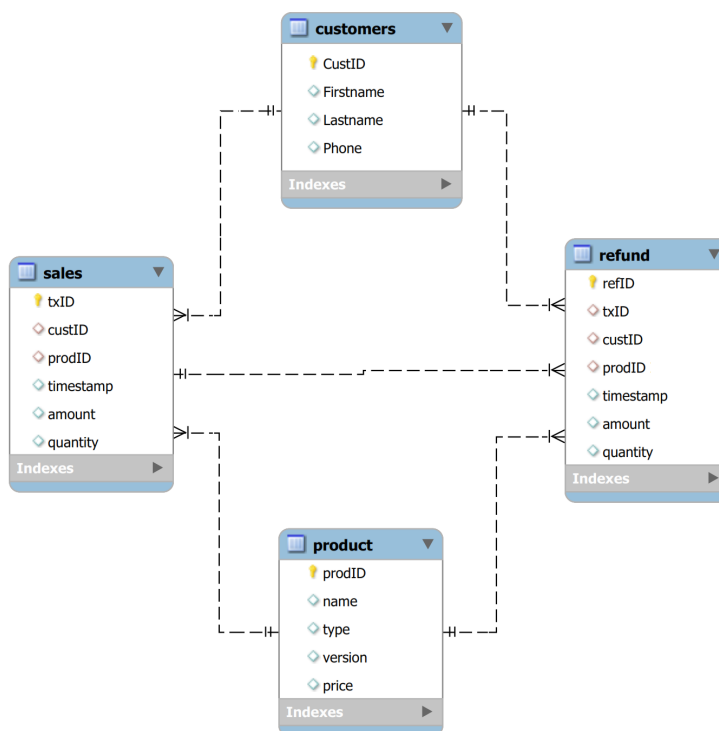


Spark using Scala

Figure bellow, present a data warehouse for a company, we would like to make some analysis about (Sales, Customers, products ..)



Spark Resilient Distributed Dataset

Create a 3 projects using sbt (scala build tools) :

- **Sales Distribution** : evaluate the amount and quantity for each product name.
- **Sales 2013** : total amount of sales.
- **Sales 2013** : total amount of sales, except refund(remboursés).
- **User Products** : total number of products for each customer.
- Verify the generated DAG

A demonstration must be presented, **20 days after receiving this work.**