

Θέμα covid-19

Το σύνολο δεδομένων [Novel Corona Virus 2019 Dataset¹](#) ενημερώνεται καθημερινά από τον Παγκόσμιο Οργανισμό Υγείας και περιέχει για κάθε χώρα και κάθε μέρα το πλήθος των κρουσμάτων, των θανάτων και των ιαθέντων **μέχρι και εκείνη την ημέρα**. Δηλαδή τα κρούσματα μιας ημέρας δεν αναφέρονται μόνο σε αυτή την ημέρα αλλά στο άθροισμα όλων των κρουσμάτων **μέχρι και** εκείνη την ημέρα. Επίσης για κάποιες χώρες υπάρχουν περισσότερες από μια εγγραφές για την ίδια μέρα γιατί οι αριθμοί αναφέρονται σε συγκεκριμένες επαρχίες (Κίνα) ή πολιτείες (ΗΠΑ). Ένα ενδεικτικό παράδειγμα με ένα μικρό υποσύνολο των δεδομένων δίνεται παρακάτω:

ΗΜ		ΠΕΡΙΟΧΗ		Δ			ΧΩΡΑ	
27	5	Qinghai	1	18	0	18	1	Mainland China
27	5	Liaoning	1	149	2	147	2	US
27	5	Iowa	2	18369	496	0	3	Spain
27	5	La Rioja	3	4044	360	3107	4	UK
27	5	Hubei	1	68135	4512	63618		.
27	5	Hunan	1	1019	4	1015		.
27	5	Michigan	2	55608	5334	0		.
28	5	Liaoning	1	149	2	147		.
28	5	Iowa	2	18586	506	0		.
28	5	Michigan	2	56014	5372	0	100	

Τα δεδομένα (N=30000) δίνονται κατά χρονολογική σειρά από την 22/01/2020 μέχρι και 30/05/2020, δηλαδή 130 ημέρες και στη ΓΛΩΣΣΑ καταχωρούνται στους εξής πίνακες:

- ΗΜ[N,2]: στην πρώτη στήλη η ημέρα και στην δεύτερη ο μήνας.
- ΠΕΡΙΟΧΗ[N, 2]: στην πρώτη στήλη η περιοχή και στη δεύτερη ο μοναδικός αριθμός της χώρας (όχι το όνομα) που είναι από 1 έως και 100.
- Δ[N, 3]: στην πρώτη στήλη τα κρούσματα, στην δεύτερη οι θάνατοι και στην τρίτη οι ιαθέντες αθροιστικά μέχρι και εκείνη τη μέρα.
- ΧΩΡΑ[100]: Η χώρα με αριθμό 3 έχει όνομα ΧΩΡΑ[3] κλπ. Θεωρούμε ότι έχουμε δεδομένα από 100 χώρες.

¹ https://www.kaggle.com/sudalairajkumar/novel-corona-virus-2019-dataset?select=covid_19_data.csv

Να γράψετε πρόγραμμα το οποίο

- 1) Καλεί την διαδικασία Εισαγωγή_Δεδομένων η οποία διαβάζει τα δεδομένα και τα καταχωρεί στους παραπάνω πίνακες. Θεωρήστε ότι η διαδικασία έχει ήδη υλοποιηθεί άρα μπορείτε να την χρησιμοποιήσετε.
- 2) Δημιουργεί έναν πίνακα Π ο οποίος περιέχει για κάθε μέρα και για κάθε χώρα το πλήθος των κρουσμάτων αθροιστικά μέχρι εκείνη την ημέρα.
- 3) Δημιουργεί έναν πίνακα ΣΗΜ ο οποίος περιέχει για κάθε μέρα και κάθε χώρα το πλήθος των κρουσμάτων που εντοπίστηκαν μόνο εκείνη την ημέρα. Δηλαδή περιέχει το πλήθος των κρουσμάτων ανά ημέρα και όχι αθροιστικά. Παρακάτω δίνεται ένα παράδειγμα μετατροπής των δεδομένων ανά ημέρα στις περιοχές Liaoning, Iowa, Michigan. (δεν είναι ανά χώρα).

HM		ΠΕΡΙΟΧΗ		ΣΗΜ		
28	5	Liaoning	1	0	0	0
28	5	Iowa	2	217	10	0
28	5	Michigan	2	406	38	0

Για την ιστορία, παρακάτω βλέπετε ένα στιγμιότυπο των δεδομένων όπως είναι καταχωρημένα στο πραγματικό αρχείο:

ObservationDate	Province/State	Country/Region	Confirmed	Deaths	Recovered
05/27/2020	Qinghai	Mainland China	18	0	18
05/27/2020	Liaoning	Mainland China	149	2	147
05/27/2020	Iowa	US	18369	496	0
05/27/2020	La Rioja	Spain	4044	360	3107
05/27/2020	Hubei	Mainland China	68135	4512	63618
05/27/2020	Hunan	Mainland China	1019	4	1015
05/27/2020	Michigan	US	55608	5334	0
05/28/2020	Liaoning	Mainland China	149	2	147
05/28/2020	Iowa	US	18586	506	0
05/28/2020	Michigan	US	56014	5372	0