



Pontificia Universidad Javeriana

Tarea No.2 de Análisis Multivariado

Fecha de Asignación : Martes 24 de Septiembre de 2024

Fecha de entrega (física): Martes 01 de Octubre de 2024

1. Los datos `USairpollution` del paquete `HSAUR3` se refieren a la contaminación del aire en 41 ciudades de Estados Unidos. Se registraron las siguientes variables:

SO2 : Contenido de dióxido de azufre en el aire en microgramos por metro cúbico

temp : Temperatura (en Fahrenheit) media anual

manu : Número de empresas manufactureras que emplean a 20 o más trabajadores

popul : tamaño de la población (censo de 1970) en miles

wind : velocidad media anual del viento en millas por hora

precip : Precipitación anual promedio en pulgadas

predays : Promedio de días con precipitación al año

Los datos se recopilaron originalmente para investigar los determinantes de la contaminación, presumiblemente mediante una regresión del SO2 en las otras seis variables.

- a. De acuerdo con los datos ¿usted trabajaría con la matriz de covarianzas o con la matriz de correlación? Justifique.
 - b. Verifique si el ACP para las últimas 6 variables es viable, en caso de serlo, escriba las 6 componentes principales y sugiera un nombre para cada una. ¿Con cuantas componentes trabajaría? Justifique
2. Los datos del archivo `Hijos.txt` corresponden a las medidas de la cabeza (en milímetros) de cada uno de los dos primeros hijos adultos en 25 familias. Estos datos fueron recopilados por Frets (1921)¹, y la pregunta de interés era si existe una relación entre las medidas de la cabeza de los pares de hijos.
 - a. Verifique si el ACP para la longitud de la cabeza (`head 1` and `head2`) es viable, en caso de serlo, escriba las 2 componentes principales y sugiera un nombre para cada una. ¿Con cuantas componentes trabajaría? Justifique
 - b. Repita (a) para ancho de la cabeza (`breadth1` and `breadth2`)

¹Frets, G. P. (1921), "Heredity of head form in man", *Genetica*, 3, 193(384)