

## 2η Εργασία: Spotify Database

**Προθεσμία: 31/03/2025**

### Σκοπός:

Σε αυτή την εργασία, θα δημιουργήσετε τους πίνακες για μια βάση δεδομένων μουσικής, η οποία περιέχει πληροφορίες για καλλιτέχνες, άλμπουμ, τραγούδια, είδη μουσικής και χαρακτηριστικά του ήχου, όπως η δημοτικότητα, η ρυθμικότητα και άλλα. Για τον ορισμό του σχήματος της βάσης δεδομένων, θα βασιστούμε στους τύπους των δεδομένων εισόδου που περιέχονται στα αρχεία CSV που σας δίνονται. Για την εισαγωγή των δεδομένων, θα πρέπει να χρησιμοποιηθεί το SQL Server extension του Azure Data Studio.

### Περιγραφή Δεδομένων:

Τα δεδομένα της άσκησης βρίσκονται διαθέσιμα στον φάκελο της 2ης εργασίας στο αρχείο **Spotify Tables**.

Η αρχική μορφή των δεδομένων περιέχει πληροφορίες για 10,000 albums από τη βάση δεδομένων του Spotify, καλλιτέχνες, τραγούδια, χαρακτηριστικά του ήχου, και κατηγορίες μουσικής. Τα δεδομένα που θα χρησιμοποιηθούν περιλαμβάνουν τους εξής πίνακες:

- **albums:** Περιέχει τα άλμπουμ με αναγνωριστικά id, όνομα, τύπο άλμπουμ, ημερομηνία κυκλοφορίας, και δημοτικότητα.
- **artists:** Περιέχει τους καλλιτέχνες με αναγνωριστικά id, όνομα, δημοτικότητα και αριθμό ακολούθων.
- **audio\_features:** Περιέχει ηχητικά χαρακτηριστικά των τραγουδιών (tracks) όπως acousticness, danceability, energy, loudness, tempo και άλλα.
- **genres:** Περιέχει τα είδη μουσικής με το όνομα κάθε είδους.
- **tracks:** Περιέχει τα τραγούδια με αναγνωριστικά id, όνομα, διάρκεια, αριθμό τραγουδιού, δημοτικότητα και προεπισκόπηση.
- **r\_albums\_artists:** Συσχετίζει τα άλμπουμ με τους δημιουργούς τους, περιέχοντας τα πεδία album\_id και artist\_id.
- **r\_albums\_tracks:** Συσχετίζει τα άλμπουμ με τα τραγούδια τους, περιέχοντας τα πεδία album\_id και track\_id.
- **r\_artist\_genre:** Συσχετίζει τους καλλιτέχνες με τα μουσικά είδη τους, περιέχοντας τα πεδία genre\_name και artist\_id.
- **r\_track\_artist:** Συσχετίζει τα τραγούδια με τους καλλιτέχνες που τα εκτελούν, περιέχοντας τα πεδία track\_id και artist\_id.

## Τι θα φτιάξετε:

Θα δημιουργήσετε και θα εισαγάγετε δεδομένα στους πίνακες για την εφαρμογή του Spotify στη **βάση που δημιουργήσατε στην 1η εργασία**. Στην βάση αυτή θα πρέπει να δημιουργηθούν πίνακες που θα περιέχουν δεδομένα για το Spotify. Από κάθε αρχείο CSV θα προκύψει ένας πίνακας στη βάση δεδομένων. Για παράδειγμα, το αρχείο `audio_features.csv` θα μετατραπεί σε έναν πίνακα στη βάση δεδομένων, ενώ άλλα αρχεία θα δημιουργήσουν αντίστοιχους πίνακες.

Τα βήματα που πρέπει να ακολουθήσετε:

1. Χρησιμοποιώντας το SQL Server Import plugin του Azure Data Studio, να εισάγετε τα δεδομένα σε κάθε πίνακα, χρησιμοποιώντας τα αντίστοιχα CSV αρχεία.
2. Χρησιμοποιώντας την εντολή ALTER TABLE:
  - ο να δημιουργήσετε τους περιορισμούς πρωτεύοντος κλειδιού (Primary Key) για τους πίνακες: **albums**, **artists**, **audio\_features**, **genres**, **tracks**.
  - ο (Τα πρωτεύοντα κλειδιά των πινάκων `r_albums_artists`, `r_albums_tracks`, `r_artist_genre`, `r_track_artist` θα δημιουργηθούν σε μεταγενέστερη άσκηση).
3. Χρησιμοποιώντας την εντολή ALTER TABLE:
  - ο Να δημιουργήσετε τους περιορισμούς ξένου κλειδιού (Foreign Key) για τους πίνακες: **r\_albums\_artists**, **r\_albums\_tracks**, **r\_artist\_genre**, **r\_track\_artist**.
  - ο Επιπλέον, θα πρέπει να δημιουργήσετε περιορισμό ξένου κλειδιού για τον πίνακα **audio\_features**, θεωρώντας ότι αυτός ο πίνακας περιέχει πληροφορίες για κάθε τραγούδι.

Σημείωση: Κάθε πίνακας μπορεί να έχει περισσότερα από ένα ξένα κλειδιά.

Τοποθετήστε όλες τις εντολές ALTER TABLE σε ένα αρχείο, το οποίο θα ονομάσετε `alter_tables.sql`.

## Απαραίτητα εργαλεία:

- Azure data studio
- SQL Server Import plugin του Azure data studio
- ή εναλλακτικά (δεν το συνιστούμε) το πρόγραμμα bcp για αντιγραφή δεδομένων μεταξύ του SQL Server και αρχείων δεδομένων με συγκεκριμένο format.

## Χρήσιμα links:

Εντολές alter table για δημιουργία και διαγραφή πρωτεύοντος και ξένου κλειδιού:

<https://learn.microsoft.com/en-us/sql/relational-databases/tables/create-primary-keys>

<https://learn.microsoft.com/en-us/sql/relational-databases/tables/create-foreign-key-relationships>

<https://learn.microsoft.com/en-us/sql/relational-databases/tables/delete-primary-keys>

Εντολές create table (δεν θα χρειαστούν αν χρησιμοποιήσετε το SQL Server Import) :

<https://learn.microsoft.com/en-us/sql/t-sql/statements/create-table-transact-sql>

Azure Data Studio:

<https://learn.microsoft.com/en-us/sql/azure-data-studio/>

SQL Server Import extension:

<https://learn.microsoft.com/en-us/sql/azure-data-studio/extensions/sql-server-import-extension?view=sql-server-ver16>

## Παραδοτέα:

- Δημιουργήστε ένα .txt αρχείο στο οποίο θα αναγράφονται τα εξής στοιχεία: ονοματεπώνυμο και αριθμοί μητρώου των μελών της ομάδας, το endpoint του Azure instance σας, το όνομα της βάσης σας και το username και το password του χρήστη examiner ή ενός άλλου χρήστη με read-only δικαιώματα, ώστε να μπορούμε να δούμε τους πίνακες της βάσης σας. Το .txt αρχείο θα πρέπει να έχει την παρακάτω μορφή:

<Ονοματεπώνυμο 1> - <A.M. 1>

<Ονοματεπώνυμο 2> - <A.M. 2>

Endpoint: <name\_of\_the\_endpoint>

Username: <username>

Password: <password>

Database: <name\_of\_the\_database>

- Βάλτε σε ένα φάκελο
  - a. το alter\_tables.sql αρχείο,
  - b. το **αρχείο .txt**,
  - c. καθώς και μία **συνοπτική αναφορά** (~ 10 γραμμές) για το τι κάνατε σε κάθε βήμα της άσκησης.

Το όνομα του φακέλου πρέπει να αποτελείται από τους αριθμούς μητρώου σας χωρισμένους με παύλα, δηλαδή *αριθμός\_μητρώου\_1-αριθμός\_μητρώου\_2*. Δημιουργήστε ένα .zip αρχείο αυτού του φακέλου, το οποίο θα έχει το ίδιο όνομα με τον φάκελο.

- Ανεβάστε το .zip αρχείο στο eclass στην ενότητα *Εργασίες / 2η Εργασία*.