**RELAZIONE DI ANALISI DATI COVID PER AZIENDA FARMACEUTICA XYFARMA**

## **Obiettivo**

Questa relazione è stata redatta per rispondere alla richiesta di XYFARMA di analizzare l’evoluzione del COVID-19 in Italia dal 2020 ad oggi.

I punti chiave dell'indagine includono l'andamento dei contagi, il numero di guariti e deceduti, le dosi di vaccini somministrate e le interrelazioni tra questi parametri. Questa valutazione dettagliata mira a fornire informazioni cruciali per supportare le decisioni strategiche di XYFarma in merito a potenziali sviluppi nel campo dei vaccini contro il COVID-19.

## 

## **Recupero e pulizia dei dati**

### 

### **Descrizione del contenuto**

Abbiamo selezionato il Dipartimento della Protezione Civile come fonte principale dei dati. Le informazioni necessarie per la nostra analisi sono disponibili in due repository GitHub: [COVID-19](https://github.com/pcm-dpc/COVID-19), che contiene i dati sui contagi e i decessi, e [covid19-opendata-vaccini,](https://github.com/italia/covid19-opendata-vaccini) che fornisce informazioni sui guariti e i vaccinati. Il primo repository offre una panoramica completa della diffusione del virus in Italia, mentre il secondo monitora la copertura vaccinale nel paese.

### **Licenze**

Entrambi i repository sono rilasciati sotto licenza CC-BY-4.0, che permette la condivisione e la modifica (anche a scopo commerciale) dei contenuti, purché non si limitino i dati e si attribuisca la paternità alla fonte originale.

### **Classificazione dei file e descrizione dei campi**

I tre file selezionati per l’analisi rientrano tra le due e le tre stelle nella scala di Tim Berners-Lee, ovvero sono file strutturati ma non inclusi in un database.

1. Il primo file, un Excel (due stelle) denominato dpc-covid19-ita-regioni nel repository COVID-19-master\dati-regioni, ha ventuno campi, tra cui data, regioni e casi totali, rispettivamente nei formati data, stringa e numero.
2. Il file copertura\_vaccinale, sempre un file Excel (due stelle), contiene 5 campi con il nome dell’area, la fascia anagrafica, i guariti e i malati.
3. L’ultimo, somministrazioni-vaccini-summary-latest, è in formato CSV (tre stelle) ed è stato scaricato dalla repository covid19-opendata-vaccini. Esso contiene quindici campi, tra cui la data della somministrazione, le regioni e il numero totale di dosi di vaccino somministrate per giorno e regione.

### **Domande per il cliente**

Per dare contesto e precisione alla nostra analisi, riteniamo necessario avviare una fase esplorativa con XYFARMA e i suoi utenti.

Innanzitutto, ci interessa conoscere le tempistiche previste per lo sviluppo e la commercializzazione del nuovo vaccino, attraverso l’impiego di interviste strutturate. Questa modalità è dettata dalla difficoltà di soddisfare un’eventuale crescita della domanda, a cui deve seguire una risposta basata su dati precisi e quantitativi.

È altrettanto fondamentale comprendere il target di riferimento che XYFARMA intende raggiungere, per stabilire se l’obiettivo è proteggere tutte le fasce d’età o se si punta specificamente a categorie a rischio. Anche in questo caso, la consapevolezza del problema impone il ricorso alle interviste strutturate.

In ultimo, l’analisi della concorrenza e delle best practice può rivelarsi utile per capire la propria posizione nel mercato. Questa risulterebbe particolarmente preziosa nell’ottica di una strategia di business su base regionale, unita all’uso dei sondaggi al fine di coinvolgere un’ampia base di utenza.

### **Metadati**

Entrambi i repository contengono metadati descrittivi per facilitare la comprensione dei singoli dataset e dei loro campi. Per quanto riguarda i metadati amministrativi e gestionali, la repository COVID-19 ha prodotto un’infografica nella sezione “Aggiornamento e flusso dei dati”. Infine, la sezione “Struttura del repository” funge da metadato strutturale.

## 

## **Presentazione e Analisi dei Dati**

La prima pagina è dotata di una maschera interattiva per effettuare ricerche su base regionale, consentendo agli utenti di reperire il numero totale di contagiati, guariti e deceduti.

La pagina successiva presenta una tabella riepilogativa, includendo una colonna dedicata al rapporto tra contagiati e vaccinati. Il risultato è stato riportato in percentuale per dare una visione immediata sull’efficacia del vaccino.

Seguono tabelle che illustrano l'andamento annuale dei dati relativi a contagiati, guariti e deceduti, affiancate da grafici che stabiliscono relazioni con il numero di vaccini somministrati.

In questa prima parte del lavoro è stata presa in considerazione l’andamento annuale dal 2020 al oggi diviso per regione per individuare la tendenza dei contagi cumulando il totale dei contagiati.

Successivamente, per avere un quadro più dettagliato si è scelto di procedere considerando gli andamenti trimestrali regionali, in questo caso avendo dei dati di tipo progressivo per le voci guariti, deceduti e vaccinati, è stato più rappresentativo considerare il valore massimo in ogni trimestre.

## **Conclusioni**

Analizzando i dati attraverso il grafico nel foglio *Trimestri*, emerge chiaramente che in tutte le regioni si registra una diminuzione del trend dei decessi già dalla fine del 2021, grazie all'efficacia dei vaccini sviluppati, mentre si osserva un aumento nel numero dei guariti.

Analizzando il grafico nel foglio Grafico *Vax Età*, è emerso che ad oggi la fascia di età 60-90 è la popolazione più a rischio. Nel 2023 una persona su sette fra quelle che hanno richiesto una dose di vaccino corrisponde a questa fascia d’età.

Pertanto, non reputiamo strategico ad oggi lo sviluppo di un nuovo vaccino. Tuttavia, i dati a disposizione, suggeriscono che un eventuale sforzo in tale direzione dovrebbe essere ottimizzato per la fascia di età 60 e oltre e portato avanti nelle regioni del Nord Italia, come evidenziato nella tabella *Contagiati tot\_regione.*