**EVALUACIÓN 1 P2 UND 1**

**ASIGNATURA: INTELIGENCIA DE NEGOCIOS**

**PROFESOR: CRISTIAN SALAZAR**

**SECCIÓN: INY6101-002D**

|  |  |  |
| --- | --- | --- |
| **PUNTAJE MÁXIMO** | **6 puntos** | **NOTA** |
| **PUNTAJE OBTENIDO** |  |
| **FECHA** |  |
| **DURACIÓN** | **POR ENCARGO** |
| **Integrantes** | **Daniel García.**  **Ignacio Salazar.** |  |

**INDICACIONES GENERALES**

Evidencia 1.2: Entrega de Encargo. Relevancia: 30% (Entregado por AVA)

**ENTREGA:**

Las respuestas de las preguntas y la evidencia de la construcción del DW y ETL debe quedar en este informe como evidencia con imágenes de pantalla

También se debe entregar la base de datos DW en formato (.SQL) y el proyecto ETL (.rar)

Mencionar en comentario los nombres de los integrantes

**APRENDIZAJES ESPERADOS**

* Diseñar el modelo conceptual que represente los requerimientos de información de la organización según una definición escalable.
* Implementar el modelo dimensional en la base de datos generando las condiciones para el almacenamiento, procesamiento, manipulación y consulta de datos.
* Implementar un proceso de extracción, transformación y carga de datos para poblar el modelo dimensional construido.
* Trabajar en equipo para alcanzar los objetivos y soluciones a los problemas planteados.
* Reconocer los elementos de modelado y diseño para el desarrollo de software de acuerdo con los fundamentos de la Ingeniería del Software.

**1. Primera parte** Conteste las siguientes preguntas:

1. ¿Qué beneficios traería la implementación de una iniciativa de datawarehousing y OLAP en una empresa?

2. Qué desafíos presenta el mercado actual a la hora de implementar una infraestructura de BI. Considere al menos los siguientes aspectos para realizar su análisis de la situación actual y construir su respuesta:

a. Realidad de los sistemas transaccionales en términos de integración, limpieza, completitud y correctitud de los datos.

b. Existencia y madurez de una estrategia analítica a nivel empresarial.

c. Necesidades de las potenciales áreas usuarias.

d. Existencia de las competencias necesarias para implementar la plataforma (DW/DM, Cubos OLAP).

1. **Segunda parte** Diseñe el siguiente modelo estrella y realice el ETL de la dimensión indicada:
2. Kent Foods ha crecido rápidamente durante el último año, por esta razón su propietario ha decidido incorporar la Inteligencia de Negocios a su compañía. Usted ha sido escogido para liderar esta iniciativa, y permitir a los encargados tomar las mejores decisiones tácticas y estratégicas necesarias para lograr que la organización siga en el buen camino.

En base a un modelo **RELACIONAL ACTUAL** (KentFoods.bak). El cual contiene gran histórico de registros almacenados. En base esto Diseñe un diagrama de **ESTRELLA** con una tabla de hechos (**Ordenes**) que contenga claves de referencia de algún proceso importante de su empresa (**tiempo, proveedores, transportista, empleados**). **Mencione en la respuesta, qué otras tablas dimensionales serían buenas candidatas (para ordenes) a ser conformed dimensions y por qué**. (los campos de medidas de la tabla de hecho, y los campos dentro de las tablas de dimensiones son a criterio de cada grupo).

DESPUES DE CREADO EL MODELO ESTRELLA, DEBERA HACER EL SIGUIENTE ETL:

**Origen**: MODELO RELACIONAL DEBERA UNIR (CONSULTA SQL – INNER JOIN) de las tablas:

[dbo].[Regiones]

[dbo].[Territorios]

[dbo].[TerritoriosEmpleados]

[dbo].[Empleados]

De las cuales solo necesitaremos transportar vía ETL los siguientes campos a la tabla EMPLEADO (modelo estrella).

**Destino**: Campos necesarios DIM EMPLEADO MODELO ESTRELLA:

[**EmpleadoID**] [int] NOT NULL

[**Apellido**] [nvarchar](20) NOT NULL

[**Nombre**] [nvarchar](10) NOT NULL

[**Cargo**] [nvarchar](30) NULL

[**Direccion**] [nvarchar](60) NULL

[**Ciudad**] [nvarchar](15) NULL

[**Pais**] [nvarchar](15) NULL

[**Territorio**] [nchar](50) NOT NULL

[**Region**] [nchar](50) NOT NULL

**Desarrollo**

**Primera Parte**

**1. ¿Qué beneficios traería la implementación de una iniciativa de datawarehousing y OLAP en una empresa?**

Entre los principales beneficios que conlleva la implementación de una iniciativa de data warehouse y OLAP serian la mejora en la calidad de las decisiones claves tomadas por parte del negocio, mejora en el poder de procesamiento, alto nivel de utilidad a mediano y largo plazo y un retorno de inversión más rápido que otros métodos de implementación disponibles en el mercado.

**2. ¿Qué desafíos presenta el mercado actual a la hora de implementar una infraestructura de BI?**

El principal desafío que se presenta al momento de implementar una infraestructura de inteligencia de negocios en una empresa por primera vez es el estado actual de la base de datos en el negocio. En la mayoría de los casos, la información almacenada en esta no se encuentra en un estado que permita su fácil traslado a una base de datos dimensional, lo cual le puede costar a la empresa tiempo y grandes sumas de dinero en el caso que el equipo encargado no logre superar este problema rápidamente.

El educar al personal asociado es otro gran desafío al momento de implementar un modelo dimensional, ya que el poco conocimiento en las estructuras de estos tiende a generar rechazos. Es tarea del área de inteligencia de negocios el promover las ventajas y bondades que este método permite, demostrando con información real que les permita tomar decisiones analizando ya sea registros históricos de la compañía o movimientos del mercado, involucrando a los directivos a que se acerquen paulatinamente a la inteligencia de negocios.

Es responsabilidad del área de inteligencia de negocios el analizar los requerimientos de información que necesitan las personas para una óptima toma de decisiones. Al brindar información clara, precisa y concisa que los ayude, será posible que el enfoque de negocio prevalezca por sobre la incomodidad de aprender esta metodología. El área debe brindar facilidades con respecto a la forma de interpretar los datos, el cómo visualizarlos por medio de cuadros de mando unificados, que estos puedan analizar la información de la compañía o del mercado en el cual están involucrados de forma rápida, acercando cada vez más a las personas u otras áreas a adoptar este enfoque de análisis de información.

En muchas empresas existe un alto porcentaje de los miembros de una organización que no poseen competencias analíticas de información. Por lo tanto, es un gran desafío para la inteligencia de negocios lograr que los usuarios de la información utilicen esta de forma adecuada. Por este motivo, el capacitar a las personas a adoptar este enfoque es crucial. Las herramientas existentes en el mercado brindan una gran variedad de facilidades para adoptar este método, algunos ejemplos como Excel y PowerBI son amigables con los usuarios, permitiendo un aprendizaje óptimo en poco tiempo, además que permitir visualizar su progreso de forma gradual. Por último, es necesario hacer hincapié en todos los beneficios que se obtienen con la inteligencia de negocios. Tener la información de manera rápida y confiable permite a los usuarios una buena toma de decisiones, dando así más cabida de esta metodología en otras empresas del rubro.

**Segunda Parte**

**Herramientas utilizadas.**

Para la presente evaluación se utilizó el siguiente ambiente de trabajo:

* Docker versión 19.03.4, build 9013bf5.
* Motor de base de datos SQL Server versión 2019 Developer Edition. (Contenedor Docker para Linux versión Ubuntu 16.04, no existe un contenedor soportado por Microsoft para la versión de SQL Server 2012)
* SQL Management Studio versión 18.4 (español).
* Visual Studio 2019 Community Edition.

**1. Nueva dimensión candidata**

Existen 2 dimensiones candidatas que podrían ser aplicadas al modelo, estas serían, dimensión cliente y dimensión transportistas.

Sin embargo, Para el modelo presentado en esta evaluación incluimos estas dimensiones dentro de la tabla de hechos (factOrden), ya que pensamos que podrían ser de mayor utilidad para obtener rápidamente la información asociada a la orden, quien la solicito y donde se envió, ¿Cuáles clientes compraron en tal fecha?, ¿Dónde es más concurrente el envío de los productos? y varias preguntas más asociadas que pueden ser resueltas de este modo, además de acceder de forma fácil tan solo con la tabla de hechos sin necesidad de generar relaciones con otras dimensiones que no sean las expuestas en el modelo.

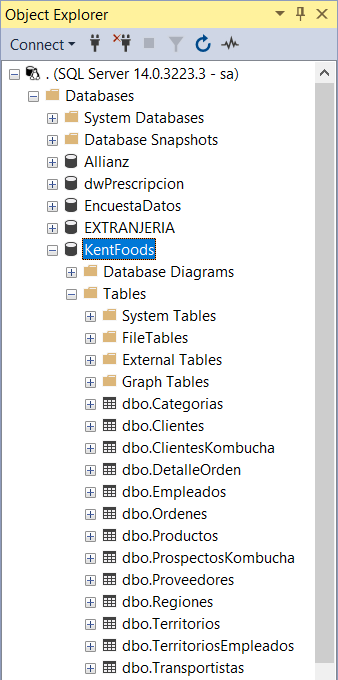
**Evidencia**  


Ilustración 1: Restauración de la base de datos KentFoods.

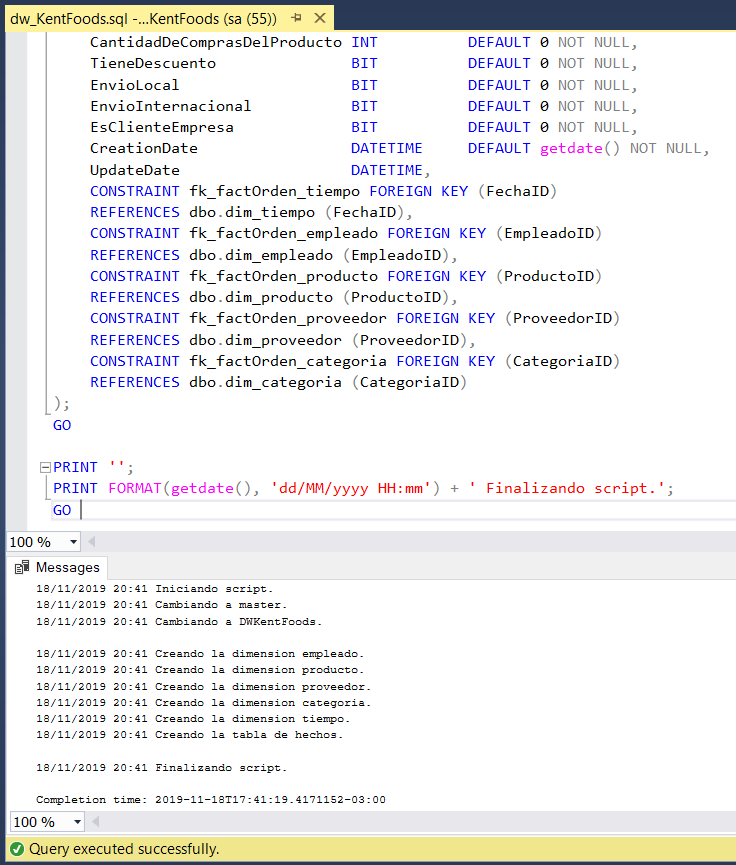


Ilustración 2: Resultado de la ejecución del script de creación de la data warehouse KentFoods.

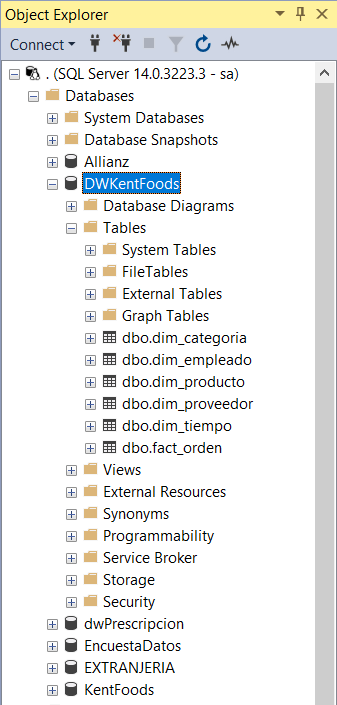


Ilustración 3: Tablas creadas por el script de la data warehouse KentFoods.

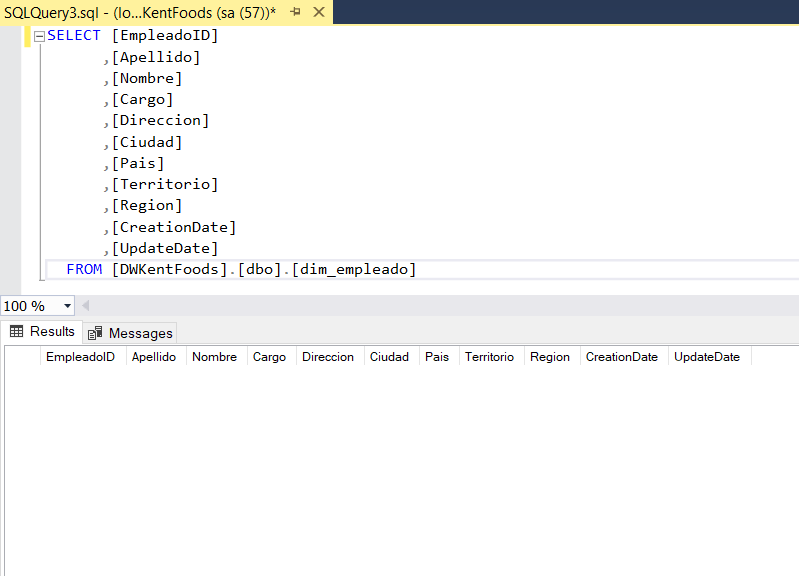


Ilustración 4: Consulta inicial a dimensión vacía de Empleados.

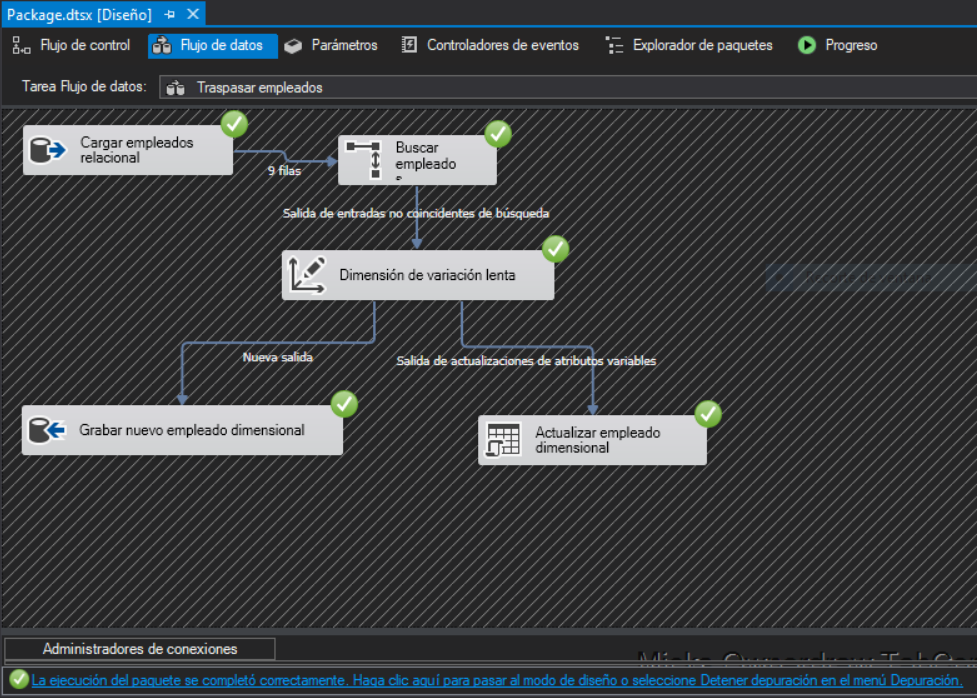


Ilustración 5: Ejecución de la carga de datos de empleados desde el modelo relacional hasta el modelo de data warehouse.

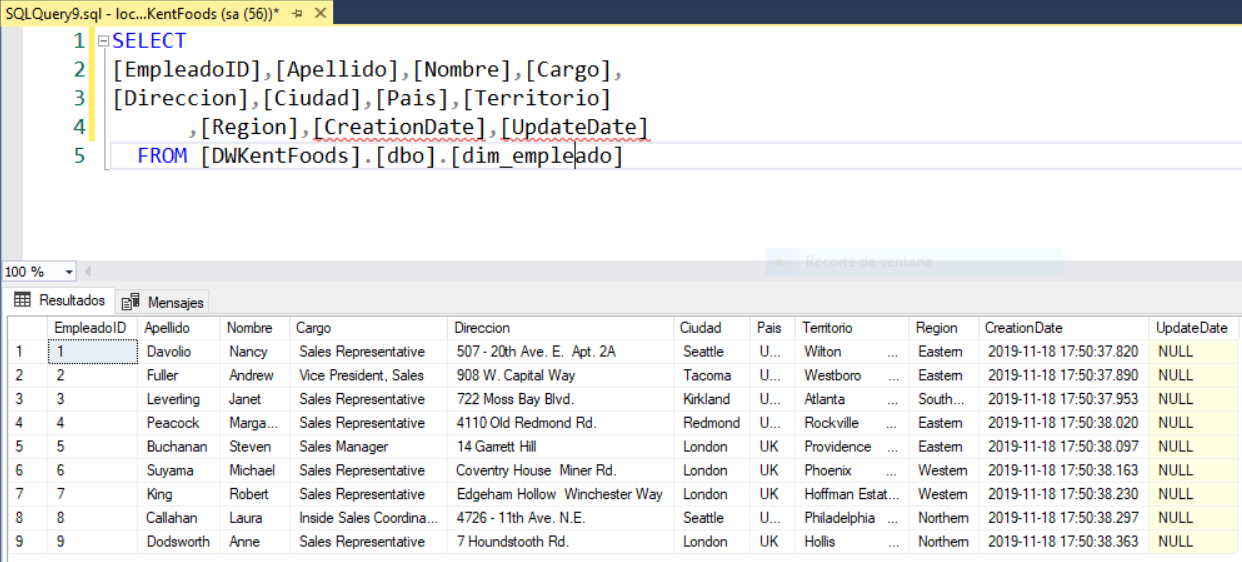


Ilustración 6: Nueva consulta a la dimensión de empleados con datos cargados desde el modelo relacional.

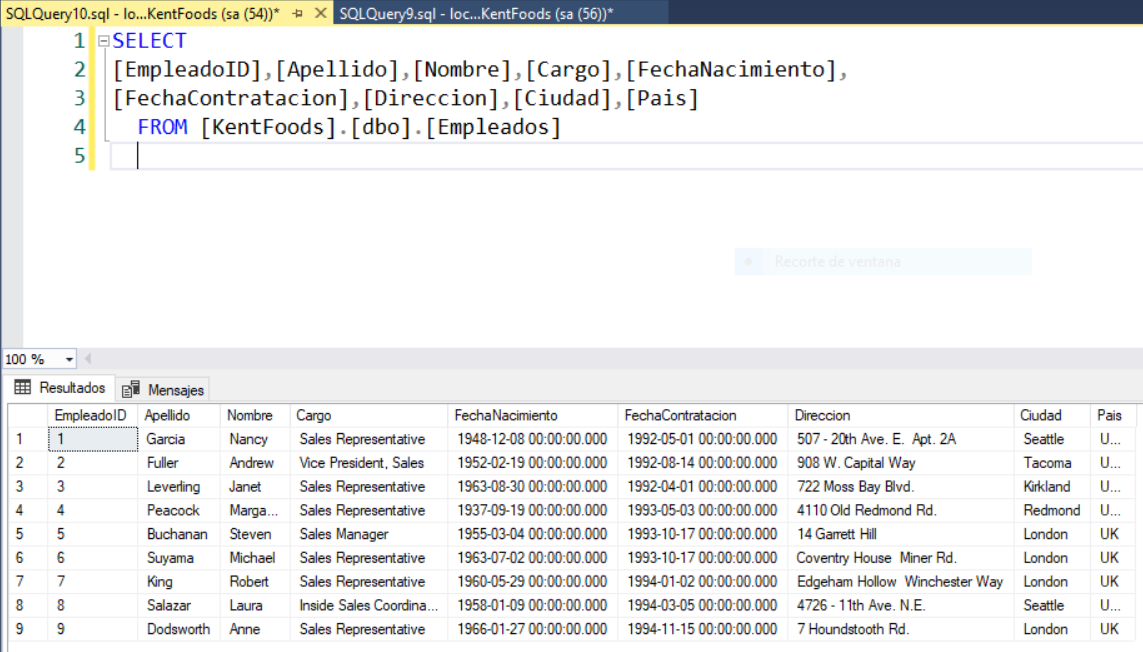


Ilustración 7: Alterando datos de empleados en el modelo relacional.

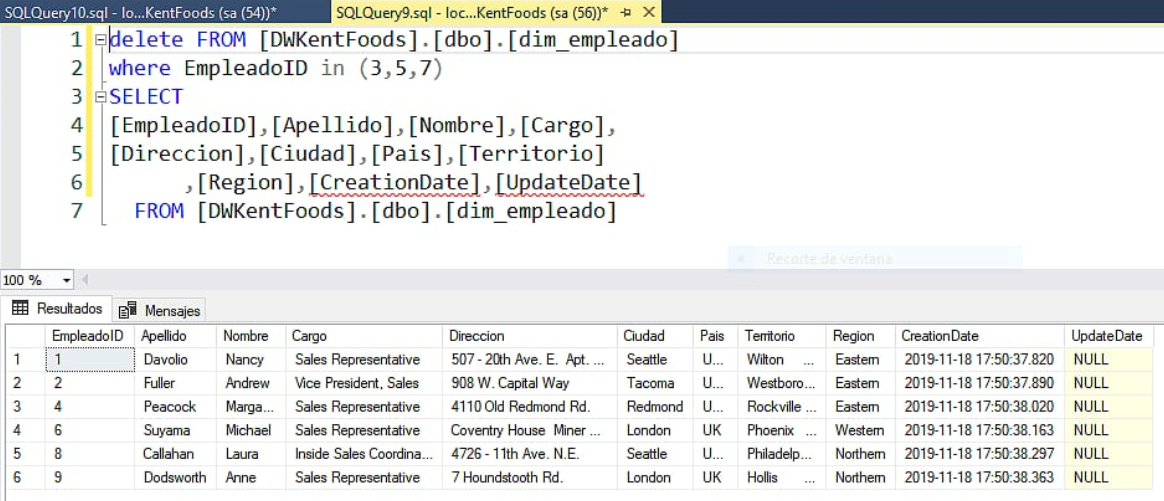


Ilustración 8: Eliminando algunos datos de empleados en el modelo data warehouse. Notese que los empleados 1 y 8 tienen un apellido distinto a los empleados con el mismo ID en el modelo relacional, además de haber eliminado los empleados con id 3,5,7.

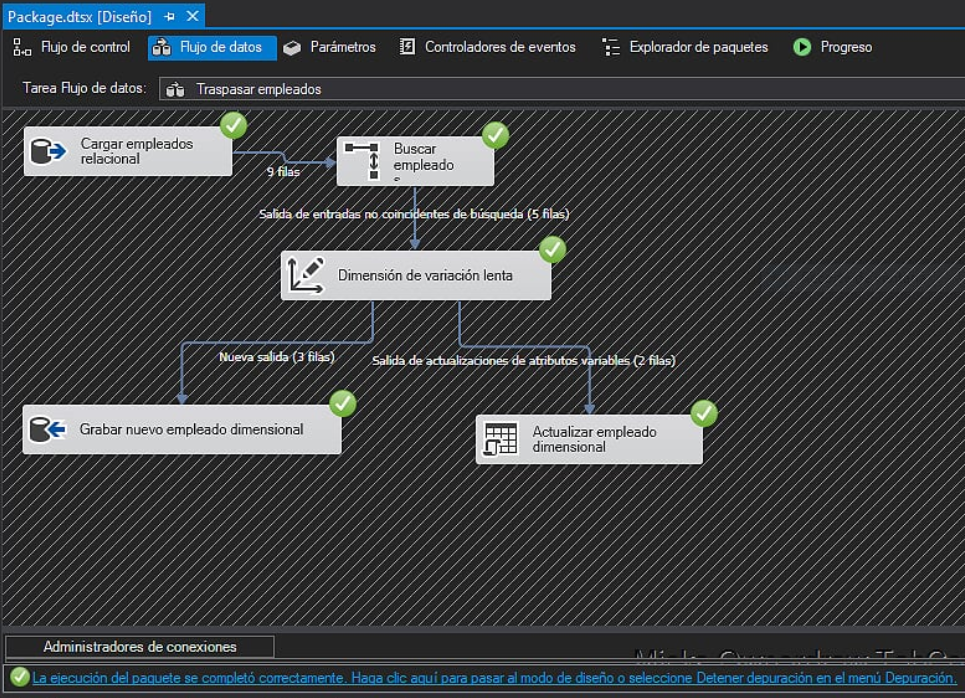


Ilustración 9: Segunda carga de datos desde el modelo relacional hasta el data warehouse.

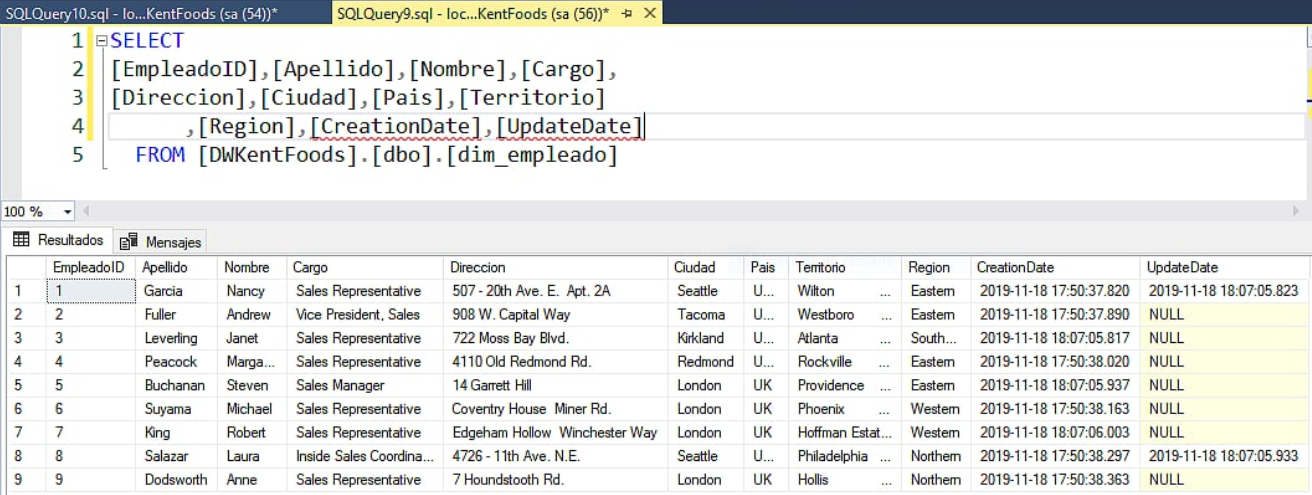


Ilustración 10: Consulta a los datos cargados desde el modelo relacional. Notese los nuevos apellidos de los empleados 1 y 8., aparte de la fecha de modificación que genero el paquete, por otra parte los empleados con id 3,5 y 7 volvieron a estar presentes.

**Estructura del ETL**

En este apartado será explicado el proceso realizado para la carga de datos desde el modelo relacional a la data warehouse, el flujo de control necesarios corresponde a “Traspasar empleados”, el cual contiene los siguientes elementos:

* Cargar empleados del modelo relacional: Componente de tipo “Origen OLEDB”, este elemento contiene las conexiones necesarias para conectar con la base de datos relacional (KentFoods) y posee la consulta requerida para obtener los datos que serán posteriormente almacenados en el modelo dimensional (dwKentFoods).
* Buscar empleados del modelo dimensional: Componente de tipo “Búsqueda”, este permite realizar comparaciones entre el componente de origen con la dimensión que recibirá los datos (dim\_empleado).
* Dimensión de variación lenta: Componente que permite obtener los datos desde la búsqueda realizada anteriormente, realiza las comparaciones según las configuraciones de esta transformación y deriva posteriormente a los componentes de inserción o actualización si fuese el caso.
* Grabar nuevo empleado dimensional: Componente de tipo “Destino OLEDB”, el cual está configurado para recibir los nuevos registros y almacenarlos en la dimensión empleado de la data warehouse.
* Actualizar empleado dimensional: Componente de tipo “Destino OLEDB”, posee las configuraciones necesarias para recibir los datos a modificar enviados por la “dimensión de variación lenta”, además de actualizar la información de los registros de la dimensión empleado, ajusta la fecha de modificación de cada registro afectado por la actualización de sus datos en el data warehouse.

**Evidencias configuración del ETL**

Cabe señalar que al momento de revisar el paquete SISS generado, primero se deben actualizar las credenciales de acceso a el motor de base de datos, los elementos a configurar en las conexiones son el nombre del servidor (localhost), puerto (1433), usuario (sa) y contraseña (su contraseña).

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 11: Conexión a base de datos relacional.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 12: Conexión a la data warehouse.

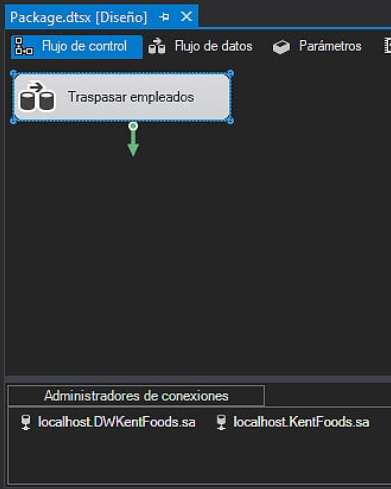


Ilustración 13: Flujo de control del paquete SISS.

Imagen que contiene captura de pantalla, monitor, negro

Descripción generada automáticamente

Ilustración 14: Estructura del flujo de control "Traspasar empleado"

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 15: Configuración del componente "Origen OLEDB" Cargar empleados en el modelo relacional.

La consulta utilizada es la siguiente:

SELECT

    DISTINCT

    dbo.Empleados.EmpleadoID,

    dbo.Empleados.Apellido,

    dbo.Empleados.nombre,

    dbo.Empleados.Cargo,

    dbo.Empleados.Direccion,

    dbo.Empleados.Ciudad,

    dbo.Empleados.Pais,

    CONVERT(NVARCHAR(100),stuff((

    SELECT

        ','+dbo.territorios.territorio

    FROM

        dbo.TerritoriosEmpleados teem

        INNER JOIN dbo.territorios ON teem.territorioID=dbo.territorios.TerritorioId

    WHERE teem.EmpleadoID = te.EmpleadoID

    FOR xml path ('')

    ),1,1,'')) AS territorio,

    CONVERT(NVARCHAR(50),stuff((

    SELECT

        DISTINCT

        ','+dbo.regiones.region

    FROM

        dbo.TerritoriosEmpleados teem

        INNER JOIN dbo.territorios ON teem.territorioID=dbo.territorios.TerritorioId

        INNER JOIN dbo.regiones ON dbo.territorios.RegionID=dbo.Regiones.RegionID

    WHERE teem.EmpleadoID = te.EmpleadoID

    FOR xml path ('')

    ),1,1,'')) AS region

FROM

    dbo.Empleados

    INNER JOIN dbo.TerritoriosEmpleados te ON dbo.Empleados.EmpleadoID = te.EmpleadoID

Esta consulta obtiene los datos desde las tablas empleados y territorios\_empleados, luego para obtener el o los territorios, fue necesario generar una subconsulta que obtuviese los territorios asociados al empleado y los consolide en un solo campo “territorio” para así, poder obtener tan solo una fila de empleados y evitar la repetición de identificadores del empleado, de la misma forma se extrae las regiones asociadas a los territorios.

Esto fue realizado de este modo, debido a que, al generar la ETL, el proceso se cae al tener dos registros con el mismo identificador, pero con diferentes territorios.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 16: Configuración general del componente "Búsqueda" - Buscar empleado dimensional.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 17: Configuración de la conexión del componente "Busqueda" - Buscar empleados dimensional.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 18: Configuración de las columnas del componente "Búsqueda" - Buscar empleado dimensional.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 19: Configuración de tablas y claves del componente “Dimension de variación lenta".

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 20: Configuración de columnas del componente "Dimensión de variación lenta".

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 21: Configuración de atributos fijos y variables del componente de "Dimensión de variación lenta".

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 22: Configuración de miembros deducidos del componente de "Dimensión de variación lenta".

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 23: Resultados del componente "Dimensión de variación lenta".

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 24: Configuración de conexión del componente "Destino OLEDB" - Grabar nuevo empleado dimensional.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 25: Asignaciones de columnas del componente "Destino OLEDB" - Grabar nuevo empleado dimensional.

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 26: Configuración del componente "Destino OLEDB" - Actualizar empleado dimensional.

La consulta generada para actualizar los registros es la siguiente:

UPDATE [dbo].[dim\_empleado]

SET

    [Apellido] = ?,

    [Cargo] = ?,

    [Ciudad] = ?,

    [Direccion] = ?,

    [Nombre] = ?,

    [Pais] = ?,

    [Region] = ?,

    [Territorio] = ?,

    UpdateDate=getdate()

WHERE [EmpleadoID] = ?

Imagen que contiene captura de pantalla

Descripción generada automáticamente

Ilustración 27: Asignaciones de columnas del componente "Destino OLEDB" - Actualizar empleado dimensional.