

PRÁCTICA Nº 2.5

Impala



- Nombre y apellidos: Alvaro Lucio-Villegas de Cea



Índice

Ejercicio 1	3
Ejercicio 2	4
Ejercicio 3	5

Ejercicio 1

Utilizando Impala mediante HUE, carga una tabla con los datos aparcamiento bicicleta, cuyo almacenamiento sea en formato CSV y el directorio donde se almacene se encuentre dentro de /user/cloudera para que pueda ser encontrado fácilmente.

Observamos la tabla una vez subida a HDFS.

The screenshot shows the HUE web interface. On the left, there's a sidebar with navigation icons and a list of files in the /user/cloudera directory. The main area displays the details of the file 'aparcamiento-bicicleta(1).csv'. It shows the file's location, size, and a preview of its contents. The preview shows a CSV file with columns: id, title, tipo, plazas, anclajes, lastUpdated, geometry, and icon. The data rows represent different bicycle parking locations in Zaragoza.

Asignamos permisos para que pueda escribir en el directorio correspondiente.

```
[cloudera@quickstart ~]$ hdfs dfs -chmod 777 /user/cloudera
[cloudera@quickstart ~]$
```

Comenzamos la creación de la tabla en Impala

Asignamos el fichero donde se encuentran todos los registros. Tendremos que ajustar los parámetros para tener el formato de la tabla correctamente.

The screenshot shows the HUE web interface for creating a table. The 'SOURCE' tab is active, showing the file path '/user/cloudera/aparcamiento-bicicleta(1).csv'. The 'FORMAT' section is configured with 'Field Separator' as comma, 'Record Separator' as 'New line', and 'Quote Character' as 'Double Quote'. The 'PREVIEW' section shows the data from the CSV file, which matches the data seen in the previous screenshot.



Una vez les asignamos los tipos correspondientes a cada columna.

FIELDS				
Name	id	Type	string	https://www.zaragoza...
Name	title	Type	string	IES Pilar Lorengar
Name	tipo	Type	string	Cerrado
Name	plazas	Type	bigint	16
Name	anclajes	Type	bigint	8
Name	lastUpdated	Type	timestamp	2022-05-13T23:06:15Z
Name	geometry	Type	string	-0.8611007934151447,...
Name	icon	Type	string	//www.zaragoza.es/co...

Comprobamos que se ha creado correctamente la tabla.

<

Ejercicio 2

Revisa mediante código los metadatos generados para ver cómo se tratan los datos almacenados en la tabla Impala.

Nos dirigimos a los detalles de la tabla desde la HUE.

DETAILED TABLE INFORMATION		
Database:	default	
Owner:	cloudera	
CreateTime:	Tue Apr 18 08:17:30 PDT 2023	
LastAccessTime:	UNKNOWN	
Protect Mode:	None	
Retention:	0	
Location:	hdfs://quickstart.cloudera:8020/user/hive/warehouse/aparcamiento_bicicleta	
Table Type:	MANAGED_TABLE	
Table Parameters:	skip.header.line.count	1
	transient_lastDdlTime	1681831050
STORAGE INFORMATION		
SerDe Library:	org.apache.hadoop.hive.serde2.lazy.LazySimpleSerDe	
InputFormat:	org.apache.hadoop.mapred.TextInputFormat	
OutputFormat:	org.apache.hadoop.hive ql.io.HiveIgnoreKeyTextOutputFormat	
Compressed:	No	
Num Buckets:	0	
Bucket Columns:	[]	
Sort Columns:	[]	
Storage Desc Params:	field.delim	;
	serialization.format	;

Ejercicio 3

Haz una consulta que devuelva la información referente a un registro concreto.

Modifica el archivo CSV manualmente desde el sistema de archivos HDFS y repite la consulta para verificar que Impala trabaja directamente sobre el archivo de texto

```
SELECT * FROM aparcamiento_bicicleta WHERE title like 'Sin Nombre'
```

Impala							
Add a name... Add a description...							
0s default text ?							
Query History Saved Queries Results (1)							
id	title	tipo	plazas	anclajes	lastupdated	geometry	
1	https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/16 Sin Nombre	Abierto	10	5	2022-05-13 23:06:15	-0.842436388	

Modificamos el fichero manualmente.

File Browser

View file

Last modified: 04/18/2023 4:03 PM

User: cloudera

Group: cloudera

Size: 12.39 KB

Mode: 100644

Home / user / hive / warehouse / aparcamiento_bicicleta / aparcamiento-bicicleta.csv

```
id;title;tipo;plazas;anclajes;lastUpdated;geometry;icon
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/13;IES Pilar
Lorenaga;Cerrado;16;8;2022-05-13T23:06:15Z;-0.8611067934151447,41.66017411919716;https://www.zaragoza.es/contenidos/iconos/aparcabici_recinto.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/14;Colegio Don
Bosco;Abierto;12;6;2022-05-13T23:06:15Z;-0.855383575580339,41.65723005705326;https://www.zaragoza.es/contenidos/iconos/aparcabici_recinto.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/15;Campo de Fútbol
Vadocrrsy;Abierto;16;8;2022-05-13T23:06:15Z;-0.8555865919973428,41.653264768024066;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/16;Sin Nombre;Cerrado;10;5;2022-05-13T23:06:15Z;-
0.8424363884475,41.66017411919716;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/17;Plaza España
S/N;Abierto;6;3;2022-05-13T23:06:15Z;-0.8414630245067362,41.717492620203465;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/18;Plaza España
S/N;Abierto;6;3;2022-05-13T23:06:16Z;-0.8413885712953965,41.71754868796037;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/19;Campo de Fútbol San
Juan;Abierto;10;5;2022-05-13T23:06:15Z;-0.8325554092106907,41.70836142409008;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/20;Centro Cívico esquivolas del
Esquirol;Abierto;11;10;2022-05-13T23:06:25Z;-0.90735948384784,41.6468696944697;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/21;Parque;Abierto;10;5;2022-05-13T23:06:15Z;-
0.8559342293766957,41.656863705166906;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/22;Parque;Abierto;10;5;2022-05-13T23:06:15Z;-
0.8571739277045953,41.655983015141835;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/113;Colegio Público Cortes de
Aragón;Abierto;16;8;2022-05-13T23:06:20Z;-0.8818054810976093,41.66533816154004;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/114;Federaciones Deportivas
Aragonesas;Abierto;6;1;2022-05-13T23:06:19Z;-0.8819725619292345,41.665728424570815;https://www.zaragoza.es/contenidos/iconos/aparcabici.png
https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/116;C/ Río
```

Save Save as

Ejecutamos de nuevo la consulta y vemos que ha cambiado.

Impala

Add a name... Add a description...

0s default text ?

```
1 SELECT * FROM aparcamiento_bicicleta WHERE title like 'Sin Nombre';
```

Query History Saved Queries Results (1)

	id	title	tipo	plazas	anclajes	lastupdated	geometry
1	https://www.zaragoza.es/sede/servicio/urbanismo-infraestructuras/equipamiento/aparcamiento-bicicleta/16	Sin Nombre	Cerrado	10	5	2022-05-13 23:06:15	-0.842436388