

ΠΡΟΧΩΡΗΜΕΝΑ ΘΕΜΑΤΑ
ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΕΦΑΡΜΟΓΩΝ
ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΕΡΓΑΣΙΑ ΓΙΑ ΤΟ
ΑΚΑΔΗΜΑΪΚΟ ΈΤΟΣ 2020-2021

ΟΜΑΔΑ ΠΥΡΑΥΛΟΣ

ΜΠΑΡΜΠΑΛΙΑΣ ΕΥΑΓΓΕΛΟΣ, 2766

ΖΕΡΒΑΣ ΔΗΜΗΤΡΙΟΣ, 2894

ΚΑΪΤΗΣ ΣΩΤΗΡΙΟΣ, 3236

ΤΕΛΙΚΗ ΑΝΑΦΟΡΑ

ΙΣΤΟΡΙΚΟ ΠΡΟΗΓΟΥΜΕΝΩΝ ΕΚΔΟΣΕΩΝ

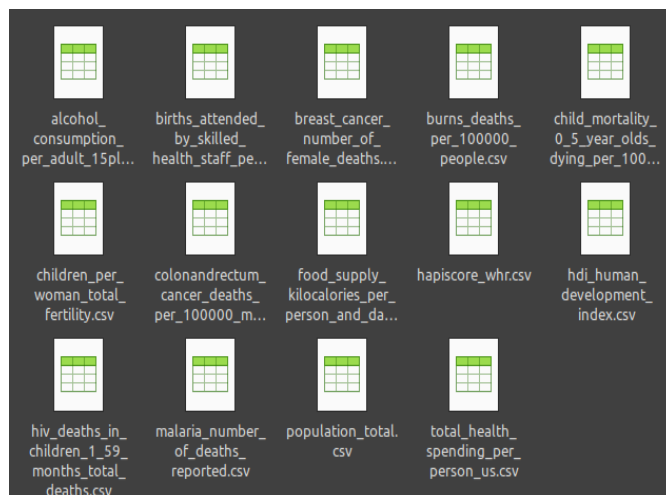
Ημερομηνία	Έκδοση	Περιγραφή	Συγγραφέας
2021/04/20	1.1		Μπαρμπαλιάς Ευάγγελος Ζέρβας Δημήτριος Καΐτης Σωτήριος
2021/05/5	1.2		Μπαρμπαλιάς Ευάγγελος Ζέρβας Δημήτριος Καΐτης Σωτήριος
2021/05/16	1.3		Μπαρμπαλιάς Ευάγγελος Ζέρβας Δημήτριος Καΐτης Σωτήριος
2021/05/27	1.4 (final)		Μπαρμπαλιάς Ευάγγελος Ζέρβας Δημήτριος Καΐτης Σωτήριος

1 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

1.1 ΣΧΕΣΙΑΚΟ ΣΧΗΜΑ ΣΕ ΛΟΓΙΚΟ ΕΠΙΠΕΔΟ

Από το script /CSV_PARSER/loadCSVtoDB.py δημιουργούμε μια νέα βάση δεδομένων **AdvancedDBs_v01** και ένα νέο table **Data** με primary keys τα πεδία (country,years).

Ο πίνακας Data περιέχει τα παρακάτω πεδία, όπου country είναι οι χώρες, όπου years τα χρόνια, και τα 1 VARCHAR(30) έως 14 VARCHAR(30), είναι ο κάθε index με την σειρά που έχει αποθηκευτεί στον φάκελο που γίνεται το parsing, δηλαδή
1 = alcohol_consumption_per_adult_15plus_litres
2 = births_attended_by_skilled_health_staff_percent_of_total
...



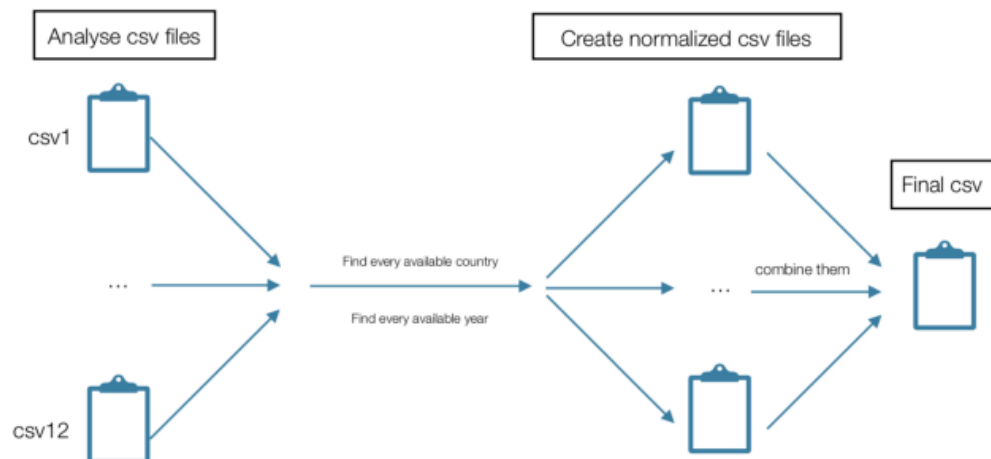
1.2 ΣΧΕΣΙΑΚΟ ΣΧΗΜΑ ΣΕ ΦΥΣΙΚΟ ΕΠΙΠΕΔΟ

Κατα την υλοποίηση της βάσης μας δε κάναμε καμία ιδέτερη ρύθμιση, παρά οτι χρησιμοποιήσαμε InnoDB τύπου αποθήκευση. Τη πρόσβαση στη βάση την έχει ο οποιοσδήποτε την έχει υλοποιήσει στο σύστημα. Από τον root Administrator μέχρι και τον απλό user.

2 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΛΟΓΙΣΜΙΚΟΥ

2.1 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΔΟΜΗ ETL

Για την εξαγωγή των δεδομένων από τα csv αρχεία καθώς και τη δημιουργία της βάσης, δημιουργήθηκαν 2 scripts.



- normalizeCSVfiles.py** : Το script αυτό χρησιμοποιεί τα csv αρχεία που είναι αποθηκευμένα στον φάκελο `CSV_PARSER/row_csv_data` προκειμένου να κανονικοποιήσει τα δεδομένα και να εξάγει ένα τελικό csv αρχείο. Πιο συγκεκριμένα, φέρνει όλα τα csv στην ίδια 'μορφή' δηλαδή να περιέχουν όλες τις δυνατές χώρες και χρονολογίες. Τα κανονικοποιημένα αρχεία παράγονται στον φάκελο `CSV_PARSER/NormalizedData`. Με βάση αυτά πλέον, παράγεται και το τελικό csv αρχείο, το οποίο έχει όλους δυνατούς συνδυασμούς χώρας-χρονολογίας (όπως φαίνεται στο επόμενο σχήμα).

```

country years  1  2  3  4  5  6  7  8  9 10 11 12 13 14
0  Afghanistan 1800 NaN NaN NaN NaN 469.0 7.00 NaN NaN NaN NaN NaN NaN 3280000 NaN
1  Afghanistan 1801 NaN NaN NaN NaN 469.0 7.00 NaN NaN NaN NaN NaN NaN 3280000 NaN
2  Afghanistan 1802 NaN NaN NaN NaN 469.0 7.00 NaN NaN NaN NaN NaN NaN 3280000 NaN
3  Afghanistan 1803 NaN NaN NaN NaN 469.0 7.00 NaN NaN NaN NaN NaN NaN 3280000 NaN
4  Afghanistan 1804 NaN NaN NaN NaN 469.0 7.00 NaN NaN NaN NaN NaN NaN 3280000 NaN
...
58690 Zimbabwe 2096 NaN NaN NaN NaN 14.1 1.84 NaN NaN NaN NaN NaN NaN 30900000 NaN
58691 Zimbabwe 2097 NaN NaN NaN NaN 13.9 1.83 NaN NaN NaN NaN NaN NaN 31000000 NaN
58692 Zimbabwe 2098 NaN NaN NaN NaN 13.8 1.83 NaN NaN NaN NaN NaN NaN 31000000 NaN
58693 Zimbabwe 2099 NaN NaN NaN NaN 13.7 1.83 NaN NaN NaN NaN NaN NaN 31000000 NaN
58694 Zimbabwe 2100 NaN NaN NaN NaN 13.7 1.83 NaN NaN NaN NaN NaN NaN 31000000 NaN

[58695 rows x 16 columns]

*** EXPORT FINAL CSV ***
    
```

- **loadCSVtoDB.py** : Το script αυτό, δημιουργεί τη βάση, εξάγει τα δεδομένα απ'το τελικό csv αρχείο και 'γεμίζει' τη βάση με αυτά. Πιο αναλυτικά, δημιουργεί έναν μοναδικό πίνακα στη βάση, όπως αυτός αναφέρεται στην ενότητα 1 και τα δεδομένα από το τελικό csv αρχείο εισάγονται σε αυτόν. Το μοντέλο της βάσης βρίσκεται στον φάκελο SQL_Models.

The screenshot shows the SQL Enterprise Manager interface. On the left, the 'SCHEMAS' pane displays a tree view with 'advanceddb_v01' expanded, showing 'Tables' (data), 'Views', 'Stored Procedures', and 'Functions'. Below this, 'mydatabase', 'sakila', 'sys', and 'world' are listed. The 'Administration' pane shows 'Schemas' selected, with 'Connection Details' for 'Local instance M' (Host: localhost, Port: 3306, Login User: root, Current User: root@localhost, SSL cipher: TLS_AES_256_GCM_SHA384). The main pane shows a query window with the following SQL code:

```
1 • USE advanceddb_v01;
2 • SELECT * FROM data;
```

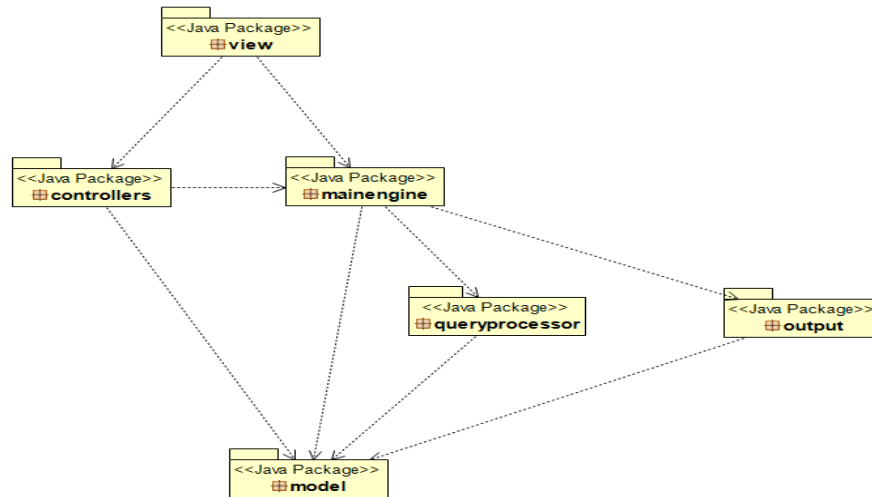
The 'Result Grid' shows the following data:

	country	years	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Congo, Rep.	1982					103.0	6.05		213...					1880000	
	Congo, Rep.	1983					99.7	5.96		211...					1940000	
	Congo, Rep.	1984					96.7	5.87		213...					2000000	
	Congo, Rep.	1985					94.0	5.77		213...					2050000	
	Congo, Rep.	1986					91.7	5.68		212...					2110000	
	Congo, Rep.	1987					90.0	5.59		216...					2170000	
	Congo, Rep.	1988					89.3	5.5		214...					2230000	
	Congo, Rep.	1989					89.2	5.42		210...					2290000	
	Congo, Rep.	1990			167.0	9.49	90.2	5.35	14.0	200...		0.531	185.0		2360000	
	Congo, Rep.	1991			166.0	9.11	91.8	5.29	13.6	200...		0.53	233.0		2420000	
	Congo, Rep.	1992			165.0	8.92	94.3	5.24	13.6	201...		0.53	283.0		2490000	
	Congo, Rep.	1993			171.0	8.9	97.3	5.2	13.8	200...		0.524	334.0		2560000	
	Congo, Rep.	1994			189.0	9.08	101.0	5.17	13.8	198...		0.521	382.0		2630000	
	Congo, Rep.	1995			200.0	9.11	105.0	5.14	13.9	198...		0.505	426.0		2710000	24.3
	Congo, Rep.	1996			214.0	9.09	108.0	5.13	13.8	197...		0.503	462.0		2790000	25.7
	Congo, Rep.	1997			221.0	9.04	112.0	5.12	13.7	199...		0.506	493.0		2870000	23.4
	Congo, Rep.	1998			219.0	8.68	115.0	5.11	13.3	205...		0.511	517.0		2950000	23.9
	Congo, Rep.	1999			226.0	8.48	115.0	5.11	12.9	210...		0.5	534.0		3040000	23.5
	Congo, Rep.	2000			226.0	8.03	114.0	5.1	12.2	218...		0.495	543.0		3130000	21.7
	Congo, Rep.	2001			220.0	7.5	111.0	5.09	11.6	223...		0.501	546.0		3220000	21.0
	Congo, Rep.	2002			217.0	7.06	106.0	5.08	11.1	226...		0.504	541.0		3310000	22.4
	Congo, Rep.	2003			227.0	6.92	100.0	5.07	11.1	224...		0.504	531.0		3410000	27.0
	Congo, Rep.	2004			236.0	6.71	93.2	5.05	10.9	226...		0.508	512.0		3510000	34.2

Τα δεδομένα στη βάση όπως αυτά εμφανίζονται μετά από *SELECT * FROM data;*

2.2 ΔΙΑΓΡΑΜΜΑΤΑ ΠΑΚΕΤΩΝ / ΥΠΟΣΥΣΤΗΜΑΤΩΝ ΚΕΝΤΡΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

Το μοντέλο αρχιτεκτονικής που χρησιμοποιήθηκε για τη δημιουργία της εφαρμογής είναι το Model-View-Controller (MVC). Στο μοντέλο αυτό η εφαρμογή διαιρείται σε τρία διασυνδεδεμένα μέρη ώστε να διαχωριστεί η παρουσίαση της πληροφορίας στον χρήστη από την μορφή που έχει αποθηκευτεί στο σύστημα.



- view package : Περιέχει τις boundary classes, υπεύθυνες για την αναπαράσταση της πληροφορίας που περιέχει το model, δημιουργώντας γραφική παρουσίαση στον χρήστη.
- mainengine package : Κεντρική business logic engine, με interface προς υλοποίηση.
- queryprocessor package : Υποσύστημα αλληλεπίδρασης με τη βάση δεδομένων.
- output package : Υποσύστημα για την δημιουργία διαγραμματικών αναπαραστάσεων των δεδομένων.
- model package : Domain classes του συστήματος.

2.3 ΔΙΑΓΡΑΜΜΑ(ΤΑ) ΚΛΑΣΕΩΝ ΚΕΝΤΡΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

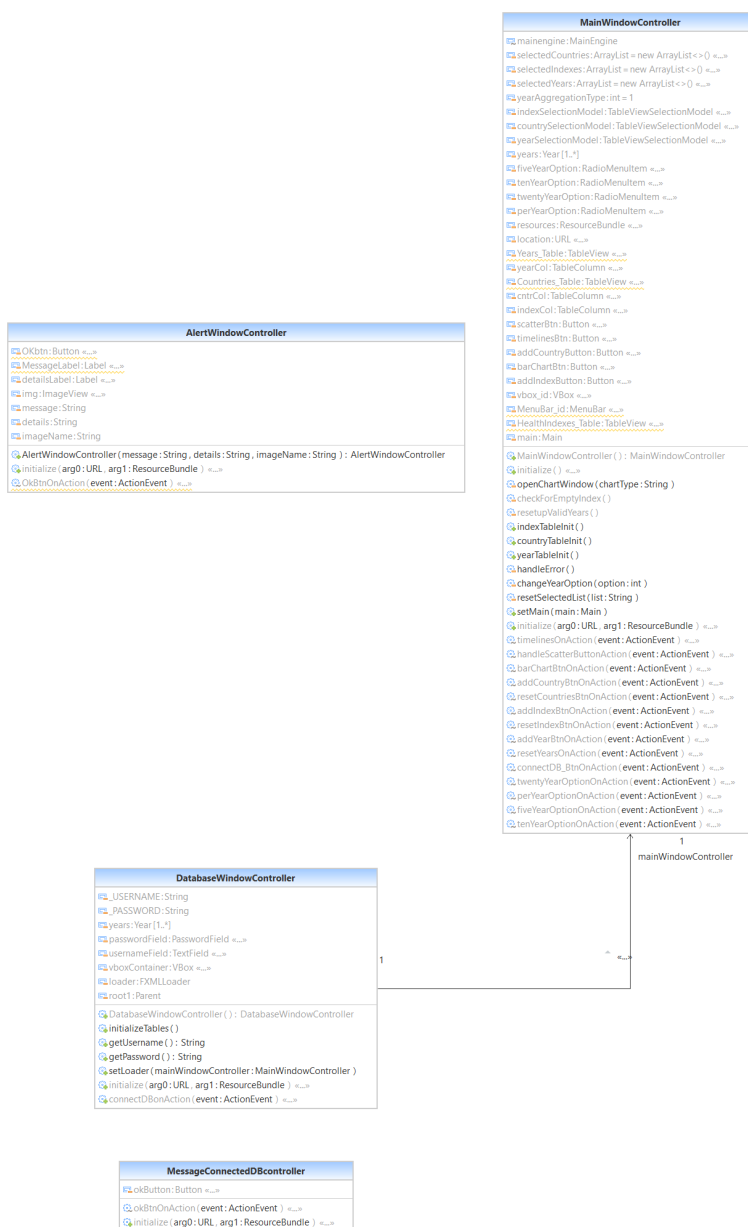
Η εφαρμογή οργανώνεται σε 6 πακέτα. Πιο συγκεκριμένα η εσωτερική λειτουργία της εφαρμογής και των πακέτων της είναι η εξής :

I. VIEW PACKAGE

Το παραπάνω πακέτο είναι υπεύθυνο για το γραφικό περιβάλλον. Περιλαμβάνεται η main καθώς και τα .fxml αρχεία για τα διάφορα windows της desktop εφαρμογής.

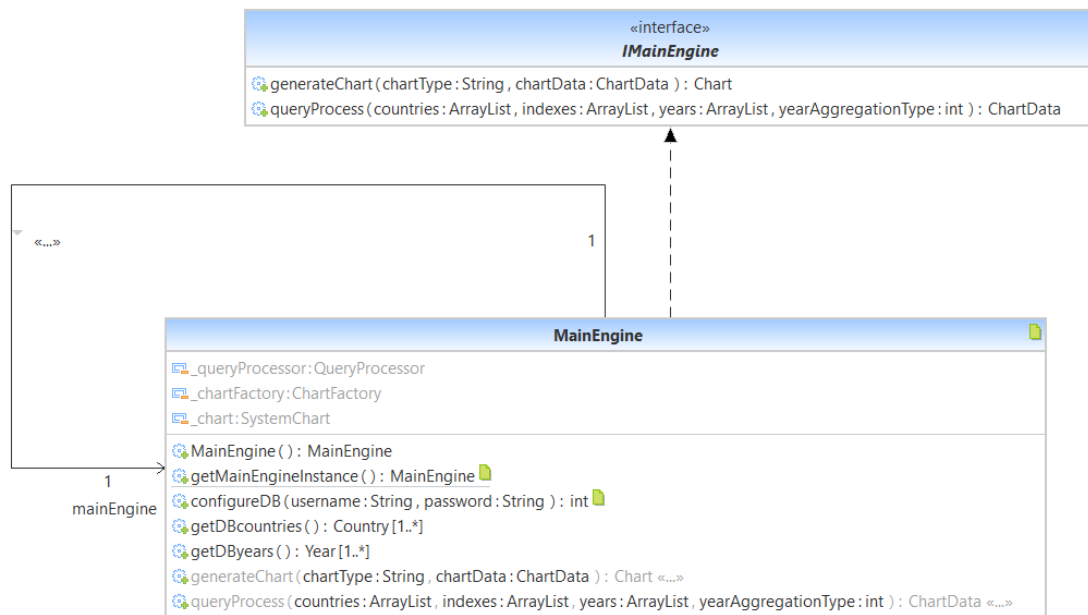
II. CONTROLLERS PACKAGE

Το controllers πακέτο περιέχει κλάσεις-controllers για τον χειρισμό των διάφορων components που απαρτίζουν το UI, καθώς και των διάφορων ενεργειών που υλοποιούνται κατά την αλληλεπίδραση του χρήστη με το UI.



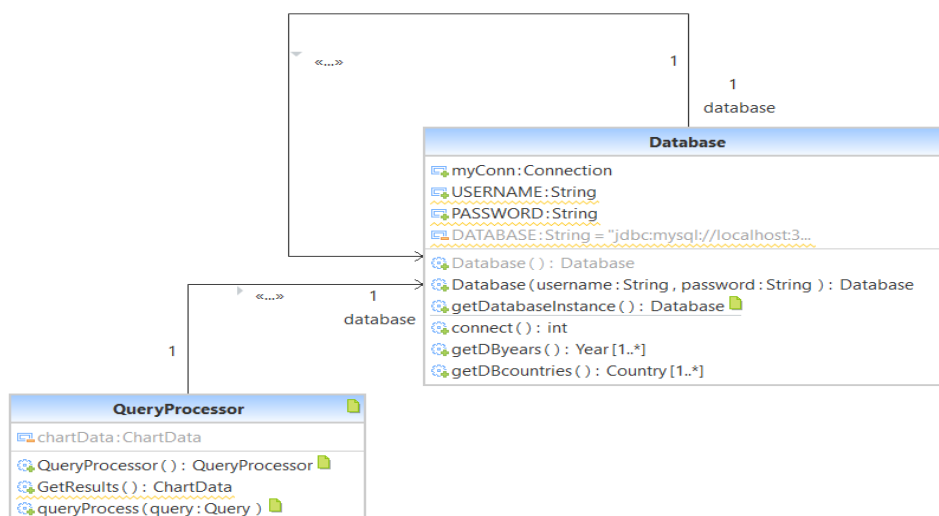
III. MAINENGINE PACKAGE

Το πακέτο που συνδέει το frontend με το backend. Οποιαδήποτε επιλογή στο UI μεταφέρεται στο backend για επεξεργασία μέσω των κλάσεων του mainengine πακέτου.



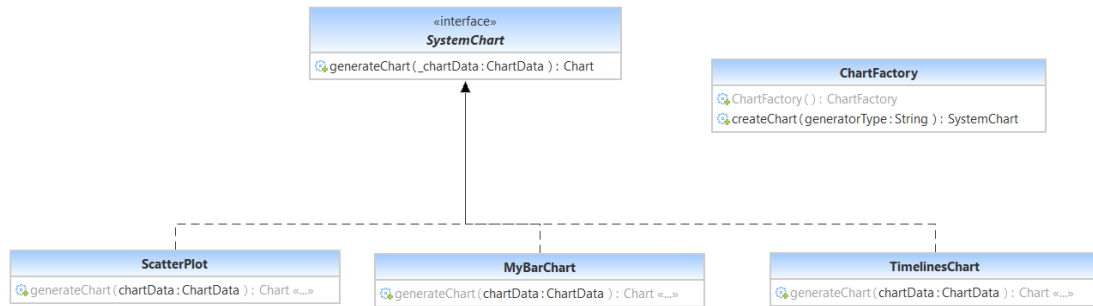
IV. QUERYPROCESSOR PACKAGE

Εδώ υλοποιείται η σύνδεση στη βάση δεδομένων, καθώς και η δημιουργία των queries για ερωτήσεις προς τη βάση. Επίσης εδώ επεξεργάζονται τα δεδομένα για διάφορες χρονολογικές περιόδους.



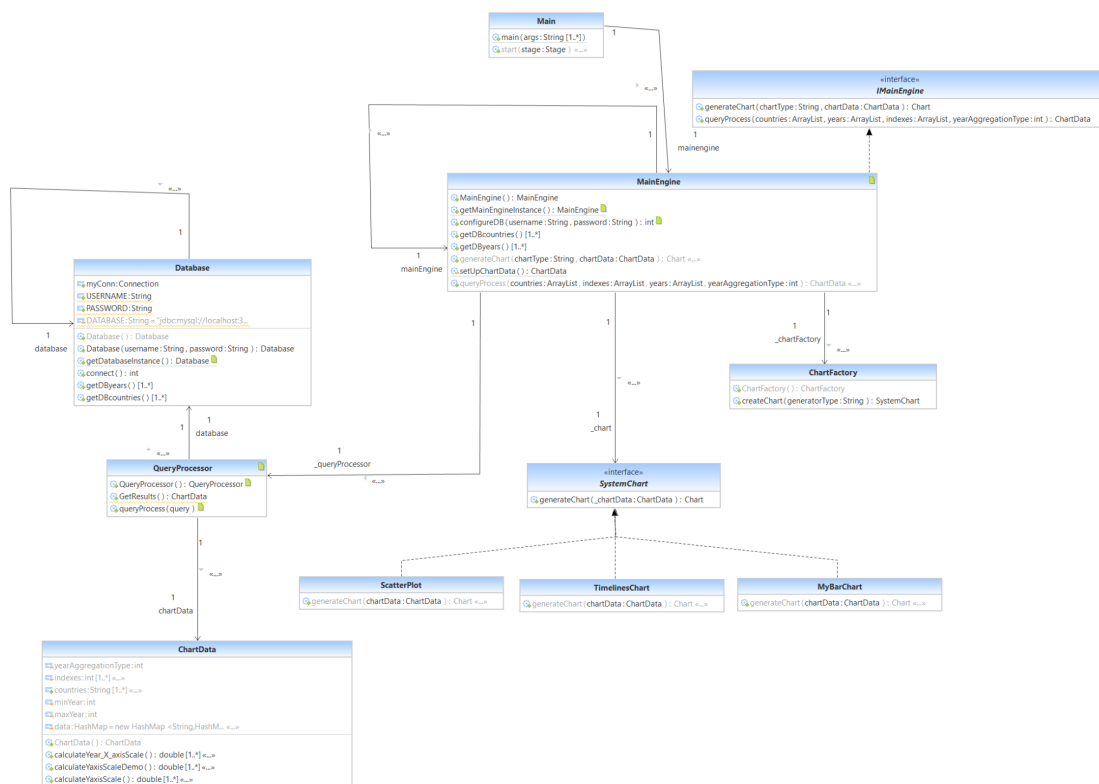
V. OUTPUT PACKAGE

Στο παραπάνω πακέτο υπάρχουν οι κλάσεις που είναι υπεύθυνες για την οπτικοποίηση των δεδομένων σε διάφορα είδη charts. Τα δεδομένα, μέσω της mainengine, φτάνουν από το πακέτο queryprocessor στο output και οπτικοποιούνται στα αντίστοιχα charts, ανάλογα με την επιλογή του χρήστη.



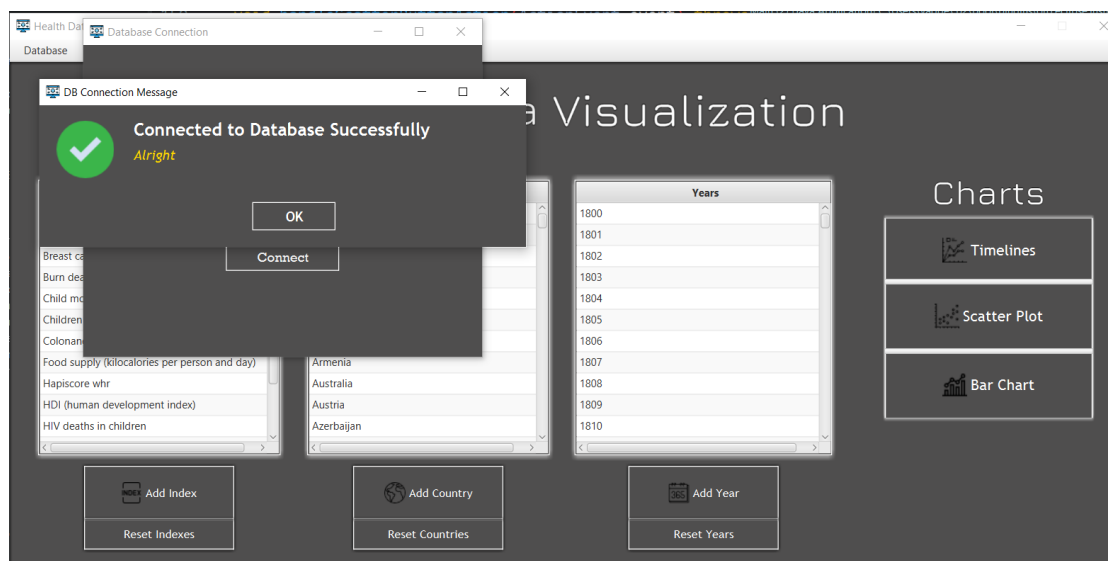
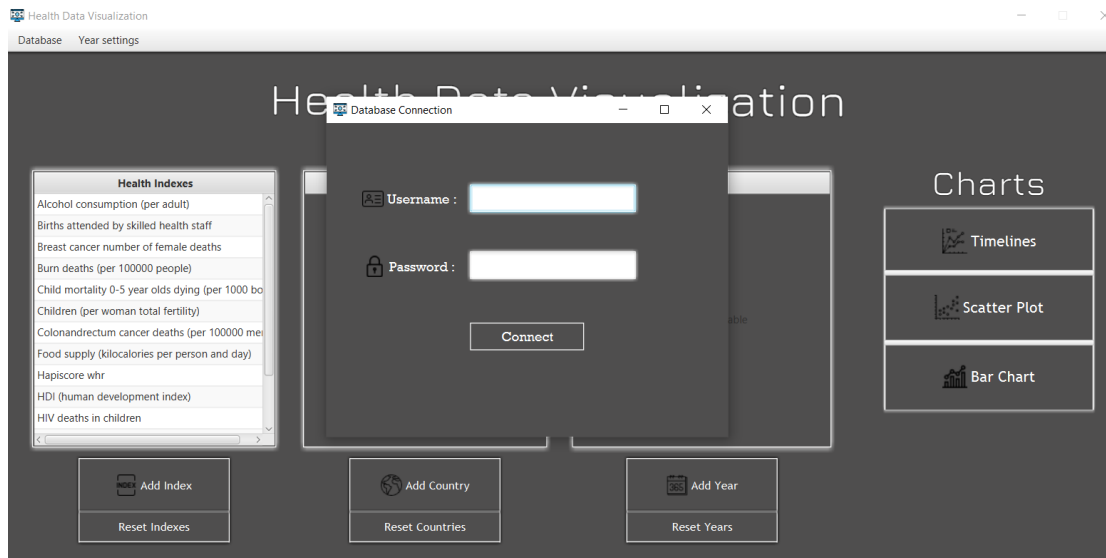
VI. MODEL PACKAGE

ΔΙΑΦΟΡΑ UMLs



3 ΥΠΟΔΕΙΓΜΑΤΑ ΕΡΩΤΗΣΕΩΝ ΚΑΙ ΑΠΑΝΤΗΣΕΩΝ

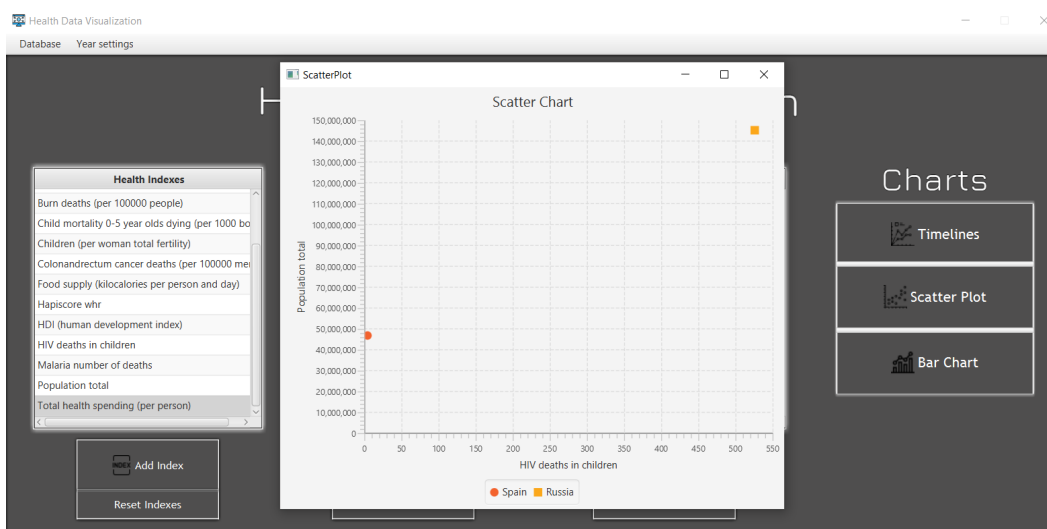
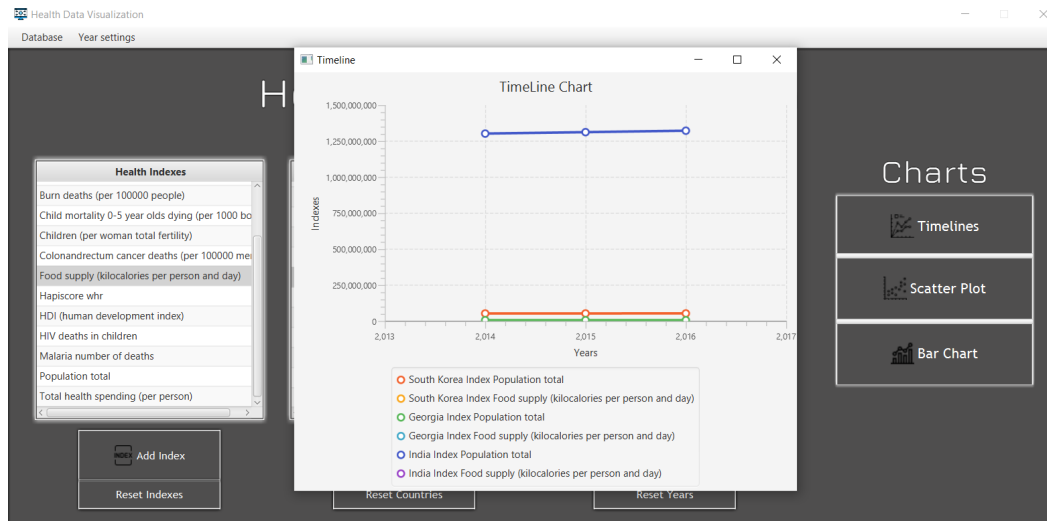
- ❖ Σύνδεση στη βάση : **Database -> Connect to DB**



- ❖ Επιλογή δεδομένων : Επιλογή index -> Add Index (Ένα ή περισσότερα indexes)
Επιλογή χωρών -> Add Country Επιλογή χρόνων -> Add Year
- ❖ Επιλογή εμφάνισης διαφορετικών χρονολογικών περιόδων :
Year Settings -> per year, per 5 years, per 10 years, per 20 years

❖ Οπτικοποίηση δεδομένων :

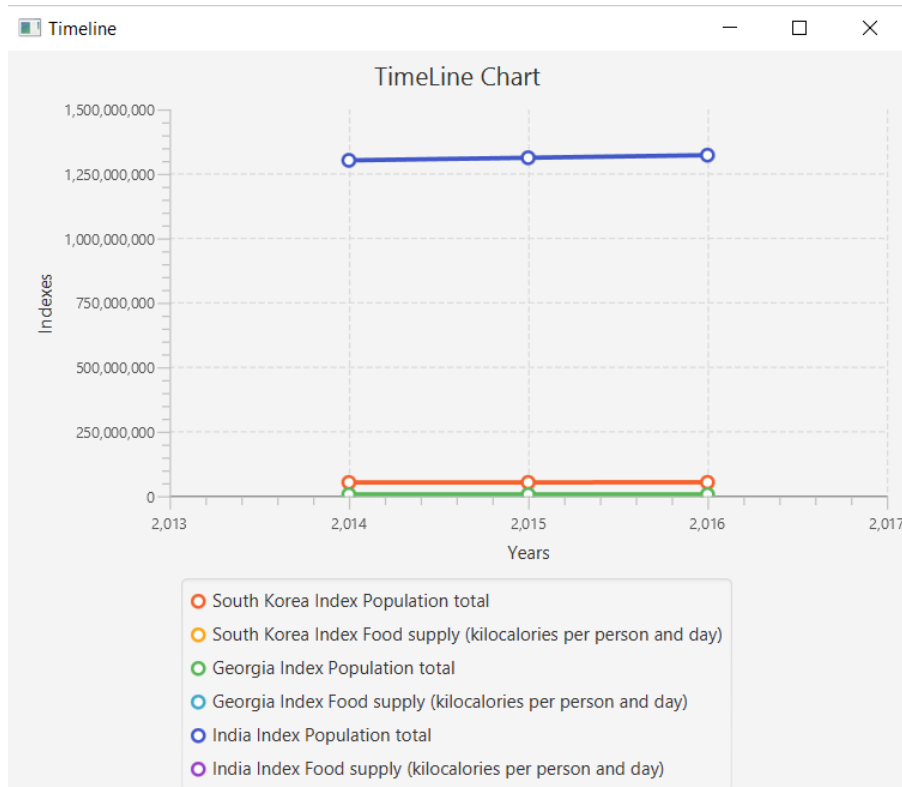
- Timelines : επιλογή κουμπιού Timelines
- Scatter : επιλογή κουμπιού Scatter Plot
- Bar Chart : επιλογή κουμπιού Bar Chart



4 ΤΕΚΜΗΡΙΩΣΗ ΚΑΙ ΛΟΙΠΑ ΣΧΟΛΙΑ

Αν ένα ολόκληρο γράφημα ή ένας δείκτης δεν εμφανιστεί σημαίνει ότι δεν υπάρχουν δεδομένα για τα δοθέντα δείκτη/χώρα/ημερομηνία.

Παράδειγμα κενών δεδομένων για μερικούς δείκτες:



Παράδειγμα τελείως κενών γραφημάτων :

