

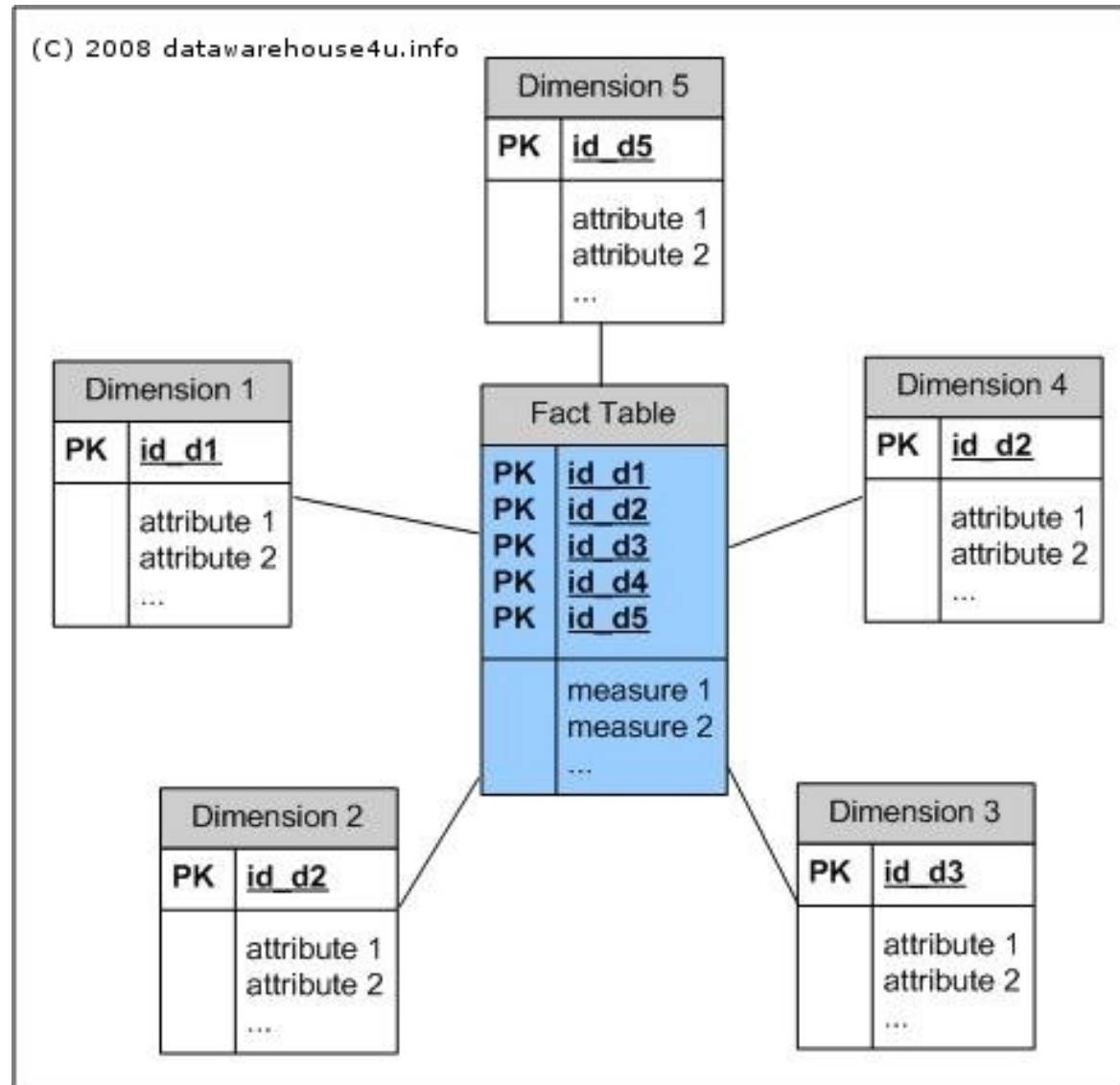
DISEÑO MULTIDIMENSIONAL

Ing. Julio Paciello

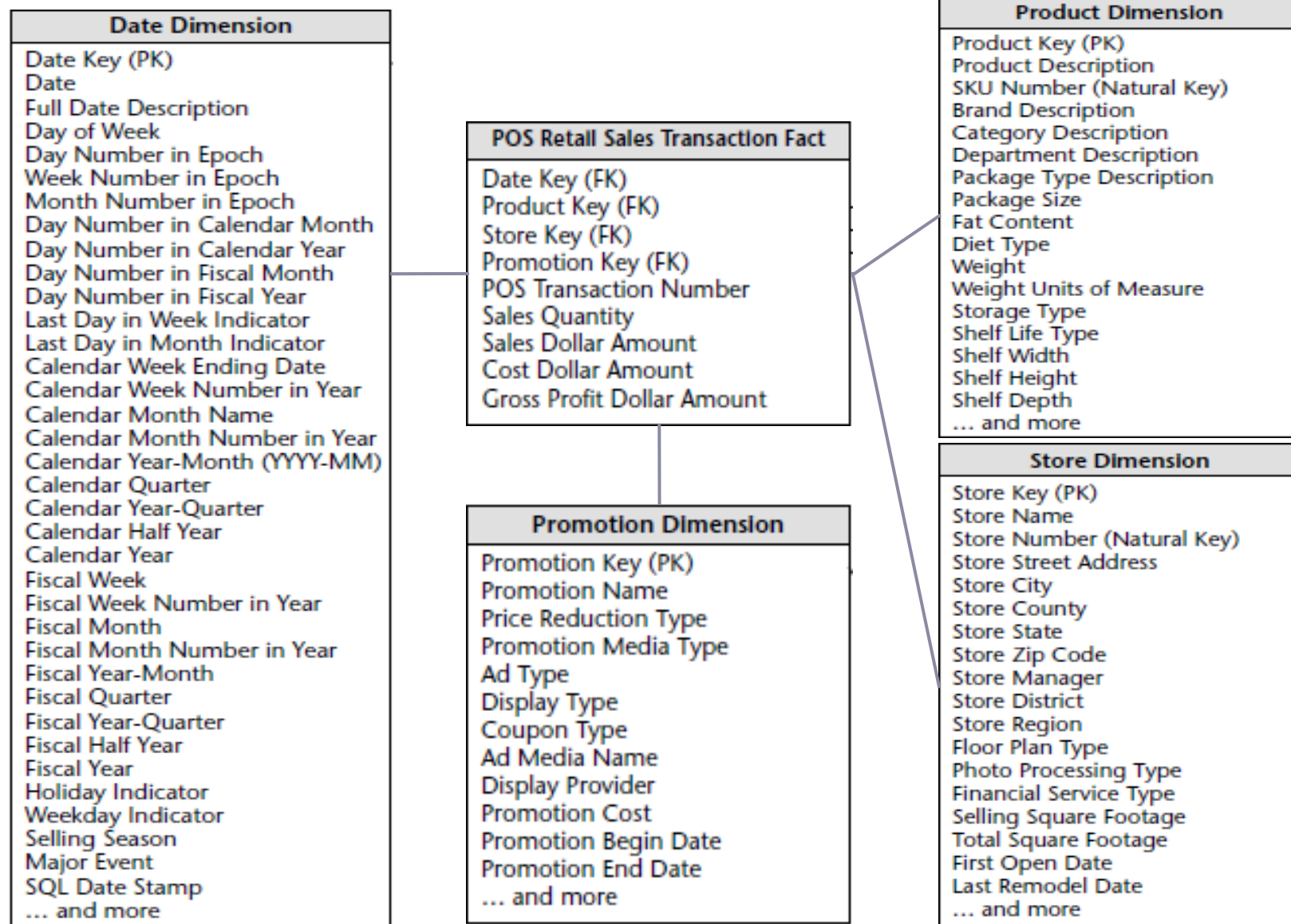
juliopaciello@cds.com.py

Diseño Multidimensional

Star scheme



Ejemplo Práctico: Ventas



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	Friqid	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

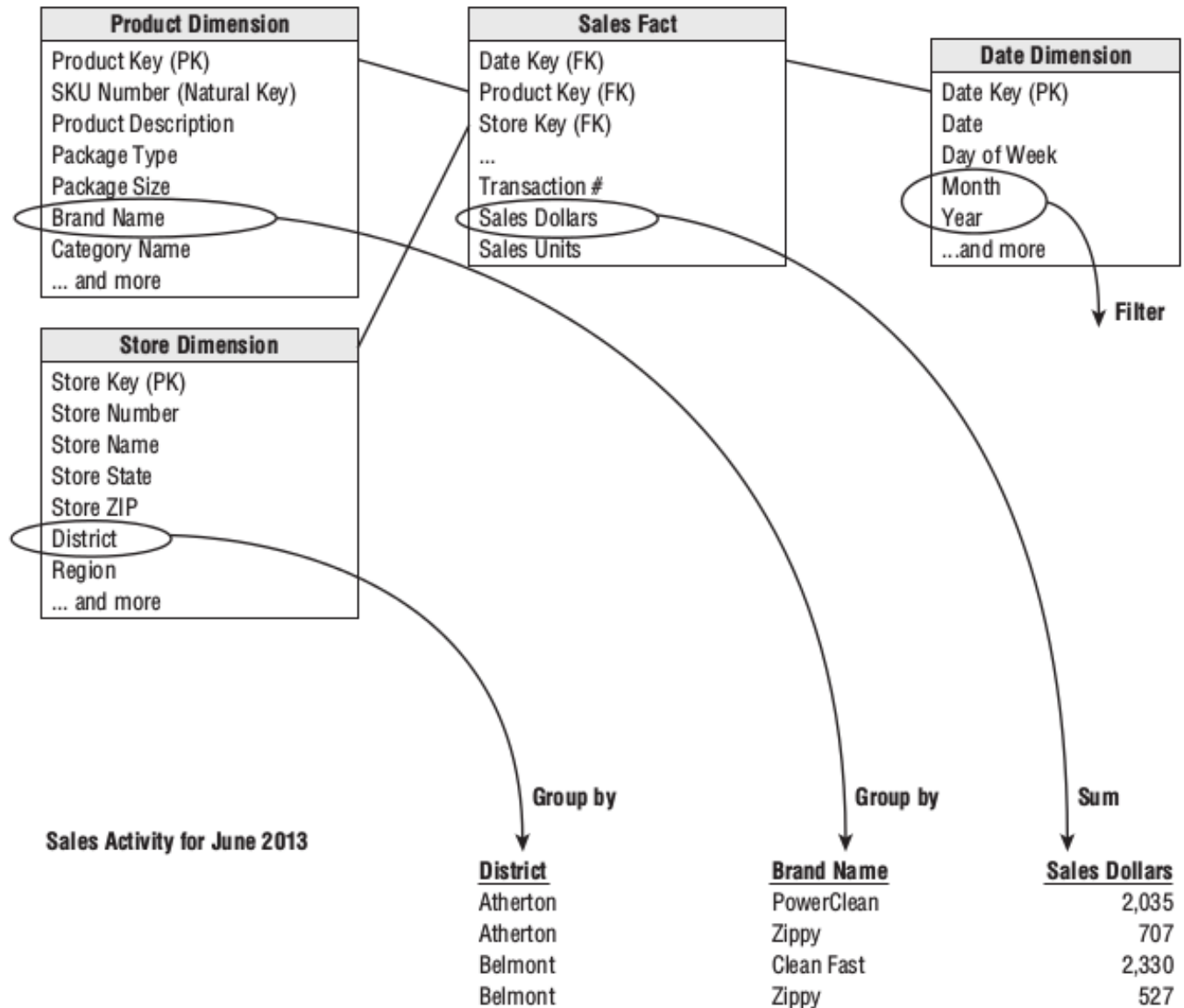
Hechos, ejemplos *drill-down*

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



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:

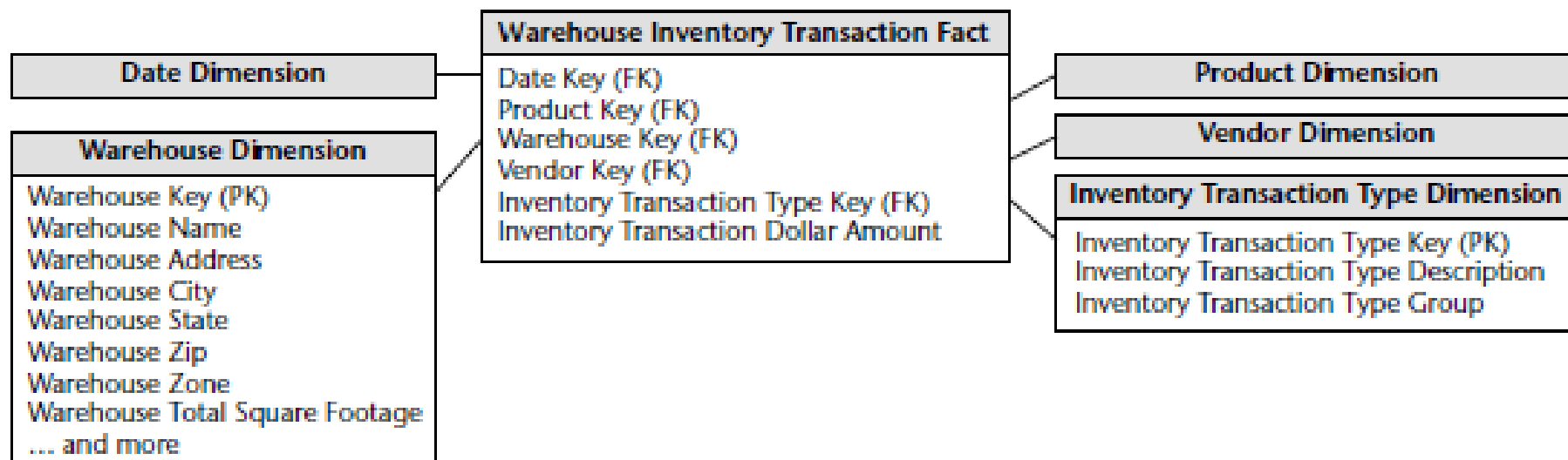
Product Key	SKU (NK)	Product Description	Department Name	Prior Department Name
12345	ABC922-Z	IntelliKidz	Strategy	Education

Tablas de Hechos

- Transactional
- Periodic Snapshot
- Accumulating Snapshot

Diseño: Transactional

Ej: Inventario

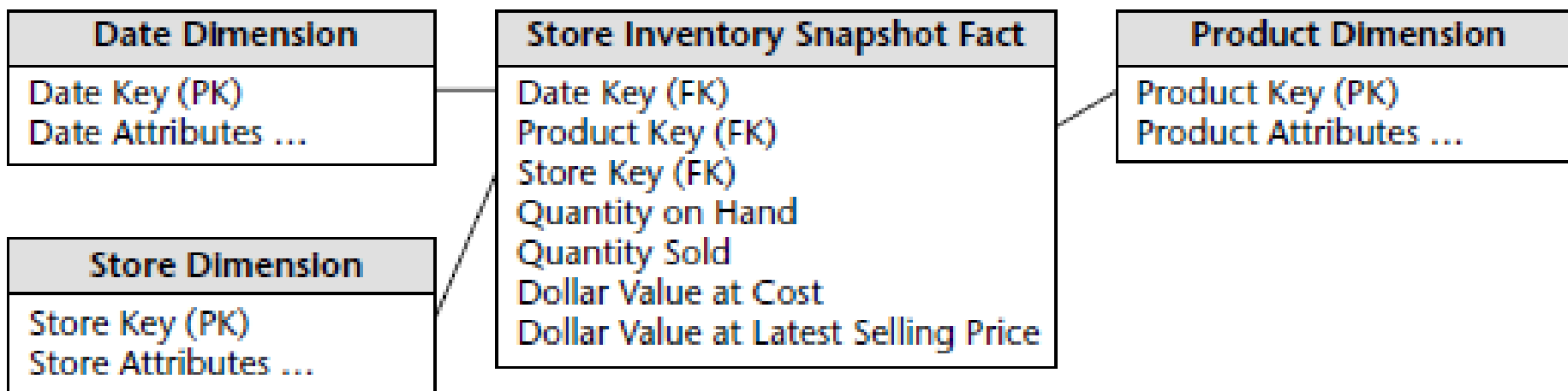


Características:

- * Orientado a transacciones
- * 1 hecho por transacción

Diseño: Periodic Snapshot

Ej: Inventario

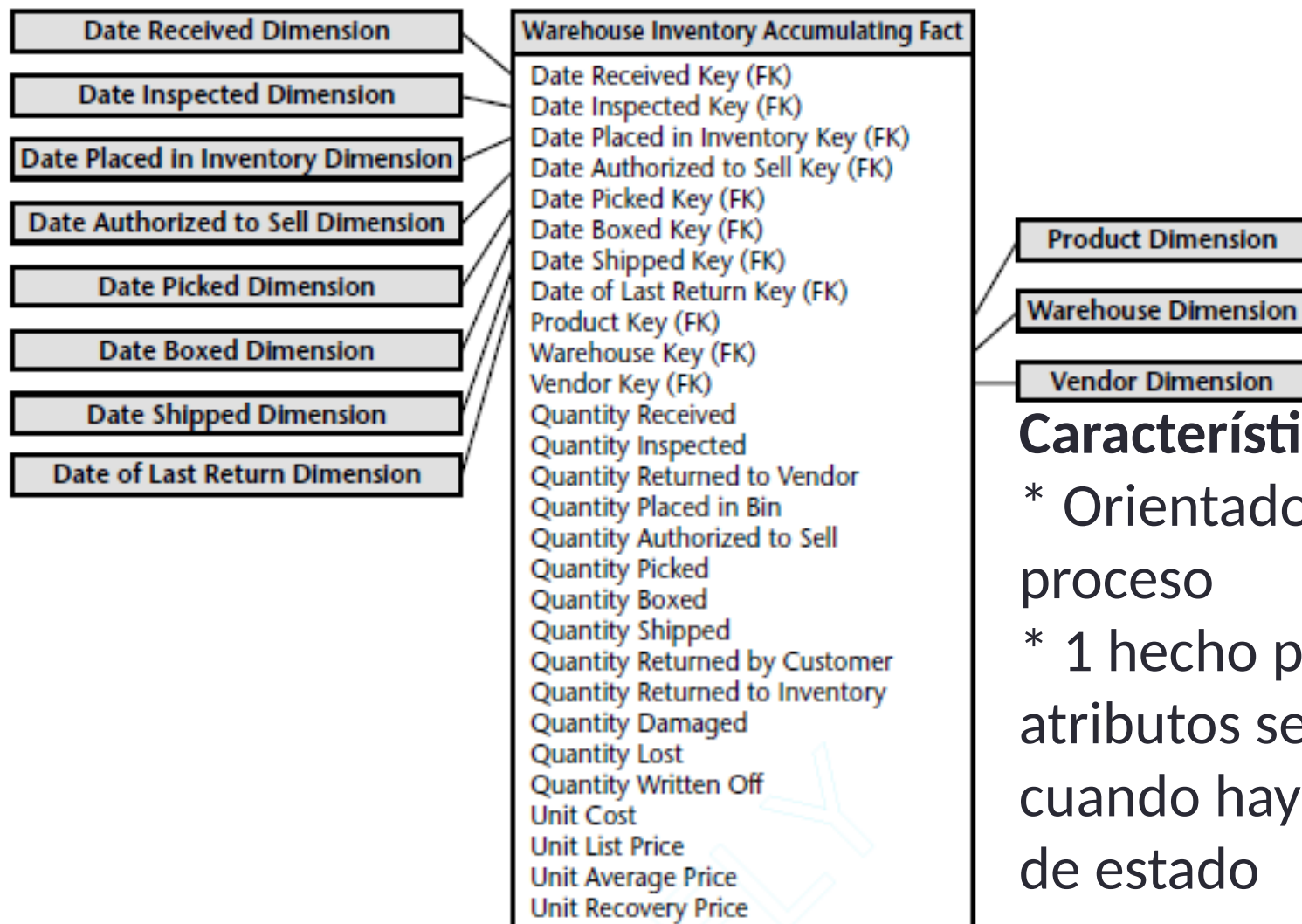


Características:

- * Orientado a medir niveles periódicamente
- * 1 hecho por medición (ej: medición diaria de stock de productos)

Diseño: Accumulating Snapshot

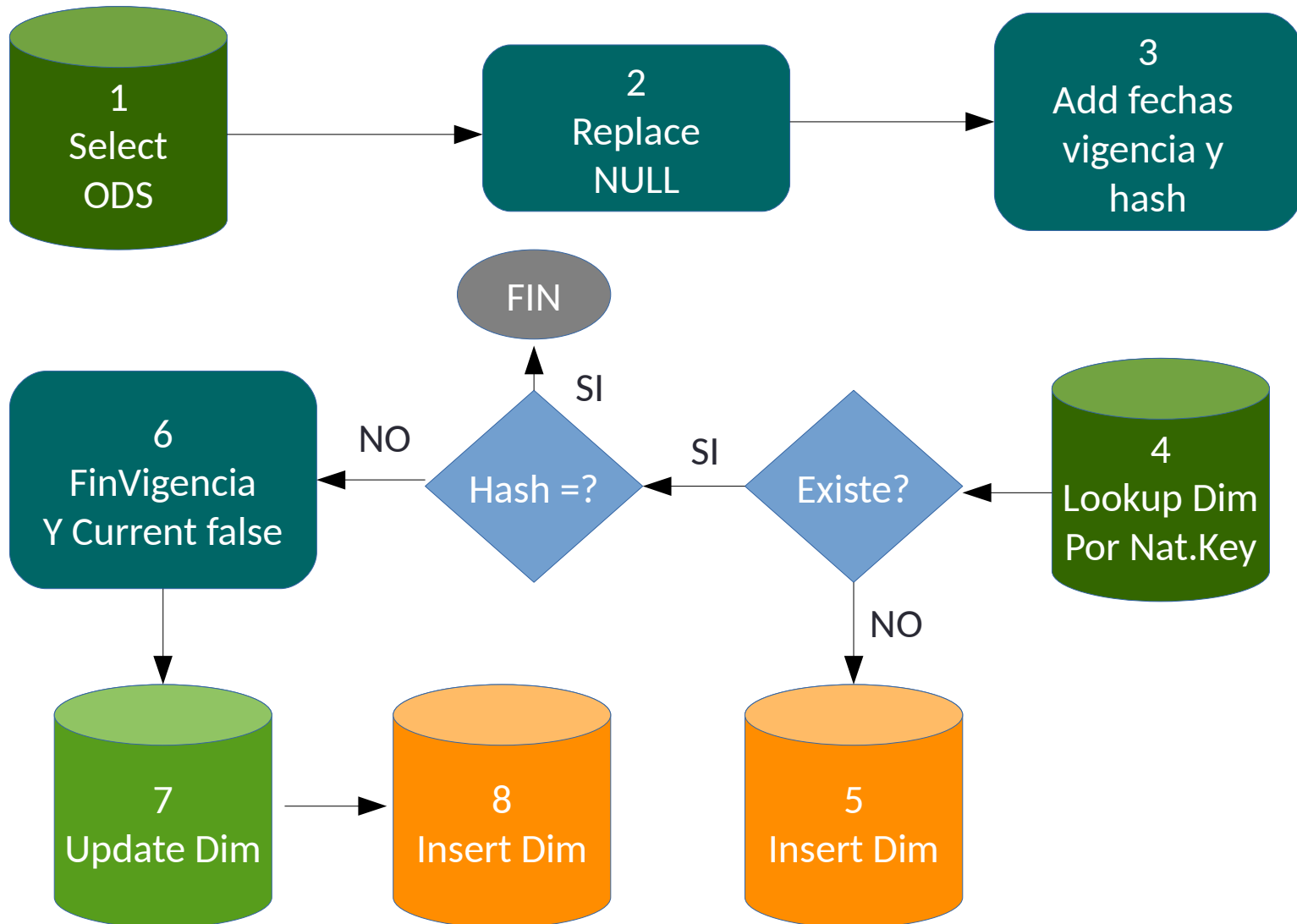
Ej: Inventario



Características:

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

ETL para Dimensiones



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-are-as-datos-nuestro-almacen-corporativo>