

# Seminario sobre GeoKettle

## Quinta actividad

### Enunciado del problema.

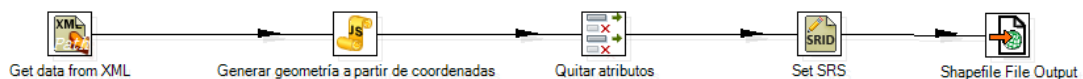
Disponemos de un **archivo de texto xml** que contiene etiquetas, atributos y datos en este formato. A continuación se muestra la primera parte del archivo para observar la estructura.

```
<markers>
<marker idcentro="2134" name="ESCUELA DE MONTA&#209;A DE ALQU&#201;ZAR" address="CL. SAN GREGORIO, 30 22145 Alqu&#233;zar Huesca" lat="42.173035" lng="0.02508" type="Privado" denom="Centro Privado Autorizado de Ense&#241;anzas Deportivas" piloncho="3"/>
<marker idcentro="2024" name="ENCANTARIA" address="CL. RAM&#211;N Y CAJAL, 1 22300 Barbastro Huesca" lat="42.036283" lng="0.132931" type="Privado" denom="Centro Privado de Educaci&#243;n Infantil" piloncho="3"/>
<marker idcentro="1762" name="ESCUELA DE MONTA&#209;A DE BENASQUE" address="Cl. Campalet, S/N 22440 Benasque Huesca" lat="42.603497" lng="0.518408" type="Privado" denom="Centro Privado Autorizado de Ense&#241;anzas Deportivas" piloncho="3"/>
<marker idcentro="1146" name="ESCUELA ESPA&#209;OLA DE ESQU&#205;" address="Ctra. De Francia, S/N 22889 Candanch&#250; Huesca" lat="42.76724" lng="-0.512321" type="Privado" denom="Centro Privado Autorizado de Ense&#241;anzas Deportivas" piloncho="3"/>
```

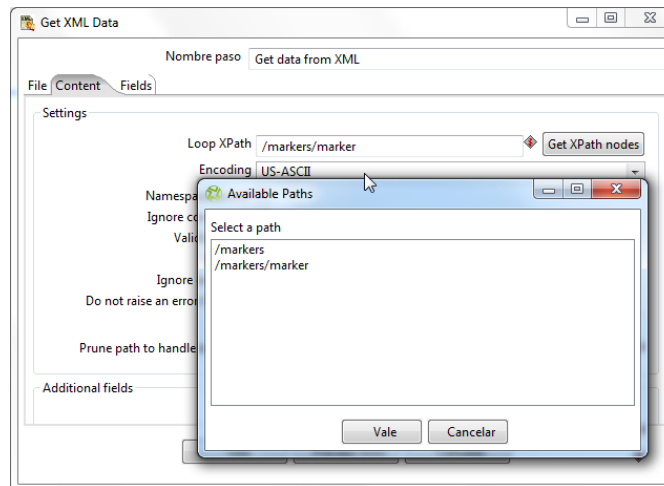
Se desea **convertir esta información en una capa shapefile** que preserve los atributos almacenados en el documento xml.

### Solución basada en GeoKettle:

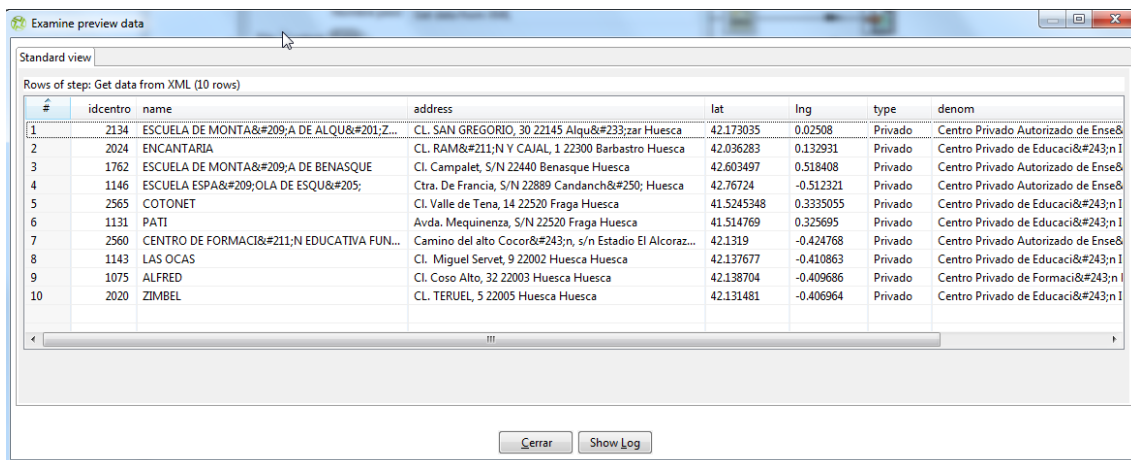
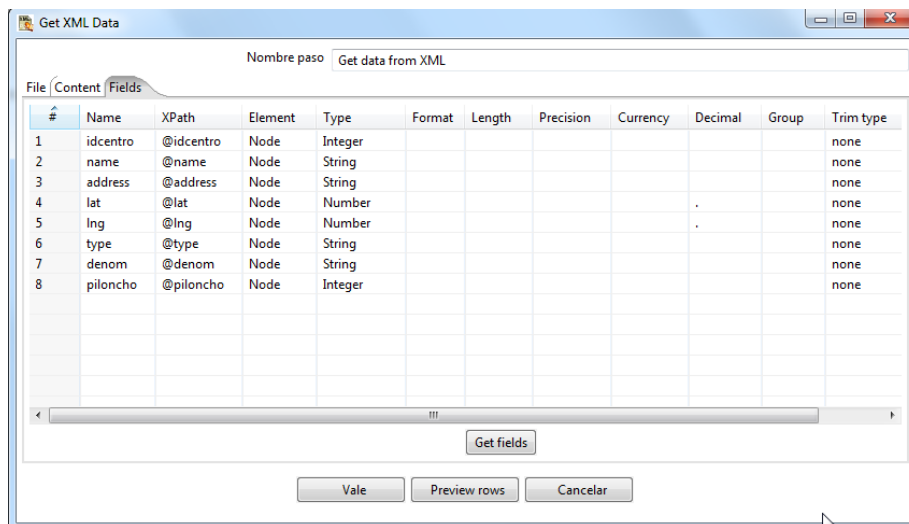
Crear una transformación que abra un fichero xml, seleccione los elementos del árbol xml con el patrón /markers/marker y seleccione los atributos (en este caso todos son atributos del elemento marker). Seguidamente se genera el atributo con la geometría a partir de las coordenadas geográficas del xml. Se adapta el tipo de dato del atributo generado para la geometría a formato Geometría y se eliminan algunos atributos innecesarios. En cuarto lugar establecer el sistema de coordenadas (SRID de la geometría) y finalmente almacenar el contenido en el documento shapefile.



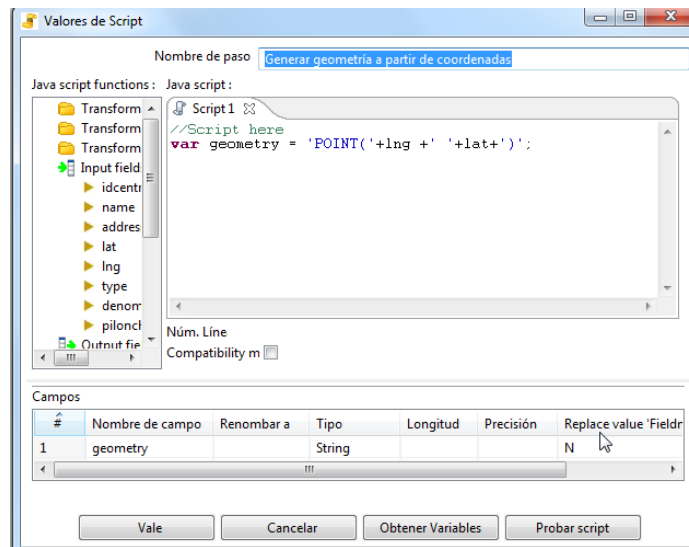
La primera etapa de la transformación, además de conectar con el fichero xml de los datos iniciales, hay que indicar en **Content** el patrón de árbol de etiquetas xml a utilizar. Para conocer dicho patrón se dispone de la herramienta GetXpathNodes, como se puede ver en la siguiente figura. Una vez identificado el patrón de etiquetas a usar se selecciona y se indica el tipo de codificación del texto usado en el documento xml. En nuestro caso hemos seleccionado US-ASCII.



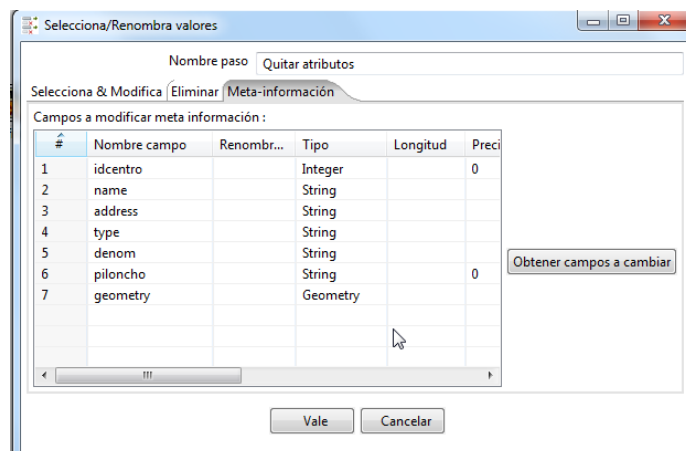
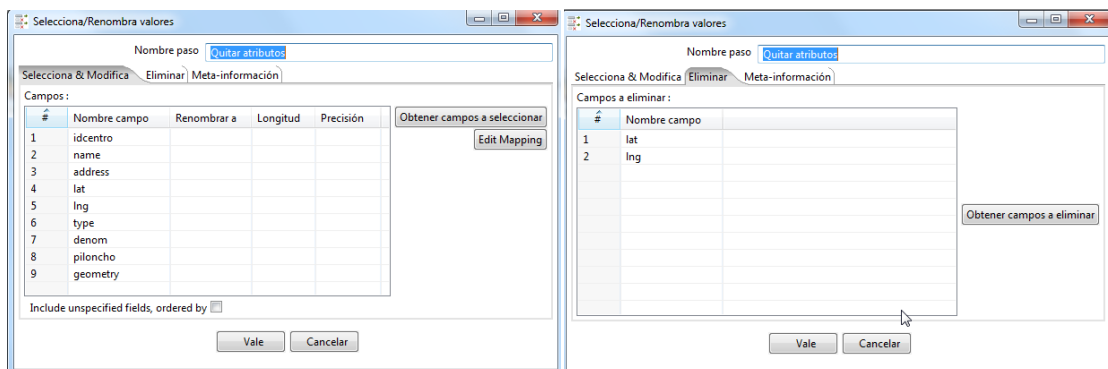
En tercer lugar hay que indicar en la solapa **Fields** los nodos, atributos y tipos de datos se pretende recuperar. Se dispone de la herramienta Get Fields que ayuda en esta tarea. No obstante antes de aceptar los resultados que la herramienta propone se recomienda echar una ojeada a los datos. Para esto se dispone de la herramienta Preview rows a la que se puede indicar cuantas filas mostrar. Por ejemplo 10. Ver ilustración.



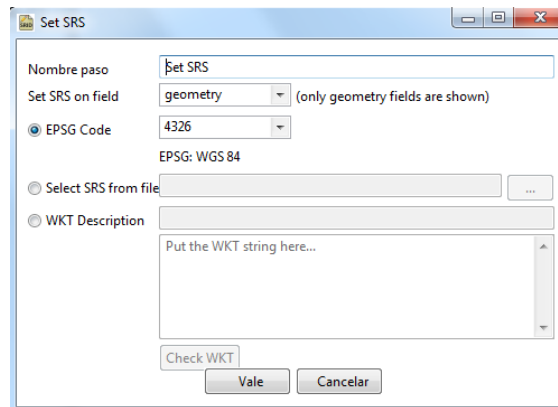
Además de reemplazar el punto por la coma, se han generado dos nuevos atributos latitud y longitud que se usarán en la siguiente etapa para generar el objeto geométrico. Ver siguiente figura. Se usará la herramienta Valor Java Script Modificado.



Si se desea eliminar como atributo textual las coordenadas en el shapefile, usaremos la herramienta Selecciona/renombrar atributos como se muestra en la siguiente figura. Además se aprovecha para convertir al vuelo de String a Geometry el tipo de dato del atributo geometry generado con el Script de java anterior.



Una vez generada la geometría, hay que establecer el sistema de coordenadas. Ver siguiente etapa de transformación tipo Set SRS.



Finalmente la etapa de almacenamiento en fichero shapefile, estableciendo el directorio y nombre de archivo a generar (no olvidad establecer la extensión shp al archivo).

