

# Βάσεις Δεδομένων Project

ΓΙΑΝΝΗΣ ΜΠΑΝΟΣ 3140126

ΜΑΡΙΝΑ ΝΤΟΓΚΑ 3160119

Αρχικά, χρησιμοποιήσαμε το python πρόγραμμα `gen_ddl_python3.py`, που μας έχετε δώσει, για να κάνουμε τα `create table` αμέσως και φτιάξαμε τους πίνακες στη βάση μας.

Έπειτα, ανεβάσαμε τα αρχεία `.csv` για κάθε πίνακα και διαγράψαμε τα διπλότυπα \*\*\*\*

Στη συνέχεια, φτιάξαμε νέους πίνακες τον `Links`, `Keywords`, `Ratings` και `Credits` με συνένωση στον `Movies_metadata`, στη οποία συνένωση συμβαίνει το εξής: η πλειάδα από τη σχέση `Movies_metadata (id)` με τη πλειάδα από τη σχέση `πχ του Links (movieid)` συσσωματώνονται και συμπεριλαμβάνονται στο αποτέλεσμα μόνο αν συμπίπτουν στο `id` και `movieid`. Οπότε αποτέλεσμα έχει να διαγραφθούν τα δεδομένα του `Links` που δεν αναφέρονται στον `Movies_metadata`. Αυτό το κάνουμε και για τους άλλους 3 πίνακες ενώ ταυτόχρονα βάζουμε κανονικά όλες τις πλειάδες των πινάκων.

Τέλος, ορίζουμε τα `foreign` και `primary keys`, διαγράφουμε τους αρχικούς πίνακες (`links`, `keywords`, `ratings`, `credits`), και κάνουμε το ER όπου ενδεικτικά για τον `Movies_metadata` έχουμε αναφέρει μόνο 4 πεδία.