



MERGE CONFLICT

Measuring the Madness

Pierre Quereuil, Krisztian Meszaros, Jaden
Rodriguez, Srijan Oduru, Sivan Nemirof



Introduction to Topic and Motivation

Research Question: "How have college basketball seeds performed in March Madness relative to their expected outcomes between 1985 and 2019?"

Motivations:

- ▶ Cameron Crazies!
- ▶ Most existing published research simply tries to predict the tournament results for a given year.
- ▶ Large Scope



Data



▶ 2,205 different March Madness games from 1985-2019

▶ Added **Ranking** calculation

- Based on # of wins in tournament and the seed of their opponents

▶ Used ranking to calculate **differential**

- Differential = Seed - (Ranking / 4)

The screenshot shows a portion of a March Madness tournament bracket. It features three main sections: two 'FINAL FOUR*' games and one 'CHAMPIONSHIP' game. The 'FINAL FOUR*' games are: 5 San Diego St. (72) vs 9 FAU (71) on 4/1, and 5 Miami (FL) (59) vs 4 UConn (72) on 4/1. The 'CHAMPIONSHIP' game is 5 San Diego St. (59) vs 4 UConn (76) on 4/3. A central banner for the Huskies is also visible, with the number 20 on the left and 23 on the right, and the text 'HUSKIES NATIONAL CHAMPS'.

Game	Seed	Team	Score	Date
FINAL FOUR*	5	San Diego St.	72	4/1
FINAL FOUR*	9	FAU	71	4/1
CHAMPIONSHIP	5	San Diego St.	59	4/3
CHAMPIONSHIP	4	UConn	76	4/3
FINAL FOUR*	5	Miami (FL)	59	4/1
FINAL FOUR*	4	UConn	72	4/1

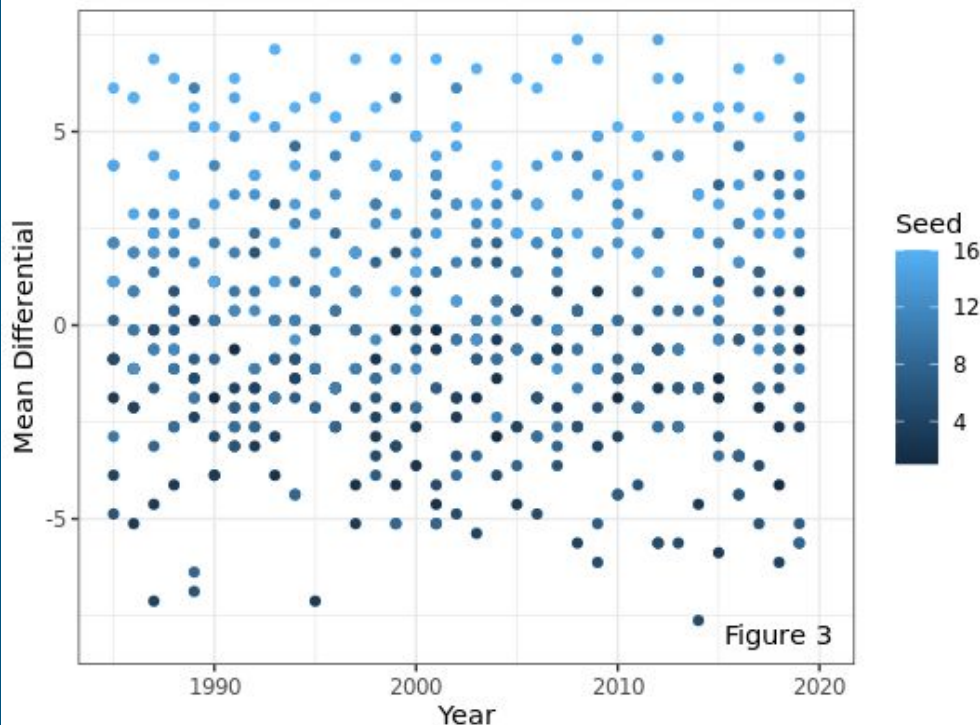
1. UConn
2. SDSU
3. Miami (lost to 4 seed)
4. FAU (lost to 5 seed)

Highlights from EDA

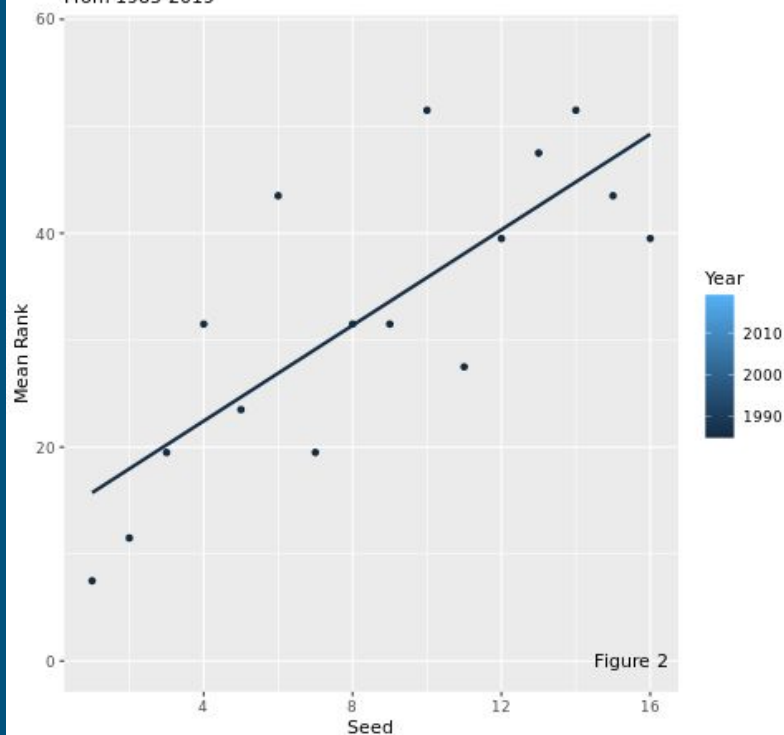


MERGE CONFLICT

Mean Differential For Each Seed Across Year



Mean Ranking by Seed Across Years From 1985-2019





MERGE CONFLICT

Linear Regression Models

$$\widehat{Ranking} = 12.23 + Seed \times 2.03$$

- ▶ Seed + 2.03 \Leftrightarrow Ranking + 1
- ▶ Ranking for 1-Seed: 14.26
- ▶ “High” Seeds Perform Well Compared to Average

$$\widehat{Differential} = -3.8 + Seed \times 0.5$$

- ▶ Seed + 0.5 \Leftrightarrow Differential + 1
- ▶ Differential for 1-Seed: -3.3
- ▶ BUT “High” Seeds Underperform Compared to Expectation



MERGE CONFLICT

Further Analysis

- ▶ Low seeds often overperform - upsets are more likely than you would think!
- ▶ Extreme variability: 5 and 3 seeds worst performers, not 1 seeds

Seed	MDifferential
16	6.0107143
15	3.9821429
14	2.9821429
12	2.2107143
13	2.1250000
11	1.4107143
10	0.5250000
9	0.2107143
7	-0.7607143
8	-1.0178571
1	-1.1035714
6	-1.5321429
2	-1.9321429
4	-2.1892857
3	-2.4178571
5	-2.5035714

