



COMPUTACIONAL ESTADÍSTICA

LABORATORIO 7

Base sociodemográfica

Unidad de análisis: regiones o provincias de un país.

Variables

Dependiente: *esperanza de vida (esp_vida)*.

Independientes: habitantes, ingresos, analfabetismo, asesinatos, universitarios, heladas, área, densidad poblacional.

Enunciado:

Un investigador en salud pública desea analizar los factores sociodemográficos y ambientales que influyen en la esperanza de vida de las provincias de un país. Se dispone de información de 50 provincias, donde se registran variables de población, ingresos, nivel educativo, seguridad ciudadana y condiciones climáticas.

Use el archivo de excel: *dataset1.xlsx*

Preguntas para regresión múltiple:

1. ¿Qué variables explican de manera significativa la variación en la esperanza de vida entre provincias?
2. ¿Cuál es el impacto del nivel de educación universitaria sobre la esperanza de vida, controlando por otras variables?
3. ¿En qué medida la densidad poblacional y los ingresos afectan conjuntamente a la esperanza de vida?
4. ¿Existen problemas de multicolinealidad entre las variables socioeconómicas?
5. Estime e interprete el coeficiente de determinación R^2 y discuta si el modelo es adecuado.



Base Retail/Marketing

Unidad de análisis: tiendas de una cadena de retail.

Variables

Dependiente: *ventas*.

Independientes: precio, publicidad (en miles), competencia, ingresos de la zona, tráfico de clientes, satisfacción de clientes.

Enunciado:

Un gerente de marketing desea conocer qué factores determinan las ventas mensuales de sus tiendas. Para ello, se dispone de información de 300 sucursales sobre precios, inversión en publicidad, nivel de competencia, ingresos promedio de la zona, tráfico de clientes y satisfacción.

Use el archivo de excel: *dataset2.xlsx*

Preguntas para regresión múltiple:

6. ¿Cuál es la relación entre la inversión en publicidad y las ventas, controlando por las demás variables?
7. ¿Qué tanto afecta el nivel de competencia en el mercado local a las ventas de la tienda?
8. ¿Las tiendas ubicadas en zonas con mayores ingresos y mayor tráfico de clientes muestran un efecto combinado significativo en las ventas?
9. Evalúe si el precio tiene un efecto negativo y significativo sobre las ventas.
10. ¿Qué porcentaje de la variabilidad de las ventas se logra explicar con el modelo de regresión múltiple?



Base Vivienda Urbana

Unidad de análisis: viviendas ofertadas en el mercado inmobiliario.

Variables

Dependiente: *precio por metro cuadrado (precio_m²).*

Independientes: distancia al centro, número de habitaciones, antigüedad, área total, índice de criminalidad, colegios cercanos, ingresos promedio del barrio.

Enunciado:

Una empresa inmobiliaria está interesada en determinar cuáles son los factores que más influyen en el precio por metro cuadrado de las viviendas en una ciudad. Para ello se analizan 250 viviendas con información sobre ubicación, características físicas, nivel de seguridad y contexto socioeconómico del barrio.

Use el archivo de excel: *dataset3.xlsx*

Preguntas para regresión múltiple:

11. ¿Cuál es el efecto de la distancia al centro de la ciudad sobre el precio de las viviendas, manteniendo constantes las demás variables?
12. ¿Qué características físicas de la vivienda (área y número de habitaciones) tienen mayor influencia en el precio por metro cuadrado?
13. ¿Los ingresos promedio del barrio explican de manera significativa la variación en los precios?
14. ¿En qué medida la criminalidad y la cercanía de colegios afectan los valores inmobiliarios?
15. Estime y analice los residuos del modelo para verificar los supuestos de regresión (normalidad, homocedasticidad, independencia).