

СЕТ СТИВЕНС-ДАВИДОВИЦ

бывший специалист по Data Science в Google

18+

НЕ ЛГУ СЕБЕ

НОВИНКА
ОТ АВТОРА
БЕСТСЕЛЛЕРА
«ВСЕ ЛГУТ»

ПОЧЕМУ
BIG DATA
ЗНАЕТ ТЕБЯ ЛУЧШЕ,
ЧЕМ ТЫ САМ, И КАК
ИСПОЛЬЗОВАТЬ ЭТО,
ЧТОБЫ ДОБИТЬСЯ УСПЕХА



БОМБОРА
ИЗДАТЕЛЬСТВО

Сет Стивенс-Давидовиц

**Не лги себе. Почему Big
Data знает тебя лучше, чем
ты сам, и как использовать
это, чтобы добиться успеха**

«ЭКСМО»

2022

УДК 316.3:004.738.5
ББК 60.5+32.973.202

Стивенс-Давидовиц С.

Не лги себе. Почему Big Data знает тебя лучше, чем ты сам, и как использовать это, чтобы добиться успеха / С. Стивенс-Давидовиц — «Эксмо», 2022 — (Библиотека ИТ. Главные книги о современных технологиях)

ISBN 978-5-04-192172-9

За последнее десятилетие ученые изучили гигантские наборы данных, чтобы найти новые подходы к решению самых важных жизненных вопросов. Исследователь данных Сет Стивенс-Давидовиц проанализировал множество научных исследований об удовольствии и счастье, чтобы понять, чего мы хотим от жизни на самом деле. В формате PDF A4 сохранен издательский макет книги.

УДК 316.3:004.738.5

ББК 60.5+32.973.202

ISBN 978-5-04-192172-9

© Стивенс-Давидовиц С., 2022

© Эксмо, 2022

Содержание

Введение	6
Moneyball для вашей собственной жизни	8
Передвижения в игровом поле жизни	11
От «Все лгут» к «Не лги себе»	13
От Бога к чувствам, от чувств – к данным	16
Глава 1	18
Конец ознакомительного фрагмента.	21

Сет Стивенс-Давидовиц
Не лги себе. Почему Big Data знает
тебя лучше, чем ты сам, и как
использовать это, чтобы добиться успеха

Джули
Если данные говорят, что любить тебя неправильно, я не хочу
правильного.

Don't Trust Your Gut: Using Data to Get What You Really Want in Life
Seth Stephens-Davidowitz
© 2022 by Seth Stephens-Davidowitz



© Вантух К. А., перевод на русский язык, 2023
© Оформление. ООО «Издательство «Эксмо», 2023

Введение

Учебник самопомощи для верящих в данные

Вам вполне по силам улучшить качество решений, касающихся вашей жизни. Большие данные могут вам в этом помочь.

На наших глазах радикально, хотя и незаметно на первый взгляд, меняются наши представления о самых важных областях человеческой жизни. Причиной тому служат Интернет и данные, которые он сгенерировал. В последние годы ученые проводили исследование различных, но всегда огромных массивов данных – от сообщений OkCupid до биографических статей на Wikipedia и статуса отношений в социальных сетях. В этих миллиардах записей они смогли найти – вероятно, впервые – достоверные ответы на основные жизненные вопросы. В частности, на такие:

- По каким признакам можно определить хорошего родителя?
- Кто богат, а мы не догадываемся об этом? И почему?
- Каковы шансы стать знаменитостью?
- Почему некоторые люди необычайно удачливы?
- По каким признакам можно предсказать счастливый брак?
- Если посмотреть на вопрос шире, что вообще делает людей счастливыми?

Часто ответы, найденные при помощи данных, не соответствуют нашим интуитивным догадкам и требуют не тех решений, которые мы приняли бы. Дело обстоит просто: из этих гор данных можно сделать выводы, которые дают возможность вам – или кому-нибудь из ваших знакомых – улучшить качество своих решений.

Вот три примера, взятых из исследовательских работ, касающихся далеких друг от друга областей жизни.

Пример № 1: предположим, вы холостяк или незамужняя женщина и личная жизнь у вас обстоит хуже, чем хотелось бы. Вы пытаетесь улучшить себя всеми способами, которые вам советуют окружающие. Одеваетесь лучше. Отбеливаете зубы. Делаете новую дорогую стрижку. Все впустую. Женихи (или невесты) не толпятся под вашей дверью.

Здесь могут оказаться полезными выводы из больших данных.

Математик и литератор Кристиан Раддер исследовал миллионы случаев выбора на OkCupid в поисках свойств, обеспечивающих наибольший успех у противоположного пола¹. Он обнаружил – и это совсем не удивительно, – что популярнее всего оказываются красавцы и красавицы, живущие среди нас Брэды Питты и Натали Портман.

Но в этих горах данных он обнаружил и другой тип людей, показавший неожиданно высокую популярность: обладателей экстремальной внешности. Представьте себе, например, синие волосы, боди-арт, безумные очки или бритый череп.

В чем же дело? Ключ к успеху обладателей необычной внешности лежит в том, что, хотя многим она не слишком нравится, всегда есть группа людей, которых она привлекает *очень*. А в вопросах отношений с противоположным полом это решает все.

Здесь, если ваша внешность отличается от кинематографического идеала красоты, лучшей стратегией будет, по словам Раддера, набрать «много “да”, много “нет”, но как можно меньше “э...”». Подобная тактика, как открыл Раддер, способна увеличить число сообщений на 70 %. Создайте крайний вариант себя, говорят нам данные, и найдутся люди, для которых вы будете в высшей степени желанны.

¹ Christian Rudder, *Dataclysm: Who We Are (When We Think No One's Looking)* (New York: Broadway Books, 2014). – *Здесь и далее прим. авт., если не указано иное.*

Пример № 2: предположим, у вас только что родился ребенок (кстати, мазаль тов²). Нужно выбрать район, где вы собираетесь его растить. Вы советуетесь с несколькими друзьями, выясняете у Google некоторые основные факты – и, собственно, все. Вы счастливый обладатель дома для своей семьи. Не бином Ньютона, правда?

В действительности сегодня дело обстоит ровно наоборот.

Ученые воспользовались недавно оцифрованными данными налоговой службы, чтобы исследовать жизненные траектории сотен миллионов американцев. Они обнаружили, что возможность вырасти в определенном городе – более того, в определенном квартале этого города – резко повышает шансы в жизни. Причем эти кварталы могут оказаться вовсе не теми, которые люди считают наиболее престижными. И не теми, жилье в которых стоит дороже всего. Сейчас существуют карты, созданные на основе тщательного анализа данных, которые показывают родителям перспективность любого квартала в Соединенных Штатах.

И это не все. Ученые внимательнейшим образом просеяли данные в поисках свойств, общих для всех перспективных районов. По ходу работы они откинули значительную часть житейской мудрости, касающейся воспитания детей. Благодаря большому набору данных мы наконец можем просветить родителей насчет того, что действительно имеет значение для успеха ребенка (намек: пример взрослых), а что имеет гораздо меньший вес (намек: лучшие школы).

И пример № 3: предположим, вы начинающий художник, которому никак не удается поймать удачу за хвост. Вы покупаете все книги по технике живописи, какие можете. Вы просите оценок и советов у друзей. Вы переделываете свои работы снова, снова и снова. И при этом все кажется бесполезным. Вы не можете понять, что делаете не так.

Большие данные указали на вероятную ошибку.

Недавнее исследование карьеры тысяч художников под руководством Сэмюэла П. Фрайбергера³ открыло прежде остававшуюся незамеченной закономерность, почему некоторые из них добиваются успеха, а некоторые нет. Итак, в чем же секрет отличия великих имен?

Часто он заключается в том, как выставлять свои работы. Данные говорят нам, что художники, не добившиеся ничего серьезного, имеют тенденцию выставлять свои работы все время в одних и тех же немногочисленных галереях. Те же, кто сделал большую карьеру, как правило, выставляются в гораздо большем числе точек, увеличивая тем самым вероятность встречи со своим шансом.

О важности демонстрации себя ради хорошей карьеры говорили многие. Но ученые, занятые исследованием данных, показали, что важно демонстрировать себя много где.

Я вовсе не хочу сказать, что эта книга будет источником советов исключительно для одиноких, молодых родителей и начинающих художников (хотя полезные указания для каждой из этих групп в ней еще будут). Моя задача – подчеркнуть выводы из больших массивов данных, которые были бы полезны именно для вас, независимо от того, на каком этапе жизни вы находитесь. Ниже последуют недавно разработанные указания, как быть счастливее, выглядеть лучше, продвинуться в карьере – и многое другое. А идея книги пришла ко мне как-то вечером, когда я... смотрел бейсбольный матч.

² «Мазаль тов» – фраза на иврите, которая используется для поздравления в честь какого-либо события в жизни человека. – Прим. пер.

³ Samuel P. Fraiberger et al., “Quantifying reputation and success in art”, Science 362(6416) (2018): 825–29.

Moneyball для вашей собственной жизни

И я, и другие любители бейсбола не могли не заметить, что он стал совсем не той игрой, что тридцать лет назад. Когда я был мальчишкой и болел за New York Mets, бейсбольные команды выбирали тактику, опираясь на собственное понимание игры и интуицию. Они решали, сделать ли им бант или украсть базу, в зависимости от того, как смотрел на это менеджер команды. Они выбирали игроков для приобретения, опираясь на впечатления скаута.

Однако во второй половине XX века стали появляться признаки, указывавшие на существование более разумной тактики. В моем детстве отец каждый год приносил домой новую книгу Билла Джеймса. Джеймс, который работал охранником на заводе консервированных бобов со свининой, был одержим бейсболом. И у него был нестандартный метод анализа игры – с помощью недавно появившихся компьютеров и оцифрованных данных. Джеймс и его коллеги – они называли себя сайберметристами – при помощи анализа данных выяснили, что большинство решений, принимавшихся командами под влиянием интуиции, были полностью ошибочными.

Как часто команде нужно исполнять бант? Значительно реже, чем сейчас, говорили сайберметристы. А как часто следует красть базы? Почти никогда. Сколько должны стоить игроки, приносящие много пробежек? Больше, чем думали команды. Кого следовало приобретать? Больше питчеров из университетских команд.

Работа Джеймса производила захватывающее впечатление не только на моего отца. Билли Бин, который начинал карьеру в качестве игрока, а впоследствии переквалифицировался в бейсбольного администратора, тоже был его горячим сторонником. И став генеральным директором клуба Oakland Athletics, он решил управлять им в соответствии с принципами сайберметрики.

Идея принесла выдающиеся результаты. В книге Moneyball приводится довольно известный факт: в Oakland Athletics платили очень скромные зарплаты, но при этом команда выходила в плей-офф в 2002 и 2003 годах⁴. С тех пор роль аналитики в бейсболе резко возросла. Клуб Tampa Bay Rays, о котором говорили, что он больше следует Moneyball, чем сама команда Oakland Athletics из Moneyball⁵, вышел в World Series 2020, несмотря на третий с конца уровень зарплат в бейсболе.

Принципы Moneyball и лежащая в их основе здравая идея, что когнитивные искажения могут быть компенсированы данными, повлияли на многие учреждения и виды спорта. Команды NBA все больше используют аналитику, прослеживающую траекторию каждого броска⁶. В данных о 300 миллионах бросков были найдены значительные отклонения от оптимальной техники. Оказывается, что для среднего игрока NBA, выполняющего бросок в прыжке, вероятность пропустить бросок с недолетом вдвое выше, чем бросок с перелетом. А когда он выполняет бросок из угла, он скорее промахнется в сторону, противоположную щиту, потому что может опасаться попасть в него. Игроки воспользовались подобными данными, чтобы и корректировать когнитивные искажения, и одновременно делать больше бросков.

Фирмы Кремниевой долины в значительной степени опираются на принципы, изложенные в Moneyball. Google, где я в прошлом работал аналитиком данных, определенно верит в полезность данных при принятии важных решений. Была довольно известная история, когда оттуда уволился дизайнер, недовольный тем, что компания предпочитала данные, а не интуи-

⁴ Michael Lewis “Moneyball: The Art of Winning an Unfair Game” (New York: Norton, 2004).

⁵ Jared Diamond, “How to succeed in baseball without spending money”, Wall Street Journal, October 1, 2019.

⁶ Ben Dowsett, “How shot-tracking is changing the way basketball players fix their game”, FiveThirtyEight, August 16, 2021, <https://fivethirtyeight.com/features/how-shot-tracking-is-changing-the-way-basketball-players-fix-their-game/>.

цию квалифицированных дизайнеров. Последней каплей для него стал эксперимент, в котором компания испытывала сорок один оттенок синего⁷ для гиперссылок в Gmail, чтобы выяснить на практике, какой из них даст больше всего кликов. Возможно, дизайнер и был недоволен, но эксперимент принес Google 200 миллионов долларов дополнительного дохода в год⁸. Google ни разу не поколебался в своей вере в данные – и со временем превратился в компанию ценой в 1,8 триллиона долларов. Как сказал ее бывший исполнительный директор Эрик Шмидт: «В Бога мы верим. Все остальные должны предоставлять данные»⁹.

Джеймс Симонс, математик мирового класса и основатель компании Renaissance Technologies, принес строгий анализ данных на Уолл-стрит. Он и его группа количественных аналитиков создали беспрецедентный массив данных, содержащий одновременно курсы акций и события реального мира, и подвергли его анализу на предмет закономерностей. Какова тенденция изменения курсов после того, как компания-эмитент объявляет о прибылях? А при дефиците хлеба? А после упоминания компании в газете?

С момента основания Renaissance ее флагманский инвестиционный фонд Medallion¹⁰, который в своей торговой стратегии опирается исключительно на закономерности в данных, всегда приносил 66 % прибыли до вычета налогов и сборов. В тот же период S&P 500 приносил 10 % до вычета. Экономист Кеннет Френч (его имя связывают с гипотезой эффективного рынка, говорящей о практической невозможности обеспечить показатели существенно выше S&P 500) так объясняет успех Renaissance: «Видимо, они просто лучше всех остальных»¹¹.

Но как нам принимать важные решения, касающиеся личной жизни? Как выбрать партнера для брака, как ходить на свидания, как проводить время, соглашаться ли на то или иное предложение о работе?

На кого мы больше похожи – на Oakland Athletics в 2002 году или на прочие бейсбольные команды в то же время? На Google или на привычный магазин? На Renaissance Technologies или на обычного управляющего инвестиционным фондом?

Я бы сказал, что большинство из нас принимают важнейшие решения, опираясь на интуицию. Может быть, мы посоветуемся с кем-то из друзей, родственников или самозванных гуру по части искусства жить. Может быть, прочитаем какие-то ни на чем не основанные советы. Бросим беглый взгляд на самую базовую статистику. И затем просто сделаем то, что кажется нам правильным.

«Что бы произошло, начни мы решать самые важные жизненные вопросы с опорой на данные?» – спрашивал я себя, смотря бейсбольный матч по телевизору. Если бы мы администрировали свою жизнь так же, как Билли Бин – клуб Oakland Athletics?

Я знаю, что в наши дни подобный подход становится все более осуществимым. В своей предыдущей книге «Все лгут» я показывал, как новые данные, которые стали доступны нам благодаря Интернету, меняют наши представления об обществе и работе человеческого ума. Может быть, статистическая революция началась с бейсбола именно из-за статистической информации, которую собирали и на которую создавали спрос сумасшедшие болельщики. Так сказать, «революция Moneyball для нашей жизни» стала возможной благодаря данным, которые собрали наши компьютеры и смартфоны.

Давайте зададимся не таким уж тривиальным вопросом: что делает людей счастливыми?

⁷ Douglas Bowman, “Goodbye, Google”, <https://stopdesign.com/archive/2009/03/20/goodbye-google.html>, March 20, 2009.

⁸ Alex Horn, “Why Google has 200m reasons to put engineers over designers”, Guardian, February 5, 2014.

⁹ “Are we better off with the internet?” YouTube, uploaded by the Aspen Institute, July 1, 2012, <https://www.youtube.com/watch?v=djVrLNaFvIo>.

¹⁰ Gregory Zuckerman, *The Man Who Solved the Market* (New York: Penguin, 2019).

¹¹ Amy Whyte, “Famed Medallion fund ‘stretches... explanation to the limit,’ professor claims”, Institutional Investor, January 26, 2020, <https://www.institutionalinvestor.com/article/b1k2fymby99nj0/Famed-Medallion-Fund-Stretches-Explanation-to-the-Limit-Professor-Claims>.

Данные, необходимые для строгого и систематического ответа на этот вопрос, в XX веке были недоступны.

Когда революция Moneyball потрясла мир бейсбола, в распоряжении сайберметристов были аккуратнейшим образом зарегистрированные данные по каждой игре и им было что анализировать. Но аналитики данных тогда не располагали подобными сведениями относительно существенных жизненных решений и настроений обычных людей. В те времена счастье, в отличие от бейсбола, не поддавалось строгому анализу.

Но теперь такая возможность есть.

Блестящие специалисты из Google, Джордж Маккеррон и Сюзанна Мурато, при помощи аппаратов iPhone сформировали не имеющий аналогов массив данных о счастье и назвали свой проект Mappiness¹². Они привлекли к работе десятки тысяч пользователей, которых опрашивали по несколько раз в течение дня. Им задавали простые вопросы: что они делают в данный момент, с кем они, насколько при этом счастливы. Таким образом они получили массив данных более чем из трех миллионов «замеров счастья». Это нельзя даже сравнивать с десятками измерений, на которые опирались исследования счастья в прошлом.

Некоторые из скрытых в этих миллионах точек результатов наводят на размышления. Например, болельщики получают больше страданий от проигрыша своей команды, чем радости от ее победы. Иногда они противоречат нашим интуитивным представлениям: так, употребление алкоголя во время исполнения рутинных обязанностей в среднем доставляет больше удовольствия, чем во время общения с друзьями. Иногда результаты представляются здравыми: работа имеет тенденцию раздражать – если только мы не работаем вместе с друзьями.

Но полезны эти результаты всегда. Вы никогда не задавались вопросом, как в точности погода влияет на настроение? Какие занятия в среднем чаще всего обманывают нас в смысле ожидаемого удовольствия? Насколько деньги действительно важны для счастья? В какой мере настроение зависит от среды? Благодаря Маккеррону и Мурато у нас теперь есть достоверные ответы на эти вопросы – и они будут предметом восьмой и девятой глав. Я даже завершу эту книгу надежной формулой счастья, выведенной из замеров на тысячах смартфонов. Я называю ее «ответом на главный вопрос жизни, полученным при помощи данных».

Итак, последние четыре года я, вдохновившись примером бейсбола, погрузился в напряженную научную работу. Я говорил со специалистами. Читал академические публикации. Рассматривал приложения к публикациям под таким углом, который – я совершенно уверен в этом – еще не приходил в голову ни одному ученому. Провел несколько собственных исследований и интерпретировал их результаты. Свою задачу я видел в том, чтобы найти своих Биллов Джеймсов в таких областях, как брак, воспитание детей, спортивные достижения, финансовое благосостояние, удача, стиль и счастье, – и дать каждому из вас возможность стать Билли Бином своей жизни. Я готов поделиться всем, что узнал.

Называйте это «Moneyball вашей собственной жизни».

¹² Дополнительную информацию о проекте Mappiness можно найти по адресу <http://www.mappiness.org.uk>.

Передвижения в игровом поле жизни

Прежде чем приступить к работе, я задал себе несколько вопросов. Как могла бы выглядеть жизнь, в основу которой положены принципы Moneyball? Как мог бы выглядеть наш процесс принятия решений, если бы мы, подобно Oakland Athletics и Tampa Bay Rays, следовали данным, а не инстинктам? Одно из бросающихся в глаза свойств бейсбола после Moneyball заключается в том, что некоторые решения опирающихся на аналитику команд выглядят... скажем так, немного странными. Вот вам пример – расположение инфилдеров¹³.

В эпоху после Moneyball бейсбольные команды все активнее смещают положение полевых игроков. Они группируют многих своих защитников в одной и той же части поля, оставляя его обширные участки совершенно незащищенными, куда бьющему игроку ничего не стоит направить мяч. Такое смещение игроков на игровом поле кажется болельщикам традиционного бейсбола чистым безумием. Но от безумия оно предельно далеко. Подобное смещение оправдывается огромными массивами данных, предсказывающими, куда именно конкретный игрок, скорее всего, пошлет мяч¹⁴. Числа говорят бейсбольным командам, что такая тактика верна, пусть и кажется неверной на первый взгляд.

Если мы применим подход Moneyball к нашей жизни, то можем ожидать, что некоторые кажущиеся странными решения – назовем их жизненными смещениями на игровом поле – на самом деле оправданны.

Мы уже говорили о поиске пары. Побриться наголо или покрасить волосы в синий цвет, чтобы тебя чаще приглашали на свидания, – это аналоги передвижения на игровом поле, только в качестве поля выступает сама жизнь. А вот еще один аналог, только найденный в больших данных по продажам.

Предположим, вы пытаетесь что-нибудь продать. Этот опыт все больше становится повсеместным. Как пишет Дэниел Пинк в книге *To Sell Is Human*, «неважно, выступаем ли перед коллегами, пытаемся ли убедить тех, от кого зависит финансирование, или развлекаем детей... мы все сейчас занимаемся продажами»¹⁵.

В любом случае, что бы вы ни пытались продать, вы вкладываете в это дело всю свою душу.

Вы пишете текст своего выступления (это хорошо!). Репетируете его (хорошо!). Ночью как следует высыпаетесь (хорошо!). Съедаете основательный завтрак (хорошо!). Справляетесь с нервами и начинаете говорить (хорошо!).

И вот, произнося речь, которая призвана продать ваш товар, вы вспоминаете, что нужно передать аудитории свою уверенность в нем широкой белозубой улыбкой (а вот это, как ни странно... не хорошо).

В недавно проведенном исследовании сопоставляются проявление эмоций агентами по продажам и результаты их работы.

Массив данных – 99 451 выступление на розничной платформе с живым потоковым видео. (Сейчас люди все чаще приобретают товары и услуги на платформах типа Amazon Live, позволяющих продавцам рекламировать свой товар при помощи видео.) Исследовали получали видео каждого такого рекламного выступления и данные о последовавших продажах. (Они также располагали информацией о продаваемом товаре, его цене и наличии бесплатной доставки.)

¹³ Игрок защищающейся стороны, находящийся на игровом поле. – *Прим. пер.*

¹⁴ Rob Arthur and Ben Lindbergh, “Yes, the infield shift works. Probably”, June 30, 2016, <https://fivethirtyeight.com/features/yes-the-infield-shift-works-probably/>.

¹⁵ Daniel H. Pink, *To Sell is Human* (New York: Penguin, 2012).

Методы – искусственный интеллект и глубокое обучение. Исследователи преобразовали 62,32 миллиона видеок кадров в цифровые данные. В частности, искусственный интеллект оказался способен закодировать эмоции, выражаемые продавцом в видео. Выглядел ли он раздраженным? Испытывающим отвращение? Испуганным? Удивленным? Грустным? Или довольным?

Результат: исследователи обнаружили, что эмоции продавца – важнейший фактор для предсказания количества проданного товара. Неудивительно, что, когда продавец выражал отрицательные эмоции, такие как раздражение или отвращение, он продавал меньше. Злость продается плохо. А вот что как раз удивительно, так это то, что, когда он показывал положительные эмоции, такие как довольство или удивление, он тоже продавал меньше! Радость тоже продается плохо. Если речь заходит об увеличении продаж, сдерживание эмоций продавцом – иными словами, когда на лице у него покерфейс вместо широкой улыбки, – сказывается на результатах продаж примерно вдвое лучше, чем бесплатная доставка¹⁶.

Иногда, чтобы продать товар, не стоит слишком демонстрировать энтузиазм на его счет. Может быть, это кажется странным – но данные говорят, что дело обстоит именно так.

¹⁶ Neeraj Bharadwaj et al., “EXPRESS: A New Livestream Retail Analytics Framework to Assess the Sales Impact of Emotional Displays”, *Journal of Marketing*, September 30, 2021.

От «Все лгут» к «Не лги себе»

Дальше последует краткая пауза, на протяжении которой я попытаюсь оправдать свою вторую книгу в глазах читателей первой, «Все лгут»¹⁷. Некоторые из вас, возможно, обратили внимание на эту книгу, поскольку вам понравилась та. Если же вас к этой книге привели другие причины, может, следующие абзацы убедят вас купить и первую. Я постараюсь.

Книга «Все лгут» была посвящена моему исследованию того, как при помощи поисковых запросов Google определить действительные мысли и поступки людей. Я назвал Google «цифровой сывороткой правды», потому что люди честны с поисковой машиной.

Кроме того, я назвал поисковые запросы Google самым важным набором данных о внутреннем мире человека.

Я показал, что:

- Расистские запросы Google предсказывали, где именно Барак Обама покажет слабые результаты на выборах 2008 и 2012 годов.

- Люди часто печатают в поисковой строке Google полные предложения – например, «Я ненавижу своего начальника», «Я пьян» или «Мне нравятся сиськи моей подружки».

- Самый популярный запрос Google, начинающийся со слов «мой муж хочет...», в Индии выглядит как «мой муж хочет, чтобы я кормила его грудью». Вообще в Индии запросов, посвященных кормлению мужа грудью, почти столько же, сколько посвященных грудному вскармливанию младенцев.

- Запросы Google о том, как сделать аборт самостоятельно, почти идеально концентрируются в тех частях Соединенных Штатов, где легально сделать аборт сложно.

- Мужчины чаще ищут, как увеличить собственный член, чем как настроить гитару, сменить колесо или приготовить омлет. Один из самых популярных запросов к Google на тему члена звучит: «Насколько большой у меня член?»

В конце предыдущей книги я предположил, что следующая будет называться «Все лгут (по-прежнему)» и что в ней я буду продолжать разбираться, что мы можем узнать из запросов Google. Кажется, я солгал насчет этого, прошу прощения. Чего еще ожидать от автора книги под названием «Все лгут»?

На первый взгляд моя вторая книга сильно отличается от первой. И если вы надеялись на продолжение анализа того, какие запросы делают мужчины относительно своих детородных органов, вы будете горько разочарованы. Впрочем, ладно. Один пример я приведу. Вы не знали, что мужчины часто набирают в поисковой строке полные предложения, сообщающие длину их пениса?¹⁸ Они печатают, например: «Длина моего члена – 5 дюймов»¹⁹. И если сопоставить данные по всем подобным запросам, мы получим близкое к нормальному распределение длины пениса, сообщаемой Google его владельцем, с пиком около 5 дюймов.

Но давайте оставим мои исследования и вернемся в сумасшедший мир поисковых запросов Google, о котором можно прочесть больше в моей книге «Все лгут».

Большинство научных работ, о которых говорится в этой книге, в отличие от «Все лгут», написаны не мной, а другими людьми. Книга носит более практический характер, больше сосредоточена на самосовершенствовании, чем на погружении в различные аспекты современной жизни. Кроме того, в ней заметно меньше внимания уделяется сексу. Говоря о нем, мы не будем сосредотачиваться на человеческих тайных желаниях и страхах – им уделено достаточно

¹⁷ «Все лгут. Поискники, Big Data и Интернет знают о вас все». Москва: Бомбора, 2022.

¹⁸ Данные о длине члена, которую мужчины сами сообщают Google, можно найти здесь: <https://clck.ru/34BU6P>.

¹⁹ 12,7 см. – *Прим. ред.*

места в предыдущей книге. Обсуждение секса в этой книге сводится к вопросу, делает ли он людей счастливее (спойлер: да).

И тем не менее я уверен, что эта книга служит естественным продолжением предыдущей – по двум причинам.

Во-первых, одним из главных побудительных мотивов к ее написанию для меня стали данные о том, чего на самом деле хотят читатели, а не о том, что они говорят, будто хотят. Закончив работу над «Все лгут», я, как любой приличный маркетолог, задал читателям вопрос, что заинтересовало их больше всего. Большинство ответили мне, что их сильнее всего заинтересовали главы, посвященные важнейшим проблемам нашего мира и путям их решения – например, те, которые касаются преступлений в отношении детей и проблем неравенства.

Но как автор «Все лгут» я отнесся к словам людей скептически и захотел посмотреть на какие-нибудь другие данные – может быть, на «цифровую сыворотку правды». И я стал выяснять, какие места чаще всего подчеркивают в цифровой версии книги для Kindle. Я заметил, что люди часто подчеркивают отрывки, где говорится, как они могли бы улучшить собственную жизнь, и редко – о том, как улучшить мир. Людей интересует самопомощь – независимо от того, готовы ли они в этом признаться.

Более тщательно исследовав данные Kindle, я пришел к этому же выводу. Обработав большой массив книг, исследователи определили, что вероятность обнаружить слово «вы» в подчеркнутых предложениях в двадцать раз выше, чем в неподчеркнутых. Людям, таким образом, действительно нравятся предложения со словом «вы»²⁰.

Именно поэтому первый абзац книги «Не лги себе» построен так, а не иначе:

«Вам вполне по силам улучшить качество решений, касающихся *вашей* жизни. Большие данные могут *вам* в этом помочь».

Данные, а не интуиция, определили содержание этого абзаца. Он открывает для вас книгу, которая призвана помочь вам получить от жизни больше тех вещей, которые именно вы хотите. Он вам понравился?

Популярность литературы, способной предложить читателю помощь, подтверждается анализом книг, которые пользовались наибольшим спросом на протяжении человеческой истории²¹. Я выяснил, какие книги лучше всего продавались. Наиболее многочисленная категория бестселлеров вне категории художественной литературы – это книги по самопомощи (примерно 42 % от общего количества). Вслед за ними идут мемуары знаменитостей (28 %). На третьем месте – исследования в области секса (8 %).

Я пытаюсь сказать, что, ориентируясь на данные, напишу сначала тот самый учебник самопомощи, который вы держите в руках. Затем перейду к книге «Секс: с точки зрения данных». Надеюсь, что к этому моменту стану достаточно известным, чтобы написать «Сет: мемуары автора, прославившегося благодаря данным о том, какие книги лучше продаются».

Второе качество, общее для обеих книг, – в обеих при помощи данных раскрываются секреты современной жизни. Одна из причин такой полезности данных для принятия более качественных решений – основополагающие факты об окружающем нас мире скрыты от нас. Существуют секреты получения от жизни желаемого, которые открываются при помощи больших данных.

Возьмите, например, такой секрет: кто богат? Очевидно, что его знание поможет любому, кто хочет заработать больше денег. Знание осложняется тем обстоятельством, что многие богатые не хотели бы, чтобы об их богатстве знали другие.

²⁰ Ariana Orwell, Ethan Kross, and Susan A. Gelman, “‘You’ speaks to me: Effects of generic-you in creating resonance between people and ideas”, PNAS 117(49) (2020): 31038–45.

²¹ https://en.wikipedia.org/wiki/List_of_best-selling_books.

Недавнее исследование, в котором использовались оцифрованные в последнее время налоговые декларации, было наиболее подробным из тех, что касались богатых людей²². Ученые определили, что типичный богатый американец – отнюдь не кудесник от высоких технологий, корпоративный магнат или нечто подобное, чего обычно ждет публика. Типичный богатый американец, по словам авторов, – это владелец «регионального бизнеса» типа автосалона или компании по оптовой торговле напитками. Кто бы мог подумать? В главе 4 мы поговорим, почему это так – и как это влияет на выбор карьеры.

Медиа тоже лгут нам – или, по крайней мере, создают у нас ложное впечатление об устройстве мира, выбирая для нас только истории определенного свойства. Прибегнуть к данным, чтобы развеять эту ложь, полезно для принятия решений.

Вот пример – связь возраста и предпринимательского успеха. Медиа дают нам искаженную картину возраста бизнесменов. Недавнее исследование показало, что медианный возраст предпринимателей, о которых пишут в бизнес-журналах, – 27 лет²³. Пресса обожает рассказывать нам завлекательные истории о вундеркиндах, создавших крупные компании.

Но сколько же лет капиталисту в действительности? Недавнее исследование данных обо всех предпринимателях Америки показало, что успешному их представителю в среднем 42 года²⁴. И шансы создать успешный бизнес увеличиваются с возрастом, достигая максимума в 60 лет²⁵. Более того, возраст является преимуществом даже для предпринимательства в сфере высоких технологий²⁶ – то есть там, где, по мнению большинства, должна работать исключительно молодежь, поскольку только она способна к освоению нового.

Разумеется, тот факт, что возраст является преимуществом для предпринимательства в любой области, полезен для всех, кто достиг среднего возраста и уверен, что их шанс создать бизнес миновал. В главе 5 мы развеем несколько мифов о предпринимательском успехе и поговорим о выведенной из данных надежной формуле, которая обеспечивает максимум шансов создать успешное предприятие.

Когда вы знаете, как на самом деле устроен мир, – и избегаете лжи как людей, так и медиа, – вы готовы улучшить качество решений, касающихся вашей жизни.

²² Matthew Smith, Danny Yagan, Owen Zidar, and Eric Zwick, “Capitalists in the Twenty-First Century”, *Quarterly Journal of Economics* 134(4) (2019): 1675–1745.

²³ Pierre Azoulay, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda, “Age and High-Growth Entrepreneurship”, *American Economic Review* 2(1) (2020): 65–82.

²⁴ Там же.

²⁵ Там же.

²⁶ Pierre Azoulay, Benjamin F. Jones, J. Daniel Kim, and Javier Miranda, “Age and High-Growth Entrepreneurship”, *American Economic Review* 2(1) (2020): 65–82.

От Бога к чувствам, от чувств – к данным

В последней главе своей книги «Homo Deus» Юваль Ной Харари пишет, что мы проходим через «религиозную революцию гигантского масштаба, подобной которой человечество не видело с XVIII столетия». Новая религия, говорит Харари, это вера в данные, «датаизм»²⁷.

Как же мы пришли к этому?

На протяжении почти всей человеческой истории наиболее образованные люди наделяли высшей властью Бога. Как пишет Харари, «когда люди не знали, на ком жениться, чем заняться и начинать ли войну, они читали Библию и следовали ее советам».

Гуманистическая революция, которую Харари относит к XVIII веку, поставила под вопрос мировоззрение, в центре которого находится Бог. Философы наподобие Вольтера, Джона Локка и моего любимого Дэвида Юма утверждали, что Бог – плод человеческого воображения, а библейские догмы ложны. Устранив внешний авторитет, способный управлять нами, философы стали утверждать, будто люди управляют своей жизнью сами. Харари называет методы принятия важных решений в эпоху гуманизма: «прислушиваться к себе», «любоваться закатом», «вести личный дневник», «беседовать по душам с близким другом».

Революция датаизма, которая только началась и, по словам Харари, может завершиться только через десятки лет, поставила под вопрос мировоззрение гуманистов, в центре которого находятся наши чувства. Тот квазирелигиозный статус, которым они были наделены, был поставлен под сомнение биологами. Они открыли, что организмы – это алгоритмы, а наши чувства – просто «процессы биохимических вычислений»²⁸.

Более того, такие легендарные специалисты в области человеческого поведения, как Амос Тверски и Даниэль Канеман, открыли, что чувства часто вводят нас в заблуждение. Наш ум, говорят они, глубоко поражен разного рода когнитивными искажениями²⁹.

Вы считаете собственную интуицию надежным проводником? Это совсем не так. Зачастую мы излишне оптимистичны, переоцениваем значимость легко запоминающихся историй, отбираем только информацию, согласующуюся с тем, во что нам хотелось бы верить, ложно заключаем, что могли бы объяснить события, считавшиеся непредсказуемыми на момент, когда они произошли, и так далее.

Фраза «прислушиваться к себе» могла звучать освобождающе и романтично для гуманистов. Но, честно говоря, «прислушиваться к себе» звучит просто страшно после прочтения последнего выпуска *Psychological Review* или великолепной статьи «Список когнитивных искажений» в Википедии.

Наконец, революция больших данных предлагает нам альтернативу тому, чтобы прислушиваться к себе. Наша интуиция или советы таких же людей, как мы, могли казаться гуманистам единственными источниками мудрости, оставшимися нам в лишенной Бога вселенной. Но сейчас ученые, работающие в области анализа данных, формируют и анализируют огромные массивы информации, способные освободить нас от когнитивных искажений.

Еще процитирую Харари: «В XXI веке чувства утратили положение лучших алгоритмов в мире. Мы создаем новые, превосходящие их алгоритмы, опирающиеся на беспрецедентную вычислительную мощь и гигантские базы данных». Согласно датаизму, теперь ответы на вопросы «на ком жениться, чем заняться и начинать ли войну» заключаются в «алгоритмах, которые знают нас лучше, чем мы сами знаем себя».

²⁷ Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (New York: Random House, 2016).

²⁸ Yuval Noah Harari. "Organisms Are Algorithms. Body Is Calculator. Answer = Sensation~Feeling~Vedan?", YouTube, uploaded by Rashid Kapadia, June 13, 2020, <https://www.youtube.com/watch?v=GrQ7nY-vevY>.

²⁹ Daniel Kahneman, *Thinking, Fast and Slow* (New York: Farrar, Straus & Giroux, 2011).

Я не настолько дерзок, чтобы утверждать, будто книга «Не лги себе» – библия датаизма, и не пытаюсь написать новых десять заповедей. (Хотя мне бы очень хотелось, чтобы вы воспринимали тех исследователей, о работе которых я буду говорить, как пророков датаизма: их работа носит новаторский характер именно до такой степени.)

Но я надеюсь, что эта книга покажет, каково новое мировоззрение датаизма, и предложит вам несколько алгоритмов, которые помогут принять значительное решение. «Не лги себе» состоит из глав; каждая из которых посвящена тому, что данные могут сказать нам о какой-то одной важной области жизни. И первая касается, вероятно, самого важного решения в жизни³⁰.

Итак, приверженцы датаизма и потенциальные новообращенные, давайте посмотрим, помогут ли вам алгоритмы с ответом на вопрос: «На ком жениться?»

³⁰ <https://www.wesmoss.com/news/why-who-you-marry-is-the-most-important-decision-you-make/>.

Глава 1

Брак с участием искусственного интеллекта

С кем вам следует вступить в брак?

Может быть, именно этот вопрос – самый важный в жизни с точки зрения последствий ответа на него. Инвестор и миллиардер Уоррен Баффет, во всяком случае, считает именно так. Он называет выбор супруга «самым важным решением из всех, что вам предстоит принять».

И тем не менее люди редко искали помощи у науки в этом отношении. Честно признаться, здесь наука не могла помочь практически ничем.

Ученые, занимающиеся человеческими отношениями, пытались найти ответы. Но найти большие выборки пар оказалось трудной и дорогостоящей задачей. Исследования в этой области, как правило, опирались на очень маленькие выборки и зачастую приходили к противоречащим друг другу выводам. В 2007 году выдающийся ученый Гарри Рейс из Университета Рочестера сравнил науку о человеческих отношениях с подростком, назвав ее «растущей, временами непокорной – а может, и более таинственной, чем хотелось бы»³¹.

Но несколько лет назад молодая, энергичная, сверхлюбопытная и блестяще талантливая ученая из Канады Саманта Джоэл задалась целью изменить такое положение вещей. Джоэл, как и многих ее коллег, интересовало, что же предсказывает удачные отношения. Но ее подход был необычным. Она не стала строить очередную скудную выборку пар – вместо этого она решила свести вместе данные прежних исследований. Джоэл рассуждала так: если ей удастся объединить небольшие массивы данных из научных работ прошлого, может получиться один большой массив. И его будет достаточно, чтобы надежно определить, что предсказывает успех отношений, а что нет.

План Джоэл сработал³². Она собрала в свою группу всех профессоров, располагавших данными об отношениях. В конечном счете ее группа состояла из 85 ученых и смогла аккумулировать массив данных об 11 196 парах³³.

Размер массива был впечатляющим, как и содержащаяся в нем информация.

Для каждой пары у Джоэл и ее группы были результаты опросов, где каждый из партнеров давал численную оценку степени удовлетворенности отношениями. При этом данные у них были практически по всем вообразимым аспектам совместной жизни пары³⁴, в частности:

- демографические сведения (например, возраст, образование, доход, раса);
- внешний вид (например, насколько высоко они оценивали привлекательность своего партнера);
- предпочтения в сексе (например, насколько часто каждый из партнеров хочет секса? Насколько традиционным должен быть этот секс?);
- интересы и хобби;
- умственное и физическое здоровье;
- ценности (например, взгляды на политику, отношения и воспитание детей);
- и многое-многое другое.

³¹ Harry T. Reis, “Steps toward the ripening of relationship science”, *Personal Relationships* 14 (2007): 1–23.

³² Samantha Joel et al., “Machine learning uncovers the most robust self-report predictors of relationship quality across 43 longitudinal couples studies”, *PNAS* 117(32): 19061–71.

³³ Исследование рассматривало только гетеросексуальные пары.

³⁴ Переменные, по которым проводилось исследование, можно найти здесь: <https://osf.io/8fzku/>. Соответствующий файл – Master Codebook With Theoretical Categorization, Final.xlsx, который находится в разделе “Master Codebook with Theoretical Categorization”. Выражаю благодарность Джоэл, указавшей мне на него.

Кроме того, у Джоэл и ее группы не просто было больше данных, чем у других полевых исследователей. Они использовали более прогрессивные статистические методы. Джоэл и некоторые ее коллеги превосходно овладели машинным обучением – подмножеством искусственного интеллекта, позволяющим современным исследователям обнаруживать неочевидные закономерности в больших массивах данных. Проект Джоэл можно было бы назвать браком с участием искусственного интеллекта: он был одним из первых, где на службу задаче прогнозирования успешных отношений были поставлены столь современные методики.

Если вам нравятся викторины и другие игры на отгадывание, можете попытаться предсказать ее результаты. Как вам кажется, какой фактор лучше всего предсказывает успешные отношения? Действительно ли общие интересы важнее общих ценностей? Насколько важна сексуальная совместимость в долгосрочном плане? Будете ли вы счастливее в союзе с человеком одного с вами происхождения?

Джоэл подобрала людей в исследовательскую группу, собрала данные и проанализировала их – и наконец была готова предъявить миру один из самых важных результатов в науке об отношениях.

На октябрь 2019 года Джоэл назначила свое выступление в Университете Ватерлоо в Канаде³⁵. Оно было озаглавлено незамысловато: «Можно ли помочь людям подобрать более подходящего партнера для романтических отношений?»

Итак, могла бы Саманта Джоэл, объединив усилия с восьмьюдесятью пятью самыми известными учеными мира, собрав воедино данные из сорока трех научных исследований, получив значения сотен переменных путем опроса свыше десяти тысяч пар, применив находящиеся на острие прогресса модели машинного обучения, – могла бы она, проделав все это, помочь людям выбрать более подходящего романтического партнера?

Нет.

Первый – и самый удивительный – урок, извлеченный ею из данных, как сама Саманта сказала мне во время интервью в Zoom³⁶, «заключается в том, насколько непредсказуемыми представляются отношения».

Джоэл и ее соавторы обнаружили, что демографические показатели, предпочтения и ценности двух людей практически ни в какой мере не позволяют предсказать, насколько те будут счастливы в романтических отношениях.

Вот так обстоят дела, друзья мои. Сегодня искусственный интеллект может:

- обыгрывать даже самых одаренных людей в шахматы и го;
- надежно предсказывать общественные беспорядки за пять дней до их начала, опираясь только на переписку в Интернете³⁷;
- сообщать людям о том, какие проблемы со здоровьем их ожидают, исходя из естественного запаха их тела³⁸.

Попросите при этом искусственный интеллект определить, смогут ли два человека счастливо жить вместе – и он окажется беспомощным, как и все мы.

Постойте... но ведь это выглядит настоящим провалом. К тому же это кошмарно плохое начало для главы моей книги, дерзко утверждающей, что наука о данных способна произвести революцию в принятии нами решений относительно своей жизни. Неужели ей в самом деле

³⁵ <https://www.psychology.uwo.ca/pdfs/cvs/Joel.pdf>.

³⁶ Интервью с Самантой Джоэл я провел 24 сентября 2020 года.

³⁷ Ed Newton-Rex, “59 impressive things artificial intelligence can do today”, Business Insider, May 7, 2017, <https://www.businessinsider.com/artificial-intelligence-ai-most-impressive-achievements-2017-3#security-5>.

³⁸ Bernard Marr, “13 mind-blowing things artificial intelligence can already do today”, Forbes, November 11, 2019, <https://www.forbes.com/sites/bernardmarr/2019/11/11/13-mind-blowing-things-artificial-intelligence-can-already-do-today/#4736a3c76502>.

нечего предложить нам в части выбора романтического партнера – самого важного решения в жизни?

Не совсем. На самом деле из проекта машинного обучения Джоэл и ее соавторов можно извлечь полезные уроки, пусть даже способность компьютеров предсказывать успех в любви оказалась хуже предположений.

Джоэл и ее коллектив обнаружили крайне малое влияние собранных ими переменных на перспективы успеха пары – но при этом наличие у партнера определенных параметров все же чуть увеличивает шансы, что вам с ним будет хорошо. А что еще важнее, из удивительной трудности предсказания успеха в любовных делах вытекают странные следствия, касающиеся выбора партнера.

Давайте задумаемся. Многие считают, что параметры, которые рассматривали Джоэл и ее коллектив, важны при выборе романтического партнера. Они яростно конкурируют за потенциального партнера, обладающего определенными свойствами, думая, что таким образом борются за собственное счастье. Если же свойства, за которые на рынке разворачивается ожесточенная конкуренция, не коррелируют с романтическим успехом – значит, многие выстраивают свою линию поведения в этой области неверно.

Это приводит нас к другому бесконечно древнему вопросу, к решению которого недавно тоже пытались применить принципиально новые данные: как именно люди выбирают романтического партнера?

В последние несколько лет другие исследовательские коллективы предприняли анализ данных с сайтов знакомств. Они рассматривали новые большие массивы данных о чертах характера и привычках десятков тысяч претендентов, чтобы выяснить, что создает привлекательность для романтических отношений. И этот анализ, в отличие от исследования счастья в отношениях, дал вполне осязаемый результат. Если работающие с данными ученые выяснили, что указать качества партнера, предвещающие счастье в отношениях, удивительно сложно, то определить качества, делающие нас такими же привлекательными для противоположного пола, как кошачья мята для кошек, оказалось поразительно просто.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «ЛитРес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на ЛитРес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.