
MÁSTER EN BIG DATA & BUSINESS ANALYTICS

MÓDULO V: INTELIGENCIA DE NEGOCIO Y VISUALIZACIÓN

Documentación del caso práctico

de Octubre del 2018

Autor

José María Sánchez Salas
josemaria.sanchezsalas@gmail.com

Índice

Módulo V: Inteligencia de Negocio y Visualización

1	Introducción	5
2	Enunciado	5
3	Solución	6
3.1	Análisis de fuentes	6
3.2	Análisis funcional y diagrama de arquitectura de flujo de datos	6
3.3	¿Qué arquitectura de referencia usaría? Justifique la respuesta	6
3.4	¿Qué tecnología OLAP usaría? Justifique la respuesta	6
3.5	Si se utiliza ROLAP, ¿cuál de estos dos modelos se ajustaría mejor: modelo de estrella o el de copo de nieve?	6
3.6	Si se utiliza ROLAP, hay que identificar y justificar si existe algún proceso de desnormalización de información que se deba realizar	6
3.7	Si se utiliza ROLAP, se debe incluir un diseño conceptual a modo explicativo junto con un diagrama	6
3.8	Si se utiliza ROLAP, se debe incluir un diseño modelo lógico	6
3.9	Si se utiliza ROLAP, se debe incluir un diseño modelo físico	6
3.10	Realizar la implementación del proceso ETL para generar y poblar el modelo multidimensional diseñado en los apartados anteriores	6
3.11	Implementación de modelo multidimensional diseñado mediante los puntos anteriores	6
3.12	Análisis de modelo	6
4	Conclusiones	6

1 Introducción

Este documento contiene la explicación de mi solución propuesta al caso práctico del Módulo V: Inteligencia de Negocio y Visualización, del M. Big Data & Business Analytics impartido por IMF Business School. El documento contiene dos secciones: uno para cada uno de los ejercicios del caso práctico. En cada ejercicio, se muestra previamente su enunciado y posteriormente la solución propuesta.

2 Enunciado

El departamento antifraude de una compañía de Mystery Shopping desea hacer un seguimiento y analizar la información relativa a las encuestas que realiza en los distintos centros de sus clientes. Para ello, el cliente solicita:

- Un análisis y diseño del Data Warehouse que daría respuesta a los usuarios analíticos del departamento antifraude, suponiendo que los usuarios aún no tienen claro el tipo de análisis que quieren realizar.
- Partiendo del análisis y diseño previo realizado y usando Pentaho Data Integration, se debe realizar la implementación del proceso ETL con el objetivo de:
 - Identificar y extraer los datos de las fuentes.
 - Procesar los datos y aplicar procesos de limpieza y calidad del dato.
 - Generar y cargar los datos en el modelo físico de estrella identificado en la fase de diseño.
- Posteriormente, partiendo del análisis y diseño previo realizado y conociendo ya la tecnología seleccionada, en este caso Pentaho Business Analytics, ha de realizarse una implementación ágil del modelo multidimensional.

El objetivo en este caso es la implementación del modelo multidimensional sobre diseño del Data Warehouse que daría respuesta a los usuarios analíticos del departamento antifraude, suponiendo que los usuarios aún no tienen claro el tipo de análisis que quieren realizar.

3 Solución

- 3.1 Análisis de fuentes
- 3.2 Análisis funcional y diagrama de arquitectura de flujo de datos
- 3.3 ¿Qué arquitectura de referencia usaría? Justifique la respuesta
- 3.4 ¿Qué tecnología OLAP usaría? Justifique la respuesta
- 3.5 Si se utiliza ROLAP, ¿cuál de estos dos modelos se ajustaría mejor: modelo de estrella o el de copo de nieve?
- 3.6 Si se utiliza ROLAP, hay que identificar y justificar si existe algún proceso de desnormalización de información que se deba realizar
- 3.7 Si se utiliza ROLAP, se debe incluir un diseño conceptual a modo explicativo junto con un diagrama
- 3.8 Si se utiliza ROLAP, se debe incluir un diseño modelo lógico
- 3.9 Si se utiliza ROLAP, se debe incluir un diseño modelo físico
- 3.10 Realizar la implementación del proceso ETL para generar y poblar el modelo multidimensional diseñado en los apartados anteriores
- 3.11 Implementación de modelo multidimensional diseñado mediante los puntos anteriores
- 3.12 Análisis de modelo

4 Conclusiones