

Máster TECI. Práctica Análisis Multivariante 2019/2020

Una **compañía eléctrica** quiere ofrecer productos a sus clientes en función del uso que hacen de la electricidad. Para ello, se plantean analizar sus patrones de consumo y personalizar las ofertas a realizar en función de los mismos. A tal fin, se proporciona una tabla con una muestra de 1.000 CUPS de una zona costera de España. La tabla contiene para cada CUPS (campo *CUPS*), 365 registros (uno por valor del campo *fecha*) y 24 columnas representativas del porcentaje que supone el consumo de cada una de las 8.784 horas de 1 año (bisiesto) respecto al consumo total anual (variables *pct_h:*).

Práctica por parejas:

El objetivo de la práctica es realizar **2 Análisis Cluster horarios** para saber si el consumo de un día concreto se encuentra focalizado en ciertas franjas horarias del día (consumos en horarios nocturnos, consumos constantes a lo largo del día, acentuado en las horas de las comidas, etc.). Las dos segmentaciones a realizar serán:

- Una para detectar patrones de consumo horario en días festivos.
- Otra para detectar patrones de consumo horario en días laborables.

Para ello, se deberán dar los siguientes pasos:

- Reconvertir los porcentajes proporcionados de manera que representen el porcentaje de consumo de cada hora respecto al día (la suma de los porcentajes por registro deberá valer 1).
- Calcular la media de dichos porcentajes por CUPS para días laborables y festivos de manera que se obtengan dos conjuntos con 1.000 CUPS cada uno de ellos: *work.consumoHorasFestivos* y *work.consumoHorasLaborables*.
- Realizar una segmentación sobre una y otra tabla **PERO NO** a partir de los 24 valores porcentuales originales, sino de las variables que resulten tras un **Análisis de Componentes Principales** realizado sobre dichas variables porcentuales.

Si la práctica se hace de forma individual, solo será necesario hacer una de las dos segmentaciones.

Si el grupo de prácticas se compone de 3 personas, se deberá realizar **adicionalmente un Análisis Cluster diario** para saber, por ejemplo, si el consumo de una semana se concentra en los días laborables o en los festivos. Será necesario:

- Construir una tabla con 52 registros (1 por semana completa) por CUPS y 7 columnas donde cada una de ellas representa el porcentaje de consumo de cada día respecto a la semana en cuestión.
- Calcular la media de dichos porcentajes por CUPS para acabar teniendo una tabla con 1.000 registros que refleje por CUPS el comportamiento medio semanal: *work.consumoSemanalPorDias*.
- Realizar una segmentación respecto a dichas variables. **No hacer Componentes Principales para esta segmentación.**

Si en el grupo de prácticas se compone de 4 personas, se deberá realizar **adicionalmente un Análisis Cluster mensual** para saber, por ejemplo, si el consumo de un año se concentra en ciertos meses (está condicionado por la temperatura, casa en la playa, etc.). Será necesario:

- Construir una tabla con 1 registro por CUPS (1.000 registros nuevamente) y 12 columnas donde cada una de ellas represente el consumo mensual de cada mes respecto al año: *work.consumoAnualPorMeses*.
- Realizar una segmentación respecto a dichas variables. **No hacer Componentes Principales para esta segmentación.**

Se deberá entregar un informe en formato .PDF o .WORD con los **pasos dados para la construcción de las tablas de entrada a los modelos**, la **justificación de las decisiones tomadas en cada uno de los modelos analíticos desarrollados (Análisis de Componentes Principales y Análisis Cluster)** así como los **resultados obtenidos con cada uno de ellos y una propuesta final de grupos para la compañía eléctrica**. Fecha límite de entrega de la práctica: **domingo 15 de diciembre de 2019 a las 23:59.**