**U**niversidad Técnica Nacional, sede San Carlos

Proyecto 1

Minería de Datos

Esteban Enrique Salas Guzmán

2021

Como consultor de la empresa, Grupo Innova S.A, usted fue contratado para definir los requerimientos para la construcción de un datawarehouse que permita conocer información de temas relevantes como ventas mensuales, rendimiento de las sucursales, rendimientos de los empleados, márgenes de utilidad neta, entre otros, con la finalidad de basar sus decisiones estratégicas en datos.

**Metodología a usar**

…

**Modelos de William H. Inmon y Ralph Kimball**

El datawarehouse de Kimball está orientado a la consulta de la información, por lo que su estructura interna está especialmente diseñada para garantizar una explotación de los datos rápida y sencilla, no requiriendo usuarios especializados para ello.

El datawarehouse de Inmon persigue la integración de todos los datos de la compañía, estando orientado hacia el almacenaje de grandes volúmenes de datos, por lo que su estructura interna normalizada se diseña para evitar la redundancia de datos, simplificar las labores de mantenimiento, etc. cuestiones que complican las consultas de la información, requiriendo que los usuarios finales estén mucho más especializados.

Así, es posible decir que el enfoque de Kimball se ajusta más a proyectos pequeños en los que se persiga un sistema fácilmente explotable y entendible por el usuario y de rápido desarrollo, siendo el modelo de Inmon más apropiado para sistemas complejos de mayor envergadura.

|  |  |  |
| --- | --- | --- |
|  | Inmon | Kimball |
| Presupuesto | Coste inicial alto | Coste inicial bajo |
| Plazos | Requiere más tiempo de desarrollo | Tiempo de desarrollo inferior |
| Expertise | Equipo con especialización alta | Equipo con especialización media |
| Alcance | Toda la compañía | Departamentos individuales |
| Mantenimiento | Fácil mantenimiento | Mantenimiento más complejo |

**Base de datos original (Oracle 11g)**

-- SUCURSALES

CREATE TABLE PROYECTO1.SUCURSALES (

    ID INT NOT NULL,

    CODIGO INT NOT NULL,

    NOMBRE VARCHAR(50) NOT NULL,

    PRIMARY KEY (ID)

);

CREATE SEQUENCE PROYECTO1.SUCURSALES\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.SUCURSALES\_ID

    BEFORE INSERT ON PROYECTO1.SUCURSALES FOR EACH ROW

BEGIN

    SELECT PROYECTO1.SUCURSALES\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

-- PERSONAS

CREATE TABLE PROYECTO1.PERSONAS (

    ID INT NOT NULL,

    CEDULA INT NOT NULL,

    NOMBRE VARCHAR(50) NOT NULL,

    APELLIDOS VARCHAR(50) NOT NULL,

    CORREO VARCHAR(50) NOT NULL,

    FEC\_NACIMIENTO DATE NOT NULL,

    ESTADO\_CIVIL INT NOT NULL,

    SEXO INT NOT NULL,

    OCUPACION INT NOT NULL,

    EDUCACION INT NOT NULL,

    HIJOS INT NOT NULL,

    PRIMARY KEY (ID)

);

CREATE SEQUENCE PROYECTO1.PERSONAS\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.PERSONAS\_ID

    BEFORE INSERT ON PROYECTO1.PERSONAS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.PERSONAS\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

--EMPLEADOS

CREATE TABLE PROYECTO1.EMPLEADOS (

    ID INT NOT NULL,

    PERSONA\_ID INT NOT NULL,

    CODIGO INT NOT NULL,

    SALARIO INT NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT EMPLEADO\_PERSONA\_ID\_FK FOREIGN KEY (PERSONA\_ID)

        REFERENCES PROYECTO1.PERSONAS (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.EMPLEADOS\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.EMPLEADOS\_ID

    BEFORE INSERT ON PROYECTO1.EMPLEADOS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.EMPLEADOS\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

CREATE SEQUENCE PROYECTO1.EMPLEADOS\_CODIGO\_SEQ;

CREATE TRIGGER PROYECTO1.EMPLEADOS\_CODIGO

    BEFORE INSERT ON PROYECTO1.EMPLEADOS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.EMPLEADOS\_CODIGO\_SEQ.nextval

    INTO :new.CODIGO

    FROM dual;

END;

-- CLIENTES

CREATE TABLE PROYECTO1.CLIENTES (

    ID INT NOT NULL,

    PERSONA\_ID NOT NULL,

    CODIGO INT NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT CLINTE\_PERSONA\_ID\_FK FOREIGN KEY (PERSONA\_ID)

        REFERENCES PROYECTO1.PERSONAS (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.CLIENTES\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.CLIENTES\_ID

    BEFORE INSERT ON PROYECTO1.CLIENTES FOR EACH ROW

BEGIN

    SELECT PROYECTO1.CLIENTES\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

CREATE SEQUENCE PROYECTO1.CLIENTES\_CODIGO\_SEQ;

CREATE TRIGGER PROYECTO1.CLIENTES\_CODIGO

    BEFORE INSERT ON PROYECTO1.CLIENTES FOR EACH ROW

BEGIN

    SELECT PROYECTO1.CLIENTES\_CODIGO\_SEQ.nextval

    INTO :new.CODIGO

    FROM dual;

END;

--PROVINCIAS

CREATE TABLE PROYECTO1.PROVINCIAS (

    ID INT NOT NULL,

    CODIGO INT NOT NULL,

    NOMBRE VARCHAR(50) NOT NULL,

    PRIMARY KEY (ID)

);

CREATE SEQUENCE PROYECTO1.PROVINCIAS\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.PROVINCIA\_ID

    BEFORE INSERT ON PROYECTO1.PROVINCIAS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.PROVINCIAS\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

--CANTONES

CREATE TABLE PROYECTO1.CANTONES (

    ID INT NOT NULL,

    PROVINCIA\_ID INT NOT NULL,

    CODIGO INT NOT NULL,

    NOMBRE VARCHAR(50) NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT CANTON\_PROVINCIA\_ID\_FK FOREIGN KEY (PROVINCIA\_ID)

        REFERENCES PROYECTO1.PROVINCIAS (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.CANTONES\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.CANTONES\_ID

    BEFORE INSERT ON PROYECTO1.CANTONES FOR EACH ROW

BEGIN

    SELECT PROYECTO1.CANTONES\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

--DISTRITOS

CREATE TABLE PROYECTO1.DISTRITOS (

    ID INT NOT NULL,

    CANTON\_ID INT NOT NULL,

    CODIGO INT NOT NULL,

    NOMBRE VARCHAR(50) NOT NULL,

    CODIGO\_POSTAL INT NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT DISTRITO\_CANTON\_ID FOREIGN KEY (CANTON\_ID)

        REFERENCES PROYECTO1.CANTONES (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.DISTRITOS\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.DISTRITOS\_ID

    BEFORE INSERT ON PROYECTO1.DISTRITOS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.DISTRITOS\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

-- DIRECCIONES

CREATE TABLE PROYECTO1.DIRECCIONES (

    ID INT NOT NULL,

    PERSONA\_ID INT NOT NULL,

    DISTRITO\_ID INT NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT DIRECCION\_PERSONA\_ID FOREIGN KEY (PERSONA\_ID)

        REFERENCES PROYECTO1.PERSONAS (ID) ON DELETE CASCADE,

    CONSTRAINT DIRECCION\_DISTRITO\_ID FOREIGN KEY (DISTRITO\_ID)

        REFERENCES PROYECTO1.DISTRITOS (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.DIRECCIONES\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.DIRECCIONES\_ID

    BEFORE INSERT ON PROYECTO1.DIRECCIONES FOR EACH ROW

BEGIN

    SELECT PROYECTO1.DIRECCIONES\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

--PROVEEDORES

CREATE TABLE PROYECTO1.PROVEEDORES (

    ID INT NOT NULL,

    DISTRITO\_ID INT NOT NULL,

    CODIGO INT NOT NULL,

    NOMBRE VARCHAR(50) NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT PROVEEDOR\_DISTRITO\_ID FOREIGN KEY (DISTRITO\_ID)

        REFERENCES PROYECTO1.DISTRITOS (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.PROVEEDORES\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.PROVEEDORES\_ID

    BEFORE INSERT ON PROYECTO1.PROVEEDORES FOR EACH ROW

BEGIN

    SELECT PROYECTO1.PROVEEDORES\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

CREATE SEQUENCE PROYECTO1.PROVEEDORES\_CODIGO\_SEQ;

CREATE TRIGGER PROYECTO1.PROVEEDORES\_CODIGO

    BEFORE INSERT ON PROYECTO1.PROVEEDORES FOR EACH ROW

BEGIN

    SELECT PROYECTO1.PROVEEDORES\_CODIGO\_SEQ.nextval

    INTO :new.CODIGO

    FROM dual;

END;

-- PRODUCTOS

CREATE TABLE PROYECTO1.PRODUCTOS (

    ID INT NOT NULL,

    PROVEEDOR\_ID INT NOT NULL,

    CODIGO INT NOT NULL,

    NOMBRE VARCHAR (50) NOT NULL,

    PRECIO\_COMPRA INT NOT NULL,

    PRECIO\_VENTA INT NOT NULL,

    INVENTARIO INT NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT PRODUCTO\_PROVEEDOR\_ID FOREIGN KEY (PROVEEDOR\_ID)

        REFERENCES PROYECTO1.PROVEEDORES (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.PRODUCTOS\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.PRODUCTOS\_ID

    BEFORE INSERT ON PROYECTO1.PRODUCTOS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.PRODUCTOS\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

-- FECHAS

CREATE TABLE PROYECTO1.FECHAS (

    ID INT NOT NULL,

    FECHA DATE NOT NULL,

    MES INT NOT NULL,

    NOMBRE\_MES VARCHAR (20) NOT NULL,

    ANNIO INT NOT NULL,

    DIA\_SEMANA INT NOT NULL,

    DIA\_ANNIO INT NOT NULL,

    PRIMARY KEY (ID)

);

CREATE SEQUENCE PROYECTO1.FECHAS\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.FECHAS\_ID

    BEFORE INSERT ON PROYECTO1.FECHAS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.FECHAS\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

-- VENTAS

CREATE TABLE PROYECTO1.VENTAS (

    ID INT NOT NULL,

    SUCURSAL\_ID INT NOT NULL,

    PRODUCTO\_ID INT NOT NULL,

    CLIENTE\_ID INT NOT NULL,

    EMPLEADO\_ID INT NOT NULL,

    FECHA\_ID INT NOT NULL,

    CODIGO INT NOT NULL,

    CANTIDAD INT NOT NULL,

    TOTAL INT NOT NULL,

    PRIMARY KEY (ID),

    CONSTRAINT VENTA\_SUCURSAL\_ID FOREIGN KEY (SUCURSAL\_ID)

        REFERENCES PROYECTO1.SUCURSALES (ID) ON DELETE CASCADE,

    CONSTRAINT VENTA\_PRODUCTO\_ID FOREIGN KEY (PRODUCTO\_ID)

        REFERENCES PROYECTO1.PRODUCTOS (ID) ON DELETE CASCADE,

    CONSTRAINT VENTA\_CLIENTE\_ID FOREIGN KEY (CLIENTE\_ID)

        REFERENCES PROYECTO1.CLIENTES (ID) ON DELETE CASCADE,

    CONSTRAINT VENTA\_EMPLEADO\_ID FOREIGN KEY (EMPLEADO\_ID)

        REFERENCES PROYECTO1.EMPLEADOS (ID) ON DELETE CASCADE,

    CONSTRAINT VENTA\_FECHA\_ID FOREIGN KEY (FECHA\_ID)

        REFERENCES PROYECTO1.FECHAS (ID) ON DELETE CASCADE

);

CREATE SEQUENCE PROYECTO1.VENTAS\_ID\_SEQ;

CREATE TRIGGER PROYECTO1.VENTAS\_ID

    BEFORE INSERT ON PROYECTO1.VENTAS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.VENTAS\_ID\_SEQ.nextval

    INTO :new.ID

    FROM dual;

END;

CREATE SEQUENCE PROYECTO1.VENTAS\_CODIGO\_SEQ;

CREATE TRIGGER PROYECTO1.VENTAS\_CODIGO

    BEFORE INSERT ON PROYECTO1.VENTAS FOR EACH ROW

BEGIN

    SELECT PROYECTO1.VENTAS\_CODIGO\_SEQ.nextval

    INTO :new.CODIGO

    FROM dual;

END;

**Transferencia al Staging Area**

…

**Transferencia al Datawarehouse**

…

**Estructuras OLAP**

* **Multidimensional**

…

* **Tabular**

…

**Comparativa de las estructuras OLAP**

…