

Taller de análisis exploratorio

Pregunta 1

El Ministerio de Salud ha solicitado un informe sobre los tiempos de espera para ser atendido en consultorios externos de medicina general en centros de salud públicos en Lima. Se realizó una recolección de datos a partir de observaciones en distintos días y horarios. A continuación, se presentan los datos registrados en una muestra de 90 pacientes, y almacenados en el archivo [datos1.csv](#). Las variables consideradas son:

- **Sede**
- **Tiempo de espera** (en minutos)
- **Tipo de atención** (cita agendada regular, cita agendada prioritaria, cita adicional, reconsulta)
- **Monto de gasto en farmacia posterior a la consulta** (en soles)

Responder las siguientes preguntas:

1. ¿Cuál es la unidad experimental del estudio?
2. Identificar la naturaleza o tipo de cada variable.
3. Interpretar la mediana del tiempo de espera en cada una de las sedes.
4. ¿Cuál es el monto mínimo de gasto del 20% de los pacientes que más gastan en farmacia?
5. Interpretar la desviación estándar del monto de gasto en farmacia.
6. ¿En qué sede los tiempos de espera son más dispersos?
7. Interpretar y comparar la asimetría en el 50% central y el 100% de los montos de gasto en farmacia en la sede de San Luis.
8. Construir la tabla de frecuencias para el tipo de atención y presentar una conclusión.
9. Elaborar una gráfica que permita visualizar la distribución del monto gastado en farmacia posterior a la consulta. Comentar la gráfica obtenida.

Pregunta 2

Una universidad pública de Lima realizó una encuesta a 80 de sus estudiantes matriculados en el ciclo 2025-2 en la modalidad presencial o semipresencial, elegidos al azar, sobre su desplazamiento hacia la universidad. Se registraron las siguientes variables:

- **Tiempo de viaje** (en minutos) desde su casa hasta la universidad para llegar a su última clase.
- **Principal medio de transporte** que usa para llegar a la universidad.
- **Gasto** (en soles) en transporte en su último recorrido hacia la universidad.
- **Sede en la que estudia** (Lima centro, Lima este, Lima norte, Lima sur).

Los datos se encuentran en el archivo [datos2.csv](#), responder las siguientes preguntas:

1. ¿Cuál es la unidad experimental del estudio?
2. Identificar la naturaleza o tipo de cada variable.
3. Interpretar la media aritmética de la variable Gasto.
4. Interpretar la moda del principal medio de transporte en cada sede.
5. ¿Cuál es el mayor tiempo que tarda el 22% de los estudiantes que llegan más rápido a la universidad?
6. ¿En qué sede los gastos en transporte son más homogéneos?
7. ¿Qué variable es menos asimétrica: tiempo de viaje o gasto?
8. Construir una tabla de frecuencias para el gasto en transporte (sin desagregar por sedes) y presentar una conclusión
9. Elaborar una gráfica que permita visualizar, para cada sede, el tiempo que demoran en llegar los estudiantes a la universidad, así como posibles valores atípicos. Comentar la gráfica obtenida.