

ΠΡΟΧΩΡΗΜΕΝΑ ΘΕΜΑΤΑ
ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΕΦΑΡΜΟΓΩΝ
ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΕΡΓΑΣΙΑ ΓΙΑ ΤΟ
ΑΚΑΔΗΜΑΪΚΟ ΈΤΟΣ 2022 - 2023

ΟΜΑΔΑ 4388_4409_4590

ΚΩΝΣΤΑΝΤΙΝΟΣ ΚΪΤΣΙΟΣ, 4388

ΑΝΑΣΤΑΣΙΟΣ ΛΙΟΝΤΟΣ, 4409

ΠΑΝΑΓΪΩΤΗΣ ΒΑΡΔΆΡΗΣ, 4590

ΤΕΛΙΚΗ ΑΝΑΦΟΡΑ

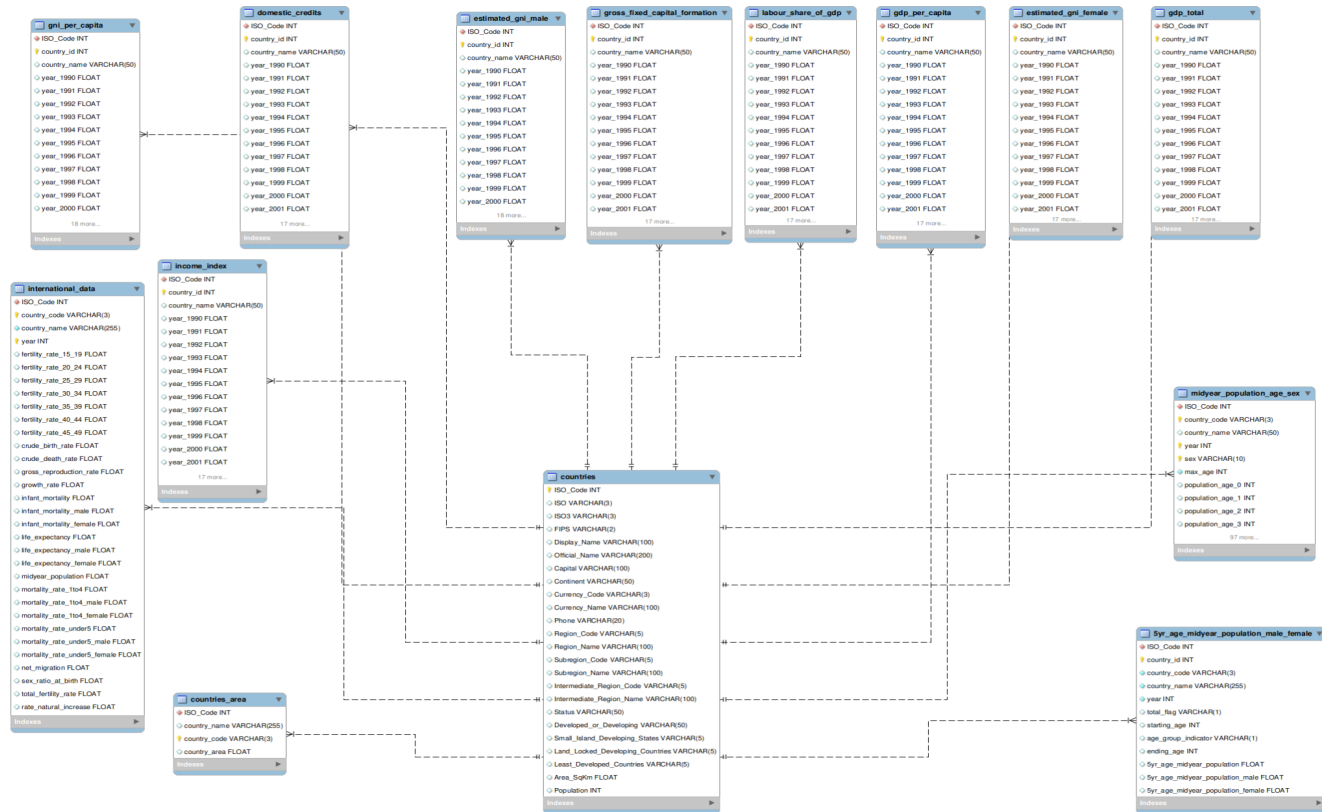
ΙΣΤΟΡΙΚΟ ΠΡΟΗΓΟΥΜΕΝΩΝ ΕΚΔΟΣΕΩΝ

Ημερομηνία	Έκδοση	Περιγραφή	Συγγραφέας
29/05/2023	0.1	Ανάλυση της διαδικασίας ETL & USE CASES	4388,4409,4590
30/05/2023	0.2	UML's, backup of database	4388,4409,4590
31/05/2023	1.0	Manual of the application, UML's complete	4388,4409,4590

1. ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Στην παρούσα ενότητα περιγράφονται τα σχήματα της βάσης (ή βάσεων, αν είναι παραπάνω από μία) δεδομένων που χρησιμοποιούνται στο project.

1.1. ΣΧΕΣΙΑΚΟ ΣΧΗΜΑ ΣΕ ΛΟΓΙΚΟ ΕΠΙΠΕΔΟ



Σχήμα 1.1 Σχεσιακό σχήμα της βάσης δεδομένων του συστήματος

2. ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΛΟΓΙΣΜΙΚΟΥ

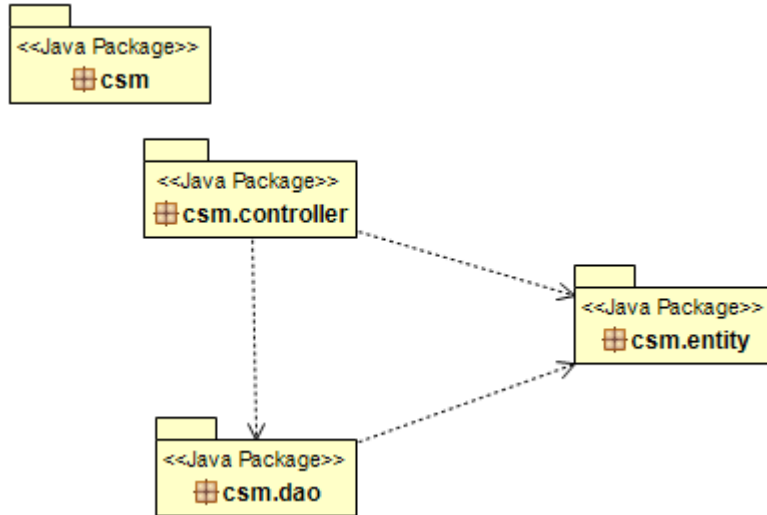
2.1. ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΔΟΜΗ ETL

Αρχικά πρέπει να φέρουμε τα δεδομένα μέσα στη βάση μας για περαιτέρω επεξεργασία. Σε πρώτο στάδιο, φτιάξαμε ETL scripts σε python, τα οποία είναι υπεύθυνα για τον μετασχηματισμό των αρχείων του dataset. Αρχικά, για κάθε αρχείο του dataset βρήκαμε ποιες χώρες ανήκουν σε όλα τα αρχεία πλην του countries.csv με διαφορετικό όνομα για να τις διορθώσουμε. Έπειτα κατά τη φόρτωση των δεδομένων που θα δούμε παρακάτω, εκτελέσαμε SQL ερωτήματα της μορφής: `SELECT * FROM table_name EXCEPT SELECT * FROM countries`. Έπειτα ελέγχουμε αν υπάρχουν data τα οποία πρέπει να αγνοήσουμε (ονόματα ηπείρων αντί χωρών κλπ) και τέλος παράγουμε τόσα csv αρχεία όσοι και οι δείκτες που περιλαμβάνονται στο αρχικό αρχείο. Τα αρχεία αυτά έχουν στήλες της μορφής: `country_code, country_name, year, metric`. Στο αρχείο `Income by Country.xlsx` κάνουμε ακριβώς την ίδια διαδικασία και επιπλέον σε όποιες στήλες δεν υπάρχουν data (αντί για τιμή υπάρχει “.”) κάνουμε interpolation (`interpolation.py`) για να υπάρχουν τιμές για κάθε στήλη. Από το συγκεκριμένο αρχείο παράγονται 9 csv αρχεία, όσα και τα φύλλα του αρχικού αρχείου.

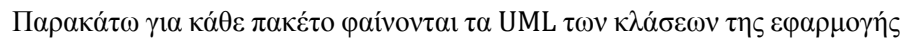
Στη συνέχεια δημιουργούμε προσωρινούς πίνακες στη βάση που έχουν τη μορφή των παραγόμενων αρχείων στη φάση του μετασχηματισμού (`aux_tmp_table_name`). Έπειτα με `insert` ερωτήματα βάζουμε τα περιεχόμενα των `aux_tmp tables` σε `aux tables` με τη διαφορά πως οι `aux` πίνακες έχουν ως `primary key` το `country_id`. Για κάθε πίνακα `aux` κάνουμε στη συνέχεια `join` με τον πίνακα `countries` με συνθήκη το `Display_Name` του πίνακα `countries` να είναι ίσο με το `country_name` του εκάστοτε πίνακα και προσθέτουμε επίσης στο `primary key` το `year` και το `ISO_Code` ως `foreign key` στον πίνακα `countries`. Τέλος στους πίνακες που προκύπτουν από το `join` κάνουμε `load infile` το αντίστοιχο αρχείο csv του dataset. Τα αρχεία csv που θα φορτωθούν με `load infile` πρέπει να βρίσκονται στο φάκελο `/var/lib/mysql-files/`

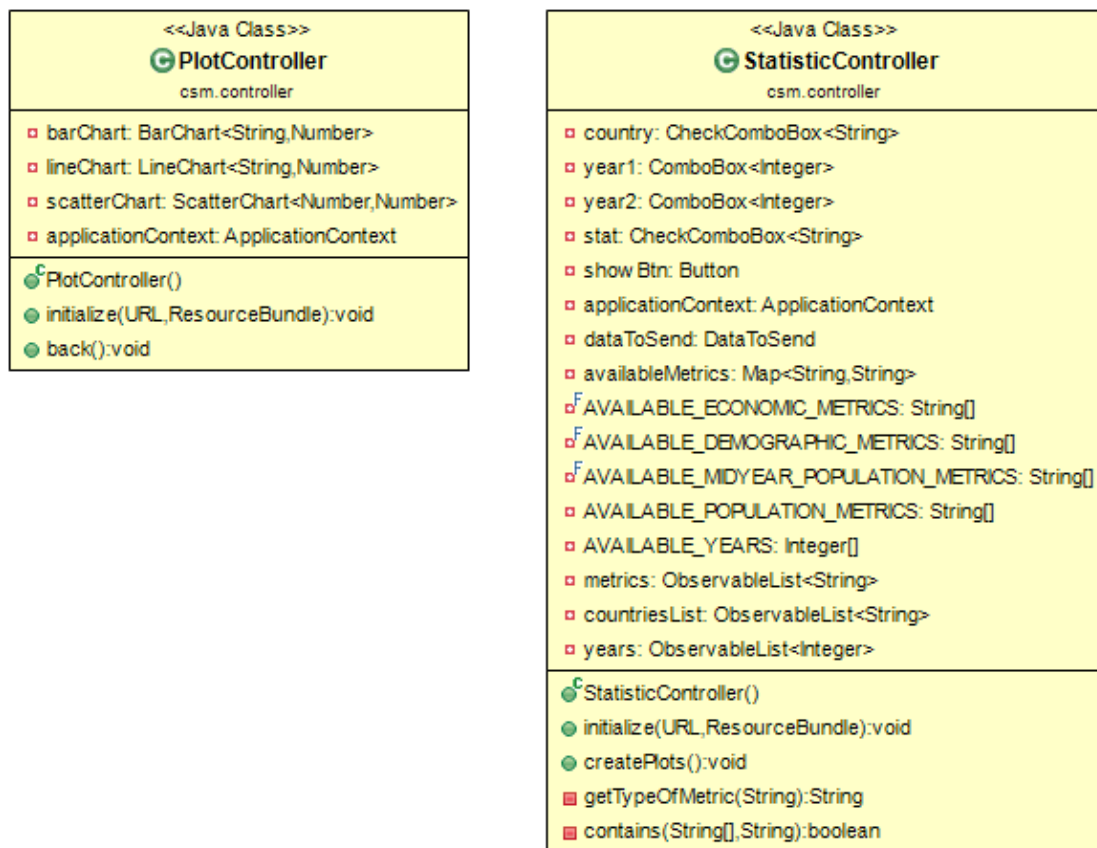
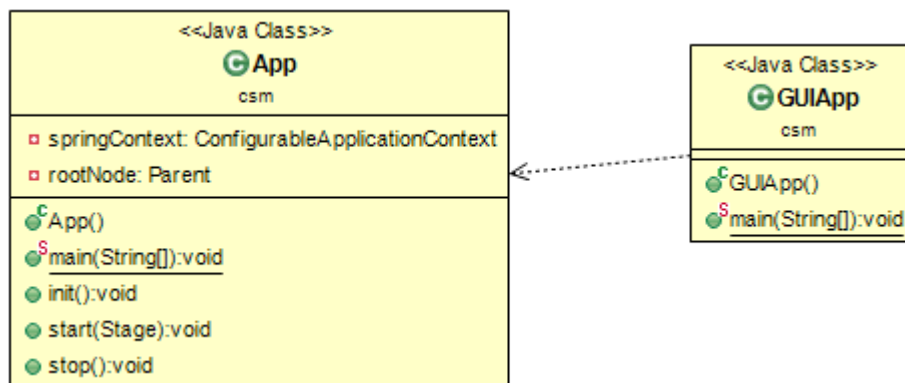
2.2. ΔΙΑΓΡΑΜΜΑΤΑ ΠΑΚΕΤΩΝ / ΥΠΟΣΥΣΤΗΜΑΤΩΝ ΚΕΝΤΡΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

Το διάγραμμα για τα υποσυστήματα / πακέτα του λογισμικού που κατασκευάσαμε ως κεντρική εφαρμογή επερώτησης. Ο στόχος είναι να φανεί η high-level αρχιτεκτονική του συστήματος, χωρίς λεπτομέρειες των επί μέρους κλάσεων και η μεταξύ τους σύνδεση.



Σχήμα 2.2 UML για τα πακέτα της εφαρμογής.



Σχήμα 2.3.3: UML για το πακέτο `csm.Controller`Σχήμα 2.3.4: UML για το πακέτο `csm.App`

3. USE CASES ΕΦΑΡΜΟΓΗΣ

Παρακάτω περιγράφονται τα use cases της εφαρμογής.

Use case ID	UC0: Show Plots
Actors	User
Pre conditions	1. The user must select the year range, country (or countries) and metric (or metrics).
Main flow of events	1. The use case starts when the user clicks on show plots button on the app. 2. The system displays the plots based on user's input
Alternative flow 1	1. If the user inputs years that don't have data for the selected country(countries) and metric(metrics) 1.1. The system displays an error message. 1.2. The system displays the data selection window.
Post conditions	-

Use case ID	UC1: Back
Actors	User
Pre conditions	1. The use must be in plots window.
Main flow of events	1. The use case starts when the user clicks on "Back" button on the plots window. 2. The system displays the first window of the app.
Post conditions	-

Use case ID	UC1: Exit
Actors	User
Pre conditions	-
Main flow of events	1. The use case starts when the user at any time clicks on the "x" button on the top of the window 2. The system closes the application.
Post conditions	-

4. ΥΠΟΔΕΙΓΜΑΤΑ ΕΡΩΤΗΣΕΩΝ ΚΑΙ ΑΠΑΝΤΗΣΕΩΝ

Καλωσήρθατε στο εγχειρίδιο χρήσης της εφαρμογής Στατιστικών Διαγραμμάτων. Αυτή η εφαρμογή σας επιτρέπει να επιλέγετε χρονιές, χώρες και διάφορα στατιστικά και να εμφανίζονται διαγράμματα ως αποτελέσματα. Ακολουθήστε τα παρακάτω βήματα για να ξεκινήσετε:

Βήμα 1: Επιλογή χρονιάς: Ανοίξτε την εφαρμογή και επιλέξτε την επιθυμητή χρονιά από τη διαθέσιμη λίστα. Μπορείτε να επιλέξετε ένα εύρος χρόνων ανάλογα με τις ανάγκες σας.

Βήμα 2: Επιλογή χώρας: Στη συνέχεια, επιλέξτε τη χώρα για την οποία θέλετε να προβάλετε τα στατιστικά δεδομένα. Μπορείτε να περιηγηθείτε στη λίστα των διαθέσιμων χωρών και να επιλέξετε αυτή που σας ενδιαφέρει.

Βήμα 3: Επιλογή Στατιστικών: Στο επόμενο βήμα, επιλέξτε τα στατιστικά δεδομένα που θέλετε να αναλύσετε και να παρουσιάσετε με τη μορφή διαγραμμάτων. Μπορείτε να διαλέξετε ένα ή περισσότερα στατιστικά

Βήμα 4: Προβολή Διαγραμμάτων: Αφού ολοκληρώσετε τις παραπάνω επιλογές, πατήστε το κουμπί "Show Plots" για να δημιουργήσετε και να εμφανίσετε τα επιλεγμένα διαγράμματα. Οι διαγραμματικές παραστάσεις θα εμφανιστούν στην οθόνη σας, παρέχοντάς σας μια οπτική αναπαράσταση των στατιστικών δεδομένων που επιλέξατε.

Βήμα 5: Εξερεύνηση και Ανάλυση Μελετήστε τα διαγράμματα που δημιουργήθηκαν και αναλύστε τα στατιστικά δεδομένα που αντιπροσωπεύουν. Εξερευνήστε τις τάσεις, τις διαφορές και τα μοτίβα που εμφανίζονται στα διαγράμματα, και αποκτήστε μια καλύτερη κατανόηση των στατιστικών δεδομένων που σας ενδιαφέρουν.

Αυτά είναι τα βασικά βήματα για να χρησιμοποιήσετε την εφαρμογή Στατιστικών Διαγραμμάτων. Επαναλάβετε τα βήματα για να εξερευνήσετε περισσότερα δεδομένα ή να αλλάξετε τις επιλογές σας. Απολαύστε την ανάλυση των στατιστικών δεδομένων σας μέσω οπτικών παραστάσεων και αποκτήστε νέα εισαγωγή και κατανόηση των πληροφοριών σας.

5. ΛΟΙΠΑ ΣΧΟΛΙΑ

Για το backup, χρησιμοποιούμε την εντολή `mysqldump -u root -p world_data > db.sql` για να δημιουργήσουμε ένα SQL script το οποίο περιέχει εντολές για την δημιουργία του σχήματος, δηλαδή των τελικών πινάκων και την φόρτωση των δεδομένων. Στην αρχή του `db.sql` υπάρχει εντολή δημιουργίας της βάσης, σε περίπτωση που δεν υπάρχει.

Για την ανάπτυξη της εφαρμογής χρησιμοποιήθηκαν οι εξής τεχνολογίες:

- SpringBoot
- Maven
- JavaFX

Χρησιμοποιήσαμε επίσης την αρχιτεκτονική MVC και η βάση δεδομένων στήθηκε σε MySQL.