

Proyecto 1 de Inteligencia de Negocios

Tema: Diseño e Implementación de un sistema de Data Warehousing

José Limaico - Jairo Vera



Tabla	de	conte	nido
-------	----	-------	------

1.	Obj	jetivos	3
1	.1.	Objetivo General	3
1	.2.	Objetivos Específicos	3
2.	Intr	oducción	3
3.	Mé	todo	3
3	.1.	Diseño Del Esquema en Estrella	3
	3.2	¿Quiénes son los clientes clave?	5
	3.3	¿Qué productos son los más rentables?	6
	3.4	¿Qué mercado es el más rentable?	7
	3.5	Dimensión DimDate [1]	8
	3.6	Tabla de Hecho FactVentas	9
	3.7	Esquema Estrella del almacén de datos	10
4.	Dic	cionario de Datos	11
5.	Arc	quitectura	11
6.	Cre	eación del cubo OLAP	12
6	.1.	Creación de la Vista de Datos	13
6	.2.	Creación del Cubo OLAP	15
6	.3.	Configurar los atributos y jerarquías de las dimensiones	17
	6.3	.1. DimDate	17
	6.3	.2. DimCliente	17
	6.3	.3. DimPuntoVenta	18
	6.3	.4. DimProducto	19
7. Pow	-	olementación del Esquema Estrella utilizando la base de datos dimensional co vot	
8.	Res	sultado	22
8	.1.	¿Quiénes son los clientes clave?	22
8	.2.	¿Qué productos son los más rentables?	
8	.3.	¿Qué mercado es el más rentable?	23
9.	Dis	cusión	
10.	C	onclusiones y Trabajo Futuro	24
11.		erencias	
12.	A	nexos	25
1	2.1.	Diccionario de Datos de Ecuavinos_DW	25



1. Objetivos

1.1. Objetivo General

 Diseñar un sistema de data warehousing para dar solución al caso de estudio de EcuaVinos.

1.2. Objetivos Específicos

- Diseñar un esquema estrella para almacenar datos multidimensionales de los sistemas de EcuaVinos.
- Desarrollar un diccionario de datos del esquema estrella diseñado.
- Presentar reportes que permitan visualizar la información cargada al data warehouse

2. Introducción

EcuaVinos es una mediana empresa situada en el valle de Guayllabamba especializada en la elaboración de vinos de alta calidad. Produce tres variedades: pinot noir y merlot (vino tinto), y pinot grigio (vino blanco). Tres tipos de pinot noir son vendidos bajo marcas separadas y se producen dos tipos de pinot grigio y dos tipos de merlot. Actualmente, las tres variedades principales de vino se venden dentro del país (en Quito y en algunas provincias) e internacionalmente en el Reino Unido y Europa. La producción y las ventas de EcuaVinos han crecido dramáticamente en los últimos 2 años, por lo tanto, el gerente general de la empresa Fabricio Carrera, cree que la existencia de cualquier crecimiento futuro necesita ser manejado basándose en una mejor comprensión de las tendencias de crecimiento de ventas.

Adicionalmente, el dueño de la empresa Gustavo Cifuentes considera que se pueden vender los vinos a precios superiores debido a que casi todas las ventas se las realiza a comerciantes de vino en Quito y también a comerciantes del Reino Unido y Europa, siendo estos lugares en donde las ventas han aumentado últimamente. En la actualidad la empresa cuenta con dos sistemas separados de bases de datos relacionales que se encargan de gestionar la producción y ventas de los vinos.

Por los antecedentes mencionados se necesita crear un sistema de data warehouse que permita administrar tanto la producción como las ventas de los vinos de manera unificada para proveer información útil que permita al gerente de la empresa tomar decisiones para cualquier crecimiento futuro basándose en el análisis del crecimiento de ventas y de producción.

3. Método

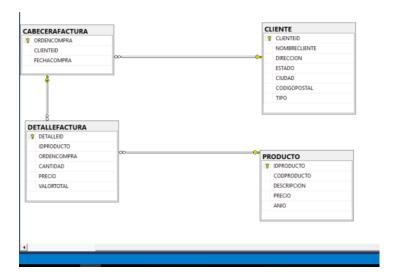
3.1. Diseño Del Esquema en Estrella

Para el desarrollo del data warehouse se tomaron en cuenta las decisiones sobre el crecimiento del negocio, en términos de los volúmenes de producción del vino tinto y blanco, la



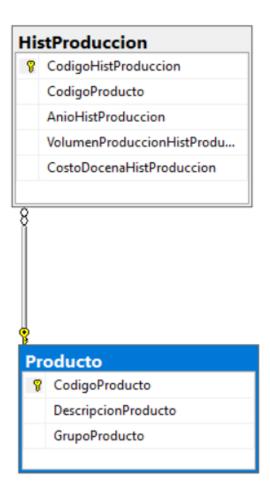
determinación de los clientes más importantes y los mercados en los que debe centrarse Ecuavinos para su crecimiento económico.

Por lo tanto, a partir de los sistemas tanto de producción como de ventas de la empresa implementados en los motores de bases de datos Oracle y SQL Server respectivamente podemos obtener la información necesaria para realizar el esquema Estrella y poder responder a las tres preguntas planteadas en los requerimientos.



Esquema de la base de datos Ventas de Ecuavinos



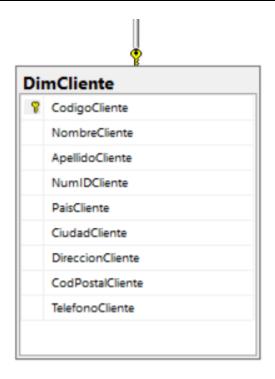


Esquema de la base de datos de Producción de Ecuavinos

3.2. ¿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 varios períodos de tiempo (incluyendo año, trimestre y mes). Por lo tanto, se debe incluir una dimensión Cliente que proporcione toda esta información.

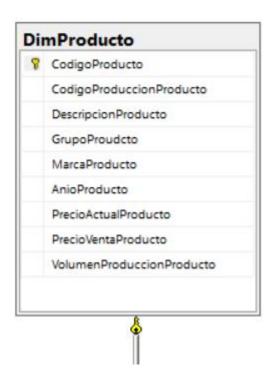




3.3. ¿Qué productos son los más rentables?

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). Por lo tanto, en el almacén de datos se debe incluir una dimensión producto que abarque la información tanto de la producción de los productos como de las ventas de los mismos.





3.4. ¿Qué mercado es el más rentable?

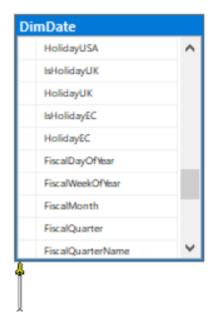
El almacén de datos debe proporcionar información sobre las ventas unitarias y las ventas en dólares para cada mercado para cada mes del año anterior. Debido a que actualmente Ecuavinos tiene tres mercados los cuales son: Quito, Ecuador e Internacional definidos por la dirección de entrega de la orden de venta. Por lo tanto, se debe crear una dimensión Punto de Venta para que contenga información útil con respecto a la ubicación geográfica de los clientes que pueda ser utilizada más adelante.





3.5. Dimensión DimDate [1]

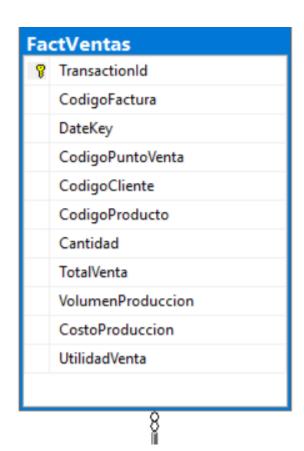
Debido a que se necesita presentar la información de manera específica por un periodo de tiempo se debe crear una dimensión DimDate, la cual permite obtener la información de manera ordenada por fechas en las que se han realizado las ventas.





3.6. Tabla de Hecho FactVentas

La granularidad de la tabla de hechos es a nivel de ítem debido a que se necesita información por cada transacción de venta que se realiza en la empresa. Las métricas utilizadas son: cantidad, costo de producción, costo de venta, y utilidad.



3.7. Esquema Estrella del almacén de datos

El esquema estrella está conformado por cuatro tablas de dimensión y una tabla de hechos obtenidas a partir de las entidades pertenecientes a las bases de datos de producción y Ventas.



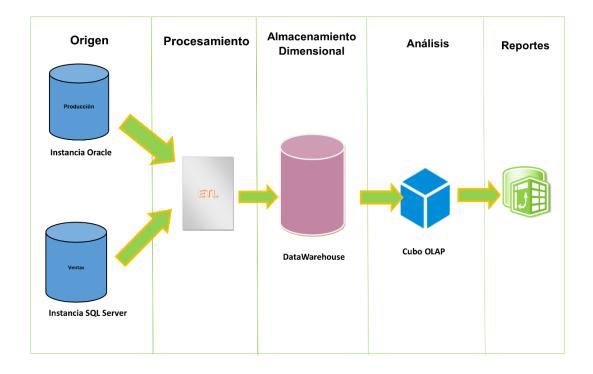


Tabla componente Dimension			
Producto DimProducto			
Cliente	liente DimCliente		
- DimPuntodeVenta			
Tabla Transaccion Tabla de Hecho			
Detalle Factura FactVentas			
Dimension Temporal			
DimDate			

4. Diccionario de Datos

Ver Anexo 1

5. Arquitectura



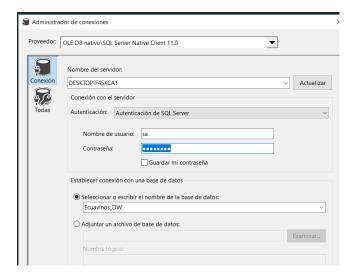
La arquitectura del almacén de datos consta de cinco fases:



- Origen
- Procesamiento
- Almacenamiento Dimensional
- Análisis
- Reportes

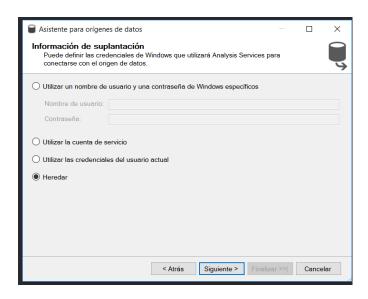
6. Creación del cubo OLAP

Primero hay que crear un nuevo origen de datos para lo cual debemos crear una nueva conexión a la base de datos con el administrador de conexiones.



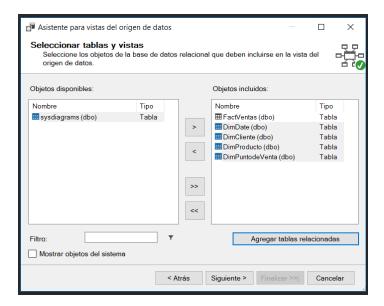
Luego, debemos seleccionar la opción Heredar





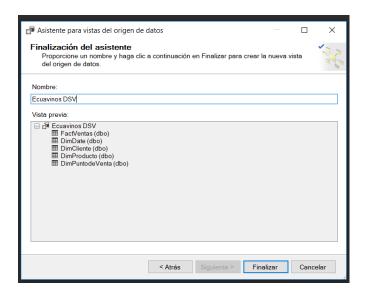
6.1. Creación de la Vista de Datos

Primero se debe seleccionar la tabla de hechos y pasarla a la sección de Objetos incluidos, luego hacer clic en el botón "Agregar tablas relacionadas".

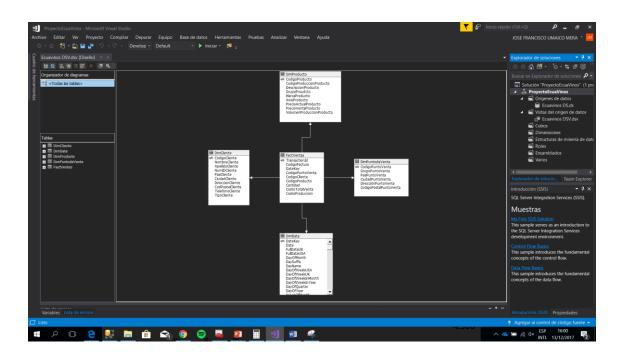




Luego, seleccionar un nombre para la vista de origen de datos y dar clic en finalizar.

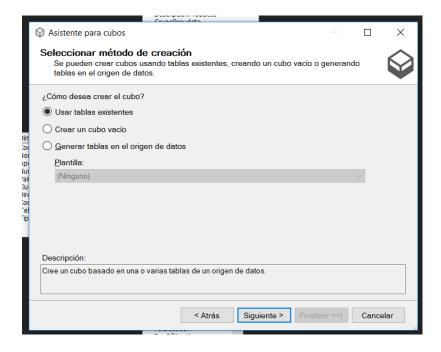


Finalmente, el diseño obtenido es el siguiente:

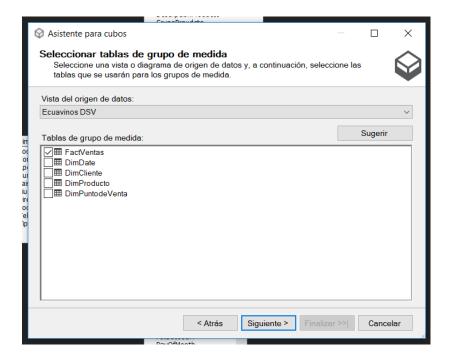


6.2. Creación del Cubo OLAP

En el asistente se debe escoger la opción usar tablas existentes

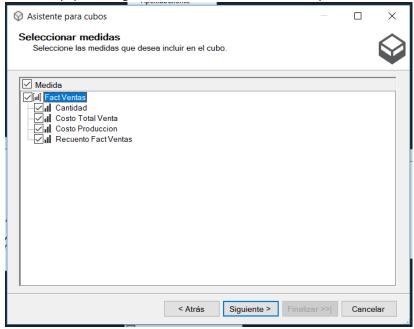


Luego de dar clic en siguiente se debe escoger la tabla de hechos desde la vista de origen de datos

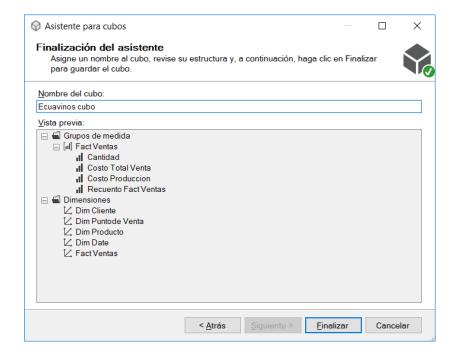




Ahora hay que escoger todas las dimensiones del esquema estrella creado previamente



Finalmente escoger un nombre para el cubo y dar clic en finalizar

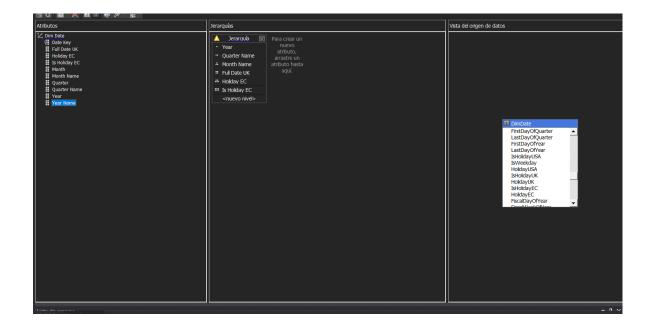




6.3. Configurar los atributos y jerarquías de las dimensiones

6.3.1. DimDate

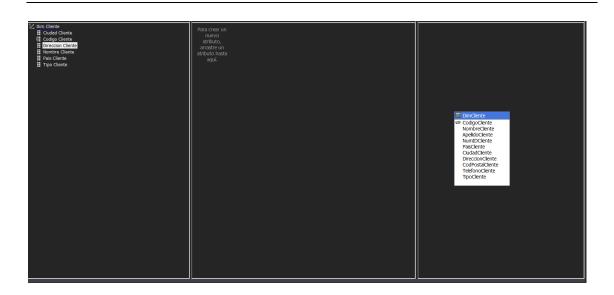
Para la dimensión Date se deben escoger los campos como se muestra en la figura debido a que los requisitos para la construcción del almacén de datos son reportes por año, trimestre y mes.



6.3.2. DimCliente

Para la dimensión cliente se deben escoger los atributos como se muestra en la figura para poder realizar el análisis de datos requerido.





6.3.3. DimPuntoVenta

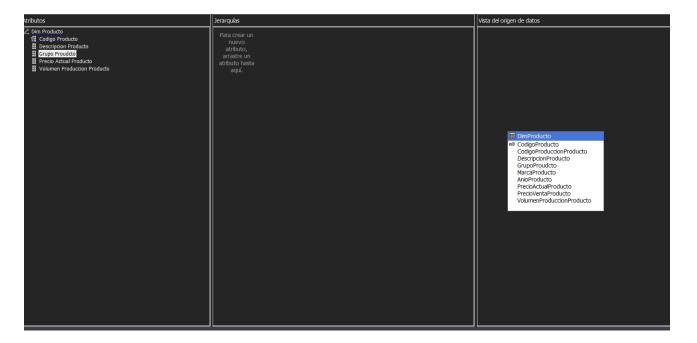
La dimensión Punto de Venta debe contener los siguientes atributos incluyendo el atributo de dirección debido a que se necesita información específica del lugar de procedencia de los compradores para saber que mercados de los existentes son los más lucrativos.



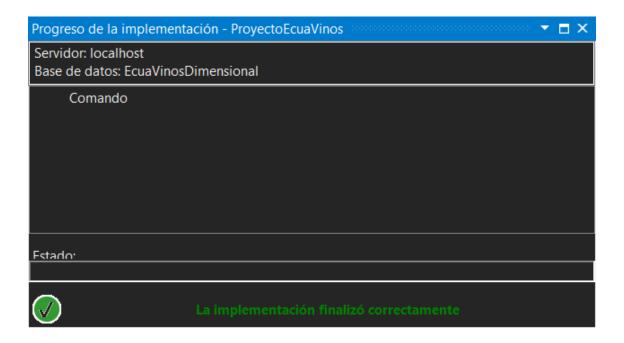


6.3.4. DimProducto

La información del producto debe contener los atributos que indiquen tanto el costo de fabricacion como el precio de venta del producto.

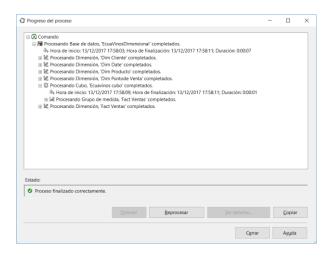


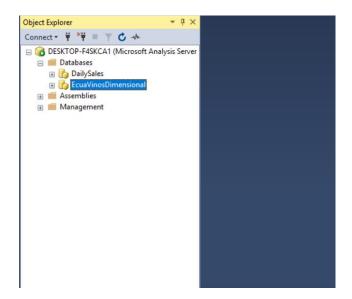
Una vez terminadas las configuraciones de las dimensiones con respecto a sus atributos y jerarquías se debe implementar la solución dando clic derecho en el nombre de la misma y seleccionando la opción "implementar".





Luego de implementar la solución se debe procesarla y se obtienen los siguientes resultados:

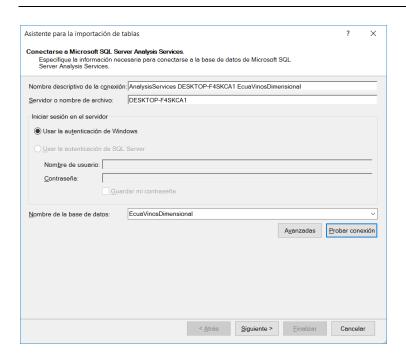




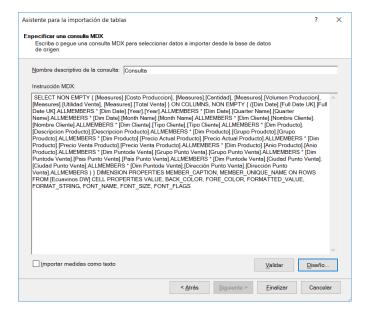
7. Implementación del Esquema Estrella utilizando la base de datos dimensional con PowerPivot.

Primero, hay que crear una conexión con la base de datos que acabamos de crear para lo cual utilizamos la opción Administrar>De Base de Datos> de Analysis Server o PowerPivot Y seleccionamos la base de datos creada (EcuaVinosDimensional).





Una vez que hayamos seleccionado la base de datos seleccionamos las diferentes métricas a evaluar según los requerimientos especificados anteriormente.







8. Resultado

8.1. ¿Quiénes son los clientes clave?

Etiquetas de fila 🔑 Suma de Meas	uresCosto Produccion	Suma de MeasuresTotal Venta	Suma de MeasuresUtilidad Venta
■ London Wines	4400	6750	2350
= 2014	4400	6750	2350
■ First	4400	6750	2350
■ February	4400	6750	2350
Reino Unido	4400	6750	2350
∃ Guayas Wines	1950	3600	1650
= 2014	1950	3600	1650
⊟ First	1950	3600	1650
■ March	1950	3600	1650
Ecuador	1950	3600	1650
∃ International Wines	3000	4500	1500
■ 2014	3000	4500	1500
⊟ First	3000	4500	1500
■ February	3000	4500	1500
Francia	3000	4500	1500
∃ Prestige Wines	3000	4000	1000
■ 2014	3000	4000	1000
⊟ First	3000	4000	1000
■ February	3000	4000	1000
Ecuador	3000	4000	1000
Total general	12350	18850	6500

8.2. ¿Qué productos son los más rentables?



Etiquetas de fila	Suma de MeasuresCantidad	Suma de Measures Volumen Produccion	Suma de MeasuresTotal Venta	Suma de MeasuresUtilidad Venta
= 2014	135	2640	18850	6500
⊟ First	135	2640	18850	6500
■ February	105	2390	15250	4850
■ Downunder Merlot	50	1100	7500	2500
Francia	30	550	4500	1500
Reino Unido	20	550	3000	1000
■ Downunder Pinot Noir	30	690	3750	1350
Reino Unido	30	690	3750	1350
■ McDonell Pinot Noir	25	600	4000	1000
Ecuador	25	600	4000	1000
■ March	30	250	3600	1650
■ Mornington Pinot Grigio	30	250	3600	1650
Ecuador	30	250	3600	1650
Total general	135	2640	18850	6500

8.3. ¿Qué mercado es el más rentable?

Etiquetas de fila	Suma de MeasuresCosto Produccion	Suma de MeasuresCantidad	Suma de MeasuresTotal Venta	Suma de MeasuresUtilidad Venta
■ Europa	7400	80	11250	
☐ Francia	3000	30	4500	1500
■ Paris	3000	30	4500	1500
2014	3000	30	4500	1500
■ First	3000	30	4500	1500
Februa	ary 3000	30	4500	1500
■ Reino Unido	4400	50	6750	2350
■ Londres	4400	50	6750	2350
□ 2014	4400	50	6750	2350
■ First	4400	50	6750	2350
Februa	ary 4400	50	6750	2350
■ Americas	4950	55	7600	2650
■ Ecuador	4950	55	7600	2650
■ Guayaquil	1950	30	3600	1650
■ 2014	1950	30	3600	1650
■ First	1950	30	3600	1650
March	1950	30	3600	1650
■ Quito	3000	25	4000	1000
= 2014	3000	25	4000	1000
■ First	3000	25	4000	1000
Februa	ary 3000	25	4000	1000
Total general	12350	135	18850	6500

9. Discusión

Según los resultados obtenidos, se puede apreciar que los clientes clave se encuentran en el exterior (Europa). Siendo London Wines, para el primer trimestre del año 2014 la empresa que ha realizado más compras a Ecuavinos generando una utilidad de venta de 2350 dólares para ese periodo de tiempo.

Con respecto a los productos que son más rentables, el vino de la marca Downunder de tipo Merlot (Tinto) es el producto mas demandado, con 50 docenas vendidas en lo que concierne al primer trimestre del año 2014, en el mes de febrero, en Reino Unido y Francia con 30 y 20 docenas respectivamente. El producto que le sigue es Downunder Pinot Noir (Igual Tinto) con un total de 30 docenas vendidas en el mismo periodo de tiempo, pero solo en Reino Unido.

Los mercados más rentables son países de Europa con respecto al exterior y provincias de nuestro país con relación a un alcance local. Para el primer trimestre del año 2014 con relación al mes de febrero, Francia y Reino Unido fueron los mercados mas rentables en donde Ecuavinos pudo obtener una utilidad neta de 3850 dólares. Siendo este mercado el que mas aporta con el crecimiento económico de la empresa.

10. Conclusiones y Trabajo Futuro

- Ecuavinos debe invertir en el exterior ya que es un mercado de alto crecimiento económico y ayudaría a que las ventas de la empresa aumenten en el futuro.
- Se debe invertir en la producción del tipo de vino Merlot ya que es el producto más rentable últimamente.
- EL aumento de precios permitirá el crecimiento económico de la empresa, pero solo a nivel de mercado internacional ya que este es el mercado fuerte.

11. Referencias

[1] M. M. Shaikh, «Create First Data WareHouse,» CodeProject, 14 Septiembre 2013. [En línea]. Available: Create First Data WareHouse. [Último acceso: 24 Noviembre 2016].

12. Anexos

12.1. Diccionario de Datos de Ecuavinos_DW

Tabla	Atributo	Data Type	Descripción
DimCliente	CodigoCliente	varchar	Primary Key
DimCliente	NombreCliente	varchar	Nombre del cliente
DimCliente	ApellidoCliente	varchar	Apellido del cliente
DimCliente	NumIDCliente	varchar	Número de Identificación del cliente
DimCliente	PaisCliente	varchar	Pais natal del cliente
DimCliente	CiudadCliente	varchar	Ciudad natal del cliente
DimCliente	DireccionCliente	varchar	Dirección de residencia del cliente
DimCliente	CodPostalCliente	varchar	Código postal de la residencia del cliente
DimCliente	TelefonoCliente	varchar	Número telefónico del cliente
DimProducto	CodigoProducto	varchar	Código referencial del producto
DimProducto	CodigoProduccionProducto	int	Primary Key
DimProducto	DescripcionProducto	varchar	Detalle general del producto
DimProducto	GrupoProudcto	varchar	Sección o Departamento del producto
DimProducto	MarcaProducto	varchar	Marca del producto
DimProducto	AnioProducto	varchar	Año de Elaboracióon del producto
DimProducto	PrecioActualProducto	money	Precio del producto para socios
DimProducto	PrecioVentaProducto	money	Precio valor público
DimProducto	VolumenProduccionProducto	int	Capacidad de producción del producto
DimPuntodeVenta	CodigoPuntoVenta	varchar	Primary Key
DimPuntodeVenta	GrupoPuntoVenta	varchar	Sección del punto de venta
DimPuntodeVenta	PaisPuntoVenta	varchar	Pais donde se ubica el punto de venta
DimPuntodeVenta	CiudadPuntoVenta	varchar	Ciudad donde se ubica el punto de venta
DimPuntodeVenta	DireccionPuntoVenta	varchar	Dirección de residencia del cliente

DimPuntodeVenta	CodigoPostalPuntoVenta	varchar	Codigo postal de la ubicación del punto de venta
DimDate	DateKey	int	Primary Key
DimDate	Date	datetime	Fecha
DimDate	FullDateUK	char	Fecha en notación británica
DimDate	FullDateUSA	char	Fecha en notación americana
DimDate	DayOfMonth	varchar	Día del mes correspondiente
DimDate	DaySuffix	varchar	Sufijo del día
DimDate	DayName	varchar	Nombre del día
DimDate	DayOfWeekUSA	char	Día de la semana en notacíon americana
DimDate	DayOfWeekUK	char	Día de la semana en notacíon británica
DimDate	DayOfWeekInMonth	varchar	Día de semana en el mes
DimDate	DayOfWeekInYear	varchar	Día del semana en el año
DimDate	DayOfQuarter	varchar	Día de Inicio del cuarto de año
DimDate	DayOfYear	varchar	Día del año
DimDate	WeekOfMonth	varchar	Semana del mes
DimDate	WeekOfQuarter	varchar	Semana del cuarto de año
DimDate	WeekOfYear	varchar	Semana del año
DimDate	Month	varchar	Mes
DimDate	MonthName	varchar	Nombre del Mes
DimDate	MonthOfQuarter	varchar	Mes del cuarto de año
DimDate	Quarter	char	Cuarto de año
DimDate	QuarterName	varchar	Nombre del cuarto de año
DimDate	Year	char	Año
DimDate	YearName	char	Nombre del año
DimDate	MonthYear	char	Mes del año
DimDate	MMYYYY	char	Mes/Año
DimDate	FirstDayOfMonth	date	Primer día del mes
DimDate	LastDayOfMonth	date	Ultimo día del mes
DimDate	FirstDayOfQuarter	date	Primer día del cuarto de año
DimDate	LastDayOfQuarter	date	Ultimo día del cuarto de año
DimDate	FirstDayOfYear	date	Primer día del año
DimDate	LastDayOfYear	date	Ultimo día del año

DimDate	IsHolidayUSA	bit	Feriado en USA
DimDate	IsWeekday	bit	Día entresemana
DimDate	HolidayUSA	varchar	Nombre de feriado de USA
DimDate	IsHolidayUK	bit	Feriado de UK
DimDate	HolidayUK	varchar	Nombre de feriado de UK
DimDate	IsHolidayEC	bit	Feriado de Ecuador
DimDate	HolidayEC	varchar	Nomre de feriado en Ecuador
DimDate	FiscalDayOfYear	varchar	Día fiscal del año
DimDate	FiscalWeekOfYear	varchar	Semana fiscal del año
DimDate	FiscalMonth	varchar	Mes fiscal
DimDate	FiscalQuarter	char	Cuarto fiscal
DimDate	FiscalQuarterName	varchar	Nombre de cuarto fiscal
DimDate	FiscalYear	char	Año fiscal
DimDate	FiscalYearName	char	Nombre del año fiscal
DimDate	FiscalMonthYear	char	Mes/Año fiscal
DimDate	FiscalMMYYYY	char	Mes/Año fiscal
DimDate	FiscalFirstDayOfMonth	date	Primer día del mes del año fiscal
DimDate	FiscalLastDayOfMonth	date	Ultimo día del mes del año fiscal
DimDate	FiscalFirstDayOfQuarter	date	Primer día del quarto del año fiscal
DimDate	FiscalLastDayOfQuarter	date	Ultimo día del quarto del año fiscal
DimDate	FiscalFirstDayOfYear	date	Primer día del año fiscal
DimDate	FiscalLastDayOfYear	date	Ultimo día del año fiscal
FactVentas	TransactionId	bigint	Primary Key
FactVentas	CodigoFactVentas	varchar	Foreign key
FactVentas	DateKey	int	Foreign key
FactVentas	CodigoPuntoVenta	varchar	Foreign key
FactVentas	CodigoCliente	varchar	Foreign key
FactVentas	CodigoProducto	varchar	Foreign key
FactVentas	Cantidad	int	Mètrica
FactVentas	CostoTotalVenta	money	Mètrica
FactVentas	CostoProduccion	float	Mètrica
FactVentas	UtilidadVentas	money	Métrica