

ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

ΤΜΗΜΑ ΜΗΧ. Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ

ΠΡΟΧΩΡΗΜΕΝΑ ΘΕΜΑΤΑ
ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΕΦΑΡΜΟΓΩΝ
ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΕΡΓΑΣΙΑ ΓΙΑ
ΤΟ ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ 2022-2023

ΟΜΑΔΑ XXX

ΜΗΤΣΗΣ ΓΕΩΡΓΙΟΣ 4258

ΤΕΛΙΚΗ ΑΝΑΦΟΡΑ

ΜΑΪΟΣ 2023

ΙΣΤΟΡΙΚΟ ΠΡΟΗΓΟΥΜΕΝΩΝ ΕΚΔΟΣΕΩΝ

Ημερομηνία	Έκδοση	Περιγραφή	Συγγραφέας
2023/02/20	0.2	Μελέτη και δημιουργία ΒΔ	
2023/02/22	0,3	Σχεδιασμός Κλάσεων για SpringBoot	
2023/03/22	0.5	Δημιουργία Κλάσεων και πρώτης επαφής με την ΒΔ	
2023/04/01	1,0	Επανασχεδιασμός Με MVC model project σε c#	
2023/04/22	1.5	Μια πρώτη ολοκληρωμένη μορφή	
2023/05/20	2.0	Η τελική μορφή του project	

Λίγα λόγια για το project και το πως εργάστηκα.

Η βάση δεδομένων μου πήρε γύρω στις 2 εβδομάδες δεν δυσκολεύτηκα καθόλου.

Εκεί που δυσκολεύτηκα ήταν στην αρχική υλοποίηση σε Java και Spring Boot ενώ είχα μπορέσει να βλέπω δεδομένα από την βάση στον explorer δεν μπόρεσα να συνεχίσω και να το φτιάξω έτσι όπως ήθελα(είχα κάνει και procedures που γεννούσαν έτοιμες κλάσεις που με ένα copy paste ήταν έτοιμα τα models). Γι' αυτό το λόγο άρχισα από την αρχή σε C# επειδή είχε καλύτερο documentation.

Η αρχή μέχρι δηλαδή να φτιάξω τα models και τι από όλα να ακολουθήσω μου πήρε κοντά στον μήνα αλλά τελικά τα κατάφερα. Δημιούργησα ένα solution με δύο projects ένα library που έχει την πληροφορία της βάσης που την πέρασα με ADO και ένα άλλο .NET που περιέχει τους controllers και τις views.

Ο χρόνος μέσα στην μέρα που έχω ελεύθερος είναι μετά το βράδυ λόγω παιδιών και δουλειάς γι' αυτό και τα τόσο μεγάλα διαστήματα.

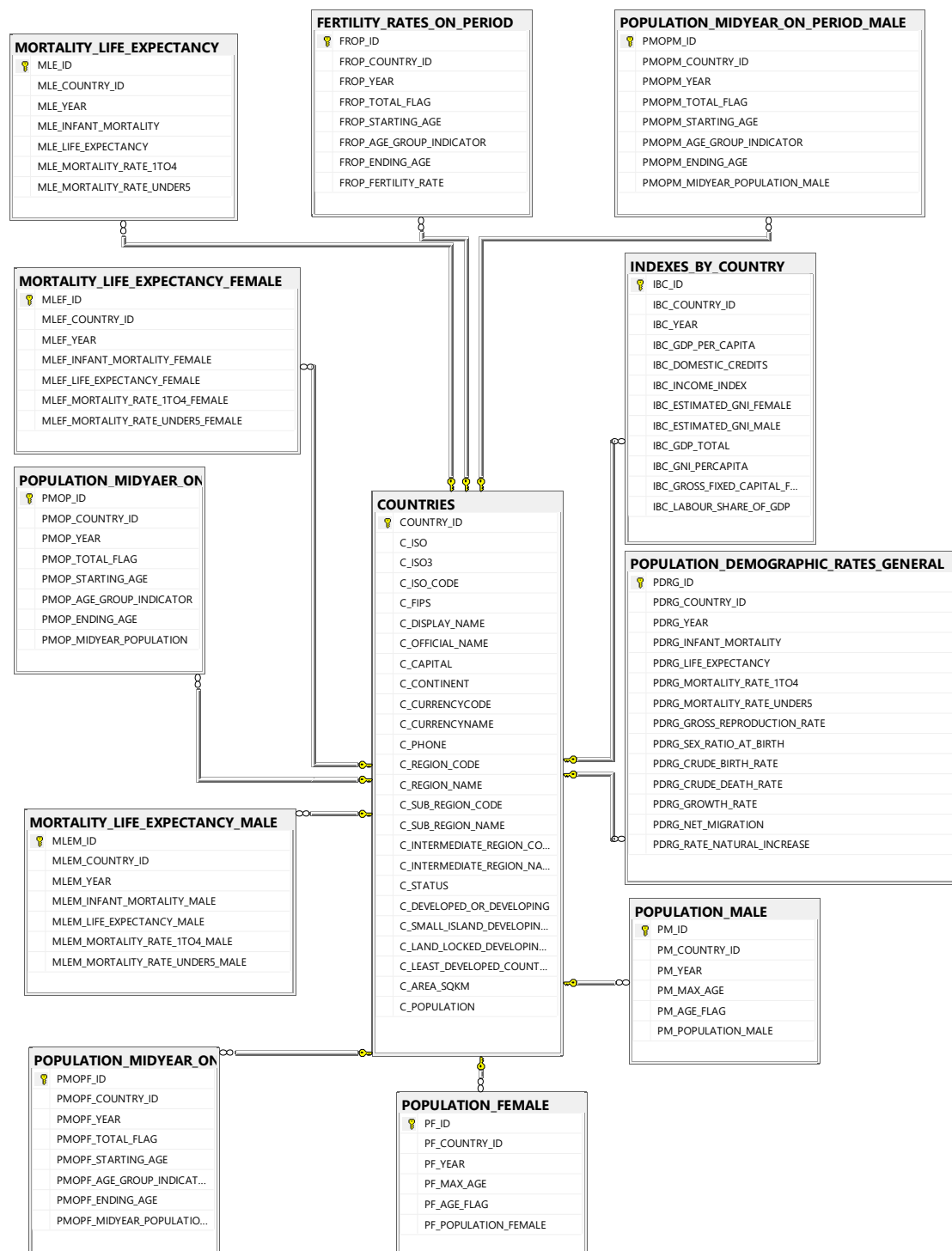
Καλό θα ήταν να έχω ομάδα αλλά σε δύο υποχρεωτικά μαθήματα τα παράτησε η ομάδα μου και έχασα εξάμηνα και έτσι επέλεξα να την κάνω μόνος μου.

Σε αυτή την εργασία πολέμησε το απόλυτο και το σωστό του μαθηματικού με το μοτίβο του τεχνικού ότι βάλτο μια τιμή να τελειώνουμε. Αν παραλάμβανα αυτά τα δεδομένα σε συνθήκες δουλειάς θα τα έστελνα πίσω μέχρι να τα στείλουν σωστά. Αν ήμουν στην αντίπερα όχθη δηλ αυτός που θα τα έστελνε, ναι θα έβαζα μια τιμή να τελειώνουμε. Βρήκα μια ισορροπία.

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

Section-break (continuous)

```
USE [DWH]
GO
/***** Object: Table [dbo].[COUNTRIES]
Script Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[COUNTRIES](
    [COUNTRY_ID] [int] IDENTITY(1,1)
    NOT NULL,
    [C_ISO] [varchar](2) NULL,
    [C_ISO3] [varchar](3) NULL,
    [C_ISO_CODE] [smallint] NULL,
    [C_FIPS] [varchar](2) NULL,
    [C_DISPLAY_NAME] [varchar](50)
    NULL,
    [C_OFFICIAL_NAME] [varchar](100)
    NULL,
    [C_CAPITAL] [varchar](50) NULL,
    [C_CONTINENT] [varchar](50) NULL,
    [C_CURRENCYCODE] [varchar](50)
    NULL,
    [C_CURRENCYNAME] [varchar](50)
    NULL,
    [C_PHONE] [varchar](50) NULL,
    [C_REGION_CODE] [tinyint] NULL,
    [C_REGION_NAME] [varchar](50) NULL,
    [C_SUB_REGION_CODE] [smallint]
    NULL,
    [C_SUB_REGION_NAME] [varchar](50)
    NULL,
    [C_INTERMEDIATE_REGION_CODE]
    [smallint] NULL,
    [C_INTERMEDIATE_REGION_NAME]
    [varchar](50) NULL,
    [C_STATUS] [varchar](50) NULL,
    [C_DEVELOPED_OR_DEVELOPING]
    [varchar](50) NULL,
    [C_SMALL_ISLAND_DEVELOPING_STATES_S
    IDS] [bit] NULL,
    [C_LAND_LOCKED_DEVELOPING_COUNTRIES
    _LLDC] [bit] NULL,
    [C_LEAST_DEVELOPED_COUNTRIES_LDC]
    [bit] NULL,
    [C_AREA_SQKM] [decimal](18, 4)
    NULL,
    [C_POPULATION] [int] NULL,
    CONSTRAINT [PK_COUNTRIES] PRIMARY KEY
    CLUSTERED
    (
        [COUNTRY_ID] ASC
    )WITH (PAD_INDEX = OFF,
    STATISTICS_NORECOMPUTE = OFF,
    IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
    ALLOW_PAGE_LOCKS = ON,
    OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
    [PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table [dbo].[ErrorLog]
Script Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[ErrorLog](
    [ID] [int] IDENTITY(1,1) NOT NULL,
    [ReferenceDate] [datetime2](7)
    NULL,
    [ErrorNumber] [int] NULL,
    [ErrorSeverity] [int] NULL,
    [ErrorState] [int] NULL,
```

```
    [ErrorProcedure] [varchar](200)
    NULL,
    [ErrorLine] [int] NULL,
    [ErrorMessage] [varchar](500) NULL
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[FERTILITY_RATES_ON_PERIOD] Script
Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE
[dbo].[FERTILITY_RATES_ON_PERIOD](
    [FROP_ID] [int] IDENTITY(1,1) NOT
    NULL,
    [FROP_COUNTRY_ID] [int] NOT NULL,
    [FROP_YEAR] [int] NOT NULL,
    [FROP_TOTAL_FLAG] [varchar](1)
    NULL,
    [FROP_STARTING_AGE] [int] NULL,
    [FROP_AGE_GROUP_INDICATOR]
    [varchar](1) NULL,
    [FROP_ENDING_AGE] [int] NULL,
    [FROP_FERTILITY_RATE] [decimal](18,
    4) NULL,
    CONSTRAINT [PK_FERTILITY_RATES_ON_PERIOD]
    PRIMARY KEY CLUSTERED
    (
        [FROP_ID] ASC
    )WITH (PAD_INDEX = OFF,
    STATISTICS_NORECOMPUTE = OFF,
    IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
    ALLOW_PAGE_LOCKS = ON,
    OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
    [PRIMARY],
    CONSTRAINT [UC_FERTILITY_RATES_ON_PERIOD]
    UNIQUE NONCLUSTERED
    (
        [FROP_COUNTRY_ID] ASC,
        [FROP_YEAR] ASC,
        [FROP_STARTING_AGE] ASC,
        [FROP_ENDING_AGE] ASC
    )WITH (PAD_INDEX = OFF,
    STATISTICS_NORECOMPUTE = OFF,
    IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
    ALLOW_PAGE_LOCKS = ON,
    OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
    [PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[INDEXES_BY_COUNTRY] Script Date:
12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[INDEXES_BY_COUNTRY](
    [IBC_ID] [int] IDENTITY(1,1) NOT
    NULL,
    [IBC_COUNTRY_ID] [int] NOT NULL,
    [IBC_YEAR] [int] NOT NULL,
    [IBC_GDP_PER_CAPITA] [decimal](18,
    4) NULL,
    [IBC_DOMESTIC_CREDITS]
    [decimal](18, 4) NULL,
    [IBC_INCOME_INDEX] [decimal](18, 4)
    NULL,
    [IBC_ESTIMATED_GNI_FEMALE]
    [decimal](18, 4) NULL,
    [IBC_ESTIMATED_GNI_MALE]
    [decimal](18, 4) NULL,
```

```

        [IBC_GDP_TOTAL] [decimal](18, 4)
NULL,
        [IBC_GNI_PERCAPITA] [decimal](18,
4) NULL,
        [IBC_GROSS_FIXED_CAPITAL_FORMATION]
[decimal](18, 4) NULL,
        [IBC_LABOUR_SHARE_OF_GDP]
[decimal](18, 4) NULL,
        CONSTRAINT [PK_INDEXES_BY_COUNTRY] PRIMARY
KEY CLUSTERED
(
        [IBC_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
        CONSTRAINT [UC_INDEXES_BY_COUNTRY] UNIQUE
NONCLUSTERED
(
        [IBC_COUNTRY_ID] ASC,
        [IBC_YEAR] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[INDEXES_BY_REGION]    Script Date:
12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[INDEXES_BY_REGION](
        [IBR_ID] [int] IDENTITY(1,1) NOT
NULL,
        [IBR_REGION] [varchar](256) NULL,
        [IBR_YEAR] [int] NOT NULL,
        [IBR_GDP_PER_CAPITA] [decimal](18,
4) NULL,
        [IBR_DOMESTIC_CREDITS]
[decimal](18, 4) NULL,
        [IBR_INCOME_INDEX] [decimal](18, 4)
NULL,
        [IBR_ESTIMATED_GNI_FEMALE]
[decimal](18, 4) NULL,
        [IBR_ESTIMATED_GNI_MALE]
[decimal](18, 4) NULL,
        [IBR_GDP_TOTAL] [decimal](18, 4)
NULL,
        [IBR_GNI_PERCAPITA] [decimal](18,
4) NULL,
        [IBR_GROSS_FIXED_CAPITAL_FORMATION]
[decimal](18, 4) NULL,
        [IBR_LABOUR_SHARE_OF_GDP]
[decimal](18, 4) NULL,
        CONSTRAINT [PK_INDEXES_BY_REGION] PRIMARY
KEY CLUSTERED
(
        [IBR_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
        CONSTRAINT [UC_INDEXES_BY_REGION] UNIQUE
NONCLUSTERED
(

```

```

        [IBR_REGION] ASC,
        [IBR_YEAR] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[MORTALITY_LIFE_EXPECTANCY]    Script
Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY](
        [MLE_ID] [int] IDENTITY(1,1) NOT
NULL,
        [MLE_COUNTRY_ID] [int] NOT NULL,
        [MLE_YEAR] [int] NOT NULL,
        [MLE_INFANT_MORTALITY]
[decimal](18, 4) NULL,
        [MLE_LIFE_EXPECTANCY] [decimal](18,
4) NULL,
        [MLE_MORTALITY_RATE_1TO4]
[decimal](18, 4) NULL,
        [MLE_MORTALITY_RATE_UNDE5]
[decimal](18, 4) NULL,
        CONSTRAINT [PK_MORTALITY_LIFE_EXPECTANCY]
PRIMARY KEY CLUSTERED
(
        [MLE_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
        CONSTRAINT [UC_MORTALITY_LIFE_EXPECTANCY]
UNIQUE NONCLUSTERED
(
        [MLE_COUNTRY_ID] ASC,
        [MLE_YEAR] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[MORTALITY_LIFE_EXPECTANCY_FEMALE]
Script Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY_FEMALE](
        [MLEF_ID] [int] IDENTITY(1,1) NOT
NULL,
        [MLEF_COUNTRY_ID] [int] NOT NULL,
        [MLEF_YEAR] [int] NOT NULL,
        [MLEF_INFANT_MORTALITY_FEMALE]
[decimal](18, 4) NULL,
        [MLEF_LIFE_EXPECTANCY_FEMALE]
[decimal](18, 4) NULL,
        [MLEF_MORTALITY_RATE_1TO4_FEMALE]
[decimal](18, 4) NULL,

```

```

        [MLEF_MORTALITY_RATE_UNDE5_FEMALE]
[decimal](18, 4) NULL,
        CONSTRAINT
[PK_MORTALITY_LIFE_EXPECTANCY_FEMALE]
PRIMARY KEY CLUSTERED
(
        [MLEF_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
        CONSTRAINT
[UC_MORTALITY_LIFE_EXPECTANCY_FEMALE]
UNIQUE NONCLUSTERED
(
        [MLEF_COUNTRY_ID] ASC,
        [MLEF_YEAR] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[MORTALITY_LIFE_EXPECTANCY_MALE]
Script Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY_MALE](
        [MLEM_ID] [int] IDENTITY(1,1) NOT
NULL,
        [MLEM_COUNTRY_ID] [int] NOT NULL,
        [MLEM_YEAR] [int] NOT NULL,
        [MLEM_INFANT_MORTALITY_MALE]
[decimal](18, 4) NULL,
        [MLEM_LIFE_EXPECTANCY_MALE]
[decimal](18, 4) NULL,
        [MLEM_MORTALITY_RATE_1TO4_MALE]
[decimal](18, 4) NULL,
        [MLEM_MORTALITY_RATE_UNDE5_MALE]
[decimal](18, 4) NULL,
        CONSTRAINT
[PK_MORTALITY_LIFE_EXPECTANCY_MALE] PRIMARY
KEY CLUSTERED
(
        [MLEM_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
        CONSTRAINT
[UC_MORTALITY_LIFE_EXPECTANCY_MALE] UNIQUE
NONCLUSTERED
(
        [MLEM_COUNTRY_ID] ASC,
        [MLEM_YEAR] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
```

```

/***** Object: Table
[dbo].[POPULATION_DEMOGRAPHIC_RATES_GENERAL]
Script Date: 12/3/2023 5:51:05 μμ
*****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE
[dbo].[POPULATION_DEMOGRAPHIC_RATES_GENERAL]
(
        [PDRG_ID] [int] IDENTITY(1,1) NOT
NULL,
        [PDRG_COUNTRY_ID] [int] NOT NULL,
        [PDRG_YEAR] [int] NOT NULL,
        [PDRG_INFANT_MORTALITY]
[decimal](18, 4) NULL,
        [PDRG_LIFE_EXPECTANCY]
[decimal](18, 4) NULL,
        [PDRG_MORTALITY_RATE_1TO4]
[decimal](18, 4) NULL,
        [PDRG_MORTALITY_RATE_UNDE5]
[decimal](18, 4) NULL,
        [PDRG_GROSS_REPRODUCTION_RATE]
[decimal](18, 4) NULL,
        [PDRG_SEX_RATIO_AT_BIRTH]
[decimal](18, 4) NULL,
        [PDRG_CRUDE_BIRTH_RATE]
[decimal](18, 4) NULL,
        [PDRG_CRUDE_DEATH_RATE]
[decimal](18, 4) NULL,
        [PDRG_GROWTH_RATE] [decimal](18, 4)
NULL,
        [PDRG_NET_MIGRATION] [decimal](18,
4) NULL,
        [PDRG_RATE_NATURAL_INCREASE]
[decimal](18, 4) NULL,
        CONSTRAINT
[PK_POPULATION_DEMOGRAPHIC_RATES_GENERAL]
PRIMARY KEY CLUSTERED
(
        [PDRG_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[POPULATION_FEMALE]
Script Date:
12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[POPULATION_FEMALE](
        [PF_ID] [int] IDENTITY(1,1) NOT
NULL,
        [PF_COUNTRY_ID] [int] NOT NULL,
        [PF_YEAR] [int] NOT NULL,
        [PF_MAX_AGE] [int] NULL,
        [PF_AGE_FLAG] [int] NULL,
        [PF_POPULATION_FEMALE]
[decimal](18, 4) NULL,
        CONSTRAINT [PK_POPULATION_FEMALE] PRIMARY
KEY CLUSTERED
(
        [PF_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
```

```
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
    CONSTRAINT [UC_POPULATION_FEMALE] UNIQUE
NONCLUSTERED
(
    [PF_COUNTRY_ID] ASC,
    [PF_YEAR] ASC,
    [PF_AGE_FLAG] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[POPULATION_MALE]    Script Date:
12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE [dbo].[POPULATION_MALE](
    [PM_ID] [int] IDENTITY(1,1) NOT
NULL,
    [PM_COUNTRY_ID] [int] NOT NULL,
    [PM_YEAR] [int] NOT NULL,
    [PM_MAX_AGE] [int] NULL,
    [PM_AGE_FLAG] [int] NULL,
    [PM_POPULATION_MALE] [decimal](18,
4) NULL,
    CONSTRAINT [PK_POPULATION_MALE] PRIMARY
KEY CLUSTERED
(
    [PM_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[POPULATION_MIDYAER_ON_PERIOD]
Script Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE
[dbo].[POPULATION_MIDYAER_ON_PERIOD](
    [PMOP_ID] [int] IDENTITY(1,1) NOT
NULL,
    [PMOP_COUNTRY_ID] [int] NOT NULL,
    [PMOP_YEAR] [int] NOT NULL,
    [PMOP_TOTAL_FLAG] [varchar](1)
NULL,
    [PMOP_STARTING_AGE] [int] NULL,
    [PMOP_AGE_GROUP_INDICATOR]
[varchar](1) NULL,
    [PMOP_ENDING_AGE] [int] NULL,
    [PMOP_MIDYEAR_POPULATION]
[decimal](18, 4) NULL,
    CONSTRAINT
[PK_POPULATION_MIDYAER_ON_PERIOD] PRIMARY
KEY CLUSTERED
(
    [PMOP_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
```

```
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
    CONSTRAINT
[UC_POPULATION_MIDYAER_ON_PERIOD] UNIQUE
NONCLUSTERED
(
    [PMOP_COUNTRY_ID] ASC,
    [PMOP_YEAR] ASC,
    [PMOP_STARTING_AGE] ASC,
    [PMOP_ENDING_AGE] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_FEMALE]
Script Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
CREATE TABLE
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_FEMALE]
(
    [PMOPF_ID] [int] IDENTITY(1,1) NOT
NULL,
    [PMOPF_COUNTRY_ID] [int] NOT NULL,
    [PMOPF_YEAR] [int] NOT NULL,
    [PMOPF_TOTAL_FLAG] [varchar](1)
NULL,
    [PMOPF_STARTING_AGE] [int] NULL,
    [PMOPF_AGE_GROUP_INDICATOR]
[varchar](1) NULL,
    [PMOPF_ENDING_AGE] [int] NULL,
    [PMOPF_MIDYEAR_POPULATION_FEMALE]
[decimal](18, 4) NULL,
    CONSTRAINT
[PK_POPULATION_MIDYEAR_ON_PERIOD_FEMALE]
PRIMARY KEY CLUSTERED
(
    [PMOPF_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
    CONSTRAINT
[UC_POPULATION_MIDYEAR_ON_PERIOD_FEMALE]
UNIQUE NONCLUSTERED
(
    [PMOPF_COUNTRY_ID] ASC,
    [PMOPF_YEAR] ASC,
    [PMOPF_STARTING_AGE] ASC,
    [PMOPF_ENDING_AGE] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
/***** Object: Table
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_MALE]
Script Date: 12/3/2023 5:51:05 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
```



```

CREATE TABLE
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_MALE](
    [PMOPM_ID] [int] IDENTITY(1,1) NOT
NULL,
    [PMOPM_COUNTRY_ID] [int] NOT NULL,
    [PMOPM_YEAR] [int] NOT NULL,
    [PMOPM_TOTAL_FLAG] [varchar](1)
NULL,
    [PMOPM_STARTING_AGE] [int] NULL,
    [PMOPM_AGE_GROUP_INDICATOR]
[varchar](1) NULL,
    [PMOPM_ENDING_AGE] [int] NULL,
    [PMOPM_MIDYEAR_POPULATION_MALE]
[decimal](18, 4) NULL,
    CONSTRAINT
[PK_POPULATION_MIDYEAR_ON_PERIOD_MALE]
PRIMARY KEY CLUSTERED
(
    [PMOPM_ID] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY],
    CONSTRAINT
[UC_POPULATION_MIDYEAR_ON_PERIOD_MALE]
UNIQUE NONCLUSTERED
(
    [PMOPM_COUNTRY_ID] ASC,
    [PMOPM_YEAR] ASC,
    [PMOPM_STARTING_AGE] ASC,
    [PMOPM_ENDING_AGE] ASC
)WITH (PAD_INDEX = OFF,
STATISTICS_NORECOMPUTE = OFF,
IGNORE_DUP_KEY = OFF, ALLOW_ROW_LOCKS = ON,
ALLOW_PAGE_LOCKS = ON,
OPTIMIZE_FOR_SEQUENTIAL_KEY = OFF) ON
[PRIMARY]
) ON [PRIMARY]
GO
ALTER TABLE
[dbo].[FERTILITY_RATES_ON_PERIOD] WITH
CHECK ADD CONSTRAINT
[FK_FERTILITY_RATES_ON_PERIOD] FOREIGN
KEY([FROP_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE
[dbo].[FERTILITY_RATES_ON_PERIOD] CHECK
CONSTRAINT [FK_FERTILITY_RATES_ON_PERIOD]
GO
ALTER TABLE [dbo].[INDEXES_BY_COUNTRY]
WITH CHECK ADD CONSTRAINT
[FK_INDEXES_BY_COUNTRY] FOREIGN
KEY([IBC_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE [dbo].[INDEXES_BY_COUNTRY]
CHECK CONSTRAINT [FK_INDEXES_BY_COUNTRY]
GO
ALTER TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY] WITH
CHECK ADD CONSTRAINT
[FK_MORTALITY_LIFE_EXPECTANCY] FOREIGN
KEY([MLEF_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY] CHECK
CONSTRAINT [FK_MORTALITY_LIFE_EXPECTANCY]
GO
ALTER TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY_FEMALE]

```

```

WITH CHECK ADD CONSTRAINT
[FK_MORTALITY_LIFE_EXPECTANCY_FEMALE]
FOREIGN KEY([MLEF_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY_FEMALE]
CHECK CONSTRAINT
[FK_MORTALITY_LIFE_EXPECTANCY_FEMALE]
GO
ALTER TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY_MALE]
WITH CHECK ADD CONSTRAINT
[FK_MORTALITY_LIFE_EXPECTANCY_MALE] FOREIGN
KEY([MLEM_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE
[dbo].[MORTALITY_LIFE_EXPECTANCY_MALE]
CHECK CONSTRAINT
[FK_MORTALITY_LIFE_EXPECTANCY_MALE]
GO
ALTER TABLE
[dbo].[POPULATION_DEMOGRAPHIC_RATES_GENERAL
] WITH CHECK ADD CONSTRAINT
[FK_POPULATION_DEMOGRAPHIC_RATES_GENERAL]
FOREIGN KEY([PDRG_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE
[dbo].[POPULATION_DEMOGRAPHIC_RATES_GENERAL
] CHECK CONSTRAINT
[FK_POPULATION_DEMOGRAPHIC_RATES_GENERAL]
GO
ALTER TABLE [dbo].[POPULATION_FEMALE] WITH
CHECK ADD CONSTRAINT
[FK_POPULATION_FEMALE] FOREIGN
KEY([PF_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE [dbo].[POPULATION_FEMALE] CHECK
CONSTRAINT [FK_POPULATION_FEMALE]
GO
ALTER TABLE [dbo].[POPULATION_MALE] WITH
CHECK ADD CONSTRAINT [FK_POPULATION_MALE]
FOREIGN KEY([PM_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE [dbo].[POPULATION_MALE] CHECK
CONSTRAINT [FK_POPULATION_MALE]
GO
ALTER TABLE
[dbo].[POPULATION_MIDYAER_ON_PERIOD] WITH
CHECK ADD CONSTRAINT
[FK_POPULATION_MIDYAER_ON_PERIOD] FOREIGN
KEY([PMOP_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE
[dbo].[POPULATION_MIDYAER_ON_PERIOD] CHECK
CONSTRAINT
[FK_POPULATION_MIDYAER_ON_PERIOD]
GO
ALTER TABLE
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_FEMALE]
WITH CHECK ADD CONSTRAINT
[FK_POPULATION_MIDYEAR_ON_PERIOD_FEMALE]
FOREIGN KEY([PMOPF_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
ALTER TABLE
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_FEMALE]
CHECK CONSTRAINT
[FK_POPULATION_MIDYEAR_ON_PERIOD_FEMALE]

```



```
GO
ALTER TABLE
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_MALE]
WITH CHECK ADD CONSTRAINT
[FK_POPULATION_MIDYEAR_ON_PERIOD_MALE]
FOREIGN KEY([PMOPM_COUNTRY_ID])
REFERENCES [dbo].[COUNTRIES] ([COUNTRY_ID])
GO
```

```
ALTER TABLE
[dbo].[POPULATION_MIDYEAR_ON_PERIOD_MALE]
CHECK CONSTRAINT
[FK_POPULATION_MIDYEAR_ON_PERIOD_MALE]
GO
```

Section-break (continuous)

Σχόλιο για τα παραπάνω script: από την στιγμή που υπάρχουν τα ddl στο github νομίζω αυτό είναι περιττό, τα έβαλα επειδή υπήρχαν στο αρχικό παράδειγμα.

Αρχικό σχόλιο

«Υπαρχει ένας πίνακας με δεδομένα από regions του μεγάλου excel που ακόμα δεν ξέρω πως θα τον χρησιμοποιήσω στα ερωτήματα. Έλεγα αν τα join με τις ιδιότητες αυτών έχουν τους ίδιους αριθμούς αλλά δεν είναι ίσα

Πχ. Αν το GNI όλων των developed countries μια χρονιά ισούται με την τιμή της γραμμής σε αυτόν τον πίνακα αλλά δεν ισχύει. Επομένως είναι data τα οποία δεν γνωρίζω αν είναι χρήσιμα.

«

Τον regions τον θεώρησα σκορπίδια και τον άφησα στην άκρη γιατί την πληροφορία την έχω στον countries και από εκεί μπορώ να πάρω τις χώρες που ανήκουν σε κάποια region

Στην συνέχεια δημιουργήθηκαν και άλλοι δύο πίνακες POPULATION_FEMALE_PERIOD, POPULATION_MALE_PERIOD που είναι στο ίδιο μοτίβο με αυτούς που έχουν τα flags.

1.1.1 ΣΧΕΔΙΑΣΗ ΣΧΗΜΑΤΟΣ

Σε οποιονδήποτε πίνακα, εκτός του excel income by country, που υπήρχε η κατηγοριοποίηση male, female δημιουργήθηκαν ξεχωριστοί με την ανάλογη κατάληξη και ένας που είχε τα γενικά χαρακτηριστικά. Όπως φαίνεται παρακάτω οι τριάδες. Με αυτό τον τρόπο γνωρίζουμε από πριν τι θα χρησιμοποιήσουμε αν θέλουμε στατιστικά ανα φύλλο. Από 3 πίνακες πήγαμε στους 9 αλλά δεν χρειάζεται να ψάχνουμε σε στήλες κάθε φορά που θέλουμε κάτι ειδικό.

```
MORTALITY_LIFE_EXPECTANCY
MORTALITY_LIFE_EXPECTANCY_FEMALE
MORTALITY_LIFE_EXPECTANCY_MALE

POPULATION_MIDYAER_ON_PERIOD
POPULATION_MIDYEAR_ON_PERIOD_FEMALE
POPULATION_MIDYEAR_ON_PERIOD_MALE

POPULATION_FEMALE
POPULATION_MALE
POPULATION_DEMOGRAPHIC_RATES_GENERAL
```

Στον POPULATION_DEMOGRAPHIC_RATES_GENERAL έχουμε στήλες από ότι έχει να κάνει με γεννήσεις, θανάτους, προσδόκιμο ζωής..

Στο excel εκτός από χώρες υπήρχαν και περιοχές, ήπειροι γενικά δεδομένα όπου στην πρώτη φάση απλά έλεγα αν είναι ίσα με τα αντίστοιχα αναλυτικά όπου δεν είναι και έτσι χωρίς να ξέρω πως θα τα χρησιμοποιήσω τα έριξα σε έναν πίνακα INDEXES_BY_REGION χωρίς κλειδιά.

Στον πίνακα INDEXES_BY_COUNTRY έβαλα όλη την άλλη πληροφορία του excel ανα χώρα και έτος άρα τα εννιά φύλλα έγιναν εννιά στήλες.

Η γενική γραμμή ήταν ότι όπου υπήρχαν στήλες 1999,2000,2001 έγιναν με την χρήση UNPIVOT σειρές και μπήκαν σε μια στήλη με το λεκτικό year. Το κυριότερο θετικό αυτού που αναγνωρίζω είναι ότι σε περίπτωση διαγραφής και ισαγωγής έτους δεν χρειάζεται να κάνουμε ALTER TABLE όπου αυτό σε παραγωγικό περιβάλλον ισοδυναμεί με διαδικασία update.

1.2 ΣΧΕΣΙΑΚΟ ΣΧΗΜΑ ΣΕ ΦΥΣΙΚΟ ΕΠΙΠΕΔΟ

1.2.1 ΡΥΘΜΙΣΗ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΤΟΥ DBMS

Server Memory Options

- Minimum server memory in MB : 0
- Maximum server memory in MB : 10240

Other Memory Options : Default

I am using all the processors of my PC

1.2.2 ΡΥΘΜΙΣΗ ΤΟΥ ΦΥΣΙΚΟΥ ΣΧΗΜΑΤΟΣ ΤΗΣ ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ

Δημιουργήθηκαν INDEXES σε όλα τα PRIMARY KEYS και σε όλα τα FOREIGN KEYS που στην εργασία μας είναι τα country_id που υπάρχουν σε κάθε πίνακα. Αυτό γιατί για κάθε ερώτημα που θα κάνουμε για το webapp θα υπάρχει τουλάχιστον ένα join με το country_id επομένως θα είναι αισθητή η διαφορά.

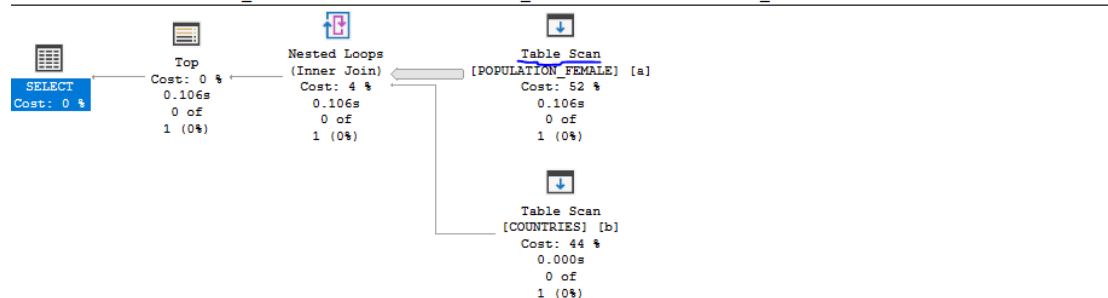
Ένα παράδειγμα:

```
SET STATISTICS IO ON;
SET STATISTICS TIME ON;
select
    top 1 *
from DWH_INTERMEDIATE.dbo.POPULATION_FEMALE a
inner join DWH_INTERMEDIATE.dbo.COUNTRIES b
    on a.COUNTRY_ID = b.COUNTRY_ID
where a.COUNTRY_ID= 1200
select
    top 1 *
from DWH.dbo.POPULATION_FEMALE a
inner join DWH.dbo.COUNTRIES b
    on a.PF_COUNTRY_ID = b.COUNTRY_ID
```

where a.PF_COUNTRY_ID= 1500

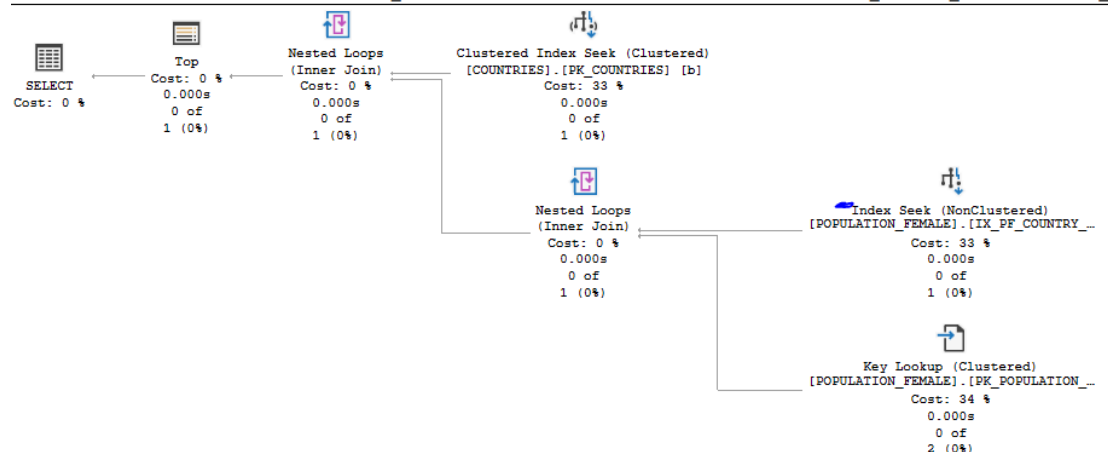
Query 1: Query cost (relative to the batch): 62%

select top 1 * from DWH_INTERMEDIATE.dbo.POPULATION_FEMALE a inner join DWH_INTERMEDIATE.dbo.COUNTRIES b on a.C



Query 2: Query cost (relative to the batch): 38%

select top 1 * from DWH.dbo.POPULATION_FEMALE a inner join DWH.dbo.COUNTRIES b on a.PF_COUNTRY_ID = b.COUNTRY_I



Από το STATISTICS έχουμε

Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads 0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead reads 0.

Table 'POPULATION_FEMALE'. Scan count 1, logical reads 6425, physical reads 0, page server reads 0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead reads 0.

SQL Server Execution Times:

Table 'Worktable'. Scan count 0, logical reads 0, physical reads 0, page server reads 0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead reads 0.

Table 'COUNTRIES'. Scan count 0, logical reads 2, physical reads 0, page server reads 0, read-ahead reads 0, page server read-ahead reads 0, lob logical reads 0, lob physical reads 0, lob page server reads 0, lob read-ahead reads 0, lob page server read-ahead reads 0.

Σε αυτό το παράδειγμα ζητάμε να φέρουμε το top ένα αποστέλεσμα ενός country_id που δεν υπάρχει δηλαδή περιμένουμε τον κενό πίνακα.

Στην βάση DWH_INTERMEDIATE που έχει οτς ίδιους πίνακες με την DWH αλλά χωρίς indexes βλέπουμε από το execution plan ότι κάνει full table scan και στοτς δύο πίνακες, όμως στην άλλη περίπτωση χρησιμοποιεί τους indexes. Από τα στατιστικά βλέπουμε

ότι στην δεύτερη περίπτωση ότι κάνει μόνο 2 logical reads και κανένα scan count στον πίνακα COUNTRIES.

Επίσης όλα τα δεδομένα είναι ανά έτος πλέον ότι δεν είχε τιμή πήρε του επόμενου έτους που είχε. Ότι είχε αριθμό εκτός του έτους έγινε decimal(18,4) γιατί το float έχει αρκετά προβλήματα το κυριότερο είναι του σε float $1/3+2/3$ δεν κάνει 1.

Στους πίνακες με στήλες country_id, year δημιουργήθηκαν UNIQUE KEYS έτσι ώστε να υποχρεώσουμε τα δεδομένα να είναι μοναδικά ανα χώρα και χρονιά.

Στους πίνακες με στήλες country_id, year, starting_age, ending_age, age_flag επίσης για την διασφάλιση της μοναδικότητας φτιάξαμε UNIQUE KEYS.

1.2.3 ΡΥΘΜΙΣΗ ΑΣΦΑΛΕΙΑΣ

Δημιουργήθηκε ένας χρήστης στην στον server ο project_user που έχει δικαιώματα datareader, datawriter και ddladmin στις βάσεις του server και sysadmin στον server είναι υποτίθεται ο χρήστης που κάνει την συντήρηση στην βάση.

Επίσης δημιουργήθηκε και ο WebAppUser που έχει δικαιώματα μόνο select , execute στα σχήματα dbo, rpt και στην βάση έχει δικαίωμα View Definition για να βλέπει το σχήμα INFORMATION_SCHEMA το οποίο χρησιμοποίησα για την παραγωγή των δεδομένων των γραφικών παραστάσεων. Με τα credentials αυου του χρήστη θα συνδέεται το application με την βάση.

Η λογική είναι ότι sysadmin είναι μόνο αυτός που χειρίζεται την βάση έστω IT και ο χρήστης του user έχει μόνο select , execute στα σχήματα dbo . rpt(report).

2 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΛΟΓΙΣΜΙΚΟΥ

2.1 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΔΟΜΗ ETL

Δημιουργήθηκαν τρεις βάσεις

- DWH_STAGING

Η βάση δημιουργήθηκε σαν data lake έτσι έχουμε τόσους πίνακες όσα τα csv αρχεία και άλλους εννιά όσα τα φύλλα του excel.

Η διαδικασία γεμίσματος γίνεται από την εκτέλεση της procedure :exec etl.InsertDataMain η οποία δεν έχει ορίσματα. Παίρνει το path και το όνομα των αρχείων από τον πίνακα dbo.DataFiles που αποθηκεύονται. Οι πίνακες έχουν όλοι στήλες του τύπου VARCHAR ανάλογης διάστασης και δεν έχει γίνει καμία ενέργεια χωρισμού των αρχικών δεδομένων.

Σε περίπτωση που τα path είναι διαφορετικά από αυτά που είναι φορτωμένα στην βάση θα πρέπει να γίνει update στον πίνακα dbo.DataFiles

```
use DWH_STAGING
declare @path varchar(max) = ''
begin transaction
update a set a.Path = @path from DataFiles a
--commit
Rollback
```

Μετά από αυτό η διαδικασία είναι απλή exec DWH_STAGING.etl.InsertDataMain

- DWH_INTERMEDIATE

Είναι η βάση που είναι ίδια με την παραγωγική αλλά δεν έχει κλειδιά και ευρετήρια.

Η διαδικασία ETL αυτής γίνεται από την εκτέλεση της procedure etl.ImportAMain το «Α» μπήκε για να είναι μόνο και μόνο στην αρχή της λίστας των ομόνυμων procedures.

Η διαδικασία είναι :

- Εισαγωγή του πίνακα COUNTRIES με την δημιουργία country_id
- Εισαγωγή fertility rates όπου γίνεται το inner join με τον countries και αντικαθιστούνται τα county_code και country_name από το country_id και η αλλαγή των στηλών fertility_rate_num1_num2 στην μορφή που έχει και το midyear_population_5yr_age.csv. Έτσι έχουμε τον πίνακα FERTILITY_RATES_PER_PERIOD.
- Εισαγωγή των INDEXES BY COUNTRY OR REGION(excel). Για να γίνει αυτό δημιουργήσαμε έναν πίνακα COUNTRY_REGION_OFFICIAL_MAPPING(άρα inner join με αυτόν) όπου έχουμε μέσα σχέσεις με το country_id των Official_Names και Display_Names και μερικά άλλα που βρήκα σκάβοντας. Από εδώ έχουμε από τους εννιά της προηγούμενης βάσης σε 2 τους INDEXES_BY_COUNTRY και INDEXES_BY_NAME. Με στήλες τις ονομασίες των φύλλων του excel.

- Εισαγωγή των δημογραφικών στοιχείων (όλα είναι τέτοια απλά είναι μια γενική διαδικασία και την ονόμασα έτσι) , εδώ έχουμε πάλι inner join με τον countries. Δημιουργούνται τρεις πίνακες οι MORTALITY_LIFE_EXPECTANCY_MALE, MORTALITY_LIFE_EXPECTANCY_FEMALE, POPULAION_DEMOGRAPHIC_RATES_GENERAL.
- Εισαγωγή των Population male female είναι δύο procedures που διαχωρίζουν τα στοιχεία του πίνακα midyear_population_age_sex σε δύο υποπίνακες. Πάλι inner join με τον COUNTRIES.
- Τέλος εκυελείται η ImportPopulationMidyearOnPeriod όπου δημιουργεί τρεις πίνακες MALE, FEMALE και γενικός.

Για να εκτελεστούν τα παραπάνω αρκεί να τρέξουμε το exec DWH_INTERMEDIATE.etl.ImportAMain 0;

- DWH
Είναι η τελική βάση.
Η διαδικασία ETL αυτής γίνεται από την εκτέλεση της procedure etl.ImportAMain το «Α» μπήκε για να είναι μόνο και μόνο στην αρχή της λίστας των ομόνυμων procedures.
Σε περίπτωση που δωθεί κενή η βάση έχει και μια μεταβλητή όπου αν την θέσουμε 1 θα δημιουργήσει τους πίνακες και τα indexes, primary keys, unique keys.

Για να εκτελεστούν τα παραπάνω αρκεί να τρέξουμε το exec DWH.etl.ImportAMain 0;

Συνοπτικά η διαδοχή των εντολών

```
use DWH_STAGING
declare @path varchar(max) = 'path don't forget the -> \'
begin transaction
update a set a.Path = @path from DataFiles a
--commit
Rollback

exec DWH_STAGING.etl.InsertDataMain;

exec DWH_INTERMEDIATE.etl.ImportAMain 0;

exec DWH.etl.ImportAMain 0;
```

2.1.1 ERROR HANDLING

Σε κάθε βάση έχουμε και τους πίνακες errorLogs που γράφεται η πληροφορία αν κάποιο insert σκάσει

Στην DWH επίσης σε αυτόν τον πίνακα γράφεται και ότι exemption έχουμε και στο πρόγραμμα. Θα ανεβάσω τις βάσεις άδειες εκτός από αυτόν τον errorLogs.

2.1.2 ΒΑΣΙΚΗ PROCEDURE REPORTING

Πιστεύω πρέπει να αφιερώσω μια ενότητα σε αυτή την procedure γιατί ότι δεδομένα ζητάμε αντλούνται από εκεί.

Δεν συνιστώ να φτιάχνουμε τέτοια εκτρώματα σε παραγωγικό περιβάλλον γιατί τα παντρευόμαστε δηλ ότι και αν τύχει μόνο εμείς που την κάναμε ξέρουμε να την διαχειριστούμε.

Έχει input ένα string της μορφής

1. /TypeOfChart/line – τύπος του chart
2. /CountriesList/1,3 – λίστα με τις χώρες (είναι none μόνο όταν έχω επιλέξει ChoosingGroups)
3. /columnIndexes/PMOP_MIDYEAR_POPULATION,PMOPF_MIDYEAR_POPULATION_FEMALE – λίστα με indices
4. /AggregateFunction/none –αν θέλω δεδομένα ανά τριετία, πενταετία ... *ετία τότε μπορώ να επιλέξω αν θέλω να βλέπω average, sum, max, min
5. /ChooseYear/40-44 – κάποιον πίνακες έχουν μέσα και εύρος ηλικίας συμπληρώνεται μόνο αν έχω επιλέξει κάποια στήλη με αυτή την πληροφορία
6. /ChoosingGroups/none – αν θέλω να προβάλω τις χώρες της Ασίας ή αυτές που είναι developed ή να είναι κάποια επιλογή από τον πίνακα countries
7. /YearAggregation/None–αριθμός χρόνων για να προβάλω grouparισμένα δεδομένα είναι κείμενο βάζω ότι θέλω αρκεί να υπάρχει αυτό το εύρος.

Γεννάει unions της μορφής

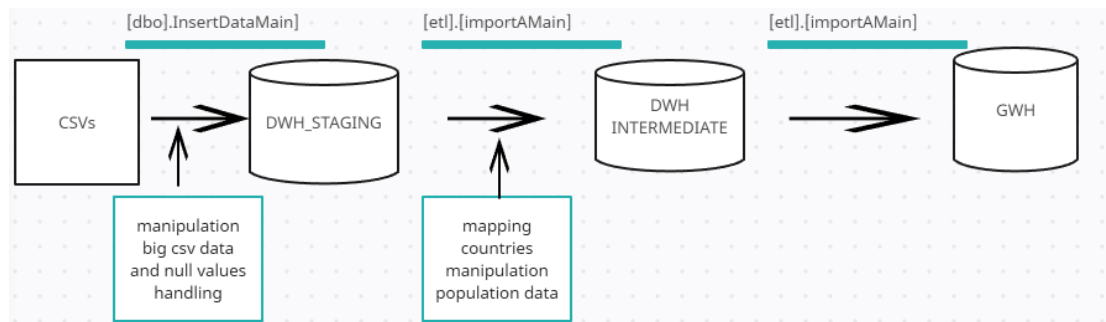
```
select
'line'as [text] ,
c.C_DISPLAY_NAME as [name],
PMOP_YEAR as x,
ISNULL(PMOP_MIDYEAR_POPULATION,0) AS y
from DWH.dbo.POPULATION_MIDYAEAR_ON_PERIOD
inner join countries c on PMOP_COUNTRY_ID = c.country_id where PMOP_COUNTRY_ID
IN (1,3)
and PMOP_STARTING_AGE = 40 and PMOP_ENDING_AGE = 44
union
select
'line'as [text] ,
c.C_DISPLAY_NAME as [name],
PMOPF_YEAR as x,
ISNULL(PMOPF_MIDYEAR_POPULATION_FEMALE,0) AS y
from DWH.dbo.POPULATION_MIDYEAR_ON_PERIOD_FEMALE
inner join countries c on PMOPF_COUNTRY_ID = c.country_id where
PMOPF_COUNTRY_ID IN (1,3)
and PMOPF_STARTING_AGE = 40 and PMOPF_ENDING_AGE = 44
order by [name] asc,x asc
```

αυτό δημιουργήθηκε από το 1..7 και δόθηκε στην view για προβολή.

και το όνομα της είναι [rpt].[ProduceResults]

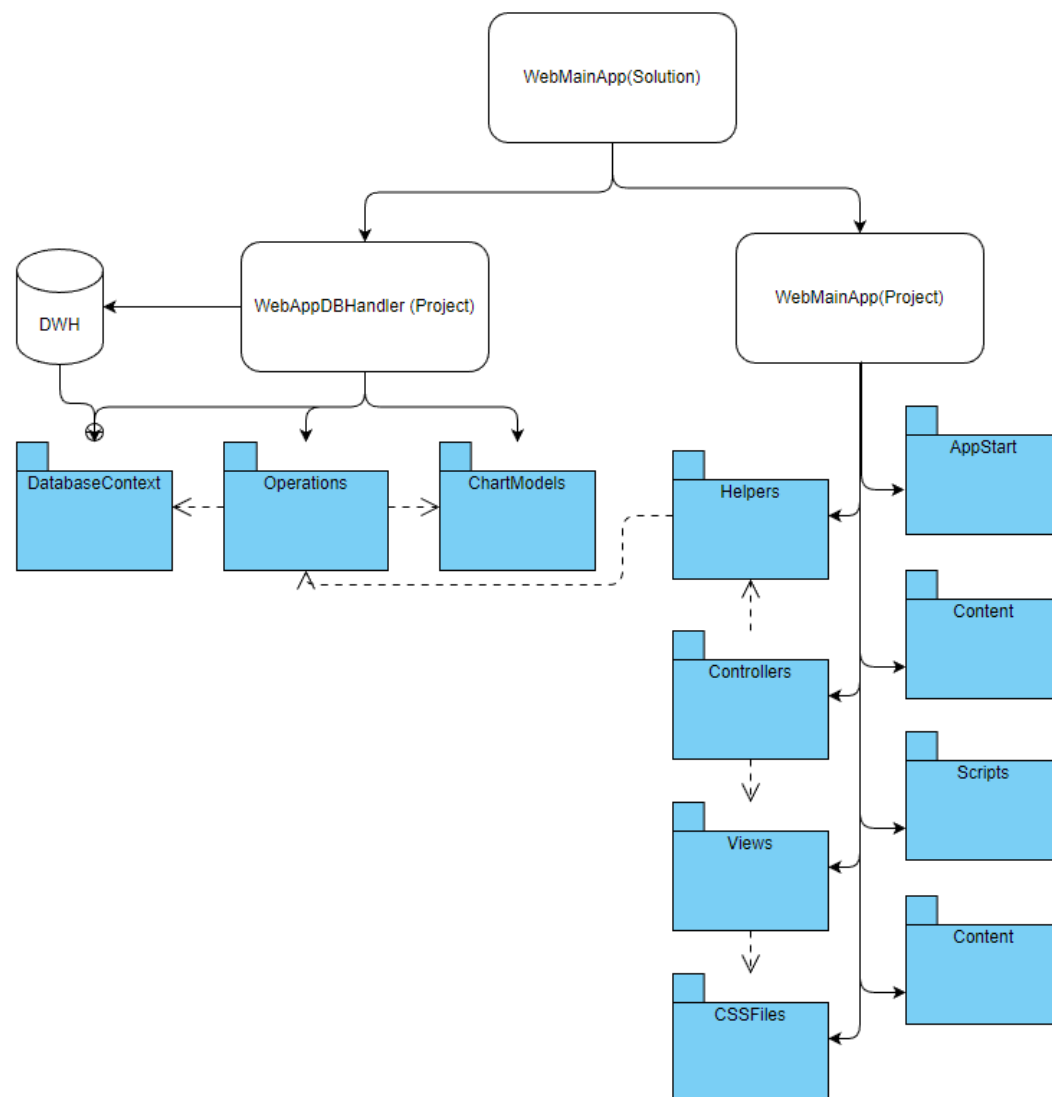
2.2 ΔΙΑΓΡΑΜΜΑΤΑ ΠΑΚΕΤΩΝ / ΥΠΟΣΥΣΤΗΜΑΤΩΝ ΚΕΝΤΡΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

Σχετικά με τις βάσεις ένα διάγραμμα είναι (δυσλεξία στην τελευταία βάση αντί για GWH = DWH)



2.3 ΔΙΑΓΡΑΜΜΑ(ΤΑ) ΚΛΑΣΕΩΝ ΚΕΝΤΡΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

Package diagram :



Μια μικρή ανάλυση

Το Solution WebMainApp έχει 2 projects

1. WebAppDBHandler

Το οποίο έχει τα εξής packages

- a. DatabaseContext : ευθύνεται για το mapping των πινάκων της βάσης στο πρόγραμμα τα λεγόμενα models
- b. ChartModels : είναι τα models/classes που δεν υπάρχουν στην βάση αλλά τα φτιάξαμε για να μας βοηθήσουν στην δημιουργία του json που χρειάζεται το canvasJS
- c. Operations : είναι το πακέτο που έχει τις κλάσεις οι οποίες τροφοδοτούν τα models με δεδομένα από την βάση είναι αυτό που ανοίγει connection με την βάση .

2. WebMainApp

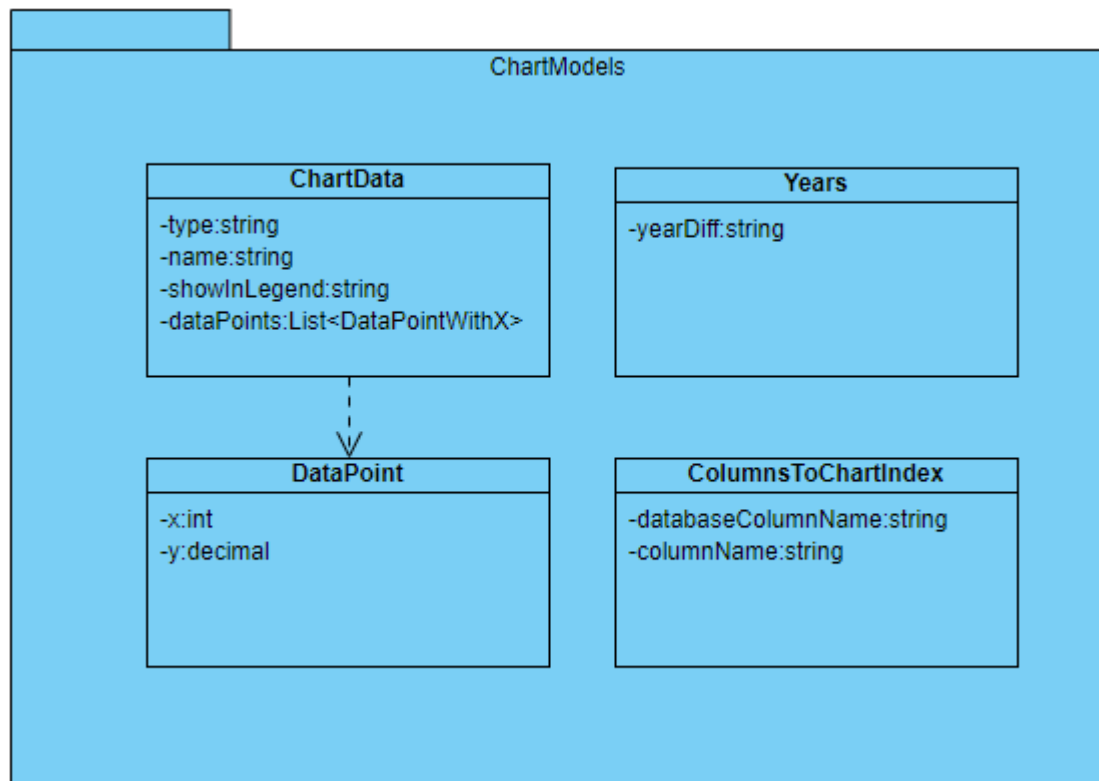
Το οποίο έχει τα εξής packages

- a. AppStart: classes για να ξεκινήσει το πρόγραμμα auto generated άλλαξα μόνο το rootConfig
- b. Content : autoGenerated bootstrap css stuff χρησιμεύουν στα views
- c. Scripts : autoGenerated bootstrap css stuff and JQuery χρησιμεύουν στα views
- d. Controllers : έχει τους δύο controllers που χρειαζόμαστε για το project οι οποίοι γεννάνε τα views
- e. Views: είναι οι «σελίδες» του browser.
- f. CSSFiles: είναι ένα αρχείο css που περιέχει τις κλάσεις που κάναμε για να έχουμε λίγο ομοιομορφία στα dropdowns ένα πιο ωραίο χρώμα από το λευκό στο background και καλύτερη στοίχιση.

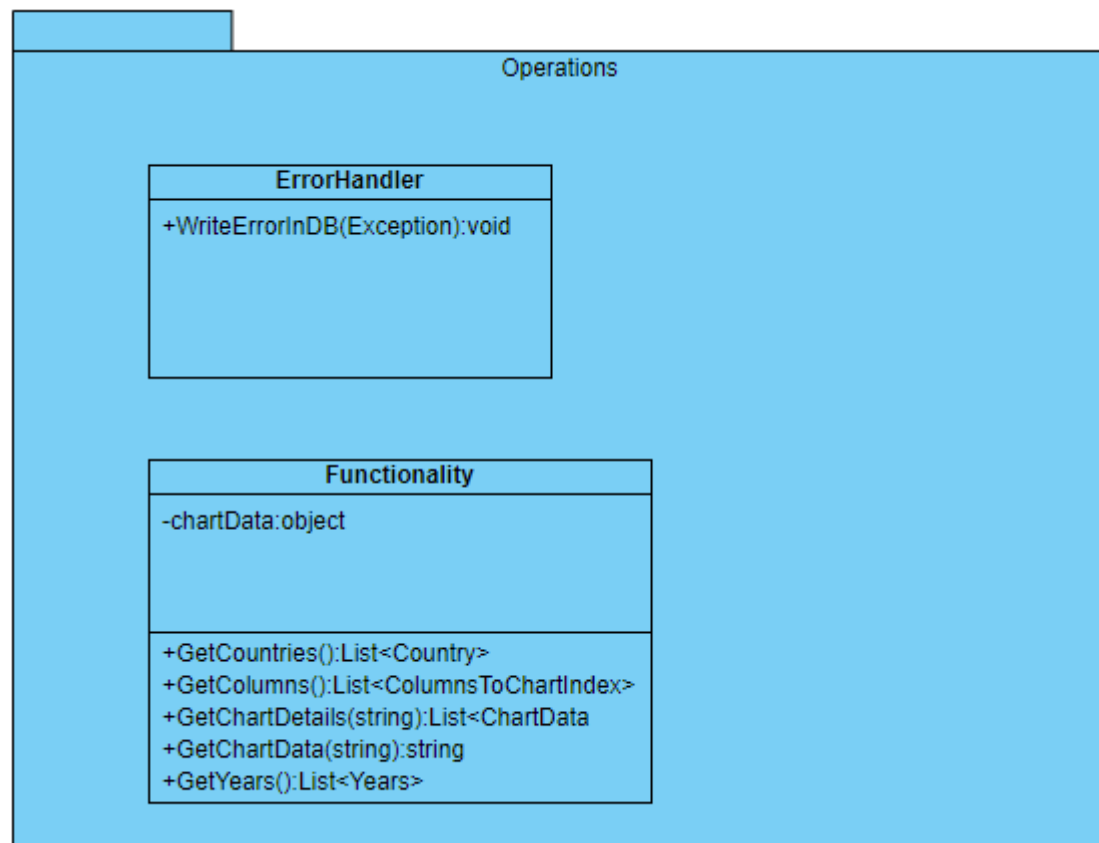
Πιο αναλυτικά

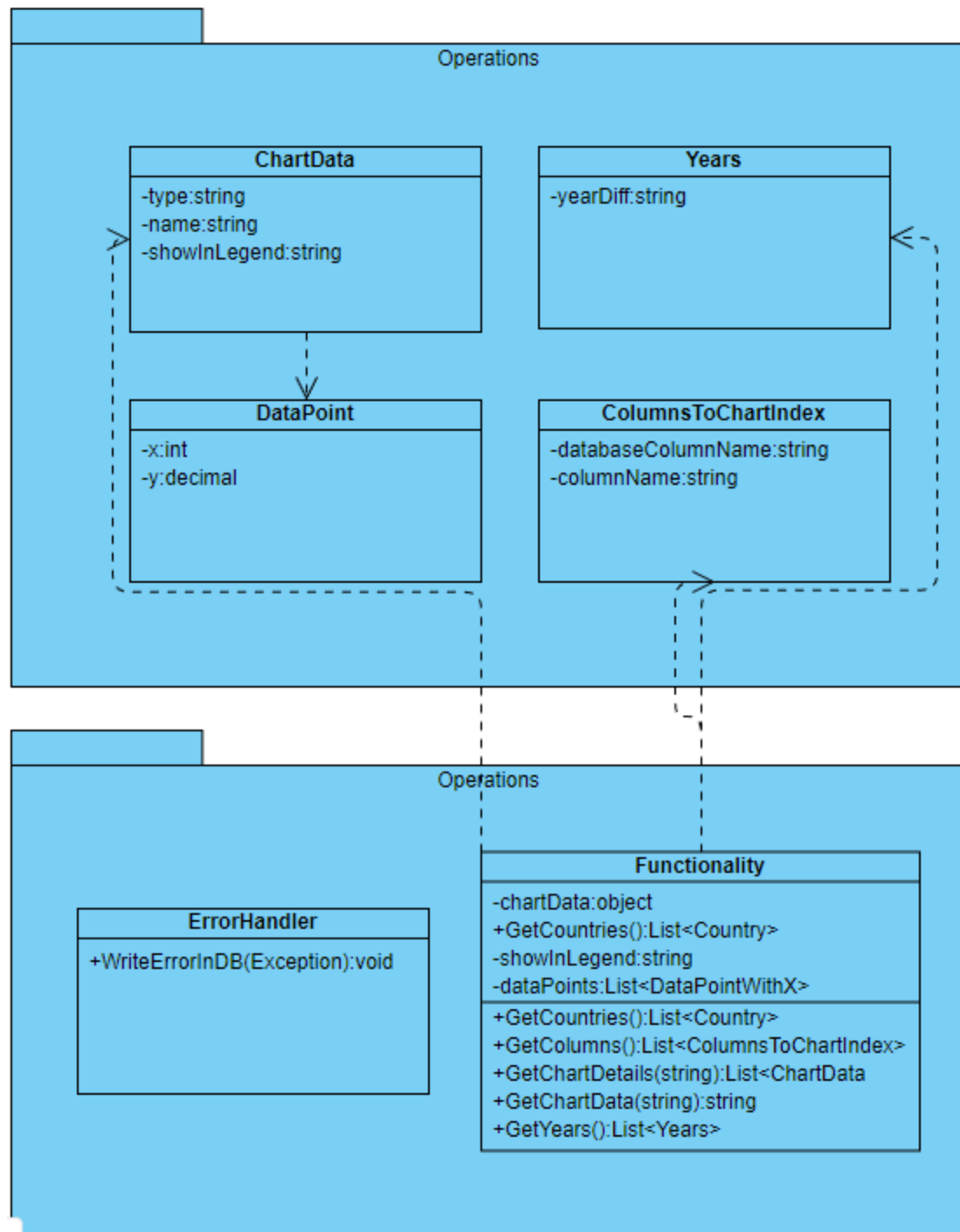
Το DatabaseContext δεν θα το αναλύσω γιατί περιέχει όπως είπαμε τους πίνακες της βάσης που δίνονται στην ενότητα 1

Το ChartModels

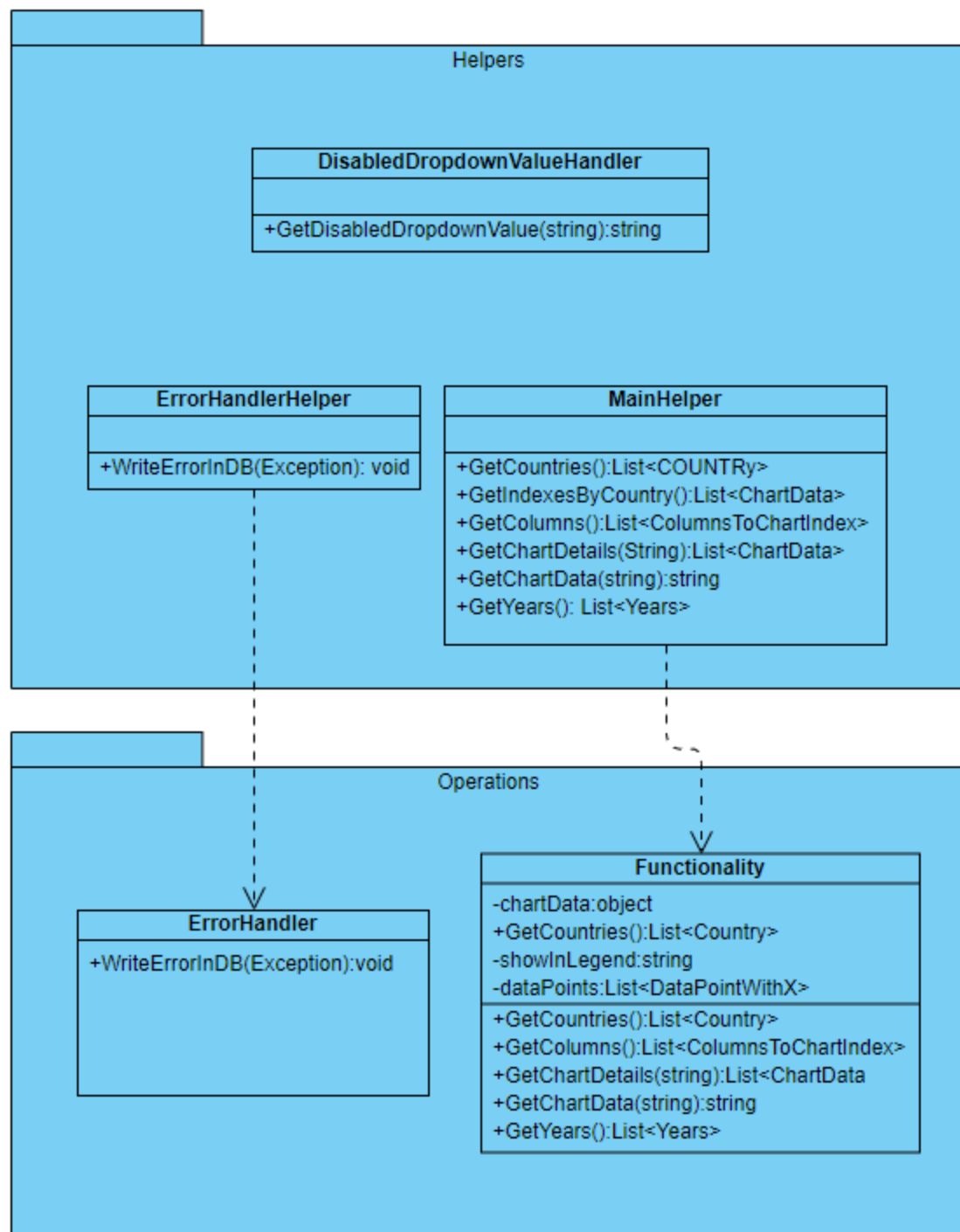


Functionality : όπως δείξαμε και στα packages είναι αυτή που θέλει στις μεθόδους της τις κλάσεις **Years**, **ColumnsToChartIndex**, **ChartData**. Επομένως θα κάνω και διάγραμμα και των δύο πακέτων

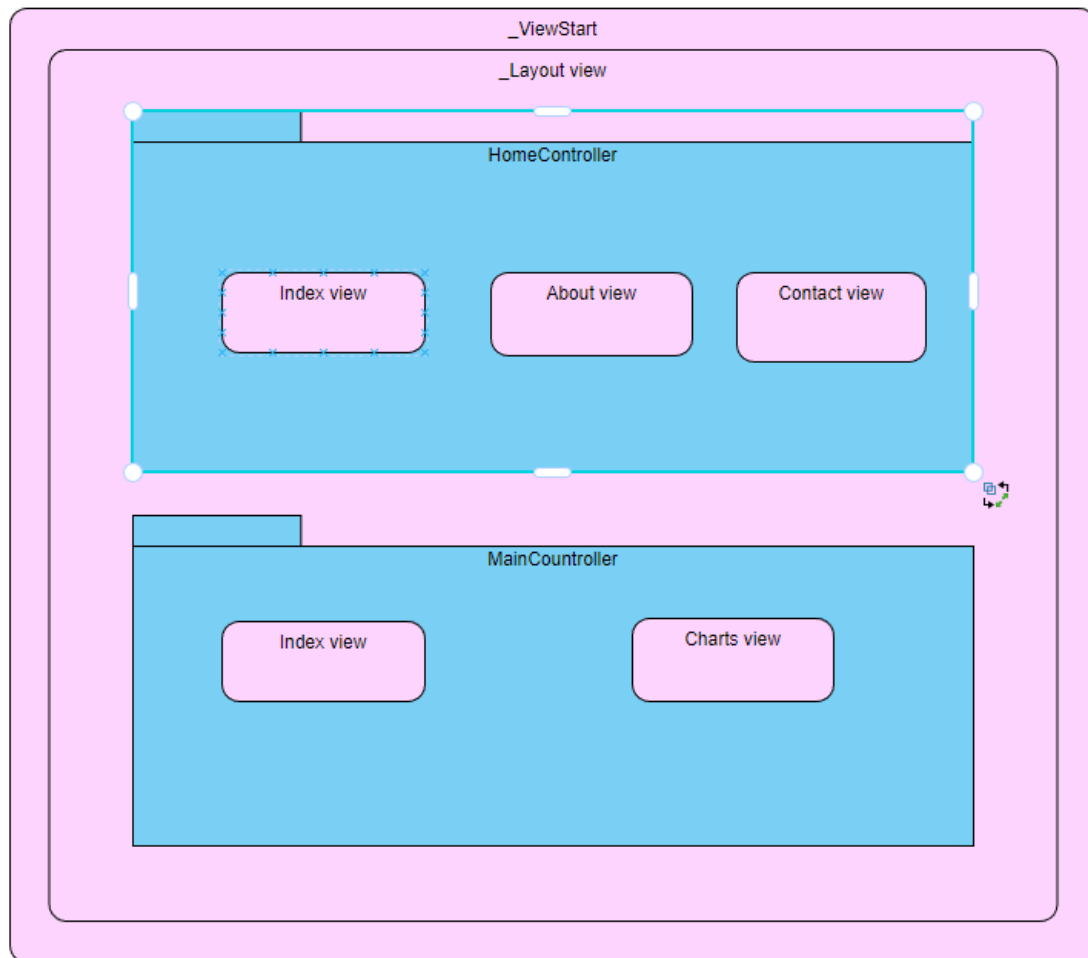




Helpers



Controllers views



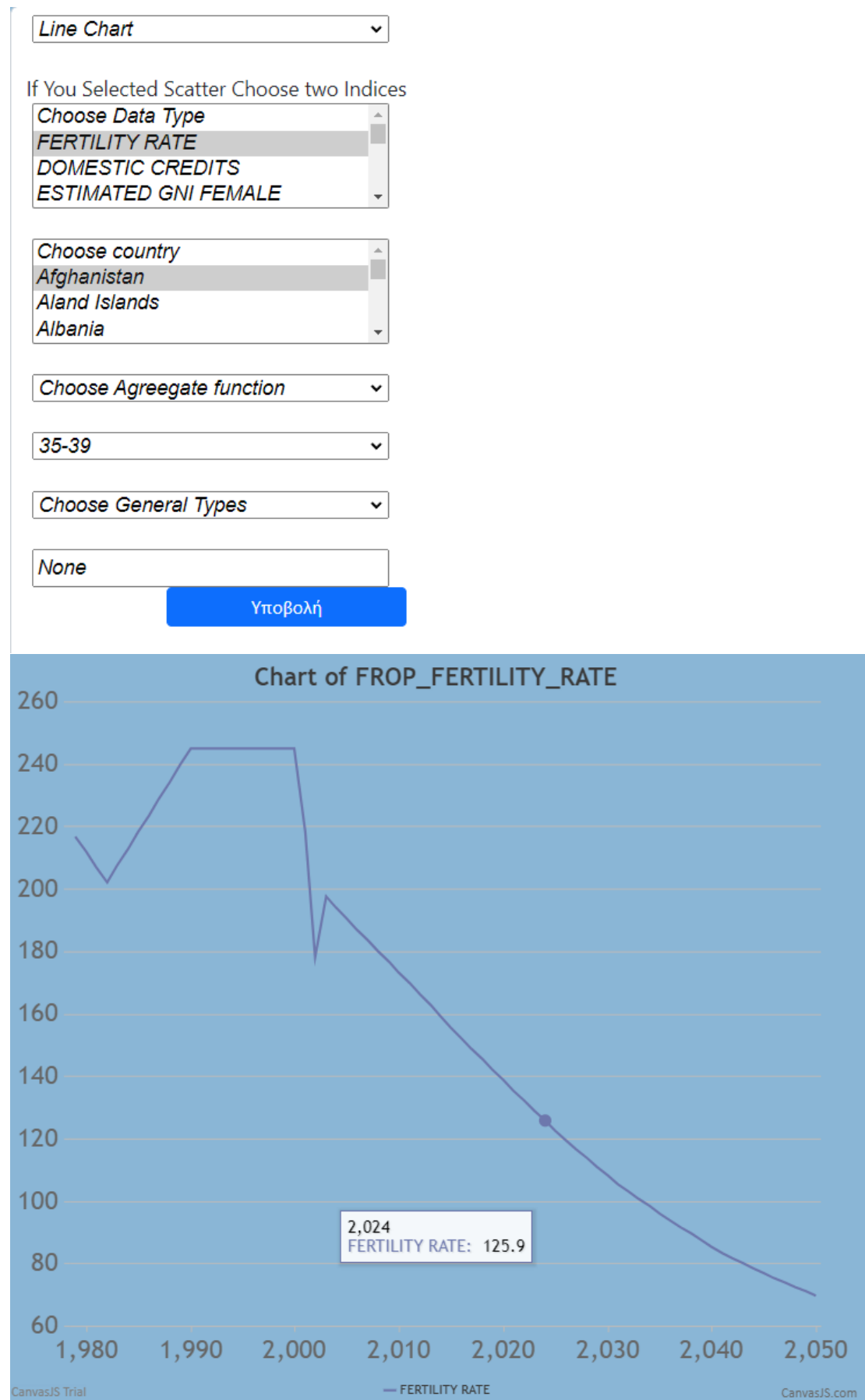
3 ΥΠΟΔΕΙΓΜΑΤΑ ΕΡΩΤΗΣΕΩΝ ΚΑΙ ΑΠΑΝΤΗΣΕΩΝ

3.1 LINE CHART

Μια Χώρα Ένα τύπο δεδομένων

Είσοδος στην procedure

/TypeOfChart/line/CountriesList/1/columnIndexes/FR0P_FERTILITY_RATE/Agreeeg
ateFunction/none/ChooseYear/35-39/ChoosingGroups/none/YearAggregation/None



Κ Χώρες κ τύποι δεδομένων

Είσοδος στην procedure

/TypeOfChart/line/CountriesList/1,3,47,60/columnIndexes/IBC_GDP_PER_CAPITA, IBC_GDP_TOTAL, IBC_GNI_PERCAPITA/AggregateFunction/none/ChooseYear/none/ChoosingGroups/none/YearAggregation/None

Line Chart

If You Selected Scatter Choose two Indices

Choose Data Type
FERTILITY RATE
DOMESTIC CREDITS
ESTIMATED GNI FEMALE

Choose country
Afghanistan
Aland Islands
Albania

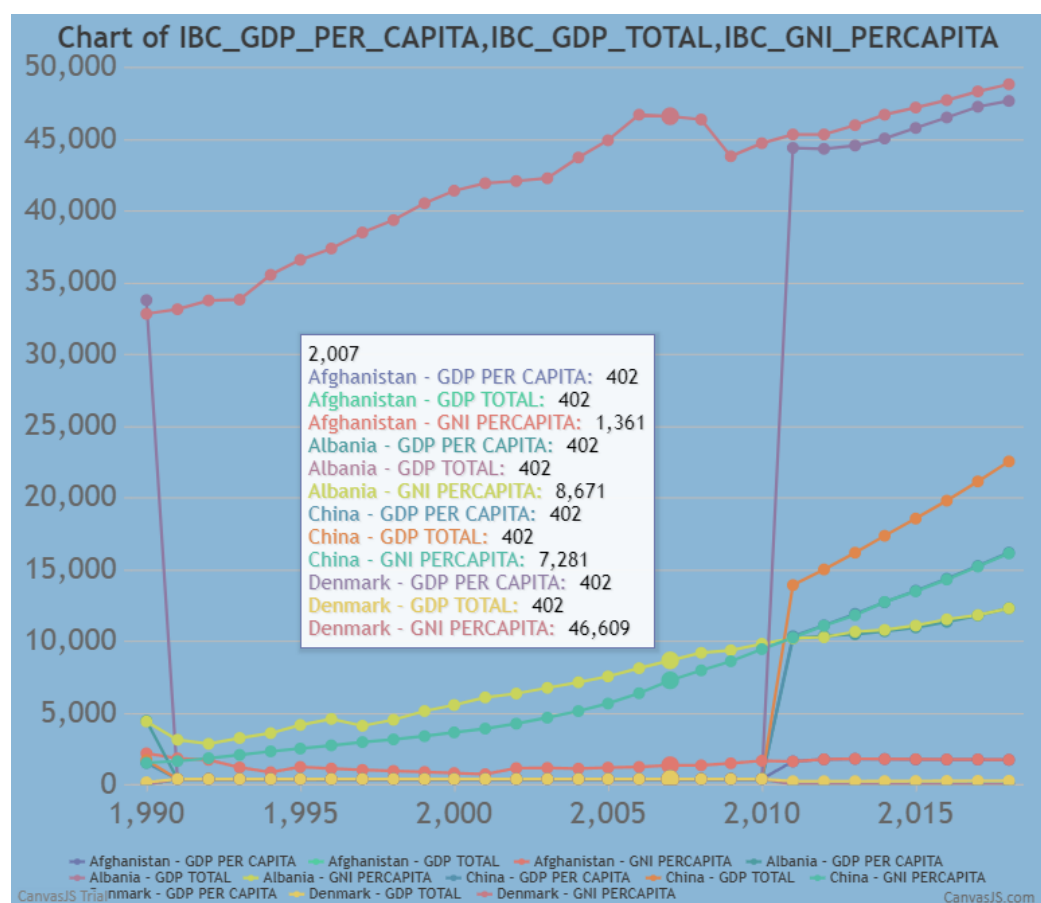
Choose Aggregate function

35-39

Choose General Types

None

Υποβολή



3.2 BAR CHART

Μια Χώρα Ένα τύπο δεδομένων

Είσοδος στην procedure

/TypeOfChart/bar/CountriesList/60/columnIndexes/FROP_FERTILITY_RATE/Agreeg
ateFunction/none/ChooseYear/35-39/ChoosingGroups/none/YearAggregation/None

Bar Chart

If You Selected Scatter Choose two Indices
 Choose Data Type
 FERTILITY RATE
 DOMESTIC CREDITS
 ESTIMATED GNI FEMALE

Cyprus
 Czechia
 Democratic Republic of the Congo
 Denmark

Choose Agreegate function

Choose Year

Choose General Types

None

Υποβολή



Κ Χώρες κ τύποι δεδομένων

Είσοδος στην procedure

/TypeOfChart/bar/CountriesList/47,60/columnIndexes/IBC_GDP_PER_CAPITA,IBC_GDP_TOTAL,IBC_GNI_PERCAPITA/AgreegateFunction/none/ChooseYear/none/ChooseingGroups/none/YearAggregation/None

Bar Chart

If You Selected Scatter Choose two Indices

ESTIMATED GNI PER

GDP PER CAPITA

GDP TOTAL

GNI PERCAPITA

GROSS FIXED CAPITAL FORMATION

Democratic Republic of the Congo

Denmark

Djibouti

Dominica

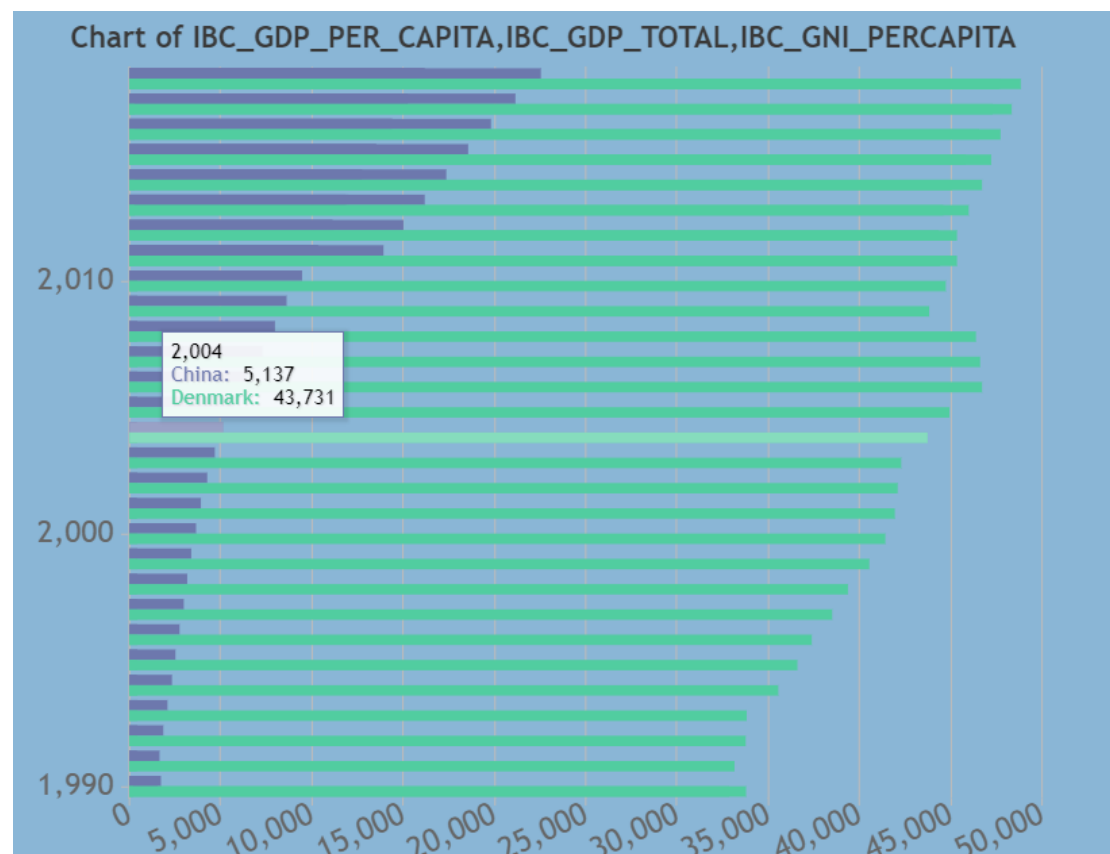
Choose Agreegate function

Choose Year

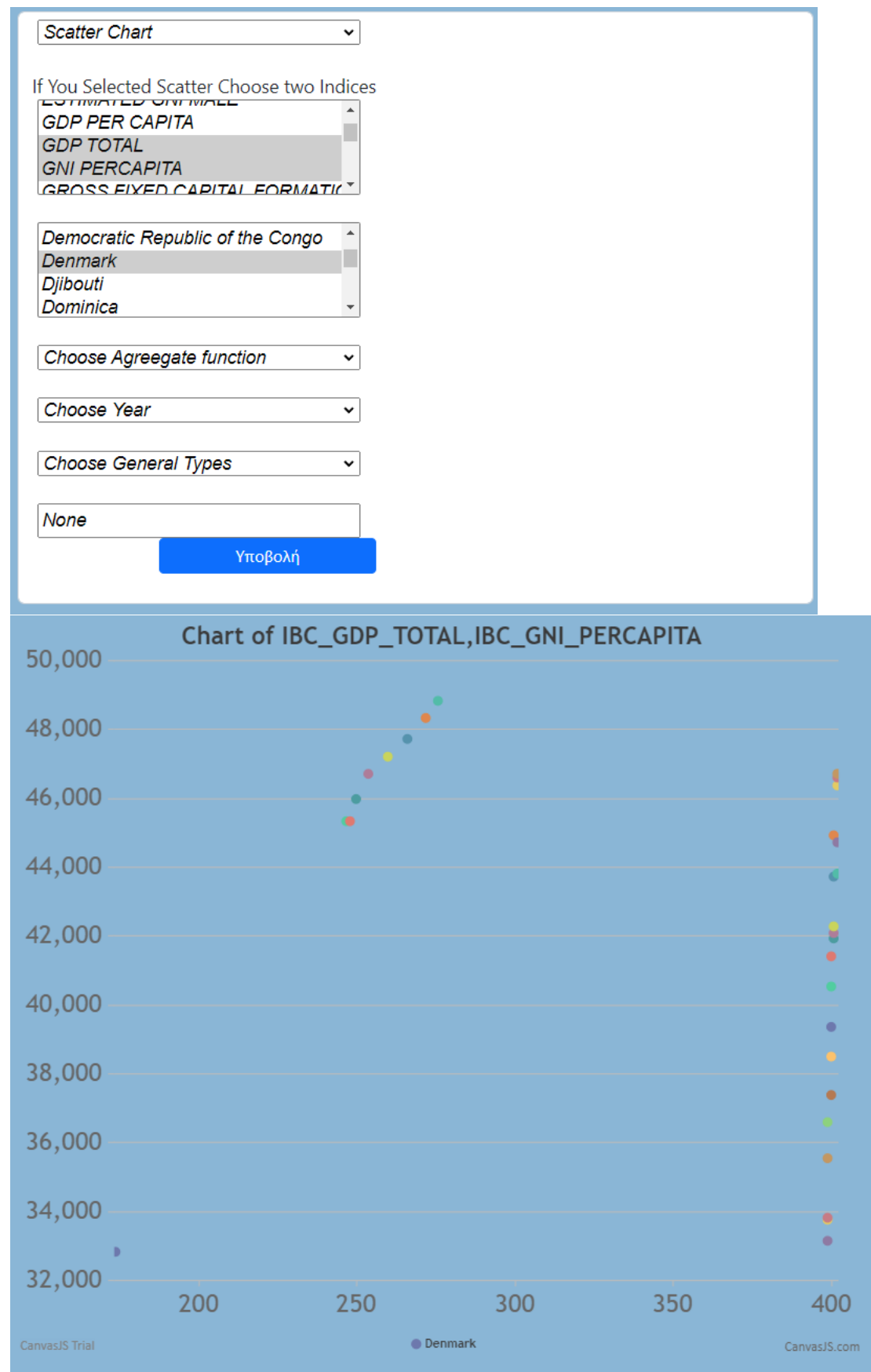
Choose General Types

None

Υποβολή



3.3 SCATTER PLOT

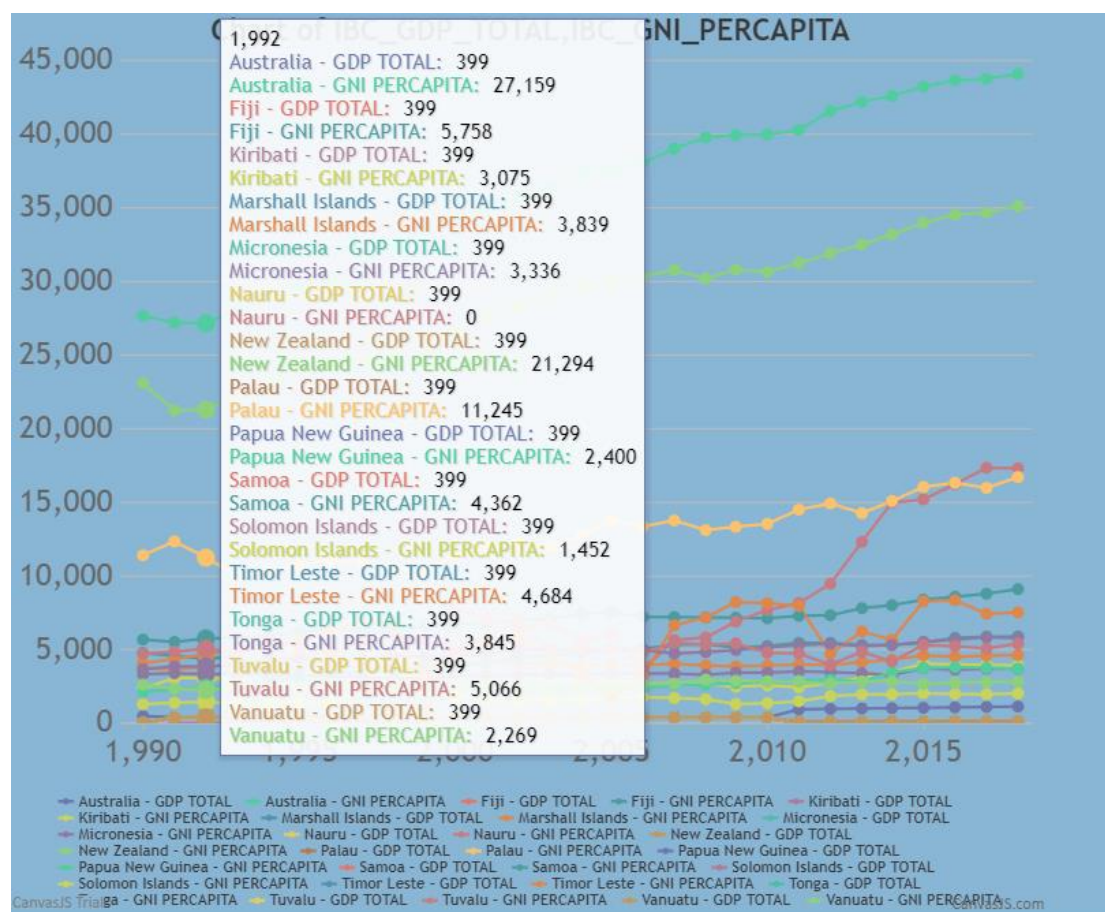


3.4 ΛΟΙΠΕΣ ΕΠΙΛΟΓΕΣ

Μπορούμε να επιλέξουμε αντί για χώρες σπαστά και περιοχές ή διάφορες ιδιότητες των χωρών όπως φαίνονται παρακάτω.

Choose General Types	
Continent Africa Antarctica Asia Europe North America Oceania South America	Sub-Saharan Africa Western Asia Western Europe
Region Name Africa Americas Asia Europe Oceania	Intermediate Region Name Caribbean Central America Channel Islands Eastern Africa Middle Africa South America Southern Africa Western Africa
Sub Region Name Australia and New Zealand Central Asia Eastern Asia Eastern Europe	Developed Or Developing #N/A Developed Developing
Other Small Island Developing Stages Land Locked Developing Countries Least Developed Countries	
Choose General Types	Choose General Types
None	None
Υποβολή	Υποβολή

Παράδειγμα για Ωκεανία είναι για επιληψία βέβαια αλλά λειτουργεί.



Τέλος για aggregate functions είναι ελεύθερη η επιλογή των ετών που θέλουμε να γκρουπάρουμε δεδομένα και επίσης ο χρήστης μπορεί να επιλέξει μεταξύ τεσσάρων όπως φαίνεται παρακάτω.

Line Chart

If You Selected Scatter Choose two Indices

Choose Data Type
FERTILITY RATE
DOMESTIC CREDITS
ESTIMATED GNI FEMALE

Antarctica
Antigua and Barbuda
Argentina
Armenia

Choose Aggregate function
SUM
MIN
MAX
AVG

3

Υποβολή

Line Chart

If You Selected Scatter Choose two Indices

Choose Data Type
FERTILITY RATE
DOMESTIC CREDITS
ESTIMATED GNI FEMALE

Choose country
Afghanistan
Aland Islands
Albania

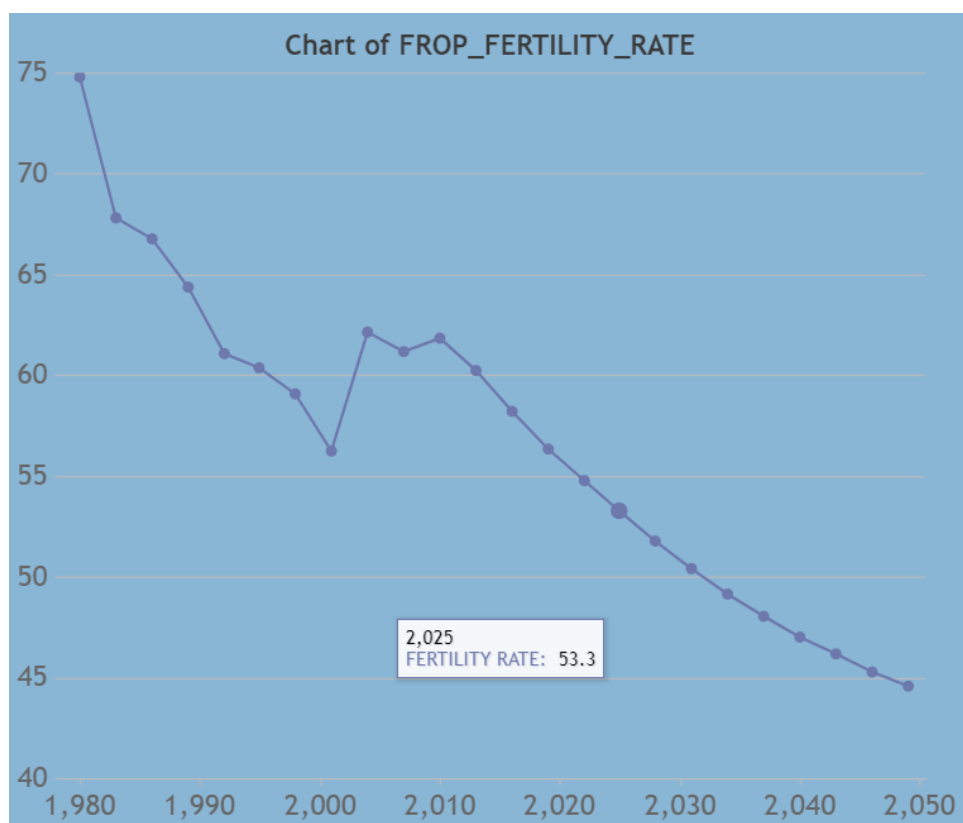
AVG

35-39

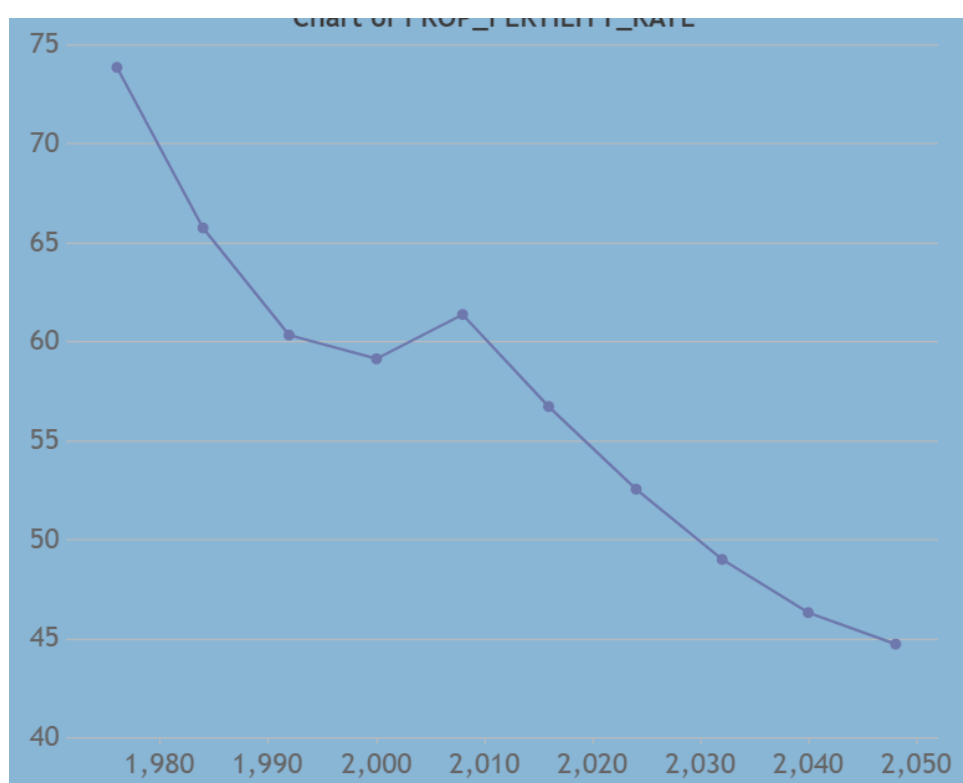
Choose General Types

8

Υποβολή



Το πάνω είναι ανά τρία χρόνια και το κάτω ανα οκτώ με τα ίδιες τις υπόλοιπες επιλογές



4 ΛΟΙΠΑ ΣΧΟΛΙΑ

Προβλήματα τα οποία δεν έχω διορθώσει

1. Όταν πατάμε το home και πάμε στην αρχική δεν κρατάει την επιλογή choose year που εμφανίζεται μόνο όταν επιλέγουμε indices με αυτόν τον διαχωρισμό πχ 4-9.
2. Στον πίνακες population male , female δεν είχα κάνει σωστά με τα flag τον διαχωρισμό 4-9,9-14.. και παραμένουν με 0,1,2,3,4,,,100 τα έφτιαξα μετά αλλά δεν κοίταξα πως μπορώ να τα ενσωματώσω αυτόματα στην library της βάσης.
3. Δεν έκανα συναρτήσεις validation κατά την εισαγωγή δεδομένων γιατί θεώρησα ότι δεν περάσει δηλαδή δεν είχε σωστή γραμμογράφιση δεν με ενδιαφέρει.
4. Το scatter plot με στήλες από διαφορετικούς πίνακες δεν παίζει προσπαθώ να το φτιάξω.
5. Δεν έχω αλλάξει τις αρχικές views about, contact.

Θα μπορούσε να γίνει πιο παραμετροποιήσιμο δηλαδή με άλλες 5 κλάσεις θα μπορούσα να φέρνω όλο το json του chart από την βάση, το προσπάθησα αλλά δεν το δέχονταν η canvasJS έχασα μια εβδομάδα εκεί και το παράτησα φέρνω μόνο τα δεδομένα.

Είναι δύο τα link για github

1. <https://github.com/siavasiatras/MYY030 Project>
2. <https://github.com/siavasiatras/WebMainApp>

Το πρώτο είναι η αρχική αντιμετώπιση με το spring boot έχει μόνο τις βάσεις και το δεύτερο το τελικό

Επειδή γράφτηκε σπαστά η εργασία ελπίζω να είχε μια συνοχή και να έβγαλε κάποιο νόημα, όσες φορές και αν την διάβασα δεν είδα κάτι κουλό. Αλλιώς όμως είναι να την βλέπει και κάποιος άλλος.