AN EFFICACY RATING FOR MARCH MADNESS TOURNAMENT SEEDING
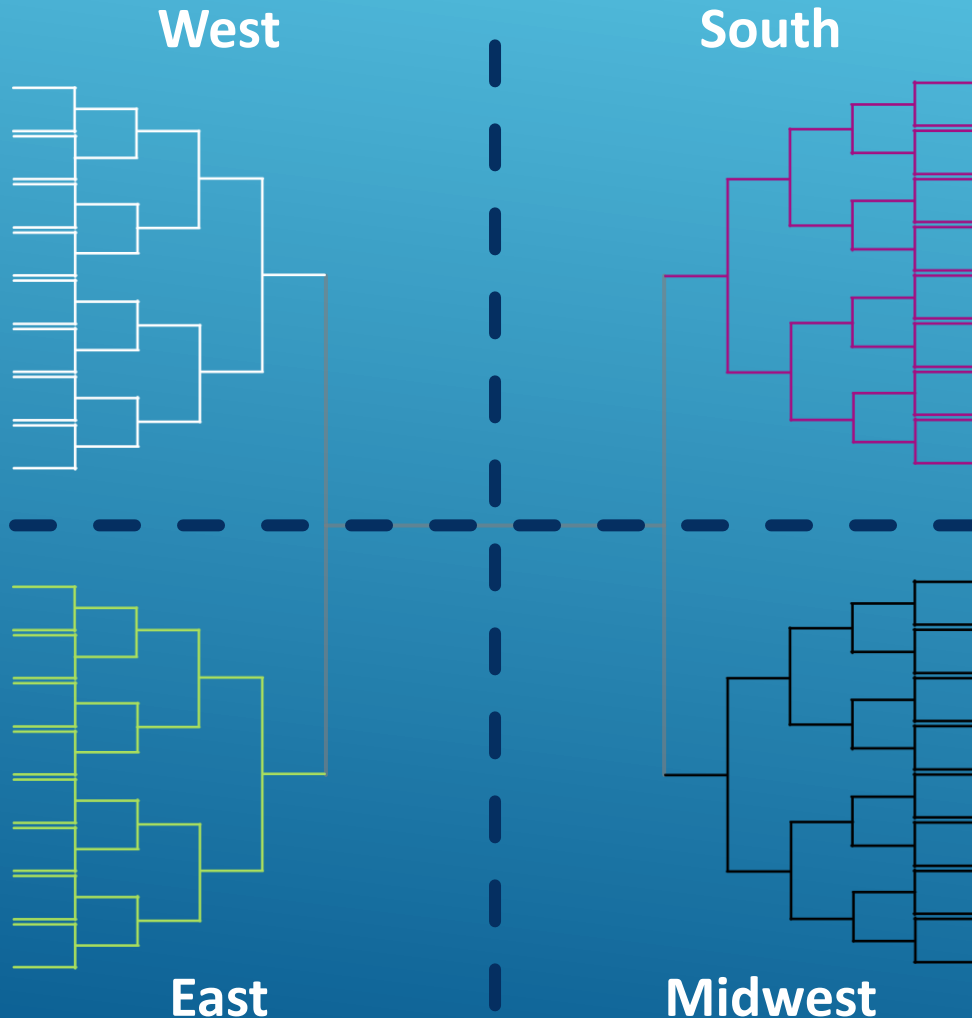
Dr. Joe DeMaio and Nathalie Jones

BACKGROUND

West

South

East

Midwest

March Madness & NCAA Seeding

- Tournament began seeding in 1979
  → Initially, 32 Team bracket

- 64-team bracket began in 1985:
  → Teams split into 4 regions
  → Ranked 1 – 16 within each region

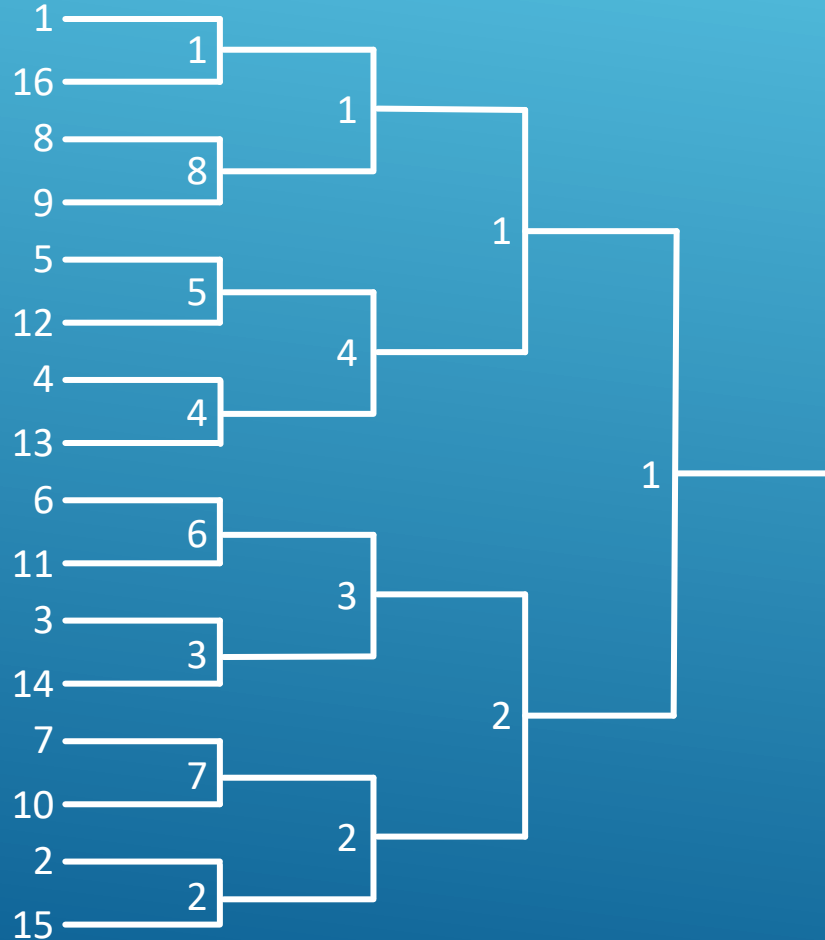- Lower seeds expected to place higher in the tournament

BACKGROUND

March Madness & NCAA Seeding

- Tournament began seeding in 1979
  → Initially, 32 Team bracket

- 64-team bracket began in 1985:
  → Teams split into 4 regions
  → Ranked 1 – 16 within each region

- Lower seeds expected to place higher in the tournament

PROJECT PURPOSE

• How accurate are team seedings?

• Has seeding efficacy increased or decreased over time?

• What influential factors exist?

• Does seeding efficacy vary across tournaments?

ABOUT THE DATA

Data Collected

- NCAA bracket information
  - 1985 – 2023 Men's tournament
  - 1994 – 2023 Women's tournament

- Player Salaries
  - **36 years** of NBA player salaries
  - **5 years** of WNBA player salaries

- Early Entrant Players
  - **38 years** of NBA early entrant players
  - **2 years** of WNBA early entrant players

ABOUT THE DATA

Data Scraping Methods

- Python libraries: "requests," "beautifulsoup"

- NBA data available but scattered across various websites and webpages
    - → Wikipedia
    - → ESPN
    - → Basketballreference.com

- Limited sources of WNBA data
    - → Spotrac (player salaries)
    - → DraftKings (early entrants)

PROPOSED METRIC

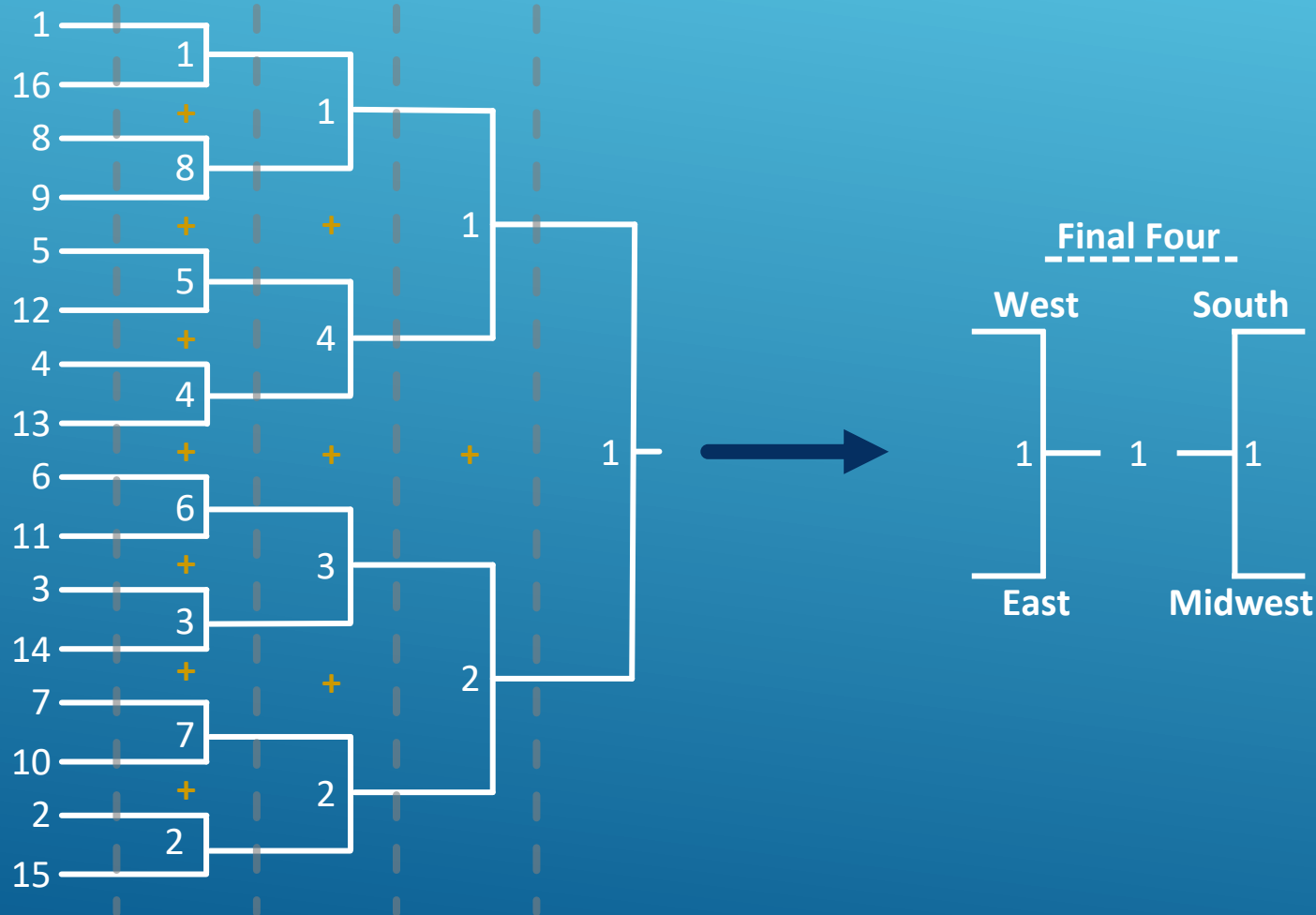- Proposed metric sums seed value for each team at a fixed round, i:

$$E(year, gender) = \sum_i R(year, round(i), gender)$$

- The error rate can then be calculated by:

$$e = \frac{E(year, gender) - Best\ Score}{Worst\ Score - Best\ Score}$$

MEASURING TEAM SEEDING EFFICACY

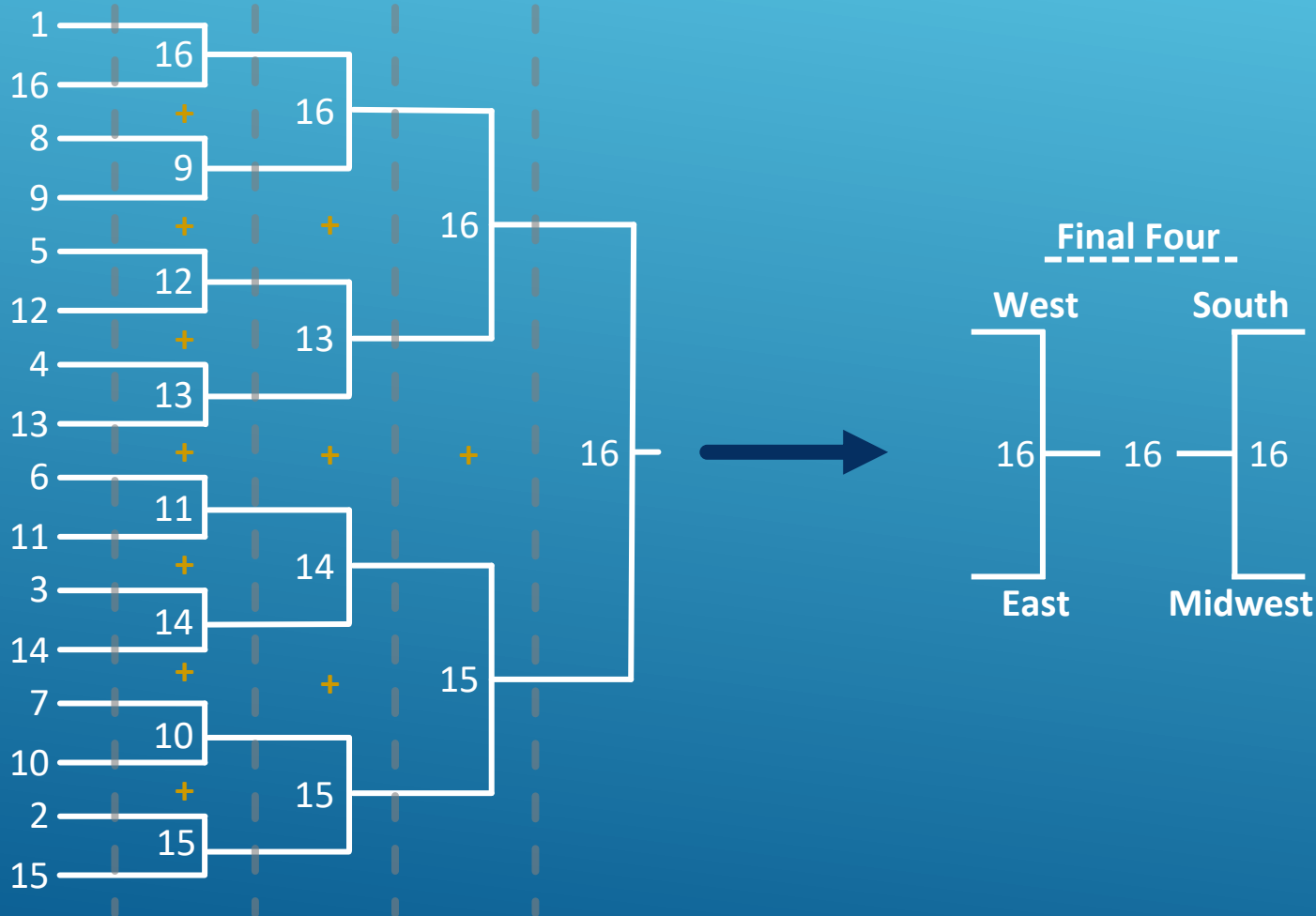(36 + 10 + 3 + 1) x 4 Regions + 1 + 1 + 1 = 203

→ If the tournament goes as expected,

$$E(year, gender) = 203$$

$$e = \frac{203 - 203}{868 - 203} = \frac{0}{665} = 0.0$$

MEASURING TEAM SEEDING EFFICACY

$(100 + 58 + 31 + 16) \times 4 \text{ Regions} + 16 + 16 + 16 = 868$

→ In the worst case,

$$E(year, gender) = 868$$

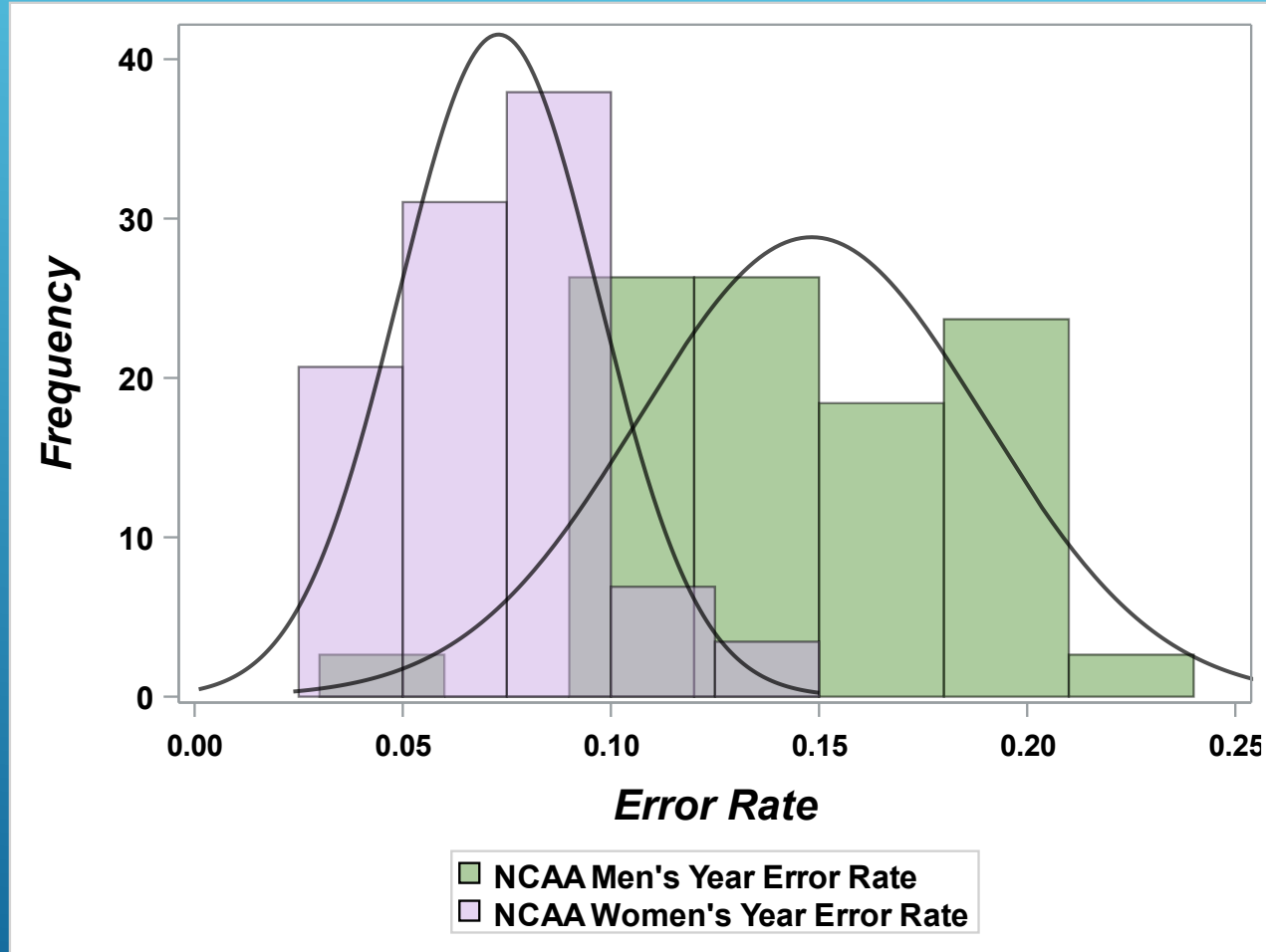$$e = \frac{868 - 203}{868 - 203} = \frac{665}{665} = 1.0$$

ERROR RATE EXAMPLE

Round Number	$E(2023, m)$	Best Score	Worst Score
1	191	144	400
2	78	40	232
3	37	12	124
4	23	4	64
5	9	2	32
6	4	1	16
Sum Total	342	203	868

$$e = \frac{E(2023, m) - Best\ Score}{Worst\ Score - Best\ Score} = \frac{342 - 203}{868 - 203} = \frac{139}{665} = 0.201$$

ERROR RATE ACROSS TOURNAMENTS

Distribution of Year Error Rates for NCAA Men's & Women's Leagues
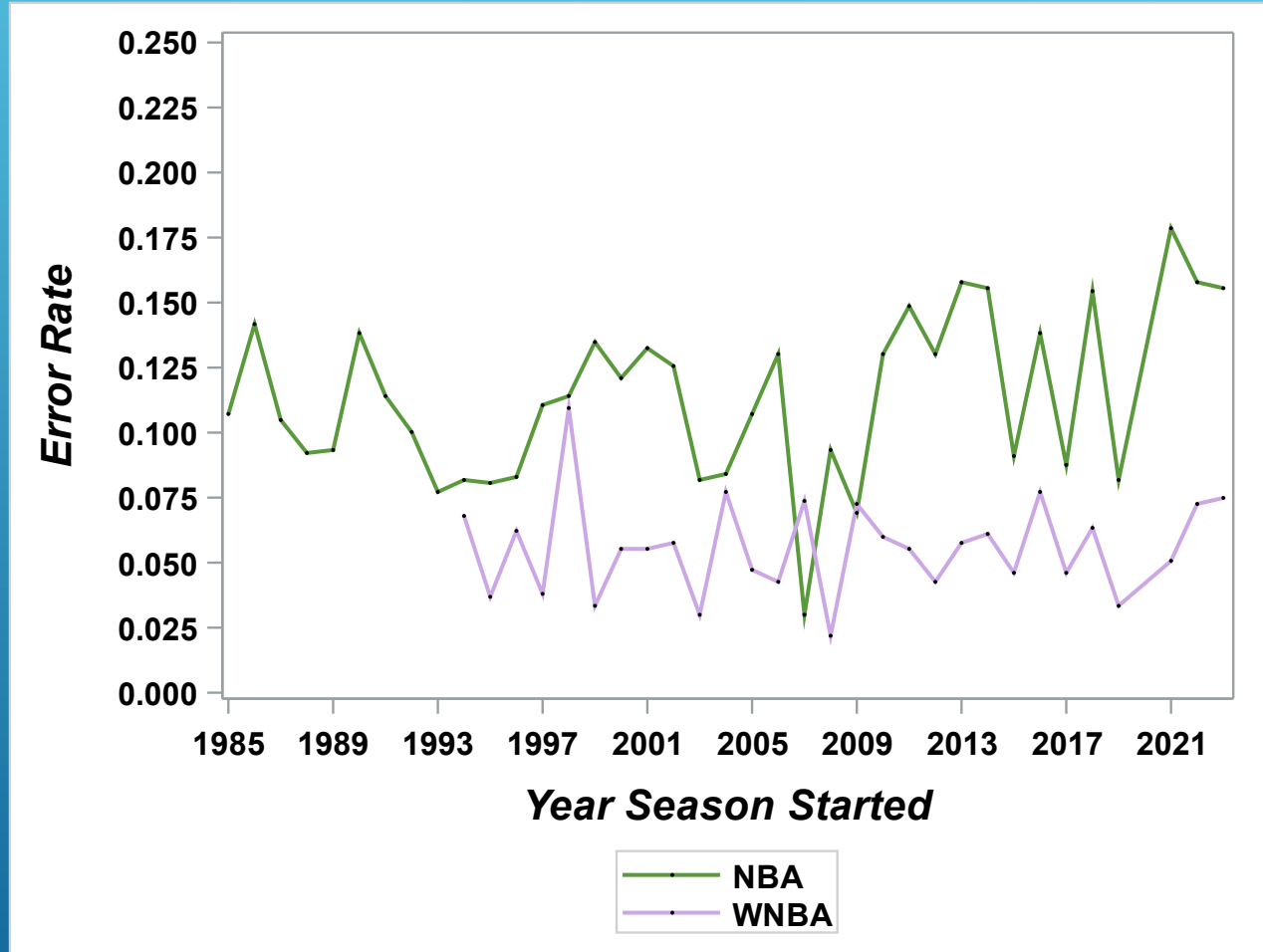
**Average Error Rates Overall**

Men's	Women's
15 %	7 %

- Slightly less variation in women's tournaments error rates

SESUG 2023
- OCTOBER, 22-24 -

ERROR RATE BY GENDER OVER TIME

Year Error Rates for Men's & Women's Tournaments Over Time

- Lower error rate in women's tournaments every year except 2006

- Slight overall increase in men's tournaments

- 1998 Largest error rate in women's history

- Harvard (16) vs. Stanford (1), 1998
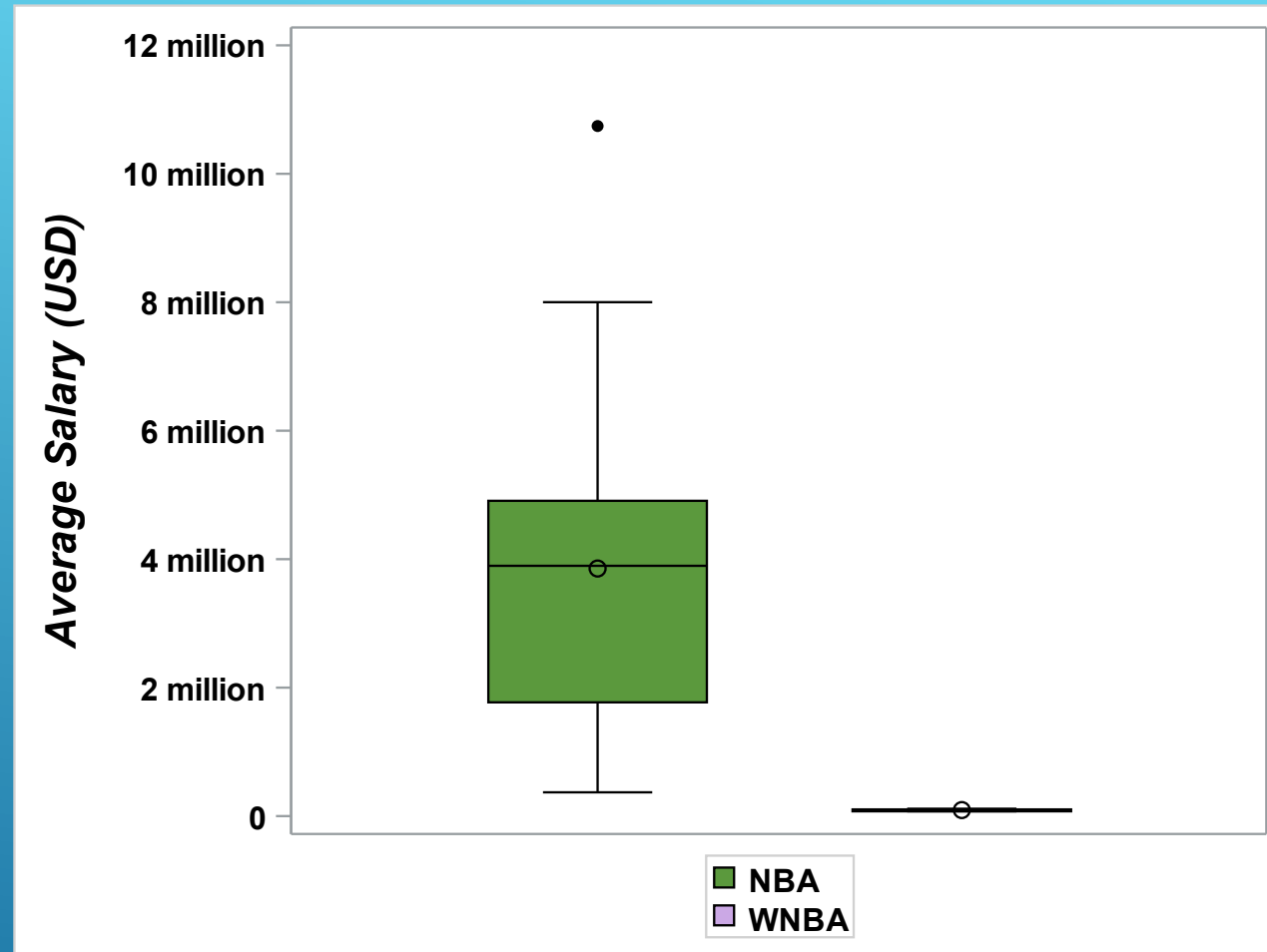
Average Salary for NBA and WNBA Players

EXTREME PAY DISPARITY

INFLUENCING FACTORS: SALARY

Name	Position	2022 Salary
	Average WNBA Player	$97,381
	Average NBA Player	$8.2 Million
Steph Curry	Highest paid NBA player	$51.9 Million
Diana Taurasi	Highest paid WNBA player	$235,936
Rocky the Mountain Lion	NBA Mascot	$625K
Harry the Hawk	NBA Mascot	$600K
Benny the Bull	NBA Mascot	$400K
Go the Gorilla	NBA Mascot	$200K

CONCLUSION

Project Purpose	NCAA March Madness Tournament	
	Women's Tournament	**Men's Tournament**
1. How Accurate are Team Seedings?	7% error rate (overall)	15% error rate (overall)
2. Has Seeding Efficacy Increased or Decreased Over Time?	Relatively steady	Slightly increasing
3. What Influential Factors Exist?	Men's & Women's salaries significantly disparate	
4. Does Seeding Efficacy Vary Across Leagues?	Women's error rate significantly lower	

QUESTIONS

THANK YOU!

APPENDIX

FUTURE WORK

Tournament Simulation

- Aim to study relative distribution of e to different probability estimators

- Fix win probabilities at 50% for each team

- Use historical probabilities for the likelihood that seed i beats seed j for $i \neq j$
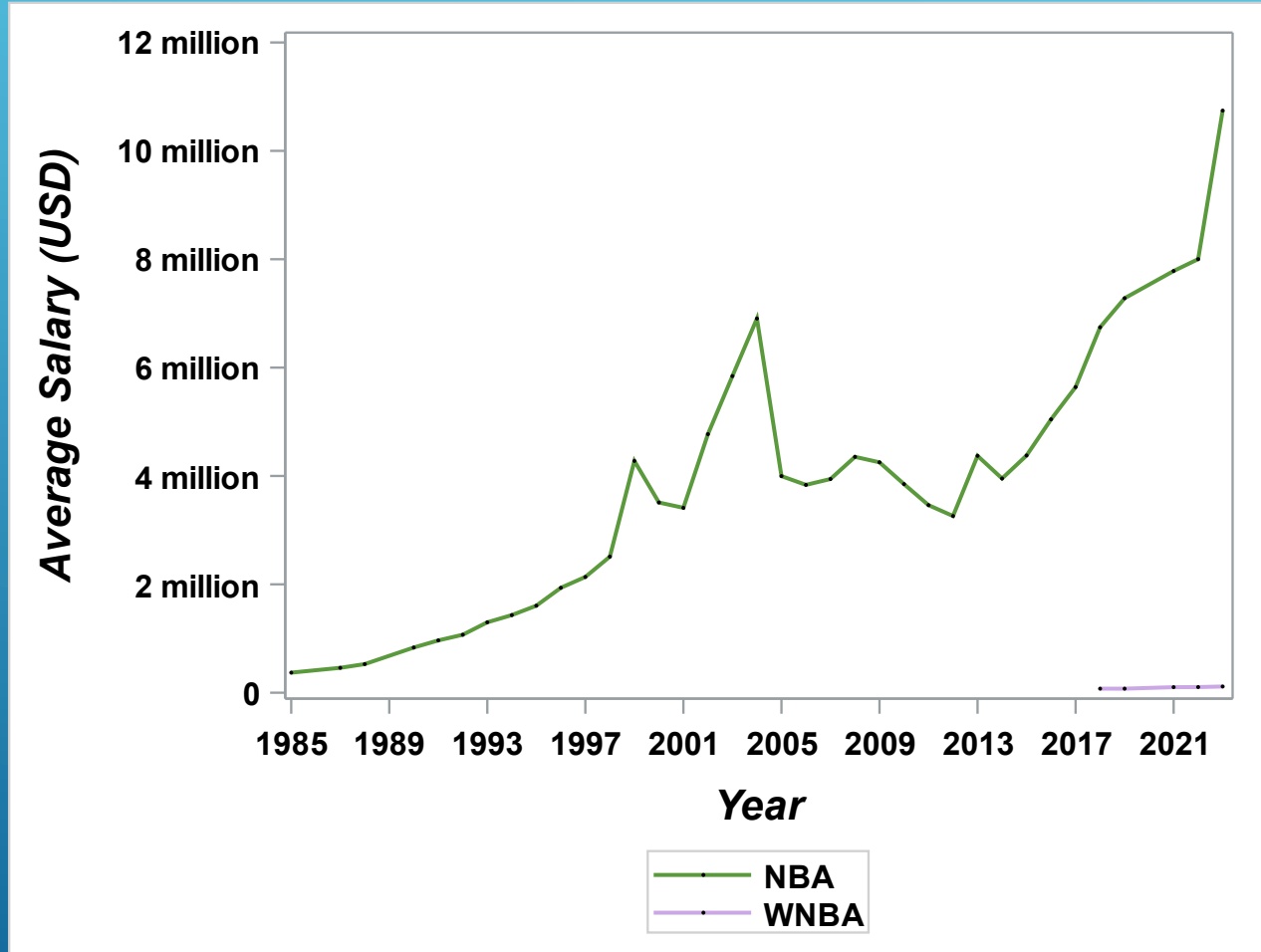
Men's vs. Women's Tournament Comparisons

- Analyze time spent playing for top earners within each league

- Analyze the number of top-performing players not drafted

- Analyze placement of top players during the draft

NBA & WNBA SALARIES

AVERAGE NBA PLAYER SALARY (1985 – 2023)

Average Salary for NBA and WNBA Players Over Time

- Little data on WNBA players, only 5 years publicly available

- NBA player salaries display a positive, increasing relationship

- Time accounts for 72.27% of the variation in NBA player salaries

- Salaries should increase by $176,262 – $213,890 per year
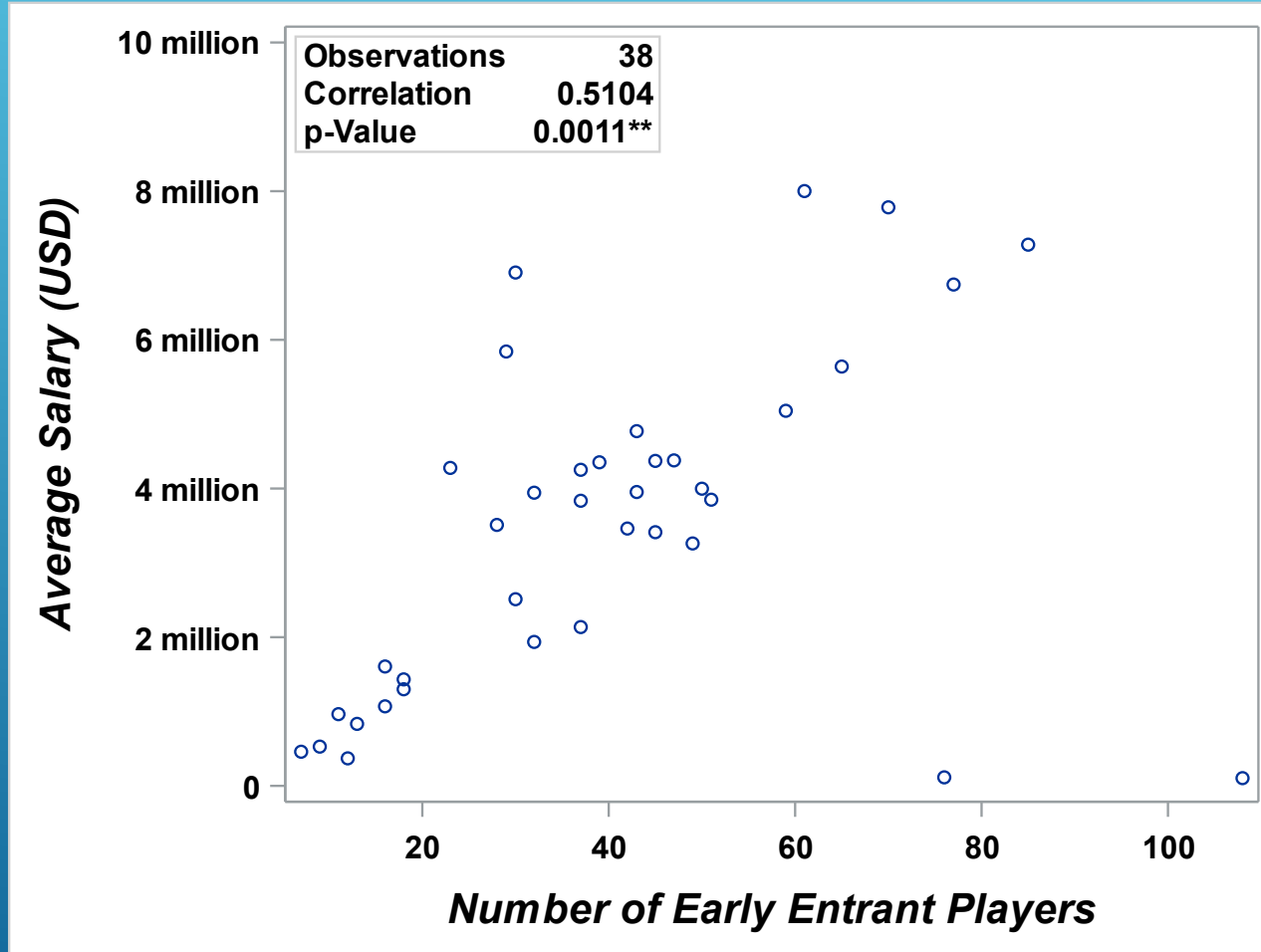
NUMBER OF EARLY ENTRANT PLAYERS
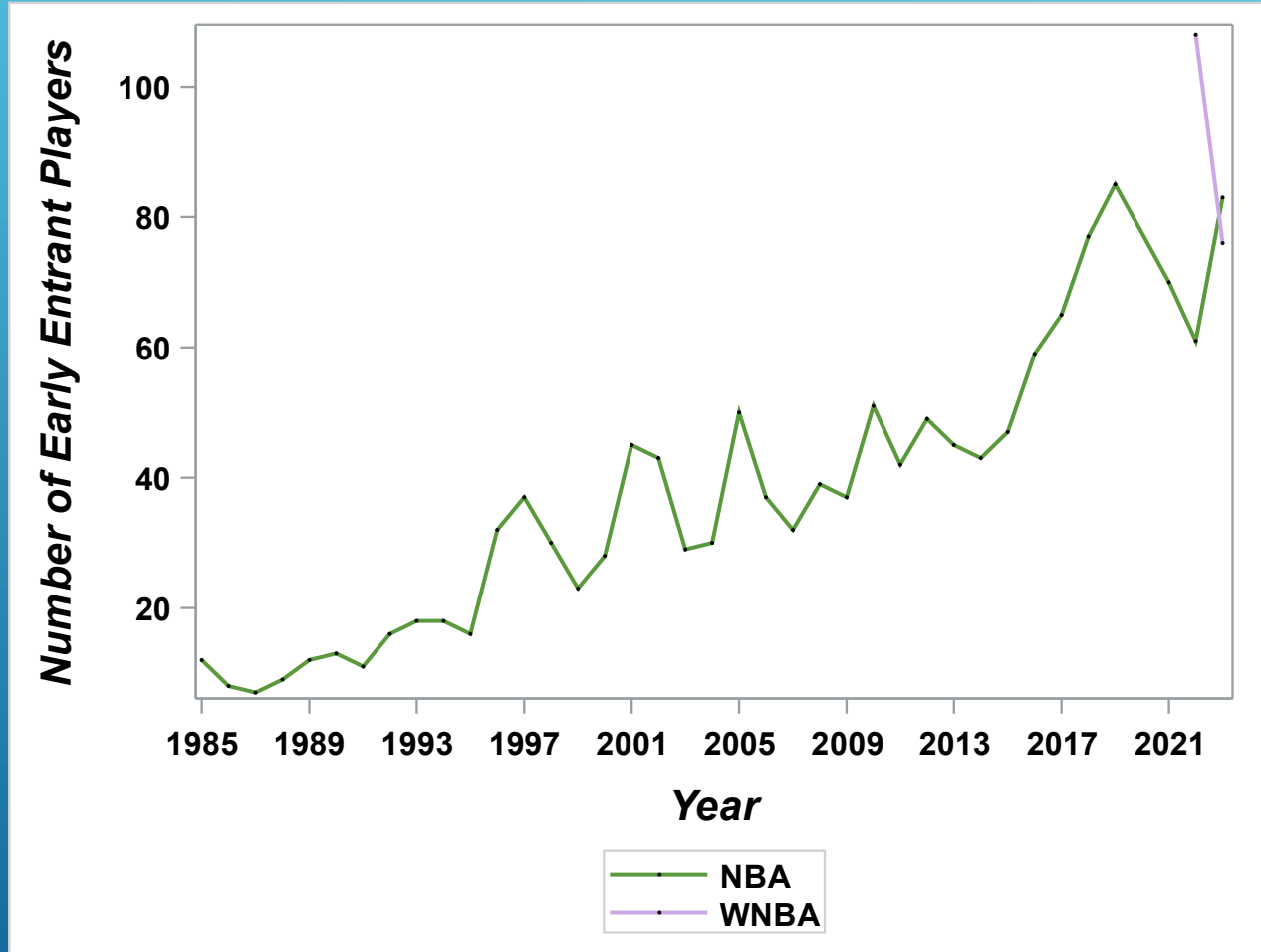
- Little data on WNBA early entrant players

- Positive, increasing relationship

- The number of early NBA entrants should increase by 1 – 2 players each year

RELATIONSHIP BETWEEN FACTORS

INFLUENCING FACTORS' RELATIONSHIP

Correlation Between Average NBA Salaries and Number of NBA Early Entrants

- Correlation: 51%, p-value: 0.0011**

- Relatively strong, positive relationship

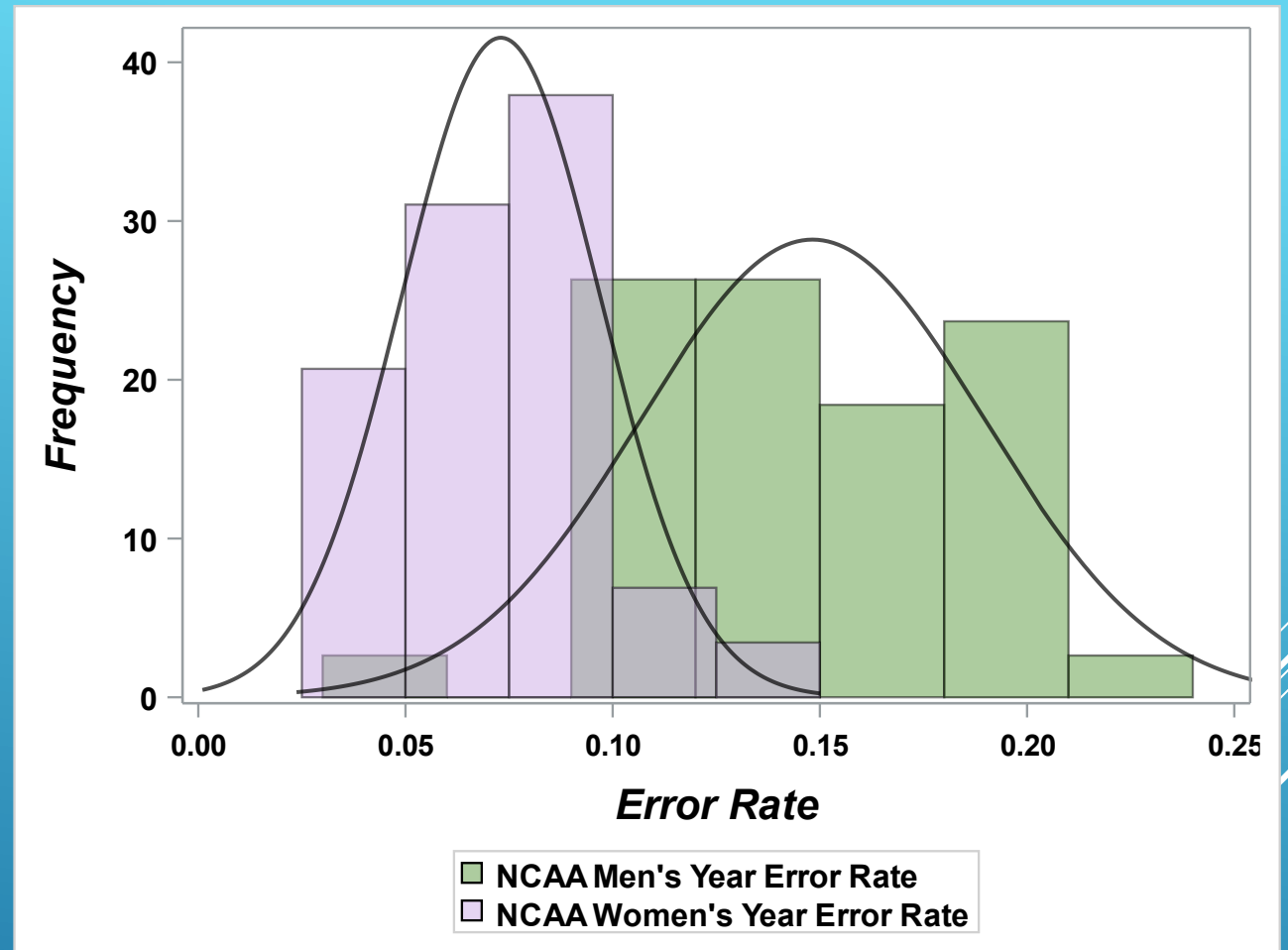- As salaries increase, so should the number of NBA early-entrant players

CODE USED

```
ODS _all_ close;
ODS html path = "&filepat" file = "visuals.html";
ODS graphics / outputfmt = svg;
PROC sgplot
  DATA = out.yr_clean;
  HISTOGRAM NBA_YEAR_ERROR /
    fillattrs = (color = "#5B993D")
    transparency = 0.5 name = "men"
    legendlabel = "NCAA Men's Year Error Rate";
  HISTOGRAM WNBA_YEAR_ERROR /
    fillattrs = (color = "#CBA8E5")
    transparency = 0.5 name = "women"
    legendlabel = "NCAA Women's Year Error Rate";
  DENSITY NBA_YEAR_ERROR / lineattrs = (color = black)
    transparency = 0.3;
  DENSITY WNBA_YEAR_ERROR / lineattrs = (color = black)
    transparency = 0.3;
  KEYLEGEND "men" "women" / autoitemsize down = 2
    valueattrs = (size = 12 weight = Bold);
  XAXIS label = "Error Rate"
    values = (0.00 to 0.25 by 0.05)
    labelattrs = (size = 16 style = Italic weight = Bold)
    valueattrs = (size = 11 weight = Bold);
  YAXIS label = "Frequency"
    labelattrs = (size = 16 style = Italic weight = Bold)
    valueattrs = (size = 12 weight = Bold);
RUN;
ODS html close;
ODS listing;
```
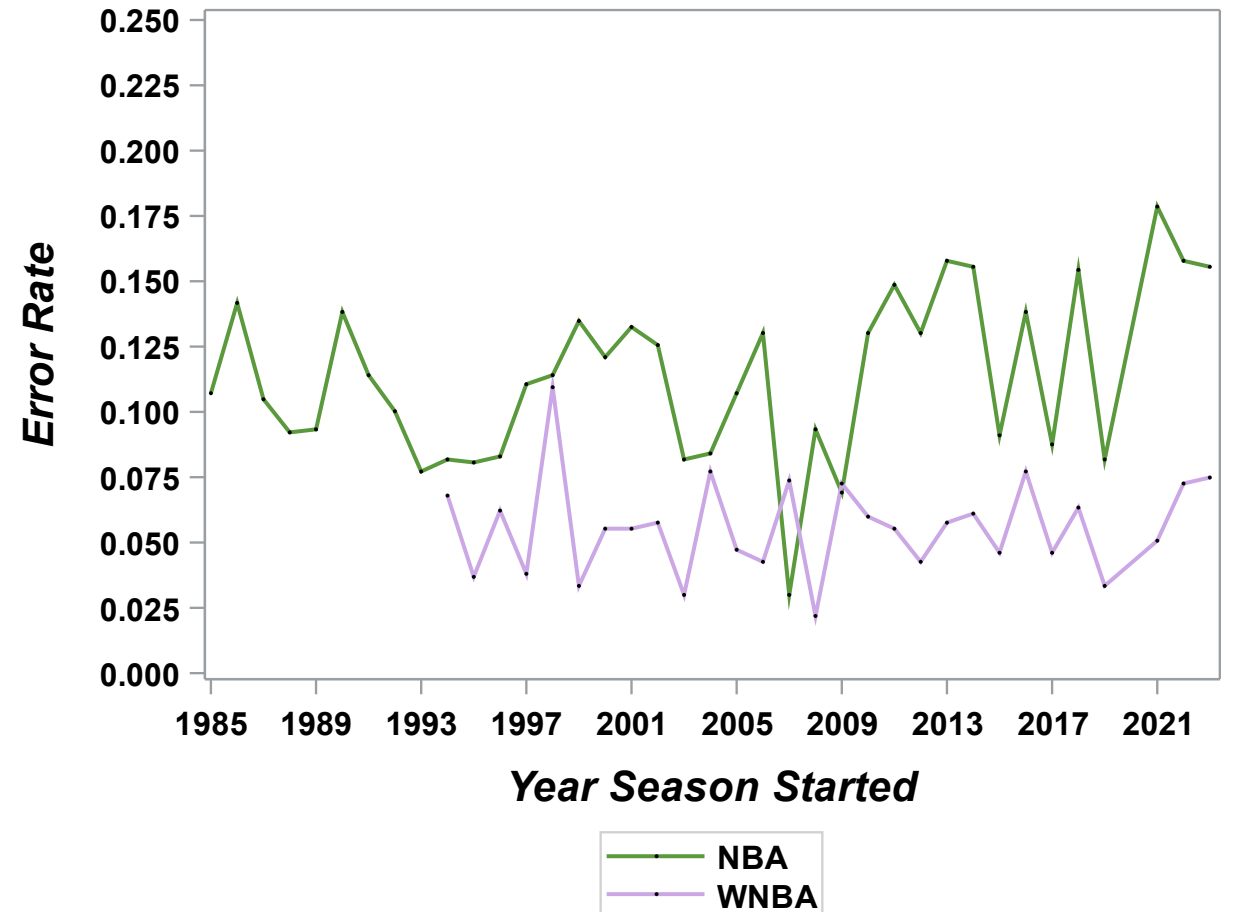

Distribution of Year Error Rates for NCAA Men's & Women's Leagues

ERROR RATE ACROSS TOURNAMENTS

```
ODS _all_ close;
ODS html path = "&filepat" file = "visuals.html";
ODS graphics / outputfmt = svg;
PROC sgplot
  DATA = out.yrlg_clean;
  STYLEATTRS datacontrastcolors = ("#5B993D" "#CBA8E5");
  SERIES x = YEAR y = YEAR_ERROR / group = LEAGUE_ID
    markers lineattrs = (thickness = 2) name = "lines"
    markerattrs = (symbol = circlefilled color = black
      size = 3);
  KEYLEGEND "lines" / down = 2
    valueattrs = ( size = 12 weight = Bold);
  XAXIS label = "Year Season Started"
    values = (1985 to 2023 by 1)
    labelattrs = (size = 16 style = Italic weight = Bold)
    valueattrs = (size = 12 weight = Bold);
  YAXIS label = "Error Rate"
    values = (0.000 to 0.250 by 0.025)
    labelattrs = (size = 16 style = Italic weight = Bold)
    valueattrs = (size = 12 weight = Bold);
RUN;
ODS html close;
ODS listing;
```
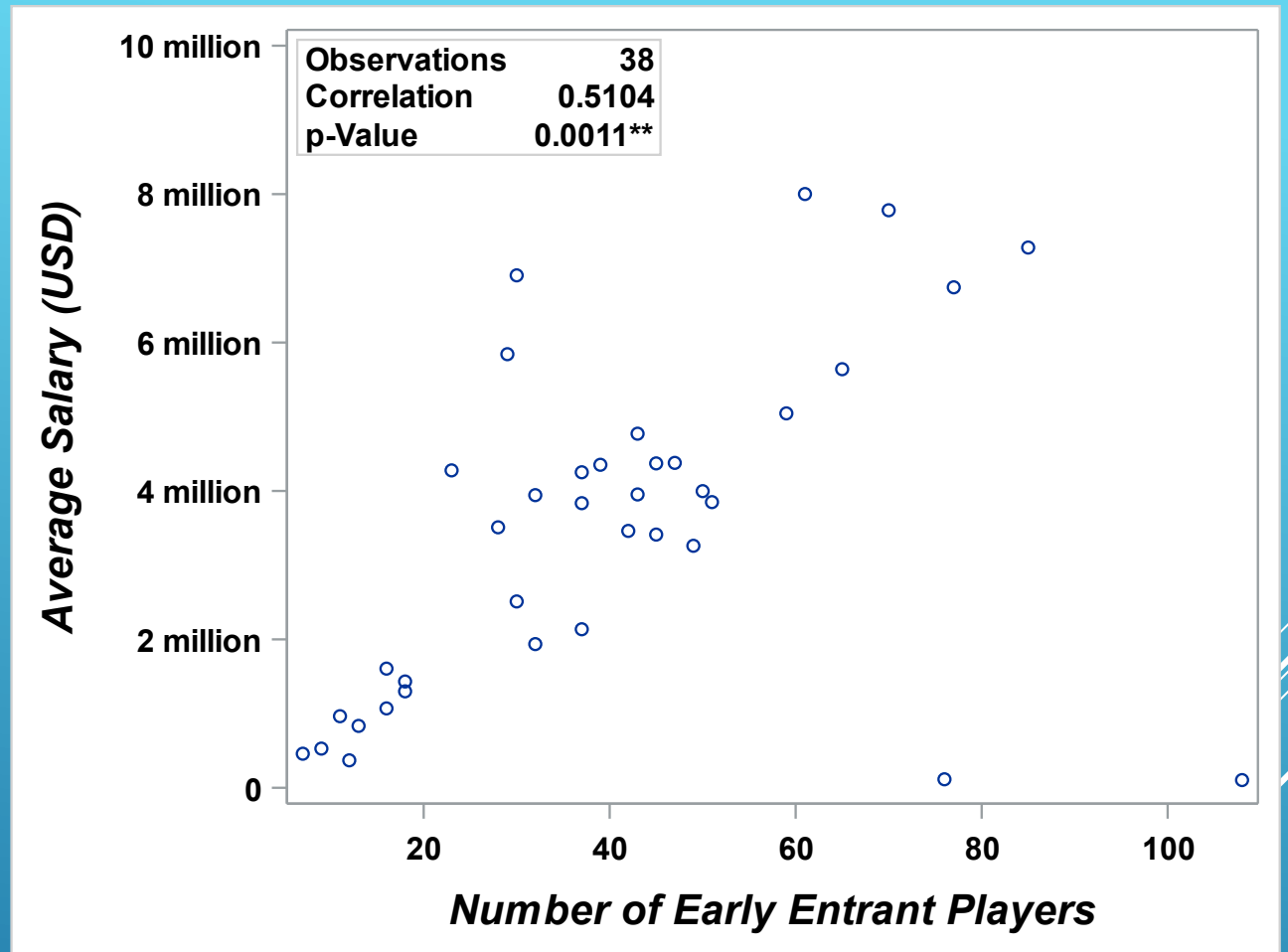
Year Error Rates for Men's & Women's Tournaments Over Time

ERROR BY GENDER OVER TIME

Correlation Between Average NBA Salaries and Number of NBA Early Entrants

INFLUENCING FACTORS' RELATIONSHIP

REFERENCES

REFERENCES

▶ ESPN Internet Ventures. (n.d.). NBA Player Salaries - 2023-2024. ESPN. https://www.espn.com/nba/salaries

▶ HoopsHype. (n.d.). *NBA Salaries.* HoopsHype. https://hoopshype.com/salaries/

▶ McDermott, G. (2023, April 10). *Which NCAA players have declared for the 2023 WNBA draft?.* DraftKings Network. https://dknetwork.draftkings.com/2023/4/10/23674042/wnba-draft-player-list-2023-early-entrants-eligibility-criteria-ncaa-renounce

▶ NCAA.com. (2023, April 4). *Latest bracket, schedule and scores for 2023 NCAA Men's Tournament.* https://www.ncaa.com/news/basketball-men/mml-official-bracket/2023-04-04/latest-bracket-schedule-and-scores-2023-ncaa-mens-tournament

REFERENCES

▶ Sporting News Editorial Team. (2021, September 18). *How to get free lunch at Little Caesars Today for UMBC's historic win.* Sporting News. https://www.sportingnews.com/us/ncaa-basketball/news/free-lunch-little-caesars-april-2-umbcs-march-madness-win/1bybkpmaoo2a91b05myvtornhy

▶ Spotrac.com. (n.d.). *WNBA rankings.* Spotrac.com. https://www.spotrac.com/wnba/rankings/

▶ Wikimedia Foundation. (2023a, October 7). *NCAA Division I men's basketball tournament.* Wikipedia. https://en.wikipedia.org/wiki/NCAA_Division_I_men%27s_basketball_tournament

▶ Wikimedia Foundation. (2023b, October 5). *2023 NBA draft.* Wikipedia. https://en.wikipedia.org/wiki/2023_NBA_draft

REFERENCES

▶ Wikimedia Foundation. (2023c, October 5). 2023 NBA draft: Early Entrants. Wikipedia. https://en.wikipedia.org/wiki/2023_NBA_draft#Early_entrants

▶ Wikimedia Foundation. (2023d, September 27). NCAA Division I women's basketball tournament. Wikipedia. https://en.wikipedia.org/wiki/NCAA_Division_I_women%27s_basketball_tournament