

Pràctica 4: Gestió de Vehicles amb HBase

Autor: Jordi Osarenkhoe

Descripció

En aquesta pràctica treballarem amb HBase per gestionar una taula de vehicles de segona mà. La taula `vehicles` conté informació detallada sobre els vehicles, incloent dades generals, estat de conservació, informació del propietari i ubicació.

Contingut

- 1. Creació de la taula `vehicles` en HBase.
- 2. Inserció de dades a la taula `vehicles` mitjançant ordres `put`.
- 3. Importació de dades des d'un fitxer CSV a HBase.
- 4. Consultes HBase per obtenir dades específiques.
- 5. Modificació i esborrat de dades en HBase.
- 6. Creació d'una taula externa en Hive vinculada a HBase.
- 7. Consultes en HiveQL i Impala per analitzar les dades dels vehicles.

Instruccions

En aquesta tasca anam a treballar amb HBase. Tenim una taula `vehicles` per gestionar un sistema de venda de vehicles de segona mà. La taula té 4 famílies de columnes: `info_general`, `conservacio`, `propietari` i `ubicacio`.

Aquesta taula mostra les dades que ha de contenir la taula `vehicles`:

clau	info_general	conservacio	propietari	ubicacio
	marca	model	matricula	preu
1	Toyota	Corolla	1234ABC	12000
2	Ford	Focus	5678DEF	15500
3	Honda	Civic	9012GHI	18000
4	Mercedes-Benz	A-Class	3456JKL	30000
5	BMW	X5	7890JKL	25000
6	Audi	A4	1122XYZ	20000

Crea la taula `vehicles` en HBase.

```
create 'vehicles', 'info_general', 'conservacio', 'propietari', 'ubicacio'
```

```
base(main):001:0> status
active master, 0 backup masters, 1 servers, 0 dead, 2.0000 average load

base(main):002:0> create 'vehicles', 'info_general', 'conservacio', 'propietari', 'ubicacio'
row(s) in 1.3550 seconds

> Hbase::Table - vehicles
```

Insereix mitjançant ordres put les tres primeres files de la taula.

```
put 'vehicles', '1', 'info_general:marca', 'Toyota'
put 'vehicles', '1', 'info_general:model', 'Corolla'
put 'vehicles', '1', 'info_general:matricula', '1234ABC'
put 'vehicles', '1', 'info_general:preu', '12000'
put 'vehicles', '1', 'conservacio:any', '2015'
put 'vehicles', '1', 'conservacio:km', '120000'
put 'vehicles', '1', 'conservacio:estat', 'bo'
put 'vehicles', '1', 'propietari:nif', '12345678A'
put 'vehicles', '1', 'propietari:nom', 'Joan Martínez'
put 'vehicles', '1', 'propietari:telefon', '623456789'
put 'vehicles', '1', 'propietari:mail', 'joan.martinez@email.com'
put 'vehicles', '1', 'ubicacio:ciutat', 'Palma'
put 'vehicles', '1', 'ubicacio:illa', 'Mallorca'
put 'vehicles', '1', 'ubicacio:longitud', '2.650160'
put 'vehicles', '1', 'ubicacio:latitud', '39.569600'

put 'vehicles', '2', 'info_general:marca', 'Ford'
put 'vehicles', '2', 'info_general:model', 'Focus'
put 'vehicles', '2', 'info_general:matricula', '5678DEF'
put 'vehicles', '2', 'info_general:preu', '15500'
put 'vehicles', '2', 'conservacio:any', '2018'
put 'vehicles', '2', 'conservacio:km', '55000'
put 'vehicles', '2', 'conservacio:estat', 'molt bo'
put 'vehicles', '2', 'propietari:nif', '87654321B'
put 'vehicles', '2', 'propietari:nom', 'Ana García'
put 'vehicles', '2', 'propietari:telefon', ''
put 'vehicles', '2', 'propietari:mail', 'ana.garcia@email.com'
put 'vehicles', '2', 'ubicacio:ciutat', 'Palma'
put 'vehicles', '2', 'ubicacio:illa', 'Mallorca'

put 'vehicles', '3', 'info_general:marca', 'Honda'
put 'vehicles', '3', 'info_general:model', 'Civic'
put 'vehicles', '3', 'info_general:matricula', '9012GHI'
put 'vehicles', '3', 'info_general:preu', '18000'
put 'vehicles', '3', 'conservacio:any', '2020'
put 'vehicles', '3', 'conservacio:km', '20000'
put 'vehicles', '3', 'conservacio:estat', 'bo'
put 'vehicles', '3', 'propietari:nif', '11223344C'
put 'vehicles', '3', 'propietari:nom', 'David López'
put 'vehicles', '3', 'propietari:telefon', '655123456'
put 'vehicles', '3', 'ubicacio:ciutat', 'Eivissa'
put 'vehicles', '3', 'ubicacio:illa', 'Eivissa'
```

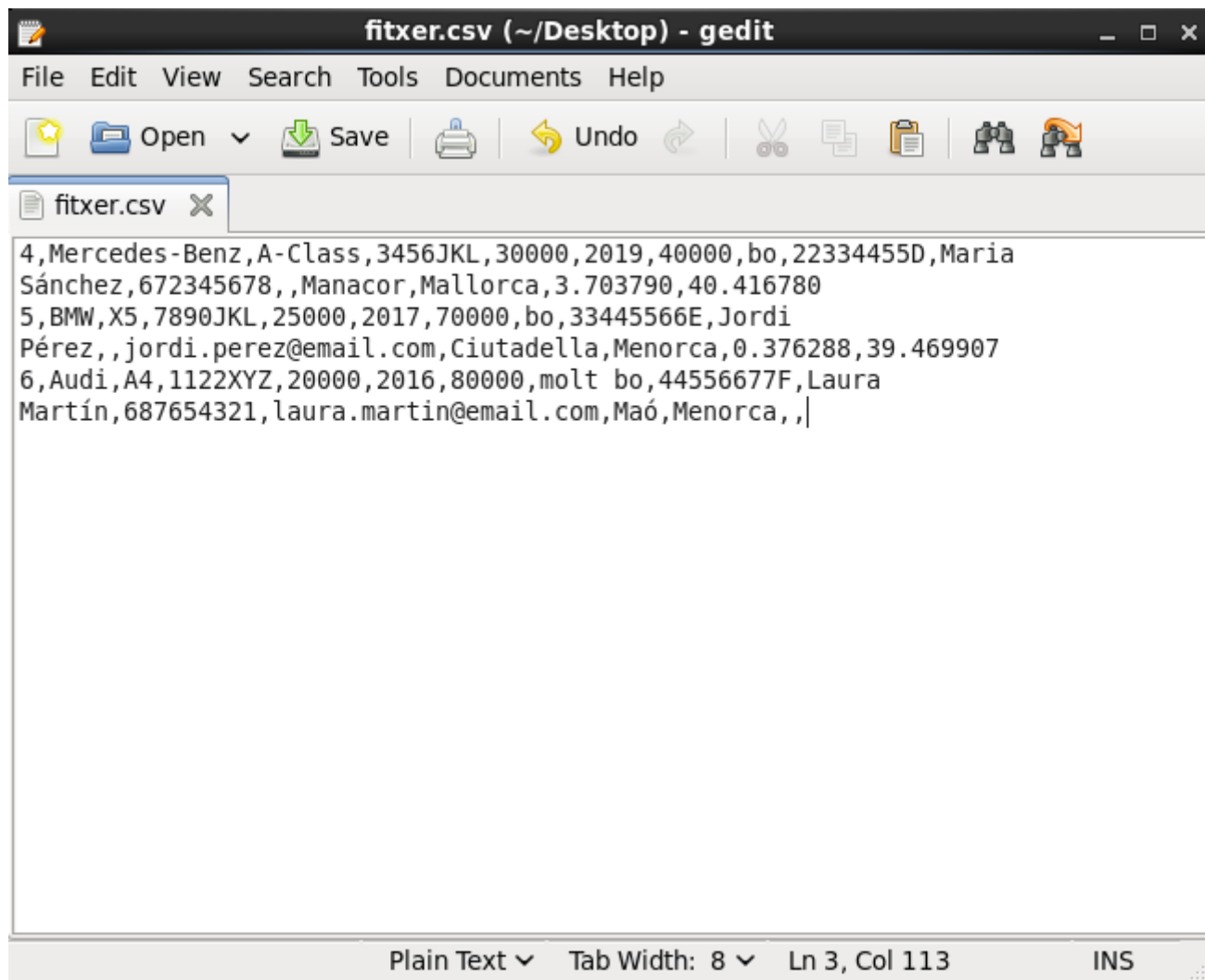
```
put 'vehicles', '3', 'ubicacio:longitud', '1.452850'  
put 'vehicles', '3', 'ubicacio:latitud', '38.916280'
```

```
scan 'vehicles'
```

```
hbase(main):048:0> scan 'vehicles'  
ROW      COLUMN+CELL  
1        column=conservacio:any, timestamp=1736730583119, value=2015  
1        column=conservacio:estat, timestamp=1736730583213, value=bo  
1        column=conservacio:km, timestamp=1736730583160, value=120000  
1        column=info_general:marca, timestamp=1736730582822, value=Toyota  
1        column=info_general:matricula, timestamp=1736730582975, value=1234ABC  
1        column=info_general:model, timestamp=1736730582912, value=Corolla  
1        column=info_general:preu, timestamp=1736730583059, value=12000  
1        column=propietari:mail, timestamp=1736730583349, value=joan.martinez@email.com  
1        column=propietari:nif, timestamp=1736730583256, value=12345678A  
1        column=propietari:nom, timestamp=1736730583286, value=Joan Mart\xC3\xADnez  
1        column=propietari:telefon, timestamp=1736730583319, value=623456789  
1        column=ubicacio:ciutat, timestamp=1736730583383, value=Palma  
1        column=ubicacio:illa, timestamp=1736730583431, value=Mallorca  
1        column=ubicacio:latitud, timestamp=1736730583479, value=39
```

Fes un fitxer CSV amb les 3 darreres files i carrega'l a la taula.

Creem un fitxer vehicles.csv amb les dades de les 3 darreres files.



Copiam el fitxer al HDFS

```
hdfs dfs -put vehicles.csv /user/cloudera
```

```
[cloudera@quickstart Desktop]$ hdfs dfs -put fitxer.csv /user/cloudera
[cloudera@quickstart Desktop]$ █
```

I ho importam amb la següent comanda:

```
hbase org.apache.hadoop.hbase.mapreduce.ImportTsv -Dimporttsv.separator=', ' \
-
Dimporttsv.columns=HBASE_ROW_KEY,info_general:marca,info_general:model,info_general:matricula,info_general:preu,info_general:any,conservacio:km,conservacio:estat,propietari:nif,propietari:nom,propietari:telefon,propietari:mail,ubicacio:ciutat,ubicacio:illa,ubicacio:longitud,ubicacio:latitud \
vehicles \
hdfs:///user/cloudera/fitxer.csv
```

I ho verificam fent un scan:

```
5 column=info_general:marca, timestamp=1736731117682, value=BMW
5 column=info_general:matricula, timestamp=1736731117682, value=7890JKL
5 column=info_general:model, timestamp=1736731117682, value=X5
5 column=info_general:preu, timestamp=1736731117682, value=25000
5 column=propietari:mail, timestamp=1736731117682, value=jordi.perez@email.com
5 column=propietari:nif, timestamp=1736731117682, value=33445566E
5 column=propietari:nom, timestamp=1736731117682, value=Jordi P\xC3\xA9rez
5 column=propietari:telefon, timestamp=1736731117682, value=
5 column=ubicacio:ciutat, timestamp=1736731117682, value=Ciudadella
5 column=ubicacio:illa, timestamp=1736731117682, value=Menorca
5 column=ubicacio:latitud, timestamp=1736731117682, value=39.469907
5 column=ubicacio:longitud, timestamp=1736731117682, value=0.376288
6 column=conservacio:estat, timestamp=1736731117682, value=molt bo
6 column=conservacio:km, timestamp=1736731117682, value=80000
6 column=info_general:any, timestamp=1736731117682, value=2016
6 column=info_general:marca, timestamp=1736731117682, value=Audi
6 column=info_general:matricula, timestamp=1736731117682, value=1122XYZ
6 column=info_general:model, timestamp=1736731117682, value=A4
6 column=info_general:preu, timestamp=1736731117682, value=20000
6 column=propietari:mail, timestamp=1736731117682, value=laura.martin@email.com
6 column=propietari:nif, timestamp=1736731117682, value=44556677F
6 column=propietari:nom, timestamp=1736731117682, value=Laura Mart\xC3\xADn
6 column=propietari:telefon, timestamp=1736731117682, value=687654321
6 column=ubicacio:ciutat, timestamp=1736731117682, value=Ma\xC3\xB3
6 column=ubicacio:illa, timestamp=1736731117682, value=Menorca
6 column=ubicacio:latitud, timestamp=1736731117682, value=
6 column=ubicacio:longitud, timestamp=1736731117682, value=
6 row(s) in 0.5470 seconds
```

Mostra totes les dades de la fila 4 amb un get.

```
get 'vehicles', '4'
```

```
hbase(main):002:0> get 'vehicles', '4'
COLUMN                                CELL
conservacio:estat                     timestamp=1736731117682, value=bo
conservacio:km                        timestamp=1736731117682, value=40000
info_general:any                      timestamp=1736731117682, value=2019
info_general:marca                    timestamp=1736731117682, value=Mercedes-Benz
info_general:matricula                timestamp=1736731117682, value=3456JKL
info_general:model                    timestamp=1736731117682, value=A-Class
info_general:preu                     timestamp=1736731117682, value=30000
propietari:mail                       timestamp=1736731117682, value=
propietari:nif                       timestamp=1736731117682, value=22334455D
propietari:nom                        timestamp=1736731117682, value=Maria S\xC3\xA1nchez
propietari:telefon                    timestamp=1736731117682, value=672345678
ubicacio:ciutat                       timestamp=1736731117682, value=Manacor
ubicacio:illa                         timestamp=1736731117682, value=Mallorca
ubicacio:latitud                      timestamp=1736731117682, value=40.416780
ubicacio:longitud                     timestamp=1736731117682, value=3.703790
15 row(s) in 0.0200 seconds
```

Mostra les dades d'informació general de la taula amb un scan.

```
scan 'vehicles', {COLUMNS => 'info_general'}
```

```
hbase(main):003:0> scan 'vehicles', {COLUMNS => 'info_general'}
ROW          COLUMN+CELL
 1          column=info_general:marca, timestamp=1736730582822, value=Toyota
 1          column=info_general:matricula, timestamp=1736730582975, value=1234ABC
 1          column=info_general:model, timestamp=1736730582912, value=Corolla
 1          column=info_general:preu, timestamp=1736730583059, value=12000
 2          column=info_general:marca, timestamp=1736730583565, value=Ford
 2          column=info_general:matricula, timestamp=1736730583612, value=5678DEF
 2          column=info_general:model, timestamp=1736730583589, value=Focus
 2          column=info_general:preu, timestamp=1736730583637, value=15500
 3          column=info_general:marca, timestamp=1736730583966, value=Honda
 3          column=info_general:matricula, timestamp=1736730584013, value=9012GHI
 3          column=info_general:model, timestamp=1736730583992, value=Civic
 3          column=info_general:preu, timestamp=1736730584066, value=18000
 4          column=info_general:any, timestamp=1736731117682, value=2019
 4          column=info_general:marca, timestamp=1736731117682, value=Mercedes-Benz
 4          column=info_general:matricula, timestamp=1736731117682, value=3456JKL
 4          column=info_general:model, timestamp=1736731117682, value=A-Class
 4          column=info_general:preu, timestamp=1736731117682, value=30000
 5          column=info_general:any, timestamp=1736731117682, value=2017
 5          column=info_general:marca, timestamp=1736731117682, value=BMW
 5          column=info_general:matricula, timestamp=1736731117682, value=7890JKL
 5          column=info_general:model, timestamp=1736731117682, value=X5
 5          column=info_general:preu, timestamp=1736731117682, value=25000
 6          column=info_general:any, timestamp=1736731117682, value=2016
 6          column=info_general:marca, timestamp=1736731117682, value=Audi
 6          column=info_general:matricula, timestamp=1736731117682, value=1122XYZ
 6          column=info_general:model, timestamp=1736731117682, value=A4
 6          column=info_general:preu, timestamp=1736731117682, value=20000
5 row(s) in 0.1030 seconds
```

Mostra totes les dades dels vehicles amb estat bo.

```
scan 'vehicles', {FILTER => "ValueFilter(=, 'binary:bo')"}

```

```
hbase(main):004:0> scan 'vehicles', {FILTER => "ValueFilter(=, 'binary:bo')"}
ROW          COLUMN+CELL
 1          column=conservacio:estat, timestamp=1736730583213, value=bo
 3          column=conservacio:estat, timestamp=1736730584141, value=bo
 4          column=conservacio:estat, timestamp=1736731117682, value=bo
 5          column=conservacio:estat, timestamp=1736731117682, value=bo
4 row(s) in 0.1150 seconds
```

Mostra les dades dels propietaris dels vehicles de Mallorca.

```
scan 'vehicles', {FILTER => "ValueFilter(=, 'substring:Mallorca')"}

```

```
hbase(main):005:0> scan 'vehicles', {FILTER => "ValueFilter(=, 'substring:Mallorca')"}
ROW          COLUMN+CELL
 1          column=ubicacio:illa, timestamp=1736730583431, value=Mallorca
 2          column=ubicacio:illa, timestamp=1736730583928, value=Mallorca
 4          column=ubicacio:illa, timestamp=1736731117682, value=Mallorca
3 row(s) in 0.0450 seconds
```

Mostra la informació general dels vehicles que siguin posteriors al 2017 (inclòs) amb menys de 50000 km.

```
scan 'vehicles', {FILTER => "SingleColumnValueFilter('info_general', 'any', >=, 'binary:2017') AND SingleColumnValueFilter('conservacio', 'km', <, 'binary:50000')"}

```



```

hbase(main):014:0> scan 'vehicles', {FILTER => "SingleColumnValueFilter('info_general', 'any', >=, 'binary:2017') AND SingleColumnValueFilter('conservacio', 'km', <, 'binary:50000')"}
COLUMN+CELL
1 column=conservacio:any, timestamp=1736730583119, value=2015
1 column=conservacio:estat, timestamp=1736730583213, value=bo
1 column=conservacio:km, timestamp=1736730583160, value=120000
1 column=info_general:marca, timestamp=1736730582822, value=Toyota
1 column=info_general:matricula, timestamp=1736730582975, value=1234ABC
1 column=info_general:model, timestamp=1736730582912, value=Corolla
1 column=info_general:preu, timestamp=1736730583059, value=12000
1 column=propietari:mail, timestamp=1736730583349, value=joan.martinez@email.com
1 column=propietari:nif, timestamp=1736730583256, value=12345678A
1 column=propietari:nom, timestamp=1736730583286, value=Joan Mart\xC3\xADnez
1 column=propietari:telefon, timestamp=1736730583319, value=623456789
1 column=ubicacio:ciutat, timestamp=1736730583383, value=Palma
1 column=ubicacio:illa, timestamp=1736730583431, value=Mallorca
1 column=ubicacio:latitud, timestamp=1736730583479, value=39.569600
1 column=ubicacio:longitud, timestamp=1736730583461, value=2.650160
3 column=conservacio:any, timestamp=1736730584102, value=2020
3 column=conservacio:estat, timestamp=1736730584141, value=bo
3 column=conservacio:km, timestamp=1736730584123, value=20000
3 column=info_general:marca, timestamp=1736730583966, value=Honda
3 column=info_general:matricula, timestamp=1736730584013, value=9012GHI
3 column=info_general:model, timestamp=1736730583992, value=Civic
3 column=info_general:preu, timestamp=1736730584066, value=18000
3 column=propietari:nif, timestamp=1736730584168, value=11223344C
3 column=propietari:nom, timestamp=1736730584194, value=David L\xC3\xB3pez
3 column=propietari:telefon, timestamp=1736730584225, value=655123456
3 column=ubicacio:ciutat, timestamp=1736730584239, value=Eivissa
3 column=ubicacio:illa, timestamp=1736730584261, value=Eivissa
3 column=ubicacio:latitud, timestamp=1736730584295, value=38.916200
3 column=ubicacio:longitud, timestamp=1736730584274, value=1.452850
4 column=conservacio:estat, timestamp=1736731117682, value=bo
4 column=conservacio:km, timestamp=1736731117682, value=40000
4 column=info_general:any, timestamp=1736731117682, value=2019
4 column=info_general:marca, timestamp=1736731117682, value=Mercedes-Benz
4 column=info_general:matricula, timestamp=1736731117682, value=3456JKL
4 column=info_general:model, timestamp=1736731117682, value=A-Class
4 column=info_general:preu, timestamp=1736731117682, value=30000
4 column=propietari:mail, timestamp=1736731117682, value=
4 column=propietari:nif, timestamp=1736731117682, value=22334455D
4 column=propietari:nom, timestamp=1736731117682, value=Maria S\xC3\xA1nchez
4 column=propietari:telefon, timestamp=1736731117682, value=672345678
4 column=ubicacio:ciutat, timestamp=1736731117682, value=Manacor
4 column=ubicacio:illa, timestamp=1736731117682, value=Mallorca
4 column=ubicacio:latitud, timestamp=1736731117682, value=40.416780
4 column=ubicacio:longitud, timestamp=1736731117682, value=3.703790
3 row(s) in 0.0810 seconds

```

Modifica la fila 2 i posa-li el telèfon 656123321

```
put 'vehicles', '2', 'propietari:telefon', '656123321'
```

Escaneja la taula per comprovar que s'ha fet el canvi.

```
scan 'vehicles'
```

```

2 column=conservacio:any, timestamp=1736730583674, value=2018
2 column=conservacio:estat, timestamp=1736730583713, value=molt bo
2 column=conservacio:km, timestamp=1736730583700, value=55000
2 column=info_general:marca, timestamp=1736730583565, value=Ford
2 column=info_general:matricula, timestamp=1736730583612, value=5678DEF
2 column=info_general:model, timestamp=1736730583589, value=Focus
2 column=info_general:preu, timestamp=1736730583637, value=15500
2 column=propietari:mail, timestamp=1736730583842, value=ana.garcia@email.com
2 column=propietari:nif, timestamp=1736730583735, value=87654321B
2 column=propietari:nom, timestamp=1736730583751, value=Ana Garc\xC3\xA0a
2 column=propietari:telefon, timestamp=1736731464646, value=656123321
2 column=ubicacio:ciutat, timestamp=1736730583895, value=Palma
2 column=ubicacio:illa, timestamp=1736730583928, value=Mallorca

```

Esborra totes les dades la fila 5.

```
deleteall 'vehicles', '5'
```



```
hbase(main):013:0> scan 'vehicles'
```

```
ROW
1      column=conservacio:any, timestamp=1736730583119, value=2015
1      column=conservacio:estat, timestamp=1736730583213, value=bo
1      column=conservacio:km, timestamp=1736730583160, value=120000
1      column=info_general:marca, timestamp=1736730582822, value=Toyota
1      column=info_general:matricula, timestamp=1736730582975, value=1234ABC
1      column=info_general:model, timestamp=1736730582912, value=Corolla
1      column=info_general:preu, timestamp=1736730583059, value=12000
1      column=propietari:mail, timestamp=1736730583349, value=joan.martinez@email.com
1      column=propietari:nif, timestamp=1736730583256, value=12345678A
1      column=propietari:nom, timestamp=1736730583286, value=Joan Mart\xC3\xADnez
1      column=propietari:telefon, timestamp=1736730583319, value=623456789
1      column=ubicacio:ciutat, timestamp=1736730583383, value=Palma
1      column=ubicacio:illa, timestamp=1736730583431, value=Mallorca
1      column=ubicacio:latitud, timestamp=1736730583479, value=39.569600
1      column=ubicacio:longitud, timestamp=1736730583461, value=2.650160
2      column=conservacio:any, timestamp=1736730583674, value=2018
2      column=conservacio:estat, timestamp=1736730583713, value=molt bo
2      column=conservacio:km, timestamp=1736730583700, value=55000
2      column=info_general:marca, timestamp=1736730583565, value=Ford
2      column=info_general:matricula, timestamp=1736730583612, value=5678DEF
2      column=info_general:model, timestamp=1736730583589, value=Focus
2      column=info_general:preu, timestamp=1736730583637, value=15500
2      column=propietari:mail, timestamp=1736730583842, value=ana.garcia@email.com
2      column=propietari:nif, timestamp=1736730583735, value=87654321B
2      column=propietari:nom, timestamp=1736730583751, value=Ana Garc\xC3\xADa
2      column=propietari:telefon, timestamp=1736731464646, value=656123321
2      column=ubicacio:ciutat, timestamp=1736730583895, value=Palma
2      column=ubicacio:illa, timestamp=1736730583928, value=Mallorca
3      column=conservacio:any, timestamp=1736730584102, value=2020
3      column=conservacio:estat, timestamp=1736730584141, value=bo
3      column=conservacio:km, timestamp=1736730584123, value=20000
3      column=info_general:marca, timestamp=1736730583966, value=Honda
3      column=info_general:matricula, timestamp=1736730584013, value=9012GHI
3      column=info_general:model, timestamp=1736730583992, value=Civic
3      column=info_general:preu, timestamp=1736730584066, value=18000
3      column=propietari:nif, timestamp=1736730584168, value=11223344C
3      column=propietari:nom, timestamp=1736730584194, value=David L\xC3\xB3pez
3      column=propietari:telefon, timestamp=1736730584225, value=655123456
3      column=ubicacio:ciutat, timestamp=1736730584239, value=Eivissa
3      column=ubicacio:illa, timestamp=1736730584261, value=Eivissa
3      column=ubicacio:latitud, timestamp=1736730584295, value=38.916280
3      column=ubicacio:longitud, timestamp=1736730584274, value=1.452850
4      column=conservacio:estat, timestamp=1736731117682, value=bo
4      column=conservacio:km, timestamp=1736731117682, value=40000
4      column=info_general:any, timestamp=1736731117682, value=2019
4      column=info_general:marca, timestamp=1736731117682, value=Mercedes-Benz
4      column=info_general:matricula, timestamp=1736731117682, value=3456JKL
4      column=info_general:model, timestamp=1736731117682, value=A-Class
4      column=info_general:preu, timestamp=1736731117682, value=30000
4      column=propietari:mail, timestamp=1736731117682, value=
4      column=propietari:nif, timestamp=1736731117682, value=22334455D
4      column=propietari:nom, timestamp=1736731117682, value=Maria S\xC3\xA1nchez
4      column=propietari:telefon, timestamp=1736731117682, value=672345678
4      column=ubicacio:ciutat, timestamp=1736731117682, value=Manacor
4      column=ubicacio:illa, timestamp=1736731117682, value=Mallorca
4      column=ubicacio:latitud, timestamp=1736731117682, value=40.416780
4      column=ubicacio:longitud, timestamp=1736731117682, value=3.703790
4      column=conservacio:km, timestamp=1736731117682, value=40000
4      column=info_general:any, timestamp=1736731117682, value=2019
4      column=info_general:marca, timestamp=1736731117682, value=Mercedes-Benz
4      column=info_general:matricula, timestamp=1736731117682, value=3456JKL
4      column=info_general:model, timestamp=1736731117682, value=A-Class
4      column=info_general:preu, timestamp=1736731117682, value=30000
4      column=propietari:mail, timestamp=1736731117682, value=
4      column=propietari:nif, timestamp=1736731117682, value=22334455D
4      column=propietari:nom, timestamp=1736731117682, value=Maria S\xC3\xA1nchez
4      column=propietari:telefon, timestamp=1736731117682, value=672345678
4      column=ubicacio:ciutat, timestamp=1736731117682, value=Manacor
4      column=ubicacio:illa, timestamp=1736731117682, value=Mallorca
4      column=ubicacio:latitud, timestamp=1736731117682, value=40.416780
4      column=ubicacio:longitud, timestamp=1736731117682, value=3.703790
6      column=conservacio:estat, timestamp=1736731117682, value=molt bo
6      column=conservacio:km, timestamp=1736731117682, value=80000
6      column=info_general:any, timestamp=1736731117682, value=2016
6      column=info_general:marca, timestamp=1736731117682, value=Audi
6      column=info_general:matricula, timestamp=1736731117682, value=1122XYZ
6      column=info_general:model, timestamp=1736731117682, value=A4
6      column=info_general:preu, timestamp=1736731117682, value=20000
6      column=propietari:mail, timestamp=1736731117682, value=laura.martin@email.com
6      column=propietari:nif, timestamp=1736731117682, value=44556677F
6      column=propietari:nom, timestamp=1736731117682, value=Laura Mart\xC3\xADn
6      column=propietari:telefon, timestamp=1736731117682, value=687654321
6      column=ubicacio:ciutat, timestamp=1736731117682, value=Ma\xC3\xB3
6      column=ubicacio:illa, timestamp=1736731117682, value=Menorca
6      column=ubicacio:latitud, timestamp=1736731117682, value=
6      column=ubicacio:longitud, timestamp=1736731117682, value=
```

Crea una taula externa en Hive vinculada a la taula vehicles d'HBase.

```
CREATE EXTERNAL TABLE vehicles (  
  id string,  
  marca string,  
  model string,  
  matricula string,  
  preu int,  
  any int,  
  km int,  
  estat string,  
  nif string,  
  nom string,  
  telefon string,  
  mail string,  
  ciutat string,  
  illa string,  
  longitud double,  
  latitud double  
)  
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'  
WITH SERDEPROPERTIES (  
  "hbase.columns.mapping" = ":key,  
  info_general:marca,  
  info_general:model,  
  info_general:matricula,  
  info_general:preu,  
  info_general:any,  
  conservacio:km,  
  conservacio:estat,  
  propietari:nif,  
  propietari:nom,  
  propietari:telefon,  
  propietari:mail,  
  ubicacio:ciutat,  
  ubicacio:illa,  
  ubicacio:longitud,  
  ubicacio:latitud"  
)  
TBLPROPERTIES("hbase.table.name" = "vehicles");
```

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

19.17s
default
text

```

25     info_general:preu,
26     info_general:any,
27     conservacio:km,
28     conservacio:estat,
29     propietari:nif,
30     propietari:nom,
31     propietari:telefon,
32     propietari:mail,
33     ubicacio:ciutat,
34     ubicacio:illa,
35     ubicacio:longitud,
36     ubicacio:latitud"
37 )
38 TBLPROPERTIES("hbase.table.name" = "vehicles");

```

Success.

Query History
Saved Queries

a few seconds ago

```

CREATE EXTERNAL TABLE vehicles ( id string, marca string,
model string, matricula string, preu int, any int, km int,
estat string, nif string, nom string, telefon string, mail string,
ciutat string, illa string, longitud double, latitud double )
STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
WITH SERDEPROPERTIES ( "hbase.columns.mapping" = ":key,
info_general:marca, info_general:model, info_general:matricula,
info_general:preu, info_general:any, conservacio:km,
conservacio:estat, propietari:nif, propietari:nom,
propietari:telefon, propietari:mail, ubicacio:ciutat, ubicacio:illa,
ubicacio:longitud, ubicacio:latitud" )
TBLPROPERTIES("hbase.table.name" = "vehicles")

```

Fes una consulta en HiveQL que retorni les dades d'informació general dels vehicles de Palma.

SELECT * FROM vehicles WHERE ciutat = 'Palma';

1|SELECT * FROM vehicles WHERE ciutat = 'Palma';

Query History
Saved Queries
Results (2)

	vehicles.id	vehicles.marca	vehicles.model	vehicles.matricula	vehicles.preu	vehicle
1	1	Toyota	Corolla	1234ABC	12000	NULL
2	2	Ford	Focus	5678DEF	15500	NULL

Fes una consulta en Impala que retorni totes les dades dels 3 vehicles més cars.

Impala

↺

Add a name...

Add a description...

💾

📄

⚙️

📁

0s default text

1

SELECT * FROM vehicles ORDER BY preu DESC LIMIT 3;

▶

📖

Query History

Saved Queries

Results (3)

	id	estat	km	any	marca	matricula	model	preu	mail
1	4	bo	40000	2019	Mercedes-Benz	3456JKL	A-Class	30000	
2	6	molt bo	80000	2016	Audi	1122XYZ	A4	20000	laura.martin@email.co
3	3	bo	20000	NULL	Honda	9012GHI	Civic	18000	NULL

Fes una consulta en HiveQL que agrupi els vehicles per estat i mostri la mitjana de km de cada estat.

Hive

↺

Add a name...

Add a description...

💾

📄

⚙️

📁

2m, 1s default text

1

SELECT estat, AVG(km) FROM vehicles GROUP BY estat;

▶

📖

Query History

Saved Queries

Results (2)

	estat	_c1
1	bo	60000
2	molt bo	67500

Fes una consulta en Impala que agrupi els vehicles per illa i mostri el nombre de vehicles de cada illa.

SELECT illa, COUNT(*) FROM vehicles GROUP BY illa;

Impala

Add a name...

Add a description...

1.36s

default

text

1

SELECT illa, COUNT(*) FROM vehicles GROUP BY illa;

Query History

Saved Queries

Results (3)

	illa	count(*)
1	Menorca	1
2	Mallorca	3
3	Eivissa	1

Fes un document de text on, per a cada un dels 15 apartats, aparegui l'ordre corresponent. En el cas de l'apartat 3, si fas servir l'eina gràfica per fer la importació, no cal. Fes un document amb Word, Write o Docs on s'inclogui almenys una captura de pantalla que mostri l'execució de cada apartat. Exporta'l a PDF quan hakis acabat. A continuació, fes un fitxer zip (o tar o tar.gz) que contengui els dos documents anteriors, a més del fitxer CSV que has carregat a l'apartat 3. El nom del fitxer ha de ser de la forma tasca_4_Nom_Llinatge.zip. Per exemple tasca_4_Toni_Navarrete.zip. No posis accents ni caràcters no anglesos (ç, ñ, etc.) i si tens espais en blanc, substitueix-los per guions baixos. Finalment, envia el teu fitxer comprimit (amb el botó "Afegeix la tramesa").