

Report Canonaco Martina 01-test

mercoledì 9 novembre 2022 16:06

Task 2.

Creazione e utilizzo del database.

```
create database companybigdata2022;  
show databases; //per vedere se effettivamente è stato creato  
use companybigdata2022;
```

Creazione delle tabelle City, Entity, Language.

Tabelle con attributi di tipo complesso da json format.

```
create table if not exists employee(name string, surname string, age int, livingplace  
struct<name:string, province:string>, programmingskills array<struct<name:string,  
description:string, conflevel:int>>) row format row format serde  
'org.apache.hive.hcatalog.data.JsonSerDe' stored as textfile;
```

Load data into the table.

```
load data inpath 'user/martina/employees.json' overwrite into table employee
```

Per creare lo schema bisogna estrapolare gli oggetti complessi City e Language a partire dagli attributi complessi livingplace e programmingskills.

Per creare la tabella City:

```
create if not exists city(name string, province string) as select distinct(employee.livingplace) from  
employee
```

Per creare la tabella Language:

```
create if not exists language(name string, description string) as select distinct  
(employee.programmingskills) from employee
```

Per creare la tabella Skills:

```
create view skill as select distinct(employee.programmingskills.name) as languagename from  
employee
```