SCHOOL OF INFORMATION
UNIVERSITY OF MICHIGAN

SIADS 593: Team Project Feedback

1.0 Project

Does your college basketball program matter? An analysis of NCAA Division 1 college basketball programs based on the NBA players they produce.

Team: 24-toobrite-highad

Mentor: Kyle Balog

2.0 Feedback

2.1 Background and Motivation (15/15)

Descriptive background and project goals, detailing the salary cap provided helpful context. Also nice job citing previous studies and clearly stating what your project aims to further contribute while supporting this hypothetical scenario!

2.2 Data Sources (30/30)

Links, format, retrieval method, important variables, and relevant time period are all provided along with the size/granularity of the final data retrieved. Great work!

2.3 Cleaning and Manipulation (100/100)

Solid descriptions of both the individual data processing steps as well as the join methodology. Handling different character conventions for names between sources definitely sounded like a challenge, if only there were a more globally unique identifier somehow available.

2.4 Analysis (58/60)

For the turnover analysis it could be more interesting to normalize by playing time or some other metric that accounts and adjusts for this strong positive correlation observed which is likely due to the underlying variable of playing time. Also should list the specific correlation coefficient beyond just describing it as a moderately strong positive correlation, since people's assessment of strength is not always standardized (though it should be). For the regression analysis it is generally encouraged to provide the model output itself in a technical report, potentially in an appendix if not directly on the slide. The interpretation of variable coefficients was useful in building understanding though, but R^2 is typically always interpreted in terms of what percent of the variance is captured instead of describing 1-R^2 as the percent of variance remaining. Its interesting that the NBA model was so much more performant than college, could be useful to have hypothesized why with the analysis.

2.5 Visualizations (48/50)

Great job constructing these *python* visualizations :) They all successfully supported your narrative while communicating the key points that were being discussed. The only feedback I have is for the Distribution of contract values between 1 or more seasons to have gone for an overlaid histogram rather than stacked, as stacking hides the true distribution of the more than one data here. By adjusting the transparency and overlaying you could have potentially captured both distributions more effectively. Also the right histogram on page 7 had the Count label cut off, but otherwise nice titles, legends, and information density on the visualizations used.

2.6 Statement of Work and References (15/15)

Detailed responsibilities and collaboration methodologies listed. Excellent work and collaboration team!

2.7 Spelling, Grammer, Styling, etc. (29/30)

Could introduce figure labels to make referencing particular visualizations more convenient. In text citations are also typically useful to more easily link the subsequently referenced work by embedding the author's last name.

2.8 Project Code (149/150)

Extensive use of comments and markdown to guide reader, thank you!Requirements.txt also provided, was able to replicate environment and run code. Great use of functional programming to preform the separate steps for generating the data and performing analysis where appropriate, though your functions could use more parameters to pass input variables/data which make them more generalizable. Descriptive variable names. Visualizations were mostly all labeled, titled, and thoughtfully designed. Nice code and workflow!

3.0 Total Points Earned

Project Report: 295/300 (98.33%)

Project Code: 149/150 (99.33%)

Total: 444/450 (98.67%)

Timestamp: 2024-03-29 08:44:06.198471+00:00 UTC