**ΟΜΑΔΙΚΗ ΕΚΘΕΣΗ - ΟΜΑΔΑ 2 ( COVID 19 )**

Το project Covid 19 Greece Data Analysis έχει στόχο την αξιοποίηση των δεδομένων που διατίθεται σε αρχείο csv στη διεύθυνση

https://raw.githubusercontent.com/Sandbird/covid19-Greece/master/cases.csv

,τα οποία δεδομένα ενημερώνονται καθημερινά απο το Υπουργείο Υγείας και η παρουσίασή τους με τρόπο που το καθιστά ευανάγνωστο και κατανοητό προς ειδικούς και μη-ειδικούς.

Η παρουσίαση των στοιχείων ώφειλε να γίνει με αναλυτικό τρόπο παρουσιάζοντας τα νούμερα των μεγεθών ,αλλά και με περιγραφικό τρόπο μέσω διαφόρων διαγραμμάτων Επίσης έπρεπε να κατασκευαστεί και ένα γραφικό περιβάλλον που θα μπορούσε να υποστηρίξει τις υπόλοιπες λειτουργίες του προγράμματος.

Η κύριες βιβλιοθήκες που χρησιμοποιήθηκαν είναι οι:

-tkinter: Για τη δημιουργία του γραφικού περιβάλλοντος

-pandas: Για την αξιοποίηση των διαφόρων δεδομένων του csv αρχείου

-matplotlib: Για τη δημιουργία των γραφικών παραστάσεων του προγράμματος

-plotly: Σαν εναλλακτικό τρόπο παρουσίασης παραστάσεων

-sklearn: για το μοντελο προβλεψης

Το project μοιράστηκε σε 4 μέρη:

Tέταρτο Μέρος: **Γραφικό Περιβάλλον Tkinter**

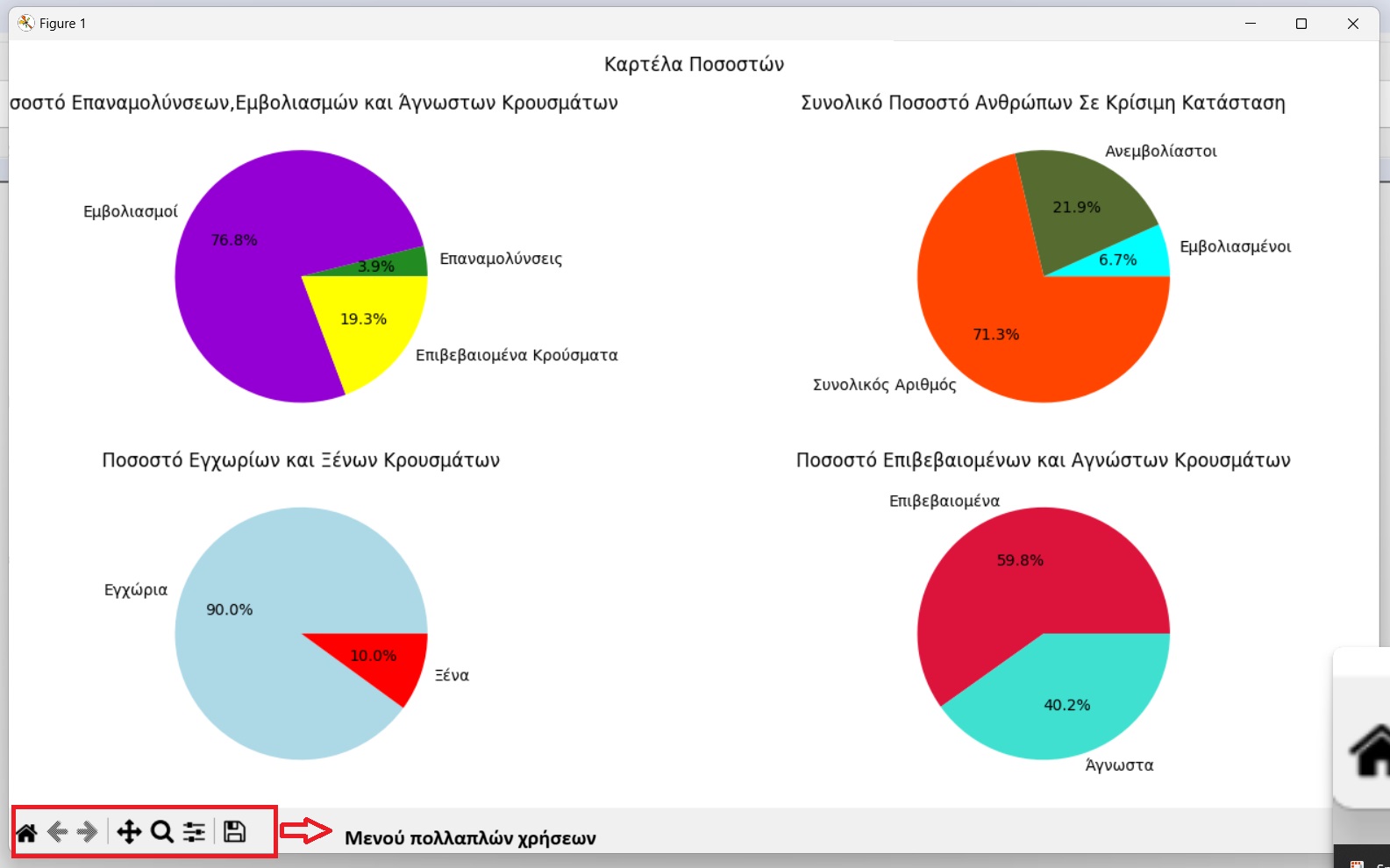
Ο στόχος και ο σκοπός μας για το γραφικό περιβάλλον είναι να παρέχει στους χρήστες ένα εύχρηστο και οπτικά ελκυστικό εργαλείο για την παρακολούθηση των δεδομένων της πανδημίας, επιτρέποντας την εύκολη πρόσβαση σε σημαντικές πληροφορίες όπως είδατε και παραπάνω.

Το γραφικό περιβάλλον του προγράμματος έχει αναπτυχθεί με τη χρήση της βιβλιοθήκης Tkinter της Python, η οποία επιτρέπει τη δημιουργία διαδραστικών και ευέλικτων γραφικών εφαρμογών. Ακολουθούν οι βασικές λειτουργίες και τα χαρακτηριστικά του GUI:

1. **Κεντρικό Παράθυρο**: Το κύριο παράθυρο του προγράμματος ανοίγει με έναν απλό και κατανοητό σχεδιασμό, προσφέροντας διάφορα κουμπιά για την πλοήγηση και απεικόνιση των δεδομένων. Το παράθυρο είναι αρκετά ευρύχωρο για να φιλοξενήσει τα κουμπιά των γραφημάτων.
2. **Κουμπιά Ενεργειών**: Υπάρχει μια σειρά από κουμπιά τα οποία επιτρέπουν στο χρήστη να επιλέξει ποιο γράφημα ή πίνακα δεδομένων θέλει να δει. Κάθε κουμπί αντιστοιχεί σε μια συγκεκριμένη ανάλυση ή γράφημα, όπως η πορεία των νέων κρουσμάτων, οι θάνατοι, οι αναρρώσεις, οι εμβολιασμοί και άλλα.
3. **Απεικόνιση Δεδομένων**: Όταν ο χρήστης επιλέγει ένα κουμπί, το πρόγραμμα εμφανίζει σε νέο παράθυρο το αντίστοιχο γράφημα ή πίνακα δεδομένων. Τα γραφήματα δημιουργούνται με τη χρήση της βιβλιοθήκης Matplotlib, η οποία επιτρέπει τη δημιουργία επαγγελματικών και καλαίσθητων διαγραμμάτων.
4. **Εύκολη Πλοήγηση**: Ο χρήστης μπορεί εύκολα να περιηγηθεί ανάμεσα στις διαφορετικές αναλύσεις χωρίς να χρειάζεται να κλείσει ή να επανεκκινήσει το πρόγραμμα. Αυτό καθιστά την εμπειρία χρήσης ομαλή και ευχάριστη. Επίσης έχει φτιαχτεί και ένα scrolldown bar για την διευκόλυνση της πλοήγησης μεταξύ των κουμπιών στο κεντρικό μενού
5. Παρακάτω μπορείτε να δείτε τα Screenshots του Menu:

A screenshot of a computer

Description automatically generated

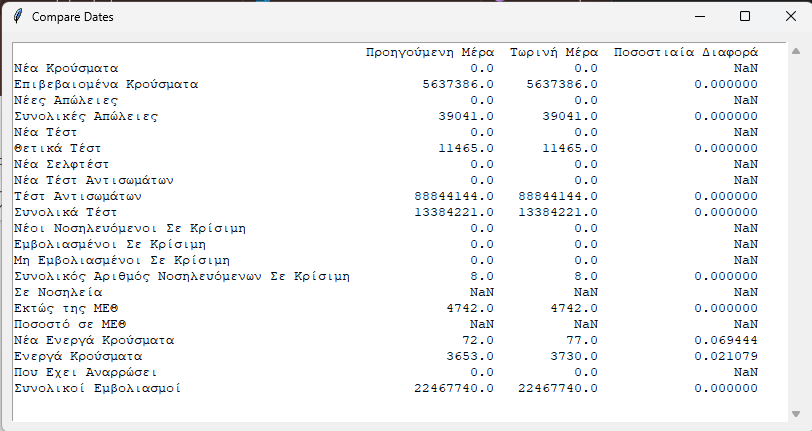
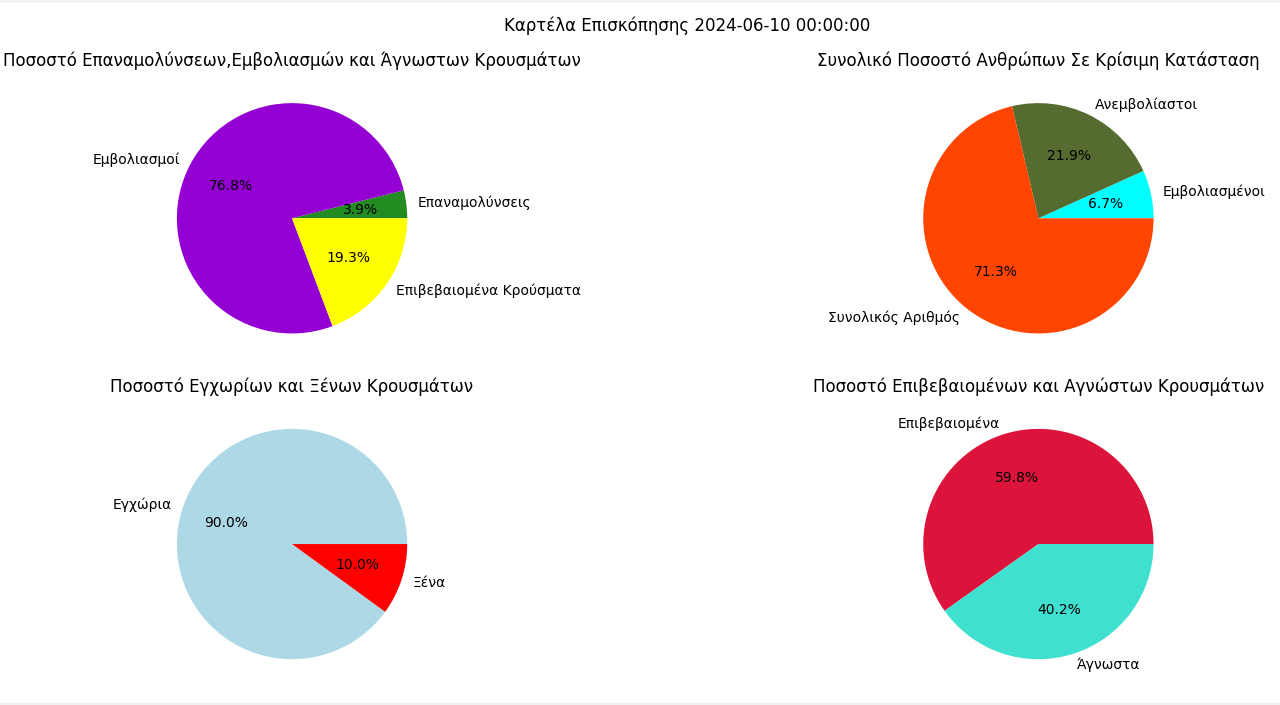
**Πλεονεκτήματα του Γραφικού Περιβάλλοντος**

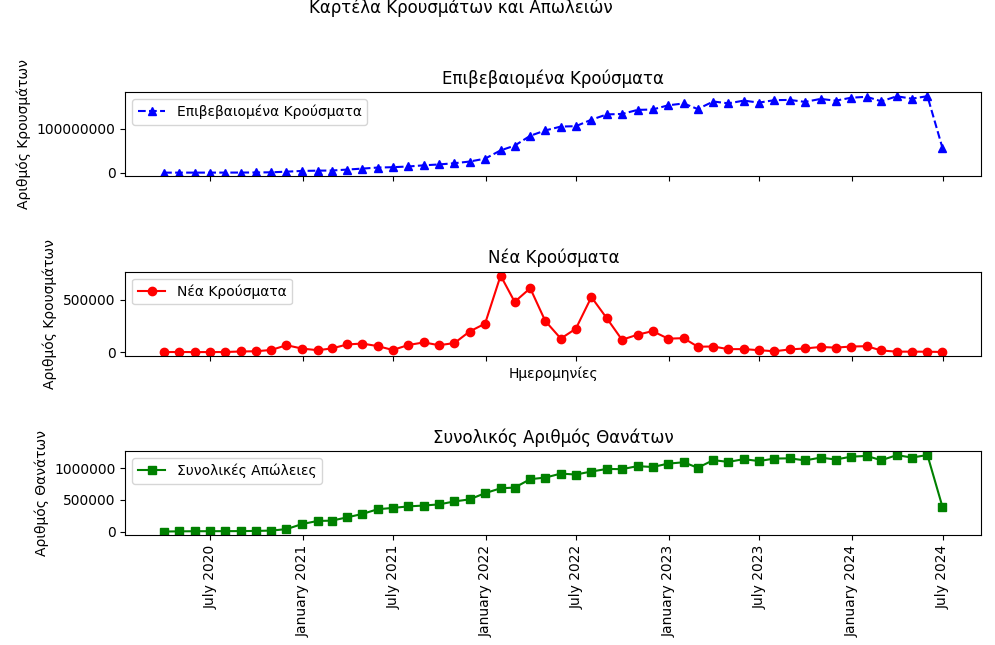
1. **Φιλικότητα προς τον Χρήστη**: Το γραφικό περιβάλλον είναι σχεδιασμένο με γνώμονα την ευκολία χρήσης, επιτρέποντας σε χρήστες με βασικές γνώσεις πληροφορικής να το χρησιμοποιήσουν αποτελεσματικά.
2. **Διαδραστικότητα**: Η δυνατότητα διαδραστικής οπτικοποίησης των δεδομένων καθιστά την ανάλυση πιο κατανοητή.
3. **Προσαρμοστικότητα**: Οι επιλογές φιλτραρίσματος και προσαρμογής των γραφημάτων επιτρέπουν την προσαρμογή του εργαλείου στις ανάγκες κάθε χρήστη.
4. **Αμεσότητα**: H άμεση οπτικοποίηση των μεταβολών παρέχουν έγκαιρη πληροφόρηση για την πορεία της πανδημίας.

Το γραφικό περιβάλλον του προγράμματος ανάλυσης δεδομένων Covid-19 στην Ελλάδα αποτελεί ένα ισχυρό εργαλείο για την κατανόηση και την παρακολούθηση της πανδημίας. Η διαδραστικότητα, η φιλικότητα προς τον χρήστη και η προσαρμοστικότητα του GUI το καθιστούν ιδανικό για την κατανόηση και μελέτη της πανδημίας.

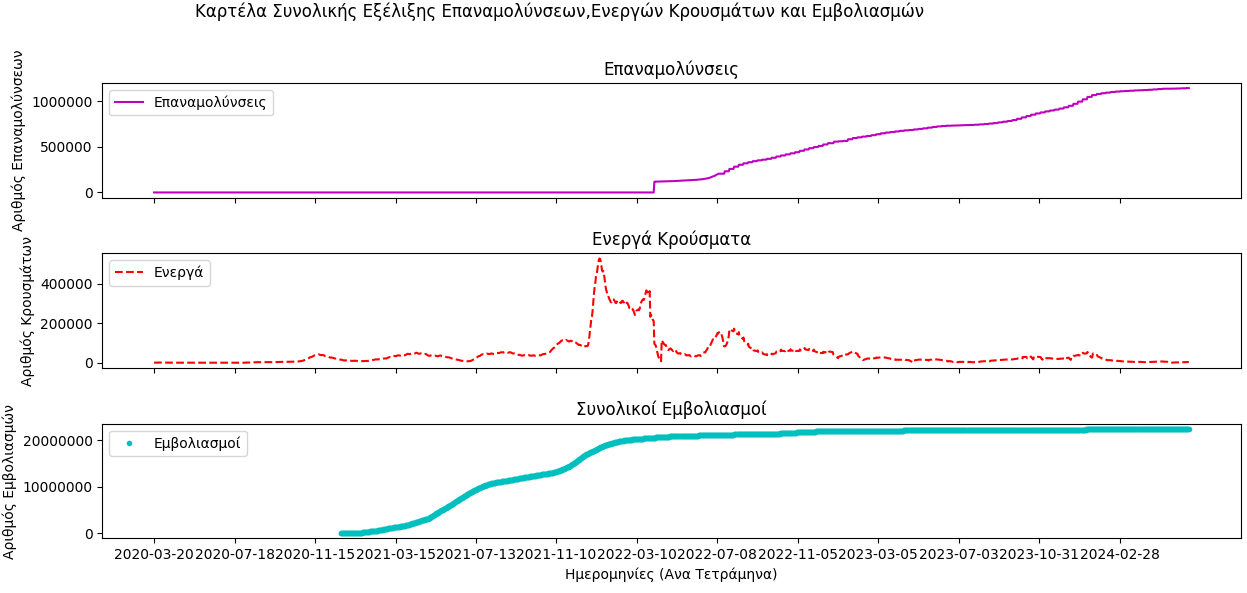
Δεύτερο Μέρος:Σκόπος του Δευτέρου Μέρους είναι η παρουσίαση της ημερήσιας επισκόπησης αλλά και συνολικής εξέλιξης βασικών στοιχείων της πανδημίας (όπως κρούσματα,απώλειες,εμβολιασμοί,νοσηλευόμενοι,επαναμολύνσεις)

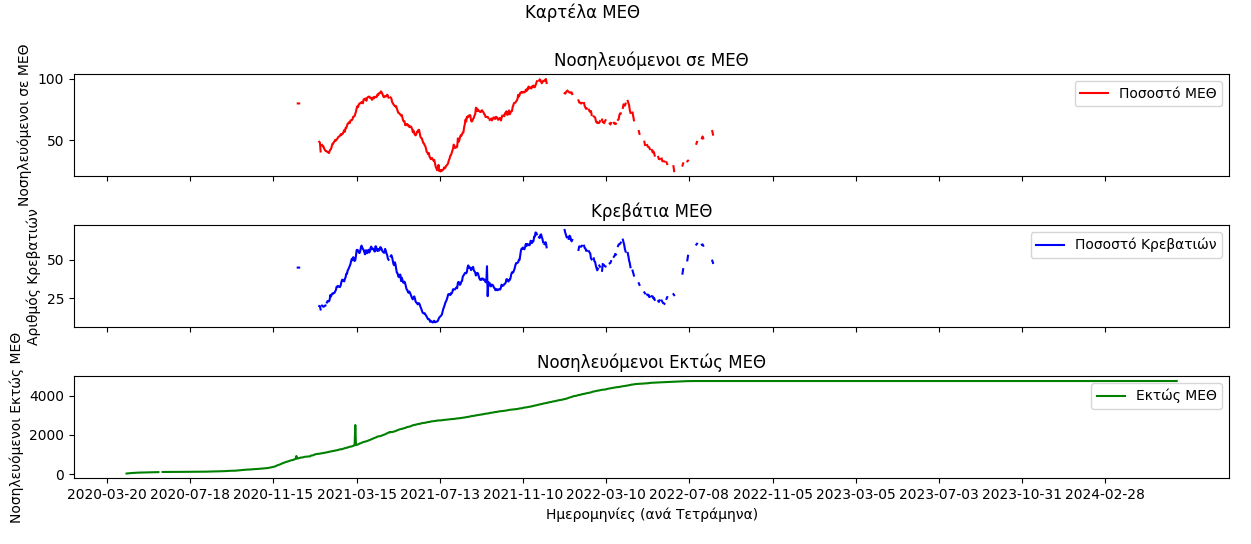
Περιλαμβάνει 6 συναρτήσεις,1 για την αναλυτική παρουσίαση των δεδομένων και 5 συναρτήσεις με παρουσιάσεις μέσω διαγραμμάτων.Ή πρώτη συνάρτηση ‘compare dates’ παρουσιάζει τα περισσότερα στοιχεία του dataframe (με τη διαγραφή ορισμένων για να μην γίνει κουραστική η ανάγνωσή τους) στις δύο τελευταίες μέρες ,όπως και τη ποσοστιαία διαφορά των δύο ημερών για την καλύτερη κατανόηση της ημερήσιας επισκόπησης.Πρώτα δημιουργήθηκε ένα dataframe μόνο των δύο τελευταίων ημερών,μετά ένα δεύτερο dataframe που απεικονίζει τη ποσοστιαία διαφορά και μετά ενώθηκαν και τα δύο σε έναν πίνακα.

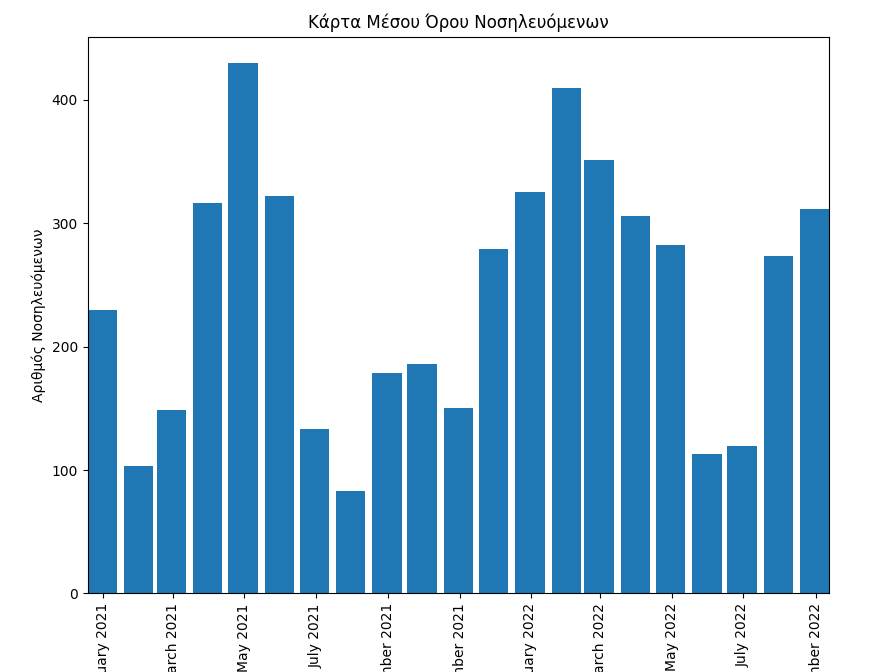
Εικόνα από την εκτέλεση της ‘compare dates: Η δεύτερη συνάρτηση ‘pie chart 1’ απεικονίζει 4 pies τα τρία απο τα οποία αφορούν επιδημιολογικά στοιχεία της τελευταία ημέρας και χρησιμοποιήθηκε η iloc για την αξιοποίηση των δεδομένων. Εικόνα απο το αποτέλεσμα:Η τρίτη συνάρτηση ‘cases and deaths’ απεικονίζει τη συνολική εξέλιξη των επιβεβαιομένων και νέων κρουσμάτων ,αλλά και των θανάτων.Αρχικά χρησιμοποιώ την pd to datetime για να αναγνωρίσει την ημερομηνία η βιβλιοθήκη και μετά με την resample ομαδοποιώ τα στοιχεία σε μήνες.Επίσης προστέθηκαν και markers για την καλύτερη κατανόηση των αλλαγών. Εικόνα Εκτέλεσης:



Η τέταρτη συνάρτηση ‘vaccinations and actives’ παρουσιάζει τη συνολική εξέλιξη των επαναμολύνσεων, και η πέμπτη συνάρτηση ‘ICU’ 3 μεγέθη που έχουν συσχέτιση με τις ΜΕΘ.Και στις δύο συναρτήσεις το δείγμα είναι ημερήσιο. Εικόνα Εκτέλεσης:

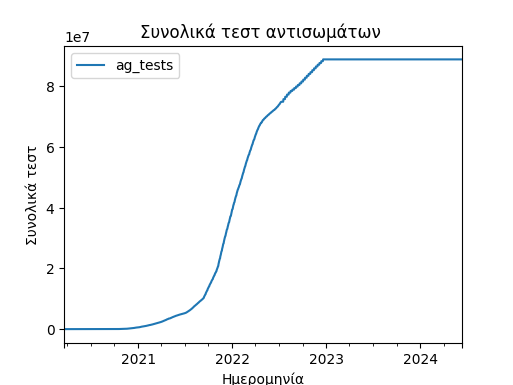
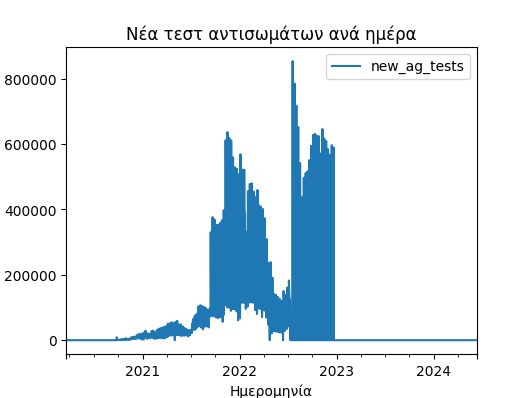


Η τελευταία συνάρτηση ‘hospitalized’ δείχνει τον μέσο όρο νοσηλευόμενων ανά μήνα.Η κατηγοριοποίηση του δείγματος γίνεται μέσω της resample και δημιουργείται ένα bar plot με τα αθροίσματα του δείγματος. Αποτέλεσμα Εκτέλεσης:

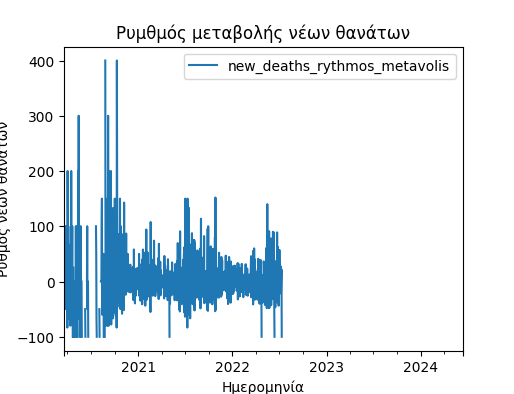
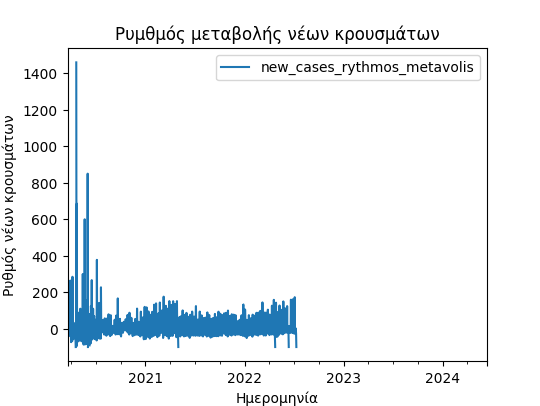


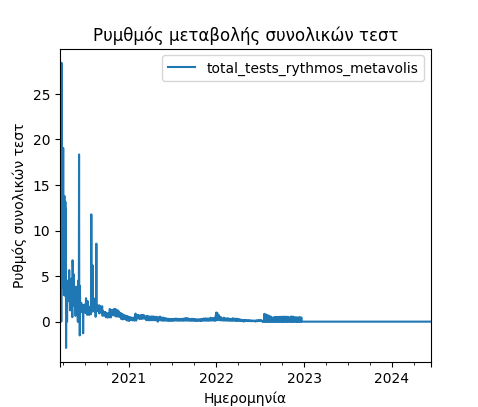
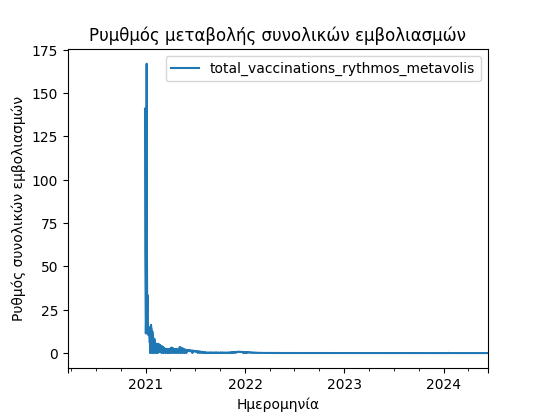
Τρίτο Μέρος:

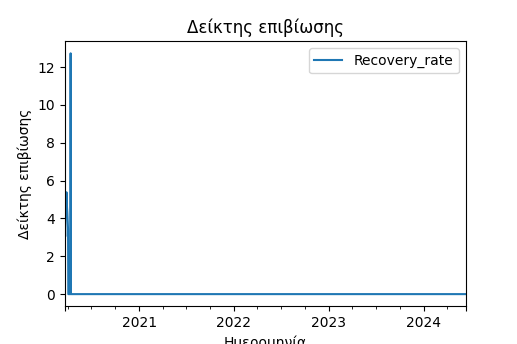
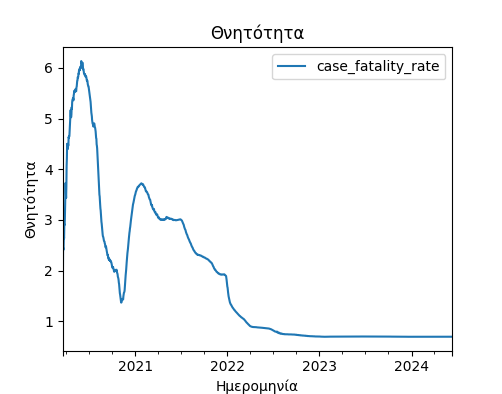
Στο τρίτο μέρος ακολουθεί η ημερήσια απεικόνιση των ελέγχων των αντισωμάτων καθώς και η συνολική εξέλιξη αυτών.



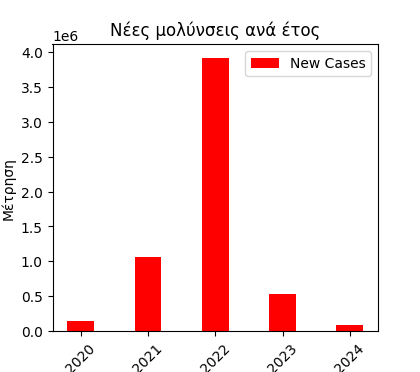
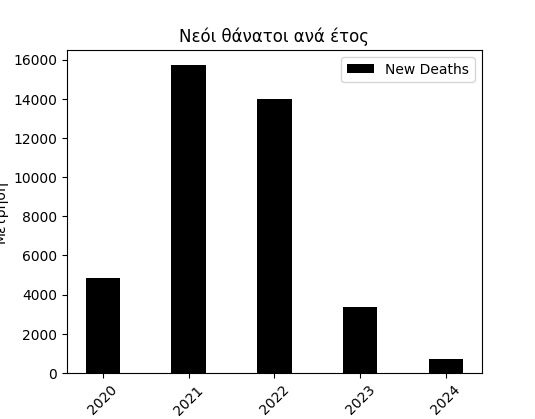
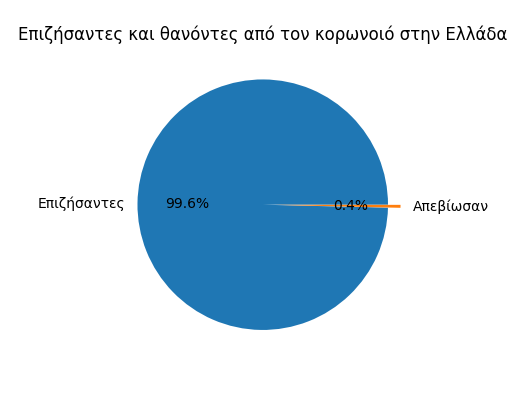
Στην συνέχεια ακολουθεί η απεικόνιση των παρακάτω ρυθμών μεταβολής:





Έπειτα παρουσιάζονται οι επιδημιολογικοί δείκτες (Case Fatality Rate, Recovery Rate):

Στο τέλος παρουσιάζονται ένα Pie Chart που δείχνει τοις συνολικές απώλειες σε ποσοστό των κατοίκων της Ελλάδας καθώς και οι νέοι θάνατοι και οι νέες μολύνσεις ανά έτος.



**ΤΕΤΑΡΤΟ ΜΕΡΟΣ**

Στο τέταρτο μέρος ο χρήστης μπορεί εναλλακτικά να πατήσει το πλαισιο με την εικόνα του plotly το οποίο χρησιμοποιεί ένα κουμπί από tkinter το οποίο είναι ενωμένο με τη συνάρτηση plotly\_btn. Έπειτα η συνάρτηση plotly button περιέχει πολλά κουμπιά από τα οποία το καθένα είναι δεμένο με τις συναρτήσεις που θα δούμε παρακάτω. Κάνοντας χρήση την command=lambda ενώνουμε το κουμπί ώστε όταν ο χρήστης πατήσει οποιοδηποτε κουμπι να ανοίγει την αντίστοιχη συνάρτηση.

A screenshot of a computer

Description automatically generated

Η πρώτη επιλογή (και συνάρτηση) που έχει ο χρήστης είναι ένα pie chart που αναδεικνύει σημαντικές πληροφορίες για τους συνολικούς εμβολιασμούς, τα άγνωστα κρούσματα και τις συνολικές επαναμολυνσεις. Ο χρήστης μπορεί επίσης να απομονώσει στο γράφημα κομμάτια από τα δεδομένα αν το επιθυμεί πατώντας πάνω δεξιά .

A pie chart with a number of percentages

Description automatically generated

**ΤΕΤΑΡΤΟ ΜΕΡΟΣ**

Η δεύτερη επιλογή που έχει ο χρήστης είναι μια συνάρτηση bar chart που παρουσιάζει τους ανθρώπους σε κρίσιμη κατάσταση (εμβολιασμένους και μη ) επίσης για το συγκεκριμένο γράφημα εχουμε και κάποια χρήσιμα εργαλεία όπως το κουμπί για ζουμ η το κουμπί για να κατεβάσεις το γράφημα στον υπολογιστή σου σαν png. Για την υλοποίηση χρειάστηκαν τα στοιχεία τα οποία αντλήθηκαν με την iloc και μετά με το plot express φτιάχνουμε ένα bar chart με την go.figure.A screenshot of a computer

Description automatically generated

A graph of a bar

Description automatically generated with medium confidence

**ΤΕΤΑΡΤΟ ΜΕΡΟΣ**

Η τρίτη επιλογή είναι ένα line chart με στοιχεία τα κρούσματα και τις απώλειες με τα δεδομένα από την πρώτη χρονική στιγμή του dataframe από το excel. Παρομοίως χρησιμοποιώντας το plotly express και την iloc έχουμε:A graph with a line

Description automatically generated

**ΤΕΤΑΡΤΟ ΜΕΡΟΣ**

Η τέταρτη επιλογή που έχει ο χρήστης είναι ένα γράφημα με linear regression όπου περιέχει ένα trendline με τα είδη υπάρχων δεδομένα (κρούσματα) και επίσης προβλέπει ένα χρόνο μπροστά. Πρώτα αντλούμε τα σωστά στοιχεία date και κρούσματα. Έπειτα διαχωρίζουμε το df για να το χρησιμοποιήσουμε με το panda datetime. Μετά τη μετατροπή διαχωρίζουμε με το .dt χρόνο μήνα και μέρα φτιάχνουμε το σωστό φορμάτ και δημιουργούμε ένα 2d array για το μοντέλο πρόβλεψης. Τέλος προσθέτουμε στο plotly graph τα στοιχεία.A graph with a line going up

Description automatically generated

**ΤΕΤΑΡΤΟ ΜΕΡΟΣ**

Στην τελευταία επιλογή ο χρήστης μπορεί επίσης να διαλέξει το κουμπί Mortality rate όπου είναι ένας δείκτης θνησιμότητας (αριθμός νεκρών με κορόνα ιό σε ποσοστό τις εκατό) το οποίο υλοποιήθηκε με μια iloc, μια πράξη και το plotly express.

A graph with a line going up

Description automatically generated