**Homework Assignment #1**

**Data Management for Data Science**

**Student**

**Vigèr Durand AZIMEDEM TSAFACK (1792126)**

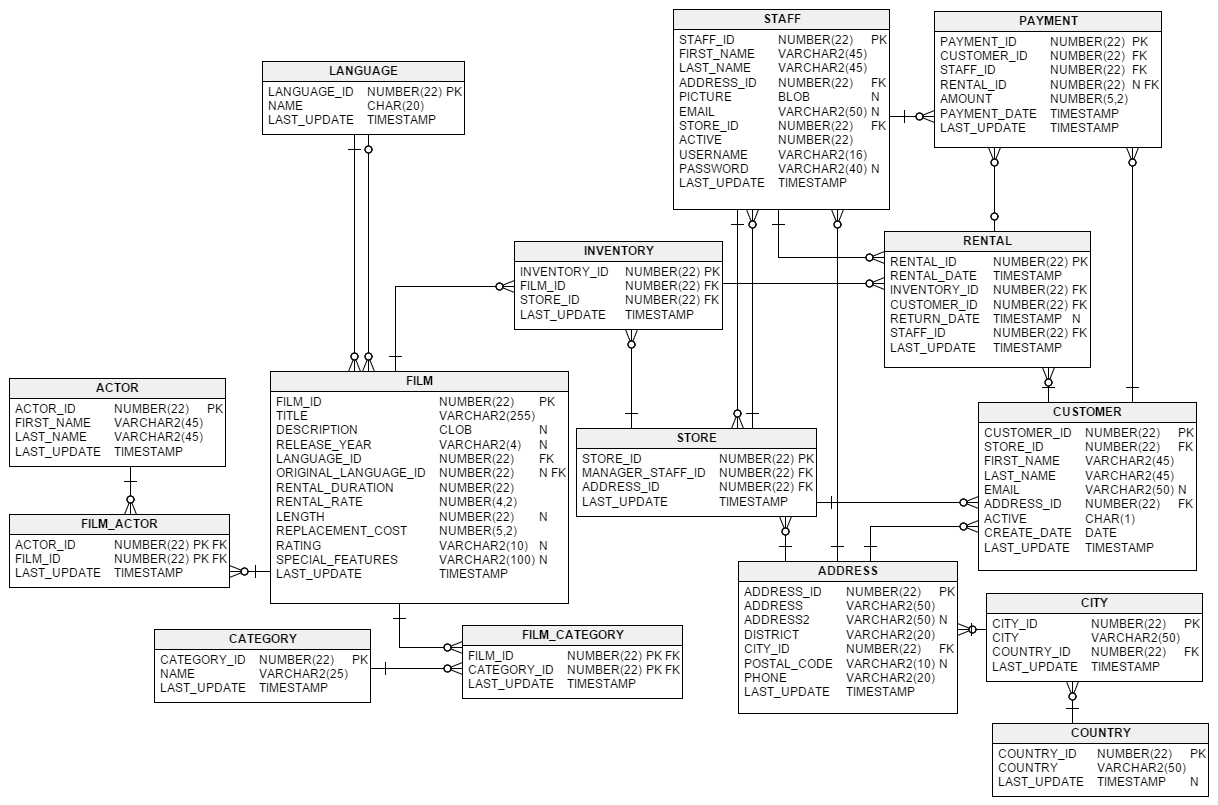
**Data**

We’ll use the **Sakila** schema, which can be found on the following link. <http://dev.mysql.com/doc/index-other.html> ("sakila database")

The MySQL's Sample Salika (DVD Rental) Database is a complex database with 16 tables. It also illustrates features such as Views, Stored Procedures and Triggers.

We will dig into a deep understanding of this database and then we will formulate (and solve) a set of 10 SQL queries on it. We are dealing with a nicely normalized schema modelling a DVD rental store, featuring things like films, actors, film-actor relationships, and a central inventory table that connects films, stores, and rentals.

Follows a graph showing relationships between all the tables of the database:



**Queries**

1. What are the names of all the languages in the database (sorted alphabetically)?
2. Return the full names (first and last) of actors with “SON” in their last name, ordered by their first name.
3. Find all the addresses where the second address is not empty (i.e., contains some text), and return these second addresses sorted.
4. Return the first and last names of actors who played in a film involving a “Crocodile” and a “Shark”, along with the release year of the movie, sorted by the actors’ last names.
5. How many films involve a “Crocodile” and a “Shark”?
6. What is the average running time of films by category?
7. Which actor has appeared in the most films?
8. When is ‘Academy Dinosaur’ due?
9. Find all the film categories in which there are between 55 and 65 films. Return the names of these categories and the number of films per category, sorted by the number of films.
10. In how many film categories is the average difference between the film replacement cost and the rental rate larger than 17?