

Дарелл Хафф
Как лгать при помощи статистики

ДАРЕЛЛ ХАФФ
**КАК ЛГАТЬ
ПРИ ПОМОЩИ
СТАТИСТИКИ**



Полезная книга
и для людей, далеких
от статистики, и для
тех, кто в теме.
Билл Гейтс

[ПО ДАННЫМ СТАТИСТИКИ, В МИРЕ ПРОДАНО
7 299 792 458 ЭКЗЕМПЛЯРОВ ЭТОЙ КНИГИ]

 альпина
паблишер

«Как лгать при помощи статистики / Дарелл Хафф»: Альпина
Паблишер; Москва; 2015
ISBN 978-5-9614-3984-6

Аннотация

В этой всемирно известной книге Дарелл Хафф рассказывает о различных способах злоупотребления статистикой в целях обмана аудитории и манипулирования ее мнением. Каждый день на вас пытаются повлиять, чтобы сподвигнуть на покупку какого-то «нужного» продукта или на выбор «правильного» кандидата: «Благодаря пасте “Чистые зубы” образование кариеса снижается на 23 %!»; «Политика N поддерживает 85 % граждан»... Как понять, насколько достоверны те или иные данные? Каким образом происходят подсчеты? Что учитывается, а что остается за кадром? Автор раскрывает секретные инструменты статистиков и вооружает читателя знаниями, которые помогут разобраться во всех хитросплетениях этой науки и не позволят ввести в заблуждение.

Дарелл Хафф Как лгать при помощи статистики

Переводчик Е. Лалаян

Редактор А. Черникова

Научный редактор В. Ионов

Руководитель проекта А. Деркач

Корректор Е. Аксёнова

Компьютерная верстка К. Свищёв

Дизайн обложки Ю. Буга

© Darrell Huff and Irving Geis, 1954

© Издание на русском языке, перевод, оформление. ООО
«Альпина Паблишер», 2015

Все права защищены. Произведение предназначено
исключительно для частного использования. Никакая часть

электронного экземпляра данной книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами, включая размещение в сети Интернет и в корпоративных сетях, для публичного или коллективного использования без письменного разрешения владельца авторских прав. За нарушение авторских прав законодательством предусмотрена выплата компенсации правообладателя в размере до 5 млн. рублей (ст. 49 ЗОАП), а также уголовная ответственность в виде лишения свободы на срок до 6 лет (ст. 146 УК РФ).

* * *

Существуют три вида лжи: ложь, наглая ложь и статистика.

Бенджамин Дизраэли

Придет время, и статистическое мышление станет таким же необходимым качеством для истинного гражданина, как умение читать и писать.

Герберт Уэллс

Нам досаждают не столько те вещи, о которых мы не знаем, сколько те, о которых мы знаем, что с ними что-то не так.

Артемус Уорд

Круглые числа всегда лгут.

Сэмюэл Джонсон

У меня есть обширная тема [статистика] и есть много, что написать по этой теме, но со всей остротой я осознаю, что мне не хватит литературных талантов, чтобы изложить ее просто и доходчиво, не жертвуя при этом точностью и основательностью.

Сэр Фрэнсис Гальтон

К читателю

Будь моя воля, я бы назвал эту книгу еще короче – «Как лгать», потому что ложь приобрела убедительность, логику и, что еще важнее, цифры, за которыми может скрываться все что угодно в «умелых руках». А «умелых рук» очень много.

В наше время вопросы лжи и правды все так же актуальны. Помимо прямого обмана появилось множество способов «огибать правду» или же показывать реальность таким образом, что даже знающему человеку сложно распознать скрывающуюся за ней ложь.

В искажении статистики заинтересованы все, кто ищет способы исказить общественное мнение и воспользоваться этим в целях собственного обогащения. Немало и таких, кто хочет скрыть настоящие цифры, потому как они отражают крайне неприглядные факты. Наконец, статистика оказывается объектом прямого манипулирования во всех случаях, когда она является частью процессов принятия решений государственного масштаба.

В России ситуация со статистикой никогда не была столь печальной, как сейчас. Если в 80-х и 90-х гг. прошлого столетия официальная статистика в России страдала от тотального недофинансирования, то сегодня сами принципы государственного и муниципального управления в России таковы, что статистика стремительно превращается из инструмента доверия в инструмент распределения государственных средств.

В последних исследованиях¹ Фонда «Хамовники»² Ольга Моляренко очень подробно разобрала примеры искажения муниципальной статистики в России. Из-за отсутствия кооперации между органами власти, использования статистических данных как основы для принятия решений о выделении бюджетных средств и многих других российских особенностей мы оказываемся перед острой необходимостью реорганизации сбора государственной статистики в целом.

Книга Даррелла Хаффа хороша не выводами и даже не огромным

¹ Муниципальная статистика и проблемы сбора информации местной властью// Вестник Новосиб. гос. ун-та. Серия: Социально-экономические науки. 2014. Т. 14, вып. 4. С. 125–140

² Фонд поддержки социальных исследований, <http://khamovniky.ru/>

числом примеров, а тем, что она учит критическому мышлению, она учит отношению к цифрам не как к «сакральному знанию», а как к инструменту, с помощью которого осуществляется манипулирование нашим мнением.

И я могу сказать, что именно критического взгляда нам остро не хватает в последние годы. Вот лишь один пример. Одна общественная организация в России решила публиковать свой рейтинг восприятия коррупции. Дабы придать этому рейтингу «научности», в качестве критериев было решено использовать абсолютные статистические показатели, такие как статистика преступлений, публикуемая МВД и Генеральной прокуратурой. Хотя благое намерение отслеживать ситуацию с коррупцией в нашей стране можно только приветствовать, сам подход является ошибочным, потому как ошибочны изначально заложенные в нем метрики.

В исследовательском отчете «Криминальная статистика: механизмы формирования, причины искажения, пути реформирования»³, подготовленном сотрудниками Института проблем правоприменения, очень подробно разобраны и описаны проблемы и последствия искажения правовой статистики в Российской Федерации. Для человека, вооруженного выводами этого исследования, совершенно очевидно, что рейтинги, построенные на современной правовой статистике, не могут быть достоверными.

А ведь рейтинги – это лишь один из тысяч продуктов и примеров применения статистических показателей.

Официальная статистика используется как минимум:

- при определении бюджетных субсидий;
- при оценке деятельности публичных компаний;
- при выделении финансирования бюджетным учреждениям;
- в городском и районном планировании;
- в финансово-экономическом обосновании государственных и частных инвестиций;
- в оценке эффективности государственных программ;
- в построении различных рейтингов.

³ Исследовательский отчет «Криминальная статистика: механизмы формирования, причины искажения, пути реформирования» (М. Шклярук, Д. Скугаревский), <http://www.enforce.spb.ru/products/books/6499–2015-mar-18-12-24-21>

В результате даже малая ошибка в показателях превращается в просчет в фундаменте, на котором строится огромное число умозаключений и решений.

Другой важной проблемой существующей практики статистического учета является технологическое отставание современной статистики. В ситуации, когда государство все более автоматизируется и сбор многих показателей не требует выборочных опросов организаций или их тотальной переписи, все острее встает вопрос автоматизации извлечения статистических данных из государственных информационных систем.

Книгу Даррелла Хаффа можно порекомендовать как тем, кто хочет докопаться до правды, так и тем, кто ищет способы ее скрыть или осознанно исказить. Она – подспорье для всех, кто часто участвует в дискуссиях о достоверности цифр. Она поможет и тем, кто ищет способ придать недостоверным данным больше «легальности».

Хотя в ней отсутствует многое из того, к чему мы привыкли в современном мире, – нет ни слова об информационных системах, о возможностях Интернета, о больших данных, тем не менее эта классическая книга все так же актуальна.

Статистика по-прежнему остается областью интерпретаций и относительных оценок, описание цифр и форма их представления имеют колossalное значение, а примеры, приводимые в этой книге, не потеряют своего значения еще не один десяток лет.

Я рекомендую эту книгу не только тем, кто хочет разбираться в статистике, но и тем, кто интересуется правдой и ложью, а также логикой, на которой они могут быть построены. Книгу, без сомнения, стоит прочитать журналистам, пишущим об экономике и любой отрасли, где важны цифры и оценки, сотрудникам пресс-служб и всем, кто работает со статистикой в любой форме.

*Иван Бегтин,
директор некоммерческого партнерства
«Информационная культура», член общественного
совета при Федеральной службе государственной
статистики*

Предисловие к русскому изданию

Человеку свойственно ошибаться, и особенно серьезные ошибки случаются, когда умозаключения строятся на основе статистических данных, «холодных цифр». Наше сознание странным образом признает за математикой право на абсолютную истину.

Статистика, как рассказывает нам Даррелл Хафф в своем бестселлере «Как лгать при помощи статистики», это такая хитрая, закамуфлированная отрасль математики. С одной стороны, она оперирует цифрами, пользуется четкой логикой и понятными методами расчетов. С другой стороны, предметом той статистики, которая нас обманывает, всегда является поведение человека (или отношение человека к чему-то, вплоть до отношения к другим людям). Цифры продают нам нас же самих, завернутых в формулы, статистические распределения и байесовские множества.

Мы видим цифры, мы видим математический авторитет тех, кто ими пользуется, и мы беззащитны перед статистикой и манипуляторами, как беззащитен первоклассник перед учителем, который доказывает, что $2 + 2 = 5$.

Статистика (и это, наверное, самая интересная часть книги Хаффа) настолько злокозненна, что регулярно обманывает и теоретиков, и прикладных исследователей, и тем более политиков, которые очень любят оперировать ею. Используя исторические примеры (от «соломенных опросов», чуть не разрушивших карьеру Дж. Гэллапа в 1930-х гг., до дискуссий о глобальном потеплении), Даррелл Хафф не только помогает читателю разобраться в прошлых обманах, но и дает ему инструмент проверки на будущее.

За последние 25 лет в большинстве ведущих университетов мира статистика и ее методы стали обязательным компонентом любого образования, включая самое что ни на есть гуманитарное, по той причине, что ученому, практику, юристу и даже филологу нужно иметь ту самую «бритву Оккама», которой рассекается любая путаница. В нашем веке «больших данных» это особенно важно – в бесконечном океане собираемых цифр очень важно избегать как базовых ошибок (о чем подробно рассказывает книга Даррелла Хаффа), так и некритического подхода к любому анализу, представляемому как мнение большинства. Большинство, увы,

не ведет нас вперед, а упорно держится за то немногое, что у него есть, оправдывая свой «консерватизм» именно ложным выводом из ложной же статистики.

*Василий Гатов,
приглашенный исследователь Центра
коммуникационного лидерства и политики (Школа
коммуникаций и журналистики Университета Южной
Калифорнии)*

*Моей супруге посвящаю. С полным на то
основанием*

Благодарности

Милые примеры откровенных нелепостей и надувательских ухищрений, которыми, как перчинками, приправлена эта книга, я собирал где только можно и не без посторонней помощи. Откликнувшись на мой призыв, посланный через Американскую статистическую ассоциацию, несколько профессиональных статистиков – а они, уж вы мне поверьте, оплакивают ненадлежащее использование статистики не менее искренне, чем все прочие, – поделились со мной примерами из своих личных коллекций. Эти достойные люди, думается мне, будут только рады, если их имена останутся неназванными на этих страницах. Ценные примеры я почерпнул также из целого ряда книг, главным образом из следующих: «Бизнес-статистика» (Business Statistics) Мартина Брумбауха и Лестера Келлогга, «Как измерять общественное мнение» (Gauging Public Opinion) Хэдли Кэнтрила, «Графическое представление данных» Уилларда Бринтона, «Практическая бизнес-статистика» (Practical Business Statistics) Фредерика Крокстона и Дадли Коудена, «Основы статистики» (Basic Statistics) Джорджа Симпсона и Фрица Кафки, «Простейшие статистические методы» (Elementary Statistical Methods) Хелен Уокер.



Введение

«Что-то было много преступности в этих краях», – заметил мой свекор вскоре после того, как переехал из Айовы в Калифорнию. Так оно и было – если верить газете, которую он читал. Газета была того сорта, что не пропустит ни единого преступления в собственной округе, и еще она славилась тем, что какому-нибудь убийству в Айове уделяла гораздо больше внимания, чем ведущая ежедневная газета той местности, где собственно и произошло убийство.

Вывод моего свекра был статистического свойства, если не по форме, то по существу, и основывался на выборке, причем необычайно предвзятой. Как и многие образцы куда более изощренно сфальсифицированных статистических данных, его вывод грешил мнимой обоснованностью и исходил из той посылки, что место, отведенное газетой под криминальную хронику, и есть мерило уровня преступности.

Несколько лет назад с дюжину исследователей независимо друг от друга опубликовали данные об антигистаминных препаратах. Во всех упоминалось о значительном проценте излечившихся от простудных заболеваний после приема этих лекарств. Началась большая шумиха (по крайней мере рекламная), и спрос на эти препараты резко вырос. Он был основан на вечных упованиях на чудо и еще на удивительном нежелании отвлечься от статистики и взглянуть в глаза давно известной истине. Как когда-то говорил писатель-юморист Генри Фелсен (весьма далекий от медицины), при должном лечении простуда проходит через семь дней,

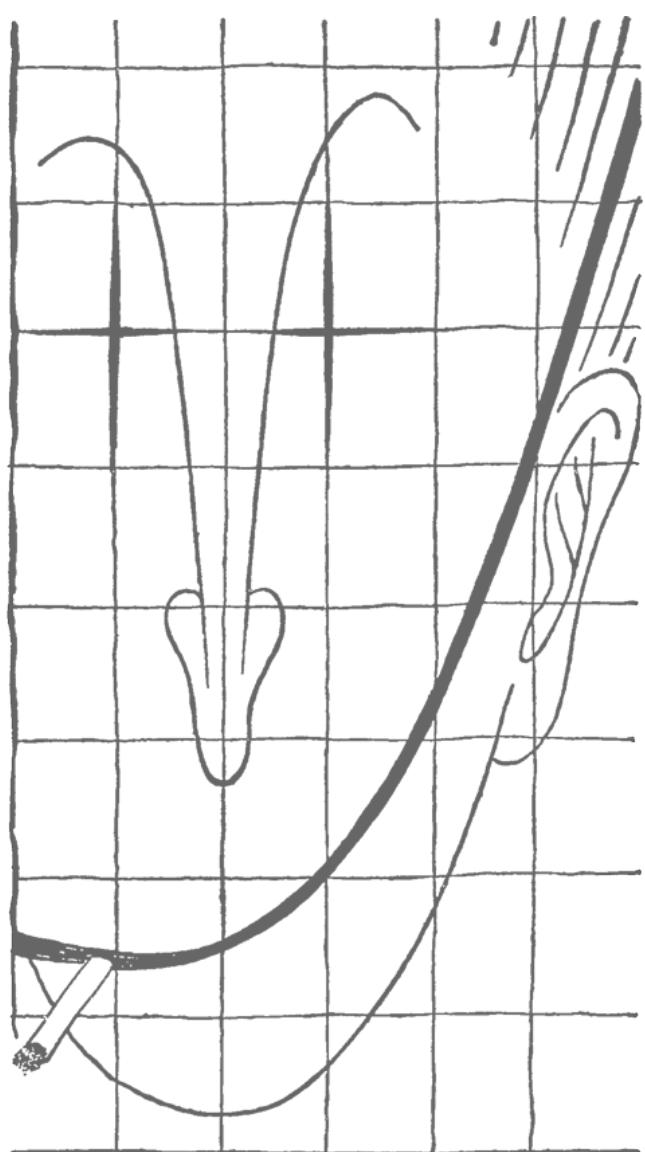
в противном случае она сама собой пройдет через неделю.

Точно так же обстоят дела со многим из того, что вы читаете и слышите. Средние величины, зависимости, тенденции и графики не всегда есть то, чем кажутся. Подчас в них таится много больше интересного, чем видно на первый взгляд, а иногда и куда как меньше.

Таинственный язык статистики, столь притягательный в условиях культуры, которая ставит во главу угла факты, используют для того, чтобы создавать сенсации, преувеличивать, сбивать с толку и чрезмерно упрощать. Статистические методы и термины необходимы, когда дело касается массовых данных о социальных и экономических тенденциях, деловой конъюнктуре, опросах общественного мнения, переписях населения. Но в отсутствие авторов, которые используют статистическую терминологию добросовестно и со знанием дела, равно как и читателей, понимающих, что означают все эти термины, результатом может стать та еще ахинея.

В современной научно-популярной литературе ругаемый на все корки статистик почти вытеснил образ самоотверженного героя-труженика в белом халате, который дни и ночи корпит над своими пробирками в неверном свете лабораторных ламп, даже не помышляя о плате за переработки. Подобно тому, как «немножко туши да щепотка пудры превратят в красавицу любую лахудру», так и статистика способна выдавать многие весьма немаловажные факты совсем не за то, что они есть в реальности. Искусно преподнесенная статистика куда лучше, чем гитлеровский прием «большой лжи»: она вводит в заблуждение, но с вас взятки гладки, и никто не подкопается.

Эта книга – своего рода руководство для начинающих, в котором изложены азы применения статистики в целях обмана. У читателя могут возникнуть подозрения, что издание слишком уж смахивает на инструкцию для мошенников. Полагаю все же, что смогу оправдать ее в манере бывшего грабителя, опубликовавшего мемуары, в сущности представляющие собой учебный курс на тему о том, как подобрать отмычку к замку и научиться ступать бесшумно: жуликам и ворам все эти трюки и так давно известны, а порядочные люди должны узнать о них, чтобы уметь защитить свой дом от непрошеных гостей.





Глава 1

Выборка изначально необъективна

«Средний выпускник Йельского университета 1924 г. зарабатывает \$25 111 в год» – это было опубликовано однажды в журнале *Time* в ответ на какой-то материал, вышедший в нью-йоркской газете *Sun*.

Ну что ж, зарабатывает – вот и молодец!

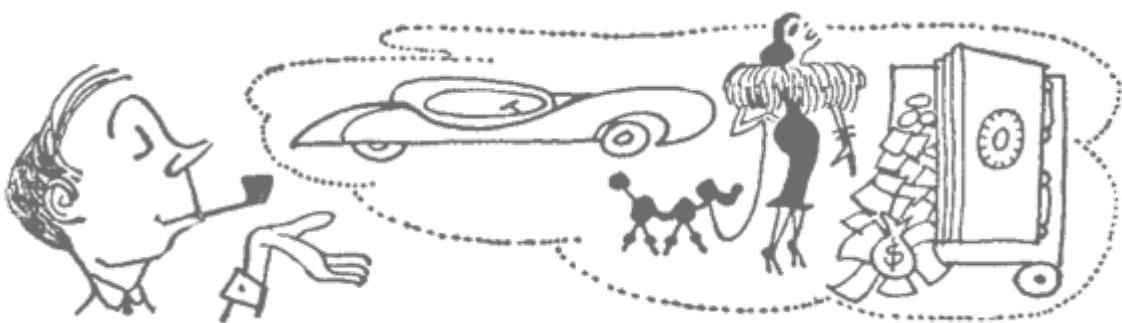
Но погодите-ка. А что, собственно, означает эта впечатляющая цифра? Будет ли она, как это кажется на первый взгляд, свидетельством того, что, если вы отправите своего отпрыска учиться в Йельский университет, вам уже не придется работать на старости лет (да и ему тоже)?

Уже при первом настороженном взгляде на эту цифру бросаются в глаза две особенности. Сама цифра на удивление точна. И потом, она неправдоподобно велика.

Маловероятно, чтобы средний доход любой сколько-нибудь обширной группы был бы известен с точностью до последнего доллара. Не так уж вероятно, чтобы вы с такой же точностью могли бы сказать, каким был ваш собственный доход в прошлом году, разве что весь он был получен исключительно за счет зарплаты. А годовой доход в размере \$25 000 редко когда складывается только

из заработной платы; люди с доходом такого уровня склонны делать инвестиции, причем их вложения предусмотрительно и с умом распределены между несколькими «корзинами».

Кроме того, этот восхитительный средний показатель рассчитан, несомненно, на основании тех сумм, которые, *как сообщили сами выпускники Йеля*, они зарабатывают. Даже если в 1924 г. у них там в Нью-Хейвене и имелась система доверия, то где гарантия, что и сегодня, четверть века спустя, она по-прежнему существует и все данные, представленные этими выпускниками, соответствуют действительности? Иные, отвечая на вопрос о личных доходах, склонны преувеличивать их – то ли из тщеславия, то ли потому, что настроены оптимистически. Другие же намеренно снижают свои доходы, особенно (и на то есть причины) в своих налоговых декларациях, а сделавши это, боятся, как бы данные, указанные ими где-нибудь еще, не опровергали тех, что значатся у них в декларации. Ведь неизвестно, какие сведения могут попасться на глаза чиновникам налогового управления. Вполне могло быть так, что эти две тенденции (заносить свои доходы или занижать их) нейтрализуют одна другую, но едва ли это вероятно. Одна из двух может быть намного сильнее другой, а вот которая – нам не известно.



Мы начали с того, что проанализировали цифру, которая, как подсказывает здравый смысл, вряд ли соответствует действительности. А теперь давайте разберемся с возможной причиной серьезной ошибки: почему эти самые \$25 111 указываются в качестве среднего дохода неких людей, чей фактический средний доход с таким же успехом может быть и вполовину меньше названной суммы?

Причина тому – процедура составления выборки, и именно она представляет собой сердцевину большей части статистических данных, которые встречаются в самых разнообразных сферах. Основа

этой процедуры довольно проста, а вот старания усовершенствовать ее на поверку только заводили на всяческие глухие окольные тропы, иногда весьма малопочтенные. Если у вас есть мешок бобов, часть из которых красного, а часть белого цвета, то единственный способ точно определить, сколько у вас белых бобов и сколько красных, – это пересчитать их. Однако есть и более простой способ приблизительно оценить количество красных и белых бобов – зачерпнуть пригоршню и сосчитать, сколько в ней будет белых бобов и сколько красных, исходя из предположения, что и в общем объеме белые и красные бобы содержатся в такой же пропорции. Если взятая вами для исследования пригоршня бобов, то есть выборка, достаточно велика и правильно отобрана, то для большинства надобностей она будет вполне репрезентативной. В противном случае выборка даст вам значительно менее точное представление о целом, чем сколько-нибудь обоснованные прикидки, а ее единственным достоинством будет разве что иллюзорное впечатление научной точности. Как ни печально, а выводы на основе такого рода выборок (необъективных или слишком малых, чтобы верно отразить свойства целого, или страдающих обоими этими изъянами) как раз и лежат в основе большинства из того, о чём нам доводится читать, или того, что мы, как нам представляется, знаем.

Упомянутые газетой сведения о доходах выпускников Йельского университета основаны на выборке. В этом можно не сомневаться, поскольку, как подсказывает здравый смысл, невозможно опросить всех выпускников 1924 г. Наверняка среди них довольно много людей, чье место проживания сейчас, спустя четверть века после выпуска, неизвестно.



А среди тех, чьи адреса известны, многие не стали бы заполнять анкету, тем более с вопросами такого щекотливого свойства.

Для некоторых анкет, рассылаемых по почте, 5–10 % ответивших уже считается достаточно высоким результатом. Данная анкета, надо полагать, добилась большего успеха, но ее результат явно далек от стопроцентного.

Итак, мы выяснили, что размер дохода вычислен на основе выборки, составленной из всех выпускников, адреса которых были известны и которые ответили на анкету. Репрезентативная ли это выборка? Иными словами, можно ли считать эту группу выпускников равной с точки зрения доходов группе выпускников, не представленных в выборке, то есть тех, чьи адреса не удалось раздобыть, и тех, кто не пожелал заполнить анкету?



Так кто же эти заблудшие овечки из числа выпускников Йельского университета, которые прошли по категории «адрес неизвестен»? Возможно ли, что они хорошо зарабатывают – ну, скажем, это дельцы с Уолл-стрит, директора компаний, руководители разных сортов? Нет, установить адреса людей состоятельных не составило бы труда. Большинство из самых преуспевающих выпускников того курса можно было бы найти в справочнике «Кто есть кто в Америке» или в других изданиях, даже если сами эти люди не пожелали поддерживать контакты с ассоциацией выпускников. Вполне правдоподобной представляется догадка, что найти не удалось имена тех выпускников, кто двадцать пять лет назад покинул стены Йельского университета с дипломом бакалавра гуманитарных наук, но так и не сумел заявить о себе чем-нибудь выдающимся. Это простые клерки, механики, бродяги, безработные алкоголики или перебивающиеся с хлеба на воду писатели и художники... в общем, те, кто только в шестером, если не больше, могли бы общими усилиями наскрести те самые \$25 111 в год. Люди подобного сорта не так уж часто изъявляют желание встретиться

со своими однокашниками, хотя бы по той причине, что не могут позволить себе подобную поездку.



Так кто они, те люди, что выкинули в ближайшую мусорную корзину анкету выпускника, присланную по почте? Знать наверняка мы не можем, но будет резонно предположить, что многие из них просто не могут похвастаться своими заработками. Они чем-то напоминают парня, который получил свою первую зарплату и увидел, что к чеку пришилена записка. В ней выражалась уверенность, что он считает размер своей заработной платы конфиденциальной информацией и не станет обсуждать ее в разговорах с коллегами. «Не беспокойтесь, – говорит этот парень своему боссу, – я стыжусь размера этой суммы не меньше вас».

Итак, нам уже ясно, что в выборку не были включены две группы выпускников, чьи доходы, скорее всего, уменьшили бы средний показатель годового заработка. И вот сумма \$25 111 получает наконец свое объяснение. Если эта цифра и вправду в чем-то верна, то она относится всего лишь к определенной группе йельских выпускников 1924 г. – к тем, чьи адреса известны и кто пожелал открыто заявить, сколько зарабатывает в год. Но даже этот вывод следует основывать на том допущении, что все эти достойные господа сказали правду.

А такое предположение не следует с ходу принимать за само собой разумеющееся. Как показывает опыт проведения выборочного исследования одной из категорий, а именно изучения рыночной

конъюнктуры, такое допущение едва ли вообще имеет право на существование. Как-то раз был проведен сплошной опрос населения, который имел целью изучить читательскую аудиторию популярных журналов. Основной вопрос, который задавали исследователи, один за другим обходя дома, был сформулирован так: «Какие журналы читают члены вашей семьи?» Когда результаты опроса свели в таблицы и проанализировали, выяснилось, что огромное количество американцев обожают *Harper's*⁴, а вот журнал *True Story*⁴ читают очень немногие. Между тем у издателей имелись в то время данные, которые очень четко показывали, что *True Story* выпускается миллионными тиражами, а *Harper's* – в сотни тысяч экземпляров. «Должно быть, мы опрашивали не тот контингент, какой следовало бы», – сказали себе организаторы опроса. Но нет: опросы проводились в самых разных районах по всей стране. В таком случае единственное разумное объяснение таково: значительная часть респондентов (так называют людей, когда они принимают участие в подобных опросах) попросту сказала неправду. В итоге практически единственное, что удалось выявить при помощи данного опроса, – порядочный сnobизм населения.

В итоге стало понятно, что, если требуется определить, что читает определенная группа людей, нет смысла спрашивать их об этом. Можно собрать намного больше сведений, если обходить дома этих людей под тем предлогом, что вы хотите купить старые журналы, и спрашивать, найдется ли у них что-нибудь в этом роде. А затем вам останется всего лишь пересчитать добытые экземпляры научного журнала *Yale Review* и душеписательного чтива *Love Romances*. Однако даже такой отчасти сомнительный способ, безусловно, не даст представления о том, какие издания читает ваш контингент, а только укажет, какие издания попадают в руки этим людям.

Подобным образом в следующий раз, когда вы прочитаете, что средний американец (в наши дни вы узнаете об этом субъекте много всякой всячины, причем по большей части все это слегка неправдоподобно) чистит зубы 1,02 раза в день – эту цифру я придумал прямо сейчас, но она ничем не уступает любой другой, – задайте себе вопрос: каким образом кому-то удалось собрать такие

⁴ *Harper's Magazine* – один из наиболее уважаемых ежемесячных журналов в США, посвященный литературе, политике, культуре, экономике и искусству. *True Story* – ежемесячный журнал, который печатает сентиментальные рассказы о любовных переживаниях и рассчитан на невзыскательный вкус. – Прим. пер.

сведения? Неужели женщина, начитавшаяся бесчисленных рекламных объявлений, где утверждается, что люди, не чистящие зубов, оскорбляют общественные устои, сознается совершенно незнакомому человеку, что делает это нерегулярно? Такого рода статистика может представлять интерес только для тех, кто хочет определить, что говорят люди о чистке зубов, но она мало что скажет о том, как часто щетка соприкасается с зубами респондентов.

Река, как нас учат, не может подняться выше своих истоков. Это верно, такое было бы возможно, только если где-нибудь неподалеку от нее находилась бы насосная станция. И в такой же степени верно утверждение, что результат выборочного исследования не может быть лучше выборки, на которой оно основано. К тому моменту, когда собранные данные, пройдя сквозь все процедуры статистических манипуляций, сведены к средним показателям, выраженным с точностью до десятых долей, они уже приобретают некий ореол убедительности, от которой не останется и следа, если внимательнее взглянуть на сам процесс выборочного исследования.



Правда ли, что ранняя диагностика рака сохраняет жизни пациентов? Очень может быть. Однако, основываясь на цифрах, часто используемых для подкрепления этого тезиса, можно сказать, что они вовсе не подтверждают это. Цифры эти, представленные

в Онкологическом реестре Коннектикута, относятся к 1935 г. и на первый взгляд указывают, что в период с того самого 1935 г. по 1941 г. существенно улучшился показатель выживаемости больных в течение пяти лет с момента диагностики у них ракового заболевания. На самом деле фиксация таких данных началась в 1941 г., а данные за предшествующие годы были получены путем ретроспективного исследования. Многие пациенты уехали из Коннектикута, и невозможно было установить, живы они или скончались. Возникшая из-за этого изначальная необъективность выборки «была достаточной, чтобы практически полностью объяснить заявленное улучшение показателя выживаемости», считает журналист Леонард Энджел, специализирующийся на медицинской тематике.

Чтобы данные выборочного исследования имели значительную ценность, они должны основываться на репрезентативной выборке, то есть на выборке, из которой устраниены все возможные источники предвзятости. Вот где наша цифра, обозначающая доходы выпускников Йеля, показывает свою несостоятельность. По этим же соображениям огромное количество сведений, встречающихся на страницах газет и журналов, лишены какого бы то ни было смысла.

Один психиатр заявил, что практически любой человек – неврастеник. Оставим в стороне тот факт, что подобное обращение с термином «неврастеник» лишает его всякого смысла, и посмотрим на выборку, послужившую основой для такого вывода. Иными словами, спросим себя: каких именно людей наблюдал данный психиатр? Оказывается, он пришел к такому поучительному выводу, изучая своих пациентов, а они более чем неподходящие кандидатуры на роль выборки из всего населения. Если человек был вполне нормален, у нашего психиатра не было никаких шансов увидеть его у себя на приеме.



Подвергайте такому осмыслению все прочитанное, и тогда вы сумеете оградить себя от великого множества сведений, не имеющих под собой реальной почвы.

Полезно помнить и о том, что скрытые источники необъективности способны с такой же легкостью подорвать надежность выборки, как и очевидные. Я имею в виду, что, даже если вам не удается обнаружить явный источник необъективности, позвольте себе некоторую долю сомнений и не доверяйте выводам безоговорочно, если имеется хоть какая-то вероятность, что они предвзяты. А это, поверьте, всегда возможно. В доказательство достаточно вспомнить президентские выборы в 1948-м и 1952 г., и всякие сомнения в этом отпадут⁵.

Для вящей убедительности давайте вернемся в 1936 г., к временам, когда влиятельнейший журнал *Literary Digest*, общепризнанный в то время лидер изучения предпочтений американских избирателей, потерпел приснопамятное фиаско. Те десять миллионов опрошенных телефонных абонентов и подписчиков *Literary Digest*, которые уверили редакцию злополучного журнала, что победителем в президентской гонки выйдет республиканец Альфред Лэндон с 370 голосами выборщиков против 161 голоса за Франклина Рузельята, были из того же списка

⁵ На президентских выборах 1948 г. Гарри Трумэн вопреки прогнозам авторитетных исследователей общественного мнения уверенно обошел кандидата-республиканца Томаса Дьюи, а в 1952 г. Дуайт Эйзенхаэр с большим перевесом одержал победу над демократом Эдлаем Стивенсоном и тем самым положил конец двадцатилетнему правлению демократической партии. – Прим. пер.

рассылки, каким журнал воспользовался в 1932 г., когда блестяще предсказал итоги президентских выборов.



Разве можно было заподозрить в предвзятости людей из списка, который в прошлом так хорошо себя зарекомендовал? Но, разумеется, предвзятость имела место, и список был нерепрезентативен, что и установили авторы диссертаций и прочие любители изысканий постфактум. Контингент населения, который в 1936 г. мог себе позволить иметь телефон и подписываться на *Literary Digest*, не был срезом всей совокупности избирателей. В экономическом плане это была особая категория населения, то есть нерепрезентативная выборка, поскольку она изобиловала теми, кто поддерживал Республиканскую партию. Данная выборка и отдала предпочтение Лэндону, тогда как избиратели в массе своей имели другое мнение на этот счет.

Базовая выборка относится к категории случайной (вероятностной) выборки. Она отбирается произвольным образом из генеральной совокупности, под которой статистики понимают весь обследуемый массив. Например, выбирается каждое десятое имя в картотеке индексных карточек. Или из шляпы, полной свернутых бумажек, наугад выбираются пятьдесят штук. Или интервьюируется каждый двадцатый человек на главной улице Сан-Франциско Маркет-стрит. (К вашему сведению, последняя из упомянутых не будет выборкой ни населения всего мира, ни Соединенных Штатов, ни самого Сан-Франциско, это всего лишь выборка из всей

массы людей, находящихся в это конкретное время на Маркет-стрит. Одна дама, проводившая опрос общественного мнения, рассказывала, что находила респондентов на железнодорожной станции, поскольку «на вокзалах можно обнаружить людей всех возможных категорий». Следовало бы указать ей, что матери малолетних ребятишек, например, могли быть недостаточно представлены в вокзальной толпе.)

Проверить, действительно ли выборка имеет случайный (произвольный) характер, можно с помощью такого вопроса: каждое ли имя или предмет из обследуемой совокупности имеют равный шанс попасть в выборку?

Безупречно случайная – единственный тип выборки, которую можно исследовать при помощи статистических методов с полной уверенностью в надежности результата. Но у нее имеется один недостаток. Получить такую выборку для множества надобностей настолько трудно и дорого, что чисто материальные соображения заставляют отказаться от этой идеи. Более экономной заменой, повсеместно используемой в таких сферах, как изучение общественного мнения и рыночной конъюнктуры, будет стратифицированная случайная выборка.

Чтобы получить стратифицированную выборку, вы должны разбить генеральную совокупность на несколько групп (страт) пропорционально известному показателю их распространенности в совокупности. Вот тут-то и начнутся трудности: сведения о том, каково соотношение групп в генеральной совокупности, могут быть некорректны. Вы инструктируете интервьюеров, которые будут проводить опрос, и наказываете им проследить, чтобы среди опрошенных было столько-то чернокожих, такой-то процент людей, относящихся к нескольким группам населения по размеру доходов, определенное число фермеров и т. п. Но вместе с тем в группе должно быть представлено равное количество людей в возрасте старше и моложе сорока лет.

Все это выглядит убедительно, но что происходит на деле? В том, что касается цвета кожи респондентов – белый это или чернокожий, интервьюеры в большинстве случаев не ошибутся. Но они допустят больше ошибок в оценке размера доходов опрашиваемых. А если говорить о фермерах, то как вы классифицируете человека, который часть времени трудится на ферме, но вдобавок имеет работу в городе? Даже такой вопрос, как возрастная категория респондента,

может создать некоторые трудности, но интервьюеры преодолевают их самым простым способом – выбирают респондентов, которые явно старше или значительно моложе сорока лет. Правда, в таких случаях выборка будет предвзятой ввиду фактического отсутствия в ней лиц в возрасте под сорок и тех, кому сорок с небольшим. Так что, как ни крути, хорошего решения все равно нет.

А кроме всего прочего, как на условиях стратификации получить вероятностную выборку? Самое очевидное решение – сначала переписать всех, кто входит в страту, а затем найти и опросить выбранных из этого списка случайным образом. Но это слишком уж дорогостоящая процедура. И тогда вы просто выходите на улицу – и сами искаjаете свою выборку, поскольку в ней не будут представлены люди, которые сидят в это время по домам. Если вы будете стучаться в двери днем – значит, не охватите большинство тех, кто работает. Решив проводить опросы по вечерам, вы упустите любителей кинематографа и завсегдатаевочных клубов.

В итоге проведение опроса сводится к стараниям побороть источники необъективности, и эту битву ведут все до единой почтенные организации, занимающиеся проведением опросов. Те, кто читает их доклады, должны помнить: эту битву никто и никогда не выигрывает. Всякий раз, когда вам где-нибудь встречается вывод, что «67 % американцев против» того-то или того-то, вам следует задаться вопросом: 67 % каких именно американцев?

То же самое относится и к «трактату о женской сексуальности» доктора Альфреда Кинси⁶.

⁶ Речь об одном из двух «Отчетов Кинси» на тему сексуального поведения – «Половое поведение самки человека» (*Sexual Behavior in the Human Female*) (1953); до этого, в 1948 г., Кинси опубликовал аналогичный труд о сексуальном поведении мужчин «Половое поведение самца человека» (*Sexual Behavior in the Human Male*). Альфред Чарлз Кинси – американский биолог и сексолог, профессор энтомологии и зоологии, основатель института по изучению секса, пола и воспроизведения, называемого ныне Институтом Кинси. Кинси считается отцом сексологии и систематического изучения сексуальности человека. Данные для своих исследований Кинси собирали посредством интервьюирования многих тысяч людей. – *Прим. пер.*



Проблема с этим трудом (как и со всеми прочими, в основу которых положены выборки) заключается в том, как ознакомиться с ним (или его кратким изложением) и при этом не намотать на ус слишком много суждений, которые не обязательно истинны. В труде доктора Кинси задействованы выборки как минимум на трех уровнях. Сделанные самим Кинси выборки из всего населения страны (один уровень) далеки от вероятностных и могут быть не особенно репрезентативны, но они колоссальны в сравнении с тем, что было сделано в этой области раньше. Приведенные Кинси цифры следует воспринимать как данные, на многое проливающие свет и значимые, даже если они и не вполне точны. Наверное, важнее иметь в виду, что любой вопросник или анкета представляют собой всего лишь выборку (еще один уровень) из всего множества возможных вопросов и что ответы женщин на эти вопросы – тоже не более чем выборка (вот вам третий уровень) из их личных взглядов и жизненного опыта по каждому заданному вопросу.

Сама личность интервьюера также способна довольно любопытным образом повлиять на ответы респондентов в ходе опроса. Несколько лет назад, еще во время войны⁷, Национальный центр изучения общественного мнения направил в один город на юге

⁷ Имеется в виду Вторая мировая война. – Прим. ред.

страны две группы интервьюеров с поручением задать три вопроса пятистам чернокожим горожанам. Одна группа состояла из белых, а вторая – из чернокожих интервьюеров.

Один из вопросов формулировался так: «Лучше или хуже обращались бы у нас с чернокожими, если бы японцы завоевали США?» По данным группы чернокожих интервьюеров, ответ «лучше» дали 9 % опрошенных ими респондентов. А белые интервьюеры зафиксировали всего 2 % таких ответов. И если чернокожие интервьюеры обнаружили всего 25 % респондентов, полагавших, что обращение с чернокожими ухудшится, то у белых интервьюеров доля респондентов, придерживающихся такой точки зрения, неожиданно оказалась равной 45 %.

В ответ на второй вопрос, аналогичный первому за тем исключением, что в формулировке значились не «японцы», а «нацисты», были получены похожие результаты.

Третий вопрос был призван прояснить взгляды респондентов, которые могли сформироваться на основе мнений, высказанных ими при ответах на первые два вопроса. «На чем, по вашему мнению, важнее сосредоточить усилия: на том, чтобы победить страны “оси”⁸, или на том, чтобы здесь, у себя дома, усилить действенность демократии?» Ответ «победить страны “оси”», по данным чернокожих интервьюеров, дали 39 % респондентов, а у белых интервьюеров этот показатель составил 62 %.

Перед нами пример предвзятости, обусловленной некими неизвестными факторами. Как представляется, самым сильнодействующим фактором будет тенденция, которую никогда не следует сбрасывать со счетов, когда знакомишься с результатами социологических опросов: желание респондента угодить интервьюеру. Стоит ли удивляться, что, отвечая на вопрос с подтекстом, намекающим на возможную нелояльность своей стране в военное время, чернокожие жители Юга скорее предпочли ответить белому интервьюеру так, чтобы их ответ выглядел достойно, чем сообщить ему, что они думают в действительности? Также нельзя исключать, что разные группы интервьюеров выбирали себе в респонденты разных людей.

Как бы там ни было, а результаты опроса, безусловно, до такой

⁸ Страны «оси» – агрессивный военный союз Германии, Италии, Японии и других государств, которому противостояла во время Второй мировой войны антигитлеровская коалиция. – Прим. ред.

степени предвзяты и необъективны, что это их практически обесценивает. Можете сами судить, сколь многие сделанные на основе опросов выводы и умозаключения до такой же степени предвзяты, равно как и ни к чему не годны – разве что у нас нет никакой возможности проверить их и изобличить их несостоятельность.



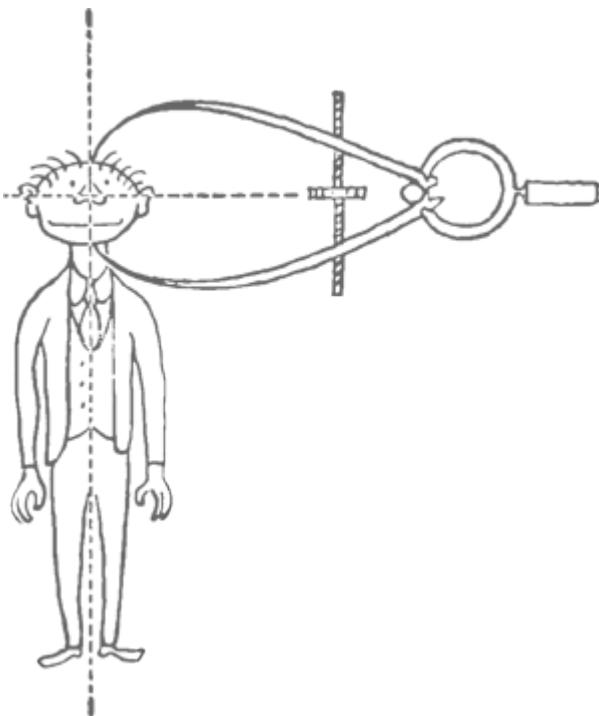
Теперь у нас достаточно оснований сделать следующий шаг и предположить, что все социологические опросы в целом грешат предвзятостью – того же сорта, что и допущенный *Literary Digest* просчет. Это крен в сторону людей более состоятельных, более образованных, более информированных и осторожных, с более пристойным внешним видом, общепринятым поведением и с более устоявшимися привычками, чем у того среднестатистического гражданина, представлять которого их выбрали.

Не так уж трудно понять, отчего такое происходит. Давайте представим, что вам поручили встать на углу улицы и задать ряд вопросов какому-то одному человеку. Вам попадаются на глаза двое прохожих, и оба вроде бы подходят под ту категорию, представителя которой вам предписано опросить, а именно – горожанина в возрасте старше сорока лет. Один прохожий одет чисто и аккуратно, другой же – во что-то непотребное, да к тому же угрюм и неприветлив на вид. Вы, чтобы выполнить свое задание, естественно, обратитесь к тому из двоих, кто выглядит приличнее,

и точно так же поступают ваши коллеги-интервьюеры по всей стране.

Среди прочих сильнейшая неприязнь к опросам общественного мнения наблюдается в кругах либералов и сторонников левых идей, где довольно-такиочно укоренилось мнение, что социологические опросы в большинстве своем – подтасовки и надувательство. Это мнение зиждется на том факте, что результаты социологических опросов очень часто не сообразуются с убеждениями и чаяниями тех, кто не разделяет консервативных взглядов. Если верить опросам, указывают они, то получается, что победят республиканцы, даже когда через считанные дни избиратели делают иной выбор.

На самом деле, как мы уже убедились, это вовсе не означает, что социологические опросы фальсифицируются – иными словами, что кто-то намеренно искажает их результаты, чтобы создать ложное впечатление. Свойственная выборке тенденция к систематическому смещению в сторону более состоятельных респондентов может исказить результаты любого опроса.



Глава 2 **Грамотно выбранное среднее**

Надеюсь, вы не страдаете снобизмом, а я определенно не имею никакого отношения к рынку недвижимости. Но давайте считать, что вы все-таки сноб, а я – агент по продаже недвижимости и что вы

как раз сейчас присматриваете земельный участок поблизости от Калифорнийской долины, где я обитаю.

Я уже разобрался, что вы за человек, и теперь, не жалея сил, убеждаю вас, что в интересующем вас месте проживают люди со средним доходом порядка \$15 000 в год. Может быть, именно этот довод и сыграл решающую роль в вашем решении поселиться в этих краях, но, как бы там ни было, вы совершаете покупку, а названная мною приятная сумма годового дохода местных жителей западает вам в душу. Более чем уверен (уж коли мы договорились, что сейчас вы немножечко сноб), что в разговорах с друзьями вы нет-нет, да и ввернете эту цифру, описывая место, где теперь живете.

Примерно через год мы с вами снова встречаемся. Как член некого комитета налогоплательщиков, я собираю подписи под петицией, требующей не повышать налоговую ставку, оценочную стоимость налогооблагаемого имущества или плату за проезд в автобусах. Я привожу такой довод: мы, местные жители, не можем позволить себе платить больше, ведь, в конце концов, средний доход жителей нашей округи составляет всего лишь \$3500 в год. Не исключено, что в этом вы соглашаетесь со мной и нашим комитетом – вы не только сноб, но еще и довольно прижимисты по натуре, – и все же вас не может не удивлять мое упоминание об этих жалких \$3500 среднего дохода. Вру я сейчас или соврал тогда, год назад?

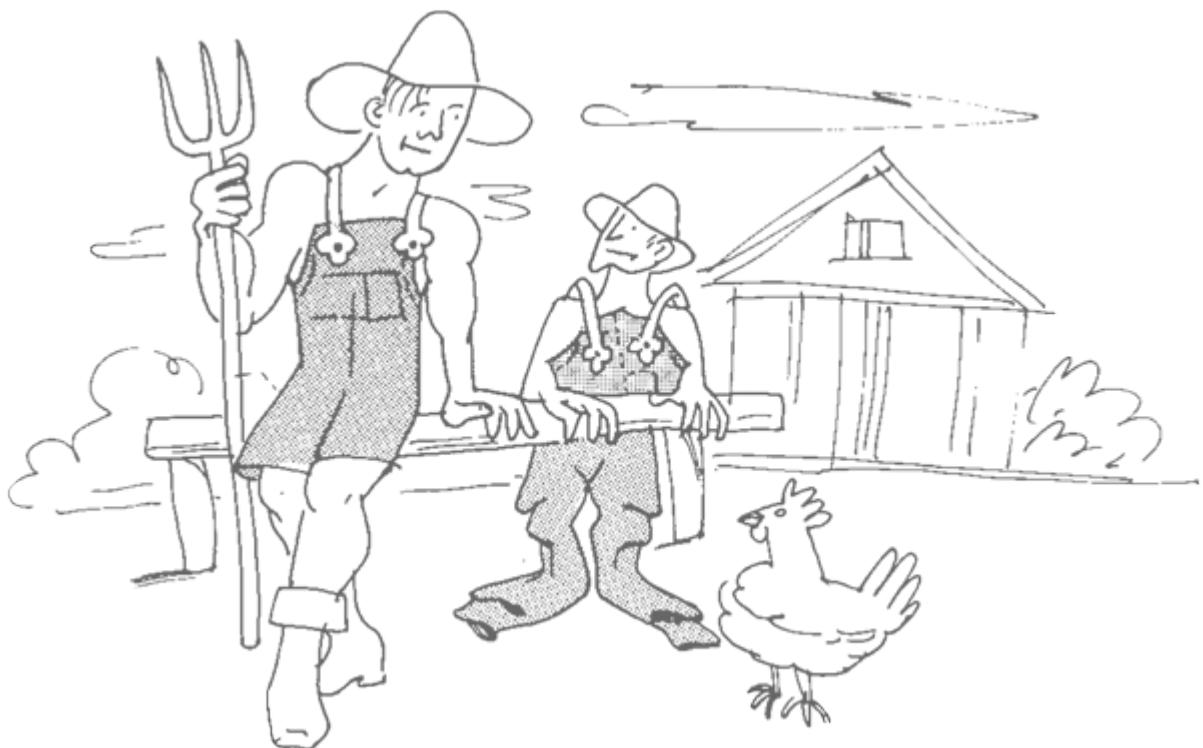
Вам не удастся дважды обвинить меня во лжи. В этом и состоит вся прелесть обмана при помощи статистики. Обе эти цифры представляют собой вполне правомочные среднестатистические показатели, обе вычислены с соблюдением всех правил. Обе отражают один и тот же массив данных, характеристики одной и той же группы людей, одни и те же показатели дохода. И тем не менее совершенно очевидно, что по крайней мере одна из цифр должна быть настолько недостоверной, что ничем не уступает самой отъявленной лжи.

Моя уловка состояла в том, что в первом и втором случае я воспользовался среднестатистическими показателями разного вида, ведь термин «средний» имеет очень расплывчатое толкование. Подобный трюк широко используют субъекты, желающие повлиять на общественное мнение или продать рекламное место – бывает, что по простоте душевной, но чаще вполне осознанно. Когда вам рассказывают, что некое число представляет собой среднюю

величину, это мало о чём вам скажет, пока вы не разберетесь, какой из трех основных видов среднего перед вами – среднее арифметическое, медиана или мода.

Когда мне требовался показатель побольше, я упомянул \$15 000. Это было простое среднее, то есть среднее арифметическое доходов всех семей, проживающих в той местности. Для его расчета требуется сложить доходы всех семей и разделить получившуюся сумму на число семей. Спустя год я воспользовался средним показателем меньшей величины – он представляет собой медиану и означает, что половина семей в рассматриваемой местности имеет годовой доход выше \$3500, а вторая половина – ниже \$3500. Я мог бы пустить в ход и моду, то есть чаще всего встречающееся значение в числовом ряду, составленном из доходов семей в интересующей нас местности. Если у большей части проживающих там семей годовой доход составляет \$5000, это значение и будет модой, или модальным доходом.

В данном случае (и, как правило, это справедливо в отношении показателей доходов) некое абстрактное «среднее» – без уточнения, какого оно вида, – в сущности, лишено всякого смысла. Добавляет путаницы и еще один фактор: во многих случаях все три средних показателя настолько близки по значению, что нет никакой необходимости делать различие между ними, если требуется в общих чертах охарактеризовать положение вещей.



Если вы где-то прочитали, что у людей, относящихся к какой-то группе, средний рост составляет всего лишь полтора метра, это дает вам вполне ясное представление об их росте. Вам нет необходимости уточнять, будет ли этот показатель средним арифметическим значением, медианой или модой – эти величины примерно одинаковы. (Разумеется, если вы связаны с пошивом спецодежды для этого контингента, вам понадобится больше данных, чем те, что способен дать любой среднестатистический показатель. Вам потребуется информация о диапазоне и отклонениях, и этим мы займемся в следующей главе.)

Разные виды среднего имеют близкие значения, когда дело касается данных наподобие тех, что относятся ко многим характеристикам человека. Они настолько любезны, что изволят тяготеть к тому, что называется нормальным распределением. Если начертить кривую нормального распределения, то по форме она будет напоминать колокол, а среднее арифметическое значение, медиана и мода попадут в одну и ту же точку.

Следовательно, один вид среднего ничуть не хуже другого, когда требуется охарактеризовать рост группы мужчин. Но все совсем не так, когда стоит задача описать размер их доходов. Если вам требуется свести в один список показатели годового дохода всех жителей исследуемого города, то может оказаться, что эти величины варьируются в пределах от довольно скромных до, вероятно, \$50 000 или около того. Кроме того, в списке могут фигурировать всего несколько показателей действительно очень высокого дохода. Порядка 95 % всех показателей будут ниже, чем \$10 000, и они займут место в левой части кривой. В итоге вместо симметричной, как колокол, кривой вы получите кривую, скошенную в одну сторону. С одной стороны – крутая горка, с другой – постепенный плавный спуск. Среднее арифметическое окажется на некотором расстоянии от медианы. Сами посудите, как это могло бы оказаться на справедливости любого сравнения между «средним» (средним арифметическим) показателем за один год и «средним» (медианой) за другой год.

В той местности, где я продал вам кое-какую недвижимость, два средних показателя особенно сильно расходятся по значению, поскольку распределение доходов имеет явное смещение. Так уж получилось, что большинство ваших соседей – мелкие фермеры,

наемные работники в близлежащем поселке или люди, отошедшие от дел и живущие на пенсию. Однако трое – миллионеры, они наведываются в здешние дома только по выходным, и именно за счет их миллионов суммарный годовой доход по вашей округе достигает такой значительной величины (и, соответственно, неимоверно увеличивает средний арифметический доход жителей). Из-за этих троих показатель среднего дохода приобретает огромный размер, какого и близко не имеет почти никто из остальных жителей местности. Это тот самый случай, когда шутка «Практически все имеют доход ниже среднего» становится реальностью.



И потому, если вам попадается заявление главы корпорации или владельца компании, где говорится, что у его сотрудников средняя зарплата достаточно высока, это может означать нечто конкретное, а может и не означать ничего. Если упомянутый средний показатель представляет собой медиану, он скажет вам нечто существенное, а именно, что половина сотрудников зарабатывает больше указанной суммы, а другая половина – меньше. Но если перед вами среднее арифметическое (а можете мне поверить, так оно и бывает, если вид среднего не уточняется), эта цифра не даст вам никакой полезной информации. Это всего лишь среднее арифметическое, которое складывается из одного показателя дохода в размере \$45 000 (владельца бизнеса) и зарплат его низкооплачиваемых сотрудников. Если где-то говорится, что «средняя заработная плата составляет \$5700», за этим могут скрываться заработки сотрудников размером \$2000 и прибыли владельца, представленные в виде дохода колоссальной величины.

Давайте чуть подольше задержим взгляд на последнем примере. В первую очередь мы видим, кто сколько получает. Босс, вероятно, предпочел бы охарактеризовать ситуацию с оплатой труда как «средний заработка в размере \$5700», то есть прибегнуть

к обманчивому среднему арифметическому значению. Однако мода сказала бы нам намного больше: самый распространенный размер заработка в этой компании составляет \$2000 в год. Однако медиана, как это часто бывает, лучше проясняет картину, чем любой отдельно взятый показатель: половина сотрудников зарабатывает больше \$3000, а другая половина – меньше.

Очень легко использовать этот прием, чтобы одним махом изобразить красивую картинку (и чем непригляднее ситуация на самом деле, тем краше она будет). Это хорошо иллюстрируют официальные отчеты некоторых компаний. Давайте попытаемся разобраться с одним из таких примеров.

Допустим, вы – один из троих партнеров, владеющих небольшим производством. Сейчас конец года, причем весьма для вас удачного. Вы выплатили в качестве зарплаты \$198 000 своим 90 сотрудникам, которые производят и отгружают заказчикам стулья или какую-то другую продукцию. Вы с партнерами выплатили себе по \$11 000 в качестве оклада. Но у вас остается еще и прибыль в размере \$45 000, которую вы должны как-то распределить. Как вы отразите эту ситуацию в отчете? Чтобы она была понятнее, вы облекаете ее в форму средних показателей. Ввиду того, что все ваши сотрудники выполняют примерно одинаковую работу за одинаковую плату, не составит никакой разницы, какого вида средний показатель вы используете – среднее арифметическое или медиану. Вот какой результат у вас получится:

Средняя сотрудников.....	зарплата
	\$2200
Средний владельцев.....	оклад
	и прибыль \$26 000

Сильно, правда? А давайте попробуем подойти к делу с другой стороны.



\$45 000



\$15 000



\$10 000



← СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ

\$5700



\$5000



\$3700



← МЕДИАНА

(этот парень будет посередине:
дюжине сотрудников платят больше,
чем ему, а другой дюжине — меньше.)

\$3000



\$2000

← МОДА

(величины зарплаты,
чище всего встречающиеся
в этой компании)

Берем из прибыли \$30 000 и распределяем между тремя партнерами в качестве премиальных. И на сей раз, рассчитывая средний размер заработка в компании, включите в расчет себя и своих партнеров. И позаботьтесь о том, чтобы воспользоваться средним арифметическим:

Средняя зарплата или оклад.....	\$2806,45
Средняя прибыль владельцев.....	\$5000,00

Ага! Вот так уже лучше. Эти данные можно и улучшить, но сейчас все выглядит куда пристойнее. Менее 6 % от фонда заработной платы и прибыли пошло владельцам, и вы можете пойти дальше и наглядно показать это, если захотите. В любом случае теперь у вас в распоряжении имеются цифры, которые вполне годятся, чтобы обнародовать их, поместить на доске объявлений или использовать при заключении трудовых договоров с сотрудниками.



Этот трюк выглядит довольно грубым, поскольку сам пример у нас упрощенный, но это ничто в сравнении с тем, что делают во имя отчетности. Если взять корпорацию со сложной структурой и несколькими уровнями иерархии сотрудников – от начинающей машинистки до президента, чьи премии исчисляются сотнями тысяч долларов, то точно таким же манером можно скрыть какие угодно реальные цифры.

Так что, когда вам встречается средний показатель зарплаты, первым делом задайтесь вопросом: это среднее чего? Кого оно включает? Был случай, когда Американская сталелитейная корпорация (United States Steel Corporation) заявила, что средняя недельная зарплата ее сотрудников возросла на 107 % в период с 1940 по 1948 г. Так оно и было – правда, впечатление от этого чудесного

прироста несколько меркнет, когда вы замечаете, что в показатель за 1940 г. включено значительно большее число частично занятых сотрудников. Если в каком-то году вы работали по полнедели, а в следующем перешли на полную занятость, ваш заработка удвоится, но это ничего не говорит о размере вашей заработной платы.

Вы могли прочитать в газете, что доход средней американской семьи в 1949 г. составил \$3100. Не стоит и пытаться делать из этого каких-либо серьезных выводов, пока вы не узнаете, что за «семьи» принимались в расчет для вычисления среднего и к какому виду относится это среднее. (И еще – кто это говорит, откуда он это взял и насколько точна сама цифра.)

Выясняется, что данную цифру опубликовало Бюро переписи населения США. Если у вас под рукой есть доклад бюро, вам не составит труда найти на его страницах всю сопутствующую информацию: опубликованная цифра представляет собой медиану; «семья» определяется как «двоє или больше человек, состоящих в родстве друг с другом и проживающих вместе». (Если в группу включить тех, кто живет один, медиана сместится к отметке \$2700, а это уже существенно отличается от первоначальных \$3100.) Если пойти назад, к исходным табличным данным, то можно заметить, что показатель дохода рассчитан на основе выборки такого размера, что с вероятностью девятнадцать из двадцати данный показатель (а он до округления равнялся \$3107) имеет погрешность в пределах $\pm \$59$.

Благодаря такой степени вероятности и такому уровню погрешности оценка получилась более или менее точной. Сотрудники бюро достаточно профессиональны, и их работа финансируется достаточно хорошо, что и позволило провести выборочное исследование с такой приличной степенью точности. И надо полагать, выполняя эту работу, сотрудники бюро не были заинтересованы подгонять оценку под какую-то конкретную цифру. Не все статистические показатели, которые вам встречаются, появляются на свет при таких счастливых обстоятельствах, да и далеко не все стандартные сопровождаются вообще какой-либо информацией, позволяющей судить, в какой степени точными или неточными они могут быть. Мы подробнее обсудим эту тему в следующей главе.



Между тем вам предоставляется случай испытать свою недоверчивость на некоторых утверждениях из колонки «От издателя» в журнале *Time*. Про новых подписчиков журнала там говорится, что «их медианный возраст составляет 34 года, а средний семейный доход равен \$7270 в год». Более раннее по времени исследование «давних подписчиков» установило, что их «медианный возраст достигает 41 года... Средний доход составляет \$9535...». Возникает резонный вопрос: почему в обоих случаях про возраст подписчиков уточняется, что это медианное значение, а вот о том, какого вида средним выражен средний доход, журнал предусмотрительно умалчивает? Может ли быть так, что среднее арифметическое для обозначения среднего дохода выбрано по той причине, что его величина больше других средних (и, судя по всему, преследуется цель приманить рекламодателей тем, что у журнала весьма состоятельная аудитория)?

Вы могли бы также поиграть в игру «Какого вида это среднее?» на примере предполагаемого преуспевания выпускников Йельского университета 1924 г. выпуска, о чём говорилось в начале первой главы.



Глава 3 Нюансы, о которых скромно умалчивают

«Потребители отмечают, что благодаря зубной пасте компании

Doakes у них образуется на 23 % меньше кариеса», – гласит набранный аршинными буквами заголовок. Вам интересно, как на 23 % уменьшить причину зубных болей, и вы читаете дальше. Выясняется, что эти результаты поступили (и это выглядит весьма обнадеживающе) из совершенно «независимой» лаборатории. Мало того, точность лабораторных расчетов удостоверена дипломированным бухгалтером⁹. Спрашивается, чего же еще?

И все же, если вы не совсем легковерны и не отъявленный оптимист, жизненный опыт говорит вам, что одна зубная паста редко бывает намного лучше другой. Но тогда на каком основании люди, пользующиеся пастой компании Doakes, отрапортовали о таком результате? Неужели они попросту позволили себе нагло солгать, да еще чтобы их вранье напечатали такими крупными буквами? Вовсе нет, да у них и не было такой надобности. На то придуманы куда более простые и действенные способы.

Главная уловка в данном случае заключается в некорректной выборке – статистически некорректной. Ну а для целей производителя зубной пасты она очень даже хороша. Группа испытуемых, как явствует из приведенного ниже текста, набранного мелким шрифтом, состояла всего из дюжины человек. (Однако следует отдать должное компании Doakes – помимо прочего, она честно оставила вам шанс проверить свою сообразительность. Иные из рекламодателей ни за что не раскроют такого рода сведения и оставят даже самых подкованных в статистике читателей теряться в догадках, какие именно махинации стоят за красивыми рекламными цифрами. Выборка в дюжину испытуемых не так уж и плоха в данных обстоятельствах. За несколько лет до этого на рынок был выставлен некий продукт под названием «зубной порошок доктора Корниша», причем под уверения, что он, дескать, продемонстрировал «значительный успех в устраниении… кариеса». Суть состояла в том, что в порошке содержится мочевина, а она, как предполагали лабораторные исследования, доказала свою эффективность в качестве противокариесного средства. Беспочвенность этого вывода обусловливалась тем, что эксперименты носили не более чем предварительный характер, и успех был зафиксирован всего в шести случаях.)

⁹ В США сертифицированный бухгалтер официально пользуется правом выступать в качестве независимого аудитора. – Прим. пер.

Однако давайте вернемся к вопросу, почему компании Doakes так легко удалось, не прибегая к вранью, добиться широкого освещения в прессе, да еще и подкрепить все это заключениями независимых экспертов. Предположим, некая немногочисленная группа потребителей в течение полугода ведет учет состояния своих зубов, а потом переключается на пасту от Doakes. Далее можно ожидать одного из трех вариантов: кариеса станет больше, кариеса станет ощутимо меньше или никаких изменений не последует. Если события пойдут по первому или последнему варианту, производитель пасты просто зафиксирует эти показатели (где-нибудь у себя, вдали от глаз общественности) и предпримет новые попытки. Рано или поздно в дело вмешается случай, и у испытуемых зафиксируют-таки значительное улучшение, достойное газетных заголовков, а то и целой рекламной кампании. И случится это независимо от того, пользуются ли испытуемые пастой Doakes, питьевой содой или своим привычным средством по уходу за зубами.

Малочисленную группу испытуемых важно воздействовать вот почему: при многочисленной группе любой случайный сдвиг в лучшую сторону будет, скорее всего, довольно скромным и потому не заслужит упоминания в прессе. Очень сомнительно, что заявления о двухпроцентном улучшении подстегнут продажи зубной пасты.

Каким образом результат, который ни о чем не говорит, можно получить по чистому везению (притом что число исследуемых случаев достаточно мало), вы можете проверить на себе, не истратив ни цента. Возьмите монетку и подбросьте ее несколько раз. Как часто она будет падать решкой вверх? В половине случаев, конечно. Это всякий скажет.

А давайте-ка проверим и поглядим, так ли это... Лично я только что десять раз подбросил монетку, и в восьми случаях она упала решкой вверх. Это доказывает, что в 80 % случаях при подбрасывании монетки она падает решкой вверх. Ну да, согласно методам, которыми получена статистика по зубной пасте, так оно и есть.

ПО ФАКТУ (одна попытка)



А теперь проделайте это сами. У вас может получиться пятьдесят на пятьдесят, но по всей вероятности это будет не так. Более вероятно, что ваш результат, как и мой, окажется довольно далек от половины наполовину. Но если у вас хватит терпения на тысячу попыток, то вы почти наверняка (хотя обещать вам этого не стану) получите результат очень близкий к тому, чтобы монетка падала решкой вверх в половине случаев. Такой результат и представляет собой реальную вероятность. Только при достаточно большом количестве попыток закон средних чисел позволяет получить значимую характеристику или прогноз.



А достаточно большое количество – это сколько? Довольно каверзный вопрос! Помимо прочего, все зависит от того, насколько обширна и вариативна группа населения, которую вы изучаете методом выборки. И случается, что количество человек, попавших в выборку, совсем не так обоснованно, как представляется.

Примечательный пример этого появился в связи с испытанием противополиомиелитной вакцины, которое проводилось несколько

лет назад. Это выглядело впечатляющим масштабным экспериментом, как свойственно подобным медицинским испытаниям: в некой местности были вакцинированы 450 детей, а 680 детей остались непривитыми (в качестве контрольной группы). Вскоре после этого в той местности случилась эпидемия полиомиелита. Ни у одного из вакцинированных детей не было выявлено полиомиелита.

Как не было его выявлено и у детей из контрольной группы. Что проглядели экспериментаторы (или просто не поняли), когда планировали свое испытание, так это редкость паралитического полиомиелита. В обычном случае в группе такой численности можно ожидать всего двух случаев заражения, так что испытание с самого начала было совершенно бессмысленным. Потребовалась бы группа численностью раз в пятнадцать, а то и в двадцать пять больше, чтобы получить сколько-нибудь значимый результат.

Подобным же образом совершались многие из выдающихся (пускай и недолго продержавшихся) открытий в медицинской области. Один врач так высказался по этому поводу: «Поторопитесь воспользоваться новым препаратом, пока он не перестал действовать».



Впрочем, не во всех случаях стоит возлагать вину на одних

только медиков. Настоятельный запрос общества и чрезмерная поспешность журналистов нередко приводят к тому, что на рынке появляются лекарства, не прошедшие всех положенных испытаний, особенно когда спрос очень велик, а исходные статистические данные неопределенны и не дают ясной картины. Так случилось с прививкой от простуды, чрезвычайно популярной несколько лет назад, а не так давно – с антигистаминными средствами. Популярность этих бесполезных «лекарств» во многом проистекала из неясности происхождения самого заболевания и изъяна логики. Дайте время, и простуда вылечивается сама собой.

Так как же не дать обмануть себя результатами каких-нибудь исследований, неокончательных и неубедительных? Должен ли каждый из нас стать сам себе статистиком и лично изучать исходные данные любого исследования? В принципе, все не так уж плохо, тем более что есть такая штука, как критерий значимости, суть которого несложно понять. Это просто способ показать, насколько вероятно, что полученная в ходе испытаний цифра отражает реальный результат, а не что-то случайное. Это тот самый нюанс, о котором обычно умалчивают – на том основании, что вы, несведущий читатель, все равно не поймете, о чем идет речь. Или наоборот, непременно поймете, если кто-то кровно заинтересован в определенном результате.

Если ваш источник сведений сообщает и о степени их значимости, у вас будет более ясное представление о том, насколько эта информация заслуживает доверия. Степень значимости проще всего выразить в виде вероятности, как это делает Бюро переписи населения, когда прямо говорит, что в девятнадцати случаях из двадцати их цифры имеют указанную степень точности. Для большинства случаев сойдет все, что не хуже этого пятипроцентного уровня. Для некоторых целей требуемый уровень точности составляет 1 %, а это означает, что в девяноста девяти случаях из ста информация верна. Подобное иногда характеризуют как «практически точные» данные.

Есть еще одного сорта нюанс, который предпочитают не указывать, но его отсутствие способно не меньше дискредитировать заявленные данные. Речь идет о размахе исследуемого признака или диапазоне отклонения от указанного среднего. Часто бывает, что среднее – будь то среднее арифметическое значение или медиана, с уточнением

или без уточнения, какого вида это среднее, – представляет собой такое чрезмерное упрощение, что оно даже хуже, чем бесполезно. Ничего не знать о предмете зачастую гораздо лучше, чем знать то, что не соответствует действительности, а малые познания – штука подчас весьма опасная.



Слишком многое в американском жилом строительстве, например, планировалось таким образом, чтобы соответствовать размеру среднестатистической семьи из 3,6 человека. В переводе на язык реальности это означает семью из трех или четырех человек, что, в свою очередь, предполагает необходимость в доме двух спален. А семья такого размера, какой бы «среднестатистической» она ни считалась, в Америке находится в меньшинстве. «Мы строим среднестатистические дома для среднестатистических семей», – говорят застройщики – и пренебрегают большинством семей большего или меньшего размера. Некоторые районы, как следствие, застроены избыточным количеством домов с двумя спальнями, и при этом ощущается нехватка домов для семей как меньшего, так и большего размера. Перед нами тот самый случай статистики, неполнота которой, вводя в заблуждение, оборачивается дорогостоящими последствиями. По данному поводу Американская ассоциация работников здравоохранения высказываеться так: «Если мы отвлечемся от среднего арифметического значения и изучим фактический диапазон данного показателя, который этим средним искается, то обнаружим, что семьи из трех и четырех человек составляют лишь 45 % от общего. 35 % – это семьи из одного или двух человек, а в 20 % семей больше четырех человек».

По каким-то неведомым причинам здравый смысл капитулировал перед точной и авторитетной цифрой в 3,6 человека. Этот показатель чудесным образом перевесил тот факт, что всем известен из личных наблюдений: многие семьи маленькие, а вот больших не так уж много.



Сходным образом мелкие опущенные детали в труде под названием «Нормы развития Гезелла» ввергли в панику папочек и мамочек. Дай только родителю прочитать раздел воскресной газеты, где говорится, что в возрасте стольких-то месяцев ребенку уже полагается сидеть, и он сейчас же примерит это к собственному малышу. И боже упаси, если чадо достигло указанного возраста, но все еще не умеет сидеть – родитель должен заподозрить, что его отпрыск «запаздывает в развитии», «не совсем нормален» или еще что-нибудь столь же жуткое и беспочвенное. А поскольку примерно половина детей к указанному возрасту все еще не научилась сидеть, это сделало несчастными многих и многих родителей. Разумеется, говоря языком математики, их страдания уравновешиваются радостью другой половины родителей, обнаруживших, что у них вполне «развитые» дети. Зато большой вред могут причинить старания несчастных родителей подстегнуть развитие своего ребенка, чтобы он соответствовал норме и больше не считался недоразвитым.

Все это нисколько не бросило тень на доктора Арнольда Гезелла и его методы. Причина погрешности такова: на пути от исследователя к читателю (через руки охочего до сенсаций или невежественного автора) исходные данные отчасти отфильтровываются, а читатель не в состоянии заметить отсутствие кое-каких цифр, испарившихся по ходу дела. Этого недоразумения во многом удалось бы избежать, если бы наряду с показателем «нормы» или среднего значения

был бы указан диапазон этой самой нормы. Тогда родители увидели бы, что их дети попадают в пределы нормы и прекратили бы беспокоиться по поводу мелких и ничего не значащих отклонений. Едва ли найдется человек, которого можно было бы считать строго нормальным в каком-то отношении, точно так же, как и сто раз подброшенная монета редко когда упадет решкой и орлом поровну.

Положение еще больше ухудшается, когда «нормальное» путают с «желательным». Доктор Гезелл просто сообщил результаты своих наблюдений. Родители сами, читая книги и статьи, сделали вывод, что ребенок, начавший ходить на день или месяц позже положенного, не иначе как отстает в развитии.

Львиная доля глупой критики хорошо известного (но едва ли внимательно прочитанного) доклада доктора Альфреда Кинси проистекает от того, что нормальное многие восприняли как эквивалент приличного, правильного и желательного. Доктора Кинси обвиняли в том, что он развращает молодежь, сея в неокрепших умах определенного рода идеи, а еще чаще в том, что он объявил нормальными всякого рода распространенные, хотя и не одобряемые обществом формы полового поведения. Но Кинси всего лишь констатировал, что, по его наблюдениям, такие формы поведения обычны, а это и означает «нормальны», однако он не спешил их одобрять. Определять, предосудительны они или нет, никак не входило в сферу деятельности доктора Кинси в том виде, как он определял ее для себя. Вот он и напоролся на нечто, на чем погорело множество исследователей до него: опасно затрагивать крайне щекотливые в глазах общества темы, не объявляя, одобряете вы их или осуждаете.



Мелкие нюансы, о которых нам не сообщают, больше всего способствуют обману тем, что их отсутствие зачастую остается незамеченным. В этом, несомненно, и заключается секрет их успеха. Критики журналистской профессии в том виде, как она практикуется сегодня, сокрушаются, что старый добрый принцип «журналиста ноги кормят» у нынешних не в чести, и мечут стрелы в адрес «واشنطنских кабинетных репортеров», которые живут тем, что, не задаваясь лишними вопросами, покорно переписывают выпускаемые правительством пресс-релизы. В качестве примера журналистики ленивой и бездеятельной предлагаю на ваш суд следующий пассаж из разряда «новые промышленные разработки», обнаруженный в перечне достижений на страницах новостного журнала *Fortnight* : «...в компании Westinghouse разработали ванну для низкотемпературной закалки, которая позволяет втрое увеличить прочность стали».

Да уж, выглядит прямо как открытие... пока вы не попытаетесь разобраться, что же имеется в виду. И тогда выяснится, что смысл сказанного ускользает от понимания, и уловить его так же трудно, как поймать шарик ртути. Новая ванна, она что, любой тип стали, какой ни возьми, сделает в три раза прочнее, чем он был до обработки? Или она производит сталь, втрое более прочную, чем сталь любого уже существующего типа? Или что она там делает? Такое впечатление, что корреспондент просто переписал некоторый набор слов, не дав себе труда поинтересоваться, что они означают. Ну а от вас ожидают, что вы проглотите новость как есть, без лишних

вопросов, и удовольствуетесь иллюзией, что узнали нечто новое. Уж больно это напоминает старое определение лекционного способа обучения: процесс, посредством которого содержание учебника в руках учителя плавно перекочевывает в тетрадь учащегося, минуя сознание обоих.

Да вот только что, просматривая кое-какие материалы о докторе Кинси в старом номере *Time*, я наткнулся на еще один образчик утверждений того сорта, что при более пристальном рассмотрении рассыпаются в прах. Это утверждение нашлось в рекламе группы электроэнергетических компаний, напечатанной в 1948 г. «Сегодня электричество доступно более чем трем четвертям американских ферм...» Звучит довольно неплохо. Эти компании и впрямь знают свое дело. Разумеется, если бы вам захотелось побрюзжать, вы перефразировали бы данное заявление так: «Сегодня почти четверть ферм США не имеют доступа к электроэнергии». Но истинная уловка заключается в формулировке «доступно». Прибегая к этому словечку, компании могут рассказывать нам обо всем, что их душе угодно. И так понятно, что утверждение про три четверти ферм вовсе не означает, что все эти фермеры действительно пользуются электроэнергией, иначе авторы рекламного объявления уж точно не преминули бы заявить об этом напрямую. Они же написали просто «доступно» – и это, насколько я могу судить, может означать, что линии электропередачи проходят в стороне от ферм или на расстоянии в десяток, а то и в сотню миль от них.

Доступность книги «Как лгать при помощи статистики» в общемировом масштабе



■ Районы, расположенные в радиусе 25 миль от железнодорожных путей, автодорог и пригодных для навигации водных путей (маршруты для движения на собачьих упряжках не учитываются)

Позвольте, я приведу название статьи из номера *Collier's* 10 за 1952 г.: «Теперь вы можете узнать, НАСКОЛЬКО ЕЩЕ ВЫРАСТЕТ ВАШ РЕБЕНОК». Как откроешь, в глаза сразу бросается два графика, отдельно для мальчиков и девочек, где показано, на какой процент от окончательного взрослого роста увеличивается за год рост ребенка. «Чтобы определить, какого роста будет ваш ребенок во взрослом возрасте, — гласит подпись, — найдите на диаграмме его нынешний рост и посмотрите, какой показатель соответствует этому значению».

Забавный момент: в самой статье — если озабочиться ее прочтением — разъясняется, в чем состоит недостаток графика (что напрочь разрушает к ней доверие). Оказывается, не все дети растут одинаковыми темпами. Одни растут сначала еле-еле, а потом делают рывок; другие какое-то время стремительно прибавляют в росте, а после замедляются; но все остальные действительно растут более или менее равномерно. А данные графики, как можно догадаться, основаны на средних показателях, рассчитанных на основе большого

¹⁰ *Collier's Weekly* (Кольеровский еженедельник) — иллюстрированный еженедельный журнал, основан Питером Фенелоном Кольером, одним из пионеров жанра журналистского расследования. Вскоре журнал приобрел авторитет и репутацию поборника социальных реформ; издавался с 1888 по 1957 г.; выпуск возобновлен в феврале 2012 г. — Прим. пер.

числа измерений. В общем и целом, то бишь в среднем, рост сотни случайным образом выбранных подростков, несомненно, с достаточной точностью укладывается в график. Но дело-то в том, что каждого родителя интересует всего один конкретный показатель, а для этого график подобного рода в сущности бесполезен. Если вам так уж хочется узнать, насколько вымажает ваше чадо, то, по всей видимости, более или менее точно об этом скажет рост его родителей, бабушек и дедушек. Это метод не такой научный и скрупулезный, как график, но по крайней мере точный.

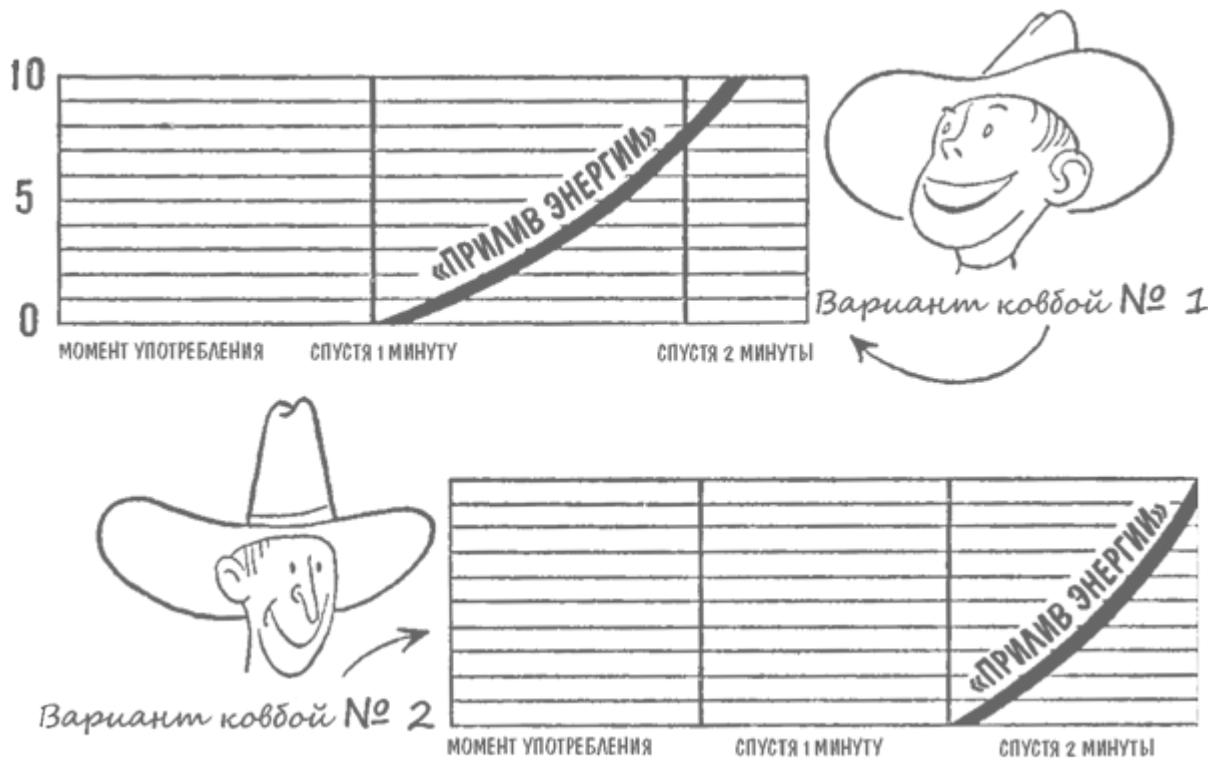


Меня, кстати, развеселило, когда я проверил себя по этому графику, взяв за основу свой рост в 14 лет, когда в средней школе приступил к курсу начальной военной подготовки. Так вот, тогда со своим ростом я был в числе последних. Судя по графику, мне суждено было вырасти всего до 173 сантиметров. На самом же деле во мне все 180. Согласитесь, что ошибка в 7 сантиметров, когда речь идет о росте человека, низводит оценку до разряда очень неточных.

Передо мной две упаковки из-под виноградно-ореховых хлопьев. Коробки оформлены немного по-разному – на них разные хвалебные рекомендации, призванные убедить покупателя в ценности продукта. На одной коробке изображен Двухревольверный Пит, на второй, где нарисован ковбой Хоппи¹¹, говорится: «Если хочешь стать

¹¹ Хопалонг Кессиди (Хоппи) – главный персонаж серии рассказов, а позже и романов о Диком Западе Кларенса Малфорда, по мотивам которых к началу 1950-х гг. было снято множество кинофильмов. Хоппи, в первых рассказах представленный как персонаж опасный, задира и грубиян, позже был несколько

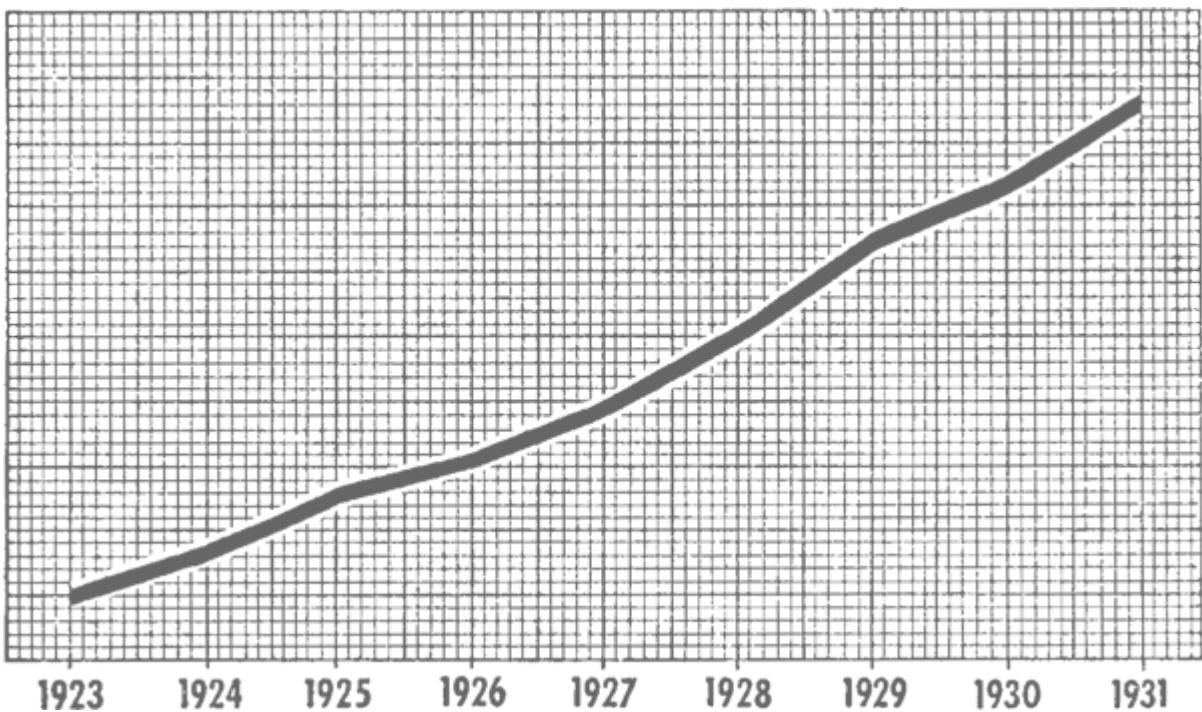
как Хоппи... ешь то же, что и Хоппи!» На обеих упаковках помещены графики, убеждающие («Ученые доказали, что так оно и есть!»), что от этих хлопьев «ты ощущаешь прилив энергии уже через две минуты!» На одной упаковке график, малозаметный на фоне целого леса восклицательных знаков, снабжен делениями на вертикальной оси; на второй упаковке никаких делений на графике нет. Впрочем, это ничего не меняет, поскольку отсутствует какой-либо намек на то, что могли бы означать эти цифры. Зато на обоих графиках красная линия (изображающая «прилив энергии») круто взбирается вверх, но в одном случае она берет начало у отметки «первая минута с момента употребления хлопьев», а во втором — у отметки «вторая минута», то есть двумя минутами позже, чем на первом графике. При этом один график вдвое круче уходит вверх, чем второй, и это подсказывает, что даже художник-оформитель был далек от мысли, что эти графики могут хоть что-то означать.



Подобные нелепости, разумеется, можно встретить только на материалах, предназначенных для глаз подростков или их измученных систематическим недосыпанием родителей. Да никто

облагорожен автором, чтобы больше соответствовать экранному образу. Ковбой Хоппи давно стал частью американского фольклора. — *Прим. пер.*

и не решился бы оскорбить разум крупного бизнесмена такими дрянными статистическими поделками... или все же решился бы? Тут я, с вашего позволения, хотел бы поговорить о графике, который использовали, чтобы прорекламировать рекламное агентство (надеюсь, этот каламбур не сбивает вас с толку) в одном из разделов журнала *Fortune*, адресованном главным образом специалистам. Кривая на данном графике показывает, что год за годом дела агентства идут все лучше и лучше. Никаких цифр, как вы заметили, на графике нет. Он мог бы с равным успехом отражать как колossalный рост доходов, которые за год удваиваются или на миллионы долларов увеличиваются, так и более чем скромный прогресс, сравнимый разве что с черепашьим шагом, когда доходы увеличиваются всего-то на доллар-другой в год. Однако график этот производил изрядное впечатление.

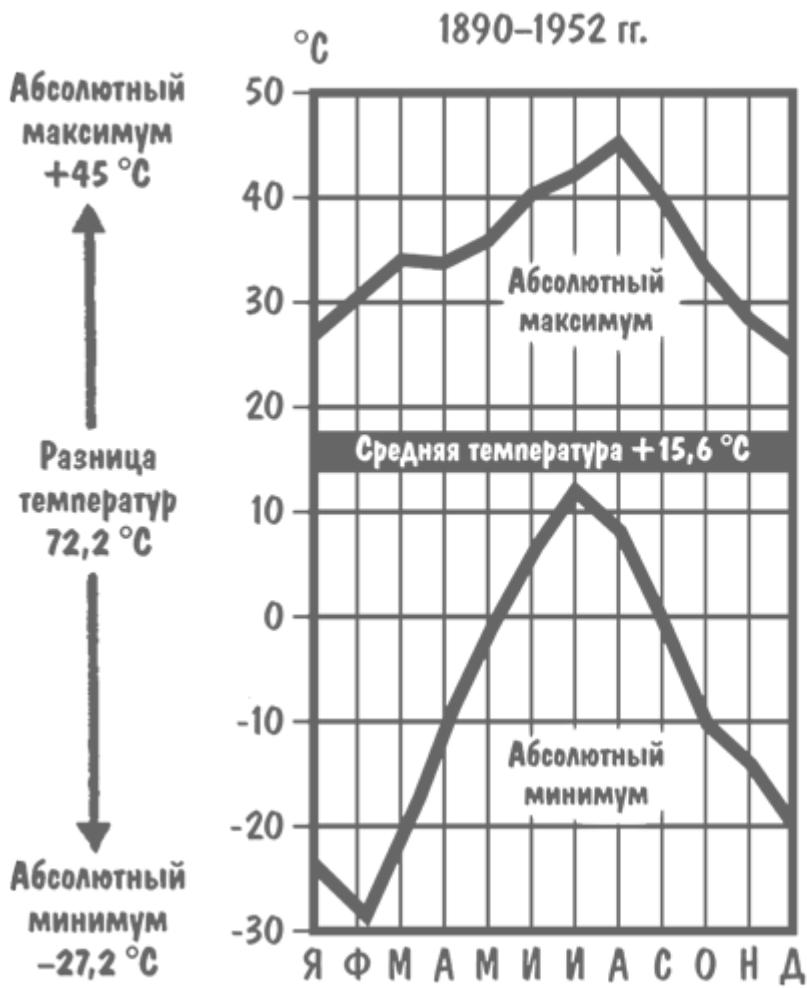


Не доверяйте особо среднестатистическим показателям, графикам и тенденциям, когда вам предъявляют их без тех важных цифр, что могли бы прояснить смысл. Иначе вы будете слепы, как тот чудак, что присматривает себе место для вылазки на природу, руководствуясь лишь сводкой средней температуры. Допустим, вас устраивает величина среднегодовой температуры 16,1 °C. В таком случае вы можете выбирать между такими калифорнийскими местечками, как расположенная вдали от моря пустыня и остров Сан-Николас, что возле южного побережья. Но вы рискуете

заледенеть от холода или свариться от жары, если не учитываете диапазон температурных колебаний в этих местах. На острове Сан-Николас температура колеблется в пределах от +8,3 до +30,5 °C, тогда как в пустыне амплитуда колебаний составляет от -9,4 до +40 °C.

Все права претендовать на близкий к вышеназванному среднегодовому показателю температуры есть и у Оклахома-Сити – за последние шестьдесят лет он держится в этом городе на уровне +15,6 °C. Но, как вы видите на приведенной ниже диаграмме, за этой комфортной прохладой скрывается амплитуда колебаний в 72,2 °C.

Рекордные температурные показатели в Оклахома-Сити



В среднем кого дка
ну просто
восхитительная





Глава 4

Много шума практически из ничего

Если вы не против, начнем с того, что одарим вас двумя ребятишками – сыном и дочкой. Питер и Линда (если уж на то пошло, мы могли бы подобрать им имена и помоднее) недавно прошли тест на уровень умственного развития, как это случается в школе с множеством детей. В наши дни из всякого рода тестов на умственные способности сотворили прямо-таки фетиш и разводят вокруг них шаманские пляски. И не исключено, что вам придется выдержать небольшую словесную баталию, чтобы добыть результаты теста. Это сведения весьма эзотерического свойства, недоступные для понимания непосвященных, и потому принято считать, что они будут сохраннее, оставаясь в руках одних только психологов и педагогов. Вероятно, это не лишено смысла. Как бы там ни было, но вы выяснили, что у Питера коэффициент умственного развития (IQ) составил 98, а у Линды – 101. А вам, разумеется, известно, что в тесте на IQ коэффициент 100 принят за средний, то есть нормальный уровень.

Ага. Линда у нас одареннее Питера. Это означает, что ее умственное развитие выше среднего. А у Питера – ниже среднего, но не будем придавать этому особого значения.

Умозаключения подобного толка – сущая чепуха.

Просто чтобы расставить все по своим местам, отметим прежде всего, что не играет особой роли, что там измеряет тест на умственное развитие, поскольку это в любом случае не совсем то, что мы обычно подразумеваем под интеллектом. Подобного рода тесты упускают из виду такие важные свойства, как инициативность и творческое воображение. Они не принимают в расчет и сложившиеся на социальной почве суждения, музыкальные и художественные способности испытуемых, не говоря уже о таких личностных качествах, как прилежание и уравновешенность. Ну а главное, в школах чаще всего проводятся быстрые и дешевые групповые тесты, результаты которых во многом зависят от умения читать: не важно, насколько одарен ученик, – если он плохо читает, шанса блеснуть у него нет.



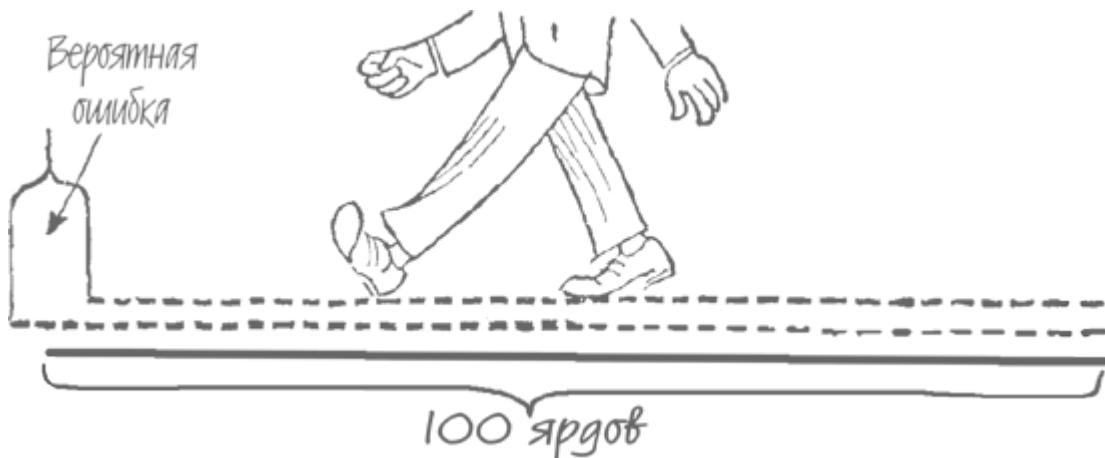
Давайте считать, что мы учитываем все вышесказанное и соглашаемся воспринимать IQ всего лишь как меру несколько туманно определенного умения оперировать предложенными отвлечеными понятиями. Интеллектуальные способности Питера и Линды проверялись с помощью усовершенствованного теста Стэнфорда – Бине, а его принято считать самым лучшим в этой области, в том числе и потому, что он проводится индивидуально и не требует каких-то особых способностей к чтению.

Идем дальше. Как заявляется, тест на IQ производит выборочную проверку интеллекта. Как и любой другой показатель, полученный методом выборки, IQ содержит статистическую погрешность, отражающую степень точности или достоверности данного показателя.

Задавать предусмотренные тестом вопросы – это все равно как если вы, чтобы оценить качество кукурузных початков, бродите по полю и тут и там очищаете от листьев початки у произвольно выбранных растений. К тому моменту, когда вы раскурочите и обследуете, скажем, с сотню початков, вы составите довольно точное представление о положении дел на поле в целом. Полученные данные будут достаточно точны, чтобы сопоставить состояние этого поля с каким-нибудь другим – при условии, что эти два поля не слишком похожи. Потому что в этом случае вам пришлось бы обследовать большее число початков, причем оценивать каждый, исходя из некоего точно определенного критерия качества.

То, насколько точной может считаться ваша выборка, призванная дать представление обо всем поле, есть мера, которую можно выразить количественно: это вероятная ошибка и стандартная ошибка.

Предположим, вам требуется определить размеры немалого числа полей, причем измерять предстоит шагами, следя вдоль изгородей. Первое, что вам следует сделать, – это проверить, насколько точна ваша система измерения, и для этого нужно несколько раз промерить шагами расстояние, длиною, как вы считаете, 100 ярдов. Вполне возможно, что в среднем погрешность подобного измерения составит 3 ярда. Иными словами, половина ваших промеров даст результат, отличающийся от 100 ярдов на 3 ярда в ту или другую сторону, а в другой половине случаев вы ошибетесь больше, чем на 3 ярда.



Тогда вероятная ошибка ваших измерений составит 3 ярда на 100 ярдов, или 3 %. Значит, длину каждой изгороди, которую вы измерили шагами и определили равной 100 ярдам, можно будет

занести в реестр как 100 ± 3 ярда.

(Большинство статистиков на сегодняшний день отдают предпочтение другому, но сопоставимому параметру, называемому стандартной ошибкой. За основу берется порядка двух третей случаев вместо ровно половины, и потом, стандартной ошибкой пользоваться значительно удобнее с точки зрения математики. Но для наших целей мы можем так и продолжить оперировать вероятной ошибкой, которая до сих пор в ходу применительно к тесту Стэнфорда – Бине.)

Как и в случае с нашим гипотетическим измерением полей, вероятная ошибка теста на IQ Стэнфорда – Бине определяется как 3 %. Это ни в коей мере не говорит о том, насколько в основе своей хорош данный тест, а просто указывает, с какой надежностью он измеряет то, что призван измерять. Итак, определенный у Питера IQ можно было бы полнее выразить в виде 98 ± 3 , а коэффициент IQ Линды – в виде 101 ± 3 .

Это означает, что у IQ Питера равные шансы оказаться где-то в диапазоне от 95 до 101: коэффициент его интеллекта может быть с равной вероятностью как выше этих 98, так и ниже. Аналогично и у Линды вероятность попасть в интервал от 98 до 104 ничуть не лучше, чем пятьдесят на пятьдесят. Из этого вы можете легко сделать вывод, что в одном случае из четырех IQ Питера может действительно быть выше 101, и существует такая же вероятность, что IQ Линды ниже 98. Но тогда Питер не хуже, а лучше, и притом с преимуществом порядка трех пунктов, а то и больше.

Все сказанное подводит нас к тому выводу, что единственно правильным будет рассматривать IQ и результаты многих других выборочных исследований не сами по себе, а с учетом размаха отклонений. Тогда «нормальным» будет считаться показатель не 100 пунктов, а в пределах, скажем, от 90 до 110. В этом случае будет некоторый смысл сравнивать ребенка, чей IQ попадает в эти рамки, с ребенком, чей коэффициент интеллекта выше или ниже данного диапазона. А вот проводить сравнения между цифрами, имеющими маленькую разницу, бессмысленно. Вам следует постоянно помнить об этом плюсе или минусе, то есть ошибке в ту или другую сторону, даже (или особенно) если ее пределы не указаны.

Те, кто пренебрегает ошибкой, которая изначально присуща любым исследованиям на основе выборки, рискуют совершить поразительно глупые поступки. Иные редакторы журналов носятся с результатами опросов читателей так, словно это истина в последней

инстанции, и главным образом по той причине, что совершенно не понимают, как их трактовать. Если им скажут, что одну статью читает 40 % мужской аудитории, а другую – только 35 %, они немедленно требуют больше статей наподобие первой.

Для журнала разница между показателями 35 и 40 % читательской аудитории действительно может быть значимой, а что касается самого опроса, то на деле этой разницы может и не быть вовсе. По финансовым соображениям выборка из читательской аудитории сводится к нескольким сотням респондентов, особенно после того, как отсеивают тех, кто вообще не читает данный журнал. Для журнала, адресованного главным образом женской аудитории, число мужчин в выборке может быть крайне мало. К тому моменту, когда опрошенные разделятся на категории тех, кто сообщил, что «прочитал всю», «прочитал почти всю», «прочитал часть» или «не читал вовсе» статью, ставшую предметом исследования, может оказаться, что эти 35 % рассчитаны на основе всего-то горстки ответов. Вероятная ошибка, что кроется за этим внушительным показателем 35 %, может быть столь велика, что для редактора полагаться на него – все равно что хвататься за тоненькую соломинку.

Случается, что большую шумиху разводят по поводу разницы с точки зрения математики реальной и доказуемой, но столь крохотной, что ее значение ничтожно. Кто так поступает, явно пренебрегает старой доброй поговоркой, что разница разнице рознь и имеет значение только та, что делает погоду. Наглядный пример тому – сыр-бор фактически на пустом месте, который с таким успехом и с такой выгодой для себя затеяли производители сигарет Old Gold.

Все начиналось довольно невинно – сидел себе редактор журнала *Reader's Digest*, покуривал сигареты, но тем не менее к курению относился с большим неодобрением. И вот его журнал развил бурную деятельность и привлек целые полчища усердных лаборантов, чтобы провести анализ дыма от сигарет нескольких разных марок. Далее журнал опубликовал на своих страницах результаты исследований, показывающие, сколько никотина и всякого прочего содержится в дыме сигарет разных марок. Вывод, который сделал журнал и который явно следовал из подробно расписанных результатов, состоял в том, что все марки сигарет практически одинаковы по содержанию вредных веществ и потому не имеет никакого

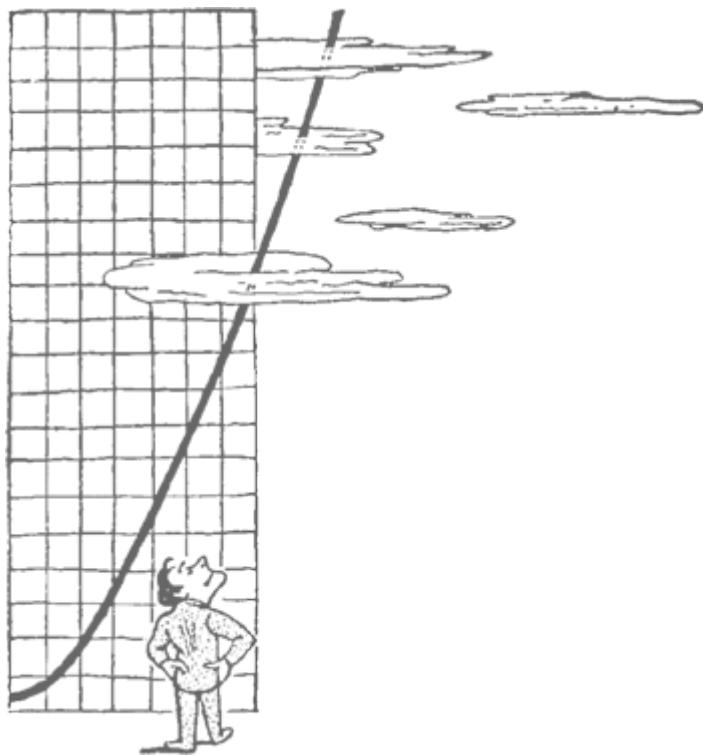
значения, сигареты какой марки вы курите.

Тут вам могла бы прийти мысль, что это стало огромным ударом для производителей сигарет и ловкачей из рекламных агентств, которые придумывают новые ухищрения для рекламных текстов. Казалось бы, опубликованные в *Reader's Digest* сведения начисто опровергают посулы табачной рекламы, что сигаретный дым, дескать, воздействует успокаивающе на горло и щадит кожу лица.

Но нашелся кое-кто, чей зоркий глаз заприметил кое-что интересное. В перечнях содержания вредных веществ, почти одинаковых для всех сигаретных марок, какая-то одна волей-неволей должна была оказаться в самом конце, и этой маркой была Old Gold. Вот оно! Во все концы тут же полетели телеграммы, и на страницах всех газет разом появились рекламные объявления, набранные самым крупным из возможных шрифтов. Заголовки и сам текст рекламы сообщали, что из всех марок сигарет, протестированных не кем-нибудь, а крупнейшим общеноциональным журналом *Reader's Digest*, дым сигарет Old Gold содержит меньше всего вредных веществ. При этом в рекламе отсутствовали какие бы то ни было цифры или хотя бы намеки, что разница между марками сигарет ничтожна.



Кончилось тем, что ушлым производителям Old Gold было велено, если следовать формулировкам судебных постановлений и административных приказов, «прекратить и воздерживаться впредь» от такой вводящей в заблуждение рекламы. Впрочем, особого толка от этого не было, ибо кто надо уже успел погреть руки на этой славной идее. Не зря еженедельник *New Yorker* говорит, что уж кто-кто, а рекламщик всегда сущется.



Глава 5 График – лучше не бывает

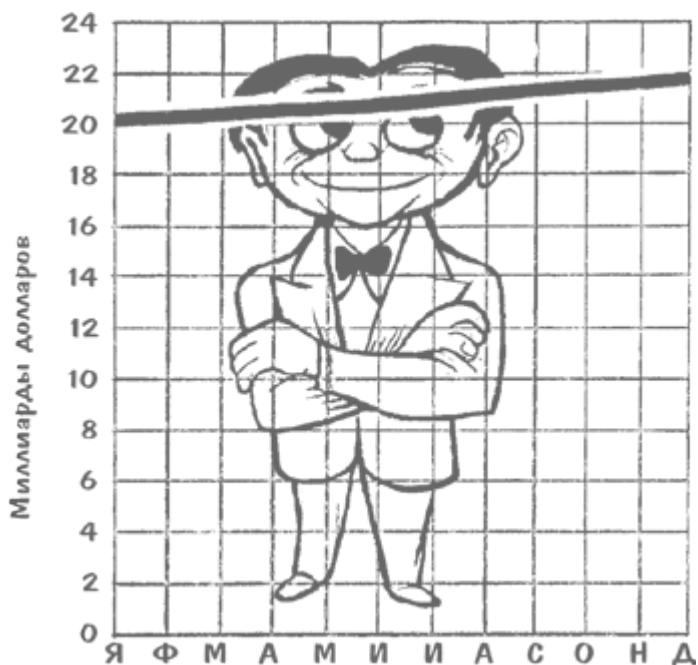
Цифры способны наводить ужас. Многие люди не отважились бы распространить на цифры ту уверенность, с какой Шалтай-Болтай поведал Алисе, что он хозяин слов, какие употребляет, и они ему подчиняются как миленькие. Должно быть, сказывается травма, нанесенная нам изучением арифметики в начальной школе.

В чем бы ни была причина, но числа действительно становятся камнем преткновения для писателя (который мечтает, чтобы его читали), рекламиста (который рассчитывает, что его рекламный текст обеспечит продажи товара) и издателя (который жаждет, чтобы его книги или журналы завоевали популярность). Когда цифры приводить в табличной форме категорически не допускается, а слова, как это нередко случается, бессильны, чтобы в точности обрисовать

картину, остается единственный выход – делать рисунки.

Пожалуй, простейшей разновидностью статистической картинки или графика будут всевозможные кривые. Они весьма полезны, когда нужно продемонстрировать те или иные тенденции, а практически все у нас заинтересованы в том, чтобы эти самые тенденции показать (а также узнать о них, посетовать на них или спрогнозировать). Сейчас мы сделаем так, что наш график наглядно покажет, как национальный доход США ежегодно увеличивается на 10 %.

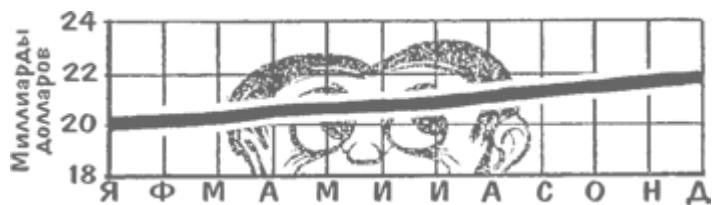
Для начала возьмите лист бумаги в клеточку. Напишите вдоль нижнего края названия месяцев. Сбоку расположите цифры, обозначающие миллиарды долларов. Теперь на получившейся координатной сетке отметьте точки, отражающие данные за каждый месяц, и соедините их линией. У вас получится график такого примерно вида:



Ну вот, теперь все достаточно наглядно. График показывает, как шли дела в течение года, и это можно проследить месяц за месяцем. Всякий сможет посмотреть и сразу понять, что к чему, поскольку график выполнен с соблюдением пропорций и внизу для сравнения имеется линия нулевой отметки. Ваши 10 % *выглядят* именно как 10 % – как восходящая тенденция, существенная, хотя и не сказать, чтобы такая уж впечатляющая.

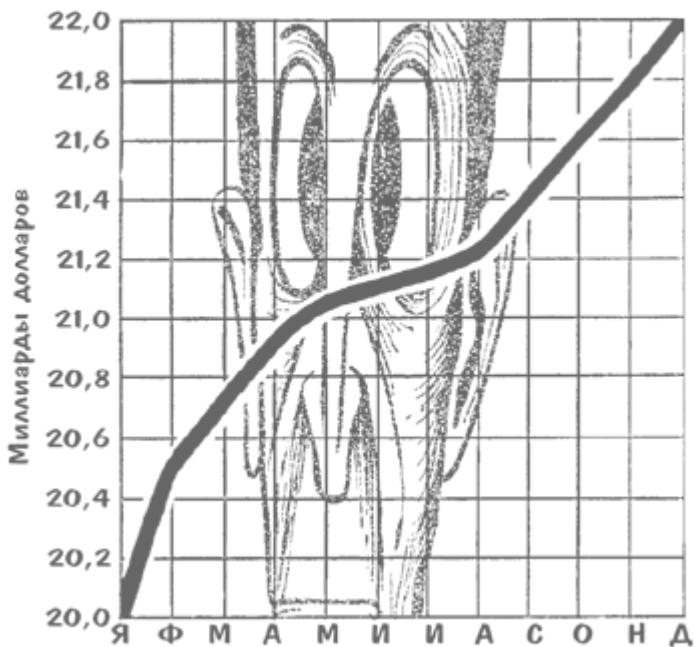
Этого вполне достаточно, если ваша задача только в том, чтобы передать информацию. А давайте предположим, что вы хотите одержать верх в споре, потрясти читателей, побудить их к действию

или что-то им продать. Но для этого вашему графику не хватает забористости, как-то он не впечатляет. А вы возьмите да и отрежьте нижнюю часть.



Ну вот, уже лучше. (К тому же вы сэкономили бумагу и можете указать на это, если какой-нибудь въедливый субъект попеняет вам за очковтирательский график.) Цифры на графике те же, и кривая выглядит совершенно аналогично. Это та же самая кривая. Никаких фальсификаций – разве что впечатление она производит другое. Но беглый читательский взгляд увидит теперь, что изображающая национальный доход кривая за двенадцать месяцев на рисунке преодолела полпути вверх, а все потому, что большая часть графика попросту отсутствует. Подобно опущенным частям предложения, с каковыми случаями вы наверняка сталкивались на уроках грамматики, отсутствующая часть графика «подразумевается». Глаз же, конечно, не в силах «подразумевать» то, что ему не видно, и скромный прирост благополучно превращается (в зрительном восприятии) в крупный.

Итак, вы уже попрактиковались в обмане, и почему бы не пойти дальше по пути усечений? Есть один хитрый трюк, который стоит дюжины таких, как вышеописанный. Он придаст вашему скромному десятипроцентному росту такой шикарный вид, какой не полагается и стопроцентному. Просто измените пропорции между осью ординат и осью абсцисс. Никакими правилами это не запрещается, зато чудо преобразит ваш график. Все, что от вас требуется, – задать единицу деления на оси ординат в десять раз меньшую, чем миллиарды долларов на исходном графике.



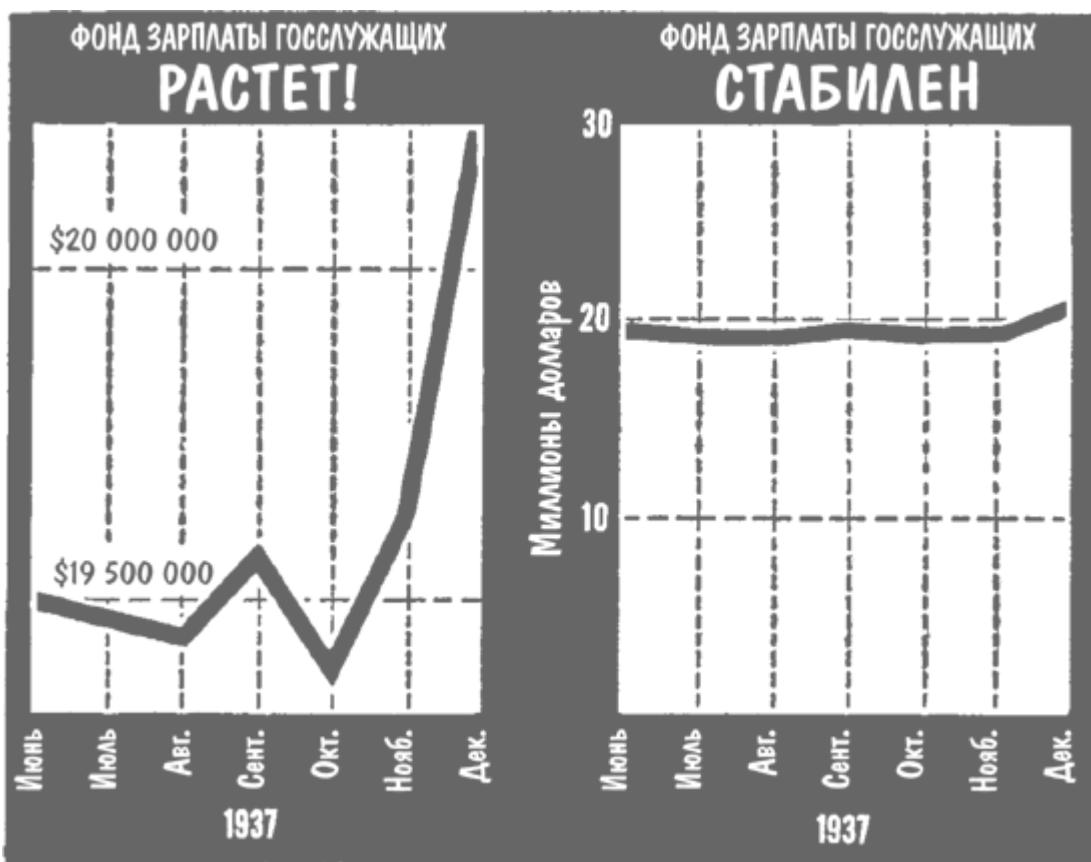
Вот это уже впечатляет, не правда ли? Всякий, глядя на этот график, может ощутить, что благотворные токи благополучия прямо-таки пульсируют в артериях страны. Это такой более утонченный способ отредактировать фразу «национальный доход возрос на 10 %», чтобы она звучала как «...взлетел на колоссальные 10 %». Однако графический способ намного эффективнее, поскольку не содержит прилагательных и наречий, способных разрушить впечатление, что информация имеет объективный характер. Так что никто не бросит в вас камень.

И к тому же вы окажетесь в очень славной или на худой конец респектабельной компании. Журнал *Newsweek*, например, прибег именно к такому методу, желая продемонстрировать, что в 1951 г. «акции достигли 21-летнего максимума», – редакция обрезала график на отметке 80, и дело с концом. А вот реклама газодобывающей компании *Columbia Gas System* в журнале *Time* за 1952 г. приводит выдержку из, как там указывается, «нашего свежего годового отчета». Если прочитать и проанализировать напечатанные мелким шрифтом числовые показатели, то обнаружится, что за десятилетний период стоимость жизни возросла примерно на 60 %, а стоимость газа понизилась на 4 %. Это вполне благоприятная картина, но ее явно не сочли таковой в *Columbia Gas System*. Они обрезали свои графики на отметке 90 % (причем на графике не наблюдается ни самого этого разрыва, ни каких-либо указаний, которые могли бы предупредить вас об этом). И потому ваши глаза видят вот что: стоимость жизни выросла больше чем втрое, тогда как стоимость газа на треть

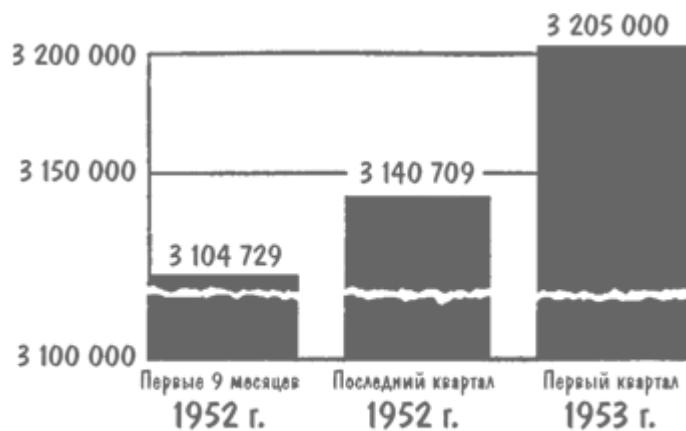
уменьшилась!

Сталелитейные компании пускают в ход аналогичные очковтирательские методы в попытках настроить общественность против идеи повышения заработной платы рабочим. И все же метод далеко не нов, и его несостоятельность задолго до сегодняшнего дня была изобличена – и далеко не в одних только узкоспециальных технических публикациях, адресованных статистикам. В редакционной статье *Dun's Review*¹² за 1938 г. автор воспроизвел рекламное объявление, которое пропагандировало размещение рекламы в столице США. Убедительный аргумент в пользу этого отлично передавал предварявший соответствующий график заголовок: «ФОНД ЗАРПЛАТЫ ГОССЛУЖАЩИХ РАСТЕТ!» О том же кричала и линия на графике, хотя стоявшие за ней цифры никак не оправдывали этих восторгов. Цифры всего лишь показывали, что расходы на зарплаты госслужащим с \$19 500 000 возросли до \$20 200 000. Зато красная кривая устремляясь от подножия графика прямиком к самому верху, отчего прирост в 4 % смотрелся как более чем в 400 %. Рядом журнал поместил свою версию этого графика – красного цвета кривая честно передавала рост в 4 %, а заголовок над ней гласил: «ФОНД ЗАРПЛАТЫ ГОССЛУЖАЩИХ СТАБИЛЕН».

¹² *Dun's Review* – обзор, который издавала в свое время американская компания Dun & Bradstreet, которая сейчас специализируется на том, что собирает, каталогизирует и анализирует информацию о субъектах бизнеса и составляет кредитные рейтинги, а также ведет крупнейший в мире реестр сведений о частных компаниях. Основана в 1841 г. – *Прим. пер.*



Журнал *Collier's* применил такой же метод, выводя на чистую воду столбиковые диаграммы в рекламных газетных объявлениях. Призываю обратить особое внимание на то, что на диаграмме середина столбиков вырезана:



Из рекламного газетного объявления от 24 апреля 1953 г., предоставлено журналу *Collier's*



Глава 6

Схематичная картинка

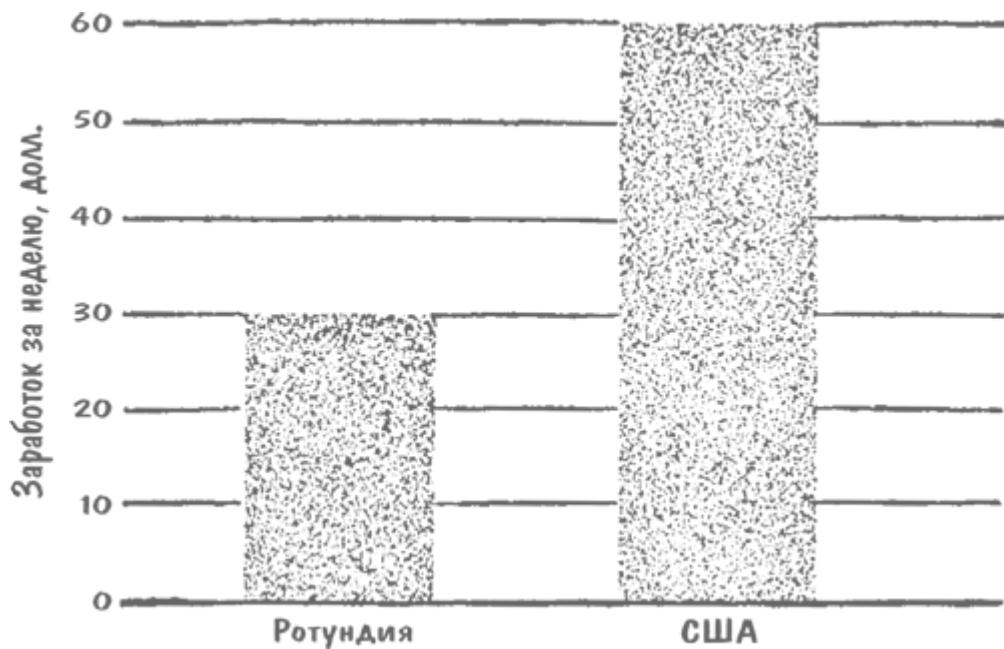
Лет десять назад вам наверняка довелось наслушаться всякого разного о маленьком человеке, под каковое определение подходим практически мы все. Позже это определение приобрело черты чрезмерной покровительственности, словно до нас снисходят, и тогда нас переименовали в простых людей. Вскоре, однако, об этом определении забыли, что, вероятно, не так уж плохо. Но маленький человек по-прежнему с нами. Он выступает персонажем разнообразных графиков и диаграмм.

Наглядной подачей информации называют графики или диаграммы, на которых фигурка человечка изображает миллион человек. Она помещается рядом с кучей денег на тысячу или миллиард долларов или контурным рисунком бычка, показывающим, сколько говядины вы потребите в следующем году. Это вещь полезная. И, боюсь, она радует глаз. Но наглядная графика способна стать красноречивым, нечистым на руку и не знающим поражений лжецом.

Породила наглядную графику, или пиктограмму, обычная столбиковая диаграмма – простой и распространенный прием подачи количественной информации, когда требуется сопоставить две или больше величины. Столбиковая диаграмма тоже умеет врать. Убедиться в этом легко. Достаточно внимательнее взглянуть на любой образец такой диаграммы, где столбики увеличиваются и в ширину, и в высоту, хотя и отображают одномерный фактор,

или когда изображаются трехмерные объекты, объемы которых трудно с ходу сопоставить. Столбиковая диаграмма с обрезанными посередине столбиками заслуживает точно такой же репутации, как и обрезанный график кривой, который мы обсуждали в предыдущей главе. Места обитания столбиковых диаграмм (их еще называют гистограммами) – книги по географии, корпоративные отчеты и новостные журналы. То же относится и к радирующему глаз отприску гистограммы – пиктограмме.

Предположим, мне понадобилось бы показать, как соотносятся две величины – средний недельный заработок плотников в США и в какой-нибудь, скажем, Ротундии. Цифры эти могли бы составлять \$60 и \$30 соответственно. Я хочу привлечь ваше внимание к этим данным, и потому вариант просто написать эти цифры меня не устраивает. И тогда я строю гистограмму. (Кстати говоря, если эти \$60 – почти ничего в сравнении с той огромной суммой, которую вам пришлось отвалить плотнику в прошлом году, когда потребовалось обновить перильца на веранде, примите во внимание, что тому плотнику не каждую неделю выпадает такая удача, как ваш заказ. И потом, я же не уточнил, какое среднее имею в виду или как я его вычислил, так что это не дает вам никаких оснований возражать. Видите теперь, насколько просто не подставиться, когда приводишь самую что ни на есть липовую статистику, если не указывать никаких других сведений? Вы, по всей видимости, уже догадались, что эти цифры взяты с потолка и я назвал их только для примера. Но ручаюсь, вам нипочем бы не догадаться об этом, если бы вместо \$60 я назвал \$59,83.)



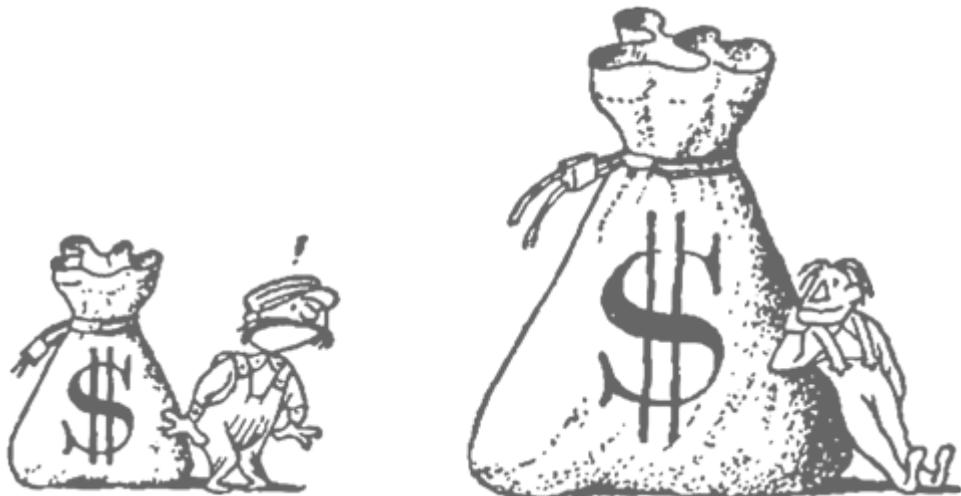
А вот и моя гистограмма: по вертикальной оси отложена величина недельного заработка в долларах. Картина ясная и правдиво показывает положение дел. Столбик на диаграмме, отображающий вдвое большую сумму, вдвое больше по размеру, таковым он и выглядит.



Только вот диаграмма эта как-то не особенно радует глаз, верно? А я могу запросто восполнить этот недостаток, если вместо столбиков использую другой графический объект, который лучше всего ассоциируется со звонкой монетой – мешки с деньгами. Одним мешком я изобразжу жалкие заработки плотника в Ротундии, а двумя мешками – заработки его американского собрата. А можно изобразить заработки в Ротундии тремя мешками, а в США – шестью. В любом случае рисунок по-прежнему будет правдив и ясен, а беглый взгляд на него никого не введет в заблуждение. Таким способом и создаются честные пиктограммы.

Способ этот и меня вполне устроил бы, если бы моя цель заключалась в том, чтобы представить эту информацию. Но мне-то этого мало. Я желаю заявить, что в Америке рабочий человек обеспечен несравненно лучше, чем в Ротундии, и чем сильнее я сгущу краски, преувеличивая разницу между \$30 и \$60, тем убедительнее будет это утверждение. По правде говоря (а этого я, разумеется, вовсе не собираюсь делать), я хочу навести вас на определенную мысль, создать у вас преувеличенное впечатление от моих данных, но при этом мне вовсе не хочется, чтобы меня уличили в подтасовке. Для этого есть свой прием, и его ежедневно пускают в ход, чтобы вас одурачивать.

А делается это так: сначала я нарисую мешок денег, изображающий тридцатидолларовый недельный заработок плотника в Ротундии, а затем нарисую второй мешок, размерами вдвое больше, и он будет изображать \$60, которые зарабатывает за неделю американский плотник. Пропорция соблюдена, не правда ли? Такая диаграмма создает как раз то впечатление, к которому я стремился. Рядом с внушительным мешком денег, которые зарабатывает американский плотник, заработка иностранца выглядит особенно ничтожным и жалким.



В этом, разумеется, и состоит подвох. Поскольку второй мешок в два раза выше первого, он еще и в два раза шире. И на бумаге он занимает не вдвое, а вчетверо больше места. Цифры свидетельствуют, что размеры заработков соотносятся как два к одному, тогда как глаза говорят нам (а зрительное впечатление в большинстве случаев преобладает над всеми остальными), что один заработка вчетверо больше другого. А могло быть и хуже. Ввиду

того, что на рисунках изображены предметы, которые в реальности трехмерны, второй мешок должен быть вдвое толще первого. А как написано в учебнике по геометрии, объемы подобных тел различаются на величину, равную одному из аналогичных параметров, возведенных в третью степень. Два умножить на два умножить на два будет восемь. Если в один мешок вмещается \$30, то второй, в восемь раз большего объема, должен вмещать не \$60, а \$240.

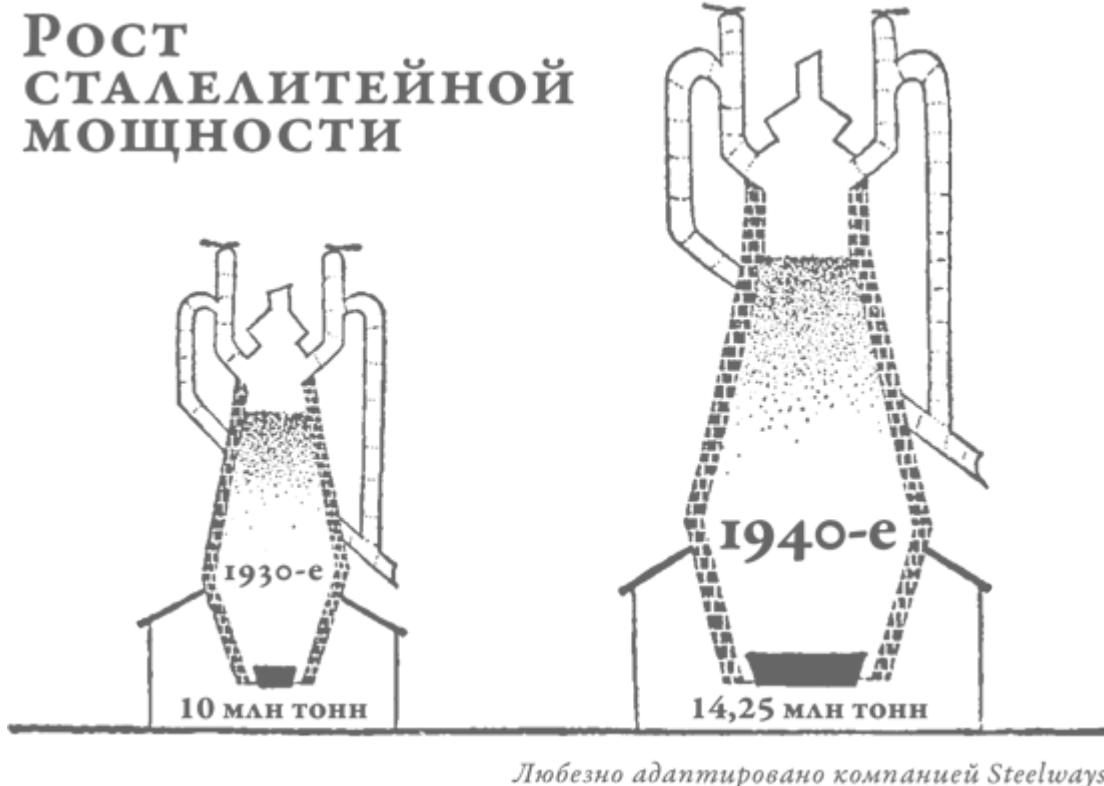
И именно такое впечатление создает моя хитроумно подправленная пиктограмма. Хотя слова у меня говорят «вдвое», у зрителя создается стойкое впечатление о подавляющем превосходстве – в соотношении восемь к одному.

К тому же вам трудно будет приписать мне какой-либо преступный умысел. Я всего лишь делаю то же, что практикуют и многие другие. Журнал *Newsweek* однажды проделал такое – и тоже использовал картинки мешков с деньгами.

Американский институт сталелитейной промышленности тоже прибег к подобному трюку, только его пиктограммы изображали доменные печи. Задача состояла в том, чтобы показать, какой бум роста сталелитейных мощностей наблюдался в 1930–1940-е гг., и тем самым продемонстрировать, что отрасль добилась этого своими силами и какого бы то ни было вмешательства государства в ее дела не требуется. Сам этот принцип достоин большего уважения, чем манера, в какой он был подан. Доменную печь, которая изображала дополнительные сталелитейные мощности размером 10 миллионов тонн, появившиеся за 1930-е гг., нарисовали высотой в две трети от второй доменной печи, обозначавшей прирост мощностей в 14,25 миллиона тонн за 1940-е гг. При взгляде на рисунок видишь две доменные печи, причем одна из них примерно втрое больше другой. Говоришь «почти в полтора раза», а воспринимается как «втрое» – вот каких чудес можно добиться с помощью схематичного изображения.

Вышеупомянутый шедевр, сотворенный руками сталелитейщиков, интересен еще в ряде аспектов. Правая домна невесть почему раздалась в ширину по сравнению с пропорциями своей соседки, а черная полоса в чреве правой домны, призванная изображать расплавленный чугун, стала в два с половиной раза длиннее, чем у домны, изображающей предыдущее десятилетие. Исходные цифры указывают на увеличение мощностей на 50 %,

а изобразили его как все 150 %, чтобы создавалось зрительное впечатление увеличения более чем на 1500 % – если только мы с моей логарифмической линейкой не ошиблись в расчетах. Арифметика становится фантазией.



(Было бы, наверное, совсем уж немилосердно упомянуть, что та же красочная глянцевая страница с двумя домами предлагает еще и прямо-таки первосортный образчик обрезанного линейного графика. Кривая преувеличивает показатель роста сталелитейных мощностей на душу населения, красиво устремляясь вверх, но нижняя часть диаграммы отсутствует. Это экономит бумагу и удваивает темп прироста мощностей.)

Отчасти в этом может быть повинна всего лишь небрежность чертежника. Но скорее всего, тут не обошлось без плутовства: если кассир все время ошибается в свою пользу, поневоле заподозришь неладное.

Newsweek как-то раз показал, что «пожилые американцы делаются старше» при помощи диаграммы, на которой помещались фигурки двух мужчин: одна изображала сегодняшнюю ожидаемую продолжительность жизни в 68,2 года, вторая – продолжительность жизни американца в 34 года в 1879–1889 гг. Ну что сказать, та же старая история: одна фигурка была вдвое выше другой, а значит,

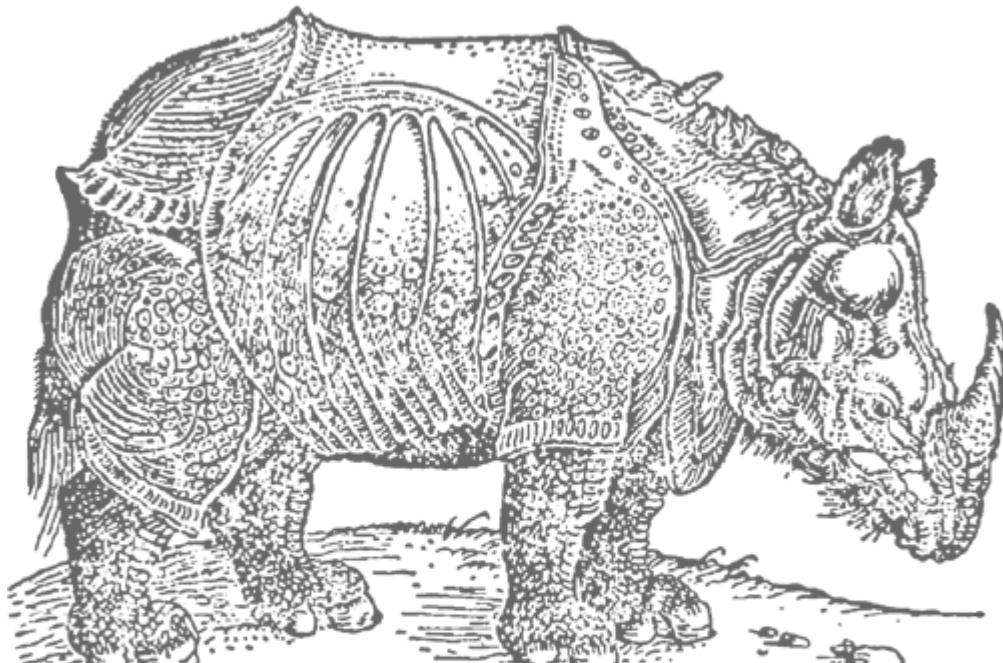
раз в восемь больше по величине. Картинка намеренно раздувала разницу, придавая ей сенсационности, чтобы приукрасить истинное положение дел. Я бы назвал это формой бульварной журналистики. Тот же выпуск *Newsweek* содержал чудо как обрезанный линейный график.

КОРОВА РАСТЕТ ПРЯМО НА ГЛАЗАХ



Попытки играть с размерами графических объектов чреваты и тем, что могут привести к недоразумению еще и другого рода. В 1860 г. поголовье молочных коров в США составляло более 8 миллионов, а в 1936 г. оно превысило 25 миллионов. Графическая подача этого факта с помощью рисунка, где изображены две коровы (причем вторая втрое выше первой в холке), несомненно, значительно усилит впечатление в той самой манере, о какой мы говорили выше. Но беглый взгляд на эту картинку может создать эффект неожиданный и даже странный: читателю может показаться, будто коровы теперь куда больше, чем были в прежние времена.

НОСОРОГИ МЕЛЬЧАЮТ



1515 г.



1936 г.

Стоит только применить эту надувательскую методу для демонстрации того, как уменьшается популяция носорогов, и смотрите, что у вас получится. Сатирик Огден Нэш¹³ в одном из своих стихотворений срифмовал «*rhinostorous*» с «*preposterous*», «носорожью стать» с чем-то до невозможности нелепым. Такое название подойдет и к вышеописанному методу – он тоже завирается порой до полного абсурда.

¹³ Огден Нэш (1902–1971) – знаменитый американский поэт-сатирик, прославился афористическими стихами, в которых остроумно, но не зло высмеивал напыщенность, тупость, нелепые пристрастия и причуды. В упомянутом стихотворении сделан акцент на несообразность «конструкции» слона, в том числе «снабжен удобными ручками с двух сторон» и «кожа у него вся в складках, никакой моли не по зубам». А если вам, обращается Нэш к читателю, зверюга под названием «слон» кажется нелепой, значит, вы еще не видали носорожьей стати. – *Прим. пер.*



Глава 7

Псевдообоснованная цифра

Если не получается доказать то, что вы хотите доказать, продемонстрируйте нечто другое и настаивайте, что это то же самое. В ошеломлении, которое всегда сопровождает столкновение статистики с человеческим разумом, едва ли хоть кто-нибудь заметит подмену. Привязать цифру, отражающую какой-то факт, к другому факту – прием известный и всегда сослужит вам добрую службу. Действует безотказно.

Допустим, вы не можете доказать, что ваше замечательное лекарство излечивает от простуды, но никто не мешает вам напечатать (крупным шрифтом) результаты настоящего лабораторного исследования: полкиапли лекарства, помещенные в пробирку, через 11 секунд уничтожают 31 108 бактерий. И если уж взялись делать это, убедитесь, что упоминаемая вами лаборатория имеет надежную репутацию или внушительное название. Отчет воспроизведите полностью. Сделайте фото человека в белом халате и поместите рядом с отчетом.

Но не следует упоминать в своей истории несколько каверзных нюансов. Не вам же – или все-таки вам? – уточнять, что антисептические капли, так прекрасно показавшие себя в лабораторной пробирке, возможно, не будут действовать столь же эффективно в горле у больного, особенно когда он разбавит их,

как того требует инструкция, чтобы не обжечь себе ткани носоглотки. Не наводите тень на плетень уточнением, бактерии какого рода пали жертвой вашего средства. Кто ж его знает, какие бактерии вызывают простуду, особенно если возбудители простуды – вовсе не бактерии?

По правде говоря, не существует известной связи между разными видами бактерий, используемыми для лабораторного эксперимента, и тем, что вызывает простуду, но люди обычно не размышляют так уж серьезно над подобным вопросом, особенно когда у них хлюпает в носу.

Возможно, данный прием слишком очевиден, и люди нет-нет да и начинают догадываться, в чем тут подвох, хотя едва ли он так уж бросается в глаза, когда читаешь рекламу. Тем не менее найдется прием и похитрее.

Будем считать, что во времена, когда нарастают расовые предрассудки, вас наняли доказать, что ничего подобного не происходит. Задачка эта не особенно трудна. Устройте опрос, а еще лучше – пусть его проведет какая-нибудь организация, пользующаяся добной репутацией. Следует сделать представительный подбор общества и провести опрос: имеют ли, по мнению респондентов, чернокожие такие же шансы получить работу, как белые. Опросов проведите несколько, через определенные промежутки времени, чтобы отразить в своем отчете тенденцию.

Однажды данный вопрос исследовало Бюро по изучению общественного мнения при Принстонском университете. То, что обнаружили специалисты Бюро, представляет собой любопытное свидетельство, что некоторые вещи, особенно при опросах общественного мнения, на самом деле совсем не таковы, какими кажутся. Каждому, кого спрашивали о шансах чернокожих на трудоустройство, попутно задавали несколько вопросов, чтобы определить, не имеет ли респондент сильного расового предубеждения. Оказалось, что респонденты с самыми сильными расовыми предрассудками на вопрос о возможностях трудаустройства чернокожих чаще всего давали ответ «Да». (Было установлено также, что около двух третей респондентов, благосклонно настроенных по отношению к чернокожим, не считают, что у тех шансы трудастроиться такие же, как у белых. А около двух третей тех, кто выказал предубеждение против чернокожих, заявили, что чернокожим предоставляются такие же хорошие отпуска, как белым.) Очевидно, что из этого опроса мало что узнаешь

об условиях трудоустройства чернокожих, зато можно почерпнуть много интересного о межрасовых отношениях.

Теперь вы видите, что, если расовые предрассудки в период проведения опроса нарастают, у вас получается больше ответов, указывающих, что у чернокожих не меньше возможностей трудоустройства, чем у белых. И вот вы объявляете результаты: согласно опросу, отношение к чернокожим неизменно справедливо.

Вы добились поистине выдающегося результата за счет добросовестного использования псевдообоснованной цифры. Чем хуже положение дел в реальности, тем больше приукрашивает его ваш опрос.

А вот еще пример: «27 % опрошенных знаменитых врачей курят сигареты Throaties – больше, чем любую другую марку». Сама цифра может быть липовой во множестве отношений, но в сущности это ничего не меняет. Единственное, что можно сделать, когда вам сообщают такую бесполезную информацию, это поинтересоваться: «Ну и что?» При всем уважении к врачебной профессии хочу усомниться: неужели докторам действительно известно о марках сигарет намного больше вашего? Неужели врачи располагают какой-то информацией не для посторонних глаз, которая и помогает им выбрать самые безвредные из всех сигареты? Разумеется, никакой такой специальной информации у врачей нет, и ваш доктор первым подтвердит это. И все же эти «27 %» почему-то умудряются производить такое впечатление, будто они действительно что-то означают.



А теперь откинем 1 % и рассмотрим случай с соковыжималкой. Ее в свое время широко рекламировали как приспособление,

позволяющее «выжать на 26 % больше сока», что «подтверждают лабораторные испытания» и за что «ручается Институт домашнего хозяйства».

Все это отлично звучит. Если можно приобрести соковыжималку, которая на 26 % эффективнее, то зачем покупать другие? Пока что оставим в стороне тот факт, что «лабораторные испытания» (и особенно «независимые лабораторные испытания») подтверждали порой самые невероятные вещи, и просто поинтересуемся, что означает приведенная цифра. На 26 % больше чего? Когда докопались до сути, оказалось, что данная соковыжималка позволяет извлечь на 26 % больше сока, чем стародавняя ручная соковыжималка. Согласитесь, эта цифра не имеет ничего общего с информацией, которой вы могли бы заинтересоваться, собираясь сделать покупку. Эта соковыжималка с таким же успехом может быть самой худшей на рынке в данной категории товаров. Помимо того, что сама цифра уж очень точная и одним этим внушает подозрения, она еще и не имеет никакого отношения к делу.

Рекламодатели не единственные, кто горазд дурачить при помощи цифр, если вы им это позволяете. Статья о безопасности езды, опубликованная в журнале *This Week*, бесспорно, из соображений самой искренней заботы о вас, рассказала, что может случиться, если вы «мчитесь по шоссе на скорости 70 миль в час, виляя из стороны в сторону». В этом случае, рассказывается в статье, у вас в четыре раза больше шансов остаться в живых, если дело происходит в семь часов утра, а не в семь вечера. А вот и подтверждение: «В семь часов вечера на автострадах происходит вчетверо больше аварий со смертельным исходом, чем в семь утра». Так оно приблизительно и есть, но вывод, который делается в статье, из этих данных вовсе не следует. По сравнению с утренними часами, в вечернее время в дорожных авариях погибает больше людей по той причине, что вечером на дорогах в принципе больше народа – вот и жертв больше. Вы же, отдельно взятый водитель, действительно можете подвергаться большей опасности на дороге в вечернее время, но ничто в приведенных цифрах не свидетельствует ни в пользу, ни против этого.

При помощи такой же логической бессмыслицы, какую использовал автор статьи, можно доказать, что в ясную погоду ездить по дорогам опаснее, чем в туманную. В ясную погоду аварий происходит значительно больше, потому что она выдается чаще

туманной. Но все равно, в туман езда на автомобиле куда опаснее.



Изучая статистику аварий, вы можете запугать себя до смерти относительно любого вида транспорта... если не сообразите, насколько некорректно привязаны приводимые цифры к явлениям, которые они призваны характеризовать.

В 1953 г. в авиакатастрофах погибло больше людей, чем в 1910 г. Должен ли отсюда следовать вывод, что авиаперелеты стали более опасны? Чепуха. Просто люди стали летать в сотни раз больше, чем раньше, вот и всё.

Или вот сообщают, что за один из прошлых годов в происшествиях на железной дороге погибли 4712 человек. Это, согласитесь, веский аргумент, чтобы отказаться от железнодорожных поездок в пользу передвижения на личном автомобиле. Но когда вы решите разобраться, что на самом деле означает указанная цифра, то увидите, что дело тут совсем в другом. Почти половину жертв составляют те, кто ехал в автомобилях, столкнувшись с поездами на железнодорожных переездах. А что до большей части остальных жертв, то эти люди погибли из-за того, что ехали на сцепках между вагонами. И только 132 человека из 4712 были пассажирами поездов. Но даже эта цифра немногого стоит, если она не привязана к сведениям об общем пассажирообороте на железной дороге.

Допустим, вас беспокоят ваши шансы погибнуть при пересечении страны от побережья до побережья. Вы не получите сколько-нибудь пригодной для выводов информации, если начнете наводить справки, на каком из видов транспорта (железнодорожном,

авиационном или автомобильном) в прошлом году погибло больше всего людей. Определить степень вероятного риска позволяют данные о количестве смертельных жертв на миллион пассажиромиль по каждому из трех видов транспорта. Эти показатели точнее других подскажут вам, на каком виде транспорта вы больше всего рискуете жизнью.

Существует много других способов для расчета показателя чего-нибудь, с тем чтобы впоследствии выдать это за что-то другое. Общий метод состоит в том, чтобы взять две вещи, которые довольно похожи, но на деле вовсе не одинаковы. В качестве начальника отдела кадров компании, у которой возникли трения с профсоюзом, вы «проводите опрос» сотрудников, чтобы выяснить, у скольких из них имеются жалобы на профсоюз. Если только сам профсоюз не есть сорвище ангелов с архангелом во главе, можете спокойно задавать свой вопрос и честно записывать ответы, а потом представить это за доказательство, что у большинства сотрудников действительно есть какие-то жалобы. На основе собранных данных вы составляете доклад, где говорится, что «подавляющее большинство (78 %) сотрудников настроены против профсоюза». По сути вы сделали вот что: собрали в одну кучу все без разбора жалобы и мелкие конфликты, а затем выдали их за нечто другое, что выглядит примерно так же. Вы так и не доказали того, что требовалось, а подано все так, словно доказали, не правда ли?

Но в каком-то смысле все было по-честному. Со своей стороны профсоюз может с такой же легкостью «доказать», что практически все сотрудники возражают против порядка управления компанией.

Если хотите поохотиться на якобы обоснованные цифры, могу предложить вам изучить корпоративные финансовые отчеты. Высматривайте прибыли, которые могли выглядеть слишком высокими и потому были замаскированы под что-то другое. Вот как описывает этот прием журнал *Ammunition*, рупор Объединенного профсоюза рабочих автомобильной промышленности Америки:

В отчете говорится, что за прошлый год прибыль компании составила \$35 миллионов. Всего каких-то полтора цента с каждого полученного от продаж доллара! Вам уже жалко эту компанию. Вот перегорит у них в уборной лампочка. Чтобы заменить ее, придется потратить 30 центов. А это как-никак прибыль от выручки в 20 долларов. Отсюда и до экономии

на бумажных полотенцах недалеко!

Но, конечно же, истина заключается в том, что в статье «Прибыль» компания показывает лишь половину, а то и треть своих истинных доходов. Та часть, о которой умалчивается, скрыта в амортизации, специальной амортизации и в резервах на случай непредвиденных расходов.

Похожая штука и с процентами. Компания General Motors, отчитываясь за девятимесячный период, показала относительно скромную прибыль (после вычета налогов) – 12,6 % от объема продаж. Но за тот же период прибыли компании от инвестиций составили 44,8 %, что выглядит куда как хуже – или лучше, в зависимости от предмета спора, в котором вы желаете одержать верх.

Подобным же образом читатель журнала *Harper's* в разделе «Письма от читателей» принял заступаться за сеть продовольственных магазинов самообслуживания A&P¹⁴, указывая, что у них низкий уровень чистой прибыли – всего каких-то 1,1 % от объема продаж. «Возможно ли, чтобы какой-нибудь американский гражданин испугался, что общество его заклеймит как спекулянта... если он выручает на толику больше, чем \$10, с каждой инвестированной в течение года \$1000?» – спрашивал читатель.

Ну да, навскидку эти 1,1 % выглядят прямо-таки прискорбно малой величиной. А сравните их с 4–6 % (а то и более), которые у большинства из нас на слуху в связи с ипотечными кредитами Федерального управления жилищного строительства, банковскими кредитами и прочим. Разве не лучше зажилось бы A&P, если бы она бросила торговаться продовольственными товарами, поместила свои капиталы в банк да и жила бы поживала на проценты?

Подвох в том, что годовая рентабельность инвестиций – совсем не тот коленкор, что прибыль на общий объем продаж. В одном из следующих номеров *Harper's* другой читатель возразил на сетования по поводу бедственного положения A&P: «Если я каждое утро буду покупать какой-то товар за 99 центов, а днем продавать его за \$1, я заработаю на общем объеме продаж всего 1 %,

¹⁴ A&P, The Great Atlantic and Pacific Tea Company – старейшая в мире сеть супермаркетов, как мы бы сказали сегодня, однако сам термин «супермаркет» вошел в оборот примерно в середине 1960-х гг. – позже, чем была написана эта книга. – Прим. пер.

но при этом 365 % на инвестированном капитале за год».

Любую количественную величину несложно выразить множеством разнообразных способов. Вы можете, например, представить один и тот же факт, называя его доходностью продаж в 1 %, рентабельностью инвестиций в 15 %, десятимиллионной прибылью, ростом прибылей на 40 % (по сравнению со средним показателем за 1935–1939 гг.) или сокращением на 60 % по сравнению с предыдущим годом. Суть в том, чтобы выбрать формулировку, которая лучше всего подходит для текущих надобностей. А после остается уповать на то, что лишь единицы, читая эту информацию, сообразят, насколько она искажает реальное положение дел.

Не все псевдообоснованные цифры порождены умышленным желанием обмануть. Многие статистические данные, в том числе из области медицины (а они довольно важны для всех нас), искажаются из-за непоследовательности и рассогласованности в учете информации на местах. Так, обнаруживаются поразительные расхождения в показателях по таким щекотливым вопросам, как abortionы, внебрачное деторождение и заболеваемость сифилисом. А если бы вам потребовалась самая свежая и доступная статистика по гриппу и пневмонии, цифры могли бы подтолкнуть вас к диковинному выводу, что эти заболевания распространены только в трех южных штатах, поскольку на них приходится порядка 80 % всех зарегистрированных случаев. Но в действительности высокий процент объясняется тем фактом, что в этих трех штатах все еще действует официальное требование представлять сведения по заболеваемости гриппом и пневмонией, тогда как остальные штаты этого больше не делают.

Столь же мало смысла и в некоторых статистических данных по малярии. В тех районах Американского Юга, где до 1940 г. отмечались сотни тысяч случаев малярии в год, на сегодняшний день зарегистрированы лишь единичные случаи, что воспринимается как благотворные и явно значимые перемены к лучшему, произошедшие за какие-то несколько лет. В действительности изменились лишь правила регистрации заболеваний – сейчас учитываются (и отражаются в статистике) только подтвержденные случаи малярии, тогда как прежде на большей части Юга малярией в просторечии называли также простуду и лихорадку.

Уровень смертности в военно-морском флоте США в период

Испано-Американской войны¹⁵ составлял девять человек на тысячу. За тот же период уровень смертности среди гражданского населения Нью-Йорка достигал шестнадцати человек на тысячу. Позже эти цифры использовали вербовщики, чтобы показать: служить в ВМС безопаснее, чем находиться за его пределами. Допустим, что сами эти цифры точны (вероятно, так оно и есть). Давайте остановимся на мгновение и проверим, сообразите ли вы, что лишает практически всякого смысла сами эти цифры, или хотя бы заключение, которое выводили из них вербовщики.

Все дело в том, что группы, к которым относятся вышеуказанные цифры, несопоставимы. В рядах ВМС служат главным образом молодые мужчины, признанные здоровыми. Гражданское же население состоит среди прочего из малых детей, старииков и больных, и для этих категорий населения уровень смертности выше, где бы они ни находились. Приведенные цифры вообще никак не доказывают, что мужчины, признанные годными к службе в соответствии со стандартами ВМС, находясь в его рядах, проживут дольше, чем если бы не служили на флоте. Впрочем, эти цифры не доказывают и обратного.



Вам, должно быть, приходилось слышать удручающую новость, что 1952 г. был худшим по ситуации с полиомиелитом за всю историю медицинских наблюдений. Этот вывод основывался на исчерпывающих, казалось бы, свидетельствах, какие только можно вообразить: в том году было зарегистрировано намного больше случаев заболевания полиомиелитом, чем в каком-либо другом.

¹⁵ Военный конфликт между США и Испанией в 1898 г. – Прим. ред.

Но когда эксперты взялись изучать исходные цифры, вскрылись и кое-какие более обнадеживающие факты. Во-первых, численность детей в возрасте, когда восприимчивость к заболеванию максимальна, в 1952 г. была так велика, что, если бы сохранился прежний уровень заболеваемости, количество случаев достигало бы рекордного значения. Другой отрадный факт таков: на фоне всеобщей бдительности по поводу полиомиелита его стали диагностировать чаще и фиксировать даже неострые формы заболевания. Наконец, в тот год усилились финансовые стимулы к большему вниманию к этой болезни, что выразилось в возросшем предложении страхования от полиомиелита, а также в увеличении возможностей получить поддержку и помочь со стороны Национального фонда борьбы с детским параличом. Выявленные факты позволяют серьезно усомниться в истинности утверждения, что распространение полиомиелита достигло нового пика, а данные о смертности от данного заболевания только усиливают эти сомнения.

Любопытный феномен: такие показатели, как уровень смертности или количество смертей, зачастую служат более точным мерилом распространенности заболевания, чем данные о количестве заболевших. А все потому, что, когда дело касается летального исхода, регистрация и учет поставлены на более качественный уровень. Это тот самый случай, когда цифра откровенно псевдообоснованная лучше той, что в первом приближении выглядит совершенно обоснованной.

В Америке псевдообоснованные цифры переживают бум раз в четыре года. Впрочем, это не свидетельствует о циклической природе таких цифр, а просто напоминает, что именно с такой периодичностью проходят выборы. Предвыборное заявление, обнародованное Республиканской партией в октябре 1948 г., целиком и полностью построено на цифрах. Создается видимость, что эти цифры связаны друг с другом, но это не так:

Когда Дьюи в 1942 г. был избран на пост губернатора, минимальный размер зарплаты учителей в некоторых районах составлял такую малость, как \$900 в год. Сегодня школьные учителя в штате Нью-Йорк получают самые высокие зарплаты в мире. По рекомендации губернатора Дьюи, которая основывалась на сведениях, полученных в ходе работы

назначенного им комитета, легислатура¹⁶ штата выделила из бюджета штата \$32 000 000 на обеспечение немедленного повышения заработной платы школьным учителям. В результате минимальный размер зарплаты учителя в Нью-Йорке варьируется в пределах от \$2500 до \$5325.

Совершенно не исключено, что мистер Дьюи проявил себя как друг учителей, да только приведенные цифры об этом не свидетельствуют. Это старый как мир трюк с «было» и «стало», когда для показа разительных перемен втихомолку приводят в действие ряд факторов, а потом представляют дело так, будто эти факторы ни при чем. Здесь у нас имеется «было» \$900 и «стало» от \$2500 до \$5325. Это, бесспорно, создает впечатление, что положение улучшилось. Но меньшая цифра отражает нижний порог зарплаты учителя в каком-нибудь сельском районе штата, а цифры побольше – диапазон заработных плат учителей в самом Нью-Йорке. Может быть, при губернаторе Дьюи улучшения действительно произошли, а может быть, и нет.

Вышеприведенное заявление иллюстрирует статистическую форму картинки «было – стало», которая представляет собой знакомый всем нам по журнальным статьям и рекламе трюк. Гостиную дважды фотографируют, чтобы показать вам, какие крупные перемены к лучшему может сотворить слой краски на стенах. Но между этими двумя демонстрациями в гостиной появляется новая мебель. А еще бывает, что фотография «было» меньше размерами, сделана при слабом освещении и вообще черно-белая, а вот фотография «стало» – цветная и гораздо крупнее. Или пара картинок показывает вам, что бывает, когда юная особа начинает применять ополаскиватель для волос. И – бог ты мой! – ее шевелюра действительно выглядит значительно лучше после, чем до. Но при внимательном изучении вы замечаете, что перемен добились главным образом за счет того, что девушку заставили улыбнуться, а ее волосы сзади подсвечены ярким светом. Тут скорее следует отдать должное мастерству фотографа, чем чудодейственной силе ополаскивателя.

¹⁶ Легислатура штата – так называется законодательный орган штата Нью-Йорк. – *Прим. пер.*





Вот две часы, они очень точно показывают время.

Когда на часах А стрелка указывает на цифру, бьют часы В. Вызван ли бой часов В часами А?

Глава 8

И снова это «после – значит вследствие»

Однажды некто взвалил на себя тяжкий труд определить, отстают ли в учебе студенты-курильщики от некурящих студентов. Как выяснилось – да, отстают. Очень многим этот вывод пришелся по душе, и они пошли еще дальше. Хочешь получать высокие оценки – бросай курить, так оно вроде бы верно. А следующий обоснованный вывод – курение пагубно влияет на умственные способности.

Вышеупомянутое исследование, как мне верится, было проведено по всем правилам: объем выборки был достаточно велик, подбрали ее добросовестно и тщательно, величина корреляции действительно оказалась очень значимой, и все прочие условия были соблюдены.

И все же не обошлось без одного заблуждения. И пускай оно древнее как мир, но почти всегда обнаруживается в статистических данных, замаскированное в гуще внушительных цифр. Это заблуждение таково: если событие В следует за событием А, значит, событие А является причиной события В. В вышеупомянутом исследовании делается неоправданное предположение, что поскольку курение и низкие оценки сопутствуют друг другу, значит, курение и есть причина плохой успеваемости. Но разве все это нельзя

с таким же успехом перевернуть наоборот? Может быть, именно скверные оценки заставили студентов искать утешения, но не в выпивке, а в курении? Если уж на то пошло, данный вывод столь же вероятен и не хуже подкреплен фактами. Вот только всяких пропагандистов он не слишком устраивает.



Представляется куда более вероятным, что ни одно из этих двух явлений не обуславливает другого. Скорее оба они следствие какого-то третьего фактора. Может ли быть так, что компанейские парни, которые не слишком утружают себя учебой, чаще имеют привычку к курению? Или секрет в том, что имеется корреляция между экстраверсией и низкой успеваемостью – связь очевидно более тесная, чем между оценками и умственными способностями? Очень может быть, что экстраверты более склонны к курению, чем интроверты. Это я к тому, что, когда имеется множество правдоподобных объяснений, у вас едва ли есть право выбрать то из них, которое вам больше нравится, и настаивать на его истинности. И все же многие поступают именно так.

Чтобы не поддаваться заблуждению «после – значит вследствие» и тем самым не уверовать в правоту многих ложных истин, следует подвергать любое утверждение самому тщательному анализу. Корреляция, эта убедительно точная зависимость, которая на первый взгляд показывает, что одно событие происходит вследствие другого,

бывает нескольких типов.

Существует корреляция, обусловленная случайными причинами. Возможно, у вас получилось установить корреляцию между двумя рядами чисел, чтобы доказать некое маловероятное утверждение. Но если вы снова попробуете проделать расчет, но уже на других цифрах, никакого доказательства не получится. Подобно производителю зубной пасты, которая, какказалось, препятствует развитию кариеса, вы просто отбрасываете неугодные вам результаты и широко тиражируете те, что подходят для ваших целей. Если выборка невелика, то, скорее всего, вы обнаружите существенную корреляцию между двумя характеристиками или событиями, которые представляют для вас интерес.

Распространенный случай ковариации¹⁷ – взаимосвязь действительно существует, но нельзя сказать, какая из переменных выступает причиной, а какая следствием. В ряде подобных случаев причина и следствие время от времени могут меняться ролями или даже обе одновременно будут и причиной, и следствием. К ковариации такого рода можно отнести корреляцию между доходом и владением ценными бумагами. Чем больше денег вы зарабатываете, тем больше акций покупаете, а чем больше у вас акций, тем больше доход. Так что утверждать, что одно влечет за собой другое, было бы некорректно.

Пожалуй, самый коварный тип ковариации представляет собой тот широко распространенный случай, когда ни одна из рассматриваемых переменных не оказывает никакого воздействия на другую, но при этом значимая корреляция между ними действительно подтверждается расчетами. Скольким недостойным делам послужил этот тип корреляции! К этой категории относятся и утверждения о низкой успеваемости курильщиков, равно как и множество прочих статистических данных из области медицины, которые обычно приводят, забывая уточнить, что, хотя корреляция действительно существует, выведенные на ее основе причинно-следственные связи взяты с потолка. В качестве примера такой чепухи, или фиктивной корреляции, которая сама по себе есть непреложный статистический факт, некто ради смеха называл такой: существует тесная корреляция между жалованьем пресвитерианских священников в Массачусетсе и ценой на ром в Гаване.

¹⁷ Ковариация – мера линейной зависимости двух случайных величин. – Прим. ред.

Что здесь причина, а что следствие? Иными словами, священники ли наживаются на продаже рома или они ей способствуют? Ну ладно. Все это сильно притянуто за уши, и абсурдность утверждения видна с первого взгляда. Но осторегайтесь случаев, когда используется та же самая логика (после – значит вследствие), только в отличие от вышеприведенного примера ее применяют более искусно и тонко. В случае со священниками и ромом легко понять, что обе цифры растут под действием третьего фактора – исторического и общемирового роста цен практически на все.

А возьмите цифры, указывающие, что уровень самоубийств достигает максимума в июне. Это что, самоубийства порождают такое количество новобрачных – или столь распространенный обычай сочетаться браком именно в июне провоцирует тех, кто отвергнут возлюбленными, свести счеты с жизнью? Несколько более убедительное (хотя тоже недоказанное) объяснение может быть таково: некто отчаявшийся всю зиму борется со своим угнетенным состоянием в надежде, что весной тучи рассеются, но окончательно сдается, когда наступает июнь, а никакого просвета нет.



Не помешает держать ухо востро и в отношении суждений, полученных в результате распространения корреляции за пределы данных, на которых она продемонстрирована. Проще простого показать: чем больше в данной местности выпадает дождей, тем выше вырастают зерновые или даже что урожай их будет тем больше. Дожди – вроде бы дарованное небесами благо. Однако очень дождливый сезон может навредить посевам или вовсе погубить урожай. Положительная корреляция сохраняется до определенной

точки, а затем быстро превращается в отрицательную. Выше такого-то количества миллиметров выпавших осадков в силу вступает обратная зависимость – чем больше дождей, тем меньше урожая вы получите.

Мы хотим обратить немного внимания на данные о ценности образования. Давайте предположим, что было доказано: выпускники университетов впоследствии зарабатывают денег больше, чем те, кто бросил учебу, и что каждый год преддипломного обучения в колледже или университете на некую сумму увеличивает последующий доход выпускника. Общий вывод будет таков: чем дольше учишься, тем больше денег в будущем заработаешь. Обратите внимание, что истинность этого вывода применительно к обучению после получения степени бакалавра не доказана и, может быть, все к нему неприменима. Многие обладатели докторской степени идут преподавать в университеты и потому не пополняют ряды состоятельных людей.

Корреляция, безусловно, указывает на тенденцию, которая не часто бывает идеальной взаимосвязью, называемой взаимно-однозначной. Рослые мальчики в среднем весят больше невысоких; таким образом, это положительная корреляция. Но не составит труда найти верзилу ростом за 180 сантиметров, чей вес будет меньше, чем у иного коротышки ростом 152 сантиметра, так что корреляция в данном случае меньше единицы. Отрицательная корреляция говорит лишь о том, что по мере увеличения одной из двух переменных вторая склонна уменьшаться. В физике это называется обратной пропорциональностью: чем сильнее вы отдаляетесь от электрической лампочки, тем меньше света падает на вашу книгу; по мере увеличения расстояния между вами и источником света интенсивность света падает. В физике взаимосвязи между явлениями, как правило, приводят к превосходной корреляции, но в других областях, начиная с бизнеса и заканчивая социологией или медициной, цифры редко когда снисходят до такой точности. Даже если образование в целом и повышает будущие доходы, оно может запросто обернуться для какого-нибудь малого полным разорением. Имейте в виду, что корреляция действительно может существовать и притом основываться на реальной причинно-следственной связи – и тем не менее не представлять почти никакой пользы, когда надо определиться с действиями в каком-нибудь конкретном случае.

Скопились кипы сплошь испещренных цифрами бумаг, призванных продемонстрировать, насколько ценно высшее образование с точки зрения будущих доходов. Отпечатаны не меньшие кипы проспектов, чтобы донести эти цифры – равно как и утверждения, в большей или меньшей степени основанные на этих цифрах, – до сведения потенциальных студентов. Нет, я не собираюсь придиরаться к этому намерению. Я и сам – горячий сторонник образования, особенно если оно включает курс элементарной статистики. Так вот, эти цифры вполне убедительно доказывают, что люди, в свое время обучавшиеся в университете, зарабатывают больше денег, чем те, кто не учился. Разумеется, из этого правила есть многочисленные исключения, но тенденция сильна и ясна.



Единственная ошибка тут такова: на основе цифр и фактов делается совершенно беспочвенный вывод. Это заблуждение по поводу причин и следствий в самом классическом виде. Оно гласит, что, как явствует из этих цифр, если вы лично посещаете университет, то, вероятно, будете зарабатывать больше денег, чем если решите посвятить эти четыре года каким-нибудь другим занятиям. В основу этого абсолютно беспочвенного вывода положено столь же беспочвенное предположение, что коль университетские выпускники зарабатывают больше денег, то это обусловлено именно тем, что они получили высшее образование. Хотя мы не можем этого знать, но вообще-то все это люди, которые и так зарабатывали бы

больше денег, даже если бы и не учились в университетах. Имеется парочка факторов, на это указывающих. В стенах университетов наблюдается непропорционально высокая численность двух категорий студентов: одаренных и богатых. Одаренные ребята благодаря своим способностям могут много зарабатывать и без университетских знаний. А что касается отпрысков богатых семейств... ну, деньги приходят к деньгам путями всем известными. Не так уж много сыновей богачей попадают в категорию лиц с низкими доходами – и не важно, учились они в университете или нет.

Нижеследующий пассаж взят из статьи, написанной в форме вопросов-ответов и опубликованной в журнале *This Week*, воскресном приложении с колоссальными тиражами. Возможно, вас позабавит (как позабавил меня) тот факт, что автор этого материала в свое время написал статью под названием «Расхожие истины: правдивы или ложны?».

Вопрос : Как влияет учеба в университете на шансы остаться несемейным человеком?

Ответ : Если вы – особа женского пола, то из-за учебы ваши шансы остаться в старых девах резко возрастают. А если вы мужчина, эффект будет прямо противоположным – учеба сводит к минимуму ваши перспективы остаться холостяком.

Корнеллский университет опросил 1500 типичных выпускников среднего возраста. Среди мужской части 93 % состояли в браке (по сравнению с 83 % мужского населения в целом).

Но среди окончивших университет женщин среднего возраста замужними были только 65 %. По сравнению с женской частью населения в целом среди выпускниц университетов старых дев оказалось в три раза больше.

Прочтет какая-нибудь семнадцатилетняя Сьюзи Браун эту статью и уяснит, что, если поступить в университет, у нее будет меньше шансов найти себе мужа. Такое ведь написали в статье, вот и соответствующие статистические данные из авторитетного источника приводятся. Так-то оно так, но только эти данные совсем не доказывают того, что утверждает автор статьи. И потом, не забудьте обратить внимание, что, хотя статистика представлена Корнеллским университетом, выводы сделаны не им. Однако невнимательный читатель вполне может заключить, что этот вывод

сделали именно специалисты известного университета.

Здесь мы снова наблюдаем тот случай, когда реально существующую корреляцию использовали, чтобы подкрепить недоказанную причинно-следственную связь. Не исключено, что выявленная корреляция действует наоборот и те незамужние женщины все равно остались бы незамужними, даже если бы и не учились в университете. Возможно, еще большее число этих женщин не смогли бы найти себе мужей. Если эти вероятности не больше, чем та, на которой настаивает автор, то ценность соответствующих выводов точно такая же – иными словами, все это не более чем догадки.

Несомненно, имеются кое-какие свидетельства в пользу той гипотезы, что предрасположенность к стародевичеству сама вполне может подтолкнуть к решению поступить в университет. Доктор Кинси, судя по всему, обнаружил некоторую корреляцию между сексуальностью и образованием, причем данная личностная особенность, по всей видимости, закрепляется еще в школьном возрасте. Сказанное только добавляет сомнительности утверждению, что учеба в университете служит помехой для замужества.

Сьюзи Браун на заметку: это не обязательно так.

Одна медицинская статья в великой тревоге указывала на рост случаев онкологических заболеваний у тех, кто пьет молоко. Представлялось так, что рак становится все более частым заболеванием в Новой Англии, штатах Миннесота и Висконсин, а также в Швейцарии, где молоко производится и потребляется в огромных количествах. А на Цейлоне, где молоко в дефиците, это заболевание редкость. В качестве дополнительного свидетельства в статье отмечалось, что в некоторых южных штатах, где потребляется меньше молока, раковых заболеваний тоже меньше. Подчеркивалось также, что регулярно употребляющие молоко женщины Новой Англии в восемнадцать раз чаще болеют некоторыми видами рака, чем японки, которые молоко пьют довольно редко.

Небольшие изыскания могли бы вскрыть не одну, а даже несколько причин, обусловивших приведенные в статье цифры. Но думаю, один фактор сам по себе достаточно ясно укажет на все эти причины. Онкологические заболевания развиваются обычно не раньше среднего возраста. Швейцария, как и упомянутые наряду с ней штаты, имеют одну общую особенность – у населения тут

относительно большая продолжительность жизни. На момент исследования женщины Новой Англии жили лет на двенадцать дольше японок.



Профессор Хелен Уокер придумала занятную иллюстрацию, чтобы изобличить всю глупость предположения, что если два явления изменяются одновременно, то одно непременно должно быть причиной, а другое – следствием. В ходе изучения взаимосвязи между возрастом и некоторыми физическими особенностями женщин начните с измерения угла, который образуют стопы ног при ходьбе. Вы установите, что у женщин постарше этот угол чаще всего больше, чем у женщин помоложе. Для начала стоит задуматься, а не указывает ли это на то, что женщины старятся оттого, что выворачивают стопы мысками наружу. Но глупость этой гипотезы вам сразу бросится в глаза. Так что выходит наоборот: это из-за возраста увеличивается угол между стопами, и большинство женщин по мере старения должны сильнее выворачивать стопы наружу.



Любое заключение подобного толка, скорее всего, должно и, безусловно, лишено оснований. Прийти к нему законным путем можно, только если вы будете изучать одних и тех же женщин – или, возможно, группы с эквивалентными характеристиками – на протяжении определенного времени. Это позволит исключить действие фактора, который и повинен в данном феномене. А суть его такова: женщины постарше воспитывались в те дни, когда юным леди прививали привычку ходить, ставя ноги носками наружу. А женщины помоложе вырабатывали походку во времена, когда это требование устарело.

Если сталкиваетесь с кем-то (как правило, это лицо заинтересованное), кто поднимает шум из-за корреляции, убедитесь прежде всего, не того ли она сорта, что образуется под влиянием самого течения событий или стародавних веяний моды. В наши дни легче легкого продемонстрировать положительную корреляцию между любой парой явлений вроде следующих: количество студентов в университетах, число пациентов в психиатрических больницах, потребление сигарет, частота сердечных заболеваний, использование рентгеновских аппаратов, производство искусственных зубов, зарплаты учителей в Калифорнии, прибыли в игорных домах Невады. Назвать один пункт из этого списка причиной другого пункта будет откровенной глупостью. И тем не менее такое происходит каждый божий день.



Позволять статистическим манипуляциям и гипнотическому мороку приведенных чисел, рассчитанных с точностью до десятых долей, наводить туман на причинно-следственные связи – немногим лучше суеверия. Но зачастую гораздо сильнее вводит в заблуждение. Это сродни наивной вере, распространенной среди жителей островов Новые Гебриды, что головные вши способствуют крепкому здоровью. Многовековые наблюдения научили их, что у людей крепкого здоровья обычно бывают вши, а вот у хворых и больных их чаще всего нет. Само по себе это наблюдение точно и здраво, какими на удивление часто бывают накапливающиеся годами бытовые наблюдения. Но этого никак не скажешь о выводе, который это первобытное племя вывело из своих наблюдений: вши укрепляют здоровье. Они должны быть у каждого.

Как мы уже заметили, свидетельства и менее скучные, чем эти (перемолотые в статистической дробилке до такой кондиции, что в них не проникнет и тень здравого смысла) помогли множеству людей озлотиться на медицине и дали пищу для бесчисленного количества статей в журналах, в том числе и специализированных. Более углубленные исследования в конечном итоге прояснили причину поверьяaborигенов с Новых Гебрид. Как оказалось, почти все члены островного племени на протяжении почти всей жизни были заражены вшами. В их представлении вшивость была, можно сказать, нормальным состоянием человека. Когда же кто-то из членов

племени заболевал лихорадкой (вполне возможно, вызванной этими самыми вшами), температура тела становилась слишком высокой для комфортного обитания вшей, и они покидали своего хозяина. Это, как видите, тот случай, когда причина и следствие исказились, поменялись местами, да еще и спутались.



Глава 9

Как производить статистикуляции

Когда публику вводят в заблуждение при помощи статистических материалов, это можно назвать статистическими манипуляциями, а если в одно слово (пускай оно и немного неуклюжее), то статистикуляциями.

Название этой книги и некоторые из приведенных в ней фактов могут создать впечатление, что подобного рода вещи есть плод чьих-то злостных намерений вас обмануть. Помнится, как-то раз президент одного из отделений Американской статистической ассоциации даже выбранил меня за это. В большинстве случаев никакие это не махинации, заявил он, а всего лишь некомпетентность. Допускаю – в его словах что-то есть¹⁸,

¹⁸ Говорят, писатель Луис Бромфилд заготовил дежурный ответ на письма от особо придирчивых корреспондентов на те случаи, когда его почта переполнялась настолько, что он был не в состоянии содержательно ответить на каждое послание. Ответ этот, хотя и не выражал согласия ни с чем и никак не поощрял к дальнейшей переписке, тем не менее удовлетворял практически каждого. Ключевая фраза звучала так: «Допускаю, что в сказанном вами что-то есть».

Это живо напоминает мне некого священника, который завоевал горячие симпатии среди молодых мамаш своей конгрегации сладкими и льстивыми речами о младенцах, которых приносили ему на крещение. Но когда

но не уверен, что первое из вышеназванных предположений звучит менее оскорбительно для статистиков, чем второе. Думается, нам важнее иметь в виду, что искажение статистических данных и манипуляции с ними ради достижения определенной цели не всегда есть дело рук профессиональных статистиков. Бывает так, что с материалом, который сходит со стола добродетельного статистика, в дальнейшем случаются всяческие перипетии – кто угодно может безбожно его извратить, утрировать, чрезмерно упростить или по собственному усмотрению переиначить его смысл. Скажем, это может быть какой-нибудь агент по продажам, специалист по связям с общественностью, журналист или автор рекламных текстов.

Впрочем, кто бы ни был повинен в том или ином случае искажения статистики, этот субъект едва ли может претендовать на звание невинного младенца, который если и напортачил, то без злого умысла. Лживые диаграммы в журналах и газетах чаще всего раздувают сенсации за счет откровенных преувеличений, а вот преуменьшения за ними почти не водится. Как подсказывает мне опыт, те, кто представляет подкрепленные статистикой доводы от лица отрасли, редко когда предлагают своим сотрудникам или клиентам условия лучше тех, что обусловлены фактами. Чаще бывает, что условия куда хуже. Видано ли, чтобы профсоюз нанял статистика настолько некомпетентного, чтобы тот умудрился представить позицию профсоюза в худшем свете, чем она есть на самом деле?

До тех пор, пока подобные статистические ошибки допускаются в пользу только одной стороны, не так уж легко списывать их только лишь на неумелость да досадные случайности.

Один из самых искусных способов извращать статистические данные связан с использованием географических карт. Карта предполагает введение изрядного числа переменных, а в них можно скрыть факты и представить в нужном свете любые соотношения. Мой любимый трофей из этой области – карта под заголовком «Мрачная тень наползает». Не так давно эту карту обнародовал Первый национальный банк Бостона, а затем ее широко

мамочки взялись сопоставить расточенные им похвалы, ни одна не смогла в точности припомнить, какие слова он произнес, а только что это было «нечто прелестное». Как позже выяснилось, его неизменный восторг выражался так: «Вот это да! (Тут следовала лучезарная улыбка.) Нет, ну вы только гляньте, это у нас младенчик!» – *Прим. авт.*

растиражировали, в том числе союзы налогоплательщиков, газеты и журнал *Newsweek*.

На карте показано, какую долю национального дохода на сегодняшний день отбирает и расходует федеральное правительство. Для этого использован прием затемнения территории штатов к западу от Миссисипи (исключая Луизиану, Арканзас и часть штата Миссури). Так оказывается, что федеральные расходы сравнялись по величине с совокупными доходами жителей этих штатов.

Обманный трюк состоит в том, что намеренно подобраны штаты, огромные по территории, но малонаселенные. Совокупные доходы жителей этих штатов относительно невелики. С такой же долей честности (или нечестности) составитель карты мог бы затенить штат Нью-Йорк или Новую Англию, и тогда тень правительства, пожирающего доходы граждан, получилась бы намного меньше, а значит, и впечатление производила бы не такое внушительное. С тем же набором данных на руках этот деятель мог бы создать совсем иное впечатление у всех, кто увидит его карту. Тем не менее никто не взял на себя труд продемонстрировать общественности эту другую карту. По крайней мере мне не известна ни одна сколько-нибудь влиятельная группа, заинтересованная в том, чтобы государственные расходы казались меньше, чем они есть.

Если бы от составителя карты требовалось всего лишь донести информацию до аудитории, это можно было бы сделать весьма простым способом. Он мог бы подобрать группу штатов среднего размера, общая территория которых соотносится с территорией страны так же, как их доходы с национальным доходом страны.

Что делает эту карту в особенности возмутительной попыткой одурачить публику, так это вот что: использованный прием – вовсе не какое-нибудь пропагандистское новшество. Перед нами нечто из области классики, иными словами, обманный прием с длиннющей бородой. Тот же Первый национальный банк Бостона много лет назад уже публиковал версии этой карты, изображавшие федеральные расходы в 1929 и 1937 гг., а вскоре эти карты можно было увидеть в классическом учебном пособии Уилларда Бrintона «Графический образ» (*Graphic Presentation*) в качестве одного из примеров отвратительной графической подачи информации. Подобный метод, как прямо и открыто говорит автор, «искажает факты». Однако банк как ни в чем не бывало продолжил в том же духе и снова

опубликовал такую карту. А *Newsweek* и те, кому полагается разбираться в подобных вещах (и кто, по-видимому, действительно в них разбирается), как и прежде, распространили эту карту, не потрудившись сопроводить ее ни предупреждениями, ни оправданиями.

Каков средний доход американских семей? Как мы отмечали выше, Бюро переписи населения США указывает, что в 1949 г. «доход средней семьи составлял \$3100». Но если прочитать газетную публикацию на тему «благотворительных пожертвований» (из тех, что распространяет Фонд Рассела Сейджа¹⁹), то можно узнать, что в том году этот показатель достигал внушительных \$5004. Очень может быть, вы порадуетесь, узнав, что люди в нашей стране живут так хорошо, но, возможно, вас до глубины души поразит, насколько сильно эта цифра не сопрягается с вашими личными наблюдениями. Впрочем, вы, должно быть, просто знакомы не с теми людьми.

¹⁹ Фонд Рассела Сейджа расположен на Манхэттене, финансирует исследования в области общественных наук и публикует их. Начинался как благотворительный фонд, был назван именем Рассела Сейджа (1816–1906), железнодорожного инвестора и политика, примечательного тем, что он «сделал себя сам», выбившись из самых низов к вершинам американского истеблишмента. Хотя сам он прослыл редким скрягой, его вдова Оливия Слокам потратила значительную часть многомиллионного наследства Сейджа на учреждение фонда его имени для осуществления разного рода благотворительных проектов. – *Прим. пер.*

МРАЧНАЯ ТЕНЬ НАПОЛЗАЕТ

Иллюстрация на примере западных штатов

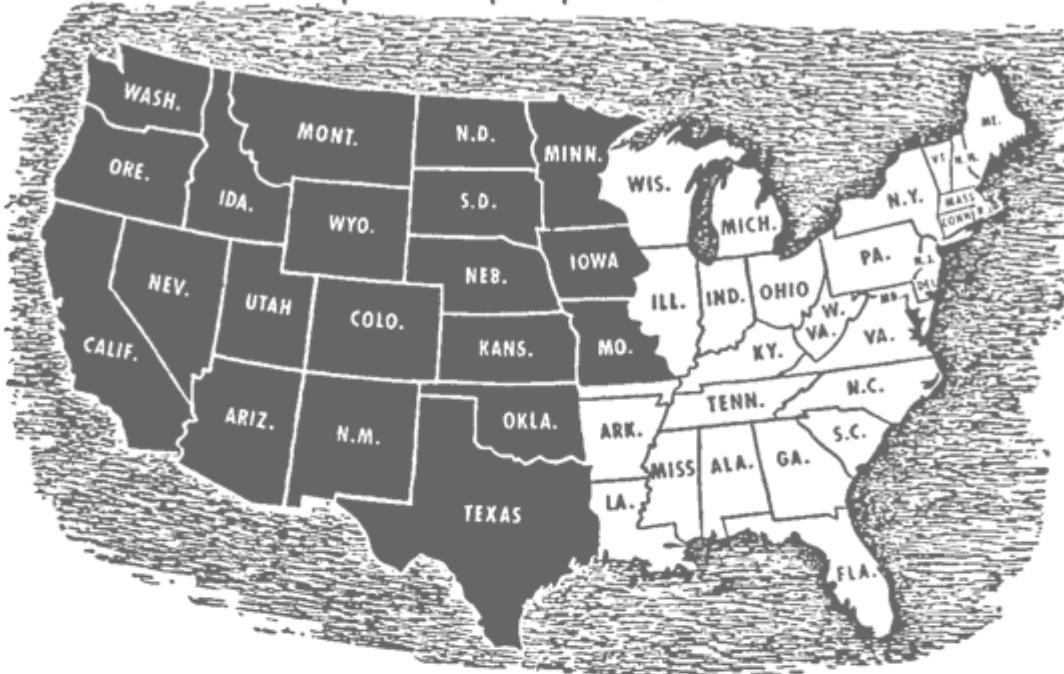


Иллюстрация на примере восточных штатов



Дабы убедить вас, что мы не жульничаем, мы обозначили также
штаты Мэриленд, Делавэр и Род-Айленд

И потом, виданное ли дело, чтобы Фонд Рассела Сейджа и Бюро переписи населения США выдали цифры, которые до такой степени расходятся? Бюро оперирует медианными показателями, и,

разумеется, так оно и должно быть, но даже если сотрудники фонда применяют средние арифметические значения, разница не должна быть настолько большой. Как выясняется, выявить такой отрадный уровень благосостояния Фонд Рассела Сейджа умудрился за счет создания того, что иначе как фиктивной семьей не назовешь. Использованный метод, как пояснили в фонде (когда специалистов все же попросили дать объяснения), состоял в том, чтобы разделить совокупный личный доход американцев на 149 000 000 с целью установить средний доход на душу населения, каковой составил \$1251. «А это, – добавили в фонде, – и превращается в доход \$5004 на семью из четырех человек».



Этот несравненный образчик статистических манипуляций преувеличивает реальное положение дел в двух направлениях. Использовано среднее арифметическое, а не меньшая по величине, но более информативная медиана (если помните, мы разбирали ее в одной из предыдущих глав). А затем строится предположение, что доход семьи прямо пропорционален количеству членов семьи. У меня на сегодняшний момент четверо ребятишек, и я очень бы желал, чтобы дела обстояли именно так, но увы. Семья из четырех человек не будет вдвое состоятельнее, чем семья из двух человек.

Будем справедливы к статистикам Фонда Рассела Сейджа. Возможно, они не пытались никого обмануть. В первую очередь им

требовалось обрисовать общее положение дел с пожертвованиями, а не с их получением. И нелепая цифра среднего дохода семьи стала всего лишь побочным продуктом работы. Хотя из-за этого она не менее успешно распространяет ложь, а потому и остается первейшим примером, показывающим, почему мы должны так мало доверять голословным заявлениям о среднем без уточнения, какой это вид среднего.

Как получить годовой доход в размере \$22 500

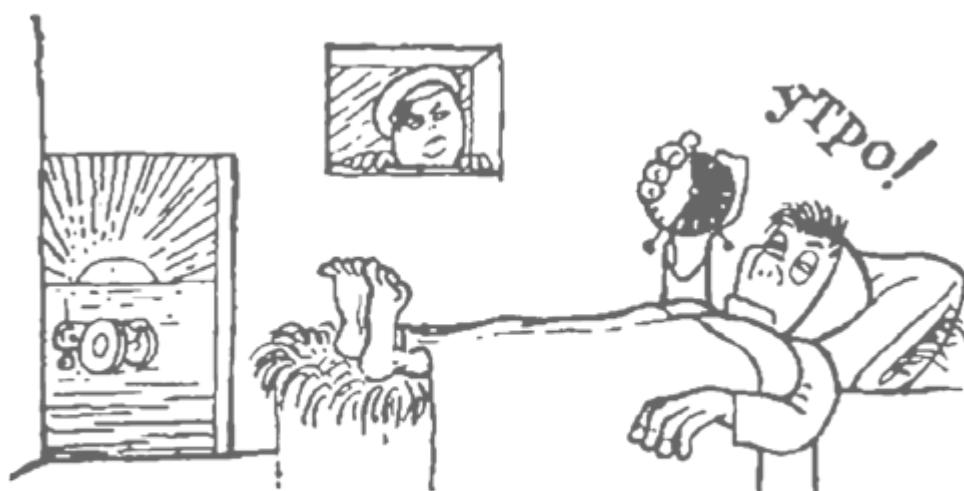
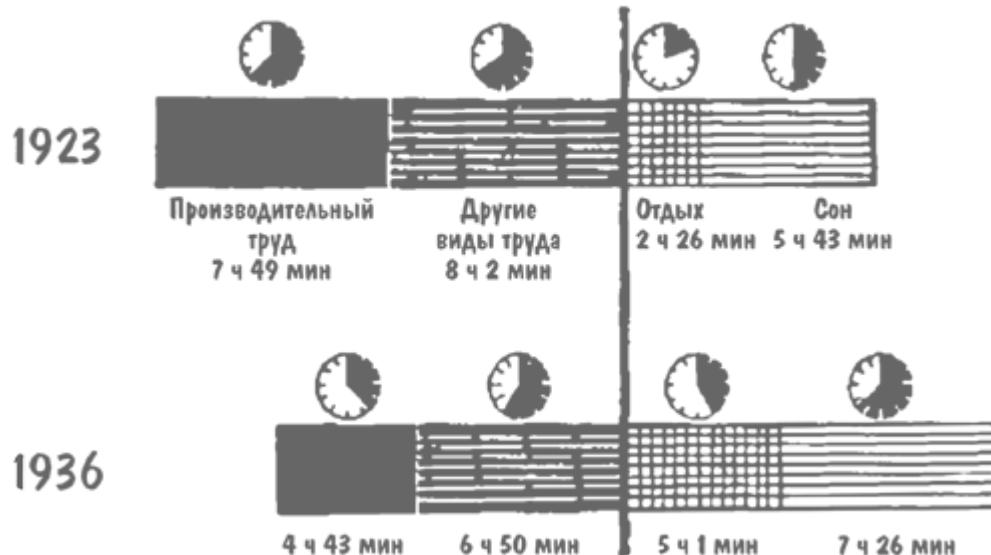
1. Возьмите как минимум 1 (одну) жену и 13 детей.
2. Рассчитайте душевой доход для населения США.
(Правильный ответ — \$1500 в год.)
3. Умножьте получившееся число на 15.
(Ответ: $15 \times \$1500 = \$22\,500$.)



Что касается обманчивого впечатления точности, способного придать вес статистическому вранью и почище, то предлагаю вам задуматься о такой штуке, как десятые доли. Спросите у сотни наших сограждан, сколько часов они проспали прошлой ночью. Предположим, что, сложив вместе названные цифры, вы получили 783,1 часа. Начнем с того, что данные такого рода далеки от точности. Большинство ответивших ошибутся в своих оценках минут на пятнадцать или больше, а кроме того, у нас нет уверенности, что ошибки в ту и другую сторону скомпенсируют друг друга. Тем более что у каждого из нас найдется знакомый, который про пять минут, проведенных без сна, скажет, что полночи проворочался в постели и глаз не сомкнул. А вы пойдите дальше и произведите арифметическое действие, а потом объявите, что в среднем продолжительность ночного сна человека составляет 7,831 часа. Это ваше заявление выглядит так, словно вы точно знаете,

о чём говорите. Если вы будете настолько глупы, чтобы заявить, что ночной сон человека составляет 7,8 часа (или «почти восемь часов»), никакого особого эффекта это не произведет. Ваши данные будут выглядеть тем, что они есть, то есть весьма приблизительными прикидками, пользы от которых не больше, чем от каких-нибудь домыслов.

ТРУД И ОТДЫХ ЖЕНЩИНЫ-КРЕСТЬЯНКИ



*Схема взята из советского источника
(Всесоюзный научно-издательский
институт изобразительной статистики)*

Немногим выше этого поднялся и Карл Маркс, когда таким же манером пытался придать своим данным вид правдоподобия и точности. Вычисление «нормы прибавочной стоимости» на примере прядильной фабрики «с 10 000 мюльных веретен» он

начинает с роскошного букета предположений, догадок и округленных цифр: мы исходим из того, что «угары составляют 6 %... хлопок [сырье] стоило окруженно 342 фунта стерлингов. Эти 10 000 веретен... стоят, будем считать, 1 фунт стерлингов на веретено... Ежегодный износ их составляет 10 %... Аренда фабричного здания – 300 фунтов стерлингов...» Далее он говорит: «Приведенные в тексте совершенно точные данные сообщены мне одним манчестерским фабрикантом». Отталкиваясь от этих приблизительных данных, Маркс производит следующие вычисления: «Таким образом, норма прибавочной стоимости = $80/52$ = $153\frac{11}{13}\%$ ». «При десятичасовом среднем рабочем дне» у Маркса получается: «необходимый труд = $33\frac{1}{3}$ часа и прибавочный труд = $62\frac{2}{3}$ часа».



Возникает восхитительное ощущение безошибочной точности этих $2/33$ часа, но все это фикция.

Проценты предлагают благодатную почву для введения в заблуждение. И подобно цифрам с неизменно впечатляющими десятыми долями, проценты способны придавать вид точности цифрам далеко не точным. В издаваемом Министерством труда США

журнале *Monthly Labor Review*²⁰ однажды написали, что на территории Вашингтона в таком-то месяце среди всех предложений по частичной занятости в домашнем хозяйстве (с оплачиваемым проездом к месту работы на общественном транспорте) в 4,9 % случаев предлагается оплата \$18 в неделю. Как позже выяснилось, указанные проценты были выведены на основе каких-то двух предложений, а в общей сложности таковых было всего-то 41. Выраженный в процентах показатель, если он рассчитан на основе ограниченного числа случаев, скорее всего, искажает реальную картину. Будет полезнее и информативнее привести сами показатели. А уж когда проценты вычисляются и потом подаются с точностью до десятых, знайте, что вы делаете шаг на пути от глупости к жульничеству.

«Купите подарки к Рождеству сегодня и сэкономите 100 %», – призывает реклама. Да, звучит как предложение, достойное самого старика Санты, а на поверку здесь просто перепутана основа, от которой эти 100 % взяты. Реально расходы сократятся на 50 %. Действительно, экономия составит 100 % от сниженной или новой цены, но в рекламном предложении говорится другое.

Аналогично этому, когда президент ассоциации цветоводов заявил в газетном интервью, что «цветы стали на 100 % дешевле, чем четыре месяца назад», он не подразумевал, что теперь цветочники раздают свой товар задаром. Но сказал-то он именно это.

Писатель и журналист Ида Тарбелл²¹ в своей книге «История нефтяной компании Standard Oil» (*History of the Standard Oil Company*) пошла еще дальше. Она пишет, что «снижение цен на юго-западе... было в диапазоне от 14 до 220 %». Будь это и в самом деле так, то предполагало бы, что продавец должен заплатить покупателю изрядную сумму, только бы он вывез прочь эту маслянистую жижу.

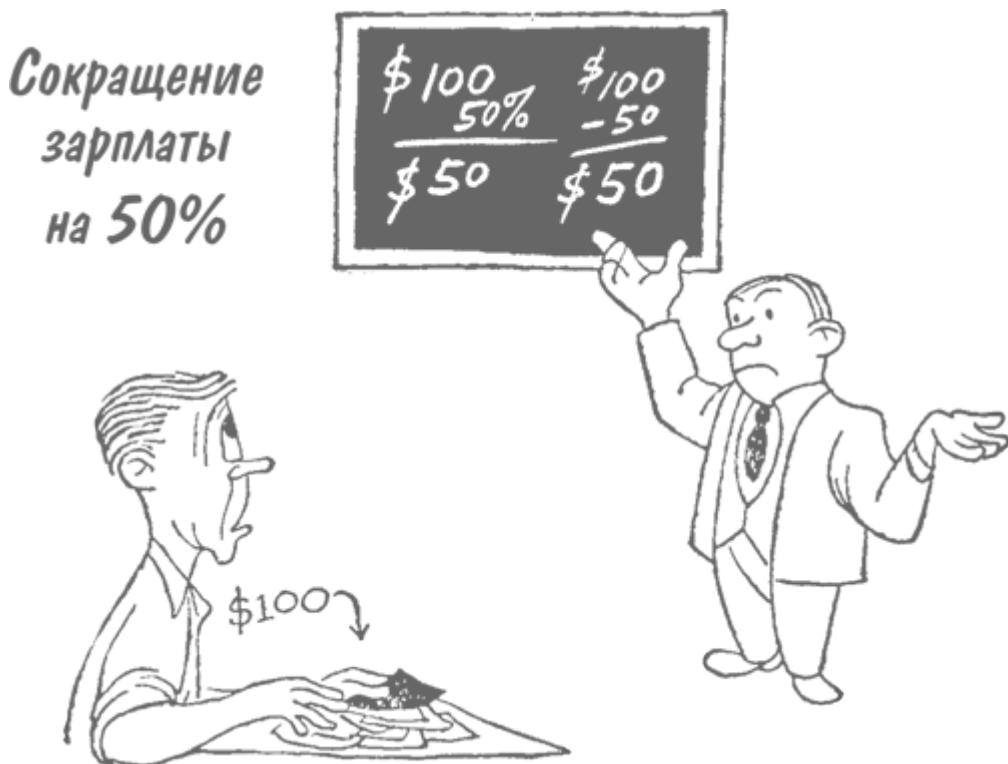
Газета *The Columbus Dispatch*²² заявляет на своих страницах, что некий промышленный продукт продавался с прибылью 3800 %,

²⁰ *Monthly Labor Review* – журнал по вопросам профсоюзного движения, анализа рынка труда и изменений в трудовом законодательстве. – *Прим. пер.*

²¹ Ида Тарбелл (1857–1944) вошла в историю Америки и американской журналистики благодаря именно этому расследованию деятельности могущественной нефтяной компании Standard Oil, которая монополизировала нефтяную индустрию штата Пенсильвания и северо-востока США. – *Прим. пер.*

²² Американская ежедневная газета, базируется в г. Колумбус, основана в 1871 г. – *Прим. пер.*

принимая за основу своих расчетов себестоимость продукта \$1,75 и его отпускную цену \$40. При расчете прибыли в процентах можно воспользоваться несколькими способами (и вы обязаны указать, какой именно применяете). Если за основу берется себестоимость, то прибыль в нашем случае составит 2185 %; а если считать исходя из отпускной цены, то прибыль будет 95,6 %. Очевидно, что газета использовала свой собственный метод, и как это водится, получила завышенную цифру, которую и опубликовала.



И даже сама *The New York Times* потерпела поражение в битве с изменчивой основой расчетов, когда поместила у себя на страницах представленную агентством Associated Press новость из Индианаполиса:

Здесь депрессия получила мощный удар в челюсть. Водопроводчики, штукатуры, плотники, маляры и все прочие, кого объединяет Индианаполисский профсоюз работников строительной промышленности, получили пятипроцентное повышение заработной платы. Тем самым им на четверть компенсировали введенное прошлой зимой сокращение зарплаты на 20 %.

*Возмещение
сокращенной
зарплаты
на 50%*

\$50	$\frac{50\%}{\$25}$	\$50	$\frac{25}{\$75}$
------	---------------------	------	-------------------



Казалось бы, все в этом сообщении выглядит разумно – не придерешься. Но если процент сокращения рассчитывался на основе первоначального размера зарплаты, то процент роста рассчитали на основе меньшего показателя – уровня заработной платы, каким он стал после сокращения.

Можете сами проверить, как получился этот маленький статистический промах, а для простоты эксперимента примите, что первоначальная зарплата строительных рабочих составляла \$1 в час. Урезанная на 20 %, она составит 80 центов. При увеличении этой зарплаты на 5 % прирост будет равен 4 центам. А по отношению к первоначальному сокращению на 20 % это будет не четвертая, а пятая часть. Как и множество добросовестных ошибок, которые предположительно допущены без всякого умысла, данная позволила преувеличить эффект, а значит, представить ситуацию в более выгодном свете.

Отсюда вы видите, почему для того, чтобы компенсировать сокращение зарплаты на 50 %, вам должны увеличить ее на 100 %.

И все та же *Times* как-то сообщила, что за финансовый год авиапочта «утратила в результате возгораний отправления общим весом в 2,206 тонн, что составляет ничтожную долю в 0,00063 %». В заметке сообщалось, что за год самолетами было доставлено 3499,801 тонн почтовых отправлений. Страховая компания, взьмись она рассчитывать ставки в таком же духе, ограбила бы целую кучу неприятностей. Попробуйте посчитать сами, и у вас получится

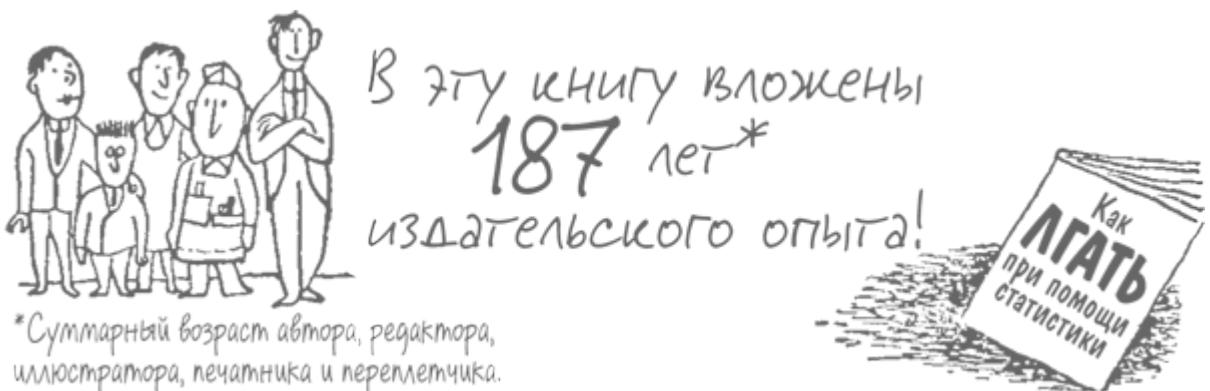
0,063 %, то есть в сто раз больше, чем написала газета.

Именно обманчивость, которую создает изменчивая основа процентных расчетов, позволяет ловчить со скидками. Когда оптовый торговец скобяными товарами предлагает вам «Скидку 50 % и еще 20 % от прейскурантной цены», он не имеет в виду, что даст вам скидку 70 %. В действительности скидка составит 60 %, поскольку 20 % берутся от меньшей цены, а именно от той, что получилась после того, как первоначальная цена уменьшилась на 50 %.

Изрядная доля несуразицы и надувательств проистекает от манеры соединять вещи несовместимые, которые только на первый взгляд кажутся таковыми. Уже не одно поколение школьников применяет разновидность этого приема, чтобы доказать, что на самом деле они вовсе и не ходят в школу.

Полагаю, и вам припоминается этот нехитрый трюк. Итак, всего в году 365 дней, из этого можно вычесть 122 дня, поскольку треть жизни мы посвящаем ночному сну. Затем вычтем еще 45 суток, ведь каждый день мы по три часа тратим на еду. Из оставшихся 198 дней вычитаем 90 – это летние каникулы, а кроме того, надо вычесть еще 21 день, ибо столько составляют праздничные дни на Рождество и Пасху. А того, что останется, уже не хватит даже на все субботние и воскресные дни.

«Прием слишком древний и очевидный, чтобы его использовали в серьезных делах», – скажете вы. Однако, как уверяет Объединенный профсоюз рабочих автомобильной промышленности Америки на страницах своего ежемесячного издания *Ammunition*, против них применяется именно такой трюк.



*Суммарный возраст автора, редактора, иллюстратора, печатника и переплетчика.

В эту книгу вложены
187 лет*
издательского опыта!



Во время каждой забастовки высказывает ужасающая, грубая и бессовестная ложь. Всякий раз, когда проводится забастовка, Торговая палата заявляет, что каждый день простоя обходится в миллионы долларов.

Эти убытки высчитываются путем сложения количества всех автомобилей, которые были бы выпущены, если бы забастовщики полностью отрабатывали свое рабочее время. Затем таким же манером подсчитываются убытки поставщиков. В ход идет все возможное, в том числе плата за проезд в городском транспорте и потери автоторговцев на продажах.

Аналогичная этой и столь же дикая идея, будто проценты можно складывать с такой же легкостью, как яблоки в корзинку, применяется и против авторов. Взгляните только, как убедительно выглядит пример подобного трюка на страницах еженедельного книжного обозрения *The New York Times Book Review*:

Разрыв между растущими ценами на книги и авторскими заработками, как выясняется, обусловлен существенным ростом производственных и материальных затрат. Вот по статьям: одни только редакционные и производственные затраты за последние десять лет ощутимо выросли – на 10–12 %; расходы на материалы возросли на 6–9 %, торговые издержки и расходы на рекламу поднялись на 10 %. В общей сложности рост затрат достиг как минимум 33 % (для крупного издательского дома) и почти 40 % для ряда издательств меньшего размера.

Но если каждая из статей, из которых складывается себестоимость данной книги, увеличится примерно на 10 %, общие затраты на книгу должны, по идее, возрасти примерно на те же 10 %. А логика, допускающая суммирование процентов роста отдельных статей, способна привести еще и не к такому полету фантазии. Купите двадцать разных товаров, и вы увидите, что каждый за год подорожал на 5 %. В сумме это «образует» стопроцентное увеличение, вот вам и удвоение стоимости жизни. Абсурд!



Все это несколько смахивает на байку про хозяина придорожной харчевни, которого попросили объяснить, как ему удается так дешево продавать сэндвичи с крольчатиной. «Очень просто, – отвечал тот. – Мне приходится добавлять некоторое количество конины. Но я смешиваю их поровну: одна лошадь и один кролик».

Профсоюзное издание поместило у себя карикатуру в знак протеста против еще одной разновидности необоснованного сложения несовместимых объектов. Карикатура изображает, как босс суммирует ставку за один час работы \$1,5 со ставкой за сверхурочную работу \$2,25 и двойной ставкой \$3,0 и таким манером выводит «среднюю» почасовую ставку \$2,25. Не думаю, что найдется много примеров подобного рода, где средний показатель настолько бессмысленен.

Благодатной почвой для обдуривания становится и путаница между процентами и процентными пунктами. Если вашим прибылям надлежит вырасти с 3 % на вложенные инвестиции в одном году до 6 % в следующем, то вы можете создать впечатление скромного увеличения прибыли, указав, что ее рост составил три процентных пункта. Точно так же, не погрешив против истины, вы могли бы описать это как стопроцентный рост. Если хотите найти еще примеры такого вольного обращения с двумя этими показателями, советую особенно присмотреться к тем, кто проводит опросы общественного мнения.

Настолько же обманчивыми могут быть и процентили. Когда вам сообщают, как выглядит успеваемость Джонни по алгебре или еще какому-нибудь предмету на фоне успехов его однокашников, то соответствующий показатель может оказаться процентилем. Это означает ранг Джонни среди каждой сотни учащихся. Например, если на курсе три сотни студентов, то у троих лучших по успеваемости процентиль будет 99; у следующих троих – 98 и т. д. У процентилей имеется одна странность: студент, имеющий процентиль 99, вероятно, успевает несколько лучше, чем тот, у кого процентиль 90, тогда как обладатели процентилей 40 и 60 могут на деле иметь почти одинаковые баллы. Это объясняется свойством, присущим множеству разнообразных характеристик, – образовывать кластеры вокруг среднего значения, что и формирует колоколообразную кривую «нормального» распределения, уже упоминавшуюся в одной из предыдущих глав.

Временами состязание между статистиками выходит наружу, и тогда даже самые неискушенные наблюдатели не могут не заподозрить неладное. Честные граждане всегда выигрывают, когда между статистиками разгорается свара. Так, Комитет по делам сталелитейной промышленности²³ указал на мошеннические уловки, которыми злоупотребляли как компании, так и профсоюзы. Дабы продемонстрировать, как прекрасно шли дела в отрасли в 1948 г. (в подтверждение, что компании могли бы позволить себе повысить зарплаты рабочим), профсоюз сравнивал производительность в указанном году с уровнем 1939 г. – когда объем производства в отрасли упал до особенно низкого уровня. Сталелитейные компании, не желая отстать в этой гонке вранья, настаивали, что основой для сравнений должны служить полученные рабочими деньги, а не средний показатель почасового заработка. Здесь хитрость крылась в том, что в предшествующем году огромное число рабочих работали неполный день, так что в текущем году, когда их перевели на полный рабочий день, их заработки все равно бы выросли и без какого-либо повышения ставок.

Журнал *Time*, примечательный тем, что неизменно безупречен в публикуемых графических данных, поместил у себя диаграмму, которая представляет собой забавный образец того, что статистика

²³ Комитет, созданный президентом Трумэном в конце 1940-х гг. для скорейшего разрешения обострившихся в тот период трудовых конфликтов в сталелитейной отрасли. – Прим. пер.

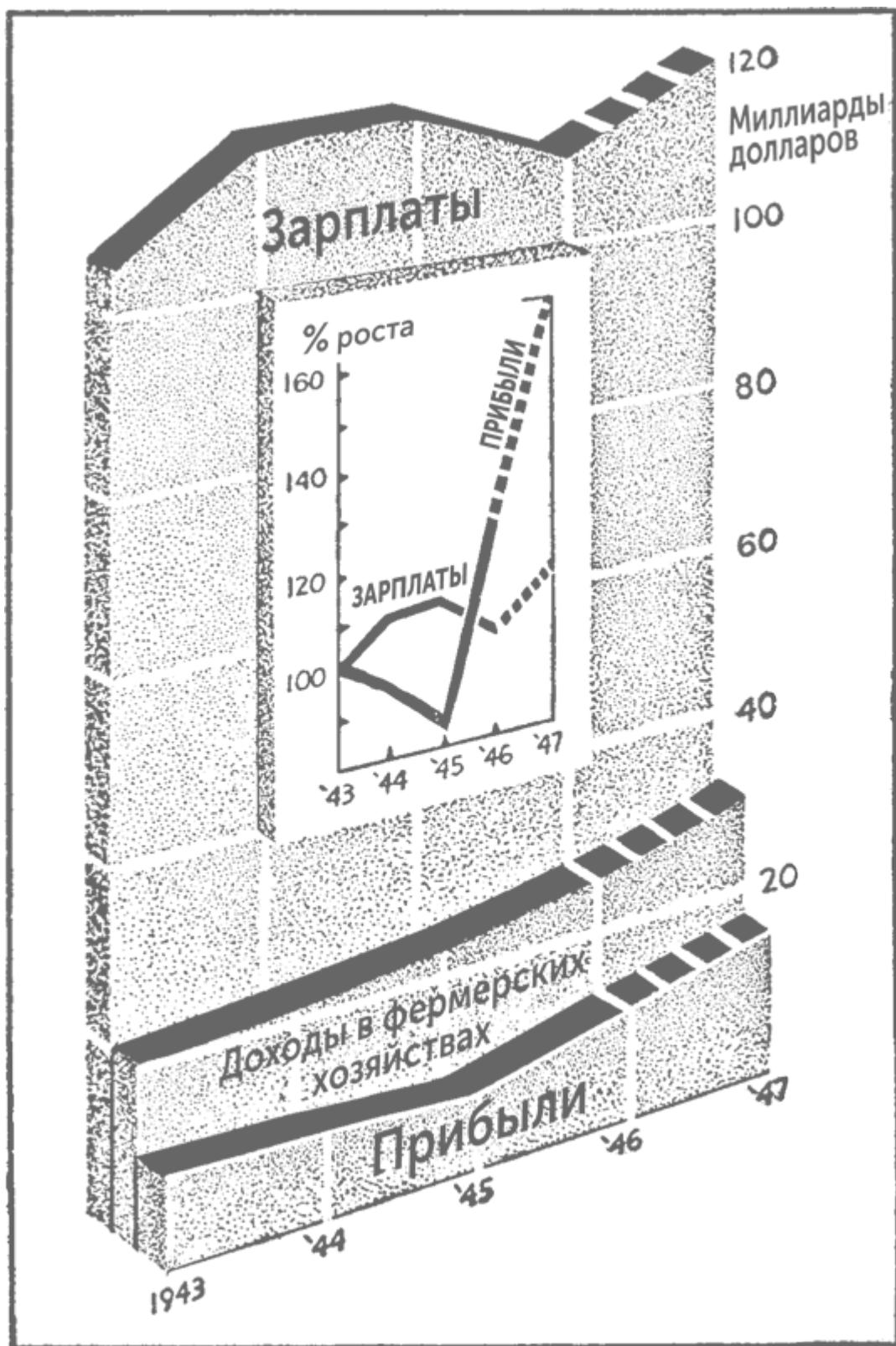
готова вынуть из шляпы любые данные, какие ей только понадобятся. Оказавшись перед выбором между двумя равноценными методами (один из них был выгоден руководству сталелитейных компаний, а другой – профсоюзам), *Time*, недолго думая, применил оба. Опубликованная журналом диаграмма состояла из двух – одна была наложена на другую. Притом обе были построены на основе одного и того же набора данных.

Одна диаграмма изображала зарплаты и прибыли в миллиардах долларов. Как из нее явствовало, и то, и другое росло и более или менее на одну и те же величину. Только вот на зарплаты приходилось раз в шесть больше долларов, чем на прибыли. Как видно из диаграммы, это зарплаты рабочих, оказывается, создали огромное инфляционное давление.

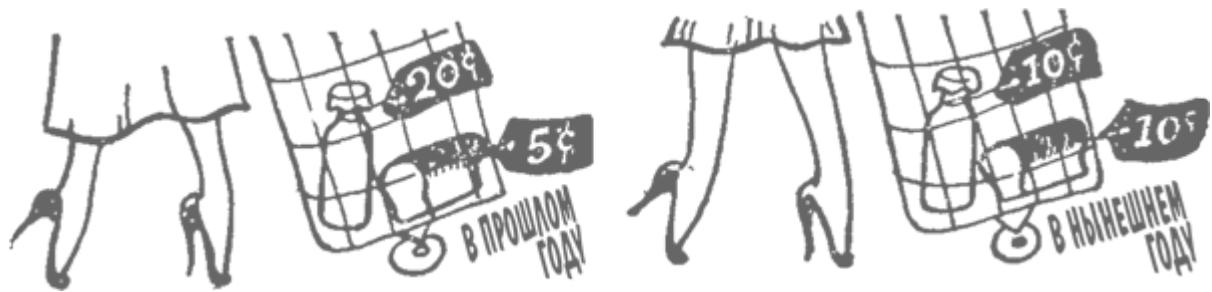
Вторая часть этой двойственной диаграммы отображала те же самые изменения, но в процентном увеличении. Кривая роста заработных плат демонстрировала относительно плавную динамику, тогда как кривая прибылей резко устремлялась вверх. Из этого можно было заключить, что инфляцию порождают главным образом прибыли.

Вам предлагалось самим выбрать, какой вывод следует из этой диаграммы. Или, что, вероятно, было бы еще лучше, вы могли заметить, что ни один из факторов нельзя выделить в качестве виновника тяжелого положения. Иногда очень важно просто показать, что вопрос, вокруг которого разгорелись споры, не такой уж простой, каким его стараются выставить.

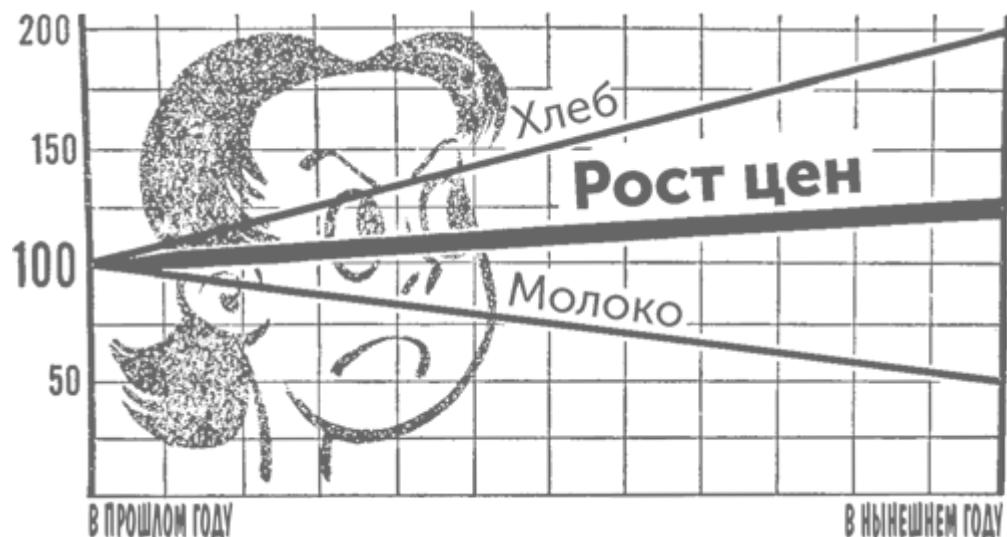
Индексы приобрели жизненно важное значение для многих людей, поскольку к ним теперь нередко привязывают ставки заработной платы. Вероятно, стоит упомянуть, какими способами можно заставить индексы плясать под чью угодно дудку.



Воспроизведется с любезного разрешения журнала Time
как пример диаграммы, которая не обманывает.



Возьмем самый упрощенный пример и предположим, что в прошлом году молоко стоило 20 центов за кварту, а хлеб – 5 центов за буханку. В этом году цена на молоко снизилась на 10 центов, а на хлеб – на столько же поднялась. Теперь решайте, какой тезис хотите доказать. Что стоимость жизни растет? Или падает? Или что никаких изменений не произошло?



Возьмем прошлый год за исходный период и, следовательно, прошлогодние цены примем за 100 %. Цена на молоко вполовину снизилась (на 50 %), а цена на хлеб удвоилась (рост на 200 %). Среднее от 50 и 200 равняется 125. Это означает, что в среднем цены выросли на 25 %.

Теперь попробуем еще раз, но за исходный период возьмем нынешний год. Молоко раньше стоило 200 % по сравнению с ценой текущего года, а хлеб продавался за 50 % его нынешней цены. Среднее по-прежнему составляет 125 %. Вывод: цены в прошлом году были на 25 % выше, чем сейчас.



А чтобы доказать, что уровень цен вообще не изменился, мы просто от среднего арифметического перейдем на среднее геометрическое. Это несколько иная величина, чем простое среднее, то есть среднее арифметическое, которым мы оперировали в двух предыдущих случаях. Среднее геометрическое – совершенно законная величина, а в некоторых случаях она наиболее полезна и хорошо раскрывает суть вещей. Чтобы вычислить среднее геометрическое трех чисел, их следует перемножить и из произведения извлечь кубический корень. Если чисел четыре, из их произведения извлекается корень четвертой степени, если два – то квадратный корень. В общем, вы поняли.

Берете за исходный период прошлый год, а уровень прошлогодних цен – за 100 %. Тогда фактически вы перемножаете 100 на 100 и извлекаете квадратный корень, который тоже равен 100. Теперь рассмотрим текущий год: цена на молоко составляет 50 % от прошлогодней, а на хлеб – 200 % от прошлогодней. Перемножаете 50 и 200, получается 10 000. Квадратный корень, который и будет средним геометрическим в данном случае, равен 100. Вывод: по сравнению с прошлым годом цены не продемонстрировали *ни* роста, *ни* снижения.

Суть в том, что, несмотря на свою математическую основу, статистика это не только наука, но и искусство. Можно проделать великое множество манипуляций и даже что-нибудь извратить, не выходя при этом за рамки приличий. Статистик зачастую вынужден сам выбирать, какой метод подачи данных ему использовать, что само по себе – процесс субъективный. В условиях коммерческой практики статистик настолько же не склонен выбирать

неблагоприятный метод подачи данных, насколько автор рекламного текста не склонен описывать товар как «хлипкий и дешевый», когда он может назвать его «не оттягивающим руку и предназначенным бережливым хозяйствам».



Даже ученый муж и тот в своих исследовательских изысканиях может проявить (возможно, и неосознанно) пристрастность в пользу тезиса, который намерен обосновать, или отрезать нарушающую его логические построения часть оси.

Все это подсказывает нам, что любые статистические данные, факты и цифры, которые приводятся в газетах и книгах, журналах и рекламных объявлениях, следует принимать к сведению только после того, как вы вдумчиво и основательно их изучите. Порой осторожная недоверчивость только помогает сфокусироваться на сути. Но отвергать статистические методы просто так, без достаточных к тому оснований, не особенно умно. Это все равно как совсем не читать книги под предлогом, что писатели при помощи слов иногда утаивают факты вместо того, чтобы обнажать их. В конце концов, был же совсем недавно случай, когда один политик на выборах во Флориде заработал порядочный капитал, обвинив

своего соперника в том, что тот «практиковал целибат»²⁴. А вспомните, как нью-йоркский кинотеатр, готовясь к прокату фильма «Камо грядеши» (Quo Vadis)²⁵, повсюду афишировал его, громадными буквами сообщая всем, что фильм «псевдоисторический». И не рекламировали ли производители патентованного лекарственного средства Crazy Water Crystals²⁶ свой продукт как «обеспечивающий немедленное мимолетное облегчение»?



²⁴ В 1950 г. во Флориде проходили первичные выборы кандидата от демократической партии в сенат. Слух о целибате Клода Пеппера запустил, как считается, его соперник Джордж Сматерс, хотя сам он позже это отрицал. Таким «обвинением» Сматерс намеревался выставить соперника в смешном свете перед сельскими избирателями, в чем и преуспел. Это стоило Пепперу места в сенате и поставило крест на его президентских амбициях. – *Прим. пер.*

²⁵ Имеется в виду очередная экранизация романа Генрика Сенкевича «Камо грядеши», выпущенная в 1951 г. с Ричардом Бартоном в одной из главных ролей. – *Прим. пер.*

²⁶ Порошок, ложка которого, растворенная в стакане воды из-под крана, придает ей, как уверяли производители, качества природной минеральной воды. В основу идеи положена огромная популярность, которую в начале XX века приобрел минеральный источник в городке Минерал-Уэллс в Техасе. – *Прим. пер.*



Глава 10

Как поставить статистика на место

До сих пор я обращался к вам так, словно вы какой-нибудь пират только и мечтающий, чтобы его подковали по части тонкостей владения абордажной саблей. В заключительной главе я позволю себе отойти от этого литературного приема. Я готов приступить к важной миссии, которая, как мне хочется думать, таится под обложкой этой книги. Пришло время растолковать вам, как смело глядеть в глаза фиктивной статистике и изобличать ее и, что не менее важно, как распознавать добросовестные и полезные данные в непролазных чащобах обмана и фальсификаций, описанию которых главным образом и посвящены предыдущие главы.

Не все статистические данные можно проверить с той же степенью надежности, какую гарантирует химический анализ или какое-то другое действие, осуществляемое в стенах лаборатории. Но что мешает вам прощупать подозрительные данные с помощью пяти простеньких вопросов? Ответив на них, вы оградите себя от невероятной массы сведений, которые не содержат и крупицы правды.

Кто это говорит?

Первое, на что следует обратить внимание, – это предвзятость статданных. Вдруг они исходят от научно-исследовательской

лаборатории, которой требуется подтвердить какую-нибудь теорию или сохранить репутацию? А еще ей просто могли за это заплатить. Или сведения представила газета, чья цель – интересная статья. А может, источником стал профсоюз или руководство компании и на кону стоит размер заработной платы персонала.

Присмотритесь: нет ли сознательного искажения сведений? Один из приемов – откровенно лживое утверждение; другой – когда утверждение сформулировано туманно, но не хуже вводит в заблуждение, а уличить автора в обмане не представляется возможным. Это может быть и подбор благоприятных фактов при одновременном скрытии неблагоприятных. Иногда намеренно подменяют точку отсчета. Такое практикуется, когда для одного сравнения за основу берется какой-то один год, а для другого – другой год, более подходящий. Бывают случаи, когда специально выбирают ненадлежащий статистический показатель: например, среднее арифметическое там, где медиана была бы более показательной (и, вероятно, чересчур показательной). Этот трюк маскируют, называя приведенный показатель «средним» – без уточнения, какой это вид среднего.

Присматривайтесь внимательно, возможна ли непреднамеренная предвзятость данных. Зачастую это куда опаснее. Именно такого рода предвзятость в 1928 г. сыграла злую шутку со многими статистиками и экономистами, когда, пробравшись в их графики и диаграммы, «помогла» им доказать вещи совершенно невероятные. Дефекты в структуре экономики они на радостях прошляпили, зато привели самые разнообразные свидетельства, причем подкрепленные статистическими выкладками, дабы продемонстрировать, что страна вступила в полосу процветания²⁷.

Иногда достаточно если не первого, то хотя бы второго взгляда, чтобы определить, кто это говорит. Этот кто-то может скрываться за тем, что Стивен Поттер²⁸ назвал бы «какими-надо-именами». Нечто хотя бы намекающее на причастность к медицинской профессии – это «какое-надо-имя». Научно-исследовательские

²⁷ А в 1929 г. началась Великая депрессия, которая продолжалась целых десять лет. – *Прим. ред.*

²⁸ Стивен Поттер (1900–1969) – английский писатель, автор книг-пародий на руководства из разряда «помоги себе сам», где с юмором поучает читателя, как надо жить и действовать в разных жизненных сферах, чтобы всегда быть выше других, не прибегая к низостям. Слово «lifemanship» впервые употреблено в книге Some Notes on Lifemanship (1950), в вольном переводе – «Некоторые заметки по поводу умения жить, все время оставаясь на высоте». Многие книги Поттера экранизированы. – *Прим. пер.*

лаборатории тоже имеют «какие-надо-имена». Таковыми обладают высшие учебные заведения, особенно университеты, а еще в большей мере те, что известны своим техническими разработками. Автор статьи, упоминавшейся несколько глав назад, который доказывал, что высшее образование подрывает шансы девушки выйти замуж, ловко воспользовался «каким-надо-именем» Корнеллского университета. Еще раз обращаю ваше внимание, что хотя статистические данные предоставлены Корнеллским университетом, выводы из них целиком и полностью на совести автора статьи. Но благодаря упоминанию «какого-надо-имени» может сложиться неверное впечатление, будто это «Корнеллский университет утверждает, что...».



Когда упоминается «какое-надо-имя», удостоверьтесь, что авторитет его обладателя действительно стоит за данной информацией, а не просто приплетается ради пущей убедительности.

В чикагском *Journal of Commerce* вам могло попасться на глаза некое горделивое заявление. Журнал провел опрос на тему манипуляций с ценами и затоваривания. Из 169 корпораций, согласившихся принять в нем участие, две трети отрапортовали, что из своего кармана компенсируют увеличение цен, вызванное

войной в Корее. «Опрос показывает [держите ухо востро, когда бы и где бы вам ни встретились эти слова!], – утверждал *Journal of Commerce*, – что корпорации сделали прямо противоположное тому, на что так рассчитывали враги американского бизнеса». Вот тут самое время задаться вопросом «Кто это говорит?», поскольку *Journal of Commerce* вполне можно счесть заинтересованной стороной. Это еще и самое подходящее время задать себе второй вопрос:

Откуда ему это известно?

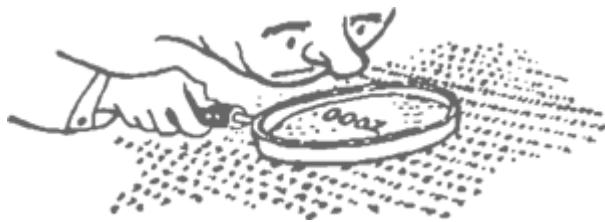
В нашем случае выяснилось, что журнал начал с того, что разослал вопросник в 1200 крупных компаний. Из них на вопросы ответили лишь 14 %. Остальные 86 % не удосужились сделать вообще никаких публичных заявлений относительно того, манипулируют ли они ценами и придерживают ли свою продукцию.

Таким образом, журнал умудрился сделать поразительно хорошую мину при довольно плохой игре, но сам факт от этого не изменился: по большому счету хвастаться было особо нечем. Все успехи сводились к следующему: из 1200 привлеченных к опросу компаний 9 % сообщили, что не поднимали цены, 5 % – что подняли, а 86 % промолчали. Ответившие компании образуют выборку, которую можно заподозрить в необъективности.

Приглядывайтесь, нет ли свидетельств тому, что выборка смещенная, то есть отобрана ненадлежащим образом или, как в вышеупомянутом случае, сформировалась сама собой. Задайте себе вопрос, который мы обсуждали в одной из первых глав: достаточно ли велика выборка, чтобы на ее основе сделать сколько-нибудь надежный вывод?

Аналогичным способом поступайте в случае, когда сообщается о корреляции: достаточно ли она велика, чтобы что-то означать? Достаточно ли случаев рассмотрено, чтобы выявленная корреляционная зависимость имела хоть какую-то значимость? Как неподготовленный читатель, вы лишены возможности применить какие-либо критерии значимости или составить себе однозначное суждение о степени адекватности выборки. Но что касается множества публикуемых данных, вы способны оценить с первого взгляда – возможно, довольно долгого и пристального, – что количество рассмотренных случаев явно недостаточно, чтобы

убедить в чем-либо человека думающего.



Чего не хватает?



Далеко не всегда сообщают, сколько случаев было взято для изучения. Отсутствия такой цифры достаточно, чтобы бросить тень подозрения на все сообщение в целом, особенно если оно исходит от заинтересованного источника. Точно так же не следует принимать слишком серьезно информацию о корреляции между двумя величинами, если не указана степень достоверности этого показателя (вероятная ошибка, стандартная ошибка).

Будьте начеку, если вам называют среднее без уточнения его вида, во всех случаях, когда можно заподозрить, что среднее арифметическое и медиана существенно различаются.

Множество цифр утрачивают всякий смысл из-за неправомерных сравнений. Так, статья в журнале *Look*, касаясь темы синдрома Дауна, сообщает, что «как показало одно исследование, в 2800 случаях более половины матерей были 35 лет или старше». Чтобы данная информация имела для вас хоть какой-то смысл, вы должны

иметь общее представление, в каком возрастном диапазоне женщины в массе своей рожают детей. А такими знаниями мало кто из нас владеет.

Ниже приводится выдержка из рубрики «Нам сообщают из Лондона» в журнале *New Yorker* от 31 января 1953 г.

Министерство здравоохранения недавно обнародовало данные, что за неделю, когда держался густой туман, уровень смертности в Лондоне подскочил на 28 %. Эти сведения глубоко потрясли общественность, привыкшую считать последствия неприветливого британского климата скорее досаждающими, нежели губительными... Небывалая смертельная опасность...

Но насколько смертоносной на самом деле *была* эта неделя? Считать ли исключительным случаем, что уровень смертности настолько превысил обычный показатель? Такого рода вещам вообще свойственно варьироваться. А как насчет последующих недель? Упал ли уровень смертности ниже среднего уровня, указывая, что если туман и убивал, то по большей части людей нездоровых, которых болезни в любом случае вскоре свели бы в могилу? Сама приведенная цифра выглядит весьма внушительно, но отсутствие других цифр, которые позволили бы оценить ее значение, сводит почти на нет весь смысл сообщения.

Бывает и так, что в источнике приводятся проценты, а стоящие за ними исходные цифры отсутствуют, и это тоже способно ввести в заблуждение. Давным-давно, когда Университет Джонса Хопкинса только начал принимать девушек, некто, не испытывавший особых восторгов по поводу совместного обучения, обнародовал данные, ставшие для многих потрясением: оказывается, 33 1/3 % студенток университета выходили замуж за преподавателей! Однако исходные цифры позволяли точнее оценить картину «бедствия». На тот момент в списке учащихся числились три девушки-студентки, и одна из них действительно вышла замуж за преподавателя.

Пару лет назад Торговая палата Бостона взялась составить свой список «Самых успешных американок года». Про шестнадцать из них, также вошедших в справочник «Кто есть кто», сообщалось, что в общей сложности на них приходится «60 ученых степеней и 18 детей». Эти сведения выглядят вроде бы достаточно информативно, чтобы вы могли составить себе представление о самых успешных американках, но лишь до тех пор, пока не становится известно,

что в составе этой группы женщин числятся декан Вирджиния Гилдерслив²⁹ и миссис Лилиан Гилбрет³⁰. На этих двух дам приходится добная треть упомянутых шестидесяти ученых степеней. А миссис Гилбрет, как известно, воспитала две трети от общего числа упомянутых детей.

Некая корпорация с полным на то основанием объявила, что держатели ее акций – 3003 человека и что в среднем у каждого акционера во владении находится по 660 акций. И это было правдой. Как был правдой и тот факт, что три четверти общего количества акций, которых насчитывалось два миллиона, принадлежали трем лицам, а оставшуюся четверть делили между собой остальные 3000 акционеров.

Если вам подсовывают индекс, вы вправе спросить, какие сведения относительно этого индекса вам недодали. Это может быть основа расчетов индекса, основа, выбранная так, чтобы представить реальную картину в превратном виде. Одна общенациональная организация рабочих как-то раз продемонстрировала, что после Депрессии индексы прибылей и производства росли опережающими темпами по сравнению с индексом заработной платы. Эта иллюстрация, предназначенная служить аргументом в пользу повышения заработной платы рабочим, утратила всю свою убедительность, когда нашелся некто, кто раскопал недостающие данные. Тогда стало ясно, что прибыли были почти «обречены» на более быстрый рост в процентах, чем заработка, по той тривиальной причине, что исходно находились на более низкой отметке, что и обусловило меньшую основу для расчета процентного роста.

²⁹ Вирджиния Гилдерслив (1877–1965) – широко известная в свое время женщина, обладательница звания академика, активный общественный и политический деятель, поборница равноправия женщин и многолетний декан Барнард-колледжа, а кроме того – единственная женщина, назначенная президентом Рузвельтом в состав делегации США на Сан-Францисской конференции 1945 г., в ходе которой был подписан устав ООН. – *Прим. пер.*

³⁰ Лилиан Гилбрет (1878–1972) – американский ученый, консультант по менеджменту, производственный психолог, эксперт по кадровым ресурсам предприятий и эргономике. Вместе с супругом Фрэнком Гилбретом стала основоположником науки о движениях, автором ряда методов научной организации труда. У четы Гилбрет было двенадцать детей. – *Прим. пер.*



В некоторых случаях не называют фактор, который и спровоцировал перемены. Подобное замалчивание позволяет создать впечатление, что перемены обусловил другой фактор, более желательный для целей тех, кто эти данные обнародует. В каком-то году были опубликованы цифры, призванные продемонстрировать, что дела у данного бизнеса идут в гору: особо подчеркивалось, что в апреле объем розничных продаж превысил прошлогодний. Что «позабыли» упомянуть авторы сообщения, так это что в прошлом году Пасха приходилась на март, а в рассматриваемом году была в апреле.



Сообщение о том, что за последнюю четверть века наблюдается резкий рост смертности от раковых заболеваний, создаст у вас ложное впечатление о положении дел, только если вам неизвестно, что это следствие целого ряда факторов. В тех случаях, когда сегодня констатируют смерть от рака, раньше применили бы формулу «смерть от неизвестных причин»; аутопсию в наши времена делают чаще, чем раньше, что позволяет более точно поставить диагноз; учет, регистрация и обобщение медицинских статданных на сегодняшний день осуществляются более полно; и наконец, сегодня люди чаще доживают до возраста, в котором шансы заболеть раком увеличиваются. И если вы изучаете абсолютные показатели смертности, а не просто уровень смертности, не пренебрегайте тем фактом, что в наши дни численность населения больше,

чем в прежние времена.

Не подменен ли объект исследования?

Когда изучаете статистические показатели, особенно внимательно следите, не произошло ли подмены в процессе перехода от исходных данных к выводам. Прискорбно часто бывает, что исследуют одно, а, сообщая результаты, называют это другим.

Как уже отмечалось, рост зарегистрированных случаев заболевания – не всегда то же самое, что рост самих случаев заболевания. Если кандидат выходит в победители по данным «соломенных опросов»³¹, это не всегда то же самое, что результат самих выборов. Если «типичная представительная группа» аудитории издания назвала в качестве предпочтительных статьи на международные темы, это еще не доказывает, что эти люди непременно будут читать подобные статьи.

Количество случаев заражения энцефалитом в Калифорнийской долине за 1952 г. оказалось втрое выше, чем в самом неблагоприятном году из предыдущих. Многие встревоженные жители поспешили отправить своих детей куда-нибудь подальше. Но когда все данные были сведены воедино, они не показали сколько-нибудь существенного роста смертности от летаргического энцефалита. Причиной тревожной статистики послужило следующее: значительные силы сотрудников органов здравоохранения штата и федерального ведомства были привлечены, чтобы найти решение проблемы в долгосрочном плане, и в результате их усилий было зарегистрировано и учтено множество случаев легкой формы заболевания, каковые в прошлые годы не фиксировались и, возможно, даже не были выявлены.

Все это живо напоминает историю о том, как двое нью-йоркских репортеров, Линкольн Стеффенс и Джейкоб Риис, однажды создали волну преступности. Криминальная хроника в газетах разрослась до таких масштабов (по числу упоминавшихся преступных деяний, отведенному им газетному месту и величине шрифта, каким печатались эти сообщения), что общественность потребовала решительных действий. Теодор Рузвельт, в те поры лично возглавлявший реформирование полицейского управления, попал

³¹ То есть такой опрос, в основе которого нет никакой научной методики. – *Прим. ред.*

в незавидное положение. Однако волну преступности он утихомирил самым простым способом: потребовал, чтобы Стеффенс и Риис прекратили нагнетать ситуацию. Корень бед, как выяснилось, состоял в том, что с подачи этой парочки газетные репортеры развернули соревнование, кто сумеет накопать больше всего информации о кражах со взломом и прочих преступлениях. Между тем официальная полицейская статистика не зафиксировала вообще никакого роста преступности.

«Британец старше 5 лет отмачивает себя в горячей ванне в среднем 1,7 раза в неделю зимой и 2,1 раза в летний сезон, – говорится в газетной заметке. – Британка же за неделю принимает ванну в среднем 1,5 раза в неделю зимой и 2,0 раза летом». Источником сведений стал проведенный Министерством общественных работ Великобритании опрос о пользовании горячей водой, в котором участвовали «6000 британских семей». Выборка отвечает требованиям репрезентативности, говорилось в итоговом докладе, и представляется вполне подходящей по размеру. Что оправдывает вывод, который сан-францисская газета *Chronicle's* сформулировала следующим забавным заголовком: «БРИТАНЦЫ ПРИНИМАЮТ ВАННУ ЧАЩЕ, ЧЕМ БРИТАНКИ».



Сведения эти могли бы больше сказать о положении дел, если бы отчет содержал хотя бы намек на указание, будут ли приведенные цифры средним или медианным значением. Но самое слабое место в этих сведениях – это что по ходу исследований произошла подмена понятий. Добытые министерством сведения в реальности говорят

о том, насколько часто британцы, по их собственным словам, принимают ванну, а не о том, насколько часто они делают это на самом деле. Когда вопрос касается такой сугубо интимной темы, да еще в свете британской традиции часто принимать ванну, нельзя исключать, что ответы респондентов и то, как они поступают в реальности, совсем не одно и то же. В жизни британец с равным успехом может принимать ванну как чаще, так и реже своей соотечественницы. Так что единственное, о чем можно судить с уверенностью, это что британцы говорят, что дела обстоят именно так.

Приведу еще несколько разновидностей подмены объекта исследований, чтобы вы были начеку и при случае не пропустили чего-то подобного.

После того как перепись населения показала, что в 1935 г. в США стало на полмиллиона больше ферм, чем пятью годами раньше, многие усмотрели в этом тенденцию к возвращению на фермы. Однако два этих подсчета имели в виду далеко не одно и то же. За те пять лет само определение фермерского хозяйства, которым руководствовалось Бюро переписи населения США, изменилось, и в итоге при новой переписи к категории фермерских были отнесены как минимум 300 000 хозяйств, которые, согласно действовавшему в 1930 г. определению, таковыми не считались и потому учтены не были.

Порой возникают и настоящие нелепости, если цифры основываются на том, что говорят сами люди, – даже когда речь идет об объективных вроде бы фактах. Так, перепись населения выявила большее количество людей в возрасте, скажем, 35 лет, чем тех, кому 34 и 36. В подобных случаях картина искажается оттого, что кто-то из членов семьи, сообщая о возрасте домочадцев и не будучи в нем точно уверенным, часто следует привычке округлять года до величины, кратной пяти. Один из способов обойти подобные ошибки – просить, чтобы респонденты называли не возраст, а дату рождения.

Или такой случай: сообщалось, что население крупного района Китая составляет 28 миллионов человек. Спустя пять лет численность населения была уже 105 миллионов. Однако реальный рост населения составлял лишь малую долю этого почти четырехкратного увеличения. Огромную разницу между двумя этими показателями можно объяснить, только принимая во внимание цели переписей

населения в том районе, и степень готовности самих жителей быть переписанными. Как выяснилось, первая перепись проводилась в интересах налогообложения и военных нужд, а вторая – для того, чтобы разработать меры по борьбе с голодом.

Нечто подобное случилось и в США. Перепись населения 1950 г. выявила больше граждан в возрасте 65–70 лет, чем было зафиксировано десять лет назад в возрастной группе 55–60 лет. Такую разницу явно нельзя списать на иммиграцию. Думается, причиной расхождения послужило желание многих получить социальное страхование, и потому люди в массовом порядке завышали свой возраст. Возможно также, что некоторые представители старшей возрастной категории (за 70) из тщеславия занизили свой реальный возраст.

Еще один образчик подмены объекта исследования явил сенатор Уильям Лангер, когда возопил, что «мы могли бы взять заключенного из “Алькатраса” и поместить на содержание в “Уолдорф-Асторию” – дешевле бы обошлось...» Дело в том, что сенатор от Северной Дакоты ссылался на ранее опубликованные данные, что содержать узника в тюрьме «Алькатрас» стоит \$8 в сутки, а «это стоимость номера в хорошем сан-францисском отеле». Здесь произошла подмена общих затрат на содержание (в «Алькатрасе») на одну только стоимость номера в отеле.



Разновидность приема «после – значит вследствие» в связке с какой-нибудь пафосной бессмыслицей представляет собой еще один способ незаметно подменить объект. Изменение чего-либо *наряду* с чем-либо другим преподносят как перемены *вследствие*. Журнал *Electrical World* однажды предложил вниманию читателей составную диаграмму к редакционной статье под названием «Какое значение для Америки имеет электричество». Глядя на диаграмму, вы могли бы заключить, что по мере увеличения «электрической мощности на фабриках» росла и «средняя почасовая заработка плата». А «среднее число рабочих часов за неделю» сокращалось. Все три явления представляют собой долгосрочные тенденции, и нет никаких свидетельств, позволяющих утверждать, что любая из трех тенденций стала причиной любой другой.

А кроме того, полно и тех, кто сделал что-то первым. Почти каждый может заявить, что стал самым первым в *чем-нибудь*, если не слишком конкретизирует, в чем именно. Так, в конце 1952 г. две нью-йоркские газеты оспаривали друг у друга пальму первенства в рекламировании продовольственных товаров. И каждая была до известной степени права. В обоснование своих претензий газета *World-Telegraph* объясняла, что лидирует по рекламе с полным охватом, то есть помещает ее во всех своих выпусках, и что засчитывается только этот параметр и никакой другой. Ее соперник, газета *Journal-American*³², настаивала, что принимать во внимание следует общее число строк в размещаемых рекламных объявлениях, а по этому показателю она, безусловно, держит первое место. Перед нами та самая уловка, применяемая в погоне за первенством в чем угодно, которая побуждает радиокомментатора в новостях о погоде именовать обычный летний денек не иначе, как «самым жарким вторым июня с 1949 г.».

Подмена объекта исследования затрудняет сопоставление расходов, когда вы размышляете, выгоднее ли вам взять ссуду для приобретения крупной вещи или совершить покупку в рассрочку. Когда вам называют ставку 6 %, она и воспринимается как 6 % – хотя в действительности все может обстоять совсем иначе.

Если вы берете в банке ссуду \$100 под 6 % и в течение года выплачиваете ее каждый месяц равными долями, то за пользование

³² *New-York Journal-American*, вечерняя газета, просуществовавшая с 1937 по 1966 г.; газета *New York World-Telegram* выходила в 1867–1966 гг. – Прим. пер.

деньгами вы платите около \$3. А другая ссуда, которую могут охарактеризовать как \$100 под \$6, обойдется вам вдвое дороже. Таким способом в основном и формируются автомобильные кредиты. Очень хитроумно придумано.

Суть в том, что вы не пользуетесь этими заемными \$100 в течение всего года. По прошествии шести месяцев вы уже выплатите половину этой суммы. Если с вас берут по \$6 со \$100, или 6 % от заемной суммы, то в реальности это обходится вам почти вдвое дороже, то есть порядка 12 %.

Но это еще цветочки по сравнению с той неприятностью, что постигла легкомысленных американцев, которые в 1952–1953 гг. польстились на программы оптового снабжения морожеными мясными продуктами. Когда им объясняли, во что это обойдется, то назывались цифры в пределах от 6 до 12 %. Эти цифры и воспринимались как процент, то есть со всей суммы покупок и в расчете на год. А оказалось, что это проценты с каждого потраченного на продукты доллара, и притом из расчета не за год, а за полгода. Получалось, что в течение полугода следовало регулярно платить по \$12 с каждого \$100, и в действительности это приравнивалось к банковскому проценту порядка 48 % за год. Немудрено, что множество клиентов не выдержали таких трат, отчего большинство программ прогорело.

Иногда, чтобы подменить объект интереса, практикуют семантический подход. Вот пример подобного со страниц журнала *BusinessWeek*:

Бухгалтеры пришли к выводу, что слово «излишки» выглядит отвратительно. Они предлагают исключить его из балансовых отчетов корпораций. Комитет по учетным процедурам Американского института бухгалтеров советует: «...используйте описательные термины, такие как “нераспределенная прибыль” или “удорожание основных средств”».

А вот еще один образец, но уже из газетной статьи, где сообщалось, что Standard Oil побила все рекорды по выручке и ее чистая прибыль достигла миллиона долларов в сутки.

Возможно, директора время от времени подумывают о дроблении акций, поскольку это может принести некоторые

преимущества... если прибыль на акцию будет выглядеть не такой большой...

Есть ли в этом смысл?

«Есть ли в этом смысл?» – такой вопрос почти всегда поставит на место многое возомнившего о себе статистика, если все его маловразумительные построения основаны на недоказанном исходном допущении. Возможно, вам приходилось слышать о формуле удобочитаемости Рудольфа Флеша. Считается, что она позволяет измерить, насколько легок для прочтения изложенный прозой текст, при помощи таких простых и объективных параметров, как длина слов и предложений. Эта идея весьма привлекательна, как и прочие подобные ухищрения, придуманные для того, чтобы свести нечто трудноуловимое к цифрам и подменить суждения чистой арифметикой. Во всяком случае формула Флеша пришлась ко двору тем, кто дает работу авторам (например, издателям газет), пусть даже многие литераторы от нее не в восторге. Формула Флеша строится на допущении, что такие параметры, как длина слов, и определяют удобочитаемость текста. Это допущение, если уж вредничать по полной программе, пока еще только ожидает доказательств.

Некто по имени Роберт Дюфор взялся проверить формулу Флеша на некоторых литературных произведениях, которые оказались у него под рукой. Как свидетельствуют его вычисления, «Легенда о сонной лощине» Вашингтона Ирвинга³³ читается в полтора раза труднее, чем «Государство» Платона. А роман Синклера Льюиса³⁴ «Кэсс Тимберлейн» – труднее, чем эссе Жака Маритена³⁵ «О духовной ценности искусства». В общем, та же история.

Уже при первом взгляде многие статистические данные внушают подозрения в своей истинности. И все же публика принимает их,

³³ Вашингтон Ирвинг (1783–1859) – выдающийся американский писатель-романтик, которого часто называют «отцом американской литературы». – *Прим. пер.*

³⁴ Синклер Льюис (1881–1951) – американский писатель, первый в США лауреат Нобелевской премии по литературе (1930), поздние произведения относятся к жанру социальной литературы, и потому Льюиса заслуженно относят к целой плеяде писателей, которых называли «людьми 1920-х годов». – *Прим. пер.*

³⁵ Жак Маритен (1882–1973) – французский философ, теолог, основатель неотомизма. – *Прим. пер.*

поскольку магия цифр на какое-то время заставляет умолкнуть здравый смысл. Леонард Энджел в своей статье для журнала *Harper's* приводит несколько тому примеров из области медицины.

Примером служат подсчеты знаменитого уролога, согласно которым в США отмечается 8 миллионов случаев рака предстательной железы – этого было бы вполне достаточно, чтобы на каждого мужчину в возрасте наибольшей восприимчивости к этому заболеванию приходилось бы по 1,1 карциномы предстательной железы! Другой пример: видный невропатолог высчитал, что каждый двенадцатый американец страдает мигренями, а поскольку мигрени – причина трети случаев хронических головных болей, это должно было бы означать, что четверть из нас мучаются от изнурительных головных болей, чреватых потерей трудоспособности. А вот и еще одна цифра: 250 000. Ее часто приводят как количество случаев заболевания рассеянным склерозом. По счастью, как показывает уровень смертности, у нас в стране насчитывается не более 30 000–40 000 случаев этой болезни.

На слушаниях по внесению поправок в Закон о социальном страховании часто вспыпал, хотя и в разных формах, аргумент, который выглядел разумным, только если к нему не слишком присматриваться. Сводился он примерно к следующему: если ожидаемая продолжительность жизни составляет только 63 года, было бы обманом и профанацией учреждать программы социального страхования, предусматривающие выход на пенсию в 65 лет, ведь до этого возраста никто не доживет.

Вы и сами можете опровергнуть этот аргумент, если припомните, сколько среди ваших знакомых людей преклонных лет. Однако главная ошибка здесь в том, что данная цифра характеризует ожидаемую продолжительность жизни при рождении, так что резонно ожидать, что примерно половина родившихся проживет дольше. И, кстати говоря, так уж пришлось, что эта цифра взята из самой поздней по сроку официальной таблицы продолжительности жизни и действительна она для периода 1939–1941 гг. Очень возможно, что это породит еще один аргумент на ту же тему, не уступающий глупостью вышеназванному: что практически все сегодня доживают до 65 лет.

Несколько лет назад планирование послевоенного развития в одной крупной компании по производству бытовой техники шло

на всех парах, исходя из той посылки, что уровень рождаемости падает, причем многие годы это считалось само собой разумеющимся. Планы ориентировали делать особый упор на маломощные бытовые приборы, в том числе на небольшие холодильники. Потом на одного из плановиков вдруг накатил приступ здравого смысла. Он вынырнул из своих графиков и диаграмм на достаточно долгое время, чтобы сообразить одну простую вещь: сам он, его коллеги, знакомые, соседи, а также бывшие однокашники за редким исключением воспитывают по трое-четверо ребятишек или планируют завести еще детей. Это подвигло отдел планирования провести кое-какие изыскания со всеми полагающимися диаграммами и графиками, на основе подходов более широких и непредвзятых. И вскоре компания перенесла акцент на бытовые приборы для больших семей, что и стало самым прибыльным для нее направлением.

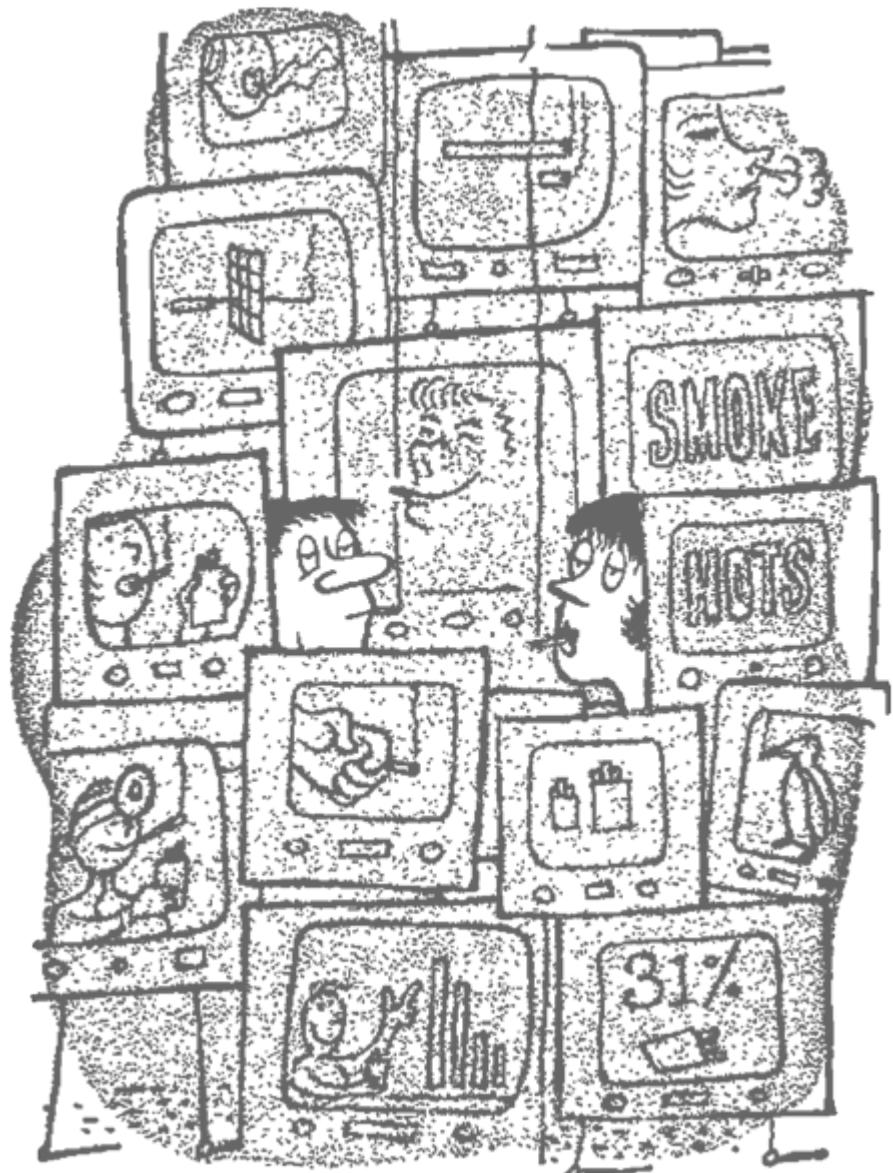
Убедительность точных цифр – еще один фактор, порой вступающий в противоречие со здравым смыслом. Согласно исследованию, о котором писали нью-йоркские газеты, работающей женщине, проживающей со своей семьей, еженедельно требуется \$40,13. Любой, у кого при чтении газет не атрофируется здравый смысл, способен сообразить, что расходы на поддержание души в теле невозможно рассчитать с точностью до последнего цента. И все равно трудно устоять перед чертовским соблазном уверовать в эту цифру, ведь сама точность этих \$40,13 намекает на солидную осведомленность источника и внушает больше уважения, чем формулировка «около \$40».

С таким же подозрением надлежит воспринимать обнародованное несколько лет назад заявление Американского комитета нефтяной промышленности, что средний размер уплачиваемого ежегодного налога на автомобиль составляет \$51,13.

В подобном же смысле полезны и экстраполяции, особенно в той форме пророчеств, что называют прогнозированием тенденций. Но разглядывая прогнозные данные или построенные на их основе диаграммы, крепко помните об одной вещи: тенденция вплоть до сегодняшнего дня вполне может быть фактом реальности, а что касается перспектив на будущее, то они не больше, чем догадки на базе прошлого опыта и имеющихся знаний. К тому же подразумевается, что прогнозы могут реализоваться только «при прочих равных условиях» и «сохранении нынешних

тенденций». Почему-то эти «прочие условия» не желают оставаться равными и все норовят измениться, но ведь иначе жизнь превратилась бы в сплошную скуку.

Дабы продемонстрировать образец абсурда, которым чреваты бездумные экстраполяции, предлагаю рассмотреть тенденцию развития телевидения. В период 1947–1952 гг. число телевизоров в домах американцев возросло почти на 10 000 %. Попробуйте спроектировать эту тенденцию на следующие пять лет, и у вас получится, что в ближайшем будущем страну заполонят – страшно подумать! – миллиарда два телевизоров. Если вам неймется выставить себя в еще более глупом свете, предлагаю за базовый год принять не 1947-й, как мы это сделали, а какой-нибудь предшествующий, и вы запросто «докажете», что в скором времени у каждой американской семьи будет не по сорок, а по сорок тысяч телевизоров.



Моррис Хансен, проводящий научные исследования по заказу правительства, называл сделанный Гэллапом в 1948 г. электоральный прогноз «самой широко разрекламиированной статистической ошибкой всех времен». Однако Гэллап со своим промахом – просто образец прогностических добродетелей, если сравнить его выводы с некоторыми из имеющих самое широкое хождение прогнозов численности населения, которые рассмешили всю Америку. Не далее как в 1938 г. президентская комиссия, составленная сплошь из именитых специалистов, усомнилась, что население США когда-нибудь достигнет отметки 140 миллионов. Но всего через дюжину лет оно превысило эту сакральную цифру на 12 миллионов. В учебниках, изданных так недавно, что ими еще вовсю пользуются в колледжах и университетах, говорится, что пиковое значение численности населения США не превысит

150 миллионов человек, и отмечается, что реально эта цифра достижима не ранее, чем примерно к 1980 г. Такая чудовищная недооценка численности населения проистекала из того, что прогноз строился на следующем допущении: тенденция роста населения останется неизменной. Аналогичное допущение сто лет назад сослужило такую же дурную службу. Прогноз численности населения исходил из того, что оно и дальше будет прирастать теми же темпами, что и в 1790–1860 гг. В своем втором послании к конгрессу Авраам Линкольн предсказал, что в 1930 г. население США достигнет 251 689 914 человек.

Прошло немного времени, и в 1874 г. Марк Твен обобщил абсурдную сторону всяких экстраполяций в произведении, называемом «Жизнь на Миссисипи»³⁶:

За сто семьдесят шесть лет Нижняя Миссисипи укоротилась на двести сорок две мили, то есть в среднем примерно на милю и одну треть в год. Отсюда всякий спокойно рассуждающий человек, если только он не слепой и не совсем идиот, сможет усмотреть, что в древнюю силурийскую эпоху, – а ей в ноябре будущего года минет ровно миллион лет, – Нижняя Миссисипи имела свыше миллиона трехсот тысяч миль в длину и висела над Мексиканским заливом наподобие удочки. Исходя из тех же данных, каждый легко поймет, что через семьсот сорок два года Нижняя Миссисипи будет иметь только одну и три четверти мили в длину, а улицы Каира и Нового Орлеана сольются, и будут эти два города жить да поживать, управляемые одним мэром и выбирая общий городской совет. Все-таки в науке есть что-то захватывающее. Вложишь какое-то пустяковое количество фактов, а берешь колossalный дивиденд в виде умозаключений. Да еще с процентами.



Я счастливый человек, ведь у меня такая прекрасная работа: искать и издавать умные книги, общаться с их авторами, узнавать от них много нового и интересного.

Издав несколько сотен деловых и развивающих книг, могу

³⁶ В переводе В. Райт-Ковалевой. – *Прим. пер.*

уверенно сказать, что книга для автора почти всегда – не цель, а результат. В какой-то момент автор понимает, что обладает уникальным опытом, рассказ о котором поможет другим людям стать лучше и узнать о жизни что-то полезное. Через некоторое время это понимание становится настолько осознанным, что в прямом смысле слова доводит автора до ручки (или до клавиатуры), заставляя написать книгу.

Вполне возможно, что Вы, читающий эти строки сейчас, – потенциальный автор книги, которая станет бестселлером и даст людям нужные знания и навыки.

Мы будем очень рады стать издателем Вашей книги! Наша креативная команда приложит все усилия, чтобы Ваша книга получилась красивой и качественной, чтобы она была заметна в магазинах, чтобы ее активно обсуждали.

**Присылайте нам Ваши рукописи,
Вам понравится работать с нами!**

С уважением,



*Сергей Турко,
кандидат экономических наук,
главный редактор издательства «Альпина
Паблишер»*

Заходите сюда alpina.ru/a