

Оглавление

<i>Предисловие. Пандемия чисел</i>	9
<i>Вступление. Плененные числами</i>	17
1. Цифры могут спасти	27
2. Бессмысленный спор об IQ и цвете кожи	47
3. Что говорит о выборках одно сомнительное сексологическое исследование	79
4. Курение вызывает рак легких (но аисты детей не приносят)	107
5. Не стоит слишком увлекаться цифрами в будущем	141
6. Значимость чисел определяет наша психология	169
<i>Послесловие. Как поставить числа на место</i>	182
<i>Инструкция: что делать при встрече с числом</i>	187
Дополнительная литература	190
Выражение благодарности	193
Примечания	197

Предисловие

Пандемия чисел

Президент Дональд Трамп взял со стола пачку бумаг и передал ее своему гостю:

— Посмотрите на эти графики.

Это происходило в Белом доме. Напротив Трампа сидел Джонатан Суон, политический обозреватель новостного сайта Axios, получивший в конце июля 2020 года возможность взять у президента получасовое интервью. Вот уже почти пятнадцать минут они говорили о коронавирусе и о кризисе, который пришел вместе с ним.

— Вот, например, — сказал Трамп, досконально изучив одну из страниц. — Тут видно: Соединенные Штаты ниже всех во многих категориях. Мы ниже всего мира.

— Ниже всего мира?

Суон нахмурился, сильно сощурился глазами. Озадаченное выражение его лица впоследствии стало популярным мемом в интернете.

— Мы ниже Европы.

— Что это значит? В чем именно?

Трамп передал листок Суону:

— Смотрите. Вот тут. Это случаи смерти.

Пока Трамп искал следующий график, Суон смотрел на страницу, которую дал ему президент.

— А, так вы о смертях среди всех заболевших! А я говорю о смертности в пропорции к численности населения, — поднял брови Суон. — И по этому параметру у США очень плохие показатели — хуже, чем у Южной Кореи, у Германии, у многих других.

— Нет, так нельзя, — покачал головой Трамп.

В 2018 году, когда эта книга вышла в свет в Нидерландах, я и понятия не имела, что ждет нас в будущем. Я подумать не могла, что к 2020 году некий вирус сокрушит не только Соединенные Штаты, но и весь мир и что десятки миллионов заразятся, а более 1,8 миллиона — и это число все еще растет — умрут.

Через полтора года после первой публикации моей книги числа неожиданно оказались в центре внимания. Все только и говорили, что про экспоненциальный рост, коэффициент воспроизводства и выход на плато. В группах WhatsApp публиковались расчеты, цифры фигурировали в каждом выпуске новостей, а знание статистики могло сделать вас знаменитостью.

Числа господствуют в современных обществах уже довольно давно. Собственно, об этом я и писала. Будь то экономическая статистика, результаты экзаменов, опросы общественного мнения, «большие данные» — числа влияют на мир. Я надеялась, что из моей книги станет предельно ясно: даже если числа вам не интересны, они правят вашей жизнью.

В 2020 году все и так уже это понимали. Влияние чисел на нашу жизнь никогда не было очевиднее, чем во время пандемии. Численные данные, связанные с коронавирусом, определяли, позволят ли вам встречаться с родными, покупать алкоголь, работать в офисе, ходить в школу, приглашать гостей, путешествовать, участвовать в футбольных матчах, ходить в театр, присутствовать на похоронах... Жизнь стала совершенно иной, и причиной тому были числа. Точнее, причиной, конечно, стал вирус, а цифровизация была совершенно необходима для того, чтобы держать вирус в поле зрения.

Представьте себе мир без чисел. Сколько остается свободных коек в реанимации? Помогают ли принимаемые властями меры предотвратить распространение инфекции? Действуют ли вакцины? Мы бы не знали ответов на эти вопросы без чисел. Они приобрели жизненно важное значение.

Вот почему так важно уметь в них разбираться. Роль чисел в обществе велика настолько, что возникает соблазн злоупотреблений.

Обвинять Дональда Трампа во лжи — в этом нет ничего оригинального. За годы своего президентского срока он высказал более 30 000 ложных утверждений. Одно из особенно памятных — что инъекции дезинфицирующих средств якобы могут помочь в борьбе с последствиями COVID-19. А еще Трамп говорил, что при его администрации американская экономика стала лучшей за всю мировую историю, и уверял, что стена на границе с Мексикой уже практически построена.

Не будем забывать и о том, что, едва став президентом, он уже солгал о количестве зрителей на инаугурации. «Там, кажется, был миллион или даже полтора миллиона», — сказал Трамп в первый же день своего президентства. Трудно сказать точно, сколько людей пришло на Национальную аллею¹ — фотографии со спутников не получились из-за плотных облаков, — но специалисты согласны, что там и близко не было ни миллиона, ни тем более полутора. А советник президента Келлиэнн Конуэй, защищавшая его ложь по телевидению, произнесла при этом историческую фразу. Это не ложь, заявила она, а «альтернативные факты».

Как обозреватель по вопросам числовой грамотности я нечасто сталкивалась с подобными злоупотреблениями — когда числа просто выдумывают. Но некоторым хватает дер-

¹ Национальная аллея (англ. National Mall) — комплекс парков, памятников и музеев в Вашингтоне, в восточном конце которого расположено здание Капитолия. — *Здесь и далее, если не указано иное, постраничные примеч. перев.*

зости оперировать цифрами «с потолка» и — что еще серьезнее — утверждениями, которые всегда обладают большей убедительностью, если кажутся научными. Тут-то и пригождаются числа.

Это можно увидеть в разговоре Трампа с Джонатаном Суоном. Очевидно, кто-то в Белом доме распечатал пачку графиков статистической платформы Our World in Data¹ при Оксфордском университете, известной своими превосходными данными, и положил их на стол рядом с креслом Трампа. И даже Дональд Трамп, олицетворение «альтернативных фактов», подкрепляет свою позицию числами — реальными числами.

Числа помогают нам убеждать других в том, что мы правы. С ними наши суждения обретают впечатление научной обоснованности. Числа же не лгут, не правда ли? Измерить — значит познать. Слова подвержены предвзятости, а числа, как нам кажется, отражают реальность как она есть. Но какими бы важными ни оказались те или иные числа, моя книга показывает, что они никогда не бывают объективными. Число утрачивает свою объективность уже в момент измерения. Решения о том, что и как вы измеряете, субъективны по определению.

Взять, к примеру, действия нидерландского правительства по борьбе с коронавирусом. 19 мая 2020 года министр здравоохранения Уго де Йонг объявил о создании сводной панели численных данных по коронавирусу, которая должна была служить «основой для принятия решений». Регистрировать следовало все — положительные результаты тестирования, факты госпитализации, случаи заражения в домах престарелых... Исходя из такой медицинской статистики, принимались новые меры — закрывались магазины, торгующие товарами не первой необходимости, рестораны и школы. И борьба с коронавирусом приводила к тому, что жители Нидерландов начинали больше курить, теряли работу и впадали в депрессию.

¹ «Наш мир в данных» (англ.).

В сводной панели не отражались цифры, описывающие образ жизни, уровень занятости или психическое здоровье. Предпочтение отдавалось другим числам, чтобы избежать перегрузки системы здравоохранения и обеспечить защиту наиболее уязвимых. Выбор понятен — но все же это лишь выбор. И речь здесь не только о статистике, но и об этике.

Это не значит, что числа — лишь чье-то мнение. Они подобны законам: это субъективные соглашения, которые можно подвергнуть тщательным испытаниям. Если договориться о том, что прогресс, интеллект или здоровье можно измерить, и о том, как это следует делать, то мы можем рассматривать изменения этих измерений с течением времени, факторы, которые на них влияют, и результаты, которых достигает наша страна по сравнению с другими.

Если числа используются правильно, они помогают нам познавать мир и изменять его. Но что означает «правильно используются»? Здесь в дело вступает статистика — методы стандартизации, сбора и анализа численных данных.

В этой книге я надеюсь убедить вас в том, что случайная выборка лучше, чем опрос пользователей твиттера; что корреляция — это не причинно-следственная связь; что погрешности необходимо учитывать и что утверждение Трампа о смертности от коронавируса в США было абсолютно неверным.

— Почему так нельзя? — спросил президента Суон.

— Нужно смотреть по... Вот, смотрите. — Трамп протянул журналисту другую страницу. На ней была диаграмма с ярко раскрашенными столбцами. — Вот Соединенные Штаты. Нужно смотреть по случаям заболевания. Вот эти случаи.

Трамп продолжал говорить, но Суон попытался спросить еще раз:

— А почему не в пропорции к численности населения?

Утверждение Трампа не было чистой выдумкой, как полтора миллиона присутствовавших на его инаугурации. Он смотрел на коэффициент летальности (Case Fatality Rate, CFR), то

есть на процентную долю умерших от болезни среди всех тех, кто получил положительные результаты тестирования. Это хорошо известный эпидемиологический параметр, позволяющий приблизительно оценить смертность от того или иного заболевания. К концу июля 2020 года, когда состоялось это интервью, значение CFR для Европы составляло 7,0 процента, а для всего мира — 3,9 процента. А для Соединенных Штатов? 3,5 процента — действительно «ниже», как и утверждал Трамп.

И все же что-то в его рассуждениях было не так. К счастью, Суону хватило проницательности сразу же это заметить. Низкое значение CFR не означало, что в США все в порядке. Даже если эта доля мала, количество умерших будет большим, если большим будет количество инфицированных. Поэтому Суон хотел увидеть, как соотносится число умерших с численностью всего населения. В конце июля смертность в Европе находилась на уровне 264 случая на миллион жителей, а в среднем по миру — 84. В США смертность составляла 453 случая на миллион человек. Она превышала европейскую более чем в полтора раза, а общемировую — более чем в пять раз. Если вкратце, то ситуация была прямо противоположной тому, что подразумевает слово «ниже».

Эта книга разоблачает злоупотребления именно такого рода. Если лучшая ложь содержит крупицу истины, то и числовой обман точно так же труднее всего раскрыть, если он связан с реальными цифрами. Такая фальшь бывает достаточно тонкой, чтобы остаться незамеченной, и достаточно убедительной, чтобы ее можно было распространять.

В своей книге я предложу вам средства, позволяющие распознать такой обман. Для этого вам не понадобится ученая степень по статистике — хватит и толики любопытства и здравого смысла. Не мешает и готовность признать свои недостатки.

— А некоторые говорят, что слишком много тестировать вредно. Вы ведь знаете об этом? — сказал Трамп Суону несколько раньше в том же интервью.

— Кто так говорит?

— Да почитайте книги, почитайте учебники.

— Учебники?

Через несколько минут Трамп сформулировал ту мысль, которую хотел высказать:

— Дело в том, что мы тестируем гораздо лучше, чем любая другая страна. Поэтому у нас больше случаев заболевания.

Здесь Трамп сделал с темой тестирования на коронавирус то же, что и с темой смертности: он начал подбирать доводы, которые подтверждали его точку зрения. Числа связаны не только со статистикой, но и с психологией. И этим грешит не один Трамп. Все мы так делаем.

Возможно, вы с надеждой ловите каждую новость о количестве вакцинированных, мечтая, чтобы все это наконец закончилось; может быть, вас раздражает исследование эффективности масок, потому что вам кажется, что это ограничение вашей свободы; или же вас печалит статистика смертности от коронавируса в вашей стране, потому что вы привыкли гордиться своим отечеством.

Все мы встречаем и такие факты, которые радостно принимаем на веру, и такие, которые мы предпочли бы немедленно отвергнуть.

Даже если уловки Трампа в этом интервью вас не обманули, вполне вероятно, что в каком-нибудь другом случае вы на них попадетесь. В марте 2020 года в Нидерландах получила широкое распространение статистическая таблица, в которой смертность от коронавируса в Италии, сильно пострадавшей в то время, сравнивалась с аналогичными данными по Нидерландам. Вывод ужасал: мы идем тем же путем, что и Италия!

Однако при анализе данных оказалось, что они отобраны и подогнаны именно к такому выводу. Каждый из двух наборов чисел начинался с некой произвольной даты, выбранной так, что при экстраполяции итальянских данных и их сравнении с положением в Нидерландах картина действительно

получалась тревожной. Но стоило взять данные с другими, более логичными начальными датами, как траектория развития пандемии в Нидерландах оказывалась совершенно отличной от итальянской и менее пугающей. Именно такой тенденциозный подбор данных и использовал Трамп в своем интервью.

А как же довод о том, что число выявленных заболеваний увеличивается из-за более интенсивного тестирования? Он был широко распространен в Нидерландах летом 2020 года, когда пришла вторая волна пандемии коронавируса. Этот довод был не уловкой, а вполне обоснованным аргументом: то, как проводится тестирование, действительно влияет на статистику. Однако это не причина не принимать вторую волну всерьез; просто нужно быть еще осторожнее в истолковании данных. Я сама писала об этом в своих статьях. Но, услышав подобное из уст Трампа, я просто рассмеялась и не стала разбираться в его доводах. Ох уж этот мне чудак со своими учебниками! Мне тоже нужно держать себя в руках.

Пандемия показала нам числа как с лучшей, так и с худшей стороны, еще раз продемонстрировав и то, насколько они важны, и то, как ими можно злоупотреблять ради личной выгоды. Я привожу ряд примеров в этом дополненном издании.

Числа дают нам опору и уверенность в собственных силах. Это подбадривает, особенно в такие неопределенные времена. Но у чисел есть и границы: они не могут ни точно предсказывать будущее, ни говорить нам, какие принимать решения. Надеюсь, эта книга поможет вам увидеть цифры в истинном свете, поскольку они играют важную роль в нашей жизни сегодня и будут играть ее завтра и еще долго после окончания пандемии.

Утрехт, декабрь 2020 года

Вступление

Плененные числами

Она зашла в пыльный кабинет сквозь раздвижные двери и пожала мне руку:

— Хуанита ^[1].

Полинявший свитер был ей велик, отчего она, и без того невысокая, казалась еще меньше. Когда Хуанита села на складной стул напротив меня, я объяснила по-испански, что приехала в Боливию по заданию нидерландского университета, провожу исследование уровня счастья и имущественного расслоения и хочу задать ей несколько вопросов о ее собственной жизни и стране, в которой она живет.

Я произносила эту заученную речь уже не в первый раз. Десять дней я беседовала с жителями боливийского города Тарихи, расположенного недалеко от аргентинской границы. Я разговаривала с рыночными торговками, пила пиво с фермерами, выращивающими клубнику, ходила на семейные барбекю — лишь бы собрать как можно больше данных. Теперь я добралась со своей пачкой анкет до офиса женской организации, руководство которой предложило связать меня с *empleadas domésticas* — домработницами. Женщинами вроде Хуаниты.

— Начнем, — сказала я. — Сколько вам лет?

— Пятьдесят восемь.

— К какой этнической группе вы принадлежите?

— Аймара.

«Ага, — подумала я, — значит, она из одного из индейских племен». С такими людьми я раньше не встречалась.

— Семейное положение?

— Не замужем.

— Вы умеете читать?

— Нет.

— А писать?

— Нет.

Я продолжила задавать вопросы в том же духе — профессия, образование, есть ли у нее мобильный телефон, холодильник или телевизор.

— Я зарабатываю двести боливиано в месяц, — ответила она, когда я спросила о ее зарплате. Это было гораздо меньше минимальной заработной платы — 815 боливиано, которую установил незадолго до того президент Эво Моралес.

— Я боюсь, что, если попрошу добавить мне денег, начальник меня уволит. Я живу в *carpita*.

Я записала это слово, но не знала, что оно означает, пока не посмотрела в словаре. Оказалось, она жила в палатке.

В конце концов я добралась до главной темы своего исследования — уровня счастья и имущественного расслоения. Еще сидя за своим столом на одиннадцатом этаже в Роттердамском университете Эразма, я подготовила в PowerPoint пять диаграмм. Они соответствовали разным вариантам распределения доходов.

В первый же день исследовательской работы в Боливии я заметила, что мои вопросы об имущественном расслоении подходят не всем. Рыночные торговки, с которыми я беседовала, не понимали, что означают мои диаграммы. Могла ли я ожидать, что Хуанита, которая не умела ни читать, ни писать, поймет мой вопрос? Я решила пропустить эту часть анкеты.

Однако она, опередив меня, заговорила.

— Знаете, что не так с Боливией? — Хуанита выпрямилась на стуле. — Есть очень большая группа бедных и очень маленькая группа богатых. И разница между ними только растёт. Неужели вас удивляет, что в этой стране больше никто никому не доверяет?

Сама того не зная, моя собеседница описала диаграмму А. При этом она еще и ответила на два других моих вопроса — о ее прогнозах на будущее и уровне доверия в ее стране. Я очень сильно недооценила ее. Я покраснела, но продолжила разговор как ни в чем не бывало. Мы дошли до последнего вопроса.

— Оцените по шкале от одного до десяти, насколько вы счастливы.

— Один.

— А насколько, как вы считаете, вы будете счастливы через пять лет?

— Один.

Мне кажется, именно во время этого разговора в 2012 году у меня зародились первые сомнения относительно чисел. До тех пор я в основном была потребителем чисел. Они встречались мне, когда я читала газеты или смотрела телевизор. Когда я училась в университете и преподаватели давали мне задания по эконометрике, я получала файлы с числами или скачивала официальные данные с сайтов разных организаций, например Всемирного банка.

Но на этот раз у меня не было заранее заполненной таблицы. Теперь данные собирала я. У меня уже был за плечами первый год аспирантуры; числа стали моей специальностью. Но разговор с Хуанитой поколебал мою уверенность. Я занималась исследованием уровня ее счастья, но обнаружила, что никакое число не может описать ее жизнь в палатке. Я слушала ее рассуждения об имущественном неравенстве, но могла только выбрать одну из диаграмм, А, В, С, D или Е. Зна-

чительную часть того, что она мне рассказала, нельзя было подсчитать, но было необходимо учесть.

Хуанита помогла мне понять и еще одну вещь. Я сильно влияла на то, как выглядели мои числа. Это я решила, что уровень счастья — важный параметр и его можно выразить в числах. Это я придумала задавать этот абстрактный вопрос, используя диаграммы. Это я подумала, что Хуанита недостаточно умна, чтобы сказать что-нибудь полезное относительно имущественного неравенства. Я, я, я. Кто-нибудь другой, используя те же исследовательские вопросы, но придерживающийся другого мировоззрения или мнения, вероятно, получил бы другие результаты. Считается, что числа объективны, но я внезапно поняла, насколько сильно они связаны с личностью исследователя.

После разговора с Хуанитой я внесла ее данные в 80-ю строку своей таблицы в Excel: возраст — 58, зарплата — 200, уровень счастья — 1. Таблица выглядела такой же ясной и аккуратной, как все остальные таблицы, с которыми я работала целыми годами. Но я внезапно осознала, насколько обманчивы мои упорядоченные строки и столбцы чисел.

Я пристрастилась к числам еще в раннем детстве. Едва начавшись считать, я особенно полюбила книжки, где нужно было соединять пронумерованные точки. Одно из моих первых воспоминаний относится к каникулам, на которые мы поехали в немецкий Шварцвальд: я проводила бесчисленные часы, соединяя точки, чтобы из них получались бесконечные снеговики и облачка. Несколько лет спустя дедушка с бабушкой подарили мне будильник с радиоприемником. По ночам я лежала в кровати, глядя на светодиодный дисплей часов и составляя в уме всевозможные арифметические операции с этими четырьмя цифрами. В школе математика была моим любимым предметом; потом я продолжила заниматься ею и в конце концов защитила диссертацию по эко-

нометрике. Я узнала все что можно о статистических основах экономических моделей. Я вычисляла, анализировала, программировала. И, как оказалось, я снова занималась тем же, что делала в свое время, соединяя точки: выискивала закономерности.

Но числа играли в моей жизни и другую роль. Они помогали мне найти мое место. С пяти до двадцати шести лет, в школе и в университете, мне ставили оценки. Они служили мне мерой моих успехов. Если я получала низкую оценку, я вешала нос. Высокая оценка приводила меня в восторг. Меня не беспокоило, что всего через несколько дней я забывала выученное, если только у меня получался достаточно хороший средний балл. Числа помогали мне ориентироваться и вне школы. Вернувшись из Боливии, я встала на весы: 56 килограммов. Я знала, что это соответствует индексу массы тела, равному 18,3, — и была очень довольна собой.

Не я одна опиралась в жизни на числа. Мои коллеги по университету продвигались по службе, если им удавалось напечатать в научных журналах много статей. Персонал больницы, в которой работала моя мама, с трепетом ожидал публикации ежегодного рейтинга 100 лучших больниц. А моего папу отправили на пенсию в тот самый день, когда ему исполнилось шестьдесят пять.

Лишь потом я поняла, что разговор с Хуанитой открыл мне нечто важное относительно чисел такого рода. Если я оказываю влияние на те числа, которые я собираю, то и другие точно так же влияют на числа, которыми руководствуются в жизни окружающие меня люди. Учителя рассчитывают проходной балл для экзаменов, врачи вычисляют оптимальное значение ИМТ¹, власти определяют, в каком возрасте вам следует перестать работать.

¹ Индекс массы тела. — *Примеч. ред.*

Научно-популярное издание

Блау Санне

Во власти цифр

Как числа управляют нашей жизнью
и вводят в заблуждение

Ответственный редактор *Н. Галактионова*

Редактор *В. Измайлов*

Художественный редактор *М. Левыкин*

Технический редактор *Л. Синицына*

Корректоры *Н. Соколова, Т. Дмитриева*

Верстка *Т. Коровенковой*

ООО «Издательская Группа «Азбука-Аттикус» —
обладатель товарного знака «КоЛибри»
115093, Москва, ул. Павловская, д. 7, эт. 2, пом. III, ком. № 1
Тел. (495) 933-76-01, факс (495) 933-76-19
E-mail: sales@atticus-group.ru

Филиал ООО «Издательская Группа «Азбука-Аттикус»
в г. Санкт-Петербурге
191123, Санкт-Петербург, Воскресенская набережная,
д. 12, лит. А
Тел. (812) 327-04-55
E-mail: trade@azbooka.spb.ru

www.azbooka.ru; www.atticus-group.ru

Знак информационной продукции
(Федеральный закон № 436-ФЗ от 29.12.2010 г.)

18+

Подписано в печать 22.03.2022. Формат 60×90/16.

Бумага офсетная. Гарнитура «Charter».

Печать офсетная. Усл. печ. л. 14,0.

Тираж 3000 экз. В-СНМ-27248-01-Р. Заказ №

Отпечатано в соответствии с предоставленными материалами
в ООО «ИПК Парето-Принт». 170546, Тверская область,
Промышленная зона Боровлево-1, комплекс № 3А
www.pareto-print.ru