

CIPFP <a href="http://www.fpmislata.com">www.fpmislata.com</a>							
<b>Actividad:</b>	<b>Iniciación en Pentaho Data Integración (PDI)</b>						
<b>Ciclo:</b>	IABD	<b>Modulo:</b>	BDA	<b>Curso:</b>		<b>Agrupación:</b>	1
<b>Alumno/a:</b>						<b>Grupo:</b>	

## 1.- Enunciado

La jefatura de estudios de un instituto quiere crear un cuadro de mando con datos de la satisfacción de los alumnos y los profesores con distintos aspectos del funcionamiento y las instalaciones del centro.

El cuadro de mando debe tener algunos filtros como:

- Los departamentos a los que pertenecen los profesores
- Familia profesional de los estudios del alumno
- Los estudios que realizan los alumnos

Para ello nos facilita los datos extraídos de las encuestas gestionadas a través de una plataforma de e-learning.

Los datos están en formato csv. Una empresa de ámbito internacional nos proporciona datos de sus ventas en distintos formatos y nos pide que los reunamos todos y construyamos un único fichero con toda la información reunida.

## 2.- Tareas

### Estudio de los datos:

- Data Profiling
- Limpieza de datos

### Almacenamiento

- Plantea una estructura de almacenamiento de los datos limpios y preparados para hacer el análisis y crear el cuadro de mando.
- Data Lake o Data Warehouse o Ambos
- ....

Se entrega al alumno lo siguiente:

- Ficheros en formato csv
- Fichero SQL para la creación de la tablas
- Fichero SQL para la inserción de datos en tablas auxiliares
- Plantilla de NiFi con la solución casi completa

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

### 3.- Ayuda

Estas son las imágenes del procesador y contralador de servicio necesarios para completar la tarea:

Configure Processor | LookupRecord 1.27.0

Invalid

SETTINGS

SCHEDULING

PROPERTIES

RELATIONSHIPS

COMMENTS

Required field

Property	Value
Record Reader	CSVReader_PuntoyComa →
Record Writer	CSVRecordSetWriter_PuntoyComa →
Lookup Service	SimpleCsvFileLookupService_Familias →
Result RecordPath	/id_familia
Routing Strategy	Route to 'success'
Record Result Contents	Insert Entire Record
Record Update Strategy	Use Property
Cache Size	0
key	/id_encuesta

CANCEL

APPLY

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

Configure Controller Service | SimpleCsvFileLookupService 1.27.0

X DISABLED

SETTINGS
PROPERTIES
COMMENTS

Required field

🔄
+

Property	Value
CSV File	🔍 /home/nifi/Temp/Familias_tmp
CSV Format	🔍 Custom Format
Character Set	🔍 UTF-8
Lookup Key Column	🔍 id_encuesta
Ignore Duplicates	🔍 true
Value Separator	🔍 ;
Quote Character	🔍 "
Quote Mode	🔍 Quote Minimal
Comment Marker	🔍 No value set
Escape Character	🔍 \
Trim Fields	🔍 true
Lookup Value Column	🔍 familia

CANCEL
APPLY

#### 4.- Entrega:

El alumno debe entregar una plantilla de NiFI con la solución completa