Zachary Bowyer

Data 512

11/11/2021

**Extension Plan for Wildfire smoke impact on Leavenworth, Kansas**

**Motivation/problem statement:**

Leavenworth, Kansas is a small city [1] with a population of 37,086 as of 2022. While the first part of this data science project has shown that the air quality of Leavenworth has historically remained at healthy levels, the extension to that work plans on supplementing previous data with demographic data and research results to come up with a policy suggestion regarding the impact wildfire smoke has on the city’s healthcare sector. In specific, this work will address the relationship between exposure to wildfire smoke, elderly people, health insurance, and persons with long term health issues. The final results of this work will be an outlined report which will include findings, conclusions, and a final recommendation to the city of Leavenworth on whether or not this issue is currently important enough to warrant action.

The issue of how wildfire smoke impacts healthcare sector is specifically important to the community of Leavenworth Kansas is because according to [4], the healthcare industry makes up 16.4% of civilian employment for Leavenworth, which is slightly above the national average. Additionally, data from the US Census Bureau shows that the old age dependency ratio [5] has increased steadily from 2010 to 2021 [3], which is relevant because older people are more likely to have health conditions [6] which may be exasterabated by wildfire smoke [7]. Overall, with these facts in mind the authors believes that city officials would be interested in knowing whether or not the the future impact of wildfire smoke is an issue they should be concerned with due to the relative in increase in older populations. To give a potential example of an issue arising from smoke, if the smoke trend continues to increase, then the healthcare sector may have to increase its numbers of employees to keep up with the number of elderly people. However, if the proportion of elderly people is too high, then the city may have to hire externally.

This analyis is potentially interesting to the author because it will help improves the author’s understanding of the relationships between wildfire smoke and health, provide the author with an opportunity to work with US Census Bureau data [3], and be a good exercise for the author in trying to create a policy recommendation based on data and statistical techniques. With that being said, in this work the author hopes to learn how to use the US Census Bureau API and create a good policy recommendation.

**Impact focus:**

The specific focus area of this study will be on health care. As stated in the problem statement section of this document, the analysis will be focused on how the increasing elderly population who are more likely to have health issues that can increase the negative effects of wildfire smoke, could affect the healthcare industry in Leavenworth, Kansas.

**Data or model to be used**:

In terms of the data being used for this work, the two datasets [8, 9] from the first part of this project will obviously be reused for smoke and air quality analyis. The first dataset comes from [8], which in short contains wildfire data (polygons and metadata) that occurred from the 1800s to 2021. This wildfire dataset is planned to be used to estimate and project the smoke impact on Leavenworth using historical fire data and statistical techniques such as linear regression modeling and autoregressive integrating moving average (ARIMA) modeling. The reason the estimate smoke impact metric is important is to be able to show policymakers whether or not Leavenworth is at risk in the future of having a lower averaged air quality. In terms of permissable use, this dataset is “freely redistributable with proper metadata and source attribution” and has one requirement stated to be “The U.S. Geological Survey requests to be acknowledged as originator of these data in future products or derivative research.“ [8]

The second dataset comes from [9], which is an API designed “to obtain row-level data from the EPA’s Air Quality System database” [9].

The last dataset is the us census bureau [3].

Will look around for models that estimate health conditions based on age, and then smoke impacts based on health conditions and come up with health care impact metrics from there.

**Unknowns and dependencies**:

The smoke estimates suck

The aqi estimates suck

Assuming census data is mostly correct regardless of stated uncertainties

What basis am I going to use for health care impacts? Research papers? Arbitrary scale? Etc?

Would my proposed policies actually help with anything? Proof?

**Timeline to completion**:

Collect demographic data for the questions we want from US census of bureau

Estimate health care impacts on general population from smoke estimates

Estimate health care impacts on elderly population from smoke estimates

Predict and project health care impact to 2050

Summarize findings

Come up with policy recommendations for the city

Presentation slides

Final report

**References:**

[1] <https://nces.ed.gov/programs/edge/docs/locale_classifications.pdf>

[2] <https://lvcountyed.org/site-selectors/leading-employers/>

[3] <https://data.census.gov/>

[4] <https://statisticalatlas.com/place/Kansas/Leavenworth/Industries>

[5] <https://data.oecd.org/pop/old-age-dependency-ratio.htm#:~:text=age%20dependency%20ratio-,The%20old%2Dage%20to%20working%2Dage%20demographic%20ratio%20is%20defined,rates%2C%20fertility%20rates%20and%20migration>.

[6] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5732407/>

[7] <https://www.epa.gov/wildfire-smoke-course/why-wildfire-smoke-health-concern>

[8] <https://www.sciencebase.gov/catalog/item/61aa537dd34eb622f699df81>

[9] <https://aqs.epa.gov/aqsweb/documents/data_api.html>