T Leah spencer  
osha/dwpp analysis

08/24/20

# Research available public reports for fy2014-fy2019

## Project Background and Description

|  |  |
| --- | --- |
|  | Thinkful’s DAI Cap3 Project: Identify dataset(s) of interest, complete analysis and research, and present findings.  The Occupational Safety and Health Administration’s (OSHA) Directorate of Whistleblower Protection Programs (DWPP) operates under the United States Department of Labor. The data for this project was first found at <https://www.whistleblowers.gov/factsheets_page/statistics> and related links in the Resources menu. I used Excel to extract data from .pdf’s and to scrape data from the embedded charts to create .csv files, viewable at <https://github.com/TLeahSpencer/DWPP_Analysis/tree/master/DWPP%20CSVs> |

## Project Scope

|  |  |
| --- | --- |
|  | This project will test several different combinations of the data with pearson r and t tests to find correlations in the totals of defined groups and statistical differences in their means. One overview that will be taken is comparing outcomes under a Democratic President 2014-2016 and a Republican President 2017-2019. I will also be looking at how outcomes relate to the Whistleblower Laws themselves, partnering agencies plus regulation requirements (e.g., provisions; allowable remedies; burden of proof; appeal venue; days to: file, complete, appeal). Other factors to be consider are a) CAA, CERCLA, FWPCA, SDWA, SWDA, TSCA statutes are all reported in the summary outcomes under EPA and b) TFA statute is only reported for FY2019. Finally, the Statutes themselves will be ranked by percentage of positive and negative outcomes. |

## Deliverables

|  |  |
| --- | --- |
|  | A Jupyter notebook for this project will include all the combinations of data tested with explanations of how and why they were combined. The t test hypotheses tested will be for no difference in the means of several groups (i.e., different political party presidents and other categories defined in the Project Scope). I expect to find some differences, but I am going to let the data tell me what they are, rather than guess at this point. The correlation analysis will be exploratory also, grouping and graphing to find correlations, and when found, they will be tested for statistical significance of p-value<0.05 and confidence level of 95% with pearson r tests. Significant difference in the means will be tested with t tests using the same criteria of significant difference being p-value<0.05 with a 95% confidence level.  A 5-10 minute presentation accompanied by a slide deck will summarize the methods and findings of this research. |

## Affected Parties

|  |  |
| --- | --- |
|  | My hope is to provide insight to whomever may want it: researchers with their own projects looking for supportive data, the public wanting to know more about how their government functions, and potential whistleblowers themselves looking to understand the odds of what the outcome of their cases may be if they decide to speak out about a situation they find themselves privy to. |