Predicting the NCAA 2019 Men's Basketball Championship

A UVA Data Science Case Study by Catherine Schuster, 2022

Prompt: It is early November of 2018 and you are a freshly starting data analytics recruit for your dream company, ESPN. The NCAA college basketball tournament is about to begin, and your manager has tasked you with creating an algorithm to model and predict the entire 2019 NCAA bracket. As an alumni from the greatest school ever, you already have a team to root for (GO HOOS!), but as a data scientist, you know you must keep an unbiased perspective in your hunt for the most plausible outcomes.

Deliverable: Produce a methodology to predict the winner of each game in the NCAA tournament. You have been given three datasets to work with for this project that contain individual game statistics and average across-season statistics from 2003-2019. You may collect additional data if you feel that you need it. Your work will be assessed by the thoughtfulness of your methodology, your choice of model, and how the features used align with the goal of predicting the outcome of basketball games. You will NOT be graded by the accuracy of your predictions.

Example Deliverable:

2019 NCAA DIVISION I MEN'S BASKETBALL CHAMPIONSHIP BRACKET

NCAA.com

| First Round MARCH 21-22 | Second Round MARCH 23-24 | Regional Semifinals MARCH 28-29 | Regional Finals MARCH 30-31 | National Semifinals APRIL 6 | | National Semifinals APRIL 6 | Regional Finals MARCH 30-31 | Regional Semifinals MARCH 28-29 | Second Round MARCH 23-24 | First Round MARCH 21-22 |

FIRST FOUR

DAYTON MARCH 19-20

Watch On truᵀⱽ

16 Prairie View (22-12) 76
16 F. Dickinson (20-13) 82 — W

11 Belmont (26-5) 81
11 Temple (23-9) 70 — E

78 N. Dakota St. (18-15) 16
74 NC Central (18-15) 16 — E

74 Arizona St. (22-10) 11
65 St. John's (21-12) 11 — W

FINAL FOUR MINNEAPOLIS

EAST WASHINGTON, D.C.

SOUTH LOUISVILLE

WEST ANAHEIM

MIDWEST KANSAS CITY

FINAL FOUR MINNEAPOLIS APRIL 6 AND 8

NATIONAL CHAMPIONSHIP APRIL 8

3 Texas Tech 77 | Virginia | 85 Virginia 1

#MarchMadness

Watch the tournament on these networks or online at NCAA.COM/MARCHMADNESS

tbs CBS TNT truᵀⱽ

March 21 and 23 first-/second-round sites: Hartford, Salt Lake City, Des Moines, Jacksonville. March 22 and 24 first-/second-round sites: Tulsa, Columbus, Columbia, San Jose. March 28 and 30 regional sites: Louisville, Anaheim. March 29 and 31 regional sites: Washington, D.C., Kansas City.

DON'T BET ON IT
The NCAA opposes all forms of sports wagering

EAST region:
1 Duke (29-5) 85
16 North Dakota St. 62 — 1 Duke 77
8 VCU (25-7) 58 — Columbia
9 UCF (23-8) 73 — 9 UCF 76
1 Duke 75
5 Mississippi St. (23-10) 76
12 Liberty (28-6) 80 — 12 Liberty 58
San Jose
4 Virginia Tech (24-8) 66 — 4 VT 67
13 Saint Louis (23-12) 52 — 4 VT 73
1 Duke 67
6 Maryland (22-10) 79 — 6 Maryland 67
11 Belmont 77 — Jacksonville
3 LSU (26-6) 79 — 3 LSU 69
14 Yale (22-7) 74 — 3 LSU 63
2 Mich. St. 51
7 Louisville (20-13) 76
10 Minnesota (21-13) 86 — 10 Minnesota 50
Des Moines
2 Michigan St. (28-6) 76 — 2 Mich. St. 70
15 Bradley (20-14) 65 — 2 Mich. St. 80
2 Mich. St. 68

1 Gonzaga (30-3) 87
16 Fairleigh Dickinson 49 — 1 Gonzaga 83
Salt Lake City
8 Syracuse (20-13) 69 — 9 Baylor 78
9 Baylor (19-13) 78 — 1 Gonzaga 72
1 Gonzaga 69
5 Marquette (24-9) 64
12 Murray St. (27-4) 83 — 12 Murray St. 62
Hartford
4 Florida St. (27-7) 76 — 4 Florida St. 90
13 Vermont (27-6) 69 — 4 Florida St. 58
3 Texas Tech 61
6 Buffalo (31-3) 91
11 Arizona St. 74 — 6 Buffalo 58
Tulsa
3 Texas Tech (26-6) 72 — 3 Texas Tech 63
14 Northern Ky. (26-8) 57 — 3 Texas Tech
3 Texas Tech 75
7 Nevada (29-4) 61
10 Florida (19-15) 70 — 10 Florida 49
Des Moines
2 Michigan (28-6) 74 — 2 Michigan 64
15 Montana (26-8) 55 — 2 Michigan 44

SOUTH / MIDWEST region:
71 Virginia (29-3) 1
63 Virginia 1 — 56 Gardner-Webb (23-11) 16
Columbia
53 Virginia 1 — 72 Mississippi (20-12) 8
51 Oklahoma 9 — 95 Oklahoma (19-13) 9
80 Virginia 1
73 Oregon 12 — 54 Wisconsin (23-10) 5
San Jose — 72 Oregon (23-12) 12
49 Oregon 12 — 64 Kansas St. (25-8) 4
54 UC Irvine 13 — 70 UC Irvine (30-5) 13
63 Virginia 1
61 Villanova 6 — 61 Villanova (25-9) 6
Hartford — 57 Saint Mary's (22-11) 11
99 Purdue 3 — 61 Purdue (23-9) 3
87 Purdue 3 — 48 Old Dominion (26-8) 14
75 Purdue 3
77 Iowa 10 — 72 Cincinnati (28-6) 7
Columbus — 79 Iowa (22-11) 10
94 Tennessee 2 — 77 Tennessee (29-5) 2
83 Tennessee 2 — 70 Colgate (24-10) 15

88 North Carolina (27-6) 1
81 N. Carolina 1 — 73 Iona (17-15) 16
Columbus
80 N. Carolina 1 — 61 Utah St. (28-6) 8
59 Washington 9 — 78 Washington (26-8) 9
77 Auburn 5
89 Auburn 5 — 78 Auburn (26-9) 5
Salt Lake City — 77 New Mexico St. (30-4) 12
97 Auburn 5 — 87 Kansas (25-9) 4
75 Kansas 4 — 53 Northeastern (23-10) 13
62 Auburn 5
59 Ohio St. 11 — 59 Iowa St. (23-11) 6
Tulsa — 62 Ohio St. (19-14) 11
58 Houston 3 — 84 Houston (31-3) 3
74 Houston 3 — 55 Georgia St. (24-9) 14
71 Kentucky 2
56 Wofford 7 — 84 Wofford (29-4) 7
Jacksonville — 68 Seton Hall (20-13) 10
62 Kentucky 2 — 79 Kentucky (27-6) 2
62 Kentucky 2 — 44 Abilene Christ. (27-6) 15