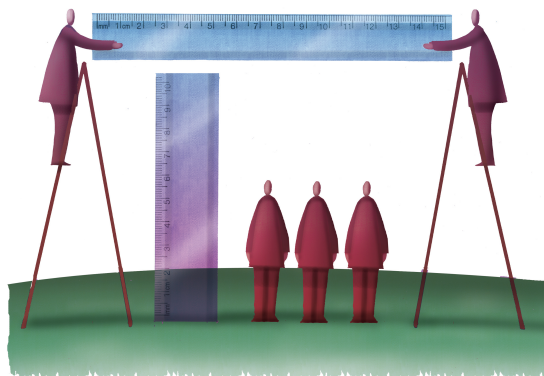


DinoParques: renovarse o morir

Los sistemas que utilizaba la empresa para llevar los datos se habían quedado casi en la era jurásica: era necesario implementar cambios para no seguir desaprovechando oportunidades.



Con cinco años en el mercado de la entretención educativa, la empresa “Parques Temáticos S.A.” opera ya ocho parques a lo largo del país, con el nombre de fantasía “DinoParques”. Estos están enfocados a dar a conocer a los escolares las maravillas del mundo jurásico. Gran parte de la operación se realiza a través del sistema escolar, tanto público como privado, con excursiones guiadas, pero durante el verano también están abiertos al público general.

Aprovechando el éxito del *merchandising* relacionado con los dinosaurios – derivado del éxito de películas y videojuegos- la empresa logró desarrollar fuertemente su unidad de negocios de venta de productos y comidas dentro de los parques. Esto, junto a la venta de

entradas representa más del 90% de los ingresos de la compañía.

Pero en cada reunión gerencial surgen los mismos problemas: los datos sobre las ventas de entradas y productos no parecen confiables ni están listos en tiempo y forma para que el equipo de gerentes pueda tomar decisiones acertadas. Esto ocurre porque los distintos gerentes (ventas, marketing, control de gestión, etc.) se manejan con diferentes fuentes de información, en algunos casos usando planillas de cálculo hechas con ingreso manual de datos sólo minutos antes de las reuniones. Y además, porque cada vez que se requiere un informe de mediana complejidad, hay que solicitarlo al área de sistemas que demora por lo menos una semana en confeccionarlo ya que están demasiado ocupados con la migración de los servidores corporativos y cuentan con poco personal.

En resumen: cada reunión es un caos por las diferentes versiones de la verdad que cada gerente tiene y, en algunos casos, porque ni siquiera se llega a tener la información necesaria o está demasiado desactualizada.

Diseñando el sistema

Debido a esta situación, “Parques” decide contratar a la consultora “Modelling” para ayudarla a modernizar los sistemas y mejorar el flujo de información. El directorio ha escuchado hablar de palabras como sistemas de ayuda a la decisión, Business Intelligence (BI), DataWarehouse y otras, pero lo que realmente necesita es una solución de negocio para afianzar el crecimiento de la empresa.

En principio se quiere trabajar con la información de ventas, tanto de entradas como de productos dentro de los parques, para luego ir extendiendo el alcance del proyecto a otras áreas.

Luego de realizar un relevamiento inicial, se llega a los siguientes requerimientos:

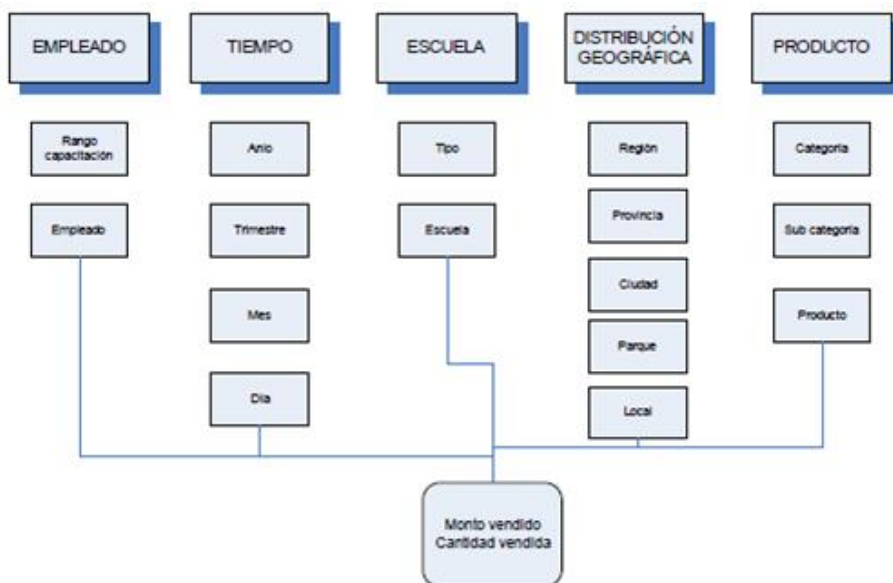
1. Los orígenes de datos que alimentarán el DataWarehouse serán:
 - a. El sistema emisor y facturador de entradas, que está presente en todas las boleterías de cada parque. (Ver anexo I Material de Profundización)
 - b. Dos sistemas distintos de facturación de productos y comidas para los locales de los parques (ya que los más nuevos tiene un sistema que se contrató hace un año a una consultora, pero los parques más antiguos tienen todavía un sistema desarrollado por el sobrino de uno de accionistas de la empresa). (Ver anexo II Material de Profundización)
 - c. El sistema de administración de RRHH y liquidación de haberes. (Ver anexo III Material de Profundización)
2. Se detecta que los analistas de negocio de la empresa necesitan analizar la información de ventas de un producto o de las entradas desde diferentes perspectivas. Por ejemplo, la venta de entradas vista por mes o por año, por parque, por región, por empleado y demás combinaciones entre las perspectivas.
3. El mínimo nivel de detalle que se quiere tener disponible para el análisis de las ventas (monto vendido y unidades

ventas) es el de la línea de los tickets.

4. Es necesario poder hacer análisis de ventas de cada empleado considerando también el impacto de las horas de capacitación recibidas.
5. Es necesario conocer también de qué manera influye, en las ventas de productos, la zona geográfica en la que están ubicados los parques.
6. De cada escuela se sabe si es privada o pública, y se desea conocer cuáles son las escuelas que generan mayores ingresos a la empresa.
7. Se necesitará hacer análisis diarios, mensuales, trimestrales y anuales.

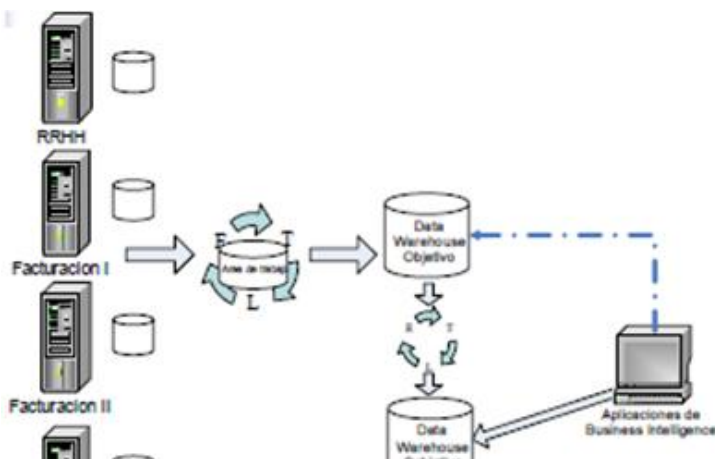
Abordando el caso

“Modelling” decide empezar por modelar el negocio de DinoParques, identificando las dimensiones de análisis y también los niveles. Con esto se llegó al siguiente Modelo multidimensional conceptual.



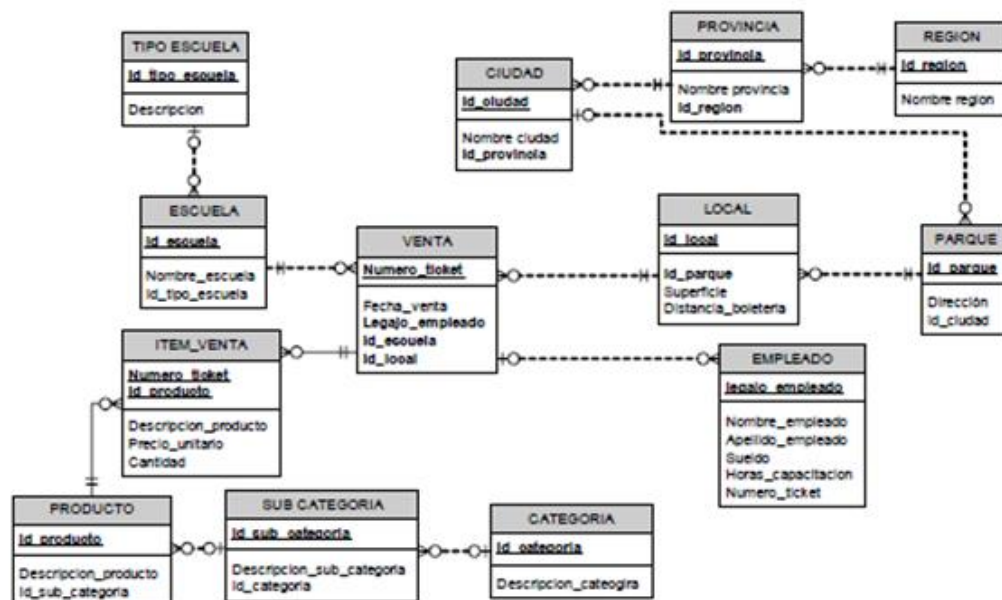
Luego, se debía definir el modelo de hardware y las fuentes que alimentarán al DataWarehouse. De acuerdo a los orígenes de datos detectados, las fuentes serán: RR.HH. Facturación I, Facturación II y Boleterías. Además, se deciden crear dos tipos de DataWarehouse: Objetivo y subjetivo:

- **DHW Objetivo:** Contendrá todos los datos transaccionales de las diversas fuentes
- **DWH Subjetivo:** Será un DataMart que posea la información agrupada y relevante para el análisis de BI.

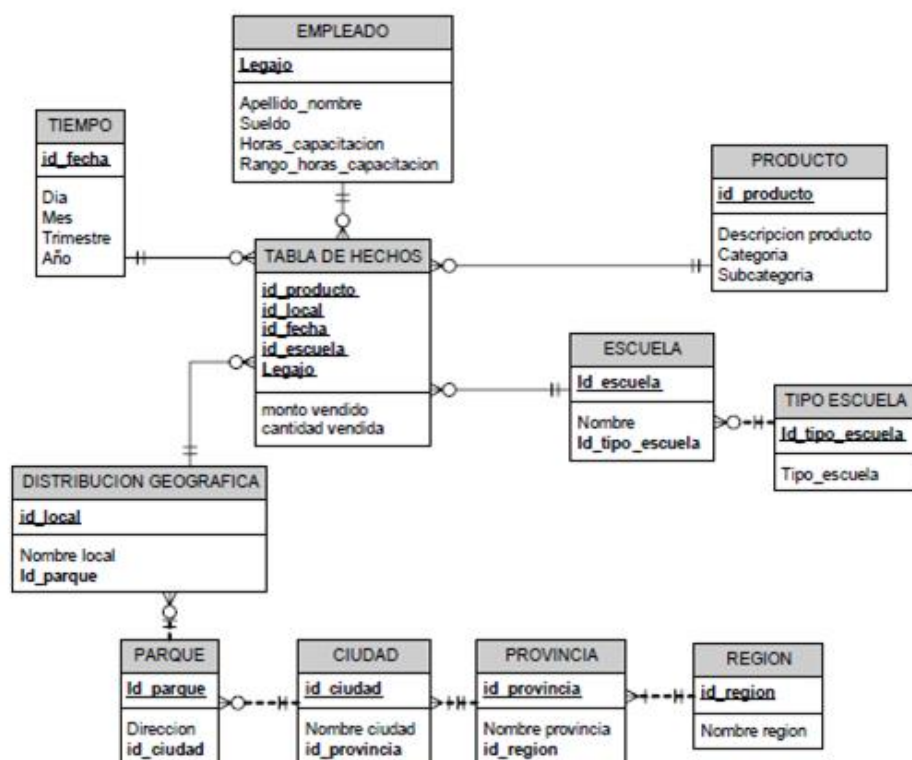




Con el diseño multidimensional conceptual y los modelos de datos de las fuentes se diseña el DataWarehouse objetivo:



También se diseña el DataWarehouse subjetivo, mediante un modelo mixto llamado “copo de nieve” (snowflake)



Extraer, transformar y cargar

Modelling explica que el diseño del DataWarehouse Subjetivo dependerá del análisis de datos que quiera orientar el área de Business Intelligence. Por ello, este diseño está **propuesto** y puede modificarse a criterio del negocio.

Finalmente, se debe tomar en cuenta la periodicidad en la carga de datos al DataWarehouse. Para ello se propone un proceso de ETL (**Extract, Transform and Load**) que extraiga los datos de los sistemas fuente, los transforme si es necesario, y luego los cargue.

Para ello, se considera la decisión basándose en la necesidad del negocio (con qué frecuencia es necesario tener datos actualizados para el análisis) y, también, la disponibilidad de los recursos, tanto las fuentes de datos como el DataWarehouse, ya que verán afectado su rendimiento en gran medida durante su ejecución. Inclusive en el DataWarehouse no estará disponible durante el proceso.

En base a ello se decide correr el proceso de ETL para la carga del DataWarehouse una vez por semana - extrayendo todas las transacciones de la semana anterior- los días domingo a las 22 horas. Esta decisión se toma porque este nivel de actualización es suficiente para los analistas y, además, porque ese es un momento en que hay disponibilidad de recursos de procesamiento. También se toma en cuenta que la mayoría de los análisis se hace el día lunes con la información de la semana anterior.

A modo de ejemplo, la empresa consultora describe tres de las transformaciones que realizará el proceso ETL:

a) Categorías y subcategorías de productos: en los tres sistemas en los que se usan tienen campos con nombres distintos. Los mismos serán integrados en el DataWarehouse en una única tabla, dejando un único valor para cada categoría y subcategoría de los productos en los sistemas de facturación I y II (ya que existe la misma codificación en ambos sistemas) y se inventarán subcategorías ficticias para cada categoría del sistema de entradas, ya que en dicho sistema no existe el concepto de subcategoría.

b) Cada sistema de facturación cuenta con su propia codificación de empleados y escuelas: esto significa, por

ejemplo, que un mismo código de empleado puede ser usado para dos empleados distintos de un parque a otro. Esto debe ser solucionado mediante el proceso de ETL, usando tablas de conversión o equivalencia.

b.1. En el caso de los empleados:

Se usará una tabla de equivalencias, donde se unificará la codificación usando el número de legajo del sistema de RRHH.

Tabla de equivalencias

ID_EMPLEADO Sistema facturación I	COD_EMPLEADO Sistema facturación II	Codigo_empleado sistema de entradas	LEGAJO
1	54	5	1
2	6	8	70
3	1	2	41
4	10	34	53

Esta tabla sirve para ser usada al momento de la extracción de las ventas de los diferentes sistemas de facturación. En cada venta se extrae el código del empleado del sistema fuente y, luego de consultar la tabla de equivalencias y obtener el legajo correspondiente, podrá cargar el registro de la venta en forma homogénea en el DataWarehouse.

b.2. En el caso de las escuelas:

La solución es similar, usando la siguiente tabla de conversión entre el sistema fuente y el DataWarehouse, pero es un poco más sencillo ya que los sistemas de facturación I y II tienen los mismos códigos, por lo tanto solo es necesaria una tabla que establezca las equivalencias con el sistema de entradas. En este caso se deja como codificación base en el DataWarehouse al de los sistemas de facturación I y II

Tabla de equivalencias

COD_ESCUELA de sistema de entradas	ID_ESCUELA Sistemas I y II
1	24
2	21
3	23
4	45

Con todo esto, el sistema quedó diseñado para cumplir los requerimientos definidos y de forma suficientemente flexible para ir incorporando nuevos requerimientos, con lo cual se solucionaron los problemas básicos de información que tenían los gerentes.

Ver anexos en Material de Profundización