FINAL

Emmalyn Holmquist

2024-12-13

In examining the question "If people are attending the market regularly, what brings them?", we decided to consider the following questions from the data to inform our analysis: 1) How do the survey respondents rank the importance of quality, variety, and quantity at the market on a scale of 1-3? 2) What role are the respondents attending the market? 3) Would the respondents agree or disagree that the market is an important resource for them?

In our initial phases of exploratory data analysis, we explored relationships between role and agreement of the market as an important resource, role and rank of quality, variety and quantity, and the distribution of roles of the respondents who attended the market.

Of the respondents who attend the market, we discovered that most respondents are either a student or faculty/staff member, with the majority being a staff member or an on-campus student. From the survey alone, we can see that approximately 32% of the market goers are staff, while approximately 28% are on campus students. Table C.1 below illustrates the proportion of various roles of survey-takers, along with the frequencies of each role. Additionally, Figure C.2 (found in the Appendix) illustrates a visualization of this distribution.

Table C.1: Frequency and Proportion of the Role of Survey Respondents

Role	Count	Proportion
Faculty	11	0.1549
Staff	23	0.3239
On-Campus Student (Resident)	20	0.2817
Off-Campus Student (Commuter)	13	0.1831
Other	4	0.0563

In addressing our question, we first analyze the relationship between importance of quantity, variety, quality, and attendance frequency, to see if a customer's perceived importance on quantity, variety, or quality effects the frequency of market visits, or visa versa. From this, we would hope to gather insight as to if customers had different expressed importance on variety, quality, or quantity of produce based on their frequency of market attendance, which might inform us as to why they come to the market. However, we discovered that the distribution of attendance groups among rank of importance for each group (quality, quantity, and variety) were represented relatively equally to each other across ranks (about 0.50 for Weekly and about 0.20 for Bi-Weekly), as visualized in the stacked bar graphs of their conditional proportions (Figures C.6-C.8), meaning that there didn't appear to be a difference in preference based on frequency of attendance. We will also note here that while there does appear to be a difference in the proportion of groups among ranks for variety, the difference is not that large. Further, the difference in proportions for variety is greatest in rank 2, which isn't very useful (it suggests that perhaps a larger portion of Weekly market attenders value variety the most middlest, whereas we are looking for data that shows a group values something the most or least). In addition, while the proportion for the Once group does not appear equal among all ranks, we

can comfortably lump this data into the Occasionally group, as we are mainly looking at the difference in those who attend the market regularly vs. not regularly. Based on the knowledge that the proportion of representation of attendance groups was largely equal among ranks, we can deem the visualizations of the bar graphs alone as inconclusive. Further, we would desire to perform a Chi-Square Test of independence, to see if there is a relationship between attendance frequency and perceived importance of quantity, variety, and quality of produce (since we sampled from one population and are examining the relationship between two categorical variables, a Chi-Square Test for Independence is most appropriate). However, these rankings were not independent, and so such a test would be inappropriate, as independence is one of the assumptions necessary for a Chi-Square. This poses a limitation on our available tests, in which the survey should be redesigned to accommodate this. Our suggestion would be for the survey to allow survey respondents to give each group (quantity, quality, and variety) a rank from 1-10 individually of how strongly they felt the importance of each group is. In this way, we could also compute the average score among each perceived importance.

Figure C.3: Attendance vs. Quality Importance

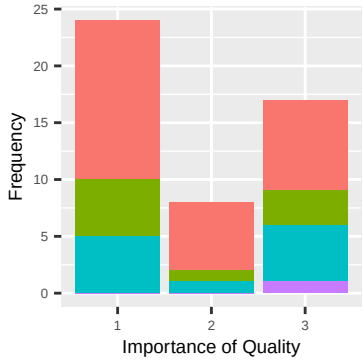


Figure C.4: Attendance vs. Variety Importance

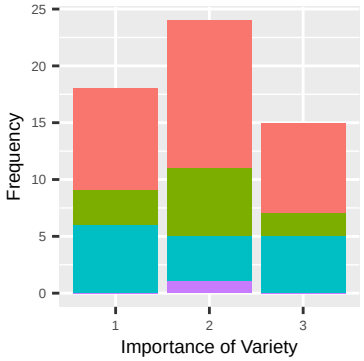


Figure C.5: Attendance vs. Quantity Importance

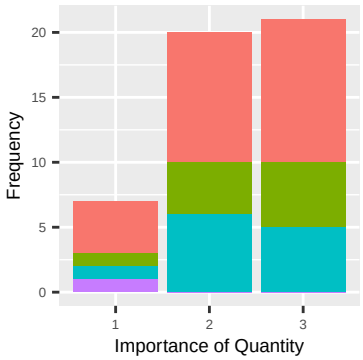


Figure C.8: Conditional Proportions of Attendance vs. Quality Importance

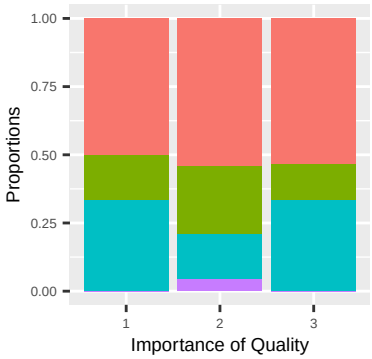


Figure C.6: Conditional Proportions of Attendance vs. Variety Importance

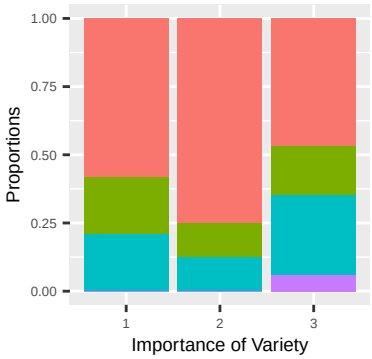
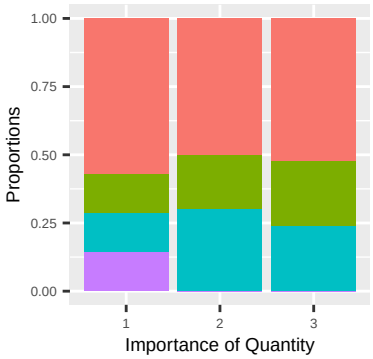
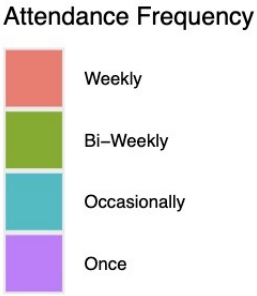


Figure C.7: Conditional Proportions of Attendance vs. Quantity Importance



Legend for Graphs in Figure C.3-C.8



Despite this, in examining Figures C.3-C.5, we can see that there does appear to be a slight perceived importance on quality, and a slight lack of perceived importance on quantity among all groups combined, which gives us some insight as to why people may attend the market in general. Further, the total votes for most important among each preference were 18 for variety, 24 for quality, and 7 for quantity, making the order of importance quality, variety, then quantity. However, it is also important to examine the counts of the rank for least important. The alternative counts for least importance were 15 for variety, 17 for quality, and 21 for quantity; notably, the counts for the least important and most important ranks for quality and variety individually are relatively similar. Thus, to decipher if there is a significant difference of perceived importance on variety, quality, and quantity of goods, we would desire to perform an ANOVA test to examine if there is a statistically significant difference in preference; but, as previously discussed, the rankings of importance were not independent, thus an ANOVA test would be inappropriate to perform. To accommodate for this limitation, we might suggest a similar amendment as above. The importance ranks of quality, variety, and quantity summed across all groups are summarized in the table below.

Table C.9: Frequency of Perceived Importance of Quality, Variety, and Quantity

Perceived Importance	1	2	3
Quality Importance	24	8	17
Variety Importance	18	24	15
Quantity Importance	7	20	21

However, perceived importance and satisfaction of quantity of produce, quality of produce, and variety of produce are not the only factors that would determine why someone would want to attend the market frequently versus infrequently. Thus, we chose to examine if respondents see the market as an important resource for them based on their different attendance groups and role within the Saint Mary's Community to see if those who attend the market more frequently value the market as an important resource more than those who don't, and if those who see the market as more of an important resource are concentrated within certain roles.

Figure C.10: Important Resource by Role

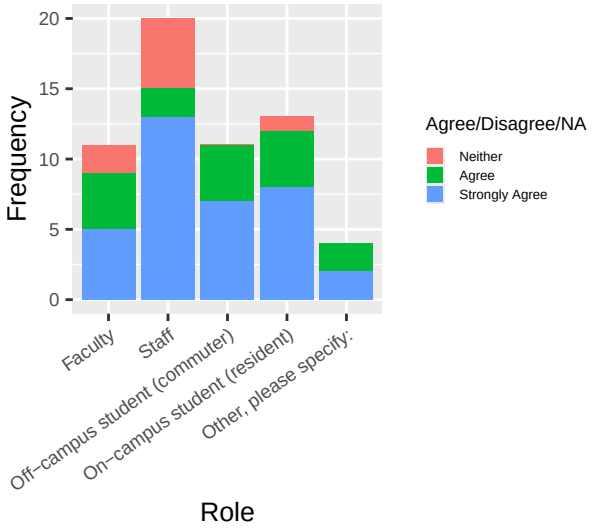
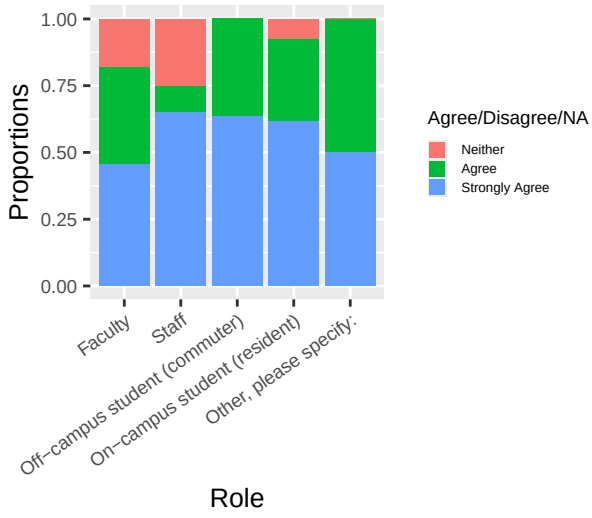


Figure C.11: Conditional Proportions of Important Resource by Role



First, in examining the market as an important resource based on role, we can gather from the frequency graph (Figure C.10) that the majority among all roles are in agreement or strong agreement that the market is an important resource for them personally. Further, from the conditional proportions graph (Figure C.11), we can see that while almost half of the survey takers strongly agree that the market is an important resource for them, nearly all of the survey takers agree that the market is an important resource for them. Additionally, there is no data that suggests that the market is not an important resource (nobody disagreed). From Figure C.11, we can also see that each role appears to hold relatively the same proportion of equal agreement, although there were a few respondents that felt indifferent or unsure among the Faculty, Staff, and On-Campus Student groups. This gives us some insight that overall, those who attend the market may do so because the market is an important resource for them, and that Off-Campus Students may feel the most strongly that the market is an important resource for them. However, the survey doesn't provide us with what a clear definition of "important resource" means. In a survey redesign, we might suggest that the survey include a precise definition, or allow the user to select multiple: "the market is an important resource for me financially" "the market is an important resource for me as an inclusive environment" etc. Additionally, those in the Other role identified as a retired professor, parent of a student, and parent of someone who works on campus. We only have three responses in this category, thus we do not have enough information to make conclusions about this group, especially as the "Other" category is extremely variable. Additionally, while we have some evidence to suggest that Off-Campus Students felt the most strongly that the market was the greatest resource, we don't have enough principle to base this off on, since none of the other groups showed evidence of disagreeing that the market was an important resource. Further, it may be a waste to test if there is a group that values the market as more of an important resource than the others, as overall nobody disagreed that it was important and the majority of all groups strongly agreed that it was.

Moving forward, we can examine whether the market is an important resource based on attendance frequency to see if those who attend the market more frequently see the market as more of a resource versus those who don't.

Figure C.12: Important Resource by Attendance Frequency

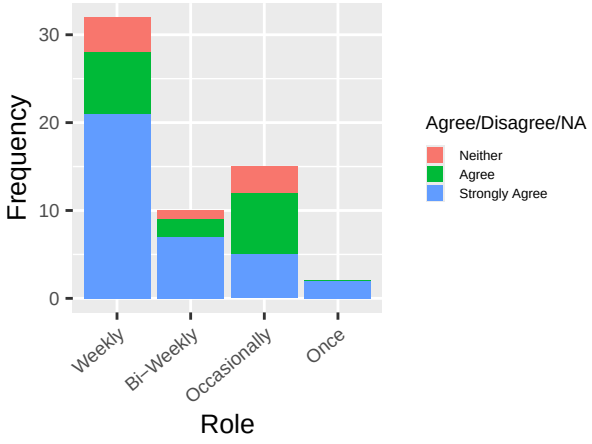
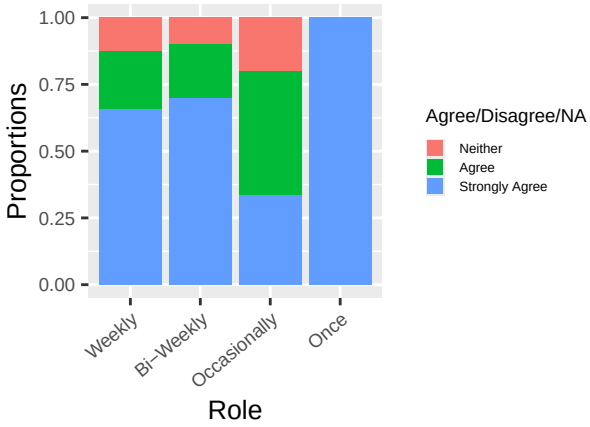


Figure C.13: Conditional Proportions of Important Resource by Attendance Frequency



Based on our graph of conditional proportions (Figure C.13), we can see that the majority of all groups agree or strongly agree that the market is an important resource, although there appears to be slightly stronger agreement among the Weekly and Bi-Weekly groups than the Occasionally attendance group. Further, it is important to note that while we see 100% strong agreement for the Once group, there are only one or two observations in this group which are causing the data to be highly sensitive. While we anticipate that those who attend the market more frequently would agree more strongly that the market is an important resource for them, this is not supported by our findings that 100% of the group that only attended the market once strongly agree that the market is an important resource for them. Further, one possibility for this response is that this response is a parent of a Saint Mary's student, who has only attended once, and sees the farm as an important resource for them personally because it ensures their student has access to good quality produce. However, this doesn't necessarily give insight as to why they have only attended the market once, thus more data needs to be collected on this group. Further, I would predict that when getting more data on this group, we would see a more diverse range of responses (strongly agree, agree, disagree, strongly disagree).

If we take this into account, there does appear to be a heightened sense of strong agreement among the Weekly and Bi-Weekly groups. To determine if those that come to the market more frequently see the market as more of an important resource to them personally, we could conduct a Chi-Square test of Independence to see if there is a relationship between attendance frequency and strong agreement that the market is an important resource. This test would be most appropriate as our data comes from one sample, and is testing the relationship between two categorical variables (attendance frequency and agreement level that the market is an important resource). Although our data is independent, we fail one assumption of the Chi-Square Test, since not all of our expected counts are above 5 (as shown in Table C.14 in the Appendix). Thus, we will proceed with permutation re-sampling to estimate our Chi-Square test statistic.

First we state our hypothesis:

H_0 : attendance and the level of agreement that the market is an important resource are independent;

H_A : attendance is related to the level of agreement that the market is an important resource

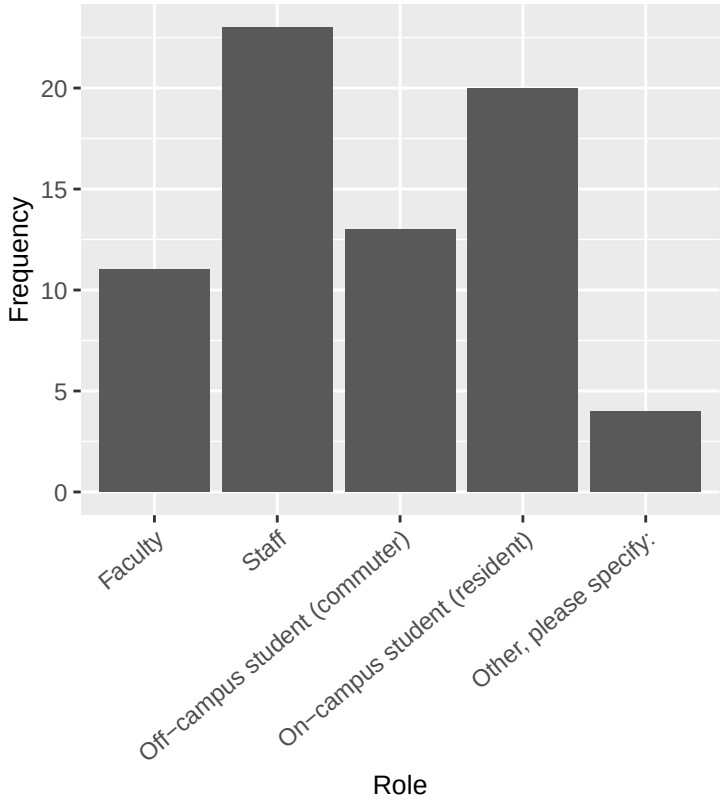
After conducting our test, we observe a test statistic of 6.6895 with p-value 0.3437. This means that we would expect a 6.6895 magnitude of discrepancy between observed and expected data. Since our p-value is not very small, at 0.3437, and if we evaluate at alpha level 0.05, $0.3437 > 0.05$; thus, we do not have sufficient evidence to reject the null hypothesis, and we therefore do not have evidence to support a relationship between attendance and the level of agreement that the market is an important resource.

In summary, while there were many tests we were not able to perform because of limitations of the survey (lack of independent data and small sample size), we were able to analyze a possible relationship between attendance and the level of agreement that the market is an important resource (through a Chi-Square Test of independence) in order to give us insight as to why survey respondents may attend the market more regularly versus those that don't. While our Chi-Square Test suggested evidence to rule out this possible relationship, we discovered other insights through process of exploratory data analysis, in which we learned that the majority of the people that come to the market come on a regular basis, and that the majority of these people are staff. Additionally, we learned that in general, based on the responses to the survey regarding importance of quality, variety, and quantity of produce, market attendees may value quality the most, and quantity the least. Further, attendance groups don't appear to be biased on their perceived importance of variety, quantity, and quality of produce, because each group is relatively equally represented among them. In addition, across all roles, we see that nearly every group agrees or agrees strongly that the market is an important resource for them. Finally, in a survey re-design, suggestions were made to allow for more independent responses, including the re-framing of survey questions, that would allow for more possible tests to be ran.

Technical Appendix

```
##Table C.1 and Figure C.2
#get info on the roles
role <- survey$Q1_text
role <- factor(role, levels = c("Faculty",  "Staff", "Off-campus student (commuter)",
                                "On-campus student (resident)", "Other, please specify:"))

library(ggplot2)

#get ranks for quality and variety
quality_importance <- survey$Q15_1
variety_importance <- survey$Q15_3

#make df
df3 <- data.frame(role, quality_importance)
no_NA_df3 <- na.omit(df3)

#make a contingency table of df
table3 <- table(no_NA_df3)

#ggplot - distribution of market goers based on role (no omitted survey)
ggplot(df3, aes(x = role)) + geom_bar() + scale_x_discrete(guide = guide_axis(angle = 40))+
  ggtitle("Figure C.2: Distribution of Market Respondents") +
  xlab("Role") + ylab("Frequency") + theme(
                                        plot.title = element_text(size = 10),
                                        axis.title = element_text(size = 10),
                                        aspect.ratio = 4/5)
```

Figure C.2: Distribution of Market Respondents




```
#frequency and proportion of roles
table(role)
```

```
## role
##                         Faculty                           Staff
##                              11                              23
## Off-campus student (commuter)  On-campus student (resident)
##                              13                              20
##        Other, please specify:
##                               4
```

```
proportions(table(role))
```

```
## role
##                         Faculty                           Staff
##                      0.15492958                      0.32394366
## Off-campus student (commuter)  On-campus student (resident)
##                      0.18309859                      0.28169014
##        Other, please specify:
##                      0.05633803
```

```
##Figures C.3-C.5
#of those that attend frequently, how do they rank stuff
#get attendance variable loaded
```

```
attendance_freq <- survey$Q9_code
#get ranks for quality and variety
quality_importance <- survey$Q15_1
amount_importance <- survey$Q15_2
variety_importance <- survey$Q15_3

###quality
df7 <- data.frame(quality_importance, attendance_freq)
no_NA_df7 <- na.omit(df7)
#make a contingency table of df
table7 <- table(no_NA_df7)

#this helps us to make the ggplot graph; get the counts, row/col names
table.df7 <- data.frame(expand.grid(rownames(table7), colnames(table7)), c(table7))
colnames(table.df7) <- c("importance", "attendance", "count")

#ggplot - make graph based on counts
c_quality <- ggplot(table.df7, aes(x = importance, y = count, fill = attendance)) +
  geom_bar(stat = "identity") +
  scale_fill_discrete(name = "Attendance Frequency",
                      labels = c("Weekly","Bi-Weekly", "Occasionally", "Once")) +
  ggtitle("Figure C.3: Attendance vs. \nQuality Importance") +
  xlab("Importance of Quality") + ylab("Frequency") +
  theme(plot.title = element_text(size = 7), axis.title = element_text(size = 7),
        axis.text = element_text(size = 5), aspect.ratio = 1, legend.position = "none")

###variety
df8 <- data.frame(variety_importance, attendance_freq)
no_NA_df8 <- na.omit(df8)
#make a contingency table of df
table8 <- table(no_NA_df8)


#this helps us to make the ggplot graph; get the counts, row/col names
table.df8 <- data.frame(expand.grid(rownames(table8), colnames(table8)), c(table8))
colnames(table.df8) <- c("importance", "attendance", "count")

#ggplot - make proportions graph based on counts
c_variety <- ggplot(table.df8, aes(x = importance, y = count, fill = attendance)) +
  geom_bar(stat = "identity") +
  scale_fill_discrete(name = "Attendance Frequency",
                      labels = c("Weekly", "Bi-Weekly", "Occasionally", "Once")) +
  ggtitle("Figure C.4: Attendance vs. \nVariety Importance") +
  xlab("Importance of Variety") + ylab("Frequency") +
  theme(plot.title = element_text(size = 7), axis.title = element_text(size = 7),
        axis.text = element_text(size = 5), aspect.ratio = 1, legend.position = "none")


###quantity
df9 <- data.frame(amount_importance, attendance_freq)
no_NA_df9 <- na.omit(df9)
#make a contingency table of df
table9 <- table(no_NA_df9)
```

```r
#this helps us to make the ggplot graph; get the counts, row/col names
table.df9 <- data.frame(expand.grid(rownames(table9), colnames(table9)), c(table9))
colnames(table.df9) <- c("importance", "attendance", "count")

#ggplot - make proportions graph based on counts
c_quantity <- ggplot(table.df9, aes(x = importance, y = count, fill = attendance)) +
  geom_bar(stat = "identity") +
  scale_fill_discrete(name = "Attendance Frequency",
                      labels = c("Weekly", "Bi-Weekly", "Occasionally", "Once")) +
  ggtitle("Figure C.5: Attendance vs. \nQuantity Importance") +
  xlab("Importance of Quantity") + ylab("Frequency") +
  theme(plot.title = element_text(size = 7), axis.title = element_text(size = 7),
        axis.text = element_text(size = 5), aspect.ratio = 1, legend.position = "none")
```

```r
##Figures C.6-C.8
#importance among each grouping proportions
##variety
#ggplot - make proportion table based on conditional probabilities - conitional probabilities graph
proportions_variety <- proportions(table7, 1)

#make condiitonal probabilities graph
proportions7.df <- data.frame(expand.grid(rownames(proportions_variety),
                                          colnames(proportions_variety)),
                              c(proportions_variety))
colnames(proportions7.df) <- c("variety_importance", "attendance", "Proportion")
p_variety <-ggplot(proportions7.df, aes(x = variety_importance, y = Proportion, fill = attendance)) +
  geom_bar(stat = "identity") +
  scale_fill_discrete(name = "Attendance Frequency",
                      labels = c("Weekly","Bi-Weekly", "Occasionally", "Once"))  +
  ggtitle("Figure C.6: Conditional Proportions of\nAttendance vs. Variety Importance") +
  xlab("Importance of Variety") + ylab("Proportions") +
  theme(, plot.title = element_text(size = 7), axis.title = element_text(size = 7),
        axis.text = element_text(size = 5), aspect.ratio = 1, legend.position = "none")

##quantity
#ggplot - make proportion table based on conditional probabilities - conitional probabilities graph
proportions_quantity <- proportions(table9, 1)

#make condiitonal probabilities graph
proportions9.df <- data.frame(expand.grid(rownames(proportions_quantity),
                                          colnames(proportions_quantity)),
                              c(proportions_quantity))
colnames(proportions9.df) <- c("amount_importance", "attendance", "Proportion")
p_quantity <- ggplot(proportions9.df, aes(x = amount_importance, y = Proportion, fill = attendance)) +
  geom_bar(stat = "identity") +
  scale_fill_discrete(name = "Attendance Frequency",
                      labels = c("Weekly", "Bi-Weekly", "Occasionally", "Once"))  +
  ggtitle("Figure C.7: Conditional Proportions of\nAttendance vs. Quantity Importance") +
  xlab("Importance of Quantity") + ylab("Proportions") +
  theme(plot.title = element_text(size = 7), axis.title = element_text(size = 7),
        axis.text = element_text(size = 5), aspect.ratio = 1, legend.position = "none")
```

```r
##quality
#ggplot - make proportion table based on conditional probabilities - conitional probabilities graph
proportions_quality <- proportions(table8, 1)

#make condiitonal probabilities graph
proportions8.df <- data.frame(expand.grid(rownames(proportions_quality),
                                          colnames(proportions_quality)),
                              c(proportions_quality))
colnames(proportions8.df) <- c("quality_importance", "attendance", "Proportion")

p_quality <- ggplot(proportions8.df, aes(x = quality_importance, y = Proportion, fill = attendance)) +
  geom_bar(stat = "identity") +
  scale_fill_discrete(name = "Attendance Frequency",
                      labels = c("Weekly", "Bi-Weekly", "Occasionally", "Once"))  +
  ggtitle("Figure C.8: Conditional Proportions of\nAttendance vs. Quality Importance") +
  xlab("Importance of Quality") + ylab("Proportions") +
  theme(plot.title = element_text(size = 7), axis.title = element_text(size = 7),
        axis.text = element_text(size = 5), aspect.ratio = 1, legend.position = "none")

library("gridExtra")
grid.arrange(c_quality, c_variety, c_quantity, p_quality, p_variety, p_quantity, ncol=3, nrow=2)
```

Figure C.3: Attendance vs. Quality Importance

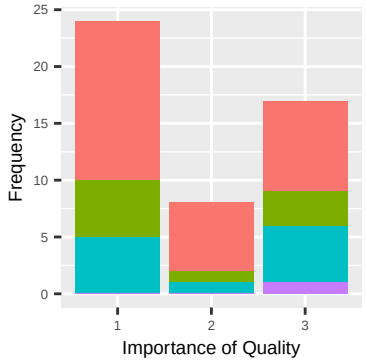


Figure C.4: Attendance vs. Variety Importance

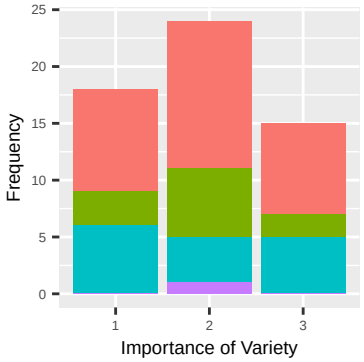


Figure C.5: Attendance vs. Quantity Importance

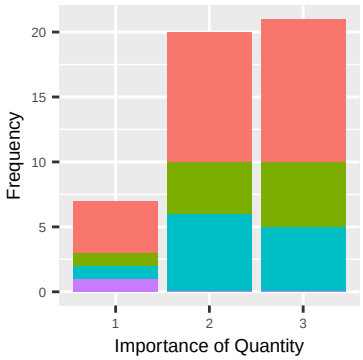


Figure C.8: Conditional Proportions of Attendance vs. Quality Importance

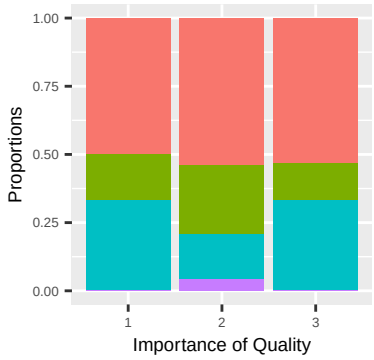


Figure C.6: Conditional Proportions of Attendance vs. Variety Importance

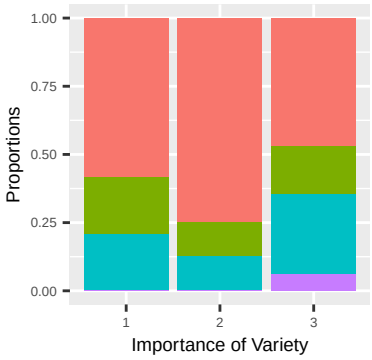
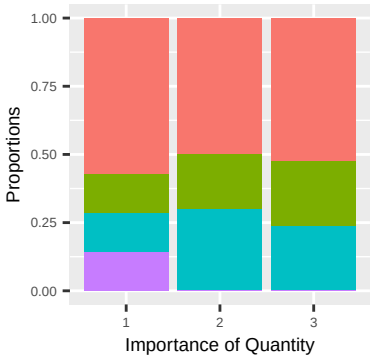


Figure C.7: Conditional Proportions of Attendance vs. Quantity Importance




```r
##Table C.9
#find row sums for total counts among each feature
#calculate what the total rankings were among each category
```

```
rowSums(table9)
```

```
##  1  2  3
##  7 20 21
```

```
rowSums(table8)
```

```
##  1  2  3
## 18 24 15
```

```
rowSums(table7)
```

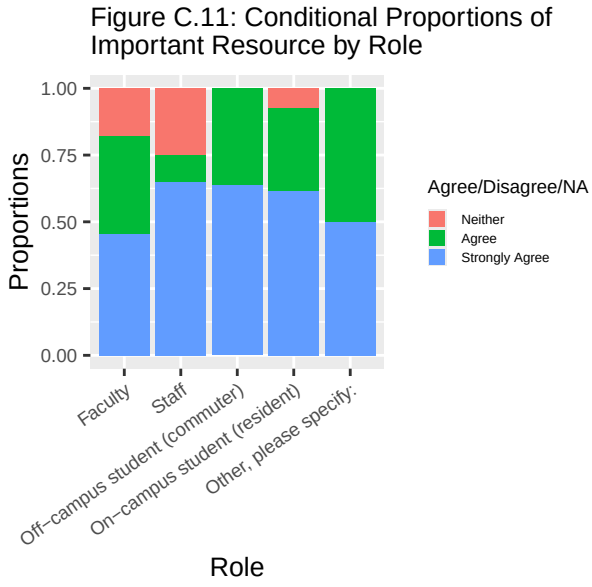
```
##  1  2  3
## 24  8 17
```

Table C.9: Frequency of Perceived Importance of Quality, Variety, and Quantity

Perceived Importance	1	2	3
Quality Importance	24	8	17
Variety Importance	18	24	15
Quantity Importance	7	20	21

```
##Figures C.10-C.11
#how do people rank that the market is an important resource for them and role
resource <- survey$Q22_9
role_all <- survey$Q1_text
role_all <- factor(role_all, levels = c("Faculty",  "Staff", "Off-campus student (commuter)",
                                        "On-campus student (resident)", "Other, please specify:"))

df10 <- data.frame(role_all, resource)
no_NA_df10 <- na.omit(df10)

#make a contingency table of df
table10 <- table(no_NA_df10)

#this helps us to make the ggplot graph; get the counts, row/col names
table.df10 <- data.frame(expand.grid(rownames(table10), colnames(table10)), c(table10))
colnames(table.df10) <- c("Role", "resource", "count")

#ggplot - make proportions graph based on counts
c_resource <- ggplot(table.df10, aes(x = Role, y = count, fill = resource)) +
  geom_bar(stat = "identity") + scale_fill_discrete(name = "Agree/Disagree/NA",
                                                    labels = c("1" = "Strongly Disagree",
                                                               "2" = "Disagree", "3" = "Neither",
                                                               "4" = "Agree", "5" = "Strongly Agree")) +
  ggtitle("Figure C.10: Important Resource by Role") + xlab("Role") +
  ylab("Frequency") + scale_x_discrete(guide = guide_axis(angle = 35)) +
  theme(legend.text = element_text(size = 5), plot.title = element_text(size = 10),
        legend.title = element_text(size = 7), axis.title = element_text(size = 10),
```

```
        axis.text = element_text(size = 7), legend.key.size = unit(0.25, 'cm'), aspect.ratio = 1)

#ggplot - make proportion table based on conditional probabilities - conitional probabilities graph
proportions <- proportions(table10, 1)

#make conditional probabilities graph
proportions.df <- data.frame(expand.grid(rownames(proportions), colnames(proportions)),
                             c(proportions))
colnames(proportions.df) <- c("Role", "resource", "Proportion")
p_resource <- ggplot(proportions.df, aes(x = Role, y = Proportion, fill = resource)) +
  geom_bar(stat = "identity") + scale_fill_discrete(name = "Agree/Disagree/NA",
                                                    labels = c("1" = "Strongly Disagree",
                                                               "2" = "Disagree", "3" = "Neither",
                                                               "4" = "Agree", "5" = "Strongly Agree")) +
  ggtitle("Figure C.11: Conditional Proportions of\nImportant Resource by Role") +
 xlab("Role") + ylab("Proportions") + scale_x_discrete(guide = guide_axis(angle = 35)) +
  theme(legend.text = element_text(size = 5), plot.title = element_text(size = 10),
        legend.title = element_text(size = 7), axis.title = element_text(size = 10),
        axis.text = element_text(size = 7), legend.key.size = unit(0.25, 'cm'), aspect.ratio = 1)

grid.arrange(c_resource, p_resource, ncol = 2)
```




```
##Figures C.12-C.13
#how do people rank that the market is an important resource for them and role proportions
```

```r
df11 <- data.frame(attendance_freq, resource)
no_NA_df11 <- na.omit(df11)

#make a contingency table of df
table11 <- table(no_NA_df11)


#this helps us to make the ggplot graph; get the counts, row/col names
table.df11 <- data.frame(expand.grid(rownames(table11), colnames(table11)), c(table11))
colnames(table.df11) <- c("attendance", "resource", "count")

#ggplot - make proportions graph based on counts
c_resource <- ggplot(table.df11, aes(x = attendance, y = count, fill = resource)) +
  geom_bar(stat = "identity") + scale_fill_discrete(name = "Agree/Disagree/NA",
                                                    labels = c("1" = "Strongly Disagree",
                                                               "2" = "Disagree", "3" = "Neither",
                                                               "4" = "Agree", "5" = "Strongly Agree")) +
  ggtitle("Figure C.12: Important Resource by \nAttendance Frequency") +
 xlab("Role") + ylab("Frequency") + scale_x_discrete(labels= c("1" = "Weekly", "2" = "Bi-Weekly",
                                                               "4" = "Occasionally", "5" = "Once"),
                                                     guide = guide_axis(angle = 40)) +
  theme(legend.text = element_text(size = 5), plot.title = element_text(size = 10),
        legend.title = element_text(size = 7), axis.title = element_text(size = 10),
        axis.text = element_text(size = 7), legend.key.size = unit(0.25, 'cm'), aspect.ratio = 1)

#ggplot - make proportion table based on conditional probabilities - conditional probabilities graph
proportions11 <- proportions(table11, 1)

#make conditional probabilities graph
proportions11.df <- data.frame(expand.grid(rownames(proportions11), colnames(proportions11)),
                               c(proportions11))
colnames(proportions11.df) <- c("attendance", "resource", "Proportion")
p_resource <- ggplot(proportions11.df, aes(x = attendance, y = Proportion, fill = resource)) +
  geom_bar(stat = "identity") + scale_fill_discrete(name = "Agree/Disagree/NA",
                                                    labels = c("1" = "Strongly Disagree",
                                                               "2" = "Disagree", "3" = "Neither",
                                                               "4" = "Agree", "5" = "Strongly Agree")) +
  ggtitle("Figure C.13: Conditional Proportions of \nImportant Resource by \nAttendance Frequency") +
 xlab("Role") + ylab("Proportions") + scale_x_discrete(labels= c("1" = "Weekly", "2" = "Bi-Weekly",
                                                                 "4" = "Occasionally", "5" = "Once"),
                                                       guide = guide_axis(angle = 40)) +
  theme(legend.text = element_text(size = 5), plot.title = element_text(size = 10),
        legend.title = element_text(size = 7), axis.title = element_text(size = 10),
        axis.text = element_text(size = 7), legend.key.size = unit(0.25, 'cm'), aspect.ratio = 1)

grid.arrange(c_resource, p_resource, ncol = 2)
```

Figure C.12: Important Resource by Attendance Frequency

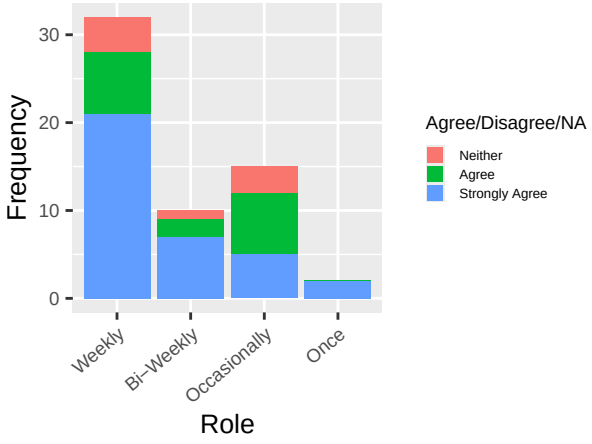
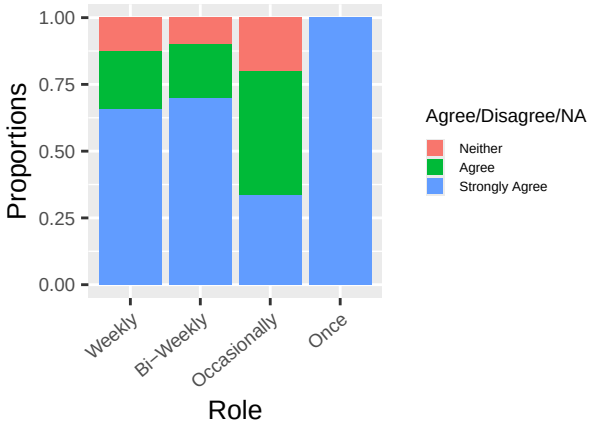


Figure C.13: Conditional Proportions of Important Resource by Attendance Frequency




```
## #find expected counts - must be at least 5 in each category for the test to be deemed appropriate
chisq.test(table11)$expected
```

```
## Warning in chisq.test(table11): Chi-squared approximation may be incorrect
```

```
##                resource
## attendance_freq         3         4         5
##               1 4.3389831 8.6779661 18.983051
##               2 1.3559322 2.7118644  5.932203
##               4 2.0338983 4.0677966  8.898305
##               5 0.2711864 0.5423729  1.186441
```

```
#proceed with permutation resampling
chisq.test(table11, simulate.p.value = TRUE, B = 10^5)
```

```
##
##  Pearson's Chi-squared test with simulated p-value (based on 1e+05
##  replicates)
##
## data:  table11
## X-squared = 6.6895, df = NA, p-value = 0.3441
```

Table 4: Table C.14: Expected Counts of Chi-Square Test

Expected Counts	NA	Agree	Strongly Agree
Weekly	4.3389	8.6779	18.9831
Bi-Weekly	1.3559	2.7119	5.9322
Occasionally	2.0339	4.0678	8.8983
Once	0.2712	0.5424	1.1864