Covid 19 data analysis



## Department of Computer Science and Engineering

(B.E. Computer Science & Engineering Program Accredited by NBA from 2018-19 to 2021-25)

Report on Mini Project

Covid 19 data analysis

### Course Code : 21CSA31

### Course Name : R Programming

Semester: III SEM Section: D

**Submitted To:**

Mr. Shashank Shetty

Assistant Professor [GD-II]

Department of Computer Science and Engineering

**Submitted By:**

Srujan Acharya

Srivatsa R Upadhya

Sonith Shetty

### Date of submission: 21-01-2023

**Signature of Course Instructor**

# TABLE OF CONTENTS

**Title Page ................................................................................................................................... i**

**Abstract .................................................................................................................................... ii**

**Table of Contents ................................................................................................................... iii**

**Introduction ...................................................................................................................... 1**

**Problem Statement .................................................................................................................... 2**

**Objectives ................................................................................................................................... 3**

**Methodology .............................................................................................................................. 4**

**Implementation Details ............................................................................................................. 5**

**Results ........................................................................................................................................ 6**

**Conclusion and Future Scope................................................................................................. References ................................................................................................................................. 8**

Covid 19 data analysis

# ABSTRACT

Data and information visualization presumes that "visual representations and interaction techniques take advantage of the human eye’s broad bandwidth pathway into the mind to allow users to see, explore, and understand large amounts of information at once. Information visualization focused on the creation of approaches for conveying abstract information in intuitive ways.It will benefit any field of study that requires innovative ways of presenting large, complex information. The advent of computer graphics has shaped modern visualization.

Visualization is the use of computer-supported, visual representation of data. Unlike static data visualization, interactive data visualization allows users to specify the format used in displaying data. Common visualization techniques are:

* Line graph: This shows the relationship between items. It can be used to compare changes over a period of time.
* Bar chart: This is used to compare quantities of different categories. • Scatter plot: This is a two-dimensional plot showing variation of two items.
* Pie chart: This is used to compare the parts of a whole.

Covid 19 data analysis

# INTRODUCTION

In brief, a Nodejs server is used to create plots by R and display it over a website. Requirements: you should have before setting up a server install Nodejs.

Begin by creating a javascript file, install libraries express and rintegration.

Set up API which takes three parameters start-date, end-date and country for the scatter plot. The r-integration function is created to run the R-script in the server, which takes in the parameter of passed through the APL

Here Rscript in the server helps to create/overwrite the images of the plot. Plots will be sorted by India and from the beginning date of the data set to the ending date by default.

On successful plot the payload is sent back to the web page with the message "success"\"fail" with image URL. The image is then populated in the webpage for visualization.

Covid 19 data analysis

# PROBLEM STATEMENT

Statistical analysis and visualization of date-wise covid-19 data using R and present in easy to read format.

**OBJECTIVES**

* Get data from a reliable source
* Sort/filter the data visualize and analyze using scatter plots in date vs new cases and date vs Deaths.
* Present the data in a website.

Covid 19 data analysis

# METHODOLOGY

# As the data is large it is more convenient to represent and visualize the data in graphical/ pitorial forms rather than having data in large spreadsheets.

# The raw data is gathered as a csv file from WHO.

# Raw data is imported to an R script and data is sorted/filtered and scatterplot is generated. The scatter plot is displayed on a website for easy access at anytime.

Covid 19 data analysis

# IMPLEMENTATION

In brief, a Nodejs server is used to create plots by R and display it over a website. Requirements: you should have before setting up a server install Nodejs.

Begin by creating a javascript file, install libraries express and r-integration.

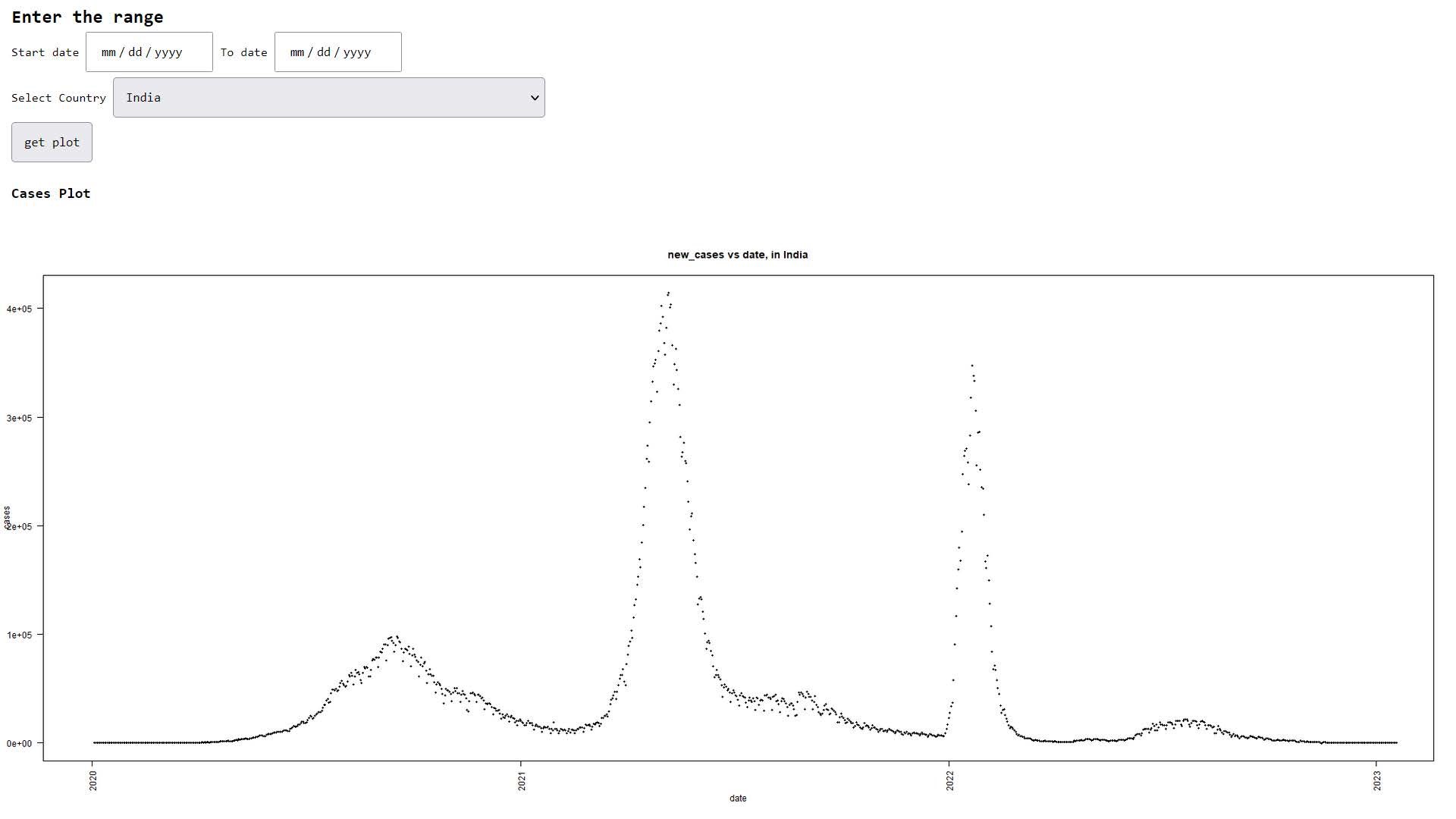
Set up API which takes three parameters start-date, end-date and country for the scatter plot. The r-integration function is created to run the R-script in the server, which takes in the parameter of passed through the APL

Here Rscript in the server helps to create/overwrite the images of the plot. Plots will be sorted by India and from the beginning date of the data set to the ending date by default.

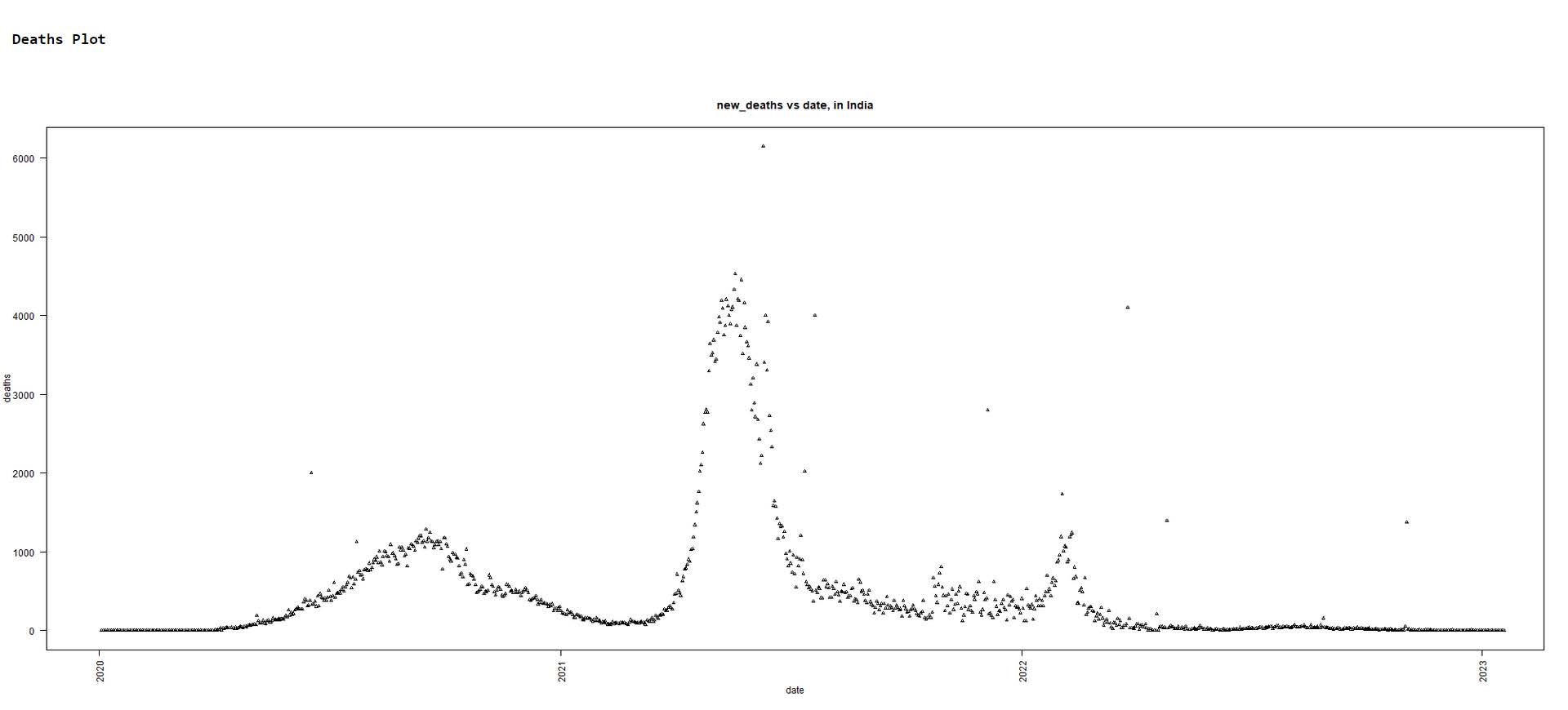
On successful plot the payload is sent back to the web page with the message "success"\"fail" with image URL. The image is then populated in the webpage for visualization.

# RESULTS AND DISCUSSIONS

Covid 19 data analysis

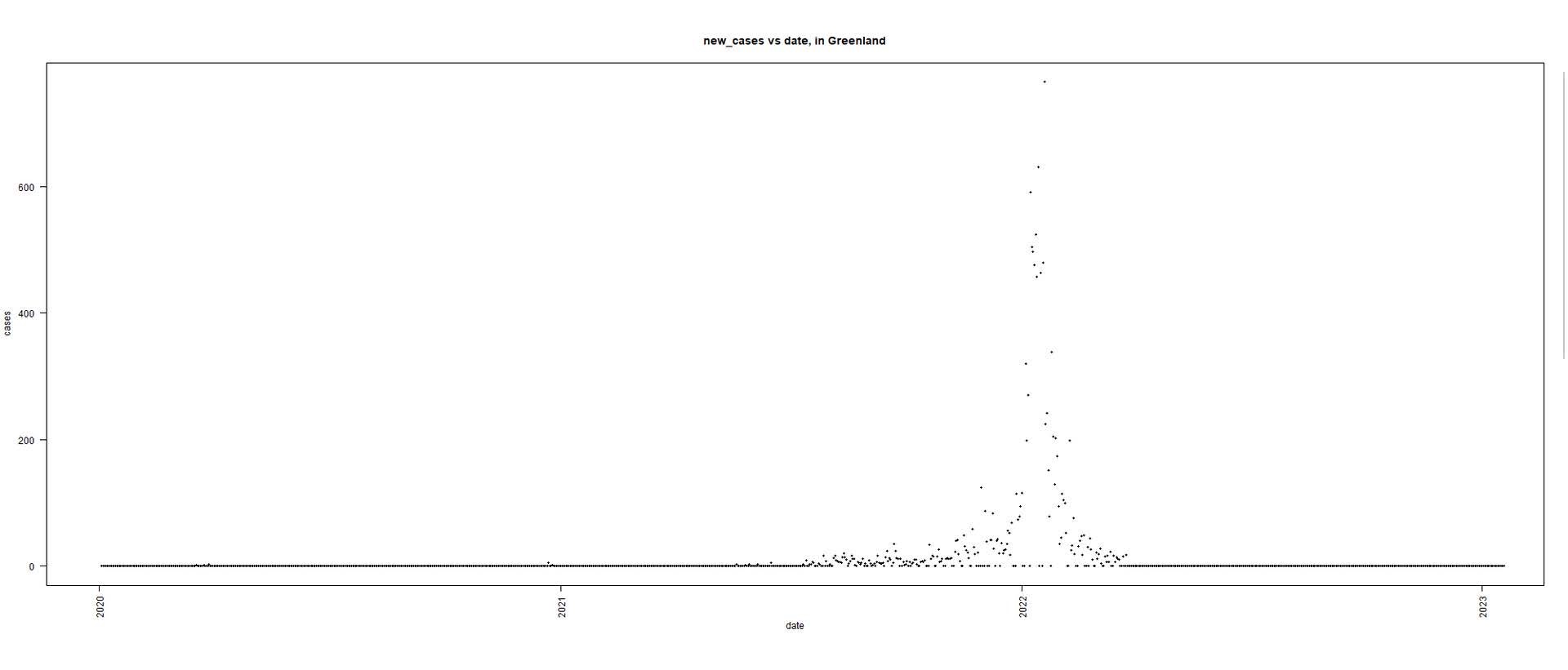


The above graph represents the data of number of covid cases against date. We can observe the rise and fall in above graph.



The above graph represents the data of number of death cases against date. We can observe the rise and fall in above graph.

Covid 19 data analysis



# CONCLUSION AND FUTURE SCOPE

Covid 19 data analysis

Data visualization is used for processing large set of data in a graphical or pictorial way in an organized fashion. It has emerged as a powerful and widely applicable tool for analyzing and interpreting

complex and large data. It is able to communicate the clarity, accuracy and efficiency of the data. This can be helpful in may fields like medical, research.

Implementation of Regression model for the plotting and prediction. Making it available for unique data sets. Web scrapping to get real time data for more reliable visualization and regression model

Covid 19 data analysis

# REFERENCES

# [1] Link for covid data –

# https://covid19.who.int/WHO-COVID-19-global-table-data

# [2] R-plot - A blog on data visualization, <https://www.digitalocean.com/community/tutorials/plot-function-in-r>

# [3] r-integration - Reference to library for R integration