

COMP47250: Team Software Project

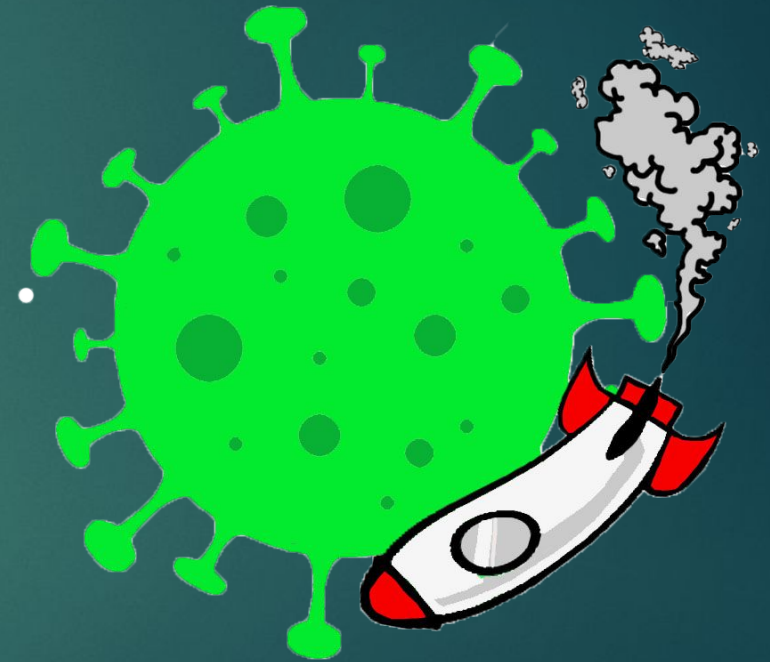
PRESENTATION 3

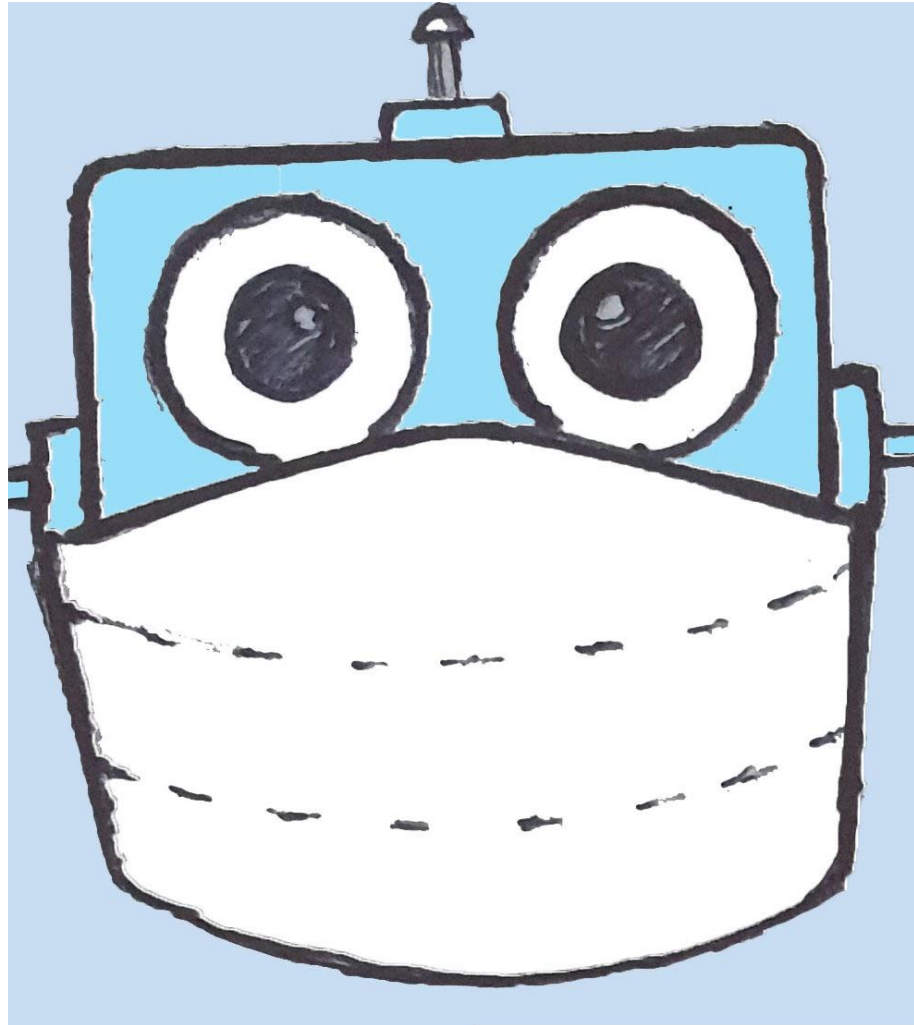
22ND JUNE, 2020



Team Rocket

▶ Aditya Krishnan	15316048
▶ Apurva Deore	19200283
▶ Ashwin Manikandan	19201000
▶ Minaz Shaikh	19200277
▶ Nripendra Singh	19200326
▶ Sumit Kumar	19200701
▶ Vishal Padma	19200174





Project objectives

Challenge 1: COVID-19

COVID-19 INTERACTIVE VISUALIZER

AIM

Creation of a **Web-based Chatbot Application and Map Visualizer** which would enable us to solve user queries regarding the COVID-19 pandemic across the world. We aim to create more awareness among people regarding the effects of the pandemic specifically in the health sector, i.e. Infected rates, Death rates, Recovery rates, etc, by letting users directly interact with the chatbot. The web page will use trusted sources to gather relevant information in the aforementioned categories which will be used to drive the responses of the bot and power the map based visualizations of the world which will be also deployed on the same application.

We plan to achieve two paramount features

CHATBOT

A chatbot which would answer user's queries about COVID-19.

Eg: Symptoms of COVID-19?

INTERACTIVE MAP

An interactive map-based feature deploying visualizations that provide statistics about COVID-19 of various countries.

E.g. Total cases? Active cases?

Project Plan

- ▶ Version 1.0 (8th June – 15th June'20)

First view of the rough front-end page of the web-page along with main statistical summarized data collected.

- ▶ Version 1.1 (16th June – 22nd June'20)

Integration of Django to the front-end and a basic implementation of the Chat-Bot. With finer implantation of previous work done.

- ▶ Version 1.2 (24th June – 1st July'20) □ Milestone 1/MVP

Core implementation of the paramount tasks (House the Chat-Bot and prepare the map visualizations) before the interim presentation and subsequent preparation of the report.

- ▶ After having achieved V1.2 and the delivery of MVP , we aim to further breakdown tasks and reach goals into sprints and release versions. As part of the main non-trivial extension we would like to integrate the two paramount features of the application in a manner where the Visualizer responds dynamically to queries made in the chatbot.

Roles

▶ **Aditya and Minaz (Front-End Team)**

Research on tools for development, color palettes, color themes, research and design logo and bot icons.
Decide on the UI design and layout and functionality of the Webpage with the subsequent creation of Front-End.
Integration of the Front-End with the Back-End.

▶ **Apurva and Vishal (Data and Visualization Team)**

Research on potential APIs required and write scripts for Data Collection.
Data Mining and Data Cleaning/Pre-processing as and when required.
NLP Processing for Text and research/deploy visualizations.

▶ **Ashwin (Cloud Deployment Team)**

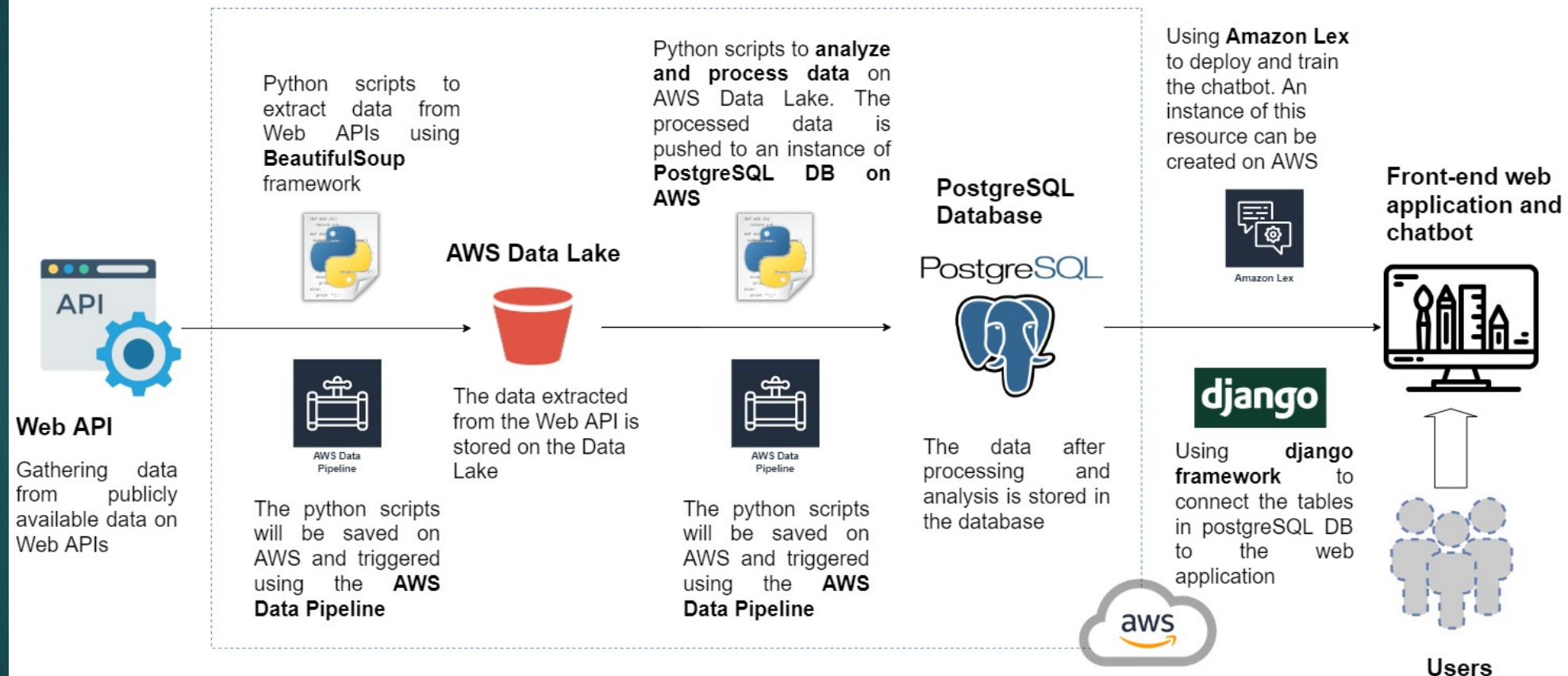
Test/Deploy the required resources (Data Lake/ PostgreSQL/Data Factory/DB).
Create dummy scripts and trigger them using data factory pipelines.
Research on how to connect PostgreSQL to Front-end.

▶ **Sumit and Nripendra (Django Team)**

Understanding Django and identify issues for implementation.
Analyze for the risks involved and prepare .Net Backup
Creation of Back-End and understanding Dialogue Flow for further integration.

Architecture

Team Rocket - Covid-19 Project Architecture



► Data to be considered

We are considering data from various APIs; We scoured the web for COVID-19 API's or other implementations of centralized world data regarding the pandemic. We plan to consider two broad types : DYNAMIC and STATIC data. Once the data is collected and curated, we aim to filter the type of queries the bot is able to answer.

DYNAMIC / LIVE DATA

► -Statistics of COVID-19 cases Per Country [Real Time Data, Updating regularly]

- Total cases
- Active cases
- Recovered cases
- Deceased cases
- Tests done

❖ <https://api.covid19api.com/summary>

❖ <https://api.covid19api.com/countries>

STATIC DATA






► -Data already loaded in and doesn't need frequent updates.

- Symptoms of COVID-19?
- Healthy habits to be followed?
- Whom to contact for testing related situations (some helpline numbers)

DATA PLAN

GitHub

- ▶ Work done until now; like ideas proposed, minutes of meetings, any research done so far related to the project, has been pushed onto our team's repository.
- ▶ We've created different folders and are constantly push-committing the work as and when it is completed.
- ▶ Just to give a gist...

 Code	Update Amazon Lex.docx	8 hours ago
 EvaluationTest	Test Script	yesterday
 Minutes	Add files via upload	22 hours ago
 Presentation	Folder Updation	3 days ago
 Project Objective_Draft.docx	Add files via upload	6 days ago

Team Management

- ❑ Team progress managed by Sumit and Aditya:
- ❑ Shared google drive folder which contains all the different folders for each team consisting of basically all of the data cumulated at one place for enhanced availability and accessibility, including the log of each MOM and documenting the progress of the team at each stage.
- ❑ Conducting daily meetings via **Google Meet** and logging MOM of each meeting
- ❑ We're using **Jira** to outline each team member's work, define sprints, plan the work to be done and the corresponding deadlines for them.
- ❑ Just to give you a gist, you can see on the right; our Jira board with tasks allocated and our shared folder on Google Drive.

The screenshot shows a Jira sprint board for 'WBCDDRCP Sprint 2' with 14 issues. The board is organized into columns for different teams: Project Management, Django Team, Data Team, Cloud and Deployment, and FrontEnd Team. Each issue is represented by a card with a title, assignee, and status. The issues are as follows:

Issue Title	Assignee	Team	Issue ID
Keep track of members and have daily morning meetings	Project Management T...	Project Management	WBCDDRCP-58
Create sprint plans for subsequent weeks	Project Management T...	Project Management	WBCDDRCP-56
Backend framework setup, development and integration	V1.1 Django Team	Django Team	WBCDDRCP-50
Building backend framework	V1.1 Django Team	Django Team	WBCDDRCP-53
preparation of global summary data of all countries until the date of visit to the website	V1.1 Data Team	Data Team	WBCDDRCP-24
Decision on the types of visualizations to do(Once the data has been finely prepared and curated)	V1.1 Data Team	Data Team	WBCDDRCP-33
Provide snapshot of the final raw data and processed data to the cloud team for collaboration	Data Team	Data Team	WBCDDRCP-24
create dummy scripts and trigger them using data pipelines to asses data retrieval effectiveness	Cloud and Deployment...	Cloud and Deployment	WBCDDRCP-62
Test deploy the required resources and learn its functionalities	Cloud and Deployment...	Cloud and Deployment	WBCDDRCP-61
Prepare information section data	V1.1 FrontEnd Team	FrontEnd Team	WBCDDRCP-40
Prepare the main colour themes and the colour palette for visualizations	FrontEnd Team	FrontEnd Team	WBCDDRCP-35
Clean web page design and layout	V1.1 FrontEnd Team	FrontEnd Team	WBCDDRCP-65
Weekly presentation to log progress of the team	Project Management T...	Project Management	WBCDDRCP-30
preparation of 5 min video explaining the rough overview of the progress	Project Management T...	Project Management	WBCDDRCP-27

The screenshot shows a Google Drive folder named 'NL Software Project' with 7 people. The folder contains a presentation titled 'COMP47250: Team Software Project' and a video script titled 'video script'. The video script was edited by Apurva Arun Deore today. Below the files, there is a section for 'Folders' with the following folders listed:

- Data
- management and documentation
- Minutes doc
- PPTs
- Python dev
- Rough roles and tasks

Progress So Far..

Front-End

- ✓ Preparation of Information Section Data
- ✓ Selecting/Preparing the main colour themes and the colour palette for visualizations.
- ✓ Clean the front-end of the web application design and structuring the layout of the pages accordingly.

Data

- ✓ Researched/Finalized the potential API's to use (main focus on the John Hopkins API)
- ✓ Scripts written for the collection of raw data from the desired sources (CSV format).
- ✓ Snapshot of the raw and processed data provided to the Cloud team for collaboration.
- ✓ Decided on the two main visualizations and deployed them (used Tableau for the same)

Test Evaluation Scripts

- ✓ Creation of a User Test Case script template and a peer review template.

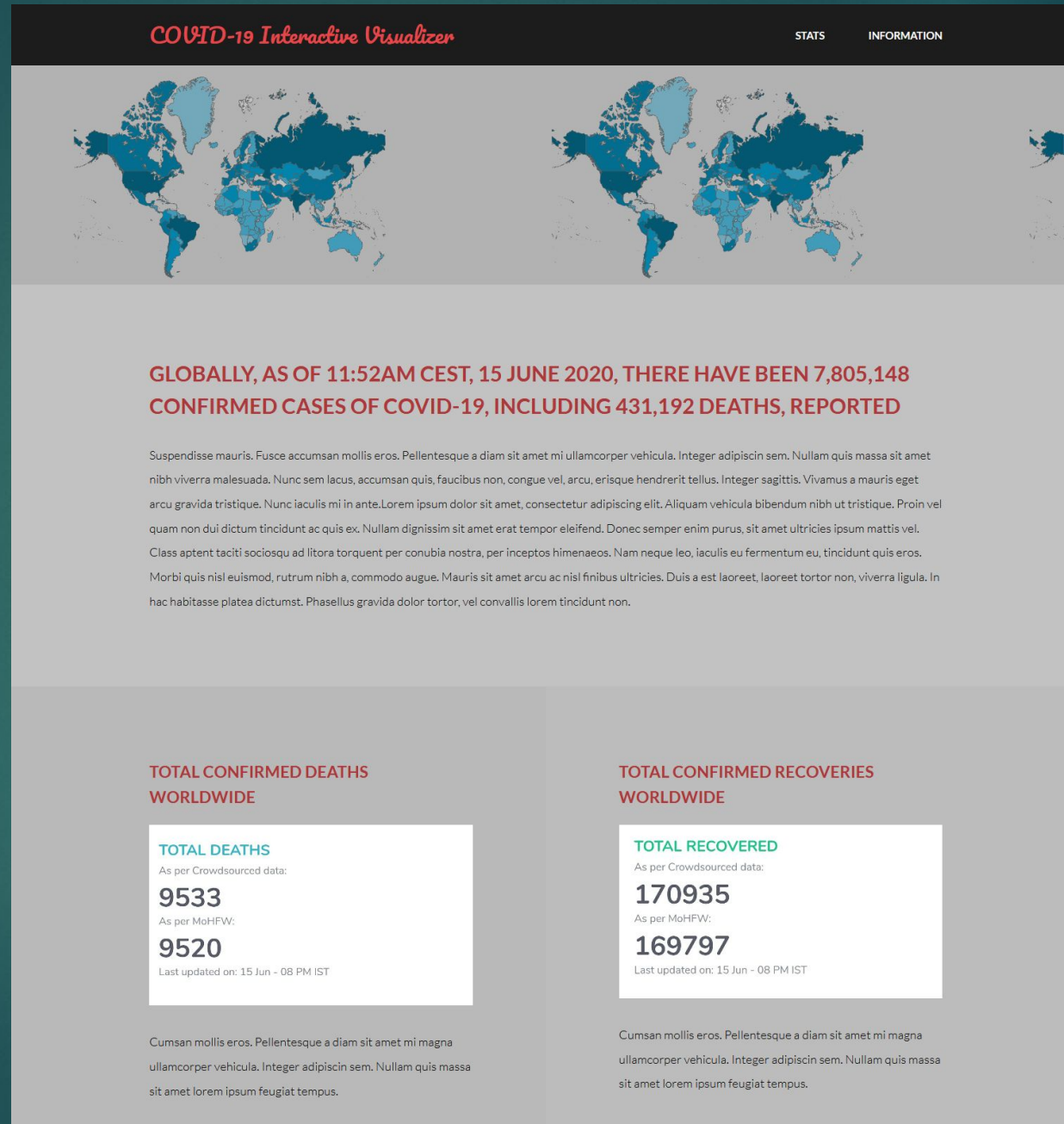
Cloud Architecture

- ✓ Test deploy the required resources (Data lake, PostgreSQL DB, Data Factory) and learn their functionalities.
- ✓ Creating dummy scripts and triggering them using data factory pipelines.
- ✓ Assess their effectiveness in data retrieval and storage.

Django Framework

- ✓ Analysis of Django completed.
- ✓ Framework has been setup and home page has been integrated with Django.
- ✓ Analyze risks involved if any.

Just a gist of what our
web-page would look like!



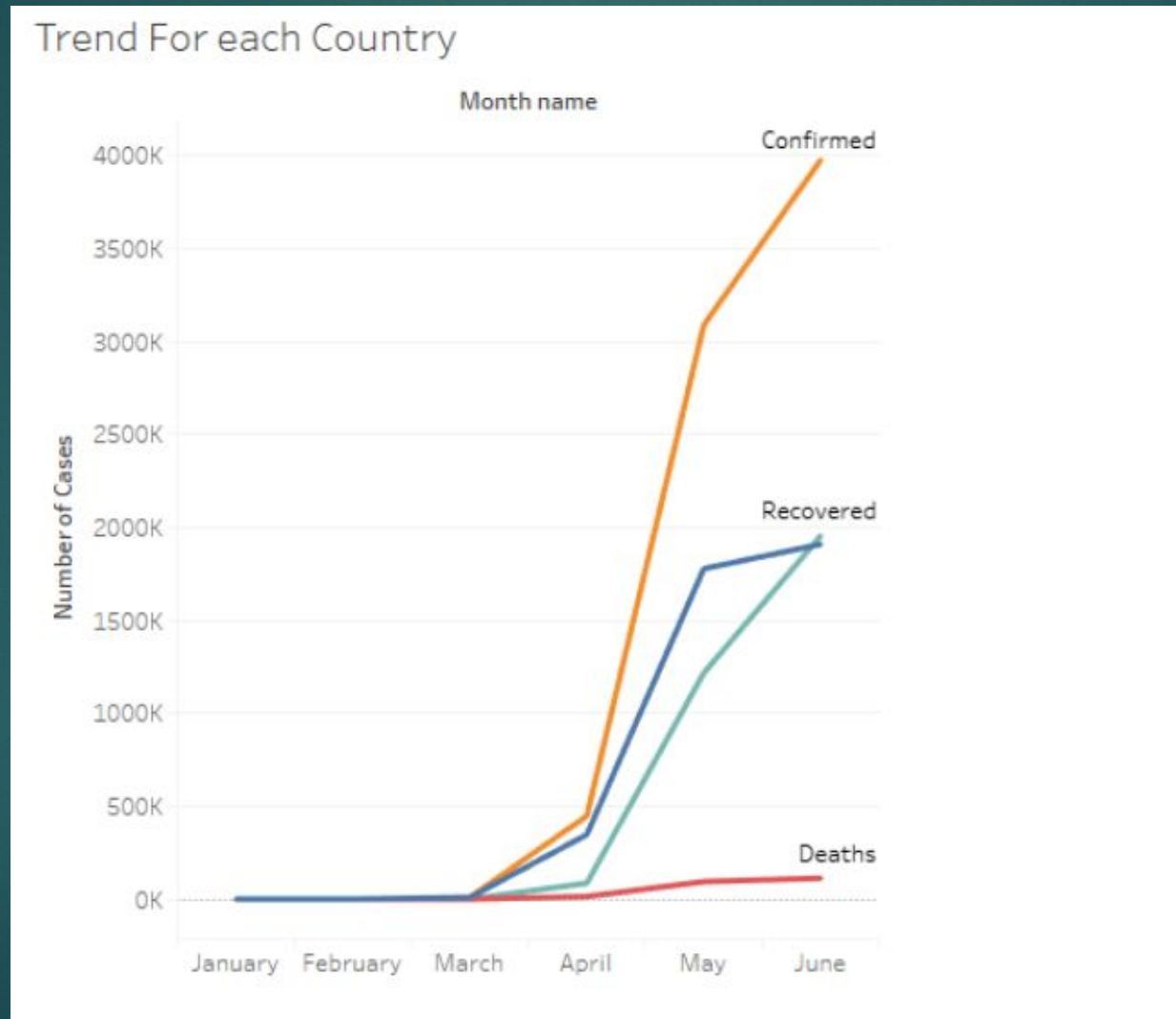
Data & Visualizations

```
Active, City, CityCode, Confirmed, Country, CountryCode, Date, Deaths, Lat, Lon, Province, Recovered
0,,,0,Afghanistan,,2020-01-22,0,0,0,,0
0,,,0,Afghanistan,,2020-01-23,0,0,0,,0
0,,,0,Afghanistan,,2020-01-24,0,0,0,,0
0,,,0,Afghanistan,,2020-01-25,0,0,0,,0
0,,,0,Afghanistan,,2020-01-26,0,0,0,,0
0,,,0,Afghanistan,,2020-01-27,0,0,0,,0
0,,,0,Afghanistan,,2020-01-28,0,0,0,,0
0,,,0,Afghanistan,,2020-01-29,0,0,0,,0
0,,,0,Afghanistan,,2020-01-30,0,0,0,,0
0,,,0,Afghanistan,,2020-01-31,0,0,0,,0
```

This is the TIME_SERIES data which would be mainly used for visualizations to show trends that each country has observed so far for the counts of confirmed/Recovered/Active/ Deceased cases.

16169,,,18054,Afghanistan,,2020-06-04,300,0,0,,1585
16898,,,18969,Afghanistan,,2020-06-05,309,0,0,,1762
17394,,,19551,Afghanistan,,2020-06-06,327,0,0,,1830
18110,,,20342,Afghanistan,,2020-06-07,357,0,0,,1875
18377,,,20917,Afghanistan,,2020-06-08,369,0,0,,2171
18424,,,21459,Afghanistan,,2020-06-09,384,0,0,,2651
18724,,,22142,Afghanistan,,2020-06-10,405,0,0,,3013
19138,,,22890,Afghanistan,,2020-06-11,426,0,0,,3326
0,,,0,Albania,,2020-01-22,0,0,0,,0
0,,,0,Albania,,2020-01-23,0,0,0,,0
0,,,0,Albania,,2020-01-24,0,0,0,,0
0,,,0,Albania,,2020-01-25,0,0,0,,0
0,,,0,Albania,,2020-01-26,0,0,0,,0
0,,,0,Albania,,2020-01-27,0,0,0,,0

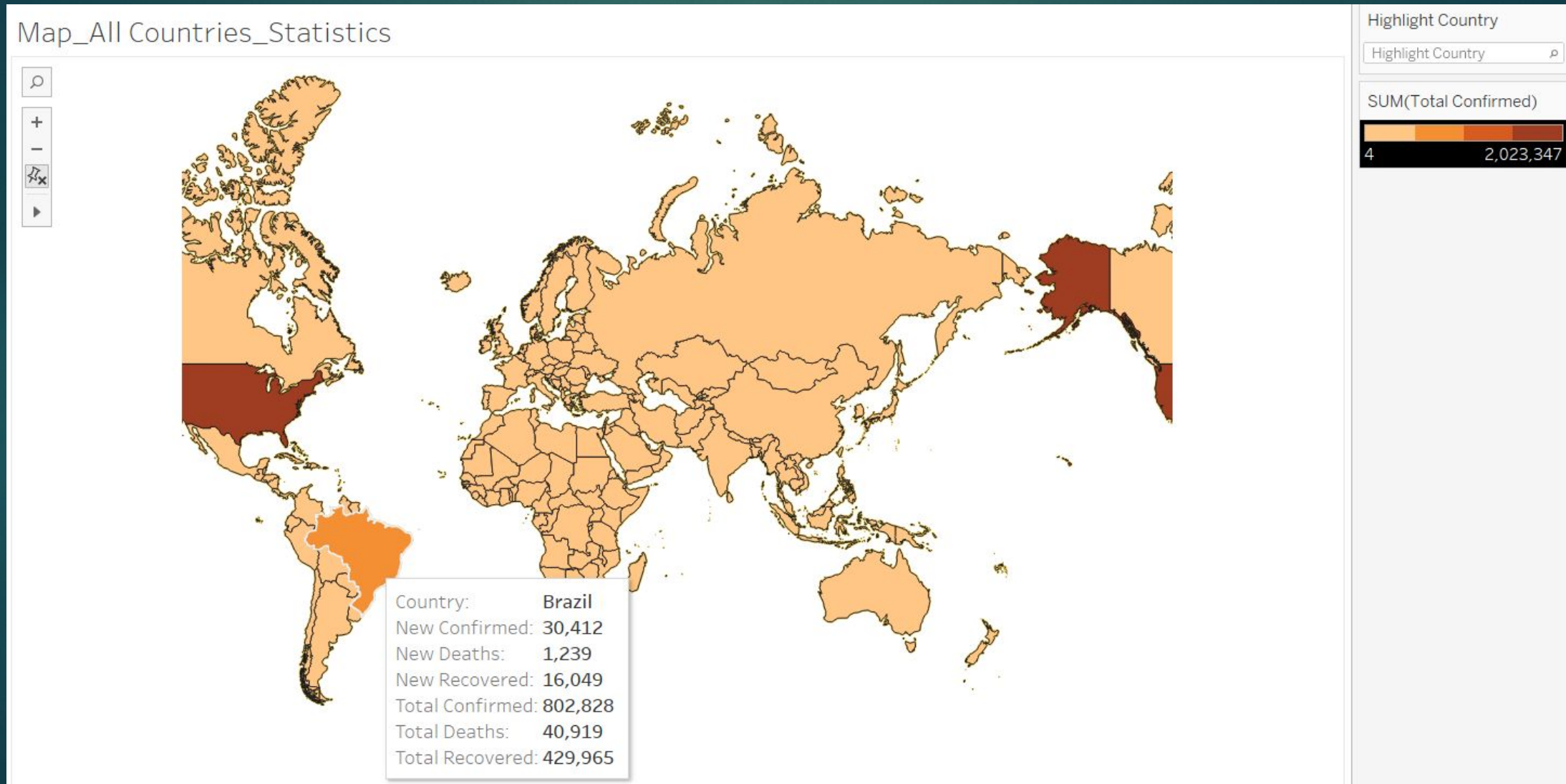
The previous slide's Data would be used for Trend Visualizations as much mentioned before.



This is the summarized data which provides us with 6 counts: New Confirmed, New Deaths, New Recovered, Total Confirmed, Total Deaths, Total Recovered for all countries on the date of visit.

	Country	CountryCode	Date	NewConfirmed	NewDeaths	NewRecovered	TotalConfirmed	TotalDeaths	TotalRecovered
0	Afghanistan	AF	6/19/2020	658	42	1502	27532	546	7660
1	Albania	AL	6/19/2020	66	1	9	1788	39	1086
2	Algeria	DZ	6/19/2020	117	12	135	11385	811	8078
3	Andorra	AD	6/19/2020	1	0	1	855	52	792
4	Angola	AO	6/19/2020	11	1	0	166	8	64
5	Antigua and Barbuda	AG	6/19/2020	0	0	0	26	3	22
6	Argentina	AR	6/19/2020	1958	35	209	37510	948	10721
7	Armenia	AM	6/19/2020	665	7	746	18698	309	7560
8	Australia	AU	6/19/2020	18	0	1	7409	102	6878
9	Austria	AT	6/19/2020	20	1	2	17223	688	16101
10	Azerbaijan	AZ	6/19/2020	338	6	117	11329	139	6192
11	Bahamas	BS	6/19/2020	0	0	2	104	11	74
12	Bahrain	BH	6/19/2020	469	6	511	20430	55	14696
13	Bangladesh	BD	6/19/2020	3803	38	1975	102292	1343	40164
14	Barbados	BB	6/19/2020	0	0	0	97	7	85
15	Belarus	BY	6/19/2020	625	7	1288	56657	331	34023
16	Belgium	BE	6/19/2020	104	8	40	60348	9683	16724
17	Belize	BZ	6/19/2020	0	0	0	22	2	16
18	Benin	BJ	6/19/2020	25	2	1	597	11	238
19	Bhutan	BT	6/19/2020	0	0	1	67	0	25
20	Bolivia	BO	6/19/2020	814	18	0	21499	697	4002
21	Bosnia and Herzegovina	BA	6/19/2020	33	0	22	3174	168	2219
22	Botswana	BW	6/19/2020	0	0	0	79	1	25
23	Brazil	BR	6/19/2020	22765	1238	13534	978142	47748	534580
24	Brunei Darussalam	BN	6/19/2020	0	0	0	141	3	138
25	Bulgaria	BG	6/19/2020	132	6	61	3674	190	1941
26	Burkina Faso	BF	6/19/2020	0	0	1	899	53	810
27	Burundi	BI	6/19/2020	0	0	0	104	1	75

When we hover over a country, we'd be provided with the statistics for that country. Also, the intensity of the colour used increases as the count of Total Confirmed cases increase. We can see, USA has been assigned the darkest shade possible in the range. You can see the colour shade used on the top-right most side of the picture.



STAY SAFE!

THANK YOU!

-Team Rocket