**Swarthmore Students’ Predicted and Actual Weekly Matchbox Attendance**

Econ 31: Introduction to Econometrics

Eugene Afranie, Daniel Banko, Maggie Christ, Patrick Holland, Sam Wang, and Liz Whipple

December 2015

**Introduction**

One of Swarthmore’s newest additions to campus is the Matchbox, designed to support the college’s fitness and wellness needs. We decided to observe Swarthmore students’ weekly Matchbox attendance to determine whether the frequency of going to the Matchbox was correlated with student majors, dorm proximity, political views, membership on a varsity or club athletic team, and other variables. We were also interested in students’ perceptions of average Matchbox attendance. Our choice to address these questions arose from our desire to understand the exercise habits of Swarthmore’s student body. We believed that varsity athletes would likely use the Matchbox at a higher rate than the rest of the student body, but we did not know what to expect from the rest of the variables we tested.

Our initial intuition turned out to have some merit. Although many of the variables we collected ultimately had insignificant correlations with personal and predicted Matchbox attendance, we found a strong correlation between membership on athletic teams and Matchbox attendance. Additionally, we were surprised to discover that females visit the Matchbox less than the rest of the student body.

We first explain our methodology and give background on some of the problems we faced while making our survey. Next we share our results from hypothesis testing. After that, we show the results of running regressions on our data set. Finally, the conclusion includes a discussion of further research questions and a summary of our findings.

**Methodology**

To determine which questions to include in our study we created a list of fifteen potential variables that might have been correlated to time spent at the Matchbox and perceptions of average Matchbox attendance. After deciding which questions to ask we created a survey using Google forms (see Appendix 1) and posted it in the class year Facebook groups. This strategy was relatively effective at obtaining results, as we received 89 usable responses to the survey. Once we closed our survey to responses, we had to transfer our dataset into a format acceptable for STATA. To do this we assigned number values to all of the different values of our variables. For example, we converted the original gender variable (0-2) to a dummy variable that assigned 1 for female and 0 for non-female. To create the difference variable we subtracted predicted Matchbox attendance from the time students reported they themselves spent in the Matchbox (diff = personalMatchboxHours - predictedMatchboxHours). We are relatively confident in our methodology, but this is not to say that we did not run into problems. We oversampled the class of 2018 and females using our method. This oversampling may have to do with the demographic of individuals who spend time on Facebook class groups. We also recognize that there may be a self reporting bias in our results. Swatties may over or underestimate the amount of time they spend in the Matchbox.

**Hypothesis Testing**

Before conducting particular hypothesis tests, we looked at the observed means across different groups. Varsity athletes reported spending on average 4.19 hours at the Matchbox per week, as opposed to 2.09 hours on average for non-varsity athletes. More generally, the average Swattie exercised 0.36 hours more per week than they believed they did (a difference variable mean of 0.36), revealing a perception that understates reality. The average Swattie seems to be unaware of how often other Swatties go to the Matchbox relative to their own weekly usage, and tend to believe that other Swatties go less often than they do.

Then, to analyze overall trends of student Matchbox attendance, we ran multiple hypothesis tests on the personal, predicted, and difference variables for the whole student body. First, we hypothesized that Swarthmore average Matchbox attendance would be approximately equal to two hours per week. Since the sample size n=87 is larger than 30 and population parameters are unknown, we ran a one-sided t-test with the null hypothesis of personalMatchboxHours=2. We calculated that we could reject the null hypothesis and accept that the true mean value of personal hours was above 2 with 99% confidence because the p-value in this test was 0.008. Second, we were interested in the mean number of hours students predicted that the average Swattie uses the Matchbox each week. Using the same one-sided t-test, we calculated a p-value of 0.0514 for the alternative hypothesis mean>2. Thus, we can reject the null hypothesis in favor of the alternative hypothesis with 90% confidence, but not with 95% confidence.

We also wanted to see whether on average Swarthmore students overestimate or underestimate their peers’ weekly hours of exercise at the Matchbox relative to their own. We ran a two-sided t-test on the difference variable with the null hypothesis diff=0 to reflect the hypothesis that, on average, students think that the rest of the campus community exercises as much as they do. With a 0.099 p-value, we were able to reject the null hypothesis with slightly more than 90% confidence. Thus, we were able to say that Swarthmore students predict that the campus average is different from their own attendance.

**Regressions**

To answer our research questions, we ran a series of regressions examining the relationships between variables related to Matchbox attendance – both students’ personal weekly hours and their predicted averages – with other variables we collected. Although we regressed most of the variables we collected with personal and predicted Matchbox hours, and with the difference between the two, there were certain variables that we expected to have stronger correlations than others. Logically, we anticipated that whether a student was a varsity athlete or not would have the largest effect on how much time they spent at the Matchbox. We also expected that dorm, ordered in distance from the Matchbox, would be correlated with average use, and that variables like stress level, major discipline and whether the student usually went with a friend to work out could potentially offer us insight into why some students work out more than others.

Our results were largely disappointing. We found that many of the variables had coefficients close to zero with no statistical significance, including ones that we thought would have a strong correlation with Matchbox use.

**Variables Regressed with Personal Matchbox Hours with P-Values Greater than .200**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Coefficient** | **P-Value** |
| Breakfast Frequency (Increasing, 1-3) | .212 | .459 |
| Start Time of Earliest Class | .246 | .282 |
| Class Year (1-4) | .126 | .652 |
| Hours of Sleep on Average Night | -.076 | .720 |
| Dorm (Ranked in order of distance to matchbox) | .016 | .932 |
| Stress Level (Increasing, 1-10) | .134 | .715 |
| Natural Science Major (1 if yes, 0 if no) | .514 | .760 |

We also regressed our variables with the difference between personal and predicted Matchbox hours to get a sense of how they were correlated with whether students thought themselves above or below average. Again, we found statistically insignificant results for most variables.

**Variables Regressed with Difference Variable with P-Values Greater than .200**

|  |  |  |
| --- | --- | --- |
| **Variable** | **Coefficient** | **P-Value** |
| Start Time of Earliest Class | .033 | .901 |
| Class Year | -.123 | .705 |
| Hours of Sleep on Average Night | -.035 | .889 |
| Dorm (Ranked in order of distance to matchbox) | -.100 | .635 |
| Average Number of Weekly Meals | .014 | .857 |
| Natural Science Major (1 if yes, 0 if no) | .108 | .858 |

Some of our results did not meet the 0.1 p-value level of significance but were below 0.15 and are worth mentioning as they may be significant if our sample size were bigger. The first was the correlation between a Swattie’s predicted number of weekly hours at the Matchbox and the start time of their earliest class this semester. We found that for each hour later that a Swattie’s first class of the day started, the predicted number of hours that the average Swatties goes to the Matchbox per week increased on average by 0.21 hours, with 89.7% confidence (p-value = 0.103). One possible explanation for this result is that feeling less tired might influence how one perceives other Swatties’ exercise habits, even if it does not change their own usage of the Matchbox. The second nearly significant result, at an 88.2% confidence level, was the correlation between Matchbox use with a friend and the difference variable. Students who reported going with friends to the Matchbox had a 0.91-hour smaller difference, on average, than Swatties who reported going alone.

Twenty-one (self-identified) varsity athletes participated in our survey, comprising 24 percent of our sample size. Running a simple two-variable regression on the results from our survey, we found that being a varsity athlete was significantly correlated with the number of weekly hours that a student spent at the Matchbox. Our findings were that, on average, the number of hours that a Swarthmore student spends at the Matchbox per week is 2.10 hours greater for varsity athletes than it is for non-varsity athletes (club sports team members included), with 99.9% confidence. This result is relatively unsurprising, as we would expect varsity athletes to exercise more at the Matchbox for both extrinsic and intrinsic reasons. For example, coaches expect their athletes to exercise more outside of practice, and athletes might already enjoy exercising more than non-athletes.

Being a varsity athlete was also positively correlated with the difference variable. On average, this difference was 2.52 more hours per week for varsity athletes than it was for the rest of our sample. In other words, varsity athletes believe that they go to the Matchbox significantly more hours per week than the average student does; or, conversely, varsity athletes believe that the average student goes to the Matchbox significantly less than they do. Their intuition is correct based on our results, but it is slightly overestimated. We see that the actual difference (based on the observed means) between the number of hours that a varsity athlete exercises and that the average Swattie exercises was 1.59 hours per week, which was 0.93 fewer hours than varsity athletes believed it to be.

Our second set of statistically significant regression results involved gender differences. Specifically, we regressed the personal Matchbox attendance variable and the difference variable on the female variable, and we found that both relationships were significant. With a p-value of 0.031, and thus a 96.9% confidence level, we found that being female was associated with, on average, going to the Matchbox 1.04 hours fewer per week. Even more interesting to us, however, was the statistically significant result (with a p-value of 0.051) that being female was, on average, associated with a smaller difference variable (by a factor of -1.10). The interpretation of this result is more complicated, as, depending on the individual, this could either translate into a less positive difference variable (in which women still estimated that the average Swattie went less than they did but by a smaller amount) or a more negative difference variable (in which women were more likely to think that the average Swattie went to the Matchbox *more* than they did). We were disappointed to see that both of these statistically significant results seem to reinforce gender stereotypes surrounding female athleticism and self-perceptions relative to the mean.

After we observed these significant relationships involving gender and varsity athlete status, we decided to run two multiple regressions that would allow us to test these factors’ joint significance and correct for possible bias. First, we regressed the personal Matchbox attendance variable on the female and the varsity athlete variables, and, second, we regressed the difference variable on the same two independent variables. In both regressions, the F-test p-values remained very low (0.0001 in both), and thus, unsurprisingly, the female and varsity athlete variables were jointly significant. However, for both regressions we found that the female coefficients became slightly less negative compared to the simple regressions, which indicated a negative bias in our original simple regression models. Given that we knew that being a varsity athlete was positively correlated with the difference and personal attendance variables, we concluded that the covariance between the female and varsity athlete variables must have been negative. In other words, our original coefficients in the simple models were too negative because they did not account for the fact that, in our sample, females were less likely to be varsity athletes.

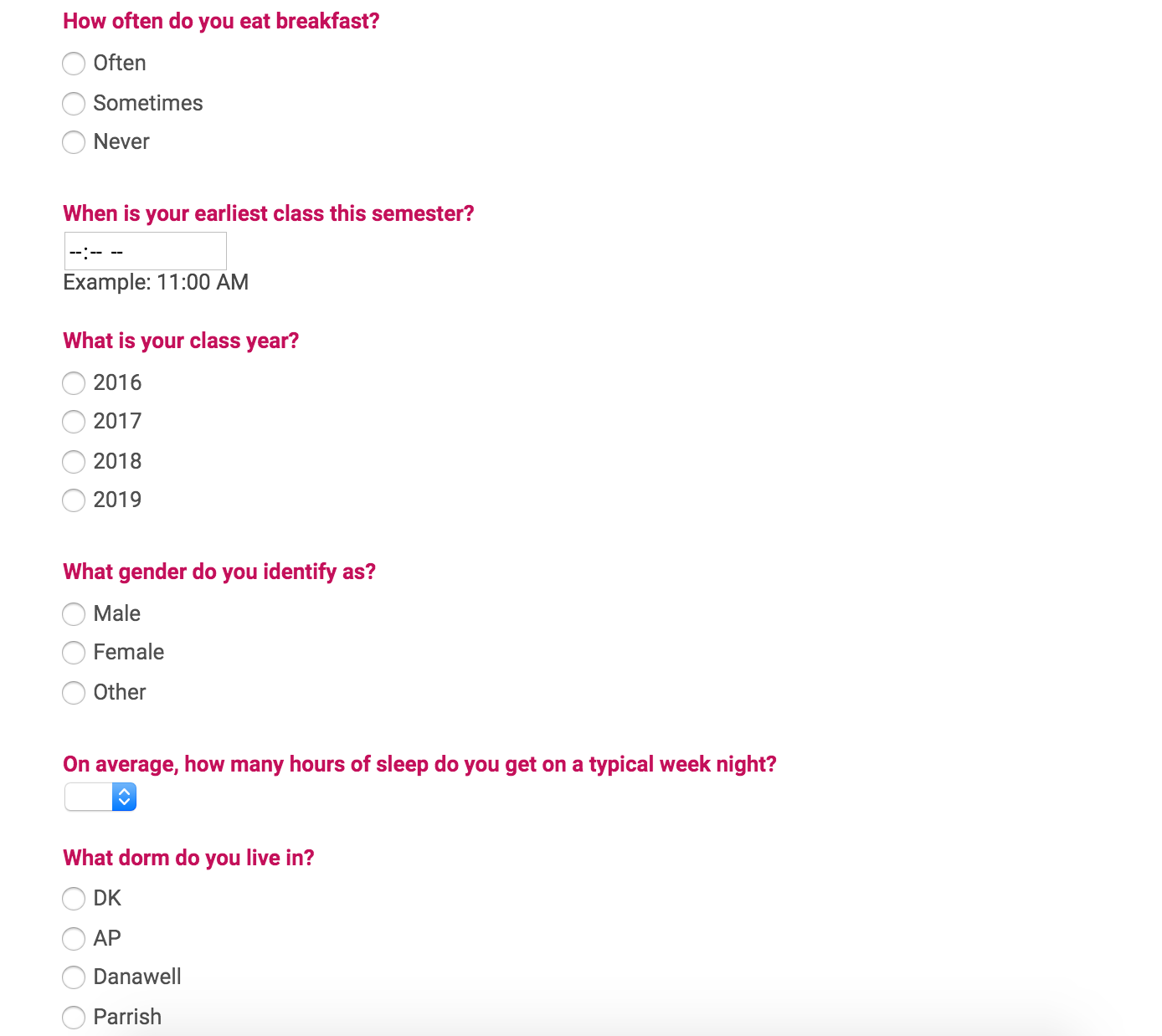
In the context of these multiple regressions, we also wondered whether there might be an interaction effect. Specifically, we wondered if female varsity athletes had lower personal Matchbox “returns” to being varsity athletes than non-female varsity athletes did. Perhaps, for example, female varsity athletes were required or encouraged to go to the Matchbox less than their male counterparts were. We tested this possibility by adding a “female x varsityAthlete” interaction variable to the personal Matchbox regression above. While the coefficient (-0.145) was negative, as we expected, the p-value was quite high (0.889), so we were unable to draw statistically significant conclusions on this hypothesized interaction effect. However, a larger sample size, particularly one that included more female varsity athletes, might have produced more definitive conclusions.

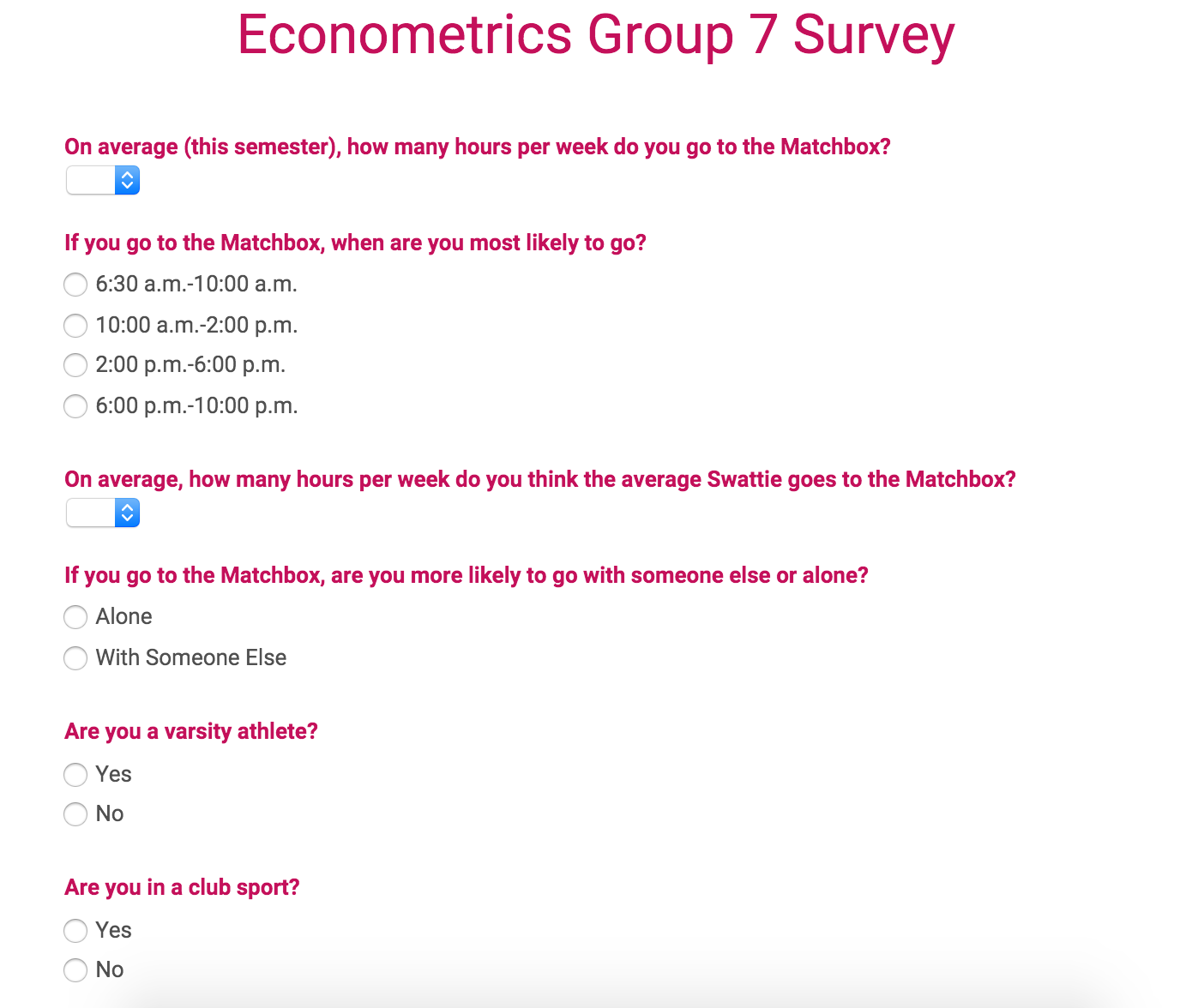
Third and finally, we decided to regress presidential candidate preference with the other variables we collected. In light of the large number of individuals currently running for president, we decided to simplify the original variable to a dummy variable, “hillary,” for which we assigned a 1 for support for Hillary Clinton and a 0 for support of any other candidate. We were amused to discover that at a p-value significance level of 0.053, Hillary Clinton supporters, on average, had a 0.81-point higher self-reported stress level than non-Clinton supporters. Given that respondents reported stress level on a 10-point scale, this amounted to an 8.1% difference in stress. Of course, in this simple regression model, omitted variable bias is a legitimate concern, but none of the variables that we collected seemed to be both associated with support of Clinton *and* impactful on stress levels. For example, identifying as female was associated with a 23% greater likelihood of being a Clinton supporter but did not impact stress levels. Similarly, one year more of class seniority was associated with a 6.3% higher stress level but had no statistically significant correlation with being a Clinton supporter. Further research that collected potential omitted variables (such as, perhaps, general optimism about the state of U.S. politics) would be needed in order to draw definitive conclusions on this correlation.

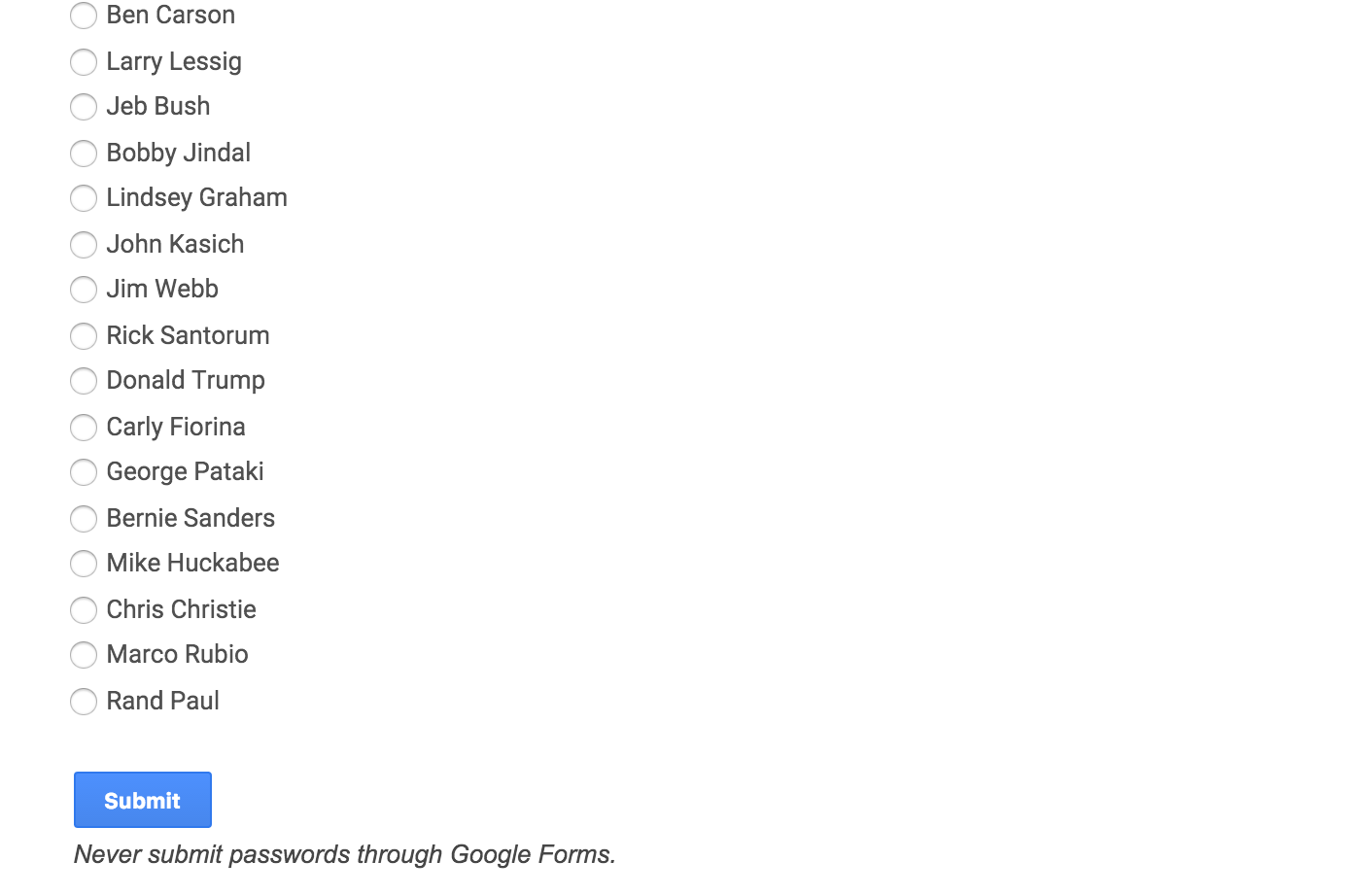
**Conclusion**

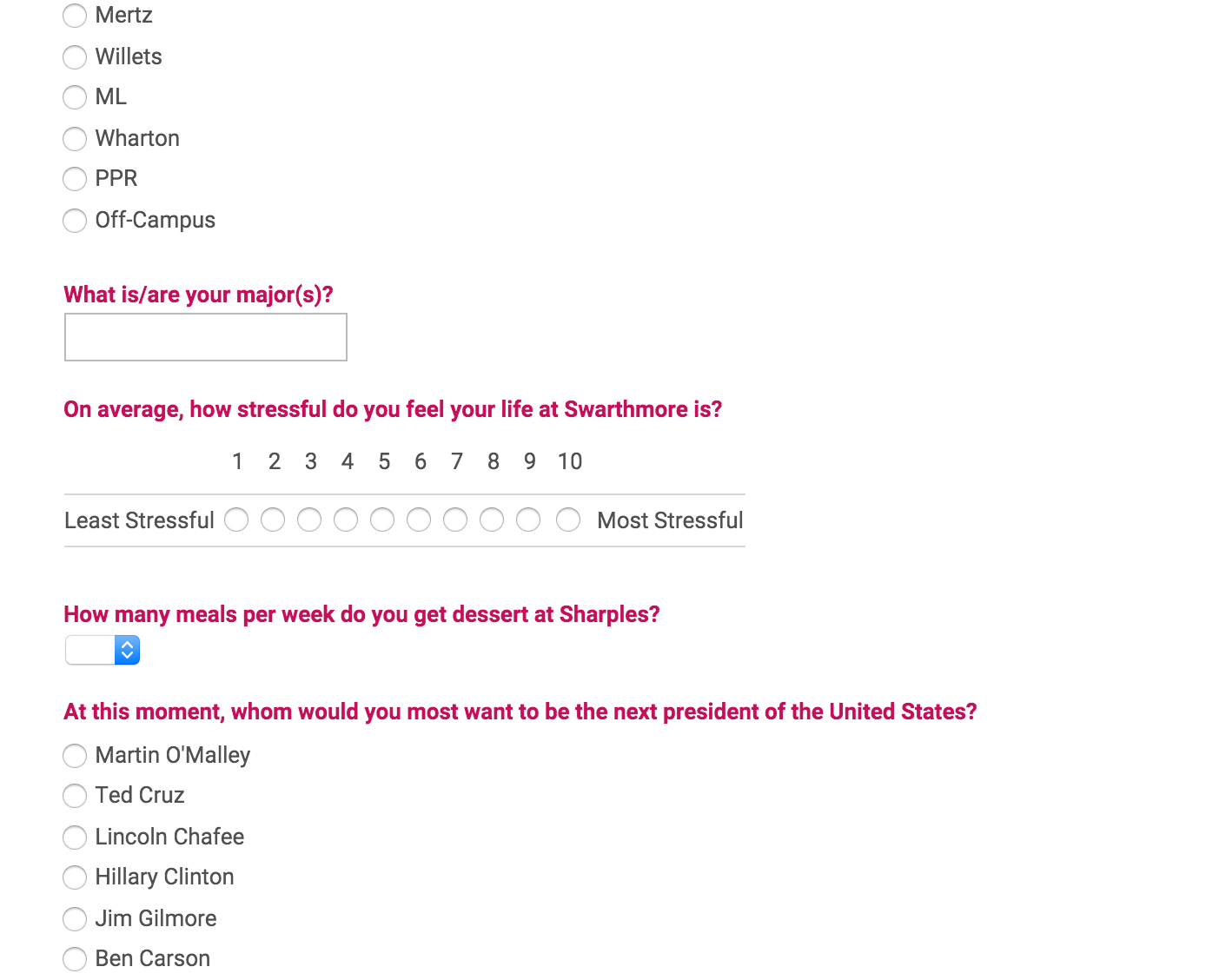
The goal of this project was to show correlations between Swarthmore students’ personal and predicted Matchbox attendance and students’ majors, genders, average meals, dorm proximities, political views, athlete statuses, and other variables. After running the regressions, we found statistical insignificance for most of the variables. However, our regressions did reveal that being a varsity athlete was significantly correlated with spending more time at the Matchbox. As discussed, we also found significant correlations between gender and Matchbox attendance, gender and the difference variable, and Clinton support and stress level.

This project leaves a lot of possibilities for further research. For example, future surveys could distinguish between athletes’ required Matchbox time and personal Matchbox time. There are also other variables, such as current course load, that could have been correlated with the Matchbox variables but that we did not collect in this project. With respect to the Clinton-stress level correlation, as discussed, future research should aim to collect potential omitted variables. Finally, as frequently alluded to, a larger sample size could have yielded a greater number of significant regression results and increased our confidence in the already significant regressions.

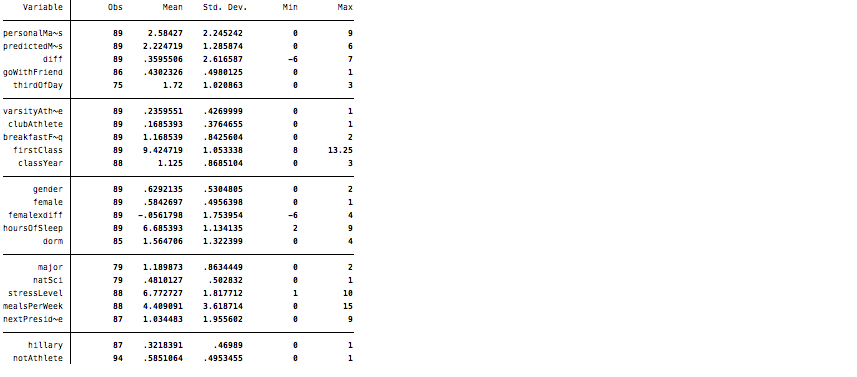
**Appendix 1: Google Survey**







**Appendix 2: Data Summary**



**Appendix 3: STATA Hypothesis Tests**

**ttest personalMa~s==2**

One-sample t test

------------------------------------------------------------------------------

Variable | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]

---------+--------------------------------------------------------------------

person~s | 89 2.58427 .2379952 2.245242 2.111304 3.057235

------------------------------------------------------------------------------

mean = mean(personalMatchboxHours) t = 2.4550

Ho: mean = 2 degrees of freedom = 88

Ha: mean < 2 Ha: mean != 2 Ha: mean > 2

Pr(T < t) = 0.9920 Pr(|T| > |t|) = 0.0161 Pr(T > t) = 0.0080

**ttest predictedM~s==2**

One-sample t test

------------------------------------------------------------------------------

Variable | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]

---------+--------------------------------------------------------------------

predic~s | 89 2.224719 .1363024 1.285874 1.953847 2.495592

------------------------------------------------------------------------------

mean = mean(predictedMatchboxHours) t = 1.6487

Ho: mean = 2 degrees of freedom = 88

Ha: mean < 2 Ha: mean != 2 Ha: mean > 2

Pr(T < t) = 0.9486 Pr(|T| > |t|) = 0.1028 Pr(T > t) = 0.0514

**ttest diff==0**

One-sample t test

------------------------------------------------------------------------------

Variable | Obs Mean Std. Err. Std. Dev. [95% Conf. Interval]

---------+--------------------------------------------------------------------

diff | 89 .3595506 .2773577 2.616587 -.1916395 .9107406

------------------------------------------------------------------------------

mean = mean(diff) t = 1.2963

Ho: mean = 0 degrees of freedom = 88

Ha: mean < 0 Ha: mean != 0 Ha: mean > 0

Pr(T < t) = 0.9009 Pr(|T| > |t|) = 0.1982 Pr(T > t) = 0.0991

**ttest personalMatchboxHours, by(varsityAthlete) level(99)**

Two-sample t test with equal variances

------------------------------------------------------------------------------

Group | Obs Mean Std. Err. Std. Dev. [99% Conf. Interval]

---------+--------------------------------------------------------------------

0 | 68 2.088235 .2563806 2.114169 1.408514 2.767957

1 | 21 4.190476 .417584 1.91361 3.002308 5.378645

---------+--------------------------------------------------------------------

combined | 89 2.58427 .2379952 2.245242 1.957662 3.210877

---------+--------------------------------------------------------------------

diff | -2.102241 .5167209 -3.463039 -.7414423

------------------------------------------------------------------------------

diff = mean(0) - mean(1) t = -4.0684

Ho: diff = 0 degrees of freedom = 87

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0

Pr(T < t) = 0.0001 Pr(|T| > |t|) = 0.0001 Pr(T > t) = 0.9999

**Appendix 4: STATA Regression Outputs**

1. Insignificant Regressions With Very High P-Values

**regress personalMa~s breakfastFreq**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 0.55

Model | 2.80430846 1 2.80430846 Prob > F = 0.4589

Residual | 440.813669 87 5.06682378 R-squared = 0.0063

-------------+------------------------------ Adj R-squared = -0.0051

Total | 443.617978 88 5.04111338 Root MSE = 2.251

---------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+-------------------------------------------------------------

breakfastF~q | .2118705 .2847905 0.74 0.459 -.3541815 .7779225

\_cons | 2.336691 .4094863 5.71 0.000 1.522792 3.150589

---------------------------------------------------------------------------

**regress personalMa~s firstClass**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 1.17

Model | 5.89670237 1 5.89670237 Prob > F = 0.2820

Residual | 437.721275 87 5.03127902 R-squared = 0.0133

-------------+------------------------------ Adj R-squared = 0.0020

Total | 443.617978 88 5.04111338 Root MSE = 2.2431

------------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

firstClass | .2457512 .2270023 1.08 0.282 -.2054404 .6969427

\_cons | .268134 2.152604 0.12 0.901 -4.010399 4.546667

------------------------------------------------------------------------------

**regress personalMa~s classYear**

Source | SS df MS Number of obs = 88

-------------+------------------------------ F( 1, 86) = 0.20

Model | 1.03714286 1 1.03714286 Prob > F = 0.6521

Residual | 435.826494 86 5.06774992 R-squared = 0.0024

-------------+------------------------------ Adj R-squared = -0.0092

Total | 436.863636 87 5.02142111 Root MSE = 2.2512

------------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

classYear | .1257143 .27789 0.45 0.652 -.4267128 .6781413

\_cons | 2.472208 .3941107 6.27 0.000 1.688742 3.255674

------------------------------------------------------------------------------

**regress personalMa~s hoursOfSleep**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 0.13

Model | .659569746 1 .659569746 Prob > F = 0.7198

Residual | 442.958408 87 5.09147595 R-squared = 0.0015

-------------+------------------------------ Adj R-squared = -0.0100

Total | 443.617978 88 5.04111338 Root MSE = 2.2564

------------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

hoursOfSleep | -.0763351 .2120879 -0.36 0.720 -.4978828 .3452125

\_cons | 3.0946 1.437923 2.15 0.034 .2365724 5.952628

------------------------------------------------------------------------------

**regress personalMa~s dorm**

Source | SS df MS Number of obs = 85

-------------+------------------------------ F( 1, 83) = 0.01

Model | .036939254 1 .036939254 Prob > F = 0.9317

Residual | 414.716002 83 4.99657834 R-squared = 0.0001

-------------+------------------------------ Adj R-squared = -0.0120

Total | 414.752941 84 4.93753501 Root MSE = 2.2353

------------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

dorm | .0158578 .1844311 0.09 0.932 -.3509683 .3826838

\_cons | 2.551658 .376911 6.77 0.000 1.801997 3.301319

------------------------------------------------------------------------------

**regress personalMa~s stressLevel**

Source | SS df MS Number of obs = 88

-------------+------------------------------ F( 1, 86) = 0.13

Model | .686281697 1 .686281697 Prob > F = 0.7152

Residual | 440.393264 86 5.1208519 R-squared = 0.0016

-------------+------------------------------ Adj R-squared = -0.0101

Total | 441.079545 87 5.06987983 Root MSE = 2.2629

------------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

stressLevel | .0488615 .1334709 0.37 0.715 -.2164698 .3141928

\_cons | 2.271347 .9355953 2.43 0.017 .4114454 4.131249

------------------------------------------------------------------------------

**regress personalMa~s natSci**

Source | SS df MS Number of obs = 79

-------------+------------------------------ F( 1, 77) = 0.09

Model | .487683008 1 .487683008 Prob > F = 0.7604

Residual | 400.930039 77 5.20688362 R-squared = 0.0012

-------------+------------------------------ Adj R-squared = -0.0118

Total | 401.417722 78 5.14638105 Root MSE = 2.2819

------------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

natSci | .1572529 .5138292 0.31 0.760 -.8659117 1.180417

\_cons | 2.658537 .3563666 7.46 0.000 1.94892 3.368153

------------------------------------------------------------------------------

**regress diff stressLevel**

Source | SS df MS Number of obs = 88

-------------+------------------------------ F( 1, 86) = 0.36

Model | 2.48507849 1 2.48507849 Prob > F = 0.5503

Residual | 594.378558 86 6.91137858 R-squared = 0.0042

-------------+------------------------------ Adj R-squared = -0.0074

Total | 596.863636 87 6.86050157 Root MSE = 2.629

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

stressLevel | .0929791 .1550593 0.60 0.550 -.2152685 .4012268

\_cons | -.2433586 1.086924 -0.22 0.823 -2.404092 1.917375

------------------------------------------------------------------------------

**regress diff thirdOfDay**

Source | SS df MS Number of obs = 75

-------------+------------------------------ F( 1, 73) = 0.86

Model | 4.88018672 1 4.88018672 Prob > F = 0.3577

Residual | 415.78648 73 5.6957052 R-squared = 0.0116

-------------+------------------------------ Adj R-squared = -0.0019

Total | 420.666667 74 5.68468468 Root MSE = 2.3866

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

thirdOfDay | -.251556 .2717629 -0.93 0.358 -.7931788 .2900668

\_cons | 1.36601 .5426193 2.52 0.014 .2845711 2.447448

------------------------------------------------------------------------------

**regress diff classYear**

Source | SS df MS Number of obs = 88

-------------+------------------------------ F( 1, 86) = 0.14

Model | 1.00595238 1 1.00595238 Prob > F = 0.7050

Residual | 599.619048 86 6.97231451 R-squared = 0.0017

-------------+------------------------------ Adj R-squared = -0.0099

Total | 600.625 87 6.90373563 Root MSE = 2.6405

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

classYear | -.1238095 .3259521 -0.38 0.705 -.7717809 .5241619

\_cons | .5142857 .4622736 1.11 0.269 -.4046839 1.433255

------------------------------------------------------------------------------

**regress diff firstClass**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 0.02

Model | .108113108 1 .108113108 Prob > F = 0.9008

Residual | 602.386269 87 6.9239801 R-squared = 0.0002

-------------+------------------------------ Adj R-squared = -0.0113

Total | 602.494382 88 6.84652707 Root MSE = 2.6313

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

firstClass | .0332759 .2662986 0.12 0.901 -.4960214 .5625732

\_cons | .0459343 2.525241 0.02 0.986 -4.973255 5.065124

------------------------------------------------------------------------------

**regress diff dorm**

Source | SS df MS Number of obs = 85

-------------+------------------------------ F( 1, 83) = 0.23

Model | 1.48403859 1 1.48403859 Prob > F = 0.6345

Residual | 541.104197 83 6.51932767 R-squared = 0.0027

-------------+------------------------------ Adj R-squared = -0.0093

Total | 542.588235 84 6.45938375 Root MSE = 2.5533

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

dorm | -.1005126 .2106683 -0.48 0.635 -.5195234 .3184982

\_cons | .5690373 .4305305 1.32 0.190 -.2872705 1.425345

------------------------------------------------------------------------------

**regress diff hoursOfSleep**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 0.02

Model | .136629392 1 .136629392 Prob > F = 0.8886

Residual | 602.357753 87 6.92365233 R-squared = 0.0002

-------------+------------------------------ Adj R-squared = -0.0113

Total | 602.494382 88 6.84652707 Root MSE = 2.6313

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

hoursOfSleep | -.0347429 .2473214 -0.14 0.889 -.5263209 .4568351

\_cons | .5918205 1.6768 0.35 0.725 -2.741002 3.924643

------------------------------------------------------------------------------

**regress diff mealsPerWeek**

Source | SS df MS Number of obs = 88

-------------+------------------------------ F( 1, 86) = 0.03

Model | .227265473 1 .227265473 Prob > F = 0.8568

Residual | 596.636371 86 6.93763222 R-squared = 0.0004

-------------+------------------------------ Adj R-squared = -0.0112

Total | 596.863636 87 6.86050157 Root MSE = 2.6339

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

mealsPerWeek | .0141238 .0780354 0.18 0.857 -.1410054 .1692531

\_cons | .3240903 .4440919 0.73 0.468 -.5587352 1.206916

------------------------------------------------------------------------------

**regress diff natSci**

Source | SS df MS Number of obs = 79

-------------+------------------------------ F( 1, 77) = 0.03

Model | .229310541 1 .229310541 Prob > F = 0.8581

Residual | 548.985879 77 7.12968674 R-squared = 0.0004

-------------+------------------------------ Adj R-squared = -0.0126

Total | 549.21519 78 7.04122038 Root MSE = 2.6701

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

natSci | .1078306 .6012638 0.18 0.858 -1.089439 1.3051

\_cons | .3658537 .417007 0.88 0.383 -.4645133 1.196221

------------------------------------------------------------------------------

**regress predictedM~s goWithFriend**

Source | SS df MS Number of obs = 86

-------------+------------------------------ F( 1, 84) = 1.07

Model | 1.76775613 1 1.76775613 Prob > F = 0.3044

Residual | 139.081081 84 1.65572716 R-squared = 0.0126

-------------+------------------------------ Adj R-squared = 0.0008

Total | 140.848837 85 1.65704514 Root MSE = 1.2868

------------------------------------------------------------------------------

predictedM~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

goWithFriend | .2895753 .2802494 1.03 0.304 -.2677314 .846882

\_cons | 2.142857 .1838215 11.66 0.000 1.777308 2.508406

------------------------------------------------------------------------------

1. Regressions With P-Values between 0.1 and 0.2

**regress predictedM~s firstClass**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 2.72

Model | 4.40792998 1 4.40792998 Prob > F = 0.1028

Residual | 141.097688 87 1.62181251 R-squared = 0.0303

-------------+------------------------------ Adj R-squared = 0.0191

Total | 145.505618 88 1.65347293 Root MSE = 1.2735

------------------------------------------------------------------------------

predictedM~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

firstClass | .2124752 .1288817 1.65 0.103 -.0436911 .4686416

\_cons | .2221997 1.222152 0.18 0.856 -2.20696 2.651359

------------------------------------------------------------------------------

**regress diff goWithFriend**

Source | SS df MS Number of obs = 86

-------------+------------------------------ F( 1, 84) = 2.50

Model | 17.2290627 1 17.2290627 Prob > F = 0.1178

Residual | 579.526751 84 6.89912799 R-squared = 0.0289

-------------+------------------------------ Adj R-squared = 0.0173

Total | 596.755814 85 7.02065663 Root MSE = 2.6266

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

goWithFriend | -.9040265 .5720677 -1.58 0.118 -2.041646 .2335929

\_cons | .7959184 .3752313 2.12 0.037 .0497298 1.542107

------------------------------------------------------------------------------

1. Statistically Significant Regressions

**regress personalM~s varsityAthlete**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 16.55

Model | 70.9092941 1 70.9092941 Prob > F = 0.0001

Residual | 372.708683 87 4.28400786 R-squared = 0.1598

-------------+------------------------------ Adj R-squared = 0.1502

Total | 443.617978 88 5.04111338 Root MSE = 2.0698

-------------------------------------------------------------------------------

personalMatc~s | Coef. Std. Err. t P>|t| [95% Conf.Interval]

---------------+---------------------------------------------------------------

varsityAthlete | 2.102241 .5167209 4.07 0.000 1.075202 3.12928

\_cons | 2.088235 .2509982 8.32 0.000 1.589349 2.587121

-------------------------------------------------------------------------------

**regress diff varsityAthlete**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 17.72

Model | 101.973374 1 101.973374 Prob > F = 0.0001

Residual | 500.521008 87 5.75311504 R-squared = 0.1693

-------------+------------------------------ Adj R-squared = 0.1597

Total | 602.494382 88 6.84652707 Root MSE = 2.3986

-------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf.Interval]

---------------+---------------------------------------------------------------

varsityAthlete | 2.521008 .5988009 4.21 0.000 1.330827 3.71119

\_cons | -.2352941 .2908688 -0.81 0.421 -.8134273 .342839

-------------------------------------------------------------------------------

**regress personalMa~s female**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 4.80

Model | 23.1730711 1 23.1730711 Prob > F = 0.0312

Residual | 420.444906 87 4.83270007 R-squared = 0.0522

-------------+------------------------------ Adj R-squared = 0.0413

Total | 443.617978 88 5.04111338 Root MSE = 2.1983

------------------------------------------------------------------------------

personalMa~s | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

female | -1.035343 .4728108 -2.19 0.031 -1.975106 -.0955804

\_cons | 3.189189 .3614049 8.82 0.000 2.470858 3.907521

------------------------------------------------------------------------------

**regress diff female**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 1, 87) = 3.92

Model | 25.9751513 1 25.9751513 Prob > F = 0.0509

Residual | 576.519231 87 6.62665782 R-squared = 0.0431

-------------+------------------------------ Adj R-squared = 0.0321

Total | 602.494382 88 6.84652707 Root MSE = 2.5742

------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

female | -1.096154 .5536557 -1.98 0.051 -2.196604 .0042967

\_cons | 1 .4232007 2.36 0.020 .1588427 1.841157

------------------------------------------------------------------------------

**regress stressLevel hillary**

Source | SS df MS Number of obs = 86

-------------+------------------------------ F( 1, 84) = 3.87

Model | 12.0904612 1 12.0904612 Prob > F = 0.0526

Residual | 262.711864 84 3.1275222 R-squared = 0.0440

-------------+------------------------------ Adj R-squared = 0.0326

Total | 274.802326 85 3.23296854 Root MSE = 1.7685

------------------------------------------------------------------------------

stressLevel | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

hillary | .8079096 .4109051 1.97 0.053 -.0092203 1.625039

\_cons | 6.525424 .2302365 28.34 0.000 6.067573 6.983274

------------------------------------------------------------------------------

***Looking for possible omitted variables in above regression:***

regress hillary female

Source | SS df MS Number of obs = 87

-------------+------------------------------ F( 1, 85) = 5.39

Model | 1.13283007 1 1.13283007 Prob > F = 0.0226

Residual | 17.8556757 85 .210066773 R-squared = 0.0597

-------------+------------------------------ Adj R-squared = 0.0486

Total | 18.9885057 86 .220796578 Root MSE = .45833

------------------------------------------------------------------------------

hillary | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

female | .2308108 .0993922 2.32 0.023 .0331924 .4284292

\_cons | .1891892 .0753491 2.51 0.014 .0393751 .3390033

------------------------------------------------------------------------------

regress stressLevel classYear

Source | SS df MS Number of obs = 87

-------------+------------------------------ F( 1, 85) = 8.56

Model | 26.2864965 1 26.2864965 Prob > F = 0.0044

Residual | 261.115802 85 3.07195062 R-squared = 0.0915

-------------+------------------------------ Adj R-squared = 0.0808

Total | 287.402299 86 3.3418872 Root MSE = 1.7527

------------------------------------------------------------------------------

stressLevel | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

classYear | .6329713 .2163838 2.93 0.004 .2027423 1.0632

\_cons | 6.057113 .3077666 19.68 0.000 5.44519 6.669035

------------------------------------------------------------------------------

regress stressLevel female

Source | SS df MS Number of obs = 88

-------------+------------------------------ F( 1, 86) = 0.30

Model | .982897336 1 .982897336 Prob > F = 0.5884

Residual | 286.471648 86 3.33106568 R-squared = 0.0034

-------------+------------------------------ Adj R-squared = -0.0082

Total | 287.454545 87 3.30407524 Root MSE = 1.8251

------------------------------------------------------------------------------

stressLevel | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

female | .2140964 .3941368 0.54 0.588 -.5694216 .9976145

\_cons | 6.648649 .300048 22.16 0.000 6.052173 7.245124

------------------------------------------------------------------------------

regress hillary classYear

Source | SS df MS Number of obs = 86

-------------+------------------------------ F( 1, 84) = 0.83

Model | .183775785 1 .183775785 Prob > F = 0.3662

Residual | 18.6999451 84 .222618395 R-squared = 0.0097

-------------+------------------------------ Adj R-squared = -0.0021

Total | 18.8837209 85 .222161423 Root MSE = .47182

------------------------------------------------------------------------------

hillary | Coef. Std. Err. t P>|t| [95% Conf. Interval]

-------------+----------------------------------------------------------------

classYear | .0537575 .0591665 0.91 0.366 -.0639015 .1714166

\_cons | .2649479 .0839169 3.16 0.002 .0980699 .4318259

------------------------------------------------------------------------------

1. Joint Regressions

**regress diff female varsityAthlete**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 2, 86) = 10.38

Model | 117.144257 2 58.5721285 Prob > F = 0.0001

Residual | 485.350125 86 5.6436061 R-squared = 0.1944

-------------+------------------------------ Adj R-squared = 0.1757

Total | 602.494382 88 6.84652707 Root MSE = 2.3756

-------------------------------------------------------------------------------

diff | Coef. Std. Err. t P>|t| [95% Conf.Interval]

---------------+---------------------------------------------------------------

female | -.8440089 .514778 -1.64 0.105 -1.867354 .1793358

varsityAthlete | 2.401618 .5975282 4.02 0.000 1.213771 3.589464

\_cons | .2860055 .4290537 0.67 0.507 -.5669251 1.138936

-------------------------------------------------------------------------------

**regress personalM~s female varsityAthlete**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 2, 86) = 10.26

Model | 85.4717195 2 42.7358597 Prob > F = 0.0001

Residual | 358.146258 86 4.16449137 R-squared = 0.1927

-------------+------------------------------ Adj R-squared = 0.1739

Total | 443.617978 88 5.04111338 Root MSE = 2.0407

-------------------------------------------------------------------------------

personalMatc~s | Coef. Std. Err. t P>|t| [95% Conf.Interval]

---------------+---------------------------------------------------------------

female | -.8269104 .4422039 -1.87 0.065 -1.705983 .0521619

varsityAthlete | 1.985269 .5132879 3.87 0.000 .9648864 3.005652

\_cons | 2.598974 .3685652 7.05 0.000 1.866291 3.331657

-------------------------------------------------------------------------------

**gen femaleVarsityAthlete = female \* varsityAthlete**

(5 missing values generated)

**regress personalMa~s female varsityAthlete femaleVarsityAthlete**

Source | SS df MS Number of obs = 89

-------------+------------------------------ F( 3, 85) = 6.77

Model | 85.5549406 3 28.5183135 Prob > F = 0.0004

Residual | 358.063037 85 4.21250632 R-squared = 0.1929

-------------+------------------------------ Adj R-squared = 0.1644

Total | 443.617978 88 5.04111338 Root MSE = 2.0524

-------------------------------------------------------------------------------

personalMatchboxHo~s | Coef. Std. Err. t P>|t| [95% Conf.Interval]

---------------------+--------------------------------------------------------

female | -.7912088 .5121692 -1.54 0.126 -1.809538 .2271207

varsityAthlete | 2.059441 .7382239 2.79 0.007 .5916536 3.527228

femaleVarsityAthlete | -.1451548 1.032726 -0.14 0.889 -2.198491 1.908181

\_cons | 2.576923 .4025164 6.40 0.000 1.776613 3.377234

------------------------------------------------------------------------------