ANALYSIS OF GLOBAL COVID-19 DATA

INTERACTIVE VISUALISATION, COUNTRY COMPARISON, CLUSTER AND BASELINE ANALYSIS

Nacabodi Nacaskul
Triam Udom Suksa School
Bangkok, Thailand
nacabodi.nacaskul@gmail.com

September 26, 2021

ABSTRACT

This article demonstrates interactive visualisation tools, country/continent comparison, cluster analysis, and comparison of country/continent incidence numbers (Covid-related cases and deaths) against baseline fit using simple *Supervised Learning* method, using **Wolfram Mathematica**.

1 Introduction

Since early 2020, the Covid-19 pandemic has been a topic of discussion all around the world. Statistics regarding the spread of the virus are displayed and updated on a day-to-day basis by various websites, namely `https://coronavirus.jhu.edu`, `https://www.worldometers.info/coronavirus`, and `https://ourworldindata.org/coronavirus`.

As a data science enthusiast, I spent a considerable amount of time delving deep into those websites. Not long after, something piqued my curiosity. It became clear that different countries around the world performed differently in combating the Covid-19 pandemic. Some did exceptionally well, while others struggled.

This article will explain the use of **Wolfram Mathematica** functions, both built-in and user-defined, in visualizing and analysing data from the website ourworldindata.org. The functions are archived in the **AnalysisCovidData** library in Github
`https://github.com/Nacabodi-Nacaskul/AnalysisCovidData`.

2 Data

2.1 Source

This article is based on data retrieved from `https://raw.githubusercontent.com/owid/covid-19-data/master/public/data/owid-covid-data.csv` on 2021.9.26 (6:00 GMT)[1].

Data and statistics from the aforementioned websites include:

1. incidence numbers (i.e. daily/total cases/deaths attributed to Covid-19),

2. national demographic/health statistics (i.e. median age, prevalence of diabetes), and

3. socio-economic factors (i.e. GDP per capita, Human Development Index).

They cover individual countries, continents, and global aggregates. Covid incidence numbers are updated daily, while other statistics are based on most recent available.

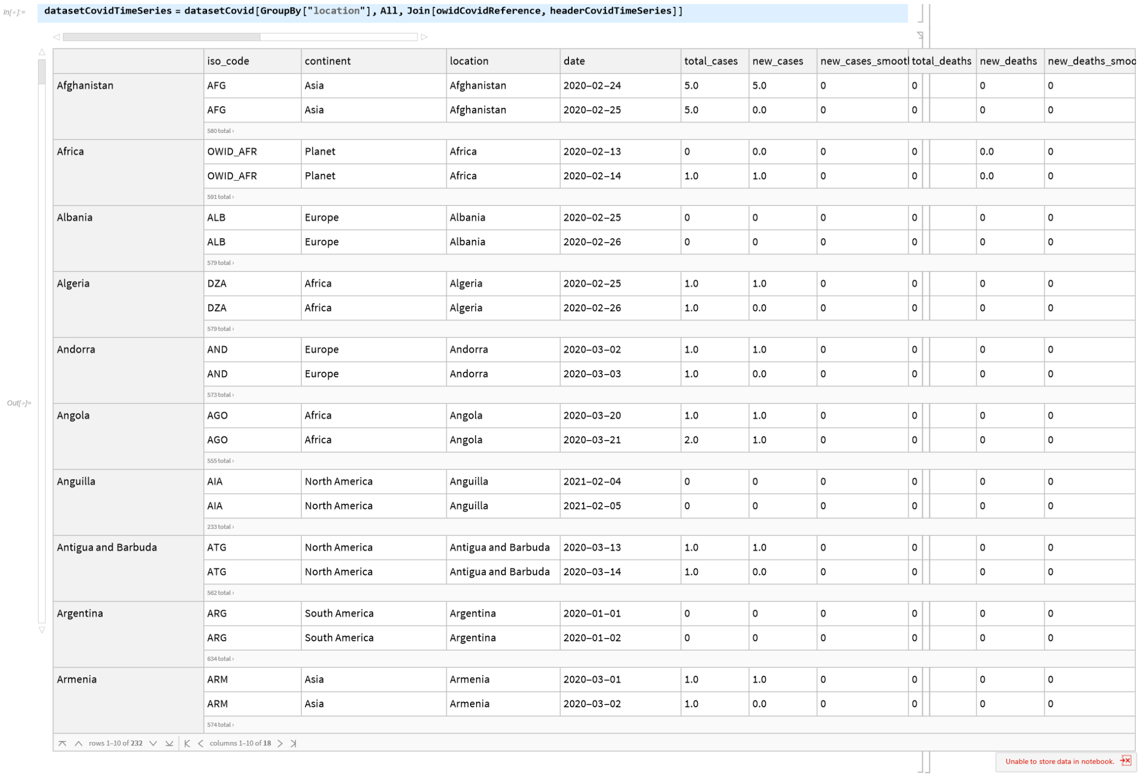
The size of this data table is given by **Dimensions[]**, displaying the number of rows, followed by the number of columns (here 119,245 rows, including 1 row for "header", and 64 columns, of which 4 are "reference" fields, namely "iso code", "continent", "location", and "date").

2.2 Data Structure

We can use the Mathematica function **ListPlot[]** and **DateListPlot[]** to view data over a period of time(Time Series Plot). However, we will take into consideration only the latest data.

The data is first imported into Mathematica as a **List**. We then use **Dataset[]** to represent grouped data. For convenience, we have created 3 datasets:

1. **datasetCovidTimeSeries** stores all incidence numbers since the beginning of the pandemic, see figure 1.



```
datasetCovidTimeSeries = datasetCovid[GroupBy["location"], All, Join[owidCovidReference, headerCovidTimeSeries]]
```

	iso_code	continent	location	date	total_cases	new_cases	new_cases_smootl	total_deaths	new_deaths	new_deaths_smoo
Afghanistan	AFG	Asia	Afghanistan	2020–02–24	5.0	5.0	0	0	0	0
	AFG	Asia	Afghanistan	2020–02–25	5.0	0.0	0	0	0	0
	580 total ›									
Africa	OWID_AFR	Planet	Africa	2020–02–13	0	0.0	0	0	0.0	0
	OWID_AFR	Planet	Africa	2020–02–14	1.0	1.0	0	0	0.0	0
	591 total ›									
Albania	ALB	Europe	Albania	2020–02–25	0	0	0	0	0	0
	ALB	Europe	Albania	2020–02–26	0	0	0	0	0	0
	579 total ›									
Algeria	DZA	Africa	Algeria	2020–02–25	1.0	1.0	0	0	0	0
	DZA	Africa	Algeria	2020–02–26	1.0	0.0	0	0	0	0
	579 total ›									
Andorra	AND	Europe	Andorra	2020–03–02	1.0	1.0	0	0	0	0
	AND	Europe	Andorra	2020–03–03	1.0	0.0	0	0	0	0
	573 total ›									
Angola	AGO	Africa	Angola	2020–03–20	1.0	1.0	0	0	0	0
	AGO	Africa	Angola	2020–03–21	2.0	1.0	0	0	0	0
	555 total ›									
Anguilla	AIA	North America	Anguilla	2021–02–04	0	0	0	0	0	0
	AIA	North America	Anguilla	2021–02–05	0	0	0	0	0	0
	233 total ›									
Antigua and Barbuda	ATG	North America	Antigua and Barbuda	2020–03–13	1.0	1.0	0	0	0	0
	ATG	North America	Antigua and Barbuda	2020–03–14	1.0	0.0	0	0	0	0
	562 total ›									
Argentina	ARG	South America	Argentina	2020–01–01	0	0	0	0	0	0
	ARG	South America	Argentina	2020–01–02	0	0	0	0	0	0
	634 total ›									
Armenia	ARM	Asia	Armenia	2020–03–01	1.0	1.0	0	0	0	0
	ARM	Asia	Armenia	2020–03–02	1.0	0.0	0	0	0	0
	574 total ›									

Figure 1: datasetCovidTimeSeries

2. **datasetCovidLatestStats** stores statistics we wish to analyse available on the latest date (i.e. 2021.9.25) for each location (country, continent, or "World"), see figure 2.

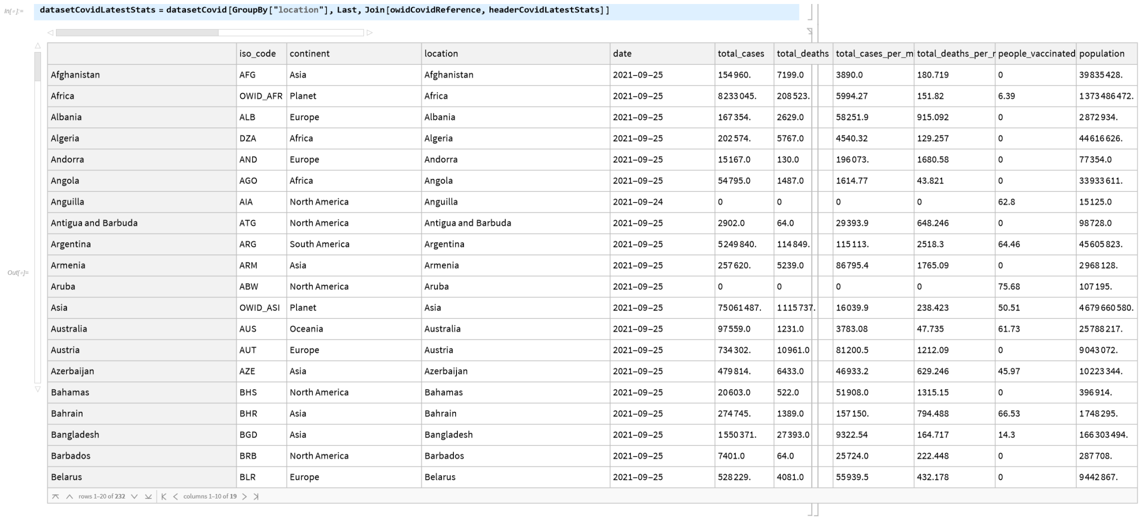
3. **datasetCovidLatestStatsContinentAve** then groups **datasetCovidLatestStats** by continents, and compute averages, see figure 3.

2.3 Visualisation

There are several methods to visualize data, but each of them bear different pros and cons. The Scatter Plot method can display many countries at a time, but only in 2-3 dimensions. The Radar Plot method can display many dimensions at a time, but only show a few countries.

2.3.1 Time Series

We made a function called **viewCovidTimeSeries[]**, which is capable of displaying data based on the location and series of our choices over a period of time. See figure 4 and figure 5 for ["World", "new cases"] and ["Thailand", "new cases smoothed per million"], respectively.

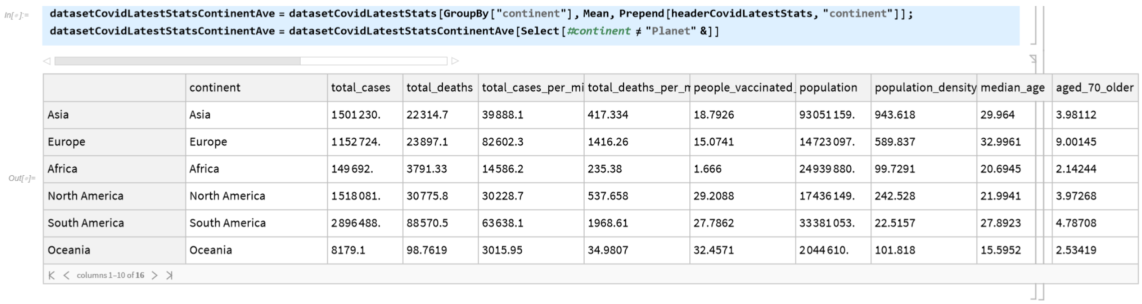


```
datasetCovidLatestStats = datasetCovid[GroupBy["location"], Last, Join[owidCovidReference, headerCovidLatestStats]]
```

	iso_code	continent	location	date	total_cases	total_deaths	total_cases_per_m	total_deaths_per_r	people_vaccinated	population
Afghanistan	AFG	Asia	Afghanistan	2021–09–25	154 960.	7199.0	3890.0	180.719	0	39 835 428.
Africa	OWID_AFR	Planet	Africa	2021–09–25	8 233 045.	208 523.	5994.27	151.82	6.39	1 373 486 472.
Albania	ALB	Europe	Albania	2021–09–25	167 354.	2629.0	58 251.9	915.092	0	2 872 934.
Algeria	DZA	Africa	Algeria	2021–09–25	202 574.	5767.0	4540.32	129.257	0	44 616 626.
Andorra	AND	Europe	Andorra	2021–09–25	15 167.0	130.0	196 073.	1680.58	0	77 354.0
Angola	AGO	Africa	Angola	2021–09–25	54 795.0	1487.0	1614.77	43.821	0	33 933 611.
Anguilla	AIA	North America	Anguilla	2021–09–24	0	0	0	0	62.8	15 125.0
Antigua and Barbuda	ATG	North America	Antigua and Barbuda	2021–09–25	2902.0	64.0	29 393.9	648.246	0	98 728.0
Argentina	ARG	South America	Argentina	2021–09–25	5 249 840.	114 849.	115 113.	2518.3	64.46	45 605 823.
Armenia	ARM	Asia	Armenia	2021–09–25	257 620.	5239.0	86 795.4	1765.09	0	2 968 128.
Aruba	ABW	North America	Aruba	2021–09–25	0	0	0	0	75.68	107 195.
Asia	OWID_ASI	Planet	Asia	2021–09–25	75 061 487.	1 115 737.	16 039.9	238.423	50.51	4 679 660 580.
Australia	AUS	Oceania	Australia	2021–09–25	97 559.0	1231.0	3783.08	47.735	61.73	25 788 217.
Austria	AUT	Europe	Austria	2021–09–25	734 302.	10 961.0	81 200.5	1212.09	0	9 043 072.
Azerbaijan	AZE	Asia	Azerbaijan	2021–09–25	479 814.	6433.0	46 933.2	629.246	45.97	10 223 344.
Bahamas	BHS	North America	Bahamas	2021–09–25	20 603.0	522.0	51 908.0	1315.15	0	396 914.
Bahrain	BHR	Asia	Bahrain	2021–09–25	274 745.	1389.0	157 150.	794.488	66.53	1 748 295.
Bangladesh	BGD	Asia	Bangladesh	2021–09–25	1 550 371.	27 393.0	9322.54	164.717	14.3	166 303 494.
Barbados	BRB	North America	Barbados	2021–09–25	7401.0	64.0	25 724.0	222.448	0	287 708.
Belarus	BLR	Europe	Belarus	2021–09–25	528 229.	4081.0	55 939.5	432.178	0	9 442 867.

rows 1–20 of 232 columns 1–10 of 19

Figure 2: datasetCovidLatestStats



```
datasetCovidLatestStatsContinentAve = datasetCovidLatestStats[GroupBy["continent"], Mean, Prepend[headerCovidLatestStats, "continent"]];
datasetCovidLatestStatsContinentAve = datasetCovidLatestStatsContinentAve[Select[#continent ≠ "Planet" &]]
```

	continent	total_cases	total_deaths	total_cases_per_mi	total_deaths_per_m	people_vaccinated_	population	population_density	median_age	aged_70_older
Asia	Asia	1 501 230.	22 314.7	39 888.1	417.334	18.7926	93 051 159.	943.618	29.964	3.98112
Europe	Europe	1 152 724.	23 897.1	82 602.3	1416.26	15.0741	14 723 097.	589.837	32.9961	9.00145
Africa	Africa	149 692.	3791.33	14 586.2	235.38	1.666	24 939 880.	99.7291	20.6945	2.14244
North America	North America	1 518 081.	30 775.8	30 228.7	537.658	29.2088	17 436 149.	242.528	21.9941	3.97268
South America	South America	2 896 488.	88 570.5	63 638.1	1968.61	27.7862	33 381 053.	22.5157	27.8923	4.78708
Oceania	Oceania	8179.1	98.7619	3015.95	34.9807	32.4571	2 044 610.	101.818	15.5952	2.53419

columns 1–10 of 16

Figure 3: datasetCovidLatestStatsContinentAve

2.3.2 Scatter Plot

Scatter plot is capable of displaying multiple countries, but it can only show 2 dimensions at a time (as seen in x-axis vs y-axis). See figure 6 and figure 7, for example.

While we present 2D scatter plots here for clarity, **Mathematica** does provide **ListPlot3D[]** function which allows users to interactively control viewing perspective via mouse.

2.3.3 Radar Plot

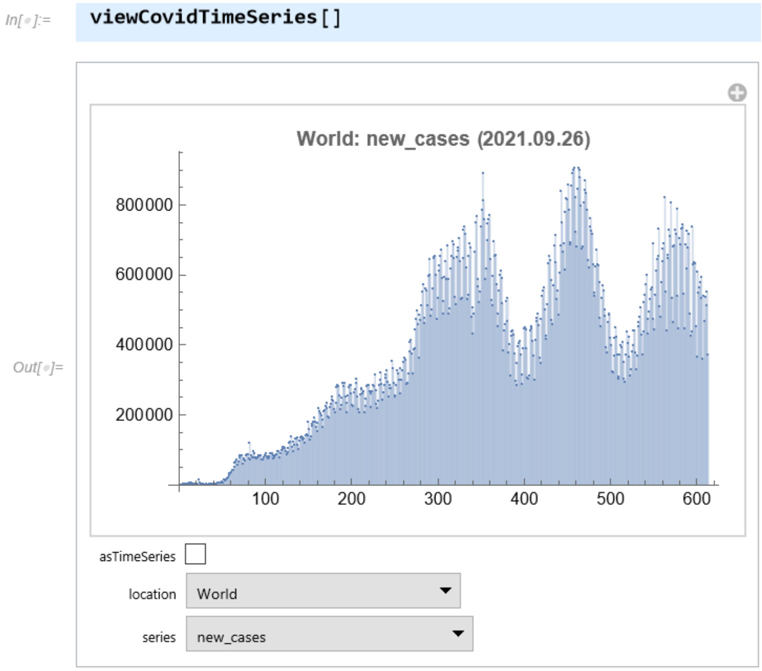
Radar plot, implemented using **Mathematica**'s built-in function **RadialAxisPlot[]**, is capable of displaying multiple dimensions. While it can technically display multiple countries at once, the result will appear tangled beyond what we can comprehend. Thus, we resort to displaying average data of continents instead of countries, as they are fewer in numbers, see figure 8.

We can also view a subset of continents in terms of a subset of attributes, see figure 9.

3 Methodology/Analysis

3.1 Cluster Analysis

We wondered if there are any clusters formed among these 12 dimensions, specifically "total cases per million", "total deaths per million", "people vaccinated per hundred", "population density", "median age", "aged 70 older", "gdp per capita", "cardiovasc death rate", "diabetes prevalence", "hospital beds per thousand", "life expectancy", and "human development index".



```
In[•]:= viewCovidTimeSeries[]
Out[•]=
```

asTimeSeries □

location World

series new_cases

Figure 4: viewCovidTimeSeries ["World", "new cases"]

Here we used **Mathematica**'s built-in function **FindClusters[]** to perform *Cluster Analysis* [2][3]. The result is shown in figure 10.

FindClusters[] found 4 clusters, with 19, 183, 12, and 18 members, respectively.

We can then use **Mathematica**'s built-in function **MemberQ[]** to query which cluster a particular country belongs to. For example, "Thailand" and "World" both belong to the second cluster.

For each cluster, we can use a radar plot to display the *Centroid* of each cluster. See figure 11

We can also view any 2 dimensions of individual locations by the cluster they belong to on a scatter plot, but the picture would be highly misleading, as *Cluster Analysis* is based on taking account all 12 dimensions simultaneously.

To get around this issue, we can perform *Dimensionality Reduction* [4] [5] [6] , essentially "compressing" 12 real dimensions into 2 "artificial" dimensions but in such a way that points that are close together in the original 12D space are also close together in the 2D space of artificial dimensions. And similarly, points that are further apart in the original 12D space are also further apart in the 2D space of artificial dimensions.

Mathematica implements *Dimensionality Reduction* as a built-in function **DimensionReduce[]**.

With that, we can then use **ListPlot[]** to visualise the clustered data points on a (2D) scatter plot. See figure 12 and 13 for the first and second cluster, respectively.

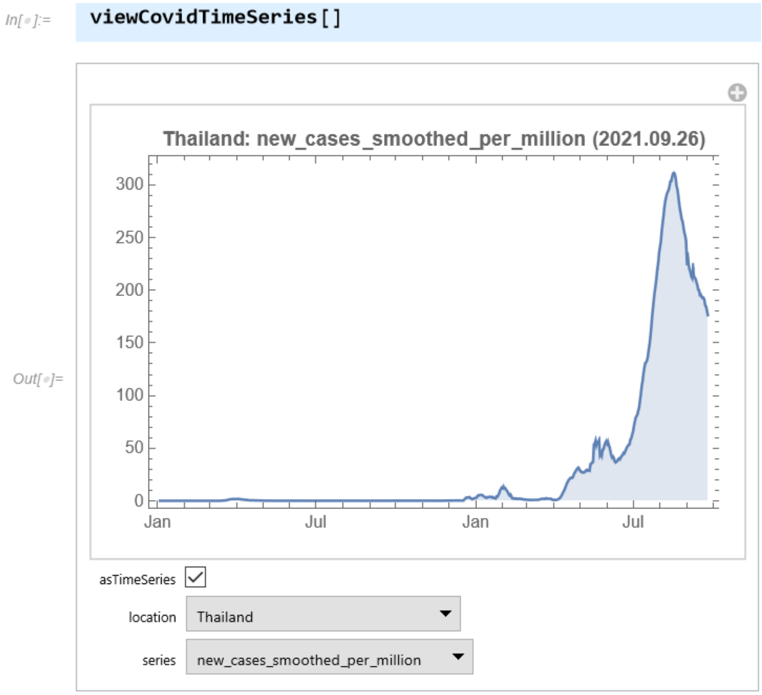
3.2 Baseline Analysis

I was also curious to see if "total cases per million" and "total deaths per million" (target variables) can be explained by the other 10 dimensions (explanatory variables).

Here **Mathematica**'s built-in function **Predict[]** attempts to build basic *Supervised Learning* model, choosing from common methodologies, i.e. "LinearRegression", "NearestNeighbors", "DecisionTree", "RandomForest", "GradientBoostedTrees", "NeuralNetwork", etc.

With our dataset, "RandomForest" [7][8] turned out the be the method **Predict[]** used to fit both "total cases per million" and "total deaths per million" targets.

Based on this "baseline fit" 14 , we can assess whether a particular country/continent exceeded expectations in terms of Covid incidence.



```
In[•]:= viewCovidTimeSeries[]
```

Out[•]=

asTimeSeries ☑

location Thailand

series new_cases_smoothed_per_million

Figure 5: viewCovidTimeSeries ["Thailand", "new cases smoothed per million"]

Table 1: Defined Datasets

datasetCovid
datasetCovidTimeSeries
datasetCovidLatestStats
datasetCovidLatestStatsContinentAve

For example, based on this "baseline fit", Thailand should have about 60,646 "total cases per million" (cpmEst) and 1,097 "total deaths per million" (dpmEst), but actually suffered much less, i.e. 22,148 "total cases per million" (cpm) and 231 "total deaths per million" (dpm), or 37 percent and 21 percent against "baseline fit". Asia as a whole suffered 59 percent cpm/cpmEst and 40 percent dpm/dpmEst.

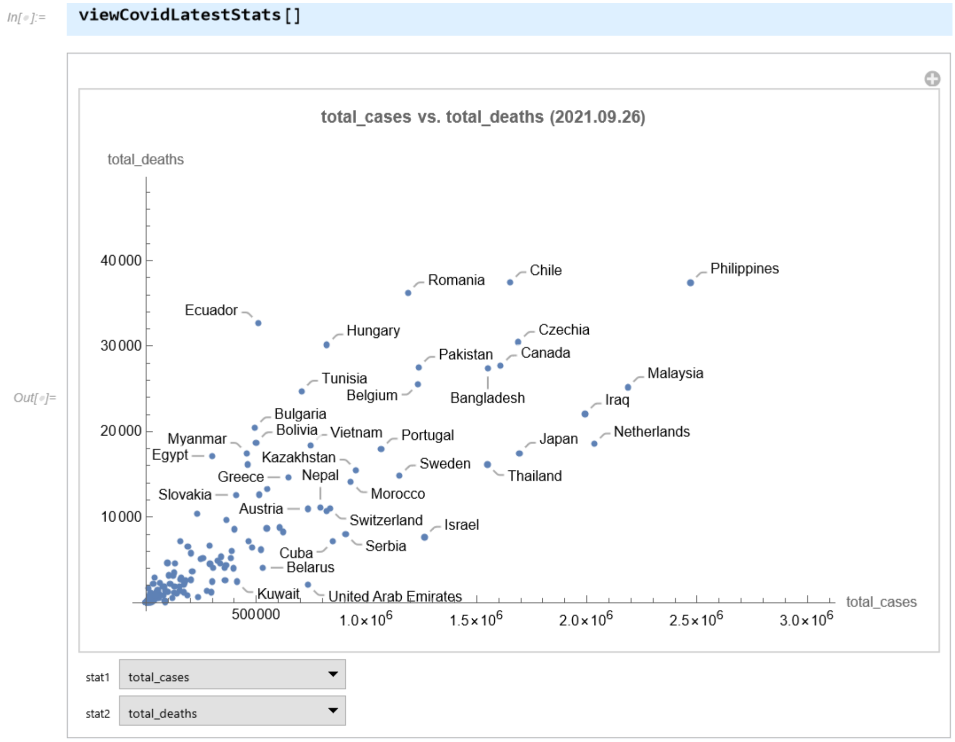
By such criteria, of all the continents, "Oceania" and "South America" had the best and worst outcomes, respectively.

4 Conclusion

We have explained the **Mathematica** function library in details, highlighting interactive visualisation functionalities. Relevant built-in functions, our user-defined **Dataset** objects, and our user-defined functions are listed in 5, 1, and 2, respectively.

Table 2: Defined Functions

viewCovidTimeSeries
viewCovidLatestStats
viewCovidLatestStatsByContinentAve
viewCovidLatestStatsByCluster
viewClustersByCentroidAttributes
viewReducedDimensionRepresentation
viewCovidIncidenceActualVsBaselineFit



In[∘]:= viewCovidLatestStats[]

Out[∘]=

stat1 total_cases

stat2 total_deaths

Figure 6: viewCovidLatestStats. Here x-axis and y-axis represent total cases and total deaths, respectively.

We found 4 different statistical clusters, taking into account "total cases per million", "total deaths per million", "people vaccinated per hundred", "population density", "median age", "aged 70 older", "gdp per capita", "cardiovasc death rate", "diabetes prevalence", "hospital beds per thousand", "life expectancy", and "human development index" statistics.

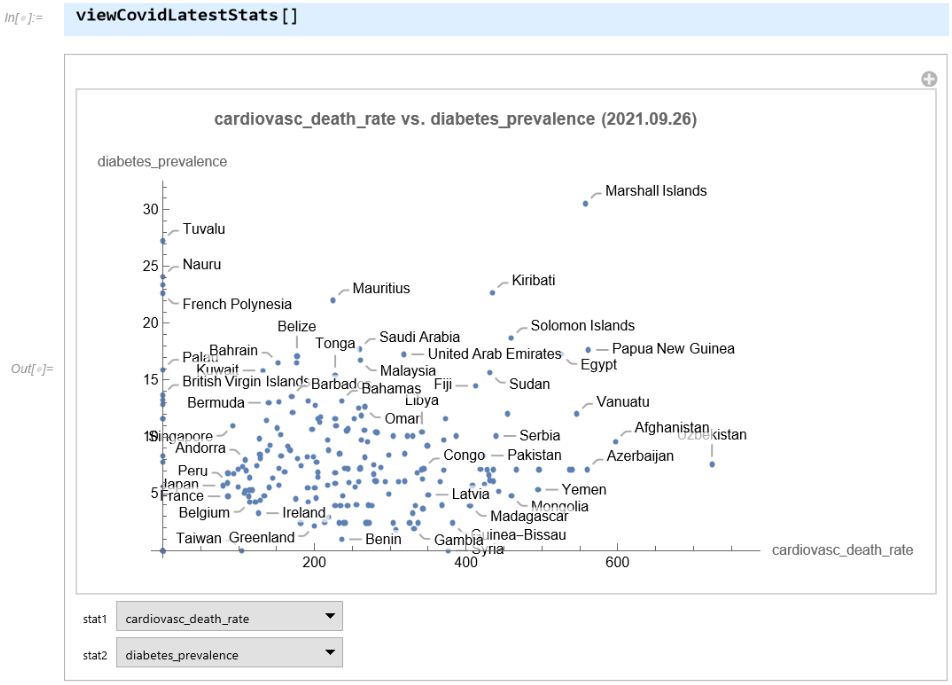
We were also able to construct "baseline fit" for "total cases per million" as well as for "total deaths per million", and can therefore highlight actual country/continent performance vs. such "baseline fit".

5 Glossary

5.1 Mathematica's Built-In Functions relevant to this study

- **Data Handling**
  `https://reference.wolfram.com/language/ref/Import.html`
  `https://reference.wolfram.com/language/ref/Export.html`
  `https://reference.wolfram.com/language/ref/Dimensions.html`
  `https://reference.wolfram.com/language/ref/Length.html`
  `https://reference.wolfram.com/language/ref/Append.html`
  `https://reference.wolfram.com/language/ref/Prepend.html`
  `https://reference.wolfram.com/language/ref/Complement.html`
  `https://reference.wolfram.com/language/ref/Replace.html`
  `https://reference.wolfram.com/language/ref/ReplaceAll.html`
  `https://reference.wolfram.com/language/ref/DeleteCases.html`
  `https://reference.wolfram.com/language/ref/DeleteDuplicates.html`
  `https://reference.wolfram.com/language/ref/Transpose.html`
  `https://reference.wolfram.com/language/ref/Thread.html`
  `https://reference.wolfram.com/language/ref/Last.html`

- **Data Structure**
  `https://reference.wolfram.com/language/ref/Table.html`
  `https://reference.wolfram.com/language/ref/Association.html`
  `https://reference.wolfram.com/language/ref/Dataset.html`



In[ ]:=
```
viewCovidLatestStats[]
```

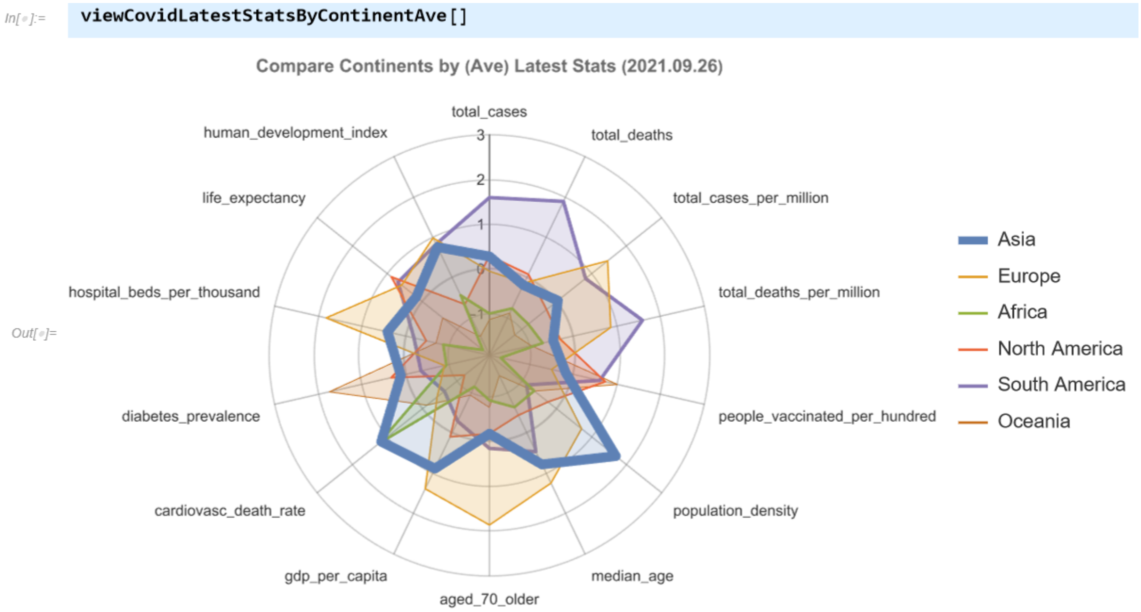
Out[ ]=

stat1 cardiovasc_death_rate

stat2 diabetes_prevalence

Figure 7: viewCovidLatestStats. Here x-axis and y-axis represent cardiovascular death rate and diabetes prevalence, respectively.

- **Dataset Query**
  `https://reference.wolfram.com/language/ref/Keys.html`
  `https://reference.wolfram.com/language/ref/Values.html`
  `https://reference.wolfram.com/language/ref/Select.html`
  `https://reference.wolfram.com/language/ref/GroupBy.html`
  `https://reference.wolfram.com/language/ref/Counts.html`
  `https://reference.wolfram.com/language/ref/CountsBy.html`
  `https://reference.wolfram.com/language/ref/MemberQ.html`
- **Numerical Computation**
  `https://reference.wolfram.com/language/ref/N.html`
  `https://reference.wolfram.com/language/ref/Total.html`
  `https://reference.wolfram.com/language/ref/Mean.html`
  `https://reference.wolfram.com/language/ref/Norm.html`
  `https://reference.wolfram.com/language/ref/Normalize.html`
  `https://reference.wolfram.com/language/ref/Standardize.html`
- **Plotting Functions**
  `https://reference.wolfram.com/language/ref/ListPlot.html`
  `https://reference.wolfram.com/language/ref/DateListPlot.html`
  `https://reference.wolfram.com/language/ref/RadialAxisPlot.html`
- **Function Shorthands**
  `https://reference.wolfram.com/language/ref/Function.html`
  `https://reference.wolfram.com/language/ref/Map.html`
- **Interactive Display**
  `https://reference.wolfram.com/language/ref/Manipulate.html`
  `https://reference.wolfram.com/language/ref/Style.html`
- **Data Analytics**
  `https://reference.wolfram.com/language/ref/FindClusters.html`
  `https://reference.wolfram.com/language/ref/DimensionReduce.html`
  `https://reference.wolfram.com/language/ref/Predict.html`



```
In[ ]:= viewCovidLatestStatsByContinentAve[]
```

Figure 8: **viewCovidLatestStatsByContinentAve** shows the performance of all continents. Each line represents a continent. Notice how the graph will become incomprehensible if we include all countries as there will be over a hundred lines.

6 Reference

References

[1] Our World In Data (2021). "Coronavirus Pandemic (COVID-19)" `https://ourworldindata.org/coronavirus`

[2] Xu, R. & Wunsch II, D. (2005). "Survey of Clustering Algorithms" *IEEE Transactions on Neural Networks* vol. 16, no. 3 (May), pp.645-678.

[3] Wikipedia (2021). "Cluster analysis" `https://en.wikipedia.org/wiki/Cluster_analysis`

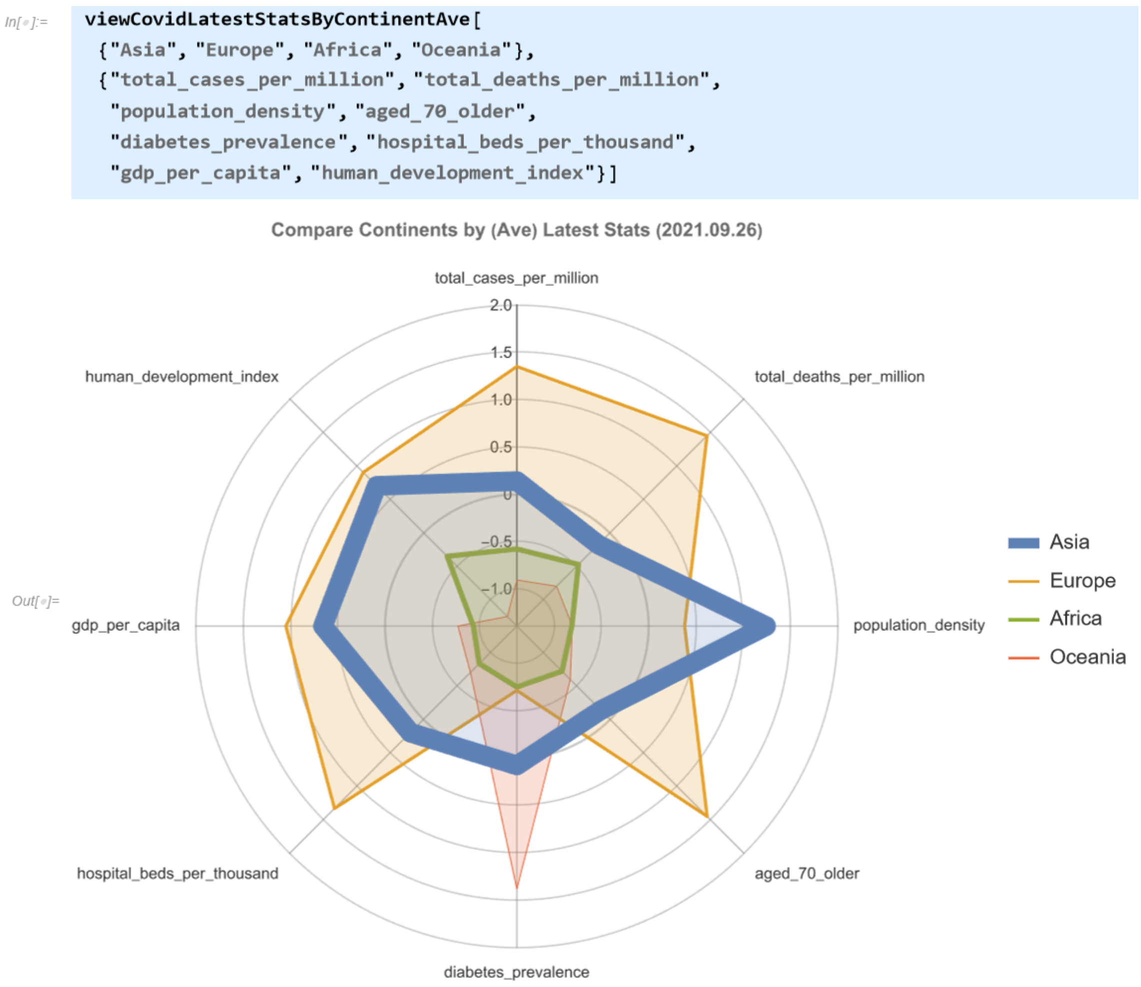
[4] van der Maaten, Postma & van den Herik (2007). "Dimensionality Reduction: A Comparative Review" *Journal of Machine Learning Research* vol. 10, no. 1.

[5] Sorzano, Vargas, Pascual-Montano (2014). "A survey of dimensionality reduction techniques" `https://arxiv.org/abs/1403.2877`

[6] Wikipedia (2021). "Dimensionality reduction" `https://en.wikipedia.org/wiki/Dimensionality_reduction`

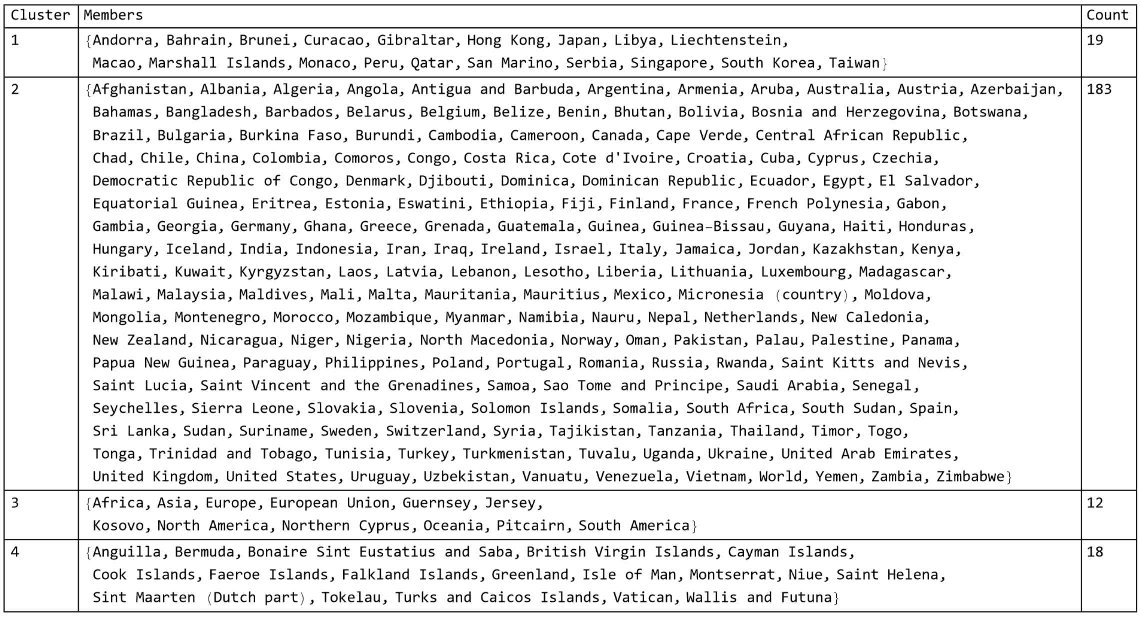
[7] Biau, G. & Scornet, E. (2015). "A Random Forest Guided Tour" `https://arxiv.org/abs/1511.05741`

[8] Wikipedia (2021). "Random forest" `https://en.wikipedia.org/wiki/Random_forest`



```
In[∘]:= viewCovidLatestStatsByContinentAve[
         {"Asia", "Europe", "Africa", "Oceania"},
         {"total_cases_per_million", "total_deaths_per_million",
          "population_density", "aged_70_older",
          "diabetes_prevalence", "hospital_beds_per_thousand",
          "gdp_per_capita", "human_development_index"}]
```

Figure 9: **viewCovidLatestStatsByContinentAve** depicting just 4 countinents and 8 attributes



Cluster	Members	Count
1	{Andorra, Bahrain, Brunei, Curacao, Gibraltar, Hong Kong, Japan, Libya, Liechtenstein, Macao, Marshall Islands, Monaco, Peru, Qatar, San Marino, Serbia, Singapore, South Korea, Taiwan}	19
2	{Afghanistan, Albania, Algeria, Angola, Antigua and Barbuda, Argentina, Armenia, Aruba, Australia, Austria, Azerbaijan, Bahamas, Bangladesh, Barbados, Belarus, Belgium, Belize, Benin, Bhutan, Bolivia, Bosnia and Herzegovina, Botswana, Brazil, Bulgaria, Burkina Faso, Burundi, Cambodia, Cameroon, Canada, Cape Verde, Central African Republic, Chad, Chile, China, Colombia, Comoros, Congo, Costa Rica, Cote d'Ivoire, Croatia, Cuba, Cyprus, Czechia, Democratic Republic of Congo, Denmark, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Eritrea, Estonia, Eswatini, Ethiopia, Fiji, Finland, France, French Polynesia, Gabon, Gambia, Georgia, Germany, Ghana, Greece, Grenada, Guatemala, Guinea, Guinea-Bissau, Guyana, Haiti, Honduras, Hungary, Iceland, India, Indonesia, Iran, Iraq, Ireland, Israel, Italy, Jamaica, Jordan, Kazakhstan, Kenya, Kiribati, Kuwait, Kyrgyzstan, Laos, Latvia, Lebanon, Lesotho, Liberia, Lithuania, Luxembourg, Madagascar, Malawi, Malaysia, Maldives, Mali, Malta, Mauritania, Mauritius, Mexico, Micronesia (country), Moldova, Mongolia, Montenegro, Morocco, Mozambique, Myanmar, Namibia, Nauru, Nepal, Netherlands, New Caledonia, New Zealand, Nicaragua, Niger, Nigeria, North Macedonia, Norway, Oman, Pakistan, Palau, Palestine, Panama, Papua New Guinea, Paraguay, Philippines, Poland, Portugal, Romania, Russia, Rwanda, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Samoa, Sao Tome and Principe, Saudi Arabia, Senegal, Seychelles, Sierra Leone, Slovakia, Slovenia, Solomon Islands, Somalia, South Africa, South Sudan, Spain, Sri Lanka, Sudan, Suriname, Sweden, Switzerland, Syria, Tajikistan, Tanzania, Thailand, Timor, Togo, Tonga, Trinidad and Tobago, Tunisia, Turkey, Turkmenistan, Tuvalu, Uganda, Ukraine, United Arab Emirates, United Kingdom, United States, Uruguay, Uzbekistan, Vanuatu, Venezuela, Vietnam, World, Yemen, Zambia, Zimbabwe}	183
3	{Africa, Asia, Europe, European Union, Guernsey, Jersey, Kosovo, North America, Northern Cyprus, Oceania, Pitcairn, South America}	12
4	{Anguilla, Bermuda, Bonaire Sint Eustatius and Saba, British Virgin Islands, Cayman Islands, Cook Islands, Faeroe Islands, Falkland Islands, Greenland, Isle of Man, Montserrat, Niue, Saint Helena, Sint Maarten (Dutch part), Tokelau, Turks and Caicos Islands, Vatican, Wallis and Futuna}	18

Figure 10: All 4 clusters are displayed in the table. Both continents and countries are included in the clusters.

```
In[∘]:= viewClustersByCentroidAttributes[]
```

Out[∘]=

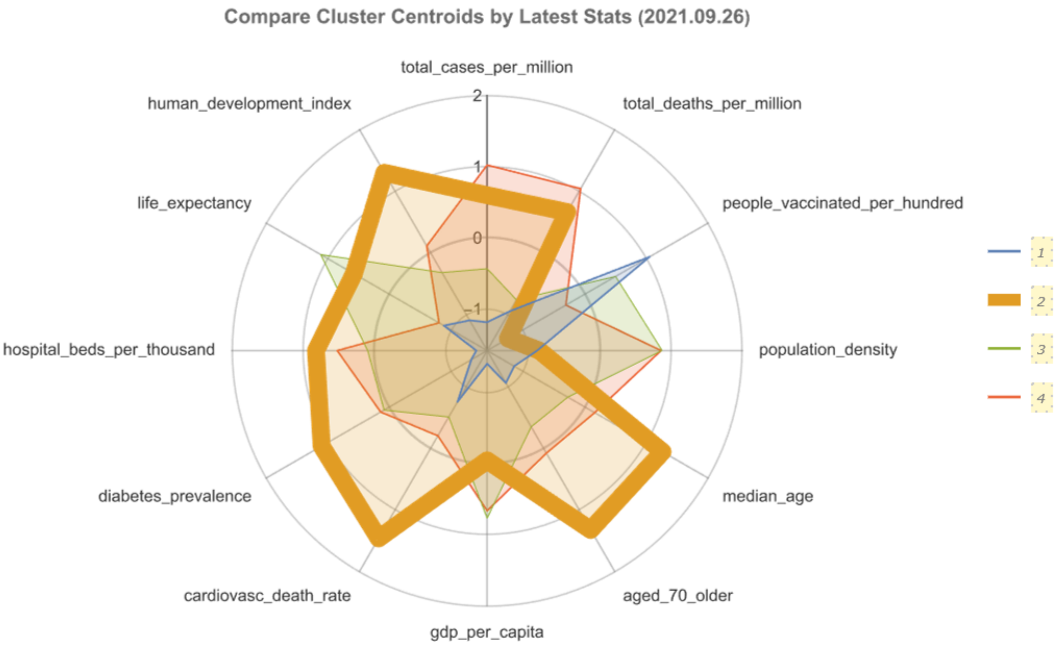


Figure 11: viewClustersByCentroidAttributes

```
In[∘]:= viewReducedDimensionRepresentation[]
```

Out[∘]=

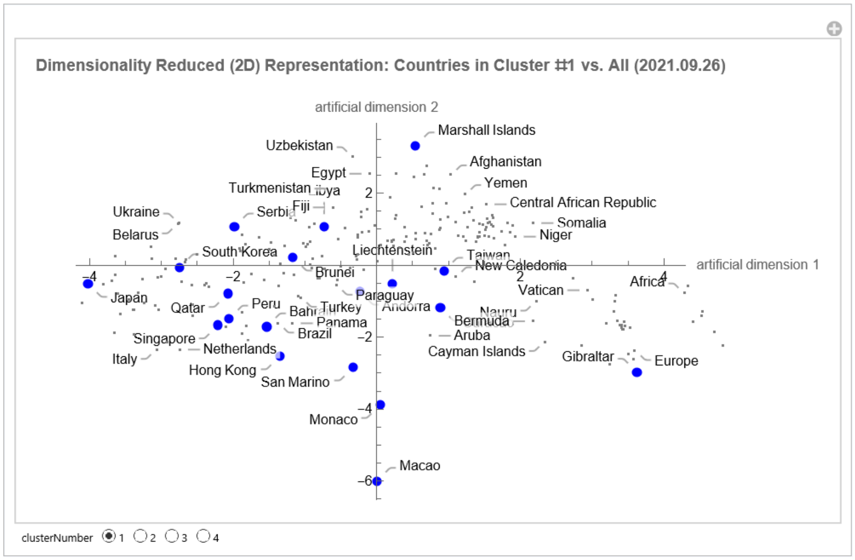
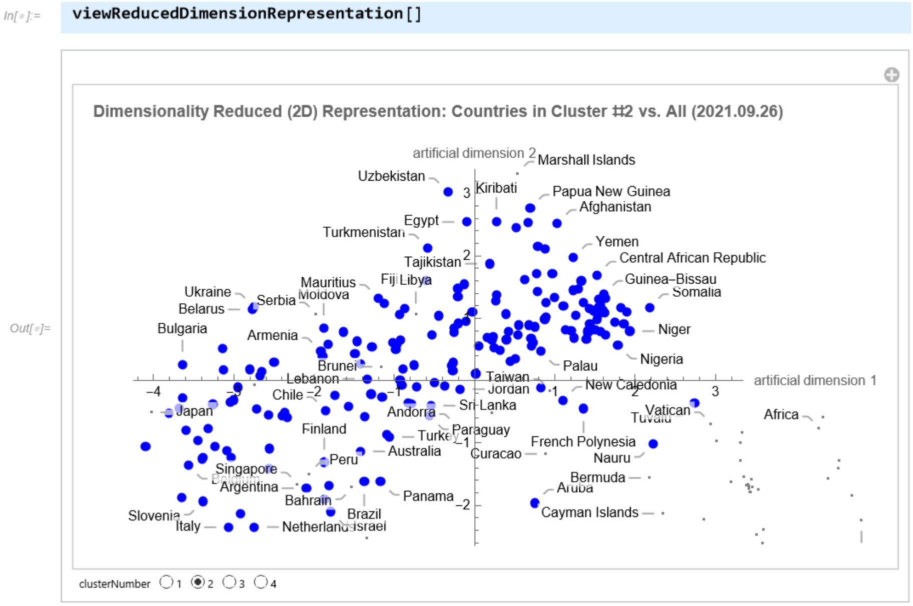


Figure 12: viewReducedDimensionRepresentation1.



In[∘]:=

```
viewReducedDimensionRepresentation[]
```

Out[∘]=

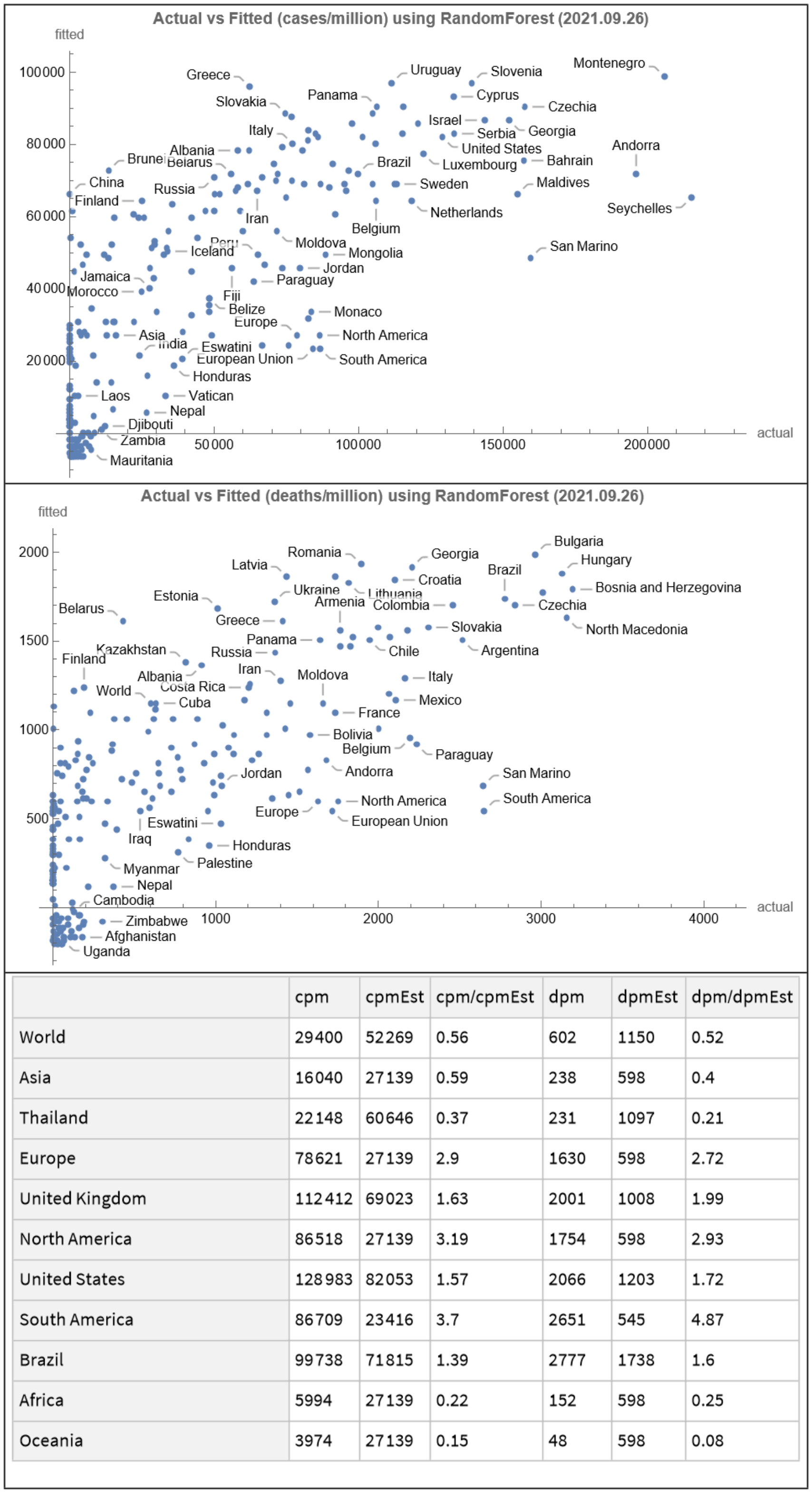
Dimensionality Reduced (2D) Representation: Countries in Cluster #2 vs. All (2021.09.26)

clusterNumber ○1 ●2 ○3 ○4

Figure 13: All 11 dimensions have been merged into two artificial dimensions, displaying the most accurate classification.

```
In[∘]:= viewCovidIncidenceActualVsBaselineFit[]
```

Out[∘]=



	cpm	cpmEst	cpm/cpmEst	dpm	dpmEst	dpm/dpmEst
World	29400	52269	0.56	602	1150	0.52
Asia	16040	27139	0.59	238	598	0.4
Thailand	22148	60646	0.37	231	1097	0.21
Europe	78621	27139	2.9	1630	598	2.72
United Kingdom	112412	69023	1.63	2001	1008	1.99
North America	86518	27139	3.19	1754	598	2.93
United States	128983	82053	1.57	2066	1203	1.72
South America	86709	23416	3.7	2651	545	4.87
Brazil	99738	71815	1.39	2777	1738	1.6
Africa	5994	27139	0.22	152	598	0.25
Oceania	3974	27139	0.15	48	598	0.08

Figure 14: The graph above shows how well countries/continents are performing. The y axes represent the predicted numbers, while the X axes represent the actual figures.