DISEÑO MULTIDIMENSIONAL
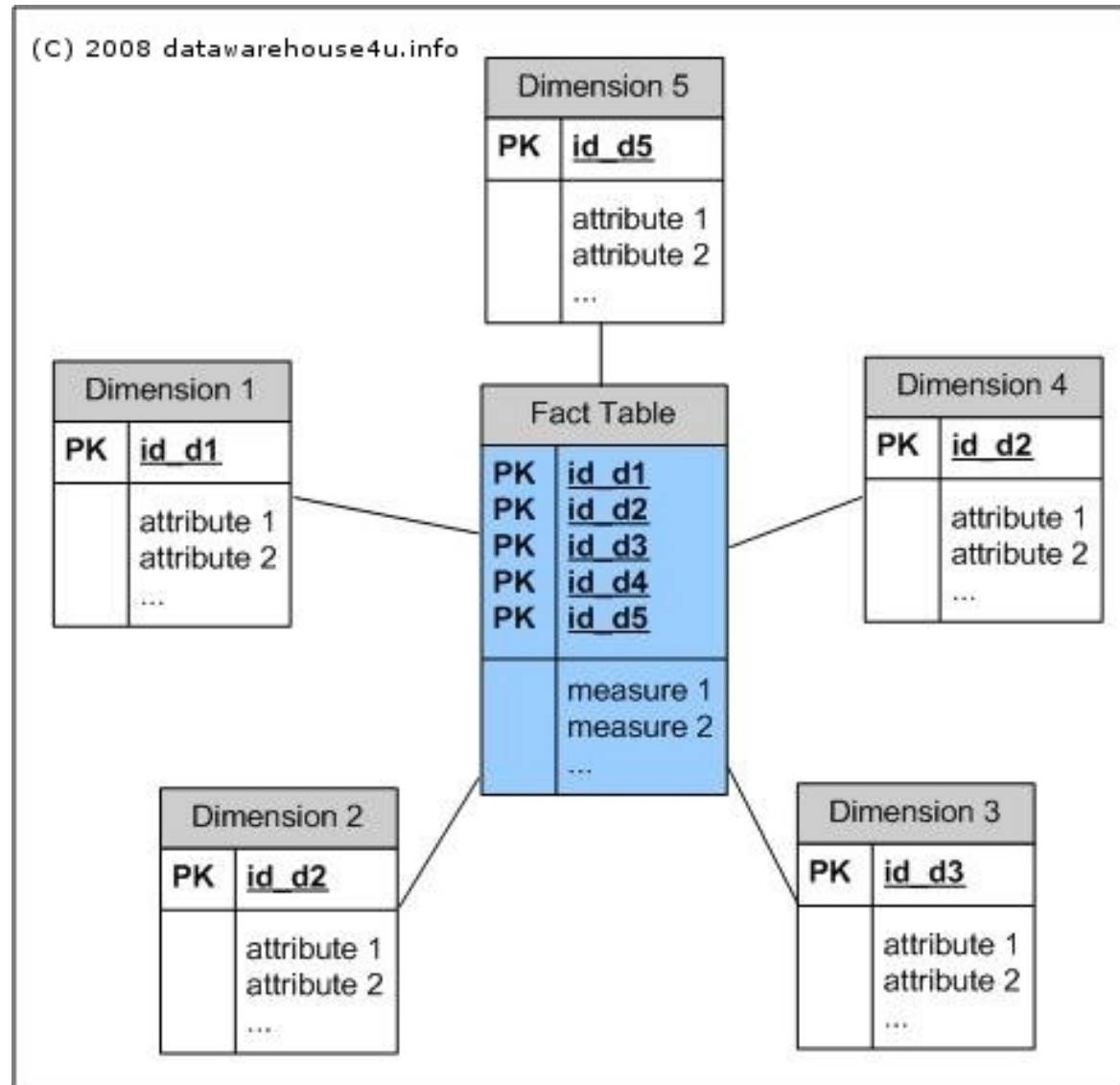
Ing. Julio Paciello

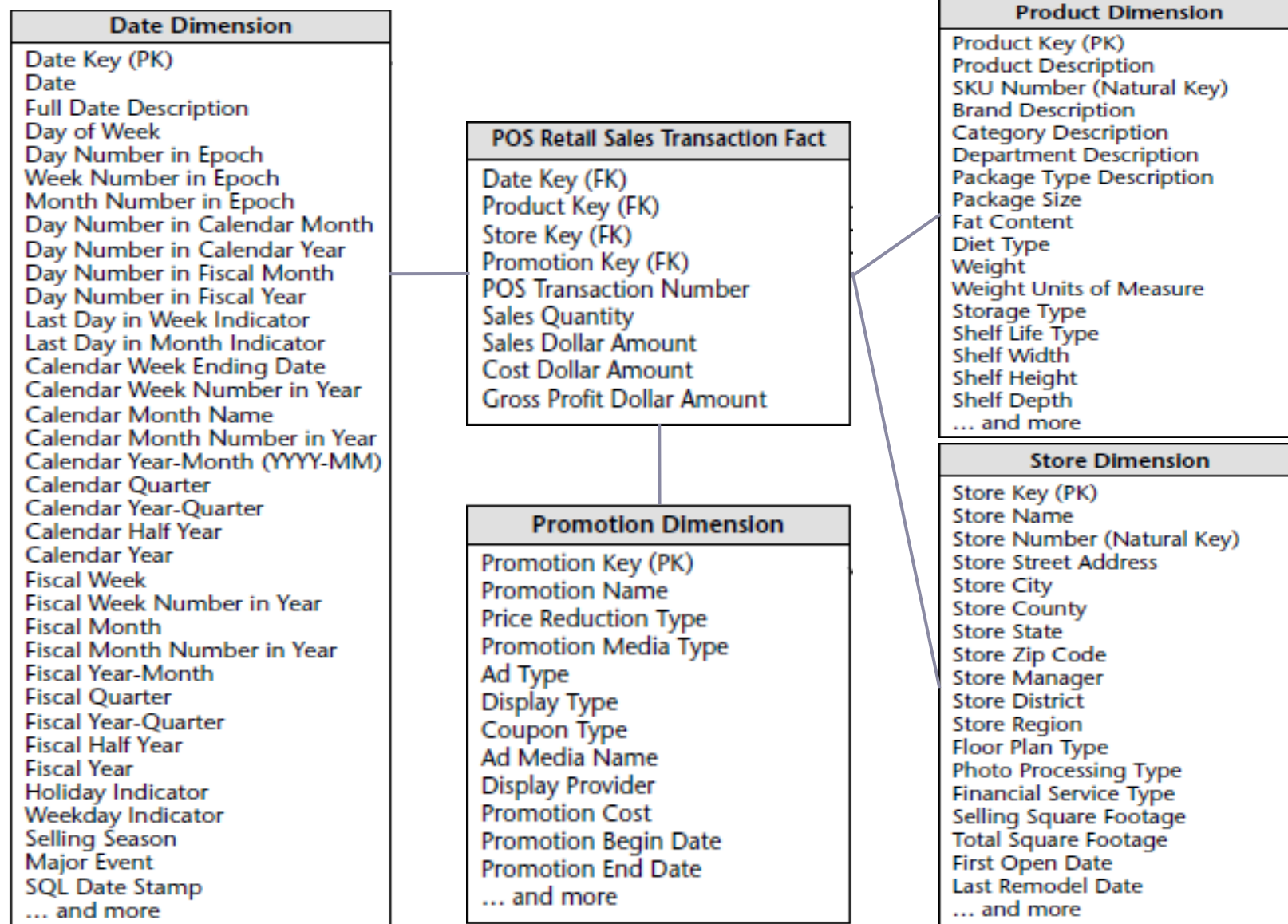juliopaciello@cds.com.py

Diseño Multidimensional
Star scheme

Ejemplo Práctico: Ventas

Date Dimension

Date Key (PK)
Date
Full Date Description
Day of Week
Day Number in Epoch
Week Number in Epoch
Month Number in Epoch
Day Number in Calendar Month
Day Number in Calendar Year
Day Number in Fiscal Month
Day Number in Fiscal Year
Last Day in Week Indicator
Last Day in Month Indicator
Calendar Week Ending Date
Calendar Week Number in Year
Calendar Month Name
Calendar Month Number in Year
Calendar Year-Month (YYYY-MM)
Calendar Quarter
Calendar Year-Quarter
Calendar Half Year
Calendar Year
Fiscal Week
Fiscal Week Number in Year
Fiscal Month
Fiscal Month Number in Year
Fiscal Year-Month
Fiscal Quarter
Fiscal Year-Quarter
Fiscal Half Year
Fiscal Year
Holiday Indicator
Weekday Indicator
Selling Season
Major Event
SQL Date Stamp
… and more

POS Retail Sales Transaction Fact

Date Key (FK)
Product Key (FK)
Store Key (FK)
Promotion Key (FK)
POS Transaction Number
Sales Quantity
Sales Dollar Amount
Cost Dollar Amount
Gross Profit Dollar Amount

Promotion Dimension

Promotion Key (PK)
Promotion Name
Price Reduction Type
Promotion Media Type
Ad Type
Display Type
Coupon Type
Ad Media Name
Display Provider
Promotion Cost
Promotion Begin Date
Promotion End Date
… and more

Product Dimension

Product Key (PK)
Product Description
SKU Number (Natural Key)
Brand Description
Category Description
Department Description
Package Type Description
Package Size
Fat Content
Diet Type
Weight
Weight Units of Measure
Storage Type
Shelf Life Type
Shelf Width
Shelf Height
Shelf Depth
… and more

Store Dimension

Store Key (PK)
Store Name
Store Number (Natural Key)
Store Street Address
Store City
Store County
Store State
Store Zip Code
Store Manager
Store District
Store Region
Floor Plan Type
Photo Processing Type
Financial Service Type
Selling Square Footage
Total Square Footage
First Open Date
Last Remodel Date
… and more

Dimensiones, ejemplos de valores

Date Key	Date	Full Date Description	Day of Week	Calendar Month	Calendar Year	Fiscal Year-Month	Holiday Indicator	Weekday Indicator
1	01/01/2002	January 1, 2002	Tuesday	January	2002	F2002-01	Holiday	Weekday
2	01/02/2002	January 2, 2002	Wednesday	January	2002	F2002-01	Non-Holiday	Weekday
3	01/03/2002	January 3, 2002	Thursday	January	2002	F2002-01	Non-Holiday	Weekday
4	01/04/2002	January 4, 2002	Friday	January	2002	F2002-01	Non-Holiday	Weekday
5	01/05/2002	January 5, 2002	Saturday	January	2002	F2002-01	Non-Holiday	Weekend
6	01/06/2002	January 6, 2002	Sunday	January	2002	F2002-01	Non-Holiday	Weekend
7	01/07/2002	January 7, 2002	Monday	January	2002	F2002-01	Non-Holiday	Weekday
8	01/08/2002	January 8, 2002	Tuesday	January	2002	F2002-01	Non-Holiday	Weekday

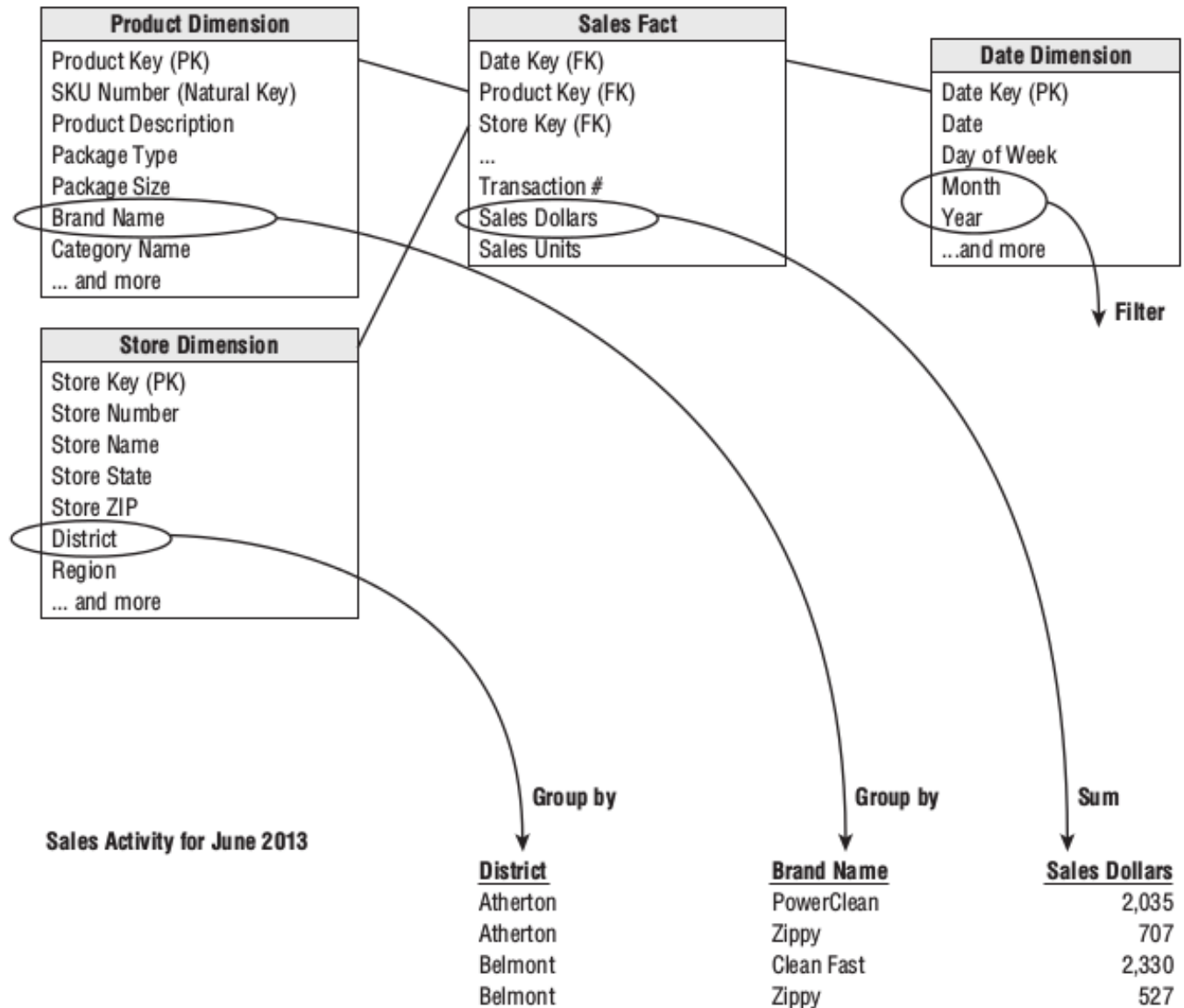| Product Key	Product Description	Brand Description	Category Description	Department Description	Fat Content
1	Baked Well Light Sourdough Fresh Bread	Baked Well	Bread	Bakery	Reduced Fat
2	Fluffy Sliced Whole Wheat	Fluffy	Bread	Bakery	Regular Fat
3	Fluffy Light Sliced Whole Wheat	Fluffy	Bread	Bakery	Reduced Fat
4	Fat Free Mini Cinnamon Rolls	Light	Sweeten Bread	Bakery	Non-Fat
5	Diet Lovers Vanilla 2 Gallon	Coldpack	Frozen Desserts	Frozen Foods	Non-Fat
6	Light and Creamy Butter Pecan 1 Pint	Freshlike	Frozen Desserts	Frozen Foods	Reduced Fat
7	Chocolate Lovers 1/2 Gallon	Frigid	Frozen Desserts	Frozen Foods	Regular Fat
8	Strawberry Ice Creamy 1 Pint	Icy	Frozen Desserts	Frozen Foods	Regular Fat
9	Icy Ice Cream Sandwiches	Icy	Frozen Desserts	Frozen Foods	Regular Fat

Hechos, ejemplos *drill-down*

Department Description	Sales Dollar Amount	Sales Quantity
Bakery	$12,331	5,088
Frozen Foods	$31,776	15,565

Department Description	Brand Description	Sales Dollar Amount	Sales Quantity
Bakery	Baked Well	$3,009	1,138
Bakery	Fluffy	$3,024	1,476
Bakery	Light	$6,298	2,474
Frozen Foods	Coldpack	$5,321	2,640
Frozen Foods	Freshlike	$10,476	5,234
Frozen Foods	Frigid	$7,328	3,092
Frozen Foods	Icy	$2,184	1,437
Frozen Foods	QuickFreeze	$6,467	3,162

Ejemplo de reporte simple

Product Dimension
- Product Key (PK)
- SKU Number (Natural Key)
- Product Description
- Package Type
- Package Size
- Brand Name
- Category Name
- ... and more

Sales Fact
- Date Key (FK)
- Product Key (FK)
- Store Key (FK)
- ...
- Transaction #
- Sales Dollars
- Sales Units

Date Dimension
- Date Key (PK)
- Date
- Day of Week
- Month
- Year
- ...and more

Filter

Store Dimension
- Store Key (PK)
- Store Number
- Store Name
- Store State
- Store ZIP
- District
- Region
- ... and more

Group by   Group by   Sum

Sales Activity for June 2013

District	Brand Name	Sales Dollars
Atherton	PowerClean	2,035
Atherton	Zippy	707
Belmont	Clean Fast	2,330
Belmont	Zippy	527

Diseño dimensional en 4 pasos

1. Seleccionar el proceso de negocio
2. Definir la granularidad
   - *Atomic grain*
   - Granularidades distintas de análisis no deben ir en una misma tabla de hechos
3. Identificar las dimensiones
4. Identificar los hechos

Dimensiones

- *"who, what, where, when, why, and how"*

- Atributos descriptivos utilizados para el filtrado y la agrupación de los hechos (*facts*)

- Clave subrogada y clave natural

- Diferentes tipos de dimensiones, según:
    - Periodicidad de cambios
    - Diseño funcional

Clave subrogada y Natural

- *Subrogada:* PK único de cada registro en una tabla de dimensión
- *Natural:* PK en origen de datos

Clave subrogada y Natural

- *Subrogada:* PK único de cada registro en una tabla de dimensión
- *Natural:* PK en origen de datos

Ventajas:

- Integración con múltiples orígenes de datos
- Desempeño, PKs y FKs simples, no compuestas
- Valores especiales para casos nulos o de error
- Soporta tracking de cambios en los registros

Tipos de Dimensiones

- Slowly Changing
  - Tipo 1: update del registro (overwrite)
  - Tipo 2: versionado (insert más atributo nuevo versión)
    - Campos adicionales (created, updated)
  - Tipo 3: nuevo atributo
  - Tipo 4: tabla actual y tabla histórica
- Rapid Changing
  - Separar en una mini dimensión
- Junk Dimension

Tipo 1: overwrite

Original row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name
12345	ABC922-Z	IntelliKidz	Education

Updated row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name
12345	ABC922-Z	IntelliKidz	Strategy

Tipo 2: versionado

Original row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name	...	Row Effective Date	Row Expiration Date	Current Row Indicator
12345	ABC922-Z	IntelliKidz	Education	...	2012-01-01	9999-12-31	Current

Rows in Product dimension following department reassignment:

Product Key	SKU (NK)	Product Description	Department Name	...	Row Effective Date	Row Expiration Date	Current Row Indicator
12345	ABC922-Z	IntelliKidz	Education	...	2012-01-01	2013-01-31	Expired
25984	ABC922-Z	IntelliKidz	Strategy	...	2013-02-01	9999-12-31	Current

Tipo 1 y 2: update + versionado

Original row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name	Introduction Date	...	Row Effective Date	Row Expiration Date	Current Row Indicator
12345	ABC922-Z	IntelliKidz	Education	2012-12-15	...	2012-01-01	9999-12-31	Current

Rows in Product dimension following type 2 change to Department Name and type 1 change to Introduction Date:

Product Key	SKU (NK)	Product Description	Department Name	Introduction Date	...	Row Effective Date	Row Expiration Date	Current Row Indicator
12345	ABC922-Z	IntelliKidz	Education	2012-01-01	...	2012-01-01	2013-01-31	Expired
25984	ABC922-Z	IntelliKidz	Strategy	2012-01-01	...	2013-02-01	9999-12-31	Current

Tipo 3: nuevo atributo

Original row in Product dimension:

Product Key	SKU (NK)	Product Description	Department Name
12345	ABC922-Z	IntelliKidz	Education

Updated row in Product dimension:

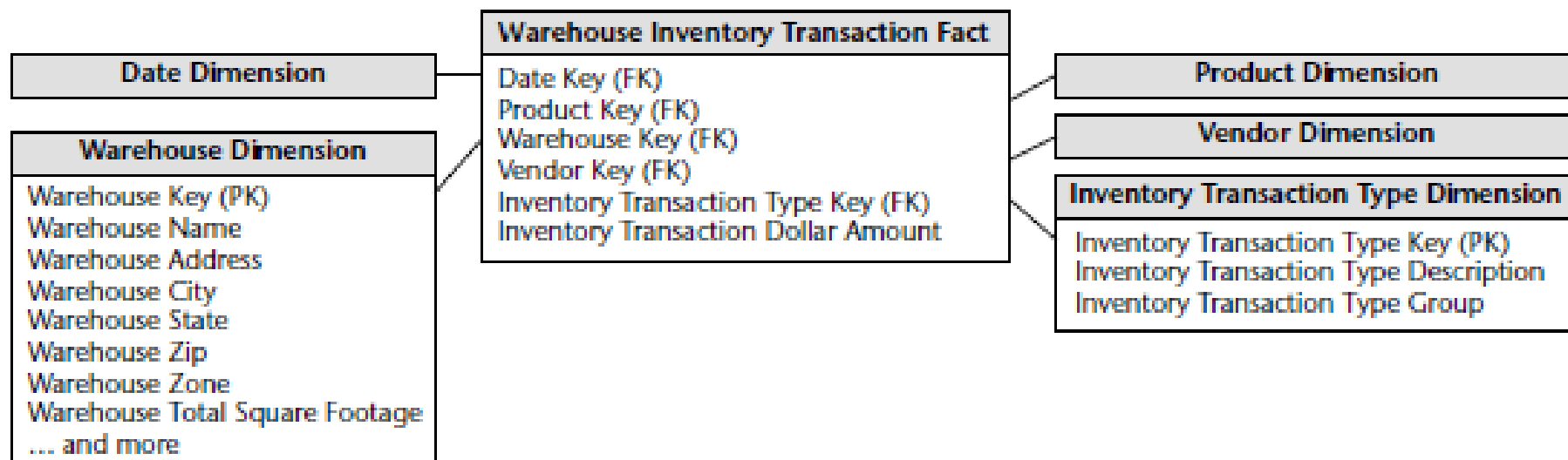| Product Key	SKU (NK)	Product Description	Department Name	Prior Department Name
12345	ABC922-Z	IntelliKidz	Strategy	Education

Tablas de Hechos

- Transactional

- Periodic Snapshot

- Accumulating Snapshot
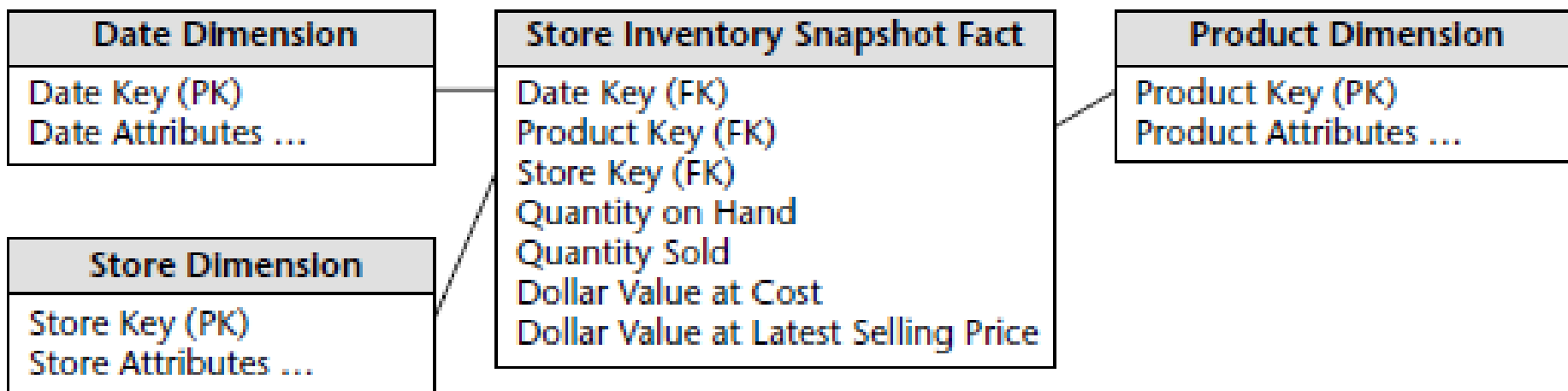
Diseño: Transactional
Ej: Inventario

Características:
* Orientado a transacciones
* 1 hecho por transacción

Diseño: Periodic Snapshot
Ej: Inventario

Date Dimension
Date Key (PK)
Date Attributes ...

Store Inventory Snapshot Fact
Date Key (FK)
Product Key (FK)
Store Key (FK)
Quantity on Hand
Quantity Sold
Dollar Value at Cost
Dollar Value at Latest Selling Price

Product Dimension
Product Key (PK)
Product Attributes ...

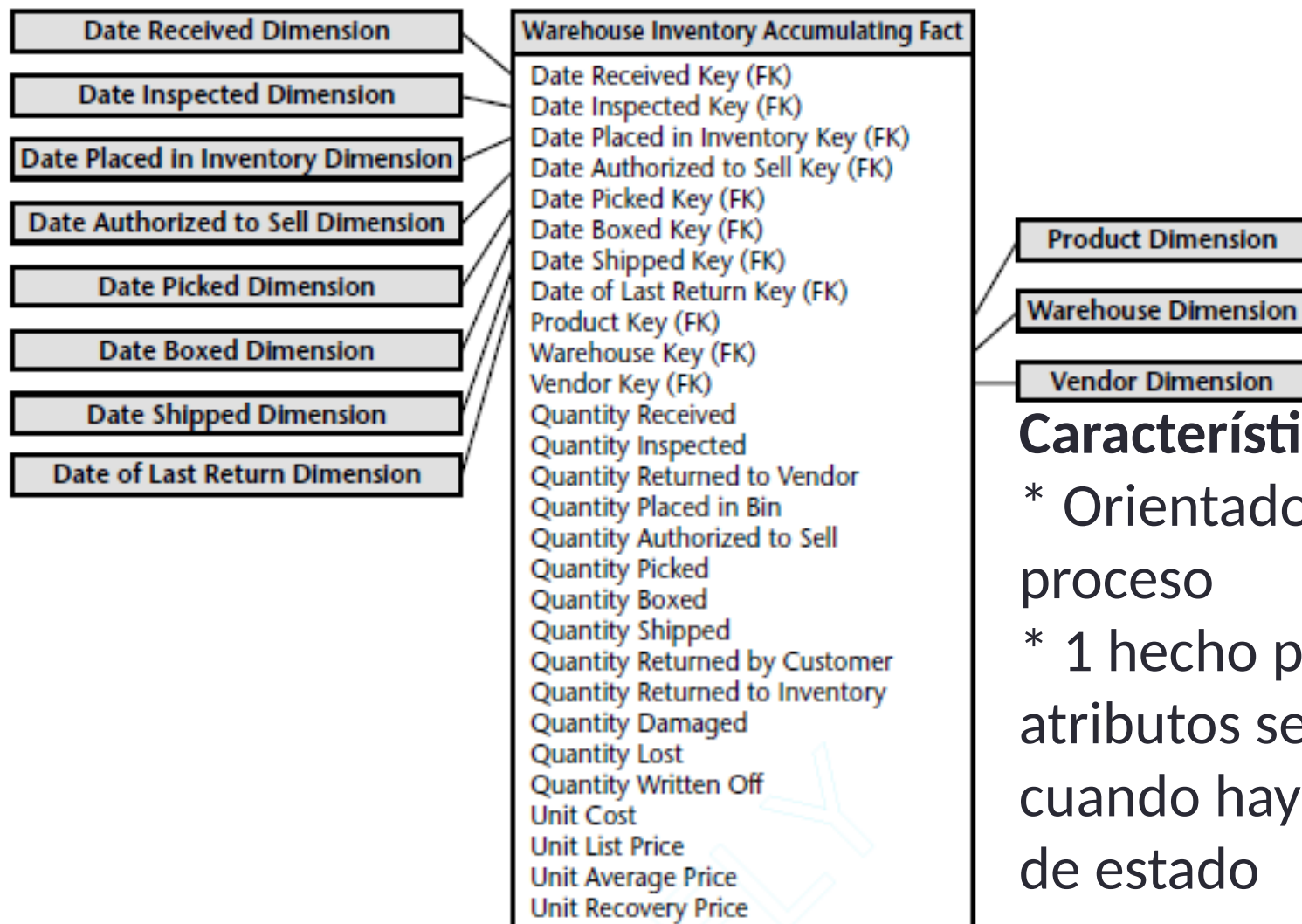| Store Dimension |
| --- |
| Store Key (PK) |
| Store Attributes ... |

Características:
* Orientado a medir niveles periódicamente
* 1 hecho por medición (ej: medición diaria de stock de productos)
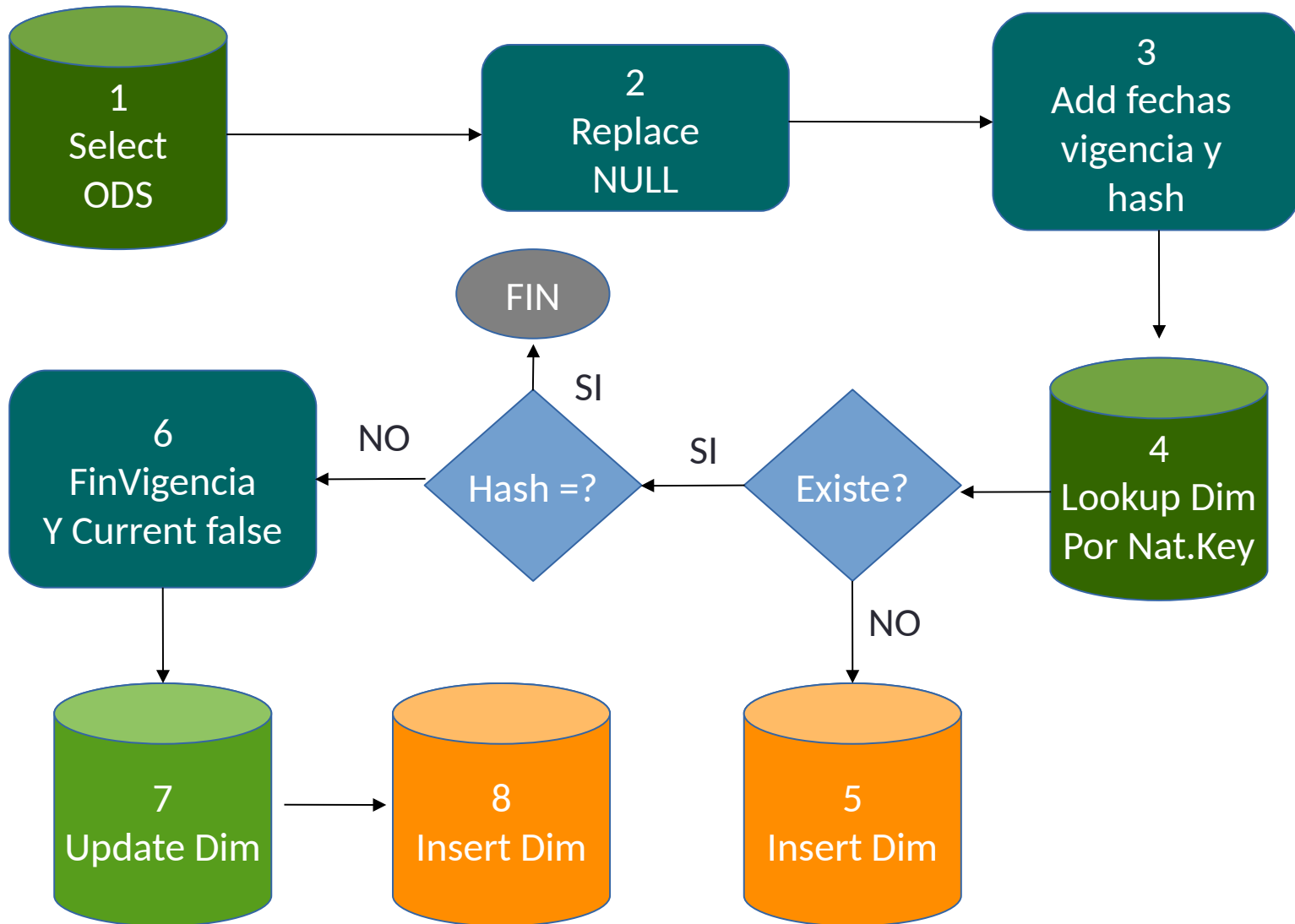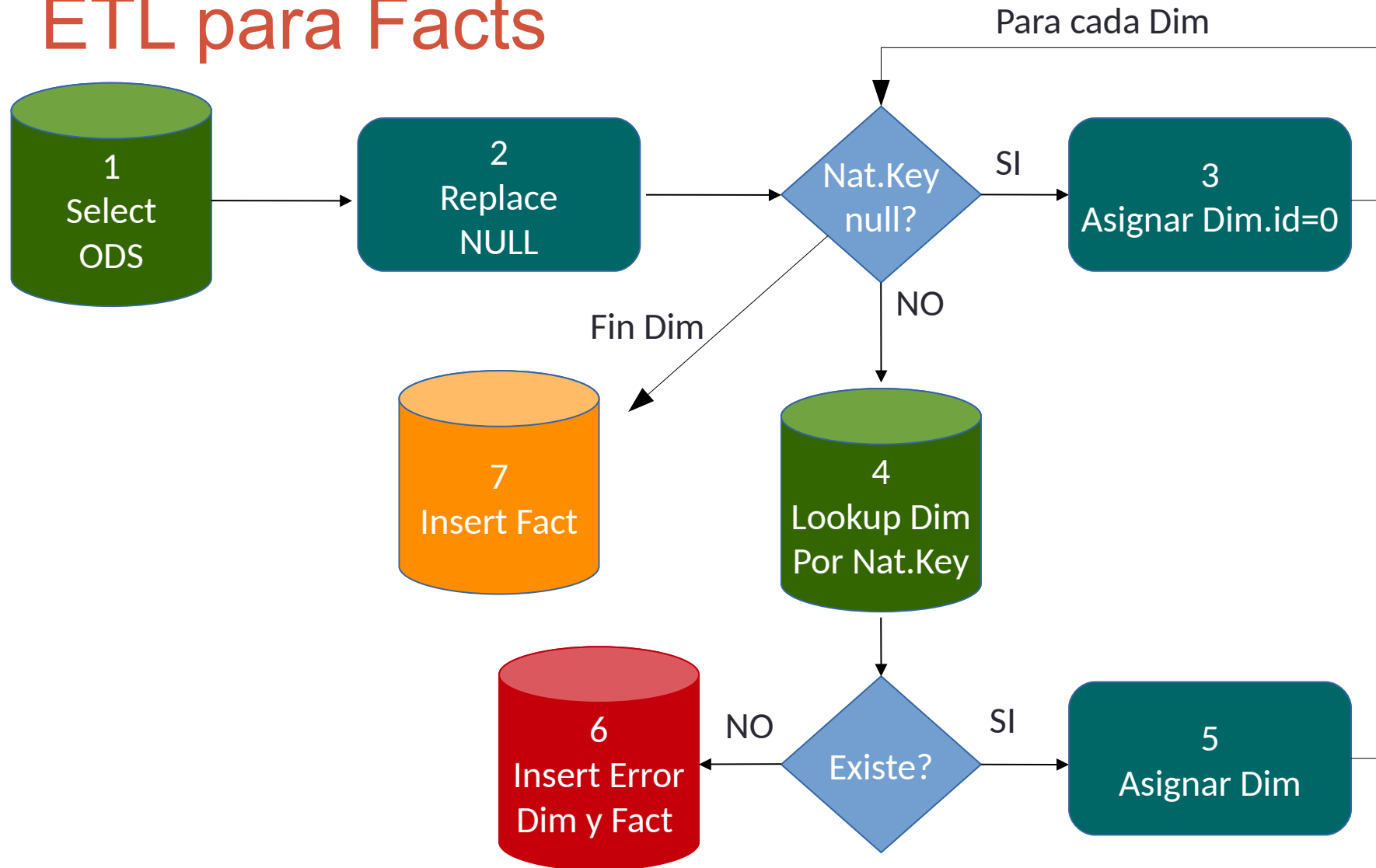
Diseño: Accumulating Snapshot
Ej: Inventario

| Date Received Dimension |

| Date Inspected Dimension |

| Date Placed in Inventory Dimension |

| Date Authorized to Sell Dimension |

| Date Picked Dimension |

| Date Boxed Dimension |

| Date Shipped Dimension |

| Date of Last Return Dimension |

Warehouse Inventory Accumulating Fact

Date Received Key (FK)
Date Inspected Key (FK)
Date Placed in Inventory Key (FK)
Date Authorized to Sell Key (FK)
Date Picked Key (FK)
Date Boxed Key (FK)
Date Shipped Key (FK)
Date of Last Return Key (FK)
Product Key (FK)
Warehouse Key (FK)
Vendor Key (FK)
Quantity Received
Quantity Inspected
Quantity Returned to Vendor
Quantity Placed in Bin
Quantity Authorized to Sell
Quantity Picked
Quantity Boxed
Quantity Shipped
Quantity Returned by Customer
Quantity Returned to Inventory
Quantity Damaged
Quantity Lost
Quantity Written Off
Unit Cost
Unit List Price
Unit Average Price
Unit Recovery Price

| Product Dimension |

| Warehouse Dimension |

| Vendor Dimension |

Características:
* Orientado a un proceso
* 1 hecho por proceso, atributos se actualizan cuando hay un cambio de estado

ETL para Dimensiones

ETL para Facts

Referencias

- Kimball, R. Ross, M. The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. 3rd Ed, Wiley. 2013

- Datawarehouse4u.info. URL: http://datawarehouse4u.info/

- Dataprix.com. URL: http://www.dataprix.com/arquitectura-data-warehouse-areas-datos-nuestro-almacen-corporativo