Comparison of State Performance Across Periods

Instead of visualizing the data with graphs (due to the large number of states, 52 in total), the following method was used to determine how states performed during different periods:

- Difference in death rates between periods was calculated for each state:
  - Difference1 = Death rate in Period2 - Death rate in Period1
  - Difference2 = Death rate in Period3 - Death rate in Period2

Issue with Alaska (AK) Death Data in Period 3

For Alaska (AK), the data in Period 3 contains numerous NA values, with only 3 data points being zero. This led to an outcome where the calculated death rate for Alaska in Period 3 resulted in 0 deaths.

67	44	2022	*NA*	AK	Wave_3	Alaska	733276	10	Pacific Northwe
68	45	2022	0	AK	Wave_3	Alaska	733276	10	Pacific Northwe
69	46	2022	*NA*	AK	Wave_3	Alaska	733276	10	Pacific Northwe
70	47	2022	*NA*	AK	Wave_3	Alaska	733276	10	Pacific Northwe
71	48	2022	0	AK	Wave_3	Alaska	733276	10	Pacific Northwe
72	49	2022	*NA*	AK	Wave_3	Alaska	733276	10	Pacific Northwe
73	50	2022	*NA*	AK	Wave_3	Alaska	733276	10	Pacific Northwe
74	51	2022	*NA*	AK	Wave_3	Alaska	733276	10	Pacific Northwe
75	52	2022	*NA*	AK	Wave_3	Alaska	733276	10	Pacific Northwe
76	1	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe
77	2	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe
78	3	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe
79	4	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe
80	5	2023	0	AK	Wave_3	Alaska	733406	10	Pacific Northwe
81	6	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe
82	7	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe
83	8	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe
84	9	2023	*NA*	AK	Wave_3	Alaska	733406	10	Pacific Northwe

The issue with Alaska's death data in Period 3 is likely caused by a combination of data collection challenges, small population size, or even weather issue. Further investigation into the data collection process and reporting systems would help clarify these problem.

Death Rate Formula

I use average death rate per 100,000 population per month for the death rate. We assume that the population remains relatively constant over a year. The calculation involves the following steps:

1. Single-Year Period:
   If a period occurs entirely within one year, the average population will simply equal the population of that year. For example, if the population in 2020 is constant at P, the average population is:
   $$\text{avg\_population} = P$$
2. Two Years Period:
   If a period spans two years, the average population will automatically be weighted by the length of time spent in each year (e.g., more weeks in one year contribute

more to the average). For example, P1 is population in year 1 and P2 in year 2. Let n1 and n2 be the number of weeks in year 1 and year 2, the average population is:

avg_population = (P1 * n1 + P2 * n2) / (n1 + n2)

3. Average Death:
   After having average population, we can calculate the average death in a period by:
   avg_death_rate = (Total death / avg_population) × 100000 ÷ months in a period

Data Aggregation

The cases and death data were aggregated by summing the total number of cases and deaths for each MMWR week to ensure consistency in reporting intervals. This weekly aggregation allowed for clearer trend analysis and visualization of the data over time. The resulting processed data were stored in two datasets: national_data for aggregated case counts and national_deaths for aggregated death counts. These datasets were then used to generate plots for further analysis.

Trend Plot in Month

Here is the monthly trend of the cases, which to help better select pandemic period.