# An Analysis of and Recommendation for NCAA Division 1 College Football Coaching Salaries

By: Nate Hoffelmeyer
August 4, 2018

## I. Introduction

### I.II Background

NCAA Division I College Football is widely popular in the United States today. In 2017, the median audience of TV viewers per game was nearly 1.4 million people[1]; and the total physical attendance at stadiums across all games in that season was almost fifty million people[2]. Thanks to lucrative television contracts, sold out stadiums, apparel sales, and more - tapping into these millions of fans has become a boon for college athletic departments, with the top thirty-one reporting over one hundred million USD in yearly revenue[3].

Given how many millions of dollars are at stake, this means that athletic programs across DI Football must compete to put the best "product" on the field; and like many major corporations in the U.S., production quality is usually attributed to strong leadership. As the leader of a football program, the head coach is typically paid like one would imagine a corporate CEO to be, with salaries at top programs well into the millions of dollars (U.S.) per year. For this reason, both schools at large and academic departments in particular have a vested interest in making evidence-informed decisions about who to hire as head coach, and how much to pay them.

### I.II Data and Problem Specification

This paper will look at a few of the various factors that might impact a college football coaching salary, and attempt to help drive an evidenced based decision on how much to pay the head football coach of one program in particular: Syracuse University. The data (described in Table 1) for this study has been collected from a variety of sources, and merged into one final data set for analysis. The final analysis contained 116 records across 18 variables, from four various sources, outlined in the data sources section at the end of this document.

[1] https://www.statista.com/statistics/748033/college-football-tv-ratings/
[2] https://www.statista.com/statistics/254672/attendance-at-division-i-ii-und-iii-college-football-games/
[3] http://sports.usatoday.com/ncaa/finances/

*Table 1*

| Variable Name | Description | Comments |
|---|---|---|
| Name | Name of head football coach | |
| School | Name of universty | |
| Conference | Name of athletic conference | |
| TotalSalary | Coach's salary in $ | Does not include bonus pay |
| Wins2012 | Number of wins in 2012 season | |
| Losses2012 | Number of losses in 2012 season | |
| WinPct | Percent of games won in 2012 season | |
| CoachTenure | Total number of years experience as head coach | |
| CoachRecord | Total percentage of games one over career by coach | |
| ProgramTenure | Number of years program has fielded a team | |
| SchoolWinPct | Total percentage of games won over entire program tenure | |
| Rank2012 | Overall team rank for 2012 season | |
| FGR | Federal graduation rate | Student-athletes who received aid grad rates[4] |
| GSR | Graduation success rate | Student-athletes who enrolled and graduated[5] |
| StadiumSize | Size of stadium attendance capacity | |
| StudentPop | Total undergraduate students enrolled at university | |

Using the above mentioned data, the study sets out to provide answers to, and guide Syracuse University on, the following issues:

- What is the recommended Salary for the Syracuse (SU) football coach?
- How does the conference the SU coach is in change the salary recommendation?
- What schools were dropped from the data, and why?
- What effect does graduation rate have on salary?

[4] http://grfx.cstv.com/photos/schools/stjo/genrel/auto_pdf/definitions.pdf
[5] http://grfx.cstv.com/photos/schools/stjo/genrel/auto_pdf/definitions.pdf

- How accurate is the model?
- What variable has the single biggest impact on salary?

To provide answers to these questions, the study here focused on obtaining, scrubbing / exploring, modeling, and interpreting model results of the data outlined above. What follows is a discussion of this process, and the conclusions this study reached – followed by a final recommendation for the salary of the SU head football coach.

## II. Obtain

Coaching data was initially provided by Professor John Fox of Syracuse University for the purposes of this study. The initial data set contained the Name, School, Conference, and total compensation including salary, "other pay", and "max bonus" of 124 NCAA Division I college football head coaches.

Athletic performance records for all 124 coaches and schools was obtained through scrubbing a sports reference site[6]. Data obtained from the site included the number of wins & lossess the team sustained in the 2012 season, a derived win percentage from that same win/loss data, the toal number of years experience by the coach and how well (win percentage) that coach has performed over those years, the program tenure and win percentage, as well as the ranking of schools to end of the 2012 season.

Graduation rates were obtained from the NCAA[7] for all 124 schools, as noted in the table above.

Finally, stadium size and undergratued student population were obtained through the scrubbing of two online sources: a college data site[8], and Wikipedia[9].

## III. Explore and Scrub

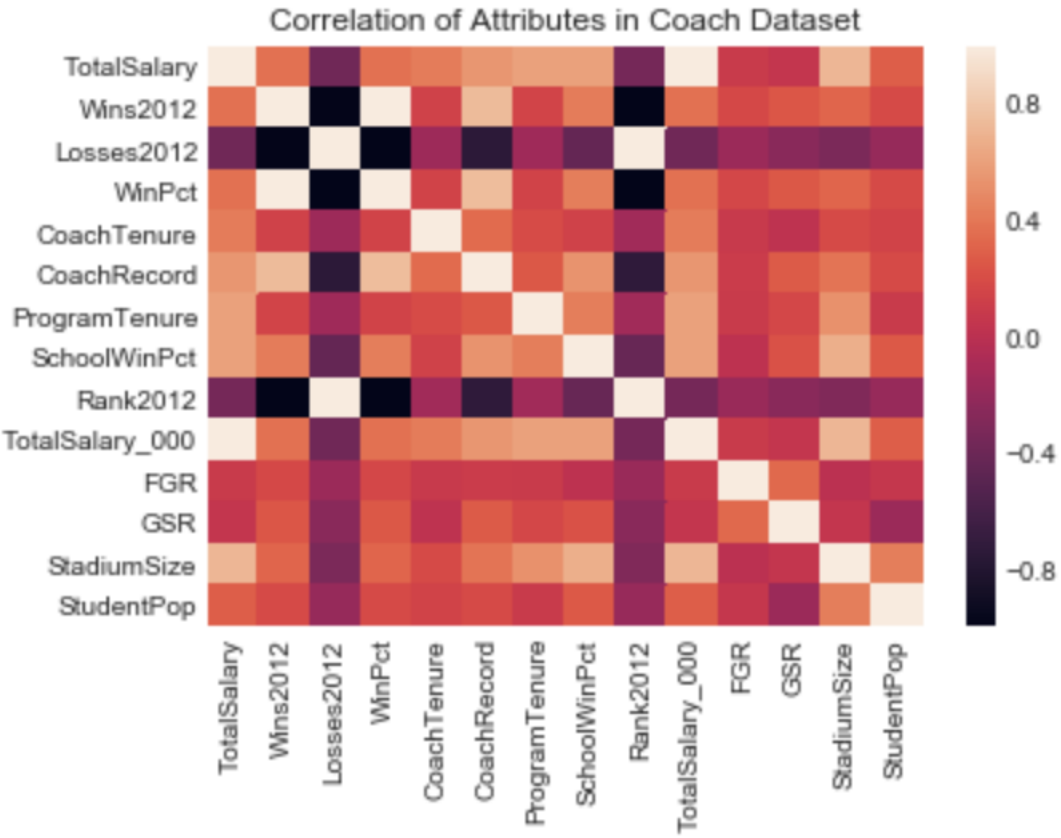
After obtaining the data described previously, exploration and scrubbing commenced using statistical analysis tools and libraries from the programming language Python. The first task was to melt the various data sources together, and explore the impact of the variables on coaches salary.

In looking at a correlation matrix (Figure 1) for all variables, it is easy to note that there are certain variables – such as CoachTenure, StadiumSize, ProgramTenure – that are more correlated to TotalSalary than others. This analysis helps to determine what factors might impact the model that will be used to attempt to make a salry recommendation, and can also

[6] https://www.sports-reference.com/cfb/schools/
[7] https://web3.ncaa.org/aprsearch/gsrsearch
[8] https://www.collegedata.com/cs/content/content_choosearticle_tmpl.jhtml?articleId=10006
[9] https://en.wikipedia.org/wiki/List_of_NCAA_Division_I_FBS_football_stadiums

be used to help athletic staff at universities learn what factors to pay attention to when scouting new coaching talent in the future.

*Figure 1*




Further exploration of the data showed strong trends across conferences for salary given a coaches overall performance record across their career (Figure 2), and given a program's tenure (Figure 3).

*Figure 2: Coach Salary related to overall Coach Record*

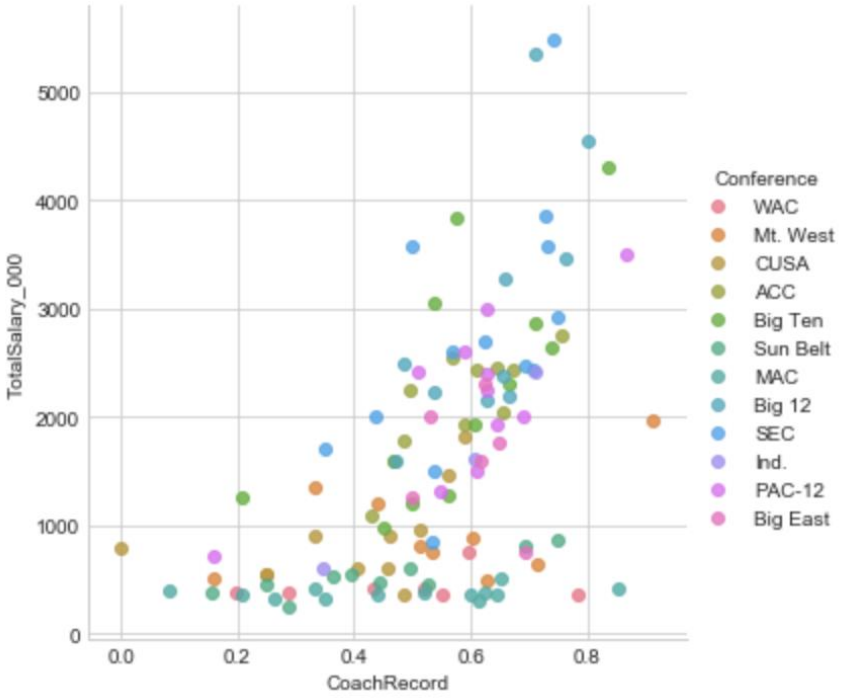
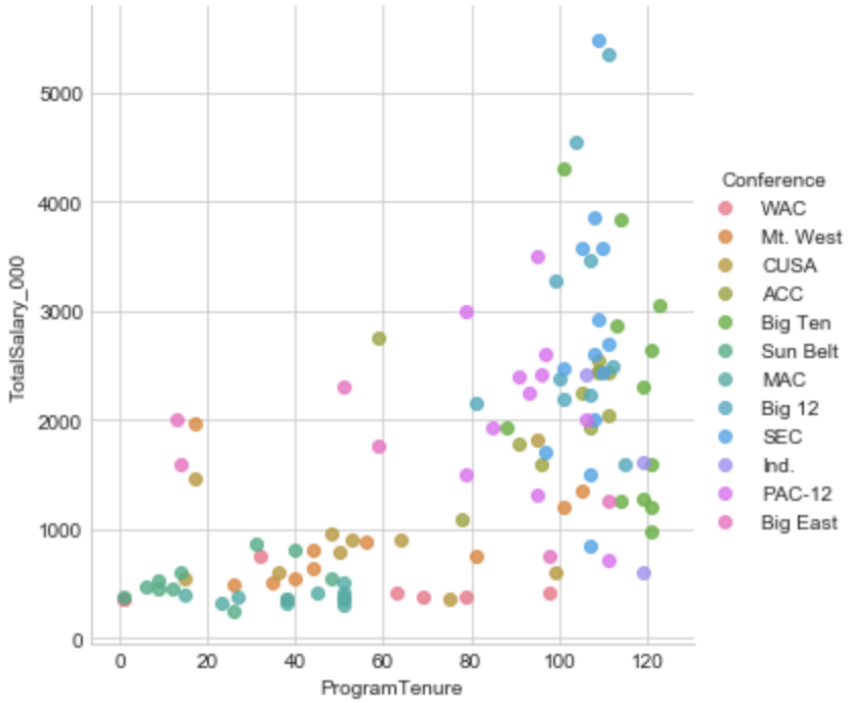
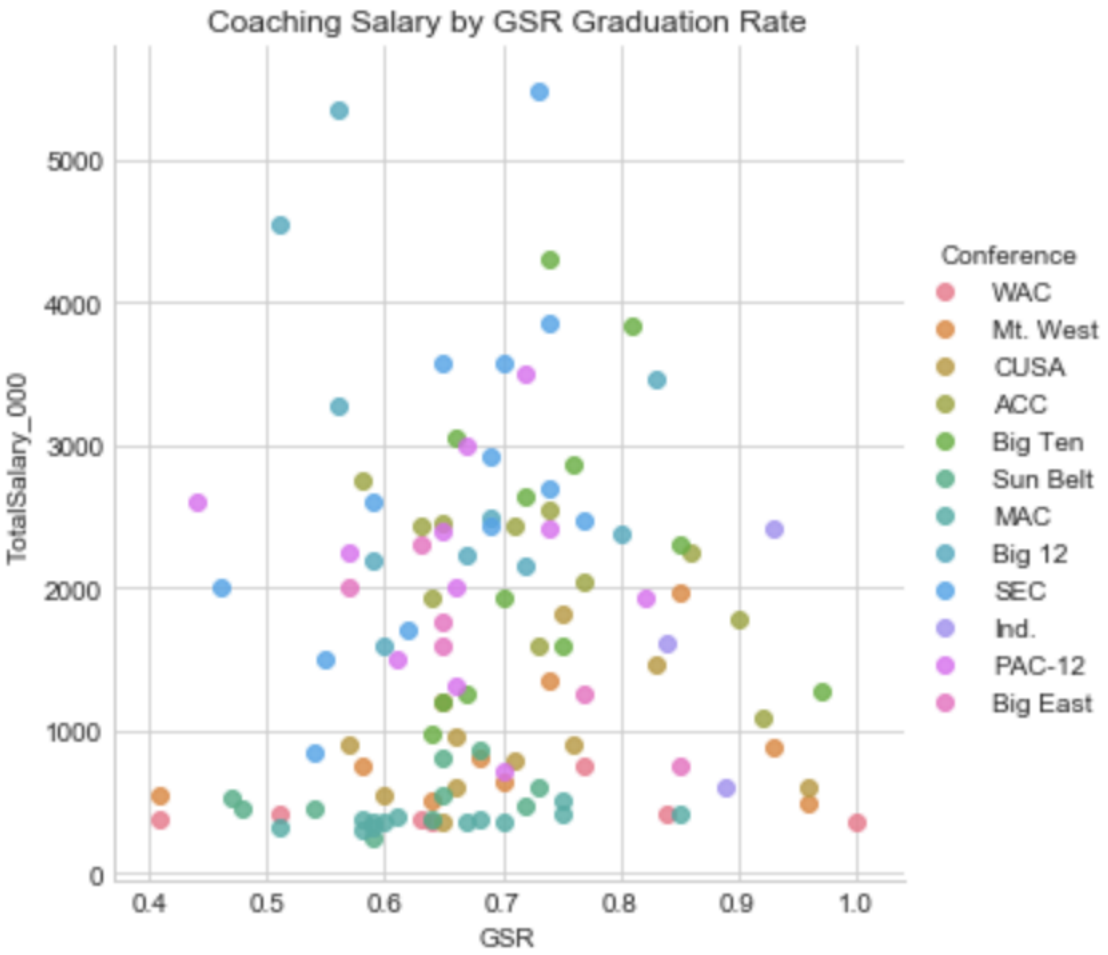



*Figure 3: Coach Salary related to Program Tenure*




During analysis, it was also particularly interesting to note the relationships between graduation rate (specifically GSR) related to a coach's salary, and also the impact of student population on coach's salary.

*Figure 4*




There isn't a strong relationship where an assumption could be made that graduation rate should be used as a predictor, or part of the reason for, a coaching salary. However, as one might expect – there is a strong relationship between stadium size and student population, and a coache's salary. This may lead to overfitting and multicolinearity in a technical model, so it is important to note that schools with bigger populations also have larger stadiums, which in turn generally make more money from both athletic events and related sales, as well as tuition and other student related costs. That, in turn, would naturally lead to the assumptino that larger schools with bigger stadiums are able to pay coaches more. What it doesn't account for is that, also, school's with larger stadiums seem to have higher percentages of games won.

*Figure 5*

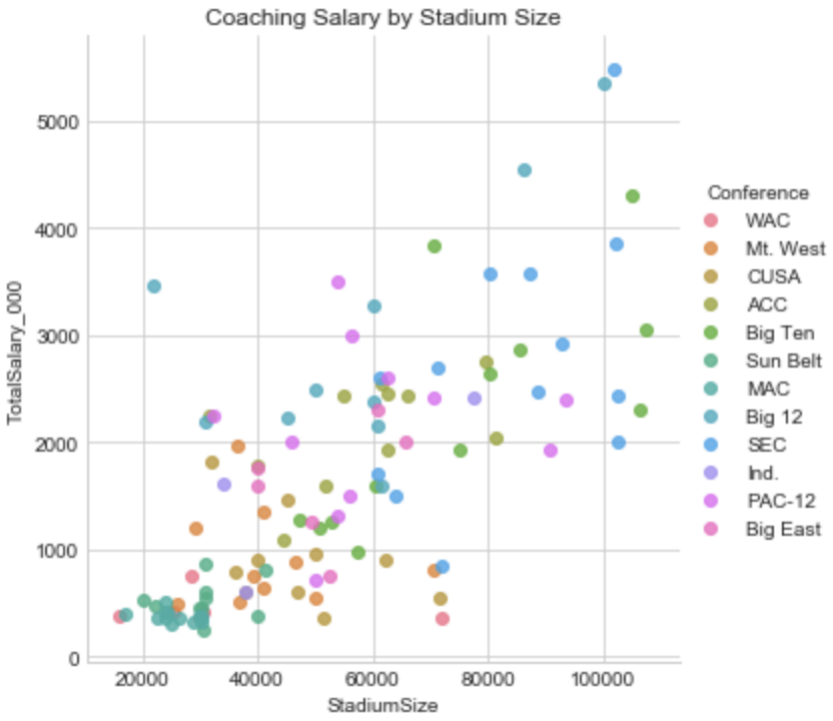
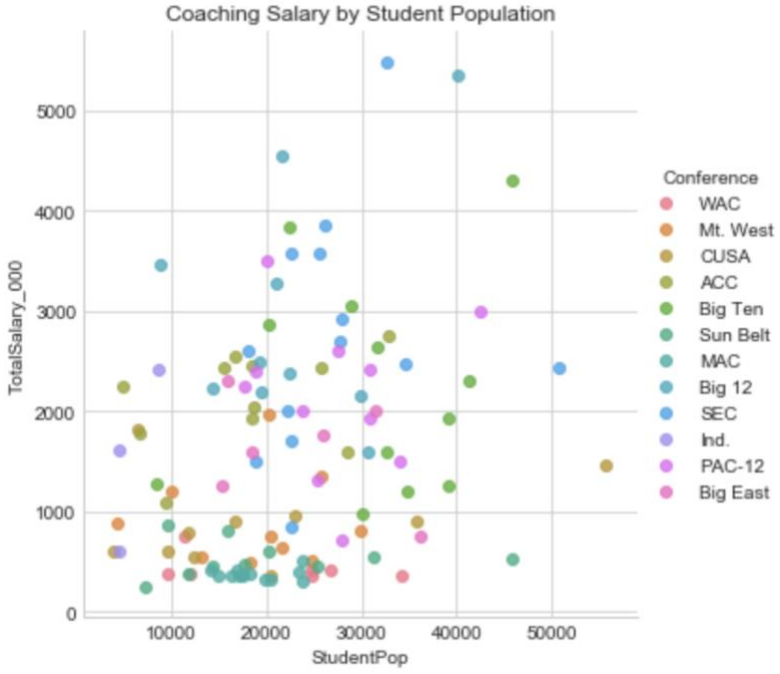



*Figure 6*



**IV. Model & Interpret**

Once the data has been sufficiently explored, modeling can commence. After running various models accounting for a range of variability in the prediction of TotalSalary of between seventy-three and seventy-six percent, a simple Ordinary Least Squares (OLS) regression model using Conference, Wins2012, Losses2012, WinPct, CoachTenure, Rank2012, and StadiumSize to predict TotalSalary.

Using this model, and with Syracuse conference set to today's known membership in the ACC, the model is able to predict Salary, with 73.2% variability accounted for and with all predictor variables significant at a 95% confidence level – except for some specific conferences which seem to have no significance, for which later models will attempt to account for.

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | TotalSalary | R-squared: | 0.771 |
| Model: | OLS | Adj. R-squared: | 0.732 |
| Method: | Least Squares | F-statistic: | 19.45 |
| Date: | Sat, 04 Aug 2018 | Prob (F-statistic): | 3.74e-24 |
| Time: | 12:51:30 | Log-Likelihood: | -1698.9 |
| No. Observations: | 116 | AIC: | 3434. |
| Df Residuals: | 98 | BIC: | 3483. |
| Df Model: | 17 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P>\|t\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 7.337e+06 | 4.05e+06 | 1.812 | 0.073 | -6.98e+05 | 1.54e+07 |
| Conference[T.Big 12] | 9.265e+05 | 2.65e+05 | 3.499 | 0.001 | 4.01e+05 | 1.45e+06 |
| Conference[T.Big East] | -2.391e+05 | 3.29e+05 | -0.728 | 0.469 | -8.91e+05 | 4.13e+05 |
| Conference[T.Big Ten] | 5.233e+04 | 2.59e+05 | 0.202 | 0.840 | -4.61e+05 | 5.66e+05 |
| Conference[T.CUSA] | -8.052e+05 | 2.64e+05 | -3.045 | 0.003 | -1.33e+06 | -2.8e+05 |
| Conference[T.Ind.] | -3.849e+05 | 3.95e+05 | -0.973 | 0.333 | -1.17e+06 | 4e+05 |
| Conference[T.MAC] | -9.082e+05 | 2.83e+05 | -3.214 | 0.002 | -1.47e+06 | -3.47e+05 |
| Conference[T.Mt. West] | -5.8e+05 | 2.7e+05 | -2.145 | 0.034 | -1.12e+06 | -4.34e+04 |
| Conference[T.PAC-12] | 1.743e+05 | 2.55e+05 | 0.683 | 0.496 | -3.32e+05 | 6.8e+05 |
| Conference[T.SEC] | 8.726e+04 | 2.66e+05 | 0.328 | 0.743 | -4.4e+05 | 6.14e+05 |
| Conference[T.Sun Belt] | -8.103e+05 | 2.84e+05 | -2.856 | 0.005 | -1.37e+06 | -2.47e+05 |
| Conference[T.WAC] | -1.054e+06 | 3.16e+05 | -3.339 | 0.001 | -1.68e+06 | -4.27e+05 |
| Wins2012 | 6.255e+05 | 2.67e+05 | 2.346 | 0.021 | 9.65e+04 | 1.15e+06 |
| Losses2012 | -1.01e+06 | 3.52e+05 | -2.872 | 0.005 | -1.71e+06 | -3.12e+05 |
| WinPct | -1.359e+07 | 6.2e+06 | -2.190 | 0.031 | -2.59e+07 | -1.28e+06 |
| CoachTenure | 3.449e+04 | 1.05e+04 | 3.284 | 0.001 | 1.36e+04 | 5.53e+04 |
| Rank2012 | 3.965e+04 | 1.91e+04 | 2.077 | 0.040 | 1772.166 | 7.75e+04 |
| StadiumSize | 16.4908 | 3.928 | 4.198 | 0.000 | 8.696 | 24.286 |

| | | | |
|---|---|---|---|
| Omnibus: | 0.578 | Durbin-Watson: | 1.864 |
| Prob(Omnibus): | 0.749 | Jarque-Bera (JB): | 0.396 |
| Skew: | 0.142 | Prob(JB): | 0.821 |
| Kurtosis: | 3.034 | Cond. No. | 7.41e+06 |

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.41e+06. This might indicate that there are
strong multicollinearity or other numerical problems.

Second, a model was constructed to predict what the SU coaching salary should be were Syracuse still a member of the Big East Conference. Similarly, the model shows that the

variables selected as predictors account for 73.3% of the change in Total Salary, and all variables except some of the specific conferences are significant at a 95% level of confidence.

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | TotalSalary | R-squared: | 0.772 |
| Model: | OLS | Adj. R-squared: | 0.733 |
| Method: | Least Squares | F-statistic: | 19.53 |
| Date: | Sat, 04 Aug 2018 | Prob (F-statistic): | 3.22e-24 |
| Time: | 14:11:31 | Log-Likelihood: | -1698.7 |
| No. Observations: | 116 | AIC: | 3433. |
| Df Residuals: | 98 | BIC: | 3483. |
| Df Model: | 17 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P>\|t\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 7.384e+06 | 4.03e+06 | 1.831 | 0.070 | -6.18e+05 | 1.54e+07 |
| Conference[T.Big 12] | 8.939e+05 | 2.7e+05 | 3.305 | 0.001 | 3.57e+05 | 1.43e+06 |
| Conference[T.Big East] | -2.9e+05 | 3.15e+05 | -0.920 | 0.360 | -9.16e+05 | 3.36e+05 |
| Conference[T.Big Ten] | 2.173e+04 | 2.64e+05 | 0.082 | 0.935 | -5.02e+05 | 5.45e+05 |
| Conference[T.CUSA] | -8.386e+05 | 2.7e+05 | -3.102 | 0.003 | -1.38e+06 | -3.02e+05 |
| Conference[T.Ind.] | -4.188e+05 | 3.99e+05 | -1.049 | 0.297 | -1.21e+06 | 3.73e+05 |
| Conference[T.MAC] | -9.482e+05 | 2.91e+05 | -3.262 | 0.002 | -1.52e+06 | -3.71e+05 |
| Conference[T.Mt. West] | -6.16e+05 | 2.77e+05 | -2.221 | 0.029 | -1.17e+06 | -6.57e+04 |
| Conference[T.PAC-12] | 1.418e+05 | 2.61e+05 | 0.544 | 0.588 | -3.76e+05 | 6.59e+05 |
| Conference[T.SEC] | 5.999e+04 | 2.69e+05 | 0.223 | 0.824 | -4.74e+05 | 5.94e+05 |
| Conference[T.Sun Belt] | -8.483e+05 | 2.91e+05 | -2.915 | 0.004 | -1.43e+06 | -2.71e+05 |
| Conference[T.WAC] | -1.093e+06 | 3.22e+05 | -3.391 | 0.001 | -1.73e+06 | -4.53e+05 |
| Wins2012 | 6.163e+05 | 2.67e+05 | 2.312 | 0.023 | 8.74e+04 | 1.15e+06 |
| Losses2012 | -1e+06 | 3.5e+05 | -2.855 | 0.005 | -1.7e+06 | -3.05e+05 |
| WinPct | -1.348e+07 | 6.19e+06 | -2.179 | 0.032 | -2.58e+07 | -1.21e+06 |
| CoachTenure | 3.392e+04 | 1.05e+04 | 3.221 | 0.002 | 1.3e+04 | 5.48e+04 |
| Rank2012 | 3.871e+04 | 1.91e+04 | 2.024 | 0.046 | 759.460 | 7.67e+04 |
| StadiumSize | 16.3510 | 3.930 | 4.161 | 0.000 | 8.553 | 24.149 |

| | | | |
|---|---|---|---|
| Omnibus: | 0.635 | Durbin-Watson: | 1.871 |
| Prob(Omnibus): | 0.728 | Jarque-Bera (JB): | 0.420 |
| Skew: | 0.145 | Prob(JB): | 0.811 |
| Kurtosis: | 3.057 | Cond. No. | 7.39e+06 |

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.39e+06. This might indicate that there are strong multicollinearity or other numerical problems.

Third, a model was constructed to determing the impact of the SU coaching salary on membership in the Big Ten conference – a conference that is similar to the old Big East in school sizes and often geographies, and competes with the ACC in many ways for audiences, athletes, and students alike. This model, similar to the previous two, also indicates that 73.2% of the variability in coaching salary is accounted for by the predictor variables, with all variables except a select few conferences being significant at a 95% confidence level.

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | TotalSalary | R-squared: | 0.771 |
| Model: | OLS | Adj. R-squared: | 0.732 |
| Method: | Least Squares | F-statistic: | 19.44 |
| Date: | Sat, 04 Aug 2018 | Prob (F-statistic): | 3.81e-24 |
| Time: | 12:51:31 | Log-Likelihood: | -1698.9 |
| No. Observations: | 116 | AIC: | 3434. |
| Df Residuals: | 98 | BIC: | 3483. |
| Df Model: | 17 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P>\|t\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | 7.292e+06 | 4.05e+06 | 1.801 | 0.075 | -7.42e+05 | 1.53e+07 |
| Conference[T.Big 12] | 8.972e+05 | 2.71e+05 | 3.312 | 0.001 | 3.6e+05 | 1.43e+06 |
| Conference[T.Big East] | -2.684e+05 | 3.34e+05 | -0.803 | 0.424 | -9.31e+05 | 3.95e+05 |
| Conference[T.Big Ten] | -1.247e+04 | 2.58e+05 | -0.048 | 0.962 | -5.25e+05 | 5e+05 |
| Conference[T.CUSA] | -8.337e+05 | 2.71e+05 | -3.080 | 0.003 | -1.37e+06 | -2.97e+05 |
| Conference[T.Ind.] | -4.137e+05 | 4e+05 | -1.035 | 0.303 | -1.21e+06 | 3.8e+05 |
| Conference[T.MAC] | -9.32e+05 | 2.9e+05 | -3.215 | 0.002 | -1.51e+06 | -3.57e+05 |
| Conference[T.Mt. West] | -6.084e+05 | 2.78e+05 | -2.192 | 0.031 | -1.16e+06 | -5.76e+04 |
| Conference[T.PAC-12] | 1.425e+05 | 2.61e+05 | 0.545 | 0.587 | -3.76e+05 | 6.61e+05 |
| Conference[T.SEC] | 5.175e+04 | 2.69e+05 | 0.192 | 0.848 | -4.83e+05 | 5.86e+05 |
| Conference[T.Sun Belt] | -8.36e+05 | 2.91e+05 | -2.875 | 0.005 | -1.41e+06 | -2.59e+05 |
| Conference[T.WAC] | -1.079e+06 | 3.22e+05 | -3.351 | 0.001 | -1.72e+06 | -4.4e+05 |
| Wins2012 | 6.212e+05 | 2.67e+05 | 2.326 | 0.022 | 9.12e+04 | 1.15e+06 |
| Losses2012 | -1.005e+06 | 3.52e+05 | -2.853 | 0.005 | -1.7e+06 | -3.06e+05 |
| WinPct | -1.347e+07 | 6.21e+06 | -2.169 | 0.033 | -2.58e+07 | -1.14e+06 |
| CoachTenure | 3.412e+04 | 1.05e+04 | 3.234 | 0.002 | 1.32e+04 | 5.5e+04 |
| Rank2012 | 3.969e+04 | 1.91e+04 | 2.079 | 0.040 | 1797.314 | 7.76e+04 |
| StadiumSize | 16.7330 | 3.885 | 4.307 | 0.000 | 9.023 | 24.443 |

| | | | |
|---|---|---|---|
| Omnibus: | 0.714 | Durbin-Watson: | 1.876 |
| Prob(Omnibus): | 0.700 | Jarque-Bera (JB): | 0.490 |
| Skew: | 0.157 | Prob(JB): | 0.783 |
| Kurtosis: | 3.058 | Cond. No. | 7.41e+06 |

Warnings:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.
[2] The condition number is large, 7.41e+06. This might indicate that there are
strong multicollinearity or other numerical problems.

Lastly, a model was run with Conference removed from the equation to determine if it is possible to normalize across college football, as well as attempt to control for those specific conferences that are not significant. The predictor variables that ended up being significant at a 95% confidence level differed slight from above, and are shown in the figure below. However, the predictor variables decresed the amount of variablity our model accounts for in Total Salary to 60.5%.

OLS Regression Results

| | | | |
|---|---|---|---|
| Dep. Variable: | TotalSalary | R-squared: | 0.615 |
| Model: | OLS | Adj. R-squared: | 0.605 |
| Method: | Least Squares | F-statistic: | 59.60 |
| Date: | Sat, 04 Aug 2018 | Prob (F-statistic): | 4.19e-23 |
| Time: | 14:28:53 | Log-Likelihood: | -1729.2 |
| No. Observations: | 116 | AIC: | 3466. |
| Df Residuals: | 112 | BIC: | 3477. |
| Df Model: | 3 | | |
| Covariance Type: | nonrobust | | |

| | coef | std err | t | P>\|t\| | [0.025 | 0.975] |
|---|---|---|---|---|---|---|
| Intercept | -7.539e+05 | 1.99e+05 | -3.785 | 0.000 | -1.15e+06 | -3.59e+05 |
| Wins2012 | 5.045e+04 | 2.24e+04 | 2.254 | 0.026 | 6094.734 | 9.48e+04 |
| CoachTenure | 5.772e+04 | 1.18e+04 | 4.875 | 0.000 | 3.43e+04 | 8.12e+04 |
| StadiumSize | 30.2627 | 3.088 | 9.800 | 0.000 | 24.144 | 36.381 |

| | | | |
|---|---|---|---|
| Omnibus: | 7.079 | Durbin-Watson: | 1.967 |
| Prob(Omnibus): | 0.029 | Jarque-Bera (JB): | 8.159 |
| Skew: | 0.370 | Prob(JB): | 0.0169 |
| Kurtosis: | 4.068 | Cond. No. | 1.69e+05 |

As such, predictive results are obtained that allow for a recommendation to be made for the coaching salary of the SU football coach.

Initially, the SU coach's salary was reported to be $1,259,276. In using the models outlined above, if Syracuse were to remain in the ACC – a predicted salary of $1,588,514 is obtained and is also recommended.

Should Syracuse still be in the Big East conference, the model assuming conference recommends a salary of $1,338,192 – and should Syrcause consider moving into the Big Ten conference – the model assuming conferences recommends a salary of $1,606,387.

Rounding out the analysis, if conference is removed and a salary recommendation is to be made – the model indicates that an appropriate figure would be $1,370,987.

**V. Final Recommendation**

After careful study, it is the recommendation of this study's analysis that the head coach of SU Football be given a salary in the range of $1.3 million annually, and $1.6 million annually. Specifically, for upcoming negotiations, assuming that SU remains a member of the ACC conference – this study recommends an annual salary of approximately $1.59 million annually.

Graduation rate, as was seen from the exploratory plots, and as is indicated by the lack of its use as a predictor variable of Salary – does not seem to be significant in determining what a

coach is paid. That is consisent with other studies[10] that have been conducted on the matter, and it is therefore the recommendation of this study that it not be used to impact coaching salary.

Finally, it is important to note that the models account for roughly only three quarters of the variability (or change) in coaching salary. Future studies might consider the collection of additional data to increase the amount of variation accounted for. However, the models can all be confident that the variables are significant at a 95% confidence level – with the exception being some conferences likely not having a significant impact of coaching salary. This too would make sense, as larger conferences like the ACC, SEC, and Pac12 are not likely to be impacted by coaching salaries in very small conferences like the MAC and WAC.

It is also important to note, in closing, that the variable that appears to have the single biggest impact in controlling for salary variations is conference, but – given conferences as a predictor of salary the variable that impacts salary the most is the Wins2012 predictor variable[11]. Success, and in particular what can be assumed as recent success, fuels a "what have you done for me lately" dynamic that drives the immediate salary of NCAA Division I men's football head coaches.

[10] https://www.insidehighered.com/news/2016/03/29/teams-academic-success-not-likely-advance-coachs-career-study-finds

[11] Note: to determine the variable with the largest impact, the study looks at the coefficients of the variables as they are outlined in the models.

## Appendix

*Table 2: Schools Removed from Analysis*

| School | Reason Removed |
|---|---|
| Temple | Private School – no data |
| Tulsa | Private School – no data |
| Pittsburgh | Release of records pending court decision |
| Vanderbilt | Private School – no data |
| Miami | Private School – no data |
| Tulane | Private School – no data |
| Brigham Young | Private School – no data |
| Stanford | Private School – no data |