

Who's Bigger?

Where Historical Figures *Really* Rank

Is Napoleon bigger than Hitler? Washington bigger than Lincoln? Picasso bigger than Einstein?

Quantitative analysts are finding homes in cultural domains from finance to politics. What about history? In this fascinating book, Steven Skiena and Charles B. Ward bring quantitative analysis to bear on ranking and comparing historical reputations. They evaluate each person by aggregating the traces of millions of opinions, similar to how Google ranks web pages.

Did you know:

- Women remain significantly underrepresented in the historical record compared to men and have long required substantially greater achievement levels to be equally noted for posterity?
- The long-term prominence of Elvis Presley rivals that of the most famous classical composers? Roll over Beethoven, and tell Tchaikovsky the news!
- If you've got a spare billion dollars, and want to be remembered forever, your best investment is to get a university named after you?

Along the way, the authors present the rankings of more than one thousand of history's most significant people in all areas of human endeavor, including science, politics, and entertainment. Anyone interested in history or biography can see where their favorite figures place in the grand scheme of things. While revisiting old historical friends and making new ones, you will come to understand the forces that shape historical recognition in a whole new light.

About the Authors

Steven Skiena serves as Distinguished Teaching Professor of Computer Science at Stony Brook University. He is the author of four well-regarded books: *The Algorithm Design Manual* (2008); *Calculated Bets: Computers, Gambling, and Mathematical Modeling to Win* (2001); *Programming Challenges* (with Miguel Revilla, 2003); and *Computational Discrete Mathematics* (with Sriram Pemmaraju, 2003). Skiena heads the Data Science Laboratory at Stony Brook, using large-scale text analysis to chart the frequency, sentiment, and relationships among millions of people, places, and things. This technology forms the foundation of General Sentiment (<http://www.generalsentiment.com>), where he serves as co-founder and chief scientist. His news analysis has been applied to research projects from financial forecasting to presidential election analysis. The rankings underlying *Who's Bigger?* derive from this analysis.

Charles B. Ward works as an engineer on the ranking team at Google. He is the author of more than a dozen scholarly papers, including research in text analysis, computational social science, computational biology, and graph theory. Ward worked as a lead developer with the Data Science Laboratory during his four years of postdoctoral studies at Stony Brook University. In his spare time, he is an accomplished solo pianist. He is also an authority on historical strategy games, and spends perhaps too much time playing and designing them.

WHO'S BIGGER?

*Where Historical Figures
Really Rank*

STEVEN SKIENA

Stony Brook University, New York

CHARLES B. WARD



CAMBRIDGE
UNIVERSITY PRESS

Cambridge University Press & Assessment
 978-1-107-04137-0 — Who's Bigger?
 Steven Skiena, Charles B. Ward
 Frontmatter
[More Information](#)



CAMBRIDGE
UNIVERSITY PRESS

Shaftesbury Road, Cambridge CB2 8EA, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

103 Penang Road, #05–06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment,
 a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of
 education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107041370

© Steven Skiena and Charles B. Ward 2014

This publication is in copyright. Subject to statutory exception and to the provisions
 of relevant collective licensing agreements, no reproduction of any part may take
 place without the written permission of Cambridge University Press & Assessment.

First published 2014

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication data

Skiena, Steven S., author.

Who's bigger? : where historical figures really rank / Steven Skiena, Charles B. Ward.

pages cm

Includes bibliographical references and index

ISBN 978-1-107-04137-0 (hardback)

1. Quantitative research. I. Ward, Charles, 1984– author. II. Title.

Q180.55.Q36S55 2013

920.02–dc23 2013032173 CIP

ISBN 978-1-107-04137-0 Hardback

Additional resources for this publication at <http://www.whoisbigger.com>

Cambridge University Press & Assessment has no responsibility for the persistence
 or accuracy of URLs for external or third-party internet websites referred to in this
 publication and does not guarantee that any content on such websites is, or will
 remain, accurate or appropriate.

Contents

<i>Rationale</i>	<i>page</i> ix
Part I Quantitative History	
1 History’s Most Significant People	3
1.1 <i>People as Memes</i>	4
1.2 <i>Our 100</i>	4
1.3 <i>Other People’s Rankings</i>	7
1.4 <i>History vs. Historiography</i>	11
1.5 <i>Overview</i>	14
2 Ranking Historical Figures	15
2.1 <i>Traditional Ranking Methodologies</i>	15
2.2 <i>Measurements of Historical Reputation</i>	17
2.3 <i>Celebrity vs. Gravitas</i>	28
2.4 <i>The Half-Life of Fame</i>	30
2.5 <i>Evaluation Results</i>	35
3 Who Belongs in Bonnie’s Textbook?	38
3.1 <i>The Significance of Significance</i>	39
3.2 <i>The People in Bonnie’s Textbook</i>	41
3.3 <i>Political Leaders</i>	44
3.4 <i>People by Time and Place</i>	49
3.5 <i>Cultural and Social Innovators</i>	54
3.6 <i>Diversity</i>	58

3.7	<i>Educational Standards vs. Editorial Decision Making</i>	62
3.8	<i>A Nation of Fourth Graders?</i>	64
4	Reading through the Past	67
4.1	<i>Google Ngrams</i>	67
4.2	<i>Visualizing Ngram Data</i>	70
4.3	<i>Ngrams and Significance</i>	74
4.4	<i>Posthumous Fame</i>	77
4.5	<i>Faded Glory</i>	79
5	Great Americans and the Process of Canonization	82
5.1	<i>The Hall of Fame for Great Americans</i>	82
5.2	<i>Who's In?</i>	83
5.3	<i>Electoral History</i>	87
5.4	<i>Decline and Fall</i>	108
6	The Baseball Hall of Fame	111
6.1	<i>Performance and Posterity</i>	112
6.2	<i>The Electoral Record</i>	114
6.3	<i>The Glory of Their Times</i>	120
6.4	<i>Who's In?</i>	122
6.5	<i>The Negro Leagues</i>	131
6.6	<i>Who's Out?: Missing Players</i>	133
6.7	<i>The Market for Collectibles</i>	135
7	Historical Timescales	140
7.1	<i>Life-Span Analysis</i>	140
7.2	<i>Birth Date Effects</i>	147
7.3	<i>Ancient History</i>	152
7.4	<i>The Middle Ages</i>	156
7.5	<i>The Modern Era</i>	158
7.6	<i>The Historical Representation of Women</i>	164
Part II Historical Rankings		
8	American Political Figures	171
8.1	<i>Presidents of the United States</i>	172
8.2	<i>Congressmen</i>	180

CONTENTS		vii
8.3	<i>Governors</i>	188
8.4	<i>Mayors</i>	190
8.5	<i>Judges and Justices</i>	192
8.6	<i>Founding Mothers and Fathers</i>	193
9	Modern World Leaders	196
9.1	<i>Analyzing Political Leadership</i>	196
9.2	<i>The Supreme Leaders</i>	202
9.3	<i>Dictators and Despots</i>	215
9.4	<i>Military Leaders</i>	216
10	Science and Technology	218
10.1	<i>Scientists</i>	218
10.2	<i>Inventors</i>	230
10.3	<i>Business Leaders</i>	235
10.4	<i>Explorers and Adventurers</i>	240
11	Religion and Philosophy	244
11.1	<i>Biblical Figures</i>	245
11.2	<i>The Catholic Church</i>	248
11.3	<i>Other Religious Denominations</i>	253
11.4	<i>Philosophers</i>	258
12	Sports	260
12.1	<i>Individual Sports</i>	260
12.2	<i>Team Sports</i>	266
12.3	<i>Olympic Sports Stars</i>	272
13	The Arts	277
13.1	<i>Literature</i>	277
13.2	<i>Popular Writers</i>	287
13.3	<i>Painters and Sculptors</i>	292
13.4	<i>Architects</i>	297
14	The Performing Arts	300
14.1	<i>Before Recording Technology</i>	300
14.2	<i>Western Classical Music</i>	302
14.3	<i>Popular Music</i>	307

viii	CONTENTS	
14.4	<i>Motion Pictures</i>	310
14.5	<i>Radio and Television</i>	318
15	Devils and Angels	322
15.1	<i>The Killers</i>	322
15.2	<i>Law and Order</i>	327
15.3	<i>Peacemakers</i>	329
15.4	<i>Social Activists</i>	331
15.5	<i>Civil Rights Leaders</i>	335
Appendix A	Ranking Methodology	339
A.1	<i>Feature Set</i>	339
A.2	<i>Ranking Methodology</i>	342
A.3	<i>The Decay of Fame</i>	343
Appendix B	Resources	347
B.1	<i>Websites and Databases</i>	347
B.2	<i>Birth/Death Calendars</i>	348
B.3	<i>The Who's Bigger App</i>	353
Appendix C	Biographical Dictionary	355
	<i>Bibliography</i>	361
	<i>Index</i>	367

Rationale

Quantitative analysts (“quants”) are finding new homes in the social and cultural domains. Finance? Hedge funds employing quantitative trading strategies now rule Wall Street, to the extent that more than 80 percent of all buy-sell decisions are made by algorithms, not people. Sports? The best-selling book *Moneyball* detailed how successful major league teams now hire general managers who value statistical analysis above gut instinct – and get to be played by Brad Pitt in the movie version of the story. Politics? Prediction markets and the meta-analysis of polling data work so reliably that it hardly seems necessary to hold the actual election.

In this book, we bring quantitative analysis to bear on ranking and comparing historical reputations. Who’s bigger: Washington or Lincoln? Hitler or Napoleon? Picasso or Michelangelo? Charles Dickens or Jane Austen? Did you realize that:

- Although *Paul Revere* (1735–1818) [627] and *Betsy Ross* (1752–1836) [2430] are well known to all American schoolchildren, they fell into complete obscurity for several generations after their contributions to the American Revolution. Their rediscoveries, completely independent of their actual achievements, tell us much about the capricious forces of history.
- Women remain significantly underrepresented in the historical record compared to men. We can prove that women have long required substantially greater achievement levels (analogous to about 4 IQ points in the mean) than men to get equally noted for posterity.

- Got a spare billion dollars, and want to be remembered forever? Your best investment is to get a university named after you.
- The long-term historical significance of *Elvis Presley (1935–1977)* [69] rivals that of the most famous classical composers. Roll over *Beethoven* [27], and tell *Tchaikovsky* [63] the news!

This is all good fun, but our quantitative approach also enables us to address important cultural questions in a rigorous manner:

- Are the historical figures discussed in school texts *really* worth teaching?
- How good are committees at recognizing historical greatness?
- Why do reputations change posthumously? Have copyright laws affected the literary canon, determining which authors get revived and which are forgotten?

We believe this book will be of interest to a broad readership and that anyone who likes history or biography will be curious about our rankings, eager to see where their favorite figures place in the grand scheme of things. Our book provides an opportunity to revisit old historical friends and make new ones. The reader will come to understand the forces that shape historical recognition in quite a different way after reading it. We detail the difference between fame and significance, the cultural biases in how we think about the past, and how rapidly reputations decay with time.

We also expect our share of right-brained readers, attracted by our quantitative methods and intrigued by our analysis. No mathematical or computational background is necessary to understand our rankings or conclusions. But we include a technical discussion for readers interested in the details of our methods.

We anticipate that professional historians may be skeptical of computer scientists intruding on their turf. But what we are really trying to do here is study what shapes the process of historical recollection. We predict that many scholars will be taken with the power of our methods, and will seek to employ these tools in their own studies. We have collaborated on research projects with political scientists and sociologists: watching them change from aloof to curious and then finally excited. We are making our massive datasets available to the academic community (and the curious public) through our website, <http://www.whoisbigger.com>. Computational

Social Science is a new and growing area of academic scholarship, and we hope to bring the field greater visibility and acceptance.

Textual Notes

We employ certain textual conventions in this book that may confuse readers, until they understand our rationale. We explain them here briefly.

More than 1,000 historical figures are mentioned by name within the body of this book. At each such reference, each person's name will be accompanied by their dates of birth and death, as well as their significance rank as assessed by our methods. Thus, Brad Pitt is *Brad Pitt (1963–)* [1845] and Aristotle is *Aristotle (384–322 B.C.)* [8]. We sort people from most significant to least significant, and assign ranks by counting down from the top. Thus the most significant person gets rank 1, and smaller numbers are better than bigger ones. Embedding our ranking with every reference will help you appreciate each individual's historical magnitude, and implicitly encourage you to evaluate the performance of our methods.

Describing first-hand experiences and opinions within a coauthored text is somewhat problematic. The royal “We” isn't always accurate here, since we are distinct individuals: one raised a Southern gentleman while the other was toilet-trained in the Bronx. And so we have adopted the following convention. We both will lay claim to “I,” distinguishing which I when necessary. Thus “I (Steve)” will be used to represent bits of Bronx cheer. But we will mostly be “we,” since we developed the methods, analysis, and contents of this text together, as a team.

Acknowledgments

I (Steve) dedicate this book to my second daughter, Abby. Fate has determined that she will never get to experience things like the fifth grade ahead of her older sister. Abby: don't worry, because you will get your own opportunities to make history. You are second only chronologically, but equal where it counts, in our hearts. And to my wife Renee: you know you will always be the most significant person in my book.

I (Charles) dedicate this book to my parents and Srivani. One could not wish for better parents or a better partner.

We acknowledge the help of several people, who should be famous even if they have not gotten there yet. Bala Mundiam built much of the infrastructure underlying our data collection and display, making this project possible. Ajeesh Elikkottil and Dhruv Matani updated our rankings engine. Vincent Tseui and Wenbin Lin built the www.whoisbigger.com site into what it is today. Qi Zhou contributed background research, data curation, and other bits of analysis. Goutham Bhat polished our Ngram data to perfection. The entire Data Science Laboratory (most recently, Rami al-Rfou, Bryan Perozzi, and Yanqing Chen) contributed to general discussions and generated some of the data underlying our analysis. We thank the National Science Foundation for their interest in our work, which was partially supported by NSF Grant IIS-1017181.

We thank the dozen book agents who passed on this project, driving us safely back into the arms of Lauren Cowles at Cambridge University Press. It has always been a pleasure to work with her. We particularly appreciate her help in shaping this manuscript. We also thank other Cambridge people who helped to get our book into your hands: Rachel Ewen, Josh Penney, Melissanne Scheid, Katy Strong, and Karen Verde. Finally, we thank other friends who read preliminary versions of our book, including Michael Holloway, Theo Pavlidis, Scott Perlman, Marshall Poe, Arnout van de Rijt, Eran Shor, and Jerry and Suzanne Trajan.

Sir *Walter Raleigh* (1552–1618) [313] wrote his history of the world while in prison, awaiting execution. We wrote ours at Stony Brook University. On balance, we feel more blessed by circumstances. In particular, we have enjoyed working with social scientists here to help make sense of our data, including Arnout van de Rijt, Eran Shor, Leonie Huddy, Matt Lebo, and Ellen Key.