Title: Exploratory Data Analysis of "2012 Workplace Fatalities by State"

Author: Nikki Lynn

Institution: University of Tennessee Professional Boot Camp through QuickStart - Data Science & Analytics Boot Camp

Instructor: James Meredith

Active Project Dates: 4/22/2024-5/5/2024

Introduction and Problem: The average person will spend approximately 100,000 hours at work over a lifetime. In 2012, approximately 116,000 persons were employed across 19 major employment sectors in the U.S., and 4,617 workers died in workplace incidents. In 1884, the post-Reconstruction and Industrial Revolution of the United States gave way to the rise of labor unions and The Labor Movement, resulting in the creation of the Nation Bureau of Labor and, ultimately, the U.S. Department of Labor (DOL.)   The DOL collects information about working people and the "means of promoting their material, social, intellectual, and moral prosperity." Thus, it is essential for the U.S. Department of Labor and Employers to track trends in workplace injuries/illnesses and fatalities and their relation to geographic location, penalties for failed working standards, and incident and compliance inspections.

Data: The dataset analyzed for this study is the "2012 Workplace Fatalities by State" provided by QuickStart Course Data Science and Analytics Boot Camp. The course creators curated the dataset from public data that is collected and reported annually through the Survey of Occupational Injuries and Illnesses (SOII) and the Census of Fatal Occupational Injuries (CFOI). The SOII and CFOI are created through the mandate of the U.S. Department of Labor, Bureau of Labor Statistics, Injuries, Illnesses, and Fatalities (IIF) program. Methods on the creation, conduction, coding, and data quality of the SOII and CFOI can be found on the Methods section of the IIF program site (https://www.bls.gov/iif/methods-overview.htm.)

The dataset used during the study includes both quantitative and qualitative data points from 2012 on the following: state name (including latitude and longitude coordinates), number of fatalities, rate of fatalities per capita, state rank based on the rate of fatalities per capita, number of non-fatal injuries/illnesses, rate of non-fatal injuries/illnesses per capita, penalties in FY2013 (average $), state rank of penalties in FY2012, number of inspectors in the state, years to inspect each workplace once, and whether this state has an Occupational Safety and Health Administration (OSHA)- approved State Plan or if they are a Federal OSHA state. The state data includes all 50 US states but does not include US districts or territories (Washington, DC., Puerto Rico, and Guam. State Plans are OSHA-approved workplace safety and health programs operated by individual states or U.S. territories. State Plans are monitored by OSHA and must be at least as effective as OSHA in

protecting workers and preventing work-related injuries, illnesses, and deaths. The SOII and CFOI, comprehensive in their coverage, are created through the mandate of the U.S. Department of Labor, Bureau of Labor Statistics, Injuries, Illnesses, and Fatalities (IIF) program. Methods on the creation, conduction, coding, and data quality of the SOII and CFOI can be found on the Methods section of the IIF program site (https://www.bls.gov/iif/methods-overview.htm.) The calculation method for the rate of fatality is found on the IIF web page here (https://www.bls.gov/iif/state-data/fatal-injury-rates-by-state-2007-2022.htm#BLStable_2023_12_4_17_42_footnotes.) This calculation method is also used to determine the rate of non-fatal injuries/illnesses per capita. This dataset includes eight states that do not have data points for the number of non-fatal injuries/illnesses and the rate of non–fatal injuries/illnesses per capita.

Methods:

Our analysis focused on relationships between State and Federal OSHA plans and the rate of injury/illness and fatality.   This was done by downloading the dataset into Microsoft Excel. Then, the structure of the set was reviewed with the dataset's data dictionary and initial descriptive statistics. From this information, the data quality of the set can be determined, and cleaning up the data can occur. Once the data is clean, the set's descriptive statistics are run again to help identify the type of visualization that will best answer the problems outlined in the introduction.

Results:

The analysis's results revealed that states whose safety programs are managed by OSHA's Federal Safety Plan had the highest rate of fatalities in 2012.
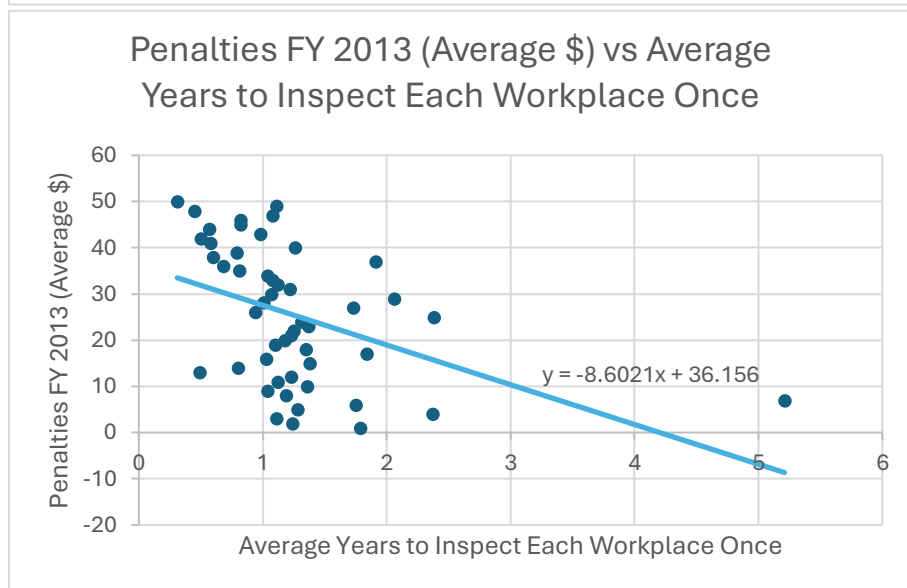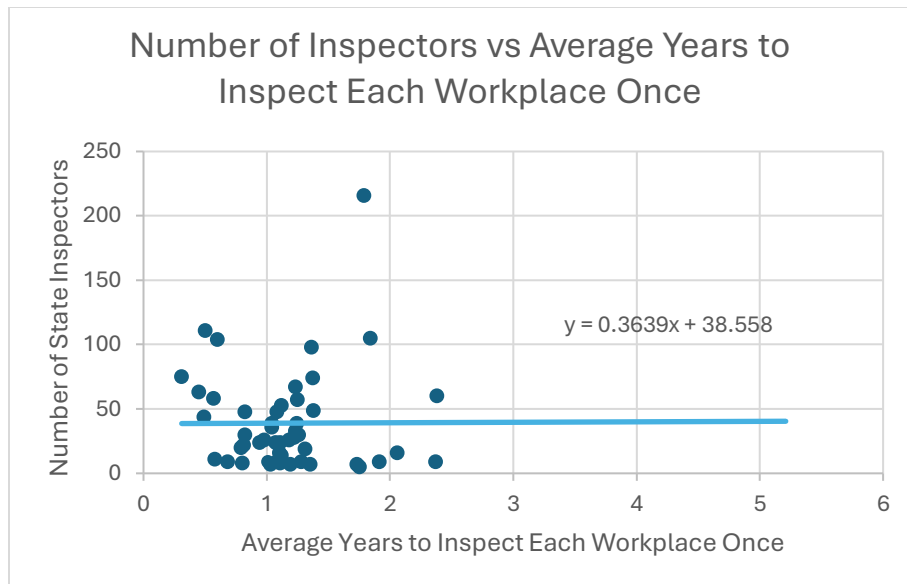

Total Rate of Fatalities by State or Federal Program
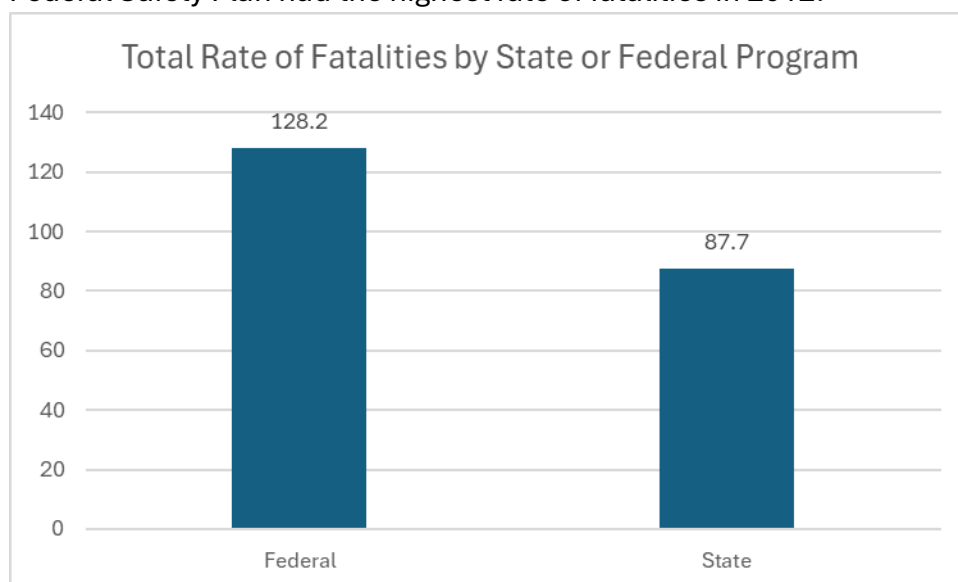
California had the highest number of injuries/illnesses of all the OSHA State Safety Plans in 2012.

Number of Injuries/Illnesses in 2012 by State Saftey Plans

89,300.00
42,900.00
6,500.00
67,5...
10...
45,600.00
32,400.00
27,700.00
77,...
345,4...
48,9...
66,2...
65,10...
75,900.00
54,400.00
19,900.00
36,...
9,700....

Powered by Bing
© GeoNames, Microsoft, TomTom

The data shows a positive relationship between the "Average of Years to Inspect Each Workplace Once" and the "Rate of Fatalities."

Rate of Fatalities vs Average Years to Inspect Each Workplace Once

$y = 0.0322x + 1.0999$

(y-axis: Rate of Fatalities in 2012, 0 to 6)
(x-axis: Average Years to Inspect Each Workplace Once, 0 to 20)
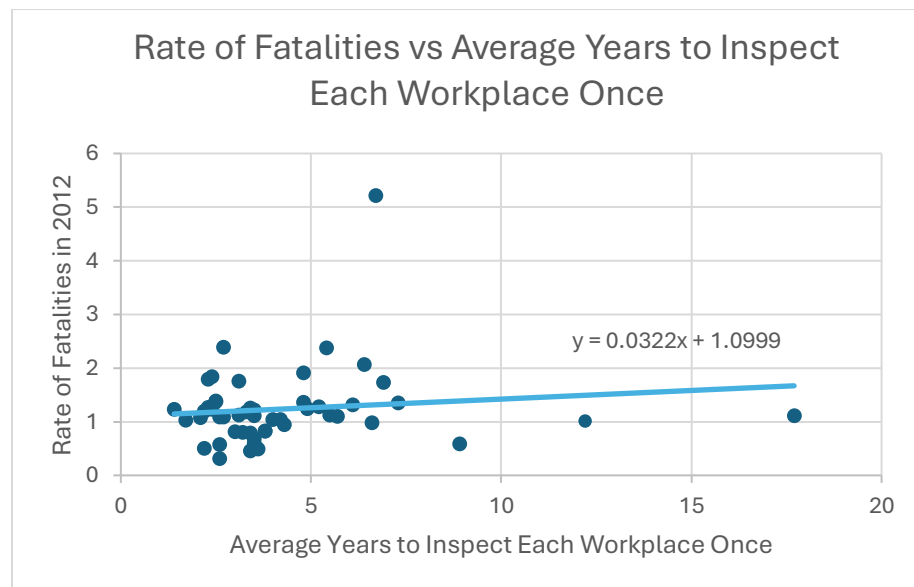
A positive correlation exists between the "Number of Inspectors vs. Average Years to Inspect Each Workplace Once" and a significant negative correlation between the "Number of Inspectors vs. Average Years to Inspect Each Workplace Once."

Number of Inspectors vs Average Years to Inspect Each Workplace Once

$y = 0.3639x + 38.558$

Number of State Inspectors (y-axis)
Average Years to Inspect Each Workplace Once (x-axis)

Penalties FY 2013 (Average $) vs Average Years to Inspect Each Workplace Once

$y = -8.6021x + 36.156$

Penalties FY 2013 (Average $) (y-axis)
Average Years to Inspect Each Workplace Once (x-axis)

The three charts above show that the number of penalties decreases as time passes. However, the X coefficient shows that the rate of inspectors is lower than the rate of fatalities.

Penalties FY 2013 (Average $) vs Average Years to Inspect Each Workplace Once

$y = -8.6021x + 36.156$
$y = 0.3639x + 38.558$
$y = 0.4653x + 3.7416$
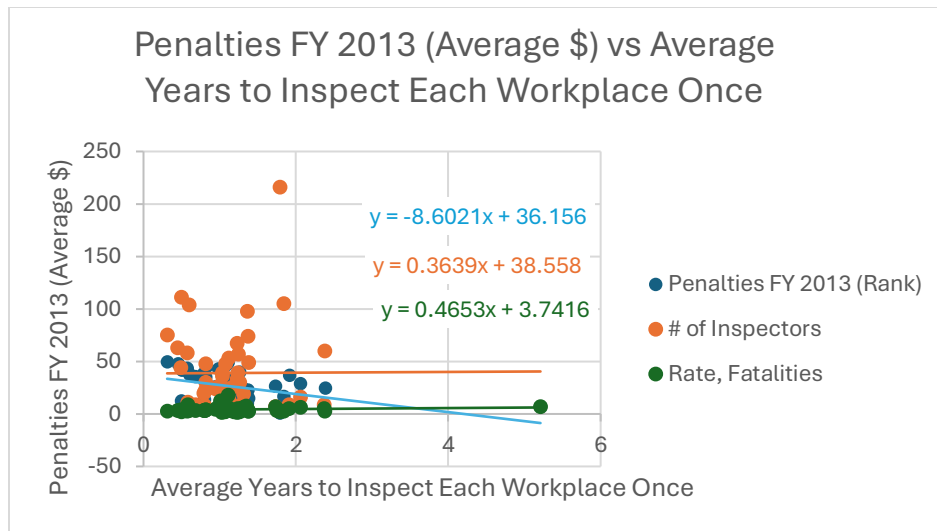
The "Rate of Fatality and Injury/Illness in 2012" chart shows the frequency of comparison between the two safety violation events on a proportional scale. The Rate of Fatality has a higher standard deviation, range, and confidence interval than the Rate of Injury/Illness. This led to the thought that in a prediction model, Injury/Illness would be more accurately predicted, and the variable could help determine a cause-factor correlation.

Rate of Fatality and Injury/Illness in 2012
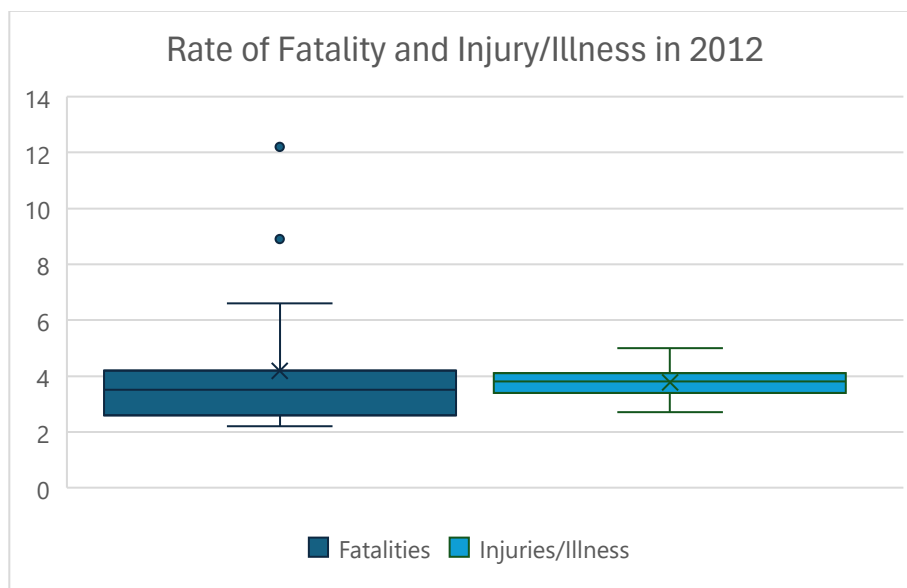
Conclusions:

Based on the analysis, the author has concluded that further study should be done to determine which of the OSHA Safety Plans are more effective, State vs. Federal. The data suggested from the initial analysis that no matter the safety plan, an increase in the number of inspectors in each state would reduce the rate of fatalities in reported safety events. More annual datasets should be reviewed to increase the study's sample size. Based on the analysis, the author would suggest further studies to determine which type of

safety event would best determine causation, fatality, or injuries in the larger population set.

For more information:

Please review the complete analysis in

For any additional questions, please get in touch with Nikki Lynn at nlynn@tvceh.org or on [LinkedIn]