

PROYECTO FINAL DE INGENIERÍA

**BIGP-Buenas prácticas asociadas a BI trabajando con modelos
asociativos**

De Merduch, Guillermo Jorge – LU118776

Ingeniería Informática

De Perez de Antueno, Gustavo – LU116266

Ingeniería Informática

Tutor/es:

Ferre, Ricardo Luis, UADE

Febrero 11, 2014



UNIVERSIDAD ARGENTINA DE LA EMPRESA

FACULTAD DE INGENIERÍA Y CIENCIAS EXACTAS

Resumen

El presente trabajo se enfoca en dos temas a tratar.

En primer lugar se centrará en la definición de un estándar con pautas de diseño a considerar en el desarrollo de modelos asociativos de forma de evitar los inconvenientes que surgen al implementar los mismos con una herramienta de Business Intelligence.

En segundo lugar será un estándar a tener en cuenta para definir y calcular distintos tipos de medidas de negocio. Se analizarán los tipos de medidas (totalmente aditivas, semi-aditivas ó no aditivas, así mismo los conceptos aplicados a la conversión en el tiempo de las mismas).

En ambos casos se darán ejemplos reales que son producto de la experiencia profesional en el área y casos probados en aplicaciones comerciales.

Por un lado las medidas de negocio son la base para tomar decisiones en una compañía y nos permiten visualizar indicadores del negocio, por ejemplo si la estrategia de ventas está funcionando correctamente ó si las ventas están cayendo respecto a la competencia. Todo esto planteado anteriormente es de vital importancia en una empresa a la hora de tomar decisiones estratégicas, ya que partirán de tener medidas de negocio que nos den información extra a la que conocemos de una compañía y de forma correcta, para luego tomar decisiones claves sobre qué hacer en cada caso. Los errores en el cálculo y exposición de las medidas de negocio pueden llevar a tomar de forma imprecisa información representada como válida, y así por ejemplo cambiar una estrategia de negocio ó incluso despedir empleados en una compañía.

Por otra parte se debe escoger un correcto periodo temporal para trabajar las medidas, algo sobre lo que no se suele discutir mucho y que va unido a nuestro modelo de negocio. A veces analizar indicadores mensuales, no nos dan una aproximación a la realidad, dado a que una empresa puede tener días en que los incrementa como otros en los que disminuye y necesitamos verlos con mas detalle (horas, días, semanas, etc.).

Abstract

This work focuses on two issues to deal with.

First we focus on the definition of a standard design guidelines to consider in the development of associative models in order to avoid the problems that arise when implementing them with a business intelligence tool.

Secondly will be considered a standard to define and compute different types of business measures. We will analyze the types of measures (full additive, semi - additive or not additive, also the concepts applied to the conversion in the same time).

In both cases we will give real examples product of professional experience in the field and tested cases in commercial applies.

On one hand the measures are the bases for company decisions and allow us to visualize business indicators, for example if the sales strategy is working correctly or if sales are falling comparing to the competition. All this mentioned before is vital in a company when it's taking strategic decisions, because it will allow us to have business measures that give us extra information to what we know about the company correctly, for then take key decisions on what to do in each case. The calculation errors and the measures presentation can carry us to take invalid information as valid and in this way, for example change business strategy or even fire company`s employees.

Moreover we must choose a correct time period to work the business measures, something that is not usually discuss much and which is attached to our business model. Sometimes analyze monthly indicators do not give us an approximation to reality, considering that a company can have days when increases as others that diminishes and we need to see them deeply (hours, days, weeks, etc.).

Agradecimientos

A los profesionales de Business Intelligence de Telecom Argentina por brindarnos información sobre sus experiencias en trabajar en el área de Business Intelligence usando modelos asociativos.

Al profesor Ricardo Luis Ferré por ser nuestro tutor, guiarnos en el camino y respondernos todas las consultas que tuvimos en el trayecto de la tesis.

A todos nuestros familiares, en especial a la abuela Lorenza Inés Puglia de Ursic por ser el sostén de los estudios de Guillermo Merduch en sistemas y a la mamá Maria Inés Ursic por ayudarnos en la corrección de la tesis.

Contenido

1. Introducción	6
2. Conceptos a saber	6
2.1. ¿Qué es Business Intelligence?	6
2.2. ¿Por qué Business Intelligence?	7
2.3. Herramientas de BI	8
2.4. ¿Qué es QlikView? ¿Como funciona?	13
2.5. Modelo Asociativo	17
2.6. ¿Qué es una Medida?	20
2.6.1. Tipos de medidas	21
3. Desarrollo	27
3.1. Pautas de diseño al trabajar con modelos asociativos	27
3.1.1. Claves sintéticas	28
3.1.2. Claves circulares	33
3.1.3. Uso de variables	34
3.1.4. Claves optimizadas	37
3.1.5. Renombrar campos	37
3.2. Medidas	37
3.2.1. Estándar metodológico para definir medidas de negocio	38
3.2.2. Aditivas	44
3.2.3. Semi-Aditivas	58
3.2.4. No Aditivas	69
4. Conclusión	79
5. Bibliografía	80

1. Introducción

Nuestro objetivo consiste en desarrollar un estándar metodológico para definir medidas de negocio y brindar soluciones de diseño a inconvenientes más comunes que surgen al implementar modelos de datos asociativos en herramientas de Business Intelligence.

Este trabajo va a contar con una sección en donde se explicarán los conceptos necesarios a tener en cuenta para el entendimiento de la tesis, otra sección en donde se van a dar pautas de diseño para trabajar con modelos asociativos y finalmente definiciones de medidas de negocio junto a sus buenas prácticas. Ambas cosas serán abordadas mediante casos reales de negocio producto de nuestra experiencia de haber trabajado en el área de Inteligencia de Negocios de diferentes compañías.

2. Conceptos a saber

2.1. ¿Qué es Business Intelligence?

En primer lugar partiremos de la definición de Business Intelligence (BI) del glosario de términos de Gartner¹:

“BI es un proceso interactivo para explorar y analizar información estructurada sobre un área (normalmente almacenada en un datawarehouse), para descubrir tendencias o patrones, a partir de los cuales derivar ideas y extraer conclusiones. El proceso de Business Intelligence incluye la comunicación de los descubrimientos y efectuar los cambios. Las áreas incluyen clientes, proveedores, productos, servicios y competidores.” (Fuente: Gartner)

Permite entender, rastrear y administrar sus negocios para incrementar al máximo el desempeño de la compañía. A partir de esto, podrán mejorar la eficiencia en las operaciones, hacer relaciones mucho más rentables con sus clientes y desarrollar nuevos productos que se diferencian del resto.

En segundo lugar en lo que hace a la estructura empresarial, proporcionará a los empleados información indispensable y necesaria para optar por las mejores decisiones. Puede utilizarse en varios ambientes que van desde pequeños grupos de trabajo hasta implementaciones de empresas de varios empleados. Si hablamos de la extranet, la forma de implementar BI se realiza mediante aplicaciones que permiten a nuestras empresas brindar

¹ Gartner es una consultora internacional especializada en Tecnologías de Información y Comunicación.

nuevos servicios y también para construir relaciones más sólidas con nuestros clientes, proveedores y socios mediante Internet.

En tercer lugar en lo que hace a la administración del desempeño empresarial, las organizaciones deben entender y tener una visibilidad constante de sus indicadores clave de desempeño en toda la empresa, con esto estaremos asegurando que nuestra estrategia siempre permanezca alineada en todo momento con la empresa, desde la mercadotecnia hasta las ventas como también desde la manufactura hasta los recursos humanos.

Finalmente nos debe permitir una visión global de nuestra empresa, es la función clave de Business Intelligence.

2.2. ¿Por qué Business Intelligence?

Hoy en día y como viene posicionándose el BI en el mercado, las empresas necesitan tomar decisiones mucho más rápidas para poder lograr ser aun más competitivas y eficientes. Este limitante de tiempo afecta a las personas en su capacidad de tomar buenas decisiones.

Existen herramientas como EUROWIN DECISIÓN EMPRESARIAL que es un software que ayuda a tomar decisiones, y también herramientas (CRM² 4, ERP³ 3 y SCM⁴ 5) que son utilizadas en las compañías y tienen una estructura muy inflexible para dicho fin, se adaptan a permitir el manejo de los datos, pero no permiten obtener información de ellos. Lo que vemos entonces es que los sistemas de información que poseen las empresas, no las están ayudando a tomar decisiones rápidas y eficaces.

Para dar soporte a las necesidades analíticas, el BI se apoya en herramientas que nos permitirán la extracción, depuración y almacenamiento de toda la información de la empresa, proporcionando luego una alta velocidad de recuperación para generar el conocimiento y ayudarnos a tomar decisiones rápidamente, objetivo principal de Business Intelligence.

En conclusión la solución BI nos permite mirar qué está pasando, comprender por qué ocurre, extraer el conocimiento de estas informaciones para poder llegar a predecir qué pasará y con esto seleccionar qué camino debemos seguir.

² CRM (Software para la administración de la relación con los clientes). Son sistemas informáticos de apoyo a la gestión de las relaciones con los clientes, a la venta y al marketing. Con este significado CRM se refiere al sistema que administra un almacén de datos con la información de la gestión de ventas y de los clientes de la empresa.

³ ERP (Sistemas de planificación de recursos empresariales) son sistemas de gestión de información que automatizan muchas de las prácticas de negocio asociadas con los aspectos operativos o productivos de una empresa.

⁴ SCM (Gestión de configuración de software) es una especialización de la gestión de configuración a todas las actividades en el sector del desarrollo de software.

2.3. Herramientas de BI

Se pretende dar una breve introducción sobre algunas de las diferentes alternativas que se utilizan en el mercado respecto a soluciones de Business Intelligence, tanto licenciadas como Open Source, según el cuadrante mágico de Gartner (Ver Figura 1). Con el análisis de la importante consultora a nivel mundial en donde creó una representación gráfica de la situación en el mercado de un producto tecnológico en un determinado momento.

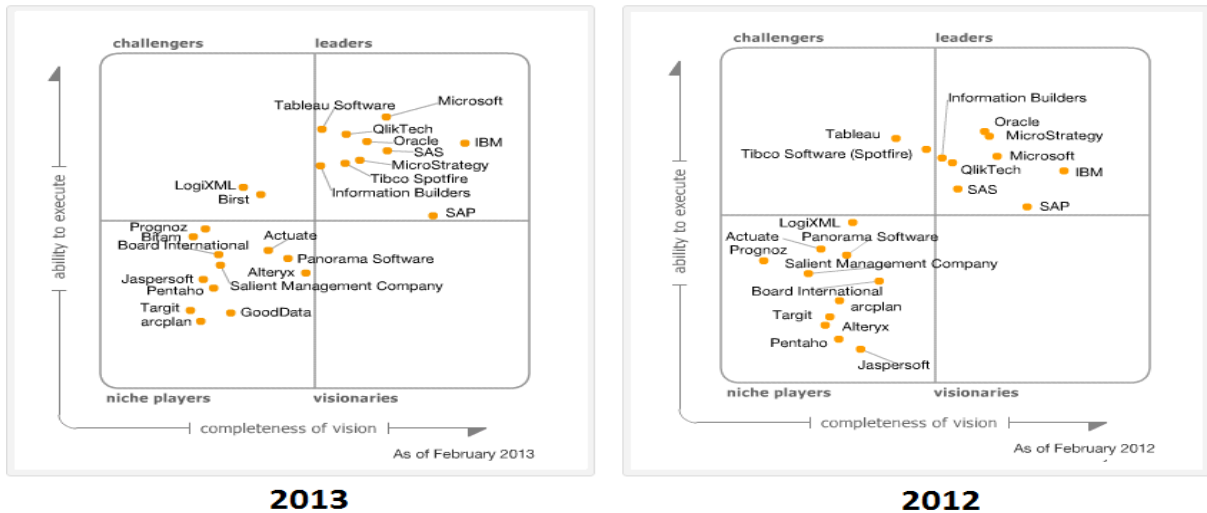


Figura 1: Arquitectura Comparativa de los años 2013, 2012 (Fuente: Gartner, Febrero 2013).

En la explicación de las herramientas, producto de la información proporcionada por sus sitios web oficiales, se mencionan muchos términos técnicos, los cuales son explicados brevemente a continuación:

- ETL: Hace referencia a todo el proceso que abarca la extracción, transformación y carga de datos.
- OLAP-Procesamiento analítico en línea: presenta una solución que se utiliza en un campo llamado Inteligencia empresarial (Business Intelligence), su objetivo principal es agilizar consultas de gran cantidad de datos. Para lograr esto se utilizan estructuras multidimensionales (llamados también Cubos OLAP), estos contienen los datos resumidos de Bases de datos grandes o de Sistemas Transaccionales (OLTP). Es utilizado para informes de marketing, de negocios de ventas, como también para informes de dirección, de minería de datos entre otras áreas.
- ROLAP- Relational OLAP: Se define como una implementación OLAP en la cual se almacenan datos en un motor relacional. En primer lugar, los datos son detallados, lo

cual evita las agregaciones y las tablas están desnormalizadas. Se usan esquemas, los más comunes son estrella y copo de nieve, pero también se puede trabajar sobre otras bases de datos relacionales. Usa una arquitectura que se compone de un servidor como el banco de datos relacional y un motor denominado OLAP que se aloja en un servidor dedicado. Una de las principales ventajas que ofrece esta arquitectura es permitirnos el análisis de enormes cantidades de datos.

- MOLAP-Multidimensional OLAP: en primer lugar podemos decir que dicha implementación OLAP en este caso almacena datos en una base de datos multidimensional. Con esto lo que se logra es la optimización de los tiempos de respuesta, ya que el resumen de la información se calcula antes. Los valores precalculados o las agregaciones son la base de las ganancias en el desempeño de este tipo de sistema. Es común ver que algunos de estos sistemas usan técnicas para compresión de datos con lo cual logran disminuir el espacio de almacenamiento en disco debido a los valores que son precalculados.

Pentaho

Este software nos ofrece, en primer lugar, todo el espectro de recursos para poder desarrollar, explotar y mantener nuestro proyecto de BI. Empezando por el ETL con el producto Data Integration hasta terminar con los cuadros de mando con el producto Dashboard Designer ó el Comunity Dashboard Framework.

La forma en que Pentaho construyó una solución de BI fue integrando los diferentes proyectos que ya existían con solvencia reconocida, como por ejemplo Data Integration (permite el ETL de datos) y Mondrian (es un motor híbrido OLAP que combina la flexibilidad del los motores ROLAP con una caché que le proporciona velocidad).

En consecuencia podemos decir que la plataforma Open Source Pentaho Business Intelligence abarca amplias necesidades de Análisis de Datos y de Informes empresariales. Estas soluciones de Pentaho se desarrollaron e implementaron en Java. Por este motivo podemos afirmar que es una de las soluciones más flexibles para poder cubrir una gran cantidad de necesidades empresariales de negocio, sumado a la gran cantidad de clientes que tiene en Argentina.

Tableau

Tableau es un software que contribuye a que las personas puedan ver y comprender sus datos. Por un lado permite visualizar, analizar, y compartir información rápidamente. Por otra parte muchas organizaciones obtienen resultados rápidos con Tableau ya sea en la oficina o con dispositivos móviles, ya que la herramienta esta disponible para smartphones, tablets, y computadoras de escritorio. Además, los usuarios de la herramienta también pueden usar Tableau Public para compartir sus datos en sitios web. Teniendo un repositorio común de ideas y sugerencias.

A nivel producto se está afianzando en el mercado Argentino, intentando competir con el producto QlikView.

Cognos

El software Cognos es utilizado en la actualidad para ayudar a las empresas a obtener una mejor comprensión de sus debilidades y fortalezas en su desempeño. Su principal objetivo es ayudar a establecer mejores objetivos y asignar de forma eficaz los recursos necesarios para alcanzarlos.

Está diseñado para ayudar a todos en la organización a tomar las decisiones que permitan alcanzar mejores resultados de negocio. Es una herramienta de IBM, por más que actualmente no tenga tanto peso en el mercado Argentino, tiene el sello de una empresa grande y reconocida mundialmente.

Oracle Hyperion

Hyperion Financial Management es utilizado para simplificar la consolidación de datos y la generación de informes de un estricto cumplimiento normativo. Se utiliza para reducir el esfuerzo necesario, y aumentar la fiabilidad de los datos ofrecidos. Es una herramienta de uso muy sencillo que permite la consolidación de datos y generación de informes sin que se implique al personal de TI.

Está diseñado de una manera específica para las tareas de presupuestación, consolidación, cierre y gestión financiera. Esta aplicación es una base sólida para asentar una cultura que se basa en la transparencia. Hyperion Financial Management nos permite eliminar ciertas transacciones de empresas, realizar cálculos de los intereses minoritarios y conversiones multdivisa entre sus diversas funcionalidades. Además este software incorpora las normas internacionales para la realización de informes financieros (IFRS) y otros

requisitos legales para que dichos informes sean precisos y cumplan con todas las reglamentaciones de hoy en día.

Además se podrá hacer rápidamente los ciclos de consolidación y generación de informes, así mejorará la comprensión de la estructura financiera del negocio. También se reducirá la necesidad de introducir y verificar datos por parte de su equipo, reemplazando los procedimientos de control de forma manual por controles automatizados, los cuales ofrecen prevención y reducen los costos de auditoría.

Hyperion es utilizado para obtener una visión coherente de la realidad de nuestra empresa, a partir de la integración de distintos sistemas y fuentes de datos en base a algunos criterios comunes. Se podrá presentar resultados operativos y financieros de manera rápida y sencilla. Visto que se consumirá menos tiempo y recursos por parte de su equipo.

Se podrán gestionar los escenarios hipotéticos con mayor flexibilidad, lo que permite generar y consolidar informes de todas las previsiones, presupuestos y planes financieros en una forma dinámica.

En Resumen con Oracle Hyperion Financial Management podremos poner la base hacia un nuevo modelo de gestión más ágil e inteligente.

SAP- Business Object

SAP BusinessObjects comprende un conjunto de herramientas que permitirán a las empresas estar mejor preparadas para analizar, detectar y prever lo que ocurre en la empresa.

El objetivo es convertir los datos de nuestra empresa en información significativa y útil, para explotarla y ser distribuida a aquellos que la necesitan, para que puedan tomar decisiones oportunas.

También le permite a las empresas que integren sus datos procedentes de una amplia variedad de orígenes, así obtendrán una nueva vista actualizada y unificada.

AP BusinessObjects contiene una gran variedad de herramientas cada una adecuada para una necesidad concreta de explotación de datos, las mismas son:

- Reporting operativo y listado: Crystal Reports.
- Informes interactivos y análisis avanzado: Universe Designer y WebIntelligence (Herramientas analíticas).
- Cuadros de mando y Dashboards de indicadores: Xcelsius Dashboard Designer.

Todas estas herramientas permiten la publicación de informes y aplicarlos a los formatos más comunes como PDF, Web, Microsoft Office, dispositivos móviles. Se encuentran integradas en una infraestructura en común (BusinessObjects Enterprise) que incluye los componentes necesarios para la distribución de la información a los usuarios, interfaces personalizadas, además de las funciones administrativas de seguridad, gestión y auditoría de los datos.

BusinessObjects Enterprise fue diseñado para integrarse con las infraestructuras informáticas propias de cada organización, teniendo en cuenta las inversiones ya realizadas. Tiene una gran capacidad de integración que permite responder rápidamente a las nuevas necesidades de análisis y acceso de datos empresariales.

Entre los beneficios SAP BusinessObjects ofrece una herramienta adecuada para cada una de las necesidades de Business Intelligence, todas integradas en una plataforma única.

Por medio de la centralización y de la integración de los datos analíticos de su empresa se podrá publicar, crear y administrar informes, también cuadros de mando, dashboards de indicadores o realizar un análisis avanzado de una forma más eficiente, segura y controlada. Esto permite a los empleados de toda nuestra organización medir resultados de procesos, realizar análisis de datos y facilitar a una toma de decisiones acertadas.

MicroStrategy

MicroStrategy es una plataforma de Business Intelligence, permite analizar y crear informes de los datos almacenados en una base de datos relacional y/o otras fuentes, transformando a los mismos en cuadros de mando e informes para un mayor conocimiento y análisis de los mismos para la posterior toma de decisiones. Se describe al software de informes núcleo como "ROLAP" u "OLAP Relacional" para remarcar el uso de tecnología en base de datos relacional y que se distinga del OLAP tradicional, por lo que también soporta tecnología MOLAP.

El software más reciente se denomina MicroStrategy 9. En esta versión se incluyen mejoras importantes, proporcionando a los negocios la capacidad de construir sus propios cuadros de mando en cuestión de unos minutos.

En la actualidad cuenta con una fuerte estrategia en el área de las aplicaciones móviles con su producto Microstrategy Mobile, el cual es una plataforma para las aplicaciones móviles que permiten a las empresas construir una gran variedad de apps móviles

proporcionando inteligencia de negocios, de transacciones y contenidos multimedia en tablets y en smartphones.

MicroStrategy Cloud es un producto que permite a las empresas aprovechar los beneficios de la plataforma BI de una manera muy fácil, permitiendo implementar aplicaciones a miles de usuarios en pocas semanas en vez de años. Nos ayuda a reducir los riesgos del proyecto y los costos operativos para eliminar los gastos de capital.

MicroStrategy Wisdom Professional ofrece una gran variedad de estudios de marketing sin precedentes en la demografía, así mismo intereses y gráficos sociales de millones de usuarios de Facebook. Esta inteligencia colectiva nos ayuda en actividades con respecto al marketing, ofrece mejores productos y servicios, con lo cual ayuda a competir mejor en el mercado.

QlikView

Qlik View es una herramienta de BI que tiene todo integrado, tanto el ETL como el reporting en la misma aplicación. De este software se va a hablar más adelante en detalle, dado a que es el único de los mencionados anteriormente que trabaja exclusivamente con modelos asociativos, el tema central de nuestra tesis.

2.4. ¿Qué es QlikView? ¿Como funciona?

Es una solución de Business Intelligence creada por la empresa QlikTech⁵. Trabaja con una nueva tecnología de lógica asociativa en memoria, la cual permite llevar a cabo cálculos en tiempo real y navegar por la información de una manera más amigable que con otras soluciones.

En la actualidad la memoria RAM (Random-Access Memory: Memoria de acceso aleatorio) no supone un problema y cada vez tiene más capacidad. Tomando en cuenta esta afirmación varias compañías, entre ellas QlikView, han utilizado esta forma de trabajo en memoria compitiendo y sacando ventaja a otras arquitecturas mucho más caras y complejas, ya que requieren otros tipos de recursos como gran capacidad de memoria, buffers, entre otras cosas. Los datos se almacenan de forma pre-calculada en memoria.

⁵ QlikTech es la empresa que se encuentra detrás de QlikView, la plataforma líder de Business Discovery que aporta una verdadera inteligencia de negocio (Business Intelligence, BI) dirigida por el usuario. Nuestra misión, "Simplificar las decisiones para todos, en cualquier lugar," busca aprovechar plenamente todo el potencial de los datos, grandes o pequeños, y liberarlo para poder cambiar el mundo.

Su Lógica Asociativa

QlikView trabaja sus datos con una lógica denominada AQL (*Associative Query Logic*). Es una tecnología de desarrollo de bases de datos aplicada en QlikView y patentada por QlikTech, que la define como sigue:

“Cuando se comprime la información, se comprime y normaliza al máximo para que no haya información redundante. Cada valor único para todos los datos se almacena una sola vez y se referencia a través de punteros. Por ejemplo, si el primer registro de una fuente de datos incluye el campo ‘coche rojo’ y la segunda incluye el valor ‘coche negro’, se almacena ‘coche’ una sola vez. En lugar de almacenar dos veces ‘coche’, un contador asociado a un puntero referencia el incremento de ese valor.” (QlikTech, 2006)

Esta tecnología consiste en mantener en memoria RAM una base de datos no relacional, sino asociativa. La lógica asociativa permite a los usuarios manejar los datos de la misma manera que piensan, asociando hechos y permitiéndoles ver al instante las conexiones y relaciones entre los datos, por sí mismos.

La principal ventaja de la lógica asociativa es que los datos fuente son retenidos y puestos a disposición inmediatamente. El resultado en QlikView es su potente capacidad analítica que permite explorar y descubrir relaciones no obvias en los datos, aportando mayor conocimiento al usuario sobre éstos.

Arquitectura QlikView

La arquitectura de los sistemas de Business Intelligence tradicionales extraen los datos y los almacenan en datawarehouses y en estructuras multidimensionales, para que finalmente sean accedidos por diversas herramientas, generando los reportes necesarios y para ser analizados por los usuarios.

A continuación mostramos un esquema de cómo es la arquitectura de los sistemas de BI tradicionales:

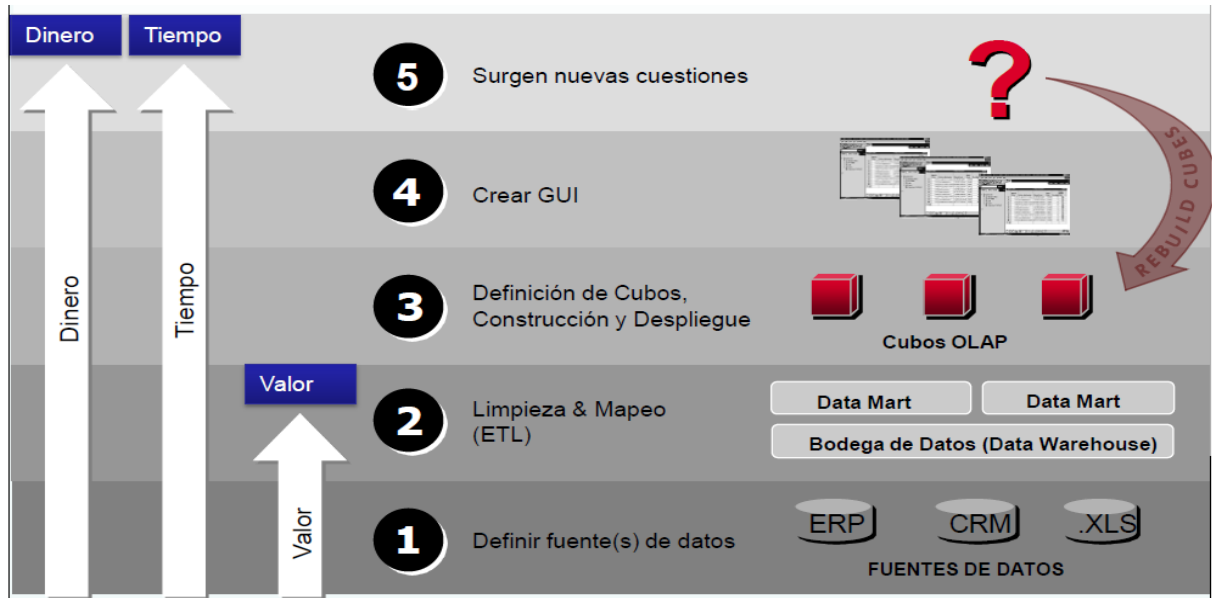


Figura 2: Arquitectura de los sistemas de BI tradicionales (Fuente: IteaSOLUTIONS/QlikTech)

QlikView cambia el concepto de la arquitectura antes mencionada y se basa en un modelo asociativo, el cual reside en memoria. Cada dato del conjunto analítico está asociado a todos los demás datos del conjunto total de datos. Por conjuntos de datos entendemos cientos de tablas, con miles de campos. De esta manera obtiene un modelo dinámico, ágil y veloz a diferencia de otras herramientas. QlikView abarca todos los pasos de ETL, extrae, transforma y carga los datos de diferentes orígenes de datos directamente a memoria.

A continuación mostramos un esquema de cómo es la arquitectura de QlikView:



Figura 3: Arquitectura de los sistemas de BI 2.0 (Fuente: IteaSOLUTIONS/QlikTech)

Los datos se encuentran precargados en memoria, prescindiendo de crear cubos OLAP multidimensionales y sin la necesidad de un datawarehouse, estas son estructuras costosas, lentas de crear y cuyo mantenimiento lleva tiempo. Si llegan a cambiar los datos, habría que actualizar el datawarehouse y crear de nuevo el cubo multidimensional, en QlikView debemos agregar solamente un campo al modelo y adaptarlo al mismo.

El tipo de arquitectura residente en memoria propuesto por QlikView es posible gracias a las soluciones hardware que existen hoy en día de 64 bits, con gran capacidad de memoria y procesadores multinúcleo que permiten realizar cálculos paralelos con un gran rendimiento. Esta ha sido ampliamente aceptada y así queda reflejado en el cuadrante mágico de Gartner.

Al conservarse los datos en memoria, se eliminan las limitaciones de los cubos en disco y se puede acceder más rápidamente a los datos, así por ejemplo la respuesta a cualquier selección que utilice un cálculo se realiza de manera mucho más rápida, ya que los datos se encuentran precargados.

Aprovechando la velocidad de residir en memoria, QlikView permite realizar tareas que las soluciones tradicionales sobre disco no permiten como:

- Poder modificar el análisis a gusto del usuario, pudiendo realizar cálculos en tiempo real.
- Entender las relaciones asociativas disponiendo del conjunto total de los datos.
- Realizar análisis complejos como por ejemplo de conjuntos, las soluciones tradicionales consumirían mucho tiempo para realizarlos.
- Llevar a cabo análisis y presentarlos con rapidez y de forma simultánea a usuarios conectados en diferentes localizaciones sin problemas de carga.
- Se eliminan problemas que traen las soluciones tradicionales de disco como son la incapacidad de crear análisis con rapidez. y el acceso lento a los datos.

Resumiendo, QlikView se trata de una solución BI completa, soportando las herramientas BI tradicionales en una misma arquitectura. Lo que permite consolidar los datos, realizar búsquedas asociadas, realizar cálculos y presentar visualmente los resultados mediante cuadros de mando, análisis e informes, todo desde una misma versión.

Las razones que nos llevan a elegir la plataforma BI QlikView para dar nuestros ejemplos son las siguientes:

- Innovadora arquitectura que rompe con los sistemas de BI tradicionales, los cuales utilizan modelo dimensionales.
- Trabaja exclusivamente con modelos asociativos, el tema principal de nuestra tesis.
- Por tercer año consecutivo ocupa el primer puesto en la encuesta BI Survey 9⁶, superando incluso a las plataformas Hyperion (Oracle), Business Object (SAP) y Cognos (IBM).
- Se trata de una solución completa, se encuentra disponible desde su sitio web de manera gratuita para trabajar de formar personal, requisito indispensable en la elección de la herramienta para este proyecto, ya que las licencias de este tipo de productos son altamente caras.

2.5. Modelo Asociativo

El modelo asociativo es un modelo lógico de base de datos que se basa en la unión de tablas de acuerdo a hechos relacionados, es decir por un campo particular, independientemente de cuál sea, relacionar tablas en el modelo y hacer filtros en el mismo.

Ayuda a los usuarios a encontrar datos empleando las palabras claves más relevantes. Por ejemplo: Si busco la palabra **casa**, muestra todos los datos asociados a **casa**, siempre y cuando estén las tablas en el mismo modelo y unidas por el mismo nombre de atributo. De esta forma se pueden explorar valores conociendo únicamente hechos relacionados, incluso cuando no se conoce la estructura de los mismos o la sintaxis.

La búsqueda asociativa permite realizar la búsqueda de información asociada en todo el modelo a un valor deseado. Permitiendo a que los usuarios busquen e interactúen con sus datos de la misma forma en que piensan, es decir, de forma asociativa. Al filtrar un concepto en una tabla se filtren de manera asociativa el resto de los datos vinculados a ese concepto.

Según lo dicho por Simon Williams⁷, creador y autor del modelo asociativo: *“si clasificamos los objetos del mundo real como entidades y asociaciones, el modelo asociativo separa dos ideas. Por un lado, la idea que cualquier elemento existe de forma discreta e independiente; y por otro lado, la idea de que existen diferentes maneras de que un elemento*

⁶ La encuesta BI Survey es la mayor encuesta realiza a usuarios de Business Intelligence. Es realizada cada año por Business Application Research Center (BARC). BI Survey 9 se realizó. mediante un estudio online sobre un espacio muestral de 2000 usuarios y analizando 23 productos BI.

⁷ Simon Williams fue el creador de la idea de modelo asociativo y fundador de la empresa Lazysoft, es una alternativa a los modelos de datos utilizados tradicionalmente en los sistemas de bases de datos que normalmente están basados en filas.

interactúe con otros elementos. Cada interacción es un elemento en sí que se quiere conservar. Una asociación en el mundo real es representada como una asociación entre dos elementos, donde cada uno probablemente será un elemento o una asociación. Por lo tanto, este modelo es parecido al funcionamiento del cerebro humano, capaz de relacionar elementos (información) entre sí”.

Partiendo de esta base consideramos que el objetivo del modelo asociativo es no preocupar a los desarrolladores con la estructura lógica de los datos, viendo al instante las conexiones y relaciones entre los datos residentes en diferentes aplicaciones, sistemas, organizaciones y regiones, todos por sí mismos, siempre y cuando se encuentren dentro del mismo modelo.

Si pensamos en un directivo de ventas, volviendo a su oficina tras asistir a una reunión de avance en la empresa y desea recordar el nombre de los productos que más piden y cuáles no, pero él solo recuerda la fecha y las sucursales donde se venden esos productos, buscando por los valores que recuerda podría llegar a encontrar el nombre de los productos mencionados.

Para entender mejor el concepto de modelo asociativo vamos a un ejemplo práctico:

- Tenemos el siguiente **modelo de datos** compuesto por cinco tablas:

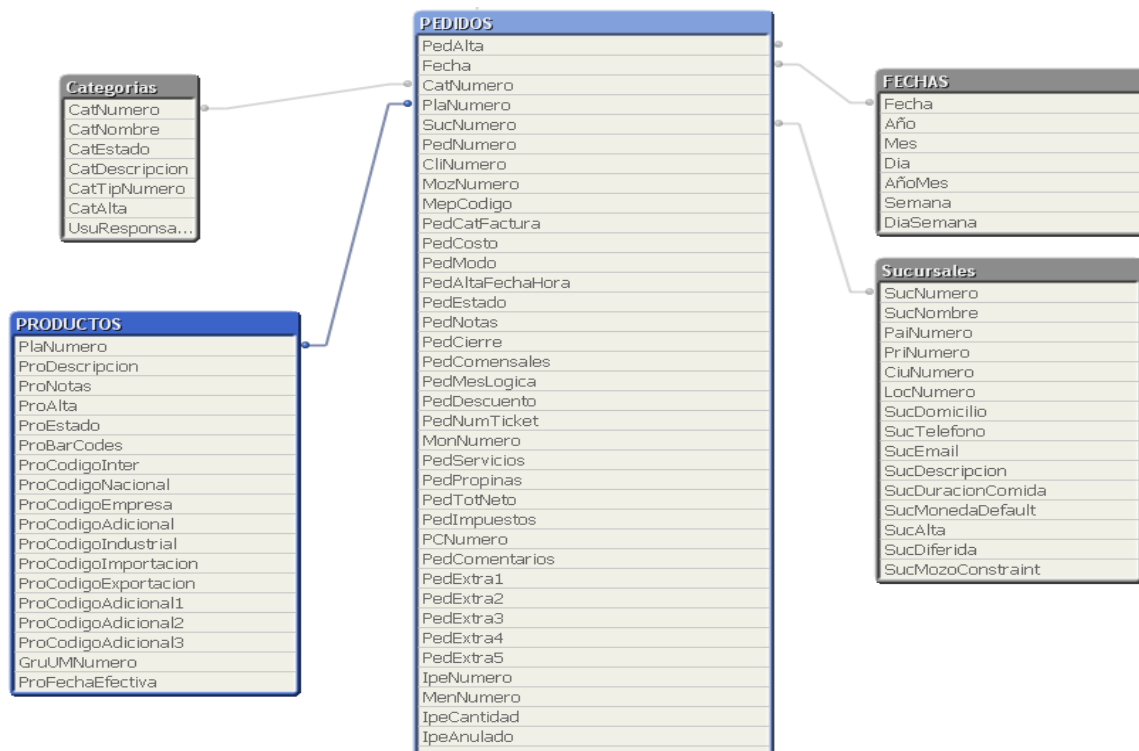


Figura 4: Modelo asociativo compuesto por cinco tablas.

- Se denomina **campo** a cada atributo de una tabla. Por ejemplo:

Tabla: Pedidos.

Campos: PedAlta, Fecha, CatNumero, PlaNumero, SucNumero, etc.

- Las tablas se encuentran unidas entre ellas por algún campo clave. Por ejemplo:

La tabla **Pedidos** se asocia con la tabla **Sucursales** con sólo tener el campo con el mismo nombre en ambas tablas, en este caso **SucNumero**.

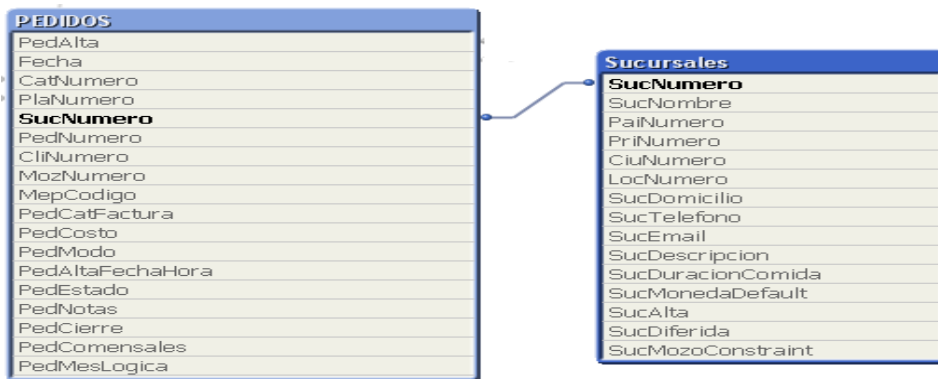


Figura 5: Relación entre dos tablas del modelo asociativo por número de sucursal.

- Los campos claves pueden ser de cualquier índole (Numérico, Texto, etc), sin necesidad de ser exclusivamente numéricos.
- Al relacionar las tablas, la columna de una tabla debe estar referida al mismo concepto en la otra, en QlikView se hace simplemente con nombrar un campo igual, de esta forma se corre el riesgo de asociar tablas que no tienen relación entre sí por solo tener un campo con el mismo nombre. Más adelante se va a tratar aspectos de diseño a tener en cuenta a la hora de realizar un modelo asociativo.
- Si uno quiere ver los pedidos que tiene una sucursal basta con seleccionar el campo: **SucNumero** o cualquier campo de la tabla **Sucursales** para que traiga la información. Por ejemplo: en este caso seleccionamos el campo **SucTeléfono** y nos muestra todos los pedidos de esa sucursal sin necesidad de seleccionar el código de sucursal ó el nombre de la misma.



The image shows two overlapping database window screenshots. The left window, titled 'Sucursales', contains a table with three columns: 'SucNumero', 'SucNombre', and 'SucTelefono'. It displays one record for 'Pilar' with phone number '(02322) 47-3507'. The right window, titled 'PEDIDOS', contains a table with four columns: 'SucNumero', 'Fecha', 'Numero Pedido', and 'ProDescripcion'. It displays 18 records of orders from Sucursal 1 on 01/01/2008, listing various items like 'Kilo', 'Jumbo Check', 'Capelina', 'Americano', 'Vaso III', 'Espresso', and 'lagrima'.

Figura 6: Información de los pedidos que tiene la sucursal Pilar.

Lo mencionado anteriormente es la ventaja de tener un modelo asociativo, con que una tabla se encuentre unida por algún valor válido con otra, se puede buscar por cualquier campo de la misma, ya sea por número de pedido, teléfono, fecha, etc. Todos los datos se encuentran asociados con todo el universo presente en el modelo.

2.6. ¿Qué es una Medida?

Una medida es un atributo numérico de un hecho que representa la performance o comportamiento del negocio relativo a la dimensión.

Identifican a una serie de valores de datos en particular (ventas en \$, cantidad de productos, total de transacciones, margen, costo, etc.).

Las posibles combinaciones de valores de datos, están determinadas por las dimensiones que definen (cruzan) a una medida y su periodicidad. La periodicidad se refiere a si esta medida esta expresada en forma anual, trimestral, mensual o diaria.

Las medidas están dimensionadas en forma Independiente, las complejas, resultan de cálculos.

La medida o hecho es un dato numérico que representa una actividad específica de un negocio, mientras que una dimensión representa una perspectiva de los datos. Cada dimensión está descrita por un conjunto de atributos (datos agregados). A su vez pueden intersectar estas dimensiones para obtener un valor, llamado medida.

Los hechos representan el patrón de interés o el evento dentro de una empresa que necesita ser analizado. Los hechos son implícitamente definidos para la combinación de valores de las dimensiones. Un Data warehouse comúnmente maneja tres tipos de hechos:

- Eventos: con la granularidad mas fina, típicamente modela eventos del mundo real.
- Fotos fijas (snapshots): Modelan entidades en un punto dado en el tiempo.
- Fotos fijas acumulativas: Modelan actividades en un punto dado en el tiempo.

Una medida contiene una propiedad numérica y una fórmula. Existen tres clases de medidas:

- a. Medidas aditivas: Pueden ser combinadas a lo largo de cualquier dimensión.
- b. Medidas semi-aditivas: No pueden ser combinadas a lo largo de una o más dimensiones.
- c. Medidas no aditivas: No pueden ser combinadas a lo largo de ninguna dimensión.

2.6.1. Tipos de medidas

Totalmente Aditivas

Una medida aditiva, también denominada medida completamente aditiva, son todas aquellas que puedo sumarlas tomando cualquier dimensión, se pueden agregar a todas las dimensiones de la tabla de hechos (es la tabla central de un esquema dimensional y contiene los valores de las medidas de negocio). Representa el tipo más común de medida y se utiliza en varias dimensiones con el fin de realizar sumas.

Este tipo de medida se define como una medida que al aplicarle la función de agregación SUM (Suma) sobre las dimensiones siempre se obtienen valores semánticamente correctos. Es uniforme, por lo cual las agregaciones se aplican de la misma forma a todas las dimensiones. Ej.: ventas, suma de las ventas de la semana o de todos los locales.

Ejemplo para explicar una medida aditiva

Los datos de una consulta se obtienen al agregar las medidas sobre niveles de las jerarquías, generalmente la función de agregación utilizada es la función SUM.

Las medidas son cantidades que representan el número de unidades de un producto vendidas en una tienda en un día determinado, número de clientes que representa el número de clientes distintos que han comprado un producto en una tienda un día determinado.

En el conjunto de datos de la Tabla I, el valor de la medida cantidad representa la cantidad de artículos vendidos de un producto en un día en un local, de esta forma durante el día 1 el local A vendió 15 unidades del producto X.


Producto	Día\Local	A	B	C	Totales
X	1	15	15	15	45
	2	15	16	13	44
	3	17	45	10	72
	4	12	37	18	67
	Total	59	113	56	228
Y	1	11	53	12	76
	2	18	78	19	115
	3	14	14	32	64
	4	13	13	22	49
	Total	56	163	85	304
Total	115	276	141	532

Total por producto y local
(Agregación sobre día)

Total por producto y día
(Agregación sobre local)

Total por producto
(Agregación sobre día y local)

Total de ventas



Total por local (agregación sobre producto y día)

Tabla I: Totales de la medida cantidad.

El valor obtenido de los totales en cada celda fue producto de realizar la función de agregación SUM sobre las dimensiones (tiempo, localización y producto), vemos que el valor es semánticamente correcto, por ejemplo el total para el día 1, producto X igual a 45 es semánticamente correcto, al igual que para los demás casos por lo que la medida cantidad es aditiva respecto a las tres dimensiones del esquema.

Semi-Aditivas

Las medidas semi-aditivas se pueden agregar a algunas dimensiones, pero no a todas. Por ejemplo, el recuento de personas y el inventario se consideran medidas semi-aditivas

Este tipo de medidas son aquellas que en algunos casos no son aditivas, nos permiten sumar cualquier dimensión menos para el tiempo. Se usa mucho por ejemplo para el Stock de depósitos (no aditivo con respecto al tiempo). Por ejemplo quiero ver la cantidad de

empleados que tiene la consultora por día, no es lo mismo que sacarlo por mes ni por año, ya que solo debo tomar en cuenta lo que hay activo al día de hoy.

Una medida semi-aditiva se puede agregar en algunas dimensiones, pero no todas las incluidas en el grupo de medida que contiene la medida. Por ejemplo, una medida que representa la cantidad disponible para inventario puede agregarse en una dimensión de geografía para generar una cantidad total disponible para todos los almacenes, pero la medida no se puede agregar en una dimensión de tiempo porque representa una instantánea periódica de las cantidades disponibles. Agregar dicha medida en una dimensión de tiempo generaría resultados incorrectos.

Ejemplo para explicar una medida Semi-Aditiva

La agregación de las medidas puede ser incorrecta sobre algunas de las dimensiones, estas medidas son denominadas semi-aditivas. Por ejemplo, el número de clientes se denomina una medida semi-aditiva ya que al realizar la función de agregación SUM sobre la dimensión producto nos arroja valores incorrectos, esto se pueden presentar por razones semánticas es decir por el significado de la medida o por el tipo de jerarquía sobre la cual se esta realizando la función de agregación.

La Tabla II, expresa el número de clientes que compraron cada día en cada local, por ejemplo, 13 clientes compraron durante el día 1 en el local A.

DíaLocal	A	B	C
1	13	17	12
2	17	11	17
3	14	10	15
4	16	14	12

Tabla II: Número de clientes por local y día.

En la Tabla III se muestran los artículos que adquirió cada cliente por compra (sólo se consideran los clientes del local A y el día 1), se puede observar que de los 13 clientes del local A tuvo en el día 1, 10 clientes adquirieron el producto X y Y en la misma compra y 3 clientes adquirieron el producto Y y no adquirieron el producto X. (Esta información no está en el AD).

Ciiente	Producto
1	X,Y,...
2	X,Y,...
3	X,Y,...
...
10	X,Y,...
11	Y,Z,...
12	Y,Z,...
13	Y,W,...





Cientes que adquirieron el producto X y Y en la misma compra, el día 1.

Cientes que adquirieron el producto X pero que no adquirieron el producto Y en la misma compra, el día 1.

Tabla III: Relación de productos adquiridos por los clientes, en el local A, el día 1.

En la Tabla IV se muestra la relación de clientes que hubo por producto, local y día y se destaca que 10 clientes compraron durante el primer día el producto X en el local A, y que esos mismos 10 clientes compraron el producto Y en el mismo local el mismo día, pero además 3 clientes distintos compraron el producto Y en el local 1 el mismo día, así el total de clientes que el local A tuvo durante el día 1 fue realmente 13, (el mismo razonamiento debe hacerse para los demás casos).

	Día\Local	A	B	C	Totales
Producto					
X	1	10	15	10	35
	2	15	9	10	34
	3	10	9+2	10	30
	4	10	10	10	30
	Total	45	44	40	129
Y	1	10+3	15+2	10+2	42
	2	15+2	9+2	10+7	45
	3	10+4	8	10+5	37
	4	10+6	10+4	10+2	42
	Total	60	50	56	166
Total	,,,,,,	60	52	56	168

 Total por local (correcto)

Tabla IV: Totales esperados para la medida número de clientes.

Este tipo de medidas son aquellas que no se pueden agregar a ninguna de las dimensiones. No se pueden agregar lógicamente entre registros o filas de hechos. Generalmente son el resultado de proporciones u otros cálculos matemáticos. El único cálculo que se puede realizar para dicha medida es obtener un recuento del número de filas de tales medidas.

Una medida no aditiva no se puede agregar en ninguna dimensión en el grupo de medida que contiene la medida. En su lugar, la medida debe calcularse de forma individual para cada celda del cubo que representa la medida. Por ejemplo, una medida calculada que devuelve un porcentaje, por ejemplo, un margen de beneficio, no se puede agregar a partir de los valores de porcentaje de los miembros secundarios en cualquier dimensión.

Ejemplo para explicar una medida no aditiva

Con este ejemplo podemos decir que cuando una medida no es aditiva usaremos el siguiente conjunto de datos (Tabla VI), que es la representación del total de estudiantes registrados en la universidad, los mismos clasificados por departamento y año.

Depto\Año	2001	2002	2003	Totales
Sistemas	15	17	13	45
Química	10	15	11	36
Total	25	32	24	81

Tabla VI: Relación de estudiantes registrados en la Universidad.

Si queremos obtener un total de estudiantes que se atendió en cada departamento en los últimos tres años. La agregación de datos ($15+17+13=45$) es totalmente incorrecta ya que suponemos que los estudiantes pueden ser atendidos por un mismo departamento durante varios años. Podemos ver que en el año 2001, el departamento de Sistemas atendió a 15 estudiantes que permanecieron en un mismo departamento hasta el año 2002, ya que en ese año 2002 se registraron sólo dos nuevos estudiantes en sistemas, los dos estudiantes continuaron hasta el 2003, de esta manera en el año 2003 se registraron solo 11 estudiantes nuevos en el departamento, así el número de estudiantes total atendidos por el departamento de sistemas en los tres años es $15+2+11=28$.

La Tabla VII nos muestra que el número de estudiantes es una medida que puede agregarse sobre la dimensión departamento (si suponemos que un mismo estudiante no puede

estar registrado en más de un departamento un mismo año) pero que no puede agregarse sobre la dimensión año.

Depto\Año	2001	2002	2003	Totales
Sistemas	15	15+2	2+11	28
Química	10	10+5	5+6	21
Total	25	32	24	49

Tabla VII: Medida semi-aditiva sobre la dimensión Tiempo.

Podemos analizar un conjunto de datos que se muestra en la Tabla VIII suponiendo que los estudiantes pueden elegir materias de diferentes departamentos en un mismo año. En este ejemplo, un estudiante puede ser atendido por un departamento de sistemas y por un departamento de química y luego continuar dos años en dicho programa por lo que el alumno se sumaría dos veces sobre la dimensión Departamento y dos veces sobre la dimensión año. Esta medida, que denominamos “número de estudiantes registrados”, sería en conclusión una medida no aditiva.

Depto\Año	2001	2002	2003	Totales
Sistemas	14+1	17	13	28
Química	9+1	15	11	21
Total	25	32	24	49

Tabla VIII: Resultados incorrectos por la medida no aditiva.

3. Desarrollo

3.1. Pautas de diseño al trabajar con modelos asociativos

Cuando trabajemos con modelos asociativos hay varios puntos a tener en cuenta, ya que no debemos pensarlo como un modelo dimensional. A continuación se darán ciertas pautas a tener en cuenta para encarar el modelo de forma correcta en la herramienta Qlikview.

Construir un modelo de datos es un arte y no una ciencia. En muchos casos un modelo Des-Normalizado es mejor a uno normalizado. Tenemos varios factores claves que afectan el modelo:

- a. Número de filas y columnas.
- b. Número de valores distintos x columnas.
- c. Estructura de datos.
- d. Número de objetos en cada hoja.

3.1.1. Claves sintéticas

¿Qué son las claves sintéticas?

Cuando dos tablas se relacionan a través de dos o más campos clave, QlikView automáticamente genera una clave sintética y de esta forma puede relacionar ambas tablas en forma asociativa. Hace una codificación de los valores de los campos clave que asigna a un nuevo campo. QlikView asigna el nombre \$Syn1 para la primer clave sintética, \$Syn2 para la segunda, \$Syn3 para la tercera y así sucesivamente.

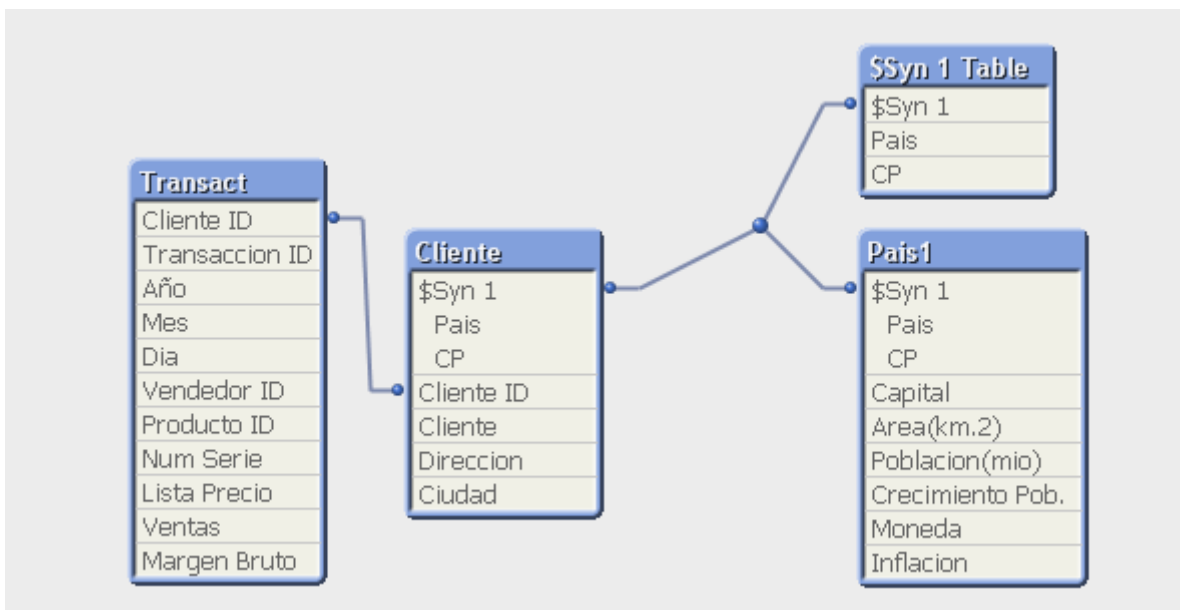


Figura 7: Modelo asociativo con claves Sintéticas-\$Syn1.

¿Por qué conviene evitar las claves sintéticas?

Las claves sintéticas en general son un recurso pesado y pueden retardar los cálculos y, en casos extremos, sobrecargar una aplicación desarrollada en alguna herramienta de Business Intelligence. También hacen que los documentos sean más difíciles de entender y mantener.

Existen tres formas de evitar las claves sintéticas en un modelo asociativo:

1. Concatenación: Uniendo varias tablas en una sola.

2. Clave compuesta: Generando una clave producto de la unión de varios campos.
3. Link Table: Tabla intermedia que nos permita llegar de una tabla a la otra.

A continuación se da un ejemplo real para entender lo mencionado anteriormente.

Ejemplo

Se tienen dos tablas **Facturas** y **Presupuestado**. Ambas tienen campos con el mismo nombre, lo cual nos va a generar una **clave sintética**, ya que el modelo no va a saber porque clave univoca relacionar ambas tablas por tener tres campos iguales.

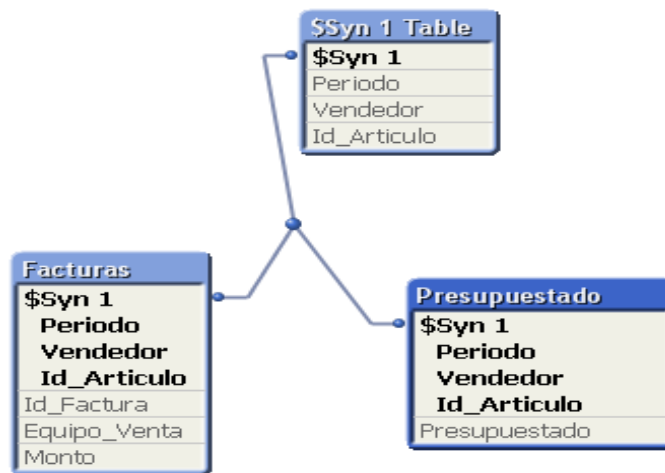


Figura 8: Modelo Asociativo con la relación entre la tabla de facturas y la tabla presupuestado.

```

Facturas:
load

    InvoiceID as Id_Factura,
    SalesTeam as Equipo_Venta,
    Amount as Monto,
    Period as Periodo,
    Salesman as Vendedor,
    ArticleID as Id_Articulo
resident Invoices_tmp;

Presupuestado:
load

    Budget as Presupuestado,
    Period as Periodo,
    Salesman as Vendedor,
    ArticleID as Id_Articulo
resident Budget_tmp;

DROP TABLES Budget_tmp, Invoices_tmp;
    
```

Script 1: Código para cargar la tabla Facturas y Presupuestado en QlikView.

Esto podemos evitarlo si utilizamos cualquiera de las tres técnicas mencionadas anteriormente. Las cuales vamos a explicar una por una.

1. Concatenación

Se unen dos o más tabla en una, formando una tabla principal con todos los campos. Con esto se obtiene una sola tabla en el modelo, ganando tiempo de proceso ya que todos los datos vinculados se encuentran en una sola tabla indexada. No es recomendable si la cantidad de registros es mucho, en este caso en recomendable tener las tablas separadas con la utilización de una tabla link que vincule todas las claves.

Concatenacion_Facturas_Presupuesto
Id_Factura
Periodo
Vendedor
Equipo_Venta
Id_Articulo
Monto
Presupuestado

Figura 9: Unión de la tabla Facturas y Presupuestado.

```
Concatenacion_Facturas_Presupuesto:
Load
    InvoiceID as Id_Factura,
    Period as Periodo,
    Salesman as Vendedor,
    SalesTeam as Equipo_Venta,
    ArticleID as Id_Articulo,
    Amount as Monto
FROM Invoices3.xls (biff, embedded labels, table is [Invoices$]);

Concatenate

    LOAD Period as Periodo,
        Salesman as Vendedor,
        ArticleID as Id_Articulo,
        Budget as Presupuestado
FROM
Budget.xls
(biff, embedded labels, table is Budget$);

EXIT Script;
```

Script 2: Código para concatenar la tabla Facturas y Presupuestado en QlikView.

2. Clave compuesta

Levantamos las dos tablas de forma independiente, pero los campos con el mismo nombre los ponemos dentro de una clave compuesta. En este caso en vez de relacionar a las tablas por Periodo, Vendedor y Id_Articulo de manera independiente, lo hacemos por un campo nuevo generado a partir de la unión de estos tres, el cual se llama Key.

Key= Periodo+ Vendedor + Id_Articulo.

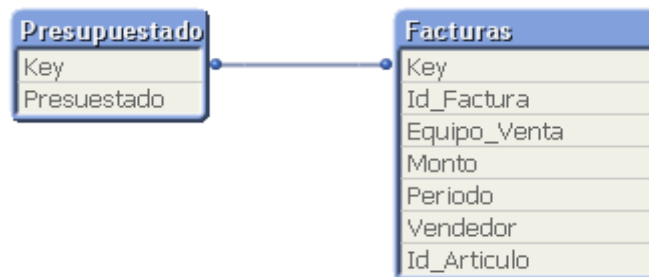


Figura 10: Modelo Asociativo con la relación entre la tabla de Facturas y la tabla Presupuestado mediante una clave compuesta.

```

Facturas:
load
  autonumber(Period & '-' & Salesman & '-' & ArticleID) as Key,
  InvoiceID as Id_Factura,
  SalesTeam as Equipo_Venta,
  Amount as Monto,
  Period as Periodo,
  Salesman as Vendedor,
  ArticleID as Id_Articulo
resident Invoices_tmp;

Presupuestado:
load
  autonumber(Period & '-' & Salesman & '-' & ArticleID) as Key,
  Budget as Presupuestado

resident Budget_tmp;

DROP TABLES Budget_tmp, Invoices_tmp;
  
```

Script 3: Código para unir la tabla Facturas y Presupuestado mediante una clave compuesta en QlikView.

3. Link Table

Generamos una tabla nueva, Link, la cual es una tabla intermedia que relaciona a la tabla Presupuestado con Facturas.

Una tabla link permite vincular dos o más tablas relacionadas, administrando todos los campos claves entre las tablas, la utilización de tablas link genera modelos estrella, con lo cual todo está relacionado con esta tabla link. La utilización de tablas link permite ganar tiempo de procesos en grandes modelos y permite la escalabilidad del modelo.

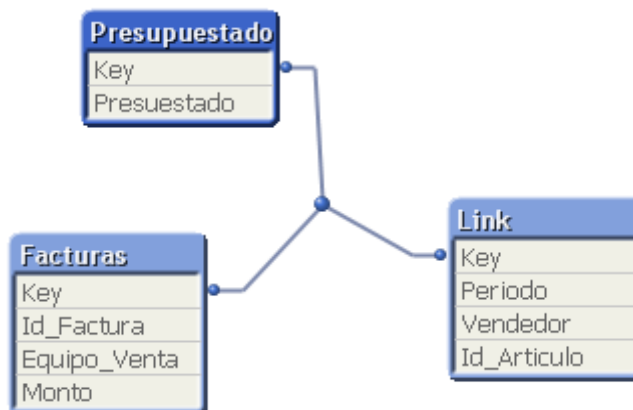


Figura 11: Modelo Asociativo con la relación entre la tabla de Facturas y la tabla Presupuestado mediante una tabla Link.


```

Link:
Load DISTINCT
  Period as Periodo,
  Salesman as Vendedor,
  ArticleID as Id_Articulo,
  Key
RESIDENT Invoices_tmp;
Load DISTINCT
  Period as Periodo,
  Salesman as Vendedor,
  ArticleID as Id_Articulo,
  Key
RESIDENT Budget_tmp;

Facturas:
load
  InvoiceID as Id_Factura,
  SalesTeam as Equipo_Venta,
  Amount as Monto,
  Key
resident Invoices_tmp;

Presupuestado:
load
  Budget as Presupuestado,
  Key
resident Budget_tmp;

DROP TABLES Budget_tmp, Invoices_tmp;
  
```

Script 4: Código para unir la tabla Facturas y Presupuestado mediante un tabla Link en QlikView.

3.1.2. Claves circulares

¿Qué son las claves circulares?

Cuando tres tablas se relacionan a través uno o más campos clave entre las todas las tablas involucradas del modelo asociativo, QlikView automáticamente genera una clave circular, produciendo un bucle entre las mismas.

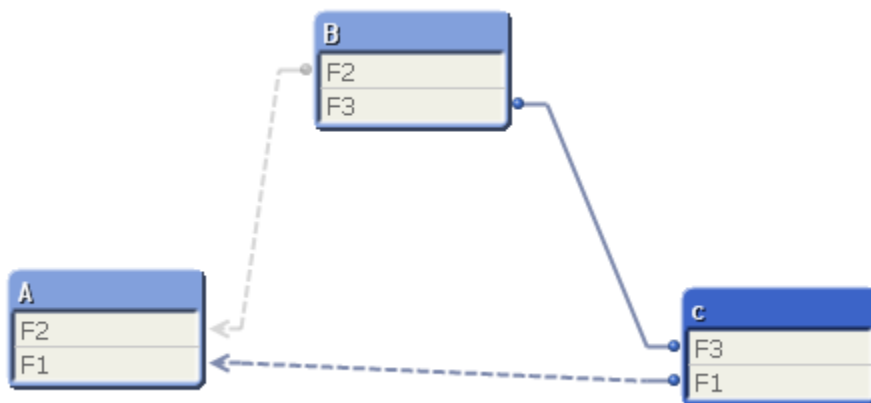


Figura 12: Unión de tablas formando una clave circular.

¿Por qué hay que eliminar las claves circulares?

A diferencia de las sintéticas, las claves circulares generan inconsistencias en la aplicación evitando el uso del modelo. Como las selecciones son persistentes en toda la aplicación, al aplicar un filtro, filtra automáticamente el resto de las tablas relacionadas. Al tener una clave circular haría que se estén filtrando las tablas una y otra vez en un bucle infinito.

En cuanto a la solución de esto es igual a lo mencionado en el ejemplo de claves sintéticas, con la salvedad de que las claves sintéticas pueden convivir en el modelo asociativo pero las circulares no. Lo debemos evitar por algunas de las tres alternativas antes mencionadas.

3.1.3. Uso de variables

Se recomienda utilizar variables que contengan expresiones que representan cálculos, si en un tablero uno utiliza varias veces las mismas fórmulas para diferentes indicadores, uno puede copiar esa fórmula en una variable y luego solo poner la variable en cada grafico, en caso de modificar la expresión se haría una sola vez.

Podemos mencionar una serie de beneficios que nos otorga esto:

- a. Promueve el reuso de la lógica.
- b. Reduce el mantenimiento cuando cambia la lógica o la fuente de datos.
- c. Permite manejo mas centralizado de las medidas de negocio.

Ejemplo

En una empresa de logística se necesita saber lo contratado, pedido, recepcionado y facturado. Cada expresión representa los cálculos necesarios para obtener dicha información según lógica de QlikView:

- **Contratado=** `sum({<EKKO.Sociedad_BUKRS = {'M660', 'M650'}>} If(match(left([EKKO.Documento compras_EBELN],2), '45', 'OJ', 'OC') and ([EKKO.Entrega final_ELIKZ]='X' or [EKKO.Indicador de borrado_LOEKZ]='S'), EKBE.Importe_MD_105_101* [EKKO.Tipo cambio moneda_WKURS]))`
`+ (sum({<EKKO.Sociedad_BUKRS = {'M660', 'M650'}>} if(EKKO.FLAG_CLASE_DOC=1 and (([EKKO.Indicador de borrado_LOEKZ] = 'L' or [EKKO.Indicador de borrado_LOEKZ] = 'S' OR EKKO.FLAG_VENCIDO=1)or EKKO.Ind.liberación_FRGKE='A'), [EKAB.Importe_MD_WE_OE] * [EKKO.Tipo cambio moneda_WKURS]))`
`+ sum({<EKKO.Ind.liberación_FRGKE={'C', ' '}, EKKO.Sociedad_BUKRS = {'M660', 'M650'}>} aggr(NODISTINCT if(EKKO.FLAG_CLASE_DOC=1 and ([EKKO.Indicador de borrado_LOEKZ] = 'L' or [EKKO.Indicador de borrado_LOEKZ] = 'S' OR EKKO.FLAG_VENCIDO=1), 0, [EKKO.Valor previsto_KTWRT]* [EKKO.Tipo cambio moneda_WKURS]), [EKKO.Documento compras_EBELN]))`
`+ sum({<EKKO.Sociedad_BUKRS = {'M660', 'M650'}>} if(EKKO.FLAG_CLASE_DOC=0 and ([EKKO.Indicador de borrado_LOEKZ] = 'L'), (EKPO.Importe_MD_NP* [EKKO.Tipo cambio moneda_WKURS]), 0)))`
`+ sum({<EKKO.Sociedad_BUKRS = {'M660', 'M650'}>} if(EKKO.FLAG_CLASE_DOC=0 and [EKKO.Indicador de borrado_LOEKZ] <> 'L' and [EKKO.Indicador de borrado_LOEKZ] <> 'S' and [EKKO.Entrega final_ELIKZ]<>'X', [EKKO.Valor neto de pedido_NETWR]* [EKKO.Tipo cambio moneda_WKURS]))`
- **Pedido=** `SUM({<[EKAB.Fecha de pedido_BEDAT]={ Year=,Month=>} EKAB.Importe_MD_WE_OE * $(vtcambiopedido))`
- **Recepcionado=** `sum({<EKBE.Fe.contabilización_BUDAT={ '>=$(vpxaini)<=$(vpxafin)' }, Year=,Month=>} EKBE.Importe_MD_105_101 * $(vtcambiorecepcionado))`
- **Facturado=** `(if(TDC='T-NPA', sum({<EKBE_FACTURA.Fe.contabilización_BUDAT={ '>=$(vpxaini)<=$(vpxafin)' }, Year =,Month=>} aggr(SUM({<EKBE_FACTURA.Fe.contabilización_BUDAT={ '>=$(vpxaini)<=$(vpxafin)' }, Year=,Month=>} EKBE_FACTURA.Importe_WRBTR), [EKKO.Tipo cambio moneda_WKURS], [EKKO.Documento compras_EBELN]) * $(vtcambiofacturado)), SUM({<EKBE_FACTURA.Fe.contabilización_BUDAT={ '>=$(vpxaini)<=$(vpxafin)' }, Year =,Month=>} [EKBE_FACTURA.Importe ML_DMBTR])))`

Por ende para mostrar en diferentes tablas o gráficos estos valores, se debe codificar indicador por indicador esto. Como buena práctica se recomienda guardar las formulas en variables, así cada vez que se debe modificar algo, solo tocamos lo que está dentro de la variable. Y la variable ponerla como expresión en cada grafico o indicador.

En este caso la Variable “Var1” representa la expresión de lo Contratado:

```

Var1= sum({<EKKO.Sociedad_BUKRS = {'M660','M650'}>}If(match(left([EKKO.Documento
compras_EBELN],2),'45','0J','0C') and ([EKPO.Entrega final_ELIKZ]='X' or
[EKPO.Indicador de borrado_LOEKZ]='S'),EKBE.Importe_MD_105_101* [EKKO.Tipo cambio
moneda_WKURS]))
+
(sum({<EKKO.Sociedad_BUKRS = {'M660','M650'}>} if(EKKO.FLAG_CLASE_DOC=1 and
(( [EKPO.Indicador de borrado_LOEKZ] = 'L' or [EKPO.Indicador de borrado_LOEKZ] =
'S' OR EKKO.FLAG_VENCIDO=1)or EKKO.Ind.liberación_FRGKE='A' ),
[EKAB.Importe_MD_WE_OE] * [EKKO.Tipo cambio moneda_WKURS]))
+
sum({<EKKO.Ind.liberación_FRGKE={'C',' '},EKKO.Sociedad_BUKRS = {'M660','M650'}>}
aggr(NODISTINCT if(EKKO.FLAG_CLASE_DOC=1 and ([EKPO.Indicador de borrado_LOEKZ] =
'L' or [EKPO.Indicador de borrado_LOEKZ] = 'S' OR EKKO.FLAG_VENCIDO=1),0,
[EKKO.Valor previsto_KTWRT]* [EKKO.Tipo cambio moneda_WKURS]) ,[EKKO.Documento
compras_EBELN]) )
+
sum({<EKKO.Sociedad_BUKRS = {'M660','M650'}>} if(EKKO.FLAG_CLASE_DOC=0 and
([EKPO.Indicador de borrado_LOEKZ] = 'L'),(EKPO.Importe_MD_NP* [EKKO.Tipo cambio
moneda_WKURS]),0)))
    
```

La misma la utilizamos en los indicadores que queremos incorporar a nuestro tablero de control, por ejemplo en la Figura 13 y Figura 14:

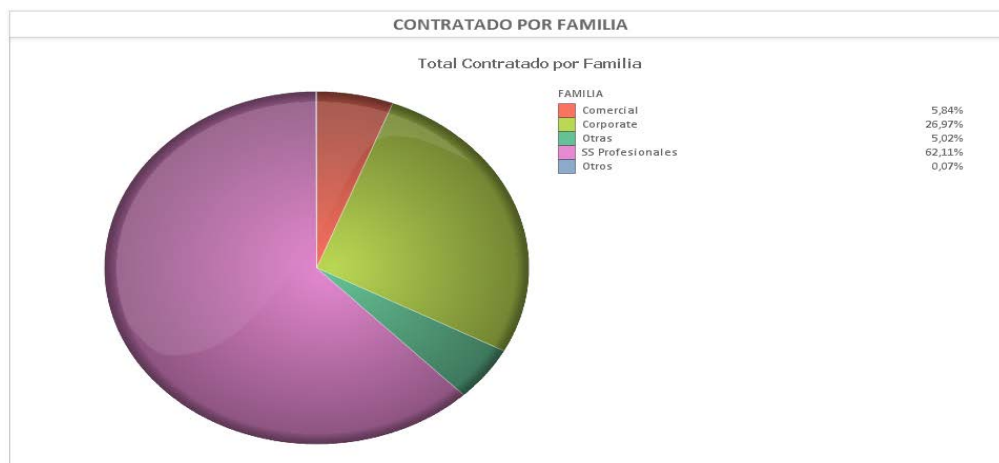


Figura 13: Gráfico de barras que representa lo contratado abierto por familia.

MONTOS POR FAMILIA				
Familia	Contratado	Pedido	Recepcionado	Facturado
SS Profesionales	175.942.298.975	2.409.399	1.483.452	60.000
Corporate	76.407.298.865	49.135.583.626	166.792.040	2.199.862
Comercial	16.531.233.344	34.836.248.675	1.792.831.374	111.124.745
Otras	14.221.722.002	1.529.368.933	1.517.834.545	395.917.721
Tecnología	137.518.575	2.192.377.289	89.466.764	705.125
Medios	29.766.226	279.915.955	406.934.494	180.010
Real Estate	1.066.657	2.166.290	902.075	7.675
Agentes	462.158	922.936	121.403	0
	283.121.366.802	87.978.993.103	3.976.366.147	510.195.137

Figura 14: Representa lo contratado, pedido, recepcionado y facturado por familia.

3.1.4. Claves optimizadas

Una buena práctica a tener cuenta a la hora de diseñar un modelo asociativo es convertir todas las claves alfanuméricas del modelo en numéricas. De esta manera se logra que las claves pesen menos en memoria y sean más optimizadas a la hora de hacer consultas en línea.

En la herramienta Qlikview contamos con la función **AutoNumber** (**'campo'**), la cual nos hace una conversión directa de alfanumérico a numérico del campo establecido dentro de los paréntesis. Ganando velocidad y procesamiento en memoria, dadas las características de los modelos asociativos.

3.1.5. Renombrar campos

A veces es necesario renombrar los campos de las tablas con el fin de obtener las asociaciones correctas.

Suele pasar que dos campos se llamen de forma diferente aunque representen una misma cosa, por ejemplo: **Número de documento** en la tabla **Personas** y **Número de identidad** en la tabla **Registros**. Obviamente, representan una identificación unívoca de la persona y deberían llamarse **NroDocumento** en ambas tablas ó algo similar para poder relacionarse de forma asociativa en el modelo de datos. A su vez, es posible que haya dos campos con el mismo nombre pero que denoten cosas diferentes, por ejemplo: **Fecha** en la tabla **Personas** y **Fecha** en la tabla **Registros**. Estos deberían renombrarse como **FechaPersona** y **FechaRegistro** o algo parecido a ser posible.

Finalmente pueden surgir errores de escritura en los datos que poseen las tablas de las bases de datos, o diferentes convenciones en cuanto al uso de mayúsculas o minúsculas. La herramienta QlikView distingue entre mayúsculas y minúsculas, por ende es importante corregir estas cosas.

3.2. Medidas

Todas las medidas que vamos a tratar serán desarrolladas en la herramienta de Business Intelligence QlikView, la cual trabaja exclusivamente con modelos asociativos. Se expondrá un estándar metodológico, el cual define las fases a tener en cuenta al momento de crear y/o calcular una medida de negocio partiendo de un modelo asociativo.

3.2.1. Estándar metodológico para definir medidas de negocio

En la Figura 15 se muestran los pasos a seguir a la hora de querer llevar a cabo la creación y post implementación de una medida de negocio en una herramienta de Business Intelligence que utiliza exclusivamente modelos asociativos. A su vez se exponen los actores principales que intervienen en cada fase.

Cabe aclarar que los actores representados en la Figura 15 pueden ir variando de acuerdo a cada empresa y a la nomenclatura que use la misma.

Los actores involucrados son los siguientes:

- **Analista Funcional:** persona encargada de hacer todo el relevamiento frente a una necesidad que tiene un usuario, se encarga de plasmar los requerimientos solicitados por el usuario en un documento denominado ERS⁸.
- **Analista Técnico:** persona encargada de transformar los requerimientos plasmados por el analista funcional en términos técnicos. (campos, tablas, sentencias de programación, etc.)
- **Usuario:** persona que define una necesidad.
- **Desarrollador de BI:** persona encargada de desarrollar la aplicación de Business Intelligence acorde al ERS.

⁸ERS (Especificación de Requisitos Software) es una descripción completa del comportamiento del sistema que se va a desarrollar. Incluye un conjunto de casos de uso que describe todas las interacciones que tendrán los usuarios con el software. Los casos de uso también son conocidos como requisitos funcionales. Además de los casos de uso, la ERS también contiene requisitos no funcionales (o complementarios). Los requisitos no funcionales son requisitos que imponen restricciones en el diseño o la implementación (Como por ejemplo restricciones en el diseño o estándares de calidad).

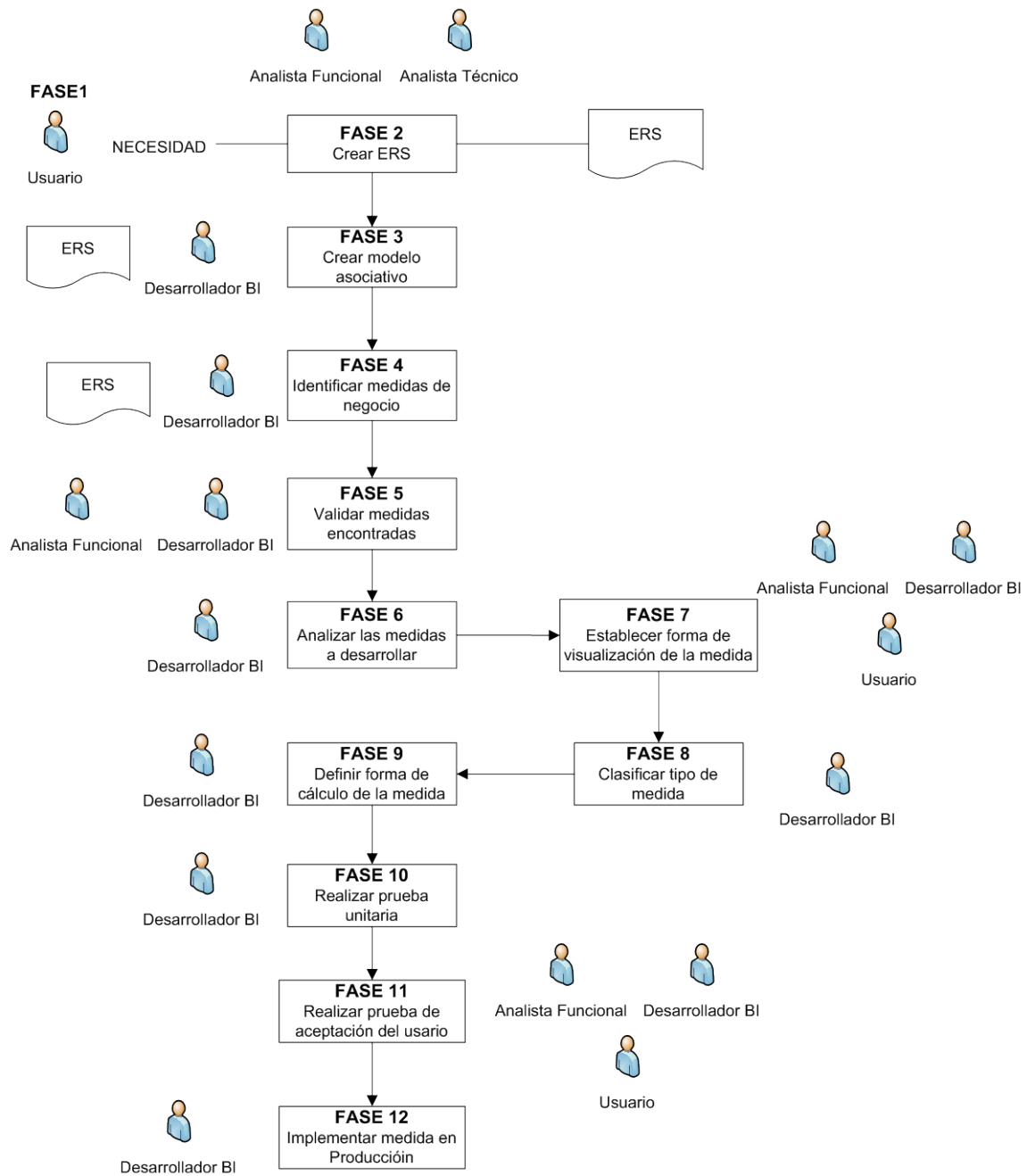


Figura 15: Fases para crear y/o calcular una medida de negocio.

A continuación se detallan las fases representadas en la Figura 15:

Fase 1

Partimos de que el usuario tiene una necesidad y por este motivo carga y/o genera un requerimiento. Esto se va a registrar en un sistema propio de la empresa y/o forma que tenga

la misma para registrarlos. Los requerimientos del usuario determinan el programa de medidas que se va a tener que realizar.

Fase 2

El analista funcional y analista técnico hacen la especificación técnica y funcional llamada ERS.

Para poder llevarla a cabo, se debe tener en cuenta que los usuarios piensan que son ellos saben lo que el software tiene que hacer, por este motivo se necesita contar con mucha experiencia y habilidad para reconocer los requisitos que no estén completos ó sean contradictorios.

El resultado del análisis de requisitos con el cliente es el que se arma en el documento denominado ERS, especificación de requisitos de software. Su estructura la vamos a obtener de la planilla de formulario estándar de cada empresa. En el mismo, el analista técnico debe agregar los campos y/ó tablas necesarias para que luego un desarrollador de BI pueda crear el modelo asociativo que cumpla con lo detallado en el documento.

Fase 3

A partir del requerimiento planteado en el ERS, el desarrollador de BI debe crear el modelo asociativo que cumpla con lo solicitado. En esta fase se plasman las tablas junto a sus relaciones mediante los campos de las mismas en base a la especificación anteriormente mencionada.

En caso de necesitar incorporar algún campo ó tabla al modelo asociativo producto de un error de definición ó cambio de alcance, el modelo se puede realizar tantas veces como sea necesario y de esta forma lograr el más adecuado en base a la necesidad del usuario.

Fase 4

A partir del ERS, identificar las medidas de negocio requeridas para luego validar que sean las adecuadas en la fase 5.

Fase 5

Validar con el analista funcional que las medidas de negocio que se identificaron son las que realmente necesita el usuario, en caso contrario el analista funcional debe comunicarse

con el usuario y redefinir algún punto del ERS. Esta fase puede repetirse las veces que sea necesaria, hasta tanto tener bien identificadas todas las medidas.

Fase 6

Analizar las medidas a realizar, identificar los campos del modelo asociativo que entran en juego para cada medida de negocio. En caso de que haya algún campo y/o tabla faltante se debe volver a la fase 2 para que el analista funcional y/o técnico lo incorporen al ERS.

Fase 7

Establecer la forma de visualizar la medida de negocio. Se debe seleccionar la dimensión por la cual queremos analizar la información y el tipo de gráfico más adecuado para exponerlo (de barras, tabla, torta, dispersión, etc.). Esto queda al sano juicio del Desarrollador de Business Intelligence según el gráfico ó tipo de tabla que crea más conveniente para mostrar la información solicitada por el Usuario. En esta fase puede intervenir el Analista Funcional y/o el Usuario.

La idea es contar con la totalidad de las de dimensiones que definen la información que se quiere visualizar. En cada dimensión debemos definir la operación que se utilizara para agregar la información, por ejemplo, establecer si las ventas diarias al pasar a mes se deben sumar, contar, promediar, etc.

Fase 8

Clasificar el tipo de medida en aditivas, semi-aditivas y no aditivas.

- a. Aditivas:** son las que pueden ser combinadas en cualquier dimensión, es decir, pueden sumarse por cualquiera de ellas.

Los casos más comunes en donde aparece este tipo de medida son:

- venta de productos en general para una determinada fecha.
- medidas de tiempo (Ej. Cantidad de horas trabajadas).
- movimientos bancarios, provenientes de las transacciones de una cuenta.
- importes de artículos.

Para ver este tipo de medida en detalle, ir al punto 3.2.2 en donde están los ejemplos de medidas aditivas hechos en QlikView.

b. Semi-aditivas: son las que se pueden agregar a algunas dimensiones, pero no a todas.

Los casos más comunes en donde aparece este tipo de medida son:

- conteo de personas.
- inventarios y/o stocks.
- todos los hechos que registran un nivel estático (niveles de inventario, saldos contables, medidas de intensidad como temperatura, etc.) son no aditivas en la dimensión tiempo y posiblemente en otras dimensiones.

En conclusión una medida semi-aditiva es aquella que no es sumable en todas las dimensiones.

Para ver este tipo de medida en detalle, ir al punto 3.2.3 en donde están los ejemplos de medidas semi-aditivas hechos en QlikView.

c. No aditivas: son las que no pueden ser combinadas a lo largo de ninguna dimensión, en otras palabras son las que no se suman.

Los casos más comunes en donde aparece este tipo de medida son:

- rentabilidad por producto / venta total – No puede ser sumado en todas las dimensiones.
- precio ó costo unitario – No puede ser sumado en todas las dimensiones.
- porcentajes e indicadores, tales como margen de ventas y el retorno sobre activos son No aditivos.
- datos demográficos (como precios unitarios) deben ser almacenados para obtener promedios, porcentajes o similares de las sumas.

Para ver este tipo de medida en detalle, ir al punto 3.2.4 en donde están los ejemplos de medidas no aditivas hechos en QlikView.

Fase 9

En base a la clasificación definida en la fase anterior, se debe definir la forma de cálculo de cada medida, es decir la fórmula a poner.

En esta fase se identifica el comportamiento de las medidas a calcular, es decir de los valores numéricos que tengo que calcular. Cada medida contiene una propiedad numérica y una formula, por ende debo analizar el comportamiento de las mismas para lograr la fórmula adecuada.

Fase 10

Realizar la prueba unitaria de la medida de negocio. El desarrollador de BI debe verificar que lo desarrollado tenga coherencia con lo solicitado por el usuario, habitualmente cada empresa tiene un formulario estándar para hacer las pruebas unitarias, esta planilla suele estar compuesta por dos partes:

- a. definir casos de pruebas
- b. verificar resultados

En la columna a. se definen todos los casos de pruebas necesarios para contemplar el requerimiento y en la columna b. se especifican los posibles resultados, para cuando sean probados, saber si esta prueba cumplió correctamente con dicha funcionalidad.

También se recomienda tener una planilla de integración, en la cual se van a probar todos los casos definidos en la planilla unitaria. En una columna se seleccionará si el caso paso o no paso ó si hay que volver a ejecutarlo con un lote nuevo de datos.

En caso de encontrar algún error en el cálculo o en la clasificación de la medida se debe volver a la fase 8 ó 9 según corresponda.

Fase 11

Hacer la prueba de aceptación del usuario, en la cual se valida con el usuario si realmente el requerimiento y las medidas desarrolladas están completas, consistentes y corresponden a las necesidad que se está buscando. En conclusión esta fase consiste en formalizar la aceptación de los entregables⁹ del proyecto que se han completado. En caso de ser necesario puede intervenir el analista funcional.

⁹ Se entiende por entregable a cada medida de negocio y/o indicador definido en la especificación de requisitos de software.

Fase 12

Realizar la implementación de la medida de negocio en producción, cumpliendo con todo lo establecido anteriormente. De esta forma se deja a la misma disponible en el entorno productivo de la plataforma de BI correspondiente.

3.2.2. Aditivas

3.2.1.1 Medida: Horario Real Vs Teórico por Dependencia.

Comparación entre el horario que realmente trabaja el empleado en la empresa contra el que debería trabajar según el contrato acordado en su ingreso.

a. Contexto:

En una empresa de telecomunicaciones se desea ver las horas que realmente trabajan sus empleados de la compañía versus las que realmente deberían trabajar tanto por día como mes a nivel dependencia.

Se considera:

- Dependencia: Sector al que depende el empleado a nivel jerárquico.
- Horario Teórico: Es el horario del empleado que se encuentra cargado en el sistema que tiene la empresa.
- Horario Neto: Es el horario que se registra al empleado cuando llega a la empresa mediante los molinetes de la compañía.
- Ausentismo: Horas cargadas en el sistema que tiene la empresa por día frente a un ausentismo del empleado.
- Evento: Horas cargadas en el sistema que tiene la empresa referidas a eventos de empleados (capacitación, home working, etc.)

b. Modelo Asociativo:

Para realizar la Medida mencionada anteriormente, contamos con el siguiente modelo asociativo:

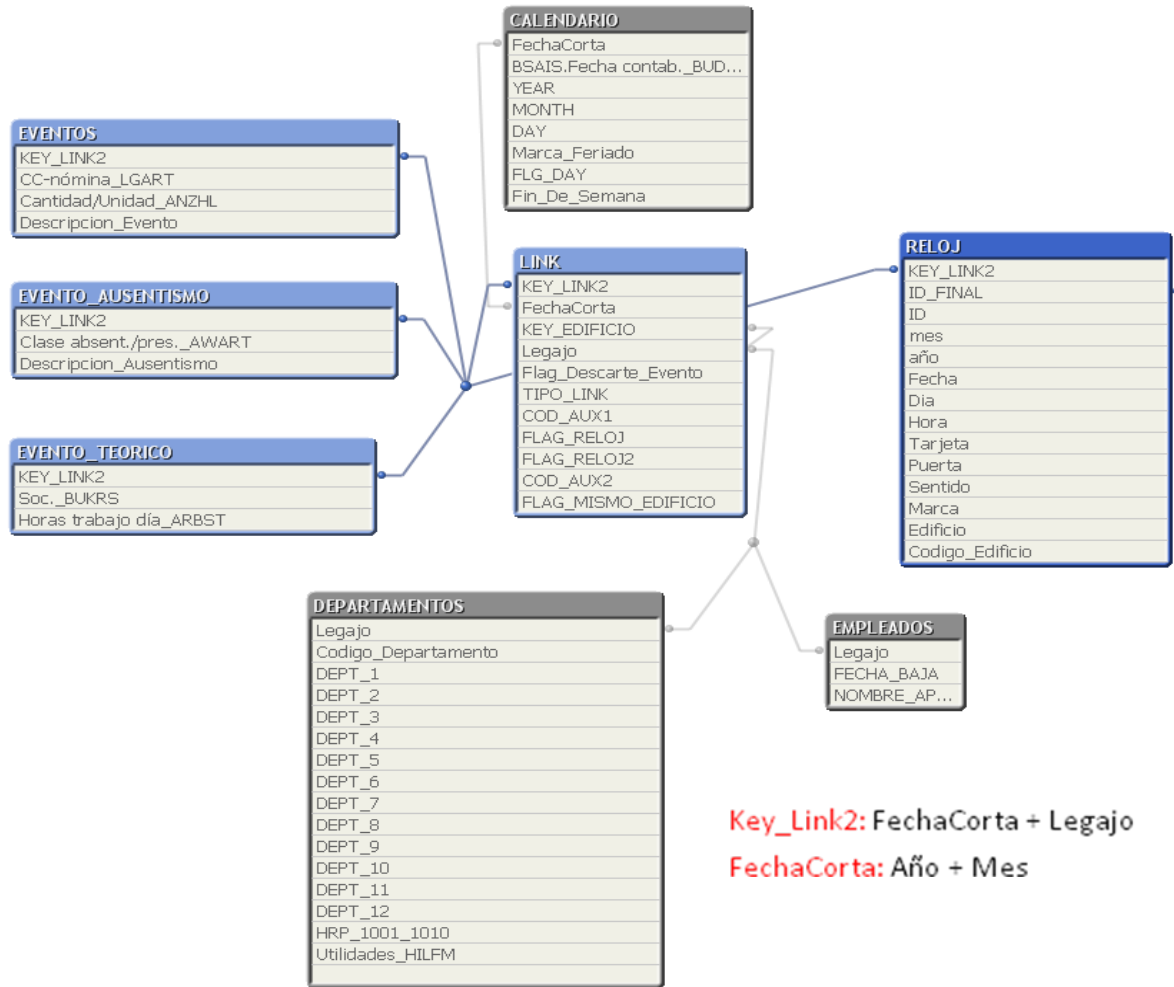


Figura 16: Modelo asociativo referente al control de accesos de los empleados en QlikView.

El mismo está compuesto por las siguientes tablas:

- Eventos: Representa las horas cargadas en un sistema transaccional como eventos.
- Evento_Ausentismo: Representa las horas cargadas en un sistema transaccional como ausentismo.
- Evento_Teórico: Representa las horas registradas en los molinetes como Eventos.
- Link: Tabla clave que relaciona a todas las tablas del modelo mediante un campo común a cada tabla, por ejemplo: Asocia la tabla 'Evento_Ausentismo' con la tabla 'Eventos' por la clave: KEY_LINK.
- Calendario: Representa la fecha.
- Departamentos: Representa los departamentos y/o dependencias de los empleados.
- Reloj: Representa la información proveniente de los molinetes.
- Empleados: Representa información de los empleados cargados en un sistema transaccional.

c. Solución:

Se hará un indicador en donde se pueda ver la cantidad de horas trabajadas (*horario real*) por dependencias de los empleados versus las horas que debería haber trabajado (horas teóricas).

A su vez se detallarán las horas que representan eventos y/o ausentismos, ya que estas se deben sumar a las horas netas para sacar las horas reales.

Las horas trabajadas son las horas reales que marcó el empleado.

- *Horario Real = Horario neto + Ausentismo + Evento.*

Al ser una medida aditiva se demuestra cómo se puede ver por diferentes dimensiones/campos del modelo haciendo siempre el mismo cálculo, una simple suma de horas y arrojando el mismo valor.

En la primera tabla tenemos la cantidad de horas por la **dependencia** que representa la **dirección** a nivel **mensual**. Mientras que en la segunda tabla las tenemos a nivel **diario**.

AÑO	2013	DIA	01 04 07 10 13 16 19 22 25 28 31 02 05 08 11 14 17 20 23 26 29 03 06 09 12 15 18 21 24 27 30						
MES	Ene Feb Mar Abr May Jun Jul Ago Sep Oct Nov Dic								
Real Vs Teórico X Dependencia-Mensual									
Dependencia	Mes	Hs Netas Reloj	Hs Teórica	Ausentismo	Eventos	Hs Real	Real Vs Teórico	Real Vs Teórico	
Dirección	Jul	3447:15:14	11550:00:00	1219:00:00	00:00:00	4666:15:14	-6883:44:46		
Total		3447:15:14	11550:00:00	1219:00:00	00:00:00	4666:15:14	-6883:44:46		
Real Vs Teórico X Dependencia-Diario									
Dependencia	Día	Hs Netas Reloj	Hs Teórica	Ausentismo	Eventos	Hs Real	Real Vs Teórico	Real Vs Teórico	
Dirección	01	00:00:00	525:00:00	24:00:00	00:00:00	24:00:00	-501:00:00		
	02	00:00:00	525:00:00	24:00:00	00:00:00	24:00:00	-501:00:00		
	03	00:00:00	525:00:00	32:00:00	00:00:00	32:00:00	-493:00:00		
	04	00:00:00	525:00:00	32:00:00	00:00:00	32:00:00	-493:00:00		
	05	00:00:00	525:00:00	40:00:00	00:00:00	40:00:00	-485:00:00		
	08	00:00:00	525:00:00	40:00:00	00:00:00	40:00:00	-485:00:00		
	10	254:40:54	525:00:00	24:00:00	00:00:00	278:40:54	-246:19:06		
	11	239:56:16	525:00:00	24:00:00	00:00:00	263:56:16	-261:03:44		
	12	204:36:00	525:00:00	24:00:00	00:00:00	228:36:00	-296:24:00		
	15	249:07:59	525:00:00	63:00:00	00:00:00	312:07:59	-212:52:01		
	16	234:39:36	525:00:00	71:00:00	00:00:00	305:39:36	-219:20:24		
	17	214:08:14	525:00:00	71:00:00	00:00:00	285:08:14	-239:51:46		
	18	198:35:38	525:00:00	79:00:00	00:00:00	277:35:38	-247:24:22		
	19	193:18:14	525:00:00	79:00:00	00:00:00	272:18:14	-252:41:46		
	22	218:59:16	525:00:00	95:00:00	00:00:00	313:59:16	-211:00:44		
	23	190:17:33	525:00:00	87:00:00	00:00:00	277:17:33	-247:42:27		
	24	203:02:28	525:00:00	87:00:00	00:00:00	290:02:28	-234:57:32		
	25	194:09:27	525:00:00	87:00:00	00:00:00	281:09:27	-243:50:33		
	26	189:26:54	525:00:00	87:00:00	00:00:00	276:26:54	-248:33:06		
	29	206:13:14	525:00:00	55:00:00	00:00:00	261:13:14	-263:46:46		
	30	259:23:16	525:00:00	55:00:00	00:00:00	314:23:16	-210:36:44		
	31	196:40:15	525:00:00	39:00:00	00:00:00	235:40:15	-289:19:45		
	Total		3447:15:14	11550:00:00	1219:00:00	00:00:00	4666:15:14	-6883:44:46	

Figura 17: Indicador que representan el Horario Real Vs el Teórico por dependencia.

Lo mismo si queremos verlo a nivel **empleado**, en vez de poner como dimensión la **dependencia**. Ponemos el **empleado** y seleccionamos la dirección para corroborar que nos de la misma cantidad que en el indicador de arriba.

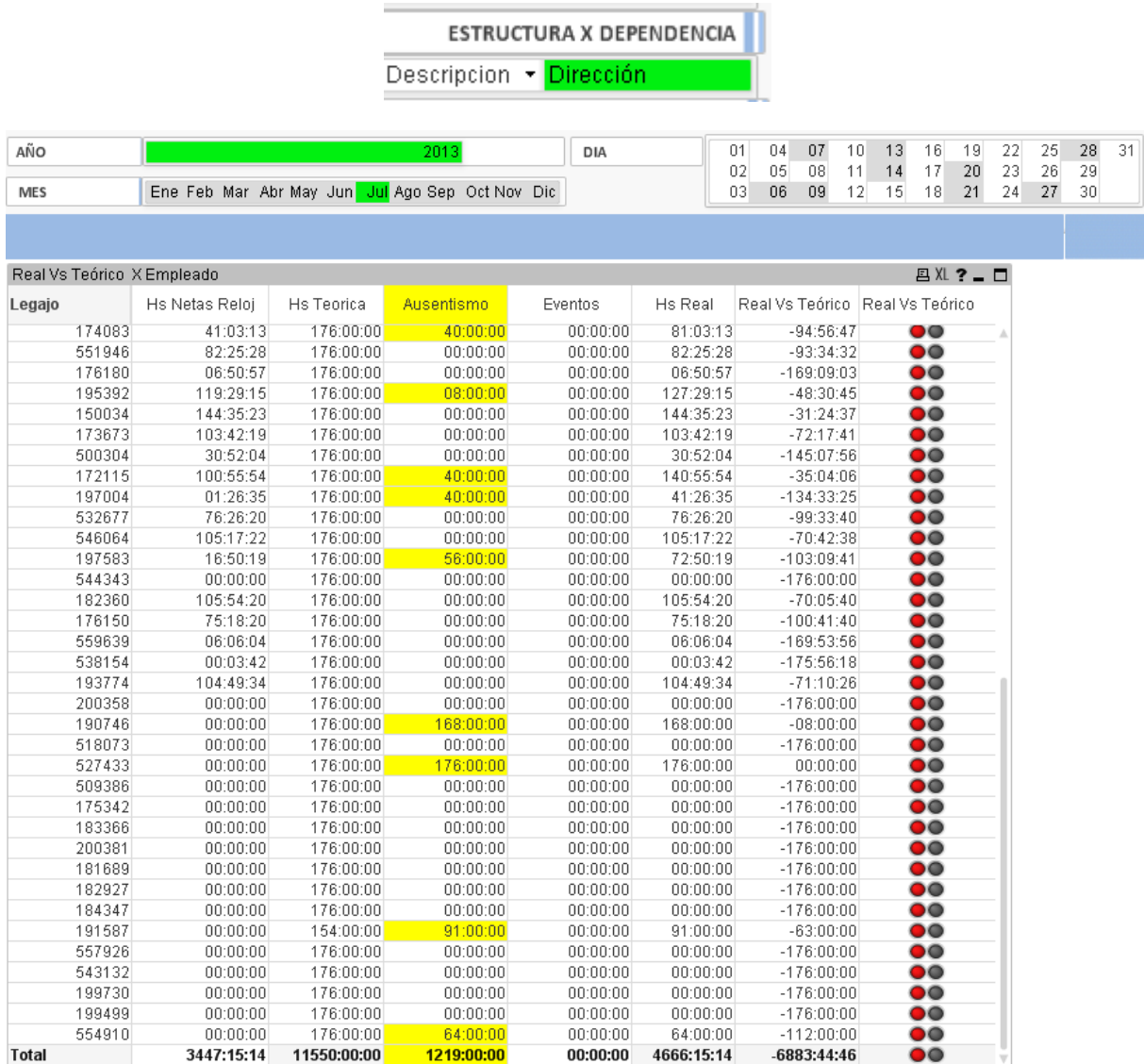


Figura 18: Indicador que representan el Horario Real Vs el Teórico por empleado.

d. Solución Técnica

Para poder calcular el siguiente indicador se utilizaron las siguientes expresiones:

- *Horario Neto*: Sum (Hs Netas) proveniente de la tabla Reloj.
- *Horario Teórico*: Sum (Hs Netas) proveniente de la tabla Evento_Teórico.
- *Ausentismo*: Sum (Hora_Ausente_x_Dia) proveniente de la tabla Evento_Ausentismo.
- *Evento*: Sum (Horas_Eventos) proveniente de la tabla eventos

- *Horario Real: $Sum(Hs\ Netas) + Sum(Hora_Ausente_x_Dia) + Sum(Horas_Eventos)$* proveniente de la información de los campos mencionados anteriormente.

Todas son válidas tanto para sacar el indicador por **día** como por **mes**. Lo mismo sucede si quiero verlo agrupado por **dirección** como si quiero verlo por los **empleados** que representan la **dirección**.

3.2.1.2 Medida: Evolución de Ventas en \$.

Es la evolución que tuvo la empresa por ventas de productos que se llevaron a cabo en moneda local.

a. Contexto:

En una empresa del rubro consumo masivo se desea saber cómo es la evolución de ventas en moneda local mediante las siguientes dimensiones de tiempo: año, mes, día y hora.

b. Modelo Asociativo:

Para realizar la Medida mencionada anteriormente, contamos con el siguiente modelo asociativo:

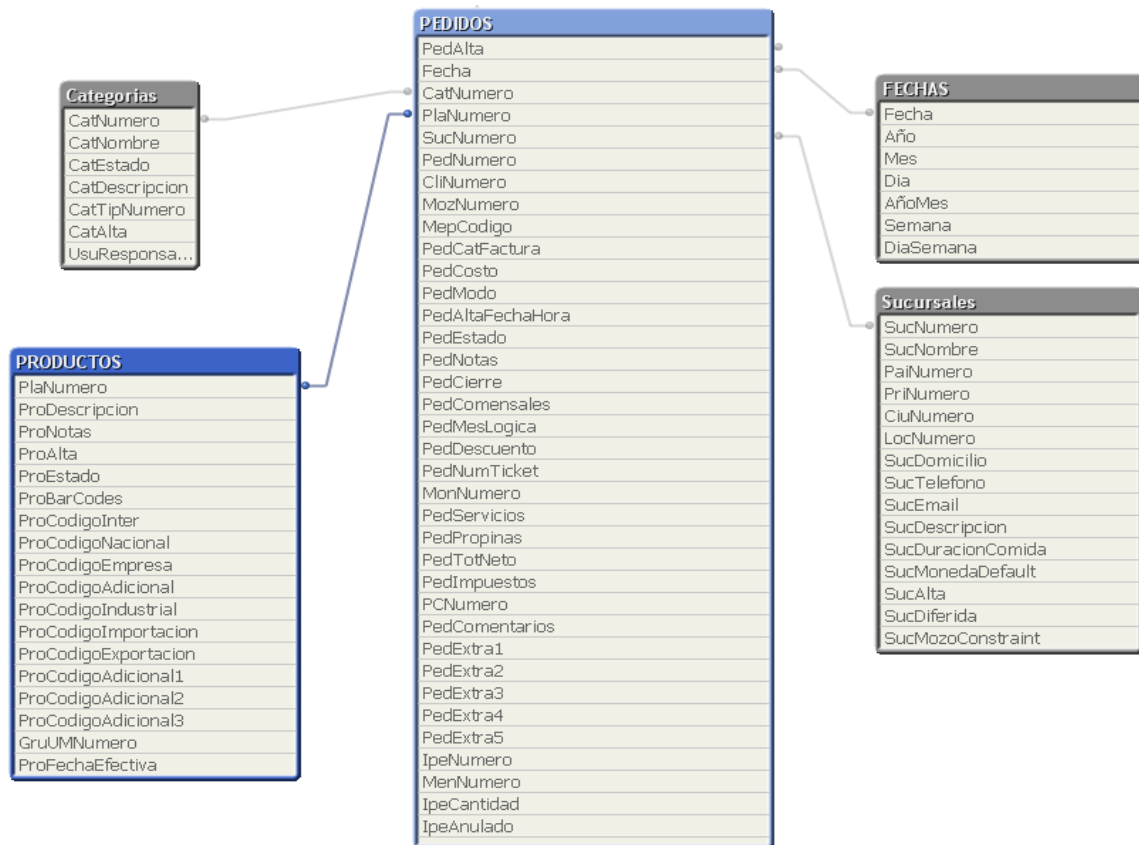


Figura 19: Modelo asociativo referente a las ventas en la compañía en QlikView.

El mismo está compuesto por las siguientes tablas:

- Pedidos: Representa los pedidos que tuvo la compañía.
- Fechas: Representa la fecha.
- Sucursales: Representa las sucursales que posee la compañía.
- Productos: Representa los productos que vende la compañía.
- Categorías: Representa la categoría de los pedidos.

c. Solución:

Se hará un indicador en donde se pueda ver por diferentes dimensiones de **tiempo** las **ventas** que tuvo la compañía en **pesos**.

Al ser una evolución de ventas se trata de una Medida puramente aditiva, siempre debemos sumarla independientemente de la dimensión que se utilice.

En la Figura 20 tenemos la evolución por **año**:

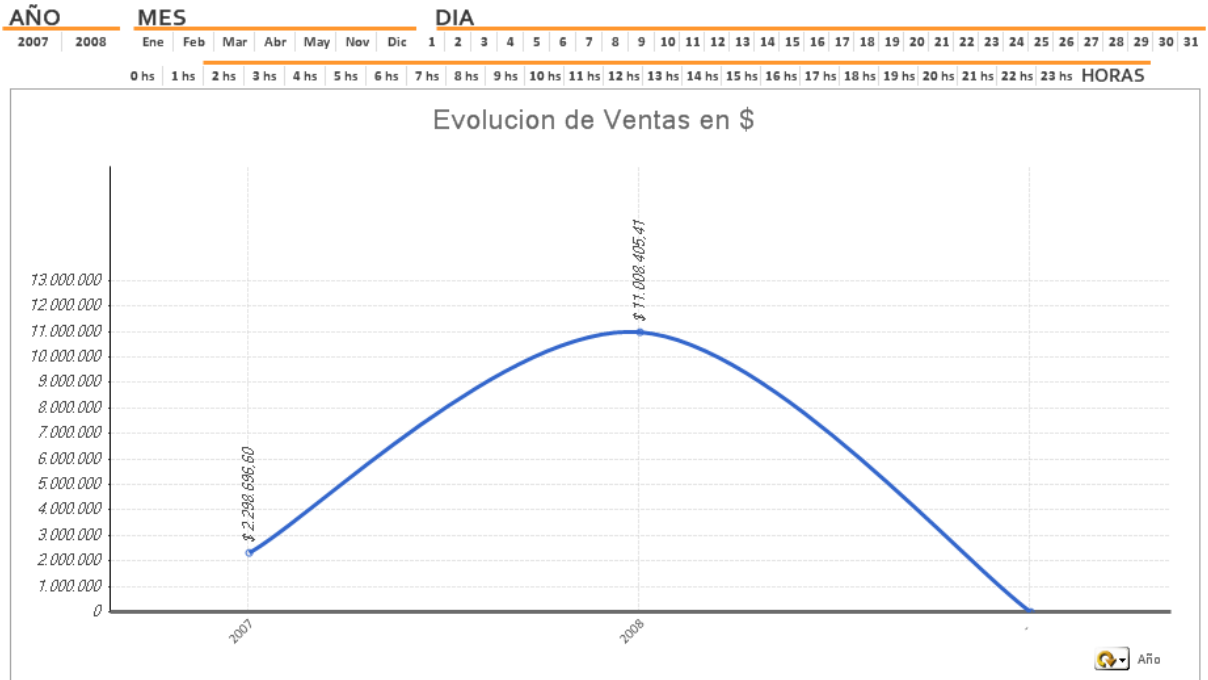


Figura 20: Evolución de ventas en pesos por año.

En la Figura 21 tenemos la evolución por mes:

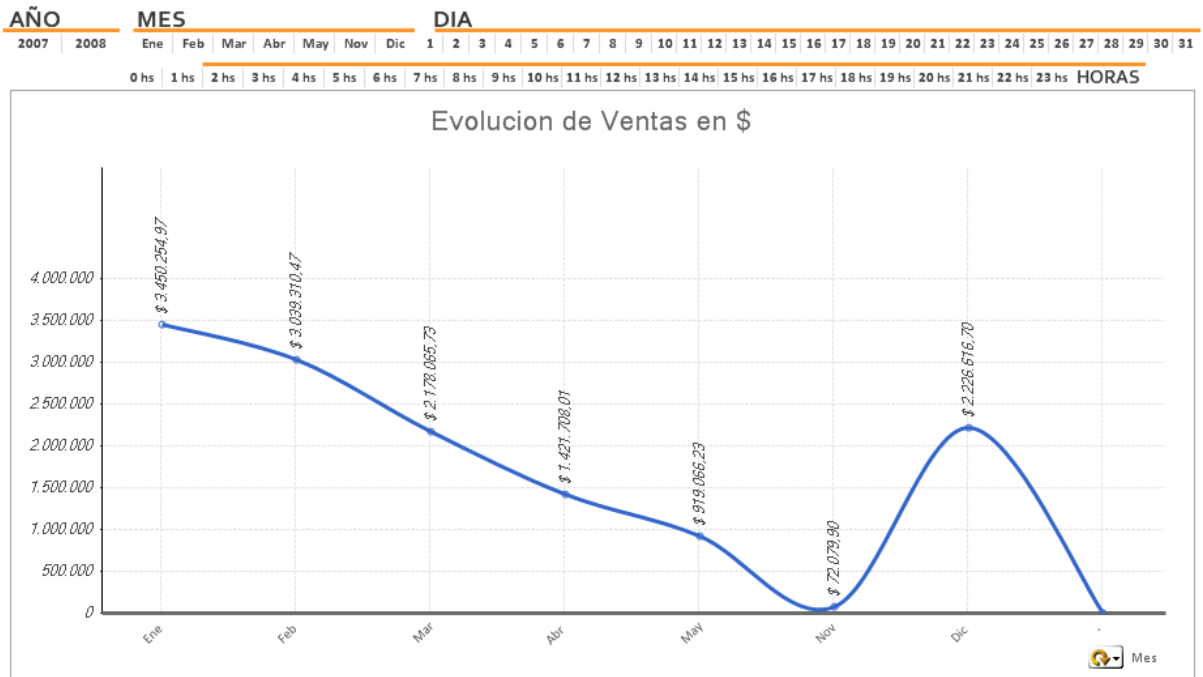


Figura 21: Evolución de ventas en pesos por mes.

En la Figura 22 tenemos la evolución por día:

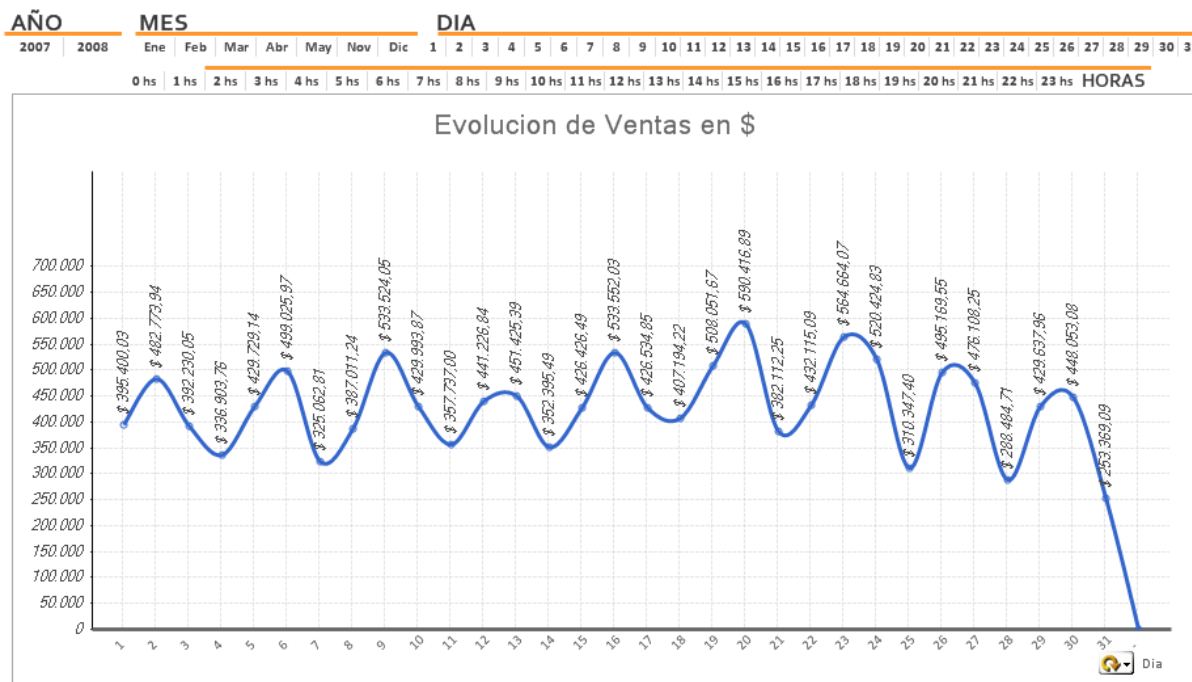


Figura 22: Evolución de ventas en pesos por día.

Y finalmente en la Figura 23 vemos la evolución por hora:

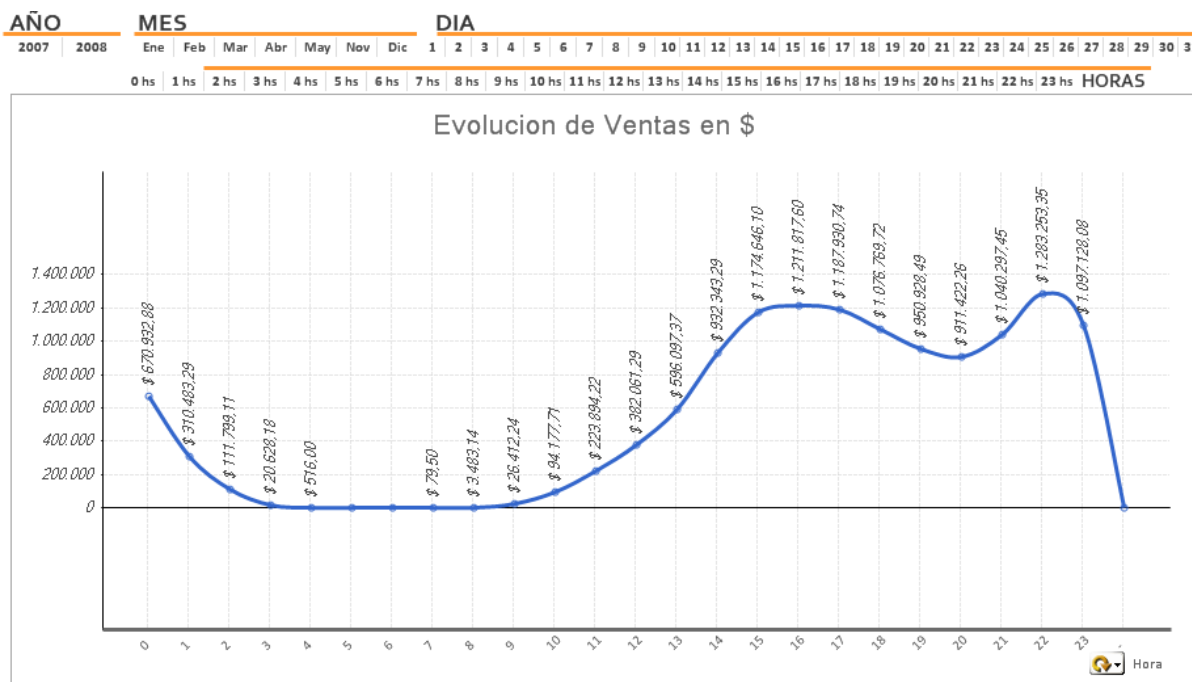


Figura 23: Evolución de ventas en pesos por hora.

Con los indicadores expuestos arriba se expone de forma clara, como mediante una misma expresión, una simple suma, puedo ver la evolución de ventas por diferentes dimensiones presentes en el modelo asociativo realizado para el ejemplo.

d. Solución Técnica

Para calcular la Medida se utilizó la siguiente fórmula:

- *Ventas*: sum (IpeCantidad* IpeCosto).

En este caso las ventas están calculadas por la suma de la multiplicación de los siguientes campos: **IpeCantidad**, representa la cantidad vendida e **IpeCosto**, representa el precio de venta.

3.2.1.3 Medida: Empleados que entraron por Año.

Es la cantidad de empleados que ingresaron a la compañía por año.

a. Contexto:

En una empresa del rubro logística se desea saber cuántos empleados fueron ingresando por año en el periodo del 1993 al 1997 a nivel empleado y a nivel oficina de venta.

Se deberá tomar en cuenta la fecha de ingreso que figura en el contrato acordado con cada empleado.

b. Modelo Asociativo:

Para realizar la Medida mencionada anteriormente, contamos con el siguiente modelo asociativo:



Figura 24: Modelo asociativo que representa la relación entre la tabla Oficinas y Empleados en QlikView.

El mismo está compuesto por las siguientes tablas:

- Oficinas: Representa las oficinas de ventas que tiene la empresa.
- Empleados: Representa los empleados que posee de la compañía.

c. Solución

Se hará un indicador en donde se pueda ver por diferentes dimensiones (**empleado** y **oficina**) las personas que ingresaron en la empresa por **año**. Para esto se cuenta con los siguientes datos provenientes del modelo asociativo:

Año Ingreso	IDEmpleado	Nombre	Año Ingreso
1993	2	Elvis Presley	1993
1994	19	David Letterman	1993
1995	22	Binh Protzmann	1994
1996	23	Cezar Sandu	1994
1997	43	Bob Park	1994
	44	Clark Kent	1994
	5	Ingrid Hendrix	1995
	9	Helen Brolin	1995
	24	Cindy Crawford	1995
	25	James Bond	1995
	30	Ira Baumel	1996
	33	Ann Lindquist	1996
	35	Charles Ingvar Jönsson	1996
	34	Bill Yang	1997
	36	Jacques Clouseau	1997

Figura 25: Representa los ingresos de empleados en la compañía por año.

Si bien para realizar el indicador vamos a utilizar como expresión un conteo de IDEmpleado, para sacar el totalizado por cualquier dimensión deberá hacerse una suma. Es decir que siempre debemos sumar la cantidad de legajos independientemente de la dimensión que se utilice, razón por la cual estamos en presencia de una Medida puramente aditiva.

Al hacer la Medida por año tomando en cuenta la dimensión **empleado**:

Año Ingreso	Empleados X Oficina							
	Nombre	Año Ingreso	1993	1994	1995	1996	1997	Total
1993	Ann Lindquist		-	-	-	1	-	1
1994	Bill Yang		-	-	-	-	1	1
1995	Binh Protzmann		-	1	-	-	-	1
1996	Bob Park		-	1	-	-	-	1
1997	Cezar Sandu		-	1	-	-	-	1
	Charles Ingvar...		-	-	-	1	-	1
	Cindy Crawford		-	-	1	-	-	1
	Clark Kent		-	1	-	-	-	1
	David Letterman		1	-	-	-	-	1
	Elvis Presley		1	-	-	-	-	1
	Helen Brolin		-	-	1	-	-	1
	Ingrid Hendrix		-	-	1	-	-	1
	Ira Baumel		-	-	-	1	-	1
	Jacques Clou...		-	-	-	-	1	1
	James Bond		-	-	1	-	-	1
	Total		2	4	4	3	2	15

Figura 26: Representa los ingresos de empleados en la compañía por año.

Al hacer la Medida por año tomando en cuenta la dimensión **oficina**:

Año Ingreso	Empleados X Oficina							
	Oficina	Año Ingreso	1993	1994	1995	1996	1997	Total
1993	Lund		1	3	2	-	-	6
1994	Paris		-	-	1	2	2	5
1995	Seattle		-	1	-	1	-	2
1996	Stockholm		1	-	1	-	-	2
1997	Total		2	4	4	3	2	15

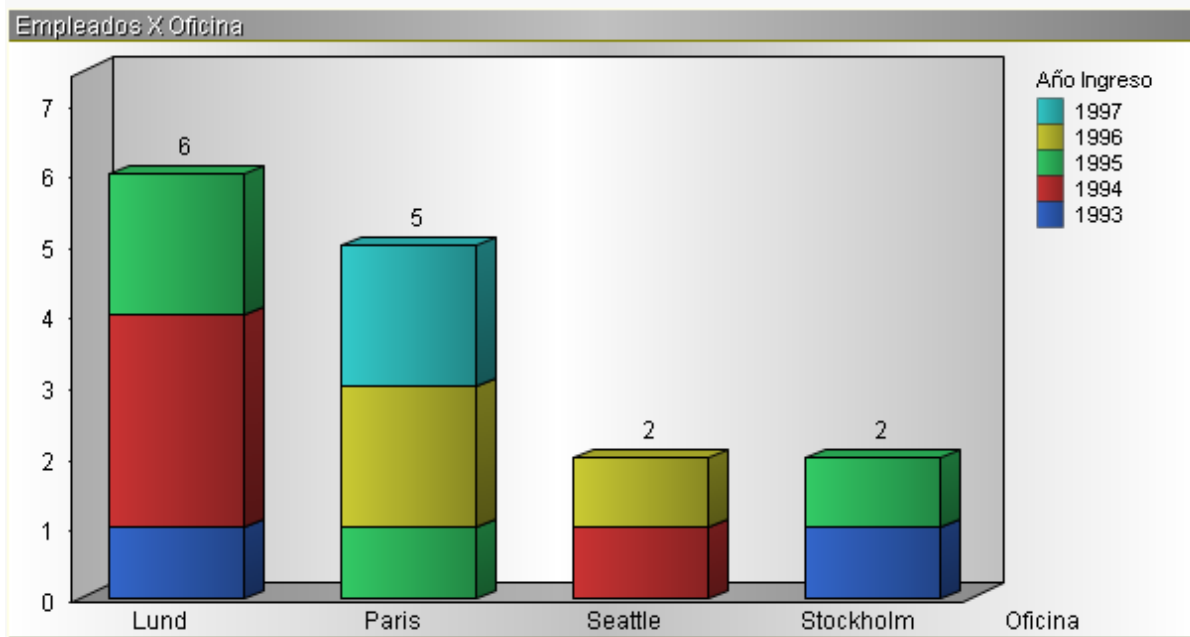


Figura 27: Representa los ingresos de empleados en la compañía por año en cada oficina de venta.

En los dos ejemplos mostrados anteriormente vemos como el resultado es similar para ambos casos.

d. Solución Técnica

Para calcular la medida se utilizó la siguiente fórmula:

- *Num. de Empleados:* Count (DISTINCT Nombre)

Si bien se utiliza la expresión “Count” para mostrar el total, se hace una suma de la cantidad de empleados.

3.2.1.4 Medida: Resultado Operativo en 2013

Es el beneficio que obtiene la empresa por el desarrollo de su actividad.

a. Contexto:

En una empresa de ventas de productos al por menor se necesita saber el resultado operativo que tuvo en el año fiscal 2013.

El resultado operativo es la utilidad neta antes del impuesto, es decir el beneficio que obtiene la empresa por el desarrollo de su actividad.

b. Modelo Asociativo:

Para realizar la medida mencionada anteriormente, contamos con el siguiente modelo asociativo:



Figura 27: Modelo asociativo que representa la relación entre la tabla Calendario y Balance en QlikView.

El mismo está compuesto por las siguientes tablas:

- Balance: Representa todos los movimientos que tuvo la empresa para todos sus años fiscales.
- Calendario: Representa la fecha desagregada por año fiscal, mes, día.

c. Solución:

Se cuenta con la siguiente información proveniente del modelo asociativo referido al balance con todos los movimientos que tuvo la empresa para el periodo fiscal 2013, de enero a diciembre.

Año Fiscal: 2013	
Balance General	
Movimientos	Valores
Ingresos por Ventas	90
Costo Directos de Producción	44
Contribución	46
Gastos Generales	24
Depreciación	6
RESULTADO OPERATIVO	18
Intereses	4
Resultado antes de Impuestos	14
Impuestos	3
RESULTADO NETO	11
Caja y Bancos	16
Cuentas por cobrar	20
Stock de Materiales	6
Productos en Proceso	18
Productos en Terminado	15
Activo Corriente	75
Equipos y Maquinaria	20
Inmuebles	5
Activo Fijo	25
TOTAL ACTIVO	100
A largo plazo	36
A corto plazo	24
Pasivo	60
Capital social	10
Resultado acumulado	30
Total Capital	40
TOTAL PASIVO Y CAPITAL	100

Figura 28: Balance respecto al año fiscal 2013.

En base a la información disponible en la Figura 28, se debe sacar el resultado operativo que tuvo la compañía para el año fiscal 2013, el cual ya lo tenemos calculado en el balance.

Al ser una medida aditiva debe sumarse siempre bajo cualquier circunstancia, al igual que si quiero saber el resultado operativo para dos años, debo hacer una suma entre los valores obtenidos de ambos años. Entonces para sacar el valor correcto del **resultado operativo** debemos hacer una suma entre la **contribución** (ingresos por ventas - costos directos de producción) menos los **costos comunes** (gastos generales + depreciación). En base al balance puesto a continuación el **resultado operativo** sería= $(90-44) - (24+6) = 16$.

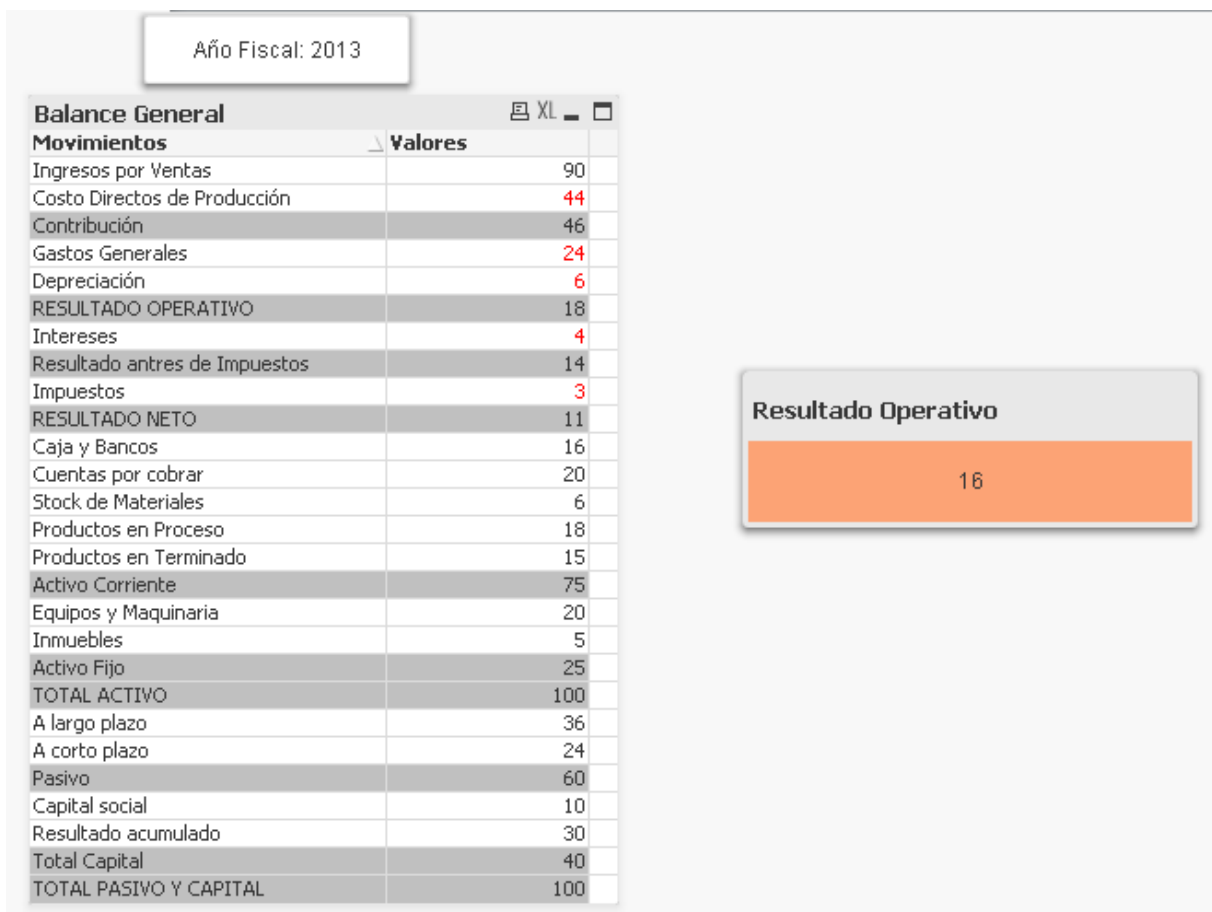


Figura 29: Resultado operativo según balance del año fiscal 2013.

d. Solución Técnica

Para calcular la medida se utilizó la siguiente fórmula:

- *Resultado Operativo:* $\text{sum}(\{\langle \text{Movimientos} = \{ \text{'Resultado Operativo'} \} \rangle \text{Valores}\}) / (\text{sum}(\{\langle \text{Movimientos} = \{ \text{'Total Activo'} \} \rangle \text{Valores}\}))$.

Se hace una suma de los movimientos que pertenezcan al **resultado operativo** dividido la suma de los movimientos que representan **el total de los activos**.

3.2.3. Semi-Aditivas

3.2.2.1 Medida: Stock de computadoras en depósito.

Computadoras que se encuentran al momento en el depósito de la empresa, producto de los ingresos y egresos que se generan en el mismo.

a. Contexto:

Empresa de tecnología informática desea saber el stock de computadoras por día y mes.

b. Modelo Asociativo:

Para realizar la medida mencionada anteriormente, contamos con el siguiente modelo asociativo:

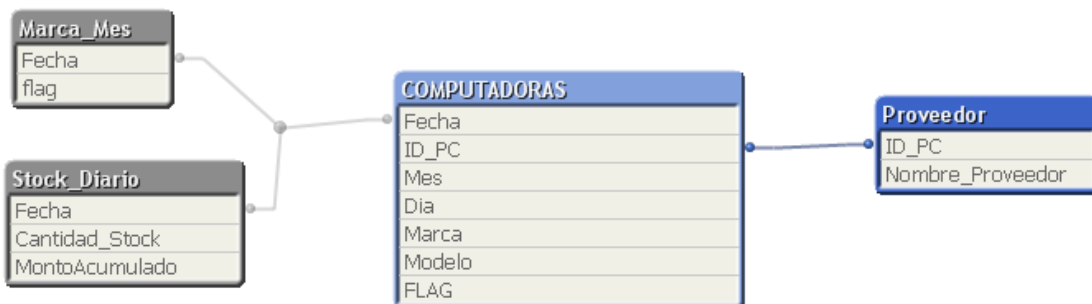


Figura 30: Modelo asociativo referente al stock de computadores.

El mismo está compuesto por las siguientes tablas:

- Proveedor: Contiene el nombre del Proveedor y el ID de la PC.
- COMPUTADORAS: Tabla representa las computadoras que ingresan a depósito.
- Stock_Diario: Representa el stock de computadoras en depósito.
- Marca_Mes: Permite saber cuál es el último día hábil por mes.

c. Solución:

Al hablar de stock, estamos frente a una medida semi-aditiva, no se debe hacer el mismo cálculo para sacar el indicador por día que por mes.

Para demostrar esto, se hará una tabla en donde se pueda ver el stock de computadoras en depósito por mes y otra por día. Es decir si uno quiere ver el indicador por mes no se debe pensar igual que si lo quiere ver por día.

Supongamos que la información que tenemos para el mes de julio y agosto es la siguiente:

- Computadoras que ingresaron y egresaron de depósito por día en julio:

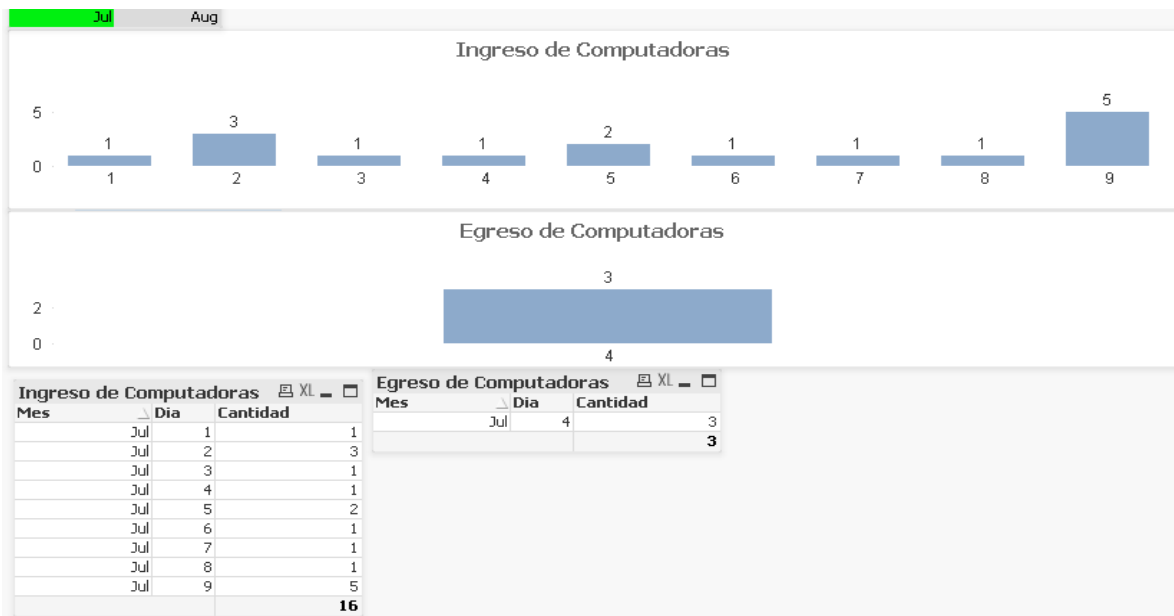


Figura 31: Computadores que ingresaron/egresaron a depósito en el mes de julio.

- Computadoras que ingresaron y egresaron de depósito por día en agosto:

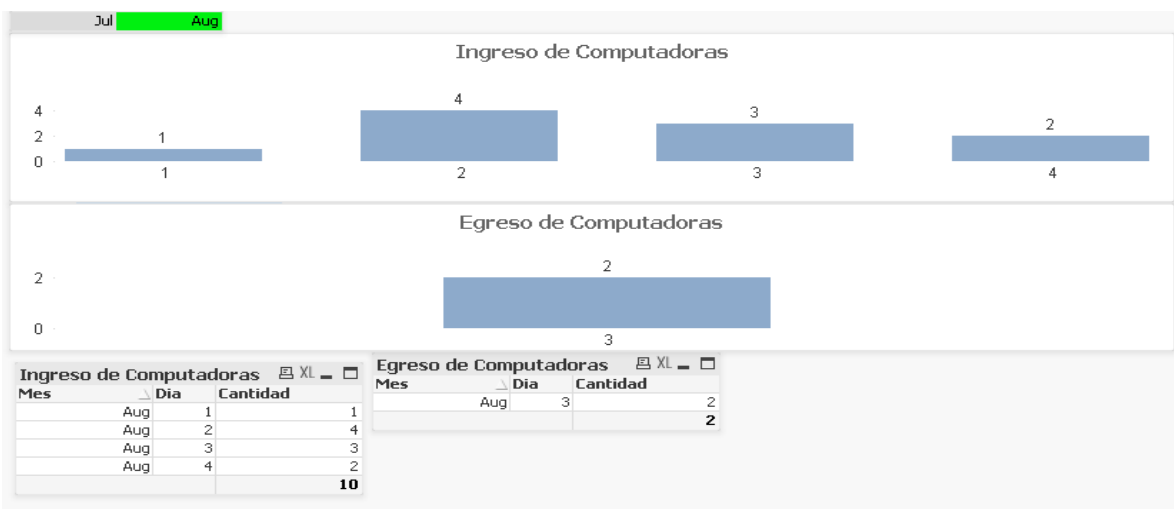
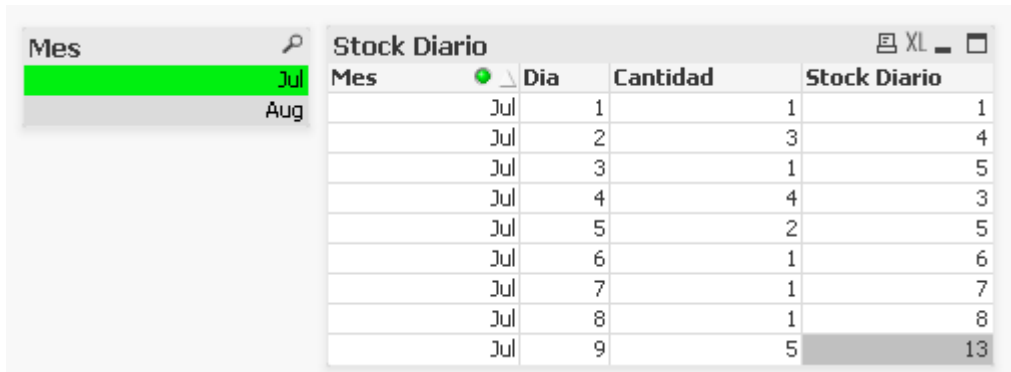


Figura 32: Computadores que ingresaron/egresaron a depósito en el mes de agosto.

Si queremos ver el stock de computadoras por día para un mes, debemos ir acumulando la cantidad de computadoras que tengo por día más las que quedaron en depósito del mes anterior:



Mes	Dia	Cantidad	Stock Diario
Jul	1	1	1
Jul	2	3	4
Jul	3	1	5
Jul	4	4	3
Jul	5	2	5
Jul	6	1	6
Jul	7	1	7
Jul	8	1	8
Jul	9	5	13

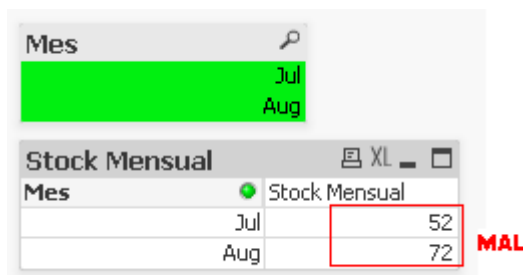
Figura 33: Stock de computadores para el mes de julio abierto por día.



Mes	Dia	Cantidad	Stock Diario
Aug	1	1	14
Aug	2	4	18
Aug	3	5	19
Aug	4	2	21

Figura 34: Stock de computadores para el mes de agosto abierto por día.

Pero si queremos ver el stock de computadoras por mes usando la fórmula del indicador anterior nos da información de forma errónea:



Mes	Stock Mensual
Jul	52
Aug	72

MAL

Figura 35: Stock erróneo de computadores para los meses de julio y agosto.

Esto se debe a que se está sumando el stock acumulado por día y no se tiene en cuenta solamente la cantidad de computadoras que tengo en depósito para el último día del mes.

Por ejemplo:

En julio nos da 52 computadoras, dado a que suma todas las filas en vez de tomar la información del 9 de Julio, la cual es 13 computadoras:

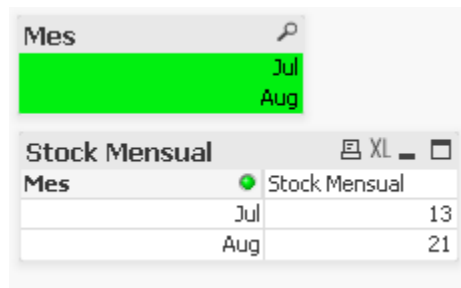


Mes	Dia	Cantidad	Stock Diario
Jul	1	1	1
Jul	2	3	4
Jul	3	1	5
Jul	4	4	3
Jul	5	2	5
Jul	6	1	6
Jul	7	1	7
Jul	8	1	8
Jul	9	5	13
			52

Figura 36: Stock erróneo de computadores a nivel mensual abierto por los días del mes.

Por lo tanto debemos tomar solamente la cantidad de computadoras que se encuentran en depósito para el último día hábil del mes.

Usando la lógica mencionada anteriormente vemos como el indicador da la información diaria correctamente tanto para el mes de julio como agosto:



Mes	Stock Mensual
Jul	13
Aug	21

Figura 37: Stock de computadores a nivel mensual.

Esto es útil para cualquier medida relacionada a stocks de cualquier índole.

d. Solución Técnica :

Antes de ponernos a desarrollar los indicadores, debemos calcular el valor del stock en base a los ingresos y egresos de computadoras que hubo en depósito para tenerlo disponible en el modelo asociativo, el cual no se puede sacar automáticamente en base a una expresión del modelo.

```
//1. Levanto la tabla COMPUTADORAS.
COMPUTADORAS:
LOAD Fecha_Alta as Fecha,
    Month(Fecha_Alta) as Mes,
    Day(Fecha_Alta) as Dia,
    [ID PC] AS ID_PC,
    Marca,
    Modelo,
    'ALTA/BAJA' as FLAG
FROM
[Pc_altas_Bajas.xls]
(biff, embedded labels, table is Hojal$);

//2. Voy agrupando por fecha y FLAG para saber si es un ingreso o egreso de
computadoras.
PC_tmp2:
NoConcatenate
load Fecha, count( distinct ID_PC)*if(FLAG='BAJA',-1,1) as ALTA_DIARIA,FLAG
Resident COMPUTADORAS Group by Fecha,FLAG;

//3. Averiguo el valor neto de la cantidad de computadoras en depósito.
PC_tmp:
NoConcatenate
load Fecha, sum(ALTA_DIARIA) as ALTA_DIARIA Resident PC_tmp2 Group By Fecha; DROP
Table PC_tmp2;

//4. Calculo el stock de las computadoras.
Stock_Diario:
NoConcatenate
load ALTA_DIARIA as Cantidad_Stock,
Fecha,
//Si la fecha anterior (Ej: 01/07/2013) a la que estoy parado (Ej: 02/07/2013) no
es nula, es decir que hay datos para esa fecha, debemos sumar las computadoras que
tengo en el stock junto a las que ingresaron esa fecha.
Caso contrario tomar la cantidad de computadoras que tengo en esa fecha.
if (not isnull(previous(Fecha)),
    ALTA_DIARIA+peek(MontoAcumulado),ALTA_DIARIA) as MontoAcumulado

Resident PC_tmp Order by Fecha ASC;
DROP Table PC_tmp;

//5. Creo un marcador para poder sacar el stock a nivel mensual.
Marca_Mes:
load max(Fecha) as Fecha, 1 as flag Resident Stock_Diario Group by month(Fecha);
```

Script 5: Script para sacar el stock.

- El campo **MontoAcumulado** representa las computadoras acumuladas por fecha.
- El campo **flag** representa el último día hábil del mes, solo se usa para ver por mes el stock.
- Luego vemos como cambian las fórmulas que se deben usar en el indicador por mes, que en el indicador por día.
- *Fórmula para sacar el stock mensual:* sum ({<flag= {1}>} MontoAcumulado).
- *Fórmula para sacar el stock diario:* sum (MontoAcumulado)

3.2.2.2 Medida: Dotación de empleados activa.

Es la cantidad de empleados activos que tiene la empresa.

a. Contexto:

En una empresa de logística se necesita saber la cantidad de empleados que se encuentran trabajando por mes y por sector.

b. Modelo Asociativo:

Para realizar la medida mencionada anteriormente, contamos con el siguiente modelo asociativo.



Figura 38: Modelo asociativo referente a la dotación activa de empleados.

El mismo está compuesto por las siguientes tablas:

- Sector: Contiene la descripción y gerente responsable del sector.
- Dotación_Activa: Contiene la información de los empleados activos en la compañía.

c. Solución:

Al hablar de dotación, estamos frente a una medida semi-aditiva, no se debe hacer el mismo cálculo si queremos ver la medida por una dimensión que si la queremos ver por otra.

Para demostrar esto, se hará una tabla en donde se pueda ver la cantidad de empleados perteneciente al sector de Business Intelligence totalizados a nivel mensual y otra a la fecha actual, es decir a la fecha en la cual estoy parado.

Se cuenta con la siguiente información proveniente del modelo asociativo:

Mes		Empleados Activos X Mes	
Mes	LEGAJO	Nombre_Completo	
Ene	510998	Geraldine Alexis Girbau	
Ene	521204	Inti Camillato	
Ene	538983	SANTIAGO GARCIA MILAN	
Ene	539837	JORGE RAUL RANNI	
Ene	993081	Ilse Fabiola Velazquez Vargas	
Feb	191642	ROBERTO DAMIAN YEBARA	
Feb	195939	MARIA MAS FEIJOO	
Feb	500103	LUCIANA FIORELLA ALUIGI	
Feb	510998	Geraldine Alexis Girbau	
Feb	521204	Inti Camillato	
Feb	522470	MARIA JOSE VELAZQUEZ	
Feb	538983	SANTIAGO GARCIA MILAN	
Feb	539590	EDGARDO DARIO IVAN SALBA	
Feb	539837	JORGE RAUL RANNI	
Feb	993081	Ilse Fabiola Velazquez Vargas	
Mar	191642	ROBERTO DAMIAN YEBARA	
Mar	195939	MARIA MAS FEIJOO	
Mar	500103	LUCIANA FIORELLA ALUIGI	
Mar	510998	Geraldine Alexis Girbau	
Mar	521204	Inti Camillato	
Mar	522470	MARIA JOSE VELAZQUEZ	
Mar	538983	SANTIAGO GARCIA MILAN	
Mar	539590	EDGARDO DARIO IVAN SALBA	
Mar	539837	JORGE RAUL RANNI	
Mar	544374	ESTEBAN EZEQUIEL BESSIERE	
Mar	544528	EDUARDO ANDRES LORENZO	
Mar	544771	EZEQUIEL BONIN	
Mar	545051	PABLO EDUARDO CANSECO DIAZ	
Mar	545062	JOSE MARTIN FERREYRA	
Mar	993081	Ilse Fabiola Velazquez Vargas	
Abr	191642	ROBERTO DAMIAN YEBARA	
Abr	195939	MARIA MAS FEIJOO	
Abr	500103	LUCIANA FIORELLA ALUIGI	
Abr	510998	Geraldine Alexis Girbau	
Abr	521204	Inti Camillato	
Abr	522470	MARIA JOSE VELAZQUEZ	
Abr	538983	SANTIAGO GARCIA MILAN	
Abr	539590	EDGARDO DARIO IVAN SALBA	
Abr	539837	JORGE RAUL RANNI	
Abr	544374	ESTEBAN EZEQUIEL BESSIERE	
Abr	544528	EDUARDO ANDRES LORENZO	
Abr	544771	EZEQUIEL BONIN	
Abr	545051	PABLO EDUARDO CANSECO DIAZ	
Abr	545062	JOSE MARTIN FERREYRA	
Abr	993081	Ilse Fabiola Velazquez Vargas	


Figura 39: Empleados activos para los meses de enero, febrero, marzo y abril.

En la Figura 40 vemos la cantidad de empleados activos que tiene el sector de Business Intelligence por mes, para este caso no es más que un simple conteo de personas activas:

Mes		Dotación de Empleados Activa X Mes			
Sector	Mes	Ene	Feb	Mar	Abr
Business Intelligence		5	10	20	15
Total		5	10	20	15

Figura 40: Cantidad de empleados activos abierto por mes.

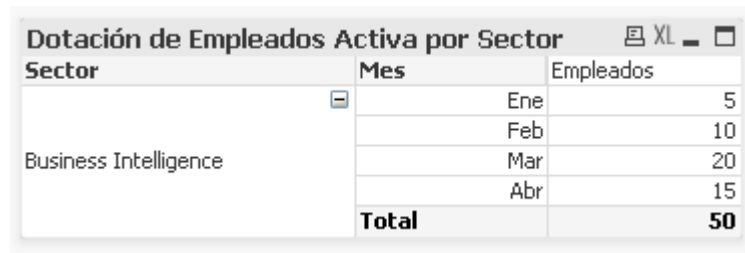
En la Figura 41 vemos la cantidad de empleados activos que tiene el sector de Business Intelligence a la fecha actual (abril), si acá utilizamos la misma fórmula que la de arriba estamos frente a un problema:



Sector	Empleados
Business Intelligence	50 MAL

Figura 41: Cantidad errónea de empleados activos para el sector Business Intelligence en enero, febrero, marzo y abril.

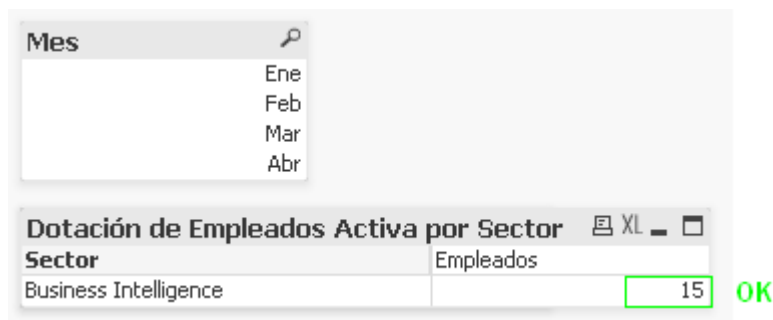
Si analizamos el resultado nos damos cuenta que es erróneo, ya que al cambiar la dimensión que queremos totalizar falla la expresión por sumar la cantidad de empleados activos que tuvo el sector por mes:



Sector	Mes	Empleados
Business Intelligence	Ene	5
	Feb	10
	Mar	20
	Abr	15
	Total	50

Figura 42: Cantidad errónea de empleados activos para el sector Business Intelligence abierto por los meses de enero, febrero, marzo y abril.

Para evitar este inconveniente debemos utilizar la siguiente lógica: tomar el último valor del mes que estamos analizando y mostrar el mismo en el total de la siguiente forma:



Sector	Empleados
Business Intelligence	15 OK

Figura 43: Cantidad de empleados activos para el sector Business Intelligence en enero, febrero, marzo y abril.

Esto es útil para cualquier medida relacionada a dotaciones de cualquier índole.

d. Solución Técnica :

- *Formula Dotación Activa por Mes:* Count (LEGAJO).
- *Formula Dotación Activa a Fecha Actual:* Count (Distinct (LEGAJO)).

Esta última fórmula también se puede obtener de tomar el último registro del mes que estoy analizando y poner el mismo valor en el total, de la siguiente forma:

- *Formula Dotación Activa a Fecha Actual:*
if(Dimensionality()=1,count(if(num(Mes)=\$(Variable1)',LEGAJO)),count(LEGAJO))

En donde:

- El campo \$(Variable1)= max (Mes) representa el último mes.
- Dimensionality ()=1 representa el total de la tabla.

3.2.2.3 Medida: Clientes en mora por mes/ trimestre

Es la cantidad de clientes que tiene la compañía en mora tanto por mes como trimestre.

a. Contexto:

En una empresa financiera necesita saber la cantidad de clientes que se encuentran en moratoria.

b. Modelo Asociativo:

Para realizar la medida mencionada anteriormente, contamos con el siguiente modelo asociativo.

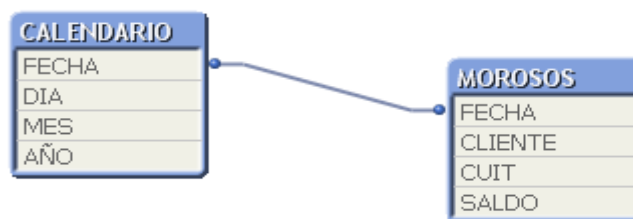


Figura 44: Modelo asociativo referente a clientes morosos.

El mismo está compuesto por las siguientes tablas:

- Morosos: Contiene el listado de los deudores morosos identificados por nombre y CUIT.
- Calendario: Representa la fecha desagregada por año, mes, día.

c. Solución:

Al hablar de empleados en moratoria, estamos frente a una medida semi-aditiva, no se debe hacer el mismo cálculo si queremos ver la medida por una dimensión que si la queremos ver por otra.

Para demostrar esto, se hará una tabla en donde se pueda ver la cantidad de empleados totalizados a nivel mensual y trimestral.

En la figura 45 se cuenta con la siguiente información proveniente del modelo asociativo:

CANTIDAD MOROSOS			
FECHA	CLIENTE	CUIT	CANTIDAD MOROSOS
ENE	Aluminio Garci...	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S....	30-12455678-8	1
FEB	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S....	30-12455678-8	1
MAR	Aluminio Garci...	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S....	30-12455678-8	1

Figura 45: Clientes en mora del trimestre.

En la Figura 46 vemos la cantidad de clientes en mora abierto por mes:

CANTIDAD MOROSOS			
MES	CLIENTE	CUIT	CANTIDAD MOROSOS
ENE	Aluminio Garcia S.A.	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		3
FEB	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		2
MAR	Aluminio Garcia S.A.	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		3

Figura 46: Clientes en mora del trimestre totalizado por mes.

En la Figura 47 vemos la cantidad de clientes en mora abierto por trimestre, si acá utilizamos la misma fórmula que la de arriba estamos frente a un problema, ya que el total nos va a dar erróneo, dado a que cuenta más de un empleado por mes en caso de que siga en mora:

CANTIDAD MOROSOS			
MES	CLIENTE	CUIT	CANTIDAD MOROSOS
ENE	Aluminio Garcia S.A.	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		3
FEB	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		2
MAR	Aluminio Garcia S.A.	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		3
Total			8 MAL

Figura 47: Cantidad errónea de clientes en el primer trimestre.

Para evitar este inconveniente debemos utilizar la siguiente lógica: contar la cantidad de clientes que se encuentran en mora siempre y cuando no se repitan en los meses del trimestre, si sucede esto, debe contarse una única vez como muestra la Figura 48:

CANTIDAD MOROSOS			
MES	CLIENTE	CUIT	CANTIDAD MOROSOS
ENE	Aluminio Garcia S.A.	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		3
FEB	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		2
MAR	Aluminio Garcia S.A.	30-45687884-3	1
	Colina Sur S.A.	30-15235789-9	1
	EuroViajes S.R.L.	30-12455678-8	1
	Total		3
Total			3 OK

Figura 48: Cantidad de clientes en el primer trimestre.

d. Solución Técnica :

- *Formula Mensual:* if(sum(SALDO)>0,count(CLIENTE))
- *Formula Trimestral:* if(sum(SALDO)>0,count(Distinct(CLIENTE)))

En ambas formulas se encuentra la expresión: if (sum (SALDO)>0, la cual permite saber si el cliente tiene saldo deudor, caso de ser así contarlos. El Distinct nos permite evitar contar repetidos.

3.2.4. No Aditivas

3.2.3.1 Medida: Cotización promedio del dólar por sucursal.

Es el valor promedio en el que se venden dólares a los clientes por día.

a. Contexto:

Empresa financiera quiere saber la cotización del dólar promedio que tiene en sus tres sucursales para cada día determinado según las operaciones realizadas en las mismas.

- Sucursales:
 1. José C. Paz
 2. Pilar
 3. San Miguel

b. Modelo Asociativo:

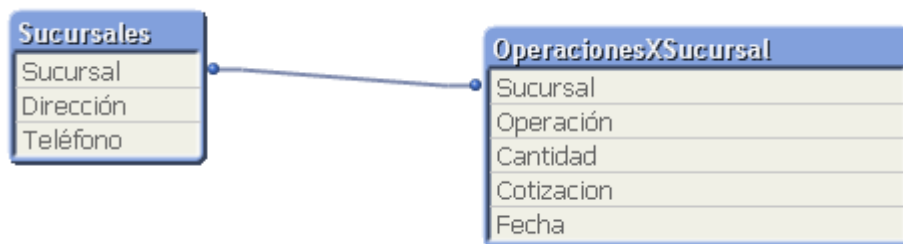


Figura 49: Modelo asociativo referente a la cotización del dólar por sucursal.

El mismo está compuesto por las siguientes tablas:

- Sucursales: Información de las sucursales de la empresa.
- OperacionesXSucursal: Representa las operaciones efectuadas por cada sucursal

c. Solución:

Para realizar la medida se cuenta con la siguiente información proveniente de tres sucursales de la compañía para un día determinado:

Operaciones X Sucursal			
Sucursal	Operación	Monto	Cotización Dolar
Jose C Paz	1	100	5,8
Jose C Paz	2	10000	5,6
Jose C Paz	3	200	6
Jose C Paz	4	100000	5,5
Pilar	1	100	6
Pilar	2	200	6
Pilar	3	300	6
Pilar	4	500	6
Pilar	5	900	6
San Miguel	1	500	6
San Miguel	2	100	6
San Miguel	3	1000	5,7

Figura 50: Operaciones realizadas por sucursal en un día determinado con su respectiva cotización.

Monto Cotizado					
Sucursal	Operación	Monto a Cotizar	Cotizacion	Monto Cotizado	
Jose C Paz	1	100	5,8	580	
Jose C Paz	2	10000	5,6	56000	
Jose C Paz	3	200	6	1200	
Jose C Paz	4	100000	5,5	550000	
Pilar	1	100	6	600	
Pilar	2	200	6	1200	
Pilar	3	300	6	1800	
Pilar	4	500	6	3000	
Pilar	5	900	6	5400	
San Miguel	1	500	6	3000	
San Miguel	2	100	6	600	
San Miguel	3	1000	5,7	5700	

Figura 51: Operaciones realizadas por sucursal en un día determinado con su respectiva cotización y monto cotizado.

Si queremos sacar la **cotización promedio** por sucursal la forma habitual de hacerlo es **sumar** las cotizaciones que tuvo la sucursal por operación y **dividirla** por la cantidad de las mismas, un simple cálculo del promedio.

Por ejemplo:

- Sucursal: José C. Paz
- Operaciones: 4
- Suma de todas las Cotizaciones: 22.9
- Cotización: 5.72

Es decir que la cotización promedio por sucursal es la siguiente:

Cotizacion Promedio	
Sucursal	Cotizacion Promedio
Jose C Paz	5,725
Pilar	6,000
San Miguel	5,900

Figura 52: Cotización Promedio por Sucursal.

Si vamos más al detalle, nos damos cuenta que el valor **cotización promedio** no es representativo para cada sucursal, dado a que hay operaciones de mayores volúmenes que se hacen a una cotización diferente a lo tomado como promedio, generando un pequeño desvío.

Para estos casos es recomendable sacar un **promedio ponderado** de las cotizaciones hechas por cada sucursal. Es decir debemos sumar los montos cotizados por sucursal y dividirlos por la suma de los montos de todas las operaciones hechas por cada sucursal.

Por ejemplo:

- Sucursal: José C. Paz
- Operaciones: 4
- Suma de montos cotizados: 607.780
- Suma de montos de todas las operaciones: 110.300
- Cotización: 5.51

Entonces la **cotización promedio ponderado** por sucursal es la siguiente:

Cotizacion Promedio Ponderado	
Sucursal	Cotizacion Promedio Ponderado
Jose C Paz	5,510
Pilar	6,000
San Miguel	5,813

Figura 53: Cotización promedio ponderado por sucursal.

Sale de la siguiente información:

Cantidad/ Cotización			
Sucursal	Monto a Cotizar	Monto Cotizado	Cotizacion Promedio Ponderado
Jose C Paz	110300	607780	5,51
Pilar	2000	12000	6,00
San Miguel	1600	9300	5,81

Figura 54: Cotización promedio ponderado, monto a cotizar y monto cotizado por sucursal.

Esto es más real a lo mostrando en el primer indicador.

d. Solución Técnica :

- *Formula Promedio:* $\text{sum (Cotización)} / \text{Count (Cotización)}$.
- *Formula Promedio Ponderado:* $\text{sum (Monto_A_Cotizar * Cotización)} / \text{sum (Monto_A_Cotizar)}$.

3.2.3.2 Medida: Porcentaje de participación en el mercado por sucursal a nivel trimestral.

Es el porcentaje que representa la participación que tienen las sucursales de la compañía en base a la cantidad de kilo gramos vendidos en el mercado en el trimestre.

a. Contexto:

En una empresa de venta de helados al público se necesita saber el porcentaje de participación en las ventas que tienen sus sucursales en el mercado a nivel trimestral.

- Trimestre: Primer trimestre del año 2013.
- Sucursales:
 1. José C. Paz
 2. Pilar
 3. San Miguel

b. Modelo Asociativo:

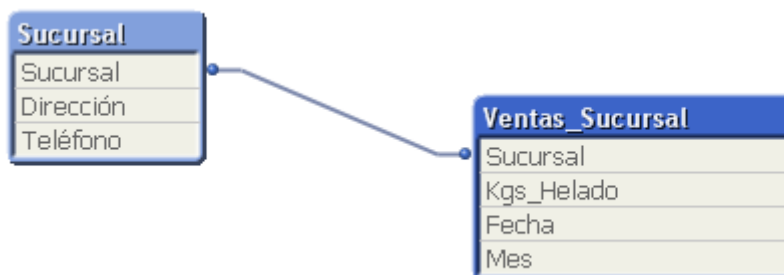


Figura 55: Modelo asociativo referente a la participación de las sucursales en referencia a las ventas de la compañía.

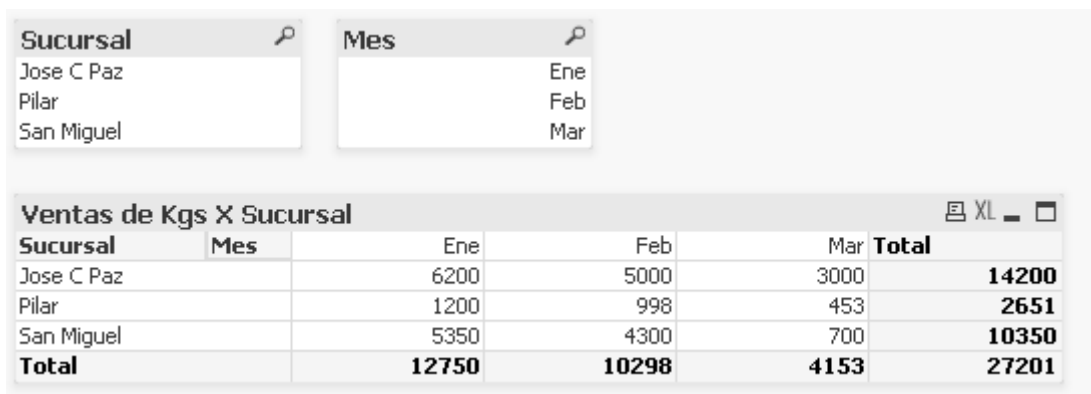
El mismo está compuesto por las siguientes tablas:

- Sucursales: Información de las sucursales de la empresa.
- Ventas_Sucursal: Representa las ventas efectuadas por cada sucursal a nivel kilo gramos.

c. Solución:

El porcentaje de participación representa una medida no aditiva, ya que bajo ninguna circunstancia se puede sumar directamente mes a mes por ninguna dimensión del modelo para darnos el porcentaje trimestral.

En base al modelo asociativo expuesto en la Figura 54 se cuenta con los siguientes datos, los cuales representan los kilos gramos vendidos por sucursal en los meses de enero, febrero y marzo del 2013, es decir el primer trimestre del año:



The screenshot shows a BI tool interface with two filter boxes at the top. The 'Sucursal' filter is set to 'Jose C Paz', 'Pilar', and 'San Miguel'. The 'Mes' filter is set to 'Ene', 'Feb', and 'Mar'. Below the filters is a pivot table titled 'Ventas de Kgs X Sucursal' with the following data:

Sucursal	Mes	Ene	Feb	Mar	Total
Jose C Paz		6200	5000	3000	14200
Pilar		1200	998	453	2651
San Miguel		5350	4300	700	10350
Total		12750	10298	4153	27201

Figura 56: Kilo Gramos vendidos por sucursal.

Si queremos ver el porcentaje de participación por sucursal a nivel trimestral, debemos primero calcular el mensual sumando los kilo gramos vendidos por esa sucursal para el mes dividido los kilo gramos vendidos por las tres sucursales en el mes, esto lo vemos en la figura de abajo.

Sucursal: Jose C Paz, Pilar, San Miguel
Mes: Ene, Feb, Mar

TOTAL ENERO J. C. PAZ

Ventas de Kgs X Sucursal

Sucursal	Mes	Ene	Feb	Mar	Total
Jose C Paz		6200	5000	3000	14200
Pilar		1200	998	453	2651
San Miguel		5350	4300	700	10350
Total		12750	10298	4153	27201

TOTAL ENERO 3 SUCURSALES

Figura 57: Kilo Gramos vendidos por sucursal a nivel mensual.

Al hacer lo mismo para las tres sucursales obtenemos el porcentaje de participación de cada sucursal para los tres meses que representan el trimestre:

Sucursal: Jose C Paz, Pilar, San Miguel
Mes: Ene, Feb, Mar

Participación Trimestral en el Mercado

Mes	Ene		Feb		Mar	
Sucursal	Kgs Vendidos	%	Kgs Vendidos	%	Kgs Vendidos	%
Jose C Paz	6200	48,63%	5000	48,55%	3000	72,24%
Pilar	1200	9,41%	998	9,69%	453	10,91%
San Miguel	5350	41,96%	4300	41,76%	700	16,86%

Figura 58: Porcentaje de participación de cada sucursal para cada mes del trimestre.

Finalmente debemos sacar el porcentaje de participación trimestral para cada sucursal, esto no lo podemos sumar, ya que es una medida no aditiva y nos daría información errónea como mostramos en el ejemplo a continuación:

Sucursal: Jose C Paz, Pilar, San Miguel
Mes: Ene, Feb, Mar

Participación Trimestral en el Mercado

Mes	Ene		Feb		Mar	
Sucursal	Kgs Vendidos	%	Kgs Vendidos	%	Kgs Vendidos	%
Jose C Paz	6200	48,63%	5000	48,55%	3000	72,24%
Pilar	1200	9,41%	998	9,69%	453	10,91%
San Miguel	5350	41,96%	4300	41,76%	700	16,86%

Figura 59: Porcentaje de participación de cada sucursal para cada mes del trimestre.

Sucursal: José C. Paz

Porcentaje por trimestre: Enero (48.63%), Febrero (48.55%) y Marzo (72.24%)

Porcentaje trimestral de la suma: 169.42%

Porcentaje trimestral Real: 52.50%

Para evitar este inconveniente y ver correctamente el porcentaje totalizado a nivel trimestral debemos hacer un agrupamiento por mes en donde se permita al modelo asociativo desagregar los kilos gramos vendidos de una sucursal para los meses del trimestre sobre el total vendido de todas las sucursales en los meses del trimestre. En la Figura 60 se muestra los datos de manera correcta:

Sucursal		Mes	
Jose C Paz	Ene	Pilar	Feb
Pilar	Mar	San Miguel	

Mes	Ene		Feb		Mar		Total	
Sucursal	Kgs Vendidos	%	Kgs Vendidos	%	Kgs Vendidos	%	Kgs Vendidos	%
Jose C Paz	6200	48,63%	5000	48,55%	3000	72,24%	14200	52,20%
Pilar	1200	9,41%	998	9,69%	453	10,91%	2651	9,75%
San Miguel	5350	41,96%	4300	41,76%	700	16,86%	10350	38,05%

Figura 60: Porcentaje de participación de cada sucursal a nivel trimestral.

Si no llegamos a hacer el agrupamiento para todos los meses, si bien vemos bien el total trimestral, internamente por mes se va a ver de manera incorrecta:

Sucursal		Mes	
Jose C Paz	Ene	Pilar	Feb
Pilar	Mar	San Miguel	

Mes	Ene		Feb		Mar		Total	
Sucursal	Kgs Vendidos	%	Kgs Vendidos	%	Kgs Vendidos	%	Kgs Vendidos	%
Jose C Paz	6200	22,79%	5000	18,38%	3000	11,03%	14200	52,20%
Pilar	1200	4,41%	998	3,67%	453	1,67%	2651	9,75%
San Miguel	5350	19,67%	4300	15,81%	700	2,57%	10350	38,05%

MAL

Figura 61: Porcentaje erróneo de participación de cada sucursal a nivel mensual pero bien a nivel trimestral.

d. Solución Técnica :

- *Formula Trimestral:* $\text{sum (Kgs_Helado)} / \text{sum (aggr (NODISTINCT sum (Kgs_Helado), Mes))}$

- Aggr: Es una función que utiliza el software QlikView para hacer el agrupamiento.

3.2.3.3 Medida: Retorno sobre los activos en 2013.

Es rendimiento de los recursos invertidos en la actividad del negocio.

a. Contexto:

En una empresa de ventas de productos al por menor se necesita saber el porcentaje del retorno sobre los activos que tiene la compañía durante el año fiscal 2013.

El ROA es un ratio de rendimiento financiero que sirve para medir el rendimiento de los recursos invertidos en la actividad del negocio. Cuanto mayor es el ROA mejor es la situación de la empresa.

b. Modelo Asociativo:

Para realizar la medida mencionada anteriormente, contamos con el siguiente modelo asociativo:

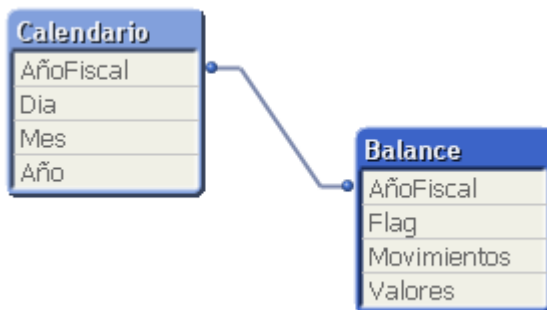


Figura 62: Modelo asociativo que representa la relación entre la tabla Calendario y Balance en QlikView.

El mismo está compuesto por las siguientes tablas:

- Balance: Representa todos los movimientos que tuvo la empresa para todos sus años fiscales.
- Calendario: Representa la fecha desagregada por año fiscal, mes, día.

c. Solución:

Se cuenta con la siguiente información proveniente del modelo asociativo referente al balance con todos los movimientos que tuvo la empresa para el periodo fiscal 2013, de enero a diciembre.

Año Fiscal: 2013

Movimientos	Valores
Ingresos por Ventas	90
Costo Directos de Producción	44
Contribución	46
Gastos Generales	24
Depreciación	6
RESULTADO OPERATIVO	18
Intereses	4
Resultado antes de Impuestos	14
Impuestos	3
RESULTADO NETO	11
Caja y Bancos	16
Cuentas por cobrar	20
Stock de Materiales	6
Productos en Proceso	18
Productos en Terminado	15
Activo Corriente	75
Equipos y Maquinaria	20
Inmuebles	5
Activo Fijo	25
TOTAL ACTIVO	100
A largo plazo	36
A corto plazo	24
Pasivo	60
Capital social	10
Resultado acumulado	30
Total Capital	40
TOTAL PASIVO Y CAPITAL	100

Figura 63: Balance del año fiscal 2013.

En base a la información disponible en la Figura 63, se debe sacar el retorno sobre los activos que tuvo la compañía para el año fiscal 2013.

Al ser una medida no aditiva, dado que bajo ninguna circunstancia se pueden sumar los valores por ninguna dimensión de tiempo y/o otra dimensión. Debemos hacer una simple división entre los valores correspondientes al resultado operativo y los valores correspondientes al total activo.

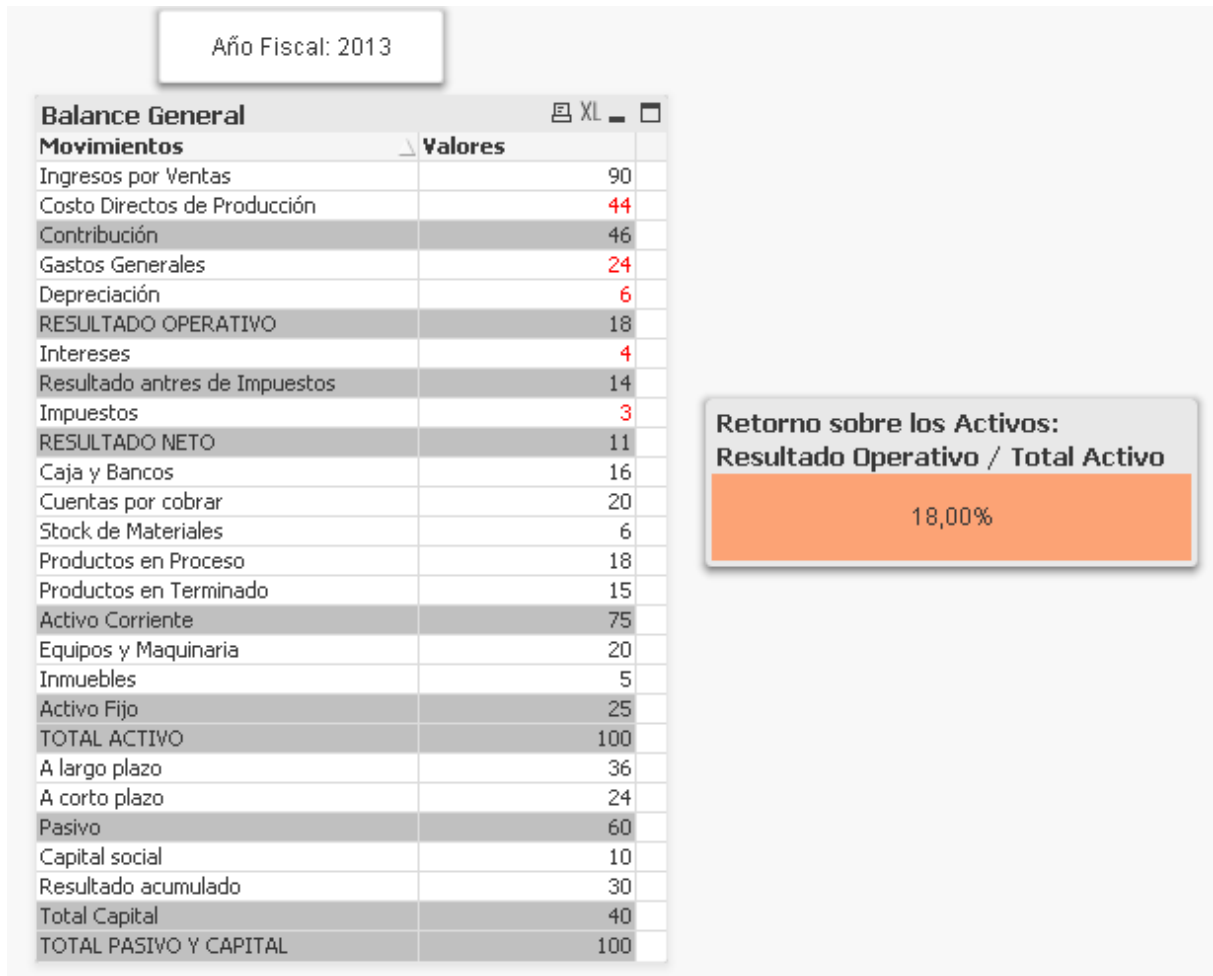


Figura 64: Retorno sobre los activos según balance del año fiscal 2013.

d. Solución Técnica

Para calcular la medida se utilizó la siguiente fórmula:

- *Retorno sobre los Activos:* $\text{sum}(\{\langle \text{Movimientos}=\{\text{'RESULTADO OPERATIVO'}\}\rangle\} \text{Valores}) / (\text{sum}(\{\langle \text{Movimientos}=\{\text{'TOTAL ACTIVO'}\}\rangle\} \text{Valores}))$.

Se hace una suma de los movimientos que pertenezcan al **resultado operativo** dividido la suma de los movimientos que representan **el total de los activos**.

4. Conclusión

En este trabajo presentamos una guía de buenas prácticas de diseño asociadas a la implementación de soluciones de Business Intelligence basadas en modelos asociativos, con el fin de permitir a los desarrolladores que cuenten con herramientas para implementar y calcular las medidas de negocio de los modelos analíticos en forma consistente y simple.

En primer lugar, debemos conocer las pautas de diseño para elaborar nuestro modelo asociativo, como también así definir qué medidas utilizaremos, a partir de los datos de nuestra empresa, esto fundamental para no cometer errores.

En segundo lugar, aplicar estas medidas de forma correcta teniendo como guía nuestro estándar metodológico, es de vital importancia, ya que los modelos Business Intelligence son los que agilizan y facilitan el acceso a la información, mostrándola de forma directa para la toma de decisiones por parte de los usuarios.

Para finalizar, todo lo mencionado anteriormente nos garantiza la disminución en los tiempos de desarrollo, mejora el diálogo y la comunicación con el usuario final, así como también la disminución del tiempo en la realización de los procesos de desarrollo de software al suprimir los tiempos de revisiones por errores.

Conceptualmente, nuestro trabajo viene a cubrir un espacio que no se encuentra desarrollado hasta el día de hoy en los manuales técnicos de la herramienta.

5. Bibliografía

- Dorian Pyle (1999), Data preparation for Data Mining, Morgan Kaufmann Publishers Inc., San Francisco, USA.
- Salvador Figueras M. y Gargallo.P, (2003) Análisis Exploratorio de Datos (en línea) 5campus.com, Estadística <http://www.5campus.com/leccion/aed>
- Kimball, Ralph, Ross, Margy. The data warehouse toolkit: the complete guide to dimensional modeling - 2da edición- Willey Computing Publishing
- Jiawei Han, Micheline Kamber.Data Mining, Concepts and Techniques- 2da edición- The Morgan Kaufmann Series in Data Management Systems Artículos
- Fayyad, U., Uthurusamy, R., Data Mining and Knowledge Discovery in Databases. Communications of the ACM 39, 11 (Nov. 1996), 24–26
- Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., The KDD Process for Extracting Useful Knowledge from Volumes of Data. Communications of the ACM 39, 11 (Nov. 1996), 27–34
- Glymour, C., Madigan, D., Pregibon, D., Smyth, P., Statical Inference an Data Mining. Communications of the ACM 39, 11 (Nov. 1996), 35–41
- Brachman, R., Khabaza, T., Kloesgen, W., Piatetsky-Shapiro, G., Simoudis, E., Mining Business Databases. Communications of the ACM 39, 11 (Nov. 1996), 42–48
- Inmon, W, The Data Warehouse and Data Mining. Communications of the ACM 39, 11 (Nov.1996), 49–50
- Fayyad, U., Haussler, D., Stolorz, P., Mining Scientific data. Communications of the ACM 39, 11 (Nov. 1996), 51–57
- Imelienski, T, Mannila, H, A database perspective on Knowledge Discovery. Communications of the ACM 39, 11 (Nov. 1996), 57–64
- Blaha, M., Data Warehouses and Decision Support Systems. Computer (Diciembre 2001), 38–39
- Pederson, T., Jensen, C., Multidimensional Database Technology. Computer (Diciembre 2001), 40–46

- Berndt, D., Fisher, J., Hevner, A., Studnicki, J., Healthcare Data Warehousing and Quality Assurance. Computer (Diciembre 2001), 56–65
- Trujillo, J., Palomar, M., Gomez, J., Song, I., Healthcare Designing Data Warehousing with Conceptual Models. Computer (Diciembre 2001), 66–75
- Armstrong, R., Seven Steps to Optimizing Data Warehouse Performance. Computer (Diciembre 2001), 76–79
- Chaudhuri S., Dayal U., An Overview of Data Warehousing and Olap Technology ACM SIGMOD Record 26, Marzo 1997 Complementaria
- César Pérez López, Daniel Santín González (2007), Minería de datos – Técnicas y Herramientas, Thomson, Madrid
- Joseph Bigus (1996), Data Mining with Neural Networks, Mc Graw Hill, USA.
- Michael Berry, Gordon Linoff (1997), Data Mining Techniques for Marketing, Sales and Customer Support, Wiley, USA,
- Robert Groth (1998), Data Mining: a hands-on approach for business professionals, Prentice Hall, USA
- Hair, Anderson, Tatham, Black (1999), Análisis Multivariante, Prentice Hall, Madrid.
- Michael Berry, Gordon Linoff (2000), Mastering Data Mining, Wiley, USA.
- A. Berson, S. Smith, K. Thearling, Data Mining applications for CRM (2000), Mc Graw Hill, USA.
- C. Seidman (2001), Data Mining with Microsoft SQL Server 2000, Microsoft Press, USA
- Simon Williams, The Associative Model of Data, Second Edition ed. Great Britain: Lazy Software Ltd, marzo 2004.

Enlaces web

BI- Definición- Glosario de Gartner:

<http://www.gartner.com>. Última fecha de acceso 20/12/2013.

Gartner- Cuadrante Mágico sobre plataformas de BI 2013:

<http://www.gartner.com/technology/reprints.do?id=1-1DYXHVT&ct=130206&st=sg>. Última fecha de acceso 14/11/2013.

Herramientas de BI - Pentaho:

http://www.intellego.com.mx/es/pentaho_bi. Última fecha de acceso 10/11/2013.

<http://www.pentaho.com>. Última fecha de acceso 10/11/2013.

Herramientas de BI – Tableau:

<http://www.tableausoftware.com/es-es/products/public>. Última fecha de acceso 02/11/2013.

<http://www.tableausoftware.com>. Última fecha de acceso 02/11/2013.

Herramientas de BI – Cognos:

<http://www-01.ibm.com/software/analytics/cognos>. Última fecha de acceso 05/11/2013.

Herramientas de BI – Oracle Hyperion:

<http://www.oracle.com/us/solutions/business-analytics>. Última fecha de acceso 05/11/2013.

Herramientas de BI – SAP- Business Object:

http://www.seidor.es/wps/portal/sei.Home/Soluciones/Inteligencia_negocio/SAP_Business_Objects.

Última fecha de acceso 08/11/2013.

Herramientas de BI – MicroStrategy:

<http://www.microstrategy.com/platforms/analytics/business-intelligence?CID=>. Última fecha de acceso 08/11/2013.

<http://prezi.com/bzlh5xnyuuqt>. Última fecha de acceso 08/11/2013.

<http://www.gesfor.com.mx>. Última fecha de acceso 08/11/2013.

Herramientas de BI – QlikView:

<http://www.qlikview.com/>. Última fecha de acceso 08/11/2013.

Peores prácticas en BI:

<http://www.mrc-productivity.com/blog/2013/09/7-worst-practices-of-business-intelligence/>.

Última fecha de acceso 13/10/2009.

BI- Arquitecturas:

<http://www.iteasolutions.com.co/index.php/servicios/bi2>. Última fecha de acceso 13/10/2013.