ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ ΤΜΗΜΑ ΜΗΧ. Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ

ΠΡΟΧΩΡΗΜΕΝΑ ΘΕΜΑΤΑ ΤΕΧΝΟΛΟΓΙΑΣ ΚΑΙ ΕΦΑΡΜΟΓΩΝ ΒΑΣΕΩΝ ΔΕΔΟΜΕΝΩΝ

ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΗ ΕΡΓΑΣΙΑ ΓΙΑ ΤΟ ΑΚΑΔΗΜΑΪΚΟ ΕΤΟΣ 2024

ΣΙΑΚΑΒΆΡΑ ΘΕΜΙΣΤΟΚΛΕΊΑ, 4786

ΤΕΛΙΚΗ ΑΝΑΦΟΡΑ

ΜΑΪΟΣ 2024

ΙΣΤΟΡΙΚΟ ΠΡΟΗΓΟΥΜΕΝΩΝ ΕΚΔΟΣΕΩΝ

Ημερομηνία	Έκδοση	Περιγραφή	Συγγραφέας
2024/04/09	1.0	Τροποποίηση αρχείων csv	4786
2024/04/13	2.0	Τελειοποίηση csv Αρχείων	4786
2024/05/10	3.0	Δημιουργία βάσης - DDL scripts	4786
2024/05/13	4.0	Δημιουργία ETL scripts - Έγινε backup της βάσης	4786
2024/05/20	5.0	Στήσιμο προγραμματιστικού περιβάλλοντος – Σύνδεση βάσης και Πειράματα	4786
2024/05/25	6.0	Δημιουργία Κλάσεων για κάθε αναζήτηση – Δημιουργία main menu	4786
2024/05/30	7.0	Αναφορά - github	4786

1 ΒΑΣΗ ΔΕΔΟΜΕΝΩΝ

Στην παρούσα ενότητα περιγράφεται το σχήμα βάσης δεδομένων που χρησιμοποιείται στο Project.

1.1 APXEIA CSV

Αρχικά, τα δεδομένα πάρθηκαν από τα links: https://www.kaggle.com/datasets/sshashankrajak/countries

https://climatedata.imf.org/pages/climatechange-data

Αποτελούνται από 5 αρχεία csv, τα οποία τροποποιήθηκαν ώστε:

- 1. Όλα τα αρχεία να έχουν τις ίδιες χώρες.
- 2. Η στήλη Iso_Code να είναι σε όλα τα αρχεία.
- 3. Στα αρχεία με τις χρονολογίες, να υπάρχουν πολλές γραμμές για κάθε χώρα, και μία στήλη 'Year" για τη χρονολογία και μία "Value" για την τιμή.

Επιλέχθηκε να υπάρχουν πολλοί πίνακες μετρήσεων αντί για ένας.

Η τροποποίηση έγινε μέσω του αρχείου "modifyCSV.java", το οποίο παίρνει τα πρωτότυπα αρχεία και τα αλλάζει όταν την καλέσουμε στο τερματικό.

1.2 ΔΗΜΙΟΥΡΓΙΑ ΒΑΣΗΣ

Αφού τα αρχεία είναι πλέον στην επιθυμητή μορφή, φτιάχτηκε η βάση δεδομένων.

Για τη δημιουργία της, χρησιμοποιήθηκε το εργαλείο "MySQL WorkBench". Η βάση ονομάστηκε "countriesData" και αποτελείται από 5 tables (όσα και τα αρχεία).

Το primary κλειδί για τον πίνακα των χωρών είναι το isoCode. Το primary κλειδί για τους υπόλοιπους πίνακες είναι ένα id που δημιουργεί η βάση και αυξάνεται αυτόματα. Και foreign κλειδί είναι το isoCode του πίνακα χωρών. Τα DDL scripts που χρησιμοποιήθηκαν είναι τα εξής:

CREATE TABLE IF NOT EXISTS countries(

```
iso VARCHAR(10) NOT NULL,
iso3 VARCHAR(10) NOT NULL UNIQUE,
isoCode INT PRIMARY KEY,
fips VARCHAR(255) NOT NULL,
displayName VARCHAR(255) NOT NULL,
officialName VARCHAR(255) NOT NULL,
capital VARCHAR(255) NOT NULL,
continent VARCHAR(255) NOT NULL,
currencyCode VARCHAR(255) NOT NULL,
currencyName VARCHAR(255) NOT NULL,
phone VARCHAR(50) NOT NULL,
regionCode INT NOT NULL,
regionName VARCHAR(255) NOT NULL,
subRegionCode INT NOT NULL,
subRegionName VARCHAR(255) NOT NULL,
intermediateRegionCode VARCHAR(255),
intermediateRegionName VARCHAR(255),
statusC VARCHAR(255) NOT NULL,
development VARCHAR(255) NOT NULL,
sids VARCHAR(10),
lldc VARCHAR(10),
ldc VARCHAR(10),
areaSqKm INT NOT NULL,
population INT NOT NULL
```

)ENGINE=InnoDB;

```
CREATE TABLE IF NOT EXISTS temperatureChange(
        id INT AUTO_INCREMENT PRIMARY KEY,
        isoCode INT,
        country VARCHAR(255),
        iso2 VARCHAR(10),
        iso3 VARCHAR(10),
        indicatorC VARCHAR(255),
        unit VARCHAR(255),
        sourceC VARCHAR(255),
        ctsCode VARCHAR(10),
        ctsName VARCHAR(255),
        ctsFullDescriptor VARCHAR(255),
        yearChange VARCHAR(255),
        valueChange FLOAT,
        FOREIGN KEY (isoCode) REFERENCES countries(isoCode)
)ENGINE = InnoDB;
CREATE TABLE IF NOT EXISTS climateDisasters(
        id INT AUTO_INCREMENT PRIMARY KEY,
        isoCode INT,
        country VARCHAR(255),
        iso2 VARCHAR(10),
        iso3 VARCHAR(10),
        indicatorC VARCHAR(255),
        unit VARCHAR(255),
        sourceC VARCHAR(255),
        ctsCode VARCHAR(10),
        ctsName VARCHAR(255),
        ctsFullDescriptor VARCHAR(255),
        yearChange VARCHAR(255),
        valueChange INT,
        FOREIGN KEY (isoCode) REFERENCES countries(isoCode)
)ENGINE = InnoDB;
```

```
CREATE TABLE IF NOT EXISTS landCover(
        id INT AUTO_INCREMENT PRIMARY KEY,
        isoCode INT,
        country VARCHAR(255),
        iso2 VARCHAR(10),
        iso3 VARCHAR(10),
        indicatorC VARCHAR(255),
        unit VARCHAR(255),
        sourceC VARCHAR(255),
        ctsCode VARCHAR(10),
        ctsName VARCHAR(255),
        ctsFullDescriptor VARCHAR(255),
        climateInfluence VARCHAR(255),
        yearChange VARCHAR(255),
        valueChange FLOAT,
        FOREIGN KEY (isoCode) REFERENCES countries(isoCode)
)ENGINE = InnoDB;
CREATE TABLE IF NOT EXISTS forestCarbon(
        id INT AUTO_INCREMENT PRIMARY KEY,
        isoCode INT,
        country VARCHAR(255),
        iso2 VARCHAR(10),
        iso3 VARCHAR(10),
        indicatorC VARCHAR(255),
        unit VARCHAR(255),
        sourceC VARCHAR(255),
        ctsCode VARCHAR(10),
        ctsName VARCHAR(255),
        ctsFullDescriptor VARCHAR(255),
        yearChange VARCHAR(255),
        valueChange FLOAT,
        FOREIGN KEY (isoCode) REFERENCES countries(isoCode)
)ENGINE = InnoDB;
```

Μετά την εκτέλεση των παραπάνω scripts, μπορούμε να της φορτώσουμε τα δεδομένα των αρχείων csv.

1.3 ETL SCRIPTS

Αρχικά, θα χρειαστεί να μεταφερθούν τα τροποποιημένα αρχεία csv στο φάκελο ProgramData/MySQL/MySQL Server 8.0/Uploads, όπου δείχνει η εντολή

```
SHOW VARIABLES LIKE "secure_file_priv";
```

Τα ETL scripts είναι τα εξής:

```
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/countries.csv'
INTO TABLE countries
FIELDS TERMINATED BY ','
ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 1 ROWS
(iso, iso3, isoCode, fips, displayName, officialName, capital, continent, currencyCode, currencyName, phone,
regionCode, regionName, subRegionCode, subRegionName, intermediateRegionCode,
intermediateRegionName, statusC, development, sids, lldc, ldc, areaSqKm, population);
CREATE TEMPORARY TABLE temp_temperatureChange (
 isoCode INT,
 country VARCHAR(255).
 iso2 VARCHAR(10),
 iso3 VARCHAR(10),
 indicatorC VARCHAR(255),
 unit VARCHAR(255),
 sourceC VARCHAR(255),
 ctsCode VARCHAR(10),
 ctsName VARCHAR(255),
 ctsFullDescriptor VARCHAR(255),
 yearChange VARCHAR(255),
 valueChange VARCHAR(255) -- Use VARCHAR to load all data without errors
) ENGINE = InnoDB;
LOAD DATA LOCAL INFILE 'C:/ProgramData/MySQL/MySQL Server
8.0/Uploads/Annual_Surface_Temperature_Change.csv'
INTO TABLE temp temperatureChange
FIELDS TERMINATED BY ','
ENCLOSED BY ""
LINES TERMINATED BY '\n'
IGNORE 1 ROWS
(isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName, ctsFullDescriptor, yearChange,
valueChange);
INSERT INTO temperatureChange (isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName,
ctsFullDescriptor, yearChange, valueChange)
SELECT isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName, ctsFullDescriptor,
yearChange,
 CASE
   WHEN valueChange REGEXP '^-?[0-9]+(\.[0-9]+)?$' THEN CAST(valueChange AS DOUBLE)
   ELSE NULL
 END
FROM temp_temperatureChange;
DROP TEMPORARY TABLE temp_temperatureChange;
```

```
CREATE TEMPORARY TABLE temp_climateDisasters (
 isoCode INT,
 country VARCHAR(255),
 iso2 VARCHAR(10),
 iso3 VARCHAR(10),
 indicatorC VARCHAR(255),
 unit VARCHAR(255),
 sourceC VARCHAR(255),
 ctsCode VARCHAR(10),
 ctsName VARCHAR(255),
 ctsFullDescriptor VARcHAR(255),
 yearChange VARCHAR(255),
 valueChange VARCHAR(255) -- Use VARCHAR to load all data without errors
) ENGINE = InnoDB;
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/Climate-
related_Disasters_Frequency.csv'
INTO TABLE temp_climateDisasters
FIELDS TERMINATED BY ','
ENCLOSED BY ""
LINES TERMINATED BY '\n'
IGNORE 1 ROWS
(isoCode, country, iso2, iso3, indicatorC, unit, source, ctsCode, ctsName, ctsFullDescriptor, yearChange,
valueChange);
INSERT INTO climateDisasters (isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName,
ctsFullDescriptor, yearChange, valueChange)
SELECT isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName, ctsFullDescriptor,
yearChange,
 CASE
   WHEN valueChange REGEXP '^-?[0-9]+$' THEN CAST(valueChange AS SIGNED)
   ELSE NULL
 END
FROM temp_climateDisasters;
DROP TEMPORARY TABLE temp_climateDisasters;
CREATE TEMPORARY TABLE temp_landCover (
 isoCode INT,
 country VARCHAR(255),
 iso2 VARCHAR(10),
 iso3 VARCHAR(10),
 indicatorC VARCHAR(255),
 unit VARCHAR(255),
 sourceC TEXT,
 ctsCode VARCHAR(10),
 ctsName VARCHAR(255),
 ctsFullDescriptor VARCHAR(255),
 climateInfluence VARCHAR(255),
 vearChange VARCHAR(255),
 valueChange VARCHAR(255) -- Use VARCHAR to load all data without errors
) ENGINE = InnoDB;
LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/Land_Cover_Accounts.csv'
INTO TABLE temp landCover
FIELDS TERMINATED BY ','
ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 1 ROWS
(isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName, ctsFullDescriptor,
climateInfluence, yearChange, valueChange);
```

IGNORE 1 ROWS

INSERT INTO landCover (isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName, ctsFullDescriptor, climateInfluence, yearChange, valueChange)

SELECT isoCode, country, iso2, iso3, indicatorC, unit, sourceC, ctsCode, ctsName, ctsFullDescriptor, climateInfluence, yearChange,

CASE

WHEN valueChange REGEXP '^-?[0-9]+(\.[0-9]+)?\$' THEN CAST(valueChange AS FLOAT)

ELSE NULL

END

FROM temp_landCover;

LOAD DATA INFILE 'C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/Forest_and_Carbon.csv'
INTO TABLE forestCarbon
FIELDS TERMINATED BY ','
ENCLOSED BY '"'

LINES TERMINATED BY '\n'

(iso Code, country, iso 2, iso 3, indicator C, unit, source C, cts Code, cts Name, cts Full Descriptor, year Change, value Change);

Η δημιουργία προσωρινών πινάκων γίνεται καθώς κάποιες τιμές στην στήλη Value των αρχείων είναι null, και επομένως δεν μπορούν να περαστούν απευθείας στην βάση. Πρώτα θα πρέπει να γίνει cast. Αυτό το συμπεραίνουμε αν προσπαθήσουμε να τα φορτώσουμε όπως κάνουμε για το αρχείο "countries.csv" ή "forest_and_carbon.csv".

Καταλαβαίνουμε ότι φορτώθηκαν σωστά από τα μηνύματα στο MySQL Workbench.



Τα errors που φαίνονται είναι η προσπάθεια απευθείας φόρτωσης.

1.4 ΒΑСΚΟΡ ΒΑΣΗΣ

Τέλος, έγινε backup της βάσης δεδομένων με τον εξής τρόπο:

Ανοίγουμε το τερματικό ως administrator και μπαίνουμε στον φάκελο του server, και συγκεκριμένα στο bin. Έπειτα, τρέχουμε το εξής:

```
C:\Program Files\MySQL\MySQL Server 8.0\bin>mysqldump -u root -p countriesdata > backup.sql
Enter password: *****************
C:\Program Files\MySQL\MySQL Server 8.0\bin>
```

2 ΠΡΟΓΡΑΜΜΑΤΙΣΤΙΚΌ ΠΕΡΙΒΑΛΛΟΝ

Το προγραμματιστικό περιβάλλον που χρησιμοποιήθηκε είναι το eclipse 2020-09, γλώσσα java 17.0.10 και για την οπτικοποίηση των δεδομένων JavaFX.

Για την επιβεβαίωση σωστής λειτουργίας, δημιουργήθηκε η κλάση "MainApp" όπου χρησιμοποιεί τη JavaFX.

Για να τρέχουν σωστά οι κλάσεις που δημιουργήθηκαν, πρώτα επιλέγουμε την επιλογή:

Run as → Run Configurations

Στο παράθυρο που ανοίγει, επιλέγουμε την καρτέλα "Arguments" και στα VM arguments, επιλέγουμε Variables. Πατάμε edit variables → New: --module-path "C:/Users/Siak/Desktop/vasiliadis/javafx-sdk-22.0.1/lib" --add-modules javafx.controls,javafx.fxml

Διαλέγοντας δηλαδή το μονοπάτι που έχουμε κατεβάσει τη βιβλιοθήκη javafx.

Τέλος, Apply- Close.

3 ΥΠΟΔΕΙΓΜΑΤΑ ΕΡΩΤΗΣΕΩΝ ΚΑΙ ΑΠΑΝΤΗΣΕΩΝ

Η κλάση "MainMenu" περιέχει τη main συνάρτηση του προγράμματός μας, και όταν την τρέχουμε εμφανίζει το εξής:



Η πρώτη κατηγορία είναι το είδος indicator.

Η δεύτερη επιλογή είναι η αρχική χρονολογία.

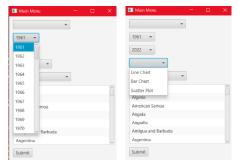
Η Τρίτη κατηγορία είναι η τελική χρονολογία.

Η τέταρτη είναι το είδος του γραφήματος.

Η Πέμπτη είναι το είδος της κλιματικής καταστροφής.(είναι διαθέσιμη επιλογή μόνο όταν indicator = Disasters)

Τέλος, το παράθυρο με όλες τις διαθέσιμες χώρες, από τις οποίες ο χρήστης μπορεί να επιλέξει μία ή πολλές. (αν θέλει πολλές κρατάει το CTRL πατημένο)

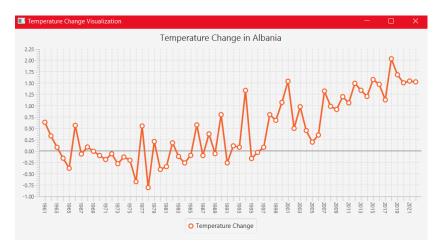




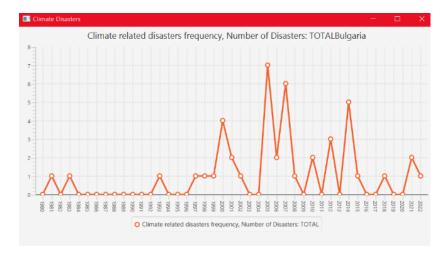


Μερικά Παραδείγματα:

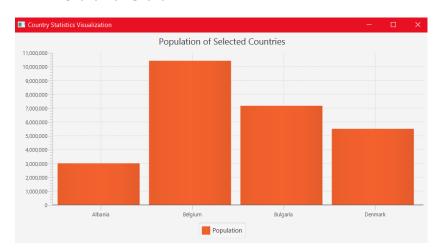
 Ερώτηση: Χώρα Albania, Query: Temperature Change, Chart: Line Chart, Years: 1961-2022



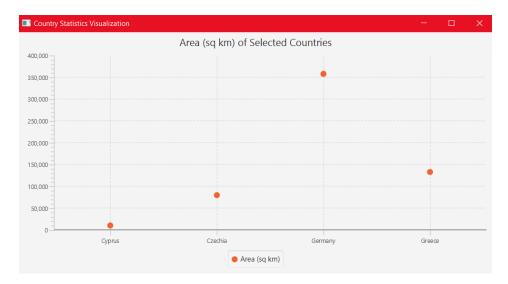
 Ερώτηση: Χώρα Bulgaria, Query: Disasters, Chart: Line Chart, Years: 1961-2022, Disaster: TOTAL



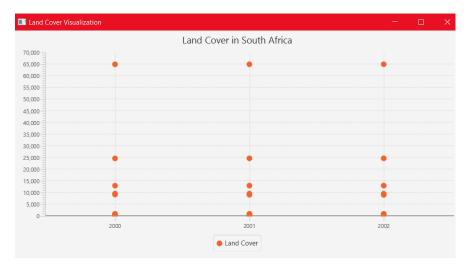
3. Ερώτηση: Χώρες Albania- Belgium- Bulgaria- Denmark, Query: Population, Chart: Bar Chart



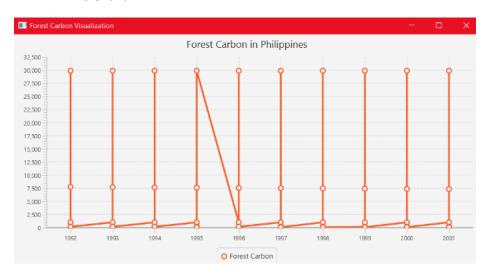
4. Ερώτηση: Χώρες Cyprus- Czechia- Germany- Greece, Query: Area (sq km), Chart: Scatter Plot

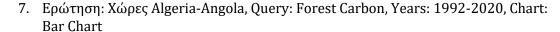


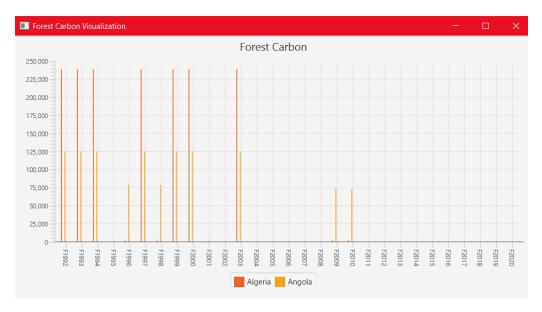
5. Ερώτηση: Χώρα South Africa, Query: Land Cover, Chart: Scatter Plot



6. Ερώτηση: Χώρα Philippines, Query: Forest Carbon, Years: 1992-2001, Chart: Line Chart





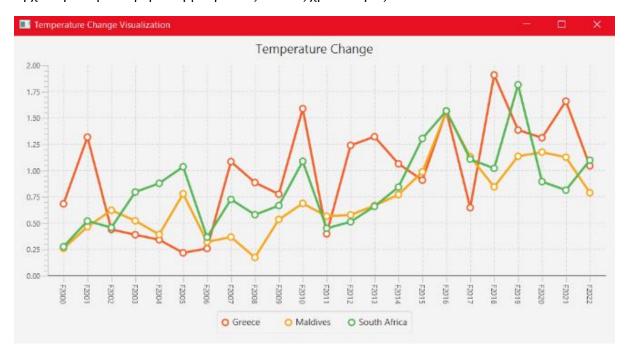


4 ΜΕΛΕΤΗ

Σε αυτή την ενότητα θα μελετήσουμε τις διαφορές ανάμεσα σε 3 χώρες: Greece, Maldives, South Africa.

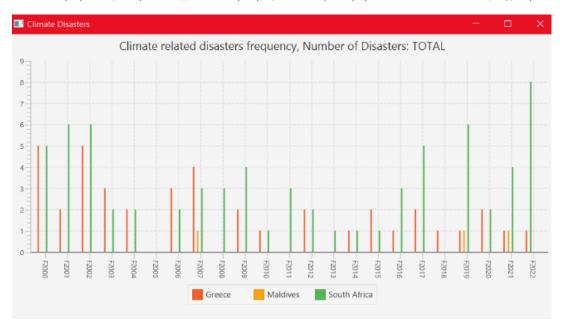
Επιλέχθηκαν οι συγκεκριμένες γιατί βρίσκονται σε διαφορετικά γεωγραφικά μέρη, οπότε θα υπάρχουν διαφορές.

Αρχικά για την διαφορά θερμοκρασίας από τις χρονολογίες 2000-2022:



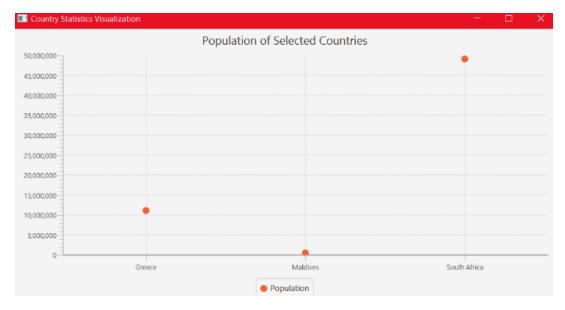
Παρατηρούμε ότι στην Ελλάδα υπάρχουν μεγαλύτερες αλλαγές σε σχέση με τις άλλες χώρες. Γενικά οι τιμές διαφέρουν. Με κάποιες εξαιρέσεις όπως πχ το 2013, όπου έχουν και οι 3 την ίδια διαφορά.

Όσων αφορά τις κλιματικές καταστροφές, και συγκεκριμένα το σύνολό τους, έχουμε:



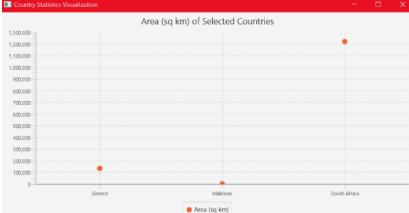
Παρατηρούμε ότι η Ελλάδα δεν έχει πολλές κλιματικές καταστροφές, καθώς έχουμε αποτέλεσμα ελάχιστες φορές. Σε αντίθεση με την Ελλάδα, η Νότια Αφρική έχει κάθε χρόνο, με αποκορύφωμα το έτος 2022. Οι Μαλδίβες έχουν επίσης κάθε χρονιά, αρκετές καταστροφές.

Για τον πληθυσμό κάθε χώρας, δεν χρειαζόμαστε πλέον χρονολογίες.



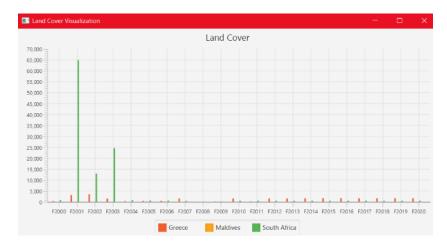
Η Νότια Αφρική εμφανέστατα έχει περισσότερο πληθυσμό από όλες. Οι Μαλδίβες τον λιγότερο.





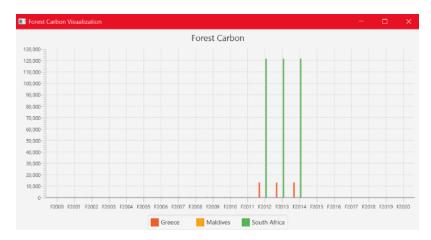
Εμφανέστατα πάλι η Νότια Αφρική αποτελείται από περισσότερα χιλιόμετρα, σε αντίθεση με τις Μαλδίβες.

Όσων αφορά το Land Cover, από το 2000 μέχρι το 2020 έχουμε:



Παρατηρούμε ότι η Νότια Αφρική έχει τεράστια διαφορά το 2001 σε σχέση με το 2020. Γενικά η τιμή του 2001 είναι εμφανέστατα διαφορετική από ολες τις τιμές των άλλων χρονολογιών για την ίδια χώρα, αλλά και διαφορετική από τις άλλες χώρες.

Τέλος για το CO2:



Υπάρχουν τιμές μόνο για τις χρονολογίες 2012-2013-2014 για τις Ελλάδα και Νότια Αφρική. Οι Μαλδίβες δεν έχουν δεδομένα. Για άλλη μια φορά η Νότια Αφρική έχει μεγαλύτερες τιμές.