

*Soluciones de Inteligencia de  
Negocios a su alcance:  
Fundamentos y Casos de Aplicación*

**Autor: Jonathan David Nima Ramos**

**Piura – Perú**

## *DEDICATORIA*

*A Jehová Dios, el sublime legislador y a mis padres que con esfuerzo y dedicación han logrado inculcar en mi corazón principios de moral elevados que llevo consigo para toda la vida.*

*El autor*

# INDICE

INTRODUCCION	5
PARTE I. INTRODUCCION A LAS SOLUCIONES INTELIGENTES DE NEGOCIOS	6
SISTEMAS TRANSACCIONALES OLTP	
CARACTERISTICAS DE UN SISTEMA OLTP	
SISTEMAS DE INFORMACION	
DATA WAREHOUSE AL RESCATE	
1. PROCESO DE EXTRACCIÓN DE DATOS	
1.1 PRODUCTIVIDAD	
1.2 CALIDAD DE LOS DATOS	
1.3 EXTRACCIÓN DE DATOS AL DATAWAREHOUSE DSS	
PARTE II. DATAWAREHOUSE	14
CONSIDERACIONES SOBRE UN DATAWAREHOUSE	
1. ¿PORQUÉ UTILIZARLO?	
2. VENTAJAS DEL DATAWAREHOUSE	
3. DEFINICIONES IMPORTANTES	
TECNOLOGIA HOLAP	
TECNOLOGIA ROLAP	
TECNOLOGIA MOLAP	
MULTI-DIMENSIONAL VS. BASES DE DATOS RELACIONALES	
MODELO DE DATOS MULTIDIMENSIONAL	
MODELO RELACIONAL DE BASE DE DATOS	
TECNOLOGÍA OLAP VS. TECNOLOGÍA OLTP	
PROCESAMIENTO ANALÍTICO	
PROCESAMIENTO OPERACIONAL	
ALMACEN DE DATOS OPERACIONALES (OPERATIONAL DATA STORE (ODS))	

DATA WAREHOUSES VS DATA MARTS  
CARACTERISTICAS DE UN DATA MART  
DATAWAREHOUSE EMPRESARIAL  
CARACTERISTICAS DEL DW  
PROPIEDADES DEL DATA WAREHOUSE

**PARTE III. ELABORANDO UN DATAWAREHOUSE 24**

**PASOS PARA ELABORAR UN DATAWAREHOUSE**

Identificar los Sistemas Fuentes  
Modelar los datos  
Diseñar la Base de Datos del Data Warehouse  
Mapear los datos  
Extraer los datos  
Limpiar los datos  
Transformar los datos  
Cargar el DataWarehouse  
Liberar la Metadata

**PARTE IV. Caso de Aplicación: Solución Inteligente de Negocios para Compras y Ventas en una Empresa 29**

DESCRIPCION DE LA BASE DE DATOS EN ESTUDIO:

DIAGRAMA DE LA BASE DE DATOS

OBTENCION DEL DATAWAREHOUSE Ventas

OBTENCIÓN DE DIMENSIONES:

IDENTIFICACIÓN DE LAS MEDIDAS:

OBTENCION DE LA TABLA DE HECHOS

OBTENCION DEL DATAWAREHOUSE COMPRAS

OBTENCION DE LOS CUBOS DEL DATA WAREHOUSE

**ANEXOS 50**

## INTRODUCCION

El principal problema dentro de una empresa es que existen muchos datos y muy poca información. Eso significa un alto volumen de datos y necesidades de respuesta rápida que exige un análisis complejo con consultas imposibles de prever y un ambiente de negocio en cambio constante.

El soporte a las decisiones en el modelo tradicional ha generado un impacto negativo sobre los sistemas operacionales y los analistas de negocios se vuelven "PROGRAMADORES", originando que las reglas de negocio no sean compartidas y los datos no están disponibles.

En lugar de construir sistemas para el apoyo del negocio, se dedican muchas veces sólo a resolver problemas operacionales del mismo y operar sus sistemas con más eficiencia, soportando más usuarios concurrentes, sin pensar en los sistemas y datos como importantes activos de la empresa que podrían ayudarla a ser más competitiva en el mercado mediante la explotación de los mismos. Este último punto que no se le hubiera escapado a un mercadólogo, no fue considerado por los ingenieros de sistemas que construyeron los sistemas de soporte operacional del negocio.

Por este motivo el presente libro presenta las ventajas significativas de utilizar sistemas de inteligencia de negocios teniendo como base fundamental las bodegas de datos; detalla los pasos necesarios para diseñar un datawarehouse que es la herramienta fundamental de los sistemas de procesamiento analítico y concluye con un caso de aplicación desarrollado por el autor que afianza los conceptos explicados.

Les invito a navegar por el texto a fin de enriquecer vuestros conocimientos en soluciones inteligentes de negocios, esperando sea de su completo agrado.

*El Autor*

## PARTE I. INTRODUCCION A LAS SOLUCIONES INTELIGENTES DE NEGOCIOS

### SISTEMAS TRANSACCIONALES OLTP

Los sistemas OLTP son bases de datos orientadas al procesamiento de transacciones. Una transacción genera un proceso atómico (que debe ser validado con un *commit*, o invalidado con un *rollback*), y que puede involucrar operaciones de inserción, modificación y borrado de *datos*. *El proceso transaccional es típico de las bases de datos operacionales.*

- ✓ El acceso a los datos está optimizado para tareas frecuentes de lectura y escritura. (Por ejemplo, la enorme cantidad de transacciones que tienen que soportar las BD de bancos o hipermercados diariamente).
- ✓ Los datos se estructuran según el nivel aplicación (programa de gestión a medida, ERP o CRM implantado, sistema de información departamental...).
- ✓ Los formatos de los datos no son necesariamente uniformes en los diferentes departamentos (es común la falta de compatibilidad y la existencia de islas de datos).
- ✓ El historial de datos suele limitarse a los datos actuales o recientes.

### CARACTERISTICAS DE UN SISTEMA OLTP

Se describen como características de un Sistema OLTP las que se mencionan a continuación:

- Operaciones Típicas: Update, Insert
- Un nivel bajo de requerimientos de análisis
- Screens o Pantallas No deben cambiar
- Cantidad pequeña de datos por transacción
- Nivel de los datos debe ser detallado
- Data tiene que estar actualizada
- Registrar la Orientación

*¿Porque no es adecuada esta tecnología para sistemas complejos y sistemas de soporte de decisiones?*

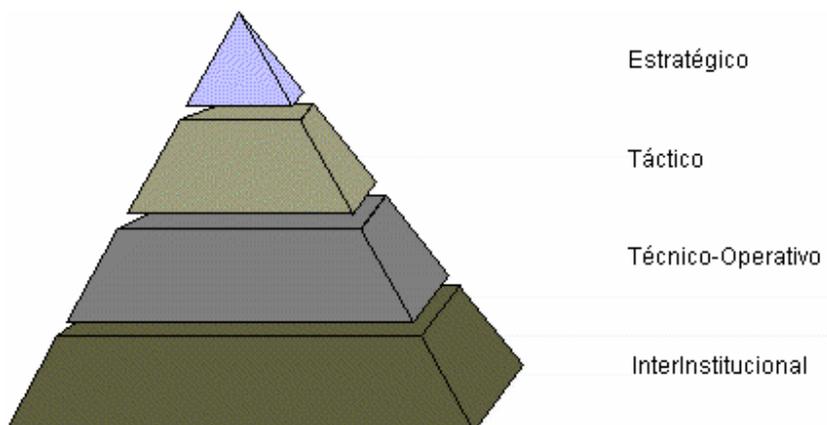
<i>OLTP</i>	<i>SISTEMA COMPLEJO</i>
<i>Información para soportar los servicios del día a día.</i>	Requiere información histórica para el análisis
<i>Data almacenada en niveles de transacción.</i>	Datos necesitan ser integrados
<i>Diseño de Base de Datos Normalizado</i>	Diseño de Base de Datos desnormalizado. Star Model

Características principales del análisis de los datos desde sistemas operacionales:

- Estructura de Datos Complejas
- La Data está dispersa
- Los sistemas OLTP no están preparadas para consultas intensivas y complejas.

## **SISTEMAS DE INFORMACION**

Los sistemas de información se han dividido de acuerdo al siguiente esquema:



**Sistemas Estratégicos**, orientados a soportar la toma de decisiones, facilitan la labor de la dirección, proporcionándole un soporte básico, en forma de mejor información, para la toma de decisiones. Se caracterizan porque son sistemas sin

carga periódica de trabajo, es decir, su utilización no es predecible, al contrario de los casos anteriores, cuya utilización es periódica.

Destacan entre estos sistemas: los Sistemas de Información Gerencial (MIS), Sistemas de Información Ejecutivos (EIS), Sistemas de Información Georeferencial (GIS), Sistemas de Simulación de Negocios (BIS y que en la práctica son sistemas expertos o de Inteligencia Artificial-AI).

**Sistemas Tácticos**, diseñados para soportar las actividades de coordinación de actividades y manejo de documentación, definidos para facilitar consultas sobre información almacenada en el sistema, proporcionar informes y, en resumen, facilitar la gestión independiente de la información por parte de los niveles intermedios de la organización.

Destacan entre ellos: los Sistemas Ofimáticos (OA), Sistemas de Transmisión de Mensajería (E-mail y Fax Server), coordinación y control de tareas (Work Flow) y tratamiento de documentos (Imagen, Trámite y Bases de Datos Documentarios).

**Sistemas Técnico-Operativos**, que cubren el núcleo de operaciones tradicionales de captura masiva de datos (Data Entry) y servicios básicos de tratamiento de datos, con tareas predefinidas (contabilidad, facturación, almacén, presupuesto, personal y otros sistemas administrativos). Estos sistemas están evolucionando con la irrupción de sensores, autómatas, sistemas multimedia, bases de datos relacionales más avanzadas y data warehousing.

**Sistemas Interinstitucionales**, este último nivel de sistemas de información recién está surgiendo, es consecuencia del desarrollo organizacional orientado a un mercado de carácter global, el cual obliga a pensar e implementar estructuras de comunicación más estrechas entre la organización y el mercado (Empresa Extendida, Organización Inteligente e Integración Organizacional), todo ésto a partir de la generalización de las redes informáticas de alcance nacional y global (INTERNET), que se convierten en vehículo de comunicación entre la organización

y el mercado, no importa dónde esté la organización (INTRANET), el mercado de la institución (EXTRANET) y el mercado (Red Global).

Sin embargo, la tecnología data warehousing basa sus conceptos y diferencias entre dos tipos fundamentales de sistemas de información en todas las organizaciones: los sistemas técnico-operacionales y los sistemas de soporte de decisiones. Este último es la base de un data warehouse.

### **1.2.1 Sistemas técnico-operacionales**

Como indica su nombre, son los sistemas que ayudan a manejar la empresa con sus operaciones cotidianas. Estos son los sistemas que operan sobre el "backbone" (columna vertebral) de cualquier empresa o institución, entre las que se tiene sistemas de ingreso de órdenes, inventario, fabricación, planilla y contabilidad, entre otros.

Debido a su volumen e importancia en la organización, los sistemas operacionales siempre han sido las primeras partes de la empresa a ser computarizados. A través de los años, estos sistemas operacionales se han extendido, revisado, mejorado y mantenido al punto que hoy, ellos son completamente integrados en la organización.

Desde luego, la mayoría de las organizaciones grandes de todo el mundo, actualmente no podrían operar sin sus sistemas operacionales y los datos que estos sistemas mantienen.

### **1.2.2 Sistemas de Soporte de Decisiones**

Por otra parte, hay otras funciones dentro de la empresa que tienen que ver con el planeamiento, previsión y administración de la organización. Estas funciones son también críticas para la supervivencia de la organización, especialmente en nuestro mundo de rápidos cambios.

Las funciones como "planificación de marketing", "planeamiento de ingeniería" y "análisis financiero", requieren, además, de sistemas de información que los soporte. Pero estas funciones son diferentes de las operacionales y los tipos de sistemas y la información requerida son también diferentes. Las funciones basadas en el conocimiento son los sistemas de soporte de decisiones.

Estos sistemas están relacionados con el análisis de los datos y la toma de decisiones, frecuentemente, decisiones importantes sobre cómo operará la empresa, ahora y en el futuro. Estos sistemas no sólo tienen un enfoque diferente al de los operacionales, sino que, por lo general, tienen un alcance diferente.

Mientras las necesidades de los datos operacionales se enfocan normalmente hacia una sola área, los datos para el soporte de decisiones, con frecuencia, toma un número de áreas diferentes y necesita cantidades grandes de datos operacionales relacionadas.

Son estos sistemas sobre los se basa la tecnología data warehousing.

## **DATA WAREHOUSE AL RESCATE**

Afortunadamente las empresas se han dado cuenta de que se deben desarrollar Sistemas que ayuden a la toma de decisiones.

Este es un esfuerzo importante debido a que el desarrollo de significados en los sistemas no está estandarizado; es decir, un programador puede almacenar un 1 en una base de datos para indicar sexo masculino y 0 para indicar sexo femenino, mientras que otro programador puede lograrlo de manera parecida guardando la letra "M" para indicar masculino y la letra "F" para indicar femenino, imposibilitando el cruce de la información de ambos sistemas porque no son compatibles entre sí. Otros problemas de los sistemas de base de datos, llamados por los técnicos de sistemas "Sistemas de Transacciones en Línea", (OLTP en el lenguaje de los técnicos) es que están orientados a la máquina, son

adecuados para la actualización rápida de la información y están diseñados para el análisis de las operaciones del día a día porque contienen mucho detalle, pero no poseen herramientas de análisis para tomar decisiones a largo plazo.

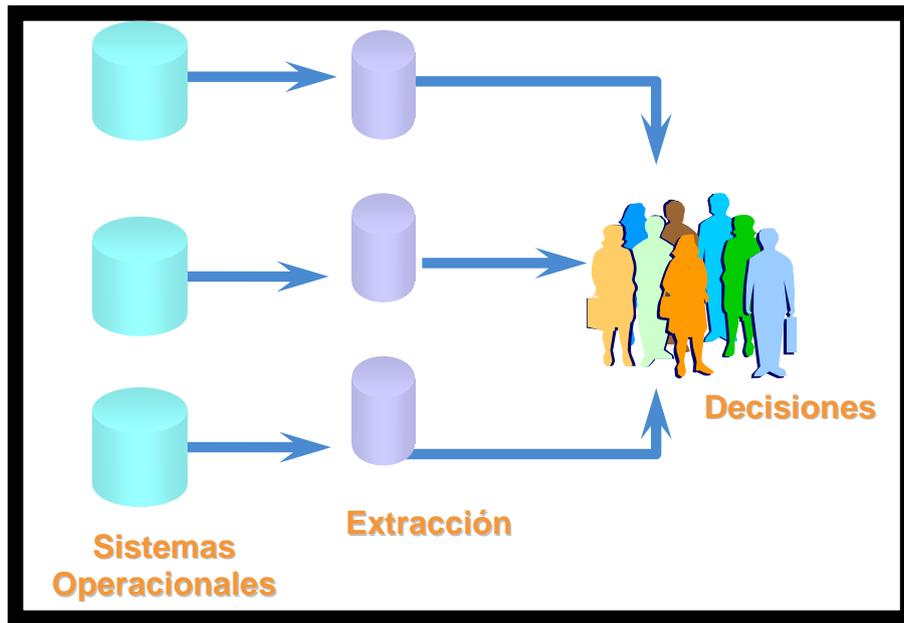
El Data warehousing es un esfuerzo para llevar esta torre de Babel de sistemas a un formato orientado al análisis inmediato de la información. Debido a que no tiene mucho nivel de detalle en los niveles elevados, es fácil para hacer análisis de negocio a nivel macro, pudiéndose aumentar el nivel de detalle mediante una técnica llamada "Drill Down" hasta llegar al nivel atómico o más.

Veamos a continuación en forma genérica una vista del proceso de Extracción de Datos de un sistema transaccional a una bodega de datos que permita una Solución Inteligente de Negocios.



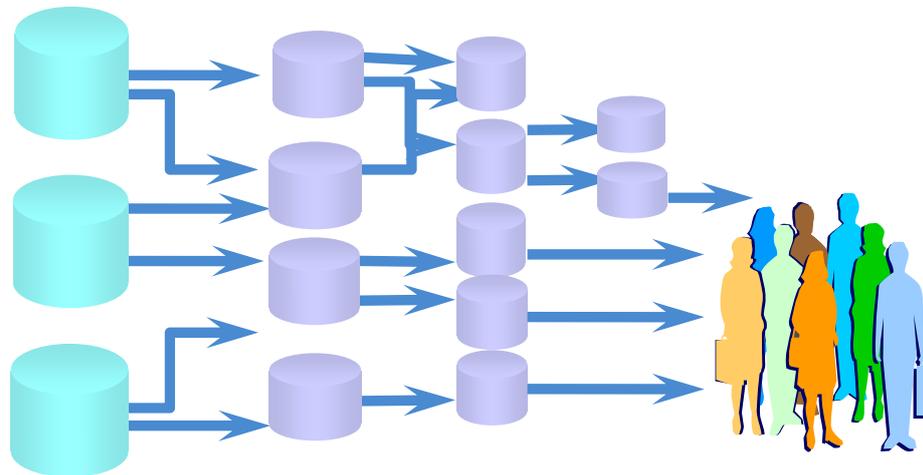
## 2. PROCESO DE EXTRACCIÓN DE DATOS

Los usuarios finales acceden a los datos fuera de línea de entornos operacionales siendo estos los propietarios de la data.



### 1.3 PRODUCTIVIDAD

- Duplicación de esfuerzos
- Utilización de múltiples tecnologías
- Reportes obsoletos
- No tienen metadata



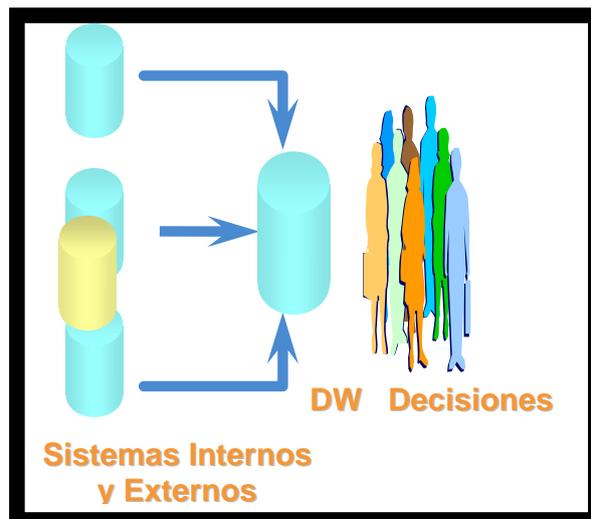
### 1.4 CALIDAD DE LOS DATOS

- Diferentes calculos de algoritmos

- Diferentes niveles de extracción
- Diferentes niveles de granularidad
- Diferentes nombres de campos de datos
- Diferentes significado de campos de datos
- Información perdida
- No posee capacidad de drill-down

#### 1.4 EXTRACCIÓN DE DATOS AL DATAWAREHOUSE DSS

- Controlada
- Confiable
- Calidad de la información
- Unica fuente de datos



## PARTE II. DATA WAREHOUSE

### CONSIDERACIONES SOBRE UN DATAWAREHOUSE

#### a. ¿PORQUÉ UTILIZARLO?

Para transformar la data en información, ante el alcance limitado de los sistemas tradicionales y a la vez brindar mayor credibilidad en la transformación de los datos operacionales en base a la unificación de criterios y así aumentar la productividad en los procesos de consulta Ante la necesidad de integrar datos dispersos en distintas BD.

#### b. VENTAJAS DEL DATAWAREHOUSE

- No duplica esfuerzos
- No necesita el soporte a muchas herramientas y tecnologías
- No hay diferencia en el significado y la representación de los datos
- No hay conflictos con los sistemas de producción
- No hay confusión de algoritmos
- No drill-down restricciones

#### c. DEFINICIONES IMPORTANTES

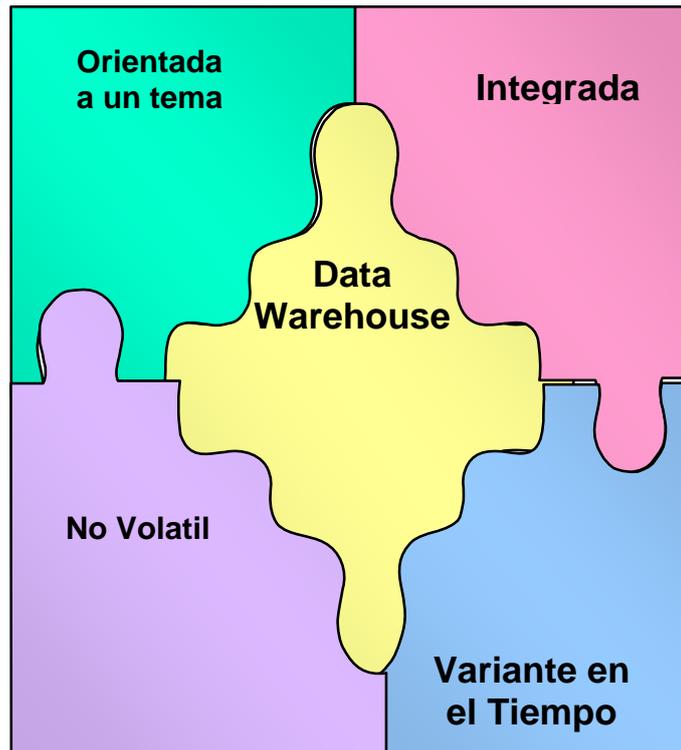
- Bill Inmon

“El DWH es una colección de datos integrada en una Base de Datos, orientada según un tema, diseñadas para soportar un Sistema de Soporte a las Decisiones (DSS), donde cada unidad de dato es relevante en algún momento del tiempo.”

- Ralph Kimball

“Un DWH es una copia de Data Transaccional, específicamente diseñada para realizar queries y análisis.”

d. PROPIEDADES DEL DATA WAREHOUSE



4.1 Orientada a Un tema

La data es categorizada y almacenada por áreas de negocio en lugar de aplicaciones

4.2 Integrada

La data es definida como única.

4.3 Variante en el Tiempo

La data es almacenada como serie de fotos asociadas al tiempo.

4.4 No-Volatil

La data en el DW típicamente No cambia.

e. **CARACTERISTICAS DEL DW**

- Las Bases de Datos son creadas específicamente para dar soporte a las decisiones.
- La información es extraída desde los sistemas originales, transformada e integrada
- La estructura del DWH es simplificada y en términos comunes del negocio, haciéndola más fácil de usar y entender
- Un DWH contiene información basada en el tiempo
- Los datos son analizados básicamente para detectar patrones y tendencias

f. **DATAWAREHOUSE EMPRESARIAL**

Permite la implementación a escala grande y alcance de todo el negocio. Se relacionan los datos desde todas las áreas mediante niveles atómicos y el desarrollo incremental.

Los usuarios de toda la organización son involucrados en este tema. Punto de distribución de los DATA MARTS dependientes

g. **CARACTERISTICAS DE UN DATA MART**

- Es un subconjunto de un DataWarehouse existente
- Optimizado para consultas específicas
- Altamente resumizado
- Específicas funciones del negocio
- Data Histórica
- Orientada a un grupo de usuarios

#### h. DATA WAREHOUSES VS DATA MARTS

<i>Propiedades</i>	<i>Data Warehouse</i>	<i>Data Mart</i>
<i>Alcance</i>	Empresarial	Departamental
<i>Áreas</i>	Múltiples	Único
<i>Fuentes de Datos</i>	Muchas	Pocas
<i>Tamaño (típico)</i>	100 GB a > 1 TB	< 100 GB
<i>Implementación</i>	Meses a años	Meses

#### ALMACEN DE DATOS OPERACIONALES (OPERATIONAL DATA STORE (ODS))

Almacena data táctica desde los sistemas de producción, orientada a un tema e integrada con fines operacionales.

Características:

- Mixtura de un DATAWAREHOUSE y un sistema operacional
- Atiende necesidades operacionales
- Estructura desnormalizada
- Levemente sumariada, datos de detalle
- Contiene algunos datos operacionales
- Data actualizada o casi actualizada

#### PROCESAMIENTO OPERACIONAL

Es el conjunto de sistemas transaccionales *on-line* que ejecutan las operaciones diarias del negocio.

**PROCESAMIENTO ANALÍTICO**

Es el conjunto de sistemas que ofrecen información usada para el análisis de un problema o situación.

**TECNOLOGÍA OLAP VS. TECNOLOGÍA OLTP**

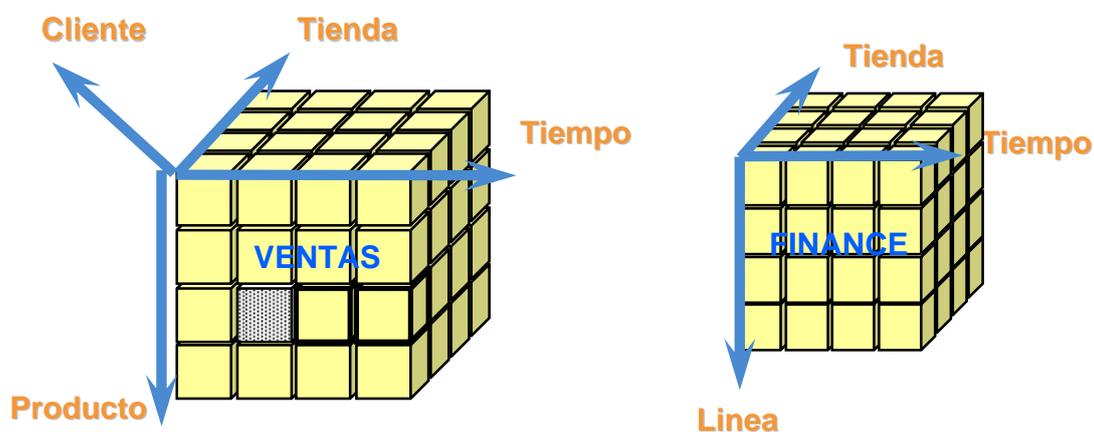
	<i>Tecnología OLTP</i>	<i>Tecnología OLAP</i>
<i>Objetivo</i>	Control de los Proc. Oper.	Toma de Decisiones
<i>Cliente</i>	Personal Operacional	Gestores del Negocio
<i>Datos</i>	Atómicos, Actualizados Consolidados y Dinámicos	Históricos y Estables
<i>Estructura</i>	Normalizada	Dimensional
<i>Tiempo resp.</i>	Segundos	De Seg. A Minutos
<i>Orientación</i>	Orientado a Aplicaciones	Orient. A Información
<i>Acceso</i>	Alto	Moderado a bajo
<i>Actualización</i>	Continuamente	Periódicamente
<i>Aplicación</i>	Estructurada y Procesos No Estructurada y Repetitivos	Procesos Analíticos

**MODELO RELACIONAL DE BASE DE DATOS**

	<i>Atributo1</i>	<i>Atributo2</i>	<i>Atributo3</i>	<i>Atributo4</i>
	<i>Nombre</i>	<i>Edad</i>	<i>Genero</i>	<i>EMP. No</i>
<i>Row1</i>	Anderson	31	F	1001
<i>Row 2</i>	Green	42	M	1007
<i>Row 3</i>	Lee	22	M	1010

Esta tabla ilustra la relación de empleados

### MODELO DE DATOS MULTIDIMENSIONAL

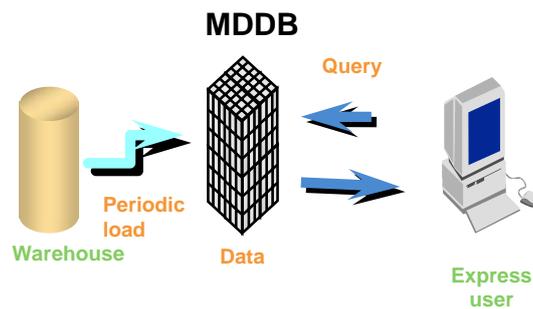


*La data se encuentra en la intersección de las dimensiones.*

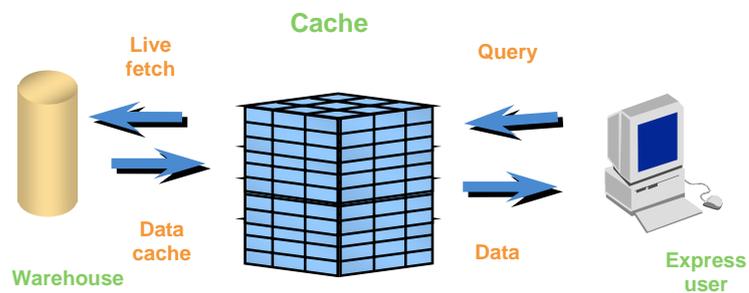
### MULTI-DIMENSIONAL VS. BASES DE DATOS RELACIONALES

Multi-Dimensional	Relacional
<ul style="list-style-type: none"> <li>- Acceso mas rápido</li> <li>- Multiplicidad de vistas de los datos</li> <li>- Slice and dice</li> <li>- Problemas de performance con base de datos grandes</li> <li>- Usada especialmente para Data Marts.</li> </ul>	<ul style="list-style-type: none"> <li>- Es conocida y entendida</li> <li>- Fuerte procesos de back-up y de restore</li> <li>- Mejor performance en base de datos grandes</li> </ul>

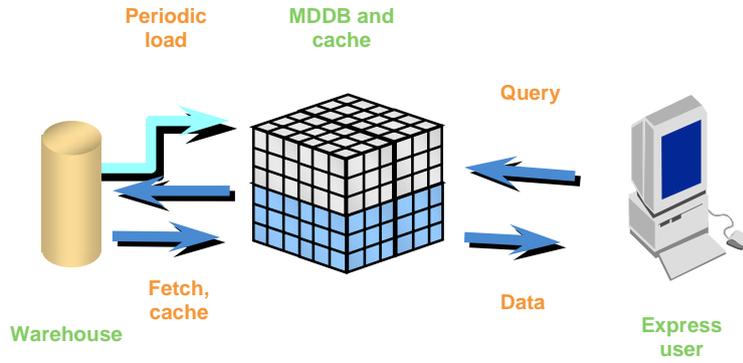
### TECNOLOGIA MOLAP



### TECNOLOGIA ROLAP



## TECNOLOGIA HOLAP



i.



## PORTE III. ELABORANDO UN DATAWAREHOUSE

### PASOS PARA ELABORAR UN DATAWAREHOUSE

#### 3.1 Identificar los Sistemas Fuentes

- Estudiar y entender la arquitectura IT
- Inventario de los sistemas transaccionales existentes
- Inventario de los sistemas de análisis existentes
- Investigar fuentes potenciales del Data Warehouse
- Explorar e investigar fuentes externas a la empresa
- Explorar los temas de calidad de datos
- Entender la administración de cambios de los sistemas fuentes



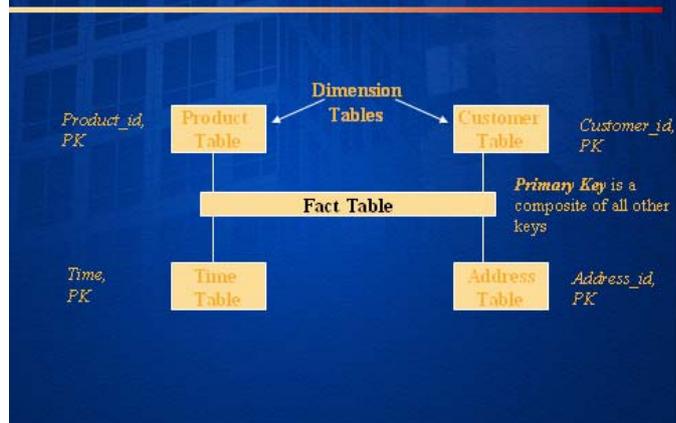
#### 3.2 Modelar los datos

- Modelar la data
  - ✓ Escoger la herramienta de modelado de datos
  - ✓ Establecer estándares de nombres
  - ✓ Determinar la granularidad
  - ✓ Escoger un esquema
  - ✓ Desarrollar los modelos de datos de alto nivel para las áreas de negocio a ser consideradas en este Data Warehouse
- Validar el modelo de datos

## Modelo Lógico



## El Star Schema



## El Snowflake Schema

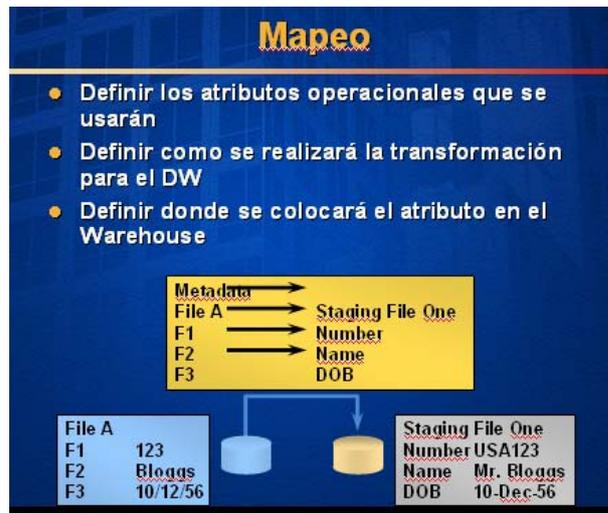


### 3.3 Diseñar la Base de Datos del data Warehouse

- Planear un nivel de staging
- Estimar volúmenes
- Escoger un DBMS
- Identificar las necesidades de los datos derivados
- Generar scripts

### 3.4 Mapear los datos

- Establecer mapeo de los requerimientos del negocio
- Mapear requerimientos a las necesidades de datos
- Crear el mapeo destino
- Mapear los datos



### 3.5 Extraer los datos

- Alinear los procesos de extracción al mapeo de datos
- Determinar el rol del staging area
- Considerar actividades de transformación y limpieza
- Escoger la data a extraer y el software de transformación
- Extraer los datos requeridos y colocarlos en el staging area (o direccionarlos en el data warehouse destino)
- Validar y probar los procesos de extracción de datos



### 3.6 Limpiar los datos

- Considerar modificaciones en las reglas de la data operacional.
- Documentar las fuentes.
- Crear un programa de calidad de datos.
- Diseñar claramente los procesos de limpieza.
- Los procesos de limpieza iniciales y los de los incrementos pueden ser diferentes
- Definir Estándares de Calidad



### 3.7 Transformar los datos

- Revisar la visión de los procesos de transformación de datos
- Detallar y describir las derivaciones necesarias, sumalizaciones y/o otras operaciones
- Escoger el software de transformación de Datos
- Transformar la Data
- Validar y probar los procesos de transformación y los datos

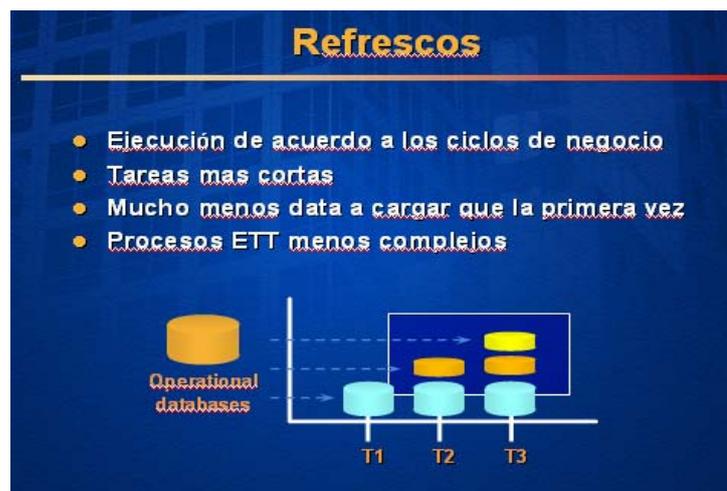
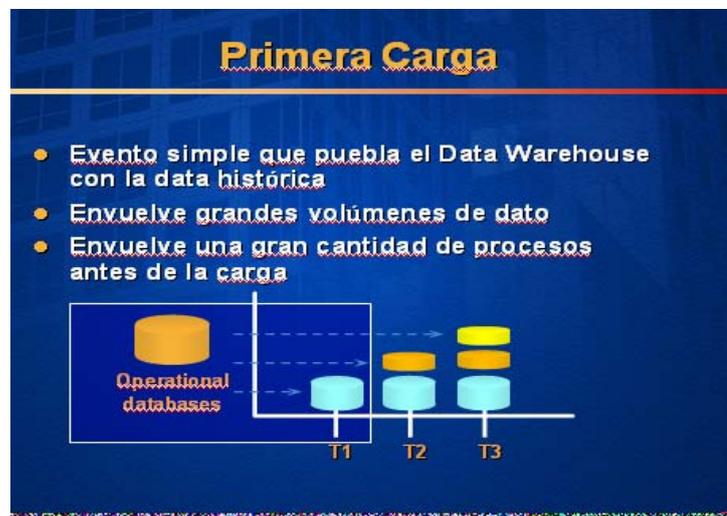


Rutinas de transformación:

- Limpieza de datos
- Eliminar inconsistencias
- Agregar elementos
- Juntando datos
- Integrando datos
- Transformando datos antes de la carga

### 3.8 Cargar el DataWarehouse

- Desarrollar el plan de carga
  - Calcular el tiempo
  - Preparar la infraestructura técnica
  - Preparar el software y los datos
  - Desarrollar el plan de contingencia
- Considerar el rol del staging area
- Cargar los datos
  - Desarrollar y probar la carga inicial
- Validar la data cargada



### 3.9 Liberar la Metadata

- La metadata son los datos acerca de los datos
- Tan importante como el propio dato del warehouse
- Es generado en todos los pasos del proceso

*En resumen, ¿ Que contiene un repositorio de metadata?*

- Nombres de campos y definiciones
- Mapeo de los datos
- Tablas
- Indices
- Cronogramas de extracción, carga, etc.
- Criterios de selección
- Calculos de los datos derivados
- Transformación de los datos

## PARTE IV. Caso de Aplicación: Solución Inteligente de Negocios para Compras y Ventas en una Empresa

### DESCRIPCION DE LA BASE DE DATOS EN ESTUDIO:

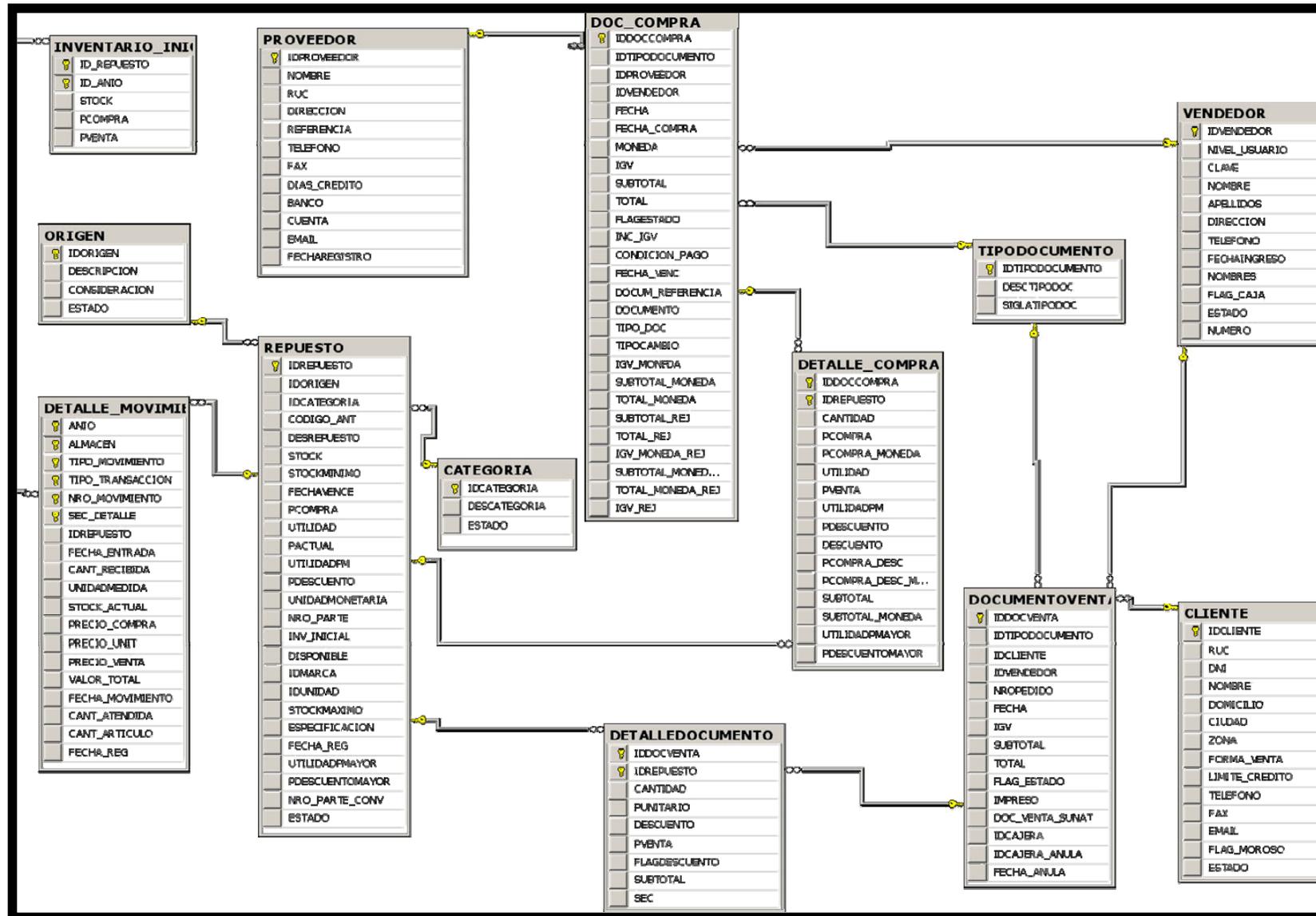
La base de datos que se ha tomado en cuenta para la aplicación de las técnicas de DATA WAREHOUSE. Esta base corresponde a un Sistema de Compras y Ventas de Repuestos en una empresa de motos.

Esta dividida en tres partes importantes:

- 1.- Manejo de las Compras: se trabaja con tablas de PROVEEDORES, VENDEDOR (Comprador de la empresa), CABECERA DE COMPRAS, DETALLE DE COMPRAS, REPUESTOS y/o otros.
- 2.- Manejo de VENTAS: se trabaja con tablas de CLIENTES, VENDEDOR, CABECERA DE VENTA, DETALLE DE VENTA, REPUESTOS y/o otros.
- 3.- Manejo de tablas MAESTRAS (MOVIMIENTO y DETALLE DE MOVIMIENTO), además de una tabla llamada TIPO DE MOVIMIENTO que pueden ser Compras, Ventas, Devoluciones, Regalos, Donaciones, etc.

Sacando esas deducciones podemos decir que tenemos tres diferentes tipos de DATA WAREHOUSE: DWCompras, DWVentas, DWMaestras.

## DIAGRAMA DE LA BASE DE DATOS



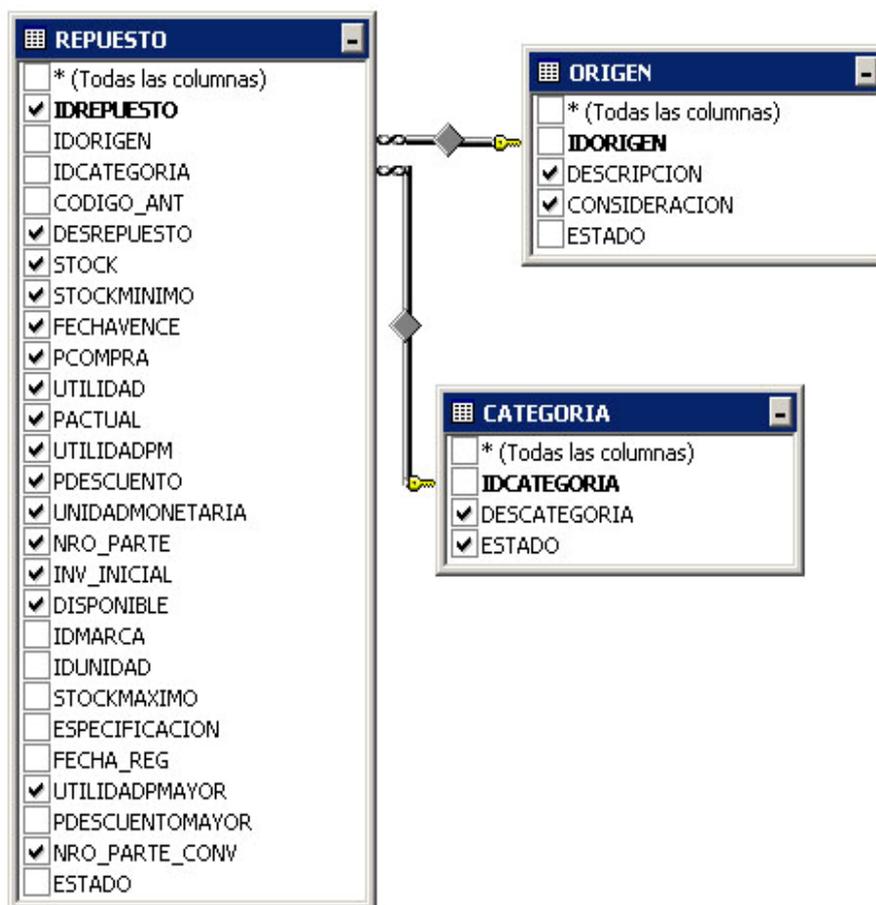
## ELABORACION DEL DATAWAREHOUSE Ventas

### OBTENCIÓN DE DIMENSIONES:

Para esta DATA DATAWAREHOUSE se trabajara con las dimensiones: Repuesto\_dim, Cliente\_dim, Time\_venta\_dim, Vendedor\_dim.

Aquí se muestra la obtención de la tabla de dimensión Repuesto\_dim, la cual se obtuvo de la creación de una vista, en la Base de Datos Original.

### Repuesto\_dim



De igual manera, con la ayuda de la Herramienta: **Proyecto de Integration Services** del SQL 2005 se logro crear la tabla de dimensiones en el DATA DATAWAREHOUSEmencionado, con la siguiente consulta:

```
CREATE TABLE [Repuesto_dim] (  
[Repuesto_dim_key] numeric(18,0) identity(1,1) primary key,  
    [IDREPUESTO] NUMERIC(18),  
    [DESREPUESTO] VARCHAR(120),  
    [STOCK] INTEGER,  
    [STOCKMINIMO] INTEGER,  
    [FECHAVENCE] DATETIME,  
    [PCOMPRA] NUMERIC (18,4),  
    [PACTUAL] NUMERIC (18,3),  
    [UTILIDAD] NUMERIC (18,4),  
    [UTILIDADPM] NUMERIC (18,4),  
    [PDESCUENTO] NUMERIC (18,3),  
    [UNIDADMONETARIA] VARCHAR(10),  
    [INV_INICIAL] INTEGER,  
    [NRO_PARTE] VARCHAR(50),  
    [DISPONIBLE] VARCHAR(1),  
    [DESCATEGORIA] VARCHAR(100),  
    [ESTADO] VARCHAR(1),  
    [DESCRIPCION] VARCHAR(50),  
    [CONSIDERACION] VARCHAR(100),  
    [UTILIDADPMAYOR] NUMERIC (18,4),  
    [NRO_PARTE_CONV] VARCHAR(50)  
)
```

Aquí se muestran las otras consultas para la creación de las otras tablas de dimensión, empleando las Herramienta del SQL SERVER 2,005 antes mencionada.

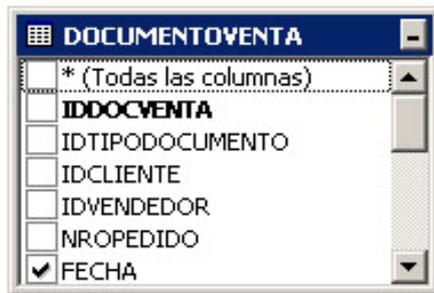
### Cliente\_dim



```
CREATE TABLE [Cliente_dim] (  
[Cliente_dim_key] numeric(18,0) identity(1,1) primary key,  
[IDCLIENTE] VARCHAR(8),  
[RUC] VARCHAR(11),  
[DNI] VARCHAR(8),  
[NOMBRE] VARCHAR(140),  
[DOMICILIO] VARCHAR(200),  
[CIUDAD] VARCHAR(12),  
[ZONA] VARCHAR(12),  
[FORMA_VENTA] VARCHAR (12),  
[LIMITE_CREDITO] NUMERIC (18,2),  
[TELEFONO] VARCHAR(12),
```

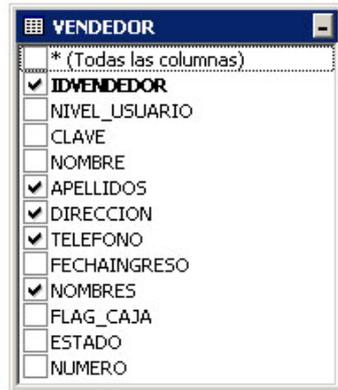
```
[FAX] VARCHAR(12),  
[EMAIL] VARCHAR(12),  
[FLAG_MOROSO] VARCHAR(2)  
)
```

Time\_venta\_dim



```
CREATE TABLE [Time_venta_dim] (  
[Time_Venta_dim_key] numeric(18,0) identity(1,1) primary key,  
[full_date_venta] DATETIME,  
[daymonth] INTEGER,  
[dayyear] INTEGER,  
[dayname] NVARCHAR(30),  
[Weeknumber] INTEGER,  
[monthname] NVARCHAR(30),  
[monthnumber] INTEGER,  
[quarternumber] INTEGER,  
[calendaryear] VARCHAR(4)  
)
```

## Vendedor\_dim



```
CREATE TABLE [Vendedor_dim] (  
  [Vendedor_dim_key] numeric(18,0) identity(1,1) primary key,  
  [IDVENDEDOR] VARCHAR(5),  
  [APELLIDOS] VARCHAR(50),  
  [NOMBRES] VARCHAR(50),  
  [DIRECCION] VARCHAR(100),  
  [TELEFONO] VARCHAR(50)  
)
```

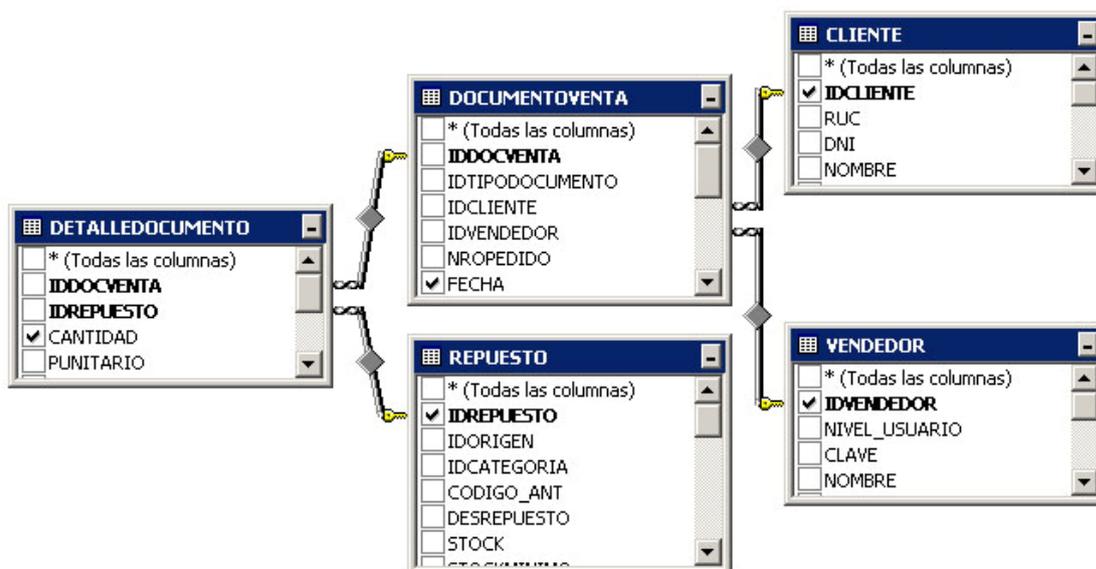
## IDENTIFICACIÓN DE LAS MEDIDAS:

Las MEDIDAS obtenidas en esta DATAWAREHOUSE Compras serán:

- Cantidad
- SubTotal

## VENTAS\_AUXILAR

Para poder conservar la integridad de los datos del DATAWAREHOUSE y evitar que estos se dupliquen o se muestren datos erróneos o incorrectos, se utilizará una tabla, la cual se obtendrá de la creación de una vista en nuestra base de datos, la que estará formada por las claves originales de las tablas que actúan en esta vista y que son campos de las tablas de dimensiones, mas no llaves principales de estas. Además, de tener las medidas del DATA WAREHOUSE.



Luego de esto con la ayuda de la herramienta: **Proyecto de Integration Services** se creara la tabla auxiliar: **VENTAS\_AUXILIAR**, la cual como se dijo anteriormente nos sirve de ayuda mas no forma parte del DATA DATAWAREHOUSEVentas.

```
CREATE TABLE [VENTAS_AUXILAR] (  
    [IDCLIENTE] VARCHAR(8),  
    [IDVENDEDOR] VARCHAR(5),  
    [IDREPUESTO] NUMERIC(18),  
    [FECHA] DATETIME,  
    [CANTIDAD] NUMERIC(5),  
    [SUBTOTAL] NUMERIC (18,2)  
)
```

#### OBTENCION DE LA TABLA DE HECHOS

En la tabla de hechos, se han identificado sus claves foráneas, las cuales son claves principales de las claves de dimensiones. Además se le agrego los campos de las medidas (todo esto se ha logrado gracias a la creación de una vista donde se mezclan todas las tablas de dimensiones con la Tabla Auxiliar).

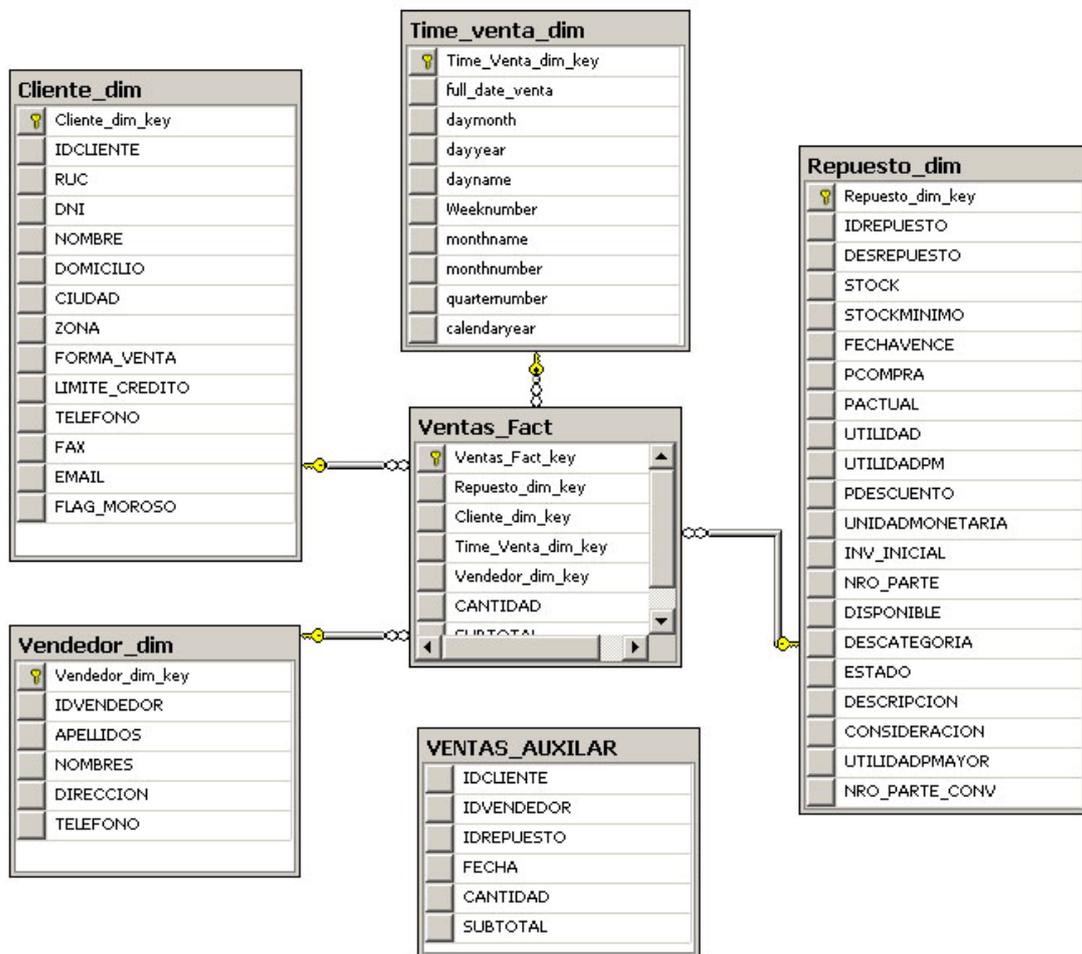
Aquí se muestra la sentencia SQL SERVER 2 005 que se utilizo en la herramienta **Proyecto de Integration Services** para la creación de la tabla de hechos obteniéndose al final el diagrama del DATA WARE HOUSE Ventas de nuestra base de datos en estudio.

#### Ventas\_Fact

```
CREATE TABLE [Ventas_Fact] (  
    [Ventas_Fact_key] numeric(18,0) identity(1,1) primary key,  
    [Repuesto_dim_key] NUMERIC(18),
```

[Cliente\_dim\_key] NUMERIC(18),  
 [Time\_Venta\_dim\_key] NUMERIC(18),  
 [Vendedor\_dim\_key] NUMERIC(18),  
 [CANTIDAD] NUMERIC(5),  
 [SUBTOTAL] NUMERIC (18,2)

)

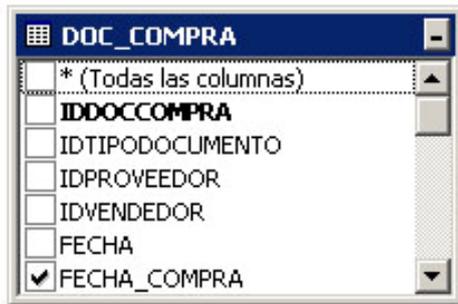


Proveedor\_dim



```
CREATE TABLE [Proveedor_dim] (  
[Proveedor_dim_key] numeric(18,0) identity(1,1) primary key,  
    [IDPROVEEDOR] VARCHAR(8),  
    [RUC] VARCHAR(12),  
    [NOMBRE] VARCHAR(120),  
    [DIRECCION] VARCHAR(100),  
    [REFERENCIA] VARCHAR(80),  
    [TELEFONO] VARCHAR(12),  
    [FAX] VARCHAR(12),  
    [DIAS_CREDITO] INTEGER,  
    [BANCO] VARCHAR(12),  
    [CUENTA] VARCHAR(12),  
    [EMAIL] VARCHAR(20)  
)
```

time\_compra\_dim



```
CREATE TABLE [time_compra_dim] (
[Time_Compra_dim_key] numeric(18,0) identity(1,1) primary key,
[fecha_full_compra] DATETIME,
[daymonth] INTEGER,
[dayyear] INTEGER,
[dayname] NVARCHAR(30),
[Weeknumber] INTEGER,
[monthname] NVARCHAR(30),
[monthnumber] INTEGER,
[quarternumber] INTEGER,
[calendaryear] VARCHAR(4)
)
```

#### 4. ELABORACION DEL DATAWAREHOUSE Compras

##### Obtención tabla de Dimensiones

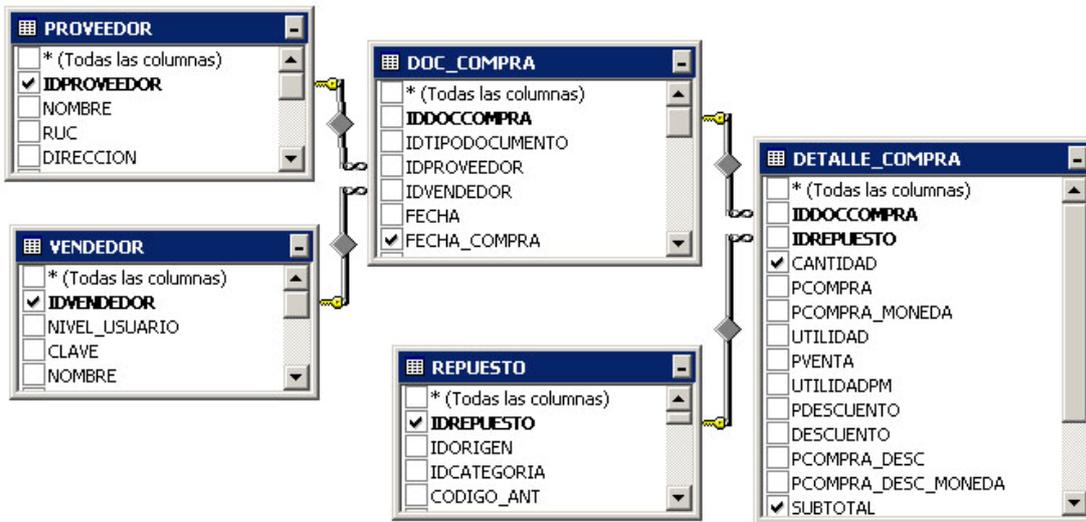
Proveedor, Vendedor (Comprador de la empresa), Repuesto,  
Time\_compra\_dim (Dimensión de Tiempo).

##### Obtención medidas.

- Cantidad.
- Subtotal.

## Compras\_Auxiliar

```
CREATE TABLE [Compras_Auxiliar] (  
    [IDPROVEEDOR] VARCHAR(8),  
    [IDVENDEDOR] VARCHAR(5),  
    [IDREPUESTO] NUMERIC(18),  
    [FECHA_COMPRA] DATETIME,  
    [CANTIDAD] NUMERIC(18),  
    [SUBTOTAL] NUMERIC (18,2)
```



Obtención tabla de hechos :Compras\_Fact

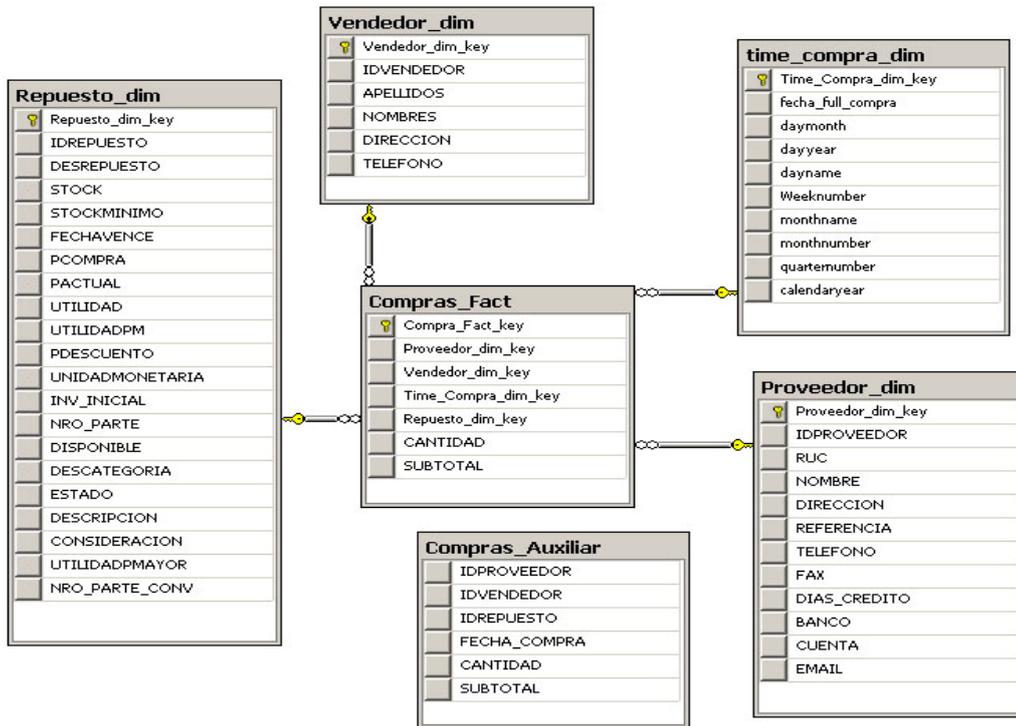


```

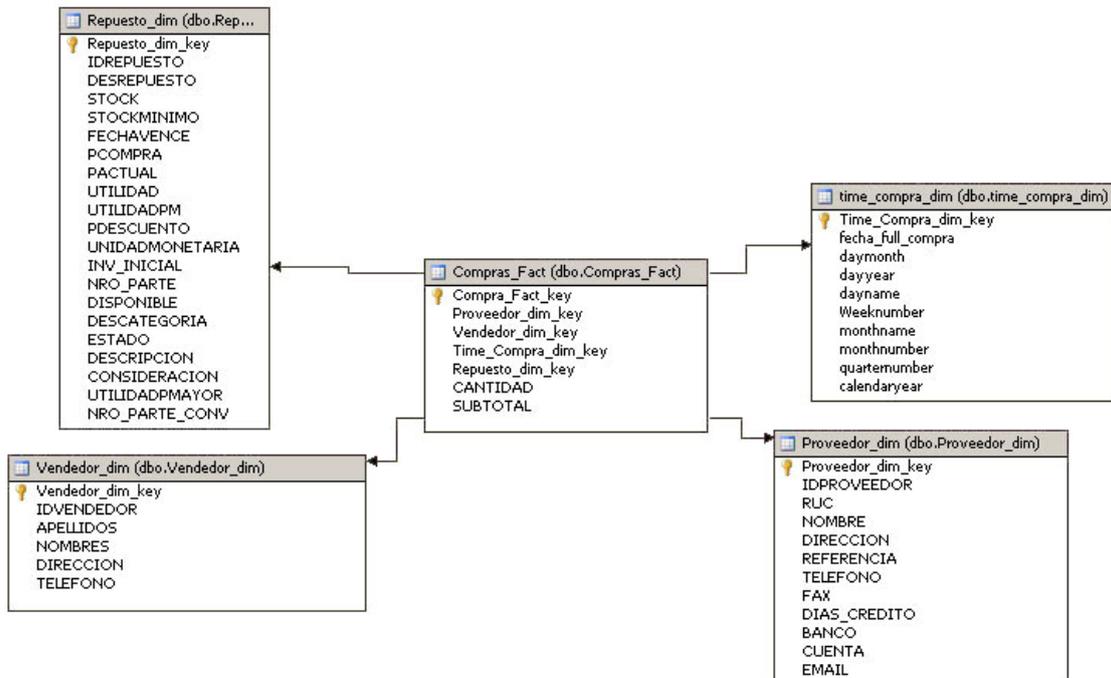
CREATE TABLE [Compras_Fact] (
[Compra_Fact_key] numeric(18,0) identity(1,1) primary key,
  [Proveedor_dim_key] NUMERIC(18),
  [Vendedor_dim_key] NUMERIC(18),
  [Time_Compra_dim_key] NUMERIC(18),
  [Repuesto_dim_key] NUMERIC(18),
  [CANTIDAD] NUMERIC(18),
  [SUBTOTAL] NUMERIC (18,2)
)

```

## OBTENCION DEL DATAWAREHOUSE COMPRAS



## OBTENCION DE LOS CUBOS DEL DATA WAREHOUSE



En el recuadro que presento en la página siguiente se observa la Solución Inteligente implementada mediante CUBOS OLAP.

Se ha utilizado Visual Studio 2005 Business Intelligence para implementar la solución y procesar los datos en modo de almacenamiento HOLAP utilizando Cubos.

**DW Racing Ventas.cube [Diseño]**

Estructura d... | Uso de dimen... | Cálculos | KPIs | Acciones | Particiones | Perspectivas | Traducciones | Examinador

Perspectiva: DW Racing Ventas | Idioma: Predeterminado

Measures

- Ventas Fact
  - CANTIDAD
  - Recuento Ventas
  - SUBTOTAL
- Cliente Dim
- Repuesto Dim
  - CONSIDERACION
  - DESCATEGORIA
  - DESCRIPCION
  - DESREPUESTO
  - DISPONIBLE
  - ESTADO
  - FECHAVENCE
  - IDREPUESTO
  - INV INICIAL
  - NRO PARTE
  - NRO PARTE CONV
  - PACTUAL
  - PCOMPRA
  - PDESCUENTO
  - Repuesto Dim
  - STOCK
  - STOCKMINIMO
  - UNIDADMONETARIA
  - UTILIDAD
  - UTILIDADPM
  - UTILIDADPMAYOR
- Time Venta Dim

Dimensión	Jerarquía	Operador	Expresión de filtro
<Seleccionar dimensión>			

Coloque campos de filtro aquí

		DESCRIPCION							Total general
	RUC	AM	BR	CH	JP	JPGN	NA	CANTIDAD	
Monthname		CANTIDAD	CANTIDAD	CANTIDAD	CANTIDAD	CANTIDAD	CANTIDAD	CANTIDAD	
Enero	10005174126			1			4	5	
	10008406788				2			2	
Total				1	2		4	7	
Febrero	10002019022			5			3	8	
	10005174126	1		1	1		2	5	
Total		1		6	1		5	13	
Marzo	10005174126		1	5	2		15	23	
	10008406788					1		1	
Total			1	5	2	1	15	24	
Abril			3	1			16	20	
Mayo	10005174126		1	4	3	1		9	
Total			1	4	3	1		9	
Junio	10005174126						1	1	
	10008406788		1					1	
Total			1				1	2	
Julio	10005174126			19			58	77	
Total				19			58	77	
Agosto	10005174126			1	2		25	28	
Total				1	2		25	28	
Septiembre	10005174126			2			24	26	
Total				2			24	26	
Total general		1	6	39	10	2	148	206	

Lista de errores

Explorador de solución... | Proyecto de Anal... | DW Racing Ventas Cube

Propiedades

**DW Racing Ventas Cube**

- ErrorConfigur (predetermina)
- EstimatedRow 0
- Language
- ProcessingMo Regular
- ProcessingPric 0
- Source DW Racing\_ve
- Visible True

**Básica**

- Description
- ID DW Racing Ver
- Name **DW Racing V**

Se observa la granularidad que nos permite un Reporte OLAP y que es de mucha ayuda al momento de analizar la información por parte de los Gerentes de una organización pues ellos pueden simplemente con arrastrar una columna o fila ver un resultado que llamamos información y que tardaría talvez mucho tiempo obtener si lo desarrolláramos con consultas simples.

Es importante acotar la importancia de la forma en que se visualización los datos pues es una gran ayuda para las personas que toman las decisiones en la Organización tener reportes que engloben información relevante, consistente y fácil de entender.

**DW Racing.cube [Diseño]**

Estructura d... | Uso de dimen... | Cálculos | KPIs | Acciones | Particiones | Perspectivas | Traducciones | Examinador

Perspectiva: DW Racing | Idioma: Predeterminado

Dimensiones: <Seleccionar dimensión>

Coloque campos de filtro aquí

		DESCRIPCION							Total ger
		IP	JPGN	NA	TH	TL	TW	SUBTOTAL	
<b>NOMBRE</b>	<b>Monthname</b>	SUBTOTAL	SUBTOTAL	SUBTOTAL	SUBTOTAL	SUBTOTAL	SUBTOTAL	SUBTOTAL	
+	ACCESORIOS IRVIN			4503,5				4503,5	
+	AUTOPARTES S.A.		55,14					105,39	
+	BARBACCI MOTORS S.A.	24400,23	65,56	86,28	3006,88	2782,44	25929,47	66437,1	
+	BELLCORP REPRESENTACIONES S.A.C.			683,42			322	4324,62	
	Total general	24400,23	120,7	5273,2	3006,88	2782,44	26251,47	75370,6	

Lista de errores

**Explorador de solución...**

- Solución 'ProyectedAnalysis5
- ProyectedAnalysis5
  - Orígenes de datos
  - Vistas de origen de da
  - Cubos
    - DW Racing.cube
  - Dimensiones
  - Estructuras de minería
  - Funciones
  - Ensamblados
  - Varios

**Propiedades**

DW Racing Cube

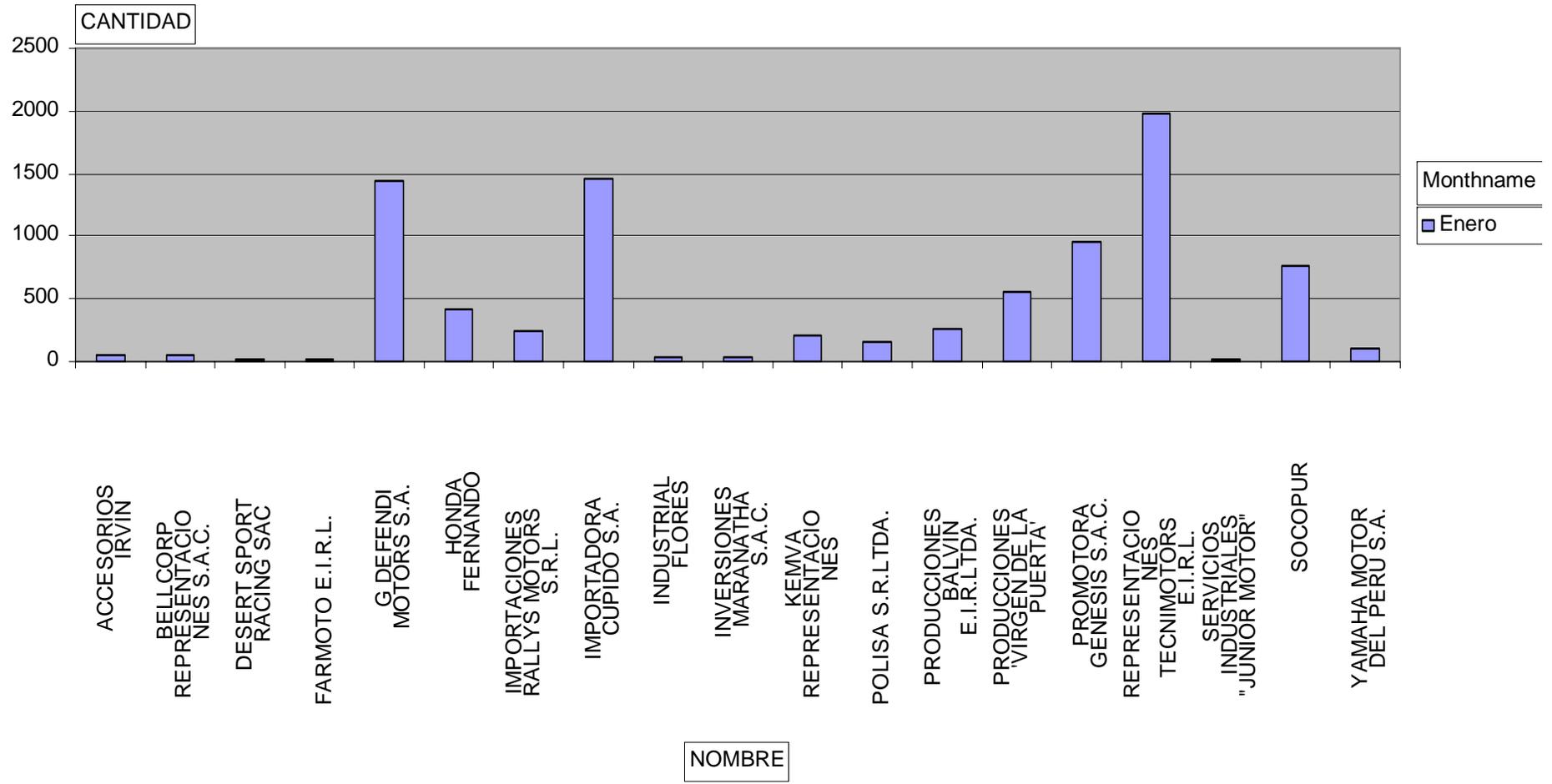
**Almacenamiento**

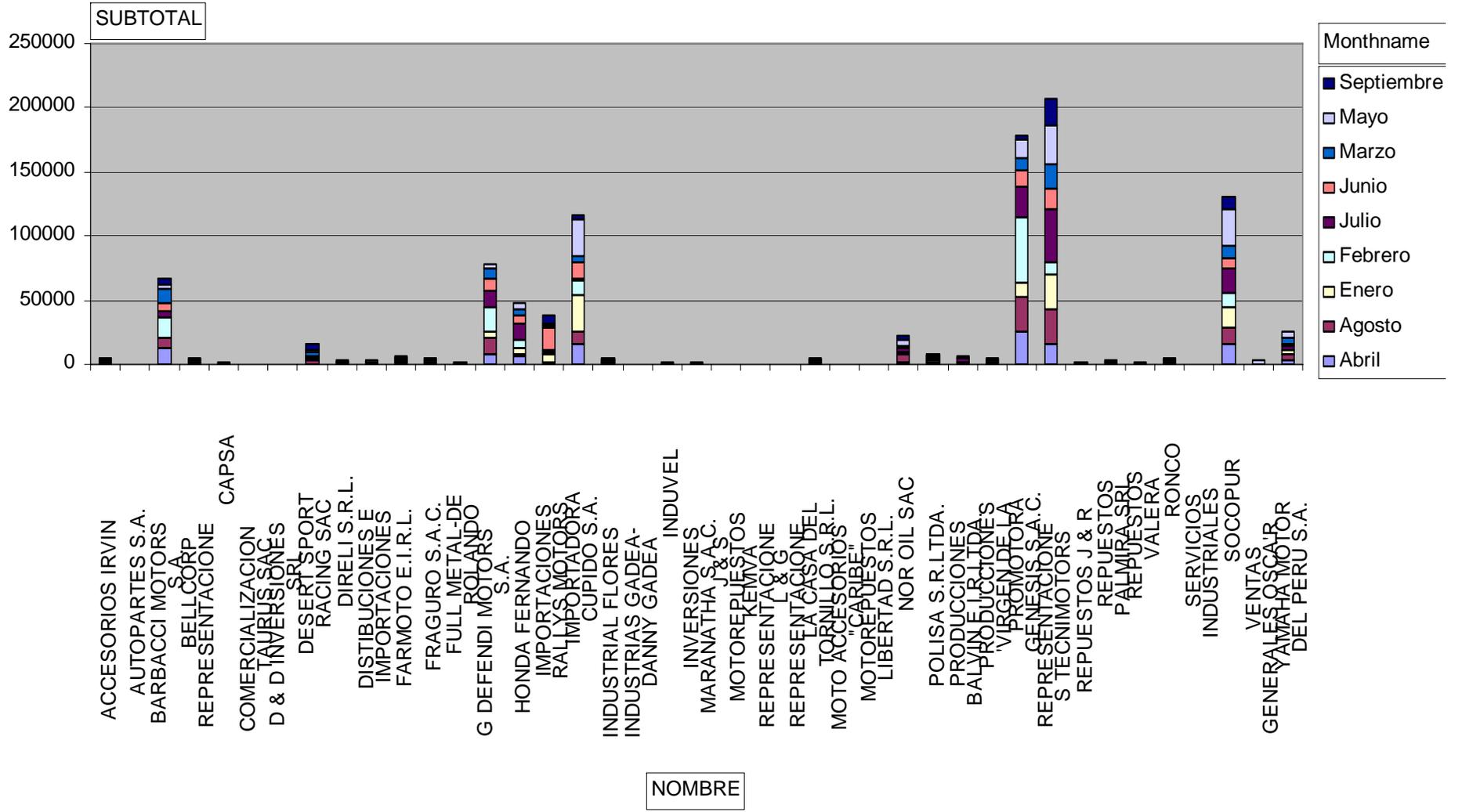
- ProactiveCach (ninguno)
- StorageMode Molap

**Avanzada**

- AggregationP
- DefaultMeasL
- ErrorConfigur (predeterminac
- EstimatedRow 0
- Language
- ProcessingMo Regular
- ProcessingPri 0

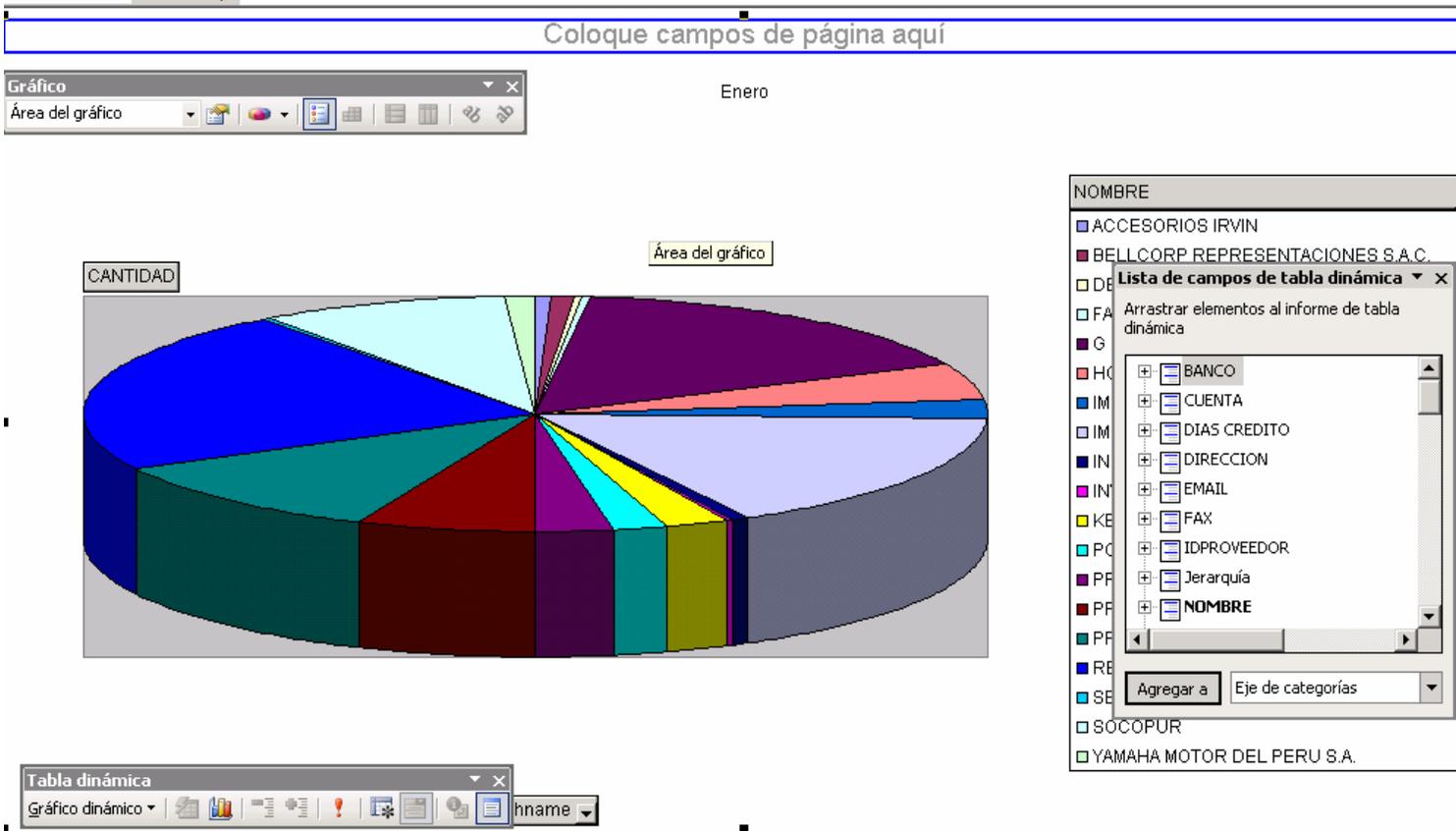
Enero



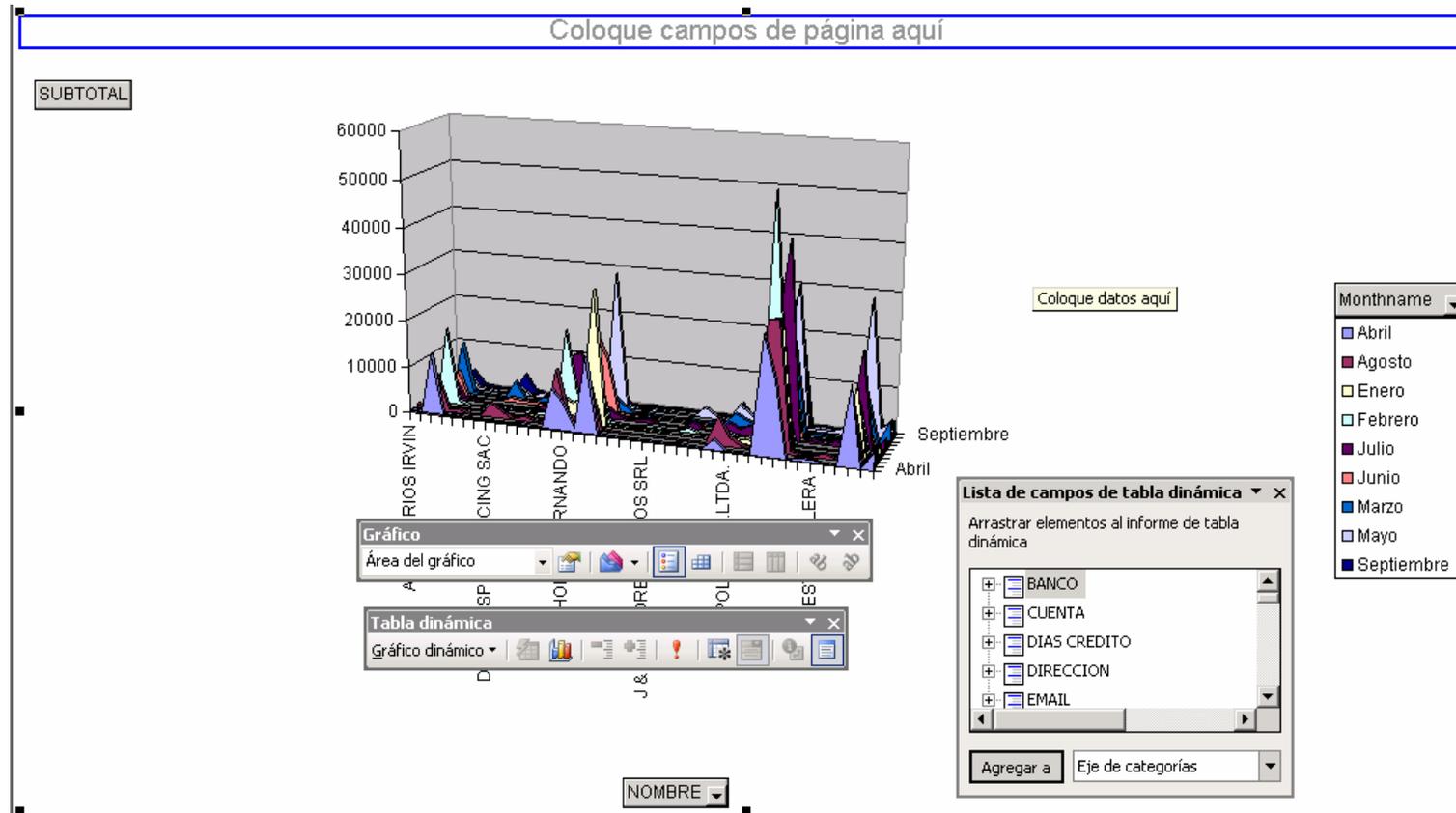


# ANEXOS

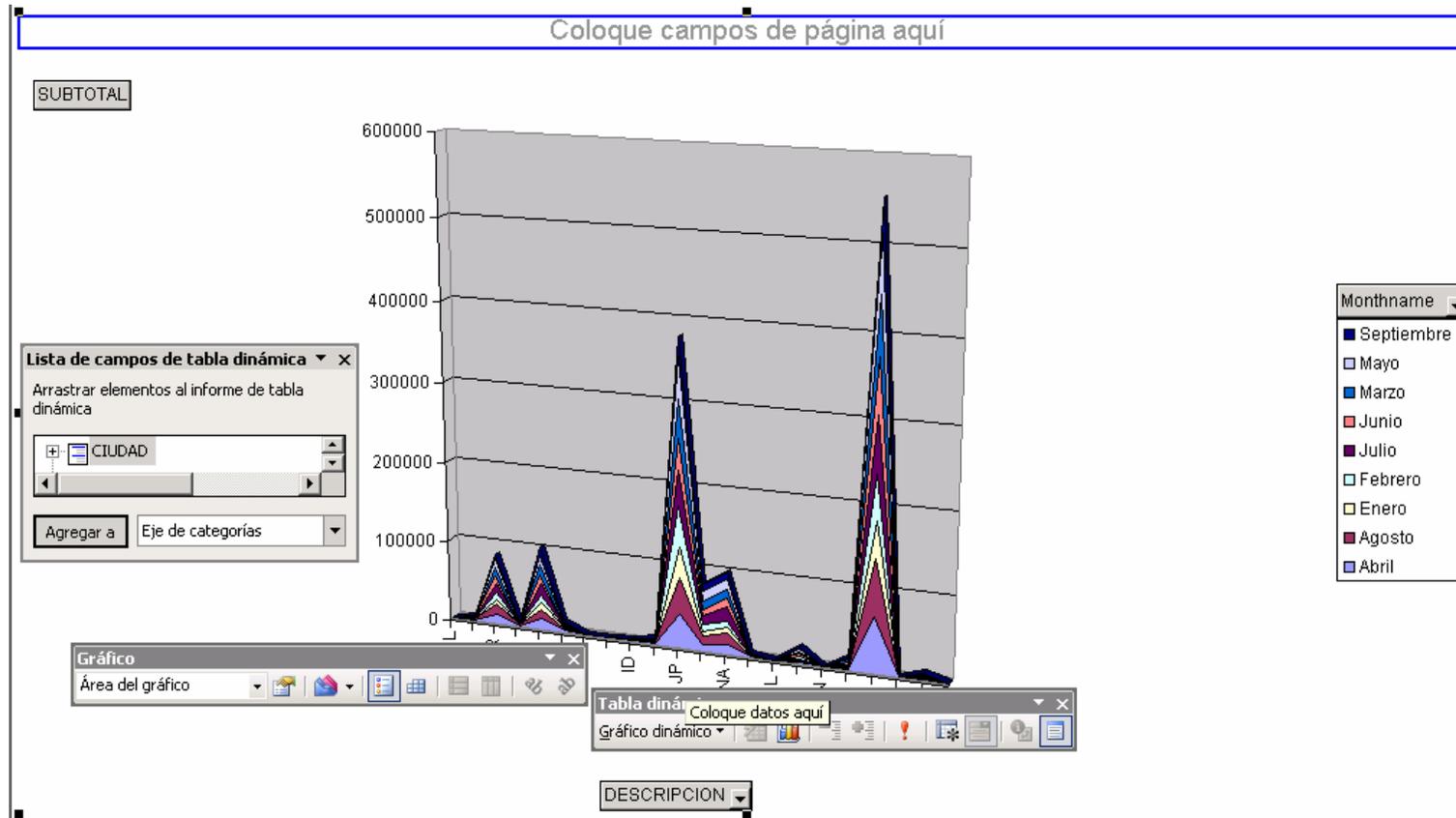
Grafico: Cantidad Comprada por Proveedor el mes Enero



## Grafico: Cantidad Comprada por Proveedor por Mes



## Grafico: Monto Comprado por Repuesto por Mes



## VISTA DE LA SOLUCION UTILIZANDO MICROSOFT EXCEL PARA INTERACTUAR CON EL USUARIO

Excel interface showing a PivotTable and a PivotTable Field List.

**Tabla dinámica**

Tabla dinámica

Coloque campos de página aquí

SUBTOTAL	Monthname						
DESCRIPCION	Abril	Agosto	Enero	Febrero	Julio	Junio	Marzo
AL	2					1	3
AM	562.6	588	365.75	652.16	795	644	431.3
BR	11274.31	12284.33	7132.95	8698.26	10751	10042.81	10617.18
BRGN	620.6	672.44	265.15	276.5	248.5	529.35	728.1
CH	13078.61	11673.88	8812.8	10487.66	13283.71	12038.32	12089.12
CHGN	1773.85	2220.14	1092.17	1400.44	1668.69	1646.45	999.02
CL	5.5	13.1			6.6		
DR	30			10		10	19
ID	311.87	525.06	275.02	122.36	428	190.78	239.88
IDGN	717.93	1779.99	529.26	728.09	585.2	737.5	1076.25
JP	41310.26	46903.21	37235.34	50178.75	43288.48	39449.52	47791.49
JPGN	8162.67	10673.47	5477.63	9123.5	9947.09	8114.35	11037.66
NA	11906.41	13807.87	7515.95	8388.22	17292.53	11399.08	10918.16
NAGN	379	1504.8	128.02	157	384.5	469.3	372.5
SL							16
TH	1647.7	2185.58	2303.39	2301.44	1492.22	2591.76	2464.34
THGN		228			5	1	
TL	1560.94	1044.51	1119.24	1276.84	2533.56	2779.19	2685.28
TW	69487.48	69947.11	45360.36	55009.15	68875.71	66161.46	70357.55
TWGN			47.22	45			
USA	922.6	1122.48	517.6	1067.4	713.8	1016.6	1398
USGN			61				
Total general	163754.33	177173.97	118238.85	149922.77	172299.59	157822.47	173243.83

**Lista de campos de tabla dinámica**

Arrastrar elementos al informe de tabla dinámica

CIUDAD

Agregar a Área de filas

