ISYE 4133 March Madness Project

Nishu Pawar, Cameron Moore, Anna Whitford, Carson Letts

4/25/2022

Part A

See: "PartA.py" and "Results-Dataframe.xlsx"

	A	B	C	D	E	F	G
1		First Round	Second Round	Third Round	Fourth Round	Fifth Round	Sixth Round
2	Duke	0.9987	0.8959412	0.722881593	0.466946458	0.27416909	0.160804207
3	N Dakota St	0.0013	9.4105E-05	3.5546E-06	4.36278E-08	3.29616E-10	2.03047E-12
4	VA Commonwealth	0.5807	0.07030431	0.028226197	0.006778107	0.00129381	0.000241221
5	UCF	0.4193	0.033660385	0.010969772	0.001991456	0.000277509	3.73602E-05
6	Mississippi St	0.8054	0.423802124	0.102529983	0.034398128	0.009584306	0.002634625
7	Liberty	0.1946	0.044130882	0.002781135	0.000286654	2.12641E-05	1.48285E-06
8	Virginia Tech	0.802	0.475315469	0.128294715	0.04711193	0.01454077	0.004437486
9	St Louis	0.198	0.056751524	0.004313051	0.000529848	4.77307E-05	4.08029E-06
10	Maryland	0.7706	0.40516962	0.115688269	0.032535056	0.009006234	0.002459431
11	Belmont	0.2294	0.060364082	0.005666307	0.000560022	5.11036E-05	4.42604E-06
12	LSU	0.9007	0.518902619	0.154747008	0.045394849	0.013129671	0.00374933
13	Yale	0.0993	0.015563678	0.000548593	2.23568E-05	7.73453E-07	2.41408E-08
14	Louisville	0.7299	0.223005559	0.132894055	0.046836617	0.016374572	0.005673506
15	Minnesota	0.2701	0.033818281	0.012268766	0.002060905	0.000332106	5.19794E-05
16	Michigan St	0.9924	0.742673546	0.578148364	0.314546661	0.171446582	0.09326079
17	Bradley	0.0076	0.000502615	3.86371E-05	9.08164E-07	1.71093E-08	2.79438E-10
18	Gonzaga	0.9988	0.879462917	0.720129677	0.485696733	0.28242469	0.166993203
19	F Dickinson	0.0012	6.7266E-05	2.75562E-06	3.59591E-08	2.36735E-10	1.38611E-12
20	Syracuse	0.5637	0.076709468	0.034099445	0.009630366	0.001967679	0.000409931
21	Baylor	0.4363	0.043760349	0.016782419	0.003872121	0.000623789	0.000102206
22	Marquette	0.7598	0.344945978	0.068346357	0.020706391	0.004594931	0.001040416
23	Murray St	0.2402	0.051767404	0.003226284	0.000357103	2.47565E-05	1.7065E-06
24	Florida St	0.8633	0.566742016	0.155562516	0.061316879	0.01861376	0.005770054
25	Vermont	0.1367	0.036544602	0.001850545	0.000168675	9.3955E-06	5.16184E-07
26	Buffalo	0.646	0.233799596	0.06326142	0.014349289	0.003047842	0.000660298
27	Arizona St	0.354	0.083555583	0.013520901	0.001801799	0.000214288	2.57637E-05
28	Texas Tech	0.9208	0.66448556	0.288499739	0.108711856	0.039636018	0.014749328
29	N Kentucky	0.0792	0.018159261	0.001210416	6.67212E-05	2.96112E-06	1.28428E-07
30	Nevada	0.5207	0.124468128	0.053338788	0.013289312	0.003121597	0.00074844
31	Florida	0.4793	0.10665537	0.043606099	0.010182365	0.002231408	0.000498926
32	Michigan	0.987	0.767645547	0.53647399	0.269847923	0.133211297	0.06697391
33	Montana	0.013	0.001230955	8.86475E-05	2.43211E-06	4.90674E-08	9.30758E-10

Figure 1: First 32 values from part A

34	Virginia	0.9963	0.884685986	0.699373474	0.493220642	0.292936673	0.158029436
35	Gardner Webb	0.0037	0.000394146	2.29313E-05	7.40744E-07	1.16632E-08	1.02936E-10
36	Mississippi	0.5153	0.061315964	0.023036082	0.006506499	0.001243804	0.000184604
37	Oklahoma	0.4847	0.053603904	0.019371512	0.005221996	0.000943611	0.000131358
38	Wisconsin	0.6876	0.480059464	0.158856611	0.077898534	0.029260577	0.009382124
39	Oregon	0.3124	0.16211506	0.029276798	0.009041199	0.001925201	0.000323286
40	Kansas St	0.7751	0.320590816	0.067727831	0.023637131	0.005840839	0.001162123
41	UC Irvine	0.2249	0.03723466	0.002334762	0.000295761	2.19169E-05	1.1059E-06
42	Villanova	0.4877	0.160873054	0.049153753	0.010242797	0.002129354	0.000347915
43	St Mary's CA	0.5123	0.174560262	0.055192763	0.011984319	0.002599766	0.000445865
44	Purdue	0.8884	0.634242912	0.328552621	0.130701794	0.052575508	0.018221675
45	Old Dominion	0.1116	0.030323772	0.004166119	0.00034353	2.68316E-05	1.43686E-06
46	Cincinnati	0.5055	0.129889236	0.047488827	0.009905583	0.002061327	0.000337126
47	Iowa	0.4945	0.124895371	0.045036205	0.009219664	0.00188234	0.000301167
48	Tennessee	0.9575	0.737251829	0.469526524	0.211738225	0.096653263	0.038668262
49	Colgate	0.0425	0.007963565	0.000883188	4.15854E-05	1.76676E-06	4.75634E-08
50	North Carolina	0.998	0.890102947	0.722949051	0.500777977	0.304930203	0.166095069
51	Iona	0.002	0.000165624	8.3632E-06	1.67306E-07	1.97984E-09	1.15584E-11
52	Utah St	0.4609	0.045408605	0.016779336	0.004030341	0.000715791	9.31658E-05
53	Washington	0.5391	0.064322824	0.02614441	0.007115376	0.001456768	0.00022372
54	Auburn	0.7672	0.497320353	0.142887491	0.061302062	0.021206348	0.005993495
55	New Mexico St	0.2328	0.088099896	0.009004088	0.001669092	0.00022137	2.05353E-05
56	Kansas	0.7879	0.372031694	0.079717232	0.027206679	0.007220825	0.001499655
57	Northeastern	0.2121	0.042548057	0.002510028	0.000283953	2.17825E-05	1.07294E-06
58	Iowa St	0.63	0.292175843	0.109425956	0.032374148	0.01006641	0.002513487
59	Ohio St	0.37	0.126176468	0.032908682	0.006455038	0.001325605	0.000204334
60	Houston	0.9257	0.571292052	0.251741319	0.089882726	0.033679672	0.010444391
61	Georgia St	0.0743	0.010355637	0.00072225	3.32886E-05	1.46969E-06	3.83413E-08
62	Wofford	0.5481	0.090556512	0.026987666	0.004260567	0.00070007	8.32302E-05
63	Seton Hall	0.4519	0.060844882	0.015615042	0.002045422	0.00027681	2.62776E-05
64	Kentucky	0.9922	0.847664381	0.562566196	0.262562674	0.128074079	0.053929222
65	Abilene Chr	0.0078	0.000934225	3.28891E-05	4.88163E-07	6.16355E-09	3.88207E-11

Figure 2: Remaining 32 values from part A

Part B

Objective

Maximize expected number of points awarded based off a team's win probability.

Decision Variables

$$x_{ik} = \begin{cases} 1 & \text{if team i wins round k,} \\ 0 & \text{o/w} \end{cases}$$

$$p_{ik} = \text{probability that team i wins round k } \forall i \in [0, 63],\ k \in [0, 5]$$

Objective Function

$$\max \sum_{k=0}^{5} (2^k \sum_{i=0}^{63} p_{ik} x_{ik})$$

Constraints

$$\sum_{k=n}^{5} x_{ik} = (6-n) x_{i,n-1} \quad \forall\ i \in [0, 63], n \in [1, 5] \qquad \text{cannot win future rounds if loss in previous round}$$

$$\sum_{a=0}^{1} x_{2p+a,0} = 1 \quad \forall p \in [0, 31] \qquad \text{winners for Round of 64}$$

$$\sum_{a=0}^{3} x_{4p+a,1} = 1 \quad \forall p \in [0, 15] \qquad \text{winners for Round of 32}$$

$$\sum_{a=0}^{7} x_{8p+a,2} = 1 \quad \forall p \in [0, 7] \qquad \text{winners for Sweet Sixteen}$$

$$\sum_{a=0}^{15} x_{16p+a,3} = 1 \quad \forall p \in [0, 3] \qquad \text{winners for Elite Eight}$$

$$\sum_{a=0}^{31} x_{32p+a,4} = 1 \quad \forall p \in [0, 1] \qquad \text{winners for Final Four}$$

$$\sum_{i=0}^{63} x_{i,5} = 1 \qquad \text{winners for Championship Game}$$

Part C

See: "PartC.py"

Objective Value: 97.337068321

Part D

We would have scored **78 points**, when comparing our optimal solution to the 2019 bracket.

Part E

$$\sum_{i,k} x_{i,k} \leq 62 \text{ with } x_{ik} \text{ from previous iteration's decision variables}$$

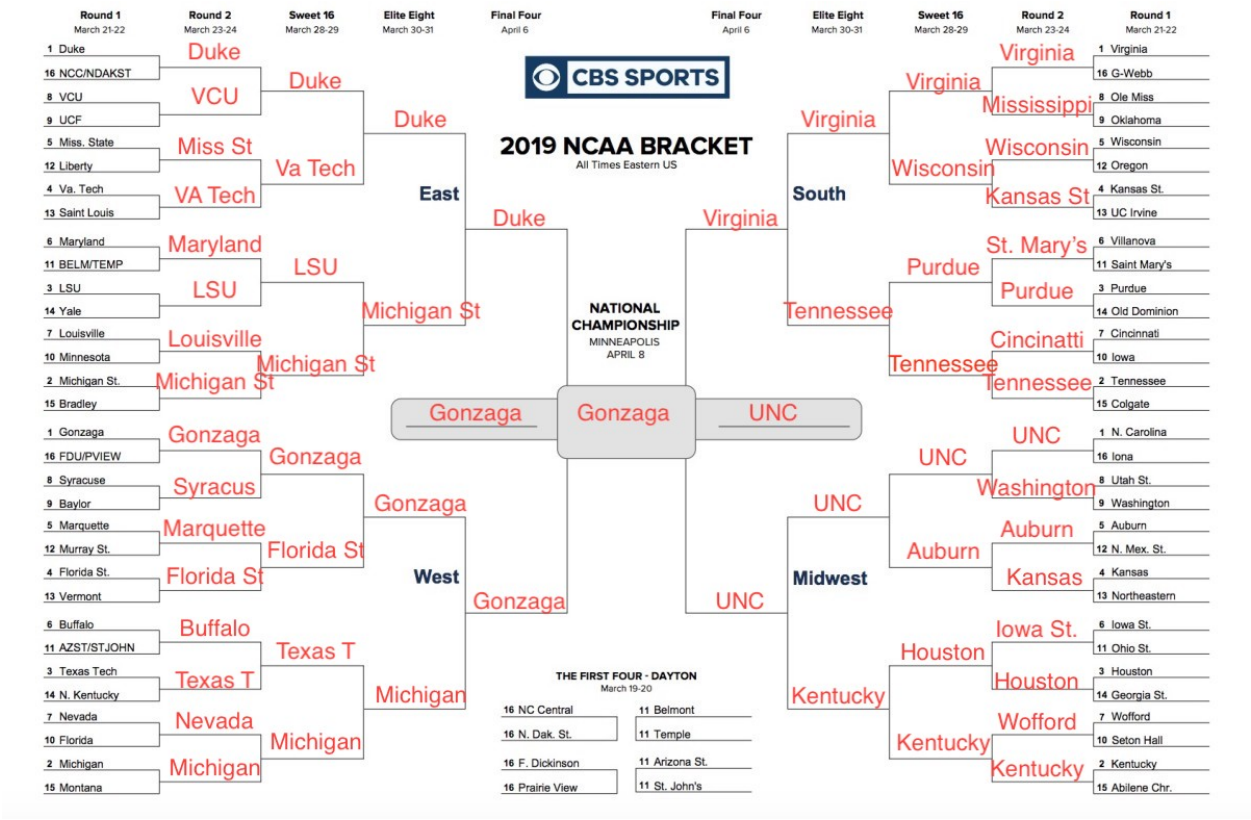
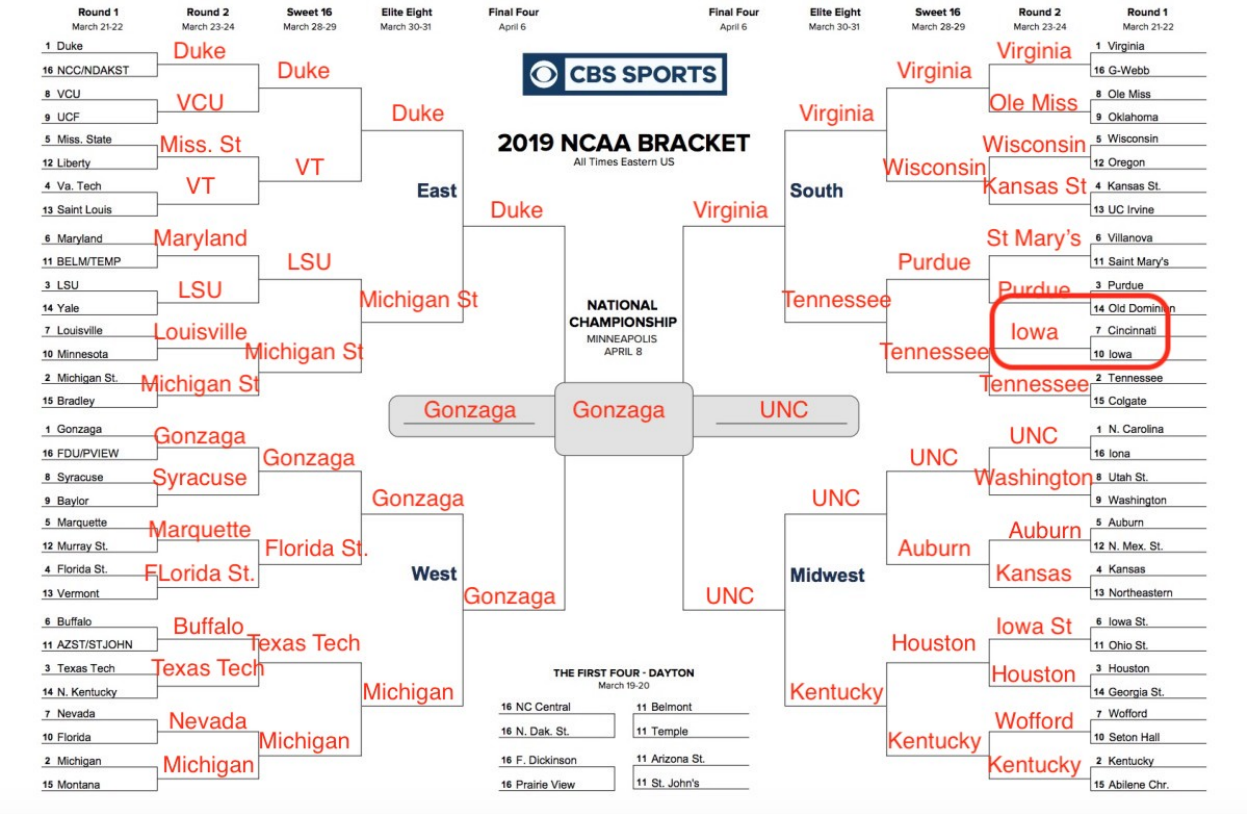


Figure 3: Optimal Solution Bracket

Part F

See "PartF.py"



In this bracket, Iowa wins in the first round, instead of Cincinatti.

The new bracket has an Objective Value of 97.326068321 and a 2019 comparison score of 79.

Part G

See "PartG.py"

We can see here that there are only minute differences in the brackets. It is usually the result of changing the outcome of a game in the Round of 64 that does not impact any of the big future games that score the most points for our objective. Even after changing our optimal solution 10 times, we only dropped .053 points off the objective value, indicating that there are a lot of games that can be changed and hardly affect the outcome.

Iterations	Actual Points	Optimal Value
1	78	97.337068321
2	79	97.326068321
3	79	97.312468321
4	78	97.308328041
5	80	97.301468321
6	79	97.306468321
7	79	97.297328041
8	79	97.295668321
9	80	97.295468321
10	80	97.284668321

Figure 4: 10 Best Optimal Solutions

Part H

See "PartH.py"

To diversify our brackets and increase the chance of having a bracket with a higher actual score, we employed two methods.

The first one focused on generating upsets. Upsets are very common in the first round, as a lower seed can sometimes shock a higher seed and win a game they weren't supposed to. We found some data describing the probability of upsets occurring, with a 50.7% chance for 9 vs. 8, 39.6% for 10 vs. 7, 37.5% for 11 vs. 6, and 35.4% for 12 vs. 5.Based off this information, we created 3 new constraints that guarantee that at least one 12 seed, 11 seed, and 10 seed upset their higher ranked opponents and also a constraint that guarantees at least two upset for the 9th seeds. This improved the average score of our bracket, with most of the brackets now reaching scores of 81 and one hitting 82. A table of scores is included in the presentation.

The second method also used historical data to create constraints. First seeds generally do well, but they aren't guaranteed to make it to the championship. For example, there has only been one instance of all Final Four team seeds being a 1 seed. To account for this, we added a constraint stating that at most 3 1-seed teams can make the Final Four. This greatly improved the average scores of our brackets to around 87 with a high of 88. This constraint had great success in predicting brackets that were closer to the 2019 bracket. A table of scores is included in the presentation.

Added Upset and 1 seed constraints:

$$x_{5,\ 0} + x_{21,\ 0} + x_{37,\ 0} + x_{53,\ 0} \geq 1$$
$$x_{9,\ 0} + x_{25,\ 0} + x_{41,\ 0} + x_{57,\ 0} \geq 1$$
$$x_{13,\ 0} + x_{29,\ 0} + x_{45,\ 0} + x_{61,\ 0} \geq 1$$
$$x_{3,\ 0} + x_{19,\ 0} + x_{35,\ 0} + x_{51,\ 0} \geq 2$$
$$x_{0,4} + x_{15,4} + x_{31,4} + x_{47,4} \leq 3$$