Community's Personal Relationship v.s. Overall Relationship towards Campus

Project Team 4:
Weishen Zou
PJ Mu
Tiantian Qiu
Xiaonan Xu

Miami University
Department of Statistics
311 Upham Hall
100 Bishop Circle
Oxford, OH 45056

STA475
March 10th, 2018

Contents

Summary

It is desired to analyze the town-gown relationships between Miami University and local community. The problems are to find the determinants within the own personal relationship with various campus groups; find answer toward the overall alcohol misuse question. Many factors may influence how personal perspective from the community towards the relationships with board members, students, faculty, staff and administration. So it is desired to narrow down which factors are predictive of a personal perspective. In addition, the gap between personal perspective and the community perspective as a whole is considered in order to analyze the relationship situation affected by different perspective positions. The purpose of this study is to use multiple linear regression and relevant importance of features to find out the ways to improve town-gown relationships. Personal perspective and whole community perspective are similar for the group of board members but different for the groups of students, faculty, staff and administration. In addition, there are different determinants that influence community members' feeling towards different campus group and many of them are related to student drinking.

Introduction

The data being used in this report is collected from the Optimal College Town Assessment. OCTA is a survey tool designed to assess the quality of the relationship between a University(Miami) and its host community(Oxford).There are several objective factors that may be related to personal perspective and community perspective as a whole towards the board members and students group, the faculty group and the staff and administration group. The aim of the investigation is to use statistical analytics to figure out possible solutions of improving the town-gown relationships. The data was started to collect from all Oxford community and Miami community from October 3rd, 2017. The survey is conducted due to the high concern on the alcohol misuse in Oxford with regards to broader community and community-campus reaction. There are several variables in the survey, including the attitudes towards the relationships of each groups, the attitudes of potential problems, the typical party nights and so on. People's attitudes is scaled from 1 to 5. "1" mostly indicates the most positive attitudes and "5" has the opposite meaning.

Methods

1.Analysis

200 participants of OCTA survey who did not complete the all questions in the survey are considered missing values and not included in the analysis. The computations are completed in R (R Core Team, 2017) with the packages readxl (Wickham and Bryan, 2017), MASS (Braun, 2015), relaimpo (Grömping, 2006), tidyverse (Wickham, 2017), and ggplot2 (Wickham, 2009).

1.1 Preliminary analysis

There are 2789 observations in the data with 113 variables in the overall dataset. Due to the consideration of concentration on the community population and missing values, 693 observations and 71 variables are taken into account in the final.

Table 1 shows that the means and variance between personal perspective and community perspective as a whole towards different groups. There are 9965 observations for personal perspective data collected and 8700 for community perspective data collected. It is easy to find that personal perspective's ratings are usually larger than those of community perspective as a whole towards all the groups of people in both means and variances. This means people have more positive attitudes towards the relationships between all groups of people with regards to the whole community than to themselves. The ratings they give from both of the two perspectives are mostly from neutral to slightly positive towards the student, faculty, staff and administration groups but slightly negative to neutral towards the board members group.

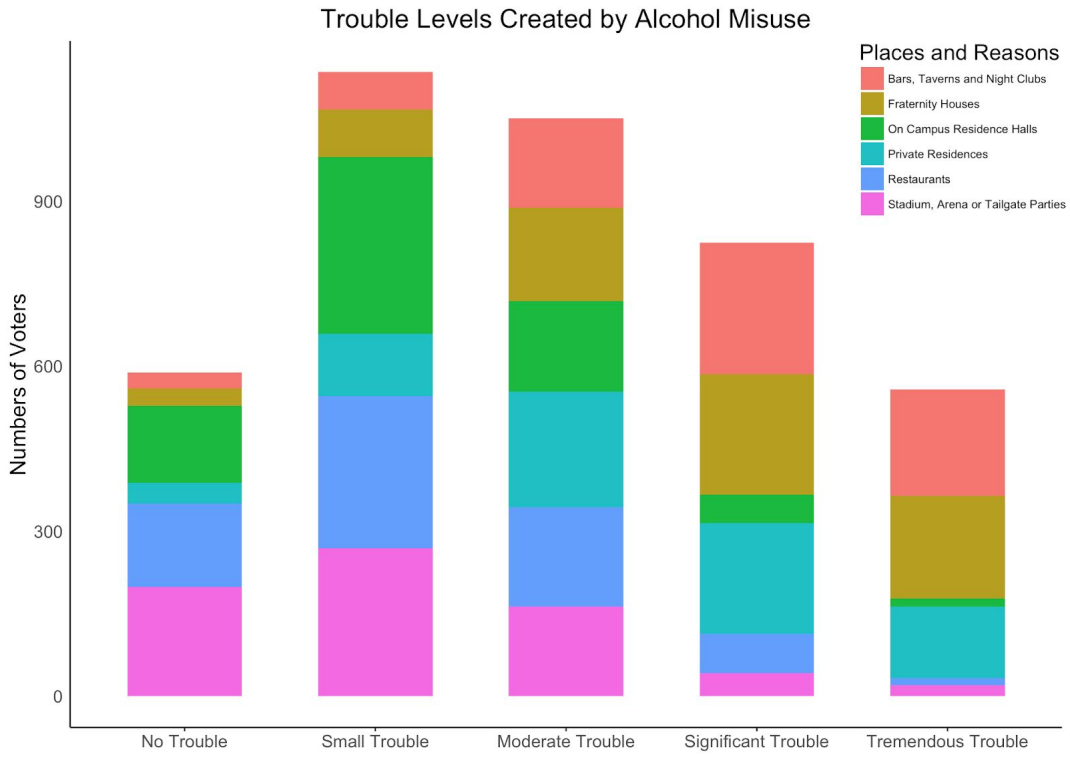
Difference between community members' personal relationship with people on campus and their relationship as a whole based on sample data from the OCTA survey doesn't tell the whole story. As always, we want to draw conclusion concerning the overall population in Oxford, not just from the sample.

Table 1. Community member's rating gaps between personal perspective and community perspective towards different groups

Groups of people	Perspective	N	Means	SEs
Students	Personal	2567	3.70	1.42
	Whole community	1821	2.63	1.41
Faculty	Personal	2763	3.99	0.92
	Whole community	2475	3.57	0.83
Staff/ Administration	Personal	2619	3.77	1.14
	Whole community	2388	3.45	0.91
Board members	Personal	2021	2.92	0.87
	Whole community	2016	2.91	0.92

Figure 2 shows that the how the trouble levels of alcohol (student partying) made by different places or reasons that the community members consider. It is obvious to find that the number of people who think the alcohol misuse problems is only small trouble or moderate trouble in Oxford is highest compared to others. On campus residence halls is the place that make least trouble to the community. However, community members regard bars, taverns, night clubs and fraternity houses as the worst trouble makers due to the alcohol misuse.

Figure 2. Community members' considerations on trouble effect of students' partying and alcohol misuse



1.2 Developmental Methodology

According to the community survey launched by OCTA, college members are divided into four groups: students, faculty and staff, administrators, and board members. Intuitively each group will have its own distinct determinants, though some of which could be the same across groups. To start with, four linear regression models are fit. The first one is to measure personal relationship with students, the second one is to measure personal relationship with faculties and staff, the third one is to measure personal relationship with administrators, and the last one is to measure the personal relationship with board members. Potential determinants for each model varies.

All potential survey questions from the OCTA community perspective survey that could influence community member's personal relationship with one campus group are identified, and then a linear regression model is constructed using these as predictor variables. The significance of each predictor is examined and the multicollinearity is checked. To further tune the model, stepwise model selection method is used. Once the linear model is tuned, the relative importance of each factor is plotted by calculating several relative importance metrics for this linear model. Such procedure is repeated for the other three campus groups. In building the model, there is no need to check assumptions of each model fit because the goal is to tune model and look for important variables that explain the response.

To understand if single member in the community perceives issues or problems same as the other members in the community, it is necessary to compare the attitude between the personal rating, or feeling and their rating when it comes to whole community. The gap or agreement between the personal perspective and community perspective as a whole is worth analyzing due to conclusion of whether the community members themselves recognize the community problems as a whole. Further inference about the difference between the "personal" relationship and relationship "as a whole" is drawn using paired t-test methods.

1.3 Full Scale Analysis

1.3.1 Community personal relationship with campus students

By carefully examining community perspective survey, all questions that could potentially result in affecting the personal relationship with students are 28 questions in list 1. Question 1,2,3, and 4 are about student party, which are undoubtedly important factors that affects community's "person relationship" towards campus students. For question 5, all the miscellaneous concerns about campus could be potential determinants of community's personal relationship with campus students.

List 1. All questions that potentially result in affecting the personal relationships with students

1) (from 1 no trouble to 5 tremendous amount of trouble) how much trouble is created in the community as a result of students "partying" in
    a) Private residence,
    b) Fraternity houses,
    c) Bars, taverns, and nightclubs,
    d) Restaurants,
    e) Residence Halls,
    f) Stadium, Arena or Tailgate Parties.
2) (from 1 not occur to 5 occur all the time) how often these problems seem to occur
    a) Excessive noise
    b) Trash/Littering
    c) Violence/fights/riots
    d) Public Urination
    e) Sexual Assault
    f) Vandalism
    g) Traffic accidents/DUI
    h) Injuries due to falls or crowds
3) (from 1 no strain to 5 many strains) To what extent has student partying put strain on campus-community ("town/gown") relationships?
4) (from 1 no impact to 5 many impact)To what extent has student partying had a financial impact on your community?
5) (from 1 strongly disagree to 5 strongly agree)Miscellaneous concerns and problem about campus
    a) There are not enough events on campus that interest me.
    b) Coursework offered on campus is too expensive.
    c) The campus seems difficult to get to.
    d) Not enough classes and degrees are offered on the campus.
    e) Campus representatives do not readily share their expertise with the community.
    f) Campus representatives are not well connected to the local school districts.
    g) It's hard to know what's happening on campus.
    h) People from the campus don't live in our community.
    i) The campus does not contribute as much as it could to the local economy.
    j) People from the campus do not get involved in volunteer efforts in the community.
    k) Campus representatives are not very visible at community events.
    l) Student misbehavior in the community is a real problem.

With all the potential predictors, A full linear regression model is built and it is named PRstudent.mod. The Multicollinearity of the full model, which is the excessive correlation among explanatory variables, is checked by using VIF function. All predictors are greater than 1 and lower than 5 are considered correlated, Hence there is no correlations between each independent variables.

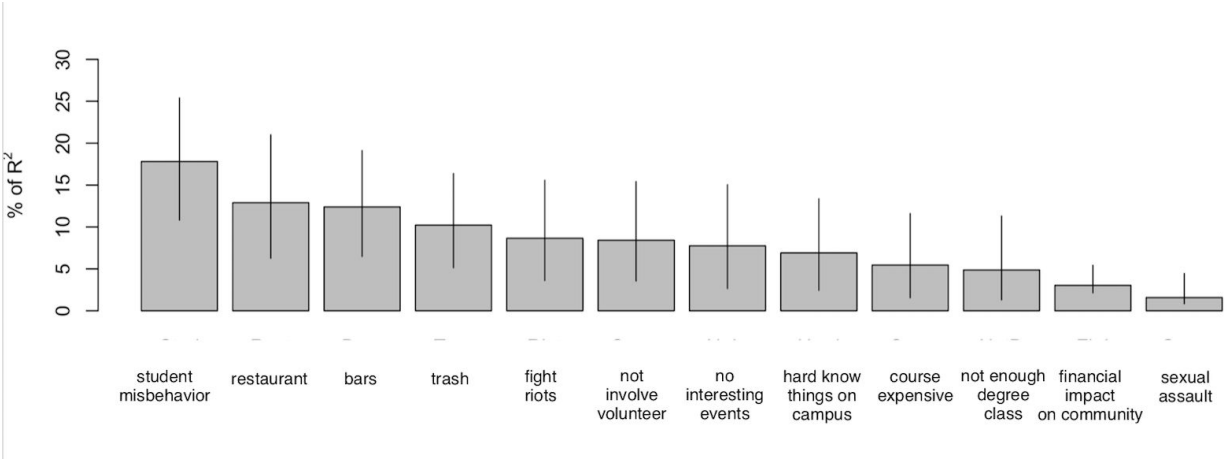
To erase the insignificant variables and further tune the model. The stepwise AIC from both sides is used to conduct the model selection. The final model contains only 12 survey questions that determine the community's "personal relationship" with campus students in table 2.

Table 2. Determinants of Personal relationship with students

variables	estimate	std.error	statistic	p.value
Bars, Night clubs	-0.12	0.05	-2.48	0.01
Restaurants	-0.15	0.05	-3.31	< 0.001
Trash Littering	-0.16	0.05	-3.08	< 0.001
Riots	-0.11	0.06	-2.04	0.04
SexualAst	0.10	0.05	2.20	0.03
FinImpact	0.13	0.04	3.10	< 0.001
No Interesting Events	-0.08	0.03	-2.43	0.02
Course Expensive	-0.07	0.03	-2.07	0.04
NotDegreeOnTheCampus	-0.04	0.03	-1.54	0.12
Hard To Know What Happening On Campus	-0.07	0.03	-2.03	0.04
Campus Not Involved In Volunteer	-0.07	0.03	-2.45	0.01
Student Misbehavior	-0.18	0.04	-4.10	< 0.001

After the model selection, It is important to rank these determinants from most important to least important. The relaimpo package (Ulrike Grömping, 2006) is used and "lmg" method is used to measure the relative importance of each feature. This method partition the R^2 into different features to display important of each determinant. To visualize the relative importance, Bootstrapping these metrics can help display the relative importance in figure 3.

Figure 3. Relative importance for determinants of Personal relationship with students



1.3.2 Community personal relationship with campus faculty

The same procedure is used to explore personal relationship with campus faculty. In total, there are 15 survey questions, which could be potential determinants, to fit the model.

List 2. The potential determinant of personal relationships with campus faculty

1. (from 1 strongly disagree to 5 strongly agree)Miscellaneous concerns and problem about campus
    a. There are not enough events on campus that interest me.
    b. Coursework offered on campus is too expensive.
    c. The campus seems difficult to get to.
    d. Not enough classes and degrees are offered on the campus.
    e. Campus representatives do not readily share their expertise with the community.
    f. Campus representatives are not well connected to the local school districts.
    g. It's hard to know what's happening on campus.
    h. People from the campus don't live in our community.
    i. The campus does not contribute as much as it could to the local economy.
    j. People from the campus do not get involved in volunteer efforts in the community.
    k. Campus representatives are not very visible at community events.
    l. Student misbehavior in the community is a real problem.

2. (from 1 never to 5 all the time)How often does following groups visit shops and stores near campus
    a. Campus staff
    b. Faculty

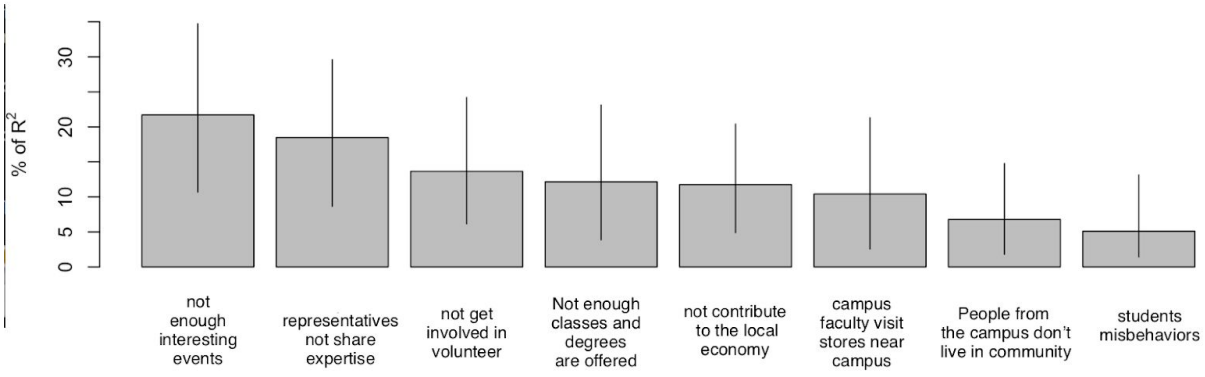
A full linear regression model is built and it is named PRfaculty.mod. The Multicollinearity of the full model is checked using VIF function and backwards stepwise AIC is used to conduct the model selection. The final model contains only 8 survey questions that determine the community's "personal relationship" with campus faculty in table 3.

Table 3. Determinants of Personal relationships with faculty

variables	estimate	std.error	statistic	p.value
No Interesting Events	-0.12	0.02	-4.28	< 0.001
NotDegreeOnTheCampus	-0.06	0.02	-2.60	<0.001
Represent Not Share Expertise	-0.08	0.02	-2.85	<0.001
People From The Campus Not Live In Community	-0.04	0.02	-1.57	0.11
The Campus Not Contribute Economy	-0.04	0.02	-1.86	0.06
Campus Not Involved In Community	-0.05	0.02	-1.83	0.06
Student Misbeh	-0.0	0.03	-1.73	0.08
Faculty Store Visit	0.11	0.03	3.11	0.001

Same as the previous model, the relative importance of each variable is measured by the relaimpo package and and visualized in figure 4.

Figure 4. Relative importance for determinants of Personal relationship with faculty



1.3.3 Community personal relationship with campus staff and administrators

The same procedure is used to explore personal relationship with campus faculty. The potential determinant list of survey questions are the same as the list 2 because campus staff and campus faculty has many in common.

List 3. The potential determinant of personal staff and administrators

1. (from 1 strongly disagree to 5 strongly agree)Miscellaneous concerns and problem about campus
    a. There are not enough events on campus that interest me.
    b. Coursework offered on campus is too expensive.
    c. The campus seems difficult to get to.
    d. Not enough classes and degrees are offered on the campus.
    e. Campus representatives do not readily share their expertise with the community.
    f. Campus representatives are not well connected to the local school districts.
    g. It's hard to know what's happening on campus.
    h. People from the campus don't live in our community.
    i. The campus does not contribute as much as it could to the local economy.
    j. People from the campus do not get involved in volunteer efforts in the community.
    k. Campus representatives are not very visible at community events.
    l. Student misbehavior in the community is a real problem.
2. (from 1 never to 5 all the time)How often does following groups visit shops and stores near campus
    a. Campus staff
    b. Faculty

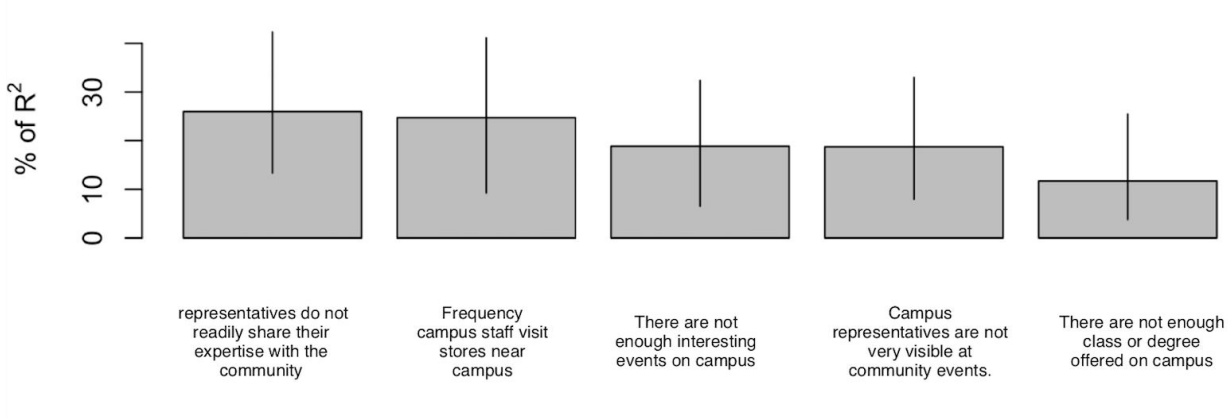
A full linear regression model is built and it is named PRstaff.mod. The Multicollinearity of the full model is checked using VIF function and backwards stepwise AIC is used to conduct the model selection. The final model contains only 5 survey questions that determine the community's "personal relationship" with campus faculty in table 4.

Table 4. Determinants of Personal relationships with staff and administrators

variable	estimate	std.error	statistic	p.value
No Interesting Events	-0.09	0.03	-2.94	0.001
Not Degree Offered	-0.05	0.02	-2.16	0.03
Represent Not Share Expertise	-0.09	0.03	-3.08	0.002
Represent Not Visible At Community Events	-0.07	0.03	-2.44	0.01
Campus Staff Store Visit	0.16	0.03	4.39	<0.001

Same as the previous model, the relative importance of each variable is measured by the relaimpo package and the plotted in figure 5. The determinants for personal relationship with staff and administrators thus are ordered from the most important to least important.

Figure 5. Relative importance for determinants of Personal relationship with staff and administrators



1.3.4 Community personal relationship with campus board members

For the community's personal relationship with campus board members, potential determinants could be miscellaneous concerns that arise in town-gown relationship. Moreover, burden caused by student in the community could way that influence community's personal relationship with board members. All the collect potential determinants are in list 4.

List 4. potential determinants lists of personal relationships with board members

1. (from 1 no strain to 5 many strains) To what extent has student partying put strain on campus-community ("town/gown") relationships?
2. (from 1 no impact to 5 many impact)To what extent has student partying had a financial impact on your community?
3. (from 1 strongly disagree to 5 strongly agree)Miscellaneous concerns and problem about campus
    a. There are not enough events on campus that interest me.
    b. Coursework offered on campus is too expensive.
    c. The campus seems difficult to get to.
    d. Not enough classes and degrees are offered on the campus.
    e. Campus representatives do not readily share their expertise with the community.
    f. Campus representatives are not well connected to the local school districts.
    g. It's hard to know what's happening on campus.
    h. People from the campus don't live in our community.
    i. The campus does not contribute as much as it could to the local economy.
    j. People from the campus do not get involved in volunteer efforts in the community.
    k. Campus representatives are not very visible at community events.
    l. Student misbehavior in the community is a real problem.

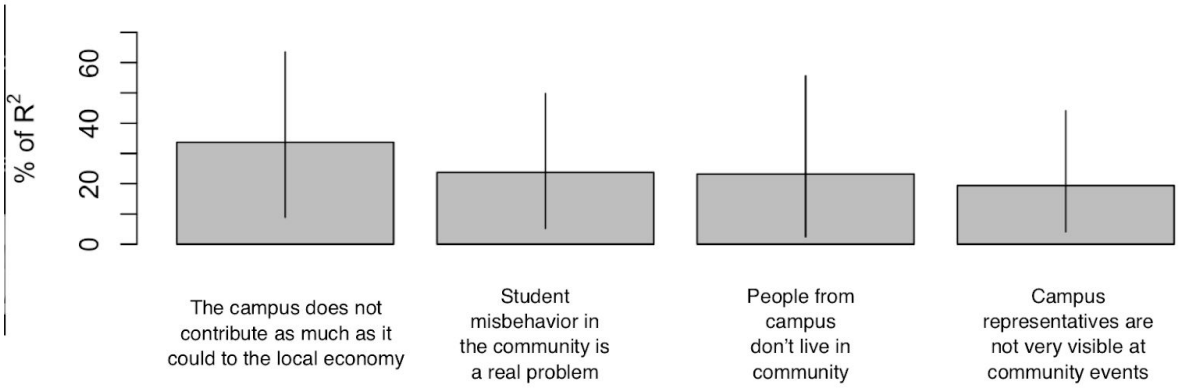
A full linear regression model is built and it is named PRboard.mod. The Multicollinearity of the full model is checked using VIF function and backwards stepwise AIC is used to conduct the model selection. The final model contains only 5 survey questions that determine the community's "personal relationship" with campus faculty in table 5.

Table 5. Determinants of Personal relationship with campus board members

variable	estimate	std.error	statistic	p.value
People From The Campus Not Live In Community	-0.04	0.02	-1.66	0.09
The Campus Not Contribute Economy	-0.06	0.02	-2.49	0.01
Represent Not Visible	-0.05	0.02	-1.94	0.05
Student Misbehavior	-0.065	0.02	-2.13	0.02

Same as the previous model, the relative importance of each variable is measured by the relaimpo package and the plotted as figure 6. The determinants for personal relationship with campus board members thus are ordered from the most important to least important.

Figure 6. Relative importance for determinants of Personal relationship with board members



1.3.5 Community "personal relationship" and relationship "as a whole"

To fully analyze the difference between community's personal perception and perception as a whole towards campus issues, paired t test is used. It is computed with the comparison and difference between the population true mean of personal perspective and community perspective as a whole. The assumption is there no relationships between the means of personal perspective and community perspective as a whole. P-value is considered as an important indicator to judge how the gap of the perspective between the two groups with regards to students, faculty, board members and staff and administration. If p-value is smaller than 0.05 (statistical alpha level), the assumptions fail and the differences between the mean of the two groups towards specific groups need to be significantly considered otherwise, it should be ignored.

Result and Conclusion

From the full scale analysis, important survey questions that determine community members' personal relationship with students, faculty, staff / administrators, and board members are displayed in table 6, table 7, table 8, and table 9. Determinants of each table are ranked from most important to least important and the direction of each determinant are labeled in survey questions in the second column. From table 6, for example, if community members' rating on "Student misbehavior in the community is a real problem" is very high, community members' personal relationship with be low. Column three stands for the partition R^2 of each determinant. It can be also considered as "importance proportion". That is, how much each determinant explain community's personal relationship with certain campus group.

Table 6. Determinants of Personal relationship with students (From most important to least important)

Determinants	R^2 (importance proportion)	Determinants	R^2 (importance proportion)
Student misbehavior	0.18	There is no interesting events on campus	0.08
Trouble created in the community as a result of students "partying" in Restaurants	0.13	Hard to know what's happening on campus	0.07
Trouble created in the community as a result of students "partying" in Bars, taverns, and nightclubs	0.12	Courses offered on campus are expensive	0.05
Trash/Littering caused by student partying	0.10	Not enough degree or class offered on campus	0.05
Violent fights caused by student partying	0.09	To what extent has student partying had a financial impact on your community?	0.03
People from the campus do not get involved in volunteer efforts in the community.	0.08	Sexual assault caused by student partying	0.02

Table 7. Determinants of Personal relationship with faculty (From most important to least important)

Determinants	R^2 (importance proportion)
Not enough interesting events on campus	0.22
Campus representatives do not readily share their expertise with the community.	0.18
People from the campus do not get involved in volunteer efforts in the community	0.14
Not enough classes and degrees are offered on the campus	0.12
The campus does not contribute as much as it could to the local economy	0.12
Frequency campus faculty visit stores near campus	0.10
People from the campus don't live in community	0.07
Student misbehaviors in community	0.05

Table 8. Determinants of Personal relationship with staff and administrators (From most important to least important)

Determinants	R^2 (importance proportion)
Campus representatives do not readily share their expertise with the community.	0.26
Frequency campus staff visit stores near campus	0.25
There are not enough interesting events on campus	0.19
Campus representatives are not very visible at community events.	0.19
There are not enough class or degree offered on campus	0.11

Table 9. Determinants of Personal relationship with board members (From most important to least important)

Determinants	R^2 (importance proportion)
The campus does not contribute as much as it could to the local economy	0.34
Student misbehavior in the community is a real problem	0.23
Campus representatives are not very visible at community events	0.24
People from campus don't live in community	0.19

Table 10. The gaps between personal perspective and whole community perspective

The group that two aspects towards	T test	Degree of freedom	P-value	CI	Conclusion
Students	-24.78	692	< 0.0001	(-1.16, -0.99)	Personal - (0.99, 1.66) = Whole community
Faculty	-2.58	692	0.0101	(-0.23, -0.03)	Personal - (0.03, 0.23) = Whole community
Staff and administration	-9.58	692	< 0.0001	(-0.40, -0.27)	Personal - (0.27, 0.40) = Whole community
Aboard members	-0.23	692	0.8145	(-0.07, 0.05)	Personal ≈ Whole community

From the four t tests, only the p-value for aboard members is larger than alpha level (0.05). The assumptions that personal perspective and community perspective as a whole is the same is rejected for students, faculty, staff and administration group. For the three groups, the personal perspective rating levels are larger than community perspective as a whole. The Confidence intervals are the different ranges between personal perspective and whole community perspective. This indicates that community members don't recognize the community issues have about the campus.There is 95 % confidence that the personal perspective ratings is from 0.99 to 1.16 larger than the whole community perspective ratings towards the student group. Also, there is 95 % confidence that the personal perspective ratings is from 0.03 to 0.23 larger than the whole community perspective ratings towards the group of the faculty. In addition, there is 95 % confidence that the personal perspective ratings is from 0.27 to 0.4 larger than the whole community perspective ratings towards staff and administration group. However, according to the aboard members, the personal and community hold the similar ideas or attitudes and the community members recognize the community issues towards this group.

General Discussion

The goal of the report is achieved through various aspects of analysis and the result of the analysis suggests that the student misbehaviour, not enough interesting events on campus, the represent not share expertise, and the campus not contribute economy in terms of the community's "personal relationship" with students, faculty, staff, and board members might be the most important factors that cause negative impact toward the relations. In additions, along with the part of community vs student, another important factor is the troubles created from students partying. Both results indicates that the community might have more concerns about the student behaviour. For the partying aspect, it might implies that the alcohol misuse is a potential problem that cause the negative impact on town-gown relations. To control and improve the alcohol use environment could be a appropriate way to reduce the concerns and frustration from the community. Besides, it is worth to suggest that future study could focusing more on the student misbehavior, the interaction between community and staff, faculty. Besides, the relations of community vs. the staff and administration reveal a fact shows campus representatives do not readily share their expertise with the community. A suggestion is the campus could communicate more with the community to gain better understanding and appreciation for each other, also create more corporations for both sides. Another suggestion for the community vs. faculty that might help to improve the relations is that the campus could host more events on community and communicate with each other. About the aspect of campus not contribute economy to the community; the suggestion is consider to care more on local business around Oxford area and contribute more economy to the community from the campus.

Since the data is obtained from the survey, the shortcoming of the data might exist as participants lost patience and start filling the questions randomly or even the participants end/quit the survey in half way through. This might create skewness, unexpected errors, incompleteness in the dataset generated by the survey.

References

Multiple linear regression: Lai, T. L_etc, Herbert Robbins, and C. Zi Wei. "Strong consistency of least squares estimates in multiple regression II." Journal of Multivariate Analysis 9.3 (1979): 343-361.

Robk@statmethods.net, R. K. (n.d.). Multiple (Linear) Regression. Retrieved March 11, 2018, from https://www.statmethods.net/stats/regression.html

Hadley Wickham and Jennifer Bryan (2017). readxl: Read Excel Files. R package version 1.0.0. https://CRAN.R-project.org/package=readxl

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2009.

Venables, W. N. & Ripley, B. D. (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0

Ulrike Grömping (2006). Relative Importance for Linear Regression in R: The Package relaimpo. Journal of Statistical Software, 17(1), 1--27.

Hadley Wickham (2017). tidyverse: Easily Install and Load the 'Tidyverse'. R package version 1.2.1. https://CRAN.R-project.org/package=tidyverse

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Appendix A

Models

```
library(car)
community <- read_excel("~/Downloads/Target Data.xlsx")
community<-na.omit(community)
PRstudent.mod<-lm(PR_Students~ .-PR_Faculty -
PR_StaffAdministration-PR_BoardMembers,data =community )

ctable <- coef(summary(PRstudent.mod))
ctable<-as.data.frame(ctable)
ctable<-ctable %>%tibble::rownames_to_column(var = "Variables")
%>%filter( `Pr(>|t|)`<0.05)
ctable<-ctable[-1,c("Variables","Pr(>|t|)")]
Ctable
vif(PRstudent.mod)

step <- stepAIC(PRstudent.mod, direction="both")
step$anova
```

##fit multiple linear regression with student aspect

```
library(relaimpo)
final.student<-lm(Prstudents ~ BarsNightc + Restaurants +
TrashLittering + Riots +
                  SexualAst + FinImpact + NoInterestingEvents +
CourseExpensive +
                  NotDegreeOnTheCampus +
HardToKnowWhatSHappeningOnCampus +
                  CampusNotInvolvedInCommunity +
StudentMisbehaviorInTheCommunity,
                  data = students)


calc.relimp(final.student,type=c("lmg"),
            rela=TRUE)
```

```
# Bootstrap Measures of Relative Importance (1000 samples)
boot <- boot.relimp(final.student , b = 1000, type = c("lmg"),
rank = TRUE, diff = TRUE, rela = TRUE)
booteval.relimp(boot)
plot(booteval.relimp(boot,sort=TRUE))
```

model faculty

```
community_fac <- read_excel("~/Downloads/Target Data.xlsx")
community_fac<-na.omit(community_stu)

PRfaculty.mod<-lm(PR_StaffAdministration~ .-PR_Students -
PR_StaffAdministration-PR_BoardMembers,data =community_fac )

ctable <- coef(summary(PRfaculty.mod))
ctable<-as.data.frame(ctable)
ctable<-ctable %>%tibble::rownames_to_column(var = "Variables")
%>%filter( `Pr(>|t|)`<0.05)
ctable<-ctable[-1,c("Variables","Pr(>|t|)")]
ctable
vv<-vif(PRfaculty.mod )

step <- stepAIC(PRfaculty.mod, direction="backward")
step$anova
library(relaimpo)
final.faculty<-lm(PRFaculty ~ NoInterestingEvents +
NotDegreeOnTheCampus + RepresentNotShareExpertise +
                    PeopleFromTheCampusDonTLiveInOurCommunity +
TheCampusNotContributeEconomy +
                    CampusNotInvolvedInCommunity +
StudentMisbeh + StoreVisitFaculty,data = faculty)


calc.relimp(final.faculty,type=c("lmg"),
            rela=TRUE)
```

model staff and administrators

```
PRstaff.mod<-lm(PR_StaffAdministration~ .-PR_Students -
PR_Faculty-PR_BoardMembers,data =community_fac )
```

```
ctable <- coef(summary(PRfaculty.mod))
ctable<-as.data.frame(ctable)
ctable<-ctable %>%tibble::rownames_to_column(var = "Variables")
%>%filter( `Pr(>|t|)`<0.05)

ctable<-ctable[-1,c("Variables","Pr(>|t|)")]
ctable
vv<-vif(PRstaff.mod )
step <- stepAIC(PRstaff.mod, direction="backward")
step$anova



library(relaimpo)
final.staff<-lm(PRStaffAdministration ~ NoInterestingEvents +
NotDegreeOnTheCampus +
                    RepresentNotShareExpertise +
RepresentNotVisibleAtCommunityEvents +
                    StoreVisitCampusStaff,data = staff)

calc.relimp(final.staff,type=c("lmg"),
            rela=TRUE)
# Bootstrap Measures of Relative Importance (1000 samples)
boot <- boot.relimp(final.staff , b = 500, type = c("lmg"), rank
= TRUE,
                    diff = TRUE, rela = TRUE)
plot(booteval.relimp(boot,sort=TRUE))
```

model board members

```
PRboard.mod<-lm(PR_BoardMembers ~ .-PR_Students -
PR_Faculty-PR_StaffAdministration,data =community_fac )
ctable <- coef(summary(PRfaculty.mod))
ctable<-as.data.frame(ctable)
ctable<-ctable %>%tibble::rownames_to_column(var = "Variables")
%>%filter( `Pr(>|t|)`<0.05)

ctable<-ctable[-1,c("Variables","Pr(>|t|)")]
ctable
vv<-vif(PRboard.mod )
```

```
step <- stepAIC(PRboard.mod , direction="backward")
step$anova

library(relaimpo)
final.board<-lm(Prboard ~
PeopleFromTheCampusDonTLiveInOurCommunity +
TheCampusNotContributeEconomy +
                  RepresentativesNotVisibleAtCommunityEvents +
StudentMisbehaviorInTheCommunity
                ,data = board)

calc.relimp(final.board,type=c("lmg"),
            rela=TRUE)
# Bootstrap Measures of Relative Importance (1000 samples)
boot <- boot.relimp(final.board , b = 500, type = c("lmg"), rank
= TRUE,
                    diff = TRUE, rela = TRUE)
plot(booteval.relimp(boot,sort=TRUE))
```

Appendix B

Data clean, calculation and data visualization for preliminary part

Calculate the means

```
studentcom<-mean(community$OnAScaleRangingFromVeryNegativeToVery
PositiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroups
OfPeopleAndTHECOMMUNITYASAWHOLEStudents)
studentper<-mean(community$YOUROWNPERSONALRELATIONSHIPSWithTheFo
llowingGroupsOfPeopleStudents)
studentcom
studentper
facultycom<-mean(community$OnAScaleRangingFromVeryNegativeToVery
PositiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroups
OfPeopleAndTHECOMMUNITYASAWHOLEFaculty)
facultyper<-mean(community$YOUROWNPERSONALRELATIONSHIPSWithTheFo
llowingGroupsOfPeopleFaculty)
facultycom
facultyper
staffcom<-mean(community$OnAScaleRangingFromVeryNegativeToVeryPo
sitiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOf
PeopleAndTHECOMMUNITYASAWHOLEStaffAdministration)
staffper<-mean(community$YOUROWNPERSONALRELATIONSHIPSWithTheFoll
owingGroupsOfPeopleStaffAdministration)
staffcom
staffper
boardcom<-mean(community$OnAScaleRangingFromVeryNegativeToVeryPo
sitiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOf
PeopleAndTHECOMMUNITYASAWHOLEBoardMembers)
boardper<-mean(community$YOUROWNPERSONALRELATIONSHIPSWithTheFoll
owingGroupsOfPeopleBoardMembers)
boardcom
boardper
```

Calculate the variance

```
studentcom1<-var(community$OnAScaleRangingFromVeryNegativeToVery
PositiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroups
OfPeopleAndTHECOMMUNITYASAWHOLEStudents)
```

```
studentper1<-var(community$YOUROWNPERSONALRELATIONSHIPSWithTheFo
llowingGroupsOfPeopleStudents)
studentcom1
studentper1
facultycom1<-var(community$OnAScaleRangingFromVeryNegativeToVery
PositiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroups
OfPeopleAndTHECOMMUNITYASAWHOLEFaculty)
facultyper1<-var(community$YOUROWNPERSONALRELATIONSHIPSWithTheFo
llowingGroupsOfPeopleFaculty)
facultycom1
facultyper1
staffcom1<-var(community$OnAScaleRangingFromVeryNegativeToVeryPo
sitiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOf
PeopleAndTHECOMMUNITYASAWHOLEStaffAdministration)
staffper1<-var(community$YOUROWNPERSONALRELATIONSHIPSWithTheFoll
owingGroupsOfPeopleStaffAdministration)
staffcom1
staffper1
boardcom1<-var(community$OnAScaleRangingFromVeryNegativeToVeryPo
sitiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOf
PeopleAndTHECOMMUNITYASAWHOLEBoardMembers)
boardper1<-var(community$YOUROWNPERSONALRELATIONSHIPSWithTheFoll
owingGroupsOfPeopleBoardMembers)
boardcom1
boardper1
```

Calculate the numbers of voters in each section

```
studentcom2<-sum(community$OnAScaleRangingFromVeryNegativeToVery
PositiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroups
OfPeopleAndTHECOMMUNITYASAWHOLEStudents)
studentper2<-sum(community$YOUROWNPERSONALRELATIONSHIPSWithTheFo
llowingGroupsOfPeopleStudents)
studentcom2
studentper2
facultycom2<-sum(community$OnAScaleRangingFromVeryNegativeToVery
PositiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroups
OfPeopleAndTHECOMMUNITYASAWHOLEFaculty)
facultyper2<-sum(community$YOUROWNPERSONALRELATIONSHIPSWithTheFo
llowingGroupsOfPeopleFaculty)
```

```
facultycom2
facultyper2
staffcom2<-sum(community$OnAScaleRangingFromVeryNegativeToVeryPo
sitiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOf
PeopleAndTHECOMMUNITYASAWHOLEStaffAdministration)
staffper2<-sum(community$YOUROWNPERSONALRELATIONSHIPSWithTheFoll
owingGroupsOfPeopleStaffAdministration)
staffcom2
staffper2
boardcom2<-sum(community$OnAScaleRangingFromVeryNegativeToVeryPo
sitiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOf
PeopleAndTHECOMMUNITYASAWHOLEBoardMembers)
boardper2<-sum(community$YOUROWNPERSONALRELATIONSHIPSWithTheFoll
owingGroupsOfPeopleBoardMembers)
boardcom2
boardper2
```

Calculate the whole numbers of voters

```
comsum<-studentcom2+facultycom2+staffcom2+boardcom2
persum<-studentper2+facultyper2+staffper2+boardper2
```

Calculate the numbers of voters in each section and make preparations to the data visualization

-----No trouble-----

```
NoTrouble<-sum(community$PrivateResidencesIncludingHousesApartme
ntsAndMiniDorms=="1")+
sum(community$FraternityHouses=="1")+sum(community$BarsTavernsAn
dNightclubs=="1")+
sum(community$Restaurants=="1")+sum(community$OnCampusResidenceH
alls=="1")+
  sum(community$StadiumArenaOrTailgateParties=="1")
```

-----Small Trouble-----

```
SmallTrouble<-sum(community$PrivateResidencesIncludingHousesApar
tmentsAndMiniDorms=="2")+
sum(community$FraternityHouses=="2")+sum(community$BarsTavernsAn
dNightclubs=="2")+
sum(community$Restaurants=="2")+sum(community$OnCampusResidenceH
alls=="2")+
  sum(community$StadiumArenaOrTailgateParties=="2")
```

-----Moderate Trouble-----

```
ModerateTrouble<-sum(community$PrivateResidencesIncludingHousesA
partmentsAndMiniDorms=="3")+

sum(community$FraternityHouses=="3")+sum(community$BarsTavernsAn
dNightclubs=="3")+

sum(community$Restaurants=="3")+sum(community$OnCampusResidenceH
alls=="3")+
  sum(community$StadiumArenaOrTailgateParties=="3")
```

-----Significant Trouble-----

```
SignificantTrouble<-sum(community$PrivateResidencesIncludingHous
esApartmentsAndMiniDorms=="4")+
sum(community$FraternityHouses=="4")+sum(community$BarsTavernsAn
dNightclubs=="4")+
sum(community$Restaurants=="4")+sum(community$OnCampusResidenceH
alls=="4")+sum(community$StadiumArenaOrTailgateParties=="4")
```

-----Tremendous Trouble-----

```
TremendousTrouble<-sum(community$PrivateResidencesIncludingHouse
sApartmentsAndMiniDorms=="5")+
sum(community$FraternityHouses=="5")+sum(community$BarsTavernsAn
dNightclubs=="5")+
sum(community$Restaurants=="5")+sum(community$OnCampusResidenceH
alls=="5")+
  sum(community$StadiumArenaOrTailgateParties=="5")
```

Collect all the data in a new excel file

```
trouble<-read_xlsx("/Users/Qiu/Documents/STA 475/Project
2/trouble.xlsx")
View(trouble)
trouble$Trouble <- factor(trouble$Trouble, levels = c("No
Trouble","Small Trouble","Moderate Trouble","Significant
Trouble","Tremendous Trouble"))
```

ggplot building: “trouble level created by alcohol misuse”

```
ggplot(trouble, aes(x=Trouble, y=Numbers, fill=`Places and
Reasons`)) +
  geom_bar(stat="identity",width=0.6) +
  theme_classic()+
  theme(axis.title.x = element_blank(),
        axis.ticks.y = element_blank(),
        legend.position = c(0.9,0.87),
        legend.text=element_text(size=8),
        text=element_text(size=15),
        plot.title = element_text(hjust = 0.5))+
  ggtitle("Trouble Levels Created by Alcohol Misuse")
```

Appendix C

t-test procedure
choose pearson t test to check if there is some correlation between the two(personal and whole)
pearson t-test to check the correlation coefficients among variables

------student aspect---------

```
student<-cor(community$
OnAScaleRangingFromVeryNegativeToVeryPositiveHowWouldYouRateTheR
elationshipsBetweenTheFollowingGroupsOfPeopleAndTHECOMMUNITYASAW
HOLEStudents,
community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleStudents,method="pearson")
```

-------Faculty aspect-------

```
faculty<-cor(community$OnAScaleRangingFromVeryNegativeToVeryPosi
tiveHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOfPe
opleAndTHECOMMUNITYASAWHOLEFaculty,
community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleStudents,method="pearson")
```

------Staff aspect-----

```
staff<-cor(community$OnAScaleRangingFromVeryNegativeToVeryPositi
veHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOfPeop
leAndTHECOMMUNITYASAWHOLEStaffAdministration,
community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleStaffAdministration,method="pearson")
```

-----Board aspect-----

```
board<-cor(community$OnAScaleRangingFromVeryNegativeToVeryPositi
veHowWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOfPeop
leAndTHECOMMUNITYASAWHOLEBoardMembers,
community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleBoardMembers,method="pearson")
```

--------t-test comparing population mean------

```
t.test(community$
OnAScaleRangingFromVeryNegativeToVeryPositiveHowWouldYouRateTheR
elationshipsBetweenTheFollowingGroupsOfPeopleAndTHECOMMUNITYASAW
HOLEStudents,
community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleStudents,paired=TRUE)

t.test(community$OnAScaleRangingFromVeryNegativeToVeryPositiveHo
wWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOfPeopleAn
dTHECOMMUNITYASAWHOLEFaculty,

community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleStudents,paired=TRUE)

t.test(community$OnAScaleRangingFromVeryNegativeToVeryPositiveHo
wWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOfPeopleAn
dTHECOMMUNITYASAWHOLEStaffAdministration,

community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleStaffAdministration,paired=TRUE)
t.test(community$OnAScaleRangingFromVeryNegativeToVeryPositiveHo
wWouldYouRateTheRelationshipsBetweenTheFollowingGroupsOfPeopleAn
dTHECOMMUNITYASAWHOLEBoardMembers,

community$YOUROWNPERSONALRELATIONSHIPSWithTheFollowingGroupsOfPe
opleBoardMembers,paired=TRUE)
```