

Rowan University

Rowan Digital Works

Faculty Scholarship for the College of Science & Mathematics

College of Science & Mathematics

4-5-2020

A Data Scientist Looks at Covid-19 Part I: Local and National Statistics

Anthony Breitzman

Rowan University, breitzman@rowan.edu

Follow this and additional works at: https://rdw.rowan.edu/csm_facpub



Part of the [Computer Sciences Commons](#), and the [Data Science Commons](#)

Let us know how access to this document benefits you - share your thoughts on our feedback form.

Recommended Citation

Breitzman, Anthony, "A Data Scientist Looks at Covid-19 Part I: Local and National Statistics" (2020).

Faculty Scholarship for the College of Science & Mathematics. 179.

https://rdw.rowan.edu/csm_facpub/179

This Technical Report is brought to you for free and open access by the College of Science & Mathematics at Rowan Digital Works. It has been accepted for inclusion in Faculty Scholarship for the College of Science & Mathematics by an authorized administrator of Rowan Digital Works. For more information, please contact brush@rowan.edu.

A Data Scientist Looks at Covid-19

Part I: Local and National Statistics

Anthony Breitzman, PhD

Associate Professor of Data Science and Computer
Science

Rowan University

4/5/2020

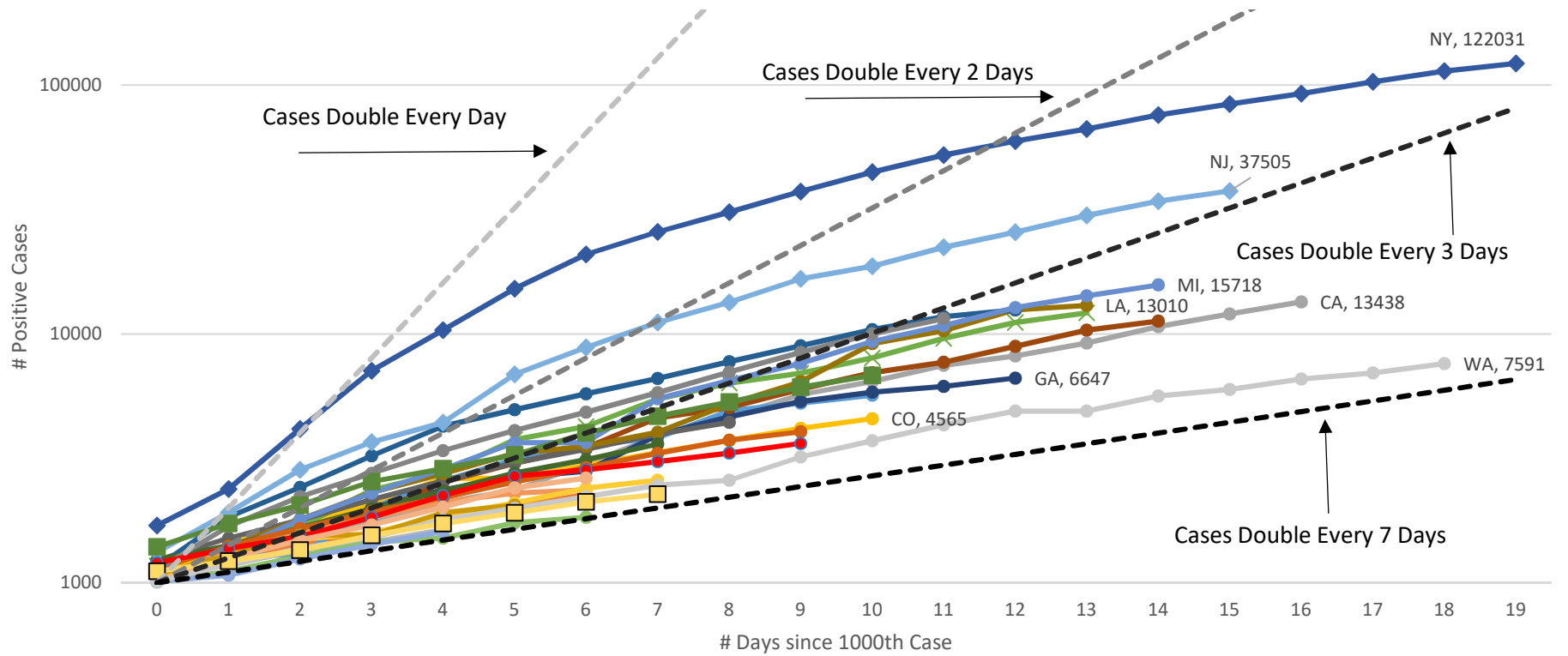
Why?

- I was curious to see if the curve was flattening in certain states and couldn't find what I wanted
- So I found the raw data and created the following figures
- State data is current as of 4pm 4/5/2020
- It comes from <https://covidtracking.com/about-tracker/> which is a site that aggregates data from each state health authority
- NJ County data is from <https://www.insidernj.com/insidernj-covid19-information-update-center/> as of 1pm 4/5/2020 and confirmed with <http://www.covid19.nj.gov> at 4/5/2020 3pm

The News is dire, but is there any good news?

- US totals and NY, NJ are bad... but the rate of growth is decreasing
- The next graph shows evidence of **“flattening-the-curve”** in all of the states with large outbreaks

Cumulative Number of COVID-19 Cases by Number of Days Since 1000th Case (Through 4/5/2020 4PM Eastern Time)



- | | | | | | | | |
|----|----|----|----|--------|--------|--------|--------|
| AL | AZ | CA | CO | CT | FL | GA | IL |
| IN | LA | MA | MD | MI | MO | MS | NC |
| NJ | NV | NY | OH | PA | SC | TN | TX |
| UT | VA | WA | WI | Db1Day | Db2Day | Db3Day | Db7Day |

Next Slide is Hard to read, but Shows Growth is Slowing in Every State

- Red is bad; blue is good.
- NY Cases are doubling every 7 days. Two weeks ago they were doubling every 1.6 days.
- NJ Cases are doubling every 5.4 days compared with every 2.2 days two weeks ago. (Note this still means 75,000 cases by next weekend unless more progress is made)
- Pennsylvania cases are growing faster than the rest of the country
- Delaware has few cases but is not significantly reducing its growth
- Oregon is doing the best of any state in terms of reducing growth

Cumulative Positive Cases of COVID-19 by State

Table with 35 columns (19-Mar to 5-Apr) and 51 rows (All US* to AK) showing cumulative COVID-19 cases by state.

Current Growth

Table with 3 columns (1 Day Growth Rate, 3 Day Growth Rate, Cases Double Every N Days) and 51 rows (All US* to AK) showing current growth metrics.

1 Week Ago

Table with 3 columns (1 Day Growth Rate, 3 Day Growth Rate, Cases Double Every N Days) and 51 rows (All US* to AK) showing growth metrics 1 week ago.

2 Weeks Ago

Table with 3 columns (1 Day Growth Rate, 3 Day Growth Rate, Cases Double Every N Days) and 51 rows (All US* to AK) showing growth metrics 2 weeks ago.

data from https://covidtracking.com/about-tracker/ 4/5/2020 4PM Eastern Time
Breitbartman 4/6/2020 1:50 AM
*All US includes US territories not shown

Testing Numbers as of 4pm 4/5/2020

State	Positive	Negative	Pending	Positive + Negative	Positive / (Positive + Negative)	Deaths / Positive Cases
NY	122031	180249		302280	40%	3.4%
CA	13438	103095	15000	116533	12%	2.4%
FL	12151	101253	1129	113404	11%	1.8%
WA	7591	80327		87918	9%	4.1%
NJ	37505	44661		82166	46%	2.4%
PA	11510	66261		77771	15%	1.3%
MA	12500	59437		71937	17%	1.8%
TX	6812	64126		70938	10%	1.9%
LA	13010	47315		60325	22%	3.7%
IL	11256	47727		58983	19%	2.4%
MI	15718	30030		45748	34%	3.9%
TN	3633	41667		45300	8%	1.2%
OH	4043	39713		43756	9%	2.9%
NC	2585	37460		40045	6%	1.2%
UT	1605	29287		30892	5%	0.5%
MD	3609	24728		28337	13%	1.9%
GA	6647	21185		27832	24%	3.2%
WI	2267	25169		27436	8%	3.0%
AZ	2269	25141		27410	8%	2.8%
MO	2367	24882		27249	9%	1.4%
MN	935	25842		26777	3%	3.1%
CO	4565	19335		23900	19%	2.8%
VA	2637	21034		23671	11%	1.9%
CT	5675	17595		23270	24%	3.3%
IN	4411	18241		22652	19%	2.9%
SC	2049	16927		18976	11%	2.1%
OR	999	17926		18925	5%	2.6%
NV	1836	14995		16831	11%	2.5%
NM	543	16285		16828	3%	2.0%
KY	917	15746		16663	6%	4.4%
AL	1796	11282		13078	14%	2.5%
HI	351	12604		12955	3%	1.1%
AR	830	10412		11242	7%	1.9%
IA	868	9973		10841	8%	2.5%
ID	1077	9184		10261	10%	0.9%
WV	324	8514		8838	4%	0.9%
KS	747	7476		8223	9%	2.9%
RI	922	7181		8103	11%	2.7%
NH	621	7411	93	8032	8%	1.4%
MS	1638	5580		7218	23%	2.6%
DE	673	6321		6994	10%	2.1%
DC	998	5836		6834	15%	2.2%
ND	207	6580		6787	3%	1.4%
MT	286	6317		6603	4%	2.1%
VT	512	6070		6582	8%	4.3%
ME	470	6088		6558	7%	2.1%
AK	185	6099		6284	3%	3.2%
SD	240	5353		5593	4%	0.8%
WY	197	3040		3237	6%	0.0%
OK	1252	1401		2653	47%	3.7%

- Each state has more negative tests than positive tests which is better than last week
- Delaware is doing much better than last week
- California has an awful lot of pending tests but only 1/3 as many as last week.
- South Korea has extensive testing and has a mortality rate of 1.8%. Japan's is 2.4%. From this we can infer that **NY probably has another 100,000 untested cases**
- Source: <https://covidtracking.com/about-tracker/>

NJ COVID-19 Cases by County

County	3/20 7:35am ...	3/23 2:01pm	3/24 3:43pm	3/25/ 2:00pm	3/26 2:57pm	3/27 2:03pm	3/28 1:00pm	3/29 2:48pm	3/30 2:32pm	3/31 1:01pm	4/1 1:00pm	4/2 1:00pm	4/3 1:05pm	4/4 1:03pm	4/5 1:15pm	1 Day Growth Rate	3 Day Average Growth Rate	Cases Double Every N Days
Bergen County:	195	609	701	819	1206	1505	1838	2169	2482	2909	3494	4099	4866	5760	6187	7.4%	14.7%	5.1
Essex County:	63	273	342	381	609	826	1086	1227	1564	1900	2262	2617	3067	3584	4082	13.9%	16.0%	4.7
Hudson County:	55	190	234	260	441	594	771	974	1314	1606	1910	2270	2835	3491	3924	12.4%	20.0%	3.8
Passaic County:	38	141	216	255	399	504	608	831	1091	1294	1494	1750	2216	2856	3227	13.0%	22.6%	3.4
Union County:	29	189	246	262	432	587	742	896	1213	1418	1661	2010	2487	2916	3216	10.3%	17.0%	4.4
Middlesex County:	64	210	277	316	505	657	808	938	1123	1277	1493	1766	2125	2578	2950	14.4%	18.7%	4.1
Monmouth County:	43	238	288	313	501	634	781	870	1030	1140	1301	1458	1743	2065	2354	14.0%	17.3%	4.3
Ocean County:	33	144	180	222	389	507	624	759	874	1022	1209	1371	1685	2003	2177	8.7%	16.7%	4.5
Morris County:	26	177	204	223	315	379	442	556	720	841	942	1082	1298	1618	1800	11.2%	18.5%	4.1
Somerset County:	21	67	102	117	179	219	258	295	349	413	472	549	641	765	833	8.9%	14.9%	5.0
Mercer County:	20	50	58	82	111	131	168	202	249	268	333	386	484	586	654	11.6%	19.2%	3.9
Camden County:	13	33	51	61	73	95	123	163	200	228	289	343	406	481	556	15.6%	17.5%	4.3
Burlington County:	14	36	42	48	64	88	115	142	178	202	255	294	367	469	547	16.6%	23.0%	3.3
Sussex County:	2	15	18	27	49	65	81	93	113	132	158	179	210	236	267	13.1%	14.3%	5.2
Gloucester County:	3	13	19	23	33	40	51	72	89	114	149	169	183	215	248	15.3%	13.6%	5.4
Warren County:	3	12	15	18	31	41	51	56	68	76	96	116	149	182	195	7.1%	18.9%	4.0
Hunterdon County:	8	18	25	25	39	52	61	66	79	97	117	130	148	171	189	10.5%	13.3%	5.6
Atlantic County:	3	6	6	9	10	14	17	24	29	31	40	50	72	98	121	23.5%	34.3%	2.4
Cape May County:	1	2	3	4	6	7	7	9	9	12	22	34	44	50	77	54.0%	31.3%	2.5
Cumberland County:		1	2	3	4	9	11	11	12	18	27	31	36	40	54	35.0%	20.3%	3.7
Salem County:		1	1	1	2	2	3	3	3	12	19	20	25	25	26	4.0%	9.1%	7.9
Unassigned NJ	118	419	645	933	1478	1984	2478	3020	3847	3686	4512	4866	4808	3935	3821	-2.9%	-7.7%	-8.6
All NJ	752	2844	3675	4402	6876	8938	11124	13376	16636	18696	22255	25590	29895	34124	37505	9.9%	13.6%	5.4

Source: <https://www.insidernj.com/insidernj-covid19-information-update-center/>

Data confirmed at covid19.nj.gov at 4/5/2020 5:21pm

Data is provisional and subject to revision.

NJ Total confirmed at JHU at 4/5/2020 5:22 pm

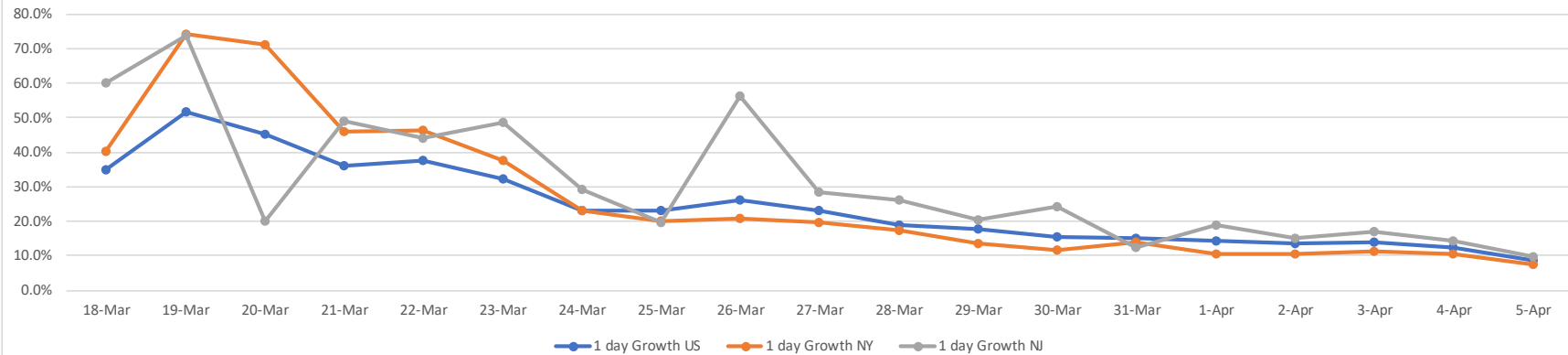
Breitzman 4/5/2020 5:22pm

Note: On March 27 at 2:03 Middlesex, Morris, Passaic, Ocean, Somerset, Sussex, Union, and Warren Counties had not updated their counts from the previous day. For these entries we take the average of counts for 3/26 and 3/28. (March 27 is not involved in any of the projections.)

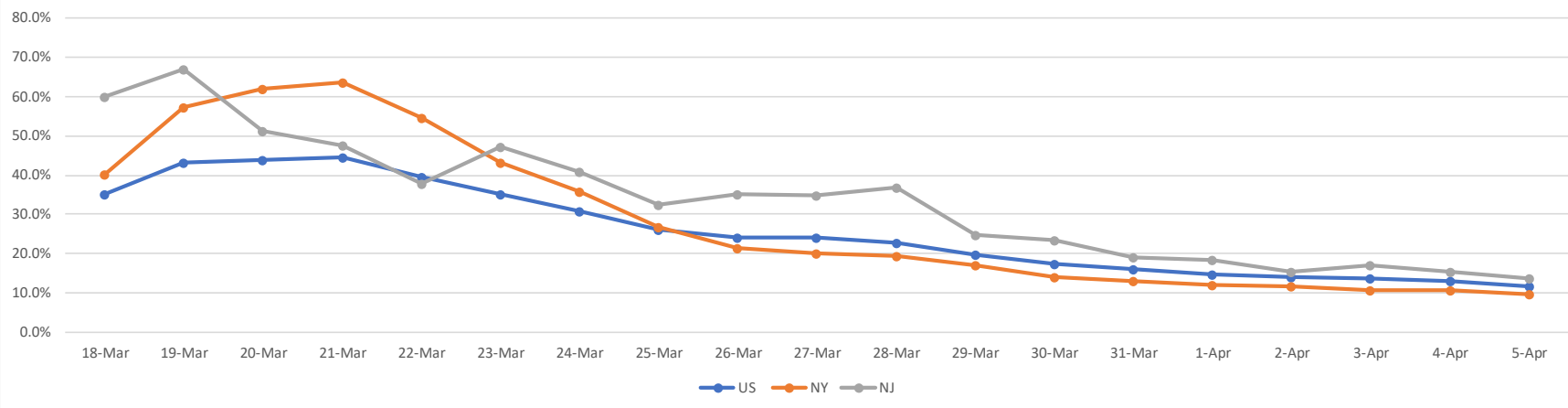
Notes on Previous Slide

- Bergen county hit the worst, but growth rates starting to slow
- South Jersey many fewer cases than North Jersey, but growth rates are worse
- Atlantic and Cape May counties growing at twice the state rate
- Last week it appeared NJ would hit 50,000 cases by today, but the growth rate slowed. Current rate suggests 75,000 cases by next weekend, but if the rate continues to slow we may see 55,000 cases instead (which is still 22,000 more than today)

COVID-19 Growth finally Slowing Change in 1-Day Growth Rates for NJ, NY, US



1 Day Growth Rates Smoothed (3-day Running Averages of Above)



COVID 19 Cases by Date

	18-Mar	19-Mar	20-Mar	21-Mar	22-Mar	23-Mar	24-Mar	25-Mar	26-Mar	27-Mar	28-Mar	29-Mar	30-Mar	31-Mar	1-Apr	2-Apr	3-Apr	4-Apr	5-Apr
All US	7730	11719	17033	23197	31879	42152	51954	63928	80735	99413	118234	139061	160530	184683	210816	239099	271988	305755	332308
NY	2382	4152	7102	10356	15168	20875	25665	30811	37258	44635	52318	59513	66497	75795	83712	92381	102863	113704	122031
NJ	427	742	890	1327	1914	2844	3675	4402	6876	8825	11124	13386	16636	18696	22255	25590	29895	34124	37505

1 Day Growth Rates by Day

	18-Mar	19-Mar	20-Mar	21-Mar	22-Mar	23-Mar	24-Mar	25-Mar	26-Mar	27-Mar	28-Mar	29-Mar	30-Mar	31-Mar	1-Apr	2-Apr	3-Apr	4-Apr	5-Apr
1 day Growth US	35.1%	51.6%	45.3%	36.2%	37.4%	32.2%	23.3%	23.0%	26.3%	23.1%	18.9%	17.6%	15.4%	15.0%	14.2%	13.4%	13.8%	12.4%	8.7%
1 day Growth NY	40.1%	74.3%	71.1%	45.8%	46.5%	37.6%	22.9%	20.1%	20.9%	19.8%	17.2%	13.8%	11.7%	14.0%	10.4%	10.4%	11.3%	10.5%	7.3%
1 day Growth NJ	59.9%	73.8%	19.9%	49.1%	44.2%	48.6%	29.2%	19.8%	56.2%	28.3%	26.1%	20.3%	24.3%	12.4%	19.0%	15.0%	16.8%	14.1%	9.9%

data from <https://covidtracking.com/about-tracker/> 4/5/2020 4PM Eastern Time

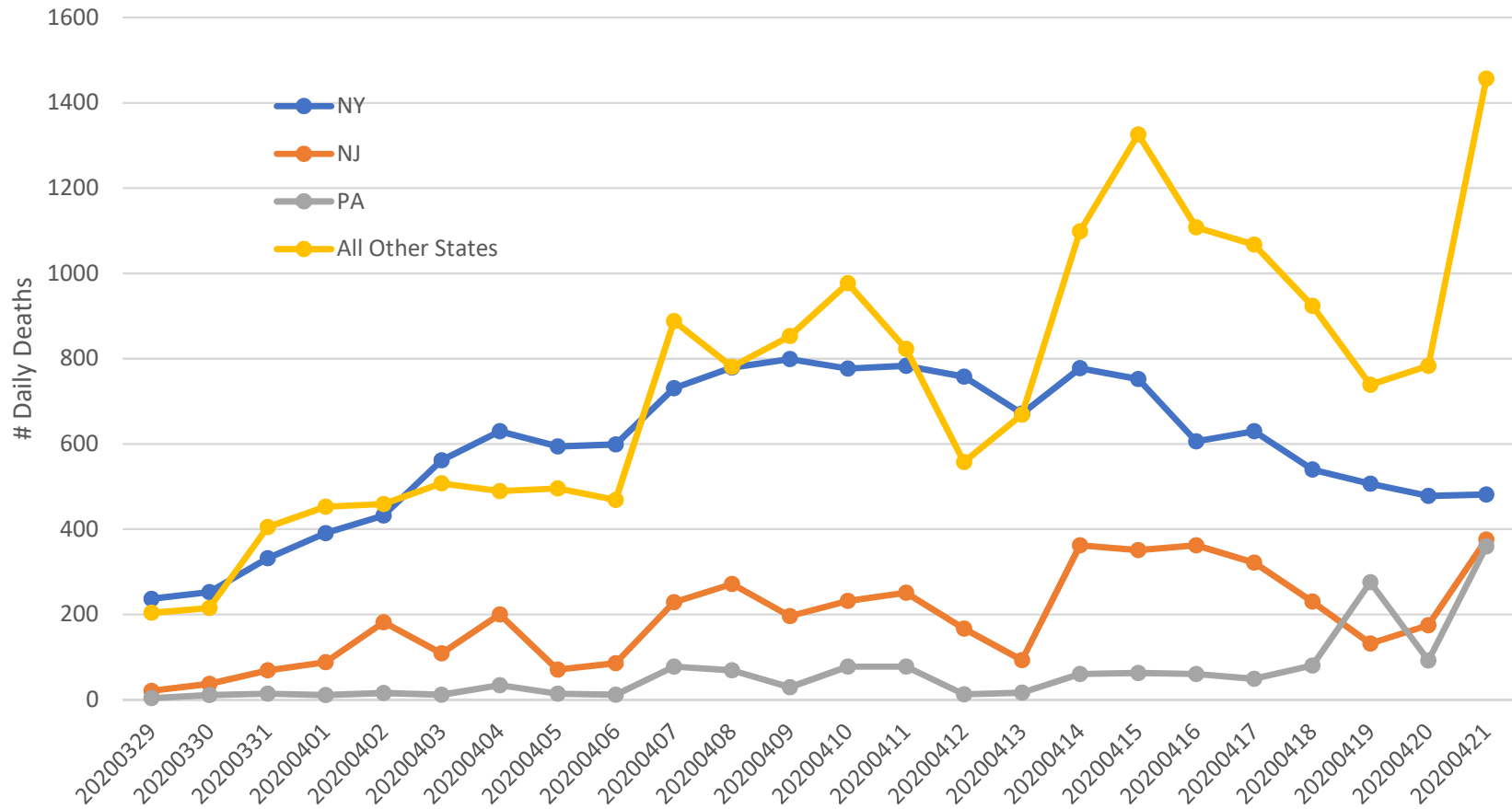
Breizman 4/6/2020 1:50 AM

Summary

- Things are bad, but there is evidence that social distancing and staying home is starting to slow down the rate of infection
- Cases are still growing rapidly, but at a rate much slower than last week and the week before
- Pennsylvania and Delaware have not flattened the curve as well as other states
- Testing is much better than it was, but evidence from South Korea and Japan suggests it is still lagging in certain states

Addendum from 4/21/2020

Scary Upturn in Deaths the Last Few Days (Number of Daily Deaths for 3-States and Rest of US)



Breiztman 4/22/2020 2:12 AM Source:
<https://covidtracking.com/api> 4/21/2020 4PM

=====Tests=====

state	positive	negative	pending	Total Tests	Deaths	Deaths/Positive Cases	Potential Cases*	Potential/Actual
MI	32967	84259		117226	2700	0.082	195652	5.93
CT	20360	43832		64192	1423	0.070	103116	5.06
MN	2567	45130		47697	160	0.062	11594	4.52
NY	251690	397635		649325	14828	0.059	1074493	4.27
OK	2807	41537		44344	164	0.058	11884	4.23
LA	24854	117576		142430	1405	0.057	101812	4.10
WA	12085	128926		141011	652	0.054	47246	3.91
KS	2025	17076		19101	107	0.053	7754	3.83
WI	4620	47841	219	52461	242	0.052	17536	3.80
IN	12097	55167		67264	630	0.052	45652	3.77
NJ	92387	92439		184826	4753	0.051	344420	3.73
KY	3050	29770		32820	154	0.050	11159	3.66
VT	818	12293		13111	40	0.049	2899	3.54
MA	41199	134173		175372	1961	0.048	142101	3.45
PA	34528	132323		166851	1564	0.045	113333	3.28
CO	10106	37360		47466	449	0.044	32536	3.22
IL	33059	121938		154997	1468	0.044	106377	3.22
OH	13250	80989		94239	557	0.042	40362	3.05
NV	3937	29118		33055	163	0.041	11812	3.00
MD	14193	59442		73635	584	0.041	42319	2.98
ME	888	14076		14964	36	0.041	2609	2.94
GA	19881	68259		88140	799	0.040	57899	2.91
AZ	5251	49901		55152	208	0.040	15072	2.87
OR	2002	39126		41128	78	0.039	5652	2.82
MS	4716	47648		52364	183	0.039	13261	2.81
CA	33261	266839		300100	1268	0.038	91884	2.76
DC	3098	11841		14939	112	0.036	8116	2.62
VA	9451	48903	477	58354	324	0.034	23478	2.48
AL	5231	43295		48526	177	0.034	12826	2.45
MO	5941	51179		57120	189	0.032	13696	2.31
FL	27495	249564	1203	277059	856	0.031	62029	2.26
RI	5500	33833		39333	171	0.031	12391	2.25
NC	6951	76380		83331	213	0.031	15435	2.22
NM	1971	36784		38755	58	0.029	4203	2.13
WV	914	21849		22763	26	0.028	1884	2.06
NH	1491	12848	244	14339	42	0.028	3043	2.04
DE	2931	13725		16656	82	0.028	5942	2.03
SC	4439	36838		41277	124	0.028	8986	2.02
ID	1736	15924		17660	48	0.028	3478	2.00
MT	437	10804		11241	12	0.027	870	1.99
AK	329	10790		11119	9	0.027	652	1.98
TX	20196	185203		205399	517	0.026	37464	1.86
IA	3641	23974		27615	83	0.023	6014	1.65
TN	7394	100788		108182	157	0.021	11377	1.54
ND	644	14343		14987	13	0.020	942	1.46
AR	2227	25214		27441	43	0.019	3116	1.40
WY	320	7301		7621	6	0.019	435	1.36
HI	584	24112		24696	10	0.017	725	1.24
UT	3296	69062		72358	32	0.010	2319	0.70
SD	1755	11060		12815	8	0.005	580	0.33
50 State Total	796570	3330287	2143	4126857	39888	0.050	2890435	3.63

The Lancet Infectious Diseases medical journal estimates that about 0.66 percent of patients who become infected with the virus will die. When undetected infections aren't taken into account, researchers found the coronavirus death rate was 1.38 percent.

*Potential Cases is based on the more conservative mortality rate of 1.38 percent

This suggests NJ may have about 3.5 times as many cases as it thinks it does. The same is true for the country as a whole.

Of course if we take the higher estimate of 0.66% then the numbers double from the projections

Source: <https://covidtracking.com/api> 4/21/2020 4PM

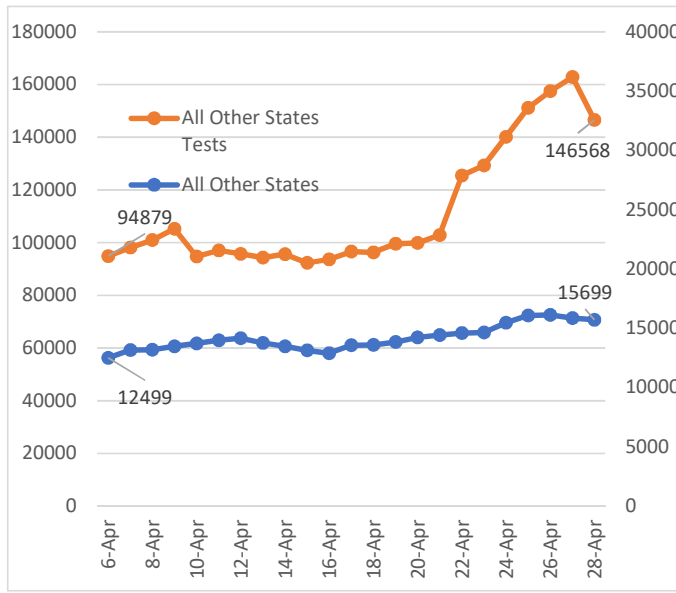
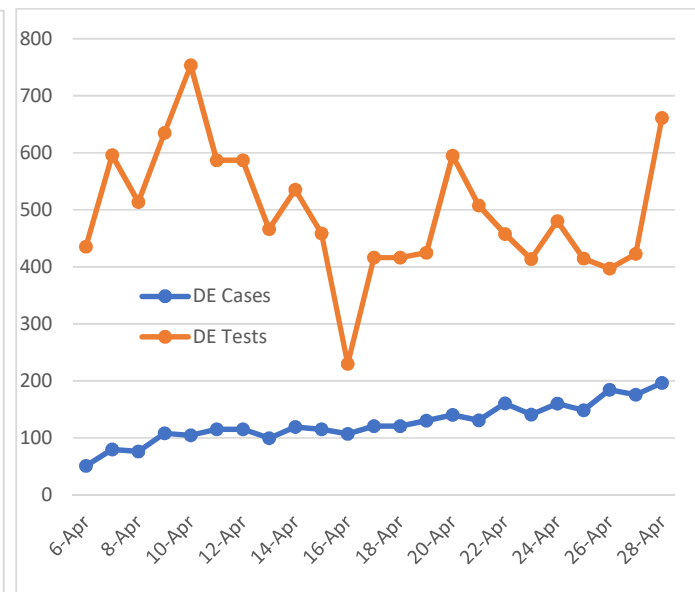
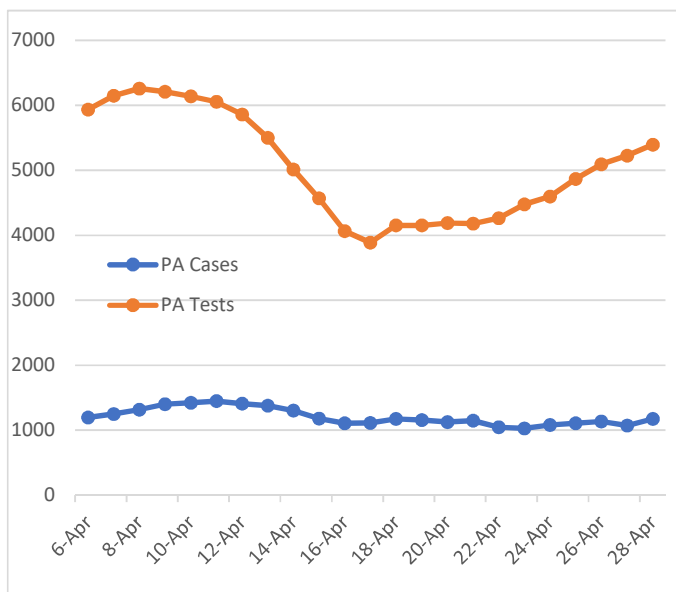
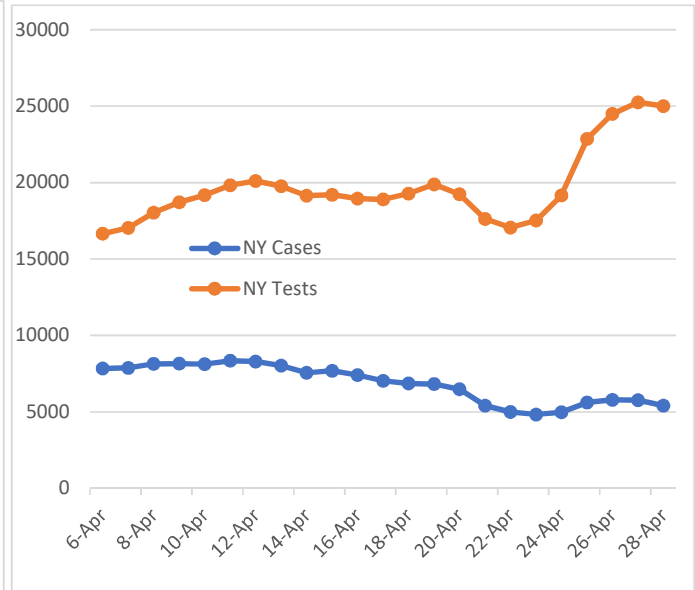
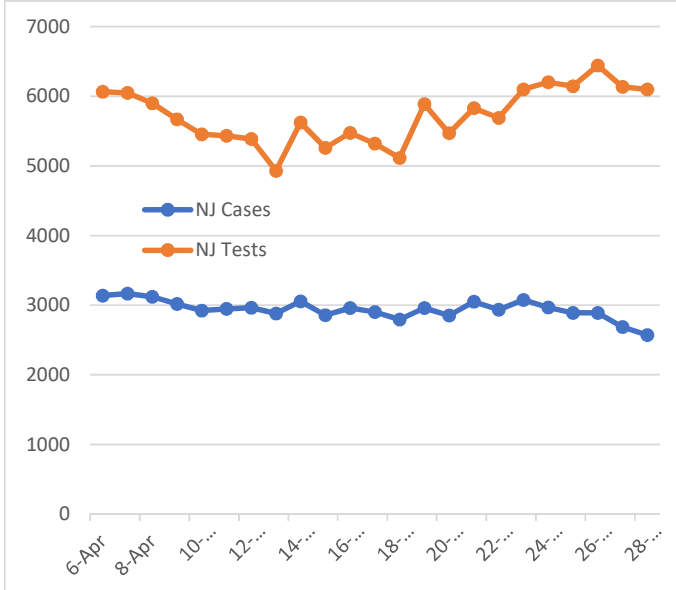
Breizman 4/22/2020 2:08 AM

Addendum from 4/29/2020

COVID-19 Daily Cases versus Daily Testing for 4 States and Rest of Country

(7 Day Rolling Averages to Smooth Out Curves)

(Note: Each State Axis to different Scale)



Source:

<https://covidtracking.com/api> as of 4pm 4/28/2020

Notes:

1. All graphs are 7 day rolling averages
2. Cases still rising in US and Delaware
3. Note secondary axis in Other State graph
4. Only NY is increasing testing significantly while cases are decreasing.
5. Delaware testing all over the place. Perhaps reporting is not consistent

7-Day Rolling averages consist of (cumulative cases in day x - cumulative cases in day x - 6)/7

Breizman: 4/29/2020 2:44 AM