# Seminario sobre GeoKettle

## Tercera actividad

#### Enunciado del problema.

Se dispone de un *archivo de texto con columnas separadas por punto y coma (;)* y entre los datos están las *coordenadas del punto que representa el objeto georreferenciado*. Los datos de ejemplo son de estaciones meteorológicas que registran datos en toda España y los datos son cada 10 minutos. Se pretende *generar una capa en formato shapefile*, aunque también se podría almacenar en cualquier formato soportado por Geokettle, entre ellos PostGIS.

## Solución inmediata:

Utilizar una herramienta de SIG, tipo QGis y abrir dicha capa como una de texto delimitado y una vez abierto salvarlo como un archivo Shapefile.

Requiere por tanto usar un programa distinto del objetivo de este seminario.

## Solución basada en GeoKettle:

Crear una transformación simple que abra el archivo de un directorio, diferencie los distintos campos separados por punto y coma en el archivo, genere una geometría a partir de las coordenadas geográficas (longitud y latitud) para lo cual habrá que convertir el separador decimal coma en punto. Después seleccionar y transformar el formato WKT de la geometría de tipo punto a tipo geometry, establecer el sistema de coordenadas de las geometrías y finalmente almacenar en el archivo Shapefile.



En la etapa de entrada de la transformación, se ha de seleccionar el archivo a abrir, el tipo de archivo y los campos que se van a identificar.

Entrada Fichero de Texto		
chero Contenido Maneio de Erro	Nombre paso Entrada Fichero de Testo	
	Tipo de fichero CSV	
	Separador de campos	Insertar TABulación
	Separador de texto	
Property of the Arrest		
Grennin baitos de anes	Escape	
	Cabacara 🐱 Número de líneas de cabacara 🗔	
D	Pie C Numero de lineto de tate 1	
	JLineas cortadas? 🗌 Número de veces que se conta 🛐	
	Paginado (impresión)? 🗌 Número de líneas por página [80	
	Lineas en cabecera documento	
	Comprimido (Zip) None	
	Eliminar filas vacias 🔽	
de	cluir nombre del fichero en salida? [] Compo con el nombre del fichero	
	"Número de fila en salida? 🗖 Campo con el número de fila	5 7
	Nimero de Ela por fichero?	
	formate 1005	
	Politico DOS	•
	Coancación	-
	Limite 0	
	¿Ser flexible al leer fechas?	
	Ublizar el formato de fecha local(e) es_es	<u>-</u>
Result filenames		
	Add filenames to result 🗭	
	Vale Cancelar Previsualizar filas	1

En la solapa Campos, se selecciona Traer Campos y luego cancelar para que los atributos sean todos de tipo texto.

E .A.	Nombre	Tipo	Formato	Posición	Longitud	Precisión	Moneda	Decimal	Grupo	Nulo si	Por defecto	Tipo de poda	Repetir	1 +
1	INDICATIVO	String										ninguno	N	-
2	AÑO	String										ninguno	N	
3	MES	String										ninguno	N	
	DIA	String										ninguno	N	
	HORA	String										ninguno	N	
	MINUTO	String										ninguno	N	
	NOMBRE	String										ninguno	N	
	ALTITUD	String										ninguno	N	
	c_X	String										ninguno	N	
0	CY	String										ninguno	N	
1	NOM_PROV	String										ninguno	N	
2	LONGITUD	String										ninguno	N	
3	LATITUD	String										ninguno	N	
4	PREC	String										ninguno	N	
5	TA	String										ninguno	N	
6	TAMAXIOM	String										ninguno	N	
7	TAMINIOM	String										ninguno	N	
8	TAMAXIH	String										ninguno	N	11.2
9	TAMINIH	String										ninguno	N	
0	HTAMAXIH	String										ninguno	N	
1	HTAMINIH	String										ninguno	N	
2	VV10M	String										ninguno	N	
3	DV10M	String										ninguno	N	
4	VMAX10M	String										ninguno	N	
		Sec. 1												

Como segundo paso una transformación de tipo Replace in String. Se trata de cambiar la coma de la longitud y latitud por un punto y generar dos nuevos atributos derivados.

AB Repla	ce in string								
		Step name <u> </u>	eplace in string		И				
Fields str	ing								
# 🔺	In stream field	Out stream field	use RegEx	Search	Replace with	Whole Word	Case sensitive		
1	LONGITUD	longitud1	N	,		N	N		
2	LATITUD	latitud1	Ν			Ν	N		
<ul> <li>( )</li> </ul>									
		V	ale Get	fields	Cancelar				

Como tercer paso, un Script que genera una variable de tipo String por concatenación de texto fijo y valores de las variables que contienen la longitud y latitud con separador decimal punto.

Jalores de Script	ıbre de paso 🚺	lor Java Script M	odificado			
Java script functions : Java	script: Script 1 & Script here r geometry h. Línea mpatibility mor []	= 'POINT('+	longitud1	+' '+1&	atitudl+')';	<ul> <li>▲</li> <li>▲</li> </ul>
Campos						
# 🔺 Nombre de campo	Renombar a	Tipo	Longitud	Precisión	Replace value 'Fieldna	me' or 'Ren
1 geometry		String			Ν	
Vale		Cancelar	Obtener V	ariables	Probar script	

Si introducimos un nombre de variable de latitud y longitud distinto al que hemos definido en el paso anterior (longitud1, latitud1), al pulsar el botón "Probar Script" y en "Vale" veremos que se nos genera un error de la variable que se ha introducido no ha sido definida:

8	ERROR	- 🗆 🗡
Han ocurrido uno o más errores durante la	a previsualización! Examina el fichero de mensajes para averiguar lo que ha ido mal.	
10 10:07:12 (, U=0, E=0 0m 2013-02-04 17.31.06 by jenkins) 0m 2	<pre>Error desconocido : org.pentaho.di core.exception.KettleValueException: Error Javaecript: ReferenceError: 'longitud" is not defined. (script#3) at org.pentaho.di trans.steps.scriptvalues_mod.ScriptValuesMod.addValues(ScriptValuesMod at org.pentaho.di trans.steps.metripvalues_mod.ScriptValuesMod run(ScriptValuesMod.java Caused by: org.mozilla.javascript.ScriptValuesMod.ScriptValuesMod.run(ScriptValuesMod.java Caused by: org.mozilla.javascript.ScriptValues.mod.ScriptValuesMod.run(ScriptValuesMod.java Caused by: org.mozilla.javascript.ScriptValuesMod.ScriptValuesMod.ins.java.3550; at org.mozilla.javascript.ScriptValuesmodError(ScriptValuesMod.scriptValuesMod.java Caused by: org.mozilla.javascript.ScriptValuesModError(ScriptValuesMod.java.3550; at org.mozilla.javascript.ScriptValuesmodError(ScriptValuesMod.java.3560) at org.mozilla.javascript.Genc2col(script]] at org.mozilla.javascript.GontextFactory.dorpCall(ScriptValuesMod.java.3620; at org.mozilla.javascript.gen.c2.col(script]] at org.mozilla.javascript.ScriptValuesModTopCall(ScriptValuesMod.java.3003] at org.mozilla.javascript.ScriptValuesModTopCall(ScriptValuesMod.addValues(ScriptValuesMod</pre>	
<		>
	Cerrar	

En probar script podemos asignar valores a las variables que estamos empleando en el script (latitud1 y longitud1) para comprobar que el script que hemos definido funciona como queremos:

<b>7</b> 目					Gene	rar Filas				-	×
			Nombre paso	## TEST DA	TA ##						
			Límite	10							+
Cam	pos :										
<del>ŵ</del>	Nombre	Tipo	Formato	Longitud	Precisión	Moneda	Decimal	Grupo	Valor		^
29	VIS	String							test value test val		
30	TPRE	String							test value test val		
31	INSO	String							test value test val		
32	RAGLOB	String							test value test val		
33	latitud1	String		100					42.6		
34	longitud1	String		100			,		2.7		~
				<u>V</u> ale	Ca	incelar	<u>P</u> revisual	izar			

Al pulsar "Vale" nos aparece la nueva variables geometry con la información correcta:

	VIS	TPRE	INSO	RAGLOB	latitud1	longitud1	geometry
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			
val	test value test val	42.6	2.7	POINT(2.7 42.6			

Como cuarto paso, una transformación de tipo selecciona Renombra para que el nuevo campo generado para la geometría que aún está en formato WKT (texto) se convierta en tipo de dato Geometry.

leccio	na & Modifica (Elimi	nar Meta-inform	sción	Kenombra val	ores		
ampo	s a modificar meta inf	formación :					
# 🔺	Nombre campo	Renombrar a	Tipo	Longitud	Precisión	Binary to Norma	
26	HR		-			N	
27	PRES		-			N	
28	NIEVE		-			N	
29	VIS		-			N	
30	TPRE		-			N	Obtener campos a cambia
31	INSO		-			N	
32	RAGLOB		-			N	
33	geometry		Geometry			N 👻	
<u>ا</u>						•	

Como quinto paso, establecer el sistema de coordenadas de las geometrías. Como las coordenadas son geográficas con datum WGS84, se usa el código EPSG 4326.

Newber area	ket SPS	
Nombre paso	persits	
Set SRS on field	geometry (only geometry fields are shown)	
EPSG Code	4326	
	EPSG: WGS 84	
C Select SRS from file		
C WKT Description		
	Put the WKT string here	
		-
	Check WKT	

Finalmente la etapa de almacenamiento de datos en formato Shapefile.

Shapefile File Output		
Nombre paso	Shapefile File Output	
Nombre fichero	C:\Users\Miguel A. Manso\Desktop\curso geokettle postgis\salida-geor	Examinar
Filename is defined in a field?		
Step to read filenames from		~
Get filename from field		-
Encoding		•
	Vale Cancelar	