LET'S RECLAIM RANKINGS:
WHY ONE-SIZE-FITS-ALL RANKINGS FAIL US, AND HOW TO USE WEIGHTED SUMS TO MAKE MORE THOUGHTFUL LIFE DECISIONS
Noah Giansiracusa

Late Monday afternoon on September 18, 2023,

students, staff, and faculty at Brandeis received a distressing email from the president of the university. Earlier that day, *U.S. News & World Report* had released its annual Best Colleges rankings, and Brandeis had plummeted sixteen crucial spots, from number forty-four to number sixty. The president's email reassured readers that Brandeis had received the same "raw score" as in the previous year, so the drop wasn't the university's fault. It was due instead to *U.S. News* implementing "significant changes in methodology this year that led to many dramatic shifts in the rankings." As President Liebowitz went on to explain, the report had "removed or decreased the weighting of indicators that were favorable to private institutions like Brandeis." The email concluded with a resolute proclamation: "We maintain that no ranking can define what school is the best option for any individual student."

Another victim that day was the University of Chicago, which lost its top-ten status as it fell from number six to number twelve. Here, too, the blame was cast on methodological changes. An official response from the university declared, "We believe in and remain committed to academics and the fundamentals that have long defined the UChicago experience—such as our smaller class size and the educational level of instructors, considerations that were eliminated from this year's *U.S. News & World Report* ranking metrics." College rankings are a zero-sum game, so when someone goes down, someone else must go up. California's Fresno State jumped a whopping sixty-four spots on that fateful Monday.

U.S. News said the 2023 shakeup resulted from the largest overhaul of the system underlying its annual Best Colleges list in the publication's forty-year history. Top-100 colleges typically jostle up or down around two or three spots, but in 2023 the movement was nearly four times greater. To understand what happened, we must dig into *U.S. News*' mysterious and mercurial methodology.

The first step in *U.S. News*' process is choosing a collection of things to measure across all schools. (These are the "indicators" mentioned by Brandeis's president, Ronald Liebowitz, but I'll do what many others do and call them *factors*.) Next, *U.S. News* assigns a weight to each factor, indicating how strongly the magazine wants the factor to influence the rankings. The factor with the biggest weight, clocking in at 20 percent of the total weight both before and after the recent methodological change, is peer assessment. In *U.S. News*' words, this is "a measure of how a school is regarded by top administrators at other institutions." That's right: the biggest factor in the Best Colleges lists is not a measure of cold, hard, objective data. It's a survey of the subjective opinions of some higher-ups who aren't at the school in question. In other words, it's a popularity contest.

The fifth-highest-weighted factor, currently at 6 percent, is faculty salaries. It's hard not to see this as a measure of institutional wealth rather than academic quality. Yet *U.S. News* weights faculty salaries more heavily than factors that are probably more important to students—such as first-year retention rate (5 percent), borrower debt (5 percent), and student-faculty ratio (3 percent). Even measures we typically think of as indicating how competitive a school is or how strong its students are (such as standardized tests, worth 5 percent), are given lower weight than faculty salaries.

U.S. News' massive methodological overhaul in 2023 was simply a reweighting of the factors in its recipe. Alumni giving, which ostensibly shows how favorably alums feel about their alma mater but really measures how wealthy the school's alums are, was reduced to zero. The weight on graduation rates for students who received need-based Pell Grants was increased. These are both good changes, in my opinion, as they move the needle away from wealth and toward helping less affluent students. But by keeping peer assessment at the top, and faculty salaries near the top, *U.S. News* ensured that the rich and famous schools largely kept their clout.

The change that seems to have had the biggest impact is that the weight given to small class size went from 8 percent to 0. Wealthy schools can afford small class sizes, so this does help level the playing field a bit. But small classes aren't *just* a measure of a university's financial resources; many students really do prefer them. At only seventy-five years old, Brandeis is much less wealthy than many peer institutions that have had hundreds of years to grow their endowments. Harvard's endowment is \$53 billion; Columbia's is \$13 billion; even Dartmouth's \$8 billion dwarfs Brandeis's comparatively meager \$1 billion. Brandeis has managed to maintain small class sizes despite its relative "poverty," but this achievement is no longer reflected in the *U.S. News* rankings.

U.S. News described its choice to cut class size from the factors as a strategic decision to "place greater focus on outcomes measures." I have a hard time understanding how faculty salaries are more of an "outcomes measure" than class size. As we'll see, the decision to cut class size may have been strategic in a different sense: damage control for a massive scandal that rocked *U.S. News'* college rankings the previous year. I'll let you be the judge.

Seeing the arbitrariness in choosing which factors to include and how much to weight them—and seeing the "dramatic shifts," as President Liebowitz called them, that result when these arbitrary choices are varied—might make you more sympathetic to his remark that "no ranking can define what school is the best option for any individual student." If so, good: you're making progress toward rethinking rankings. But I do believe that rankings can help students choose colleges. We just need to personalize the rankings by letting each person choose factors and weights that reflect their values.

To do this, we must look under the numerical hood of rankings to see in more detail how they work—and how they can be refashioned to better suit our needs. The benefits reach far beyond higher education. Rankings, and the weighted sums they're built from, appear in settings ranging from the stock market and inflation to consumer credit scores. We'll learn the math to make sense of these and to personalize many of them, helping you compare options and make thoughtful life decisions.

We must look under the numerical hood of rankings to see in more detail how they work—and how they can be refashioned to better suit our needs.

THE MATH OF RANKINGS

The first challenge with building a ranking system is that the different factors are typically measured on such wildly different scales that it doesn't work well to compare them directly. The difference between a huge class of 250 and a tiny class of ten looks insignificant compared with tuitions that might range from $15,000 to $80,000. To address this, *U.S. News* uses a popular statistical technique called *standardizing* the data. This converts all the numbers to a measure of how far above or below average they are.

In the US, the average male height is five feet, ten inches. Heights, like many things in life, are distributed in a familiar bell-shaped curve. The average indicates where this curve is centered, while the standard deviation indicates how wide it is. Consequently, the number of standard deviations above or below the average is a useful measure of how rare something is. For American male heights, the standard deviation is about three inches. So someone who is six foot four is two standard deviations above average, while a six foot seven person is three standard deviations above average.

To standardize a factor, we first compute the average of the factor across all the entities being ranked, then we record how many standard deviations above or below average each entity is. This is called a Z-score. For example, the average graduation rate across all colleges in the US is around 55 percent with a standard deviation of twenty-three percentage points, so a school with an 89.5 percent graduation rate gets a Z-score of 1.5 while a school with a 32 percent graduation rate gets a Z-score of −1. Converting all factors to Z-scores puts them on the same scale. Instead of comparing dollars, days, donations, and whatever else, we're just comparing standard deviations. A score of one for tuition means the same thing as a score of one for class size: one standard deviation above average.

The next step is to combine the Z-scores across all the different factors to produce a single overall score for each entity. The most straightforward approach is to simply add up the Z-scores, but this puts all the factors on equal footing. It's more common to use a *weighted sum*, which allows some factors to influence the rankings more than others. With this approach, a weight is chosen for each factor, then when adding up the Z-scores, each is first multiplied by the factor's weight. For example, if a school has Z-scores of +2 and –1 and the first factor is weighted at 70 percent while the second factor is weighted at 30 percent, then the school's overall score is 0.7 × 2 + 0.3 × (–1) = 1.1. These weighted sums of Z-scores boil each entity's data down to a single number so that they can be ranked from highest to lowest. This is the math behind most ranking systems, and it is how *U.S. News* decides which are the so-called best colleges. *Times Higher Education* in the UK has a similar university ranking system based on a weighted sum of Z-scores; it uses thirteen factors organized into five groups.

When Brandeis's president said the school fell sixteen spots despite reporting the same raw score as the previous year, he probably meant that the school's unweighted sum of Z-scores hadn't changed much—Brandeis was above and below average on the same measures in roughly the same amounts. But by changing the weights, *U.S. News* caused the rankings to shift considerably.

Thankfully, math enables us to break free from *U.S. News*' one-size-fits-all ranking system. We can create personalized rankings that include whatever factors we want and weight them however we wish. Since not everyone is comfortable using spreadsheets and computing standard deviations, I'll show you an easy way to make flexible rankings by hand. You'll even be able to include quirky non-numerical factors like how much you like a school's cafeteria food and how willing you are to support the football team. And you can use this math to rank anything, not just colleges.

PERSONALIZED RANKINGS

Creating your own ranking system is surprisingly simple. First, pick the things you want to compare. Next, choose the factors you'd like to include. Pretty much anything is allowed as a factor. You don't need actual data, and you certainly don't need Z-scores. For each factor, you just need to decide which of the things you're ranking is the best, which is the second best, and so on. Next, choose how much weight you'd like each factor to count for; it's common to use percentages that add up to 100 percent, but any numbers will work. There are no rules or magic recipes for this step; just as *U.S. News* chooses the weights however it wants, so can you. The final step, and the only one involving any math, is to compute the weighted sum of your rankings across all your factors.

Let's try this with college rankings. It's fine to start with a small list of schools you came up with the old-fashioned way, using things like word of mouth and intuition. Suppose you're trying to decide between Tufts, UCLA, and Georgetown. For the factors, let's go with how much you like the nearest city, how much you like the weather, student-faculty ratio, and how good the basketball program is. The first two factors are entirely subjective, but that's OK: if they matter to you, it's fine to include them in your personalized ranking.

There are no rules or magic recipes for this … just as *U.S. News* chooses the weights however it wants, so can you.

All three schools are near large, vibrant cities, but let's say you like DC the best since you're interested in politics and international affairs. Boston is second. Los Angeles is third: it's OK, but not too convenient for a college student without a car. For weather, Los Angeles is the clear winner and Boston the loser (sorry, Boston fans, but you know it's true). The US Department of Education's College Navigator website has lots of useful data that helps with college comparisons. From that site we find that Tufts has the best student-faculty ratio at 10:1, Georgetown is a close second at 11:1, and UCLA comes in third at 18:1. Since Tufts and Georgetown are so close in this regard, I'm going to call it a tie for first between them and place UCLA in third. You're in charge of your personalized ranking, so you get to make executive decisions like this—if you feel that ranking them 1, 1, 3 more closely matches the data than ranking them 1, 2, 3, then go ahead and do so. Finally, UCLA and Georgetown both have storied basketball teams, but I'll give the win here to UCLA since it has eleven national championships to Georgetown's lone win in 1984. Tufts does well in its division, but compared with these two basketball powerhouses, we've got to place it third. Let's summarize all these rankings in a table:

	City	Weather	Student ratio	Basketball
Tufts	2	3	1	3
UCLA	3	1	3	1
Georgetown	1	2	1	2

Now that we have the schools ranked for each factor separately, we can choose weights to combine these rankings into a single overall ranking. First, let's try using equal weights across all four factors, meaning 25 percent each:

Tufts: $0.25 \times 2 + 0.25 \times 3 + 0.25 \times 1 + 0.25 \times 3 = 2.25$

UCLA: $0.25 \times 3 + 0.25 \times 1 + 0.25 \times 3 + 0.25 \times 1 = 2$

Georgetown: $0.25 \times 1 + 0.25 \times 2 + 0.25 \times 1 + 0.25 \times 2 = 1.5$

Since lower scores are better, like in golf, we have Georgetown on top of our personalized ranking, followed by UCLA and then Tufts. Now let's try adjusting the weights by giving more weight to weather (35 percent) and basketball (35 percent) and less weight to city (15 percent) and student ratio (15 percent):

Tufts: $0.15 \times 2 + 0.35 \times 3 + 0.15 \times 1 + 0.35 \times 3 = 2.55$

UCLA: $0.15 \times 3 + 0.35 \times 1 + 0.15 \times 3 + 0.35 \times 1 = 1.6$

Georgetown: $0.15 \times 1 + 0.35 \times 2 + 0.15 \times 1 + 0.35 \times 2 = 1.7$

This change is enough to propel UCLA to the top, since it put more weight on UCLA's strengths and less weight on its weaknesses. Tufts came out in third place both times, and it did so when I tried a few other sets of weights, too. My take is that based on these factors, if weather and basketball are really important to you, then UCLA is the place for you. If you're not sure how important these are to you compared with the other factors in consideration, then it's safe to place Tufts in third, but I wouldn't use these rankings to choose between UCLA and Georgetown. There's too much dependence on the choice of weights to reliably do so.

You might feel awkward choosing factors and weights whimsically like this, but it's no less scientific than what *U.S. News* has been doing for forty years. Seeing how the sausage is made, and making some sausage of your own, might make you less willing to rely on rankings at all, and that's perfectly OK. My advice is to view rankings as one data point among many when making decisions like which college to attend. And when you do make personalized rankings, try playing around with the weights to see how robust your winners are. If something tops your rankings for a range of weights, that should give you more confidence that it really is your top choice.

Feel free to experiment and have some fun with personalized rankings. Try ranking movies you've seen or restaurants you've eaten at—not just according to reviews by critics and the public (although those could be included as factors) but according to things that matter to you personally. If you're deciding where to move next, try ranking cities. Don't be afraid to get creative with the factors you use: it's OK to consider things like weather, music scene, restaurant scene, public transit, affordability, and job prospects. It's usually not too hard to rank cities for each of these factors separately, and weighted sums allow you to combine a bunch of separate rankings into a single holistic ranking.

Upending *U.S. News*' hegemony would restore some of the distinctness and individuality of yore. Personalized rankings are a key to doing this.

One-size-fits-all rankings usually aren't as useful as they would be if their priorities, reflected in the choice of factors and weights, more closely aligned with yours. But their problems run much deeper than this. In her bestselling book *Weapons of Math Destruction*, Cathy O'Neil argues that when different institutions focus their efforts on the same set of metrics, they come out looking kind of the same.

Take higher education, for example. Instead of each school defining its own strengths and drawing students who appreciate them, universities know what factors they need to score well on, and they all do essentially the same things to boost their scores on these factors. Upending *U.S. News*' hegemony would restore some of the distinctness and individuality of yore. Personalized rankings are a key to doing this. And as you've just seen, the math needed to make them isn't hard at all.

Another problem O'Neil discusses is that the more influential a ranking is, the greater is the incentive to game it. Baylor paid the exam fee for students to retake the SAT after they were admitted, hoping to boost its Z-score on that factor. Bucknell and Claremont McKenna found a cheaper approach: they sent false data to *U.S. News* to inflate the SAT scores of their incoming students. Iona College (now Iona University) went further and was caught fudging nearly all the data it submitted. Some of the factors and weights *U.S. News* uses have changed since these scandals, but what hasn't changed is the pressure colleges feel to do whatever it takes to come out on top. And with the math of rankings in hand, we can now see how they are gamed.

GAMING THE COLLEGE RANKINGS

When the *U.S. News* college rankings were launched forty years ago, Columbia was ranked eighteenth. Over the years it steadily climbed, reaching a pinnacle in the 2022 ranking, when it was listed as second in the nation—tied with Harvard and MIT and topped only by Princeton. The next year, Columbia dropped sixteen spots all the way back down to number eighteen. This was a single-year fall from grace as large as Brandeis's in the subsequent year, but for a school so close to the top to plummet so abruptly was even more staggering. And in the case of Columbia, it wasn't caused by reweightings within *U.S. News*. It was triggered by one of the university's own faculty members.

Michael Thaddeus is a highly regarded math professor at Columbia. I've known him for years because he and I happen to be in the same research community. I even used a result from his PhD thesis in my own dissertation—a result that, ironically, is all about adjusting weights, but not the kind *U.S. News* adjusted. As accomplished as he is in theoretical mathematics, Thaddeus doesn't normally do the kind of work that lands you in the newspaper, so I was surprised when I found his familiar face (albeit sporting an unfamiliar pandemic hairdo) staring at me from the pages of *The New York Times* on the morning of March 17, 2022. It turns out that in his spare time, Thaddeus had done some numerical detective work and uncovered a massive scandal. To boost its national ranking, Columbia had allegedly submitted data to *U.S. News* that was, in his words, "inaccurate, dubious, or highly misleading."

Thaddeus detailed these allegations in a methodical twenty-one-page analysis he posted on his webpage. He framed his report as a critique not just of Columbia but of the college ranking industry more broadly. Word quickly spread.

After looking into the matter, *U.S. News* changed Columbia's number-two ranking in 2022 to "unranked." The following year, with more scrutiny placed on Columbia's data, the university received its number-eighteen ranking.

I'll highlight just a few of Thaddeus's findings here. First up, instructional spending. Columbia reported $3.1 billion, the highest among the six thousand institutions of higher learning reporting to the government. That's more than Harvard, Yale, and Princeton combined. This staggering figure helped Columbia to score very well on the financial resources per student factor, then weighted heavily at 10 percent. But Thaddeus found Columbia's $3.1 billion to be the result of creative bookkeeping, in which dollars spent at the university hospital were counted as instructional spending on the shaky grounds that medical students are educated there.

The more influential a ranking is, the greater is the incentive to game it.

Next, the percentage of full-time faculty with a terminal degree, meaning the highest degree in the profession. Columbia reported an impressive 100 percent here. But Thaddeus argued that this literally could not be possible, pointing to sixty-six faculty members who appear to counter Columbia's claim. Among these are distinguished scholars, artists, and even a Nobel Prize winner. These eminent faculty members surely add tremendous value to Columbia, but they do not have terminal degrees.

And now, class size. Columbia said 82.5 percent of its undergraduate classes had fewer than twenty students—an astounding figure that surpassed all other schools in the top one hundred. Thaddeus ran the numbers and found that the actual value Columbia should have reported was somewhere between 62.7 percent and 66.9 percent. It was hard for him to pin down the exact number. He had to painstakingly go through online course directories, some of which were available only through the Internet Archive, a service that saves older versions of websites. If it was this difficult for someone at Columbia to verify the reported class size figure, what chance did officials at *U.S. News* have of doing so? And if Columbia was fudging this figure, how many other schools were, too?

One year after this 2022 ranking that Thaddeus lambasted, *U.S. News* published the 2023 ranking that opened this article. That was the ranking with "dramatic shifts," in President Liebowitz's words, stemming from the biggest reweighting of factors in the forty-year history of Best Colleges. Five factors were eliminated entirely that year. One of these, as we discussed earlier, was class size, which dropped from 8 percent to 0. Another was percentage of full-time faculty with a terminal degree, which dropped from 3 percent to 0. That's right: two of the five factors that *U.S. News* dropped in the aftermath of the Columbia scandal were factors that featured prominently in that scandal.

Correlation isn't causation, as they say. But all this makes me wonder how much *U.S. News'* recent methodological overhaul was really about "outcomes measures," as the magazine claims. From where I sit, dropping these factors sure looks like a convenient albeit hamfisted way to address the manner in which these factors were apparently gamed.

NOT JUST RANKINGS

Weighted sums are a popular ingredient in rankings, because they turn a collection of separate scores across a range of factors into a single aggregate score. And they do so in a flexible way that allows the factors to have varying amounts of influence. Weighted sums are used for this purpose in many settings, not just those involving rankings. The lessons you've learned so far—how rankings are made, how they can be used and abused, and how you can personalize them—directly carry over to these other applications.

Have you ever wondered what it means when they say on the news that inflation is 3 percent (or 5 percent, or …)? Surely this means that prices are going up 3 percent per year—but which prices? Housing prices, gas prices, food prices, consumer electronics prices: these are all important, but they often move at different rates and sometimes even in different directions. To get an overall sense of how much prices are changing, we'd need a measure that combines all these price factors. The weighted sum heeds the call of duty.

The Consumer Price Index (CPI), for example, is a weighted sum of prices across eighty thousand goods and services in the US. The weights are not chosen arbitrarily; they are meant to capture how important each item is to a typical American household. Americans currently spend about ten times more on meat products than on tofu products, so meat is weighted ten times more heavily than tofu in the CPI. The collection of goods and services in this weighted sum is fixed from year to year (except for infrequent periodic updates), so the percent the CPI goes up annually can be viewed as the inflation rate in America.

But what if you're an American who doesn't eat ten times more meat than tofu? Just as you can personalize college rankings, you could create a personalized consumer price index if you wished. No need to use eighty thousand items—just think of a handful of things you tend to spend money on each year, then compute a weighted sum of their inflation rates with weights that reflect what fraction of your budget each item occupies. If you get a cost of living adjustment to your salary each year, it's nice to know how it stacks up against the rise in prices for the things you buy, and the amounts of them you buy, rather than naively pretending that rising prices impact us all equally. Your neighbors don't spend money on the same things as you, so their inflation rates are not the same as yours.

Even when personalization isn't an option, it can be interesting just to recognize that a weighted sum is behind an important number. And it can be helpful to see the weights involved.

Now that you've spent some time with weighted sums, you'll be better equipped to navigate a world that relies on them to rank and score a wide variety of things.

Did you know that credit scores are weighted sums? FICO is the most popular credit score in the US, and while the full details are a closely guarded corporate secret, FICO revealed that there are five main factors that are combined via a weighted sum. Payment history (including bankruptcies and late payments) carries 35 percent of the weight; amounts owed (including number of accounts and balances) has 30 percent; length of credit history (including the average age of accounts and the oldest account) has 15 percent; credit mix (having different types of credit, such as mortgage and consumer finance) has 10 percent; and new credit (opening new lines of credit and credit inquiries resulting from applications for new credit) has 10 percent. FICO won't tell us many details beyond this, but already this weighted-sum description helps dispel some of the mystery—and suggests what to focus on to improve your credit score. For instance, it's more important to avoid late payments than credit inquiries. Outside the US, lending agencies tend to use their own formulas rather than relying on a centralized one like FICO. These formulas are almost always weighted sums; some lenders reveal the weights and factors while others do not.

If you follow financial news, you've likely come across various stock market indices around the world. Most of these are in essence weighted sums. The S&P 500 and Nasdaq in the US, the FTSE in the UK, the DAX in Germany, the CAC in France, the KOSPI in South Korea, the Hang Seng in Hong Kong, the Ibovespa in Brazil, and the NIFTY in India, to name a few, are known as capitalization-weighted indices. Here's what that means. The market capitalization of a company is the stock price of the company multiplied by the number of shares traded on the market. Capitalization-weighted indices are fractions where the numerator is the total market capitalization of the companies in the index and the denominator is a bookkeeping device meant to keep the index on a reasonable scale and to ensure that its value doesn't fluctuate too much when companies are added or removed.

The numerator is what's most important, and one way to think about it is as the sum of the stock prices of the companies in the index weighted by the number of shares available for each. (The Dow Jones in the US and the Nikkei in Japan sum up stock prices without weighting them by the number of available shares. Some experts have argued that these indices are less meaningful because a big swing in a company's stock price has a big impact on the index, even if very few shares of the company are traded on the market.)

Now that you've spent some time with weighted sums, you'll be better equipped to navigate a world that relies on them to rank and score a wide variety of things. **Most important, instead of trusting the big one-size-fits-all rankings that everyone tries to game, I hope you'll try making some personalized rankings. And when you can't avoid being scored by a weighted sum, I encourage you to look into the factors and weights behind the score so that you can play this numbers game as strategically and successfully as possible.** ⌂

Info
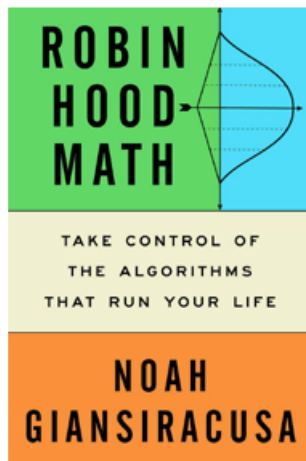
Ready to dig deeper into the book?
Buy a copy of Robin Hood Math.

Want copies for your organization or for an event?
We can help: customerservice@porchlightbooks.com
800-236-7323

ABOUT THE AUTHOR

Noah Giansiracusa is an associate professor of mathematics at Bentley University, a visiting scholar at Harvard University, and the author of *How Algorithms Create and Prevent Fake News*. His writing has appeared in *Scientific American*, *Time*, *Wired*, *Slate*, and *The Washington Post*, among others, and he has been featured as a guest on CNN, BBC Radio 4, and Newsmax. Giansiracusa lives in Acton, Massachusetts, with his wife, two kids, two dogs, and twelve chickens.

SHARE THIS
Pass along a copy of this manifesto to others.

SUBSCRIBE
Sign up for e-news to learn when our latest manifestos are available.

Porchlight

Curated and edited by the people of Porchlight, ChangeThis is a vehicle for big ideas to spread. Keep up with the latest book releases and ideas at porchlightbooks.com.

This document was created on August 6, 2025 and is based on the best information available at that time.