

Ejercicios de Hive:

Ejercicio 1:

- Generar una base de datos
- Crear las tablas en Hive con las estructuras correctas. Debemos generar diferentes tipos de tablas (external,default..)

Estructuras de los datos a cargar:

customers

Field	Type	Null	Key	Default	Extra
cust_id	int(11)	NO	PRI	NULL	
fname	varchar(15)	YES		NULL	
lname	varchar(20)	YES		NULL	
address	varchar(40)	YES		NULL	
city	varchar(30)	YES		NULL	
state	char(2)	YES		NULL	
zipcode	char(5)	YES		NULL	

order_details

Field	Type	Null	Key	Default	Extra
order_id	int(11)	NO		NULL	
prod_id	int(11)	NO		NULL	

orders

Field	Type	Null	Key	Default	Extra
-------	------	------	-----	---------	-------

Field	Type	Null	Key	Default	Extra
order_id	int(11)	NO	PRI	NULL	
cust_id	int(11)	NO		NULL	
order_date	datetime	NO		NULL	

products

Field	Type	Null	Key	Default	Extra
prod_id	int(11)	NO	PRI	NULL	
brand	varchar(20)	YES		NULL	
name	varchar(75)	YES		NULL	
price	int(11)	YES		NULL	
cost	int(11)	YES		NULL	
shipping_wt	smallint(6)	YES		NULL	

Ejercicio 2

Una vez generada las tablas, hay que realizar la carga de los datos a través del comando LOAD DATA.

Ejercicio 3

- Resolver las siguientes questiones:

- 1) Contar el número de registros en la tabla customer
- 2) Obtener los 10 estados con el mayor número de clientes
- 3) Top de los 3 productos más vendidos
- 4) Ingresos totales en Mayo de 2013
- 5) Beneficio en Mayo de 2013 (precio de venta menos el costo(cost))