



Statistics for Data Science

INTRODUCTION TO DATA MANAGEMENT AND ENGINEERING
ASSIGNMENT

Stylianos Giagkos | f3352410

Συνοπτική Αναφορά Βημάτων με Χρήση του pgAdmin (μέσω Queries)

1. **Δημιουργία Βάσης Δεδομένων:** Δημιουργήθηκε νέα βάση δεδομένων για το MovieLens μέσω του pgAdmin με τη χρήση της εντολής:

```
sql
CREATE DATABASE MovieLens;
```

2. **Δημιουργία Πινάκων:** Εκτελέστηκε το script create_tables.sql μέσω του pgAdmin για τη δημιουργία των παρακάτω πινάκων με τα πρωτεύοντα κλειδιά:

- Movie (id)
- Genre (id)
- Collection (id)
- ProductionCompany (id)
- Keyword (id)
- Movie_Cast (cid)
- Movie_Crew (cid)
- Ratings (user_id, movie_id)

Οι εντολές δημιουργίας πινάκων περιλαμβάνουν τη χρήση των εξής SQL queries:

```
sql
CREATE TABLE Movie (
    id INT PRIMARY KEY,
    adult BOOLEAN NOT NULL,
    budget DECIMAL(15, 2) NOT NULL,
    homepage VARCHAR(255),
    original_language VARCHAR(10),
    original_title VARCHAR(255) NOT NULL,
    title VARCHAR(255) NOT NULL,
    tagline TEXT,
    overview TEXT,
    popularity DECIMAL(10, 2),
    release_date DATE,
    revenue DECIMAL(15, 2),
    runtime FLOAT
);
-- (ανάλογες εντολές για τους υπόλοιπους πίνακες)
```

3. **Δημιουργία Πινάκων Σχέσεων:** Δημιουργήθηκαν οι πίνακες σχέσεων χωρίς πρωτεύοντα κλειδιά για τη σύνδεση των οντοτήτων:

- HasKeyword (movie_id, keyword_id)
- BelongsToCollection (movie_id, collection_id)
- HasGenre (movie_id, genre_id)
- HasProductionCompany (movie_id, pc_id)

Οι εντολές δημιουργίας σχέσεων περιλαμβάνουν:

```
sql
CREATE TABLE HasKeyword (
    movie_id INT NOT NULL,
    keyword_id INT NOT NULL
);
-- (ανάλογες εντολές για τους υπόλοιπους πίνακες σχέσεων)
```

4. **Εισαγωγή Δεδομένων:** Τα δεδομένα φορτώθηκαν στους πίνακες μέσω του pgAdmin, χρησιμοποιώντας εντολές INSERT INTO για να γεμίσουν οι πίνακες με τις απαιτούμενες πληροφορίες:
5. **Προσθήκη Περιορισμών Ξένου Κλειδιού:** Εκτελέστηκε το script alter_tables.sql για την προσθήκη περιορισμών ξένου κλειδιού και τη σύνδεση των πινάκων σχέσεων με τους κύριους πίνακες μέσω των εξής queries:

```
sql
ALTER TABLE BelongsToCollection
ADD CONSTRAINT fk_belongstocollection_movie FOREIGN KEY
(movie_id) REFERENCES Movie(id),
ADD CONSTRAINT fk_belongstocollection_collection FOREIGN KEY
(collection_id) REFERENCES Collection(id);
-- (ανάλογες εντολές για τους υπόλοιπους περιορισμούς)
```

6. **Επαλήθευση Ακεραιότητας μέσω Queries:** Η ακεραιότητα των δεδομένων και οι περιορισμοί ξένου κλειδιού ελέγχθηκαν μέσω SQL queries, επιβεβαιώνοντας τη σωστή σύνδεση των πινάκων:

 - Έγινε έλεγχος για τη σωστή εφαρμογή των περιορισμών ξένου κλειδιού με την εκτέλεση του εξής query:

```
sql
SELECT * FROM information_schema.table_constraints
WHERE constraint_type = 'FOREIGN KEY';
```

- Ελέγχθηκε αν όλοι οι περιορισμοί ξένου κλειδιού υπάρχουν στους πίνακες σχέσεων και αν οι συνδέσεις μεταξύ των πινάκων είναι σωστές.