**ΠΡΟΦΗΣ ΙΩΑΝΝΗΣ : p3180161**

**ΠΑΛΑΙΟΛΟΓΟΣ ΚΛΕΑΝΘΗΣ : p3180136**

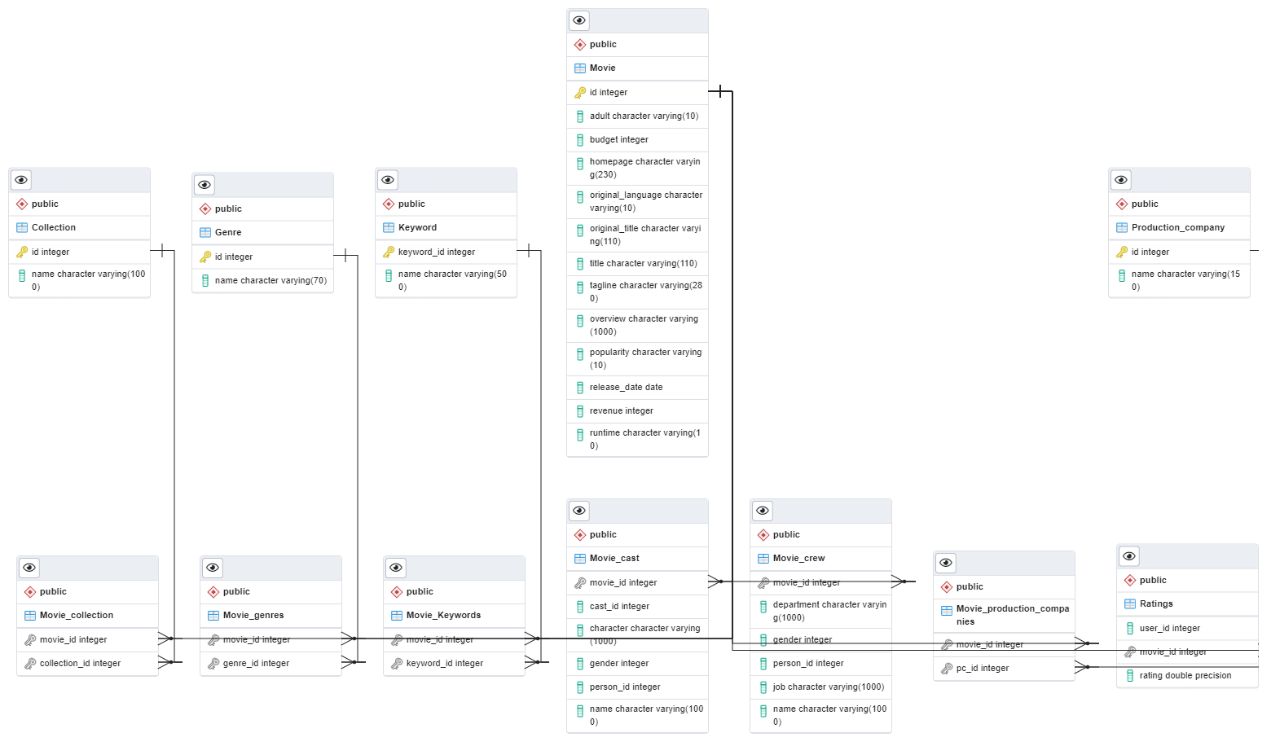
**Δημιουργία, import των .csv και προσθήκη primary keys :** Οι πίνακες δημιουργήθηκαν με την εντολή create table. Για την δημιουργία του Movie λόγω πολλών columns χρησιμοποιήθηκε το gen\_ddl\_python3.py . Για τους πίνακες Collection, Genre, Keyword, Movie, Production\_company, Ratings προστέθηκαν primary keys και έγινε η κατάλληλη αφαίρεση διπλοτύπων. Επίσης από μερικούς πίνακες αφαιρέθηκαν ορισμένες σειρές οι οποίες προκαλούσαν error κατά την προσπάθεια import τους.

**Δημιουργία Movie\_Keywords.csv και Keyword.csv :**

Πιο συγκεκριμένα, για την δημιουργία του Movie\_Keywords αρχικά διαβάζω το keywords.csv και δημιουργώ μια λίστα movie\_id και μια json οι οποίες περιέχουν τα movie\_id και μια λίστα που κάθε στοιχείο της είναι μια λίστα από λεξικά. Στην συνέχεια, χρησιμοποιώντας την json φτιάχνω μια λίστα id η οποία περιέχει μόνο το 1ο value κάθε λεξικού, δηλαδή το id του. Μετατρέπω τις λίστες movie\_id και id σε λίστες που περιέχουν λίστες και έπειτα ανοίγω προσωρινό csv όπου γράφω για κάθε σειρά το movie\_id όσες φορές όσες και το μέγεθος του json της σειράς και στη συνέχεια γράφω και το κατάλληλο id. Π.χ η πρώτη σειρά έχει 8 στοιχεία json άρα θα γράψω το movie\_id 8 φορές και ταυτόχρονα το εκάστοτε id. Τέλος, δημιουργώ το τελικό csv στο οποίο γράφω όλες τις μη κενές σειρές του προσωρινού.

Για την δημιουργία του Keyword ακολουθώ παρόμοια βήματα. Χρησιμοποιώ και πάλι την βοηθητική λίστα json που κάθε στοιχείο της είναι μια λίστα από λεξικά. Γράφω σε προσωρινό csv τα values του κάθε λεξικού και στο τελικό keyword.csv που δημιουργώ γράφω μόνο τις μη κενές σειρές του προσωρινού.

**Προσθήκη foreign keys :** Η προσθήκη των foreign keys έγινε με την εντολή alter table. Αφαιρέθηκαν περίπου 10 σειρές από τον πίνακα Movie έτσι ώστε να μπορέσουν να προστεθούν τα foreign keys.

Παρακάτω βρίσκεται το ER διάγραμμα που παράχθηκε αυτόματα από το pgΑdmin για να φανούν καλύτερα οι συσχετίσεις.