

**Universidad Tecnológica de Chihuahua Tecnologías de la
Información**



**Universidad Tecnológica
de Chihuahua**

Diagnostico Unidad 2: Preparación de datos

Alumno:

Erick Adrian Sanchez Cervantes

Grupo:

IDGS91N

Materia:

Extracción de
Conocimiento
en Bases de
Datos

Docente:

Enrique Mascote

Contenido

Diagnostico.....	3
Respuesta Corta	4

Diagnostico

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”.

Es un valor atípico que se desvía significativamente del resto de los datos.

6. Menciona dos fuentes comunes de datos semiestructurados.
- Archivos XML
 - Documentos JSON
7. ¿Para qué sirve particionar tablas en un data warehouse?

Para mejorar el rendimiento de las consultas