

Βάσεις Δεδομένων 2025 - 1^η Εργασία

Διαχείριση Δεδομένων Ταινιών

Στόχος των εργαστηριακών εργασιών είναι η εξοικείωση με τη σχεδίαση, ανάπτυξη και διαχείριση βάσεων δεδομένων μέσα από ένα χρηστικό παράδειγμα. Στην πρώτη εργασία θα ασχοληθούμε με την ανάλυση των απαιτήσεων και τη σχεδίαση της βάσης δεδομένων, τόσο με χρήση του μοντέλου Οντοτήτων-Συσχετίσεων όσο και με το Σχεσιακό μοντέλο.

A – Απαιτήσεις 1^{ης} εργασίας

Καλείστε να φτιάξετε μια ΒΔ για 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/>

Σε μία βασική θεώρηση, μπορείτε να υποθέσετε ότι:

A) Σε κάθε ταινία συμμετέχουν ένας ή περισσότεροι ηθοποιοί, ο καθένας παίζοντας έναν ή μπορεί και περισσότερους ρόλους. Την ταινία σκηνοθετούν ένας ή περισσότεροι σκηνοθέτες. Μπορεί ακόμη να αποτελεί την συνέχεια (Sequel) κάποιας άλλης ταινίας. Για κάθε ταινία μπορούμε να καταγράψουμε ορισμένες λέξεις-κλειδιά (keywords), καθώς και τα είδη στα οποία ανήκει (πχ κωμωδία, θρίλερ, μιούζικαλ), τον προϋπολογισμό (budget) της, τα έσοδά της, τον τίτλο της, το αναγνωριστικό της, την διάρκειά της, το έτος της, καθώς και τον αριθμό ψήφων που έχει λάβει από τους χρήστες και τη μέση βαθμολογία των χρηστών (vote_avg).

B) Κάθε συντελεστής έχει ένα αναγνωριστικό αριθμό και ημ/νία γέννησης, ενώ μπορεί να έχει ακόμα ημερομηνία θανάτου και κύριες επαγγελματικές ιδιότητες (π.χ. σκηνοθέτης, ηθοποιός, σεναριογράφος). Ένας συντελεστής μπορεί να συμμετέχει ως ηθοποιός και ως σκηνοθέτης σε πολλές ταινίες.

Γ) Κάθε εταιρεία παραγωγής χαρακτηρίζεται από το όνομα και τη χώρα της (μπορεί να υπάρχουν διαφορετικές εταιρείες με το ίδιο όνομα σε διαφορετικές χώρες). Γνωρίζουμε επίσης το έτος ιδρύσεώς της και τον ιδρυτή της. Ο ίδιος άνθρωπος μπορεί να έχει ιδρύσει διαφορετικές εταιρείες.

Δ) Ένας ρόλος συνοδεύεται από τον τίτλο του, μια σύντομη περιγραφή και έναν χαρακτηρισμό (καλός, κακός, ουδέτερος).

Η βάση θα υποστηρίξει μια εφαρμογή μέσα από την οποία:

- Θα μπορεί κανείς να εισάγει νέες ταινίες, μαζί με τους συντελεστές της.
- Θα μπορεί να εξαγει χρήσιμα στατιστικά για κάθε ταινία, που μπορεί να φανούν χρήσιμα σε ένα σύστημα προτάσεων ταινιών ή σε μία πλατφόρμα προβολής ταινιών.
- Θα μπορεί να εντοπίζει ηθοποιούς που ερμήνευσαν τον ίδιο ρόλο σε διαφορετικές ταινίες, όπως είναι οι ρόλοι των υπερηρώων ή των κακών σε ταινίες δράσης.

1^ο μέρος – Διάγραμμα Οντοτήτων-Συσχετίσεων

Στο πρώτο μέρος, θα αναλύσετε τις απαιτήσεις, θα λάβετε υπόψη σας τις πληροφορίες και παραδοχές της εκφώνησης και θα δώσετε το διάγραμμα οντοτήτων-συσχετίσεων για τη βάση δεδομένων που θα αναπτύξετε. Μπορείτε να προσθέσετε και δικές σας οντότητες αν το κρίνετε σκόπιμο. Σίγουρα όμως θα πρέπει να εμφανίζονται οι βασικές οντότητες (με τα γνωρίσματά τους) και συσχετίσεις που προσδιορίζει η εκφώνηση.

Το διάγραμμα Ο-Σ να συνοδεύεται από ένα πίνακα όπου θα εξηγείτε με ποιον τρόπο λαμβάνετε υπόψη την κάθε παραδοχή της εκφώνησης καθώς και όποιες άλλες παραδοχές κάνετε.

2^ο μέρος – Σχεσιακό μοντέλο

Με βάση το διάγραμμα Ο-Σ του 1^{ου} μέρους, και ακολουθώντας τα βήματα που παρουσιάστηκαν στη θεωρία:

- α) να δώσετε και τελικό σχεσιακό σχήμα που προκύπτει και
- β) να περιγράψετε τις ενέργειες που θα κάνετε για να απεικονίσετε το διάγραμμα Ο-Σ σε σχεσιακό.

B – Οδηγίες Παράδοσης

Η εργασία θα υλοποιηθεί από ομάδες των 3 ατόμων, αν και επιτρέπεται να εργαστεί κανείς ατομικά. Το διάγραμμα Ο-Σ μπορεί να σχεδιαστεί σε οποιοδήποτε πρόγραμμα σχεδίασης ή και σε εξειδικευμένα προγράμματα και στη συνέχεια να επικολληθεί στο κείμενο της εργασίας ή ως ξεχωριστό αρχείο. Το σχεσιακό μοντέλο και τα ενδιάμεσα βήματα της απεικόνισης επίσης. Θα πρέπει τελικά να ανεβάσετε ένα zip αρχείο με ονομασία τους ΑΜ των μελών της ομάδας:

π.χ. **AM1-AM2-AM3.zip**

- Το zip θα περιλαμβάνει:
 - ένα αρχείο readme.txt
 - με τα ονοματεπώνυμα και τους ΑΜ των φοιτητών της ομάδας
 - ένα αρχείο κειμένου με την τελική εργασία, σε μορφή pdf

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 9-4-2025**

Όσες εργασίες παραδοθούν μετά το πέρας της ημερομηνίας και μέχρι τις 11-4-2025 θα έχουν μείωση 2 μονάδων στο βαθμό.