# Memoria

# PL2-Parte 2

Álvaro Camacho 48083824R

Fátima C. de la Morena 71961234F

Comandos usados durante la segunda parte de la PL2

* Crear tablas intermedias:

CREATE TABLE actores\_pelis\_intermedio(año int, pelicula text, actor text);

CREATE TABLE actores\_ intermedio (nombre text, fecha\_nacimiento date);

CREATE TABLE directores\_ intermedio (nombre text, fecha\_nacimiento date);

CREATE TABLE películas\_ intermedio (año int, titulo text, generos text, puntuacion float, duracion text, idioma text, calificacion text);

CREATE TABLE pelis\_comentarios\_ intermedio (año int, titulo text, puntuacion float, autor text, contenido text, hash text, pagina\_web text);

CREATE TABLE pelis\_directores\_ intermedio (titulo text, año int, nombre text);

* Copiar datos de los CSV a las tablas intermedias:

COPY actores\_pelis\_ intermedio (año, pelicula, actor) from '/Users/alvarocamacho/Desktop/Bideoklub/Ficheros csv/actores\_pelis.csv' csv header delimiter E'\t' NULL 'NULL' encoding 'UTF8';

COPY actores\_ intermedio (nombre, fecha\_nacimiento) from '/Users/alvarocamacho/Desktop/Bideoklub/Ficheros csv/actores.csv' csv header delimiter E'\t' NULL 'NULL' encoding 'UTF8';

COPY pelis\_directores\_ intermedio (titulo, año, nombre) from '/Users/alvarocamacho/Desktop/Bideoklub/Ficheros csv/pelis\_directores.csv' csv header delimiter E'\t' NULL 'NULL' encoding 'UTF8'

COPY directores\_ intermedio (nombre, fecha\_nacimiento) from '/Users/alvarocamacho/Desktop/Bideoklub/Ficheros csv/directores.csv' csv header delimiter E'\t' NULL 'NULL' encoding 'UTF8'

COPY películas\_ intermedio (año, titulo, generos, puntuacion, duracion, idioma, calificacion) from '/Users/alvarocamacho/Desktop/Bideoklub/Ficheros csv/peliculas.csv' csv header delimiter E'\t' NULL 'NULL' encoding 'UTF8';

COPY pelis\_comentarios\_ intermedio (año, titulo, puntuacion, autor, contenido, hash, pagina\_web) from '/Users/alvarocamacho/Desktop/Bideoklub/Ficheros csv/pelis\_comentarios.csv' csv header delimiter E'\t' NULL NULL'

* Meter datos a las tablas finales

INSERT INTO peliculas.personal(nombre, año\_nacimiento)

SELECT nombre, fecha\_nacimiento

FROM actores\_intermedio

*Solo metemos en las tablas finales las columnas comunes como en este caso q la tabla peliculas.personal también contiene nacionalidad,pero como en la tabla intermedia no,no la añadimos.*

INSERT INTO peliculas.peliculas(año, titulo, duracion, idioma, calificacion) SELECT distinct año, titulo, duracion, idioma, calificacion FROM peliculas\_intermedio;

INSERT INTO peliculas.actua(nombre\_personal\_actor, año\_peliculas, titulo\_peliculas) SELECT distinct actor, año, pelicula FROM actores\_pelis\_intermedio;

INSERT INTO peliculas.actor(nombre\_personal) SELECT distinct actor FROM actores\_pelis\_intermedio;

INSERT INTO peliculas.criticas(critico, puntuacion, texto, año\_peliculas, titulo\_peliculas, url\_pag\_web) SELECT distinct autor, puntuacion, contenido, año, titulo, pagina\_web FROM pelis\_comentarios\_intermedio;

*Falta fecha porque en la tabla intermedia no tienen fecha las criticas*

INSERT INTO peliculas.director(nombre\_personal) SELECT distinct nombre FROM directores\_intermedio;

INSERT INTO peliculas.generos(genero, año\_peliculas, titulo\_peliculas) SELECT distinct generos, año, titulo FROM peliculas\_intermedio;

*Falta caratulas y pag\_web porque no hay como rellenarlos con las tablas intermedias*

* Género

En las peliculas tenemos el atributo género,el cual es multievaluado,este tipo de de atributos crean una tabla a mayores,nosotros para tratarlos usamos:

SELECT distinct regexp\_split\_to\_table(genero, '\s+'),

año::int,titulo

FROM public.peliculas

Este comando separa los generos en diferentes tuplas ,de esta manera podemos cargar la tabla género.