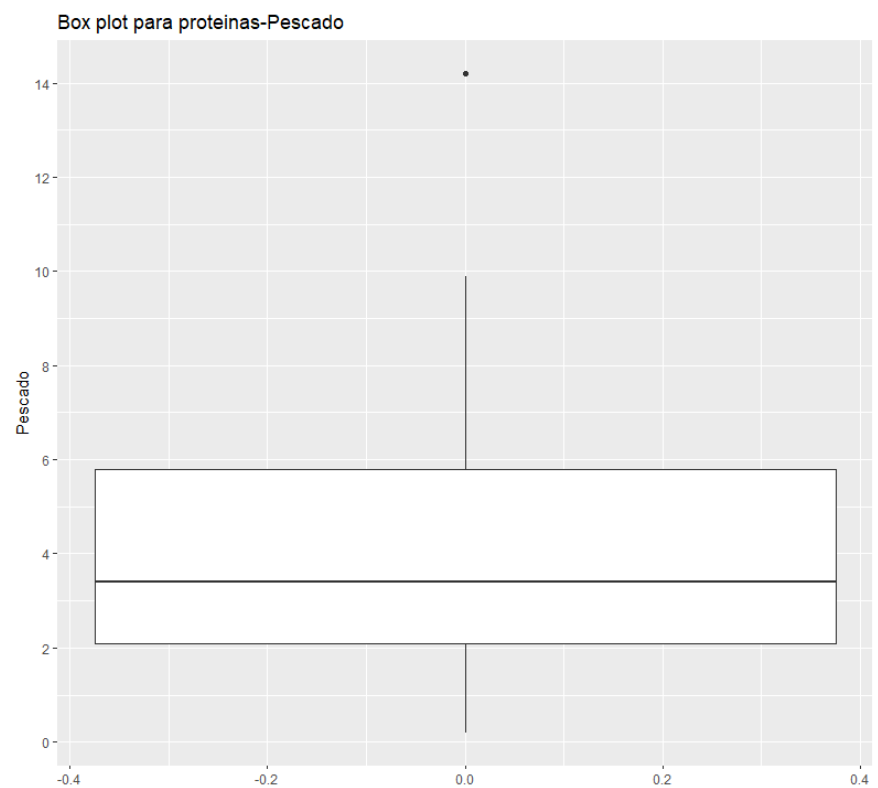
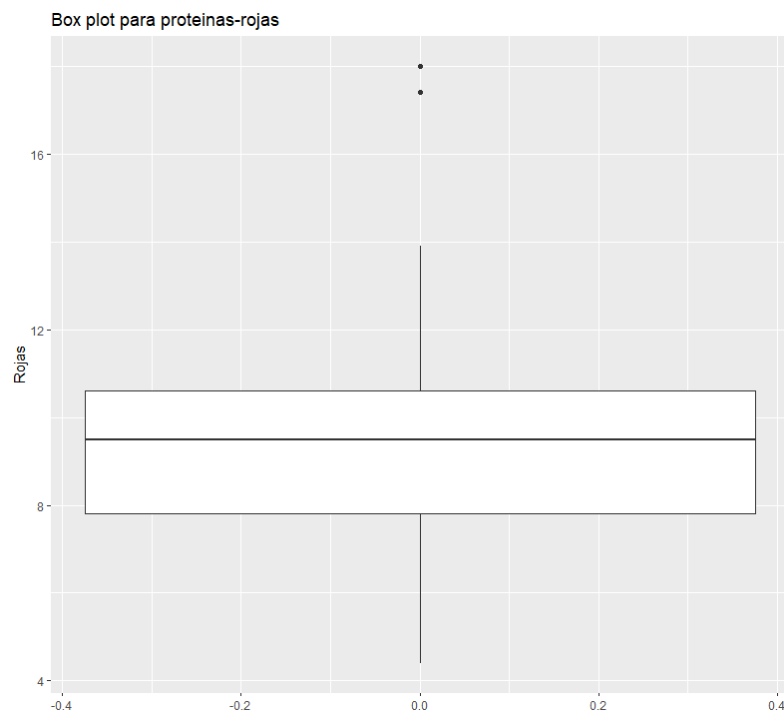


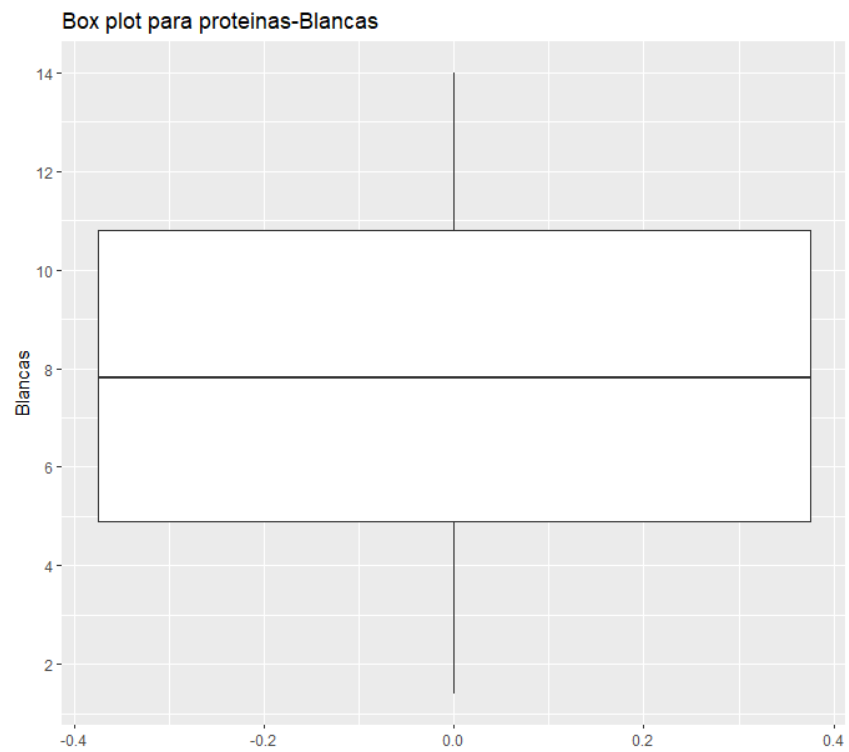
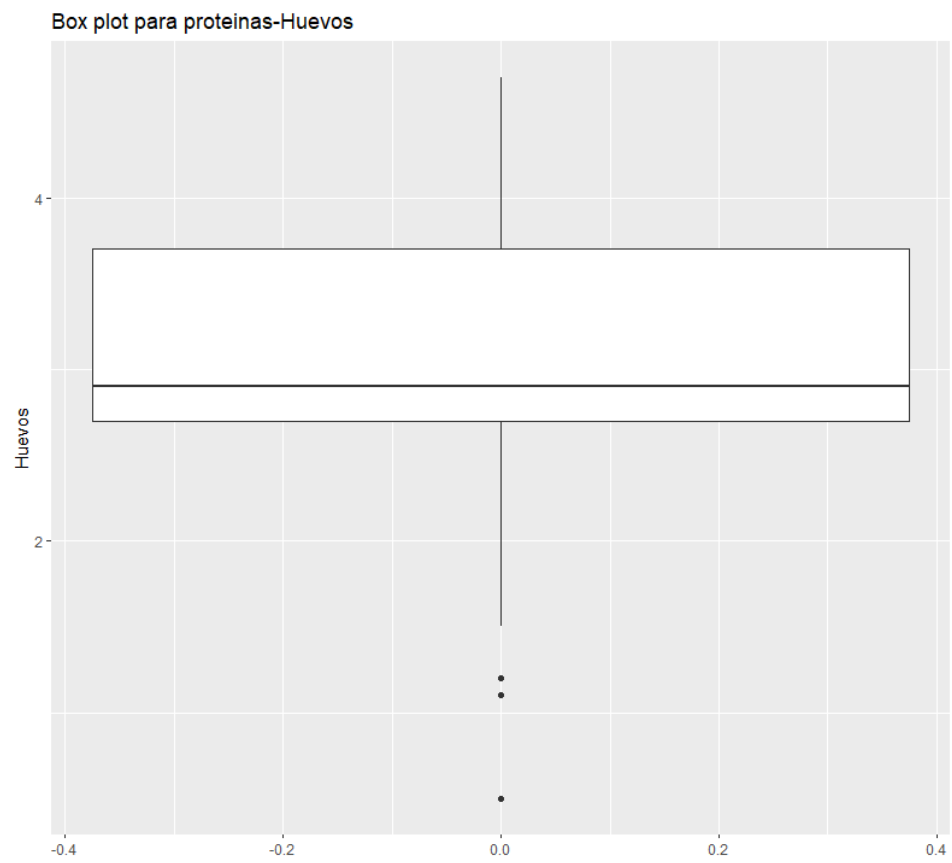
Alumno: Rodrigo Jeldres Carrasco

Control Multivariado

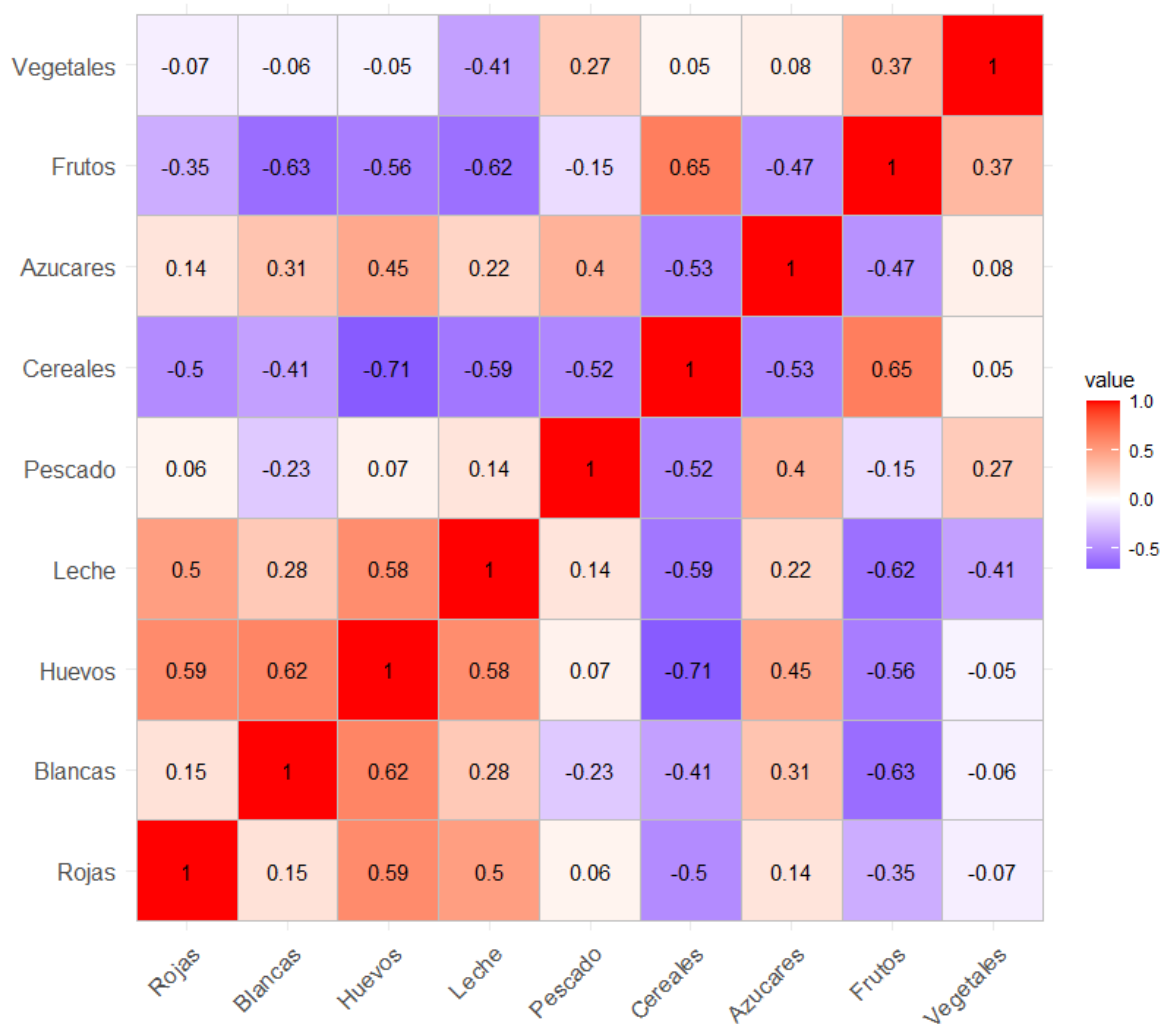
1. Describir las variables. Comente (comportamiento, anomalías, etc.)

Graficando los boxplot:





R: se visualizan algunos datos atípicos para las “proteínas”: rojas, pescado, huevos



R: Según el gráfico, se identifica que las proteínas de origen animal como Pescado, Leche, Huevos, Blancas y Rojas presentan una fuerte correlación entre ellas (en el grafico se observa valores entre 0.59, 0.5, 0.58, etc). Además, existe un grupo de proteínas vegetales compuesto por Frutos, Azúcares y Cereales que también muestran una alta asociación. Por último, otro grupo que correlaciona bien entre sí incluye las proteínas Pescado, Azúcares y Cereales.

2 a)

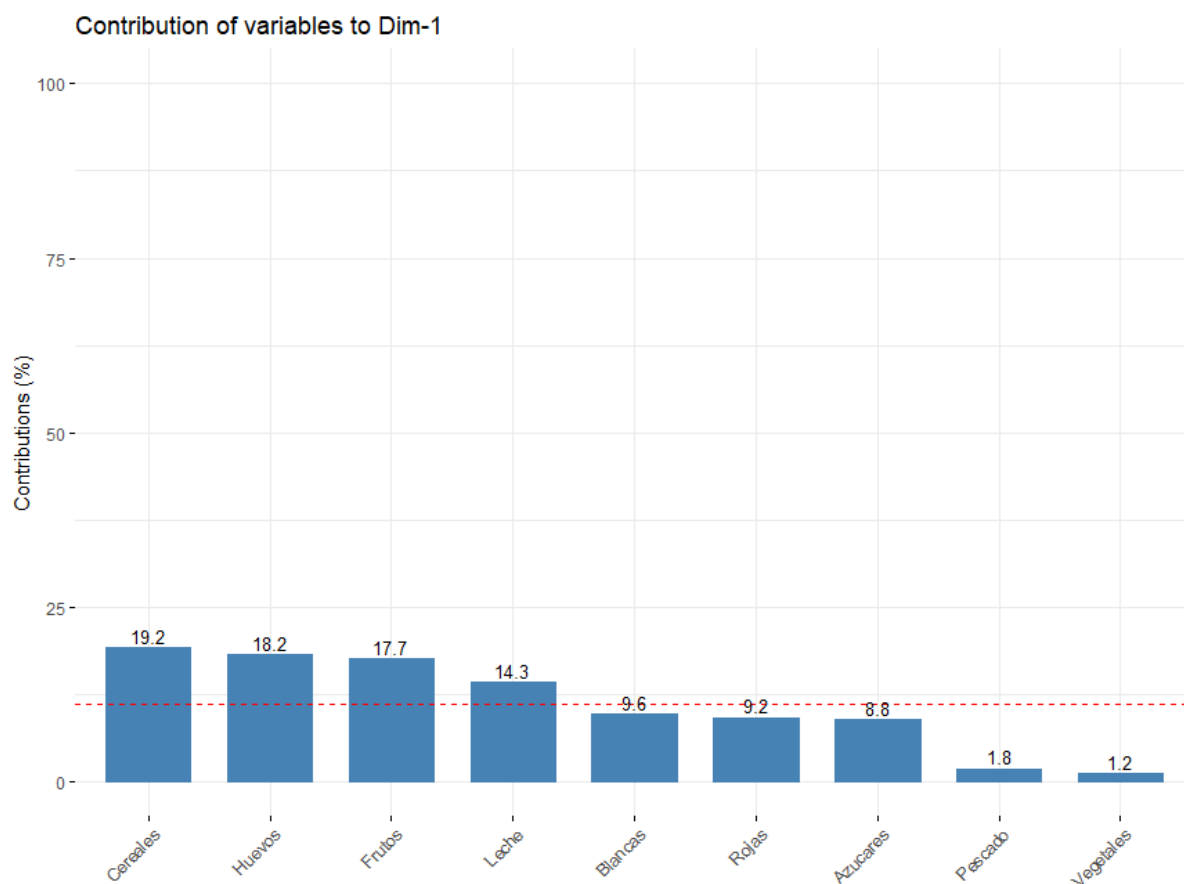
```
> summary(pca)
Importance of components:
               Comp.1  Comp.2  Comp.3  Comp.4  Comp.5  Comp.6
Standard deviation  2.0016087 1.2786710 1.0620355 0.9770691 0.6810568 0.57020257
Proportion of Variance 0.4451597 0.1816666 0.1253244 0.1060738 0.0515376 0.03612566
Cumulative Proportion 0.4451597 0.6268263 0.7521507 0.8582245 0.9097621 0.94588776
               Comp.7  Comp.8  Comp.9
Standard deviation  0.52115865 0.34101599 0.31482043
Proportion of Variance 0.03017848 0.01292132 0.01101243
Cumulative Proportion 0.97606624 0.98898757 1.00000000
```

Composición de las dimensiones:

	Comp.1	Comp.2
Rojas	0.3026094	0.05625165
Blancas	0.3105562	0.23685334
Huevos	0.4266785	0.03533576
Leche	0.3777273	0.18458877
Pescado	0.1356499	-0.64681970
Cereales	-0.4377434	0.23348508
Azucares	0.2972477	-0.35282564
Frutos	-0.4203344	-0.14331056
Vegetales	-0.1104199	-0.53619004

Respuesta:

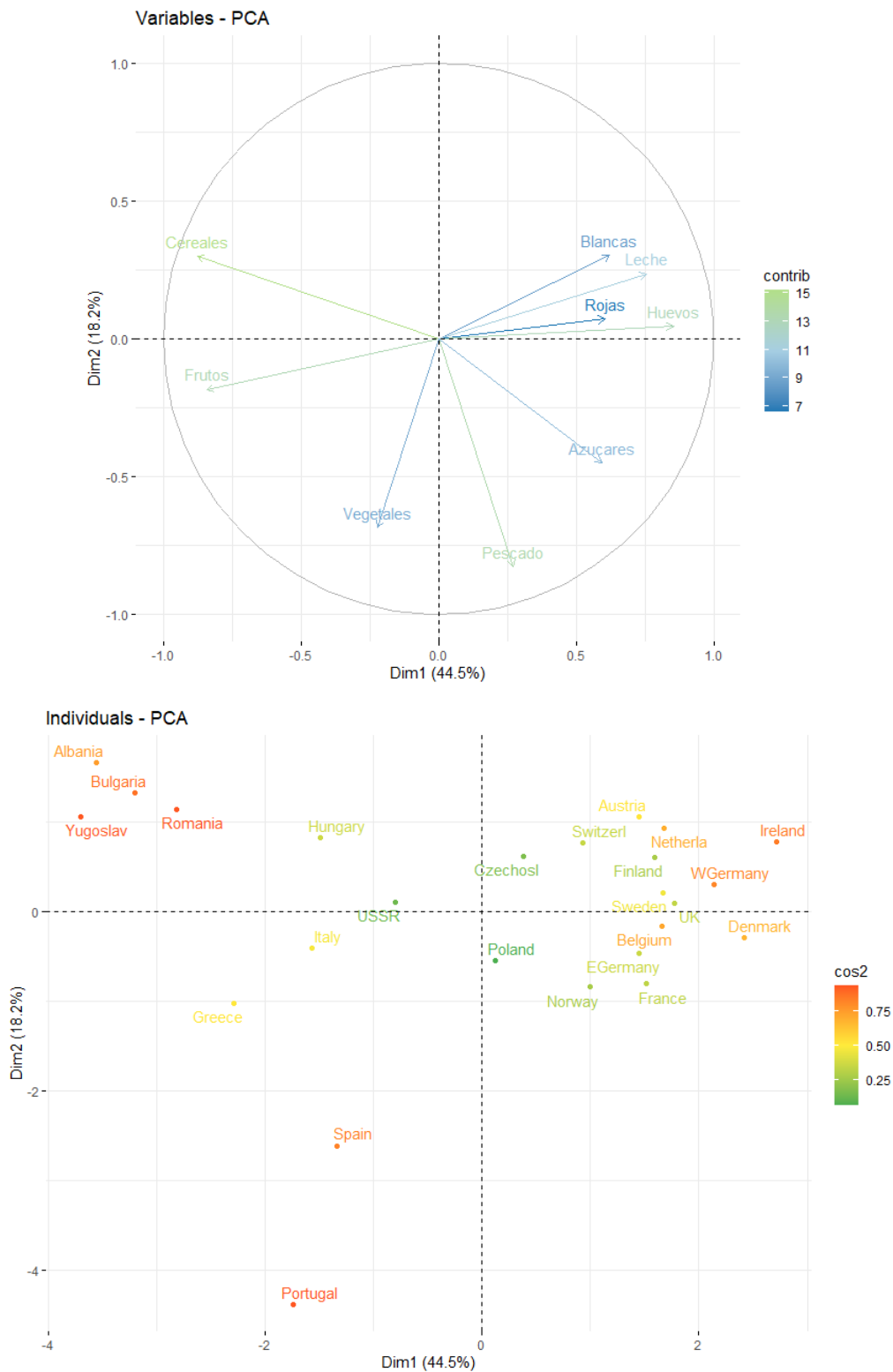
Número de componentes que explican al menos el 60% de la varianza: 2 componentes principales (Comp.1 y Comp.2). Varianza explicada por estos dos componentes: 62.68%.



Resp: podemos ver que las variables con mayor contribución al primer componente (CP1) son Cereales, Huevos, Frutos, y Leche. Esto sugiere que el CP1 está representando un eje de variabilidad relacionado principalmente con estos alimentos.

Componente 1: representa las proteínas de: cereales, huevos, frutos, leche

B ¿Es posible construir una “agregación” que permita identificar regiones?



R: Del gráfico anterior, los cuadrantes 1 y 2 (Primer cuadrante: PC1 positivo, PC2 positivo; Segundo cuadrante: (PC1 negativo, PC2 positivo) respectivamente), son los países que más incorporan proteínas en su dieta. Los países son cercanos entre sí (comparando con un viejo mapa de la URSS). También se visualiza que países cercanos tienen similar consumo de comida.



3 AFAC: Lleve a cabo un Análisis Factorial.

a. ¿Qué representa el primer factor? ¿cuántos factores proponen?

```
> psych::KMO(proteinas)
Kaiser-Meyer-Olkin factor adequacy
Call: psych::KMO(r = proteínas)
Overall MSA = 0.65
MSA for each item =
```

	Rojas	Blancas	Huevos	Leche	Pescado	Cereales
	0.63	0.49	0.72	0.81	0.37	0.77
Azúcares						
	0.76	0.67	0.38			

Considerando la escala de:

Una forma de interpretar el índice KMO es la siguiente:

1.0 >= KMO >= 0.9	muy bueno
0.9 >= KMO >= 0.8	meritorio
0.8 >= KMO >= 0.7	mediano
0.7 >= KMO >= 0.6	mediocre
0.6 >= KMO > 0.5	bajo
KMO <= 0.5	inaceptable

Y un KMO global mediocre (0.65), se obtiene el siguiente resumen:

- Rojas: 0.63 (bajo)
- Blancas: 0.49 (bajo)
- Huevos: 0.72 (mediano)
- Leche: 0.81 (meritorio)
- Pescado: 0.37 (inaceptable)
- Cereales: 0.77 (meritorio)
- Azúcares: 0.76 (meritorio)
- Frutos: 0.67 (bajo)
- Vegetales: 0.38 (inaceptable)

R: para las siguientes “proteínas”

Blancas: = 0.49 (bajo)

Pescado: = 0.37 (inaceptable)

Vegetales: = 0.38 (inaceptable)

Se recomendaría quitar del análisis factorial

Loadings:

	Factor1	Factor2
Rojas	0.561	
Blancas	0.593	-0.432
Huevos	0.839	
Leche	0.679	
Cereales	-0.902	
Azucares	0.542	
Frutos	-0.760	
Pescado		0.951
Vegetales		0.325

	Factor1	Factor2
ss loadings	3.629	1.395
Proportion Var	0.403	0.155
Cumulative Var	0.403	0.558

R: hay dos factores propuestos.

Factor 1: Estos representan a las proteínas: huevos, cereales, frutos, leche

Factor 2: representa proteínas huevos, cereales, frutos, leche, rojas, blancas, azucares.

b.) Con la solución de dos factores, realice una rotación varimax y grafique.

Grafico manual contrib

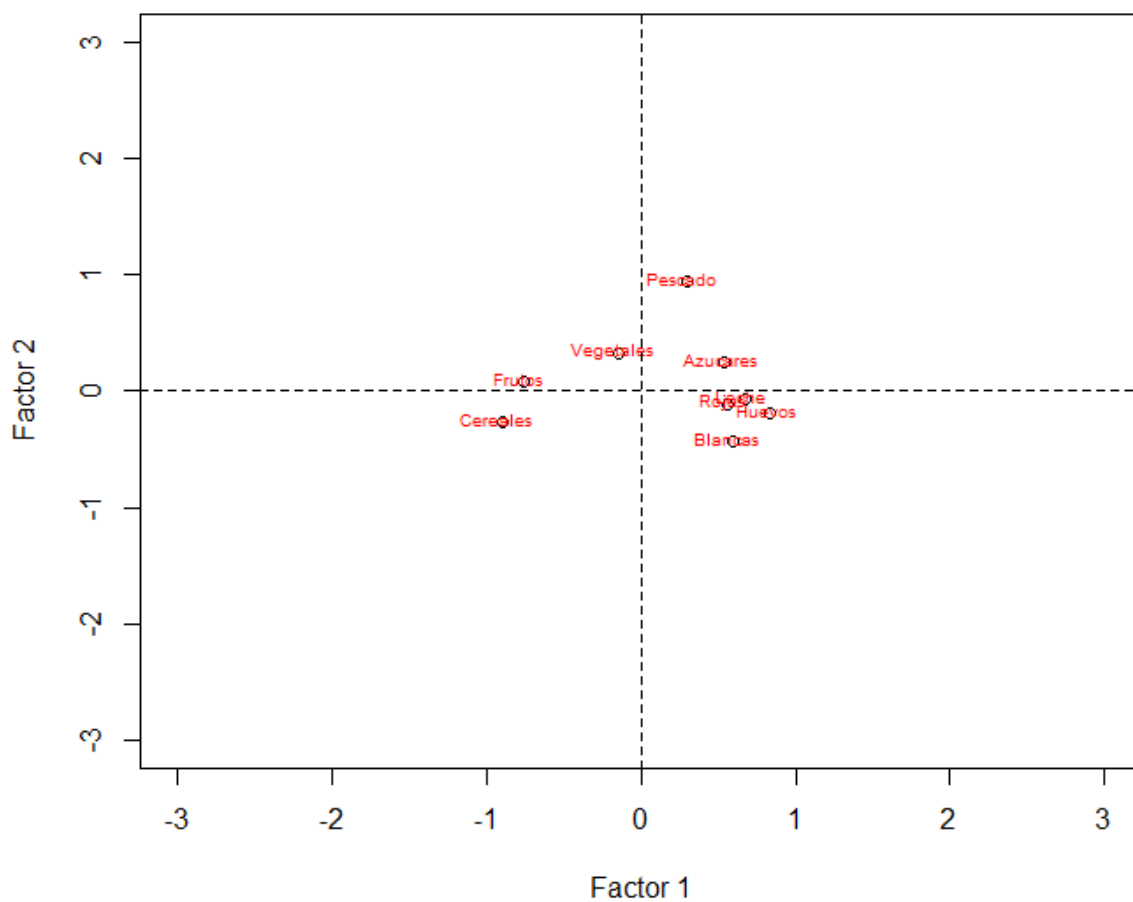
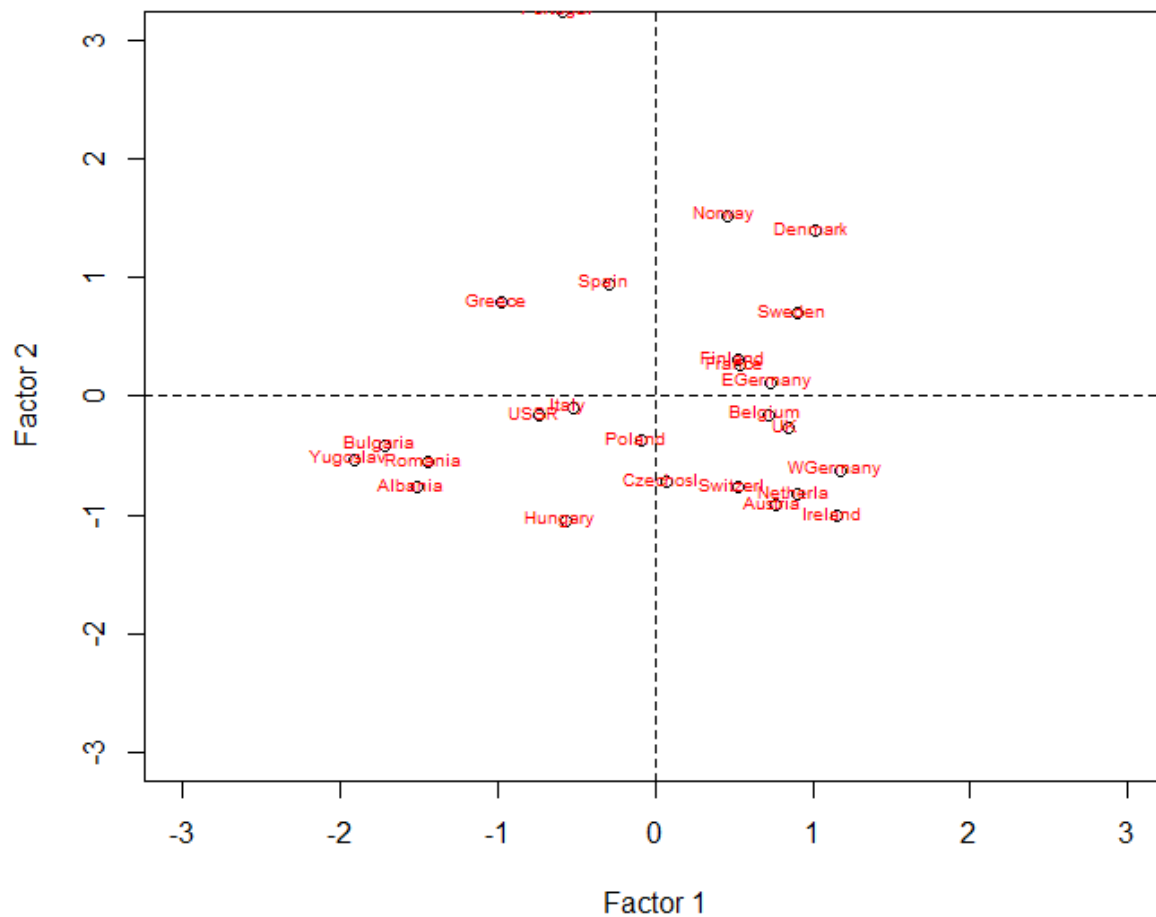
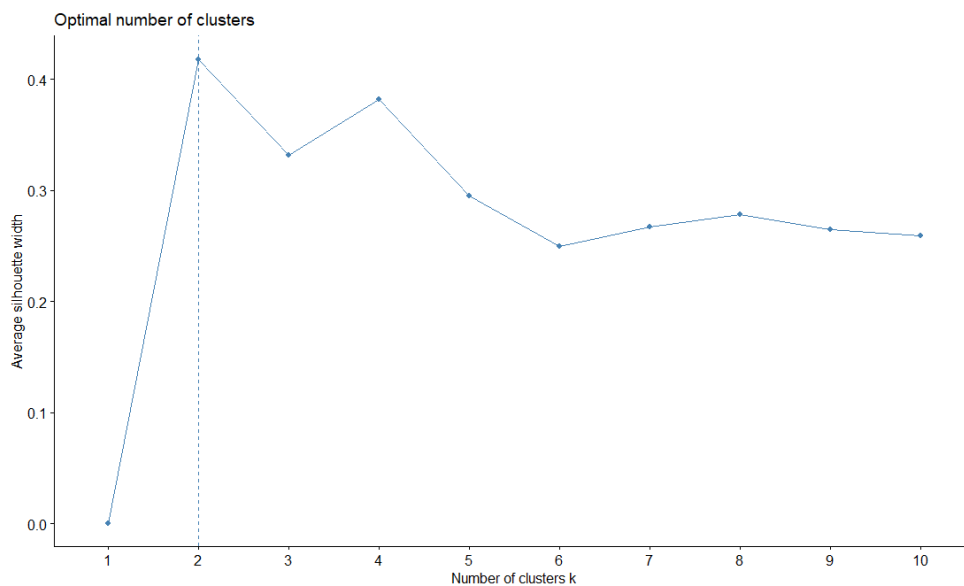


Grafico manual scores

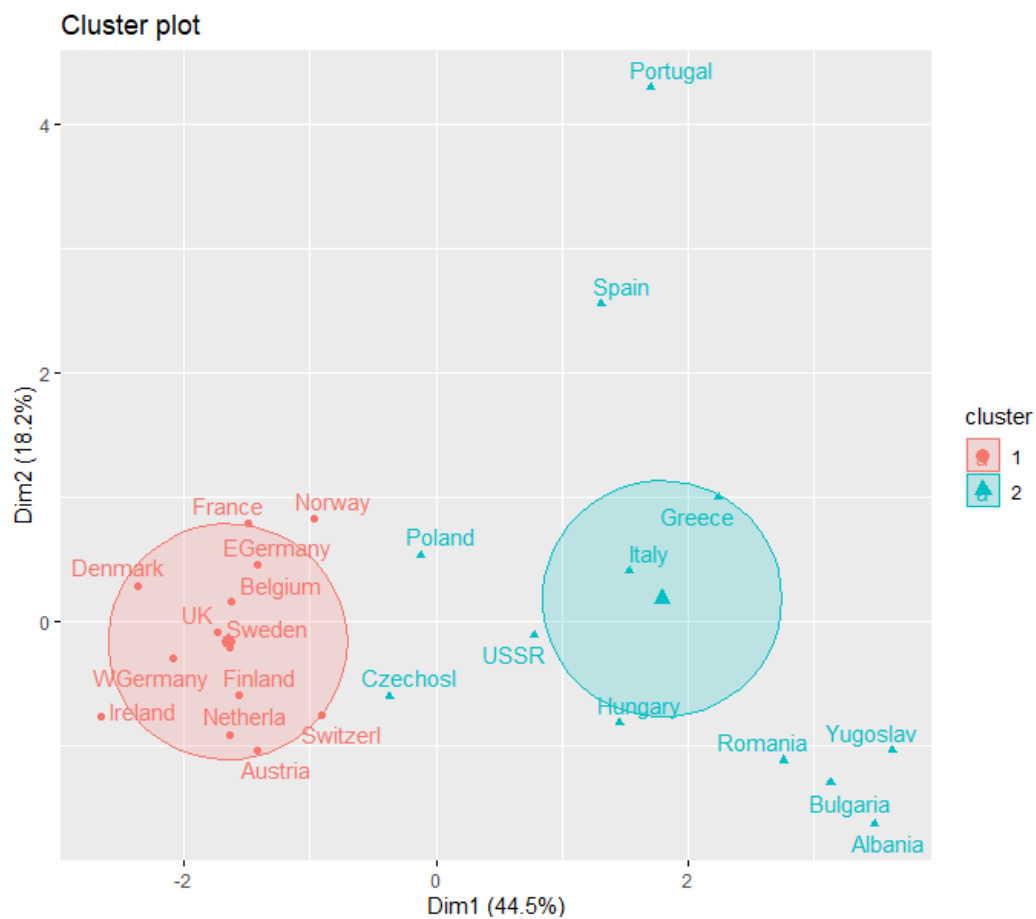


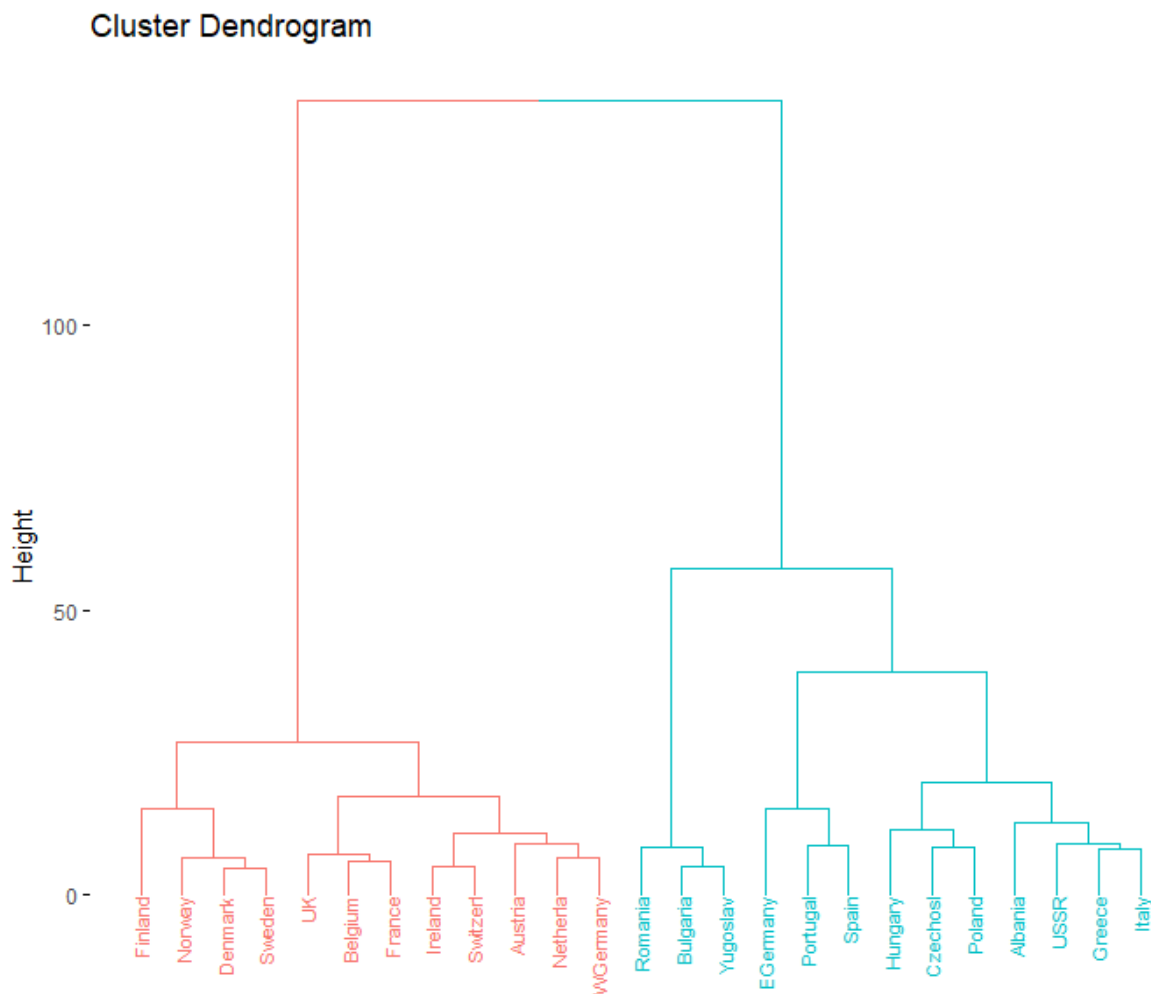
4 Clúster: Obtenga un número apropiados de “Cluster”, ¿cómo se agrupan los países?



R: Se sugieren 2 cluster

Forma 1: Cluster - No jerárquico





¿cómo se agrupan los países? Se identifica que se agrupan por su cercanía en fronteras, además de que se agrupan por la posición en Europa, ya sea el norte o el sur.

5) En no más de dos párrafos, contraste y discuta los resultados obtenidos por cada uno de los métodos

R: Los análisis de clustering muestran una clara separación entre países de Europa occidental y norte (Cluster 1) y países del sur y este (Cluster 2), lo cual refleja diferencias económicas en el consumo de proteínas. El grupo de Europa occidental y norte incluye países como Alemania, Francia y Reino Unido, que podrían tener dietas más ricas en proteínas animales debido a mayores ingresos y acceso a una variedad de alimentos. En cambio, el grupo del sur y este, que incluye a Grecia, Italia y algunos países ex-soviéticos, posiblemente refleja patrones dietéticos influenciados por la geografía y la tradición agrícola de estas regiones, con una mayor dependencia de proteínas de origen vegetal.

La estructura jerárquica del dendrograma refuerza esta división, mostrando relaciones cercanas entre países dentro de cada grupo y permitiendo observar subgrupos interesantes. En conjunto, los resultados sugieren que el perfil de consumo de proteínas en Europa está profundamente influenciado por factores culturales y económicos, destacando cómo la ubicación geográfica y el desarrollo económico impactan en la dieta de cada región.