COVID-19 Confirmed Case Correlation Analysis Based on Spearman and Kendall Correlation

Tora Fahrudin
School of Applied Sciences
Telkom University
Bandung
torafahrudin@telkomuniversity.ac.id

Dedy Rahman Wijaya
School of Applied Sciences
Telkom University
Bandung
dedyrw@tass.telkomuniversity.ac.id

Anak Agung Gde Agung
School of Applied Sciences
Telkom University
Bandung
agung@tass.telkomuniversity.ac.id

***Abstract*—COVID-19, which was declared by WHO as a pandemic in march 2020, is a severe problem which is faced by 175 countries in the world. This virus is very easily transmitted by close contact and via respiratory droplets produced when people cough or sneeze. Started from Wuhan, it spreads fast to many other countries. This paper investigates the confirmed case analysis of a COVID-19 in all countries based on Spearman and Kendall correlation and grouping those countries which have the same level correlation. There are 134 countries for the shortest period and 16 for the longest period have been evaluated. In the shortest period, 39 and 27 countries have high correlation value based on Spearman and Kendall, respectively. In the longest period, 11 and 6 countries have high correlation value based on Spearman and Kendall, respectively. Those groups indicate that the strength of association of confirmed case between them are similar. The darker colour shows there is a high correlation value among those countries. So, by using those groups, the decision-maker can analyze the characteristic of those countries and make the decision better.**

Keywords—COVID-19, spread, Spearman, Kendall, correlation

I. INTRODUCTION

By the end of 2019, infection of novel coronavirus to humans firstly confirmed in Wuhan city, Hubei Province, China [1]. Most of the patients visited the Huanan seafood market [2] and wild animal market [1] last month in Wuhan. Although COVID-19 firstly reported a link between animal-to-human transmission, studies have increasingly demonstrated human-to-human transmission of COVID-19 through droplets or direct contact [3].

The transmission of COVID-19 not only reported transmitted on local transmission in China but also spread outside of China, such as Italy, Spain, Germany, and the United States [4]. Figure 1 shows the confirmed case of COVID-19 in the world on March 28th, 2020. Nine countries have COVID-19 confirmed cases more than 10,000 cases, such as the US, Italy, China, Spain, Germany, France, Iran, the United Kingdom, and Switzerland.

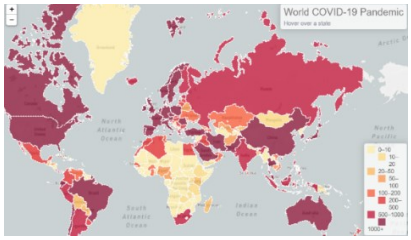


Fig. 1. World COVID-19 Confirmed Case [5]

Figure 2 shows the confirmed case trend of each country. From that figure, we can see that the line pattern of the confirmed case for most countries is more than the "doubling every ten days" line. Only Sri Lanka, Kuwait, and Singapore which have lower than "doubling every ten days" line.

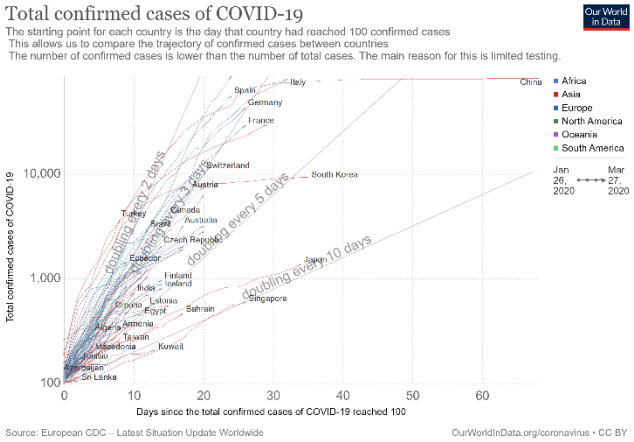


Fig. 2. World COVID-19 Confirmed Case Trend [6]

So, according to the background, we need to investigate COVID-19 confirmed case correlation for all countries and see which country has the same pattern. In this research, we present the correlation analysis of confirmed cases for 175 countries around the world.

The contributions of this paper can be summarized as follows:

- Detailed analysis of confirmed case correlation by using the Spearman and Kendall correlation method.
- Determined the group of each country based on confirmed case correlation analysis by using Spearman and Kendall correlation method.

The rest of this paper is organized as follows: section 2 is related works, including COVID-19 and correlation analysis. Section 3 explains our methodology and experimental setting. Section 4 discusses our results. Finally, section 5 concludes this work.

II. RELATED WORKS

This section contains related works which can be split into two subsections. First, explanation about COVID-19. Second, explanation about Correlation analysis.

A. COVID-19

Coronavirus disease 2019 (COVID-19) is firstly identified in Wuhan city, Hubei Province, China. First, confirm that you have the correct template for your paper size. COVID-19 is an infectious disease caused by severe acute respiratory syndrome coronavirus 2. COVID-19 patients mostly had respiratory signs and symptoms such as fever, flanked by cough, and primarily dry [7].

COVID-19 primarily spread from person-to-person through close contact by respiratory droplets and transmitted with subsequent contact with eyes, nose, or mouth [4].

B. Correlation analysis

Correlation is a measure of the association between two variables. The value of the correlation can be positive and negative (ranging from -1 to +1). A positive value means if the value of one variable increase, the other variable also increase. Otherwise, a negative value means if the value of one variable increases, the other variable decrease. There are two approaches: parametric and non-parametric rank correlation. Parametric tests need the assumption that the data complies with some normal, linear, and homoscedastic distribution [8], while Non-Parametric no need to comply with that assumption [9].

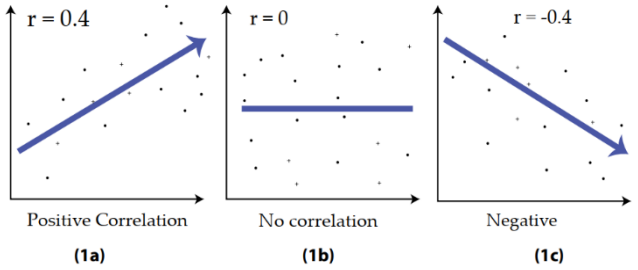


Fig. 3. Three Types of Correlation Value [10]

Pearson correlation is one of the examples of a Parametric rank correlation (r). Unlike Pearson's correlation, Spearman's and Kendall, as the examples of a Non-parametric rank-order correlation does not make any assumptions about the distribution of the data [8]. In this paper, we use a Non-parametric approach using the Spearman and Kendall correlation method. Benefit of using Non-parametric: First, nonparametric correlations can work on incomplete data (where only ordinal information of the data is available). Second, Spearman's ranks correlation (SR) and Kendalls' tau correlation (KT) equal 1 when Y is a monotonically increasing function of X.Third, Spearman and Kendall are more robust against outliers than Pearson Correlation [11]. Kendall is more robust than Spearman, which indicated on Gross Error Sensitivity (GES) value is lower than Spearman [12].

Spearman rank-order correlation coefficient (denoted ρ_s) is a rank-based version of the Pearson correlation coefficient [13]. The formulas of Spearman can be seen in equation (1). Meanwhile, Kendall's formula (denoted τ) can be expressed in equation (2).

$$\rho_s = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(\left(R(x_i)-\overline{R(x)}\right)\left(R(y_i)-\overline{R(y)}\right)\right)}{\sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\left(R(x_i)-\overline{R(x)}\right)^2\right)\left(\frac{1}{n}\sum_{i=1}^{n}\left(R(y_i)-\overline{R(y)}\right)^2\right)}} \quad (1)$$

$$\tau = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} sign(x_i-x_j)\, sign(y_i-y_j)}{n(n-1)} \quad (2)$$

III. Materials and Methods

This section contains materials and methods. Data set and correlation analysis library will be explained in the material subsection. The experimental design step will be described in the methods subsection.

A. Materials

Data set in this research is provided by the John Hopkins Center for Systems Science and Engineering (JHU CSSE) [14], which is updated periodically. That data is combined from WHO, CDC, ECDC, NHC, DXY, 1point3acres, Worldometers.info, BNO, state, and national government health departments, and local media reports data [15]. In this research, we focus on the confirmed case data set for all affected countries. Confirmed cases include presumptive positive cases of COVID-19. The confirmed case is served by time-series data, which is derived from the daily case reports and regular updates by two times.

The correlation function in the pandas 1.0.2 dataframe in python is used to analyze the data set. Python 3.7.6 was used to analyze the data.

B. Methods

The methodology design of this research can be seen in Figure 4.

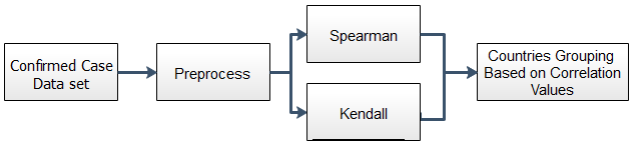


Fig. 4. Methodology Design

The confirmed case data set was needed to be preprocessed first. That is because, the data set is in time series form, started in Jan 22th 2020 until March 27th 2020, which means the emergence of confirmed cases for each country is different. So we need to use the first day of confirmed cases for each country as the starting point, not based on the date. Two scenarios will be conducted, the shortest period and the longest period. We can see the data before preprocessed and after preprocessed in Table 1 and Table 2. Figure 5 shows the number of countries which reported the confirmed case (y-axis) and the number of accumulated days confirmed for each country since it was first discovered in that country (x-axis).

TABLE I. CONFIRMED CASE DATA FOR EACH COUNTRY

Dates	Afghanistan	..	Cambodia	Canada	⋯
…	…	…	…	…	…
26/1/20	0	...	0	1	⋯
27/1/20	0	...	1	1	⋯
…	…	…	…	…	…

TABLE II. STARTED CONFIRMED CASE DATA FOR EACH COUNTRY

Day	Afghanistan	..	Cambodia	Canada	⋯
1	1	…	1	1	…
2	1	...	1	1	⋯

Day	Afghanistan	..	Cambodia	Canada	…
…	…	…	…	…	…

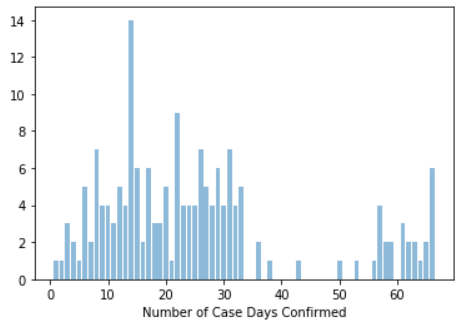


Fig. 5. Number of Case Days Confirmed all Countries

From Figure 5, we choose 14 days as the shortest day period because it can capture most countries and 60 days as the longest period. Section IV will explain the result of the shortest period scenarios and the longest period. One hundred thirty-four countries will be evaluated in the shortest period, and six teen countries will be assessed for the longest period scenarios.

IV. Experimental Results

Spearman and Kendall's correlation was conducted in confirmed cases number between one country to another country after preprocess has been done. This correlation is conducted by using pairwise correlation, which means it compares the confirmed case of one country to another country one by one.

A. *Shortest Period Results*

Figure 6 shows the pairwise correlation results of some countries using Spearman and Kendall in Figure 7. The average Spearman pairwise correlation of all country is 0.908, which means it has strong positive correlation, but may still sensitive with outlier than Kendall. Meanwhile, an average of Kendall pairwise correlation of all country is 0.845; it means that lower sensitive than Spearman.

	Albania	Algeria	Antigua ar	Argentina	Armenia	Australia	Austria	Azerbaijar	Bahrain	Banglades	Belarus
Albania	1	0.95848	0.79282	0.99338	0.71867	0.99117	0.9978	0.97778	0.9967	0.93644	0.93644
Algeria	0.95848	1	0.82717	0.95678	0.73066	0.94851	0.96059	0.95681	0.9559	0.97211	0.89988
Antigua and Barbud	0.79282	0.82717	1	0.7981	0.85246	0.77612	0.79457	0.76832	0.78989	0.84664	0.75043
Argentina	0.99338	0.95678	0.7981	1	0.72345	0.97991	0.99114	0.97184	0.98669	0.94032	0.94268
Armenia	0.71867	0.73066	0.85246	0.72345	1	0.72198	0.72025	0.69434	0.71798	0.76745	0.76418
Australia	0.99117	0.94851	0.77612	0.97991	0.72198	1	0.98669	0.98649	0.99112	0.91637	0.92821
Austria	0.9978	0.96059	0.79457	0.99114	0.72025	0.98669	1	0.97993	0.99669	0.93851	0.93851
Azerbaijan	0.97778	0.95681	0.76832	0.97184	0.69434	0.98649	0.97993	1	0.98101	0.92172	0.92172
Bahrain	0.9967	0.9559	0.78989	0.98669	0.71798	0.99112	0.99669	0.98101	1	0.93483	0.93601
Bangladesh	0.93644	0.97211	0.84664	0.94032	0.76745	0.91637	0.93851	0.92172	0.93483	1	0.86216
Belarus	0.93644	0.89988	0.75043	0.94268	0.76418	0.92821	0.93851	0.92172	0.93601	0.86216	1

Fig. 6. Spearman Correlation Results Samples on the shortest period

	Afghanist	Albania	Algeria	Antigua a	Argentina	Armenia	Australia	Austria	Azerbaija	Bahrain	Bangladesh	Belarus
Afghanistan	1	0.37796	0.41914	0.54984	0.3888	0.60093	0.36314	0.38219	0.37928	0.35479	0.430945804	0.44721
Albania	0.37796	1	0.90177	0.68741	0.97214	0.62897	0.96077	0.98895	0.91987	0.98338	0.877058019	0.84515
Algeria	0.41914	0.90177	1	0.76229	0.91508	0.67811	0.87517	0.91184	0.90084	0.89223	0.944809024	0.85071
Antigua and Barbuda	0.54984	0.68741	0.76229	1	0.70711	0.83874	0.6822	0.69509	0.69515	0.68277	0.783763813	0.71877
Argentina	0.3888	0.97214	0.91508	0.70711	1	0.647	0.94124	0.97157	0.92165	0.95409	0.889305227	0.86938
Armenia	0.60093	0.62897	0.67811	0.83874	0.647	1	0.63647	0.636	0.62678	0.62184	0.717137166	0.72354
Australia	0.36314	0.96077	0.87517	0.6822	0.94124	0.63647	1	0.94837	0.95743	0.96538	0.847665863	0.8526
Austria	0.38219	0.98895	0.91184	0.69509	0.97157	0.636	0.94837	1	0.93014	0.98307	0.886857854	0.8546
Azerbaijan	0.37928	0.91987	0.90084	0.69515	0.92165	0.62678	0.95743	0.93014	1	0.93541	0.85811633	0.86224
Bahrain	0.35479	0.98338	0.89223	0.68277	0.95409	0.62184	0.96538	0.98307	0.93541	1	0.866400225	0.84622
Bangladesh	0.43095	0.87706	0.94481	0.78376	0.88931	0.71714	0.84767	0.88686	0.85812	0.8664	1	0.81537
Belarus	0.44721	0.84515	0.85071	0.71877	0.86938	0.72354	0.8526	0.8546	0.86224	0.84622	0.815374248	1

Fig. 7. Kendalls Correlation Results Samples on the shortest period

Based on the heat map color in Figure 6 and Figure 7, we make a group of all countries which have a strong correlation (darker color means stronger correlation value) which can be seen in Figure 8 and Figure 9. Table 3 and Table 4 show a list of countries that have a high correlation value for Spearman and Kendall, respectively. There are thirty-six countries for Spearman and twenty-seven for Kendall.

TABLE III. Countries Group Based on Spearman Correlation

Correlation value	Number of Countries	Countries
1-0.98	39	Albania, Argentina, Austria, Bahrain, Bolivia, Bosnia & Herzegovina, Brunei, Bulgaria, China, Costa Rica, Croatia, Cyprus, Czechia, Diamond Princess, Ghana, Guatemala, Hungary, Iceland, Iran, Iraq, Kazakhstan, Moldova, Netherlands, Norway, Panama, Peru, Poland, Portugal, Rwanda, Singapore, Slovakia, Slovenia, Switzerland, Trinidad and Tobago, Turkey, Uruguay, Venezuela, West Bank and Gaza

	Albania	Argentina	Austria	Bahrain	Bolivia	Bosnia and Herzegovina	Brunei	Bulgaria	China	Costa Rica
Albania	1	0.99338	0.9978	0.9967	0.99559	0.994490316	0.9978	0.99449	1	0.997799777
Argentina	0.99338	1	0.99114	0.98669	0.99222	0.985540098	0.99114	0.99666	0.99338	0.993357883
Austria	0.9978	0.99114	1	0.99669	0.99447	0.996683241	1	0.99115	0.9978	0.996688742
Bahrain	0.9967	0.98669	0.99669	1	0.99336	0.993350557	0.99669	0.98781	0.9967	0.993370772
Bolivia	0.99559	0.99222	0.99447	0.99336	1	0.994451223	0.99447	0.99445	0.99559	0.997790053
Bosnia and Herzegovina	0.99449	0.98554	0.99668	0.99335	0.99445	1	0.99668	0.98889	0.99449	0.996683241
Brunei	0.9978	0.99114	1	0.99669	0.99447	0.996683241	1	0.99115	0.9978	0.996688742
Bulgaria	0.99449	0.99666	0.99115	0.98781	0.99445	0.988888889	0.99115	1	0.99449	0.996683241
China	1	0.99338	0.9978	0.9967	0.99559	0.994490316	0.9978	0.99449	1	0.997799777
Costa Rica	0.9978	0.99336	0.99669	0.99337	0.99779	0.996683241	0.99669	0.99668	0.9978	1

Fig. 8. Highest Spearman Correlation Result on Shortest Period (0.98 – 1)

TABLE IV. Countries Group Based on Kendall Correlation

Correlation value	Number of Countries	Countries
1-0.97	27	Albania, Austria, Bosnia & Herzegovina, Brunei, China, Costa Rica, Czechia, Denmark, Diamond Princess, Ghana, Guatemala, Iceland, Iran, Iraq, Kazakhstan, Netherlands, Norway, Panama, Peru, Poland, Portugal, Rwanda, Slovakia, Switzerland, Trinidad & Tobago, Turkey, Uruguay

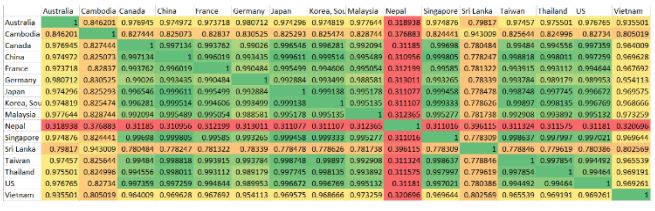
	Albania	Austria	Bosnia ar	Brunei	China	Costa Ric	Czechia	Denmark	Diamond	Ghana
Albania	1	0.98895	0.97778	0.98895	1	0.98895	0.98895	0.97778	0.98338	0.98895
Austria	0.98895	1	0.9887	1	0.98895	0.98876	0.98876	0.97734	0.98307	0.97753
Bosnia and Herzegovir	0.97778	0.9887	1	0.9887	0.97778	0.9887	0.97734	0.97701	0.98287	0.97734
Brunei	0.98895	1	0.9887	1	0.98895	0.98876	0.98876	0.97734	0.98307	0.97753
China	1	0.98895	0.97778	0.98895	1	0.98895	0.98895	0.97778	0.98338	0.98895
Costa Rica	0.98895	0.98876	0.9887	0.98876	0.98895	1	0.98876	0.9887	0.99437	0.98876
Czechia	0.98895	0.98876	0.97734	0.98876	0.98895	0.98876	1	0.97734	0.98307	0.97753
Denmark	0.97778	0.97734	0.97701	0.97734	0.97778	0.9887	0.97734	1	0.98287	0.97734
Diamond Princess	0.98338	0.98307	0.98287	0.98307	0.98338	0.99437	0.98307	0.98287	1	0.98307
Ghana	0.98895	0.97753	0.97734	0.97753	0.98895	0.98876	0.97753	0.97734	0.98307	1

Fig. 9. Highest Kendall Correlation Result on Shortest Period (0.97 – 1)

B. *Longest Period Results*

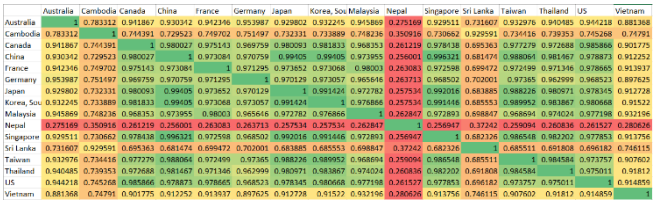
The average Spearman pairwise correlation of all countries in the longest period scenarios is 0.873, while the Kendall pairwise correlation of all countries is 0.834. The difference among Spearman and Kendall is smaller than in the shortest period scenarios, which means the confirmed case's number is relatively the same for COVID-19-infected countries over a long period. The spread of virus is depend on

how effective the government battle the virus, such as lock down, social distancing, punishment and etc.



	Australia	Cambodia	Canada	China	France	Germany	Japan	Korea, Sou	Malaysia	Nepal	Singapore	Sri Lanka	Taiwan	Thailand	US	Vietnam
Australia	1	0.846201	0.976945	0.97972	0.973718	0.980712	0.971296	0.974819	0.977644	0.318938	0.974876	0.79817	0.97457	0.975501	0.976765	0.935501
Cambodia	0.846201	1	0.827444	0.825073	0.82837	0.830525	0.825293	0.825474	0.828744	0.376883	0.824441	0.943009	0.825644	0.824996	0.82734	0.805019
Canada	0.976945	0.827444	1	0.997134	0.993762	0.99026	0.996546	0.996281	0.992094	0.31185	0.99698	0.780484	0.99484	0.994556	0.997359	0.964009
China	0.97972	0.825073	0.997134	1	0.996019	0.993435	0.999611	0.999514	0.995489	0.310956	0.999805	0.778247	0.998818	0.998011	0.997259	0.969628
France	0.973718	0.82837	0.993762	0.996019	1	0.990484	0.995499	0.994606	0.995054	0.312196	0.99585	0.781322	0.993915	0.993112	0.994644	0.967697
Germany	0.980712	0.830525	0.99026	0.993435	0.990484	1	0.992884	0.993499	0.995135	0.313011	0.993765	0.78339	0.993784	0.989179	0.989953	0.954113
Japan	0.974296	0.825293	0.996546	0.999611	0.995499	0.992884	1	0.999138	0.995178	0.311077	0.999458	0.778478	0.998748	0.997745	0.996672	0.969575
Korea, Sou	0.974819	0.825474	0.996281	0.999514	0.994606	0.993499	0.999138	1	0.995135	0.311107	0.999333	0.778626	0.99897	0.998135	0.996769	0.968666
Malaysia	0.977644	0.828744	0.992094	0.995489	0.995054	0.988581	0.995178	0.995135	1	0.312365	0.995277	0.781738	0.992908	0.993892	0.995132	0.973259
Nepal	0.318938	0.376883	0.31185	0.310956	0.312199	0.313011	0.311077	0.311107	0.312365	1	0.311019	0.396115	0.311324	0.311575	0.31181	0.320696
Singapore	0.974876	0.824441	0.99698	0.999805	0.99585	0.993765	0.999458	0.999333	0.995277	0.311019	1	0.778309	0.998637	0.997997	0.997021	0.969644
Sri Lanka	0.79817	0.943009	0.780484	0.778247	0.781322	0.78339	0.778478	0.778626	0.781738	0.396115	0.778309	1	0.778846	0.779619	0.780186	0.802569
Taiwan	0.97457	0.825644	0.99484	0.998818	0.993915	0.993784	0.998748	0.99897	0.992908	0.311324	0.998637	0.778846	1	0.997854	0.994492	0.965539
Thailand	0.975501	0.824996	0.994556	0.998011	0.993112	0.989179	0.997745	0.998135	0.993892	0.311575	0.997997	0.779619	0.997854	1	0.99464	0.969191
US	0.976765	0.82734	0.997359	0.997259	0.994644	0.989953	0.996672	0.996769	0.995132	0.31181	0.997021	0.780186	0.994492	0.99464	1	0.969261
Vietnam	0.935501	0.805019	0.964009	0.969628	0.967697	0.954113	0.969575	0.968666	0.973259	0.320696	0.969644	0.802569	0.965539	0.969191	0.969261	1

Fig. 10. Spearman Correlation Results on the longest period



	Australia	Cambodia	Canada	China	France	Germany	Japan	Korea, Sou	Malaysia	Nepal	Singapore	Sri Lanka	Taiwan	Thailand	US	Vietnam
Australia	1	0.783312	0.941867	0.930352	0.942346	0.953987	0.929802	0.932245	0.945869	0.275169	0.929511	0.731607	0.932976	0.940485	0.944218	0.881368
Cambodia	0.783312	1	0.744391	0.729523	0.749702	0.751697	0.732331	0.733889	0.748236	0.350916	0.730682	0.929591	0.734516	0.728353	0.745268	0.74791
Canada	0.941867	0.744391	1	0.980027	0.975143	0.969759	0.980093	0.981833	0.968353	0.261219	0.978438	0.695363	0.977279	0.972688	0.985866	0.901775
China	0.930352	0.729523	0.980027	1	0.973084	0.970759	0.99405	0.973955	0.973855	0.256001	0.996321	0.681474	0.988064	0.981467	0.978873	0.912252
France	0.942346	0.749702	0.975143	0.973084	1	0.971295	0.973652	0.973068	0.98003	0.263083	0.972598	0.699472	0.972499	0.971346	0.978665	0.913937
Germany	0.953987	0.751697	0.969759	0.970759	0.971295	1	0.970129	0.973057	0.965646	0.263713	0.968502	0.702001	0.97365	0.962999	0.968523	0.897625
Japan	0.929802	0.732331	0.980093	0.99405	0.973652	0.970129	1	0.991424	0.972782	0.257534	0.992018	0.683885	0.988226	0.980971	0.978345	0.912728
Korea, Sou	0.932245	0.733889	0.981833	0.99405	0.973068	0.973057	0.991424	1	0.976866	0.257534	0.991446	0.685553	0.989952	0.983867	0.980668	0.91522
Malaysia	0.945869	0.748236	0.968353	0.973955	0.98003	0.965646	0.972782	0.976866	1	0.262847	0.972893	0.698847	0.968694	0.974024	0.977198	0.932196
Nepal	0.275169	0.350916	0.261219	0.256001	0.263083	0.263713	0.257534	0.257534	0.262847	1	0.256947	0.37242	0.259094	0.260836	0.261527	0.280626
Singapore	0.929511	0.730682	0.978438	0.996321	0.972598	0.968502	0.992018	0.991446	0.972893	0.256947	1	0.682326	0.986548	0.982202	0.977853	0.913756
Sri Lanka	0.731607	0.929591	0.695363	0.681474	0.699472	0.702001	0.683885	0.685553	0.698847	0.37242	0.682326	1	0.685511	0.691808	0.696182	0.746115
Taiwan	0.932976	0.734516	0.977279	0.988064	0.972499	0.97365	0.988226	0.989952	0.968694	0.259094	0.986548	0.685511	1	0.984584	0.973757	0.907602
Thailand	0.940485	0.728353	0.972688	0.981467	0.971346	0.962999	0.980971	0.983867	0.974024	0.260836	0.982202	0.691808	0.984584	1	0.975011	0.91812
US	0.944218	0.745268	0.985866	0.978873	0.978665	0.968523	0.978345	0.980668	0.977198	0.261527	0.977853	0.696182	0.973757	0.975011	1	0.914859
Vietnam	0.881368	0.74791	0.901775	0.912252	0.913937	0.897625	0.912728	0.91522	0.932196	0.280626	0.913756	0.746115	0.907602	0.91812	0.914859	1

Fig. 11. Kendall Correlation Results on the longest period

	Canada	China	France	Germany	Japan	Korea, Sou	Malaysia	Singapore	Taiwan	Thailand	US
Canada	1	0.997134	0.993762	0.99026	0.996546	0.996281	0.992094	0.99698	0.99484	0.994556	0.997359
China	0.997134	1	0.996019	0.993435	0.999611	0.999514	0.995489	0.999805	0.998818	0.998011	0.997259
France	0.993762	0.996019	1	0.990484	0.995499	0.994606	0.995054	0.99585	0.993915	0.993112	0.994644
Germany	0.99026	0.993435	0.990484	1	0.992884	0.993499	0.988581	0.993265	0.993784	0.989179	0.989953
Japan	0.996546	0.999611	0.995499	0.992884	1	0.999138	0.995178	0.999458	0.998748	0.997745	0.996672
Korea, Sou	0.996281	0.999514	0.994606	0.993499	0.999138	1	0.995135	0.999333	0.99897	0.998135	0.996769
Malaysia	0.992094	0.995489	0.995054	0.988581	0.995178	0.995135	1	0.995277	0.992908	0.993892	0.995132
Singapore	0.99698	0.999805	0.99585	0.993265	0.999458	0.999333	0.995277	1	0.998637	0.997997	0.997021
Taiwan	0.99484	0.998818	0.993915	0.993784	0.998748	0.99897	0.992908	0.998637	1	0.997854	0.994492
Thailand	0.994556	0.998011	0.993112	0.989179	0.997745	0.998135	0.993892	0.997997	0.997854	1	0.99464
US	0.997359	0.997259	0.994644	0.989953	0.996672	0.996769	0.995132	0.997021	0.994492	0.99464	1

Fig. 12. Heat Map of Highest Spearman Correlation Result on the longest period (0.98 – 1)

	China	Japan	Korea, South	Singapore	Taiwan	Thailand
China	1	0.99405	0.994050096	0.996321	0.988064	0.981467
Japan	0.99405	1	0.991423671	0.992016	0.988226	0.980971
Korea, South	0.99405	0.991424	1	0.991446	0.989952	0.983867
Singapore	0.996321	0.992016	0.991445821	1	0.986548	0.982202
Taiwan	0.988064	0.988226	0.989951907	0.986548	1	0.984584
Thailand	0.981467	0.980971	0.983866747	0.982202	0.984584	1

Fig. 13. Heat Map of Highest Kendall Correlation Result on the longest period (0.98 – 1)

Figure 10 and Figure 11 show that among all countries with have more than 60 days infected by COVID-19, only Nepal did not show the rapid increase of confirmed cases. Besides Nepal, Sri Lanka, Cambodia, Vietnam, and Australia, we make a group of all countries which have a strong correlation based on the heat map color, which can be seen in Figure 12 and Figure 13.

Eleven countries have been chosen in Spearman based on colour; meanwhile, six-country have been chosen in Kendall. The list countries for Kendall are China, Japan, South Korea, Singapore, Taiwan, and Thailand. Besides those six countries, there are five additional countries for Spearman such as Canada, France, Germany, Malaysia and US.

The groups generated based on the colour of the heat map showing the pattern of confirmed case correlation between them are similar. That is shown by the results of strong correlations among countries which are shown in a darker colour.

V. Conclusion

COVID-19 is spreading so fast and very easily transmitted by close contact and via respiratory droplets. 175 country has been infected by this virus rapidly. So, this research is design to investigate the confirmed case correlation analysis of all affected countries and try to group them by considering the correlation value by using Spearman and Kendall.

Spearman and Kendall's correlation was successfully conducted in COVID-19 confirmed cases using the John Hopkins Center for Systems Science and Engineering data set. The experiment has been done and evaluated in two scenarios data set: 14 days confirmed case data and 60 days confirmed case data. There are 134 countries have been assessed in the shortest period, and 16 countries for the longest period.

In the shortest path, there are thirty-nine countries collected into one group because they have a high correlation value by using Spearman and twenty-seven countries by using Kendall correlation. The heatmap also showed a darker colour among those countries. Meanwhile, in the longest path, there are eleven countries obtained by using Spearman and six countries obtained by using Kendall.

From all countries which have been observed, Nepal, Sri Lanka, Cambodia, Vietnam, and Australia show the lower correlation values, which mean that those countries can keep the number of confirmed case not to be grown rapidly.

Lastly, it needs to be considered that different countries have different situation and strategies to fight COVID-19. So, it makes the situation is complicated and difficult to compare them perfectly.

References

[1] H. Lu, C. W. Stratton, and Y. W. Tang, "Outbreak of pneumonia of unknown etiology in Wuhan, China: The mystery and the miracle," J. Med. Virol., vol. 92, no. 4, pp. 401–402, 2020, doi: 10.1002/jmv.25678.

[2] R. Lu et al., "Genomic Characterisation and Epidemiology of 2019 Novel Coronavirus: Implications for Virus Origins and Receptor Binding," Lancet, vol. 395, pp. 565–574, 2020, doi: 10.1016/S0140-6736(20)30251-8.

[3] C. C. Lai, T. P. Shih, W. C. Ko, H. J. Tang, and P. R. Hsueh, "Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges," Int. J. Antimicrob. Agents, vol. 55, no. 3, p. 105924, 2020, doi: 10.1016/j.ijantimicag.2020.105924.

[4] C. M. Amado, C. J. Minahk, E. Cilli, G. Oliveira, and F. G. Dupuy, "Coronavirus Disease (COVID-19): A primer for emergency physicians," Am. J. Emerg. Med., p. 183135, 2020, doi: 10.1016/j.bbamem.2019.183135.

[5] D. R. Wijaya, A. A. G. Agung, T. Fahrudin, and Suryatiningsih, "COVID-19 Virus," 2020. [Online]. Available: https://ais-rg.com/corona/peta. [Accessed: 28-Mar-2020].

[6] N. Routley, "Infection Trajectory: See Which Countries are Flattening Their COVID-19 Curve," Visual Capitalist, 2020. [Online]. Available: https://www.visualcapitalist.com/infection-trajectory-flattening-the-covid19-curve/.

[7] T. Lupia, S. Scabini, S. M. Pinna, G. Di Perri, F. G. De Rosa, and S. Corcione, "2019-novel coronavirus outbreak: A new challenge.," J. Glob. Antimicrob. Resist., vol. 21, pp. 22–27, 2020, doi: 10.1016/j.jgar.2020.02.021.

[8] K. F. Weaver, V. C. Morales, S. L. Dunn, K. Godde, and P. Weaver, An Introduction to Statistical Analysis in Research: With Applications in the Biological and Life Sciences. Wiley, 2017.

[9] Justin, "Selecting Between Parametric and Non-Parametric Analyses," Statistics Solutions, 2020. [Online]. Available:

https://www.statisticssolutions.com/selecting-between-parametric-and-non-parametric-analyses/. [Accessed: 28-Mar-2020].

[10] N. J. Gogtay and U. M. Thatte, "Principles of correlation analysis," J. Assoc. Physicians India, vol. 65, no. March, pp. 78–81, 2017.

[11] Y. Kim, T. H. Kim, and T. Ergün, "The instability of the Pearson correlation coefficient in the presence of coincidental outliers," Finance. Res. Lett., vol. 13, pp. 243–257, 2015, doi: 10.1016/j.frl.2014.12.005.

[12] C. Croux and C. Dehon, "Influence functions of the Spearman and Kendall correlation measures," Stat. Methods Appl., vol. 19, no. 4, pp. 497–515, 2010, doi: 10.1007/s10260-010-0142-z.

[13] N. S. Chok, "Pearson Versus Spearman And Kendall Correlation Coefficients For Continuous Data," 2010.

[14] Johns Hopkins University, "2019 Novel Coronavirus COVID-19 (2019-nCoV) Data Repository by Johns Hopkins CSSE," Github, 2020. [Online]. Available: https://github.com/CSSEGISandData/COVID-19. [Accessed: 28-Mar-2020].

[15] Johns Hopkins University, "Coronavirus COVID-19 Global Cases by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University," 2020. [Online]. Available: https://coronavirus.jhu.edu/map.html. [Accessed: 28-Mar-2020].