

# Collaborative Data Science and COVID-19

Assoc Prof Peter Julian Cayton, PhD

School of Statistics, University of the Philippines Diliman

Member, UP COVID-19 Pandemic Response Team

Member, LEADS for Health Security and Resilience



# Flow of the Presentation

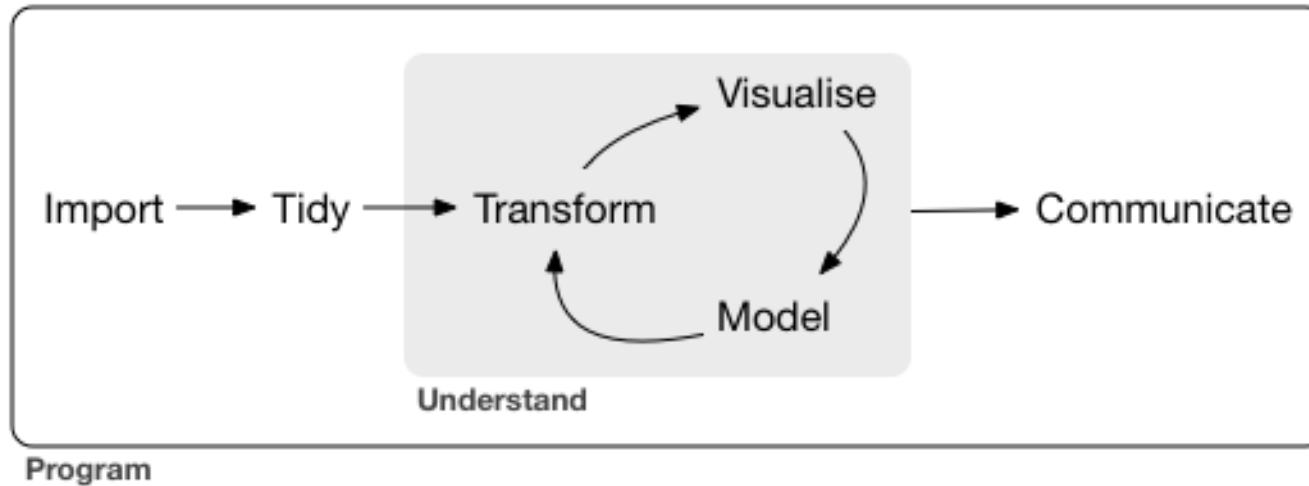
- ▶ A Typical Data Science/Analytics Workflow
- ▶ Membership to Data-Driven Teams
- ▶ Example Projects
- ▶ Demonstrating the Workflow with a Team using Example Projects
- ▶ Closing Remarks



# A Typical Data Science/Analytics Workflow

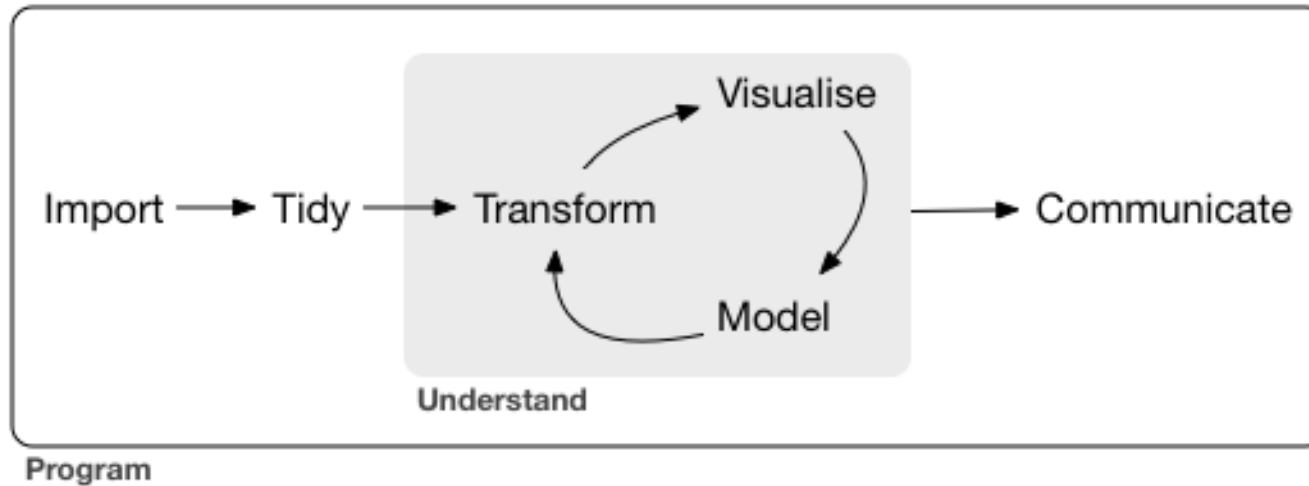


# A Typical Data Science/Analytics Workflow



- Source: Wickham, H and Grolemund, G (2017). R for Data Science. O'Reilly.  
<https://r4ds.had.co.nz/introduction.html>

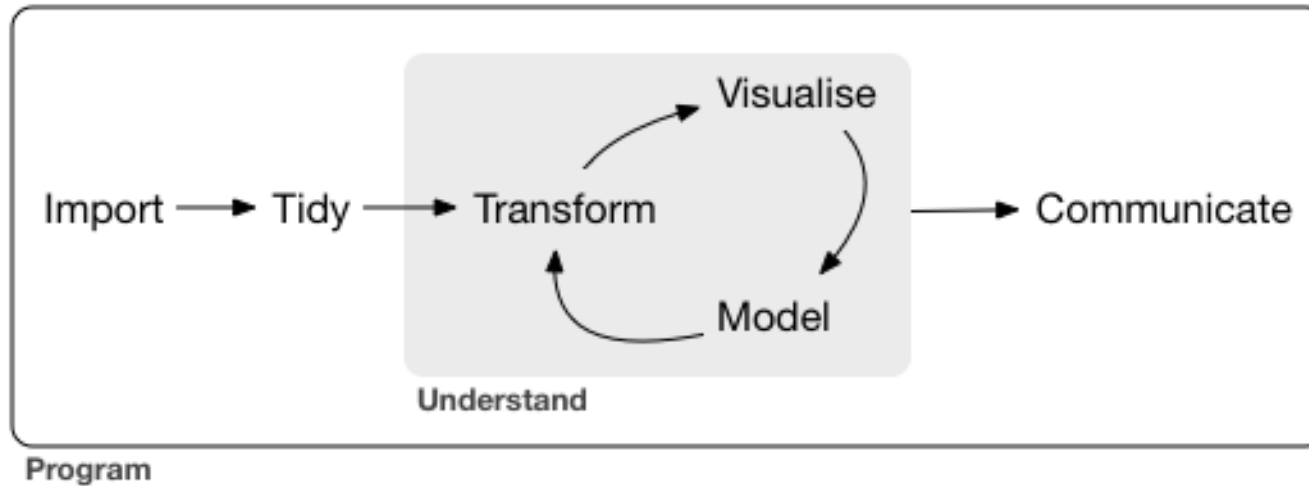
# A Typical Data Science/Analytics Workflow



## ► Import

- Extracting the data from an internal database, a file, an online website, or thru a web application programming interface (API), to be loaded in R/RStudio

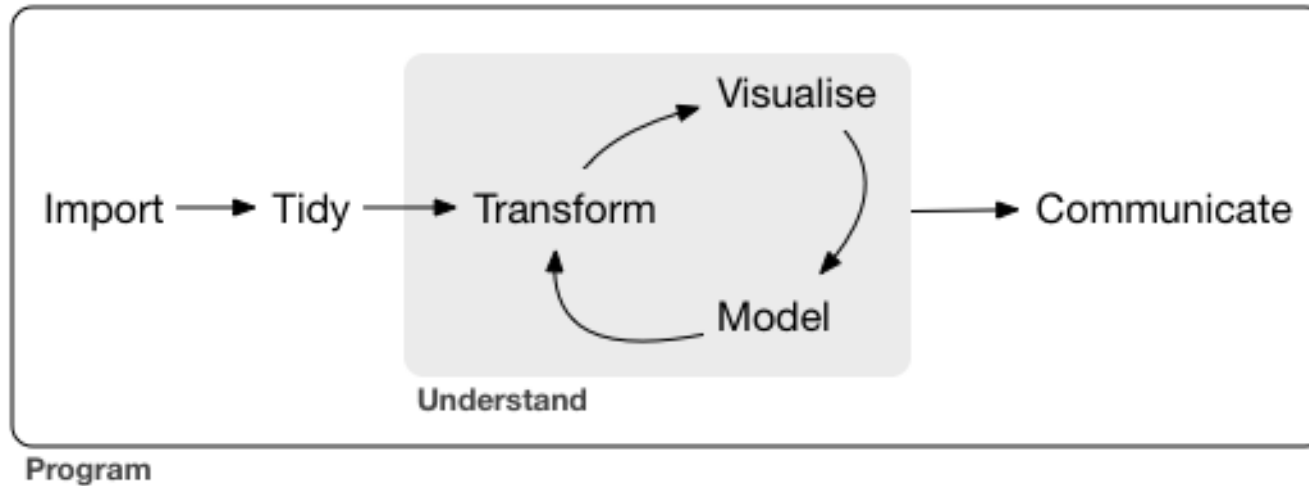
# A Typical Data Science/Analytics Workflow



## ► Tidy

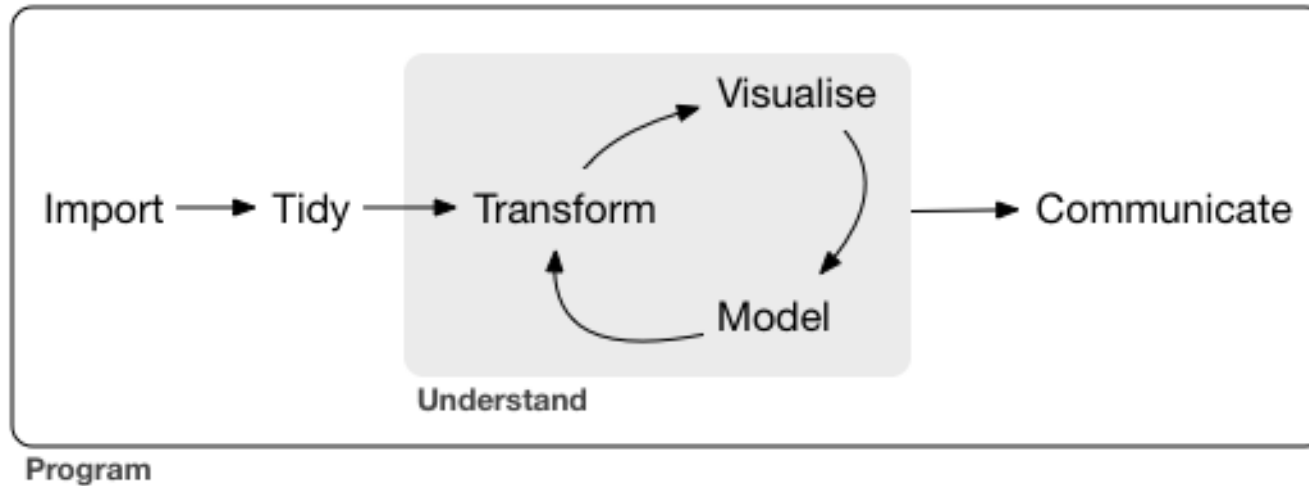
- Arranging the data into a neat data structure, with variable as columns and data points as rows.
- Included in this step would be data cleaning, data augmentation, missing data imputation, and others

# A Typical Data Science/Analytics Workflow



- Understand
  - Generally, the steps to extract insights from data after tidying up.

# A Typical Data Science/Analytics Workflow

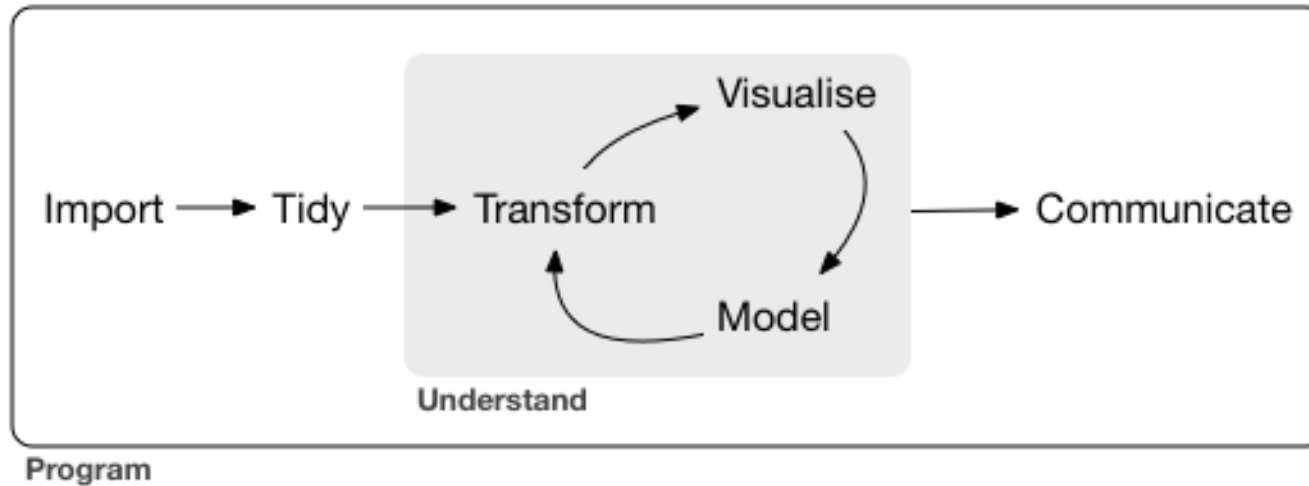


## ► Transform

- Processing the data in preparation for further steps. Examples are:
  - Narrowing the data, e.g., by region or by age,
  - Computing new variables, e.g., length of days until recovery, or delays in reporting cases
  - Aggregating data, e.g., counting cases or solving rates/means

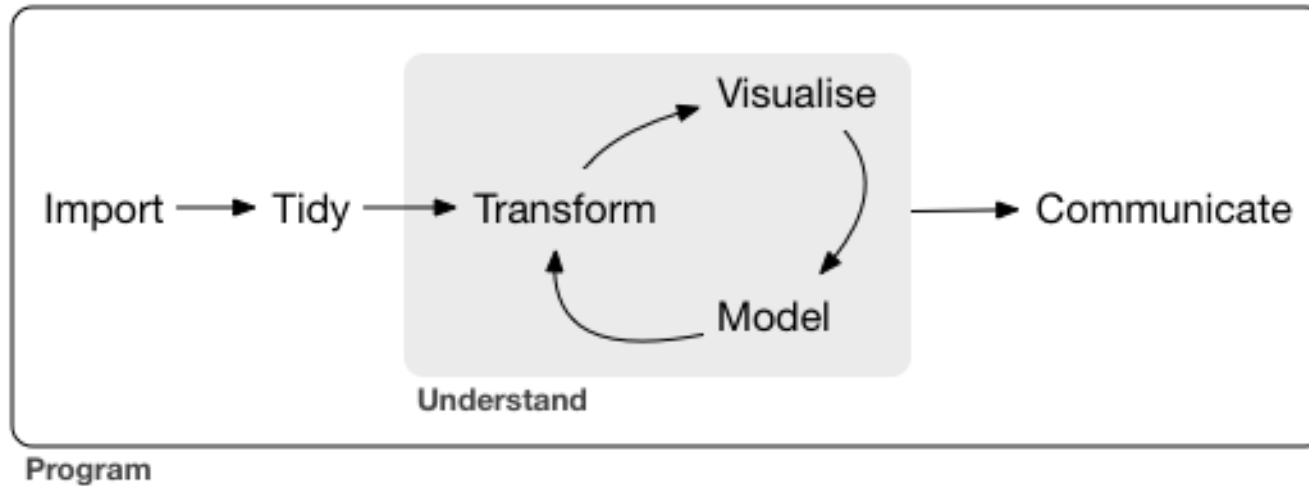


# A Typical Data Science/Analytics Workflow



- Visualize
  - Plot data into graphs so that patterns and features may be explored and insights be extracted from what is seen.

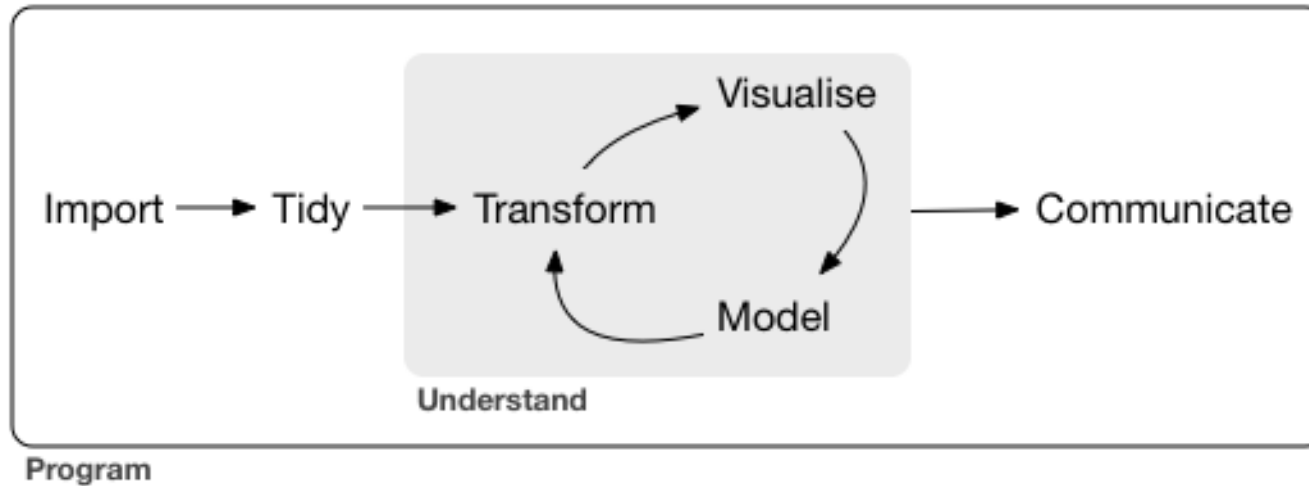
# A Typical Data Science/Analytics Workflow



## ► Model

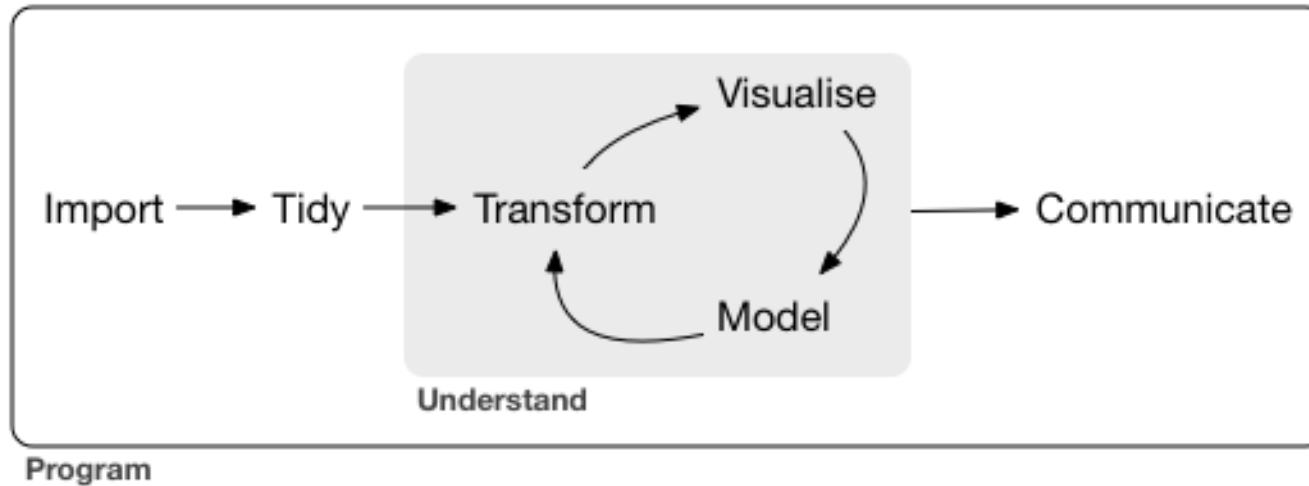
- When necessary, models help in summarizing the complex relationships and the patterns found from visualizations
- Designing models for prediction or forecasting

# A Typical Data Science/Analytics Workflow



- Communicate
  - These include:
    - Writing reports, creating dashboards, making presentations, compilations, etc.

# A Typical Data Science/Analytics Workflow



## ► Program

- All these processes to be encapsulated in a data science project plan
- Possible to be encapsulated in one software, but it's not impossible to use more than one depending on team members' capabilities to process and analyze data.

# Membership to Data-Driven Teams





## About the UP Pandemic Response Team

# About UP COVID-19 Pandemic Response Team

- Leaders:
  - Dr Teodoro Herbosa, UP System
  - Dr Alfredo Mahar Lagmay, Executive Director, UP Resilience Institute
- Who we are
  - 200+ experts and volunteers from the whole UP System, from Baguio to Davao
  - spanning multiple fields: political scientists, statisticians, mathematicians, geographers, geologists, medical doctors, linguists, economists, etc.
- Next slides from Dr Lagmay: Our Portfolio





University of the Philippines

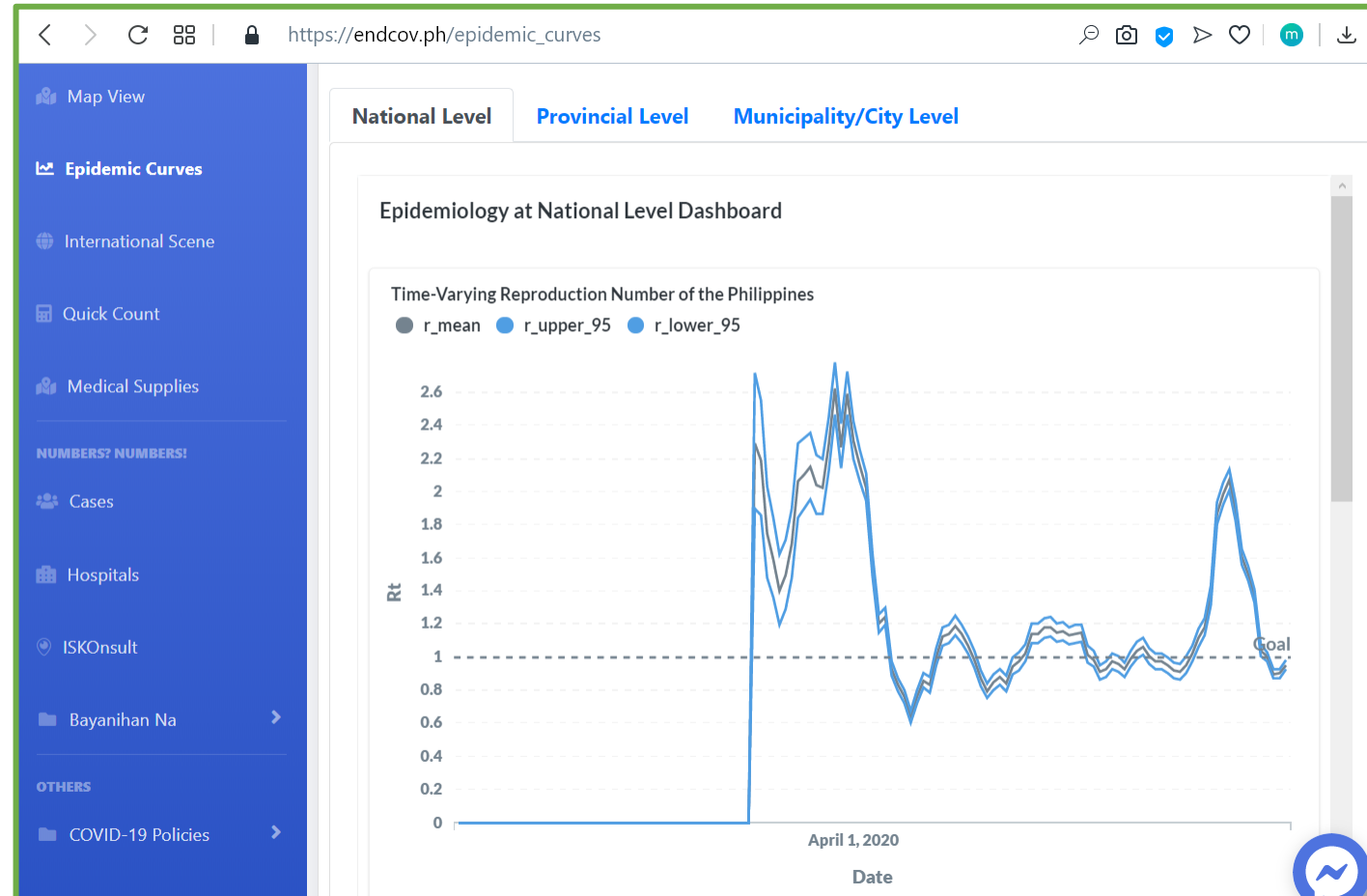
April 2 · 🌐

The UP COVID-19 Pandemic Response Team has created a web portal (<http://endcov.ph/>) mapping out COVID-19 cases in the country in clear, useful detail.

Read more here: <https://www.up.edu.ph/up-launches-web-portal-map-for-covid.../>

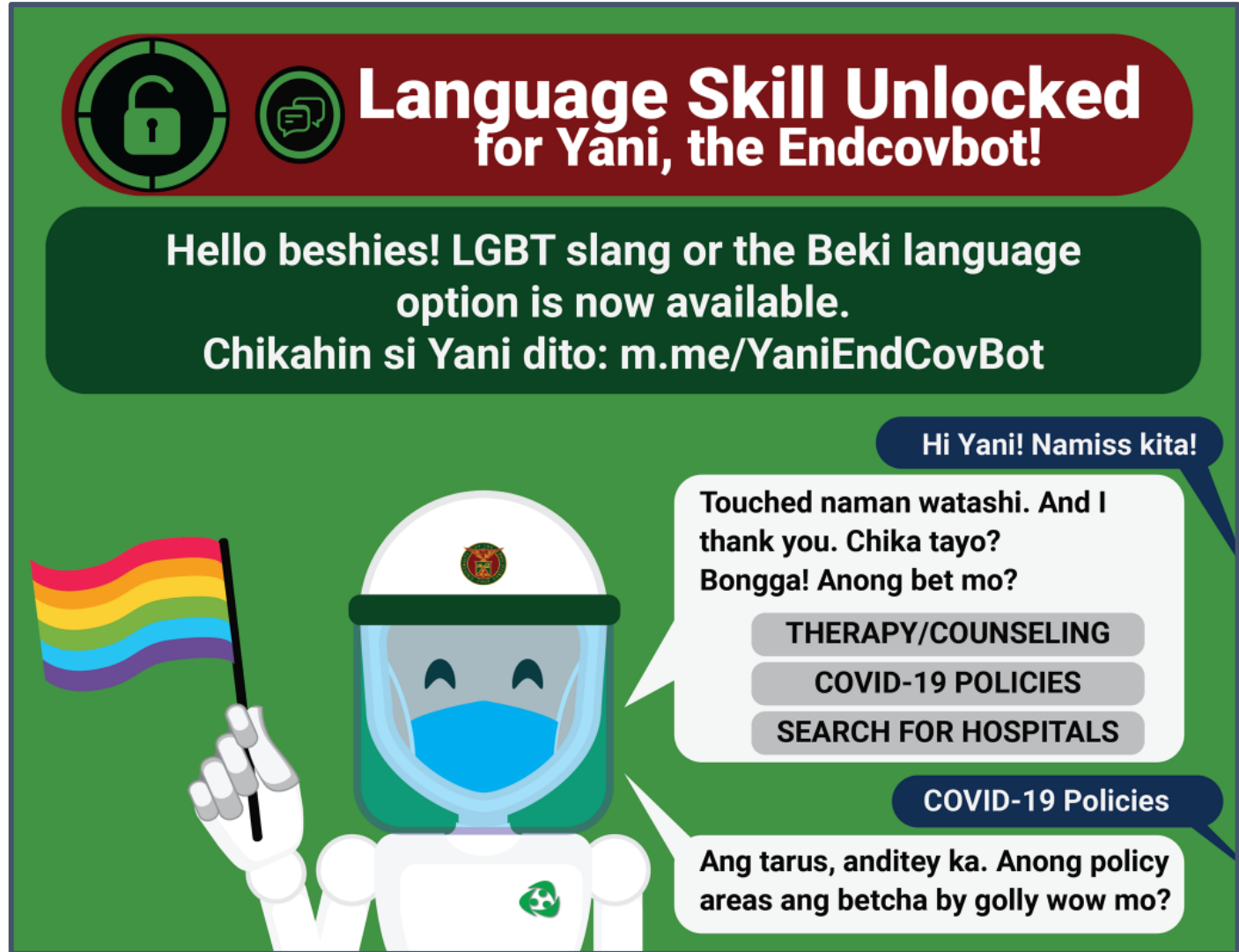


# endcov.ph





# Yani, the Endcov bot



**Language Skill Unlocked**  
for Yani, the Endcovbot!

Hello beshies! LGBT slang or the Beki language option is now available.  
Chikahin si Yani dito: [m.me/YaniEndCovBot](https://m.me/YaniEndCovBot)

Hi Yani! Namiss kita!

Touched naman watashi. And I thank you. Chika tayo? Bongga! Anong bet mo?

- THERAPY/COUNSELING
- COVID-19 POLICIES
- SEARCH FOR HOSPITALS

COVID-19 Policies

Ang tarus, anditey ka. Anong policy areas ang betcha by golly wow mo?

## Mga Potensiyal na Tagapagpakalat ng COVID-19 Batay sa Uri ng Trabaho

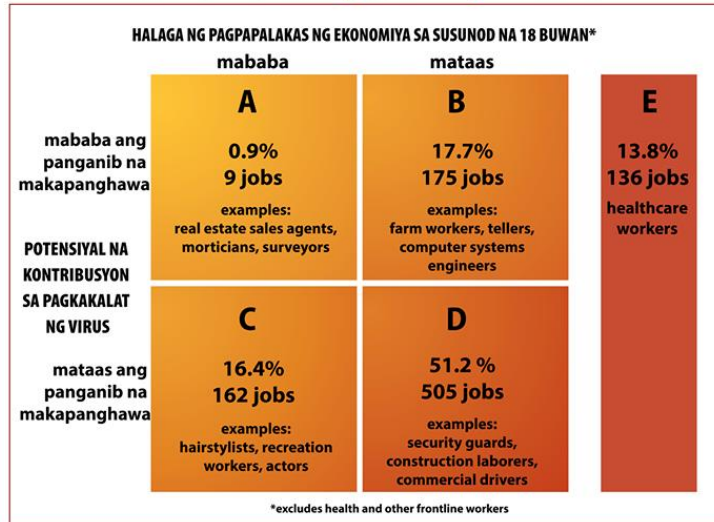


Figura 1: Klasipikasyon ng trabaho sa Pilipinas batay sa halaga sa ekonomiya para sa susunod na 18 buwan at ang potensiyal na kontribusyon sa pagkakalat ng virus.

**816 mula sa 987** na uri ng trabaho (82%) sa Pilipinas ang napakahalaga sa pagpapalakas ng ekonomiya sa susunod na 18 buwan.

Mayorya ng mga trabahong ito (65%) ang may mataas na potensiyal na maikalat ang sakit na COVID-19 dahil nangangailangan ito ng lapit-lapit o regular na interaksyon sa mga tao.



**Sanggunian:** Policy Note 7 / 20 May 2020: Post-ECQ Job Risks - Analysis and Recommendations  
UP COVID-19 Pandemic Response Team

upri.covid19@up.edu.ph

2

## Panganib sa Trabaho Batay sa Antas ng Kita

**Walang korelasyon** ang average na buwanang suweldo at antas ng panganib.

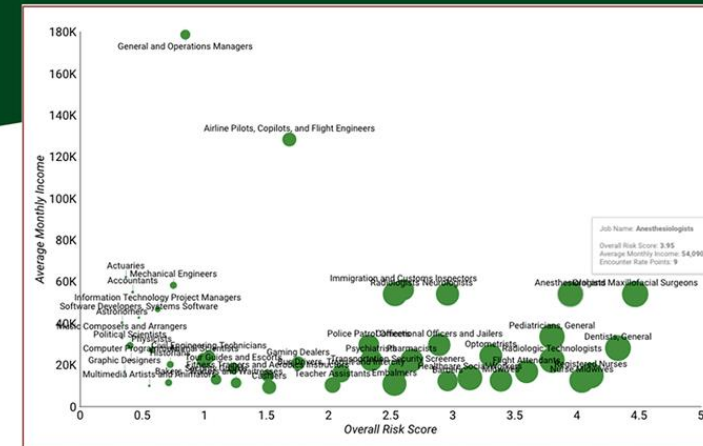


Figura 2: Average na Buwanang Kita ng mga Piling Trabaho at ang kanilang mga Iskor ng Panganib. Ang laki ng bilog ay batay sa average na bilang ng nakakasalamuhang mga tao sa bawat oras.

Gayunman, yaong mga nasa **braket ng mababang kita** ay maaaring mas mahirapang tustusan ang mga gastos sa pagpapaospital o mas mahirapang makabangon sa nawalang kita dahil sa pagkasuspindi sa trabaho.



**Sanggunian:** Policy Note 7 / 20 May 2020: Post-ECQ Job Risks - Analysis and Recommendations  
UP COVID-19 Pandemic Response Team

upri.covid19@up.edu.ph

4

To better inform the general public, the team has also redesigned their studies into **animation, video formats and flashcards** to be more accessible.





UP Resilience Institute

June 3 at 8:41 PM · 🌐

Antonaa i maana a "flattening the curve" on a miyapantar so curv  
Inibegay a UP COVID-19  
"Manga bababa a Semb  
Curve".

Batiyaangka saya h  
Published May 8, 2020  
Initogalin saki (Meranaw  
Noroniah O. Taher  
Musmira C. Bantuas-Ra  
Mona Miscille T. Domato  
Rayhana P. Ali

Layout i Ian Villanueva  
#UPCOVID19Response

**Antonaa i r  
flattening t**

Giyai so madidiang  
o isaka inged ka ad  
kapephakadake  
COVID-19 para ma  
ago so mga gumag  
kagampasi iran ko  
miyatanged a ade  
iyan a COVID 19.

(Flattening the curve refers  
measures that keep the dai  
cases at a manageable leve



Phoon sa: Ang Mga S



UP Resilience Institute

May 30 at 9:22 PM · 🌐

Ano ba iton "epidemic wave"? (Isinalin sa Waray-Leyte)

Ginhahatag inin nga *brief* sa *Waray-Leyte* (Kasaysayan sa Baybayin)

Nahinanabo iton Usa

Epidemic Wave" han

orihinal nga English r

ngan ginhubad ha Fil

Basaha an briefer ha

Basaha an briefer ha

**Ano ng**

Ito ang kurl  
ng bilang r  
marating

Number of cases

Sa pagbawas  
bababa ang ku



Mula sa: "Paliw  
Isinulat ni Mary  
UP COVID-19 Par  
upci.covid19

## Paliwanag Kung Paano Nagaganap ang Isang Epidemya at Ano ang Ibig Sabihin ng Epidemic Wave

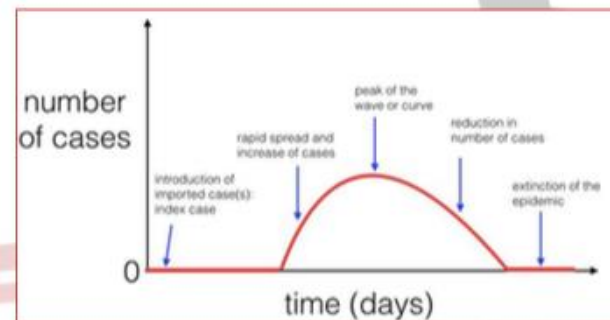
Mary Grace Dacuma, Ph.D.

University of the Philippines, Los Baños

(salin sa Filipino ng orihinal na Ingles)

Ipinagpapalagay na nagsimula ang epidemyang COVID-19 sa Pilipinas nang magkaroon ng imported na kaso mula sa isang tao o mga taong may impeksiyon na pumasok sa Pilipinas. Ang tao o mga taong may virus na ito ay ang pinakaunang natukoy na kaso ng nakahahawang sakit (index case) na nagkalat ng virus sa iba dito sa bansa. Hindi kabilang ang (mga) index case, lalo na yaong hindi naman naging dahilan ng lokal na transmisyon, sa bugso ng epidemya (epidemic wave).

Kapag nagkaroon ng lokal na transmisyon ng virus sa ibang tao, may panahon ng ingkubasyon (ibig sabihin nito na nahawa na ng virus ang isang tao ngunit wala pang anumang sintomas o klinikal na mga senyales) na karaniwang 5.2 araw hanggang 14 na araw lumalabas, lyon ang dahilan kung bakit may patag na linya pagkaraang magkaroon ng index case. (tingnan ang Pigura).



**Policy notes** are now available in English as well as Tagalog, Ilokano, Bikol Sentral, Waray, Cebuano, Hiligaynon, Aklanon, Kapampangan, Itawis, Chavacano de Zamboanga, Meranaw, and Bahasa Sug.

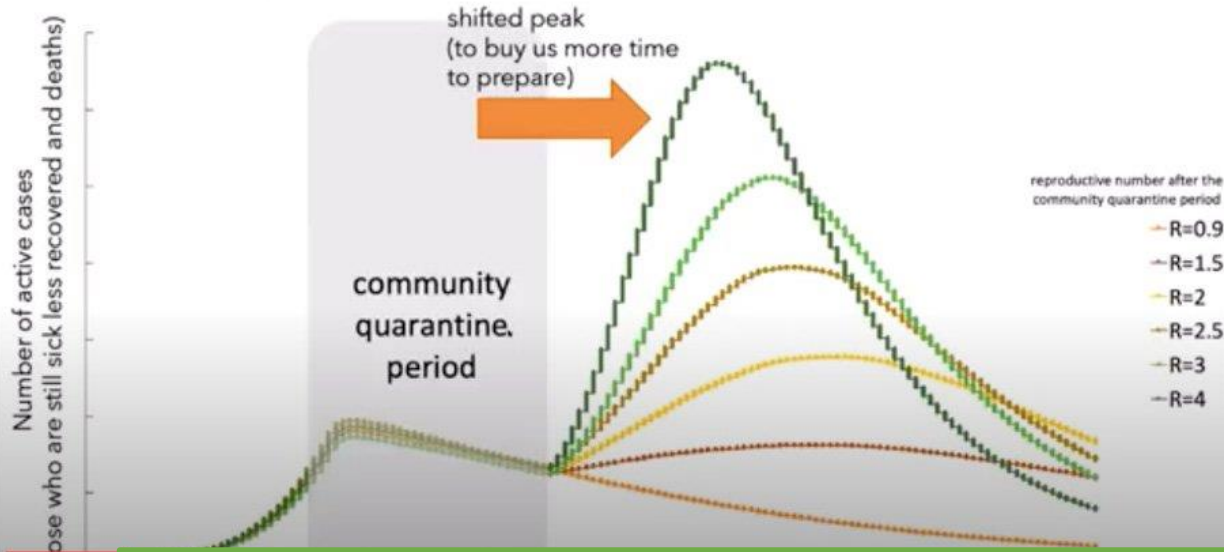




# COVID-19 Threshold Model

The response team has drafted a number of **policy notes** based on the results of their studies, including recommendations for a graduated activation of the ECQ that depends on the level of risk per area.

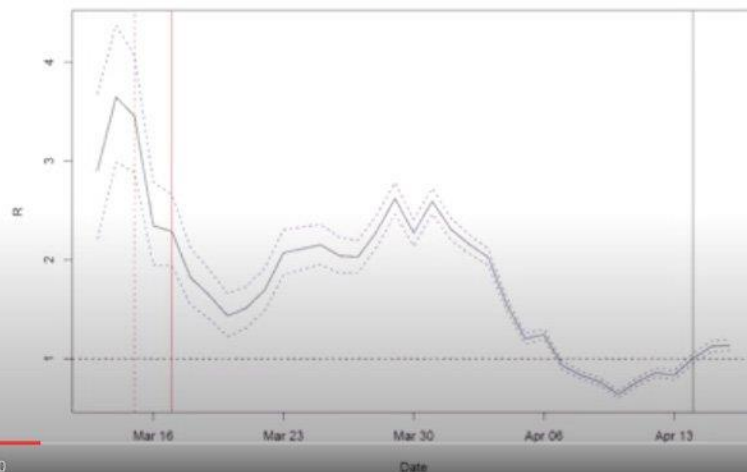
## EPIDEMIC PEAK DYNAMICS



## Time-Varying Reproduction Number $R_t$

Some Results: National  $R_t$  as of 16 April 2020

Philippines  $R_t$



# Working with the national government

Team members have also **given presentations to the President and the Inter-Agency Task Force on Emerging Infectious Disease (IATF)**. Given their highly specialized competencies, members have been called to join the IATF Technical Working Group on Anticipatory and Forward Planning.



# Coordination with the LGUs

COVID-19 Updates: Philippines, as of Wednesday, June 10, 2020 9:39:51 PM

This is a platform for LGUs to report COVID-19-related data per barangay. Data collected can help LGUs monitor and gain insights on the spread of COVID-19. With barangay-level real-time data, we can #EndCovidAsOne.


**Sign In**

Email Address \*

Password \*

**SIGN IN**

[Forgot password?](#) [Don't have an account? Sign Up](#)

 [Link to DILG Memorandum](#)

**COVID-19 INFORMATION**


- Dashboard View
- Map View
- Epidemic Curves
- International Scene
- Quick Count**
- Medical Supplies


**NUMBERS? NUMBERS!**

- Cases
- Hospitals
- ISKONSult
- Bayanihan Na

**OTHERS**

- COVID-19 Policies
- Advisories

  
Republic of the Philippines  
DEPARTMENT OF THE INTERIOR AND LOCAL GOVERNMENT  
DILG-NAPOLCOM Center, EDSA corner Quezon Avenue, West Triangle, Quezon City  
<http://www.dilg.gov.ph>

  
OFFICIAL RELEASED  
Signature: *[Signature]*  
Date: 04-28-2020  
Time: 2:50 PM  
RECORDS SECTION  
DEPARTMENT OF THE INTERIOR AND LOCAL GOVERNMENT

**MEMORANDUM**

**TO :** ALL PROVINCIAL GOVERNORS, CITY AND MUNICIPAL MAYORS, DILG REGIONAL DIRECTORS, BARMM MINISTER OF LOCAL GOVERNMENT, AND ALL OTHERS CONCERNED

**SUBJECT :** CORONAVIRUS DISEASE 2019 (COVID-19) INFORMATION FOR REAL-TIME MODELING AND ANALYTICS

**DATE :** APR 28 2020

The making of sound decisions requires accurate and timely information. And the decisions that the Philippine government must take in addressing the COVID-19 pandemic are crucial to the lives and livelihood of every Filipino. The considerations to improve the quality of this decision-making process depends largely on the quality of the data used in official reports, testing accuracy, monitoring, and the faithful accounts of cases, among others. The model estimates would greatly improve if nationwide barangay-level COVID-19 data are available to the decision-makers on a daily basis.

These barangay-level data and metrics are important in accurately projecting the rate of spread, identifying the locations of hotspots and outbreaks and the future trends on a daily basis. These are critical in helping the government reach sound policy decisions, such as when and where to restart economic activities.


Thus, in the best interest of the Filipino nation, all city and municipal local government units are hereby directed to generate and report the required barangay-level data using the automated online LGU data collection system developed by the University of the Philippines COVID19 Pandemic Response Team, which is accessible through the <https://endcov.ph> dashboard. The reporting shall be assigned to the city/municipal disaster risk reduction and management officer who must fill in the required barangay-level cumulative data not later than 6:00 PM every day.

Attached to this Memorandum is a short instructional video on how to sign up for an account and input the required cumulative data.


Provincial Governors are requested to support their component LGUs in ensuring the success of this very important endeavor.

All DILG Regional Directors, and the BARMM Minister of Local Government, are hereby directed to cause the widest dissemination of this Memorandum.

For information and immediate compliance.

  
**EDUARDO M. AÑO**  
Secretary

Attachment:  
<https://youtu.be/mWYV6gBscKw>

  
DILG-ORIG-ANDRES-001  
22

Aside from working with national and local government units, the UP response team has been collaborating with academics from the National University of Singapore, University of California Davis and University College London as well as local academic institutions.

On the international front, the team has joined the Forecast-based Warning, Analysis and Response Network (FOREWARN), an organization of academics, scientists and humanitarian workers.









# About UP COVID-19 Pandemic Response Team

## ► Projects I am Involved In:

1. Epidemic Curves of the UP COVID-19 Pandemic Response Team
2. ENDCOV Map on current active cases and probability of outbreak
3. LGU Data Analytics Group
4. Compendium of COVID-19 Statistics for Island Groups, Regions, Provinces, Cities, and Municipalities
5. PSPHP Graphs, Time Varying-R Dashboard with the LEADS 4 Health Security and Resilience



# About UP COVID-19 Pandemic Response Team

- ▶ Platform:
  - ▶ <https://endcov.up.edu.ph>
- ▶ Contact Us:
  - ▶ [upri.covid19@up.edu.ph](mailto:upri.covid19@up.edu.ph)





# About the L4H Consortium

# About the L4H Consortium

- ▶ Leading Evidence-based Actions through Data Science for Health Security and Resilience
- ▶ A consortium of data scientists, physicians, mathematicians, and epidemiologists, convened by the Philippine Society of Public Health Physicians (PSPHP)
- ▶ website: <https://leads4health.org/l4h/>

# About the L4H Consortium

## Epidemic Indicators

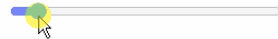
Select KPI Date:

2020-06-24

Select Trend Chart:

Time Varying R (Rt)

Chart Time Period: Past 29 days



### Time Varying R (Rt)

Related to the basic reproductive number ( $R_0$ ), the effective reproductive number is the transmission rate of the virus at time ( $t$ ). Threshold for  $R_t$  is 1.0, if for consecutive days  $R_t$  is below 1.0 it suggests that the community has started to effectively control the spread of the disease.



## CO-INFORM Risk Maps

Select Area:

Select Area:  
Region  
Province  
NCR

Heatmap indicator to display heatmaps.



Select Table:

# About the L4H Consortium

## Our Members

- **Jason V. Alacapa, MD, MBA, MPH, MHM** - Consortium Co-convenor, Philippine Society of Public Health Physicians; CEO, metaHealth Insights and Innovation, Inc.
- **Geminn Louis C. Apostol, MD, MBA** - Assistant Professor and Environmental Health Specialist, Ateneo School of Medicine and Public Health
- **DJ Darwin R. Bandy, DVD, PhD(c)** - Assistant Professor, University of the Philippines
- **Peter Julian A. Cayton, PhD** - Associate Professor II, University of the Philippines Diliman
- **Lester Sam Araneta Geroy, MD, MPH, MSc** - President, Philippine Society of Public Health Physicians
- **Dominic Ligot** - Chief Technology Officer, CirroLytix Research Services; Board Member, AAP; Board Member, PCIJ; Lecturer, University of Asia and the Pacific
- **Jason D. Ligot, MD** - Director, Organic Intelligence
- **Robert Neil F. Leong, MSc, PhD(c)** - Chief Data Officer, metaHealth Insights and Innovation, Inc.; PhD Researcher, The University of New South Wales - Sydney; Assistant Professor 2, Mathematics and Statistics Dept. - De La Salle University
- **Lionel A. Peters, MD, MPM** - Member, Philippine Society of Public Health Physicians
- **Michael Angelo B. Promentilla, PhD** - Professor and Research Lead, Waste and Resource Management Unit, Center for Engineering and Sustainable Development Research, De La Salle University
- **Jomar F. Rabajante, PhD** - Professor, University of the Philippines Los Baños and University of the Philippines Open University
- **Miguel Antonio S. Salazar, MD, MSc, Dr.sc.hum.(c)** - Consortium Co-Convenor, Philippine Society of Public Health Physicians
- **Xerxes T. Seposo, MENRM, MPH, PhD** - Assistant Professor, Nagasaki University
- **Jan Gil G. Sarmiento, MSc (ongoing)** - Instructor, University of the Philippines Diliman
- **Theresa Rosario Tan, MABA (ongoing)** - Associate Consultant, CirroLytix Research Services
- **April Anne S. Tigue, MSc, PhD(c)** - Research staff, Waste and Resource Management Unit, Center for Engineering and Sustainable Development Research, De La Salle University

# About the L4H Consortium

## Our Institutions



# Example Projects





# Example Projects

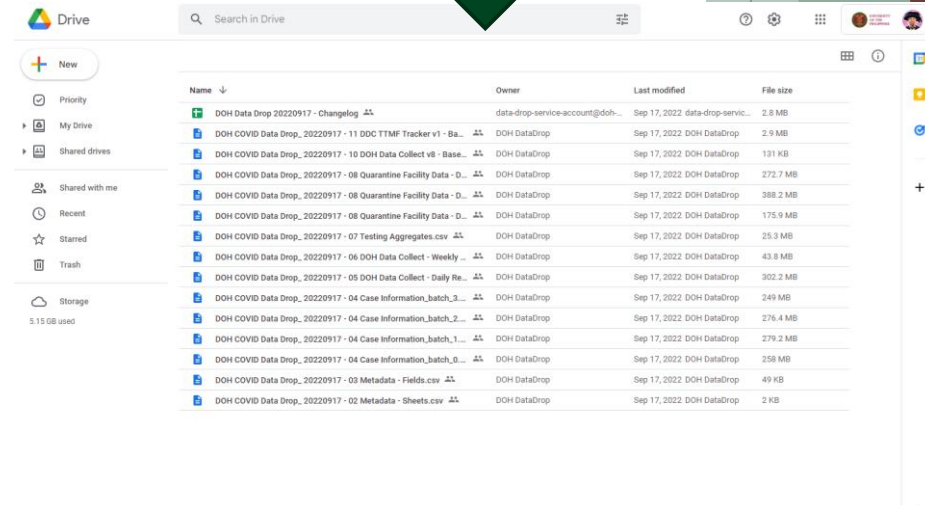
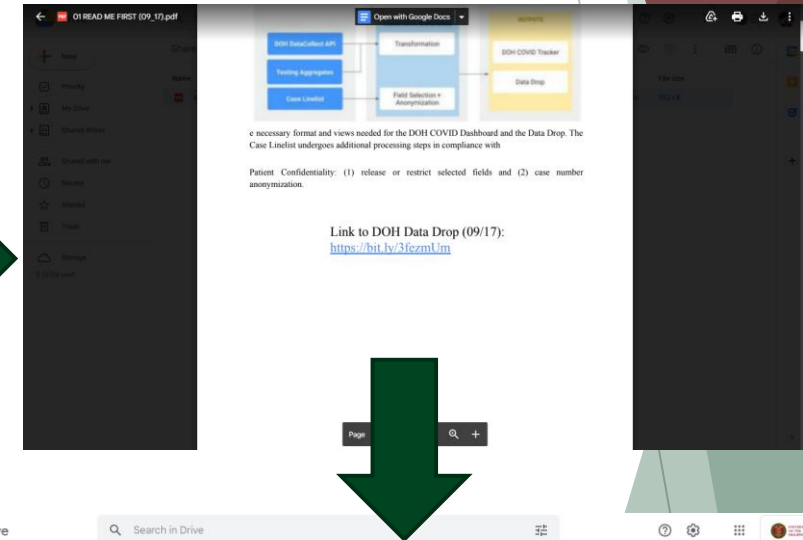
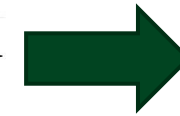
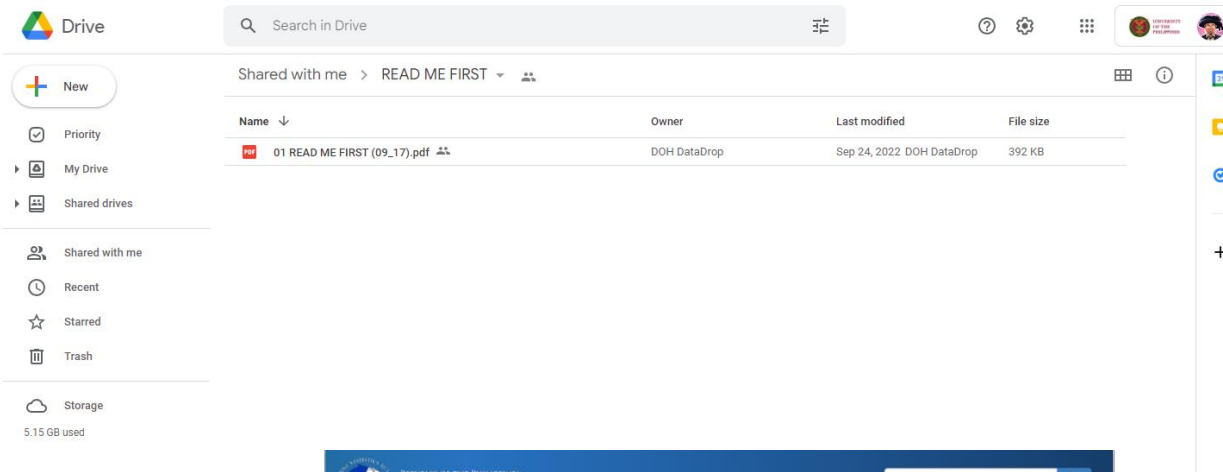
1. Epidemiological Statistics for COVID-19 (w/UP PRT)
  - > Compiling epidemiological statistics for website and PDF statistical reports outputs
2. Short-term Forecasting of COVID-19 Cases (w/ L4H)
  - > Forecasting for national and regional reported case counts of COVID-19
3. Mapping COVID-19 Cases (w/ UP PRT)
  - > Mapping the distribution of active cases and case per 100,000 population by barangay in the Philippines

# Demonstrating the Workflow with a Team using Example Projects

# Demonstrating the Workflow with a Team using Example Projects

| Project  | Import   |
|--|--|
| Epidemiological Statistics for COVID-19 (w/UP PRT) | Primary Data Source:<br><br>DOH Data Drop for the Case Information Dataset [ <a href="https://bit.ly/DataDropPH">https://bit.ly/DataDropPH</a> ]<br><br>Secondary Data Source:<br><br>PSGC Data with 2020 Population Counts [ <a href="https://psa.gov.ph/classification/psgc/downloads/2_PSGC%20Q-2022-Publication-Datafile.xlsx">https://psa.gov.ph/classification/psgc/downloads/2_PSGC%20Q-2022-Publication-Datafile.xlsx</a> ]<br><br>Researchers:<br>Trixie Delmendo, UPRI;<br>Peter Julian Cayton, UP PRT |
| Short-term Forecasting of COVID-19 Cases (w/L4H)   |  |
| Mapping COVID-19 Cases (w/ UP PRT)                 |  |

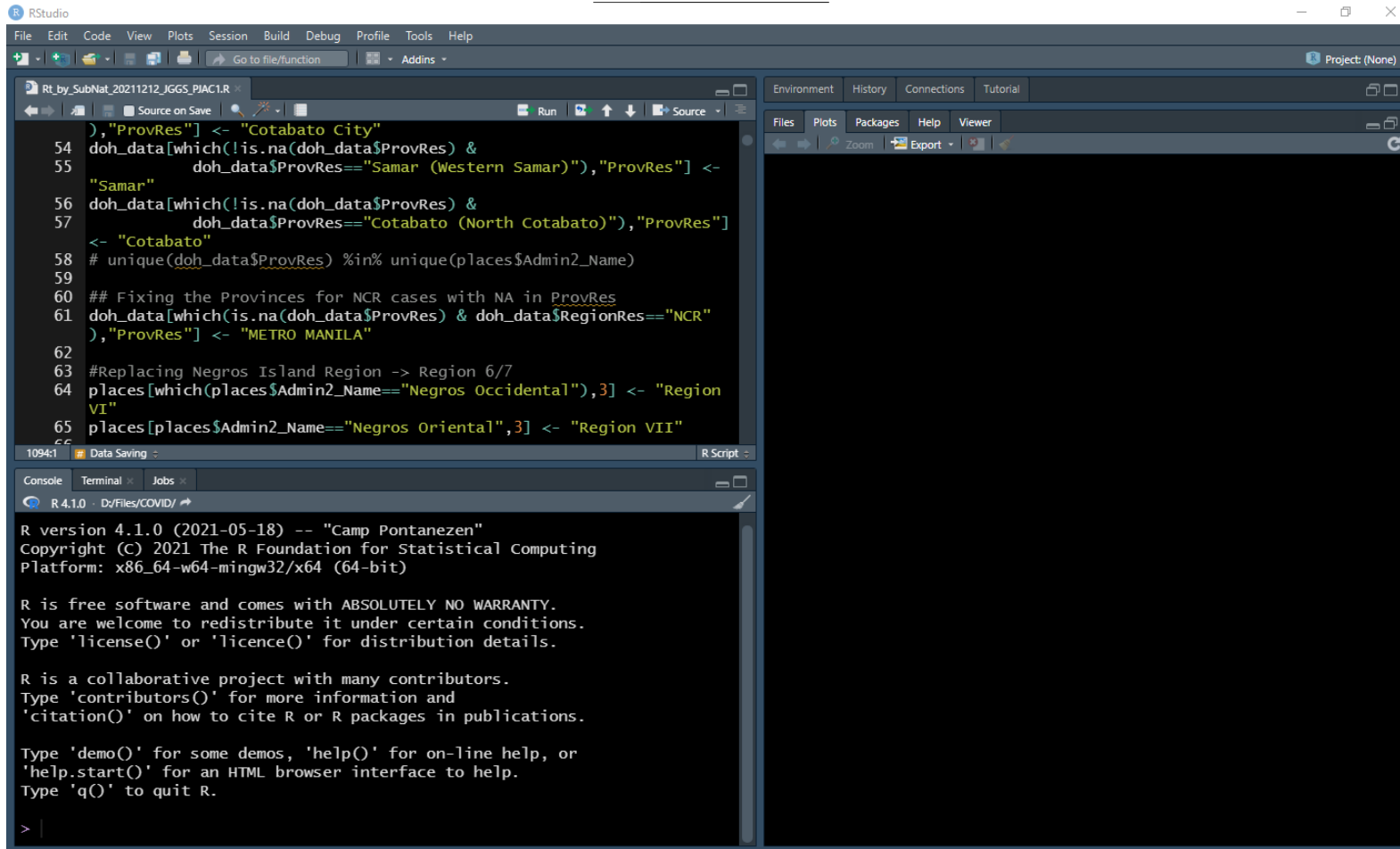
# Demonstrating the Workflow with a Team using Example Projects



# Demonstrating the Workflow with a Team using Example Projects

| Project  | Tidy   |
|--|--|
| Epidemiological Statistics for COVID-19 (w/UP PRT) | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Reconcile and correct residence variable matched with PSGC data, assuming Barangay PSGC as supreme location variable</li><li>➤ Reconcile between dataset versions to add missing recovery dates</li></ul> <p>Researchers:</p> <p>Trixie Delmendo, UPRI;<br/>Jan Gil Sarmiento, UPD SS<br/>Peter Julian Cayton, UP PRT</p> |
| Short-term Forecasting of COVID-19 Cases (w/L4H)   |  |
| Mapping COVID-19 Cases (w/ UP PRT)                 |  |

# Demonstrating the Workflow with a Team using Example Projects



The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains R code for data manipulation. The code includes comments and assignments for 'ProvRes' and 'Region' based on specific conditions.
- Environment:** Shows the current environment with no objects loaded.
- Files:** Shows the file explorer with a project named 'Project: (None)'.
- Console:** Displays the R version (4.1.0) and the R Foundation for Statistical Computing logo. It also shows the R license and a list of contributors.

```
), "ProvRes"] <- "Cotabato City"
54 doh_data[which(!is.na(doh_data$ProvRes) &
55 doh_data$ProvRes=="Samar (Western Samar)", "ProvRes"] <-
  "Samar"
56 doh_data[which(!is.na(doh_data$ProvRes) &
57 doh_data$ProvRes=="Cotabato (North Cotabato)", "ProvRes"]
  <- "Cotabato"
58 # unique(doh_data$ProvRes) %in% unique(places$Admin2_Name)
59
60 ## Fixing the Provinces for NCR cases with NA in ProvRes
61 doh_data[which(is.na(doh_data$ProvRes) & doh_data$RegionRes=="NCR"
  ), "ProvRes"] <- "METRO MANILA"
62
63 #Replacing Negros Island Region -> Region 6/7
64 places[which(places$Admin2_Name=="Negros Occidental"),3] <- "Region
  VI"
65 places[places$Admin2_Name=="Negros Oriental",3] <- "Region VII"
66
```

R version 4.1.0 (2021-05-18) -- "Camp Pontanezen"  
Copyright (C) 2021 The R Foundation for Statistical Computing  
Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.

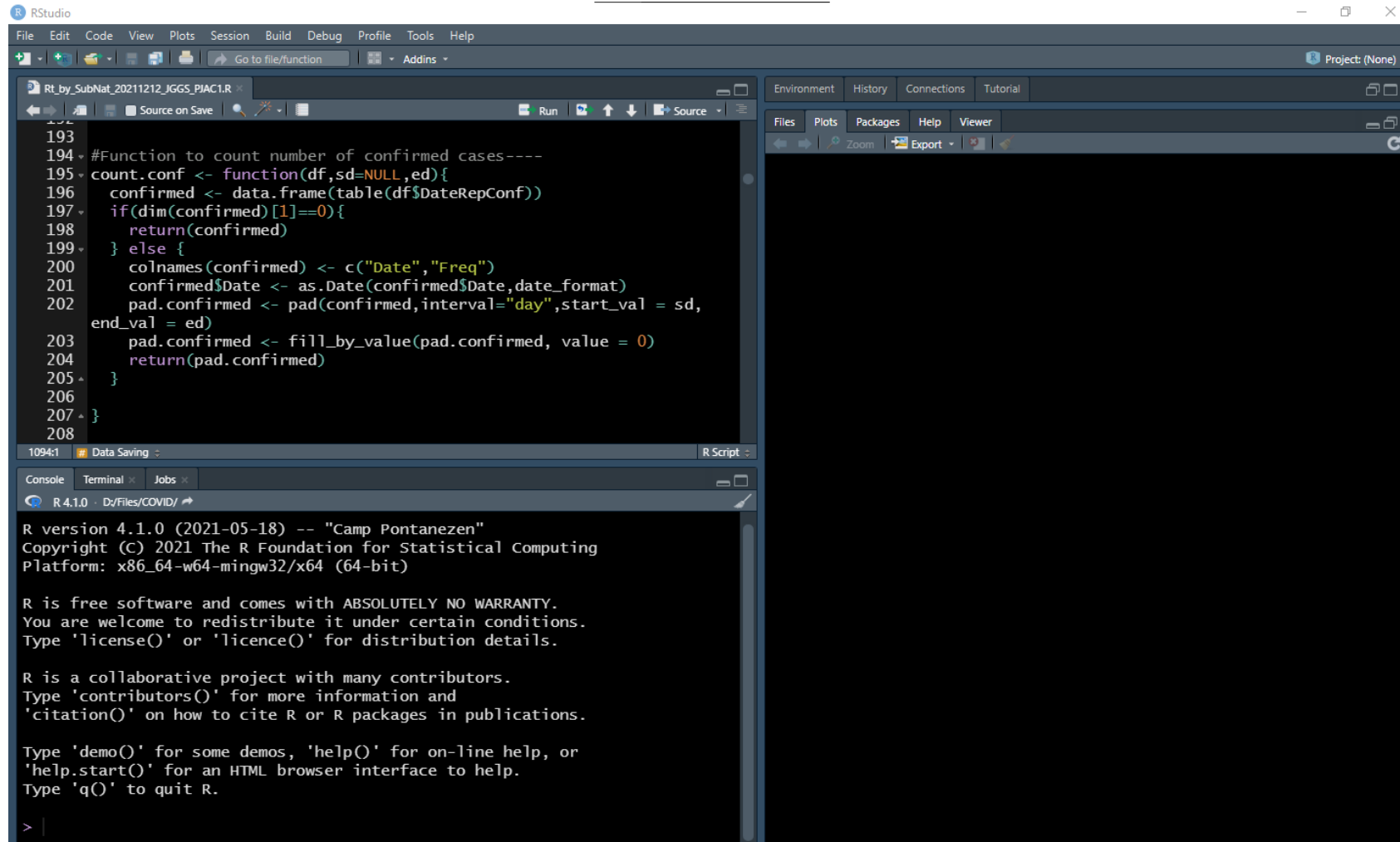
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

> |

# Demonstrating the Workflow with a Team using Example Projects

| Project  | Transform   |
|--|---|
| Epidemiological Statistics for COVID-19 (w/UP PRT) | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Group data by geographic hierarchy: 1) National, 2) Island Group, 3) Regional, 4) Provincial, 5) City/Municipality</li><li>➤ Compute for cumulative and new counts for cases, recoveries, deaths, and active cases per date per hierarchy</li><li>➤ Compute for case rates such as recovery rates, fatality rates, and doubling times</li></ul> <p>Researchers: Jan Gil Sarmiento, UPD SS; Peter Julian Cayton, UP PRT</p> |
| Short-term Forecasting of COVID-19 Cases (w/ L4H)  | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Group data by geographic hierarchy: 1) National and 2) Regional</li><li>➤ Compute for new counts for cases per date per hierarchy</li></ul> <p>Researchers: Jan Gil Sarmiento, Simon Bismonte, Maryliz Zubiri, and Nicole Uy</p>   |
| Mapping COVID-19 Cases (w/ UP PRT)                 | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Group data by geographic location: Barangay via PSGC</li><li>➤ Compute for total counts for cases, recoveries, deaths, and active cases per barangay</li><li>➤ Compute for case per 100,000 population using PSGC population data and case counts by barangay.</li></ul> <p>Researchers: Trixie Delmendo, UP RI; Peter Julian Cayton, UP PRT</p>   |

# Demonstrating the Workflow with a Team using Example Projects



The screenshot displays the RStudio interface. The main editor window shows an R script with a function `count.conf` designed to count confirmed cases. The function takes a data frame `df`, an optional standard deviation `sd`, and an end date `ed`. It creates a date-frequency table, pads the data to the end date, and fills missing values with zero. The environment pane on the right is currently empty, showing 'Project: (None)'. The console at the bottom shows the R startup message for version 4.1.0, dated 2021-05-18, on a Windows 64-bit platform.

```
193  
194 #Function to count number of confirmed cases----  
195 count.conf <- function(df,sd=NULL,ed){  
196   confirmed <- data.frame(table(df$DateRepConf))  
197   if(dim(confirmed)[1]==0){  
198     return(confirmed)  
199   } else {  
200     colnames(confirmed) <- c("Date","Freq")  
201     confirmed$Date <- as.Date(confirmed$Date,date_format)  
202     pad.confirmed <- pad(confirmed,interval="day",start_val = sd,  
203   end_val = ed)  
204     pad.confirmed <- fill_by_value(pad.confirmed, value = 0)  
205     return(pad.confirmed)  
206   }  
207 }  
208
```

1094:1 Data Saving - R Script -

Console Terminal Jobs

R 4.1.0 · D:/Files/COVID/

R version 4.1.0 (2021-05-18) -- "Camp Pontanezen"  
Copyright (C) 2021 The R Foundation for Statistical Computing  
Platform: x86\_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.  
You are welcome to redistribute it under certain conditions.  
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.

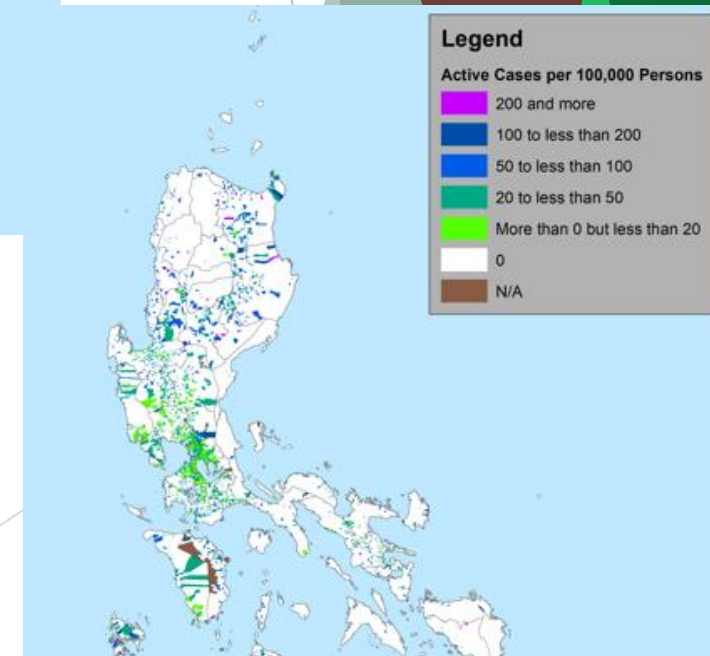
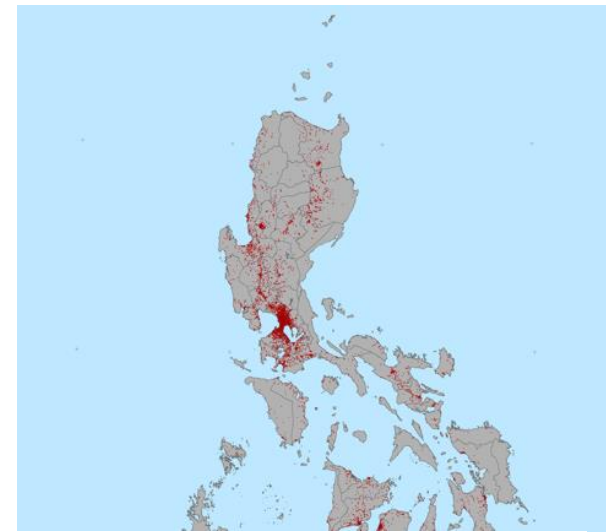
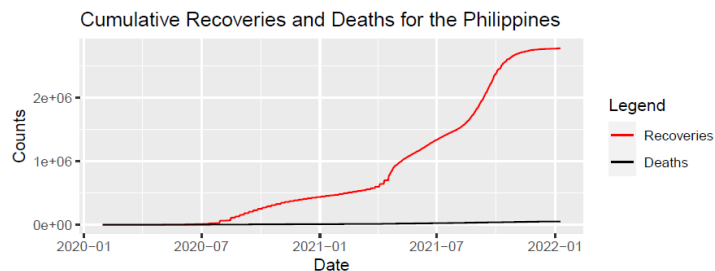
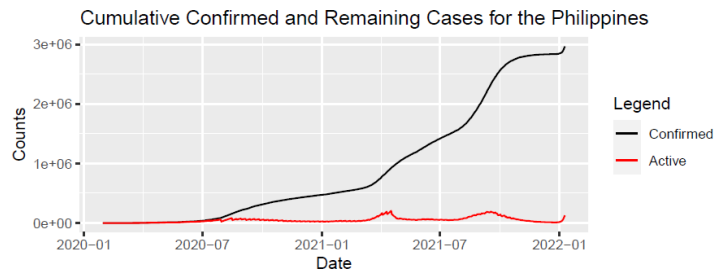
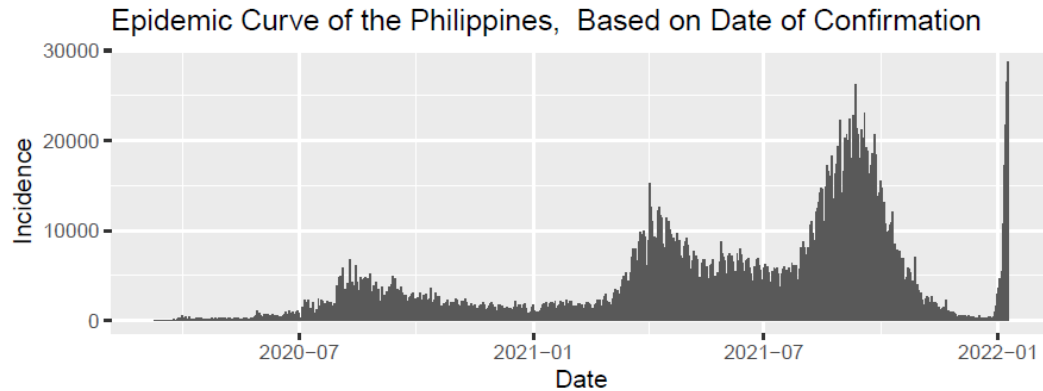
> |



# Demonstrating the Workflow with a Team using Example Projects

| Project  | Visualize  |
|--|--|
| Epidemiological Statistics for COVID-19 (w/UP PRT) | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Generate plots of cases, recoveries, deaths, and active cases per hierarchy</li></ul> <p>Researchers: Jan Gil Sarmiento, UPD SS; Peter Julian Cayton, UP PRT</p>  |
| Short-term Forecasting of COVID-19 Cases (w/L4H)   | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Generate plots of cases per hierarchy</li></ul> <p>Researchers: Jan Gil Sarmiento, Simon Bismonte, Maryliz Zubiri, and Nicole Uy</p>  |
| Mapping COVID-19 Cases (w/ UP PRT)                 | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Produce a map of active case distribution in which case dots are randomly distributed within their barangay of residence</li><li>➤ Produce a color map of case per 100,000 by barangay of residence with the barangay colors done by belonging into an interval</li></ul> <p>Researchers:<br/>Trixie Delmendo, UP RI; Steffanie Chua, UP RI; Jake Mendoza, UP RI; Feye Andal, UP RI; UP Resilience Youth Mappers; Peter Julian Cayton, UP PRT</p> |

# Demonstrating the Workflow with a Team using Example Projects

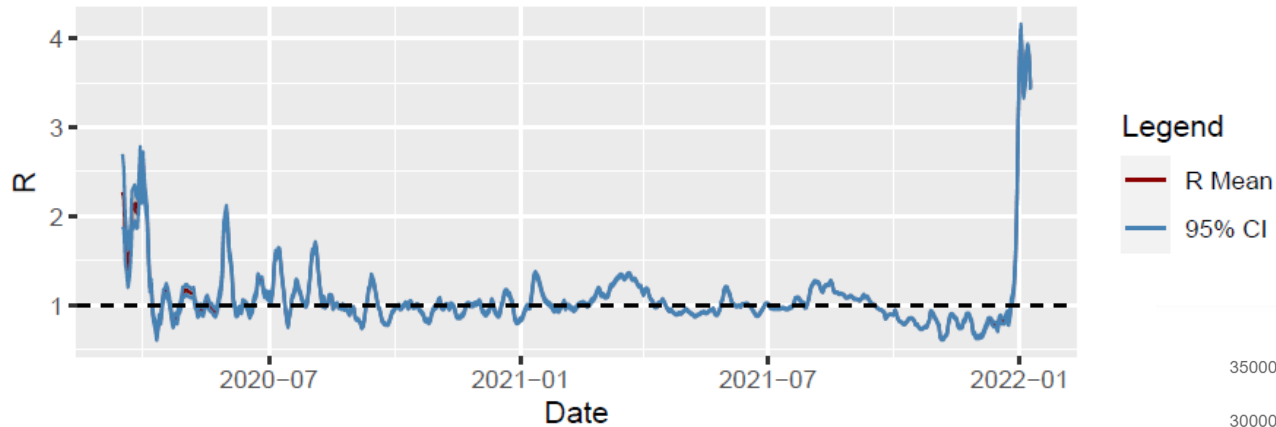


# Demonstrating the Workflow with a Team using Example Projects

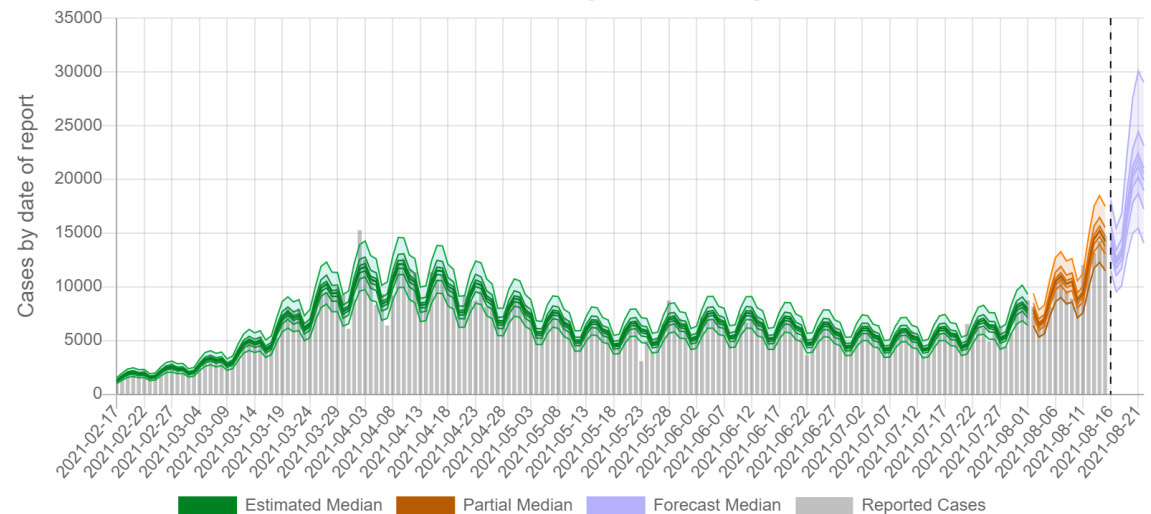
| Project  | Model   |
|--|---|
| Epidemiological Statistics for COVID-19 (w/UP PRT) | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Generate the time-varying reproduction number based on an existing Bayesian estimation technique</li></ul> <p>Researchers: Jan Gil Sarmiento, UPD SS; Peter Julian Cayton, UP PRT</p>                            |
| Short-term Forecasting of COVID-19 Cases (w/ L4H)  | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Generate the short-term forecasts assuming a model based on the Epiforecasts methodology</li></ul> <p>Researchers: Jan Gil Sarmiento, Simon Bismonte, Maryliz Zubiri, Peter Julian Cayton, Robert Neil Leong</p> |
| Mapping COVID-19 Cases (w/ UP PRT)                 | (none)  |

# Demonstrating the Workflow with a Team using Example Projects

Rt of the Philippines, Based on Date of Confirmation



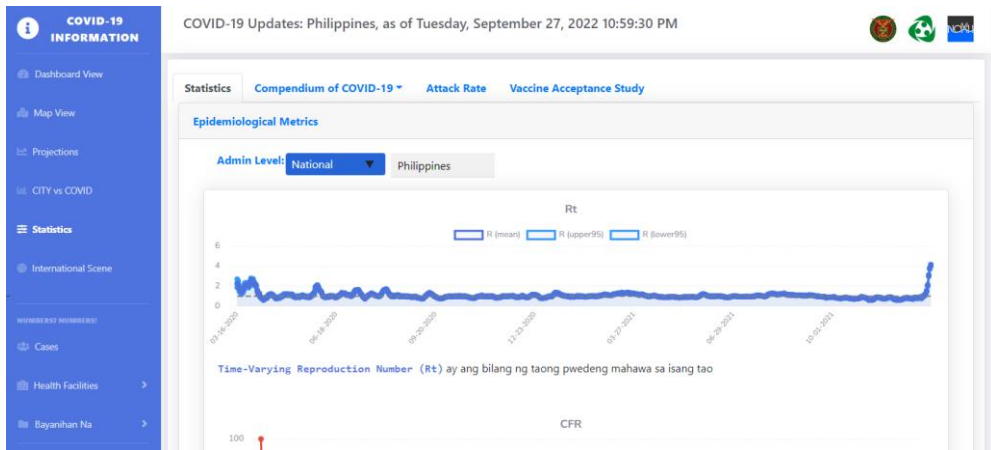
National Cases by Date of Report



# Demonstrating the Workflow with a Team using Example Projects

| Project  | Communicate   |
|--|---|
| Epidemiological Statistics for COVID-19 (w/UP PRT) | <p>Procedures:</p> <ul style="list-style-type: none"><li>➤ Publish a website page for uploading the statistics</li><li>➤ Generate a PDF report file</li></ul> <p>Researchers: Feye Andal and developers from the UP RI; Peter Julian Cayton, UP PRT</p> |
| Short-term Forecasting of COVID-19 Cases (w/ L4H)  | <p>Procedure:</p> <ul style="list-style-type: none"><li>➤ Publish the results through the L4H website</li></ul> <p>Researchers: Dominic Ligot, Mark Toledo, and Angelica Mhay Salazar</p>   |
| Mapping COVID-19 Cases (w/ UP PRT)                 | <p>Procedure:</p> <ul style="list-style-type: none"><li>➤ Produce maps based on the results</li></ul> <p>Researchers: Trixie Delmendo, UP RI; Steffanie Chua, UP RI; Jake Mendoza, UP RI; mappers and research of the UP RI and UP RI Youthmappers,</p> |

# Demonstrating the Workflow with a Team using Example Projects



## Compendium of Philippine COVID-19 Statistics

as of Jan 9, 2022

### Volume I: Island Groups, Regions, and Provinces

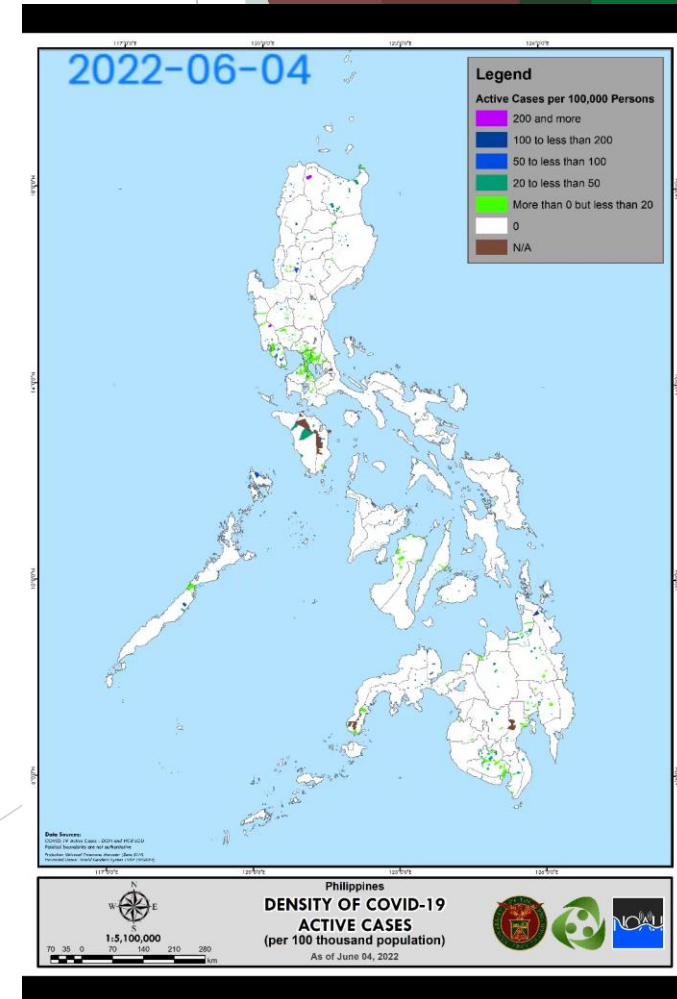
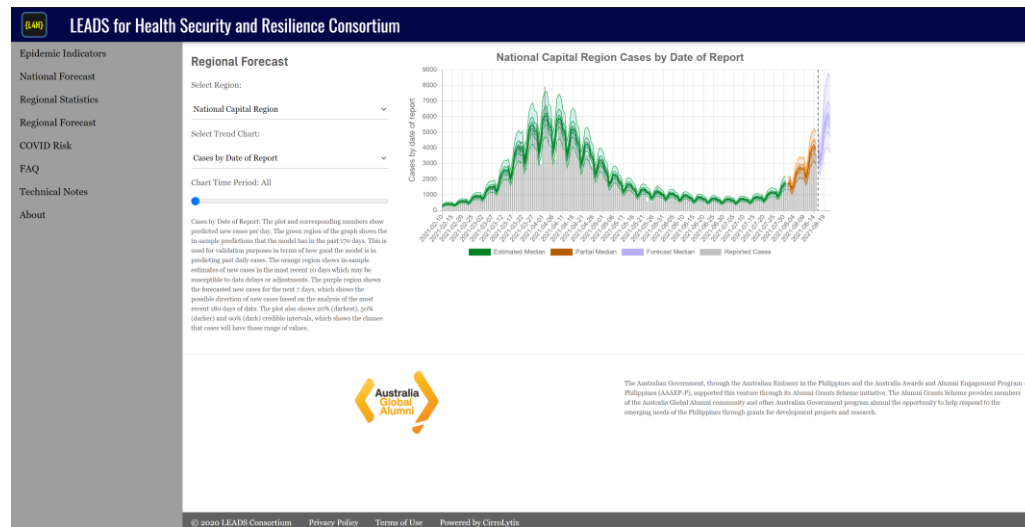
Peter Julian Cayton, Jan Gil Sarmiento, Trixie Delmendo, UP COVID-19 PRT, and L4H

2022-01-10

#### About the Authors

Peter Julian Cayton is an associate professor at the School of Statistics, University of the Philippines Diliman. His Google Scholar page is at [https://scholar.google.com.au/citations?user=KtH\\_mGEAAAAJ](https://scholar.google.com.au/citations?user=KtH_mGEAAAAJ). His email is [pacayton\(at\)up.edu.ph](mailto:pacayton(at)up.edu.ph).

Jan Gil Sarmiento is a Master of Science (Statistics) student and former instructor at the School of Statistics, University of the Philippines Diliman.



# Closing Remarks

# Closing Remarks

- ▶ Data science involves not only technical skills such as programming, statistics, and subject matter knowledge, but also communication and team collaboration
- ▶ By making data science projects in teams, the burden per person is reduced and high-quality data products are produced.
- ▶ Utilize each member's comparative advantage skills to maximize quality of output
- ▶ Build cohesion and empathy as a team of analysts and scientists producing data products



Thanks and stay safe always!