¿Qué es un conjunto de datos?

Noticias 24hs | 16/02/2023



Un conjunto de datos es una colección de elementos discretos relacionados de datos relacionados a los que se puede acceder individualmente o en combinación o administrar como una entidad completa

Un conjunto de datos está organizado en algún tipo de estructura de datos. En una base de datos, por ejemplo, un conjunto de datos puede contener una colección de datos comerciales (nombres, salarios, información de contacto, cifras de ventas, etc.). La base de datos en sí puede considerarse un conjunto de datos, al igual que los conjuntos de datos que contiene relacionados con un tipo particular de información, como los datos de ventas de un departamento corporativo en particular.

El término conjunto de datos se originó con IBM, donde su significado era similar al de archivo. En un sistema operativo de mainframe de IBM, un conjunto de datos es una colección de datos con nombre que contiene unidades de datos individuales organizadas (formateadas) de una manera específica prescrita por IBM y a las que se accede mediante un método de acceso específico basado en la organización del conjunto de datos.

¿Qué es un conjunto de datos?

Un conjunto de datos para empresas es una colección organizada de datos que pueden ser recogidos en la organización y operaciones de la empresa . Por lo general, se asocian con un cuerpo de trabajo único y, por lo general, cubren un tema a la vez. Los elementos de información dentro de un conjunto de datos se relacionan entre sí, y los analistas a menudo clasifican los tipos de datos para crear conjuntos de datos relevantes que respalden procesos comerciales importantes, como métricas financieras o transacciones de ventas.

En las profesiones científicas y estadísticas, los conjuntos de datos pueden ayudar a profesionales como los biólogos a analizar información sobre el medio ambiente o el clima de un área. En el comercio minorista, una empresa puede almacenar información relacionada con sus clientes en un conjunto de datos para su análisis. Los investigadores, científicos, matemáticos y analistas de finanzas, economía, ventas y marketing suelen utilizar conjuntos de datos con regularidad en sus trabajos.

Diferencia entre conjunto de datos y base de datos.

Los conjuntos de datos son diferentes de las bases de datos. Esencialmente, una base de datos es una colección de conjuntos de datos. Por lo tanto, las bases de datos suelen ser más grandes y contienen mucha más información que un conjunto de datos. Las bases de datos pueden cubrir un rango de enfoque más amplio, mientras que un conjunto de datos generalmente solo almacena información sobre un tema. Para acceder y manipular bases de datos, los científicos de datos confían en sistemas informáticos sofisticados.

¿Cuáles son los tipos de conjuntos de datos?

Hay varios tipos de conjuntos de datos. Lo que determina el tipo de conjunto de datos es la información que contiene. A continuación se muestran los tipos de conjuntos de datos que puede ver:

Numérico

Un conjunto de datos numéricos es aquel en el que todos los datos son números. También puede referirse a este tipo como un conjunto de datos cuantitativos, ya que los valores numéricos pueden aplicarse a cálculos matemáticos cuando sea necesario. Muchos procesos de análisis financiero también se basan en conjuntos de datos numéricos, ya que los valores del conjunto pueden representar números en cantidades en dólares. Los ejemplos de un conjunto de datos numéricos pueden incluir:

El número de cartas en una baraja.

Medidas de altura y peso de una persona.

Las medidas de los espacios interiores.

El número de páginas de un libro.

Categórico

Los conjuntos de datos categóricos contienen información relacionada con las características de una persona u objeto. Los científicos de datos también se refieren a conjuntos de datos categóricos como conjuntos de datos cualitativos porque contienen información relacionada con las cualidades de un objeto. Hay dos tipos de conjuntos de datos categóricos: dicotómicos y politómicos.

En un conjunto de datos dicotómicos, cada variable solo puede tener uno de dos valores. Por ejemplo, un conjunto de datos que contiene respuestas a preguntas verdaderas y falsas es dicotómico porque solo proporciona un resultado u otro. En un conjunto de datos politómicos, puede haber más de dos valores posibles para cada variable. Por ejemplo, un conjunto de datos que contiene el color de ojos de una persona puede brindarle múltiples resultados.

Autor: Redacción