

ESCUELA POLITÉCNICA NACIONAL
MAESTRÍA EN CALIDAD DE SOFTWARE
PROYECTO DE INTELIGENCIA DE NEGOCIOS

TEMA: Diseño e Implementación de un sistema de Data Warehousing

OBJETIVO GENERAL: Diseñar un sistema de data warehousing para dar solución al caso de estudio de Ecuavinos

OBJETIVOS ESPECIFICOS:

- Diseñar un esquema estrella para almacenar datos multidimensionales de los sistemas de Ecuavinos
- Implementar procesos ETL que permitan extraer, transformar y cargar datos automáticamente en el repositorio centralizado (data warehouse)
- Desarrollar un diccionario de datos del esquema estrella diseñado
- Presentar reportes que permitan visualizar la información cargada al data warehouse

CASO DE ESTUDIO DE ECUAVINOS

Ecuavinos es una bodega de tamaño medio situada en el valle de Guayllabamba y se especializa en vinos de alta calidad para lo cual produce tres variedades: pinot noir, merlot (ambos tintos), y pinot grigio (vino blanco). Tanto los vinos tintos como los blancos se venden actualmente en Quito, y en algunas provincias del país e internacionalmente en el Reino Unido y Europa. La producción y las ventas de Ecuavinos ha crecido dramáticamente en los últimos 2 años, y el gerente general, Fabricio Carrera, cree que cualquier crecimiento futuro necesita ser mejor manejado basándose en una mejor comprensión de las tendencias de ventas. Últimamente, ha escuchado de los vendedores que el análisis de datos y de negocios le proporcionará la capacidad de tomar decisiones más informadas.

Usted ha sido contratado como consultor para proponer una solución de data warehousing a Fabricio.

Por lo tanto, es necesario considerar que las operaciones de Ecuavinos incluyen: el cultivo de las uvas, la producción del vino y su venta. Actualmente hay 30 hectáreas de pinot noir, 30 hectáreas de merlot y 25 hectáreas de pinot grigio. La siembra adicional es posible, pero implica una planificación a largo plazo.

El enólogo Gustavo Cifuentes, dueño de Ecuavinos, se dedica especialmente a la producción de vino de calidad y por lo tanto es posible vender los vinos de Ecuavinos a precios superiores. Tres tipos de pinot noir son embotellados y vendidos bajo marcas separadas y se producen dos tipos de pinot grigio y dos tipos de merlot. Para aumentar el volumen de producción, Cifuentes compra uvas de alta calidad adicionales de los viñedos. Una pequeña cantidad de vino se vende directamente a través de una pequeña tienda de vinos ubicada cerca de Guayllabamba (ventas al por menor), pero casi todas las ventas son a los clientes comerciantes de vino en Quito, provincialmente y cada vez más al Reino Unido y Europa (ventas al por mayor). Las ventas a los comerciantes de vino se hacen por cartones de 12 botellas cada uno y las ventas al por menor, se hacen por botella.

La fábrica cuenta con dos sistemas de información separados para ayudar a administrar la producción y las ventas a los comerciantes de vino. La información de estos sistemas es difícil

de agregar ya que tienen diferentes formatos de datos y utilizan diferentes sistemas de gestión de bases de datos o también denominados Relational Database Management Systems(RDBMS).

PROBLEMAS DEL NEGOCIO

Las decisiones sobre el crecimiento del negocio, en términos de los volúmenes de producción del vino tinto y blanco, qué clientes son los más importantes y los mercados en los que centrarse han sido hechas por Fabricio, con la ayuda de Gustavo. Las decisiones que se han tomado hasta ahora han sido en base a la experiencia y al instinto, lo cual ha sido efectivo antes, pero ahora que el negocio ha crecido, Fabricio quiere evidencia basada en datos para tomar decisiones mejor informadas relativas a tres problemas importantes del negocio: clientes, productos y mercados.

1. ¿Quiénes son los clientes clave?

El almacén de datos debe proporcionar información sobre ventas unitarias, ventas en dólares, costo y utilidad para cada cliente (sólo comerciantes), para cada producto (producto base y tipo de vino), para varios períodos de tiempo (incluyendo año, trimestre y mes).

[Costo = promedio de la producción de un cartón de vino]

[Utilidad = Ventas en dólares - Costo]

2. ¿Qué productos son los más rentables?

El almacén de datos debe proporcionar información sobre las ventas unitarias y las ventas en dólares, el costo y la utilidad de cada producto (producto base y tipo de vino), para varios periodos de tiempo (incluyendo año, trimestre y mes).

3. ¿Qué mercado es el más rentable?

El almacén de datos debe proporcionar información sobre las ventas unitarias y las ventas en dólares para cada mercado para cada mes del año anterior. Actualmente hay tres mercados (Quito, Ecuador e Internacional), pero esto puede cambiar en el futuro. Los mercados se definen por la dirección de entrega en la orden de venta pues algunos clientes son globales.

SISTEMAS DE INFORMACIÓN DE ECUAVINOS

Los datos para el almacén de datos se obtendrán de dos de los sistemas de operaciones existentes: el sistema de producción y el sistema de ventas comerciales. Las ventas directas en Guayllabamba tienen volúmenes bajos y no se incluirán en el almacén de datos.

1. Sistema de Producción

El sistema de producción es un paquete basado en una base de datos Oracle y se ejecuta en una computadora basada en Windows. Un código de producto de 5 dígitos se utiliza para identificar productos de nivel base. También se pueden definir grupos de productos. El sistema admite también el historial de costes del producto.

Ejemplos de datos en las tablas relevantes de Oracle se muestran a continuación

PRODUCTO (datos de ejemplo)

Código	Descripción	Grupo
12765	McDonell Pinot Noir	Red
12766	Mornington Pinot Noir	Red
12767	Downunder Pinot Noir	Red
12821	Mornington Merlot	Red
14821	Downunder Merlot	Red
14823	Mornington Pinot Grigio	White
14827	Downunder Pinot Grigio	White

HISTORIAL DE PRODUCCIÓN (datos de ejemplo, registros incompletos)

Código	Descripción	Año	Volumen de Producción	Costo por docena
12765	McDonell Pinot Noir	2013	600	\$120
12765	McDonell Pinot Noir	2012	580	\$110
12765	McDonell Pinot Noir	2011	510	\$90
14823	Mornington Pinot Grigio	2013	400	\$70
14823	Mornington Pinot Grigio	2012	250	\$65
14821	Mornington Merlot	2013	550	\$100
14821	Mornington Merlot	2012	400	\$100
12767	Downunder Pinot Noir	2013	780	\$80
12767	Downunder Pinot Noir	2012	690	\$85
14827	Downunder Pinot Grigio	2013	440	\$70

2. Sistema de Ventas a Comerciantes

El sistema de ventas comerciales es un paquete basado en una base de datos de SQL Server que también se ejecuta en un equipo basado en Windows. Los datos de las órdenes de compra, productos y clientes se incluyen en el sistema. Ejemplos de datos en tablas relevantes de SQL Server (Cliente y Producto) y formularios de órdenes de compra se muestran a continuación:

CLIENTE (datos de ejemplo, incompleto)

Cliente ID	Nombre	Dirección	Código Postal	Tipo
C478W	Prestige Wines	Av. 12 de octubre, Quito.	3053	Wholesale (ventas al por mayor)
C567R	Acme Wine Imports	High St, Fullham, London	SW6	Retail (ventas al por menor)
C121R	Oz Wines	Av. Patria, UIO.	3121	Retail
C478W	The Wine Club	Po Box 184 Av. Gaspar villarroel, Quito	3051	Wholesale
C128R	London Wines	The Strand, London	EC4	Retail
C342W	International Wines	PO Box 324, Paris	75008	Wholesale

PRODUCTO (datos de ejemplo, incompleto)

<i>Producto ID</i>	<i>Descripción</i>	<i>Precio por unidad (docena)</i>	<i>Año</i>
P12766	Mornington Pinot Noir	\$180	2013
P12766	Mornington Pinot Noir	\$150	2012
P14823	Mornington Pinot Grigio	\$140	2013
P14823	Mornington Pinot Grigio	\$120	2012
P12767	Downunder Pinot Noir	\$125	2013
P12767	Downunder Pinot Noir	\$115	2012

ORDEN DE COMPRA (datos de ejemplo, incompleto)**Orden de Ventas** S135 12/Feb/2014**Cliente** C478 The Wine Club**Dirección de entrega** The Wine Club, Av. América E11-257, Quito

<i>Producto</i>	<i>Cantidad (docena)</i>	<i>Precio (por docena)</i>	<i>Precio Total</i>
McDonell Pinot Noir 2013	25	\$160	\$4.000
Mornington Pinot Grigio 2012	30	\$110	\$3.300

Orden de Compra S140 15/Feb/2014**Cliente** C128 London Wines

<i>Producto</i>	<i>Cantidad (docena)</i>	<i>Precio (por docena)</i>	<i>Precio Total</i>
Downunder Pinot Noir 2013	30	\$80	\$2.400

Orden de Compra S168 16/Feb/2014**Cliente** C478 The Wine Club**Dirección de Entrega** The Wine Club, Fleet Street, London

<i>Producto</i>	<i>Cantidad (docena)</i>	<i>Precio (por docena)</i>	<i>Precio Total</i>
Downunder Pinot Noir 2012	30	\$95	\$2.850
Downunder Pinot Grigio 2013	24	\$100	\$2.400

ENTREGABLES

1. Diseño del esquema en estrella que responda a las 3 preguntas planteadas en el problema del negocio junto con su explicación de cada criterio considerado en el diseño. En esta fase es importante indicar todos los criterios de diseño y su justificación respectiva de acuerdo a los pasos indicados en clase
2. Diccionario de datos del esquema en estrella en donde se indique la o las fuentes de datos (tablas o vistas) de los sistemas transaccionales, el significado de cada campo en cada tabla, sus relaciones, etc.
3. Arquitectura de la solución (diagrama) donde se especifiquen claramente los pasos que siguieron para extraer datos de la/s fuentes de datos, el proceso ETL, el repositorio central y la visualización de los datos.
4. Implementación del esquema en estrella utilizando alguna herramienta de data warehousing.

INFORME ESCRITO

1. Máximo de palabras

El documento escrito debe contener un máximo de 2500 palabras más o menos el 5%. Se excluyen tablas, figuras, referencias y apéndices en formato paper.

2. Presentación del documento

Contenido esperado del trabajo escrito

- **Introducción**
- **Método**
- **Resultado**
- **Conclusiones y trabajo futuro**
- Anexos (diccionario de datos)

El documento escrito deberá ser entregado a través de la herramienta de detección de plagio Turnitin en un enlace abierto en el aula virtual.

3. Penalidades por entrega atrasada de trabajos

Con el fin de garantizar igualdad para todos los estudiantes, las tareas deben ser completadas dentro del plazo especificado. Las presentaciones tardías significarán un 20% de reducción en la nota por cada día retraso. Los trabajos entregados después de dos días de la fecha final, ya no serán receptados y tendrán una nota de cero.

Los trabajos que excedan los límites de palabras en un 5%, incluyendo notas al pie, tendrán una penalidad del 10% de la nota total.

Los trabajos que excedan el límite de palabras en un 25% o más tendrán una penalidad del 25% de la nota total.

4. Plagio

El material de otras fuentes presentación sin el pleno reconocimiento (plagio) está fuertemente penalizado. Las sanciones por plagio pueden incluir una calificación de cero en el trabajo o una calificación de cero al semestre dependiendo de la gravedad de la acción.

5. Criterios de Calificación

Criterios de calificación escrito	Posible calificación
Introducción, presentación, ortografía y gramática	5
Método	10
Resultados y análisis	5
Conclusiones	5
Diccionario de datos	5
Total	30

PRESENTACIÓN

Cada pareja tendrá 15 minutos para realizar la exposición de su solución. En la presentación de debe evidenciar el funcionamiento de la solución y se realizarán pruebas para verificar las funcionalidades solicitadas.

Criterios de calificación presentación	Posible calificación
Explicación clara de los criterios de diseño del esquema	5
Explicación clara de problemas y soluciones	5
Manejo de recursos ¹	5
Interacción con el público	5
Respuesta correcta a preguntas	5
Organización y estructura de la presentación.	5
Manejo apropiado del tiempo	5
Total	35

FECHA DE ENTREGA Y PRESENTACIÓN: Viernes 05 de enero del 2018 a partir de las 18:00

1

http://www.eoi.es/wiki/index.php/El_manejo_de_recursos_en_la_presentaci%C3%B3n_en_Direcci%C3%B3n_de_personas_y_habilidades