Relazione Dati

### Intro

Il focus per il cliente è ricavare dei segnali per la creazione di un nuovo vaccino anti-covid. Con questo obbiettivo abbiamo scelto come fonte dati il Dipartimento della Protezione Civile. I dati sono stati recuperati sul repository proprietario del [Dipartimento](https://github.com/pcm-dpc) in formato CSV.

### Dati raw

1. [*covid-19-ita-regioni.csv*](https://github.com/pcm-dpc/COVID-19/blob/master/dati-regioni/dpc-covid19-ita-regioni.csv)

Il dataset contiene i dati inerenti alla situazione COVID per regione in Italia, raccolti in modo giornaliero dal 2020-02-24.

#### Campi raw

|  |  |  |
| --- | --- | --- |
| Campo | Descrizione | Tipo Dato |
| data | Data dell'informazione | YYYY-MM-DDTHH:MM:SS (ISO 8601) Ora italiana |
| stato | Stato di riferimento | XYZ (ISO 3166-1 alpha-3) |
| codice\_regione | Codice della Regione (ISTAT 2019) | Numero |
| denominazione\_regione | Denominazione della Regione | Testo |
| lat | Latitudine | WGS84 |
| long | Longitudine | WGS84 |
| ricoverati\_con\_sintomi | Ricoverati con sintomi | Numero |
| terapia\_intensiva | Ricoverati in terapia intensiva | Numero |
| totale\_ospedalizzati | Totale ospedalizzati | Numero |
| isolamento\_domiciliare | Persone in isolamento domiciliare | Numero |
| totale\_positivi | Totale attualmente positivi (ospedalizzati + isolamento domiciliare) | Numero |
| variazione\_totale\_positivi | Variazione del totale positivi (totale\_positivi giorno corrente - totale\_positivi giorno precedente) | Numero |
| nuovi\_positivi | Nuovi attualmente positivi (totale\_casi giorno corrente - totale\_casi giorno precedente) | Numero |
| dimessi\_guariti | Persone dimesse guarite | Numero |
| deceduti | Persone decedute | Numero |
| casi\_da\_sospetto\_diagnostico | Casi positivi al tampone emersi da attività clinica **Non più popolato** | Numero |
| casi\_da\_screening | Casi positivi emersi da indagini e test, pianificati a livello nazionale o regionale **Non più popolato** | Numero |
| totale\_casi | Totale casi positivi | Numero |
| tamponi | Totale tamponi (processati con test molecolari) | Numero |
| casi\_testati | Totale dei soggetti sottoposti al test | Numero |
| note | Note | Testo |
| ingressi\_terapia\_intensiva | Ingressi giornalieri in terapia intensiva | Numero |
| note\_test | Note sui test effettuati | Testo |
| note\_casi | Note sui casi testati | Testo |

#### Classificazione

I dati in formato CSV hanno una classificazione .

1. [*somministrazioni-vaccini-latest.csv*](https://github.com/italia/covid19-opendata-vaccini/blob/master/dati/somministrazioni-vaccini-latest.csv)

Il dataset contiene tutti i dati raccolti in modo giornaliero dal 2020-12-27 riguardo al processo di vaccinazione per le diverse fasce di età e di richiamo.

#### Campi raw

|  |  |  |
| --- | --- | --- |
| Campo | Descrizione | Tipo Dato |
| area | Sigla della regione di consegna. | Testo |
| forn | nome completo del fornitore del vaccino. | Testo |
| data | Data di somministrazione. | YYYY-MM-DDTHH:MM:SS (ISO 8601) Ora italiana |
| eta | Identifica la fascia anagrafica alla quale appartengono i soggetti a cui è stato somministrato il vaccino. | Testo |
| m | Totale dei soggetti di sesso maschile a cui è stato somministrato il vaccino per giorno, regione e fascia d’età. | Numero |
| f | Totale dei soggetti di sesso femminile a cui è stato somministrato il vaccino per giorno, regione e fascia d’età. | Numero |
| d1 | Numero prime somministrazioni. | Numero |
| d2 | Numero seconde somministrazioni. | Numero |
| dpi | Numero di somministrazioni effettuate a soggetti con pregressa infezione da covid-19 nel periodo 3-6 mesi e che, pertanto, concludono il ciclo vaccinale con un'unica dose | Numero |
| db1 | Numero somministrazioni dose addizionale/richiamo (booster). | Numero |
| dbi | Numero somministrazioni dose booster a soggetti che hanno già completato il ciclo vaccinale primario con tre dosi (di cui la terza addizionale), dopo un intervallo minimo di almeno quattro mesi (120 giorni) dalla dose addizionale stessa. | Numero |
| db2 | Numero di dosi di second booster somministrate a soggetti che abbiano già completato il ciclo vaccinale comprensivo di dose booster, dopo un intervallo minimo di almeno quattro mesi (120 giorni) dalla dose booster stessa. | Numero |
| N1 | Classificazione europea delle unità territoriali NUTS: livello NUTS 1. | Testo |
| N2 | Classificazione europea delle unità territoriali NUTS: livello NUTS 2. | Testo |
| ISTAT | Codice ISTAT della Regione. | Numero |
| reg | Denominazione standard dell'area (dove necessario denominazione bilingue). | Testo |

#### Classificazione

I dati in formato CSV hanno una classificazione .

### Dati set

Dai dati raw analizzati sono stati selezionati solo quelli di interesse per l’indagine e aggiunte colonne per evidenziarne i punti focali.

1. *Dataset\_Contagiati*:

Il dataset è stato costruito elaborando i dati grezzi presenti in [covid-19-ita-regioni.csv](https://github.com/pcm-dpc/COVID-19/blob/master/dati-regioni/dpc-covid19-ita-regioni.csv).

L’entità che si viene a formare avrà una primary-key a due dimensioni: DATA e REGIONE;

#### Campi set

|  |  |  |
| --- | --- | --- |
| Colonna | Descrizione | Tipo Dato |
| DATA | Data dell'informazione | YYYY-MM-DD |
| COD\_REGIONE | Codice della Regione (ISTAT 2019) | Numero |
| REGIONE | Denominazione della Regione | Testo |
| GUARITI | Persone dimesse guarite | Numero |
| DECEDUTI | Persone decedute | Numero |
| CONTAGIATI | Totale casi positivi | Numero |

1. *Dataset\_Vaccini*

Il dataset è stato costruito elaborando i dati grezzi presenti in [somministrazioni-vaccini-latest.csv](https://github.com/italia/covid19-opendata-vaccini/blob/master/dati/somministrazioni-vaccini-latest.csv).

L’entità che si viene a formare avrà una primary-key a due dimensioni: DATA e REGIONE;

#### Campi set

|  |  |  |
| --- | --- | --- |
| Colonna | Descrizione | Tipo Dato |
| DATA | Data dell'informazione | YYYY-MM-DD |
| COD\_REGIONE | Codice della Regione (ISTAT 2019) | Numero |
| REGIONE | Denominazione della Regione | Testo |
| VACCINI | Dosi di vaccino somministrate | Numero |

1. *Dataset\_Regioni*

Il dataset è stato costruito mantenendo la codifica ISTAT-2019.

L’entità conterrà una enumerazione specifica della relazione di ogni regione col proprio codice ISTAT, quindi ogni riga conterrà solo valori univoci per tutte le colonne.

#### Campi set

|  |  |  |
| --- | --- | --- |
| Colonna | Descrizione | Tipo Dato |
| COD\_REGIONE | Codice della Regione (ISTAT 2019) | Numero |
| REGIONE | Denominazione della Regione | Testo |