Alfonso Rodriguez

Cohort-19A

2012 Workplace Fatality Analysis

Demand Overview

The request for this project was to complete exploratory data analysis on 2012 workplace fatality data and create an interactive visual for the user to explore the data for themselves. In addition, provide a summary of descriptive statistics.

The visual must be able to answer the following:

- Which program, state or federal, has the highest rate of fatalities?
- Which state with a state program has the highest number of injuries/illnesses?
- What is the relationship, if any, between "Average of Years to Inspect Each Workplace Once" and "Rate of Fatalities"?

Data Cleansing & Transformation

Utilized Excel to clean data by deleting null spaces. Added column 'Number of Businesses' from 2012 US Census report via manual data entry.

Power BI was used to create measures to calculate Businesses per Inspector. A measure was also used to make a treemap into a slicer.

Data Model:

This is a one table data model using variables:

State	State in the USA
Number of Fatalities	*Total # of Fatalities*
Rate of Fatalities	*Rate of Fatalities per 100k*
Rate of Fatalities State Rank	*State rank for Rate of Fatalities*
Number of Injuries/Illnesses	*Total # of Injury and Illnesses*

Injuries/Illnesses Rate	*Rate of Injury and Illness per 100k*
Injuries/Illnesses Rate State Rank	*State Rank for Rate of Injury/Illnesses*
Penalties FY 2013 (Average $)	*Avg $ of imposed penalties for workplace violations*
Penalties FY 2013 (Rank)	*State Rank of Avg $ Penalties Assessed in 2013*
Inspectors	*Total # of inspectors*
Years to Inspect Each Workplace Once	*Total time in years it would take to inspect each workplace once*
State or Federal Program	*Type of Regulation Program*
Number of Businesses	*Total # of Businesses*

Insights:

We are only working with one year of data so we are unable to see trend or make comparisons over time. We are unable to find many strong correlations between variables. Seeing as this is the case, I utilized means for relative comparison's sakes. Although the correlations were not strong there are some questions that we might ask that could allow to dig deeper if we had more dimensions to our data.

State programs have less variance in the number of Inspectors when compared to the Federal program and they also have a lower rate of fatality. It might seem that Federal programs would have more uniformity but based on this limited data that is not clear.

In order to further expand on this analysis, I would want to breakdown the data by industry type. I would also want to take a deeper look at more detailed dimensions of the outliers and see if there are any patterns that arise that we could then take and utilize to create and avoid standardized practices across both program types.