## Trabajo Práctico Nº 2

### Ejercicio 1.

El archivo "ine.dta" es una base de datos que contiene los gastos promedio, en euros, de los hogares españoles, por grandes rubros y comunidad autónoma, correspondientes a los relevamientos de la encuesta de presupuestos familiares del año 2005 realizada por el Instituto Nacional de Estadísticas de España (INE). La descripción de las variables y de la base puede consultarse en las etiquetas asociadas a cada una.

### (a) Realizar un análisis descriptivo de los datos.

Va	riable			Obs	M	lean	Std.	dev.	N	Min	Max	X
v vi	alybnh estcal vagelo bymant			18 18 18 18	4200. 1751 7863. 1131.	.38 757 067	389. 1846 2	.833 05.8	3406. 1244. 4833. 724.	. 85 . 29 . 82	6561.2° 2792.2° 12291.9° 1500.5°	5 4 6
	salud	 -+		18 	570.0		151.	5053 	302.	.85 	942.2	_
	transp	i I		18 18	2561. 678.1		411. 97.7		1680. 496.		3148. 897.1	
	ocio	i		18			352.		746.		2080.1	
	educ	i		18			98.8		87.		438.9	
	esparc			18	2178.		349.		1670.		2840.5	
		-+										_
	otros	i		18	1510.	961	263.	4181	898	3.3	1925.8	9
Stats	alybnh	vestcal			salud				educ	esparc	otros	
Variance	4200.572 507861.5 .1696541	151770	7863.757 3410793	1131.067 42353.62	22953.85	2561.397 169101.5	9562.441	1485.072 124215.1	9773.491 12	1839.1 6	9389.08	
	a.	ybnh	vestcal ·	vivagelo	mobymant	salud	transp	comu	ocio	edu	c esparc	otros
alv	bnh   1	.0000										
			1.0000									
	relo   0		1.0000									
-	ant   0		0.2817	1.0000								
		4255	0.2817 0.5484	0.6729	1.0000							
	lud   0	.4255 .5876	0.2817 0.5484 0.5515	0.6729 0.3730	0.4359							
	lud   0	.4255 .5876 .1807	0.2817 0.5484 0.5515 -0.0512	0.6729 0.3730 0.2390	0.4359 0.3747	0.2431						
C	lud   0 nsp   0 omu   0	.4255 .5876 .1807 .4905	0.2817 0.5484 0.5515 -0.0512 0.4005	0.6729 0.3730 0.2390 0.6784	0.4359 0.3747 0.7074	0.2431 0.6655	0.5336	1.0000				
0	lud   0 nsp   0 comu   0	.4255 .5876 .1807 .4905 .3118	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689	0.6729 0.3730 0.2390 0.6784 0.8845	0.4359 0.3747 0.7074 0.7746	0.2431 0.6655 0.5416	0.5336 0.2545	1.0000 0.7954	1.0000	1 000	0	
o e	lud   0   nsp   0   comu   0   cio   0   duc   0	.4255 .5876 .1807 .4905 .3118	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761	0.4359 0.3747 0.7074 0.7746 0.6143	0.2431 0.6655 0.5416 0.3818	0.5336 0.2545 0.1906	1.0000 0.7954 0.6814	1.0000 0.8939			
o e esp	lud   0	.4255 .5876 .1807 .4905 .3118	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366	0.6729 0.3730 0.2390 0.6784 0.8845	0.4359 0.3747 0.7074 0.7746 0.6143	0.2431 0.6655 0.5416 0.3818 0.5586	0.5336 0.2545 0.1906 0.2354	1.0000 0.7954 0.6814 0.7587	1.0000 0.8939 0.8548		5 1.0000	1.0000
o e esp	1   1   0   0   1   1   0   0   1   0   0	.4255 .5876 .1807 .4905 .3118 .1116 .3450 .2994	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632	0.5336 0.2545 0.1906 0.2354 0.4942	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548	0.779 0.572	5 1.0000 9 0.4650	1.0000 otros
c o e esp ot	llud   0	4255 5876 1807 4905 3118 1116 3450 2994	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632	0.5336 0.2545 0.1906 0.2354 0.4942	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
esp ot aly	1ud   0   0   0   0   0   0   0   0   0	.4255 .5876 .1807 .4905 .3118 .1116 .3450 .2994	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632	0.5336 0.2545 0.1906 0.2354 0.4942	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
c o e esp ot aly vest	1ud   0   0   0   0   0   0   0   0   0	4255 .5876 1807 4905 3118 1116 .3450 .2994 ybnh  07862 80526	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632	0.5336 0.2545 0.1906 0.2354 0.4942	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
esp ot  aly vest vivag	1ud   0   0   0   0   0   0   0   0   0	4255 5876 1807 4905 3118 1116 3450 2994 	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766 vestcal	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744 vivagelo	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632	0.5336 0.2545 0.1906 0.2354 0.4942	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
eesp ot aly vest vivag mobym	1.1ud   0   0.nsp   0   0.omu   0   0.oto	4255 5876 1807 4905 33118 1116 3450 2994 	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766 vestcal 	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744 vivagelo	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579 mobymant	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632	0.5336 0.2545 0.1906 0.2354 0.4942	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
esp ot aly vest vivag mobym sa	Delta	4255 5876 1807 4905 3118 1116 3450 2994  77862 00526 11455 106.2	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766 vestcal  151770 202659 43970.8 32552.6	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744 vivagelo	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579 mobymant 	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632 salud	0.5336 0.2545 0.1906 0.2354 0.4942	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
esp ot aly vest vivag mobym sa tra	One   One	4255 5876 1807 4905 3118 1116 33450 22994 	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766 vestcal 	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744 vivagelo 	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579 mobymant 	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632 salud	0.5336 0.2545 0.1906 0.2354 0.4942 transp	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
eesp ot aly vest vivag mobym sa tra	Llud   0   0   0   0   0   0   0   0   0	4255 5876 1807 4905 3118 1116 3450 2994 	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766 vestcal 	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744 vivagelo 	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579 mobymant 	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632 salud 22953.8 15144.9 9859.84	0.5336 0.2545 0.1906 0.2354 0.4942 transp	1.0000 0.7954 0.6814 0.7587 0.6145	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650	
eesp ot aly vest vivag mobym sa tra	Llud   0   0   0   0   0   0   0   0   0	4255 5876 1807 4905 3118 1116 3450 2994 ybnh  17862 00526 11455 106.2 140.3 53.1 -	0.2817 0.5484 0.5515 -0.0512 0.4005 0.3366 0.4988 0.3766 vestcal 	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744 vivagelo 	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579 mobymant 42353.6 13591.6 31707.6 14235.9 56184.1	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632 salud 	0.5336 0.2545 0.1906 0.2354 0.4942 transp	1.0000 0.7954 0.6814 0.7587 0.6145 comu	1.0000 0.8939 0.8548 0.6348	0.779 0.572	5 1.0000 9 0.4650 C esparc	
esp ot aly vest vivag mobym sa tra	One   One	4255 5876 1807 4905 3118 1116 3450 2994 2994 290526 11455 106.2 10526 83.7 83.7 83.7 83.7	0.2817 0.5484 0.5515 -0.0512 0.4005 0.4689 0.3366 0.4988 0.3766 vestcal 	0.6729 0.3730 0.2390 0.6784 0.8845 0.8761 0.8082 0.5744 vivagelo 	0.4359 0.3747 0.7074 0.7746 0.6143 0.7159 0.6579 mobymant 	0.2431 0.6655 0.5416 0.3818 0.5586 0.2632 salud 22953.8 15144.9 9859.84 28917.1 579.04 29541.2	0.5336 0.2545 0.1906 0.2354 0.4942 transp 169101 21457.3 36890.3 7749.14	1.0000 0.7954 0.6814 0.7587 0.6145 comu 	1.0000 0.8939 0.8548 0.6348	0.779 0.572 edu 9773.4 26897.	5 1.0000 9 0.4650 C esparc	otros

#### Maestría en Econometría UTDT - Análisis Estadístico Multivariado | 2

#### Juan Menduiña

Variable	Partial corr.	Semipartial corr.	Partial corr.^2	Semipartial corr.^2	Significance value
vestcal	0.8688	0.5915	0.7549	0.3498	0.0024
vivagelo	0.1226	0.0416	0.0150	0.0017	0.7534
mobymant	-0.2715	-0.0951	0.0737	0.0090	0.4798
salud	-0.1229	-0.0417	0.0151	0.0017	0.7528
transp	0.2801	0.0983	0.0785	0.0097	0.4654
comu	0.6013	0.2536	0.3615	0.0643	0.0868
ocio	0.0837	0.0283	0.0070	0.0008	0.8305
educ	-0.4532	-0.1714	0.2054	0.0294	0.2205
esparc	-0.2692	-0.0942	0.0725	0.0089	0.4837
otros	-0.2278	-0.0788	0.0519	0.0062	0.5556

### (b) Realizar un análisis de componentes principales.

### Basado en matriz de correlaciones:

Principal components/correlation

Number of obs = 18

Number of comp. = 11

Trace = 11

Rotation: (unrotated = principal)

Rho = 1.0000

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	6.31799	4.63452	0.5744	0.5744
Comp2	1.68348	.518886	0.1530	0.7274
Comp3	1.16459	.4527	0.1059	0.8333
Comp4	.711894	.376163	0.0647	0.8980
Comp5	.33573	.0888713	0.0305	0.9285
Comp6	.246859	.0686718	0.0224	0.951
Comp7	.178187	.0371071	0.0162	0.967
Comp8	.14108	.0155581	0.0128	0.980
Comp9	.125522	.0650827	0.0114	0.991
Comp10	.0604393	.0262194	0.0055	0.996
Comp11	.0342199		0.0031	1.000

Principal components (eigenvectors)

Variable   alybnh	Comp1 0.2073 0.2441	Comp2 0.6146 0.5185	Comp3 0.1074	Comp4 0.0974	Comp5	Comp6	Comp7	Comp8	Comp9	Comp10	Comp11	Unexplained
. 2			0.1074	0 0974								
					-0.0291	0.4075	-0.2469	0.1482	0.2589	-0.1293	0.4740	0
vestcal			-0.2140	0.3225	0.0077	0.1046	0.3670	0.0381	-0.1907	0.1589	-0.5595	0
vivagelo	0.3315	-0.3155	-0.2008	-0.0142	-0.0053	0.1415	0.1154	0.1745	0.7708	0.2617	-0.1536	0
mobymant	0.3387	-0.0037	0.0613	0.2925	-0.5949	-0.5657	-0.1077	0.2229	-0.0724	0.1447	0.1859	0
salud	0.2635	0.3311	0.0166	-0.6080	0.3392	-0.5052	0.2374	0.0221	0.0754	0.0557	0.1226	0
transp	0.1628	-0.1258	0.7924	-0.1381	-0.1746	0.2416	0.4188	0.1691	-0.0608	-0.0674	-0.0976	0
comu	0.3549	-0.0061	0.2045	-0.2695	0.0288	0.1126	-0.6905	-0.1208	-0.1672	0.2743	-0.3892	0
ocio	0.3716	-0.1529	-0.1714	-0.0178	0.0565	-0.0395	-0.0991	0.2073	-0.0471	-0.8409	-0.2046	0
educ	0.3285	-0.2921	-0.2482	-0.0056	0.2639	0.2693	0.1127	0.3735	-0.4946	0.2721	0.3634	0
esparc	0.3507	-0.0735	-0.2080	-0.1840	-0.3619	0.2056	0.2194	-0.7228	-0.0861	-0.0665	0.1970	0
otros	0.2863	-0.1195	0.3072	0.5524	0.5432	-0.2004	-0.0105	-0.3845	0.0748	-0.0081	0.1241	0

## Basado en matriz de varianzas y covarianzas:

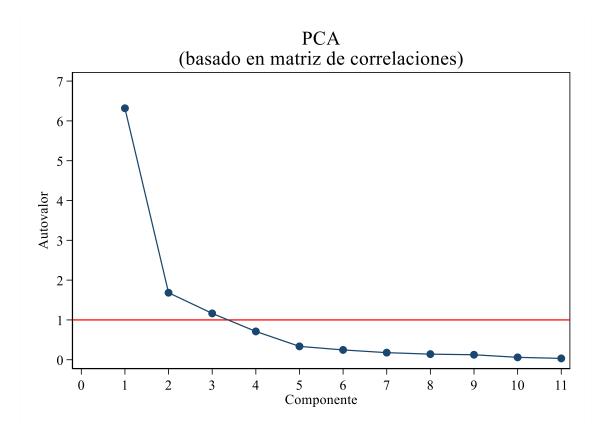
Principal components/covariance	Number of obs	=	18
	Number of comp.	=	11
	Trace	=	4639612
Rotation: (unrotated = principal)	Rho	=	1.0000

Component	   	Eigenvalue	Difference	Proportion	Cumulative
Comp1		3687478	3054180	0.7948	0.7948
Comp2		633298	456210	0.1365	0.9313
Comp3		177088	130083	0.0382	0.9694
Comp4		47004.7	5090.15	0.0101	0.9796
Comp5		41914.5	23807.5	0.0090	0.9886
Comp6		18107	5406.33	0.0039	0.9925
Comp7		12700.7	510.343	0.0027	0.9953
Comp8		12190.3	4652.97	0.0026	0.9979
Comp9		7537.36	6015.16	0.0016	0.9995
Comp10		1522.2	750.196	0.0003	0.9998
Comp11		772.001	•	0.0002	1.0000

Principal components (eigenvectors)

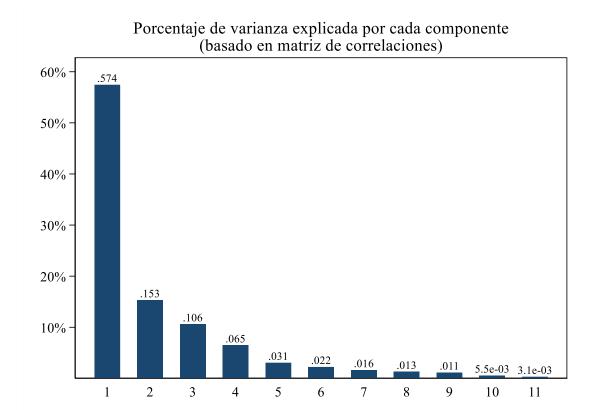
Variable	Comp1	Comp2	Comp3	Comp4	Comp5	Comp6	Comp7	Comp8	Comp9	Comp10	Comp11	Unexplained
alybnh	0.0638	0.8748	0.0036	-0.3835	-0.0667	0.1843	-0.1913	-0.0431	-0.0210	-0.0181	0.0771	0
vestcal	0.0674	0.4068	-0.2632	0.4366	-0.1017	-0.5842	0.4014	0.1016	-0.1445	0.0458	-0.1550	0
vivagelo	0.9596	-0.1403	-0.0612	-0.1929	-0.1053	-0.0776	0.0174	0.0039	0.0315	-0.0062	-0.0122	0
mobymant	0.0756	0.0907	0.0793	0.3368	0.0122	-0.0716	-0.1282	-0.8100	0.4121	0.0663	0.1192	0
salud	0.0325	0.1034	0.0181	-0.0021	0.2610	0.2465	0.5642	0.2636	0.6688	-0.0669	0.1393	0
transp	0.0559	0.0724	0.9261	-0.0395	0.1198	-0.2521	0.1785	0.0056	-0.1481	-0.0271	-0.0018	0
comu	0.0363	0.0463	0.0764	0.0249	0.0980	0.2043	-0.0221	0.0149	0.1007	0.6585	-0.7031	0
ocio	0.1662	0.0812	0.0016	0.3749	0.1859	0.6086	0.2876	-0.1949	-0.4906	-0.2321	-0.0707	0
educ	0.0455	-0.0020	-0.0093	0.0953	0.0272	0.0785	0.0867	0.0706	-0.2086	0.7023	0.6578	0
esparc	0.1514	0.1031	-0.0276	0.3115	0.7140	-0.1199	-0.5073	0.2790	0.0444	-0.0676	0.0522	0
otros	0.0823	0.0777	0.2368	0.5126	-0.5803	0.2381	-0.2953	0.3757	0.2055	-0.0525	0.0391	0

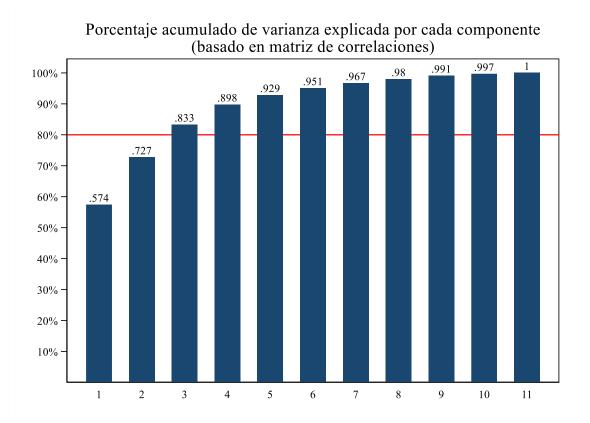
- **(c)** Efectuar la selección de los componentes principales de acuerdo con los siguientes modos:
- (i) Búsqueda del "codo".



Mediante este modo, se seleccionarían los tres primeros componentes principales.

(ii) Búsqueda por umbral de varianza a explicar, considerando un 80%.





Mediante este modo, se seleccionarían los tres primeros componentes principales.

(iii) Búsqueda por tope mínimo al valor de los eigenvalores, considerando la varianza media.

Mediante este modo, se seleccionarían los tres primeros componentes principales, ya que estos tienen un autovalor mayor a 1, correspondiente a la varianza media.

(d) Interpretar los componentes seleccionados. Para ello, se puede emplear el archivo "renta.csv". Aclaración: El archivo "renta.csv" contiene información referida al ingreso promedio por hogar, por comunidad; relevado por el INE para el año 2005.

Principal componed Rotation: (un	nts/correlation rotated = princ	ipal)	Number of obs Number of comp. Trace Rho	= 18 = 3 = 11 = 0.8333
Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	6.31799	4.63452	0.5744	0.5744
Comp2	1.68348	.518886	0.1530	0.7274
Comp3	1.16459	.4527	0.1059	0.8333
Comp4	.711894	.376163	0.0647	0.8980
Comp5	.33573	.0888713	0.0305	0.9285
Comp6	.246859	.0686718	0.0224	0.9510
Comp7	.178187	.0371071	0.0162	0.9672
Comp8	.14108	.0155581	0.0128	0.9800
Comp9	.125522	.0650827	0.0114	0.9914
Comp10	.0604393	.0262194	0.0055	0.9969
Comp11	.0342199	•	0.0031	1.0000

Principal components (eigenvectors)

alybnh   0.2073	Variable		 Comp1	Comp2	Comp3		Unexplained
	vestcal vivagelo mobymant salud transp comu ocio educ esparc	-+-               	0.2441 0.3315 0.3387 0.2635 0.1628 0.3549 0.3716 0.3285 0.3507	0.5185 -0.3155 -0.0037 0.3311 -0.1258 -0.0061 -0.1529 -0.2921 -0.0735	-0.2140 -0.2008 0.0613 0.0166 0.7924 0.2045 -0.1714 -0.2482 -0.2080	-+-           	.1178 .0913 .2709 .3763 .07458 .1554 .05395 .103 .1633

El primer autovector pondera con signo positivo a todas las variables, por lo que el primer componente puede ser interpretado como una medida global de gasto.

El segundo autovector pondera con un signo positivo a la primera, segunda y quinta variable y con signo negativo a las restantes, por lo que el segundo componente tomará

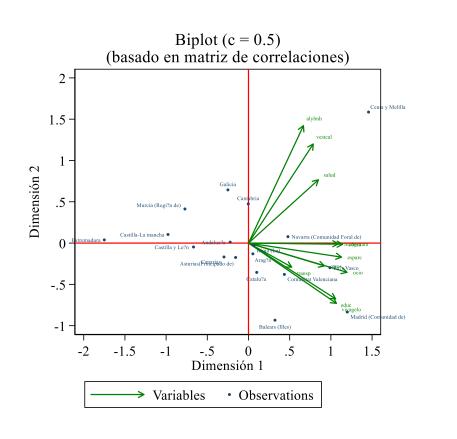
Juan Menduiña

valores altos en aquellas comunidades cuyos gastos de necesidad primaria (alimentos y bebidas no alcohólicas, artículos de vestir y calzado, y salud) resulten más importantes, en términos relativos, a los restantes.

El tercer autovector pondera con un signo negativo a la segunda, tercera, octava, novena y décima variable y con signo positivo a las restantes, por lo que el tercer componente tomará valores altos en aquellas comunidades cuyos gastos en transporte, en comunicaciones y otros resulten más importantes, en términos relativos, a los restantes.

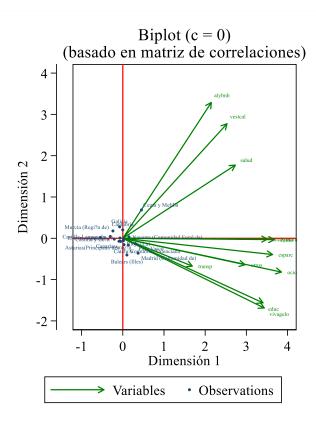
# Ejercicio 2.

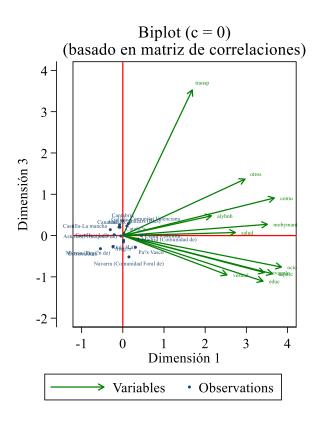
En relación al ejercicio anterior, realizar el biplot asociado a las dos dimensiones principales, considerando un parámetro c=0,5.



## Ejercicio 3.

Teniendo en cuenta la misma base de datos del primer punto, efectuar los biplots correspondientes a las dos dimensiones principales y a la primera y tercera, considerando, en ambos casos, un parámetro c=0. Comentar, brevemente, el resultado en relación a lo obtenido mediante el análisis de componentes principales efectuado en el primer ejercicio.





## Ejercicio 4.

La base "wb.dta" contiene información relevada por el Banco Mundial en 77 países durante el año 2007 (los detalles de las variables relevadas se pueden observar en las etiquetas asignadas a las mismas). A partir de estos datos, se pide:

(a) Realizar un análisis normado de componentes principales a partir de los datos.

### Basado en matriz de correlaciones:

Principal components/correlation	Number of obs	=	77
	Number of comp.	=	9
	Trace	=	9
Rotation: (unrotated = principal)	Rho	=	1.0000

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1   Comp2   Comp3   Comp4	3.76621 2.62649 .883587	1.13972 1.7429 .217687 .115556	0.4185 0.2918 0.0982 0.0740	0.4185 0.7103 0.8085 0.8825
Comp5   Comp6   Comp7	.550344 .230906 .178371	.319438 .0525345 .0821595	0.0611 0.0257 0.0198	0.9436 0.9693 0.9891
Comp8   Comp9	.0962117	.0942351	0.0107 0.0002	0.9998 1.0000

Principal components (eigenvectors)

Unexplained	Comp9	Comp8	Comp7	Comp6	Comp5	Comp4	Comp3	Comp2	Comp1	Variable
0	-0.0050	0.0780	-0.2748	0.2654	0.6232	-0.3706	0.3471	-0.4207	0.1573	var4
0	0.0276	0.7690	0.0536	-0.3818	0.0836	-0.0344	-0.0045	0.1900	0.4636	var5
0	-0.0016	0.0585	-0.1262	0.0165	-0.1591	0.3210	0.8765	0.2650	-0.1164	var6
0	0.0077	0.0775	0.6789	0.2174	0.0303	0.4485	0.1023	-0.4836	0.1984	var7
0	-0.7216	-0.0804	-0.0934	0.4218	-0.1126	0.0394	-0.0633	0.2387	0.4593	var8
0	0.0093	0.0504	0.6188	0.2391	0.1469	-0.5155	0.1307	0.4571	-0.2097	var9
0	0.0086	-0.6102	0.1922	-0.5361	0.2079	-0.0692	0.1710	0.1062	0.4634	var10
0	0.6916	-0.1077	-0.1196	0.4577	-0.1291	0.0375	-0.0641	0.2327	0.4583	varl1
0	0.0075	-0.0136	-0.0441	0.0654	0.6959	0.5327	-0.2151	0.3892	-0.1666	var12

### Basado en matriz de varianzas y covarianzas:

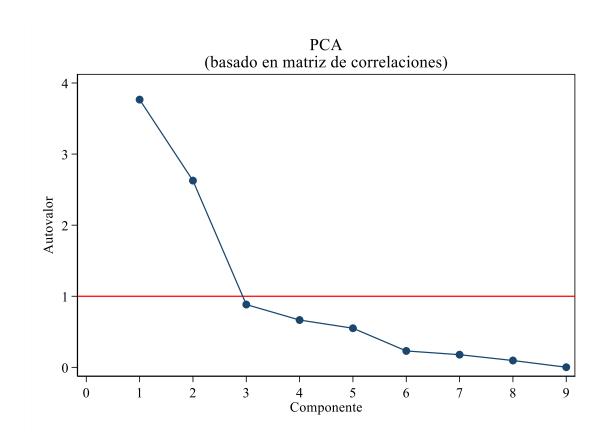
Principal components/covariance	Number of obs	=	77
	Number of comp.	=	4
	Trace	=	3.43e+16
Rotation: (unrotated = principal)	Rho	=	1.0000

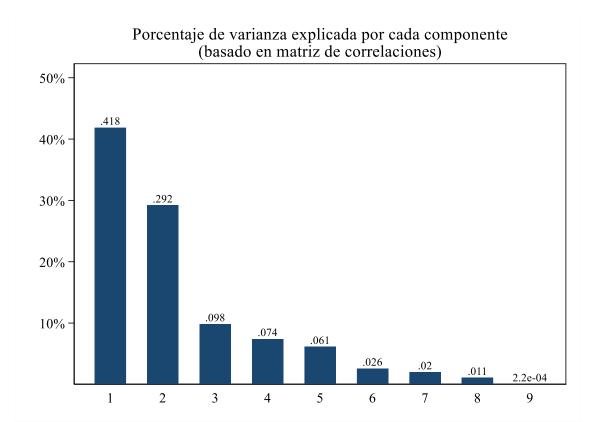
Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	3.37471e+16	3.32396e+16	0.9837	0.9837
Comp2	5.07523e+14	4.75686e+14	0.0148	0.9985
Comp3	3.18369e+13	1.13473e+13	0.0009	0.9994
Comp4	2.04896e+13	2.04896e+13	0.0006	1.0000
Comp5	0	0	0.0000	1.0000
Comp6	0	0	0.0000	1.0000
Comp7	0	0	0.0000	1.0000
Comp8	0	0	0.0000	1.0000
Comp9	0		0.0000	1.0000

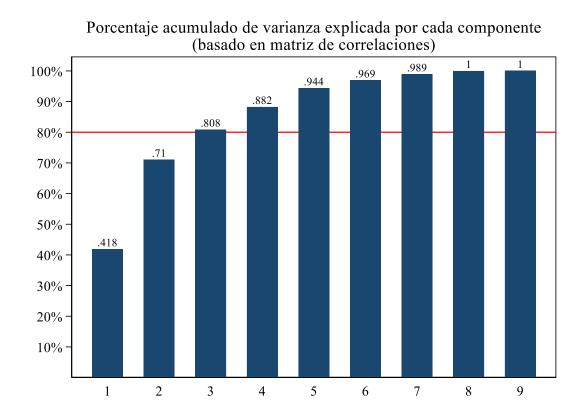
Principal components (eigenvectors)

Variable	  -	Comp1	Comp2	Comp3	Comp4		Unexplained
var4 var5 var6 var7 var8 var9 var10 var11	+-	-0.0000 0.0625 -0.0000 0.0000 0.8535 -0.0000 0.1607 0.4917 -0.0000	0.0003 0.1785 0.0000 0.0000 -0.1211 -0.0000 0.9679 -0.1287 -0.0000	-0.0005 0.5969 -0.0000 -0.0000 0.3820 0.0000 -0.1538 -0.6886 0.0000	0.0001 0.7797 -0.0000 -0.0000 -0.3331 -0.0000 -0.1167 0.5172 -0.0000	-+-           	199727543 6.625 45.19 88.41 -4 1.515 27.25 35 15.35

### (b) ¿Cuántos componentes se sugiere extraer?







Por lo tanto, se sugiere extraer los tres primeros componentes principales.

(c) ¿Cuál es el porcentaje de variabilidad total explicado por las componentes seleccionadas?

El porcentaje de variabilidad total explicado por las componentes seleccionadas es 80,85%.

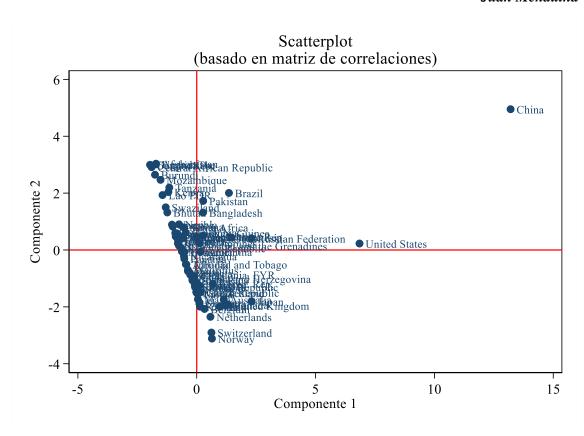
(d) ¿Qué interpretación se sugiere de las componentes, con arreglo a las correlaciones existentes con las variables originales? Para este punto, se puede ayudar con la estructura de análisis presente en el libro "P2.xlsx".

El primer autovector pondera con signo positivo a todas las variables excepto a "Real Interest Rate", "Life Expectancy at Birth" y "MFN Tariff Rate", por lo que el primer componente principal tomará valores bajos en aquellos países cuyos valores en estas variables resulten más importantes, en términos relativos, a los restantes.

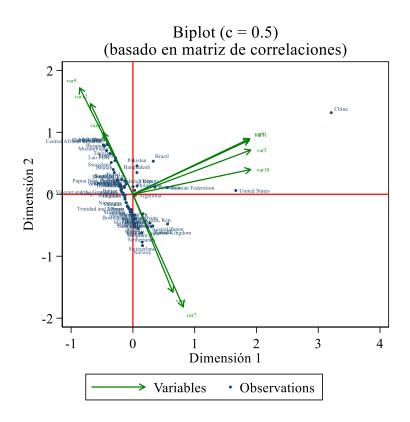
El segundo autovector pondera con signo positivo a todas las variables excepto a "GNI per cápita" y "Life Expectancy at Birth", por lo que el segundo componente principal tomará valores bajos en aquellos países cuyos valores en estas variables resulten más importantes, en términos relativos, a los restantes.

El tercer autovector pondera con signo positivo a todas las variables excepto a "Population", "Labor Force" y "MFN Tariff Rate", por lo que el tercer componente principal tomará valores bajos en aquellos países cuyos valores en estas variables resulten más importantes, en términos relativos, a los restantes.

(e) Clasificar a los países de acuerdo con las dos componentes principales.



(f) Realice un biplot considerando las dos dimensiones principales de análisis para un parámetro c=0.5. ¿Qué se puede decir sobre la posición de los Estados Unidos y China y su influencia en el análisis, a la luz de los resultados obtenidos?



A la luz de los resultados obtenidos, lo que se puede decir sobre la posición de los Estados Unidos y China es que ambos tienen, respecto al resto de los países, un alto valor del componente 1 y China, además, del componente 2, reflejando, por ejemplo, en el primer caso, el alto nivel de GNI per cápita y, en el segundo caso, el alto nivel de población.

## Ejercicio 5.

La base "hspendusa2009.csv" contiene información relevada durante los años 2008-2009 por el Instituto de Estadísticas de los Estados Unidos, correspondiente a los gastos medios de los hogares por capítulos y por área metropolitana (los detalles de las variables se pueden observar en el archivo "usa.xlsx"). Efectuar un análisis de componentes principales.

### Basado en matriz de correlaciones:

Principal components/correlation	Number of obs	=	18
	Number of comp.	=	12
	Trace	=	12
Rotation: (unrotated = principal)	Rho	=	1.0000

Component		Eigenvalue	Difference	Proportion	Cumulative
Comp1		5.75656	3.47503	0.4797	0.479
Comp2	İ	2.28152	.811626	0.1901	0.669
Comp3		1.4699	.560801	0.1225	0.792
Comp4		.909096	.345173	0.0758	0.868
Comp5		.563923	.185157	0.0470	0.915
Comp6		.378766	.0882703	0.0316	0.946
Comp7		.290495	.154762	0.0242	0.970
Comp8		.135733	.0151273	0.0113	0.982
Comp9		.120606	.065335	0.0101	0.992
Comp10		.0552706	.0325018	0.0046	0.996
Comp11		.0227689	.00740786	0.0019	0.998
Comp12	1	.015361		0.0013	1.000

Principal	components	(eigenvectors)

Unexplain	Comp12	Comp11	Comp10	Comp9	Comp8	Comp7	Comp6	Comp5	Comp4	Comp3	Comp2	Comp1	Variable
1	-0.1932	0.0672	-0.0399	-0.1863	-0.2547	-0.5954	-0.4483	-0.0847	0.4039	0.1179	0.1082	0.3249	s1
1	-0.2054	0.2057	0.2614	-0.2384	0.5530	0.2928	-0.2630	-0.4197	-0.0526	-0.0891	-0.1079	0.3622	s2
1	0.6790	0.0173	0.1808	-0.2400	0.0835	-0.1371	0.2124	0.0321	0.0860	0.4079	-0.3628	0.2664	s3
1	0.0702	0.1525	-0.0194	0.5343	-0.4145	0.4060	-0.3517	0.0199	0.0198	0.0866	-0.3461	0.3125	84
1	0.0885	-0.4413	-0.0157	-0.1494	0.1939	0.1522	-0.2134	0.6302	0.1789	-0.3934	-0.0343	0.2889	s5
1	0.0592	0.0680	-0.2671	-0.5335	-0.4110	0.2830	-0.0027	0.0120	-0.4400	0.0667	0.3189	0.2949	s6
1	-0.1083	-0.2872	0.6132	0.2105	-0.2149	-0.2605	0.2392	-0.0746	-0.3747	-0.1814	0.0996	0.3543	87
1	-0.2739	0.3941	-0.3629	0.1111	0.2117	-0.3263	0.2489	0.3305	-0.3259	-0.0560	-0.3184	0.3040	s8
1	0.1427	-0.4322	-0.5086	0.3488	0.2557	-0.0393	0.1077	-0.3529	0.0110	0.1281	0.2693	0.3469	s9
1	-0.4352	-0.0253	0.1770	0.0344	0.0633	0.2699	0.2879	0.3044	0.2782	0.6064	0.2279	0.1669	s10
1	0.3828	0.5103	0.1527	0.2712	0.1823	-0.0393	-0.0972	0.2171	0.0189	-0.0966	0.6215	0.0943	s11
1	0.0002	0.2111	-0.0668	-0.0736	-0.2244	0.1506	0.5408	-0.1972	0.5255	-0.4642	-0.0180	0.2219	s12

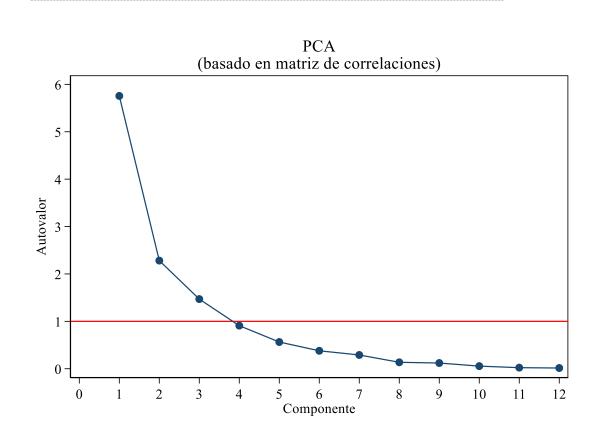
## Basado en matriz de varianzas y covarianzas:

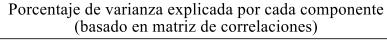
Principal components/covariance Number of obs = 18 Number of comp. = 12 Trace = 1.11e+07 Rotation: (unrotated = principal) Rho = 1.0000

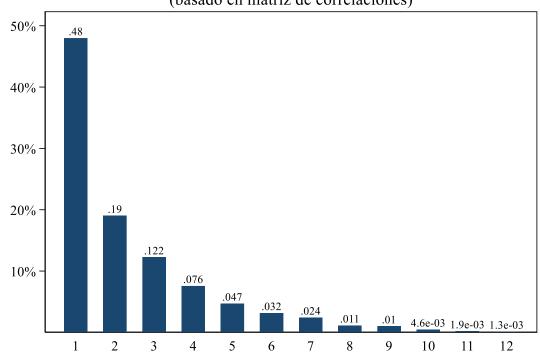
Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	8841316	7450143	0.7964	0.7964
Comp2	1391172	974398	0.1253	0.9217
Comp3	416774	174838	0.0375	0.9592
Comp4	241937	132052	0.0218	0.9810
Comp5	109885	63204.4	0.0099	0.9909
Comp6	46680.9	18101.4	0.0042	0.9951
Comp7	28579.5	9397.67	0.0026	0.9977
Comp8	19181.8	14531.8	0.0017	0.9994
Comp9	4649.94	3549.22	0.0004	0.9998
Comp10	1100.72	486.454	0.0001	0.9999
Comp11	614.268	497.431	0.0001	1.0000
Comp12	116.837	•	0.0000	1.0000

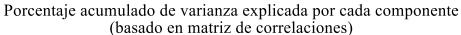
rincipal components (eigenvectors)

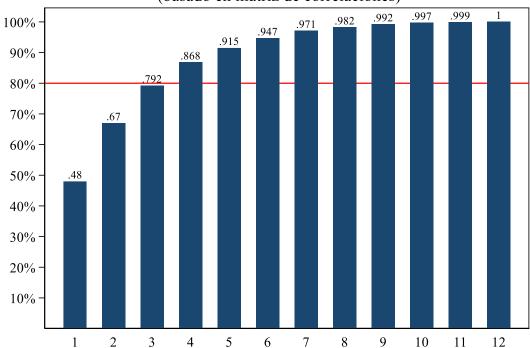
Unexplained	Comp12	Comp11	Comp10	Comp9	Comp8	Comp7	Comp6	Comp5	Comp4	Comp3	Comp2	Comp1	Variable
0	-0.0189	-0.0675	0.0351	-0.0775	-0.0759	-0.1722	0.1399	0.8776	0.2551	0.2331	0.1877	0.0895	s1
0	-0.1380	-0.1600	0.1149	0.9142	0.2168	0.1400	0.1349	0.0850	-0.0893	0.0374	0.0633	0.0307	s2
0	0.0031	0.0396	-0.0254	-0.0012	0.0402	-0.0827	-0.0301	-0.0350	-0.0449	-0.0750	-0.1659	0.9757	s3
0	-0.0133	0.0802	0.0199	-0.1596	-0.2176	0.7225	0.6025	0.0360	-0.1080	-0.0638	0.0965	0.0937	84
0	0.0228	-0.0121	-0.0464	0.0228	-0.0283	-0.1095	-0.0134	-0.1831	0.2253	-0.3679	0.8667	0.1136	s5
0	-0.0029	0.0067	0.0060	-0.1441	0.3776	-0.2943	0.4613	-0.2815	-0.1579	0.6171	0.2246	0.0416	s6
0	0.0006	-0.0589	-0.1254	0.0190	-0.3278	0.2103	-0.4759	0.0605	-0.5521	0.4250	0.3266	0.0808	87
0	-0.1535	-0.1625	0.9443	-0.1266	-0.1416	-0.0656	-0.0456	-0.0621	-0.0950	-0.0183	0.0420	0.0295	s8
0	0.9586	0.1325	0.2046	0.1234	0.0328	0.0434	-0.0110	0.0301	0.0006	0.0464	0.0158	0.0064	89
0	-0.0222	-0.1403	0.0338	0.0570	-0.1359	0.3131	-0.2078	-0.2626	0.7197	0.4695	-0.0240	0.0882	s10
0	-0.1846	0.9444	0.1460	0.1476	-0.0671	-0.0264	-0.0693	0.0156	0.0748	0.1118	0.0486	-0.0145	s11
0	-0.0530	0.0518	0.1027	-0.2341	0.7767	0.4173	-0.3295	0.1724	-0.0110	-0.0491	0.1035	0.0131	s12











Por lo tanto, se sugiere extraer los cuatro primeros componentes principales.

