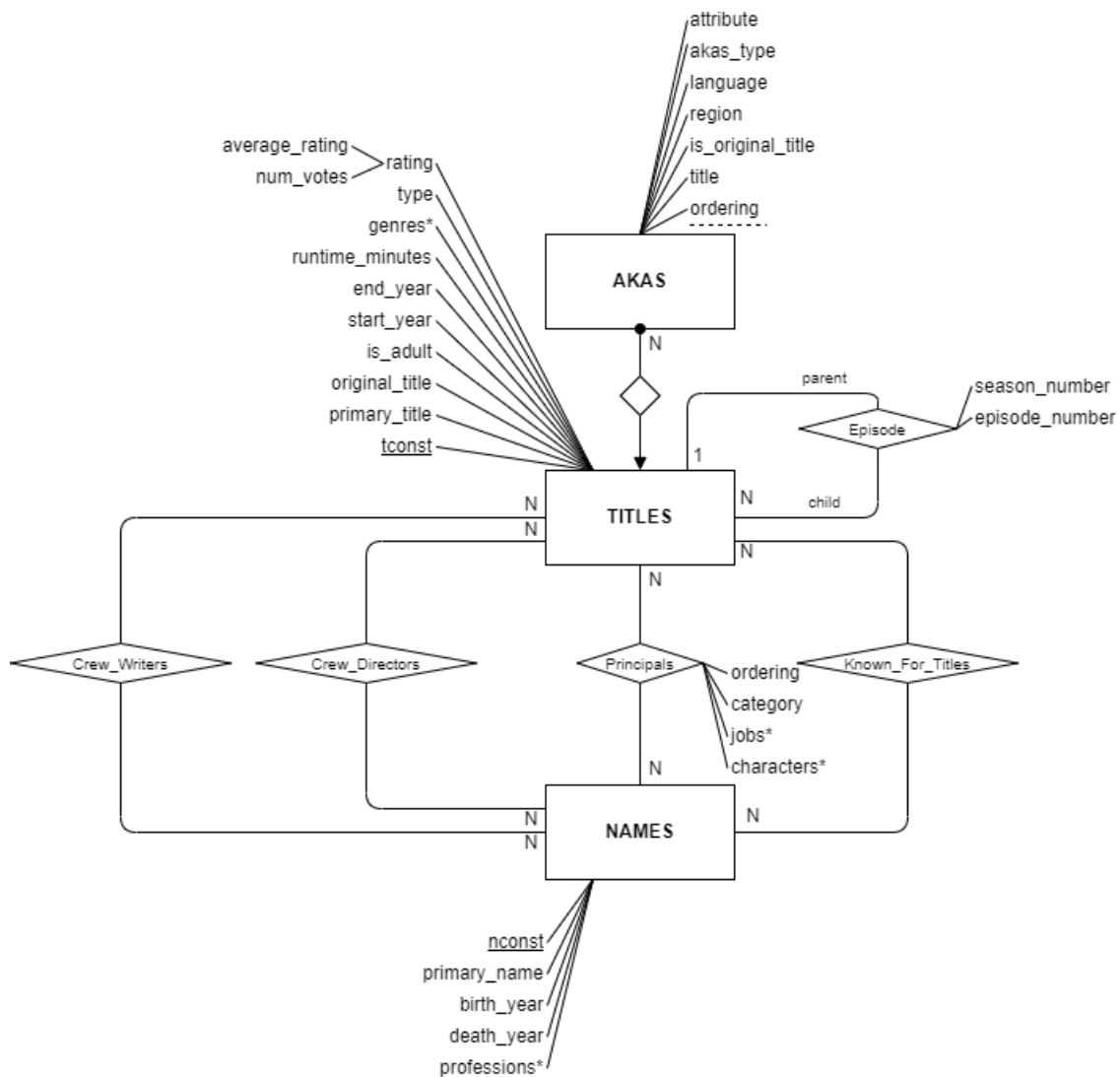


Tarea 1: Modelado de datos y modelo relacional

La IMDb (Internet Movie Database) es una base de datos en línea que almacena información relacionada con títulos (películas, series y contenidos digitales), y de sus correspondientes protagonistas, actores, directores, escritores y el resto del personal del equipo de producción.

Con el objetivo de almacenar en una base de datos la información que publica el IMDb¹ se generó el siguiente modelo conceptual (de entidad-relación):



Del análisis de los datos no se derivaron restricciones no estructurales que complementen el modelo anterior.

¹ <https://www.imdb.com/interfaces/>

Parte 1 - Diseño de base de datos

Partiendo del modelo de entidad relación anterior se diseñó una base de datos relacional en PostgreSQL. Para esta tarea se proveen los siguientes archivos:

- `mpgvd_2022_postgres_ddl.sql`: contiene las sentencias necesarias para la creación de la base de datos
- `mpgvd_2022_postgres_data.zip`: contiene los datos que deben cargarse en las distintas tablas de la base de datos

Describe el mapeo que permitió obtener como resultado la base de datos en PostgreSQL partiendo del modelo de entidad relación presentado en la sección anterior.

Parte 2 - Consultas

Diseñar las consultas que permitan obtener:

1. los títulos en los que 'Jim Carrey' tuvo un rol protagónico (principal) y los personajes que interpretó. El resultado debe contener para cada título (`primary_title`, `character`).
2. para cada título la información de sus episodios (`episodes`) relacionados directa o indirectamente (descendientes). El resultado debe contener para cada episodio (`parent_primary_title`, `depth`, `primary_title`).

El valor de `depth` indica la cantidad de niveles de separación entre ambos títulos, iniciando en 0 si están directamente relacionados.

3. la cantidad de personajes principales y el título primario de cada película ('movie') dirigida por 'Alfred Hitchcock'. El resultado debe contener para cada título (`primary_title`, `count_characters`).

Se debe entregar:

- código SQL de cada una de las consultas
- si creó índices en la base de datos para resolver de forma más eficiente las consultas, incluir sus respectivos scripts de creación y justificar esta decisión

Parte 3 - Modelado

¿Identifica ventajas y desventajas al modelar esta realidad en una base de datos relacional? Justifique su respuesta.

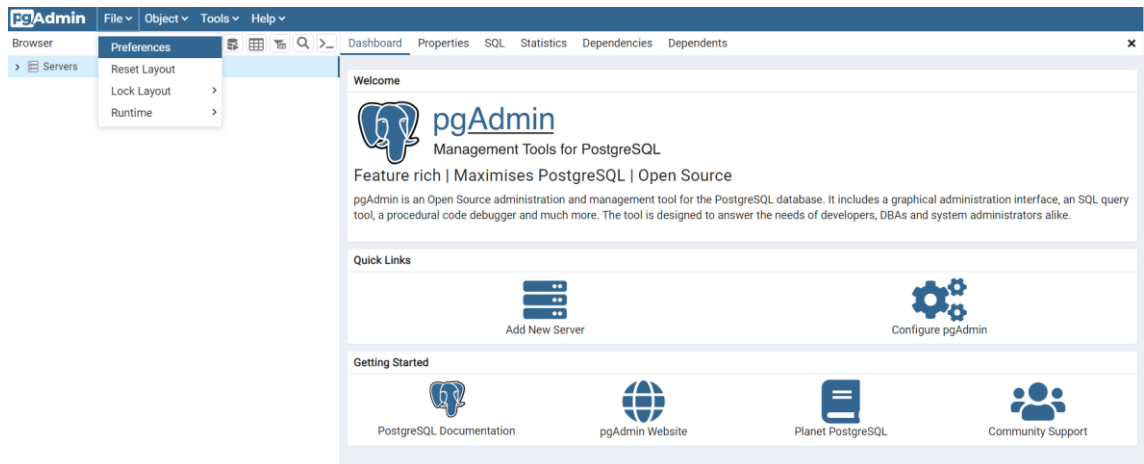
Entregable

Se deberá entregar un informe en formato PDF que contenga lo solicitado en las partes descritas en las secciones anteriores.

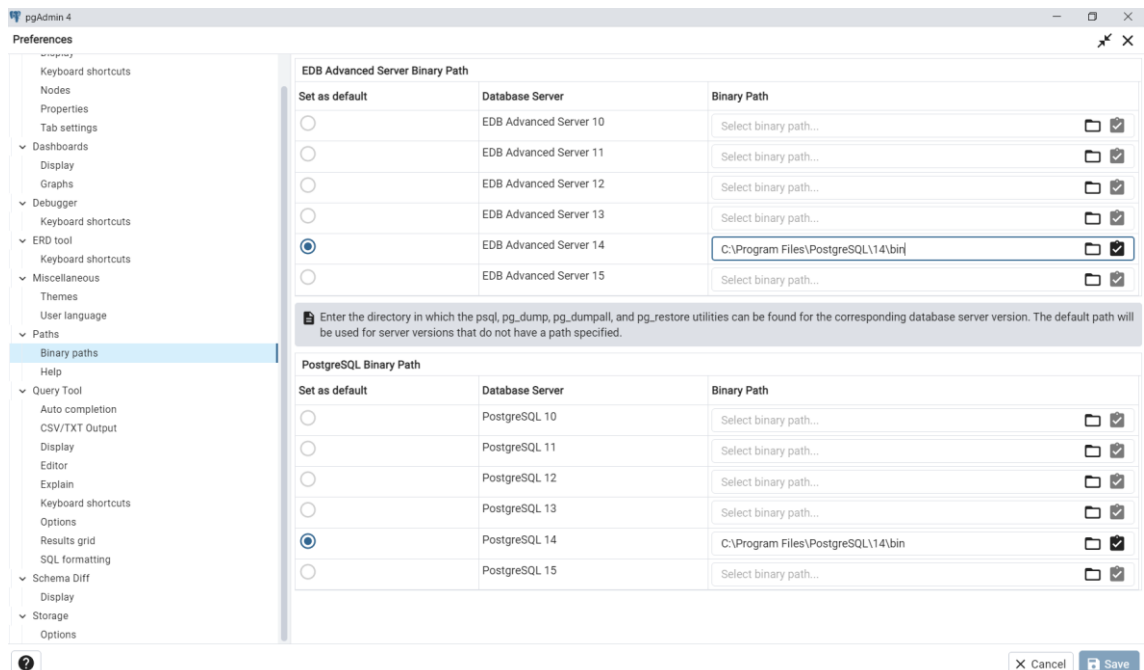
Anexo A: Instalación de ambiente

Los pasos a seguir para instalar el ambiente de trabajo y crear la base de datos son los siguientes:

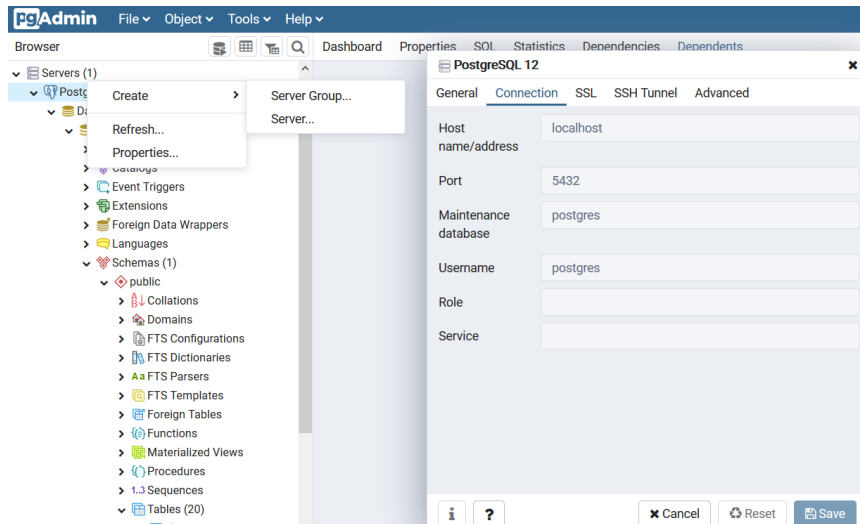
1. Descargar e instalar PostgreSQL server y pgAdmin del siguiente link:
<https://www.postgresql.org/download/>
2. Abrir la herramienta pgAdmin e ingresar en File > Preferences.



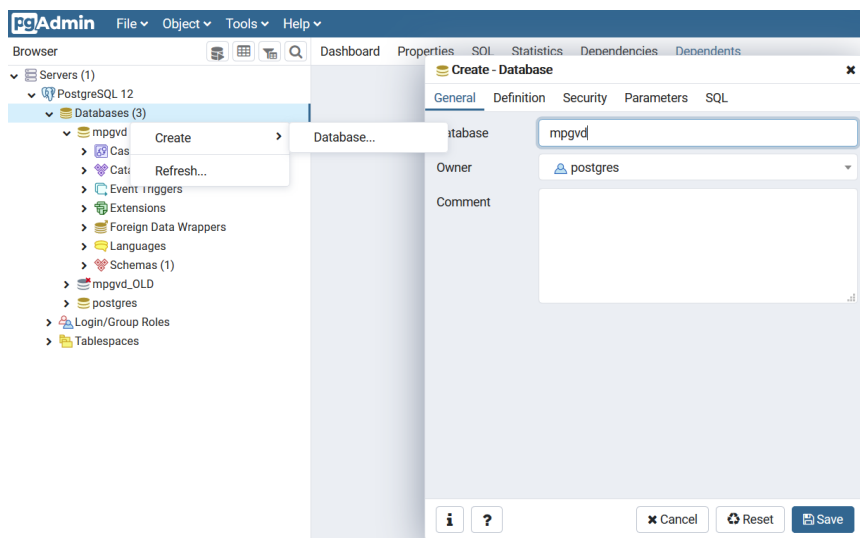
En la nueva ventana que se abrirá dirigirse a Paths > Binary paths e ingresar allí las rutas en dónde se encuentran los archivos ejecutables del servidor (ej: C:\Program Files\PostgreSQL\14\bin)



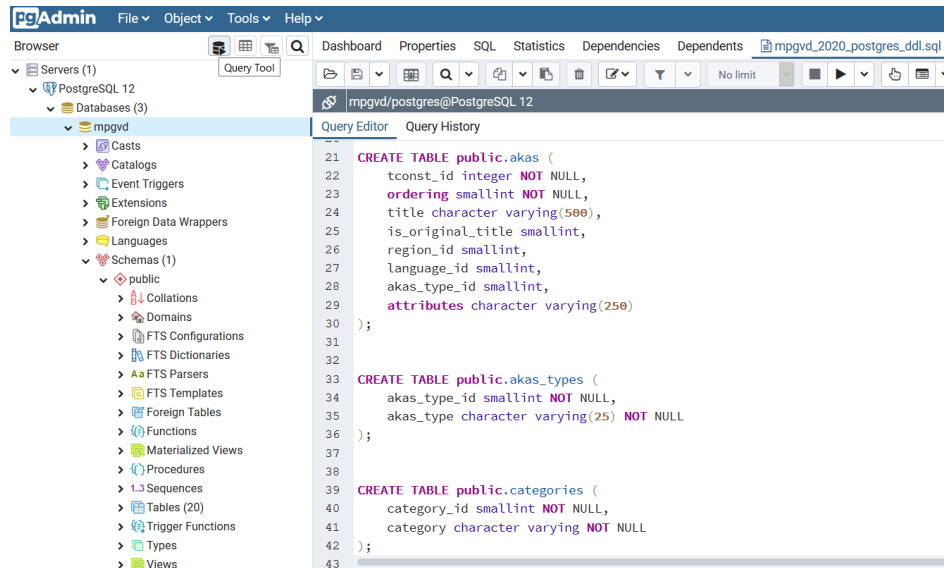
3. Utilizando la herramienta pgAdmin crear una nueva conexión servidor PostgreSQL local con los siguientes parámetros:



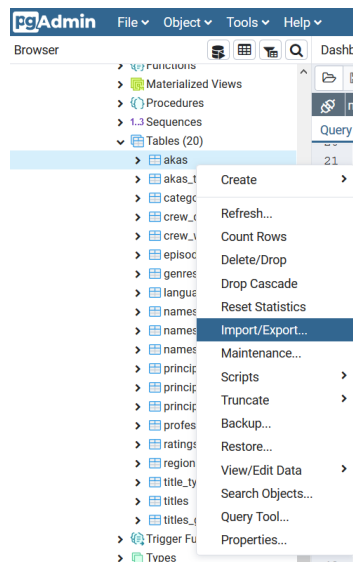
4. Crear una nueva base de datos (por ejemplo: de nombre mpgvd):



5. Accediendo al Query Tool dentro de pgAdmin abrir el archivo `mpgvd_2022_postgres_ddl.sql` y ejecutar las sentencias DDL que contiene:



6. Para importar los datos en cada una de las tablas seleccionar la opción Import/Export:



Los parámetros a utilizar para la importación son los siguientes:

- Format: CSV
- Encoding: UTF8
- Header: Yes
- Delimiter: ,
- Quote: ``
- Escape: ``

Import/Export data - table 'titles' ↗ ✕

Options Columns

Import/Export Import Export

File Info

Filename 📁

Format | ▾

Encoding ✕ | ▾

Miscellaneous

OID

Header

Delimiter | ▾

Specifies the character that separates columns within each row (line) of the file. The default is a tab character in text format, a comma in CSV format. This must be a single one-byte character. This option is not allowed when using binary format.

Quote ✕ | ▾

Specifies the quoting character to be used when a data value is quoted. The default is double-quote. This must be a single one-byte character. This option is allowed only when using CSV format.

Escape ✕ | ▾

Specifies the character that should appear before a data character that matches the QUOTE value. The default is the same as the QUOTE value (so that the quoting character is doubled if it appears in the data). This must be a single one-byte character. This option is allowed only when using CSV format.

ℹ ? ✕ Close ↺ Reset

- Una vez importadas todos los archivos de datos pueden utilizarse el Query Tool para escribir consultas y ejecutar sentencias SQL sobre la base de datos

Anexo B: Referencias

A continuación, se incluyen algunos links de referencia que pueden ser de utilidad:

- PostgreSQL: Tutorial
<https://www.postgresqltutorial.com/>
- pgAdmin 4
<https://www.pgadmin.org/docs/pgadmin4/6.10/index.html>
- PostgreSQL: Manual
<https://www.postgresql.org/docs/manuals/>