MONITORING THE GLOBAL COVID-19 IMPACT ON TOURISM: The COVID19tourism index

Yang Yang, Ph.D.*

Associate Professor

Department of Tourism and Hospitality Management

Temple University, Philadelphia, Pennsylvania, 19122, USA

Benjamin Altschuler, Ph.D.

Assistant Professor

Department of Tourism and Hospitality Management

Temple University, Philadelphia, Pennsylvania, 19122, USA

Zhengkang Liang

Ph.D. student

Department of Statistical Science

Temple University, Philadelphia, Pennsylvania, 19122, USA

Xiang (Robert) Li, Ph.D.

Professor and Washburn Senior Research Fellow

Department of Tourism and Hospitality Management

Temple University, Philadelphia, Pennsylvania, 19122, USA

* indicates the corresponding author

Please cite as:

Yang, Y., Altschuler, B., Liang, Z., and Li, X. (accepted). Monitoring the global COVID-19 impact on tourism. *Annals of Tourism Research*. doi: 10.1016/j.annals.2020.103120

Keywords: COVID-19; impact analysis; time-series clustering; pandemic curve; tourism recovery

Author Bio

Yang Yang, Ph.D. is an associate professor at the School of Sport, Tourism and Hospitality Management (STHM), Temple University. His research interest includes tourism and hospitality analytics.

Benjamin Altschuler, Ph.D. is an assistant professor at STHM, Temple University. His research interest includes sustainable tourism.

Zhengkang Liang is a Ph.D. student of Department of Statistical Science, Temple University. His research interest includes Bayesian statistics.

Xiang (Robert) Li, Ph.D. is a professor at STHM, Temple University. His research interest includes tourism marketing.

MONITORING THE GLOBAL COVID-19 IMPACT ON TOURISM: The COVID19tourism index

Introduction

The COVID-19 virus is highly transmittable, wreaking havoc on the world's economy with the travel and tourism industry one of the most ravaged sectors. Given the impacts of COVID-19 and the intricacies of the global tourism economy, it is critical for policymakers to monitor recovery across numerous destinations. In this research note, we describe the development and calibration of an analytical tool named the "COVID19tourism index" to monitor the pandemic's tourism effects. As a powerful numerical and visual tool, the index provides important information related to the potential travel and tourism recovery at the global, regional, and country levels. Compared to a benchmark of "normal" levels (equal to 100 in this case), the COVID19tourism index offers insight into the tourism industry's recovery process along with the pandemic's impacts on numerous aspects of tourism.

Index development

The COVID19tourism index consists of an overall index and five sub-indices, and the overall index is constructed with three major sub-indices. The aviation and hotel sub-indices reflect performance in two major sectors of the tourism industry. Based on the timeliness of daily data, we used air flight departure data to construct the aviation sub-index and took revenue per available room (RevPAR) as a vital measure of hotel performance to construct the hotel sub-index. By default, a tourism recovery index must reflect pandemic conditions. Regardless of how strongly the tourism industry is performing, if the pandemic is not well controlled, then industry recovery will be incomplete. As such, the third sub-index—the pandemic sub-index—considers the number of confirmed COVID-19 cases relative to the inbound tourist population. The fourth sub-index, the interest sub-index, examines latent travel demand to a given country based on global online search trends from Google. The last index, the mobility sub-index, indicates the level of tourist/traveler mobility within a country. The data were collected from Google users opting-in to location history service and then aggregated based on different categories of places (https://www.google.com/covid19/mobility/). The raw data and data sources for each sub-index are presented in Table 1.

(Please insert Table 1 about here)

The Joint Research Centre-European Commission (2008) has recommended checking the statistical coherence of diverse data sources. Statistical definitions of all data in this study remained consistent throughout the research period and across countries. Raw data were 7-day moving-averaged to iron out the effect of the day of the week and to compensate for missing values. We defined baseline hotel and aviation performance outside the COVID-19 pandemic as:

$$y_{ijt}^{*} = y_{ij(t-364)} * W_{ij} \tag{1}$$

where y indicates the raw measure for each sub-index, i indicates the country, j indicates the type of sub-index, and t indicates the date. The "normal"-level value $y*$ incorporates an adjustment factor W_{ij}, defined as the sum of daily values from January 1 to 28, 2020 over those from January 1 to 28, 2019. With this method, we assumed that the first 28 days of 2020 in each country would reflect a normal level of tourism activity, based upon which we calculated the year-to-year change over this period. We could not access daily historical values of raw data for the interest and mobility sub-indices; therefore, we referred to Google's COVID-19 Community

Mobility Reports and took the 5 weeks of January 3–February 6, 2020 as a baseline. The sub-indices of aviation, hotel, interest, and mobility were then calculated as

$$I_{ijt} = \min\left(100\,, 100 * y_{ijt}/y^{*}_{ijt}\right) \tag{2}$$

Therefore, these four sub-indices reflect the actual level y compared to the normal level y^*. Also, the sub-index was truncated at a value of 100 because a value beyond that reflects a normal level in our research context. Lastly, the pandemic index was defined as

$$I_{ijt} = \min\left(100\,, 100/\sqrt[4]{z_{it}}\right) \tag{3}$$

where z_{it} indicates the number of daily confirmed COVID-19 cases per 1 million inbound tourists.

The overall index was calculated as the geometric mean of the aviation, hotel, and pandemic sub-indices. These sub-indices were less substitutable to each other to reflect the recovery level, hence geometric aggregation was deemed superior to linear aggregation (i.e., arithmetic mean) (Joint Research Centre-European Commission, 2008). The interest and mobility sub-indices were excluded due to data unavailability in countries with a low market share of Google products. The overall index represents the general recovery potential at global, regional, and country levels. The closer this number is to 100, the greater the potential for tourism recovery. Likewise, the numerical value of a sub-index reflects the potential extent of tourism recovery from different perspectives.

We conducted three sensitivity analyses on our index methodology (Joint Research Centre-European Commission, 2008). First, we generated the overall index using alternative aggregation methods: the arithmetic mean and the harmonic mean. Second, we used a different method to calculate the adjustment factor, which specifies separate factors adjusted for weekend and weekday data per country. Third, we constructed the overall index with all five sub-indices. High R^2 values (between, within, and overall) between the COVID19tourism index and the index using alternative construction convey the general robustness of our index methodology. Detailed results can be found in the supplementary materials.

We developed a web-GIS dashboard to present the index interactively. The dashboard is hosted on the ESRI ArcGIS web service and supplies current and retrospective data. Figure 1 provides a snapshot of the dashboard. The left panel displays the country list in descending order of the latest overall index. In the center, the dashboard presents the average index score for the overall index and its sub-indices as well as statistical graphs. Within the map view of each (sub-)index, users can click any country of interest to view the latest data of each country.

(Please insert Figure 1 about here)

Index statistics

Table 2 provides the average of the index and sub-indices as of October 24, 2020. The world average of the COVID19tourism index was 27.483 based on a sample of 100 countries. The recovery level was highest in Oceania and lowest in South America according to the index average. Regarding sub-indices, Asia recorded the lowest aviation recovery, while Western and Eastern Europe witnessed substantially less recovery in hotels compared with other parts of the world. Notably, although the overall recovery remained low in Eastern Europe, its average interest sub-index, which was highest among all regions, exhibited strong recovery in potential tourists' information search behavior for this destination region. Figure 2 presents the time-series curve of the global average. The overall index bottomed out on April 18 and climbed slowly thereafter. After July 11, it started to decline moderately over time. We also clustered the COVID19tourism index curve of different countries to recognize the pattern of tourism recovery. The results can be found in the supplementary materials.

(Please insert Table 2 about here)

(Please insert Figure 2 about here)

Discussion and conclusion

The COVID19tourism index has far-reaching utility for travel and tourism practitioners, researchers, travelers, and government entities. The instrument provides easily digestible travel and tourism information that can be broken down to shed light on multiple travel and tourism sectors. Our index also provides scholars a unique dataset; daily data capture high-frequency tourism industry performance throughout the pandemic. Numerous econometric models can be adopted to explore factors leading to fluctuations in performance. Moreover, the index and its sub-indices can serve as independent and control variables in various econometric models. Lastly, our data can generate statistics as parameter input for future economic simulation analysis, such as macro-economic modeling efforts (Yang, Zhang, & Chen, 2020).

For travelers, having timely access to data related to aviation, hotels, the pandemic, and mobility is essential when planning domestic and international trips. Industry uses are even broader: the index enables practitioners to benchmark and compare data points with possible competitors. The COVID19tourism index also allows for comparisons with countries in the same geographical area, such that users can identify the strengths of a given destination versus a potential competitor. Destinations can refer to the interest sub-index to assess their visibility to potential travelers even as the pandemic endures. At present, COVID-19 cases are continuing to rise at a staggering rate—making it all the more important for destinations to assess their recovery status and devise appropriate forecasts. Strong tourism recovery relies on stakeholders' ability to predict the future (Ritchie & Jiang, 2019). Based on other countries' trajectories in different stages of recovery, it is possible to foresee future patterns in tourism recovery. Put simply, the COVID19tourism index produces accessible graphs for all sub-indices, affording destinations the chance to trace trends and accurately project the future.

At the governmental level, the COVID19tourism index can contextualize travel and tourism policies. The World Travel and Tourism Council (2020) recently declared that developing travel "bubbles" of countries with similarly low COVID-19 cases can partially support the travel and tourism industry. The larger these travel bubbles, the greater the benefits for included countries and for the travel and tourism industry at large. As the pandemic continues to unfold, national governments can also use index information to continually re-evaluate and update their travel and tourism policies. Also, based on the index, destination governments can allocate resources to aid the tourism industry under various scenarios while prioritizing any sectors requiring immediate attention. Amid growing geopolitical tension among some countries during the pandemic, evidence-based decision making is crucial to effective governance.

Some limitations of this index should be noted. First, all sub-indices were constructed using a single measure, which may lead to potential biases. For example, the quality of RevPAR data depends on the representativeness of surveyed hotels, the quality of COVID-19 case data depends on the testing capacity of the country, and the quality of Google data depends on the market share of Google products. In particular, the Google search volume based on a country's name disregard the volume pattern of major regions within the country, some of which could be more popular in Google search than the country itself. Second, we only incorporated the air flight volume through aviation sub-index, which overlooks other major transport modes, including land transport that play a significant role in mobilizing international tourists in regions like Europe. Third, the baseline was constructed using January data over two years without thoroughly considering the seasonality of tourism activities. This baseline issue can be particularly impactful for small countries with a strong seasonality pattern in tourism, such as those in the Caribbean. Lastly, the exclusion of mobility and interest sub-indices from the overall index may disregard the information on within-country mobility and latent demand for travel.

Table 3 guides index users by summarizing the suitability level of the index under different scenarios. The overall index can be less suitable to understand domestic than international tourism as it does not encompass land transport volume. Furthermore, because VFR tourists are less likely to stay in hotels, the overall index, as well as hotel and interest sub-indices, are not necessarily suitable to analyze this segment compared to others. Additionally, constrained by the

COVID-19 testing capacity and the hotel sample size reporting the performance data, the overall index can be less applicable in developing countries compared to developed ones.

(Please insert Table 3 about here)

References

Joint Research Centre-European Commission. (2008). *Handbook on Constructing Composite Indicators: Methodology and User Guide*: OECD publishing.

Ritchie, B. W., & Jiang, Y. (2019). A review of research on tourism risk, crisis and disaster management: Launching the annals of tourism research curated collection on tourism risk, crisis and disaster management. *Annals of Tourism Research, 79*, 102812.

World Travel and Tourism Council. (2020). WTTC lays out guidelines for a safe and seamless travel in the 'new normal'. Retrieved from https://wttc.org/News-Article/WTTC-lays-out-guidelines-for-a-Safe-and-Seamless-Travel-in-the-new-normal

Yang, Y., Zhang, H., & Chen, X. (2020). Coronavirus pandemic and tourism: Dynamic stochastic general equilibrium modeling of infectious disease outbreak. *Annals of Tourism Research, 83*, 102913.

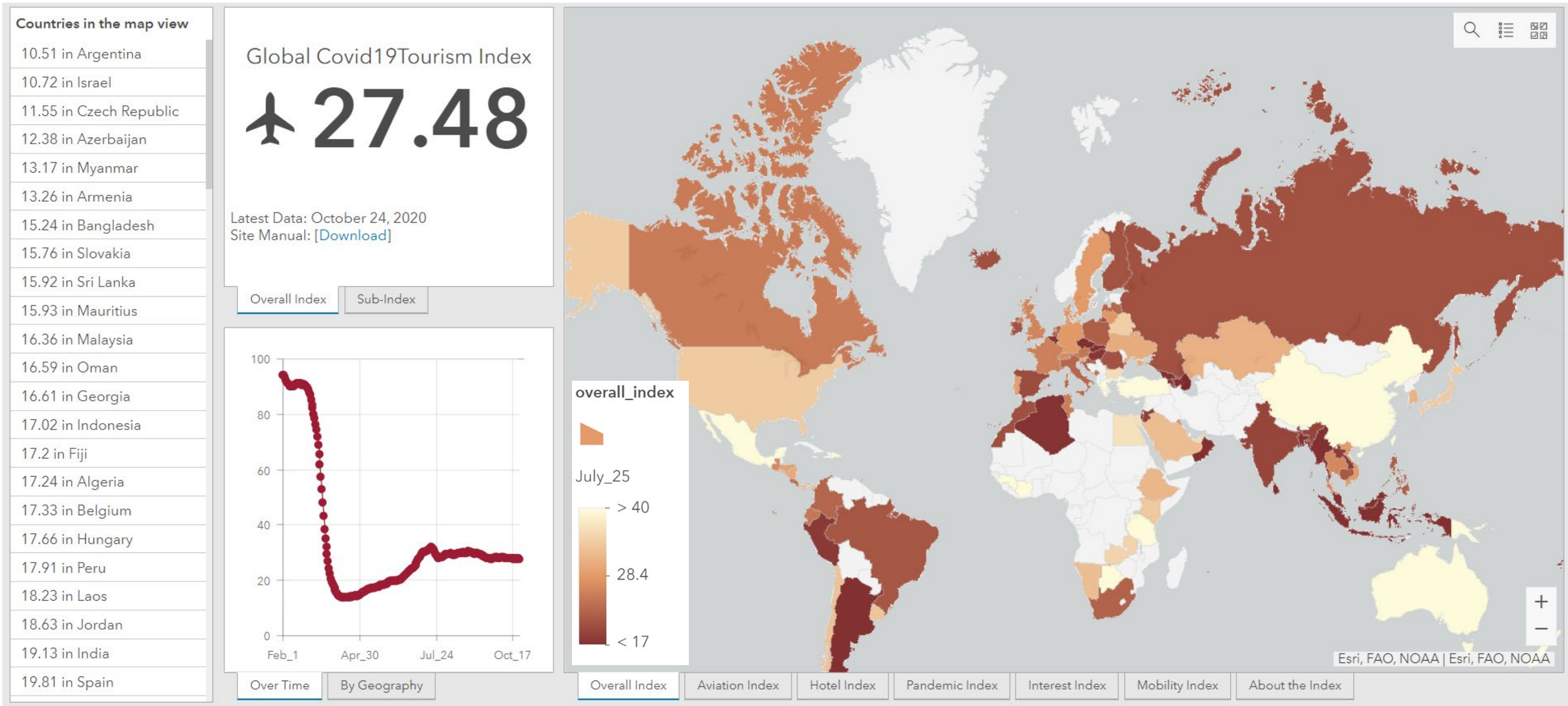


Figure 1. Web-GIS platform of COVID19tourism index

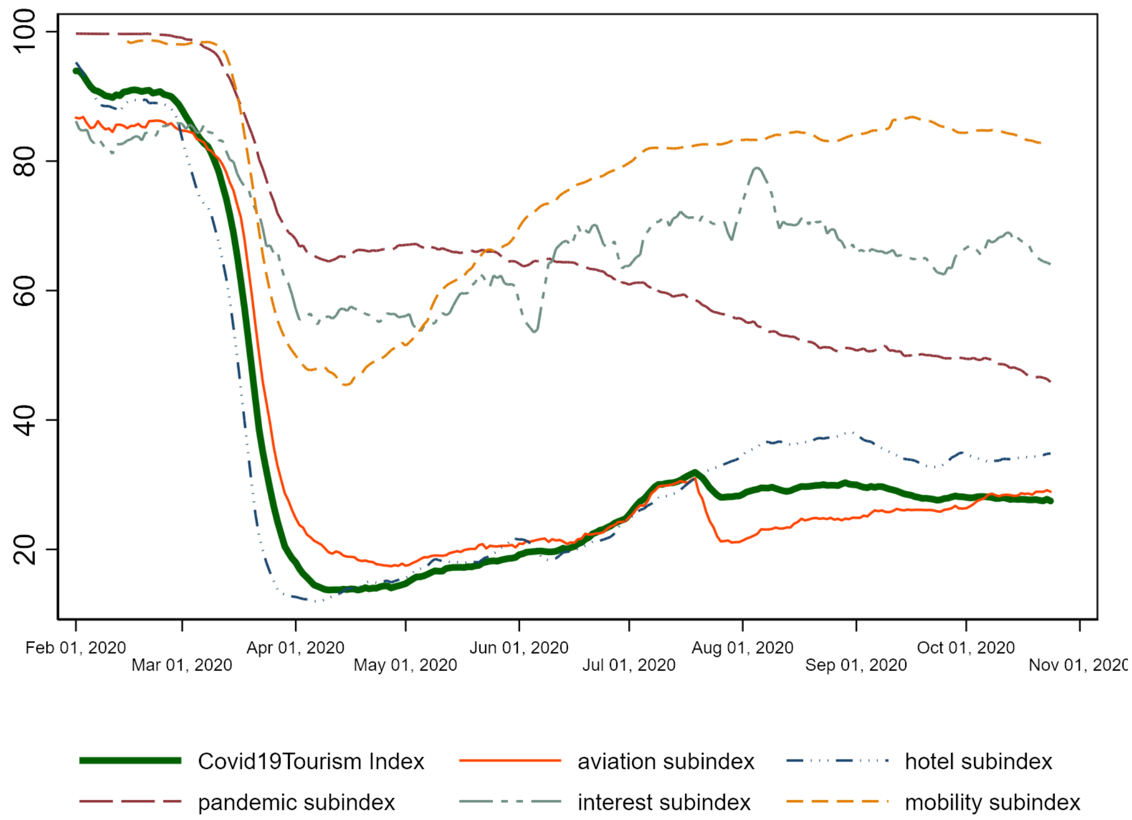


Figure 2. Time-series graph of overall index and sub-indices

Table 1. Operationalization of sub-indices

Sub-indices	Raw Data	Baseline (equal to 100)	Data Source	Included in overall index
Aviation	Daily departure data from a country's international airports	Year-over-year adjusted data (Equation 1)	https://www.icao.int/safety/Pages/COVID-19-Airport-Status.aspx	Yes
Hotel	Daily data on a country's hotel RevPAR	Year-over-year adjusted data (Equation 1)	STR, LLC	Yes
Pandemic	Daily number of confirmed COVID-19 cases per 1 million inbound tourists in 2018	1 confirmed case per 1 million inbound tourists or less	World Development Indicators; https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide	Yes
Interest	Daily Google search volume for keyword "[country name] hotel"	Jan 3–Feb 6, 2020	Google Trends	No
Mobility	Daily Google community mobility data in locations of "Retail and recreation," "Parks," and "Public transport", which are directly related to travel and tourism.	Jan 3–Feb 6, 2020	https://www.google.com/covid19/mobility/	No

Table 2. Average of index and sub-indices over regions as of October 24, 2020

Region	Overall	Aviation	Hotel	Pandemic	Interest	Mobility
Africa	33.705	39.608	43.707	46.855	62.629	90.418
	(15)	(50)	(18)	(40)	(55)	(28)
Asia	24.425	12.649	33.973	61.208	69.992	78.681
	(17)	(23)	(17)	(23)	(23)	(17)
Central America	32.580	35.993	36.415	54.342	50.509	68.543
	(9)	(20)	(11)	(27)	(32)	(15)
Eastern Europe	23.197	19.993	25.217	27.090	95.516	89.564
	(11)	(13)	(11)	(13)	(13)	(13)
Middle East	27.183	23.890	39.041	28.704	65.920	83.936
	(13)	(22)	(13)	(18)	(24)	(17)
North America	28.763	27.146	40.463	24.552	64.230	85.562
	(2)	(2)	(2)	(2)	(2)	(2)
Oceania	38.804	39.629	52.514	86.722	37.244	87.564
	(4)	(14)	(6)	(17)	(20)	(4)
South America	23.106	18.556	33.485	28.546	66.868	66.565
	(8)	(13)	(10)	(15)	(16)	(12)
Western Europe	25.149	25.590	25.422	31.044	75.839	89.647
	(21)	(26)	(23)	(31)	(35)	(23)
Total	27.483	28.920	34.823	45.871	64.055	82.982
	(100)	(183)	(111)	(186)	(220)	(131)

(notes: parenthesis indicates the number of countries in the sample)

Table 3. Suitability of index application under different scenarios

	Overall	Aviation	Hotel	Pandemic	Interest	Mobility
Market composition						
-international	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
-domestic	☆☆	☆☆	☆☆☆	☆☆☆	☆☆	☆☆☆
Tourist type						
-leisure	☆☆☆	☆☆☆	☆☆	☆☆☆	☆☆☆	☆☆☆
-VFR	☆☆	☆☆☆	☆	☆☆☆	☆	☆☆☆
-business	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
Economy level						
-developed	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
-developing	☆☆	☆☆☆	☆☆	☆☆	☆☆☆	☆☆
Means of arrival						
-land	☆☆	☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆
-air	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆☆	☆☆

(Note: ☆☆☆ indicates highly suitable with little limitations, ☆☆ indicates reasonably suitable with notable limitations, and ☆ indicates marginally suitable with significant limitations)

Supplementary materials

1. Sensitivity analysis on index construction

Table S1 indicates a very high level of correlation between the COVID19tourism index and alternative indexes. Within R^2 indicates this correlation over time for a country, whereas between R^2 indicates the correlation across different countries. The only minor discrepancy can be found in the between R^2 of 0.843 with the index based on the arithmetic mean. We can further check this in the following scatterplot matrix.

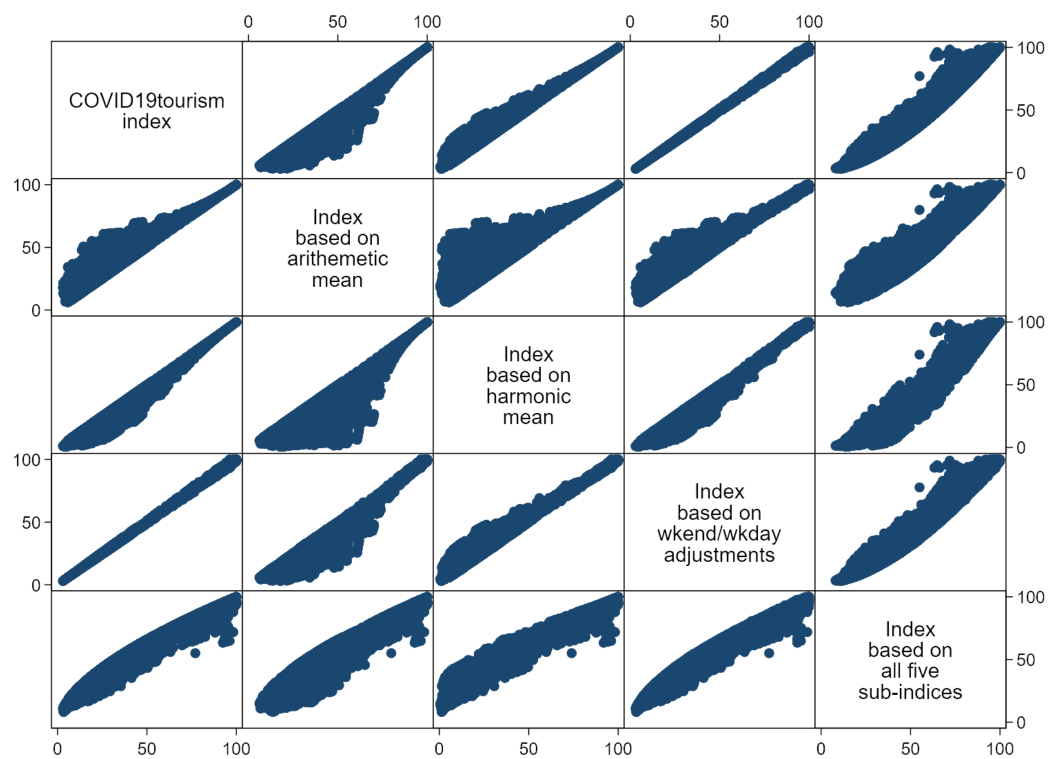
Table S1. Correlation between COVID19tourism index and alternative indexes for sensitivity analysis

	Index based on arithmetic mean	**Index based on harmonic mean**	**Index based on wkend/weekday adjustments**	**Index based on all five sub-indices***
Within R^2	0.965	0.992	1.000	0.971
Between R^2	0.843	0.937	1.000	0.867
overall R^2	0.944	0.985	1.000	0.954

*(note: * indicates only a part of countries are available for the index based on all five sub-indices)*

As shown in Figure S1, the COVID19tourism index is highlight correlated with indices using alternative aggregation methods. In particular, we found that if we use a linear aggregation (arithmetic mean), the index value increases at lower values. This is because the geometric aggregation we used penalized the sub-index has a low value more than that in the linear aggregation.

Figure S1. Scatterplot matrix between different indexes for sensitivity analysis



2. Clustering analysis

The time-series curve of retrospective index data offers vital information to track the pandemic's impacts over time and predict future trends. We performed a clustering analysis of the index curve using the overall index for 84 countries. China was excluded because it saw the initial emergence of COVID-19, and the sample period failed to capture the early decline in the country's index curve. Partitional clustering was employed by choosing the dynamic time-warping distance, a distance measure that reflects similarities in the shapes of time-series curves (Petitjean, Ketterlin, & Gançarski, 2011). The prototype (centroid) of each cluster was calculated by taking the average of each time point across all time series in that cluster. After a trial with different numbers of clusters, we determined that a three-cluster solution was satisfactory. Figure S2 illustrates the synchronized prototype curve for each cluster. The shapes of different countries' curves share similarities, and all countries witnessed a rapid descent to their respective minimums. The detailed shape of a given curve can be distinguished by its tail. Based on the curve shape, Clusters 1 and 3 underwent a U-shaped recovery, while Cluster 2 experienced an L-shaped recovery. Cluster 2 possessed a flat tail that was the lowest of several clusters, indicating that countries in this cluster have consistently suffered from substantial COVID-19 losses. The curve of Cluster 1 began to decline after a recovery in late July, while that of Cluster 3 continued to rise after an initial recovery. Figure S3 presents index curves of countries in different clusters, and Table S2 lists country's cluster membership

Figure S2. Prototype curves of different clusters

Figure S3. Index curves of countries in different clusters

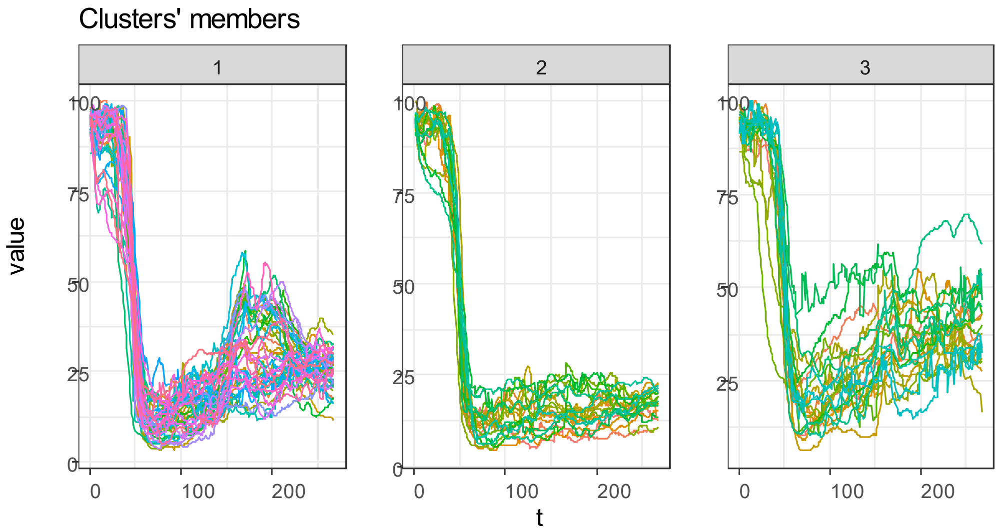


Table S2. List of country's cluster membership

country	**cluster ID**
Austria	1
Bahrain	1
Belgium	1
Cambodia	1
Canada	1
Chile	1
Czech Republic	1
Denmark	1
Ecuador	1
Egypt	1
El Salvador	1
Finland	1
France	1
Hungary	1
Ireland	1
Italy	1
Kazakhstan	1
Kenya	1
Kuwait	1
Latvia	1
Lithuania	1
Malaysia	1

Maldives	1
Malta	1
Netherlands	1
Nicaragua	1
Panama	1
Poland	1
Portugal	1
Romania	1
Saudi Arabia	1
Singapore	1
Spain	1
Sri Lanka	1
Sweden	1
Switzerland	1
Thailand	1
Trinidad / Tobago	1
Tunisia	1
Ukraine	1
United Kingdom	1
Vietnam	1
Algeria	2
Argentina	2
Armenia	2
Azerbaijan	2
Bangladesh	2
Brazil	2
Colombia	2
Costa Rica	2
Fiji	2
India	2
Indonesia	2
Israel	2
Jordan	2
Morocco	2
Myanmar	2
Oman	2
Peru	2
Philippines	2
Russia	2
South Africa	2
Australia	3
Belarus	3
Bulgaria	3
Cote D`Ivoire	3

Dominican Republic	3
Ethiopia	3
Georgia	3
Germany	3
Guinea	3
Japan	3
Lebanon	3
Mexico	3
New Zealand	3
Papua New Guinea	3
Qatar	3
South Korea	3
Tanzania	3
Turkey	3
United Arab Emirates	3
United States	3
Uruguay	3
Zambia	3