

Descripción de las características de la base de datos “Titanic-Train”

Realizado por:

- Laura Daniela Rojas Ardila
- Stiven Alexis Villamizar Buitrago

En el presente documento se detallan las características de la base de datos “Titanic-Train”, la cual, como su nombre indica contiene, datos relacionados con los pasajeros del Titanic. A continuación, se describe cada una de las características contenidas en el archivo:

PassengerID: Identificador único asignado a cada pasajero, es decir, el número de pasajero en el dataset. Toma valores de 1 a 891. Es un número entero y único para cada pasajero.

Tipo de dato: Numérico entero

Categorico: No

Survived: Indica si el pasajero sobrevivió al hundimiento del Titanic.

1: Pasajero sobrevivió

0: Pasajero no sobrevivió

Tipo de dato: Numérico entero

Categorico: Sí (binario)

Pclass: Clase del boleto del pasajero, que indica el nivel socioeconómico de su cabina 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: Cadena de texto

Categorico: Sí (nominal)

Sex: Género del pasajero. Los valores posibles son "male" (masculino) y "female" (femenino).

Tipo de dato: Cadena de texto

Categorico: Sí (nominal)

Age: Edad del pasajero en años. Puede contener valores decimales y algunos datos faltantes.

Tipo de dato: Numérico continuo

Categorico: No

SibSp: Número de hermanos o cónyuges 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 del pasajero a bordo.

Tipo de dato: Numérico entero

Categorico: No (es un conteo, aunque a veces se considera discreto)

Ticket: Código del boleto del pasajero. Puede contener números y letras, y no es necesariamente único.

Tipo de dato: Texto

Categorico: Sí (nominal)

Fare: Precio del boleto pagado por el pasajero. El valor varía según la clase y la ubicación de la cabina.

Tipo de dato: Numérico continuo

Categorico: No

Cabin: Código de la cabina asignada al pasajero. Algunos pasajeros, especialmente los de tercera clase, no tienen cabina asignada, por lo que este campo puede estar vacío.

Tipo de dato: Texto

Categorico: Sí (nominal)

Embarked: Puerto donde el pasajero abordó el Titanic.

S: Southampton

C: Cherburgo

Q: Queenstown

Tipo de dato: Texto

Categorico: Sí (nominal)