

Data Warehouse: Análisis Multidimensional de BAFICI utilizando Power Pivot

Data Warehouse: Multidimensional Analysis of BAFICI using Power Pivot

Wayner Xavier BUSTAMANTE- Granda [1](#); Estela María MACAS- Ruiz [2](#); Fanny Beatriz CEVALLOS- Macas [3](#)

Recibido: 25/04/2018 • Aprobado: 30/05/2018

Contenido

[1. Introducción](#)

[2. Metodología](#)

[3. Resultados](#)

[4. Conclusiones](#)

[Referencias bibliográficas](#)

RESUMEN:

Este artículo de investigación tiene como objetivo construir un Data Warehouse del Festival Internacional de Cine Independiente de Buenos Aires (BAFICI), aplicando Kimball como metodología de desarrollo, ETL para la extracción, transformación y carga de datos y Power Pivot como la herramienta para la construcción, análisis y presentación de resultados. El análisis multidimensional que se realizó de las dimensiones películas, directores, países, ventas, secciones y lugares, permitió obtener indicadores que ayudan a tomar decisiones precisas sobre el desempeño de BAFICI.

Palabras-Clave: BAFICI, Power Pivot, Data Warehouse, Kimball.

ABSTRACT:

This research article aims to build a Data Warehouse of the International Independent Film Festival of Buenos Aires (BAFICI), applying Kimball as a development methodology, ETL for the extraction, transformation and loading of data and Power Pivot as the tool for the construction, analysis and presentation of results. The multidimensional analysis that was carried out of the dimensions films, directors, countries, sales, sections and places, it allowed obtaining indicators that help to make precise decisions about BAFICI's performance

Keywords: BAFICI, Power Pivot, Data Warehouse, Kimball

1. Introducción

Se considera Business Intelligence como el método científico para resolver problemas de negocios, las herramientas Data Warehouse y Data Mining, permiten gestionar los datos para obtener conocimientos que aporten satisfactoriamente al modelo de negocio de las organizaciones para la toma de decisiones de alto nivel, con criterios racionales.

Según Wang (2015) "Los datos son la base del negocio digital, cada touch, cada click, cada bit, son parte de la explotación de información."

Los volúmenes de datos, entonces, son considerados en las empresas como instrumentos imprescindibles; su uso radica en el estudio, integración, depuración y el procesamiento de información existente para la toma de decisiones, convirtiéndolos en hitos históricos para la implantación de soluciones mejoradas e integrales.

La importancia de gestionar conocimiento en empresas, radica en que las implementaciones de gestión de conocimiento han generado una gran variedad de acciones, es así que Rubikey (2011) indica que el 76% ha generado una estrategia de conocimiento, el 64% ha adoptado el entrenamiento, el 58% ha establecido compartir mejores prácticas, el 57% ha instaurado políticas de conocimiento y el 50% ha establecido redes formales de gestores de conocimiento.

Por lo tanto, con el crecimiento en la generación de los datos dentro y fuera de las entidades, y con el aumento de las posibilidades de acceso a los mismos, cada vez más organizaciones apuestan a la

importancia de gestionar y administrar sus datos como un recurso valioso, que perfeccionen sus procesos (Fu, Wojak, Neagu, Ridley, & Kim, 2011; Niemi, 2011), de ahí que sean apreciados como otro tipo de activo clave que genera valor en el capital financiero y humano (DGI, 2015a; Kooper, Maes, & Lindgreen, 2011; Rifaie, Alhajj, & Ridley, 2009).

Este proyecto de investigación recopila información de una base de datos abierta del Gobierno Autónomo de Buenos Aires, específicamente del Festival Internacional de Cine Independiente de Buenos Aires (BAFICI); información que será procesada con técnicas y herramientas (ETL, Data Warehouse, Data Mining). Los datos obtenidos serán analizados e interpretados para generar conocimiento. Resultante de ello, se logra obtener tendencias de los eventos, películas, obras de teatro y otros eventos presentados en el Festival BAFICI.

1.1. Bafici

Bafici es la audiencia masiva en cinematográfica más significativa de América Latina. Anualmente reúne a directores dedicados junto a talentos nuevos, nacionales e internacionales.

Es el lugar que difunde un cine renovador e innovador, de alta vanguardia. La programación incluye secciones para todos los gustos y edades, con sedes en toda la ciudad, muchas de ellas al aire libre, organizado por el Ministerio de Cultura de Argentina.

El Festival BAFICI es reconocido por ser un espectáculo internacional, lugar en el que se concentran las mejores películas y los más selectos productores previo a una prolija selección, su organización se basa en una amplia clasificación en el cual conjugan el arte cineasta y la cultura de cada uno de los países participantes (Bafici, 2017).

Los números finales de la edición 2017 evidencian el incremento de la convocatoria y confirman, una vez más, que el Festival es un verdadero acontecimiento en la agenda cultural de la Ciudad, en 11 días, más de 380000 personas recorrieron 32 sedes en las que se presentaron más de 1000 funciones (Bafici, 2017).

Figura 1
Actividades Gratuitas Espectadores
de BAFICI 2017 (Visión del Cine, 2018)



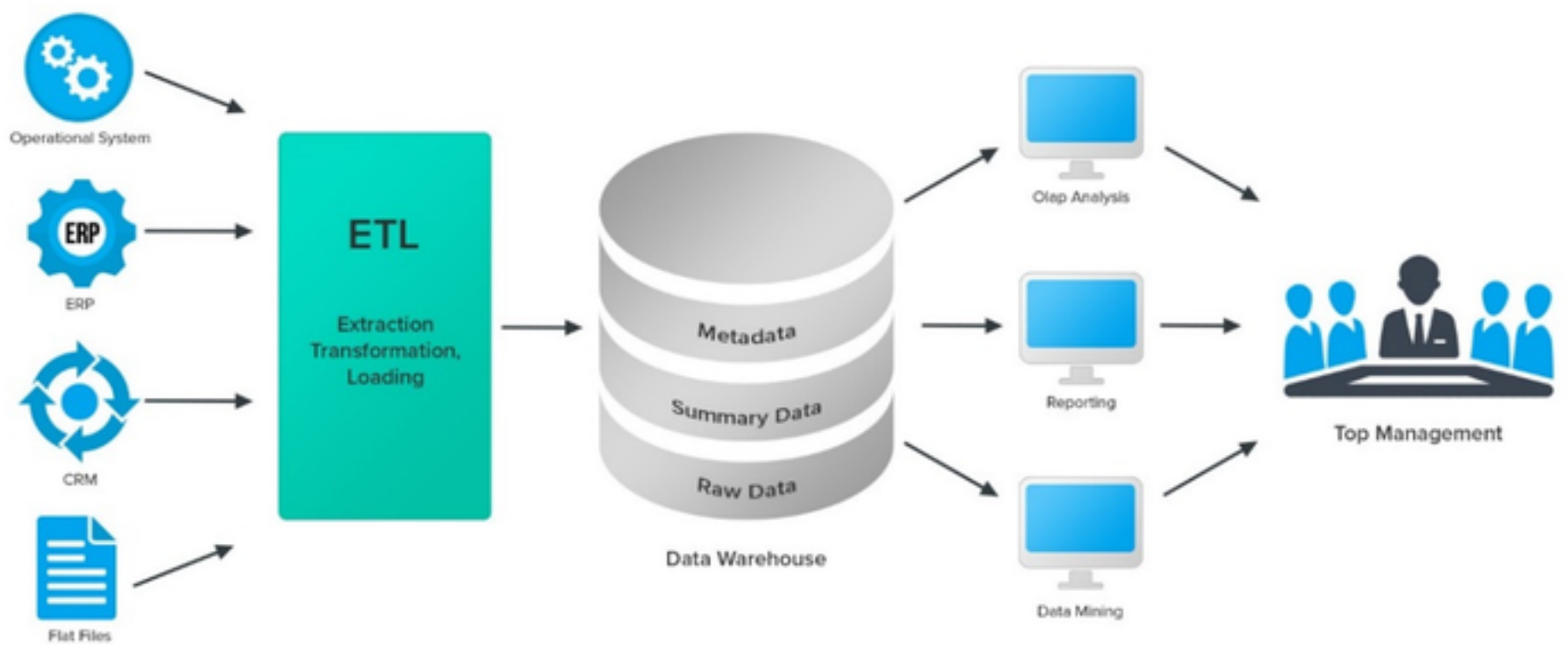
El crecimiento del Festival se traduce también en su presencia en la ciudad y la respuesta de propios y extraños amantes del arte cinematográfico, que año a año marcan un nuevo estilo e innovación de nuevas promesas del arte cinematográfico que muestran sus productos realizados para el deleite de todos.

El Festival Internacional de Cine Independiente Número 19 (BAFICI) se realizó del 19 al 30 de abril, la programación fue ampliamente difundida y publicitada de manera local, nacional e internacional, el cual cubrió alrededor de 410 películas, 100 de ellas tipo premier con carácter universal y 100 bajo contexto latinoamericano (Proyector Fantasma, 2017).

1.2. Business Intelligence

Business Intelligence concentra alto rendimiento, capacidades y beneficios como la analítica acelerada e información, BI utiliza técnicas como ETL para analizar de forma rápida grandes volúmenes de datos, dando paso al análisis en tiempo real, al procesamiento de información, la explotación de datos y la reducción de tiempos de espera para reportes, convirtiendo a BI en un instrumento tecnológico de alta gama puntal para la toma de decisiones.

Figura 2
ETL & Business Intelligence (Technologies, 2017)



La entrega oportuna, efectividad y eficiencia para las organizaciones son puntos de atención y dependerá del estudio estratégico, la administración asertiva, la aplicación de Business Process Management (BPM) la apropiada business Intelligence para llegar al cumplimiento exitoso (Shahzad & Zdravkovic, 2011).

Agostino & Solberg (2013) frente a la falta de adopción de Cloud BI por las empresas mayoritarias, manifiestan que los elementos de criterios de decisión para la adopción de una solución práctica son, la utilización de técnicas de evaluación a través de encuestas a proveedores y usuarios, y posterior a ello identificar y formular mejoras a la organización existente apoyadas de business Intelligence.

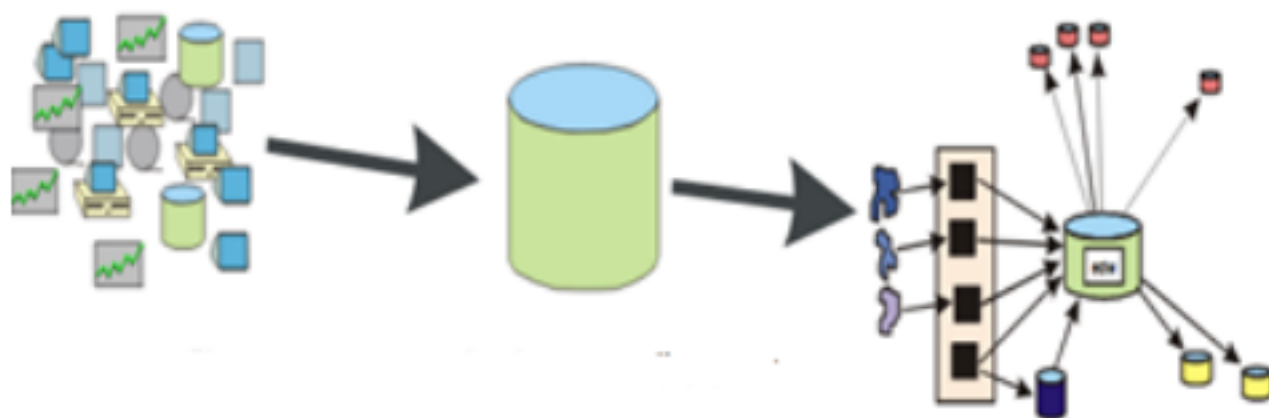
(Sakulsorn, 2011) manifiesta que las pruebas resultantes del análisis con herramientas de Business Intelligence in Memory muestran el rendimiento de los métodos BI convencionales, optimizando el tiempo de carga del software y el tiempo de respuesta.

1.3. Data Warehouse

Data Warehouse es reconocida en TI como la base de datos corporativa que se identifica por componer y refinar datos de una o más fuentes de información homogéneas o distintas.

Inmon (2010) indica que desde el principio la evolución del almacenamiento de datos ha sido modelada de acuerdo al grado de necesidades, en primera instancia existió la necesidad de acceder a los datos; luego de integrarlos; posterior a ello surgió la necesidad de una única versión refinada considerando los requerimientos departamentales y el acceso por niveles de requerimiento al entorno de almacén de datos de las organizaciones, a todo esto se lo ha denominado como la evolución histórica y general de las primeras etapas del entorno de depósito de los datos .

Figura 3
Evolución del entorno del almacén de datos (Inmon, 2010).



Inmon (2010) manifiesta que las características de Data Warehouse están conformadas por: la Integración, factor temático, factor histórico y no volátil; y se justifica en base al almacenamiento de información en el Data Warehouse y su organización sólida tienen la finalidad de apoyar con los procesos resultantes para la toma de decisiones.

La arquitectura de un Data Warehouse puede tener varias estructuras en varias ejecuciones. Por ejemplo un almacén de datos operacionales, otras pueden tener múltiples Data Marts, o un mínimo número de fuentes de información, mientras que otras pueden tener múltiples de fuentes de datos, por lo tanto existirán diversos modelos con estructuras de construcción en su arquitectura.

Figura 4
Modelo de Arquitectura Data Warehouse (Chowdhury, 2014)

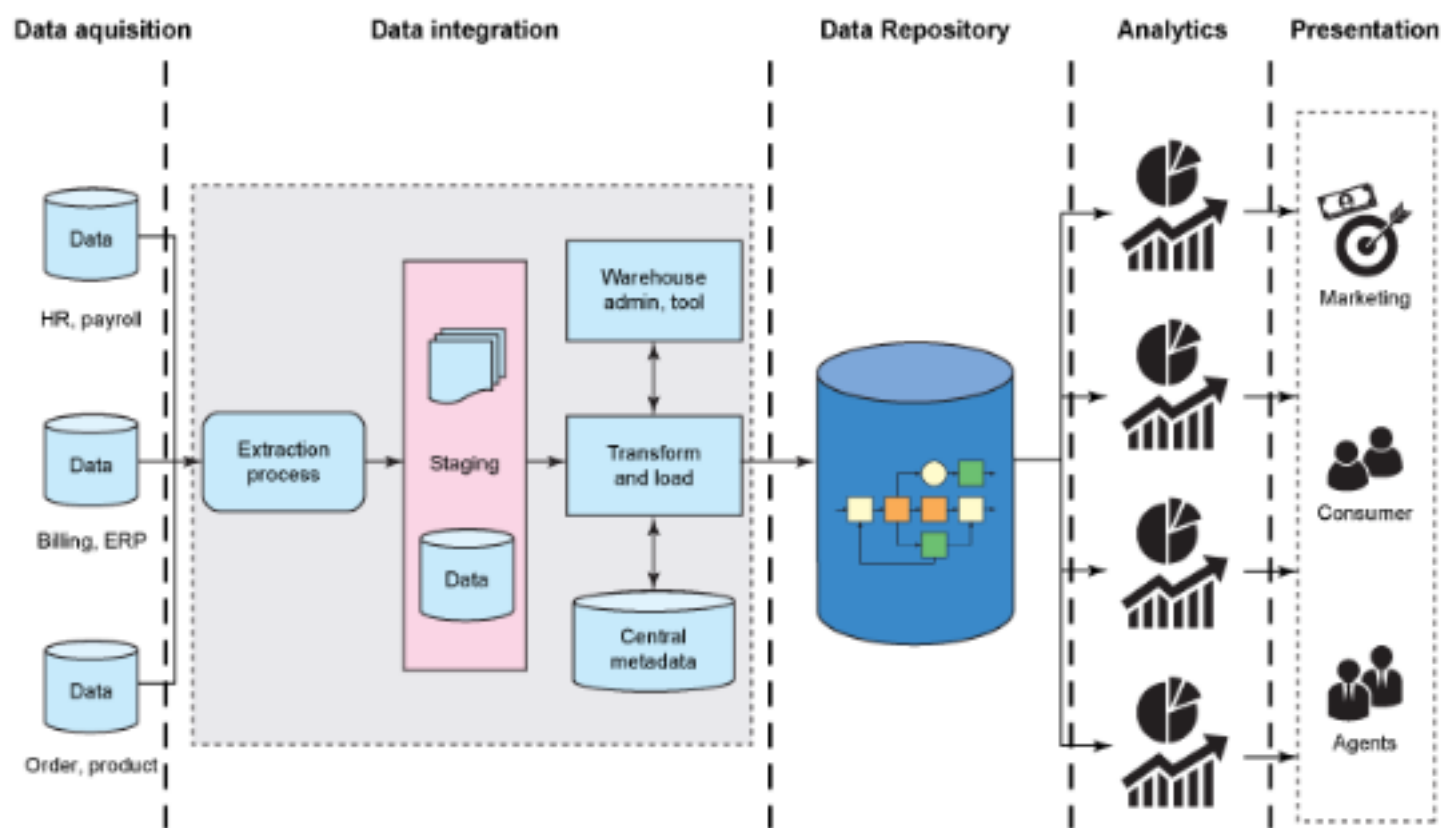
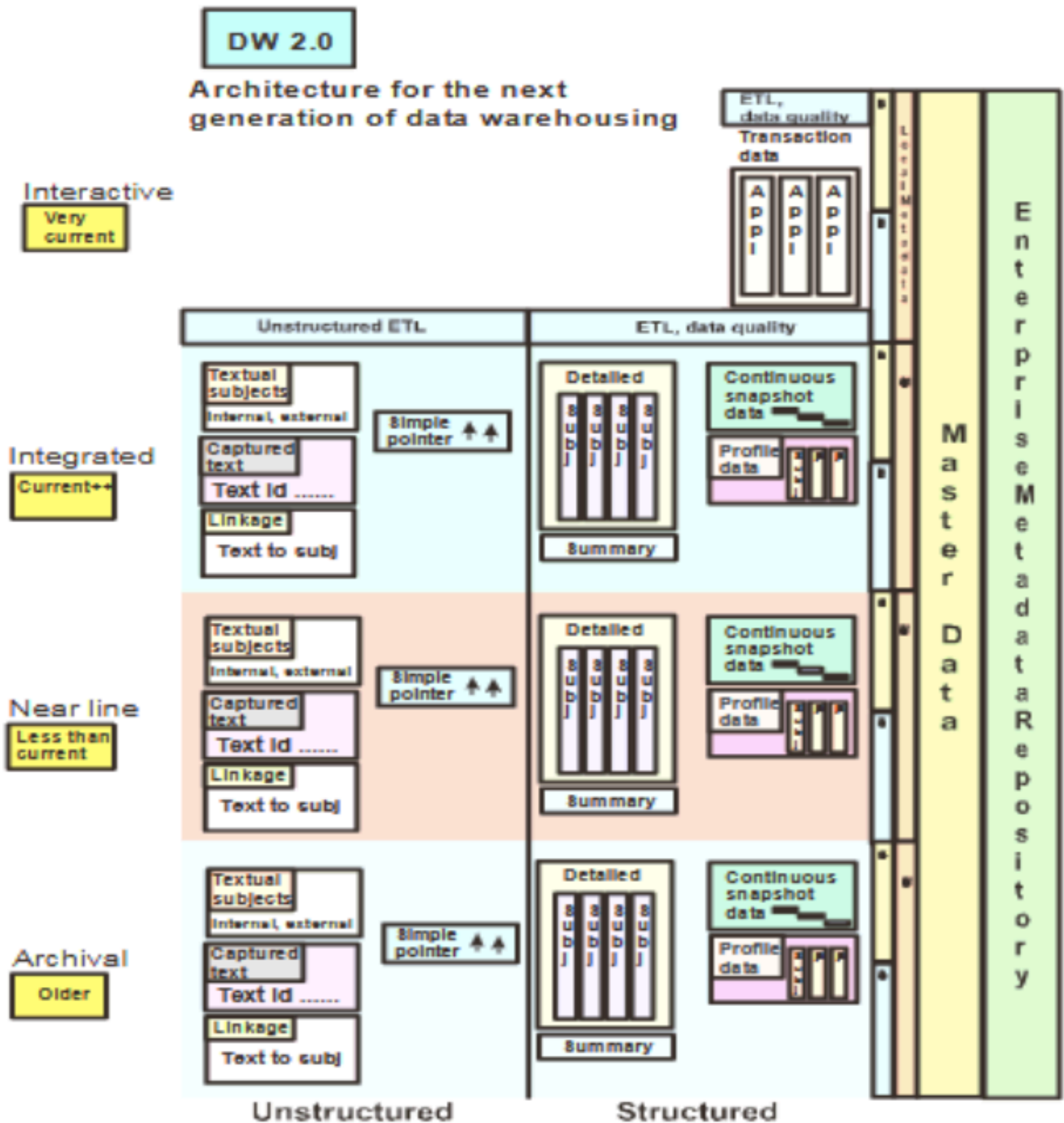


Figura 5
Modelo de Arquitectura de almacén de datos (Inmon, 2010)

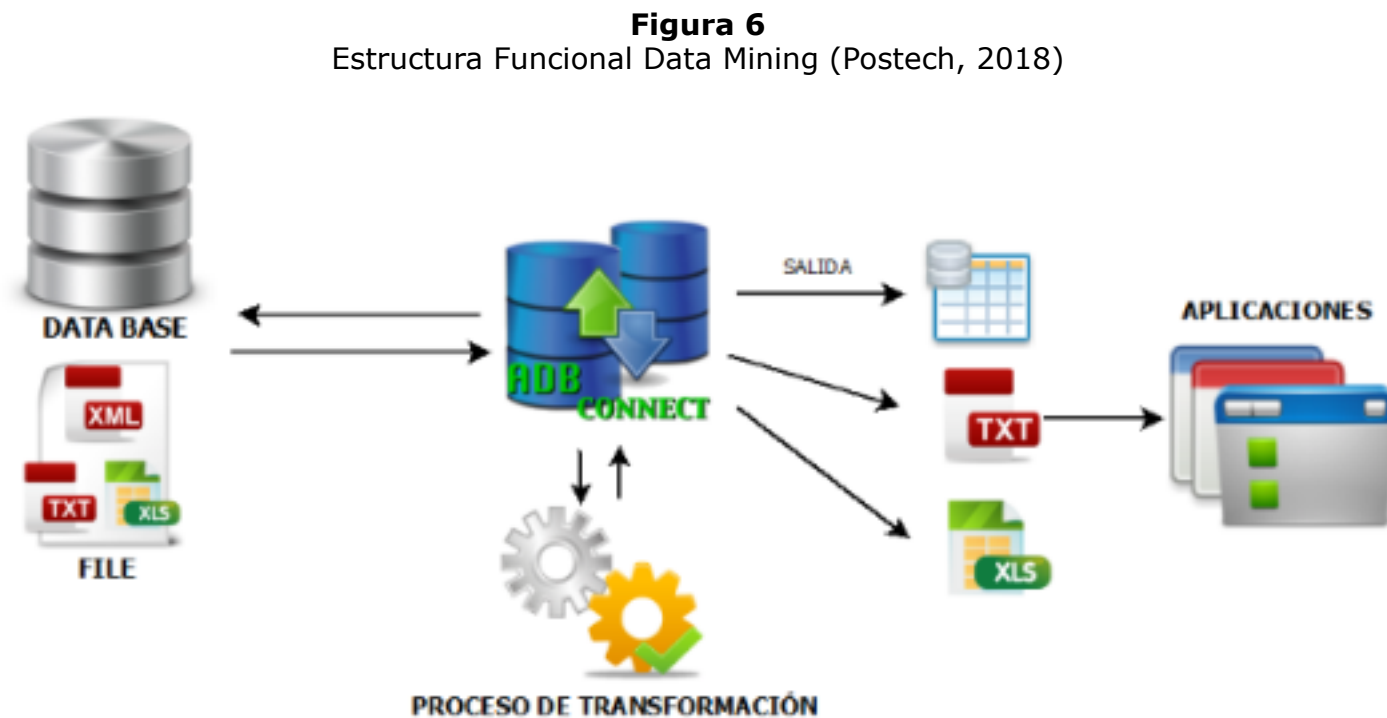


En la figura 5 muestra la arquitectura básica de un Data Warehouse versión 2.0; sobresale la interacción de datos, integración, la proximidad y el archivo resultante.

1.4. Data Mining

La Minería de datos, MD, es el proceso de extraer información válida, previamente desconocida, comprensible y útil de una base de datos de gran tamaño y utilizar dicha información para tomar decisiones de negocio crucial.

Los módulos principales descritos en la Figura 6, corresponden a la georeferenciación para la localización de registros, generación de gráficas, business Intelligence y las reglas de negocio; y la utilización de ETL para la extracción, transformación y carga de datos.



De acuerdo a la figura 6, se puede visualizar la conexión de base de datos mediante accesos de administración tipo usuario y acceder a los archivos existentes; la inspección y extracción de datos; procesamientos de información y generación de datos de salida; la información resultante organizada en tablas y/o archivos planos.

Entre las ventajas de Data Mining se puede evidenciar: la comunicación entre bases de datos, lectura de información, integración de soluciones organizacionales a través del ETL, y el reajuste de procesos significativos a través de la migración de información.

Por lo tanto la definición de Minería de Datos implica actuar sobre bases de datos de gran tamaño (a mayor volumen de datos mayor fiabilidad en las conclusiones) y complejidad elevada, para extraer el conocimiento que no es evidente o posible conocer mediante las herramientas de recuperación tradicionales.

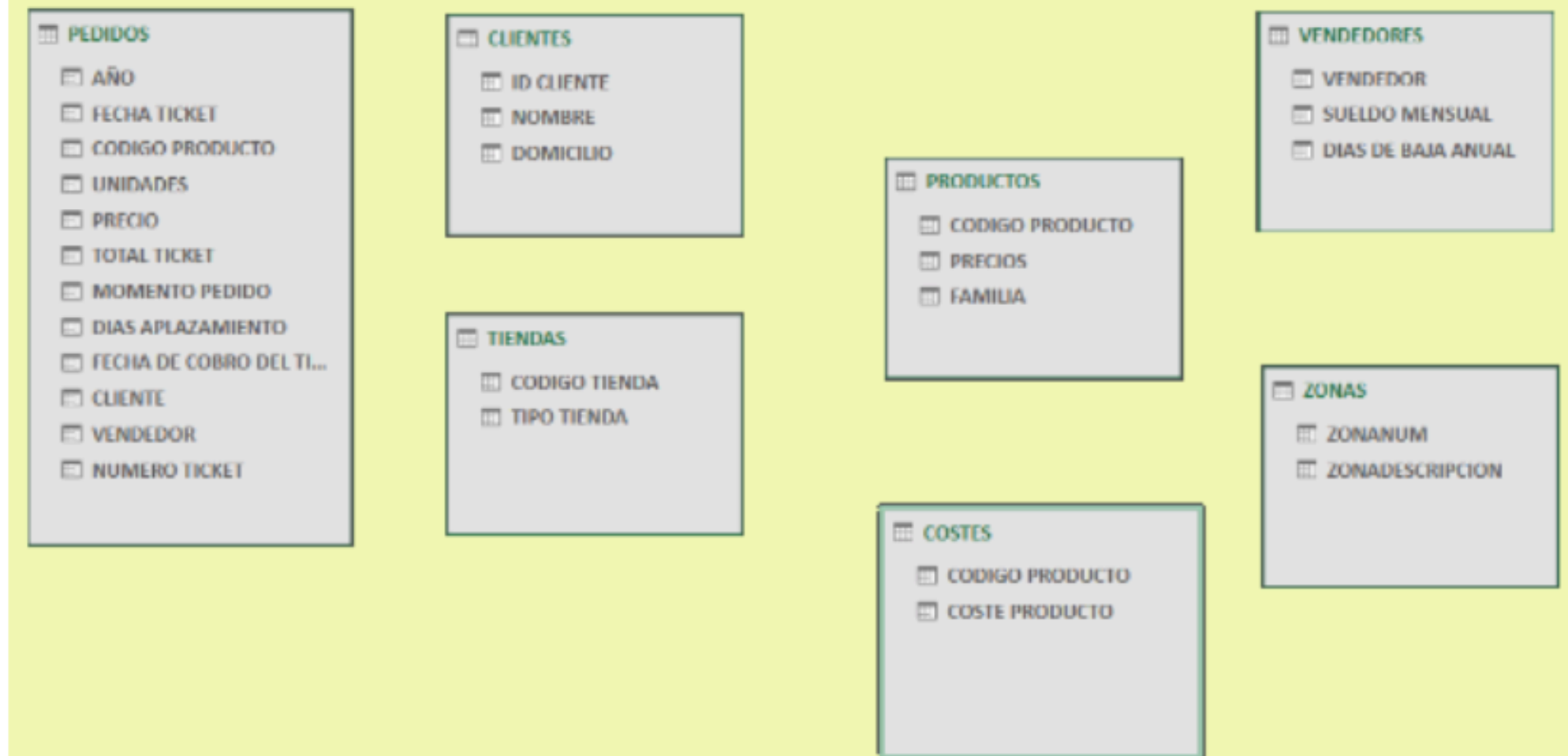
MD, es una parte o componente de KDD. KDD, es el proceso completo de extracción de información, que se encarga además de la preparación de los datos y de la interpretación de los resultados obtenidos; aunque modelos de MD como CRISP-DM, Two-Crows y SEMMA, consideran etapas referentes a estos procesos.

1.5. Power Pivot

Considerando que la indagación de información se cumple mediante la introspección de datos brutos y, posterior a ello, manifestar mediante conclusiones claras qué es lo que se obtuvo y qué factores, de acuerdo a la necesidad de la organización, se pueden ocupar como aporte para la simplificación de procesos existentes. La información resultante apoya de singular manera a la toma de decisiones de las organizaciones en relación a la parte científico-académico, este tipo de análisis da lugar a la generación juicios de comprobación o desaprobación de modelos o teorías existentes.

Power Pivot de Excel crea un análisis de datos y ejemplifica los cambios que se conciben en éstos, admitiendo analizar e identificar sucesos de escenarios futuros de fraudes y buscando métodos de prevención de la información resultante para, posterior a ello, comunicar al cliente por medio de indicadores tabulados y calculados; el complemento de Power Pivot permite crear tablas dinámicas, indicadores y gráficos interactivos a través de la herramienta de análisis de datos (DAX).

Figura 7
Generación Tablas Power Pivot (Autores, 2018)



Power Pivot actúa entonces como una herramienta que permite, posterior al análisis, representar el impacto de las decisiones utilizando como línea base los datos disponibles de toda la data del usuario petionario o del contexto de toda la información que se pretende indagar para el estudio respectivo, por lo tanto, Power Pivot cumple con la función de analizar datos sin la necesidad de la presencia y apoyo de técnicos especializados en la materia.

2. Metodología

Según Rivadera (2010) el mercado se imponen dos metodologías para el diseño de almacenes de datos: Kimball y Inmon, la diferencia desde un punto de vista arquitectónico radica en que Kimball empieza su proceso desde la fase de diseño y construcción de los Data Marts y se dirige a la construcción del DW (Botton-up, Kimball); mientras que Inmon tiene un proceso inverso, empieza desde el diseño y construcción del DW y luego se traslada hasta los Data Marts (Top-Down, Inmon).

Analizadas las metodologías existentes se determina que la que se debe aplicar en este trabajo de investigación es la metodología Kimball, teniendo en cuenta que la misma basa su diseño en modelos multidimensionales, no normalizados.

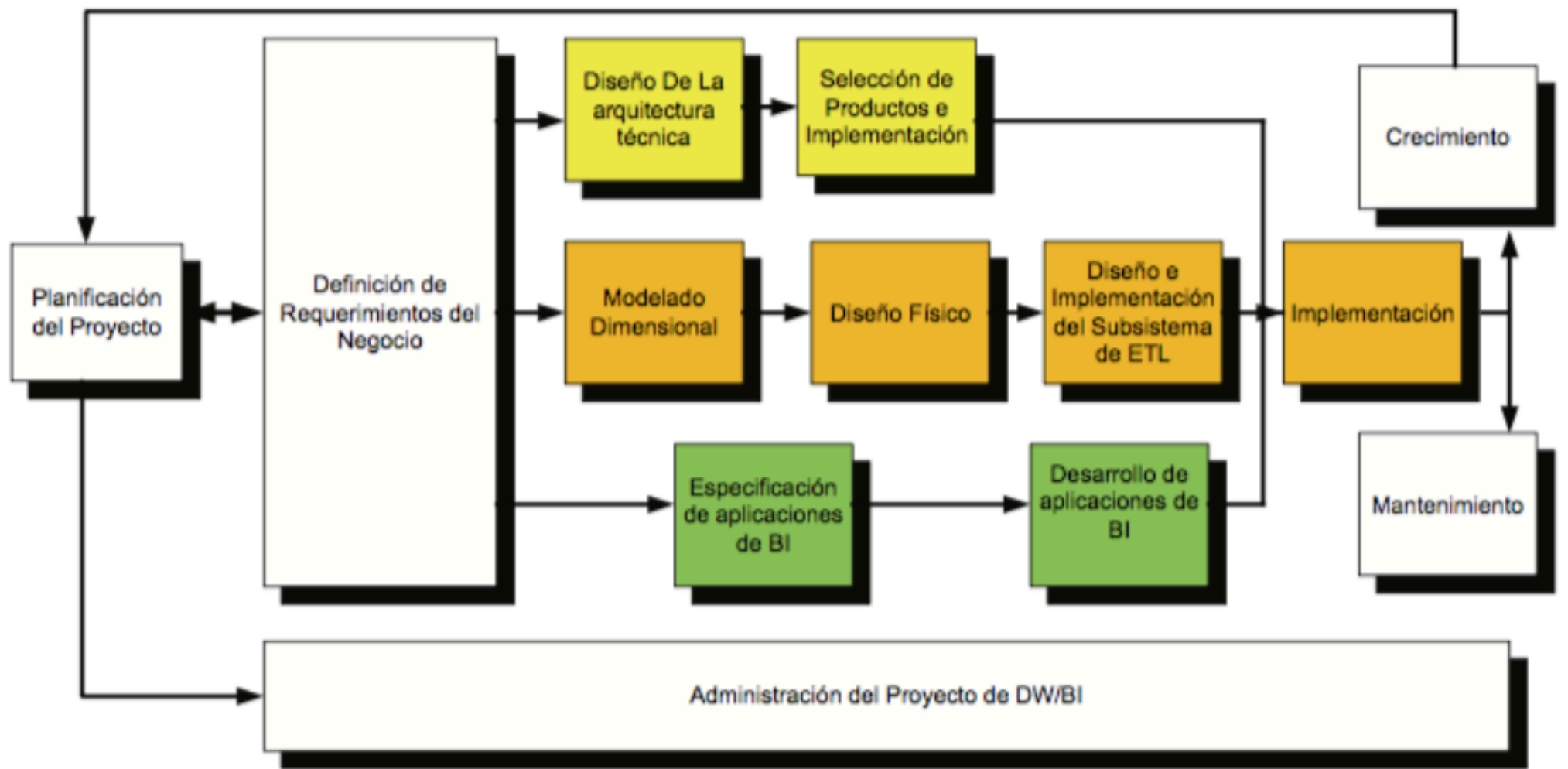
El Ciclo de Vida Dimensional del Negocio de Kimball (Business Dimensional Lifecycle) (Kimball et al 98, 08, Mundy & Thornthwaite 06), está basado en cuatro principios básicos:

- Centrarse en el negocio, núcleo del ciclo de vida de todo proyecto de Data Warehouse, porque concentra su mayor esfuerzo en la identificación de los requerimientos del negocio. Este principio permitió definir claramente el alcance de este proyecto de investigación, además se estableció como regla de negocio el uso de datos abiertos; datos que fueron obtenidos de las bases de datos del Gobierno Autónomo de Buenos Aires que corresponden al año de 2012 y que se encuentran disponibles en la dirección electrónica: <https://data.buenosaires.gob.ar/> del año 2012.
- Construir una infraestructura de información adecuada: Con la finalidad de cumplir con los requisitos establecidos, en este principio se diseña una base de información que sea fácil de usar, única, integrada y de alto rendimiento. Además, permite definir las herramientas, técnicas, modelos que se pueden aplicar para la construcción eficiente y de alto rendimiento del Data Warehouse. La herramienta seleccionada es Power Pivot, que permitió construir el modelo multidimensional usando una arquitectura tipo copo de nieve.
- Realizar entregas en incrementos significativos: con este principio y usando Scrum como metodología ágil se realizaron incrementos de entregables con plazos máximos de 15 días.
- Ofrecer la solución completa: Cuya finalidad es entregar valor a los usuarios de negocios, ofreciendo un almacén de datos que sea integradoras, con alto rendimiento y con datos sólidos. Este principio implica entregar una solución lista para usar inmediatamente (llave en mano); incluyendo funciones de reportes, manuales de usuario y documentación.

Kimball propone una metodología que nos ayuda a simplificar la compleja tarea de construcción de una Data Warehouse. En la siguiente figura se aprecian las tareas que Kiball propone como ciclo de vida.

Figura. 8

Metodología Kimball (Kimball et al 98, 08, Mundy & Thornthwaite 06)

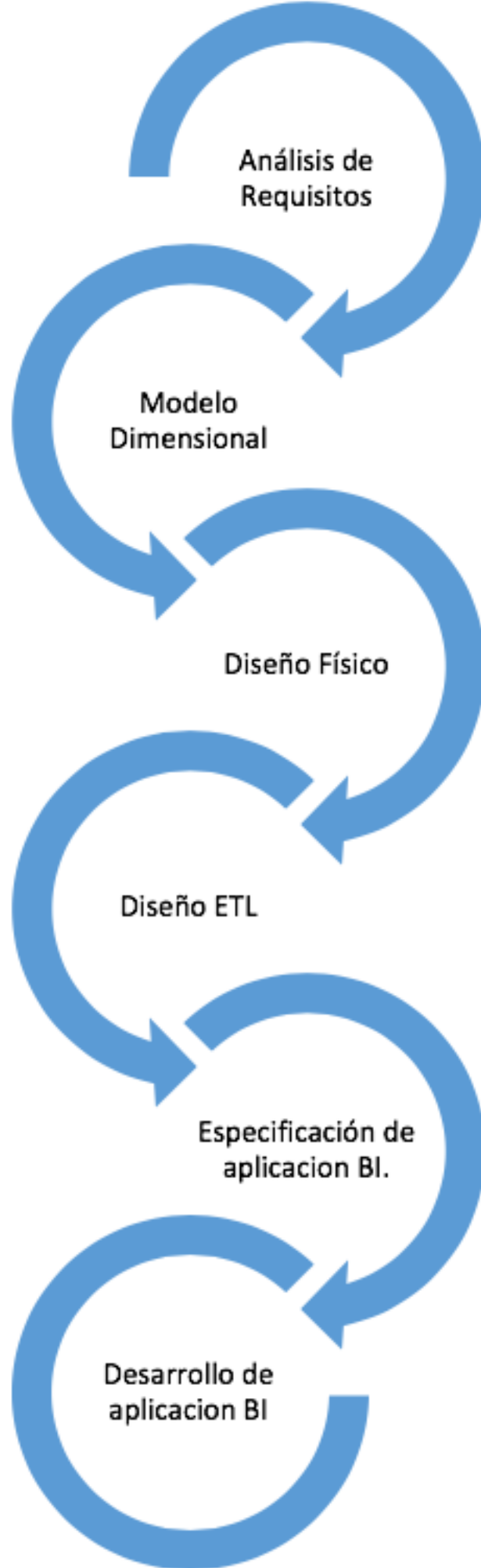


A continuación, se describen las tareas propias de esta metodología:

Planificación:

Las actividades desarrolladas en esta etapa de planificación, se enfocan principalmente en el área de definición del alcance del proyecto, requerimientos, necesidades, tareas, uso de recursos, etc.; tal como lo muestra la siguiente figura.

Figura 9
Planificación del Data Warehouse



Análisis de Requerimientos:

Para esta etapa del proyecto se utilizaron técnicas y herramientas de Recolección datos como: entrevistas, encuestas, cuestionarios, lluvia de ideas; de las cuales se pudieron obtener necesidades y requerimientos puntuales de interés común tales como:

- Conocer de forma precisa los eventos que se presentan en el festival BAFICI.
- Nivel de ventas de los eventos que se desarrollan en el festival de BAFICI.
- Metadata de los eventos, películas, lugares, que se registran del festival BAFICI.
- Valores cuantitativos de proyecciones de eventos, películas, por países, lugares, secciones.
- Fechas de presentación de eventos más concurridos, identificando lugares, países, secciones.

Modelo Dimensional:

Es importante conceptualizar el tema de dimensión, Rivadera G. (2010) menciona que *"Una dimensión es una forma o vista o criterio por medio del cual se puede sumarizar, cruzar o cortar datos numéricos a analizar, datos que se denominan medidas"*. Teniendo en cuenta esta definición, las dimensiones para este proyecto de investigación son: 1. Películas proyectadas. 2. Eventos Realizados. 3. Países en donde se proyectaron películas. 4. Lugares en donde se realizaron eventos. 5. Directores de películas. 6. Secciones en donde se realizaron eventos. Tiempo o fechas en las que se realizaron estos eventos.

Finalmente, se utilizó otra herramienta que ofrece la metodología Kimball, llamada matriz de procesos/dimensiones, tal como se aprecia en la tabla 1.

Tabla 1
Matriz procesos/dimensiones

	Dimensiones						
Proceso de Negocio	Tiempo	Películas	Eventos	Países	Lugares	Directores	Secciones
BAFICI	X	X	X	X	X	X	X
Ventas	X	X	X	X	X		X

Esta matriz permite identificar los procesos del negocio y las dimensiones que intervienen, mostrando de forma clara cómo interactúan la una con la otra, además permite proyectar resultados más precisos, acorde a las necesidades del negocio.

En los siguientes puntos, se detallan los procesos de negocio y las dimensiones que intervienen en el diseño del Data Warehouse:

Elegir el proceso de negocio:

Los dos procesos identificados en este proyecto de investigación son Negocio (BAFICI) y Ventas.

Elegir Dimensiones:

Las dimensiones son: 1. Películas proyectadas. 2. Eventos realizados. 3. Países en donde se proyectaron películas. 4. Lugares en donde se realizaron eventos. 5. Directores de películas. 6. Secciones en donde se realizaron eventos.

Identificar las Tablas de hechos y medidas:

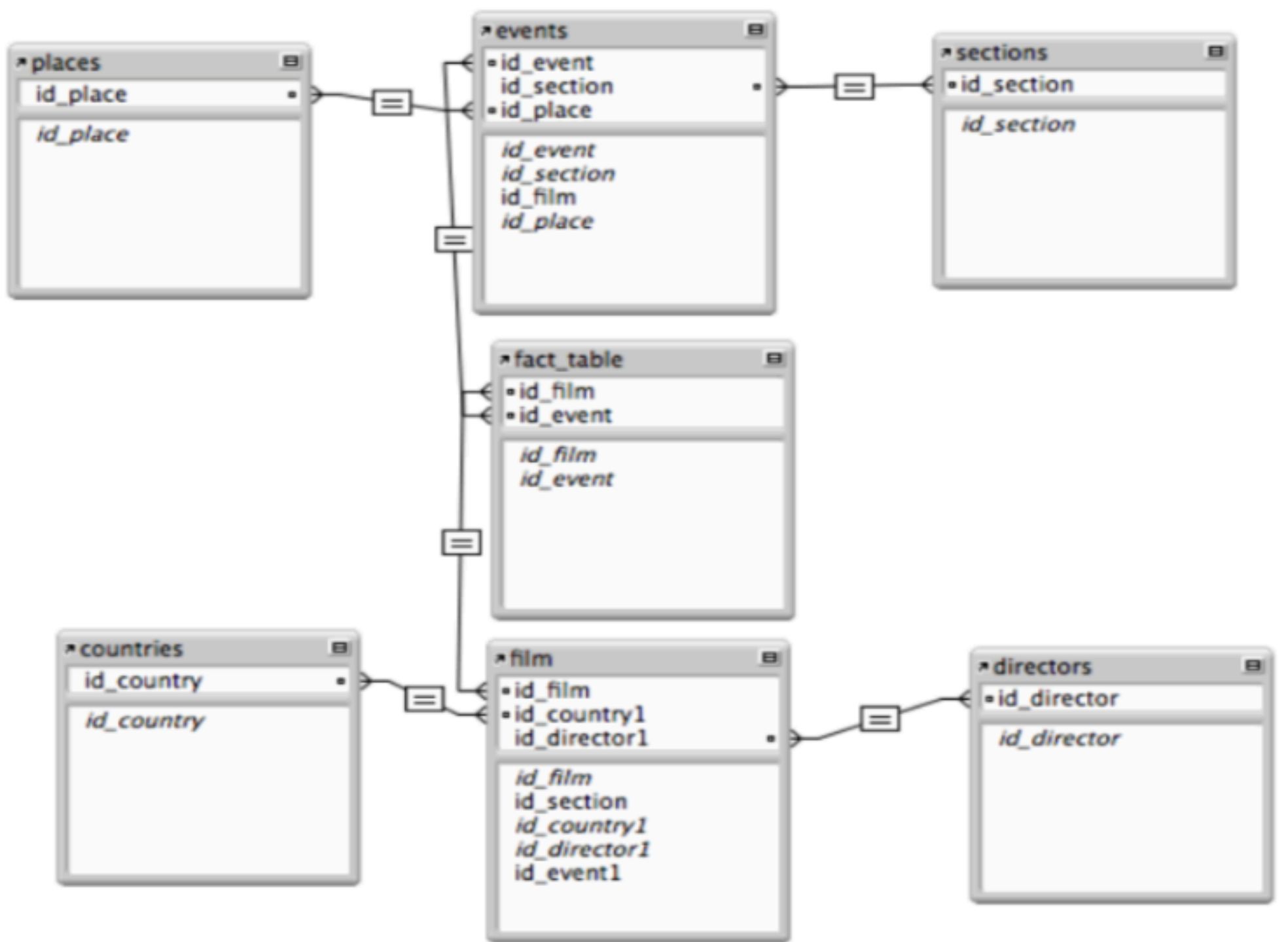
Las medidas son: 1. Cantidad de películas. 2. Cantidad de eventos. 3. Cantidad de directores. 4. Ventas de los eventos. 5. Ventas de eventos por fechas.

La tabla de hechos se conforma de las dimensiones películas y eventos, que se mostrará en el Diseño Físico.

Diseño Físico:

Para el diseño del Data Warehouse, se utiliza un modelo tipo copo de nieve, porque ofrece: mayor flexibilidad al cambio y permite que los datos puedan ser cargados más rápidamente. La figura 10 representa el diseño físico del Data Warehouse, donde se aprecian la integración de la tabla de hecho con las dimensiones.

Figura 10
Diseño Físico



Diseño del sistema de Extracción, Transformación y Carga (ETL).

El proceso de ETL, se realizó de la siguiente manera:

Una vez descargada la información de las dimensiones mencionadas anteriormente, se hizo un análisis del contenido, para establecer las reglas de negocio que permitan definir el alcance que tiene este proyecto, a continuación, se observa la información de cada una de las dimensiones.

Figura 11
Contenido de las dimensiones



Es necesario recalcar que las tablas se descargaron en formato .csv y el análisis de información se hizo utilizando Excel; por este motivo se decidió usar el complemento de esta misma herramienta llamada Power Pivot, para armar el Data Warehouse.

Se procede a aplicar el proceso ETL, el cual consiste en:

Extraer la información de cada una de las dimensiones, hacer el análisis de la información, levantar las reglas de negocio, las cuales son:

- Limitar la cantidad de directores por películas, se tomó esta decisión porque en el análisis realizado se

encontraron que existían algunas películas que tenían más de 4, para efectivizar esta condición en la base de datos se eliminaron los datos de los directores en exceso.

- Uso del Modelo tipo copo de nieve.
- En un evento se pueden presentar máximo cuatro películas.
- Un evento tiene una sección.
- Una película puede ser presentada en máximo cuatro países.

Transformar la información aplicando las reglas de negocio definidas anteriormente, con la finalidad de limpiar datos, eliminar información basura y la depuración de datos no representativos para la muestra de la población.

Load (Cargar) la información en Power Pivot, con la información ya procesada.

Especificación y desarrollo de aplicaciones BI

Constituye la etapa final del proceso de desarrollo del Data Warehouse, cuyo proceso es el siguiente:

- Inicia con la carga de los datos.
- Creación de dimensiones.
- Creación de métricas.
- Creación de fact table.
- Generación de relaciones e identificación de llaves primarias, secundarias.
- Definen que operaciones se pueden realizar con el cubo.
- Finalmente, con Power Pivot se definen las consultas que se pretenden obtener de los datos y al mismo tiempo se establece la forma gráfica en la que se va a mostrar los resultados.

3. Resultados

Producto del levantamiento de la información y la aplicación de la metodología Kimball, incluyendo los procesos y herramientas, en este apartado se presentan las consultas que se realizaron al Data Warehouse construido.

3.1. Relación Eventos-Películas

El resultado de esta consulta permite contabilizar en cuantos eventos se ha presentado una película. También permite conocer los eventos y las veces en las que se han repetido las películas. La siguiente figura muestra lo indicado.

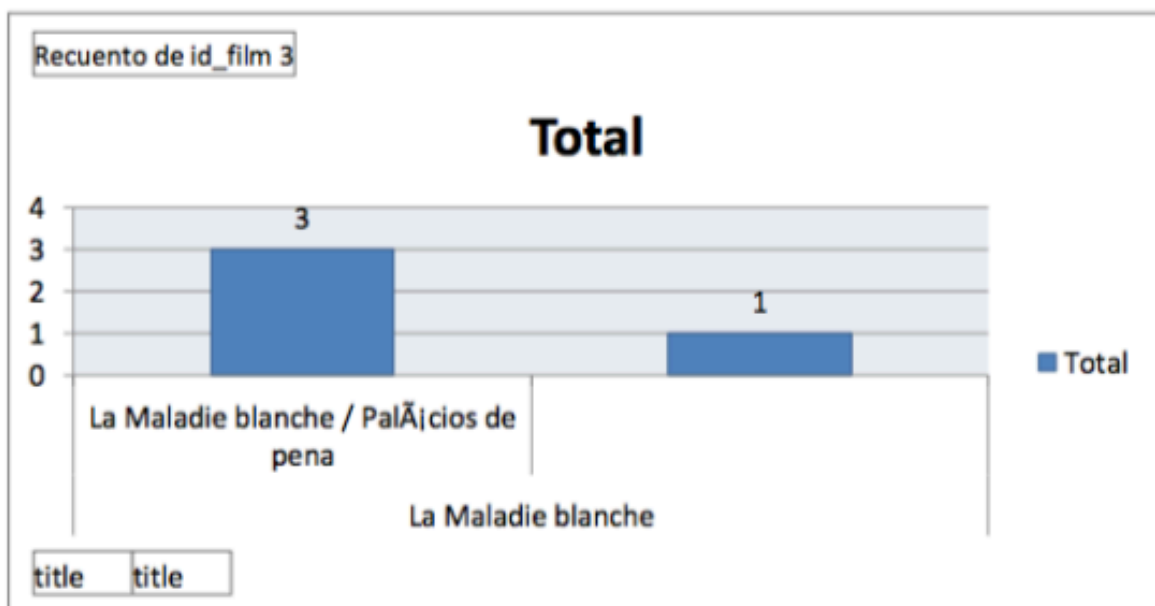
Figura 12

Consulta eventos – películas

id_film					
625	626	627	628	629	630
631	632	633	634	635	636
637	638	639	640	641	642
643	644	645	646	647	648
649	650	651	652	653	654
655	656	657	658	659	660
661	662	663	664	665	666

title
La Maladie blanche
(Mr.) Fred Zentner's Cinema Bookshop...
... (Puntos suspensivos)
17 filles
17 monumentos
17:15hs. Rosa Chérubín
24 passions

Etiquetas de fila	Recuento de id_film 3
La Maladie blanche	4
La Maladie blanche / Palacios de pena	3
	1
Total general	4



3.2. Relación ventas de los eventos.

El producto de esta relación muestra las ventas de los eventos que se han dado por secciones. Además, permite ver si el evento se ha vendido completamente (SI =1) o parcialmente (NO = 0).

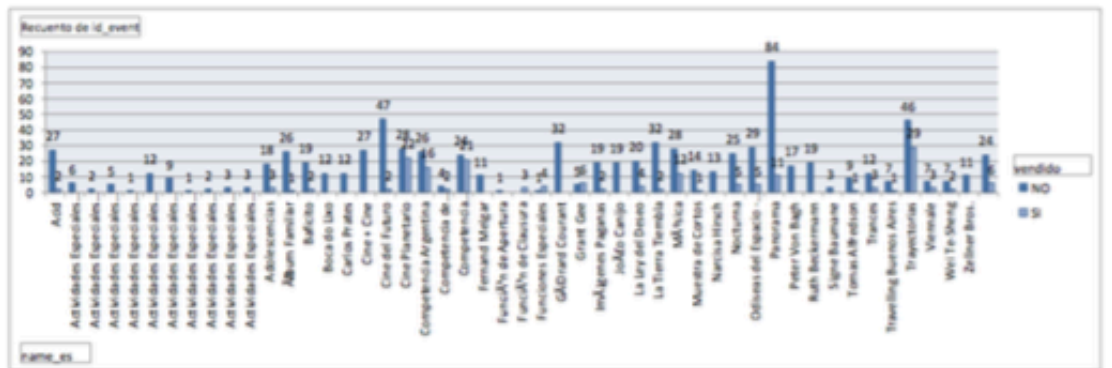
Se puede consultar secciones con eventos con ventas completas (SI=1) o secciones con eventos con ventas parciales (NO =0).

Figura 13
Consulta ventas – eventos

sol... ven...

0 NO
1 SI

id_event		Recuento de id_event		Etiquetas de columna		Total general
Etiquetas de fila		NO	SI			
2244	2245	Acid	27	2	29	
2246	2247	Actividades Especiales / BAFICO al Aire Libre	6	0	6	
2248	2249	Actividades Especiales / BAFICITO al Aire Libre	2	0	2	
2250	2251	Actividades Especiales / Diálogos	5	0	5	
2252	2253	Actividades Especiales / Entrega de Premios	1	0	1	
2254	2255	Actividades Especiales / Mesas	12	0	12	
2256	2257	Actividades Especiales / Pequeña Música Nocturna	9	0	9	
2258	2259	Actividades Especiales / Presentaciones de Concursos	1	0	1	
2260	2261	Actividades Especiales / Presentaciones de libros	2	0	2	
2263	2264	Actividades Especiales / Proyecciones	3	0	3	
2265	2266	Actividades Especiales / Seminarios	3	0	3	
2267	2268	Adolescencias	18	3	21	
2269	2270	Album Familiar	26	1	27	
2271	2272	Baficito	19	2	21	
2273	2274	Boca do Lixo	12	0	12	
		Carlos Prates	12	0	12	
		Cine + Cine	27	0	27	
		Cine del Futuro	47	2	49	
		Cine Planetario	28	22	50	
		Competencia Argentina	26	16	42	
		Competencia de Cortos Argentinos	4	2	6	
		Competencia Internacional	24	21	45	
		Fernand Melgar	11	0	11	
		Función de Apertura	1	0	1	
		Función de Clausura	3	0	3	
		Funciones Especiales	1	4	5	
		GARDard Courant	32	0	32	
		Grant Gee	5	6	11	
		Imágenes Paganas	19	2	21	
		João Canijo	19	0	19	
		La Ley del Deseo	20	4	24	
		La Tierra Tiembla	32	2	34	
		Música	28	12	40	
		Muestra de Cortos	14	1	15	
		Narcisa Hirsch	13	0	13	
		Nocturna	25	5	30	
		Odiseas del Espacio - Cine y Arquitectura	29	5	34	
		Panorama	84	11	95	
		Peter Von Bagh	17	0	17	
		Ruth Beckermann	19	0	19	
		Signe Baumane	3	0	3	
		Tomas Alfredson	9	1	10	
		Trances	12	3	15	
		Travelling Buenos Aires	7	1	8	
		Trayectorias	46	29	75	
		Viennale	7	3	10	
		Wei Te-Sheng	7	2	9	
		Zelner Bros.	11	0	11	
		Zelner Bros.	24	6	30	
		Total general	809	171	980	



3.3. Relación Películas–Directores.

En esta consulta se observan las películas por cantidad de directores. También se puede obtener un resumen de la cantidad de películas que tienen 0, 1, 2, 3 o 4 directores, esta representación se muestra porcentualmente.

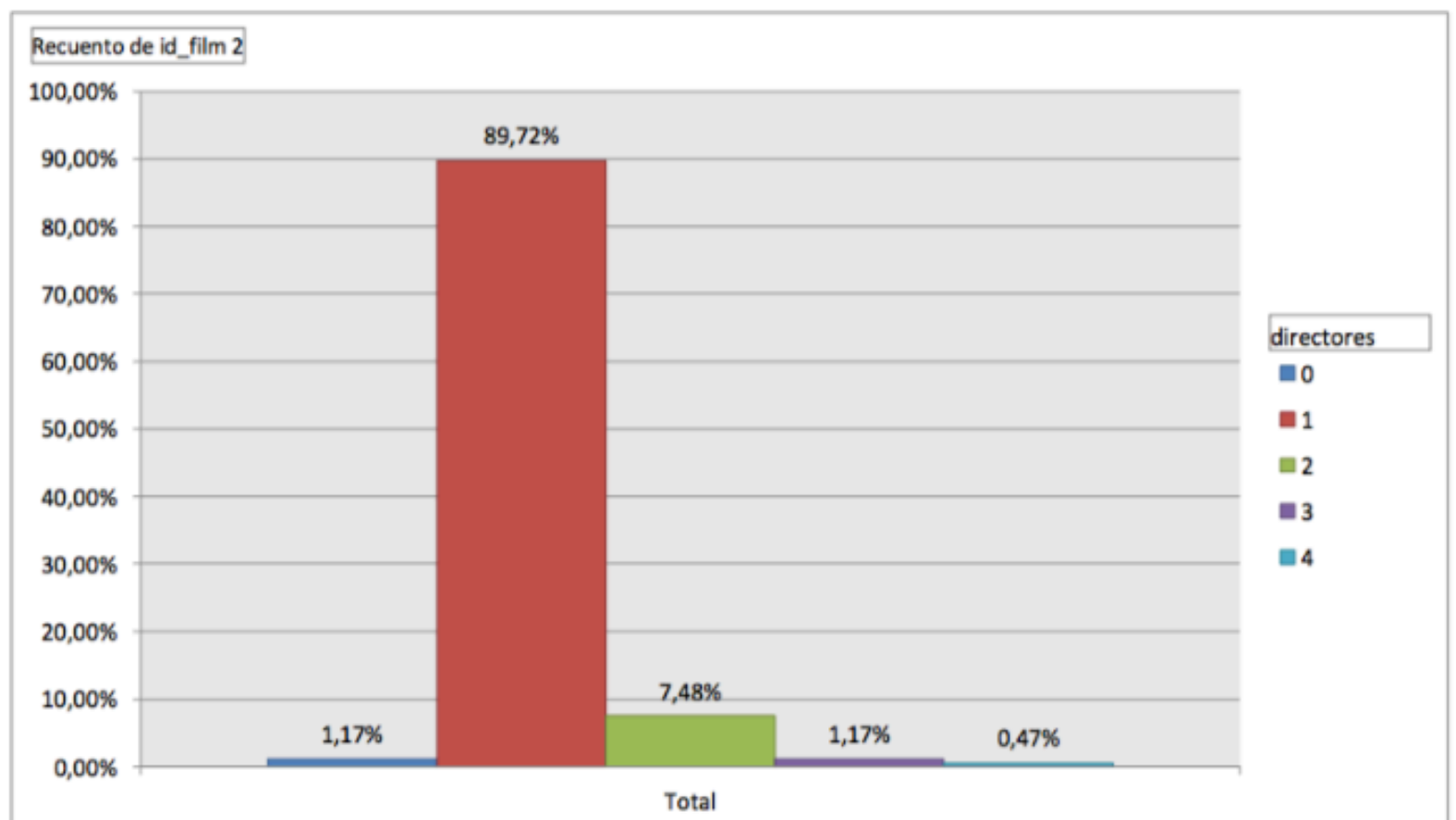
Figura 14
Consulta películas – directores

directores

0 1 2 3

4

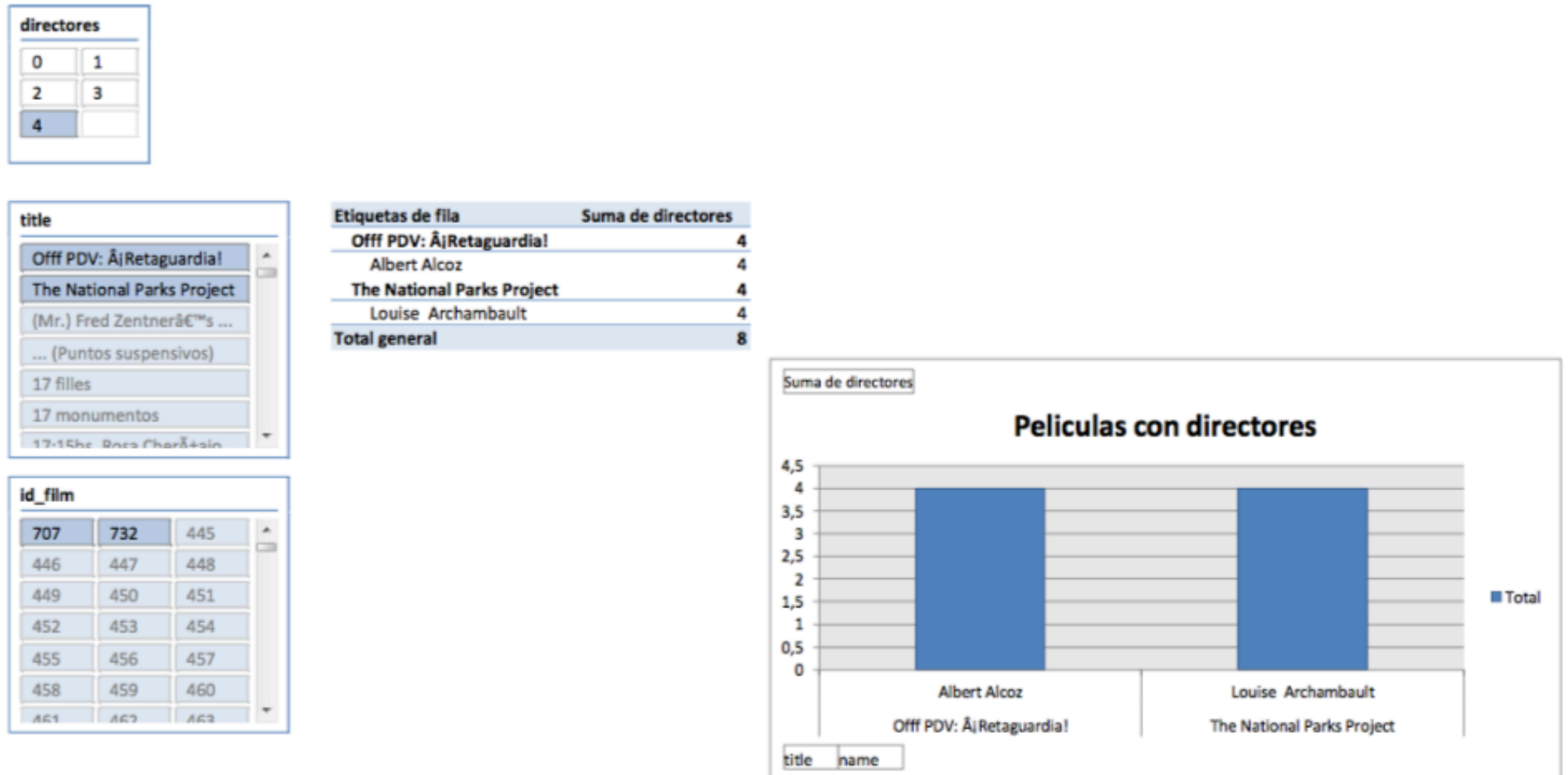
Etiquetas de columna		0	1	2	3	4	Total general
Recuento de id_film 2		1,17%	89,72%	7,48%	1,17%	0,47%	100,00%



La siguiente consulta permite seleccionar una película y ver cuál ha sido su director principal y cuántos directores tiene. En la siguiente figura se configuró un caso específico, que corresponde a seleccionar 4 directores, además de las películas que cumplen esta condición y el director principal de la misma.

Figura 15

Consulta películas – directores principales.



3.4. Relación Películas–País

En esta consulta se puede apreciar la cantidad de películas proyectadas por países. Además, se presenta el ranking de los 10 países con más cantidad de películas proyectadas.

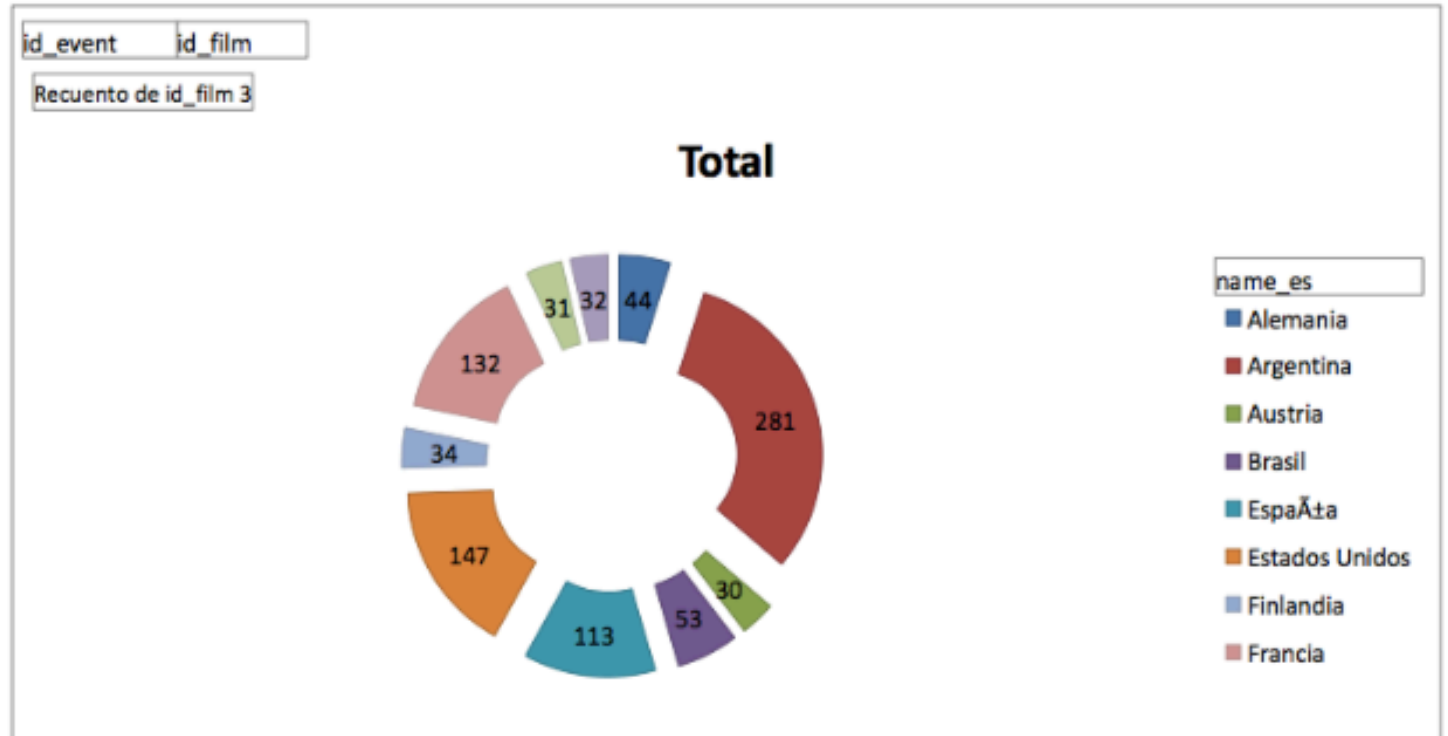
Figura 16

Consulta películas – países

id_film			
0	445	446	447
448	449	450	451
452	453	454	455
456	457	458	459
460	461	462	463
464	465	466	467
468	469	470	471
472	473	474	475
476	477	478	479
480	481	482	483
484	485	486	487
488	489	490	491
492	493	494	495
496	497	498	499
500	501	502	503
504	505	507	508

id_event (Varios elementos)
id_film (Varios elementos)

Etiquetas de fila	Recuento de id_film 3
Alemania	44
Argentina	281
Austria	30
Brasil	53
España	113
Estados Unidos	147
Finlandia	34
Francia	132
Portugal	31
Suiza	32
Total general	897



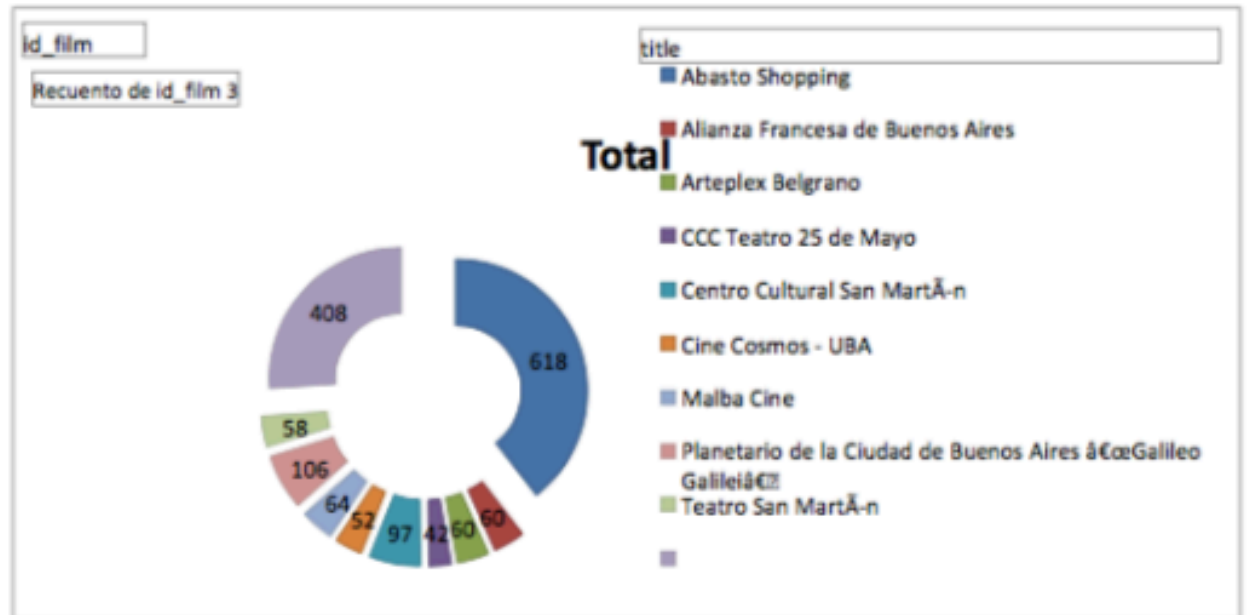
3.5. Relación Películas–Lugares

Esta consulta es similar a la anterior, pero en este caso se usa la dimensión lugares para determinar la cantidad de películas proyectadas por lugares. Además, se puede apreciar el ranking de los 10 lugares con más cantidad de películas proyectadas.

Figura 17
Consulta películas – lugares

id_film						
0	445	446	447	448	449	450
451	452	453	454	455	456	457
458	459	460	461	462	463	464
465	466	467	468	469	470	471
472	473	474	475	476	477	478
479	480	481	482	483	484	485
486	487	488	489	490	491	492
493	494	495	496	497	498	499
500	501	502	503	504	505	506
507	508	509	510	511	512	513
514	515	516	517	518	519	520
521	522	523	524	525	526	527
528	529	530	531	532	533	534
535	537	538	539	540	541	542
543	544	545	546	547	548	549

id_film	(Varios elementos)
Etiquetas de fila	
Abasto Shopping	618
Allianza Francesa de Buenos Aires	60
Arteplex Belgrano	60
CCC Teatro 25 de Mayo	42
Centro Cultural San Mart�n	97
Cine Cosmos - UBA	52
Malba Cine	64
Planetario de la Ciudad de Buenos Aires "Galileo Galilei"	106
Teatro San Mart�n	58
Total general	1565



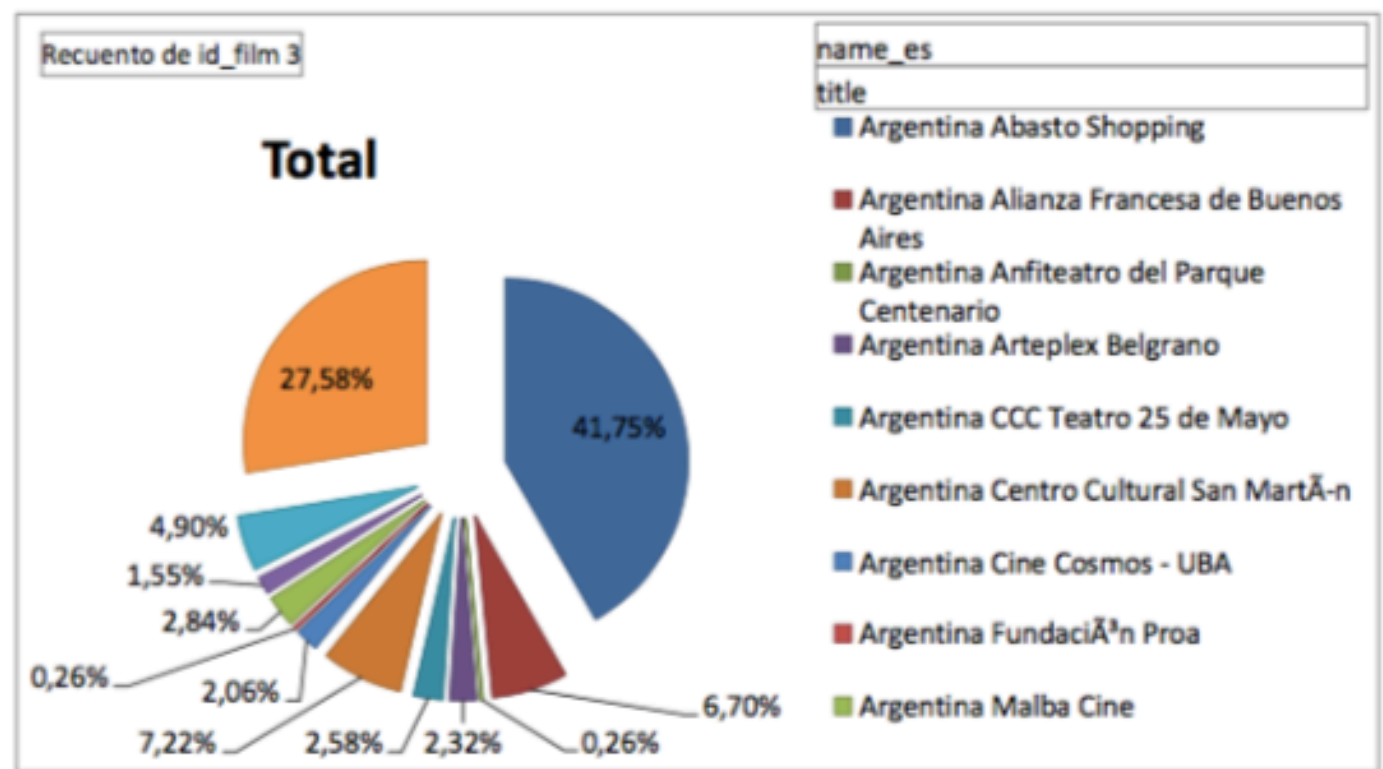
3.6. Relaci n Eventos-Lugares

En esta consulta se puede apreciar la cantidad de eventos proyectados en Argentina y segmentados por lugares; es decir se observan todos los lugares de Argentina en donde se han presentado eventos.

Figura 18
Consulta eventos - lugares

id_event		
0	1329	1336
1341	1344	1346
1352	1367	1375
1377	1378	1386
1389	1392	1404
1407	1409	1412
1421	1428	1439
1440	1445	1449
1450	1452	1454
1456	1457	1460
1470	1477	1479
1484	1485	1489
1492	1510	1511
1512	1518	1521
1534	1537	1538

Etiquetas de fila	Recuento de id_film 3
Argentina	100,00%
Abasto Shopping	41,75%
Alianza Francesa de Buenos Aires	6,70%
Anfiteatro del Parque Centenario	0,26%
Arteplex Belgrano	2,32%
CCC Teatro 25 de Mayo	2,58%
Centro Cultural San Mart�n	7,22%
Cine Cosmos - UBA	2,06%
Fundaci�n Proa	0,26%
Malba Cine	2,84%
Punto de Encuentro BAFICI	1,55%
Teatro San Mart�n	4,90%
	27,58%
Total general	100,00%



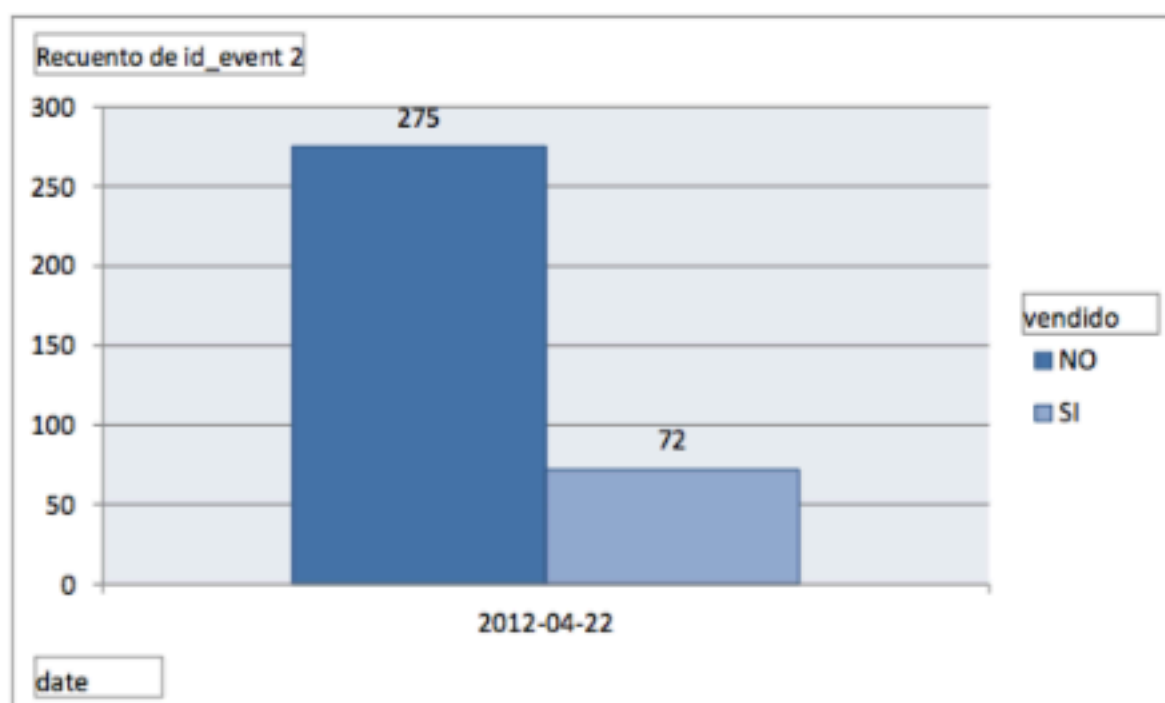
3.7. Relaci n Fecha de Eventos-Ventas

Presenta las fechas en las que se han realizado eventos y si dichos eventos se han vendido completamente (SI) o parcialmente (NO).

Figura 19
Consulta fecha de eventos - ventas.

date
2012-03-23
2012-03-25
2012-04-12
2012-04-13
2012-04-14
2012-04-15
2012-04-16
2012-04-17
2012-04-18
2012-04-19
2012-04-20
2012-04-21
2012-04-22

Recuento de id_event 2		Etiquetas de columna	
Etiquetas de fila	NO	SI	Total general
2012-04-22	275	72	347
Total general	275	72	347



4. Conclusiones

Definir las reglas de negocio es el principio fundamental para crear con éxito un Data Warehouse; hacer un análisis eficiente permite que el Data Warehouse sea robusto, de alto rendimiento con datos sólidos. Obtener reglas de negocio claras y bien definidas, orientan y clarifican los objetivos que se pretenden alcanzar con el Data Warehouse.

En este proyecto de investigación se decidió usar la metodología Kimball, porque nos permitió construir Data Warehouse de forma escalonada, siguiendo un procedimiento similar a la cascada; empezando su proceso desde la fase de diseño y construcción de los Data Marts hasta la construcción del Data Warehouse.

El proceso de Extracción, Transformación y Cargar (ETL) es fundamental para la construcción del DW ya que permite hacer un análisis efectivo de los datos, una limpieza adecuada de la información y se definen procesos de carga y creación de dimensiones para el DW.

Gracias a la aplicación de Open Data, el Gobierno Autónomo de Buenos Aires pone a disposición, en forma libre y sin restricciones, bases de datos para que los investigadores hagan buen uso de ellas. Este proyecto de investigación resulta del análisis de dichos datos disponibles en el sitio web oficial.

El uso de la herramienta Power Pivot, facilitó la construcción del Data Warehouse en este escenario; permitiendo obtener resultados óptimos que se pueden visualizar amigablemente en tablas y gráficas precisas.

Referencias bibliográficas

A. Agostino, K. S. Soilen y B. Gerritsen, (2013). Cloud solution in Business Intelligence for SMEs—vendor and customer perspectives. Disponible en:

https://www.researchgate.net/publication/268151560_Cloud_solutions_in_Business_Intelligence_for_SMEs_vendor_and_customer_perspectives

CANO, José; GOMEZ, Rodrigo; SALAZAR, Fernando. **Routing policies in multi-parallel warehouses: an analysis of computing times.** Revista Espacios. 38 (Nº 51) Año 2017. Page 23. Recuperado de: <http://www.revistaespacios.com/a17v38n51/17385123.html>

CHOWDHURY, Sandip (2014) Aumento de almacén de datos, Parte 1: Big data and data warehouse augmentation. IBM Corporation Disponible en: <https://www.ibm.com/developerworks/ssa/library/ba-augment-data-warehouse1/ba-augment-data-warehouse1-pdf.pdf>

GIRALDO, Juan Camilo; JIMÉNEZ, Jovani; TABARES, Marta Silvia. Modelo para optimizar el proceso de gestión de negocio combinando minería de procesos con inteligencia de negocios desde almacenes de datos. Revista Espacios. Vol. 38 (Nº 02) Año 2017. Pág. 9. Recuperado de: <http://www.revistaespacios.com/a17v38n02/17380209.html>

INMON, Bill (2010). Data Warehousing 2.0 Modeling and Metadata Strategies for Next Generation Architectures. Forest Rim Technology, LLC Disponible en: <http://devgear.co.kr/pdf/bill-inmon-data-warehousing-2-0-whitepaper.pdf>

Kimball & Ross. (2010). The Kimball Group Reader; Relentlessly Practical Tools for Data Warehousing and Business Intelligence, Indianapolis, Wiley.

Leonard Brizuela, E., & Castro Blanco, Y. (2013). Metodologías para desarrollar Almacén de Datos. Revista de Arquitectura e Ingeniería, 7 (3), 1-12

RAMIREZ, Catalina; OLIVIERA, Renato. **Avaliando sistemas de inteligência de negócios em uma empresa farmacêutica: um estudo de caso.** Revista Espacios. Vol. 35 (Nº 7) Año 2014.

Pág. 2. Recuperado de: <http://www.revistaespacios.com/a14v35n07/14350702.html>

Rivadera, G. R. (2010). La metodología de Kimball para el diseño de almacenes de datos (Data Warehouses). Cuadernos de la Facultad, 56-71.

SANTOS, Bruno; STEINER, Maria; RAMOS Luis Enrique; MARTINS Luis; ANDRADE Pedro. **Data Mining: A theoretical approach and its applications.** Revista Espacios. Vol. 37 (Nº 05) Año 2016. Pág. 23.

Recuperado de:

<http://www.revistaespacios.com/a16v37n05/16370523.html>

SCHEIDL Herbert; SIMON, Alexandre; CAMPOS Fernando. **Conhecimento Strategic and operational Information Management: the logistic operators relationship management from the perspective of Business Intelligence and knowledge management.** Revista Espacios. Vol. 37 (Nº 11) Año 2016. Pág. 11. Recuperado de:

<http://www.revistaespacios.com/a16v37n11/16371111.html>

<http://www.revistaespacios.com/a16v37n11/16371111.html>

Shahzad, K., & Zdravkovic, J. (2011, May). Towards goal-driven access to process warehouse: Integrating goals with process warehouse for business process analysis. In *Research Challenges in Information Science (RCIS), 2011 Fifth International Conference on* (pp. 1-11). IEEE.

SOUZA, Jovani; FRANCISCO Antonio; KOVALESKI, Luis; OLIVEIRA, Bruno; BATISTA Álamo; CANTERI, Maria. Knowledge creation through Data Mining: A study of association rules in a Weka database. Revista Espacios. Vol. 37 (Nº 06) Año 2016. Pág. 24. Recuperado de:

<http://www.revistaespacios.com/a16v37n06/16370624.html>

VANTIO. Proposed reconfiguration of strategic financial business intelligence module in the context of governance of information technology (GTI): a study in plastic industry sector. Revista Espacios. Vol. 32 (1) 2011. Pág. 41. Recuperado de: <http://www.revistaespacios.com/a11v32n01/113201171.html>

1. Ingeniero en Sistemas Informáticos y Computación. Magister en Tecnologías de la Información - Grupo de Investigación UIDE. Escuela de Informática y Multimedia. Universidad Internacional del Ecuador. wabustamantegr@uide.edu.ec

2. Ingeniera en Sistemas. Magister en Evaluación y Auditoría Informática. Grupo de Investigación UIDE. Informática y Multimedia. Escuela de Informática y Multimedia. Universidad Internacional del Ecuador. esmacasru@uide.edu.ec

3. Licenciada en Análisis de Sistemas. Diplomado en Gerencia Estratégica en Desempeño Empresarial. Departamento de Ciencias de la Computación y Electrónica. Universidad Técnica Particular de Loja. fbcevallos@utpl.edu.ec

Revista ESPACIOS. ISSN 0798 1015
Vol. 39 (Nº 34) Año 2018

[Índice]

[En caso de encontrar un error en esta página notificar a webmaster]

©2018. revistaESPACIOS.com • @Derechos Reservados