



ΔΙΕΘΝΕΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΗΣ ΕΛΛΑΔΟΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ
ΣΥΣΤΗΜΑΤΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΕΥΦΥΕΙΣ ΤΕΧΝΟΛΟΓΙΕΣ ΔΙΑΔΙΚΤΥΟΥ - WEBINTELLIGENCE

**Σύστημα παρακολούθησης και οπτικοποίησης δεδομένων
τιμών υγρών καυσίμων**

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Παπαδημητρίου Νικολάου του Δημοσθένη

Επιβλέπων : Ουγιάρογλου Στέφανος
Επίκουρος Καθηγητής, ΔΙ.ΠΑ.Ε.

Θεσσαλονίκη, Σεπτέμβριος 2023

Η σελίδα αυτή είναι σκόπιμα λευκή.



ΔΙΕΘΝΕΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΤΗΣ ΕΛΛΑΔΟΣ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ
ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΕΥΦΥΕΙΣ ΤΕΧΝΟΛΟΓΙΕΣ ΔΙΑΔΙΚΤΥΟΥ – WEB
INTELLIGENCE

Σύστημα παρακολούθησης και οπτικοποίησης δεδομένων τιμών υγρών καυσίμων

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

Παπαδημητρίου Νικολάου του Δημοσθένη

Επιβλέπων : Ουγιάρογλου Στέφανος
Επίκουρος Καθηγητής, ΔΙ.ΠΑ.Ε.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή στις Choose a date.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Όνομα Επώνυμο
Choose an item. ΔΙ.ΠΑ.Ε.

.....
Όνομα Επώνυμο
Choose an item. ΔΙ.ΠΑ.Ε.

.....
Όνομα Επώνυμο
Choose an item. ΔΙ.ΠΑ.Ε.

Θεσσαλονίκη, Σεπτέμβριος 2023

(Υπογραφή)

.....

Click here to enter text.

Click here to enter text.

© Choose a date– Allrightsreserved

«Στην σύζυγό μου Νικολέττα και τον γιο μου Δημοσθένη»

Παπαδημητρίου Νικόλαος

Περίληψη

Η παρούσα διπλωματική εργασία έχει ως σκοπό την οπτικοποίηση των δεδομένων των τιμών υγρών καυσίμων της Ελλάδος, χρησιμοποιώντας σύγχρονες τεχνολογίες και μεθόδους. Τα δεδομένα που χρησιμοποιεί η εργασία αναρτώνται από το υπουργείο Ανάπτυξης και Επενδύσεων, το οποίο καταγράφει καθημερινά τις τιμές των υγρών καυσίμων στη χώρα και τα δημοσιεύει μέσω ενός δελτίου επισκόπησης τιμών στην ιστοσελίδα του, <http://www.fuelprices.gr/>. Το δελτίο είναι διαθέσιμο ως αρχείο pdf διαφόρων μορφών, γεγονός που καθιστά την περαιτέρω χρήση και επεξεργασία του απαγορευτική. Τα δεδομένα αυτά περιλαμβάνουν τις καθημερινές παρατηρήσεις των τιμών υγρών καυσίμων ανά νομό της Ελλάδος και ανά τύπο καυσίμου (5 τύποι καυσίμου). Η παρούσα πτυχιακή εργασία προέκυψε από την ανάγκη για βαθύτερη κατανόηση των παραγόντων που διαμορφώνουν τις τιμές των υγρών καυσίμων. Ο σκοπός της εργασίας είναι διπλός: Μέσω τεχνικών scraping, η δημιουργία ενός ανοικτού συνόλου δεδομένων με τις τιμές υγρών καυσίμων και μέσω του λογισμικού Grafana οπτικοποίηση των εν λόγω δεδομένων. Με τη χρήση της ανάλυσης των δεδομένων, παρέχεται μια πιο άμεση και με μεγαλύτερη πληρότητα εικόνα του ενεργειακού τοπίου. Στο πλαίσιο της εργασίας, εξετάζονται τα εργαλεία και οι τεχνικές που απαιτούνται για την ανάλυση και την επεξεργασία των δεδομένων. Παρουσιάζονται συγκεκριμένες τεχνολογίες όπως η γλώσσα προγραμματισμού Python και οι συναφείς βιβλιοθήκες που χρησιμοποιήθηκαν, ενώ εξετάζεται και η χρήση εργαλείων όπως η μηχανή σχεσιακών βάσεων δεδομένων (RDBMS) Microsoft SQL Server για τη διαχείριση των δεδομένων. Η οπτικοποίηση των δεδομένων αποτελεί το κεντρικό σημείο της εργασίας αυτής. Με τη χρήση της πλατφόρμας Grafana και μέσω γραφημάτων και διαγραμμμάτων, παρουσιάζεται μια καλά δομημένη απεικόνιση των ενεργειακών τάσεων, αποτυπώνοντας βασικά πρότυπα, τόσο ανά νομό όσο και ανά χρονική διάρκεια. Τέλος, η εργασία συνοψίζεται με συμπεράσματα που αναδεικνύουν τη σημασία της ανάλυσης δεδομένων για την ενεργειακή πολιτική. Παράλληλα, προτείνονται μελλοντικές επεκτάσεις που μπορούν να ενισχύσουν την κατανόηση των ενεργειακών δεδομένων και να συμβάλουν στην ανάπτυξη βιώσιμων ενεργειακών πρακτικών.

Λέξεις Κλειδιά: Ανοιχτά δεδομένα, Τιμές υγρών καυσίμων, Οπτικοποίηση δεδομένων, Grafana

Η σελίδα αυτή είναι σκόπιμα λευκή.

« Information system for monitoring and visualization of liquid fuel price data »

Papadimitriou Nikolaos

Abstract

The present thesis aims to visualize the data of liquid fuel prices in Greece, using modern technologies and methods. The data used by the Ministry of Development and Investment, which is the source of the data used in the thesis, which records daily the prices of liquid fuels in the country and then publish them through a price overview bulletin on its website, <http://www.fuelprices.gr/>. The bulletin is available as a pdf file in various formats, which makes its further use and processing prohibitive. These data include the daily observations of liquid fuel prices per prefecture of Greece and per fuel type (5 fuel types). This thesis arose from the need for a deeper understanding of the factors that factors that shape the prices of liquid fuels. The aim of the thesis is twofold: With the use of scraping techniques, to create an open dataset with the data of liquid fuel prices and through the Grafana software the visualization of the data This data includes daily observations of liquid fuel prices by prefecture in Greece and by fuel type (5 fuel types). This thesis was born out of the need for a deeper understanding of the factors that liquid fuel prices. By using the analysis of data, it provides a more direct and more complete picture of the fuel prices. In the context of this work, the following examined are the tools and techniques required for the analysis and processing of the data. Specific technologies are presented such as the Python programming language and the associated libraries and the use of tools such as the relational database management system (RDBMS) Microsoft SQL Server for data management. The visualization of the data is the central point of this work. Using the Grafana platform and through graphs and diagrams, a well-structured and visualization of energy trends, capturing key patterns, both in terms of by county and by time period. Finally, the paper is summarized with conclusions that highlight the importance of data analysis for energy policy. At the same time, future directions are suggested, future extensions that can enhance the understanding of energy policy data and contribute to the development of sustainable energy policies and practices.

Keywords: Open data, Liquid fuel prices, Data visualization, Grafana

Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή της διπλωματικής εργασίας κ. Ουγιάρογλου Στέφανο για την καθοδήγησή του σε όλα τα στάδια της εκπόνησης αυτής της εργασίας. Θα ήθελα επίσης να δώσω πολλές ευχαριστίες στην οικογένειά μου, τη σύζυγο μου Νικολέττα και τον υιό μου Δημοσθένη, για την υποστήριξη που έλαβα σε όλη τη διάρκεια των μεταπτυχιακών μου σπουδών.

Πίνακας περιεχομένων

1	Εισαγωγή.....	1
1.1	Υγρά Καύσιμα.....	1
1.2	Η παγκόσμια αγορά υγρών καυσίμων.....	2
1.3	Η Ελληνική Αγορά Υγρών Καυσίμων.....	3
1.4	Ανοιχτά Δεδομένα (Τιμών Υγρών Καυσίμων).....	4
1.5	Κίνητρο.....	5
1.6	Συνεισφορά.....	7
1.7	Οργάνωση της Εργασίας.....	9
2	Τεχνολογίες που χρησιμοποιήθηκαν στην εργασία.....	10
2.1	Οι τεχνολογίες επιγραμματικά.....	11
2.2	Python.....	12
2.3	Βιβλιοθήκες τη Python.....	13
2.4	PyCharm.....	15
2.5	Microsoft SQL Server.....	16
2.6	MS SQL Server Management Studio.....	18
2.7	Brevo transactional email API.....	19
2.8	Grafana.....	20
3	Διαδικασία ανάκτησης και επεξεργασίας των δεδομένων.....	22
3.1	Στάδια ανάκτησης και επεξεργασίας δεδομένων.....	22
3.2	Ανάκτηση Αρχείων PDF.....	24
3.3	Μετατροπή Αρχείων PDF σε Αρχεία Κειμένου.....	25
3.4	Εισαγωγή Δεδομένων στη Βάση Δεδομένων.....	26
3.5	Ενημέρωση στον Διαχειριστή.....	28
3.6	Αυτοματοποίηση με τη χρήση Cron Jobs.....	29
3.7	Πηγαίος κώδικας της εφαρμογής.....	32
3.8	Σχήμα βάσης δεδομένων.....	38
3.9	Επικαιροποίηση δεδομένων.....	42
4	Οπτικοποίηση των δεδομένων.....	47
4.1	Εγκατάσταση του Grafana, Παραμετροποίηση και Σύνδεση με τη Βάση Δεδομένων.....	47
4.2	Δημιουργία SQL Views.....	49
4.3	Δημιουργία Dashboards.....	51
4.4	Παρουσίαση των δεδομένων υγρών καυσίμων.....	52

5	Επίλογος	59
5.1	Συμπεράσματα.....	59
5.2	Μελλοντικές Επεκτάσεις και Προοπτικές	60
6	Βιβλιογραφία.....	62

1

Εισαγωγή

1.1 Υγρά Καύσιμα

Τα υγρά καύσιμα αποτελούν ζωτικό συστατικό της σύγχρονης κοινωνίας, παρέχοντας την ενέργεια που απαιτείται για την κίνηση οχημάτων, τη θέρμανση κτιρίων και τη λειτουργία των βιομηχανιών. Από τη βενζίνη και το ντίζελ μέχρι το πετρέλαιο θέρμανσης και το πετρέλαιο κίνησης, τα υγρά καύσιμα συμβάλλουν σημαντικά στην οικονομική και κοινωνική λειτουργία της σύγχρονης κοινωνίας.

Τα υγρά καύσιμα αποτελούνται κυρίως από υδρογονάνθρακες, που προέρχονται από την αποδόμηση οργανικών υλικών, όπως πετρέλαιο, φυτικά έλαια και φυσικό αέριο. Η σύνθεση των υγρών καυσίμων εξαρτάται από την πηγή προέλευσης τους και τη διαδικασία επεξεργασίας που υφίστανται προκειμένου να ανταποκρίνονται στα πρότυπα ποιότητας και εκπομπών που ορίζονται από τη κεντρική διοίκηση.

Οι χρήσεις των υγρών καυσίμων είναι πολύπλευρες, από την καθημερινή κίνηση των αυτοκινήτων μέχρι την παραγωγή ενέργειας. Επιπλέον, εξαιτίας της ευρείας γκάμας καυσίμων που χρησιμοποιούνται, τα υγρά καύσιμα έχουν εφαρμογές σε διάφορους τομείς όπως η ναυτιλία, η αεροπορία, η γεωργία και η βιομηχανία.

Παρά τον κρίσιμη ρόλο που διαδραματίζουν, τα υγρά καύσιμα αντιμετωπίζουν προκλήσεις σχετικά με την εξάντληση των φυσικών πόρων, τις εκπομπές αερίων του θερμοκηπίου και την εξάρτηση από τις περιοχές εισαγωγής. Ωστόσο, οι τεχνολογικές εξελίξεις, όπως η ανάπτυξη εναλλακτικών καυσίμων και η βελτίωση της αποδοτικότητας καύσης, δημιουργούν ευκαιρίες για τη μείωση των περιβαλλοντικών επιπτώσεων και την ανακατανομή των ενεργειακών πηγών.

Τα υγρά καύσιμα διαδραματίζουν ουσιαστικό ρόλο στην καθημερινή ζωή και την οικονομία, αλλά και στην περιβαλλοντική βιωσιμότητα. Η αντιμετώπιση των προκλήσεων που συνδέονται με τη χρήση τους απαιτεί συνεχή έρευνα, καινοτομία και δράση προς την κατεύθυνση της ενεργειακής ανανέωσης και της βελτίωσης της αποδοτικότητας.

1.2 Η παγκόσμια αγορά υγρών καυσίμων

Η παγκόσμια αγορά υγρών καυσίμων διαμορφώνεται σημαντικά από πολιτικές και γεωπολιτικές παραμέτρους. Οι συγκρούσεις και η αστάθεια σε περιοχές παραγωγής, όπως η Μέση Ανατολή, μπορούν να έχουν δραστικές επιπτώσεις στην παραγωγή και, ως εκ τούτου, στις τιμές των καυσίμων. Ενδεχόμενες διακοπές στην παραγωγή πετρελαίου λόγω συγκρούσεων μπορούν να οδηγήσουν σε αύξηση των τιμών. Αντίθετα, η συνεργασία μεταξύ παραγωγικών χωρών και οι πολιτικές αποφάσεις που αφορούν την παραγωγή και εξαγωγή καυσίμων μπορούν να έχουν σταθεροποιητική επίδραση στις τιμές.

Η οικονομική κατάσταση παίζει επίσης κρίσιμο ρόλο στην παγκόσμια αγορά υγρών καυσίμων. Οικονομικές υφέσεις μπορούν να μειώσουν τη ζήτηση για καύσιμα και ως αποτέλεσμα να προκαλέσουν πτώση των τιμών. Αντίθετα, η οικονομική ανάκαμψη και ανάπτυξη συνήθως συνοδεύονται από αυξημένη κατανάλωση ενέργειας συνεπώς αύξηση της ζήτησης και των τιμών.

Η ισορροπία ανάμεσα στη προσφορά και στη ζήτηση αποτελεί κρίσιμο παράμετρο για τις τιμές των καυσίμων. Η αύξηση της παραγωγής πετρελαίου από μη παραδοσιακές πηγές μπορεί να επηρεάσει την προσφορά και να οδηγήσει σε αυξημένο ανταγωνισμό, με αποτέλεσμα χαμηλότερες τιμές. Αντίθετα, περιορισμοί στην παραγωγή ή εμπορικές κυρώσεις μπορεί να μειώσουν την προσφορά και να αυξήσουν τις τιμές.

Οι επενδύσεις στην ανάπτυξη υποδομών επηρεάζουν επίσης την διαμόρφωση των τιμών των υγρών καυσίμων. Η κατασκευή νέων αγωγών, τερματικών και εγκαταστάσεων αποθήκευσης μπορεί να βελτιώσει την αποδοτικότητα και την ασφάλεια της παραγωγής, μεταφοράς και

διανομής καυσίμων. Αυτό μπορεί να έχει ως αποτέλεσμα μειωμένα κόστη και μείωση των τιμών.

Συνολικά, η παγκόσμια αγορά υγρών καυσίμων είναι υπό τη συνεχή επίδραση πολιτικών, γεωπολιτικών, οικονομικών και υποδομικών παραγόντων, που συνδυάζονται για να διαμορφώνουν τις τιμές και την δυναμική της αγοράς των καυσίμων.

1.3 Η Ελληνική Αγορά Υγρών Καυσίμων

Η Ελληνική αγορά υγρών καυσίμων αντιπροσωπεύει ένα σημαντικό κομμάτι της ενεργειακής υποδομής της χώρας, αποτελώντας κρίσιμο στοιχείο για τη λειτουργία της κοινωνίας και την οικονομία. Καθημερινά, εκατομμύρια οχήματα, από ιδιωτικά αυτοκίνητα μέχρι φορτηγά, λεωφορεία και άλλα, εξαρτώνται από την παροχή υγρών καυσίμων για την κίνηση τους.

Ο τομέας των υγρών καυσίμων στην Ελλάδα χαρακτηρίζεται από πολυπλοκότητα και σημαντικές προκλήσεις. Παρότι η χώρα διαθέτει μια πλούσια γεωγραφική θέση και πόρους στον τομέα της ενέργειας, εξακολουθεί να εξαρτάται σε μεγάλο βαθμό από την εισαγωγή πετρελαίου και φυσικού αερίου για την κάλυψη των ενεργειακών της αναγκών. Οι διακυμάνσεις των τιμών των πετρελαϊκών προϊόντων στις παγκόσμιες αγορές επηρεάζουν άμεσα την τιμολόγηση των καυσίμων στην Ελλάδα, δημιουργώντας προκλήσεις για τους καταναλωτές και τους επιχειρηματίες.

Ενδιαφέρουσα διάσταση προσδίδεται από τον φορολογικό καθεστώς της χώρας, το οποίο επιδρά σε μεγάλο βαθμό στην τελική τιμή των καυσίμων. Ο ειδικός φόρος κατανάλωσης (ΕΦΚ) και ο φόρος προστιθέμενης αξίας (ΦΠΑ) αναδεικνύονται ως σημαντικοί παράγοντες που διαμορφώνουν την οικονομική επιβάρυνση των καταναλωτών.

Η αναβίωση της ελληνικής οικονομίας, σε συνέχεια της χρηματοπιστωτικής κρίσης της προηγούμενης δεκαετίας, διαδραματίζει επίσης καθοριστικό ρόλο στην αύξηση της ζήτησης για καύσιμα. Παρόλα αυτά, παρατηρείται ότι οι τιμές των καυσίμων παραμένουν επιμόνως σε υψηλά επίπεδα. Αυτή η κατάσταση φαίνεται να προκύπτει ως αποτέλεσμα της συνεχιζόμενης οικονομικής αβεβαιότητας και των περιορισμών που αντιμετωπίζει η Ελλάδα στη προσπάθειά της για ανάκαμψη.

Επιπλέον, οι καθοριστικοί παράγοντες που αναπτύχθηκαν προηγουμένως, όπως το φορολογικό πλαίσιο και οι πολιτικές επιλογές, συνυπάρχουν ως σημαντικοί παράγοντες που προσδίδουν περαιτέρω πολυπλοκότητα στην εικόνα της αγοράς καυσίμων στην Ελλάδα.

1.4 Ανοιχτά Δεδομένα (Τιμών Υγρών Καυσίμων)

Η ανάπτυξη της ψηφιακής εποχής και η εξέλιξη της τεχνολογίας έχουν επιφέρει αλλαγές σε πολλούς τομείς της κοινωνίας μας, συμπεριλαμβανομένου και του τομέα των υγρών καυσίμων. Η έννοια των "Ανοιχτών Δεδομένων" (Open Data) αποτελεί ένα από τα βασικά εργαλεία που επιτρέπουν την αξιοποίηση και την κοινή χρήση πληροφοριών με σκοπό τη βελτίωση της διαφάνειας, της συνεργασίας και της καινοτομίας.

Τα Ανοιχτά Δεδομένα αναφέρονται σε πληροφορίες που είναι ελεύθερα διαθέσιμες σε όλους, χωρίς περιορισμούς πνευματικής ιδιοκτησίας, πνευματικών δικαιωμάτων ή άλλων περιορισμών. Αυτή η ελευθερία πρόσβασης και χρήσης επιτρέπει σε οποιονδήποτε να αξιοποιήσει αυτά τα δεδομένα για τη δημιουργία νέων εφαρμογών, αναλύσεων και εξαγωγής συμπερασμάτων, προάγοντας την καινοτομία και την επίλυση κοινωνικών προβλημάτων.

Στο πλαίσιο του τομέα των υγρών καυσίμων, τα Ανοιχτά Δεδομένα αποτελούν μια πολύτιμη πηγή πληροφοριών για τις τιμές των καυσίμων. Οι τιμές των υγρών καυσίμων είναι υψηλά ευαίσθητες στις παγκόσμιες αγορές και υπόκεινται σε συνεχείς διακυμάνσεις. Η πρόσβαση σε αυτές τις πληροφορίες μπορεί να παίξει ένα καθοριστικό ρόλο στην λήψη αποφάσεων των καταναλωτών και των επιχειρήσεων σχετικά με την αγορά και τη χρήση των καυσίμων.

Τα Ανοιχτά Δεδομένα των τιμών των υγρών καυσίμων παρέχουν τη δυνατότητα στους ενδιαφερόμενους να αποκτήσουν μια πλήρη εικόνα των τιμών που ισχύουν σε διάφορες περιοχές και χρονικά διαστήματα. Με την ανάλυση αυτών των δεδομένων, είναι δυνατόν να εντοπιστούν τάσεις, πρότυπα και εξελίξεις που αφορούν την αγορά των καυσίμων.

Στο πλαίσιο της παρούσας εργασίας, τα Ανοιχτά Δεδομένα αξιοποιήθηκαν για τη δημιουργία αναλυτικών εργαλείων που επιτρέπουν την παρακολούθηση και την ανάλυση των τιμών των υγρών καυσίμων. Η συλλογή, επεξεργασία και οπτικοποίηση αυτών των δεδομένων παρέχει στους ενδιαφερόμενους μια πλήρη εικόνα του τρέχοντος καυσίμου της αγοράς και των εξελίξεων που τη διαμορφώνουν.

Το γεγονός ότι οι τιμές των καυσίμων είναι διαθέσιμες ως Ανοιχτά Δεδομένα ενισχύει τη διαφάνεια, ενισχύοντας την ενημέρωση των καταναλωτών και την δυνατότητα προσαρμογής στις αλλαγές της αγοράς. Παράλληλα, επιτρέπει την ανάπτυξη νέων εφαρμογών και υπηρεσιών που μπορούν να εξυπηρετήσουν τους καταναλωτές και τις επιχειρήσεις με πιο ορθολογικό τρόπο.

Συνοψίζοντας, τα Ανοιχτά Δεδομένα των τιμών υγρών καυσίμων αντιπροσωπεύουν ένα ισχυρό εργαλείο για τη βελτίωση της ενημέρωσης, της διαφάνειας και της ανταγωνιστικότητας στην αγορά. Η ενσωμάτωση τους σε αναλυτικά εργαλεία όπως αυτά που αναπτύχθηκαν στο πλαίσιο αυτής της εργασίας δίνει νέες δυνατότητες για την παρακολούθηση και την ανάλυση των δεδομένων, με στόχο τη βελτίωση της λήψης αποφάσεων και την επίτευξη βιώσιμης ανάπτυξης στον τομέα των υγρών καυσίμων.

1.5 Κίνητρο

Στο ευρύτερο πλαίσιο της αγοράς ενέργειας και των διαρκών ενεργειακών προκλήσεων, η ανάγκη για βιώσιμες, αποτελεσματικές και διαφανείς ενεργειακές πολιτικές είναι επιτακτική. Το κίνητρο που ωθεί την αξιοποίηση των ανοιχτών δεδομένων στον τομέα των υγρών καυσίμων απορρέει από την ανάγκη για βελτιωμένη κατανόηση της αγοράς αυτής και την ανάπτυξη δεξιοτήτων που θα οδηγήσουν σε βιώσιμες ενεργειακές επιλογές.

Το κύριο κίνητρο που διαμόρφωσε την κατεύθυνση της παρούσας εργασίας είναι η ανάγκη για εμβάθυνση και κατανόηση των δεδομένων τιμών των υγρών καυσίμων. Οι τιμές αυτές αποτελούν αναπόσπαστο μέρος της καθημερινότητας μας, επηρεάζοντας τον προϋπολογισμό των νοικοκυριών και των επιχειρήσεων, και έχουν σημαντικό αντίκτυπο στην οικονομία γενικότερα. Η εμβάθυνση στην ανάλυση αυτών των δεδομένων μπορεί να αποκαλύψει τάσεις, πρότυπα και συμπεράσματα που ενδεχομένως να προκαλούν αλλαγές στην κοινωνική και οικονομική μας δομή.

Μια σημαντική πτυχή του κινήτρου μας είναι η ανάγκη για εκμετάλλευση των Ανοιχτών Δεδομένων. Οι πληθώρα των διαθέσιμων δεδομένων από δημόσιες πηγές ανοίγει νέες προοπτικές για αναλύσεις που προηγουμένως ήταν δύσκολο να πραγματοποιηθούν. Η δυνατότητα πρόσβασης σε ένα πλούσιο σύνολο δεδομένων ενδυναμώνει την ικανότητά μας να εξάγουμε ενδιαφέρουσες πληροφορίες και να εξάγουμε συμπεράσματα από αυτές.

Στο πλαίσιο της τρέχουσας κατάστασης, όπου τα δεδομένα παρέχονται από το Υπουργείο σε μορφή PDF, η πληροφορία δεν είναι εύκολα διαχειρίσιμη και αξιοποιήσιμη, και γι' αυτό η αντιμετώπιση της πρόκλησης απαιτεί καινοτόμες προσεγγίσεις. Η μετατροπή των δεδομένων από μη δομημένη μορφή σε δομημένη πληροφορία αποτελεί βήμα προς τη σωστή κατεύθυνση.

ΘΕΜΑ: ΚΑΘΗΜΕΡΙΝΟ ΔΕΛΤΙΟ ΕΠΙΣΚΟΠΗΣΗΣ ΤΙΜΩΝ ΥΓΡΩΝ ΚΑΥΣΙΜΩΝ

Μέσες τιμές λιανικής πώλησης καυσίμων ανά Νομό, για την

Πέμπτη, 31 Αυγούστου 2023

(τιμές σε €/λτ, συμπ. ΦΠΑ)

ΝΟΜΟΣ	Αμόλυβδη 95 οκτ.	Αμόλυβδη 100 οκτ.	Diesel Κίνησης	Υγραέριο κίνησης (Autogas)
ΝΟΜΟΣ ΑΤΤΙΚΗΣ	1,956	2,154	1,735	0,795
ΝΟΜΟΣ ΑΙΤΩΛΙΑΣ ΚΑΙ ΑΚΑΡΝΑΝΙΑΣ	2,028	2,223	1,798	0,921
ΝΟΜΟΣ ΑΡΓΟΛΙΔΟΣ	1,998	2,173	1,761	0,876
ΝΟΜΟΣ ΑΡΚΑΔΙΑΣ	2,025	2,202	1,789	0,893
ΝΟΜΟΣ ΑΡΤΗΣ	2,003	2,184	1,778	0,909
ΝΟΜΟΣ ΑΧΑΪΑΣ	1,999	2,186	1,781	0,903
ΝΟΜΟΣ ΒΟΙΩΤΙΑΣ	1,987	2,188	1,763	0,909
ΝΟΜΟΣ ΓΡΕΒΕΝΩΝ	2,031	2,205	1,788	0,900
ΝΟΜΟΣ ΔΡΑΜΑΣ	1,985	2,156	1,750	0,827
ΝΟΜΟΣ ΔΩΔΕΚΑΝΗΣΟΥ	2,154	2,289	1,882	1,090
ΝΟΜΟΣ ΕΒΡΟΥ	2,001	2,167	1,769	0,908
ΝΟΜΟΣ ΕΥΒΟΙΑΣ	2,017	2,188	1,780	0,882
ΝΟΜΟΣ ΕΥΡΥΤΑΝΙΑΣ	2,069	2,275	1,838	0,951
ΝΟΜΟΣ ΖΑΚΥΝΘΟΥ	2,028	2,182	1,794	0,994
ΝΟΜΟΣ ΗΛΕΙΑΣ	2,007	2,182	1,769	0,845
ΝΟΜΟΣ ΗΜΑΘΙΑΣ	1,969	2,154	1,744	0,819
ΝΟΜΟΣ ΗΡΑΚΛΕΙΟΥ	2,046	2,188	1,815	0,950
ΝΟΜΟΣ ΘΕΣΠΡΩΤΙΑΣ	2,000	2,205	1,782	0,932
ΝΟΜΟΣ ΘΕΣΣΑΛΟΝΙΚΗΣ	1,957	2,147	1,738	0,831
ΝΟΜΟΣ ΙΩΑΝΝΙΝΩΝ	1,990	2,146	1,772	0,877
ΝΟΜΟΣ ΚΑΒΑΛΑΣ	2,003	2,175	1,768	0,883
ΝΟΜΟΣ ΚΑΡΔΙΤΣΗΣ	2,003	2,195	1,764	0,893

(Δείγμα αρχείου τιμών)

Επιπλέον, η επιλογή να χρησιμοποιήσουμε προηγμένα εργαλεία όπως το Grafana αποδεικνύει την προσήλωσή μας στην αποτελεσματική αξιοποίηση των δεδομένων. Μέσω τέτοιων εργαλείων, μπορούμε να ξεπεράσουμε τους περιορισμούς της μορφής PDF και να δημιουργήσουμε ένα περιβάλλον που επιτρέπει την οπτικοποίηση, την ανάλυση και την αντιληπτή αναπαράσταση των δεδομένων.

Η προσέγγισή μας δεν περιορίζεται μόνο στην επίλυση των τρεχουσών προκλήσεων, αλλά ανοίγει τον δρόμο για μελλοντικές επεκτάσεις και βελτιώσεις. Μέσω της βαθύτερης κατανόησης και αξιοποίησης των δεδομένων, μπορούμε να ανακαλύψουμε πρότυπα, νέες προοπτικές και πιθανότητες που θα συνεισφέρουν στη βελτίωση της αποτελεσματικότητας, της προοπτικής και της λήψης αποφάσεων στην ελληνική αγορά υγρών καυσίμων.

Συνολικά, η κίνητρο μας για τη μετατροπή των δεδομένων, τη χρήση προηγμένων εργαλείων και την καινοτόμο προσέγγιση αποτελεί την προσπάθειά μας να υπερβούμε τα όρια της τωρινής

κατάστασης και να δημιουργήσουμε ένα περιβάλλον που ενθαρρύνει την εξερεύνηση και την καινοτομία στον τομέα της ανάλυσης δεδομένων.

1.6 Συνεισφορά

Η παρούσα εργασία αναδεικνύει και επισημαίνει τη σημαντική συνεισφορά που εντάσσεται στο πλαίσιο της ενεργειακής ανάλυσης και οπτικοποίησης των δεδομένων των τιμών των υγρών καυσίμων. Συνδυάζοντας μεθόδους αποτελεσματικής εξαγωγής, μετασχηματισμού και αναπαράστασης πληροφοριών, η εργασία διευκολύνει την πρόσβαση των ενδιαφερομένων φορέων σε πολύτιμα δεδομένα.

Η κεντρική συνεισφορά της εργασίας εστιάζεται στην ολοκληρωμένη ανάλυση και αναπαράσταση των συλλεγμένων δεδομένων των τιμών των υγρών καυσίμων, ακολουθώντας μια επιστημονική μεθοδολογία. Κατ' αρχάς, πραγματοποιείται η συλλογή των δεδομένων από την επίσημη ιστοσελίδα του υπουργείου ανάπτυξης της Ελλάδος. Στη συνέχεια, προβαίνουμε στον μετασχηματισμό αυτών των δεδομένων σε ένα ευανάγνωστο και δομημένο σύνολο, διευκολύνοντας την εμπειριστατωμένη ανάλυση και ερμηνεία τους.

Βασικός πυλώνας της μελέτης αποτελεί η οπτικοποίηση των δεδομένων. Μέσω της χρήσης εξειδικευμένων εργαλείων οπτικοποίησης, όπως γραφήματα και διαγράμματα, αποκαλύπτουμε τις δομές και τα μοτίβα των τιμών των καυσίμων. Αυτά τα οπτικά εργαλεία παρέχουν ολοκληρωμένη εικόνα των ενεργειακών τάσεων και αναδεικνύουν πλήρως τις συνθήκες της αγοράς.

Η συνεισφορά αυτή έχει σημαντικές επιπτώσεις στο ευρύ φάσμα των ενδιαφερομένων. Ενισχύει τη διαφάνεια και την ενημέρωση των πολιτών, προσφέροντας εργαλεία για μια καλύτερη διαχείριση των καυσίμων. Η ανανεωμένη προσέγγιση παρέχει εύκολη πρόσβαση στις πληροφορίες, ενώ ταυτόχρονα εξασφαλίζει την πλήρη ενημέρωση του κοινού για τις ενεργειακές εξελίξεις.

Επιπροσθέτως, αυτή η προσέγγιση προάγει την επιστημονική έρευνα, ενθαρρύνοντας την ανάπτυξη νέων ερευνητικών προσεγγίσεων και την εξερεύνηση προηγμένων μεθοδολογιών ανάλυσης. Τα αποτελέσματα παρέχουν πολύτιμες γνώσεις για τη πολυπλοκότητα της αγοράς υγρών καυσίμων και τον αντίκτυπο της στη κοινωνία και το περιβάλλον.

Καταλήγοντας, η παρούσα εργασία αντιπροσωπεύει ένα σημαντικό ορόσημο προς την κατεύθυνση της προώθησης της ενεργειακής ευαισθησίας και της επιστημονικής προόδου. Η συνεισφορά της ανοίγει νέους ορίζοντες για την εξαγωγή γνώσης, την αξιοποίηση δεδομένων και την ανάπτυξη βιώσιμων λύσεων στο πεδίο της ενέργειας και των καυσίμων.

1.7 Οργάνωση της Εργασίας

Το υπόλοιπο της παρούσας διπλωματικής εργασίας είναι οργανωμένο ως εξής: στο δεύτερο κεφάλαιο, θα εξετάσουμε μια σειρά από τεχνολογίες που διαδραματίζουν ζωτικό ρόλο στην υλοποίηση της εργασίας μας. Θα εστιάσουμε σε βασικές τεχνολογίες όπως η γλώσσα προγραμματισμού Python, με την οποία δημιουργήθηκε ο κώδικας του συστήματός μας, και θα αναλύσουμε τις βιβλιοθήκες που χρησιμοποιούμε για να διευκολύνουμε τη διαχείριση και την ανάλυση των δεδομένων. Επίσης, θα εξετάσουμε εργαλεία όπως το Pycharm μέσω του οποίου ολοκληρώθηκε η ανάπτυξη του κώδικα, καθώς και βάσεις δεδομένων όπως η Microsoft SQL Server και το διαχειριστικό της περιβάλλον SQL Server Management Studio, τα οποία χρησιμοποιήθηκαν για την αποθήκευση και διαχείριση των πληροφοριών μας. Τέλος, θα εξετάσουμε το API της Brevo μέσω του οποίου είναι εφικτή η αποστολή των transactional emails προς τους διαχειριστές και το Grafana που είναι υπεύθυνο για την οπτικοποίηση των δεδομένων.

Στο τρίτο κεφάλαιο, θα εξετάσουμε τη διαδικασία ανάκτησης και επεξεργασίας των αρχικών δεδομένων που θα χρησιμοποιήσουμε στην εργασία μας. Θα περιγράψουμε τα στάδια που ακολουθούμε για την ανάκτηση αρχείων pdf, τη μετατροπή τους σε αρχεία κειμένου και την εισαγωγή των δεδομένων στη Βάση Δεδομένων μας. Θα εξετάσουμε επίσης τη χρήση εργαλείων όπως οι Cron jobs για την αυτοματοποίηση των διαδικασιών.

Στο τέταρτο κεφάλαιο, θα εξετάσουμε τη χρήση του Grafana για τη δημιουργία των οπτικών αναπαραστάσεων των δεδομένων μας. Θα αναλύσουμε τη διαδικασία εγκατάστασης και παραμετροποίησης του Grafana, καθώς και τον τρόπο σύνδεσής του με τη Βάση Δεδομένων μας. Θα παρουσιάσουμε τη δημιουργία SQL Views και Dashboards, ενώ θα εξηγήσουμε πώς το Grafana χρησιμοποιήθηκε για την παρουσίαση των ενεργειακών τάσεων και την ανάδειξη των σημαντικών πληροφοριών.

Στο πέμπτο και τελευταίο κεφάλαιο, θα παρουσιάσουμε τα συμπεράσματα που προκύπτουν από την εργασία μας και θα αξιολογήσουμε τη συνεισφορά μας στο πεδίο της ενεργειακής ανάλυσης και οπτικοποίησης. Θα εξετάσουμε τη συνολική πορεία της εργασίας και τον τρόπο με τον οποίο κάθε κομμάτι συνδέεται για την επίτευξη των στόχων μας. Τέλος, θα

διερευνήσουμε πιθανές μελλοντικές επεκτάσεις της εργασίας μας και πώς οι εξελίξεις στον τομέα της τεχνολογίας μπορούν να επηρεάσουν και να εμπλουτίσουν την περαιτέρω έρευνα σε αυτό το πεδίο.

2

Τεχνολογίες που χρησιμοποιήθηκαν στην εργασία

2.1 Οι τεχνολογίες επιγραμματικά

Στο παρόν κεφάλαιο, θα εξετάσουμε τις κυριότερες τεχνολογίες που επιλέχθηκαν για την ανάπτυξη του συστήματος μας. Καθεμία από αυτές τις τεχνολογίες επιλέχθηκε με προσοχή, λαμβάνοντας υπόψη τις ανάγκες και τους στόχους του έργου μας. Παρακάτω, παρουσιάζουμε και αναλύουμε κάθε μία από αυτές τις τεχνολογίες και τον τρόπο με τον οποίο συνεισέφερε στην υλοποίηση του συστήματος μας.

Python: Η γλώσσα προγραμματισμού Python επιλέχθηκε ως η κύρια γλώσσα ανάπτυξης λόγω της ευελιξίας και της ευκολίας στην ανάπτυξη. Η συντακτική της απλότητα και η μεγάλη ποικιλία βιβλιοθηκών την καθιστούν ιδανική για τη δημιουργία των εφαρμογών μας του συστήματός μας. Επιπλέον, η κοινότητα του Python παρέχει σταθερή υποστήριξη και συνεχείς βελτιώσεις.

Βιβλιοθήκες Python: Κατά τη διάρκεια της ανάπτυξης, αξιοποιήθηκαν πολλές βιβλιοθήκες της Python, όπως το Beautiful Soup για την ανάκτηση των δεδομένων από την σελίδα του Υπουργείου, και το pyodbcL για τη σύνδεση με τη βάση δεδομένων MS SQL Server. Αυτές οι βιβλιοθήκες ενίσχυσαν την παραγωγικότητα και επέτρεψαν την άμεση αντιμετώπιση της πιθανής πολυπλοκότητας.

PyCharm: Το περιβάλλον ανάπτυξης PyCharm επιλέχθηκε για την ανάπτυξη του κώδικα της Python, παρέχοντας προηγμένες δυνατότητες αντιστοίχισης κώδικα, αναλυτή σφαλμάτων και ολοκληρωμένο σύστημα ελέγχου εκδόσεων. Επίσης, κάτι που πρέπει να αναφερθεί είναι ότι κατά την υλοποίηση του συστήματός μας διαπιστώσαμε ότι το PyCharm είναι ένα εξαιρετικά σταθερό περιβάλλον ανάπτυξης εφαρμογών.

SQL Server: Το Microsoft SQL Server επιλέχθηκε ως το σύστημα διαχείρισης βάσεων δεδομένων (DBMS) λόγω της αξιοπιστίας, της απόδοσης και της ασφάλειάς του. Η ευέλικτη δομή του SQL Server επέτρεψε την αποτελεσματική εισαγωγή και ανάκτηση των πληροφοριών. Να σημειωθεί ότι για το παρόν έργο χρησιμοποιήθηκε η Developer Edition το Microsoft SQL Server, η οποία διανέμεται δωρεάν από την δημιουργό εταιρεία για σκοπούς ανάπτυξης που δεν περιλαμβάνουν εμπορική χρήση.

SQL Server Management Studio: Το SQL Server Management Studio (SSMS) είναι το εργαλείο που χρησιμοποιήθηκε για τη διαχείριση, τη συντήρηση και τον έλεγχο της βάσης δεδομένων SQL Server. Με το SSMS, μπορέσαμε να εκτελέσουμε ερωτήματα SQL, να διαχειριστούμε τη δομή της βάσης δεδομένων και να επιτυγχάνουμε αποτελεσματικά τον έλεγχο και τη διαχείριση των δεδομένων μας.

Brevo Transactional Email API: Η Brevo Transactional Email API χρησιμοποιήθηκε για την αποστολή αυτοματοποιημένων email ειδοποιήσεων προς τους χρήστες. Μέσω της συγκεκριμένης υπηρεσίας κατέστη εφικτή η αξιόπιστη και ασφαλής αποστολή μηνυμάτων ηλεκτρονικού ταχυδρομείου προς τους διαχειριστές του συστήματος. Να σημειωθεί ότι για τη παρόν έργο χρησιμοποιήθηκε η δωρεάν έκδοση της υπηρεσίας, η οποία κατά τη στιγμή της υλοποίησης, επιτρέπει την αποστολή 300 ηλεκτρονικών μηνυμάτων ανά ημέρα, άνευ χρέωσης.

Grafana: Το Grafana επιλέχθηκε για την οπτικοποίηση των δεδομένων και τη δημιουργία προσαρμοσμένων πινάκων ελέγχου. Μέσω του συγκεκριμένου εργαλείου κατέστη δυνατή η εύκολη δημιουργία γραφημάτων και πινάκων ελέγχου που βοήθησαν στην οπτικοποίηση των δεδομένων του συστήματός μας.

Καθένα από τα παραπάνω τεχνολογικά εργαλεία και πλατφόρμες συνέβαλαν στην αποτελεσματική ανάπτυξη, υλοποίηση και λειτουργία του συστήματος μας, διασφαλίζοντας όχι μόνο την υψηλή ποιότητα και αξιοπιστία του, αλλά και την επεκτασιμότητα και την ανταγωνιστικότητά του στο μελλοντικό περιβάλλον.

2.2 Python

Η Python είναι μία δυναμική, υψηλού επιπέδου γλώσσα προγραμματισμού που τα τελευταία χρόνια έχει κατακτήσει τον προγραμματιστικό κόσμο με την ευκολία χρήσης, την ευελιξία και τις πολλαπλές δυνατότητες που προσφέρει. Αναπτύχθηκε από τον Guido van Rossum και

κυκλοφόρησε για πρώτη φορά το 1991, και από τότε έχει κατακτήσει τη θέση της ως μία από τις πιο δημοφιλείς και ευρέως χρησιμοποιούμενες γλώσσες προγραμματισμού παγκοσμίως.

Η επιλογή να χρησιμοποιηθεί η Python στο πλαίσιο του έργου μας βασίστηκε σε πολλούς παράγοντες, που περιλαμβάνουν την ευανάγνωστη σύνταξή της, την εκτεταμένη βιβλιοθήκη που προσφέρει και την ευελιξία της για ευρύ φάσμα εφαρμογών. [2] Στο παρόν κεφάλαιο, θα εξετάσουμε λεπτομερώς τη γλώσσα προγραμματισμού Python, τη σύνταξή της, τα χαρακτηριστικά της και πώς εφαρμόστηκε στο έργο μας για την ανάπτυξη και τη λειτουργία του συστήματος μας.

Η Python ξεχωρίζει για την απλότητά της, καθιστώντας την κατανοητή και προσιτή. Η σύνταξή της είναι καθαρή και ευανάγνωστη, με έμφαση στον ευανάγνωστο κώδικα, καθιστώντας έτσι την ανάπτυξη πιο αποτελεσματική και συντηρήσιμη. [3]

Ένα βασικό χαρακτηριστικό της Python είναι η δυνατότητα της για δυναμική και αυτόματη διαχείριση μνήμης, η οποία απαλλάσσει τον προγραμματιστή από την ανάγκη να χειρίζεται λεπτομερώς τη μνήμη του συστήματος. Επιπλέον, η Python διαθέτει μία μεγάλη βιβλιοθήκη που παρέχει προκατασκευασμένους κώδικες για ποικίλες εργασίες, όπως το χειρισμό συμβολοσειρών, την επεξεργασία αρχείων, τη διαχείριση εργασιών σύνδεσης στο διαδίκτυο, τη βάση δεδομένων, την μηχανική μάθηση και πολλά άλλα. [1]

Η Python υποστηρίζει πολλά παραδείγματα προγραμματισμού, συμπεριλαμβανομένου του προγραμματισμού αντικειμένων και του διαδικαστικού προγραμματισμού. Αυτό το ευέλικτο μοντέλο προγραμματισμού επιτρέπει στους προγραμματιστές να επιλέξουν την προσέγγιση που ταιριάζει καλύτερα στις απαιτήσεις τους και στη φύση του έργου.

Στο έργο μας, επιλέξαμε να χρησιμοποιήσουμε την Python λόγω της ευκολίας και της ευελιξίας της για την ανάπτυξη του συστήματός μας. Η Python μας επέτρεψε να αναπτύξουμε γρήγορα και αποτελεσματικά τον κώδικα, μειώνοντας το χρόνο ανάπτυξης και επιτρέποντάς μας να επικεντρωθούμε στην περαιτέρω βελτίωση της λειτουργίας και της απόδοσης του συστήματος.

Η χρήση των βιβλιοθηκών της Python μας επέτρεψε να επιταχύνουμε την ανάπτυξη, αφού μπορούσαμε να αξιοποιήσουμε προκατασκευασμένους κώδικες για βασικές λειτουργίες. [5]

Για παράδειγμα, χρησιμοποιήσαμε τις βιβλιοθήκες Requests και Beautiful Soup για την αλληλεπίδραση με την ιστοσελίδα του υπουργείου και την εξαγωγή πληροφοριών.

Επιπλέον, για την αποστολή ενημερώσεων στους διαχειριστές μέσω email, χρησιμοποιήσαμε το Brevo Transactional Email API, το οποίο ενσωματώνεται εύκολα με την Python και μας επέτρεψε να διαχειριστούμε την αποστολή email με αξιοπιστία και ασφάλεια.

Η επιλογή να χρησιμοποιήσουμε την Python στο έργο μας αποδείχθηκε εξαιρετικά επιτυχημένη. Η ευκολία χρήσης, η εκτεταμένη βιβλιοθήκη της και η ευελιξία της μας επέτρεψαν να αναπτύξουμε ένα αξιόπιστο και αποτελεσματικό σύστημα που ανταποκρίνεται στις απαιτήσεις μας. Η Python συνέβαλε σημαντικά στην επίτευξη των στόχων μας και στην ανάπτυξη ενός υψηλής ποιότητας λογισμικού.

2.3 Βιβλιοθήκες τη Python

Η ανάπτυξη λογισμικού είναι μια πολύπλοκη διαδικασία που απαιτεί τη χρήση διάφορων εργαλείων και τεχνικών για να διευκολύνει τον προγραμματιστή και να βελτιώσει την απόδοση του λογισμικού. Μια από τις πιο ισχυρές επιλογές που προσφέρει η γλώσσα προγραμματισμού Python είναι η δυνατότητα χρήσης βιβλιοθηκών. Οι βιβλιοθήκες είναι σύνολα προκατασκευασμένων λειτουργιών και εργαλείων που βοηθούν τους προγραμματιστές να επιτελέσουν συγκεκριμένες εργασίες με αποδοτικό και αξιόπιστο τρόπο.

Στο παρόν κεφάλαιο, θα εξετάσουμε πέντε σημαντικές βιβλιοθήκες της Python που χρησιμοποιήθηκαν στο έργο μας για την ανάπτυξη και τη λειτουργία του συστήματός μας. Κάθε βιβλιοθήκη έχει τη δική της χρησιμότητα και προσφέρει μοναδικά πλεονεκτήματα που συνέβαλαν στην ομαλή εκτέλεση του έργου μας.

BeautifulSoup

Η βιβλιοθήκη BeautifulSoup αποτελεί ένα από τα επικρατέστερα εργαλεία για εξαγωγή περιεχομένου από τον ιστό. Με τη χρήση της BeautifulSoup, είναι δυνατή η εξαγωγή πληροφορίας από δομές HTML και XML όπως είναι τα δεδομένα που είναι αποθηκευμένα σε ιστοσελίδες. [4]

Στο πλαίσιο του έργου μας, χρησιμοποιήσαμε την BeautifulSoup για την εξαγωγή δεδομένων από την σελίδα του υπουργείου (<http://www.fuelprices.gr/>). Με τη χρήση της συγκεκριμένης βιβλιοθήκης κατέστη δυνατή η αυτοματοποίηση της διαδικασίας εξαγωγής δεδομένων, εξοικονομώντας χρόνο και προσπάθεια και διευκολύνοντας τη διαχείριση του περιεχομένου.

Pyodbc

Η βιβλιοθήκη Pyodbc αποτελεί ένα εξαιρετικά χρήσιμο εργαλείο για τη σύνδεση και την αλληλεπίδραση με βάσεις δεδομένων Microsoft SQL Server. Είναι μία βιβλιοθήκη που παρέχει δυνατότητες σύνδεσης, ανάκτησης, ενημέρωσης και διαγραφής δεδομένων από τη βάση δεδομένων.

Στο πλαίσιο του έργου μας, χρησιμοποιήσαμε τη βιβλιοθήκη pyodbc για να δημιουργήσουμε τη σύνδεση της εφαρμογής με τη βάση δεδομένων Microsoft SQL Server που χρησιμοποιήσαμε. Με τη χρήση της βιβλιοθήκης κατέστη δυνατή η εκτέλεση SQL ερωτημάτων, η ανάκτηση δεδομένων και η ενημέρωση της βάσης μας με ασφάλεια και αξιοπιστία.

Requests

Η βιβλιοθήκη Requests αποτελεί ένα απαραίτητο εργαλείο για την αλληλεπίδραση με ιστοσελίδες και την αποστολή HTTP αιτημάτων. Απλή στη χρήση της και ισχυρή, η βιβλιοθήκη Requests μας παρέχει τη δυνατότητα να λαμβάνουμε δεδομένα από ιστοσελίδες, να αναλύουμε αποκρίσεις και να επικοινωνούμε με ανοιχτές πηγές πληροφοριών.

Στο πλαίσιο της εφαρμογής μας, χρησιμοποιήσαμε τη Requests για να συλλέξουμε τα δεδομένα από την σελίδα του υπουργείου (<http://www.fuelprices.gr/>). Με τη χρήση της συγκεκριμένης βιβλιοθήκης κατέστη δυνατή η μεταφορά τοπικά της σελίδας του υπουργείου, σε μορφή HTML, ώστε να υποστεί περαιτέρω επεξεργασία.

Fitz

Η βιβλιοθήκη Fitz, γνωστή και ως PyMuPDF, αποτελεί μία ισχυρή βιβλιοθήκη για την ανάγνωση και την επεξεργασία αρχείων PDF. Παρέχει δυνατότητες ανάγνωσης κειμένου, εικόνων και διανυσμάτων από αρχεία PDF, καθώς και δυνατότητες επεξεργασίας και αντιγραφής περιεχομένου.

Στο πλαίσιο του έργου, η βιβλιοθήκη Fitz χρησιμοποιήθηκε για την ανάγνωση και την εξαγωγή δεδομένων από αρχεία PDF. Με τη βοήθεια της Fitz, κατέστη δυνατή η εξαγωγή της πληροφορίας από τα έγγραφα PDF, γεγονός που διευκόλυνε την επεξεργασία και την ανάλυση των δεδομένων.

Shutil

Η βιβλιοθήκη Shutil αποτελεί ένα εργαλείο που παρέχει λειτουργίες για τη διαχείριση αρχείων και φακέλων στο σύστημα αρχείων. Επιτρέπει την αντιγραφή, τη μετακίνηση, τη μετονομασία και τη διαγραφή αρχείων με απλό και ευανάγνωστο τρόπο.

Στο πλαίσιο της εφαρμογής μας, χρησιμοποιήσαμε τη βιβλιοθήκη Shutil για τη διαχείριση και τη μετακίνηση αρχείων και φακέλων. Με αυτήν τη βιβλιοθήκη, μπορέσαμε να οργανώσουμε και να διαχειριστούμε τα δεδομένα μας με αποδοτικό και αξιόπιστο τρόπο.

Συνοψίζοντας, οι παραπάνω βιβλιοθήκες αποτέλεσαν εξαιρετικά εργαλεία για την ανάπτυξη και τη λειτουργία του συστήματός μας. Κάθε μία εξ αυτών παρείχε μοναδικές δυνατότητες και λειτουργίες που συνέβαλαν στην ομαλή λειτουργία του συστήματος, στην αξιοπιστία των διαδικασιών και στην αποτελεσματική διαχείριση των δεδομένων.

2.4 PyCharm

Το PyCharm είναι ένα εξαιρετικά πανίσχυρο και ευέλικτο περιβάλλον ανάπτυξης εφαρμογών που στοχεύει στη της Python για δημιουργία εξαιρετικής ποιότητας λογισμικού. Στο πλαίσιο της υλοποίησης του έργου μας, το PyCharm υπηρέτησε ως το κεντρικό εργαλείο ανάπτυξης, παρέχοντας απαραίτητες δυνατότητες και λειτουργίες για την ομαλή και αποτελεσματική διαδικασία κατασκευής του λογισμικού μας.

Η χρήση του PyCharm στο έργο μας αποτέλεσε έναν κρίσιμο παράγοντα για την επιτυχή υλοποίηση και ανάπτυξη του λογισμικού μας. Κατά τη διάρκεια της διαδικασίας ανάπτυξης, αξιοποιήσαμε τις εξής βασικές λειτουργίες του PyCharm:

Αρχικά, χρησιμοποιήσαμε το PyCharm για τη δημιουργία νέων έργων Python. Η δυνατότητα επιλογής του κατάλληλου περιβάλλοντος εργασίας μας επέτρεψε να προσαρμόσουμε τις ρυθμίσεις σύμφωνα με τις απαιτήσεις του έργου μας.

Ο μηχανισμός αυτόματης συμπλήρωσης κώδικα του PyCharm αποδείχθηκε ιδιαίτερος χρήσιμος κατά την επεξεργασία του κώδικα. Η λειτουργία αυτή βοήθησε στην αποφυγή τυπογραφικών λαθών και στην επιτάχυνση της διαδικασίας πληκτρολόγησης, ενώ παρείχε επίσης προτάσεις για κατάλληλες συναρτήσεις και μεθόδους.

Ο ενσωματωμένος διαχειριστής εξαρτήσεων του PyCharm διευκόλυνε τη διαχείριση των απαραίτητων βιβλιοθηκών. Μέσω του διαχειριστή, κατέστη δυνατή η εγκατάσταση,

ενημέρωση και διαχείριση των εξαρτήσεων που απαιτούνται για τη σωστή λειτουργία του προγράμματος.

Ο ενσωματωμένος αποσφαλματωτής του PyCharm αποτέλεσε ζωτικό εργαλείο κατά τη διάρκεια της ανάπτυξης. Η δυνατότητα ανίχνευσης, ανάλυσης και επίλυσης σφαλμάτων στον κώδικα συνέβαλε στη διασφάλιση της σταθερότητας και της αξιοπιστίας του λογισμικού μας.

Μέσω των ενσωματωμένων εργαλείων οπτικής ανάλυσης κώδικα, είχαμε τη δυνατότητα να εξερευνήσουμε τη δομή του προγράμματος μας με ευκολία. Η απεικόνιση των συσχετίσεων μεταξύ κλάσεων και συναρτήσεων βοήθησε στην κατανόηση του κώδικα και την εντοπισμό πιθανών βελτιώσεων.

Το PyCharm παρέχει ένα ολοκληρωμένο περιβάλλον ανάπτυξης, συμπεριλαμβανομένων εργαλείων διαχείρισης εκδόσεων, ενσωματωμένων συστημάτων ελέγχου κώδικα και δυνατοτήτων δοκιμής και αποσφαλμάτωσης. Αυτό επέτρεψε την αποτελεσματική διαχείριση του κύκλου ζωής του λογισμικού μας.

Συνοψίζοντας, το PyCharm αποδείχθηκε αναντικατάστατο εργαλείο κατά τη διάρκεια της ανάπτυξης του λογισμικού μας. Η ισχυρή του λειτουργικότητα, σε συνδυασμό με τις προηγμένες δυνατότητες ανάλυσης και αποσφαλμάτωσης, διευκόλυναν την ομάδα ανάπτυξης και συνέβαλαν στην ομαλή εξέλιξη του έργου μας. Η ενσωμάτωση του PyCharm στην εργασιακή μας ροή ενίσχυσε την αποτελεσματικότητα, την ποιότητα και την ταχύτητα της ανάπτυξης, καθιστώντας το ένα κρίσιμο εργαλείο για την επιτυχή υλοποίηση του έργου.

2.5 Microsoft SQL Server

Ο Microsoft SQL Server είναι μια ισχυρή, επιχειρησιακή βάση δεδομένων που παρέχει αξιόπιστη αποθήκευση, διαχείριση και επεξεργασία δεδομένων. [7] Στο πλαίσιο του έργου μας, ο Microsoft SQL Server χρησιμοποιήθηκε ως η κεντρική βάση δεδομένων για την αποθήκευση και ανάκτηση των πληροφοριών που αφορούν τα καύσιμα και τις συναλλαγές. Στο παρόν έργο χρησιμοποιήθηκε η Developer Edition το Microsoft SQL Server, η οποία διανέμεται δωρεάν από την Microsoft για σκοπούς ανάπτυξης που δεν περιλαμβάνουν εμπορική χρήση.

Ο Microsoft SQL Server επιτρέπει τη δημιουργία και τη διαχείριση σύνθετων βάσεων δεδομένων, παρέχοντας προηγμένες δυνατότητες για την οργάνωση και την αποθήκευση των πληροφοριών. [8] Εκτός από την κλασική διαχείριση δεδομένων, ο Microsoft SQL Server προσφέρει και δυνατότητες όπως:

Προηγμένες λύσεις ασφαλείας για την προστασία των δεδομένων. Ορισμένες από αυτές περιλαμβάνουν την κρυπτογράφηση δεδομένων σε αναπαραστάσεις όταν τα δεδομένα βρίσκονται σε κατάσταση αδράνειας και τη δυνατότητα περιορισμού της πρόσβασης σε συγκεκριμένα δεδομένα μόνο από εξουσιοδοτημένους χρήστες.

Ενσωματωμένες υπηρεσίες ανάλυσης δεδομένων, όπως το SQL Server Analysis Services (SSAS). Αυτό επιτρέπει στους χρήστες να δημιουργούν πολυδιάστατες προβολές, αναλύσεις δεδομένων αλλά και δομές δεδομένων τύπου OLAP που αποσκοπούν στη βέλτιστη λήψη αποφάσεων βασισμένων στα δεδομένα. [6]

Εργαλεία διαχείρισης επιδόσεων που βοηθούν στον εντοπισμό και την επίλυση προβλημάτων απόδοσης. Οι διαχειριστές μπορούν να παρακολουθούν τις επιδόσεις της βάσης και να προβαίνουν σε βελτιστοποιήσεις προκειμένου να διασφαλίσουν την απρόσκοπτη λειτουργία του συστήματος.

Η επιλογή του Microsoft SQL Server ως βάσης δεδομένων για το έργο μας αποδείχθηκε ουσιώδης. Οι παρακάτω τομείς αναδεικνύουν τον τρόπο με τον οποίο χρησιμοποιήσαμε τον Microsoft SQL Server:

Ο Microsoft SQL Server χρησιμοποιήθηκε για τη δημιουργία μιας πλήρους και αξιόπιστης βάσης δεδομένων που αποθηκεύει πληροφορίες σχετικά με τα καύσιμα και τις συναλλαγές. Η δομή της βάσης σχεδιάστηκε με προσοχή, επιτρέποντας την αποθήκευση, αναζήτηση και ανάκτηση δεδομένων με αποδοτικό τρόπο.

Χρησιμοποιώντας τις δυνατότητες του Microsoft SQL Server, αναπτύξαμε διαδικασίες επεξεργασίας και επικύρωσης δεδομένων. Αυτό εξασφάλισε τη συνέπεια και την ακρίβεια των πληροφοριών που καταγράφονται στη βάση μας, συμβάλλοντας στη βελτίωση της ποιότητας των δεδομένων.

Τέλος, μέσω του Microsoft SQL Server, δημιουργήθηκαν εξειδικευμένες αναλύσεις και αναφορές βασισμένες στα δεδομένα μας. Στη συνέχεια κάποιες από αυτές τις αναφορές χρησιμοποιήθηκαν από το Grafana για την οπτικοποίηση των δεδομένων.

Συμπερασματικά, ο Microsoft SQL Server αποδείχθηκε ένα ισχυρό εργαλείο στην ανάπτυξη του έργου μας. Η υψηλή απόδοση, η αξιοπιστία και οι προηγμένες δυνατότητες διαχείρισης

και ανάλυσης δεδομένων του Microsoft SQL Server ενίσχυσαν την αποτελεσματικότητα και την ποιότητα του έργου μας. Η ενσωμάτωση του Microsoft SQL Server στην υποδομή μας βελτίωσε τη διαχείριση των δεδομένων μας, επιτρέποντάς μας να αντιμετωπίσουμε αποτελεσματικά τις απαιτήσεις του έργου μας.

2.6 MS SQL Server Management Studio

Το Microsoft SQL Server Management Studio (SSMS) αποτελεί ένα εξαιρετικά ισχυρό εργαλείο διαχείρισης και ανάπτυξης για τον Microsoft SQL Server. Αυτό το εργαλείο παρέχει μια ευρεία γκάμα δυνατοτήτων που επιτρέπουν στους διαχειριστές βάσεων δεδομένων και τους προγραμματιστές να αλληλεπιδρούν με τις βάσεις δεδομένων με αποτελεσματικότητα και ευκολία. Το SSMS παρέχει μια σειρά σημαντικών λειτουργιών που βοηθούν στη διαχείριση και ανάπτυξη των βάσεων δεδομένων.

Μέσω του SSMS, οι διαχειριστές μπορούν να συνδεθούν με τις βάσεις δεδομένων του Microsoft SQL Server, να διαχειριστούν συνδέσεις και να εκτελούν ερωτήματα SQL. Η δυνατότητα εκτέλεσης ερωτημάτων επιτρέπει την ανάκτηση, τον έλεγχο και την επεξεργασία δεδομένων.

Οι χρήστες μπορούν να δημιουργούν νέες βάσεις δεδομένων, πίνακες, SQL procedures, συναρτήσεις και άλλα αντικείμενα της βάσης δεδομένων. Μπορούν επίσης να πραγματοποιούν τροποποιήσεις σε υπάρχοντα αντικείμενα ώστε να ανταποκρίνονται στις μεταβαλλόμενες απαιτήσεις της εφαρμογής.

Το SSMS επιτρέπει την εισαγωγή και εξαγωγή δεδομένων από και προς τις βάσεις δεδομένων. Αυτό είναι ιδιαίτερα χρήσιμο κατά τη μεταφορά δεδομένων από ένα περιβάλλον ανάπτυξης σε ένα περιβάλλον παραγωγής ή κατά τη μεταφορά δεδομένων μεταξύ διαφορετικών βάσεων δεδομένων.

Το SSMS αποτέλεσε κρίσιμο εργαλείο κατά τη διάρκεια της ανάπτυξης και της διαχείρισης της βάσης δεδομένων του έργου μας. Οι ακόλουθοι τομείς αναδεικνύουν τον τρόπο με τον οποίο χρησιμοποιήσαμε το SSMS:

Χρησιμοποιήσαμε το SSMS για να δημιουργήσουμε, να τροποποιήσουμε και να διαχειριστούμε τη δομή της βάσης δεδομένων μας. Μέσω του εργαλείου, δημιουργήσαμε πίνακες, προσαρμόσαμε τα πεδία τους και δημιουργήσαμε σχέσεις μεταξύ τους.

Επίσης χρησιμοποιήσαμε το εργαλείο για να εκτελέσουμε πολύπλοκα ερωτήματα SQL προκειμένου να ανακτήσουμε συγκεκριμένα δεδομένα από τη βάση μας. Επίσης, δημιουργήσαμε προσαρμοσμένες αναφορές που παρείχαν αναλυτική πληροφορία για την κατάσταση των καυσίμων και των συναλλαγών.

Το SSMS μας επέτρεψε να προβαίνουμε σε αλλαγές στη δομή της βάσης δεδομένων όποτε απαιτούνταν λόγω αλλαγών στις απαιτήσεις του έργου. Επίσης, μέσω του εργαλείου, εντοπίσαμε και επιλύσαμε προβλήματα απόδοσης που αφορούσαν τη βάση δεδομένων.

Συνοψίζοντας, το Microsoft SQL Server Management Studio αποδείχθηκε κρίσιμο για την αποτελεσματική διαχείριση και ανάπτυξη της βάσης δεδομένων του έργου μας. Η ικανότητα άμεσης και απλής διεπαφής με τις βάσεις δεδομένων, η δυνατότητα εκτέλεσης ερωτημάτων SQL, η δημιουργία αναφορών και γραφημάτων, καθώς και η ευελιξία στη διαχείριση αλλαγών, συνέβαλαν στην ομαλή εξέλιξη του έργου. Μέσω του SSMS, επιτεύχθηκε η αποτελεσματική αξιοποίηση και ανάλυση των δεδομένων, ενισχύοντας την ευρύτερη επιτυχία του έργου μας.

2.7 Brevo transactional email API

Το Brevo Application Protocol Interface (API) αποτελεί ένα εξαιρετικά παραμετροποιήσιμο εργαλείο που επιτρέπει τη διαχείριση, την ανάλυση και την εκτέλεση επικοινωνιακών καμπανιών με μεγάλη ευκολία. Σχεδιασμένο για να καλύπτει τις ανάγκες της σύγχρονης επιχειρηματικής επικοινωνίας, το API προσφέρει ένα φάσμα λειτουργιών που καλύπτουν την πλήρη διαδικασία αποστολής email, από τη δημιουργία προτύπων μέχρι την αναλυτική παρακολούθηση των αποτελεσμάτων.

Στο πλαίσιο του έργου μας, η χρήση του API αποτελεί τον πυρήνα της αποστολής email στους διαχειριστές του συστήματος. Ο μηχανισμός αυτοματοποιημένης αποστολής που υλοποιήθηκε επέτρεψε την άμεση ειδοποίηση των διαχειριστών για οποιαδήποτε κρίσιμη εξέλιξη ή πιθανό σφάλμα στο έργο. [13] Αυτό προσέδωσε ευελιξία και ακρίβεια στην αποστολή ειδοποιήσεων, καθιστώντας το έργο αξιόπιστο και δραστικό.

Η αυτοματοποίηση αποδείχθηκε κρίσιμη για την επιτυχή διαχείριση του έργου. Κατά τη διάρκεια της υλοποίησης, η αποστολή email στους διαχειριστές κατέστη απαραίτητη για να αντιμετωπιστούν άμεσα ενδεχόμενα προβλήματα και να διασφαλιστεί η συνεχής πρόοδος του έργου. Η επικοινωνία μεταξύ της ομάδας διαχείρισης ήταν άμεση και αποτελεσματική, επιτρέποντας την έγκαιρη λήψη αποφάσεων και την αποτελεσματική αντιμετώπιση προβλημάτων.

Η χρήση του Brevo API στο έργο μας αποτελεί ένα παράδειγμα της τεχνολογίας που μετατρέπει την επικοινωνία και τη διαχείριση σε μια ολοκληρωμένη και αποτελεσματική διαδικασία. Η ικανότητα να ενημερώνουμε τους διαχειριστές με αυτοματοποιημένες ειδοποιήσεις ενισχύει τη δυναμική του έργου και εξασφαλίζει την άμεση ανταπόκριση σε κάθε πρόκληση. Εν κατακλείδι, η ενσωμάτωση του Brevo API ενίσχυσε την επιτυχή πορεία του έργου μας προς την επιτυχή υλοποίηση του.

2.8 Grafana

Το Grafana αποτελεί ένα ισχυρό εργαλείο ανάλυσης και οπτικοποίησης δεδομένων [12], το οποίο χρησιμοποιήθηκε, πλαίσιο του έργου μας, για την οπτικοποίηση των τιμών των υγρών καυσίμων.

Η οπτικοποίηση των τιμών των υγρών καυσίμων αποτελεί σημαντική πτυχή της διπλωματικής εργασίας, καθώς μας επιτρέπει να παρακολουθούμε και να αναλύουμε τις μετρήσεις που σχετίζονται με τα καύσιμα. Το Grafana παρέχει τη δυνατότητα δημιουργίας προσαρμοσμένων πινάκων ελέγχου (dashboards), όπου μέσω αυτών είναι δυνατή η παρακολούθηση των δεδομένων από διάφορες πηγές, συμπεριλαμβανομένων των βάσεων δεδομένων μας.

Η οπτικοποίηση των τιμών των υγρών καυσίμων μέσω του Grafana επιτρέπει τη κατανόηση των τάσεων και των μοτίβων που σχετίζονται με την κατανάλωση καυσίμων. Μέσω γραφημάτων και διαγραμμάτων, είναι δυνατή η ανάλυση των διακυμάνσεων των τιμών των καυσίμων και ο εντοπισμός ενδεχόμενων προβλημάτων ή ανωμαλιών.

Ένα από τα ισχυρά χαρακτηριστικά του Grafana είναι η δυνατότητα δημιουργίας προσαρμοσμένων πινάκων ελέγχου. Μπορούμε να επιλέξουμε ποια δεδομένα θέλουμε να παρακολουθούμε και να τα οπτικοποιήσουμε με τον τρόπο που μας βολεύει περισσότερο. Μέσω της δημιουργίας γραφημάτων, διαγραμμάτων και πινάκων ελέγχου, μπορούμε να έχουμε μια συνολική εικόνα των τιμών των υγρών καυσίμων και να εξάγουμε σημαντικές πληροφορίες.

Η εφαρμογή του Grafana στο έργο μας μας έχει προσφέρει μια ισχυρή εργαλειοθήκη για την οπτικοποίηση και ανάλυση των τιμών των υγρών καυσίμων. Μέσω της δημιουργίας προσαρμοσμένων πινάκων ελέγχου και της επιλογής των κατάλληλων γραφημάτων και διαγραμμάτων, μπορούμε να παρακολουθούμε την απόδοση των υγρών καυσίμων και να λαμβάνουμε έγκαιρες ειδοποιήσεις για τυχόν προβλήματα. Χάρη στη χρήση του Grafana η

οπτικοποίηση των δεδομένων κατέστη δυνατή με τρόπο άμεσο, αξιόπιστο και παραμετροποιήσιμο.

Συνοψίζοντας, η χρήση του Grafana αποδείχθηκε αναπόσπαστο μέρος της επιτυχημένης υλοποίησης του έργου μας, παρέχοντας μια αξιόπιστη και ολοκληρωμένη λύση για την οπτικοποίηση και ανάλυση των δεδομένων των τιμών των υγρών καυσίμων.

3

Διαδικασία ανάκτησης και επεξεργασίας των δεδομένων

3.1 Στάδια ανάκτησης και επεξεργασίας δεδομένων

Στο παρόν κεφάλαιο, θα επικεντρωθούμε στα στάδια της ανάκτησης και επεξεργασίας των δεδομένων που απαιτούνται για την περαιτέρω ανάλυση και οπτικοποίηση των τιμών των υγρών καυσίμων. Η διαδικασία αυτή αποτελεί βασικό βήμα για τη δημιουργία μιας αξιόπιστης και ενημερωμένης βάσης δεδομένων που θα υποστηρίξει την ανάλυση και την οπτικοποίηση των δεδομένων.

Ανάκτηση Αρχείων PDF: Στο πρώτο στάδιο, πραγματοποιούμε τη συλλογή των δεδομένων από την ιστοσελίδα του υπουργείου. Τα δεδομένα είναι σε μορφή αρχείων PDF, ένα για κάθε ημέρα. Σε αυτό το στάδιο γίνεται ανάγνωση της σελίδας, εύρεση των αρχείων PDF και μεταφορά τους τοπικά. Η διαδικασία εισάγει μια εγγραφή στη βάση δεδομένων για κάθε αρχείο που κατεβάζει ώστε να μην τα ξανακατεβάσει, δηλαδή η διαδικασία λειτουργεί διαφορεικά.

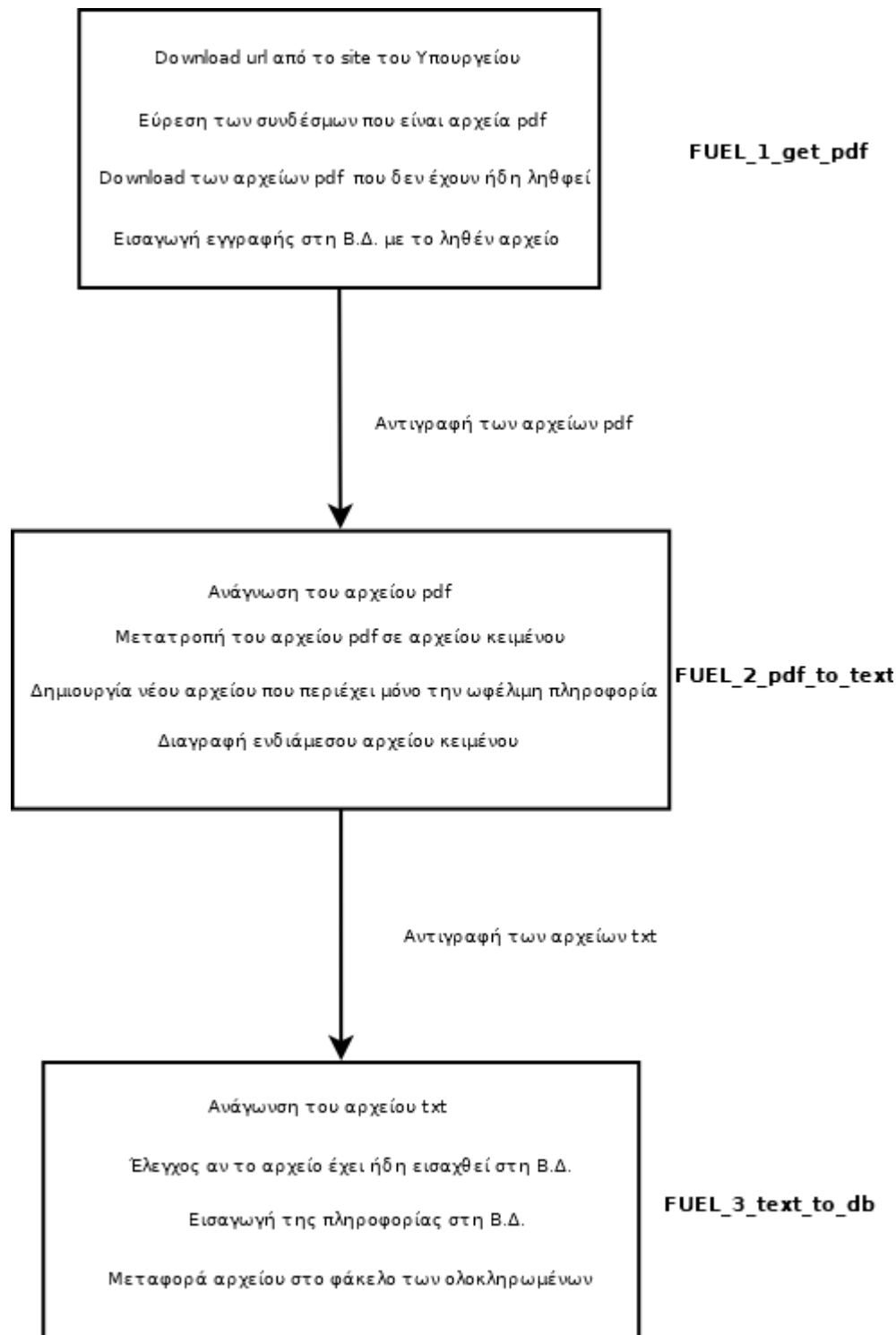
Μετατροπή Αρχείων PDF σε Αρχεία Κειμένου: Σε αυτό το στάδιο, προχωρούμε στη μετατροπή των δεδομένων από τη μορφή PDF σε μορφή κειμένου, ώστε στη συνέχεια να εισαχθούν στη βάση δεδομένων, ως δομημένη πλέον πληροφορία. Σε αυτό το στάδιο, τα δεδομένα υπόκεινται σε διάφορες διαδικασίες επεξεργασίας. Αυτές περιλαμβάνουν την αφαίρεση περιττών πληροφοριών, την εξαγωγή των απαιτούμενων δεδομένων και τη διόρθωση ενδεχόμενων σφαλμάτων.

Εισαγωγή Δεδομένων στη Βάση Δεδομένων: Μετά την επιτυχημένη μετατροπή των δεδομένων σε αρχεία κειμένου συγκεκριμένης γραμμογράφησης, είμαστε έτοιμοι να εισάγουμε

τα δεδομένα στη Βάση Δεδομένων. Τα δεδομένα αυτά θα χρησιμοποιηθούν στα επόμενα στάδια για την ανάλυση και οπτικοποίηση.

Ενημέρωση στον Διαχειριστή: Σε αυτό το στάδιο γίνεται η αυτοματοποιημένη αποστολή ενημερώσεων στους διαχειριστές του συστήματος. Οι ενημερώσεις αυτές έχουν ως σκοπό να ενημερώσουν τους διαχειριστές για πιθανά λάθη επί τις διαδικασίας ώστε να προβούν στις απαραίτητες ενέργειες διόρθωσης και αποσφαλμάτωσης.

Συνοψίζοντας, αυτό το κεφάλαιο παρουσιάζει τα στάδια της ανάκτησης και επεξεργασίας δεδομένων που επιτρέπουν τη δημιουργία μιας ενημερωμένης και αξιόπιστης βάσης δεδομένων για την ενεργειακή ανάλυση των τιμών των υγρών καυσίμων. Αυτός ο ενδιάμεσος σταθμός είναι κρίσιμος για την εξασφάλιση υψηλής ποιότητας και ακρίβειας στα δεδομένα που θα χρησιμοποιηθούν στα επόμενα στάδια της εργασίας μας.



(διάγραμμα ροής διαδικασιών της εφαρμογής)

3.2 Ανάκτηση Αρχείων PDF

Στο πλαίσιο της εργασίας μας, η ανάκτηση των αρχείων PDF αποκτά ζωτική σημασία καθώς αποτελεί το πρώτο βήμα προς την αξιοποίηση και ανάλυση των δεδομένων που περιέχονται σε

αυτά. Στο παρόν κεφάλαιο, θα εξετάσουμε αναλυτικά τη διαδικασία ανάκτησης αρχείων PDF από την ιστοσελίδα του υπουργείου ανάπτυξης (<http://www.fuelprices.gr/>), εστιάζοντας στις τεχνολογίες που χρησιμοποιούμε για την επίτευξη αυτού του στόχου. Πρόκειται για μια διαδικασία που περιλαμβάνει την ανάγνωση της ιστοσελίδας, την εντοπισμό των αρχείων PDF και την αποθήκευσή τους σε τοπικό διακομιστή, προετοιμάζοντάς τα για περαιτέρω επεξεργασία και ανάλυση.

Η διαδικασία ανάκτησης αρχείων PDF ενσωματώνει πολλές τεχνολογίες και βιβλιοθήκες που συνεργάζονται για την επίτευξη του επιθυμητού αποτελέσματος. Στο πλαίσιο αυτό, η βιβλιοθήκη Beautiful Soup της Python αποδεικνύεται ζωτικής σημασίας. Χρησιμοποιώντας τη Beautiful Soup, πραγματοποιούμε την ανάλυση του HTML κώδικα της ιστοσελίδας, εντοπίζοντας τα αρχεία PDF. Αυτή η επεξεργασία είναι καίριας σημασίας για τον ακριβή εντοπισμό των αρχείων που μας ενδιαφέρουν, καθώς η ιστοσελίδα περιέχει και αρκετή πληροφορία η οποία δεν μας είναι χρήσιμη στο πλαίσιο της εργασίας.

Κατά τη διαδικασία ανάκτησης, επιδιώκεται η αποφυγή επαναληπτικής λήψης αρχείων που έχουν ήδη ανακτηθεί. Αυτό επιτυγχάνεται μέσω της αλληλεπίδρασης με μια βάση δεδομένων SQL. Με τη χρήση της pyodbc βιβλιοθήκης, πραγματοποιούμε τη σύνδεση με τη βάση δεδομένων και προσθέτουμε μία εγγραφή για κάθε ανακτημένο αρχείο PDF. Αυτή η διαδικασία εξασφαλίζει τη διατήρηση της συνολικής ακεραιότητας των δεδομένων και αποτρέπει την περαιτέρω ανάκτηση των αρχείων που έχουν ήδη αποθηκευτεί.

Η παρούσα διαδικασία ανάκτησης αρχείων PDF αποτελεί κρίσιμο κομμάτι της εργασίας μας, καθώς εξασφαλίζει την πρώτη επαφή με τα δεδομένα που θα αναλυθούν στη συνέχεια. Η χρήση εξειδικευμένων βιβλιοθηκών και τεχνολογιών εξασφαλίζει την ακρίβεια, την αποτελεσματικότητα και την ασφάλεια της διαδικασίας, ενισχύοντας έτσι την συνολική ποιότητα της εργασίας μας. Με αυτόν τον τρόπο, δημιουργούμε μια βάση αξιόπιστων και υψηλής ποιότητας δεδομένων που θα αποτελέσουν τη βάση για την επόμενη φάση της εργασίας μας, την ανάλυση και την απόδοση σημασίας στα συγκεντρωμένα δεδομένα.

3.3 Μετατροπή Αρχείων PDF σε Αρχεία Κειμένου

Η μετατροπή αρχείων PDF σε αρχεία κειμένου αποτελεί βασικό βήμα για την αξιοποίηση των πληροφοριών που περιέχονται στα PDF αρχεία. Τα PDF αρχεία αποτελούν ένα δημοφιλές μέσο αποθήκευσης δεδομένων, αλλά η ανάγνωση και ανάλυσή τους με προγραμματιστικό τρόπο είναι πολύ πιο πολύπλοκη σε σχέση με αρχεία κειμένου. Η χρήση της Python βιβλιοθήκης Fitz

μας επιτρέπει την ανάκτηση πληροφοριών από τα PDF αρχεία και τη μετατροπή τους σε μορφή κειμένου για περαιτέρω επεξεργασία και ανάλυση.

Η Fitz βιβλιοθήκη προσφέρει ένα πλούσιο σετ εργαλείων για την ανάκτηση πληροφοριών από τα PDF αρχεία. Χρησιμοποιώντας τη Fitz, μπορούμε να εξάγουμε κείμενο από διάφορα στοιχεία του PDF, όπως κείμενα, τίτλους, παραγράφους και εικόνες.

Η διαδικασία της ανάκτησης κειμένου από PDF περιλαμβάνει τα εξής βήματα:

- Φόρτωση του PDF αρχείου με τη χρήση της Fitz βιβλιοθήκης.
- Ανάκτηση των διαφόρων στοιχείων του PDF, όπως κείμενα και εικόνες.
- Αποθήκευση του ανακτηθέντος κειμένου σε αρχεία κειμένου για περαιτέρω επεξεργασία.
- Μετατροπή σε Αρχεία Κειμένου

Αφού ανακτήσουμε το κείμενο από τα PDF αρχεία, το επόμενο βήμα είναι η μετατροπή του σε αρχεία κειμένου. Η Pythοn παρέχει ενσωματωμένες λειτουργίες για την εγγραφή κειμένου σε αρχεία, και με τη χρήση αυτών των λειτουργιών μπορούμε εύκολα να αποθηκεύσουμε το ανακτηθέν κείμενο σε αρχεία κειμένου στον τοπικό διακομιστή.

Η διαδικασία της μετατροπής σε αρχεία κειμένου περιλαμβάνει τα εξής βήματα:

- Λήψη του ανακτηθέντος κειμένου από τα PDF αρχεία.
- Χρήση των ενσωματωμένων λειτουργιών της Pythοn για την εγγραφή του κειμένου σε αρχεία κειμένου.
- Αποθήκευση των αρχείων κειμένου στον τοπικό διακομιστή για περαιτέρω επεξεργασία.

Το επόμενο σημαντικό βήμα μετά την μετατροπή των PDF αρχείων σε αρχεία κειμένου είναι η αποθήκευση των δεδομένων σε μια βάση δεδομένων. Η βάση δεδομένων λειτουργεί ως αποθηκευτικός χώρος για τα δεδομένα που έχουμε συλλέξει, επιτρέποντάς μας να τα διατηρούμε οργανωμένα και προσβάσιμα για περαιτέρω ανάλυση.

3.4 Εισαγωγή Δεδομένων στη Βάση Δεδομένων

Η εισαγωγή των δεδομένων που έχουμε ανακτήσει και επεξεργαστεί από τα αρχεία PDF στη βάση δεδομένων αποτελεί σημαντικό βήμα προς την ολοκλήρωση του επεξεργαστικού κύκλου. Αυτή η διαδικασία επιτρέπει την οργανωμένη αποθήκευση, διαχείριση και ανακτησιμότητα των πληροφοριών, προσφέροντας μια πλατφόρμα για επιπλέον αναλύσεις και οπτικοποίηση.

Για την εισαγωγή των δεδομένων στη βάση δεδομένων, χρησιμοποιούμε την Pyodbc βιβλιοθήκη, η οποία παρέχει μια διασύνδεση με ODBC-συμβατές βάσεις δεδομένων. Ακολουθούμε τα εξής βήματα για την εισαγωγή των δεδομένων:

Σύνδεση με τη Βάση Δεδομένων: Αρχικά, δημιουργούμε μια σύνδεση με τη βάση δεδομένων MS SQL Server, χρησιμοποιώντας την Pyodbc. Παρέχουμε τις απαραίτητες πληροφορίες σύνδεσης, όπως το όνομα του διακομιστή, τη βάση δεδομένων, το όνομα χρήστη και τον κωδικό.

Δημιουργία SQL Εντολών: Δημιουργούμε SQL εντολές INSERT για κάθε καταχώρηση δεδομένων που επιθυμούμε να εισαγάγουμε στη βάση. Οι εντολές αυτές δημιουργούνται δυναμικά μέσα στα Python Scripts, λαμβάνοντας υπόψιν τα δεδομένα που έχουμε εξάγει από τα αρχεία PDF. Για παράδειγμα το SQL ερώτημα εισαγωγής τιμών για κάποια ημέρα του χειμώνα είναι διαφορετικό από μια καλοκαιρινή ημέρα καθώς το καλοκαίρι απουσιάζει καταγραφή για το καύσιμο Diesel Θέρμανσης.

Εισαγωγή Δεδομένων: Χρησιμοποιούμε τις δημιουργημένες SQL εντολές για την εισαγωγή των δεδομένων στη βάση δεδομένων. Εκτελούμε τις εντολές χρησιμοποιώντας τη σύνδεση που δημιουργήσαμε προηγουμένως.

Διαχείριση Διπλοεγγραφών: Πριν εισάγουμε δεδομένα στη βάση, ελέγχουμε αν τα δεδομένα έχουν ήδη εισαχθεί στο παρελθόν. Αυτό γίνεται για την αποφυγή διπλοεγγραφών στη βάση, διασφαλίζοντας την ακεραιότητα των δεδομένων.

Κλείσιμο Σύνδεσης: Αφού ολοκληρωθεί η εισαγωγή των δεδομένων, κλείνουμε τη σύνδεση με τη βάση δεδομένων για να απελευθερώσουμε πόρους.

Η διαδικασία εισαγωγής δεδομένων στη βάση δεδομένων αποτελεί βασικό στοιχείο για την ορθή λειτουργία του συστήματος. Κατά τη διάρκεια της διαδικασίας, τα αρχεία PDF που έχουν ανακτηθεί και μετατραπεί σε κείμενο εισάγονται στη βάση δεδομένων, δημιουργώντας ένα αξιόπιστο αποθετήριο για τα δεδομένα.

Ένα από τα βασικά πλεονεκτήματα αυτής της διαδικασίας είναι η δυνατότητα αναζήτησης και ανάκτησης πληροφοριών. Αποθηκεύοντας τα δεδομένα σε μια δομημένη βάση δεδομένων, μπορούμε να εκτελέσουμε αποτελεσματικές ερωτήσεις SQL για την εξαγωγή επιθυμητών

πληροφοριών. Επιπλέον, η βάση δεδομένων προσφέρει δυνατότητες για αναλύσεις, στατιστικά και άλλες μορφές επεξεργασίας δεδομένων που συμβάλλουν στην εξαγωγή συμπερασμάτων.

Το επόμενο βήμα μετά την εισαγωγή των δεδομένων είναι η ενημέρωση του διαχειριστή για τις νέες προσθήκες στη βάση. Αυτό γίνεται με την αποστολή ειδοποιήσεων ηλεκτρονικού ταχυδρομείου μέσω της Brevo transactional email API. Ο διαχειριστής ενημερώνεται για τις νέες καταχωρήσεις και μπορεί να παρακολουθεί την πρόοδο του συστήματος.

Συμπερασματικά, η διαδικασία εισαγωγής δεδομένων στη βάση δεδομένων αποτελεί κρίσιμο στάδιο της ανάκτησης και επεξεργασίας πληροφοριών από τα αρχεία PDF. Με την χρήση των προαναφερθεισών τεχνολογιών και βιβλιοθηκών, καθίσταται δυνατή η αυτοματοποίηση και η οργανωμένη διαχείριση των δεδομένων, επιτρέποντας την εξαγωγή ουσιαστικών πληροφοριών από τα αρχεία PDF και την ανάπτυξη ενός αποτελεσματικού πληροφοριακού συστήματος.

3.5 Ενημέρωση στον Διαχειριστή

Η επικοινωνία και ενημέρωση του διαχειριστή αποτελεί ένα ζωτικό στοιχείο στην αυτοματοποιημένη διαδικασία ανάκτησης και αποθήκευσης δεδομένων από αρχεία PDF. Μέσω της αποστολής ενημερωτικών μηνυμάτων ηλεκτρονικού ταχυδρομείου, ο διαχειριστής πληροφορείται για την πρόοδο και την κατάσταση του συστήματος, επιτρέποντας την άμεση αντίδραση σε πιθανά προβλήματα ή ανωμαλίες. Το email που αποστέλλεται περιέχει πληροφορίες σχετικά με την εκτέλεση του καθημερινού αυτοματοποιημένου σεναρίου ελέγχου δεδομένων και τις αναφορές των τελευταίων ημερών.

Ενημερωτικό Email

Το ενημερωτικό email περιλαμβάνει τις ακόλουθες ενότητες:

Στατιστικά Αρχείων

Αρχικά, το email παρέχει στον διαχειριστή πληροφορίες σχετικά με τον αριθμό των αρχείων που δεν έχουν εισαχθεί με επιτυχία στη βάση δεδομένων. Αυτό προσφέρει μια γρήγορη εικόνα της κατάστασης της εισαγωγής δεδομένων και επιτρέπει στον διαχειριστή να εντοπίζει πιθανά προβλήματα.

Τελευταίες Ενημερωμένες Εγγραφές

Ένα ακόμα σημαντικό κομμάτι του email είναι η λίστα με τις τελευταίες 15 ενημερωμένες εγγραφές ανά νομό. Κάθε εγγραφή περιλαμβάνει την ημερομηνία της ενημέρωσης, τον κωδικό επαλήθευσης, καθώς και διάφορα στατιστικά στοιχεία όπως οι τιμές για τα προϊόντα Un195,

Unl100, Motor Oil, Heating Oil και Autogas για κάθε νομό. Αυτή η λίστα προσφέρει μια επισκόπηση των πρόσφατων αλλαγών στα δεδομένα και επιτρέπει στον διαχειριστή να παρακολουθεί τις ενημερώσεις και την ακρίβεια των δεδομένων.

Επισκόπηση Εγγραφών ανά Νομό

Μια σημαντική πτυχή του ενημερωτικού email είναι η επισκόπηση του αριθμού των εγγραφών ανά νομό. Με αυτόν τον τρόπο, ο διαχειριστής μπορεί να παρατηρήσει την κατανομή των δεδομένων και να εντοπίσει πιθανές ανωμαλίες ή απουσίες δεδομένων σε συγκεκριμένους νομούς.

Δεδομένα των Τελευταίων Ημερών

Τέλος, το email παρέχει δεδομένα για την τρέχουσα ημερομηνία και τα τελευταία 3 ημερολογιακά ημέρες. Αυτά τα δεδομένα περιλαμβάνουν την ημερομηνία, τον νομό, τον κωδικό επαλήθευσης, και τις τιμές για τα προϊόντα Unl95, Unl100, Motor Oil, Heating Oil και Autogas. Αυτή η παρουσίαση προσφέρει μια συνολική εικόνα των πρόσφατων αλλαγών στα δεδομένα και επιτρέπει στον διαχειριστή να αντιλαμβάνεται τις τάσεις και τις μεταβολές των τιμών.

Συμπερασματικά, η αποστολή ενημερωτικών μηνυμάτων στο διαχειριστή είναι ζωτικής σημασίας για τη διατήρηση της λειτουργικότητας και της αξιοπιστίας του συστήματος ανάκτησης και αποθήκευσης δεδομένων από αρχεία PDF. Μέσω αυτών των ενημερωτικών email, ο διαχειριστής είναι σε θέση να παρακολουθεί την πρόοδο των διαδικασιών, να εντοπίζει πιθανά προβλήματα, και να λαμβάνει έγκαιρα μέτρα για την επίλυση των προκλήσεων. Μέσω αυτής της διαδικασίας, διασφαλίζεται η ομαλή και αποτελεσματική λειτουργία του συστήματος και η ακρίβεια των δεδομένων που ανακτώνται από τα αρχεία PDF.

3.6 Αυτοματοποίηση με τη χρήση Cron Jobs

Σε αυτό το κεφάλαιο, θα εξετάσουμε πώς χρησιμοποιούμε την τεχνολογία των Cron Jobs για να αυτοματοποιήσουμε την εκτέλεση των βασικών εργασιών της πλατφόρμας μας για την αποθήκευση και ανάκτηση δεδομένων. Μια αποτελεσματική αυτοματοποίηση είναι ουσιώδης για τη διατήρηση της ακρίβειας, της συνέπειας και της αποτελεσματικότητας των εργασιών μας, χωρίς να απαιτείται επέμβαση από τον χρήστη.

Εκτέλεση Εργασιών με τη Βοήθεια των Cron Jobs

Οι Cron Jobs αποτελούν ένα σύστημα αυτοματοποιημένων εργασιών σε λειτουργικά συστήματα Unix-like, όπως το Linux. Με τη χρήση των Cron Jobs, μπορούμε να ορίσουμε

εργασίες που εκτελούνται αυτόματα σε συγκεκριμένες χρονικές στιγμές ή με βάση κάποιο χρονοδιάγραμμα. Αυτό μας επιτρέπει να εκτελούμε τις εργασίες μας με τακτικότητα και ακρίβεια.

Το Script των Cron Jobs

Στην περίπτωση της πλατφόρμας μας για αποθήκευση και ανάκτηση δεδομένων, έχουμε δημιουργήσει ένα σύνολο από Python scripts για την εκτέλεση των βασικών εργασιών μας. Αυτά τα scripts είναι υπεύθυνα για τις παρακάτω ενέργειες:

- 1.Λήψη των αρχείων PDF από τον ιστότοπο.
- 2.Μετατροπή των αρχείων PDF σε κείμενο.
- 3.Εισαγωγή των δεδομένων από τα αρχεία κειμένου σε μια βάση δεδομένων.
- 4.Εκτέλεση επαλήθευσης των δεδομένων και αποστολή ειδοποιήσεων μέσω email.

Για την αυτοματοποίηση της εκτέλεσης αυτών των εργασιών, έχουμε δημιουργήσει ένα cron job script που καλεί αυτά τα Python scripts σε συγκεκριμένες χρονικές στιγμές.

```
bash
Copy code
#!/bin/sh
cd /var/www/html/fuel/FUEL_1_get_pdf/
python3 main.py
sleep 10m
cd /var/www/html/fuel/FUEL_2_pdf_to_text/
python3 main.py
sleep 10m
cd /var/www/html/fuel/FUEL_3_text_to_db/
python3 main.py
sleep 10m
cd /var/www/html/fuel/FUEL_5_Email_validation/
python3 main.py
```

Ανάλυση του Cron Job Script

Το παραπάνω script αποτελεί ένα συνολικό cron job που εκτελεί τις παραπάνω εργασίες σειριακά. Ας αναλύσουμε το script βήμα προς βήμα:

Πρώτα, το script μεταβαίνει στον κατάλογο `/var/www/html/fuel/FUEL_1_get_pdf/`, όπου βρίσκονται τα αρχεία που κατεβάζουν τα PDF από τον ιστότοπο. Χρησιμοποιεί την εντολή `cd` για να αλλάξει τον τρέχοντα κατάλογο σε αυτόν.

Έπειτα, το script εκτελεί το `main.py` του καταλόγου `FUEL_1_get_pdf/`, το οποίο είναι υπεύθυνο για τη λήψη των αρχείων PDF από τον ιστότοπο του υπουργείου.

Μετά την ολοκλήρωση της πρώτης εργασίας, το script περιμένει για 10 λεπτά (`sleep 10m`) πριν συνεχίσει με την επόμενη εργασία. Αυτό δίνει αρκετό χρόνο στην πρώτη εργασία να ολοκληρωθεί πριν ξεκινήσει η επόμενη.

Έπειτα, το script μεταβαίνει στον κατάλογο `/var/www/html/fuel/FUEL_2_pdf_to_text/`, όπου βρίσκονται τα αρχεία που μετατρέπουν τα αρχεία PDF σε αρχεία κειμένου. Χρησιμοποιεί την εντολή `cd` για να αλλάξει τον τρέχοντα κατάλογο σε αυτόν.

Στη συνέχεια, το script εκτελεί το `main.py` του καταλόγου `FUEL_2_pdf_to_text/`, το οποίο είναι υπεύθυνο για τη μετατροπή των αρχείων PDF.

Μετά την ολοκλήρωση αυτής εργασίας, το script περιμένει για 10 λεπτά (`sleep 10m`) πριν συνεχίσει με την επόμενη εργασία. Αυτό δίνει αρκετό χρόνο αυτή την εργασία να ολοκληρωθεί πριν ξεκινήσει η επόμενη.

Έπειτα, το script μεταβαίνει στον κατάλογο `/var/www/html/fuel/FUEL_3_text_to_db/`, όπου βρίσκονται τα αρχεία που εισάγουν αρχεία κειμένου στη βάση δεδομένων. Χρησιμοποιεί την εντολή `cd` για να αλλάξει τον τρέχοντα κατάλογο σε αυτόν.

Στη συνέχεια, το script εκτελεί το `main.py` του καταλόγου `FUEL_3_text_to_db/`, το οποίο είναι υπεύθυνο για την εισαγωγή των αρχείων κειμένου στη βάση δεδομένων.

Μετά την ολοκλήρωση αυτής εργασίας, το script περιμένει για 10 λεπτά (`sleep 10m`) πριν συνεχίσει με την επόμενη εργασία. Αυτό δίνει αρκετό χρόνο αυτή την εργασία να ολοκληρωθεί πριν ξεκινήσει η επόμενη.

Έπειτα, το script μεταβαίνει στον κατάλογο `/var/www/html/fuel/FUEL_5_Email_validation/`, όπου βρίσκονται τα αρχεία που είναι υπεύθυνα για την αποστολή της ειδοποίησης προς τους διαχειριστές. Χρησιμοποιεί την εντολή `cd` για να αλλάξει τον τρέχοντα κατάλογο σε αυτόν.

Στη συνέχεια, το script εκτελεί το main.py του καταλόγου FUEL_5_Email_validation/, το οποίο είναι υπεύθυνο για την αποστολή της ειδοποίησης προς τους διαχειριστές.

Συμπερασματικά, μέσω της χρήσης των Cron Jobs και του σχετικού script, μπορούμε να επιτύχουμε μια αυτοματοποιημένη διαδικασία αποθήκευσης, επεξεργασίας και διαχείρισης δεδομένων. Οι Cron Jobs εξασφαλίζουν την τακτική και ακριβή εκτέλεση των εργασιών, ενώ το σχετικό script διευκολύνει την οργανωμένη και αποτελεσματική εκτέλεσή τους. Αυτή η αυτοματοποίηση ενισχύει την αξιοπιστία και τη συνέπεια του συστήματός μας.

3.7 Πηγαίος κώδικας της εφαρμογής

Ακολουθούν αποσπάσματα από σημεία του πηγαίου κώδικα της εφαρμογής όπου αποτυπώνονται και επεξηγούνται οι κύριες λειτουργικότητες της εφαρμογής.

FUEL_get_pdf

Το πρώτο script λαμβάνει και αποθηκεύει τοπικά τα αρχεία pdf.

Στην αρχή λαμβάνει τη σελίδα του υπουργείου και από αυτήν αναζητεί όλα τα links που είναι αρχεία τύπου pdf.

```
main.py x
14 # requests url and get response object
15 response = requests.get(url)
16
17 # Parse text obtained
18 soup = BeautifulSoup(response.text, 'html.parser')
19
20 # Find all hyperlinks present on webpage
21 links = soup.find_all('a')
22
23 i = 0
24
25 # Elegxei ta link gia pdf
26 # An vrei pdf to katevazei
27 for link in links:
28     file_date = link.text.strip()
29
30     if (link == '<a class="bluelink" href="/manual.pdf" target="_blank">06nyics xp6nsq</a>'):
31         continue
32
33 # Gets links filename
34 s = link.get('href')
35 pdf_name = s.split('/')[-1]
36
```

Αν το αρχείο pdf δεν βρεθεί ως εγγραφή στη βάση δεδομένων, τότε το αποθηκεύει τοπικά.

```
IMERISIO_DELTIO_ANA_NOMO_17_08_2023.pdf
Downloading file: 8
File 8 downloaded
16/08/2023
http://www.fuelprices.gr/./files/deltia/IMERISIO\_DELTIO\_ANA\_NOMO\_16\_08\_2023.pdf
IMERISIO_DELTIO_ANA_NOMO_16_08_2023.pdf
Downloading file: 9
File 9 downloaded
15/08/2023
http://www.fuelprices.gr/./files/deltia/IMERISIO\_DELTIO\_ANA\_NOMO\_15\_08\_2023.pdf
IMERISIO_DELTIO_ANA_NOMO_15_08_2023.pdf
Downloading file: 10
```

Αν το αρχείο βρεθεί ως εγγραφή στη βάση δεδομένων, τότε το αγνοεί.

```
Run: main x
↑ http://www.fuelprices.gr/./files/deltia/IMERISIO\_DELTIO\_ANA\_NOMO\_07\_09\_2018.pdf
↓ IMERISIO_DELTIO_ANA_NOMO_07_09_2018.pdf
⇅ IMERISIO_DELTIO_ANA_NOMO_07_09_2018.pdf <--- Exists
☰ 06/09/2018
☰ http://www.fuelprices.gr/./files/deltia/IMERISIO\_DELTIO\_ANA\_NOMO\_06\_09\_2018.pdf
☰ IMERISIO_DELTIO_ANA_NOMO_06_09_2018.pdf <--- Exists
☰ 05/09/2018
☰ http://www.fuelprices.gr/./files/deltia/IMERISIO\_DELTIO\_ANA\_NOMO\_05\_09\_2018.pdf
☰ IMERISIO_DELTIO_ANA_NOMO_05_09_2018.pdf <--- Exists
☰ 04/09/2018
☰ http://www.fuelprices.gr/./files/deltia/IMERISIO\_DELTIO\_ANA\_NOMO\_04\_09\_2018.pdf
```

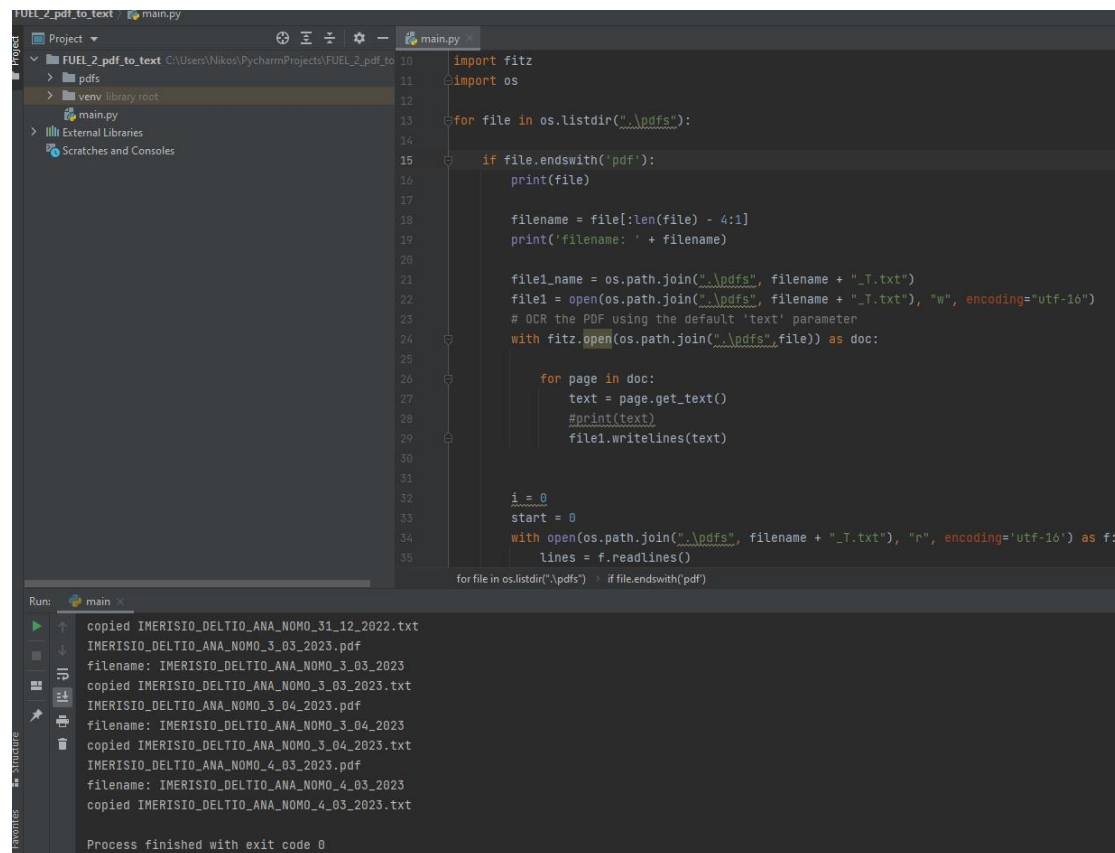
Τέλος μεταφέρει τα αρχεία στον φάκελο που τα αναμένει το επόμενο script

```
Run: main x
↑ copied IMERISIO_DELTIO_ANA_NOMO_30_11_2022.pdf
↓ copied IMERISIO_DELTIO_ANA_NOMO_30_12_2022.pdf
⇅ copied IMERISIO_DELTIO_ANA_NOMO_31_01_2023.pdf
☰ copied IMERISIO_DELTIO_ANA_NOMO_31_03_2023.pdf
☰ copied IMERISIO_DELTIO_ANA_NOMO_31_05_2023.pdf
☰ copied IMERISIO_DELTIO_ANA_NOMO_31_07_2023.pdf
☰ copied IMERISIO_DELTIO_ANA_NOMO_31_12_2022.pdf
☰ copied IMERISIO_DELTIO_ANA_NOMO_3_03_2023.pdf
☰ copied IMERISIO_DELTIO_ANA_NOMO_3_04_2023.pdf
☰ copied IMERISIO_DELTIO_ANA_NOMO_4_03_2023.pdf

Process finished with exit code 0
```

FUEL_pdf_to_text

Το δεύτερο script μετατρέπει τα αρχεία pdf σε txt αρχεία.



```
main.py
10 import fitz
11 import os
12
13 for file in os.listdir("pdfs"):
14
15     if file.endswith('pdf'):
16         print(file)
17
18         filename = file[:len(file) - 4:1]
19         print('filename: ' + filename)
20
21         file1_name = os.path.join("pdfs", filename + ".T.txt")
22         file1 = open(os.path.join("pdfs", filename + ".T.txt"), "w", encoding="utf-16")
23         # OCR the PDF using the default 'text' parameter
24         with fitz.open(os.path.join("pdfs", file)) as doc:
25
26             for page in doc:
27                 text = page.get_text()
28                 #print(text)
29                 file1.writelines(text)
30
31
32 i = 0
33 start = 0
34 with open(os.path.join("pdfs", filename + ".T.txt"), "r", encoding='utf-16') as f:
35     lines = f.readlines()
36
37 for file in os.listdir("pdfs") if file.endswith(pdf)
```

```
Run: main
copied IMERISIO_DELTIO_ANA_NOMO_31_12_2022.txt
IMERISIO_DELTIO_ANA_NOMO_3_03_2023.pdf
filename: IMERISIO_DELTIO_ANA_NOMO_3_03_2023
copied IMERISIO_DELTIO_ANA_NOMO_3_03_2023.txt
IMERISIO_DELTIO_ANA_NOMO_3_04_2023.pdf
filename: IMERISIO_DELTIO_ANA_NOMO_3_04_2023
copied IMERISIO_DELTIO_ANA_NOMO_3_04_2023.txt
IMERISIO_DELTIO_ANA_NOMO_4_03_2023.pdf
filename: IMERISIO_DELTIO_ANA_NOMO_4_03_2023
copied IMERISIO_DELTIO_ANA_NOMO_4_03_2023.txt
Process finished with exit code 0
```

Όταν ολοκληρωθεί η μετατροπή, μεταφέρει τα αρχεία στο φάκελο όπου τα αναμένει το επόμενο script.

```
# Antigrafoume ta text sto folder pou ta perimenei to trito script, oste na ginoun import sti vasi
# *****
import os
import shutil

source_folder = r"C:\Users\Nikos\PycharmProjects\FUEL_2_pdf_to_text\pdfs\"
destination_folder = r"C:\Users\Nikos\PycharmProjects\FUEL_3_text_to_db\docs\"

# fetch all files
for file_name in os.listdir(source_folder):
    if file_name.endswith('txt'):
        # construct full file path
        source = source_folder + file_name
        destination = destination_folder + file_name
        # copy only files
        if os.path.isfile(source):
            shutil.copy(source, destination)
            print('copied', file_name)
        # Delete temporary file
        os.remove(source)

# print("All PDF files downloaded")
```

FUEL_3_text_to_db

Το τρίτο script εισάγει τα δεδομένα στη βάση δεδομένων.

```
cursor = conn.cursor()
cursor.execute('''
    IF NOT EXISTS
    (
        SELECT county_code
        FROM prices
        WHERE county_code = ?
        and price_date = ?
    )
    BEGIN
        INSERT INTO [dbo].[prices]
            ([county_code]
            ,[unleaded_95]
            ,[unleaded_100]
            ,[motor_oil]
            ,[autogas]
            ,[price_date]
            ,[validation]
            ,[record_date]
            )
        VALUES
            (?
            ,?
            )
    END
''')
```

Η εισαγωγή των δεδομένων γίνεται με την εκτέλεση των κατάλληλων SQL ερωτημάτων.

```

main.py x
2 from datetime import datetime
3 import pyodbc
4 conn = pyodbc.connect('Driver={SQL Server};
5                       'Server=DESKTOP-NRAMQ1P\SQL;'
6                       'Database=Fuel;'
7                       'Trusted_Connection=yes;')
8
9 for file in os.listdir("..\docs"):
10
11     if file.endswith('txt'):
12
13         filename_pdf = file[:len(file) - 4:1] + '.pdf'
14         print('*****')
15         print(filename_pdf)
16         print('*****')
17         i = 0
18         start = 0
19         line_count = 0
20         with open(os.path.join("..\docs", file), "r", encoding='utf-16') as f:
21             lines = f.readlines()
22             i = 0
23             start = 0
24             end = 0
25             thermansi = ''
26             souper = ''
27             # SQL Script pou epistrefei an to arxeio exei ginei import
28             cursor = conn.cursor()
29             cursor.execute('''
30
31                 select top 1 imported
32                 from downloads
33                 where name = ?

```

Στη περίπτωση που έχουμε επιτυχή εισαγωγή των δεδομένων, ενημερώνεται αντίστοιχα η βάση δεδομένων

```

# SQL script pou kanei update to imported ston downloads
cursor = conn.cursor()
cursor.execute('''
update downloads set imported = 1 where doc_date = ?
'''
(conn.cursor(),
(doc_date))
conn.commit()

```

FUEL_4_Email_validation

Το τέταρτο script αποστέλλει το αυτοματοποιημένο email προς τους διαχειριστές

```
1 from __future__ import print_function
2 import time
3 import sib_api_v3_sdk
4 from sib_api_v3_sdk.rest import ApiException
5 from pprint import pprint
6 from datetime import date
7 import os
8 import pyodbc
9
10 conn = pyodbc.connect('Driver={SQL Server};'
11                      'Server=DESKTOP-NRAM01\SQL;'
12                      'Database=Fuel;'
13                      'Trusted_Connection=yes;')
14
15
16 # Email Configuration
17 # *****
18 configuration = sib_api_v3_sdk.Configuration()
19 configuration.api_key['api-key'] = 'xkeysib-c5b7a159da4fd0fe817d42df0ff6b24534b64fc1
20 api_instance = sib_api_v3_sdk.TransactionalEmailsApi(sib_api_v3_sdk.ApiClient(configu
21
22
Run: main x
C:\Users\Nikos\PycharmProjects\FUEL_5_Email_validation\venv\Scripts\python.exe C:/Users/Nikos/PycharmProjects/FUEL_5_Email_validation/main.py
{'message_id': '<202308270803.26927029235@smtp-relay.mailin.fr>',
 'message_ids': None}
Process finished with exit code 0
```

Ελέγχει ποια αρχεία δεν έχουν εισαχθεί με επιτυχία στη βάση δεδομένων

```
# SQL Scripts for Data Validation
# *****
# SQL Script που epistrefei poia arxeia den exoun ginei import
cursor = conn.cursor()
cursor.execute(''
               select count(doc_date)
               from downloads
               where imported = 0 and year = 2022
               ''
               )
result = cursor.fetchone()
imported = result[0]
```

Δημιουργεί το email και το αποστέλλει σε συγκεκριμένη λίστα αποδεκτών

```

# Email Recipients
# *****
sender = {"name": "FUEL Administration", "email": "formertangerine@aol.com"}
to = [{"email": "npa@gravitylabs.gr", "name": "Nikos"}]
#cc = [{"email": "stoug@ihu.gr", "name": "Stefanos"}]
cc = None
#bcc = [{"name": "Nikol", "email": "nikolman@gmail.com"}]
bcc = None
#reply_to = {"email": "replyto@domain.com", "name": "John Doe"}
reply_to = None
result1 = date.today().strftime("%d/%m/%Y")

# Email Subject
# *****
subject = "FUEL Data Validation - " + str(result1)

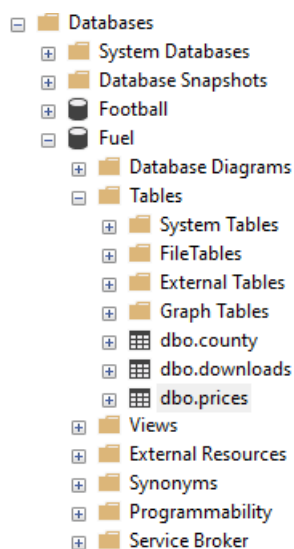
# Email Body
# *****
html_content = "<html><body><h3>Daily automated script for Data Validation</h3></body></html>"

# Not imported files
html_content += "\n" + "Not imported files: " + str(imported) + "<br>" + "<br>"
#for row in cursor.fetchall():
#    html_content += "\n" + row[0]

```

3.8 Σχήμα βάσης δεδομένων

Σε αυτό το κεφάλαιο, θα αναλύσουμε το σχήμα της βάσης δεδομένων μας, με όνομα FUEL. Η βάση δεδομένων αποτελεί την κεντρική αποθήκη πληροφοριών που σχετίζονται με τις τιμές των υγρών καυσίμων στην Ελλάδα. Το σχήμα της βάσης δεδομένων FUEL περιλαμβάνει τρεις πίνακες: τον πίνακα county (νομοί), τον πίνακα downloads (λήψεις αρχείων PDF), και τον πίνακα prices (τιμές καυσίμων).

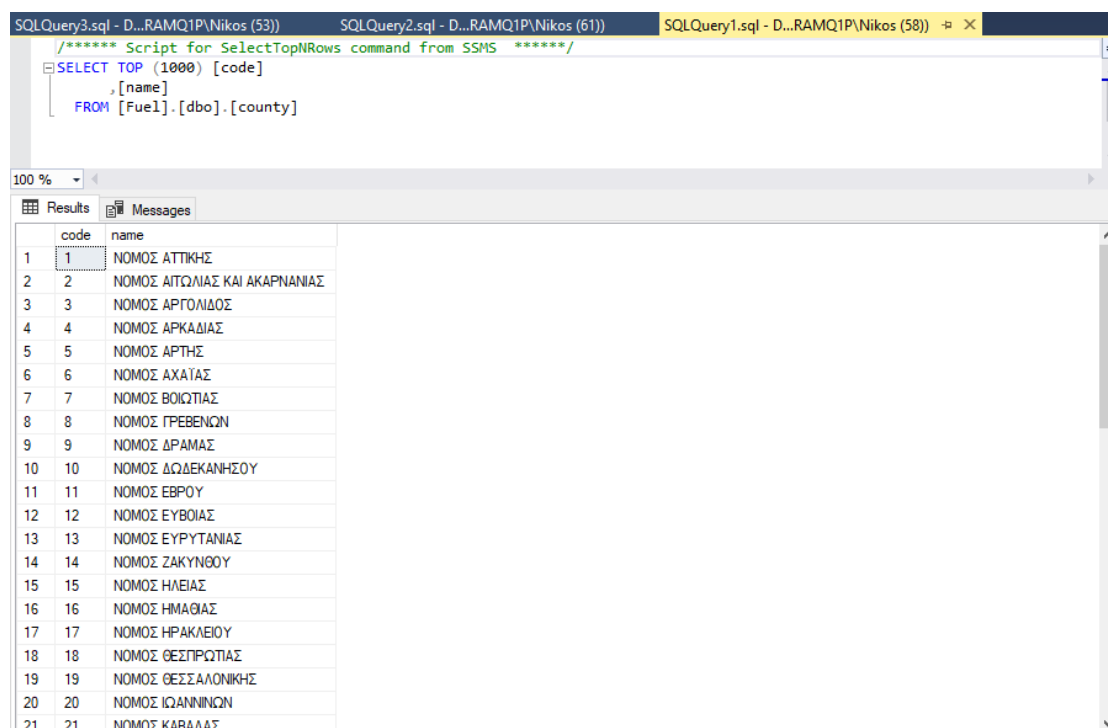


Πίνακας county

Ο πίνακας county περιλαμβάνει τις περιγραφές των νομών της Ελλάδος. Κάθε εγγραφή σε αυτόν τον πίνακα αντιπροσωπεύει έναν νομό και περιέχει τα παρακάτω πεδία:

code: Ένας μοναδικός αριθμητικός αναγνωριστικός αριθμός για κάθε νομό.

name: Το όνομα του νομού.



```
/****** Script for SelectTopNRows command from SSMS *****/
SELECT TOP (1000) [code]
, [name]
FROM [Fuel].[dbo].[county]
```

	code	name
1	1	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
2	2	ΝΟΜΟΣ ΑΙΤΩΛΙΑΣ ΚΑΙ ΑΚΑΡΝΑΝΙΑΣ
3	3	ΝΟΜΟΣ ΑΡΓΟΛΙΔΟΣ
4	4	ΝΟΜΟΣ ΑΡΚΑΔΙΑΣ
5	5	ΝΟΜΟΣ ΑΡΤΗΣ
6	6	ΝΟΜΟΣ ΑΧΑΪΑΣ
7	7	ΝΟΜΟΣ ΒΟΙΩΤΙΑΣ
8	8	ΝΟΜΟΣ ΓΡΕΒΕΝΩΝ
9	9	ΝΟΜΟΣ ΔΡΑΜΑΣ
10	10	ΝΟΜΟΣ ΔΩΔΕΚΑΝΗΣΟΥ
11	11	ΝΟΜΟΣ ΕΒΡΟΥ
12	12	ΝΟΜΟΣ ΕΥΒΟΙΑΣ
13	13	ΝΟΜΟΣ ΕΥΡΥΤΑΝΙΑΣ
14	14	ΝΟΜΟΣ ΖΑΚΥΝΘΟΥ
15	15	ΝΟΜΟΣ ΗΛΕΙΑΣ
16	16	ΝΟΜΟΣ ΗΜΑΘΙΑΣ
17	17	ΝΟΜΟΣ ΗΡΑΚΛΕΙΟΥ
18	18	ΝΟΜΟΣ ΘΕΣΣΑΛΟΝΙΚΗΣ
19	19	ΝΟΜΟΣ ΘΕΣΣΑΛΟΝΙΚΗΣ
20	20	ΝΟΜΟΣ ΙΩΑΝΝΙΝΩΝ
21	21	ΝΟΜΟΣ ΚΑΒΑΛΑΣ

Πίνακας downloads

Ο πίνακας downloads καταγράφει πληροφορίες σχετικά με τα αρχεία PDF που έχουν ληφθεί και αποθηκευτεί στη βάση δεδομένων. Αυτός ο πίνακας περιλαμβάνει τα παρακάτω πεδία:

id: Ένας μοναδικός αριθμητικός αναγνωριστικός αριθμός για κάθε λήψη αρχείου PDF.

name: Το όνομα του αρχείου PDF.

year: Η χρονολογία των τιμών που περιέχει το αρχείο.

imported: Πεδίο ελέγχου, το οποίο έχει τις εξής τιμές: 0 = Δεν έχει ολοκληρωθεί η εισαγωγή των τιμών του αρχείου στη βάση δεδομένων, 1 = Έχει ολοκληρωθεί με επιτυχία η εισαγωγή των τιμών του αρχείου στη βάση δεδομένων.

doc_date: Η ημερομηνία των τιμών που περιέχει το αρχείο.

SQLQuery3.sql - D:\RAMQ1P\Nikos (53) | SQLQuery2.sql - D:\RAMQ1P\Nikos (61) | SQLQuery1.sql - D:\RAMQ1P\Nikos (58)

```

/***** Script for SelectTopNRows command from SMS *****/
SELECT TOP (1000) [id]
      , [name]
      , [year]
      , [imported]
      , [doc_date]
FROM [Fuel].[dbo].[downloads]

```

	id	name	year	imported	doc_date
1	4E35889A-5BA7-4C12-B1DF-857FE3D3DA2B	IMERISIO_DELTIO_ANA_NOMO_29_08_2022.pdf	2022	1	29/08/2022
2	86333A94-BFF8-47F9-8BFD-F66028E6D013	IMERISIO_DELTIO_ANA_NOMO_09_07_2022.pdf	2022	1	09/07/2022
3	25FFB6E8-7DCD-4DFA-9AD4-9DE8F9993F2B	IMERISIO_DELTIO_ANA_NOMO_19_05_2022.pdf	2022	1	19/05/2022
4	4431DD2A-688A-4231-B9A4-4E40DA9EA253	IMERISIO_DELTIO_ANA_NOMO_29_03_2022.pdf	2022	1	29/03/2022
5	1CC08618-7182-444F-A446-A3B8808732F4	IMERISIO_DELTIO_ANA_NOMO_06_02_2022.pdf	2022	1	06/02/2022
6	F3835AFD-D9EA-4E5F-9FBE-F2BF6132F820	IMERISIO_DELTIO_ANA_NOMO_16_12_2021.pdf	2021	1	16/12/2021
7	BEF61D61-A586-40BB-B3AB-E5C17FE3FA7A	IMERISIO_DELTIO_ANA_NOMO_26_10_2021.pdf	2021	1	26/10/2021
8	C7FA3935-3363-4E7C-908D-F9EC59F58829	IMERISIO_DELTIO_ANA_NOMO_05_09_2021.pdf	2021	1	05/09/2021
9	ADD5E463-FA9A-4C3D-AB3A-2BE1DC65E990	IMERISIO_DELTIO_ANA_NOMO_16_07_2021.pdf	2021	1	16/07/2021
10	89E85843-77EE-4C4C-93B0-98563A5D063F	IMERISIO_DELTIO_ANA_NOMO_26_05_2021.pdf	2021	1	26/05/2021
11	BFF81ECD-0CEF-421C-A0CB-DBB397393933	IMERISIO_DELTIO_ANA_NOMO_05_04_2021.pdf	2021	1	05/04/2021
12	93541F6E-C6A7-4F6C-B01F-6AR2FAC06E28	IMERISIO_DELTIO_ANA_NOMO_13_02_2021.pdf	2021	1	13/02/2021

Πίνακας prices

Ο πίνακας prices περιέχει πληροφορίες σχετικά με τις τιμές των υγρών καυσίμων στην Ελλάδα. Κάθε εγγραφή σε αυτόν τον πίνακα αντιπροσωπεύει την τιμή ενός καυσίμου για μια συγκεκριμένη ημερομηνία και σε έναν συγκεκριμένο νομό. Ο πίνακας περιέχει τα παρακάτω πεδία:

id: Ένας μοναδικός αριθμητικός αναγνωριστικός αριθμός για κάθε τιμή.

county_code: Ο αναγνωριστικός αριθμός του νομού όπου ισχύει η τιμή.

unleaded_95: Η τιμή του καυσίμου unleaded_95 σε μια συγκεκριμένη ημερομηνία και νομό.

unleaded_100: Η τιμή του καυσίμου unleaded_100 σε μια συγκεκριμένη ημερομηνία και νομό.

motor_oil: Η τιμή του καυσίμου motor_oil σε μια συγκεκριμένη ημερομηνία και νομό.

heating_oil: Η τιμή του καυσίμου heating_oil σε μια συγκεκριμένη ημερομηνία και νομό.

autogas: Η τιμή του καυσίμου autogas σε μια συγκεκριμένη ημερομηνία και νομό.

price_date: Η ημερομηνία για την οποία καταγράφεται η τιμή.

validation: Ένα πεδίο που χρησιμοποιείται για τον έλεγχο της εγκυρότητας των δεδομένων.

SQLQuery4.sql - D:\RAMQ1P\Nikos (68))* × SQLQuery3.sql - D:\RAMQ1P\Nikos (53))* SQLQuery2.sql - D:\RAMQ1P\Nikos (61)) SQLQuery1.sql -

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT TOP (1000) [id]
, [county_code]
, [unleaded_95]
, [unleaded_100]
, [motor_oil]
, [heating_oil]
, [autogas]
, [price_date]
, [validation]

FROM [Fuel].[dbo].[prices]

```

100 %

Results Messages

	id	county_code	unleaded_95	unleaded_100	motor_oil	heating_oil	autogas	price_date	validation
1	78BE5D0E-2ABF-4DOC-89BD-38301A3E6E8F	9	1.471	1.770	1.316	0.983	0.783	01/01/2019	ΝΟΜΟΣ ΔΡΑΜΑΣ
2	5B60E809-8096-457D-9FA7-C632899CF4F8	18	1.635	1.880	1.436	1.120	0.849	01/01/2020	ΝΟΜΟΣ ΘΑΛΠΡΩΣΗΔΑ
3	C1C852E1-6B89-4F7A-A4C8-898F3CA43039	27	1.426	1.592	1.162	0.849	0.745	01/01/2021	ΝΟΜΟΣ ΚΟΖΑΝΗΣ
4	4FF85F1E-15E2-4699-86F9-73A8443BD0E1	36	1.760	1.924	1.483	1.146	1.045	01/01/2022	ΝΟΜΟΣ ΜΕΣΣΗΝΙΑΣ
5	D85B5004-8413-4A0B-9FFB-70A5D9815B1B	45	1.575	1.738	1.344	1.003	0.813	01/02/2018	ΝΟΜΟΣ ΣΠΙΚΑΛΩΝ
6	CAD00E99-F746-4506-A55D-7B36E90951E1	3	1.594	1.806	1.366	1.044	0.812	01/02/2020	ΝΟΜΟΣ ΑΡΓΟΛΙΔΟΣ
7	29D32C18-71C5-4B9B-BF29-D09E12BF8916	12	1.487	1.665	1.239	0.880	0.793	01/02/2021	ΝΟΜΟΣ ΕΥΒΟΙΑΣ
8	DD9B1E32-4B35-4D9A-9E13-2473838E340B	21	1.831	1.975	1.578	1.227	1.068	01/02/2022	ΝΟΜΟΣ ΚΑΒΑΛΑΣ
9	5D0AEA36-FC2C-48D4-93CF-5965BAE60B55	30	1.544	1.757	1.314	0.994	0.821	01/03/2018	ΝΟΜΟΣ ΛΑΚΩΝΙΑΣ
10	8CA86391-A105-4788-9C5D-A0BCB3C19BAD	39	1.527	1.715	1.365	1.045	0.815	01/03/2019	ΝΟΜΟΣ ΠΙΕΡΙΑΣ
11	47AD12DE-66E2-4F7F-8483-2FCBDF21517A	48	1.641	1.863	1.394	1.062	0.881	01/03/2020	ΝΟΜΟΣ ΦΩΚΗΓΩ
12	3D2040EC-3F26-4FF9-9C95-15F866A443A3	6	1.921	2.083	1.657	1.299	1.107	01/03/2022	ΝΟΜΟΣ ΑΧΑΪΑΣ
13	8B3B1098-6C69-42CB-A914-62722762E44C	15	1.550	1.723	1.315	0.998	0.809	01/04/2018	ΝΟΜΟΣ ΗΛΕΙΑΣ

Το πεδίο validation παίζει κρίσιμο ρόλο στη διασφάλιση της ορθότητας των δεδομένων στον πίνακα prices και στον έλεγχο της διαδικασίας εισαγωγής τους. Αυτό το πεδίο, που εισάγεται δυναμικά κατά την εισαγωγή των δεδομένων από το αρχείο txt, λειτουργεί ως ένα είδος ελέγχου συμφωνίας μεταξύ των δεδομένων και των κωδικών νομών.

Αν το πεδίο validation δεν αντιστοιχεί στη σωστή τιμή για έναν δεδομένο κωδικό, αυτό υποδηλώνει ότι υπάρχει κάποιο σφάλμα στη διαδικασία εισαγωγής των δεδομένων για αυτόν τον συγκεκριμένο νομό. Με αυτό το τρόπο αποτελεί ένα πολύ χρήσιμο μέσο για την ανίχνευση και τη διόρθωση πιθανών σφαλμάτων στα δεδομένα.

SQLQuery4.sql - D...RAMQ1P\Nikos (68)* SQLQuery3.sql - D...RAMQ1P\Nikos (53)* SQLQuery2.sql - D...RAMQ1P\Nikos (61)) SQLQuery1.sql - D...RAMQ1P\Nikos (60)*

```

/***** Script for SelectTopNRows command from SSMS *****/
SELECT TOP (1000) [id]
, [county_code]
, [unleaded_95]
, [unleaded_100]
, [motor_oil]
, [heating_oil]
, [autogas]
, [price_date]
, [validation]
FROM [Fuel].[dbo].[prices]
where county_code = 1

```

100 %

Results Messages

	id	county_code	unleaded_95	unleaded_100	motor_oil	heating_oil	autogas	price_date	validation
3	9F69C5F3-34E9-4C10-9E19-552AB05852DF	1	1.812	2.029	1.764	1.123	0.893	01/01/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
4	89135400-34E3-4039-ADA2-9AC5436F47DD	1	1.822	2.041	1.767	1.205	0.892	03/01/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
5	2F5B11AE-9881-4FE7-95E8-4EF2D9CD5C20	1	1.920	2.150	1.857	1.158	0.932	02/12/2022	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
6	564EB5AF-BBE0-444A-A9E5-CE5B7BA91E63	1	1.832	2.052	1.544	NULL	0.852	03/05/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
7	5803DEB5-4974-462F-8BE3-85525A95B27C	1	1.833	2.053	1.528	NULL	0.761	03/06/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
8	55C0176E-7BB7-47D2-A193-642306F16ADE	1	1.825	2.044	1.546	NULL	0.717	03/07/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
9	890740B2-7D35-4A09-AA88-E8F25DB68415	1	1.929	2.160	1.664	NULL	0.771	03/08/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
10	50599CD8-8FB5-4707-82E0-70AF1279A10C	1	2.029	2.272	2.059	1.329	0.945	03/11/2022	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
11	7264708A-B679-451C-B4E4-996222DA7654	1	1.918	2.148	1.854	1.159	0.931	03/12/2022	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
12	0E617470-9DF4-4A8E-B9D1-E75DB597DB12	1	1.825	2.044	1.767	1.219	0.892	04/01/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
13	2647B704-513D-4ECA-A6E0-445C5983F2F5	1	1.890	2.117	1.618	1.190	0.921	03/04/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
14	E96E32AF-8270-4E54-ADDD-0F982570D63C	1	1.867	2.091	1.677	1.182	0.954	04/03/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ
15	D87B9141-073B-4227-BDE8-04A6EC576A5A	1	1.864	2.088	1.673	1.178	0.951	03/03/2023	ΝΟΜΟΣ ΑΤΤΙΚΗΣ

3.9 Επικαιροποίηση δεδομένων

Στο παρόν κεφάλαιο πραγματοποιείτε η επεξήγηση της επικαιροποίησης των αρχικών δεδομένων. Η επικαιροποίηση των δεδομένων έγινε με τη χρήση των κατάλληλων SQL ερωτημάτων. Για την ολοκλήρωση της επικαιροποίησης ελέγχθηκαν οι τιμές ανά έτος.

2018

Υπάρχουν οι τιμές όλων των ημερομηνιών του έτους

SQLQuery3.sql - D...RAMQ1P\Nikos (51))* SQLQuery2.sql - D...RAMQ1P\Nikos (63))*

```

select LEFT(CAST(dbo.prices.record_date AS DATE), 4) AS etos
, count(record_date) count_of_days
FROM [Fuel].[dbo].[prices]
where county_code = 1
and LEFT(CAST(dbo.prices.record_date AS DATE), 4) = 2018
group by LEFT(CAST(dbo.prices.record_date AS DATE), 4)

```

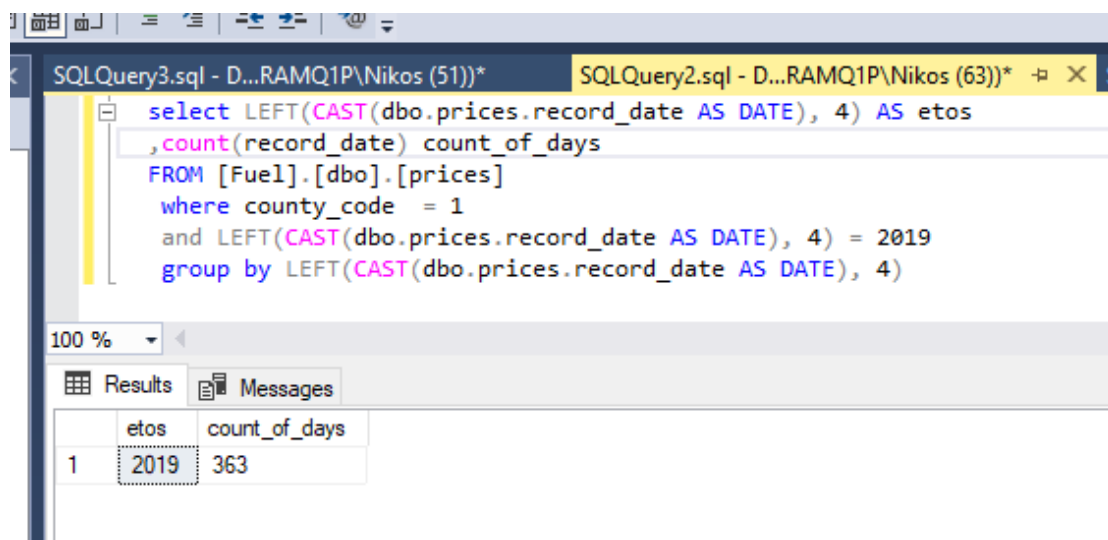
100 %

Results Messages

	etos	count_of_days
1	2018	365

2019

Απουσιάζουν δεδομένα για δύο ημέρες, λόγω του ότι δεν έχουν ανέβει τα αντίστοιχα αρχεία για τις δύο αυτές ημέρες στη σελίδα του υπουργείου



The screenshot shows a SQL query window with the following query:

```
select LEFT(CAST(dbo.prices.record_date AS DATE), 4) AS etos
, count(record_date) count_of_days
FROM [Fuel].[dbo].[prices]
where county_code = 1
and LEFT(CAST(dbo.prices.record_date AS DATE), 4) = 2019
group by LEFT(CAST(dbo.prices.record_date AS DATE), 4)
```

The results pane shows a single row:

etos	count_of_days
2019	363



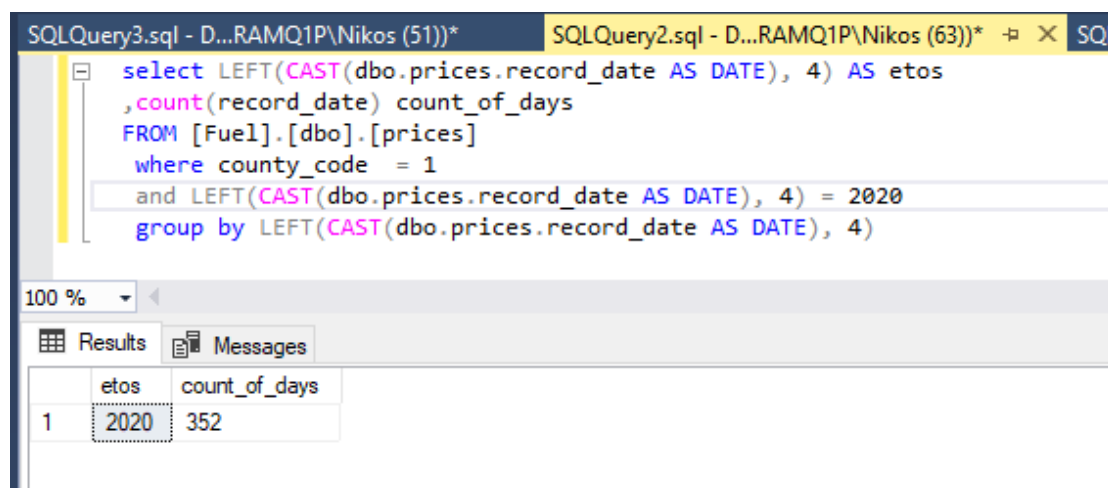
The screenshot shows a list of dates for December 2019, with a red box highlighting the dates from 31/12/2019 to 01/12/2019. The list is as follows:

Δεκέμβριος 2019:
31/12/2019, 30/12/2019, 27/12/2019, 26/12/2019, 25/12/2019, 24/12/2019, 23/12/2019, 22/12/2019, 21/12/2019, 20/12/2019, 19/12/2019, 18/12/2019, 17/12/2019, 16/12/2019, 15/12/2019, 14/12/2019, 13/12/2019, 12/12/2019, 11/12/2019, 10/12/2019, 09/12/2019, 08/12/2019, 07/12/2019, 06/12/2019, 05/12/2019, 04/12/2019, 03/12/2019, 02/12/2019, 01/12/2019

Νοέμβριος 2019:

2020

Απουσιάζουν τα δεδομένα για 14 ημέρες, λόγω του ότι τα αντίστοιχα αρχεία έχουν ανέβει στη σελίδα ως φωτογραφίες και έτσι δεν είναι δυνατή η εξαγωγή τους



The screenshot shows a SQL query window with the following query:

```
select LEFT(CAST(dbo.prices.record_date AS DATE), 4) AS etos
, count(record_date) count_of_days
FROM [Fuel].[dbo].[prices]
where county_code = 1
and LEFT(CAST(dbo.prices.record_date AS DATE), 4) = 2020
group by LEFT(CAST(dbo.prices.record_date AS DATE), 4)
```

The results pane shows a single row:

etos	count_of_days
2020	352

Name	Date modified	Type	Size
IMERISIO_DELTIO_ANA_NOMO_01_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	423 KB
IMERISIO_DELTIO_ANA_NOMO_02_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	426 KB
IMERISIO_DELTIO_ANA_NOMO_03_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	423 KB
IMERISIO_DELTIO_ANA_NOMO_04_07_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	485 KB
IMERISIO_DELTIO_ANA_NOMO_04_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	447 KB
IMERISIO_DELTIO_ANA_NOMO_05_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	450 KB
IMERISIO_DELTIO_ANA_NOMO_06_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	435 KB
IMERISIO_DELTIO_ANA_NOMO_07_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	512 KB
IMERISIO_DELTIO_ANA_NOMO_08_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	476 KB
IMERISIO_DELTIO_ANA_NOMO_09_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	476 KB
IMERISIO_DELTIO_ANA_NOMO_10_08_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	469 KB
IMERISIO_DELTIO_ANA_NOMO_29_07_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	409 KB
IMERISIO_DELTIO_ANA_NOMO_30_07_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	413 KB
IMERISIO_DELTIO_ANA_NOMO_31_07_2020.pdf	27/8/2023 10:35 πμ	Adobe Acrobat D...	430 KB



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΥΠΟΥΡΓΕΙΟ
ΑΝΑΠΤΥΞΗΣ ΚΑΙ ΕΠΕΝΔΥΣΕΩΝ

Αθήνα, 03-08-2020

Διυηρεσιακή Μονάδα Ελέγχου Αγοράς

Αριθ. Πρωτ. : 82417

Ταχ. Δ/ση : Πλ. Κάνιγγος, Αθήνα
Ταχ. Κώδικας : 106 77
Πληροφορίες : Νικολοπούλου Μαρία
Τηλέφωνο : 213 1514 284
Fax : 210 3815 317
Email : fuelprices@gge.gr
Ιστοσελίδα : www.mindev.gov.gr

ΠΡΟΣ : ΑΝΑΡΤΗΤΕΟ ΣΤΟ FUELPRICES

ΘΕΜΑ: ΚΑΘΗΜΕΡΙΝΟ ΔΕΛΤΙΟ ΕΠΙΣΚΟΠΗΣΗΣ ΤΙΜΩΝ ΥΓΡΩΝ ΚΑΥΣΙΜΩΝ

Μέσες τιμές λιανικής πώλησης καυσίμων ανά Νομό, για την

Σάββατο, 1 Αυγούστου 2020

(τιμές σε €/лт, συμπ. ΦΠΑ)

ΝΟΜΟΣ	Αμόλυβδη 95 οκτ.	Αμόλυβδη 100 οκτ.	Super	Diesel κίνησης	Υγραέριο κίνησης (Autogas)
ΝΟΜΟΣ ΑΤΤΙΚΗΣ	1,404	1,615	1,724	1,135	0,679
ΝΟΜΟΣ ΑΙΤΩΛΙΑΣ ΚΑΙ ΑΚΑΡΝΑΝΙΑΣ	1,455	1,697	-	1,174	0,773
ΝΟΜΟΣ ΑΡΓΟΛΙΔΟΣ	1,422	1,654	1,564	1,149	0,745
ΝΟΜΟΣ ΑΡΚΑΔΙΑΣ	1,454	1,674	-	1,177	0,750
ΝΟΜΟΣ ΑΡΤΗΣ	1,434	1,645	-	1,163	0,733
ΝΟΜΟΣ ΑΧΑΪΑΣ	1,428	1,651	1,678	1,167	0,756

2021

Υπάρχουν οι τιμές όλων των ημερομηνιών του έτους.

SQLQuery3.sql - D...RAMQ1P\Nikos (51))* SQLQuery2.sql - D...RAMQ1P\Nikos (63))* X S

```

select LEFT(CAST(dbo.prices.record_date AS DATE), 4) AS etos
,count(record_date) count_of_days
FROM [Fuel].[dbo].[prices]
where county_code = 1
and LEFT(CAST(dbo.prices.record_date AS DATE), 4) = 2021
group by LEFT(CAST(dbo.prices.record_date AS DATE), 4)

```

100 %

Results Messages

	etos	count_of_days
1	2021	365

2022

Απουσιάζουν δεδομένα για μία ημέρα, λόγω του ότι δεν έχει ανέβει το αντίστοιχο αρχείο για τη συγκεκριμένη ημέρα στη σελίδα του υπουργείου

SQLQuery3.sql - D...RAMQ1P\Nikos (51))* SQLQuery2.sql - D...RAMQ1P\Nikos (63))* X SQLQ

```

select LEFT(CAST(dbo.prices.record_date AS DATE), 4) AS etos
,count(record_date) count_of_days
FROM [Fuel].[dbo].[prices]
where county_code = 1
and LEFT(CAST(dbo.prices.record_date AS DATE), 4) = 2022
group by LEFT(CAST(dbo.prices.record_date AS DATE), 4)

```

100 %

Results Messages

	etos	count_of_days
1	2022	364

Ιανουάριος 2022:
 31/01/2022 , 30/01/2022 , 29/01/2022 , 28/01/2022 , 27/01/2022 , 26/01/2022 , 24/01/2022 , 23/01/2022 , 22/01/2022 ,
 21/01/2022 , 20/01/2022 , 19/01/2022 , 18/01/2022 , 17/01/2022 , 16/01/2022 , 15/01/2022 , 14/01/2022 , 13/01/2022 ,
 12/01/2022 , 11/01/2022 , 10/01/2022 , 09/01/2022 , 08/01/2022 , 07/01/2022 , 06/01/2022 , 05/01/2022 , 04/01/2022 ,
 03/01/2022 , 02/01/2022 , 01/01/2022

2023

Υπάρχουν οι τιμές όλων των ημερομηνιών του έτους, μέχρι και την ημερομηνία που έγινε η τελευταία ενημέρωση

SQLQuery4.sql - D:\RAMQ1P\Nikos (64))* SQLQuery2.sql - D:\RAMQ1P\Nikos (63))* SQLQuery3.sql - D:\RAMQ1P\Nikos (51))*

```
select LEFT(CAST(dbo.prices.record_date AS DATE), 4) AS etos , MONTH(CAST(dbo.prices.record_date AS DATE))
--, record_date
, count(record_date) count_of_days
FROM [Fuel].[dbo].[prices]
where county_code = 1
and LEFT(CAST(dbo.prices.record_date AS DATE), 4) = 2023
group by LEFT(CAST(dbo.prices.record_date AS DATE), 4), MONTH(CAST(dbo.prices.record_date AS DATE))
-- order by record_date
```

100 %

Results Messages

	etos	minas	count_of_days
1	2023	1	31
2	2023	2	28
3	2023	3	31
4	2023	4	30
5	2023	5	31
6	2023	6	30
7	2023	7	31
8	2023	8	24

4

Οπτικοποίηση των δεδομένων

4.1 Εγκατάσταση του Grafana, παραμετροποίηση και σύνδεση με τη βάση δεδομένων

Η εγκατάσταση, παραμετροποίηση και σύνδεση του Grafana με τη βάση δεδομένων αποτελούν σημαντικά βήματα για τη δημιουργία ενός λειτουργικού συστήματος οπτικοποίησης και ανάλυσης δεδομένων. Στο πλαίσιο αυτού του κεφαλαίου, θα αναλύσουμε τη διαδικασία εγκατάστασης του Grafana, την παραμετροποίηση του και τον τρόπο σύνδεσής του με τη βάση δεδομένων που χρησιμοποιούμε στο έργο μας.

Εγκατάσταση του Grafana

Η εγκατάσταση του Grafana αποτελεί ένα σημαντικό βήμα για την αρχικοποίηση του εργαλείου. Αρχικά, πρέπει να λάβουμε το πακέτο εγκατάστασης από την επίσημη ιστοσελίδα του Grafana (<https://grafana.com/>) και να ακολουθήσουμε τις οδηγίες εγκατάστασης που παρέχονται. Οι περισσότερες πλατφόρμες υποστηρίζονται, συμπεριλαμβανομένων Linux, Windows και macOS.

Παραμετροποίηση του Grafana

Μετά την εγκατάσταση, πρέπει να προχωρήσουμε στην παραμετροποίηση του Grafana. Η πρώτη είσοδος πραγματοποιείται μέσω του περιηγητή, όπου ορίζουμε τα απαραίτητα στοιχεία όπως το όνομα χρήστη και ο κωδικός πρόσβασης. Αξίζει να σημειώσουμε ότι ο Grafana προσφέρει πολλαπλά επίπεδα πρόσβασης και ασφάλειας, επιτρέποντάς μας να περιορίσουμε την πρόσβαση σε συγκεκριμένους χρήστες ή ομάδες.

Σύνδεση με τη Βάση Δεδομένων

Για να επιτύχουμε την οπτικοποίηση των δεδομένων από τη βάση δεδομένων, πρέπει να συνδέσουμε το Grafana με αυτήν. Στο πλαίσιο του έργου μας, χρησιμοποιούμε τη βάση δεδομένων SQL Server, οπότε πρέπει να παρέχουμε τις σωστές πληροφορίες σύνδεσης. Αυτές περιλαμβάνουν τη διεύθυνση IP ή το όνομα του διακομιστή, το όνομα της βάσης δεδομένων, το όνομα χρήστη και τον κωδικό πρόσβασης.

Δημιουργία Πίνακα Ελέγχου στο Grafana

Μετά τη σύνδεση με τη βάση δεδομένων, μπορούμε να δημιουργήσουμε πίνακες ελέγχου (dashboards) στο Grafana. Ο πίνακας ελέγχου είναι ο χώρος όπου θα τοποθετήσουμε τα γραφήματα, τα διαγράμματα και τα άλλα οπτικά στοιχεία που θέλουμε να παρακολουθούμε. Το Grafana παρέχει μια ευέλικτη διεπαφή χρήστη για τη δημιουργία αυτών των πινάκων ελέγχου, με δυνατότητα προσαρμογής τους στις ανάγκες μας.

Οπτικοποίηση των Δεδομένων των Υγρών Καυσίμων

Στο πλαίσιο του έργου μας, επικεντρωνόμαστε στην οπτικοποίηση των δεδομένων των τιμών των υγρών καυσίμων. Αυτό περιλαμβάνει τη δημιουργία γραφημάτων που απεικονίζουν τις παραλλαγές των τιμών των καυσίμων με την πάροδο του χρόνου ή σε σχέση με άλλα στοιχεία. Χρησιμοποιώντας τα διαθέσιμα εργαλεία του Grafana, μπορούμε να δημιουργήσουμε προσαρμοσμένα γραφήματα που αντιπροσωπεύουν απτά τα δεδομένα μας.

Επιπλέον, μπορούμε να χρησιμοποιήσουμε τις δυνατότητες του Grafana για να ορίσουμε ειδοποιήσεις όταν οι τιμές των υγρών καυσίμων υπερβαίνουν ή κάτω από συγκεκριμένα κατώτατα ή ανώτατα όρια. Με αυτόν τον τρόπο, μπορούμε να διατηρούμε την εποπτεία και να αντιδρούμε άμεσα σε περιπτώσεις που απαιτείται ενέργεια.

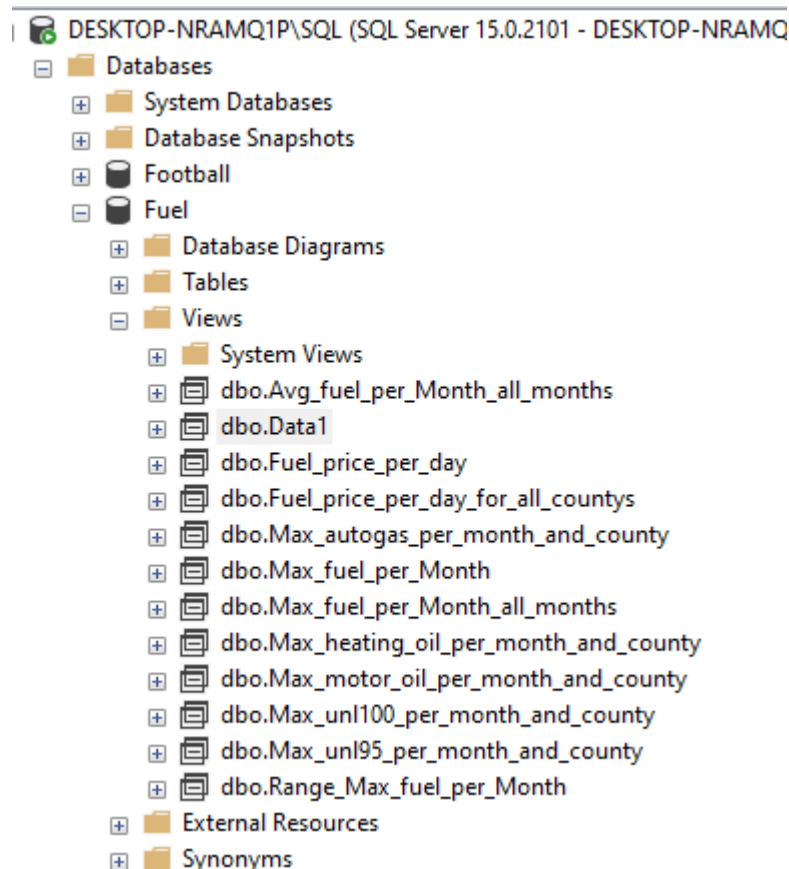
Συμπερασματικά, το Grafana αποτελεί ένα ισχυρό εργαλείο για την οπτικοποίηση και ανάλυση δεδομένων, και η χρήση του στο έργο μας έχει αποδειχθεί εξαιρετικά ωφέλιμη. Μέσω της εγκατάστασης, παραμετροποίησης και σύνδεσής του με τη βάση δεδομένων, καθώς και μέσω της δημιουργίας πινάκων ελέγχου και της οπτικοποίησης των δεδομένων των τιμών των υγρών καυσίμων, μπορούμε να παρακολουθούμε και να αναλύουμε την απόδοση του συστήματός μας με ακρίβεια και αξιοπιστία. Συνολικά, η ενσωμάτωση του Grafana στο έργο μας έχει βελτιώσει την ικανότητά μας να λαμβάνουμε αποφάσεις βασισόμενες σε δεδομένα και να διατηρούμε το σύστημά μας σε βέλτιστη λειτουργία.

4.2 Δημιουργία SQL Views

Η δημιουργία SQL Views αποτελεί ένα σημαντικό βήμα στη διαδικασία προετοιμασίας των δεδομένων για οπτικοποίηση από το Grafana. Σε αυτό το κεφάλαιο, θα εξετάσουμε πώς μπορούμε να δημιουργήσουμε SQL Views [9] που θα αποτελούν τα ερωτήματα για την εξαγωγή και μετασχηματισμό των δεδομένων που θέλουμε να οπτικοποιήσουμε στο Grafana.

Η Σημασία των SQL Views

Οι SQL Views είναι εικονικοί πίνακες που προσφέρουν μια επικάλυψη πάνω από τα υπάρχοντα δεδομένα στη βάση μας. Αυτές οι προβολές επιτρέπουν τη δημιουργία προσαρμοσμένων προβολών των δεδομένων χωρίς να αλλάζουμε τη βάση δεδομένων καθαυτή. Επιπλέον, επιτρέπουν την απλοποίηση των πολύπλοκων ερωτημάτων, δημιουργώντας προκαθορισμένες προβολές που συνήθως χρησιμοποιούμε. Στο πλαίσιο της εργασίας μας υλοποιήθηκε πλήθος από διάφορα SQL Views, με διαφορετικά χαρακτηριστικά. Στη παρακάτω εικόνα φαίνονται τα SQL Views που χρησιμοποιήθηκαν.



Δημιουργία SQL Views

Η διαδικασία δημιουργίας ενός SQL View απαιτεί τη σύνταξη ενός SQL ερωτήματος που θα καθορίζει τα δεδομένα που θέλουμε να επιστρέφονται από την προβολή. Ας υποθέσουμε ότι έχουμε μια βάση δεδομένων που περιλαμβάνει πληροφορίες για τις τιμές των υγρών καυσίμων και θέλουμε να δημιουργήσουμε μια προβολή που θα επιστρέφει τις τιμές μαζί με τα αντίστοιχα ονόματα των προϊόντων.

Χρήση SQL Views στο Grafana

Αφού έχουμε δημιουργήσει τις απαραίτητες SQL Views, μπορούμε να τις χρησιμοποιήσουμε στο Grafana για να οπτικοποιήσουμε τα δεδομένα μας. Κατά τη δημιουργία ενός πίνακα ελέγχου, μπορούμε να επιλέξουμε την προβολή που δημιουργήσαμε ως πηγή δεδομένων. Αυτό επιτρέπει στο Grafana να επικοινωνήσει με τη βάση δεδομένων μας μέσω των προβολών και να ανακτήσει τα δεδομένα που χρειάζεται για τη δημιουργία των γραφημάτων και των άλλων οπτικών στοιχείων. Στη παρακάτω εικόνα αποτυπώνεται ο πηγαίος κωδικός ενός από τα SQL Views της εφαρμογής μας.

```
SQLQuery4.sql - D...RAMQ1P\Nikos (66)*  SQLQuery3.sql - D...RAMQ1P\Nikos (65)  SQLQuery2.sql - D...RAMQ1P\Nikos (63)  SQLQuery1.sql - D...RAMQ1P\Nikos (56)
USE [Fuel]
GO
/***** Object: View [dbo].[Max_un195_per_month_and_county]    Script Date: 9/9/2023 1:22:06 μμ *****/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER VIEW [dbo].[Max_un195_per_month_and_county]
AS
SELECT
    iso_date,
    unleaded_95,
    county_code,
    name,
    etos,
    month_num,
    Minas
FROM (
    SELECT
        Convert(nvarchar(10), record_date, 103) AS iso_date,
        unleaded_95 as Unleaded_95,
        county_code,
        county_name,
        LEFT(CAST(record_date AS DATE), 4) as etos,
        MONTH(CAST(record_date AS DATE)) as month_num,
        ROW_NUMBER() OVER (PARTITION BY county.name, MONTH(record_date), LEFT(CAST(record_date AS DATE), 4) ORDER BY unleaded_95 DESC) as rank,
        CASE
            WHEN MONTH(CAST(record_date AS DATE)) = 1 THEN 'ΙΑΝΟΥΑΡΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 2 THEN 'ΦΕΒΡΟΥΑΡΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 3 THEN 'ΜΑΡΤΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 4 THEN 'ΑΠΡΙΛΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 5 THEN 'ΜΑΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 6 THEN 'ΙΟΥΝΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 7 THEN 'ΙΟΥΛΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 8 THEN 'ΑΥΓΟΥΣΤΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 9 THEN 'ΣΕΠΤΕΜΒΡΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 10 THEN 'ΟΚΤΩΒΡΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 11 THEN 'ΝΟΕΜΒΡΙΟΣ'
            WHEN MONTH(CAST(record_date AS DATE)) = 12 THEN 'ΔΕΚΕΜΒΡΙΟΣ'
        END as Minas
    FROM dbo.prices
    INNER JOIN county ON county.code = prices.county_code
) AS ranked_prices
WHERE rank = 1
GO
```

Συμπερασματικά, οι SQL Views αποτελούν ισχυρό εργαλείο για την προετοιμασία και την οργάνωση των δεδομένων προς οπτικοποίηση στο Grafana. Με τη δημιουργία προβολών που αντικατοπτρίζουν τα πολύπλοκα ερωτήματα και τις συνδέσεις που απαιτούνται, καθιστούμε τη διαδικασία οπτικοποίησης αποδοτική και ευκολότερη στη συντήρηση. Η χρήση των SQL Views σε συνδυασμό με το Grafana ενισχύει την αναλυτική διαδικασία και παρέχει στους χρήστες ένα περιβάλλον οπτικοποίησης που βασίζεται σε αξιόπιστες πηγές δεδομένων.

4.3 Δημιουργία Dashboards

Τα dashboards αποτελούν το επίκεντρο της οπτικοποίησης δεδομένων στο Grafana. [10] Είναι προσαρμοσμένες σελίδες όπου μπορούμε να οργανώσουμε και να εμφανίσουμε τα γραφήματα, τα πίνακες και τις πληροφορίες που θέλουμε να αναλύσουμε. Σε αυτό το κεφάλαιο, θα εξετάσουμε τη διαδικασία δημιουργίας dashboards στο Grafana και πώς μπορούμε να προσαρμόσουμε τις εμφανίσεις μας για ακριβή ανάλυση δεδομένων.

Βήμα 1: Δημιουργία Dashboard

Για να δημιουργήσουμε ένα νέο dashboard, ακολουθούμε αυτά τα βήματα:

Στην αριστερή πλευρά του παραθύρου περιήγησης, επιλέξτε το "Διαχείριση" και στη συνέχεια "Dashboards".

Επιλέξτε "Δημιουργία Dashboard".

Επιλέξτε το πλαίσιο "Προσθήκη Panel" για να προσθέσετε τα πρώτα σας γραφήματα και πίνακες.

Βήμα 2: Προσθήκη Panels

Κάθε panel αντιπροσωπεύει ένα γράφημα, έναν πίνακα ή μια άλλη παρουσίαση δεδομένων.

Για να προσθέσετε ένα νέο panel:

Επιλέξτε το πλαίσιο "Προσθήκη Panel" στο dashboard.

Επιλέξτε τον τύπο του panel που θέλετε να προσθέσετε, όπως γράφημα χρονοσειράς, πίνακας ή μετρητή.

Στη συνέχεια, μπορείτε να προσαρμόσετε τις ρυθμίσεις του panel, όπως το ερώτημα SQL που θα χρησιμοποιηθεί για τη λήψη των δεδομένων, τον τύπο του γραφήματος και άλλα εμφανισιακά χαρακτηριστικά.

Βήμα 3: Επεξεργασία εμφάνισης

Ένα σημαντικό κομμάτι της δημιουργίας dashboards είναι η επεξεργασία της εμφάνισης για να επιτύχουμε την ακριβή ανάλυση δεδομένων που χρειαζόμαστε. Μπορούμε να προσαρμόσουμε τα εξής στοιχεία:

Τίτλος Dashboard: Προσθέστε έναν περιγραφικό τίτλο που αντικατοπτρίζει το περιεχόμενο του dashboard.

Διάταξη Panels: Αλλάξτε τη θέση των panels για να βρίσκονται στη βέλτιστη διάταξη για την ανάλυση δεδομένων.

Χρήση Filters: Δημιουργήστε φίλτρα που επιτρέπουν στους χρήστες να αλλάζουν δυναμικά τα δεδομένα που εμφανίζονται στο dashboard.

Χρήση Templates: Ορίστε παραμέτρους που μπορούν να παραμετροποιηθούν για να δημιουργήσετε ποικίλες προβολές των ίδιων δεδομένων.

Βήμα 4: Αποθήκευση και Δημοσίευση

Όταν έχουμε ολοκληρώσει τη δημιουργία και την προσαρμογή του dashboard, μπορούμε να το αποθηκεύσουμε και να το δημοσιεύσουμε:

Επιλέξτε το "Αποθήκευση" για να αποθηκεύσετε το dashboard και να δώσετε ένα όνομα.

Επιλέξτε το "Δημοσίευση" για να καταστήσετε το dashboard προσβάσιμο σε άλλους χρήστες του Grafana.

Με τη δημιουργία και την προσαρμογή dashboards, μπορούμε να αξιοποιήσουμε τις δυνατότητες του Grafana για την οπτικοποίηση και ανάλυση των δεδομένων μας. Προσαρμόζοντας τα panels, τις διατάξεις και τις ρυθμίσεις, μπορούμε να δημιουργήσουμε προσαρμοσμένες προβολές που αντικατοπτρίζουν ακριβώς την ανάλυση που απαιτούμε για το έργο μας.

4.4 Παρουσίαση των δεδομένων υγρών καυσίμων

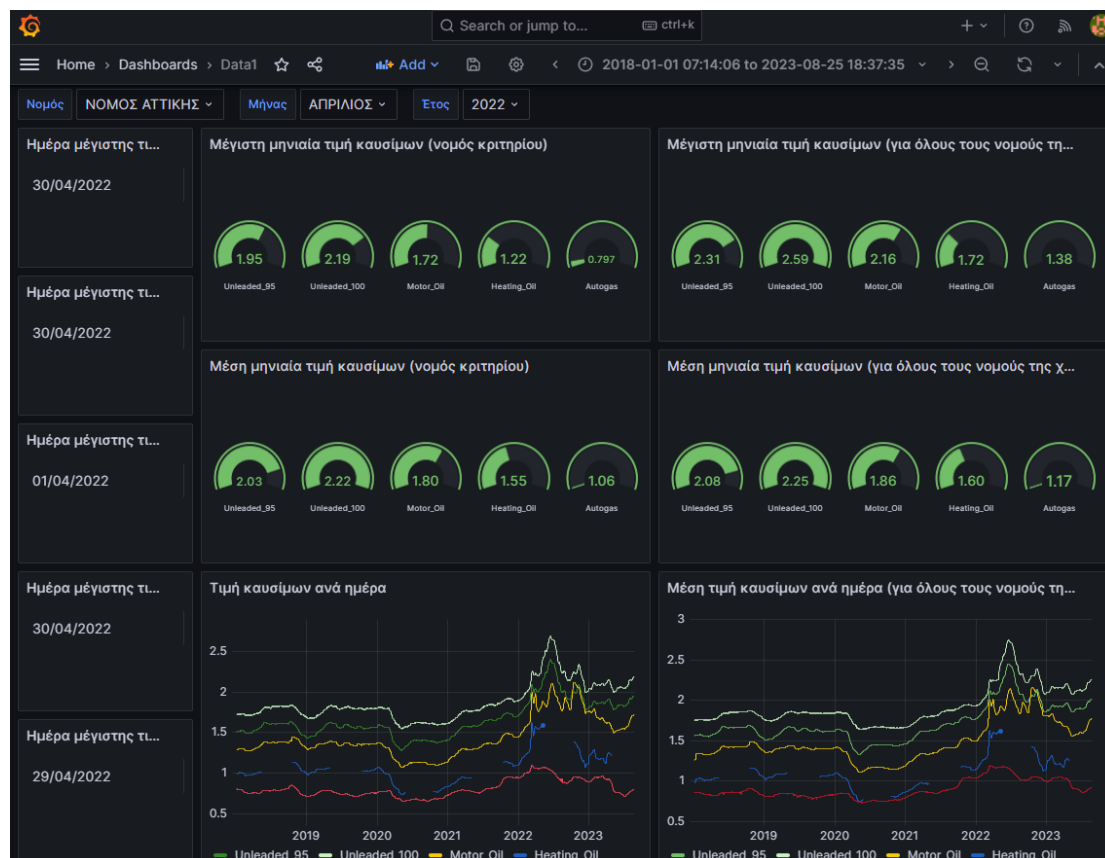
Κεφάλαιο 4.4: Παρουσίαση των Δεδομένων Υγρών Καυσίμων στο Grafana

Σε αυτό το κεφάλαιο, θα εξετάσουμε την παρουσίαση των δεδομένων υγρών καυσίμων μέσω των dashboards στο Grafana. Τα dashboards αποτελούν το κομμάτι του Grafana [11] που επιτρέπει την αναλυτική οπτικοποίηση και παρουσίαση των δεδομένων, επιτρέποντάς μας να παρακολουθούμε την απόδοση, τις τάσεις και τα μοτίβα των υγρών καυσίμων.

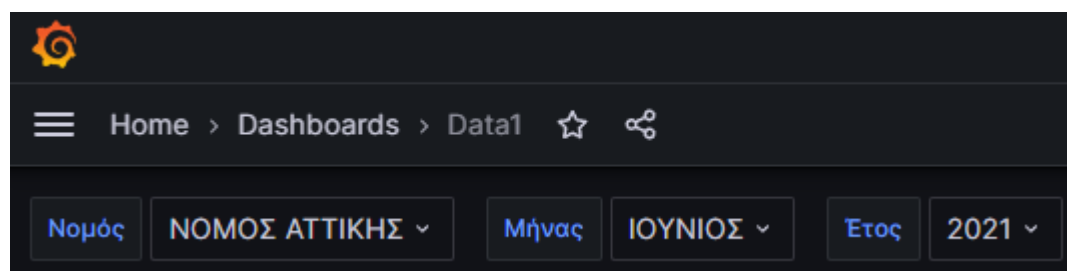
Έχουμε δημιουργήσει ποικίλα dashboards για να καλύψουμε διάφορες πτυχές των δεδομένων υγρών καυσίμων. Κάθε dashboard εστιάζει σε συγκεκριμένες μετρικές και παρέχει ένα

ολοκληρωμένο παράθυρο για την παρακολούθηση. Παρακάτω, παρουσιάζουμε μερικά από τα κυριότερα dashboards που δημιουργήσαμε:

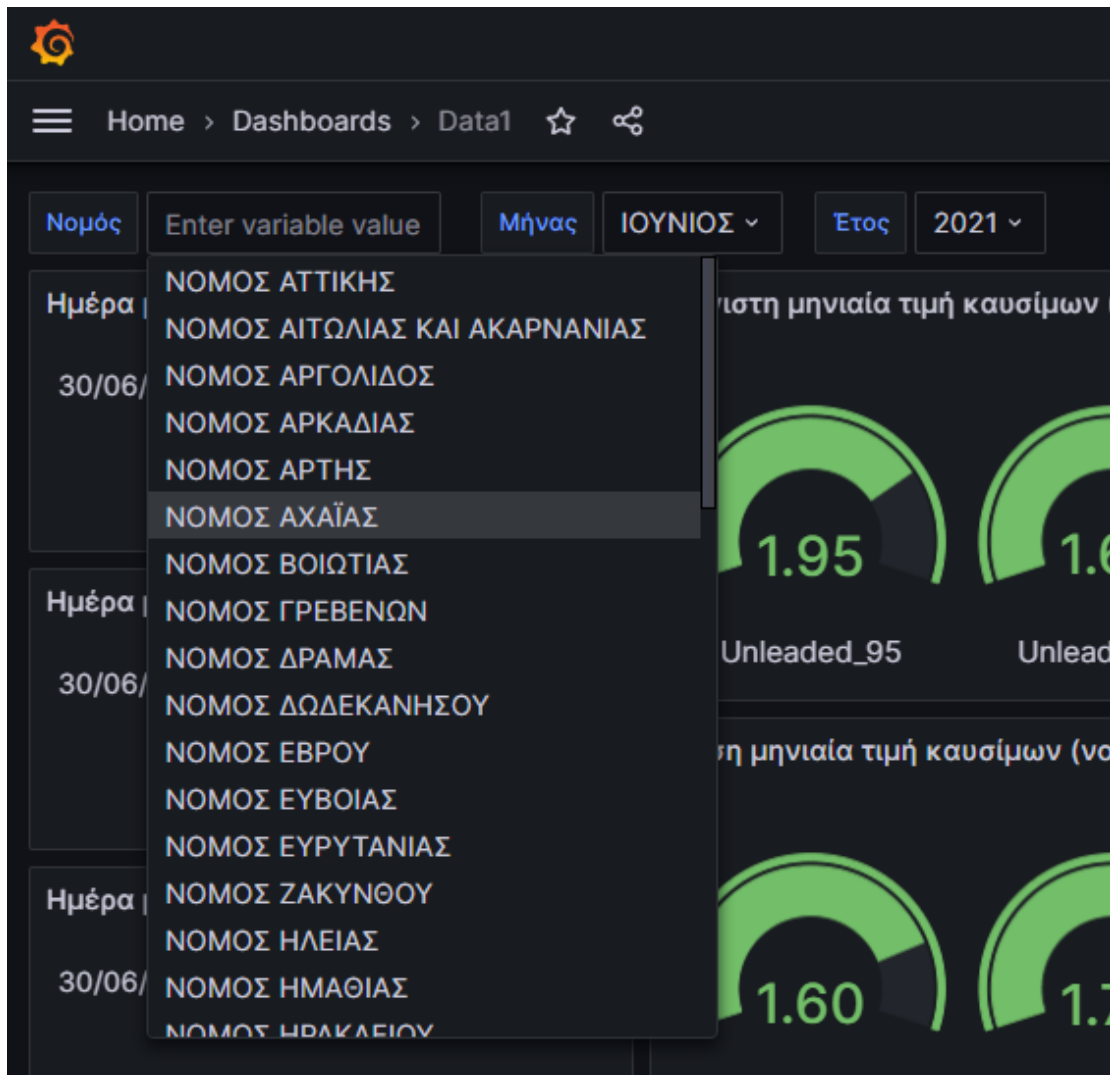
Συνολική εικόνα οπτικοποίησης



Κριτήρια επιλογής δεδομένων. Τα συγκεκριμένα κριτήρια φιλτράρουν τα δεδομένα κάποιων από τα dashboards, ενώ κάποια dashboards εμφανίζουν τα δεδομένα ασχέτως της επιλογής του χρήστη επί των κριτηρίων

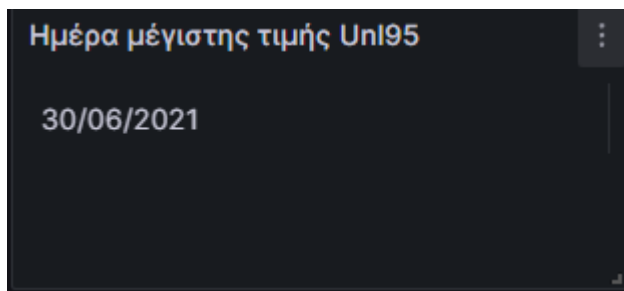


Η επιλογή των επιθυμητών τιμών των κριτηρίων από το χρήστη γίνεται μέσω drop-down menu.

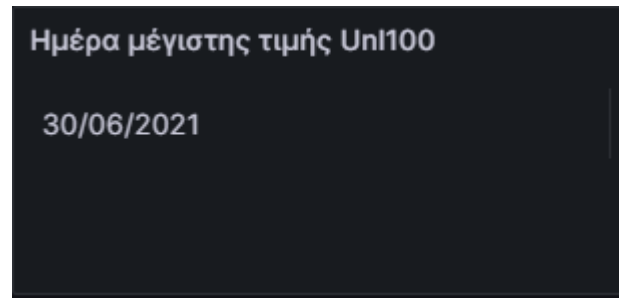


Αναλυτική επεξήγηση των dashboards

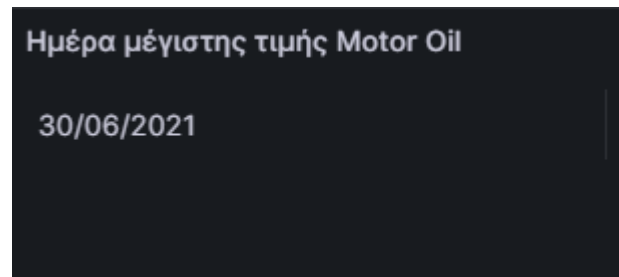
Ποια ημέρα εμφανίστηκε η μέγιστη τιμή για το καύσιμο Unleaded 95, για το μήνα έτος και νομό που έχει επιλεγεί στα κριτήρια



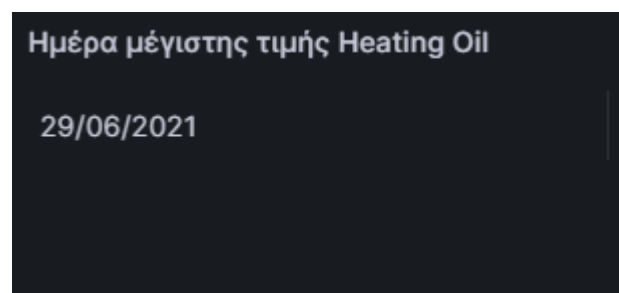
Ποια ημέρα εμφανίστηκε η μέγιστη τιμή για το καύσιμο Unleaded 100, για το μήνα έτος και νομό που έχει επιλεγεί στα κριτήρια



Ποια ημέρα εμφανίστηκε η μέγιστη τιμή για το καύσιμο Motor Oil, για το μήνα έτος και νομό που έχει επιλεγεί στα κριτήρια



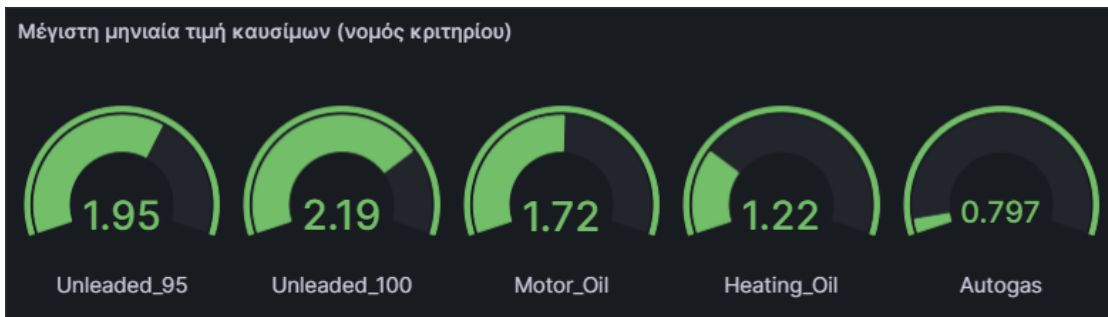
Ποια ημέρα εμφανίστηκε η μέγιστη τιμή για το καύσιμο Heating Oil, για το μήνα έτος και νομό που έχει επιλεγεί στα κριτήρια



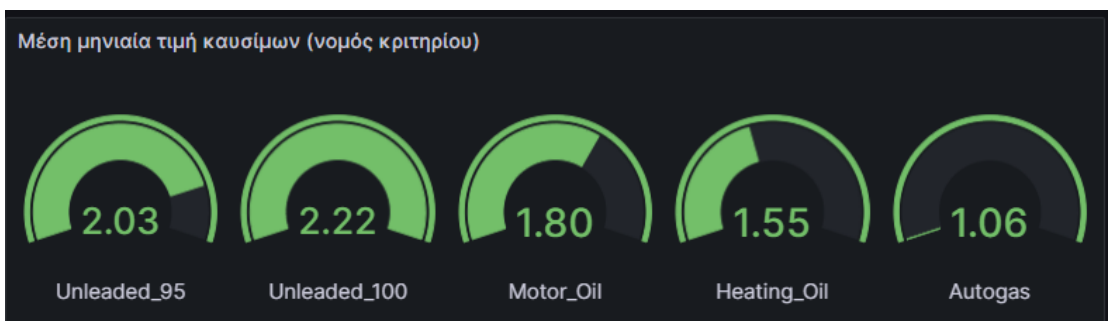
Ποια ημέρα εμφανίστηκε η μέγιστη τιμή για το καύσιμο Autogas, για το μήνα έτος και νομό που έχει επιλεγεί στα κριτήρια

Ημέρα μέγιστης τιμής Autogas
30/06/2021

Η μέγιστη τιμή ανά καύσιμο, για το μήνα, έτος και νομό που επέλεξε ο χρήστης στα κριτήρια



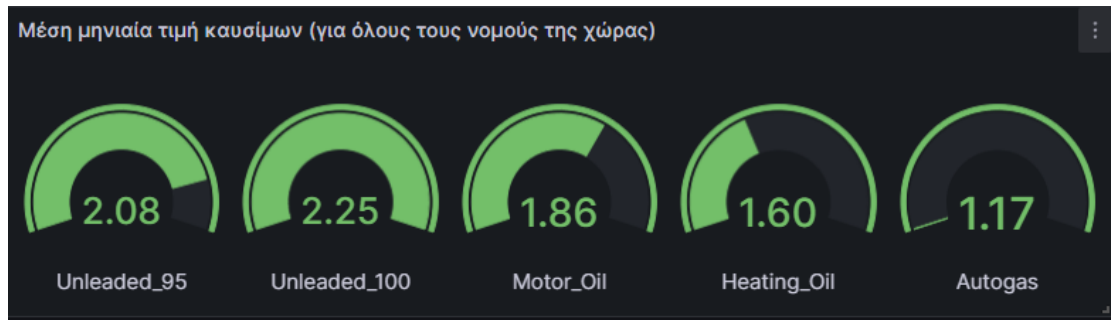
Η μέση τιμή ανά καύσιμο, για το μήνα, έτος και νομό που επέλεξε ο χρήστης στα κριτήρια



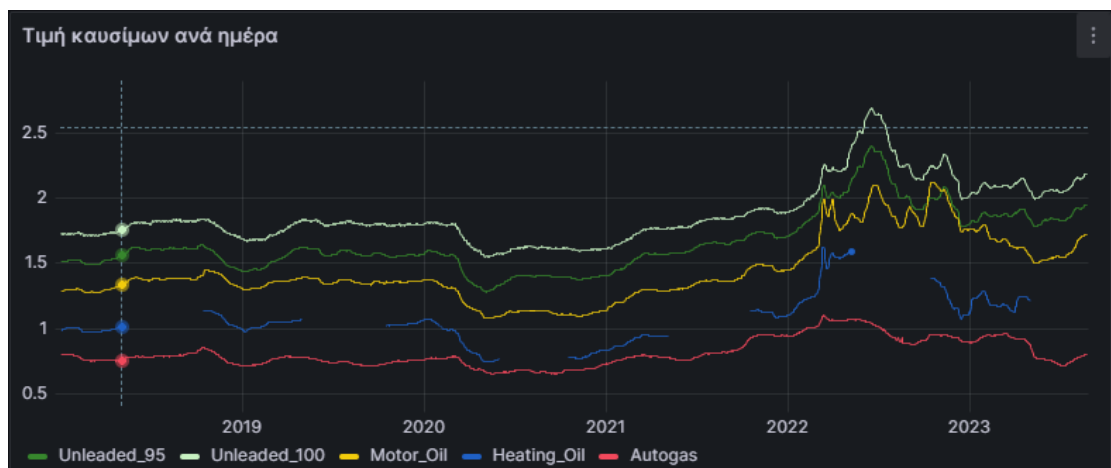
Η μέγιστη τιμή ανά καύσιμο, για το μήνα και έτος που επέλεξε ο χρήστης στα κριτήρια. Η τιμή είναι η μέση τιμή όλων των νομών της χώρας



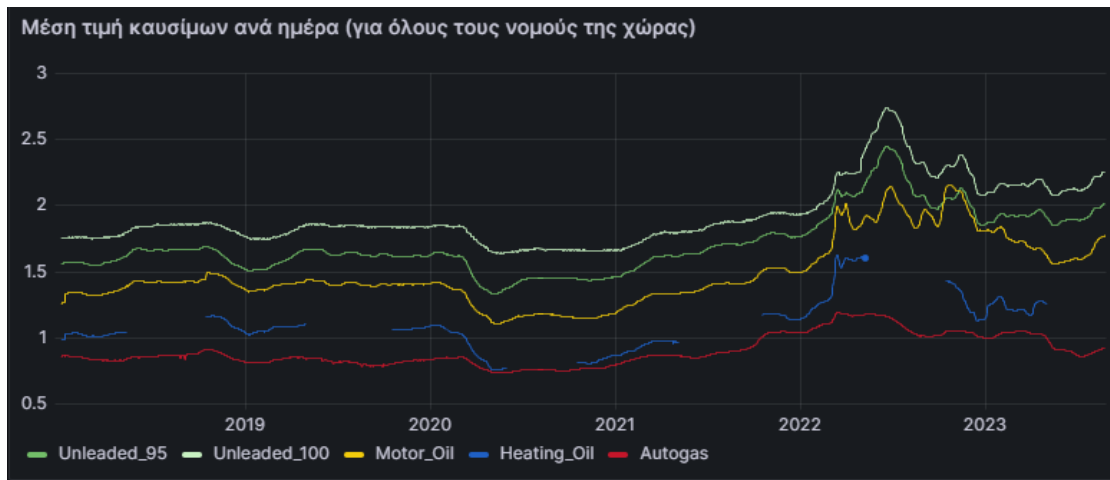
Η μέση τιμή ανά καύσιμο, για το μήνα και έτος που επέλεξε ο χρήστης στα κριτήρια. Η τιμή είναι η μέση τιμή όλων των νομών της χώρας



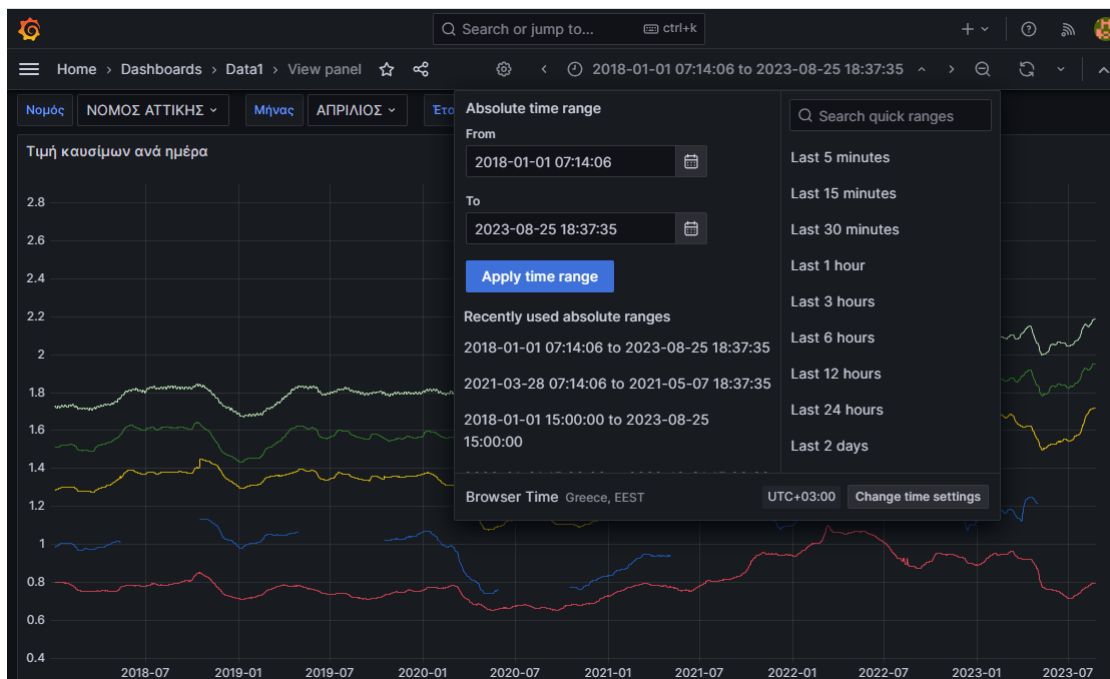
Οι τιμές καυσίμων σε μορφή χρονοσειράς, για το νομό που έχει επιλεγεί στα κριτήρια



Οι τιμές καυσίμων σε μορφή χρονοσειράς, σε πανελλαδικό επίπεδο



Ειδικά για τα δύο dashboards που είναι σε μορφή χρονοσειράς, είναι δυνατή η αλλαγή του χρονικού διαστήματος αναφοράς μέσω του ειδικού drop-down menu του Grafana.



5

Επίλογος

5.1 Συμπεράσματα

Από το κίνητρο της εργασίας προκύπτει η επιτακτική ανάγκη για την ανάπτυξη ενός αυτοματοποιημένου μηχανισμού επεξεργασίας και παρουσίασης δεδομένων υγρών καυσίμων. Οι πολυδιάστατες προκλήσεις της διαχείρισης τέτοιων δεδομένων απαιτούν μια εντελώς ολοκληρωμένη προσέγγιση, συνδυάζοντας την αυτοματοποίηση με την ακρίβεια της ανάλυσης.

Η συνεισφορά αυτής της εργασίας εκδηλώθηκε μέσω της ανάπτυξης ενός ολοκληρωμένου συστήματος επεξεργασίας και οπτικοποίησης δεδομένων, βασιζόμενο σε μια πολυεπίπεδη αρχιτεκτονική. Η χρήση προηγμένων τεχνολογιών όπως το Python για την αυτοματοποίηση, οι βιβλιοθήκες Python για την ανάλυση και την επεξεργασία δεδομένων, το PyCharm για τον ανάπτυξη των σεναρίων, το Microsoft SQL Server για την αποθήκευση και διαχείριση των δεδομένων, και το Grafana για την οπτικοποίηση, δημιούργησαν ένα ενιαίο οικοσύστημα που ανταποκρίνεται στις ανάγκες της εφαρμογής μας.

Κατά την διάρκεια της υλοποίησης, διαπιστώθηκε ότι η αυτοματοποίηση και η οπτικοποίηση δεδομένων διαδραμάτισαν καθοριστικό ρόλο στη βελτιστοποίηση της διαχείρισης υγρών καυσίμων. Τα δεδομένα, όπως παρουσιάζονται από τη σελίδα του υπουργείου, είναι σε μη ωφέλιμη μορφή για εξαγωγή συμπερασμάτων. Μέσω του συγκεκριμένου έργου κατέστη δυνατή τόσο η μετατροπή των δεδομένων σε ωφέλιμη μορφή όσο και η οπτικοποίησή τους, η οποία παρείχε ενισχυμένη κατανόηση και ανάλυση της πληροφορίας. Η συγκεκριμένη προσέγγιση ενθαρρύνει τη δυνητική λήψη ενημερωμένων αποφάσεων και την επίτευξη βέλτιστων αποτελεσμάτων.

Αξίζει επίσης να αναφερθεί ότι η συνεισφορά αυτής της εργασίας δεν περιορίστηκε μόνο στον τομέα της διαχείρισης δεδομένων υγρών καυσίμων, αλλά μπορεί μελλοντικά να εφαρμοστεί ευρύτερα σε διάφορους τομείς της βιομηχανίας και της έρευνας. Οι τεχνολογίες και οι μεθοδολογίες που αναπτύχθηκαν σε αυτήν την εργασία ανοίγουν νέες δυνατότητες για την αξιοποίηση των δεδομένων και την πραγματοποίηση προηγμένων αναλύσεων, συμβάλλοντας στην εξέλιξη των τεχνολογικών καινοτομιών.

Συνολικά, η εργασία αυτή παρουσιάζει μια περιεκτική και συνολική προσέγγιση για τη διαχείριση και την οπτικοποίηση δεδομένων υγρών καυσίμων, προσφέροντας μια συνεκτική λύση που μπορεί να ανταποκριθεί στις προκλήσεις του σύγχρονου περιβάλλοντος. Η συμβολή της εργασίας αυτής στην προώθηση της τεχνολογικής καινοτομίας και τη βελτίωση της ενεργειακής αποδοτικότητας είναι πολύτιμη και έχει ευρύτερες εφαρμογές στον τομέα της βιομηχανίας και της έρευνας.

5.2 Μελλοντικές Επεκτάσεις και Προοπτικές

Η παρούσα εργασία ανέδειξε μόνο το ξεκίνημα μιας ευρύτερης εξέλιξης και ανάπτυξης που μπορεί να εφαρμοστεί στον τομέα της διαχείρισης των δεδομένων των υγρών καυσίμων. Τα επιτεύγματα που επιτεύχθηκαν αποτελούν τη βάση για μελλοντικές επεκτάσεις και προοπτικές που μπορούν να αξιοποιηθούν προκειμένου να ενισχυθεί η αποδοτικότητα, η ακρίβεια και η λειτουργικότητα του συστήματος. Κάποιες πιθανές μελλοντικές επεκτάσεις είναι οι:

1. Πρόβλεψη Τιμών Υγρών Καυσίμων με Χρήση Τεχνητής Νοημοσύνης

Μία πολλά υποσχόμενη μελλοντική επέκταση είναι η χρήση τεχνητής νοημοσύνης, όπου με τη χρήση αλγορίθμων μηχανικής μάθησης θα είναι δυνατή η πρόβλεψη των τιμών των υγρών καυσίμων. Ενδεικτικά, μπορούν να αναπτυχθούν μοντέλα πρόβλεψης που θα βασίζονται στα δομημένα της συγκεκριμένης εργασίας καθώς και σε δεδομένα ιστορικών τιμών, μετεωρολογικά δεδομένα, κοινωνικές και πολιτικές εξελίξεις και άλλες παραμέτρους που επηρεάζουν τις τιμές των καυσίμων. Αυτή η προσέγγιση θα μπορούσε να συνδυαστεί με την αυτοματοποίηση της ανάκτησης δεδομένων και της ενημέρωσης της βάσης δεδομένων, προσφέροντας προβλέψεις που θα συντελούν στη βελτίωση της λήψης αποφάσεων.

2. Δημιουργία Ενός Application Protocol Interface (API) για ευέλικτη διαχείριση των δεδομένων από εφαρμογές τρίτων

Η δημιουργία ενός προσαρμόσιμου και ευέλικτου API αποτελεί μία άλλη ενδιαφέρουσα προοπτική. Το API θα επιτρέπει τη διασύνδεση του συστήματος διαχείρισης δεδομένων υγρών καυσίμων με άλλες εφαρμογές ή συστήματα που απαιτούν πρόσβαση σε αυτά τα δεδομένα. Με αυτόν τον τρόπο, μπορούμε να επιτύχουμε τη διασύνδεση με άλλα συστήματα διαχείρισης, ενισχύοντας την ευελιξία και την αξιοποίηση των δεδομένων σε ποικίλες εφαρμογές.

Σε συνεχή αναζήτηση για τη βελτίωση της αποδοτικότητας και της ακρίβειας των διαδικασιών διαχείρισης δεδομένων υγρών καυσίμων, η επέκταση των λειτουργιών και η εφαρμογή προηγμένων τεχνολογιών αναμένονται να συμβάλουν στην εξέλιξη του συστήματος. Αυτές οι προοπτικές ανοίγουν νέα πεδία έρευνας και ανάπτυξης, προκειμένου να αξιοποιηθούν πλήρως οι δυνατότητες που παρέχει το σύστημα διαχείρισης δεδομένων υγρών καυσίμων και να επιτευχθούν ακόμα πιο εντυπωσιακά αποτελέσματα.

Συνολικά, οι μελλοντικές επεκτάσεις και προοπτικές που αναπτύχθηκαν σε αυτό το κεφάλαιο διαφωτίζουν τον δρόμο για μία αναπτυξιακή πορεία με πολλές δυνατότητες. Ο συνδυασμός των επιστημονικών αρχών και των τεχνολογικών εφαρμογών που παρουσιάστηκαν σε αυτήν την εργασία αποτελεί το θεμέλιο για την επίτευξη αναγκαίων βελτιώσεων και καινοτομιών στον τομέα της διαχείρισης των δεδομένων των τιμών των υγρών καυσίμων.

6

Βιβλιογραφία

- [1] McKinney, W. (2017). Python for data analysis. O'Reilly Media.
- [2] Lutz, M. (2013). Learning Python, 5th Edition. O'Reilly Media.
- [3] Jake VanderPlas (2016). Python Data Science Handbook. O'Reilly Media.
- [4] Ayat Abodayeh, Reem Hejazi, Ward Najjar "Web Scraping for Data Analytics: A BeautifulSoup Implementation." IEEE Sixth International Conference of Women in Data Science at Prince Sultan University (WiDS PSU), 2023.
- [5] Travis E. Oliphant "Python for Scientific Computing" IEEE Computing in Science & Engineering, 2007.
- [6] Milos Radivojevic, Dejan Sarka, William Durkin (2018). Mastering SQL Server 2017. Packt Publishing.
- [7] Itzik Ben-Gan (2012). SQL Server 2012 T-SQL Fundamentals. O'Reilly Media.
- [8] Sikha Saha Bagui, Richard Walsh Earp (2006). Learning SQL on SQL Server 2005. O'Reilly Media.
- [9] Dmitri Korotkevitch (2022). SQL Server Advanced Troubleshooting and Performance Tuning. O'Reilly Media.
- [10] Ronald McCollam (2022). Getting Started with Grafana: Real-Time Dashboards for IT and Business Operations. O'Reilly Media.
- [11] Eric Salituro (2020). Learn Grafana 7.0. O'Reilly Media.

- [12] Rodrigo Juan Hernandez (2022). Building IoT Visualizations using Grafana: Power up your IoT projects and monitor with Prometheus, LibreNMS, and Elasticsearch. Packt Publishing.
- [13] Jaime Buelta (2018). Python Automation Cookbook. O'Reilly Media.