

# Αναφορά εξαμηνιαίας εργασίας στις Βάσεις Δεδομένων

Η παρούσα εργασία πραγματοποιήθηκε από τους φοιτητές:

Ρεϊζή Εμμανουήλ με ΑΜ: 03121067

Δημητρακόπουλο Δημήτρη με ΑΜ: 03121066

Σκότη Απόστολο με ΑΜ: 03121212

Το git repo με τα αρχεία της βάσης βρίσκεται στον παρακάτω σύνδεσμο:

<https://github.com/Apostolos24/Database-Semester-Project/tree/main>

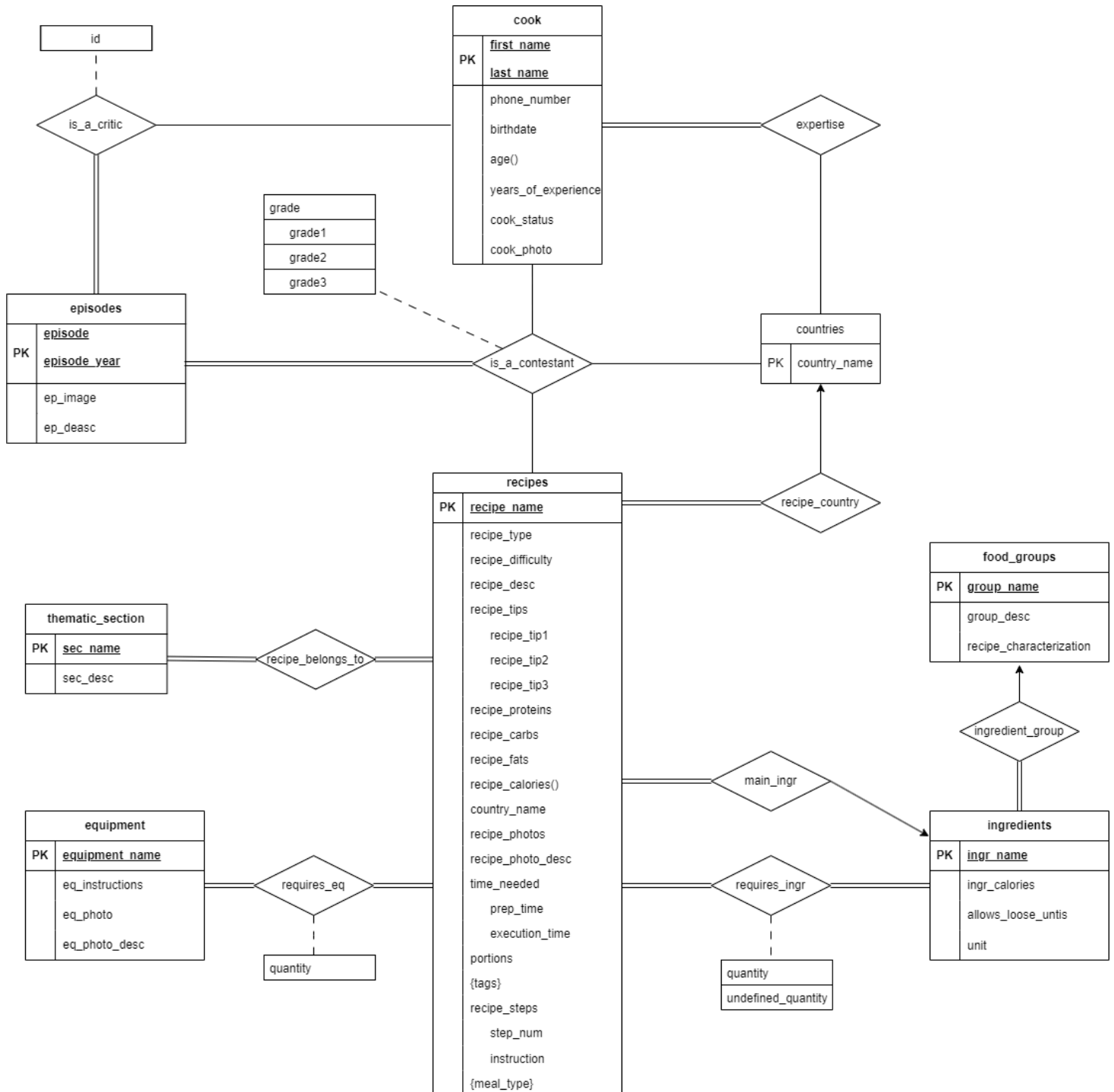
στο main branch. Το αρχείο DDL.sql περιέχει τον ορισμό των tables, τα indexes, triggers, stored functions και stored procedures. Το DML.sql περιέχει τα δεδομένα που εισάγουμε στην βάση και τα views που ορίζουμε για τους χρήστες της εφαρμογής. Επίσης, το queries.sql περιέχει τις απαντήσεις στα ερωτήματα 3.1 ως 3.15 σε μορφή ερωτημάτων select. Τέλος, υπάρχουν εικόνες που δείχνουν το ER και το σχεσιακό διάγραμμα, στα αρχεία er.drawio και relational-model.png αντίστοιχα.

Για να εγκατασταθεί σωστά η βάση ακολουθούμε τα εξής βήματα:

- Τρέχουμε πρώτα το αρχείο DDL.sql το οποίο εγκαθιστά την βάση
- Έπειτα τρέχουμε το αρχείο DML.sql το οποίο εισάγει δεδομένα στην βάση και δημιουργεί views και users
- Τέλος μπορούμε να τρέξουμε το αρχείο queries.sql που περιέχει τα queries που μας ζητήθηκαν

Στην επόμενη σελίδα φαίνεται το ER διάγραμμα της βάσης:

# ER διάγραμμα



(υπάρχει και στο github)

Στο μοντέλο οντοτήτων συσχετίσεων της βάσης περιλαμβάνουμε τις κάτωθι οντότητες:

1. Συνταγές, με τα ακόλουθα attributes: όνομα συνταγής, είδος συνταγής (δηλαδή συνταγή μαγειρικής, ζαχαροπλαστικής ή ρόφημα), βαθμός δυσκολίας (ακέραιος μεταξύ 0 και 5), περιγραφή, 3 tips (με την μορφή τριών διαφορετικών attributes), πλήθος πρωτεϊνών, υδατανθράκων, λίπους και θερμίδων ανά μερίδα, αριθμός μερίδων, φωτογραφία και περιγραφή της φωτογραφίας καθώς και χρόνους εκτέλεσης και προετοιμασίας. Επιπλέον, διαθέτουμε σύνθετα attributes. Αυτά είναι τα tips της συνταγής που αποτελείται από 3 tips, ο χρόνος που απαιτείται για την συνταγή, που αποτελείται από τον χρόνο προετοιμασίας και τον χρόνο εκτέλεσης και τα βήματα της συνταγής που χωρίζονται στον αριθμό του βήματος και την περιγραφή του βήματος. Ακόμη, υπάρχουν και οι ιδιότητες πολλαπλών τιμών, οι οποίες είναι το είδος γεύματος και οι ετικέτες, καθώς κάθε συνταγή μπορεί να ανήκει σε περισσότερα από ένα είδη γεύματος (πρωινό, μεσημεριανό κλπ) και σε περισσότερες από μία ετικέτες (υγιεινό, κρύο, γρήγορο κλπ). Το πρωτεύον κλειδί της οντότητας είναι το όνομα της συνταγής.
2. Υλικά, με τα ακόλουθα attributes: όνομα υλικού, θερμίδες, μονάδες μέτρησης (πχ γραμμάρια ή ml) και ένα boolean attribute που δηλώνει αν το υλικό επιτρέπει ασαφείς ποσότητες όπως λίγο, αρκετό κλπ. Οι θερμίδες, αν το υλικό μετριέται σε γραμμάρια ή ml, δίνονται ανά 100 gr/ml, διαφορετικά δίνονται ανά μονάδα. Τέλος, πρωτεύον κλειδί είναι το όνομα του υλικού.
3. Ομάδες τροφίμων, με τα ακόλουθα attributes: όνομα ομάδας, περιγραφή και χαρακτηρισμός συνταγής. Το τελευταίο αναφέρεται στο ζητούμενο ότι μία συνταγή χαρακτηρίζεται με βάση την ομάδα τροφίμου του βασικού της υλικού. Πρωτεύον κλειδί είναι το όνομα της ομάδας.

4. Εξοπλισμός, με τα ακόλουθα attributes: όνομα εξαρτήματος, οδηγίες χρήσης, φωτογραφία και περιγραφή της φωτογραφίας. Πρωτεύον κλειδί είναι το όνομα του εξαρτήματος
5. Θεματικές ενότητες, με τα ακόλουθα attributes: όνομα θεματικής ενότητας και περιγραφή της. Πρωτεύον κλειδί είναι το όνομα.
6. Χώρες, με μόνο attribute το όνομα της χώρας, το οποίο είναι και πρωτεύον κλειδί.
7. Επεισόδια, με τα εξής attributes: έτος επεισοδίου, αριθμός επεισοδίου, φωτογραφία και λεζάντα. Το πρωτεύον κλειδί είναι το έτος και ο αριθμός του επεισοδίου. Ο αριθμός δεν αρκεί ως πρωτεύον κλειδί εφόσον κάθε χρόνο επαναλαμβάνονται τα επεισόδια 1 έως 10.
8. Τέλος, οι μάγειρες με τα εξής attributes: όνομα και επίθετο, αριθμός τηλεφώνου, ημερομηνία γέννησης, χρόνια εξειδίκευσης, επαγγελματική κατάρτιση και φωτογραφία του μάγειρα με την περιγραφή της. Επιπλέον, η ηλικία είναι ένα παράγωγο attribute που εξαρτάται από την ημερομηνία γέννησης, δηλαδή είναι η διαφορά της ημερομηνίας γέννησης από την τρέχουσα ημερομηνία. Το πρωτεύον κλειδί είναι το (όνομα,επίθετο).

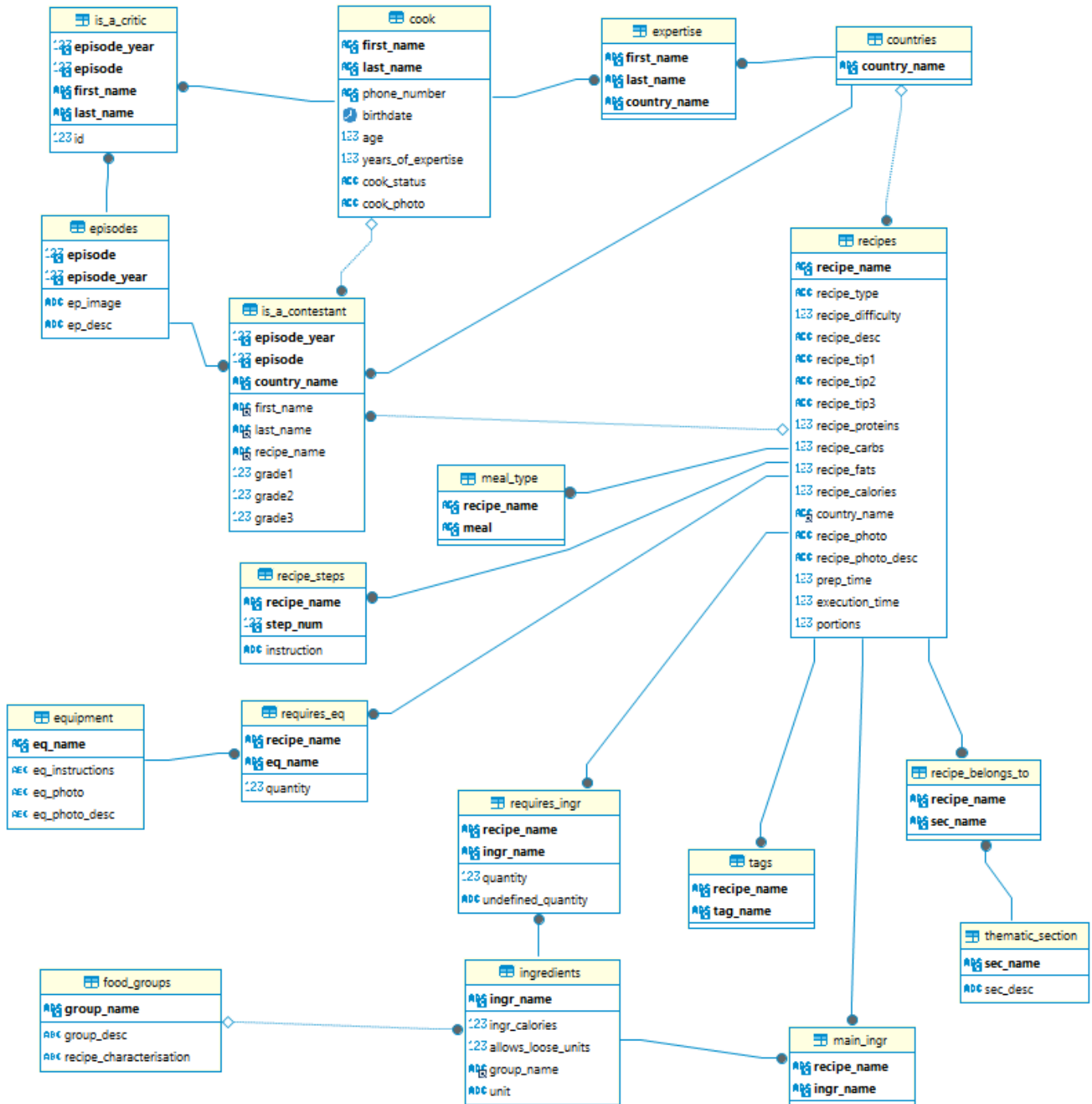
Το ER διάγραμμα περιλαμβάνει επίσης, τις παρακάτω σχέσεις:

1. Η σχέση `requires_ingr` που συνδέει τις συνταγές με τα υλικά τους. Επιπλέον attribute της σχέσης είναι η ποσότητα του υλικού στην συγκεκριμένη συνταγή και το `undefined_quantity`, που χρησιμοποιείται όταν θέλουμε να δηλώσουμε ότι ένα υλικό χρησιμοποιείται με ακαθόριστη ποσότητα σε μία συνταγή (λίγο, πολύ κλπ). Η σχέση αυτή είναι many-to-many και έχει ολική συμμετοχή των συνταγών και των υλικών.
2. Η σχέση `main_ingr` που συνδέει τις συνταγές με το κύριο υλικό τους. Η σχέση αυτή είναι many-to-one από τις συνταγές προς τα υλικά και έχει ολική συμμετοχή των συνταγών και μερική των υλικών.
3. Η σχέση `recipe_belongs_to` που συνδέει τις συνταγές με τις θεματικές ενότητες. Η σχέση αυτή είναι many-to-many και έχει ολική συμμετοχή συνταγών και θεματικών ενοτήτων.
4. Η σχέση `requires_eq` που συνδέει τις συνταγές με τα εξαρτήματά τους. Επιπλέον attribute της σχέσης είναι το `quantity` που δείχνει το πλήθος του κάθε εξαρτήματος που απαιτεί η συνταγή. Η σχέση αυτή είναι many-to-many και έχει ολική συμμετοχή των συνταγών.
5. Η σχέση `recipe_country` που συνδέει τις συνταγές με την χώρα τους. Η σχέση είναι many-to-one από τις συνταγές στις χώρες και έχει ολική συμμετοχή των συνταγών.
6. Η σχέση `ingredient_group` που συνδέει τα υλικά με την ομάδα τροφίμων που ανήκουν. Η σχέση αυτή είναι many-to-one από τα υλικά στις ομάδες, με ολική συμμετοχή των υλικών.

7. Η σχέση expertise που συνδέει τους μάγειρες με τις χώρες στις οποίες ειδικεύονται. Η σχέση είναι many-to-many με ολική συμμετοχή των μαγείρων.
8. Η σχέση is\_a\_critic που συνδέει τους μάγειρες με τα επεισόδια και δηλώνει ότι στο εκάστοτε επεισόδιο ο συγκεκριμένος μάγειρας είναι κριτής. Η σχέση έχει ως επιπλέον attribute ένα id που δείχνει τον αριθμό του κριτή (1ος, 2ος ή 3ος κριτής του επεισοδίου). Η σχέση είναι many-to-many με ολική συμμετοχή των επεισοδίων.
9. Η σχέση is\_a\_contestant. Αυτή δηλώνει ότι στο δεδομένο επεισόδιο έχουμε επιλέξει την εξής χώρα, την εξής συνταγή καθώς και τον μάγειρα που θα εκτελέσει την συνταγή. Η σχέση έχει επιπλέον attribute τα grade1, grade2 και grade3 που δείχνουν την βαθμολογία του μάγειρα στην συγκεκριμένη συνταγή από τους κριτές 1,2 και 3 αντίστοιχα. Πρόκειται για τετραδική σχέση μορφής many-to-many-to-many-to-many. Ολική συμμετοχή έχουν μόνο τα επεισόδια.

Τα παραπάνω περιγράφουν το μοντέλο οντοτήτων-συσχετίσεων της βάσης. Με βάση αυτό, προχωράμε στην ανάπτυξη του σχεσιακού μοντέλου. Στην επόμενη σελίδα φαίνεται το σχεσιακό διάγραμμα.

## Σχεσιακό διάγραμμα



(υπάρχει και στο github)

Κάθε οντότητα του μοντέλου ER αντιστοιχεί σε έναν πίνακα στο σχεσιακό σχήμα με attributes ίδια με αυτά των οντοτήτων. Η μόνη εξαίρεση προκύπτει στα σύνθετα attributes. Για τα είδη γεύματος μιας συνταγής κατασκευάζουμε τον πίνακα meal\_type με foreign key το όνομα της συνταγής και επιπλέον attribute το είδος του γεύματος. Το primary key είναι το (recipe\_name,meal).

Ομοίως για τις ετικέτες κατασκευάζουμε τον πίνακα tags με foreign key το recipe\_name και primary key το (recipe\_name,tag\_name).

Τέλος, για τα βήματα εκτέλεσης της συνταγής κατασκευάζουμε τον πίνακα recipe\_steps με foreign key το recipe\_name, primary key το (recipe\_name,step\_num) και επιπλέον attribute το instruction.

Κάθε σχέση του μοντέλου ER αντιστοιχεί σε έναν πίνακα στο σχεσιακό με ένα foreign key για κάθε οντότητα που συμμετέχει σε αυτήν (το οποίο ταυτίζεται με το primary key της οντότητας). Εξαίρεση αποτελούν οι σχέσεις recipe\_country και ingredient\_group. Οι σχέσεις αυτές είναι many-to-one με ολική συμμετοχή της πλευράς many. Σε αυτές τις περιπτώσεις μπορούμε να απαλείψουμε τον πίνακα τη σχέσης και να προσθέσουμε απλώς ως foreign key στην οντότητα ολικής συμμετοχής (εδώ οι πίνακες ingredients και recipes) το primary key της άλλης οντότητας.

Το primary key των πινάκων που προκύπτουν από σχέσεις του ER είναι ο συνδυασμός των foreign keys. Εξαίρεση αποτελεί ο πίνακας is\_a\_contestant στον οποίο αρκεί ως primary key το (episode\_year,episode,country\_name).



## Περιορισμοί ακεραιότητας

Στο DDL.sql στο οποίο αναπτύσσουμε την βάση έχουμε προσθέσει και αρκετούς περιορισμούς ακεραιότητας. Πέραν από τις δηλώσεις πρωτευόντων και ξένων κλειδιών που έχουμε ήδη αναφέρει, έχουμε τους εξής ακόλουθους περιορισμούς:

- Ο βαθμός δυσκολίας κάθε συνταγής είναι ακέραιος αριθμός μεταξύ του 0 και του 5 και δεν πρέπει να είναι NULL
- Η περιγραφή μίας συνταγής δεν πρέπει να είναι NULL
- Το είδος της συνταγής πρέπει να είναι ένα εκ των τριών επιλογών: Cooking, Pastry ή Beverage
- Οι θερμίδες, τα λιπαρά, οι υδατάνθρακες και οι πρωτεΐνες κάθε συνταγής πρέπει να είναι αριθμοί μεγαλύτεροι ή ίσοι του μηδενός
- Ο αριθμός του κάθε εξαρτήματος που χρησιμοποιεί μία συνταγή είναι ακέραιος μεγαλύτερος ή ίσος του μηδενός
- Η ποσότητα των υλικών που χρησιμοποιεί κάθε συνταγή είναι αριθμός μεγαλύτερος ή ίσος του μηδενός
- Οι μονάδες μέτρησης ενός υλικού είναι μία των κάτωθι: gr, ml, tsp, tbsp, unit και cups
- Η ακαθόριστη ποσότητα ενός υλικού μπορεί να είναι ένα από τα παρακάτω: Some, Little, Much, A lot, pinch, to taste, as needed, to garnish, slices, stalks, cloves
- Η ηλικία ενός μάγειρα πρέπει να είναι μεγαλύτερη του 15 και μικρότερη του 100
- Το χρόνια εμπειρίας ενός μάγειρα πρέπει να είναι λιγότερα από την ηλικία του
- Η ημερομηνία γέννησης του κάθε μάγειρα πρέπει να είναι έγκυρη (μετά το 1900)
- Ο αριθμός τηλεφώνου ενός μάγειρα πρέπει να είναι μοναδικός

- Η επαγγελματική κατάρτιση ενός μάγειρα δεν πρέπει να είναι NULL και πρέπει να είναι μία από τις επιλογές A Cook, B Cook, C Cook, Sous Chef και Chef. Το default status είναι C Cook
- Ο αριθμός του κάθε επεισοδίου πρέπει να είναι μεταξύ 1 και 10
- Το έτος των σεζόν πρέπει να είναι μετά το 2000
- Το id του κριτή πρέπει να είναι 1,2 ή 3
- Η βαθμολογία που παίρνει ένας μάγειρας από τους κριτές πρέπει να είναι ακέραιος μεταξύ 0 και 5

## Ευρετήρια

Στην βάση προσθέτουμε μερικά ευρετήρια ώστε τα queries να εκτελούνται πιο αποδοτικά. Αποφύγαμε να τοποθετήσουμε πολλαπλά ευρετήρια σε κάθε table καθώς είναι κοστοβόρα τόσο από άποψη χώρου όσο από άποψη χρόνου, δεδομένου ότι όταν κάνουμε αλλαγές στην βάση πρέπει και αυτά να αλλάζουν. Συγκεκριμένα, προσθέσαμε τα εξής ευρετήρια:

- Στον πίνακα recipes πάνω στο attribute recipe\_name δεδομένου ότι τα περισσότερα queries θα θέλουν να αναζητήσουν συγκεκριμένες συνταγές και τα χαρακτηριστικά τους, ή να κάνουν join χρησιμοποιώντας το όνομα της συνταγής
- Στον πίνακα tags πάνω στο recipe\_name εφόσον το όνομα της συνταγής χρησιμοποιείται για joins. Επιπλέον, τα περισσότερα queries ενδέχεται να αναζητούν τις ετικέτες συγκεκριμένων συνταγών
- Στους πίνακες requires\_eq, requires\_ingr και main\_ingr πάνω στο recipe\_name δεδομένου ότι, όπως και προηγουμένως, τα περισσότερα queries κάνουν select και join πάνω σε ονόματα συνταγών

- Στον πίνακα ingredients πάνω στο ingr\_name εφόσον η αναζήτηση υλικών γίνεται συνήθως με το όνομά τους. Επιπλέον, όπως και πριν, μπορεί να χρησιμοποιηθεί για join operations
- Στους πίνακες cook και is\_a\_critic πάνω στο επίθετο του μάγειρα επειδή μας είναι χρήσιμο για τα queries που ζητούνται
- Στον πίνακα is\_a\_contestant πάνω στους μάγειρες και στις χώρες, εφόσον πολλά ερωτήματα απαιτούν στατιστικά πάνω σε αυτά τα attributes

## Triggers

Στο DDL.sql προσθέσαμε μερικούς triggers που βοηθάνε στην σωστή λειτουργία της βάσης. Συγκεκριμένα:

1. Το trigger cook\_age ενεργοποιείται σε κάθε εισαγωγή ενός νέου μάγειρα προκειμένου να θέσει την ηλικία του ως την διαφορά της ημερομηνίας γέννησής του από την τρέχουσα ημερομηνία
2. Το trigger initialize\_calories ενεργοποιείται όταν προσθέτουμε μία συνταγή και αρχικοποιεί τις θερμίδες του σε 0
3. Το trigger recipe\_calories ενεργοποιείται όταν προσθέτουμε ένα νέο υλικό σε μία συνταγή προκειμένου να ενημερώσουμε τις θερμίδες της συνταγής. Με άλλα λόγια, πολλαπλασιάζουμε την ποσότητα του υλικού με τις θερμίδες του και προσθέτουμε στις ήδη υπάρχουσες θερμίδες της συνταγής. Αν η ποσότητα είναι σε gr ή ml τότε πρέπει πρώτα να διαιρέσουμε και με 100

4. Το trigger `some_quantity` θέτει σε NULL την ποσότητα ενός υλικού αν χρησιμοποιούμε απροσδιόριστες ποσότητες

## Stored functions

Στο `DDL.sql` περιλαμβάνεται και ένα stored function, το `status_to_int()` που μετατρέπει την επαγγελματική κατάρτιση ενός μάγαιρα σε ακέραιο. Η συνάρτηση αυτή χρησιμοποιείται στο ερώτημα 3.13

## Stored procedures

Τέλος, το `DDL.sql` περιέχει stored procedures με τα οποία κάνουμε την κλήρωση των επεισοδίων.

Το `create_episode1()` χρησιμοποιείται για να συμπληρώσει ένα επεισόδιο στους πίνακες `is_a_contestant` και `is_a_critic`. Συγκεκριμένα λαμβάνει ως παράμετρο το έτος και τον αριθμό του επεισοδίου. Έπειτα, επιλέγουμε 10 τυχαίες χώρες και με την βοήθεια ενός κέρσορα διατρέχουμε σε έναν βρόχο κάθε μία από αυτές. Για κάθε μία από τις 10 χώρες επιλέγουμε τυχαία μία συνταγή αυτή της χώρας και έναν μάγαιρα που ειδικεύεται σε αυτήν και προσθέτουμε αυτές τις τιμές στον πίνακα `is_a_contestant`. Παράλληλα, τοποθετούμε τυχαία τρεις αριθμούς από το 0 ως 5 για κάθε μία από τις βαθμολογίες του διαγωνιζόμενου. Τέλος, για το `is_a_critic` αρκεί να επιλέξουμε πάλι τυχαία τρεις μάγαιρες και να τους τοποθετήσουμε στο συγκεκριμένο επεισόδιο.

Το `create_episode2()` είναι ίδιο με το `create_episode1()` με την διαφορά ότι αποτρέπει την εμφάνιση ενός κριτή/μάγαιρα/χώρας/συνταγής πάνω από τρεις συνεχόμενες φορές σε επεισόδια. Για αυτόν τον σκοπό, τα queries που χρησιμοποιούνται από τους κέρσορες και τις τυχαίες επιλογές δεδομένων

είναι πιο πολύπλοκα και αποκλείουν δεδομένα που έχουν εμφανιστεί στα δύο προηγούμενα επεισόδια του διαγωνισμού.

Τέλος, το `create_season()` είναι το stored procedure που παράγει μία καινούρια σεζόν. Πρώτα, προσθέτει τα 10 καινούρια επεισόδια στο `episodes table`. Έπειτα, καλεί τις συναρτήσεις `create_episode1()` και `create_episode2()` για να γεμίσει τους πίνακες `is_a_contestant` και `is_a_critic`. Συγκεκριμένα, εκτελείται ένας βρόχος 10 φορές (μία για κάθε επεισόδιο). Στις δύο πρώτες επαναλήψεις καλούμε την `create_episode1()` για να δημιουργήσουμε τα δύο πρώτα επεισόδια, ενώ στις υπόλοιπες επαναλήψεις καλούμε την `create_season2()` ώστε να αποφύγουμε πολλαπλές εμφανίσεις δεδομένων. Να σημειωθεί ότι ο πίνακας `current_year` είναι βοηθητικός και αποθηκεύει το τελευταίο έτος στο οποίο έγινε ο διαγωνισμός ξεκινώντας από το 2023. Σε κάθε κλήση της `create_season()` ο αριθμός αυτός αυξάνεται κατά 1.

## Παραδοχές

Λόγω της ελευθερίας στην επιλογή διαφόρων σχεδιαστικών θεμάτων της βάσης κάναμε τις παρακάτω παραδοχές:

- Ο κάθε μάγειρας έχει εξειδίκευση σε συγκεκριμένες εθνικές κουζίνες και μπορεί να εκτελέσει συνταγές μόνο από τις κουζίνες αυτές και καμία άλλη.
- Για κάθε συνταγή το `recipe_type` εκτός από το συνταγή μαγειρικής ή ζαχαροπλαστικής μπορεί να είναι και ποτό (`beverage`), το οποίο προσθέσαμε για λόγους πληρότητας.
- Θεωρούμε ότι στο `is_a_critic` ο μάγειρας με `id = 1`, είναι αυτός που αναθέτει το `grade1`, αυτός με `id = 2`, αυτός που αναθέτει το `grade2` και αυτός με `id = 3`, το `grade3`.
- Εάν ένα υλικό σε μια συνταγή δεν έχει σαφώς ορισμένη ποσότητα (`Some`, `Little`, `Much`, `A lot`, `pinch`, `to taste`, `as needed`, `to garnish`, `slices`, `stalks`, `cloves`), τότε δεν το συμπεριλαμβάνουμε στον υπολογισμό των συνολικών θερμίδων (ανά μερίδα) της συνταγής.
- Τα `grades` που ανατίθενται σε κάθε μάγειρα δημιουργούνται στις `create_episode1` και `create_episode2` και είναι τυχαίοι αριθμοί.
- Έχουμε επιπλέον δημιουργήσει τα `views episode_winner` και `season_winner` που επιστρέφουν τον μάγειρα με την υψηλότερη βαθμολογία, σε κάθε επεισόδιο και κάθε σεζόν ανάλογα. Σε περίπτωση ισοβαθμίας νικητής είναι ο μάγειρας με την υψηλότερη επαγγελματική κατάρτιση. Σε περίπτωση νέας ισοβαθμίας, ο νικητής ανακηρύσσεται τυχαία. Τα `views` αυτά βρίσκονται στο `DDL.sql`.