From Normalcy to Pandemic (2017-2020): Analyzing Life Expectancy Trends and Their Correlations in the United States, with a Case Study on COVID-19's Impact in Texas

In the realm of public health, life expectancy serves as a vital barometer, offering insights into a population's overall health and well-being. The study delves into the intricate tapestry of life expectancy trends in the United States from 2017 through 2020—a period marked by significant societal and health transitions. It meticulously analyzes these trends, both at national and state levels, underlining the profound implications of the COVID-19 pandemic, especially in the state of Texas.

Motivated by the urgent need to understand the pandemic's effects on life expectancy, the research provides a dual perspective. It encompasses a broad analysis of national trends and a focused case study of Texas, a state with unique demographic and characteristics. The question driving the investigation is: "How has the COVID-19 pandemic affected life expectancy in the United States, and what are its specific impacts in Texas?"

Public data from reputable sources are utilized in the analysis, which employs robust statistical methods such as paired t-tests, Pearson correlation coefficients, and regressions. The exploration focuses on the stark changes in life expectancy and its correlations with various social and health determinants. In Texas, the localized impact of the pandemic is observed. This state-specific investigation offers a nuanced understanding of how different regions within the U.S. have been uniquely affected, highlighting the need for tailored public health strategies.

In summary, the paper, through its rigorous quantitative analysis, aims to contribute significantly to the ongoing discourse on public health, particularly in the context of major crises like the COVID-19 pandemic. It seeks to inform policy decisions and public health interventions, fostering a better understanding of the intricate relationship between life expectancy and various

social and health factors. As we step into a post-pandemic world, the insights gleaned from this research could prove instrumental in shaping a resilient and health-optimized future.

Method

Life Expectancy Trends at the National and State Levels

Life expectancy data is utilized for a longitudinal analysis, tracking changes in life expectancy at both national and state levels. Paired t-tests are adopted to assess the significance of year-on-year changes. This method aids in pinpointing the precise moments of significant shifts in life expectancy, particularly highlighting the temporal impact of the COVID-19 pandemic.

Life Expectancy Correlations at the National Level

Investigating the relationships between life expectancy and various factors involves categorizing these factors into six groups. This categorization is based on and extends the Social Determinants of Health (SDOH) framework. The six categories are: Economic Stability, Education and Youth Engagement, Healthcare Access and Quality, Health Behaviors and Conditions, Neighborhood and Built Environment, and Social and Community Context. To quantify the strength and direction of these associations, Pearson’s correlation coefficient is used, providing insights into the complex interplay of factors that influence life expectancy.

Case Study - COVID-19's Impact in Texas

The study focuses on a detailed case analysis of Texas, offering insights into the localized effects of the COVID-19 pandemic. Segmented regression is used to analyze the impact of the COVID-19 pandemic on life expectancy in Texas. Additionally, the research features a comparative analysis of various factors before and after the pandemic's onset.

Data Sources and Description

The study delves into an analysis of life expectancy and its associated metrics, focusing on both national and state levels. The core of this research is anchored in an extensive dataset spanning from 2017 to 2020, as detailed in Table A1 of Appendix A. This dataset not only includes life expectancy figures but also encompasses a diverse range of indicators, categorized into six distinct areas. The data are sourced from the County Health Rankings & Roadmaps by the University of Wisconsin Population Health Institute, although the data originate from various external sources. For the Texas-specific data, it's important to note that it incorporates additional indicators, such as demographic variables, thereby providing a more comprehensive understanding of the state's health landscape. The six categories, along with life expectancy, used for national correlation analysis are listed below:

- Life Expectancy: The life expectancy is calculated using a three-year average, drawing on data from the National Vital Statistics System's Mortality Files. This method provides an age-adjusted measure, which is grounded in current mortality rates. It takes into consideration the number of deaths as well as the population at risk.
- Economic Stability: Economic factors such as median household income, income ratio, unemployment, and children in poverty are derived from the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE), American Community Survey (ACS) 5-year estimates, and Bureau of Labor Statistics.
- Education and Youth Engagement: English proficiency, education (% Some College), disconnected youth are derived from the American Community Survey (ACS), Bureau of Labor Statistics, and other census data.
- Healthcare Access and Quality: Data on preventable hospitalizations and screenings (mammogram and flu vaccinations) from the Mapping Medicare Disparities Tool is

integrated. Uninsured rates for adults and children from the U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE) are also included.

- Health Behaviors and Conditions: This category includes various health measures such as obesity, diabetes, physical inactivity, adult smoking, frequent mental distress, excessive drinking, drug overdose mortality, low birthweight (based on a 7-year average), teen births, chlamydia, and HIV prevalence. These are assessed through data collection systems like the CDC Behavioral Risk Factor Surveillance System (BRFSS), the National Center for Health Statistics' National Vital Statistics System, and the National Center for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP).
- Neighborhood and Built Environment: Data on long commutes, the population living in rural areas, and food environment are from the American Community Survey (ACS) 5-year estimates, Census Population Estimates, USDA Food Environment Atlas, and Feeding America's Map the Meal Gap.
- Social and Community Context: Children in single-parent households, social associations, homicide, firearm fatality, injury death are included to account for social and community context.

Data Preprocessing

Data preprocessing involves several steps: Excluding measures that have more than 20% missing data and those with calculation changes during the research years, combining data, checking for accuracy and comparability, and addressing the remain missing values through median imputation.

Statistical Tools and Software

In the study, statistical analysis and data visualization were conducted using Python, specifically through Jupyter Notebooks provided by the Anaconda distribution, leveraging its extensive libraries and tools for robust and efficient data handling.

Ethical Considerations

All data used in the study are secondary and publicly available.

Results

Life Expectancy Trends at the National and State Levels

Nationally, life expectancy experienced a slight increase from 2017 to 2019, followed by a sharp decline in 2020, coinciding with the onset of the COVID-19 pandemic. At the state level, data reveal varied outcomes. For instance, Hawaii has the highest life expectancy, while Mississippi has the lowest. The gap between them is approximately 7.64 years, highlighting the importance of region-specific health strategies.

Figure 1

Life Expectancy Trend in the U.S. (2017-2020)

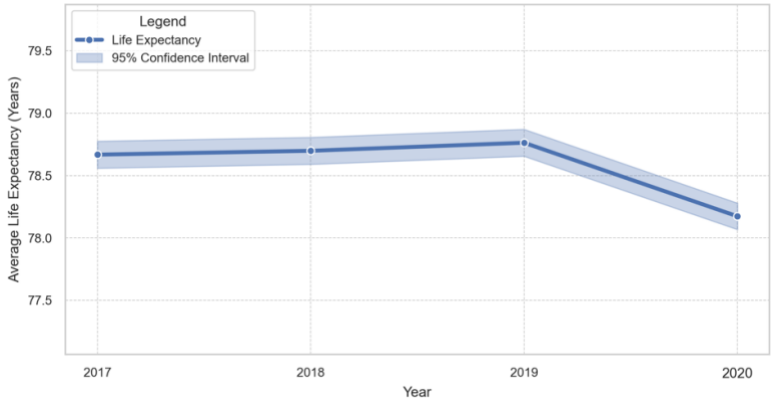
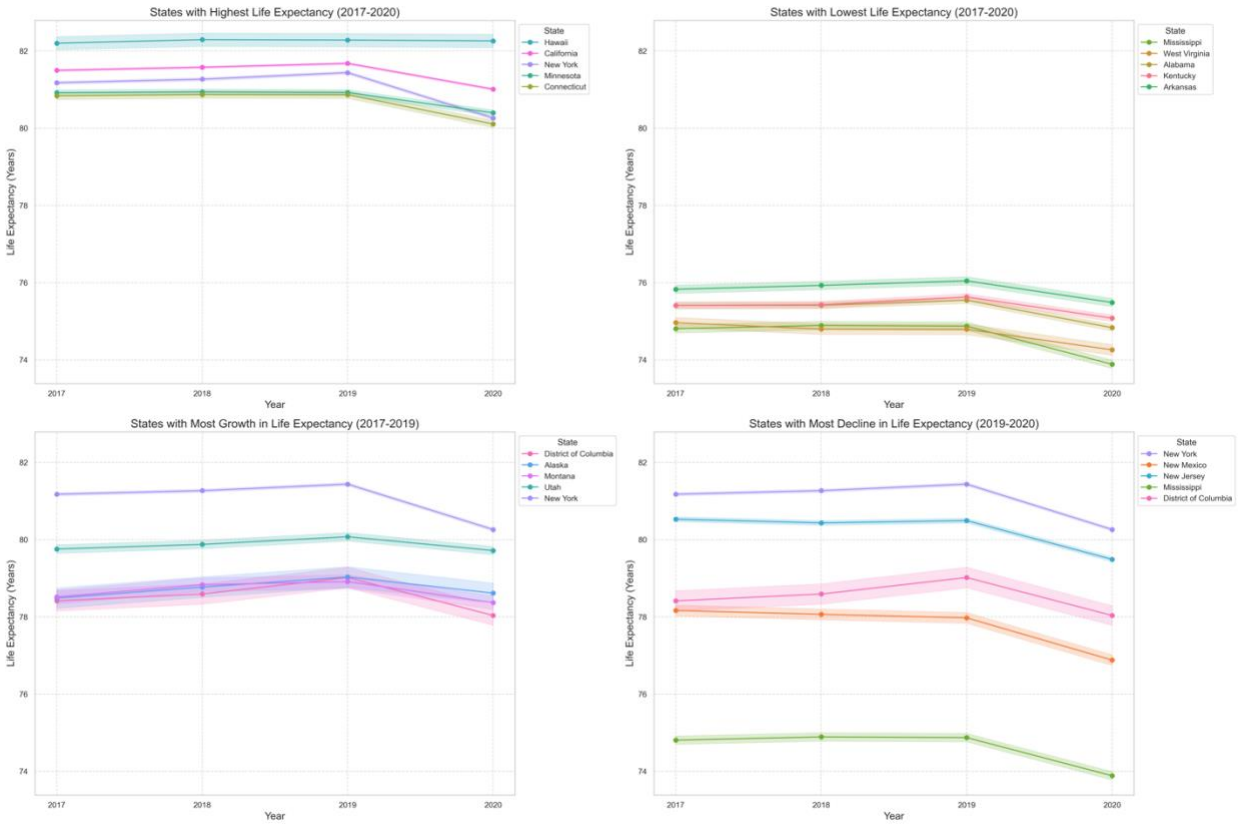


Figure 1 illustrates the national trend in age-adjusted life expectancy, calculated as a three-year average, from 2017 to 2020. Utilizing paired t-tests, the data indicate that life expectancy initially increased marginally, from an average of 78.67 years in 2017 to 78.70 years

in 2018 ($p = 0.0273$), and then to 78.76 years in 2019 ($p < 0.001$); however, a notable decline occurred in 2020, with life expectancy dropping to 78.17 years ($p < 0.001$). Consequently, the cumulative change in life expectancy across this four-year span amounted to an overall reduction of approximately 0.63%. This trend coincides with the global COVID-19 pandemic, which, as detailed in Chan, Cheng, and Martin's (2021), has had a substantial impact on life expectancy.

Figure 2

Life Expectancy Trends Across U.S. States (2017-2020)



A state-level analysis of life expectancy trends in the United States from 2017 to 2020 reveals significant variations (Figure 2). Hawaii exhibited the highest average life expectancy at 82.26 years, followed by California and New York. In contrast, the lowest was recorded in Mississippi at 74.62 years, alongside West Virginia and Alabama. The disparity between the highest and lowest life expectancies was approximately 7.64 years. Temporal changes in life

expectancy were also examined and supported by paired t-tests. From 2017 to 2019, the District of Columbia and Alaska experienced the most notable increases in life expectancy, with increments of 0.60 and 0.55 years, respectively. However, a marked decrease in life expectancy occurred in all states during the 2019 to 2020 period, coinciding with the onset of the COVID-19 pandemic. Notably, New York and New Mexico experienced the sharpest declines, losing 1.18 and 1.09 years respectively. These varying trends in life expectancy emphasize the need for region-specific health interventions.

Figure 3

Annual Changes in Life Expectancy Across the U.S. (2017-2020)

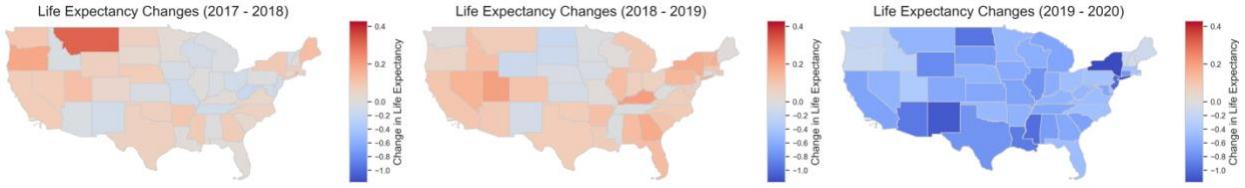


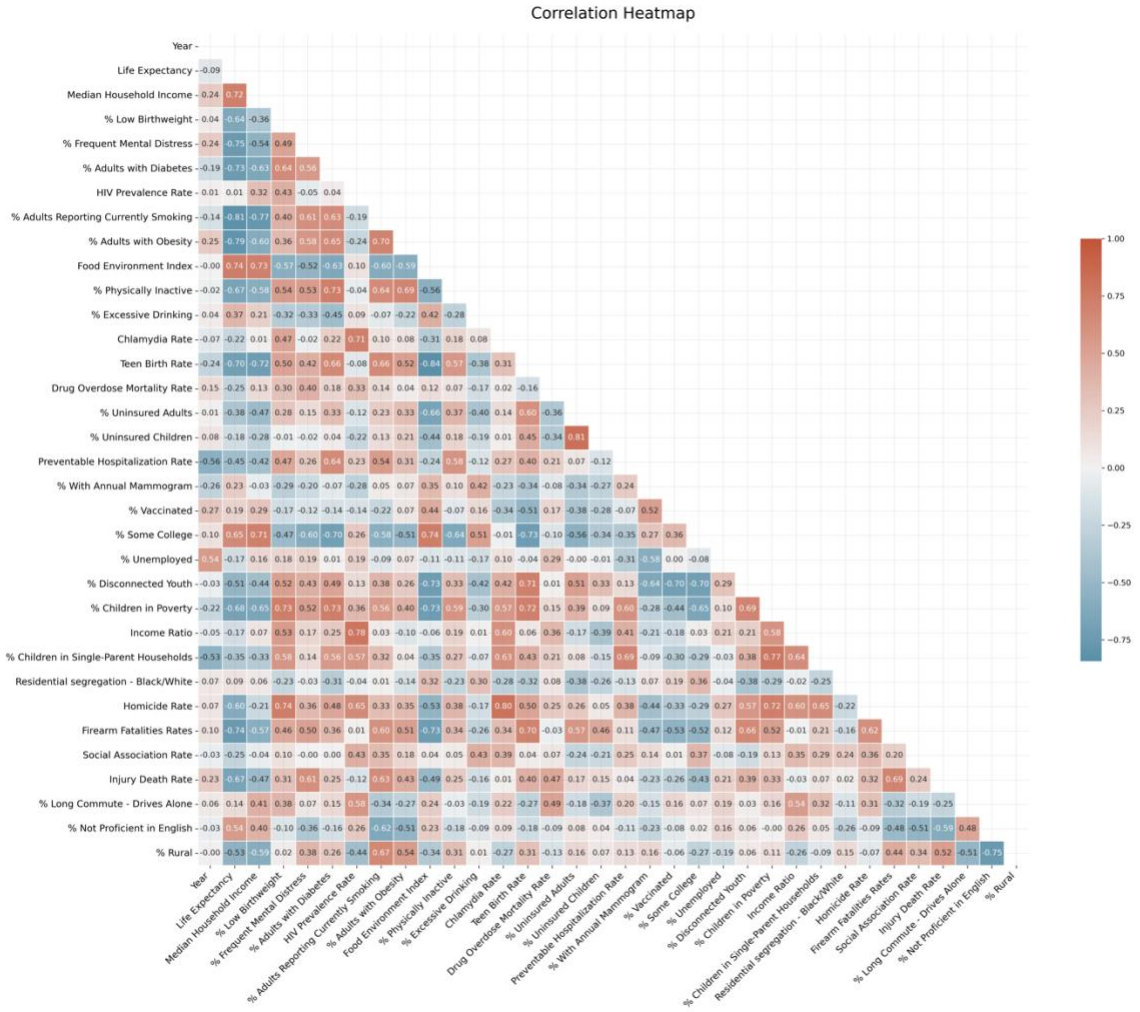
Figure 3 visualizes annual life expectancy fluctuations across the United States from 2017 to 2020. Between 2017 and 2018, most portion of the country experienced marginal increases in life expectancy, with exceptions in the central northern states, which registered slight declines. In 2018-2019, the trends were more varied, with both increases and decreases observed nationwide; the northeastern region largely showed a trend of improvement. The 2019-2020 map markedly contrasts with earlier years, displaying widespread reductions in life expectancy, especially in the southern and eastern states. This trend may reflect the impacts of significant health events, such as the COVID-19 pandemic, or policy changes during that period. These maps collectively underscore the variability and complexity of life expectancy changes over time, influenced by a multitude of regional factors and policies.

Life Expectancy Correlations at the National Level

The correlations between life expectancy in the United States (2017-2020) and factors in Figure 4 are explored across six domains: Economic Stability, Education and Youth Engagement, Healthcare Access and Quality, Health Behaviors and Conditions, Neighborhood and Built Environment, and Social and Community Context. Pearson's correlation coefficient assesses the strength and direction of these associations. The full list of statistically significant correlations with life expectancy is provided in Table B1 of Appendix B.

Figure 4

Life Expectancy Correlations at the National Level



Key findings at the national level reveal substantial negative correlations between life expectancy and health behaviors and conditions, notably smoking ($r = -0.814$, $p < 0.001$), obesity ($r = -0.794$, $p < 0.001$), mental distress ($r = -0.748$, $p < 0.001$), and diabetes ($r = -0.735$, $p < 0.001$). Additionally, firearm fatalities (-0.743, $p < 0.001$) are also significantly linked to shorter life expectancy. On the other hand, there is a positive correlation between life expectancy and socio-economic measures such as median household income (0.717, $p < 0.001$), aligning with the known connection between economic status and health outcomes, as highlighted in Chetty et al.'s study.

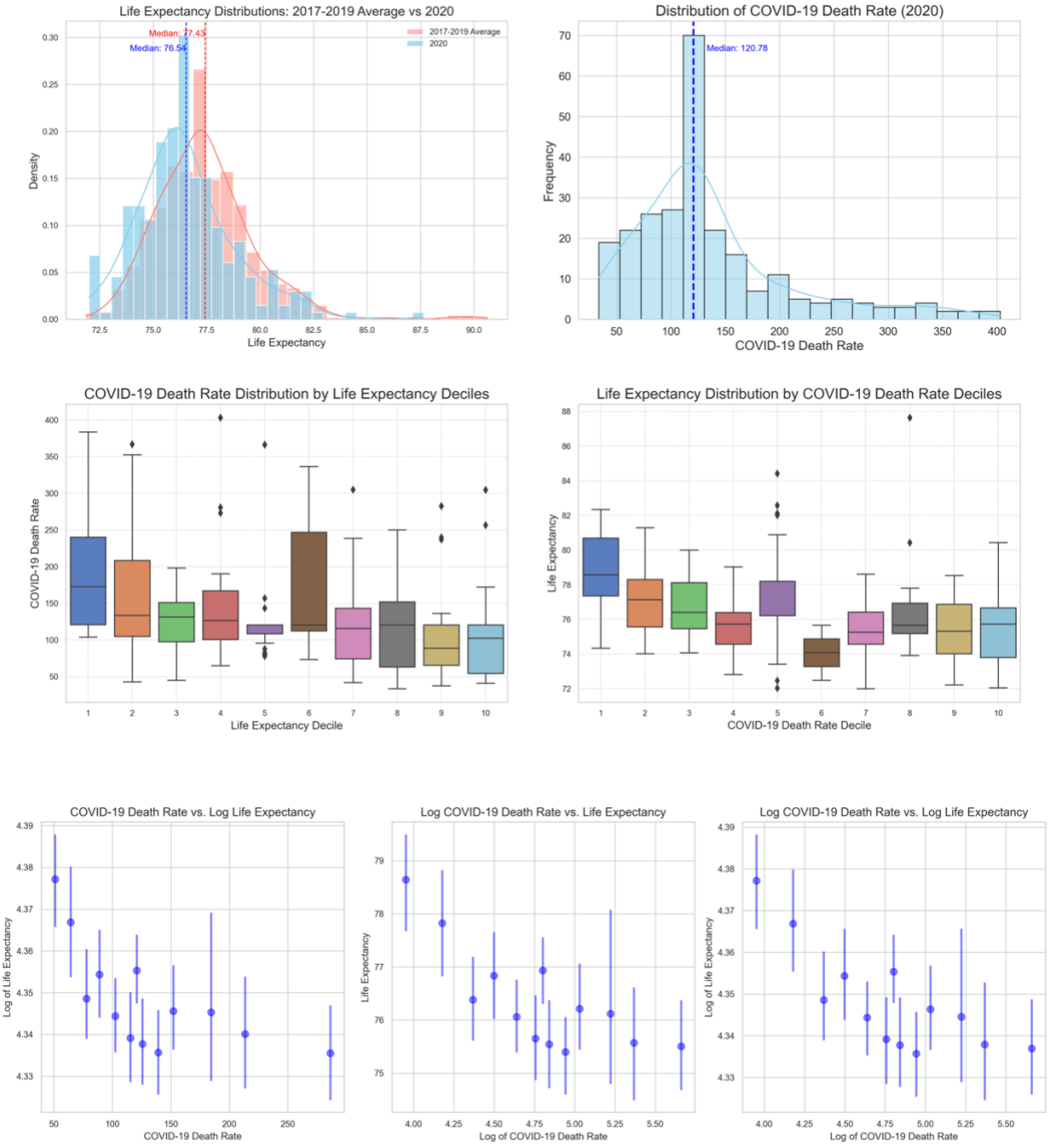
Intriguingly, traditionally impactful factors like HIV prevalence ($r = 0.008$, $p > 0.05$) and racial residential segregation ($r = 0.091$, $p > 0.05$) show unexpectedly weak correlations, indicating a shift in the dynamics influencing life expectancy, possibly due to advancements in healthcare, socio-economic transformations, or underreporting and reduced testing due to the stigma associated with HIV.

Case Study - COVID-19's Impact in Texas

The investigation focuses on the changes in life expectancy and COVID-19’s impact in Texas. The pandemic has significantly transformed many facets of life in Texas. These alterations include significant increases in unemployment and mental health issues, shifts in healthcare service utilization, and changes in lifestyle choices. The research also highlights the pandemic's effect on the correlation strength between life expectancy and crucial health and social factors. Prior to the pandemic, these factors had a different impact on life expectancy compared to their influence during the pandemic. The findings indicate an intensification of existing socio-economic disparities due to the pandemic.

Figure 5

Life Expectancy Trends and COVID-19 Mortality in Texas



The life expectancy in Texas has seen a notable disparity, with a significant difference of approximately 16.93 years (2017-2020 average). The year 2020, marked by the onset of the COVID-19 pandemic, witnessed a considerable decline in life expectancy. Specifically, the median life expectancy decreased from an average of 77.43 years during 2017-2019 to 76.54 years in 2020. This decrease is depicted in Figure 5, which shows the life expectancy distribution shifting leftward. Although the overall shape of the distribution remains similar to previous years, it now represents a lower life expectancy. The histogram details the distribution of COVID-19 death rates across Texas for the year 2020. The distribution's pronounced peak at 120.78 deaths per 100,000 population suggests a widespread impacting across various demographics and regions within the state, reflecting the pandemic's extensive reach.

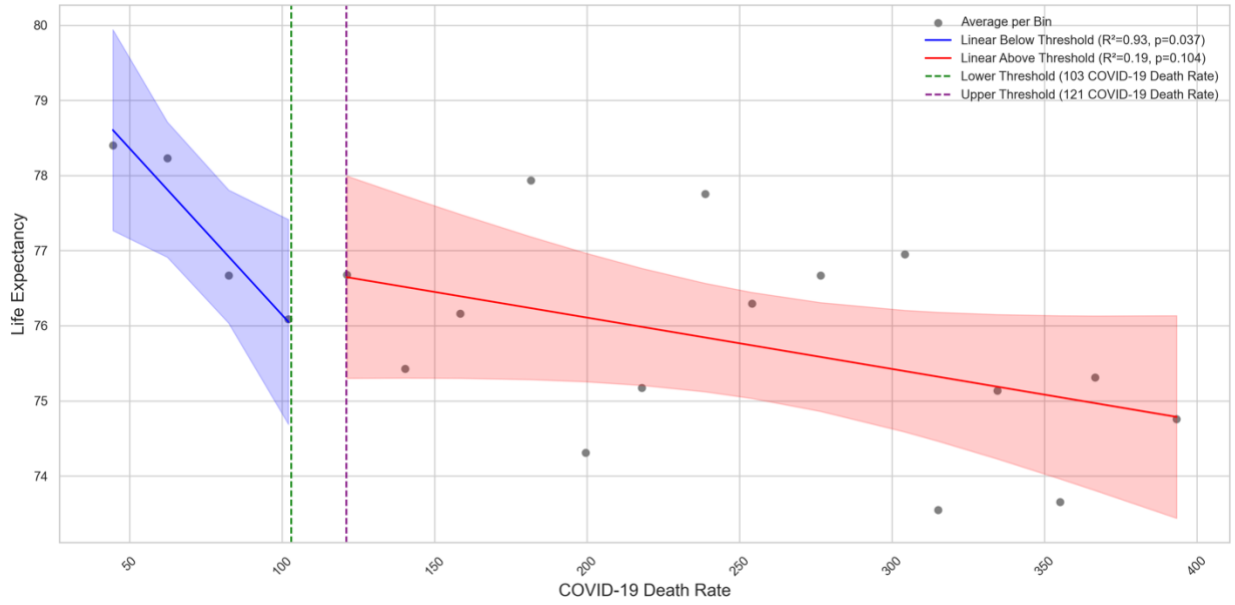
The two boxplots in Figure 5 analyze the relationship between counties' life expectancy and COVID-19 death rates. The left boxplot stratifies counties based on life expectancy deciles and their corresponding COVID-19 death rates. A potential trend is observable where higher life expectancy deciles tend to have lower median COVID-19 death rates. However, there is considerable variation within deciles, as evidenced by the range of the boxplots and the presence of outliers, which are points that deviate significantly from the rest of the data in their decile. The right boxplot contrasts this by displaying life expectancy across COVID-19 death rate deciles. It reveals overlapping interquartile ranges in some deciles, indicating limited variance in life expectancy within these groups. Despite some deciles showing outliers, which indicate significant deviations in life expectancy, the median life expectancy across the deciles appears relatively consistent.

In Texas, life expectancy shows a potential downward trend as the COVID-19 death rate increases up to the clustered region from 103 to 121 death rates (Figure 6). The segmented

regression analysis demonstrates a strong inverse correlation between life expectancy and COVID-19 death rate below this clustered region. Specifically, for every unit increase in the COVID-19 death rate, life expectancy decreases by estimated 16.19 days. This correlation is robustly supported by a high coefficient of determination (R^2) of 0.93, alongside a statistically significant p-value of 0.037, indicating that the observed relationship is unlikely to be a result of random variation. However, it is crucial to emphasize that while the R^2 value denotes a strong correlation, it should not be interpreted as causation.

Figure 6

Binned Scatter Plot with Segmented Linear Regression: Life Expectancy vs. COVID-19 Death Rate



Above the clustered region where various life expectancy figures are densely grouped, making the relationship ambiguous, the relationship between COVID-19 death rates and life expectancy weakens, as evidenced by a lower R^2 of 0.19 and a non-significant p-value of 0.104. This suggests that in regions with exceptionally high COVID-19 mortality rates, other factors may play a more dominant role in determining life expectancy.

Table 1

Indicators with the Most Significant Changes in Texas (Pre-Pandemic vs. Pandemic Periods)

Measure	**2017** **Mean ± SD**	**2018** **Mean ± SD**	**2019** **Mean ± SD**	**2020** **Mean ± SD**	**2019-2020** **Change (%)**	**p-value**
% Unemployed	4.47 ± 1.53	3.87 ± 1.26	3.52 ± 1.10	6.87 ± 3.06	95.35	< 0.001
Preventable Hospitalization Rate	5267.45 ± 1537.53	5108.43 ± 1609.46	4447.95 ± 1419.40	3237.14 ± 1037.25	-27.22	< 0.001
% Physically Inactive	27.29 ± 4.67	25.38 ± 4.90	32.82 ± 4.59	28.21 ± 4.59	-14.04	< 0.001
% With Annual Mammogram	33.17 ± 6.36	32.96 ± 6.59	34.14 ± 7.15	29.65 ± 7.55	-13.16	< 0.001
Teen Birth Rate	45.27 ± 14.98	42.24 ± 14.16	39.68 ± 13.46	36.49 ± 12.75	-8.02	0.006
% Frequent Mental Distress	12.60 ± 1.11	14.71 ± 1.24	14.99 ± 1.34	16.11 ± 0.98	7.46	< 0.001
% Children in Poverty	24.41 ± 7.84	23.81 ± 7.74	22.24 ± 7.51	20.84 ± 7.23	-6.30	0.033
% Adults Reporting Currently Smoking	14.94 ± 1.56	19.00 ± 2.33	18.42 ± 2.42	19.40 ± 3.09	5.34	< 0.001
% Adults with Diabetes	11.36 ± 4.62	11.64 ± 3.84	12.99 ± 2.31	12.35 ± 2.21	-4.92	0.002
% Vaccinated	36.41 ± 9.77	37.85 ± 9.93	37.87 ± 9.75	39.66 ± 10.17	4.723673	< 0.001
% Excessive Drinking	17.95 ± 1.83	18.88 ± 1.60	19.16 ± 1.65	20.00 ± 1.48	4.37	< 0.001

A comparative analysis was conducted for pre-pandemic and pandemic periods (Table 1). The observed trends in unemployment, healthcare, and socio-economic indicators provide a multifaceted perspective on the pandemic's impact on the life in Texas:

- Unemployment: There was a significant increase in unemployment rates, jumping from 3.52% in 2019 to 6.87% in 2020. This 95.35% rise clearly shows the immediate economic effects of the pandemic.

- Preventable Hospitalization Rate: The preventable hospitalization rate saw a notable decrease of 27.22%, possibly reflecting the public's hesitancy to seek hospital care amidst the pandemic.
- Physical Inactivity: Physical inactivity levels initially increased in 2019 to 32.82% but slightly decreased in 2020 to 28.21%, suggesting a complex interplay of factors influencing lifestyle choices during the pandemic.
- Annual Mammogram: The percentage of women receiving annual mammograms also decreased by 13.16%, highlighting potential disruptions in routine healthcare.
- Teen Birth Rates: Teen birth rates continued their downward trend, moving from 39.68 in 2019 to 36.49 in 2020, suggesting that long-term public health efforts might be sustaining their impact even during the pandemic.
- Frequent Mental Distress: There was a measurable increase in frequent mental distress, rising by 7.46% in 2020, indicating the pandemic's significant psychological toll.
- Children in Poverty: The percentage of children living in poverty continued to decline, dropping from 22.24% in 2019 to 20.84% in 2020, a decrease of 6.30%. This trend might reflect ongoing socio-economic improvements or the impact of relief measures during the pandemic.
- Smoking: The percentage of adults reporting currently smoking increased, from 18.42% in 2019 to 19.40% in 2020, possibly reflecting increased stress or changes in habits during lockdowns.
- Diabetes: The percentage of adults with diabetes showed a slight decrease in 2020, moving from 12.99% in 2019 to 12.35% in 2020, but the complexity of this change warrants further investigation.

- Vaccination Rates: There was an increase in the percentage of people vaccinated, moving from 37.85% in 2019 to 39.66% in 2020. This 4.723673% rise could be indicative of heightened public health initiatives, possibly in response to the pandemic.
- Drinking: The changes in excessive drinking rates, with a small increase from 19.16% in 2019 to 20.00% in 2020, also point to altered health behaviors under pandemic stress.

Overall, the multifaceted impact of the COVID-19 pandemic on health and social indicators reflects both immediate and potentially enduring changes in Texas's life. These findings are crucial for understanding the broader implications of the pandemic and for guiding future public health and policy decisions.

Table 2

Shifts in Life Expectancy Correlations with Health and Social Variables in Texas During the COVID-19 Pandemic (Pre-Pandemic vs. Pandemic Periods) ($r > |0.3|$, $p < 0.05$)

Pre-Pandemic Period		**Pandemic Period**	
Measure	**Correlation Coefficient**	**Measure**	**Correlation Coefficient**
Injury Death Rate	-0.453	% Frequent Mental Distress	-0.451
% Not Proficient in English	0.384	% Adults Reporting Currently Smoking	-0.433
Motor Vehicle Mortality Rate	-0.368	Median Household Income	0.429
% Asian	0.339	Injury Death Rate	-0.427
% Adults Reporting Currently Smoking	-0.330	Teen Birth Rate	-0.386
% Food Insecure	-0.323	Motor Vehicle Mortality Rate	-0.382
Median Household Income	0.323	% Asian	0.367

Table 2 presents the evolution of correlations between various health-related and socioeconomic indicators and life expectancy in Texas from the pre-pandemic era to the

pandemic period. Detailed correlations for both timeframes can be found in Figures C1 and C2 in Appendix C.

During the pre-pandemic period, a notable strong inverse relationship was observed between the injury death rate and life expectancy ($r = -0.453$), which slightly weakened during the pandemic ($r = -0.427$). The percentage of adults reporting current smoking, while negatively correlated with life expectancy before the pandemic ($r = -0.330$), reflected a more pronounced negative correlation during the pandemic ($r = -0.433$), possibly indicating an increased negative impact of smoking on health outcomes during the COVID-19 pandemic. Median household income continued to show a positive relationship with life expectancy in both periods, but with a slightly stronger correlation during the pandemic ($r = 0.429$), highlighting possibly intensified economic health disparities due to the pandemic.

Not present in the pandemic period data are the percentage of individuals not proficient in English and the rate of food insecurity, which were correlated with life expectancy before the pandemic ($r = 0.384$ and $r = -0.323$, respectively). Their absence may suggest a change in the predictive power of these variables in the face of the pandemic's influence.

Newly significant variables during the pandemic include the percentage of individuals experiencing frequent mental distress ($r = -0.451$) and the teen birth rate ($r = -0.386$), both of which have surfaced as important factors associated with life expectancy in the context of COVID-19. However, it's interesting to note that the correlation between COVID-19 death rate and life expectancy was less pronounced ($r = -0.296$) but still statistically significant ($p < 0.001$), an aspect that deserves further exploration.

Discussion

Life Expectancy Trends at the National and State Levels

The varying life expectancy across regions and time, reflecting the influence of regional policies and major events like the pandemic, highlighting the need for a targeted approach in public health planning and interventions. The sharp decline in life expectancy in 2020, both nationally and across states, particularly in New York and New Mexico, underscores the profound impact of the COVID-19 pandemic.

Life Expectancy Correlations at the National Level

The strong negative correlations between life expectancy and factors such as smoking, obesity, mental distress, and diabetes, as well as the negative correlation with firearm fatalities, alongside the positive correlation with median household income, reaffirm the critical role of health behaviors, socio-economic status, and safety conditions in determining health outcomes. The weakened correlations of traditionally impactful factors like HIV prevalence and racial residential segregation suggest a dynamic shift in the determinants of life expectancy, potentially attributable to healthcare advancements or socio-economic changes.

Case Study - COVID-19's Impact in Texas

The COVID-19 pandemic has had a profound impact on life expectancy in Texas, as the case study of the state reveals. This detailed examination sheds light on the significant disparities at the county level within the state. The findings underscore the necessity for ongoing public health surveillance and interventions, particularly in areas most affected by the pandemic. Importantly, Texas’s experiences demonstrate how the pandemic has exacerbated pre-existing socio-economic inequalities, affecting a range of health and social factors. Increases in unemployment and mental distress, alongside shifts in healthcare service utilization and lifestyle habits, paint a picture of a society grappling with the multifaceted challenges posed by the pandemic. The observed shift in correlations between various health and social measures with

life expectancy from pre-pandemic to pandemic times indicates a changing landscape of health determinants. The introduction of new variables like frequent mental distress and the changing dynamics of factors like smoking and median household income suggest an evolving context for health outcomes in the pandemic era.

Implications for Public Health and Policy

These findings highlight the necessity of adaptable and resilient public health strategies. The variability in life expectancy trends across states and the shifting determinants of health outcomes call for tailored approaches that address local needs, health behaviors, and socio-economic disparities. Policies need to focus not only on direct health interventions but also on addressing underlying socio-economic factors that significantly impact health outcomes.

Limitations

The study acknowledges several limitations. First, it uses only one year of data to analyze the pandemic's impact due to limitations in data availability. Second, the accuracy of the part of self-reported data may not always be reliable. Third, there are unmeasured factors in the correlation analysis, particularly concerning the omission of pandemic-related data at the national level, which limits the study's comprehensiveness for this period. Fourth, a significant limitation is the challenge in establishing cause-and-effect relationships from mere correlations. Fifth, although the measurement algorithm has undergone rigorous examination to ensure the suitable usage of data, there remains a risk of biases, especially given the nature of the secondary data. Furthermore, the study admits that some of its data are derived from statistical models; while these models are useful, they might not fully capture real-world scenarios.

Future Research Directions

Further research should understand the long-term implications of the pandemic on life expectancy and health outcomes. Studies should explore the indirect effects of the pandemic, such as its impact on healthcare access, mental health, and socio-economic conditions. A thorough understanding of these interconnected dynamics is vital to formulate and implement effective public health strategies aimed at mitigating the pandemic's adverse effects on life expectancy. Additionally, exploring the effectiveness of different state-level public health responses in mitigating the pandemic's impact could provide valuable insights for future health crises. Moreover, the quest for effective solutions highlights the necessity of scalable and measurable interventions, adaptable to different populations and regions. In an era marked by the proliferation of big data, improving data collection quality, obtaining real-world data, and using broader, more diverse datasets should be a research priority. Such approaches will deepen our understanding and yield more inclusive and comprehensive results.

Conclusion

In conclusion, the comprehensive analysis illustrates a significant transformation in U.S. life expectancy trends, profoundly influenced by the COVID-19 pandemic. State-by-state variations underscore the need for tailored public health strategies. The study reveals that life expectancy not only reflects the direct impacts of health crises but is also intricately tied to socio-economic factors and health behaviors. In Texas during 2020, factors such as increased unemployment, mental health issues, and changes in healthcare and lifestyle habits highlight the pandemic's broad societal impact. The changing relationship between life expectancy and various health and social indicators in Texas during the pandemic indicates evolving health determinants. Therefore, the study emphasizes the need for adaptable, resilient public health approaches that address both direct and indirect consequences of health crises. As we move forward, it is crucial

to continue exploring these trends and their implications to strengthen our preparedness and response to future health challenges.

References

Chan, E. Y. S., Cheng, D., & Martin, J. (2021). Impact of COVID-19 on excess mortality, life expectancy and years of life lost in the United States. *PLoS ONE*, 16(9), Article e0256835. https://doi.org/10.1371/journal.pone.0256835

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., & Cutler, D. (2016). The association between income and life expectancy in the United States, 2001-2014. *JAMA*. https://doi.org/10.1001/jama.2016.4226

University of Wisconsin Population Health Institute. (2023). County health rankings & roadmaps. Retrieved from www.countyhealthrankings.org

Appendices

Appendix A

Table A1

Measures for Correlation Analysis at the National Level: Definitions and Data Sources

Measure	Definition	Source
Economic Stability:		
Median Household Income	Middle household income value calculated using a combination of administrative records, census data, and the American Community Survey.	U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE), 2017-2020
Income Ratio	Disparity between the incomes at the 80th and 20th percentiles.	American Community Survey (ACS) 5-year estimates, 2017-2020
% Unemployed	Percentage of the labor force that is actively seeking employment.	Bureau of Labor Statistics, 2017-2020
% Children in Poverty	Percentage of children under 18 living in poverty, calculated using household income and poverty thresholds.	U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE), 2017-2020
Education and Youth Engagement:		
% Not Proficient in English	Percentage of individuals who do not speak English well.	American Community Survey (ACS) 5-year estimates, 2017-2020
% Some College	Proportion of individuals aged 25-44 with post-secondary education, including vocational or higher education.	American Community Survey (ACS) 5-year estimates, 2017-2020
% Disconnected Youth	Percentage of individuals aged 16 to 19 not employed and not enrolled in education.	American Community Survey (ACS) 5-year estimates, 2017-2020

Healthcare Access and Quality:		
% Uninsured Adults	Percentage of individuals aged 18 to 64 without health insurance.	U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE), 2017-2020
% Uninsured Children	Percentage of children under 19 without health insurance.	U.S. Census Bureau's Small Area Health Insurance Estimates (SAHIE), 2017-2020
Preventable Hospitalization Rate	Hospitalizations for certain conditions per 100,000 Medicare Part A enrollees, age 18 and older.	Mapping Medicare Disparities Tool, 2017-2020
% With Annual Mammogram	Percentage of women aged 65-74 on Medicare receiving an annual mammogram.	Mapping Medicare Disparities Tool, 2017-2020
% Vaccinated	Percentage of Medicare Part B enrollees who received a reimbursed flu vaccine.	Mapping Medicare Disparities Tool, 2017-2020
Health Behaviors and Conditions:		
% Adults with Obesity	Percentage of adults aged 18 or older with a BMI $\geq$ 30 kg/m², based on self-reported data.	CDC Behavioral Risk Factor Surveillance System (BRFSS), 2017-2020
% Adults with Diabetes	Age-adjusted prevalence of diabetes among adults.	CDC Behavioral Risk Factor Surveillance System (BRFSS), 2017-2020
% Physically Inactive	Percentage of adults aged 18 and over with no leisure-time physical activity.	CDC Behavioral Risk Factor Surveillance System (BRFSS), 2017-2020

% Adults Reporting Currently Smoking	Percentage of adults who report current smoking and have smoked over 100 cigarettes in their lifetime.	CDC Behavioral Risk Factor Surveillance System (BRFSS), 2017-2020
% Frequent Mental Distress	Percentage of adults reporting 14 or more days of poor mental health within a 30-day period.	CDC Behavioral Risk Factor Surveillance System (BRFSS), 2017-2020
% Excessive Drinking	Percentage of adults reporting binge or heavy drinking.	CDC Behavioral Risk Factor Surveillance System (BRFSS), 2017-2020
Drug Overdose Mortality Rate	Drug overdose deaths per 100,000 people, adjusted for population.	National Center for Health Statistics' Mortality Files, 2017-2020
% Low Birthweight	Seven-year average percentage of infants weighing less than 2,500 grams at birth, by mother's residence.	National Center for Health Statistics' National Vital Statistics System, 2017-2020
Teen Birth Rate	Number of females aged 15-19 giving birth, calculated per 1,000 females in this age group.	National Center for Health Statistics' National Vital Statistics System, 2017-2020
Chlamydia Rate	Newly diagnosed chlamydia cases expressed as a percentage of the total population.	National Center for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP), 2017-2020
HIV Prevalence Rate	Number of the population aged 13 and older diagnosed with HIV.	National Center for HIV, Viral Hepatitis, STD, and TB Prevention (NCHHSTP), 2017-2020
<u>Neighborhood and Built Environment:</u>		
Residential segregation - Black/White	Segregation index between black and white populations.	American Community Survey (ACS) 5-year estimates, 2017-2020

% Long Commute - Drives Alone	Percentage of workers commuting more than 30 minutes alone by car, truck, or van.	American Community Survey (ACS) 5-year estimates, 2017-2020
% Rural	Percentage of the population living in rural areas.	2010 Census Population Estimates
Food Environment Index	Community access to healthy foods and food insecurity, assessed by two indicators: proximity to grocery stores and food insecurity levels.	USDA Food Environment Atlas and Feeding America's Map the Meal Gap, 2017-2020
<u>Social and Community Context:</u>		
% Children in Single-Parent Households	Percentage of children under 18 living in single-parent family households.	American Community Survey (ACS) 5-year estimates, 2017-2020
Social Association Rate	Number of membership associations per 10,000 residents.	County Business Patterns, 2017-2020
Homicide Rate	Assault-related deaths per 100,000 population.	National Center for Health Statistics, 2017-2020
Firearm Fatalities Rates	Firearm-related deaths per total population.	National Center for Health Statistics, 2017-2020
Injury Death Rate	Injury-related deaths per 100,000 people.	National Center for Health Statistics' Mortality Files, 2017-2020

Appendix B

Table B1

Statistically Significant Correlations with Life Expectancy at the National Level (2017-2020)

Variable	Correlation with Life Expectancy	p-value
% Adults Reporting Currently Smoking	-0.814	< 0.001
% Adults with Obesity	-0.794	< 0.001
% Frequent Mental Distress	-0.748	< 0.001
Firearm Fatalities Rates	-0.743	< 0.001
Food Environment Index	0.735	< 0.001
% Adults with Diabetes	-0.735	< 0.001
Median Household Income	0.717	< 0.001
Teen Birth Rate	-0.695	< 0.001
% Children in Poverty	-0.680	< 0.001
Injury Death Rate	-0.669	< 0.001
% Physically Inactive	-0.668	< 0.001
% Some College	0.648	< 0.001
% Low Birthweight	-0.635	< 0.001
Homicide Rate	-0.601	< 0.001
% Not Proficient in English	0.543	< 0.001
% Rural	-0.529	< 0.001
% Disconnected Youth	-0.512	< 0.001
Preventable Hospitalization Rate	-0.449	< 0.001
% Uninsured Adults	-0.379	< 0.001
% Excessive Drinking	0.373	< 0.001
% Children in Single-Parent Households	-0.353	< 0.001
Drug Overdose Mortality Rate	-0.253	< 0.001
Social Association Rate	-0.247	< 0.001

% With Annual Mammogram	0.230	< 0.001
Chlamydia Rate	-0.220	0.002
% Vaccinated	0.193	0.006
% Uninsured Children	-0.175	0.012
% Unemployed	-0.167	0.017
Income Ratio	-0.167	0.017
% Long Commute - Drives Alone	0.139	0.048

Appendix C

Figure C1

Broader Correlations in Texas (2017-2019)

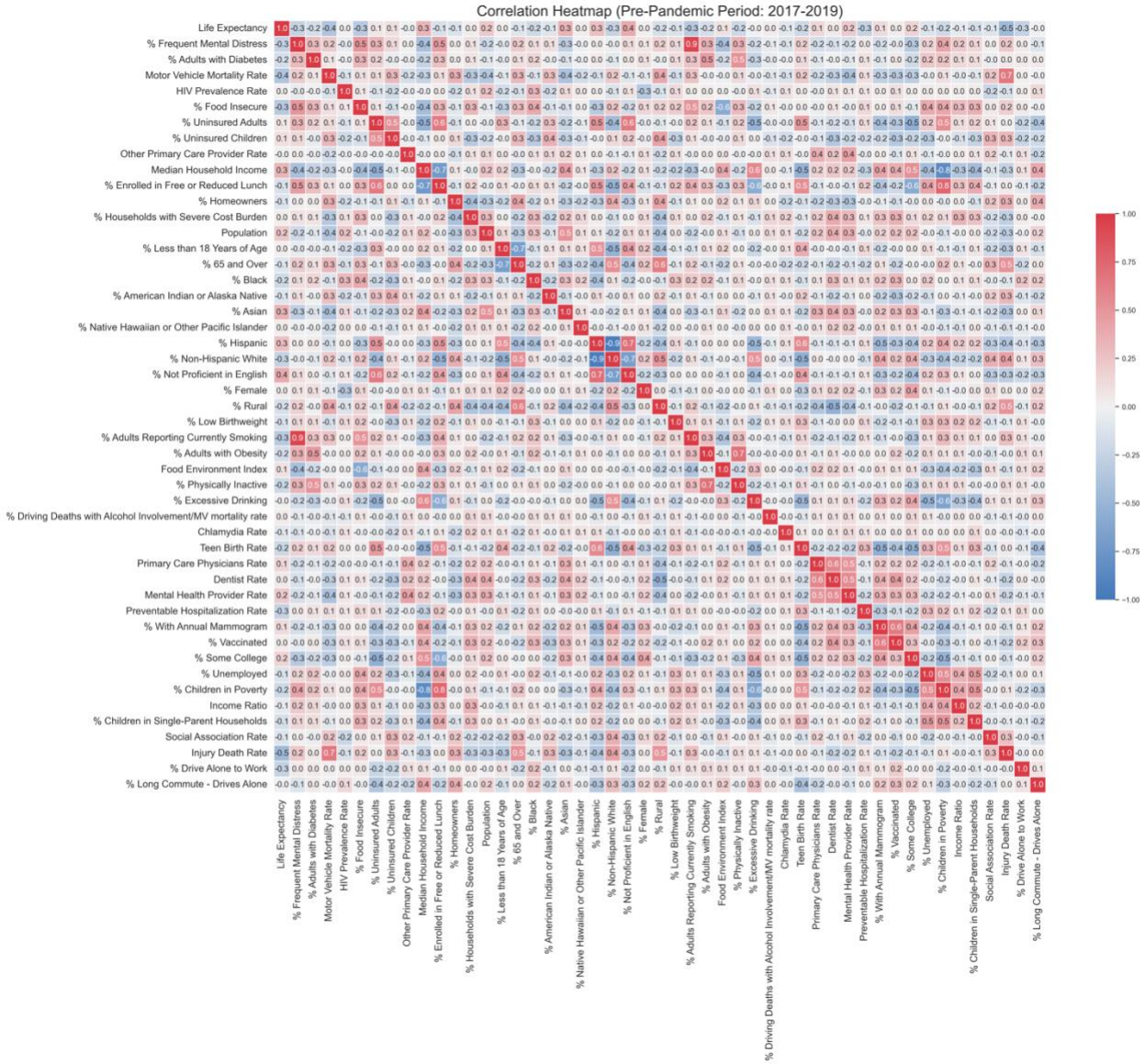


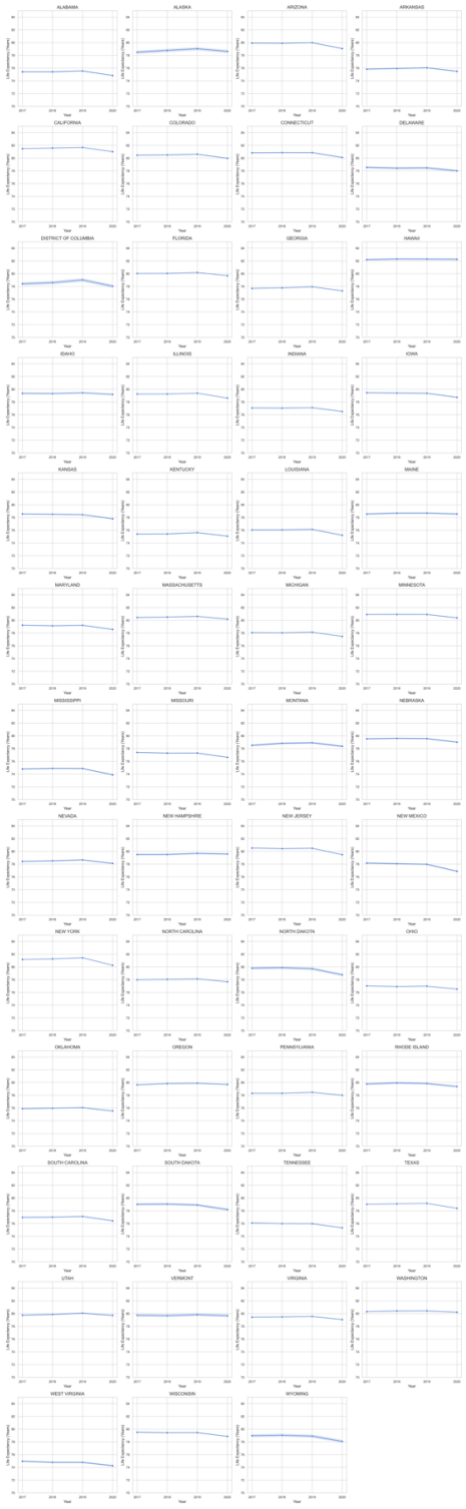
Figure C2

Broader Correlations in Texas (2020)

Appendix D

Figure D1

Life Expectancy Trends Across All U.S. States



Appendix E

Figure E1

Scatter Plots: Correlating Variables with National Life Expectancy