



# Bases en el diseño e implantación de un proyecto de Análisis de Datos y de Inteligencia de Negocio. (borrador)

Análisis de datos e inteligencia de negocio (Ad y In)  
Data Analysis and Business Intelligence (DA and BI)

*Sobre el diseño de un modelo de datos. Cómo comenzar un proyecto de Business Intelligence*

*Presentamos a continuación una serie de consideraciones a tener en cuenta en la fase inicial de desarrollo de un proyecto de Business Intelligence (BI) y donde la base crítica está en el adecuado diseño del modelo de datos que sustenta el proyecto BI.*

1. Un modelo basado en preguntas
2. Análisis de los datos necesarios y disponibles. Diseño del modelo estrella
3. Presentación de la dimensión calendario y su incorporación al modelo
4. El proceso ETL Power Query (PQ)
5. Fases y conceptos a tener en cuenta en un proyecto de análisis de datos y BI

## Un modelo basado en preguntas

Siempre un modelo de datos dirigido a la inteligencia de negocios debe ser el resultado de analizar y reflexionar sobre preguntas vinculadas con la empresa, área de negocio, departamento o sector y a la que debe dar respuesta los informes propuestos.

Es fundamental el hablar con el cliente, conocer los objetivos de las personas y sus necesidades, los KPI que utilizan (la rentabilidad de los productos o gama, análisis de mercados, de clientes, etc.), que indicadores son importantes, todo ello es la clave para el éxito en el desarrollo un proyecto de inteligencia de negocio

En resumen, si vamos a diseñar un modelo de datos lo fundamental es conocer y entender cuál es el objetivo, cuáles son las preguntas del negocio, es decir que debe contener el informe que debemos diseñar, indicadores KPis, etc.

*Muchos consultores con amplia experiencia comienzan esta primera etapa con una hoja en blanco sin abrir Excel ni Power BI e intentan pensar en cómo será el modelo de datos intentando identificar y conocer el proceso de negocio que se quiere analizar e identificar todos los filtros o dimensiones asociados.*

Este análisis previo va a determinar los datos que necesitamos y confirmar si existen para dar respuestas a las preguntas de negocio lo que supone un conocimiento de la infraestructura técnica con que cuenta la empresa (CRM, ERP, etc.) y de que información disponemos y tenemos acceso para incorporar a nuestro modelo de datos a través de una importación o vinculación y en caso necesario de un proceso una depuración y adaptación.

Señalar que es recomendable conocer los aspectos técnicos básicos relacionados con las bases de datos, como son las consultas, tipo de campos, estructuras, etc.. que están íntimamente relacionados con el proceso de diseño y construcción de los modelos de datos.

Así partiendo de que información disponemos (CRM, ERP, etc.) y tenemos acceso para incorporar a nuestro modelo de datos y siguiendo la estructura del modelo estrella debemos identificar los dos componentes siguientes:

- **Tabla de hechos** que viene a contener el procesos de negocio (por ejemplos las ventas, atenciones telefónicas, producción, etc..) es decir cualquier hecho que queramos analizar y que normalmente suceden en un momento determinado y que contienen valores que son los que deseamos analizar.
- **Tablas de dimensión** viene a ser como queremos analizar ese proceso, por ejemplo el proceso de ventas, la quiero analizar por producto, por cliente, por fecha, empleado, sucursal, etc.. esas serían las dimensiones, como vemos contiene las preguntas que queremos estudiar sobre el proceso, es decir las dimensiones vienen a ser como las preguntas que queremos hacer sobre ese proceso, quien hizo ese proceso, cuando hizo ese proceso, como lo hizo, donde, porque, etc.. esas serían las tablas de dimensión.

Por tanto nuestro modelo de datos tipo estrella va a estar configurado por una tabla de hechos y una o varias de dimensión que estarán relacionadas y que luego cualquier pregunta de negocio relacionados con esos procesos se pueda responder.

De esta forma todos los informes partan de ese modelo final tipo estrella.

En resumen, lo primero que debemos hacer es identificar el proceso de negocio y en segundo lugar todos los filtros que nos van a permitir analizar y dar respuestas a las preguntas detectadas en el modelo de datos.

### Presentación de la dimensión calendario y su incorporación al modelo

Destacar que en cualquier modelo de datos debe existir la **dimensión calendario** y para ello debemos considerar el disponer como mínimo de esta tabla en nuestro modelo de datos, es fundamental, ya que nos va a permitir analizar por cualquier dimensión de tiempo, como semestre, año, semana, día de la semana, etc.. aplicar funciones de inteligencia de tiempo con DAX (time intelligence dax), como son comparativas de tiempo, mes actual con respecto al mes anterior, semana actual con respecto a la misma semana del años pasado, etc...

## El proceso ETL Power Query (PQ)

Como hemos señalado los datos que nos interesan pueden estar (CRM, ERP, etc.) en una base de datos, en una hoja de cálculo, en una página web, en un fichero de datos, pero en cualquier caso una vez localizados necesitamos importarlos y tratarlos o adaptarlos para incorporar al modelo de datos, es decir conectar esos orígenes (importación y/o vinculación), vamos a transformarlos y a modelar y a este proceso es el que se le denomina ETL que significa Extraer, Transformar y Cargar (Extract, Transform, Load).

### Fuentes de Datos



Debemos evitar comenzar un proyecto conectando los datos y cargando las distintas fuentes de información tal y como están, es indispensable en la mayoría de los casos un proceso de transformación de los datos, pero no solo de limpieza (quitar columnas y filas, ajustar encabezados, remplazar un dato por otro, formateo de los campos, etc..) sino además cambiar la estructura de las tablas con procesos de consulta y consolidación de tablas, estableciendo relaciones, etc.. y para este proceso de depuración y transformación contamos con Power Query que es la herramienta fundamental del proceso ETL.

Es decir si tenemos 2 tablas o 20 tablas no las vamos a cargar directamente al modelo sino con PQ las vamos a limpiar y transformar e incluso en algún caso debemos consolidar algunas tablas y relacionar y así finalmente cargar el resultado de una tabla combinada al modelo de datos para construir posteriormente las métricas con DAX

## 1. Presentación del caso. Estudio del caso

- a) Análisis de la empresa y en concreto del proceso de negocio objeto de estudio
- b) Recursos necesarios y disponibles

## 2. El proceso ETL con Power Query. Preparando los datos y modelado

- a) Conectar y/o agregar fuentes de datos (tablas, ficheros planos, etc)
- b) Transformar y limpiar los datos con Power Query
  - Depuración y adaptación
  - Preparando datos, eliminando columnas y filas.
  - Jerarquías
- c) Modelado. Esquema Estrella
  - Identificación de Tabla de Hechos y de Dimensiones.
  - Tabla de calendario
  - Conexión y relaciones de tablas de datos

## 3. Creación de Medidas. Introducción a DAX

- a) Medidas y métricas básicas.
- b) Funciones básicas de inteligencia de tiempo
- c) Otras funciones.

## 4. Visualización de datos

- a) Listas, tarjetas, matriz y gráficas
- b) Visualizaciones de KPI y Medidor
- c) Aplicando filtros de tarjeta slicers
  - Interacciones y sincronizar segmentadores.
- d) Navegación de páginas