

Hoja informativa: Trabajar con valores ausentes y duplicados

Práctica

```
# Valores únicos de columna y conteos
data['column'].value_counts()
```

```
# Operaciones aritméticas con columnas
# ¡NB!: las columnas deben tener valores numéricos
data['column1'] = data['column2'] + data['column3']
data['column1'] = data['column2'] - data['column3']
data['column1'] = data['column2'] * data['column3']
data['column1'] = data['column2'] / data['column3']
```

```
# Operaciones importantes para columnas con valores numéricos
data['column'].sum()
data['column'].min()
data['column'].max()
data['column'].mean()
data['column'].median()

# Número de valores en una columna
data['column'].count()
```

Teoría

Fuentes de tráfico: canales a través de los cuales los visitantes llegan a los sitios web.

Tasa de conversión: una medición métrica de la efectividad de una fuente de tráfico (ya sea el porcentaje de visitantes que realizan acciones focalizadas o el porcentaje de

acciones focalizadas comparadas con el total de las acciones).

Tasa de cliente frecuente: la proporción de clientes que realizan más de una compra frente al número total de clientes.

Sistema de analítica web: un sistema que reúne datos automáticamente de los contadores en sitio para registrar el comportamiento del usuario.

ID de usuario: un número asignado a un visitante para distinguirlo de los demás.

Cookies: archivos especiales de texto que permanecen en una memoria del dispositivo tras la primera visita del usuario y que se envían al servidor en las visitas repetidas.

Registros: archivos de texto con datos sobre las visitas al sitio.

NaN ("no es un número"): un valor especial float que se usa cuando no se puede mostrar o realizar un cálculo (p. ej., $=0/0$).

None: un valor especial None Type ("sin tipo") usado cuando está ausente un valor.

Variable categórica: una variable que toma sus valores de un conjunto limitado.

Variable cuantitativa: una variable que toma valores numéricos dentro de un rango.

Variable lógica (booleana): una variable que toma los valores Verdadero o Falso.

Diccionario: una estructura de datos que contiene pares **clave-valor**.