

Alumno: Vargas Cristian Raul L.U. 39513

Tema 2:

Pautas para la evaluación

- El examen es individual y según el tema asignado en planilla.
- El examen se desarrolla de 16:00 a 18:45 horas, con 15 minutos adicionales (máximo) para preparar el documento final y completar el envío.
- Prever:
 - o Cámara del equipo deberá estar encendida si o si todo el tiempo.
 - o Tener a mano una identificación para mostrarla a la cámara cuando se le pida.
 - o Colocar su Nombre y Apellido en el ZOOM (no apodos, ni alias).
 - o Asegurarse en cuanto a la calidad de la/las fotos que envíen para realizar la evaluación (que sea legible).
- Durante todo el tiempo, los alumnos deberán estar conectados a una videoconferencia que se habilitará mediante Zoom (asegurarse de contar con el equipamiento y conexión de internet adecuados).
- Desarrollo completo, agregar copias de planillas electrónicas y fotos al documento que remitan.
- En el plazo establecido, convertir el archivo a formato PDF, renombrarlo con el formato BD2.Apellido.nombreSegundoParcial.pdf
- Enviar el archivo por correo electrónico, especificando en el asunto con el formato Apellido y Nombre – BD2 Segundo Parcial y utilizando la cuenta declarada en SIU Guaraní a la dirección: jose_a_carballo@hotmail.com CC a bd2.facena@gmail.com

Favor de esperar confirmación de recepción antes de desconectarse.

1- Que es un almacén de datos?

- a. Colección de datos de un dominio, integrada, no volátil X**
- b. Datos orientados a las transacciones**
- c. Datos orientados a las replicas**
- d. Datos orientados a la seguridad**
- e. Ninguna es correcta**

2- Cuales no son DSS (Sistemas de Soporte de Decisión)

- a. EIS**
- b. OLAP**
- c. OLTP X**
- d. Minería de datos**
- e. Ninguna es correcta**

3- Como se llama cuando un patrón donde uno o más atributos determina el valor de otro?

- a. Asociación**
- b. Dependencia X**
- c. Agrupamiento**
- d. Clasificación**
- e. Ninguna es correcta**

4- Que técnicas busca predecir valores de una variable continua a partir de otra continua?

- a. Tendencia / regresión X
- b. Correlación
- c. Dependencia
- d. Asociación
- e. Ninguna es correcta

5- Respecto a las herramientas de OLAP:

- A) Brindan técnicas para la integración y distribución de bases de datos.
- B) Presentan una visión tridimensional de los datos o esquema tridimensional, para cada actividad que es objeto de análisis.
- C) Se genera una consulta y la envía al Gestor de Consultas del sistema X
- D) Proporcionan independencia respecto a la red y al SGBD
- E) Ninguna es correcta.

6- En las herramientas OLAP, algunos de los operadores de refinamiento de manipulación de consultas, pueden ser:

- A) Pivot X
- B) Copy
- C) Drill X
- D) Delete
- E) Ninguna es correcta

7- Respecto a los sistemas ROLAP:

- B) Se implementa sobre tecnología relacional X
- C) Disponen de técnicas de compactación de datos
- D) Disponen de estructuras de almacenamiento específicas como Arrays
- E) Almacena físicamente los datos en estructuras multidimensionales de forma que la representación externa e interna coincidan.
- F) Ninguna es correcta.

8- Entre las áreas de aplicación de la minería de datos, podemos mencionar:

- A) Medicina X
- B) Astronomía X
- C) Políticas demográficas X
- D) Economía
- E) Ninguna es correcta

9- En el proceso KDD, algunas fases son:

- A) Minería de datos X
- B) Evaluación y validación X
- C) Normalización de datos
- D) Interpretación y difusión X
- E) Ninguna es correcta

10- Respecto a la fase de selección en KDD:

- A) Se implementan criterios de evaluación de hipótesis
- B) Comprende la recolección e integración de la información X
- C) Uso de ROLAP o MOLAP
- D) Se utiliza análisis exploratorio y grafico de los datos
- E) Ninguna es correcta

11- Respecto a las técnicas descriptivas y predictivas de clasificación de minería de datos, entre las técnicas del análisis cluster, podemos mencionar:

- A) Métodos solapados-exclusivos X
- B) Métodos directo-iterativos X
- C) Métodos ponderados-no ponderados X
- D) Métodos indirectos-iterativos
- E) Ninguna es correcta

12- Respecto a las técnicas descriptivas y predictivas de clasificación de minería de datos, entre las técnicas del análisis cluster no jerárquico, podemos agrupar los métodos como:

- A) Reasignación X**
- B) Reducción de dimensiones X**
- C) Indirectos**
- D) Búsqueda de la densidad X**
- E) Ninguna es correcta**

13) Que puede decir del trabajo de monografía: " Base de Datos Distribuida – Arquitectura Cliente/Servidor".

El trabajo de monografía Base de Datos Distribuida – Arquitectura Cliente/Servidor nos ofrece una introducción a las bases de datos distribuidas y nos explica que una DDB es una colección de múltiples bases de datos distribuidas interrelacionadas de forma lógica sobre una red de computadoras, y consta de un Sistema de administración de bases de datos distribuidas como el software encargado de administrar la base de datos distribuida.

Se explica que existen dos tipos de arquitecturas de sistema multiprocesador:

- Arquitectura de memoria compartida. Varios procesadores comparten el almacenamiento secundario y la memoria primaria.
- Arquitectura de disco compartido. Varios procesadores comparten el almacenamiento secundario (d, pero cada uno de ellos tiene su propia memoria primaria

Las Ventajas de las bases de datos distribuidas son:

- 1.** Administración de datos distribuidos con distintos niveles de transparencia.
 - a.** Transparencia de red o de distribución: Hace referencia a la autonomía del usuario de cómo opera la red.
 - b.** Transparencia de replicación. permite que el usuario no se entere de la existencia de copias.
 - c.** Transparencia de fragmentación. permite que el usuario no se entere de la existencia de fragmentos.
 - d.** La transparencia de diseño y de ejecución hace referencia poder saber cómo está diseñada la base de datos distribuida.
- 2.** Incremento de la fiabilidad y la disponibilidad. los datos o el software DBMS puede fallar, mientras el resto continúa operativo.
- 3.** Rendimiento mejorado. Un DBMS distribuido fragmenta la base de datos manteniendo la información lo más cerca posible del punto donde es más necesaria.
- 4.** Expansión más sencilla. El incremento del tamaño de las bases de datos o la incorporación de más procesadores es mucho más sencillo.

En cuanto a la arquitectura cliente-servidor de tres niveles se dan las siguientes capas:

- 1.** Capa de presentación (cliente). Proporciona al usuario la interfaz e interactúa con él. Los programas de esta capa presentan al cliente interfaces web o formularios que sirven como conexión con la aplicación
- 2.** Capa de aplicación (lógica de negocio). Esta capa programa la lógica de aplicación.
- 3.** Servidor de bases de datos. Esta capa controla las consultas y peticiones de actualizaciones procedentes de la capa de aplicación, procesa las solicitudes y envía los resultados.

Como conclusión el avance en las redes e internet y la arquitectura cliente-servidor hicieron posible la aparición de múltiples implementaciones de sistemas distribuidos y que actualmente son la tendencia a seguir

14) Sea la siguiente definición del esquema anidado Municipalidad:

Municipalidad = (Nombre, Dirección, Teléfono, Autoridades, Secretarías, Organigrama)

Autoridades = (Nombre autoridad, Función)

Secretarías = (Nombre secretaría, Oficinas, Correo electrónico)

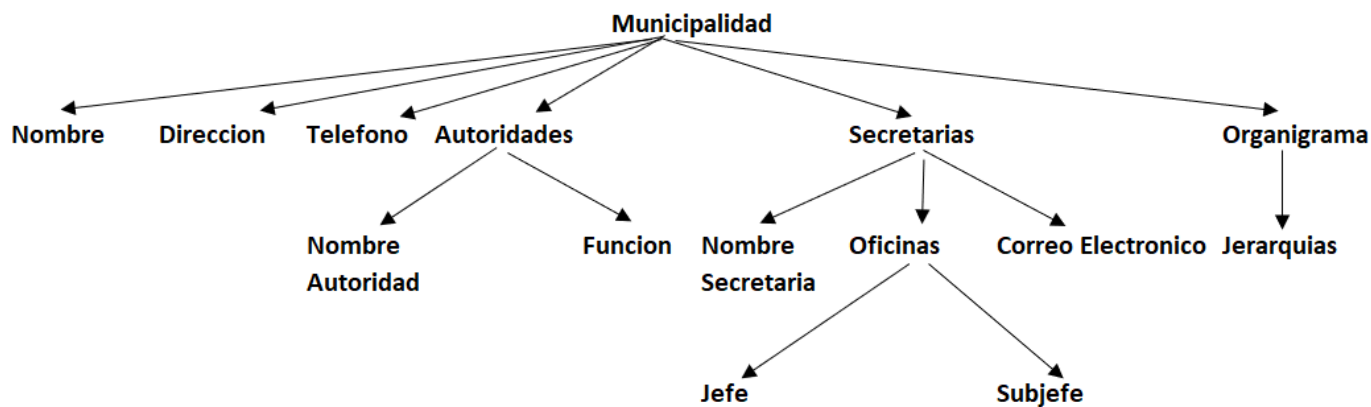
Oficinas = (Jefe, Subjefe)

Organigrama = (Jerarquías)

- a) Represente el esquema anidado Municipalidad con el formato de una tupla
- b) Realice la representación en árbol del esquema anidado Municipalidad

Municipalidad

Hemisferio Sur									
Nombre	Dirección	Teléfono	Autoridades		Secretarías			Organigrama	
			Nombre autoridad	Función	Nombre secretaría	Oficinas		Correo Electrónico	Jerarquías
						Jefe	Subjefe		



15) Considerando la siguiente consulta para el análisis de datos:

Categoría	Trimestre	Bolsas vendidas
Papas	Primero	30000
Papas	Segundo	40000
Papas	Tercero	20000
Papas	Cuarto	24000
Cebolla	Primero	60000
Cebolla	Segundo	10000
Cebolla	Tercero	16000
Cebolla	Cuarto	20000

- a) Disgregar en 2 nuevos grupos cada Categoría/Trimestre de la consulta original, considerando que las ciudades de Corrientes y Esquina conforman el universo estudiado, y las unidades vendidas se distribuyen en igual proporción en dichas ciudades.

Función Disgregación (Drill)

Categoría	Trimestre	Ciudad	Bolsas vendidas
Papas	T1	Corrientes	15000
Papas	T1	Esquina	15000
Papas	T2	Corrientes	20000
Papas	T2	Esquina	20000
Papas	T3	Corrientes	10000
Papas	T3	Esquina	10000
Papas	T4	Corrientes	12000
Papas	T4	Esquina	12000
Cebolla	T1	Corrientes	30000
Cebolla	T1	Esquina	30000
Cebolla	T2	Corrientes	5000
Cebolla	T2	Esquina	5000
Cebolla	T3	Corrientes	8000
Cebolla	T3	Esquina	8000
Cebolla	T4	Corrientes	10000
Cebolla	T4	Esquina	10000

b) Presente matricialmente los datos seleccionados.

Esquina				
Corrientes				
Papas	15000	20000	10000	12000
Cebolla	30000	5000	8000	10000
	T1	T2	T3	T4

c) Realice las operaciones necesarias para expresar la consulta obtenida en el punto a) en semestres.

Categoría	Semestre	Ciudad	Bolsas vendidas
Papas	S1	Corrientes	35000
Papas	S1	Esquina	35000
Papas	S2	Corrientes	22000
Papas	S2	Esquina	22000
Cebolla	S1	Corrientes	35000
Cebolla	S1	Esquina	35000
Cebolla	S2	Corrientes	18000
Cebolla	S2	Esquina	18000

d) Represente el punto b) luego de utilizar el operador tipo Pivot.

		Papas	
		Cebolla	
T4	10000	10000	
T3	8000	8000	
T2	5000	5000	
T1	30000	30000	
		Corrientes	Esquina

Puntajes:

Item 1 a 12	Bien contestada (solo una opción)	0,50 Puntos
	No contestada	0 "
Item 13	1 Punto	
Item 14 a 15	1,5 Puntos c/u	