Workers’ Compensation

Molly Hischke & Kyle Hancock

3/13/2020

# Introduction (4 pts - Kyle)

The brewing industry has seen substantial growth in Colorado over the past several years, going from 126 craft breweries in 2011 to 425 in 2019. In 2019, the economic impact of the craft brewing industry was over $3 billion, ranking Colorado the highest per capita in the United States (Brewers Association, 2019). The growth of the brewing industry and its economic importance has prompted the recent creation of a regional emphasis program by Colorado OSHA in 2019. This program was created to help employers identify hazards, ensure OSHA compliance, and reduce risk through the implementation of controls (OSHA, 2019). In the regional emphasis program, confined space, falls, noise, repetitive motions, lifting, and chemical exposures are identified as some of the hazards in the industry. Specifically, craft breweries tend to be small operations that rely heavily on manual material handling which has been associated with low back injuries (Lavendar, 2012). Researchers in the occupational ergonomics and safety program at Colorado State University have also taken interest in the quickly evolving brewing industry. They want to understand how and why workers are injured in brewing and actions that can be taken to reduce risk. Workers compensation data has been obtained from one of the largest providers in the state of Colorado to look at these research questions. This type of data is often used in the field of occupational health and safety to look at entire industries or regions. Very little has been published in regards to occupational health and safety for brewing and greater understanding could allow for better safety and production outcomes for the industry as a whole.

For this project, the dataset was workers’ compensation data for all craft breweries in Colorado who held policy with a specific Colorado based insurance company between 2013 and 2018. There was a total of 975 observations encompassing a total of 130 breweries. Information on adjusted payroll (proxy for size of brewery), if a workers compensation claim was made, and year was available for each brewery. Researchers were interested in the relationship between brewery size and workers compensation claims. Is there an association between the size of the brewery and if a claim is filed (binary response) or the number of claims that are filed (continuous response)? The predictor variable was adjusted payroll (continuous) and the response variables were if a claim was filed (binary) or the number of claims (continuous). This question was of interested because of the current profile of brewery size. Most tend to be small operations, producing just a few thousand barrels each year, but some are extrememly large, producing hundreds of thousands of barrels. There are few breweries in the middle. As breweries grow, many things change: the number of employees, resources, automation, etc. These changes create a new environment and change the hazards associated with the job. Understanding the relationship between size and injuries could help to identify direction for future research in the brewing industry.

* Identify observational/experimental units and number of observations
* Further details of design (make sure to include how we processed the data here)

# Summary statistics and/or graphics (4 pts - Kyle)

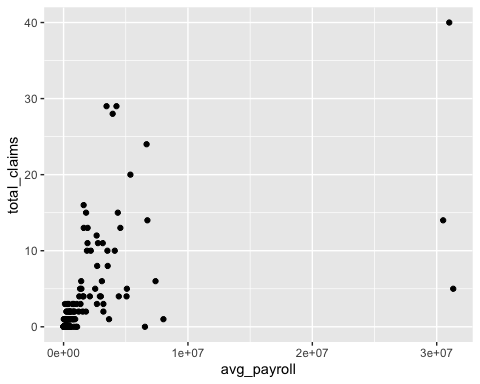
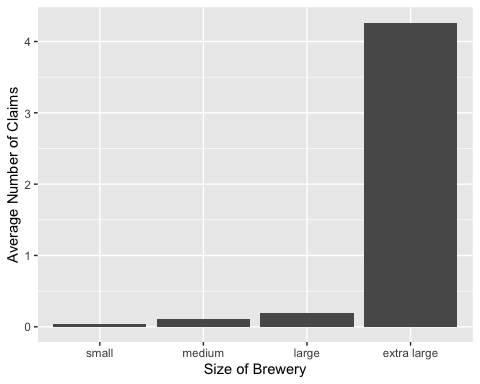
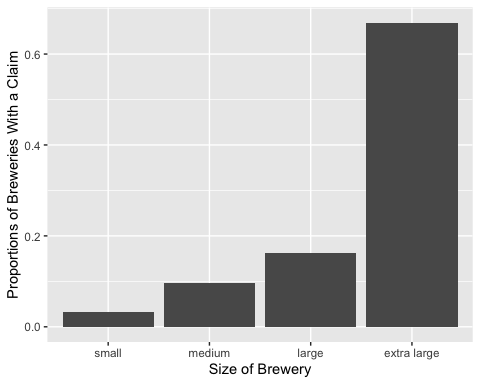
## brewery year avg\_payroll total\_claims   
## 3373385: 6 2013: 37 Min. : 6216 Min. : 0.000   
## 3505469: 6 2014: 51 1st Qu.: 71370 1st Qu.: 0.000   
## 3508453: 6 2015: 78 Median : 143400 Median : 0.000   
## 3530154: 6 2016: 93 Mean : 661117 Mean : 1.154   
## 3537607: 6 2017:107 3rd Qu.: 307722 3rd Qu.: 0.000   
## 3561828: 6 2018:129 Max. :31327727 Max. :40.000   
## (Other):459   
## claim\_filed   
## Min. :0.0000   
## 1st Qu.:0.0000   
## Median :0.0000   
## Mean :0.2404   
## 3rd Qu.:0.0000   
## Max. :1.0000   
##

Sample of the brewery worker’s compensation claim data.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Brewery | Year | Average Payroll | Total Claims | Claim Filed? |
| 3627117 | 2014 | 296366.57 | 1 | 1 |
| 3622411 | 2017 | 262451.16 | 0 | 0 |
| 3609789 | 2016 | 235805.48 | 0 | 0 |
| 3530154 | 2018 | 100766.63 | 0 | 0 |
| 3607337 | 2017 | 224039.52 | 0 | 0 |
| 3607183 | 2014 | 91146.30 | 0 | 0 |
| 3586070 | 2013 | 146991.93 | 0 | 0 |
| 3611462 | 2017 | 73162.88 | 0 | 0 |
| 3609169 | 2018 | 39484.75 | 0 | 0 |
| 3613425 | 2018 | 97220.06 | 0 | 0 |

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 6216 71370 143400 661117 307722 31327727

## # A tibble: 4 x 3  
## size n\_breweries prop\_claim  
## <fct> <int> <dbl>  
## 1 small 124 0.0323  
## 2 medium 124 0.0968  
## 3 large 123 0.163   
## 4 extra large 124 0.669



# Analysis (6 pts - Molly)

Two analyses were conducted (both using Generalized Estimating Equations) with the response being binary and continuous (count data).

## Binary Response with Repeated Measures (Model1)

Model1 analysis was done using R and the geepack package (Hojsgaard, Halekoh, & Yan 2006; Yan & Fine, 2004; Yan, 2002). Generalized Estimating Equations (GEE) was used to fit a model with the response being if a brewery filed a claim (1 = yes, 0 = no). Fixed effects include payroll (proxy for brewery size). Breweries were used as clusters and we assumed a exchangeable correlation structure.

## Count Response with Repeated Measures (Model2)

Model2 analysis was done using R and the geepack package (Hojsgaard, Halekoh, & Yan 2006; Yan & Fine, 2004; Yan, 2002). Generalized Estimating Equations (GEE) was used to fit a model with the response being the number of times a brewery filed a claim within a year. Fixed effects include payroll (proxy for brewery size). Breweries were used as clusters and we assumed a exchangeable correlation structure.

# Results and Conclusions (4 pts - Molly)

The average salary of a brewery worker in Colorado is $35,265 (ZipRecruiter, 2020) - therefore this was used to help interpret our analysis.

LIST OF THINGS TO INCLUDE:

* ANOVA style likelihood ratio tests and/or table of estimated coefficients (UNSURE IF WE CAN COMPUTE THESE)
* Other results as appropriate (ex: pairwise comparisons for categorical predictor) (N/A)
* Interpretation and discussion (even if nothing is “significant”)
* Respond to your research questions

## Binary Response with Repeated Measures (Model1)

Based on Model1 (p = 0.039), we find evidence of a difference between the odds (*KYLE CHECK*). A $35,265 increase (hiring approximately one worker) in average yearly payroll is associated with a multiplicative increase of 1.12 (12%) in the estimated odds of having (at least one) workers’ compensation claim filed. The results we found were anticipated in that we assumed that as a craft brewery hires more staff and increases in size that there would also be an increase the odds of filing a workers’ compensation claim (i.e. work-related injuries). Although part of the increase could be because smaller breweries may be less likely to file a workers’ compensation claim.

*Model1 Coefficients Table:*

## Estimate Std.err Wald Pr(>|W|)  
## (Intercept) -2.369949e+00 4.102487e-01 33.37211 7.610769e-09  
## avg\_payroll 3.210099e-06 1.554756e-06 4.26297 3.895175e-02

*Model1 Odds Ratio:*

## [1] "1.119856609635867"

## Count Response with Repeated Measures (Model2)

Based on Model2 (p = 0.0157), we find evidence of a difference between the odds (*KYLE CHECK*). A $35,265 increase (hiring approximately one worker) in average yearly payroll is associated with a *multiplicative* increase of 1.004 (0.4%) in the predicted number of workers’ compensation claims filed.

However, these results may underestimate the increase in the number of workers’ compensation claims filed with an increase in brewery size. The National Brewers Association found that 17.8% of brewery employees who experienced an injury at work did not report it. Furthermore, 20% of those who reported an injury stated their injury was not documented by the organization (Embry & Stinchfield, 2020). Especially for smaller breweries with less employees, the organizations might not even know they need to report serious injuries (Pell, 2013).

*Model2 Coefficients Table:*

## Estimate Std.err Wald Pr(>|W|)  
## (Intercept) -6.479684e-01 3.693048e-01 3.078490 0.07933415  
## avg\_payroll 1.240874e-07 5.137409e-08 5.834006 0.01571929

*Model2 Odds Ratio:*

## [1] "1.00439"

One of the limitations in this dataset is that there are only 130 breweries who had insurance policies with this particular insurance company. In 2013 (the start of the dataset), there were a total of 175 breweries in Colorado. In 2018 (the end of the dataset), there were a total of 396 breweries in Colorado (Brewers Association, 2019). We do not expect this to change the overall results (that larger breweries have and have more workers’ compensation claims).

Overall, our results aligned with what we had predicted - larger breweries file or file more workers’ compensation claims than smaller breweries.

# References

1. Højsgaard, S., Halekoh, U. & Yan J. (2006) The R Package geepack for Generalized Estimating Equations Journal of Statistical Software, 15, 2, pp1–11
2. Yan, J. & Fine, J.P. (2004) Estimating Equations for Association Structures Statistics in Medicine, 23, pp859–880.
3. Yan, J (2002) geepack: Yet Another Package for Generalized Estimating Equations R-News, 2/3, pp12-14.
4. ZipRecruiter (2020). Brewery salary in Colorado. Retrieved from <https://www.ziprecruiter.com/Salaries/How-Much-Does-a-Brewer-Make-a-Year--in-Colorado>
5. Embry, E., & Stinchfield, M. (2020). Praise and paradox: What we learned from the brewers association safety and injury survey. Retrieved from <https://www.brewersassociation.org/seminars/praise-and-paradox-what-we-learned-from-the-brewers-association-safety-and-injury-survey/>
6. Pell, M.B. (2013). Insight: Fast-growing U.S. craft brewers struggle with worker safety. Retrieved from <https://www.reuters.com/article/us-brewing-safety-idUSBRE96B0MW20130712>
7. Brewers Association. (2019). State craft beer sales & production statistics, 2019. Retrieved from <https://www.brewersassociation.org/statistics-and-data/state-craft-beer-stats/>
8. OSHA. (2019). Regional Emphasis Program for Beverage Manufacturing. Retrieved from <https://www.osha.gov/sites/default/files/enforcement/directives/CPL_20-11_04-01.pdf>
9. Lavender, S. A., Marras, W. S., Ferguson, S. A., Splittstoesser, R. E., & Yang, G. (2012). Developing physical exposure-based back injury risk models applicable to manual handling jobs in distribution centers. Journal of occupational and environmental hygiene, 9(7), 450-459.

# R code appendix

# Retain (and do not edit) this code chunk!!!  
library(knitr)  
knitr::opts\_chunk$set(echo = FALSE)  
knitr::opts\_chunk$set(message = FALSE)  
library(readr)  
library(tidyr)  
library(dplyr)  
library(ggplot2)  
library(car)  
library(emmeans)  
library(knitr)  
library(geepack)  
  
## Reading in raw data  
wc <- read\_csv("../data/craftbeer\_wc.csv")  
  
  
## Cleaning data, wc2 = final cleaned data frame to use for analysis  
wc2 <- wc %>%   
 group\_by(Policy\_Dim\_Key, year) %>%   
 summarize(avg\_payroll = mean(adj\_payroll),  
 total\_claims = sum(Claim\_Count)) %>%   
 mutate(claim\_filed = ifelse(total\_claims == 0, 0, 1)) %>%   
 ungroup() %>%   
 mutate(Policy\_Dim\_Key = as.factor(Policy\_Dim\_Key),  
 year = as.factor(year)) %>%   
 rename(brewery = Policy\_Dim\_Key)  
summary(wc2)  
  
## Adds a table to present in the report.   
data\_table <- wc2 %>%   
 rename(Brewery = brewery,  
 Year = year,  
 'Average Payroll' = avg\_payroll,  
 'Total Claims' = total\_claims,  
 'Claim Filed?' = claim\_filed)  
  
kable(sample\_n(data\_table, 10, replace = TRUE),   
 caption = "Sample of the brewery worker's compensation claim data.",  
 align = "ccccc")  
  
  
#payroll stats  
payroll\_stats <- summary(wc2$avg\_payroll)  
payroll\_stats  
  
#size categories based on quartile  
size\_categories <- wc2 %>%   
 mutate(size = cut(avg\_payroll, breaks=c(-Inf, 71370, 143400, 307722, Inf), labels=c("small","medium","large","extra large")))  
  
SumStats <- size\_categories %>%   
 group\_by(size) %>%   
 summarize(n\_breweries = n(),  
 prop\_claim = sum(claim\_filed/n\_breweries))  
SumStats  
  
SumStats2 <- size\_categories %>%   
 group\_by(size) %>%   
 summarize(n\_breweries = n(),  
 avg\_claims = mean(total\_claims))  
  
ggplot() +  
 geom\_col(data = SumStats, aes(size, y = prop\_claim)) +  
 labs(x = "Size of Brewery", y = "Proportions of Breweries With a Claim")  
   
ggplot() +  
 geom\_col(data = SumStats2, aes(size, y = avg\_claims)) +  
 labs(x = "Size of Brewery", y = "Average Number of Claims")  
  
ggplot() +  
 geom\_point(data = wc2, aes(x = avg\_payroll, y = total\_claims))  
  
SumStats3 <- wc2 %>%   
 group\_by(year) %>%  
 summarise(n = n(),  
 Prop\_Claim = sum(claim\_filed)/n)  
  
SumStats4 <- wc2 %>%   
 group\_by(year) %>%  
 summarise(n = n(),  
 avg\_claims = mean(total\_claims))  
  
# Binary Response with Repeated Measures  
Model1 <- geeglm(claim\_filed ~ avg\_payroll, id = brewery,  
 family = binomial(link = "logit"),   
 corstr = "exchangeable", data = wc2)  
# Count Response with Repeated Measures  
Model2 <- geeglm(total\_claims ~ avg\_payroll, id = brewery,  
 family = poisson(link = "log"),   
 corstr = "exchangeable", data = wc2)  
# Summary Table of Model1  
summary(Model1) %>%   
 coefficients()  
  
# Model1 Odds Ratio  
format(exp(35265 \* 3.21e-06), digits = 16)  
  
# Summary Table Model2  
summary(Model2) %>%   
 coefficients()  
# Model2 Odds Ratio  
format(exp(35265\* 1.240874e-07), digits = 6)  
  
# References   
citation("geepack")