

Estrategia y Gestión Empresarial basada en datos

Profesor de la asignatura: Gerardo Herranz

Sesión 9: Procesos ETL Datawarehouse

Temario

Clase de hoy

- ▶ Tema 1. La función empresarial y el valor de los datos
- ▶ Tema 2. Inteligencia de negocios: contenido y alcance
- ▶ Tema 3. El papel de los sistemas de información en la inteligencia de negocio
- ▶ Tema 4. Aspectos éticos y regulatorios en materia de seguridad y de privacidad
- ▶ **Tema 5. Datawarehouse o almacén de datos**
- ▶ Tema 6. La Dirección estratégica
- ▶ Tema 7. Nuevos perfiles profesionales en las corporaciones data driven
- ▶ Tema 8. La privacidad y el análisis masivo de datos en la práctica



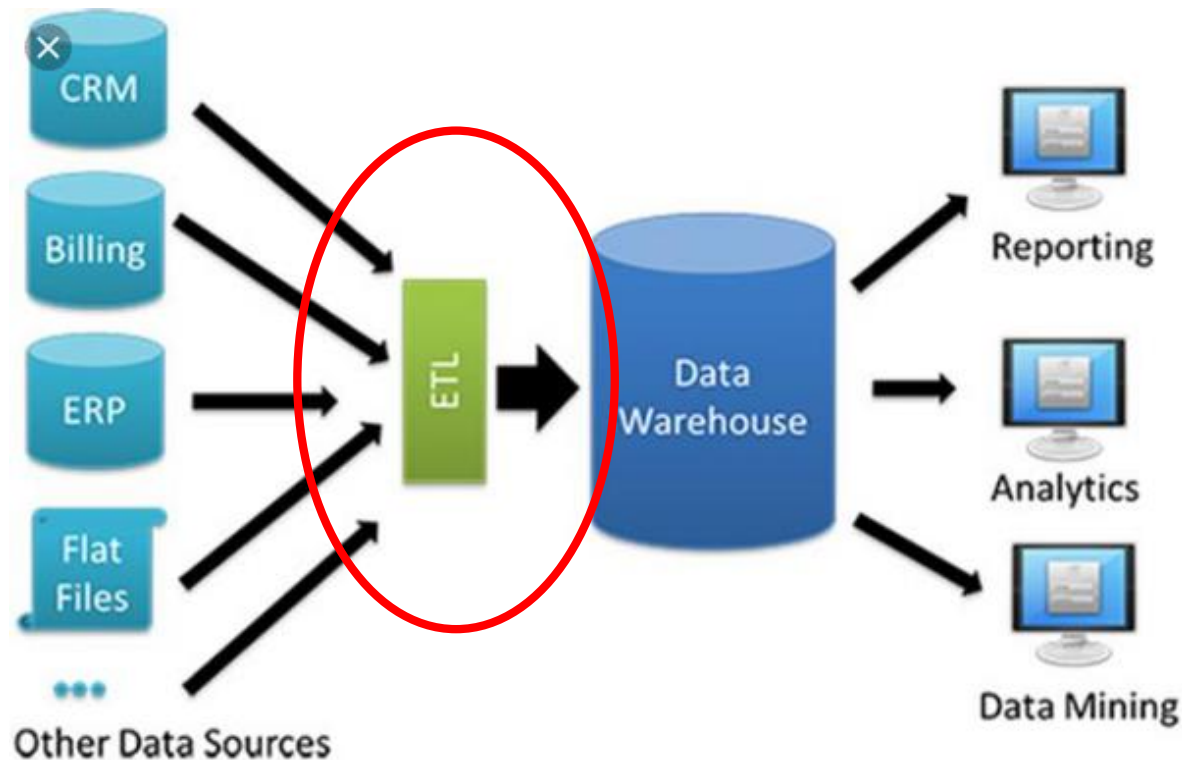
Principales conceptos:

- Proceso ETL: en qué consiste, qué beneficios aporta
- Data warehouse: concepto, proceso creación, beneficios
- Proceso para definición de una Estrategia Promocional: cálculo de escenarios de rentabilidad

Data warehouse: Proceso ETL

Información transaccional
Procesos de facturación
CRM
ERP's
Medios digitales (web, app, rss)
Fuentes externas a la empresa
.....

DTW: Centraliza la información disponible



Proceso ETL

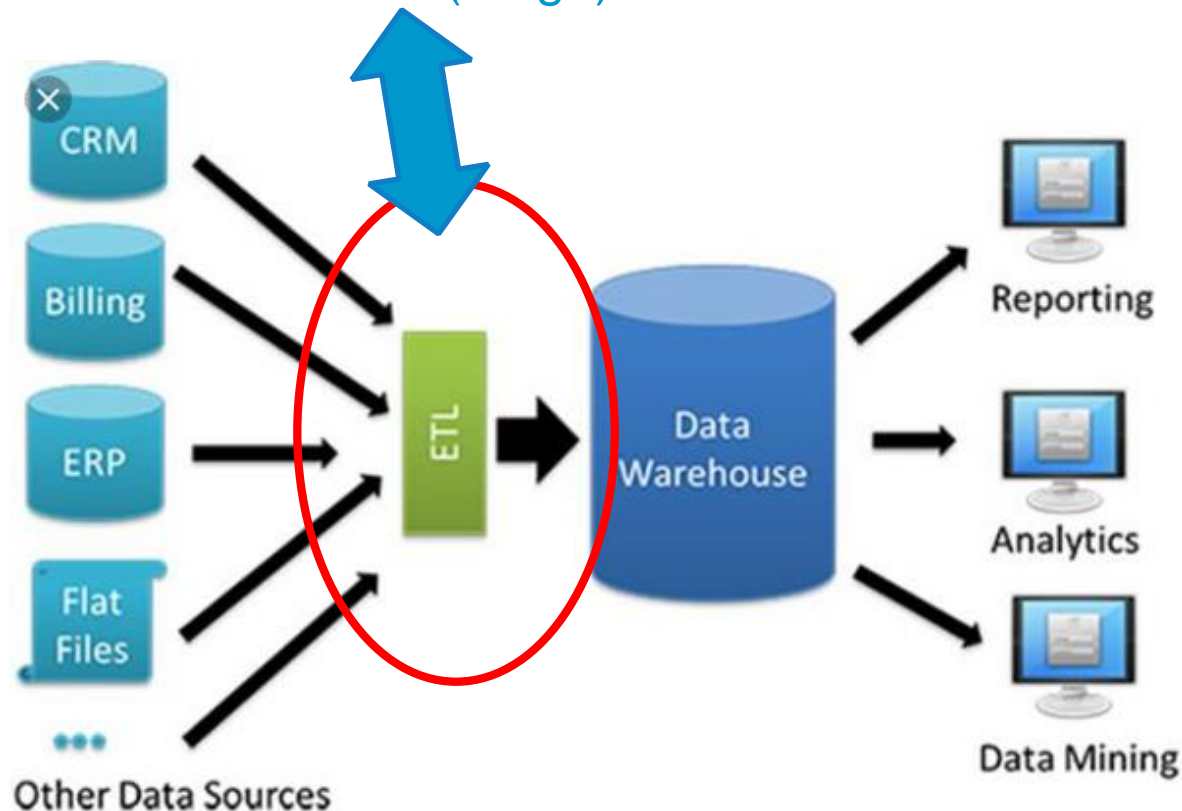
Fase de limpieza
(datos incorrectos,
inconsistentes o
duplicados)

E: extracción

T: transformación

L: load (carga)

ETL: Proceso BÁSICO que
asegura la posibilidad de analizar
la información al asegurar su
idoneidad y consistencia



Proceso ETL: consecuencias de no aplicar este proceso...

Existencia de duplicados

Pueden existir en nuestra base de datos emails que el sistema considera diferentes porque en un caso existen espacios en blanco al inicio...

email	count
carmencs2770@gmail.com	2
carmenyusplo@hotmail.com	2
chrisgkroome@btinternet.com	2
concepociof@gmail.com	2
cottonlining@icloud.com	2
d.combe@dsimail.com	2
documjose@gmail.com	2
dstanovick@gmail.com	2
elandetes@gmail.com	2
eliane.pradeis@wanadoo.fr	2
evej.marano@wanadoo.fr	2
ewa.jackiewicz@gmail.com	2
fergrv@hotmail.com	2
flo_rossignol@hotmail.com	2
fraief74@aol.com	2
garferzua@hotmail.com	2
gelinmarie@hotmail.com	2
info@matracos.es	2
iso.tardivan@yahoo.com	2
jeaniebenz@yahoo.com	2
joanfarrell0@gmail.com	2
kosky29@gmail.com	2
karolgrady@yahoo.com	2
lehnm@zoominternet.net	2
lexetiure77@hotmail.com	2

Tabla Clientes					
email	name	surname	date_nac	address	cp
carmencs2770@gmail.com	Carmen	Cortijos SÁnchez		Los jurados jaravia	4648
carmencs2770@gmail.com	Carmen	Cortijos SÁnchez		Losjuradosjaravia	4648

Tabla Pedidos							
codpedido	tienda_3	fcreeacion	email	moneda	total_preco_pedido	gastos_envio	precio_modelo
200088842	espana	29/08/2017 11:39	carmencs2770@gmail.com	EUR	188	0	99
200088842	espana	29/08/2017 11:39	carmencs2770@gmail.com	EUR	188	0	89
200088842	espana	29/08/2017 11:39	carmencs2770@gmail.com	EUR	188	0	99
200088842	espana	29/08/2017 11:39	carmencs2770@gmail.com	EUR	188	0	89

pero que realmente hacen referencia al mismo cliente

Proceso ETL: consecuencias de no aplicar este proceso...

email	name	surname	date_nac	address
midniht@mail.ru	Kristina	Kim		
verenaschmidt@web.de	Verena	Held		Lindenbergstr. 5
*ichel-taberlet0147@orange.fr	chantal taberlet		15/8/1951 0:00:00	
+cannbolt@gvtc.com	Cindy	Bolt		5735 derby way
00.snyder@web.de	Sonja	Schneider		KIOSKZWISCHENSTOPPKREUZSTR....
002info@arcor.de	Uwe	Meerz		Silbergrasweg 6
003002roma@gmail.com	María	Barco Díaz	14/2/1977 0:00:00	C/JoséMaríaPereda8,2C
0076418@gmail.com	AJ Powell		14/10/1978 0:00:00	
007lalee@gmail.com				
00cerezas@gmail.com	LOURDES MARIA	RUIZ POZO		BDA ANDALUCIA. CALLE MARCHEN...
00fhagerty@gmail.com				
00hb33@gmail.com	Jonathan	Lawson		
00jrmr@gmail.com	Jose Rodolfo	Martinez Rodriguez		
01.obivan@gmail.com				
0102.lola@gmail.com	Lola	Lopez Garcia		Calle Esteban Mora,6-3a
0102.lolas@gmail.com	Lola		1/2/1963 0:00:00	
0123amalia@gmail.com	Amalia	MUNOZ PINEIRO		Adolf Reydamslan 54

Existe gran número de registros con los campos a missing y, sobre todo

...no existe una normalización de la dirección de correspondencia

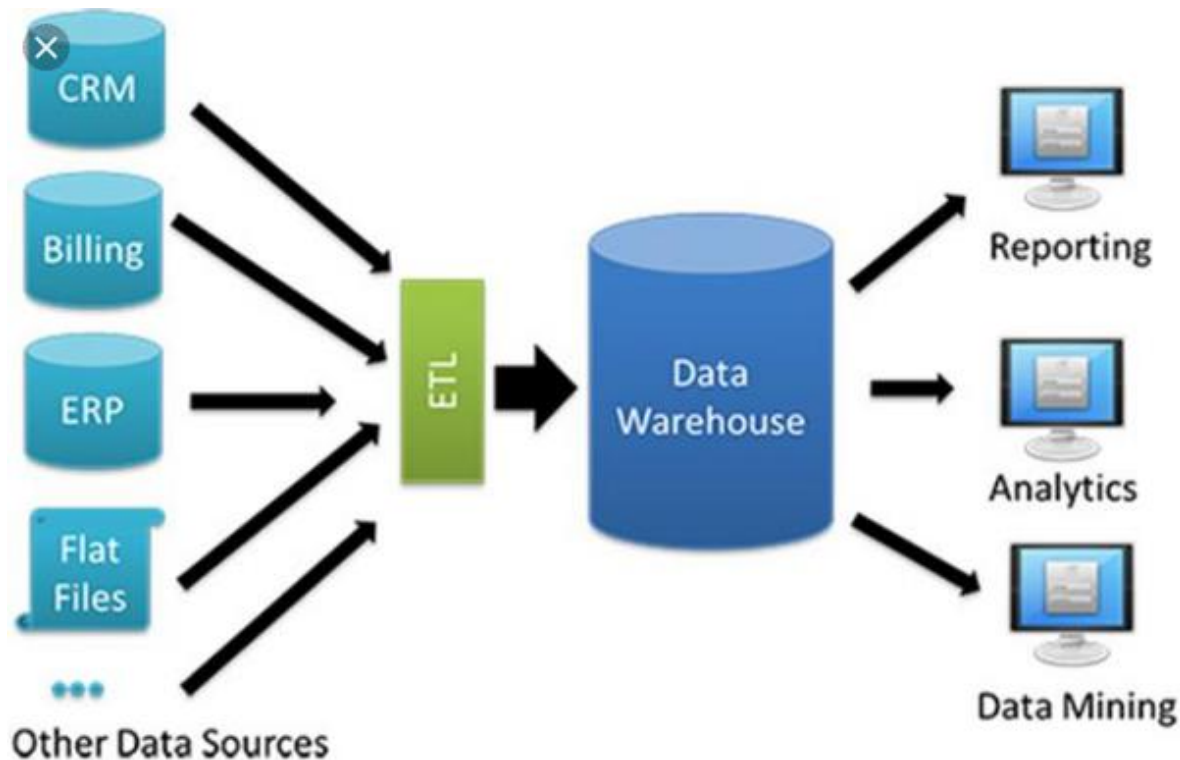


consecuencias: alto índice de devoluciones (en envíos off-line: compras, comunicaciones...), no poder llevar a cabo procesos de “enriquecimiento de la información”...

Data warehouse

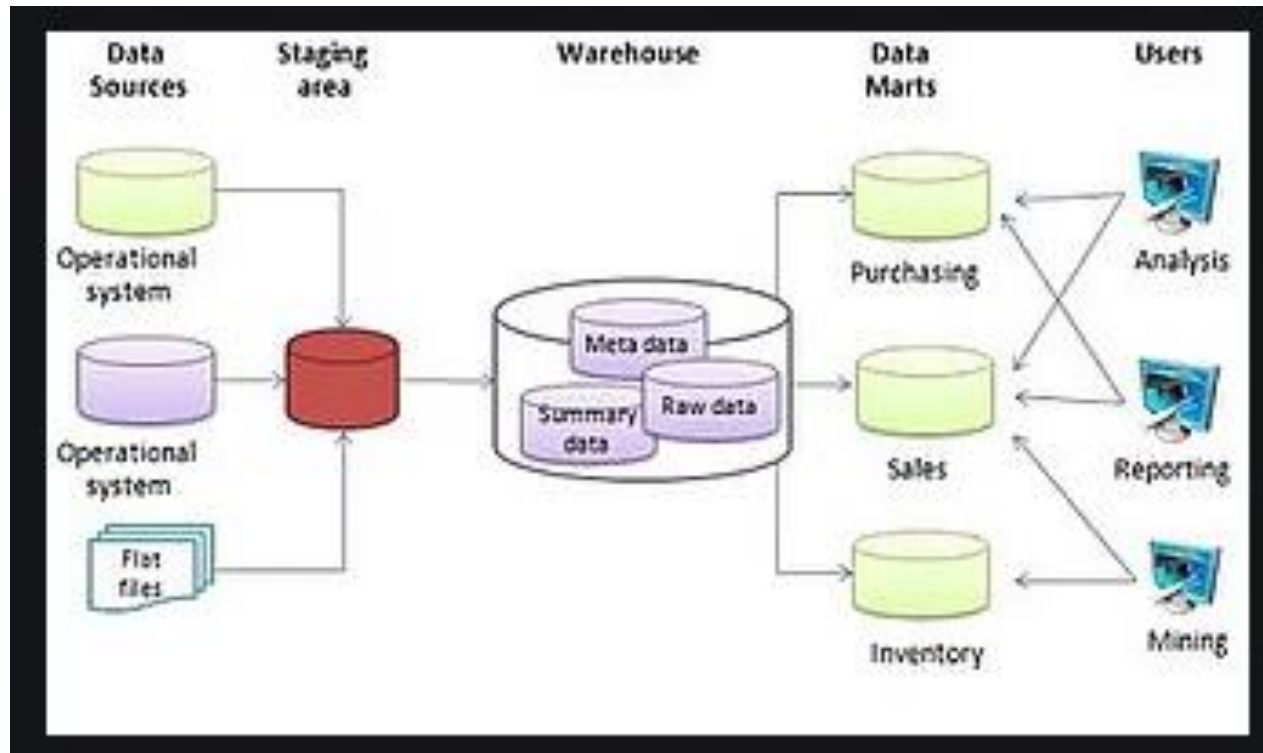
Información transaccional
Procesos de facturación
CRM
ERP's
Medios digitales (web, app, rss)
Fuentes externas a la empresa
.....

Centraliza la información disponible



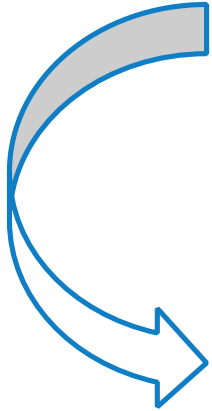
Data warehouse ↔ Data marts

En ocasiones, pueden existir data marts...que no son más que data warehouse “específicos” (compras, facturación...)



Data warehouse

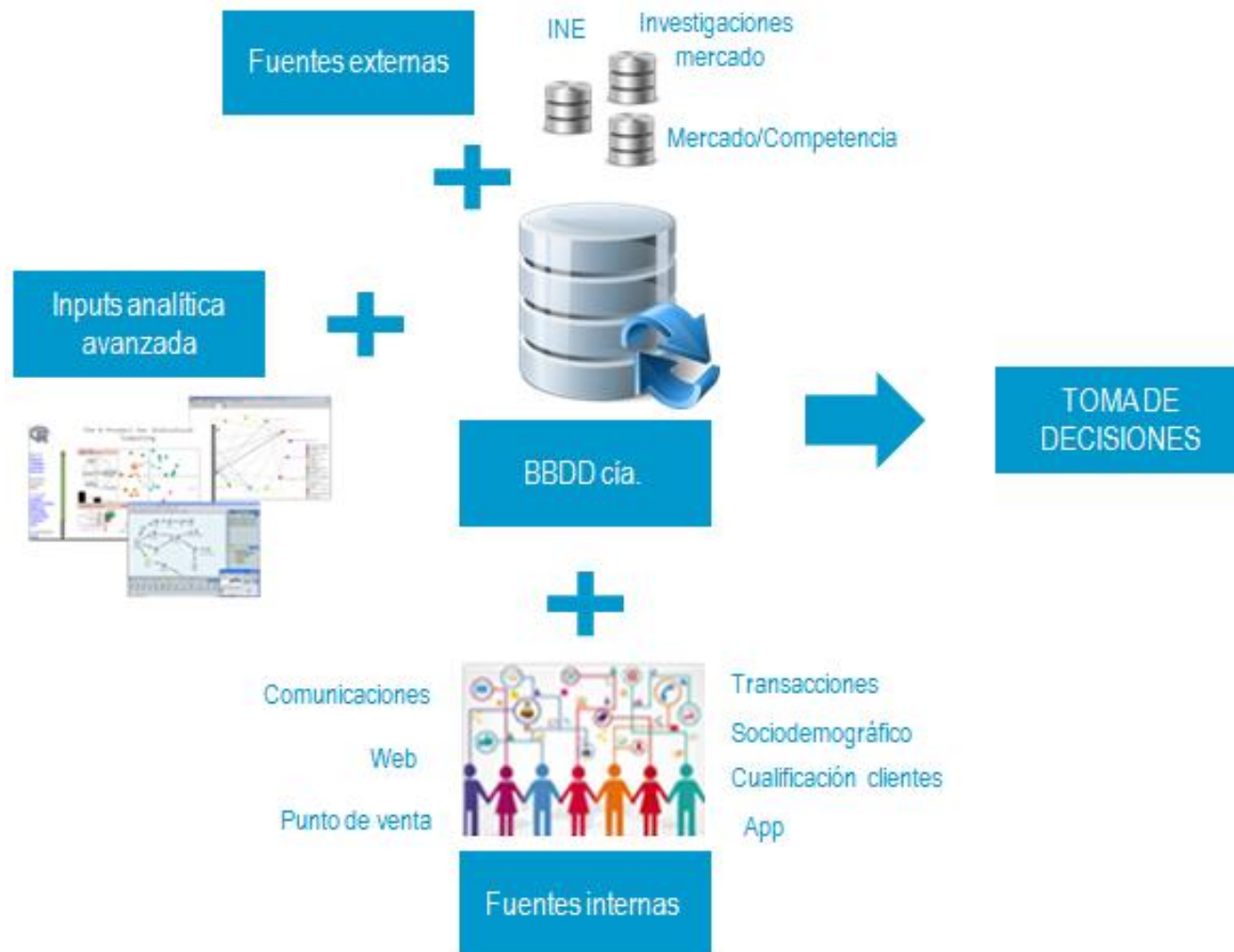
Beneficios



- Centralización de toda la información
- Eficacia y Eficiencia de procesos (sólo es necesario acceder a un repositorio de información)
- Es el motor que permite generar conocimiento e inteligencia
- Necesario en todo proceso de toma de decisiones
- Elemento necesario para generar ventaja competitiva

Data warehouse

Elemento IMPRESCINDIBLE en el
proceso de toma de decisiones



Data warehouse: modelo estrella

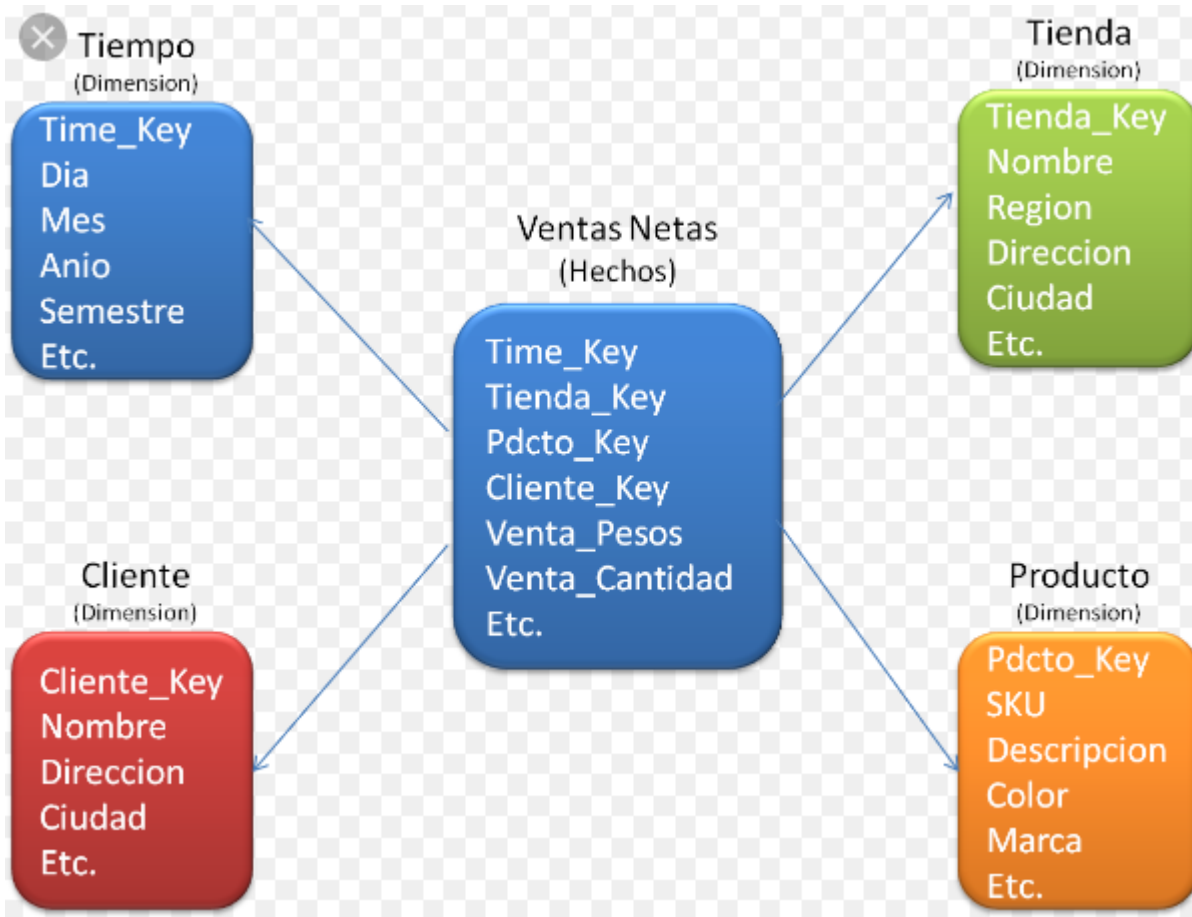
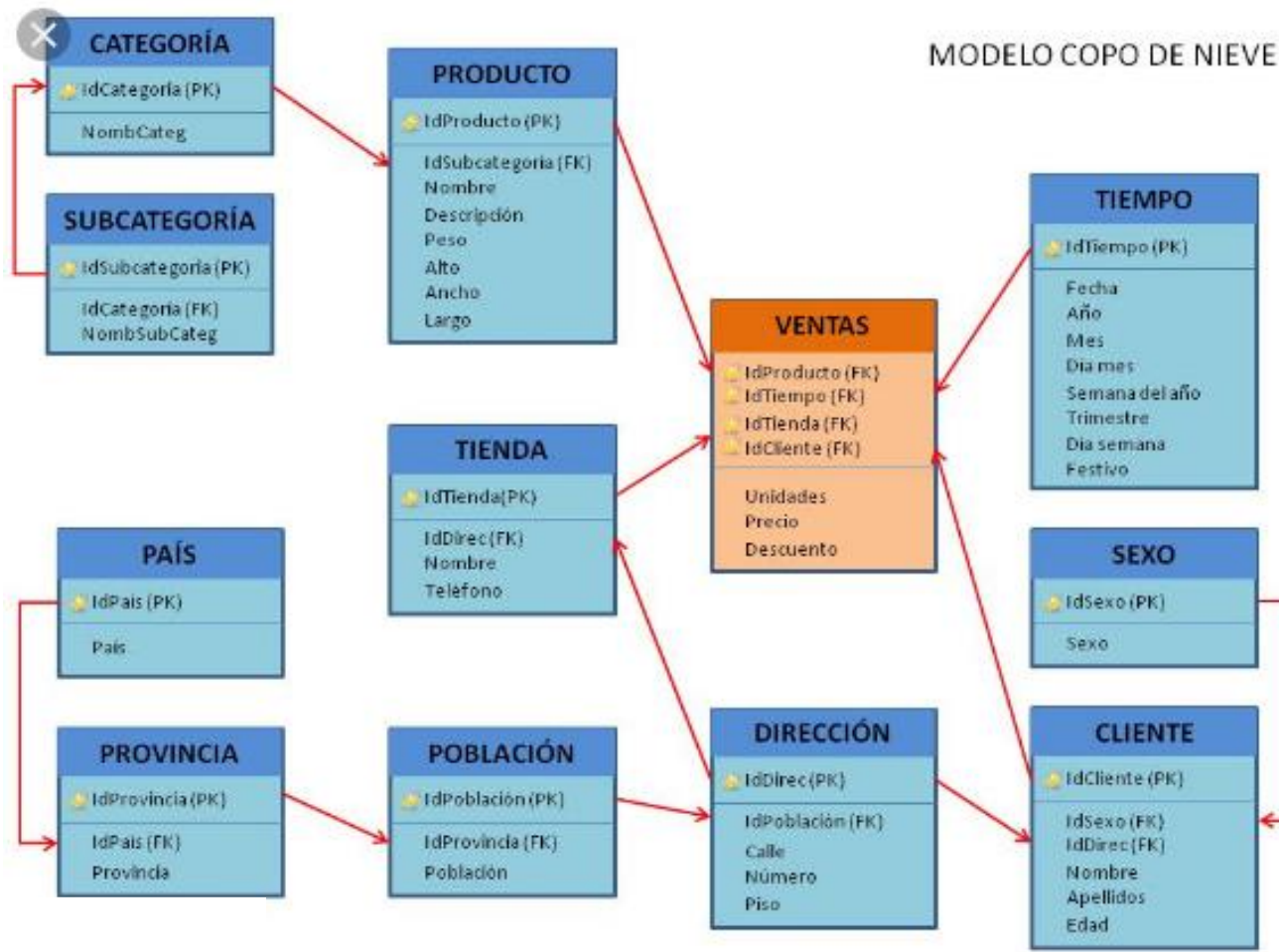


Tabla de hechos (datos) y Tablas de dimensiones (variables por las que se puede analizar los datos)

No existe relación (jerarquía) entre las tablas de dimensiones

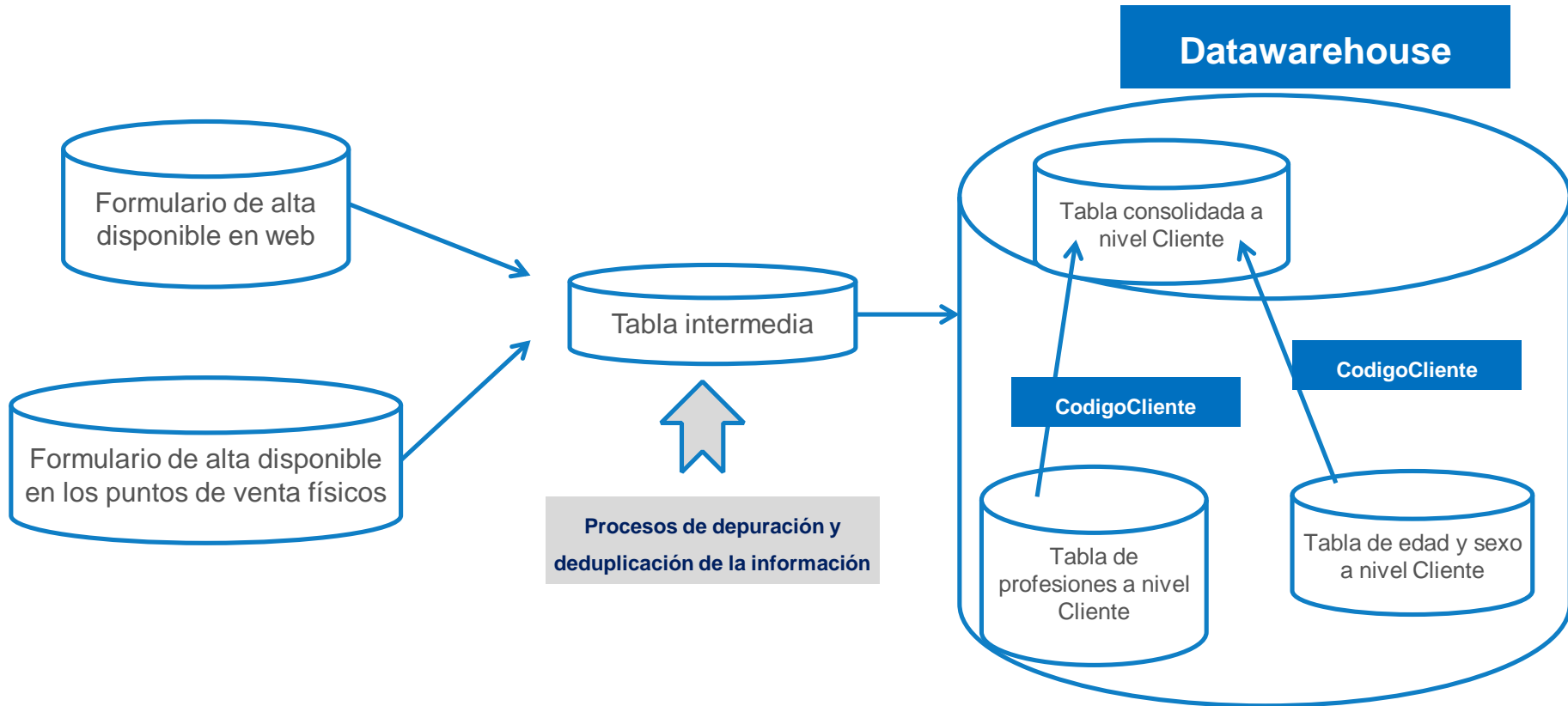
Data warehouse: modelo copo de nieve (snowflake)



Las tablas de dimensiones (variables por las que se puede analizar los datos) se relacionan entre sí (presentan jerarquías)

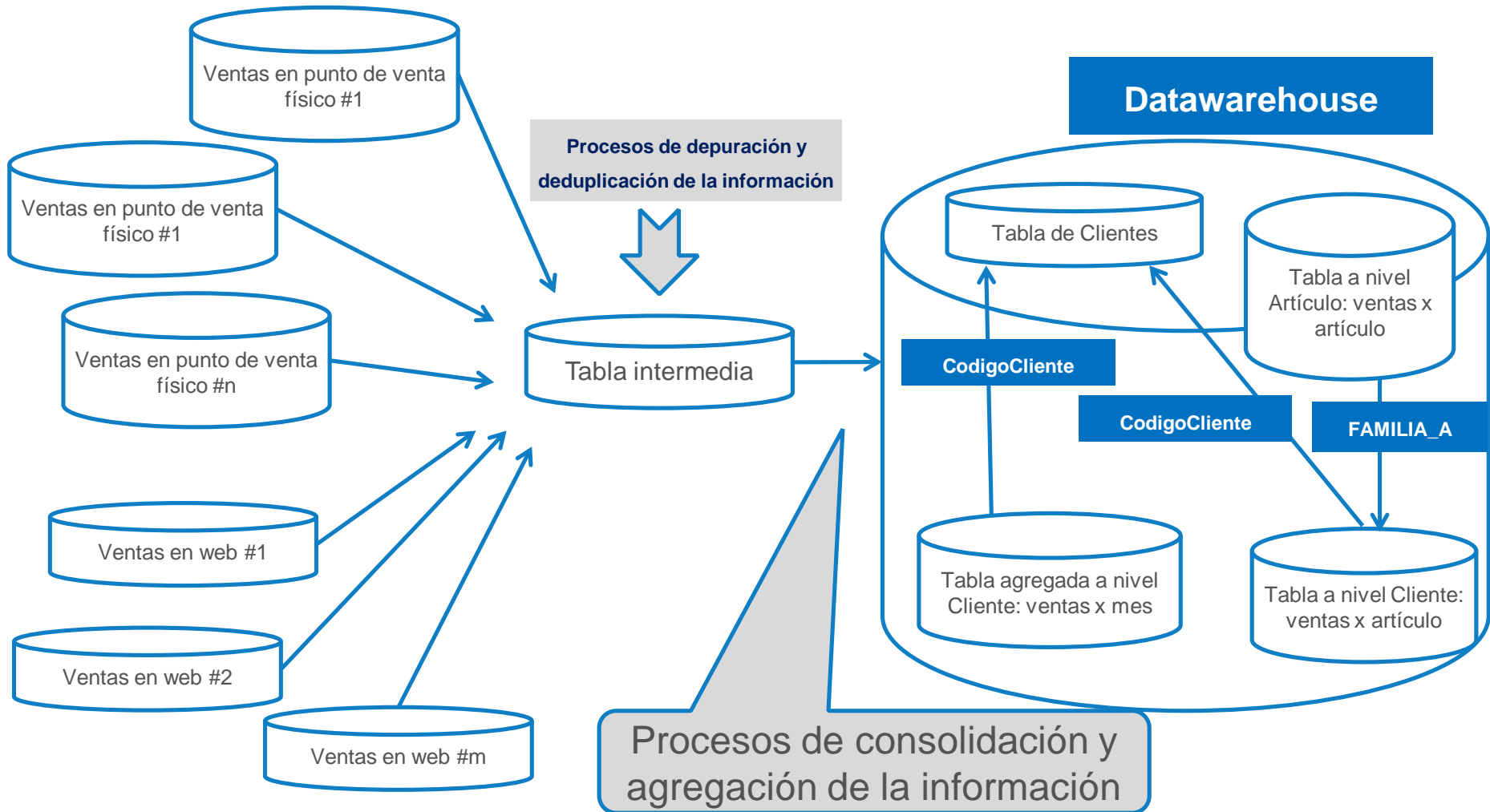
Datawarehouse: ejemplo de diseño

- Disponer de información socio-demográfica: edad, sexo, nº hijos, profesión...

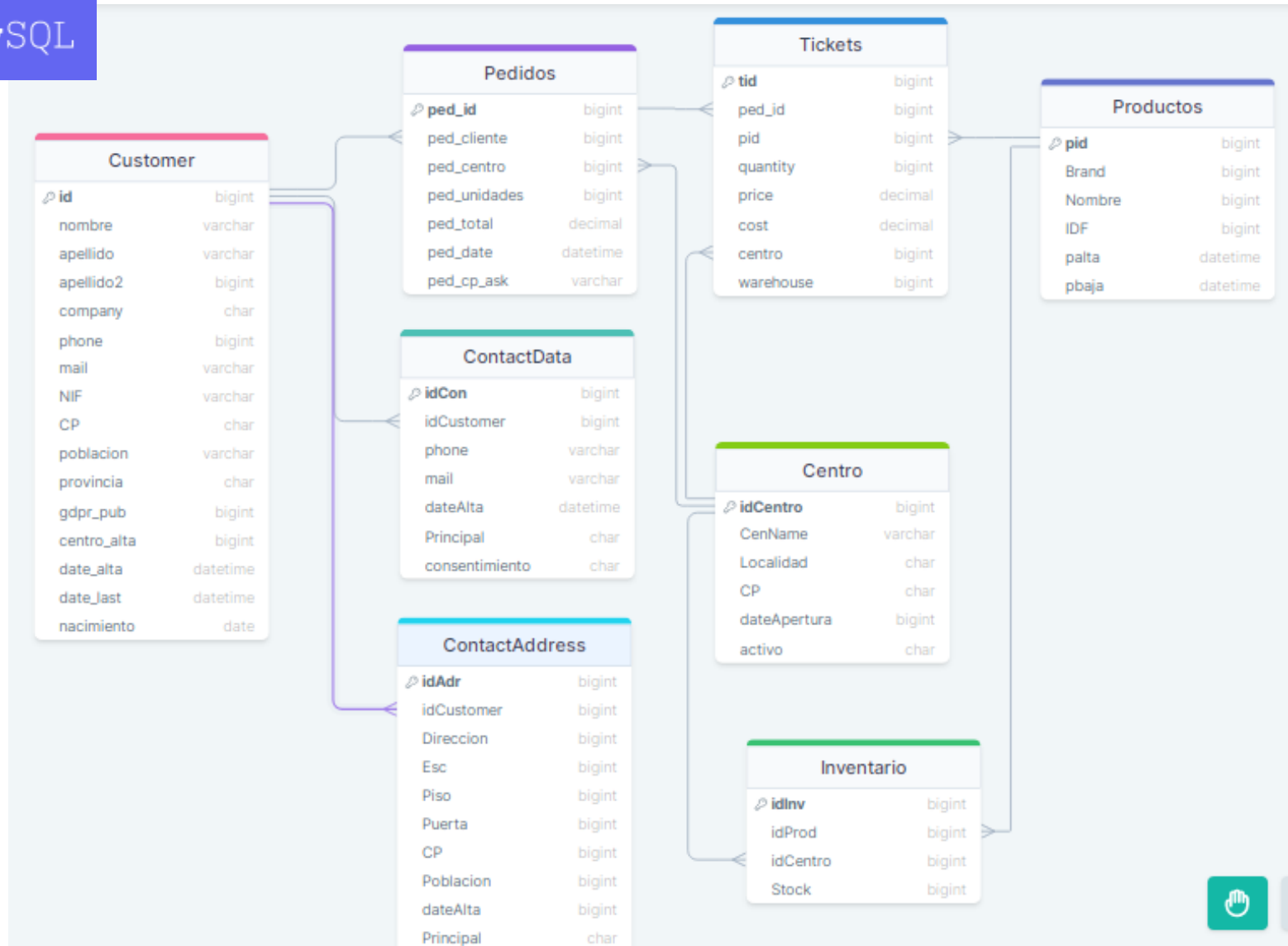


Datawarehouse: ejemplo de diseño

- Calcular agregadas x Cliente, x Mes, x Categoría...

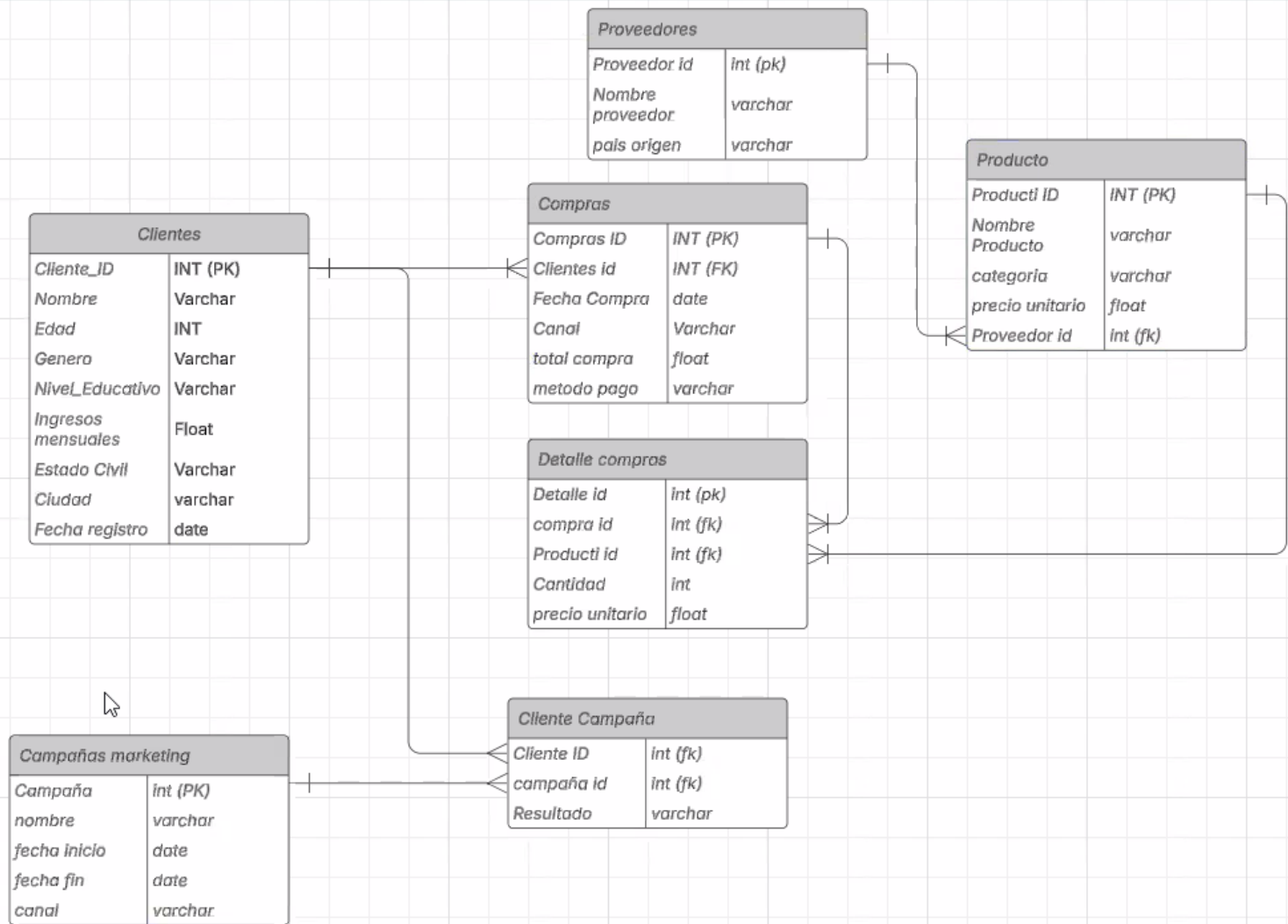


Datawarehouse: ejemplo de diseño



Cuadro de Mando Integral





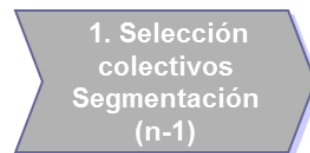
Estrategia Promocional

Objetivo: Optimizar la Rentabilidad de las Promociones a través de:

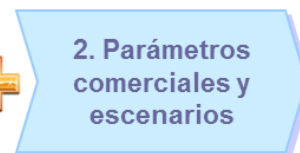
1. Estimación y cuantificación escenarios de ROI, según segmentos / ofertas
2. Testeo y elección de Niveles de Exigencia / Respuesta para minimizar el coste de dilución (respuesta espontánea)
3. Selección del mejor binomio Oferta-Respuesta que aporten mayor Contribución / ROI

Metodología

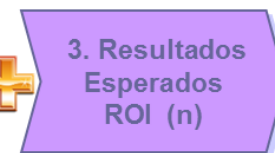
1. Definición segmentos
2. Definición periodo, oferta, exigencia
3. Simulación ROI "futuro"



Identificación colectivos de interés mismo periodo año anterior.



Cálculo diferentes niveles exigencia x socio (ej: media de los 3 Tickets máximos último trimestre)



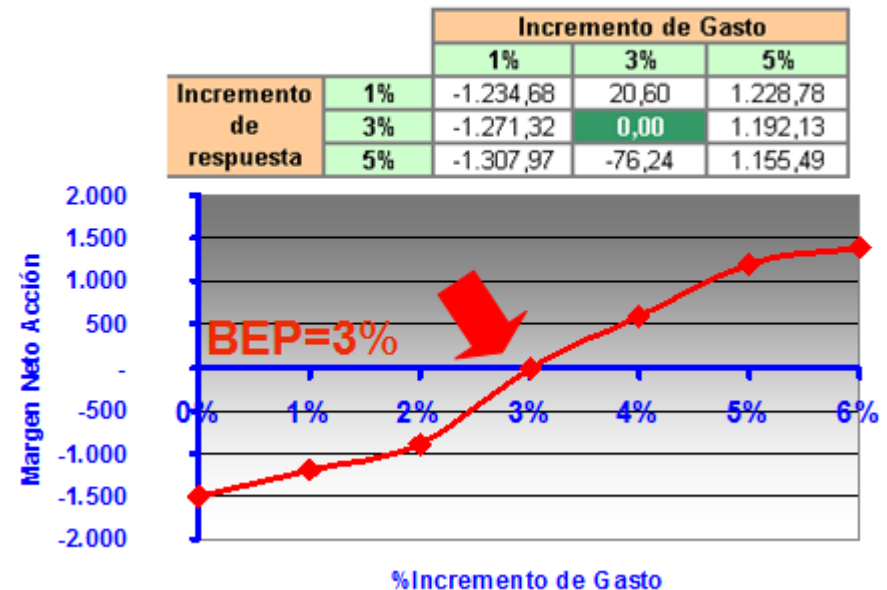
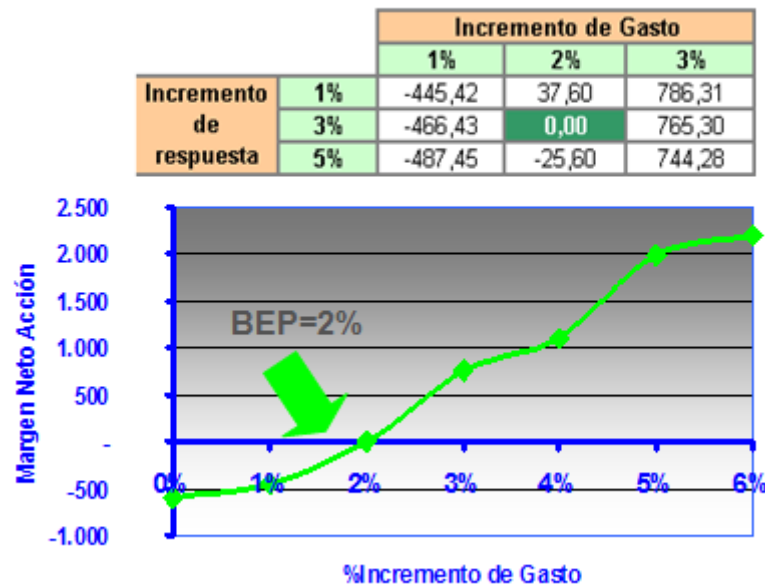
Estimación resultados: Diferentes Ofertas, Respuesta e incremento de facturación...



Estrategia Promocional

- Para los Clientes con el mismo nivel de exigencia, el margen neto de la acción es diferente dependiendo de la Oferta seleccionada
- Los escenarios permiten comparar Ofertas y determinar su BEP (break even point):

Estimando un incremento del 3% en la respuesta, la Oferta 1 necesita alcanzar un incremento de gasto del 2% para alcanzar su BEP (punto mínimo de rentabilidad) vs la Oferta 2 que necesitaría un 3%





www.unir.net