

## **Examen UD1. Big Data Aplicado. CFP Virgen de Gracia. 10/12/2024**

*El examen consta de partes:*

1. *Test:*
  - a. *Cada pregunta acertada suma 0,15 puntos.*
  - b. *A su vez, por cada 3 preguntas erróneas se resta una correcta.*
2. *Ejercicios: la puntuación se incluyen en cada ejercicio*
3. *Únicamente hay que entregar la plantilla, en formato PDF.*

### **Parte 1. Test (5,25 puntos)**

**Pregunta 1.** ¿Cuál de las siguientes características describe mejor el concepto de datos?

- A. Conocimiento aplicado para la toma de decisiones.
- B. Hechos crudos y sin procesar, sin contexto ni significado.
- C. Información organizada en un contexto útil.
- D. Reglas y patrones desarrollados a partir de la información.

**Pregunta 2.** ¿Cuál es el paso que convierte los datos en información?

- A. Organizar los datos y darles contexto para que sean significativos.
- B. Procesar la información para generar patrones y reglas.
- C. Analizar datos para identificar tendencias complejas.
- D. Asociar directamente datos con conocimiento estratégico.

**Pregunta 3.** ¿Cómo se define el conocimiento en relación con los datos y la información?

- A. Es la interpretación directa de los datos sin procesar.
- B. Es el procesamiento de la información para reconocer patrones y tomar decisiones.
- C. Es la simple organización de los datos con un contexto específico.
- D. Es la visualización gráfica de la información en un sistema Big Data.

**Pregunta 4.** ¿Cuál de las siguientes afirmaciones describe mejor el procesamiento *online de datos*?

- A. Los datos se analizan después de ser almacenados, como en la detección de fraudes bancarios.
- B. Se realiza en sistemas distribuidos, adquiriendo recursos bajo demanda.
- C. Los datos se procesan al generarse, como en sistemas de control de tráfico en tiempo real.
- D. Utiliza programación paralela y multiprocesador para optimizar tiempos.

**Pregunta 5.** ¿Cuál de los siguientes formatos de datos se analiza utilizando técnicas de procesamiento de lenguaje natural (PLN)?

- A. Imágenes y videos para el reconocimiento visual.
- B. Datos geográficos representados en grafos.
- C. Texto de fuentes como redes sociales y archivos PDF.
- D. Audio procesado para análisis acústico.

**Pregunta 6.** ¿Cuál de las siguientes características describe a un almacén de datos (Data Warehouse) según W. Inmon?

- A. Se enfoca únicamente en transacciones diarias.
- B. Es no volátil, integrado y orientado a temas.
- C. Almacena datos de manera centralizada y sin escalabilidad.
- D. Permite exclusivamente análisis transaccional en tiempo real.

**Pregunta 7.** ¿Qué característica clave diferencia a los sistemas OLAP de los sistemas OLTP?

- A. OLAP se centra en transacciones diarias; OLTP en análisis estratégico.
- B. OLTP almacena datos históricos; OLAP los modifica constantemente.

- C. OLAP se orienta al análisis y toma de decisiones estratégicas; OLTP gestiona transacciones diarias.
- D. OLTP utiliza representación multidimensional; OLAP, una estructura relacional.

**Pregunta 8.** ¿Cuál es uno de los retos principales que un almacén de datos resuelve en términos de integración?

- A. Escalar el sistema para soportar múltiples usuarios simultáneamente.
- B. Representar los datos en múltiples dimensiones para su análisis.
- C. Combinar datos de distintas fuentes, eliminando duplicidades e inconsistencias.
- D. Facilitar el acceso desde cualquier dispositivo sin comprometer el rendimiento.

**Pregunta 9.** ¿Qué función realiza la capa de extracción, transformación y carga (ETL) en un almacén de datos?

- A. Proporciona herramientas de visualización y acceso a los datos.
- B. Recoge datos de diferentes fuentes externas y transaccionales.
- C. Limpia, transforma y estandariza los datos antes de almacenarlos.
- D. Almacena los datos en modelos multidimensionales para su análisis.

**Pregunta 10.** ¿Cuál de las siguientes es una característica de la capa de almacenamiento de datos en la arquitectura de un almacén de datos?

- A. Recoge datos de sistemas transaccionales y bases de datos externas.
- B. Estandariza los datos para garantizar su calidad.
- C. Utiliza modelos multidimensionales como esquema estrella o copo de nieve.
- D. Facilita el acceso a los datos mediante consultas OLAP.

**Pregunta 11.** ¿Qué capa en la arquitectura de un almacén de datos permite que los usuarios finales accedan y analicen los datos?

- A. Capa de origen.
- B. Capa ETL.
- C. Capa de almacenamiento de datos.
- D. Capa de presentación y acceso.

**Pregunta 12.** Los datos en términos de Big Data son:

- A. Homogéneos, desestructurados y dinámicos
- B. Heterogéneos, desestructurados y dinámicos
- C. Homogéneos, estructurados y dinámicos
- D. Heterogéneos, desestructurados y estáticos

**Pregunta 13.** ¿Cuál es la principal utilidad del Big Data en relación a la toma de decisiones?

- A. Ayudar a la toma de decisiones
- B. No tiene relación en la toma de decisiones, sino en el almacenamiento de datos.
- C. El Big Data solamente toma decisiones en situaciones de estrés
- D. Ninguna de las anteriores

**Pregunta 14.** Las 3 uves del Big Data son...

- A. Volumen, Velocidad y Valentía
- B. Volumen, Velocidad y Volatilidad
- C. Volumen, Velocidad y Variedad
- D. Volumen, Volatilidad y Valor

**Pregunta 15.** Un enfoque de solución de Big Data es...

- A. Almacenes de datos

- B. Bases de datos documentales
- C. Bases de datos sobre grafos
- D. Todas las anteriores son opciones correctas.

**Pregunta 16.** Un almacén de datos tiene como una de sus cualidades principales...

- A. Rendimiento
- B. Escalabilidad
- C. Flexibilidad
- D. Todas son correctas

**Pregunta 17.** ¿Qué es el proceso ETL?

- A. Extorsión, Transformación y Carga
- B. Extracción, Transformación y Carga
- C. Estadística, Transformación y Carga
- D. Estadística, Transformación y Canibalismo

**Pregunta 18.** Durante la fase de Transformación de ETL, se...

- A. Limpian los datos para evitar anomalías.
- B. Se localizan las fuentes de datos (como archivos .csv)
- C. Se cargan en una base de datos relacional.
- D. Todas las anteriores son correctas.

**Pregunta 19.** Soluciones como MongoDB y Neo4J son bases de datos...

- A. SQL
- B. No SQL
- C. Relacionales
- D. Relacionales y No SQL.

**Pregunta 20.** Cada entrada en una base de datos MongoDB se llama...

- A. Columna
- B. Fila
- C. Documento
- D. Tabla

**Pregunta 21.** En la sentencia de MongoDB `db.usuarios.find({edad: 25}, {nombre: 1})`, ¿qué sería `{nombre: 1}`?

- A. Colección
- B. Consulta
- C. Proyección
- D. Potenciador

**Pregunta 22.** En la sentencia de MongoDB `db.usuarios.find({edad: 25}, {nombre: 1})`, ¿qué sería `{edad: 25}`?

- A. Colección
- B. Consulta
- C. Proyección
- D. Potenciador

**Pregunta 23.** En la sentencia de MongoDB `db.usuarios.find({edad: 25}, {nombre: 1})`, ¿qué implica que al lado de nombre aparezca un "1"?

- A. Solo se devolverá el campo "nombre" en los resultados de la consulta.

- B. No se devolverá el campo "nombre" en los resultados de la consulta.
- C. Se devolverán todos los campos, y "nombre" será el primero.
- D. No tiene nada que ver con los campos que se devuelven.

**Pregunta 24.** En la sentencia de MongoDB *db.usuarios.find({edad: 25}, {nombre: 1})*; ¿qué implicaría si en lugar de aparecer un "1" apareciera un "0"?

- A. Solo se devolverá el campo "nombre" en los resultados de la consulta.
- B. No se devolverá el campo "nombre" en los resultados de la consulta.
- C. Se devolverán todos los campos, y "nombre" será el primero.
- D. No tiene nada que ver con los campos que se devuelven.

**Pregunta 25.** El uso de grafos en Big Data es útil en aplicaciones como...

- A. Recomendaciones personalizadas
- B. Análisis de tendencias y opiniones
- C. Optimización de publicidad dirigida
- D. Todas las anteriores son correctas

**Pregunta 26.** Si ejecuto en Neo4J la sentencia: *MATCH (n) RETURN (n)*, ¿qué ocurre?

- A. Se muestra el grafo de forma gráfica.
- B. Se muestra el grafo en forma de tabla.
- C. Se guarda el grafo
- D. Se borra el grafo

**Pregunta 27.** Si ejecuto en Neo4J la sentencia: *CALL gds.graph.drop('myGraph1')*; ¿qué ocurre?

- A. Se muestra el grafo de forma gráfica.
- B. Se muestra el grafo en forma de tabla.
- C. Se guarda el grafo
- D. Se borra el grafo

**Pregunta 28.** Si quiere conocer el camino más corto entre dos nodos de un grafo...

- A. Hago un recorrido en anchura
- B. Hago un recorrido en profundidad
- C. Calculo el camino mínimo.
- D. Ninguna de las anteriores.

**Pregunta 29.** Si ejecuto en Neo4J la sentencia: *CREATE (c3:City {name: 'Ciudad C'})*, ¿qué ocurre?

- A. Se crea una conexión de c3 a City
- B. Se crea una conexión de City a c3
- C. Se crea un nodo llamado "Ciudad C"
- D. Ninguna de las anteriores.

**Pregunta 30.** La medida que "evalúa la importancia de un nodo en función de la cantidad de conexiones que tiene" es...

- A. Centralidad de grado
- B. Conteo de triángulos
- C. Intermediación
- D. Coeficiente local de agrupamiento

**Pregunta 31.** La medida que se basa en la idea "Dos extraños con un amigo en común tienen más posibilidad de conectarse entre sí que quienes no" es:

- A. Adhesión preferencial
- B. Vecinos comunes
- C. Intermediación
- D. Coeficiente local de agrupamiento

**Pregunta 32.** ¿Qué hace la siguiente sentencia en Neo4j?

```
MATCH (a:Node {name: 'A'}), (b:Node {name: 'B'})  
CREATE (a)-[:CONNECTED {weight: 3}]->(b);
```

- A. Crea dos nodos con las etiquetas Node y los nombres A y B.
- B. Encuentra los nodos existentes A y B con la etiqueta Node y crea una relación CONNECTED entre ellos con un atributo weight igual a 3.
- C. Actualiza el peso de una relación CONNECTED existente entre los nodos A y B a 3.
- D. Elimina la relación existente entre A y B y la reemplaza con una nueva relación CONNECTED con peso 3.

**Pregunta 33.** ¿Qué hace la siguiente sentencia en Neo4j?

```
MATCH (n)-[r:CONNECTED]->(m)  
RETURN n, r, m;
```

- A. Crea una relación CONNECTED entre los nodos n y m y devuelve los nodos y la relación.
- B. Encuentra todos los nodos n y m conectados por una relación CONNECTED y los devuelve junto con la relación.
- C. Elimina todas las relaciones CONNECTED entre los nodos n y m y las devuelve.
- D. Actualiza las relaciones CONNECTED entre los nodos n y m y devuelve los cambios realizados.

**Pregunta 34.** ¿Qué hace la siguiente sentencia en Neo4j?

```
MATCH (n:Node {name: 'H'})  
RETURN id(n) AS nodeId;
```

- A. Crea un nodo con la etiqueta Node, nombre H, y devuelve su identificador único.
- B. Encuentra un nodo con la etiqueta Node y nombre H, y devuelve su identificador único bajo el alias nodeId.
- C. Devuelve todos los nodos con la etiqueta Node y sus identificadores únicos.
- D. Elimina un nodo con la etiqueta Node y nombre H, devolviendo su identificador único.

**Pregunta 35.** ¿Qué hace la siguiente sentencia en Neo4j?

```
YIELD path
```

- A. Crea un camino a modo de array del grafo y lo devuelve organizado para analizarse
- B. Devuelve los caminos generados por una consulta previa bajo el alias path.
- C. Encuentra todos los nodos que forman parte de un camino en el grafo.
- D. Elimina un camino existente en el grafo y lo devuelve.

## Parte 2. Ejercicios (4,75 puntos)

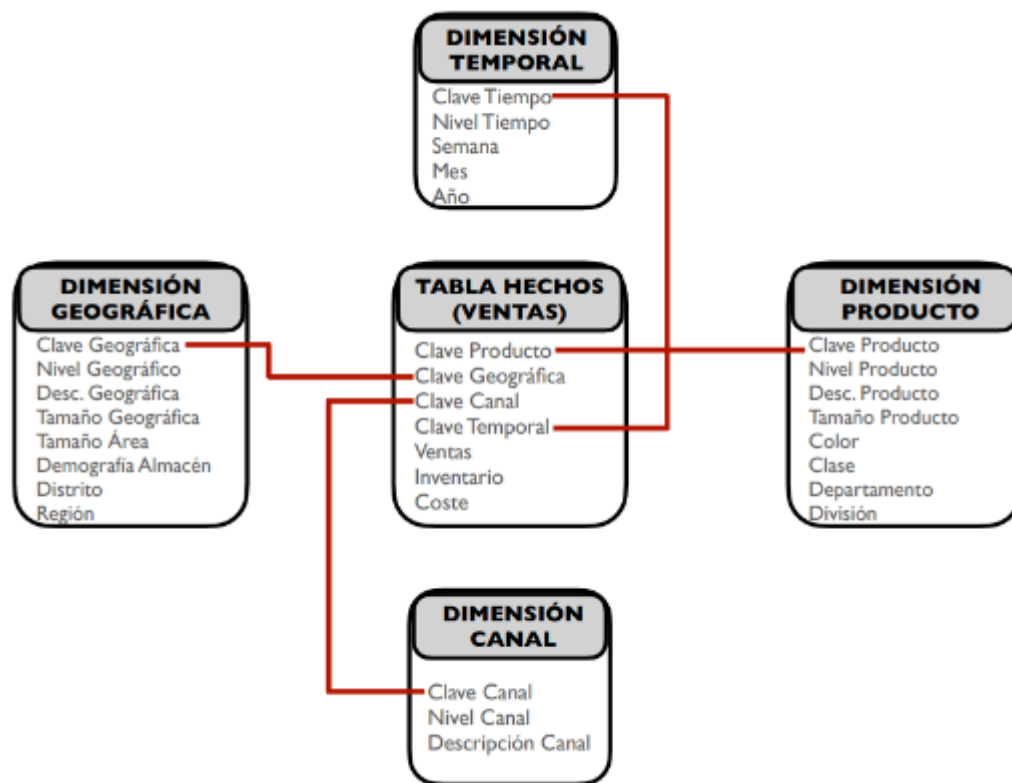
**Ejercicio 36 (1 punto).** Crear un JSON que represente la siguiente situación:

*Nueva suscripción a Netflix de María Ramírez Izquierdo. Inscripción estándar. Coste mensual 11,99€. Incluye la posibilidad de ver contenidos en dos pantallas al mismo tiempo, así como tener descargas en dos teléfonos o tabletas*

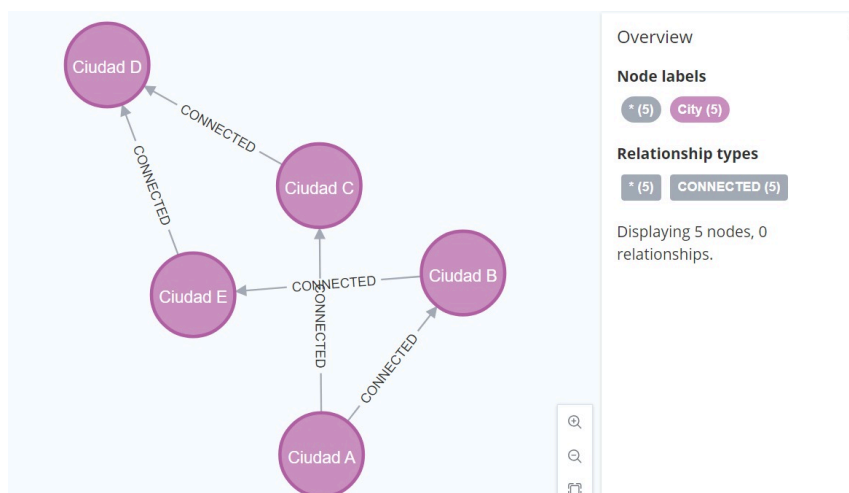
**Ejercicio 37 (0,75 puntos).** Escribe el código XML para la siguiente tabla:

PELÍCULAS		
Id_Película	Título	Año
1	Octopussy	1983
2	Hable con ella	2002
3	Ocho apellidos vascos	2014

**Ejercicio 38 (1 punto).** Explica qué es lo que se ve en la imagen, su diseño, ventajas y desventajas.



**Ejercicio 39 (1 punto).** Teniendo en cuenta el siguiente grafo:



Ejecutamos el siguiente comando para ejecutar un recorrido en profundidad empezando por el nodo A:

```
CALL gds.dfs.stream('myGraph1', { sourceNode: 0 })
YIELD path
UNWIND [ n in nodes(path) | n.name ] AS tags
RETURN tags;
```

Dibuja la tabla que se mostraría

**Ejercicio 40 (1 punto).** Con el grafo del ejercicio anterior, dibuja la tabla al ejecutar el siguiente código:

```
CALL gds.bfs.stream('myGraph1', { sourceNode: 0 })
YIELD path
UNWIND [ n in nodes(path) | n.name ] AS tags
RETURN tags;
```