componente temos carga negativa apenas para a variável *age*, as variáveis *DEXfat*, *hipcirc* e *waistcirc* receberam sinal positivos.

Já quando analisamos os componentes formados com os dados padronizados, percebemos que todas as variáveis são levadas em consideração. O primeiro componente é formado por cargas positivas em todas as variáveis, onde as variáveis *anthro3a*, *anthro3b*, *anthro3c*, *anthro4*, *Dexfat*, *hipcirc*, *kneebreadth* e *waistcirc* tem as cargas mais elevadas. O segundo componente é composto, principalmente, pelas variáveis *age*, *elbowbreadth* e *kneebreadth*.

Centróides dos componentes:

PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
50.75	30.6439	87.2789	105.0844	6.5048	9.3218	3.8605	4.2793	3.8764	5.383

Centróides dos componentes com dados padronizados:

PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
50.75	30.6439	87.2789	105.0844	6.5048	9.3218	3.8605	4.2793	3.8764	5.383

Autovalores dos componentes:

PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
415.484	162.7966	19.4313	11.2885	0.3823	0.2692	0.1336	0.0345	0.0077	0.0018

Autovalores dos componentes com dados padronizados:

PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
6.7355	1.2452	0.7084	0.7067	0.2697	0.1289	0.1054	0.0638	0.0311	0.0054

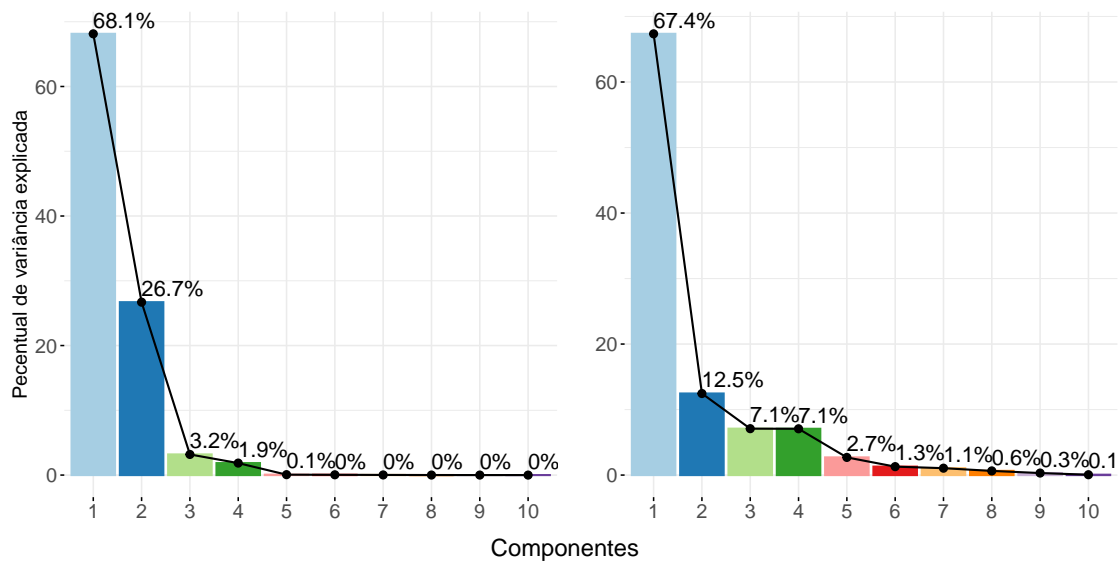


Importância dos componentes:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Standard deviation	20.3834	12.7592	4.4081	3.3598	0.6183	0.5189	0.3656	0.1859	0.088	0.0421
Proportion of Variance	0.6813	0.2670	0.0319	0.0185	0.0006	0.0004	0.0002	0.0001	0.000	0.0000
Cumulative Proportion	0.6813	0.9483	0.9801	0.9986	0.9993	0.9997	0.9999	1.0000	1.000	1.0000

Importância dos componentes com dados padronizados:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Standard deviation	2.5953	1.1159	0.8416	0.8407	0.5193	0.3590	0.3247	0.2525	0.1763	0.0733
Proportion of Variance	0.6736	0.1245	0.0708	0.0707	0.0270	0.0129	0.0105	0.0064	0.0031	0.0005
Cumulative Proportion	0.6736	0.7981	0.8689	0.9396	0.9666	0.9794	0.9900	0.9964	0.9995	1.0000

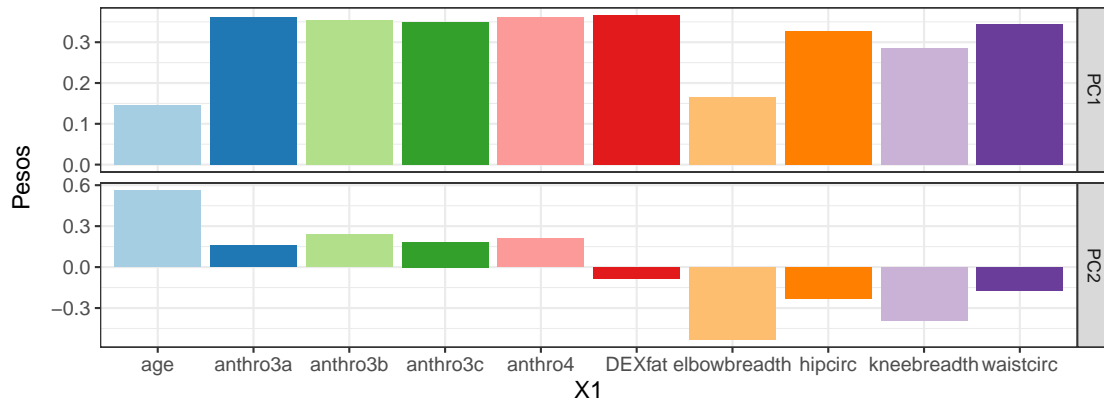
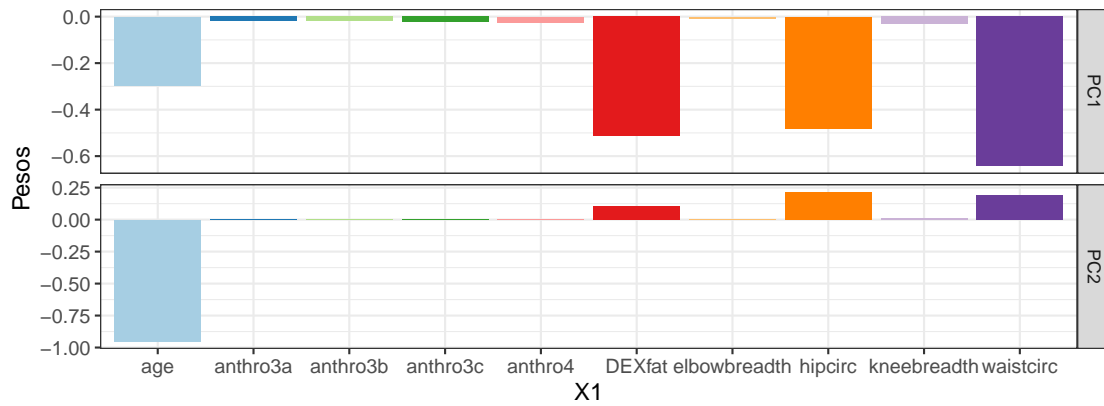


Autovetores:

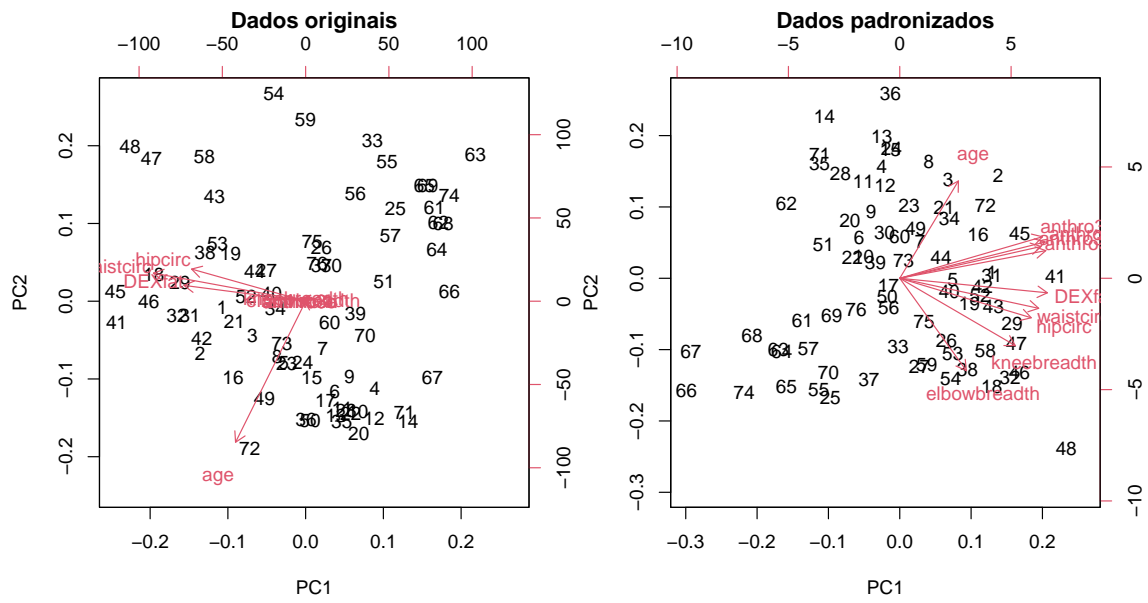
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
age	-0.2959	-0.9515	0.0355	-0.0753	0.0094	0.0023	0.0044	0.0023	-0.0003	0.0002
DEXfat	-0.5140	0.1068	0.2621	0.8026	0.0261	0.1014	0.0228	0.0014	-0.0014	-0.0024
waistcirc	-0.6428	0.1895	-0.7129	-0.2054	-0.0080	0.0151	-0.0094	-0.0039	0.0025	0.0016
hipcirc	-0.4817	0.2169	0.6491	-0.5460	-0.0312	-0.0203	0.0105	0.0016	0.0007	0.0001
elbowbreadth	-0.0078	0.0071	-0.0156	0.0082	0.2307	-0.3360	0.9058	0.1138	0.0083	0.0051
kneebreadth	-0.0327	0.0143	0.0146	0.0168	0.7956	-0.4714	-0.3769	-0.0086	0.0289	-0.0087
anthro3a	-0.0191	-0.0009	0.0016	0.0406	-0.1997	-0.3557	-0.0229	-0.4422	-0.6344	0.4827
anthro3b	-0.0187	-0.0020	0.0026	0.0479	-0.2752	-0.3689	-0.0712	-0.0442	0.7192	0.5111
anthro3c	-0.0220	0.0005	-0.0017	0.0473	-0.2914	-0.3843	-0.1690	0.8156	-0.2657	-0.0187
anthro4	-0.0254	-0.0024	0.0003	0.0573	-0.3340	-0.4936	-0.0513	-0.3525	0.0930	-0.7109

Autovetores com dados padronizados:

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
age	0.1456	0.5618	0.7324	0.3430	0.0407	-0.0149	0.0778	-0.0334	0.0094	-0.0046
DEXfat	0.3663	-0.0838	-0.0746	0.2012	0.1138	-0.0947	0.0292	0.8841	0.0780	0.0471
waistcirc	0.3447	-0.1733	-0.0219	0.2481	0.4187	0.6043	-0.4527	-0.1811	-0.0962	-0.0397
hipcirc	0.3263	-0.2279	-0.1581	0.3918	0.3615	-0.5205	0.3510	-0.3711	-0.0228	-0.0021
elbowbreadth	0.1653	-0.5331	0.5998	-0.5266	0.1871	-0.0455	0.1178	0.0189	-0.0016	-0.0044
kneebreadth	0.2866	-0.3884	0.1281	0.3342	-0.7856	0.0795	-0.0228	-0.1060	-0.0589	0.0140
anthro3a	0.3624	0.1596	-0.0591	-0.2252	-0.1031	-0.2918	-0.4290	-0.1220	0.5764	-0.4026
anthro3b	0.3541	0.2394	-0.1185	-0.2759	-0.0899	-0.0663	0.0453	0.0187	-0.7239	-0.4338
anthro3c	0.3505	0.1835	-0.1789	-0.2029	-0.0611	0.4818	0.6445	-0.0886	0.3333	0.0117
anthro4	0.3608	0.2094	-0.0829	-0.2641	-0.0695	-0.1567	-0.2215	-0.1097	-0.1152	0.8035

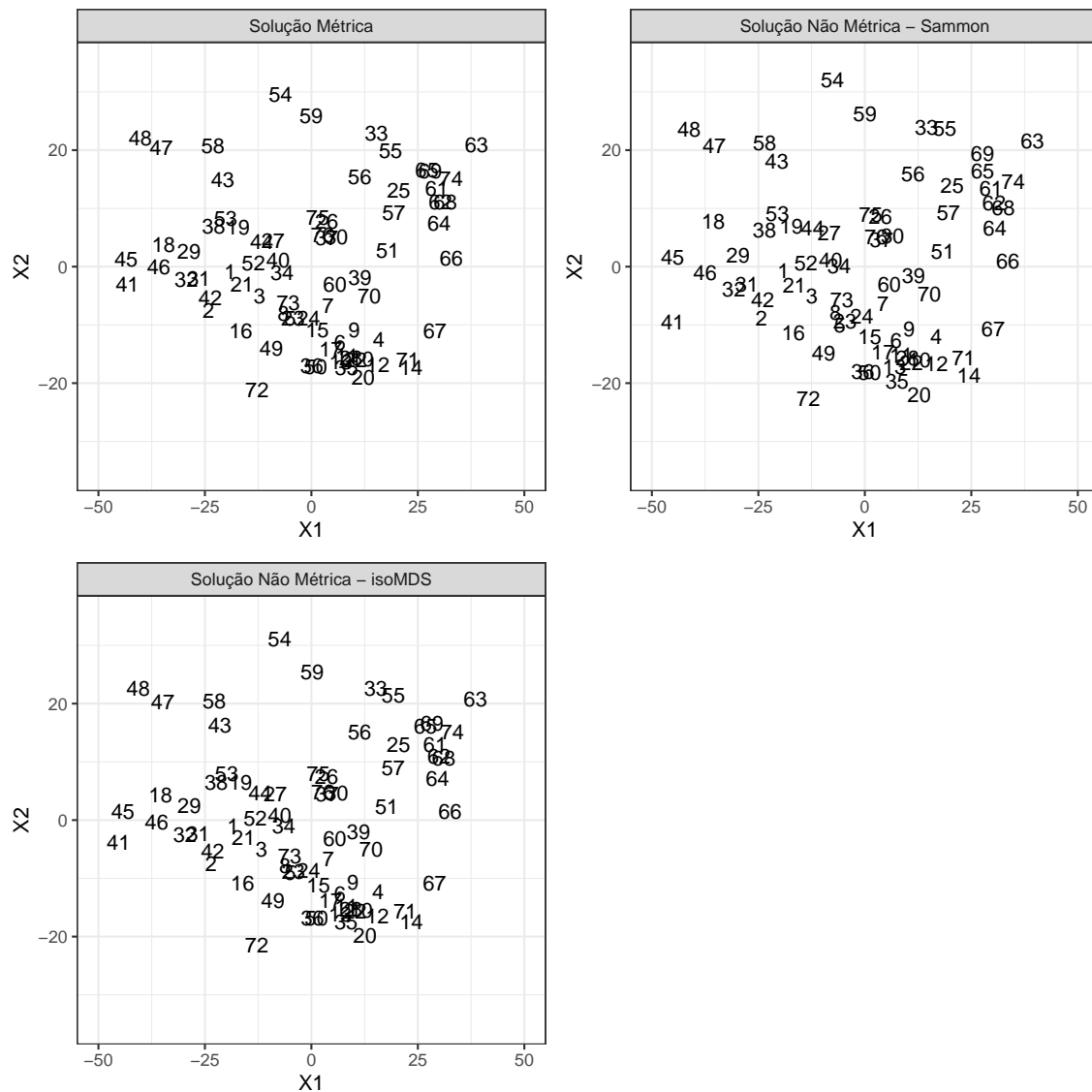


Biplots:



## 2.2 - Escalonamento Multidimensional (ou Coordenadas Principais) – Compare as soluções métricas e não-métricas.

R: Com os dados “bodyfat\_new” também foram realizadas três análises de escalonamento multidimensional. Essa técnica facilitou a representação gráfica dos dados em duas dimensões. Podemos observar que as análises com a solução métrica e a solução não métrica via isoMDS geraram representações similares para a distância entre as observações nas duas dimensões. Podemos perceber que a solução não métrica de Sammon gerou uma representação diferente das duas anteriores, onde alguns grupos de observações estão mais separados.



2.3 - Análise Fatorial Exploratória – Solução de MVS (rotacionar, se for de interesse). Em cada caso, que proporção da variância total dos dados pode ser explicada por 2 componentes? Quais variáveis mais influenciaram na redução de dimensionalidade? Represente os dados em eixos bidimensionais, identifique as observações e compare os resultados das três análises.

R: Foram feitas 3 análises fatoriais, a primeira sendo via Componentes Principais, a segunda e a terceira via Máxima Verossimilhança, sendo que na terceira rotacionamos via Promax. Em três fatores, análise fatorial via CP explicou 86,9% da variação total dos dados e a análise fatorial via MVS sem rotação e com rotação Promax explicou 74,5%.

A análise fatorial via CP teve equivalência com a análise de componentes principais com dados padronizados realizada anteriormente, entretanto na análise fatorial via CP consideramos 3 fatores ao invés de 2, onde 86,9% da variância total dos dados é explicada pelos 3 fatores. O primeiro fator é formado por cargas negativas em todas as variáveis, onde as variáveis *anthro3a*, *anthro3b*, *anthro3c*, *anthro4*, *Dexfat*, *hipcirc*, *kneebreadth* e *waistcirc* tem as cargas mais elevadas. O segundo fator é composto, principalmente, pelas variáveis *age* com sinal positivo, *elbowbreadth* e *kneebreadth* com sinais negativos. E o terceiro fator pelas variáveis *age*, *elbowbreadth* e *kneebreadth* com sinais positivos.

Para a análise fatorial via máxima verossimilhança onde não houve rotação, o primeiro fator foi composto por cargas positivas em todas as variáveis, onde *age* e *elbowbreadth* tiveram as menores cargas. O segundo fator foi composto pelas variáveis: *DEXfat*, *elbowbreadth*, *hipcirc*, *kneebreadth*, *waistcirc* e *anthro3b*, todos com cargas positivas, exceto *anthro3b*, com carga negativa. E o terceiro fator foi caracterizado pela variáveis: *elbowbreadth*, *anthro3a*, *anthro3b* e *anthro3c*.

A análise fatorial via MVS com rotação Promax conseguiu definir melhor quais variáveis compõe cada fator. O primeiro fator reduziu de 10 variáveis para 7, removendo *hipcirc*, *elbowbreadth* e *kneebreadth*. No segundo fator houve a saída da variável *anthro3b* e a entrada de *anthro3c*. E o terceiro fator foi reduzido de 4 para 3, com uma troca de variável, o terceiro fator foi composto por *elbowbreadth*, *anthro3a* e *anthro4*.

Análise Fatorial Exploratória via CP:

Autovalores e Variância explicada:

	CP01	CP02	CP03	CP04	CP05	CP06	CP07	CP08	CP09	CP10
Eigenvalues	6.736	1.245	0.708	0.707	0.270	0.129	0.105	0.064	0.031	0.005
Cumulative Proportion	0.674	0.798	0.869	0.940	0.967	0.979	0.990	0.996	0.999	1.000

Matriz de cargas:

	vecn1	vecn2	vecn3
	-0.3778	0.6269	0.6164
	-0.9508	-0.0935	-0.0627
	-0.8947	-0.1934	-0.0184
	-0.8469	-0.2543	-0.1331
	-0.4290	-0.5949	0.5048
	-0.7439	-0.4334	0.1078
	-0.9405	0.1781	-0.0497
	-0.9190	0.2672	-0.0998
	-0.9098	0.2047	-0.1505
	-0.9365	0.2337	-0.0698

Análise Fatorial Exploratória via Máxima Verossimilhança - Sem Rotação:

Cargas Fatoriais:

	Factor1	Factor2	Factor3
age	0.3845801	-0.0913021	-0.0044930
DEXfat	0.8629018	0.4494969	0.0796433
waistcirc	0.7716096	0.5146403	0.0470325
hipcirc	0.7091802	0.6036607	0.0741937
elbowbreadth	0.3310613	0.2387215	-0.1667511
kneebreadth	0.5814160	0.5647846	-0.0470744
anthro3a	0.9853906	0.0050984	-0.1401128
anthro3b	0.9834246	-0.1015120	0.1156568
anthro3c	0.9321144	0.0013199	0.1834629
anthro4	0.9944037	-0.0609075	-0.0131731

Variância Explicada:

	Factor1	Factor2	Factor3
SS loadings	6.234	1.230	0.111
Proportion Var	0.623	0.123	0.011
Cumulative Var	0.623	0.746	0.757

# Análise Fatorial Exploratória via Máxima Verossimilhança - Rotação Promax:

Cargas Fatoriais:

	Factor1	Factor2	Factor3
age	0.4170155	-0.0879643	0.0790798
DEXfat	0.2707014	0.7857773	-0.0148434
waistcirc	0.1236042	0.8388732	0.0056099
hipcirc	-0.0178822	0.9782093	-0.0503578
elbowbreadth	0.0160522	0.2286991	0.2455890
kneebreadth	-0.0973203	0.8088692	0.0923009
anthro3a	0.8095819	-0.0009900	0.3440842
anthro3b	0.9430760	0.0643860	0.0203834
anthro3c	0.8007471	0.2580595	-0.0886630
anthro4	0.8977215	0.0143857	0.1865528

Variância Explicada:

	Factor1	Factor2	Factor3
SS loadings	3.264	3.063	0.239
Proportion Var	0.326	0.306	0.024
Cumulative Var	0.623	0.746	0.757

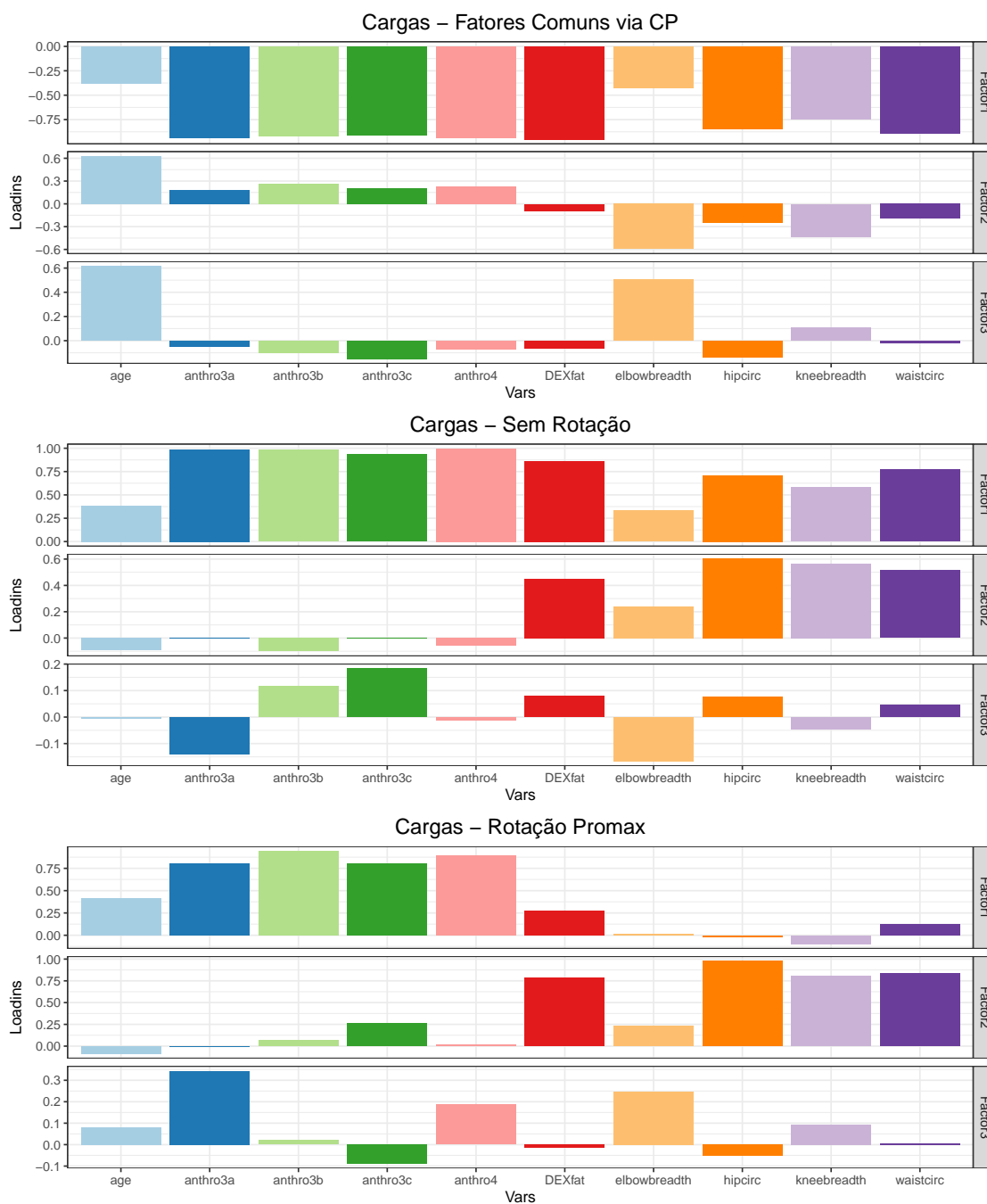
Especificidades:

age	DEXfat	waistcirc	hipcirc	elbowbreadth	kneebreadth	anthro3a	anthro3b	anthro3c	anthro4
0.844	0.047	0.138	0.127	0.806	0.341	0.01	0.01	0.098	0.01

Matrizes de correlação para a qual a aproximação é feita:

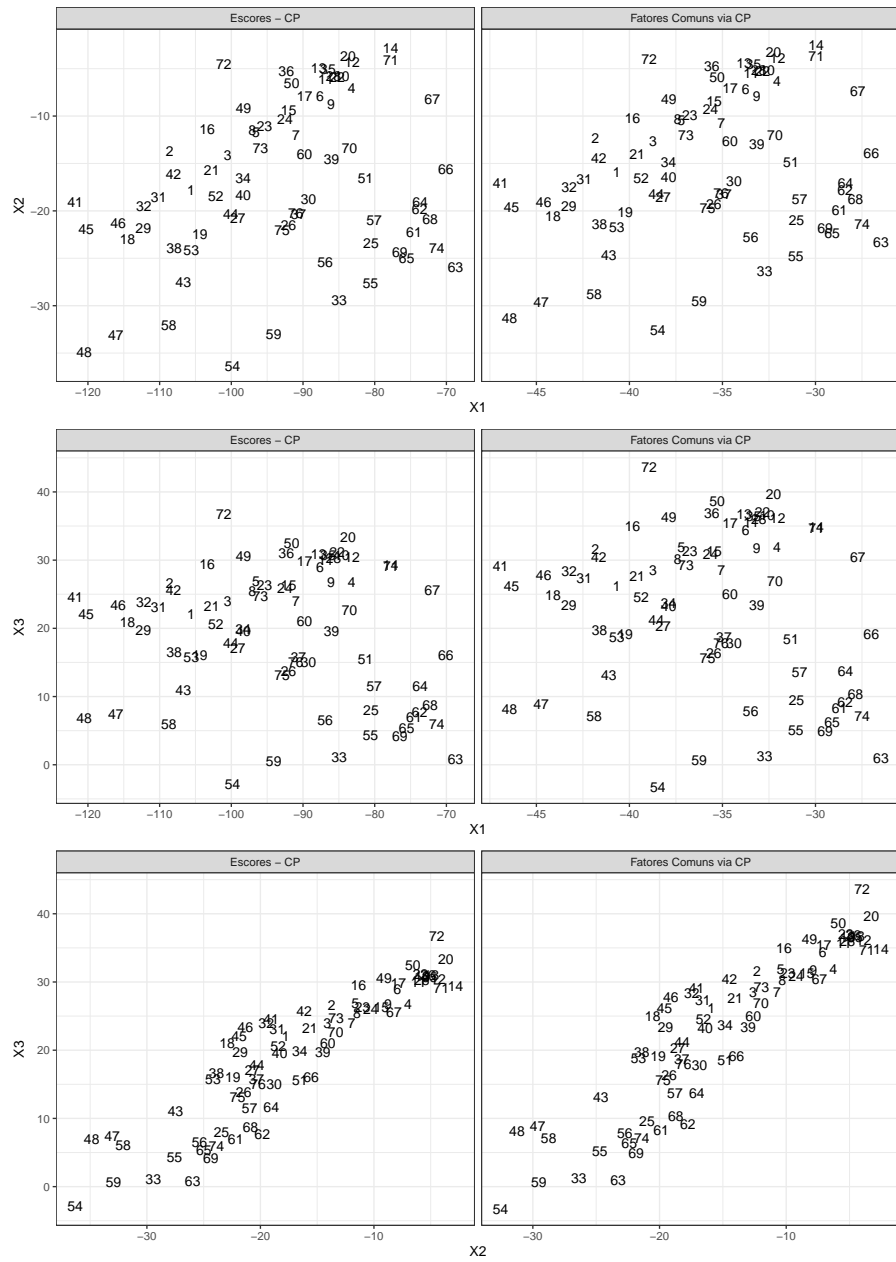
	age	DEXfat	waistcirc	hipcirc	elbowbreadth	kneebreadth	anthro3a	anthro3b	anthro3c	anthro4
age	1.0000	0.3105	0.2656	0.1820	-0.0243	0.1481	0.3780	0.3856	0.3340	0.3911
DEXfat	0.3105	1.0000	0.8987	0.8906	0.3647	0.7573	0.8422	0.8132	0.8165	0.8281
waistcirc	0.2656	0.8987	1.0000	0.8659	0.4090	0.7260	0.7543	0.7086	0.7414	0.7396
hipcirc	0.1820	0.8906	0.8659	1.0000	0.3268	0.7382	0.6916	0.6447	0.6697	0.6680
elbowbreadth	-0.0243	0.3647	0.4090	0.3268	1.0000	0.4665	0.3473	0.2841	0.2700	0.3203
kneebreadth	0.1481	0.7573	0.7260	0.7382	0.4665	1.0000	0.5836	0.5114	0.5402	0.5401
anthro3a	0.3780	0.8422	0.7543	0.6916	0.3473	0.5836	1.0000	0.9516	0.8930	0.9828
anthro3b	0.3856	0.8132	0.7086	0.6447	0.2841	0.5114	0.9516	1.0000	0.9382	0.9841
anthro3c	0.3340	0.8165	0.7414	0.6697	0.2700	0.5402	0.8930	0.9382	1.0000	0.9240
anthro4	0.3911	0.8281	0.7396	0.6680	0.3203	0.5401	0.9828	0.9841	0.9240	1.0000

Cargas ou coeficientes dos fatores comuns:

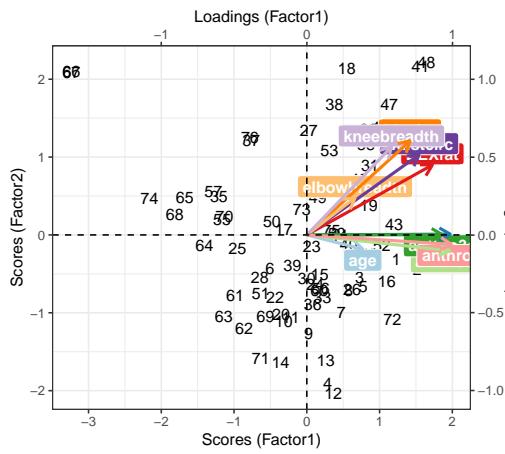




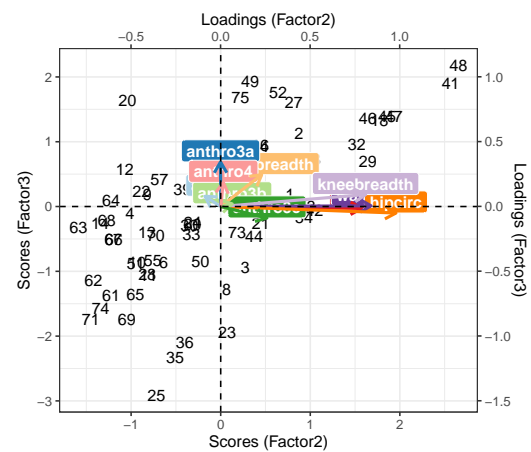
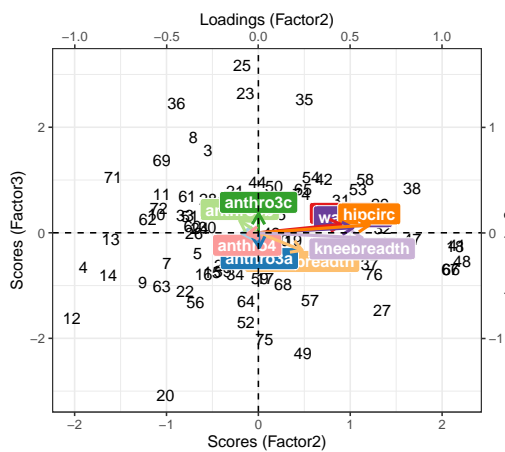
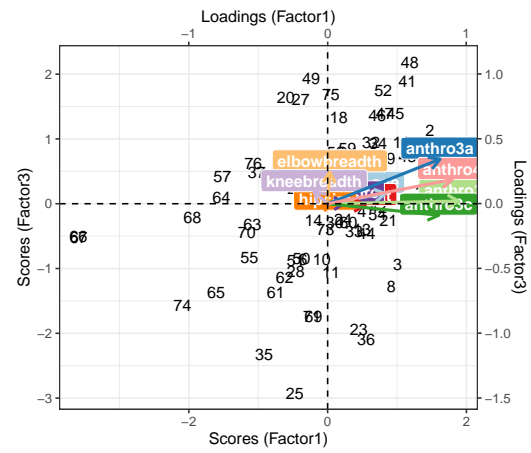
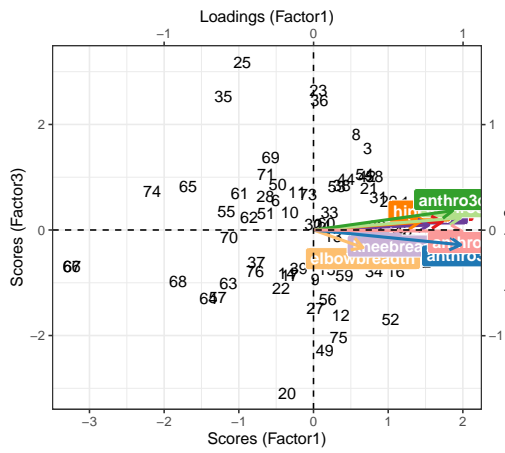
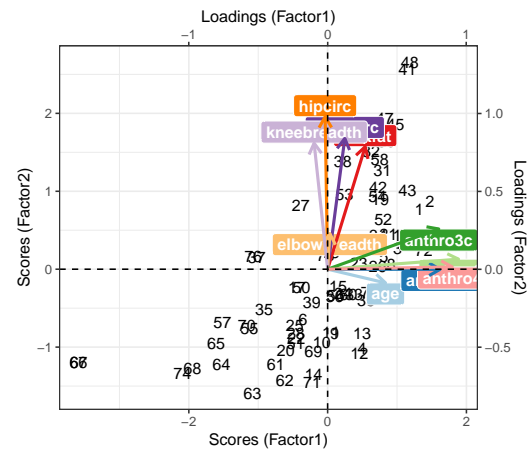
Plot de escore entre componentes e biplot entre os fatores:



AF – Bartlet – Sem Rotação



AF – Bartlet – Rotação Promax



2.4 - Escolha uma das variáveis do banco de dados e obtenha uma tabela de contingência categorizando esta variável de acordo com faixas etárias das observações. Realize uma Análise de Correspondência e comente sobre o padrão de associação presente nesses dados.

R: Temos 5 trinômios representando os perfis de frequência relativa no gráfico simplex 3D, fixando os totais de linha, dos dados totais. A representação do gráfico symetric permite verificar (no plano bidimensional) a associação das variáveis, e com estes podemos fixar as linhas e representar os pontos referentes às colunas ou também fixar as colunas e representar os pontos referentes às linhas. Já o plot Rowgreen é basicamente equivalente ao simplex. Temos também verificar o Colgreen onde damos um “zoom” nas populações da variável DEX\_fat e verificamos como se “encaixam” as categorias de “age”.

Por meio das representações gráficas podemos perceber que os indivíduos com “age” até 35 tendem a apresentar um índice DEX\_fat inferior a 20 (assim como “age” de 35 a 45 também se encontram no quadrante de DEX\_fat inferior a 20), indivíduos com “age” entre 55 e 65 tendem a apresentar uma associação um pouco mais próximas ao índice DEX\_fat entre 20 a 35, as demais categorias de “age” (de 45 a 55 e acima de 65) tendem a apresentar um índice DEX\_fat acima de 35, contudo estes dois últimos padrões são não de fácil visualização.

	DEXfat < 20	20 <= DEXfat < 35	DEXfat >= 35
age < 35	7	5	1
35 <= age < 45	2	5	4
45 <= age < 55	2	6	6
55 <= age < 65	1	15	13
age >= 65	1	4	4

DEXfat < 20	20 <= DEXfat < 35	DEXfat >= 35
0.54	0.38	0.08
0.18	0.45	0.36
0.14	0.43	0.43
0.03	0.52	0.45
0.11	0.44	0.44

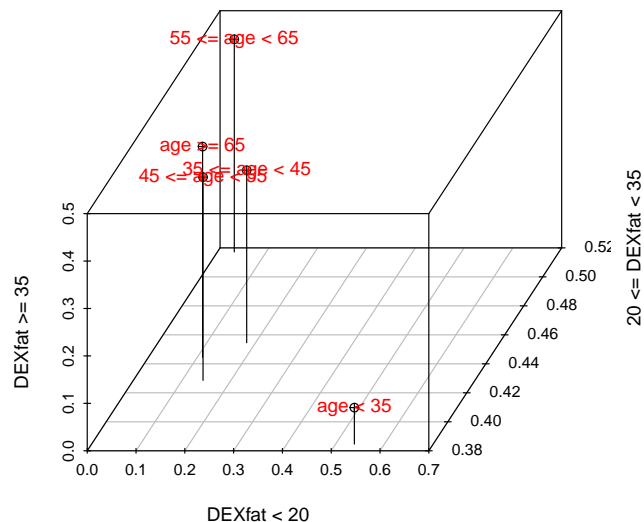
DEXfat < 20	20 <= DEXfat < 35	DEXfat >= 35
0.54	0.14	0.04
0.15	0.14	0.14
0.15	0.17	0.21
0.08	0.43	0.46
0.08	0.11	0.14

Teste para as populações:

H0: Populações homogêneas

H1: Populações heterogêneas

```
>
> Pearson's Chi-squared test
>
> data: tab2.4
> X-squared = 17.865, df = 8, p-value = 0.02226
```



```

>
> Principal inertias (eigenvalues):
>      1      2
> Value  0.232172 0.002889
> Percentage 98.77% 1.23%
>
>
> Rows:
>      age < 35  35 <= age < 45  45 <= age < 55  55 <= age < 65  age >= 65
> Mass      0.171053      0.144737      0.184211      0.381579  0.118421
> ChiDist   1.016030      0.028590      0.129172      0.365146  0.193013
> Inertia   0.176581      0.000118      0.003074      0.050877  0.004412
> Dim. 1    -2.108022     -0.057362      0.206077      0.750155  0.377299
> Dim. 2    -0.456563      0.135978      1.537128     -0.963333  1.206270
>
>
> Columns:
>      DEXfat < 20  20 <= DEXfat < 35  DEXfat >= 35
> Mass      0.171053      0.460526      0.368421
> ChiDist   1.025928      0.107194      0.367402
> Inertia   0.180038      0.005292      0.049731
> Dim. 1    -2.128254      0.188032      0.753077
> Dim. 2     0.562753     -1.065867      1.071056

```

