The need for data during a pandemic

With the current Covid-19 virus placing many countries and their people into lockdown, there is a sudden thirst for data and information that unravels what the spread of the virus means. Leaders around the world are using data to present models on how we can "flatten the curve" and how physical distancing impacts numbers, as well as what specific infection scenarios will yield, and are showing.

Clinton Scott, MD at TechSoft International

According to TechSoft International, data science and the visualisation of analytics in these trying times, is providing people with a simplified means to understand the complex scenarios that the global population is being confronted with.

"The data scientists at TIBCO have been collecting, collating and unpacking data from myriad resources to try and create scenarios for not just the layman, but governments alike, to offer a better understanding what lies in the detail," says Clinton Scott, managing director at TechSoft International.

"Critical learnings thus far, include the fact that errors around future predictions are very real. It is only through precise and ongoing data modelling, visualisation, and predictions of infections that medical professionals will be able garner a view of emerging predictions."

A lesson in data

Using visual analytics, data scientists are using modelling, simulation and analytics, inputting verified and trusted data sources from around the world. One such example includes the Johns Hopkins University, and locally the National Institute for Communicable Diseases, albeit the latter's external data is scant and merely there as a public service announcement.

"A hard lesson that the world is being faced with is to rely on only qualified data sources. Sure data can be collated from anywhere, but is it accurate, and is it qualified are key considerations. When connected to an accurate resource, and models are defined, you can gain a continually updated view of data and see the Covid-19 case trajectories by country," adds Scott.

Drawing from these data sources, the team at TechSoft International has developed an app, called Spotfire, that shows fatality trajectories by country, and tracks where peaks and troughs are being experienced in new infections.

Modelling outbreaks and interventions

"When we have data that shows what the state of a situation looks like, we can start using it to drill down into the cause and effect and then map the potential outcomes. Which is what the Spotfire app is attempting to highlight," states Scott.

When viewing the visualisation in the app, it is important to note that Epidemiologists model infectious diseases in compartment models; for example, the SEIR model where people transition from susceptible (S) to exposed (E) to infected (I) to removed (R), with S+E+I+R = N, where and R can be recovered or died, and N is the total population size. This then also leads to the reproduction number (R0), which is the average number of people infected from a person with an infection.

"When you provide the public health system with this data they can act on one of two things. Where can they slow or stop the spread and what mitigation strategies are working. It also offers the public a very real, visual view of the cause and effect of their own actions. Which at this point might be the most critical component in helping to stop the spread," adds Scott.

In the data, you will see that the initial focus of health experts is to focus on suppressing, namely reducing the RO by isolating infected people, reducing case numbers and containing it until there is a vaccine. This worked in the case of SARS and Ebola, but Covid-19 is different, some patients are 100% asymptomatic, so their data is not being collected and with that goes their ability to infect others. Nations like South Korea where mass testing has been enforced helps to identify more cases, and in turn, gives a better view of the data.

Tracking the data

"One huge curveball that Covid-19 is throwing at researchers is discrepancies in fatality rates per region. The variable factors such as age demographics, air pollution, social behaviour patterns, and access to healthcare are so varied by region that the Case Fatality Rate (CFR) is going to differ. This makes it very difficult for researchers to create a single view. Looking at the data one thing that is clear is that the models are proving that by comparison, the CFR for Flu is ~0.1%. So Covid-19 is ~10X+ more deadly than Flu," he says.

When tracking CFR data, variables make accurate figures almost impossible to provide, as this depends on the fact that every fatality has been tested. Which we currently know not to be the case, especially in places like Iran, where some deceased are only showing positive after they have passed away and a post mortem test has been conducted.

The Spotfire tool is also offering insights into the notion of "Flattening the Curve", by using a simple 3-compartment SIR

numeric model, with Susceptible, Infected and Recovered sub-populations (e.g. Jones 2008). Here the changes of specific populations over time are being measured, looking at contact between people, mobility, and the natural rate of recovery from the disease all to see the changes they are having on RO. Using a numerical model, TIBCO is able to explore scenarios for mitigation of an outbreak.

Is it working?

"I am no medical professional, but rather believe in the data. Medical professions warn that we still need to create herd immunity, making it critical to balance the timing of the introduction of social distancing, isolation and quarantine measures, with the scale of disruption imposed and the likely period over which the interventions can be maintained."

Looking at the data models presented by TIBCO, you can see when a flattening occurred in South Korea, where broad testing in South Korea (>270,000 tests) was mounted early along with social distancing, closing of schools, tracing contacts. Here you can see the curve flattening positively.

One thing is sure. The data shows that without intervention – things will get worse. Acting too early may be a risk, but not as much as acting too late. Looking at the numbers, without intervention the number of cases will double every 3-4 days. Then it becomes like chasing the dragon, as opposed to trying to curtail a pandemic.

"While the data is telling, it is still flawed, making it almost impossible to create a fool-proof Covid-19 model. Which means we are faced with a societal issue and not just a disease. People's mental health is as affected as their ability to earn an income, which will see many unravel. In short, wash your hands, stay at home if you can, and consider others. Don't become obsessed with the data, but ensure that the data you are reading speaks to the facts and doesn't feed the fiction," concludes Scott.