Análisis Multivariado

Producto académico 03

Kevin Heberth Haquehua Apaza

30 de julio del 2025

Table of Contents

# Ejercicios AFE, Cluster y análisis de correspondencia múltiple

## CASO 1: Postulantes (6 puntos)

Se tienen datos de las notas de alumnos postulantes a un colegio de alto rendimiento, se desea agrupar las notas de los cursos y ver que grupos podrían haber de cursos.

Las notas de los siguientes cursos son Razonamiento verbal, Razonamiento matemático, Matemáticas, Psicología y filosofía, Física, Lógica, Biología, Historia y Química.

Archivo a utilizar **postulantes.sav**

1. Realizar análisis factorial exploratorio
2. Decida cuantos factores retener explique el por qué.
3. Decida el método de rotación y explique el por qué.
4. Explicar los resultados y de sus conclusiones del ejercicio.

### Solución

1. **Realizar análisis factorial exploratorio**
2. **Decida cuantos factores retener explique el por qué.**
3. **Decida el método de rotación y explique el por qué.**
4. **Explicar los resultados y de sus conclusiones del ejercicio.**

## CASO 2: Agrupando clientes mayoristas (7 puntos):

El conjunto de datos se refiere a los clientes de un distribuidor mayorista de Portugal, el cual comercializa distintos tipos de productos.

Cada una de las observaciones hace referencia a un cliente distinto, el cual incluye el gasto anual en unidades monetarias (u.m.) para cada una de las categorías.

Se nos solicita realizar un análisis clúster que nos permita agrupar a nuestros clientes en función de los distintos tipos de productos que adquirieron, para lo cual contamos:

| Variable | Descripción |
| --- | --- |
| Channel | Canal de clientes: 1. Horeca (Hotel/Restaurante/Café) 2. Canal Minorista |
| Región | Región de los clientes: 1. Lisboa, 2. Oporto y 3. Otra |
| Fresh | Gasto anual en productos frescos. |
| Milk | Gasto anual en productos lácteos. |
| Grocery | Gasto anual en productos comestibles. |
| Frozen | Gasto anual en productos congelados. |
| Detergent\_Papers | Gasto anual en detergentes y productos de papel. |
| Delicatessen | Gasto anual en productos preparados (snacks y licor). |

Los datos se encuentran en el archivo “clientes.csv”.

Luego de cargar el conjunto de datos en R, realizar las 2 opciones que se presenta:

**Opción 1**:

1. Generar un nuevo dataset solo con las variables numéricas y estandarizarlas.
2. Generar el agrupamiento por particiones utilizando el método kmeans con k=4.
3. Añadir el dataset original la columna cluster, que identificará a los grupos que obtuvimos mediante esta metodología.
4. Graficar y perfilar a nuestros clientes según su agrupación.

**Opción 2**:

1. Generar un nuevo dataset solo con las variables numéricas y estandarizarlas.
2. Encuentre ahora los clusters de forma jerárquica, calculando la matriz de distancias euclidianas y seleccionando en enlace que creas mejor se ajuste a los datos.
3. Comparar los métodos de enlace y determinar cuál es el adecuado.
4. Generar el nuevo agrupamiento jerárquico con el enlace seleccionado.
5. Graficar el dendograma respectivo y determinar el número de clusters.
6. Graficar y perfilar a nuestros clientes según su agrupación jerárquica.

### Solución

**Opción 1**:

1. **Generar un nuevo dataset solo con las variables numéricas y estandarizarlas.**
2. **Generar el agrupamiento por particiones utilizando el método kmeans con k=4.**
3. **Añadir el dataset original la columna cluster, que identificará a los grupos que obtuvimos mediante esta metodología.**
4. **Graficar y perfilar a nuestros clientes según su agrupación.**

**Opción 2**:

1. **Generar un nuevo dataset solo con las variables numéricas y estandarizarlas.**
2. **Encuentre ahora los clusters de forma jerárquica, calculando la matriz de distancias euclidianas y seleccionando en enlace que creas mejor se ajuste a los datos.**
3. **Comparar los métodos de enlace y determinar cuál es el adecuado.**
4. **Generar el nuevo agrupamiento jerárquico con el enlace seleccionado.**
5. **Graficar el dendograma respectivo y determinar el número de clusters.**
6. **Graficar y perfilar a nuestros clientes según su agrupación jerárquica.**

## CASO 3: (7 puntos)

Investigar y realizar un informe monográfico sobre el ***análisis de correspondencia múltiple*** adjuntar un ejercicio aplicando R o Phyton.

### Solución

Aca va la solución