**BUILD WEEK 1**

**Gruppo 3**

**Studio di fattibilità per la creazione di un nuovo vaccino anticovid**

**Committente XYFARMA S.p.A.**

**A cura di**

**Alessio Mangiagi**

**Andrea del Bo**

**Marie Andrey Bwemba**

**Chiara Cordì**

**Federico Praticò**

**Marzo 29/03/2024**

Cliente : **XYFARMA**

Obiettivo : **studio di fattibilità per la realizzazione di un nuovo Vaccino anticovid**

Periodo analizzato : **Febbraio 2020 – Marzo 2024**

Fonte dati :

* **Sole 24 ore (**[**https://lab24.ilsole24ore.com/coronavirus/?refresh\_ce=1#**](https://lab24.ilsole24ore.com/coronavirus/?refresh_ce=1)**)**
* **Ministero della Salute – Istituto Superiore di Sanità**

Descrizione contenuto : dopo una prima fase consultiva, volta alla determinazione dei dati necessari per lo svolgimento dell’analisi , si è passati all’individuazione delle fonti : è stata pertanto condotta una ricerca approfondita sui dati ufficiali riguardanti l’andamento dei contagi, dei deceduti e dei guariti , soprattutto in ragione dei vaccini somministrati nel periodo di riferimento. I formati dei dati prescelti sono stati diversi : pdf, csv, metadati on line. Tutti i dati sono stati sottoposti al un lavoro processo di sgrezzamento e ripulitura, al fine di ottenere una data base valido e per quanto possibile, pulito.

Descrizione dei campi :

* Nel foglio excel “dati regioni definitivo” :
* **data** : campo indicante il giorno di riferimento, periodo dal febbraio 2020 al marzo 2024,
* **Regione** : campo testo , riportante la regione oggetto dell’analisi.
* **totale positivi** : campo numerico, indicante il numero totale dei positivi giornalieri.
* **variazione totale positivi** : campo numerico, indicante la variazione del numero totale dei positivi, rispetto al giorno precedente , nella regione di riferimento.
* **nuovi positivi** : campo numerico, indica il numero dei nuovi positivi accertati in quel giorno, in quella regione.
* **dimessi\_guariti** : campo numerico, accorpa in un unico dato , il numero dei soggetti dimessi e guariti dal covid., in quel dato giorno , nella regione indicata.
* **deceduti :** campo numerico, indica il numero dei morti , nel giorno e regione indicati.
* **totale casi :** campo numerico, riportante la somma algebrica tra totale positivi, dimessi\_guariti e deceduti.
* **tamponi :** campo numerico , che riporta il numero dei tamponi somministrati giornalmente.

* Nel foglio “maschera” : abbiamo riportato appunto una Maschera, nella quale è possibile, filtrando per Regione, conoscere in modo rapido, il numero dei contagiati, dei deceduti e dei guariti .

### Classificazione 5 stelle : secondo la classificazione ideata da TIM BERNERS LEE, i dati utilizzati nel presente lavoro appartengono alla categoria 5stelle,  in quanto si tratta di dati strutturati e codificati in un formato aperto, dotati di URI ( Uniform Rsource Identifier), cioè dati indirizzabili sulla rete , quindi accessibili on line, che contengono inoltre dei collegamenti ad altri dataset , come i LOD ( Linked Open Data).

### **Licenza:** [CC-BY-4.0](https://creativecommons.org/licenses/by/4.0/deed.en) - [Visualizza licenza](https://github.com/pcm-dpc/COVID-19/blob/master/LICENSE)

### Per completezza informativa abbiamo ritenuto opportuno fare un cenno anche alla licenza OPEN DATA d’uso IODL2.0, in ragione della quale i dati utilizzati in questo report appartengono alla licenza di tipo [CC BY](https://creativecommons.org/licenses/by/4.0/) . Questa licenza consente ai riutilizzatori di distribuire, remixare, adattare e sviluppare il materiale in qualsiasi mezzo o formato, a condizione che l'attribuzione sia data al creatore. La licenza consente l'uso commerciale.

**Nominare entità ed attributi** : paziente, positivi, contagiati, vaccinati, dimessi-guariti, deceduti.

**Chiave primaria di ogni tabella** : Regione.

**Domande da porre al cliente :** abbiamo pensato di sottoporre al cliente un sondaggio con domande a risposta aperta .

Le domande sono state :

* A quanto ammontano le risorse finanziarie che siete disposti ad investire?
* Quante e quali sono le risorse in termini di personale, impiegabili nel progetto?
* Valutereste la creazione di un vaccino che si rivolga solo ad una parte della popolazione?

**Metadati :** ci troviamo in presenza di metadati strutturali , in quanto le informazioni fornite permettono di localizzare il documento all’interno del sistema di conservazione.