

Analyzing Student Data as a Measurement of Success for *Boy With A Ball*

Ping Ye, Whitney Richardson, Derek Allen, Aneta Galazka, William Fayson

University of North Georgia, USA

Huijun Yi

Troy University, USA

Contact Email*: ping.ye@ung.edu

Abstract: *This paper analyzes survey responses and student outcomes from the Boy With A Ball (BWAB) program. Decision trees are applied to analyze the Hemingway survey and a set of selection criteria is constructed to identify a group of high-risk students who need help. Meanwhile, pivot tables and multiple linear regressions are utilized to evaluate the mentoring program. The goal is to measure how the program has been executed by the BWAB team and determine whether it will be of great value to the investors. Not only are student academic performances used for evaluating the BWAB mentoring program, but also internal student characteristics are considered. After analyzing student data from both the survey responses and the student academic records, it shows that a sense of connectedness with community-family plays an important role in the BWAB program.*

Keywords: Hemingway, attendance analysis, student characteristics, connectedness

1. Introduction

Boy With A Ball (BWAB) is a non-profit organization which works to make cities better places by reaching young people and equipping them to be leaders capable of transforming their communities. The BWAB program works in multiple locations, both across the nation and globally, with the goal of developing troubled youth and thus develop communities. Team members and volunteers choose a city's most hard to reach neighborhoods and come together in a number of outreach methods that build relationships with every resident in an economically disadvantaged community. Team members, volunteers, and community members then provide mentoring relationships to young people and their family members who are facing moments of

crisis. As community members are identified as facing similar situations, supportive small groups are formed to meet in community homes. Often, these groups include mothers' group, children's group, tutoring center, job skills training group, ESL group and leaders' group. As these teams and volunteers pour out continual outreach that builds on each other, the community is powerfully developed. This paper works on student data for a cross-age mentoring program such that it pairs middle and high school students together with the main goal of preventing both from dropping out of school. One success story of this program includes a school labeled as a "drop out factory" in 2007. Three years later, the BWAB program intervened, targeting students with few or no extra-curricular activities or students considered at high-risk for dropping out of school. This intervention resulted in a 95 percent graduation rate, with most students continuing their education in college.

"The Hemingway" is a questionnaire given to students both as they enter and as they exit the BWAB program. This 57-question survey aims to show the overall level of connectedness of the student. These responses along with the participants' grades, number of absences, and number of discipline referrals are meant to improve with program participation. Improvement in these four areas leads to the overall development of both the participants and their communities.

With the support of PIC Math grant (Preparation for Industrial Careers in Mathematical Sciences), the research group helps BWAB to analyze the 2009-2014 St. Antonio data for the cross-age mentoring program. BWAB relocated their global headquarter from St. Antonio Texas to Atlanta Georgia in July 2013.

2. Experimental Section

The endeavor of the authors for this study was to learn how to design experiments, clean data and utilize R/R Studio/Excel in order to generate and interpret results. Data analyzed included the survey questions, the grades of the aforementioned four major areas of study, the number of absences from school, the frequency of discipline referrals, program attendance, and a range of other tracked data regarding the program's Participants. BWAB provided the authors this data via several Microsoft Excel spreadsheets. The authors also compiled the data into a single spreadsheet for easier reading and interpretation.

2.1 Hemingway Survey

The original "The Hemingway" survey has two versions: The short version contains 57 questions, while the long version has 31 extra questions besides those 57 ones. The authors focus on the 57 questions due to the limited sample size of 220 students. Here, the missing data is treated as survey response "Unclear". Students took the Hemingway survey before entering the BWAB mentoring program and after finishing the program.

2.2 Attendance Analysis

After preparing data from table “BWAB_Att” (Attendance table), simple linear regressions are applied. The goal is to check the effectiveness of the BWAB mentoring program for different student groups. The designated student groups are represented by the following:

AA - original middle school student mentees who joined during 2011 school year

BB - middle school student mentees who joined during 2012 school year

A - original high school student mentors who joined during 2011 school year

B – high school student mentors who joined during 2012 school year

2.3 Multiple Linear Regression

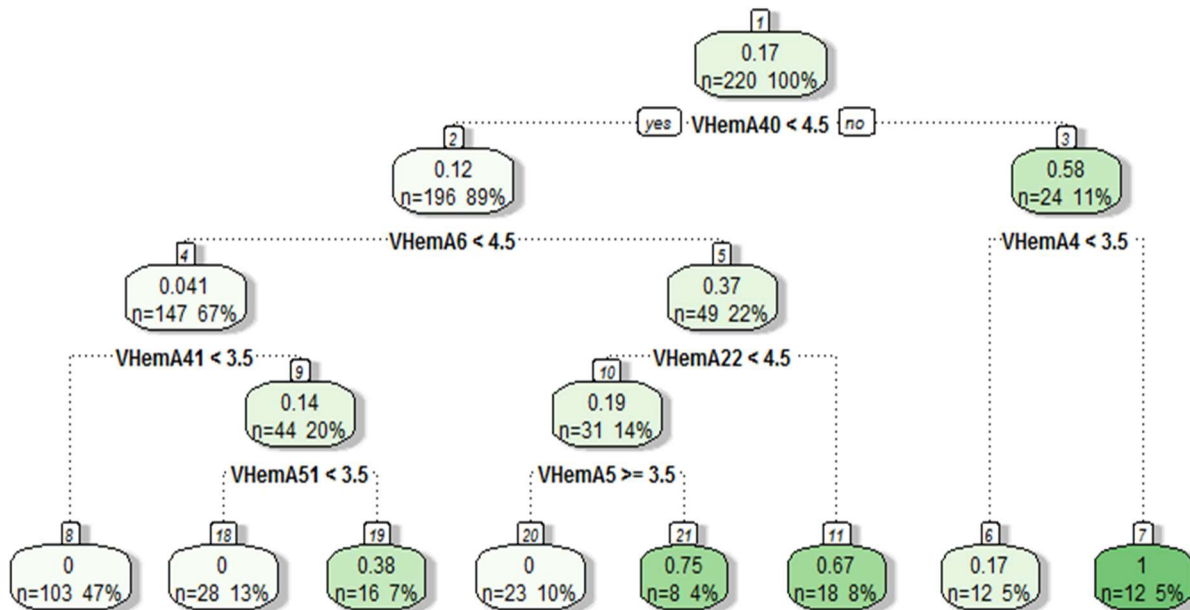
BWAB provided total 18 tables to the research team to investigate. The research team combined these tables into a master table and operated multiple linear regressions. The master table data frame includes the variables year, the grades in four disciplines (math, science, English and social science), attendance, absence and count of incidents. The goal is to identify the disciplines which affect incidents/absences the most.

3. Results and Discussion

3.1 Decision tree Results

3.1.1 Gini index decision tree for program status

The classification is binary: 0 is coded as taking the Hemingway survey before entering the BWAB program, while 1 is coded as taking the Hemingway survey after finishing the BWAB program. The Gini index tree is as follows:



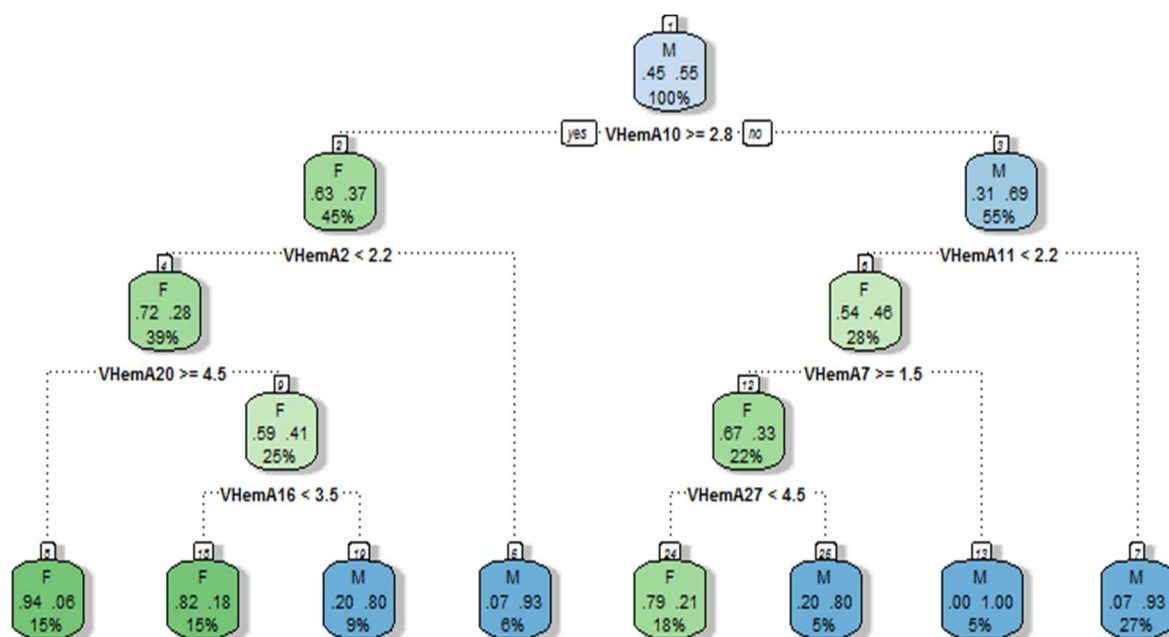
The variable selection from 57 survey questions shows the most important 7 survey questions as the following order:

1. Question # 40 “I often read when I have free time”;
2. Question #6 “I work hard at school”;
3. Question #4 “My family has fun together”;
4. Question #41 “I hang out a lot with kids in my neighborhood”;
5. Question #22 “Spending time with my friends is a big part of my life”;
6. Question # 51 “My neighborhood is boring”;
7. Question # 5 “I have a lot of fun with my brother(s) or sister(s)”.

The above result shows that after finishing the BWAB program, students are more likely to be connected with family members and friends so that they will feel more secured and being loved. The BWAB program improves internal student characteristics, such as knowledge, attitudes, beliefs, self-perception traits like self-esteem or self-efficacy. Since the most important 7 questions are shown on the above Gini index decision tree result, they provide an idea to design a short version survey to identify the high-risk student who needs help.

3.1.2 Gini index decision tree for genders

The classification is binary: male and female. The Gini index tree is as follows:



The variable selection from 57 survey questions shows the most important 7 survey questions for the differences between male and female students. The questions are in the following order:

1. Question # 10 “I enjoy spending time by myself reading”;
2. Question #2 “Spending time with friends is not so important to me”;
3. Question #11 “I spend a lot of time with kids around where I live”;
4. Question #20 “I like to read”;
5. Question #7 “My classmates often bother me”;
6. Question # 16 “I enjoy being at school”;
7. Question # 27 “I like working with my classmates”.

In general, female students who took the survey are more likely to enjoy reading alone, they prefer to have less time at school; while male students are more likely to spend time with friends and are willing to spend more time both at school and after school with other kids.

3.2 Attendance Analysis

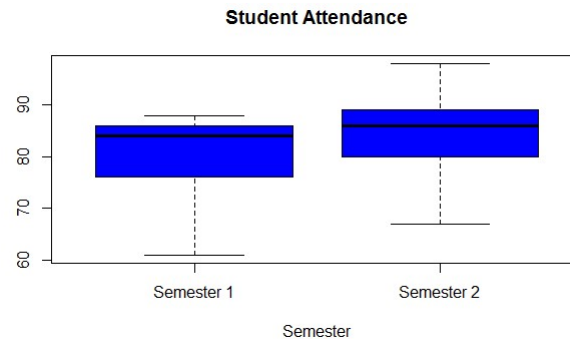
The research team tried to use simple linear regression to model the student attendance pattern.

Model:

$$y = \beta_1 x + \beta_0 ,$$

where y represents the counts of attendance, x refers to years in the program.

R-squared is found to be less than 0.01 after modelling the data, it indicates that simple linear regression model does not fit the data. Then the research team visualized the attendance data of 2009-2014 school years for both semesters by applying the following boxplot:



It shows that the attendance for semester2 is better than the attendance for semester1 in general during those school years. The research team decided to look into the details for attendance of different school years to figure out the impacts of the BWAB mentoring program.

3.2.1 Group AA

Row Labels	Average of PRES_SEM1	Average of PRES_SEM2	Increasing Rate
2009	78.00	83.23	6.7%
2010	74.85	79.28	5.9%
2011	78.00	79.69	5.9%
2012	67.60	70.84	4.8%
2013	62.28	73.17	17.5%
2014	66.98	66.38	-0.9%
Grand Total	70.68	74.92	

After using pivot table to compare the attendance results of two semesters for the group of middle school student mentees who started from 2011, it shows that this group has better attendance in general during semester2. During the school year 2014, the semester2 attendance slightly dropped due to the complexity that the BWAB headquarter was moved from St. Antonio to Atlanta. After attending the BWAB mentoring program, the attendance of middle school student mentees keeps increasing, especially during 2013 school year it increased dramatically due to new methods

adding to the program. Therefore, the BWAB mentoring program has positive impacts on the middle school student mentees. It's evaluated by the increasing of attendance.

3.2.2 Group BB

Row Labels	Average of PRES_SEM1	Average of PRES_SEM2	Increasing Rate
2009	81.25	84.08	3.5%
2010	82.92	86.42	4.2%
2011	79.64	78.29	-1.7%
2012	69.87	78.47	12.3%
2013	53.05	60.37	13.8%
2014	46.47	46.26	-4.5%
Grand Total	66.20	69.73	

The pivot table results of both semesters for the group of middle school student mentees who joined during 2012 school year show that the BWAB mentoring program has positive impacts on the middle school student mentees since student attendance increased dramatically right after the joining year 2012. It's evaluated by the double digits increasing rates. It indicates that the mentoring program works extremely well to reduce the absences for this group.

3.2.3 Group A

Row Labels	Average of PRES_SEM1	Average of PRES_SEM2	Increasing Rate
2009	82.57	83.90	1.6%
2010	79.23	84.13	6.2%
2011	85.63	86.97	1.6%
2012	84.80	86.60	2.1%
2013	69.97	83.16	18.9%
2014	69.94	68.53	-2.0%
Grand Total	77.66	81.20	

After comparing the results for two semesters for the group of original high school student mentors started from 2011 school year, it shows that there are slight increases during 2011 and 2012 school years for the attendance. While there is a significant increase during the 2013 school year due to

new methods added to the program. Therefore, the BWAB mentoring program has positive impacts on attendance of high school student mentors as well.

3.2.4 Group B

The sample size for the group of high school student mentors who joined during 2012 school year is 14, which is too small for modeling the attendance pattern. The attendance analysis for this student group is ignored.

3.3 Multiple Linear Regression Results

Model:

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_0 \quad (*)$$

3.3.1 Count of discipline incidents versus four discipline grades

For the above model (*), y is the count of incidents, x_1 to x_4 represent math, science, English and social study grades, respectively.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	***
(Intercept)	10.0075020	1.5576936	6.425	2.62e-09	
M	-0.0122680	0.0259546	-0.473	0.6373	
S	-0.0413438	0.0287558	-1.438	0.1530	
E	-0.0002648	0.0279868	-0.009	0.9925	
SS	-0.0582040	0.0254775	-2.285	0.0241	*

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.752 on 123 degrees of freedom
(53 observations deleted due to missingness)

Multiple R-squared: 0.2365, Adjusted R-squared: 0.2117

F-statistic: 9.527 on 4 and 123 DF, p-value: 9.625e-07

Here the p-value of the F-test shows that the model is statistical significant. Among four predictors the only p-value less than the significance level 0.05 is the one of the social study grades, which indicates that if the students in the BWAB program improve their social study grades, their count of discipline incidents may drop. Therefore, social study seems to have the largest impact on the count of discipline incidents. Hence BAWB may develop its mentoring program by focusing on improving students' social study grades in order to improve their behaviors at schools. The adjusted R-squared of this model is 0.2117 which indicates that it could be improved by other regression model. This will be part of the future work for the research team to investigate.

3.3.2 Absences versus four discipline grades

For the above model (*), where y is the number of absences, x_1 to x_4 represent math, science, English and social study grades, respectively.

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	***
(Intercept)	46.581928	7.106649	6.555	1.38e-09	
M	-0.184578	0.118412	-1.559	0.1216	
S	-0.246966	0.112206	-2.201	0.0296	*
E	0.000493	0.127684	0.004	0.9969	
SS	-0.048539	0.116236	-0.418	0.6770	

Signif. codes:

0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.995 on 123 degrees of freedom

(53 observations deleted due to missingness)

Multiple R-squared: 0.2298, Adjusted R-squared: 0.2048

F-statistic: 9.176 on 4 and 123 DF, p-value: 1.604e-06

Here the p-value of the F-test shows that the model is statistical significant. Among four predictors the only p-value less than the significance level 0.05 is the one of the science grades, which indicates that if the students in the BWAB program improve their science grades, their number of absences may drop. Therefore, science seems to have the largest positive impact on the class participation. As the social science, BWAB may also focus on improving students' science grades in order to improve their behaviors at schools. Here the adjusted R-squared is 0.2048 which indicates that this model could be improved by other regression model. This will be part of the future work for the research team to investigate.

4. Conclusions

By applying the decision tree, seven questions from the Hemingway survey could be used to redesign a short survey and to identify the students who need help. The results also show that the BWAB mentoring program connects students in the program to satisfy their safety and belonging needs. Attendance analysis results indicate the BWAB has positive impacts on the academic performance and behavior of the student participants. Multiple regression analysis reveals the relationship between the subjects of social science along with science and student behaviors at schools. Connectedness plays an important role once again.

Connectedness within the BWAB program allows students the potential to reach and satisfy a couple of levels through their program. Including safety and security, as well as love and belonging needs. Safety needs focus on the environment including, the home, school, and elsewhere. If a

student is placed in a bad environment or an unsafe neighborhood, such conditions will hinder academic performance. Research suggests that the student will have trouble focusing on learning when he/she does not feel secure (Burlinson and Thoron, 2014). In terms of love and belonging needs, these needs include friendships and relations. This can include a sense of belonging through clubs, volunteer groups, churches, or in this case, the BWAB program. It has been shown that students who feel connected are less likely to engage in risky behaviors, such as, alcohol or drug use as well as violence and gang involvement (McLeod, 2016).

BWAB is able to stay connected and fulfill this safety need for students within school by placing representatives from BWAB inside the schools that the students attend. These representatives act as mentors, who care for the students' best interests. For the love and belonging needs, BWAB allows these at-risk students a place where they can engage with other peers, mentors, and in various events throughout the community. This allows students the opportunity to establish relationships that can last after they have left the BWAB program. With the BWAB program being able to satisfy and guide students to meet these two needs, it gives the students a chance to reach a higher level on their own in terms of performance. These needs are referred to as the Esteem Needs.

Once students in the BWAB program feel as though they belong, they can begin to satisfy and develop positive feelings of self-worth and self-esteem. Esteem Needs involve one's confidence, self-worth, and sense of achievement (Perera, 2015). In terms of achievement, this includes school performance and the students, desire to see themselves achieve good grades. Self-esteem also goes hand in hand with motivation. If a student is not motivated, they are less likely to perform well in school or any activity for that matter. Given the record of the BWAB program, which includes increased academic performance and graduation rates for students who are part of the program, it appears as though the way in which they get these kids connected is working positively in satisfying Maslow's needs (Maslow, 1943) and overall improving the student participants' future potential.

Future Work

The student data analyzed included records from the program in one district in Texas and followed one set of students, who were in middle school and high school. It involved the first group of students participating in the program. Given that the program is expanding nationally and increasing student participation in already existing areas of operation, more data should be analyzed so that BWAB can compare consistency in the results. Meanwhile, the multiple regression shows higher social study and science grades correlate to better school behaviors. Other models using attendance/count of discipline incidents as dependent variable could be built to explore the connection between the grades and school behaviors. Additionally, a universally formatted database should be developed which will streamline the analysis process for future

statistical studies. Furthermore, alternate statistical methods such as logistic regression and cluster should be considered to analyze the information collected in order to improve the model accuracy and efficiency. Boy With A Ball fully relies on the donation from sponsors. We recommend further analysis regarding fundraising methods that could clarify and improve upon the donor recruiting process. This analysis could help the organization reach additional members and increase individual donations.

Acknowledgment

The authors would like to thank the PIC Math program, the Boy With A Ball and the University of North Georgia which made it possible for the research team to work on this project.

References

- [1] An Introduction to Statistical Learning, with Applications in R by *James, Witten, Hastie and Tibshirani*: Springer, 2013.
- [2] Maslow's Hierarchy of Needs and Its Relation to Learning and Achievement. *Burleson, Sarah E., and Andrew C. Thoron*. EDIS New Publications RSS. *Agricultural Education and Communication*, 25 Sept. 2014. Web. 21 Mar. 2017.
- [3] Maslow's Hierarchy of Needs. Simply Psychology. *Saul McLeod*. *Creative Commons*, 16 Sept. 2016. Web. 25 Mar. 2017.
- [4] Maslow and Your Self Esteem Needs. *Karl Perera*. *More-selfesteem.com*. *Better Internet Bureau*, n.d., 12 Jan. 2015. Web. 25 Mar. 2017
- [5] Data Science For Business-What You Need to Know About Data Mining and Data-Analytic Thinking. *Foster Provost and Tom Fawcett*.
- [6] Fundamentals of Machine Learning for Predictive Data Analytics Algorithms, Wored Examples, and Case Studies. *John D Kelleher, Brian Mac Namee, and Aoife D'arcy*.
- [7] Statistical Design and Analysis of Experiments. *Wiley*, NY. 1989.
- [8] Probability and Statistics. *Michael J Evans and Jerrey S Rosenthal*.