**The Influence of population size, distance from NYC and population density on the number of positive tests in a county in NJ on 4/24/2020.**
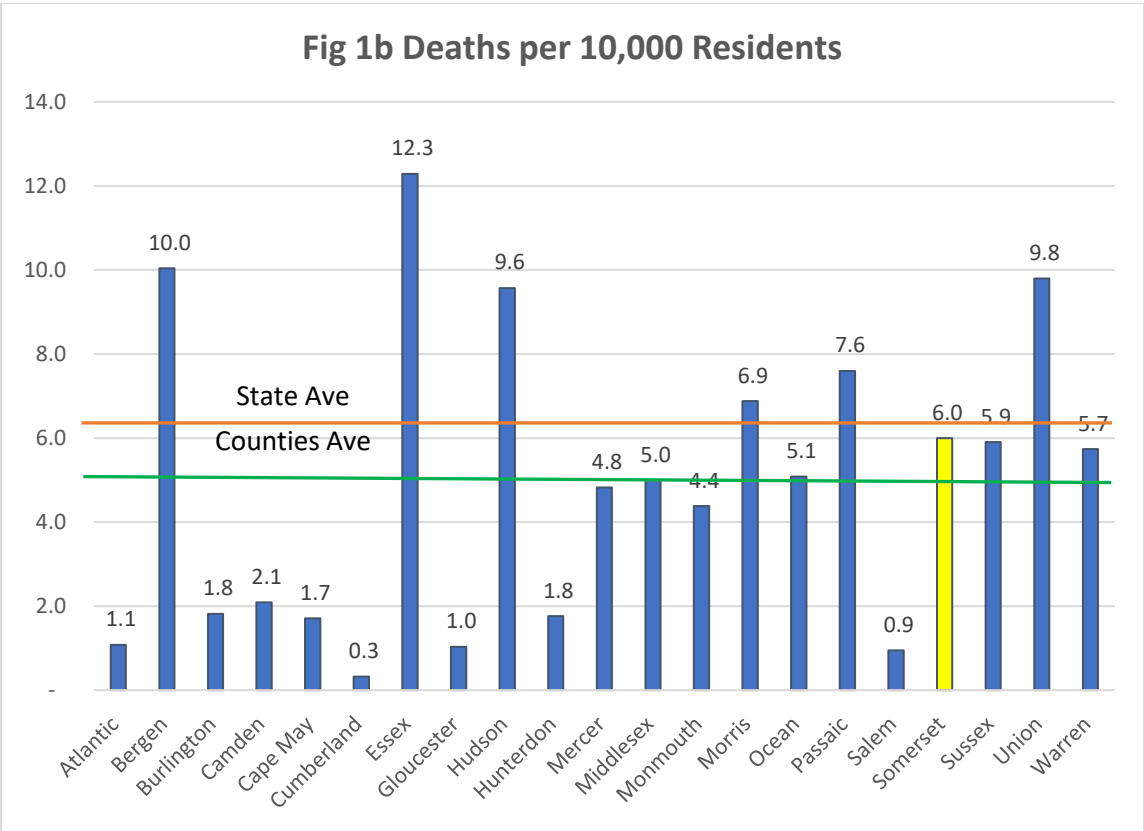
This is a technical post. If you are not into science the bottom line is the higher the population of a county and the closer the county is to NYC the higher the number of positive tests. I now find population density does have a significant influence on the number of positive cases. (See 2nd to last paragraph). This time I also look at the number of those who have died as well.

My original post on this was on 3/28/2020 and I updated it on 4/3/2020, 4/11/2020 and 4/20/2020. For the rest of this post I will use the word "cases" as shorthand for "positive tests".

Fig 1a compares the numbers of cases/1000 people county by county. I drew in the average value of counties 8.9 (green) and the state value of 11.4 (orange). Somerset County is 8.8.


Fig 1a Positive Tests per 1000 Residents
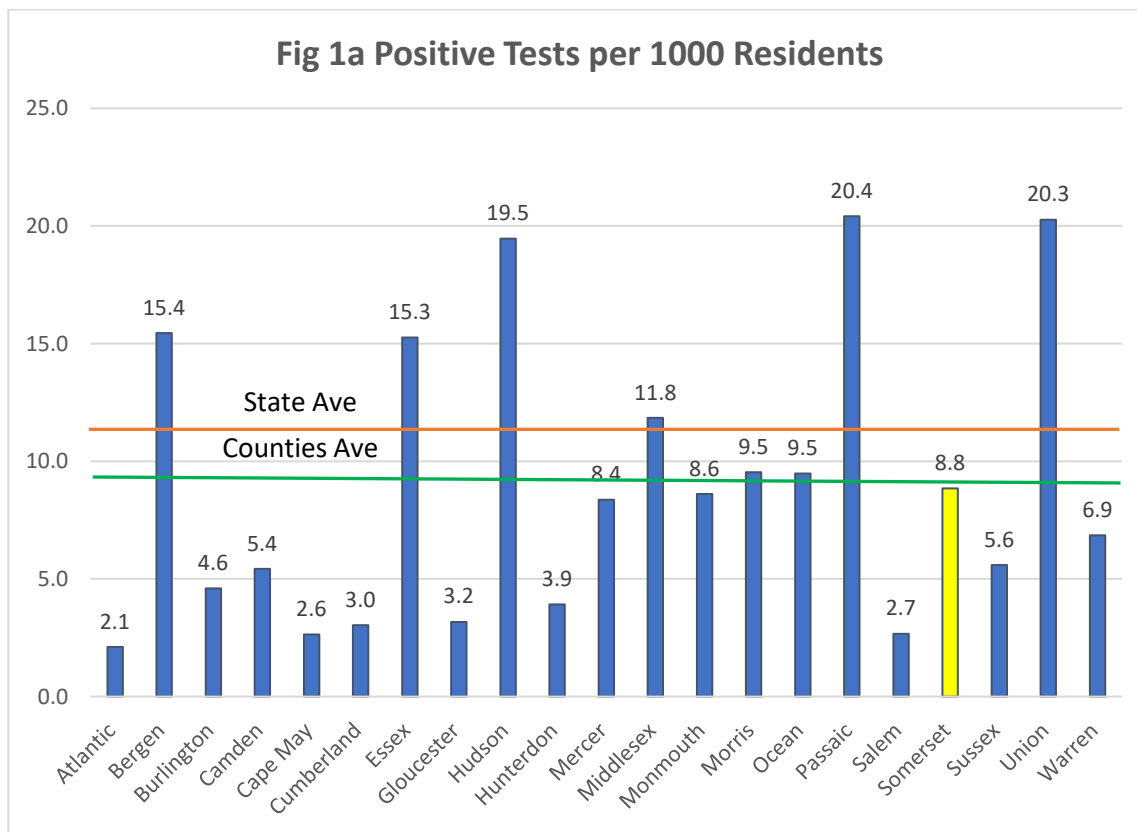
I also looked at deaths/100,000 in Figure 1b. The average values are for counties 4.9 (green) and the state 6.3 (orange). Somerset County is 6.0.

Fig 1b Deaths per 10,000 Residents

State Ave
Counties Ave

| County	Deaths per 10,000
Atlantic	1.1
Bergen	10.0
Burlington	1.8
Camden	2.1
Cape May	1.7
Cumberland	0.3
Essex	12.3
Gloucester	1.0
Hudson	9.6
Hunterdon	1.8
Mercer	4.8
Middlesex	5.0
Monmouth	4.4
Morris	6.9
Ocean	5.1
Passaic	7.6
Salem	0.9
Somerset	6.0
Sussex	5.9
Union	9.8
Warren	5.7

Does the population of a county influence the result (Fig 2a)? On the graph you will note a dotted line. That is the line that best fits (polynomial) the data. $R^2 = 0.79$. R^2 is a measure of the "goodness" of the fit. Zero is no correlation and 1 is perfect correlation. For medical issues anything around 0.8 is pretty good so this isn't bad. No surprise, larger populations have more cases. The yellow dot is Somerset County, right where you'd expect it.

Fig 2a Cases vs Population

$R^2 = 0.7926$
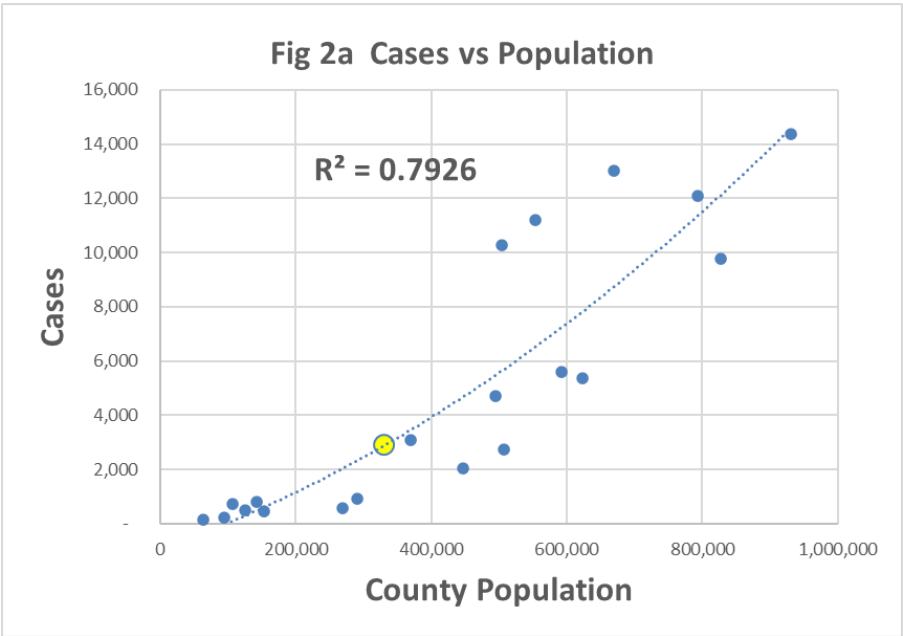
Looking at deaths vs county population (Fig 2b). It is a polynomial fit $R^2 = 0.78$. Thus the number of deaths are influenced by the size of the population.

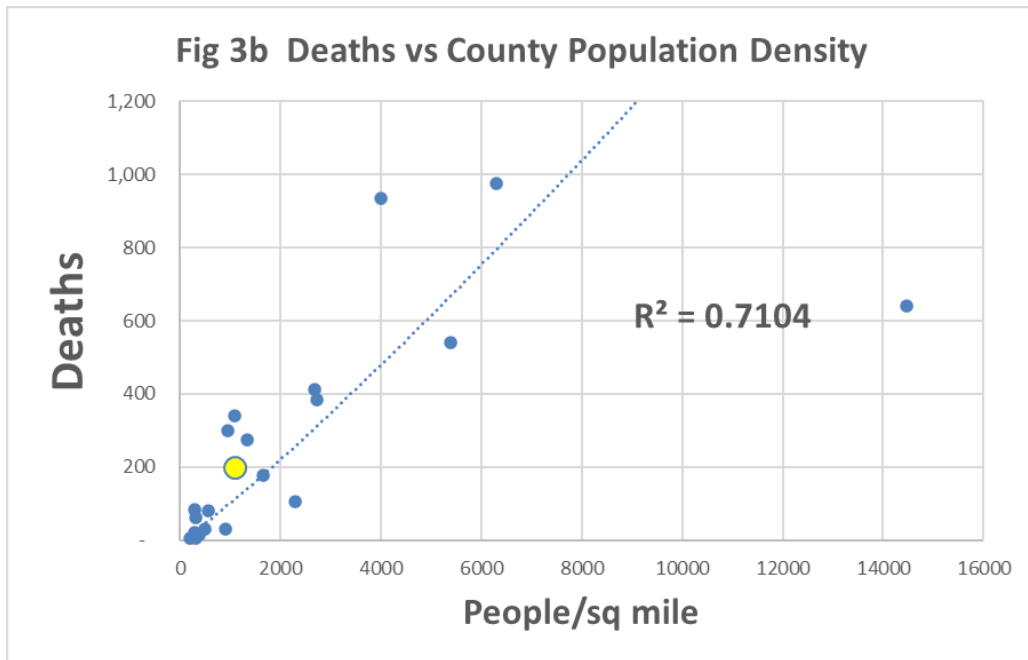Fig 2b Deaths vs Population

$R^2 = 0.7831$
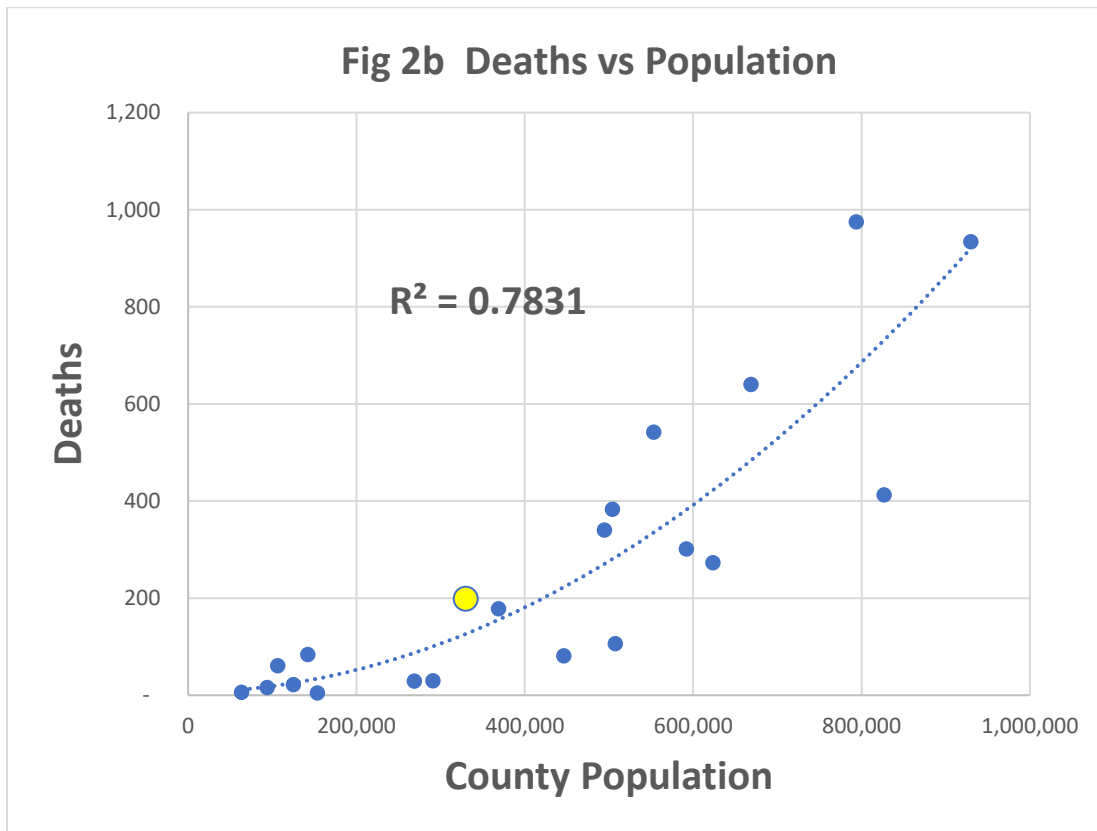
Does population density influence the number of cases? Fig 3a (log fit) addresses the question by looking at cases vs the number of people per square mile for each county. The R^2 is 0.81 so population density does influence the number of cases. This is a change. I went back and looked at log fits of this data and it turns out since April 11th there was a good log fit. I'd only been looking here at linear fits.

Fig 3a Cases vs County Population Density

$R^2 = 0.8119$
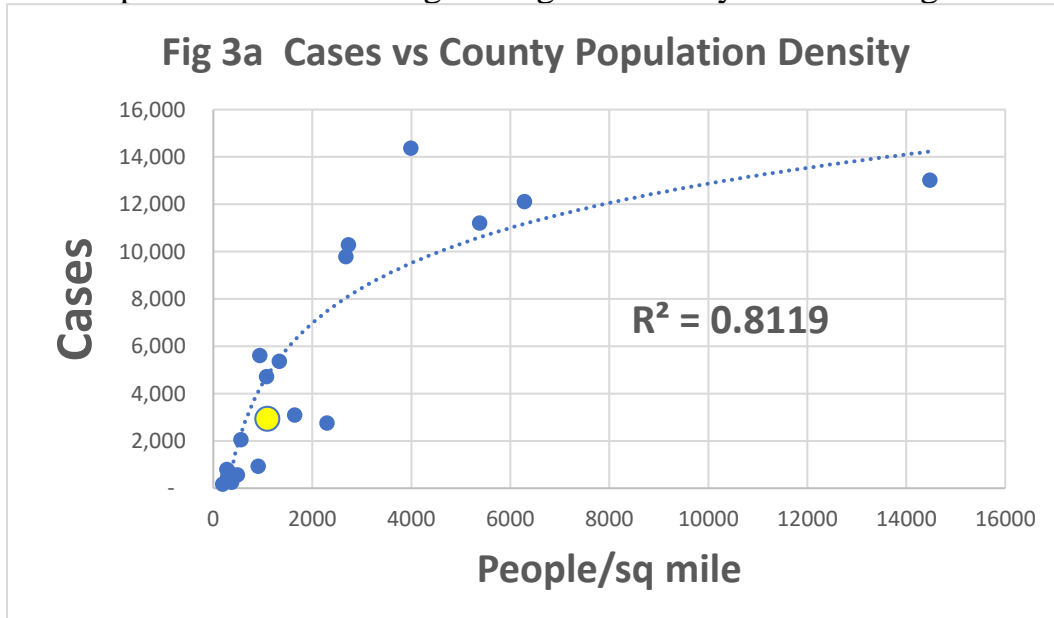
What Fig 3a is telling me is changes at low densities it matters but above densities of about 5000 people/sq mile density effects taper off. And this only seems to happen after the virus has been around a while.

Fig 3b looks at deaths vs population density. It is a power fit. The R^2 =0.71 is less impressive but population density does seem to influence the death rate.

Fig 3b Deaths vs County Population Density

$R^2 = 0.7104$

Deaths (y-axis): -, 200, 400, 600, 800, 1,000, 1,200

People/sq mile (x-axis): 0, 2000, 4000, 6000, 8000, 10000, 12000, 14000, 16000

Does proximity to NYC influence the number of cases? The distances are Google Map's distance from each county to the Empire State Building. The result is in Figure 4a. The curve fit uses a log fit. $R^2 = 0.72$ so the closer to NYC the more cases. Somerset County is where you'd expect it to be.

Fig 4a Cases Vs Distance from NYC

$R^2 = 0.7225$

Cases (y-axis): -, 2,000, 4,000, 6,000, 8,000, 10,000, 12,000, 14,000, 16,000
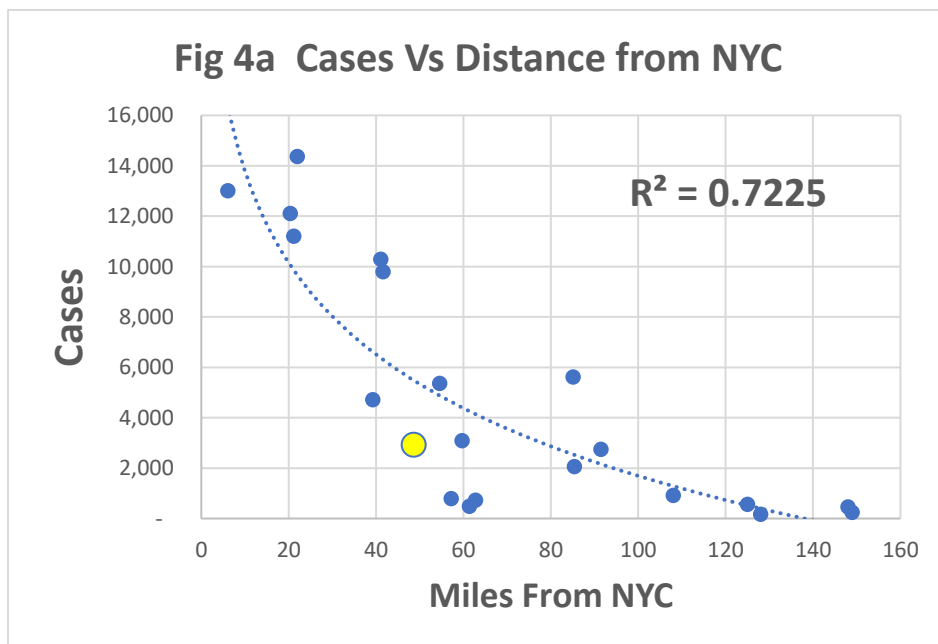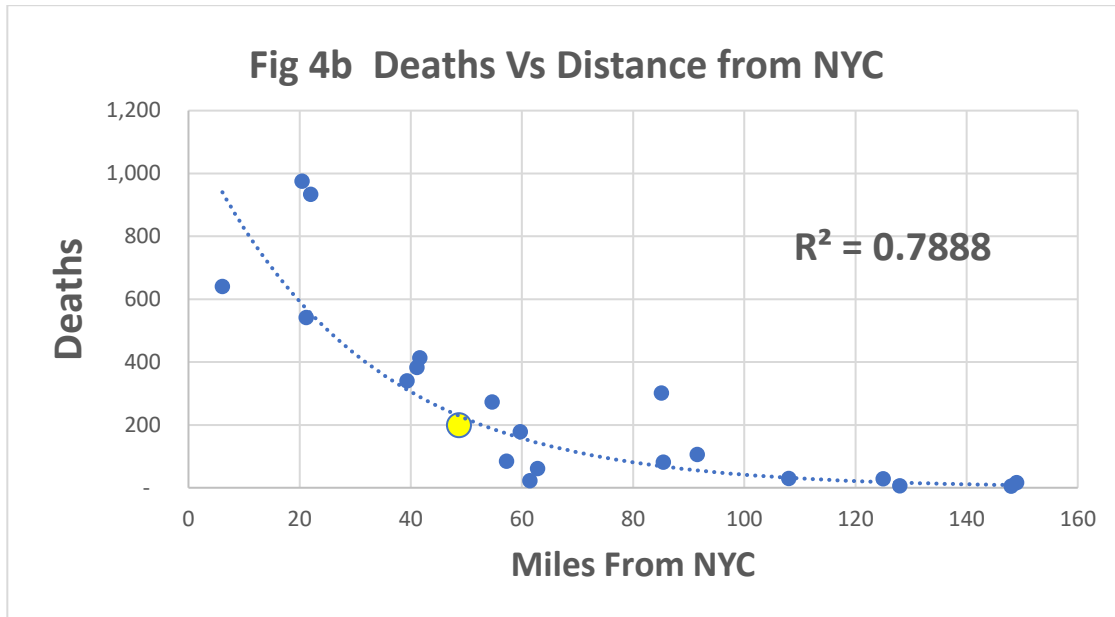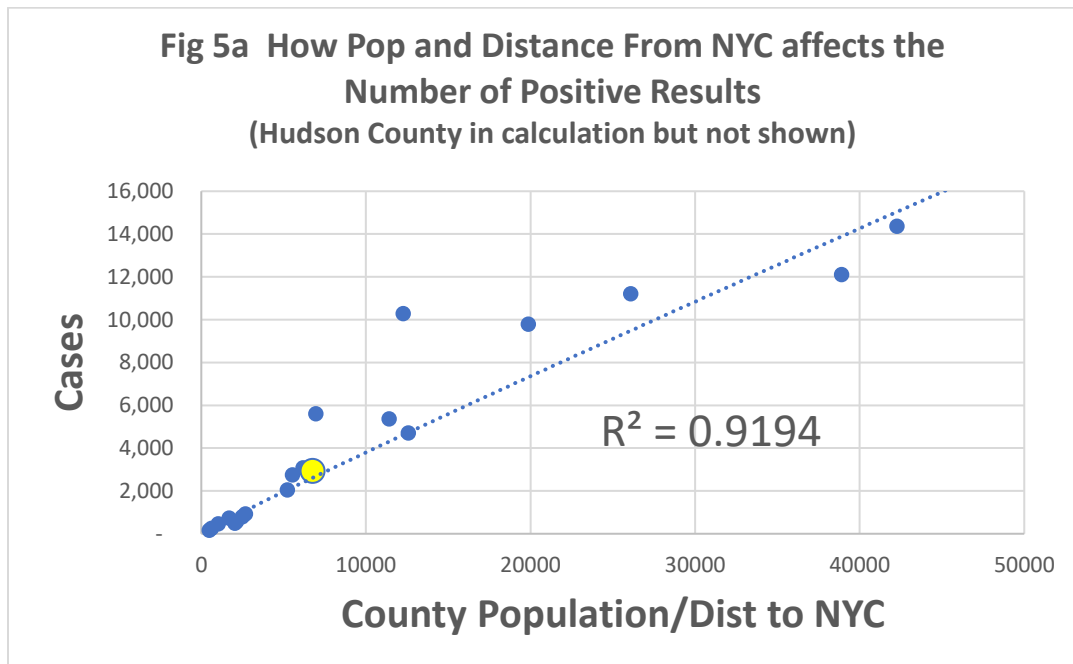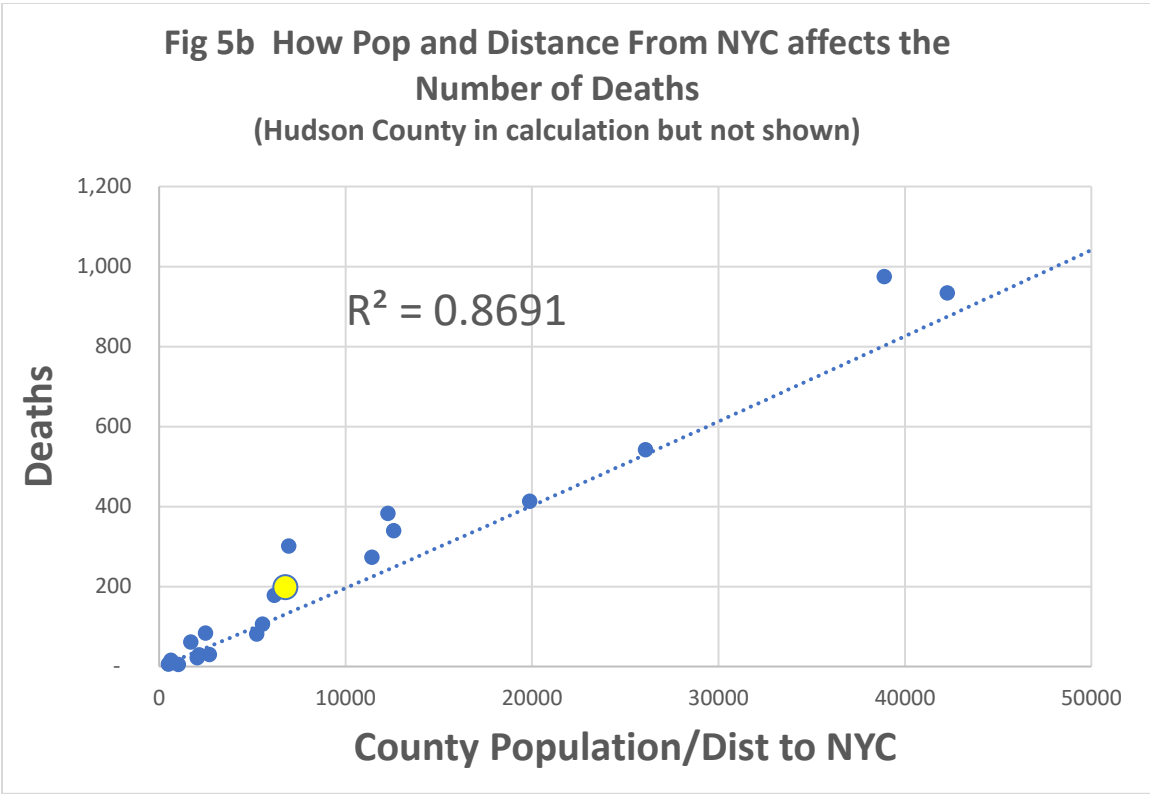
Miles From NYC (x-axis): 0, 20, 40, 60, 80, 100, 120, 140, 160

Fig 4b is an exponential fit and examines the number of deaths vs distance from NYC

Fig 4b Deaths Vs Distance from NYC

$R^2 = 0.7888$

Clearly there is an influence.

As a final exercise I examined how the population of a county together with its proximity to NYC influence the number of cases (Fig 5a) and the number of deaths (Fig 5b). Both are power fits. I plotted the number of both vs population divided by the distance to NYC for each county. This has a very good correlation. By the way, Hudson County had an X axis value of almost 110,000 so for clarity I excluded it from the graph, but it was included in the R^2 calculation.

Fig 5a How Pop and Distance From NYC affects the Number of Positive Results
(Hudson County in calculation but not shown)

$R^2 = 0.9194$

Fig 5b How Pop and Distance From NYC affects the Number of Deaths
(Hudson County in calculation but not shown)

$R^2 = 0.8691$

For the very technically minded a confounding factor is these calculations assume the population of a county is independent of its proximity to NYC. We all know that isn't true.

So, what we have learned is the greater the population of a county and the closer it is to NYC the higher the number of people who have tested positive. Population density is also a factor.

For completeness here is the data.

County	Population	Cases	Area sq mi	Dist NYC	Cases / 1000	cases/ sqmi	Pop Density	Pop/Dist
Atlantic	268,539	410	555.7	125	1.5	0.7	483	2148
Bergen	929,999	13,011	233.0	22	14.0	55.8	3991	42273
Burlington	446,367	1,663	798.6	85.4	3.7	2.1	559	5227
Camden	507,367	2,255	221.3	91.5	4.4	10.2	2293	5545

Cape May	93,705	212	251.4	149	2.3	0.8	373	629
Cumberland	153,400	338	483.7	148	2.2	0.7	317	1036
Essex	793,555	10,729	126.2	20.4	13.5	85.0	6288	38900
Gloucester	290,852	770	322.0	108	2.6	2.4	903	2693
Hudson	668,631	11,150	46.2	6.1	16.7	241.4	14476	109612
Hunterdon	125,051	419	427.8	61.4	3.4	1.0	292	2037
Mercer	368,762	2,591	224.6	59.7	7.0	11.5	1642	6177
Middlesex	826,698	8,346	308.9	41.6	10.1	27.0	2676	19873
Monmouth	623,387	4,787	468.8	54.6	7.7	10.2	1330	11417
Morris	494,383	4,236	460.2	39.3	8.6	9.2	1074	12580
Ocean	591,939	4,868	628.8	85.1	8.2	7.7	941	6956
Passaic	504,041	8,479	184.6	41.1	16.8	45.9	2731	12264
Salem	63,336	120	331.9	128	1.9	0.4	191	495
Somerset	330,176	2,488	301.8	48.7	7.5	8.2	1094	6780
Sussex	142,298	680	519.0	57.2	4.8	1.3	274	2488
Union	553,066	9,972	102.9	21.2	18.0	97.0	5377	26088
Warren	106,293	613	356.9	62.8	5.8	1.7	298	1693
Counties				Ave	7.7			
State	8,881,845	88,137		Ave	9.9			