Summer Training Report

By

Mohd Aamir Khan

MBA (Hospital and Health Management)

2019-21

IIHMR UNIVERSITY

Acknowledgements

A summer training project is an excellent opportunity for enlightening and self-development. I think I am fortunate enough for having been provided an opportunity of summer training project. I wish to take this opportunity to express my deepest gratitude towards all the people who have shown their willingness to provide their valuable inputs to help me prepare this report.

I am thankful to Dr P.R. Sodani, Dean- Academics for granting me permission to undertake the project. I would like to convey my thanks to my mentor and my project guide Dr Alok Kumar Mathur, Associate Professor, for ready assistance, keen interest and valuable suggestions whenever I sought them. You provided me the knowledge and tools that I needed to successfully complete my project.

In addition to this, I would like to thank other faculties for their active support, which was of great help during the period of my summer training project.

Finally, there are my friends, who were of great support in deliberating over our problems and findings, as well as providing happy distraction to rest my mind outside of my work.

Without the support of all above mentioned people, this project would not have been successful. I once again thank you all for your great interest and support throughout the project.

MOHD AAMIR KHAN

Table of Contents

List of Figure and Charts

List of Abbreviations

SARS	Severe Acute Respiratory Syndrome Corona Virus -2
WHO	World Health Organization
CDC	Centers for Disease Control and Prevention
PCR	Polymerase Chain Reaction
PRC	Peoples Republic of China
NCBI	National Center for Biotechnology Information
NIC	National Informatics Center
USA	United States of America
NGO	Non-governmental Organizations
IOT	Internet of Things
MIOT	Medical-Internet of Things
AI	Artificial Intelligence
VR	Virtual Reality
UV	Ultra-violet
IT	Information Technology

PART 1- RESEARCH REPORT- CONTRIVANCE OF DIGITAL TECHNOLOGY FOR COVID-19

Abstract

a. Introduction-When a disease reaches to pandemic level leading to severe disaster situation, it spontaneously alerts the health-care workers and hospitals for protecting the health of the citizens of the country. This requires adoption of quickly acting health protecting systems with immediate and distant vision, which are to be continued even for the longer period. This is an essential task since the disease in its pandemic form, spreads over mass population with a faster pace and persists for a longer period. This certainly leads to damage of the health system of mass population at one end and at other end it badly affects the economy of a country. Such pandemic also overburdens the health-care facilities. As we are having limited resources in health sector, we need to work with less and need to give the more. To enhance the capabilities of our hospitals and various health-care organizations, digital technology can play a major role in early detection of the disease as well as limiting the spread of the disease. This report provides possible solutions for the application of digital technologies which can further help in overcoming the pandemic situation, highlighting the successful ways like pandemic planning, contact tracing and surveillance, increased testing and many others in order to achieve the controllable position in the current outbreak.

b. Objective- To understand the use of digital technology to fight against covid-19 at different levels.

c. Methodology- Data was collected from various research papers, articles and studies conducted by different organizations and universities like Elsevier Covid-19 Resource Centre, Sage Pub, Lancet, University of Sheffield, University of Ghana, University of Westminster, Bahrein Institute of Banking and Finance as well as Indian Institute of Health Management Research, Jaipur.

d. Findings- In the times when research on covid-19 is providing new insights everyday and manual approach is lacking in tackling the outbreak, digital technology can be helpful in diagnosis of outbreak with the help of AI, disease tracking, forecasting the health condition of the patient, discovery of drug and generating awareness towards the outbreak.

e. Suggestions- Digital Technology can assist world in monitoring, surveillance, patient data processing, managing data in on cloud level in secure manner which can help remote patients to access their data in optimized experience. Drones can also be helpful in sending supplies, managing patient's treatment and surgeries.

These advanced technology solutions can mitigate the effects of the virus to a significant extent which will further helpful in suppressing the disease.

1. Introduction

As the year of 2020 started, a new disease came in light named novel Corona virus disease, also known as Covid-19, which was first detected in Wuhan, China. This disease is caused by Novel Severe Acute Respiratory Syndrome Corona Virus-2 (SARS-CoV-2). Transmission of virus is through droplets on close contact with patient with the incubation period of about 14 days in most of the cases. The disease spreads in phases from individual to community level with high transmissibility. As the death toll is rising rapidly in almost all the countries of the World including India, the main challenge is how to suppress the disease to the zero level when its vaccine is not yet available.

Health system of world has been collapsed because of lack of resources and medical equipment in developing countries like India and delayed response by developed countries like United States of America. Also, according to the official data, it has been observed that there are around 80 percent asymptomatic cases and symptomatic cases are around 20 percent in India. Presently, this disease is being tackled just like Severe Acute Respiratory Syndrome (SARS) in 2003 along with certain restrictions. Although, as its curative measure, The India has earlier started using Hydroxychloroquine with certain combinations which is significantly improving the condition of some patients, recently Remdesivir is being used in some cases and some cases are being treated with plasma therapy, depending on the condition of the patient. However, the measures are not that effective as success rate is far smaller than the disease growth rate. Plasma therapy although giving positive results but it cannot be considered as primary treatment. Strict lockdown applicable to whole of the population, is one of the preventive measure. Most of the governments have imposed it but it leads to major downfall in economy of countries. The implication of lockdown condition for the civilians is, however, in strict sense is not a complete protective measure but only acts as a pause to the disease for some amount of time. With such difficult situation, where human efforts do not seem sufficient, we can rely on the digital technology which can help to control the Covid-19 disease to its minimal level.

2. Review of Literature

2.1- Background- Since the time when SARS was arised in 2002 as an outbreak, the WHO along with various other National and International organizations discussed this issue and made some efforts to handle the situation. They also took the help of communities. When everyone got relaxed about the efforts to fight viruses, in the year 2009, H1N1 Virus (Swine Flu) had accentuated the efforts made by these organizations. Twenty cases of H1N1 influenza were reported in USA in the month of April, 2009 which started becoming a pandemic and resulted in 12220 deaths in over 208 countries. This virus attacks the respiratory system of body. Later-on, Middle-east respiratory syndrome was reported in 2012 from Middle-East, which was associated with Corona virus. In the year 2014, Ebola virus outbreak came in light in Republic of Cango, although it was earlier reported in 1976. Vaccine of Ebola was announced by CDC in 2015. Zika virus outbreak was also reported later in 2016. All these virus outbreaks, give us a hint that the preparedness of world against the virus outbreak is not enough, specially for the recent Covid-19 outbreak.

2.2- Existing Condition Globally- The rapid speed with which covid-19 has spread globally, requires innovation. In current scenario, we are following PCR Test, which takes up-to 48 hours to generate results. This test examines naso-pharyngeal swab or sputum sample. After taking swab, clinician looks for presence of genetic material of corona virus. Other methods include Isothermal amplification assay which has been recently approved by USA but not commonly used in India.

3. Objective

To understand the use of digital technology to fight against covid-19 at different levels.

4. Methodology

4.1- Type of study- Descriptive study

4.2- Type of data- Qualitative data

4.3- Technique- Secondary research

4.4- Data collection method- A progressive search of data related to modern technology for Covid-19 is made and assessed the various aspects of digital and latest available technologies which can be helpful in tackling Covid-19. This report is prepared by collecting data from Research papers and reports of universities and scholars, mainly includes data from United Nations, Sage Pub, NCBI, Lancet Digital Health, University of Sheffield, University of Ghana, Bahrein Institute of Banking and Finance, University of Westminster, IIHMR University, and articles published in Bennett Coleman and Company Ltd' media platforms. Situation is analyzed on the basis of data collected from December, 2019 to June, 2020.

5. Findings

As the research on novel Corona virus is progressing, new information's are available which lead to development of new strategies. The pandemic has caused extensive demand for digital technology solutions for health and has resulted in availability of successful solutions like screening, tracking, surveillance and various other solutions.

Digital technology can be helpful in diagnosis of outbreak with the help of radiology images, disease tracking, forecasting the health condition of the patient, discovery of drug and generating awareness towards the outbreak.

Figure 1: Steps to manage pandemic using digital technology

6. Analysis of Data

The rapidly emerging situation of Covid-19 disease is a wake-up call for health-care professionals, health-care service providers as well as the patients. We should understand that upgradation to digital technology starts with our-selves. This study is the review of various papers, related to digital technology and covid-19, published in distinguished platforms like lancet, Elsevier, sage-pub and many more.

6.1- Tracking and Tracing-

6.1.1- About Tracking and Tracing- Having a system which monitors online and detects location is very much helpful in guiding health-care professionals about the time schedule related to patients. One of the prime example of tracking and tracing in India is Arogya Setu app. With over 100 million downloads till May 23rd, 2020; Arogya Setu app is available in India for self-detection as well as the patient tracker.

6.1.2- About Arogya Setu App- Arogya Setu is a mobile application, developed by the Government of India. Its focus is to protect citizens during Covid-19. This app helps by informing the people about presence of active covid-19 patients in their nearby area. It also suggests best practices to stay healthy along with certain medical advices related to covid-19. This app is developed by National Informatics Center (NIC). It is based on AI.

6.1.3- Functions- It tells the people about patients of Covid-19 in the surrounding area with the help of Navigation system and Bluetooth. It's identification of covid-19 patients is based on signs and symptoms provided by user and, also data of active cases provided by health centers. So, it's accuracy and efficacy is partly dependable on data provided by users.

6.1.4- How it works- The Arogya Setu app on your phone detects other devices that have the Arogya Setu app when they come within the Bluetooth proximity of your phone. When this happens, both the phones securely exchange a digital signature of this interaction, including time, proximity, location and duration. This data is stored on the device of all individuals. In the unfortunate event that any of the people that you came in contact with during the last 14 days, tests positive for covid-19, the App calculates your risk of infection based on the number of your interactions and the proximity of your interaction and recommends suitable action. This action is displayed on your Home screen. Your updated risk of infection is analysed by Government of India, to facilitate suitable medical interventions, as and when required.

6.1.5- Suggestions- Strict monitoring and surveillance.

6.2- Artificial Intelligence-

6.2.1- About- In computer science, AI is intelligence demonstrated by machines, unlike the natural intelligence displayed by humans and animals. It is also called machine intelligence. In these times of crisis, one thing is clear that investment in digital technology can only suppress the effect of Covid-19.

6.2.2- Functions- AI has capability to copy Human Intelligence Process. It learns from experience. AI can also be useful in detecting and forecasting outbreaks. Use of AI based systems can help alleviate the load of health-

care professionals. *Chatbot* is also a type of AI which detects patient's data, temperature, records patient history, provides consultation and prevents unnecessary load to doctors

6.2.3- How it works-

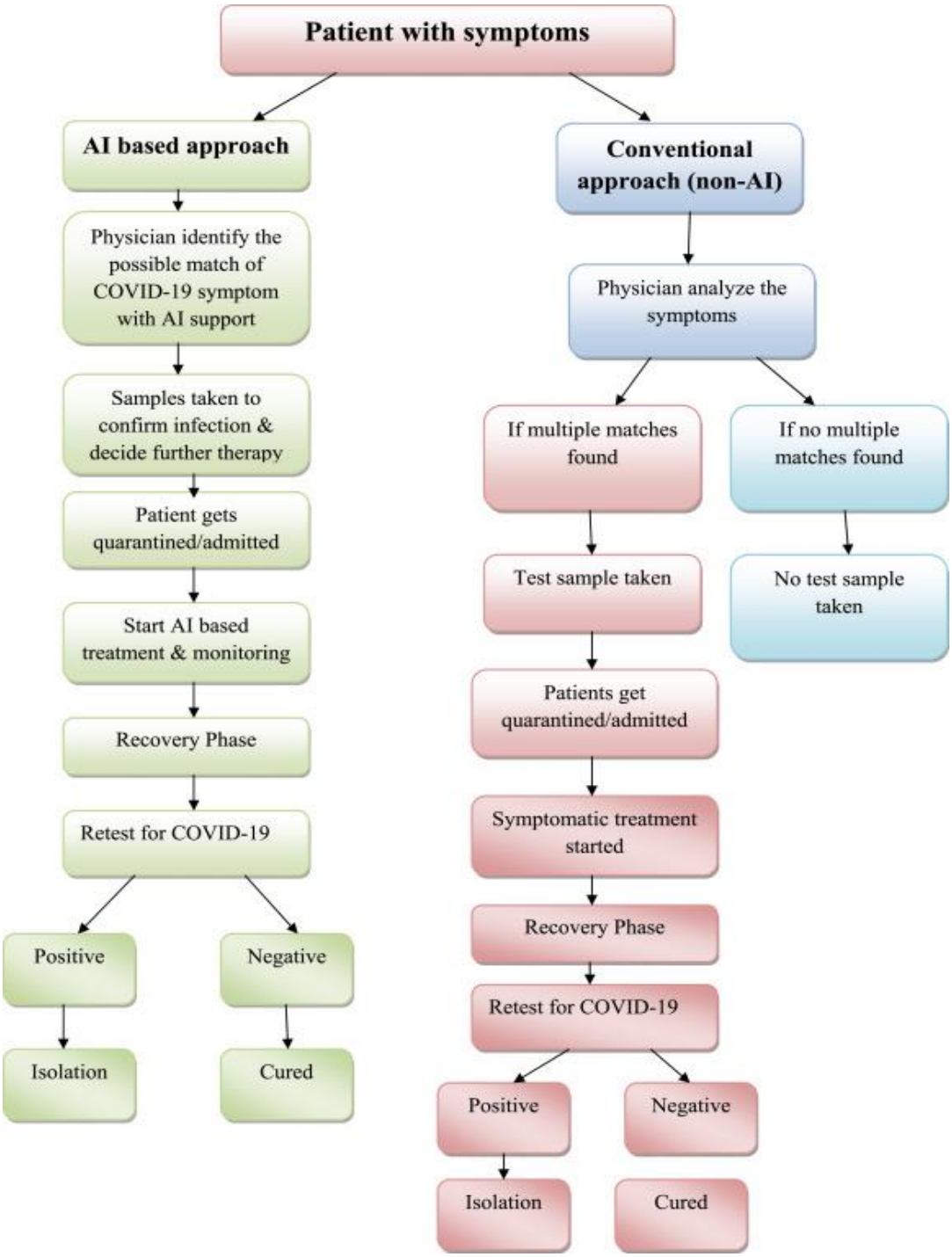

Chart 1: How AI works in Patient Management

6.2.4- Suggestions- Careful balance between data privacy and public health, and rigorous human- AI interaction required.

6.3- Blockchain-

6.3.1- About- A Blockchain is a type of diary or spreadsheet which contains information related to transactions. If a transaction is approved by a majority of the nodes then it is written into a block. Each block refers to the previous block and together forms a Blockchain.

6.3.2- Function- Blockchain protects data from tempering and prevent data loss. Blockchain provides traceability. It helps in transparency of information. It not only helps in management of patients and hospital records but also can track the supplies like medicines, sanitizers, disinfectant and other materials of the health-care organizations. For example, in February,2020; Alibaba launched a blockchain system which helped the Govt of China and others NGO's to work more efficiently and effectively in better collaborative way.

6.3.3- How it works- Blockchain platforms can be used to monitor response to pandemic situation in a fast and transparent way without the violation of user data. Blockchain can help companies to research the best available ways to combat the Covid-19 and contact-tracing applications can be built on blockchain to allow anonymity. Blockchain can bring reliability, transparency and security to medical data. Several organizations have been accused of manipulating data during this pandemic, blockchain can help to solve them by providing transparent and immutable medical data. A blockchain-based global pandemic map can be used to track the spread of the virus, the number of infected citizens and the number of recovered citizens. Also, blockchain can provide companies payment receive system through cryptocurrencies which will promote less use of cash.

6.3.4- Suggestions- Services related to Record management, Retails and Human Resources can be improved using Blockchain.

6.4- Drones-

6.4.1- About- Drone is an unmanned aircraft or ship, that can navigate autonomously without human control or beyond the line of sight.

6.4.2- Functions- Drones are used in many areas and have many applications. There is no end when it comes to their possibilities. Main areas where drones are of great use are--Search and rescue, Security, Surveillance, Science and research, Mapping, Aerial photography and video capturing.

6.4.3- How it works- It has been widely observed that the government is facing problem to track the disease suspects in narrow areas. Such areas have less accessibility. In order to gain access to these areas, health-care

workers are using drones. These drones are useful in surveillance, tracking and observing the suspects. Drones are also used to enforce social distancing by filming public areas and ordering people back to their homes during lockdown. Many time's these drones are also used in announcements and food delivery in the area which is hotspot of an outbreak.

6.4.4- Suggestions- As drones are widely being used to spread disinfectant, sending medical supplies and in thermal camera monitoring which is giving positive results, so we can broaden the horizon of its use to make sure it's been used more efficiently.

.

6.5- Internet of Things (IoT)-

6.5.1- About- The Internet of Things (IoT) is a scheme of interconnecting computing tactics, digital, and mechanical devices; possessing the capability of transmission of data over the defined network without having any human involvement at any level. In the present typical situation, most of the problems are arising because typical reachability to the patients, which is the second most considerable issue after the concern of vaccine development.

6.5.2- Functions- The use of the IoT concept makes better reachability to patients, which ultimately provide them significant care so that they can get out of this disease. IoT can be used in medical field as M-IoT. It also helps in prevention and control of Covid-19.

6.5.3- How it works- IoT is an innovative technology which ensures that all infected persons due to this virus are quarantine. During quarantine, it is helpful in a proper monitoring system. All high-risk patients are tracked easily using the internet-based network. This technology is used for the biometric measurements like blood pressure, heartbeat and glucose level.

6.5.4- Benefits of IoT in terms of Covid-19-

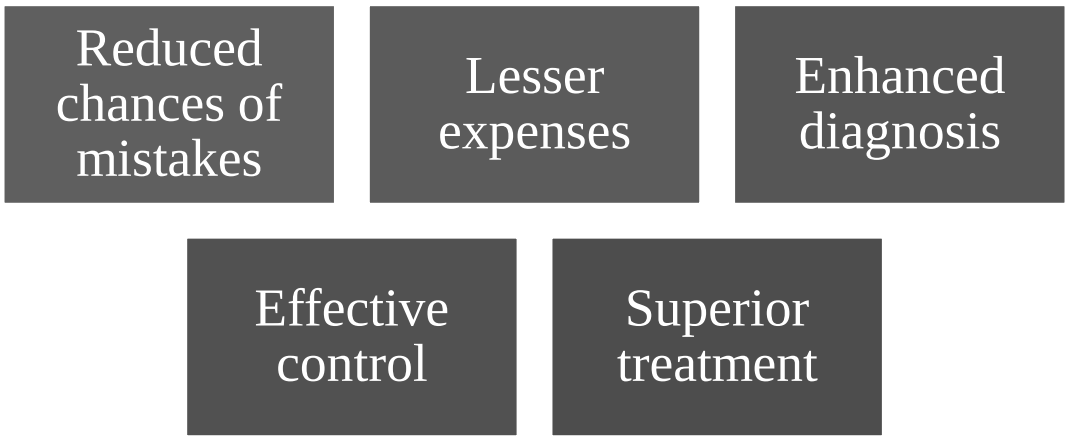

Chart 2: Benefits of IoT in Covid-19 management

6.5.6- Suggestions- It should not have security holes. For this IoT should be designed of standard level. Plus, IoT requires network to communicate with each other so networking should be less expensive. Also, it generates lots of data, so focus should be on artificial intelligence so that even large data sets can be processed easily.

6.6- Virtual Reality-

6.6.1- About- Virtual reality (VR) is defined as developing simulated expertise which is somewhat similar to the real-time situation.

6.6.2- Functions-

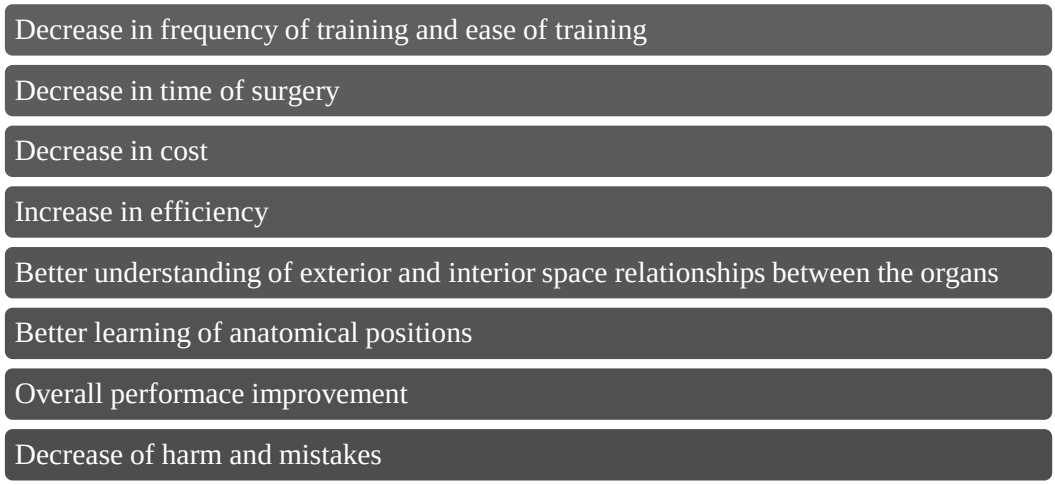

Chart 3: Functions of VR

6.6.3- How it works-

VR is beneficial for remote sites for exploring telemedicine, planning, treatment, and controlling of the infections by providing proper awareness to the people regarding this disease. VR technology develops a platform to reduce the face to face interaction of doctors with the infected covid-19 patients. Plus, through live video streaming, it helps to improve surveillance systems on the ongoing situation.

6.6.4- Suggestions- Virtual Reality is of great use as it provides better sense of place, one can learn by doing and it makes a person more creative. So, we should invest more on VR to improve efficiency and increase creativity.

6.7- Deep Learning-

6.7.1- About- Deep learning is a class of machine learning algorithms that uses multiple layers to progressively extract higher level features from the raw input. For example, in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits or letters or faces.

6.7.2- Functions- People suffering from covid-19 have more chances that their lungs will be affected which may later leads to death of the patient. With the help of deep learning, we can identify people who are having damaged lungs because of covid-19 from normal individuals or pneumonia (not infected by covid-19). The detection of covid-19 can be carried out using various deep learning models. Since it is important to detect covid-19 that spread rapidly and globally, AI techniques can be used to perform this accurately and quickly.

6.7.3- How it works- In association with deep learning, AI has been used in many applications such as image detection, data classification and image segmentation. People infected by covid-19 may suffer from pneumonia because the virus impacts the lungs. Many deep learning studies have detected the disease using a chest X-ray image data approach. The success rate was 93.73% in pneumonia from X-ray images.

6.7.4- Suggestions- We should take care of quality of data sets. Capacity can be increased so that there can be more scope collectable information.

6.8- Robotics-

6.8.1- About- The global battle against COVID-19 is playing an extremely important role in assisting humans in containing the spread of the virus and dealing with the existing cases. One of the key technologies that has made a huge difference on the ground is robotics. Many hospitals across the globe are currently using robots to aid both the healthcare staff and patients.

6.8.2- Functions- From preparing meals at hospitals, doubling up as waiters in restaurants, spraying disinfectants to vending rice and dispensing hand sanitizers, robots were on the frontline to prevent the spread of Coronavirus. In many hospitals, robots were also performing diagnosis and conducting thermal imaging. Shenzhen, PRC based company 'Multicopter' is using robots to transport medical samples.

6.8.3- How it works- A study published in the Science Robotics says, "Robots have the potential to be deployed for disinfection, delivering medications and food, measuring vital signs, and assisting border controls. As epidemics escalate, the potential roles of robotics are becoming increasingly clear."

“For disease prevention, robot-controlled non-contact ultraviolet (UV) surface disinfection is being used because COVID-19 spreads not only from person to person via close contact respiratory droplet transfer but also via contaminated surfaces,” the study further says.

6.8.4- Suggestions- As time is changing, robots are also playing an important role in fight against diseases like COVID-19. In the case of an outbreak, the robotic technology can play a crucial role in not just assisting the patients but also keeping the doctor and healthcare staff safe. So, we should work on using advanced technology of robotics in health-care organizations.

6.9- Big Data-

6.9.1- About- Big data is a term that describes the large volume of data – both structured and unstructured – that inundates a business on a day-to-day basis. But it’s not the amount of data that’s important. It is what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.

6.9.2- Functions- Big data is the accumulation of all the data points related to this disease being received from around the world. Mathematical modeling has taken that data and used it to identify geographical hotspots, make death prediction provide estimate for testing supplies, and guide decision-making among policymakers, health care providers and other key stakeholders.

6.9.3- Suggestions- Big data is powerful. But with great power comes great responsibility. While we have access to this huge database of information, there will always be some blind spots and missing variables of data that prevent us from having the full picture. In the proper context, big data can be incredibly useful. But in order to harness its full power, we need to have both people who are familiar with data models as well as people who understand the epidemiology and the medical implications of the virus to work together. It is important to keep in mind that existing big data models may be incomplete due to variables that remain unaccounted for (such as population density), either due to statistical or methodological limitations, or because the relevant variables have not yet been identified. This means that despite our best efforts, we may still be missing the entire picture on COVID-19.

6.10- Facial Recognition-

6.10.1- About- A facial recognition system is a technology capable of verifying a person from a digital image or a video frame from a source. There are multiple available methods in which facial recognition systems

work, but usually they work by comparing selected facial features from given image with faces within a database. It is also described as a Biometric Artificial Intelligence based application that can uniquely identify a person by analyzing patterns based on the person's facial texture and shape.

6.10.2- How it works- Facial recognition technology involves the processing of biometric data for the purposes of uniquely identifying an individual, which constitutes a "special category" of personal data.

6.10.3- Suggestions- Facial recognition technology is preferred because it has highly invasive potential and can infringe individuals' rights to privacy and the protection of their personal data. The use of facial recognition technology should also be fair, lawful and transparent and the purposes for which it is used should be limited.

6.11- Cloud Computing-

6.11.1- About- Cloud computing is the on-demand availability of computer system resources, especially cloud storage and computer power, without direct active management by the user. Cloud computing is the delivery of on-demand computing services from applications to storage and processing power, typically over the internet and on a pay-as-you-go basis.

6.11.2- How it works- Rather than owning their own computing infrastructure or data centres, companies can rent access to anything from applications to storage from a cloud service provider. One benefit of using cloud computing services is that firms can avoid the upfront cost and complexity of owning and maintaining their own IT infrastructure, and instead simply pay for what they use, when they use it. In turn, providers of cloud computing services can benefit from significant economies of scale by delivering the same services to a wide range of customers.

6.11.3- Examples- Cloud computing underpins a vast number of services. That includes consumer services like Gmail or the cloud back-up of the photos on your smartphone, though to the services which allow large enterprises to host all their data and run all of their applications in the cloud. Netflix relies on cloud computing services to run its video streaming service and its other business systems too, and have a number of other organizations.

6.11.4- Suggestions- Increase security, Optimized storage environment, Optimize end-user experience.

7. Results

This study provides a window to think about the technologies which are used to decrease the impact of pandemic. Presently, various technologies are being used or going to be used in various sectors including health as well as other industries.

Hospital Care	Hospital Robots
	Clinical Algorithms
	Protective Gear
	Advanced Ventilator Valves

Remote Healthcare	Remote Monitoring
	Telehealth
	Digital Stethoscope
	Wearable Sensors
	Chatbots

Diagnostic Solutions	Mobile Diagnostics
	Tele Health
	Remote Patient Diagnosis
	Disease Detection Kits
	Temperature Monitoring

Preventive Measures	Air Filtration System
	Disinfectants
	Spit Disposal
	Masks

Public Safety during Pandemic	Drone Delivery
	Specialized Isolation Units
	Mobile Support Application

	Outbreak Prediction
	Patrolling and Disinfecting Robots

Lockdown Management for Individuals	OTT
	Remote Work
	Online Learning
	Indoor Workout Platforms

Lockdown Management for Industry	Real-Time Supply Chain Monitoring
	Augmented Reality

Lockdown Management for Cities	Smart Surveillance
	Location Based Systems
	Sickness Mapping

8. Digital Ethics

Citizens may face several issues when using Digital Technology in response to current pandemic crisis. For example- issue of privacy.

Privacy is the main issue in maintaining information of users as during the process of tracing and tracking, for example in Arogya Setu app, with the help of navigation system and Bluetooth, system stores user's information of patterns and movements. When we see this from the lens of privacy, it is controversial and seems to infringe the rights of privacy. When we see it beyond the lens of privacy, it is effective measure to rapidly control and limits the spread of crisis.

Also, we must be aware about the content we received through online channels as they are regulated by AI without human in the loop. These kinds of contents are managed by automated systems. All of us should do what we can to maintain the level of solidarity during this pandemic period.

9. Conclusion

As we know that Digital technology can be of great help in pandemic but there is another side of current world also. When we talk about digital health initiatives, it amplifies socio-economic imbalance and disparities.

Government of India is focusing on tracking and tracing with use of smartphone. However, the facility of internet is still not the thing of pocket for everyone, which is far reaching the masses of lower-class people. Those who have cell phone provided by state or central government are not even having money to buy food in this pandemic period. Although, public hotspots installed by the government, that can be a temporary option but it's not feasible at all, as person can't go and remain seated on main road to access internet.

Technologies in healthcare can help doctors by providing various kind of advanced facilities which will further help them to assess the situation and handle it accordingly. With the availability of technologies, Covid-19 as well as other diseases can be eliminated by fighting misinformation, finding drugs, increasing traceability and transparency by sharing data, tracking people with facial recognition and big data, contactless movements and deliveries through autonomous vehicles, drones and robots; technology supported temperature monitoring, remote working technologies to support social distancing and maintain business continuity. Also, it can be helpful in reducing inequalities, minimizing the risk of virus spread and easy access to public services during the pandemic crisis.

In nutshell, it is **concluded** that in today's situation, the greatest risk is not the weapons, instead a worldwide covid-19 pandemic. If not cared properly it is able to kill people at large. In the line, the researchers with great dedication are working to find every possible solution available to fight Covid-19 and digital technology is the fascinating road in this research. Covid-19 made us to realize that we have all advancements but actually, we are not even geared-up to tackle a virus. Presently, India is fighting Covid-19 with around 60% rate. But as the lockdown lifted, cases started increasing at surprising rate. We should understand that lockdown is not the ultimate solution. We should move towards working on technological advancements as well as on research. After Covid-19 outbreak, it is proven that, from Artificial Intelligence to Big Data to IoT to Robotics, technological innovations are helping to manage the pandemic and better equip to fight future public health emergencies in perfect and timely manner.

10. Recommendations

a) Strict monitoring and surveillance with tracking and tracing at society level with apps like Arogya Setu and robotics at health-care organization level for procedures which involves interaction with Covid-19 patients.

b) Careful balance between data privacy and public health in AI as AI uses user's data, better data processing in IoT and better data quality in deep learning to provide enhanced care.

c) Drones are being used in major cities to supply medicals in outbreak affected area. We should use them in rural areas also where we can found issues of electricity and proper roads; monitoring through camera is already happening in areas where people are not following lockdown rules. Along with this, drones can be used to check temperature of patients in outbreak affected area; for inspection of areas to know the status of lockdown or curfew rules in narrow, affected streets and to broadcast notification released by government or local authorities.

d) Increase efficiency and creativity with blockchain in retail management and patient management and in Virtual Reality which provides facility in which one can learn things by actually doing it.

e) Cloud Computing should have Increased security, Optimized storage environment, Optimized end-user experience making the platform more user-friendly, so that it attracts more number of people.

f) Remote healthcare can be provided in form of telehealth, wearable sensors and remote monitoring and diagnosis in the areas where direct medical facility is not available or visiting health-care facility for regular check-up is not possible because of distance factor.

g) Lockdown management at different levels in the form of remote working with cloud networking; online learning through e-learning platforms like Edx, Coursera and Swayam; mapping and real-time supply-chain monitoring with navigation based system and technologies.

PART 2- ONLINE COURSE- DATA ANALYTICS FOR LEAN SIX SIGMA

ABOUT THE COURSE

1. Name of the Course- Data Analytics for Lean Six Sigma.

2. Course Description- This course helps in data analytical techniques which are useful in lean six sigma projects with the help of data analytical tools like Minitab and various interpretation methods.

3. Course Objective- To learn about data analysis and lean six sigma, understanding and visualizing data, probability distribution and testing.

4. Course Contents- Process improvement, lean six sigma as a process improvement method, DMAIC, critical to quality (CTQ), descriptive statistics, visualization of data, sampling, estimation, causality and data analysis.

5. Learning Methodology- Whole course consists of video lectures, projects, cases, reading material and some data files for practice exercise, all of them have some weightage in learning the course. It also provides learning through Minitab, although, other methods can also be used.

6. Key Learning from the course- Analysis of data and interpretation of data using various techniques and Brief about lean six sigma.

CONTENTS OF COURSE

The course content is divided into 5 weeks and 7 modules. Each module consists of videos, projects and reading materials, basis of which performance of every week is measured by attempting a graded quiz.

Module 1- Module 1 is about Data and lean six sigma

Processes- Processes are used to make a product with involvement of series of action or various steps to deliver services. Processes are present everywhere, like in industry, in encapsulation process of medicine. Here a medicine is put into capsules. Another example is in health care, where a typical process is the treatment of the patient in a hospital. There are also many other areas available in which there can be processes and services involved.

Process Improvement- Processes can be improved with the help of Lean Six Sigma. This is a stepwise method for running a process improvement project. Lean Six Sigma is based on process data and on evidence. This processed data can be used in various projects to analyse several aspects of the process.

For example, we can use the data to find the biggest problems that we should focus on. In the example of the encapsulation process, data on the major production stops showed that from all possible stops, the stop reason waiting for gelatin takes up most of the stop time. This knowledge now helps to focus the project on it's biggest problem. Data can also be use to analyse the current process. For example, in a hospital a project can aim to reduce a number of diagnostic tests that are used. Here also, data can be used in project to investigate whether there are any improvement or not? We can also see whether our handling time is influenced by giving our employees a certain training or not.

Overall, process improvement is important to organizations and Lean Six Sigma is a popular method to do this. Data can be widely used to improve the processes of any project. Data can be helpful to find the biggest problems, diagnose the current processes or to investigate the effects of the improvement actions.

DMAIC- Lean six sigma is a data-driven approach to process improvement and it helps to find-out the use of DMAIC process in data and data analytics.

DMAIC is a kind of roadmap which is offered by lean six sigma. The five phases of DMAIC process prescribes a stepwise procedure for process improvement based on measurements and evidence based improvement actions.

In the DMAIC-

D stands for Define, M stands for Measure, A for Analyse, I for Improve and C stands for Control.

In the Define phase, the belt selects a project and establishes project objectives. In the Measure phase, the belts makes our problem quantifiable, and measures the process performance. Analyse phase is the phase, in which, the belt analyses the current situation and makes a diagnosis. After diagnosis of the current process a belt develops, and implements improvement actions based on vital influence factors. In the control phase, the process controls are improved and the project is closed. This is a vital phase of any project, because no-one wants that the process slips away to its performance level before the improvement.

Critical to Quality- A CTQ or critical to quality is a property of a product or a service. In Lean Six Sigma terminology, this metric is called a CTQ. Sometimes, it is also called Y-variable or dependent variable or the outcome variable. There are often multiple ways in which we can select our CTQs. Always be assured during CTQ selection process and look whether they've captured the desired project objective or not. A CTQ is a property of a process, product, or service that is relevant to the project objective. We can find analysing of what would we want to change?

For example, if we run a project on processing address changes in an insurance company and management feels that current procedure of address change is too much time taking, both for the business as well as for the customer. What will be the relevant CTQ here? Main aspects we want to improve is the time that it takes employees to process these address changes. This can, according to example, be measured by looking at the CTQ processing time per address change.

Another aspects is the customer time which can be measured by looking at the CTQ throughput time, which is the time it takes for the address change to be processed. This includes the time a customer takes to fill in the form, the waiting time, and the processing time. In most service processes, processing time is measured in minutes, whereas through-put time is often measured in days.

Sampling- Sampling can be easily understood with an example of measurement of speed with speed camera. Using a speed camera, we can measure the speed of vehicles passing by. We cannot measure all vehicles passing by, as it is impossible to measure the speed of two cars that pass by at the same time. Or there may be one car just behind to another car, which is also not traceable. The solution to this is the sampling process. A sample should be representative of the whole cars passing by, according to this example. A representative sample means that the sample should be based on a mechanism that is independent of the question that we want to answer. We should never randomly select the vehicles by ourselves. Because as humans we are not programmed to make random selections. This kind of selecting is called sampling.

Overall, we have seen that sometimes it is necessary to sample and sometimes it is important to select a representative sample. One way to obtain a representative sample is by randomly selecting the items. However, we should never do this also by ourselves, and should always use a computer to make a random selection.

Module 2- Module 2 is about understanding and visualising the data

Types of data- There are two types of data available- Numerical and Categorical.

If we focus on numerical, we can make a subdivision between discrete and continuous data. If we focus on categorical, we can make a subdivision between ordinal, nominal and binary data.

Numerical data, also known as quantitative data, is data based on numbers that can be calculated with, such as temperature, dimension, or the number of people in a room. It can make a distinction between discrete, which is count data, or data that is rounded, and with variables that are continuous, like temperature, that can be any kind of number.

Categorical data is qualitative data. Hereby, It can divide the data into three groups-

If there is a meaningful order to the groups, it is ordinal data.

If the groups consist of unordered categories, it is nominal data.

If you only have two options available, we will call it binary data.

Numerical data contains more information than categorical data. Hence, if we have choice to choose, we should always collect the data in a numeric way.

There are two main reasons of this preference- for categorical data, we require more observations than for numerical data. As a rule of thumb, we can use that for a numerical CTQ if we have 30 observations, while for a categorical CTQ, we need a lot more, at least 300 observations. The second reason to distinguish between numerical and categorical data is a very practical one, it determines which tool we should use. If our Y variable is numerical, we can make a histogram or a boxplot. However, if the variable is categorical, we can make a pie chart, bar chart, or a Pareto chart.

There are two reasons for making a distinction between numerical and categorical variables.

First, it contain different levels of information and secondly, we can easily know which statistical tool can be applied

As we know that numerical data are numbers, while categorical data are classes and categories, descriptive statistics are only available for numerical data because they are based on mathematical operations like addition and multiplication, while, for categorical data, we can make a table.

Whenever we take data, there is always a certain degree of variation or dispersion in the data. A low amount of dispersion means that most failures are rather close to the mean. A high amount of dispersion means that most

failures are quite far away from the mean and spread out over a larger interval. To measure dispersion we should always look at the standard deviation.

Overall, we can say that the mean and the median represents something about the location of the data. And the standard deviation and the variance say something about the dispersion in our data. The range and the interquartile range do that too.

Types of graphs and charts- There are two types of graphs discussed in this course-histogram and a boxplot.

Graphs are often used to quickly and easily understand the measurements of a variable that we have. The histogram and boxplot are graphs for a numerical variable. When there is numerical variable available, the histogram and the boxplot are the most popular methods for visualization of this variable.

Histograms can be used to review interesting patterns in data. It can be used to see whether the data is symmetrical or skewed, or even whether it is bimodal. Bimodality usually occurs when two distinct populations are present in the sample and giving a mixture of distributions.

Data usually has tails, which are clearly visualized in a histogram. The larger part of the data that is concentrated together is called the bulk and the remaining parts on both sides are called the tails. Tail part of any data decreases outward.

Boxplots shows us the median, the bulk, the whiskers and possibly outliers. A boxplot reveals exact locality of data on the real line. It shows us the minimum and the maximum value. If there is one outlier, then the second highest value determines where the whisker will stop. The boxplot also shows us the interquartile range. The interquartile range is equal to the length of the box and that is equal to the difference between the third quartile and the first quartile, The boxplot shows us the total range which is equal to the deduction of minimum from maximum.

Overall, we can make a histogram or a boxplot to visualize a numerical variable.

A histogram shows us the variable and gives information about the distribution of data.

A box plot also gives us information about the spread in data and about outliers.

Other techniques includes, pie chart, which shows us how often each category occurs. And a bar chart gives us the same information, but it is easier to spot which category occurs most often. A tally table gives the exact numbers. And each of these are visualization techniques for a single categorical variable.

There is one more chart pattern, known as, Pareto Chart. 80% of all problems can be attributed to 20% of the causes. This 80/20 principle is also known as the Pareto principle, and is also known as the rule of the vital few and the trivial many.

It is this rule of thumb that can be very useful in solving problems. Pareto chart is a chart in which a Pareto principle can easily be spotted. Pareto chart with the frequencies only shows

how often each problem occurs. But most of the time we are more interested in which problem takes up most of our time. This is the 80/20 principle. However, it will never be exactly 80 and 20%. Of course, it could be that, say, 70% of your problems can be attributed to only 10% of the causes or something like this in approximation.

The main principle is, there are always a few vital issues and trivial many, and make sure to focus on the vital few. So, a Pareto chart can help to distinguish the vital few causes from the trivial many. We can do this, either by counting the number of occurrences, or by weighing them with the relevant characteristic.

When the Y variable is numerical and X variable is categorical, we make a Boxplot with groups or an individual value plot. When the Y variable is categorical and the X variable is numerical we can make a transposed Boxplot with groups.

In summary, the appropriate graph to visualize two variables depends on the type of data we have. When variable X is numerical and variable Y is also numerical, the appropriate technique for visualizing the data is a Scatterplot. If the variable X is categorical and variable Y is numerical, we can make a Boxplot. If the variable X is numerical and variable Y is categorical and, we can make a transposed Boxplot. If variable X is categorical and variable Y is also categorical, we should make a cross tabulation or Stacked bar chart.

Module 3- Module 3 is about Probability Distribution

Sampling- From the whole group of population, a small group is selected which is known as Sample. A population is a collection of all units of interest and the sample is a subset of the population.

For example, in case of analysis of customer satisfaction, we want to know whether our customers are satisfied with our company's product, and therefore we wish to measure customer satisfaction. Suppose, our organization has some 40000 customers. These 40,000 customers are all the customers that our organization sells to, so they are the population of interest. To know the satisfaction level of these customers, we can select 100 customers as sample from the population. each customer in the sample is one measurement represented by an X. So, X1 for the 1st customer, X2 for the second customer in the sample until we have 100 customers in our sample, X100. To each customer in the sample we will ask the question, are you satisfied with our product? And their answer will be our data. Then we have to select which units, or customers, you want to have in your sample, and

this should be done randomly to get a representative sample. If 75% are found satisfied with product then it is 30000 satisfied customers.

Overall, the population is all units of interest. We select the sample. We collect data on the units in sample, we can analyse this data using descriptive techniques, such as histogram, the mean, and the standard deviation. However, in the end, we want to draw conclusions about the entire population, and we will use the sample distribution to estimate the population distribution and we will use the sample statistics to estimate the population parameters. For now, it should remember that a sample is a subset of the entire population, and the sample is used to say something about the population.

Estimation and Confidence Interval- It is then when we do not know the exact answer or value, but we can make an educated guess. When we have some estimates but with uncertainty, we calculate two boundaries in-between which clearly shows that the actual population mean will lie and this is called the confidence interval. The most often selected confidence level Is 95%.

We can also construct confidence Intervals for the median or the standard deviation or any other statistic. If we would have a larger sample, the estimates will be more precise and our interval, will become smaller. Overall, we use estimates to approximate an unknown population parameter. These estimates will never be exactly equal to the population parameter because it is based on sample. Therefore, to quantity the uncertainty around the estimate, we calculate confidence intervals. A confidence interval gives us boundaries in which a population parameter will lie with a certain level of confidence.

Normal, Lognormal and Weibull Distribution- When the number of measured values is less, we need some imagination to recognize the normal distribution. However, as the number of measured values becomes larger, the distribution converges more and more to these bell shapes. The more N increases, the more it starts to look like a bell shape. We can use a normal distribution to make some simple calculations. In case of throughput times, normal distribution curve does not fits well. The histogram and a normal distribution curve have different shapes. And therefore we say that a normal distribution does not fit the data. In this case, we use second type of distributions, this is the Weibull distribution, and it is also called a skewed distribution. A distribution like this is called skewed to the right, because the tail is to the right. The Weibull distribution is often used for data, like, throughput times, and processing times, because these are often skewed variables. Many values are relatively small, and some values are very high. The Weibull distribution can take various forms. When it comes to Lognormal distribution, it can also be used in places where normal distribution does not fits well. It depends on the condition.

Though Normal Distribution occurs very often. The Normal Distribution has some interesting properties which facilitate rules of thumb to compute relevant percentages. Normal Distribution is also often called the Gauss curve because Carl Friedrich Gauss (1777) is often associated with this distribution and its form. Normal

Distribution has two parameters that determine the shape of its distribution. The first parameter is the location parameter which determines the midpoint. This is denoted by mu, which is the location or mean. The second parameter describes the shape of the distribution and is known as the dispersion parameter, denoted by sigma.

And it determines the spread of the distribution. The location of a Normal Distribution can be changed by manipulating mu. It is also possible to change the scale of the distribution by tweaking the sigma.

So, we can say that, the normal distribution is seen most often. But the Weibull distribution and the lognormal distribution are relevant in Lean Six Sigma project as well.

Probability Plot- Probability plot is a powerful tool to determine which distribution fits best, and it can be used to spot these anomalies. It can also help us to detect bimodality. A probability plot can be used to find a theoretical distribution that fits best to your data. If many data points deviate from a straight line, and do not fall between the two outside lines, then we call it a poor fit. If the data lies more or less on the straight line and many of the points lie between the two outside lines, we call it a good fit.

Empirical CDF- To calculate probabilities and percentages, a tool is there, called the empirical CDF. In DMAIC, this tool come in the analyse phase. The empirical CDF is used to determine how many percent of the cases meet our service level agreement, or our specifications?

Module 4- Module 4 is about Basics of Data Analysis and Testing and Causality

Data Analysis- Data-based testing of ideas and improvements is one of the principle of Lean Six Sigma. Also, this principle is sometimes referred to as evidence-based improvement actions. Usually, we start analysing data if we want to know if two different concepts are related to each other. In a Lean Six Sigma project, we also want to know whether our CTQ is related to an influence factor. If we look at the DMAIC phases of our project, this question will be part of the improve phase where we want to establish the effect of an influence factor on our CTQ in order to develop and implement an effective improvement action. To analyse, if two variables are related, we need to determine which variable is our influence factor or our X variable and which variable is our CTQ or our Y variable. Next, we can model the relationship using data.

The suitable method depends on the type of data that we have. It depends if our CTQ and our influence factor are categorical or numerical variables. For example, If the Y variable is numerical and X variable is also numerical, we should perform a regression analysis.

If Y variable is numerical and your X variable is categorical, we should perform an ANOVA. A special case of the ANOVA is a 2-sample t-test. If the assumptions underlying the ANOVA are not met, we can perform a

Kruskal-Wallis analysis. If the Y variable is categorical and your X variable is numerical, we can perform a logistic regression. Finally, if the Y variable is categorical and your X variable is also categorical, we can perform a chi-square analysis.

So, to analyse if two variables are related, we need to determine which variable is the influence factor, or X variable, and which variable is CTQ, or Y variable. Next, we have to determine whether our variables are categorical or numerical. Once determined what type of data we have, the tree diagram can be a very useful tool to determine which type of analysis method is suitable for our data. And these are the methods that we can use for that data-based testing of ideas and improvements in the Lean Six Sigma project.

Hypothesis Testing- We know that estimation is to describe a population as accurately as possible using our sample data. We estimate parameters that describe the entire population. An estimate is merely descriptive. Therefore, we often switch from estimation to testing. Testing means that we claim something. And we use data to test this claim. This claim is called a hypothesis. It is a procedure to arrive at a decision. We do this as objectively as possible. And it deals with uncertainty in a rational manner. More advanced statistical techniques can even be used to quantify the risk of a false decision.

We begin always with a null hypothesis, which is the hypothesis that there is no effect between the variables you are testing. And we state our alternative hypothesis, which is the hypothesis that there is an effect between variables. Next, we collect evidence, which will be mostly in the form of data. After collection of data, we can calculate our p-value. This p-value will allow to decide whether or not to reject the null hypothesis. If the p-value is small enough this means that we have found a significant effect, that H0 is unlikely and that we reject this null hypothesis. If a p-value is small enough, it's quantified as saying the p-value is smaller than 0.05 which is a common used threshold. If the p-value is larger than this threshold of 0.05, we do not reject null hypothesis. One can think of 100s of different hypotheses. Every hypothesis has its own testing procedure to calculate a p-value.

Summarizing, hypothesis testing is a procedure to arrive at a decision as objectively as possible, while dealing with uncertainty in a rational manner. We start with constructing our hypothesis, then we collect data. And finally, draw conclusion.

Causality- Correlation does not imply causation. Correlation means that two things move in the same direction, they show correlation. Whereas causation means that the one thing causes the other to change.

Causality is about cause and effect relationships. This means when we change our X variable and nothing else, Y variable will change as a result. This is often very straight forward. For example, if children grow, they also become heavier, hence growing or becoming taller causes their weight to go up. We use an arrow to show this causal relationship. However, sometimes, the causal relationship can be a little less straight forward.

There are four possible situations of causality where it can be confused with correlation-wrong direction of causality, time problem, the third variable problem and the underlying variable problem. Therefore, we need different ways to detect correct causal relationships.

One option is to study the literature that argues why, or why not, the certain relationship is causal. Or we can use logical thinking, and time order- which variable comes first, and which variable follows.

In summary, we can say that there are four different types of errors we can make when we assume causality. We can assume a causal relationship in the wrong direction, we can relate two variables to each other that are actually unrelated but change the same way over time, we can have a third variable that causes two variables to change at the same time, we can have an underlying variable which is actually the cause of the changes in Y instead of your initial X variable. We can avoid making these errors by performing a controlled experiment or using additional literature that explains to you the cause and effect relationship. Logical thinking might also help to avoid making these errors.

Module 5- Module 5 is about Testing: Numerical Y and Categorical X

ANOVA- ANOVA stands for analysis of variance. ANOVA is the appropriate statistical technique when the Y variable is numerical and the X variable is categorical. ANOVA is used to study the relationships between a numerical Y and a categorical X variable. ANOVA is a suitable statistical method if you are comparing means across various groups.

ANOVA can be performed in three steps, the first step is to organize your data. Second step is performing the ANOVA, and third step is performing residual analysis just to check if the assumptions are met.

As a first step, The X and Y variables should have their separate columns. Each row in the data set needs to contain one unit. Second step is p-value test. The statistical tests involving p-values are formally called hypothesis tests. For this ANOVA analyses the hypothesis that all group means are equal is our null hypothesis. The hypothesis that there is a significant difference between the group means is the alternative hypothesis.

And that before we can be sure that the conclusions in our second step are valid, you will have to perform a residual analysis. The residual analysis is a vital part of any statistical method.

The P-value we use in a main analysis is only valid if the assumptions are satisfied. These checks are called the residual analysis, and this is the last and final step of ANOVA.

As we know, ANOVA consists of three steps in total. To validate the assumptions, we will check if the residuals are normally distributed and if there are any outliers or other irregularities present. Please note that the residuals are calculated by subtracting the expected value from each observation.

If conclusions of step two are not found valid in step 3 then we can proceed to a Kruskal-Wallis analysis.

In summary, we can say that ANOVA is a technique to test whether a categorical influence factor X has a significant effect on a numerical Y. After organizing data in the first step, we will run the analysis in the second step and interpret the p-value for significance, and the r squared for importance. In the third step, we will validate the conclusions by checking whether the residuals are normally distributed and whether they don't contain any outliers or other strange patterns.

Kruskal-Wallis Test- This technique is an alternative to ANOVA technique, and hence, the Kruskal-Wallis test is used to analyse the relationship between a numerical CTQ and a categorical influence factor. The Kruskal-Wallis test is a non-parametric test. This means that there is no specific distribution assumed on the residuals.

As we know that ANOVA is based on the assumption that the residuals are normally distributed, and that the residuals contain no irregularities such as outliers, then, in respect to that, the Kruskal-Wallis test will be less powerful than the ANOVA analysis. This means that we will be requiring more data to show the same difference.

So, we can say that, this test should be performed when the underlying assumptions of the ANOVA analysis are not met. To interpret the output, we have to look at the medians for each group and of course, add the P-value to see if this difference is statistically significant.

Two Sample T-Test- As per the course, Two Sample T-Test is associated with William Sealy Gosset and he lived in England around 1900. He studied chemistry and mathematics and upon graduating, he started to work for the Guinness Brewery. Gosset applied his statistical knowledge, to select the best-yielding varieties of barley. He used trials. And as it is obviously very time consuming to produce barley, he could not do too many trials. Therefore, he was faced with only very small samples. To deal with this, he developed a special test to be able to draw conclusions from small samples and to be able to determine which barley had the best use. This test is now widely known. It's the two sample t-test. For confidentiality reasons Gosset wasn't allowed to publish the result under his own name. Therefore, he was forced to publish his work under the pseudonym Student. And this is why the two sample t-test is now also known as Student's t-test.

The two sample t-test is a method to compare means of two groups.

We can say that the two sample t-test is a very basic test to analyse the differences between means of two groups. The p-value can be used to calculate whether the differences of the two means is coincidental, or truly significant. Especially for small samples this is very useful.

The two sample t-test is a very old test and forms the building block of many other tests that have been developed since.

Module 6- Module 6 is about Testing: Numerical Y and Numerical X

Correlation- Correlation is an index that expresses how strong a relationship between two numerical variables is. This index is always between minus 1 and plus 1. Please note that a correlation of 0 means that there is no linear relationship. A positive correlation means that higher values of one of the variables will generally correspond to higher values of the other variables. A negative correlation means that the variables are correlated in the opposite direction. So higher values of one of the variables generally corresponds to lower values of the other variable.

Methods to interpretate the correlation coefficient- There are some rules of thumb-

A correlation coefficient from 0 to 0.4 means that there is no or a weak correlation between the variables. A correlation coefficient from 0.4 to 0.8 means that there is a moderate correlation.

A coefficient from 0.8 to 1 means that there is a very strong relationship. If the coefficient is exactly 1, we speak of a perfect correlation between the two variables. However, in reality, it rarely occurs that two variables correlate perfectly, and, as a final warning, always be careful not to mix up correlation with causality.

Regression- Regression analysis is a very popular technique within statistics. When the Y variable's numerical and the X variable's also numerical, regression analyses is the appropriate technique. Regression analysis is a statistical method to test if two numerical variables are related by calculating the best fitting line.

Regression analysis consists of four steps- First, we make a fitted line plot to visualize the data. The actual calculations or performing the main regression analysis is the second step. At this step, we also look whether the relationship is statistically significant, and if so, how strong the relationship is. The third step is about performing a residual analysis. And in the fourth step, we can calculate the prediction interval. This step is optional.

So, the fitted line plot indicates the constant a and the slope b of our line. The p-value indicates whether the findings are due to chance or whether they are a structural effect and the R squared shows you how strong the relationship is.

Before we can completely trust our conclusions however, we have to verify the underlying assumptions of regression analysis. And this is called the residual analysis. We have to check if the residuals are normally distributed and if the residuals do not show any outliers or other irregularities.

If the residuals are normally distributed and show no outliers or irregularities, this means that our assumptions are met. If the assumptions are not met, we have to be careful to interpretate our results.

There are prediction intervals also. They are used to make predictions about data that is not in our sample. The prediction interval can be used to determine the level of influence factor, which ensures a certain CTQ performance. We can use crosshairs, or the formula, to determine the required value of the influence factor.

Quadratic Regression- If the linear regression analysis does not meet the assumption, we will have to apply an alternative analysis technique. One of these alternatives is the quadratic regression analysis. This technique fits a curved line to our data instead of a straight line. After performing the regression analysis, we always have to perform the residual analysis again to see if the assumptions are met.

Quadratic regression model is one possible extension to the simple linear regression model we already know. If we perform a quadratic regression analysis, it means that we need to add an extra term to our formula. This changes the shape of the fitted line. It makes it curved.

Module 7- Module 7 is about Testing: Categorical Y

Chi-Square Analysis- If we have a Y variable or see the queue that is categorical and an X variable or influence factor that is also categorical, we have to perform a chi-square analysis.

Steps of the chi-square analysis- The first step is to make a cross tabulation. The second step is to take a look at the p-value. The p-value shows, the effect you found is significant or not.

So, we can say that A Chi-Square analysis is a suitable method if we have a categorical Y, or CTQ, and a categorical X variable, the influence factor.

Logistic Regression- To interpret the output, when we have a categorical Y variable and a numerical X variable, the suitable method is a logistic regression.

Logistic regression consists of three steps- In the first step, we may need to convert data to the correct format. In the second step, we make a fitted line plot and evaluate the quality of the fit. In the final step, we analyse the p-value and see if we have found a significant relationship.

KEY LEARNING

1. Data analytics techniques that are typically useful within Lean Six Sigma improvement projects.

2. Use of data analytics tools and the interpretation of the outcome.

11. References

1. IIHMR University Student Paper submitted to library
2. **www.un.org**
3. **www.studymode.com**
4. **https://sundial.csun.edu/156361/news/a-timeline-of-outbreaks-from-2000-to-present/** A timeline of outbreaks from 2000 to present
5. **https://www.nature.com/articles/s41591-020-0824-5** Digital technology and covid-19- Daniel Shu Wei Ting, Lawrence Carin, Victor Dzau and Tien Y. Wong
6. **https://journals.sagepub.com/doi/10.1177/2055207620920083** Covid-19-A new digital dawn?
7. Wikipedia
8. Technology can help alleviate Covid-19 crisis- Dr Chris Hobsten- Hamilton Spectator
9. **https://ww2.frost.com/frost-perspectives/ten-digital-technologies-helping-humans-in-the-fight-against-covid-19/** -Ten digital technologies helping humans in the fight against covid-19- Hiten shah, Kiran Kumar
10. **https://www.researchgate.net/publication/340574089_Digital_technologies_IoT_Big_Data_AI_and_Deep_learning_for_Covid-19** Digital technologies for Covid-19-IoT, Bigdata, AI and deep learning- Marwa F. Mohamed
11. **https://cio.economictimes.indiatimes.com/news/next-gen-technologies/covid-19-8-ways-in-which-technology-helps-pandemic-management/75139759** Covid-19: 8 ways in which technology helps pandemic management
12. **https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7204706/** - A review of modern technologies for tackling COVID-19 pandemic- Aishwarya Kumar, Puneet Kumar Gupta, and Ankita Srivastava
13. **https://www.startus-insights.com/innovators-guide/6-ways-technologies-tackle-the-global-coronavirus-pandemic/** -8 Ways Emerging Technologies Tackle The Coronavirus Pandemic
14. **https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30142-4/fulltext** Application of Digital Technology in Covid-19 Pandemic Planning and Response
15. **https://www.uclg.org/sites/default/files/eng_briefing_technology_final_x.pdf** Digital Technologies and the Covid-19 Pandemic
16. **https://www2.deloitte.com/nl/nl/pages/risk/articles/covid-19-and-technology.html** -Covid-19 and digital technologies
17. **https://www.sciencedirect.com/science/article/pii/S1871402120300771** Artificial Intelligence (AI) applications for Covid-19 pandemic

18. **https://cointelegraph.com/news/how-blockchain-technology-can-help-fighting-against-covid-19** - How Blockchain technology can help fighting against covid-19
19. **https://cointelegraph.com/bitcoin-for-beginners/how-blockchain-technology-works-guide-for-beginners** - How Blockchain technology works?
20. **https://www.mydronelab.com/blog/what-is-a-drone.html** -What is drone?
21. **https://www.sciencedirect.com/science/article/pii/S1871402120301065** - IoT Applications to fight against Covid-19 pandemic
22. **https://www.sciencedirect.com/science/article/pii/S1871402120301302** -Significant applications of virtual reality for Covid-19 pandemic
23. **https://www.sciencedirect.com/science/article/pii/S0010482520301736#abs0010** – Covid-19 detection using deep learning models
24. **https://www.geospatialworld.net/article/the-sino-approach-use-of-technology-to-combat-covid-19/** -The China way: Use of technology to combat Covid-19
25. **https://www.geospatialworld.net/blogs/covid-19-how-robots-help-in-fight-against-covid-19/** -How robots are helping in fight against Covid-19
26. **https://www.outlookindia.com/website/story/opinion-big-data-has-potential-to-save-lives-in-covid-times-but-proceed-with-caution/353390** -Big data has potential to save lives in covid-19 times, but proceed with caution
27. **https://www.lexology.com/library/detail.aspx?g=f8c441c9-ccdf-44a1-9661-b3db151fc1af** The use of facial recognition technology to combat covid-19
28. **https://www.zdnet.com/article/what-is-cloud-computing-everything-you-need-to-know-about-the-cloud/** -What is cloud computing?
29. **1. All materials are published under a creative common license and are copyright of the University of Amsterdam. All material and videos have been created and developed by IBIS UvA, the Lean Six Sigma institute of the University of Amsterdam (www.ibisuva.nl).**
30. **2. H Kieler [1], O Axelsson, S Nilsson, U Waldenström - The Length of Human Pregnancy as Calculated by Ultrasonographic Measurement of the Fetal Biparietal Diameter**