Work-related Injury and Illness

Alexandra Antonica Princewill Ayuk Mattia Sferazza Renato Mutavdzic

ABSTRACT

Workplace safety is a critical issue, and the ability to explore complex datasets effectively is essential for identifying patterns, prioritizing interventions, and fostering a safer work environment. Traditional static tools often fall short in analyzing multi-variable relationships within large datasets, limiting their utility for stakeholders such as safety officials, analysts, and policymakers. The proposed tool leverages interactive visualizations to enable users to examine key metrics such as incident rates, severity indices, and employee exposure rates. It features dynamic filtering, drill-down capabilities, and customizable dashboards, offering a user-friendly platform for identifying high-risk industries, exploring temporal trends, and detecting outlier companies. Tasks such as multivariate analysis across job categories and correlation between company size and injury types are addressed, supporting targeted safety interventions and informed decision-making. Preprocessing techniques included cleaning missing values, normalizing numerical features, and creating derived metrics such as the Severity Index and Employee Exposure Rate for enhanced analysis. The tool's implementation follows visualization design principles, utilizing Python libraries like Dash and Plotly to ensure usability and clarity. The project demonstrates the potential of interactive visualization in uncovering actionable insights from complex datasets, with applications ranging from regulatory focus to company-specific safety improvements. Future enhancements include integrating predictive analytics, improving outlier detection visualizations, and enabling comparisons with external datasets, broadening the tool's impact on occupational health and safety research.

1 INTRODUCTION

This report focuses on visualizing workplace injury and illness data from the US Occupational Safety and Health Administration (OSHA). Workplace safety is a critical concern, and understanding the prevalence and types of work-related injuries and illnesses across different industries and company sizes is crucial for developing effective safety interventions. The OSHA dataset, obtained through freedom-of-information lawsuits by Reveal and Public Citizen, provides a valuable resource for this investigation. However, raw data tables can be challenging to interpret, especially when dealing with large datasets that include multiple variables such as injury type, industry classification, and company size. Existing tools for exploring workplace safety data often lack the capability to dynamically analyze relationships between these variables, which limits their usefulness for identifying complex patterns or actionable insights. For instance, many tools provide static reports or generic charts without enabling users to explore cross-variable interactions, such as how injury frequency varies by company size within a specific industry. To address this gap, we present our visualization tool. The tool is designed to support safety officials, analysts, company safety officers and researchers in occupational health. The tool allows users to interactively examine the Work-Related Injury and Illness dataset, focusing on specific metrics such as incident rates by company size, injury categories by industry, or trends over time. Through features like dynamic filtering, drill-down capabilities, and customizable dashboards, users can uncover precise patterns and relationships, enabling them to pinpoint high-risk areas and prioritize interventions. This tool aims to address the domain problem of effectively analyzing complex workplace safety data. Visualization is the right choice for this goal because it enhances pattern recognition, facilitates comparison, promotes exploration, and improves communication. Visual representations make it easier to identify trends and outliers, while interactive features allow users to dive deeper into the data. By combining statistical analysis with interactive visualization, we can empower users to explore the data, generate hypotheses, and draw meaningful conclusions about workplace safety.

2 DATA ANALYSIS AND ABSTRACTION

Table 1 *Descriptive Statistics of the Dataset*

Feature	Mean	Median	Standard Deviation
Establishment ID	451380.2367	451770	259802.3544
Establishment Name	/	/	/
Zip Code	/	/	/
NAICS Code	25950000	53150	132100000
Annual Average Employees	496292.149	3	135696.455
Total Hours Worked	2021.0126	347	12.8735
Incident Outcome	38.2003	0.9949	2034000
DAFW Number Array	3528000	3	24140000
DJTR Number TR	0.8384	0	0.6988
Type of Incident	1.1924	0	1.3768
Time of Incident and Time Unknown	2.9735	1	10.8392
SOC Reviewed	981.0839	/	690.3626
Year of Filing	2023	2023	0
Incident Descriptions	/	/	/

For data preprocessing, we have removed the rows that contained missing or unknown values in all numerical fields to ensure the ability to use it in visualizations. This includes removing null values by removing the rows containing them, which allows us to better approach task analysis. Furthermore, date columns were converted to DateTime format, for easier manipulation. Only valid states were kept, adding a column for full state name. In addition, industry groups were extracted and incident counts grouped. Some additional features were created, such as:

Incident Rate (Total Incidents*20000 / Total Hours Worked): This provides a normalized measure of incident frequency, making it easier to compare across establishments regardless of their size. It highlights workplaces with disproportionately high incident occurrences for targeted safety interventions.

Severity Index (DAFW Number Array and DJTR Number TR): This combines two critical indicators of incident impact into a single measure, providing a more comprehensive understanding of the severity of incidents. It is useful for prioritizing high-severity workplaces for immediate attention and resource allocation.

Employee Exposure Rate (Total Hours Worked / Annual Average Employees): This feature identifies establishments where employees may be overexposed to potential hazards due to higher workload demands. It helps in detecting workload imbalance and guiding adjustments to improve employee safety.

Finally, total_hours_worked variable was scaled with a MinMaxScaler, to normalize the range for comparability. Some specific post-processing was done, based on the task at hand.

3 Task Analysis

Users of our visualization tool, including safety officials, safety analysts, company safety officers, policymakers, and researchers in occupational health, aim to extract actionable insights from the work-related Injury and Illness dataset. These users require a range of analyses, from broad overviews to detailed examinations of specific factors affecting workplace safety. Our visualization tool is designed to support these tasks, facilitating comprehensive analysis and informed decision-making.

Focusing on specific industries or regions, users might investigate the Severity Index (DAFW Number Array and DJTR Number TR) to understand the impact of incidents. This metric combines critical indicators of incident severity, allowing users, **such as safety officials** to identify not just the frequency but also the seriousness of workplace incidents. Such insights are essential for developing targeted prevention strategies and allocating resources effectively, **as for example company safety officers may be able to allocate resources differently based on the time the injured employees are unable to work.**

Company safety officers may be interested in analyzing the Employee Exposure Rate (Total Hours Worked / Annual Average Employees) to assess how employee workload correlates with injury rates. **This would help them** identify establishments where high employee exposure may contribute to higher incident rates, prompting interventions to improve workload distribution and enhance safety measures.

Temporal analysis is another critical task, where users assess how injury and illness rates have evolved over time. **Researchers in occupational health may benefit from visualizing** trends and changes, evaluating the effectiveness of past safety interventions and identifying emerging risks. **This may also help policymakers** in understanding the long-term impact of policies and guiding future safety initiatives.

Geographic distribution analysis allows users to uncover regional patterns in workplace safety. By mapping injury and illness rates alongside the Safety Inspection Rate, users can identify regions with high incident rates and low inspection rates, indicating potential areas for regulatory focus and enhanced safety programs. Identifying outlier companies with unusually high or low injury rates compared to industry averages is crucial for both regulatory purposes and the dissemination of best practices. **Policymakers may want to identify these** outliers effectively, to facilitate targeted interventions for high-risk companies and highlighting exemplary safety practices for others to adopt.

To summarize, the following tasks can be formulated:

1. Determine **geographical patterns** : Compare Incidence Rates, Severity Index, **number of companies and employee exposure rate** to aid in prioritizing regulatory attention and resource allocation, **as well as identify regional patterns.**

2. Investigate prevalent injuries and illnesses: Analyze types of injuries and illnesses common in different industries to identify hazards and develop targeted prevention strategies.

3. Correlation between company size and injury rates **and injury types**: Examine how company scale affects incident rate and **incident types over time, to** determine where support may be needed.

4. Identify outlier companies: Use the incident rate to pinpoint companies with unusually high or low injury rates for targeted regulatory attention or to highlight best practices.

5. Multivariate Analysis across Job Categories: Analyze occupational risk factors across job categories considering the relationships between incident rate, severity index, and employee exposure rate. The insights support risk assessment and decision-making for workplace safety improvements.

By addressing these tasks, users can gain comprehensive insights into workplace safety, identify areas for improvement, and make informed decisions to enhance occupational health.

4 Task Abstraction

~~Building on the task analysis, the following task abstraction has been developed. The first task identifies high-risk industries and regions by comparing Incident Rates and Severity Index, which focuses on extremes in the data to prioritize regulatory attention. The second task investigates the types of injuries and illnesses across industries, summarizing their distribution to identify common hazards. The third task examines the correlation between company size and injury rates, discovering relationships that highlight how scale affects safety outcomes.~~

~~Going more into detail, the fourth task assesses the effectiveness of safety interventions by exploring temporal trends in Incident Rates and Severity Index, helping to evaluate the impact of policies. The fifth task focuses on geographic patterns, searching and locating regions with high Incident Rates and low Safety Inspection Rates to uncover spatial safety gaps.~~

~~The sixth task identifies outlier companies with unusual injury rates by locating extremes in the Severity Index, which can indicate either significant risks or best practices. The seventh task analyzes multivariate relationships among factors like industry type, company size, and Safety Inspection Rates, exploring dependencies to uncover complex interactions. Finally, the eighth task detects emerging risks in specific injury categories by discovering new or increasing trends, enabling proactive intervention.~~

The domain-specific tasks outlined for the visualization tool can be generalized into the following abstract actions and targets based on Munzner's Task Abstraction framework:

- **Determine geographical patterns**
  - **Action:** Compare
  - **Target:** Extremes
- **Investigate prevalent injuries and illnesses**
  - **Action:** Summarize
  - **Target:** Distribution
- **Examine correlation between company size and injury rates**
  - **Action:** Discover
  - **Target:** Correlation
- **Identify outlier companies**
  - **Action:** Identify
  - **Target:** Extremes
- **Analyze multivariate relationships across Job Categories**
  - **Action:** Discover
  - **Target:** Correlation

5 Final Solution

The Work Injury Dashboard was designed to adhere Munzner's principles of visualization design and analysis and with the user in mind at every stage and to adhere to. The design follows key visualization principles, maintaining simplicity while maximizing usability, enabling users to explore data characteristics, identify patterns, and investigate specific aspects with ease. The layout supports task-driven interaction by placing the primary visual area at the center, while a structured left-hand menu (Figure 1) clearly organizes tabs, each labeled with informative titles to guide navigation. This follows the principle that "eyes beat memory," reducing cognitive load by making relevant options immediately visible.

Each tab is dedicated to a key task identified in the task analysis, allowing users to efficiently focus on their specific goals while still having the flexibility to explore other aspects of the data. The dashboard includes Plotly's interactive features, such as zooming, panning, and lasso selection, enhancing data exploration and analysis. Furthermore, compared to our original solution, we have decided to merge certain tasks into the same tab, to improve how easy the application is to use and since upon further inspection, certain tasks benefitted from interactions with each other.

5.1 Tab 1: Geographical Analysis

The interaction with our Dashboard begins on tab one (Figure 2), which showcases a choropleth map. This tab is set as the default tab to provide an immediate overview to the users - as the choropleth map displayed may be of interest to all types of users of our dashboard. The choropleth map is centrally positioned with a title and a short description, ensuring effective communication by adhering to Tufte's principle of maximizing the data-ink ratio and minimizing chart junk. It displays the U.S. states with clear borders and is color-coded based on the continued color range shown on the left. The chosen color range is designed to be perceptually effective, considering luminance contrast to enhance discriminability and accessibility for color-blind users.

Safety officials can interact via a drop-down above the map to select features such as incident rate, severity index, employee exposure rate, or number of companies. Each interaction updates the choropleth map dynamically and the present drop-down always reflects the currently selected feature to reduce the amount of information the user needs to remember.

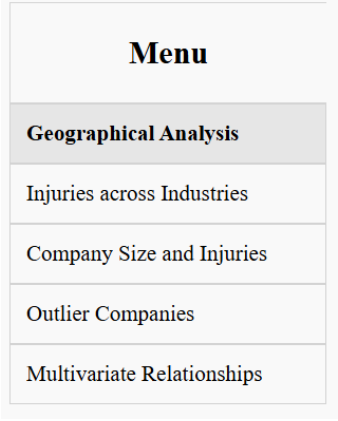

Figure 1: Menu with different tabs.

5.2 Tab 2: Injuries across Industries

The second tab on our menu features a pair of interactive stacked bar charts designed to illustrate the distribution of injuries and illnesses across different industries, as can be seen in Figure 3. The dataset categorizes companies using the North American Industry Classification System (NAICS), where the first two digits of a company's NAICS code indicate its industry group.

Users can interact with dropdown menus to filter the first chart by industry and incident type as can be seen in Figure C, while the Severity Index Slider allows for further refinement. The first chart visually encodes the number of occurrences of each incident type across selected industries, employing stacked bars to support part-to-whole comparisons—a key principle in visual encoding. Clicking on an industry's bar dynamically updates the second chart, which reveals sub-industries (full NAICS codes) while maintaining the same encoding for consistency. Both charts incorporate hover interactions, reducing cognitive load by allowing users to access specific details on demand rather than relying on memory.

5.3 Tab 3: Company Size and Injuries

This tab showcases a group of three charts: violin plot, bar plot, and line chart, which work towards discovering the correlation between company sizes and injury and illnesses types. The violin plot, illustrated in Figure 4, displays the distribution of incident rates for different bins. Each bin corresponds to a size where the companies are categorized into in the original dataset. The violin plot still maintains the standard boxes of a box plot, to be more informative for the user. The user can only choose one of the 4 size categories (Small (0–19 employees), Medium (20–99), Large (100–249), and Huge (250+), which will then display a bar chart showing the frequency of each incident type and a line graph displaying how the incident types have changed over time, for all companies within that size category. This allows users like Safety analysts and Researchers to look more in depth into the effect that size has not only incident rate but incident type. Furthermore, the line graph also allows us to see how incident types evolve over time, allowing us to spot and mitigate potential trends for specific incident types. In all of these plots, the user can also deselect certain injury types, as often not the category 'injury' tends to skew the data; note also how the injury types have been set to a standard color, so to minimize the user's attention in re identifying which incident type each color represents.

5.4 Tab 4: Outlier Companies

This page presents a set of interactive charts—a scatter plot and a line chart—designed to identify outlier companies in terms of incident rate (Figure 5). These visualizations support policymakers in detecting companies with abnormally high incident rates, allowing for targeted interventions to mitigate risks.

Users can select one or more industry groups, structured similarly to those in Tab 2, which dynamically updates the scatter plot to display the top 5% of outlier companies. The scatter plot was specifically chosen due to its effective visual encoding for quantitative comparison, ensuring that extreme values stand out clearly, by using the highest ranked channel of the spatial region. Selecting a company from the scatter plot updates the line chart, which then reveals its incident rate trends over 2023.

Both charts incorporate interactive hover functionality, allowing users to access detailed company-level information—including industry name, company name, incident rate, severity index, and incident date—without cluttering the visualization. This design adheres to Tufte's principle of maximizing data-ink ratio, ensuring that every visual element contributes to meaningful analysis while maintaining clarity.

5.5 Tab 5: Multivariate Relationships

Tab 5 explores the relationships between multiple workplace risk factors—Incident Rate, Severity Index, and Exposure Rate—across job categories using an interactive Parallel Coordinates Plot (PCP), Figure 6. The visualization categorizes jobs into 22 groups, based

on the first two digits of the Standard Occupational Classification (SOC) codes.

Users can customize the visualization by selecting specific job categories from the dropdown menu. The PCP supports brushing interactions, allowing users to filter specific value ranges by clicking and dragging vertically along any axis, dynamically updating the plot. Additionally, axes are draggable, enabling users to rearrange them and explore different perspectives on variable correlations. These interactive capabilities align with task abstraction principles, providing flexibility in exploratory data analysis.

This visualization is particularly beneficial for occupational health researchers, enabling them to identify patterns, trends, and correlations across multiple dimensions. The use of consistent color mapping for job categories enhances readability while maintaining perceptual uniformity, ensuring that users can efficiently analyze complex relationships within the dataset.

6 IMPLEMENTATION

Our tool was developed in Python, using mainly the Dash library for the interface and the Plotly library for creating the visualizations. We made use of the Dash template provided as a resource by the course professors. We structured our code in such a way that each tab is defined using a different .py file, making it easier for the development process and isolating different challenges that appeared for specific tasks. Regarding the visualization in Tab 1, the choropleth map raises an issue when computing the Employee Exposure Rate for the state 'TX' (Texas), resulting in a NA value and in a grey non-significant area on the map. Another issue that persists is in Tab 2, which concerns the unjustified double appearance of the slider.

While implementing our tool, we came across some limitations of the Plotly library, especially regarding the Parallel Coordinate Plot (PCP) in Tab 5. Despite our best efforts and desire, we were not able to create a discrete list of colors for the job categories, this resulting in a possible confusion when trying to identify the specific color corresponding to the categories. Moreover, we need to pinpoint the absence of the hovering feature in the predefined Plotly function we have used to create the PCP.

7 USE CASE

7.1 Use case 1: Determine geographical patterns

A safety regulator must identify regions with particularly high or low workplace incident rates. This information will guide decisions on where to allocate inspection teams and focus safety campaigns.The analyst opens the Geographical Analysis dashboard within the visualization tool. A dropdown menu at the top allows the user to select which metric to map, such as "Incident Rate," "Severity Index," or "Employee Exposure Rate. The analyst chooses "Incident Rate" from the dropdown. The map updates to display the United States, with each state shaded according to its incident rate. A color gradient (from yellow to purple) indicates the range, where yellow suggests lower rates and purple indicates higher rates. By visually scanning the map, the analyst quickly spots states with extreme colors. Hovering over each state provides a tooltip with the exact incident rate. The user observes that some states in the Southeastern U.S. show notably darker shades, indicating higher incident rates. Certain states in the Northeastern region appear to have comparatively lower incident rates.

The map clearly highlights these extremes, giving an immediate visual cue for where to focus attention. The analyst may switch to "Severity Index" in the same dropdown to see if the same states also have severe incidents, reinforcing the need for intervention. Identifying these regional patterns helps prioritize where immediate safety interventions are most necessary and highlights places that could serve as models for best practices (Figure 2). The color-coded approach simplifies comparison across all states at once, significantly accelerating the discovery process.

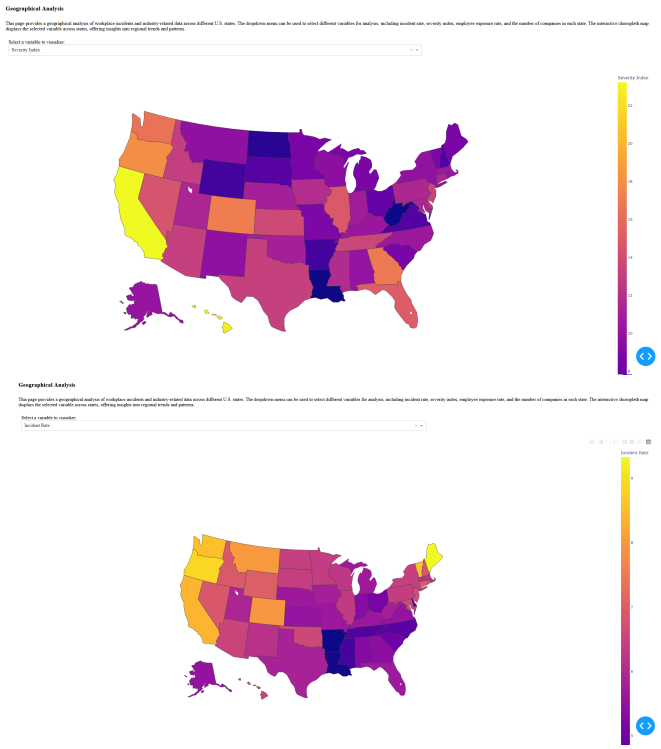

Figure 2: Choropleth map of incident rate and severity rate.

7.2 Use case 2: Investigate prevalent injuries and illnesses

A safety analyst needs to compare the frequency of different incident types across multiple industry groups and then explore specific sub-industries for deeper insight. The analyst opens the "Injury Types across Industries" dashboard. They see three filter controls at the top: Severity Index, Industry Group, and Type of Incident. The analyst selects Wholesale Trade (42), Retail Trade (44, 45) from the Industry Group list, and checks several incident types (Figure 3). A stacked bar chart shows how many incidents each selected industry has, with color-coded bars distinguishing the types. Upon clicking the bar for Retail Trade (NAICS 44), a second chart displays sub-industries and their corresponding incident counts. The Severity Index slider can further highlight high-impact cases, letting the analyst focus on the most critical incidents. The analyst notices that one sub-industry stands out with most of the incidents caused and that this sub-industry has an unusually high count for the incident type "Respiratory Condition". By summarizing these distributions across industries and zooming into sub-industries, the analyst pinpoints which sectors and incident types warrant further safety measures. Through these filters and drill-down visualizations, the analyst gains a clear overview of incident types at both the industry and sub-industry levels, supporting more targeted interventions.

7.3 Use case 3: Correlation Between Company Size and Injury Rates

An analyst wants to see if larger establishments truly have higher incident rates and whether certain incident types dominate in specific size categories. The analyst opens the "Company Size and Injuries" dashboard, which displays a violin plot of incident rates for four establishment size bands (Figure 4). Each violin plot shows how incident rates are distributed within that size category. By scanning the plots, the analyst sees that "Large" and "Huge" establishments

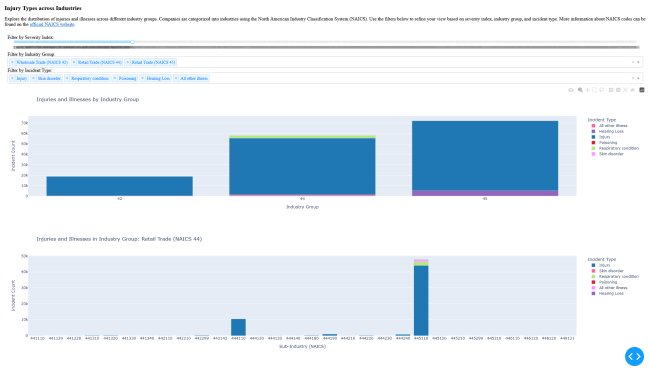

Figure 3: Use case 3: Injury Types Across Industries.

have broader distributions of incident rates, some extremely high, some moderate. Conversely, "Small" companies appear to have generally lower medians, but with a few outliers at the very high end. Intrigued by the spread in large establishments, the analyst clicks the "Large" filter in the second view. A bar chart now shows the frequency of each incident type for large companies. An accompanying line chart tracks how these incident types have trended over time. "Injury" incidents dominate among Large establishments, with counts vastly outstripping other categories like hearing loss or respiratory conditions. The trend line for "Injury" peaks around mid-2023, then begins to decline, suggesting a recent improvement or changed reporting practices.

This exploration confirms that while some large workplaces have very high rates, the problem is not uniform. Many are near the lower end. The analyst may propose more targeted interventions or additional follow-up to understand why certain Large establishments report far higher injury counts. The time-series visualization offers clues about when new safety measures might have taken effect or if seasonal factors influence these trends. By comparing distributions across company sizes and then drilling down into specific incident types and time frames, the analyst gains a nuanced view of how establishment size relates to both frequency and nature of workplace incidents.

7.4 Use case 4: Identify outlier companies

A safety analyst wants to pinpoint which companies far exceed normal incident rates within a given industry. By identifying such outliers, they can focus inspections and resources where they're needed most. The analyst opens the Outlier Companies dashboard and chooses "Agriculture, Forestry, Fishing and Hunting (NAICS 11)" from the dropdown. A chart appears showing the Top 1% of outliers in that industry based on incident rate (Figure 5). On the box plot, a few points appear far above the industry's typical range. One of those companies is LEGACY FRUIT LLC, which stands out with an extraordinarily high incident rate. Clicking on the outlier company reveals a time series graph indicating how its incident rate fluctuates (or remains stable) throughout 2023.

In this example, the line remains consistently high around 2,084, confirming that the company's rate far exceeds industry norms over the entire year. Seeing this pattern, the analyst can flag LEGACY FRUIT LLC for immediate follow-up. By highlighting extreme incident rates and then displaying a time-based breakdown, the tool makes it easy to see which outlier companies warrant top-priority attention.

Figure 4: Use case 4: Company Size and Injuries.

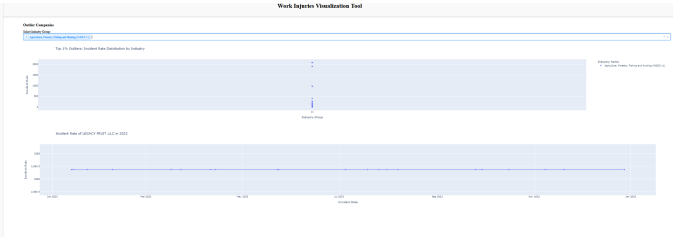

Figure 5: Use case 5: Outlier Companies.

7.5 Use case 5: Multivariate Analysis across Job Categories

A researcher wants to see how Incident Rate, Severity Index, and Exposure Rate relate to different job categories. They suspect that certain roles might have both higher incident rates and severe outcomes, and they need a visual way to confirm any correlations. In the "Multivariate Analysis Across Job" dashboard, the researcher sees a multi-axis chart displaying all job categories (Figure 6). Each line runs across three axes—Incident Rate, Severity Index, and Exposure Rate, with a color legend that assigns a unique color to each job category. By scanning the web of lines, the researcher can spot initial patterns, like certain categories (e.g., Transportation, Construction) consistently trending toward higher Incident Rates or Severity Index values.

The cluttered lines, however, make it tough to isolate specific jobs right away. The researcher filters the chart to four job categories (Figure 7). For example, Management Occupations, Business and Financial Operations, Office and Administrative Support, and Sales and Related. The visualization updates, removing other lines. Now only the four categories appear, making it easier to compare their trends across the three axes. The researcher sees that Management Occupations might have a moderate Incident Rate but a high Severity Index—indicating that while incidents aren't as frequent, they can be more serious. Conversely, Sales and Related might have a slightly higher Incident Rate yet a lower Severity Index, pointing to more frequent but less severe incidents. This targeted comparison lets the researcher quickly discover correlations between different factors.

Insights from the filtered view can guide further inquiries. For example, investigating why managerial roles lead to high-severity incidents or why certain administrative roles show particular exposure rates. The ability to toggle job categories on and off helps overcome visual clutter and focus on the relationships of interest. By discovering how job categories differ across multiple variables at once, the researcher gains clear evidence for which roles deserve more attention or targeted safety measures, thus fulfilling the task of multivariate analysis in a user-friendly, interactive way.

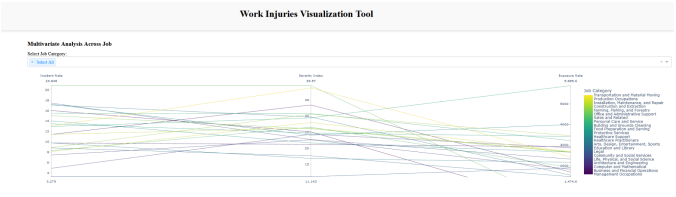

Figure 6: Multivariate Analysis Across Job.

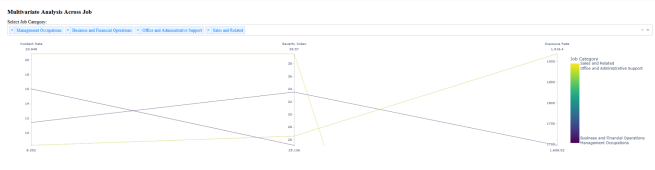

Figure 7: Use case 5: Multivariate Analysis Across Job.

8 Conclusion and Future Work

The goal of this project was to develop a visualization tool that would effectively analyze complex workplace safety data from the US Occupational Safety and Health Administration (OSHA). The tool aimed to empower users, such as safety officials, analysts, and researchers, to explore data, generate hypotheses, and draw meaningful conclusions about workplace safety.

The development process involved exploring various visualization techniques and addressing the specific needs outlined in the task analysis. We explored different visualization options. We started with eight different visualization tasks, but we discovered that certain tasks needed to be removed or merged into each other, simplifying the interface and better aligning with actual user needs.

A few of our assumptions needed adjustment too. For example, we initially underestimated the complexity of parallel coordinates, especially when dealing with many job categories. We also learned that consistent color scales and standardized tooltips significantly reduce user confusion. Additionally, technical constraints in certain plotting libraries forced compromises in how we implemented interactivity (e.g., hovering options or axis reordering).

We have successfully developed a visualization tool that provides a valuable foundation for analyzing workplace safety data. The tool equips users with the necessary tools to gain insight into workplace safety performance, identify areas for improvement, and ultimately contribute to a safer working environment. However, there are areas for improvement. Enhancing outlier visualization through clearer visual cues (e.g., color, size, shape) would make them more readily identifiable. Interactive filtering based on company size, location, or other relevant factors would allow for more targeted analysis. Incorporating contextual information, such as industry-specific safety regulations or hazard information, would better the analysis and provide a deeper understanding of the factors that contribute to the outlier status.

While the developed visualization tool has successfully addressed our primary goals, several opportunities for future work exist. These improvements could further enrich user experience, provide deeper insights, and broaden the tool's ability to visualize trends. Customizable dashboards could be implemented, where users can save and share specific views or visualizations tailored to their needs. Also, the tool could allow users to import their own datasets to compare OSHA data with internal company reports or other data. Predictive analytics is another idea for improvement, where it could be used to forecast future trends in incident rates, severity, and exposure, enabling proactive safety planning. Finally, other, more complex visualizations could be implemented, to enhance certain parts of the data that might be more useful for the user.

9 References

References

[1] Plotly. (n.d.). *Modern Analytic Apps for the Enterprise - Plotly*. Retrieved from `https://plotly.com/`

[2] Bureau of Labor Statistics. (2019, March 29). *List of SOC Occupations*. Retrieved from `https://www.bls.gov/oes/current/oes_stru.htm`

[3] U.S. Census Bureau. (n.d.). *North American Industry Classification System (NAICS)*. Retrieved from `https://www.census.gov/naics/?58967?yearbck=2022`

[4] Occupational Safety and Health Administration (OSHA). (n.d.). *Clarification on how the formula is used by OSHA to calculate incident rates*. Retrieved from `https://www.osha.gov/laws-regs/standardinterpretations/2016-08-23`