

Βάσεις Δεδομένων 2025 - 2^η Εργασία

Διαχείριση Δεδομένων Ταινιών

Στόχος των εργαστηριακών εργασιών είναι η εξοικείωση με τη υλοποίηση και διαχείριση βάσεων δεδομένων μέσα από ένα χρηστικό παράδειγμα. Στη δεύτερη εργασία θα ασχοληθούμε με την υλοποίηση του σχήματος, την καταχώρηση δεδομένων σε αυτό και την υποβολή ερωτημάτων προς τη βάση δεδομένων.

A – Απαιτήσεις 2^{ης} εργασίας

Καλείστε να φτιάξετε τη ΒΔ που σχεδιάσατε στην 1η εργασία και αφορά τη διαχείριση δεδομένων κινηματογραφικών ταινιών. Τα δεδομένα αυτά περιλαμβάνουν τα στοιχεία των ταινιών (π.χ. όνομα, διάρκεια, budget), τους συντελεστές τους (ηθοποιούς, σκηνοθέτες), τις εταιρείες παραγωγής κτλ. Αναλυτικότερα, κάποια βασικά δεδομένα είναι διαθέσιμα εδώ:

<https://www.kaggle.com/datasets/juzershakir/tmdb-movies-dataset>

και εδώ:

<https://www.kaggle.com/datasets/ashirwadsangwan/imdb-dataset/>

Καλείστε να δώσετε τις εντολές SQL για τη δημιουργία του σχήματος. Καθώς και τις εντολές για την εισαγωγή δεδομένων.

1ο μέρος – Εντολές δημιουργίας σχήματος

Στο πρώτο μέρος, θα αναλύσετε το σχήμα της 1ης εργασίας και τα δεδομένα που έχετε στη διάθεσή σας, θα λάβετε υπόψη σας τις πληροφορίες και παραδοχές που ακολουθούν και θα δώσετε το σύνολο των εντολών για τη δημιουργία του σχήματος (σε ένα αρχείο κειμένου που θα μετονομάσετε σε **create.sql**)

Για κάθε πίνακα θα πρέπει να δώσετε ένα σύντομο κείμενο με τους περιορισμούς που θα ορίσετε χωρισμένους: α) σε περιορισμούς υποχρεωτικών πεδίων (not null), β) σε περιορισμούς πρωτεύοντος κλειδιού, γ) σε περιορισμούς αναφοράς (ή ξένα κλειδιά), γ) σε περιορισμούς πεδίου τιμών ή προκαθορισμένων τιμών.

Είναι επιθυμητό να δώσετε ονόματα στους περιορισμούς σας ακολουθώντας την ονοματολογία που περιγράφηκε στα μαθήματα: π.χ. movie_pk για το πρωτεύον κλειδί στον πίνακα movie, movie_director_fk για το ξένο κλειδί από τον movie προς τον director, movie_language_chk για περιορισμό πεδίου τιμών στο πεδίο language του πίνακα movie.

Οι τύποι των πεδίων και επιπλέον περιορισμοί προκύπτουν από τα δεδομένα που ακολουθούν και που πρέπει να εισάγετε στη ΒΔ σας και από τη φύση του προβλήματος (π.χ. μια ταινία έχει πάντοτε τίτλο).

2ο μέρος – Εισαγωγή δεδομένων στο σχήμα

Στο σχήμα που δημιουργήσατε να εισάγετε τα ακόλουθα δεδομένα.

Να δώσετε εντολές για την εισαγωγή τουλάχιστον 10 ταινιών, 20 ηθοποιών που παίζουν στις ταινίες αυτές, 8 σκηνοθετών και 4 εταιρειών παραγωγής. Αν χρειάζονται να γίνουν σχετικές καταχωρήσεις σε επιπλέον πίνακες που δεν αναφέρονται εδώ, θα πρέπει να τις κάνετε και αυτές. Δείτε επίσης τα ερωτήματα του 4ου μέρους καθώς μπορεί να χρειαστεί να κάνετε συγκεκριμένες

καταχωρήσεις για να έχετε απαντήσεις στα ερωτήματα (επιθυμητό αλλά όχι υποχρεωτικό).

Να δώσετε τις εντολές σε ένα αρχείο κειμένου που θα μετονομάσετε σε **insert.sql** και θα περιέχει σχόλια όπου θα εξηγείτε τι κάνετε σε κάθε βήμα.

Σημείωση: Η σειρά με την οποία θα γεμίσετε τους πίνακες δεν είναι απαραίτητα η σειρά με την οποία σας δίνονται.

3ο μέρος – Εντολές διαγραφής σχήματος

Να δώσετε εντολές για τη διαγραφή των πινάκων που δημιουργήσατε. Δώστε τις εντολές σε ένα αρχείο **drop.sql**.

4ο μέρος – Εντολές επερωτήσεων

Να δώσετε εντολές για τα παρακάτω ερωτήματα.

- 1) Ποιοι είναι οι τίτλοι και οι διάρκειες όλων των ταινιών δράσης που έχουν διάρκεια μεγαλύτερη από 120 λεπτά;
- 2) Ποιες είναι οι 3 ταινίες με τον μεγαλύτερο αριθμό συντελεστών (ηθοποιούς και σκηνοθέτες);
- 3) Ποιοι ηθοποιοί έχουν συμμετάσχει στην ταινία με τίτλο της επιλογής σας;
- 4) Ποιες ταινίες σκηνοθέτησε ο σκηνοθέτης που θα επιλέξετε και ποια ήταν η μέση βαθμολογία τους;
- 5) Ποιες είναι οι 2 ταινίες (τίτλοι) με τα περισσότερα έσοδα που έχει παράγει η εταιρεία παραγωγής "Warner Bros";
- 6) Ποιοι ηθοποιοί έχουν παίξει σε περισσότερες από 5 διαφορετικές ταινίες;
- 7) Ποιες εταιρείες παραγωγής έχουν συνολικά έσοδα άνω των 500 εκατομμυρίων δολαρίων από τις ταινίες που έχουν παράγει;
- 8) Ποιοι είναι οι σκηνοθέτες που έχουν σκηνοθετήσει τουλάχιστον 2 ταινίες που είναι sequels και έχουν βαθμολογία πάνω από 7;
- 9) Ποιος σκηνοθέτης έχει παίξει σε όλες τις ταινίες που έχει σκηνοθετήσει;

Δώστε τις εντολές σε ένα αρχείο **select.sql**.

B – Οδηγίες Παράδοσης

Η εργασία θα υλοποιηθεί από ομάδες των 3 ατόμων, αν και επιτρέπεται να εργαστεί κανείς ατομικά. Θα πρέπει τελικά να ανεβάσετε ένα zip αρχείο με ονομασία τους AM των μελών της ομάδας:

π.χ. **AM1-AM2-AM3.zip**

- Το zip θα περιλαμβάνει:
 - ένα αρχείο readme.txt
 - με τα ονοματεπώνυμα και τους AM των φοιτητών της ομάδας
 - τα τέσσερα αρχεία .sql που περιγράφονται παραπάνω
 - ένα αρχείο pdf που να δείχνει τα αποτελέσματα της εκτέλεσης των select ερωτημάτων.

Γ – Άλλες Οδηγίες

Όσες εργασίες δεν τηρούν τις οδηγίες παράδοσης, θα έχουν επίπτωση στο βαθμό.

Όσες εργασίες κριθούν ότι είναι **αντιγραφές θα μηδενίζονται**.

Ημερομηνία παράδοσης: **Στο e-class με οριστική τελική ημερομηνία 16-5-2025**

Όσες εργασίες παραδοθούν μετά το πέρας της ημερομηνίας και μέχρι τις 18-5-2025 θα έχουν μείωση 2 μονάδων στο βαθμό.