

## Βήμα 1:

### Επιπεδοποίηση του αρχείου keywords.csv

Την υλοποίηση της εργασίας μας ξεκινήσαμε, με την επεξεργασία του αρχείου *keywords.csv*. Χρησιμοποιώντας τις εντολές για διάβασμα και γράψιμο σε *csv* αρχεία και ξεκινήσαμε να υλοποιούμε το αρχείο σε *python*. Πρώτα γράψαμε την πρώτη γραμμή στα αρχεία. Στο *Keyword.csv*, “*id*”, “*name*” και στο *hasKeyword.csv*, “*movie\_id*”, “*key\_id*”. Δημιουργήσαμε 2 λίστες, μία για κάθε αρχείο, *Keyword.csv* και *hasKeyword.csv*. Τρέχοντας όλο το αρχείο *keywords.csv* και αποκωδικοποιώντας την *json* συμβολοσειρά, ελέγχαμε αν το στοιχείο που θέλαμε να εισάγουμε, υπήρχε στην λίστα και αν υπήρχε δεν το εισαγάγαμε, ενώ αν δεν υπήρχε, το εισαγάγαμε. Αφού τελειώσαμε την επανάληψη και στην λίστα υπήρχαν μοναδικά στοιχεία, δηλαδή **χωρίς διπλότυπα**, τότε γράψαμε τις λίστες στα ανάλογα αρχεία *Keyword.csv* και *hasKeyword.csv*. Συνολικά στο αρχείο *Keyword.csv* υπάρχουν μαζί με τους τίτλους των πεδίων **11821** εγγραφές, ενώ στο αρχείο *hasKeyword.csv* υπάρχουν μαζί με τους τίτλους των πεδίων **56982** εγγραφές.

## Βήμα 2:

### SQL Server Import

Σύμφωνα με τις οδηγίες κατεβάσαμε το ανάλογο *extension* και κάναμε *update* το *Azure data studio* για να μπορούμε να ενεργοποιήσουμε την επιλογή *Import wizard*.

## Βήμα 3:

### Εισαγωγή Δεδομένων στην βάση μας

Έπειτα ασχοληθήκαμε με το *Azure data studio*. Με την βοήθεια του *Import wizard* εισαγάγαμε τους πίνακες που μας έδινε σαν δεδομένα η άσκηση, εκτός από τον *keywords.csv*, αφού εισαγάγαμε τους 2 άλλους που προήλθαν από αυτόν, τους *Keyword.csv* και *hasKeyword.csv*. Φυσικά και σύμφωνα με τα βήματα της άσκησης, όπου το *Azure* “μάντευε” λάθος τύπο δεδομένων, εμείς βλέποντας το *error*, τον αλλάζαμε κατάλληλα.

## Βήμα 4:

### Δημιουργία πρωτεύοντων και Ξένων κλειδιών

Τέλος, υλοποιήσαμε την προσθήκη των πρωτεύοντων και ξένων κλειδιών στους πίνακες που ζητάγε η άσκηση. Δημιουργώντας ένα καινούριο *query* και σύμφωνα με τον πρότυπο κώδικα που δόθηκε στις διαφάνειες, προσαρμόζοντας τον *sql* κώδικα δημιουργήσαμε τα πρωτεύοντα και ξένα κλειδιά. Πρωτεύοντα κλειδιά χρειάζονταν οι πίνακες *movie*, *genre*, *productioncompany*, *collection*, *movie\_cast*, *movie\_crew* και *Keyword*. Στους πίνακες *movie*, *genre*, *productioncompany*, *collection* και *Keyword*, σαν **πρωτεύον** κλειδί μπήκε το πεδίο **id**, καθώς το **id** είναι το χαρακτηριστικό γνώρισμα (π.χ στον πίνακα *movie*, το *id* είναι το **χαρακτηριστικό γνώρισμα για κάθε ταινία**, δηλαδή **κάθε ταινία την ορίζει με μοναδικό τρόπο το id της**). Στους πίνακες *movie\_cast*, *movie\_crew* σαν πρωτεύον κλειδί, μπήκε το **cid** καθώς προσδιορίζει μοναδικά κάθε επιλογή που έχει γίνει στο *casting/crew*. Ξένα κλειδιά χρειάζονταν οι πίνακες *belongsToCollection*, *hasGenre*, *hasProductionCompany*, *ratings*,

## ΕΡΓΑΣΙΑ 2

Μαρία Σχοινάκη: 3210191  
Ελευθέριος Βερούχης: 3200019

*movie\_cast*, *movie\_crew*, *hasKeyword*. Σε όλους τους πίνακες μπήκε ως **ξένο** κλειδί το ***movie\_id***, που είναι πρωτεύον κλειδί του πίνακα ***movie***. Ο λόγος είναι ότι κάθε πίνακας δείχνει την σχέση μιας ταινίας με κάτι άλλο(π.χ *genre*). Ο πίνακας *belongsTocollection* είχε σαν 2<sup>ο</sup> ξένο κλειδί το ***collection\_id*** του πίνακα ***collection*** καθώς δείχνει την σχέση ανάμεσα σε μία ταινία και την συλλογή που ανήκει. Ο πίνακας *hasGenre* είχε σαν 2<sup>ο</sup> ξένο κλειδί το ***genre\_id*** του πίνακα ***genre*** καθώς δείχνει την σχέση ανάμεσα σε μία ταινία και τα είδη που ανήκει. Ο πίνακας *hasProductionCompany* είχε σαν 2<sup>ο</sup> ξένο κλειδί το ***pc\_id*** του πίνακα ***productioncompany*** καθώς δείχνει την σχέση ανάμεσα σε μία ταινία και εταιρείες παραγωγής της. Ο πίνακας *hasKeyword* είχε σαν 2<sup>ο</sup> ξένο κλειδί το ***key\_id*** του πίνακα ***Keyword*** καθώς δείχνει την σχέση ανάμεσα σε μία ταινία και τις λέξεις που την χαρακτηρίζουν. Τέλος, οι πίνακες *ratings*, *movie\_cast*, *movie\_crew* **δεν** έχουν 2<sup>ο</sup> ξένο κλειδί καθώς τα *user\_id* και *person\_id* που είναι τα πιθανά ξένα κλειδιά, δεν είναι πρωτεύοντα και μάλιστα δεν υπάρχουν καν σε κάποιον άλλον πίνακα της βάσης μας.