Contents

[**INTRODUCTION** 1](#_heading=h.gjdgxs)

[**BACKGROUND & MOTIVATION FOR STUDY** 2](#_heading=h.30j0zll)

[**About Kearney** 2](#_heading=h.1fob9te)

[**Wildfire Activity Near Kearney** 2](#_heading=h.3znysh7)

[**Air Quality Near Kearney** 3](#_heading=h.2et92p0)

[**RESEARCH QUESTIONS** 3](#_heading=h.tyjcwt)

[**RELATED WORK** 4](#_heading=h.3dy6vkm)

[**DATA** 5](#_heading=h.1t3h5sf)

[**Wildfire Attribute Data** 5](#_heading=h.4d34og8)

[**Air Quality Estimates Near Kearney** 5](#_heading=h.2s8eyo1)

[**Population Data** 6](#_heading=h.17dp8vu)

[**Other Causes of Premature Mortality** 6](#_heading=h.3rdcrjn)

[**METHODS AND FINDINGS** 6](#_heading=h.26in1rg)

[**R1: Differences in PM2.5 Levels, Smoke Days vs. No Smoke Days** 6](#_heading=h.lnxbz9)

[*Methods* 6](#_heading=h.35nkun2)

[*Findings* 6](#_heading=h.1ksv4uv)

[**R2: Fire-Related PM2.5 Estimates** 8](#_heading=h.44sinio)

[*Methods* 8](#_heading=h.2jxsxqh)

[*Findings* 9](#_heading=h.z337ya)

[**R3: Estimating Health Impacts 2006-2018** 9](#_heading=h.3j2qqm3)

[*Methods* 9](#_heading=h.1y810tw)

[*Findings* 10](#_heading=h.4i7ojhp)

[**R4: Comparison of Mortality by Cause** 11](#_heading=h.2xcytpi)

[*Methods* 11](#_heading=h.1ci93xb)

[*Findings* 11](#_heading=h.3whwml4)

[**R5: Predicting Future Mortality from Fire Related PM2.5** 12](#_heading=h.2bn6wsx)

[*Methods* 12](#_heading=h.qsh70q)

[*Findings* 12](#_heading=h.3as4poj)

[**R6: Predicting Future Mortality from PM2.5 & Wildfire Attribute Data** 13](#_heading=h.1pxezwc)

[*Methods* 13](#_heading=h.49x2ik5)

[*Findings* 14](#_heading=h.2p2csry)

[**DISCUSSION** **AND RECOMMENDATIONS** 15](#_heading=h.147n2zr)

[**RELATIONSHIP TO HUMAN CENTERED DATA SCIENCE** 15](#_heading=h.3o7alnk)

[**LIMITATIONS** 16](#_heading=h.ihv636)

[**CONCLUSION** 17](#_heading=h.32hioqz)

[**APPENDIX I – SENSITIVITY ANALYSIS** 1](#_heading=h.1hmsyys)9

[**APPENDIX II - REFERENCES** 2](#_heading=h.2grqrue)2

[**APPENDIX III - DATA SOURCES** 2](#_heading=h.vx1227)4

# **INTRODUCTION**

This project investigated the health impact of landscape fires near Kearney, Nebraska. As wildfire activity continues to increase, the resulting smoke pollution is an increasing issue of public concern. [1, 4-5, 8-9, 13, 15-17]. Smoke from landscape fires – including both “wildfires” and prescribed burns – contributes to fine particulate matter pollution (PM2.5). [Id.] Although wildfires emit several types of pollution, PM2.5 pollution is, according to the US Environmental Protection Agency [8], the most common and the most toxic type of pollution omitted, and therefore the focus of this study. PM2.5 pollution increases asthma-related ER visits, hospital admissions, and premature mortality from a variety of causes [1, 4-5, 8-9, 13, 15-17].

In any assessment of wildfire health impacts, it is important to isolate PM2.5 pollution attributable to wildfires from other potential sources of PM2.5 pollution such as power plants, factories, cars, and trucks. Throughout much of the US, and including in Kearney, pollution has been steadily *decreasing* in recent decades, even as wildfire activity trends up. [4, 9]. This overall decrease likely results from pollution mitigation measures that flowed from the Clean Air Act. [9]. Unfortunately, increasing wildfire activity, if recent trends continue, is likely to erode or even reverse this progress.

This project therefore sought to detangle *fire related* PM2.5 pollution from other-cause PM2.5 pollution. Building on the work of O’Dell *et. al*. [15] this project categorized days as either “Smoke Days” or “No Smoke Days” by reference to satellite imagery used to detect whether a smoke plume was visible above Kearney on that day. The data used by this project had already made the “Smoke Day” / “No Smoke Day” determination, which I followed. The smoke day designation variable was then combined with EPA readings of PM2.5 levels near Kearney to estimate annual smoke impacts from 2006 to 2018, the period for which the satellite imagery data was available.

This project also estimated the historic asthma-related ER visits, asthma-related hospital visits, and premature mortality (from all causes) among Kearney residents that can be attributed to fire pollution. It concluded that these health impacts are significant, resulting in the premature death of approximately five people during the study period. But premature mortality from wildfire is much smaller than premature death from other leading health concerns, such as heart disease.

This project also predicted how health impacts from fire pollution might impact Kearney residents in the future. *First*, it used observed fire related PM2.5 pollution from 2006-2018 to project future fire related PM2.5 pollution and health impacts. *Second,* it combined the observed fire related PM2.5 with the longer time series of data available on wildfire attributes near Kearney. The study concluded that, while adverse health impacts appear to be on the rise, limitations in available data suggest that future predictions should be treated with skepticism.

This report concludes with policy recommendations, a discussion of how human centered data science principles influenced the project, study limitations, and directions for further research.

# **BACKGROUND & MOTIVATION FOR STUDY**

## **About Kearney**

Kearney, Nebraska is a city of over thirty thousand residents located in Buffalo County, Nebraska, in the southeastern portion of the state. [12] It is approximately 700 miles south of the Canadian border [10], and home to the University of Nebraska at Kearney. [12] Kearney is governed by a five-member city council. [12]

## **Wildfire Activity Near Kearney**

Understanding the health impacts from wildfire smoke is important because wildfire activity is on the rise. Figure 1 shows a time series of total acres burned per year within 1250 miles of Kearny. Data is from the [US Geological Survey](https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81), which contains data on known wildfires in the US. Notably, the dataset is likely to underreport fires prior to 1980. Even limiting the data to 1980+, however, still shows a clear upward trend in wildfire activity as seen in Figure 2.

**Figure 1: Total Acres Burned within 1250 Miles of Kearney, Nebraska (1963-2020)**

A blue line graph with numbers

Description automatically generated

**Figure 2: Total Acres Burned within 1250 Miles of Kearney, Nebraska (1980-2020)**

A graph showing the growth of the stock market

Description automatically generated

Due to its relative proximity to Canada, Kearney is also impacted by Canadian wildfires. [14] Thus, the US Wildfire dataset likely underestimates the relevant wildfire smoke impact.

## **Air Quality Near Kearney**

Despite increasing wild-fire activity, air quality near Kearney has been improving in recent decades. Figure 3 graphs the annual mean PM2.5 concentrations near Kearney from two different data sources: (1) unprocessed air quality readings from EPA monitoring stations within 100 miles of Kearney; and (2) air quality estimates published by O’Dell *et. al*. [15] According to both data sets, PM2.5 pollution has decreased in recent years, contrary to the expectation that increased annual fire activity would spur increased PM2.5 levels.

**Figure** 3**: Annual Mean PM2.5 Near Kearney, Nebraska**

A line graph with orange and blue lines

Description automatically generated

This apparent paradox can be reconciled by considering recent trends in PM2.5 pollution caused by sources *other than* landscape fires. PM2.5 pollution can be created or exacerbated by many anthropogenic sources, including power plants, factories, and cars and trucks. [6] Burke *et. al.*[4] noted that air pollution concentrations have been declining over several decades, thanks in part to the Clean Air Act, initially enacted in 1963 and subsequently amended. The Clean Air Act regulates pollutants, including PM2.5, from sources such as energy production, manufacturing, agriculture, and transportation. Burke *et. al.* conclude that although air pollution concentrations have been declining, increased wildfire activity threatens to slow or reversed these trends. [4]

In short, when estimating the health impacts of fire related PM2.5 pollution, it is critical to disentangle fire related PM2.5 pollution from background causes of PM2.5 pollution. Motivated by that insight, this study sought to estimate fire specific PM2.5 pollution, to use those estimates to estimate historical health impacts, and to predict future health impacts on Kearney residents.

# **RESEARCH QUESTIONS**

1. Using Satellite Imagery Data to Segregate Pollution (PM2.5) Estimates into “Smoke Days” and “No Smoke Days”, do we observe statistically significant differences in pollution levels based on the presence or absence of smoke?
2. Using the presence or absence of smoke above Kearney as a proxy for whether elevated PM2.5 levels on that day can be attributed to wildfire, what were the estimated annual PM2.5 pollution impacts from fires near Kearney in the period 2006-2018?
3. Using the results from (R2), what were the estimated health impacts of fire related PM2.5 on the residents of Kearney from 2006-2018?
4. How do the estimates of premature mortality from fire-related pollution found in (R3) compare to other causes of premature mortality impacting residents of Kearney?
5. How are health impacts from fire related PM2.5 likely to evolve in the future?
6. Can estimates/predictions based on air quality data from 2006-2018 be improved by also incorporating information on wildfire attributes from 1963-2020?

# **RELATED WORK**

This analysis builds on data and methods from two published studies. *First*, in “*Estimated mortality and morbidity attributable to smoke plumes in the United States: not just a western US problem,*" O’Dell *et. al.* [15] combined observed levels of PM2.5 pollution and estimates of gas-phase enhancements in smoke hazardous air pollutants (HAPs) across the United States, with satellite imagery data to detect smoke plumes in the atmosphere. From these two data sources, they estimated, for each 15 km x 15 km grid in the US, a daily measure of PM2.5 that can be attributed to fires, for the period 2006-2018. They then used these estimates to calculate the health impact from fire related pollution across several different measures: (1) asthma-related ER visits due fire-related PM2.5; (2) asthma-related emergency room visits due to fire related PM2.5; (3) premature mortality due to fire related PM2.5; and (4) disability-adjusted life years resulting from PM2.5 combined with HAPs. They conclude that fire related pollution contributes to increased morbidity and mortality across the US, even in regions thought to be remote from wildfire activity. As further described herein, my study leveraged the data used by O’Dell *et. al.*, as well as several aspects of their methodology. Unlike O’Dell et. al, however, I limited my analysis to PM2.5 pollution and did not consider the impact of HAPs or disability-adjusted-life-years. I also focused on one specific 15 km x 15 km grid: the one located closest to Kearney.

*Second,* in “*Future fire impacts on smoke concentrations, visibility, and health in the contiguous United States*,” Ford *et. al.* [9] use a model-based approach to simulate future PM2.5 from wildfires over the next century. The authors then use these simulated values of fire related PM2.5 to estimate impacts on air quality, visibility, and premature mortality through the end of the 21st century. Among other things, their analysis concluded that premature deaths attributable to fire related PM2.5 will double by the end of the 21st century.

Ford *et. al* and O’Dell *et al.* use similar approaches to calculating the health impacts that flow from fire related PM2.5. The main methodological differences between the two studies lie in the source for estimating fire related PM2.5 in the first instance. O’Dell *et. al*., used observed values and Ford *et. al.* used simulated values. My analysis followed the O’Dell *et. al*. approach and used observed, rather than simulated levels of PM2.5. However, as further discussed in the methods sections, in certain respects I followed the slightly less computationally difficult approach of Ford *et. al.* as an alternative to aspects of the O’Dell *et. al.* approach.

I also considered estimating PM2.5 pollution from fires using the EPA’s BlueSky fire model, as referenced in [17] but was unable to determine whether I would be able to gather necessary data specific to Kearney Nebraska in the timeframe available for this class project.

Although this study primarily draws on the work by O’Dell *et. al.* and Ford *et. al.*, it also considered and gained insights from several additional studies. For example, work by Burke *et. al.*, [4] noting the downward trends in non-fire related PM2.5 served as the inspiration for seeking methods to disaggregate PM2.5 measures into fire-related and non-fire related sources. Other work [16-17] studied prescribed fires, and concluded that prescribed fires significantly contributed to PM2.5 pollution. This informed the decision to keep prescribed burns in the wildfire dataset for purposes of creating Figures 1-2 and for modelling historical and future smoke impacts in research question 6, below.

# **DATA**

This section briefly describes data used in the analysis. More detail is available in the README file in the [project repository](https://drive.google.com/drive/folders/1xI5BOdkj-Avt1WA8ay1DfII35q5tgN6c?usp=drive_link).

## **Wildfire Attribute Data**

Wildfire Data was obtained from the wildfire dataset aggregated by the [US Geological Survey](https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81.), (“US Wildfire Dataset”). It contains data on 135061 fires in the United States, spanning from 1835 to 2020, merged from 40 data sets. Not all historical fires were reported, especially for years prior to 1980, so data may underestimate historical wildfires. The database contains data on the annual date of the fire, the number of acres burned, and location, among other attributes.

## **Air Quality Estimates Near Kearney**

The study considered pollution estimates near Kearney from two sources. First, I leveraged air quality data [made available](about:blank) by O’Dell *et. al.* under [license terms](https://lib.colostate.edu/find/csu-digital-repository/policies-guidelines-forms/research-data-terms-of-use/) that allow reuse with attribution. O’Dell *et al*. downloaded annual daily 24-hr average PM2.5 observations from the EPA Air Quality Systems Data Mart. They then “kriged” the data – that is, they applied an inverse-distance-weighted data interpolation method to create daily PM2.5 estimates for each 15km-by-15km grid in the US. They also incorporated smoke plume information from satellite imagery. For each grid cell, and each day in the study period, the authors categorized whether a smoke plume was visible above the grid on that day using data from NOAA Hazard Mapping Systems (“HMS”) Polygons. In the dataset they published, and which I leveraged, the presence or absence of smoke above a grid on a particular day is recorded as either a 1 “Smoke” or a 0 “No Smoke” in the indicator variable “HMS\_Smoke.” From this dataset, I extracted data relevant to Kearney by selecting data from the 15km-by-15km grid whose center was the closest to the center of Kearney.

Second, I obtained air quality estimates directly from the [EPA Air Quality System API](https://aqs.epa.gov/aqsweb/documents/data_api.html). For each year beginning in 1999, I queried air monitoring stations within 100 miles of Kearney (the closest stations available), that also measured particulate matter using either of the two methods relevant to PM2.5:EPA codes “88101” and “88502”, the same codes in the O’Dell *et. al.* study. For years prior to 1999, no monitoring station measuring these codes could be found within a 350-mile radius of Kearney, so no data was collected for these years.

Comparing PM2.5 concentrations from the two methods suggested that the two datasets were consistent, with a correlation coefficient of .84.

## **Population Data**

All population data is sourced from the US census. I used population estimates to calculate the absolute health impact on the citizens of Kearney based on relative risk calculations. In some cases, population data is only available at decade intervals; in these cases, annual estimates were constructed assuming smooth annual growth between decades. I also used the most recent decade’s annual growth rate to project the future Kearney population.

## **Other Causes of Premature Mortality**

Data on traffic accidents in Nebraska was sourced from the National Highway Traffic Safety Association. Data on annual firearms death in Nebraska was manually retrieved from [US gun deaths (usafacts.org)](https://usafacts.org/data/topics/security-safety/crime-and-justice/firearms/firearm-deaths/),which in turn relies on data from the CDC [WISQARS Data Visualization (cdc.gov)](https://wisqars.cdc.gov/data/explore-data/home) tool. Data on premature deaths from heart diseases in Nebraska was sourced [from the CDC](https://www.cdc.gov/nchs/data-visualization/mortality-leading-causes/index.htm), as were [homicide rates](#bookmark=id.3fwokq0). Estimates for Kearney were based on the population ratio between Kearney and Nebraska. In all cases, absolute, rather than age-adjusted, mortality rates were used. In the case of homicide rates, data was only available for the years 2014-2018, so I took the average value for those years, multiplied by 13, as the estimate for the 13-year period 2006-2018.

# **METHODS AND FINDINGS**

## **R1: Differences in PM2.5 Levels, Smoke Days vs. No Smoke Days**

### *Methods*

Following the methodology employed by O’Dell, *et. al.*, I segregated PM2.5 pollution estimates into those associated with “Smoke Days” – that is, when a smoke plume was observed above Kearney – and “No Smoke Days.” I then compared annual mean PM2.5 on Smoke Days and No Smoke Days, and used a paired two-sample t-test to determine whether the distribution in means were significantly different. I also repeated the analysis for monthly mean PM2.5 on Smoke Days and No Smoke Days.

Finally, I estimated *daily* fire related PM2.5 as:

* 0 on days when no smoke was observed above Kearney.
* Observed PM2.5 - Background PM2.5 on days when no smoke was observed.

For purposes of the *daily* fire related PM2.5 estimates only, “Background PM2.5” is the seasonal median PM2.5 from No Smoke Days as calculated by O’Dell *et. al*. I performed a one sample t-test of the null hypothesis that the mean distribution of *daily* fire related PM2.5 on Smoke Days is actually zero, as an additional check of whether classification as a Smoke Day is associated with a statistically significant increase in PM2.5.

### *Findings*

Smoke Days had, on average, significantly higher PM2.5 levels, as would be expected. The relationship held regardless of whether average PM2.5 was calculated on an annual (Figure 4) or monthly (Figure 5) basis, suggesting that seasonality alone was not causing the discrepancy.

**Figure 4: Annual Mean PM2.5 Near Kearney, Smoke Days vs. No Smoke Days**

A graph with blue and orange lines

Description automatically generated

**Figure 5: Monthly Mean PM2.5 Near Kearney, Smoke Days vs. No Smoke Days**

A graph of a graph showing smoke and smoke

Description automatically generated with medium confidence

The two-sample t-test on the difference in annual mean PM2.5 on Smoke Days vs. No Smoke Days resulted in a p-value of .0001. The two-sample t-test on the difference in monthly mean PM2.5 (relevant months only) resulted in a p-value of 1.32e-8. Both p-values suggest statistically significant differences in (annual/monthly) means between Smoke Days and No Smoke Days.

Figure 6 graphs the estimated the *daily* “Fire Related PM2.5”, calculated as the difference between the PM2.5 reading for that day and the seasonal “No Smoke Day” median PM2.5 (e.g., the average “Background” level of PM2.5).

**Figure 6: Daily Fire Related PM2.5 (Smoke Days) vs. Seasonal Background (Median Seasonal PM2.5 on No Smoke Days)**

A graph with blue and orange lines

Description automatically generated

Notably, the daily estimates of Fire Related PM2.5 contain several negative values. This suggests imperfections in the data or methodology, as one would not expect the presence of smoke to have a negative impact on pollution. The one sample t-test of the null hypothesis that the mean distribution of “Daily Fire Impact” on Smoke Days is zero resulted in a p-value of 2.23e-29. This result suggests that it is highly unlikely that the true distribution is zero or below. Nonetheless, the presence of negative “Fire Related PM2.5” values is an important reminder that there are imperfections inherent in these estimates. Notably, O’Dell *et. al.* do not mention this aspect of the data and analysis, nor do they appear to check for the presence of negative daily Fire Related PM2.5 values in their code, even though they check for negative PM2.5 values in others contexts. The failure to check for and address negative daily Fire Related PM2.5 is a possible oversight and limitation in their study.

## **R2: Fire-Related PM2.5 Estimates**

### *Methods*

For each year, I calculated the cumulative smoke impact for the year as:

(MeanPM2.5 on Smoke Days **-** MeanPM2.5 on No Smoke Days) **\*** Number of Smoke Days / 365

This method calculated smoke impact *as if the smoke pollution were amortized evenly over the entire year.* This is, of course, *not* how we experience smoke pollution. Rather, wildfire activity tends to create intense spikes during wildfire season. I nonetheless used annual methods for several reasons. First, as noted above, using daily calculations resulted in negative fire-related pollution estimates for some days, which is both implausible and makes little sense in the context of the health impact assessment equations employed in later sections. Using annual averages smoothed out this concern. Notably, the annual method approach was also used by Ford *et. al.* (whereas O’Dell *et. al.* used a daily calculation method). Second, annual estimates were better suited to joining the data with the wildfire attribute data in Research Question 6. On the other hand, using the annual method created drawbacks based on the shape of the health impact response functions. Nonetheless, sensitivity analysis – further detailed in Appendix A - suggests that the simplification to annual average fire related PM2.5 was reasonable.

### *Findings*

Using an annual estimation method, PM2.5 pollution attributable to fires is increasing over time.

**Figure 7: Average Annual Fire Pollution Impact   
(Δ Ave PM2.5 on Smoke Days \* Number of Smoke Days / 365)**

A graph with a line and a red line

Description automatically generated

## **R3: Estimating Health Impacts 2006-2018**

### *Methods*

The impact of fire related PM2.5 on asthma-related ER visits, asthma-related hospital admissions, and premature mortality were estimated using a Health Impact Equation of a form commonly employed by researchers in this area. [1, 5, 9, 13, 15].

**Equation 1**: Δ Event = Pop\*BaselineEvent(1-e- Β\_Event\*ΔPM2.5)

Where:

* Pop is the population of interest (here, the residents of Kearney).
* BaselineEvent is the annual average occurrence of the event. For example, BaselineMortality was estimated as aprox. 733 per 100,000, meaning that on average 733 of 100,000 people per year die prematurely from all causes.
* ΔPM2.5 is the average annual increase in PM2.5 due to fires, measured in ug/m3 (derived from research question 2).
* ΒEvent is calculated as ΒEvent = ln (RREvent)/10 and
* RREventis the relative risk per increase in 10 ug/m3 PM2.5 concentration, sourced from literature review.

Equation 1 is used by O’Dell *et. al.* to estimate both asthma related ER visits and asthma related hospital visits caused by fire pollution. [15] It is also used by Ford *et. al.* to estimate premature mortality from *all* causes (such as asthma, non-communicable diseases, lower respiratory infections) that can be exacerbated by PM2.5 pollution. [9] For estimating premature mortality, O’Dell *et. al.* used a somewhat more computationally challenging method than Equation 1, but I stick with the Equation 1 approach. Ford *et. al* noted that Equation 1 approach is a close approximation to more sophisticated methods in the range of PM2.5 relevant to wildfires. [9]

Baselines rates and relative risk ratios were based on empirically estimated values. With one exception, I used the same estimated values as those used by O’Dell *et. al*., [15] which in turn were based on literature review and their own monte carlo estimations. The exception is related to the relative risk ratio for premature mortality, which I source from Ford *et. al*. [9]

**Table 1 – Baseline Rates Used in Health Impact Estimates**

|  |  |
| --- | --- |
| **Event** | **Baseline Rate  (per 100K people)** |
| Asthma-Related ER Visits | 625.7 |
| Asthma-Related Hospital Admits | 129.9 |
| Premature Mortality (All causes) | 732.9 |

**Table 2 –**

**Relative Risk Ratios in Health Impact Estimates**

|  |  |  |
| --- | --- | --- |
| Event | RR | 95% CI RR |
| Asthma-Related ER Visits | 1.07 | (1.03 – 1.11) |
| Asthma-Related Hospital Admits | 1.08 | (1.03 – 1.14) |
| Premature Mortality (All causes) | 1.06 | (1.04 -1.08) |

### *Findings*

As expected, health impact estimates vary quite a bit by year and follow a cyclical pattern, just as the smoke impact estimates did. Health Impact Assessments are small in absolute numbers – typically less than one instance per year – when applied to the approximately thirty-thousand residents of Kearney, as shown in Figure 8. Gray shaded areas represent the 95% confidence interval of the estimates, calculated by using the relative risk ratios, and resulting Beta values, from the 95% CI range for estimated relative risk ratios in Table 2. On a cumulative basis, fire related PM2.5 pollution is estimated to have caused 5 premature deaths, 5 ER visits for asthma, and 1 additional hospital admissions for asthma.

**Figure 8: Annual Fire Health Impacts by Cause 2006-2018, Residents of Kearney**

A graph of a patient with a number of numbers

Description automatically generated with medium confidence

**Figure 9: Cumulative Fire Health Impacts 2006-2018  
(Estimates in Red and 95% Confidence Interval in Blue)**

A graph of a number of individuals

Description automatically generated with medium confidence

## **R4: Comparison of Mortality by Cause**

### *Methods*

To place my R3 findings in context, I also estimated premature mortality of Kearney residents from selected other causes, including heart disease, traffic accidents, fire-arm related deaths, and homicides. As this data was only available at the state level, state-wide figures were scaled for Kearney using a ratio of Kearney population to the population of all of Nebraska.

### *Findings*

Non-pollution health concerns may be of higher priority than fire pollution health concerns for the residents and city council of Kearney. For example, premature mortality due to heart disease is orders of magnitude more significant than premature mortality from fire pollution. Nonetheless, premature mortality from fire smoke is approaching the same general range as from causes, such as homicides, that often garner policymaker attention and intervention. Moreover, premature mortality from wildfires is likely to increase, as discussed in the next section.

**Figure 10: Premature Deaths by Cause, Kearney 2006-2018**

A graph with numbers and a bar

Description automatically generated with medium confidence

## **R5: Predicting Future Mortality from Fire Related PM2.5**

### *Methods*

I predicted future premature mortality from fire PM2.5 pollution in two steps. *First*, I predicted future annual fire PM2.5 impacts based on the 13-year history of PM2.5 health impacts from step R2, using a simple linear regression model:

**Equation 2**: Fire Related PM2.5 = Β0 + Β1\* Year + ϵ

Notably, the linear model assumes no autocorrelation between the observed Fire Related PM2.5 values from year to year, an assumption which is violated. Both fire data and the resulting pollution is cyclical. If a longer time series were available, more sophisticated methods, such as ARIMA models might be more appropriate for modeling future impact. Given the relative data sparsity, however, I chose the more simple, linear model. Future predictions should be viewed with skepticism given the data and model limitations.

*Second,* I used the predicted Fire Related PM2.5 estimates as inputs to the Health Impact Equations discussed in step R3. Notably, this layered on a second level of uncertainty. I estimated uncertainty by creating 5000 bootstrapped samples from the 13 years of historical fire related impact data. For each sample I performed a multi-step process. (1) I performed a regression analysis on the bootstrapped sample and used the results to predict PM2.5 in future years. (2) To select the BetaMortality to be used in Equation 1, I randomly sampled a relative risk ration from a normal distribution centered at the point estimate of the mortality relative risk ratio in Table 2 (1.06) and with standard deviations derived from the 95% Confidence interval listed in Table 2 (1.04 – 1.08). I used the sampled relative risk ratio to calculate the BetaMortality to be used with the predictions in step 1. (3) Using the combined sample data from steps (1) and (2) I calculated the premature mortality in future years. I used the results from these bootstrap samples to construct 95% confidence intervals.

### *Findings*

The linear regression revealed only a modest correlation (R2 =.33) between fire related PM2.5 and year. Using the regression model to predict future annual fire pollution and using those results to predict future health impacts added another layer of uncertainty (Figure 12). Future estimates should therefore be viewed with caution. Nonetheless, the available evidence does suggest that fire related pollution, and the resulting health impacts, are increasing over time.

**Figure 11: Average Annual Fire Related PM2.5 Regression vs. Year and a Constant**

A graph with a red line and blue dots

Description automatically generated

**Figure 12: Predicted Annual Premature Mortality from Fire Related PM2.5**

**A graph with a line going up

Description automatically generated**

## **R6: Predicting Future Mortality from PM2.5 & Wildfire Attribute Data**

### *Methods*

I explored creating a model of smoke impact by combining observed PM2.5 levels for the 13 years available with the longer time series of fire attribute data. *First*, I created a model for smoke impacts, based on reasonable assumptions regarding how factors such as size and distance impacted smoke pollution on Kearney:

**Equation3**: PollutionFire = b\*(Acres\_Burned)^(Size\_Factor)/(Distance)^(Distance\_Factor)

Where:

* b is the "discount rate" applied to fires that were a prescribed burn. It is reasonable to assume that fires that are a prescribed burn contribute less to smoke pollution because they are (one would hope) likely set at times when conditions are more favorable in terms of carrying smoke away from population centers, more often set out of "fire season", and managed for intensity. The default value is set at 0.8 based on research showing that prescribed wildfires do contribute significantly to overall pollution; however more empirical research is required to understand the appropriate discount rate. In any event, I use the data from observed PM2.5 levels to help determine this value more precisely.
* PM2.5 levels are assumed to vary linearly with the number of acres burned, thus Size\_Factor has a default value of 1. A linear relationship is a reasonable assumption but could be wrong. For example, perhaps "big" fires are hotter at their center and throw off more smoke per acre. I again use observed PM2.5 data to help estimate this parameter.
* PM2.5 levels are assumed to vary in inverse proportion to the distance from the fire; in other words, the further away, the less the smoke impact. The Distance\_Factor in the denominator has a default value of 2 on the assumption that smoke dissipates in proportion to the square of the distance from the fire. Observed Data will also be used to refine this parameter.

Smoke Impact for the year is calculated as:

**Equation 4:** Annual Pollution = Factoring\_Constant\*ΣFiresIn\_Year PollutionFire

The Factoring\_Constant scales the results to match the relevant range of observed PM2.5.

*Second*, after setting this model structure, I used the observed values of fire related PM2.5 data from O’Dell *et. al.* to select the hyper-parameters b, Size\_Factor, and Distance\_Factor in Equation 3 that best correlates with the observed values of PM2.5 from a range of possible values:

**Table 3: Hyper-parameters Tested In Equation 3**

|  |  |  |
| --- | --- | --- |
| Parameter | Default Value | Values Tested |
| b | 0.8 | [0, .4, .8, 1] |
| Size\_Factor (numerator) | 1 | [0, 0.5, 1, 1.5, 2] |
| Distance\_Factor (denominator) | 2 | [0, 1, 2, 2.5, 3] |

I initially intended to use the wildfire attribute model with the tuned hyperparameters as a second way to estimate historical PM2.5 levels for a longer time series (1980-2020), and from there to estimate future PM2.5 and the resulting health impacts. As discussed below, however, I deemed the results from this step in the analysis insufficiently reliable to warrant further estimation.

### *Findings*

The highest correlation observed between observed PM2.5 and the estimates using wildfire attribute data was .43, at the following values of the hyper-parameters:

**Table 4: Hyper-parameters Resulting in Fire Attribute   
Model Best Correlating with Observed PM2.5**

|  |  |
| --- | --- |
| Parameter | Optimal Value |
| b | 1.0 |
| size\_factor (numerator) | 0.0 |
| distance\_factor (denominator) | 3.0 |

Not only was the “optimal” correlation (.43) underwhelming, the resulting hyper-parameters seem to have little theoretical justification. In particular, the size\_factor of zero means that each fire is treated equally regardless of number of acres burned which seems an unlikely result given what I know about the physical world. Common sense therefore suggests that the choice of hyper-parameters is more likely the result of random noise than underlying causal factors. For these reasons, I determined that the fire attribute method was not sufficiently reliable to use for creating future PM2.5 estimates and the resulting health impact estimates.

There are a number of possible explanations for the poor performance of the wildfire attribute model. The wildfire attribute dataset excludes Canadian wildfires, which also have an impact on Kearney. [14] The data does not contain information on wind or weather, factors which are likely to impact how smoke dissipates into surrounding areas. The data also does not contain information such as fuel load of a particular fire, another variable that has been identified as relevant to pollution emissions. [17].

# **DISCUSSION** **AND RECOMMENDATIONS**

This study concludes that PM2.5 pollution levels are elevated during periods where smoke plumes are visible overhead Kearney. Further, such fire related PM2.5 levels have been marching upward in recent years. This has health consequences. During the period from 2006 to 2018, fire related pollution likely contributed to approximately 5 premature deaths, 5 additional asthma-related ER visits, and 1 additional asthma related hospital admissions. Further, the relative health impact of fire related PM2.5 pollution is likely to increase in future years as wildfire activity continues to trend upward. Specific future estimates, however, should be viewed with caution given the levels of uncertainty in the data and models.

Although the health impacts from fire pollution are not as large as other health concerns such as heart disease, they are nonetheless significant. Any human life lost to pollution is too many. Thus, I recommend that the residents of Kearney and its city council take the following mitigation measures based on recommendations from Airnow.gov. [3]

* **For All Residents**
  + Stay indoors during high exposure days.
  + Use air filters indoors.
  + For residents who cannot afford air filters for their entire home, create a “clean room” for sleeping. Select a room with as few windows as possible and keep them closed. Fiter the air in that room.
* **For the City Council** 
  + Consider providing indoor spaces for the homeless residents of Kearny during high exposure days.

Many of these mitigation measures are relatively low cost, as is appropriate given my findings that other public health priorities may be more pressing.

# **RELATIONSHIP TO HUMAN CENTERED DATA SCIENCE**

This study was motivated and informed by human centered data science principles. *First*, the choice of topics is grounded in the concern for human beings – their health and longevity.

*Second,* the study emphasized real-world impact that would be meaningful to the residents of Kearney. The study’s historical estimates of premature morbidity and mortality are *statistically* significant, and many analyses may have stopped at that conclusion. But human centered data science principles spurred me to probe further. Even if fire pollution impacts are “large” in a statistical sense, it may still be preferable to allocate limited resources to other public health priorities if those public health priorities have an even larger impact or are more amendable to policy intervention. These considerations caused me to consider my estimates of premature mortality from fire pollution compared to other health concerns that the city of Kearney may be facing. I concluded that other public health concerns, such as heart disease, may be of more pressing concern. This desire to contextualize the data was inspired, in part, by the reflections of Wang [18] who advocates augmenting big data with “thick” data.

*Third,* as Duarte *et. al*. [7] pointed out (albeit in the slightly different context) models work best when trained on data that is as specific to the research question as possible. Researchers should not assume that data or models applicable in one context necessarily apply equally in another context. For this reason, when locating mortality data from causes other than wildfire, I took care to use the smallest level of generality of data available – the state of Nebraska – rather than more readily attainable data from a national level.

*Fourth,* my focus on the homeless population in Kearney when making recommendations was also motivated by human centered design principles. As pointed out by Hagendorff [11], ensuring that data science projects consider the interests of society’s underprivileged is a key component of data science ethics.

*Finally*, this study endeavored to use reproducibility best-practices by making data, code, and explanations available in the [project repository](https://drive.google.com/drive/folders/1xI5BOdkj-Avt1WA8ay1DfII35q5tgN6c?usp=drive_link). This enables peer review and allows others to extend the work in new ways. My study certainly benefitted from the fact that O’Dell *et. al.* also followed reproducibility best practices, making both their data and code publicly available. This enabled me to build on their work to customize the analysis for the residents of Kearney. It also enabled me to offer a potential critique related to the presence of negative daily pollution impact estimates in their analysis.

# **LIMITATIONS**

This study had several limitations. *First,* I only had access to satellite imagery data to detect the presence of smoke above Kearney on a particular day for a 13-year period. This renders any future predictions tenuous due to data sparsity. Moreover, the linear regression model assumes no autocorrelation between the observed smoke impact values in different years when in fact wildfire activity, and the resulting pollution, follows a cyclical pattern. In future research, if longer time series of data is available, researchers should consider alternate model specifications to account for this autocorrelation.

*Second,* I attempted to augment the 13 years of satellite imagery data by estimating smoke impact in earlier years via the Wildfire Attributes data, rather than observed readings of PM2.5 pollution, to create a longer time series of data from which to estimate future smoke impact. However, limitations in the Wildfire Attributes data resulted in the conclusion that joining the Wildfire Attribute data into the model might exacerbate, rather than decrease, uncertainty. Future research should attempt to incorporate wildfire data from Canada, and, if available, more specific information on fire dates, relevant weather conditions, and fuel type/load per fire.

*Third,* even with the 13 years of satellite imagery data that was available, my estimates of fire related PM2.5 pollution were imperfect. The fact that some daily estimates of fire related PM2.5 were negative made this clear. EPA air pollution readings could contain noise, and/or the kriging method used by O’Dell *et. al.* could over or underestimate daily pollution levels in the grid cell relevant to Kearney. Moreover, the assumption that nearby fires only impact PM2.5 pollution on days where smoke is visible could be flawed. PM2.5 particles blown into town by wildfire smoke could remain in the air for several days after smoke is no longer visible. Likewise, wind and weather patterns, neither of which were explicitly considered in the model, could impact the rate at which PM2.5 dissipates over time and distance.

*Fourth,* even if the fire related PM2.5 estimates were perfect,use of the Health Impact Assessment models injected additional uncertainty into the model. Each Health Impact Assessment equation rests on the assumption that health impacts can be represented using the equation: Δ Event = Pop\*BaselineEvent(1-e- Β\_Event\*ΔPM2.5). While this approach is widely used in literature, like all models it is necessarily an imperfect representation of reality. Further, constants used in the Health Impact Assessment equations are sourced from empirical studies. Where available, I used estimates from a meta-analysis of available relevant studies, which seemed likely to be the best estimates available. Nonetheless, the fact that different studies result in differing estimates and confidence intervals is testament to the fact that the estimates are imperfect.

*Finally,* the study assumes that the toxicity from fire-related PM2.5 is similar to the toxicity from other-source PM2.5. In fact, some research suggests that fire related PM2.5 may be more toxic than other-source PM2.5. [1, 15]

# **CONCLUSION**

Informed by human-centered data science principles, this project investigated the health impacts from wildfire pollution on Kearney, Nebraska. The research addressed the following questions:

|  |  |  |
| --- | --- | --- |
|  | Research Question | Summary of Findings |
| 1 | Using Satellite Imagery Data to Segregate PM2.5 Pollution Estimates into “Smoke Days” and “No Smoke Days”, do we observe statistically significant differences in pollution levels based on the presence or absence of smoke? | Yes, PM2.5 levels are higher on “Smoke Days” than “No Smoke Days.” Results are significant regardless of whether analyzed on a yearly, monthly, or daily basis |
| 2 | Using the presence or absence of smoke above Kearney as a proxy for whether elevated PM2.5 levels on that day can be attributed to wildfire, what were the estimated annual PM2.5 pollution impacts from fires near Kearney in the period 2006-2018? | Average Annual Fire Pollution Impact Per Day (calculated as if fire impact was equally amortized over the year) ranged between -.02 (effectively, zero) in 2010 to a high of .70 in 2013. The pollution exhibited an upward trend over time. |
| 3 | Using the results from (R2), what were the estimated health impacts of fire related PM2.5 on the residents of Kearney from 2006-2018? | During the period from 2006 to 2018, fire related pollution likely contributed to approximately 5 premature deaths, 5 additional asthma-related ER visits, and 1 additional asthma related hospital admissions. |
| 4 | How do the estimates of premature mortality from fire-related pollution found in (R3) compare to other causes of premature mortality impacting residents of Kearney? | Health impacts from fire pollution are smaller than from health concerns such as heart disease. However, premature mortality from fire pollution is estimated to be on the same order of magnitude (5) as mortality from homicide (13) over the period. |
| 5 | How are health impacts from fire related PM2.5 likely to evolve in the future? | Available evidence suggests that adverse health impacts from fire related PM2.5 pollution are increasing. But limitations in available data suggest that future predictions should be treated with caution. |
| 6 | Can estimates/predictions based on air quality data from 2006-2018 be improved by also incorporating information on wildfire attributes from 1963-2020? | Due to current data limitations, using wildfire attribute data to predict future PM2.5 levels and the resulting health impacts does not appear to be a fruitful method for estimation. Future research should look to augment existing wildfire data with data from Canada, and any available data related to weather conditions and fuel mass. |

Based on these findings, residents of Kearney should take low-cost mitigation measures where available, such as staying inside during periods of high wildfire smoke exposure. The city council should also consider providing indoor accommodations for the homeless population on high exposure days. Although more invasive and costly wildfire mitigation measures may not be warranted on the basis of this study alone, these findings should also be considered alongside other research and additional adverse impacts from wildfires, including decreased visibility, economic consequences, property damage, and direct threat to human life/health to those in close proximity to the fire. Future research should also continue to study the health impacts considered herein, making use of additional available data where feasible.

# **APPENDIX I – SENSITIVITY ANALYSIS**

The health impact estimates displayed in Figures 8 and 9 are used as inputs for research question 4 and 5 are based on estimates of annual fire related PM2.5 pollution calculated *as if* fire pollution were equally distributed over the entire year. This methodology potentially overestimates impacts compared to estimates based on discrete daily readings. The function Δ Event = Pop\*BaselineEvent(1-e- Β\_Event\*ΔPM2.5) is concave; thus, assuming many small increase in PM2.5 will produce a larger impact estimate than modeling a few larger values.

I employed several methods to test the sensitivity of my analysis to the simplifying assumption that fire related PM2.5 pollution was evenly distributed over the year. I concluded that the simplifying assumption was reasonable, though it should be noted that it can cause somewhat inflated estimates, especially at higher concentration levels.

*First,* I used my annual methods to calculate the US wide estimates of asthma-related ER visits and asthma-related hospital admissions for the period 2006 – 2018 and compared my results to those published by O’Dell *et.al.*[15, Table 1]. In the table below, the reported range represents the minimum and maximum values found for any years in the study period, 2006-2018.

**Table A1: Comparison of Fire Related Asthma ER Visits, Asthma Hospital Admits, and Premature Mortality Across the US, 2006-2018, O’Dell *et. al.* and this Study**

|  |  |  |
| --- | --- | --- |
|  | **This work** | **O’Dell et. al** |
| **Fire Related Asthma ER Visits** | 1428 – 10,076 | 1,300 – 5,900 |
| **Fire Related Hospital Admits** | 337 - 2,379 | 300 – 1,400 |
| **Premature Mortality** | 1441 - 10168 | 4,800 – 7,800 |

My estimates for Asthma ER Visits and Hospital Admits are roughly comparable to the O’Dell *et. al*. estimates at the lower end of the range but overstate the upper end of the range by a factor of 1.5 to 2. For Premature Mortality, the midpoint of the ranges are similar, but my range is much wider, almost 70% lower on the low end and 1.5 times larger on the high end.

*Second*, I inspected the shape of the function Δ Event = Pop\*BaselineEvent\* (1-e- Β\_Event\*ΔPM2.5) for the range of *Beta* and ΔPM2.5values in my study. Although the function is highly convex for high values of *Beta*\*ΔPM2.5,it is near linear using the values relevant to this study. Figure A1 plots the value (1-e- Β\_Event\*ΔPM2.5) using the largest value of *Beta* relevant to this study (0.73), and the range of ΔPM2.5 observed when calculates ted daily (0-35 ug/m3).

**Figure A1: Shape of The Function (1-e- Β\_Event\*ΔPM2.5) for  
 Largest Beta in study and relevant range of ΔPM2.5**

A graph with a line going up

Description automatically generated

*Third*, I also recalculated premature mortality estimates using daily levels of fire related PM2.5

pollution, with appropriate adjustments to the baseline values in Health Impact Equation to

correspond to daily, rather than annual baseline measures. I ran the analysis two ways, both

treating negative daily PM2.5 fire pollution values as zero, and leaving the negative values in the analysis. Figure A2 displays the results.

**Figure A2: Estimate of Premature Deaths from Fire Pollution by Calculation Method** A graph of different types of method

Description automatically generated

The annual method estimates are within a factor of two of the daily estimates in all cases except

for (1) in 2010 (likely due to the estimates hugging zero) and (2) the comparison between the   
annual method in 2006 and the daily method in 2006 when the negative daily values are retained.  
(Specific values and calculations available in the report repo.) The cumulative estimates over the 13-year period are similar, as shown in Table A2.

**Table A2: Cumulative Premature Mortality from Fires in Kearney, 2006-2018, by PM2.5**

**Impact Calculation Method**

|  |  |  |
| --- | --- | --- |
| Calculation Method | Cumulative Premature Mortality Estimate | % Change Compared  to Annual Method |
| Annual | 4.9 | -- |
| Daily (negative values retained) | 4.3 | -14% |
| Daily (neg. values treated as zero) | 5.5 | 12% |

In summary, the choice of annual vs. daily calculation methods can alter the results, but the differences are in most cases within a factor of two, and usually less. Neither method is clearly superior. While the daily calculation method would seem more precise, it introduces the problem of negative fire related estimates which are difficult to interpret.

# **APPENDIX II - REFERENCES**

[1] Aguilera, R., et al*.* Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. *Nat Commun* 12, 1493 (2021). <https://doi.org/10.1038/s41467-021-21708-0>

[2] Aziz, S, “A look at Canada’s wildfires in numbers and graphics over the decades”, Global News (July, 2021), available at https://globalnews.ca/news/8045796/canada-wildfires-yearly-trends.

[3] AirNow, “Extremely High Levels of PM2.5: Steps to Reduce Your Exposure”, available at https://www.airnow.gov/aqi/aqi-basics/extremely-high-levels-of-pm25/.

[4] Burke, M., et al. "The contribution of wildfire to PM2. 5 trends in the USA." *Nature* (2023): 1-6.

[5] Burnett, R. et al. "Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter." *Proceedings of the National Academy of Sciences* 115.38 (2018): 9592-9597.

[6] Center for Disease Control and Prevention, “Particle Pollution”, available at [https://www7.cdc.gov/air/particulate\_matter.html](https://www.cdc.gov/air/particulate_matter.html).

[7] Duarte, N., et. al. (2018). [Mixed Messages? The Limits of Automated Social Media Content Analysis](https://cdt.org/wp-content/uploads/2017/12/FAT-conference-draft-2018.pdf). Proceedings of the 1st Conference on Fairness, Accountability and Transparency, 81, 106.

[8] US Environmetal Protection agency [Why Wildfire Smoke is a Health Concern | US EPA](https://www.epa.gov/wildfire-smoke-course/why-wildfire-smoke-health-concern#:~:text=Fine%2C%20inhalable%20particulate%20matter%20%28PM2.5%29%20is%20the%20air,the%20lungs%20and%20may%20even%20enter%20the%20bloodstream.), available at https://www.epa.gov/wildfire-smoke-course/why-wildfire-smoke-health-concern#:~:text=Fine%2C%20inhalable%20particulate%20matter%20%28PM2.5%29%20is%20the%20air,the%20lungs%20and%20may%20even%20enter%20the%20bloodstream.

[9] Ford, B., et al. "Future fire impacts on smoke concentrations, visibility, and health in the contiguous United States." *GeoHealth* 2.8 (2018): 229-247.

[10] Google Maps, https://www.google.com/maps/dir/Kearney,+NE/49.1473591,-97.2621327/@44.8030359,-108.5341051,5z/data=!4m9!4m8!1m5!1m1!1s0x87994ebfa9ccdb73:0xdc488d11719c9a8e!2m2!1d-99.0816784!2d40.6993303!1m0!3e0?entry=ttu.

[11] Hagendorff, T. Blind spots in AI ethics. AI Ethics (2021). <https://doi.org/10.1007/s43681-021-00122-8>

[12] Kearney, Nebraska. In *Wikipedia*, <https://en.wikipedia.org/wiki/Kearney,_Nebraska>

[13] Krewski, D, et al. “Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality.” Vol. 140. Boston, MA: Health Effects Institute, 2009.

[14] NTV News, “*Air quality alert issued for Nebraska as Canadian wildfire continues*,” (May 18, 2023), available at https://nebraska.tv/news/local/air-quality-alert-issued-for-nebraska-as-canadian-wildfire-continues.

[15] O’Dell, K. et al. "Estimated mortality and morbidity attributable to smoke plumes in the United States: not just a western US problem." *GeoHealth* 5.9 (2021): e2021GH000457. Available: <https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021GH000457>.

[16] Odman, M., “Contributions of prescribed fire to air quality and health”, Health Effects Institute Research, available at <https://www.healtheffects.org/research/ongoing-research/contributions-prescribed-fire-air-quality-and-health>.

[17] Ravi, V, et al. "Impacts of prescribed fires and benefits from their reduction for air quality, health, and visibility in the Pacific Northwest of the United States." *Journal of the Air & Waste Management Association* 69.3 (2019): 289-304

[18] Wang, T. [Why Big Data Needs Thick Data](https://medium.com/ethnography-matters/why-big-data-needs-thick-data-b4b3e75e3d7). Ethnography Matters, 2016.

# **APPENDIX III - DATA SOURCES**

The following data sources were used in the analysis. More detailed descriptions of each source, and its limitations, can be found in the README file in the repository.

|  |  |  |  |
| --- | --- | --- | --- |
| **Data Desc.** | **Download Link(s)** | **File Name** | **License Terms  (if applicable)** |
| US Pop. Data, 2000-2010 | [National Intercensal Tables: 2000-2010 (census.gov)](https://www.census.gov/data/tables/time-series/demo/popest/intercensal-2000-2010-national.html) | us-est00int-01.xls | n/a |
| US Pop. Data, 2010-2020 | [National Population Totals: 2010-2020 (census.gov)](https://www.census.gov/programs-surveys/popest/technical-documentation/research/evaluation-estimates/2020-evaluation-estimates/2010s-totals-national.html) | nst-est2020.csv | n/a |
| Kearney Pop, 2000, 2010 and 2020 | [U.S. Census Bureau QuickFacts: Kearney city, Nebraska](https://www.census.gov/quickfacts/fact/table/kearneycitynebraska/PST045222) | n/a | n/a |
| Nebraska Pop. 2020 | [U.S. Census Bureau QuickFacts: United States](https://www.census.gov/quickfacts/fact/table/NE,US/PST045222) | n/a | n/a |
| Traffic Mortality for Neb. 2006-2018 | [FARS Encyclopedia: States - Fatalities and Fatality Rates (dot.gov)](https://www-fars.nhtsa.dot.gov/States/StatesFatalitiesFatalityRates.aspx) | n/a | [Terms of Use | NHTSA](https://www.nhtsa.gov/about-nhtsa/terms-use) (no restrictions) |
| Firearm Mortality for Neb.  2006-2018 | [US gun deaths (usafacts.org)](https://usafacts.org/data/topics/security-safety/crime-and-justice/firearms/firearm-deaths/),which in turn relies on data from the CDC [WISQARS Data Visualization (cdc.gov)](https://wisqars.cdc.gov/data/explore-data/home) | n/a | [Terms and Conditions | USAFacts](https://usafacts.org/terms-and-conditions/) (no restrictions) |
| Heart Disease Mortality for Neb. 2006-2017 | https://www.cdc.gov/nchs/data-visualization/mortality-leading-causes/index.htm | NCHS\_-\_Leading\_Causes\_of\_Death\_\_United\_States (1).csv | n/a |
| Heart Disease Mortality for Neb. 2018 | https://www.cdc.gov/nchs/pressroom/sosmap/heart\_disease\_mortality/heart\_disease.htm | n/a | n/a |
| Homicide Rates for Nebraska 2014-2018 | https://www.cdc.gov/nchs/pressroom/sosmap/homicide\_mortality/homicide.htm | n/a | n/a |
| US Geological Survey Wildfire Dataset | <https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81>. | USGS\_Wildland\_Fire\_Combined\_Dataset.json  USGS\_Wildland\_Fire\_Combined\_Dataset.csv | n/a |
| O’Dell et. al., data (PM2.5 and smoke plume) | <https://doi.org/10.25675/10217/230602> | krigedPM25\_2006\_v2.nc  krigedPM25\_2007\_v2.nc … krigedPM25\_2018\_v2.nc | [Research Data Terms of Use – Colorado State University: Libraries (colostate.edu)](https://lib.colostate.edu/find/csu-digital-repository/policies-guidelines-forms/research-data-terms-of-use/) |
| EPA Air Quality System API | [AQS API | AirData | US EPA](https://aqs.epa.gov/aqsweb/documents/data_api.html) | n/a | [AQS API | AirData | US EPA](https://aqs.epa.gov/aqsweb/documents/data_api.html#terms) |