

M2.863

Diseño y construcción del data warehouse PRAC2

Objetivos

El objetivo de esta actividad consiste en construir el data warehouse que se ha definido en la anterior entrega de la práctica, aplicando los conocimientos sobre diseño multidimensional de datos estudiados en el módulo 4. Básicamente en esta entrega se pretende realizar la implementación del sistema de la FIC.

A partir del análisis de las fuentes de datos, el estudiante debe diseñar, implementar y ejecutar los procesos de extracción, transformación y carga de los datos existentes en los ficheros proporcionados, que contienen información operativa.

Se pide:

La segunda parte de la práctica consiste en:

- Revisión del diseño conceptual, lógico y físico del modelo multidimensional realizado en la Práctica 1 (solución oficial).
- Identificar y diseñar los procesos de extracción, transformación y carga de datos desde el sistema operacional o fuentes de datos, hacia el almacén de datos.
- Implementar los procesos de carga (ETL) mediante las herramientas proporcionadas.
- Realizar la carga de datos de forma efectiva.

Descripción:

Se recomienda revisar y validar la calidad de los datos de los ficheros proporcionados, pudiendo existir cuestiones a depurar en los mismos. También se deberá tener en cuenta que ésta es una carga inicial del almacén de datos.

Se evaluará la implementación sobre la Máquina Virtual, por ese motivo se debe desarrollar la PRAC2 en ella.

Se podrá modificar el almacén de datos diseñado en la solución oficial de la PRAC1, siempre que esté justificado o para la implementación del versionado de las dimensiones.

La solución de la PRAC debe incluir:

- Descripción de todas las acciones realizadas en el proceso.
- Capturas de pantalla de todas las partes del ETL, sus características y su correspondiente explicación.
- Capturas de pantalla que demuestren la correcta ejecución de la ETL y el tiempo de ejecución.
- Capturas que demuestren la correcta carga de datos (datos cargados en la BBDD).

Además de la correcta implementación se valorará también la simplicidad, repetitividad, tratamiento de errores y uso correcto de la herramienta proporcionada (PDI).

Formato y fecha de entrega:

Para hacer la entrega se debe enviar un único mensaje al buzón de entrega de actividades. En este mensaje se debe adjuntar un único documento en formato Word o pdf con la solución de la PRAC. El nombre del archivo debe ser la composición del nombre de usuario y “_DW_PRAC2” (por ejemplo: si el nombre de usuario es “bantich”, el nombre del archivo debe ser “bantich_DW_PRAC2.pdf”).

La fecha máxima de entrega es el 19/12/2018 a las 23:59 horas.