

**UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA**  
**DESARROLLO Y GESTIÓN DE SOFTWARE**



**EXTRACCIÓN PARA CONOCIMIENTOS EN BASES  
DE DATOS**

**Diagnostico Unidad 2: Preparación de datos**

DOCENTE:

ING. LUIS ENRIQUE MASCOTE CANO

PRESENTA:

DARON TARÍN GONZÁLEZ

GRUPO:

IDGS91N

Chihuahua, Chih., 30 de septiembre de 2025

1. ¿Qué tipo de dato no es 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 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.

Repositorios de internet, colecciones

7. ¿Para qué sirve particionar tablas en un data warehouse?

Para dejar fuera los datos que no sean de utilidad dentro de las consultas