

Bases de datos relacionales

Cómo administrar bases de datos en R

Denisse Fierro Arcos

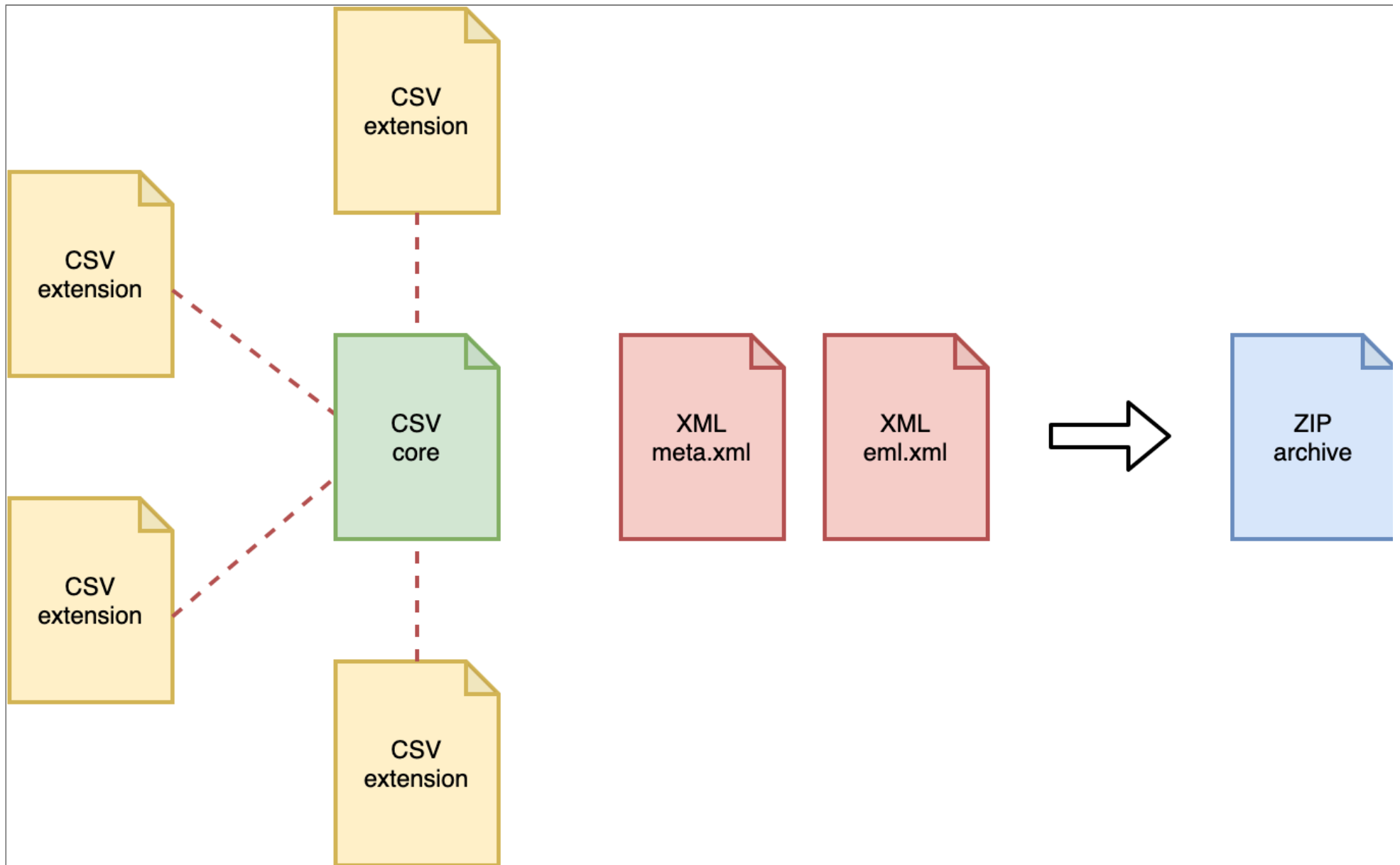
2024-07-24

Gestión de bases de datos relacionales

Esta presentación está basada en el capítulo 5 (Gestión de bases de datos relacionales) del libro “Fundamentos de ciencia de datos con R” de Gema Fernández-Avilés y José-María Montero. Puedes acceder a este libro en el siguiente enlace: <https://cdr-book.github.io/datos-sql.html>.

¿Qué es una base de datos?

Es una colección de datos organizados de tal manera que se puedan acceder, gestionar y actualizar de forma rápida y eficiente. En una base de datos relacional, los datos se estructuran en tablas que están vinculadas entre ellas por diferentes atributos.



Estructura de Darwin Core Archive

SQL y las bases de datos

SQL se refiere al lenguaje estructurado de consulta (Structured Query Language) y es un lenguaje de programación especializado en la gestión de bases de datos relacionales.

SQL es un lenguaje declarativo, lo que significa que el usuario le dice a la base de datos qué hacer, pero no cómo hacerlo.

Existen varias versiones de SQL, tanto propietarias (por ej., MySQL, PostgreSQL, SQLite), así como versiones de acceso abierto (por ej., MariaDB).

SQL en R

Existe un paquete en R llamado **RSQLite** que nos permite conectarnos a bases de datos MySQL. También usaremos el paquete **DBI** para conectarnos a bases de datos relacionales.

Para instalar estos paquetes, debes ejecutar los siguientes comandos:

```
1 install.packages("RSQLite")  
2 install.packages("DBI")
```

Recuerda que solo es necesario instalar este paquete una sola vez. Nota también que debes tener instalado SQL en tu computadora para poder usar estos paquetes.

Hora de trabajar con bases de datos en R

Puedes seguir el código en este repositorio:

https://github.com/lidefi87/bases_datos_RLadies.