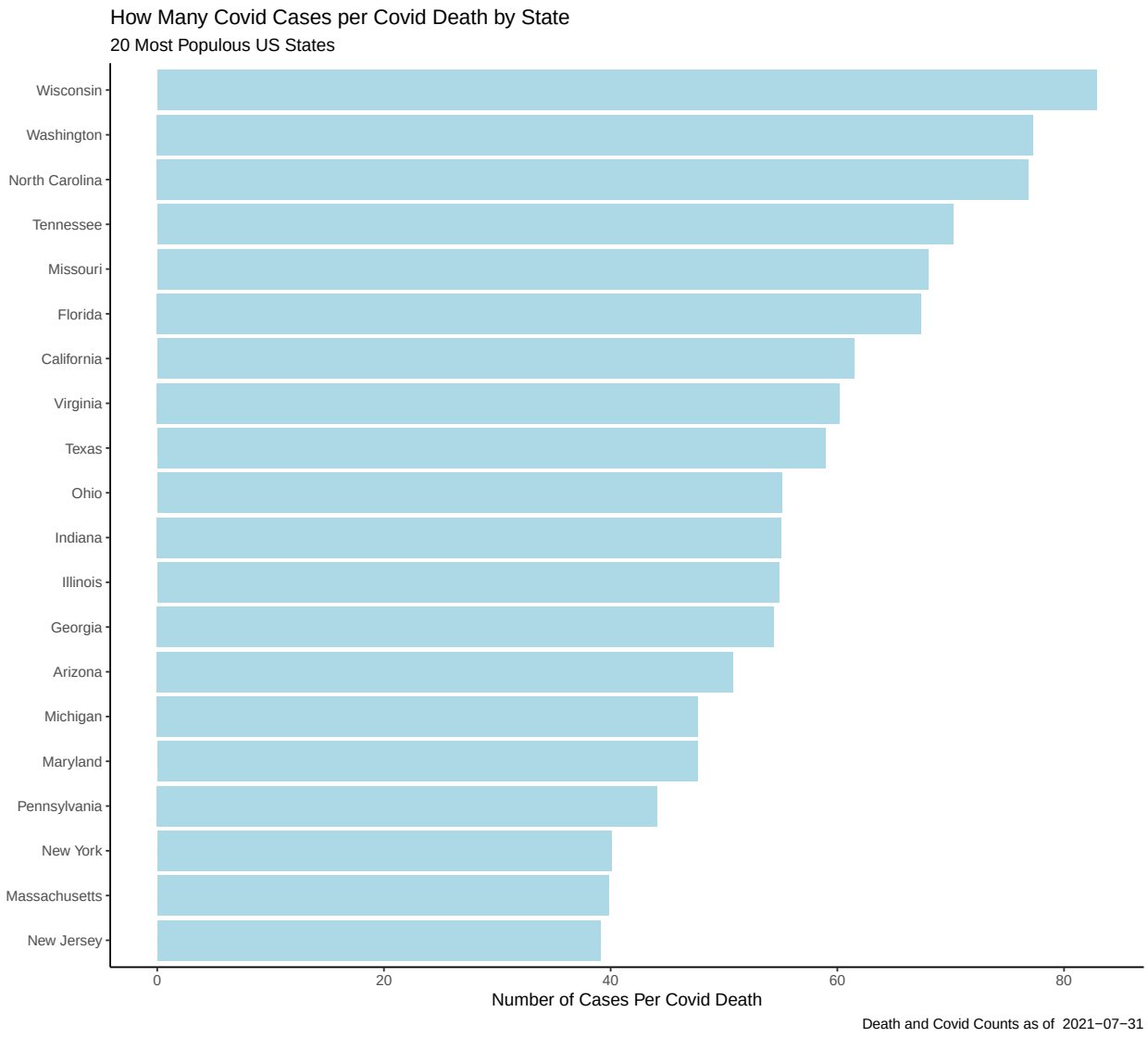
Final Project - COVID 19 DATA

US Analysis by State

How Many Covid Cases Does it Take Until a Death is Likely?



Province_State	Total_Cases	Total_Deaths	Cases_Per_Death
Alaska	75486	389	194.05141
Utah	432467	2451	176.44512
Virgin Islands	4629	37	125.10811
Nebraska	228450	2280	100.19737
Vermont	24889	260	95.72692
Northern Mariana Islands	183	2	91.50000
Idaho	200484	2197	91.25353
Wyoming	65127	776	83.92655
Wisconsin	687198	8290	82.89481
Colorado	575082	6945	82.80518
Hawaii	42410	537	78.97579
Minnesota	612701	7761	78.94614
Maine	70463	900	78.29222
Washington	473076	6122	77.27475
Oregon	219755	2858	76.89118
North Carolina	1048076	13635	76.86659
New Hampshire	100657	1387	72.57174
North Dakota	111644	1570	71.11083
Tennessee	893554	12724	70.22587
Montana	116366	1704	68.28991
Missouri	686547	10099	67.98168
Florida	2634234	39079	67.40792
Kentucky	482599	7334	65.80297
Oklahoma	480635	7485	64.21309
Kansas	333850	5261	63.45752
Arkansas	386452	6141	62.92982
South Carolina	617148	9904	62.31300
California	3961190	64425	61.48529
Iowa	378831	6183	61.26977
South Dakota	125225	2045	61.23472
Delaware	111407	1833	60.77851
Nevada	356401	5912	60.28434
Virginia	694384	11532	60.21367
Guam	8541	143	59.72727
Texas	3138837	53277	58.91542
Puerto Rico	146801	2580	56.89961
West Virginia	167016	2946	56.69246
Rhode Island	154339	2740	56.32810
Ohio	1129277	20492	55.10819
Indiana	771299	14005	55.07312
Illinois	1419611	25873	54.86843
Georgia	1179449	21676	54.41267
Arizona	927235	18246	50.81854
Alabama	585607	11536	50.76344
Louisiana	541679	10999	49.24802
Michigan	1011106	21188	47.72069
New Mexico	210416	4410	47.71338
Maryland	468548	9825	47.68936
Mississippi	343505	7543	45.53957
Pennsylvania	1229002	27850	44.12934
District of Columbia	50398	1149	43.86249
Connecticut	354335	8293	42.72700

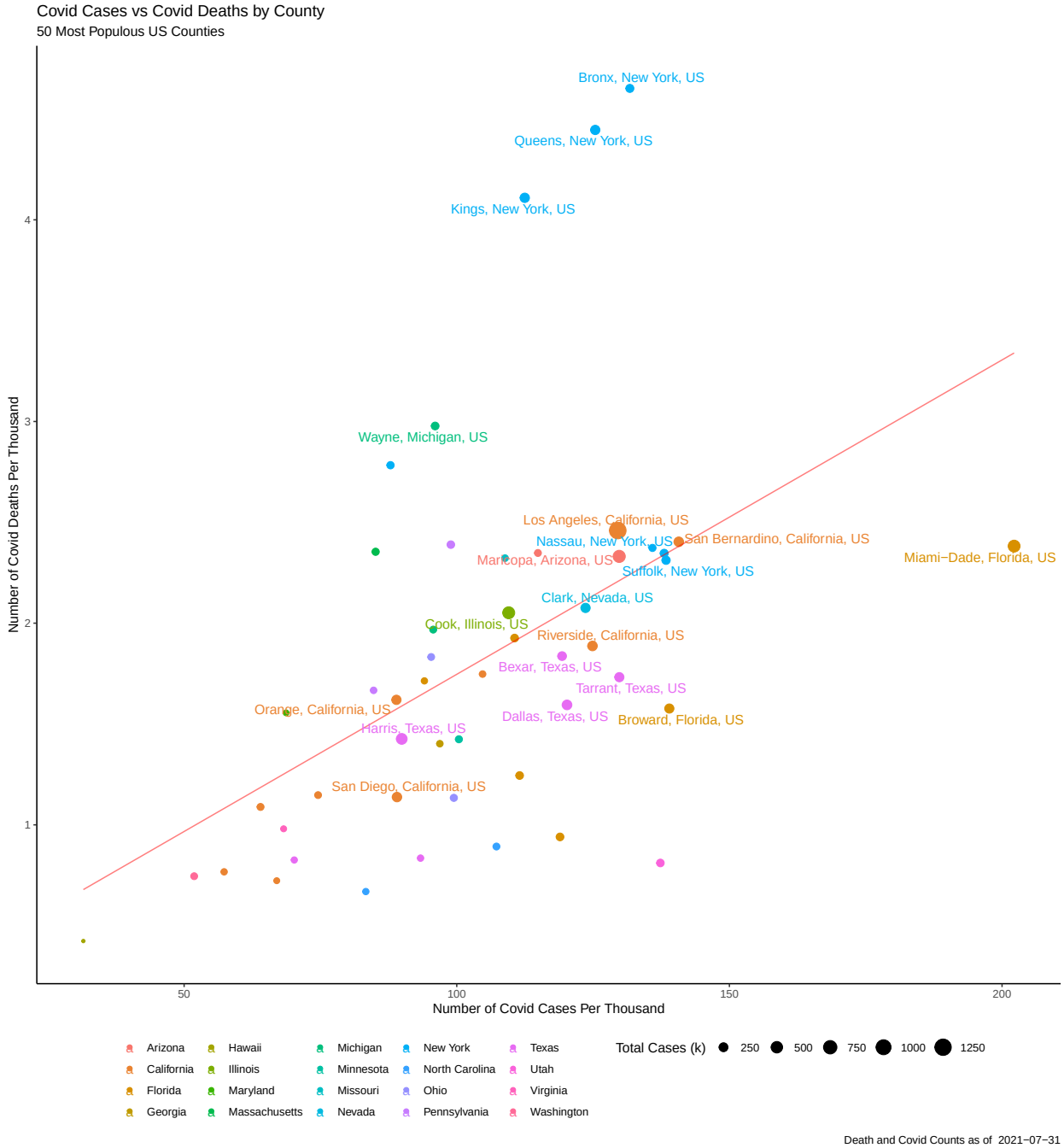
Province_State	Total_Cases	Total_Deaths	Cases_Per_Death
New York	2150390	53641	40.08855
Massachusetts	719780	18082	39.80644
New Jersey	1039353	26602	39.07048
Grand Princess	103	3	34.33333

Covid Cases and Deaths by Most Populated Counties

Modeling the Number of Deaths (Per Thousand) as a Function of Cases (Per Thousand)

```
##
## Call:
## lm(formula = Deaths_Per_Thou ~ Cases_Per_Thou, data = county_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5176 -0.4515 -0.0846  0.1588  2.4105
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.187589   0.420986   0.446 0.657893
## Cases_Per_Thou 0.015588   0.003901   3.996 0.000221 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7999 on 48 degrees of freedom
## Multiple R-squared:  0.2496, Adjusted R-squared:  0.234
## F-statistic: 15.97 on 1 and 48 DF,  p-value: 0.000221
```

Which Counties had more Deaths Per Covid Case?



Final Analysis

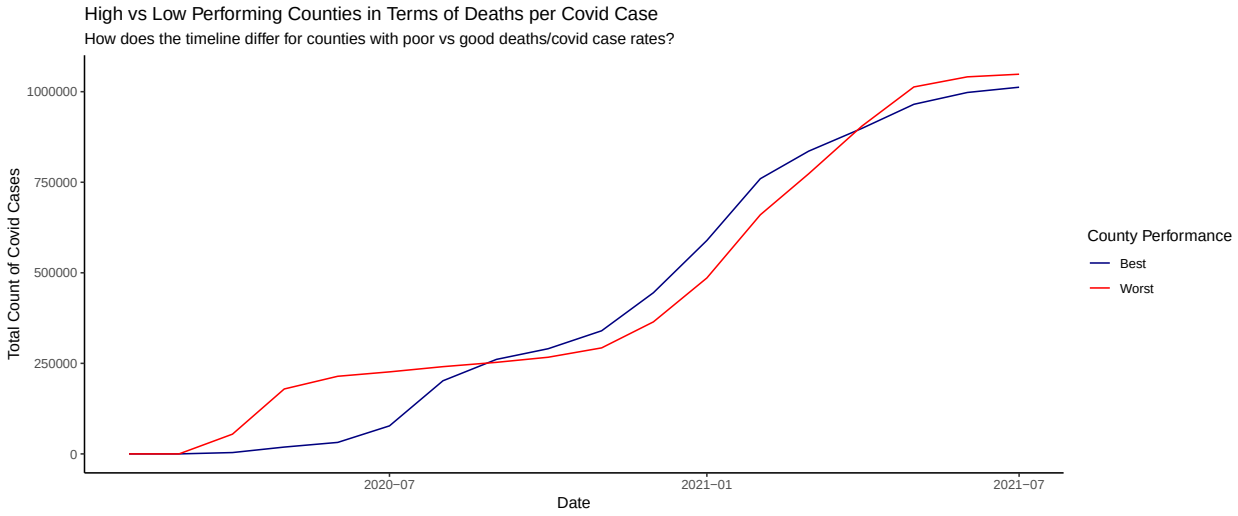
Digging Into County Data

Table 2: Counties Performing Best in Terms of Death Per Covid Case

County	Cases Per Thousand	Deaths Per Thousand	Predicted Deaths
Salt Lake, Utah, US	137.34223	0.8109014	2.328462
Orange, Florida, US	118.93628	0.9401113	2.041553
Mecklenburg, North Carolina, US	107.28001	0.8925065	1.859856
Miami-Dade, Florida, US	202.21462	2.3820916	3.339685
Wake, North Carolina, US	83.32276	0.6692086	1.486414

Table 3: Counties Performing Worst in Terms of Death Per Covid Case

County	Cases Per Thousand	Deaths Per Thousand	Predicted Deaths
Bronx, New York, US	131.74240	4.651648	2.241173
Queens, New York, US	125.39743	4.445710	2.142268
Kings, New York, US	112.46989	4.109140	1.940755
Wayne, Michigan, US	96.03434	2.977689	1.684560
New York, New York, US	87.84643	2.783191	1.556928



Difference in Timelines This final plot shows drastic differences in the volume of Covid 19 cases for counties who had low deaths per case (blue) and high deaths per case (red).

It appears counties who had high number of deaths per case reached higher volumes of cases earlier in the pandemic. This suggests that treatment was not as effective against positive Covid 19 patients early on in the pandemic versus later on. Therefore, it is likely that having a low rate of deaths per Covid case was not so much a measure of success within the county as much as it was at what stage in the pandemic did cases spike within the county.

Bias Areas

I think the conclusion that I suggested in the above paragraph was a belief that I suspected prior to doing this analysis. I attempted to combat my bias by questioning my belief head on to see if my suspicions were correct. However, there are likely numerous factors in play other than when cases spiked that attribute to high death per case rate. Perhaps vaccination rates by state or county would affect these rates later on in the pandemic. The age, race and underlying population make-up of a location is likely a big factor on the death per case rate as well.

One other belief I held prior to this analysis was that Florida would have higher rates of death per case. This is because of my belief that the age of a typical Florida resident skews older than the age of a typical American. However, when looking at the rates of death per case by state and most populous counties this actually turned out to not be the case. Overall, Florida had one of the better deaths per case ratios amongst heavily populated states. This is most evident in the first visual on page 1.