**Universidad san buenaventura**

**Manual de instalación de pentaho-uso de transformaciones**

**Mario reyes casas**

**Manuel cordoba**

**Eduardo Santamaria**

**Datawharehouse**

**Bogotá D.C**

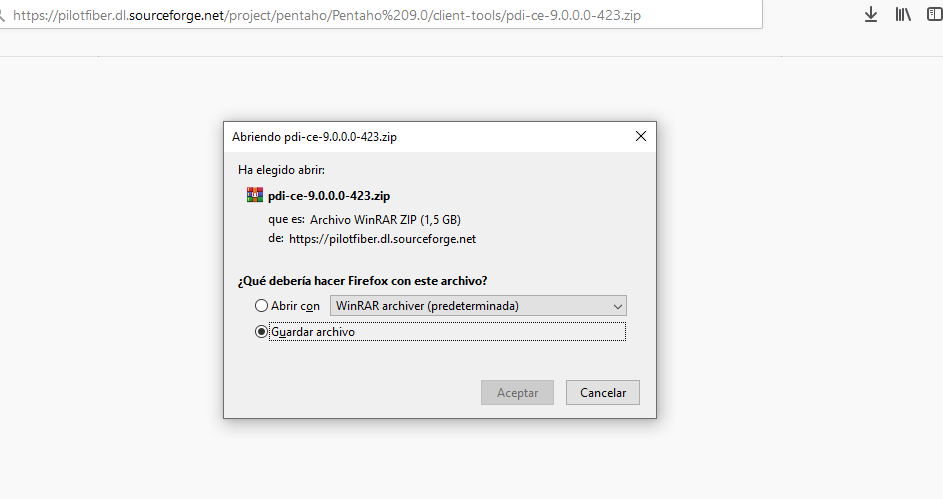
**2020**

Pentaho es un herramienta de ETL’S , se pueden realizar transformaciones en esta herramienta ,vamos a ver como se instala :

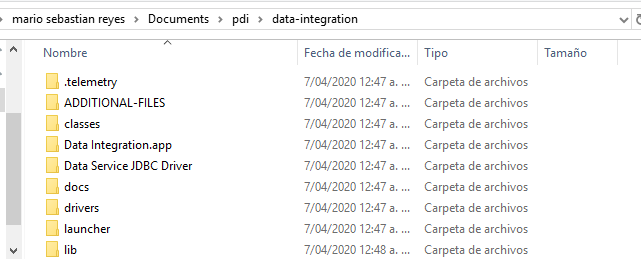
Seguir el siguiente enlace

<https://pilotfiber.dl.sourceforge.net/project/pentaho/Pentaho%209.0/client-tools/pdi-ce-9.0.0.0-423.zip>

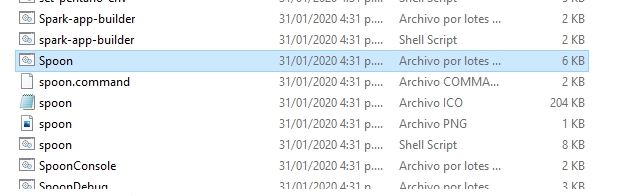
aparecerá lo siguiente



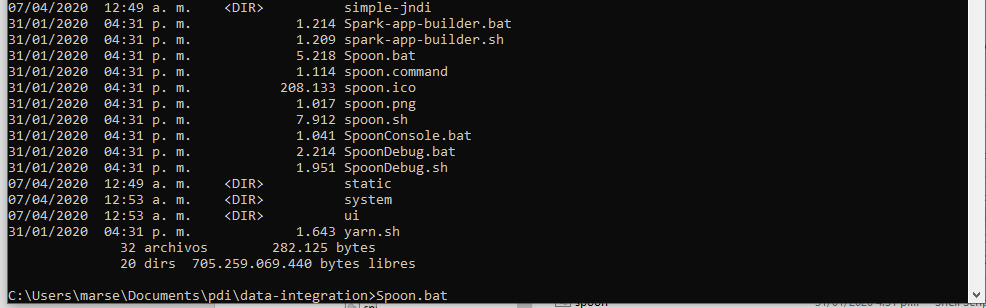
Guardamos y procedemos a descomprimir el archivo , se ha creado una carpeta y dentro de esa carpeta se guardará el .zip al descomprimir se obtiene:

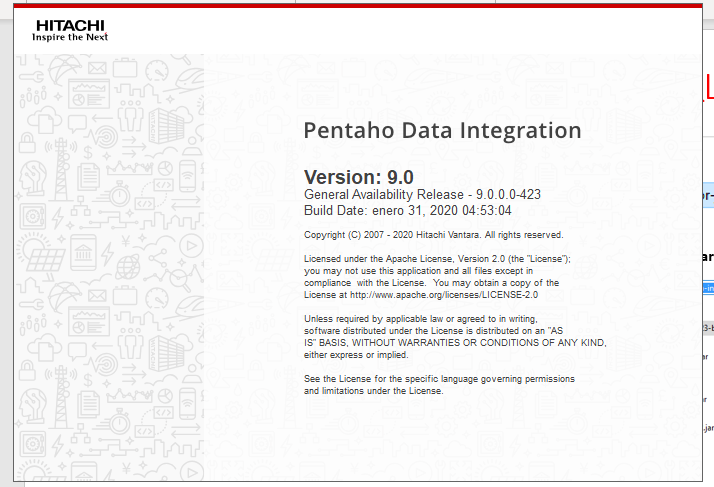


Para ejecutar hay dos opciones



Dar click en el .bat o por el contrario entrar a la cmd y activarlo

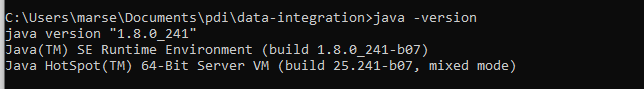




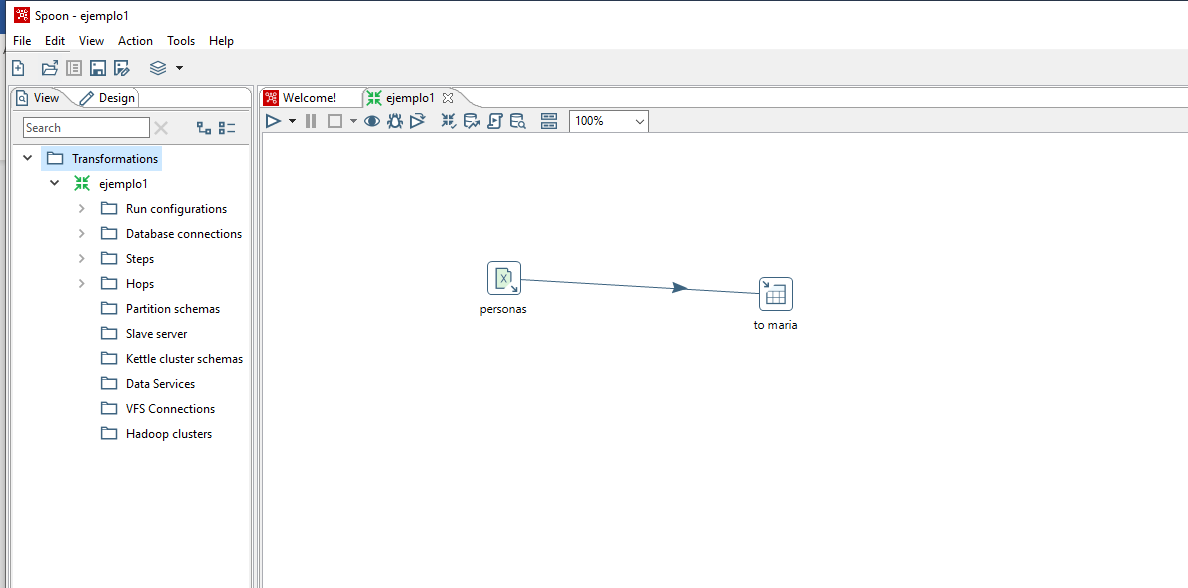
Si se quieren adjuntar nuevos conectores hay que ir a la siguiente ruta

C:\Users\marse\Documents\pdi\data-integration

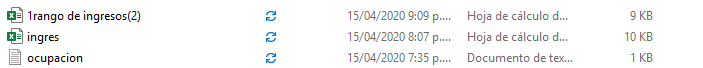
**ADVERTENCIA:**  para correr este proyecto se debe tener java , para verificar si se tiene java se ejecuta el comando



Este es el resultado que se obtiene



Una vez adentro se tienen la siguientes fuentes o inputs



También se le considera fuente a la base de datos obtenida de la peluquería

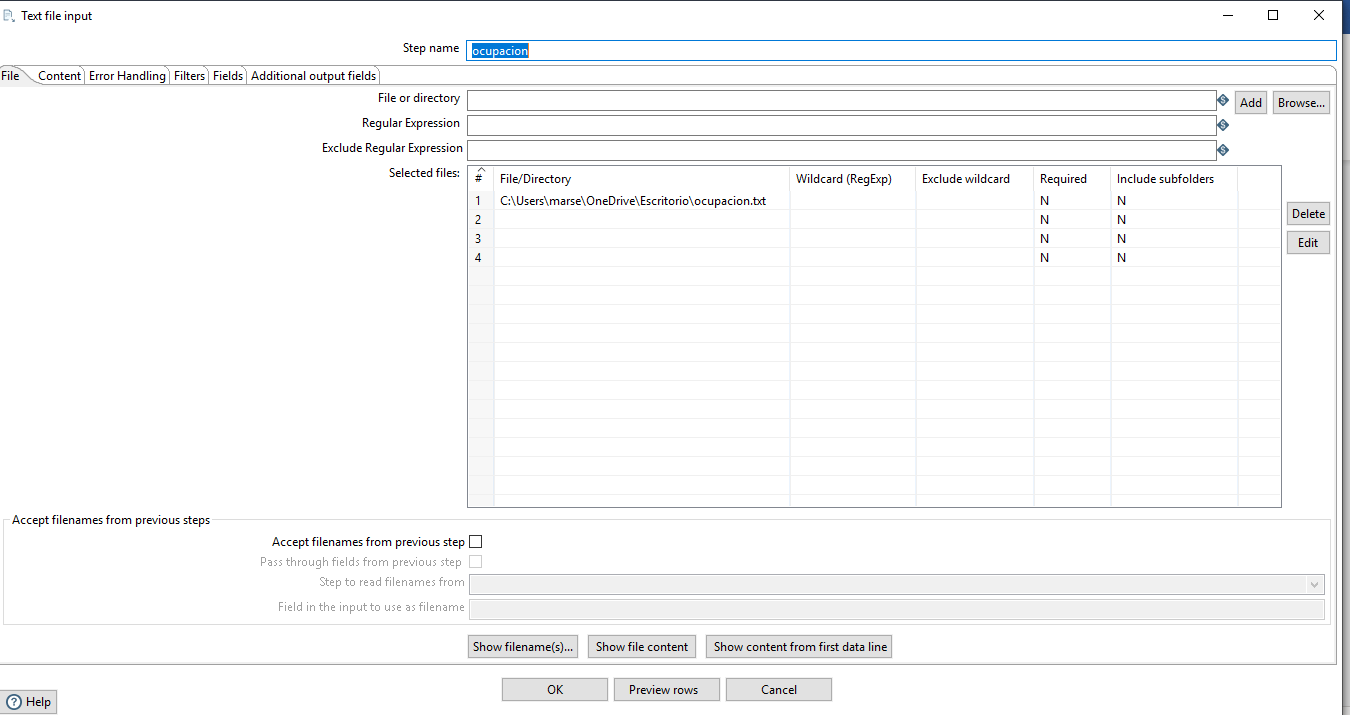


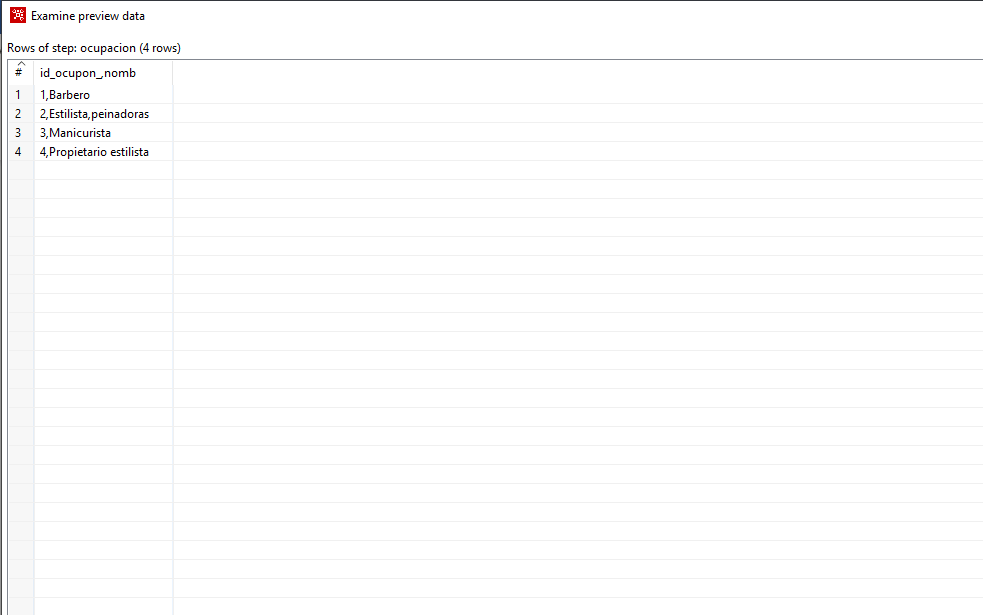
Ahora , se debe abrir el spoon

**Transformación 1**

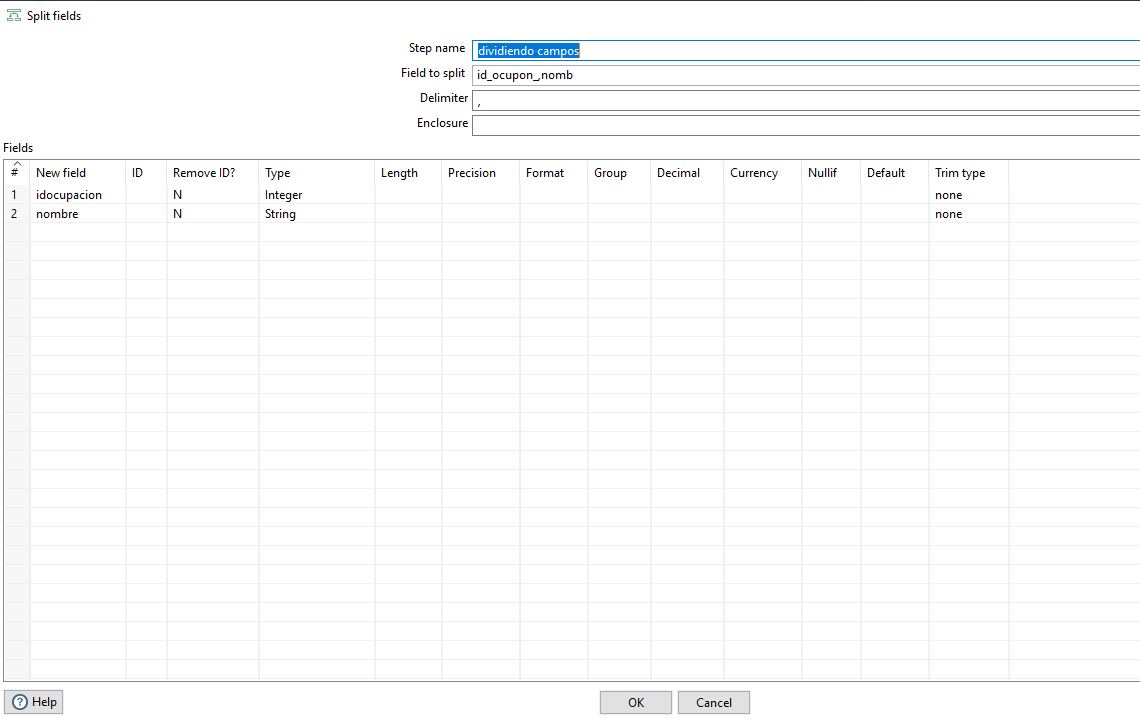


Lo primero es usar un input txt

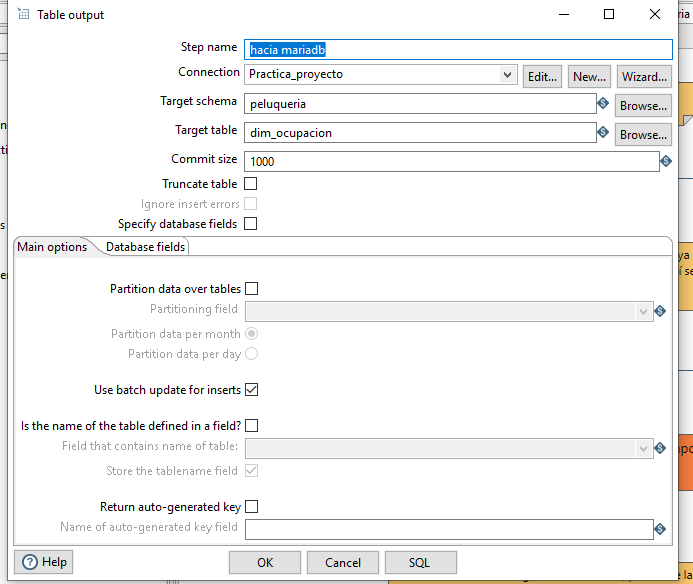




Como se ven los datos están juntos y sin campos que los identifiquen

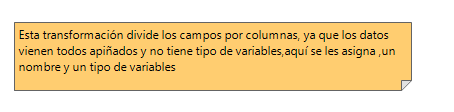


Acá se dividen y se les asigna un tipo de variable ,en field to Split se tiene el dato en bruto que se irá a dividir

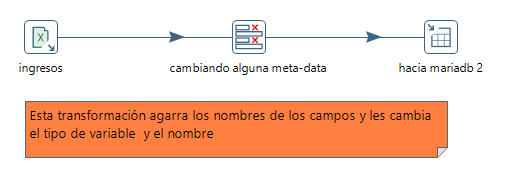


Finalmente pasa a la base de datos

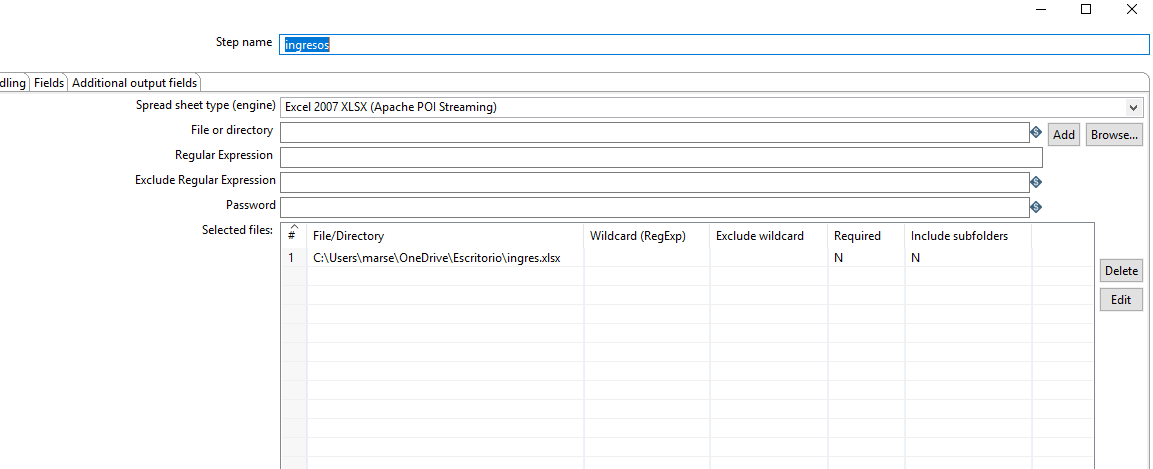
A un lado está la descripción de la misma



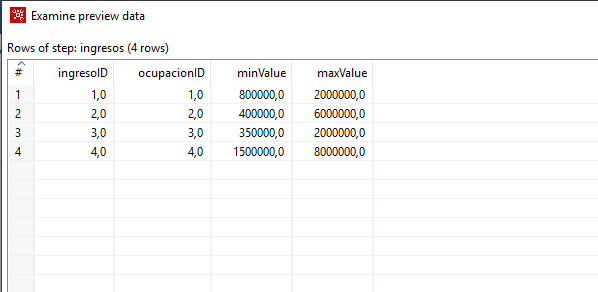
**Transformación 2**

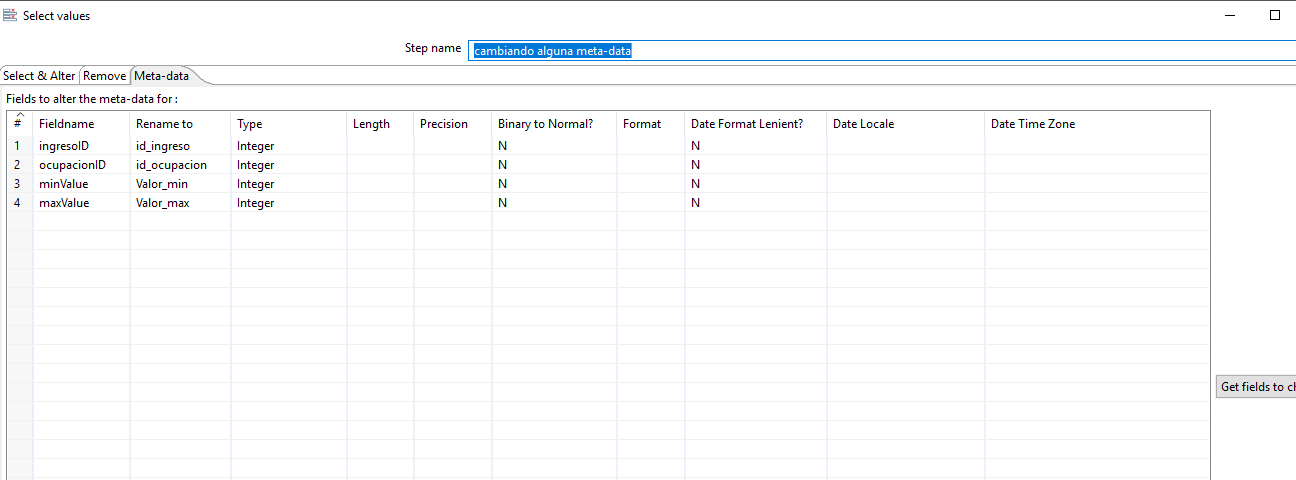


En caso de que ya se tenga un xlsx listo ,dividido por celdas con sus respectivos tipos de datos es posible que hayan datos incorrectos , para esto se utiliza el select values ,que nos permite asignarle valor a un campo y cambiarle el nombre



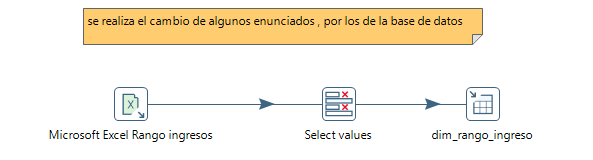
Este es el documento usado



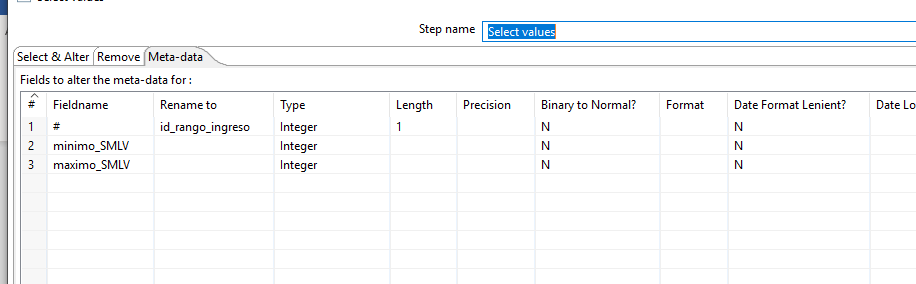
seleccionamos los datos a cambiar ,antes hay dos opciones ,una cambia el nombre nada mas , la otra solo elimina los campos ,es decir si hay 10 campos solo usar 5 , y el final ,que se esta usando, simplemente cambia la metadata de los datos ,finalmente se pasa a la base de datos

**Transformación 3**

Se replica el mismo ejercicio con un Excel



Seleccionamos los campos y les cambiamos el tipo de variable para que concuerde con la base de datos



**Posible transformación 4**

Ya que nuestra fuente primordial es la base de datos de la peluquera , pasar de sql a sql es muy sencillo y no requiere casi de controlar los datos , es decir cambiar variables , dividir celdas , aumentar y recortar strings etc.

Ahora se realizó la carga automáticamente :

