

# Everybody Lies

What the Internet can tell us about who we really are

**Seth Stephens - Davidowitz**

**This book is a New York Times bestseller**

**I am an internet data expert. Every day I track the digital trails that people leave as they make their way across the web.**

**Google Trends, a tool that was released with little fanfare in 2009, tells users how frequently any word or phrase has been searched in different locations at different times.**

**People's search for information is in itself information.**

**When and where they search for facts, jokes, quotes, places, persons, things, help can tell us about what they think, what they really desire and really fear.**

**The power in Google data is that people tell the giant search engine things they might not tell anyone.**

**In the US , the search for the word nigger , showed where trump won his votes. The more the place where this word was searched, the more they favored Trump.**



**In the 2016 election, people who searched for Trump Clinton tended to vote for Trump and those who searched for Clinton Trump voted for Hillary. The first name they used is a giveaway and they tend to vote for that person.**

**Big data is inherently vague, how big is big?**

**Is**

**18,462 big or does 18,463 make it big?**

**Searches for jokes are the lowest on a Monday, when people report they are most unhappy. Search for jokes is also low on cloudy days and rainy days and they plummet after a major tragedy.**

**Like it or not, data is playing an increasingly important role in our lives, and its role is getting larger.**

**At its core, data science is about spotting patterns and predicting how one variable will affect another. People do this all the time.**

**Just having mounds and mounds of data by itself does not automatically generate insights. Data size by itself is overrated.**

**The big data revolution is less about collecting more and more data. It is about collecting the right data.**

**In buying race horses, the old way was to buy pedigree horses, where the parentage of the horse on sale had won derby's.**



**A data scientist Seder studied the size of the nostril in horses and found there was no correlation. He then found that the size of the heart and specifically the size of the left ventricle was a massive predictor of a horse's ability to win races.**

**If you are going to try to use new data to revolutionize a field, then it is best to choose a field where old methods are lousy.**

**Before a storm, households tend to stock up strawberry pop tarts. Why? Because they do not require cooking or refrigeration, both will not be possible in a storm. Wal Mart regularly stocks up at its outlet before a storm now as a result.**

**Dating is another area where a lot of data has been checked. A man gets a second date from the woman when he laughs at her jokes, he limits the range of his speech.**

**Women on the other hand speak more softly, varies her pitch. A woman is likely to be more interested if she talks about herself. If a man asks too many questions on a first date, he does not get another date from her.**

**Women like to talk about shopping and their hair while men want to talk about football and x box in the US. Women use the tomorrow more than men.**

**Content that is positive is more likely to go more viral.**

**Four people at Brown and Berkeley did a unique study. They took pictures from school and college year books and developed the average looking girl and boy from 1900 to 2015.**



**One clear shift was that current boys and girls smile more. When this was analyzed, the reason early pictures of people did not have smiles was because they were aiming to be like paintings. In paintings, a person sat for hours and could not smile for hours and hence painters painted portraits with studious looking poses.**

**Kodak's happiness moments campaign changed that completely as Kodak said that you should take pictures when you are happy and that's how the smile came into pictures.**

**Any theory of racism has to explain a big puzzle in America. On the one hand, the overwhelming number of black Americans feel they suffer from prejudice, on the other hand very few Americans will admit to being racist.**

**The discrimination black people feel in America is fueled by explicit, hidden hostility**

**About 28 % of young girls are overweight in the USA and 35 % of young boys are overweight. However when one sees parents search, parents tend to search about girls overweight challenge twice as much as they check for boys!**

**On social media, as in surveys, you have no incentive to tell the truth. On social media, you have a large incentive to make yourself look good.**

**Facebook is digital brag to my friends about how good my life is serum. In a Facebook world, the average adult seems happy, is happily married, vacationing in the Caribbean and reading the Atlantic.**

**The digital truth is about searches, views, clicks and swipes. Digital lies are about social media posts, social media likes and dating profiles.**



**Peter Thiel an early investor in Facebook says that great businesses are built on secrets, either secrets about nature or secrets about people**

**Thiel defines secrets about people as things that people don't know about themselves or things they hide because they don't want others to know. These kinds of businesses in other words are built around people's lies.**

**Netflix learned a similar lesson about people's choice. Netflix started by asking people for a list of movies they want to see and then reminded them when to watch them. Nothing happened.**

**Poor people tend to live longer in a city with rich people because they start picking the positives, smoke less, eat less, exercise more etc.**

**A/B testing is very prevalent in a digital world where you ask selected consumers about the preferences on line.**

**Obama's second campaign showed him alone in A setting and his family in a B setting picture. The second picture 'B' won and they collected nearly \$60m million as a result of donations. Marketers must test as much as possible using A/B method.**

**Big data can help us identify significant pockets in distinct geographies which earlier data in a city would not throw up.**

**The things we often measure are not things that we care about.**



**The terms people use in their application by people who are most likely to pay back the loan ... debt free, after tax, lower interest rate, graduate and minimum payment.**

**The terms people use in their application by people who are not likely to pay back the loan .... God, will pay, promise, thank you and hospital.**

**Do we want to live in a world where people are selected and deselected based on the use of language and words?**

**The danger of data revolution is that as more of our life is quantified , these proxy judgements can get more esoteric yet more intrusive. Better predictions can lead to subtler and more nefarious discrimination.**

**Casinos need to judge the gamblers to a nicety. They have to judge the 'pain point' of a gambler. It could be \$3000 for one and \$5000 for another.**

**If a gambler loses more than the pain point, then they will not come back again and that would be a big loss to the casino.**

**Casinos scan data about their patrons and ensure that they offer them a steak dinner or something else before the gambler reached the pain point.**

