

ΠΡΟΧΩΡΗΜΕΝΑ ΘΕΜΑΤΑ
ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΕΦΑΡΜΟΓΩΝ
ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΕΡΓΑΣΙΑ ΓΙΑ
ΤΟ ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ 2021-2022

ΙΑΚΩΒΙΔΗΣ ΜΑΡΙΟΣ, ΑΜ 4063

ΤΕΛΙΚΗ ΑΝΑΦΟΡΑ

ΜΑΪΟΣ 2022

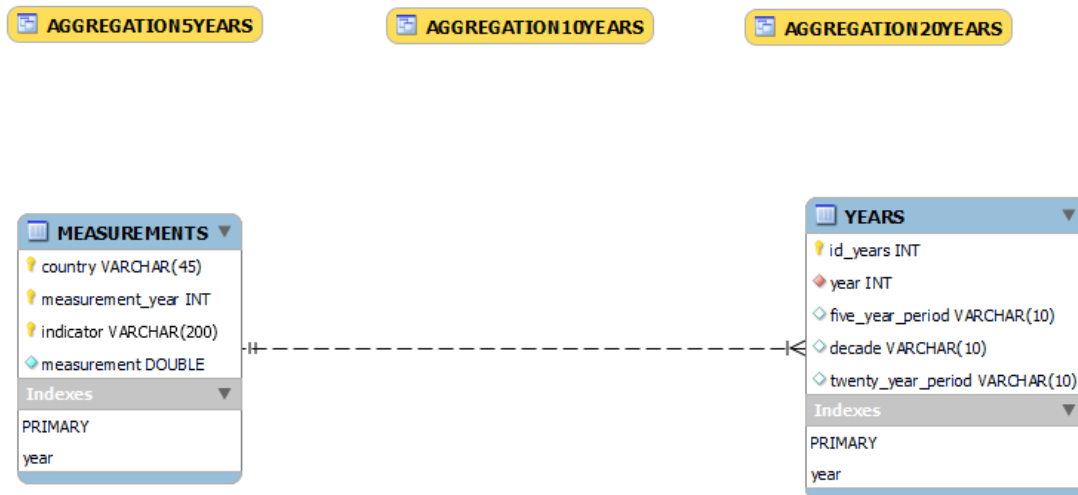
ΙΣΤΟΡΙΚΟ ΠΡΟΗΓΟΥΜΕΝΩΝ ΕΚΔΟΣΕΩΝ

Ημερομηνία	Έκδοση	Περιγραφή	Συγγραφέας
2022/03/22	1.0	Ολοκλήρωση υλοποίησης βάσης δεδομένων	Ιακωβίδης Μάριος
2022/05/16	2.0	Ολοκλήρωση εφαρμογής	Ιακωβίδης Μάριος
2022/05/21	3.0	Προσθήκη διαγραμμάτων και ολοκλήρωση αναφοράς	Ιακωβίδης Μάριος

1 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Στην παρούσα ενότητα περιγράφονται τα σχήματα της βάσης δεδομένων που χρησιμοποιείται στο Project.

1.1 ΣΧΕΣΙΑΚΟ ΣΧΗΜΑ ΣΕ ΛΟΓΙΚΟ ΕΠΙΠΕΔΟ



Σχήμα 1.1 Σχεσιακό σχήμα της βάσης δεδομένων του συστήματος

-- MySQL Workbench Forward Engineering

SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0;

SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0;

SET @OLD_SQL_MODE=@@SQL_MODE,
SQL_MODE='ONLY_FULL_GROUP_BY,STRICT_TRANS_TABLES,NO_ZERO_IN_DATE,NO_ZERO_DATE,ERROR_FOR_DIVISION
_BY_ZERO,NO_ENGINE_SUBSTITUTION';

-- Schema stats_database

-- Schema stats_database

CREATE SCHEMA IF NOT EXISTS `stats_database` DEFAULT CHARACTER SET utf8 ;

USE `stats_database` ;

-- Table `stats_database`.`MEASUREMENTS`

CREATE TABLE IF NOT EXISTS `stats_database`.`MEASUREMENTS` (

`country` VARCHAR(45) NOT NULL,

```
`measurement_year` INT NOT NULL,  
`indicator` VARCHAR(200) NOT NULL,  
`measurement` DOUBLE NOT NULL,  
PRIMARY KEY (`country`, `measurement_year`, `indicator`))  
ENGINE = InnoDB;  
  
CREATE INDEX `year` ON `stats_database`.`MEASUREMENTS` (`measurement_year`);  
  
-----  
-- Table `stats_database`.`YEARS`  
-----  
  
CREATE TABLE IF NOT EXISTS `stats_database`.`YEARS` (  
  `id_years` INT NOT NULL,  
  `year` INT NOT NULL,  
  `five_year_period` VARCHAR(10) NULL,  
  `decade` VARCHAR(10) NULL,  
  `twenty_year_period` VARCHAR(10) NULL,  
  PRIMARY KEY (`id_years`),  
  INDEX `year` (`year` ASC) INVISIBLE,  
  CONSTRAINT `year`  
    FOREIGN KEY (`year`)  
      REFERENCES `stats_database`.`MEASUREMENTS` (`measurement_year`)  
    ON DELETE NO ACTION  
    ON UPDATE NO ACTION)  
ENGINE = InnoDB;  
  
SET SQL_MODE=@OLD_SQL_MODE;  
SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;  
SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECK;
```

1.2 ΣΧΕΣΙΑΚΟ ΣΧΗΜΑ ΣΕ ΦΥΣΙΚΟ ΕΠΙΠΕΔΟ

Εδώ καταγράφονται οι ρυθμίσεις σε φυσικό επίπεδο.

1.2.1 ΡΥΘΜΙΣΗ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΤΟΥ DBMS

Η storage engine που χρησιμοποιείται και στους δύο πίνακες είναι η InnoDB. Το μέγεθος της βάσης είναι στα 5.5 MiB επομένως δώσαμε στο MySQL Server 8 MiB ως buffer pool size.

1.2.2 ΡΥΘΜΙΣΗ ΤΟΥ ΦΥΣΙΚΟΥ ΣΧΗΜΑΤΟΣ ΤΗΣ ΒΑΣΗΣ ΔΕΔΟΜΕΝΩΝ

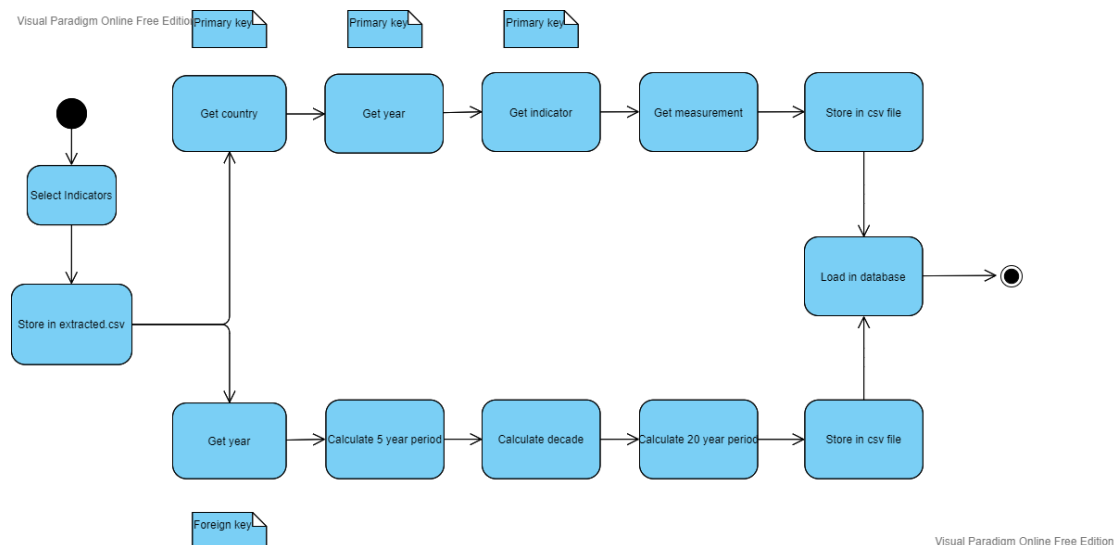
Εκτός των PRIMARY ευρετηρίων που ορίζονται αυτόματα σε κάθε πίνακα για τα πρωτεύοντα κλειδιά τους, ορίσαμε ένα ευρετήριο στο πεδίο measurement_year για τον πίνακα MEASUREMENTS και ένα ευρετήριο στο πεδίο year του πίνακα YEARS. Το παρακάτω συνέβη επειδή η εφαρμογή κάνει ερωτήσεις στη βάση οι οποίες για να απαντηθούν πρέπει να συνδυαστεί πληροφορία και από τους δύο πίνακες οπότε οι πίνακες θα συνενωθούν. Τα πεδία της συνένωσης θα είναι το πεδίο measurement_year και year με το year να είναι ξένο κλειδί στο measurement_year. Επομένως, προσθέσαμε ευρετήρια ώστε η σύγκριση στο πεδίο της συνένωσης να γίνεται πιο γρήγορα και να μην χάνετε πολύς χρόνος στην αναζήτηση των τιμών. Επίσης, ορίστηκαν τρεις όψεις (AGGREGATION5YEARS, AGGREGATION10YEARS, AGGREGATION20YEARS) οι οποίες κρατάνε τα ενδιάμεσα αποτελέσματα των ερωτήσεων οι οποίες λαμβάνουν δεδομένα για scatter plots, ώστε μετά με μία νέα ερώτηση η εφαρμογή να κρατήσει από το ενδιάμεσο αποτέλεσμα μόνο την πληροφορία που επιθυμεί και όχι την περιττή πληροφορία που προκύπτει από τον ενδιάμεσο αποτέλεσμα.

1.2.3 ΡΥΘΜΙΣΗ ΑΣΦΑΛΕΙΑΣ

Μόνο ο root user του εκάστοτε DBMS μπορεί να συνδεθεί με την βάση δεδομένων και να λάβει πληροφορία από αυτήν.

2 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΛΟΓΙΣΜΙΚΟΥ

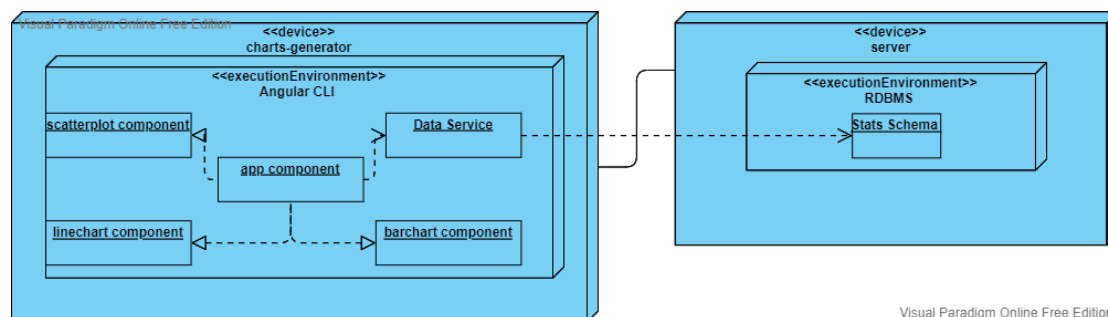
2.1 ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΚΑΙ ΔΟΜΗ ETL



Σχήμα 2. ETL με ένα UML activity diagram

Κατεβάσαμε από το data.worldbank.org τα αρχεία για τις 27 χώρες της Ευρωπαϊκής Ένωσης και επιλέξαμε 50 δείκτες που σχετίζονται με την παραγωγή ηλεκτρικής ενέργειας, την πρόσβαση σε αυτήν, τις εκπομπές CO2 καθώς και στατιστικά για ζωντανούς οργανισμούς που απειλούνται και σχετίζονται με την παραγωγή ενέργειας. Έπειτα δημιουργήσαμε τα `extracted.txt` αρχεία για κάθε χώρα όπου κρατάμε μόνο την πληροφορία για τους δείκτες που επιλέξαμε. Έπειτα οργανώσαμε την πληροφορία όπως πρέπει να είναι στα πεδία των πινάκων `MEASUREMENTS` και `YEARS` και υπολογίσαμε για κάθε έτος σε ποια πενταετία/δεκαετία/εικοσαετία ανήκει με έτος αναφοράς το 1960 που είναι το παλαιότερο έτος που διαθέτουμε δεδομένα. Έπειτα προχωρήσαμε στην φόρτωση των πινάκων στη βάση.

2.2 ΔΙΑΓΡΑΜΜΑ ΥΠΟΣΥΣΤΗΜΑΤΩΝ ΚΕΝΤΡΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

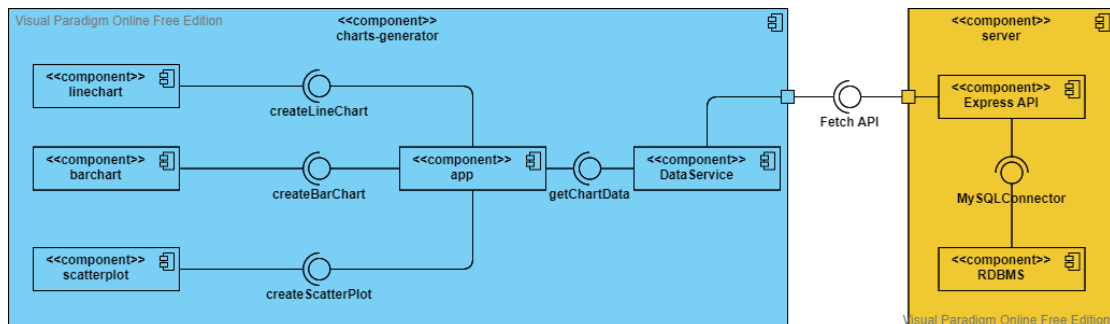


Σχήμα 2.1 Deployment diagram of the system

Τα υποσυστήματα του λογισμικού που κατασκευάσαμε είναι ο `charts-generator` και ο `server`. Ο `server` λαμβάνει από το `charts-generator` της πληροφορίες της ερώτησης που ζητά ο χρήστης, στέλνει την κατάλληλη ερώτηση στο RDBMS (έχουμε χρησιμοποιήσει την MySQL) και στέλνει τα αποτελέσματα πίσω στο `charts-generator`. Ο `charts-`

generator λαμβάνει τα δεδομένα στο Data Service του και τα δίνει στα Angular Components της εφαρμογής για να τα οπτικοποιήσουν.

2.3 ΔΙΑΓΡΑΜΜΑ ΚΛΑΣΕΩΝ ΚΕΝΤΡΙΚΗΣ ΕΦΑΡΜΟΓΗΣ

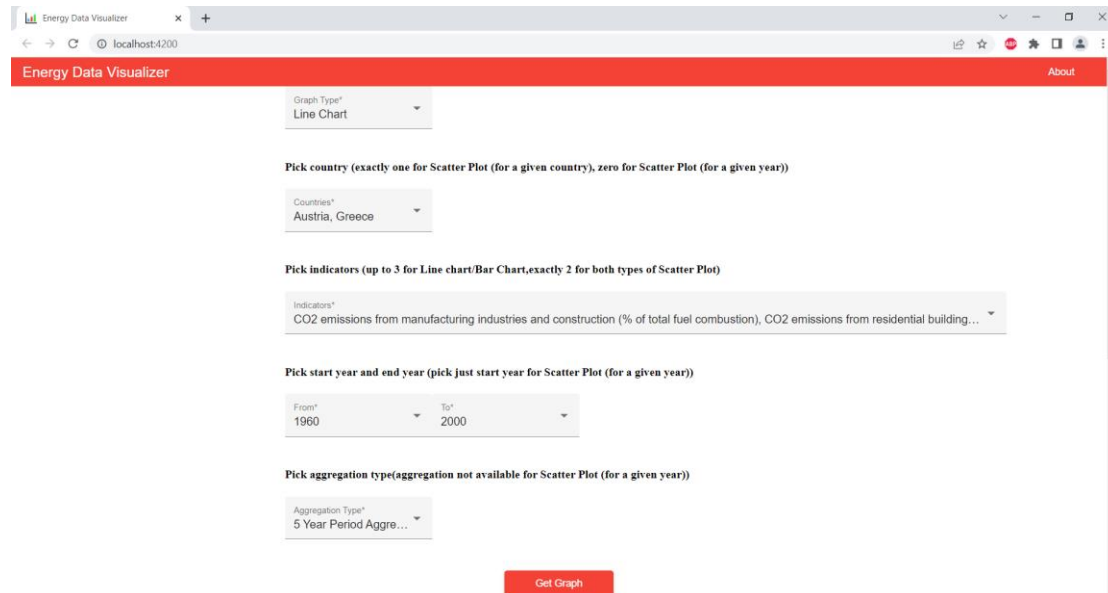


Σχήμα 2.3 Component Diagram of the system

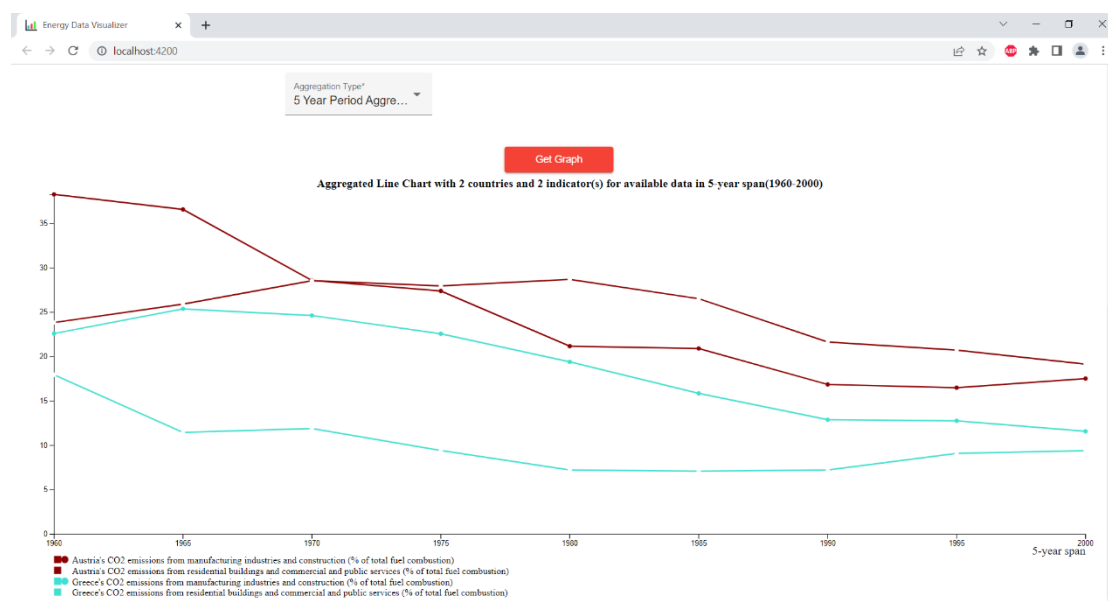
Ο server χρησιμοποιεί ένα Express API για να κάνει διαχείριση των HTTP αιτημάτων που καταφτάνουν σε αυτόν καθώς, και να στέλνει queries και να λαμβάνει απαντήσεις από το RDBMS. Ο charts-generator αποτελείται από το app component το οποίο δέχεται είσοδο από τον χρήστη και τροφοδοτείται με τα δεδομένα που ζήτησε ο χρήστης από το DataService. Έπειτα, τροφοδοτεί τα child components του (linechart, barchart, scatterplot) με αυτά, ώστε να δημιουργήσουν τα γραφήματα με την χρήση της βιβλιοθήκης D3.js. Η επικοινωνία μεταξύ charts-generator και server επιτυγχάνεται μέσω του Fetch API της Typescript το οποίο μας δίνει την δυνατότητα να αναμένουμε στον charts-generator την άφιξη των δεδομένων από τον server ώστε να μην τρέξει με undefined data (Promise data type).

3 ΥΠΟΔΕΙΓΜΑΤΑ ΕΡΩΤΗΣΕΩΝ ΚΑΙ ΑΠΑΝΤΗΣΕΩΝ

Αυτή είναι η κύρια οθόνη της εφαρμογής. Έχει 6 drop-down menus μέσω των οποίων ο χρήστης μπορεί να διαλέξει τύπο γραφήματος (graph type), χώρες (countries), δείκτες (indicators), χρονικό διάστημα (from-to), και οργάνωση του αποτελέσματος σε πενταετίες, δεκαετίες, εικοσαετίες (Aggregation type). Έστω ότι δίνουμε τις παρακάτω εισόδους:



Ζητάμε ένα Line Chart για την χώρες Αυστρία και Ελλάδα από το 1960 έως το 2000 με οργάνωση ανά πενταετία για τις εκπομπές CO2 από την κατασκευαστική βιομηχανία και τις εκπομπές CO2 από τις κατοικίες, καταστήματα και δημόσιες υπηρεσίες. Πατάμε το κουμπί Get Graph και εμφανίζεται από κάτω το εξής αποτέλεσμα:\

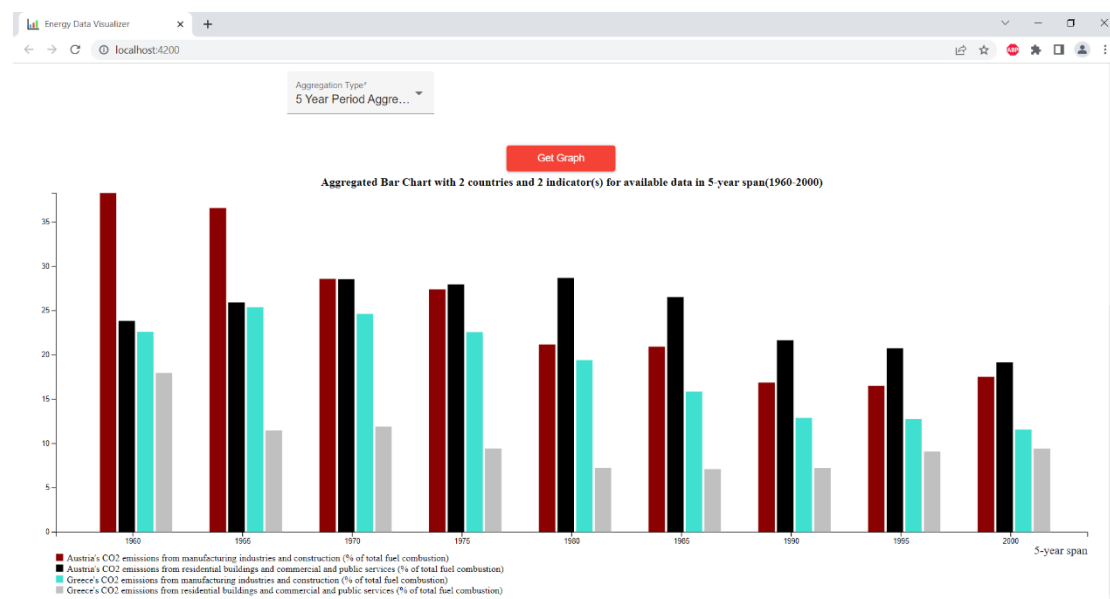


Βλέπουμε το γράφημα με τα δεδομένα που ζητήσαμε. Στο κάτω αριστερά μέρος φαίνεται το υπόμνημα με την πληροφορία που εκφράζει κάθε χρωματικός συνδυασμός.

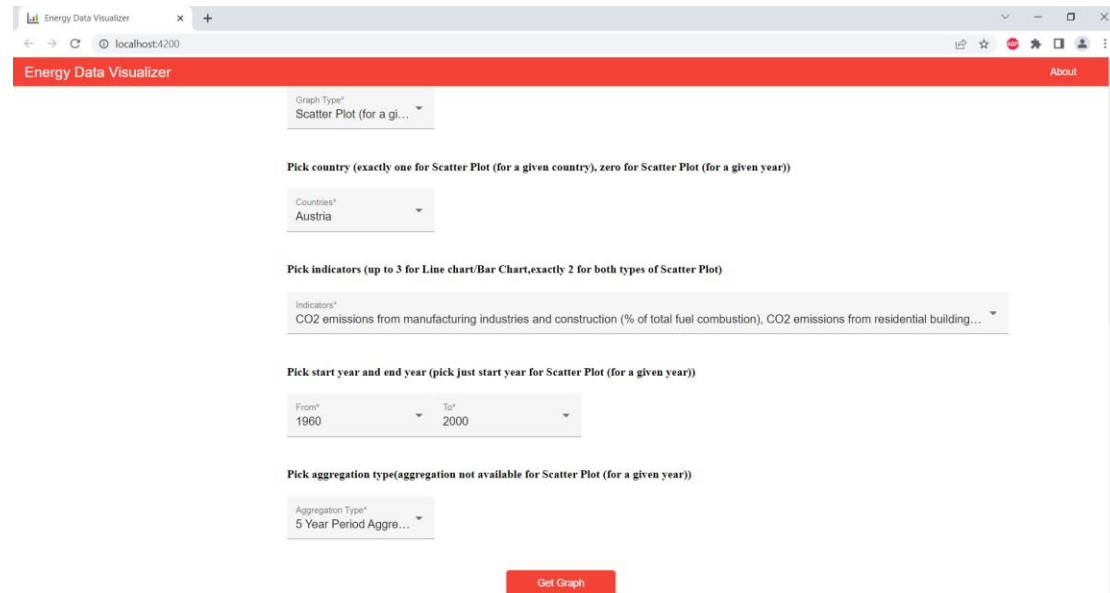
Με την ίδια διαδικασία μπορούμε να λάβουμε την παραπάνω πληροφορία και υπό την μορφή bar chart:

The screenshot shows the 'Energy Data Visualizer' web application interface. The 'Graph Type' is set to 'Bar Chart'. The 'Country' is set to 'Austria, Greece'. The 'Indicators' are set to 'CO2 emissions from manufacturing industries and construction (% of total fuel combustion), CO2 emissions from residential building...'. The 'From' year is 1960 and the 'To' year is 2000. The 'Aggregation Type' is set to '5 Year Period Aggre...'. A red 'Get Graph' button is visible at the bottom.

Λαμβάνουμε το εξής αποτέλεσμα:



Τέλος για τους δύο δείκτες μπορούμε να κάνουμε ένα scatter plot για την Αυστρία:



Λαμβάνουμε το παρακάτω αποτέλεσμα:



4 ΛΟΙΠΑ ΣΧΟΛΙΑ

Περιορίσαμε το πλήθος των γραμμών που εμφανίζονται στο line chart σε 12 και τον πλήθος των bars σε 3 στο bar chart ώστε να είναι το διάγραμμα ευανάγνωστο. Επίσης κάτω αριστερά εμφανίζεται υπόμνημα που εξηγεί τις αντιστοιχίες χρωμάτων-πληροφορία στο γράφημα. Στο line chart οι διαφορετικοί δείκτες απεικονίζονται με διαφορετικό χρώμα κουκίδων και ο αντίστοιχος χρωματισμός εμφανίζεται στο υπόμνημα.