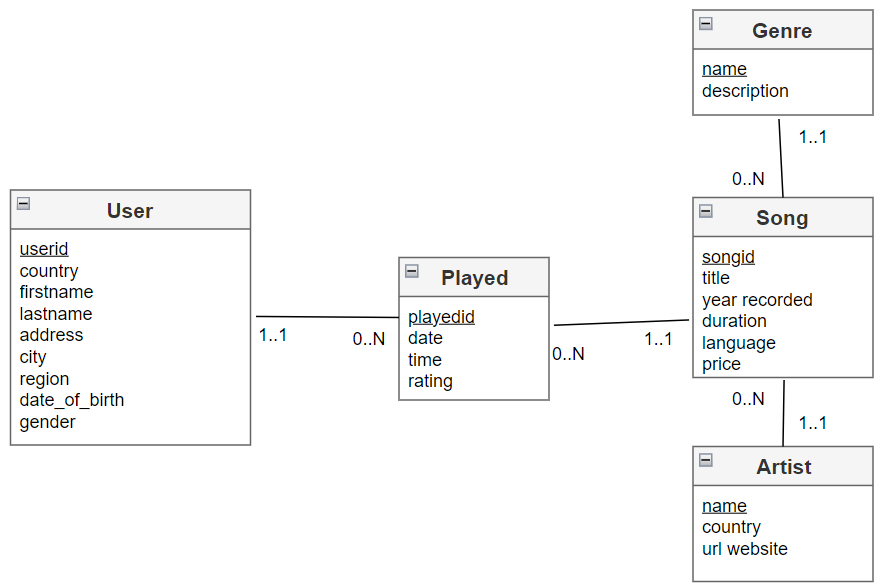
**Exercise OLTP Schema à Star schema**

Below a small ERD is shown of the OLTP database of a music streaming company.



The manager decides to build a DWH.

1. What are the potential benefits of building a DWH? This costs extra money and takes a lot of effort, why would a manager still decide on building a DWH?

If you need to create a DWH, you need to draw a starschema, according to the theory of Kimball.

A starschema consists of a fact table and dimensions.

The fact table is the table where quantities – that make a process measurable – are stored. Turnover, margin, number of units sold and costs are all measurable variables that say something about the sales process. It is the numbers that have made the process measurable. Or to put it another way, they are the facts.

Dimension tables are the tables that contain the context that gives meaning to the facts.

1. Draw a star schema of the DWH to be developed

* Almost every DWH has a Date Dimension.
  + Give some examples that illustrate why a Date Dimension is useful.
  + How does a Date Dimension looks like?
  + What are columns in a Date Dimension
* Next to the Date Dimension, we will also need a Time Dimension.
  + Why?
  + How does a Time Dimension looks like?
* Which other dimensions are you going to create?
* Are you going to save the entire address of the User in the DWH? Why or why not?
* Which dimensions will be Slowly Changing Dimensions.
  + Which fields are you going to add to implement an SCD Type 2? What is the meaning of an SCD?

1. What are the facts? How is the Fact table going to look like?

**Datawarehouse: Use case**

Below you find (part of) the receipt from a sale in supermarket Carrefour. The management wants to make long-term reports to (amongst others):

* Determine trends in the sales of specific product categories (like “Cheese”) per branch over the last 10 years.
* Calculate the total and average (per customer) allowed discount during a promotional campaign (e.g. 2+1 on certain products during one week).

**Afbeelding met tekst, menu, document, Lettertype

Automatisch gegenereerde beschrijving**

**Answer the following questions**

1. What is the ideal fact table? What does a single row in the fact table represent? What are the facts? What are the foreign keys?
2. Determine the most important dimension tables.
3. The daily number of customers in Carrefour Belgium is approximately 700.000. What’s the minimum size (in number of rows) of the fact table to fulfill the above requirements if you consider the above receipt as average.
4. Design the fact table (all fields + data types + other specifications)
5. Which fields in the fact table you designed should have an index?