

Numbers matter

#covid19 – history 3

If we consider only the generic data of a country, we can't have a complete information without any further details.

Saying that there are more than 230.000 cases in Italy doesn't give us any data to increase awareness on the virus behavior.

We are not saying anything about localization or age range of sick people.

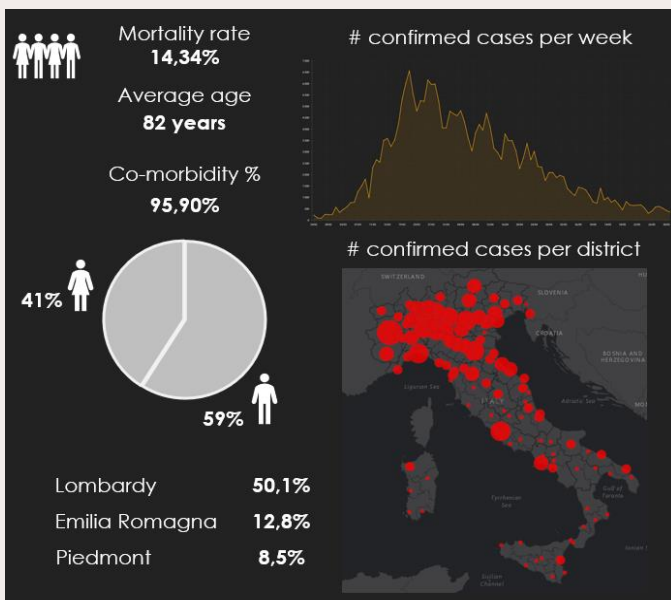
Having more detailed data by region and age would enable us to check the most affected areas and groups.

Most of the times, it may be useless to use an overlay aggregated average value as a meaningful information. Rather, inside it we can find many variations that allow us to discover additional signs on the observed reality.

After the first period of chaos, we are now able to draw a clearer picture of the situation thanks to the use of an effective segmentation on data collection.

After the first reactions to Covid19 outbreaks, healthcare and statistical institutions started to provide analyses based on several data: region, age range, gender, previous existing pathologies and so on.

Based on those details, of course, mortality rates show different values for different clusters, and new and more specific questions rise up.



“ Data must be considered, but a decision cannot be delegated to an automatic process”



Tommaso Cavalli
Partner K.Group*

I numeri contano

#covid19 – storia 3

Considerando i dati macro di un Paese non possiamo beneficiare di reali informazioni senza disporre di ulteriori elementi di dettaglio. Dire che ci sono più di 230 mila casi in Italia non ci fornisce informazioni per accrescere la nostra consapevolezza sul comportamento del virus. Nessuna informazione rispetto la localizzazione o rispetto il range di età della popolazione. Avere dati più dettagliati per regione ed età consentirebbe di controllare le aree ed i gruppi di popolazione maggiormente colpiti. Spesso potrebbe non aver senso usare il valore medio ad un livello troppo sintetico come informazione significativa.

Al suo interno si possono trovare molte variazioni che permettono di scoprire più elementi rispetto la realtà osservata.

Usando un'efficace segmentazione nel data collection, dopo il primo periodo di caos, siamo ora in grado di disegnare un quadro più chiaro della situazione.

Dopo le prime reazioni al diffondersi dell'epidemia di Covid19, le istituzioni sanitarie e statistiche hanno cominciato a pubblicare analisi basate su varie dimensioni: regioni, range di età, sesso, patologie preesistenti e così via.

In base a questi dettagli certamente il tasso di mortalità svela differenti valori per differenti classi di analisi e nuove e più specifiche domande cominciano ad emergere.

Dove sono i focolai della malattia?

Il Covid19 sta realmente colpendo più maschi che femmine?



Where are the outbreaks?
 Is Covid19 affecting more males than females?
 Is Covid19 more dangerous for the subjects aged over 70?
 To what extent is co-morbidity a cause of death?

Based on these real global data, we can consider some aspects of the **impact** of the virus. Certainly, we cannot tell why the virus is functioning that way yet, because we are still waiting for the answers of scientific studies, but our government could take some more precise actions.

Despite the usefulness of these data, some flaws in both data collection and outbreak forecast can be detected.

- First of all, the tests were carried out on symptomatic subjects, not randomly. In this case, it's hard to have a **significant view** of the whole country only with a selected sample.
- Without any objective answer from scientific studies, we are not able to carry out any **evaluation** on the progress of the virus, and forecasting methods based only on historical data cannot be taken seriously.

Despite this deep uncertainty, decisions have to be made and we cannot expect them to be totally data-driven.

This is what normally happens in **business organizations**, where data must be considered, but a decision cannot be delegated to an automatic process.

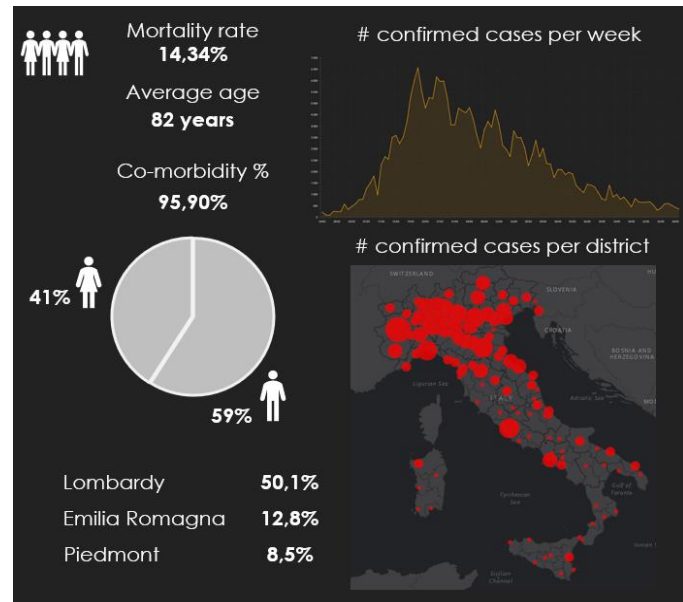
“ After the first period of chaos, we are now able to draw a clearer picture of the situation thanks to the use of an effective segmentation on data collection ”

* Bio

Tommaso

has more than 20 years of experience in performance management projects in operational and financial fields. Since 2001 partner of K.Group, he works now as knowledge worker for several organizations supporting management in data-driven decision making process.

Il Covid19 è più pericoloso per gli over 70?
 Quanto è importante la comorbidità come causa di morte?



Basandoci su questi dati del mondo reale possiamo considerare alcuni aspetti dell'**impatto** del virus.

Certamente non possiamo ancora dire perché il virus funzioni in un certo modo visto che si stanno aspettando le risposte della scienza ma il nostro Governo può adattare i propri interventi in modo più preciso.

Nonostante questi dati siano molto utili ci sono alcuni difetti sia nella raccolta dei dati che nelle previsioni dell'epidemia.

- Prima di tutto i tamponi sono stati fatti sulla popolazione con sintomi e non in maniera casuale. In questo caso, mediante un campione auto-selezionato, è difficile ottenere una **vista significativa** dell'intero Paese.
- In assenza di risposte dalla scienza non siamo in grado di produrre **valutazioni** sulla situazione del virus e le previsioni basate su metodi statistici non si possono prendere in seria considerazione.

Nonostante il forte momento di incertezza, delle decisioni devono essere prese e non ci possiamo aspettare che siano totalmente data-driven.

Questo è quello che normalmente accade nelle **organizzazioni aziendali** dove i dati devono essere presi in considerazione, ma una decisione non può essere delegata a un processo automatico.

