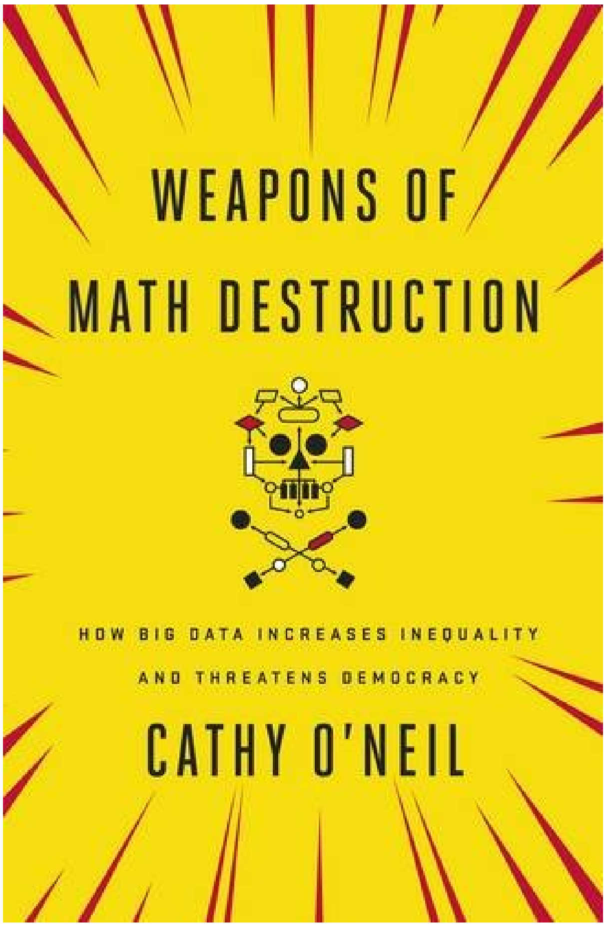
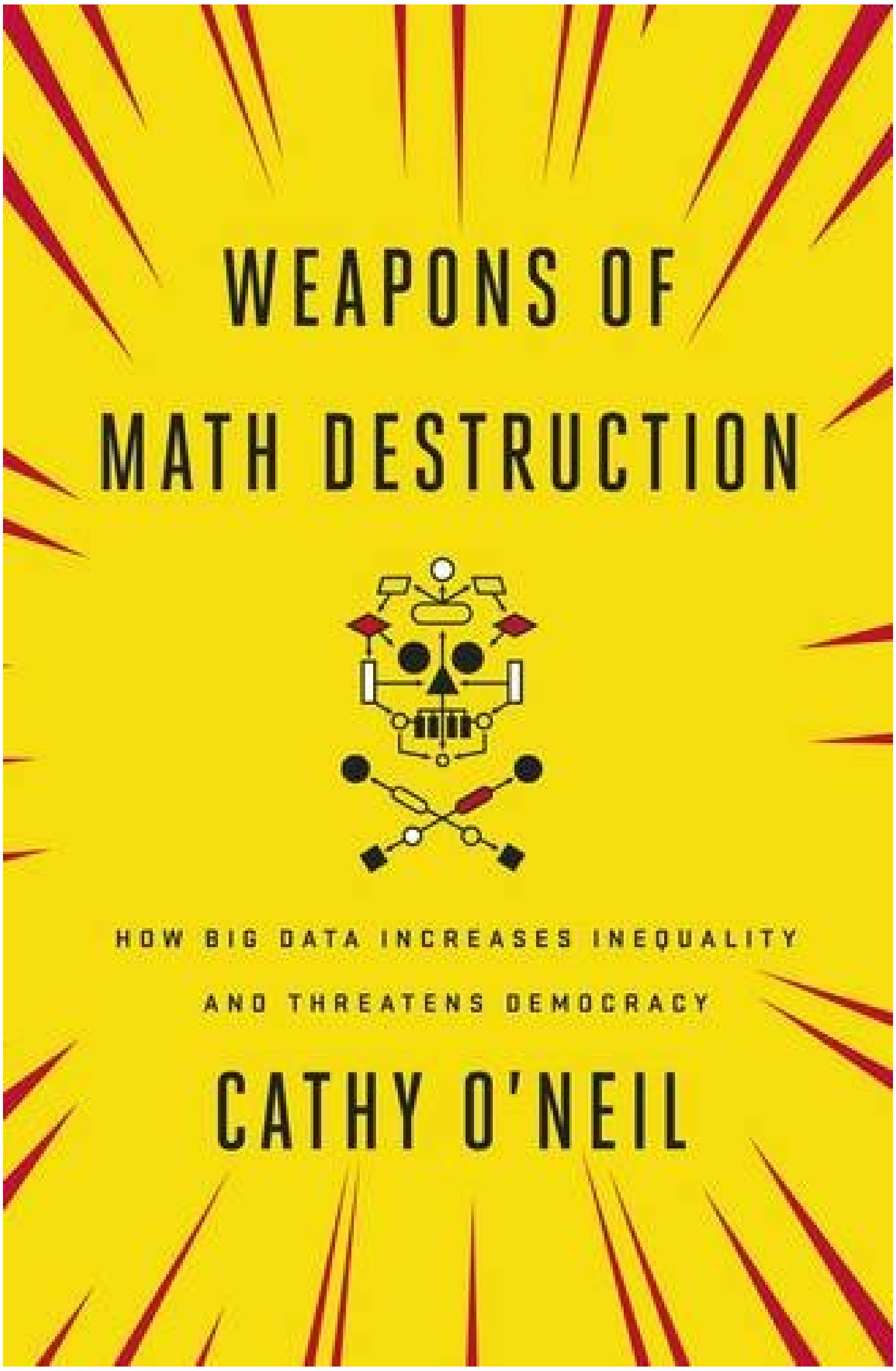
WEAPONS OF MATH DESTRUCTION BY CATHY O'NEIL




WEAPONS OF
MATH DESTRUCTION
HOW BIG DATA INCREASES INEQUALITY
AND THREATENS DEMOCRACY
CATHY O'NEIL

WEAPONS OF MATH DESTRUCTION BY CATHY O'NEIL PDF

To overcome the issue, we now give you the modern technology to download the book *Weapons Of Math Destruction By Cathy O'Neil* not in a thick published documents. Yeah, reading Weapons Of Math Destruction By Cathy O'Neil by online or obtaining the soft-file just to read can be among the methods to do. You might not really feel that reviewing a book Weapons Of Math Destruction By Cathy O'Neil will work for you. But, in some terms, May people effective are those that have reading behavior, included this type of this Weapons Of Math Destruction By Cathy O'Neil

Review
A New York Times Book Review Notable Book of 2016
A Boston Globe Best Book of 2016
One of Wired's Required Reading Picks of 2016
One of Fortune's Favorite Books of 2016
A Kirkus Reviews Best Book of 2016
A Chicago Public Library Best Book of 2016
A Nature.com Best Book of 2016
An On Point Best Book of 2016
New York Times Editor's Choice
A Maclean's Bestseller
Winner of the 2016 SLA-NY PrivCo Spotlight Award

“O’Neil’s book offers a frightening look at how algorithms are increasingly regulating people… Her knowledge of the power and risks of mathematical models, coupled with a gift for analogy, makes her one of the most valuable observers of the continuing weaponization of big data… [She] does a masterly job explaining the pervasiveness and risks of the algorithms that regulate our lives.”
—New York Times Book Review

"Weapons of Math Destruction is the Big Data story Silicon Valley proponents won't tell… [It] pithily exposes flaws in how information is used to assess everything from creditworthiness to policing tactics… A thought-provoking read for anyone inclined to believe that data doesn't lie.”
—Reuters

“This is a manual for the 21st-century citizen, and it succeeds where other big data accounts have failed—it is accessible, refreshingly critical and feels relevant and urgent.”
—Financial Times

"Insightful and disturbing."
—New York Review of Books

“Weapons of Math Destruction is an urgent critique of… the rampant misuse of math in nearly every aspect

of our lives."
—Boston Globe

"A fascinating and deeply disturbing book."
—Yuval Noah Harari, author of Sapiens; The Guardian's Best Books of 2016

"Illuminating… [O'Neil] makes a convincing case that this reliance on algorithms has gone too far."
—The Atlantic

"A nuanced reminder that big data is only as good as the people wielding it."
—Wired

"If you've ever suspected there was something baleful about our deep trust in data, but lacked the mathematical skills to figure out exactly what it was, this is the book for you."
—Salon

"O'Neil is an ideal person to write this book. She is an academic mathematician turned Wall Street quant turned data scientist who has been involved in Occupy Wall Street and recently started an algorithmic auditing company. She is one of the strongest voices speaking out for limiting the ways we allow algorithms to influence our lives… While Weapons of Math Destruction is full of hard truths and grim statistics, it is also accessible and even entertaining. O'Neil's writing is direct and easy to read—I devoured it in an afternoon."
—Scientific American

"Readable and engaging… succinct and cogent… Weapons of Math Destruction is The Jungle of our age… [It] should be required reading for all data scientists and for any organizational decision-maker convinced that a mathematical model can replace human judgment."
—Mark Van Hollebeke, Data and Society: Points

"Indispensable… Despite the technical complexity of its subject, Weapons of Math Destruction lucidly guides readers through these complex modelling systems… O'Neil's book is an excellent primer on the ethical and moral risks of Big Data and an algorithmically dependent world… For those curious about how Big Data can help them and their businesses, or how it has been reshaping the world around them, Weapons of Math Destruction is an essential starting place."
—National Post

"Cathy O'Neil has seen Big Data from the inside, and the picture isn't pretty. Weapons of Math Destruction opens the curtain on algorithms that exploit people and distort the truth while posing as neutral mathematical tools. This book is wise, fierce, and desperately necessary."
—Jordan Ellenberg, University of Wisconsin-Madison, author of How Not To Be Wrong

"O'Neil has become [a whistle-blower] for the world of Big Data… [in] her important new book… Her work makes particularly disturbing points about how being on the wrong side of an algorithmic decision can snowball in incredibly destructive ways."
—TIME

"O'Neil's work is so important… [her] book is a vital crash-course in the specialized kind of statistical knowledge we all need to interrogate the systems around us and demand better."
—Boing Boing

“Cathy O’Neil, a number theorist turned data scientist, delivers a simple but important message: Statistical models are everywhere, and they exert increasing power over many aspects of our daily lives… Weapons of Math Destruction provides a handy map to a few of the many areas of our lives over which invisible algorithms have gained some control. As the empire of big data continues to expand, Cathy O’Neil’s reminder of the need for vigilance is welcome and necessary.”
—American Prospect

“An avowed math nerd, O’Neil has written an engaging description of the effect of crunched data on our lives.”
—Hicklebee’s, San Francisco Chronicle

“By tracking how algorithms shape people's lives at every stage, O'Neil makes a compelling case that our bot overlords are using data to discriminate unfairly and foreclose democratic choices. If you work with data, or just produce reams of it online, this is a must-read.”
—ArsTechnica

“Lucid, alarming, and valuable… [O’Neil’s] writing is crisp and precise as she aims her arguments to a lay audience. This makes for a remarkably page-turning read for a book about algorithms. Weapons of Math Destruction should be required reading for anybody whose life will be affected by Big Data, which is to say: required reading for everyone. It’s a wake-up call – a journalistic heir to The Jungle and Silent Spring. Like those books, it should change the course of American society.”
—Aspen Times

"[O'Neil's] propulsive study reveals many models that are currently 'micromanaging' the US economy as opaque and riddled with bias."
—Nature

“You don’t need to be a nerd to appreciate the significance of [O’Neil’s] message… Weapons is a must-read for anyone who is working to combat economic and racial discrimination.”
—Goop

"Cathy O’Neil’s book... is important and covers issues everyone should care about. Bonus points: it’s accessible, compelling, and—something I wasn’t expecting—really fun to read.”
—Inside Higher Ed

“Often we don’t even know where to look for those important algorithms, because by definition the most dangerous ones are also the most secretive. That’s why the catalogue of case studies in O’Neil’s book are so important; she’s telling us where to look.”
—The Guardian

“O’Neil is passionate about exposing the harmful effects of Big Data–driven mathematical models (what she calls WMDs), and she’s uniquely qualified for the task… [She] makes a convincing case that many mathematical models today are engineered to benefit the powerful at the expense of the powerless… [and] has written an entertaining and timely book that gives readers the tools to cut through the ideological fog obscuring the dangers of the Big Data revolution.”
—In These Times

“In this simultaneously illuminating and disturbing account, [O’Neil] describes the many ways in which widely used mathematic models—based on ‘prejudice, misunderstanding, and bias’—tend to punish the poor

and reward the rich… She convincingly argues for both more responsible modeling and federal regulation. An unusually lucid and readable look at the daunting algorithms that govern so many aspects of our lives."
—Kirkus Reviews (starred)

"Even as a professional mathematician, I had no idea how insidious Big Data could be until I read Weapons of Math Destruction. Though terrifying, it's a surprisingly fun read: O'Neil's vision of a world run by algorithms is laced with dark humor and exasperation—like a modern-day Dr. Strangelove or Catch-22. It is eye-opening, disturbing, and deeply important."
—Steven Strogatz, Cornell University, author of The Joy of x

"This taut and accessible volume, the stuff of technophobes' nightmares, explores the myriad ways in which largescale data modeling has made the world a less just and equal place. O'Neil speaks from a place of authority on the subject… Unlike some other recent books on data collection, hers is not hysterical; she offers more of a chilly wake-up call as she walks readers through the ways the 'big data' industry has facilitated social ills such as skyrocketing college tuitions, policing based on racial profiling, and high unemployment rates in vulnerable communities… eerily prescient."
—Publishers Weekly

"Well-written, entertaining and very valuable."
—Times Higher Education

"Not math heavy, but written in an exceedingly accessible, almost literary style; [O'Neil's] fascinating case studies of WMDs fit neatly into the genre of dystopian literature. There's a little Philip K. Dick, a little Orwell, a little Kafka in her portrait of powerful bureaucracies ceding control of the most intimate decisions of our lives to hyper-empowered computer models riddled with all of our unresolved, atavistic human biases."
—Paris Review

"Through harrowing real-world examples and lively story-telling, Weapons of Math Destruction shines invaluable light on the invisible algorithms and complex mathematical models used by government and big business to undermine equality and increase private power. Combating secrecy with clarity and confusion with understanding, this book can help us change course before it's too late."
—Astra Taylor, author of The People's Platform: Taking Back Power and Culture in the Digital Age

"Weapons of Math Destruction is a fantastic, plainspoken call to arms. It acknowledges that models aren't going away: As a tool for identifying people in difficulty, they are amazing. But as a tool for punishing and disenfranchising, they're a nightmare."
—Cory Doctorow, author of Little Brother and co-editor of Boing Boing

"Many algorithms are slaves to the inequalities of power and prejudice. If you don't want these algorithms to become your masters, read Weapons of Math Destruction by Cathy O'Neil to deconstruct the latest growing tyranny of an arrogant establishment."
—Ralph Nader, author of Unsafe at Any Speed

"In this fascinating account, Cathy O'Neil leverages her expertise in mathematics and her passion for social justice to poke holes in the triumphant narrative of Big Data. She makes a compelling case that math is being used to squeeze marginalized segments of society and magnify inequities. Her analysis is superb, her writing is enticing, and her findings are unsettling."
—danah boyd, founder of Data & Society and author of It's Complicated

"From getting a job to finding a spouse, predictive algorithms are silently shaping and controlling our destinies. Cathy O'Neil takes us on a journey of outrage and wonder, with prose that makes you feel like it's just a conversation. But it's an important one. We need to reckon with technology."
—Linda Tirado, author of Hand to Mouth: Living in Bootstrap America

"Next time you hear someone gushing uncritically about the wonders of Big Data, show them Weapons of Math Destruction. It'll be salutary."
—Felix Salmon, Fusion

About the Author
Cathy O'Neil is a data scientist and author of the blog mathbabe.org. She earned a Ph.D. in mathematics from Harvard and taught at Barnard College before moving to the private sector, where she worked for the hedge fund D. E. Shaw. She then worked as a data scientist at various start-ups, building models that predict people's purchases and clicks. O'Neil started the Lede Program in Data Journalism at Columbia and is the author of Doing Data Science. She is currently a columnist for Bloomberg View.

1

BOMB PARTS
What Is a Model?

It was a hot August afternoon in 1946. Lou Boudreau, the player-manager of the Cleveland Indians, was having a miserable day. In the first game of a doubleheader, Ted Williams had almost single-handedly annihilated his team. Williams, perhaps the game's greatest hitter at the time, had smashed three home runs and driven home eight. The Indians ended up losing 11 to 10.

Boudreau had to take action. So when Williams came up for the first time in the second game, players on the Indians' side started moving around. Boudreau, the shortstop, jogged over to where the second baseman would usually stand, and the second baseman backed into short right field. The third baseman moved to his left, into the shortstop's hole. It was clear that Boudreau, perhaps out of desperation, was shifting the entire orientation of his defense in an attempt to turn Ted Williams's hits into outs.

In other words, he was thinking like a data scientist. He had analyzed crude data, most of it observational: Ted Williams usually hit the ball to right field. Then he adjusted. And it worked. Fielders caught more of Williams's blistering line drives than before (though they could do nothing about the home runs sailing over their heads).

If you go to a major league baseball game today, you'll see that defenses now treat nearly every player like Ted Williams. While Boudreau merely observed where Williams usually hit the ball, managers now know precisely where every player has hit every ball over the last week, over the last month, throughout his career, against left-handers, when he has two strikes, and so on. Using this historical data, they analyze their current situation and calculate the positioning that is associated with the highest probability of success. And that sometimes involves moving players far across the field.

Shifting defenses is only one piece of a much larger question: What steps can baseball teams take to maximize the probability that they'll win? In their hunt for answers, baseball statisticians have scrutinized every variable they can quantify and attached it to a value. How much more is a double worth than a single? When, if ever, is it worth it to bunt a runner from first to second base?

The answers to all of these questions are blended and combined into mathematical models of their sport. These are parallel universes of the baseball world, each a complex tapestry of probabilities. They include every measurable relationship among every one of the sport's components, from walks to home runs to the players themselves. The purpose of the model is to run different scenarios at every juncture, looking for the optimal combinations. If the Yankees bring in a right-handed pitcher to face Angels slugger Mike Trout, as compared to leaving in the current pitcher, how much more likely are they to get him out? And how will that affect their overall odds of winning?

Baseball is an ideal home for predictive mathematical modeling. As Michael Lewis wrote in his 2003 bestseller, Moneyball, the sport has attracted data nerds throughout its history. In decades past, fans would pore over the stats on the back of baseball cards, analyzing Carl Yastrzemski's home run patterns or comparing Roger Clemens's and Dwight Gooden's strikeout totals. But starting in the 1980s, serious statisticians started to investigate what these figures, along with an avalanche of new ones, really meant: how they translated into wins, and how executives could maximize success with a minimum of dollars.

"Moneyball" is now shorthand for any statistical approach in domains long ruled by the gut. But baseball represents a healthy case study—and it serves as a useful contrast to the toxic models, or WMDs, that are popping up in so many areas of our lives. Baseball models are fair, in part, because they're transparent. Everyone has access to the stats and can understand more or less how they're interpreted. Yes, one team's model might give more value to home run hitters, while another might discount them a bit, because sluggers tend to strike out a lot. But in either case, the numbers of home runs and strikeouts are there for everyone to see.

Baseball also has statistical rigor. Its gurus have an immense data set at hand, almost all of it directly related to the performance of players in the game. Moreover, their data is highly relevant to the outcomes they are trying to predict. This may sound obvious, but as we'll see throughout this book, the folks building WMDs routinely lack data for the behaviors they're most interested in. So they substitute stand-in data, or proxies. They draw statistical correlations between a person's zip code or language patterns and her potential to pay back a loan or handle a job. These correlations are discriminatory, and some of them are illegal. Baseball models, for the most part, don't use proxies because they use pertinent inputs like balls, strikes, and hits.

Most crucially, that data is constantly pouring in, with new statistics from an average of twelve or thirteen games arriving daily from April to October. Statisticians can compare the results of these games to the predictions of their models, and they can see where they were wrong. Maybe they predicted that a left-handed reliever would give up lots of hits to right-handed batters—and yet he mowed them down. If so, the stats team has to tweak their model and also carry out research on why they got it wrong. Did the pitcher's new screwball affect his statistics? Does he pitch better at night? Whatever they learn, they can feed back into the model, refining it. That's how trustworthy models operate. They maintain a constant back-and-forth with whatever in the world they're trying to understand or predict. Conditions change, and so must the model.

Now, you may look at the baseball model, with its thousands of changing variables, and wonder how we could even be comparing it to the model used to evaluate teachers in Washington, D.C., schools. In one of them, an entire sport is modeled in fastidious detail and updated continuously. The other, while cloaked in mystery, appears to lean heavily on a handful of test results from one year to the next. Is that really a model?

The answer is yes. A model, after all, is nothing more than an abstract representation of some process, be it a baseball game, an oil company's supply chain, a foreign government's actions, or a movie theater's attendance. Whether it's running in a computer program or in our head, the model takes what we know and uses it to predict responses in various situations. All of us carry thousands of models in our heads. They tell us what to expect, and they guide our decisions.

Here's an informal model I use every day. As a mother of three, I cook the meals at home—my husband, bless his heart, cannot remember to put salt in pasta water. Each night when I begin to cook a family meal, I internally and intuitively model everyone's appetite. I know that one of my sons loves chicken (but hates hamburgers), while another will eat only the pasta (with extra grated parmesan cheese). But I also have to take into account that people's appetites vary from day to day, so a change can catch my model by surprise. There's some unavoidable uncertainty involved.

The input to my internal cooking model is the information I have about my family, the ingredients I have on hand or I know are available, and my own energy, time, and ambition. The output is how and what I decide to cook. I evaluate the success of a meal by how satisfied my family seems at the end of it, how much they've eaten, and how healthy the food was. Seeing how well it is received and how much of it is enjoyed allows me to update my model for the next time I cook. The updates and adjustments make it what statisticians call a "dynamic model."

Over the years I've gotten pretty good at making meals for my family, I'm proud to say. But what if my husband and I go away for a week, and I want to explain my system to my mom so she can fill in for me? Or what if my friend who has kids wants to know my methods? That's when I'd start to formalize my model, making it much more systematic and, in some sense, mathematical. And if I were feeling ambitious, I might put it into a computer program.

Ideally, the program would include all of the available food options, their nutritional value and cost, and a complete database of my family's tastes: each individual's preferences and aversions. It would be hard, though, to sit down and summon all that informationoff the top of my head. I've got loads of memories of people grabbing seconds of asparagus or avoiding the string beans. But they're all mixed up and hard to formalize in a comprehensive list.

The better solution would be to train the model over time, entering data every day on what I'd bought and cooked and noting the responses of each family member. I would also include parameters, or constraints. I might limit the fruits and vegetables to what's in season and dole out a certain amount of Pop-Tarts, but only enough to forestall an open rebellion. I also would add a number of rules. This one likes meat, this one likes bread and pasta, this one drinks lots of milk and insists on spreading Nutella on everything in sight.

If I made this work a major priority, over many months I might come up with a very good model. I would have turned the food management I keep in my head, my informal internal model, into a formal external one. In creating my model, I'd be extending my power and influence in the world. I'd be building an automated me that others can implement, even when I'm not around.

There would always be mistakes, however, because models are, by their very nature, simplifications. No model can include all of the real world's complexity or the nuance of human communication. Inevitably, some important information gets left out. I might have neglected to inform my model that junk-food rules are relaxed on birthdays, or that raw carrots are more popular than the cooked variety.

To create a model, then, we make choices about what's important enough to include, simplifying the world into a toy version that can be easily understood and from which we can infer important facts and actions. We expect it to handle only one job and accept that it will occasionally act like a clueless machine, one with enormous blind spots.

Sometimes these blind spots don't matter. When we ask Google Maps for directions, it models the world as a series of roads, tunnels, and bridges. It ignores the buildings, because they aren't relevant to the task. When avionics software guides an airplane, it models the wind, the speed of the plane, and the landing strip below, but not the streets, tunnels, buildings, and people.

A model's blind spots reflect the judgments and priorities of its creators. While the choices in Google Maps and avionics software appear cut and dried, others are far more problematic. The value-added model in Washington, D.C., schools, to return to that example, evaluates teachers largely on the basis of students' test scores, while ignoring how much the teachers engage the students, work on specific skills, deal with classroom management, or help students with personal and family problems. It's overly simple, sacrificing accuracy and insight for efficiency. Yet from the administrators' perspective it provides an effective tool to ferret out hundreds of apparently underperforming teachers, even at the risk of misreading some of them.

Here we see that models, despite their reputation for impartiality, reflect goals and ideology. When I removed the possibility of eating Pop-Tarts at every meal, I was imposing my ideology on the meals model. It's something we do without a second thought. Our own values and desires influence our choices, from the data we choose to collect to the questions we ask. Models are opinions embedded in mathematics.

Whether or not a model works is also a matter of opinion. After all, a key component of every model, whether formal or informal, is its definition of success. This is an important point that we'll return to as we explore the dark world of WMDs. In each case, we must ask not only who designed the model but also what that person or company is trying to accomplish. If the North Korean government built a model for my family's meals, for example, it might be optimized to keep us above the threshold of starvation at the lowest cost, based on the food stock available. Preferences would count for little or nothing. By contrast, if my kids were creating the model, success might feature ice cream at every meal. My own model attempts to blend a bit of the North Koreans' resource management with the happiness of my kids, along with my own priorities of health, convenience, diversity of experience, and sustainability. As a result, it's much more complex. But it still reflects my own personal reality. And a model built for today will work a bit worse tomorrow. It will grow stale if it's not constantly updated. Prices change, as do people's preferences. A model built for a six-year-old won't work for a teenager.

This is true of internal models as well. You can often see troubles when grandparents visit a grandchild they haven't seen for a while. On their previous visit, they gathered data on what the child knows, what makes her laugh, and what TV show she likes and (unconsciously) created a model for relating to this particular four-year-old. Upon meeting her a year later, they can suffer a few awkward hours because their models are out of date. Thomas the Tank Engine, it turns out, is no longer cool. It takes some time to gather new data about the child and adjust their models.

This is not to say that good models cannot be primitive. Some very effective ones hinge on a single variable. The most common model for detecting fires in a home or office weighs only one strongly correlated variable, the presence of smoke. That's usually enough. But modelers run into problems—or subject us to problems—when they focus models as simple as a smoke alarm on their fellow humans.

Racism, at the individual level, can be seen as a predictive model whirring away in billions of human minds around the world. It is built from faulty, incomplete, or generalized data. Whether it comes from experience or hearsay, the data indicates that certain types of people have behaved badly. That generates a binary prediction that all people of that race will behave that same way.

Needless to say, racists don't spend a lot of time hunting down reliable data to train their twisted models.

And once their model morphs into a belief, it becomes hardwired. It generates poisonous assumptions, yet rarely tests them, settling instead for data that seems to confirm and fortify them. Consequently, racism is the most slovenly of predictive models. It is powered by haphazard data gathering and spurious correlations, reinforced by institutional inequities, and polluted by confirmation bias. In this way, oddly enough, racism operates like many of the WMDs I’ll be describing in this book.

WEAPONS OF MATH DESTRUCTION BY CATHY O'NEIL PDF

Why must choose the inconvenience one if there is very easy? Get the profit by getting guide **Weapons Of Math Destruction By Cathy O'Neil** here. You will obtain various method making a bargain and also get guide Weapons Of Math Destruction By Cathy O'Neil As recognized, nowadays. Soft file of guides Weapons Of Math Destruction By Cathy O'Neil end up being incredibly popular with the viewers. Are you one of them? And right here, we are supplying you the extra collection of ours, the Weapons Of Math Destruction By Cathy O'Neil.

Obtaining the publications *Weapons Of Math Destruction By Cathy O'Neil* now is not kind of hard means. You could not just going for publication shop or collection or borrowing from your close friends to read them. This is an extremely basic way to precisely obtain guide by on the internet. This on-line publication Weapons Of Math Destruction By Cathy O'Neil can be one of the alternatives to accompany you when having downtime. It will not lose your time. Think me, the book will reveal you new point to read. Simply spend little time to open this on the internet e-book Weapons Of Math Destruction By Cathy O'Neil as well as read them wherever you are now.

Sooner you get guide Weapons Of Math Destruction By Cathy O'Neil, faster you can appreciate checking out guide. It will certainly be your rely on keep downloading and install guide Weapons Of Math Destruction By Cathy O'Neil in given web link. In this method, you can actually choose that is served to obtain your personal publication online. Here, be the initial to obtain guide qualified Weapons Of Math Destruction By Cathy O'Neil as well as be the first to recognize just how the writer indicates the notification as well as expertise for you.

WEAPONS OF MATH DESTRUCTION BY CATHY O'NEIL PDF

Longlisted for the National Book Award
New York Times Bestseller

A former Wall Street quant sounds an alarm on the mathematical models that pervade modern life — and threaten to rip apart our social fabric

We live in the age of the algorithm. Increasingly, the decisions that affect our lives—where we go to school, whether we get a car loan, how much we pay for health insurance—are being made not by humans, but by mathematical models. In theory, this should lead to greater fairness: Everyone is judged according to the same rules, and bias is eliminated.

But as Cathy O'Neil reveals in this urgent and necessary book, the opposite is true. The models being used today are opaque, unregulated, and uncontestable, even when they're wrong. Most troubling, they reinforce discrimination: If a poor student can't get a loan because a lending model deems him too risky (by virtue of his zip code), he's then cut off from the kind of education that could pull him out of poverty, and a vicious spiral ensues. Models are propping up the lucky and punishing the downtrodden, creating a "toxic cocktail for democracy." Welcome to the dark side of Big Data.

Tracing the arc of a person's life, O'Neil exposes the black box models that shape our future, both as individuals and as a society. These "weapons of math destruction" score teachers and students, sort résumés, grant (or deny) loans, evaluate workers, target voters, set parole, and monitor our health.

O'Neil calls on modelers to take more responsibility for their algorithms and on policy makers to regulate their use. But in the end, it's up to us to become more savvy about the models that govern our lives. This important book empowers us to ask the tough questions, uncover the truth, and demand change.

- Sales Rank: #24767 in Books
- Published on: 2016-09-06
- Original language: Italian
- Number of items: 1
- Dimensions: 8.50" h x .75" w x 5.31" l, .84 pounds
- Binding: Paperback

Review
A New York Times Book Review Notable Book of 2016
A Boston Globe Best Book of 2016
One of Wired's Required Reading Picks of 2016
One of Fortune's Favorite Books of 2016
A Kirkus Reviews Best Book of 2016
A Chicago Public Library Best Book of 2016
A Nature.com Best Book of 2016
An On Point Best Book of 2016

New York Times Editor's Choice
A Maclean's Bestseller
Winner of the 2016 SLA-NY PrivCo Spotlight Award

“O’Neil’s book offers a frightening look at how algorithms are increasingly regulating people… Her knowledge of the power and risks of mathematical models, coupled with a gift for analogy, makes her one of the most valuable observers of the continuing weaponization of big data… [She] does a masterly job explaining the pervasiveness and risks of the algorithms that regulate our lives.”
—New York Times Book Review

"Weapons of Math Destruction is the Big Data story Silicon Valley proponents won't tell… [It] pithily exposes flaws in how information is used to assess everything from creditworthiness to policing tactics… A thought-provoking read for anyone inclined to believe that data doesn't lie.”
—Reuters

“This is a manual for the 21st-century citizen, and it succeeds where other big data accounts have failed—it is accessible, refreshingly critical and feels relevant and urgent.”
—Financial Times

"Insightful and disturbing."
—New York Review of Books

“Weapons of Math Destruction is an urgent critique of… the rampant misuse of math in nearly every aspect of our lives.”
—Boston Globe

“A fascinating and deeply disturbing book.”
—Yuval Noah Harari, author of Sapiens; The Guardian’s Best Books of 2016

“Illuminating… [O’Neil] makes a convincing case that this reliance on algorithms has gone too far.”
—The Atlantic

“A nuanced reminder that big data is only as good as the people wielding it.”
—Wired

“If you’ve ever suspected there was something baleful about our deep trust in data, but lacked the mathematical skills to figure out exactly what it was, this is the book for you.”
—Salon

“O’Neil is an ideal person to write this book. She is an academic mathematician turned Wall Street quant turned data scientist who has been involved in Occupy Wall Street and recently started an algorithmic auditing company. She is one of the strongest voices speaking out for limiting the ways we allow algorithms to influence our lives… While Weapons of Math Destruction is full of hard truths and grim statistics, it is also accessible and even entertaining. O’Neil’s writing is direct and easy to read—I devoured it in an afternoon.”
—Scientific American

“Readable and engaging… succinct and cogent… Weapons of Math Destruction is The Jungle of our age… [It] should be required reading for all data scientists and for any organizational decision-maker convinced

that a mathematical model can replace human judgment."
—Mark Van Hollebeke, Data and Society: Points

“Indispensable… Despite the technical complexity of its subject, Weapons of Math Destruction lucidly guides readers through these complex modelling systems… O’Neil’s book is an excellent primer on the ethical and moral risks of Big Data and an algorithmically dependent world… For those curious about how Big Data can help them and their businesses, or how it has been reshaping the world around them, Weapons of Math Destruction is an essential starting place.”
—National Post

“Cathy O’Neil has seen Big Data from the inside, and the picture isn’t pretty. Weapons of Math Destruction opens the curtain on algorithms that exploit people and distort the truth while posing as neutral mathematical tools. This book is wise, fierce, and desperately necessary.”
—Jordan Ellenberg, University of Wisconsin-Madison, author of How Not To Be Wrong

“O’Neil has become [a whistle-blower] for the world of Big Data… [in] her important new book… Her work makes particularly disturbing points about how being on the wrong side of an algorithmic decision can snowball in incredibly destructive ways.”
—TIME

“O’Neil’s work is so important… [her] book is a vital crash-course in the specialized kind of statistical knowledge we all need to interrogate the systems around us and demand better.”
—Boing Boing

“Cathy O’Neil, a number theorist turned data scientist, delivers a simple but important message: Statistical models are everywhere, and they exert increasing power over many aspects of our daily lives… Weapons of Math Destruction provides a handy map to a few of the many areas of our lives over which invisible algorithms have gained some control. As the empire of big data continues to expand, Cathy O’Neil’s reminder of the need for vigilance is welcome and necessary.”
—American Prospect

“An avowed math nerd, O’Neil has written an engaging description of the effect of crunched data on our lives.”
—Hicklebee’s, San Francisco Chronicle

“By tracking how algorithms shape people's lives at every stage, O'Neil makes a compelling case that our bot overlords are using data to discriminate unfairly and foreclose democratic choices. If you work with data, or just produce reams of it online, this is a must-read.”
—ArsTechnica

“Lucid, alarming, and valuable… [O’Neil’s] writing is crisp and precise as she aims her arguments to a lay audience. This makes for a remarkably page-turning read for a book about algorithms. Weapons of Math Destruction should be required reading for anybody whose life will be affected by Big Data, which is to say: required reading for everyone. It’s a wake-up call – a journalistic heir to The Jungle and Silent Spring. Like those books, it should change the course of American society.”
—Aspen Times

"[O'Neil's] propulsive study reveals many models that are currently 'micromanaging' the US economy as opaque and riddled with bias."

—Nature

“You don’t need to be a nerd to appreciate the significance of [O’Neil’s] message… Weapons is a must-read for anyone who is working to combat economic and racial discrimination.”
—Goop

"Cathy O’Neil’s book... is important and covers issues everyone should care about. Bonus points: it’s accessible, compelling, and—something I wasn’t expecting—really fun to read.”
—Inside Higher Ed

“Often we don’t even know where to look for those important algorithms, because by definition the most dangerous ones are also the most secretive. That’s why the catalogue of case studies in O’Neil’s book are so important; she’s telling us where to look.”
—The Guardian

“O’Neil is passionate about exposing the harmful effects of Big Data–driven mathematical models (what she calls WMDs), and she’s uniquely qualified for the task… [She] makes a convincing case that many mathematical models today are engineered to benefit the powerful at the expense of the powerless… [and] has written an entertaining and timely book that gives readers the tools to cut through the ideological fog obscuring the dangers of the Big Data revolution.”
—In These Times

“In this simultaneously illuminating and disturbing account, [O’Neil] describes the many ways in which widely used mathematic models—based on ‘prejudice, misunderstanding, and bias’—tend to punish the poor and reward the rich… She convincingly argues for both more responsible modeling and federal regulation. An unusually lucid and readable look at the daunting algorithms that govern so many aspects of our lives.”
—Kirkus Reviews (starred)

“Even as a professional mathematician, I had no idea how insidious Big Data could be until I read Weapons of Math Destruction. Though terrifying, it’s a surprisingly fun read: O’Neil’s vision of a world run by algorithms is laced with dark humor and exasperation—like a modern-day Dr. Strangelove or Catch-22. It is eye-opening, disturbing, and deeply important.”
—Steven Strogatz, Cornell University, author of The Joy of x

“This taut and accessible volume, the stuff of technophobes’ nightmares, explores the myriad ways in which largescale data modeling has made the world a less just and equal place. O’Neil speaks from a place of authority on the subject… Unlike some other recent books on data collection, hers is not hysterical; she offers more of a chilly wake-up call as she walks readers through the ways the ‘big data’ industry has facilitated social ills such as skyrocketing college tuitions, policing based on racial profiling, and high unemployment rates in vulnerable communities… eerily prescient.”
—Publishers Weekly

"Well-written, entertaining and very valuable."
—Times Higher Education

"Not math heavy, but written in an exceedingly accessible, almost literary style; [O'Neil's] fascinating case studies of WMDs fit neatly into the genre of dystopian literature. There's a little Philip K. Dick, a little Orwell, a little Kafka in her portrait of powerful bureaucracies ceding control of the most intimate decisions of our lives to hyper-empowered computer models riddled with all of our unresolved, atavistic human

biases."
—Paris Review

“Through harrowing real-world examples and lively story-telling, Weapons of Math Destruction shines invaluable light on the invisible algorithms and complex mathematical models used by government and big business to undermine equality and increase private power. Combating secrecy with clarity and confusion with understanding, this book can help us change course before it’s too late.”
—Astra Taylor, author of The People’s Platform: Taking Back Power and Culture in the Digital Age

"Weapons of Math Destruction is a fantastic, plainspoken call to arms. It acknowledges that models aren't going away: As a tool for identifying people in difficulty, they are amazing. But as a tool for punishing and disenfranchising, they're a nightmare.”
—Cory Doctorow, author of Little Brother and co-editor of Boing Boing

“Many algorithms are slaves to the inequalities of power and prejudice. If you don’t want these algorithms to become your masters, read Weapons of Math Destruction by Cathy O’Neil to deconstruct the latest growing tyranny of an arrogant establishment.”
—Ralph Nader, author of Unsafe at Any Speed

“In this fascinating account, Cathy O'Neil leverages her expertise in mathematics and her passion for social justice to poke holes in the triumphant narrative of Big Data. She makes a compelling case that math is being used to squeeze marginalized segments of society and magnify inequities. Her analysis is superb, her writing is enticing, and her findings are unsettling.”
—danah boyd, founder of Data & Society and author of It’s Complicated

"From getting a job to finding a spouse, predictive algorithms are silently shaping and controlling our destinies. Cathy O'Neil takes us on a journey of outrage and wonder, with prose that makes you feel like it's just a conversation. But it’s an important one. We need to reckon with technology.”
—Linda Tirado, author of Hand to Mouth: Living in Bootstrap America

“Next time you hear someone gushing uncritically about the wonders of Big Data, show them Weapons of Math Destruction. It’ll be salutary.”
—Felix Salmon, Fusion

About the Author
Cathy O'Neil is a data scientist and author of the blog mathbabe.org. She earned a Ph.D. in mathematics from Harvard and taught at Barnard College before moving to the private sector, where she worked for the hedge fund D. E. Shaw. She then worked as a data scientist at various start-ups, building models that predict people’s purchases and clicks. O’Neil started the Lede Program in Data Journalism at Columbia and is the author of Doing Data Science. She is currently a columnist for Bloomberg View.

1

BOMB PARTS
What Is a Model?

It was a hot August afternoon in 1946. Lou Boudreau, the player-manager of the Cleveland Indians, was having a miserable day. In the first game of a doubleheader, Ted Williams had almost single-handedly annihilated his team. Williams, perhaps the game’s greatest hitter at the time, had smashed three home runs and driven home eight. The Indians ended up losing 11 to 10.

Boudreau had to take action. So when Williams came up for the first time in the second game, players on the Indians' side started moving around. Boudreau, the shortstop, jogged over to where the second baseman would usually stand, and the second baseman backed into short right field. The third baseman moved to his left, into the shortstop's hole. It was clear that Boudreau, perhaps out of desperation, was shifting the entire orientation of his defense in an attempt to turn Ted Williams's hits into outs.

In other words, he was thinking like a data scientist. He had analyzed crude data, most of it observational: Ted Williams usually hit the ball to right field. Then he adjusted. And it worked. Fielders caught more of Williams's blistering line drives than before (though they could do nothing about the home runs sailing over their heads).

If you go to a major league baseball game today, you'll see that defenses now treat nearly every player like Ted Williams. While Boudreau merely observed where Williams usually hit the ball, managers now know precisely where every player has hit every ball over the last week, over the last month, throughout his career, against left-handers, when he has two strikes, and so on. Using this historical data, they analyze their current situation and calculate the positioning that is associated with the highest probability of success. And that sometimes involves moving players far across the field.

Shifting defenses is only one piece of a much larger question: What steps can baseball teams take to maximize the probability that they'll win? In their hunt for answers, baseball statisticians have scrutinized every variable they can quantify and attached it to a value. How much more is a double worth than a single? When, if ever, is it worth it to bunt a runner from first to second base?

The answers to all of these questions are blended and combined into mathematical models of their sport. These are parallel universes of the baseball world, each a complex tapestry of probabilities. They include every measurable relationship among every one of the sport's components, from walks to home runs to the players themselves. The purpose of the model is to run different scenarios at every juncture, looking for the optimal combinations. If the Yankees bring in a right-handed pitcher to face Angels slugger Mike Trout, as compared to leaving in the current pitcher, how much more likely are they to get him out? And how will that affect their overall odds of winning?

Baseball is an ideal home for predictive mathematical modeling. As Michael Lewis wrote in his 2003 bestseller, Moneyball, the sport has attracted data nerds throughout its history. In decades past, fans would pore over the stats on the back of baseball cards, analyzing Carl Yastrzemski's home run patterns or comparing Roger Clemens's and Dwight Gooden's strikeout totals. But starting in the 1980s, serious statisticians started to investigate what these figures, along with an avalanche of new ones, really meant: how they translated into wins, and how executives could maximize success with a minimum of dollars.

"Moneyball" is now shorthand for any statistical approach in domains long ruled by the gut. But baseball represents a healthy case study—and it serves as a useful contrast to the toxic models, or WMDs, that are popping up in so many areas of our lives. Baseball models are fair, in part, because they're transparent. Everyone has access to the stats and can understand more or less how they're interpreted. Yes, one team's model might give more value to home run hitters, while another might discount them a bit, because sluggers tend to strike out a lot. But in either case, the numbers of home runs and strikeouts are there for everyone to see.

Baseball also has statistical rigor. Its gurus have an immense data set at hand, almost all of it directly related to the performance of players in the game. Moreover, their data is highly relevant to the outcomes they are trying to predict. This may sound obvious, but as we'll see throughout this book, the folks building WMDs routinely lack data for the behaviors they're most interested in. So they substitute stand-in data, or proxies. They draw statistical correlations between a person's zip code or language patterns and her potential to pay back a loan or handle a job. These correlations are discriminatory, and some of them are illegal. Baseball models, for the most part, don't use proxies because they use pertinent inputs like balls, strikes, and hits.

Most crucially, that data is constantly pouring in, with new statistics from an average of twelve or thirteen games arriving daily from April to October. Statisticians can compare the results of these games to the predictions of their models, and they can see where they were wrong. Maybe they predicted that a left-handed reliever would give up lots of hits to right-handed batters—and yet he mowed them down. If so, the stats team has to tweak their model and also carry out research on why they got it wrong. Did the pitcher's new screwball affect his statistics? Does he pitch better at night? Whatever they learn, they can feed back into the model, refining it. That's how trustworthy models operate. They maintain a constant back-and-forth with whatever in the world they're trying to understand or predict. Conditions change, and so must the model.

Now, you may look at the baseball model, with its thousands of changing variables, and wonder how we could even be comparing it to the model used to evaluate teachers in Washington, D.C., schools. In one of them, an entire sport is modeled in fastidious detail and updated continuously. The other, while cloaked in mystery, appears to lean heavily on a handful of test results from one year to the next. Is that really a model?

The answer is yes. A model, after all, is nothing more than an abstract representation of some process, be it a baseball game, an oil company's supply chain, a foreign government's actions, or a movie theater's attendance. Whether it's running in a computer program or in our head, the model takes what we know and uses it to predict responses in various situations. All of us carry thousands of models in our heads. They tell us what to expect, and they guide our decisions.

Here's an informal model I use every day. As a mother of three, I cook the meals at home—my husband, bless his heart, cannot remember to put salt in pasta water. Each night when I begin to cook a family meal, I internally and intuitively model everyone's appetite. I know that one of my sons loves chicken (but hates hamburgers), while another will eat only the pasta (with extra grated parmesan cheese). But I also have to take into account that people's appetites vary from day to day, so a change can catch my model by surprise. There's some unavoidable uncertainty involved.

The input to my internal cooking model is the information I have about my family, the ingredients I have on hand or I know are available, and my own energy, time, and ambition. The output is how and what I decide to cook. I evaluate the success of a meal by how satisfied my family seems at the end of it, how much they've eaten, and how healthy the food was. Seeing how well it is received and how much of it is enjoyed allows me to update my model for the next time I cook. The updates and adjustments make it what statisticians call a "dynamic model."

Over the years I've gotten pretty good at making meals for my family, I'm proud to say. But what if my husband and I go away for a week, and I want to explain my system to my mom so she can fill in for me? Or what if my friend who has kids wants to know my methods? That's when I'd start to formalize my model, making it much more systematic and, in some sense, mathematical. And if I were feeling ambitious, I might put it into a computer program.

Ideally, the program would include all of the available food options, their nutritional value and cost, and a complete database of my family's tastes: each individual's preferences and aversions. It would be hard, though, to sit down and summon all that informationoff the top of my head. I've got loads of memories of people grabbing seconds of asparagus or avoiding the string beans. But they're all mixed up and hard to formalize in a comprehensive list.

The better solution would be to train the model over time, entering data every day on what I'd bought and cooked and noting the responses of each family member. I would also include parameters, or constraints. I might limit the fruits and vegetables to what's in season and dole out a certain amount of Pop-Tarts, but only enough to forestall an open rebellion. I also would add a number of rules. This one likes meat, this one likes bread and pasta, this one drinks lots of milk and insists on spreading Nutella on everything in sight.

If I made this work a major priority, over many months I might come up with a very good model. I would have turned the food management I keep in my head, my informal internal model, into a formal external one. In creating my model, I'd be extending my power and influence in the world. I'd be building an automated me that others can implement, even when I'm not around.

There would always be mistakes, however, because models are, by their very nature, simplifications. No model can include all of the real world's complexity or the nuance of human communication. Inevitably, some important information gets left out. I might have neglected to inform my model that junk-food rules are relaxed on birthdays, or that raw carrots are more popular than the cooked variety.

To create a model, then, we make choices about what's important enough to include, simplifying the world into a toy version that can be easily understood and from which we can infer important facts and actions. We expect it to handle only one job and accept that it will occasionally act like a clueless machine, one with enormous blind spots.

Sometimes these blind spots don't matter. When we ask Google Maps for directions, it models the world as a series of roads, tunnels, and bridges. It ignores the buildings, because they aren't relevant to the task. When avionics software guides an airplane, it models the wind, the speed of the plane, and the landing strip below, but not the streets, tunnels, buildings, and people.

A model's blind spots reflect the judgments and priorities of its creators. While the choices in Google Maps and avionics software appear cut and dried, others are far more problematic. The value-added model in Washington, D.C., schools, to return to that example, evaluates teachers largely on the basis of students' test scores, while ignoring how much the teachers engage the students, work on specific skills, deal with classroom management, or help students with personal and family problems. It's overly simple, sacrificing accuracy and insight for efficiency. Yet from the administrators' perspective it provides an effective tool to ferret out hundreds of apparently underperforming teachers, even at the risk of misreading some of them.

Here we see that models, despite their reputation for impartiality, reflect goals and ideology. When I removed the possibility of eating Pop-Tarts at every meal, I was imposing my ideology on the meals model. It's something we do without a second thought. Our own values and desires influence our choices, from the data we choose to collect to the questions we ask. Models are opinions embedded in mathematics.

Whether or not a model works is also a matter of opinion. After all, a key component of every model, whether formal or informal, is its definition of success. This is an important point that we'll return to as we explore the dark world of WMDs. In each case, we must ask not only who designed the model but also what that person or company is trying to accomplish. If the North Korean government built a model for my family's meals, for example, it might be optimized to keep us above the threshold of starvation at the lowest cost, based on the food stock available. Preferences would count for little or nothing. By contrast, if my kids were creating the model, success might feature ice cream at every meal. My own model attempts to blend a bit of the North Koreans' resource management with the happiness of my kids, along with my own priorities of health, convenience, diversity of experience, and sustainability. As a result, it's much more complex. But it still reflects my own personal reality. And a model built for today will work a bit worse tomorrow. It will grow stale if it's not constantly updated. Prices change, as do people's preferences. A model built for a six-year-old won't work for a teenager.

This is true of internal models as well. You can often see troubles when grandparents visit a grandchild they haven't seen for a while. On their previous visit, they gathered data on what the child knows, what makes her laugh, and what TV show she likes and (unconsciously) created a model for relating to this particular four-year-old. Upon meeting her a year later, they can suffer a few awkward hours because their models are out of date. Thomas the Tank Engine, it turns out, is no longer cool. It takes some time to gather new data about the child and adjust their models.

This is not to say that good models cannot be primitive. Some very effective ones hinge on a single variable. The most common model for detecting fires in a home or office weighs only one strongly correlated variable, the presence of smoke. That's usually enough. But modelers run into problems—or subject us to problems—when they focus models as simple as a smoke alarm on their fellow humans.

Racism, at the individual level, can be seen as a predictive model whirring away in billions of human minds around the world. It is built from faulty, incomplete, or generalized data. Whether it comes from experience or hearsay, the data indicates that certain types of people have behaved badly. That generates a binary prediction that all people of that race will behave that same way.

Needless to say, racists don't spend a lot of time hunting down reliable data to train their twisted models. And once their model morphs into a belief, it becomes hardwired. It generates poisonous assumptions, yet rarely tests them, settling instead for data that seems to confirm and fortify them. Consequently, racism is the most slovenly of predictive models. It is powered by haphazard data gathering and spurious correlations, reinforced by institutional inequities, and polluted by confirmation bias. In this way, oddly enough, racism operates like many of the WMDs I'll be describing in this book.

Most helpful customer reviews

184 of 193 people found the following review helpful.
Stop Using Math as a Weapon
By Amazon Customer
So here you are on Amazon's web page, reading about Cathy O'Neil's new book, Weapons of Math Destruction. Amazon hopes you buy the book (and so do I, it's great!). But Amazon also hopes it can sell you some other books while you're here. That's why, in a prominent place on the page, you see a section entitled:

Customers Who Bought This Item Also Bought

This section is Amazon's way of using what it knows -- which book you're looking at, and sales data collected across all its customers -- to recommend other books that you might be interested in. It's a very simple, and successful, example of a predictive model: data goes in, some computation happens, a prediction comes out. What makes this a good model? Here are a few things:

1. It uses relevant input data.The goal is to get people to buy books, and the input to the model is what books people buy. You can't expect to get much more relevant than that.

2. It's transparent. You know exactly why the site is showing you these particular books, and if the system recommends a book you didn't expect, you have a pretty good idea why. That means you can make an informed decision about whether or not to trust the recommendation.
3. There's a clear measure of success and an embedded feedback mechanism. Amazon wants to sell books. The model succeeds if people click on the books they're shown, and, ultimately, if they buy more books, both of which are easy to measure. If clicks on or sales of related items go down, Amazon will know, and can investigate and adjust the model accordingly.

Weapons of Math Destruction reviews, in an accessible, non-technical way, what makes models effective -- or not. The emphasis, as you might guess from the title, is on models with problems. The book highlights many important ideas; here are just a few:

1. Models are more than just math. Take a look at Amazon's model above: while there are calculations (simple ones) embedded, it's people who decide what data to use, how to use it, and how to measure success. Math is not a final arbiter, but a tool to express, in a scalable (i.e., computable) way, the values that people explicitly decide to emphasize. Cathy says that "models are opinions expressed in mathematics" (or computer code). She highlights that when we evaluate teachers based on students' test scores, or assess someone's insurability as a driver based on their credit record, we are expressing opinions: that a successful teacher should boost test scores, or that responsible bill-payers are more likely to be responsible drivers.

2. Replacing what you really care about with what you can easily get your hands on can get you in trouble. In Amazon's recommendation model, we want to predict book sales, and we can use book sales as inputs; that's a good thing. But what if you can't directly measure what you're interested in? In the early 1980's, the magazine US News wanted to report on college quality. Unable to measure quality directly, the magazine built a model based on proxies, primarily outward markers of success, like selectivity and alumni giving. Predictably, college administrators, eager to boost their ratings, focused on these markers rather than on education quality itself. For example, to boost selectivity, they encouraged more students, even unqualified ones, to apply. This is an example of gaming the model.

3. Historical data is stuck in the past. Typically, predictive models use past history to predict future behavior. This can be problematic when part of the intention of the model is to break with the past. To take a very simple example, imagine that Cathy is about to publish a sequel to Weapons of Math Destruction. If Amazon uses only purchase data, the Customers Who Bought This Also Bought list would completely miss the connection between the original and the sequel. This means that if we don't want the future to look just like the past, our models need to use more than just history as inputs. A chapter about predictive models in hiring is largely devoted to this idea. A company may think that its past, subjective hiring system overlooks qualified candidates, but if it replaces the HR department with a model that sifts through resumes based only on the records of past hires, it may just be codifying (pun intended) past practice. A related idea is that, in this case, rather than adding objectivity, the model becomes a shield that hides discrimination. This takes us back to Models are more than just math and also leads to the next point:

4. Transparency matters! If a book you didn't expect shows up on The Customers Who Bought This Also Bought list, it's pretty easy for Amazon to check if it really belongs there. The model is pretty easy to understand and audit, which builds confidence and also decreases the likelihood that it gets used to obfuscate. An example of a very different story is the value added model for teachers, which evaluates teachers through their students' standardized test scores. Among its other drawbacks, this model is especially opaque in practice, both because of its complexity and because many implementations are built by outsiders. Models need to be openly assessed for effectiveness, and when teachers receive bad scores without knowing why, or when a single teacher's score fluctuates dramatically from year to year without explanation, it's hard

to have any faith in the process.

5. Models don't just measure reality, but sometimes amplify it, or create their own. Put another way, models of human behavior create feedback loops, often becoming self-fulfilling prophecies. There are many examples of this in the book, especially focusing on how models can amplify economic inequality. To take one example, a company in the center of town might notice that workers with longer commutes tend to turn over more frequently, and adjust its hiring model to focus on job candidates who can afford to live in town. This makes it easier for wealthier candidates to find jobs than poorer ones, and perpetuates a cycle of inequality. There are many other examples: predictive policing, prison sentences based on recidivism, e-scores for credit. Cathy talks about a trade-off between efficiency and fairness, and, as you can again guess from the title, argues for fairness as an explicit value in modeling.

Weapons of Math Destruction is not a math book, and it is not investigative journalism. It is short -- you can read it in an afternoon -- and it doesn't have time or space for either detailed data analysis (there are no formulas or graphs) or complete histories of the models she considers. Instead, Cathy sketches out the models quickly, perhaps with an individual anecdote or two thrown in, so she can get to the main point -- getting people, especially non-technical people, used to questioning models. As more and more aspects of our lives fall under the purview of automated data analysis, that's a hugely important undertaking.

7 of 7 people found the following review helpful.
Very clear, but over-reliant on government solutions instead of more choices for consumers (competition!)
By David Zetland
I was excited to read this book as soon as I heard Cathy O'Neill, the author, interviewed on EconTalk.

O'Neill's hypothesis is that algorithms and machine learning can be useful, but they can also be destructive if they are (1) opaque, (2) scalable and (3) damaging. Put differently, an algorithm that determines whether you should be hired or fired, given a loan or able to retire on your savings is a WMD if it is opaque to users, "beneficiaries" and the public, has an impact on a large group of people at once, and "makes decisions" that have large social, financial or legal impacts. WMDs can leave thousands in jail or bankrupt pensions, often without warning or remorse.

As examples of non-WMDs, consider bitcoin/blockchain (the code and transactions are published), algorithms developed by a teacher (small scale), and Amazon's "recommended" lists, which are not damaging (because customers can decide to buy or not).

As examples of WMDs (many of which are explained in the book), consider Facebook's "newsfeed" algorithm, which is opaque (based on their internal advertising model), scaled (1.9 billion disenfranchised zombies) and damaging (echo-chamber, anyone?)

I took numerous notes while reading this book, which I think everyone interested in the rising power of "big data" (or big brother) or bureaucratic processes should read, but I will only highlight a few:

* Models are imperfect -- and dangerous if they are given too much "authority" (as I've said)
* Good systems use feedback to improve in transparent ways (they are anti-WMDs)
WMDs punish the poor because the rich can afford "custom" systems that are additionally mediated by professionals (lawyers, accountants, teachers)
* Models are more dangerous the more removed their data are from the topic of interest, e.g., models of "teacher effectiveness" based on "student grades" (or worse alumni salaries)
* "Models are opinions embedded in mathematics" (what I said) which means that those weak in math will

suffer more. That matters when "American adults... are literally the worst [at solving digital problems] in the developed world."
* It is easy for a "neutral" variable (e.g., postal code) to reproduce a biased variable (e.g., race)
* Wall Street is excellent at scaling up a bad idea, leading to huge financial losses (and taxpayer bailouts). It was not an accident that Wall Street "messed up." They knew that profits were private but losses social.
* Many for-profit colleges use online advertisements to attract (and rip off) the most vulnerable -- leaving them in debt and/or taxpayers with the bill. Sad.
* A good program (for education or crime prevention) also relies on qualitative factors that are hard to code into algorithms. Ignore those and you're likely to get a biased WMD. I just saw a documentary on urbanism that asked "what do the poor want -- hot water or a bathtub?" They wanted a bathtub because they had never had one and could not afford to heat water. #checkyourbias
* At some points in this book, I disagreed with O'Neill's preference for justice over efficiency. She does not want to allow employers to look at job applicants' credit histories because "hardworking people might lose jobs." Yes, that's true, but I can see why employers are willing to lose a few good people to avoid a lot of bad people, especially if they have lots of remaining (good credit) applicants. Should this happen at the government level? Perhaps not, but I don't see why a hotel chain cannot do this: the scale is too small to be a WMD.
* I did, OTOH, notice that peer-to-peer lending might be biased against lender like me (I use Lending Club, which sucks) who rely on their "public credit models" as it seems that these models are badly calibrated, leaving retail suckers like me to lose money while institutional borrowers are given preferential access.
* O'Neill's worries about injustice go a little too far in her counterexamples of the "safe driver who needs to drive through a dangerous neighborhood at 2am" as not deserving to face higher insurance prices, etc. I agree that this person may deserve a break, but the solution to this "unfair pricing" is not a ban on such price discrimination but an increase in competition, which has a way of separating safe and unsafe drivers (it's called a "separating equilibrium" in economics). Her fear of injustice makes me think that she's perhaps missing the point. High driving insurance rates are not a blow against human rights, even if they capture an imperfect measure of risk, because driving itself is not a human right. Yes, I know it's tough to live without a car in many parts of the US, but people suffering in those circumstances need to think bigger about maybe moving to a better place.
* Worried about bias in advertisements? Just ban all of them.
* O'Neill occasionally makes some false claims, e.g., that US employers offered health insurance as a perk to attract scarce workers during WWII. That was mainly because of a government-ordered wage freeze that incentivised firms to offer "more money" via perks. In any case, it would be good to look at how other countries run their health systems (I love the Dutch system) before blaming all US failures on WMDs.
* I'm sympathetic to the lies and distortions that Facebook and other social media spread (with the help of WMDs), but I've gotta give Trump credit for blowing up all the careful attempts to corral, control and manipulate what people see or think (but maybe he had a better way to manipulate). Trump has shown that people are willing to ignore facts to the point where it might take a real WMD blowing up in their neighborhood to take them off auto pilot.
* When it comes to political manipulations, I worry less about WMDs than the total lack of competition due to gerrymandering. In the 2016 election, 97 percent of representatives were re-elected to the House.
* Yes, I agree that humans are better at finding and using nuances, but those will be overshadowed as long as there's a profit (or election) to win. * * * Can we push back on those problems? Yes, if we realize how our phones are tracking us, how GPA is not your career, or how "the old boys network" actually produced a useful mix of perspectives.
* Businesses will be especially quick to temper their enthusiasm when they notice that WMDs are not nearly so clever. What worries me more are politicians or bureaucrats who believe a salesman pitching a WMD that will save them time but harm citizens. That's how we got dumb do not fly lists, and other assorted government failures.

* Although I do not put as much faith in "government regulation" as a solution to this problem as I put into competition, I agree with O'Neill that consumers should own their data and companies only get access to it on an opt-in model, but that model will be broken for as long as the EULA requires that you give up lots of data in exchange for access to the "free" platform. Yes, Facebook is handy, but do you want Facebook listening to your phone all the time?

Bottom Line: I give this book FOUR STARS for its well written, enlightening expose of MWDs. I would have preferred less emphasis on bureaucratic solutions and more on market, competition, and property rights solutions.

313 of 354 people found the following review helpful.
"They back up their analysis with reams of statistics, which give them the studied air of evenhanded science."
By CodeMaster Talon
I struggled with the star rating for this book. There are certainly aspects of the work that merit five stars. And it is VERY thought-provoking, like a good book should be. But there are flaws, significant ones, with the biggest flaw being a glaring over-simplification of many of the systems that O'Neil decries in the book. I don't know if O'Neil has personally ever had to take a psychology test to get a job, worked under the Kronos scheduling system, lived in a neighborhood with increased police presence due to crime rates, been victimized by insurance rates adjusted to zip codes, and endured corporate wellness programs. But all of those things have been a part of my life for years, and even I have to admit the many positive aspects of some of these systems. A few examples:

--Kronos. Despised by the rank and file of companies that I've worked for, Kronos software contains many aspects and automates things that previously were done by people, mostly managers. I hated it, but I have to admit overall it made things more fair. Why? Well, say you have a workplace policy that mandates chronically-late employees be written up for tardiness and eventually fired if they don't shape up. What tended to happen at multiple companies I worked for was that managers would look the other way when their buddies were tardy, and write up people they didn't like. Kronos changed that, because the system automatically generated write-ups for any employee that clocked in late too many times. Kronos has no buddies. Popular, habitually-late people suffered, but it was more "fair" in the true sense of the word. Some systems, like Kronos, have both aspects that level the playing field and aspects (like increased scheduling "efficiency") that can victimize workers. O'Neil tends to harp on the negative only, and if you have not personally seen both sides of system, you might not realize there was another side at all.

--Increased police presence in high-crime areas. This one really grated me the wrong way. O'Neil positions this as something that victimizes the poor. Well I have been poor, or at least this country's version of it, and I have lived in very high crime areas where if you didn't shut your window at night chances were good you would hear a murder. And believe me when I say I was DEEPLY grateful for the increased police presence. But then, I wasn't committing crimes. Now I live in a very wealthy neighborhood (though I am not wealthy) where I have not seen a single police car drive down my street in the past four months. O'Neil argues that many crimes, like drug use by affluent college students, go unpunished because the police are busy in the poorer neighborhoods. I agree, but police resources are limited and for mercy's sake they should be sent where people are being killed, not where a college student is passed out in his living room. My current neighbors many be committing as many crimes as O'Neil implies, but I'm not terrified to walk down the street, so I don't mind the lack of police presence. I know officers have to go deal with the more life-threatening stuff, and I am grateful to them. It all depends on your perspective.

--Corporate Wellness programs. These programs have never done anything for me except shower me with gift cards and educate me on behavioral sleep therapy. I love them. But, again, perspective. I am not overweight, I love to work out, and I eat healthy. The programs were a source of income for me and my family when we needed it most. I just would have liked acknowledgement that wellness programs really do have benefits for some people, instead of a chapter painting them as some sort of 1984-style nightmare where we are all forced to be thin. It's more complicated than that.

--And the best for last: The psychology tests. Those things are pretty bad. Despite winning multiple Employee and Student of the Year awards in my life, I can't pass those tests. Not to save my life. I didn't think much of it, until I heard about another star employee how couldn't pass them either. Then I met a third star employee (and I am talking about an employee who won two JD Power Awards in two years) who couldn't pass them. Why? Picture holding a hundred quarters in your hands and then throwing them at a wall. Some will go off to one side and some to another, but most will probably cluster in the middle. Those tests keep the quarters in the middle, weeding out people who aren't typical. Sometimes that's good (deadbeats) sometimes that's bad (talented employees that think different). Here O'Neil misses an opportunity to convince owners of companies that the tests can cost them highly desirable employees. Offering real, concrete ideas of how the tests could be improved to benefit both workers and company owners would have been a harder book to write, but a much more useful one.

A lot of the ominous implications made in the book have to do with what MIGHT happen in the future, if certain systems become more common. O'Neil often uses blanket statements to imply that certain outcomes are inevitable, but that is far from true. Irritate enough people, and the systems change. Legal challenges are made and won. Some companies, eager to lure star workers, throw out some of the more punishing aspects of commonly-used systems (that happened at a company where I worked, where "life-style" scheduling that forbid clopening and gave you two days off in a row was used in conjunction with Kronos. Worked great, people loved it.). The biggest weapon against abuse is, as O'Neil repeatedly states, transparency. Having been in the industry that creates these algorithms, she is in a unique position to expose the finer details of how they work. But the book is short on the kind of details I personally crave and long on blanket statements and generalizations, the same kind of generalizations she denounces companies for making. Not all automated systems victimize the poor, not even the ones spotlighted in this book. I know because I lived them and I was poor.

I hovered on the edge of a four star rating for this book, until a chance conversation with a Japanese woman a couple days ago. Her grandmother had lost most of her possessions and land after World War II because of land redistribution. My friend was not complaining, she thought the reforms overall a good thing, though her family had lost a lot from it. "Something may benefit 99 people of of 100," she told me,"but there's always that one person...". Exactly. There's always that one person. These systems that have come to permeate our culture need to be tweaked to minimize injustice. Unlawful algorithms need to be outlawed. Bad ideas need to be replaced with good ones. And Cathy O'Neil does discuss this, especially in the conclusion, but for me the focus of the book wasn't on target. It was too slanted against systems I have seen both harm AND help. It over-simplified issues, at least for me. It's a mess out there, and solutions that work for everyone wickedly hard to come by.

Because there's always that one person.

GRADE: B-
Interesting side-note: In Greek Mythology, "Kronos" is the name of the Titan who devoured his own children. My co-workers always found that hilarious.

See all 210 customer reviews...

WEAPONS OF MATH DESTRUCTION BY CATHY O'NEIL PDF

It will certainly believe when you are going to select this publication. This motivating **Weapons Of Math Destruction By Cathy O'Neil** e-book can be read entirely in certain time relying on how often you open and also review them. One to bear in mind is that every e-book has their own manufacturing to obtain by each reader. So, be the great visitor as well as be a far better individual after reading this book Weapons Of Math Destruction By Cathy O'Neil

Review
A New York Times Book Review Notable Book of 2016
A Boston Globe Best Book of 2016
One of Wired's Required Reading Picks of 2016
One of Fortune's Favorite Books of 2016
A Kirkus Reviews Best Book of 2016
A Chicago Public Library Best Book of 2016
A Nature.com Best Book of 2016
An On Point Best Book of 2016
New York Times Editor's Choice
A Maclean's Bestseller
Winner of the 2016 SLA-NY PrivCo Spotlight Award

“O’Neil’s book offers a frightening look at how algorithms are increasingly regulating people… Her knowledge of the power and risks of mathematical models, coupled with a gift for analogy, makes her one of the most valuable observers of the continuing weaponization of big data… [She] does a masterly job explaining the pervasiveness and risks of the algorithms that regulate our lives.”
—New York Times Book Review

"Weapons of Math Destruction is the Big Data story Silicon Valley proponents won't tell… [It] pithily exposes flaws in how information is used to assess everything from creditworthiness to policing tactics… A thought-provoking read for anyone inclined to believe that data doesn't lie.”
—Reuters

“This is a manual for the 21st-century citizen, and it succeeds where other big data accounts have failed—it is accessible, refreshingly critical and feels relevant and urgent.”
—Financial Times

"Insightful and disturbing."
—New York Review of Books

“Weapons of Math Destruction is an urgent critique of… the rampant misuse of math in nearly every aspect of our lives.”
—Boston Globe

“A fascinating and deeply disturbing book.”
—Yuval Noah Harari, author of Sapiens; The Guardian’s Best Books of 2016

“Illuminating… [O’Neil] makes a convincing case that this reliance on algorithms has gone too far.”
—The Atlantic

“A nuanced reminder that big data is only as good as the people wielding it.”
—Wired

“If you’ve ever suspected there was something baleful about our deep trust in data, but lacked the mathematical skills to figure out exactly what it was, this is the book for you.”
—Salon

“O’Neil is an ideal person to write this book. She is an academic mathematician turned Wall Street quant turned data scientist who has been involved in Occupy Wall Street and recently started an algorithmic auditing company. She is one of the strongest voices speaking out for limiting the ways we allow algorithms to influence our lives… While Weapons of Math Destruction is full of hard truths and grim statistics, it is also accessible and even entertaining. O’Neil’s writing is direct and easy to read—I devoured it in an afternoon.”
—Scientific American

“Readable and engaging… succinct and cogent… Weapons of Math Destruction is The Jungle of our age… [It] should be required reading for all data scientists and for any organizational decision-maker convinced that a mathematical model can replace human judgment."
—Mark Van Hollebeke, Data and Society: Points

“Indispensable… Despite the technical complexity of its subject, Weapons of Math Destruction lucidly guides readers through these complex modelling systems… O’Neil’s book is an excellent primer on the ethical and moral risks of Big Data and an algorithmically dependent world… For those curious about how Big Data can help them and their businesses, or how it has been reshaping the world around them, Weapons of Math Destruction is an essential starting place.”
—National Post

“Cathy O’Neil has seen Big Data from the inside, and the picture isn’t pretty. Weapons of Math Destruction opens the curtain on algorithms that exploit people and distort the truth while posing as neutral mathematical tools. This book is wise, fierce, and desperately necessary.”
—Jordan Ellenberg, University of Wisconsin-Madison, author of How Not To Be Wrong

“O’Neil has become [a whistle-blower] for the world of Big Data… [in] her important new book… Her work makes particularly disturbing points about how being on the wrong side of an algorithmic decision can snowball in incredibly destructive ways.”
—TIME

“O’Neil’s work is so important… [her] book is a vital crash-course in the specialized kind of statistical knowledge we all need to interrogate the systems around us and demand better.”
—Boing Boing

“Cathy O’Neil, a number theorist turned data scientist, delivers a simple but important message: Statistical models are everywhere, and they exert increasing power over many aspects of our daily lives… Weapons of Math Destruction provides a handy map to a few of the many areas of our lives over which invisible algorithms have gained some control. As the empire of big data continues to expand, Cathy O’Neil’s reminder of the need for vigilance is welcome and necessary.”

—American Prospect

“An avowed math nerd, O’Neil has written an engaging description of the effect of crunched data on our lives.”
—Hicklebee’s, San Francisco Chronicle

“By tracking how algorithms shape people's lives at every stage, O'Neil makes a compelling case that our bot overlords are using data to discriminate unfairly and foreclose democratic choices. If you work with data, or just produce reams of it online, this is a must-read.”
—ArsTechnica

“Lucid, alarming, and valuable… [O’Neil’s] writing is crisp and precise as she aims her arguments to a lay audience. This makes for a remarkably page-turning read for a book about algorithms. Weapons of Math Destruction should be required reading for anybody whose life will be affected by Big Data, which is to say: required reading for everyone. It’s a wake-up call – a journalistic heir to The Jungle and Silent Spring. Like those books, it should change the course of American society.”
—Aspen Times

"[O'Neil's] propulsive study reveals many models that are currently 'micromanaging' the US economy as opaque and riddled with bias."
—Nature

“You don’t need to be a nerd to appreciate the significance of [O’Neil’s] message… Weapons is a must-read for anyone who is working to combat economic and racial discrimination.”
—Goop

"Cathy O’Neil’s book... is important and covers issues everyone should care about. Bonus points: it’s accessible, compelling, and—something I wasn’t expecting—really fun to read.”
—Inside Higher Ed

“Often we don’t even know where to look for those important algorithms, because by definition the most dangerous ones are also the most secretive. That’s why the catalogue of case studies in O’Neil’s book are so important; she’s telling us where to look.”
—The Guardian

“O’Neil is passionate about exposing the harmful effects of Big Data–driven mathematical models (what she calls WMDs), and she’s uniquely qualified for the task… [She] makes a convincing case that many mathematical models today are engineered to benefit the powerful at the expense of the powerless… [and] has written an entertaining and timely book that gives readers the tools to cut through the ideological fog obscuring the dangers of the Big Data revolution.”
—In These Times

“In this simultaneously illuminating and disturbing account, [O’Neil] describes the many ways in which widely used mathematic models—based on ‘prejudice, misunderstanding, and bias’—tend to punish the poor and reward the rich… She convincingly argues for both more responsible modeling and federal regulation. An unusually lucid and readable look at the daunting algorithms that govern so many aspects of our lives.”
—Kirkus Reviews (starred)

“Even as a professional mathematician, I had no idea how insidious Big Data could be until I read Weapons

of Math Destruction. Though terrifying, it’s a surprisingly fun read: O’Neil’s vision of a world run by algorithms is laced with dark humor and exasperation—like a modern-day Dr. Strangelove or Catch-22. It is eye-opening, disturbing, and deeply important.”
—Steven Strogatz, Cornell University, author of The Joy of x

“This taut and accessible volume, the stuff of technophobes’ nightmares, explores the myriad ways in which largescale data modeling has made the world a less just and equal place. O’Neil speaks from a place of authority on the subject… Unlike some other recent books on data collection, hers is not hysterical; she offers more of a chilly wake-up call as she walks readers through the ways the ‘big data’ industry has facilitated social ills such as skyrocketing college tuitions, policing based on racial profiling, and high unemployment rates in vulnerable communities… eerily prescient.”
—Publishers Weekly

"Well-written, entertaining and very valuable."
—Times Higher Education

"Not math heavy, but written in an exceedingly accessible, almost literary style; [O'Neil's] fascinating case studies of WMDs fit neatly into the genre of dystopian literature. There's a little Philip K. Dick, a little Orwell, a little Kafka in her portrait of powerful bureaucracies ceding control of the most intimate decisions of our lives to hyper-empowered computer models riddled with all of our unresolved, atavistic human biases."
—Paris Review

“Through harrowing real-world examples and lively story-telling, Weapons of Math Destruction shines invaluable light on the invisible algorithms and complex mathematical models used by government and big business to undermine equality and increase private power. Combating secrecy with clarity and confusion with understanding, this book can help us change course before it’s too late.”
—Astra Taylor, author of The People’s Platform: Taking Back Power and Culture in the Digital Age

"Weapons of Math Destruction is a fantastic, plainspoken call to arms. It acknowledges that models aren't going away: As a tool for identifying people in difficulty, they are amazing. But as a tool for punishing and disenfranchising, they're a nightmare.”
—Cory Doctorow, author of Little Brother and co-editor of Boing Boing

“Many algorithms are slaves to the inequalities of power and prejudice. If you don’t want these algorithms to become your masters, read Weapons of Math Destruction by Cathy O’Neil to deconstruct the latest growing tyranny of an arrogant establishment.”
—Ralph Nader, author of Unsafe at Any Speed

“In this fascinating account, Cathy O'Neil leverages her expertise in mathematics and her passion for social justice to poke holes in the triumphant narrative of Big Data. She makes a compelling case that math is being used to squeeze marginalized segments of society and magnify inequities. Her analysis is superb, her writing is enticing, and her findings are unsettling.”
—danah boyd, founder of Data & Society and author of It’s Complicated

"From getting a job to finding a spouse, predictive algorithms are silently shaping and controlling our destinies. Cathy O'Neil takes us on a journey of outrage and wonder, with prose that makes you feel like it's just a conversation. But it’s an important one. We need to reckon with technology.”
—Linda Tirado, author of Hand to Mouth: Living in Bootstrap America

“Next time you hear someone gushing uncritically about the wonders of Big Data, show them Weapons of Math Destruction. It’ll be salutary.”
—Felix Salmon, Fusion

About the Author
Cathy O'Neil is a data scientist and author of the blog mathbabe.org. She earned a Ph.D. in mathematics from Harvard and taught at Barnard College before moving to the private sector, where she worked for the hedge fund D. E. Shaw. She then worked as a data scientist at various start-ups, building models that predict people’s purchases and clicks. O’Neil started the Lede Program in Data Journalism at Columbia and is the author of Doing Data Science. She is currently a columnist for Bloomberg View.

1

BOMB PARTS
What Is a Model?

It was a hot August afternoon in 1946. Lou Boudreau, the player-manager of the Cleveland Indians, was having a miserable day. In the first game of a doubleheader, Ted Williams had almost single-handedly annihilated his team. Williams, perhaps the game’s greatest hitter at the time, had smashed three home runs and driven home eight. The Indians ended up losing 11 to 10.

Boudreau had to take action. So when Williams came up for the first time in the second game, players on the Indians’ side started moving around. Boudreau, the shortstop, jogged over to where the second baseman would usually stand, and the second baseman backed into short right field. The third baseman moved to his left, into the shortstop’s hole. It was clear that Boudreau, perhaps out of desperation, was shifting the entire orientation of his defense in an attempt to turn Ted Williams’s hits into outs.

In other words, he was thinking like a data scientist. He had analyzed crude data, most of it observational: Ted Williams usually hit the ball to right field. Then he adjusted. And it worked. Fielders caught more of Williams’s blistering line drives than before (though they could do nothing about the home runs sailing over their heads).

If you go to a major league baseball game today, you’ll see that defenses now treat nearly every player like Ted Williams. While Boudreau merely observed where Williams usually hit the ball, managers now know precisely where every player has hit every ball over the last week, over the last month, throughout his career, against left-handers, when he has two strikes, and so on. Using this historical data, they analyze their current situation and calculate the positioning that is associated with the highest probability of success. And that sometimes involves moving players far across the field.

Shifting defenses is only one piece of a much larger question: What steps can baseball teams take to maximize the probability that they’ll win? In their hunt for answers, baseball statisticians have scrutinized every variable they can quantify and attached it to a value. How much more is a double worth than a single? When, if ever, is it worth it to bunt a runner from first to second base?

The answers to all of these questions are blended and combined into mathematical models of their sport. These are parallel universes of the baseball world, each a complex tapestry of probabilities. They include every measurable relationship among every one of the sport’s components, from walks to home runs to the players themselves. The purpose of the model is to run different scenarios at every juncture, looking for the optimal combinations. If the Yankees bring in a right-handed pitcher to face Angels slugger Mike Trout, as compared to leaving in the current pitcher, how much more likely are they to get him out? And how will that affect their overall odds of winning?

Baseball is an ideal home for predictive mathematical modeling. As Michael Lewis wrote in his 2003 bestseller, Moneyball, the sport has attracted data nerds throughout its history. In decades past, fans would pore over the stats on the back of baseball cards, analyzing Carl Yastrzemski’s home run patterns or comparing Roger Clemens’s and Dwight Gooden’s strikeout totals. But starting in the 1980s, serious statisticians started to investigate what these figures, along with an avalanche of new ones, really meant: how they translated into wins, and how executives could maximize success with a minimum of dollars.

“Moneyball” is now shorthand for any statistical approach in domains long ruled by the gut. But baseball represents a healthy case study—and it serves as a useful contrast to the toxic models, or WMDs, that are popping up in so many areas of our lives. Baseball models are fair, in part, because they’re transparent. Everyone has access to the stats and can understand more or less how they’re interpreted. Yes, one team’s model might give more value to home run hitters, while another might discount them a bit, because sluggers tend to strike out a lot. But in either case, the numbers of home runs and strikeouts are there for everyone to see.

Baseball also has statistical rigor. Its gurus have an immense data set at hand, almost all of it directly related to the performance of players in the game. Moreover, their data is highly relevant to the outcomes they are trying to predict. This may sound obvious, but as we’ll see throughout this book, the folks building WMDs routinely lack data for the behaviors they’re most interested in. So they substitute stand-in data, or proxies. They draw statistical correlations between a person’s zip code or language patterns and her potential to pay back a loan or handle a job. These correlations are discriminatory, and some of them are illegal. Baseball models, for the most part, don’t use proxies because they use pertinent inputs like balls, strikes, and hits.

Most crucially, that data is constantly pouring in, with new statistics from an average of twelve or thirteen games arriving daily from April to October. Statisticians can compare the results of these games to the predictions of their models, and they can see where they were wrong. Maybe they predicted that a left-handed reliever would give up lots of hits to right-handed batters—and yet he mowed them down. If so, the stats team has to tweak their model and also carry out research on why they got it wrong. Did the pitcher’s new screwball affect his statistics? Does he pitch better at night? Whatever they learn, they can feed back into the model, refining it. That’s how trustworthy models operate. They maintain a constant back-and-forth with whatever in the world they’re trying to understand or predict. Conditions change, and so must the model.

Now, you may look at the baseball model, with its thousands of changing variables, and wonder how we could even be comparing it to the model used to evaluate teachers in Washington, D.C., schools. In one of them, an entire sport is modeled in fastidious detail and updated continuously. The other, while cloaked in mystery, appears to lean heavily on a handful of test results from one year to the next. Is that really a model?

The answer is yes. A model, after all, is nothing more than an abstract representation of some process, be it a baseball game, an oil company’s supply chain, a foreign government’s actions, or a movie theater’s attendance. Whether it’s running in a computer program or in our head, the model takes what we know and uses it to predict responses in various situations. All of us carry thousands of models in our heads. They tell us what to expect, and they guide our decisions.

Here’s an informal model I use every day. As a mother of three, I cook the meals at home—my husband, bless his heart, cannot remember to put salt in pasta water. Each night when I begin to cook a family meal, I internally and intuitively model everyone’s appetite. I know that one of my sons loves chicken (but hates hamburgers), while another will eat only the pasta (with extra grated parmesan cheese). But I also have to take into account that people’s appetites vary from day to day, so a change can catch my model by surprise. There’s some unavoidable uncertainty involved.

The input to my internal cooking model is the information I have about my family, the ingredients I have on hand or I know are available, and my own energy, time, and ambition. The output is how and what I decide to cook. I evaluate the success of a meal by how satisfied my family seems at the end of it, how much they’ve eaten, and how healthy the food was. Seeing how well it is received and how much of it is enjoyed allows me to update my model for the next time I cook. The updates and adjustments make it what statisticians call a “dynamic model.”

Over the years I’ve gotten pretty good at making meals for my family, I’m proud to say. But what if my husband and I go away for a week, and I want to explain my system to my mom so she can fill in for me? Or what if my friend who has kids wants to know my methods? That’s when I’d start to formalize my model, making it much more systematic and, in some sense, mathematical. And if I were feeling ambitious, I might put it into a computer program.

Ideally, the program would include all of the available food options, their nutritional value and cost, and a complete database of my family’s tastes: each individual’s preferences and aversions. It would be hard, though, to sit down and summon all that informationoff the top of my head. I’ve got loads of memories of people grabbing seconds of asparagus or avoiding the string beans. But they’re all mixed up and hard to formalize in a comprehensive list.

The better solution would be to train the model over time, entering data every day on what I’d bought and cooked and noting the responses of each family member. I would also include parameters, or constraints. I might limit the fruits and vegetables to what’s in season and dole out a certain amount of Pop-Tarts, but only enough to forestall an open rebellion. I also would add a number of rules. This one likes meat, this one likes bread and pasta, this one drinks lots of milk and insists on spreading Nutella on everything in sight.

If I made this work a major priority, over many months I might come up with a very good model. I would have turned the food management I keep in my head, my informal internal model, into a formal external one. In creating my model, I’d be extending my power and influence in the world. I’d be building an automated me that others can implement, even when I’m not around.

There would always be mistakes, however, because models are, by their very nature, simplifications. No model can include all of the real world’s complexity or the nuance of human communication. Inevitably, some important information gets left out. I might have neglected to inform my model that junk-food rules are relaxed on birthdays, or that raw carrots are more popular than the cooked variety.

To create a model, then, we make choices about what’s important enough to include, simplifying the world into a toy version that can be easily understood and from which we can infer important facts and actions. We expect it to handle only one job and accept that it will occasionally act like a clueless machine, one with enormous blind spots.

Sometimes these blind spots don’t matter. When we ask Google Maps for directions, it models the world as a series of roads, tunnels, and bridges. It ignores the buildings, because they aren’t relevant to the task. When avionics software guides an airplane, it models the wind, the speed of the plane, and the landing strip below, but not the streets, tunnels, buildings, and people.

A model's blind spots reflect the judgments and priorities of its creators. While the choices in Google Maps and avionics software appear cut and dried, others are far more problematic. The value-added model in Washington, D.C., schools, to return to that example, evaluates teachers largely on the basis of students' test scores, while ignoring how much the teachers engage the students, work on specific skills, deal with classroom management, or help students with personal and family problems. It's overly simple, sacrificing accuracy and insight for efficiency. Yet from the administrators' perspective it provides an effective tool to ferret out hundreds of apparently underperforming teachers, even at the risk of misreading some of them.

Here we see that models, despite their reputation for impartiality, reflect goals and ideology. When I removed the possibility of eating Pop-Tarts at every meal, I was imposing my ideology on the meals model. It's something we do without a second thought. Our own values and desires influence our choices, from the data we choose to collect to the questions we ask. Models are opinions embedded in mathematics.

Whether or not a model works is also a matter of opinion. After all, a key component of every model, whether formal or informal, is its definition of success. This is an important point that we'll return to as we explore the dark world of WMDs. In each case, we must ask not only who designed the model but also what that person or company is trying to accomplish. If the North Korean government built a model for my family's meals, for example, it might be optimized to keep us above the threshold of starvation at the lowest cost, based on the food stock available. Preferences would count for little or nothing. By contrast, if my kids were creating the model, success might feature ice cream at every meal. My own model attempts to blend a bit of the North Koreans' resource management with the happiness of my kids, along with my own priorities of health, convenience, diversity of experience, and sustainability. As a result, it's much more complex. But it still reflects my own personal reality. And a model built for today will work a bit worse tomorrow. It will grow stale if it's not constantly updated. Prices change, as do people's preferences. A model built for a six-year-old won't work for a teenager.

This is true of internal models as well. You can often see troubles when grandparents visit a grandchild they haven't seen for a while. On their previous visit, they gathered data on what the child knows, what makes her laugh, and what TV show she likes and (unconsciously) created a model for relating to this particular four-year-old. Upon meeting her a year later, they can suffer a few awkward hours because their models are out of date. Thomas the Tank Engine, it turns out, is no longer cool. It takes some time to gather new data about the child and adjust their models.

This is not to say that good models cannot be primitive. Some very effective ones hinge on a single variable. The most common model for detecting fires in a home or office weighs only one strongly correlated variable, the presence of smoke. That's usually enough. But modelers run into problems—or subject us to problems—when they focus models as simple as a smoke alarm on their fellow humans.

Racism, at the individual level, can be seen as a predictive model whirring away in billions of human minds around the world. It is built from faulty, incomplete, or generalized data. Whether it comes from experience or hearsay, the data indicates that certain types of people have behaved badly. That generates a binary prediction that all people of that race will behave that same way.

Needless to say, racists don't spend a lot of time hunting down reliable data to train their twisted models. And once their model morphs into a belief, it becomes hardwired. It generates poisonous assumptions, yet rarely tests them, settling instead for data that seems to confirm and fortify them. Consequently, racism is the most slovenly of predictive models. It is powered by haphazard data gathering and spurious correlations, reinforced by institutional inequities, and polluted by confirmation bias. In this way, oddly enough, racism operates like many of the WMDs I'll be describing in this book.

To overcome the issue, we now give you the modern technology to download the book *Weapons Of Math Destruction By Cathy O'Neil* not in a thick published documents. Yeah, reading Weapons Of Math Destruction By Cathy O'Neil by online or obtaining the soft-file just to read can be among the methods to do. You might not really feel that reviewing a book Weapons Of Math Destruction By Cathy O'Neil will work for you. But, in some terms, May people effective are those that have reading behavior, included this type of this Weapons Of Math Destruction By Cathy O'Neil