

# Narrative.rmd

Bryce Farabaugh

12/10/2020

Summary: For this project, I found outdated datasets detailing nuclear weapon information (both arsenal size and nuclear weapons tests) and updated the datasets by adding post-2014 information. I then merged the nuclear weapon datasets into the Correlates of War National Material Capabilities (NMC) dataset which includes information on such variables as military expenditures, military personnel, etc to calculate a “Composite Index of National Capability (CINC)”, a score meant to reflect a state’s military power but which doesn’t account for nuclear weapons (likely the most powerful military deterrent in history.) I updated the data using a tool we didn’t cover in class, “`add_row`” from the Tibble package After updating and merging the datasets, I then made two plots: a line plot showing nuclear arsenals since 1945 as well as a bar plot showing which states possess the most nuclear weapons in 2020. I originally intended to somehow plot CINC as a function of the number of nuclear weapons possessed by nuclear weapon states and draw some conclusions from the resulting information but I don’t yet have the statistical training required to make such an analysis. Nevertheless, I think integrating nuclear weapon information in the Correlates of War database is an interesting and useful contribution and I intend to use the data for further research moving forward.

Challenges: The main challenges for this project were locating the original datasets, finding supplemental information to update the datasets, editing the datasets, merging the datasets, and plotting the relevant information. S

Solutions: In terms of locating the original dataset, I found a page on “Our World in Data” that hosted several datasets with outdated information on nuclear weapons pre-2014. I supplemented this data with the latest information from Hans Kristensen’s “Nuclear Notebook” series, but as the Nuclear Notebook isn’t published every year, some information is “NA.” Editing the datasets was a somewhat arduous task within R that entailed using “`add_row`” for each missing observation. Merging the datasets required ensuring variables were the same across all dataframes and using `left_join` to combine. Plotting the information required first deciding what information I wanted to focus on and then editing the code to get the desired aesthetics.

Results: the results are a “merged” dataframe and CSV that includes all information from the original NMC dataset as well as information about nuclear arsenals and tests, as well as two plots (a line plot and bar plot.)

Citations: The original datasets were pulled from “Our World in Data- Nuclear Weapons”, found at <https://ourworldindata.org/nuclear-weapons>

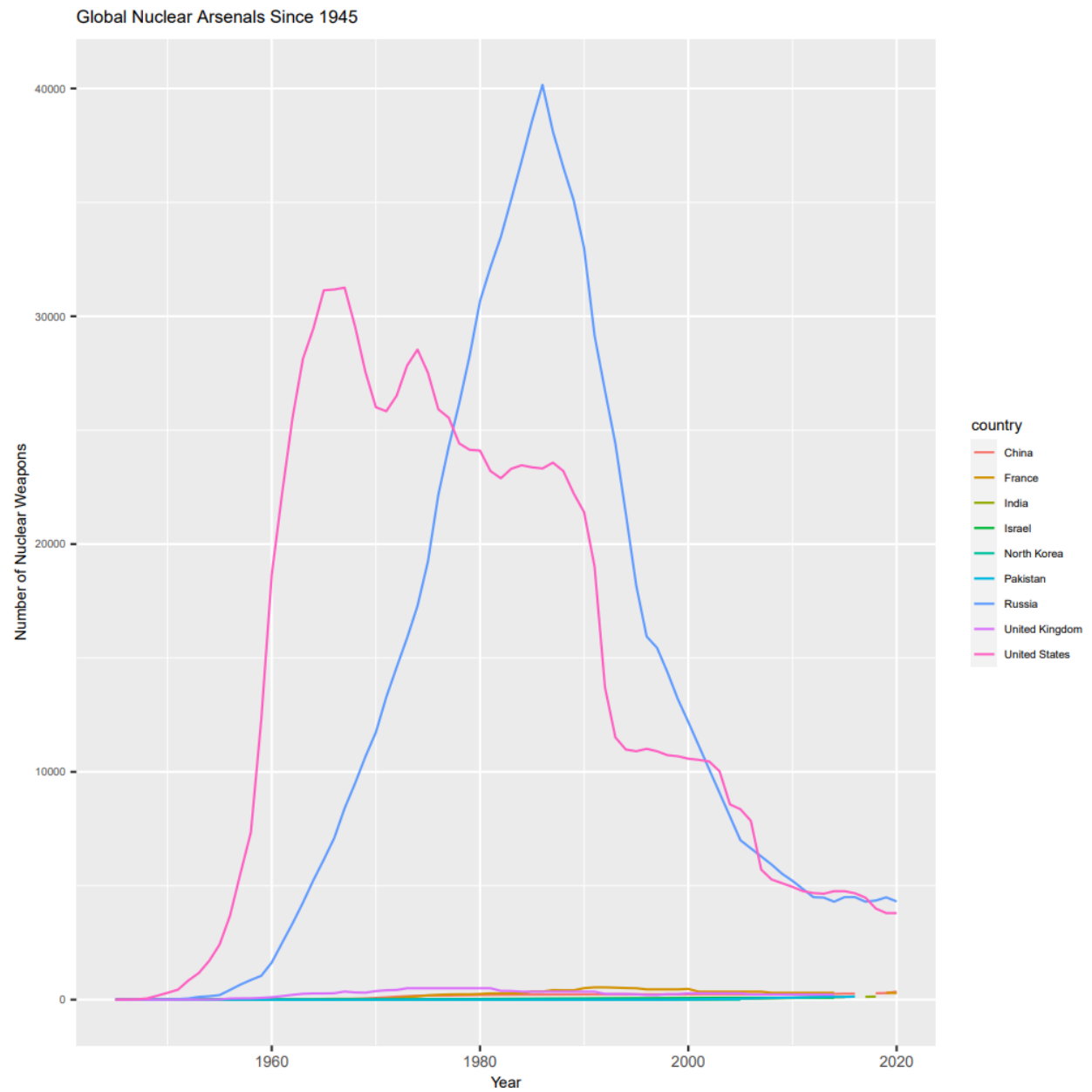
The National Material Capabilities dataset can be found on the Correlates of War website at <https://correlatesofwar.org/data-sets>

The current Nuclear Notebook can be found at <https://fas.org/issues/nuclear-weapons/status-world-nuclear-forces/> and past versions can be found at <http://explore.tandfonline.com/page/pgas/rbul-nuclear-notebooks/>

```
knitr::include_graphics(path = "Screenshot (109).png")
```

R144																	
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
1	stateabb	cocode	year	milex	milper	irst	pec	tpop	upop	cinc	version	country	arsenal	tests			
132	131 USA		2	1946 45133984	3030	60421	1188144	141389	39725	0.363988	2011	United States	9	2			
133	132 USA		2	1947 14315999	1583	77015	1315777	144126	40857	0.309942	2011	United States	13	0			
134	133 USA		2	1948 10960998	1446	80413	1396394	147208	42019	0.29466	2011	United States	50	3			
135	134 USA		2	1949 13503000	1615	70740	1199344	149767	43216	0.273166	2011	United States	170	0			
136	135 USA		2	1950 14559000	1460	87848	1241836	152271	43964	0.284443	2011	United States	299	0			
137	136 USA		2	1951 33398000	3249	95435	1325948	154878	44711	0.3195	2011	United States	438	16			
138	137 USA		2	1952 47852000	3636	84520	1309824	157553	45123	0.311367	2011	United States	841	10			
139	138 USA		2	1953 49621008	3555	101250	1347597	160184	46581	0.311158	2011	United States	1169	11			
140	139 USA		2	1954 42786000	3302	80115	1302449	163026	47007	0.28094	2011	United States	1703	6			
141	140 USA		2	1955 40518000	2935	106173	1415958	165931	47688	0.266423	2011	United States	2422	15			
142	141 USA		2	1956 41773008	2806	104522	1492506	168903	48389	0.260614	2011	United States	3692	18			
143	142 USA		2	1957 44548000	2796	102253	1494767	171984	49087	0.255049	2011	United States	5543	30			
144	143 USA		2	1958 45503008	2411	77342	1437859	174882	49733	0.23447	2011	United States	7345	73			
145	144 USA		2	1959 46614000	2328	84773	1513627	177830	50320	0.228965	2011	United States	12298	0			
146	145 USA		2	1960 45380000	2306	90067	1421218	180671	50966	0.215444	2011	United States	18638	0			
147	146 USA		2	1961 47808000	2307	88917	1447002	183691	50941	0.210655	2011	United States	22229	10			
148	147 USA		2	1962 52381008	2807	89202	1508570	186538	50916	0.209911	2011	United States	25540	96			
149	148 USA		2	1963 52295008	2700	99120	1589655	189242	50891	0.207967	2011	United States	28133	45			
150	149 USA		2	1964 51213008	2687	115281	1610682	191889	50866	0.202983	2011	United States	29463	50			
151	150 USA		2	1965 51827008	2660	119260	1700357	194303	50841	0.201539	2011	United States	31139	40			
152	151 USA		2	1966 67572000	3090	121655	1784048	196560	50815	0.20867	2011	United States	31175	49			
153	152 USA		2	1967 75448000	3380	115406	1903194	198712	50790	0.208548	2011	United States	31255	40			
154	153 USA		2	1968 80732000	3550	119262	2094353	200706	50765	0.203412	2011	United States	29561	73			
155	154 USA		2	1969 81446000	3460	128152	2198902	202677	53117	0.197951	2011	United States	27552	69			
156	155 USA		2	1970 77827008	3070	119309	2276904	205052	55468	0.180969	2011	United States	26008	66			
157	156 USA		2	1971 74862000	2720	109265	2311854	207661	55560	0.169956	2011	United States	25830	31			
158	157 USA		2	1972 77639000	2323	120875	2396418	209896	55652	0.164743	2011	United States	26516	38			
159	158 USA		2	1973 78385000	2206	136804	2490304	211909	55743	0.158819	2011	United States	27835	27			
160	159 USA		2	1974 85906000	2146	132196	2430725	213854	55736	0.152699	2011	United States	28537	26			
161	160 USA		2	1975 90948000	2098	105897	2393766	215973	55728	0.143948	2011	United States	27519	24			
162	161 USA		2	1976 91013000	2075	116121	2536257	218035	55721	0.142463	2011	United States	25914	20			
163	162 USA		2	1977 1.01E+08	2060	113701	2619481	220239	55713	0.141824	2011	United States	25542	24			
164	163 USA		2	1978 1.09E+08	2060	124314	2612982	222585	56208	0.139674	2011	United States	24418	20			
165	164 USA		2	1979 1.22E+08	2033	123688	2659868	225055	56702	0.137216	2011	United States	24138	15			
166	165 USA		2	1980 1.44E+08	2050	101456	2584210	227726	57197	0.133115	2011	United States	24104	14			
National Material Capabilities																	

```
knitr::include_graphics(path = "Screenshot (106).png")
```



```
knitr::include_graphics(path = "Screenshot (107).png")
```

