

Anthony Asilo aka pillared

<https://github.com/pillared/COVID-COMPARISON>

COVID-COMPARISON

Introduction

- **Outline**

This project highlights some major outbreaks / diseases since the year 2000 and provides a means to visualize and compare between eachother.

- **Purpose**

My hypothesis was that COVID 19 was not as detrimental to the population in comparison to other diseases spread since the year 2000 in terms of deaths... So I found data to either prove or disprove it.

- **Contents**

`COVID_COMPARISON.ipynb` contains source code for project.

`CLEAN` contains `.csv` files with data.

`Figures` contains all visualizations saved as `.pngs`

Introduction (Cont.)

- Analyzed the following outbreaks either globally or in the US.

COVID-19 Pnuemonia HIV/AIDS EBOLA SARS

- Findings

- i. In the USA, COVID-19 is a bigger problem than Pnuemonia or any other outbreak.
- ii. Globally, Pneumonia and HIV/AIDS is a bigger problem than COVID-19 and other outbreaks.

- References

https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_daily_reports

<https://www.kaggle.com/imdevskp/sars-outbreak-2003-complete-dataset>

<https://www.kaggle.com/imdevskp/ebola-outbreak-20142016-complete-dataset>

<https://github.com/ckeating19/HIV-AIDS-Deaths/blob/master/hiv-death-rates.csv>

<https://data.cdc.gov/Health-Statistics/Deaths-from-Pneumonia-and-Influenza-P-I-and-all-de/pp7x-dyj2/data>

<https://ourworldindata.org/pneumonia>

<https://data.worldbank.org/indicator/SP.POP.TOTL>

Cleaned using excel

Dataset previews

SARS WORLD

	Date	Country	Cumulative number of case(s)	Number of deaths	Number recovered
0	2003-03-17	Germany	1	0	0
1	2003-03-17	Canada	8	2	0
2	2003-03-17	Singapore	20	0	0
3	2003-03-17	Hong Kong SAR, China	95	1	0
4	2003-03-17	Switzerland	2	0	0
...
2533	2003-07-11	Switzerland	1	0	1
2534	2003-07-11	Thailand	9	2	7
2535	2003-07-11	United Kingdom	4	0	4
2536	2003-07-11	United States	75	0	67
2537	2003-07-11	Viet Nam	63	5	58

2538 rows × 5 columns

COVID_USA

	i	Province_State	Confirmed	Deaths	FIPS	Mortality_Rate	Date
0	0	Alabama	20777	692	1	3.375610	6/7/20
1	1	Alaska	543	10	2	1.841621	6/7/20
2	2	Arizona	26989	1051	4	3.894179	6/7/20
3	3	Arkansas	9426	154	5	1.633779	6/7/20
4	4	California	130615	4632	6	3.546300	6/7/20
...
11623	11623	Virginia	11594	411	51	3.544937	4/24/20
11624	11624	Washington	12977	722	53	5.563690	4/24/20
11625	11625	West Virginia	1010	32	54	3.168317	4/24/20
11626	11626	Wisconsin	5356	263	55	4.910381	4/24/20
11627	11627	Wyoming	473	7	56	1.479915	4/24/20

11628 rows × 7 columns

WORLD_POPULATIONS

	Country Name	Country Code	2000	2001	2002	2003	2004	2005	2006	2007	...	2010	2011	2012	2013
0	Aruba	ABW	90853	92898	94992	97017	98737	100031	100834	101222	...	101669	102046	102560	103159
1	Afghanistan	AFG	20779953	21606988	22600770	23680871	24726684	25654277	26433049	27100536	...	29185507	30117413	31161376	32269589
2	Angola	AGO	16395473	16945753	17519417	18121479	18758145	19433602	20149901	20905363	...	23356246	24220661	25107931	26015780
3	Albania	ALB	3089027	3060173	3051010	3039616	3026939	3011487	2992547	2970017	...	2913021	2905195	2900401	2895092
4	Andorra	AND	65390	67341	70049	73182	76244	78867	80993	82684	...	84449	83747	82427	80774
...
258	Kosovo	XKX	1700000	1701154	1702310	1703466	1704622	1705780	1719536	1733404	...	1775680	1791000	1807106	1818117
259	Yemen, Rep.	YEM	17409072	17918373	18443691	18985000	19540098	20107409	20687646	21282515	...	23154855	23807588	24473178	25147109
260	South Africa	ZAF	44967708	45571274	46150913	46719196	47291610	47880601	48489459	49119759	...	51216964	52004172	52834005	53689236
261	Zambia	ZMB	10415944	10692193	10971698	11256743	11550642	11856247	12173514	12502958	...	13605984	14023193	14465121	14926504
262	Zimbabwe	ZWE	11881477	11923914	11954290	11982224	12019912	12076699	12155491	12255922	...	12697723	12894316	13115131	13350356

263 rows × 22 columns

HIV_WORLD

	Unnamed: 0	Entity	Code	Year	Deaths - HIV/AIDS - Sex: Both - Age: Age-standardized (Rate) (per 100,000 people)
0	10	Afghanistan	AFG	2000	0.692406
1	11	Afghanistan	AFG	2001	0.688246
2	12	Afghanistan	AFG	2002	0.764333
3	13	Afghanistan	AFG	2003	0.727879
4	14	Afghanistan	AFG	2004	0.774852
...
4135	6463	Zimbabwe	ZWE	2013	278.365351
4136	6464	Zimbabwe	ZWE	2014	208.990701
4137	6465	Zimbabwe	ZWE	2015	173.987398
4138	6466	Zimbabwe	ZWE	2016	150.980630
4139	6467	Zimbabwe	ZWE	2017	130.353613

4140 rows × 5 columns

EBOLA_WORLD

	Country	Date	Cumulative no. of confirmed, probable and suspected cases	Cumulative no. of confirmed, probable and suspected deaths
0	Nigeria	8/29/14	19.0	7
1	Sierra Leone	8/29/14	1026.0	422
2	Guinea	8/29/14	648.0	430
3	Liberia	8/29/14	1378.0	694
4	Senegal	9/5/14	1.0	0
...
2374	Mali	3/23/16	8.0	6
2375	Nigeria	3/23/16	20.0	8
2376	Guinea	3/23/16	3804.0	2536
2377	Sierra Leone	3/23/16	14122.0	3955
2378	Liberia	3/23/16	10666.0	4806

2379 rows × 4 columns

PNUEMONIA_USA

	Year	Week	Percent of Deaths Due to Pneumonia and Influenza	Expected	Threshold	All Deaths	Pneumonia Deaths	Influenza Deaths
0	2013	40	6.617957	6.36132	6.77011	47492	3140	3
1	2013	41	6.652714	6.45326	6.86185	47304	3135	12
2	2013	42	6.779127	6.55439	6.96279	47602	3216	11
3	2013	43	6.622544	6.66322	7.07142	47746	3151	11
4	2013	44	6.730631	6.77811	7.18613	48777	3271	12
...
359	2020	34	9.486959	5.12719	5.46286	56857	5385	9
360	2020	35	8.912452	5.13611	5.47178	53285	4738	11
361	2020	36	8.189548	5.15856	5.49423	48220	3942	7
362	2020	37	7.088050	5.19422	5.52989	39489	2795	4
363	2020	38	5.375106	5.24261	5.57828	19981	1074	0

364 rows × 8 columns

COVID_WORLD

	Province/State	Country/Region	Lat	Long	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	...	11/16/20	11/17/20	11/18/20	11/19/20	11/20/20
0	NAN	Afghanistan	33.939110	67.709953	0	0	0	0	0	0	...	1626	1638	1645	1650	1650
1	NAN	Albania	41.153300	20.168300	0	0	0	0	0	0	...	631	637	646	657	660
2	NAN	Algeria	28.033900	1.659600	0	0	0	0	0	0	...	2168	2186	2206	2224	2240
3	NAN	Andorra	42.506300	1.521800	0	0	0	0	0	0	...	76	76	76	76	77
4	NAN	Angola	-11.202700	17.873900	0	0	0	0	0	0	...	324	328	332	333	335
...
266	NAN	West Bank and Gaza	31.952200	35.233200	0	0	0	0	0	0	...	572	580	589	598	606
267	NAN	Western Sahara	24.215500	-12.885800	0	0	0	0	0	0	...	1	1	1	1	1
268	NAN	Yemen	15.552727	48.516388	0	0	0	0	0	0	...	605	607	607	608	610
269	NAN	Zambia	-13.133897	27.849332	0	0	0	0	0	0	...	353	353	356	356	359
270	NAN	Zimbabwe	-19.015438	29.154857	0	0	0	0	0	0	...	257	260	261	265	269

271 rows × 313 columns

PNEUMONIA_WORLD

	Entity	Code	Year	Deaths - Lower respiratory infections - Sex: Both - Age: Under 5 (Number)	Deaths - Lower respiratory infections - Sex: Both - Age: 5-14 years (Number)	Deaths - Lower respiratory infections - Sex: Both - Age: 15-49 years (Number)	Deaths - Lower respiratory infections - Sex: Both - Age: 50-69 years (Number)	Deaths - Lower respiratory infections - Sex: Both - Age: 70+ years (Number)
0	Afghanistan	AFG	1990	19713.086217	337.582192	422.240838	897.360063	1466.643037
1	Afghanistan	AFG	1991	19174.882442	328.892192	433.069331	907.420453	1481.369513
2	Afghanistan	AFG	1992	19907.997133	355.842407	505.451602	932.847437	1503.142170
3	Afghanistan	AFG	1993	24732.410838	408.801264	594.080120	963.145116	1531.278059
4	Afghanistan	AFG	1994	29035.079800	436.866895	626.908993	989.667690	1563.774252
...
6463	Zimbabwe	ZWE	2013	6152.929327	308.350344	1625.366442	2791.002724	2230.502510
6464	Zimbabwe	ZWE	2014	5951.444115	304.174770	1602.593906	2769.643197	2213.654683
6465	Zimbabwe	ZWE	2015	5760.039745	291.222446	1602.323287	2768.235602	2215.907533
6466	Zimbabwe	ZWE	2016	5575.706127	289.851003	1609.758753	2780.395785	2239.435624
6467	Zimbabwe	ZWE	2017	5398.094153	285.438928	1619.340855	2795.301561	2271.479208

6468 rows × 8 columns

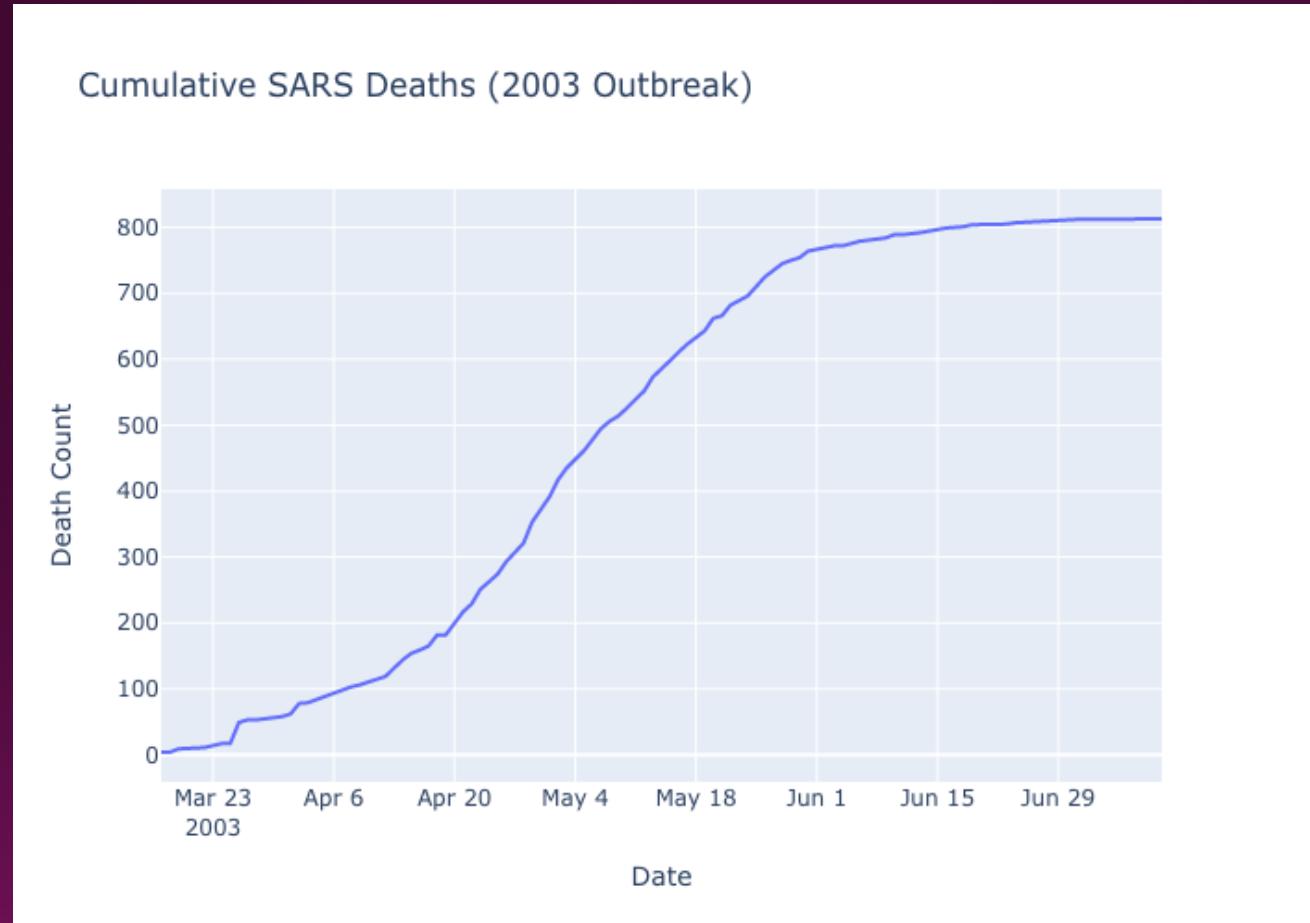
Created in python with pandas and plotly.

Visualizations (by dataset)

SARS_WORLD

Cumulative deaths in the form of a line chart

From 2003-03-17 00:00:00 to 2003-07-11 00:00:00, there have been a total of 45209 deaths in the WORLD from SARS... 116 days 00:00:00

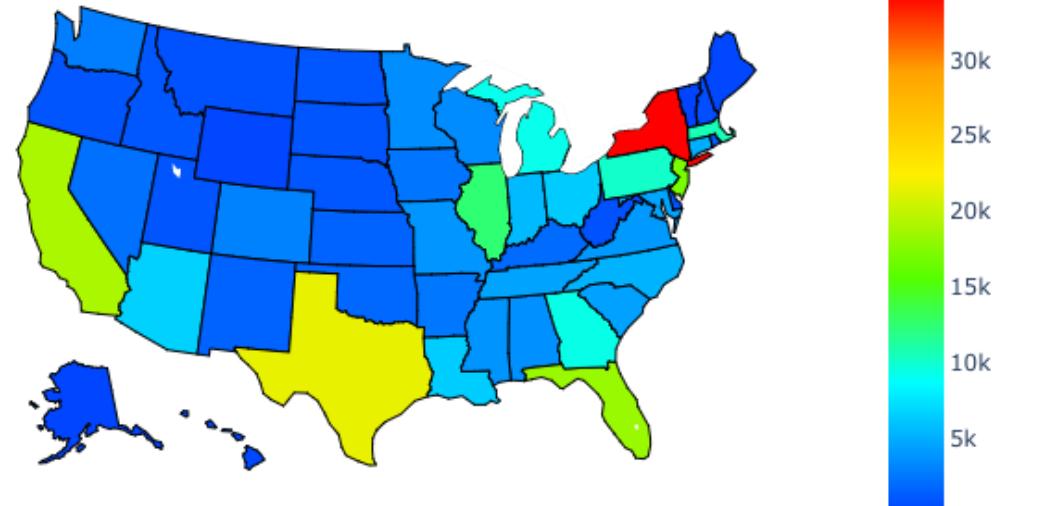


COVID_USA

Choropleth for USA COVID deaths

From 2020-04-12 00:00:00 to 2020-11-25 00:00:00, there have been a total of 261033 deaths in the USA from COVID-19... 227 days 00:00:00

COVID19 Deaths USA



WORLD_POPULATIONS

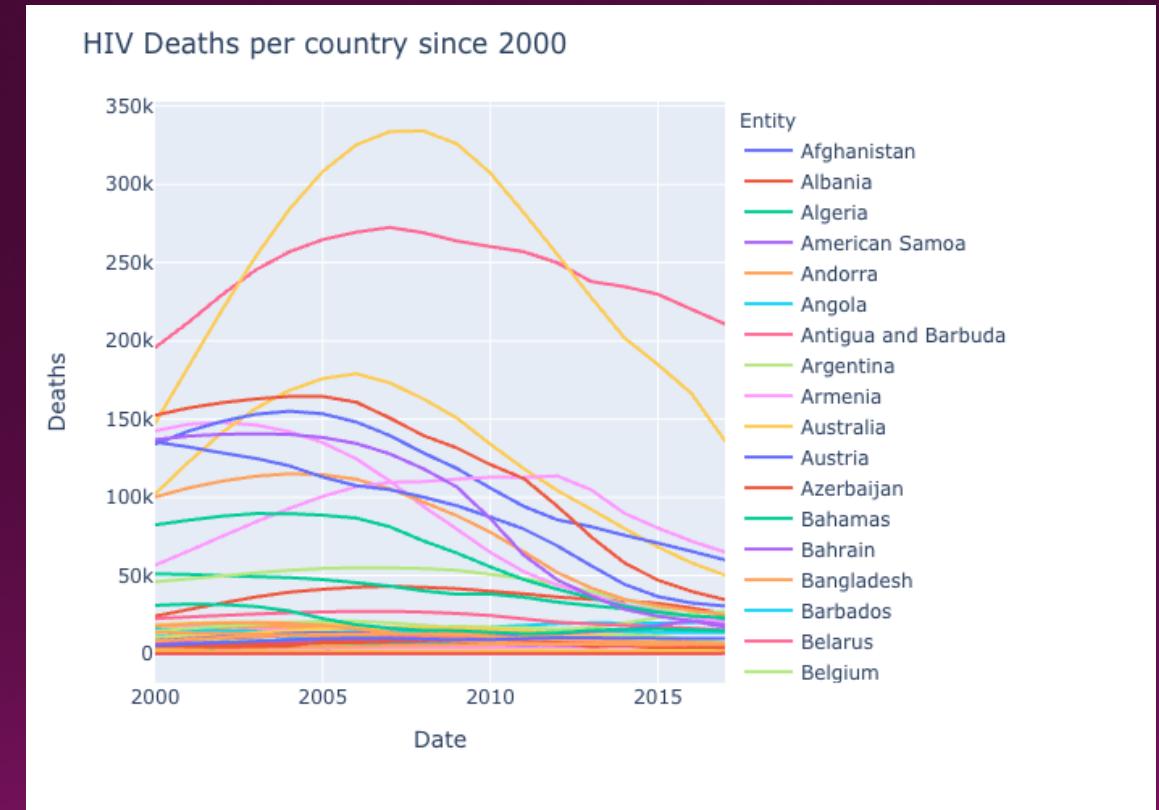
- No visualization needed as it is not necessarily crucial data to present other than the fact that is is needed to convert rates to exact numbers, like in the case of the HIV dataset.

HIV_WORLD

Choropleth



Line

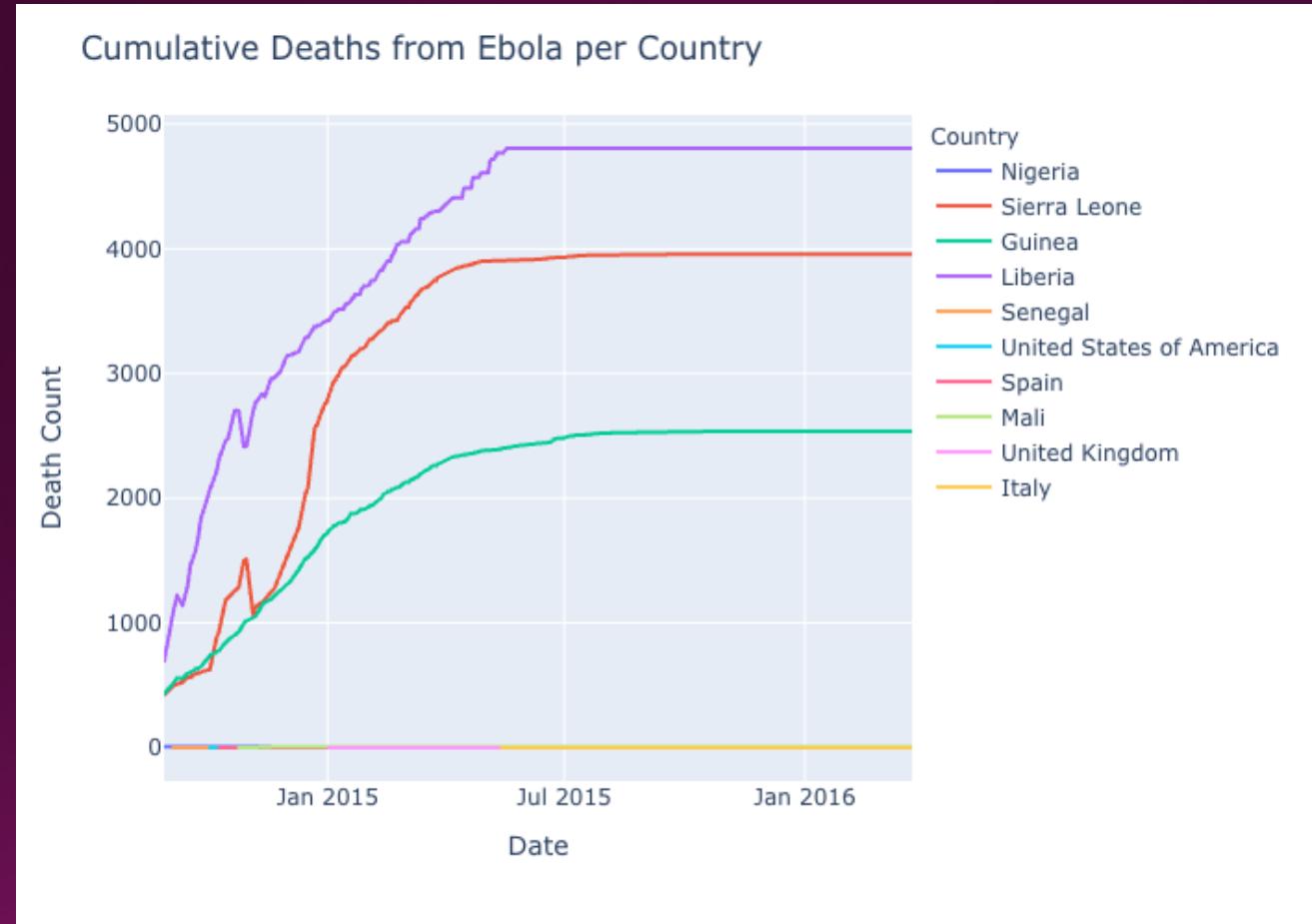


From 2020-04-12 00:00:00 to 2020-11-25 00:00:00, there have been a total of 261033 deaths in the USA from COVID-19... 227 days 00:00:00

EBOLA_WORLD

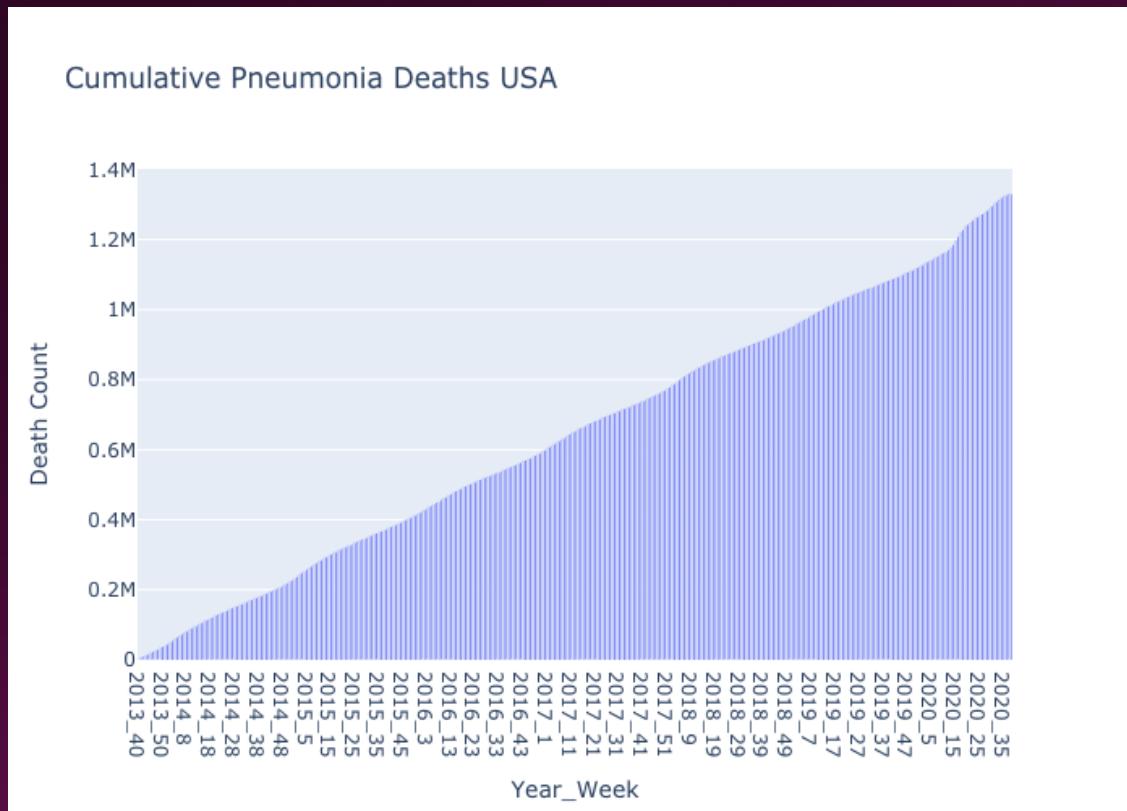
Line Graph for Cumulative Ebola deaths,
mostly in Africa.

From 2014-08-29 00:00:00 to 2016-03-23
00:00:00, there have been a total of
11312 deaths in the WORLD from
EBOLA... 572 days 00:00:00

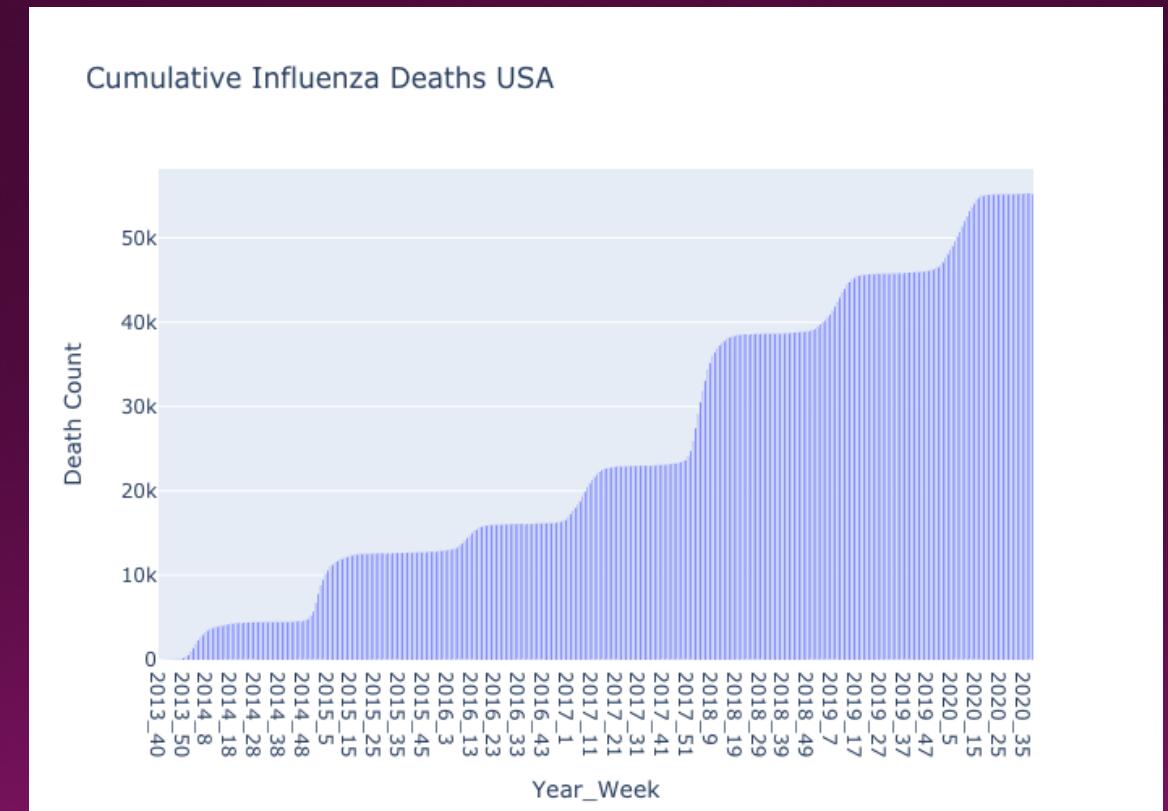


PNEUMONIA_USA (includes Influenza)

Bar Graph



Bar Graph

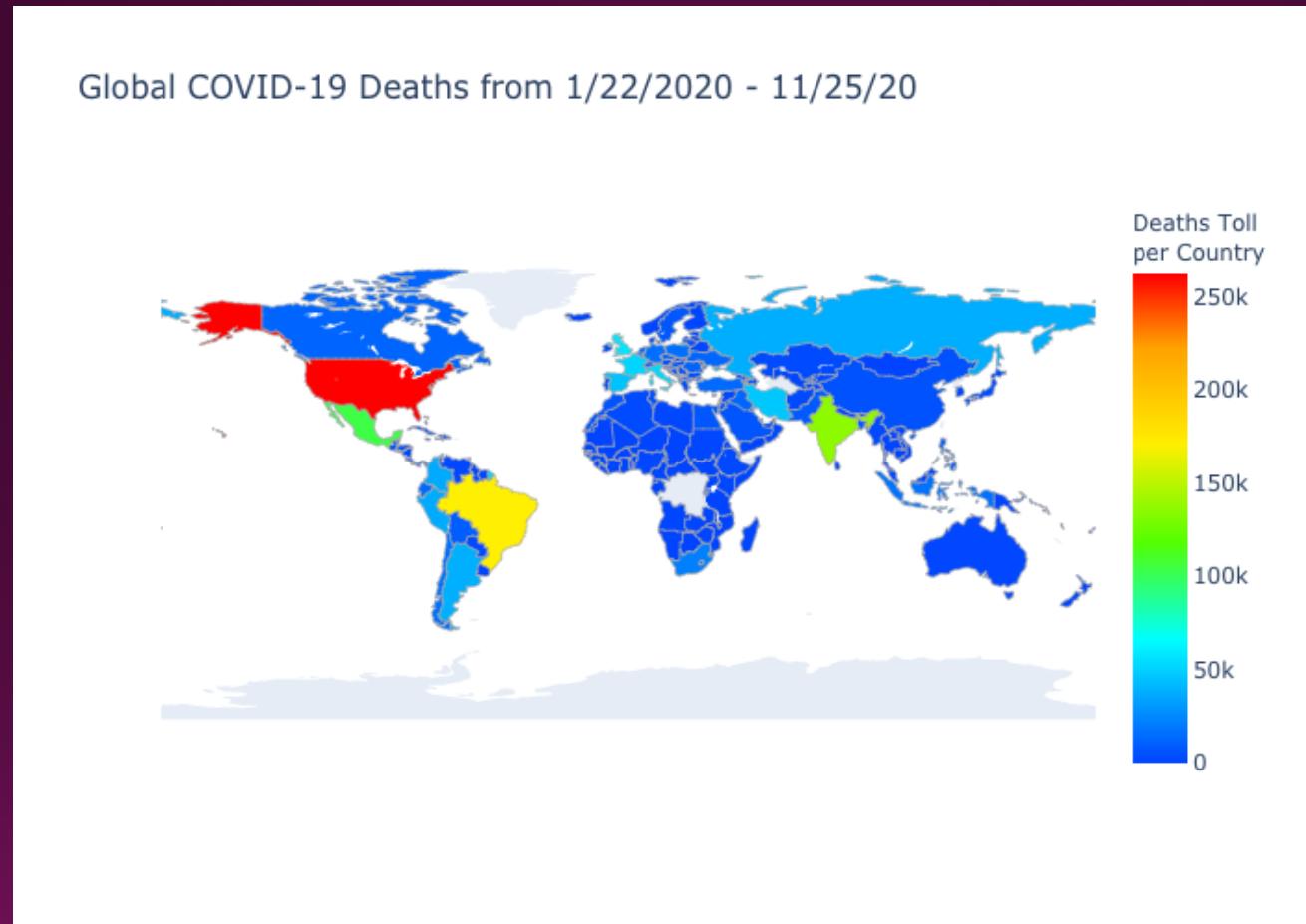


From 2013_40 to 2020_9, there have been a total of 1331647 deaths in the USA from Pneumonia... 2337 days
From 2013_40 to 2020_9, there have been a total of 55259 deaths in the USA from Influenza... 2337 days

COVID_WORLD

From 1/22/20 to 11/25/20, there have been a total of 1421308 deaths in the WORLD from COVID-19... 308 days

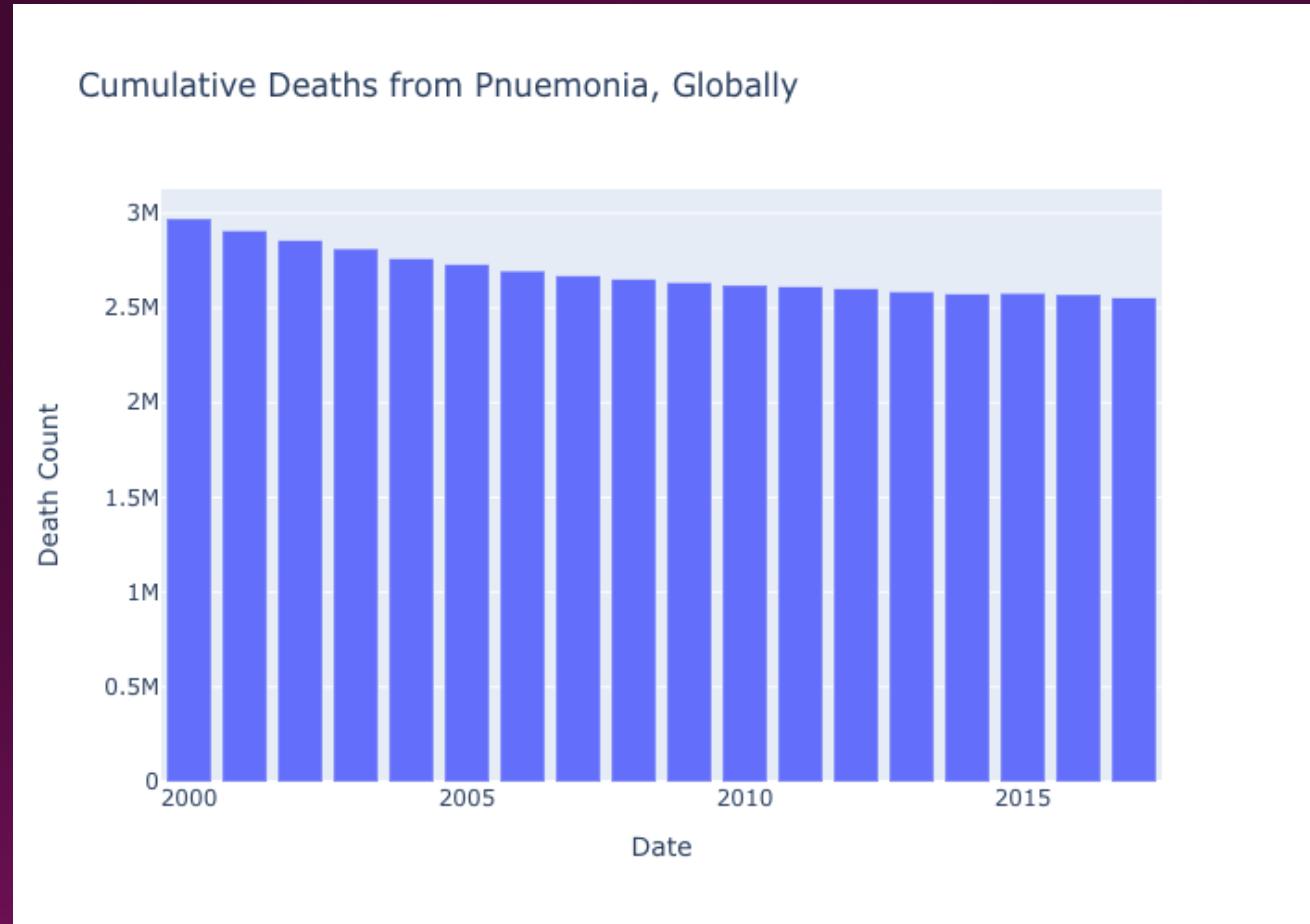
Choropleth Graph



PNEUMONIA_WORLD

From 2000 to 2017, there have been a total of 48378423.41746253 deaths in the WORLD from Pneumonia... 6570 days

Bar Graph



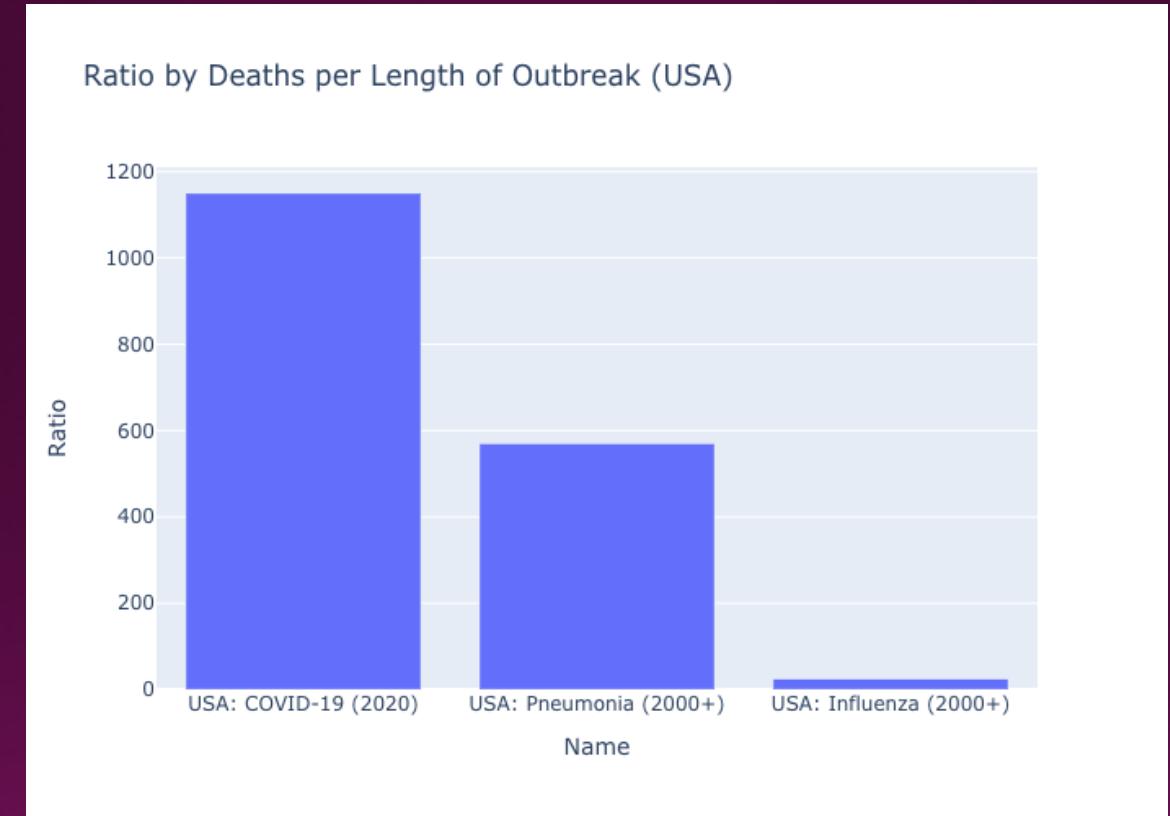
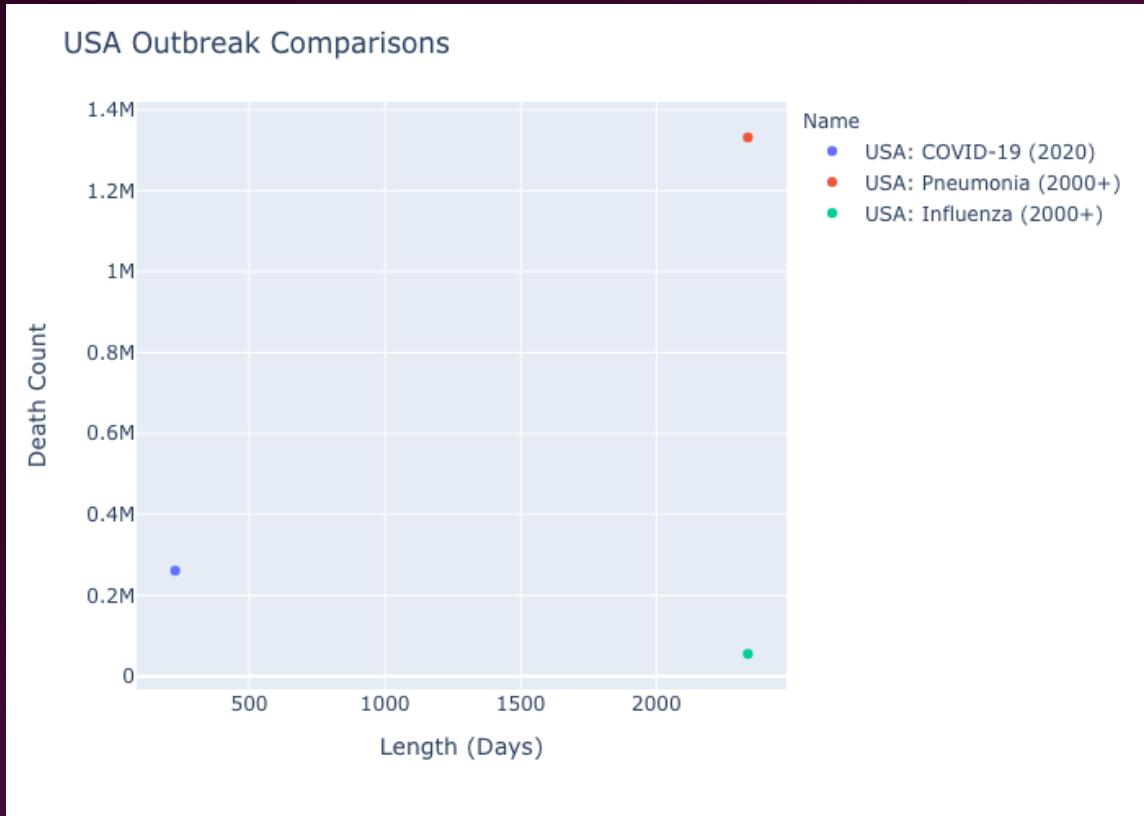
Comparing DataFrames into Visualizations!
Putting it all together...

DFX (DataFrameX)

Once I got all of the Datasets and their Total Deaths as well as Length of outbreak / disease, I divided the deaths by the days to get a ratio to provide a mean of comparison

	Location	Name	Length(Days)	TotalDeaths	Ratio
0	USA	USA: COVID-19 (2020)	227	2.610330e+05	1149.925110
1	GLOBAL	GLOBAL: SARS (2003)	116	4.520900e+04	389.732759
2	GLOBAL	GLOBAL: HIV/AIDS (2000+)	6570	3.343616e+07	5089.218280
3	GLOBAL	GLOBAL: EBOLA (2014)	572	1.131200e+04	19.776224
4	USA	USA: Pneumonia (2000+)	2337	1.331647e+06	569.810441
5	USA	USA: Influenza (2000+)	2337	5.525900e+04	23.645272
6	GLOBAL	GLOBAL: COVID-19 (2020)	308	1.421308e+06	4614.636364
7	GLOBAL	GLOBAL: Pnuemonia (2000+)	6570	4.837842e+07	7363.534767

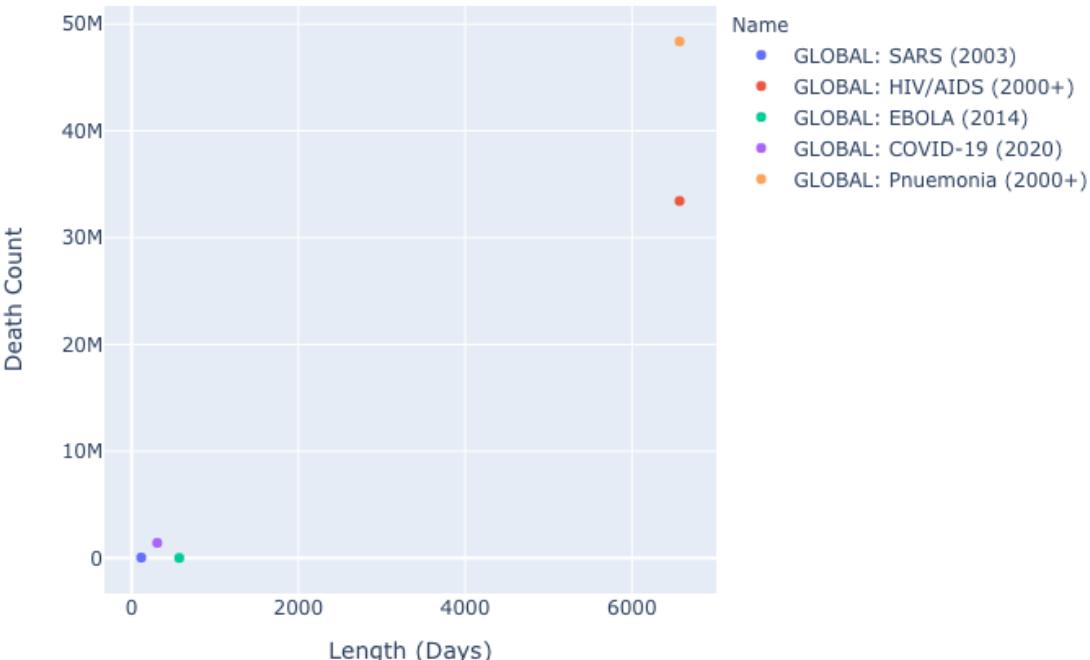
USA



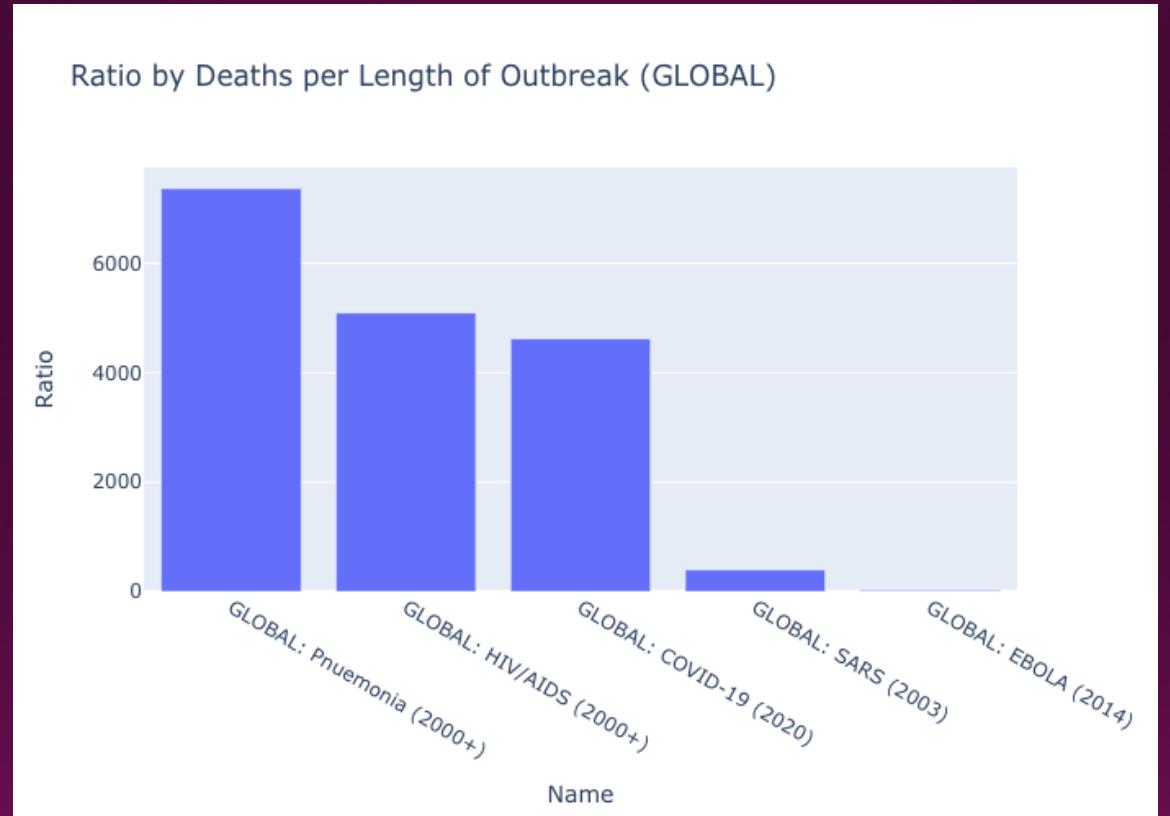
Here you can see that we do a scatter plot of USA outbreaks with the x axis as the length of disease and the y axis as the death toll. We then divide the deaths by days to get a ratio, and graph it out in a bar graph seen above. Note that in the US COVID death to length ratio is greater than pneumonia.

GLOBAL

GLOBAL Outbreak Comparisons

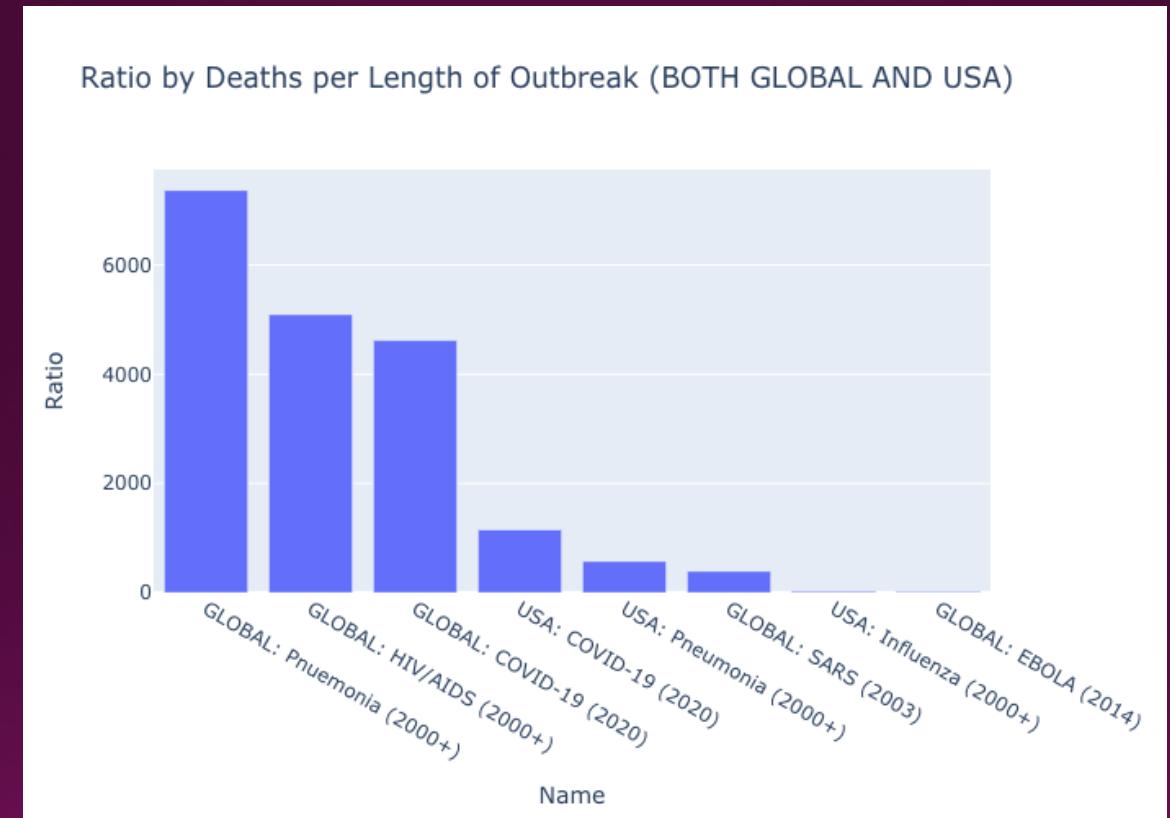
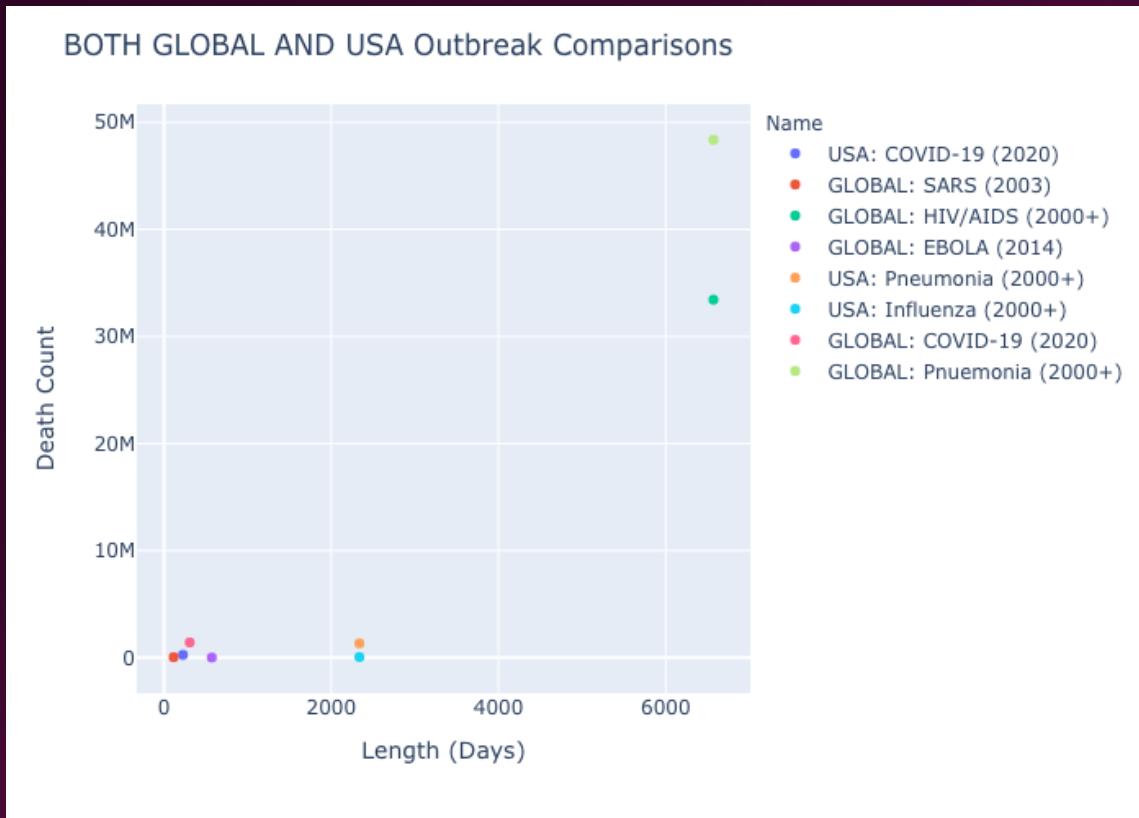


Ratio by Deaths per Length of Outbreak (GLOBAL)



Here you can see that we do a scatter plot of GLOBAL outbreaks with the x axis as the length of disease and the y axis as the death toll. We then divide the deaths by days to get a ratio, and graph it out in a bar graph seen above. Note that GLOBALLY, Pneumonia and HIV/AIDS has a greater death to length ratio compared to COVID

BOTH Global and USA



Here you can see that we do a scatter plot of Both Global and USA outbreaks with the x axis as the length of disease and the y axis as the death toll. We then divide the deaths by days to get a ratio, and graph it out in a bar graph seen above. Note that for both, Pneumonia and HIV/AIDS have a greater death to length ratio compared to COVID-19, but USA COVID is worse than GLOBAL SARS or EBOLA

Findings

- In the USA, COVID is more deadly than pneumonia,
- However, Globally, Pneumonia is deadlier than COVID.

Future Improvements

- I would like to add a tread line to see, if COVID was as long a length as Pneumonia (The dataset is from 2000-2017), if it would be just as deadly or if not more. I believe it would be more but I must continue my research