



PAC2

Presentación

La PAC2 consiste en un conjunto de preguntas con el objetivo de consolidar los conocimientos teóricos de los módulos 1 y 2 de la asignatura.

Objetivos

- Diferenciar entre almacén de datos y base de datos operacional, identificando las principales diferencias.
- Conocer el contexto del almacén de datos: la factoría de información corporativa, sus componentes y la función de cada una de las partes.

Descripción de la PAC

Pregunta 1. Indicad si la siguiente afirmación es verdadera o falsa. Justificad brevemente la respuesta:

El Data Warehouse tiene como uno de sus principales objetivos ser el repositorio central de información corporativa que puede proveer distintos sistemas. Por este motivo, una de sus principales funciones es consolidar y homogeneizar la información procedente de los diferentes sistemas de la compañía, además de depurar y limpiar los datos, asegurando su calidad.

La afirmación es verdadera. Se puede encontrar en el punto 2.1 del módulo 1, donde se añaden otras características, como ser el punto central de información, donde se haya la versión más veraz de los datos, evitando así tener distinta información en función de la fuente consultada.



Pregunta 2. Indicad si la siguiente afirmación es verdadera o falsa. Justificad brevemente la respuesta:

Dados los problemas surgidos en la construcción de un Data Warehouse corporativo si se desarrolla en paralelo con el almacén de datos departamental, se recomienda construirlo a posteriori. Se debe no empezar la planificación de la FIC hasta que los almacenes de datos departamentales estén consolidados.

La afirmación es falsa. En las páginas 20 y 21 del módulo 2 se explica que lo correcto es planificar la construcción de la FIC desde el inicio. En caso contrario se requiere de un proceso de reingeniería sobre los diferentes componentes de integración y transformación y sobre los almacenes de datos departamentales.

Pregunta 3. Seleccionad la opción correcta:

- a) El almacén de datos operacional guarda todos los datos históricos de la empresa para satisfacer las necesidades de los analistas.
- b) Utilizando el almacén de datos departamental siempre quedan cubiertas con todas las necesidades de información de la empresa.
- c) El almacén de datos operacional y corporativo se complementan.**
- d) Todas las opciones anteriores son correctas
- e) Ninguna de las anteriores
- f) Las respuestas correctas son la a) y la c).

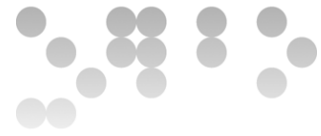
La afirmación A es falsa. Se puede comprobar en el punto 4.3 del módulo 1:

“El almacén de datos operacional es una estructura a caballo entre el mundo operacional y el de la toma de decisiones. Está orientado al tema e integrado como cualquier almacén de datos, pero en este caso no contiene ningún tipo de información temporal.”

La afirmación B es falsa. Se puede comprobar en el punto 4.3 del módulo 1

“Desgraciadamente, es posible que con los almacenes de datos departamentales y el corporativo todavía no tengamos cubiertas todas las necesidades de información de la empresa.”

La afirmación C es correcta. Se puede comprobar en el punto 4.3 del módulo 1



“De este modo, el almacén de datos operacional y el corporativo se complementan: el corporativo guarda todos los datos históricos, pero no está actualizado siempre, y el operacional siempre está actualizado, pero no contiene datos históricos.”

Pregunta 4. Seleccionad la opción correcta:

- a) La temática del almacén de datos departamental es específica.
- b) El volumen de datos del almacén de datos departamental es el mismo que el del almacén de datos corporativo.
- c) El tiempo de desarrollo del almacén de datos departamental es menos que el del almacén de datos corporativo.
- d) Todas las anteriores son correctas.
- e) Ninguna de las anteriores es correcta.

f) Las respuestas correctas son la a) y la c).

La afirmación A es correcta. Se puede ver en el apartado 4.1 del módulo 1

La afirmación B es falsa. Se puede ver en el apartado 4.3 del módulo 1

La afirmación C es correcta. Se puede ver en el apartado 4.2 del módulo 1

Tabla 1

Característica	Almacén de datos	
	Departamental	Corporativo
Temática	Específica	Genérica
Fuentes de datos	Pocas	Muchas
Tamaño	Gigabytes	Terabytes
Tiempo de desarrollo	Meses	Años
Modelo de datos	Multidimensional	Relacional



Pregunta 5. Seleccionad la opción correcta

- a) Un almacén de datos nunca contendrá entidades con información sensible, como por ejemplo: clientes, proveedores o productos.
- b) Resulta habitual que los datos maestros sean compartidos por distintos almacenes de datos departamentales.
- c) Los procesos de gestión de calidad de los datos maestros suelen tratar aspectos como: exactitud, integridad, consistencia y ser completas.
- d) Todas las anteriores son correctas.
- e) Ninguna de las anteriores es correcta.

f) Las respuestas correctas son la b) y la c).

La afirmación A es falsa. Se puede comprobar en el apartado 4.5 del módulo 1 y en la página 19 del módulo 2

“Generalmente todos los almacenes de datos tienen una serie de entidades críticas en cuanto a la información que contienen como pueden ser: clientes, Productos, proveedores o cuentas”

La afirmación es verdadera. Se puede comprobar en el apartado 4.5 del módulo 1

“Es común que dichas entidades sean compartidas por varios almacenes de datos departamentales y en ocasiones por sistemas no informacionales”

La afirmación es verdadera. Se puede comprobar en el apartado 4.5 del módulo 1

“Directamente relacionadas con las actividades de MDM están los procesos de seguimiento de la calidad del dato que se implementarán sobre los datos maestros y que permitirán monitorizar su calidad, revisando aspectos tales como la exactitud, integridad, consistencia y completitud.”

Pregunta 6. Marca qué opciones son características de un Data Warehouse

- a) Es perdurable.
- b) Garantiza la calidad de los datos.
- c) Satisface necesidades específicas de un área de información.
- d) Permite analizar la evolución temporal de los datos.
- e) Permite hacer análisis y predicciones con los datos.

Son ciertas las opciones a) – como sinónimo de no volátil –, c) como sinónimo de orientado al tema – y d)



Pregunta 7. En referencia al tiempo de respuesta, marca la/las opciones características de un Data Warehouse.

- a) Es importante optimizar los accesos con tal de permitir tiempos de respuesta razonables.
- b) Debe proporcionar una respuesta inmediata.
- c) El tiempo de respuesta no es relevante.
- d) Es más importante el tiempo de respuesta que la veracidad de los datos.
- e) Es admisible que el usuario reciba una respuesta de timeout si la consulta que realiza satura el servidor de base de datos.

Es verdadera la opción a)

Pregunta 8. Para las siguientes opciones, determinad si son características de un entorno operacional o informacional.

- a) La información está centralizada.
- b) Las actuaciones son automáticas y transaccionales.
- c) Sólo almacena información reciente.
- d) Su estructura está optimizada para la ejecución de consultas.
- e) Se utiliza para gestionar la operativa de la organización.

Son operacionales b), c) y e), informacionales a) y d)

Pregunta 9. Describe brevemente un nuevo ejemplo de base de datos operacional.

Basado en el caso Operador de Telecomunicaciones del módulo 2 (página 23)

Pregunta 10. Describe brevemente cómo el cambio en el grado de agregación puede afectar a la transformación de los datos.

El grado de agregación está relacionado con la dimensión o granularidad y determina el nivel de detalle de acceso a los datos.

Cuanto mayor sea, más grande será la capacidad analítica del modelo. Con un nivel de detalle alto se puede obtener también información de niveles más



agregados, pero no al revés. Por ejemplo, si la dimensión temporal contempla el día, se podrá obtener la información agregada por mes y año.

Recursos

Módulo 1: Introducción al Data Warehouse



Módulo 2: Construcción de la FIC

Criterios de valoración

Se valoraran todas las respuestas por igual, 1 punto.

Solo hay una respuesta correcta para cada pregunta.

Las respuestas incorrectas no descuentan puntos.

Formato y fecha de entrega

Para realizar la entrega es necesario enviar un único mensaje en el buzón de entrega de actividades.

En dicho mensaje debe adjuntarse un único archivo en formato Word o PDF con la solución a las preguntas.

El título del mensaje debe ser "Entrega PAC2".

En nombre del fichero debe ser la composición del nombre del usuario con el: "_DW_PEC2 " (por ejemplo: si el nombre de usuario es "jregolfq", entonces el archivo debe llamarse "jregolfq_DW_PAC2 ").

La entrega puede realizarse hasta el día marcado en el plan docente

Nota: **Propiedad intelectual**

Al presentar una práctica o PEC que haga uso de recursos ajenos, se tiene que presentar junto con ella un documento en que se detallen todos ellos, especificando el nombre de cada recurso, su autor, el lugar donde se obtuvo y su estatus legal: si la obra está protegida por el copyright o se acoge a alguna otra licencia de uso (Creative Commons, licencia GNU, GPL etc.). El estudiante tendrá que asegurarse que la licencia que sea no impide específicamente su uso en el marco de la práctica o PEC. En caso de no encontrar la información correspondiente tendrá que asumir que la obra está protegida por el copyright.

Será necesario, además, adjuntar los ficheros originales cuando las obras utilizadas sean digitales, y su código fuente, si así corresponde.