

## ΣΥΝΟΠΤΙΚΗ ΑΝΑΦΟΡΑ 2<sup>ης</sup> ΕΡΓΑΣΙΑΣ – ΒΑΣΕΙΣ ΔΕΔΟΜΕΝΩΝ

Μαρία Κονταράτου - Γεώργιος Κουμουνδούρος  
(3200078-3200083)

Για τη δημιουργία των αρχείων: create\_collection, create\_genre, create\_movie\_cast, create\_movie\_collection, create\_movie\_crew, create\_movie\_genre, create\_movie\_production\_company, create\_movie, create\_keyword, create\_movie\_keywords, create\_production\_company, create\_ratings, χρησιμοποιήσαμε τη λειτουργία Import/Export που περιέχει το PgAdmin για την εισαγωγή των δεδομένων. Ανάλογα με τον τύπο δεδομένων που είχαμε, χρησιμοποιήσαμε integers, boolean, character και άλλες μεταβλητές για τη δημιουργία πινάκων. Εκεί ορίσαμε και τα primary keys ανάλογα τον πίνακα. Επιπλέον, τα παραπάνω ζητούμενα sql files τα κάναμε backup μέσω του PgAdmin, οπότε περιέχουν και επιπλέον πληροφορίες σχετικά με τους πίνακες.

Μέσω του αρχείου alter\_table, κάναμε τις κατάλληλες δηλώσεις foreign keys ανάλογα τον πίνακα

Ως προς τον κώδικα python, αρχικά χρησιμοποιούμε την μεταβλητή file για το διάβασμα του αρχείου keywords.csv, η μεταβλητή f στην αρχή χρησιμοποιείται για να γράψουμε τα ζητούμενα στοιχεία στο αρχείο Movie\_Keywords.csv. Με την εντολή <writer.writerow(["movie\_id", "keyword\_id"])> γράφουμε το header του πίνακα. Έπειτα δημιουργούμε ένα άδειο set το οποίο περιέχει tuples 3 στοιχείων, τα οποία είναι το movie\_id, keyword\_id και το name. Επίσης με την if data ελέγχουμε ώστε να μην υπάρχουν null στοιχεία στο αρχείο keywords.csv. Στη συνέχεια περνάμε τα 2 πρώτα στοιχεία κάθε tuple που περιέχεται στο set στο αρχείο Movie\_Keywords.csv και κλείνουμε το συγκεκριμένο αρχείο. Τέλος ανοίγουμε το αρχείο Keywords.csv και εφόσον με την εντολή <writer.writerow(["id", "name"])> γράψουμε το header του πίνακα, με την for περνάμε στο αρχείο το 2ο και 3ο στοιχείο κάθε tuple του set, τα οποία είναι το keyword\_id και το name.

Παραθέτουμε και για καλύτερη κατανόηση το ERD:

