A Review of Job Satisfaction in West Carts

This study was conceived with the intentions of finding evidence into what causes Team Member (TM) job satisfaction. It asked 56 TM's to fill out a card and answer questions about their view on other jobs, their personal job enjoyment, and the quality of their training. The results of this study are broken down into four parts, observing the relationship between Current Job and the TM's first choice of Cross Train (when a TM trains into another job with the intention of picking up shifts and not permanently changing duties), Current Job and Job Enjoyment, Current Job and Training Effectiveness, and Job Enjoyment and Training Effectiveness. The results of these studies have the capabilities to inform leadership of which jobs have the weakest trainers, which jobs have the lowest job satisfaction, which jobs are viewed most positively by TM's, and the extent of the connection between Quality of Training and Job Enjoyment. This information can, in turn, motivate decisions on morale booting activities, job duty re-assignment, and trainer changes.

This paper is broken up into four sections. The first is about the study design. Which questions were asked, to whom were they asked, and when did the study take place will all be answered in section 1. Section 2 is devoted to technical findings, mainly tests of independence and some pretty graphs. Section 3 is a discussion on the limitations of the study, and ways to allow future studies to reach more powerful statistical conclusions. Finally, Section 4 provides some concluding remarks, taking the statistical analysis and applying it back to West Carts. It identifies problems for leadership to address that can be identified using the data, and some solutions that are in line with the data as well. The two appendices include the R-output from the study and an address to some of the comments received from TM's while conducting the study.

Study Design

61 TM's in West Carts/Pizza Pred were surveyed over the course of March 25th 2017 – March 27th 2017. Most TM's were asked in the beginning or the end of the day, to avoid interfering with peak hour operation. Team Captains were asked to administer the survey to TM's as they showed up, cashed out, or went to break. Cooks, bussers, and stockers were surveyed through sheer force of will, they running all over the venue during administrators down time to get responses. The Survey asked four questions:

1. Select Your Current Job Title:
    a. Cashier
    b. Busser
    c. Stocker
    d. Line Cook
2. If you could be Cross Trained, which job would you like to cross train into?
    a. Cashier
    b. Busser
    c. Stocker
    d. Line Cook
3. Do you enjoy working here at West Carts / Pizza Pred?

   a. Yes
   b. No
4. Do you feel you were effectively trained for your current Job?
   a. Yes
   b. No

61 survey cards were returned. 5 were thrown away due to improper responses. 9 responses to question two were thrown away do to the poor formatting of the question. Many TM’s selected multiple responses, the question should have specified “Select One”. Only one response had to be thrown out for question 4, the respondent chose to be arrogant instead of useful. In order to keep the sample size within working limits, the survey was limited to entry level TM’s, and excluded leadership positions such as Team Captains, Management, Second Cooks, and First Cooks. As Job Title was the dependent variable in three of the four datasets, any responses marked with “Other” for job title were discarded.

Findings

Job – Cross Train

Job / Cross Train	Cashier	Stocker	Busser	Line Cook	Total
Cashier	0	8	4	11	23
Stocker	1	0	1	5	7
Busser	1	1	0	3	5
Line Cook	3	9	0	0	12
Total	5	18	5	19	**47**

Table 2.1

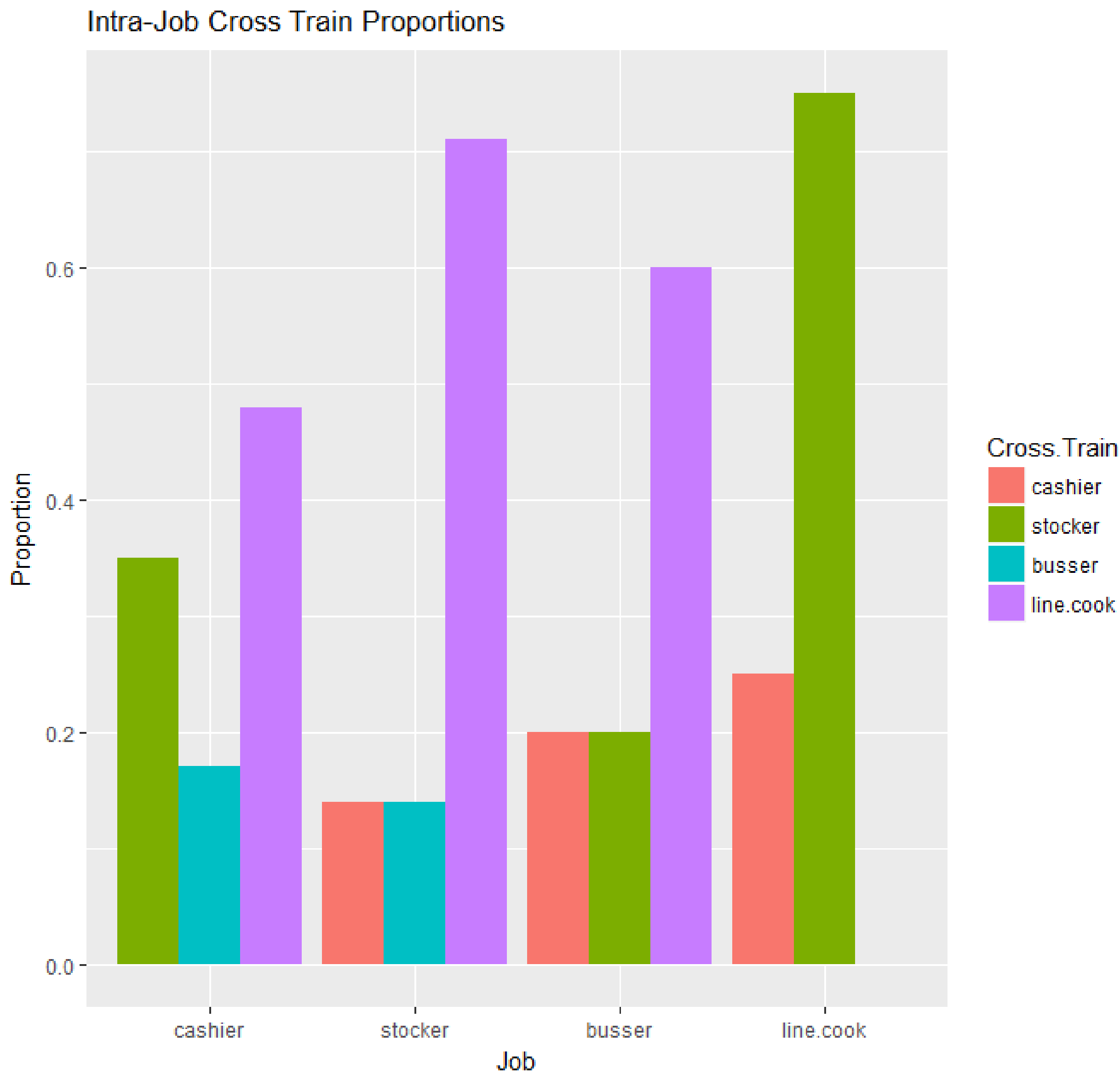
Table 1.1 displays the Raw Data for the Job – Cross Train data set. Observing the data without any statistical processes yields some very interesting insights. Most people view the jobs of Line Cook or Stocker to be their preferred cross train. Breaking this table down into row probabilities gives up table 2.2.

Job / Cross Train	Cashier	Stocker	Busser	Line Cook	Total
Cashier	0%	35%	17%	48%	1
Stocker	14%	0%	14%	71%	1
Busser	20%	20%	0%	60%	1
Line Cook	25%	75%	0%	0%	1

Table 2.2

Table 2.2 tells us that 71% of stockers surveyed would prefer to cross train into a Line Cook job over the jobs of Cashier or Busser. The same can be said for both Cashiers and Bussers. Line Cook was universally (Pun intended) preferred over any other job, Stocker came in a close second due to the influence of Line Cooks, who, interestingly enough, have no interest in Bussing.

Performing a χ^2 Test of Independence gives us a test statistic of 22.869 on 9 degrees of freedom, yielding a p-value of .006498, and allowing us to handily reject the null hypothesis of independence. But this conclusion is largely meaningless. Choice of Cross Train is dependent upon Current Job mainly because you can't cross train into your current job. That row of zeros down the center of the table is the primary source of magnitude in our test statistic. Removing these data points gives us a new χ^2 statistic of 5.99 on 9 degrees of freedom, significantly less and well outside our rejection region. This leads us to conclude that Cross Train does not depend upon Current Job as every job would prefer to be a cook, besides people who are already cooks, and can't cross train into it by nature.



Graph 2.1

Job – Enjoyment

Job / Enjoy	Yes	No	Total
Cashier	19	6	25
Stocker	9	1	10
Busser	3	3	6
Line Cook	11	4	15
Total	42	14	**56**

Table 2.3

Table 2.3 details responses to question 3, broken down by Job Title. I am happy to see most people are enjoy working at West Carts/Pizza Pred. Stockers are the happiest, with 90% of stocker respondents enjoying work, but Bussers were the least happy, only 50% of respondents saying they enjoyed working at Pred. The latter result is intimated at in the previous dataset, Busser was one of the bottom jobs for people to cross train into.

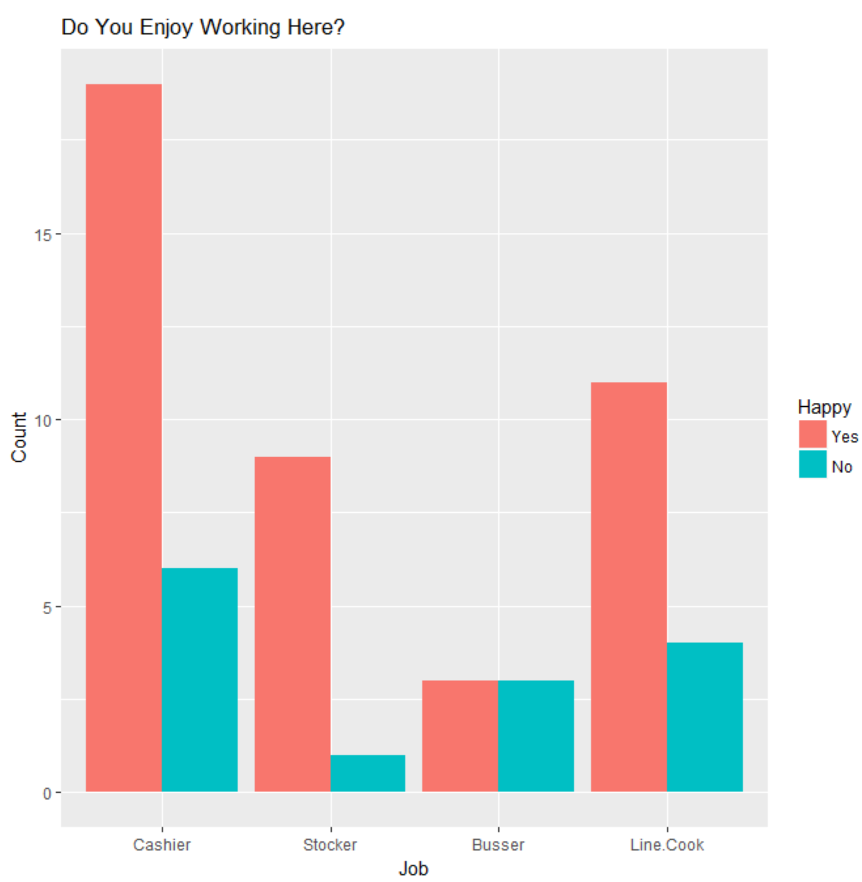
Running a χ^2 Test of Independence on this dataset tells us that there is no evidence of dependence between these two factors, allowing us to reduce the chart to table 2.4.

	Count	Proportion
Enjoy	42	75%
Do Not Enjoy	14	25%

Table 2.4

Table 2.4 does not give us much information, but does supply a nice fuzzy feeling with knowing that 75% of our TM's enjoy working here.

Graph 2.2 provides a nice comparative visual, to further illustrate the "Busser Problem".



Job – Training

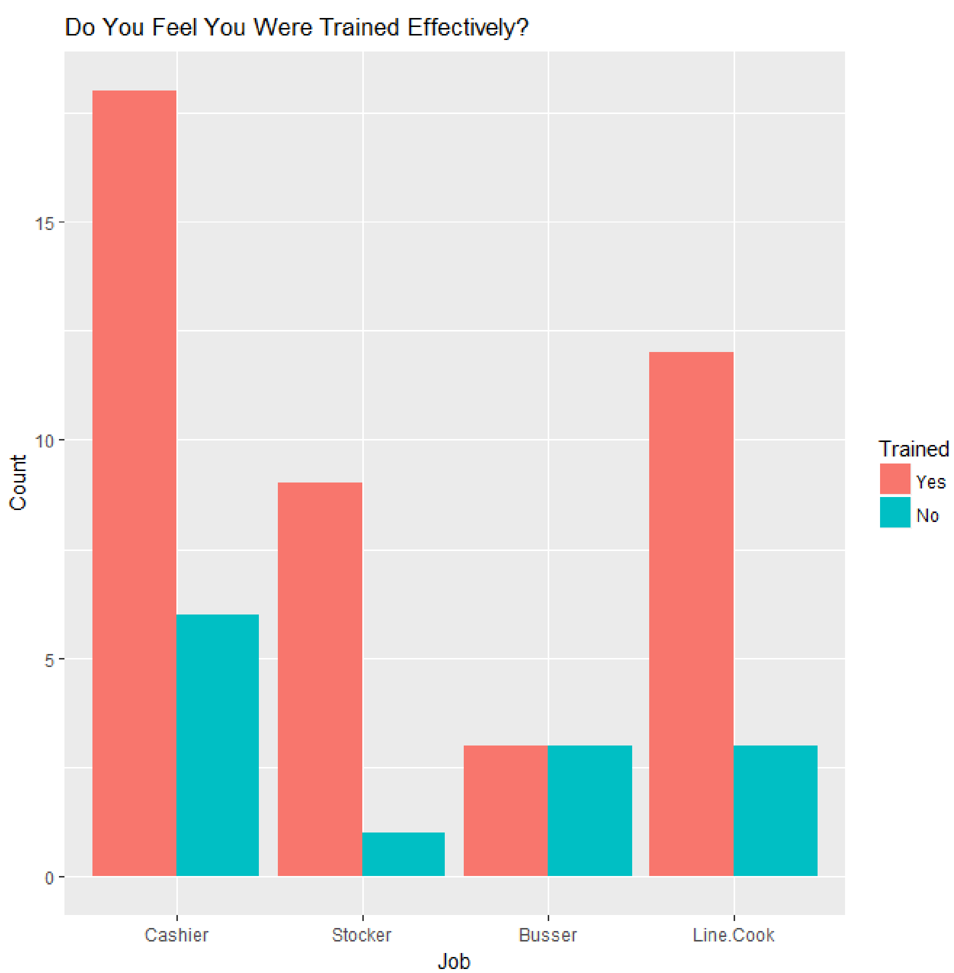
Job / Trained	Yes	No	Total
Cashier	18	6	24
Stocker	9	1	10
Busser	3	3	6
Line Cook	12	3	15
Total	42	13	**55**

Table 2.5

	Count	Prop
Trained Well	42	76%
Trained Poorly	13	24%

Table 2.6

Table 2.5 contains the raw data from question 4 broken down based on current Job Title. This dataset is uninteresting on its own, and unintuitive. Many veteran TM's were "trained" under the old program, and my hunch is that the proportion of TM's who are unhappy with the quality of their training should roughly line up with the proportion of TM's who have been here longer than three years. Unfortunately, I don't have data on that, and cannot run the tests. For the purposes of this study, Job Title and Training Effectiveness were found to be independent by a χ^2 Test and a Fisher's Exact Test, so there is no meaningful relationship found in this table. What I would like to point out is a continuation of the "Busser Problem". I would also like to point out that stockers are, once again, the most satisfied overall with their training. Though that might have something to do with the fact that the person administering the survey is also the person who trained them. You're welcome guys!



Graph 2.3

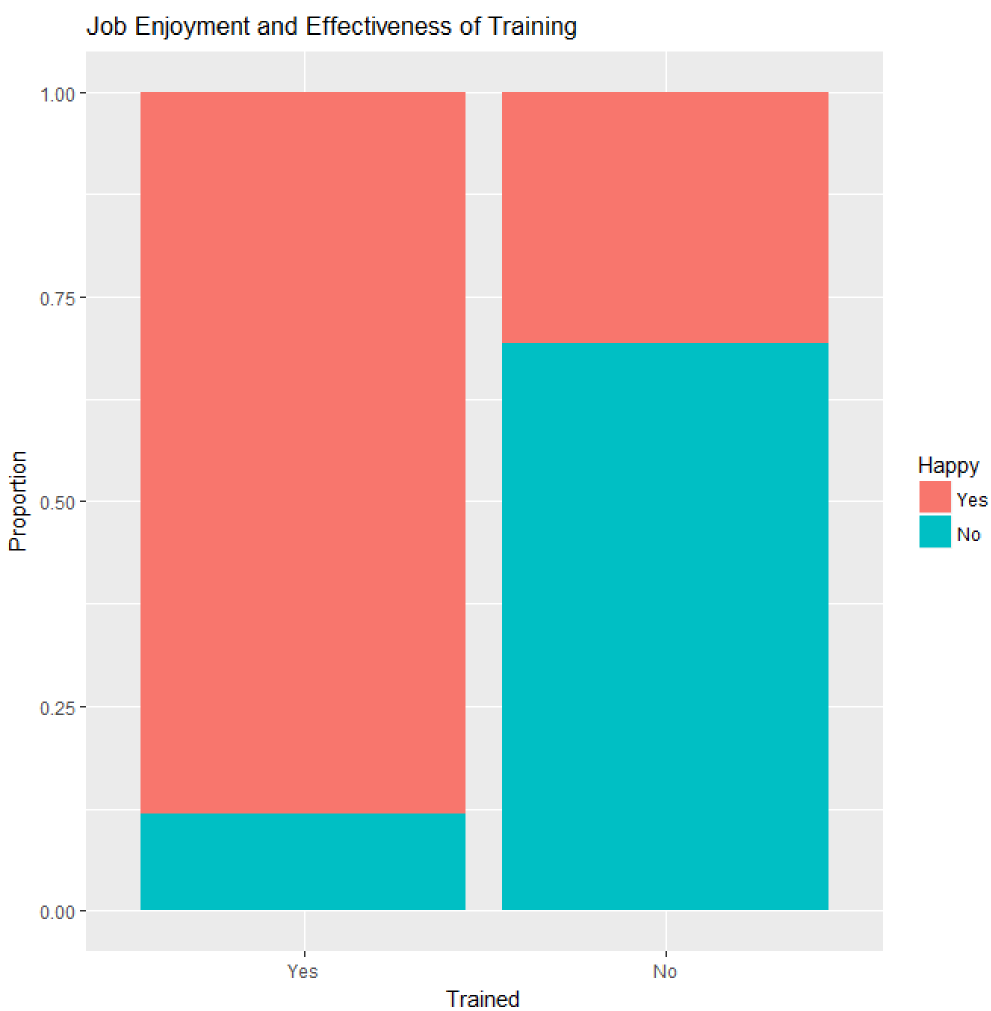
Enjoy – Trained

Happy / Trained	Yes	No	Total
Yes	37	4	41
No	5	9	14
Total	42	13	**55**

Table 2.7

Table 2.7 was one of the keystone pieces of the survey. The goal was to identify if Effective Training is a true determinant of Job Enjoyment or if there are other, more important, confounding factors. Performing Pearson's Chi Squared Test of Independence upon table 2.7 yields a p-value ≈ 0, allowing us to say with almost certainty that there is dependence between the two factors. Exploring the relationships odds ratio gives us $\theta = 16.65$, implying that the odds of an employee enjoying their job is 16.65 times higher is that employee was trained well.

The primary goal of this section was to develop a logit model for predicting how the happiness proportion will change as training improves but attempting to fit a model to this data created some of the worst models I have ever seen. On that, venture, I have not succeeded.



Graph 2.4

Here on graph 2.4, blocks of well trained and poorly trained have been equalized to better see the effect on happiness. TM's who are better trained are much more likely to enjoy working at Pizza Pred / West Carts

Study Limitations

These studies are limited by their scope. With such a small population to draw samples from, surveys are limited in the number of possible responses. I.E. with a sample size of 50, roughly half the venue, I can't ask a question with more than 5 possible answers. Such a question would give answers with an expected value of less than 10 and negate the effectiveness of parametric methods. Non-parametric methods can be used as a substitute but at a substantial loss of statistical power and interpretative ease. The second disadvantage to small sample sizes is the magnification of user error. 9 responses to question 2 had to be thrown away, that constitutes almost 20% of the sample. With a larger sample, such loss of data would be of less consequence. The third important disadvantage of a small population is that it becomes difficult to break the population down into subgroups. For example, if I were to survey cooks only, I would get a sample of approx. 20. This is too small for much besides a single binomial response. To summarize, better studies need more people, and we don't have many people. With more people, I could write more specific questions, but with less people, I have to keep things general to get enough responses in each category to power my tests.

Conclusion

This study has identified that Line Cook is the overwhelming "second favorite job" of the venue, with vastly more people preferring over the alternatives for a cross train. In close second place is Stocker with Cashier and Busser tying for last place. An interesting note from the first study is that no Line Cook said they would like to be cross trained as a Busser. The second part of the study found that 75% of the venue enjoys working here, and that a Stocker is more likely to enjoy working here than any other job title, but Bussers are less likely to enjoy working here than any other job. The third question found that 76% of the venue feels properly trained for their jobs, with Stockers more likely to feel well trained (WHOOOO), and Bussers the least likely to feel well trained. The final Part of the study identifies the strong connection between the effectiveness of a TM's training and their current job enjoyment. The odds of a TM enjoying working here are 16.65 times higher is the TM feels they were effectively trained.

The running issue discovered over and over again in this study was what I have been calling "The Busser Problem". Bussers are the least preferred job to cross train into, they are the least likely to enjoy working here, and they are the least likely to feel effectively trained. Perhaps there is something to be said in that the Job that interacts with Bussers more than anything doesn't have a single respondent saying they'd be willing to cross train as one. This paper does not offer any statistically backed advice for improving the job quality of the Busser besides improving training, but that does nothing for Bussers already at work. The advice of the author is to talk to the Bussers and figure out what they need and how they feel, then work out a solution with them. But this is a stats report not a management paper so do what you will.

Appendix A: Address to Comments Submitted on Ballots

Some TM's were thoughtful enough to share their thoughts and opinions with me on ballots submitted to the study. I figure I owe it to all of them to give them as detailed a response as they merit.

Comment 1: No One Cares

I care, ok?

Comment 2: Do you know who I am? (Written in the box with question4 – no answer supplied)

No

Comment 3: Stats is not Fun!

Don't hurt my feelings

Comment 4: You Surk

Your handwriting surks

Comment 5: I don't come to work to play your games, Devlin.

You say that now but when I bring in chutes and ladders next week, you're gonna be all over it

Appendix B: R Code and Associated Output

```
> ## This file deals with the JOB x CROSS TRAIN dataset
> ## I asked employees of west carts what their current job was and what job
> ## they would cross train into if they had to. The analysis is below
>
> setwd("D:/Files/Programing and Data/R Directory/Work Surveys")
>
> library(ggplot2)
Warning message:
package 'ggplot2' was built under R version 3.3.2
>
> # First we set up our Contingency Table
> JX <- read.csv("JobCross.csv")
> JX$Job <- factor(x = JX$Job, levels = c("cashier", "stocker", "busser", "line.cook"))
> JX$Cross.Train <- factor(x = JX$Cross.Train, levels = c("cashier", "stocker", "busser",
+ "line.cook"))
> JX.table <- xtabs(formula = Count ~ Job + Cross.Train, data = JX)
> sample.sum <- sum(JX.table)
>
>
> ## Now I want to be able to see the column sums
> ## They wont be added to the table, but are still useful to call upon
> cash.sum.x <- sum(JX.table[,1])
> stock.sum.x <- sum(JX.table[,2])
> buss.sum.x <- sum(JX.table[,3])
> cook.sum.x <- sum(JX.table[,4])
> xtrain.sums <- c(cash.sum.x, stock.sum.x, buss.sum.x, cook.sum.x)
> names <- c("cashier", "stocker", "busser", "cook")
> xtrain.props <- xtrain.sums / sample.sum
> JX.col.sums <- data.frame(names, xtrain.sums, xtrain.props)
>
> ## Here are the Row Sums
> ## This just gives me info about my sample
> cash.sum.c <- sum(JX.table[1,])
> stock.sum.c <- sum(JX.table[2,])
> buss.sum.c <- sum(JX.table[3,])
> cook.sum.c <- sum(JX.table[4,])
> job.sums <- c(cash.sum.c, stock.sum.c, buss.sum.c, cook.sum.c)
> job.props <- job.sums / sample.sum
> JX.row.sums <- data.frame(names, job.sums, job.props)
>
> ## Now I want to view the proportions on my full table
> JX.hat <- round(JX.table / sample.sum, 2)
>
> ## Now I want a table that details the proprtion of each cell divided by its row sum
> JX.r.hat <- round(JX.table / rowSums(JX.table), 2)
>
> ##################################################################
> ##             Lets Review What I Have So Far                  ##
>
> ## The Table
> JX.table
```

```
           Cross.Train
Job         cashier stocker busser line.cook
  cashier         0       8      4        11
  stocker         1       0      1         5
  busser          1       1      0         3
  line.cook       3       9      0         0
>
> ## The Row Summary
> ## Details Sample Respondants Current Job
> JX.row.sums
    names job.sums job.props
1 cashier       23 0.4893617
2 stocker        7 0.1489362
3  busser        5 0.1063830
4    cook       12 0.2553191
>
> ## The Column Summary
> ## Details Sample Respondants Cross Train Prefferences
> ## Ignores influence of current job - NOT GOOD STATS
> JX.col.sums
    names xtrain.sums xtrain.props
1 cashier           5    0.1063830
2 stocker          18    0.3829787
3  busser           5    0.1063830
4    cook          19    0.4042553
>
> ## The table of proportions
> JX.hat
           Cross.Train
Job         cashier stocker busser line.cook
  cashier      0.00    0.17   0.09      0.23
  stocker      0.02    0.00   0.02      0.11
  busser       0.02    0.02   0.00      0.06
  line.cook    0.06    0.19   0.00      0.00
>
> ## The table of intra-Row proportions
> JX.r.hat
           Cross.Train
Job         cashier stocker busser line.cook
  cashier      0.00    0.35   0.17      0.48
  stocker      0.14    0.00   0.14      0.71
  busser       0.20    0.20   0.00      0.60
  line.cook    0.25    0.75   0.00      0.00
>
> ##                                                               ##
> ###################################################################
>
> ## Lets make some observations
> ## Cashiers want to be cooks and stockers
> ## Cooks want to be stockers
> ## Bussers want to be cooks
> ## Stockers want to be cooks
>
```

```
> ## GOALS:
> ### Rank Venue Job Prefferences
> ### Test Independance
> ### Make 1 Pretty Chart in GGPlot2
>
> ## Start where its easy: Testing Independance:
> chisq.test(JX.table, correct = FALSE)

        Pearson's Chi-squared test

data:  JX.table
X-squared = 22.869, df = 9, p-value = 0.006498

Warning message:
In chisq.test(JX.table, correct = FALSE) :
  Chi-squared approximation may be incorrect
> ## BAM that what I like
> ## Easily reject the null and conclude dependance
> fisher.test(JX.table)

        Fisher's Exact Test for Count Data

data:  JX.table
p-value = 0.0002162
alternative hypothesis: two.sided

> ## these results are fairly obvious
> ## of course cross.train is dependant upon your current job
> ## because you cant cross train into your current job
> chisq.test(JX.table, correct = FALSE)$residual
           Cross.Train
Job            cashier    stocker     busser  line.cook
  cashier   -1.5642278 -0.2724172  0.9929446  0.5582135
  stocker    0.2958682 -1.6373305  0.2958682  1.2901061
  busser     0.6418060 -0.6611474 -0.7293250  0.6884099
  line.cook  1.5253182  2.0544451 -1.1298654 -2.2025131
Warning message:
In chisq.test(JX.table, correct = FALSE) :
  Chi-squared approximation may be incorrect
> ## The residuals prove my point
>
> ## Ranking Venue Job Preference
> ## Fairly Straighforward
> ### 1. Cook
> ### 2. Stocker
> ### 3. Busser
> ### 4. Cashier
> #### Why?
> ##### More cashiers than anyone else in study dilutes possible cashier responses
> ##### Still doesnt account for the proportions that are heavily in favor of stocking
and cooking
>
> ## GGPlot Graph
```

```
> JX.p <- JX
> r.hat <- as.vector(t(JX.r.hat))
> r.props <- r.hat[-c(1, 6, 11, 16)]
> JX.p[3] <- r.props
> ggplot(JX.p, aes(x = Job, y = Count, fill = Cross.Train)) +
+ geom_bar(position = "dodge", stat = "identity") +
+ labs(y = 'Proportion', title = 'Intra-Job Cross Train Proportions')
>
>
> ## This file deals with the JOB x HAPPY dataset
> ## I asked employees of west carts what their current job was and if
> ## they were happy working here. The analysis is below
>
> setwd("D:/Files/Programing and Data/R Directory/Work Surveys")
>
> library(ggplot2)
>
> # First we set up our Contingency Table
> JH <- read.csv("JobHappy.csv")
> JH$Job <- factor(x = JH$Job, levels = c("Cashier", "Stocker", "Busser", "Line.Cook"))
> JH$Happy <- factor(x = JH$Happy, levels = c("Yes", "No"))
> JH.table <- xtabs(formula = Count ~ Job + Happy, data = JH)
> h.sample.sum <- sum(JH.table)
>
> ## Now I want to be able to see the column sums
> ## They wont be added to the table, but are still useful to call upon
> happy.sum <- sum(JH.table[,1])
> no.sum <- sum(JH.table[,2])
> happynot.sums <- c(happy.sum, no.sum)
> names.h <- c("Happy", "Not")
> happy.props <- happynot.sums / h.sample.sum
> JH.col.sums <- data.frame(names.h, happynot.sums, happy.props)
>
> ## Here are the Row Sums
> ## This just gives me info about my sample
> cash.sum.h <- sum(JH.table[1,])
> stock.sum.h <- sum(JH.table[2,])
> buss.sum.h <- sum(JH.table[3,])
> cook.sum.h <- sum(JH.table[4,])
> h.job.sums <- c(cash.sum.h, stock.sum.h, buss.sum.h, cook.sum.h)
> h.job.props <- h.job.sums / h.sample.sum
> JH.row.sums <- data.frame(names, h.job.sums, h.job.props)
>
> ## Now I want to view the proportions on my full table
> JH.hat <- round(JH.table / h.sample.sum, 2)
>
> ## Now I want a table that details the proprtion of each cell divided by its row sum
> JH.r.hat <- round(JH.table / rowSums(JH.table), 2)
>
> ################################################################
> ##             Lets Review What I Have So Far                 ##
>
> ## The Table
```

```
> JH.table
           Happy
Job          Yes No
  Cashier    19  6
  Stocker     9  1
  Busser      3  3
  Line.Cook  11  4
>
> ## The Row Summary
> ## Details Sample Respondants Current Job
> JH.row.sums
    names h.job.sums h.job.props
1 cashier         25   0.4464286
2 stocker         10   0.1785714
3  busser          6   0.1071429
4    cook         15   0.2678571
>
> ## The Column Summary
> ## Details Sample Respondants Cross Train Prefferences
> ## Ignores influence of current job - NOT GOOD STATS
> JH.col.sums
  names.h happynot.sums happy.props
1   Happy            42        0.75
2     Not            14        0.25
>
> ## The table of proportions
> JH.hat
           Happy
Job           Yes   No
  Cashier   0.34 0.11
  Stocker   0.16 0.02
  Busser    0.05 0.05
  Line.Cook 0.20 0.07
>
> ## The table of intra-Row proportions
> JH.r.hat
           Happy
Job           Yes   No
  Cashier   0.76 0.24
  Stocker   0.90 0.10
  Busser    0.50 0.50
  Line.Cook 0.73 0.27
>
> ##                                                              ##
> ##################################################################
>
> ## Lets make some Observations
> ## Bussers Not Happy
> ## Stockers very happy
> ## overall, venue 75% happy
>
> ## GOALS:
> ### Test of Independance
```

```
> ### Pretty Graph
>
> ## Starting iwth a test of Independance
> chisq.test(JH.table, correct = FALSE)

        Pearson's Chi-squared test

data:  JH.table
X-squared = 3.2356, df = 3, p-value = 0.3567

Warning message:
In chisq.test(JH.table, correct = FALSE) :
  Chi-squared approximation may be incorrect
> ## Fail to reject the Null with pearsons chi-squared
> fisher.test(JH.table)

        Fisher's Exact Test for Count Data

data:  JH.table
p-value = 0.3793
alternative hypothesis: two.sided

> ## Same Result
>
> ## Graph
> ggplot(JH, aes(x = Job, y = Count, fill = Happy)) +
+ geom_bar(position = "dodge", stat = "identity") +
+ labs(y = 'Count', title = 'Do You Enjoy Working Here?')
>
>
>
>
> ## This file deals with the JOB x TRAINED dataset
> ## I asked employees of west carts what their current job was and if
> ## they felt they were trained effectively. The analysis is below
>
> setwd("D:/Files/Programing and Data/R Directory/Work Surveys")
>
> library(ggplot2)
>
> # First we set up our Contingency Table
> JT <- read.csv("JobTrain.csv")
> JT$Job <- factor(x = JT$Job, levels = c("Cashier", "Stocker", "Busser", "Line.Cook"))
> JT$Trained <- factor(x = JT$Trained, levels = c("Yes", "No"))
> JT.table <- xtabs(formula = Count ~ Job + Trained, data = JT)
> t.sample.sum <- sum(JT.table)
>
> ## Now I want to be able to see the column sums
> ## They wont be added to the table, but are still useful to call upon
> train.sum <- sum(JT.table[,1])
> not.sum <- sum(JT.table[,2])
> trainnot.sums <- c(train.sum, not.sum)
> names.t <- c("Trained", "Not")
```

```
> train.props <- trainnot.sums / t.sample.sum
> JT.col.sums <- data.frame(names.t, trainnot.sums, train.props)
>
> ## Here are the Row Sums
> ## This just gives me info about my sample
> cash.sum.t <- sum(JT.table[1,])
> stock.sum.t <- sum(JT.table[2,])
> buss.sum.t <- sum(JT.table[3,])
> cook.sum.t <- sum(JT.table[4,])
> t.job.sums <- c(cash.sum.t, stock.sum.t, buss.sum.t, cook.sum.t)
> t.job.props <- t.job.sums / t.sample.sum
> JT.row.sums <- data.frame(names, t.job.sums, t.job.props)
>
> ## Now I want to view the proportions on my full table
> JT.hat <- round(JT.table / t.sample.sum, 2)
>
> ## Now I want a table that details the proprtion of each cell divided by its row sum
> JT.r.hat <- round(JT.table / rowSums(JT.table), 2)
>
> #################################################################
> ##             Lets Review What I Have So Far                  ##
>
> ## The Table
> JT.table
           Trained
Job          Yes No
  Cashier    18  6
  Stocker     9  1
  Busser      3  3
  Line.Cook  12  3
>
> ## The Row Summary
> ## Details Sample Respondants Current Job
> JT.row.sums
    names t.job.sums t.job.props
1 cashier         24   0.4363636
2 stocker         10   0.1818182
3  busser          6   0.1090909
4    cook         15   0.2727273
>
> ## The Column Summary
> ## Details Sample Respondants Cross Train Prefferences
> ## Ignores influence of current job - NOT GOOD STATS
> JT.col.sums
  names.t trainnot.sums train.props
1 Trained            42   0.7636364
2     Not            13   0.2363636
>
> ## The table of proportions
> JT.hat
           Trained
Job          Yes   No
  Cashier   0.33 0.11
```

```
  Stocker   0.16 0.02
  Busser    0.05 0.05
  Line.Cook 0.22 0.05
>
> ## The table of intra-Row proportions
> JT.r.hat
           Trained
Job          Yes   No
  Cashier   0.75 0.25
  Stocker   0.90 0.10
  Busser    0.50 0.50
  Line.Cook 0.80 0.20
>
> ##                                                              ##
> ##################################################################
>
> ## Lets make some observations
> ## Most people feel well trained
> ## Except Bussers
> ## Cooks and Stockers recieve above average training
> ## Cashiers and Bussers recieve below average training
> ## There isnt much here thats interesting
>
> ## GOALS:
> ### Test of independance
> ### Graph
>
> ## Here is a Chi Squared Test of Independance
> chisq.test(JT.table, correct = FALSE)

        Pearson's Chi-squared test

data:  JT.table
X-squared = 3.4753, df = 3, p-value = 0.324

Warning message:
In chisq.test(JT.table, correct = FALSE) :
  Chi-squared approximation may be incorrect
> ## We fail to reject the null here, That much is obvious from looking at the data
> fisher.test(JT.table)

        Fisher's Exact Test for Count Data

data:  JT.table
p-value = 0.3646
alternative hypothesis: two.sided

> ## Same
>
> ## Graph
> ggplot(JT, aes(x = Job, y = Count, fill = Trained)) +
+ geom_bar(position = "dodge", stat = "identity") +
+ labs(y = 'Count', title = 'Do You Feel You Were Trained Effectively?')
```

```
> ## This file deals with the HAPPY x TRAINED dataset
> ## I asked employees of west carts whether they enjoy working here and if
> ## they felt they were trained effectively. The analysis is below
>
> setwd("D:/Files/Programing and Data/R Directory/Work Surveys")
>
> library(ggplot2)
>
> # First we set up our Contingency Table
> HT <- read.csv("HappyTrained.csv")
> HT$Happy <- factor(x = HT$Happy, levels = c("Yes", "No"))
> HT$Trained <- factor(x = HT$Trained, levels = c("Yes", "No"))
> HT.table <- xtabs(formula = Count ~ Happy + Trained, data = HT)
> ht.sample.sum <- sum(HT.table)
>
> ## Now I want to be able to see the column sums
> ## They wont be added to the table, but are still useful to call upon
> ht.sum <- sum(HT.table[,1])
> nht.sum <- sum(HT.table[,2])
> ht.sums <- c(ht.sum, nht.sum)
> ht.props <- ht.sums / ht.sample.sum
> HT.col.sums <- data.frame(names.h, ht.sums, ht.props)
>
> ## Here are the Row Sums
> ## This gives me info about Trained
> sum.ht <- sum(HT.table[1,])
> sum.nht <- sum(HT.table[2,])
> sums.ht <- c(sum.ht, sum.nht)
> props.ht <- sums.ht / ht.sample.sum
> HT.row.sums <- data.frame(names.t, sums.ht, props.ht)
>
> ## Now I want to view the proportions on my full table
> HT.hat <- round(HT.table / ht.sample.sum, 2)
>
> ## Now I want a table that details the proprtion of each cell divided by its row sum
> HT.r.hat <- round(HT.table / rowSums(HT.table), 2)
>
> ################################################################
> ##             Lets Review What I Have So Far                  ##
>
> ## The Table
> HT.table
     Trained
Happy Yes No
  Yes  37  4
  No    5  9
>
> ## The Row Summary
> ## Details Sample Respondants Happiness
> HT.row.sums
  names.t sums.ht  props.ht
1 Trained      41 0.7454545
2     Not      14 0.2545455
```

```
>
> ## The Column Summary
> ## Details Sample Respondants Training
> HT.col.sums
  names.h ht.sums  ht.props
1   Happy      42 0.7636364
2     Not      13 0.2363636
>
> ## The table of proportions
> HT.hat
     Trained
Happy  Yes   No
  Yes 0.67 0.07
  No  0.09 0.16
>
> ## The table of intra-Row proportions
> HT.r.hat
     Trained
Happy  Yes   No
  Yes 0.90 0.10
  No  0.36 0.64
>
> ##                                                              ##
> ################################################################
>
> ## OBSERVATIONS
> ## Happiness and Training seem to be related....
> ## Still lots of people who were well trained and are still unhappy
> ## Suggests more variables influence happiness than just training
>
> ## GOALS:
> ### Odds Ratio
> ### Chi Sq Test
> ### Graph
> ### Fit a Logit Model
>
> ## Starting with whats easy
> chisq.test(HT.table, correct = FALSE)

        Pearson's Chi-squared test

data:  HT.table
X-squared = 17.193, df = 1, p-value = 3.377e-05

Warning message:
In chisq.test(HT.table, correct = FALSE) :
  Chi-squared approximation may be incorrect
> ## Easily reject the null hypothesis of independance
>
> ## Odds Ratio
> HT.or <- round((HT.table[1,1] * HT.table[2,2]) / (HT.table[2,1] * HT.table[1,2]), 2)
> HT.or
[1] 16.65
```

```
> ## This implies that the odds of an emplyee Enjoying work at west carts
> ## are 16.65 times higher if the employee has been properly trained
> ## WOWZA
>
> ## Now to fit a Logit Model
> HT.model <- glm(Happy ~ Trained, binomial(link = 'logit'), data = HT.table)
> summary(HT.model)

Call:
glm(formula = Happy ~ Trained, family = binomial(link = "logit"),
    data = HT.table)

Deviance Residuals:
     1       2       3       4
-1.177   1.177  -1.177   1.177

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  4.441e-16  1.414e+00       0        1
TrainedNo   -4.441e-16  2.000e+00       0        1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 5.5452  on 3  degrees of freedom
Residual deviance: 5.5452  on 2  degrees of freedom
AIC: 9.5452

Number of Fisher Scoring iterations: 2

> ## this model is the worst model I have ever seen in my entire life
>
>
> ggplot(HT, aes(x = Trained, y = Count, fill = Happy)) +
+ geom_bar(position = "fill", stat = "identity") +
+ labs(y = 'Count', title = 'Job Enjoyment and Effectiveness of Training')
```