**Indonesia’s Report on:**

**Covid-19**

**Pandemic in 2020**

**Author :** Kevin Elfri

**Date created :** 18 Mar 2021

**Contact :** [kevinelfri@ymail.com](mailto:kevinelfri@ymail.com)

# Preface

In the era of data driven actions and information disclosure to fulfill the most basic human nature and rights called 'curiosity', we thrive this pandemic together and do our best to make a better future. Scientist, Engineer, Medical Personnel, Government, and all of the human work together and cooperate with respect of their expertise to save others life and survive this global pandemic. By showing an Information to the public about what have the scientists and medical personnel done to fight this pandemic, that mean the rights to know and the data transparency is fulfilled. This report will show the information and insights generated by implementing data mining techniques.

In this section we will metion the pupose and motivation behind the creation of this report paper. Please keep in mind this paper is not for further research, academic citation, and the most important motivation; this paper is intended to demonstrate data mining or data science techniques only. The insights generated from the data source may not be accurate due to the author’s limit.

# Table of content

[Preface 2](#_Toc67229901)

[Table of content 3](#_Toc67229902)

[Abstract 4](#_Toc67229903)

[Introduction 5](#_Toc67229904)

[Literature Review 6](#_Toc67229905)

[Methodology 6](#_Toc67229906)

[The Results 7](#_Toc67229907)

[Discussion 10](#_Toc67229908)

[Conclusion 10](#_Toc67229909)

[References 11](#_Toc67229910)

# Abstract

Indonesia’s Report on Covid-19 pandemic in 2020 is a report paper targeted mainly to the Indonesian citizen about the coronavirus spread cases and the government institutions reaction on the pandemic over the months of 2020. With descriptive and inferential statistic methodologies can be hypothesised and report that Indonesia is the country that has the highest covid-19 case among its neighbors Malaysia and Singapore, more severely that Indonesia’s Strictness policy related to covid-19 is not affecting the reduction of new cases (need further research). Key finding of this report; Indonesian stakeholders and every entities whose are responsible for making policy and rules are suggested to improve stricness and emphasize on making weekdays (Monday, Tuesday,Wednesday and Thursday) more strict to implement WHO daily activity,outdoor, and business protocol guide to supress covid-19 spread. For further research and data collection are highly encouraged to collect mroe possible factor leading to covid-19 mortality such as medical facilitiy features, comorbid desease, age of patient, and etc.

**Initial Plan for Data Exploration**

* Import the dataset
* Create subset of Indonesia’s data
* Create subset of Malaysia’s data
* Create subset of Singapore’s data
* Clean all of them
* Feature selection
* Perform EDA to get descriptive and inferential statistics
* Visualize
* Pearson correlation test
* Hypothesis testing

**Actions taken for data cleaning and feature engineering**

* Drop missing value by column if the missing value above 60%
* Drop missing value by row if the missing value below 3 %
* Extract days of week, days of month, year by parsing date feature
* Drop redundant features
* Fill missing value by reasonable and can be accounted for

# Introduction

Covid-19 has been declared as global pandemic by the world health organization (WHO) since 11 March 2020 and based on the data that we have, the corona virus entered Indonesia first time is on March 2020. In this report will show insights that can be an actionable manner to Indonesian public consumption. Also in this report will compare the Indonesia’s Covid-19 statistics and performances to the nearest two neighbors countries called Malaysia and Singapore.

This report will emphasize on the new covid-19 cases, total cases, governmnet stringency index and the timestamp features on three countries; Indonesia, Malaysia, and Singapore then genererate insights that can be an actionable manner.

# Literature Review

World Health Organization Situation Report – Indonesia is the previous similiar report published by World Health Organization, reporting Indonesia’s pandemic situation with more accurate results and I will highly encourage to all readers to refer and take the information for further research from the WHO release papers. Unfortunately, the date of WHO situation report mentioned as 27 July 2020, which can make the data is not actual yet not updated.

# Methodology

Indonesia’s Report on Covid-19 pandemic in 2020 use Exploratory Data Analysis, Extract Transform Load techniques, missing value treatment, feature selection, data visualization, and descriptive and inferential statistic to extract information and generate insights. The Hyphothesises stated in this paper will be tested by two test; two sample t test and ANOVA.

Tools used in this report; Python 3.8, Jupyter Lab, Scipy, Numpy, Pandas, Matplotlib, seaborn, Pywaffle, and Google Chrome.

# The Results

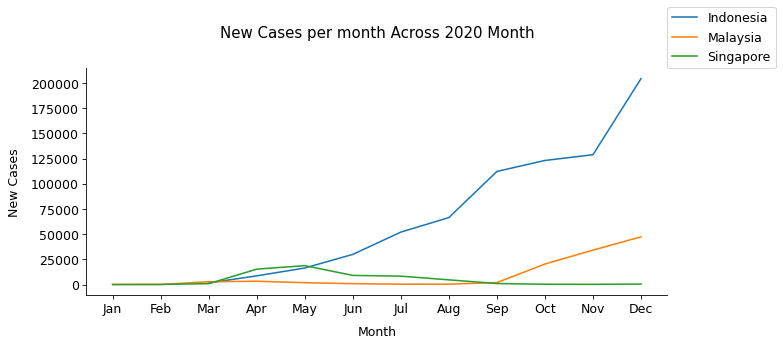
1. Indonesia’s total Covid Case



At the end of the year among Indonesia’s neighbors, **Indonesia has the highest total covid-19 case** with 743.198 cases and compared with Malaysia has 113.010 cases and Singapore has the lowest with 58.599 cases.

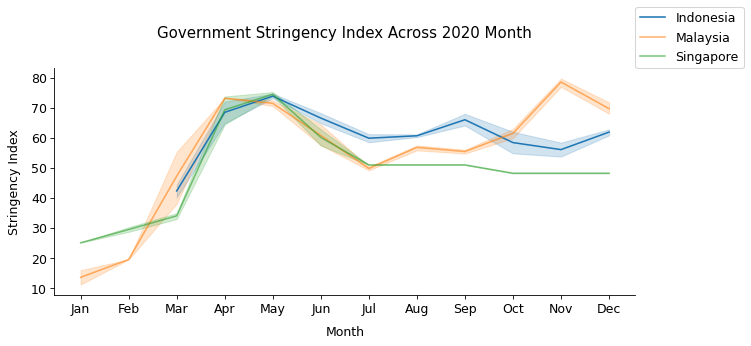
Below plot will show Indonesia, Malaysia, and Singapore virus spread behavior across 2020’s months.

1. Indonesia’s Covid spread has the most steep increases



After May 2020, Indonesia’s covid case increasing exponential significantly until at the end of 2020. Compared by Its two neighbors Malaysia and Singapore, both countries relatively can handle the virus spread and Malaysia showing significant increase on September 2020 but far below Indonesia’s cases. Indonesia’s massive spread and case increases may caused by several factor. Below graph we will view with the government stringency index point of view

1. Stringency Index of Three countries



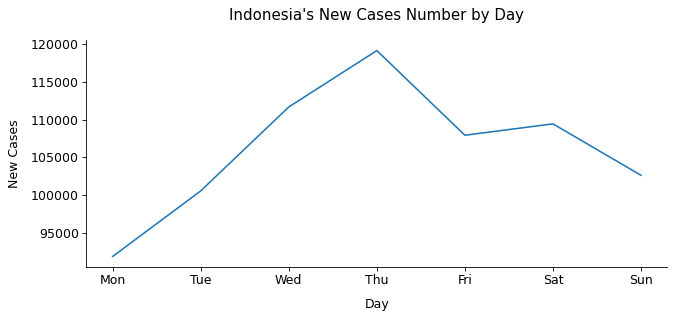
In a brief explanation, Stringency index is a quantification of governmental physical distancing policies. Data includes: the limitation of public gatherings, travel bans, movement restriction, closure of public services and non essential businesses, shelter-in-place measures and lockdowns. Each of these is weighted with a given score. The index goes from 0 (low strictness) to 100 (max stringency).

From the plot we can conclude that;

* **Corona** virus **hit Indonesia** at the first time reported **at 18 March 2020.**
* **Corona virus hit Malaysia and Singapore two months earlier** than Indonesia, despite later on Indonesia has the most Covid-19 case among them.
* Shade in the line means, dynamic changes of y value. In conclusion, **Singapore have a lot of inconsistency** of stringency index **from February to April and continuing stabe** at the rest of 2020’s month. Differ from **Indonesia,** the **stringency index went a lot of inconsistency across 2020**.
* Pearson corrleation of new cases reduction and the **Stringency Index** show 0.0081 in range of -1 to 1. Which means because the value is positive and near to zero the correlation **is almost take no effect** **and even worst** because the value **means as the stringency index grow the new cases also increases** (needs further research)

In daily activity we may have to know which day that the new infections happen and can it be related to weekend activity or weekday activity. Below plot will show the happening of new cases in day.

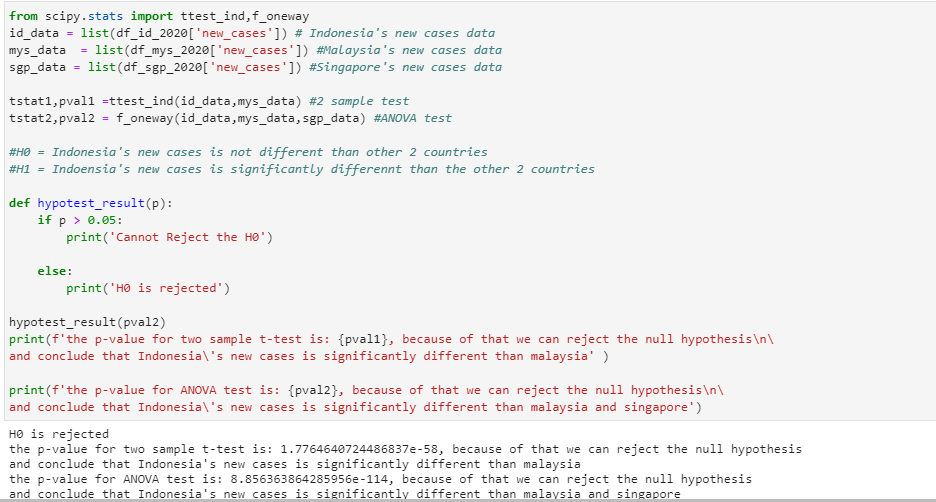
1. New cases in Indonesia per day



From the chart above we may conclude;

* Almost new cases happened in weekday

1. **Hypothesis Testing**



# Discussion

From the findings above there are several **insights** that might **need to be discussed** ;

1. Indonesia’s government policy shows more inconsistency than Malaysia and Singapore
2. Indonesia’s Stringency Index does not affecting new cases suppresses
3. It is more likely to be new cases in weekday in Indonesia

# Conclusion

1. Indonesia has the highest covid-19 case among Malaysia and Singapore with 743.198 cases
2. Indonesia has the highest new case permonth rather than Malaysia and Singapore at the peak of more than 200.000 new cases permonth.
3. Based on new cases by day, weekday is relatively higher and increasing the new cases. which means, in order to suppress the new cases increases, government and the stakeholders are suggested to improve weekdays activities's policy that lead to implementing the world health organization's standard operational procedure to operate businesses and daily life.

Possible future research development;

1. From the original data source the Indonesia’s data contain a lot of missing value that it is importan information such as icu patient and hospital patient. The author of this report highly suggested to emphasize on collecting information on these.
2. Longitude/latitude or at least the city code of Indonesia might be useful information to collect.
3. Highly suggested to add possible cause to covid-19 mortality rate or comorbid desease factor.
4. Due to the size of this dataset, ones can generate insights of ASEAN countries covid-19 data, covid -19 data per continent, research of possible factor to mortality rate, etc.

# References

<https://dataplatform.cloud.ibm.com/analytics/notebooks/v2/79dc677c-f593-47d1-a00d-b0b9c48c7191/view?access_token=f61ce5e714be3384cc054a075f8450c6695790d5d9183c01838376c28f3ea930>

<https://www.unicef.org/innovation/media/14331/file>

<https://www.who.int/docs/default-source/searo/indonesia/covid19/who-situation-report-18.pdf?sfvrsn=9fd5302_2>

<https://ourworldindata.org/covid-vaccinations>