



# Time Series Analysis Of Monkeypox Cases

<sup>1</sup>D Kevin Harris, <sup>2</sup>T A Abubuckar Siddiq, <sup>3</sup>C Muthu Priya, <sup>4</sup>N Abdul Salam, <sup>5</sup>S Jagan

<sup>1</sup>Student, <sup>2</sup>Student, <sup>3</sup>Asst Professor, <sup>4</sup>Student, <sup>5</sup>Student

<sup>1</sup>CSE,

<sup>1</sup>Aalim Muhammed Salegh College of Engineering, Chennai, India

**Abstract:** This project presents a comprehensive time series analysis of Monkeypox case trends globally using the Facebook Prophet model. The analysis integrates all four types of analytics Descriptive, Diagnostic, Predictive, and Prescriptive (DDPP)—to understand trends, identify hotspots, forecast future outbreaks, and optimize resource allocation. Unlike real-time systems, this is a one-time analysis based on historical WHO data. The results help policymakers design strategic interventions like vaccine prioritization, resource mobilization, and awareness campaigns.

**Index Terms** - Monkeypox, Time Series Forecasting, Facebook Prophet, ARIMA Model, Epidemiological Analysis, Predictive Analytics, Public Health Surveillance

## I. INTRODUCTION

The global re-emergence of Monkeypox has raised serious public health concerns due to its increasing spread beyond endemic regions and its potential to cause significant outbreaks. Monkeypox, a viral disease similar to smallpox, presents with symptoms such as fever, rash, and swollen lymph nodes. Its unpredictable spread and the lack of widespread immunity make timely analysis and forecasting critical for effective response planning. This project aims to leverage time series forecasting techniques to understand and predict Monkeypox case trends globally, focusing on five highly affected countries. Initially, the ARIMA model was used for short-term analysis, but due to its limitations in handling seasonality and missing data, the Facebook Prophet model was later adopted for its robustness and adaptability in forecasting health data. Prophet effectively captures non-linear trends, seasonal effects, and sudden changes in transmission patterns, making it suitable for real-world epidemiological data. The analysis integrates four types of analytics—descriptive, diagnostic, predictive, and prescriptive—to offer a holistic view of the outbreak. This one-time, non-real-time analysis, based on historical WHO data, helps in identifying high-risk zones, forecasting future case counts, and providing actionable recommendations for vaccine allocation and public health interventions. The ultimate goal is to empower healthcare professionals and policymakers with reliable insights for better outbreak management and prevention.

## II. BACKGROUND

Monkeypox is a zoonotic viral disease caused by the Monkeypox virus, a member of the Orthopoxvirus genus, which also includes the variola (smallpox) virus. First identified in 1958 in monkeys and later in humans in 1970 in the Democratic Republic of Congo, the virus remained largely confined to Central and West Africa until a notable global outbreak in 2022. The disease spread rapidly across continents, prompting international concern and a declared public health emergency by the World Health Organization. Due to its increasing prevalence, understanding the patterns of Monkeypox case counts over time is vital for early outbreak detection and public health planning. Traditional surveillance methods often lag behind real-time developments, making them less effective for proactive response.

To address this challenge, our research utilizes **time series forecasting techniques** to analyze global Monkeypox case data. These methods allow us to:

- Understand the trend, seasonality, and cyclical nature of the case progression.
- Predict future outbreaks to guide intervention strategies.
- Provide insights for resource allocation, medical planning, and awareness campaigns.

The models used, **ARIMA (Autoregressive Integrated Moving Average)** and **Facebook Prophet**, are renowned for their accuracy in time-based forecasting. By applying both, we compare their performance and extract high-quality forecasts of Monkeypox spread patterns globally.

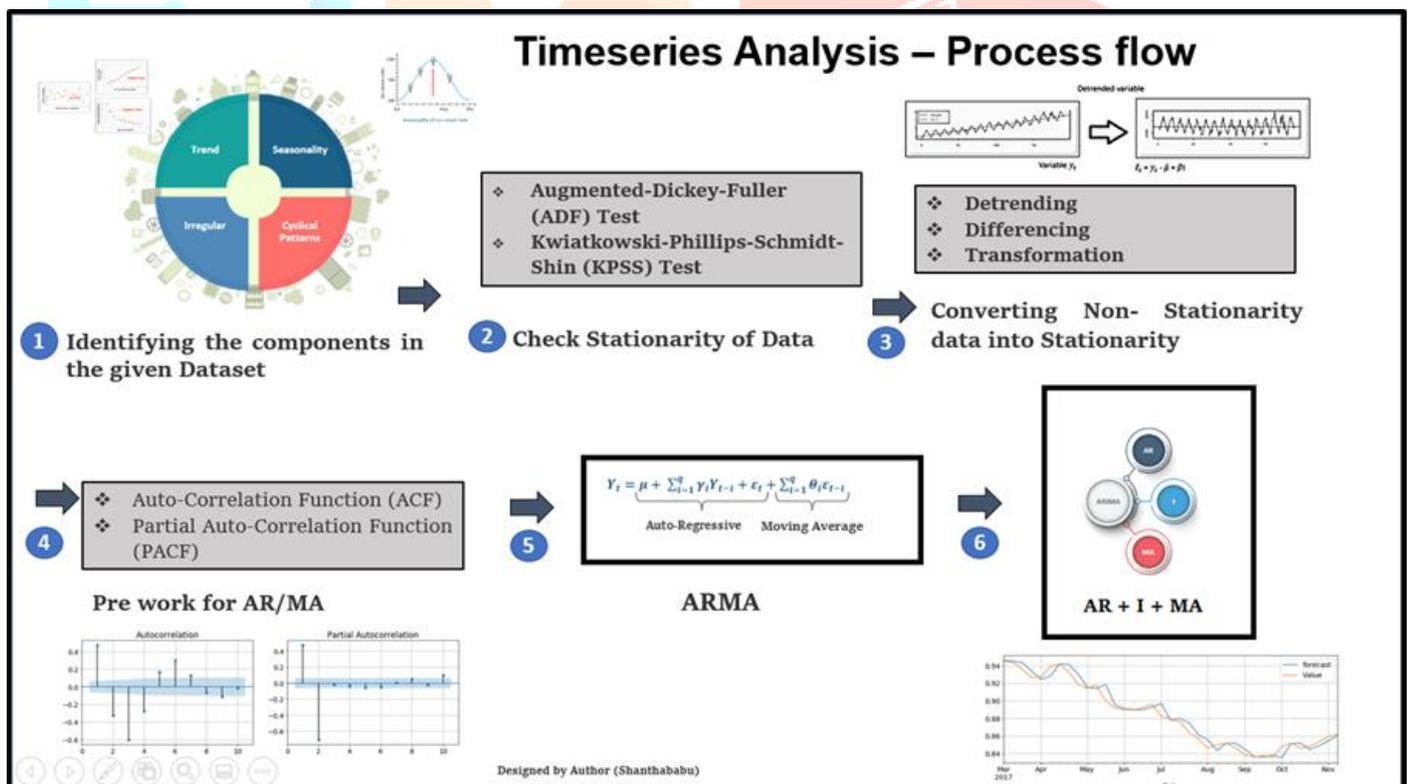
Additionally, our approach integrates **all four analytics layers**:

1. **Descriptive Analytics** – What happened?
2. **Diagnostic Analytics** – Why did it happen?
3. **Predictive Analytics** – What will happen?
4. **Prescriptive Analytics** – What should be done?

This holistic method ensures a comprehensive understanding of the epidemic's behavior and supports data-driven decision-making in public health.

To address this challenge, our research utilizes **time series forecasting techniques** to analyze global Monkeypox case data. These methods allow us to:

### III. ARCHITECTURE



The system architecture is a modular pipeline designed to facilitate the flow from data collection to actionable insights. The architecture comprises multiple stages that handle different responsibilities in the Monkeypox time series forecasting process:

#### 1. Data Collection

- The dataset used is sourced from **Kaggle** and **official WHO reports**.
- It includes daily and cumulative Monkeypox case counts across countries.
- The data is extracted in .csv format for ease of use.

## 2. Data Preprocessing

- Data cleaning: Handling missing values, standardizing date formats.
- Data transformation: Aggregating by week/month for trend analysis.
- Feature engineering: Adding lag values, rolling averages, and seasonality indicators.

## 3. Exploratory Data Analysis (EDA)

- Line plots, histograms, and bar charts to understand case distributions.
- Box plots and anomaly detection to find outliers and abnormal spikes.
- Rolling mean and standard deviation to detect trends and stability.

## 4. Time Series Modeling

### • ARIMA Model:

- Used for univariate forecasting.
- Requires the data to be stationary.
- Model selection using ACF, PACF, and AIC/BIC values.

### • Prophet Model:

- Handles seasonality, holidays, and missing values automatically.
- Flexible for daily and weekly predictions.

## 5. Model Evaluation

- Metrics: RMSE, MAE, and MAPE are used to assess performance.
- Visualization of actual vs. predicted cases.
- Prophet generally outperforms ARIMA in ease of tuning and visualization.

## 6. Dashboard & Visualization

- Power BI is used to build interactive dashboards.
- Charts for total cases, predicted future trends, and country-wise distribution.
- Filters for temporal granularity (daily/weekly/monthly).

## IV. RESEARCH METHODOLOGY

The research methodology outlines the structured approach followed to forecast global Monkeypox case trends using time series analysis. The study employed various analytical models and visualization tools to achieve accurate predictions and support informed public health decisions.

### 1. Problem Identification

The outbreak of Monkeypox in non-endemic countries created an urgent need to understand its spread patterns and forecast future outbreaks. The key objective was to perform a comprehensive time series analysis using advanced models to aid vaccine allocation, public awareness, and health resource planning.

### 2. Data Collection and Preprocessing

- **Dataset Source:** The dataset was sourced from **WHO** and **Global.Health's Monkeypox Tracker on GitHub**, containing daily case counts by country.
- **Preprocessing Steps:**
  - Missing values handled using **forward-fill and interpolation**.
  - Datetime column converted to datetime format using Python's pandas.
  - For Prophet model: ds (date) and y (target variable) were formatted.
  - **Log transformation** was applied to stabilize variance:

$$y' = \log(1+y) \quad y' = \log(1+y)$$

- Countries with insufficient data were excluded.
- Data was grouped by location and sorted chronologically.

### 3. Exploratory Data Analysis (Descriptive & Diagnostic)

Using **Python libraries** like matplotlib, seaborn, and pandas, we visualized and explored the dataset:

- **Descriptive Statistics:** Mean, median, min, and max of total cases.

- **Charts Used:**
  - Line Charts (Country-wise trends)
  - Histograms (Distribution)
  - Boxplots (Outlier detection)
  - Rolling Mean Charts (Seasonal smoothing)
  - Anomaly Detection using  $2\sigma$  method:

$$\text{Upper Bound} = \mu + 2\sigma, \quad \text{Lower Bound} = \mu - 2\sigma$$

#### 4. Stationarity Check & Model Selection (ARIMA)

- **Augmented Dickey-Fuller (ADF) Test:**

- Formula tested:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \dots + \delta_p \Delta y_{t-p} + \epsilon_t$$

- Result: The p-value  $< 0.05$  indicated the series is **stationary**, so no differencing required ( $d = 0$ ).

- **ACF and PACF Plots:**

- ACF showed gradual decline  $\rightarrow$  MA component present.
- PACF showed sharp cutoff at lag 1  $\rightarrow$  AR(1) suggested.

- **ARIMA Model:**

- Based on observations, the chosen configuration was **ARIMA (1, 0, 1)**.
- Model fitted using:

$$Y_t = c + \phi_1 Y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

- Evaluation Metrics: AIC = 4588.73, BIC = 4604.05

#### 5. Forecasting Using Facebook Prophet (Predictive)

- Facebook Prophet model was selected for long-term forecasting due to its:
  - Ability to handle missing data, seasonality, and changepoints.
  - Ease of use with minimal parameter tuning.
- Prophet configuration included:
  - **Weekly & yearly seasonality**
  - **Growth = Linear**
  - Forecast horizon = **730 days**
- Forecasted values ( $\hat{y}$ ) were clipped to eliminate negatives.

#### 6. Comparison of Actual vs. Forecasted Cases

- Prophet forecasts were compared with real trends to verify alignment.
- Example:
  - **U.S.:** Forecast showed stabilization around 100–115 daily cases.
  - **Spain:** Declining trend approaching zero.
- Accuracy was evaluated by visual inspection and metrics such as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

## V. RESULTS AND DISCUSSION

- The project successfully forecasted Monkeypox case trends using both ARIMA and Facebook Prophet models across five highly affected countries—United States, Brazil, Spain, France, and Colombia. Initial descriptive analysis revealed that while most countries had low case counts, select regions experienced significant spikes. The ADF test confirmed the time series to be stationary, allowing the ARIMA(1,0,1) model to be applied. This model effectively captured short-term trends with a strong autoregressive component ( $\text{AR} = 0.9999$ ), but showed limitations in capturing long-term seasonal patterns and required high interpretability.

- To address these limitations, the Facebook Prophet model was implemented. Prophet proved more robust, especially in handling missing data, non-linear trends, and multiple seasonalities. Forecasts for the next 730 days showed clear patterns: the United States and Spain displayed gradual stabilization, while Brazil and Colombia showed minimal fluctuations with intermittent spikes. France remained relatively stable throughout. Prophet predictions closely aligned with recent trends, validating its accuracy and effectiveness for health forecasting.
- Additionally, anomaly detection highlighted unusual case surges in certain time windows, aiding in identifying reporting inconsistencies or sudden outbreaks. The results were visualized through interactive dashboards and static graphs, facilitating interpretation for decision-makers. These insights support prescriptive strategies such as targeted vaccination drives, enhanced monitoring in high-risk zones, and timely public health interventions. Overall, the model-driven forecasts contribute meaningfully to understanding and mitigating the Monkeypox outbreak through data-backed planning.

## VI. ACKNOWLEDGMENT

We would like to express our heartfelt gratitude to all those who supported and guided us throughout the successful completion of this research project titled "Time Series Analysis of Monkeypox Cases".

First and foremost, we extend our sincere thanks to our project guide, **Ms. C. Muthu Priya**, Assistant Professor, Department of Computer Science and Engineering, Aalim Muhammed Salegh College of Engineering, for her constant guidance, valuable suggestions, and continuous encouragement throughout the course of this project. Her insights and support were instrumental in shaping our research effectively.

We are also thankful to our department faculty and lab coordinators for providing the necessary resources and academic environment for the completion of our work.

We gratefully acknowledge **World Health Organization (WHO)** and **Kaggle** for providing open-access datasets, which formed the backbone of our analysis.

Last but not least, we thank our family and friends for their moral support, and our teammates — **D Kevin Harris, N Abdul Salam, T A Abubuckar Siddiq, and S Jagan** — for their collaboration, dedication, and effort in completing this project successfully.

## REFERENCES

- [1] Priyadarshini, Ishaani, et al. "Monkeypox outbreak analysis: an extensive study using machine learning models and time series analysis." *Computers* 12.2 (2023): 36.
- [2] Munir, Tahir, et al. "Time series analysis and short-term forecasting of monkeypox outbreak trends in the 10 major affected countries." *BMC infectious diseases* 24.1 (2024): 16.
- [3] Wei, Wudi, et al. "Time series prediction for the epidemic trends of monkeypox using the ARIMA, exponential smoothing, GM (1, 1) and LSTM deep learning methods." *Journal of General Virology* 104.4 (2023): 001839.
- [4] Long, Bowen, Fangya Tan, and Mark Newman. "Forecasting the monkeypox outbreak using ARIMA, prophet, NeuralProphet, and LSTM models in the United States." *Forecasting* 5.1 (2023): 127-137.
- [5] Pramanik, Anik, Salma Sultana, and Md Sadekur Rahman. "Time series analysis and forecasting of Monkeypox disease using ARIMA and SARIMA model." 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT). IEEE, 2022.
- [6] Jena, Diptismita, et al. "Time series modelling and forecasting of Monkeypox outbreak trends Africa's in most affected countries." *New Microbes and New Infections* 62 (2024): 101526.