

PEC 1

Presentación.

La primera práctica de evaluación continua consta de dos partes diferenciadas, una **teórica (40%)** y otra **práctica (60%)**.

La primera parte consiste en una serie de preguntas que tienen el objetivo de consolidar los conocimientos teóricos de los módulos **1. Introducción a las bases de datos analíticas** y **2. La construcción de la factoría de información corporativa** de la asignatura.

La segunda parte consiste en la toma de contacto con la máquina virtual proporcionada por la UOC (VDI) y el software necesario para realizar las prácticas de la asignatura, capacitando al estudiante para iniciar el resto de las actividades prácticas del curso.

Objetivos y competencias

- Diferenciar claramente entre el almacén de datos y la base de datos operacional, identificando las principales diferencias.
- Conocer el contexto del almacén de datos, es decir, la Factoría de Información Corporativa, sus componentes y la función de cada uno de ellos.
- Validar el entorno de prácticas.

Contenido.

Parte teórica.

Se compone de 5 preguntas teóricas que tienen por objetivo comprobar la correcta comprensión del módulo 1 y 2 de la asignatura. Cada pregunta tiene una única opción válida entre las posibles.

Parte práctica.

Consta de cuatro ejercicios:

- Ejercicio 1. Configuración del entorno VDI
- Ejercicio 2. Validación de la BBDD.
- Ejercicio 3. Validación de la aplicación de ETL.
- Ejercicio 4. Microsoft SQL *Analysis Services*

El objetivo es conseguir que el estudiante entre en contacto con el entorno de prácticas (Entorno VDI) y pueda comprobar la correcta configuración del mismo, antes de iniciar el caso práctico del curso.

Cada estudiante accederá a través de su propio escritorio personal al escritorio virtual, previamente configurado, y validará la conectividad con la base de datos así como la aplicación para ejecución de procesos ETL (Extracción, Transformación y Carga). Además, también realizará un ejercicio con Visual Studio y Microsoft SQL *Analysis Services* (SSAS).

Información relevante de credenciales:

- a) Las credenciales para acceder a las máquinas VDI son las mismas que utilizáis para entrar en el campus de la UOC.
- b) Las credenciales de acceso al SQL server para cada estudiante son:

Nombre del servidor: Se facilitará vía tablero de la asignatura

Tipo de autenticación: SQL Server Authentication

Usuario: STUDENT_loginuoc

Contraseña: \$d3f4ult\$

En el primer login os solicitará cambiar el password.

Criterios de evaluación.

- Parte teórica (40%)

Cada pregunta correctamente contestada, sumará un 8% a la nota final de la PEC, y las respuestas erróneas descontarán un 2% de la parte teórica. Por tanto, las respuestas incorrectas no afectarán negativamente al 60% de la parte práctica.

- Parte práctica (60%)

Se valorará de acuerdo con las evidencias de cumplimiento de los pasos indicados y aportados por el estudiante en el documento de solución, mediante **capturas de pantalla y las explicaciones oportunas** de las actividades realizadas, de acuerdo con los objetivos establecidos. Las capturas sin

explicación no se valorarán como totalmente correctas.

La nota final estará formada por:

Pregunta 1 (8%) + Pregunta 2 (8%) + Pregunta 3 (8%) + Pregunta 4 (8%) +
Pregunta 5 (8%) + Ejercicio 1 (5%) + Ejercicio 2 (20%) + Ejercicio 3 (25%) +
Ejercicio 4 (10%).

Formato y fecha de entrega

La entrega se realizará enviando un único mensaje al buzón de entrega de actividades del aula. Dicho mensaje llevará adjunto un único documento en formato pdf con la solución de la PEC. En el documento se debe indicar obligatoriamente el nombre completo del estudiante y los estudios que está cursando. El nombre del archivo debe ser la composición del nombre de usuario y “DW_PEC1” (por ejemplo: si el nombre de usuario es “acabero” , el nombre del archivo debe ser “acabero_DW_PEC1.pdf”). Es responsabilidad única del estudiante asegurarse que entrega el documento que pretende en el lugar que la Universidad habilita con este objetivo, que es el registro EC (Evaluación Continua).

La fecha máxima de entrega es el 24/03/2021 a las 23:59 h

Formato y fecha de feedback o retorno pedagógico

El feedback o retorno pedagógico se realizará de manera individual a través del registro EC (Evaluación Continua). Además, se publicará en el foro asociado a la PEC una valoración general de los resultados donde se indicarán aquellos errores comunes u observaciones que se quieran dar con carácter general a los estudiantes.

Parte teórica (40%).

Nota : En azul la opción marcada como correcta.

Pregunta 1 (8%):

Indicad la opción que describe mejor cuales son las principales características de un Data Warehouse.

- Las principales características de un Data Warehouse son: que esté orientado al tema, que los datos estén integrados, y que la información sea dinámica y volátil
- Las principales características de un Data Warehouse son: que esté orientado al tema, que los datos estén integrados, y que la información sea histórica y no volátil
- Las principales características de un Data Warehouse son: que esté orientado a la funcionalidad, que los datos estén integrados, y que la información sea histórica y no volátil
- Las principales características de un Data Warehouse son: que esté orientado al tema, que los datos sean homogéneos, y que la información sea histórica y volátil

Pregunta 2 (8%):

Indicad la opción que describe mejor los datos Maestros.

- Los Maestros son entidades con información generalmente crítica, que intervienen en la mayoría de consultas para poder realizar análisis mucho más precisos. Algunos ejemplos de datos maestros pueden ser clientes, productos o proveedores.
- Los Maestros son entidades con información generalmente crítica, que intervienen en la mayoría de consultas para poder realizar análisis mucho más precisos. Algunos ejemplos de datos maestros pueden ser el detalle de compras, el inventario de stock, o el histórico de campañas de marketing realizado a cada uno de los clientes de la empresa.
- Los maestros son un elemento específico de la FIC que representan características de otros datos, como pueden ser las estructuras de almacenamiento de los mismos, y por lo tanto pueden considerarse como un componente por sí mismos.
- Los maestros son las estructuras menos relevantes a nivel analítico. Contienen información de clientes y productos, pero al no tener información de compras o ventas no se utilizan para análisis.

Pregunta 3 (8%):

Indicad la opción que describe mejor cuáles son las diferencias existentes entre una base de datos operacional y un *Data Warehouse* desde un punto de vista de la estructura de datos.

- Los Data Warehouses tienen una estructura relacional y bastante estática para reforzar su estabilidad, mientras que las bases de datos operacionales destacan por su estructura multidimensional y dinámica, que les permite adaptarse rápidamente a las necesidades del negocio
- Las bases de datos operacionales tienen una estructura relacional y bastante estática para reforzar su estabilidad, mientras que los Data Warehouse destacan por su estructura multidimensional y dinámica, que les permite adaptarse rápidamente a las necesidades del negocio
- Las bases de datos operacionales tienen una estructura relacional y bastante estática para poder dar respuesta a las necesidades analíticas del negocio, mientras que los Data Warehouse destacan por su estructura multidimensional y dinámica, que les permite obtener una consistencia mayor que en las bases

de datos operacionales

- Las bases de datos operacionales deben tener presente el componente temporal, mientras que los Data Warehouse muestran una única versión actualizada de los datos

Pregunta 4 (8%):

Indicad la opción correcta:

- Una dimensión conformada es una entidad del modelo de datos que unifica múltiples entidades, como pueden ser productos y clientes en una única tabla
- Una dimensión conformada es una entidad del modelo de datos de origen dudoso, y que sólo debe utilizarse en casos de extrema necesidad sabiendo que todas las partes implicadas en el análisis se “conforman” con los errores que pueden aparecer a partir de la misma
- Una dimensión conformada es una entidad que se conforma a partir de consultas sobre otras entidades, como por ejemplo al obtener el listado de clientes a partir de una consulta sobre la entidad de transacciones.
- Una dimensión conformada es una entidad del modelo de datos que puede ser compartida entre varios almacenes de datos, como puede ser la entidad calendario o cliente.

Pregunta 5 (8%):

Indicad la opción correcta:

- Debido a los problemas que aparecen a la hora de construir el almacén de datos corporativo, se recomienda hacerlo siempre después de construir los almacenes de datos departamentales, de modo que tengamos visibilidad completa del tamaño de la corporación.
- Debido a los problemas que aparecen a la hora de construir el almacén de datos corporativo, se recomienda hacerlo siempre de forma simultánea con los almacenes de datos departamentales, planificando la construcción de la FIC desde un inicio.
- Los almacenes de datos departamentales no pueden existir si previamente no se ha creado el almacén de datos corporativo
- Los almacenes de datos corporativos y departamentales son excluyentes entre ellos, por lo que las empresas deberán seleccionar cuál de los dos enfoques se adapta mejor a sus necesidades.

Parte práctica (60%).

Ejercicio 1 (5%): Configuración del entorno VDI

Objetivo:

Configurar la conexión al escritorio VDI que se utilizará durante todo el curso, tanto para realizar esta PEC (PEC1) como las siguientes actividades evaluables del curso.

Desarrollo del ejercicio:

Utilizando los documentos: **Entorno_Virtual_cast.pdf** (como guía de configuración y conexión) e **Importar base de datos, como crear vistas y cubos.pdf** (con indicaciones sobre cómo proceder para crear cubos), el estudiante debe de llevar a cabo todo el proceso hasta obtener su escritorio VDI. Se recomienda leer atentamente dichos documentos y seguir sus indicaciones.

El estudiante debe entregar, como solución a este ejercicio, una **guía paso a paso** de todo el proceso de configuración de la máquina accesible desde el escritorio de la VDI. La guía debe de incluir todas las explicaciones y capturas de pantalla que sean necesarias.

También se adjuntará una captura de pantalla del escritorio VDI y la información técnica disponible bajo el logo de la UOC (*HostName*, *UserName*, etc.), que encontraréis como fondo de escritorio.

SOLUCION: Estuve estudiando los contenidos antes de empezar con la PEC , motivo por lo cual instale CITRIX sin hacer las capturas respectiva , le he comentado la situación a el profesor David y el me autorizo como alternativa poner solo estas pantallas .

Una vez instalado el escritorio virtual citrix, cuando accedemos al link nos sale la lista de servidores.



Le damos clic al botón connect y tendremos acceso al servidos.



Ejercicio 2 (20%): Validación de la BBDD.

Objetivo:

Toma de contacto con el Sistema Gestor de Bases de Datos facilitado con la máquina VDI (SQL Server 2017), configuración de una conexión con el servidor y creación de una base de datos simple con una tabla.

Desarrollo del ejercicio:

1) Conectar al servidor de base de datos SQL Server

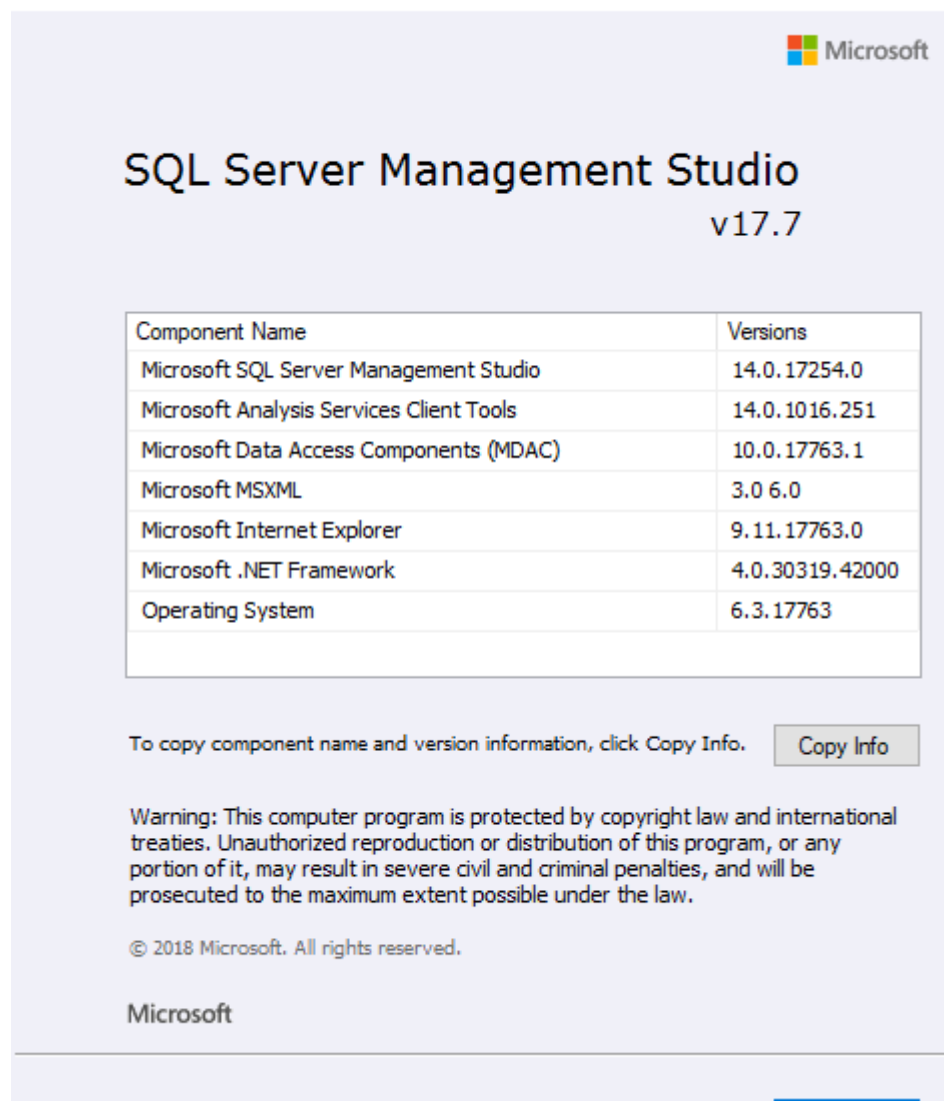
Utilizad la consola “Microsoft SQL Server Management Studio 2017 (SSMS)” instalada en la VDI para abrir una conexión con el servidor de base de datos y adjuntad las siguientes capturas de pantalla:

- Versión de SSMS

SOLUCION:

Para ver la versión le damos click a la opción de menú Help->About ...

About Microsoft SQL Server Management Studio

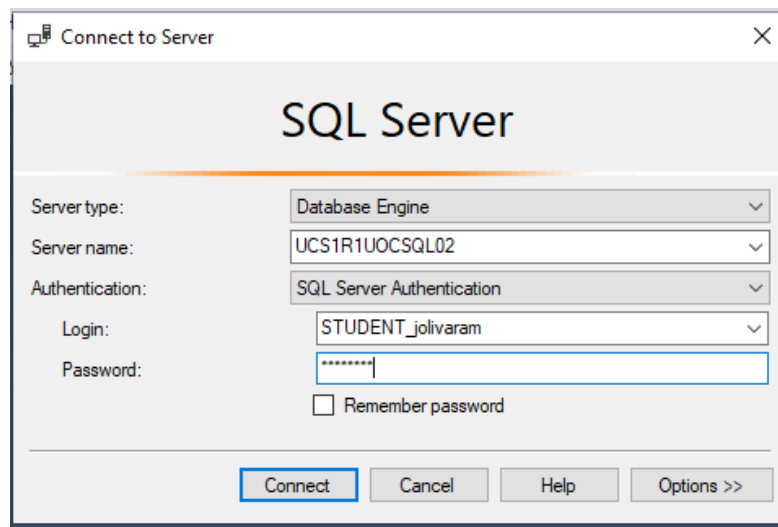


b) Pantalla de conexión inicial de la consola SQL Server Management Studio (SSMS) donde se pueda observar la configuración utilizada.

- o Server Type*
- o Server name*
- o Authentication*
- o Login*

SOLUCION:

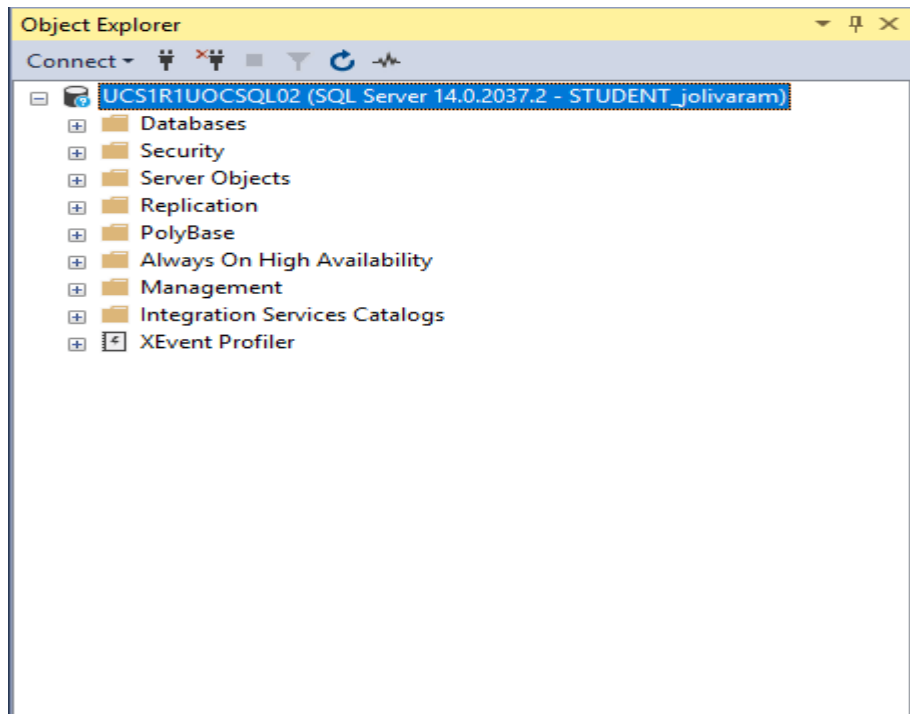
Para ver la conexión le damos click a la opción de menú File ->Connect Object Explorer



c) Explorador de objetos del SSMS donde se puede ver el usuario de conexión.

SOLUCION:

Para ver el usuario tenemos que expandir el Object explorer que tenemos en la parte superior izquierda y fijarnos en el ultimo parámetros este es el Usuario de conexión

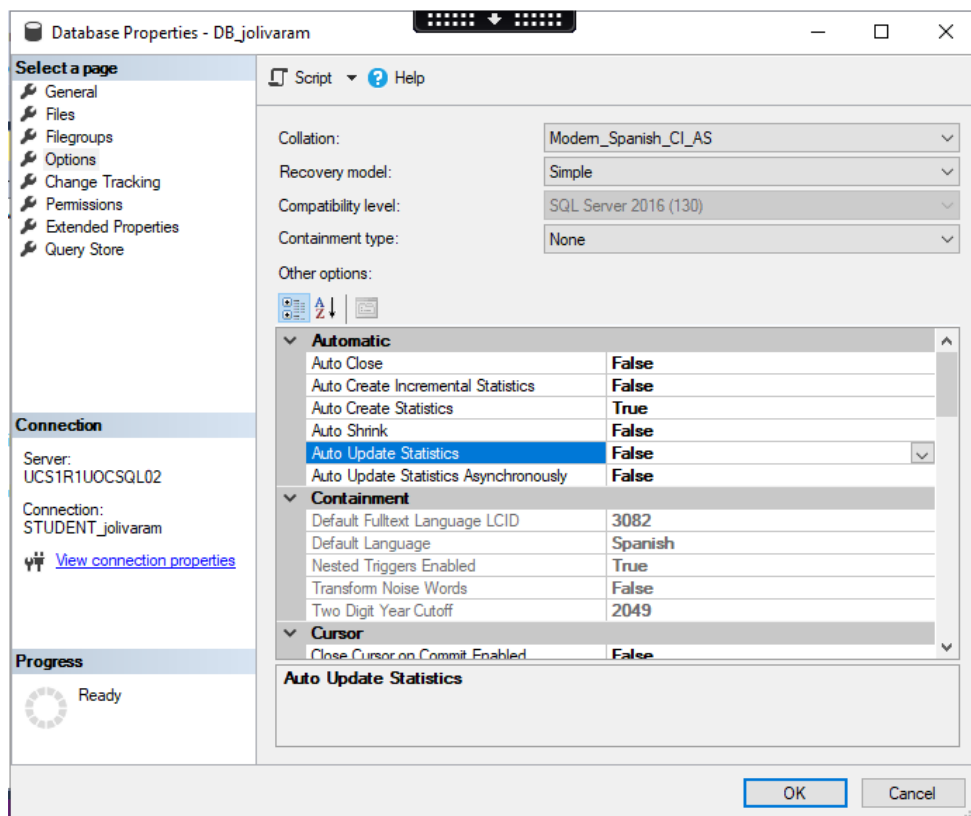


2) Configurar la base de datos DB_loginuoc.

Se debe configurar el parámetro *Recovery Model* de la BBDD como “simple” y desactivar la actualización automática de estadísticas. Adjuntad las capturas de pantalla que demuestren la correcta configuración.

SOLUCION:

Para hacer este cambio tenemos que entrar a las propiedades de la DB y entrar en Options como se ve en la imagen :



3) Crear una tabla con el nombre T_PEC1

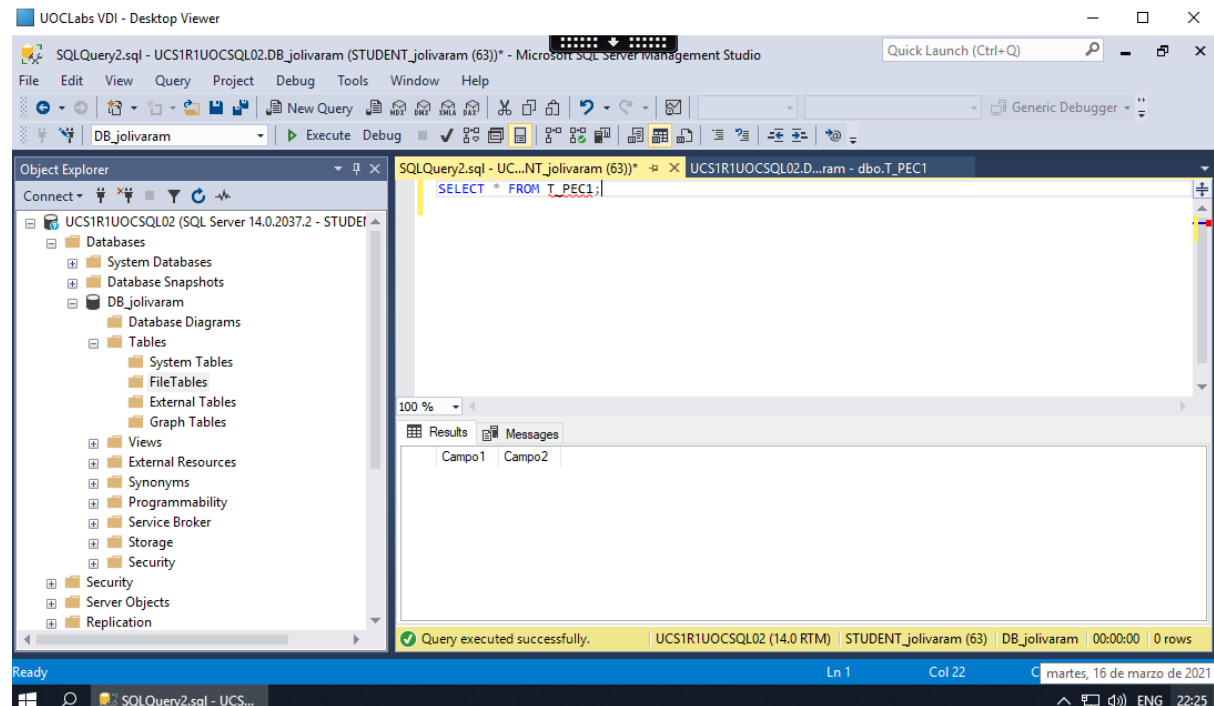
La tabla tendrá dos campos: campo1 y campo2 de la siguiente manera:

- o **Campo1** será de tipo varchar(50) y permitirá valores nulos
- o **Campo2** será de tipo int y no permitirá valores nulos.

Adjuntad las capturas de pantalla que demuestren la correcta creación de la tabla con sus campos.

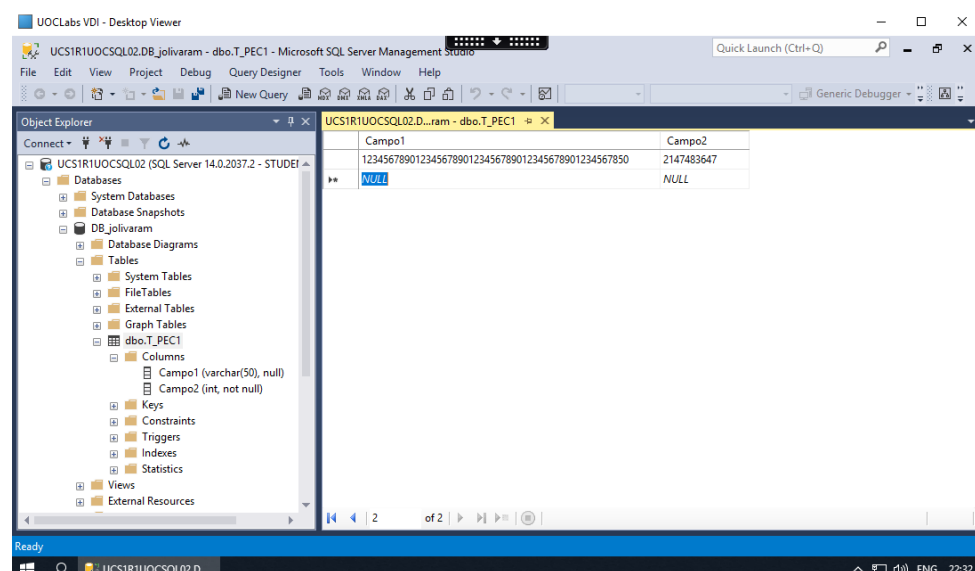
SOLUCION:

Consulta de la tabla via SQL



SOLUCION:

Consulta via Object Explorer , campos con los máximos valores para sus tipos : varchar(50) e integer.



Ejercicio 3 (25%): Validación de la aplicación de ETL.

Objetivo:

Toma de contacto con la aplicación de ETL suministrada en la máquina VDI (*Pentaho Data Integration*) y capacitar al estudiante para iniciar las actividades del curso

Desarrollo del ejercicio:

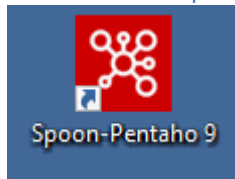
Para familiarizarse con la herramienta *Spoon* que permite la transformación de datos, se pide cargar los datos de un archivo CSV (origen de datos) en una nueva tabla de la base de datos (destino de datos).

Describid paso a paso todas las acciones realizadas aportando las capturas de pantalla y scripts que sean necesarias. Los pasos para realizar son los siguientes:

1) Abrir el componente Spoon.

SOLUCION:

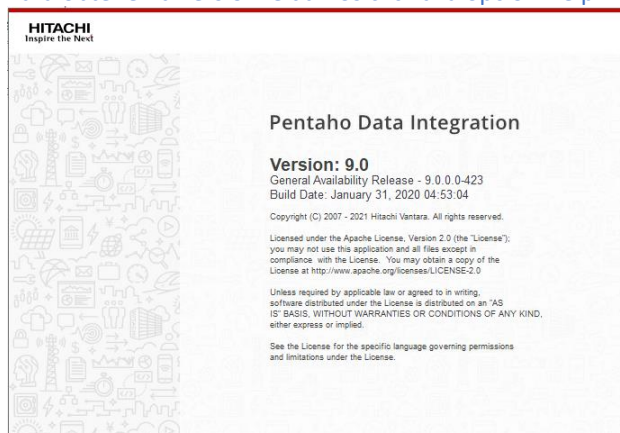
Para abrir el componente Spoon le damos click al siguiente icono del escritorio :



2) Capturar versión de PDI.

SOLUCION:

Para obtener la versión le damos click a la opción Help->About



3) Crear un repositorio de tipo file con el nombre PEC1.

SOLUCION:

Le damos click a la opción connect y abrimos esta pantalla y luego pulsamos en la opción "Other Repositories"



En la opción “Other Repositories” encontraremos la opción que nos pide el ejercicio “File Repository” la seleccionamos y damos click a “Get Started”

Other Repositories

Not supported or recommended for production use

Database Repository

Unsupported ETL storage that uses a database

File Repository

Unsupported ETL storage that uses your local file system

Pentaho Repository

Help

Get Started

Close

Se nos abrirá la siguiente pantalla

Ingresamos los datos en siguiente pantalla

Repository Details

Display Name

PEC1

X

Location

F:\PEC1\

Browse

☐ Don't modify items in this location

☐ Show hidden folders

Description

File repository


☐ Launch connection on startup

Help

Back

Finish

Le damos a “Connect Now “



Congratulations!

Your connection has been created and is ready to use.

What would you like to do?

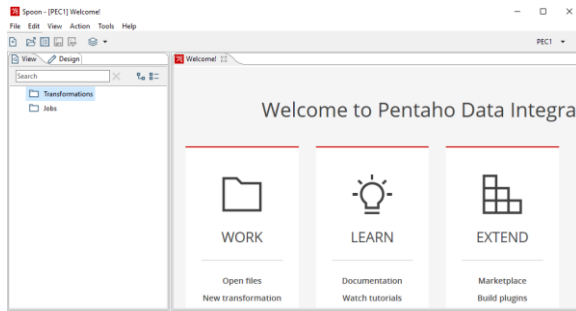
Connect Now

Manage Connections

Help

Finish

Vemos la conexión a PEC1 en el lado superior derecho



4) Crear una transformación y un Job básicos:

Anexo al enunciado de la PEC1, encontraréis un archivo **datos.csv**.

Se debe copiar el archivo de origen de datos a una de las carpetas persistentes de la máquina VDI.

A. Análisis preliminar del origen de datos

Es conveniente analizar los orígenes de datos antes de cargarlos, dado que nuestro origen es un archivo csv de muy pocos registros, podemos hacer un análisis preliminar directamente:

- Identificar los campos a cargar

SOLUCION:

En este punto hemos analizado el fichero en Excel ordenando por diferentes criterios y hemos visto que tenemos 8 campo con ciertas características

A	B	C	D	E	F	G	H
PracNo	PracticeName	Address1	Address2	Address3	Postcode	LCG	Registered Patients
657	NULL	THREE SPIRES SU	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	A
601	NULL	MOURNESIDE ME	STRABANE	NULL	BT82 9AF	Western	10868
662	NULL	DRUMRAGH FAM	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	10746
615	NULL	CLARENDON MEE	35 NORTHLAND A	LONDONDERRY	BT48 7GW	Western	10726
157	NULL	UNIVERSITY HEAL	ELMWOOD MAN	7 UNIVERSITY TEF	BT7 1NP	Belfast	10018
596	NULL	CITYVIEW MEDIC	SPENCER ROAD	LONDONDERRY	BT47 6AH	Western	9680
516	NULL	THE FRIARY SURG	DOBBIN LANE	ARMAGH	BT61 7QG	Southern	9089
616	NULL	CASTLEDERG GRC	13A LOWER STRA	CASTLEDERG	BT81 7AZ	Western	8886
493	NULL	NORTHLAND SUR	79 CUNNINGHAM	DUNGANNON	BT71 6BX	Southern	8667

- Determinar el tipo de dato de cada campo

SOLUCION:

El primero y el ultimo campo son numéricos (Integer) el resto son textos(Varchar)

A	B	C	D	E	F	G	H
PracNo	PracticeName	Address1	Address2	Address3	Postcode	LCG	Registered Patients
657	NULL	THREE SPIRES SU	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	A
601	NULL	MOURNESIDE ME	STRABANE	NULL	BT82 9AF	Western	10868
662	NULL	DRUMRAGH FAM	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	10746
615	NULL	CLARENDON MEE	35 NORTHLAND A	LONDONDERRY	BT48 7GW	Western	10726
157	NULL	UNIVERSITY HEAL	ELMWOOD MAN	7 UNIVERSITY TEF	BT7 1NP	Belfast	10018
596	NULL	CITYVIEW MEDIC	SPENCER ROAD	LONDONDERRY	BT47 6AH	Western	9680
516	NULL	THE FRIARY SURG	DOBBIN LANE	ARMAGH	BT61 7QG	Southern	9089
616	NULL	CASTLEDERG GRC	13A LOWER STRA	CASTLEDERG	BT81 7AZ	Western	8886
493	NULL	NORTHLAND SUR	79 CUNNINGHAM	DUNGANNON	BT71 6BX	Southern	8667

- Identificar campos que deben admitir valores nulos

SOLUCION:

El segundo campo PracticeName tiene que admitir nulos

A	B	C	D	E	F	G	H
PracNo	PracticeName	Address1	Address2	Address3	Postcode	LCG	Registered Patients
657	NULL	THREE SPIRES SUITE 100	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	A
601	NULL	MOURNESIDE MEDICAL CENTRE	STRABANE	NULL	BT82 9AF	Western	10868
662	NULL	DRUMRAGH FAMILY PRACTICE	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	10746
615	NULL	CLARENDON MEDICAL CENTRE	35 NORTHLAND AVE	LONDONDERRY	BT48 7GW	Western	10726
157	NULL	UNIVERSITY HEALTH CENTRE	ELMWOOD MANOR	7 UNIVERSITY TERN	BT7 1NP	Belfast	10018
596	NULL	CITYVIEW MEDICAL CENTRE	SPENCER ROAD	LONDONDERRY	BT47 6AH	Western	9680
516	NULL	THE FRIARY SURGERY	DOBBIN LANE	ARMAGH	BT61 7QG	Southern	9089
616	NULL	CASTLEDERG GRC	13A LOWER STRABANE	CASTLEDERG	BT81 7AZ	Western	8886
493	NULL	NORTHLAND SURGERY	79 CUNNINGHAM ROAD	DUNGANNON	BT71 6BX	Southern	8667

- Otras observaciones

SOLUCION:

El último campo "Register Patients" tendría que ser número pero vemos que tiene un carácter, imagino que esto lo resolveremos mas adelante con algún tipo de transformacion

A	B	C	D	E	F	G	H
PracNo	PracticeName	Address1	Address2	Address3	Postcode	LCG	Registered Patients
657	NULL	THREE SPIRES SUITE 100	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	A
601	NULL	MOURNESIDE MEDICAL CENTRE	STRABANE	NULL	BT82 9AF	Western	10868
662	NULL	DRUMRAGH FAMILY PRACTICE	MOUNTJOY ROAD	OMAGH	BT79 7BA	Western	10746
615	NULL	CLARENDON MEDICAL CENTRE	35 NORTHLAND AVE	LONDONDERRY	BT48 7GW	Western	10726
157	NULL	UNIVERSITY HEALTH CENTRE	ELMWOOD MANOR	7 UNIVERSITY TERN	BT7 1NP	Belfast	10018
596	NULL	CITYVIEW MEDICAL CENTRE	SPENCER ROAD	LONDONDERRY	BT47 6AH	Western	9680
516	NULL	THE FRIARY SURGERY	DOBBIN LANE	ARMAGH	BT61 7QG	Southern	9089
616	NULL	CASTLEDERG GRC	13A LOWER STRABANE	CASTLEDERG	BT81 7AZ	Western	8886
493	NULL	NORTHLAND SURGERY	79 CUNNINGHAM ROAD	DUNGANNON	BT71 6BX	Southern	8667

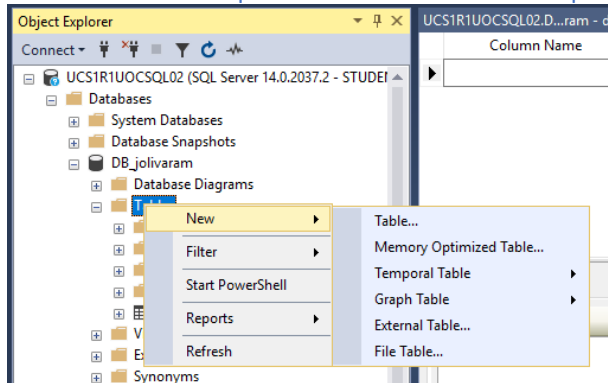
B. Crear una tabla en la BBDD con el nombre T_DATOS

Esta tabla permitirá cargar toda la información del archivo datos.csv en los campos correspondientes.

SOLUCION:

Para ello accederemos al SSMS, accedemos a nuestra DB y pulsamos la opción :

Table->New-> Table que nos habilitara en la zona superior derecha un espacio para la creación de tablas



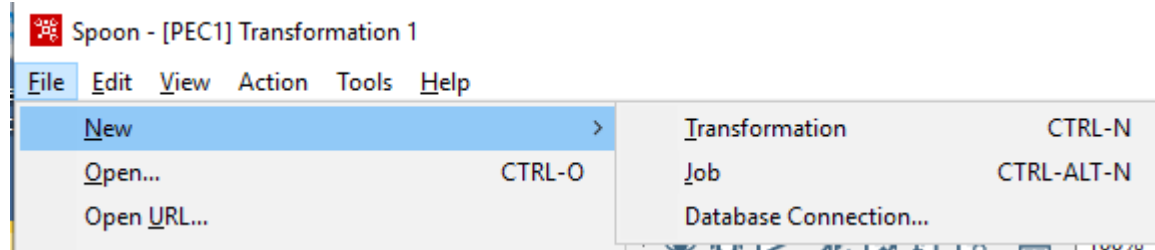
El primer campo esta claro que es la primary key motivo por lo cual evitamos los nulos y le asignamos el tipo int.

Column Name	Data Type	Allow Nulls
PracNo	int	<input type="checkbox"/>
PracticeName	varchar(50)	<input checked="" type="checkbox"/>
Address1	varchar(50)	<input checked="" type="checkbox"/>
Address2	varchar(50)	<input checked="" type="checkbox"/>
Address3	varchar(50)	<input checked="" type="checkbox"/>
PostCode	varchar(10)	<input checked="" type="checkbox"/>
LCG	varchar(15)	<input checked="" type="checkbox"/>
RegisteredPatients	int	<input checked="" type="checkbox"/>

C. Crear una transformación con *Spoon* que realice las siguientes tareas:

SOLUCION:

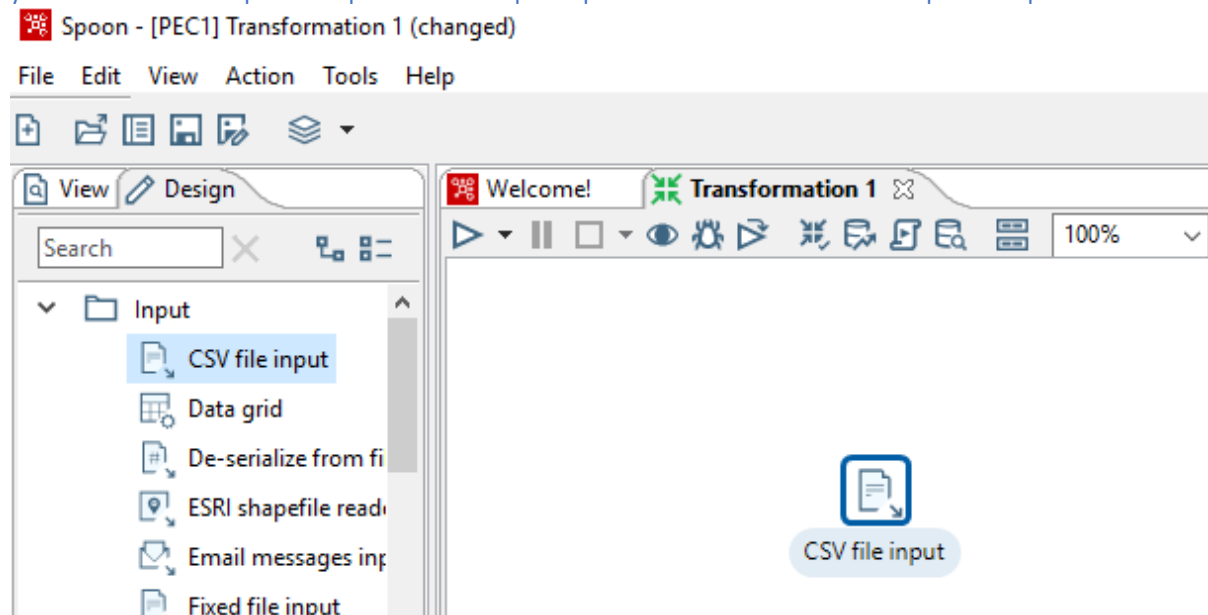
Le damos click a **File->New->Transformation** para crear la transformación :



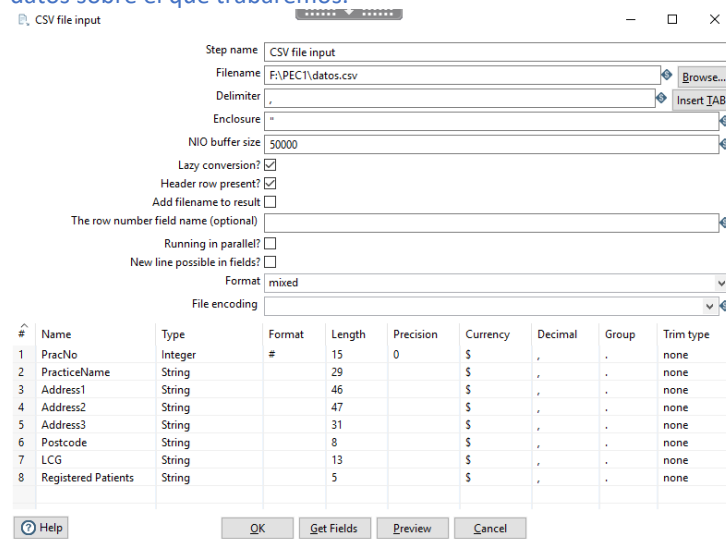
1. Extraer la información del archivo CSV

SOLUCION:

Cuando creamos una nueva transformación en la parte de diseño de la izquierda se abre una serie de opciones y le damos click a la opción "Input->CSV file input" que nos crea un icono CSV file input en la parte derecha




Sobre este icono , le damos click con el botón derecho y elegimos la opción edit , aquí es donde definiremos el origen de los datos , una vez definamos el origen le damos al botón "GetFields" para obtener un previo de los datos sobre el que trabajaremos.

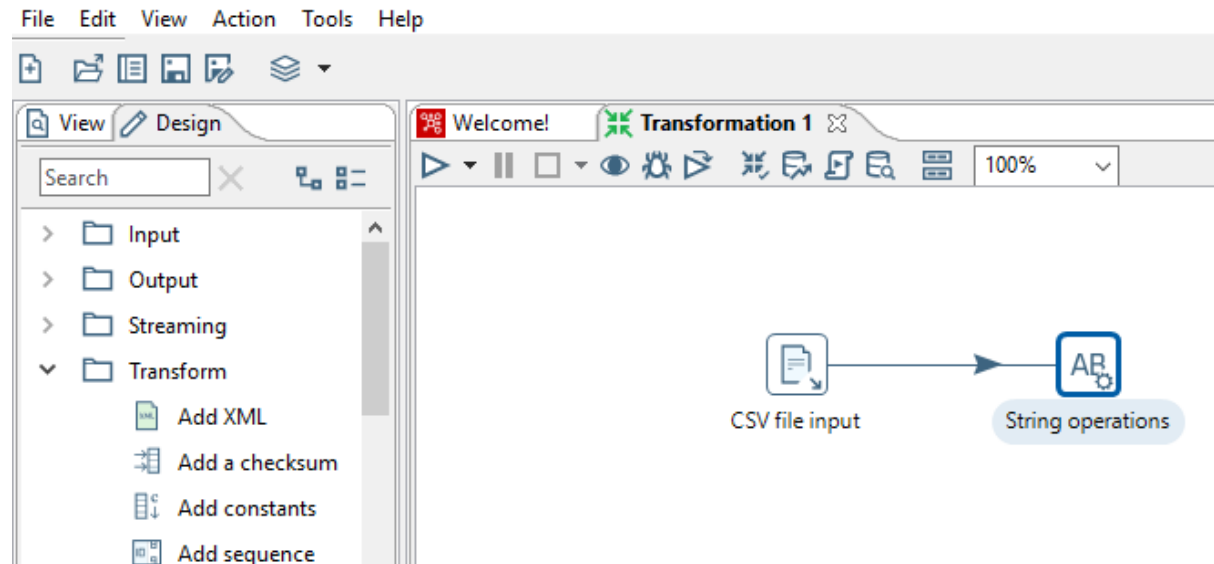


2. Transformar todos los datos de los campos de tipo texto a mayúsculas

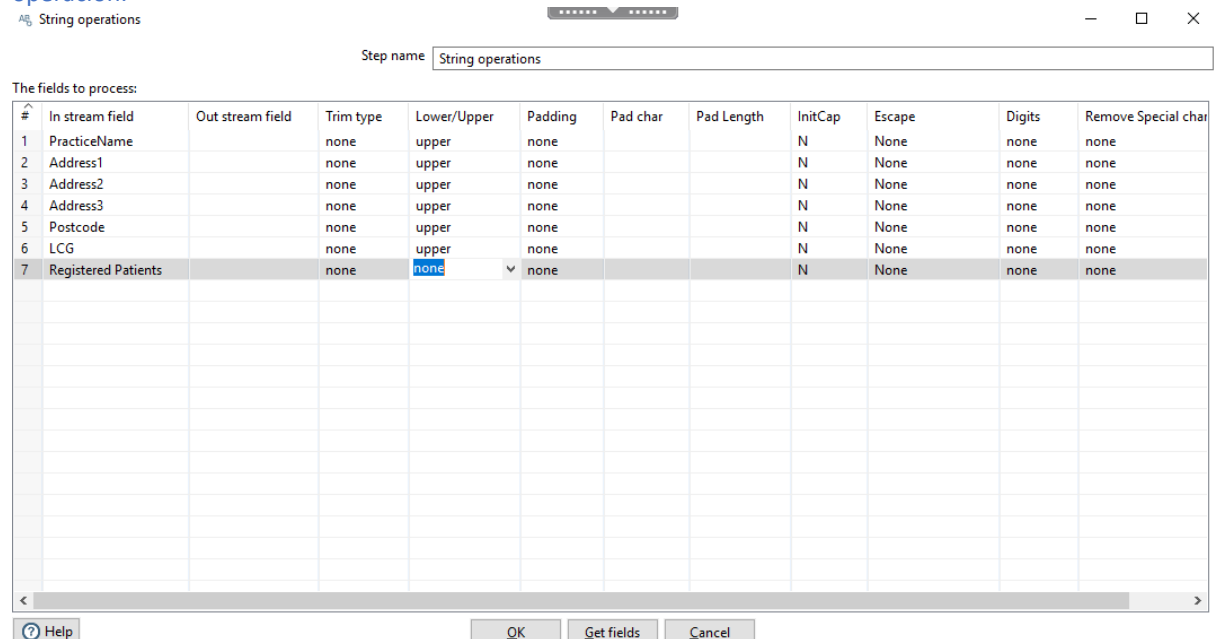
SOLUCION:

Para transformar a mayúsculas definimos otro paso dando click a “Transform->String operation”

 Spoon - [PEC1] Transformation 1 (changed)



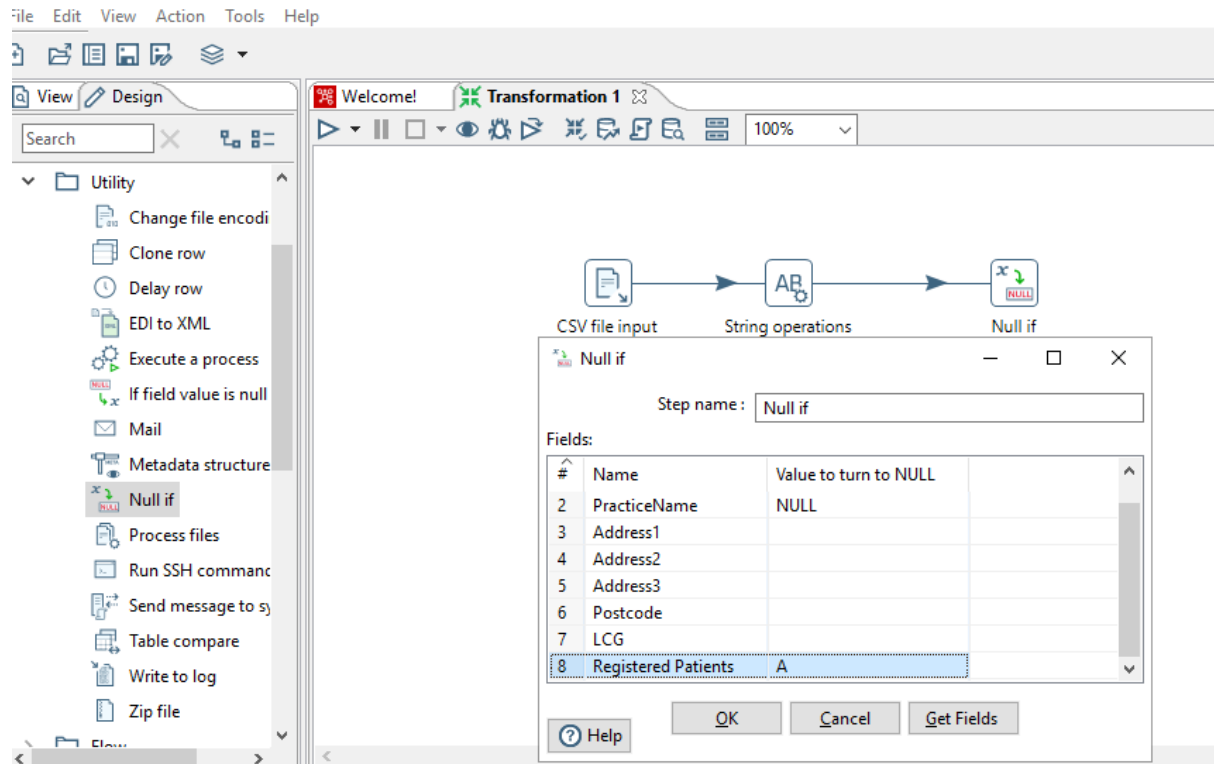
Editamos string operation y le damos click al botón “Get fields” y ponemos todas las columnas de texto en MAYUSCULAS cambiando los valores Lower/Upper a upper. Finalmente le damos click a OK para validar la operación.



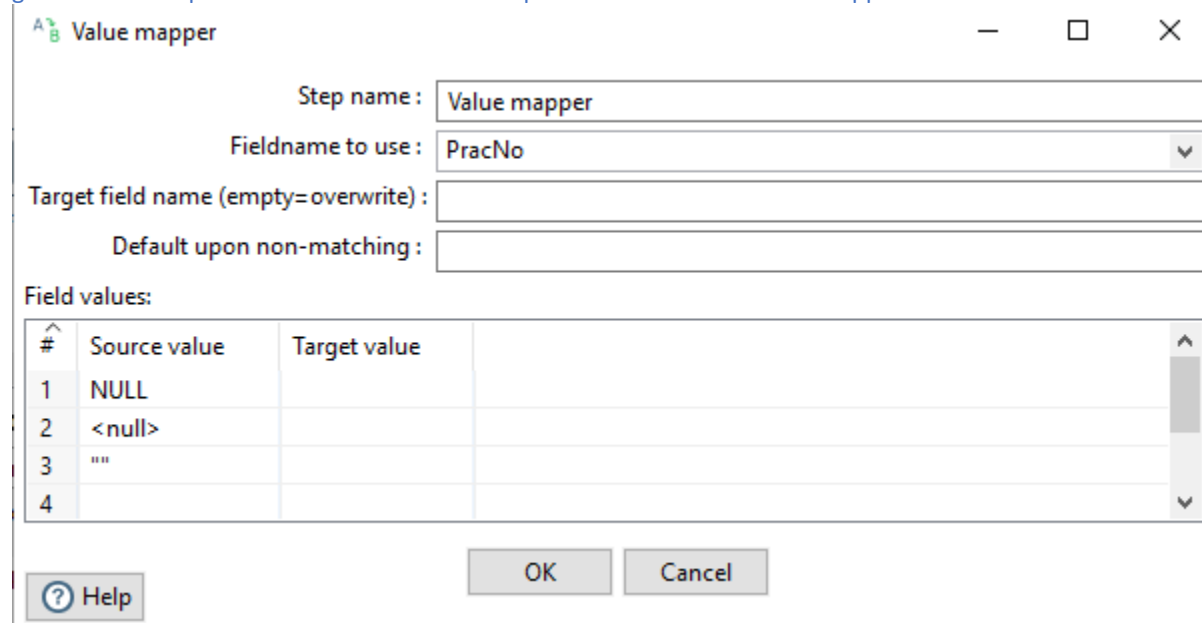
3. Corregir los errores detectados en el análisis

SOLUCION:

Le damos a la opción Utility -> Null if para convertir a null los valores NULL y A detectados en los campos "PracticeName" y "RegisterPatients"



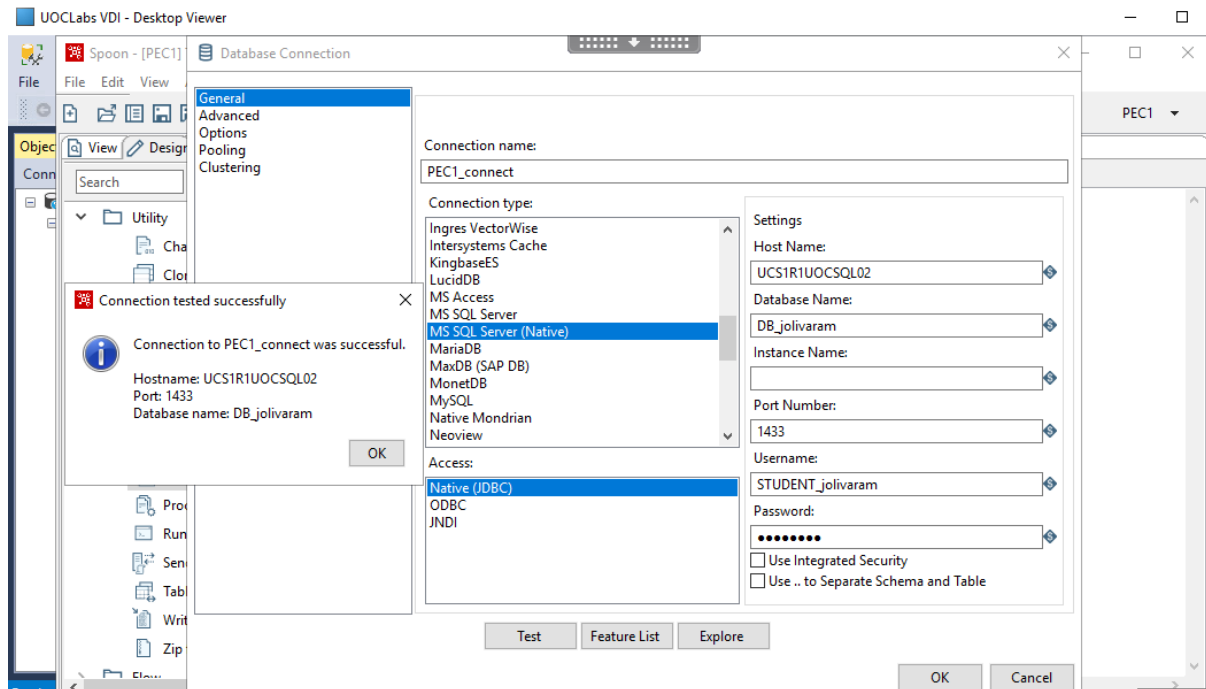
Evitamos que en nuestra posible clave primaria que no acepta nulos (PracNo) tengamos valores extraños y generemos excepciones en nuestra base de datos para ello utilizamos Value mapper :



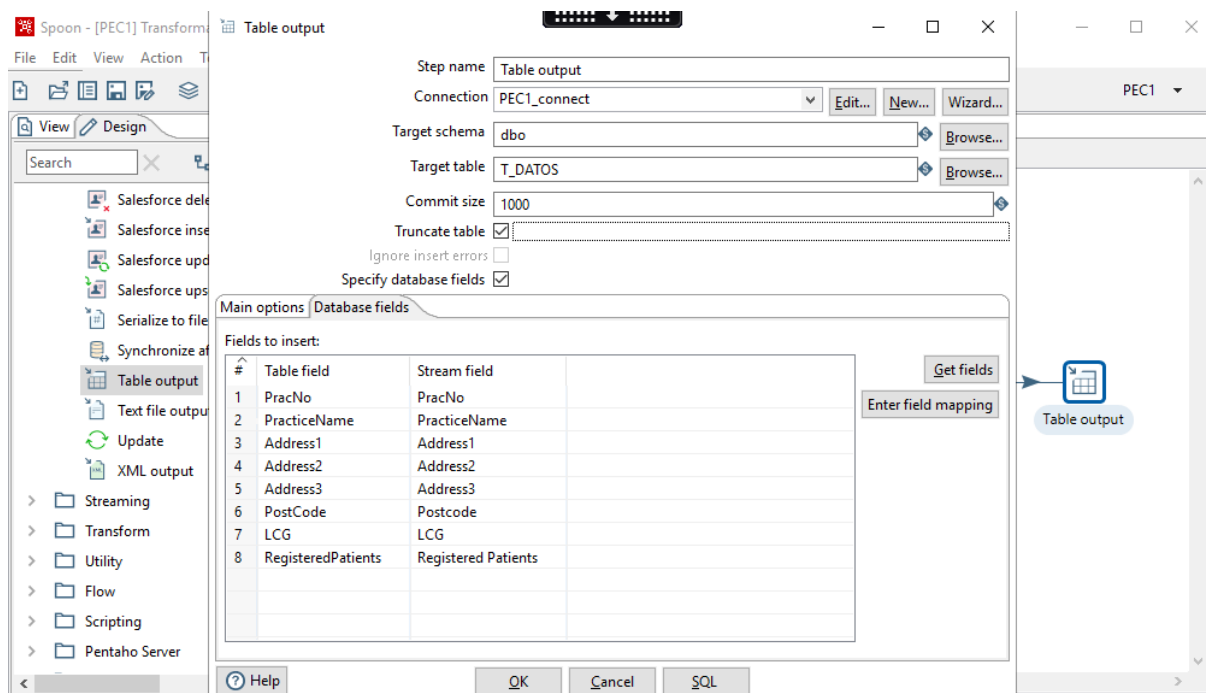
4. Cargar la información transformada en la tabla anteriormente creada.

SOLUCION:

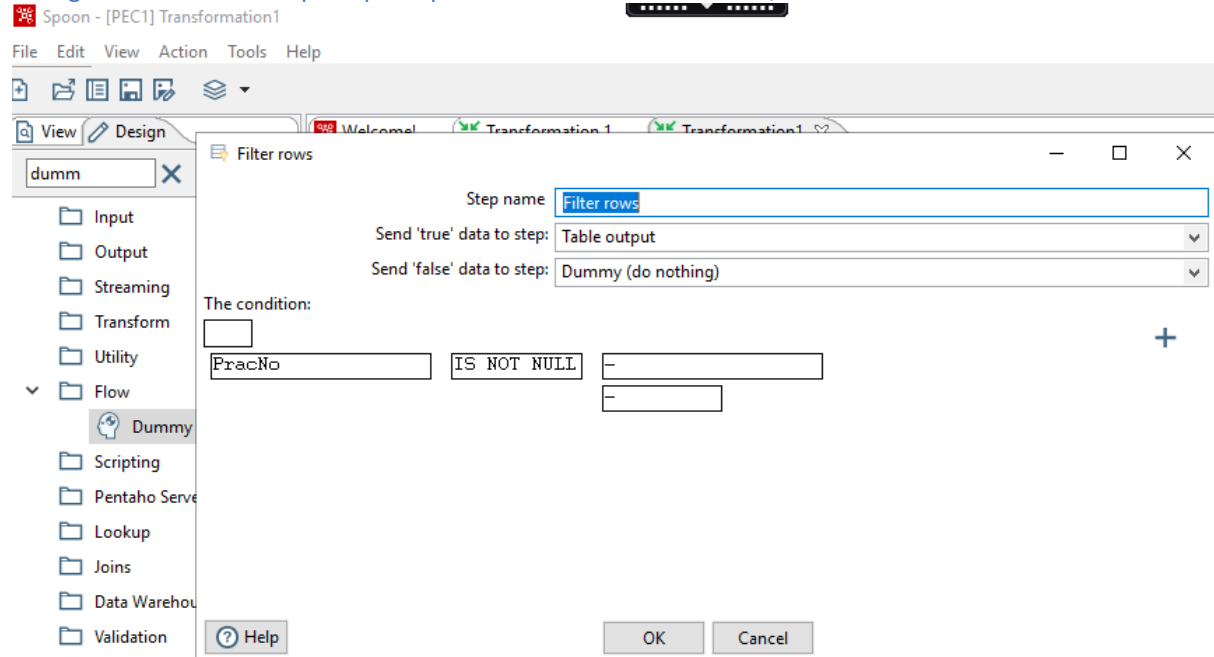
Le damos click File->New Connection y configuramos los parámetros pertinentes y testeamos la conexión , cuando tengamos la conexión testeada le damos a ok



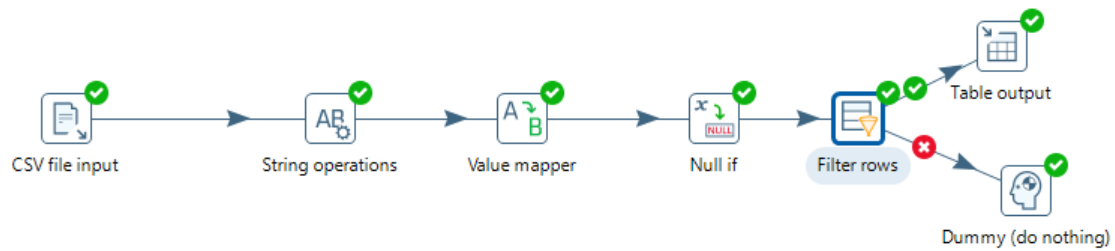
Creamos los pasos para las ultimas tres opciones Filter Row para filtrar los valores nulos , Output value donde enviaremos los valores correctos y Dummy donde enviaremos los valores nulos
Configuramos Output values:



Configuramos Filter Rows para que dependiendo del valor de PracNo , nos lo envíe a uno u otro fichero :



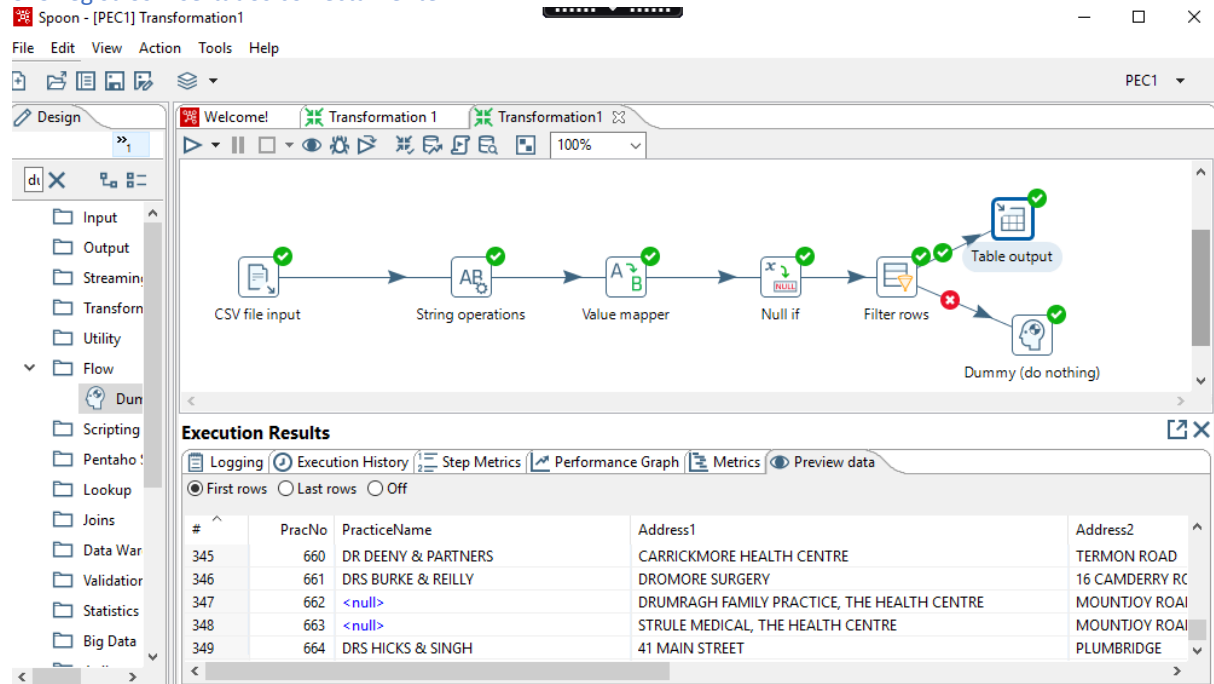
Y Obtenemos el siguiente flujo de trabajo :



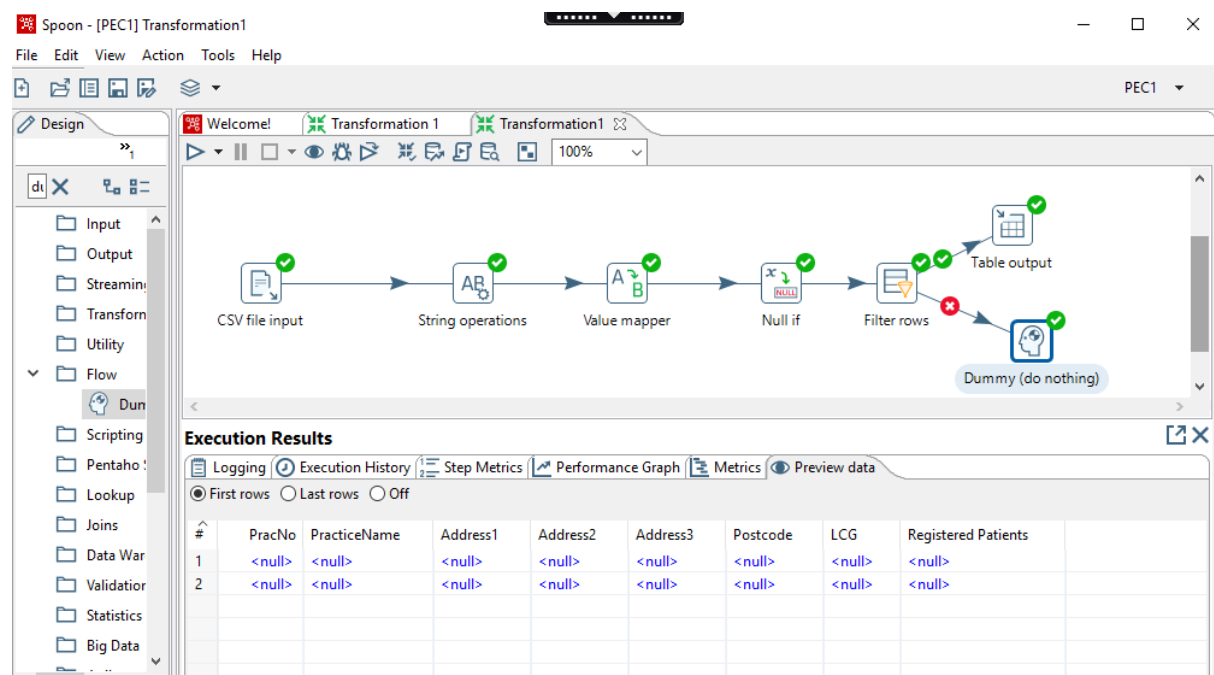
SOLUCION:

Para realizar la carga efectiva le damos al botón Run y tenemos como resultado:

349 registros insertados correctamente :



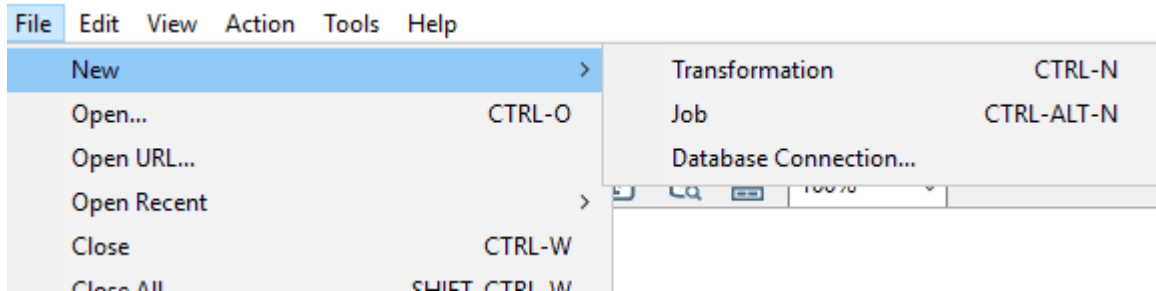
Dos registros nulos que los almacenamos en Dummy



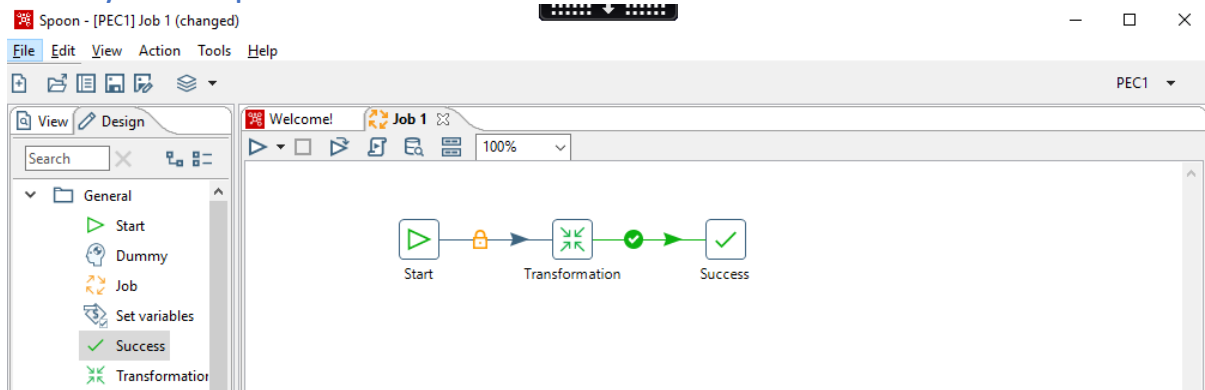
D. Crear un *job* que ejecute la transformación anterior

SOLUCION:

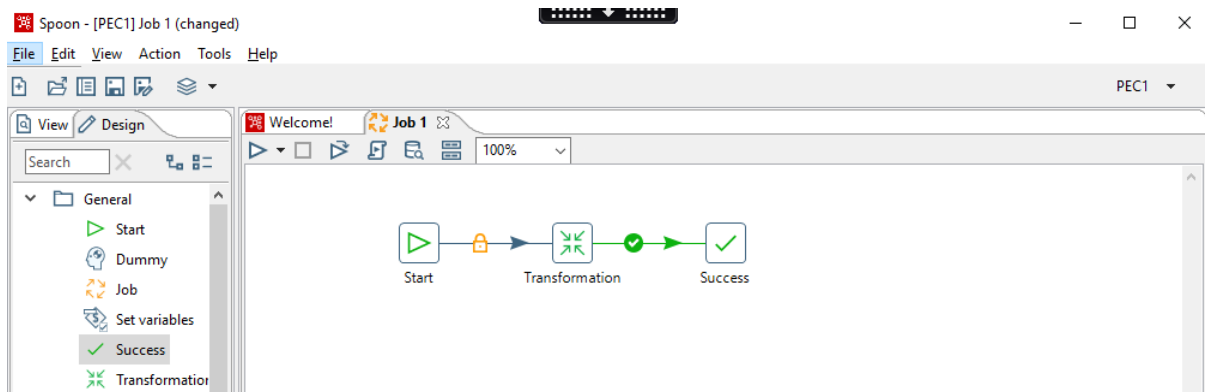
Para crear un Job le damos click a File->New->Job



Añadimos los pasos : “Start “ para empezar, “Transformation” para incorporar la transformación que incluye este Job y “Success” para finalizar



Los unimos e incorporamos en “Transformation” la transformación guarda anteriormente



Lo guardamos y lo ejecutamos

Ejercicio 4 (10%): Microsoft SQL Analysis Services

Objetivo:

Toma de contacto con la aplicación MOLAP suministrada en la máquina VDI (*Microsoft SQL Analysis Services*) y capacitar al estudiante para iniciar las actividades del curso.

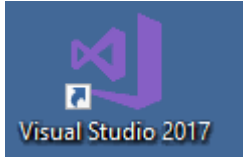
Desarrollo del ejercicio:

En este último ejercicio veremos cómo crear un cubo de SSAS y conectarlo a una base de datos.

Describid paso a paso todas las acciones realizadas, aportando las capturas de pantalla que sean necesarias. Los pasos para seguir son los siguientes:

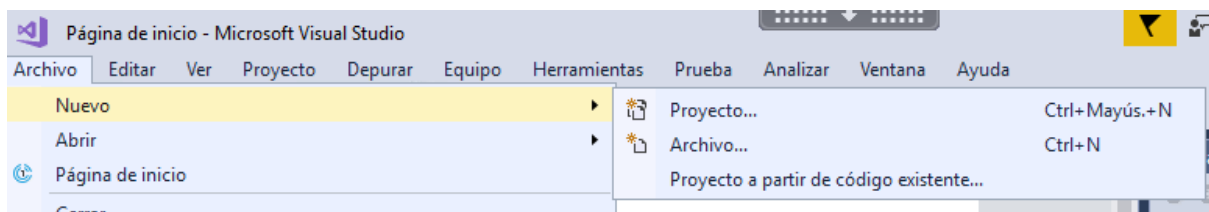
1) Abrir Visual Studio (SSAS)

Abrimos Visual Studio 2017 por medio del icono

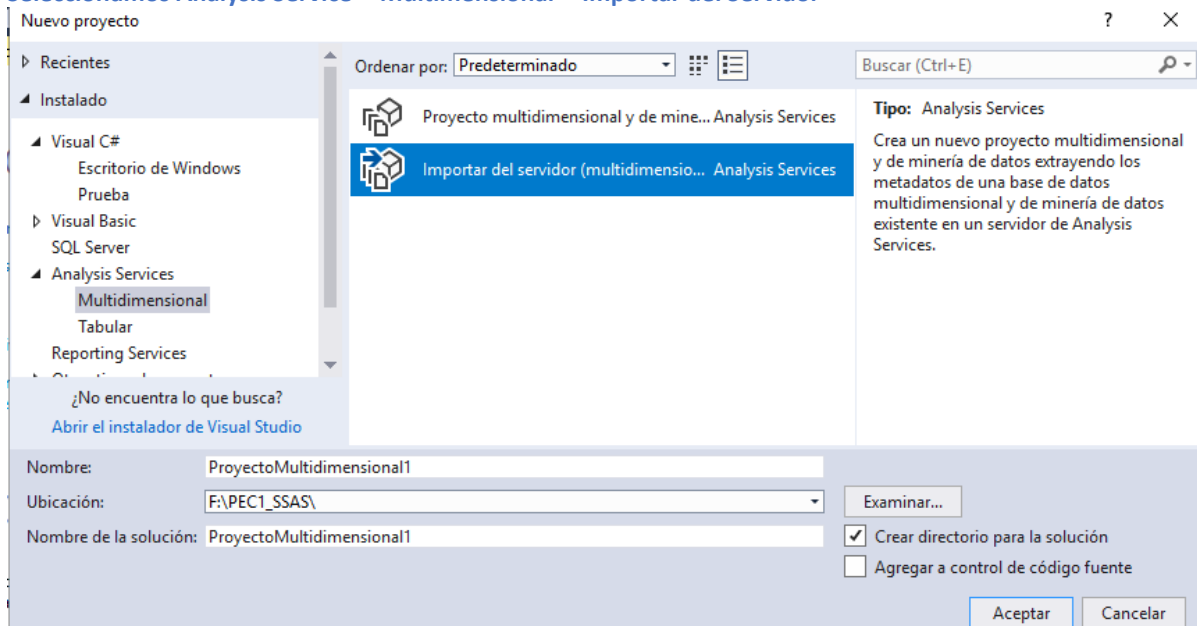


2) Crear un nuevo proyecto de Analysis Services multidimensional

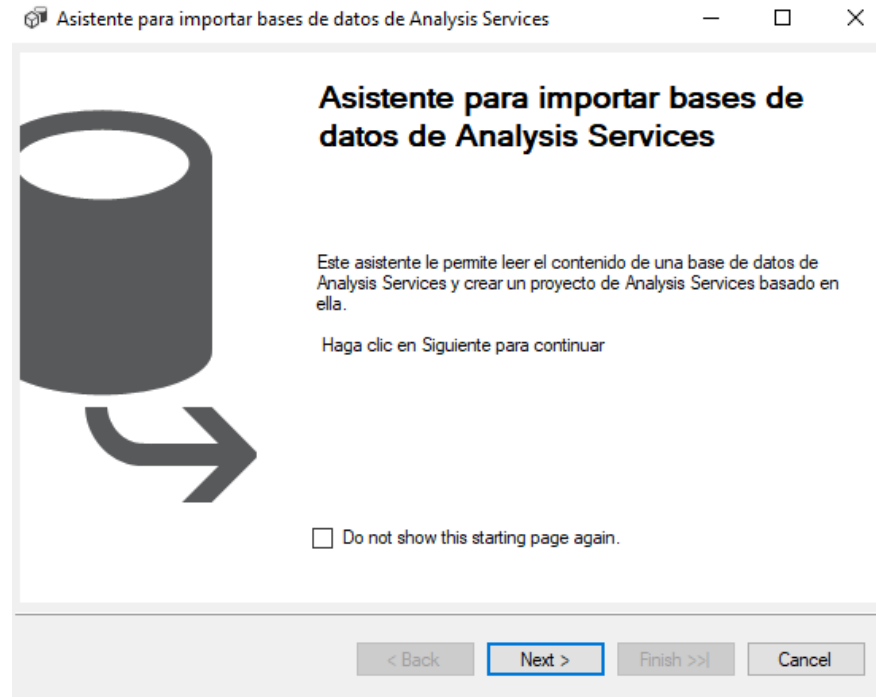
Creamos desde el Menu Archivo->Nuevo->Proyecto



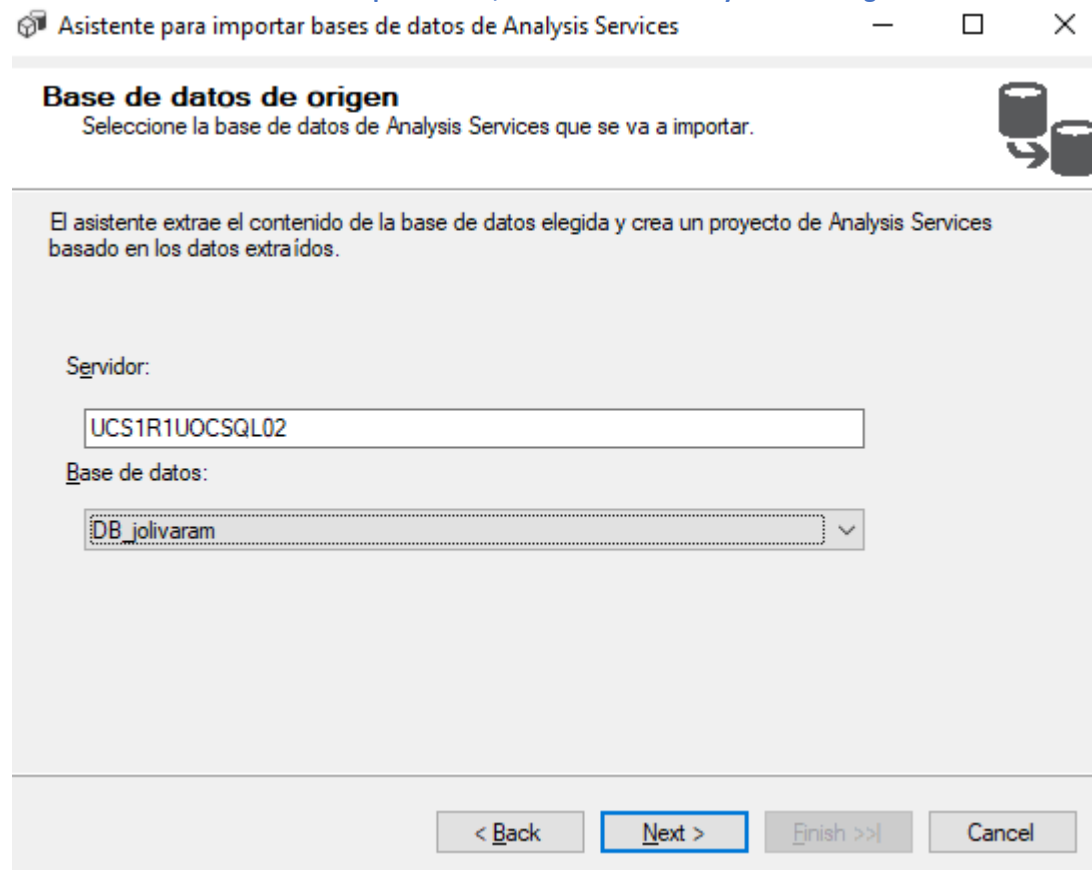
Seleccionamos Analysis Service-> Multidimensional ->Importar del Servidor



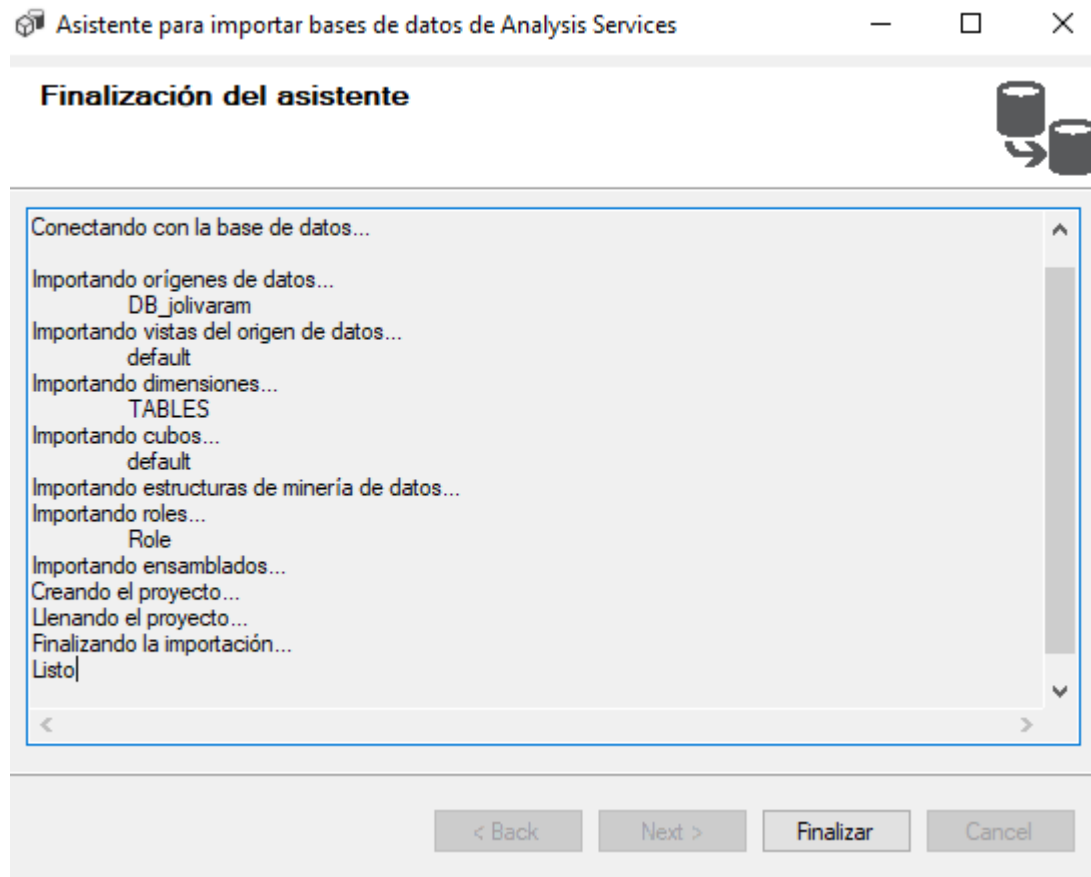
Se abre un asistente y le damos siguiente



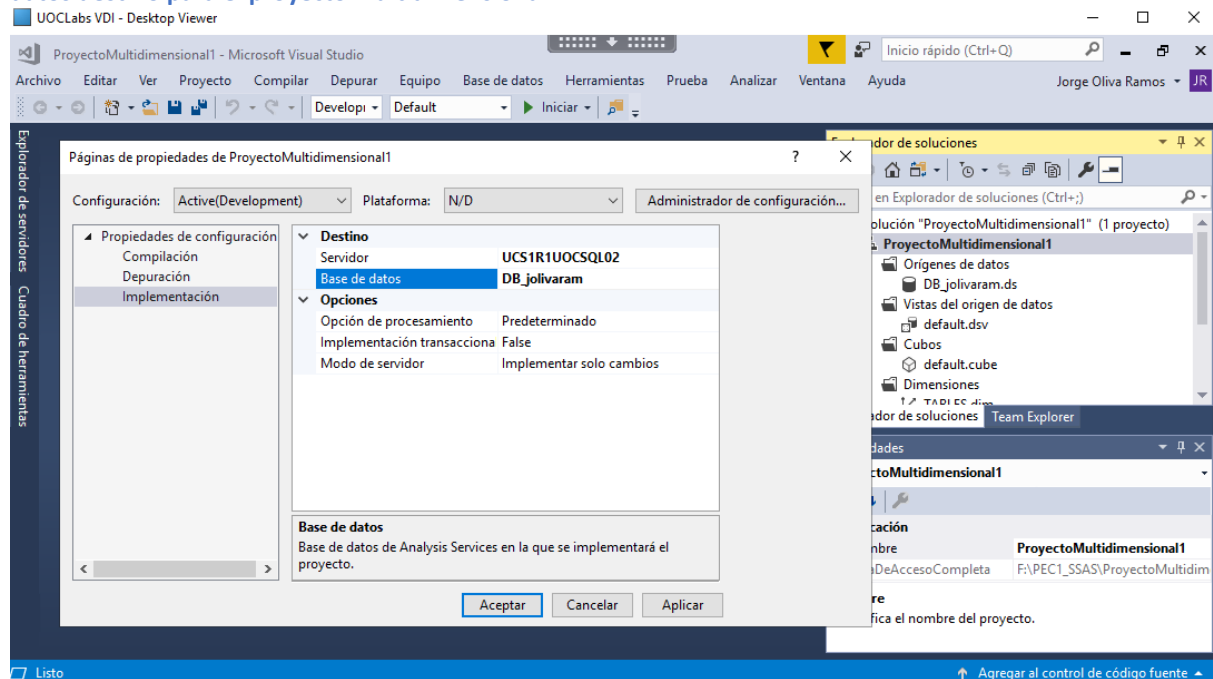
Seleccionamos el servidor correspondiente , nuestra base datos y le damos siguiente:



Por últimos nos muestra la información de conexión con nuestra DB .le damos a finalizar para crea el proyecto

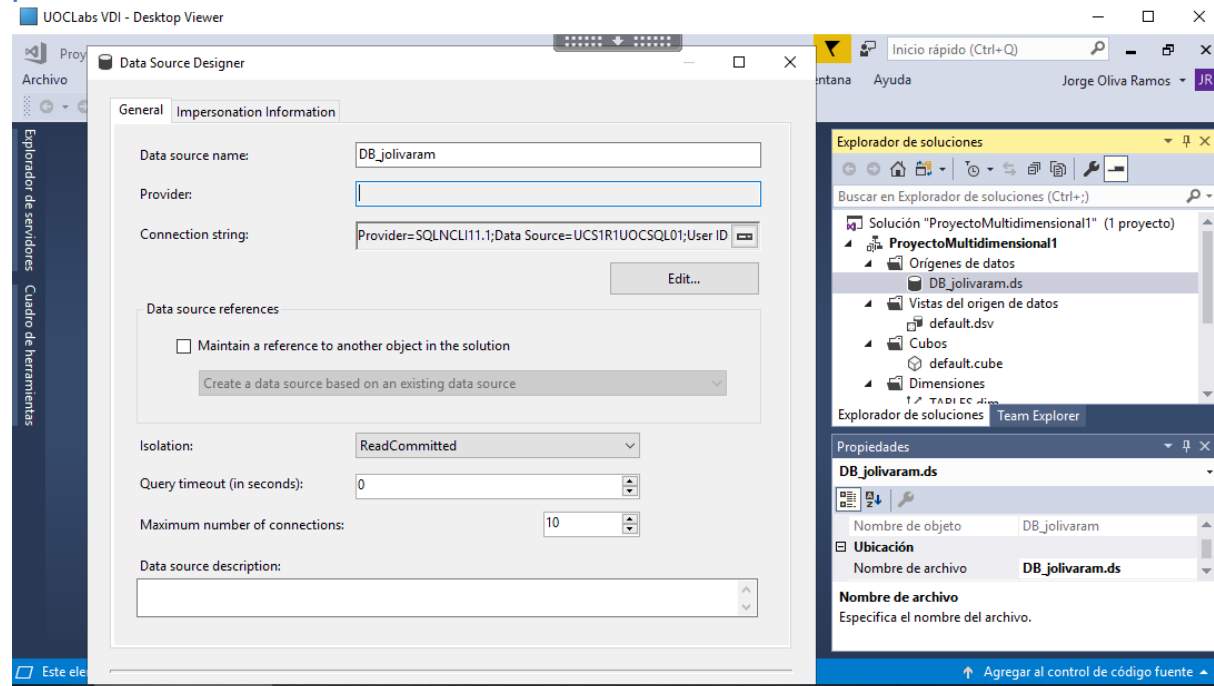


Una vez hemos creado el proyecto , vamos a propiedades del proyecto para indicarle el servidor y la base de datos destino para el proyecto multidimensional

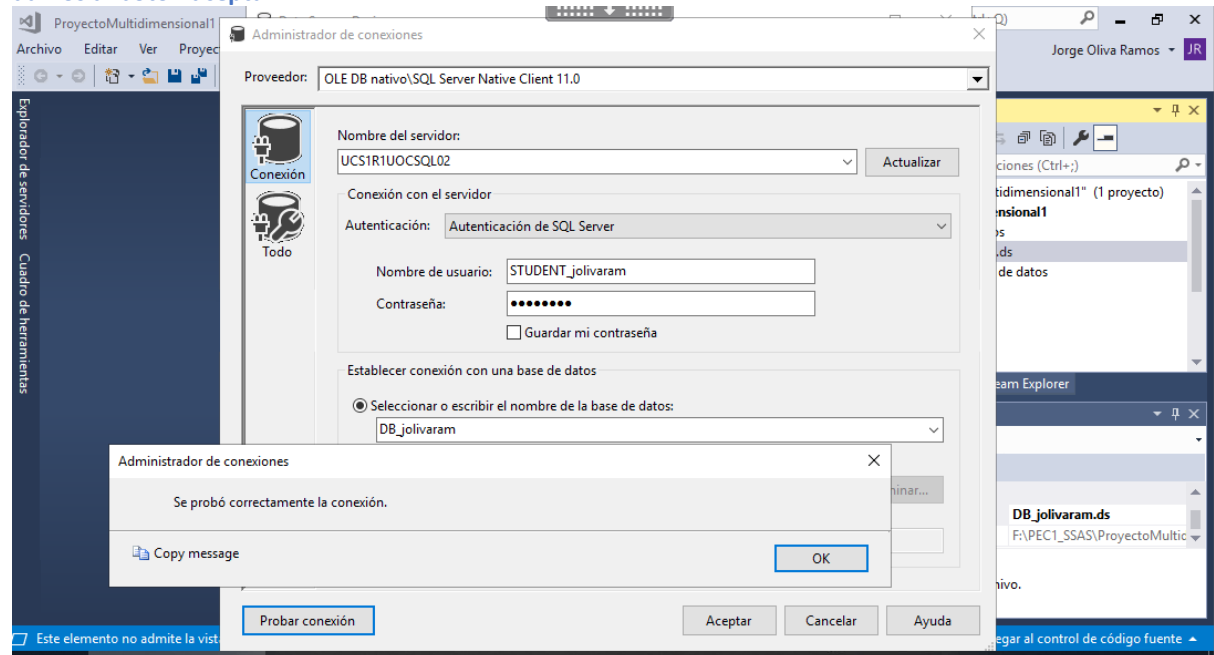


3) Crear un origen de datos (DataSource) conectado a DB_loginuoc.

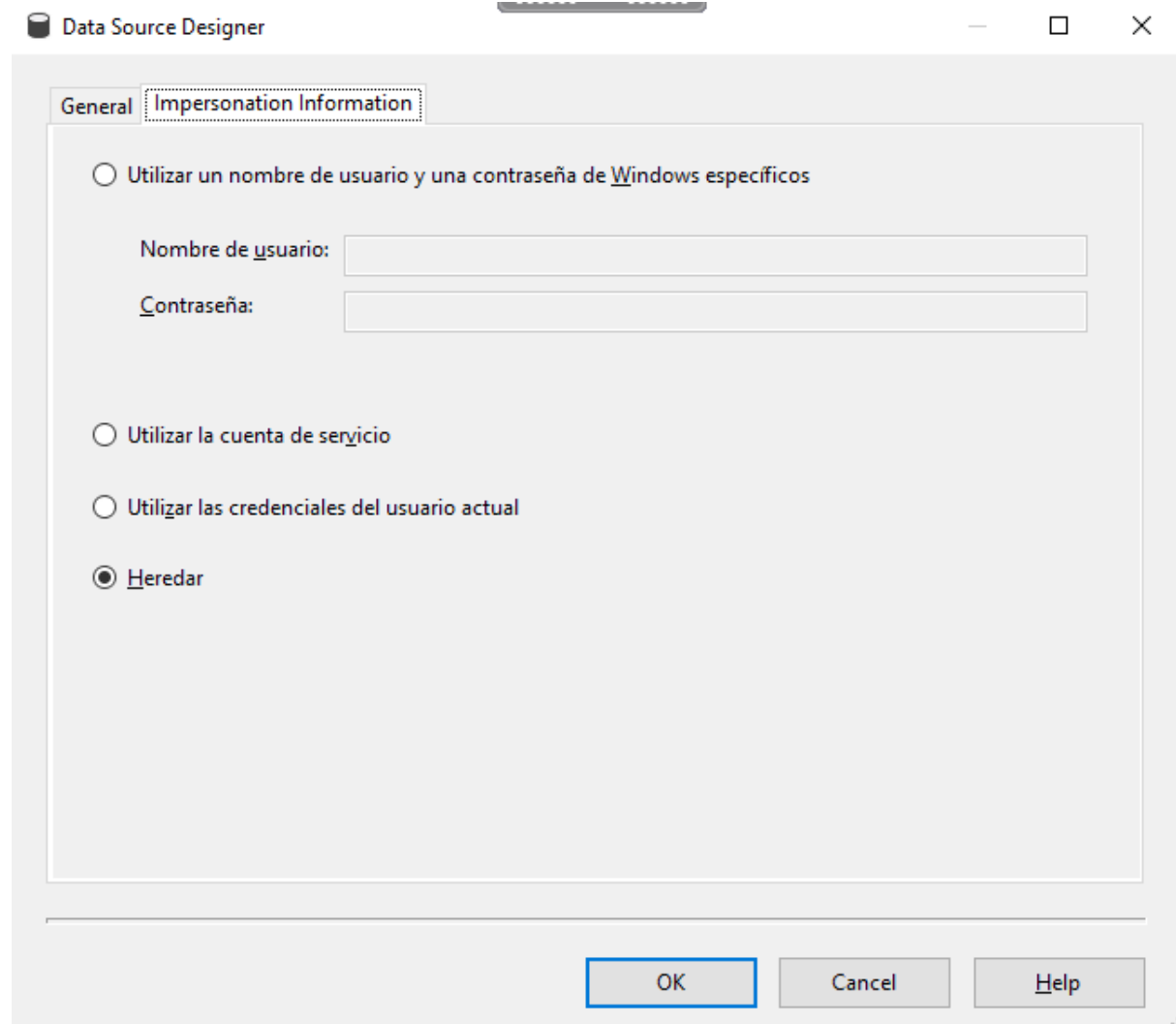
Para ello pulsamos doble click sobre la base datos origen , esto nos abre una pantalla del data source y aquí pulsamos el botón edit



Introducimos los parámetros de conexión y Probamos la conexión una vez se verifica que es correcta le damos al botón aceptar.



Pulsamos la pestaña de Impersonal Información , comprobamos que tengamos Heredar y pulsamos el botón OK



The screenshot shows the 'Data Source Designer' window with the 'Impersonation Information' tab selected. The window has a title bar with the text 'Data Source Designer' and standard window controls. The 'Impersonation Information' tab is active, showing three radio button options: 'Utilizar un nombre de usuario y una contraseña de Windows específicos', 'Utilizar la cuenta de servicio', and 'Utilizar las credenciales del usuario actual'. The 'Utilizar un nombre de usuario y una contraseña de Windows específicos' option is selected, with input fields for 'Nombre de usuario:' and 'Contraseña:'. The 'Utilizar la cuenta de servicio' and 'Utilizar las credenciales del usuario actual' options are unselected. The 'Heredar' option is selected, indicated by a filled radio button. At the bottom of the window, there are three buttons: 'OK', 'Cancel', and 'Help'. The 'OK' button is highlighted with a blue border.

General Impersonation Information

☐ Utilizar un nombre de usuario y una contraseña de Windows específicos

Nombre de usuario:

Contraseña:

☐ Utilizar la cuenta de servicio

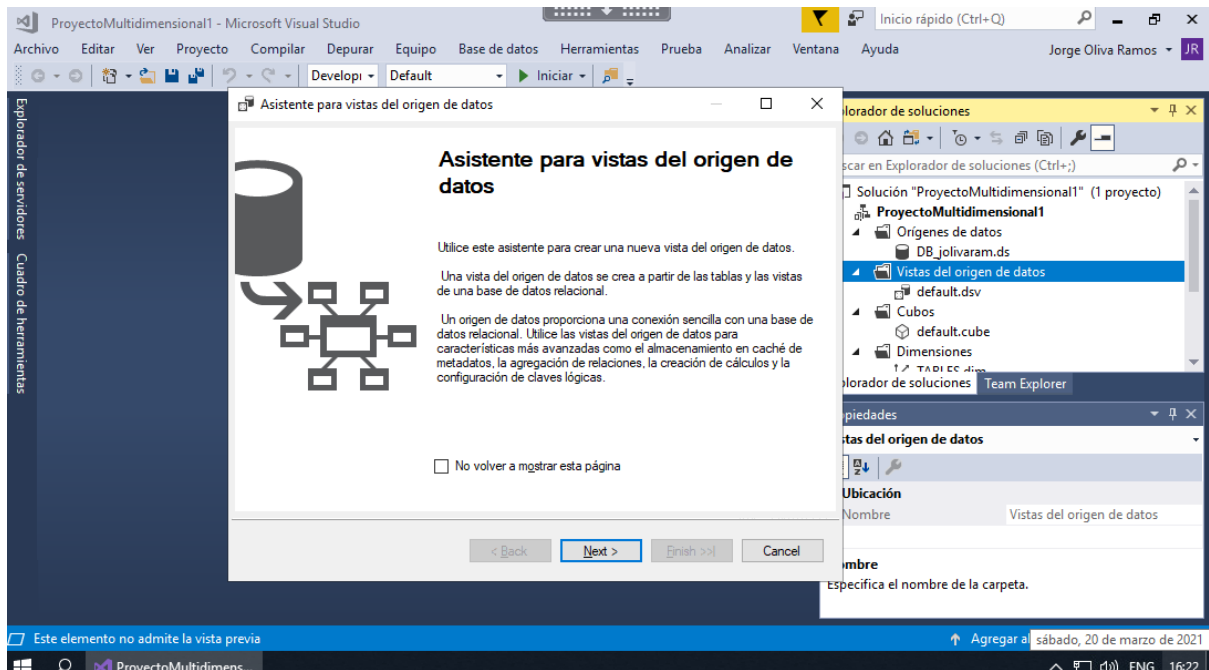
☐ Utilizar las credenciales del usuario actual

☒ Heredar

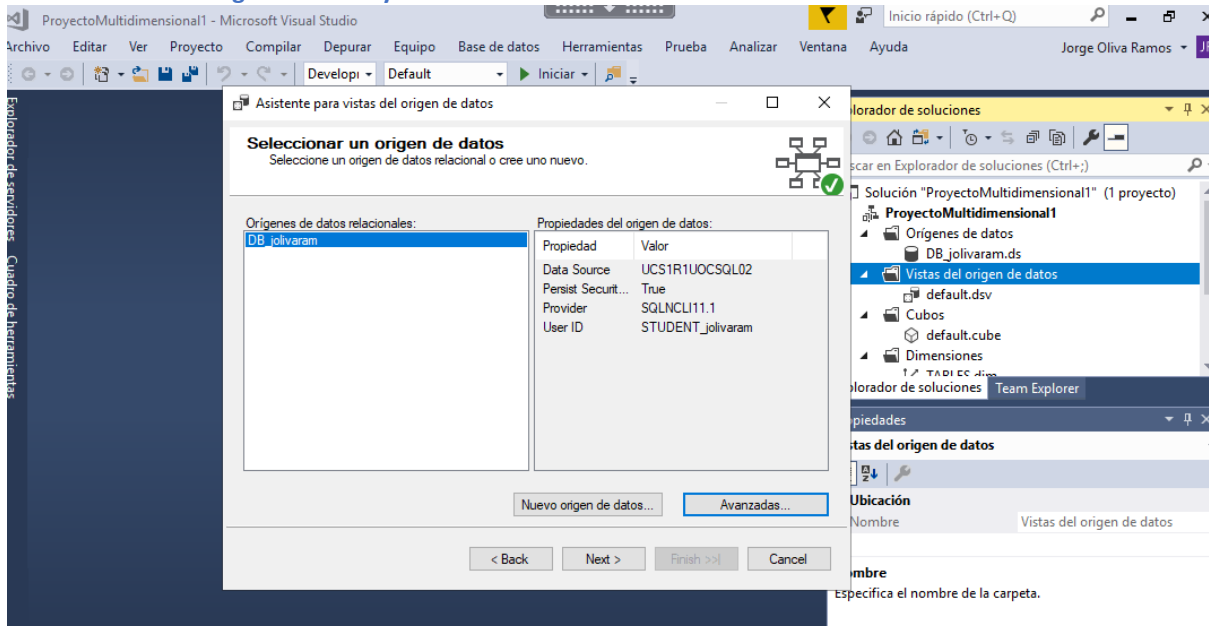
OK Cancel Help

4) Crear una vista de origen de datos (Data Source View)

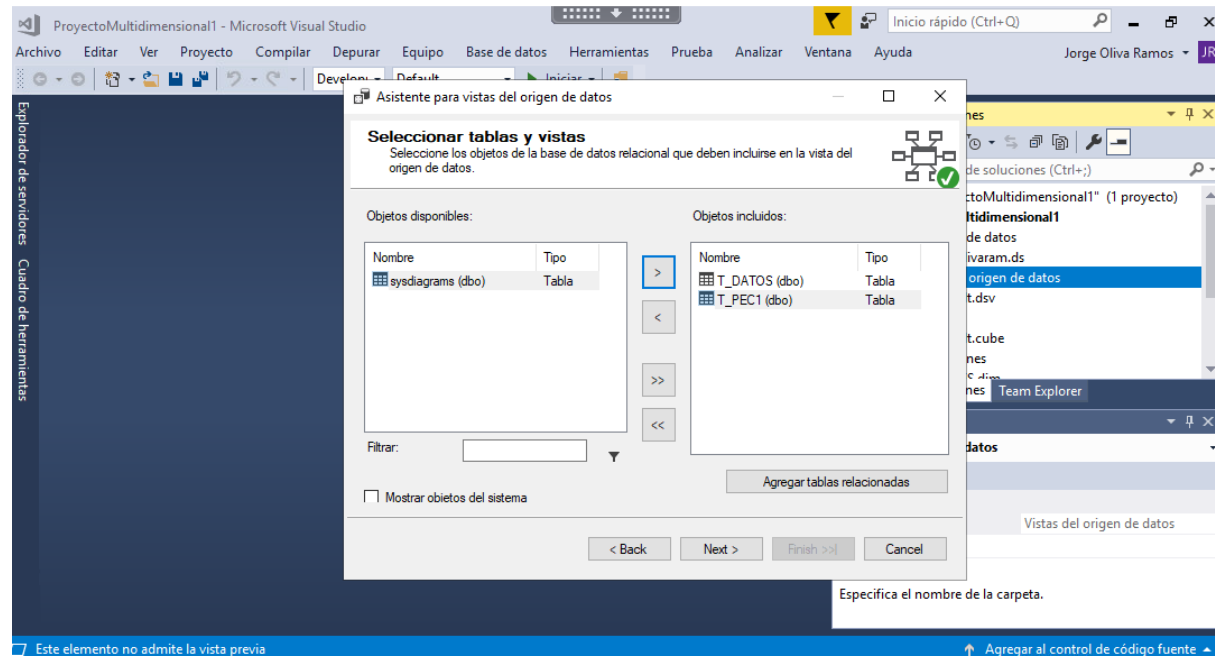
En vista origen de datos pulsamos nuevo y nos apertura un asistente al cual le damos “next”



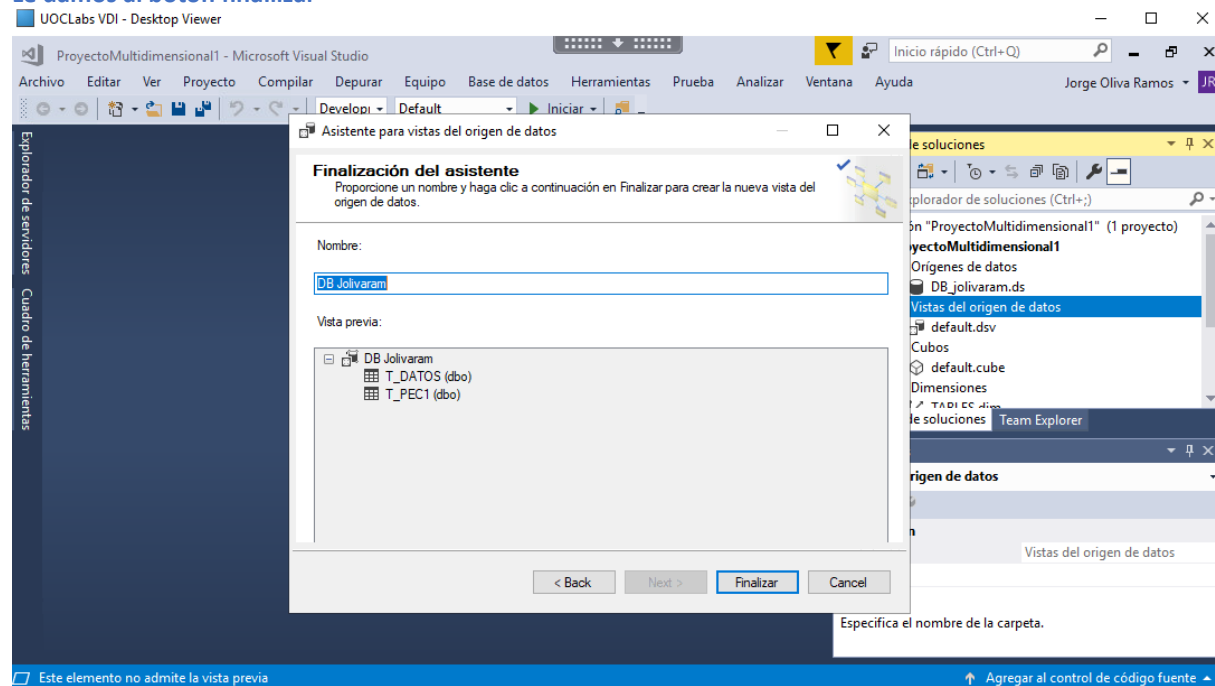
Seleccionamos el origen de datos y le damos click a next



Seleccionamos las tablas

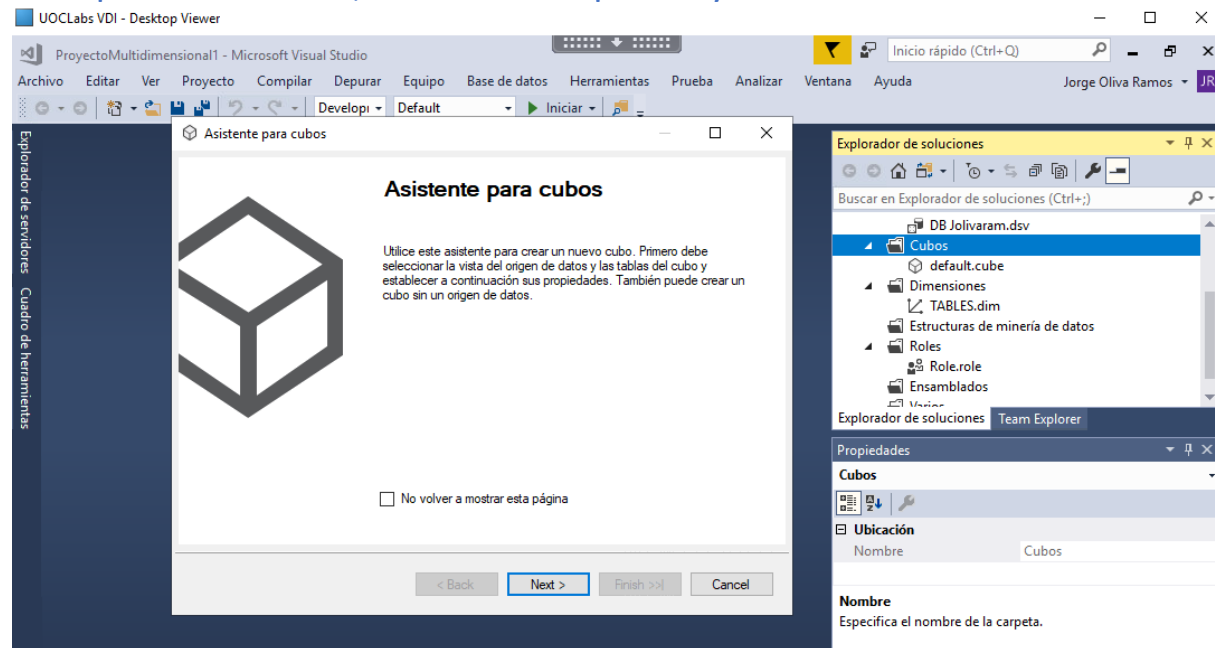


Le damos al botón finalizar

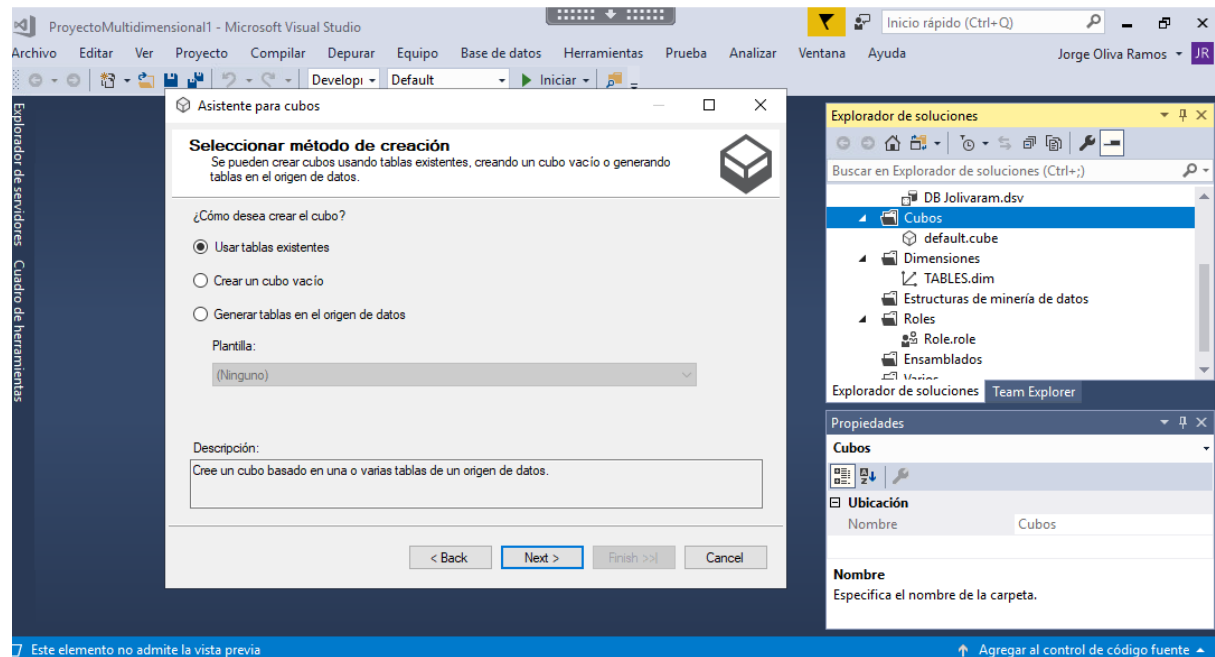


5) Crear un cubo con una única tabla de hechos T_Datos.

En el explotador de soluciones , seleccionamos la carpeta cubo y le damos a nuevo



Le damos a la opción por defecto y pulsamos siguiente



Seleccionamos la TABLA requerida

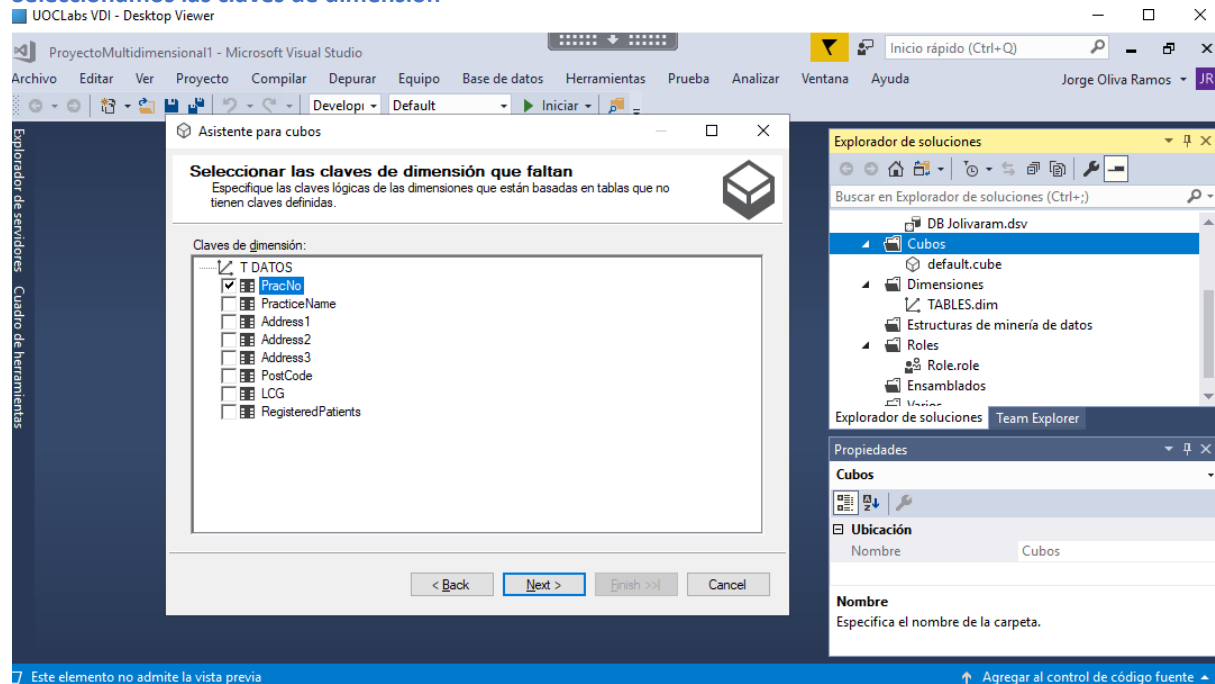
The screenshot shows the Microsoft Visual Studio interface with the 'Asistente para cubos' (Cube Wizard) open. The wizard is at the 'Seleccionar tablas de grupo de medida' (Select measure group tables) step. The 'Vista del origen de datos' (Data source view) is set to 'DB Jolivaram'. Under 'Tablas de grupo de medida' (Measure group tables), the table 'T_DATOS' is selected with a checkmark, while 'T_PEC1' is not. The 'Explorador de soluciones' (Solution Explorer) pane on the right shows the project structure, including 'DB Jolivaram.dsv', 'Cubos', 'default.cube', 'Dimensiones', 'TABLES.dim', 'Estructuras de minería de datos', 'Roles', 'Role.role', and 'Ensamblados'. The 'Propiedades' (Properties) pane below it shows the 'Cubos' (Cubes) property set, with the 'Ubicación' (Location) property set to 'Cubos'.

Seleccionamos las medidas del cubo

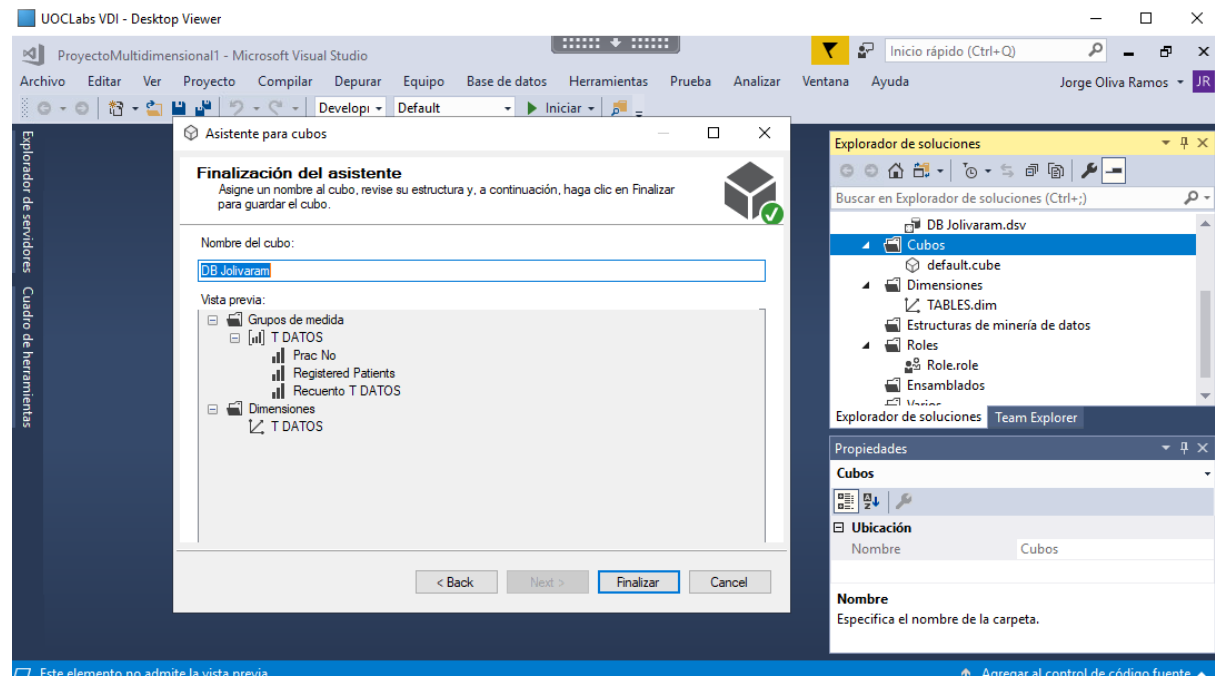
The screenshot shows the Microsoft Visual Studio interface with the 'Asistente para cubos' (Cube Wizard) open. The wizard is at the 'Seleccionar medidas' (Select measures) step. The 'Medida' (Measure) list shows 'T_DATOS' selected with a checkmark, and 'Prac No', 'Registered Patients', and 'Recuento T DATOS' are also listed. The 'Explorador de soluciones' (Solution Explorer) pane on the right shows the project structure, including 'DB Jolivaram.dsv', 'Cubos', 'default.cube', 'Dimensiones', 'TABLES.dim', 'Estructuras de minería de datos', 'Roles', 'Role.role', and 'Ensamblados'. The 'Propiedades' (Properties) pane below it shows the 'Cubos' (Cubes) property set, with the 'Ubicación' (Location) property set to 'Cubos'.

Este elemento no admite la vista previa Agregar al control de código fuente

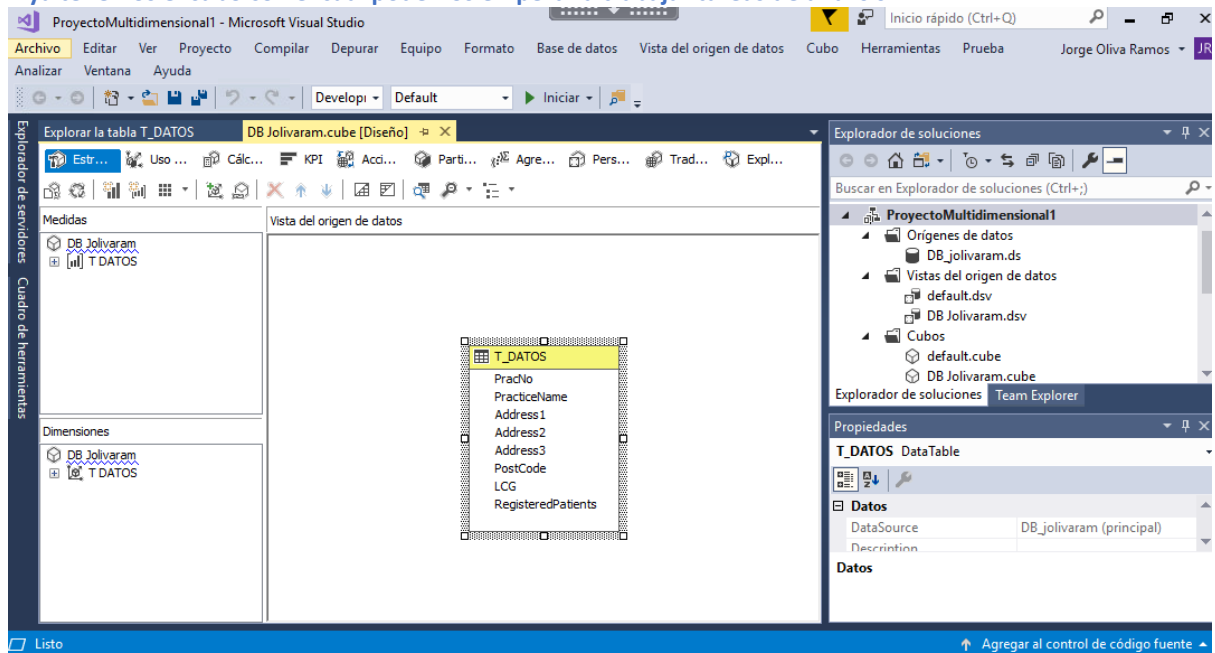
Seleccionamos las claves de dimension



Por últimos pulsamos el botón finalizar



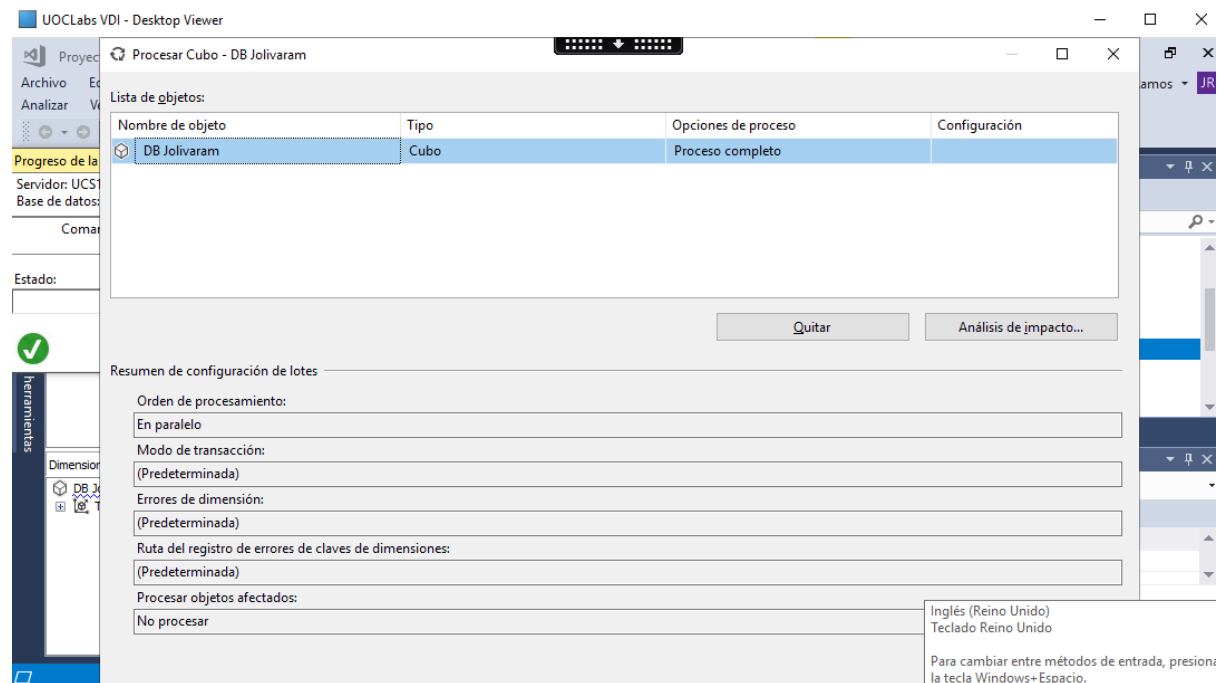
Y ya tenemos el cubo con el cual podemos empezar a trabajar tareas de análisis .



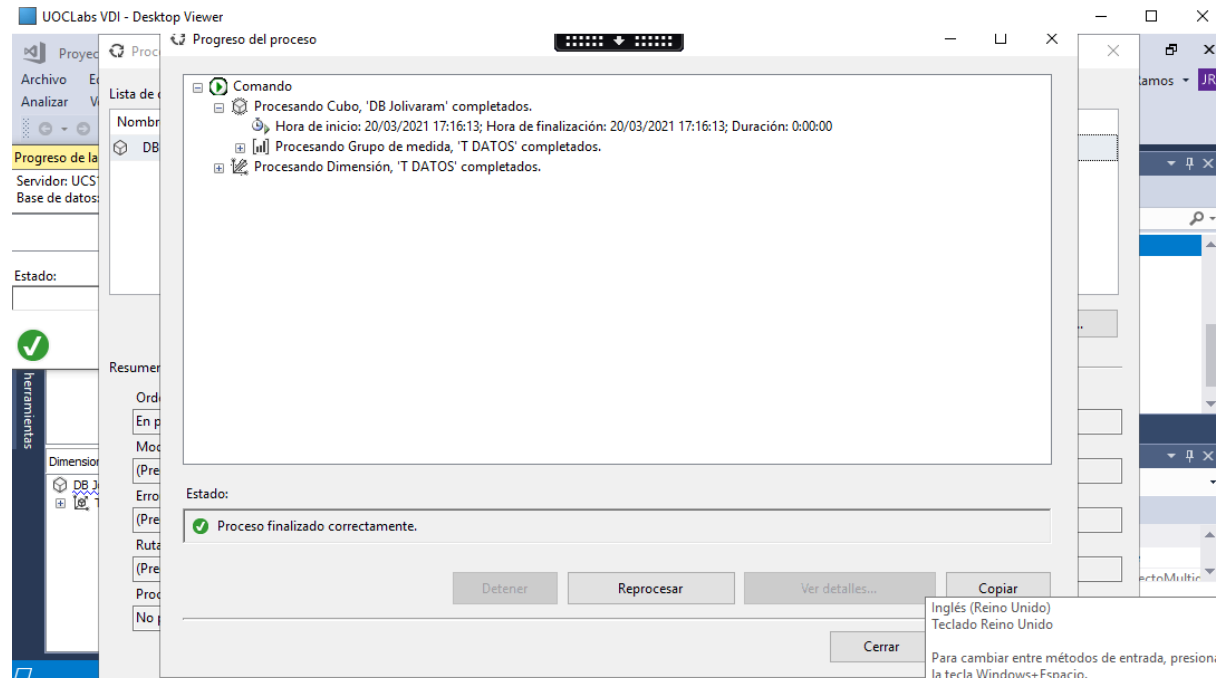
6) Procesar e implementar el cubo en el servidor

Lógicamente un cubo tan simple carece de funcionalidad, el objetivo de esta PEC únicamente es validar las herramientas disponibles.

Seleccionamos el cubo creado en el punto anterior y le damos a procesar, nos abre esta pantalla y le damos a ejecutar



Una vez procesado nos muestra un resumen del proceso , le damos cerrar



Ahora lo tenemos disponible el cubo para explorar

