Accounting for Mortality

Death Rates in the Age of COVID-19

Team Members

Julie Kirkpatrick

Kyle Tomlinson

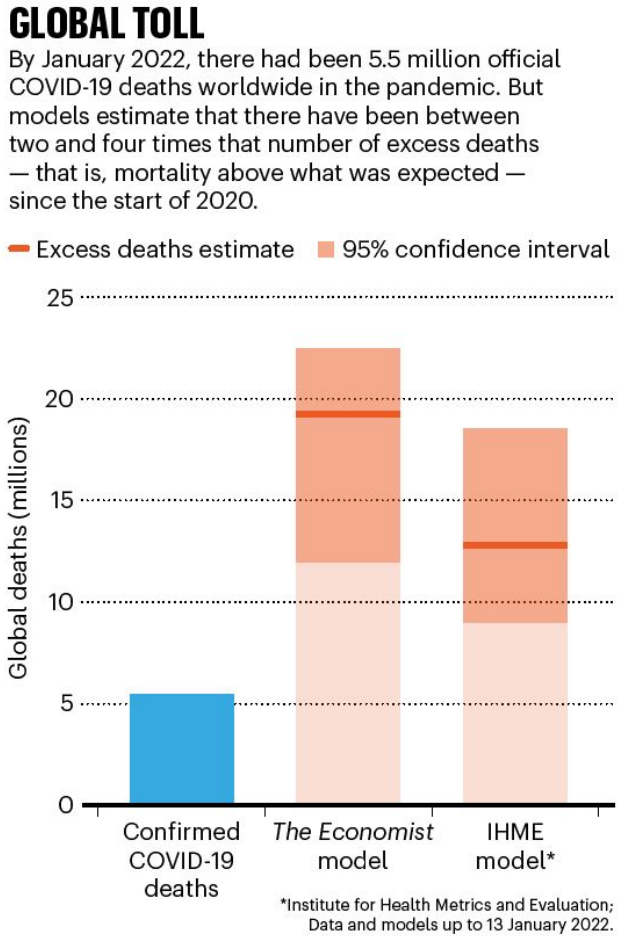
Alex Melnick

Description

An investigation into the relationship between mortality rates and COVID-19.

What are the challenges involved in extrapolating data from excess mortality rates during the COVID-19 pandemic?

What are the secondary effects of the pandemic on mortality rates?



Graph by David Adam. From "The Effort to Count the Pandemic's Global Death Toll" Nature, Vol 601, 20 January 2022.

Prior Work

- Substantial research has been done concerning the pandemic's effect on mortality rates, particularly excess mortality
  - Excess mortality: # of deaths occurring above the predicted mortality rates for a given population
- Majority of prior work focuses on either…
  - Total excess mortality
  - Excess mortality directly attributable to COVID-19
- Our project will replicate the basic findings of this prior work, then seek to investigate the changes in mortality rates from non-COVID causes during the pandemic

Datasets

- Human Mortality Database (HMD)
  - Accessed at: https://www.mortality.org/
  - Stored on Alex's machine locally and on Github
  - Mortality data by week and age group for 38 countries

3	CountryCc	Year	Week	Sex	D0_14	D15_64	D65_74	D75_84	D85p	DTotal	R0_14	R15_64	R65_74	R75_84	R85p	RTotal	Split	SplitSex	Forecast
108164	TWN	2020	51	f	6	271	227	415	538	1457	0.00022	0.00167	0.00985	0.03678	0.12298	0.00636	0	0	1
108165	TWN	2020	51	b	13	881	658	868	1087	3507	0.00023	0.00271	0.01509	0.0436	0.142	0.00772	0	0	1
108166	TWN	2020	52	m	10	633	419	451	556	2069	0.00033	0.00389	0.02039	0.05228	0.1695	0.00919	0	0	1
108167	TWN	2020	52	f	8	261	199	392	544	1404	0.00029	0.00161	0.00864	0.03474	0.12435	0.00613	0	0	1
108168	TWN	2020	52	b	18	894	618	843	1100	3473	0.00031	0.00275	0.01418	0.04234	0.1437	0.00765	0	0	1
108169	USA	2015	2	m	343.357	8912.98	6110.07	7516.47	8012.73	30895.6	0.00057	0.00439	0.02474	0.06495	0.19274	0.01018	1	1	0
108170	USA	2015	2	f	267.905	5435.76	4601.93	7281.53	13399.3	30986.4	0.00047	0.00266	0.01637	0.0479	0.17088	0.0099	1	1	0
108171	USA	2015	2	b	611.262	14348.7	10712	14798	21412	61882	0.00052	0.00352	0.02029	0.05527	0.17845	0.01004	1	0	0
108172	USA	2015	3	m	353.55	8825.92	6018.8	7413.36	7961.83	30573.5	0.00059	0.00434	0.02437	0.06406	0.19152	0.01007	1	1	0
108173	USA	2015	3	f	275.858	5382.67	4533.2	7181.64	13314.2	30687.5	0.00048	0.00263	0.01613	0.04724	0.16979	0.0098	1	1	0
108174	USA	2015	3	b	629.408	14208.6	10552	14595	21276	61261	0.00054	0.00349	0.01998	0.05451	0.17732	0.00994	1	0	0
108175	USA	2015	4	m	349.764	8487.22	5743.87	7214.76	7550.94	29346.6	0.00058	0.00418	0.02326	0.06234	0.18163	0.00967	1	1	0
108176	USA	2015	4	f	272.904	5176.11	4326.13	6989.24	12627.1	29391.4	0.00048	0.00253	0.01539	0.04598	0.16103	0.00939	1	1	0
108177	USA	2015	4	b	622.668	13663.3	10070	14204	20178	58738	0.00053	0.00335	0.01907	0.05305	0.16817	0.00953	1	0	0
108178	USA	2015	5	m	345.978	8298.85	5767.83	6986.19	7311.44	28710.3	0.00058	0.00408	0.02336	0.06037	0.17587	0.00946	1	1	0
108179	USA	2015	5	f	269.95	5061.22	4344.17	6767.81	12226.6	28669.7	0.00047	0.00248	0.01546	0.04452	0.15592	0.00916	1	1	0
108180	USA	2015	5	b	615.928	13360.1	10112	13754	19538	57380	0.00052	0.00328	0.01915	0.05137	0.16283	0.00931	1	0	0
108181	USA	2015	6	m	325.301	8596.27	5747.87	6967.39	7177.85	28814.7	0.00054	0.00423	0.02327	0.06021	0.17266	0.00949	1	1	0

Datasets

- COVID-19 Dataset
  - Accessed at: https://ourworldindata.org/explorers/coronavirus-data-explorer
  - Stored on Alex's machine locally and on Github
  - Global COVID Statistics by country and date

1	iso_code	continent	location	date	total_case	new_case	new_case	total_deat	new_deat	new_deat	total_case	new_case	new_case	total_deat	new_deat r
150818	GBR	Europe	United Kingdom	1/26/2022	16189420	102076	92151	154831	346	262.143	237356.8	1496.559	1351.047	2270.012	5.073
150819	GBR	Europe	United Kingdom	1/27/2022	16286017	96597	90683.14	155169	338	263.286	238773	1416.231	1329.526	2274.968	4.955
150820	GBR	Europe	United Kingdom	1/28/2022	16374927	88910	89662.29	155447	278	261.571	240076.5	1303.53	1314.559	2279.044	4.076
150821	GBR	Europe	United Kingdom	1/29/2022	16447070	72143	89192.86	155743	296	261.429	241134.2	1057.705	1307.677	2283.384	4.34
150822	GBR	Europe	United Kingdom	1/30/2022	16509469	62399	87421.29	155828	85	262.857	242049.1	914.846	1281.703	2284.63	1.246
150823	GBR	Europe	United Kingdom	1/31/2022	17357945			155885	57	263.286	254488.8			2285.465	0.836
150824	GBR	Europe	United Kingdom	2/1/2022	17470812	112867	76427.43	157008			256143.5	1654.769	1120.52	2301.93	
150825	GBR	Europe	United Kingdom	2/2/2022	17558231	87419	74333.57	157542	534	226.857	257425.2	1281.67	1089.821	2309.759	7.829
150826	GBR	Europe	United Kingdom	2/3/2022	17651284	93053	73827.29	157865	323	224.714	258789.5	1364.271	1082.399	2314.495	4.736
150827	GBR	Europe	United Kingdom	2/4/2022	17733624	82340	72888.71	158119	254	221.286	259996.7	1207.205	1068.638	2318.219	3.724
150828	USA	North America	United States	1/22/2020	1						0.003				
150829	USA	North America	United States	1/23/2020	1	0					0.003	0			
150830	USA	North America	United States	1/24/2020	2	1					0.006	0.003			
150831	USA	North America	United States	1/25/2020	2	0					0.006	0			
150832	USA	North America	United States	1/26/2020	5	3					0.015	0.009			
150833	USA	North America	United States	1/27/2020	5	0					0.015	0			
150834	USA	North America	United States	1/28/2020	5	0	0.571				0.015	0	0.002		
150835	USA	North America	United States	1/29/2020	6	1	0.714				0.018	0.003	0.002		

Proposed Work

- Datasets require little cleaning
- Data preprocessing reveals some mismatching attribute labels and formats
  - mm/dd/yyyy (COVID) vs. year/week(1-53) (Mortality)
  - Country ID codes
- COVID dataset missing many values
  - Iterative, automated scanning to reveal importance of missing values and potential remedies
- Datasets to be integrated
  - Separately by country
  - Together for global scale (limited to 38 countries contained in HMD)

List of Tools

- Python/Pandas
  - Powerful for data processing
  - Familiarity among team members
- Tableau
  - Visualization
  - Using Info Vis ideas from that course

Evaluation

- Exploratory regression tests
  - Which attributes affect each other
  - Will reveal further question avenues
- Clustering
  - Comparing and Contrasting on groups
- K-nearest neighbors
  - Linear algebra background on team