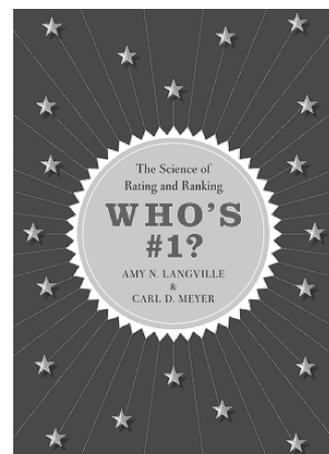


Who's #1? The Science of Rating and Ranking

Reviewed by Andrew I. Dale



Who's #1? The Science of Rating and Ranking

A. N. Langville and C. D. Meyer
Princeton University Press, February 2012
US\$29.95, 266 pages
ISBN-13: 978-06911-542-20

The XXX Olympiad of the modern era has recently taken place, and readers will be fully aware of the considerable enthusiasm, verging at times on the perfervid, with which the outcomes of each day's events were anticipated. Countries were ranked according to the number of medals won, but what would the ranking table be if the competing countries were ranked by population size or GDP¹ (Gross Domestic Product) or by the number of athletes in each team? According to an alternative medals table published in *The Guardian* (a British newspaper), the USA, which gained the greatest number of medals (104), has a GDP rank of 66 (Grenada is ranked 1), a population rank of 47 (again Grenada is 1), and a team size rank of 5 (Grenada 9).

Daily and constantly one is faced with having to choose one or more "things" (for want of a better word) from several alternatives. Such a choice may be made in an almost irrational way or it may be necessary to carefully weigh up the merits (including all relevant factors) of the different options. Our desire, however, may be seemingly simpler: to rank the alternatives according to some

Andrew I. Dale is professor emeritus at the University of Kwazulu-Natal in Durban, South Africa. His email address is dale@ukzn.ac.za.

¹Before one places too much faith in the GDP, one may want to read a recent article in *Significance* in which Marks has argued cogently that the GDP "ignores all the important measures of human life" [5, p. 27].

DOI: <http://dx.doi.org/10.1090/noti934>

measure of "importance" or "bestness", or we may even want to go so far as to choose, say, not just the "best" option, but a subset of options containing the "best" one (though this is really a question of *selection* rather than *ranking*). It is of course necessary in such a situation to know that there *exists* a best option and in such a case to be able to identify it. If we are sophisticated enough, we may well feel the need, after a statistical experiment, for statistical procedures to enlighten us to the probability (and possible consequences) of an erroneous choice.

The intended readership of *Who's #1?*, Langville and Meyer claim in their preface, comprises "sports enthusiasts, social choice theorists, mathematicians, computer scientists, engineers, and college and high school teachers of classes such as linear algebra, optimization, mathematical modeling, graph theory" [p. xiii]. I suspect that the book will be of most interest to one who is already aware of some of the many different methods used in sports to rank football, basketball, etc., teams. To the results of which of these ranking methods can the fan attach most credence, and how do the different rankings compare? Langville and Meyer have presented a most useful work on this subject, exposing the reader to many different methods of ranking, discussing their origins with reference to much original work and exploring the ranked lists obtained by different methods.

While of course the choice of the method of ranking and the input data need careful thought, Langville and Meyer note the necessity for the making of occasional snap decisions: "Evolution has ...rewarded those who make quick comparisons" [p. 3]: those who thought more slowly (or even those who were quick but incorrect) were no doubt removed from the gene pool by the swifter-thinking (or swifter-moving) predator!

Who's #1? is not a text on the *theory* of ranking methods (the authors wisely and correctly distinguish between *ranking* and *rating*). Rather, it is a discussion of various methods from the simple to the more complicated, with illustrations drawn from various fields (mainly American football), often with convenient summarizing sections. The second chapter, for example, deals with Kenneth Massey's original method for ranking college football teams, where the difference in ratings between two teams is used to predict the margin of victory in a future contest between these two. Things get more complicated and methods more refined as we go through the book, even the matter of ties being covered. The last chapter discusses not only the obtaining of but also the appropriate computer entering of data for analysis.

According to Langville and Meyer, the mathematical background required for an easy read consists mainly of elementary linear algebra and some optimization. They suggest that the book can be read for enjoyment even by those who lack these skills, who can implement the requisite techniques by the use of appropriate software, though those relying on such software must always be aware of possible pitfalls in the interpretation of results obtained in this way. Much use is expeditiously made here of the Perron-Frobenius theory to extract a unique ranking vector in a number of cases. The mathematics is by no means "heavy", and the interested reader will be able to follow Langville and Meyer without going into the deeper algebra.

While the methods discussed are, as stated, based on matrix analysis and optimization, Langville and Meyer do point out that there exist methods of more particular use in subjects such as game theory or statistics. In fact, there are techniques required in *Who's #1?* that I would classify as coming from statistics rather than linear algebra: for instance, on p. 10 Langville and Meyer write of the normal equations, minimum variance, and linear unbiased estimates, and the last of these certainly requires some knowledge of statistics for its proper understanding.

Between each two successive chapters there is a page headed "By the Numbers": for instance, we read on page 8 that the attendance at American college football games in 2010 was 50,000,000. The reader interested in such things might like to look at Watson and Yip [8] on the estimation of the sizes of crowds.

Flaws in the various methods are also indicated where necessary, and rankings obtained by different methods on the same data set are compared. For example, some methods for obtaining rankings use data different from others or perhaps weigh the same data differently (the Colley method, for

instance, ignores game scores, but as Langville and Meyer ask, is this a strength or a weakness?).

In Chapter 10 Langville and Meyer discuss user preference rankings, the sort of thing that might arise in preference scores given to products advertised on the World Wide Web. But do respondents rank from the top down (five stars, four stars, etc.) or from the bottom up? This might well make a difference, as Luce [4, §2F] has suggested.

Chapter 16 is concerned with trying to discover which of the many methods discussed thus far is best. Here we find use of the well-known statistical tools *Kendall's tau* and *Spearman's footrule*. Useful here is the bipartite graph for a quick visual assessment of the comparison between methods.

There is some discussion of Arrow's Impossibility Theorem on the nonexistence of a perfect voting system. Perhaps more correctly, Arrow showed that the existence of a perfect voting system is incompatible with the requirement that certain reasonable criteria (e.g., collective rationality, nondictatorship, independence of irrelevant alternatives) be met. (This result was given in Arrow [1, Chap. V] as the general *possibility* theorem for social welfare functions.) Incidentally, Rescher has noted that "It is one of the tasks for the *social* order embracing different individuals to find a means of resolving a unified result out of a mass of potentially divergent individual preferences" [6, p. 99]. And though we may feel *preference* and *value* to be intimately connected, Rescher has suggested that "preference is too gross an instrument to capture the subtle nuances of value" [6, p. 109].

In their discussion of Keener's Method (Chapter 4) Langville and Meyer introduce a_{ij} , the value of some statistic or attribute that is thought to be a good basis for comparing two teams i and j . If S_{ij} is the number of points scored by i against j , then one's first choice might be

$$a_{ij} = \frac{S_{ij}}{S_{ij} + S_{ji}}.$$

They go on, however, to say that Keener indicated that a more appropriate measure would be

$$a_{ij} = \frac{S_{ij} + 1}{S_{ij} + S_{ji} + 2}$$

and add "The motivation for this is Laplace's rule of succession" [p. 31]. However reminiscent this formula might be of the latter rule, I find no such identification in Keener [3], and the resemblance is perhaps more apparent than real.

There are a few points to which one may well take exception. For example, on page 1 we read "because you scored in the 95th percentile": one who prefers the correct technical use of Francis Galton's carefully defined term may like to read

Senn's [7] comments. Further, there are a few errors in the index: for instance, we have "Saaty, T. L." and "Saaty, Thomas" (the references being to different pages); there are entries for "Saverin, Eduardo" and "Eduardo Saverin"; "Bhlmann" should be "Buhlmann" (also on p. 226, and here 1962 should be 1963). There are also some inconsistencies: "Kendall's tau" or "Kendall tau" and "Spearman's footrule" or "Spearman footrule"? Finally, in the bibliography, reference [47] is out of place, and the paper cited in reference [80] in fact occupies pages 155-169. Perhaps such things are bound to happen in a jointly written book with a computer-generated index.

The reader who is not *au fait* with American football will find enough discussion here of rankings in other areas to whet his/her appetite for more. Langville and Meyer refer frequently to the Netflix system for the online renting of films whose popularity has been rated. Recently a large prize was offered to anyone who could make a ten percent improvement in the company's own recommendation system: the question has undergone statistical investigation by Feuerverger et al. [2]. As an example of other recent work we mention the results of the British Film Industry magazine *Sight & Sound's* recently published 2012 poll (846 respondents). For the first time since 1962 in the ten-yearly report, *Citizen Kane* was knocked out of first place by Alfred Hitchcock's 1958 film *Vertigo*.

When I started this book I knew very little about American football. I was little the wiser after finishing it, but I had an excellent understanding of various methods used in the obtaining of the ranking of teams and their interrelationships. Langville and Meyer are to be commended for this collection, and anyone who is more conversant with North American sports than I am will most certainly be stimulated by reading *Who's #1?*

References

- [1] K. J. ARROW, *Social Choice and Individual Values*, New York: John Wiley & Sons, 1963.
- [2] A. FEUERVERGER, Y. HE, and S. KHATRI, Statistical significance of the Netflix challenge, *Statistical Science* 27, no. 2 (2012), 202-231.
- [3] J. P. KEENER, The Perron-Frobenius theorem and the ranking of football teams, *SIAM Review* 35, no. 1 (1993), 80-93.
- [4] R. D. LUCE, *Individual Choice Behavior: A Theoretical Analysis*, New York: John Wiley & Sons, 1959.
- [5] N. MARKS, GDP RIP (1933-2010), *Significance* 7, no. 1 (2010), 27-30.
- [6] N. RESCHER, *Introduction to Value Theory*, Englewood Cliffs: Prentice-Hall, 1969.
- [7] S. SENN, Francis Galton and regression to the mean, *Significance* 8, no. 3 (2011), 124-126.
- [8] R. WATSON and P. YIP, How many were there when it mattered? Estimating the sizes of crowds, *Significance* 8, no. 3 (2011), 104-107.

About the Cover

2013 marks the 125th anniversary of the American Mathematical Society. The first meeting of the Society was held on Thanksgiving Day in 1888. When it was founded, the Society had dedicated mathematicians as members, but it had no staff.

In 2013, the Society has approximately 30,000 individual members and 570 institutional members supported by a staff of over 200. This month's cover shows staff photos from the AMS's four facilities.

The Ann Arbor staff, numbering approximately 75, creates and maintains the Mathematical Reviews database and MathSciNet. Their job is a big one, staying abreast of the mathematical research literature in nearly 2,000 journals, plus research monographs and numerous refereed conference proceedings.

The AMS staff members in Pawtucket, Rhode Island, print and distribute all of the books published by the AMS, all of the AMS research journals that have print editions, and materials that are widely distributed as part of the Society's public awareness program.

The Washington Office staff is small in number, but has a very big footprint in its leadership of government relations and public policy. Its mission is advocacy and the advancement of the research and educational interests of the sciences and all areas of mathematics.

The Society's Providence headquarters has the largest and most diverse staff. Approximately 115 staff members cover meetings, professional programs, membership services, organizational infrastructure for finance and information technology, acquisition and prepress work for AMS publications, and support of Society governance. The original focus of the Society in 1888, meetings, has grown today to include support of eight sectional meetings, the national meeting as part of the January Joint Mathematics Meetings, and one international meeting every year. The publishing program encompasses thirteen journals and a broad book program.

The Society's activities and its staff have come a long way from the inspired beginning in 1888.

—Donald E. McClure
Executive Director