



FACULTAD DE INFORMÁTICA

TESINA DE LICENCIATURA

Programa de Apoyo al Egreso de Profesionales en Actividad

TÍTULO: Implementación Tablero de Control - Procedimiento Compra del Gobierno de la Ciudad de Buenos Aires
AUTOR: Mormando, Julio Maximiliano
DIRECTOR ACADÉMICO: Lic. Amadeo, Ana Paola
DIRECTOR PROFESIONAL: Marcote, Cristian Gabriel
CARRERA: Licenciatura en Sistemas

Resumen

El presente trabajo de tesina tiene como finalidad presentar la actividad profesional desarrollada en las oficinas de Sistemas del Ministerio de Hacienda de la Ciudad de Buenos Aires.

Se describen las tareas realizadas en la elaboración del Tablero de Control utilizado por el Gobierno para su gestión, indicando cada uno de los pasos implicados para su desarrollo, desde obtención del requerimiento hasta llegar a la solución, basado como se verá en la metodología de Ralph Kimball.

Como resultado final se obtiene el Tablero de Control que le permite al Estado la toma de decisiones gerenciales.

Palabras Clave

ETL, Data Warehouse, Data Mart, OLTP, OLAP, Esquema estrella, Esquema copo de nieve, Tabla de Hechos, Tabla de Dimensiones, Métricas, Jerarquías, Panel de Control, KPI, Metadatos

Conclusiones

Se presentó un caso de éxito de Business Intelligence, llevado adelante por el Gobierno de la Ciudad de Buenos Aires, que viene siendo utilizado desde el año 2010. En este trabajo de tesina se buscó mostrar cuáles son los conocimientos aplicados para la construcción de un Tablero de Control, demostrando que BI brinda soluciones estratégicas a la hora de gestionar la administración de entidades públicas y/o privadas.

Trabajos Realizados

Se realizaron consultas a fuentes bibliográficas específicas y se correlacionaron con el objetivo de dicha tesina.

Presente la arquitectura de servidores configuradas en las oficinas de DGUIAF.

Explique la herramienta de software utilizada para la construcción de los ETL

Explique la herramienta de software utilizada para la construcción del Tablero de Control.

Mostré las distintas etapas de trabajo realizadas hasta llegar a la solución.

Trabajos Futuros

Actualizar el modelo Data Warehouse, que se adapte mejor a la Ciencia de Datos y Big Data.

Incorporar herramientas que permitan analizar grandes volúmenes de datos, como por ejemplo R.

Contenido

Contenido.....	2
Agradecimientos	4
Presentación de la Propuesta PAEPA.....	5
Capítulo 1 – Marco teórico.....	9
1. Introducción.....	9
1.1 - Qué es Business Intelligences?	9
1.2 - Definiciones.....	10
1.3 - Principales productos de Business Intelligence	12
1.3.1 Cuadros de Mando Integrales (CMI).....	12
1.3.2 Sistemas de Soporte a la Decisión (DSS).....	13
1.3.3. Sistemas de Información Ejecutiva (EIS)	16
1.4 Principales componentes de orígenes de datos.....	17
Capítulo 2 - Análisis de herramientas	22
2. Introducción:	22
2.1 Plataforma de trabajo.....	22
2.2 ETL.....	24
2.3 Microstrategy	30
¿Qué es MicroStrategy?.....	30
Componentes de Microstrategy	31
2.3.1. Conceptos y terminología	32
2.3.2. Introducción a Microstrategy Agente.....	34
Capítulo 3 – Implementación de la solución Business Intelligence.....	37
3. Introducción.....	37
3.1 Planificación del Proyecto Delimitación de Alcance del Proyecto.	40
3.2 Definición de Requerimientos del Negocio.....	40
Requerimiento Funcional	40
Requerimiento Técnico.....	40
3.3 Modelado Dimensional de Negocio (BDM).....	41
3.4 Diseño Físico.....	42
3.5 Diseño y Desarrollo de Presentación de Datos	45
3.5.1 Desarrollo ETL	45
3.5.2 ¿Cómo funciona el paquete ‘Carga diaria’?	46

3.5.3 Proceso de carga tabla de la DIM_OBJETO_GASTO	47
3.5.4 Proceso de carga tabla de la FACT_COMPRA.....	49
3.6 Desarrollo Tablero – Microstrategy	50
Creación de Atributos	50
Creación de un indicador simple.....	51
3.7 Pruebas de Calidad de Datos.....	52
Alcance.....	54
Prueba de Proceso.....	54
Prueba de Datos.....	55
Capítulo 4 - Entrega del tablero	58
4. Introducción.....	58
4.1 Tablero de Control	58
Conclusión Final	64
Trabajo Futuro.....	65
Anexo Query	66
1 - Proceso Id.....	66
2 - Actualizo A Estado 1.....	66
3 - Trunco Table Dim_Objeto_Gasto.....	66
4 - Cargo Table Dim_Objeto_Gasto.....	66
5 - Cargo Table Dim_Fecha.....	66
6 - Carga Table Fact_Compra	67
Bibliografía	68

Agradecimientos

Orgullosa de haber realizado mis estudios universitarios en una institución pública como lo es la UNLP, agradezco a la Facultad de Informática por crecer año a año, brindando lo mejor para sus alumnos.

Agradezco a mi familia por brindarme la posibilidad de estudiar esta carrera y acompañarme en cada momento.

Agradezco a la Profesora Amadeo Paola, por su dedicación y motivación recibida desde el primer día que se ofreció para ser mi directora de tesina.

En lo personal, desarrollar el trabajo de Tesina fue una gran experiencia debido a que consistió en documentar un trabajo realizado en mi ámbito profesional. Logre exponer mis conocimientos prácticos en papel, como también los conocimientos teóricos, avalado por las diferentes bibliografías consultadas, siendo algunas de ellos profesores de la Universidad Nacional de La Plata, como autores de reconocimiento mundial.

Presentación de la Propuesta PAEPA

Nombre del Trabajo	Implementación Tablero de Control - Procedimiento Compra del Gobierno de la Ciudad de Buenos Aires				
Director	Lic. Amadeo Ana Paola				
Codirector	Marcote, Cristian Gabriel				
Alumnos	Nro. Alumno	Apellido y Nombres		Dirección, teléfono y e-mail	
	7552/3	Mormando, Julio Maximiliano		15 N° 776, 2215433837, juliomormando@gmail.com	
Clasificación	Investigación Teórica	Investigación Aplicada	Desarrollo Tecnológico	Transferencia Tecnológica	Temática de RS
Plazo de Ejecución	3 meses, 90 días.				

Puntos:

- Experiencia laboral
- Motivación
- Objetivo General
- Objetivo Especifico
- Desarrollos propuestos
- Resultados esperados
- Referencias bibliográficas

• Experiencia laboral del postulante

Me encuentro trabajando en la Contaduría General de la Provincia de Buenos Aires, desde Octubre del 2017. Desempeño mi actividad laboral en al área de Business Intelligence (BI), como desarrollador, realizando la construcción y mantenimiento de los procesos de Extracción, Transformación y Carga de los datos (ETL). Así como también, la construcción y mantenimiento de los Tableros de Control usando la herramienta de software Microstrategy.

Comencé iniciando mi experiencia como desarrollador BI en el año 2014, en la Dirección General Unidad Informática de Administración Financiera (DGUIAF). La DGUIAF depende de la Subsecretaría de Gestión Operativa perteneciente al Ministerio de Economía y Finanzas del Gobierno de la Ciudad Autónoma de Buenos Aires. Proporciona soporte informático a través del diseño, desarrollo, implementación, administración y capacitación del Sistema Integrado de Gestión y Administración Financiera (SIGAF).

SIGAF, hoy, se encuentra implementado en la Provincia de Buenos Aires desde el año 2018, con motivo de esta implementación, posibilitó mi traslado del gobierno de la ciudad de Buenos Aires al de la Provincia de Buenos Aires, ya que me encuentro radicado en la ciudad de La Plata.

• **Motivación**

El gobierno de la Ciudad de Buenos Aires utiliza como sistema de Gestión y Administración Financiera el Sistema Integrado de Gestión y Administración Financiera (SIGAF), desde el año 2005, a partir de la aplicación de la “Ley 70, Decreto reglamentario N° 1000/GCBA/99”, Ley de Gestión, Administración Financiera y Control del Sector Público de la Ciudad.

SIGAF es un sistema de información que apoya la gestión presupuestaria, financiera y contable de los Ministerios del Gobierno Central, así como las funciones de rectoría y fiscalización que corresponden al Ministerio de Economía y Finanzas y a los Órganos de Control.

Este sistema se encuentra diseñado en forma modular posibilitando el ingreso, la verificación y el control de la documentación, en toda la etapa de la gestión administrativa-contable del Gobierno de la Ciudad Autónoma de Buenos Aires.

En la actualidad, personal del gobierno como ministros, secretarios, usuarios administrativos del sistema, entre otros, ya cuentan con una herramienta llamada Microstrategy, que les permite generar reportes de manera ágil para la toma de decisiones, que deben ser realizadas de forma estratégica e inmediata.

Mediante los indicadores y resultados obtenidos se puede determinar medidas de cambios y mejoras basadas en el análisis de información y datos procesados, que generan una visión más clara de la gestión del gobierno.

La implementación de una solución BI representa un apoyo de gran utilidad al momento de realizar informes, puesto que las personas encargadas de realizar los reportes, dejan de lado sus tareas asignadas para entregar los reportes a tiempo, a la persona encargada de tomar las decisiones, quienes requieren contar con información relevante para definir estrategias claves.

En reuniones gerenciales o interdepartamentales que se llevan a cabo, existe información vital, que se necesita conocer, la cual esta herramienta es de gran apoyo por su repuesta rápida en la obtención de los reportes, con grandes volúmenes de datos. La información obtenida debe ser oportuna, veraz, explícita y confiable, para que de esta manera las medidas tomadas, sean las más acertadas en beneficio de la sociedad por parte del gobierno.

• **Objetivo General**

El objetivo general es implementar un Dashboard para la gestión del sistema de soporte de decisiones (SSD), solicitado por el área de compras del Gobierno de la Ciudad de Buenos Aires.

El siguiente Dashboard, visualizara los indicadores claves (KPI's), Importe y Cantidad, los cuales serán sumarizados por los clasificadores Jurisdicción¹ y Objeto del Gasto².

Se aplicaran las mejores prácticas metodológicas desarrolladas por Kimball R, para soluciones de BI.

Objetivo Especifico

Las actividades que se describirán en el siguiente trabajo de tesina son:

- análisis de requerimientos de la organización para identificar los indicadores clave (KPI's)
- análisis de los sistemas de procesamiento transaccional (OLTP's)
- diseño del almacén de datos (Data Warehouse)
- proceso de integración de datos o ETL (Extracción, transformación y carga)

¹ La palabra Jurisdicción hace referencia a los departamentos de gobierno encargado de un área política determinada. Por ej. Ministerio de Salud, Ministerio de Educación, etc.

² El clasificador objeto del gasto ha sido diseñado con un nivel de desagregación que permite que sus cuentas faciliten el registro único de todas las transacciones con incidencia económica financiera que realiza una institución pública. Es un instrumento informativo para el análisis y seguimiento de la gestión financiera de sector público y, en consecuencia, se distingue como clasificador analítico o primario del sistema de clasificaciones presupuestarias. Desde el punto de vista estructural, tiene cuatro niveles de cuentas: el de los incisos, el de las partidas principales, el de las partidas parciales y el de las partidas subparciales.

- y finalmente, la implementación una Dashboard para el soporte de decisiones.

Se aplicara el proceso metodológico creado por Kimball, utilizando herramientas privativas como Microstrategy, Visual Studio (SSIS Integration) y el DBMS SQL Server, obteniendo como resultado un software tipo panel de control que consolida la información clave e histórica de la organización que proporciona a los empleados administrativos, la capacidad de análisis descriptivo multidimensional en línea (OLAP³) de apoyo para sus decisiones.

• **Desarrollos propuestos:**

Para el siguiente trabajo de tesina se explicaran los ambientes de trabajo utilizados, indicando cuales son los servidores que se utilizan para realizar los desarrollos, pruebas y puestas en producción. Se realizara una introducción a las herramientas de software utilizadas para llevar adelante los desarrollos, mostrando como se crea el Data Warehouse y los procesos que se realizan para realizar las cargas de datos. Posterior a este trabajo, se construirán los atributos e indicadores que darán paso a la construcción del tablero.

• **Resultados esperados**

Al finalizar el documento de tesina, se espera haber transmitido de forma clara y simple, los pasos realizados para el desarrollo del tablero de control implementado para el gobierno de la Ciudad de Buenos Aires.

Con la finalización del desarrollo del tablero de control, se buscó cumplir con las expectativas de gobierno, brindándoles una herramienta de trabajo, que le sea útil para ver el estado de las compras realizadas, como también, les sirva para tomar decisiones en base a los datos brindados.

³ Acrónimo en inglés de procesamiento analítico en línea (On-Line Analytical Processing). Es una solución utilizada en el campo de la llamada Inteligencia de negocios (o Business Intelligence) cuyo objetivo es agilizar la consulta de grandes cantidades de datos

Capítulo 1 – Marco teórico

1. Introducción

En este capítulo se describirá el concepto de Inteligencia de Negocio (BI - Business Intelligence), explicando los hechos que dieron lugar a su nacimiento, como también, presentando las distintas definiciones que expusieron los autores en base a sus estudios realizados. Se explicaran los productos Cuadro de Mando (CMI), Sistemas de Soporte de Decisión (DSS) y Sistema de Información Ejecutiva (EIS), que ofrece BI para los usuarios de acuerdo a sus necesidades, buscando obtener un mayor análisis de sus datos. Posteriormente, se explicará el significado de Data Mart, Data Warehouse, Data Mart OLAP y Data Mart OLTP, siendo los componentes principales que se requieren construir el producto.

1.1 - Qué es Business Intelligences?

Josep Lluís Cano [18], en su libro realiza la siguiente introducción de Business Intelligence.

“Vivimos en la sociedad de la información. Gracias a Internet y al desarrollo de los sistemas de información en las empresas, sus directivos pueden acceder a mucha más información, de más calidad y con mayor rapidez. El potencial que ello ofrece para mejorar la toma de decisiones y para guiar a las empresas hacia la consecución de sus objetivos es enorme. Sin embargo, muchos directivos se enfrentan a la paradoja de que “cada vez tienen más información y menos tiempo para analizarla”.

Para enfrentar estos problemas, en los últimos años han surgido una serie de técnicas que facilitan el procesamiento avanzado de los datos. La idea clave es que los datos contienen más información oculta de la que se ve a simple vista. El verdadero valor de la información se revela cuando a partir de ella somos capaces de descubrir conocimiento. Este es el verdadero objetivo de la Business Intelligence.”

En la actualidad, en cualquier organización se hace necesario la toma decisiones. Generalmente, la información que se quiere investigar sobre un cierto dominio de la organización se encuentra en bases de datos y en otras fuentes de almacenamiento.

Surge aquí la necesidad de juntar los distintos archivos y bases de datos de manera que se pueda utilizarlos para extraer conclusiones.

La estructuración de los datos no es sencilla y esto se dificulta cuando los diferentes archivos o bases de datos se encuentran en sistemas diferentes, como ilustra la imagen 1. Lo razonable sería recoger los datos (información histórica) en un sistema separado y específico. Nace el Data Warehouse (DWH), Almacén de Datos, con la necesidad de unificar los distintos archivos y bases de datos para poder comprenderlos. Por ello, se necesita de tecnologías que sirvan de guía para comprender el contenido de las Bases de Datos.

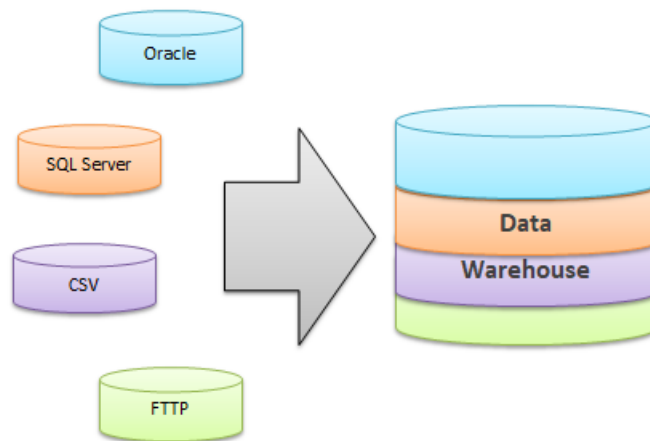


Imagen 1. Formación Data Warehouse

1.2 - Definiciones

Son muchos los expertos de la industria de la Tecnología de la Información que han dado una definición al concepto Business Intelligence. Veamos a continuación, qué puntos en común y qué diferencias presenta cada definición.

"Desde un punto de vista más pragmático, y asociándolo directamente a las tecnologías de la información, podemos definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la compañía) en información estructurada, para su explotación directa (reporting, análisis OLAP...) o para su análisis y conversión en conocimiento soporte a la toma de decisiones sobre el negocio." [19].

"Business Intelligence engloba todos los sistemas con los que se puede obtener no sólo información o conocimiento, sino una auténtica inteligencia capaz de

proporcionar una ventaja competitiva. Se trata, por lo tanto, de la habilidad que tienen las compañías para manejar, consolidar y analizar sus datos, de forma que estos se transformen en información valiosa para tomar las mejores decisiones estratégicas. Incrementar las ventas, mejorar los productos y las relaciones con los clientes, aumentar la calidad de los servicios prestados, reducir costes... son los principales beneficios que alcanzan las empresas que cuentan con soluciones de BI." [30].

"Business Intelligence es un enfoque para la gestión empresarial que permite a una organización definir qué información es útil y relevante para la toma de decisiones corporativas. Business Intelligence es un concepto polifacético que fortalece a las organizaciones para tomar mejores decisiones más rápidamente, convertir los datos en información y usar una estrategia inteligente para la gestión empresarial." [16]

Como se observa, el punto en común que presentan las distintas definiciones es sin duda la ventaja competitiva que una empresa adquiere con la implantación de un sistema inteligente de negocio. Una ventaja competitiva que se adquiere utilizando una estrategia inteligente; esto es, identificar qué información es de vital importancia para la empresa, identificar los factores clave de rendimiento, es decir, aquello que la empresa cree que es importante medir y analizar.

Dependiendo del sector desde el que se define el término Business Intelligence, éste puede presentar diferentes matices. Si se define el término desde un punto de vista de negocio, los aspectos a resaltar son, "estrategia, rendimiento, competencia, decisión". Si por el contrario, la definición viene de un entorno más tecnológico, es fácil encontrarnos con términos como "análisis, tratamiento de información, modelado, reporting".

Business Intelligence es la habilidad para transformar los datos en información, y la información en conocimiento, de forma que se pueda optimizar el proceso de toma de decisiones en los negocios [28]. La Imagen 2, grafica el concepto.

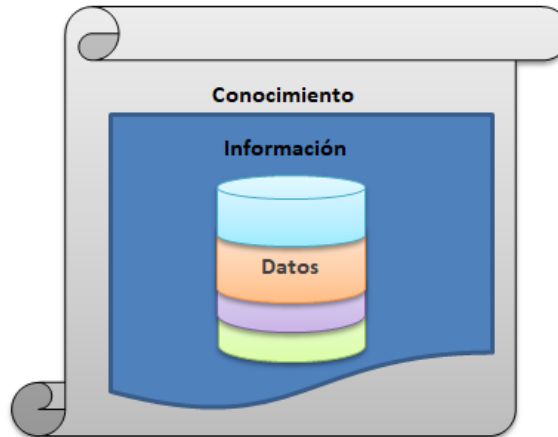


Imagen 2. Del Dato al Conocimiento

Desde un punto de vista más pragmático, y asociándolo directamente con las tecnologías de la información, podemos definir Business Intelligence como el conjunto de metodologías, aplicaciones y tecnologías que permiten reunir, depurar y transformar datos de los sistemas transaccionales e información desestructurada (interna y externa a la organización) en información estructurada, para su explotación directa (reporting, análisis OLTP / OLAP, alertas...) o para su análisis y conversión en conocimiento, dando así soporte a la toma de decisiones sobre el negocio.

La Inteligencia de Negocio actúa como un factor estratégico para una empresa u organización, generando una potencial ventaja competitiva, que no es otra que proporcionar información privilegiada para responder a los problemas de negocio: entrada a nuevos mercados, promociones u ofertas de productos, eliminación de islas de información, control financiero, optimización de costes, planificación de la producción, análisis de perfiles de clientes, rentabilidad de un producto concreto, etc..

A continuación se definen los principales productos de Business Intelligence, definidos por el portal Web sinnexus.com [28].

1.3 - Principales productos de Business Intelligence

1.3.1 Cuadros de Mando Integrales (CMI)

El Cuadro de Mando Integral (CMI), también conocido como **Balanced Scorecard (BSC)** o **Dashboard**, es una herramienta de control empresarial que

permite establecer y monitorizar los objetivos de la organización y de sus diferentes áreas o unidades.

También se puede considerar como una aplicación que ayuda a una organización a expresar los objetivos e iniciativas necesarias para cumplir con su estrategia, mostrando de forma continuada cuándo la organización y los empleados alcanzan los resultados definidos en su plan estratégico.

1.3.2 Sistemas de Soporte a la Decisión (DSS)

Un **Sistema de Soporte a la Decisión (DSS)** es una herramienta de BI enfocada al análisis de los datos de una organización.

En principio, puede parecer que el análisis de datos es un proceso sencillo, y fácil de conseguir mediante una aplicación hecha a medida o un ERP sofisticado. Sin embargo, no es así: estas aplicaciones suelen disponer de una serie de informes predefinidos en los que presentan la información de manera estática, pero no permiten profundizar en los datos, navegar entre ellos, manejarlos desde distintas perspectivas.



Imagen 3. Sistema de Soporte a la Decisión

El DSS es una de las herramientas más emblemáticas del Business Intelligence ya que, entre otras propiedades, permiten resolver gran parte de las limitaciones de los programas de gestión. Estas son algunas de sus características principales:

- **Informes dinámicos, flexibles e interactivos**, de manera que el usuario no tenga que ceñirse a los listados predefinidos que se configuraron en el momento de la implantación, y que no siempre responden a sus dudas reales.

- **No requiere conocimientos técnicos.** Un usuario no técnico puede crear nuevos gráficos e informes y navegar entre ellos, haciendo drag&drop o drill through. Por tanto, para examinar la información disponible o crear nuevas métricas no es imprescindible buscar auxilio en el departamento de informática.
- **Rapidez en el tiempo de respuesta,** ya que la base de datos subyacente suele ser un Data Warehouse corporativo o un Data Mart, con modelos de datos en estrella o copo de nieve. Este tipo de bases de datos están optimizadas para el análisis de grandes volúmenes de información (véase análisis OLTP-OLAP).
- **Integración entre todos los sistemas/departamentos de la organización.** El proceso de ETL previo a la implantación de un Sistema de Soporte a la Decisión garantiza la calidad y la integración de los datos entre las diferentes unidades de la organización. Existe lo que se llama: integridad referencial absoluta.
- **Cada usuario dispone de información adecuada a su perfil.** No se trata de que todo el mundo tenga acceso a toda la información, sino de que tenga acceso a la información que necesita para que su trabajo sea lo más eficiente posible.
- **Disponibilidad de información histórica.** En estos sistemas está a la orden del día comparar los datos actuales con información de otros períodos históricos de la organización, con el fin de analizar tendencias, fijar la evolución de parámetros de negocio... etc.

Diferencia con otras herramientas de Business Intelligence

El principal objetivo de los Sistemas de Soporte a Decisiones es, a diferencia de otras herramientas como los Cuadros de Mando (CMI) o los Sistemas de Información Ejecutiva (EIS), explotar al máximo la información residente en una base de datos corporativa (Data Warehouse o Data Mart), mostrando informes muy dinámicos y con gran potencial de navegación, pero siempre con una interfaz gráfica amigable, vistosa y sencilla, como se ilustra en la Imagen 4.



Imagen 4. Dashboard.

Otra diferencia fundamental radica en los usuarios a los que están destinadas las plataformas DSS: cualquier nivel gerencial dentro de una organización, tanto para situaciones estructuradas como no estructuradas. (En este sentido, por ejemplo, los CMI están más orientados a la alta dirección).

Por último, destacar que los DSS suelen requerir (aunque no es imprescindible) un motor OLAP subyacente, que facilite el análisis casi ilimitado de los datos para hallar las causas raíces de los problemas/pormenores de la organización.

Tipos de Sistemas de Soporte a Decisiones

- Sistemas de información gerencial (MIS)

Los sistemas de información gerencial (MIS, Management Information Systems), también llamados Sistemas de Información Administrativa (AIS) dan soporte a un espectro más amplio de tareas organizacionales, encontrándose a medio camino entre un DSS tradicional y una aplicación CRM/ERP implantada en la misma organización.

- Sistemas de información ejecutiva (EIS)

Los sistemas de información ejecutiva (EIS, Executive Information System) son el tipo de DSS que más se suele emplear en Business Intelligence, ya que proveen a los gerentes de un acceso sencillo a información interna y externa de su organización, y que es relevante para sus factores clave de éxito.

- Sistemas expertos basados en inteligencia artificial (SSEE)

Los sistemas expertos, también llamados sistemas basados en conocimiento, utilizan redes neuronales para simular el conocimiento de un experto y utilizarlo de forma efectiva para resolver un problema concreto. Este concepto está muy relacionado con el **Data Mining**.

- **Sistemas de apoyo a decisiones de grupo (GDSS)**

Un sistema de apoyo a decisiones en grupos (GDSS, Group Decision Support Systems) es "un sistema basado en computadoras que apoya a grupos de personas que tienen una tarea (u objetivo) común, y que sirve como interfaz con un entorno compartido". El supuesto en que se basa el GDSS es que si se mejoran las comunicaciones se pueden mejorar las decisiones.

1.3.3. Sistemas de Información Ejecutiva (EIS)

Un Sistema de Información para Ejecutivos o Sistema de Información Ejecutiva es una herramienta software, basada en un DSS, que provee a los gerentes de un acceso sencillo a información interna y externa de su organización, y que es relevante para sus factores clave de éxito.

La finalidad principal es que el ejecutivo tenga a su disposición un panorama completo del estado de los indicadores de negocio que le afectan al instante, manteniendo también la posibilidad de analizar con detalle aquellos que no estén cumpliendo con las expectativas establecidas, para determinar el plan de acción más adecuado.

De forma más pragmática, se puede definir un EIS como una aplicación informática que muestra informes y listados (query & reporting) de las diferentes áreas de negocio, de forma consolidada, para facilitar la monitorización de la organización o de una unidad de la misma.

El EIS se caracteriza por ofrecer al ejecutivo un acceso rápido y efectivo a la información compartida, utilizando interfaces gráficas visuales e intuitivas. Suele incluir alertas e informes basados en excepción, así como históricos y análisis de tendencias. También es frecuente que permita la domiciliación por correo de los informes más relevantes.

A través de esta solución se puede contar con un resumen del comportamiento de una organización o área específica, y poder compararla a través del tiempo. Es posible, además, ajustar la visión de la información a la teoría de Balanced Scorecard o Cuadro de Mando Integral impulsada por Norton y Kaplan, o bien a cualquier modelo estratégico de indicadores que maneje la organización.

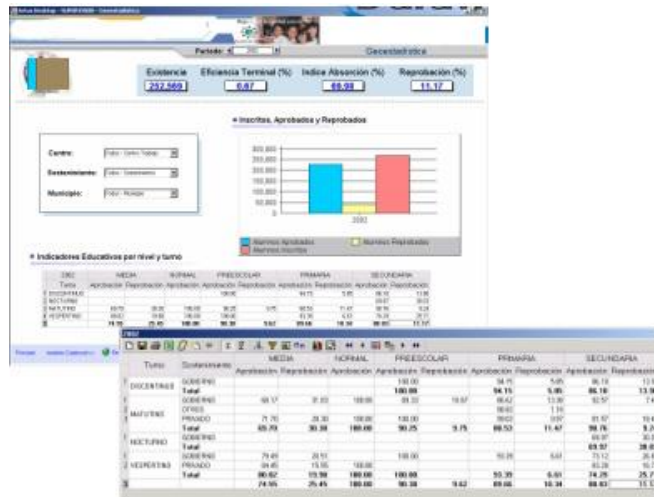


Imagen 5. Sistemas de Información Ejecutiva

1.4 Principales componentes de orígenes de datos

Datos Marts. representa un subconjunto de datos enfocados en el análisis de un departamento, área o ámbito específico en una organización [17].

Se caracteriza por disponer la estructura óptima de datos para analizar la información al detalle desde todas las perspectivas que afecten a los procesos de dicho departamento. Un Data Mart puede ser alimentado desde los datos de un Data Warehouse, o integrar por sí mismo un compendio de distintas fuentes de información.

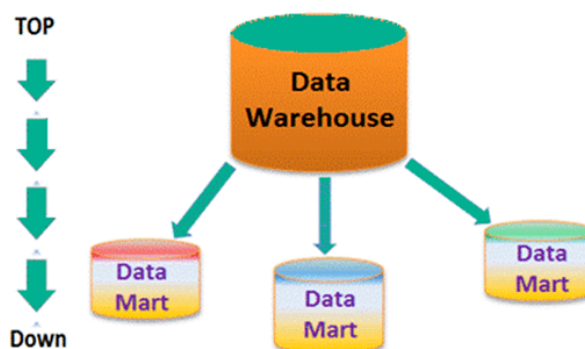


Imagen 6, Representación Data Mart

Por tanto, para crear el Data Mart de un área funcional de la organización es preciso encontrar la estructura óptima para el análisis de su información, estructura que puede estar montada sobre una base de datos OLTP, como el propio Data Warehouse, o sobre una base de datos OLAP. La designación de una u otra

dependerá de los datos, los requisitos y las características específicas de cada departamento. De esta forma se pueden plantear dos tipos de Data Marts:

Data Mart OLAP. Se basan en los populares cubos OLAP, que se construyen agregando, según los requisitos de cada área o departamento, las dimensiones y los indicadores necesarios de cada cubo relacional. El modo de creación, explotación y mantenimiento de los cubos OLAP es muy heterogéneo, en función de la herramienta final que se utilice.

Data Mart OLTP. Pueden basarse en un simple extracto del Data Warehouse, no obstante, lo común es introducir mejoras en su rendimiento (las agregaciones y los filtrados suelen ser las operaciones más usuales) aprovechando las características particulares de cada área de la organización. Las estructuras más comunes en este sentido son las tablas report, que vienen a ser fact-tables reducidas (que agregan las dimensiones oportunas), y las vistas materializadas, que se construyen con la misma estructura que las anteriores, pero con el objetivo de explotar la reescritura de queries (aunque sólo es posible en algunos SGBD avanzados, como Oracle).

Los Data Marts que están dotados con estas estructuras óptimas de análisis presentan las siguientes ventajas:

- Poco volumen de datos
- Mayor rapidez de consulta
- Consultas SQL y/o MDX sencillas
- Validación directa de la información
- Facilidad para la historización de datos

Data Warehouse. Bodega o almacén de datos, contiene información histórica, consolidada, integrada, organizada y temática (Rosado G. & Rico B., 2010).

Es una base de datos corporativa que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta. La creación de un Data Warehouse representa en la mayoría de las ocasiones el primer paso, desde el punto de vista técnico, para implantar una solución completa y fiable de Business Intelligence.

La ventaja principal de este tipo de bases de datos radica en las estructuras en las que se almacena la información (modelos de tablas en estrella, en copo de nieve, cubos relacionales... etc.). Este tipo de persistencia de la información es homogénea y fiable, y permite la consulta y el tratamiento jerarquizado de la misma (siempre en un entorno diferente a los sistemas operacionales).

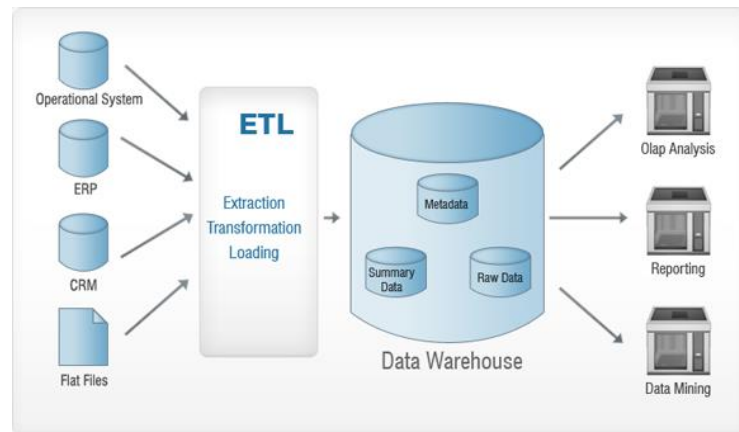


Imagen 7, Representación Data Warehouse

El término Data Warehouse fue acuñado por primera vez por Bill Inmon, y se traduce literalmente como almacén de datos. No obstante, y como cabe suponer, es mucho más que eso. Según definió el propio Bill Inmon [4], un Data Warehouse se caracteriza por ser:

- Integrado: los datos almacenados en el Data Warehouse deben integrarse en una estructura consistente, por lo que las inconsistencias existentes entre los diversos sistemas operacionales deben ser eliminadas. La información suele estructurarse también en distintos niveles de detalle para adecuarse a las distintas necesidades de los usuarios.
- Temático: sólo los datos necesarios para el proceso de generación del conocimiento del negocio se integran desde el entorno operacional. Los datos se organizan por temas para facilitar su acceso y entendimiento por parte de los usuarios finales. Por ejemplo, todos los datos sobre clientes pueden ser consolidados en una única tabla del Data Warehouse. De esta forma, las peticiones de información sobre clientes serán más fáciles de responder dado que toda la información reside en el mismo lugar.

- Histórico: el tiempo es parte implícita de la información contenida en un Data Warehouse. En los sistemas operacionales, los datos siempre reflejan el estado de la actividad del negocio en el momento presente. Por el contrario, la información almacenada en el Data Warehouse sirve, entre otras cosas, para realizar análisis de tendencias. Por lo tanto, el Data Warehouse se carga con los distintos valores que toma una variable en el tiempo para permitir comparaciones.
- No volátil: el almacén de información de un Data Warehouse existe para ser leído, pero no modificado. La información es por tanto permanente, significando la actualización del Data Warehouse la incorporación de los últimos valores que tomaron las distintas variables contenidas en él sin ningún tipo de acción sobre lo que ya existía.

Otra característica del Data Warehouse es que contiene metadatos, es decir, datos sobre los datos. Los metadatos permiten saber la procedencia de la información, su periodicidad de refresco, su fiabilidad, forma de cálculo, entre otras funciones.

Los metadatos serán los que permiten simplificar y automatizar la obtención de la información desde los sistemas operacionales a los sistemas informacionales.

Los objetivos que deben cumplir los metadatos, según el colectivo al que va dirigido, son:

- Dar soporte al usuario final, ayudándole a acceder al Data Warehouse con su propio lenguaje de negocio, indicando qué información hay y qué significado tiene. Ayudar a construir consultas, informes y análisis, mediante herramientas de Business Intelligence como DSS, EIS o CMI.
- Dar soporte a los responsables técnicos del Data Warehouse en aspectos de auditoría, gestión de la información histórica, administración del Data Warehouse, elaboración de programas de extracción de la información, especificación de las interfaces para la realimentación a los sistemas operacionales de los resultados obtenidos... etc.

Por último, destacar que para comprender íntegramente el concepto de Data Warehouse, es importante entender cuál es el proceso de construcción del mismo, denominado ETL (Extracción, Transformación y Carga), explicados en el siguiente capítulo, a partir de los sistemas operaciones de una organización:



- Extracción: obtención de información de las distintas fuentes tanto internas como externas.
- Transformación: filtrado, limpieza, depuración, homogeneización y agrupación de la información.
- Carga: organización y actualización de los datos y los metadatos en la base de datos.

Capítulo 2 - Análisis de herramientas

2. Introducción:

En el capítulo 1 se realizó una introducción a los principales términos de Inteligencia de Negocios, describiendo las principales definiciones realizadas por los grandes autores. También, se hizo mención a los principales componentes como Data Mart, Data Warehouse, Data Mart OLAP, Data Mart OLTP, que pueden ser requeridos para la creación del producto.

Este capítulo, estará dividido en tres grandes temas. En el primero se describirán los servidores en los cuales se trabaja, las herramientas que se utilizan para desarrollar el tablero. El segundo tema explica en que consiste realizar los ETL desde lo conceptual hasta como se realiza la ejecución de ellos. Por último, se realizara una introducción a Microstrategy, como herramienta principal para la creación del tablero, describiendo como están organizados los elementos de la herramienta y cuales son ellos.

2.1 Plataforma de trabajo

Para comenzar con la creación del tablero vamos a necesitar de la ayuda de varias tecnologías. Para dar comienzo a la construcción del tablero, el primer paso es tener instalado los siguientes softwares;

- SQL Server 2008,
- Integration Service 2008,
- Microstrategy Desktop / Web
- Terminal, Chrome, Internet Explorer, Mozilla FireFox

SQL Server 2008, es el motor de la base de datos, que me servirá para almacenar los datos de Data Warehouse. Esta herramienta se instala desde su página oficial, <https://www.microsoft.com/es-ar/download/details.aspx?id=30438>

Integration Service 2008, es la herramienta que me permite crear cada uno de los ETL, para luego poder realizar su ejecución y controlar su estado de actividad. Esta herramienta se instala desde su página oficial, <https://docs.microsoft.com/en-us/sql/integration-services/install-windows/install-integration-services?view=sql-server-ver15>

Microstrategy Desktop/Web, crea el frontend del tablero. Esta herramienta se instala desde su página oficial, <https://www.microstrategy.com/es>

Chrome, Internet Explorer, Mozilla Firefox, son algunas de las posibles terminales que se utilizan para consultar el tablero.

Para construir el tablero se requiere un servidor de desarrollo, un ambiente de prueba y por último un ambiente de producción. En cada uno de estos servidores se encuentra instalados los softwares mencionados anteriormente.

Una particularidad de estos ambientes de trabajo es que cuentan con dos bases de datos cada uno de ellos. En una base de datos, Data Warehouse, se almacenan los datos que van a ser visualizados en el tablero. Y la otra base de datos, Meta Data, es utilizada por Microstrategy siendo un repositorio que almacena la definición de los objetos y la información sobre el modelo de base de datos explotado.

El ambiente de Desarrollo, es utilizado para realizar la construcción de los ETL, explicado más adelante dentro de este capítulo, y del tablero. Los desarrollos se realizan con Integration Services 2008, tomando los datos desde su fuente origen (SIGAF), que se almacena en una base de datos Oracle, cargándolos el Warehouse en SQL Server 2008.

Se crean las tablas del modelo del Data Warehouse y el proceso encargado de extraer los datos fuentes, transformarlos y cargarlos en estas tablas correspondientes, llamadas con el prefijo DIM y FT. También se diseña la pantalla que mostrara los datos en el tablero. Todo lo realizado dentro de este servidor es a nivel de diseño y prueba de ejecución, comprobando que los procesos se ejecuten correctamente. Acá no se validan los datos obtenidos, ya que la fuente de datos es una base de datos sin estar actualizada.

Una vez realizado el desarrollo, se pasa a trabajar en el ambiente de Prueba. Aquí se realizan tanto las pruebas de ejecución de los procesos como la de la validez de los datos, consultando y analizando que dichos datos sean correctos, exactos, con un formato esperado. En esta etapa, se trabaja con los mismos datos que producción, para permitir que las pruebas que se realicen sean iguales a las que se van a obtener en el ambiente Productivo.

Por último, una vez que los desarrollos se testearon y se obtuvieron resultados exitosos, el proyecto es pasado al servidor de Producción. Estando el tablero en este

ambiente ya puede ser utilizado por el usuario que lo solicito, también llamado Sponsor.

De aquí en adelante sigue la etapa de Mantenimiento, que busca mantener la integridad en sus datos, refiriéndose a la validez y a la consistencia de los datos almacenados.

ESQUEMA DE SERVIDORES - WAREHOUSE Y METADATA DE APLICACIÓN

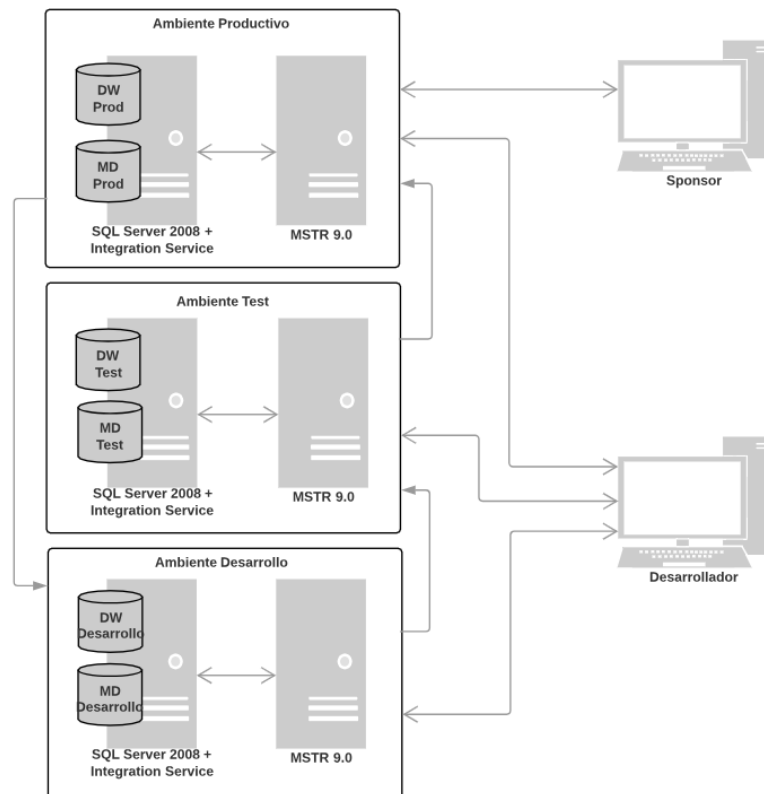


Imagen 8 – Esquemas de servidores

2.2 ETL

Conceptos:

Extract, Transform and Load («extraer, transformar y cargar», frecuentemente abreviado ETL) es el proceso que permite a las organizaciones mover datos desde múltiples fuentes, reformatearlos y limpiarlos, que luego serán cargados en otra base de datos, llamada Data Mart o Data Warehouse para su análisis, o en otro sistema operacional para apoyar un proceso de negocio.

Extraer

La primera parte del proceso ETL consiste en extraer los datos desde los sistemas de origen. La mayoría de los proyectos de almacenamiento de datos fusionan datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o ficheros planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

Una parte intrínseca del proceso de extracción es la de analizar los datos extraídos, de lo que resulta un chequeo que verifica si los datos cumplen la pauta o estructura que se esperaba. De no ser así los datos son rechazados.

Un requerimiento importante que se debe exigir a la tarea de extracción es que ésta cause un impacto mínimo en el sistema origen. Si los datos a extraer son muchos, el sistema de origen se podría ralentizar e incluso colapsar, provocando que éste no pueda utilizarse con normalidad para su uso cotidiano. Por esta razón, en sistemas grandes las operaciones de extracción suelen programarse en horarios o días donde este impacto sea nulo o mínimo.

Transformar

La fase de transformación aplica una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. Algunas fuentes de datos requerirán alguna pequeña manipulación de los datos. No obstante en otros casos pueden ser necesarias aplicar algunas de las siguientes transformaciones:

- Seleccionar sólo ciertas columnas para su carga (por ejemplo, que las columnas con valores nulos no se carguen).
- Traducir códigos (por ejemplo, si la fuente almacena una "H" para Hombre y "M" para Mujer pero el destino tiene que guardar "1" para Hombre y "2" para Mujer).
- Codificar valores libres (por ejemplo, convertir "Hombre" en "H" o "Sr" en "1").
- Obtener nuevos valores calculados (por ejemplo, $\text{total_venta} = \text{cantidad} * \text{precio}$, o $\text{Beneficio} = \text{PVP} - \text{Coste}$).
- Unir datos de múltiples fuentes (por ejemplo, búsquedas, combinaciones, etc.).

- Calcular totales de múltiples filas de datos (por ejemplo, ventas totales de cada región).
- Generación de campos clave en el destino.
- Transponer o pivotar (girando múltiples columnas en filas o viceversa).
- Dividir una columna en varias (por ejemplo, columna "Nombre: García López, Miguel Ángel"; pasar a tres columnas "Nombre: Miguel Ángel", "Apellido1: García" y "Apellido2: López").
- La aplicación de cualquier forma, simple o compleja, de validación de datos, y la consiguiente aplicación de la acción que en cada caso se requiera:
 - Datos OK: Entregar datos a la siguiente etapa (Carga).
 - Datos erróneos: Ejecutar políticas de tratamiento de excepciones (por ejemplo, rechazar el registro completo, dar al campo erróneo un valor nulo o un valor centinela).

Carga

La fase de carga es el momento en el cual los datos de la fase anterior (transformación) son cargados en el sistema de destino. Dependiendo de los requerimientos de la organización, este proceso puede abarcar una amplia variedad de acciones diferentes. En algunas bases de datos se sobrescribe la información antigua con nuevos datos. Los Data Warehouse mantienen un historial de los registros de manera que se pueda hacer una auditoría de los mismos y disponer de un rastro de toda la historia de un valor a lo largo del tiempo.

Existe una única forma de cargar los datos:

Rolling

El proceso de Rolling por su parte, se aplica en los casos en que se opta por mantener varios niveles de granularidad (jerarquías). Para ello se almacena información resumida a distintos niveles, correspondientes a distintas agrupaciones de la unidad de tiempo o diferentes niveles jerárquicos en alguna o varias de las dimensiones de la magnitud almacenada (por ejemplo, totales diarios, totales semanales, totales mensuales, etc.).

SQL Server Integration Service (SSIS)

Integration Services es básicamente una poderosa herramienta de SQL Server para realizar tareas tipo ETL (Extract Transform Load). Fue liberada por Microsoft por

primera vez junto con SQL Server 2005 y fue el reemplazo de DTS (Data Transformation Services). Ha estado presente desde entonces en todas las distribuciones de SQL Server.

SSIS ofrece una variedad de formas y utilidades para mover datos de un lugar a otro, por ejemplo si se requiere mover datos de una forma rápida de alguna fuente de datos tipo OLE DB (o compatible) SSIS ofrece el Asistente de Importación y Exportación con el cual no solo se pueden mover datos si no realizar transformaciones sobre ellos tales como convertir un tipo de datos en otro. Este asistente permite de forma rápida hacer revisiones de las tablas de origen y escribir consultas para optimizar la forma y cantidad de los datos que se quieren recuperar.

Pero SSIS es más que un asistente para mover datos, usando el ambiente BIDS (Business Intelligence Development Studio) que está basado en Visual Studio se pueden crear paquetes (colecciones de elementos de trabajo) con flujos de trabajo y tareas tan complejas como se quiera, además una de las ventajas que ofrece es que BIDS no está ligado a un servidor particular de SQL Server por lo que no se necesita estar conectado a uno para diseñar un paquete. En otras palabras se puede diseñar el paquete en modo desconectado y al terminar hacer la instalación de este en el servidor de SQL que se quiera.

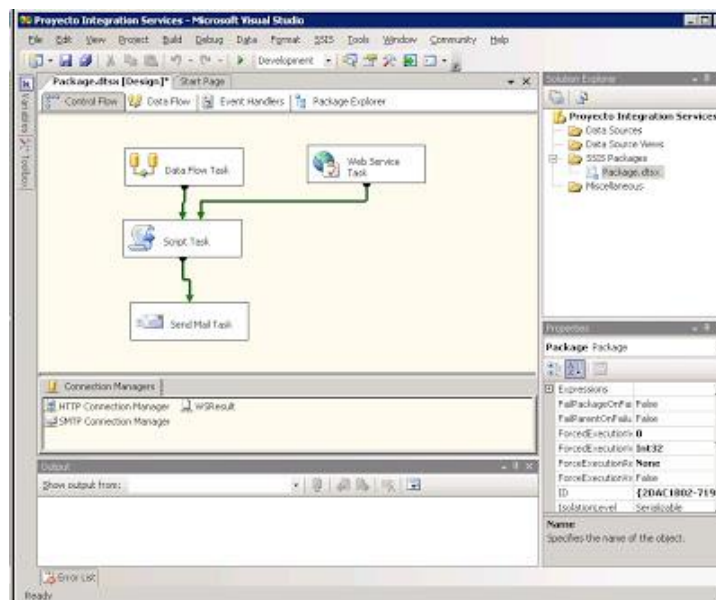


Imagen 9 - SQL Server Integration Service

En términos de arquitectura SSIS se divide en 4 componentes principales:

- El Servicio SSIS

- El Motor de ejecución de SSIS y otros ejecutables
- El Motor de Flujo de Datos de SSIS y los componentes de flujo de Datos
- Los Clientes SSIS

En más detalle el servicio SSIS maneja los aspectos operacionales de la herramienta, es básicamente un servicio de Windows que se instala junto con SQL Server y registra la ejecución de paquetes y tiene utilidades para el almacenamiento de los mismos. Realmente este servicio no es necesario para ejecutar paquetes SSIS pero si ya está corriendo y el servicio es detenido entonces todos los paquetes que se están ejecutando en ese momento se detienen por defecto.

El motor de ejecución y sus programas auxiliares, son los que realmente se encargan de correr los paquetes. Este motor es el que salva la distribución de los elementos dentro de BIDS, administra las bitácoras, depuración, configuración, conexiones y transacciones. Adicionalmente maneja los eventos de envío de correos o almacenamiento de bitácoras en la base de datos cuando un evento se dispara en el paquete. Los ejecutables auxiliares proveen otras funciones tales como: Contenedores, Tareas, Manejadores de Eventos y restricciones de precedencia.

El componente principal de un proyecto SSIS es el paquete, se puede definir como una serie de tareas que se ejecutan en un orden predeterminado, las restricciones de precedencia ayudan a definir en qué orden una tarea debe ejecutarse. Normalmente el paquete se salva como un archivo con extensión DTSX, que en esencia es un archivo XML. A su vez una tarea puede definirse como una unidad individual de trabajo, estas tareas proveen diferentes funcionalidades dentro del paquete, por ejemplo mover archivos, cargar un archivo en una base de datos, enviar un correo o incluso escribir código fuente (C# o VB.Net) de manera automática. SSIS viene con un número de tareas ya definidas que pueden ser utilizadas de inmediato, algunas de ellas son:

Bulk Insert (Inserción de volúmenes altos de datos): Carga datos en una tabla usando el comando BULK INSERT.

Data Flow (Flujo de datos): Es una de las más utilizadas, se usa para cargar y transformar datos dentro de un destino OLE DB.

Execute Package (Ejecución de Paquetes): Ejecuta paquetes dentro de otros paquetes, esto sirve para modularizar paquetes.

Execute Process (Ejecución de Procesos): Ejecuta un programa externo al paquete.

File System (Sistema de Archivos): Se encarga de operaciones tales como crear directorios, renombrarlos y borrarlos. También maneja operaciones sobre archivos tales como copiar o mover.

FTP (Protocolo de Transferencia de Archivos): Envía o recibe archivos desde servidores FTP.

Script: Ejecuta código fuente creado en Visual Studio (C# o VB.Net)

Send Mail (Envío de Correo): Envía correo a través de un servidor SMTP (Send Mail Transfer Protocol)

Analysis Services Processing (Procesamiento de servicios de análisis): Procesa cubos, dimensiones o minería de modelos de SQL Server Analysis Services.

Web Service (Servicios Web): Ejecuta métodos de servicios web.

WMI Data Reader (Lectores de Datos WMI): Con esta tarea se pueden ejecutar consultas WQL contra WMI (Windows Management Instrumentation). De esta forma se puede leer información de las bitácoras de eventos que guarda Windows u obtener información del hardware instalado por ejemplo.

WMI Event Watcher (Captura de Eventos WMI): Con esta tarea SSIS es capaz de esperar y responder por ciertos eventos que ocurren en el sistema operativo.

XML: Parsea o procesa archivos XML.

Estas son solo algunas de las tareas disponibles, algunas de ellas son contenedores de más elementos que se pueden utilizar para operaciones específicas, tal es el caso de las tareas de flujo de datos. También hay otras tareas que se pueden bajar e instalar o incluso utilizando lenguajes como C# o VB.NET y el modelo de objetos de SSIS se pueden crear nuevas tareas.

En síntesis SSIS es una de las herramientas más innovadoras de migración de datos que han aparecido en el mercado, si ya se tiene una licencia de SQL Server no tiene costo pues se incluye como parte del paquete. Con un poco de imaginación y conocimiento, SSIS puede utilizarse en un sin fin de tareas interesantes.

2.3 Microstrategy

¿Qué es MicroStrategy?

MicroStrategy [27], es una compañía que ofrece software OLAP, una solución utilizada en el campo de la llamada Inteligencia de Negocio (o Business Intelligence) cuyo objetivo es agilizar la consulta de grandes volúmenes de datos.

El software de MicroStrategy permite crear informes y analizar los datos almacenados de diferentes fuentes. MicroStrategy define el núcleo de su software de informes como un sistema “ROLAP” (Relational Online Analytical Processing) u “OLAP Relacional”, en los que normalmente los datos son detallados evitando las agregaciones, estando formada la estructura de almacenamiento por tablas des normalizadas. Los modelos de datos más comunes sobre los que se trabaja son, por su facilidad de uso y diseño intuitivo, los dimensionales (modelo de estrella o copo de nieve), aunque es posible trabajar sobre cualquier otro modelo relacional.

Este software también soporta tecnología MOLAP (Multi-dimensional Online Analytical Processing) desde la versión 7i. Por lo general, el procesamiento MOLAP ofrece un mayor rendimiento en las consultas gracias a su sistema de indexación avanzado y la tecnología de compresión de datos que emplea, lo que le confiere unas necesidades de almacenamiento inferiores en comparación con las soluciones ROLAP, siendo esta una gran ventaja competitiva.

Parte de la fuerza de Microstrategy radica en que su producto está formado por una plataforma integrada de herramientas de Inteligencia de negocio, y no por diferentes productos aislados agregados para formar una suite. Además, cabe destacar su gran escalabilidad, capacidad de adaptación y respuesta a medida que aumenta de forma significativa el número de usuarios y/o volumetría de datos con la que estemos trabajando.

La versión de software más reciente se llama MicroStrategy 10. Esta versión incluye importantes mejoras respecto a versiones anteriores, proporcionando a los usuarios la capacidad de construir sus propios cuadros de mando en cuestión de minutos mediante la herramienta de creación de informes ad-hoc Visual Insight. Este producto le aporta a los usuarios la capacidad de explorar los datos de forma intuitiva y totalmente visual, siendo en sí misma una herramienta orientada a facilitar la toma de decisiones y búsqueda de oportunidades de negocio.

Componentes de Microstrategy

Los elementos más importantes que forman la herramienta de MicroStrategy son los siguientes:

- MicroStrategy metadata: repositorio que almacena la definición de los objetos de MicroStrategy y la información sobre el modelo de base de datos explotado.
- MicroStrategy Intelligence Server: servidor analítico optimizado para el reporting empresarial y para el análisis OLAP.
- MicroStrategy Desktop: aplicación en entorno Windows que proporciona un completo abanico de funciones analíticas diseñadas para facilitar el desarrollo de informes.
- MicroStrategy Web: entorno interactivo para la ejecución de informes y análisis desde el entorno web.

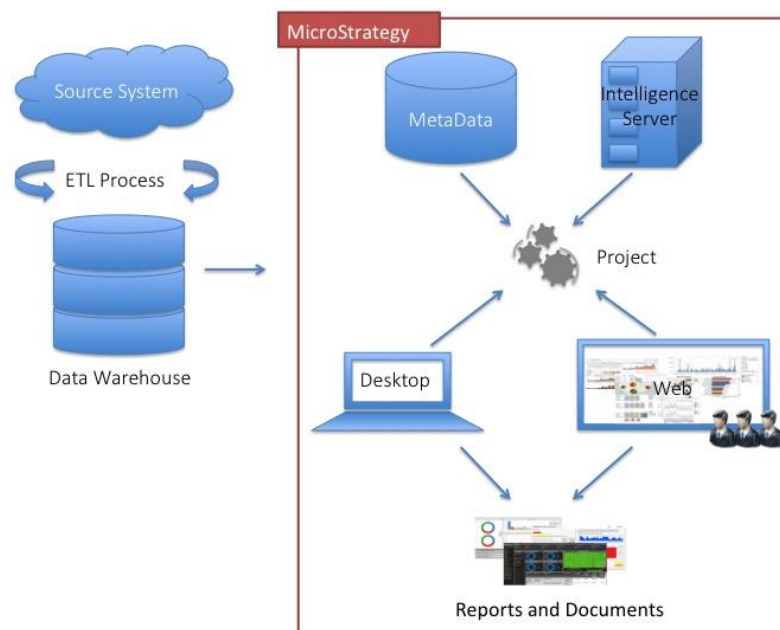


Imagen 10 – Componentes

2.3.1. Conceptos y terminología

Modelización de Datos.

Es necesario definir varios términos para poder entender cómo se configurarán conceptual y operacionalmente los datos para el soporte de la aplicación Microstrategy Agente.

Modelo de Datos.

Un modelo de datos es un diagrama lógico de las entidades de una organización. Contiene sólo aquellas entidades que serán incluidas en el Proyecto de Microstrategy y sirve como base para el diseño del Data Warehouse.

Los componentes principales de un modelo de datos son:

- Hechos
- Atributos
- Elementos de Atributos
- Jerarquías de Atributos
- Relaciones entre Atributos

Hechos

Cantidades, medidas o variables de negocio que serán base para hacer métricas. Las métricas son cálculos analíticos realizados a partir de hechos que se utilizan en los informes para analizar el rendimiento de un negocio.

Las métricas son los resultados o medidas, normalmente de naturaleza numérica, que representan el objeto de la investigación en la Toma de Decisiones. Ejemplos de métricas incluyen Ventas, Ingresos.

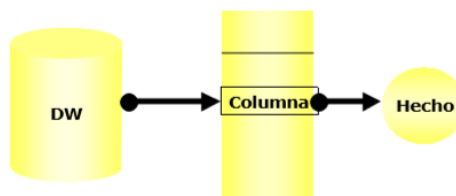


Imagen 11 – Hecho

Para el desarrollo del tablero Procedimientos Compras, se crean los hechos, Importe y Cantidad.

Atributos

Los atributos representan entidades en el modelo de negocio. Se trata de conceptos que agrupan datos que responden al mismo tipo de información. Por ejemplo, el atributo Año, engloba los datos: 1997, 1998, 1999, 2000 etc. El atributo País engloba: España, Francia, Italia, etc.

Los atributos tienen dos funciones principales:

- 1.- Definen el nivel de agregación al que los hechos deben ser presentados en el informe.
- 2.- Condicionan el conjunto de resultados presentados en un informe, esto es, pueden ser parte de las condiciones del filtro de un informe.

Para el desarrollo del tablero Procedimientos Compras, se crean los atributos Año, Trimestre, Mes, Procedimiento Compra, Estado OC, Jurisdicción, Objeto del Gasto.

Jerarquías de Atributos

Son grupos de atributos que están relacionados. Normalmente, en las jerarquías aparecen todos los atributos de una misma familia con sus relaciones padre-hijo.

Relaciones entre Atributos

Los atributos de un proyecto están relacionados entre ellos a través de relaciones padre-hijo que vienen definidas por los elementos o los valores de cada uno.

Para conocer las relaciones es útil pensar en los atributos, en grupos lógicos o jerarquías, pero la definición de relaciones surge de la relación existente en el esquema definido del proyecto.

Tipos de relaciones

- Uno a uno. Ejemplo: ID de cliente y DNI
- Uno a muchos. Ejemplo: Un pedido con un ID de cliente
- Muchos a muchos. Ejemplo: ID de cliente y Fecha de pedido

Indicadores

Los indicadores son objetos de Microstrategy que representan medidas de negocio calculadas sobre la base de los objetos almacenados en la Base de Datos.

Los indicadores pueden estar basados tanto en fórmulas como en niveles:

- Fórmula: es una expresión matemática - como suma o media - aplicada a hechos, atributos u otras métricas

- Nivel o dimensionalidad: se corresponde con el nivel de cálculo para los indicadores, como un año o un cliente.

Para crear un nuevo indicador lo único imprescindible es establecer la fórmula o definición del indicador.

Informes

Un informe es un objeto de Microstrategy que representa una respuesta para datos específicos procedentes del DW. Permiten obtener información de negocio a través del análisis de datos. Un informe contiene todas las métricas y atributos especificados en él que cumplen todas las condiciones enumeradas en el filtro. Por ejemplo, si se desea los ingresos por región a nivel mensual con el filtro del año 2.000 se obtendrá el informe mostrado a continuación en modo cuadrícula.

Filtros

Como se dijo anteriormente, los filtros permiten especificar las condiciones que los datos deben cumplir para ser incluidos en un informe. Se tratan de condiciones combinadas mediante operadores que se utilizan para reducir la cantidad de información que aparece en el informe.

Se pueden establecer condiciones sobre atributos y sobre indicadores.

2.3.2. Introducción a Microstrategy Agente

Microstrategy Agente es el interfaz que permite obtener información del Data Warehouse a través de la creación de informes en los diferentes proyectos que tengamos creados.

Tipos de Objetos dentro de un proyecto

Una vez dentro de un proyecto de Agente, nos encontramos con tres tipos de objetos con los que vamos a trabajar:

- Objetos Personales
- Objetos Públicos
- Objetos de Esquema
- Explorador de Datos

Objetos Personales

Dentro de los objetos personales cada usuario puede crear carpetas para organizar todos sus objetos (informes, métricas, filtros...) Todos los objetos que aparezcan dentro de estas carpetas sólo pueden ser visualizados por el usuario que los creó.

Objetos Públicos

Los objetos públicos son aquellos que pueden ver todos los usuarios, aunque no pueden ser modificados ni borrados a no ser que se tenga permiso específico para hacerlo. Estos objetos están agrupados en las siguientes subcarpetas por tipo de objeto

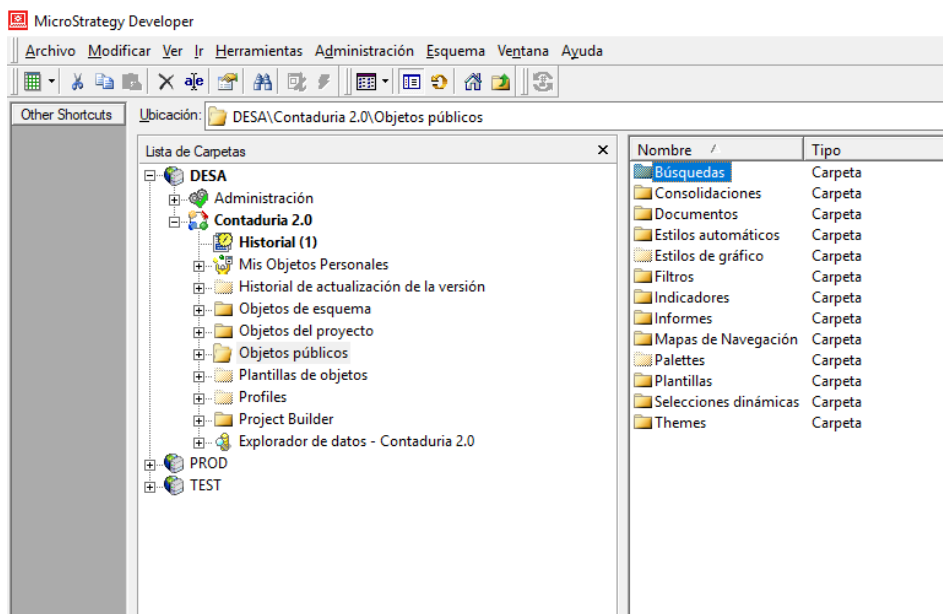


Imagen 12 – Microstrategy Desktop

Objetos de Esquema

Esta carpeta recoge todos los componentes cuya creación, modificación o eliminación afecta al Metadata del proyecto.

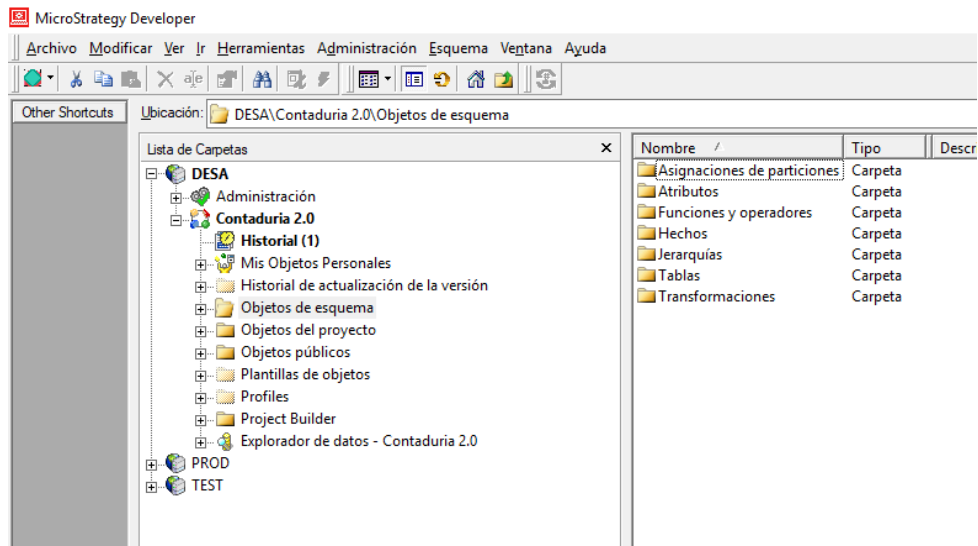


Imagen 13 - Objetos de Esquema

Explorador de Datos

El explorador de Datos contiene todos los atributos que se han definido en los proyectos agrupados en jerarquías.

Contiene la jerarquía del sistema y todas las jerarquías que se han definido en el proyecto:

La jerarquía del sistema se crea automáticamente y en ella aparece una lista con todos los atributos del proyecto. El resto de las jerarquías se crean según las necesidades y preferencias de los usuarios. Las jerarquías son simplemente una forma de agrupar los atributos de forma que el usuario pueda navegar en la información de forma sencilla.

Capítulo 3 – Implementación de la solución Business Intelligence

3. Introducción

En el capítulo 1 se realizó una breve descripción de Business Intelligence, mencionando las definiciones de los grandes autores que hablan del tema, donde explican y fundamenta la utilización de la metodología de trabajo.

En el capítulo 2, se realizó una descripción de los ambientes de trabajo que se utilizaron para llevar a producción el tablero y así poder ser utilizado por el sponsor que lo requirió. También, explique las herramientas que son utilizadas en todo el proceso de desarrollo, para llevar adelante la creación del Data Warehouse y el Dashboard con Microstrategy. Por último, se menciona la necesidad de una terminal como Chrome, que nos permita consumir el tablero.

En este capítulo, se explicará cómo llevar adelante la planificación del proyecto, cuales son los pasos de desarrollo, como también, las tareas que se van realizando hasta llegar al producto final.

El SIGAF comenzó a funcionar en el año 2005, basándose en la “Ley 70, Decreto reglamentario N° 1000/GCBA/99”, Ley de Gestión, Administración Financiera y Control del Sector Público de la Ciudad [26]. Este sistema apoya la gestión presupuestaria, financiera y contable de los Ministerios del Gobierno Central, así como las funciones de rectoría y fiscalización que corresponden al Ministerio de Hacienda y a los Órganos de Control.

SIGAF, está compuesto por 1148 Tablas relacionadas, ocupando un espacio en disco de aproximadamente de 100 GB, variando día a día, por su cantidad de operaciones.

DGUIAF (Dirección General Unidad Informática de Administración Financiera Ministerio de Hacienda), en el evento **I Jornada "La Administración Financiera de la CABA a 15 años de su autonomía"** [25], presentó los motivos que llevaron a tener la necesidad de construir sus Tableros de Control.

A medida que fueron pasando los años, SIGAF comenzó a tener mayor solides en la gestión de la Administración Financiera del Gobierno de la Ciudad de Buenos Aires, llegando a sumar nuevos módulos, como por ejemplo el módulo de Compras, Contrataciones Personales, Obras Públicas, Proveedores, entre otros. A este crecimiento como sistema, llegó en el año 2010, la necesidad de implementar su Tablero de Control, que permitiera tomar decisiones y realizar un seguimiento de la gestión.

Antes de la implementación del Tablero de Control, los datos obtenidos provenían de distintas fuentes de datos, esto provocaba no tener un control sobre ellos, generando que estuvieran repetidos, no sabiendo su veracidad. Los reportes que se generaban, contenían información escasa, limitada, siendo básicos para tomar decisiones. La Imagen 14 – Situación Inicial [25], refleja el estado de los datos, al momento de consultarlos.

Situación Inicial

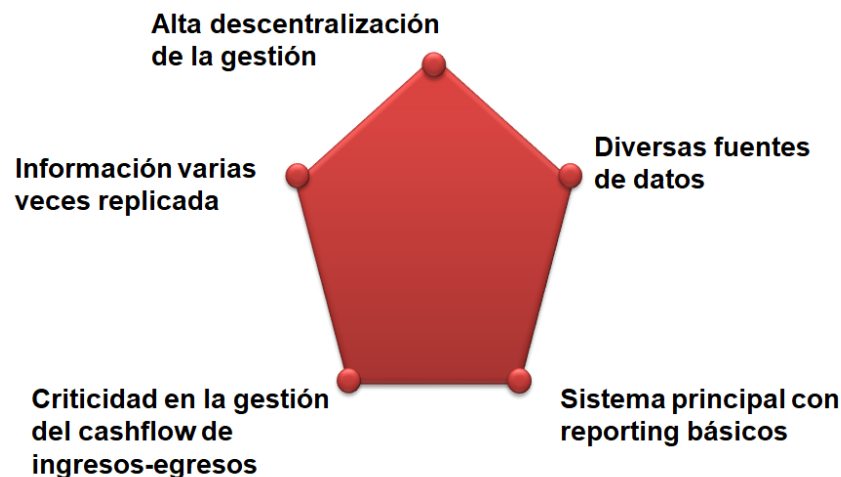


Imagen 14 – Situación Inicial

Por lo tanto, para realizar un reporte el mecanismo que se utilizaba era mediante una hoja de cálculo en Excel. Su forma de operar, requería de varias personas trabajando en distintas oficinas y espacio físicos, generando una hoja de cálculo, para luego juntarlas y obtener un único resultado. Esta tarea, implicaba contar con varias personas de dedicación exclusiva, trabajando a contra reloj, cada uno de los días.

Situación Inicial



Imagen 15 – Hoja de Calculo

Es así, que se decide mejorar la forma de trabajar, buscando optimizar los tiempos de entrega de los informes y del trabajo del personal. Al personal le implicaba muchas horas del día recolectar todos los datos, ese mismo tiempo, podría ser utilizado en otras tareas. Entonces, nació la idea de crear un Data Warehouse.

Su ventaja principal es que los datos son almacenados en un mismo repositorio, como ilustra la Imagen 16 – Dato Centralizado, repositorio, antes de ser presentados en reportes. La centralización de los datos ayuda a minimiza los errores, como también la del personal para su mantenimiento. Como resultado principal se logra presentar los informes en un menor tiempo, en conjunto con datos de calidad.

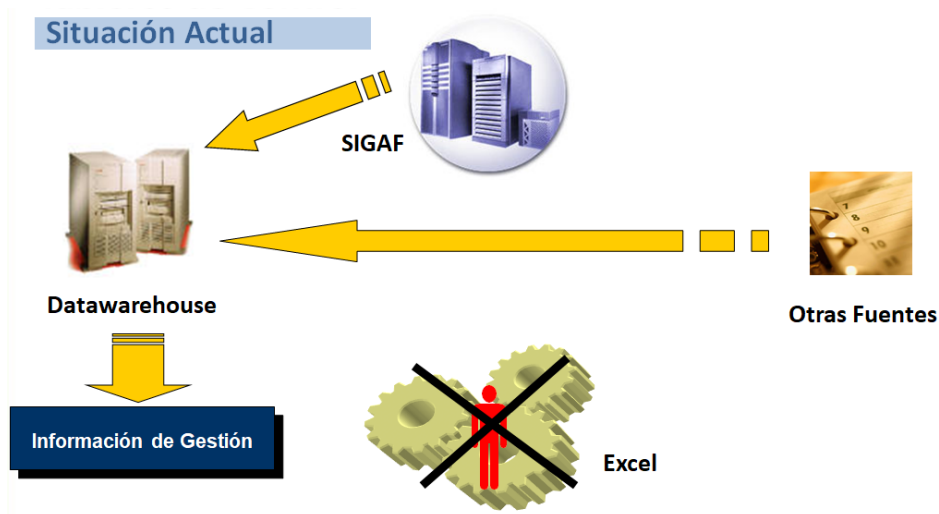


Imagen 16 – Dato Centralizado.

En resumen, la implementación de Tablero de Control se realiza por la obtención de beneficios, como ilustra la Imagen 17 – Beneficios Obtenidos.



Imagen 17 – Beneficios Obtenidos.

A continuación se describirá el desarrollo del Tablero de Control realizado para el departamento de Compras y Rendiciones.

3.1 Planificación del Proyecto Delimitación de Alcance del Proyecto.

Área: Departamento Compras y Rendiciones

Objetivo: Determinar con respecto al tiempo las compras de cada organismo por Objeto del Gasto.

3.2 Definición de Requerimientos del Negocio

Requerimiento Funcional “Se requiere conocer las compras que están realizando las Jurisdicciones desde varios enfoques, para lo cual se necesita un tablero de control que visualice las operaciones y así compararlo contra las compras de años anteriores.”

Requerimiento Técnico “Se necesita adquirir un sistema de Business Intelligence que pueda ser útil para presentar una gran diversidad de Informes a la Dirección General.”

Como ejemplo, se muestra la Imagen 18 – Datos.

Año	Trimestre	Mes	Jurisdicción	Objeto del Gasto	Procedimiento Compra	Estado	Cantidad	Importe
-----	-----------	-----	--------------	------------------	----------------------	--------	----------	---------

Imagen 18 - Datos

3.3 Modelado Dimensional de Negocio (BDM)

El Modelo Dimensional de Negocio debe de ser analizado con detenimiento, es importante conocer el modelo de la fuente de datos, porque a partir de él se formulan las interrogantes y se dan respuestas a ellas.

Matriz Dimensional

Es importante realizar una matriz dimensional ya que con esta se puede observar la relación que existe entre los Hechos y las Dimensiones. En la Imagen 19 se observa que los hechos de Importe de Compra y Cantidad, se pueden relacionar mediante las dimensiones de Fecha, Jurisdicción, Objeto del Gasto, Procedimiento Compra, Estado.

	Dimensiones				
Hechos	Fecha	Jurisdicción	Objeto del Gasto	Procedimiento compra	Estado
Importe	x	x	x	x	x
Cantidad	x	x	x	x	x

Imagen 19 – Matriz Dimensional

El BDM que se propone en este trabajo está enfocado en el área de compras, los hechos, cantidad e importe de las compras, están definidos en la dimensión Compra, la cual es el centro del modelo estrella, por lo que se puede conocer la cantidad e importe de las compras por Jurisdicción, Fecha, Objeto del Gasto. El modelo se muestra en la Imagen 20 – Modelo Data Warehouse.

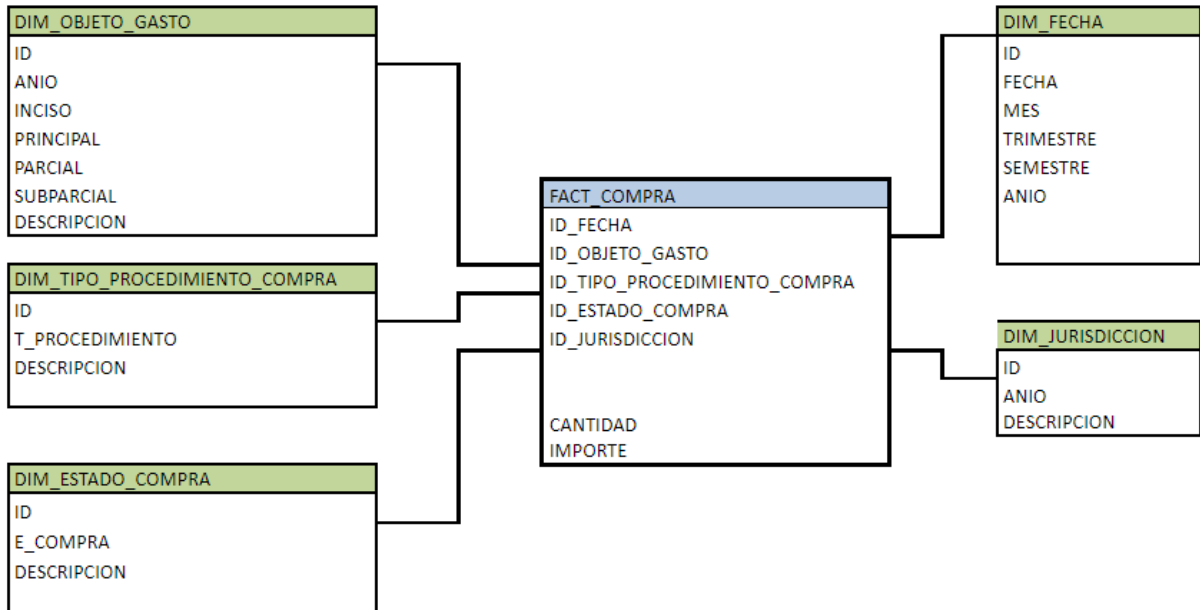


Imagen 20 – Modelo Data Warehouse

3.4 Diseño Físico

Después de conocer los requerimientos y el modelo de negocio, se continúa con el diseño físico del modelo de datos. Para construir el modelo de la base de datos o mejor llamado, Data Warehouse, se utiliza el sistema de gestión de datos, Microsoft SQL Server 2008.

Al momento de definir los nombres de las tablas que componen el Data Warehouse, se utilizan prefijos que ayudan a entender rápidamente si estamos leyendo una tabla de Hecho o de Dimensiones. Para las tablas de Hechos, se indica con el prefijo FACT, el prefijo DIM indica que es una tabla de dimensión.

A continuación se describe la estructura de cada una de las tablas, graficadas en la Imagen 20 – Modelo Data Warehouse. La estructura de la tabla está compuesta por el nombre del campo, el tipo de dato que almacena y si permite ser null su valor. Primero se mencionan las tablas de Dimensiones y por último la tabla de Hecho.

DIM_OBJETO_GASTO. La dimensión Objeto del Gasto, contiene el listado de los diferentes conceptos por los cuales se realizan las compras.

DIM_OBJETO_GASTO

COLUMN_NAME	DATA_TYPE	NULLABLE
ID	NUMBER	NO
ANIO	NUMBER	NO
C_INCISO	NUMBER	NO
C_PRINCIPAL	NUMBER	NO
C_PARCIAL	NUMBER	NO
C_SUBPARCIAL	NUMBER	NO
DESCRIPCION	VARCHAR(100)	NO

Imagen 21 –Tabla Dimensión Objeto del Gasto

DIM_TIPO_PROCEDIMIENTO_COMPRA. La dimensión Procedimiento de Compras, contiene el listado de los diferentes procesos, por los cuales se puede llevar adelante la compra, como por ejemplo, Licitación Pública, Licitación Privada, Convenio Marco, etc.

DIM_TIPO_PROCEDIMIENTO_COMPRA

COLUMN_NAME	DATA_TYPE	NULLABLE
ID	NUMBER	NO
T_PROCEDIMIENTNO	VARCHAR(30)	NO
DESCRIPCION	VARCHAR(50)	NO

Imagen 22 – Tabla Dimensión Procedimiento de Compras

DIM_ESTADO_COMPRA. La dimensión Estado de la Compra, contiene los diferentes estados por los cuales va pasando la compra durante su proceso de ejecución, como por ejemplo, Perfeccionada, Cumplida.

DIM_ESTADO_COMPRA

COLUMN_NAME	DATA_TYPE	NULLABLE
ID	NUMBER	NO
E_COMPRA	VARCHAR(1)	NO
DESCRIPCION	VARCHAR(20)	NO

Imagen 23 – Tabla Dimensión Estado de la Compra

DIM_FECHA. La dimensión fecha, contiene la lista de todas las fechas a partir del día 01/01/2000, con sus clasificadores del mes, trimestre, año.

DIM_FECHA

COLUMN_NAME	DATA_TYPE	NULLABLE
ID	NUMBER	NO
FECHA	DATE	NO
MES	NUMBER	NO
TRIMESTRE	NUMBER	NO
SEMESTRE	NUMBER	NO
ANIO	NUMBER	NO

Imagen 24 – Tabla Dimensión Fecha

DIM_JURISDICCION. La tabla de dimensión Jurisdicción, contiene la lista de todas las Jurisdicciones perteneciente a la Provincia de Buenos Aires a lo largo de los años.

DIM_JURISDICCION

COLUMN_NAME	DATE_TYPE	NULLABLE
C_ID	NUMBER	NO
ANIO	NUMBER	NO
DESCRIPCION	VARCHAR(50)	NO

Imagen 25 – Tabla Dimensión Jurisdicción

Hasta acá se describieron la estructura de las tablas de Dimensiones, a continuación se describe la estructura de la única tabla de Hecho definida para este modelo.

FACT_COMPRA. La tabla de hecho Fact_Compra, siendo la tabla central del modelo estrella del Data Warehouse, contiene los ID de todas las Dimensiones, junto con sus indicadores, Cantidad e Importe.

FACT_COMPRA

COLUMN_NAME	DATA_TYPE	NULLABLE
ID_FECHA	NUMBER	NO
ID_OBJETO_GASTO	NUMBER	NO
ID_TIPO_PROCEDIMIENTO	NUMBER	NO
ID_ESTADO_COMPRA	NUMBER	NO
ID_JURISDICCION	NUMBER	NO
CANTIDAD	NUMBER	NO
IMPORTE	NUMBER(18,4)	NO

Imagen 26 - Tabla de Hecho Compra

3.5 Diseño y Desarrollo de Presentación de Datos

Con el desarrollo del diseño del modelo Estrella elegido para nuestro modelo de Data Warehouse, se procede a explicar cómo se implementa el código de cada ETL mediante la herramienta SQL Server Integration Services (SSIS, herramienta en el capítulo anterior), quien se encargaran de insertar los datos en cada una de las tablas del modelo propuesto.

3.5.1 Desarrollo ETL

Luego de instalar el complemento de Visual Studio, se crea un nuevo proyecto del tipo “Business Intelligence/Integration Services Project”, llamando al proyecto: Procedimiento Compra

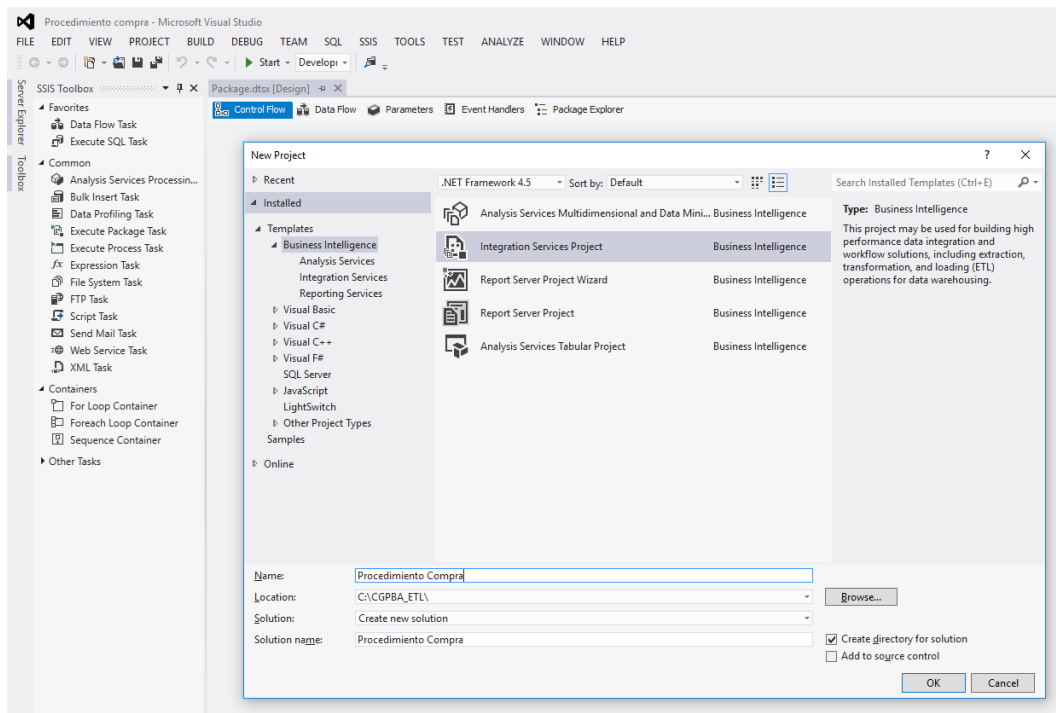


Imagen 27 – Visual Studio

Una vez creado el Proyecto, dentro de él, va a contener varios Paquetes, cada uno con una lógica diferente de acuerdo a su función.

Se crearan tres tipos de paquetes, el primero se le dará el nombre de ‘Carga Diaria’, el siguiente, cargara las tablas del tipo DIM, y por último, están los paquetes que cargan las tablas FACT.

El paquete ‘Carga Diaria’ es llamado así, porque da el inicio a la ejecución de todos los paquetes restantes, tanto del tipo DIM, como del tipo FACT. Como así también,

brinda información del estado de la ejecución de cada uno de ellos, avisando si terminó exitosamente o con fallas. En caso de existir una falla, se crea una bifurcación que informa de ese estado y dependiendo del tipo de falla, permite continuar con la ejecución de los demás paquetes. A continuación se explica con más detalle su implementación.

3.5.2 ¿Cómo funciona el paquete ‘Carga diaria’?

Cabe destacar, que este proceso de carga diaria, ya se encontraba desarrollado e implementado por un grupo de personas, antes de mi ingreso en la DGUIAF.

La primera tarea es registrar en una tabla llamada PROC_Procesos, la lista de procesos a ejecutar. Los nombres de los procesos se almacenan en la tabla PROC_Lotes, donde se corresponden uno a uno, con cada una de las tablas de hechos (FACT) y las tablas de dimensiones (DIM). Por ejemplo, se definió la tabla DIM_JURISDICCION, donde el proceso que carga los datos en ella, se llama ETL_DIM_JURISDICCION, siendo este nombre el mismo que se almacena en la tabla PROC_Lotes.

Por cada ejecución de una tarea, existe un bloque de instrucciones a intentar (try), marcado por un flujo de color verde, como también, una respuesta si se produce una excepción (catch), marcado por el flujo de color rojo.

En el principal contenedor ‘**Loop de procesos**’, se obtiene el nombre del paquete a ejecutar, se ejecuta el mismo y se vuelve a repetir el loop hasta terminar con todos los paquetes. Siempre controlando si la ejecución fue exitosa o fallida.

La Imagen 28 - Paquete Carga Diaria, ilustra el desarrollo del paquete ‘Carga Diaria’.

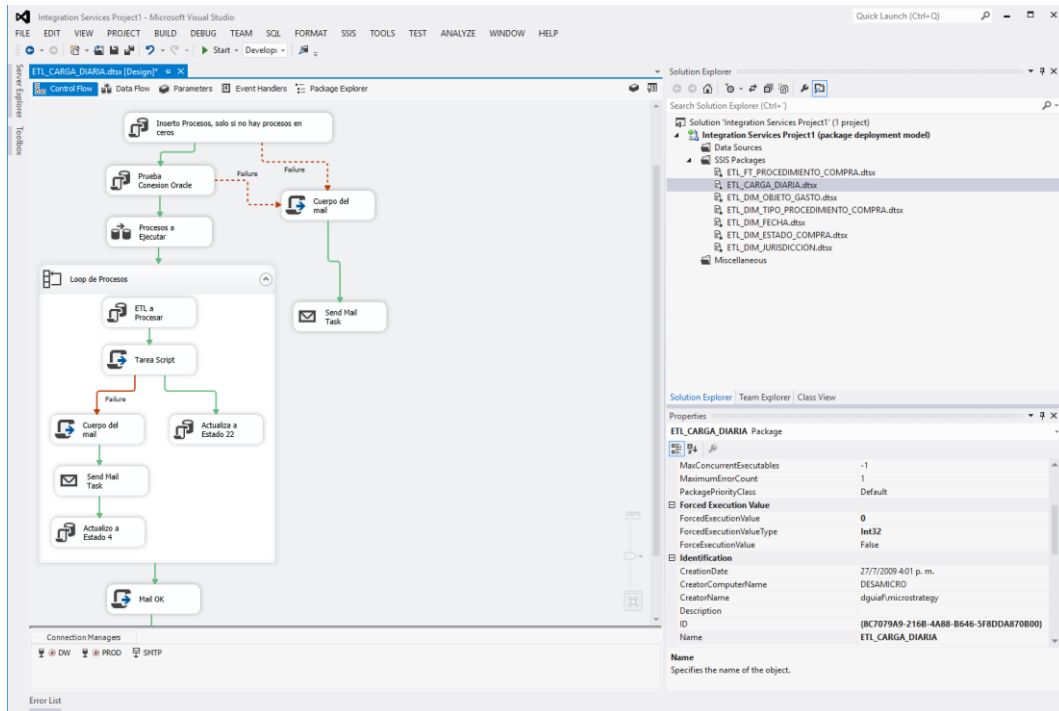


Imagen 28 - Paquete Carga Diaria

3.5.3 Proceso de carga tabla de la DIM_OBJETO_GASTO

El proceso DIM_OBJETO_GASTO, se encarga de llenar la tabla que lleva el mismo nombre que el proceso.

Este proceso trabaja de manera que borra y carga los datos de la tabla DIM, cada vez que el proceso se ejecuta. En la Imagen 29, se visualiza un contenedor de tareas llamado 'Contenedor de Secuencia'. Dentro de él se ven tareas relacionadas con el seguimiento de ejecución del proceso de carga de la tabla, como son las tareas 'Proceso Id' y 'actualiza a Estado 1', como también tareas relacionadas con la lógica funcional del ETL, que son 'truncate table DIM_OBJETO_GASTO' Y 'Cargo table DIM_Objeto_Gasto'.

Para una mejor descripción del proceso, en la imagen se detallan 6 pasos, que cumplen la siguiente función.

El paso 1 y 2, son tareas que hacen al control de la ejecución del proceso. El paso 1, obtiene el id del proceso cargado en la tabla PROC_Procesos, que luego será seteado con un numero en el Paso 2, marcando que el proceso está en un estado de ejecución. Los números utilizados para marcar el estado del proceso son 1 (se está ejecutando), 2 (finalizo correctamente), 4 y 12 (finalizo por algún tipo de error).

El paso 3 y 4, son tareas más específicas de un ETL. El paso 3, trunca los datos de la tabla que luego serán cargados en el paso 4. Dependiendo del tipo de finalización

del proceso, en el paso 5, se indica que termino bien o con error la carga de los datos, en la tabla PROC_Procesos.

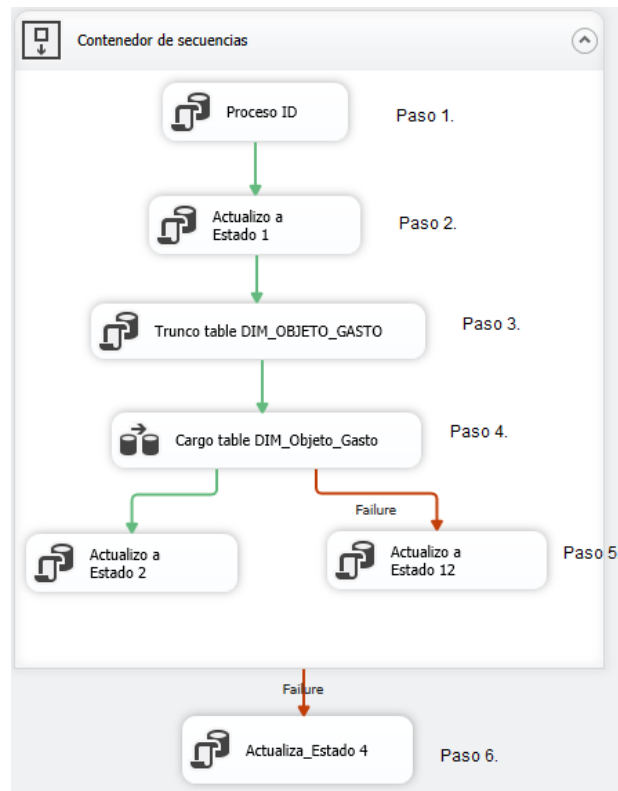


Imagen 29 - Carga tabla DIM_OBJETO_GASTO

En la sección queries (**ver Anexo Query, página 59**), se muestran las query más relevantes del proceso.

Los siguientes procesos que cargan las tablas del tipo Dimensión (Dim_Estado_Compra, Dim_Tipo_Procedimiento_Compra, Dim_Jurisdiccion), mantienen la misma lógica desarrollada para la carga explicada anteriormente, buscar el ID del proceso para setear su estado de ejecución, trunco la tabla y cargarla. El proceso de carga para ellas es diario, los procesos están programados para correr todos los días, a excepción, del proceso ETL_DIM_FECHA.

La carga de la tabla DIM_FECHA, a diferencia de la carga de las demás tablas, se optó por realizar su carga de datos una sola vez. Al ser una tabla con datos que no varían en el tiempo, no hace programarla para que se ejecute diariamente. Esta tabla con tiene datos del tipo fecha, siendo tomados entre los siguientes días, 01/01/2000 y 31/12/2100.

3.5.4 Proceso de carga tabla de la FACT_COMPRA

Unos de los procesos más importantes dentro de este desarrollo es la carga de la tabla FACT_Compras. Esta tabla contiene el resultado final de los datos procesados por los demás ETL. Todos los datos cargados en las tablas Dimensiones, cruzado con la tabla principal del sistema SIGAF, para este desarrollo, se verán reflejados en esta FACT.

No solo es importante la tabla por ser cruzada con las demás tablas de dimensiones, sino que también es la tabla la cual a partir de ella, se van a crear todos los objetos necesarios para luego ser mostrados en el tablero. Sus datos son evaluados y probados rigurosamente, validando su veracidad.

Como muestra la Imagen 30, el desarrollo del proceso es similar a los procesos de carga de las tablas de dimensión, siguiendo sus misma lógicas, se busca su id de ejecución en la tabla PROC_Procesos, se setea su estado en Ejecución, se trunca la tabla diariamente y luego es cargada con los datos actualizados, luego informa su estado de ejecución siendo exitoso o no.

El query más relevante del proceso, que se ejecuta en la tarea 'Cargo FT_PROCEDIMIENTO_COMPRA', reflejado en la imagen, se encuentra en la sección Anexo Query (**ver Anexo Query, página 59**), query N°6.

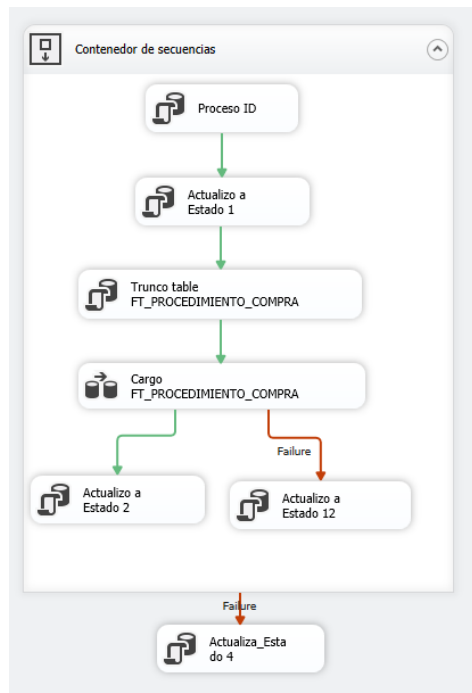


Imagen 30 - Carga tabla FT_COMPRA

3.6 Desarrollo Tablero – Microstrategy

Una vez desarrollado el modelo de DWH, y habiendo cargado los datos en cada una de las tablas mediante sus ETL, se procedió a probar y analizar los datos en búsqueda de comprobar la veracidad.

Paso siguiente, se comienza a trabajar en la última etapa, con la herramienta Microstrategy. Se procede a crear todos los objetos necesarios para la implantación del tablero. Para ello, se necesita crear, atributos, jerarquía, hechos, indicadores, informes.

A continuación, se explican cada uno de los objetos mencionado.

Creación de Atributos

Todo atributo está definido por al menos una forma, que es una representación lógica del atributo.

Comúnmente un atributo suele tener forma Descripción y forma ID (que suele ser designada como clave primaria).

Ejemplo: queremos crear el atributo “Jurisdicción” en nuestro proyecto.

Para crear un atributo vamos a utilizar el editor, de modo que nos pondremos en la parte derecha del agente y botón derecho del ratón “Nuevo Atributo”.

En este punto elegiremos la tabla de origen donde aparece la columna que nos interesa establecer como ID del atributo. En nuestro caso seleccionaremos la tabla DIM_JURISDICCION donde sabemos que se encuentra la variable ID que es la que vamos a utilizar como forma ID del atributo.

Una vez elegido el ID del atributo, Microstrategy nos enseñará la expresión a la que corresponde, las tablas de origen donde aparece esta variable y demás información sobre el mismo.

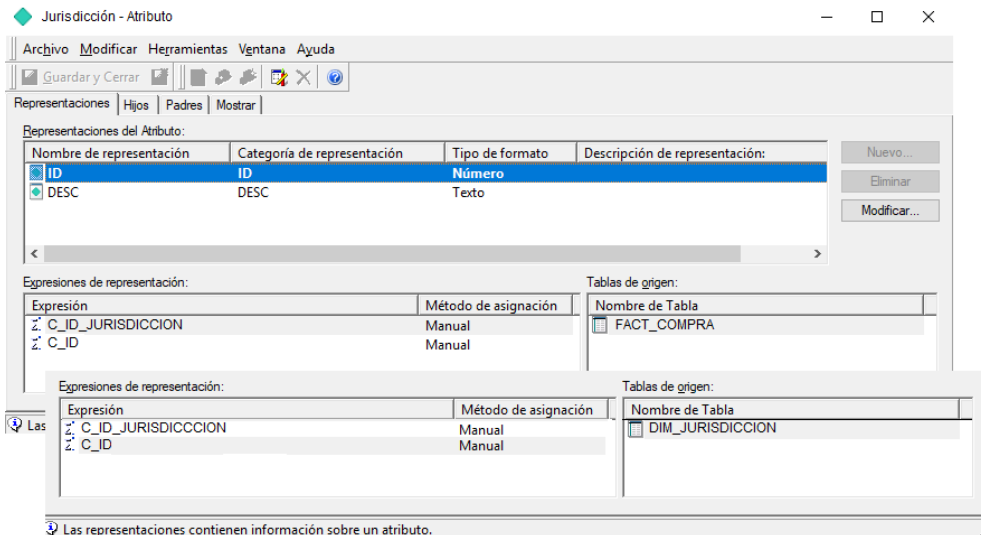


Imagen 31 - Creación de Atributos

Creación de un indicador simple

Un indicador simple es aquel que está formado por un operador como SUM, MAX, etc. y una o varias columnas de hechos, es decir, estas métricas son las que se definen directamente accediendo a columnas de las tablas de hechos. Para crear un indicador simple hay que seguir los siguientes pasos:

- 1.- Menú Archivo o botón derecho del ratón: Nuevo Indicador
- 2.- Seleccionar la plantilla Indicador vacío.
- 3.- El editor de indicadores presenta abierta la carpeta de hechos, a partir de ellos se hará la métrica.
- 4.- Seleccionar el hecho y arrastrar a la parte de definición de la métrica. La herramienta incorpora directamente el operador SUM. Este operador puede ser cambiado por cualquier otro disponible en la carpeta de operadores.
- 5.- Guardar y cerrar. Dar un nombre a la nueva métrica.

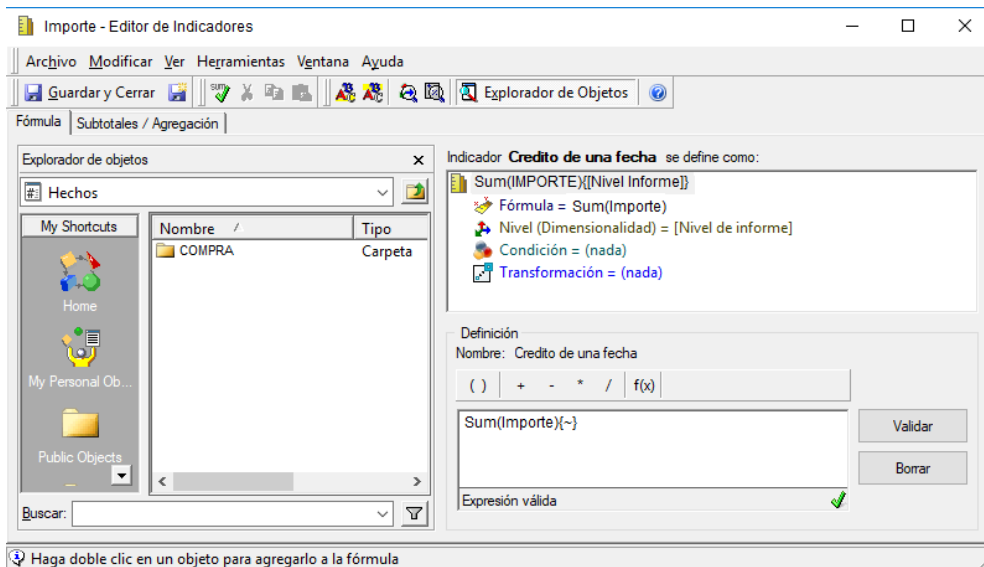


Imagen 32 - Creación de un indicador simple

3.7 Pruebas de Calidad de Datos

Hasta aquí se explicó cuáles fueron los requerimientos solicitados para la creación del Tablero de Control. Como también, las distintas etapas que se llevaron adelante para la creación del Data Warehouse. Se diseñó el modelo de datos, en conjunto con cada uno de los procesos de ETL que luego van a almacenar los datos. Finalizada la construcción del modelo de Data Warehouse y sus procesos de carga, se continúa con la comprobación de su funcionamiento, para probar que cada uno de los procesos de ETL se ejecuta y guardan los datos en sus respectivas tablas.

En este apartado, se explica cuáles son las herramientas utilizadas para comprobar que todos los resultados que se obtengan sean los esperados.

Esta etapa se considera un Hito del proyecto, porque si las pruebas realizadas dan resultados satisfactorios, se continúa con el desarrollo normal de la planificación, construyendo el tablero de control, en caso de obtener resultados no esperados, se volverá a la etapa de desarrollo para corregir los errores encontrados, para luego pasar nuevamente a la etapa de prueba, continuando de esta forma, hasta que los resultados de esta etapa sean satisfactorios.

Al destacar la importancia de que los datos obtenidos sean los esperados se está haciendo referencia a la calidad de datos. El ingeniero y profesor de la Facultad de Informática, Emilio Lorenzon [L 9], define calidad de datos de la siguiente forma.

‘Podemos decir que la calidad de datos implica que los datos capturados, procesados, almacenados y entregados son un fiel reflejo de la realidad que se

desea tratar mediante sistemas informáticos. Esto supone que los datos no contengan errores, sean veraces y estén actualizados.

La definición más aceptada de calidad de datos es “aptitud para su uso” (“fitness for use”): los datos tiene calidad si sirven para el propósito que se necesita. Podemos decir entonces, que la calidad depende tanto de los datos como del uso de los mismos. A la vista de esta definición, se deduce que la calidad de datos es un concepto subjetivo (depende del uso que vaya a hacer cada persona) y multidimensional (es factible entender la calidad desde distintos puntos de vista – al estilo del software con ISO 9126).’

Por otra parte, Ralph Kimball, también hace referencia al tema, mencionando al dato como materia prima del Data Warehouse. La desconfianza y pérdida de credibilidad del Data Warehouse serán resultados inmediatos e inevitables si el usuario choca con información inconsistente. Es por ello que la calidad de los datos es un factor determinante en el éxito de un proyecto de Data Warehousing. Es esta etapa donde deben sanearse todos los inconvenientes relacionados con la calidad de los datos fuentes.

Ralph Kimball propone un plan de 10 ítems que ayudara a guiar esta etapa.

Plan:

1. Crear un diagrama de flujo-destino esquemático, de una página y de muy alto nivel.
2. Probar, elegir e implementar una herramienta de Data Staging.
3. Profundizar en detalle por tabla destino, gráficamente describir las reestructuraciones o transformaciones complejas. Gráficamente ilustrar la generación de las claves subrogadas. Desarrollo preliminar de la secuencialidad de los trabajos.

Carga de dimensiones:

4. Construir y probar la carga de una tabla dimensional estática. La principal meta de este paso es resolver los problemas de infraestructura que pudieran surgir (conectividad, transferencia, seguridad, etc.).
5. Construir y probar los procesos de actualización de una dimensión.
6. Construir y probar las cargas de las restantes dimensiones.

Fact Tables y automatización:

7. Construir y probar la carga histórica de las fact tables (carga masiva de datos). Incluyendo búsqueda y sustitución de claves.
8. Construir y probar los procesos de carga incremental.
9. Construir y probar la generación de agregación.
10. Diseñar, construir y probar la automatización de los procesos.

El objetivo de esta etapa de prueba de calidad de los datos, es realizar validaciones en el desarrollo realizado, buscando que el tablero de control muestre calidad y veracidad de sus datos, logrando de esta manera ganar la confianza del sponsor al momento de consultar el tablero, que le servirá al momento de tomar las decisiones sobre el negocio.

Alcance

Se realizaran pruebas de los procesos de ETL creados y de los datos almacenados en el Data Warehouse.

Prueba de Proceso

Las pruebas de los procesos es la primera que se realiza debido a la necesidad de tener datos almacenados en el Data Warehouse, por ello primero es necesario que los procesos de ETL se ejecuten exitosamente. Con esta prueba, se busca que cada uno de los procesos inicie y termine su ejecución correctamente como se los programo.

Al hablar de la ejecución del ETL, aún no se observan los datos, si fueron cargados y/o si son los esperados. Sólo se busca saber si corrió el proceso sin detenerse o sin devolver algún tipo de error. Por lo tanto, al finalizar la ejecución del ETL, pueden suceder distintas situaciones, que no se cargaron los datos o que los datos se cargaron sin saber realmente la calidad de los mismos.

En el capítulo 3 se realizó una descripción del funcionamiento del proceso de carga de las tablas de Dimensiones y las tablas de Hecho. En la descripción se menciona paso a paso cada una de las tareas que se realizan para la obtención del dato fuente y su almacenamiento en la tabla, como muestra la Imagen 25 – Tabla Dimensión Jurisdicción (**página 44**).

Prueba de Datos

Finalizada la Prueba de Proceso, se continúa con la Prueba de Datos. Esta prueba tiene como objetivo principal comprobar que hayan datos cargados en el Data Warehouse (control realizado mediante una sentencia SQL, utilizando la Función 'Count()') y luego, comprobar que los datos cargados brinden calidad y veracidad.

Cabe mencionar que el proceso ETL solo va a trabajar con los datos que cumplan con el formato esperado por la tabla destino. Los datos que contienen algún tipo de error y provoquen una excepción en su carga, serán descartados sin analizarlos, mediante el proceso de bifurcación que se ha programado, como se puede observar en la Imagen 25 – Tabla Dimensión Jurisdicción (**página 44**). Para el desarrollador es muy difícil tomar una decisión sobre el dato, que permita corregir su formato o valor, para luego cargarlo en la tabla, tanto sea de dimensiones como de hecho.

Al hablar de Calidad de los datos almacenados en el Data Warehouse, significa que el formato del dato es el esperado. Un dato sin calidad, quiere decir que se encuentran caracteres (letras, números, símbolos, etc.) que no se esperaban en campo determinado. Por ejemplo, en un campo definido del tipo Date, solo se espera que el dato almacenado se corresponda con una fecha y que no sea un dato de otro estilo.

La veracidad de los datos, es el grado de confianza que se establece sobre el valor almacenado. Sobre este valor almacenado, se tomarán decisiones importantes. Por ejemplo, el valor antes de ser almacenado, puede haber sido calculado, entonces, se debe controlar que su resultado sea el correcto. Como también, puede pasar que el campo no almacene el dato, dejando el campo en null, esta situación, también estaría mostrando inconsistencia en el resultado.

Sabiendo que hay datos cargados en el Data Warehouse, se realizan distintas pruebas:

Datos nulos, se controla que los registros más importantes como aquellos que fueron definidos como campos claves no hayan quedado en null. Este control, me da seguridad de que la tabla analizada está bien y en consultas futuras se va a enlazar correctamente con datos de otras tablas, de Dimensión o de Hecho.

Registros repetidos, se controla que no existan registros repetidos. Cada registro solo debe estar registrado una sola vez en la tabla. Es importante que la tabla de

dimensiones no contenga un mismo registros almacenado más una vez, ya que los registros repetidos pueden provocar una apertura de datos no correcta.

Cruce de tablas, es una prueba general que realizamos con todas las tablas involucradas del Data Warehouse, enlazando unas con otras de acuerdo al modelo definido. Para este modelo definido del Data Warehouse, se creó una tabla de Hecho con cada una de las FOREIGN KEY de las tablas de Dimensiones.

Query Función Count

Para la siguiente query se busca saber si en la tabla seleccionada, se logró almacenar los datos obtenidos a partir del ETL. Para ello, se utiliza la función count() de SQL Server, donde devuelve como resultado el valor 0 (cero) en caso de que la tabla no tenga registros guardados, o un valor mayor a 0, en caso de tener uno o más registros almacenados en la tabla.

```
select count(1) from Tabla
```

Query Datos nulos

Con la existencia de datos en la tabla y realizada la comprobación anterior, se puede ejecutar la siguiente query, donde se busca saber si los campos de la tabla tienen valor o no (null). Para ello, en caso de no existir valores, se considera que los campos pueden estar en null. Para saber si algún campo del registro es null, se utiliza el operador IS NULL, donde devuelve como resultado los registros que cumplen con la condición.

```
select campo1, campo2  
from Tabla  
where campo1 is null or campo2 is null
```

Query Registros repetidos

Una vez realizada la validación de existencia de datos en la tabla y de que ellos contienen valores, se puede realizar la siguiente prueba de validación para comprobar que cada registro cargado en la tabla es único. La siguiente query, realiza tal comprobación de registros. En caso de encontrarse registros cargados más de una vez, indica cuales son, por medio de los campos referenciados como campo1 y campo2, en la sentencia del select, de la siguiente sentencia SQL.


```
select count(1) as cant, campo1, campo2
from Tabla
group by campo1, campo2
having count(1) > 1
```

Cruce de tablas

La última validación que se realiza es juntando todas las tablas involucradas en el modelo del Data Warehouse, tablas de Dimensiones y tabla de Hecho, para relacionarlas por medio de sus claves, buscando como resultado obtener los datos esperado para el modelo definido. A continuación se muestra un ejemplo para esta sentencia SQL.

```
select D1.campo1, D1.campo2, D2.campo1, T.Indicador1, T.Indicador2
from TablaHecho T
inner join TablaDimension1 D1 on D1.id = T.Id_Dimension1
inner join TablaDimension2 D2 on D2.id = T.Id_Dimension2
```

Luego de haber realizado todas las pruebas necesarias para llegar a tener calidad en los datos cargados en las tablas, comprobado mediante la ejecución de las query detallada anteriormente y habiéndose obtenido resultados positivos para cada una de las tablas, tanto de Dimensiones, como para la tabla de Hecho, se puede proceder con la siguiente etapa del desarrollo del proyecto, la creación del Dashboard.

Capítulo 4 - Entrega del tablero

4. Introducción

Se realizó una introducción al mundo de BI, en el capítulo 1 de este trabajo de tesis, buscando entender de qué se trata esta tecnología implantada desde hace varios años, la cual ha demostrado resultados favorables, haciendo que permanezca vigente por muchos años más. En el capítulo 2, se explicó la arquitectura utilizada para llevar adelante el desarrollo del tablero presentado en esta tesina. Como también, se introdujeron los conceptos de ETL y de la herramienta de Microstrategy. En el capítulo 3, el trabajo de tesis, se enfocó en la parte más práctica, explicando cuales fueron los pasos realizados para construir cada uno de los ETL y los objetos necesarios que Microstrategy requiere para la construcción del tablero.

En este capítulo, se muestra el tablero terminado, visualizando los datos cargados, a partir del DW construido.

Los datos visualizados en el tablero, son de prueba, ilustrando el funcionamiento del Tablero de Control.

4.1 Tablero de Control

El tablero de control reflejara los datos cargados en nuestro Data Warehouse. Se mencionó que un tablero está compuesto por atributos e indicadores. Los atributos brindan la descripción que deseamos mostrar y los indicadores reflejan los valores asociados al atributo. La descripción de atributo se obtiene de las tablas de Dimensiones, estando relacionada con la tabla de Hecho que almacena los valores de los indicadores (**Ver Imagen 20 – Modelo Data Warehouse, página 42**).

A continuación se realizara una descripción del Tablero de Control.

El Tablero de Control está compuesto por tres secciones. El Encabezado, que contiene el título y logos de la entidad. Una sección Filtro, que permite configurar nuestro criterio de visualización. Y por último, la sección Detalle, donde se visualizan los valores en formato de grafico o grilla.

La sección Filtro está compuesta por tres botones simulando solapas, que permiten organizar los datos por 'Importe y Cantidad', 'Jurisdicción' y 'Distribución por OG'.

Estas solapas comparten filtros, compuestos por los atributos Tiempo (**Mes**, **Trimestre**, **Año**), **Procedimiento** y **Estado OC**, que permiten seleccionar de forma dinámica el grado de comparación que se necesite obtener para lograr la información requerida.

A continuación, se explica la información que brida el tablero en cada una de sus solapas.

Importe y Cantidad, se llama la primera solapa del tablero y hace referencia a los indicadores (KPI) creados en la tabla FACT_Compra del DW.

Los gráficos (**ver Imagen 29 - Carga tabla DIM_OBJETO_GASTO, página 48**) muestran estos KPI sumariados de acuerdo a los parámetros seleccionados, en la sección de filtros. En los gráficos de barra ubicada en la parte superior de la imagen, se puede observar que su serie está a nivel de **Mes** en el año. Sobre la izquierda de la imagen, se ubican los gráficos determinados por el indicador **Importe** y por la derecha, los determinados por el indicador **Cantidad**. En la parte inferior de la imagen, los datos se agrupan a nivel de **Año**.

A partir de estos gráficos, el usuario, podrá realizar comparaciones de los datos obtenidos por los **Mes** del año seleccionado, como también, por los **Años** seleccionados.

La Imagen 34 – Importe y Cantidad, muestra los datos devueltos para los siguientes filtros seleccionados, Órdenes de Compras en estado Perfeccionada, realizadas mediante un proceso de Licitación Pública, durante el periodo 2014 y 2016.

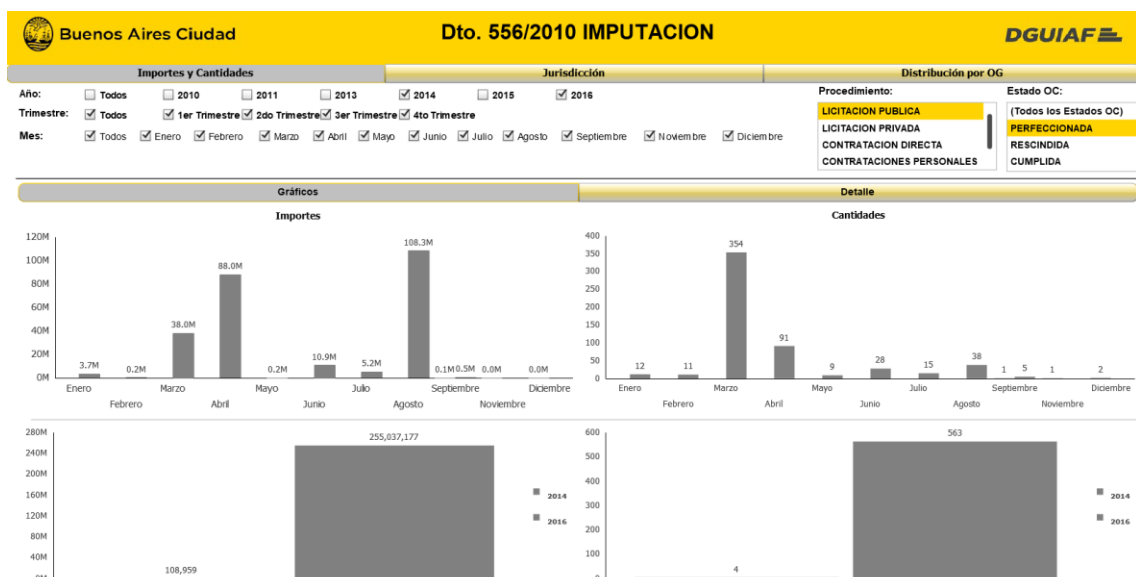


Imagen 34 – Importe y Cantidad

La Imagen 35 - Detalle, contiene gráficos de barra en la parte superior, reflejando los valores para el indicador **Importe** y **Cantidad**. En la parte inferior, se visualiza una grilla con un nivel mayor granularidad de los datos.

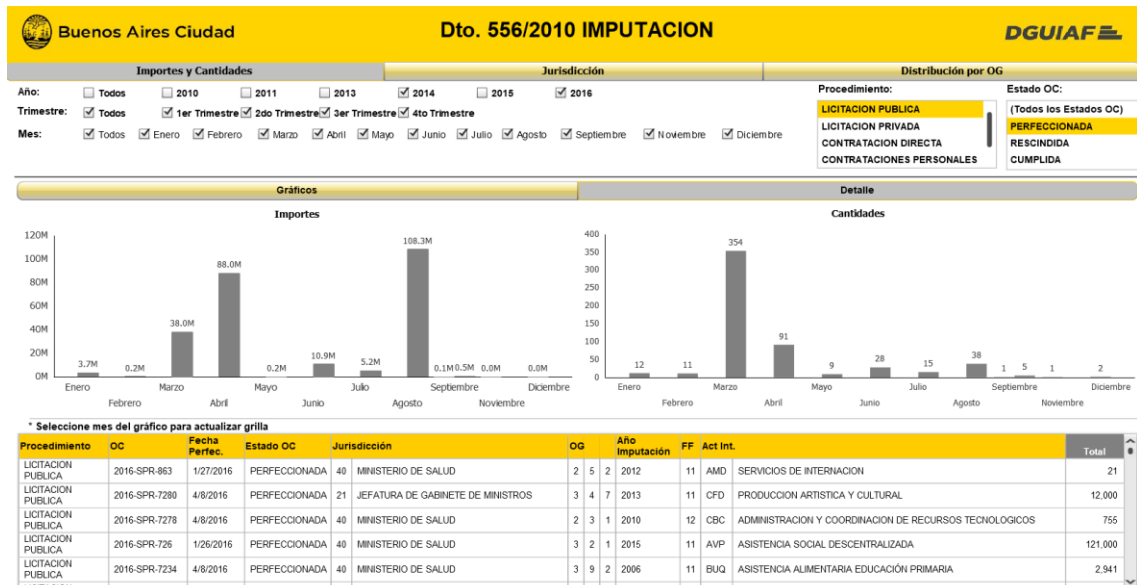


Imagen 35 - Detalle

Jurisdicción se llama la segunda solapa del tablero. La Imagen 36 - Jurisdicción, muestra en la parte superior un gráfico de barra organizado por el clasificador **Jurisdicción**, donde las barras muestran su valor de acuerdo al indicador **Importe**. En la parte inferior de la imagen, se presentan dos grillas, brindando información más detallada que el gráfico de barra superior, al igual que la grilla que se encuentra sobre su derecha.

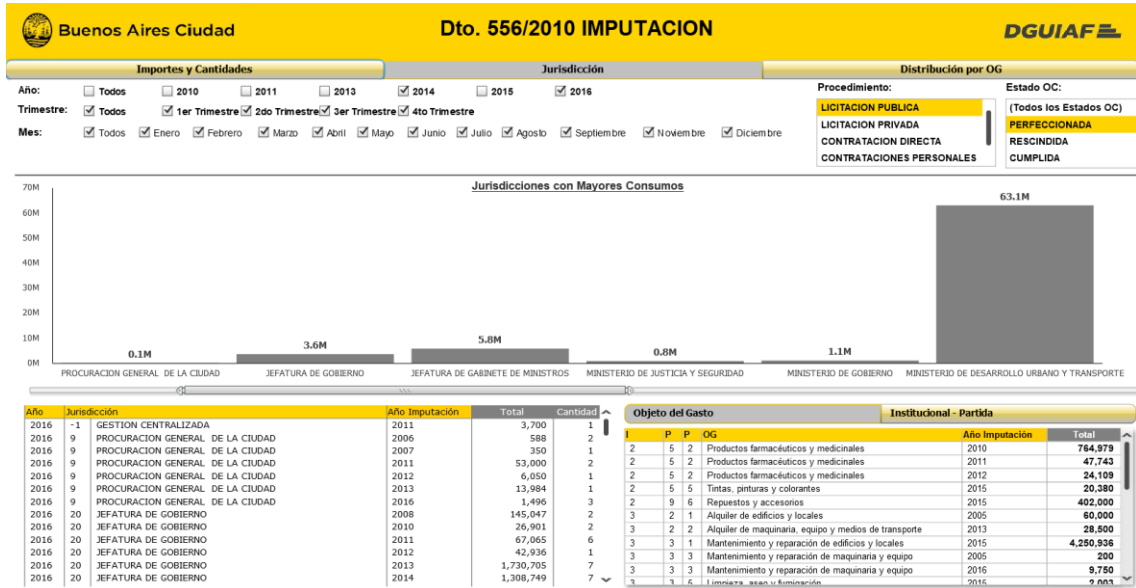


Imagen 36 - Jurisdicción

Distribución por OG, es la tercera solapa del tablero, donde se puede visualizar la pantalla del tablero en la Imagen 37 – Distribución OG. Esta solapa, muestra los datos referido al **Objeto del Gasto**, organizados en tres solapas, 'Importe'; 'Cantidad' y 'Distribución por OG y Jurisdicción', que permiten organizar la información de acuerdo a distintos criterios. Para el caso de la Imagen 37 – Distribución OG, está seleccionada la solapa 'Importe', la cual contiene un gráfico de barras sobre la izquierda de la imagen y una grilla sobre su derecha. Los datos que se visualizan están relacionados con el clasificador **Objeto del Gasto**, donde se puede observar cuanto monto se proporcionó para cada Objeto del Gasto, de acuerdo al periodo seleccionado en el filtro superior del tablero.

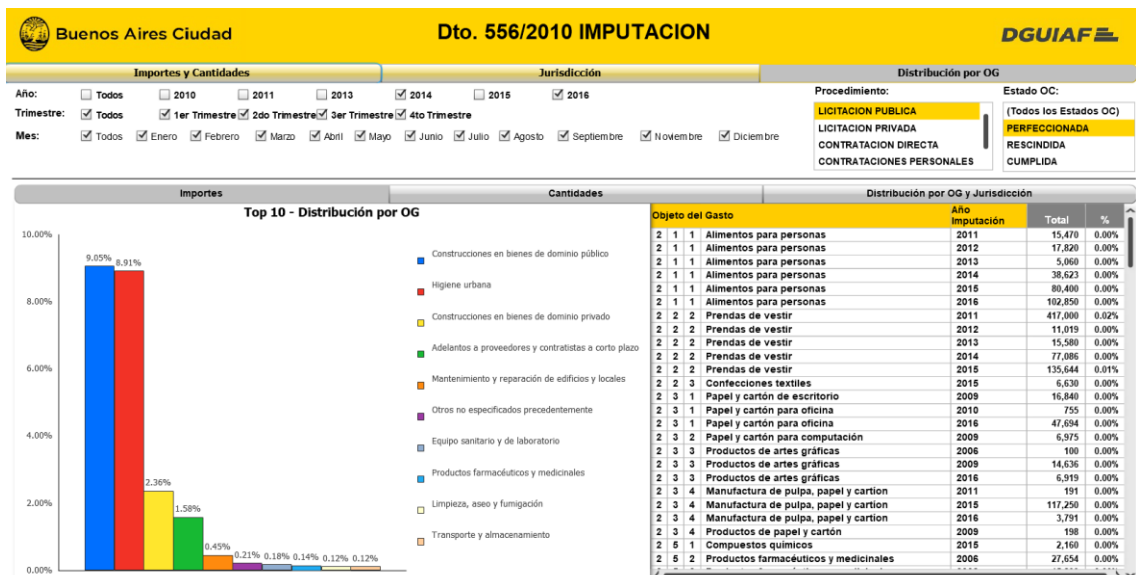


Imagen 37 – Distribución OG

La Imagen 38 – Distribución OG, Cantidad, muestra los datos presentados por el indicador **Cantidad**, mediante el grafico de barra en la izquierda de la imagen y una grilla a su derecha.

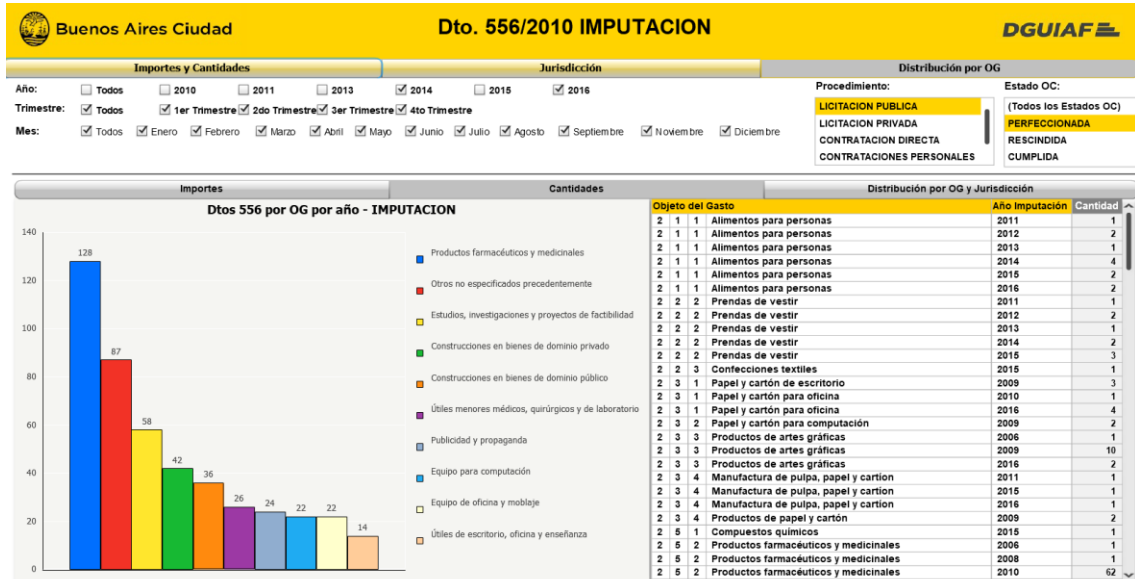


Imagen 38 – Distribución OG, Cantidad

Para la Imagen 39 – Solapa Distribución por OG y Jurisdicción, se presenta una grilla, donde brinda una apertura de datos a un mayor, juntando los atributos **Objetos del Gasto y Jurisdicción**.

Año	Inciso	Principal	Parcial	Descripción OG	Jurisdicción	Año Imputación	Total
2014	2	1	1	Alimentos para personas	2014 30 MINISTERIO DE DESARROLLO URBANO	2014	31,200
2014	3	4	1	Estudios, investigaciones y proyectos de factibilidad	2014 30 MINISTERIO DE DESARROLLO URBANO	2014	62,300
2014	4	3	6	Equipo para computación	2014 21 JEFATURA DE GABINETE DE MINISTROS	2013	15,459
2016	2	1	1	Alimentos para personas	2016 35 MINISTERIO DE AMBIENTE Y ESPACIO PUBLICO	2011	15,470
2016	2	1	1	Alimentos para personas	2016 40 MINISTERIO DE SALUD	2012	15,840
2016	2	1	1	Alimentos para personas	2016 55 MINISTERIO DE EDUCACION	2012	1,580
2016	2	1	1	Alimentos para personas	2016 21 JEFATURA DE GABINETE DE MINISTROS	2013	5,060
2016	2	1	1	Alimentos para personas	2016 20 JEFATURA DE GOBIERNO	2014	4,432
2016	2	1	1	Alimentos para personas	2016 21 JEFATURA DE GABINETE DE MINISTROS	2014	2,992
2016	2	1	1	Alimentos para personas	2016 26 MINISTERIO DE JUSTICIA Y SEGURIDAD	2015	54,000
2016	2	1	1	Alimentos para personas	2016 40 MINISTERIO DE SALUD	2015	26,400
2016	2	1	1	Alimentos para personas	2016 40 MINISTERIO DE SALUD	2016	4,950
2016	2	1	1	Alimentos para personas	2016 60 MINISTERIO DE HACIENDA	2016	97,500
2016	2	2	2	Prendas de vestir	2016 60 MINISTERIO DE HACIENDA	2011	417,000
2016	2	2	2	Prendas de vestir	2016 60 MINISTERIO DE HACIENDA	2012	11,019
2016	2	2	2	Prendas de vestir	2016 40 MINISTERIO DE SALUD	2013	15,580
2016	2	2	2	Prendas de vestir	2016 21 JEFATURA DE GABINETE DE MINISTROS	2014	9,486
2016	2	2	2	Prendas de vestir	2016 28 MINISTERIO DE GOBIERNO	2014	67,600
2016	2	2	2	Prendas de vestir	2016 20 JEFATURA DE GOBIERNO	2015	3,483
2016	2	2	2	Prendas de vestir	2016 40 MINISTERIO DE SALUD	2015	5,578
2016	2	2	2	Prendas de vestir	2016 60 MINISTERIO DE HACIENDA	2015	126,583
2016	2	2	3	Confecciones textiles	2016 20 JEFATURA DE GOBIERNO	2015	6,630
2016	2	3	1	Papel y cartón de escritorio	2016 26 MINISTERIO DE JUSTICIA Y SEGURIDAD	2009	1,511
2016	2	3	1	Papel y cartón de escritorio	2016 40 MINISTERIO DE SALUD	2009	15,329
2016	2	3	1	Papel y cartón para oficina	2016 40 MINISTERIO DE SALUD	2010	755
2016	2	3	1	Papel y cartón para oficina	2016 9 PROCURACION GENERAL DE LA CIUDAD	2016	1,209
2016	2	3	1	Papel y cartón para oficina	2016 40 MINISTERIO DE SALUD	2016	45,881

Imagen 39 – Solapa Distribución por OG y Jurisdicción

En este capítulo se presentó la solución, demostrando la importancia de los datos cargados a partir de los ETL desarrollados para su Data Warehouse. El tablero de control es la pantalla visible para el usuario final, abstrayéndose totalmente de la complejidad demostrada que lleva construir un Data Warehouse, desde la obtención de los datos, transformándolos para que se adapten al requerimiento y posteriormente siendo cargados en las tablas. El tablero de control, debe ganarse la confianza del sponsor diariamente, mediante la proyección de datos reales.

Conclusión Final

El en presente trabajo de tesina se presentó la actividad profesional desarrollada en las oficinas de Sistemas del Ministerio de Hacienda de la Ciudad de Buenos Aires. Describiendo las tareas realizadas en la elaboración del Tablero de Control utilizado por el Gobierno para su gestión.

Mediante la presentación del problema, se buscó la forma de abarcarlo a través de diferentes autores, tomando como metodología de trabajo principal propuesta por Ralph Kimball.

Se describió la arquitectura de servidores configurada para la implementación del Tablero, mencionando los ambientes de Desarrollo, Testeo y Producción.

Posteriormente, se realizó una introducción a las herramientas de software utilizadas que dan lugar a la implementación del Tablero de Control.

A partir del requerimiento funcional obtenido, se definió el requerimiento técnico con 7 clasificadores (Año, Trimestre, Mes, Jurisdicción, Objeto del Gasto, Procedimiento Compra y Estado) y dos Indicadores (Cantidad e Importe). Luego, se describió la solución para los ETL que cargarán los datos en las tablas de Dimensiones y en la tabla de Hecho, controlando que los mismos cuenten con Calidad, como se enseñó en las distintas materia de la cátedra. Por último, utilizando la herramienta Microstrategy, se diseñó el Dashboard, que da cumplimiento al requerimiento funcional.

Se puede concluir que Business Intelligence brinda soluciones estratégicas a la hora de gestionar la administración de entidades públicas y/o privadas.



Trabajo Futuro

El siguiente trabajo de tesina presentó una solución desarrollada mediante software de licencia primitiva. En el futuro se puede incorporar software de licencia libre, como lo es R, que permita analizar grandes volúmenes de datos mediante los nuevos paradigmas como la Ciencia de Datos y Big Data.

Incorporar análisis estadísticos, minería de datos, aprendizaje automático y análisis predictivo.

Anexo Query

En este apartado se muestran las query que son ejecutadas por los procesos ETL. El nombre de cada query hace referencia al nombre utilizado en la tarea de SSIS.

1 - Proceso Id

```
/* Busca el Id del Proceso, almacenado en la table Proc_Procesos, y lo guarda en la variable llamada Proceso_Id */
```

```
Select Top 1 Proceso_Id
From Proc_Procesos Pr
Inner Join Proc_Lotes Lo
  On Pr.Lote_Id = Lo.Lote_Id
Where Pr.Estado_Id = 0 And
  Lo.Nombre_Etl = 'Etl_Dim_Objeto_Gasto'
Order By Pr.Proceso_Id
```

2 - Actualizo A Estado 1

```
/* Actualiza el campo Estado_ID = 1 de la table Proc_Procesos, en el cual indica que el proceso se está ejecutando */
```

```
Update Proc_Procesos
Set Estado_Id = 1
Where Proceso_Id = @Proceso_Id
```

3 - Trunco Table Dim_Objeto_Gasto

```
Truncate Table Dbo.Dim_Objeto_Gasto
```

4 - Cargo Table Dim_Objeto_Gasto

```
Insert Into Dim_Objeto_Gasto (Anio, C_Inciso, C_Partida, C_Parcial,
C_Subparcial)
```

```
Select
  Obj.Aa_Ejervg      Anio,
  Obj.C_Inciso      C_Inciso,
  Obj.C_Pprincipal  C_Principal,
  Obj.C_Pparcial    C_Parcial,
  Obj.C_Psparcial   C_Subparcial,
  Substr(Obj.Xl_Objgto,1,50)  Descripcion,
From Slu.Bobjeto_Gasto Obj
```

5 - Cargo Table Dim_Fecha

```
Select To_Date('01/01/2000', 'Dd/Mm/Yyyy')+Rownum-1 As Fecha
      , To_Char((To_Date('01/01/2000', 'Dd/Mm/Yyyy')+Rownum-1), 'Mm') Mes
      , To_Char((To_Date('01/01/2000', 'Dd/Mm/Yyyy')+Rownum-1), 'Q')
Trimestre
      , Ceil((((To_Number (To_Char((To_Date('01/01/2000',
'Dd/Mm/Yyyy')+Rownum-1), 'Mm')) - 1) / 6) + 1) / 2) As Semestre
      , To_Char((To_Date('01/01/2000', 'Dd/Mm/Yyyy')+Rownum-1), 'Yyyy')
Anio
From Dual
Connect By Level <= (Select Trunc(To_Date('31/12/2100', 'Dd/Mm/Yyyy')-
To_Date('01/01/2000', 'Dd/Mm/Yyyy')) From Dual)
```

6 - Carga Table Fact_Compra

```

Insert Into Fact_Compra (Id Id_Fecha
, Id_Objeto_Gasto
, Id_Tipo_Procedimiento
, Id_Estado_Compra
, Id_Jurisdicción
, Cantidad
, Importe)
Select
  Fh.Id      Id_Fecha
, Ob.Id      Id_Objeto_Gasto
, Tpc.Id     Id_Tipo_Procedimiento
, Ec.Id      Id_Estado_Compra
, Jur.Id     Id_Jurisdicción
, 1 As      Cantidad
, Toc.Importe
From Slu.Torden_Compra Toc
Left Join Dim_Objeto_Gasto Ob
  On Ob.Anio = Toc.Anio
  And Ob.C_Inciso = Toc.C_Inciso
  And Ob.C_Principal = Toc.C_Principal
  And Ob.C_Parcial = Toc.C_Parcial
  And Ob.C_Subparcial = Toc.C_Subparcial
Left Join Dim_Jurisdicción Jur
  On Jur.Anio = Toc.Anio
  And Jur.Descripción = Toc.D_Jurisdicción
Left Join Dim_Fecha Fh
  On Fh.Fecha = Toc.Fh_Estado
Left Join Dim_Estado_Compra Ec
  On Ec.E_Compra = Toc.E_Compra
Left Join Dim_Tipo_Procedimiento Compra Tpc
  On Tpc.T_Procedimiento = Toc.C_Procedimiento

```

Bibliografía

- [1] Ballvé, A. M. (2007). *Tablero de Control, Información para crear valor*.
- [2] *Blog.bi-geek.com*. (s.f.). Obtenido de <https://blog.bi-geek.com/microstrategy-parte-i-introduccion/>
- [3] Brian Knight, D. K. (2012). *Knight's Microsoft SQL Server 2012 Integration Services 24-Hour Trainer*. John Wiley & Sons.
- [4] Inmon, W. H. (2002). *Building the Data Warehouse*. Third Edition.
- [5] Javier Díaz, María Alejandra Osorio y Ana Paola Amadeo. (2019). *Tecnologías para el análisis de datos basadas en software libre*. La Plata: EDULP.
- [6] Kimball, R. (2004). *The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming, and Delivering Data*.
- [7] Kimball, R. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling, 3rd Edition*.
- [8] LARA, L. S. (2014). *ANÁLISIS DE INFORMACIÓN Y TOMA DE DECISIONES PARA ADMINISTRACIÓN DE NEGOCIOS*. MÉXICO, D.F.
- [9] Lorenzon, E. (2011). *Sistemas y Organizaiones*. La Plata.
- [10] *Microstrategy*. (s.f.). Obtenido de http://www2.microstrategy.com/producthelp/archive/10.8/AnalyticsDesktop/WebUser/WebHelp/Lang_3082/Content/Creating_a_dashboard.htm
- [11] Mosquera, L. &. (2014). Data Mart Para El Sistema De Servicios Sociales Del Conadis. *Revista Politécnica*.
- [12] Mosquera, L. &. (2014). Data Mart Para El Sistema De Servicios Sociales Del Conadis. *Revista Politécnica*.
- [13] Robert S. Kaplan, D. P. (2016). *El cuadro de mando integral: The balanced scorecard (Sin colección)*.
- [14] *Sinnexus*. (s.f.). Obtenido de https://www.sinnexus.com/business_intelligence/DataMart.aspx
- [15] Wee-Hyong Tok, R. P. (2012). *Microsoft SQL Server 2012 Integration Services*. Pearson Education.
- [16] Vitt E, Luckevich M, Misner S. (2002). "Business Intelligence. Técnicas de análisis para la toma de decisiones estratégicas". McGrawHill.
- [17] Mosquera L., Hallo M. (2014) Data Mart Para El Sistema De Servicios Sociales Del Conadis.
- [18] Josep Lluís Cano (2007). BUSINESS INTELLIGENCE: COMPETIR CON INFORMACIÓN

Portales web

- [19] **Wikipedia**, (*Web consultada 03/12/2019*).
https://es.wikipedia.org/wiki/Extract,_transform_and_load
- [20] **Vanguardsoft Developers Blog**, Creando ETL con SQL Server Integration Services (SSIS), (*Web consultada 03/12/2019*).
<https://olgeferrer.wordpress.com/2013/10/14/creando-etl-con-sql-server-integration-services-ssis/>
- [21] **SQL para Humanos**, Que es SSIS (SQL Server Integration Services).
(*Web consultada 03/12/2019*)
<http://sqlparahumanos.blogspot.com/2011/08/que-es-ssis-sql-server-integration.html>
- [22] **sedici, Monitoreo de procesos y construcción de un tablero de control usando Portlets**.
<http://sedici.unlp.edu.ar/bitstream/handle/10915/59944>
- [23] **ptolomeo**,
<http://www.ptolomeo.unam.mx:8080/xmlui/bitstream/handle/132.248.52.100/3243/Tesis.pdf>
- [24] **Universidad Autónoma Metropolitana**,
http://sgpwe.izt.uam.mx/files/users/uami/sppc/Business_Intelligence/SinnexusQue_es_Busines_s_Intelligence.pdf
- [25] **I Jornada "La Administración Financiera de la CABA a 15 años de su autonomía"**,
(*Web consultada 12/2019*)
<https://www.buenosaires.gob.ar/hacienda/i-jornada-la-administracion-financiera-de-la-ciudad-autonoma-de-buenos-aires-15-anos-de-su>
- [26] **www.buenosaires.gob.ar**, (*Web consultada 12/2019*)
https://www.buenosaires.gob.ar/areas/leg_tecnica/sin/normapop09.php?id=3811&qu=c&ft=0&cp=&rl=1&rf=1&im=&ui=0&printi=1&pelikan=1&sezion=825352&primera=0&mot_da=&mot_frase=comunas&mot_alguna=
- [27] **Microstrategy**, (*Web consultada 12/2019*)
<https://www.microstrategy.com/es>
- [28] **sinnexus.com**, (*Web consultada 12/2019*)
https://www.sinnexus.com/business_intelligence/
- [29] **Ibermatica**, (*Web consultada 12/2019*)
<https://churriwifi.files.wordpress.com/2009/11/business-intelligence-ibermatica.pdf>
- [30] **Palazón, F.J. "Business Intelligence: Decisiones para el éxito". Microsoft España. 2006**, (*Web consultada 12/2019*)
http://www.microsoft.com/spain/enterprise/perspectivas/numero_18/estrategia.msp