

# Executive Summary

*Jing Heng Lim*

*September 10, 2019*

## 1. Problem and context

Over the past decade, there were a significant number of workers were injured or killed while working. Hence, it is no stranger to a global industrial manufacturing company that holds over 10,000 worker across the world. Unfortunately, due to a recent accident in South America had placed company into scrutiny. The chief executive officer of the company needs a solution to this growing crisis from our company. Hence, a model will be build to prevent injuries occur.

Objective:

1. What is the ideal safety regime for the company basedon injury prevention performance?
2. Is industry experience more important than the safety regime when it comes to prevent injuries?

## 2. Data collected

The data was collected with four variables with injuries is the outcome variable (response) and hours (working hours) as a count variables, experience had been categorized into 4 levels, it is coded as 1 =“least experienced”, 2=“less experienced”, 3 = “experienced”, 4 = " very experienced" based on overall workers experience and safety had been categorized into certification, certification + 12monthly review, no standardised process and on-site induction.

## 3. Justify choosen method

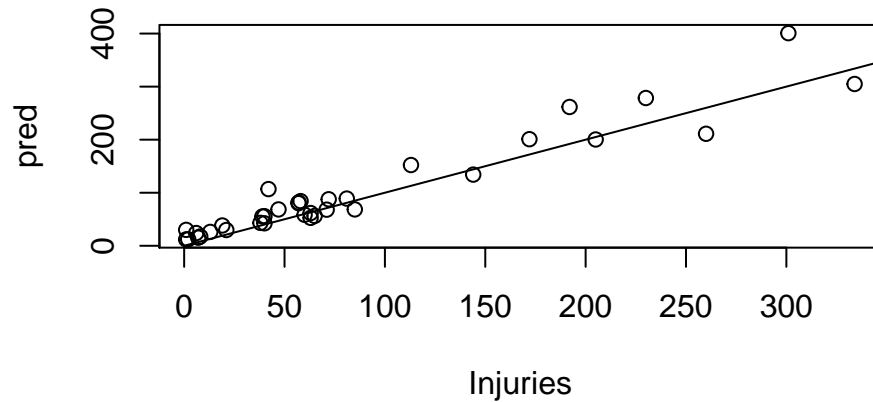
Several method had been tested in this analysis such as Poisson regression model, Quasi-Poisson model and Negative-Binomial model as they are a reasonable approach for our count data. These model would be able to identify the significant effects of each variables, optimization to reduce injuries occur and it enable to predict future unforeseen number of injuries. Through the analysis, we found out the data was overdispersed. However after couple of model fitting Negative-Binomail model (NB\_model) approach is the most ideal model compared to others.

Our final NB\_model is  $\text{Injuries} \sim \text{Experience} + \text{Safety} + \text{offset}(\log(\text{Hours}))$ . The summary of the model is that when there is an increase of 1 worker in experience 2 will reduce approx 0.48 times of injuries, experience 3 will reduce approx 1.12 times and experience 4 will reduce approx 1.977 times respectively. For safety procedure, a increase of 1 worker that had safe certification with 12 month review would reduce 0.02 times of injuries. However, no standardised process and on-site induction will increase the number of injuries by 0.24 and 0.36 respectively.

This model holds an approximately 0.13% of variability on coefficients and the confidence intervals for covariate coefficients are:

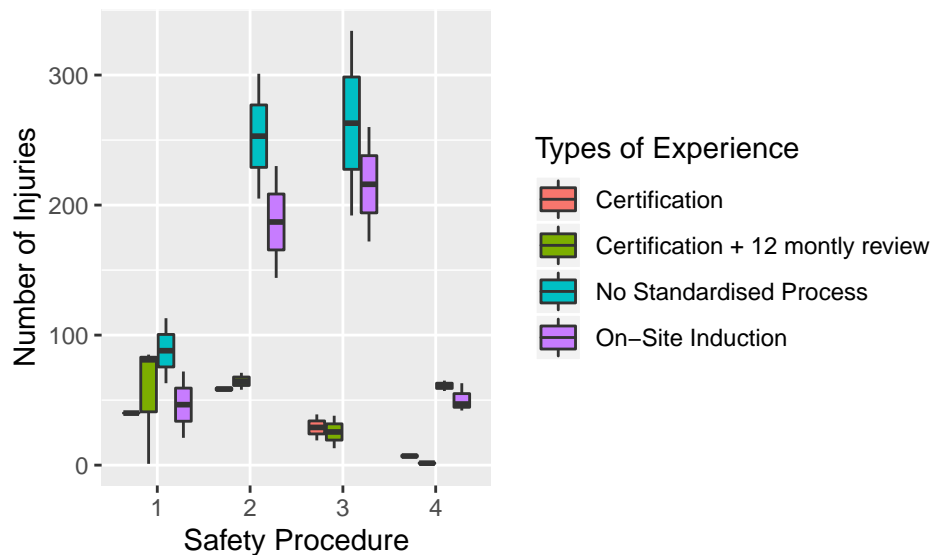
##	2.5 %	97.5 %
## (Intercept)	-8.220809299	-7.7565725
## Experience2	-0.708429159	-0.2605011
## Experience3	-1.351060415	-0.8867063
## Experience4	-2.238358422	-1.7171381
## SafetyCertification + 12 montly review	-0.289142621	0.2405599
## SafetyNo Standardised Process	0.002152297	0.4877352
## SafetyOn-Site Induction	0.120522523	0.6054524

#### 4. Validity and fit of the model



To validate our negative binomial GLM, a leave one out cross validation has been implemented. The model validity and fit can be easily justified by the plot. The plot above has been plotted against the actual and the predicted number of injuries based on our model. As we can observe that the predicted values are pretty close to the actual value and it shows a linearity between them. Therefore, the model fits perfectly here and it holds a pretty accurate predictability on unforeseen number of injuries that would occur.

#### 5. Conclusion



In conclusion, a certification or a certification with 12 monthly review is recommended as an international standard safety regime for the company. The boxplot shows that safety regime is more important than experiences when it comes to preventing injuries as it tends to have a lower count of injuries throughout 4 levels of experience of workers. Furthermore, workers tend to injure themselves when they have no standardised process and on-site induction safety regime.