COLLEGE SPORTS: A GAME OR BUSINESS?

INTRODUCTION

- College athletics are no longer about education and the value of sport (Ridpath, 2017)
- Since 1984, a free market system of revenue generation exists in the industry (Morones)
- The money generated from this free market mostly goes directly to athletic expenditures, which oftentimes surpass the amount of money the schools athletics are bringing in (Zimbalist)
- Knowing what variables predict revenue can help a school benefit their student athletes over other schools, in education and game development

RESEARCH QUESTIONS

- With the recent shuffling of the Power 5 conferences, do schools in larger conferences have higher athletic revenue?
- Do schools with more NCAA championships bring in more athletic revenue than ones with fewer NCAA championships?
- Considering that some college athletes leave school early without graduating to pursue a career in professional athletics, do schools with lower graduation rates bring in more athletic revenue?

MODEL BUILDING

STEP 1: QUANTITATIVE VARIABLES

Initial: AthlRev = B1*AthlStudAid + B2*GSR + B3*Olympic + B4*HCoachesComp + B5*UGPop + B6*NCAATotal + B7*sqrt(NCAATotal)
Revised: AthlRev = B1*AthlStudAid + B2*NCAATotal + B3*sqrt(NCAATotal)

STEP 2: QUALITATIVE VARIABLES

Initial: AthlRev = B1*AthlStudAid + B2*NCAATotal + B3*sqrt(NCAATotal) + B4*PublicPrivatePublic + B5*Religious.AffY + B6*ConferenceBig10 + B7*ConferenceBig12 + B8*ConferenceIvy + B9*ConferencePac12 + B10*ConferenceSEC
Revised: AthlRev = B1*AthlStudAid + B2*CoferenceBig10 + B3*ConferenceBig12 + B4*ConferenceIvy + B5*ConferencePac12 + B6*ConferenceSEC + B7*NCAATotal + B8*sqrt(NCAATotal)

- ConferenceBig10 = 1 if in Big10, 0 if other
- ConferenceBig12 = 1 if in Big12, 0 if other
- ConferenceIvy = 1 if in Ivy, 0 if other
- ConferencePac12 = 1 if in Pac12, 0 if other
- ConferenceSEC = 1 if in SEC, 0 if other

STEP 3: INTERACTIONS

There are no interactions to be considered in the model building process.
Final model: AthlRev = B1*AthlStudAid + B2*CoferenceBig10 + B3*ConferenceBig12 + B4*ConferenceIvy + B5*ConferencePac12 + B6*ConferenceSEC + B7*NCAATotal + B8*sqrt(NCAATotal)

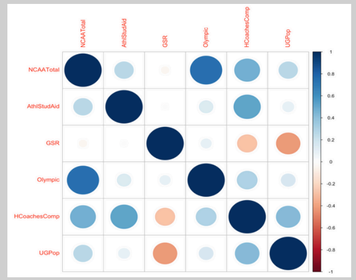
VARIABLE SCREENING: STEPWISE REGRESSION

Stepwise Summary

Step	Variable	AIC	SBC	SBIC	R2	Adj. R2
0	Base Model	1908.729	1913.442	1685.414	0.00000	0.00000
1	HCoachesComp (+)	1858.678	1865.748	1636.257	0.48692	0.48016
2	AthlStudAid (+)	1850.117	1859.544	1627.962	0.55189	0.53994
3	PublicPrivate (+)	1844.577	1856.360	1622.904	0.59318	0.57669
4	NCAATotal (+)	1843.032	1857.172	1621.730	0.61126	0.58995
5	Religious.Aff (+)	1842.158	1858.655	1621.315	0.62532	0.59930
6	Conference (+)	1840.166	1868.446	1613.746	0.67871	0.63076
7	HCoachesComp (-)	1839.559	1865.483	1612.466	0.67292	0.62963

CHECKING FOR MULTICOLLINEARITY

PAIRWISE CORRELATION PLOT



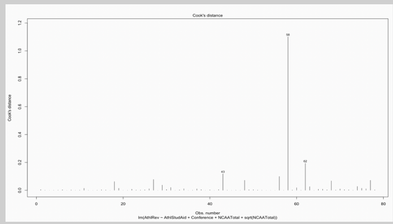
VIF VALUES

NCAATotal	AthlStudAid	GSR	Olympic	HCoachesComp	UGPop
2.649	1.475	1.332	2.309	2.077	1.428

[1] 1.878333

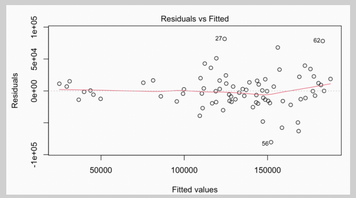
- None of the individual VIF values exceed 10.
- The average VIF of 1.878 is less than 3.

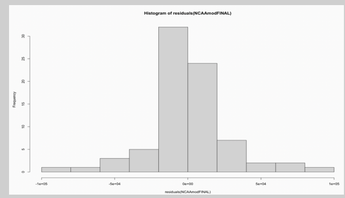
IDENTIFYING INFLUENTIAL OBSERVATIONS



- Observations 43, 58, and 62 are influential.
- These observations represent University of Notre Dame, Stanford, and University of Texas.

CHECKING MLR ASSUMPTIONS





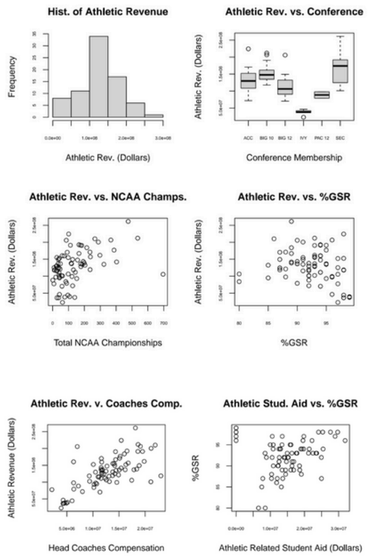
- No patterns in individual residual plots; observations are independent.
- A weighted least squares model will be fit to correct the fanning out pattern.

ADDITIONAL TECHNIQUE: WEIGHTED LEAST SQUARES REGRESSION

```
#weighted least squares
wt<- 1/lm(abs(NCAAmodFINAL$residuals)~(NCAAmodFINAL$fitted.values))$fitted.values^2
wls_model<-lm(AthlRev~AthlStudAid+Conference+NCAATotal+sqrt(NCAATotal), data = NCAAupdated, weights=wt)
summary(wls_model)
```

- Final MLR model: adjusted R^2= 0.6408; WLSR model: adjusted R^2 = 0.8221
- The WLSR accounts for more of the variability in athletic revenue than MLR.

EDA AND DATA DICTIONARY



Variable Name	Abbreviated Name	Description
Total NCAA Championships	NCAATotal	Team and individual NCAA championships won by an institution
Public/Private Status	PublicPrivate	Labeled "Public" if the institution is Public and labeled "Private" if the institution is private
Athletic Student Aid	AthlStudAid	Total amount spent on athletically related student aid, measured in thousands of dollars
Graduation Success Rate	GSR	The percentage of first-year, full-time student athletes who graduated within a six year time frame, as measured by the NCAA
Religious Affiliation	ReligiousAff	Labeled "Y" if the institution is religiously affiliated and "N" if not
Olympic Athletes	Oylmpic	Olympic Athletes from a school who participated in either the 2022 or 2024 Olympic games
Head Coaches Compensation	HCoachesComp	Average annual salary for non-volunteer head coaches, measured in thousands of dollars
Undergraduate Pop	UGPop	Undergraduate students enrolled as designated by the institution; measured in thousands
Conference Membership	Conference	The NCAA-designated conference the institution is a part of, as of 2024
Athletic Revenue	AthlRev	Response Variable; total revenue from intercollegiate athletic activities, measured in thousands of dollars

CONCLUSIONS

PREDICTION EQUATION

$$\widehat{AthlRev} = 73970 + 1.538AthlStudAid + 9608ConferenceBIG10 - 11310ConferenceBIG12 - 65680ConferenceIVY - 31210ConferencePAC12 + 29520ConferenceSEC - 98.880NCAATotal + 4575sqrt(NCAATotal)$$

- With a p-value of far less than .05, we reject the null hypothesis. There is evidence that our model is adequate at predicting athletic revenue of a school. The individual betas of AthlStudAid, square root of NCAA total, and of the dummy variables ConferenceIVY and Conference SEC were found to be significant as well
- Our model can help predict revenue for schools in this era as college athletics change for a program. For example, our model can help predict revenue for Boise State University as they will transition to the PAC 12 Conference soon.
- Example prediction: University of Iowa, with NCAATotal = 166, AthlStudAid = 13450, Conference = BIG10 --> predicted AthlRev = 146794, actual = 166887
- Our predict is only off by 20,000,000
- In the future, we hope to add more schools and more predictors to make our model more accurate and applicable to a wider range of college athletics programs.

REFERENCES

- References
- *CFR 668.47 -- Report on athletic program participation rates and financial support data.* (n.d.). eCFR. Retrieved October 27, 2024, from https://www.ecfr.gov/current/title-34/subtitle-B/chapter-VI/part-668/subpart-D/section-668.47
- *Championships Summary*. (2024, June 24). NCAA. Retrieved October 27, 2024, from http://fs.ncaa.org/Docs/stats/champs_records_book/Overall.pdf
- "Equity in Athletics Disclosure Act (EADA). (2024). U.S. Department of Education. Retrieved from https://ope.ed.gov/athletics"
- *Graduation Rates*. (2013, November 19). NCAA.org. Retrieved October 27, 2024, from https://www.ncaa.org/sports/2013/11/19/graduation-rates.aspx
- Kajak, E. (2023, December 27). *Following the Money in College Sports*. Morones Analytics. https://moronesanalytics.com/following-the-money-in-college-sports/#:~:text=The%20Big%2D10%20followed%20suit,of%20that%20revenue%20in%202022 intro
- *Shared NCAA Research Data*. (2016, December 14). NCAA.org. Retrieved October 30, 2024, from https://www.ncaa.org/sports/2016/12/14/shared-ncaa-research-data.aspx
- Ridpath, B. D. (2017, January 31). *Unwinding Madness in College Sports? A New Book Shows How!*. Forbes. https://www.forbes.com/sites/bdavidridpath/2017/01/30/unwinding-madness-in-college-sports-a-new-book-shows-how/ intro
- Zimbalist, A. (2023, March 18). *Analysis: Who is Winning in the High-revenue World of College Sports?*. PBS. https://www.pbs.org/newshour/economy/analysis-who-is-winning-in-the-high-revenue-world-of-college-sports intro

XOXO - the data driven divas