



ESTADÍSTICA DESCRIPTIVA

Actividad de transferencia

La consigna consiste en analizar dentro de su entorno laboral, datos que contengan una variable cuantitativa y una categórica con dos estados posibles. A modo de ejemplo: un dato podría ser “cliente”, y este dato contiene las variables: “importe de la compra” (cuantitativa), y si hubiera dos posibles puntos de venta (A y B) entonces la variable categórica sería “lugar en donde compró”, y tendría dos estados posibles.

- 1) Seleccionar entonces, un tipo de dato como el mencionado en el párrafo anterior, y con ese dato concreto, tomar una muestra que contenga al menos 80 observaciones.

A partir de los datos de esa muestra, calcular e indicar los resultados obtenidos para los siguientes indicadores: Los datos pueden ser encontrados en el siguiente enlace: [datos](#); y el código con el que se realizaron los cálculos se encuentra en el siguiente enlace: [código](#).

a) Media aritmética = 5,451.68

b) Mediana = 5,615.00

c) Desvío muestral = 2,862.06

d) Amplitud intercuartílica = 4,954.00

e) Coeficiente de asimetría de Fisher = -0.11

f) Explicar cuál de estos indicadores utilizaría como indicador de centro y porqué: El indicador más adecuado para representar el centro de estos datos sería la mediana. Esto se debe a que la asimetría de -0.11 indica una distribución casi simétrica, pero en casos donde la distribución puede contener valores extremos (outliers), la mediana es más robusta.

La mediana es menos sensible a valores atípicos en comparación con la media, lo que la convierte en una mejor medida en este contexto, especialmente si existe variabilidad significativa.

g) Indicar cuál emplearía como indicador de dispersión, y porqué: El desvío estándar es el indicador más adecuado para representar la dispersión en este caso, ya que mide la variabilidad promedio alrededor de la media y es



especialmente útil en distribuciones que son aproximadamente simétricas (como indica la baja asimetría).

La amplitud intercuartílica (IQR) también puede ser útil si el interés está en analizar la dispersión central sin considerar extremos, pero el desvío estándar es más interpretativo en términos de distancia promedio respecto a la media.

h) Explicar qué significado le atribuye al coeficiente de asimetría obtenido: El coeficiente de asimetría de -0.11 indica que la distribución de los datos es prácticamente simétrica, con una ligera inclinación hacia valores menores que la media. En términos prácticos, esto significa que no hay un sesgo significativo hacia valores extremos en ninguno de los extremos del rango. Este valor sugiere que las medidas de tendencia central (media y mediana) están bastante alineadas, y cualquier análisis basado en estos indicadores será representativo del conjunto de datos.

i) Redactar brevemente qué utilidad podría tener esta información en su contexto laboral: En un contexto laboral, esta información podría ser útil para tomar decisiones fundamentadas basadas en patrones de datos cuantitativos. Por ejemplo:

- **Planeación financiera:** La media y la mediana pueden ayudar a estimar valores típicos esperados, mientras que el desvío estándar y el IQR aportan información sobre la variabilidad, lo cual es crucial para evaluar riesgos.
- **Segmentación de clientes:** La asimetría cercana a cero indica una distribución balanceada en el comportamiento, lo cual permite segmentar adecuadamente en base a otras variables categóricas o cuantitativas.
- **Evaluación de rendimiento:** Si estos datos reflejan métricas de rendimiento o resultados, la baja asimetría y medidas de dispersión pueden usarse para establecer benchmarks realistas y detectar áreas de mejora.

2) A continuación, segmentar el conjunto de datos, considerando la variable categórica en cada una de sus dos alternativas (por ejemplo, clientes que compraron en A y clientes que compraron en B), y así generar entonces, dos series de datos (de modo tal que cada una tenga al menos 30 datos (por ejemplo, si originalmente había 90 datos, podría haber un segmento de 55 que compraron en A, y otro de 35 que compraron en B).

Para cada una de esas series, calcular, la media, el desvío y el coeficiente de asimetría de Fisher:

a) Media (A) = 5,368.33

d) Media (B) = 5,509.21

b) Desvío (A) = 2,709.58

e) Desvío (B) = 2,980.40



c) Asimetría (A) = -0.14

f) Asimetría (B) = -0.11

g) De acuerdo con los indicadores obtenidos, comparar ambos segmentos y discutir acerca de si observa que los indicadores están alineados, o no, con su opinión al respecto:

Los indicadores están razonablemente alineados y sugieren comportamientos similares entre los segmentos A y B. Aunque las compras en B tienen una media ligeramente mayor y una mayor dispersión, la similitud en la asimetría y la cercanía en los valores promedio reflejan que no hay diferencias drásticas entre los dos lugares de compra.

En términos de opinión, estos resultados podrían indicar que los patrones de compra no están significativamente influenciados por el lugar de compra, pero la mayor dispersión en B podría estar asociada con factores contextuales, como una mayor diversidad en los tipos de clientes o productos ofrecidos en ese lugar.

ACLARACIÓN: en caso de no contar con información de una empresa, que pueda aplicar a este caso, puede utilizar algún tipo de fuente de datos de la realidad que considere aplicable y de interés.

Nota: El entregable es este archivo con los puntos suspensivos completados.

El nombre de archivo debe contener “Estadística Descriptiva” y también “Nombre y Apellido del alumno”.