Analysis of Current Death Counts and Leading Causes in the US, Focusing on COVID-19

Yumeng Gao

2022-12-07

Introduction

COVID-19 has gradually became one of the leading causes of death in the US since its outbreak in 2020. According to US Center for Disease Control and Prevention (CDC), the top 10 leading causes of death now are: heart disease, cancer, COVID-19, accidents (unintentional injuries), stroke (cerebrovascular diseases), chronic lower respiratory diseases, Alzheimer's disease, diabetes, influenza and pneumonia, as well as nephritis, nephrotic syndrome, and nephrosis. Moreover, COVID-19 may interact with other diseases to impair the health conditions of its hosts, eventually leading to death. Thus, it is pivotal to draw a picture of COVID-19-attributed mortality with other leading causes of death, to quantify the influence of this pandemic. This study focused on the overall death counts in the US and leading causes, then narrowed down to COVID-caused death counts in different age, sex, and race groups. The primary research questions are has the pandemic affected the mortality trends in the US and how is COVID-caused mortality related to sex, age, and race?

Methods

The original dataset of this study was exported from CDC's website: AH Monthly Provisional Counts of Deaths for Select Causes of Death by Sex, Age, and Race and Hispanic Origin (https://data.cdc.gov/NCHS/AH-Monthly-Provisional-Counts-of-Deaths-for-Select/65mz-jvh5) as a csv format.

There were 3960 observations for death counts, including 14 major causes of death from Jan 2019 to Sep 2021 (33 months in total). After reading in the dataset, the major elements and the key variables (leading causes of death) were checked for missing values and outliers. Variables with long names were renamed for convenience. Then we took a closer look at the key variables (leading causes of death) to check for missing values and outliers. For each categorical variables, the name and order of categories were checked and corrected if necessary.

After data cleaning, we got 1980 males and 1980 females (1:1) and ten Age categories from "0-4 years" to "85 years and over". Race were divided into Hispanic, American Indian or Alaska Native, Asian, Black, White, and other. Luckily, Year, Total Month, Sex, Age, and Race all have same sample size for each category, so the death counts among groups could also be considered as the proportions, eligible to be compared directly.

The key variables (leading causes) were:

1. Natural Cause,
2. Septicemia,
3. Malignant neoplasms,

4. Diabetes mellitus,
5. Alzheimer disease,
6. Influenza and pneumonia,
7. Chronic lower respiratory diseases,
8. Other diseases of respiratory system ,
9. Nephritis, nephrotic syndrome and nephrosis,
10. Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified,
11. Diseases of heart,
12. Cerebrovascular diseases,
13. COVID-19 (Multiple Cause of Death),
14. COVID-19 (Underlying Cause of Death)

Since the death counts were wide-ranged integers, averages of death counts were calculated to better interpret the data. We first summarized the average death counts among different causes by year, then generated a scatter plot to visualize the tendency of death counts by leading causes form 2019 to 2021. To get more details, the average death counts of different causes by total 33 months were also created, following with paired scatter plot. To emphasize COVID-19 multiple and underlying causes, we plotted the related trends of death counts by year and by total month.

Then for the analysis narrowing down to COVID-19 multiple and underlying causes, grouping by sex, age, and race groups. A summary table along with bar charts were generated for each these three categorical variables to present the difference of COVID-caused mortality among groups.

Results

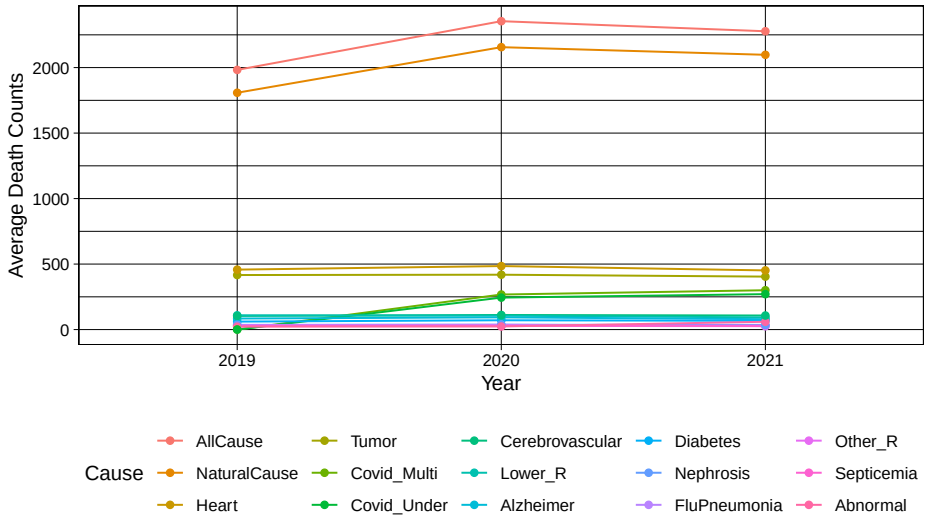
1. Overall Mortality Trends by Year and Total Month

Table 1. Average Death Counts for Leading Causes by Year

	2019	2020	2021
AllCause	1982.42	2354.16	2277.36
NaturalCause	1807.87	2155.91	2097.53
Septicemia	26.69	27.86	26.54
Tumor	416.38	418.81	404.31
Diabetes	60.86	71.04	66.95
Alzheimer	84.37	93.25	79.40
FluPneumonia	34.57	37.27	26.50
Lower_R	109.01	106.09	93.01
Other_R	30.63	31.37	30.07
Nephrosis	35.81	36.53	35.71
Abnormal	22.48	23.77	59.58
Heart	457.64	484.86	451.73
Cerebrovascular	104.17	111.47	107.40
Covid_Multi	0.00	267.54	300.28
Covid_Under	0.00	244.10	270.30

Table 1 presented the average death counts among different causes by year. Except for two COVID-19 causes, all other factors remained similar from 2019 to 2021. COVID-19-attributed mortality started at 2020 and increased slightly in 2021.

Figure 1. Average Death Counts for Leading Causes by Year



The top 3 causes of death were: natural cause, heart disease, and malignant neoplasms. Except the drastic increase of COVID-caused death counts in 2020, all the other causes' death counts have remained relatively stable.

Figure 2. Average Death Counts for Leading Causes by Total Month

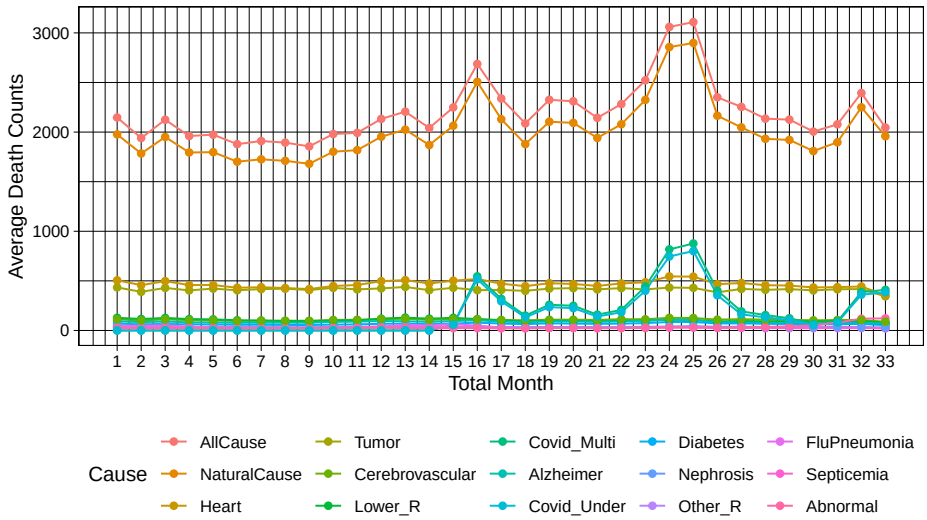


Figure 2 presented the detailed tendency of all 15 kinds of leading causes during the 33 month. The trend of natural cause line were similar to the all causes line and the difference between their death counts at all months were pretty small, indicating that natural cause was the most major cause of death. Still, Multiple and underlying COVID-19 causes showed a different pattern from all other causes.

2. COVID-Caused Mortality Trends

Figure 3. Average Death Counts for COVID by Year

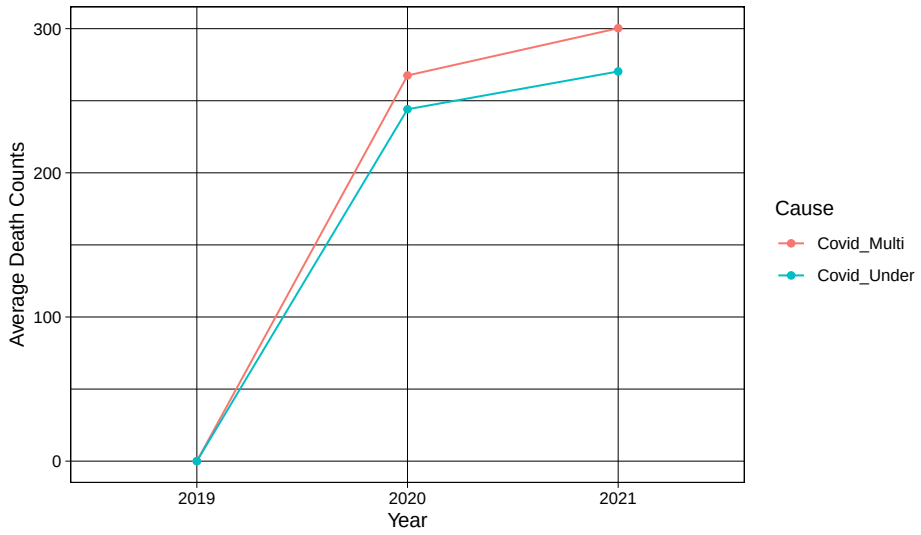
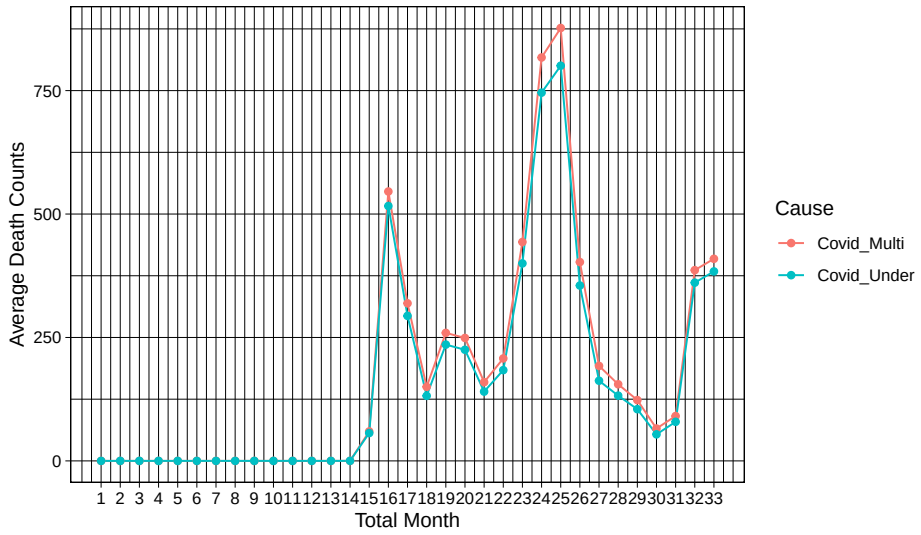


Figure 4. Average Death Counts for COVID by Total Month



Narrowing down to COVID-19 causes, before Month 14 (Feb 2020), there was no COVID-19-attributed mortality in the US. Then for both causes, average death counts started to increase from Feb 2020, the highest peak was at Month 25 (Jan 2021).

Since there was no COVID-caused death data in 2019, following interpretation would be focused on data of 2020 and 2021.

3. COVID-caused Average Death Counts by Sex

Table 2. Summary of COVID-caused Average Death Counts by Sex

Year	Sex	AllCause	Covid_Multi	Covid_Under
2019	F	1918.02	0.00	0.00
2019	M	2046.83	0.00	0.00
2020	F	2243.90	244.45	220.13
2020	M	2464.41	290.63	268.06
2021	F	2137.97	264.16	234.04
2021	M	2416.74	336.40	306.57

Figure 5. Average Counts of Death for Multiple COVID−19 Cause by Sex

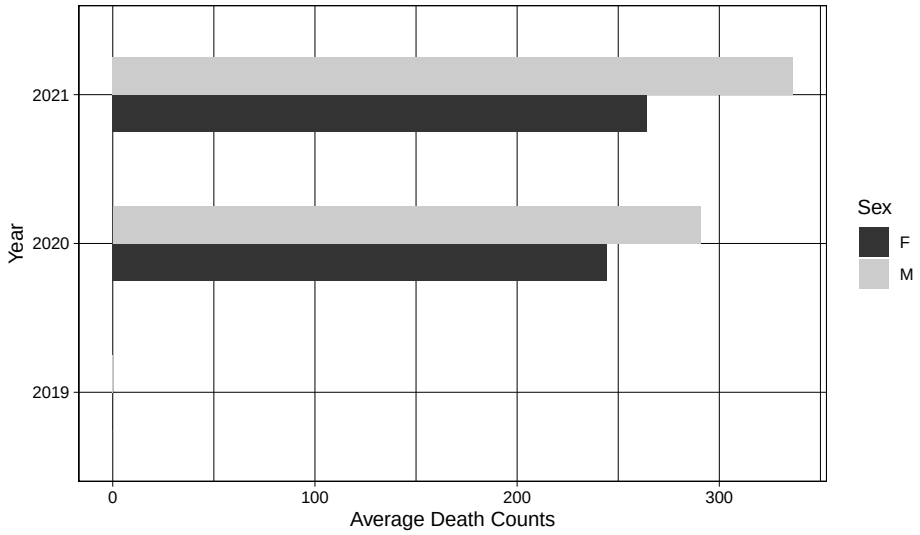
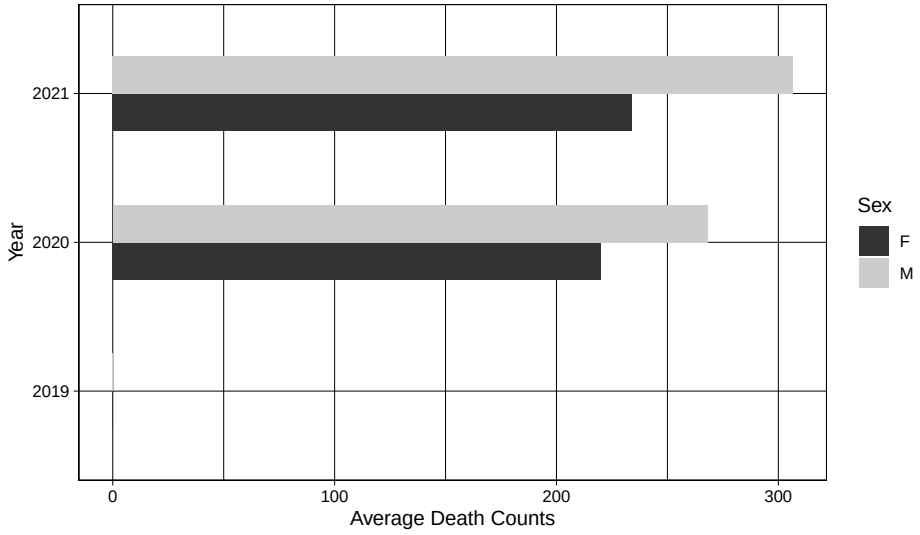


Figure 6. Average Counts of Death for Underlying COVID−19 Cause by Sex



Females had smaller COVID-caused average death counts than males in 2020 and 2021, indicating that men might have higher risk of dying from COVID-19 than women.

4. COVID-caused Average Death Counts by Age

Table 3. Summary of COVID-caused Average Death Counts by Age

Year	AgeGroup	AllCause	Covid_Multi	Covid_Under
2019	0-4	170.81	0.00	0.00
2019	5-14	38.17	0.00	0.00
2019	15-24	206.74	0.00	0.00
2019	25-34	410.96	0.00	0.00
2019	35-44	576.29	0.00	0.00
2019	45-54	1113.84	0.00	0.00
2019	55-64	2603.73	0.00	0.00
2019	65-74	3858.05	0.00	0.00
2019	75-84	4777.97	0.00	0.00
2019	>=85	6067.68	0.00	0.00
2020	0-4	161.00	0.53	0.38
2020	5-14	39.19	0.47	0.34
2020	15-24	250.10	4.26	3.48
2020	25-34	513.08	18.19	15.71
2020	35-44	729.06	47.07	42.34
2020	45-54	1333.22	127.19	118.23
2020	55-64	3069.50	316.33	293.43
2020	65-74	4693.92	571.19	530.95
2020	75-84	5715.49	737.53	675.03
2020	>=85	7037.00	852.63	761.08
2021	0-4	149.31	1.01	0.67
2021	5-14	38.60	0.90	0.62
2021	15-24	252.42	9.06	7.71
2021	25-34	538.61	39.81	35.12
2021	35-44	805.20	98.84	91.19
2021	45-54	1402.75	235.61	220.88
2021	55-64	3122.28	489.81	454.97
2021	65-74	4747.55	717.71	660.58
2021	75-84	5464.61	746.86	667.22
2021	>=85	6252.27	663.20	564.07

Figure 7. Average Counts of Death for Multiple COVID−19 Cause by Age

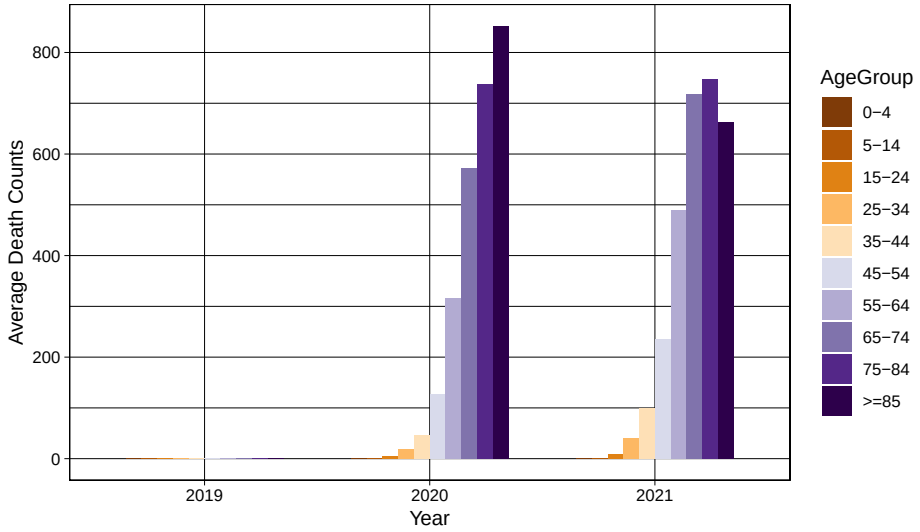
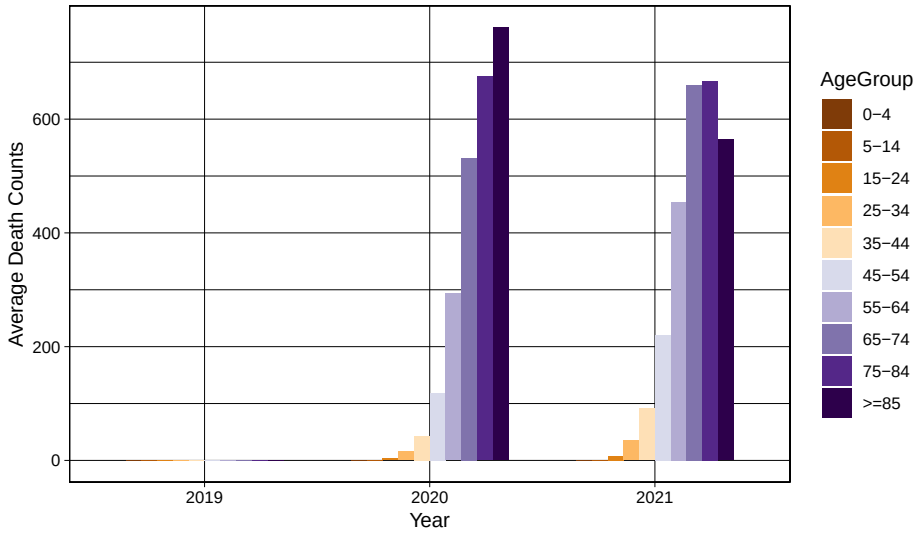


Figure 8. Average Counts of Death for Underlying COVID−19 Cause by Age



The overall pattern was: COVID-caused average death counts increased with age. In 2021, the equals to or older than 85 years old group got smaller counts than 75-84 years old group, which required future analysis.

5. COVID-caused Average Death Counts by Race

Table 4. Summary of COVID-caused Average Death Counts by Race

Year	Race	AllCause	Covid_Multi	Covid_Under
2019	Asian	293.88	0.00	0.00
2019	Black	1444.41	0.00	0.00
2019	Hispanic	884.95	0.00	0.00
2019	Indian/Alaska	75.24	0.00	0.00
2019	Other	96.95	0.00	0.00
2019	White	9099.13	0.00	0.00
2020	Asian	381.63	56.55	53.08
2020	Black	1875.12	256.14	235.21
2020	Hispanic	1281.44	289.31	273.25
2020	Indian/Alaska	103.11	19.23	17.78
2020	Other	122.26	14.26	13.18
2020	White	10361.38	969.75	872.09
2021	Asian	378.58	62.75	58.54
2021	Black	1761.60	253.45	229.17
2021	Hispanic	1290.42	314.93	296.88
2021	Indian/Alaska	100.42	17.62	16.15
2021	Other	120.27	15.23	13.81
2021	White	10012.87	1137.72	1007.28

Figure 9. Average Counts of Death for Multiple COVID−19 Cause by Race

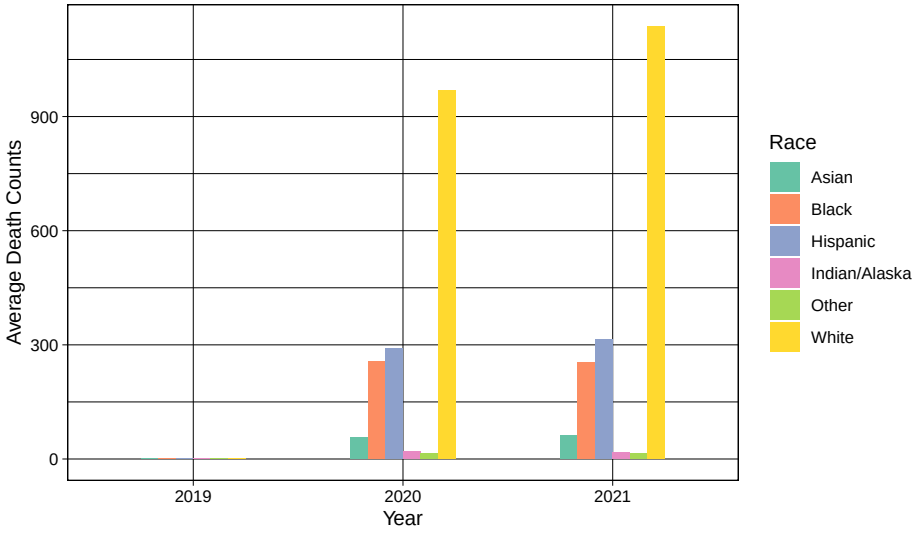
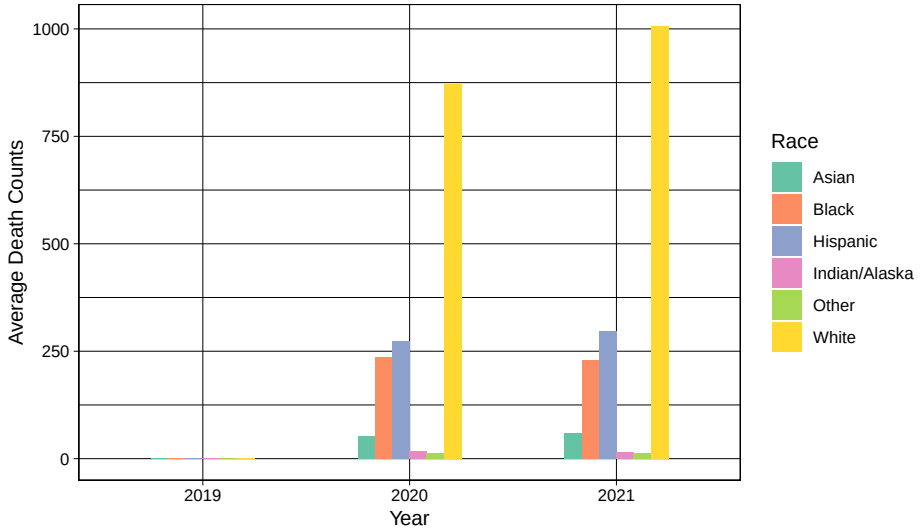


Figure 10. Average Counts of Death for Underlying COVID−19 Cause by Race



We can find the same pattern in 2020 and 2021. White people got the highest average death counts, while Indian/Alaska Native had smallest mortality (excluding Other group). From small to large average death counts, the order is: Other, Indian/Alaska Native, Asian, Black, Hispanic, White.

Conclusion and Summary

This research of mortality comparison by different causes first showed that except for two COVID-19 causes, all other leading causes of death had stable trends from 2019 to 2021. The top 3 leading causes were natural cause, heart disease, and tumor. The COVID-caused mortality started at Feb 2020 and reached it peak at Jan 2021. For different sex, age, and race groups, there were obvious difference of average COVID-caused death counts. Females tended to had smaller average death counts than males, elder people tended to had higher average mortality than young people. White people got the highest average death counts, while Indian/Alaska Native had smallest mortality. In conclusion, COVID-19 has obviously affected the mortality trend in the US.