DATA 512

Extension Plan Outline

Logan O'Brien

1. **Motivation and Focus:**
   The objective for my project extension is to dig deeper into the wildfire smoke analysis I did in Part 1 regarding Richland, Washington. I want to study how the smoke from these wildfires may impact the health of the residents of Richland, WA. This is an important topic for the city leaders to consider, as the findings, if valid, would guide their decision-making in how to allocate funds to promote the well-being of their citizens. For example, if the rate of hospitalization for respiratory illnesses is increasing, perhaps more hospital staff, beds, etc. should be provided. Additionally, the people of Richland would be very interested in learning how wildfires might impact their health in future years. Thus, there is a societal benefit to studying this topic further.

   More specifically, for my extended analysis, I wish to answer the question "How will wildfire smoke impact the number of hospitalizations for respiratory issues in Richland, WA in the coming years?" To answer this question, I plan on acquiring population data and healthcare data to supplement the wildfire data I gathered in Part 1. Then, I will find the correlation between the smoke and hospitalizations in the county that Richland is located in. I will also use a simple model (probably a regression model) to predict future hospitalization rates for the future years. Lastly, I will document my findings, visualize the results, and provide a discussion on some of the dominant limitations and assumptions of my analysis.

2. **Data or model to be used:**
   This project extension is centered around two new data sources. One, is population data. I plan on using annual population estimates from the United States Census Bureau www.census.gov for my analysis. Now, I conducted a thorough search through their website and located several datasets that would be useful:
   - First, I found annual county population totals for 2000 – 2010 (https://www2.census.gov/programs-surveys/popest/datasets/2000-2010/intercensal/county/co-est00int-tot.csv), 2010 – 2019 (https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-total.html), and for 2020 – 2022 (https://www.census.gov/data/tables/time-series/demo/popest/2020s-counties-total.html).
   - I also located some older data, for 1980 through 2000, but it is not organized as well and requires some manipulation. Additionally, the files contain population estimates

at the national level, not the county level, which is not as useful for my analysis. I may or may not use this data due to these challenges.
- I called a customer service agent and she said that the data provided is publicly available and that I am free to use it as long as I cite it.

The other data I want to utilize is healthcare data to get a measurement of the health of Benton County (which contains Richland, WA). It proved quite challenging to find publicly available healthcare data due to its sensitive nature. However, I discovered a great online tool called the Healthcare Cost and Utilization Project (HCUPnet) which provides data tools for analyzing and visualizing healthcare information, including hospitalizations and discharge diagnoses (see https://datatools.ahrq.gov/hcupnet/). This tool is provided by the Agency for Healthcare Research and Quality (AHRQ). In particular, I plan on using the tool to export the number of discharges for select conditions and diagnoses (like asthma) on a per-county level of the state of Washington. These exports contain several fields including:

- *Number of Discharges*
- *Average Length of Stay (in days)*
- *Rate of Discharges per 100,000 Population*
- *Age-Sex Adjusted Rate of Discharges per 100,000 Population*
- *Aggregate Hospital Costs (in $)*
- *Average Hospital Costs per Stay (in $)*

Note: I am unable to export data for multiple diagnoses at the same time, so I will have to manually merge the data to get annual overall counts of the number of respiratory discharges in Benton County. Additionally, I will make my decision on which conditions/diagnoses to pull for my analysis based on the categories that frequently appear in the relevant research literature on wildfire smoke impacts on health.

This data should be permitted for use. According to the webpage, users are prohibited from attempting to use the data to identify people or establishments. Additionally, users must agree to the following terms:

- *I will make no attempts to identify individuals, including by the use of vulnerability analysis or penetration testing. In addition, methods that could be used to identify individuals directly or indirectly shall not be disclosed, released, or published.*
- *I will make no attempts to identify establishments directly or by inference.*
- *I will not use deliberate technical analysis to discover or release information on small numbers of observations ≤10.*

- *I will not attempt to link this information with individually identifiable records from any other source.*
- *I will not attempt to use this information to contact any persons or establishments in the data for any purpose.*

3. **Unknowns and dependencies:**
   The only dependency that I am currently aware of is the necessity of acquiring the data for my analyses. However, as described above, I have found data that I can use and am confident that I will be able to get what I need.
4. **Timeline to completion:**

- Collect data – Target Completion Nov. 22nd
  - Gather the population estimate data and group the files into a single file.
  - Gather the hospital discharge data for the diagnoses of interest and combine them into a single file.
- Build a model – Target Completion Nov. 25th
  - Conduct a brief survey of various models that have been used in the current literature
  - Identify a model to use for my analysis. My hunch is that I will stick with a simple model.
- Visualize results – Target Completion Nov. 26th
  - Create charts illustrating the correlation between the smoke and hospital discharge counts, as well as other relevant charts depicting the predicted impact of smoke on hospital discharges.
- Create Slides for Presentation – Due Nov. 30th
- Document process, write final report, and complete project repository – Due Dec. 11

Select Research Articles I Have Consulted:

- [1]. *DeFlorio-Barker S, Crooks J, Reyes J, Rappold AG. Cardiopulmonary effects of fine particulate matter exposure among older adults, during wildfire and non-wildfire periods, in the United States 2008-2010. Environ Health Perspect 2019;127(3):37006. doi: 10.1289/ehp3860.*

- [2]. Liu JC, Wilson A, Mickley LJ, Dominici F, Ebisu K, Wang Y, et al.. 2017. Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties. Epidemiology 28(1):77–85, PMID: 27648592, 10.1097/EDE.0000000000000556. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

- [3]. Rappold AG, Stone SL, Cascio WE, Neas LM, Kilaru VJ, Carraway MS, et al.. 2011. Peat bog wildfire smoke exposure in rural North Carolina is associated with

cardiopulmonary emergency department visits assessed through syndromic surveillance. Environ Health Perspect 119(10):1415–1420, PMID: 21705297, 10.1289/ehp.1003206. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

- [4]. Delfino RJ, Brummel S, Wu J, Stern H, Ostro B, Lipsett M, et al.. 2009. The relationship of respiratory and cardiovascular hospital admissions to the southern California wildfires of 2003. Occup Environ Med 66(3):189–197, PMID: 19017694, 10.1136/oem.2008.041376. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

- [5]. Moore D, Copes R, Fisk R, Joy R, Chan K, Brauer M. 2006. Population health effects of air quality changes due to forest fires in British Columbia in 2003: estimates from physician-visit billing data. Can J Public Health 97(2):105–108, PMID: 16619995. [PMC free article] [PubMed] [Google Scholar]
    - https://link.springer.com/article/10.1007/BF03405325

- [6]. Liu JC, Pereira G, Uhl SA, Bravo MA, Bell ML. 2015. A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. Environ Res 136:120–132, PMID: 25460628, 10.1016/j.envres.2014.10.015. [PMC free article] [PubMed] [CrossRef] [Google Scholar]

- [7]. Ignotti E, Valente JG, Longo KM, Freitas SR, Hacon Sde S, Netto PA. Impact on human health of particulate matter emitted from burnings in the Brazilian Amazon region. Rev Saude Publica. 2010;44:121–130. [PubMed] [Google Scholar]
    - https://www.scielo.br/j/rsp/a/bp9BffF785sJmcC6hqX366d/?lang=en

- [8]. de Mendonca MJ, et al. Estimation of damage to human health due to forest burning in the Amazon. *J Popul Econ.* 2006;19:593–610. [Google Scholar]