Untapped Energy

Datathon For The Oil and Gas Industry
October 12 - 14, 2018

REPORT

Health & Safety

Darrell Aunger
Gabriele Ehnes-Lilly
Sheldon Wall
Lewinda Knowles
Ruth Diaz
Jin Wang
Farhad Qureshi
Jamie Mok

Overview

Energy companies have an obligation to ensure the safety and wellbeing of all its employees and contractors. While the industry has done a good job reducing Lost Time Incidents (LTI) and fatalities, it is not 0.

The objectives of the analysis were twofold:

(i) What is preventing the industry from achieving a zero-lost time incident rate?
(ii) Is it possible to create a personalized safety alert mechanism for every worker?

Through the process laid out in the Cross-Industry Standard for Data Mining model (CRISP-DM) we used the Business Understanding, Data Understanding, and Data Preparation elements to develop a case for change within the safety industry.

Initial observations and insights derived

- It was problematic to obtain raw operator data due to privacy and legality issues. While information such as WCB claims can be obtained it is not complete. It would take months of upfront work to develop relationships to obtain raw data.
- Behavioral aspects (state of mind, state of safety culture) as it relates to incident data are not contained within the WCB data. Individual companies may record this type of information but it is not typically shared through industry. Data shared with government and WCB is “lagging” type information or causal only. Leading type of information is not captured.
- While we looked at costs to the industry for LTI and fatalities, we did not look at how much safety management programs cost the industry to implement and maintain. How can these costs be minimized? (Ian MacGregor feedback)
- The metadata that is being tracked by individual companies is not complete but would provide a good start towards a personalized safety alert app.
- A consistent taxonomy and definitions relating to KPI’s and reportable data do not exist between companies or provinces. This makes comparisons difficult when accumulating data from different reporting jurisdictions.

Actions items and the path forward

- Safety is not a competitive advantage. Sharing anonymized health and safety data amongst industry players would provide immediate benefits and insights that could be provided to both individual workers and their employers. The issue of existing data silos needs to be resolved.
- A business case can be made to develop a personalized safety alert app for industry.

- Examples exist in Europe for health and safety data sharing (European Safety Reliability and Data Association-ESReDA).
- The Society of Petroleum Engineers has an initiative called ‘Getting to Zero and Beyond: The Path Forward’ based upon the International Association of Oil & Gas Producers (IOGP) member company annual safety data. This may be another template and needs to be examined.

Contents

Objective

The objective of the pod was to provide answers to the following:

i) what is preventing the industry from obtaining a zero incident rate?, and
ii) can a personalized safety alert be created with the data that is typically gathered by individual companies?

Technical Summary

External Data Used

- **WCB Claims Data 2011-2016 –** Obtained through Alberta Occupational Health and Safety. Listed 24000 energy sector related WCB claims in Alberta. Contained information on age, type of injury, number of work days missed, total worker hours, and other supplementary information.
- **OHS Stop Work Orders 2011-2016**
- **Metadata and KPI's –** Obtained through Encana and Suncor. These are data streams either being collected or desired to be collected.
- **Various Online Summary Reports** – From Alberta WCB, Alberta OHS, Association of Workers Compensation Boards of Canada (AWCBC).
- **SME Skype Call** – Waqar Mughal generously offered to do a Skype call and presentation with the group providing insights into the complexities of health safety reporting and analytics.

Other Data Obtained But Not Used – UK Offshore Hydrocarbon spill data (1992-2014), NEB Pipeline Spills.

Tools and Technologies

- Power BI
- Excel
- Tableau

Analysis Techniques

- **Power BI** – clean, merge and visualize data, creating dashboards
- **Excel** – some data collected within Excel, preliminary sorting
- **Tableau** – for dashboarding and visualization of data

Key Learnings and Takeaways

The following work was completed in each phase of CRISP-DM.

Business Understanding

Through consultation with Pod Subject Matter Experts (SME's) and an external Safety Industry SME a business model was developed with the following elements: Workers, Environment, Incident, Outcomes and Impact.

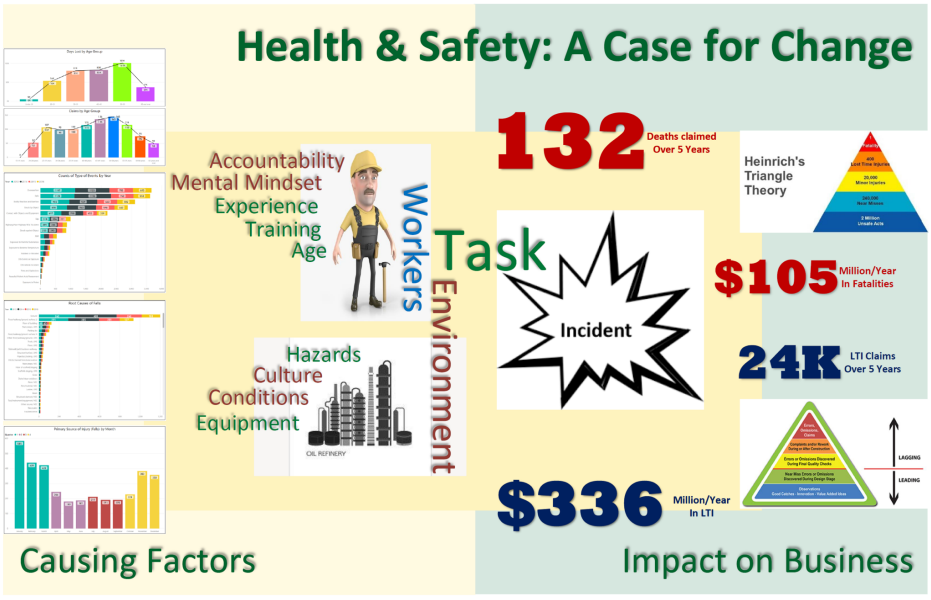

Figure 1 Business model canvas

Further analysis was then performed on these components of the business reference model.

Data Understanding

Through data analysis of the data sets provided, with SME consultation, the following insights around the data were made.

- Falls and overexertion related injuries were the most commonly reported incidents during the 2011-2016 time frame. Types of events are likely a factor that would be a predictor.

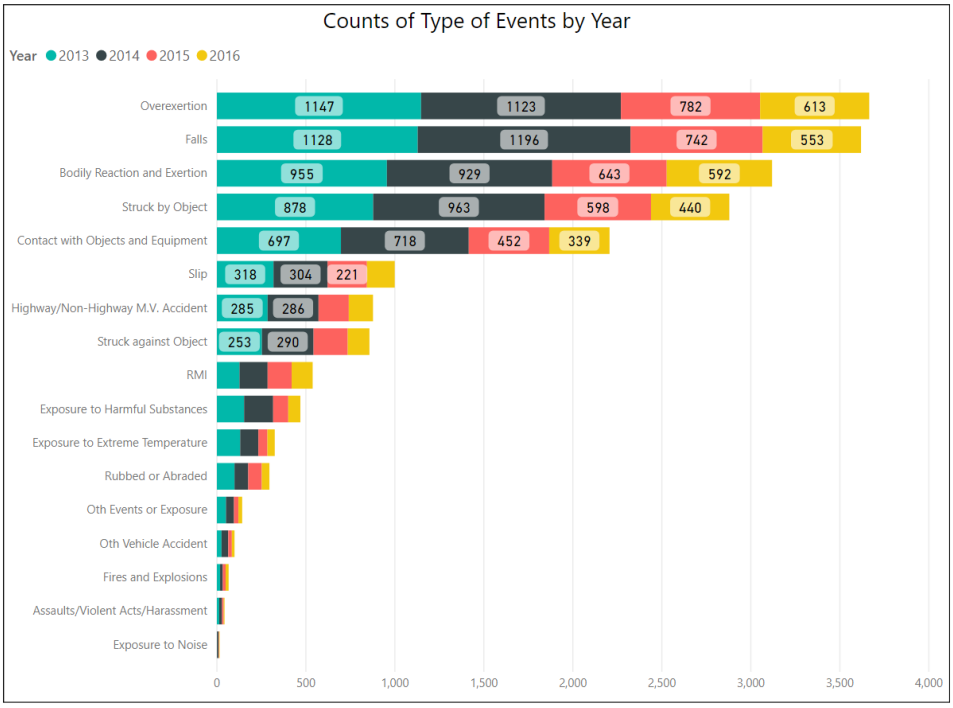

Figure 2 Counts of Type of Events by Year

- The age groups 25-29 and 50-59 had the highest claims rate with the 50-59 age group having the most injured days off per claim. Age appears to be a factor that would be a predictor.

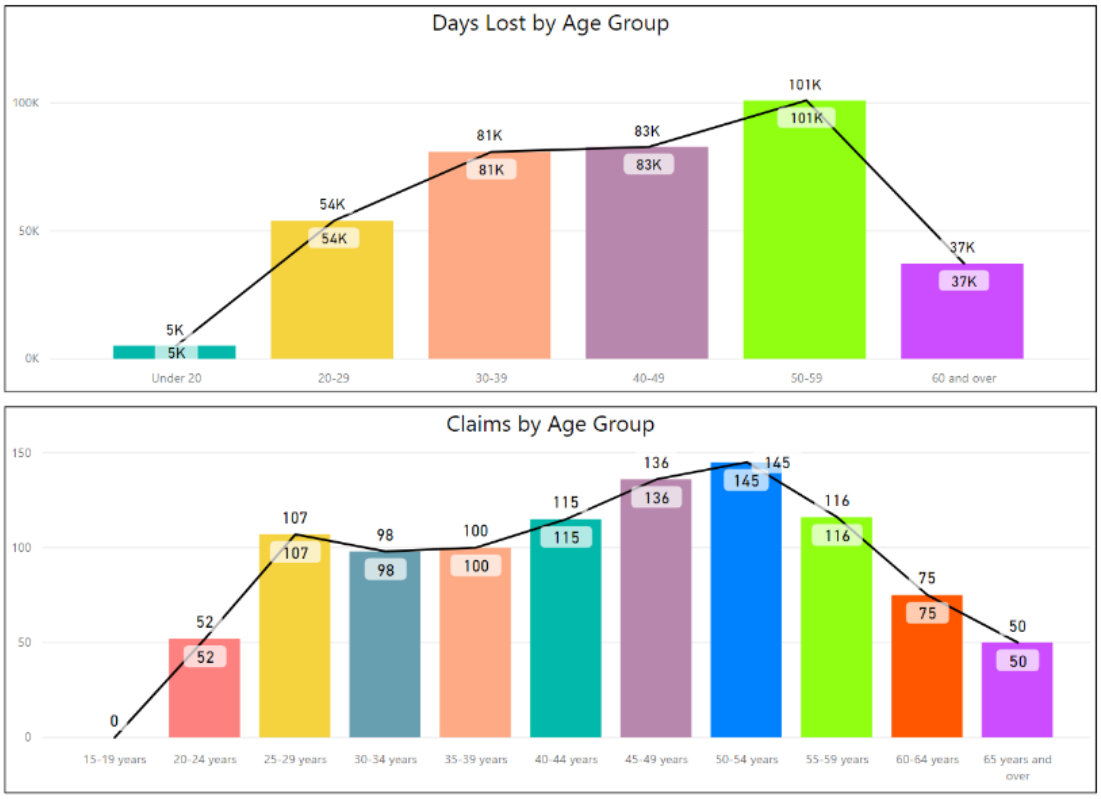

Figure 3 Days Lost and Claims by Age Group

- The initial data suggested that overall the LTI frequency has been dropping year over year. However, when normalized against total hours worked to account for staff reductions during the downturn, in fact the LTR frequency is increasing. It suggests doing more with less may not be the best.

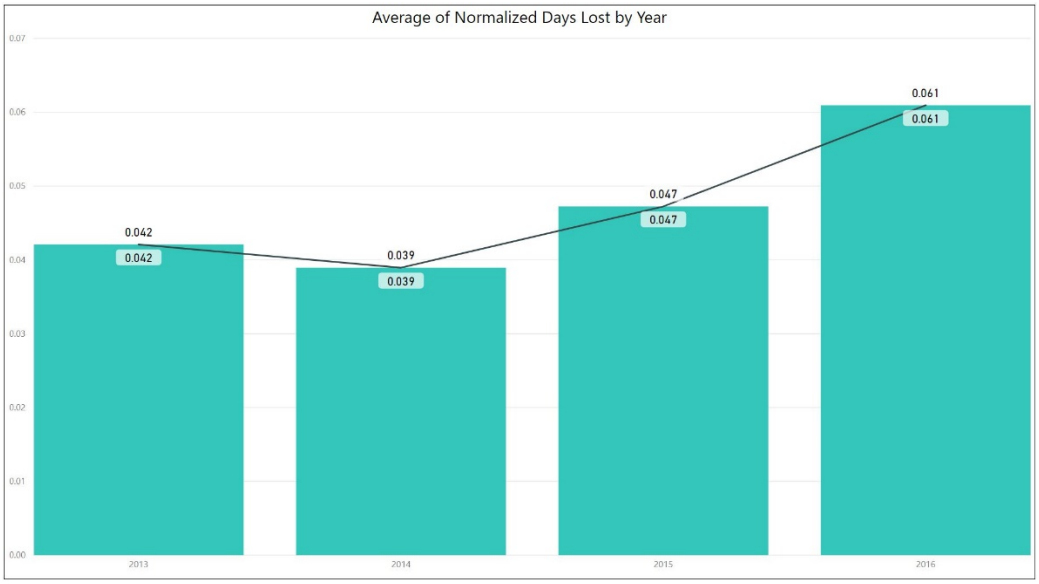

Figure 4 Average Days Lost by Year - Normalized

- The time of year was a factor in the frequency of claims with Q4 and Q1 having the highest number of reported claims. This suggests time of year (and weather) would be a predictor.

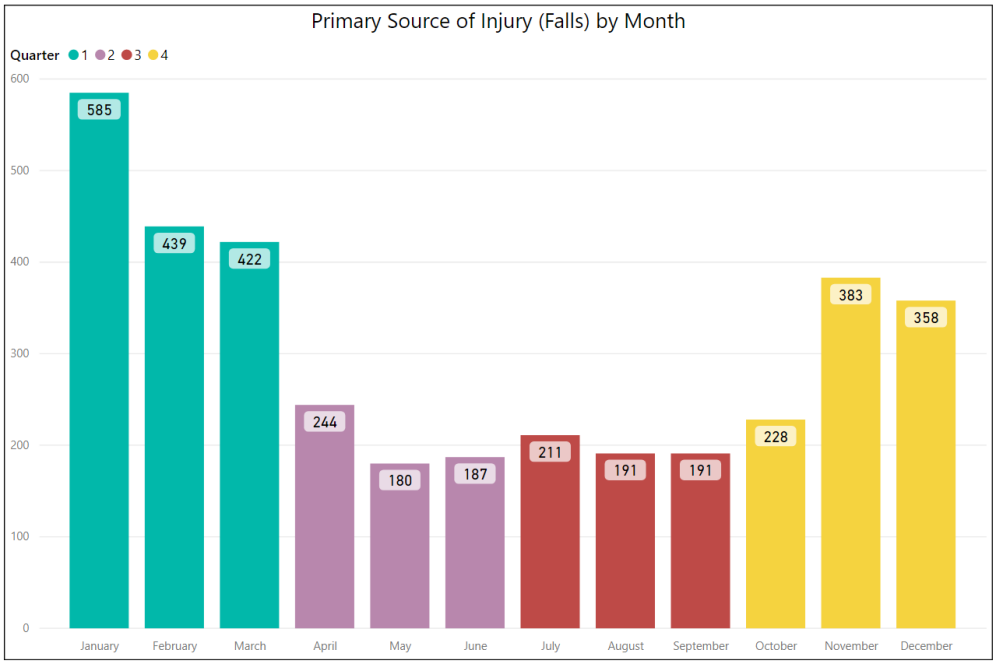

Figure 5 Incidents of Falls - by Month

- Most falls related to level ground, not from at height, with construction trades, truckers and rig workers comprising the bulk of those claims. This suggests that working conditions vary greatly by trades and that age is another factor that changes within each trade.

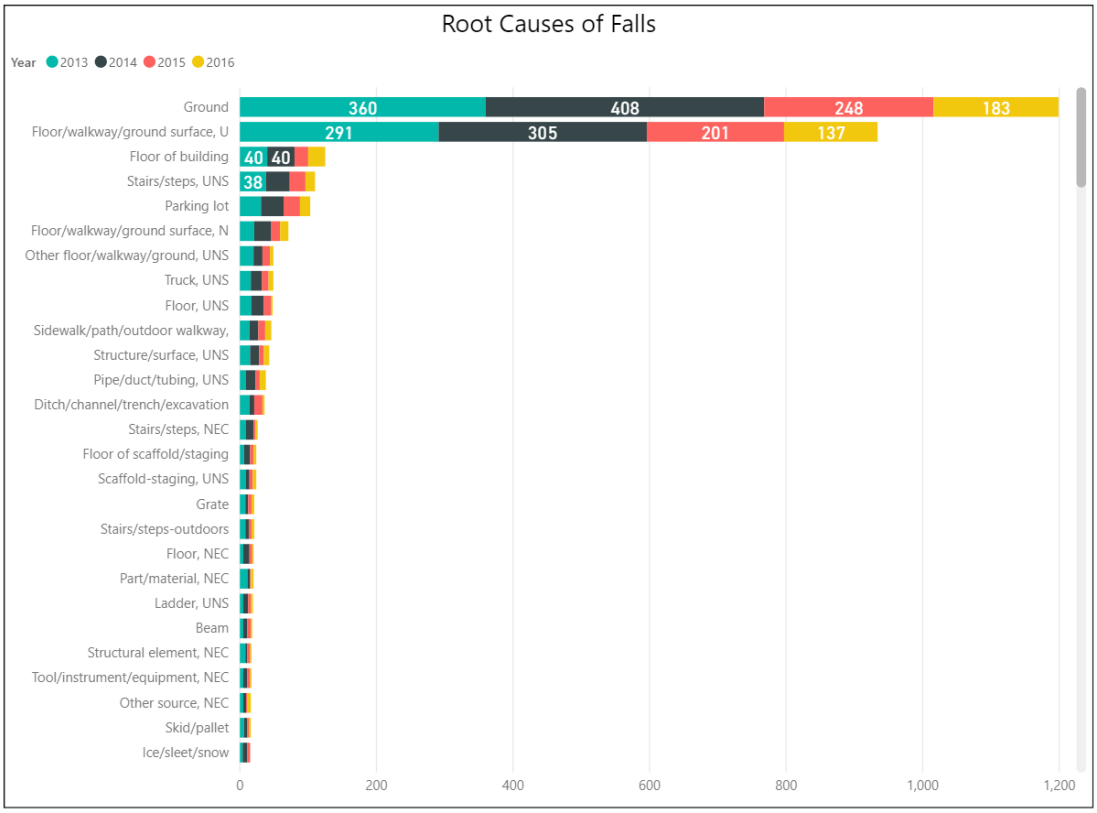

Figure 6 Root Cause of Falls

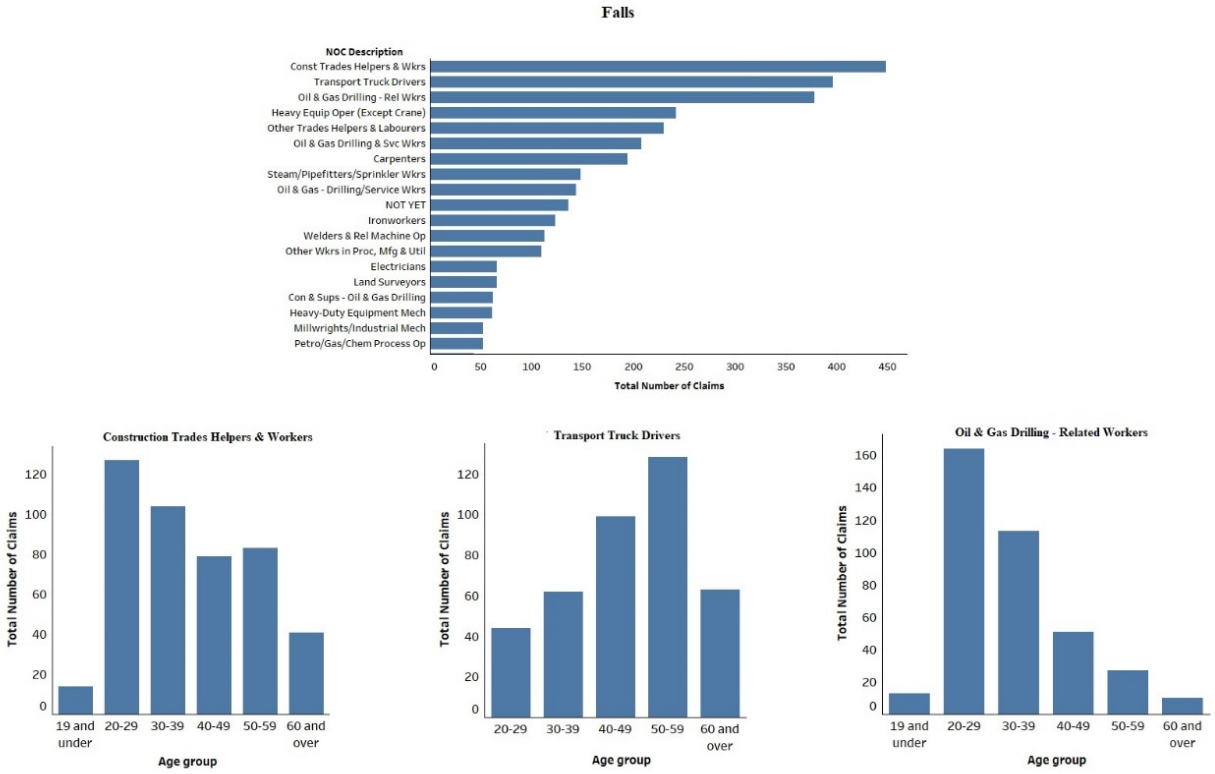

Figure 7 Trades and Ages for Falls

- The same grouping also comprised the largest grouping in terms of overexertion related injuries.

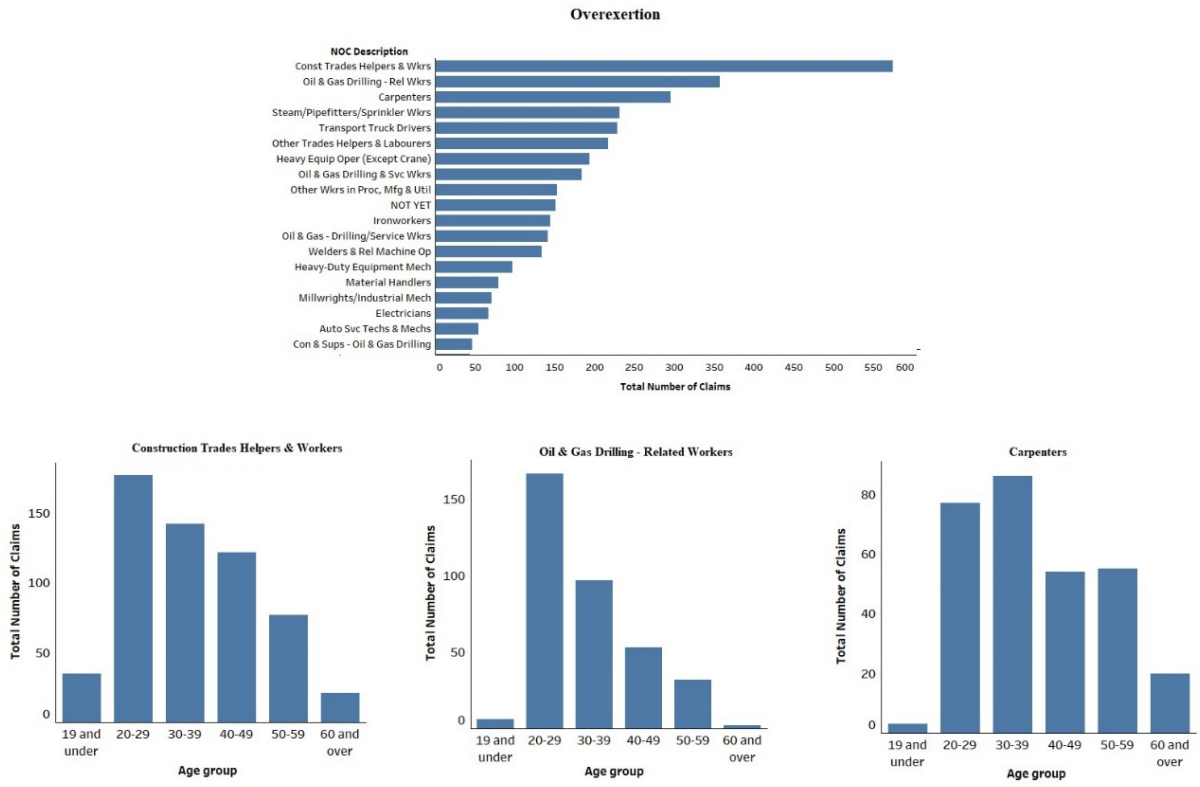

Figure 8 Trades and Ages for Overexertion

It appears from this cursory glance there are likely many features in the data that would make good predictors. Additional work needs to be performed to identify:

i) features that would make good predictors and,
ii) confounding variables that should be removed

Assessing Impact

Over the past 5 years there were 24,000 LTI claims and 132 fatalities. This cost the industry on average $440MM per year.[1] This does not include damages, repairs or lost production associated with the incidents. Based on Heinrich's triangle theory this number is understated as costs associated with minor injuries, near misses and unsafe acts have

[1] 24K claims * 5 years from data set * $10K avg. direct cost per LTI (from data analysis) multiplied by approx. 5-7x. for indirect costs.
132 fatalities from data set – with estimate of $4MM direct and indirect per fatality.
(see https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/Analytics/ca-en-analytics-workplace-safety-analytics.pdf for further indirect cost analysis)

not been analyzed. Individual organizations may collect information on these types of incidents, however, they are not shared externally.

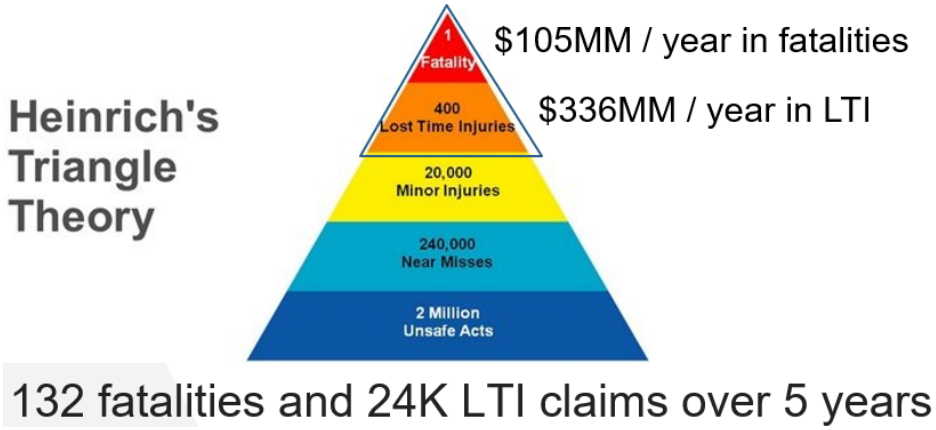

Figure 9 Impact Analysis

Data Preparation

Through the analysis of the meta data provided, with SME consultation, the following insights around the data were made.

H&S Metadata Analysis.xlsx

- A lot of information is being collected by organizations with more mature safety management plans with only a minimal subset reported to government authorities based on policy.
- More data is required to record behavioural aspects of both the worker and the environment in which they work.
- The safety management systems need to be designed with intent such that all required variables are being tracked, links are established with other databases (such as maintenance/repair databases) to allow leading indicators to be used rather than lagging indicators.
- Incidents are captured at a worker level but the 'cure' (policy, procedures, programs) is treated at a group level. Analytics on big data derived from aggregating industry safety data, including new wearable devices, provides a compelling reason to have a more personalized approach rather than a 'peanut butter spread' approach.
- Each reporting jurisdiction has differing reporting standards, definition, and taxonomy. Standardization and consistency needs to be in place to allow large datasets to be integrated and analyzed.

Product Recommendation

The following product recommendation was made:

Personalized H&S Alerts should be created based on further predictive modelling exercises considering:

- Work Activity (Task, Equipment, Complexity, Risk)
- Environment (Culture, Hazards, Conditions)
- Worker (Demographics, Training, Mindset)
- Incident (Date, Description, Testimonies, Root Cause Analysis)

Additional Improvements:

- Additional Data that is tracked (such as state of mind)
- Data Sharing throughout the industry (because more data = higher accuracy)
- Standardization (units, ontology, definitions) – because normalized data requires less interpretation that comes with transformation
- Capture root cause by incident
- Implement and monitor feedback loops for consumption and results of corrective actions
- Evaluate and adjust elements of culture towards a proactive mindset around safety

Journey Map

The conditions were present and following steps taken, during the datathon weekend, that were instrumental to performing this analysis and presentation of results.

- The team consisted of a wide variety of skillsets and knowledge:
  Business SME's: Jamie and Lewinda
  BI SME's: Jin, Farhad
  Research Analyst (epidemiology): Ruth
  Petroleum Engineering: Darrell
  Project Management: Gab
  Expert Data Scientist (& Chief Bottle Washer and Ragamuffin): Sheldon
- The team discussed and agreed upon the objectives they were trying to fulfill and then postulated questions that needed to be answered to achieve those objectives.

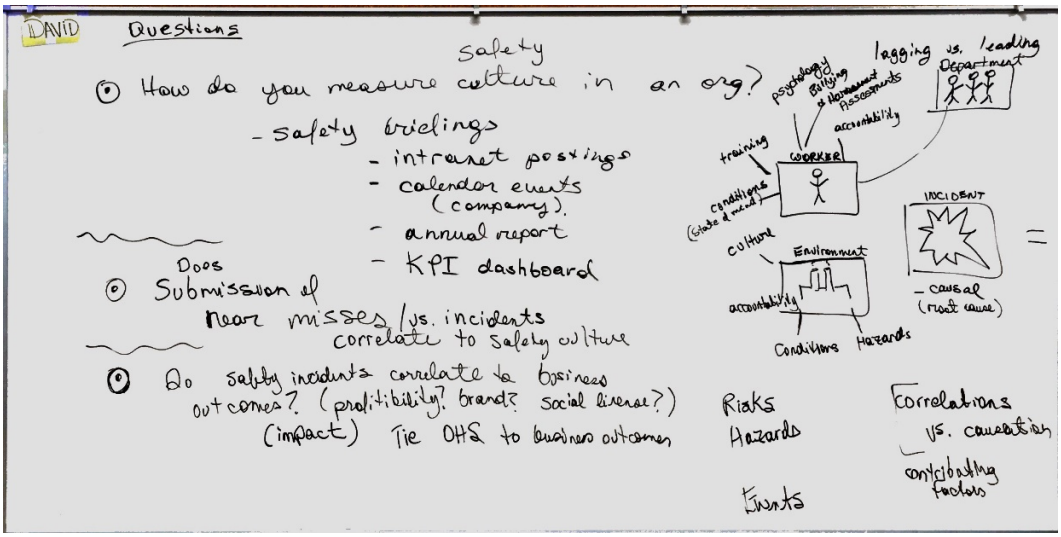

Figure 10 objective breakdown

- The team quickly established a process they were going to adopt (CRISP-DM) and which steps of that process that they were undertaking.
- A business SME presented a business model that could be used. This was adopted and then later refined after conversations with an external SME.

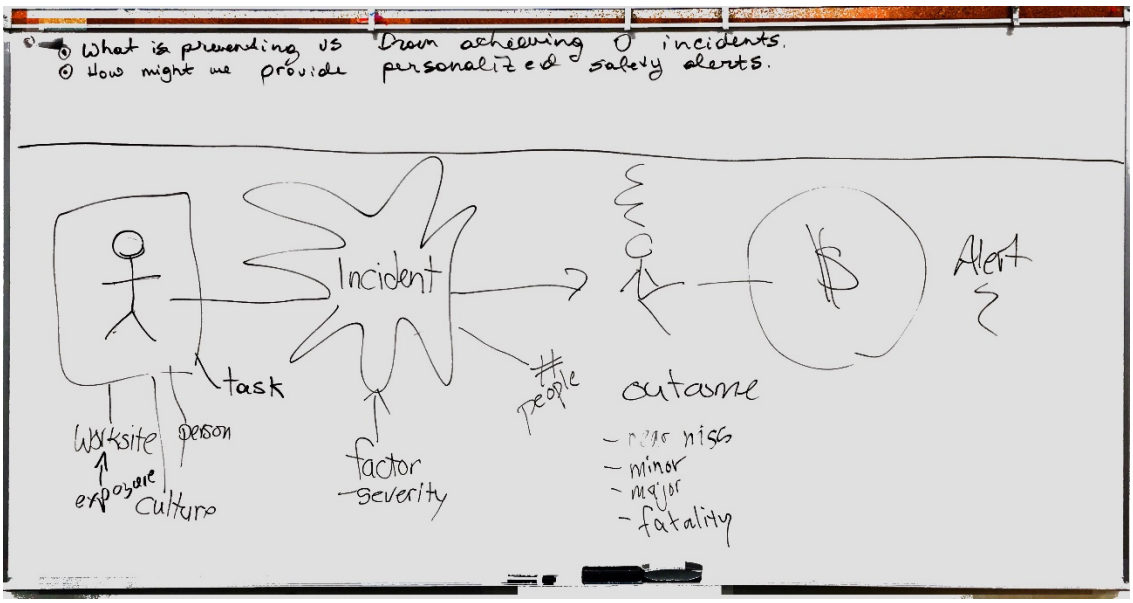

Figure 11 Initial Business Model

- An external SME was engaged to gather expert insights in the objectives the team was attempting to achieve.
- An iterative cycle was taken for:
    - refining the objectives,
    - performing analysis to understand the data and
    - reviewing the results
    - developing the “story” and preparing the results for sharing and discussion

Table of Figures