



ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ
ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΨΗΦΙΑΚΗ ΑΦΗΓΗΣΗ ΑΝΟΙΧΤΩΝ
ΔΕΔΟΜΕΝΩΝ ΜΕ ΤΗ ΧΡΗΣΗ
ΕΡΓΑΛΕΙΩΝ ΤΗΣ PLOTLY

Ζαχαρίας Σάμος

Επιβλέπουσα Καθηγήτρια: **Κολωνιάρη Γεωργία**

Θεσσαλονίκη, Οκτώβριος 2019



ΠΕΡΙΕΧΟΜΕΝΑ

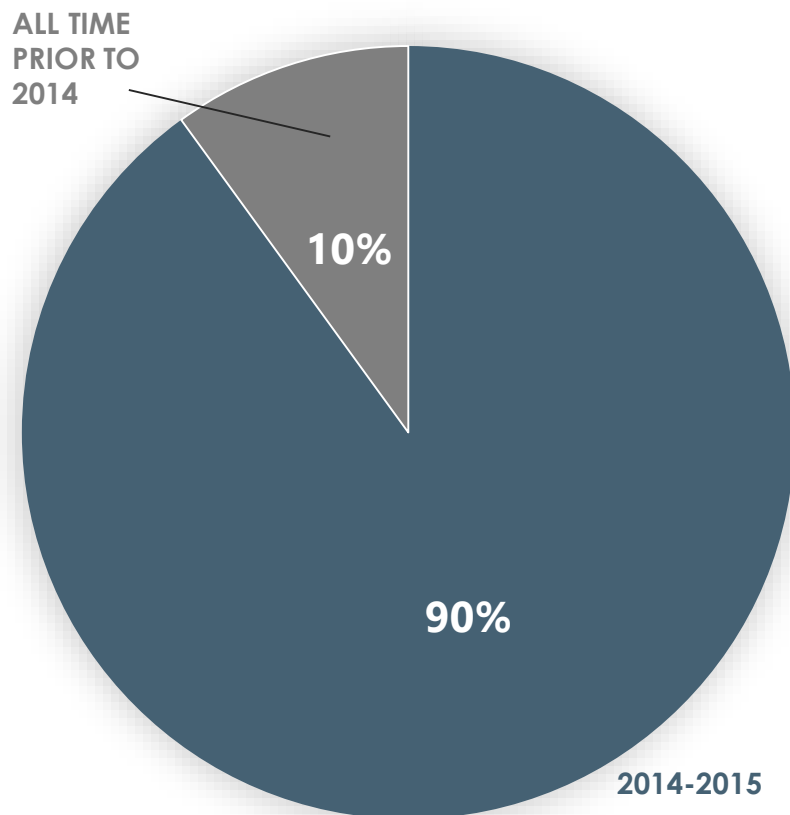
- 1. Εισαγωγή**
- 2. Θεωρητικό πλαίσιο**
 - 2.1 Data visualization
 - 2.2 Data storytelling
- 3. Plotly Chart studio**
- 4. Μελέτη περίπτωσης**
- 5. Συμπεράσματα**

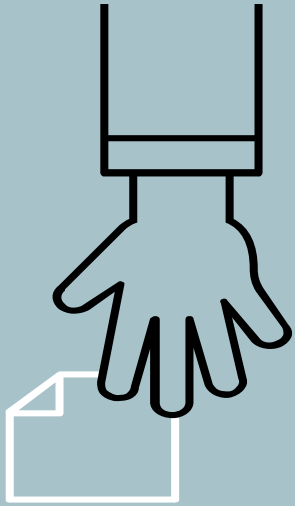


1. ΕΙΣΑΓΩΓΗ

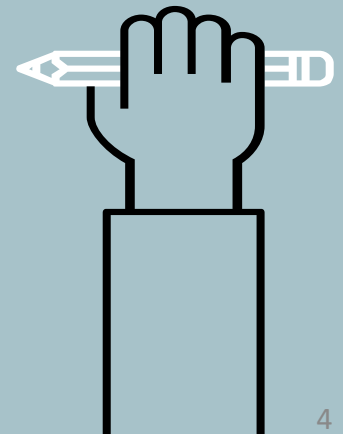
- ❖ Τα τελευταία χρόνια υπάρχει μια έκρηξη στην παραγωγή δεδομένων
- ❖ Οργανισμοί και επιχειρήσεις εξερευνούν τα δεδομένα προκειμένου να παρθούν καλύτερες αποφάσεις
- ❖ Αυτή η εξερεύνηση περιλαμβάνει την οπτικοποίηση των δεδομένων ή **“data visualization”**
- ❖ Τα αποτελέσματα από την οπτικοποίηση μεταφέρονται προς τρίτους μέσω αφήγησης **“data storytelling”**

GROWTH OF DATA IN THE WORLD





2. ΘΕΩΡΗΤΙΚΟ ΠΛΑΙΣΙΟ



2.1 Data visualization

Η οπτικοποίηση δεδομένων είναι η διαδικασία χαρτογράφησης τιμών ή πληροφοριών σε οπτικό περιεχόμενο. Η οπτικοποίηση δεδομένων μπορεί να φανερώσει τάσεις που διαφορετικά είναι δύσκολο να αναγνωριστούν

Γιατί είναι σημαντική;

- ❖ Προσφέρει μεγαλύτερη πυκνότητα στην πληροφορία
- ❖ Επεξεργάζεται πιο ευκολά από τον ανθρώπινο εγκέφαλο συγκριτικά με απλές λέξεις/αριθμούς
- ❖ Είναι πιο εύκολο να αποτυπωθούν στην ανθρώπινη μνήμη

2.1 Data visualization

Static vs Interactive

Static

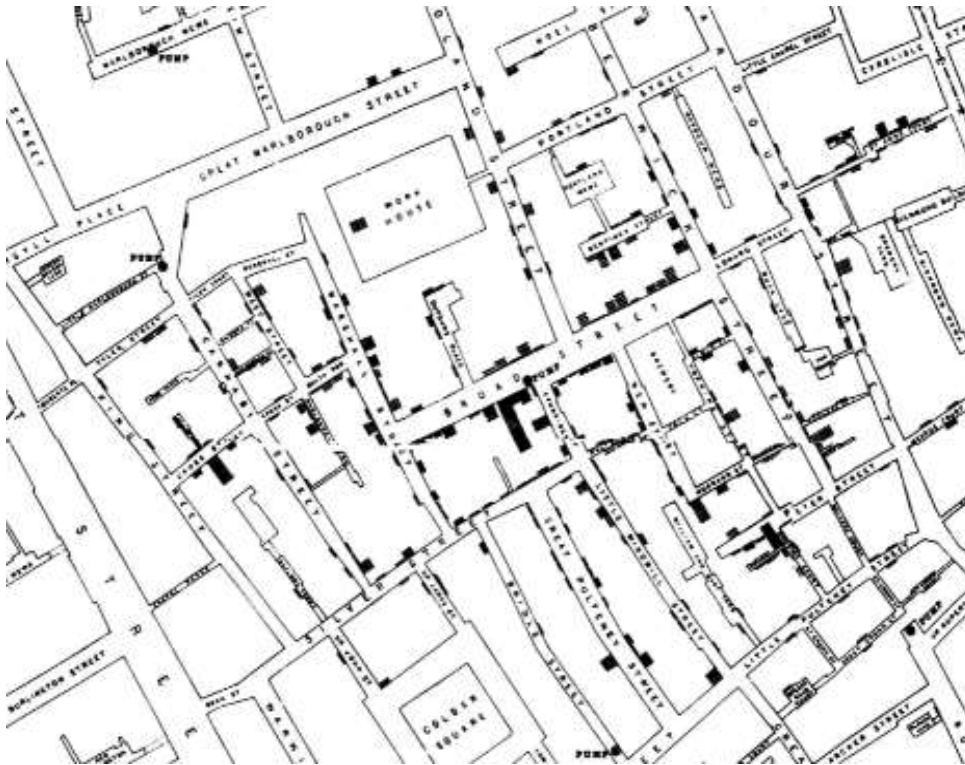
- ❖ Είναι ο πιο απλός τρόπος παρουσίασης των δεδομένων
- ❖ Χρησιμοποιούνται εικόνες συνοδευόμενες από κείμενο
- ❖ Δεν απαιτούνται δεδομένα εισόδου από τον χρήστη

Interactive

- ❖ Περιέχουν ελέγχους όπως ζουμ και φιλτράρισμα που δίνουν πολύ περισσότερες επιλογές
- ❖ Απαιτούν δεδομένα εισόδου από τους χρήστες
- ❖ Ευνοούν την παρουσίαση των αποτελεσμάτων από μεγάλα σύνολα δεδομένων

2.1 Data visualization

Σημαντικές οπτικοποιήσεις του παρελθόντος

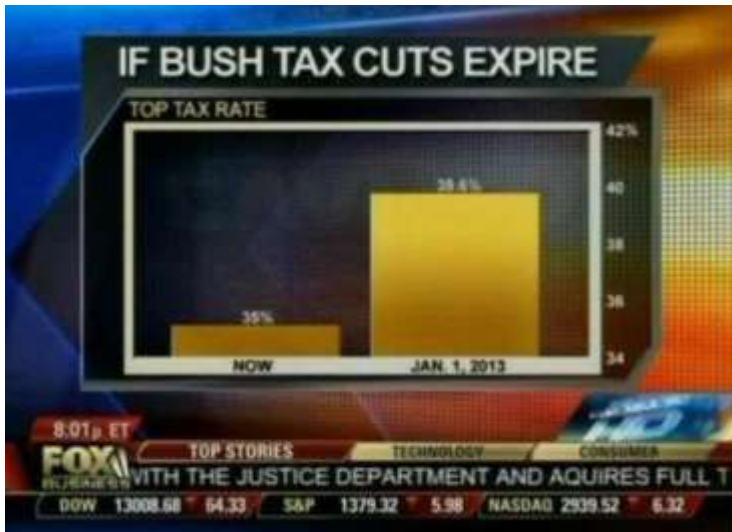


- ❖ Ο John Snow βρήκε τον λόγο εξάπλωσης της επιδημίας της χολέρας χάρη σε έναν χάρτη που έφτιαξε
- ❖ Ο χάρτης κατέγραφε την τοποθεσία των θανάτων τον βοήθησε να εντοπίσει την πηγή της επιδημίας: την δημόσια αντλία νερού στην Broad Street.
- ❖ *Μια ουσιώδης ανωμαλία στα αποτελέσματα είναι ότι στην κοντινή ζυθοποιία δεν υπήρξε κανένας θάνατος, που οφείλεται στο ότι οι εργάτες είχαν ελεύθερη κατανάλωση μπίρας και η διαδικασία της ζύμωσης σκοτώνει το βακτήριο της χολέρας*

2.1 Data visualization

Αντικειμενικότητα των οπτικοποιήσεων

Η απεικόνιση των δεδομένων θα πρέπει να γίνεται όσο το δυνατόν πιο αντικειμενικά



2.2 Data storytelling

- ❖ Η αφήγηση δεδομένων μπορεί να οριστεί ως η μεθοδολογία για την επικοινωνία πληροφοριών, προσαρμοσμένη σε ένα συγκεκριμένο κοινό. Αποτελεί το τελευταίο βήμα της ανάλυσης δεδομένων και είναι αναμφισβήτητα σημαντικό
- ❖ Η αφήγηση σε συνδυασμό με την οπτικοποίηση δεδομένων μπορούν να επικοινωνήσουν αποδοτικά ιδέες και πληροφορίες

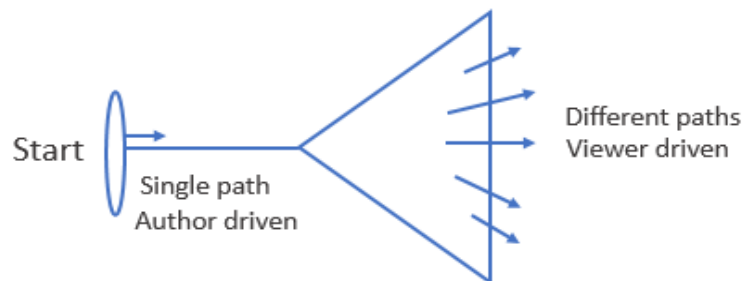
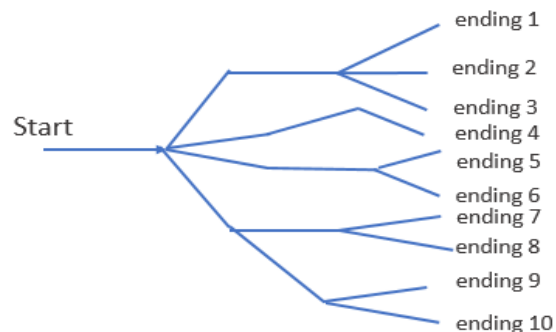
Οι ιστορίες:

- ❖ Ελκύουν την προσοχή του κοινού
- ❖ Μεταφέρουν αποδοτικά ιδέες
- ❖ Ενισχύουν την πλήρη κατανόηση των νέων πληροφοριών
- ❖ Έχουν ισχυρό αντίκτυπο στο κοινό που δεν είναι εφικτό να επιτευχθεί μόνο με δεδομένα
- ❖ Συχνά μεταδίδονται και επαναλαμβάνονται από άτομο σε άτομο

2.2 Data storytelling

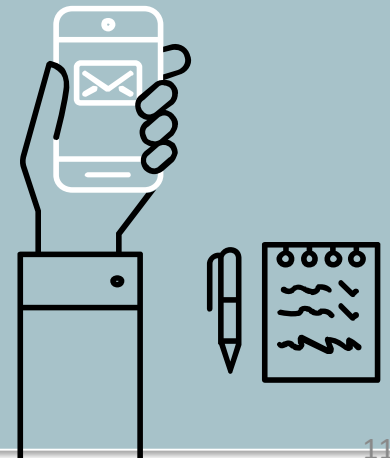
Διαφορετικοί τύποι δομών ιστορίας

- ❖ Αφήγηση του συγγραφέα
 - Προχωρεί με μια γραμμική ροή που ορίζει ο συγγραφέας
- ❖ Αφήγηση του θεατή
 - Επιτρέπει στον θεατή να επιλέξει την κατεύθυνση στην οποία προχωρεί η αφήγηση
- ❖ Αφήγηση Martini glass
 - Συνδυάζει τις δύο παραπάνω δομές





3. Plotly Chart studio



3 Plotly Chart studio

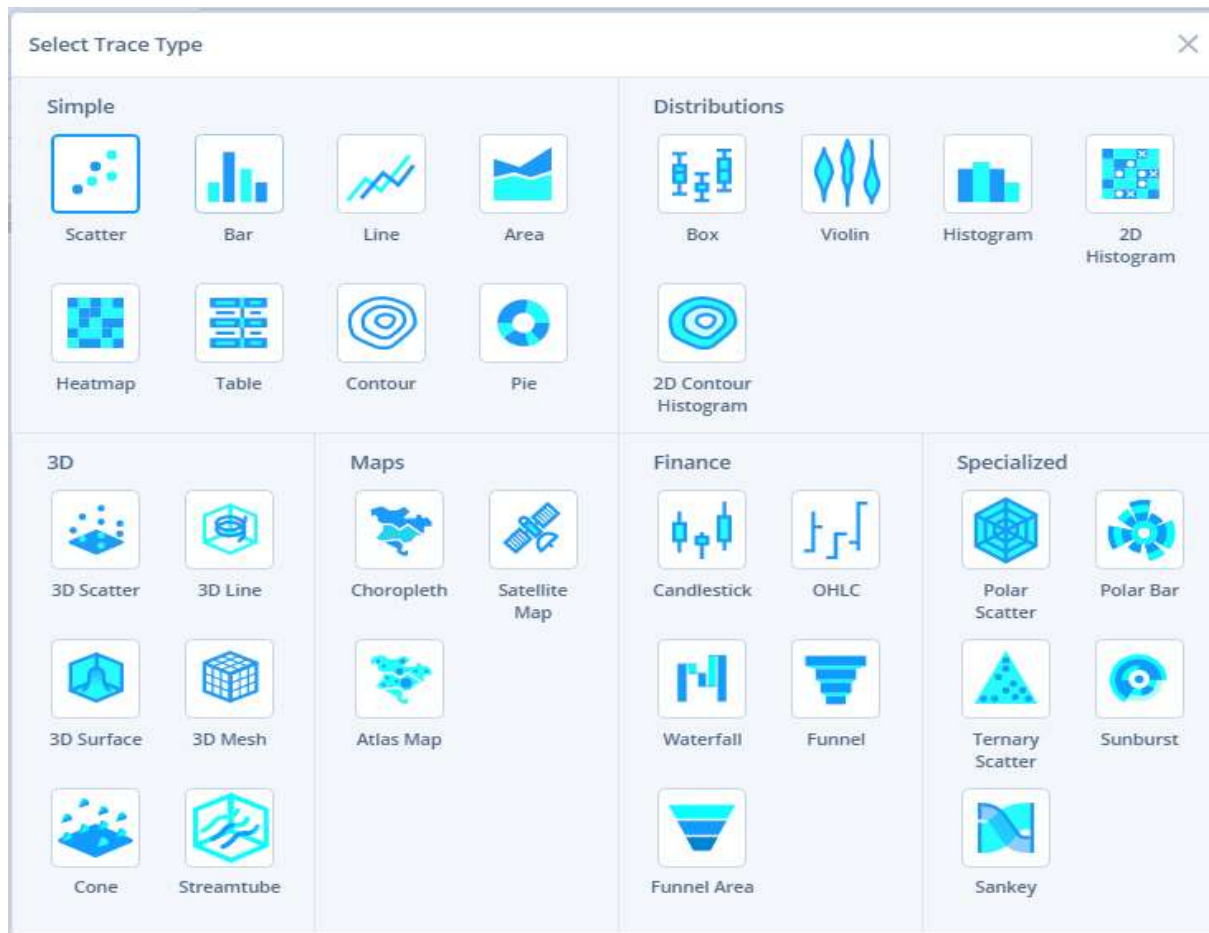
Σκοπός: Η δημιουργία ενός case study μέσω της ανάπτυξης μιας digital data story η οποία υλοποιήθηκε στο εργαλείο **Plotly Chart studio**

Το συγκεκριμένο εργαλείο:

- ❖ Υλοποιεί ένα μεγάλο πλήθος διαγραμμάτων
- ❖ Συνδυάζει τα διαγράμματα σε dashboards
- ❖ Προσφέρει μεγάλες δυνατότητες διαδραστικότητας
- ❖ Δεν χρειάζεται γνώσεις προγραμματισμού
- ❖ Δημιουργεί διαγράμματα έτοιμα να εισαχθούν σε ιστοσελίδες
- ❖ Προσφέρει ευκολία στην είσοδο δεδομένων π.χ. Import from excel, connect to SQL, drag and drop
- ❖ Είναι δωρεάν (η βασική έκδοση)

3 Plotly Chart studio

Το Plotly Chart studio προσφέρει ένα ευρύ φάσμα διαγραμμάτων



3 Plotly Chart studio

Εισαγωγή δεδομένων

Υπάρχουν οι παρακάτω επιλογές για τον χρήστη:

- ❖ Πληκτρολόγηση των δεδομένων απευθείας στα κελιά ή αντιγραφή από ένα υπολογιστικό φύλλο και επικόλληση
- ❖ Εισαγωγή αρχείου δεδομένων (.xls, .xlsx ή .csv). Μπορεί να χρησιμοποιηθεί και «drag and drop»
- ❖ Πληκτρολόγηση της διεύθυνσης URL ενός ηλεκτρονικού συνόλου δεδομένων
- ❖ Σύνδεση σε μια βάση δεδομένων SQL

3 Plotly Chart studio

Διαδραστικότητα



- ❖ Pan
- ❖ Zoom in – Zoom out
- ❖ Box select και Lasso select
 - Για να φιλτράρονται τα δεδομένα επιλέγοντας απλά μια περιοχή του γραφήματος
- ❖ Spikelines
 - Όταν ο δρομέας βρίσκεται πάνω από ένα σημείο δεδομένων, θα εμφανιστούν γραμμές που δείχνουν τις τομές με τους άξονες
- ❖ Hover options
 - Όταν ο δρομέας βρίσκεται πάνω από ένα σημείο δεδομένων, μπορούν να προβληθούν είτε οι τιμές μόνο για το συγκεκριμένο σημείο, είτε οι τιμές για όλα τα σημεία με την ίδια τιμή του x-άξονα
- ❖ Δυνατότητα φιλτραρίσματος των 'traces'

3 Plotly Chart studio

Style

Αφού δημιουργηθούν τα γραφήματα, υπάρχουν πολλές λειτουργίες που βοηθούν στο στυλιζάρισμα

❖ Themes

- 6 ενσωματωμένα διαθέσιμα θέματα στην ελεύθερη έκδοση

❖ General tab

- Επεξεργασία γενικών χαρακτηριστικών όπως χρώμα φόντου, margins, sizes κλπ.

❖ Traces

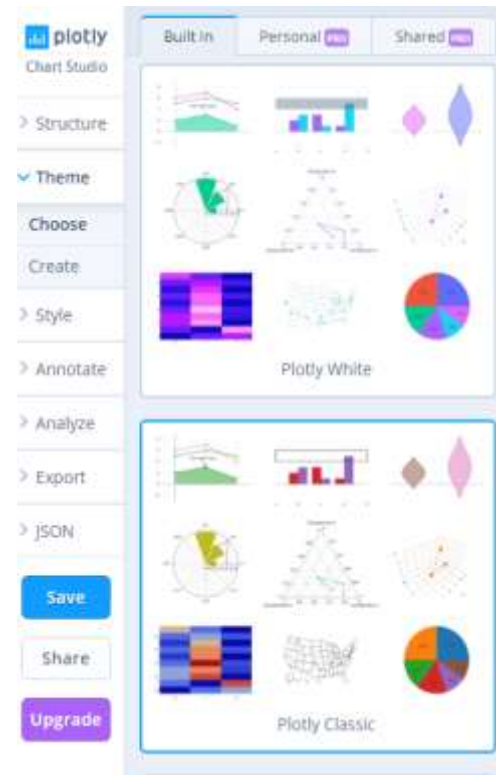
- Αλλαγή χρώματος, τύπου/πάχους γραμμών, αλλαγή συμβόλων, κλίμακας χρωμάτων κλπ.

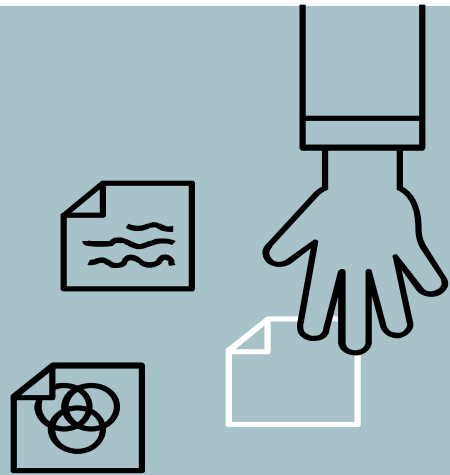
❖ Axes

- Ορισμός των ελάχιστων και μέγιστων τιμών, της μορφής της ετικέτας του άξονα κλπ.

❖ Legend

- Επιλογές που σχετίζεται με το κείμενο, το χρώμα και το μέγεθος γραμματοσειράς των ετικετών





4. CASE STUDY



4 Case study

Περιγραφή Θέματος

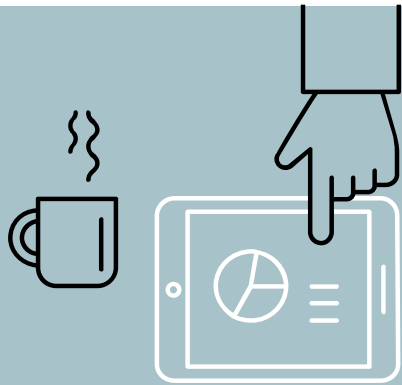
Οι καταγεγραμμένες πωλήσεις των επιβατικών αυτοκινήτων και των χαρακτηριστικών τους τα τελευταία 8 έτη στην Ευρώπη

- ❖ Δεδομένα από τον Ευρωπαϊκό οργανισμό περιβάλλοντος
 - Τα κράτη μέλη είναι υποχρεωμένα να καταγράφουν πληροφορίες για κάθε καινούριο όχημα που εισέρχεται στον στόλο τους και να υποβάλλουν τα δεδομένα στην Commission

| Field name | Field Definition | Data type | Primary key |
|-----------------------|---|--------------|-------------|
| ID | Identification number | integer | Yes |
| MS | Member state | varchar(2) | No |
| Mp | Manufacturer pooling | varchar(50) | No |
| VFN | Vehicle family identification number | varchar(25) | No |
| Mh | Manufacturer name EU standard denomination | varchar(50) | No |
| Man | Manufacturer name OEM declaration | varchar(50) | No |
| MMS | Manufacturer name MS registry denomination | varchar(125) | No |
| TAN | Type approval number | varchar(50) | No |
| T | Type | varchar(25) | No |
| Va | Variant | varchar(25) | No |
| Ve | Version | varchar(35) | No |
| Mk | Make | varchar(25) | No |
| Cn | Commercial name | varchar(50) | No |
| Ct | Category of the vehicle type approved | varchar(5) | No |
| Cr | Category of the vehicle registered | varchar(5) | No |
| m (kg) | Mass in running order Completed/complete vehicle | integer | No |
| Mt | WLTP test mass | integer | No |
| Enedc (g/km) | Specific CO ₂ Emissions (NEDC) | integer | No |
| Ewltp (g/km) | Specific CO ₂ Emissions (WLTP) | integer | No |
| W (mm) | Wheel Base | integer | No |
| At1 (mm) | Axle width steering axle | integer | No |
| At2 (mm) | Axle width other axle | integer | No |
| Ft | Fuel type | varchar(25) | No |
| Fm | Fuel mode | varchar(1) | No |
| ec (cm ³) | Engine capacity | integer | No |
| ep (KW) | Engine power | integer | No |
| z (Wh/km) | Electric energy consumption | integer | No |
| IT | Innovative technology or group of innovative technologies | varchar(25) | No |

4 Case study

<https://chart-studio.plot.ly/dashboard/samoszach:57/present#/>

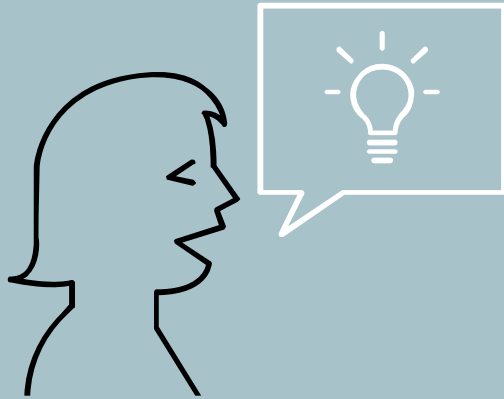


5. ΣΥΜΠΕΡΑΣΜΑΤΑ



5 Συμπεράσματα

- ❖ Το Plotly Chart studio καλύπτει ένα πολύ μεγάλο αριθμό γραφημάτων
- ❖ Συνδυάζει πολλά επιθυμητά χαρακτηριστικά διαδραστικότητας
- ❖ Από το case study έγινε εμφανές ότι παρέχει ένα πολύ φιλικό προς τον χρήστη interface
- ❖ Το ότι παρέχεται δωρεάν του δίνει ένα σημαντικό πλεονέκτημα έναντι άλλων εργαλείων
- ❖ Το Plotly Community forum προσφέρει απαντήσεις γρήγορα όποτε χρειαστεί
- ❖ Μειονέκτημα είναι ότι δεν μπορεί να εξαχθεί ο «κώδικας» πίσω από τα διαγράμματα
- ❖ Λείπει η δυνατότητα να κατασκευαστούν 'animation' διαγράμματα



Ευχαριστώ για την προσοχή σας!

