

CIPFP www.fpmislata.com							
Actividad:	Pentaho Data Integración (PDI)						
Ciclo:	IABD	Modulo:	BDA	Curso:		Agrupación:	1
Alumno/a:						Grupo:	

Actividad

Crea una transformación en Pentaho para realizar las siguientes tareas

Data Profiling

Analiza los datos con especial atención en los campos POSTALCODE y COUNTRY

Extracción

1. Con el fichero **sales_data.csv** realiza una extracción de datos

Transformación

2. Filtra las filas con el POSTALCODE distintas de NULL
3. Resuelve las filas con POSTALCODE igual a NULL con el fichero **Zipssortedbycitystate.csv** utilizando un paso *StreamLookup (Busqueda en Flujo de datos)* y seleccionado los codigos postales encontrados con un paso *Select Values*

Limpieza de datos

4. Unifica los valores del campo COUNTRY y convierte United States a USA con el paso *Value Mapper*
5. Clasifica los valores del campo SALES en pequeño, medio y grande con el paso *Number Range*

Carga

6. Carga las filas en un Tabla

Crea un job con la transformación, que compruebe la existencia de los ficheros de entrada.

Ejecuta la transformación y el job, desde la consola con *pan* y *kitchen*