

Análisis de la base de datos “Titanic-Train”

Autora: Danna Pauleth López Acosta

Este informe presenta una descripción detallada de los atributos incluidos en la base de datos "Titanic-Train", que recopila información sobre los pasajeros del Titanic. A continuación, se especifica cada variable contenida en el conjunto de datos:

- **PassengerID:** Número de identificación único asignado a cada pasajero dentro del conjunto de datos. Su rango varía de 1 a 891 y es un valor entero exclusivo para cada registro.
 - **Tipo de dato:** Numérico entero
 - **Categorico:** No
- **Survived:** Variable que indica si un pasajero logró sobrevivir al naufragio.
 - **Valor 1:** Sobrevivió
 - **Valor 0:** No sobrevivió
 - **Tipo de dato:** Numérico entero
 - **Categorico:** Sí (binario)
- **Pclass:** Clase del billete del pasajero, reflejando su nivel socioeconómico en el barco.
 - **1:** Primera clase (alta)
 - **2:** Segunda clase (media)
 - **3:** Tercera clase (baja)
 - **Tipo de dato:** Numérico entero
 - **Categorico:** Sí (ordinal)
- **Name:** Nombre completo del pasajero.
 - **Tipo de dato:** Texto
 - **Categorico:** Sí (nominal)
- **Sex:** Género del pasajero, con dos opciones posibles:
 - **"male":** Masculino
 - **"female":** Femenino
 - **Tipo de dato:** Texto
 - **Categorico:** Sí (nominal)

- **Age:** Edad del pasajero expresada en años. Puede incluir valores decimales y contener datos faltantes.
 - **Tipo de dato:** Numérico continuo
 - **Categorico:** No
- **SibSp:** Cantidad de hermanos o pareja que el pasajero tenía a bordo del Titanic.
 - **Tipo de dato:** Numérico entero
 - **Categorico:** No (es un conteo, aunque a veces se considera discreto)
- **Parch:** Número de padres o hijos que acompañaban al pasajero en el barco.
 - **Tipo de dato:** Numérico entero
 - **Categorico:** No (es un conteo, aunque a veces se considera discreto)
- **Ticket:** Código del boleto del pasajero, el cual puede contener tanto números como letras y no siempre es único.
 - **Tipo de dato:** Texto
 - **Categorico:** Sí (nominal)
- **Fare:** Precio del boleto pagado, cuyo valor depende de la clase y ubicación de la cabina.
 - **Tipo de dato:** Numérico continuo
 - **Categorico:** No
- **Cabin:** Código de la cabina asignada a cada pasajero. Muchos viajeros, en especial los de tercera clase, no tenían cabina designada, lo que ocasiona que algunos registros aparezcan vacíos en esta variable.
 - **Tipo de dato:** Texto
 - **Categorico:** Sí (nominal)
- **Embarked:** Puerto de embarque del pasajero, con las siguientes opciones:
 - **S:** Southampton
 - **C:** Cherburgo
 - **Q:** Queenstown
 - **Tipo de dato:** Texto
 - **Categorico:** Sí (nominal)