International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

PUBLIC HEALTH COVID-19 CASE TRACKER

[1] B Jyothi, [2] P Roshan, [3] L Bhuvan Chandra, [4] G Sachin
[1]AssistantProfessor, [234]Students
Department of Data Science
Siddhartha Institute of Technology & Sciences, Narapally
Bjyothi.cse@siddhartha.org.in, 23TQ1A6739@siddhartha.co.in, 23TQ1A6722@siddhartha.co.in, 23TQ1A6703@siddhartha.co.in,

ABSTRACT

The Public Health COVID-19 Case Tracker is a data-driven system designed to collect, analyze, visualize, and monitor COVID-19 case information. During a public health emergency, large amounts of information may be generated from different regions and time periods. A centralized tracking system can help organize this information and present it in an understandable format.

The proposed system processes information such as confirmed cases, recovered cases, active cases, deaths, vaccination statistics, testing information, geographical location, and reporting dates. Data preprocessing techniques are applied to handle missing values, duplicate records, inconsistent formats, and invalid entries. The cleaned information is stored and prepared for further analysis.

The system provides an interactive dashboard containing important indicators such as total cases, active cases, recovered cases, deaths, vaccination counts, and daily case changes. Users can analyze cases based on states, districts, countries, or selected time periods depending on the available dataset. Charts, graphs, maps, tables, and KPI cards make public-health information easier to understand.

The tracker can also identify case trends and changes over time through time-series analysis. Regional comparisons can help users understand differences in reported case counts and recovery patterns. Additional analytical modules can be incorporated to identify unusual increases in reported cases and provide statistical summaries.

Overall, the Public Health COVID-19 Case Tracker provides a centralized platform for organizing and visualizing COVID-19-related data. It can support researchers, public-health analysts, administrators, and other authorized users in understanding historical and current data patterns. Future enhancements can include real-time data integration, forecasting models, automated alerts, vaccination analysis, and integration with additional public-health datasets.

I. INTRODUCTION

COVID-19 created a significant need for timely collection and analysis of public-health information. Health organizations and governments collected information about reported cases, recoveries, deaths, testing, vaccination, and geographical distribution. Organizing this information effectively is important for understanding the progression of an infectious disease.

Case counts can change significantly across different regions and time periods. Analyzing these changes can help identify periods of increasing or decreasing reported cases. Historical analysis can also provide useful information for researchers studying the progression and impact of the pandemic.

Traditional public-health reporting may involve spreadsheets, static reports, and information collected from multiple sources. When datasets become large, manually processing and comparing information can be time-consuming. Differences in data formats and reporting schedules can also make analysis more difficult.

Data analytics and visualization provide effective methods for presenting public-health information. Interactive dashboards can display case counts, trends, regional

comparisons, and other indicators in a single interface. Users can apply filters to examine particular locations or periods and obtain focused analytical views.
The proposed Public Health COVID-19 Case Tracker combines data collection, preprocessing, analysis, visualization, and optional predictive capabilities. It provides a centralized dashboard for examining reported COVID-19 data and identifying important trends. The system is intended as an analytical and educational platform and should rely on authoritative health data for real-world use.

II. LITERATURE SURVEY

Public-health surveillance systems have been widely used to collect and analyze information about infectious diseases. Such systems typically maintain records related to reported cases, geographical locations, dates, testing, hospitalization, recovery, and mortality. The collected information can be used to understand disease patterns across populations and regions.
During the COVID-19 pandemic, digital dashboards became widely used for communicating case information. These dashboards commonly displayed confirmed cases, recoveries, deaths, vaccination statistics, and geographical distributions. Visualization made large datasets easier for users to understand without requiring advanced data-analysis knowledge.
Time-series analysis is an important technique for infectious-disease monitoring. Daily and weekly case data can be analyzed to identify increases, decreases, peaks, and other temporal patterns. Moving averages and other statistical techniques can be used to reduce the effect of short-term fluctuations in reported data.
Geographical analysis is also useful in public-health monitoring. Case counts can be grouped by country, state, district, or other geographical areas depending on the available data. Maps and regional charts can help users compare reported case patterns between locations.
Machine Learning and statistical forecasting techniques have also been explored for infectious-disease prediction. Historical case data, environmental information, mobility data, and other variables can be used to develop forecasting models. However, predictions depend heavily on data quality, reporting practices, assumptions, and changing conditions.
Modern public-health analytics systems increasingly combine data integration, visualization, statistical analysis, and predictive modeling. Interactive dashboards allow authorized users to examine different dimensions of health data efficiently. The proposed COVID-19 Case Tracker follows this analytical approach by integrating case-data processing, temporal analysis, regional analysis, visualization, and a foundation for future forecasting capabilities.

III. SYSTEM ANALYSIS

The Public Health COVID-19 Case Tracker is designed to process reported COVID-19 data and generate meaningful public-health insights. The system accepts information such as confirmed cases, active cases, recoveries, deaths, vaccination data, location, and reporting date. The input dataset is first validated to ensure that required fields, formats, and data types are appropriate. A preprocessing module handles missing values, duplicate records, inconsistent formats, and invalid entries. Date and geographical fields are standardized to support accurate analysis. The cleaned data is stored in a structured database for efficient retrieval. The analytics engine calculates total cases, active cases, recovered cases, deaths, and other available indicators. Time-series analysis identifies changes in reported cases across different dates and periods.

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Regional analysis compares reported case counts across available geographical areas. Visualization components present results using charts, graphs, maps, tables, and KPI cards. Optional statistical or Machine Learning models can analyze historical patterns for forecasting purposes. The final system provides an interactive dashboard through which authorized users can explore public-health data and analytical results.

Existing System

Existing COVID-19 tracking systems generally collect and display reported case information from public-health organizations and other data sources. Many systems provide basic statistics such as confirmed cases, recoveries, deaths, vaccination numbers, and regional case counts. These systems are useful for communicating pandemic information to the public and health-related users. However, datasets from different sources may have different structures, update schedules, and reporting conventions. Users may need to access multiple dashboards to compare information from different regions or periods. Static reports may provide limited filtering and analytical capabilities. Historical trends may not always be easy to examine through simple tables. Manual processing of downloaded datasets can require considerable time. Regional comparisons may require additional calculations and visualization tools. Some systems may focus primarily on current statistics rather than detailed historical analysis. Advanced analytics and forecasting may require separate applications. These limitations create a need for a unified analytical dashboard that organizes available COVID-19 data and provides interactive visualization and analysis.

Disadvantages of Existing System

- Data may be distributed across multiple sources.
- Different sources may use different data formats.
- Manual data processing can be time-consuming.
- Limited interactive historical analysis.
- Static reports may provide limited insights.
- Difficult to compare multiple regions efficiently.
- Limited customizable visualization.
- Data inconsistencies may affect analysis.
- Limited automated trend analysis.
- Forecasting may require separate tools.
- Limited integration of multiple health indicators.
- Manual report preparation may be required.

Proposed System

The proposed Public Health COVID-19 Case Tracker provides a centralized platform for analyzing reported COVID-19 case data. The system accepts datasets containing case counts, recovery information, deaths, vaccination statistics, locations, and reporting dates from authorized sources. The data validation module checks the structure, data types, required fields, and basic data quality. The preprocessing module removes duplicate records and handles missing, invalid, and inconsistent information. Date and location fields are standardized for reliable temporal and geographical analysis. The analytics module calculates important indicators such as total cases, active cases, recoveries, deaths, and available vaccination statistics. Time-series analysis identifies changes in reported cases over different periods. Regional analysis compares case patterns across states, districts, countries, or other available locations. Interactive charts, graphs, maps, tables, and KPI cards provide a clear representation of the processed information. Users can filter the dashboard by location, date, case category, or other available attributes. Optional statistical and Machine Learning modules can be incorporated for historical trend analysis and forecasting. The proposed system therefore provides a unified environment for COVID-19 data analysis, visualization, and reporting.

Advantages of Proposed System

- Provides centralized COVID-19 data analysis.
- Automates data preprocessing and cleaning.
- Displays important public-health indicators.
- Supports time-based case analysis.
- Provides regional comparison.
- Offers interactive charts and visualizations.
- Supports geographical visualization.
- Allows filtering by date and location.
- Reduces manual data-analysis effort.
- Supports historical trend analysis.
- Provides a foundation for forecasting.
- Simplifies public-health data reporting.

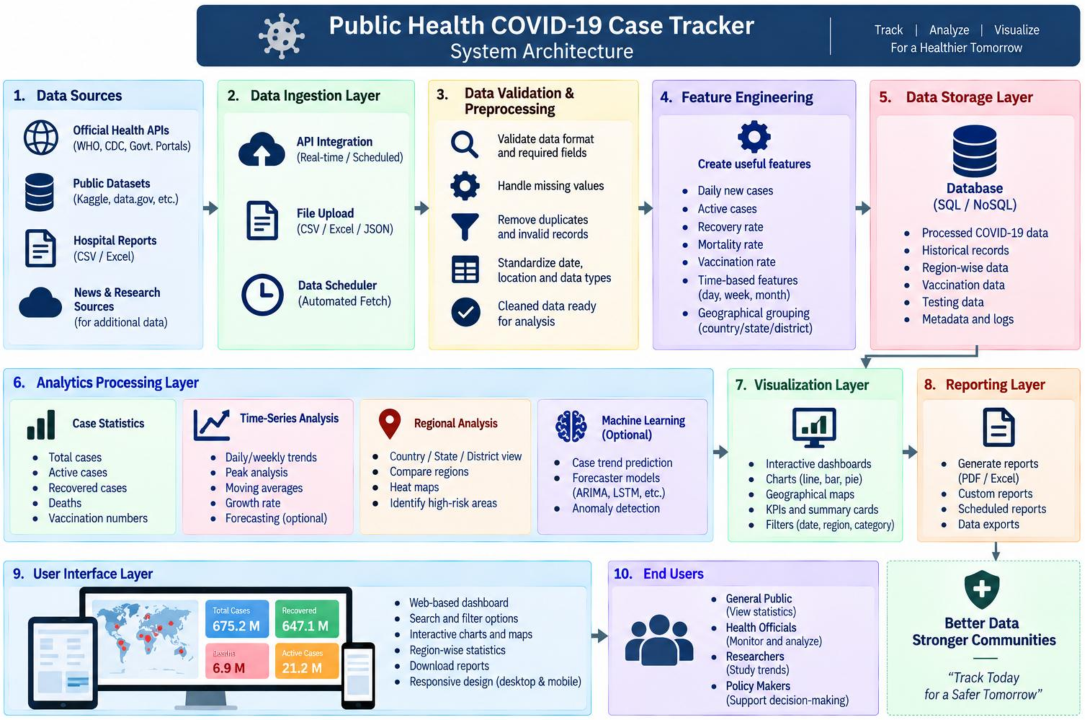
IV. METHODOLOGY

The methodology begins with collecting COVID-19 case information from authorized and reliable public-health data sources. The collected dataset is validated to check its structure, required attributes, data types, and completeness. Data preprocessing is performed to remove duplicate records and handle missing, invalid, or inconsistent values. Date fields are converted into standardized formats for time-series analysis. Geographical attributes are standardized to ensure consistent regional comparisons. Relevant features such as daily cases, active cases, recovery counts, deaths, and vaccination statistics are prepared when available. Exploratory Data Analysis is performed to understand the distribution and progression of reported cases. Time-series analysis is used to identify trends, peaks, decreases, and other changes in reported case counts. Regional analysis compares reported cases and other available indicators across different geographical areas. Visualization techniques present the results through line charts, bar charts, maps, tables, and KPI cards. Statistical analysis or Machine Learning models can optionally be trained using historical data for forecasting and trend analysis. Suitable evaluation metrics are used when a predictive model is implemented. Finally, the processed information and analytical results are displayed through an interactive dashboard for authorized users.

SYSTEM ARCHITECTURE

The system architecture consists of multiple interconnected layers that process COVID-19 data and present public-health insights. The Data Source Layer provides reported case, recovery, mortality, vaccination, testing, location, and date information from authorized public-health sources. The Data Input Layer accepts information through CSV files, Excel files, databases, or APIs. The Data Validation Layer checks file formats, required fields, data types, and basic data quality. The Data Preprocessing Layer handles missing values, duplicate records, invalid entries, and inconsistent formats. The Feature Engineering Layer creates useful attributes such as daily cases, active cases, recovery rates, reporting periods, and regional groups when supported by the data.

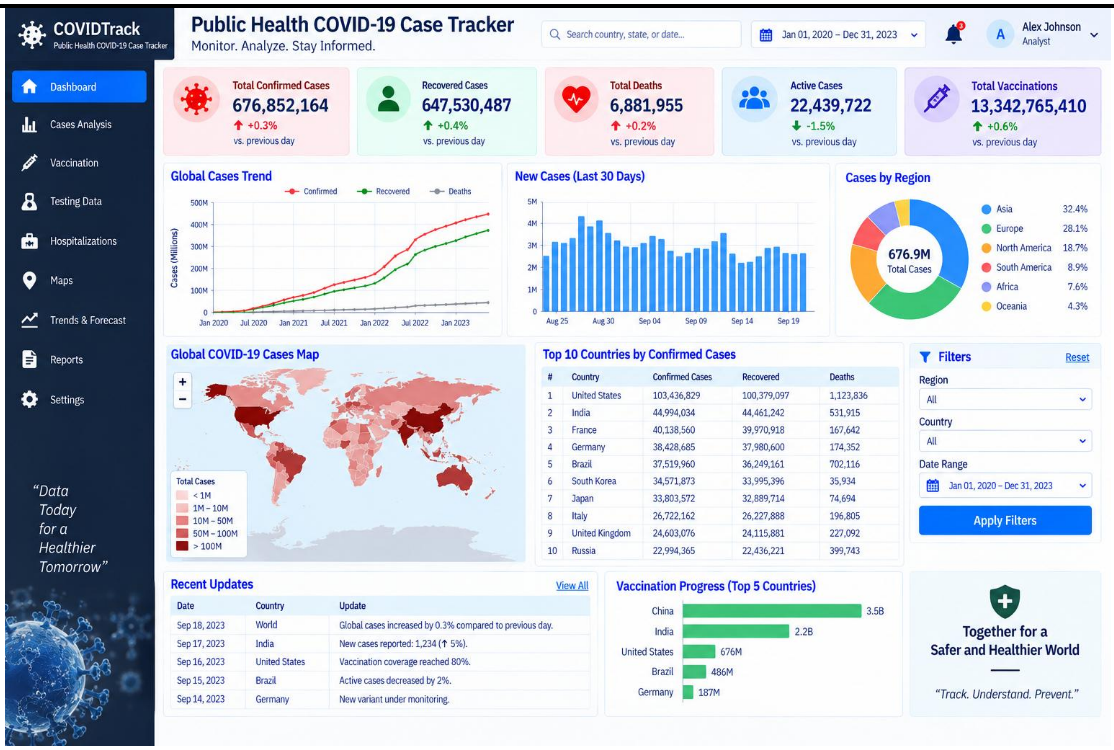
International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



The Data Storage Layer securely stores cleaned historical health records. The Analytics Processing Layer calculates case totals, active cases, recoveries, deaths, vaccination indicators, and other available metrics. The Time-Series Analysis Layer examines changes in reported cases across different dates and periods. The Regional Analysis Layer compares reported indicators across states, districts, countries, or other available geographical areas. The Visualization Layer presents results through interactive charts, graphs, maps, tables, and KPI cards. The optional Prediction Layer can use statistical or Machine Learning techniques for historical trend forecasting. The Reporting Layer generates analytical summaries and reports for authorized users. The User Interface Layer provides an interactive dashboard for exploring COVID-19 data and analytical results.

V. RESULT AND OUTPUT

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper



VI. CONCLUSION

The Public Health COVID-19 Case Tracker provides a centralized and data-driven platform for monitoring, analyzing, and visualizing COVID-19 case information. The system converts raw public-health data into meaningful insights through data preprocessing, statistical analysis, visualization, and optional predictive techniques.

The system analyzes important indicators such as confirmed cases, active cases, recovered cases, deaths, vaccination data, daily case changes, and regional distributions. This helps users understand the progression of reported cases and identify variations across different locations and time periods.

The interactive dashboard presents information through KPI cards, line charts, bar charts, maps, tables, and trend visualizations. These features make complex public-health datasets easier to understand and allow users to examine specific regions and time periods efficiently.

The system can also support historical trend analysis and, when appropriate data and validated models are available, forecasting of future case patterns. This can assist researchers and authorized public-health analysts in studying reported case trends and preparing analytical reports.

In the future, the system can be enhanced with real-time data integration, automated alerts, advanced forecasting models, vaccination analysis, geographical heatmaps, additional public-health indicators, and integration with multiple authoritative health-data sources. These enhancements can make the platform more comprehensive for public-health data monitoring and analysis.

REFERENCES

[1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine

learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.
[2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.
[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm
[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.
[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, “Real-Time Object Detection in Drone Surveillance Using YOLOv5,” in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.
[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, “Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks,” in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.
[7] R. D. Kumar, V. N. S.Manaswini, “Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology,” in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.
[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, “An advanced movie recommender using collaborative filtering and sentiment analysis,” International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.
[9] Ravi Kumar Banoth, Ramana Murthy B V, “Automatic crop recommendation system using LightGBM and decision tree machine learning models,” Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.
[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, “Smart agriculture through IoT and machine learning for analyzing carbon footprints,” in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.
[11] Ravi Kumar Banoth, B. V. Ramana Murthy, “Soil image classification using transfer learning approach: MobileNetV2 with CNN,” SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.