Re-Ranking National Universities: A Simple Exercise in Descriptive Analytics

J.L. Nazareth

August, 2015

Each year, the magazine *U.S. News* publishes an eagerly-anticipated and not uncontroversial issue that compares and contrasts the more than 1400 four-year colleges and universities of the United States. These educational institutions are subdivided into major categories---national universities, national liberal arts colleges, regional universities, and so on---and, within each category, the institutions are ranked numerically. The centerpiece of the yearly study is the ranking of the *national* (also known as *major research*) *universities*, which will be the focus of our discussion. Our conclusions are also relevant for the other categories of universities and colleges not considered here.

A Typical Recent Ranking

Consider specifically a recent *U.S. News* study [1], published in 2011, that encompassed 264 national universities, 164 being public institutions and the remainder private. Data was gathered from each such university on several key *numerically-quantifiable measures* of academic quality: undergraduate academic reputation, student retention, faculty resources, student selectivity, financial resources, alumni giving rate, and graduation rate performance.

Next, a weight was assigned---albeit subjectively---to each quality indicator, i.e., a decimal number between 0 and 1, where the weights over all indicators necessarily sum to unity. The 'undergraduate academic reputation' indicator received the largest weight of 0.225. For the remaining indicators, the weights were as follows: 0.2 for `student retention', 0.2 for `faculty resources', 0.15 for `student selectivity', 0.1 for `financial resources', 0.05 for `alumni giving', and 0.075 for `graduation rate performance'.

Finally, a total weighted score over all indicators was computed for each university and the universities were ranked and listed from highest scoring to lowest. (When weighted scores were identical then the corresponding equally-ranked universities were listed in alphabetical order.) Presumably as a face-saving measure for the lower-scoring

universities, or perhaps the result of a loss of nerve, only the top 75 percent of the universities were listed by rank in the detailed study published by the magazine, each rank being accompanied by the corresponding weighted score and details of its associated, underlying quality indicators. The remaining 25 percent of the national universities, which had lower scores than those explicitly ranked, were assigned to a so-called *second tier*, where the list was continued simply in alphabetical order, i.e., no associated numerical rank or total score was reported for universities in this lower tier. Thus at the very bottom of this full list of ranked and unranked were the unfortunate, low-scoring national universities whose names began with letters close to the end of the alphabet. At the other extreme, the top university in the numerically-ranked list would not be hard to guess---it was Harvard!

Typical Ranking from a Decade Earlier

Let us step back from the results reported in recent years to the antecedent studies of a decade earlier. Thus, consider specifically the fall issue of the magazine [2], published in 2000, which ranked a slightly smaller number of 228 national universities, 147 of which were public and the remainder private.

The overall methodology underlying the ranking was much the same as the one just described for the more recent 2011 study, save in one important respect, namely, the indicator `undergraduate academic reputation'. In the 2000 study, this indicator was identified more simply as `academic reputation' (peer assessment) and measured from a survey of top academics at universities using a 5-point scale. The averaged 5-point score (over all respondents to the survey) for this `academic reputation' indicator was reported directly for each of the universities, whereas in 2011 the corresponding indicator was `*undergraduate* academic reputation' and the associated reported numerical measure for a university was a less accessible, rescaled quantity between 1 and 100 derived, in turn, from the averaged 5-point scores from surveys of two categories of assessors---academics (as in the 2000 study) and high-school counselors---these categories then being suitably (sub-)weighted to obtain a composite number for the indicator. To distinguish between the two versions of this particular quality indicator, we will henceforth identify it as `*peer assessment'* instead of `academic reputation' when referring to the 2000 study whilst retaining the term `undergraduate academic reputation' for the 2011 study.

The weighting factor associated with the `peer assessment' indicator in the 2000 study that was used to compute the overall score of a university was 0.25 (instead of 0.225 for the 2011 indicator `undergraduate academic reputation'). The other quality indicators and their weighting factors remained unchanged with the exception of the

`graduation rate performance' indicator, which picked up the miniscule change in weight of 0.025 between the `undergraduate academic reputation and `peer assessment' indicators, i.e., in 2011, `graduation rate performance' received a weight of 0.075, whereas in 2000 its weight was 0.05.

Ranking methodology aside, a fundamental difference between the 2000 and 2011 results was the way that the final rankings were reported. In 2000, the top 50 universities comprised a *first tier* and only these were ranked. The remaining national universities were grouped into *three* additional tiers, each containing roughly the same number of universities: the second tier comprised 64 universities with the next highest weighted scores; the third tier comprised 56 universities and the last 58 universities formed the fourth tier. In these latter three tiers, the universities were listed *alphabetically* and their weighted score totals were *not* reported, although their associated peer assessment indicator numbers were provided. Additional detail on some of the other quality indicators was also reported with each lower-tier university entry, but not as extensively as the reporting for the first tier.

Drawbacks of the Studies

The description in the foregoing two sections glosses over many details of the ranking studies but gives their essence. Observe, in particular, that over the course of the decade separating the two cited studies, the magazine's ranking scheme has increasingly and explicitly focused on *undergraduate* education, but nevertheless, reducing the merit of a university to a single numerical rank, even in this limited regard, is suspect. Furthermore, the magazine's rankings are frequently interpreted by the media and the public-at-large as a measure of the overall quality, or worthiness, of the national universities themselves, and, from this latter viewpoint, many of the 2011 relative rankings [1] were bizarre. For instance, the University of Miami in Florida and Yeshiva University in New York were among the top 50, receiving numerical ranks of 47 and 50, respectively, whereas much more highly-regarded universities such as Purdue University in Indiana, the University of Minnesota at Twin-Cities, and the University of Colorado at Boulder received significantly lower ranks of 56, 64, and 86, respectively. *The top 21 universities in the ranked list were all private institutions.* One of the most highly regarded universities in the entire world, the University of California at Berkeley (UCB), which consistently ranks near the top in world-wide ranking studies, receives the lowly rank of 22 in the *U.S. News* study. (An undergraduate must surely benefit *immeasurably* from the incredibly rich *intellectual* environment coupled with the mind-broadening *social diversity* that is provided by our leading public universities, for example, UCB or the University of Michigan, regardless of the drawbacks of greater class crowding and lower faculty-to-student ratios.)

The relative rankings in the cited study [2] of 2000 were equally bizarre. There were universities in the first tier that had a relatively low academic reputation (peer assessment) number, while some of the universities in the second tier had much higher peer assessment numbers than those in the first. There were even instances of universities in the third tier with a higher peer assessment than a couple of universities in the first tier, although there were no instances of universities in the fourth tier having a higher peer assessment than any university in the first. Likewise, there were a couple of universities in the fourth tier that had a better peer assessment number than one or two in the second tier, several universities in the third tier that exceed a few in the second, and several in the fourth tier that exceed a few in the third.

Re-Ranking the National Universities

Whilst not throwing the baby out with the bathwater, how does one restore a higher level of common sense to the *U.S. News* ranking and thereby obtain a better impression of the relative stature of our major research universities than that furnished by single, composite-ranking number?

The way that the results were reported in the 2000 study provides a possible answer, one that has a qualitative and a quantitative dimension:

1. *Qualitative:* We seek to retain the tiers 1, 2, 3 and 4, hierarchy as presented in the 2000 U.S. News study, but simultaneously to blur the distinctions between them. Thus we will employ an alternative terminology, one that is derived from the social classification made in certain European countries, for example, Great Britain, and henceforth refer to the four tiers as *upper tier, upper-middle tier, lower-middle tier, and lower tier*, respectively. This brings with it a change in perspective on the hierarchical ranking, because members of human society that are outwardly labeled as middle or lower tier are not always judged as being less meritorious than some members of the upper tier---just think of the fictional, upper-tier character Bertie Wooster vis-a-vis his lower-tier butler. The same connotation and qualitative blurring of hierarchical distinctions would thus carry over to the national universities.

2. *Quantitative:* The separation into four tiers was done using the overall, weighted scores for the universities, as described in previous sections, and this feature is retained. But now let us drop use of the weighted scores for ranking *within each tier*, and thus instead within each tier let us *reorder, or re-rank,* the U.S. News list using the *peer assessment* quality indicator number reported for each university. When there are ties within the upper tier, i.e. when two universities have the same peer assessment indicator number, then further reorder

according to the overall weighted score; and if some ties still remain unresolved then reorder these ties alphabetically. In the case of the three remaining tiers---upper-middle, lower-middle, and lower---weighted scores were not provided by the magazine so that ties in our reordered lists can only be resolved alphabetically. But this could easily have been remedied by a reporting of the weighted scores of all universities.

We now have a *completely uniform scheme* from top to bottom. Within the universities in the upper tier, we find a better mix of public and private amongst the top 21 in the ranking, with public universities like UCB and Michigan taking their rightful place near the very top. *No longer would it be necessary to have a separate, ancillary ranking that is restricted to the public universities, again based purely on weighted scores, as is provided in both the 2000 and 2011 studies.* Furthermore, the reordered lists for the four tiers will immediately reveal which universities within any given tier have a *greater* peer assessment number than a subset of the universities in the immediately adjacent tier above it in the hierarchy. In other words, within each tier, the list of universities will now fall into three groups: universities that would naturally be assigned to that tier based on either peer assessment or weighted score; universities that might be viewed as belonging to the adjacent higher tier; and universities that might be viewed as belonging to the adjacent lower tier. It would be a simple matter to highlight, say by color coding, these different subsets of universities within each of the tiers.

When performing this simple re-ranking exercise, for which an app could easily be written whenever the database is digitally accessible, I don't think there would be a single university in the entire re-ranking that violates one's sense of fair play or common sense. Furthermore, the re-ranking opens up an evolutionary path from one year to the next, because the surveyed experts who view the ranking results might begin to reconsider their peer assessment numbers in the following year's survey. In other words, an iterative process would be set in motion, resulting in improved survey results from year to year.

We note also that the most recent 2015 *U.S. News* study uses an essentially identical methodology to that of the cited 2011 study [1] and differs only as follows: a small alteration is made in the weighting factors for two of the quality indicators (`graduation and retention rates' receives weight .225 and `student selectivity' receives weight .125); a change in nomenclature ('undergraduate academic reputation' is now called `assessment of excellence'); and an explicit reporting is made of the `peer assessment' and `high school counselor assessment' scores that, as before, determine a composite `assessment of excellence'. The last of these three refinements indeed serves to facilitate the qualitative-cum-quantitative re-ranking exercise advocated in this section

and permits variants on it, for example, use of the `high-school counselor assessment' or composite `assessment of excellence' scores in place of `peer assessment' to carry out the quantitative re-ranking within each of the four tiers, where these qualitative tiers---upper, upper-middle, lower-middle, and lower---are obtained, as before, from the overall weighted scores.

Broader Context

How does our simple re-ranking exercise fit into the larger context of OR/MS/Analytics? Ten years ago, during the INFORMS `Science of Better' campaign that sought to raise the profile of Operations Research (OR) and Management Science (MS), we contributed a short essay [3] in which we proposed an alternative metaphor for the field and a broadening of the scope of OR/MS under the banner *`Art and Science of Better'.* In particular, we advocated an alternative training track within a university's academic program that emphasized "practice at the grassroots level", as well as a professional certification by INFORMS. This article was slightly prescient, although we cannot claim to have foreseen the wonderful revolution coming over the horizon through the addition of Analytics (AN) to OR/MS. However, it is heartening to see the name Analytics associated nowadays with the foregoing banner, for example, by the software company SAS, one of the leading exponents of advanced Analytics.

In a recent column, the former INFORMS President Steve Robinson [4] provided a useful way to visually capture the evolution and current stabilization of the field, one that can be embellished as follows. Think of a small circle, or nucleus, corresponding to Industrial Engineering (IE), a field that is as old as the industrial era itself. From it grew the field of Operations Research (OR), also called Operational Research in Europe, which received its initial impetus from the logistical and other needs of the military during WWII and then expanded across many other "operational" fields. Thus OR can be represented pictorially by the area between the IE nucleus and a second, concentric circle that is defined by a broadening of the IE boundary perimeter. (Many academic OR departments continue to be identified by the acronym IEOR or IE&OR.) Management Science (MS) was a further outgrowth to serve the burgeoning areas of Business Administration and Management, and it can be depicted by broadening the OR outer circular perimeter into a third concentric circle. (Today, OR and MS are closely united under the OR/MS and INFORMS umbrellas.) In order to serve the decision-making needs connected to the recently-created field of Data Science, we have yet another broadening of the outermost circular perimeter into a fourth concentric circle, with the expanded area depicting the addition of (descriptive, predictive, prescriptive) Analytics to the INFORMS mission. Its aim is to derive both insight and better decision-making from a data stream, big or small. Finally, as discussed by Robinson [4], a

further broadening may be needed in the future to incorporate the needs of Systems Analysis (Systems Analytics?) and thereby address pressing global problems centered on limited resources and a deteriorating environment.

Thus we see a progressive broadening of the INFORMS mission, at each stage in response to new challenges, which can be depicted pictorially by the foregoing progression of concentric circles. (For convenience, let us hereinafter identify Analytics by the acronym AN, and Systems Analysis, or Systems Analytics, by the acronym SY.) For each of the circular boundary perimeter lines, let us associate the word `advancing' with the inner edge and `advanced' with the outer. Thus we can say that advancing OR borders on, or leads to, advanced MS; advancing OR/MS borders on advanced Analytics; and advancing OR/MS/Analytics borders on advanced Systems Analysis, the widest envelope of them all. Hopefully INFORMS can design a visually-appealing logo that captures this expanding progression from IE → OR → MS → AN → SY → <whatever trans-global issues the future might hold in store>! Drawing on this sequence of acronyms and speaking informally and playfully, we might even venture that the "language" of INFORMS is INFORMANSY, and, in this same playful spirit, we provide a sample design for the aforementioned logo at the end of this article. It remains to be seen whether the field of Operations Research as depicted in this logo will begin to fade from prominence in light of the rapid advance of Analytics (and earlier of Management Science), in much the same way that the field of Industrial Engineering began to fade as Operations Research advanced during previous decades. Two notable and recent instances of such change in the making are as follows: the stellar Department of Operations Research at Stanford University chose to rebrand itself as the Department of Management Science and Engineering; and the also highly-regarded School of Operations Research and Industrial Engineering (ORIE) at Cornell University transformed itself into the School of Operations Research and Information Engineering.

Returning now to the question posed at the start of this section, it is obvious that the cited studies [1] and [2] necessitated the collection and organization of a vast amount of data. The *U.S. News* effort in this regard is highly admirable, especially when one considers that this `big-data' has been gathered, not just over a single year, but over the course of many years. Collecting, organizing, and extracting information from data in an effective way lies within the domain of the newly-identified scientific subject called *Data Science*. But something more specific is being asked: we need to derive insight from the *U.S. News* data stream in order make decisions as to whether one university is "better" than another. This enters the INFORMS province of OR/MS/Analytics. And then it becomes immediately evident that our simple re-ranking exercise of the previous section, which is neither predictive nor prescriptive, falls squarely into the category
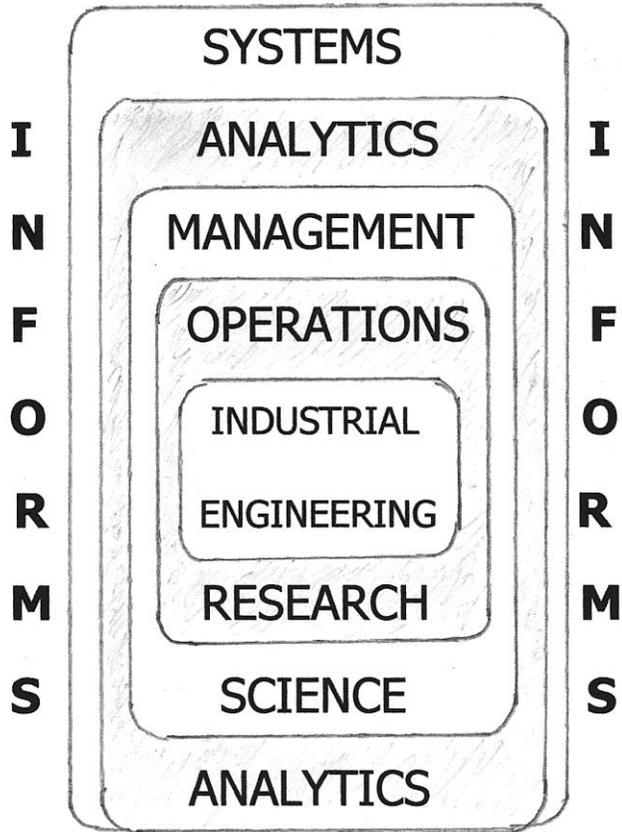
termed data-driven, descriptive Analytics. This exercise has shown that the ranking methodology employed by *U.S. News* has not necessarily improved over the course of the decade separating our two cited studies and needs to be revisited. And in closing, we note also that our discussion has sought to refine but nevertheless remains within the boundaries of this methodology and that more sophisticated ordering, or ranking, knowledge and techniques are available within the broader field of OR/MS and Advanced Analytics.

References

[1] U.S. News and World Report, *Best Colleges*: *Find the Right School for You*, 2011 Edition, pgs. 84-93. ISBN: 978-1-931469-37-1.

[2] U.S. News and World Report, *America's Best Colleges*, September 11, 2000, pgs. 104-110.

[3] Nazareth, J.L., "The art and science of better," *OR/MS Today*, Vol. 32, No. 6, December 2005, pgs. 30-34.

[4] Robinson, S., "Beyond analytics," *OR/MS Today*, Vol. 41, No. 6, December 2014, pg. 8.

Author's info: John Lawrence (Larry) Nazareth is Professor Emeritus at Washington State University and an Affiliate Professor at the University of Washington. He can be reached at nazareth@amath.washington.edu or larry_nazareth@q.com . For additional biographical information, visit the website: www.math.wsu.edu/faculty/nazareth .

INFORMANSY

SYSTEMS

ANALYTICS

MANAGEMENT

OPERATIONS

INDUSTRIAL

ENGINEERING

RESEARCH

SCIENCE

ANALYTICS

INFORMS

INFORMS

INFORMANSY