

**A Statistical Analysis of  
Changing Religious Perspectives  
Over the period of time spent in Higher Education**

Katherine Dykes  
Systems 302  
Dr. Tony Smith  
Due: 5 – 9 – 03

# Table of Contents

Table of Contents.....	2
Introduction.....	3
Variables Defined.....	4
Assumptions.....	4
Summary of Findings.....	6
Variable Distribution.....	9
Logistic Regression.....	11
Stepwise Regression and Multicollinearity.....	14
Predictions from Logistic Model.....	16
Conclusion.....	17
Sources.....	19
Appendix A: Histograms.....	20
Appendix B: Survey Form.....	23

## Introduction

The College House System staff is well trained to cope with a variety of problems that students make the transition from youth to adulthood throughout their university experiences. Considerable time is spent in training staff to deal with the stresses of academic pressures, cultural differences, racial identities, and relationship dynamics. However, due to the sensitive nature of personal beliefs, both religious and philosophical, the training does not touch on the spiritual development of the student throughout their undergraduate career.

In fact, the only mention of such issues in the RA/GA manual, is that “religion and spiritual development are often very important [to the undergraduate experience], especially in the context of a non-denominational institution such as the University of Pennsylvania, and is generally not accounted for in most student development theories” (Manual, 29). It goes on to mention that “students must be given the freedom to observe religious holidays, practice traditions, attend services, etc” (Manual, 29). Importance is often placed on religious freedom, so much so that studying its effects is a precarious matter. However, the development of belief systems is fundamental and affects a person’s behavior and actions in many other contexts.

The lack of information is intriguing, and thus this statistical analysis takes a first stab at identifying some of the factors that may affect the spiritual development of a student in an institution of higher education, specifically here at the University of Pennsylvania. 44 graduating seniors were interviewed randomly from the population of seniors at the

University of Pennsylvania. These surveys asked several questions that would hope to pin down some of the fundamental factors that influence spiritual development in the undergraduate population over a four-year period. These questions were grouped roughly by religious background, current religious activity, social activity, and living & learning environment. The results were used to try and determine which of these factors, if any, are influential in shaping whether or not the student goes through significant spiritual development while at the university.

## Variables Defined

The variables used in this study corresponded directly to the survey questions asked.

Please see appendix B: Survey Questionnaire.

Table 1: Variable names in JMP-IN with respect to Survey Questions

Variable	Question Number
Number of Majors	0
Hours / Week in Religious Group	3
Level of Activity in Living Environment	4
Conversation of same religion	5
Conversation with diff religion	6
Academic Courses	7
Books	8
Interaction with Diverse groups	9
Social Activity Level	10
Extra-curricular	11
Religious Services	12
Interest of Parents in Religion	15
High School Activity	16

## Assumptions

There were several important assumptions made when selecting the initial set of variables to test as well as the population to test them against. The most fundamental were with respect to the time period over which the questions were directed and the selection of variables which were more external in nature rather than a reflection of intrinsic

characteristics of an individual. The first had to deal with when the significant developments in student spirituality take place. Given more time and attention to an in-depth study, it might prove much more useful to take a selection of data at different points in a student's career: upon entering the university, at the end of each semester and/or academic year and then finally upon graduation.

This questionnaire assumes that the survey respondent can look back and accurately reflect and assess the changes that occurred in their beliefs over their college career without any bias. To remove such a bias, data should be collected as closely to when those changes in perspective are occurring rather than after the fact. The scope of this particular project did not allow for such a study, and so it was assumed that each graduating senior is well aware of the spiritual development process that he or she has gone through. Thus the graduating senior class was chosen as the population for study and 44 random individuals were selected as a population sample.\*

The second assumption has to deal with selection of variables themselves. Given the difficulty in addressing the more personal issues and the lack of objectivity one often has in assessing their own personality, the study did not take into account intrinsic traits of an individual that may affect their spiritual development. The study discounts the argument that some people are intrinsically more open to changing perspectives, differing viewpoints and assumes that the factors influencing spiritual development are external.

Some of these internal characteristics are taken into account indirectly when reflecting on the fact that a persons actions (i.e. discussing religion with people of same and different religions, reading books on various philosophical subject matter) are directly influenced by their intrinsic nature. This is, of course, another aspect of this assumption used by the study for variable selection. Beyond these two primary assumptions, it was decided that for a number of survey responses (see appendix B), a scale of 1 to 5 was appropriate to assess the intensity or frequency of the various independent variables.

*\* In addition, it should be noted that in order to obtain 50 random responses from seniors. The writer staked out the "Senior Week" booth on locust walk and targeted seniors that were patrons of the table. Therefore the senior population is representative only of seniors that attend Senior Week events. This study assumes that the number of seniors attending Senior Week events is roughly equivalent to the senior class student body as a whole.*

## **Summary of Findings**

The first step in the analysis was to look at the distribution of the data among the different explanatory variables. Given the nature of the scales used to obtain the data, normality was difficult to visualize. The data was often clustered towards the upper or lower limit due to two reasons: the scales imposed a limit to resolution of various survey responses and there were several variables that involved a response that was greater than or equal to zero and therefore data was clustered at this lower limit.

The distribution data showed several interesting facts about the frequency of various activities and habits of students at the university. Some of which, for instance the lack of reading students do on the subjects of philosophy and/or religion, were very surprising and in themselves merit further study. Given a larger project scope, it would be

interesting to analyze what academic factors influence the opportunities of students to engage in other activities that affect spiritual development.

After looking at the distributions, the next step was to determine which, if any, of these independent variables had an effect on the change or lack of change in a student's spiritual perspectives at the university. The variable under test was specifically the dichotomous variable of whether or not the student assessed that their beliefs had undergone any considerable changes over their university attendance. Because, of the dichotomous nature of the variable, a logistic regression had to be used in order to obtain representative results.

Two variables were tested, one that measured the intensity of the change on a scale of 1 to 5 and one that measured whether a change had occurred at all. The variable of interest was the second and was used to first narrow down the regression to the specific variables that significantly affect spiritual development. These were determined by running a logistic regression in jump-in on all 13 variables and iteratively removing the least significant ones until all the variables had a Prob>ChiSq that was less than 0.10 and until the whole model test showed a Prob>ChiSq that was less than 0.0001.

More analysis had to be done to assess why the selected variables did in fact explain the change in beliefs of the student over the other variables both from an intuitive perspective of why certain variables should have certain effects and an analytical perspective looking at any multicollinearities that might exist between the removed variables. To search for

multicollinearities, a mixed stepwise regression was performed on the variable measuring the intensity of change in religious perspective with all of the variables for explanation. Two variables, high school activity and interest of parents in religion were discounted right away due to their high levels of (Prob>ChiSq >= 0.6). Beyond that, questions of multicollinearity were proposed between academic courses and books read, social activity and extra-curricular activity, hours / week in a religious group and conversation of same religion, and conversation with different religion and interaction with diverse groups.

Significant multicollinearity was found between academic courses and books read. Social activity and extra-curricular activity had less R Squared Adjusted when plotted against each other, but the correlation on estimates showed the second highest correlation between these two variables. The last four variables did not show a huge amount of correlation between each other, but one can intuitively understand how each of these is related to one another.

Having justified the model obtained through the logistic regression method, we compared the predictions of the model to our actual results to look at a success rate. The graph of predictions would not give us too much insight because of the dichotomous variable. For a logistic regression, saving the probability formula allowed us to calculate the number of successful predictions to find a success rate. The model we created, which was based only on the test data collected, showed a success rate of 81% which is acceptable given the latitude of the assumptions stated previously.

The final variables of significance were found to be number of majors, level of religious activity in living environment, conversations with people of the same religion, interaction with diverse groups of people, reading books on philosophy and/or religion, and inversely the number of extra-curricular activities a person holds. These results are interesting for a number of reasons that will be discussed in the paper's conclusion.

## **Variable Distribution**

Very quickly, looking at the distribution for the various variables gives us an idea as to whether assumptions of normality will hold for the regression analysis. The first three variables of number of majors, hours / week spent with a religious group and level of religious activity do not offer too many surprises. The first two are clustered at the lower limit of 1 and 0 respectfully. Most students carry one major at the university and are not involved in any religious organizations. The level of religious activity in various students' living environment is spread normally around a medium level such that on average students live in a somewhat religiously oriented environment.

The next two variables regarding the frequency of conversations held on subjects of philosophy and religion with people of similar or different beliefs offer interesting results. Both are somewhat normal though the frequency of conversations held with people of similar beliefs is skewed towards the lower limit while conversations with people of different beliefs are skewed towards the upper limit. Students have more frequent discussion on religious and philosophical topics with people of differing belief systems than their own. The next variable to note is the level of interaction with people of diverse backgrounds which is also skewed towards the upper limit. This leads us to

believe we might see a correlation between interaction with diverse groups and conversations held with people from different belief systems which is what we would intuitively expect. This will be kept in mind for later analysis.

The next two variables of number of books read and the number of classes taken on philosophical and religious subject matter are both skewed heavily towards the lower limit of zero. However, we might expect another correlation here between these two variables since students taking courses on such subjects will also be reading a number of books on such subjects. The small number of students who actually dedicate a portion of their time spent at the undergraduate level to religious and philosophical studies is surprising given nature of academic institutions. One might wonder if this is an aspect more of UPenn's undergraduate curriculum and whether this is a trend across many universities. Further studies might wish to look at students from universities with a spectrum of required curriculums.

Social activity and extra-curricular activity again appear to be skewed towards the upper-limit and possibly correlated. Social activity and extra-curricular activity for the group surveyed might be higher than expected for the sample population of graduating seniors because only seniors interested in attending senior week events were interviewed, and one might expect that the more social and involved students of the class would be the more likely group to attend such functions.

Lastly, the interest of parents in the religious activities of the students was roughly normal while religious activity in high school tended to the lower limit. Given the fact that most graduating seniors were not religiously active here at the university, one might suppose that many were not very religiously active in high school either. As the RA/GA manual addressed, the non-denominational nature of the UPenn institution is probably responsible for the lack of religious involvement of its students.

## **Logistic Regression**

Keeping in mind what the interesting components of the distributions on each independent variable, a logistic regression was then run initially using all 12 of these variables as well as attendance at religious services while at the university. Including religious services, which was not of any significance when a stepwise regression was run, made it impossible to run the logistic regression in jump-in. Having removed this variable, the regression that was run showed a relatively high success rate at predicting the appropriate change or lack of change in the religious perspective of the student with an 86% success rate. This is at first a good indication that there is some significance relating the independent variables to change because in a logistic regression, the graphs of Fit Y by X and of the residuals are difficult to interpret. Therefore, the success rate of prediction is important to the model's assessment.

However, it is also important to note the Chi-Square value for each variable under test because the significance of variables in a logistic regression is represented by this column. The Prob>ChiSq represents the probability of obtaining the given beta if the actual beta were equal to zero. It is a test of the null hypothesis that the beta for a

particular independent variable equals zero. At first glance, several of the variables selected to explain the change, or lack of change, have uncomfortably high values in this column.

Table 2: Parameter Estimates for Logistic Regression (All Variables)

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	56.3378404	32.56695	2.99	0.0836
Number of Majors	-12.342674	7.4873474	2.72	0.0993
Hours / Week in Religious Group	0.10206902	0.4049504	0.06	0.8010
Level of Activity in Living Environment	-3.7561861	2.1546703	3.04	0.0813
Conversation of same religion	-2.5182859	1.559486	2.61	0.1064
Conversation with diff religion	-2.4410159	1.910765	1.63	0.2014
Academic courses	4.66903117	2.6797098	3.04	0.0814
Books	-3.3337155	1.8997448	3.08	0.0793
Interaction with Diverse groups	-0.9911796	0.8962381	1.22	0.2688
Social Activity Level	-4.2563418	2.7963564	2.32	0.1280
Extra-curricular	4.4153627	2.4069271	3.37	0.0666
High School Activity	-2.3751701	1.4995302	2.51	0.1132
Interest of Parents in Religion	-1.0054769	1.2709518	0.63	0.4289

Many of these variables have an overall Prob>ChiSq that is greater than or equal to 0.10 which will be used as a standard to compare whether or not a particular variable should be kept in the regression.

Beginning with the least significant variable Hours / Week in a Religious Group at 0.80, variables were removed until all the explanatory variables were relatively significant. The variables left at the end of this process were number of majors, level of religious activity in living environment, conversations held with people of same religion, books read on subject matter, interaction with diverse groups of people, and extra-curricular involvement with Prob>ChiSq all below 0.10 as seen below.

Table 3: Parameter Estimates for Logistic Regression (Final Variables)

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	11.7654809	4.388609	7.19	0.0073
Number of Majors	-3.1579465	1.3478146	5.49	0.0191
Level of Activity in Living Environment	-1.4188435	0.6656317	4.54	0.0330
Extra-curricular	0.80956957	0.4155559	3.80	0.0514
Books	-0.4473518	0.181731	6.06	0.0138

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Conversation of same religion	-1.1735199	0.4916936	5.70	0.0170
Interaction with Diverse groups	-0.7334752	0.4142486	3.14	0.0766

The nominal variable of Dichotomous change were given the representations of 1 and 0 such that the model is actually predicting the lack of change (given the alphabetical assignment order in JMP-IN) so these variables represent the opposite of what we might expect. It seems that interaction with diverse groups, conversations held with people of the same religion, level of religious activity in living environment, books read and the number of majors tends to decrease the probability of a student keeping their same beliefs throughout their undergraduate experience while level of extra-curricular activities augments this tendency.

This is what we might expect. If a person is really busy with extra-curricular activities, then the time they have to reflect on their spiritual development would decrease. Living in an environment that is somewhat religious in nature, reading books on such subjects and interacting with people of diverse backgrounds would all be expected to increase the likelihood of a student to change their religious or philosophical viewpoints. The two more surprising variables are the number of majors that a student has at the university and the conversations with people of their same religion.

The first can be explained by the fact that a person who is interested in studying several different areas is also someone who is open-minded and curious about a range of different academic topics, including religion and philosophy. From this perspective, the increase in likelihood of change due to the number of majors of a student makes sense. The other surprising note is that conversations held about philosophy and religion with

people of similar beliefs increases the likelihood of a person changing their religious perspective. Looking over the survey responses, it was noted that a number of people who noticed a change in their religious viewpoints actually noticed a strengthening of their beliefs in the religion they had upon entrance to the university. Thus, talking with people of the same religion helped to reinforce, in most cases, a student's beliefs in their religion.

## Stepwise Regression and Multicollinearity

Having narrowed down the variables of importance, attention was given as to why these variables, among the others, would be significant. To do this, a stepwise regression was performed against the intensity of the change experienced by students in their spiritual development.

Table 4: Stepwise Regression on Intensity of Change (All Variables)

Parameter	Estimate	nDF	SS	"F Ratio"	"Prob>F"
Intercept	0.76386843	1	0	0.000	1.0000
Number of Majors	0.89106132	1	12.57163	5.473	0.0247
Hours / Week in Religious Group	0.30998555	1	13.25642	5.771	0.0213
Level of Activity in Living Environment	0.29742132	1	3.751201	1.633	0.2090
Conversation of same religion	.	1	2.891825	1.268	0.2674
Conversation with diff religion	0.34655879	1	4.133755	1.800	0.1877
Academic courses	.	1	2.698389	1.180	0.2843
Books	.	1	2.151472	0.935	0.3398
Interaction with Diverse groups	.	1	1.364183	0.587	0.4483
Social Activity Level	.	1	0.670159	0.286	0.5958
Extra-curricular	-0.3877422	1	9.702691	4.224	0.0468
Religious Services	.	1	1.945382	0.843	0.3644
Interest of Parents in Religion	.	1	0.038339	0.016	0.8992
High School Activity	.	1	0.547061	0.233	0.6319

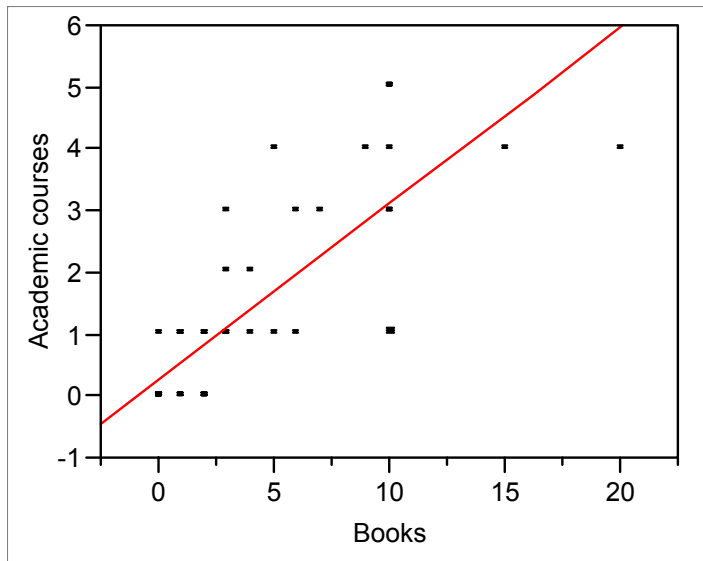
Interest of Parents in Religion, High School Activity and Social Activity Level were immediately dismissed due to their high Prob>F. Interest of parents in the religion of their children does not seem to dictate whether or not a student's beliefs change at the university. The freedom associated with the university life was cited in many survey responses to have allowed them to explore new religious viewpoints. Religious activity

in high school was insignificant and had minimal correlation with any of the other explanatory variables. These two variables were removed from the regression and the regression was rerun with attention to the Variance Inflation Factor as a first indication of multicollinearity.

Table 5: Stepwise Regression on Intensity of Change  
(Removed Least Significant Variables)

Term	Estimate	Std Error	t Ratio	Prob> t	VIF
Intercept	-0.438454	1.673445	-0.26	0.7949	.
Number of Majors	0.9547504	0.432009	2.21	0.0341	1.284138
Level of Activity in Living Environment	0.3654691	0.251228	1.45	0.1552	1.4586639
Conversation of same religion	0.2597769	0.285349	0.91	0.3692	2.2259227
Books	0.0383508	0.110319	0.35	0.7303	4.5421784
Interaction with Diverse groups	0.1644269	0.23188	0.71	0.4832	1.3102621
Extra-curricular	-0.449516	0.24142	-1.86	0.0715	1.6170349
Hours / Week in Religious Group	0.2090816	0.156458	1.34	0.1906	1.8092722
Conversation with diff religion	0.1653849	0.322662	0.51	0.6117	1.6589529
Academic courses	0.0284155	0.332729	0.09	0.9325	4.9748551
Social Activity Level	0.0682263	0.329628	0.21	0.8373	2.1182237

Academic courses and books have unusually high VIF's and so there is most likely strong multicollinearity between the two variables. Running a Fit Y by X confirms this assertion:



**Bivariate Fit of Academic Courses By Books**

**Linear Fit**

Academic courses = 0.289771 + 0.283824 Books

**Summary of Fit**

RSquare	0.669051
RSquare Adj	0.661171
Root Mean Square Error	0.932831
Mean of Response	1.386364
Observations (or Sum Wgts)	44

**Parameter Estimates**

Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	0.289771	0.184226	1.57	0.1232
Books	0.283824	0.030802	9.21	<.0001

Therefore, it is not surprising that only one of the two variables, books, is significant in our final logistic regression.

Other variables with high VIF's were social activity level and conversations with people of the same religion. The bivariate fit of social activity by extra-curricular activity did not show as good a fit as that of academic courses by books, but the Prob>[t] of extra-curricular activity was 0.0061 for a positive beta of 0.341 so multicollinearity between the two variables seems highly likely. This would explain why only extra-curricular activity appears in the final logistic regression.

## Predictions from Logistic Model

Having analyzed the interdependencies of the various explanatory variables, it is time to assess the accuracy of our final logistic model. The first step in this process is to look at the overall test for the model.

Table 7: Whole Model Test on Dichotomous Change (Sig Vars Only)

Model	-LogLikelihood	DF	ChiSquare	Prob>ChiSq
Difference	14.351208	6	28.70242	<.0001
Full	16.101798			
Reduced	30.453006			
RSquare (U)	0.4713			
Observations (or Sum Wgts)	44			

The Prob>ChiSq can be interpreted as the probability of obtaining the betas obtained (see Table 3) given the null hypothesis of each beta set to zero. In this case, with 6 degrees of freedom, the ChiSquared value was determined to be 28.7 which results in a Prob>ChiSq of <0.0001 which gives a good justification for the significance of the selected model with the given explanatory variables. The RSquare (U) or Likelihood Ratio Index (LRI)

is at 0.4713 which does not carry the same meaning as the adjusted R-squared in stepwise regression.

Therefore, the better assessment of the model accuracy is the success rate of the model for the given sample data. The probability for the formula was saved so that a predicted value for dichotomous change was found for each student. This was compared to the actual dichotomous change value and a success rate column was determined with a one for a match between predicted and actual value and a zero otherwise. Finally, the success rate was found to be 82%. This is a decently high success rate though not quite as high as we would like it to be. The fact that our initial logistic regression which included all of the variables had a success rate of 86% indicates that the removal of some of the lesser significant variable took away some of the predictive power of the model. However, the removal of these variables was necessary to improve the significance of the explanatory variables (remove multicollinearity) as well as improve the significance of the overall model (reduce the Prob>ChiSq value of the whole model).

## **Conclusion**

An 82% success rate was satisfactory given the limitations of the study. Further studies would hope to improve this success rate. A study that looks more closely at intrinsic characteristics of individuals might be able to better explain the tendency to change or not to change one's religious and philosophical views while at the university. In addition, study that compared data over different periods of time might also serve to better explain the data. Finally, surveying among different academic institutions of different focus and

religious affiliations would increase the understanding of how curriculum and the university environment affect spiritual development.

However, this brief study did offer some justification for ideas that the personal pursuit of understanding of religious and philosophical ideas through reading, an interest in a range of academic topics, conversation with people about religious and philosophical subjects, interaction with people from diverse backgrounds and living in an environment with an active religious community will all increase the likelihood of the student to go through some sort of significant spiritual development while at the university. It is interesting to note that reading books on religious and philosophical material was the single most significant variable in explaining spiritual development. Glancing again at the histogram, we note the unfortunate clustering of students who read no books on such subject matter at the university. A shift in curriculum might remedy such a state and affect overall the likelihood of students to go through significant spiritual development during their undergraduate experience.

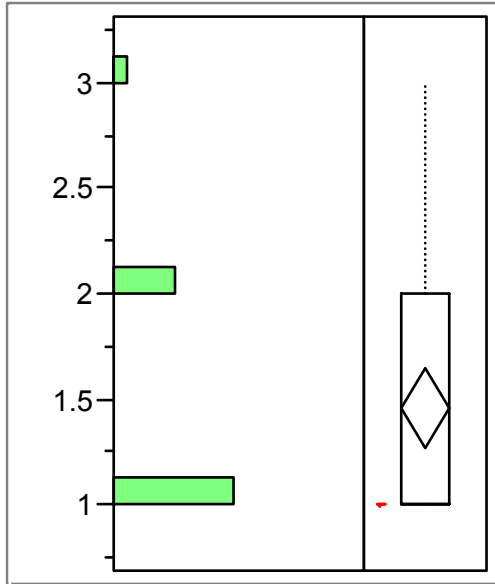
Involvement in a variety of extra-curricular activities that fill up ones time will detract from this opportunity. Therefore, if the UPenn undergraduate community is interested in creating an environment in which students can develop spiritually, they might take actions to increase the likelihood of the first four events mentioned above while taking a more critical view of students who engage in a number of activities that will not allow them such an opportunity.

## Sources

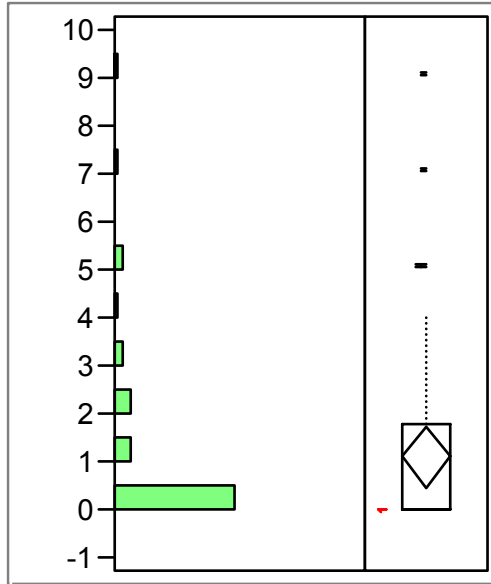
1. "Section 6: Student Development, Counseling, and Mediation." RA/GA Manual. University of Pennsylvania, 2002-2003
2. Smith, Tony E. "Notes on Logistic Regression." Systems 302:  
[http://www.seas.upenn.edu/%7Esys302/extra\\_mtls/Logistic\\_Regression.pdf](http://www.seas.upenn.edu/%7Esys302/extra_mtls/Logistic_Regression.pdf)

# Appendix A: Histograms

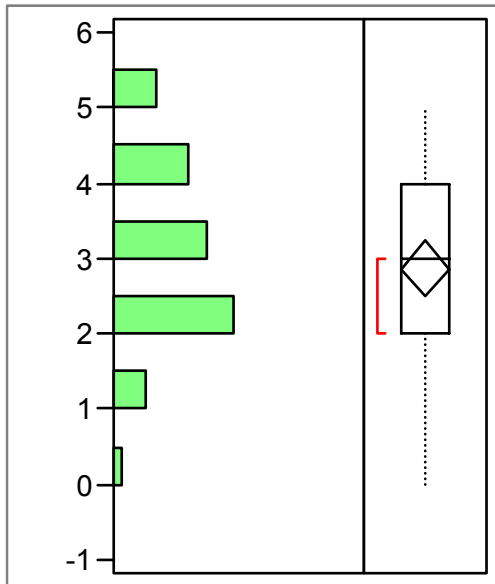
Number of Majors



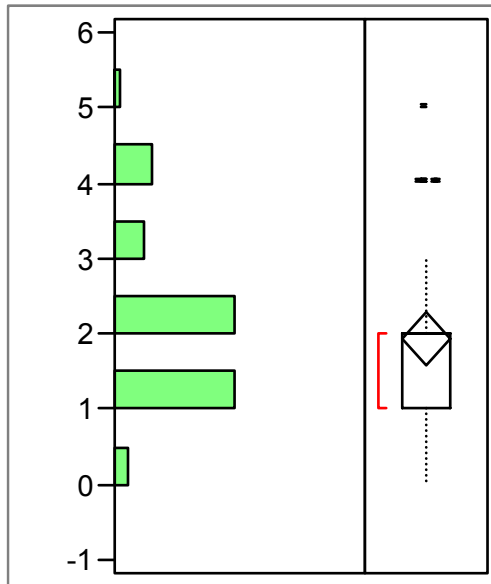
Hours / Week in Religious Group



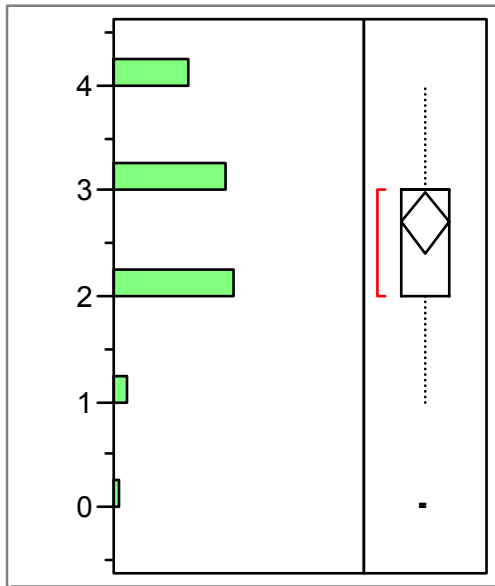
Level of Activity in Living Environment



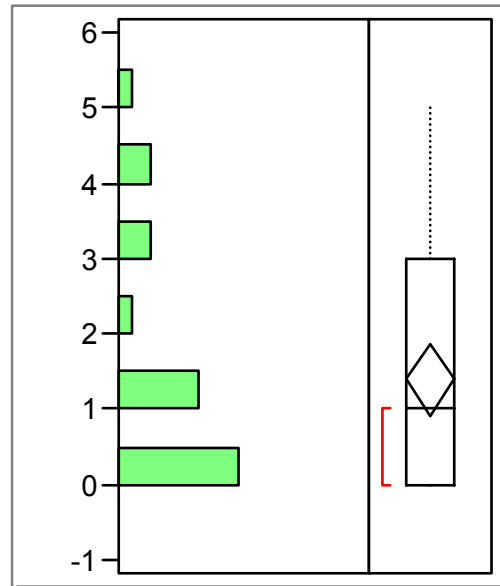
Conversation of same religion



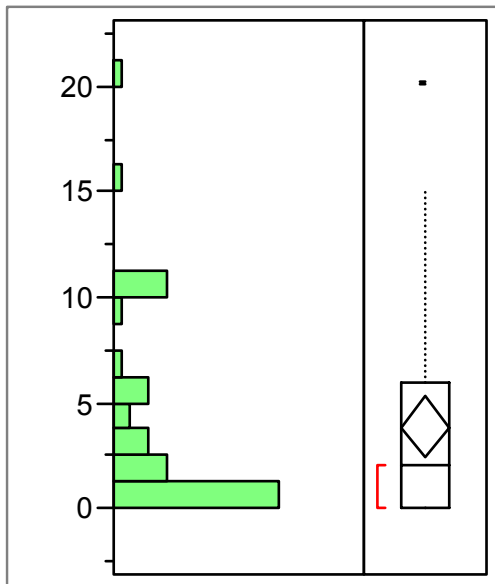
**Conversation with diff religion**



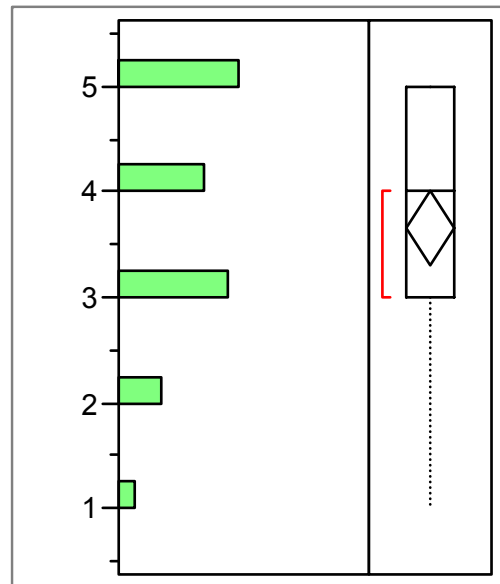
**Academic courses**



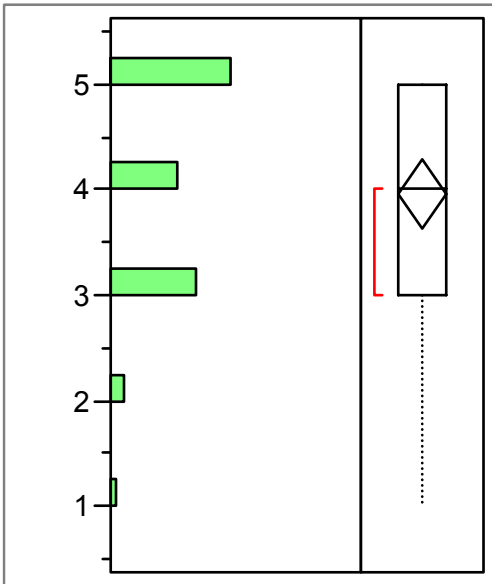
**Books**



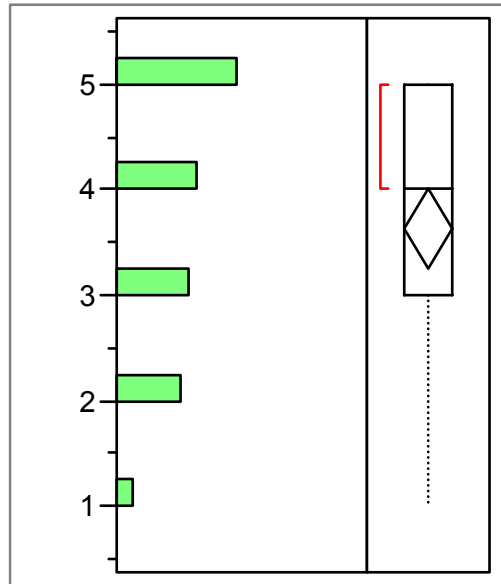
**Interaction with Diverse groups**



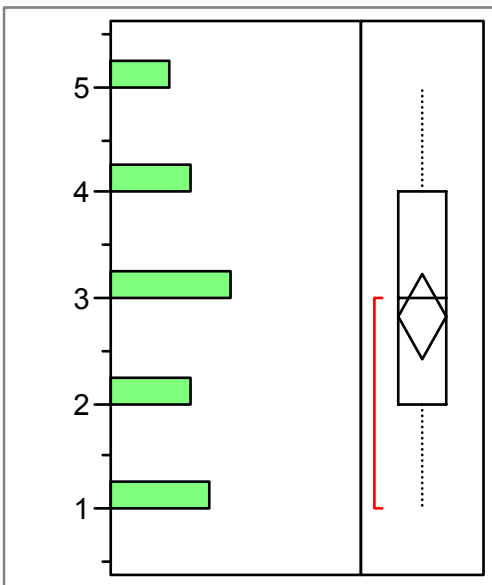
**Social Activity Level**



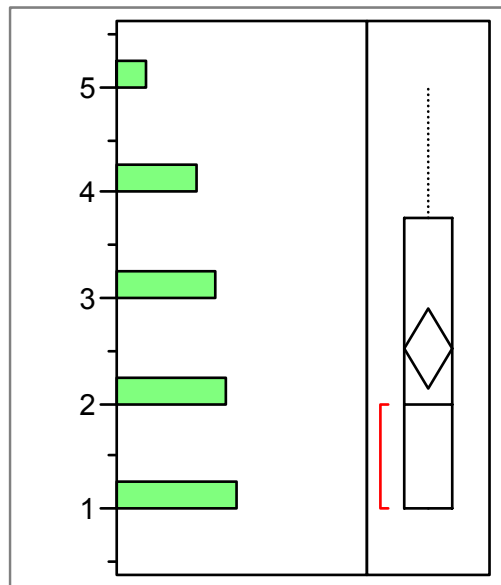
**Extra-curricular**



**Interest of Parents in Religion**



**High School Activity**



## Appendix B: Survey Form

Your School(s): \_\_\_\_\_  
Your Major(s): \_\_\_\_\_

2003 Graduating Seniors Survey on Religious Activity:  
(Please email responses to [dykesk@seas.upenn.edu](mailto:dykesk@seas.upenn.edu))

Present Day:

1. What is your religion (if any) now?
2. With what religious groups (if any) do you participate here at UPenn?
3. How many hours a week does your involvement with this group take?
4. What is the level of religious activity in your living environment? (scale of 1 to 5)
5. How frequently do you engage in conversations on philosophy and/or religion with people of your *same* religion? (daily, weekly, monthly, a few times each year, never)
6. How frequently do you engage in conversations on philosophy and/or religion with people of a *different* religion? (daily, weekly, monthly, a few times each year, never)
7. How many academic courses at the university have you taken with respect to philosophy and/or religious studies?
8. How many books have you read at the university related to this subject matter?
9. What level is your interaction with diverse groups of people here at the university? (scale of 1 to 5)
10. How active are you (socially)? (scale of 1 to 5)
11. How active are you (extra-curricular activities)? (scale of 1 to 5)
12. Do you attend religious services while at the university? (never, seldom, often)

Background:

13. What was your religion (if any) upon entrance to the university?
14. What is the religion (if any) of your parents?
15. How interested are your parents in your religious activities? (scale of 1 to 5)
16. How religiously active were you in high school? (scale of 1 to 5)

Comparison:

17. Do you believe that your beliefs have changed considerably over the 4 years?
18. If so, what would be the largest factor to affect that change?

Thanks very much for your time!