

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA

Desarrollo y Gestión de Software



Extracción de Conocimiento en Bases de Datos

Diagnóstico Unidad 2: Preparación de datos

IDGS91N

PRESENTA:

T.S.U. Hugo Uriel Chaparro Estrada

DOCENTE:

Enrique Mascote

Chihuahua, Chih., 9 sep 2025

1. ¿Qué tipo de dato no está estructurado?
 - A) CSV
 - B) JSON
 - C) Texto libre
 - D) Tabla SQL
2. Un esquema “estrella” incluye:
 - A) Varias tablas de hechos
 - B) Una tabla de hechos y varias de dimensiones
 - C) No usa dimensiones
 - D) Solo tablas de dimensiones
3. ¿Cuál técnica no sirve para la imputación de valores faltantes?
 - A) Media aritmética
 - B) Regresión
 - C) Valor constante
 - D) Análisis de clusters
4. ¿Qué parámetro de base de datos mejora el rendimiento de consultas grandes?
 - A) Indexación
 - B) Codificación de texto
 - C) Interfaz gráfica
 - D) Formato PDF

Respuesta Corta

5. Define brevemente qué es “outlier”.
6. Menciona dos fuentes comunes de datos semiestructurados.
7. ¿Para qué sirve particionar tablas en un data warehouse?