

Partial Least Squares Path Modeling to Determine the Success of College Football Teams

Matthew Emmerling
Advisor: Dr. Jebessa Mijena

May 9, 2019



Contents

1	Acknowledgments	3
2	Abstract	4
3	Introduction	5
4	Methodology	6
5	Analysis	8
6	Conclusion	23
7	Glossary	24
8	References	26

1 Acknowledgments

I would like to thank Dr. Mijena for mentoring me as well as providing essential guidance throughout this research. His knowledge of data sciences and statistics was extremely beneficial. I would also like to thank Georgia College for providing the opportunity of performing undergrad research.

2 Abstract

In this paper, we analyze college football data for the purpose of determining what makes a successful college football team. The data used to analyze success was pulled from ESPN and includes 87 initial variables and includes season 2010 to season 2018. We use Partial Least Squares Path modeling, or PLSPM, to determine linear relationships between multiple blocks of variables. PLSPM consists of an outer model and inner (structural) model. While the outer model deals with the relationships between the latent variables and their corresponding block of manifest variables. The structural model describes the relationship between latent variables. This research will show what elements of a college football team contribute most to their success.

3 Introduction

College football is an incredibly popular sport in the United States. With most of the focus being on division 1-A, which most recently consists of 130 teams. Each college football team is made up of three major sub-teams, those being offense, defense, and special teams. Each of the sub-teams contribute to the team in their own way. Knowing how these sub-teams contribute to the overall success of the team can be useful and informative information. This research will analyze the effect that offense, defense, and special teams has on the overall success of college football teams. While also breaking the sub-teams into there own parts and seeing what elements contribute the most. With the use of Partial Least Squares Path-Modeling, PLS-PM, the goal of this research is to see how each sub-team contributes to success and what parts of these sub-teams are the most important. This will be done using a PLS-PM package in RStudio, version 1.1.383. The data used for this analysis consists of 11 seasons of Division I-A college football teams, seasons 2010-2018. It contains 87 initial variables that are divided into there respective sub-teams. The data was collected from ESPN and put into an Excel spread sheet, to be used by RStudio.

There has been other research done involving the Success of college football teams. One study was done on the relationship of recruiting and winning, of a college football team (Caro, Cary A.). They used regression analysis to see the effects that recruiting had for on field success. There was also research done concerning the physical characteristics that can predict functional performance (Mayhew, J. L., et al.). In this research a stepwise multiple regression analysis was used to determine relationship between certain physical characteristics and overall performance. However no research was found that uses PLS-PM, the amount of seasons, and the amount of initial variables, that this research has, to determine what elements of a college football team most contribute to their success.

4 Methodology

Partial Least Squares Path Modeling, PLS-PM, is a useful tool in data analysis, because of its ability to solve the problem of wanting to see the relationships between variables that you can not put a directed numeric value to. These variables are called latent variables and some examples would be customer satisfaction, happiness of college students, or success of a sports team. Even though latent variable can not be directly observed they are made up of and measured by a group of variables that are able to be easily measured, these variable are called manifest variables. PLS-PM will analyze manifest variables that make up the different latent variable and then be able to show the connections that each latent variable has with each other. At its core PLS-PM is used to analyze a system of linear relationships between using multiple blocks of variables.

Every PLS-PM model will consist of an inner model, also known as a structural model, and an outer model, which is also know as a measurement model. The inner model will deal with the relationships between each latent variable. While the outer model will deal with the relationship between the latent variables and there respective block of manifest variables.

The way PLS-PM works can be broken down into three major steps. First, it gets weights in order to compute latent variable scores. Second, it estimates path coefficients for the inner model. Lastly, it obtains the loadings for the outer model. The first step is the most intense but it is vital to process of PLS-PM. To start the first part initial arbitrary weights, w_{jk} , of 1 are set. This will form the following \tilde{w}_k vectors:

$$\begin{aligned} \tilde{w}_1 &= (\tilde{w}_{11} = 1, \tilde{w}_{12} = 1, \dots, \tilde{w}_{1k} = 1) \\ &\quad \vdots \\ \tilde{w}_j &= (\tilde{w}_{j1} = 1, \tilde{w}_{j2} = 1, \dots, \tilde{w}_{jk} = 1) \end{aligned}$$

Next the weights \tilde{w}_k are used to get an external estimation which is expressing a latent variable as a weight sum of of its indicators. This is shown as:

$$Y_j = \sum_k w_{jk} \mathbf{X}_{jk},$$

where X_{jk} is the manifest variables for the score of latent variable Y_j . Then inner weights, e_{ij} , are found by by a correlation between two latent variables that are adjacent to each other. One way of doing this can be represented mathematically as:

$$e_{ij} = cor(Y_j, Y_i).$$

This method of calculating inner weights takes into account the sign direction of the path and the strength of each path in the structural model. For the third part of step 1 internal estimations, Z_j , are calculated. This is done by:

$$Z_j = \sum_{i \leftrightarrow j} e_{ij} \mathbf{Y}_i.$$

Then new outer weights are calculated. This is done based on what types of manifest variables are associated with the latent variable. If the variable are caused by the latent variable then this equation is used:

$$\tilde{w}_{jk} = (Y_j' Y_j')^{-1} (Y_j' X_{jk}).$$

While if a latent variable is formed by the manifest variable then this equation is used:

$$\tilde{w}_j = (X_j' X_j')^{-1} (X_j' Y_j).$$

For the final part the PLS-PM algorithm checks to see if the weights have sufficiently converged. This is done by:

$$|w_{jk}^{S-1} - w_{jk}^S| < 10^{-5}.$$

If the weights do not converge then all of the parts of step 1 are repeated with the new weights. Once the weights are found step 1 of three is finished. Now for step 2 is all about calculating path coefficient estimates, $\hat{\beta}_{ij} = B_{ij}$. The structural coefficients are estimated using multiple least squares regression. Note that

$$Y_j = \sum_{i \leftrightarrow j} \beta_{ji} Y_i,$$

and the least squares solution is:

$$\hat{\beta}_{ji} = (Y_i' Y_i)^{-1} (Y_i' Y_j).$$

Finally stage three is to calculate the loadings. Loadings are simply calculated as correlations between a latent variable and its indicators:

$$\widehat{\lambda}_{jk} = \text{cor}(X_{jk}, Y_j)$$

The above three steps are the heart of PLS-PM and they shows how PLS-PM calculates scores and relationships between variables that are normally unobservable.

PLSPM contains different types of models one of those being models with higher order constructs. These models consist higher order latent variables as well as lower order latent variables, which are just normal latent variables. There are two different types of higher order constructs, one being when the lower order latent variables are conceived by the higher order latent variable, this is known as a molecular model. While the other is when higher order latent variables are composed of the lower order latent variables, this is known as a molar model. The molar model is used for this research. Higher order latent variables do not consist of their own manifest variables. There manifest variables must be obtained using certain methods. The method that was used in this research is know as the repeated indicator approach. This approach takes the manifest variables of all of the lower order latent variables and uses them for the corresponding higher order latent variable. Models with higher order constructs are incredibly useful when want to break a latent variable into multiple sub groups, so that the impact of each individual subgroup can be seen.

5 Analysis

We used multiple graphs, figures, and tables in this research in order to properly interpret the data that was collected.

The first model consists of 17 latent variables 14 of those being first order latent variables and 3 being second order latent variables. Each of the first order latent variables consist of their own unique manifest variables. The manifest variables for with their corresponding first order latent variables are listed below.

Passing : PYds, Pyds.G, PassFd, PAtt, PCompPct, Yds.PA, LongPA, PTd, Int

Rushing : Ryds, RYds.G, RushFd, RushAtt, Ryds.A, LongR, RushTd

Receiving : Rec, RecYds, AvgRec, LongRec, RecTd, RecYds.G

Downs : TotalFd, PenFd, MTDC, AttTDC, PctTDC, MFDC, AttFDC, PctFDC

Penalties : TPen, TPenYds

Tackles : Sacks, Ydsl, SoloT, AstT, TotalT

Pass Defense : PD, PInt, IntYds, LongInt, IntTd

Fumbles Recovered : FRec

Kickoff Returns : KRAtt, KRYds, AvgKR

Punt Returns : PRAAtt, PRYds, AvgPR, LongPR, PRTd

Fieldgoals : Fgm, Fga, FPct, LongF, X1to19A, X1to19M, X20to29A, X20to29M, X30to39A, X30to39M, X40to49A, X40to49M, X50plusA, X50plusM

Extra Point : XPM, XPA, XPct

Punting : Punts, PuntYds, LongP, AvgP

Success : WinH, WinA, Pts, Pts.G, Yds, Yds.G.

The definitions for these manifest variables can be found in the glossary.

	C.alpha	DG.rho	eig.1st	eig.2nd
Passing	0.88	0.92	5.31	1.65
Rushing	0.94	0.96	5.42	0.89
Receiving	0.81	0.88	3.73	1.33
Downs	0.76	0.83	3.16	1.88
Penalties	0.97	0.99	1.94	0.06
Tackles	0.57	0.68	2.08	1.72
PassDefense	0.81	0.87	2.93	0.89
FumblesRecovered	1.00	1.00	1.00	0.00
KickoffReturns	0.68	0.83	1.98	1.02
PuntReturns	0.87	0.91	3.32	0.90
FieldGoals	0.84	0.87	5.00	2.40
ExtraPoint	0.71	0.85	2.06	0.94
Punting	0.45	0.66	2.00	1.07
Offense	0.90	0.89	10.66	7.92
Defense	0.76	0.82	3.67	1.74
SpecialTeams	0.79	0.82	6.01	3.45
Success	0.94	0.95	4.65	0.66

Table 1: First Model Unidimensionality

Table 1 shows three different unidimensionality tests, those being Cronbachs alpha, Dillon-Goldsteins rho, first eigenvalue, and second eigenvalue. All four of these tests are used to see how well a block of manifest variables measures their corresponding latent variable. Cronbachs alpha and Dillon-Goldsteins rho are both supposed to be above 0.7, while the first eigenvalue is supposed to be above 1 and the second eigenvalue is supposed to be below one. These test show that some of the latent variables are having problems with unidimensionality, which will be fixed in the improved model. The second eigenvalue is showing the most unidimensionality problems by far.

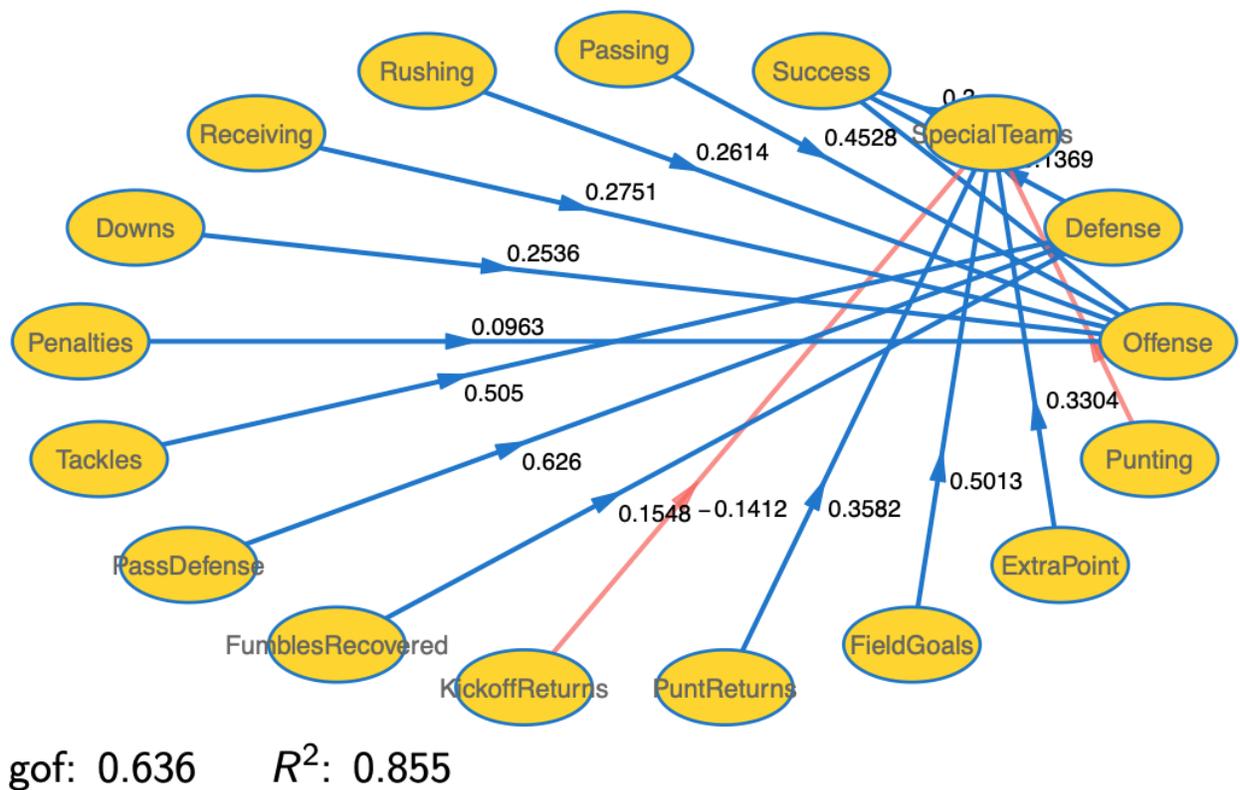


Figure 1: First Model Inner Plot

Figure 1 shows the latent variable that we chose to see the indirect and direct effect that they have on success. With Offense, Defense, and Special Teams being all second order latent variables, such that we can see the direct effect that they have on success. While the rest of the variables being first order latent variables, where they are used to see what parts of Offense, Defense, and Special Teams are most important, and see an indirect effect of success. The first order variables for Offense are Passing, Rushing, Receiving, Downs, and Penalties. Defense has the first order variables Tackles, Pass Defense, and Fumbles Recovered. While Special Teams has Kickoff Returns, Punt Returns, Field Goals, Extra Points, and Punting. This initial model has a goodness of fit, or gof, of 0.636 and an R^2 of 0.855. Which are both fairly good but can and will be improved upon.

	Estimate	Std. Error	t value	Pr(> t)
Intercept	0.00	0.01	0.00	1.00
Offense	0.64	0.01	52.42	0.00
Defense	0.14	0.01	11.78	0.00
SpecialTeams	0.30	0.01	22.41	0.00

Table 2: First model p-values for Success

	Estimate	Std. Error	t value	Pr(> t)
Intercept	-0.00	0.00	-0.00	1.00
Passing	0.45	0.00	145.23	0.00
Rushing	0.26	0.00	380.48	0.00
Receiving	0.28	0.00	87.88	0.00
Downs	0.25	0.00	302.70	0.00
Penalties	0.10	0.00	188.68	0.00

Table 3: First model p-values for Offense

	Estimate	Std. Error	t value	Pr(> t)
Intercept	-0.00	0.00	-0.00	1.00
Tackles	0.50	0.00	723.26	0.00
PassDefense	0.63	0.00	909.95	0.00
FumblesRecovered	0.15	0.00	246.05	0.00

Table 4: First model p-values for Defense

	Estimate	Std. Error	t value	Pr(> t)
Intercept	0.00	0.00	0.00	1.00
KickoffReturns	-0.14	0.00	-180.99	0.00
PuntReturns	0.36	0.00	441.54	0.00
FieldGoals	0.50	0.00	646.32	0.00
ExtraPoint	0.33	0.00	342.12	0.00
Punting	-0.22	0.00	-244.00	0.00

Table 5: First model p-values for Special Teams

Tables 2-5 are showing whether or not the path coefficient are significant or not. Since all of the p-values are less than 0.05 all of the path coefficients are significant.

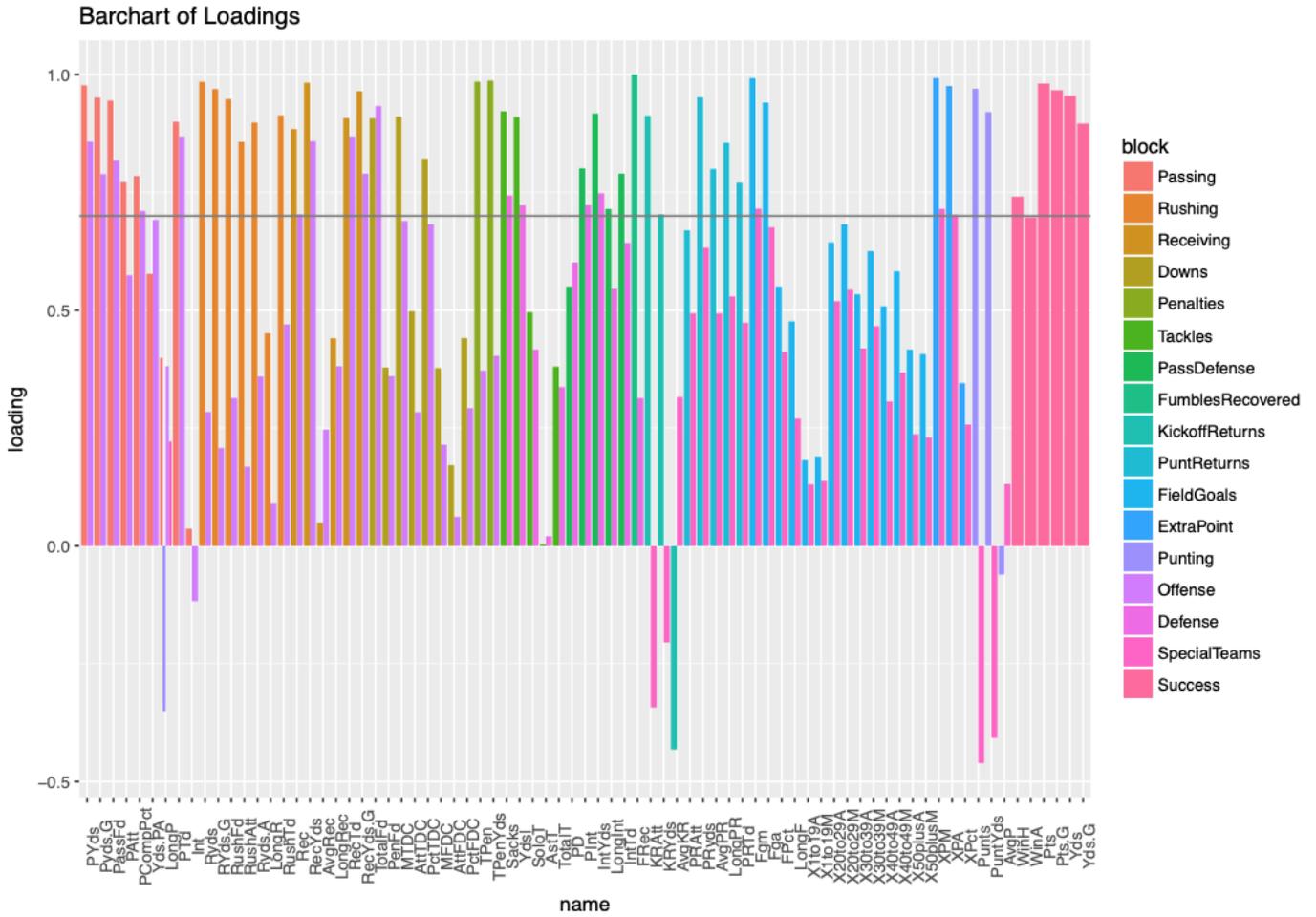


Figure 2: First Model Loadings

We can see the loadings of the first model in Figure 2. Only loadings greater than 0.7 are acceptable. So in order to improve the model we removed variables with loadings lower than 0.7 one at a time until all were greater than 0.7. Furthermore, on improving the model any variable that had a negative weight was changed in order to have positive weights. Also any first order latent variable that were insignificant, having a p-value greater than 0.05, were removed in the improved model, which will be shown next.

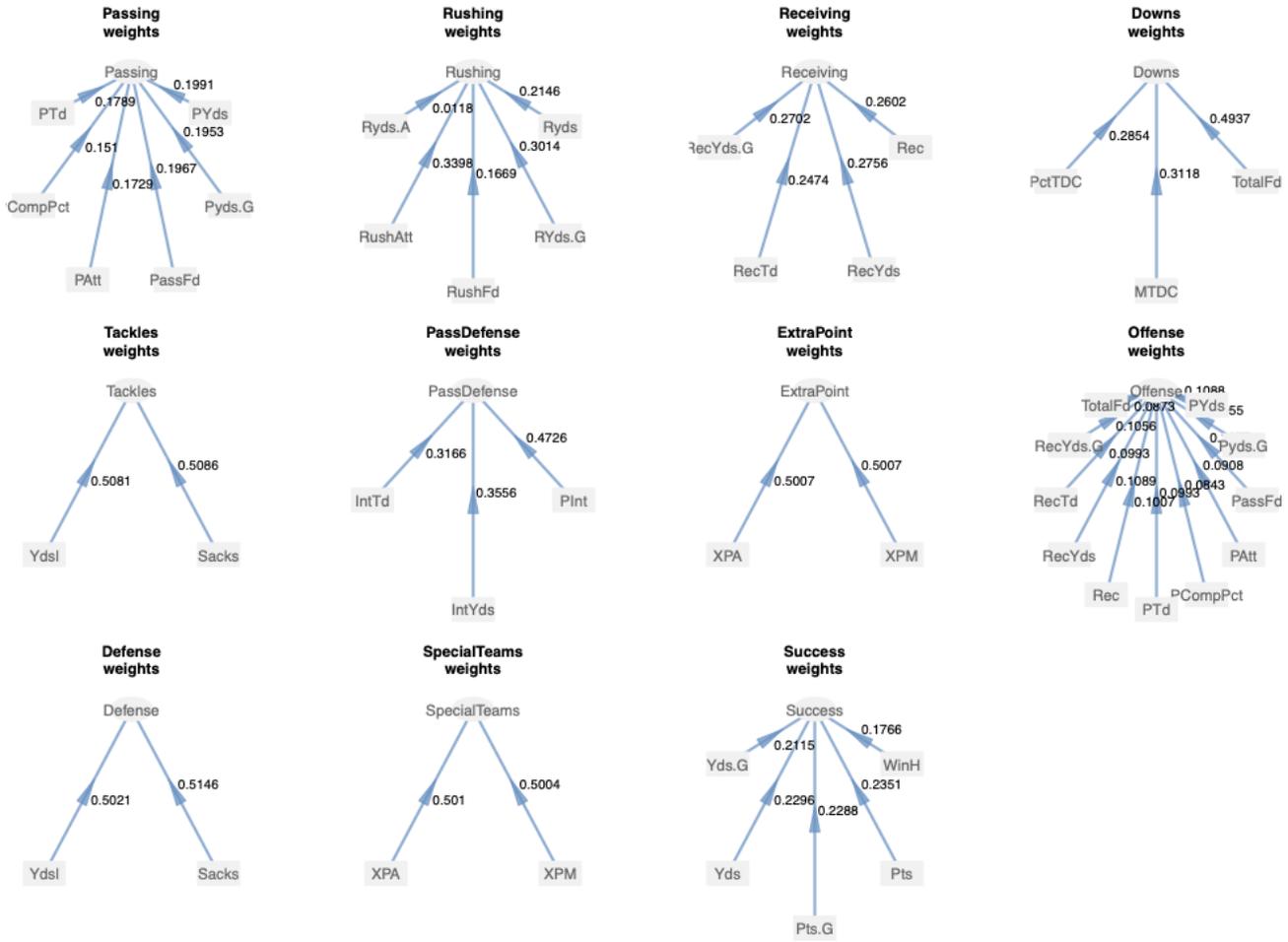


Figure 3: Improved Model Weights

Note that the first model weights could not be displayed as PLS-PM could not produce an output with the initial number of variables in this research. Figure 3 shows the important results of weights of the improved model. These weights allow the calculation of the latent variable scores, as well as show us the more important manifest variables to their corresponding latent variable. For instance TotalFD, total number of first downs, is the most important manifest variable for Downs. Most of the weight are close in value so they share similar importance their respective latent variables. Note that the definition of the manifest variables will be located in the glossary.

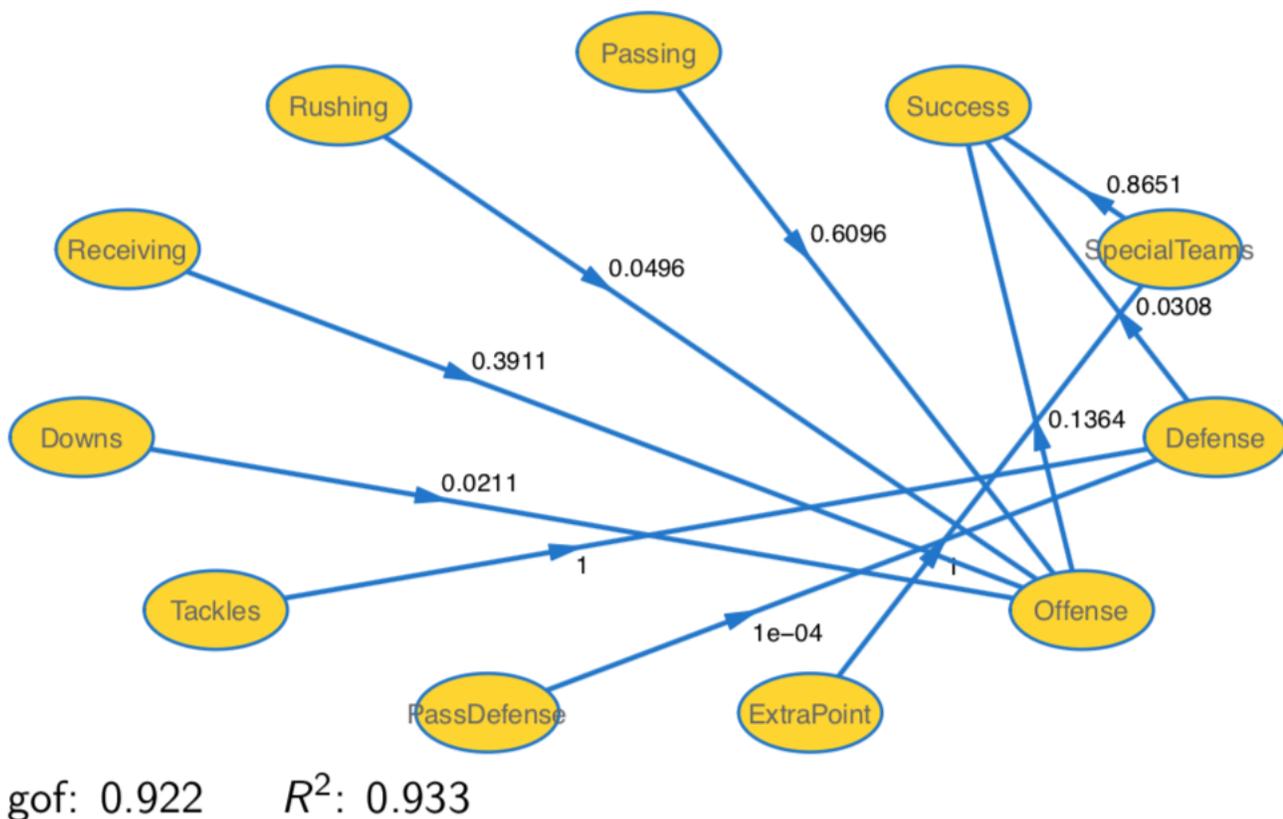


Figure 4: Improved Inner Model

For the improved inner model there were multiple first order latent variables that were removed, but the second order latent variables remain the same. The most important information that Figure 4 provides is the gof of 0.922 and R^2 of 0.933. These figures improved drastically from the first model and both show that our model does a good job of modeling success.

	Mode	MVs	C.alpha	DG.rho	eig.1st	eig.2nd
Passing	A	6	0.96	0.97	4.97	0.55
Rushing	A	5	0.97	0.98	4.48	0.43
Receiving	A	4	0.96	0.97	3.60	0.31
Downs	A	3	0.91	0.94	2.53	0.32
Tackles	A	2	0.97	0.98	1.93	0.07
PassDefense	A	3	0.84	0.91	2.29	0.50
ExtraPoint	A	2	1.00	1.00	1.99	0.01
Offense	A	11	0.98	0.98	9.06	0.90
Defense	A	2	0.97	0.98	1.93	0.07
SpecialTeams	A	2	1.00	1.00	1.99	0.01
Success	A	5	0.95	0.96	4.23	0.59

Table 6: Improved Model Unidimensionality

Note that Table 6 shows the same test as table 1. Except all of the unidimensionality test are past for each latent variable in the improve model. Thus all of the manifest variables are properly measuring there respective latent variables

	Estimate	Std. Error	t value	Pr(> t)
Intercept	0.00	0.00	0.00	1.00
Passing	0.60	0.01	106.63	0.00
Rushing	0.05	0.00	49.80	0.00
Receiving	0.40	0.01	73.88	0.00
Downs	0.02	0.00	18.25	0.00

Table 7: Improved Model p-values for Offense

	Estimate	Std. Error	t value	Pr(> t)
Intercept	-0.00	0.00	-0.00	1.00
Tackles	1.00	0.00	15295.58	0.00
PassDefense	0.00	0.00	1.92	0.05

Table 8: Improved Model p-values for Defense

	Estimate	Std. Error	t value	Pr(> t)
Intercept	0.00	0.00	0.00	1.00
ExtraPoint	1.00	0.00	1115096.17	0.00

Table 9: Improved Model p-values for Special Teams

	Estimate	Std. Error	t value	Pr(> t)
Intercept	0.00	0.01	0.00	1.00
Offense	0.14	0.01	15.23	0.00
Defense	0.03	0.01	3.80	0.00
SpecialTeams	0.87	0.01	89.31	0.00

Table 10: Improved Model p-values for Success

Just like tables 2-5, tables 7-10 also show that every path coefficient is significant. Since all p-values are lower than 0.05. These tables also give us the ability to produce formulas for each of latent variables. The formulas are as follows:

$$\begin{aligned}
 \text{Offense} &= 0.60(\text{Passing})+0.05(\text{Rushing})+0.40(\text{Receiving})+0.02(\text{Downs}), \\
 \text{Defense} &= 1.00(\text{Tackles}), \\
 \text{Special Teams} &= 1.00(\text{ExtraPoint}), \text{ and} \\
 \text{Success} &= 0.14(\text{Offense})+0.03(\text{Defense})+0.87(\text{SpecialTeams}).
 \end{aligned}$$

These equations truly show what makes up each latent variable and what elements are the most important. For Offense we can see that Passing and Receiving contribute most to the offense of a team. While Defense and Special Teams are determined by an individual variable. When analyzing Success we can see that Special Teams contributes the most with Offense coming in second. Note that in an alternate model we tested the effect of Offense on Special Teams and found that Offense had a significant effect on Special Teams. However this alternate model had a lower gof and R^2 , which is why the model is not being used. The current model is still correct in showing that Special Teams has a large effect on Success, but it is important to note that the other variables are still significant to determining Success.

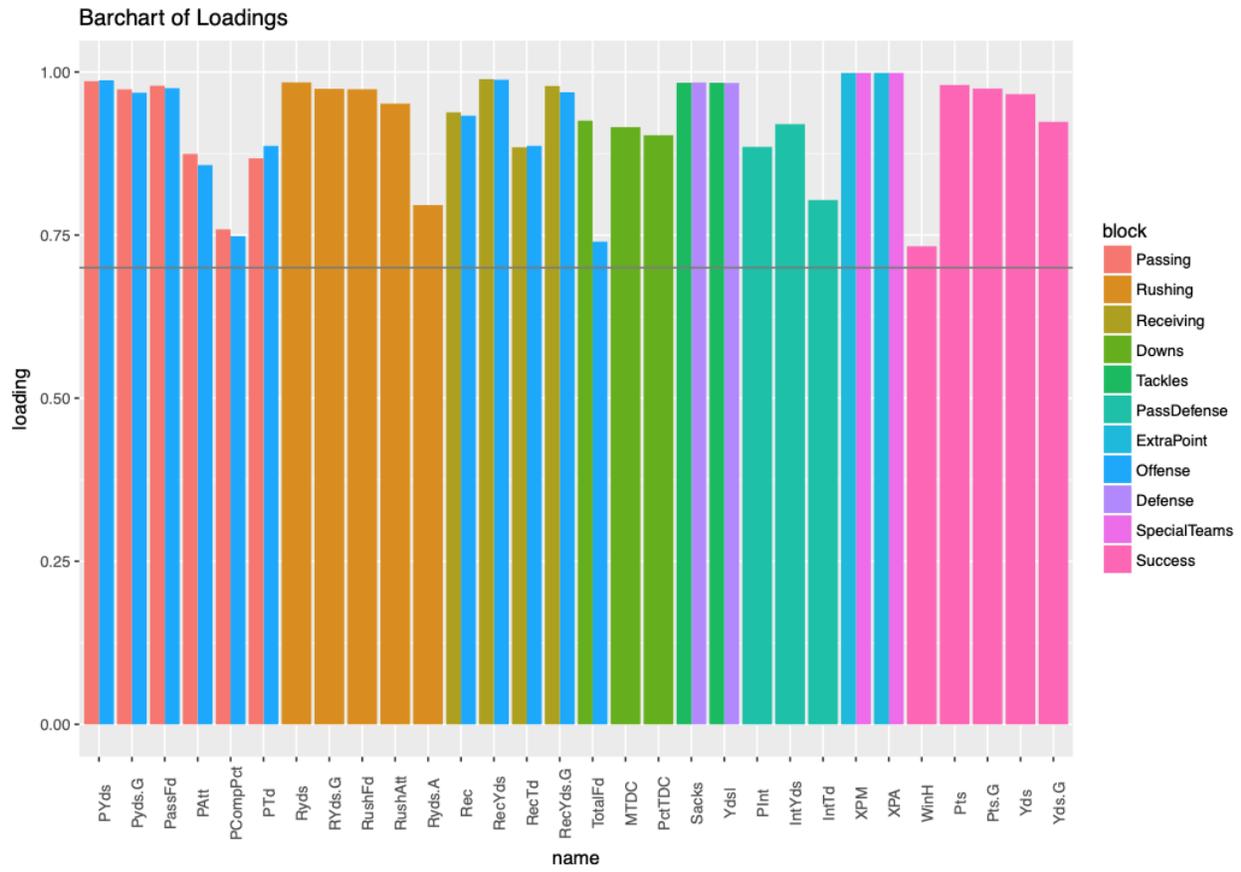


Figure 5: Loadings of Final Model

From Figure 5 we can see that all loadings are above the acceptable value of 0.7. Which is a vast improvement from the first model. This figure also show how many variables where removed in order to improve the model.

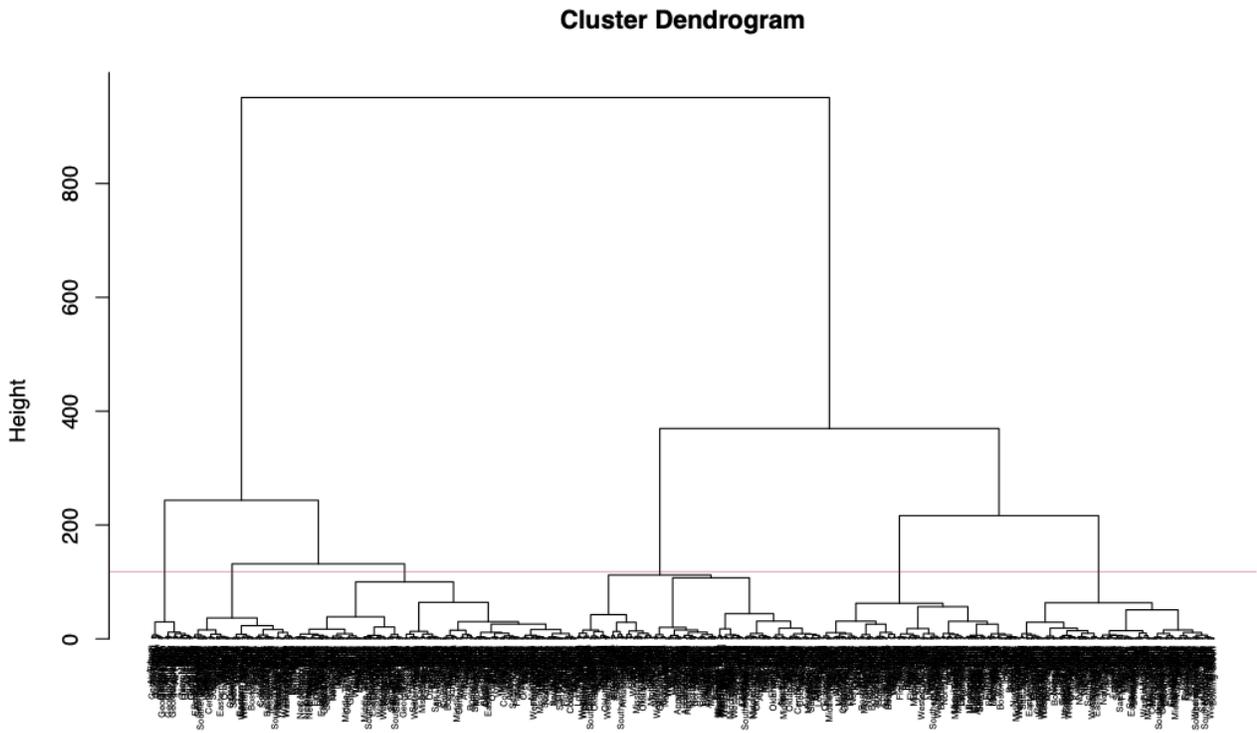


Figure 6: Success Dendrogram

The above cluster dendrogram contains 1291 teams, so that is why it appears cluttered. The dendrogram is grouping teams based on their success score. Then we separated the dendrogram into 6 different groups, using the red line. An example of 1 of the groups can be seen below.

[1]	Boston College	Miami	Pittsburgh	Syracuse	Temple
[6]	Tulane	Iowa	Penn State	Louisiana Tech	Marshall
[11]	UAB	Notre Dame	Northan Illinois	Western Michigan	Nevada
[16]	Califorina	Utah	Auburn	Florida	Kentucky
[21]	LSU	Mississippi State	Troy	Boston College	Clemson
[26]	Miami	Temple	TCU	Texas	Iowa
[31]	Michigan	Louisiana Tech	Akron	Central Michigan	Northan Illinois
[36]	Utah State	Wyoming	Arizona State	Stanford	Washington
[41]	LSU	Mississippi State	Texas A&M	Appalacian State	Troy
[46]	Boston College	Florida State	Miami	NC State	Pittsburgh
[51]	Wake Forest	SMU	Temple	UCF	Kansas State
[56]	TCU	Texas	Maryland	Michigan	Minnesota
[61]	Penn State	Wisconsin	North Texas	Old Dominion	BYU
[66]	Ohio	San Diego State	Colorado	Stanford	Utah
[71]	Florida	LSU	Mississippi State	Texas A&M	Appalacian State
[76]	Arkansas State	Troy	Florida State	Louisville	Pittsburgh
[81]	South Florida	Kansas State	Texas	Iowa	Michigan
[86]	Michigan State	Penn State	Florida Atlantic	Northan Illinois	New Mexico
[91]	San Diego State	Colorado	Utah	Washington	Florida
[96]	LSU	Texas A&M	Boston College	Clemson	Louisville
[101]	NC State	Virginia	Virginia Tech	UCF	Oklahoma State
[106]	Texas	Maryland	Minnesota	Louisiana Tech	North Texas
[111]	Rice	UAB	UTEP	Utah State	Arizona State
[116]	Stanford	Utah	Washington	Florida	Kentucky
[121]	Missouri	Ole Miss	Louisiana	Boston College	Maryland
[126]	North Carolina	Syracuse	Virginia Tech	Houston	Kansas State
[131]	Oklahoma	TCU	Texas	Michigan	Nebraska
[136]	Northwestern	Florida Atlantic	Middle Tennessee	Tulane	Buffalo
[141]	Ohio	Toledo	Arizona	Stanford	USC
[146]	Vanderbilt	Louisiana	South Alabama	Florida State	TCU
[151]	Ohio State	Wisconsin	SMU	Tulsa	BYU
[156]	Notre Dame	Bowling Green	Kent State	Toledo	Boise State
[161]	San Diego State	Stanford	Washington	Florida	LSU
[166]	Mississippi State	Ole Miss	South Carolina	Vanderbilt	Western Kentucky
[171]	Florida State	NC State	Virginia Tech	Texas	Illinois
[176]	Michigan State	Tulsa	Temple	Utah	LSU
[181]	South Carolina	Vanderbilt	Florida Intl	Louisiana	Clemson
[186]	Florida State	Maryland	Miami	North Carolina	Virginia Tech
[191]	Missouri	Nebraska	Iowa	UCF	Kent State
[196]	Miami (OH)	Toledo	Alabama	Florida	Georgia
[201]	Kentucky	Mississippi State	South Carolina	Florida Intl	Middle Tennessee
[206]	Clemson	North Carolina	Virginia Tech	Nebraska	Oklahoma State
[211]	Iowa	Ohio State	Penn State	Wisconsin	Southern Mississippi
[216]	UCF	Northan Illinois	Temple	Utah	Califorina

Figure 7: Cluster 1

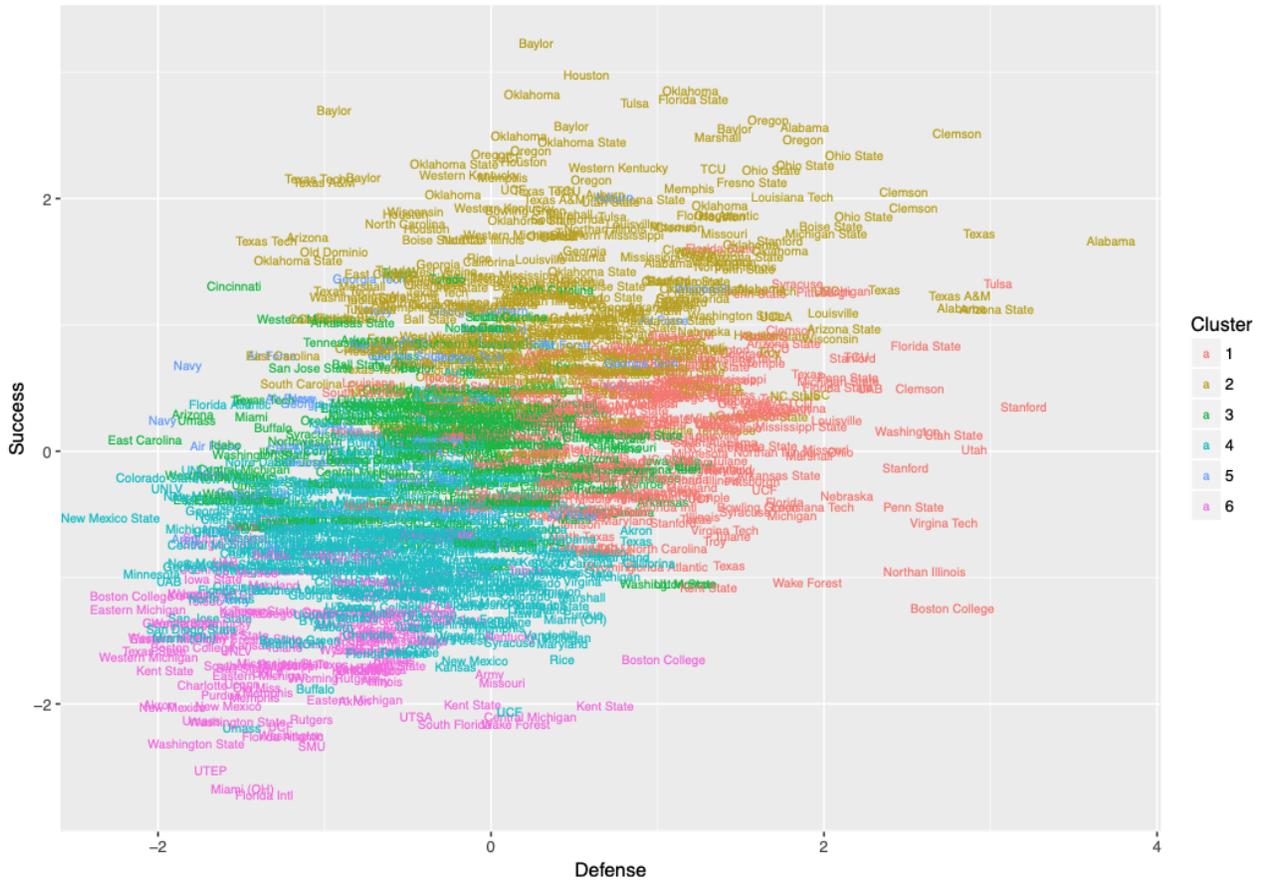


Figure 8: Defense Effect on Success

Figure 8 is plotting the individual teams score of Success and Defense. The graph shows that there is a positive linear relationship between Defense and Success. We can see that there are some outliers, teams with only high Defense but lower Success as well as teams with lower Defense but high Success. Note that for Figures 8-10 there are 11 seasons worth of teams, so there will be repeated team in each figure.

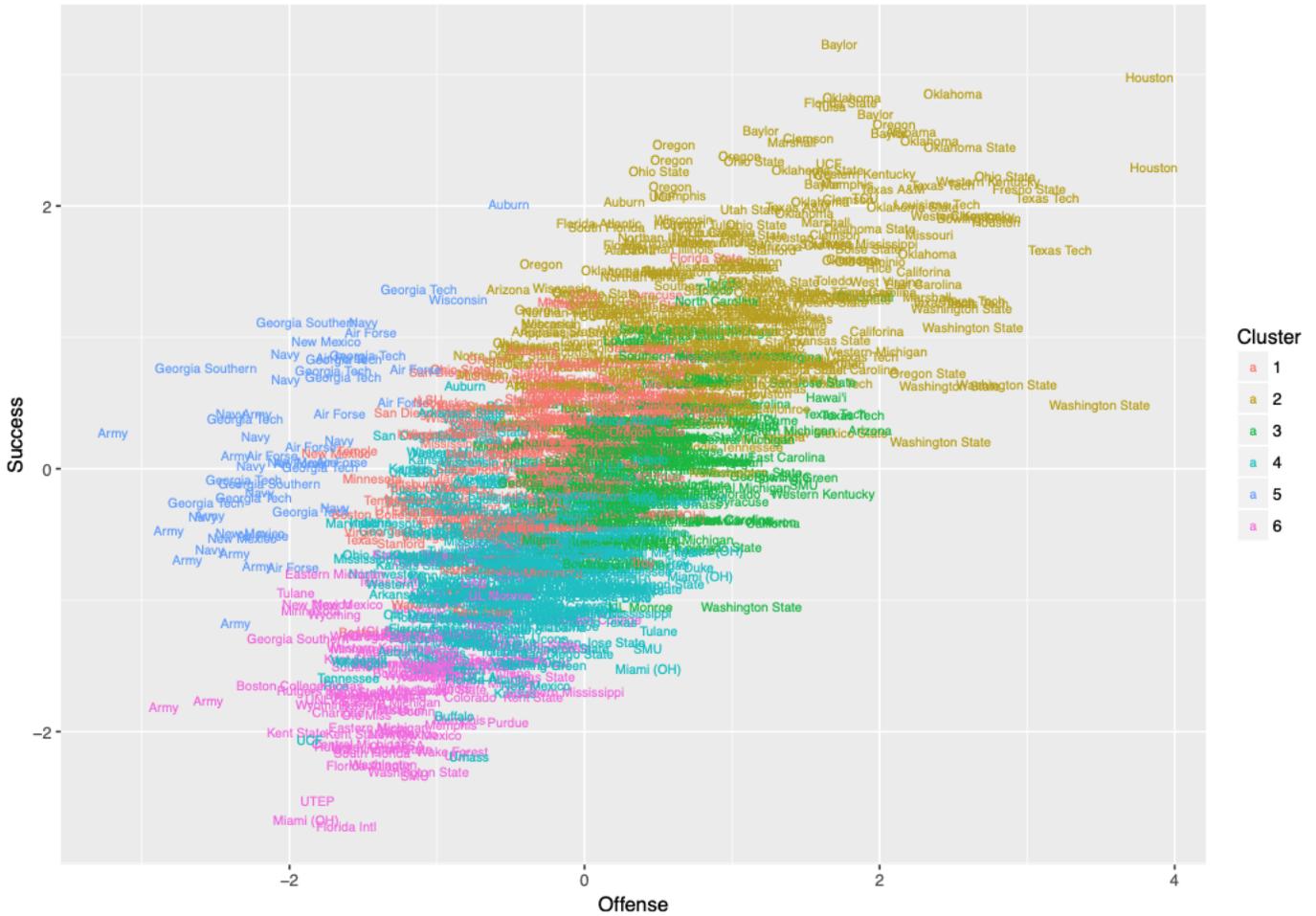


Figure 9: Offense Effect on Success

Figure 9 is plotting the individual teams score of Success and Offense. The graph shows that there is a strong positive linear relationship between Defense and Success. Which we can see is a stronger relationship than Offense on Success. However there are still some outliers.

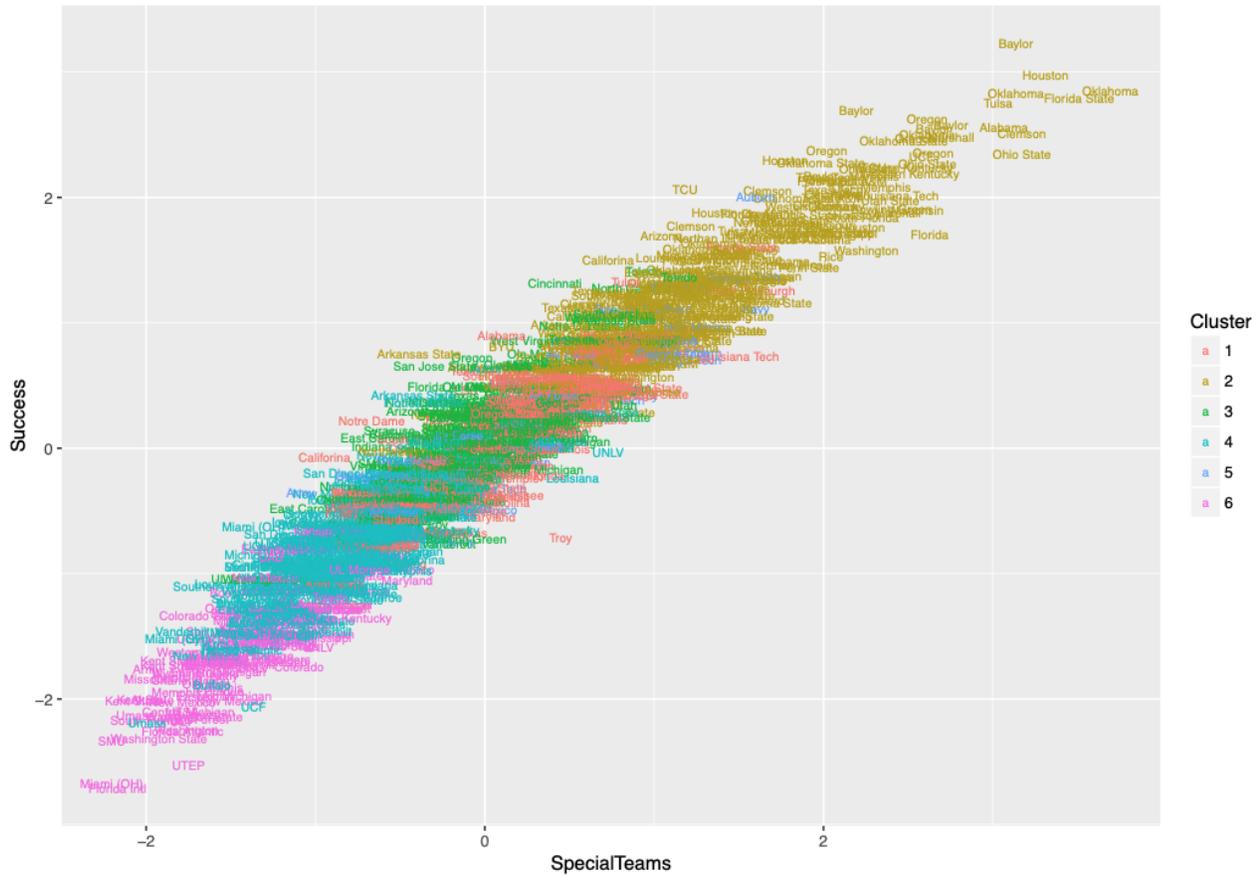


Figure 10: Special Teams Effect on Success

From Figure 10 we can see that Special Teams has a very strong positive linear relationship with Success, which is by far the strongest correlation among the three, while also having little to no outliers.

6 Conclusion

In conclusion we went through two different models for modeling the success of college football. Through the improvement of the model we remove multiple manifest variables and some first order latent variables. Throughout this improvement we were able to achieve a far better model for the success of a college football team. With the use of the improved model we were able to obtain the following relationships

$$\begin{aligned} \text{Offense} &= 0.60(\text{Passing})+0.05(\text{Rushing})+0.40(\text{Receiving})+0.02(\text{Downs}), \\ \text{Defense} &= 1.00(\text{Tackles}), \\ \text{Special Teams} &= 1.00(\text{ExtraPoint}), \text{ and} \\ \text{Success} &= 0.14(\text{Offense})+0.03(\text{Defense})+0.87(\text{SpecialTeams}). \end{aligned}$$

These equations are a large part of the research, as they show us the composition of each latent variable along with what elements are the most important. Though as stated in the analysis section even though Special Teams has the largest effect on Success, the other latent variables are still important to consider. Through these formulas we can predict the success of college football teams scores.

There are a few ways that this research can be continued. One would be to widen the model to include a few more variables. For example, the conference of each team could be added as a variable to see the effect that it has on success. The use of different methods of determining the indicators could provide interesting results. There are also many different configurations of the latent variables and other models that can be considered. Though there are some improvements to be made the process of this research was extremely thorough and we attempted to find the best possible model with the material that we used.

7 Glossary

Passing:

PTd - Passing touchdown
PCompPct - Passing completion percentage
PAtt - Passing Attempt
PassFd - Passing first down
Pyds.G - Passing Yards per game
PYds - Passing Yards
Int - Interception thrown
Yds.PA - Yards per pass attempt
LongPA - Longest pass

Rushing:

Ryds.A - Rushing Yards per attempt
RushAtt - Rushing Attempt
RushFd - Rushing First down
RYds.G - Rushing yards per game
Ryds - Rushing yards
LongR - Longest Rush
RushTd - Rushing Touchdown
Receiving:
RecYds.G - Receiving yards per game
RecTd - Receiving touchdown
RecYds - Reception yards
Rec - Receptions
AvgRec - Average yards per reception
LongRec - Longest reception
RecTd - Receiving touchdown

Downs:

PctTDC - Percentage third down conversion
MTDC - Made Third down conversion
TotalFd - Total first downs
PenFd - Penalty first down
AttTDC - Attempted third down conversion
MFDC - Made fourth down conversion
AttFDC - Attempted fourth down conversion
PctFDC - Percentage fourth down conversion

Penalties:

TPen - Total penalties
TPenYds - Total penalty yards

Tackles:

Ydsl - Yards lost on sack
Sacks - Total sacks
SoloT - Unassisted tackles
AstT - Assisted tackles
TotalT - Total tackles

Pass Defense:

IntTd - Interceptions returned for touchdowns
IntYds - Intercepted returned yards
PInt - Passes intercepted
PD - Passes defensed
LongInt - Longest interception return

Fumbles:

FRec - Fumbles recovered

Kickoff Returns:

KRAtt - Kickoff return attempts
KRYds - Kickoff return yards
AvgKR - Average yards per kickoff return

Punt Return:

PRAtt - Punt return attempts
PRyds - Punt return yards
AvgPR - Average yards per punt return
LongPR - Longest punt return
PRTd - Punt return touchdowns

Field Goals:

Fgm - Field goals made
Fga - Field goal attempts
FPct - Percentage of field goals made
LongF - Longest field goal made
1to19M - 1-19 yard field goals made
1to19A - 1-19 yard field goals attempted
1to19 - 1-19 yard field goals
20to29M - 20-29 yard field goals made
20to29A - 20-29 yard field goals attempted
20to29 - 20-29 yard field goals
30to39M - 30-39 yard field goals made
30to39A - 30-39 yard field goals attempted
30to39 - 30-39 yard field goals
40to49M - 40-49 yard field goals made
40to49A - 40-49 yard field goals attempted
40to49 - 40-49 yard field goals
50plusM - 50+ yard field goals made
50plusA - 50+ yard field goals attempted
50plus - 50+ yard field goals

Extra Point:

XPA - Extra point attempts
XPM - Extra points made

XPct - Percentage of extra point made

Punting:

Punts - Total punts

PuntYds - Gross punting yards

LongP - Longest punt

AvgP - Gross punting average

Success:

WinH Wins Home

WinA Wins away

Pts - Total points

Pts.G - points per game

Yds - Total Yards

Yds.G - Yards per game.

8 References

Caro, Cary A. College Football Success: The Relationship between Recruiting and Winning. *International Journal of Sports Science & Coaching*, vol. 7, no. 1, Mar. 2012, pp. 139-152, doi:10.1260/1747-9541.7.1.139.

<http://www.espn.com/college-football/conferences>, accessed December 2018

<http://www.espn.com/college-football/standings>, accessed December 2018

Mayhew, J. L., et al. "Contributions of speed, agility and body composition to anaerobic power measurement in college football players." *The Journal of Strength & Conditioning Research* 3.4 (1989): 101-106.

Sanchez, G. (2013) *PLS Path Modeling with R* Trowchez Editions. Berkeley, 2013. <http://www.gastonsanchez.com>