

Final Data Analysis Assignment

Course: MGSC 2301 Business Statistics

Section: 04 Summer 2026

Professor: Professor Spencer

*Examining the Relationship between Total Annual Athletic Department Spending and SEC
League Standing Using Simple Linear Regression*

Prepared By: Rachel Lee, Keshav Banka, Lokesh Addagiri, Soma Yasuda

Date: June 16, 2026



Executive Summary

This study examined the relationship between Total Athletic Department Spending and Final SEC Standing among Southeastern Conference (SEC) football programs from 2020 through 2024. As athletic spending continues to increase across collegiate athletics, it is important to understand whether greater financial investment is associated with improved competitive outcomes. The purpose of this analysis was to determine whether a statistically significant linear relationship exists between Total Athletic Department Spending and Final SEC Standing.

The study analyzed all 70 SEC football program-seasons from 2020 through 2024. Total Athletic Department Spending, measured in millions of dollars, served as the independent variable, while Final SEC Standing served as the dependent variable. Spending data were obtained from the U.S. Department of Education's Equity in Athletics Disclosure Act (EADA) database, and conference standings were obtained from official SEC records. Simple linear regression was used to evaluate the relationship between the two variables.

We hypothesize that as total athletic department spending increases, the final league standing will decrease, indicating an inverse relationship. A formal statement of our null hypothesis is H_0 : The slope of the regression line for total athletic department spending and final league standing position equals zero. This suggests that the two variables are unrelated by way of a linear relationship. Our formal statement of the alternative hypothesis is H_1 : The slope of the regression line is not zero. This suggests a statistically significant inverse linear relationship between total athletic department spending and final league standing.

The estimated regression equation was: $\hat{y} = 12.1647 - 0.0289x$

The slope coefficient indicates that for every additional \$1 million spent by an athletic department, the predicted Final SEC Standing improves by approximately 0.0289 ranking positions. Because lower ranking values represent stronger conference performance, the negative slope suggests that higher spending is associated with better SEC standings.

The analysis found a statistically significant relationship between Total Athletic Department Spending and Final SEC Standing. The t-test for the regression coefficient produced a t-statistic of -2.5517 with a p-value of 0.0130. Because the p-value was less than the significance level of $\alpha = 0.05$, the null hypothesis was rejected. The F-test for the overall regression model produced an F-statistic of 6.5112 with a p-value of 0.0130, leading to the same conclusion that the regression model is statistically significant.

Although the relationship was statistically significant, the practical strength of the model was relatively weak. The coefficient of determination (R^2) was 0.0874, indicating that Total Athletic Department Spending explained only 8.74% of the variation in Final SEC Standing. The remaining 91.26% of the variation is attributable to factors not included in the model. Additionally, the standard error of the estimate was 3.9072 ranking positions, indicating considerable variability around the regression line and limited predictive accuracy.

Several limitations should be considered when interpreting these findings. First, Total Athletic Department Spending reflects expenditures across all athletic programs rather than football spending. Second, Final SEC Standing is an ordinal-level variable, while linear

regression assumes a continuous dependent variable. Finally, the analysis was limited to SEC football programs from 2020 through 2024 and therefore may not be generalizable to other conferences or sports.

Overall, the results suggest that higher athletic department spending is associated with improved SEC football performance; however, spending alone is not a strong predictor of success. Future research should examine football-specific expenditures and incorporate additional variables such as recruiting rankings, coaching quality, coaching tenure, facilities investment, and strength of schedule. A multiple regression model including these factors would likely provide a more comprehensive explanation of competitive success in SEC football.

Results and Discussion

All SEC football program seasons from 2020 through 2024 will make up our sampling frame. The unit of analysis being measured is a program-season, with each school's performance in a given year treated as one observation. Programs that were transitioning into the SEC or facing postseason bans during the study period will be excluded. Our sampling frame consists of 70 program seasons, electing to include all SEC program seasons across 2020-2024 rather than a subset. This approach maximizes statistical power and avoids information loss associated with subsampling in an already limited frame of SEC-seasons.

Simple linear regression will be used to predict final SEC league standings based on total athletic department spending.

Descriptive Statistics

Descriptive statistics were calculated for both variables in this study. The sample consists of SEC football program seasons from 2020 through 2024, so each observation represents one program in one season rather than one program overall. Final SEC standing was regressed on Total Athletic Department Spending. Summary statistics for both variables are presented in Table 1 (Appendix A).

Athletic Department Spending varied considerably with respect to the financial resources programs allocated to athletics. Spending ranged from a minimum of \$82.76 million to a maximum of \$285.36 million. The mean spending was \$161.20 million and the median was \$159.97 million; these two values sit close together near the center of the distribution. The standard deviation was \$41.48 million, which indicates that spending was fairly spread out around the mean. The closeness of the mean and median, combined with the distance between

the minimum and maximum values, indicates that most programs cluster in the middle of the financial range while a small number spend well above the rest.

Final SEC Standing ranged from a minimum of 1 to a maximum of 14, covering the full competitive range of the conference with seasons from 2020 through 2024. The mean and median were both 7.5, and the standard deviation was 4.06. Because Final SEC Standing is ordinal level data, the median is the more appropriate measure of central tendency; a median of 7.5 falls exactly at the midpoint of the fourteen possible ranks. The mean for SEC Standing was calculated across all SEC programs in the sample and reflects the average across that group. The standard deviation of 4.06 indicates that standings were fairly evenly distributed across the ranking categories.

Distribution Analysis: Histograms and Boxplots

The histogram of Athletic Department Spending (Chart 1, Appendix A) shows that most observations fall between roughly \$120 million and \$190 million, so the bulk of the data is concentrated in that middle range. A small number of programs spent substantially more than the rest, which stretches the upper tail of the distribution to the right. The boxplot (Chart 3, Appendix A) is consistent with this. The interquartile range, which contains the middle 50% of the data, sits in a fairly narrow band around the median, and the mean is marked slightly above the median line inside the box, which is what you would expect from a distribution that is positively skewed by the high spenders. The bottom whisker reaches down to the minimum value of \$82.76 million; this value represents the low end of the data and is not an outlier, since Excel and SPSS identify outliers as separate dots or asterisks rather than as the end of a whisker. There is one observation plotted as a dot above the upper whisker, and this is the actual outlier in

this variable. It represents University of Tennessee Knoxville in 2024, with the highest athletic expenditure in the sample at \$285.36 million. The data was retained because it is a legitimate SEC program that accurately reflects the variation in spending across the conference. The Final SEC Standings for UT Knoxville in 2024 ended at a rank of number 2.

The histogram of Final SEC Standing (Chart 2, Appendix A) shows observations spread fairly evenly across the ranking categories. There are fourteen possible ranks, and as the data dealt with contains a consistent frequency of occurrence for all rankings and ordinal level data, this is not unexpected.

The boxplot of Final SEC Standing (Chart 4, Appendix A) shows the median at 7.5 and the whiskers extending to the best and worst finishes in the dataset. Since this variable is ordinal and the range of possible values only runs from 1 to 14, the distances between ranks cannot be assumed to be equal; the spread should be read with that limitation in mind. A lower rank number indicates a better finish, so a standing of 1 is the best possible outcome and a standing of 14 is the worst.

Scatterplot and Correlation Analysis

The scatterplot shown in Chart 5 (Appendix A) was used to evaluate the potential linear relationship between athletic department spending and final SEC standing. A regression line was added to visually assess the direction and strength of the relationship.

The scatterplot demonstrates a downward trend, indicating that programs with higher athletic department spending generally achieve better conference standings. Because lower-ranking values represent stronger competitive performance, the negative slope suggests that greater spending tends to be associated with improved outcomes.

However, the observations remain widely dispersed around the regression line, suggesting that athletic department spending alone does not strongly predict final SEC standing.

The coefficient of determination shown on the scatterplot is $R^2 = 0.0874$. This indicates that approximately 8.74% of the variation in final SEC standing can be explained by differences in athletic department spending. The remaining variation is attributable to other factors not included in the model.

Pearson Correlation Coefficient

The Pearson correlation coefficient was calculated to measure the strength and direction of the linear relationship between Athletic Department Spending and Final SEC Standing. Given in Table 2, Appendix A, the calculated Pearson's r is -0.296

This value indicates a weak negative linear relationship between the two variables. As athletic department spending increases, final SEC standing tends to improve, resulting in lower ranking values. While the relationship is consistent with the study hypothesis, the magnitude of the correlation suggests that many additional factors contribute to competitive success.

Examples of factors not included in this analysis include recruiting rankings, coaching quality, player development, injuries, schedule difficulty, institutional resources, and historical program strength.

Regression Equation

By regressing the dependent variable (Final SEC Standing) on the independent variable (Total Athletic Department Spending), we obtained the estimated regression equation of

$y=12.1647+(-0.0289x)$. The slope was calculated as -0.028938 , which means an additional \$1 million increase in total athletic department spending leads to the predicted final SEC standing improving(decreasing) by 0.0289 rank positions. The negative slope aligns with the hypothesized inverse relationship, with higher spending associated with better(lower number) league standings. For a program to improve its predicted standing by one position, it would need to increase spending by \$34.6 million.

The y-intercept of 12.1647 represents the final predicted SEC standing position when total spending equals zero. This value is a theoretical extrapolation completely outside the observed data range of \$82.76 million to \$285.36 million, and no SEC program has zero spending. Thus, this should not be interpreted practically.

The sum of squares regression(SSR) was observed to be 99.4015, measuring the amount of variation in the final SEC standing(dep. variable) that is explained by the regression model. Of the total variation in SEC standings, 99.40 units are attributable to the linear relationship with total athletic department spending. On the other hand, the sum of squares error(SSE)=1038.0985 measures the amount of variation in SEC standing that is not explained by spending. The large SSE relative to SSR indicates that most of the variation in standings comes from the external factors mentioned in the limitations, such as recruiting class rankings, coaching quality/spending, and historical program strength. Together, the SST(Total sum of squares)=1137.5000 represents the total variation in SEC standing, with $SSR+SSE$. With one independent variable in the observation, the MSR represents the average variation explained.

As discussed in the correlation analysis, the coefficient of determination is 0.0874, and 8.74% of the variation in final SEC standing is explained by athletic department spending, while the remaining 91.26% of the variation is due to factors outside the model. The R^2 value of 8.74%

is relatively low, meaning that knowing a program's total athletic department spending provides limited predictive power for determining its final SEC standing. The finding suggests that while total spending and program success are related, a successful season is driven quite heavily by factors beyond financial investment.

The SE value found indicates that, on average, the actual final SEC standing deviates from the standing predicted by the regression by 3.91 rank positions. This is a substantial prediction error given the 1 to 14 ranking scale, while furthering the confirmation that the model provides limited predictive power. The standard error of the slope comes out to 0.011341, which is smaller relative to the absolute value of the slope, indicating a more precise estimate.

Hypotheses:

$H_0: \beta_1 = 0$ (There is no significant linear relationship between athletic department spending and SEC standing.)

$H_1: \beta_1 \neq 0$ (There is a significant linear relationship between athletic department spending and SEC standing.)

Given the t-statistic and p-value found in Table 3, Appendix A (2.5517 > 1.9955, 0.0130 < 0.05, null hypothesis rejected), there is sufficient evidence at the $\alpha=0.05$ significance level that a statistically significant relationship between total athletic department spending and final SEC standing exists. With -2.5517 being negative, the direction of the relationship hypothesized is confirmed, in the context of the research question, supporting that financial investments in athletics are associated with better competitive season outcomes in SEC football

H_0 : The overall regression model is not significant ($\beta_1 = 0$).

H_1 : The overall regression model is significant ($\beta_1 \neq 0$).

From the F-statistic taken from Table 3, Appendix A, since $6.5112 > F\text{-critical}(3.9819)$ and $p\text{-value} < 0.05$, the null hypothesis is rejected, meaning that the regression model is statistically significant.

In both the t-Test and the F-Test, both the hypotheses arrive at the same conclusion with the null hypothesis rejected. Both tests also produce the same p-value of 0.0130, yielding the same reject/fail to reject. However, if two different conclusions were produced, a computational error would have been indicated.

Conclusion

This study examined the relationship between total annual athletic department spending and final SEC standing using simple linear regression. The estimated regression equation, $\hat{y} = 12.1647 + (-0.0289)x$, indicates that for every additional \$1 million in athletic department spending, the predicted final SEC standing improves by approximately 0.0289 rank positions. The findings are directionally consistent with expectations; programs that spend more on athletics do tend to finish higher in the conference standings. Based on both the t-test ($t = -2.5517$, $p = 0.0130 < 0.05$) and the F-test ($F = 6.5112$, $p = 0.0130 < 0.05$), the null hypothesis was rejected. Both tests reach the same conclusion: the relationship between athletic department spending and final SEC standing is statistically significant. However, with a Pearson's r of -0.296 , an R^2 of only 8.74%, and standard error of the estimate (Table 3, Appendix A) indicate that the practical strength of that relationship is weak; statistical significance and practical significance are not the same thing, and spending alone explains a relatively small share of the variation in conference outcomes.

Several limitations affected the strength of these results. The most significant is that spending was measured at the total athletic department level rather than the football program level. Because most SEC schools fund anywhere from 15 to 20 or more varsity sports, a large portion of the spending captured in the independent variable has no direct connection to football performance; this likely diluted the relationship and contributed to the low R^2 observed. Additionally, final SEC standing is an ordinal variable, meaning it reflects rank order but not equal intervals between positions. The gap between finishing first and second may differ considerably from that between finishing seventh and eighth, an assumption that linear regression does not account for. The sample is also limited to a single conference over a fixed number of seasons, which restricts the extent to which these findings can be applied.

It is unclear whether football-specific spending, such as coaching salaries, recruiting budgets, and facilities investment, would predict conference performance more strongly than total athletic spending; this distinction could meaningfully change the results. It is also unknown how much of the unexplained variance is attributable to factors like recruiting class quality, coaching tenure, or historical program strength. Future research could address these gaps by isolating football-only expenditure data, incorporating multi-year recruiting rankings, and including coaching tenure as an additional variable. Expanding the time frame would also increase sample size and statistical power. A multiple regression model that includes several of these predictors simultaneously would provide a more complete and accurate understanding of what drives success in SEC football.

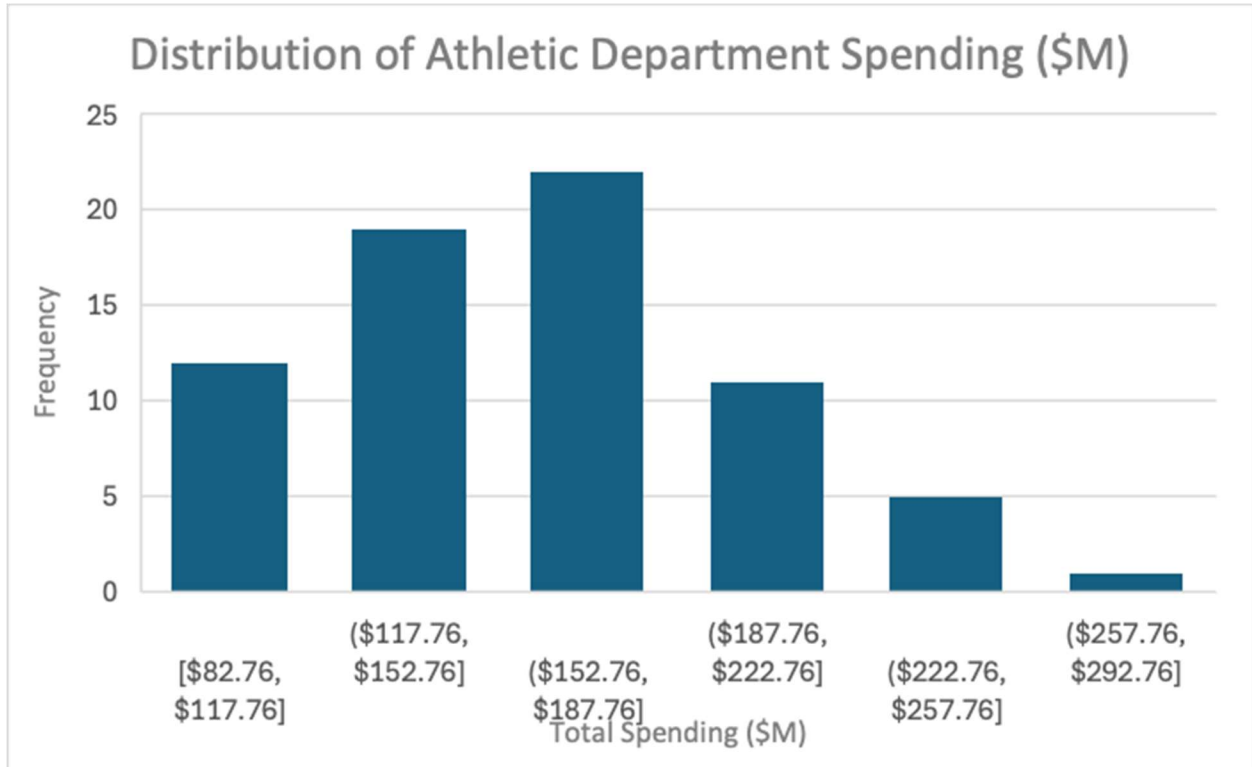
Appendix A

Table 1: Descriptive Statistics for Athletic Department Spending and Final SEC Standing

Statistic	Athletic Dept Spending (X)	Final SEC Standing (Y)
Minimum	\$82.76 million	1 (1st place)
Maximum	\$285.36 million	14 (14th place)
Mean	\$161.20 million	7.5
Median	\$159.97 million	7.5
Standard Deviation	41.48 million	4.06

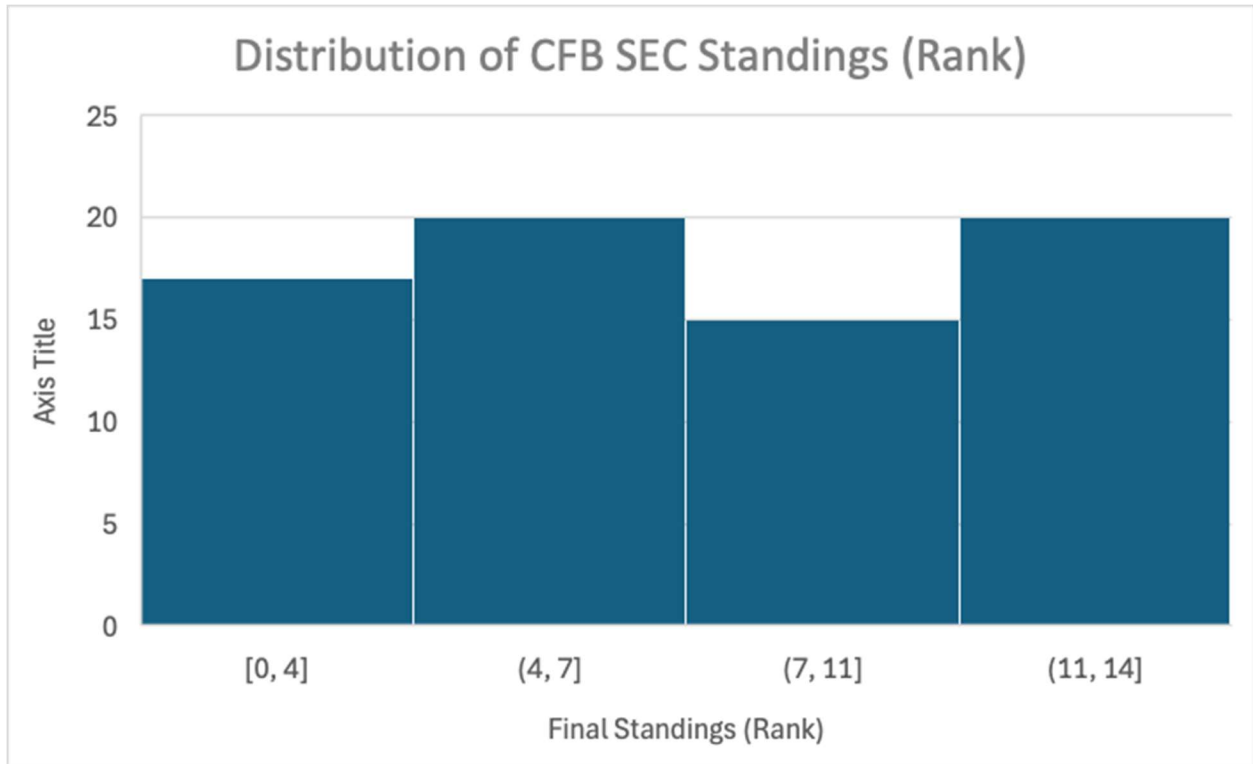
Note: This table shows the main descriptive statistics we calculated for both of our variables. Athletic Department Spending is the independent variable and Final SEC Standing is the dependent variable. We looked at the lowest and highest values, the mean, the median, and the standard deviation to get a better sense of how spread out the data is and where most of the values fall. All spending values are in millions of dollars.

Chart 1: Histogram of Athletic Department Spending



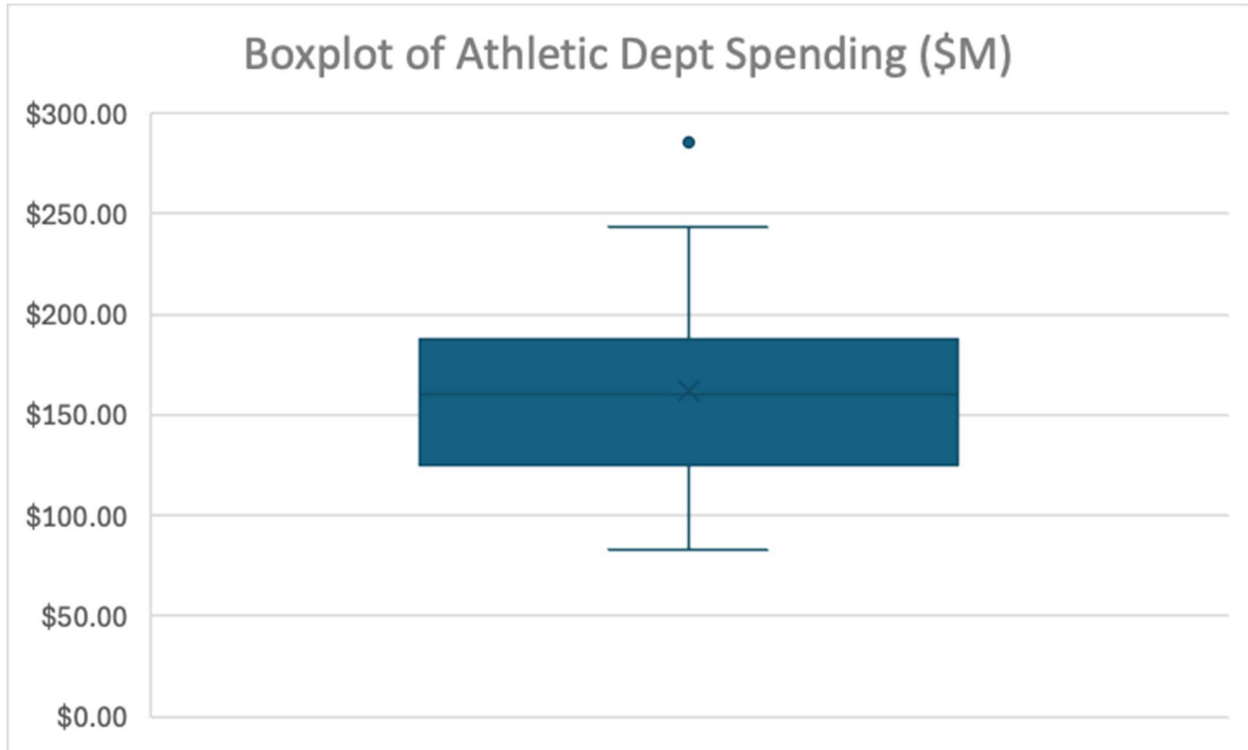
Note: This histogram shows how athletic department spending is spread out across SEC football programs. The x axis uses spending ranges in dollars and the y axis shows how many programs fall into each range. Most programs spend somewhere between \$120 million and \$190 million, so the data is fairly concentrated in that middle range. There are a few programs that spend a lot more than the rest, which pulls the distribution slightly to the right. The mean of \$161.20 million sits right in the middle of where most of the data falls.

Chart 2: Histogram of Final SEC Standing



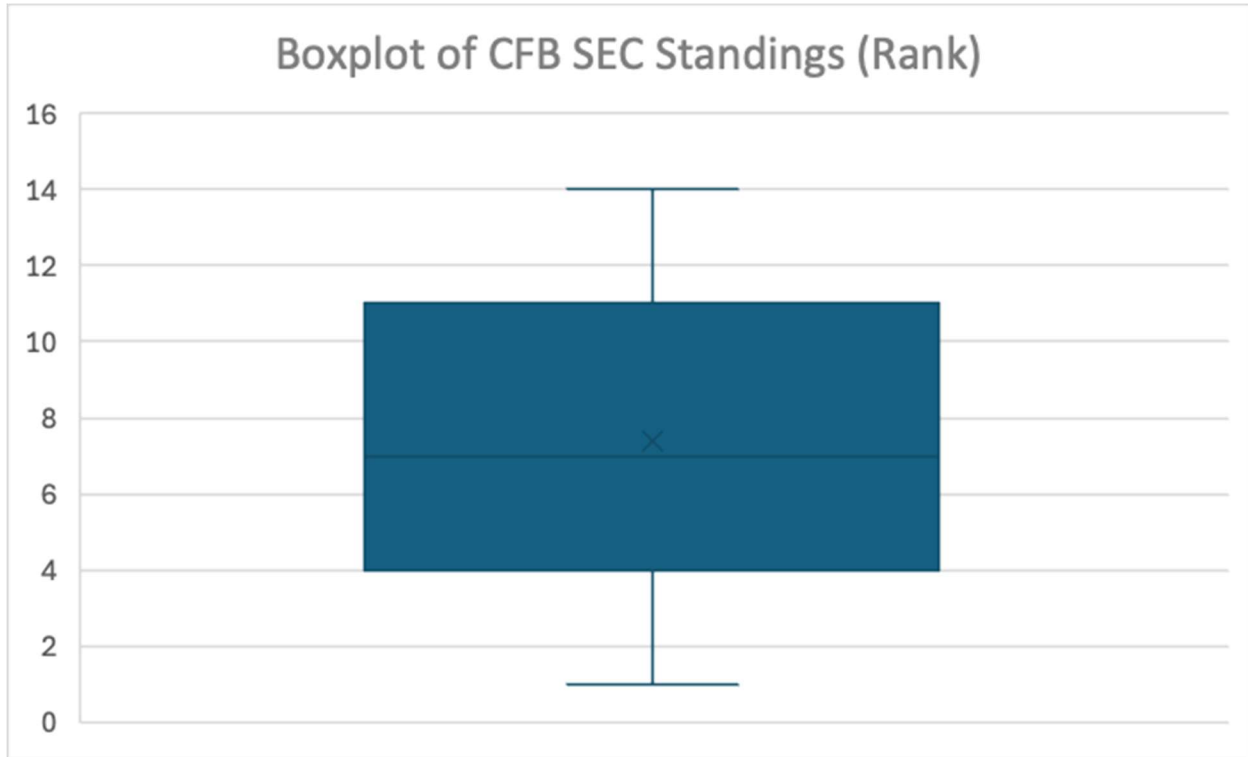
Note: This histogram shows how the final SEC standings are spread out across our sample. The x axis displays different ranking ranges, where lower numbers mean the program finished higher in the conference. The y axis shows how many observations fall into each range.

Chart 3: Boxplot of Athletic Department Spending



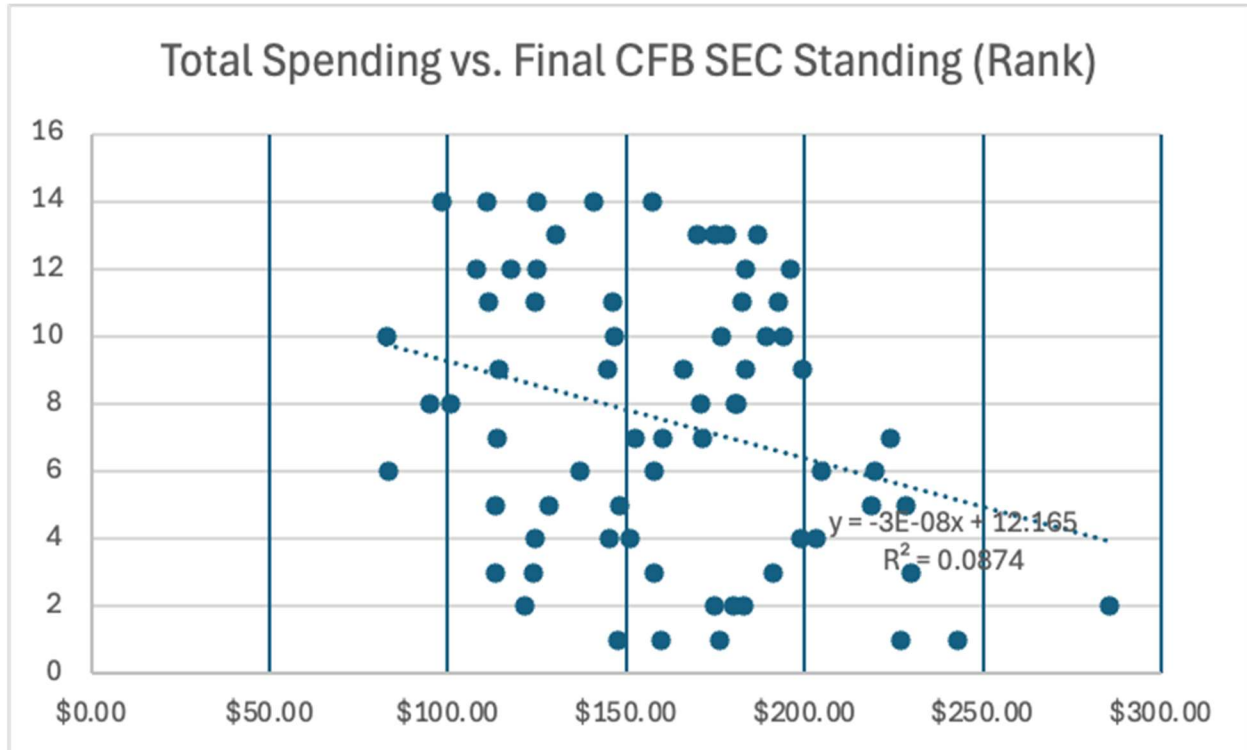
Note: This boxplot shows how athletic department spending is distributed among the programs in our sample. The box in the middle covers the interquartile range, which is where most of the data falls. The line inside the box marks the median spending amount. The whiskers stretch from the lowest value, which was about \$82.76 million, all the way up to the highest before the outlier. There is one dot above the upper whisker, which represents a program that spent significantly more than all the others. We kept this data point in the analysis because it is a real SEC program and not an error.

Chart 4: Boxplot of Final SEC Standing



Note: This boxplot shows how the final SEC standings are spread out across our sample. The box covers the middle 50 percent of the data, the line inside marks the median, and the whiskers reach out to the lowest and highest values in the dataset. Any points that fall outside the whiskers are considered potential outliers. Keep in mind that in this dataset, a lower number means a better finish, so a rank of 1 is the best and a rank of 14 is the worst in the conference.

Chart 5: Scatterplot of Athletic Department Spending and Final SEC Standing



Note: This scatterplot shows the relationship between how much each SEC program spent on athletics and where they finished in the conference standings. Spending is shown on the x axis and the final SEC standing is on the y axis, where a lower number means a better finish. The regression line on the plot slopes downward, which means that programs with higher spending tended to finish better in the conference. However, the points are spread pretty far from the line, which tells us the relationship is not very strong. The Pearson correlation value of r equals negative 0.296 confirms that there is only a weak negative relationship between the two variables.

Table 2: Pearson Correlation Coefficient

Statistic | Value
Pearson's r | -0.296

Note: This table shows the Pearson correlation coefficient we calculated to measure the strength and direction of the relationship between athletic department spending and final SEC standing. We got a value of r equals negative 0.296, which tells us there is a weak negative relationship between the two variables. The negative sign makes sense because a lower ranking number means a better finish, so as spending goes up, the standing number tends to go down. Since the value is close to zero, spending by itself is not a very strong predictor of how well a team does in the SEC

Table 3: Regression Output Summary

Statistic	Value
Slope (b_1)	-0.028938
Intercept (b_0)	12.1647
R ² (Coefficient of Determination)	0.0874
SSR (Sum of Squares Regression)	99.4015
SSE (Sum of Squares Error)	1038.0985
SST (Total Sum of Squares)	1137.5000
MSR (Mean Square Regression)	99.4015
Standard Error of Estimate	3.9072
Standard Error of b_1	0.011341
t-Statistic	-2.5517
p-Value	0.0130
F-Statistic	6.5112

Note: This table summarizes the key regression metrics from the simple linear regression of Final SEC Standing on Total Athletic Department Spending (n = 70).