

**МАТЕМАТИЧЕСКИЕ МЕТОДЫ В
ЖУРНАЛИСТИКЕ**

**ИСПОЛЬЗОВАНИЕ МАТЕМАТИКИ И
СТАТИСТИКИ В НОВОСТЯХ**

ВТОРОЕ ИЗДАНИЕ

Сара Коэн

Ассоциация журналистов-расследователей и редакторов

Редактор:

Дэн Клакстон

Дизайн:

Уэнди Грэй

ISBN 0-9766037-1-3

Одна из специальных серий книг и пособий для журналистов,
выпускаемых Ассоциацией журналистов-расследователей и
редакторов.

Свои комментарии и предложения можно отправить
по электронной почте: beatbooks@ire.org

Данное издание было подготовлено при финансовой поддержке
Фонда Сорос-Казахстан. Содержание публикации
отражает точку зрения авторов и может не совпадать
с точкой зрения Фонда Сорос-Казахстан.

Авторские права защищены 2014 ©
Ассоциация журналистов-расследователей и редакторов
и Сара Коэн

ПРЕДИСЛОВИЕ

Публикация данного пособия была бы невозможна без всесторонней помощи и советов. В подготовке рукописи принимали участие практики, такие как Нил Боровски, работавший тогда в The Philadelphia Inquirer, Дэн Браунинг из Minneapolis StarTribune, Рон Кэмбэл из Orange County Register, Крис Калахэн, ныне декан Школы Журналистики из Университета штата Аризона, Джо Крэйвэн Мак Гинти из The Wall Street Journal, Стивен Дойк из Университета штата Аризона, Форд Фессенден из The New York Times, Дэн Китинг из The Washington Post, Дженифер ЛяФлёр из Центра журналистских исследований, Энди Лерэн из The New York Times, Филип Майер из Университета Северной Каролины и Патрик Ремингтон из Университета штата Висконсин. На протяжении многих лет нам помогали и другие люди.

Спасибо всем, кто помог сформировать эту книгу, а также представил свои рекомендации. Также выражаю благодарность Уэнди Грэй за дизайн книги и графики, Лорэн Грандэштаф за редактуру и Зане Ло из IRE за поиск справочной информации для журналистов.

Если вы желаете узнать краткое содержание концепций, изложенных в этой книге, то можете посмотреть видео, отснятое Школой Журналистики Колумбийского Университета для студентов весеннего семестра 2014 года: <https://www.youtube.com/watch?v=IZjsCycсcNc>

Особая благодарность – Лен Брузес и Брэнт Хьюстон из IRE – за реализацию первоначального проекта, Дэну Клакстону и Марку Хорвиту – за его последующее продолжение.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	9
ГЛАВА 1: ЧТО ТАКОЕ ЧИСЛО?	11
СЕМЬ ПОДСКАЗОК КАК ОСВОБОДИТЬ ВАШ МАТЕРИАЛ ОТ ЧИСЕЛ, СОХРАНИВ ОБЪЕКТИВНОСТЬ	11
ПРИБРЕТАЙТЕ ХОРОШИЕ ПРИВЫЧКИ	17
ГЛАВА 2: МАТЕМАТИЧЕСКИЙ СПРАВОЧНИК ДЛЯ ОТДЕЛА НОВОСТЕЙ	20
ДРОБИ, ОТНОШЕНИЯ, ПРОЦЕНТЫ И ПОКАЗАТЕЛИ НА ДУШУ НАСЕЛЕНИЯ.....	20
<i>ДРОБИ И ПРОЦЕНТЫ</i>	21
<i>«НА ЧЕЛОВЕКА» (ИЛИ НА ДУШУ НАСЕЛЕНИЯ): УРОВЕНЬ СМЕРТНОСТИ, УРОВЕНЬ ПРЕСТУПНОСТИ И ДРУГИЕ ПОКАЗАТЕЛИ</i>	24
ДОЛИ (ПРОДОЛЖЕНИЕ).....	26
<i>СЛОЖЕНИЕ И ВЫЧИТАНИЕ ДОЛЕЙ</i>	27
<i>ОТНОШЕНИЕ ДОЛЕЙ, ИЛИ «ОТНОСИТЕЛЬНЫЙ РИСК»</i>	27
<i>ЗАМЕЧАНИЯ ПО ПРОВЕРКЕ ЗНАЧИМОСТИ</i>	30
ВЫЧИСЛЕНИЕ ИЗМЕНЕНИЯ.....	31
<i>ПРОСТЫЕ ИЗМЕНЕНИЯ, ИЛИ ФАКТИЧЕСКИЕ ИЗМЕНЕНИЯ</i>	32
<i>ОТНОСИТЕЛЬНОЕ ИЗМЕНЕНИЕ, ИЛИ ОТНОСИТЕЛЬНАЯ РАЗНИЦА В ПРОЦЕНТАХ</i>	33
АНАЛИЗ ИЗМЕНЕНИЙ: ГОДОВЫЕ ПОКАЗАТЕЛИ И КОРРЕКТИРОВКА ПОКАЗАТЕЛЕЙ С УЧЕТОМ ИНФЛЯЦИИ	39
<i>ГОДОВЫЕ ПОКАЗАТЕЛИ</i>	39
<i>КОРРЕКТИРОВКА С УЧЕТОМ ИНФЛЯЦИИ</i>	44
ПОНИМАНИЕ СРЕДНИХ – СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ, МЕДИАНА, МОДА	48

СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ, ИЛИ СРЕДНЕЕ	49
МЕДИАНЫ, ИЛИ СРЕДИННЫЕ ЗНАЧЕНИЯ	50
МОДА, ИЛИ НАИБОЛЕЕ ЧАСТО ВСТРЕЧАЮЩЕЕСЯ ЗНАЧЕНИЕ	52
ВСЕ ЛИ ЗНАЧЕНИЯ ЯВЛЯЮТСЯ СРЕДНИМИ?	53
БОЛЕЕ УГЛУБЛЕННОЕ ИЗУЧЕНИЕ ЦЕНТРАЛЬНОЙ ТЕНДЕНЦИИ: УСРЕДНЕНИЕ СРЕДНИХ ВЕЛИЧИН, ИЗМЕНЕНИЯ И ЗНАЧИМОСТЬ	55
ЗАДАЧА УСРЕДНЕНИЯ СРЕДНИХ	55
ЧИСЛА, КОТОРЫЕ ПОМОГУТ ПРОИЗВЕСТИ ПРОВЕРКУ ПРИ РАСЧЕТЕ СРЕДНИХ ВЕЛИЧИН	60
УСЕЧЕННОЕ СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ ЗНАЧЕНИЕ ..	62
ГИСТОГРАММЫ	62
СРАВНЕНИЕ СРЕДНИХ ВЕЛИЧИН ВО ВРЕМЕНИ	63
ГЛАВА 3: РАБОТА С ГРАФИКАМИ	66
ВИДЫ ДИАГРАММ	67
ПРЕДОТВРАЩЕНИЕ ЗАБЛУЖДЕНИЙ	67
МАЛЫЕ МНОЖИТЕЛИ И «ИСКРОГРАФИК» (SPARKLINES)	69
КАРТЫ	71
ПРОБЛЕМА МАСШТАБА	71
ИСПОЛЬЗОВАНИЕ ВИЗУАЛИЗАЦИИ КАК ИНСТРУМЕНТ РЕПОРТАЖА	75
ТОЧНОСТЬ, ДОСТОВЕРНОСТЬ И ИНТЕРАКТИВНАЯ ВИЗУАЛИЗАЦИЯ	76
ГЛАВА 4: ТИПОВЫЕ ПРИМЕРЫ ИНФОРМАЦИОННЫХ МАТЕРИАЛОВ	77
ОСВЕЩЕНИЕ ТЕМЫ БЮДЖЕТА – ИСПОЛЬЗОВАНИЕ ПРОЦЕНТОВ, КОРРЕКТИРОВКИ УРОВНЯ ИНФЛЯЦИИ И ВЕЛИЧИН, ФИКСИРУЮЩИХ ИЗМЕНЕНИЯ	77
ПОНИМАНИЕ БЮДЖЕТА	77
КАК ЖУРНАЛИСТУ ПОДГОТОВИТЬСЯ К СЕЗОНУ БЮДЖЕТНОГО ПЛАНИРОВАНИЯ	78

НАЛОГИ НА ИМУЩЕСТВО	79
<i>РАСЧЕТ НАЛОГОВ НА ИМУЩЕСТВО</i>	79
<i>ДВА ПРИМЕРА РАСЧЕТА НАЛОГОВ</i> <i>НА ИМУЩЕСТВО</i>	81
ОСВЕЩЕНИЕ В СМИ ЗЕМЛЕТРЯСЕНИЙ, УРАГАНОВ И ПОГОДЫ – ИСПОЛЬЗОВАНИЕ ИНДЕКСОВ	84
<i>ЗЕМЛЕТРЯСЕНИЯ И ШКАЛА РИХТЕРА</i>	84
<i>УРАГАНЫ И ТОРНАДО</i>	85
<i>ИНДЕКС ЖАРЫ (ТЕПЛА) И ВЕТРО-ХОЛОДОВЫЕ</i> <i>ПАРАМЕТРЫ</i>	86
СТАДИОНЫ, УВОЛЬНЕНИЯ И ОЦЕНКА ЭКОНОМИЧЕСКОГО ВОЗДЕЙСТВИЯ	87
<i>КАК РАСПОЗНАТЬ ПОДВОДНЫЕ КАМНИ</i> <i>В ЭКОНОМИЧЕСКИХ ОБОСНОВАНИЯХ</i>	88
ГЛАВА 5: ОБЗОРЫ И ОПРОСЫ – ЧИСЛА	
КАК ОБОБЩЕНИЯ И ДОГАДКИ	90
ЗНАЙТЕ СВОЕГО СОЦИОЛОГА	90
ФОРМУЛИРОВКА ВОПРОСОВ И ИХ ПОСЛЕДОВАТЕЛЬНОСТЬ	92
МЕТОДОЛОГИЯ ВЫБОРКИ И ПРОЦЕНТ ОПРОШЕННЫХ	93
ПРЕДЕЛ ПОГРЕШНОСТИ (ОШИБКИ В ВЫБОРКЕ)	94
ЭЛЕМЕНТЫ ПОГРЕШНОСТИ ВЫБОРКИ	95
<i>ВАШ САМОСТОЯТЕЛЬНЫЙ ПОДСЧЕТ</i> <i>ДОПУСТИМОГО ПРЕДЕЛА ПОГРЕШНОСТИ</i>	96
ОПРОСЫ: ДОПОЛНИТЕЛЬНЫЕ ПОДСКАЗКИ	100
ГЛАВА 6: 10 САМЫХ РАСПРОСТРАНЕННЫХ «ЛЯПОВ»	
В НОВОСТЯХ: ОТ ПРОСТЫХ МАТЕМАТИЧЕСКИХ	
ОШИБОК ДО ОШИБОК В СУЖДЕНИИ	102
№ 10. ОБРАТНАЯ ВЕРОЯТНОСТЬ, ИЛИ КАКОВЫ ШАНСЫ ОБНАРУЖИТЬ КЛАСТЕР БОЛЬНЫХ РАКОМ?	102
№ 9. ТРЮК С ВОЛШЕБНОЙ РЕЗИНКОЙ, ИЛИ ОШИБКА В ПОНИМАНИИ «ЗАКОНА СРЕДНИХ ЧИСЕЛ»	103

№ 8. СРЕДНЕЕ ОТ СРЕДНЕГО ЧИСЛА – ЛУЧШЕ НЕ ПЕРЕУСЕРДСТВОВАТЬ	104
№ 7. К ЧЕМУ ЭТО ПРИМЕНИМО? ИЛИ НЕ УПУСКАЙТЕ ИЗ ВИДУ ИСХОДНЫЕ ОБЪЕМЫ.....	104
№ 6. ОШИБОЧНАЯ ТОЧНОСТЬ	105
№ 5. ЕСЛИ БЫ, ТО ТОГДА ...: ПУТАНИЦА МЕЖДУ ПРИЧИНОЙ И СЛЕДСТВИЕМ.....	106
№ 4. ПРЕНЕБРЕЖЕНИЕ ВНУТРЕННИМ ГОЛОСОМ.....	108
№ 3. НЕАДЕКВАТНЫЕ ИСТОЧНИКИ	108
№ 2. ПЕРЕПОЛНЕНИЕ МАТЕРИАЛА ЧИСЛАМИ.....	109
№ 1. ПОЗВОЛИТЬ ПРИЧИНАМ ОТ 10 ДО 2 ВВЕСТИ НАС В СТУПОР	110

**ГЛАВА 7: ЛОТЕРЕИ, УДАРЫ МОЛНИИ,
ПРОДОЛЖИТЕЛЬНОСТЬ ЖИЗНИ – КОММЕНТАРИИ
НА ТЕМУ ВЕРОЯТНОСТИ.....**

ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТИ.....	111
ВЫИГРЫШ В ЛОТЕРЕЮ	113
СОВЕТЫ ДЛЯ ПОЛЬЗОВАТЕЛЕЙ ЭЛЕКТРОННЫХ ТАБЛИЦ Excel	116
МАЛЕНЬКИЕ ВЕРОЯТНОСТИ X БОЛЬШИЕ ЧИСЛА = ГРУБЫЕ ОШИБКИ.....	116
СОВПАДЕНИЯ С ТЕЧЕНИЕМ ВРЕМЕНИ – КОМБИНИРОВАНИЕ ВЕРОЯТНОСТЕЙ	117
БИРЖЕВЫЕ БРОКЕРЫ	118
ЧИСЛО СТРАЙКОВ.....	118
ВСПЫШКИ ОНКОЗАБОЛЕВАНИЙ.....	119

**ГЛАВА 8: ДОПОЛНИТЕЛЬНЫЕ РЕСУРСЫ
ДЛЯ ЖУРНАЛИСТОВ**

ОБУЧЕНИЕ	122
РЕСУРСНЫЙ ЦЕНТР.....	123
ПУБЛИКАЦИИ	123
ОНЛАЙН-РАССЫЛКИ.....	124
В ИНТЕРНЕТЕ.....	124
ДРОБИ И ПРОЦЕНТЫ – СРАВНИТЕЛЬНАЯ ТАБЛИЦА.....	125

ОБ АВТОРЕ	127
СОВЕТ ДИРЕКТОРОВ IRE (2014 г.).....	128
ШТАТ РУКОВОДИТЕЛЯ IRE	129
КОНТАКТНАЯ ИНФОРМАЦИЯ.....	129

ВВЕДЕНИЕ

*Статистика – участь людей
с высохшими слезами.
Пол Бродер.*

Два года, будучи директором тренинговых программ в Ассоциации журналистов-расследователей и редакторов, я пытался помочь журналистам, редакторам и продюсерам преодолеть страх перед числами. Это не столь занимательно, как стучать по клавиатуре. Журналисты остро осознавали, что для качественного повышения уровня своих материалов, они должны научиться работать с числами, которые встречаются повседневно и повсеместно.

Это пособие призвано помочь вам в понимании того, как устроен живой мир чисел. Изложение в этой книге основано на идеях, которые затрагиваются в материалах по навыкам счета, статистики и графики. Ведь уместное использование в материале наиболее верной цифры зависит от критического мышления, равно как и выбор цитаты, ситуации или образа.

Наша задача заключается в формировании видения будущего материала не только посредством текста или визуально, но и в цифрах. Я надеюсь, что вы с большей уверенностью станете применять в своих материалах числа, досконально изучив это пособие либо используя его как справочник.

Десять лет после выхода первого издания этой книги мы воочию наблюдали, как менялся подход к использованию числовых данных в новостных материалах: от простых диаграмм в газете до ослепляющих интерактивных дисплеев. Развитие «журналистики данных» способствовало более свободному применению и анализу журналистами статистических данных.

Так есть ли действительно потребность писать еще один талмуд по этой теме? Думаю, все же «да».

Одни настойчиво твердят, что журналисты не способны адекватно воспринимать и оценивать практически любые числа, которые публикуют в своем материале. Однако идеальных чисел нет. Другие же высмеивают журналистов, выбирая примеры некорректного использования чисел, полагая, что каждый материал, как и математическая задача, должен сводиться к поиску верного ответа. В итоге, они часто убеждают нас в том, что мы не способны использовать числа без боязни критики со стороны.

Другие пособия, однако, уделяют особое внимание статистическим проблемам, которые реже встречаются в отделах новостей: теория вероятностей, составление выборки и разработка методологии социологических опросов. Эти аспекты также отражены в книге, и для интересующихся составлен краткий справочник-шпаргалка.

Но, если ваша ежедневная работа – это освещение вопросов политики и здравоохранения, где вы часто готовите информацию о научных проектах или соцопросах, я рекомендую вам изучить «Новости и числа» Виктора Кона («News and Numbers» by Victor Cohn) и «Справочник для редакции по опросам и обзорам» («The Newsroom Guide to Polls and Surveys»).

ГЛАВА 1: ЧТО ТАКОЕ ЧИСЛО?

Зачастую мы слишком много требуем от чисел и поэтому попадаем в тупиковую ситуацию. В материале число, прежде всего, выражает определенное мнение. Оно просто скрыто за завесой научных методов и создает определенное авторитетное мнение. Некоторые мнения обоснованы годами беспристрастных исследований. Другие – нет.

Число – это краткое обобщение. Некоторые числа помогают создать точную картину живого мира, который они отображают. Другие – нет.

Число – это предположение. Президентские выборы в США 2000 года, исход которых зависел от нескольких десятков избирателей во Флориде, напомнили нам, что даже столь простое как число голосов на деле может оказаться неточным. Некоторые предположения достаточно точны, чтобы использовать их без сомнения. Другие – нет.

Если помнить эти свойства чисел, то можно избежать так называемой лжи, в распространении которой обвиняются числа. Использовать числа как мнение, краткое обобщение или предположение – один из действенных способов минимизировать отупляющее нагромождение чисел в материале.

СЕМЬ ПОДСКАЗОК КАК ОСВОБОДИТЬ ВАШ МАТЕРИАЛ ОТ ЧИСЕЛ, СОХРАНИВ ОБЪЕКТИВНОСТЬ

Ранее было допустимо, что журналисты и редакторы не умеют производить математические расчеты. Это было даже похвально, возможно в силу представления автора материала в роли искусного рассказчика или наблюдателя человеческой жизни. Но теперь все иначе. Журналисты, испытывающие страх перед числами, воспринимаются недостаточно квалифицированными, т.е. способными

выполнять лишь часть требуемой от них работы. Журналисты, которые поддаются этому страху, вряд ли смогут освещать сложные темы. Как преодолеть этот страх? Вот несколько простых подсказок, чтобы «приручить» числа и улучшить качество ваших материалов.

1. ИСПОЛЬЗУЙТЕ МЕНЕЕ ВОСЬМИ ЦИФР В ОДНОМ АБЗАЦЕ

Уменьшение количества цифр в материале может улучшить качество его освещения. Совет, которым пользуются многие: сократите количество цифр в абзаце до восьми.

Не выстраивайте абзацы с восемью числами друг за другом. Разбавьте их историями, пояснениями ваших источников, цитатами или наблюдениями.

Этот прием, позаимствованный из книги «Искусство и ремесло написания очерков» Уильяма Бланделла («Art and Craft of Feature Writing» by William Blundell), показывает, что важно выбирать числа так же тщательно, как и цитаты. Он также учит заменять числа словами, используя «вдвойне», «почти» или «более». Он также советует вам требовать от своих источников аналогичного упрощения данных для лучшего восприятия информации.

Почему именно восемь цифр и не более? Как правило, в информацию из восьми цифр можно вместить один год и два существенных числа. Однако попытайтесь сократить количество цифр, чтобы подчеркнуть наиболее важную мысль.

Ниже приведена обновленная версия примера Бланделла. Такое предложение сложно понять:

Бюджет «Офиса по делам чрезмерных избытков»¹ увеличился на 48 процентов в 2013 году, с \$700,3 миллиона до \$1,03 миллиарда.

Подумайте, как это можно переписать.

¹ Название вымышленное.

За прошлый год бюджет «Офиса по делам чрезмерных избытков» вырос почти наполовину, превысив \$1 миллиард.

Если вы будете следовать этому примеру, то убедитесь, что ваш материал будет читаться более гладко и обретет больше смысла. Вы также можете поработать совместно с корректорами или отделом графики, чтобы на сэкономленном за счет текста месте подать информацию через диаграммы.

2. ЗАПОМНИТЕ НАИБОЛЕЕ ЧАСТО ВСТРЕЧАЮЩИЕСЯ ЦИФРЫ

Самое сложное в работе с числами – это ответ на вопрос, «Это много или мало?»

Чтобы ответить на этот вопрос, вы должны знать основные цифры, наиболее часто встречающиеся в вашей работе.

Например, американские репортеры непременно должны владеть такой базовой информацией как: население США (приблизительно 317 млн. чел.), бюджет страны (в 2012 году оценен в размере \$16,2 трлн.) и ежегодный доход среднестатистической семьи (приблизительно \$ 51 000).

В каждой сфере есть присущие ей числа.

При освещении вопросов образования это может быть размер заработной платы учителей, бюджет Министерства образования, проходной балл студентов в лучшие и худшие школы. Для материалов на медицинскую тематику вы, возможно, должны знать статистику по ежегодной смертности, число курильщиков или основные причины смертности (в стране). В отчетах правительства крайне важна информация о расходах на вооружение и о местном бюджете.

Джон Аллен Паулос называет это «пониманием реальных масштабов». Вы можете считать это пониманием разницы между большим и маленьким значением, которая объясняет проблематику понятия «в сравнении с чем-либо». Как бы то ни было, понимание реальных масштабов, по крайней мере, уменьшает количество ошибок в материале.

Следующий случай произошел в 1990-е годы. Работая как-то в выходной день, редактор, не посоветовавшись с журналистом, уменьшила в написанном материале данные об объеме продаж сети магазинов Walmart с \$1,2 млрд. до \$1,2 млн., посчитав, что «миллиард» – это слишком много. Ее ошибкой стало непонимание масштабов продаж. Если бы она знала, что в 1990-е годы для сети магазинов Walmart \$1 млрд. от продажи – не является особенно большим показателем, или что в действительности объем их продаж превышает 100-кратные размеры этой величины, она бы не совершила эту ошибку.

Как вы можете запомнить нужные вам эти цифры? Это не так трудно, и вам не нужно заучивать справочник. Уделяйте повышенное внимание цифрам. Попробуйте сначала догадаться. Это поможет закрепить в памяти правильный ответ.

3. ОКРУГЛЯЙТЕ ПОСТОЯННО

Заметьте, приводимые до сих пор в тексте числа далеко не являются точными числами.

На самом деле, вы можете себе позволить быть менее точным на начальной стадии написания материала. Максимально точное число необходимо лишь для того, чтобы охарактеризовать его в материале. Причем, слово «охарактеризовать» является в этом случае ключевым. Филип Мейер в своей книге «Журналистика точности» («Precision Journalism» by Philip Meyer) говорит: «Десятичные значения чисел даны для общего понимания, но не для того, чтобы заострять на них внимание». В таком случае, дополнительная точность нужна лишь тогда, когда это имеет значение.

Действительно, при подготовке материала не помешает знать, что население США составляет более 300 млн. Этого вполне достаточно для оценочных суждений, а также облегчает анализ некоторых чисел в уме.

Подобная практика освобождает мозг от точных чисел, исключает излишнюю скрупулезность, в которую свойственно впадать многим

при работе с цифрами, а также позволяет не заикливаться на бессмысленных различиях.

4. УЧИТЕСЬ МЫСЛИТЬ СООТНОШЕНИЯМИ

Большая часть математических задач в редакции строится в какой-то степени вокруг тех или иных соотношений: процент, относительная разница в процентах, коэффициент или какие-либо показатели на одного человека.

Мы не можем четко представить себе в уме очень большое, либо очень маленькое число. Например, можно отчетливо представить число «пять»: будь то пальцы рук или ног, большинство коллегии Верховного Суда, или число очков / пробежек, выигранных в бейсбольном матче. Но как только вы попытаете представить цифру 355, восприятие становится менее четким. Хорошо, возможно, это – кинотеатр. Как только вы измените это значение на 317 млн., визуализация практически невозможна. Аналогичная ситуация происходит в том случае, когда цифры ничтожно малы. Как только значение опускается ниже $1/4$, визуализировать его еще сложнее.

Чтобы сделать числа понятными, научитесь мыслить простыми соотношениями.

Давайте рассмотрим это на примере расходов на устранение компьютерной неполадки 2000 года. Согласно распространенным данным американским компаниям эти расходы обошлись в \$50 миллиардов. Насколько это много – \$50 миллиардов? В то время это было меньше, чем собственный капитал Билла Гейтса. Это была стоимость ущерба от двух ураганов. Это был суммарный доход жителей города Портленд штата Орегон.

Мы не хотим сказать, что описанная выше ситуация не имеет новостной ценности. Однако широкое распространение информации по расходам в сочетании с неопределенностью, ее окружающей, сделало это событие более значительным и достойным новостей. Но сложившаяся ситуация финансовой катастрофой не была.

5. ИСПОЛЬЗУЙТЕ ПРИЕМЫ ИЗ ПОВСЕДНЕВНОЙ ЖИЗНИ

При работе с цифрами у спортивных фанатов есть преимущество. Баскетболисты привыкли думать из расчета «1 к 10» или «2 к 8». Бейсбольные болельщики знают, как вычислить ставку или очко, заработанное на пробежке, или от подачи питчера или его средний уровень.

Но даже если вы не фанат спорта, некоторые арифметические действия вы можете сделать интуитивно. Вы без труда можете посчитать, сколько часов осталось до окончания срока подготовки, или сколько должен стоить товар стоимостью в \$20 на распродаже с 10-процентной скидкой.

Если вы делаете покупки на распродажах, откладываете сбережения на свою пенсию, следите за спортивными событиями или оставляете чаевые в ресторане, то вы, наверняка, знаете большую часть математических расчетов, которые мы ежедневно применяем при подготовке материалов новостей. Просто преобразуйте свои вопросы под этот ракурс.

Например, если вы знаете, что латиноамериканцы составляют более 15 процентов населения США, то без труда можно вычислить численность их населения. Это также легко, как и рассчитать чаевые в ресторане.

6. ПРЕДСТАВЬТЕ ЧИСЛО СВОЕЙ МЕЧТЫ И ВЫЧИСЛИТЕ ЕГО, ЕСЛИ ОНО ВАМ НЕ ДАНО

Журналист Дрю Салливан рекомендует репортерам «представить себе успех», когда материал пишется через мучения. Эта уловка работает и в тех случаях, когда вы не за компьютером.

Сначала напишите предложение целиком без ключевой цифры. Предположим, у вас есть некая зацепка: будь это 10 процентов или 90 процентов, \$2 или \$300 000. Затем вычислите то самое число, которым необходимо заполнить пробел.

На деле вы можете обнаружить, что эти числа не отвечают вашим первичным суждениям, либо они не отвечают на поставленные вами вопросы. Это, конечно, печально. Мы зачастую хотим выплеснуть всю информацию из блокнота на страницу. Но так вы сможете заранее понять, какой информации еще не хватает.

7. УЧИТЕСЬ ДРУГ У ДРУГА

У нас у всех есть сильные стороны.

Журналисты, работающие на радио и телевидении, интуитивно понимают неспособность своих зрителей и слушателей одновременно сосредоточиться на множестве цифр. Журналисты, которые по большей части пишут тексты, зачастую работают с отделами информационной графики и веб-дизайнерами, чтобы переместить значимые цифры из текста материала в графики и диаграммы. Прибегайте к возможностям редакции и коллег из других сфер. Если вы работаете самостоятельно, обращайтесь за помощью к экспертам в различных отраслях, включая членов IRE.

ПРИБРЕТАЙТЕ ХОРОШИЕ ПРИВЫЧКИ

Отчасти наша боязнь перед цифрами – это результат хороших привычек, благодаря которым реально избежать исправлений в материале и ошибочных выводов.

Эти привычки банальны, но если вы их приобретете, то скоро заметите, что ваша работа идет значительно быстрее.

ИСПОЛЬЗУЙТЕ ДОСТУПНЫЕ ИНСТРУМЕНТЫ

Существуют инструменты, которые могут помочь нам с арифметикой. Используйте их. Все компьютеры снабжены калькулятором.

Если вы научитесь применять электронные таблицы Excel, то вскоре поймете, что это еще более мощный калькулятор, который производит множественные расчеты по формулам посредством одного щелчка мыши.

Рональд Кэмпбелл, который обучал математике журналистов в газете The Orange County Register, как-то сказал: «В любом случае математика никогда не была вашим коньком. Так используйте же легкодоступные инструменты».

ПРОВЕРЯЙТЕ СНОВА И СНОВА

Взяв в руки калькулятор, будьте готовы произвести расчеты дважды. Журналистам не хватает такой привычки.

Если у вас нет времени, то попросту не работайте с цифрами.

Те, кто работает в электронной таблице Excel, будьте готовы вручную проверить некоторые формулы, перепроверить любые набранные цифры и сравнить полученные суммы с суммами заявленных значений цифр из отчетов.

Некоторые журналисты на распечатке черновика своего материала обводят каждую цифру (и имя) для дальнейшей перепроверки. Эти цифры включают в себя года, данные, полученные из источников, и самостоятельно подсчитанные данные. Попробуйте сделать то же самое, вы удивитесь, узнав, сколько ошибок скрывалось в черновике вашего материала.

Работа над математическими расчетами может быть облегчена, если вы можете работать с кем-то, кому вы доверяете, в команде.

ВСЕГДА ПРОВЕРЯЙТЕ СЕБЯ: МОЖЕТ ЛИ БЫТЬ ТАК, А НЕ ИНАЧЕ

Используя цифры в предложении, параллельно включайте логику: поверили ли бы вы сами в легитимность этих цифр, если бы ваш коллега-журналист, а не эксперт-математик, представил бы вам эти цифры?

Если ответ «нет», то самое время обратиться к советам, изложенным в этой книге: к правилам вычисления соотношений, к способам округления цифр и к изучению базовых статистических данных по освещаемой вами тематике.

В октябре 1998 года читатель The Washington Post в своем письме редактору подверг сомнению информацию о том, что Бюро по переписи населения США намеревается опросить 27 млн. американцев из 78 000 домохозяйств. Читатель оказался прав. Если слепо верить этой ошибочной информации, то получается, что в одном домохозяйстве проживает 350 человек.

Некоторые числа просто абсурдны. Лишний раз убедитесь, что нет ошибок в расчетах, и исключены опечатки.

ГЛАВА 2: МАТЕМАТИЧЕСКИЙ СПРАВОЧНИК ДЛЯ ОТДЕЛА НОВОСТЕЙ

Теоретически любое число, которое мы используем, должно быть с чем-то сопоставлено – с другим временем, с другим местом либо с чем-то более общим. Этот математический справочник подскажет вам, как корректно сопоставлять цифры, используя три приема: сопоставление «на» нечто, изменение и среднее.

Эти приемы являются простыми арифметическими задачами – формулами школьной программы за четвертый класс. Однако, прежде чем мы стали журналистами, у нас было десять лет, чтобы их подзабыть и выработать в себе фобию к математике.

Хорошая новость заключается в том, что большинство новостных материалов далеко не требуют знаний высшей математики. Освойте простую арифметику, и практически любой журналистский материал будет вам под силу.

ДРОБИ, ОТНОШЕНИЯ, ПРОЦЕНТЫ И ПОКАЗАТЕЛИ НА ДУШУ НАСЕЛЕНИЯ

Как правило, можно существенно сократить материал, переведя числа в отношение, пропорцию, или процент. «Один из четырех» – это дробь, или отношение. «Сорок процентов» – это еще одна пропорция или отношение. Также как и 235 смертей на 100 000 человек – еще одна ее вариация.

Проценты и дроби позволяют взглянуть на числа в масштабе, сохранив их реальные пропорции.

Отношения также позволяют сравнивать два числа с разными знаменателями. Аналогично, можно сравнить два маленьких числа без искажения их действительных значений.

Если ваш материал перегружен цифрами, то замена их простым отношением «один из четырех», или 25 процентов упростит понимание читателя.

ДРОБИ И ПРОЦЕНТЫ

«Проценты – это дроби. Дроби – это проценты».

Уяснив это, вы поймете главное свойство процентов: это – пропорции или отношения, выраженные как часть от целого, равного 100.

Разбивка слова «*per cent*» на два помогает понять: *per* («на каждого») и «*cent*» (производная от англ. «*century*» – «век» или 100). Таким образом, 25 процентов буквально означает 25 из каждых 100 человек, долларов, автомобилей и т.п. Дроби более просты для восприятия читателей и зрителей, нежели проценты, поэтому чаще применяйте дроби.

<p>ВЫЧИСЛЕНИЕ ПРОЦЕНТА:</p>	<p>Шаг 1. Найдите основание – общее число элементов всех групп.</p> <p>Шаг 2. Разделите число элементов рассматриваемой группы на основание. Помните, что знак дроби (/) означает «разделить на» (÷).</p> <p>Шаг 3. Переместите запятую на две цифры вправо (или умножьте на 100), чтобы получить часть от ста или процент.</p> <p>Шаг 4. Округлите ответ с точностью до одной десятой. Еще лучше – подберите ту дробь, которая будет наиболее понятна вашим читателям.</p>
<p>ФОРМУЛА:</p>	<p>Шаг 1. Общее число элементов = основание</p> <p>Шаг 2. (Категория / основание) = пропорция или отношение</p> <p>Шаг 3. Пропорция × 100 = процент</p> <p>Шаг 4. Округлите процент</p>
<p>ПРИМЕР:</p>	<p>Если 58 человек утверждают, что будут голосовать на предстоящих выборах, и 92 говорят, что нет, как вычислить процент тех, кто пойдет на выборы?</p>

	<p>Шаг 1. Основание = общее число опрошенных = $= 92 + 58 = 150$</p> <p>Шаг 2. Отношение = 58 из 150 = $58/150 =$ $0,386666\dots$</p> <p>Шаг 3. Процент = $0,38666\dots \times 100 = 38,666666\dots$</p> <p>Шаг 4. Округлите и упростите: $= 38,7$ процентов, или больше чем одна треть.</p>
--	--

ОТ ДРОБЕЙ К ПРОЦЕНТАМ И НАОБОРОТ

В материале намного легче понять простую дробь, нежели точный процент, поэтому привыкайте переводить дроби в проценты и наоборот.

Вы знаете, что один из четырех – это одна четверть, или 25 процентов, но не знаете, как из одного вывести другое.

ПЕРЕВОД ДРОБИ В ПРОЦЕНТЫ:

$1/4 = 1 \div 4 = 0,25$. переместите запятую на два знака вправо или умножьте на 100, чтобы получить 25%.

ПЕРЕВОД ПРОЦЕНТОВ В ДРОБИ:

В справочных материалах этой книги для более сложных чисел приведен список их дробных и процентных значений. Но если вы хотите их вычислить самостоятельно, то:

1. Напишите процент в виде дроби: $25/100$
2. Числитель и знаменатель этой дроби разделите на 25
3. Упростите: $(25/25) / (100/25) = 1/4$.

ПОЛУЧИТЬ «ЧАСТЬ ОТ ЧИСЛА»:

1. Разделите процент на 100, таким образом, 25% – это 0,25.
2. Теперь разделите число 1 на 0,25.

Вы получите $1/0,25 = 4$. Таким образом, ваш ответ: четвертая часть от числа.

Совет для пользователей электронной таблицы: Excel позволяет форматировать число в виде дроби или процента. Поэкспериментируйте, чтобы подобрать наиболее легко воспринимаемое число.

«НА ЧЕЛОВЕКА» (ИЛИ НА ДУШУ НАСЕЛЕНИЯ): УРОВЕНЬ СМЕРТНОСТИ, УРОВЕНЬ ПРЕСТУПНОСТИ И ДРУГИЕ ПОКАЗАТЕЛИ

Как и в случае с процентами, показатели «на человека» или «на душу населения» используются для сравнительного анализа. Например, когда нужно сравнить преступления в городах с различной численностью населения, смертность от различных болезней или валовой внутренний продукт между разными странами.

Отношения зачастую используются при очень больших или очень маленьких числах, чтобы сделать их более осязаемыми для восприятия.

Сложно понять слишком большие числа: «миллион» во многом звучит как «миллиард». Но как отмечает Джон Аллен Паулос в книге «Innumeracy»¹, миллион секунд проходит за 11 дней, а миллиард секунд занимает около 32 лет.

Очень маленькие числа тоже создают путаницу: 0,003 выглядит как 0,3. Оба числа трудно визуализировать.

СРЕДНЕЕ НА ДУШУ НАСЕЛЕНИЯ

Число «на человека» (на душу населения) – среднее арифметическое значение. Как правило, это значение предполагает большие объемы исследуемых объектов.

Согласитесь, трудно себе представить объем валового внутреннего продукта (ВВП) в \$17 трлн. Но если разделить объем ВВП на

¹ Книга «Innumeracy: Mathematical Literacy and Its Consequences» (авт. Джон Аллен Паулос, Нью-Йорк, 1988, 135 с.) была в списке бестселлеров The New York Times.

все население США – 317 млн. то получится приблизительно \$54 000 на душу населения. Однако, это не означает, что каждый американец зарабатывает по \$54 000 в год. На деле почти половина всех семей зарабатывает меньше этой суммы. Эта сумма – материальное благосостояние страны, созданное компаниями и людьми.

Этот расчет делает немислимые числа понятными и осязаемыми. Также, если нужно провести сравнительный анализ стран, то эти расчеты сопоставляются с размерами страны.

УРОВНИ ПРЕСТУПНОСТИ, СМЕРТНОСТИ И ДРУГИЕ РЕДКИЕ СЛУЧАИ

Отношения, такие как, например, «на 1000 человек» тоже вычисляются как проценты, но путем умножения более чем на 100, перемещением запятой вправо.

Например, если в США ежегодно умирает 2,5 миллиона человек, то процент смертности ничтожно мал и составляет 0,789 процента, или 0,00789, т.е. малую десятичную дробь.

Согласитесь, эта цифра сложна для восприятия, поэтому эксперты выражают эту цифру как 789 смертей на 100 000 человек.

ВЫЧИСЛЕНИЕ УРОВНЯ:

Шаг 1. Выберите основание. Зачастую это сложно. Чтобы рассчитать уровень смертей на дорогах по марке автомобиля, вы должны сделать выбор: взять за основание либо число проданных машин этой конкретной марки, либо общий миляж, пройденный за год.

Шаг 2. Выбор верного числителя также может быть непростой задачей. Возвращаясь к примеру автокатастроф со смертельным исходом, что вы возьмете в качестве числителя: общее число случаев с летальным исходом или же число смертей на дорогах?

	<p>Эксперты, составляющие такие отчеты, как правило, негласно руководствуются следующим правилом: использовать ту цифру, которая наиболее показательна. Разделите выбранное число на основание.</p> <p>Шаг 3: Умножьте полученную дробь, например, на 1 000, 100 000 или 1 000 000, чтобы получить целое число.</p>
ФОРМУЛА:	<p>Шаг 1. За основание выберите «общее количество» («Всего»)</p> <p>Шаг 2. (Категория / Общее количество) = Пропорция или отношение</p> <p>Шаг 3. Отношение × 1 000 = Доля на 1 000</p>
ПРИМЕР:	<p>По данным ФБР, в Соединенных Штатах за 2012 год в Питтсбурге с населением 312 000 человек было зарегистрировано 13 000 крупных преступлений против собственности, в Тусоне с населением 531000 человек – 8 870 преступлений.</p> <p>Определение уровня преступности на тысячу жителей позволяет вам сравнить эти два города:</p> <p>ПИТТСБУРГ</p> <p>Шаг 1: база = 312 000 человек</p> <p>Шаг 2: 13 000 преступлений / 312 000 человек = 0,041</p> <p>Шаг 3: 0,041 X 1 000 = 41 преступление на 1 000 человек</p> <p>ТУСОН</p> <p>Шаг 1: база = 531 000 человек</p> <p>Шаг 2: 8 870 / 531 000 человек = 0,017</p> <p>Шаг 3: 0,017 X 1 000 = 17 преступлений на 1 000 человек.</p> <p>Таким образом, уровень преступности в Питтсбурге почти в 2,5 раза больше, чем в Тусоне: 41 / 17 = 2,4</p>

ВЫБОР МНОЖИТЕЛЯ

Некоторые считают, что изменение множителя от 100 до чего-то большего – обман.

Особенно, если 0,2 процента вдруг превращается в большее число – 200, когда основание меняется со 100 на 100 000!

На практике же нет ничего магического в использовании 100 в качестве основания (аналог процентов) или же числа, которое наиболее удобно для сравнения. Вот несколько вариантов:

- a) Выберите целое число – 1 000, 1 000 000 или 100 000.
- b) Выберите то же самое число, которое используют эксперты: например, преступления на 1000 человек или уровень смертности на 100 000 человек.
- c) Выберите то основание, которое будет наиболее понятно вашим читателям: как правило, это числа от 1 до 1 000.
- d) Старайтесь не использовать очень большие основания: для цифр локального масштаба не используйте основание в 1 000 000, поскольку города с населением свыше 1 000 000 – это скорее исключение.
- e) Во всем материале используйте одно и то же основание. При представлении статистики преступности не перескакивайте от 100 000 до 1 000 при подсчете уровня преступности.

Ввиду отсутствия четких требований, эти правила поддаются компромиссу.

ДОЛИ (ПРОДОЛЖЕНИЕ)

В редакции Вам, наверняка, будет достаточно базовых знаний по вычислению процентов, дробей и множителей, полученных из предыдущего раздела. Но для некоторых вычисление доли может представлять некоторую сложность.

СЛОЖЕНИЕ И ВЫЧИТАНИЕ ДОЛЕЙ

При сложении и вычитании долей или процентов они должны быть приведены к общему основанию. То есть перед тем, как складывать, или вычитать их, они должны быть представлены как части от одного числа. Как правило, это математическое действие применимо для объединения малых групп в одну общую, например, при подведении итогов голосования.

Конечный результат выражается в процентных пунктах, а не процентах: в 2012 году в ходе всенародного голосования Барак Обама получил на 4 процентных пункта больше, нежели его соперник Митт Ромни.

ОТНОШЕНИЕ ДОЛЕЙ, ИЛИ «ОТНОСИТЕЛЬНЫЙ РИСК»

Как и в примере со статистикой по преступности, порой эффективно сравнить две аналогичные доли: «Афроамериканцам отказывают в получении ипотечного кредитования в два раза чаще, нежели белым с тем же уровнем финансового дохода».

Это пример «относительного риска».

Это понятие заимствовано из медицинских исследований: риск заражения тем или иным заболеванием среди курящих и некурящих. «Относительный риск» зачастую отражает точную картину происходящего.

Это наиболее очевидно на примере с небольшими числами. Если 20 из 1 000 курящих заболели раком, то доля заболевших раком среди курящих составляет 0,02; если же 10 из 1000 некурящих заболели раком, то доля заболевших раком среди некурящих составляет 0,01. Тогда относительный риск среди курящих = 2, т.е. возможность заболевания раком среди курящих в два раза больше, чем у некурящих. Относительный риск среди некурящих составляет 1/2, но не на 200 процентов меньше, чем среди курящих.

Относительный риск вычисляется в отношении «группы риска», а не наоборот.

Следите за языком изложением этих цифр. У курильщика – двойной риск заболеть раком, но не «в два раза больший риск». Если же вы хотите сказать «больше, чем», или «меньше, чем», то необходимо вычислить процентную разницу, а это предполагает дополнительное математическое действие.

Как правило, редкие события провоцируют высокие относительные риски. Так, в материале по медицинской тематике принято использовать отношения, отражающие двойные или тройные относительные риски.

Для больших же чисел – или более обыденных событий, как, например, вычисление шансов на получение ипотечного кредита – малые относительные риски также значимы.

<p>КАК ВЫЧИСЛИТЬ ОТНОСИТЕЛЬНЫЙ РИСК:</p>	<p>Шаг 1. Определите процент людей, находящихся «под риском» по каждой группе. Например, процент отказа в ипотечном кредитовании среди афроамериканцев или аналогичный процент для белых.</p> <p>Шаг 2. Разделите два числа. В числителе должна стоять группа, по которой вы делаете материал. Например, в нашем случае числитель – «группа риска» афроамериканцев, знаменатель – аналогичный процент для белых.</p> <p>Шаг 3: Прежде чем производить какие-либо расчеты, определитесь, как представить числовое значение: как отношение («скорее всего, как») или как процентная разница («более вероятно»).</p>
<p>ФОРМУЛА:</p>	<p>Шаг 1. (Число людей «под риском»/ Итого в группе 1), (Число людей «под риском»/ Итого в группе 2)</p> <p>Шаг 2. Разделите два результата: (Результат группы 1/Результат группы 2)</p>

	<p>Шаг 3. Преобразуйте в проценты, если результат меньше чем 2; выразите в виде дроби, если результат больше чем 2.</p> <p>Если вы преобразовываете как разницу в процентах, выразите как «процент больше (или меньше) вероятнее, чем»: (относительный риск – 1) x 100.</p>
<p>ПРИМЕР:</p>	<p>В г. Нэшвилл 40 из 120 афроамериканцев с высоким уровнем дохода было отказано в предоставлении ипотечного кредита; 300 из 2 400 «белых» соискателей с тем же уровнем дохода было отказано в предоставлении кредита.</p> <p>Шаг 1. Вычислите показатели отказа.</p> <p>Для афроамериканцев: $40/120 = 0,33333$, или 33,3 процента, или 1/3.</p> <p>Для «белых»: $300/2\ 400 = 0,125$ или 12,5 процента, или 1/8.</p> <p>Шаг 2. Выясните относительный риск: $0,333/0,125 = 2,664$</p> <p>Другими словами, «Афроамериканцы получают отказ в кредитовании в 2,5 раза чаще нежели «белые»».</p> <p>Шаг 3: Перефразируйте, чтобы изложить информацию как «более вероятно» для вычисления «изменения процентной разницы» без умножения на 100. (Для подсчета процентной разницы обратитесь к главе «Вычисление изменений»): $(33,3 - 12,5) / 12,5 = 1,664$.</p> <p>Другими словами, вероятность отказа в предоставлении ипотечного кредита среди афроамериканцев в 1,7 раза выше, чем среди «белых»</p>

ЗАМЕЧАНИЯ ПО ПРОВЕРКЕ ЗНАЧИМОСТИ

Журналисты часто задаются вопросом, каким образом можно определить значительность результата, исходя из неинформативных показателей. Например, если полиция применяет силу в отношении арестованных афроамери-канцев в 6 процентах случаев, а в отношении «белых» – в 5 процентах случаев, то насколько значительна наблюдаемая разница?

Возможно, это и не так уж важно.

Во-первых, журналисты редко непосредственно работают с выборками случайных чисел. Статистическая значимость важна при проведении голосования или в социологических опросах, где респонденты выбираются посредством выборки.

В приведенном выше примере большинство журналистов отразят найденные данные как любое, а не каждое десятое по счету применение сил, исходя из того, что полученная информация действительно отражает реально происходящие события. Нет смысла анализировать найденную информацию по отношению ко всей вселенной. Просто используйте то, что вы знаете. В примере с полицией достаточно утверждения, что в городе «Х» в таком-то году полиция применяла силу в отношении афроамериканской молодежи на 20 процентов чаще чем в отношении к «белым» – не нужно обобщать полученную информацию на другие города и другие года.

Но статистические тесты все же важны для того, чтобы исключить долю чистой случайности. Они также позволяют понять, насколько правильно вы расставляете акценты и не зацикливаетесь на незначимых мелочах.

«Значительный» далеко не означает «имеющий значение», а значит, достойный освещения в новостях. Это всего-навсего означает, насколько часто такие различия имеют место быть по чистой случайности. Ввиду того, какими цифрами зачастую оперируют журналисты по роду своей профессиональной деятельности, вполне вероятно, что итоговые цифры окажутся статистически значимыми,

даже если и не несущими информационного значения. Касательно нашего примера по применению силы полицией: этот пример отражает редкое, не достойное освещения в новостях событие, приводящее к сравнительно небольшим относительным рискам, хотя само значение может оказаться статистически значимым и верным.

Причина в том, что главное число χ^2 , которое отражает насколько разность между эмпирическими и теоретическими долями – результат случайности, на деле зависит от двух аспектов: разность между самими долями и сколько фактически человек приходится на эту разницу. В новостях, как правило, практически всегда за основу берется большое число, поэтому даже незначительные расхождения между коэффициентами могут показаться значимыми.

Это пособие не ставит целью научить вычислять значение критерия «хи-квадрат». Для информации вы можете ознакомиться с пособием Филипа Мейера «Новая журналистике точности» / «New Precision Journalism». При освещении частных случаев, как аварии с летальным исходом, перестрелки с участием полицейских, а также статистика по заболеванию раком, лучше использовать распределение Пуассона, чем «хи-квадрат» распределение. Распределение Пуассона позволяет просчитать вероятность, с какой частотой то или иное событие может произойти по чистой случайности. (Одним из первых эту формулу применил прусский генерал, который хотел узнать, были ли случайными смертельные исходы от ударов лошади за единицу времени или же неслучайными.) На деле распределение Пуассона применимо при большом объеме выборки, но при этом число случаев, где событие фактически имело место, мало и дискретно.

ВЫЧИСЛЕНИЕ «ИЗМЕНЕНИЯ»

Материалы новостей зачастую отражают изменения времени или места.

ПРОСТЫЕ ИЗМЕНЕНИЯ (ИЛИ ФАКТИЧЕСКИЕ ИЗМЕНЕНИЯ)

Простая разница – это результат вычитания одного числа из другого. Если вы измеряете разницу во времени, то это разность между «новым значением» и «старым значением».

Это применимо в тех случаях, когда вероятность того, что событие произойдет, фактически мала, но существует всеобщее понимание того, что это событие, по сути, не должно произойти. Недавние примеры включают в себя 32 аварии из-за поломки в ключе зажигания у автомобилей марки «General Motors» или 64 смертельных случая в 2012, которые согласно анализу Центров по контролю и профилактике заболеваний произошли в результате ошибочного применения медикаментов.

Другой случай использования «чистого» значения разницы, когда его применение очевидно без каких-либо расчетов. Цены предметов домашнего обихода, суммы зарплат и цены на жилье – наиболее очевидные примеры чисел, для которых далеко не обязательно подсчитывать изменения в процентном соотношении.

В конце концов, наша работа – это подготовка материалов для новостей, поэтому в некоторых случаях порой целесообразнее использовать само числовое значение, нежели его производное значение (например, процентное изменение).

ВЫЧИСЛЕНИЕ «ИЗМЕНЕНИЯ»	Вычтите последующее число из предыдущего. Это не одно и то же, что вычитание меньшего числа из большего. Если последующее число меньше предыдущего, результат будет отрицательным. Если же последующее число больше предыдущего, то результат будет положительным.
ФОРМУЛА:	Последующее число – предыдущее число
ПРИМЕР:	В прошлом году руководитель компании заработал \$2,4 миллиона; этом году – \$2,9 миллиона.

	Его прибавка к зарплате составила: $\$2,9 - \$2,4 = \$0,5$ миллиона, или \$500 000, или полмиллиона долларов США.
--	--

**ОТНОСИТЕЛЬНОЕ ИЗМЕНЕНИЕ, ИЛИ
ОТНОСИТЕЛЬНАЯ РАЗНИЦА В ПРОЦЕНТАХ**

Это, пожалуй, одна из самых горячих тем для споров в отделах новостей.

Часть проблемы состоит в том, что существует пять или шесть различных способов вычисления относительного изменения в процентах. Однако, на деле безотказно работают только два способа, но я продемонстрирую вам самый легкий из них для запоминания.

Обратите внимание, что в нашем случае не имеет значения, какое число больше или меньше. Если число уменьшилось, то относительная разница в процентах будет отрицательной и наоборот. В любом случае результат является безошибочным, если даже изменение превышает 100 процентов.

ВЫЧИСЛЕНИЕ ПРОЦЕНТНОЙ РАЗНИЦЫ:	<p>Шаг 1. Вычтите из последующего значения предыдущее значение, при этом неважно, какое из этих чисел больше или меньше!</p> <p>Шаг 2. Разделите результат на «предыдущее значение».</p> <p>Шаг 3. Умножьте полученный результат на 100 или поставьте запятую на два знака вправо.</p> <p>Шаг 4. Округлите и упростите конечный результат.</p>
ФОРМУЛА:	<p>Шаг 1. Последующее – предыдущее = Разница</p> <p>Шаг 2. Разница / предыдущее = десятичное число</p>

	<p>Шаг 3. Десятичное число $\times 100$ = Процентная разница</p> <p>Шаг 4. Округлите конечный результат.</p>
ПРИМЕР:	<p>В прошлом году руководитель компании заработал \$2,4 миллиона, а в этом году – \$ 2,9 миллиона.</p> <p>Шаг 1. Разница = $\\$2,9 - \\$2,4 = \\$0,5$ миллиона.</p> <p>Шаг 2. Разница / предыдущее число = $\\$0,5 / \\$2,4 = 0,208$</p>
	<p>Шаг 3. Переместите запятую вправо на две цифры = 20,8%</p> <p>Шаг 4. Округлите и упростите:</p> <p>$21\% = 21 / 100 =$ приблизительно</p> <p>$20 / 100 = 1/5$.</p> <p>Итак, руководитель получил повышение зарплаты, эквивалентной одной пятой его первоначальной зарплаты</p>

ФАКТИЧЕСКОЕ ИЗМЕНЕНИЕ ПРОТИВ ОТНОСИТЕЛЬНОГО ИЗМЕНЕНИЯ В ПРОЦЕНТАХ

В редакциях часто возникает спор: какая мера вычисления наиболее верна – фактическое изменение либо относительное изменение в процентах? На деле – это скорее редакторское решение, нежели четкое правило.

Предлагаем вам ознакомиться с четким правилом из пособия «Статистический обман на работе» / «Statistical Deception at Work»: Проценты можно сравнивать лишь в том случае, когда у них одинаковые основания.

Большинство журналистов, которые когда-либо писали на экономическую тематику, перефразируют это правило по-своему: когда числа являются различными по масштабу, единственный способ их сравнить – проценты.

Таким образом, какое числовое значение использовать в материале новостей – фактическое изменение или относительное изменение в процентах?

Ответ: Используйте то числовое значение, которое наиболее значимо и заслуживает внимания в СМИ. Это далеко не означает, что это наибольшее число. Это скорее то число, которое именно значимо, нежели то, что вызывает тревогу. Вы должны четко для себя определить, какое число использовать: абсолютные разности, относительные изменения или что-либо еще.

В качестве примера давайте рассмотрим типичный материал новостей, публикуемый в СМИ время от времени: зарплаты и бонусы руководителей крупных компаний.

Во Флориде в 1990-е годы, когда я писал этот материал, многие генеральные директора частных компаний, будучи их собственниками, преднамеренно платили себе зарплату по заниженной ставке. Особенно на первых порах некоторые руководители получали сначала \$50 000 в год, но по мере того как бизнес набирал обороты их зарплата возрастала до \$100 000. В то же время руководители национальных компаний получали зарплату по \$2 000 000 в год, при этом из года в год их зарплата повышалась еще на \$200 000.

Что, на ваш взгляд, заслуживает освещения в СМИ: увеличение зарплаты руководителя частной компании в два раза или увеличение зарплаты руководителя нацкомпании на \$200 000?

Вопрос спорный. Это зависит от того, насколько прибыльна сама компания и на сколько зарплата первого руководителя отличается от других сотрудников.

Зачастую, во Флориде центром внимания СМИ был сам зарплатный пакет в \$2 миллиона, нежели факт повышения заработной платы первых руководителей.

ОБРАТНЫЙ ПОДСЧЕТ ОТНОСИТЕЛЬНЫХ ИЗМЕНЕНИЙ В ПРОЦЕНТАХ

Помните, что хотя число может вырасти во много раз, но уменьшиться оно может только со 100 процентов до нуля. Другими словами, если стоимость товара увеличивается с \$20 до \$40, то рост стоимости составляет 100 процентов; тройное увеличение стоимости будет равно \$60. Но если вы уменьшите стоимость с \$40 до \$20, то понижение стоимости товара составит 50 процентов (или в 1/2 раз меньше), а не 100 процентов. Другими словами, равные абсолютные разности не означают столь же равных относительных изменений, то есть процентные изменения напрямую не поддаются обратному подсчету.

Это означает, что рекламные объявления, которые утверждают, что вы будете использовать в три раза меньше мощного средства либо тот или иной продукт содержит в три раза меньше соли, по сути, являются ошибочными. Производители товаров, вероятно, имеют в виду следующее: при использовании продукта другой марки, вы затратите объем в три раза больший, чем объем рекламируемого продукта, или же объем рекламируемого продукта составляет треть от объема товара конкурентов. Давайте рассмотрим два примера:

<p>СОПОСТАВЛЕНИЕ ОТНОСИТЕЛЬНЫХ ИЗМЕНЕНИЙ В ПРОЦЕНТАХ</p>	<p>Если годовой бюджет полицейского участка сначала возрос на 15 процентов, затем понизился на те же 15 процентов, это не означает возвращения к исходному значению. Объем сниженного годового бюджета будет гораздо меньше.</p>
<p>ПОЧЕМУ ТАК ПРОИСХОДИТ?</p>	<p>Шаг 1. Один миллион долларов США увеличивается на 15 процентов:</p> $\begin{aligned} & \$1 \text{ миллион} + (15\% \text{ от } \$1 \text{ миллиона}) = \\ & \$1 \text{ миллион} + (\$1\,000\,000 \times 0,15) = \\ & \$1 \text{ миллион} + \$150\,000 = = \$1\,150\,000. \end{aligned}$

	<p>Шаг 2. Снижение \$1 150 000 на 15 процентов, или $\\$1\,150\,000 - (15\% \text{ от } \\$1\,150\,000) = \\$1\,150\,000 - \\$172\,500 = \\$977\,500$.</p> <p>Это объясняется тем, что процент вычисляется от большего числа.</p>
<p><i>ВОЗВРАЩАЯСЬ К ТОМУ ЖЕ САМОМУ УРОВНЮ, С КОТОРОГО ВЫ НАЧАЛИ:</i></p>	<p>Число, которое сначала увеличилось на одно процентное значение, затем уменьшилось на то же самое значение, не будет равно своему исходному значению. Число большее нуля не может уменьшиться более чем на 100 процентов, но в то же время число может возрасти на любое процентное значение.</p> <p>Шаг 1: \$1 миллион утраивается до \$3 миллионов. Процентное изменение $(\\$3 - \\$1) / \\$1 = \\2, или 200-процентное увеличение. (Мы часто выражаем это как трехкратное увеличение вместо того, чтобы вычислить процентное изменение).</p> <p>Шаг 2. Уменьшение значения от \$3 миллионов до \$1 миллиона.</p> <p>Процентное изменение $(\\$1 - \\$3) / \\$3 = -0,667 = -67$ процентов, или снижение исходного значения на две трети.</p>

Фактически невозможно вычислить относительное изменение в процентах для значения, которое меняет знаки. Такой расклад часто встречается в финансовых отчетах компаний, когда в течение одного года компания понесла убытки, а в следующем году сделала прибыль. Такой расчет не является допустимым.

Предположим, если доход возрос с минус \$200 до \$100, то теоретически увеличение дохода должно составить 300 процентов, но подобный подсчет неверен. Те же 300 процентов вы можете получить при увеличении дохода с \$25 до \$100, что является единственным верным значением. Считая в обратную сторону, вы получите минус 300 процентов изменения, что является также ошибочным.

РАЗНИЦА В ПРОЦЕНТНЫХ ПУНКТАХ ПРОТИВ ОТНОСИТЕЛЬНОЙ РАЗНИЦЫ В ПРОЦЕНТАХ – ОДНО СЛОВО: БОЛЬШОЕ РАЗЛИЧИЕ

Выражение числовых значений в процентах порождает споры и разногласия в редакциях. Давайте рассмотрим два примера, которые математически верны.

- Прирост численности населения замедлился с момента демографического пика в 1950 году примерно на 1,3 процентных пункта до 0,7 процента. (Простое изменение между 2 процентами и 0,7 процента, выраженное в процентных пунктах).
- Прирост населения в США замедлился почти на две трети по сравнению с 1950 годом, снизившись до 0,7 процента в 2013 году. (Процентная разница между 2 и 0,7 процентами).

А.К. Дьюдни, автор книги «200 процентов из ничего» («200% of Nothing» by A.K. Dewdney), считает последний пример «преднамеренным повышением процентного значения». Он считает это законным способом того, чтобы незначительные изменения выглядели значительными.

Эксперты в области медицины и инвестиций придерживаются другого мнения. В течение многих лет нас учили, что неправомерно считать относительную разницу в процентах между двумя процентными значениями. Однако, если вы вернетесь к предыдущему разделу, то вы поймете, что это именно то, что происходит при вычислении «относительного риска». Логарифмы и производные в математическом анализе позволяют выразить ту же идею, но другими способами.

В двух случаях вычисление относительной разности в процентах между процентами или процентного отношения может помочь прояснить число, нежели ввести в заблуждение.

Во-первых, в том случае, когда сравниваемые показатели находятся на различных уровнях. Например, в 1980 году показатель уровня тяжкой преступности в штате Мэриленд практически в три раза превышал уровень тяжкой преступности в штате Вирджиния. При сравнении тенденций в обоих штатах, вполне очевидно, что ситуация в Вирджинии не может кардинально улучшиться по сравнению с Мэрилендом, если брать за основу наших расчетов процентные пункты. Уровень преступности в Мериленде упал больше, нежели общий уровень преступности в штате Вирджиния.

Во-вторых, в тех случаях, когда речь идет о редких событиях, как в случае работы с относительными рисками.

Правильный ответ зависит от Вас. Вы можете акцентировать внимание читателей на присутствии или на отсутствии изменения. Но, принимая решение, исходите из идеи и смысла материала, а не из более значительных и тревожных показателей. Также убедитесь в том, что использованная лексика правдиво отражает то, что должно быть описано.

ПРОДОЛЖАЯ АНАЛИЗ ИЗМЕНЕНИЙ: ГODOVЫЕ ПОКАЗАТЕЛИ И КОРРЕКТИРОВКА ПОКАЗАТЕЛЕЙ С УЧЕТОМ ИНФЛЯЦИИ

ГODOVЫЕ ПОКАЗАТЕЛИ

Мы редко обращаемся к годовым показателям в отделах новостей, но их подсчет может оказаться полезным, особенно в том случае, когда вы хотите удостовериться в точности тех цифр, которые вам предоставляют другие.

Еще в 1996 году кандидат в вице-президенты США Джек Кемп утверждал, что предлагаемый им единый налог и сокращение ряда других налоговых выплат обеспечит рост экономики страны на 5 процентов в год. В качестве сравнения он взял 60-процентный рост эко-

номики с 1950 по 1960 годы. Основываясь на этом примере, он аргументировал, что предсказываемый им результат вполне достижим.

Расчеты Кемпа, возможно, основывались на интуитивно объяснимой логике вычисления 60 процентного изменения за 10 лет. А логический вывод среди многих американцев, в том числе и журналистов, заключался в том, что годовой рост процентной ставки должен составить 6 процентов.

Но на деле все обстоит иначе: при вычислении изменений на протяжении долгого периода времени важно просчитать годовые процентные показатели. Эффект наращивания процентов работает по тому же принципу, который позволяет вашей изначальной сумме на депозите значительно расти с течением времени, если даже ваш первоначальный взнос невелик.

Существует две ситуации, в которых вам необходимо будет подсчитать изменение показателя в годовом исчислении.

Во-первых, в том случае, когда вам нужно спроецировать влияние последних изменений на будущую динамику развития ситуации.

Во-вторых, при подсчете того, каким образом изменение в течение продолжительного периода времени отражается на средних годовых показателях. Эти вычисления основываются на едином принципе, который включают в себя степенные вычисления. Если вы не хотите вникать в эти вычисления на данный момент, вы можете пропустить этот параграф и вернуться к нему, когда вам будет необходимо.

ПРОГНОЗИРОВАНИЕ ГОДОВЫХ ПОКАЗАТЕЛЕЙ

ПРОЕЦИРОВАНИЕ ТЕКУЩИХ ИЗМЕНЕНИЙ В БУДУЩЕМ

При проецировании значений на много лет вперед, несущественные изменения могут быть представлены как более значительные. Подобным принципом могут воспользоваться и ваши источники. Вот в чем он заключается:

	<p>Шаг 1. Разделите текущее значение на прошлогоднее значение. Вы получите результат больше единицы, если текущее значение возросло, и меньше единицы, если оно уменьшилось. Назовем полученный результат коэффициентом.</p> <p>Шаг 2. Возведите коэффициент в степень, равную количеству прогнозируемых лет. Если количество прогнозируемых лет равно 10, то возведение в степень равносильно умножению коэффициента на самое себя 10 раз.</p> <p>Шаг 3. Для вычисления будущего значения умножьте полученный результат на исходное значение и переходите к дальнейшим необходимым вам вычислениям.</p>
<p>ФОРМУЛА:</p>	<p>Шаг 1: (Новое значение/Предыдущее значение) = Коэффициент</p> <p>Шаг 2: Коэффициент ^{Годы} = Множитель</p> <p>Шаг 3: Множитель × Предыдущее значение = Будущее значение</p> <p>Шаг 4: Используйте полученный результат, чтобы произвести остальные расчеты: например, относительное изменение в процентах с настоящего момента до желаемого момента в будущем.</p>
<p>ПРИМЕР:</p>	<p>Бюджет небольшого города вырос с \$144,3 миллионов до \$149,6 миллионов за прошлый год. Насколько может увеличиться бюджет города за 10 лет при условии сохранения текущего темпа роста?</p> <p>Шаг 1: (Новое значение/Предыдущее значение) = \$149,6/\$144,3 = 1,03672...</p> <p>Это эквивалентно возрастанию на 3,7 процента.</p>

	<p>Шаг 2: Возведите результат в 10 степень = $(1,03672)^{10} = 1,43434\dots$</p> <p>Чтобы возвести число в степень на калькуляторе, найдите кнопку, которая выглядит как «y^x» Введите число, которое вы хотите возвести в степень (1,03672, как в этом примере), нажмите эту кнопку и введите количество лет (например, 10).</p> <p>Чтобы возвести число в степень в электронной таблице Excel, используйте символ «^»: $=1,03672^{10}$.</p>
	<p>Языки программирования часто используют в этом случае «две звездочки» (**). Операция возведения в степень является одной из базовых операций, используемых в любом языке программирования. Если (**) не срабатывают, обратитесь к документации по языку.</p> <p>Шаг 3: Умножьте полученный результат на прошлогоднее значение:</p> <p>$1,43434 \times \\$144,3 = \\207 миллионов</p> <p>Шаг 4: Произведите другие вычисления, например, подсчет относительного изменения в процентах:</p> <p>$(\\$207 - \\$144,3) / \\$144,3 = 0,43$, или 43 процентное повышение.</p>

ПЕРЕСЧЕТ ИТоговых Изменений на годовые показатели

Нилл Боровски (Neill Borowski), бывший директор департамента «Отчетности и анализа при помощи компьютерных технологий» газеты «The Philadelphia Inquirer», использовал данные по годовым процентным ставкам в местном материале по стоимости недви-

мости, несмотря на то, что он не смог получить надежную информацию по этим ставкам через реестр отчислений налогов на собственность. Вместо этого он сначала определил, насколько увеличилась цена каждого дома при его продаже, а затем, исходя из полученного результата, вычислил среднюю ежегодную процентную ставку стоимости жилья.

Работая таким же образом в обратную сторону, можно проанализировать пример Джека Кемпа (Jack Kemp) (см. ниже). В целом, навык перевода любых итоговых значений за многолетний период в ежегодные процентные показатели может оказаться вам полезен. Полученные в результате значения часто кажутся менее впечатляющими.

<p>ВЫЧИСЛЕНИЕ ЕЖЕГОДНЫХ ПОКАЗАТЕЛЕЙ:</p>	<p>Шаг 1: Разделите новое значение на предыдущее значение либо преобразуйте уже известное процентное изменение в коэффициент, разделив его на 100 и прибавив 1.</p> <p>Шаг 2: Возведите полученный результат в степень (1/число лет). На деле это действие невозможно заменить более простым математическим действием, поскольку необходимо взять корень n-ой степени из полученного числа. (Другими словами, какое число, умноженное само на себя 10 раз, даст вам искомое значение?)</p> <p>Шаг 3: Преобразуйте результат обратно в процентное значение, вычитая 1 и перемещая запятую вправо.</p>
<p>ФОРМУЛА:</p>	<p>Шаг 1: Общий Коэффициент = (Новое значение/Предыдущее значение), или</p> <p>Общий коэффициент = (Общее процентное изменение/100) + 1.</p>

	<p>Шаг 2: Годовой Коэффициент = Общий коэффициент^(1/Годы)</p> <p>Шаг 3: Годовая процентная ставка = = (1 – Годовой Коэффициент) * 100</p>
ПРИМЕР:	<p>Используя пример Джека Кемпа, вы можете произвести один из двух следующих расчетов:</p> <p>Шаг 1: ВВП за 1968 год / ВВП за 1958 год = = (\$3293,9 / \$2057,5) = 1,601, или ((Процентное изменение с 1958 до 1968 / 100) + 1) = = (60,1 / 100) + 1 = 1, 601</p>
	<p>Шаг 2: Возведите в 1/10 степень = = 1,601 (1/10) = 1,048</p> <p>Шаг 3: Вычтите 1 и переместите запятую = (1,048 – 1) = 0,048, или годовая ставка 4,8 процента.</p> <p>Таким образом, даже в период наибольшего экономического роста в современной истории американская экономика не соответствовала темпам роста, предложенным Кемпом. И это еще без учета активного темпа роста населения, послевоенного роста экономики и других непредвиденных факторов, которые, как правило, не повторяются.</p>

КОРРЕКТИРОВКА С УЧЕТОМ ИНФЛЯЦИИ

При сравнении двух денежных величин на протяжении длительного периода времени необходимо, как правило, производить корректировку этих величин в соответствии с уровнем инфляции.

Большинство экспертов и журналистов солидарны во мнении, что без учета инфляции несправедливо сравнивать зарплаты в 2000-х

годах с зарплатами в 1970-х годах. Другие величины менее однозначны. Например, в государственных финансах сокращения бюджета допустимы после инфляции (что может служить признаком незначительного увеличения фактической величины), или перед инфляцией (что соответствует абсолютному снижению в денежных единицах). При этом политические предпочтения источников могут повлиять на сделанный выбор, поэтому уточнение подобных аспектов обязательно.

Журналисты должны использовать значения, сопоставимые с контекстом материала. Но если вы сравниваете величины на протяжении более чем пятилетнего периода, то в этом случае принято производить корректировку с учетом инфляции или ожидаемой инфляции. При корректировке непременно принимайте во внимание все параметры, которые вы считаете неординарными, например, небывало высокие цены на бензин.

Пусть замечания по спорности Индекса Потребительских Цен (ИПЦ) (Consumer Price Index, CPI) не останавливают Вас от использования этого простого инструмента для проверки себестоимости денег. Данные, предоставляемые ИПЦ, все же лучше, чем полное отсутствие данных. Кроме того, большинство альтернатив этому индексу выводятся из ИПЦ и могут содержать те же ошибки.

Для большей точности вы можете обратиться к индексам инфляции, которые предоставляет Бюро Экономического Анализа при Департаменте Торговли США. Эти индексы носят название Дефляторы ВВП (Implicit Price Deflators) и вычисляются посредством комбинации подробных результатов ИПЦ и величин, применяемых для вычисления ВВП. Наиболее близким аналогом к ИПЦ среди дефляторов является дефлятор «потребительских расходов». Другой индекс измеряет изменение цен государственных расходов, и применим при работе с городскими бюджетами.

Эти индексы зачастую считаются наиболее консервативными оценками инфляции, поскольку они разрешают некоторые спорные моменты, связанные с вычислением ИПЦ. Однако не стоит вдаваться в подробности. Все государственные оценки инфляции сходны в долгосрочной перспективе.

Большинство отделов новостей публикуют результаты за текущий год. Экономисты же чаще всего берут за основу «базовый год» индекса. В качестве оправдания нарушения канонов академической среды служит тот факт, что наши читатели скорее знают себестоимость доллара на сегодняшний день, нежели понимают его значение пять или десять лет назад.

Выше рассмотренные вычисления зачастую называют «реальными долларами» или «постоянными долларами». Например, Бюро Экономического Анализа (Bureau of Economic Analysis) использует такое понятие, как «реальный ВВП». В этом контексте «реальный» или «постоянный» означает, что предоставленные данные уже были скорректированы с учетом инфляции.

В статистических отчетах данные могут быть выражены, например, в «миллиардах цепных (фиксированных) индексов за 2009 год» или похожим образом. Это означает, что предоставленные данные также были скорректированы с учетом уровня инфляции, но подсчитаны они в соответствии с себестоимостью денег в 2009 году, а не на сегодняшний день.

Если же вам встретилась фраза «текущие доллары», то вам необходимо будет произвести корректировку данных с учетом инфляции самостоятельно.

**КАК ПРОИЗВЕСТИ
КОРРЕКТИРОВКУ С
УЧЕТОМ УРОВНЯ
ИНФЛЯЦИИ**

Для корректировки с учетом уровня инфляции вам необходимы следующие данные: цены первого периода (сейчас), цены второго периода (тогда) и значение величины в прошедшем периоде (тогда).

Многие используют Индекс Потребительских Цен для Всех Городских Потребителей (который включает в себя цены товаров без сезонной корректировки), предоставляемый Бюро Трудовой Статистики США, за самый последний имеющийся в наличии базовый период.

	<p>На веб-сайте Бюро Трудовой Статистики (www.bls.gov/data) есть калькулятор вычисления уровня инфляции. Для получения исторических данных (с целью применения расчета к выборке значительного размера) обратитесь к разделам сайта «Data Tools» или «Prices and Inflation» и найдите данные, обозначенные как «высшие точки».</p> <p>Шаг 1: Разделите ИПЦ последнего года на ИПЦ того года, который содержит ранние данные</p> <p>Шаг 2: Умножьте полученный результат на долларовой показатель за тот год, который вы хотите скорректировать</p>
<p>ФОРМУЛА:</p>	<p>Шаг 1: Соотношение между ценами сейчас к ценам тогда = (ИПЦ Сейчас / ИПЦ Тогда)</p> <p>Шаг 2: Умножение на предыдущее значение величины:</p> <p>Полученный Результат X Значение величины Тогда</p> <p>В усложненном варианте за основу берется другой период в качестве базового периода.</p> <p>Таким образом: Скорректированное значение = (ИПЦ Год / ИПЦ Базового Года) X Значение величин в соответствующие годы. Вам нужно будет произвести подобные расчеты, как с предыдущими, так и с текущими числовыми значениями, чтобы получить сопоставимые величины.</p>
<p>ПРИМЕР:</p>	<p>В 1970 году учителя зарабатывали около \$9 000 в год. В 2012 году их ежегодный доход составил \$54 000. Насколько значительно это увеличение?</p> <p>Шаг 1: Возьмите соотношение дохода 2012 года к доходу 1970 года, используя Индекс Потребитель-</p>

	<p>ских Цен для Всех Городских Потребителей для всех товаров и всех городов. Это соотношение составит $229,6 / 38,8 = 5,9175$.</p> <p>Это означает, что то, что стоило \$6 в 2012 году, в 1970 году можно было приобрести за \$1.</p> <p>Шаг 2: Используйте этот результат по отношению к первоначальной зарплате, умножив: $5,9175 \times \\$9000 = \\$53\,257$.</p> <p>Это означает, что с учетом уровня инфляции зарплата учителя выросла на \$54000 – \$53257, или \$750. Таким образом, за 26 лет, при учете уровня инфляции, зарплата учителей выросла совсем не намного.</p> <p>Заметьте, что все цифры отражают значения сегодняшних показателей, поэтому они максимально понятны вашим читателям.</p>
--	--

ПОНИМАНИЕ СРЕДНИХ – СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ, МЕДИАНА, МОДА

Средние величины передают сущность процесса. Если одной цитатой можно передать сущность события, одним коротким рассказом можно описать человека, использующего действия вместо слов, при помощи средних величин можно передать различные аспекты человеческой жизни – например, движение финансовых потоков, пробки на дорогах, смертность или заболеваемость – одним числом.

От тщательного выбора средней величины или любой другой величины, которая сможет отразить суть ситуации, может зависеть точность вашей передачи происходящего.

На самом деле, понимание различных типов средних величин и того, какую информацию они содержат или не содержат, является материалом базового курса статистического анализа.

Если средняя величина недостаточно хорошо передает сущность вашей информации, то не имеет смысла заниматься дальнейшим анализом. Существует три основных вида средних величин.

Среднее арифметическое – это число, которое вычисляется посредством сложения всех членов числового ряда и деления полученной суммы на число ее членов. Медиана – это число, находящееся ровно посередине выборки; журналисты часто называют его «типичным числом». Мода – это число, которое наиболее часто встречается в выборке.

СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ, ИЛИ СРЕДНЕЕ

Среднее арифметическое, также называемое арифметическим средним, чаще всего встречается в новостях. Оно наиболее информативно в том случае, когда оно усредняет значения, которые не слишком сильно варьируют внутри интервала выборочных данных. В противном случае использование среднего арифметического может ввести в заблуждение, например, когда речь идет о величинах, измеряемых в денежных единицах, как доходы или стоимость недвижимости и тому подобное (см. пример ниже).

ВЫЧИСЛЕНИЕ СРЕДНЕГО АРИФМЕТИЧЕСКОГО:	Шаг 1: Сложите все числа. Шаг 2: Разделите полученный результат на количество чисел, которые вы сложили.
ФОРМУЛА ДЛЯ СРЕДНЕГО АРИФМЕТИЧЕСКОГО:	Шаг 1: Сумма чисел Шаг 2: Сумма / Количество чисел Для пользователей электронной таблицы Excel: = AVERAGE (перечень чисел)
ПРИМЕР:	Ниже приведена информация по пяти объектам недвижимости, расположенных в одном районе:

	\$75 000
	\$75 000
	\$95 000
	\$100 000
	\$350 000
	Шаг 1: $\$75\,000 + \$75\,000 + \$95\,000 + \$100\,000 + \$350\,000 = \$695\,000$
	Шаг 2: $\$695\,000 / 5 = \$139\,000$. Таким образом, средняя стоимость жилья превышает стоимость всех объектов в списке кроме одного.

МЕДИАНЫ, ИЛИ СРЕДИННЫЕ ЗНАЧЕНИЯ

Медианы часто используются при расчетах средних величин объектов, измеряемых в денежных единицах, в особенности стоимости жилья и доходов. Причина их частого применения заключается в том, что они не чувствительны к необычайно высоким или низким значениям внутри выборки в отличие от среднего арифметического значения из предыдущего примера.

Однако рассчитать медиану сложнее, поскольку для этого необходим перечень (упорядоченный) всех значений выборки. Например, если вам известен совокупный доход жителей всего города и число людей, там проживающих, то вы можете вычислить средний уровень дохода или сумму дохода на душу населения, но не медиану.

Если же вам необходимо значение медианы, то вам надо будет либо немного разобраться в журналистике данных, либо получить эту информацию из соответствующих источников.

Иногда при работе с медианой ее называют «среднестатистическим» числовым значением или также «срединным» значением.

**ВЫЧИСЛЕНИЕ
МЕДИАНЫ:**

Чтобы вычислить медиану, сначала расположите все числа выборки по порядку, начиная с наименьшего числа и заканчивая наибольшим числом.

Посчитайте число значений в выборке и разделите его на 2.

Далее прибавьте к полученному числу 0,5. Если в сумме получается целое число (например, 13), то пересчитайте выборку до числа, с порядковым номером равным этому числу. Найденное таким образом число будет являться медианой выборки.

Если полученное число нецелое (например, 12,5), то возьмите среднее значение от двух чисел в выборке, порядковые номера которых расположены справа и слева от полученного числа. Например, если в выборке 24 числа, то, соответственно, полученное число составляет $24 / 2 + 0,5 = 12,5$. Поэтому для получения медианы необходимо рассчитать среднее арифметическое чисел, расположенных в выборке под номером 12 и 13.

Другими словами, при подсчете медианы находятся значения, которые расположены ближе всего к середине или в середине выборки. Подсчет основан на принципе сортировки, и не требует калькулятора.

Для вычисления медианы в электронной таблице Excel используйте функцию =MEDIAN().

Большинство языков программирования имеют готовую функцию медианы, но для таких языков программирования, как Python и Ruby, Вам, возможно, потребуется установка дополнительного математического пакета. У реляционных

	баз данных как Microsoft Access или MySQL нет функции медианы. Таким образом, для этих баз данных, также как и для других вариаций SQL потребуются альтернативные программные решения.
ПРИМЕР:	<p>Ниже представлен знакомый нам список цен из пяти объектов недвижимости:</p> <p>\$75 000 \$75 000 \$95 000 \$100 000 \$350 000</p> <p>Медиана составляет \$95 000, что является третьим значением в списке (5 цен/ 2 = 2,5, + + 0,5 = 3, поэтому берется третья по счету цена). Теперь сравните это значение со средним арифметическим значением в \$141 000, которое мы получили в последнем разделе. Которое из этих двух чисел наиболее корректно представляет рассматриваемую выборку одним числом?</p>

МОДА, ИЛИ НАИБОЛЕЕ ЧАСТО ВСТРЕЧАЮЩЕЕСЯ ЗНАЧЕНИЕ

Мода является наиболее часто встречающимся числом в выборке. Значение «моды» редко встречается в новостях, за исключением материалов, освещающих выборы и социологические опросы. Используйте это значение для того, чтобы найти чаще всего встречающийся элемент, но не для обобщенного представления выборки.

Например, человек, победивший на выборах является «модальным» кандидатом: он/она получил(а) «наибольшее количество» голосов. Это, однако, не эквивалентно большинству голосов – мода

может составлять менее 50 процентов, несмотря на то, что она соответствует наибольшему количеству.

В примере с выборкой цен на недвижимость мода составляет \$75 000.

ВСЕ ЛИ ЗНАЧЕНИЯ ЯВЛЯЮТСЯ СРЕДНИМИ?

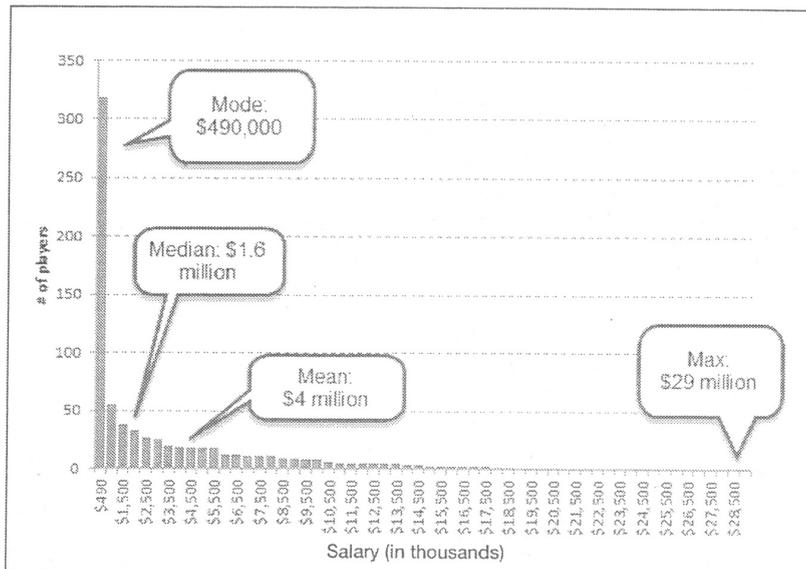
Когда читатели или телезрители слышат слово, «средний», им свойственно полагать, что речь идет об обычном человеке.

Но когда в материале идет речь о зарплатах, ценах на недвижимость, или других денежных величинах, то мы, как правило, представляем себе, что речь идет о тех людях, чьи финансовые возможности выше возможностей обычного человека.

Другими словами, ваши представления часто неточны или искаженно представляют информацию.

Основное практическое правило заключается в том, чтобы использовать среднее арифметическое в тех случаях, когда его значение схоже с медианой. Медиану, в свою очередь, используйте тогда, когда эти значения существенно разнятся. В тех случаях, когда выборка состоит из множества числовых значений, мода не совсем пригодна. Но, если вы можете выстроить эти числовые значения в кластеры, например, «менее \$100 000», то информация приобретет смысл.

Вот пример того, как были ранжированы зарплаты бейсболистов за 2013 год (значения средних величин указаны на основе фактического поименного списка с указанием зарплаты каждого бейсболиста):



Простое среднее арифметическое имеет смысл использовать, если в материале рассматриваются расходы владельцев команд на игроков. Но если речь идет о физическом состоянии (или уровне доходов) игроков, то лучше использовать медиану. Для рассмотрения наиболее часто встречающихся событий лучше обратиться к моде, которая в этом примере также является минимальной заработной платой в Главной Бейсбольной Лиге.

Порой невозможно одним числом адекватно отобразить описываемую ситуацию. Как правило, это часто происходит в случаях с так называемыми «бимодальными распределениями чисел». Такие наборы чисел были бы представлены на предыдущей диаграмме в виде двух кластеров чисел: одним на верхнем конце, одним на нижнем, со снижением посередине.

Зачастую бывает сложно обобщить какую-либо информацию, связанную с резким увеличением уровня рождаемости, как, например, информацию по родителям и детям, тенденциям строительства и тому подобное.

Например, в 1968 году набор учеников в начальную школу составил 34 миллиона человек, но в 1986 году этот показатель снизился до 27 миллионов, а в конце 1990-х снова возрос до 31 миллиона человек. При этом среднее арифметическое значение в 29,9 миллиона не отражает корректно ни один из упомянутых периодов, несмотря на то, что это значение близко к медиане в 30,2 миллиона. В этом случае для полноценного отражения ситуации, вам могут понадобиться два числа, вместо одного.

При работе с распределениями, средние величины которых сложно определяемы, можно использовать метод квартилей. При помощи квартилей числовые значения внутри выборки разделяются на четыре группы, каждая из которых содержит равное количество членов. Концептуально квартили схожи с медианой, но в тоже время они более информативны.

БОЛЕЕ УГЛУБЛЕННОЕ ИЗУЧЕНИЕ ЦЕНТРАЛЬНОЙ ТЕНДЕНЦИИ: УСРЕДНЕНИЕ СРЕДНИХ ВЕЛИЧИН, ИЗМЕНЕНИЯ И ЗНАЧИМОСТЬ

ЗАДАЧА УСРЕДНЕНИЯ СРЕДНИХ

ВЗВЕШЕННЫЕ СРЕДНИЕ ЗНАЧЕНИЯ

Как правило, при усреднении какого-то набора средних чисел не всегда возможно получить осмысленный результат. Рассмотрим следующий пример:

	<i>ЧИСЛЕННОСТЬ ПЕРСОНАЛА</i>	<i>СРЕДНЯЯ ЗАРПЛАТА</i>
УЧИТЕЛЬ	10 000	\$37 000.
УБОРЩИК (ЦА)	2 000	\$20 000
ДИРЕКТОР ШКОЛЫ	500	\$75 000
ПРОСТОЕ СРЕДНЕЕ		\$44 000.

Полученное в примере среднее значение не учитывает тот факт, что численность персонала по каждой из должностей различается.

Чтобы вычислить истинную среднюю заработную плату, необходимо взять \$37000 10000 раз, затем взять \$75000 500 раз, и \$20000 2000 раз. Сложив все результаты умножения и разделив их на общую численность персонала, вы получите истинное среднее арифметическое значение.

	ЧИСЛЕННОСТЬ ПЕРСОНАЛА		СРЕДНЯЯ ЗАРПЛАТА	
УЧИТЕЛЬ	10 000	X	\$37 000.	= \$370 000000
УБОРЩИК (ЦА)	2 000	X	\$20 000	= \$40 000000
ДИРЕКТОР ШКОЛЫ	500	X	\$75 000	= \$37 500000
СУММА	12 500			= \$447 500000
СРЕДНЕЕ-ВЗВЕШЕННОЕ			\$35 800	

При разделении \$447 500 000 / 12 500 (общая сумма денег, разделенная на общее количество людей), вы получите среднюю зарплату в \$35 800 – примерно то, что и ожидалось. Это не просто оценка среднего значения; математически – это именно то число, которое вы бы получили, если бы вы скрупулезно просчитали каждую зарплату по списку сотрудников.

НАХОЖДЕНИЕ МЕДИАННОГО ИНТЕРВАЛА

Из набора предложенных чисел в предыдущем примере невозможно определить медианный интервал. Даже для получения приблизительного его значения, вам будет необходим другой список,

содержащий количество сотрудников, сгруппированных по интервалам зарплат, вне зависимости от занимаемой ими должности.

Необходимо представить данный список в порядке возрастания с дополнительной колонкой, в которой указано число сотрудников, получающих зарплату в каждом из указанных диапазонов. В результате, вы получите общее число сотрудников, которое нужно будет разделить на 2 и прибавить 0,5, а затем найти группу, которая содержит то же самое число, что и в столбце промежуточного итога:

ДИАПАЗОН ЗАРПЛАТЫ	КОЛИЧЕСТВО СОТРУДНИКОВ	НАКОПЛЕННАЯ СУММА ПО КОЛИЧЕСТВУ СОТРУДНИКОВ
0 – \$20 000	2 000	2 000
\$21 000 – \$30 000	6 000	8 000
\$31 000 – \$40 000	4 000	12 000
\$41 000 – \$50 000	250	12 250
\$51 000 И ВЫШЕ	250	12 500
СУММА	12 500	

Эта группа включает среднее число 6 250-го и 6 251-го по счету человека

РАБОТА СО СРЕДНИМИ ВЕЛИЧИНАМИ И ПОКАЗАТЕЛЯМИ

Усреднение средних чисел порой предоставляет вам новую информацию при работе с показателями. Например, если вы располагаете списком показателей уровня преступности по городам, возможно, имеет смысл указать «средний уровень преступности» по всем городам, нежели общий уровень преступности в стране. Помните, что здесь идет речь о двух различных показателях и используйте соответствующую лексику. Ниже приведен пример, где предоставлен уровень преступности в 21 крупном городе США в 2012 году (информация составлена на основе Единого Отчета ФБР о Преступности / Uniform Crime Report).

<i>ГОРОД</i>	<i>НАСЕЛЕНИЕ</i>	<i>КОЛИЧЕСТВО ПРЕСТУПЛЕ-НИЙ</i>	<i>НА 1000 ЧЕЛ.</i>
Нью-Йорк	8 289 415	195 753	24
Лос-Анджелес	3 855 122	106 025	28
Хьюстон	2 177 273	129 288	59
Филадельфия	1 538 957	74 850	49
Феникс	1 485 509	70 235	47
Лас-Вегас	1 479 393	58 025	39
Сан-Антонио	1 380 123	89 611	65
Сан-Диего	1 338 477	37 229	28
Даллас	1 241 549	62 680	50
Сан-Хосе	976 459	32 010	33
Джексонвилл	840 660	39 863	47
Индианаполис	838 650	56 840	68
Остин	832 901	46 877	56
Сан-Франциско	820 363	44 675	54
Шарлотт-Мекленбург	808 504	37 825	47
Форт-Уэрт	770 101	37 038	48
Детройт	707 096	55 967	79
Эль-Пасо	675 536	19 270	29
Луисвилльское метро	666 200	32 595	49
Мемфис	657 436	53 010	81
Вашингтон	632 323	36 712	58
Итого	32 012 047	1 316 378	
Взвешенное среднее			41
Невзвешенное среднее			49
Медиана			49

Обратите внимание, что взвешенное среднее значение, на подсчет которого наибольшее влияние оказали наиболее крупные города, немного ниже, чем невзвешенное, что предполагает более низкий уровень преступности в крупных городах. (Чикаго, третий по вели-

чине город, не входит в выборку отчета ФБР, но уровень преступности там высок, что следует упомянуть в случае написания материала по данным примера).

Какое из подсчитанных средних использовать, зависит от сюжета вашего материала.

Если вы хотите сравнить свой город по отношению к другим городам, то целесообразно использовать невзвешенное усредненное значение уровня преступности. В противном случае, среднее значение является искусственно заниженным ввиду низкого уровня преступности в крупных городах. Но если вы хотите сравнить жителей вашего города с жителями других крупных городов, тогда важно учитывать численность населения городов, которая включена в взвешенное среднее или итоговую строку таблицы.

Выбор зависит от того, что исследователи называют «единицей анализа». Если фокусом вашего материала является работа полицейских управлений в крупных городах, то тогда имеет смысл сравнивать показатель преступности на 1000 человек по городам. Однако, если ваш основной акцент приходится на жителей больших городов, подвергающихся насилию, то возможно лучше использовать итоговые взвешенные данные по таблице.

Исследователи, как правило, используют медиану, нежели среднее арифметическое значение в тех случаях, когда им приходится работать с усредненными значениями показателей.

Работая с источниками информации, необходимо четко понимать с какими именно показателями вы имеете дело, поскольку они могут значительно отличаться:

- Итоговое значение или среднее значение по группе, часто во многом обусловлено влиянием одной или двух очень крупных составляющих, например, таких как города Лос-Анджелес и Нью-Йорк в примере выше. Во взвешенном среднем значении по городам, два этих города составляют более 1/3 от общей совокупности. Если показатели в этих крупных городах чрезмерно высоки по отношению к остальным городам в

выборке, то ваш сравнительный анализ будет основываться на сравнении с этими доминантными компонентами.

- При работе со средним значением от средних чисел возникает обратная проблема. Этот показатель рассматривает в расчетах меньшие значения или малые группы с той же степенью важности, что и большие значения и группы. Также расчеты могут быть искажены за счет одного чрезмерно большого или чрезмерно маленького значения внутри группы с меньшими значениями. Возможным примером подобной ситуации может являться Мемфис в таблице выше.
- Медиана уровней (преступности) между городами также может быть подвержена влиянию группы, содержащей меньшие значения, но в меньшей степени. По этой причине исследователи чаще используют именно эту величину.

ЧИСЛА, КОТОРЫЕ ПОМОГУТ ПРОИЗВЕСТИ ПРОВЕРКУ ПРИ РАСЧЕТЕ СРЕДНИХ ВЕЛИЧИН

Статистики используют ряд стандартных методов для проверки средних величин. Так как средние величины очень важны при любом проводимом ими тестировании, им необходимо знать, содержат ли подсчитанные средние какие-то ошибки. Цель статистиков, тем не менее, схожа с непосредственной задачей журналиста: Корректно ли отражает простая средняя величина совокупность числовых значений реального мира? Если нет, то дальнейший анализ может оказаться бессмысленным.

Вооружаясь статистическим аппаратом, можно проверять и исправлять средние значения.

СТАНДАРТНОЕ ОТКЛОНЕНИЕ

Стандартное отклонение дает вам информацию о том, как близко расположены числовые значения по отношению к среднему арифметическому значению.

Очень большое стандартное отклонение (например, значение которого превышает среднее арифметическое число для положительных чисел) означает, что данные по выборке сильно разбросаны, и ее среднее вряд ли отражает истинную ситуацию. К примеру, вполне вероятно, что внутри этой выборки существует группа людей, каждый из которых зарабатывает по \$50 000, и только один человек, который зарабатывает \$150 000. Вспомните, что в нашем предыдущем примере про бейсболистов средняя зарплата составляла \$3,3 миллиона. Стандартное отклонение при этом составляло \$4,5 миллиона, что значительно больше, чем среднее арифметическое, то есть среднее арифметическое не отражало истинную ситуацию.

Более умеренное стандартное отклонение (к примеру, когда его значение несколько меньше среднего арифметического значения для выборки неотрицательных чисел) означает, что рассматриваемая выборка вполне типична и ведет себя ожидаемым образом: разброс ее значений близок к колокольной кривой, соответствующей нормальному распределению. В этом случае среднее арифметическое значение поддается стандартным правилам статистики. Физические характеристики, такие как высота и вес следуют этому принципу. Также как и результаты таких экзаменов, как SAT (Scholastic Aptitude Test) или IQ, которые создаются таким образом, чтобы следовать принципу нормального распределения. В симметричном нормальном распределении числовых значений, которое встречается редко, медиана равна среднему арифметическому значению.

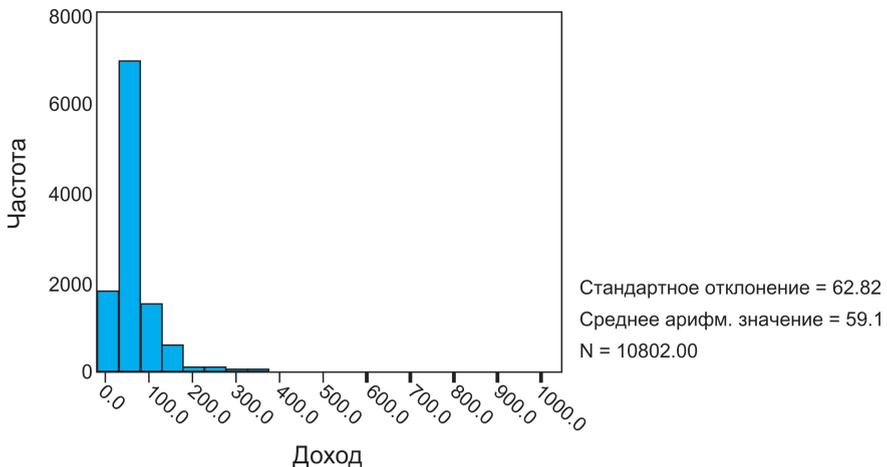
Очень маленькое стандартное отклонение (значение очень близкое к нулю) означает, что числовые значения внутри выборки не сильно разнятся между собой. Из этого может следовать, что рассматриваемые значения не столь интересны с информационной точки зрения. Например, если средний балл на школьном тестировании равен 98, и стандартное отклонение при этом равно 1, то фактически все тестовые баллы находятся в промежутке между 96 и 100. В таком случае само тестирование не отражает какой-либо значительной разницы между способностями студентов.

УСЕЧЕННОЕ СРЕДНЕЕ АРИФМЕТИЧЕСКОЕ ЗНАЧЕНИЕ

При подсчете усеченного среднего происходит «отсечение» экстремальных значений выборки. Как правило, для этого рассматривают средние 95 процентов значений выборки, игнорируя при этом экстремально высокие и экстремально низкие значения. Зачастую это значение приближается к чему-то похожему на медиану. Вы вправе определять самостоятельно, какое количество значений должно быть урезано. Некоторые урезают только по 1 или 2 процентам сверху и снизу числового ряда.

ГИСТОГРАММЫ

Диаграмму, которую мы рассматривали в примере с зарплатами игроков бейсбола, называется гистограммой, при помощи которой вы можете представить, каким образом распределяются значения выборки. Диаграммы такого типа позволяют определить, какое количество значений попадает в заданные интервалы. (Вы также можете воспользоваться программой статистического анализа, как, например, «R», чтобы произвести подобные расчеты.) Ниже приведен пример смещенного распределения, в котором все значения собраны в одной стороне графика.



Этот график отображает доходы в городе Нэшвилл, штат Теннесси. Под «частотой» подразумевается количество людей в каждой категории доходов. Это сложно разглядеть на графике, но некоторые соискатели на получение ипотечного кредита имеют доход свыше \$1 миллиона.

Но большинство соискателей – около 7 000 из них – имеют доход в размере \$40 000 (значения, которые располагаются на диаграмме в диапазоне от 0 до 100 по горизонтали).

Вы так же видите, что наша программа статистического анализа данных предоставляет числовые значения, связанные с данной выборкой, такие как среднее арифметическое значение и стандартное отклонение. Мы видим, что стандартное отклонение выше среднего арифметического значения.

Другим способом проверки средних величин, который обычно применяется при сравнении различных групп (например, при сравнении мужчин и женщин), является метод «диаграмм размаха». В этом виде диаграммы емко представлено большое количество информации и его стоит применять при работе со сгруппированными средними величинами.

Представленная информация является довольно простым изложением того, что статистики называют «одномерным анализом». Мы рекомендуем вам ознакомиться с учебником по введению в элементарную статистику или прослушать онлайн-курс по этой дисциплине.

СРАВНЕНИЕ СРЕДНИХ ВЕЛИЧИН ВО ВРЕМЕНИ

Будьте бдительны при сравнении средних величин во времени. Во многих случаях полученные средние значения отражают различные величины, создавая сравнения подобные сравнению яблок с апельсинами. В качестве примера, давайте рассмотрим медиану выборки цен на недвижимость.

Каждый месяц различные государственные ведомства и торговые ассоциации делают обзор цен на недвижимость. Порой это рынок

первичного жилья, порой – вторичного. Но никогда эти обзоры не содержат информацию по стоимости всех существующих домов, в силу ее недоступности.

Проблемы с этими данными включают в себя то, что называется взвешиванием и поправкой на качество.

Каждый объект, построенный или выставленный на продажу, считается за отдельную единицу при получении показателя (медианы или среднего арифметического значения). Но набор объектов недвижимости, проданных каждый месяц, меняется постоянно. Поэтому вполне вероятно, что обзоры, содержащие показатели за несколько месяцев, фактически измеряют изменения площадей недвижимости (поправка на качество), либо популярность недвижимости в городе по сравнению с пригородом, либо между, например, Калифорнией и Айовой (взвешивание), но при этом не содержат фактического значения стоимости типичного жилья.

Подобные проблемы возникают также и с другими данными, как, например, уровень смертности в больницах, зарплата руководителей, отчеты по успеваемости. К примеру, сравнение уровней смертностей между штатами может в конечном итоге свестись к сравнению возрастов, а не здоровья между жителями каждого штата.

В некоторых профессиональных средах существуют общепринятые методы, которые позволяют произвести соответствующую корректировку. Например, уровень смертности зачастую просчитывается по отношению к установленным возрастным категориям. Порой исследователи даже сужают свой фокус, выбирая одну возрастную категорию для изучения данной проблемы. К примеру, группа людей в возрасте 50 и выше является частым предметом сравнений. Иногда выборку целенаправленно составляют из молодых и пожилых респондентов, применяя уровни смертности к этим группам. В этом случае происходит суммирование уровней смертности для стандартного населения.

При расчете цен на недвижимость редко применяется подобный научный подход. Таким образом, мы попадаем в тупиковую

ситуацию: опубликовывать материал по теме, затрагивающей одну из крупнейших инвестиций для обычной семьи, беря в качестве основы ошибочные данные, либо не применять эти данные вообще?

Снятие материала с публикации не является выходом из ситуации для большинства журналистов. Таким образом, вопрос по трактовке данных, с большой вероятностью содержащих потенциальные ошибки, остается открытым.

Одним из возможных способов решения этой проблемы является опубликование в материале показателя как «наилучшую доступную оценку стоимости домов», включая при этом короткое пояснение того, что набор домов изменяется каждый месяц, таким образом, показатель может варьировать, хотя фактическая стоимость не изменяется.

ГЛАВА 3: РАБОТА С ГРАФИКАМИ

Мы привыкли к мысли о том, что визуализация данных зачастую является более эффективным инструментом представления информации, нежели их описание с использованием чисел. Инструменты визуализации информации, которые доступны даже для наименее технически подкованных из нас, позволяют за несколько минут преобразовать данные в графики.

Стивен Фью, консультант по визуализации бизнес-информации описывает некоторые принципы работы с графиками, которые позволяют читателю понять информацию лучше, чем просто на словах или в сухих цифрах. Итак, хороший график предоставляет качественный обзор самих данных, позволяет оперативно сравнивать цифровые значения и отслеживать или сопоставлять тенденции и закономерности.

Медиа-критики жалуются, что графика в новостях порой ввергает в заблуждение, хотя их критицизм в последнее время угасает. Это отчасти связано с доступностью журналистам не только всевозможных инструментов визуализации, включающих в себя изощренные программы, позволяющие создавать дизайн хорошего качества, но также и простых и внятных образцов по искусству создания информационной графики, как, например, веб-сайт Нейтана Яу (www.FlowingData.com). Эдвард Тафти, профессор Йельского Университета, также выпускает серию книг по инфографике. Эти книги могут оказаться вам полезными.

Однако эти ресурсы не всегда решают проблемы, с которыми сталкиваются журналисты при написании материала. В дальнейшем, я сделаю обзор некоторых из основных принципов и рекомендаций для создания статичной визуализации, которые позволят вам увидеть и понять то, что на первый взгляд не столь очевидно. Я буду намеренно избегать, насколько это возможно, упоминания конкретных инструментов или языков программирования, поскольку они очень часто меняются (помните ли вы Flash?). Многие из представленных способов легко применимы в электронной таблице Excel, но многосторонние возможности языка программирования «R», а

также и других инструментов могут стать для Вас стимулом к изучению данной тематики на более профессиональном уровне.

ВИДЫ ДИАГРАММ

Книги и примеры помогут вам понять, какую диаграмму / график лучше построить на основе имеющейся у вас информации. Книга Нейтана Яу («Искусство визуализации в бизнесе» / «Visualize This!» by Nathan Yau) – отличный учебник для начинающих. Далее – несколько других важных аспектов, которые важно знать при создании инфографики.

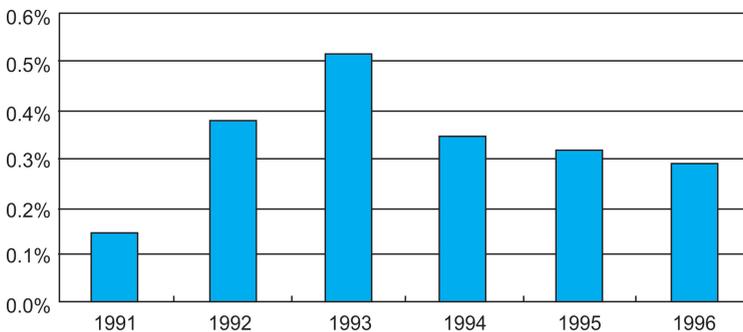
ПРЕДТВРАЩЕНИЕ ЗАБЛУЖДЕНИЙ

Легко впасть в заблуждение, когда информация о процентных изменениях представлена в виде линейных диаграмм.

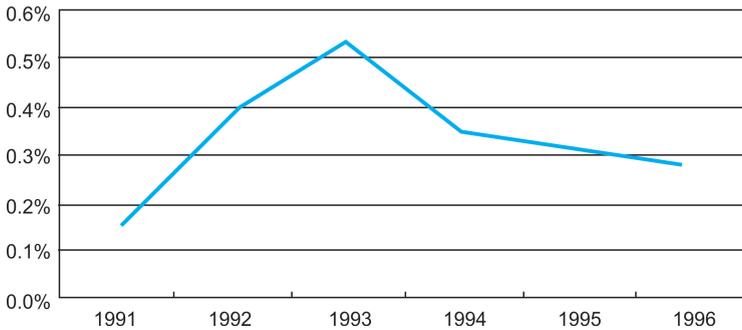
Как вы интерпретируете два ниже приведенных графика, которые показывают изменения в численности населения Нью-Йорка и его пригорода в течение шести лет?

Из первого графика – гистограммы – видно, что рост численности населения Нью-Йорка в начале 1990-х был более активным, нежели в дальнейшей перспективе.

Изменение численности населения Нью-Йорка и его пригорода



Изменение численности населения Нью-Йорка и его пригорода



Рассмотрим второй график. При первом взгляде на эту непрерывную линию, кажется, что, несмотря на четкие обозначения осей и необходимые маркировки, численность населения снизилась, так как линия на графике снижается. В отличие от столбчатого графика на Диаграмме N 1, на Диаграмме N 2 нет визуальной подсказки того, что показатели численности населения по-прежнему выше нуля. Поскольку столбчатые диаграммы четко фиксируют все изменения на оси, они понятнее отображают процессы изменений, нежели линейные.

Похожие проблемы иногда возникают с графиками, где информация разбита на различные категории. Если сами категории не сформированы согласно какому-то естественному признаку, как, например, хронологически или по степени важности, то лучше отображать их на горизонтальной оси, нежели на вертикальной, как это показано в предыдущем примере.

Создавая диаграмму, сначала продумайте ее досконально: либо самостоятельно, либо совместно с дизайнерами. Убедитесь в том, что график логически отображает информацию в материале. Покажите полученный результат без пояснительного текста своим коллегам, чтобы проверить, насколько понятно им визуальное изображение как самостоятельный элемент материала. Если при первом взгляде на график возникают сомнения уже у вас, то, представьте себе, какого будет вашим коллегам, читателям или зрителям.

МАЛЫЕ МНОЖИТЕЛИ И «ИСКРОГРАФИК» (SPARKLINES)

Эдвард Тафти советует журналистам применять два метода, которые можно использовать в начале материала. Первый заключается в создании структур, которые он называет малыми множителями и которые представляют собой диаграммы в уменьшенном масштабе, отдельно начертанные для различных категорий. В качестве категорий могут выступать, например, областные центры, виды грантовых программ или классификация рабочих. Представление данных диаграмм на одном рисунке облегчает анализ различий, так как, поняв одну диаграмму, проще считывать следующую по аналогии. Подобные графики легко создаются в «R» или других языках программирования, но не в Excel или других программных обеспечениях.

Следующий пример инфографики был использован в материале о национализации государственных законодательных реформ, в которых победу одержала одна политическая партия. Конкретнее, мы хотели понять, вырос ли объем финансовой поддержки лицами, проживающими за пределами того штата, где проводилась кампания. Научившись читать один график, вы поймете все остальные:

Вклады в законодательные кампании и



Заметьте, что масштаб на графиках не указан. Целью данной диаграммы было создание картины общей ситуации, позволяющей нам обозначить интересных участников, а не передача точной информации. (Стоит заметить, что данные в диаграмме не скорректированы с учетом уровня инфляции. Изменения в финансировании компании в течение рассматриваемого периода превысили любые последствия инфляции).

Техника, схожая с малыми множителями, также встроена в Excel – это «искрографики» Тафта, которые он называет графиками размером в одно слово. Изначально он предлагал заменить текстовое описание числовых значений этими кривыми, но вместо этого они стали стандартным способом сравнения рядов чисел внутри электронной таблицы.

Ниже приведена таблица вкладов из предыдущего примера с использованием двух видов искрографиков: в одной изображены столбцы, а в другой – линии. Заметьте, что в обоих случаях информация воспринимается по-разному, поскольку использована разная шкала.

ШТАТ	2000	2002	2004	2006	2008
Калифорния	\$2,266,183	\$1,689,853	\$1,111,774	\$1,334,752	\$1,683,304
Техас	\$1,216,272	\$876,942	\$893,773	\$998,619	\$1,439,708
Пенсильвания	\$1,070,238	\$708,271	\$830,084	\$1,327,543	\$1,535,582
Иллинойс	\$1,320,173	\$1,201,951	\$650,275	\$792,131	\$637,331
Нью-Йорк	\$481,776	\$572,375	\$736,652	\$1,003,309	\$1,115,335
Огайо	\$424,996	\$523,313	\$311,127	\$573,245	\$955,527

ШТАТ	2010	2012	ВСЕГО		
Калифорния	\$1,574,195	\$2,026,745	\$11,886,806		
Техас	\$1,504,763	\$1,882,099	\$8,812,175		
Пенсильвания	\$1,213,818	\$1,354,282	\$8,039,819		

Иллинойс	\$1,043,078	\$1,393,577	\$7,038,514		
Нью-Йорк	\$1,155,434	\$1,115,500	\$6,180,381		
Огайо	\$1,440,407	\$1,450,571	\$5,619,186		

КАРТЫ

Создавать карты географической информации – это один из самых важных навыков, которые журналисты обрели за последние два десятилетия. Ничто не может быть лучше для понимания материала, чем, например, увидеть нанесенные на карту преступления или обозначить сосредоточение обеспеченных семей в районах, где расположены школы.

Но карты не всегда являются наилучшим инструментом отображения географических данных. В предыдущем примере о финансовой поддержке законодательных инициатив для использования карты пришлось бы отказаться от использования другой существенной информации. В частности, карты США, ввиду малонаселенности в горной местности и на севере страны, могут отражать нетипичные для страны тенденции. Например, Северная Дакота, с ее единственным избирательным округом по выборам в Конгресс, на электоральных картах часто выглядит единой монолитной стеной. Во время выборов 2012 года «NPR» и «The New York Times» опубликовали подобные карты, но при этом не обратили внимания аудитории на особенности географии.

Безусловно, размещение географических данных на интерактивном графике позволяет аудитории географически переместиться в тот уголок, который вызывает у них наибольший интерес. Кроме того, карты помогают найти эпицентр событий материала.

ПРОБЛЕМА МАСШТАБА

Графики созданы для того, чтобы передавать информацию, а не скрывать ее. Некоторые из редакционных правил, однако, препятствуют достижению этой цели. Большинство проблем связано с

выбором масштаба для оси Y, или той оси, которая содержит числовые данные, а не описательную информацию.

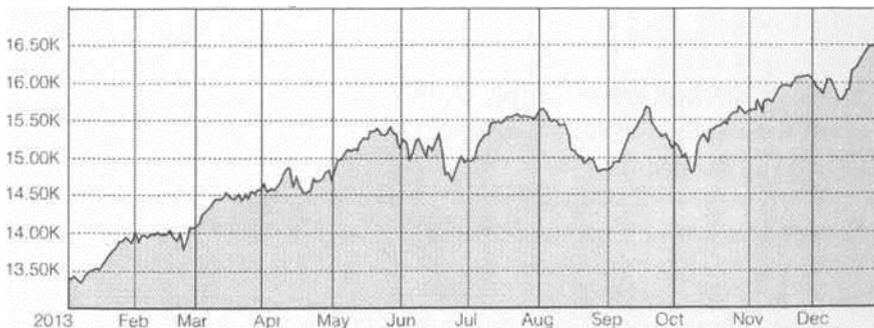
Многие авторы, занимающиеся тематикой подтасовки при помощи статистики, утверждают, что графики являются одним из инструментов, при помощи которых изоэдренные журналисты вводят в заблуждение своих читателей, приводя сенсационные версии некоторых фактов, а не правду.

Далее мы хотим привести пример одного такого типичного комментария, который был упомянут в книге «Как лгать с помощью статистики» / «How to Lie with Statistics», ставшей Библией для многих журналистов:

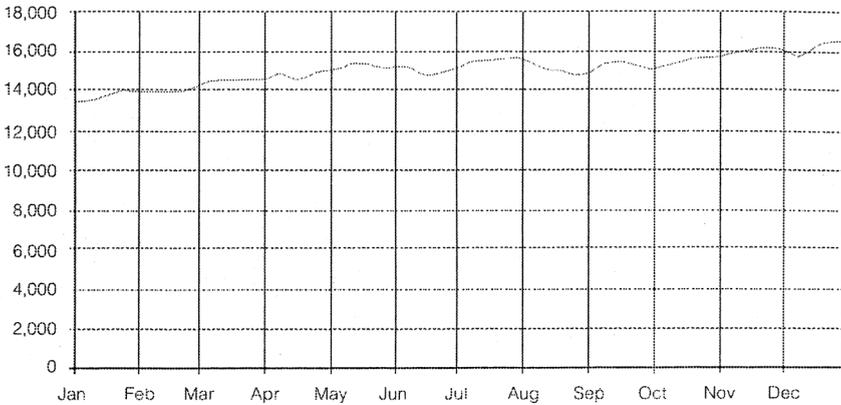
... предположим, что вы хотите выиграть спор, шокировать читателя, побудить его к действию, продать ему что-либо. Для этого... отсекаете основание графика.

Другие, напротив, не соглашаются с тем, что отсечение основания у графика является обманом или попыткой ввести в панику.

В качестве примера давайте рассмотрим «Промышленный индекс Дуу-Джонса», который в 2013 году варьировался приблизительно от 13 000 до 17 000. На приведенной ниже диаграмме мы видим один из способов представления этого индекса.



Вот аналогичная диаграмма еженедельных цен за тот же период времени, где отсчет ведется с нуля.



Подумайте об аудитории этого графика – внимательные инвесторы, которые, прежде всего, заинтересованы в эффективной информации. Соответственно, они предпочтут вариант с отсеченным основанием.

Патрик Ремингтон (Patrick Remington), эпидемиолог из университета Висконсина, также считает бессмысленным правило о том, что график должен начинаться именно с нуля. Во время презентации своей работы журналистам он продемонстрировал три графика и спросил аудиторию, который из этих графиков наиболее информативен. Как оказалось, это не тот график, который начинается с нуля.

Ремингтон предложил оставить резервное пространство у основания графика, которое составляло бы 1/3 от общего диапазона графика. Это означает, что для представления значений в диапазоне от 40 до 100, нужно оставить резервной одну треть разницы между наибольшим и наименьшим значениями (или 60/3, или 20) с каждой стороны графика. Сам график нужно начертить, начиная с точки 40-20 (то есть 20) и заканчивая точкой 100+20 (то есть 120).

Вот как будет выглядеть диаграмма движения Доу-Джонса, при использовании этого правила:



Необоснованно считать, что некоторые числовые значения снизятся до нуля. В случае Ремингтона, он знает, что рак груди, во всяком случае, в ближайшем будущем, невозможно будет исключить как причину смертности. Таким образом, он ищет точку отсчета или соответствующую нижнюю границу для стартового значения, подтвержденного результатами исследований. Если «правило одной третьей» не включает в себя эту границу, он может расширить шкалу графика.

В 1988 году Уильям Кливленд (William Cleveland) и два его соавтора создали концепцию «угла в 45° », согласно которой средний угол наклона касательной в графике должен быть равен 45° , что позволяет наиболее отчетливо визуальнo воспринимать течение изменений на графике, как это показано на диаграмме «Yahoo» выше. Концептуальные разработки Кливленда по восприятию информации встроены по умолчанию в «ggplot2» библиотеки языка «R», которая также включает в себя рекомендуемые соотношения между сторонами (соотношение горизонтального к вертикальному) и угол в 45° .

При построении графика, обязательно учитывайте ту информацию, которая, согласно мнению экспертов, важна. Если большинство экспертов считает значимым снижение показателей на рынке на 5 процентов, то ваш график непременно должен передавать этот

факт. Он должен отражать реальную значимость тенденции – либо ее отсутствие.

ИСПОЛЬЗОВАНИЕ ВИЗУАЛИЗАЦИИ КАК ИНСТРУМЕНТ РЕПОРТАЖА

Большинство книг и источников по визуализации информации рассматривают визуализацию как конечный продукт публикации. Журналистам уже известна история, которую они хотят пересказать, и они ищут эффективные и интересные формы ее подачи.

Желательно, чтобы большая часть журналистов рассматривала визуализацию как средство репортажа, которое помогает проверять истории и исправлять информацию, полученную из источников или доступных архивов. Подобную проверку можно сделать при помощи Excel, но лучше, если вы изучите азы языка программирования «R» или станете активным пользователем программы «Tableau Public» (www.Public.Tableau.com).

На ранних стадиях работы над материалом не зацикливайтесь над точностью каждого отдельного числа. Возможно, поначалу у вас будет отсутствовать одно из чисел в линейке значений, отслеживающих ту или иную тенденцию, или одно из значений может оказаться ошибочным. Не стремитесь к совершенству; на начальных стадиях можно даже начертить простой схематичный график от руки, если для вас это проще. Вы сможете внести исправления перед публикацией, но вполне возможно, что вы придете к выводу, что график не настолько необходим, чтобы оправдать потраченное на него время. Вы осознаете, каким сложным процессом является заблаговременный сбор информации для ее точной обработки в сочетании с попыткой понять происходящую ситуацию.

Вот несколько советов, которые помогут вам на ранней стадии подготовки материала:

Попробуйте сначала посмотреть на ваши графики без каких-либо обозначений – в различных размерах и при разных масштабах. Попытайтесь построить различные виды диаграмм (но, конечно, не

трехмерных). Вы убедитесь, что каждый из построенных графиков отобразит информацию по-своему.

Отобразите свою информацию в исходном виде, в процентах от общего, в качестве процентной ставки, в виде изменений по времени или относительных процентных изменений во времени, иногда даже в виде логарифмов или квадратных корней. Всякий раз вы сможете увидеть что-то новое. В поиске закономерностей не существует правил, вы можете заметить что-то интересное, что может исчезнуть при рассмотрении информации под другим углом. Однако будьте бдительны по отношению к тем закономерностям, которые прослеживаются только в одном или двух графиках.

Поэкспериментируйте с менее традиционными видами графиков, как, например, графики в виде дерева, в виде пузырьковой диаграммы или так называемыми «диф-чартами» (от англ. difference – разница), которые отражают разницу изменения.

ТОЧНОСТЬ, ДОСТОВЕРНОСТЬ И ИНТЕРАКТИВНАЯ ВИЗУАЛИЗАЦИЯ

Всем нам известно и свойственно желание опубликовать свою работу, раскрыв при этом информацию при помощи новых технологии, как, например, интерактивной графики или новостных приложений.

Но прежде чем окунаться с головой в эту сферу, убедитесь, что использованная Вами информация способна выстоять любую суровую критику и тесты. Визуализация информации может выявить все ее недостатки, в особенности те проблемы, которые можно скрыть при использовании только текста или статичной графики.

ГЛАВА 4: ТИПОВЫЕ ПРИМЕРЫ ИНФОРМАЦИОННЫХ МАТЕРИАЛОВ

ОСВЕЩЕНИЕ ТЕМЫ БЮДЖЕТА – ИСПОЛЬЗОВАНИЕ ПРОЦЕНТОВ, КОРРЕКТИРОВКИ УРОВНЯ ИНФЛЯЦИИ И ВЕЛИЧИН, ФИКСИРУЮЩИХ ИЗМЕНЕНИЯ

Редакторов СМИ часто интересует тематика улучшения качества изложения материала по вопросам местного государственного бюджета. Некоторые из журналистов, работающих с этими материалами, научены использовать электронные таблицы для создания наилучших вопросов для предстоящего интервью. Но осознанное освещение вопросов бюджета, с учетом понимания воздействия информации на читателей и на зрителей, требует на деле использования всех методов, затронутых в этой книге.

Материалы по бюджету намного проще подготовить, используя электронные таблицы Excel, поскольку это позволит вам произвести или перепроверить необходимые математические расчеты. Но работа в Excel не является обязательной. Вы можете произвести подсчеты вручную на листе бумаги.

ПОНИМАНИЕ БЮДЖЕТА

Бюджет это предмет планирования, который может быть точным или неточным, осознанным и обоснованным или наоборот.

К примеру, если я рассчитаю свой годовой бюджет, полагая, что в этом году моя зарплата вырастет на 25 процентов, то вы, наверняка, подумаете, что я просто мечтатель(ница). У меня должна быть веская причина для подобного расчета, прежде чем я использую его при пересчете размера своей арендной платы на следующий год.

Аналогично, если я начну рассчитывать свой семейный бюджет, исходя из внезапного сокращения моих расходов, скажем, снижения аренды на 15 процентов, вы тоже подумаете, что я выжил(а) из

ума. Что же нужно сделать, чтобы так расщедрить своего арендодателя?

Также, если вдруг со следующего месяца размер моей страховки возрастет на 30 процентов, то это будет безответственно по отношению к моей семье, если я не учту это повышение при планировании бюджета на следующий год.

Но государственные органы и ведомства разных уровней порой основывают свой ежегодный бюджет на аналогичных ложных предположениях. Таким образом, к вопросу бюджета следует подходить с позиции: «Есть ли основание считать эту информацию достоверно точной?»

Одна из участниц наших семинаров в штате Флорида привезла с собой бюджет на учебный год, который начинался буквально на той же самой неделе. Госорганы на тот момент запаздывали с принятием окончательной версии документа и должны были еще в тот день проголосовать. Мы просмотрели этот документ на предмет выявления проблем бюджетного планирования и обнаружили одну значительную нестыковку: за первые три дня учебного года школьный округ уже потратил больше половины своего запланированного бюджета на оплату труда замещающих учителей.

В бюджете также отражаются решения, которые принимаются политиками. Они обсуждают вопросы повышения налогов, сокращения расходов, переноса ресурсов на борьбу с преступностью или на сокращение контингента учебных классов. Мы должны подходить к вопросам бюджета аналогично. Это означает, что мы должны присматриваться к изменениям в тенденциях за последние годы, что может включать в себя также освещение решений, которые принимают политики.

КАК ЖУРНАЛИСТУ ПОДГОТОВИТЬСЯ К СЕЗОНУ БЮДЖЕТНОГО ПЛАНИРОВАНИЯ

Для составления общей картины соберите данные по численности населения, уровню инфляции, объемах бюджетного финанси-

рования в прошлом. Чтобы проследить, как бюджет изменился за определенный промежуток времени, используйте в своих расчетах такие показатели, как «на душу населения» (см. стр. 24) и корректировка уровня инфляции (см. стр. 28).

Сравните предыдущие объемы финансирования с фактическими расходами. Привлеките органы законодательной власти к ответственности, если они допустили ошибки при планировании бюджета (см. «Измерение изменений» стр. 22).

Будьте готовы проверить арифметику в бюджете. Она зачастую бывает неверной.

НАЛОГИ НА ИМУЩЕСТВО

Расчет налогов на собственность зависит от метода оценки, установленного в каждом штате. С одной стороны, штат Нью-Йорк не обновлял параметры оценки стоимости имущества уже более 50 лет; таким образом, налог на недвижимость в Нью-Йорке зависит больше от того, когда объект был построен, нежели какова его фактическая ценность. С другой стороны, штат Мэриленд обновляет эти параметры каждые три года по скользящему графику, включая стоимость дома в оценку. В середине этого спектра находятся такие штаты, как Флорида, где такие параметры обновляются регулярно и действует частичное освобождение от налога на недвижимость и потолок для ставки налога на недвижимость.

РАСЧЕТ НАЛОГОВ НА ИМУЩЕСТВО

Существует три основных способа изменения размера налога на имущество.

Самый простой путь – это поменять налоговую ставку, которую, в зависимости от ее применения, иногда называют «миляж» («тысячной»). Эти вопросы налогообложения обсуждаются на заседаниях Совета округа.

Второй способ – это откорректировать рыночную стоимость объекта недвижимости. Как правило, так поступают домовладельцы, которые хотят оспорить размер налогового взыскания, либо местные инициативы, в интересах которых сдерживать рост цен на рынке недвижимости, вызванный перерасчетом параметров оценки стоимости недвижимости. Поскольку все остальные расчеты производятся четко по формуле, у собственников недвижимости есть только один выход – это оспаривать саму оценочную стоимость жилья.

Третий вариант – это изменить корректировку, применяемую к величине, облагаемой налогом, а не саму стоимость недвижимости. За счет денежных дотаций на недвижимость эта величина уменьшается для граждан, проживающих в своих домах; все остальные облагаются налогом по полной ставке. В некоторых штатах включают в облагаемую налогом величину (налоговую стоимость) годовые повышения. В других штатах к этой величине применяют разновидности дополнительных льгот, сходных с уменьшением подоходного налога.

Итак, ключевыми цифрами для определения налога на имущество являются:

1. Стоимость дома, указанная в оценочных документах при расчете налога на собственность. При этом реальная стоимость дома неважна. В Калифорнии, например, за «рыночную стоимость» берется стоимость дома в 1978 году, либо в год его постройки или последней продажи. Обратите внимание, что в некоторых штатах текущий пункт не учитывается, и расчет налога можно начинать с пункта номер 2.

2. Стоимость недвижимости, которая облагается налогом, иногда называемая оценочной стоимостью или налоговой стоимостью. Как правило, это – рыночная стоимость дома с некоторыми корректировками.

Например, во Флориде, налоговую стоимость можно сократить двумя способами: во-первых, эта величина за год не может

вырасти более чем на 3 процента, поскольку избирательные инициативы ограничивают повышение тарифа на уплату налогов. Кроме того, денежные дотации на недвижимость в размере \$25 000, прилагаемые к рыночной стоимости имущества, действительны для граждан, проживающих в своих домах. Фактический подсчет сокращения немного более сложен, как это будет показано в последующем примере.

3. Ставка налогообложения, или долларовая величина, определяемая от оценочной стоимости. Если налог взимается со стоимости, близкой к рыночной стоимости недвижимости, то ставку часто выражают в «миляже» («тысячных»), что соответствует налогу на каждую \$1 000 оценочной стоимости.

4. Любые специальные правила для расчета налога на имущество, применяемые в вашем регионе. В большинстве случаев дополнительные исправления применяются к налоговой стоимости, а не к окончательному налогу. Здесь могут действовать так называемые кредитные пошлины. (Они в чем-то сходны с уменьшением персонального подоходного налога: выплаты по ипотечному кредиту могут быть вычтены из зарплаты. При этом перечень величин, которые можно вычесть из окончательного налога ограничен.)

ДВА ПРИМЕРА РАСЧЕТА НАЛОГОВ НА ИМУЩЕСТВО

<p>ПРИМЕР 1 МЭРИЛЭНД</p>	<p>В штате Мэриленд налоги высчитываются на основе реальной рыночной стоимости дома, индивидуально определяемой каждые три года. Вместо использования «тысячных» Мерилэнд сокращает реальную стоимость недвижимости и применяет процентные ставки.</p> <p>Шаг 1: Оценка стоимости: оценщик, сравнил дом с другими аналогичными объектами, по которым уже были заключены сделки купли-продажи, назначает его рыночную стоимость.</p> <p>При первой оценке дом был оценен в \$150 000, при следующей – в \$210 000.</p>
--	---

	<p>Шаг 2: Если стоимость недвижимости возросла, то разделите это возрастание на 3. Затем прибавьте полученную величину к значению стоимости по годам.</p> <p>$\\$60\,000 / 3 = \\$20\,000$ – повышение в оценке стоимости за каждый год: $\\$170\,000$ – первый год, $\\$190\,000$ – следующий, $\\$210\,000$ – за последний год.</p> <p>Шаг 3: Для получения налоговой (оценочной) стоимости сократите оценку стоимости следующим образом: возьмите 40 процентов от реальной рыночной стоимости объекта.</p> <p>Для расчетов на следующий год:</p> <p>$\\$170\,000 \times 0,40 = \\$68\,000$</p> <p>Шаг 4: Примите в расчет государственные и местные налоговые ставки, которые колеблются от 0,5 до 6,0 процентов от оценочной стоимости.</p> <p>В течение следующего года:</p> <p>$\\$68\,000 \times 0,06 = \\$4\,080$</p>
<p>ПРИМЕР 2 КАЛИФОРНИЯ</p>	<p>В 1970-е годы калифорнийцы приняли закон, который «заморозил» стоимость недвижимости в налоговых ведомостях.</p> <p>Так, в Нью-Йорке и в нескольких других штатах единственный способ определить налог на имущество – это знать, когда дом был построен или в последний раз продан.</p> <p>Это – тот самый случай, когда может понадобиться электронная таблица Excel, где вы можете рассчитать, как налог на имущество влияет на самые различные группы домовладельцев – тех, кто живут в своих домах с момента постройки, тех, кто только недавно</p>

	<p>построил себе дом, и тех, кто только собирается приобрести недвижимость.</p> <p>Необходимо произвести следующие расчеты:</p> <p>Шаг 1: Возьмите за базу прошлогоднюю оценочную стоимость. \$40 000</p> <p>Шаг 2: Выясните, продан ли был дом в прошлом году. Если да, то используйте его продажную стоимость.</p> <p>Если нет, тогда используйте Индекс Потребительских Цен для Калифорнии за прошлый и текущий годы для выяснения процентного изменения общих цен в этом регионе.</p> <p>Цена продажи дома составляет \$300 000.</p> <p>Индекс Потребительских Цен для Калифорнии составил 160,5 в прошлом году, 157,1 – годом раньше, таким образом, процентное изменение составляет:</p> $(160,5 - 157,1) / 157,1 = 2,2 \text{ процента}$ <p>Шаг 3: Для тех объектов недвижимости, которые были проданы или построены, вычтите \$7 000.</p> <p>Для остальных домов примените к старой оценочной стоимости недвижимости наименьшую из следующих двух величин: 1) процентное изменение, согласно Индексу Потребительских Цен</p> <p>2) 2 процента</p> <p>Проданный дом:</p> $\$300\,000 - \$7\,000 = \$293\,000$ <p>Дом, не сменивший владельца:</p> $\$40\,000 \times 1,02 = \$40\,800$
--	--

	<p>Шаг 4: Умножьте результат на размер налоговой ставки. Несколько лет назад в Сан-Франсиско налоговая ставка составляла \$1,164 на каждые \$100 от стоимости объекта недвижимости.</p> <p>Налог с проданного дома:</p> $\$293\,000 \times 1,164 \text{ процента (или } \times 0,01164) = \3410 <p>Дом, не сменивший владельца:</p> $\$40\,800 \times 0,01164 = \$475.$
--	---

ОСВЕЩЕНИЕ В СМИ ЗЕМЛЕТРЯСЕНИЙ, УРАГАНОВ И ПОГОДЫ – ИСПОЛЬЗОВАНИЕ ИНДЕКСОВ

Освещение погоды или природных катаклизмов является ключевым примером для понимания понятия «по сравнению с чем?» в новостях.

Если в некоторых районах уровень воды достигает 10 футов, насколько далек этот уровень от уровня наводнения? Если ураган переходит из первой категории во вторую, то в каких масштабах ожидается ущерб, и насколько были разрушительны другие ураганы подобной силы? Если сила землетрясения составила 5,2 балла по шкале Рихтера, то каковы последствия для жителей в эпицентре землетрясения?

Эти информационные сообщения сами по себе визуальны, поэтому нет особой необходимости заикливаться на цифрах. Но все-таки без некоторых цифр не обойтись. Мы собрали перечень полезной информации по этой теме (см. терминологию на тему: «Погода и погодные условия»).

ЗЕМЛЕТРЯСЕНИЯ И ШКАЛА РИХТЕРА

Шкала Рихтера измеряет мощность землетрясения в его эпицентре. Отсчет шкалы ведется с 1 балла, хотя фактически землетрясение ощущается, начиная с 2 баллов. Самое сильное землетрясение

было зарегистрировано 22 мая 1960 года в Чили, где сила толчков достигла 9,5 баллов.

Сложность в освещении шкалы Рихтера состоит в том, что ее показания представлены в виде логарифмических значений: с каждым новым делением амплитуда землетрясения увеличивается в 10 раз. Так, землетрясение силой в 2 балла фактически в 10 раз сильнее, чем в 1 балл; в 3 балла – в 10 раз сильнее, чем в 2.

Кроме того, по шкале Рихтера сложно определить масштабы нанесенного ущерба, поскольку замеры амплитуды производятся под землей или на дне океана, но не в самих населенных пунктах.

Существует еще одна шкала, которая может оказаться более понятной для ваших читателей. Возможно, вы ранее встречали упоминание шкалы Меркалли. При ее использовании есть два очевидных преимущества. Если в вашей местности происходит землетрясение, то по шкале Меркалли можно определить масштаб нанесенного ущерба именно в вашей местности, а не в эпицентре самого землетрясения. Что немаловажно, значения этой шкалы – не логарифмические, а понятные нам – линейные, как например, 10 баллов вдвое сильнее, чем 5 баллов.

УРАГАНЫ И ТОРНАДО

Понять, как определить силу урагана намного легче, поскольку за основу исчисления берется сила ветра и используется шкала от 1 до 5, согласно т.н. «Шкале ураганов Саффира-Симпсона».

У урагана первой категории скорость ветра достигает 74-95 миль в час, но ущерб при этом минимальный. Таким был ураган «Дэнни» в 1997 году близ устья реки Миссисипи.

При самом сильном урагане, урагане пятой категории, скорость ветра превышает 155 миль в час. При этом размеры ущерба сокрушительные: небольшие здания и трейлеры полностью разрушены, деревья вырваны с корнем, крыши многих домов снесены ветром.

Ураган «Камиль» пронесся через штаты Миссисипи и Луизиана. Ураган «Катрина» в 2005 году, в результате которого погибло более 1 800 человек, также был пятой категории, но когда он обрушился на побережье, его сила равнялась только третьей категории.

Сила торнадо измеряется по шкале Фудзиты – от F0 до F6, но торнадо силой F6 ранее не пока наблюдалось. Эта шкала встречается реже, поскольку данные определяются постфактум, а не во время самого торнадо, а в новостных сводках важна самая оперативная информация. Один из ключевых элементов расчета силы торнадо – это скорость ветра внутри воронки. Кроме того, сила торнадо не зависит от его размера, поэтому ущерб от большого торнадо может быть незначительным.

ИНДЕКС ЖАРЫ (ТЕПЛА) И ВЕТРО-ХОЛОДОВЫЕ ПАРАМЕТРЫ

Ветро-холодовой индекс, измеряемый по шкале Фаренгейта, измеряет потерю тепла человеком при учете температуры воздуха и скорости ветра. Формула расчета этого индекса довольно сложна и необходимости в ней нет.

Если в погодных сводках речь идет о резких похолоданиях, важно помнить следующее: показания, согласно ветро-холодовому индексу, как правило, ниже фактической температуры на улице.

То же самое касается индексов жары, которые выводятся из ряда подсчетов, включающих в себя оценки сравнения степени потопотделения среднестатистического человека в сухую и влажную погоду. К животным эти показатели не применимы. Они дают понять, как человек будет ощущать себя на улице при той или иной температуре в зависимости от своего географического местоположения. В некоторых случаях они являются более полезными для сравнения климатических условий в различных регионах страны, чем ветро-холодовой индекс. Индекс жары (тепла) фактически квантифицирует такое выражение, как «не жара, а влажность».

СТАДИОНЫ, УВОЛЬНЕНИЯ И ОЦЕНКА ЭКОНОМИЧЕСКОГО ВОЗДЕЙСТВИЯ

Одна из самых непростых тем для журналистов, освещающих деятельность местных государственных органов, это отчеты об оценке экономического развития. Этот всеобъемлющий документ содержит интересы различных групп: лоббистов, которые борются за послабление налоговых выплат в своих отраслях либо критикуют новые законы и постановления о минимальной заработной плате как ведущие к сокращению рабочих мест; промышленных магнатов, которые лоббируют строительство нового бейсбольного стадиона; местных активистов, лоббирующих за снижение налогов для нового бизнеса или расхваливающих «плюсы» туризма во время чемпионата США по футболу «Супербоул» («Super Bowl»). Данные этих отчетов являются самыми ненадежными в новостях, так как они окрашены мнением людей, создающих оценки, либо групп, которые их оплачивают.

На самом же деле, истина заключается в том, что никто с точностью не знает, какова степень экономического воздействия, которое может возникнуть или, напротив, исчезнуть вследствие предписаний или ослаблений. Оценки экономического воздействия проецируются, исходя из моделей, основывающихся на исторических данных развития различных отраслей и регионов.

Эти модели могут быть как консервативными, так и оптимистичными; все зависит от того, используют ли они точные данные или приблизительные расчеты «на салфетке». К примеру, администрация Обамы ошибочно применила самый консервативный метод оценки количества «созданных или сохраненных» рабочих мест в результате программы по стимулированию экономики с бюджетом \$787 млрд., принятой в разгар экономического кризиса в 2009 году. Правильнее было бы посчитать необходимое число как количество рабочих мест, которые были напрямую профинансированы средствами этой программы. Конечно, частично деньги программы были потрачены на создание дополнительных рабочих мест, но остается единственный вопрос – сколько новых рабочих мест было создано и в каких именно отраслях?

На другом примере мы видим, что некоторые из оценок являются чистой фабрикацией. В г. Тампа (Tampa) Джордж Штейнбрэннер (George Steinbrenner), в то время владелец профессиональной бейсбольной команды «Нью-Йорк Янкиз» (The New York Yankees), утверждал, что небольшой стадион, который использовался лишь в течение шести недель в году, создаст большее количество рабочих мест, нежели тот стадион, который он одновременно предлагал нью-йоркским чиновникам для использования в течение полного сезона.

В следующем примере использования оценок экономического воздействия мы видим более позитивные эффекты. Исследование деятельности предприятия Whirlpool, закрывающего свои заводы на территории США и переводящего деятельность в оффшорные страны или страны с более дешевой рабочей силой, которое было профинансировано профсоюзами и проведено лояльной исследовательской группой «Good Jobs First», является хорошим примером такой прозрачной, хотя и сделанной в собственных интересах, оценки экономического воздействия.

КАК РАСПОЗНАТЬ ПОДВОДНЫЕ КАМНИ В ЭКОНОМИЧЕСКИХ ОБОСНОВАНИЯХ

Как правило, существуют три показателя экономических воздействий, вовлеченных в оценку количества рабочих мест. Сложность заключается в том, что многие люди используют одинаковые термины для обозначения различных понятий.

Первый – это эффективность фактических затрат или сумм, вложенных в экономику: сколько человек будет трудоустроено в результате проекта? Только этот показатель был учтен при оценке эффективности программы стимулирования экономики.

Второй вид показателя, иногда называемый «прямым», а иногда «косвенным» эффектом экономического воздействия, отражает, сколько товара будет приобретено у местных производителей, которые, в свою очередь, откроют дополнительные рабочие места для поддержания проекта.

Эти два показателя являются показателями предложения: сколько поставщиков нужно привлечь, чтобы вести бизнес.

Последний показатель – наиболее сложный к пониманию: он отражает условные эффекты, иногда называемые мультипликаторами прибыли или даже (вновь) косвенными эффектами. Эти эффекты показывают сторону спроса: эффект от наличия больших денег, циркулирующих в местной экономике.

ГЛАВА 5: ОБЗОРЫ И ОПРОСЫ – ЧИСЛА КАК ОБОБЩЕНИЯ И ДОГАДКИ

Существует множество замечательных книг, написанных для журналистов, в которых разбираются тонкости освещения обзоров и опросов общественного мнения. Ни одно пособие по использованию математики в журналистике не было бы полным без обращения к этим вопросам.

Основная ошибка журналистов при освещении подобных тем – это игнорирование того, что уже известно. Несоответствие между тем, что говорится в заголовке или в самом материале и тем, что отражено в диаграмме или графике, зачастую побуждает критиков подвергать сомнению наше понимание основ статистики и математики.

ЗНАЙТЕ СВОЕГО СОЦИОЛОГА

Одним из простых способов определения истинности и беспристрастности проведенного опроса является оценка группы исследователей, которые его провели. Однако это не всегда так просто, как кажется, поскольку заинтересованные стороны могут нанять авторитетных социологов для проведения своих опросов.

Но большинство профессиональных социологических агентств следуют установленным стандартам качества. Так, «Галлап» (Gallap), центр исследования общественного мнения «Роупер» (Roper), «Мэйсон-Диксон» «Mason-Dixon», а также некоторые другие исследовательские центры при СМИ не будут рисковать своей репутацией, проводя опросы, заказанные группой лиц с сугубо личными интересами. Поводов для беспокойства меньше, если вам известно, что одна из этих организаций участвовала в создании опроса. Один из способов выяснить, был ли опрос проведен беспристрастной группой исследователей – это проверить её в базе данных Центра исследования общественного мнения «Роупер» (Roper). Если никакой информации по составителям опроса не найдено, поводов для беспокойства нет, поскольку не все исследовательские группы зарегистрированы в этой базе данных. Но если какая-либо информация присутствует, то она может оказаться полезной.

Остерегайтесь опросов, спонсированных политическими кандидатами, специальными группами и компаниями с очевидной заинтересованностью в результатах. Уточните, кто является непосредственным заказчиком самого опроса.

Далее мы хотели бы перечислить ряд преимуществ работы с авторитетным социологом или агентством.

1. СТАНДАРТНЫЕ МЕТОДЫ ВЫБОРКИ

Социологи выбирают кандидатов для опроса в соответствии с поставленными целями самого исследования. В зависимости от последующего анализа опроса, они могут составить выборку из 350, 500 или более 1 000 человек. Иногда они также идут на «перебор выборки», выбирая кандидатов, редко участвующих в опросах, как, например, молодежь, мужчин либо жителей северных регионов. Работа с авторитетным социологом во многом устранил ваше беспокойство по поводу использования соответствующих методов составления выборки опрашиваемых.

2. БЛАГОРАЗУМНОЕ ОСВЕЩЕНИЕ РЕЗУЛЬТАТОВ

Социологи, как правило, отказываются разбивать результаты опросов по таким малым группам, что смысл опроса просто теряется. Например, если вы поинтересуетесь ответами респондентов в возрасте от 20 до 25 лет, то, возможно, вы услышите: «Мы не проводим столь детальные опросы». И это хороший ответ, а не плохой. Не нужно вынуждать авторитетного социолога выходить за рамки опроса.

3. ВОПРОСЫ, ПРОВЕРЕННЫЕ ВРЕМЕНЕМ

Незначительные различия в вопросах могут привести к значительно различным результатам. Как показывают исследования, одно слово в вопросе о прошлых предпочтениях избирателя может существенно изменить прогноз социолога о возможной группе избирателей. Профессиональные социологи отслеживают подобные вопросы и применяют свой опыт при их формулировке.

ФОРМУЛИРОВКА ВОПРОСОВ И ИХ ПОСЛЕДОВАТЕЛЬНОСТЬ

Запрашивайте точную формулировку каждого вопроса. Более того, запрашивайте полный перечень использованных вопросов (или, по крайней мере, тот раздел, который представляет для Вас интерес), дабы увидеть в каком порядке был задан вопрос. Вот несколько пунктов, на которые стоит обратить внимание.

ЯСНОСТЬ. Если вопросы затруднительны для Вас, то они были непонятны и для опрашиваемых.

СДВОЕННЫЕ ВОПРОСЫ. Как и репортер в интервью, социолог может задать два вопроса в одном. В отличие же от репортера, социолог требует только одного ответа: «Поддерживаете ли вы или не поддерживаете реформы здравоохранения и строгий контроль цен?»

ПРОВОКАЦИОННЫЕ ВОПРОСЫ. Такие вопросы редко задаются авторитетными социологами, но они весьма распространены в практике организаций, преследующих свои интересы. Некоторые весьма очевидны: «Вы “за” или “против” президентской реформы в сфере здравоохранения, которая позволит снизить цены на медицинские услуги?». Некоторые – менее: «Голосовали ли вы на предыдущих выборах?» Подумайте, как бы вы лично ответили на такой вопрос.

СПОСОБНЫ ЛИ РЕСПОНДЕНТЫ ОТВЕТИТЬ НА ТАКОЙ ВОПРОС? Во время появления какого-либо значительного события в новостях социологов интересует, каков будет его эффект. Произойдет ли очередной теракт на американской земле? Будет ли мир на Ближнем Востоке? Хотя такие вопросы могут помочь в понимании общественного мнения по отношению к государственной политике, ответы на эти вопросы не имеют какого-либо существенного значения, так как большинство людей не может знать ответы на подобные вопросы.

СОЛГАЛИ ЛИ БЫ ВЫ, ЕСЛИ ВАС СПРОСИЛИ? Я живу в Вашингтоне, округ Колумбия. Проживание здесь близко к городу является

предметом гордости. Так, согласно статистическому бюро, среднее время, затрачиваемое на дорогу от моего места проживания (замеры производятся по почтовому индексу), составляет 20 минут. Но чтобы не опаздывать, и с расчетом пробок на дорогах и задержек метро, на деле мне надо выезжать на час раньше. Мой путь должен занимать 20 минут. Но такое практически никогда не случается.

ОБРАТИТЕ ВНИМАНИЕ НА САМУ ПОСЛЕДОВАТЕЛЬНОСТЬ ВОПРОСОВ. Вопрос на тему абортов, который следует сразу за вопросом об эвтаназии, может привести к определенному ответу. Тот же самый вопрос, заданный после вопроса о правах женщин, может вызвать иную реакцию и совершенно другой ответ. Уточните, каким образом был определен порядок вопросов.

ПРИМЕЧАНИЕ: Данный перечень был подготовлен по книге «Руководство по проведению опросов и анкетированию» Г. Кливленда Вилхойта и Дэвида Х. Уивера (Newsroom Guide to Polls and Surveys by G. Cleveland Wilhoit, David H. Weaver).

МЕТОДОЛОГИЯ ВЫБОРКИ И ПРОЦЕНТ ОПРОШЕННЫХ

Большое число респондентов необязательно для создания качественного исследования. Вместо этого вам будут необходимы другие важные характеристики.

1. МЕТОДОЛОГИЯ ВЫБОРКИ

Для отражения мнения большинства необходимо, чтобы каждая из участвующих в опросе групп имела равные шансы попадания в выборку. Подобная селекция не обязательно является «чисто случайной выборкой», однако она содержит множество её характеристик. В целом, статистикам необходима некоторая степень случайности. Они должны начинать работу с полного списка рассматриваемых респондентов. Как правило, это делается набором случайных цифр номера телефона, что дает каждому телефонному номеру шанс на попадание в выборку.

2. ПРОЦЕНТ ОПРОШЕННЫХ

Число респондентов, принимающих участие в телефонных опросах, зачастую составляет примерно от 60 до 70 процентов. Это означает, что многие отказываются от участия в опросе. Это не имеет большого значения только в тех случаях, когда отказавшиеся от участия в опросе люди ничем не отличаются от остальных его участников.

Например, если многие испаноговорящие респонденты не могут ответить на вопросы на английском языке, то любой опрос, который в той или иной степени затрагивает интересы этой социальной группы, может быть смещенным в пользу интересов англоязычной группы. Если в то время суток, когда проводится (телефонный) опрос, молодежь чаще находится вне дома, нежели респонденты старшего поколения, то, вполне возможно, что участие молодежи в исследовании будет не столь репрезентативным.

Уточните у социологов, каково соотношение опрошиваемых и имеет ли место какая-либо пристрастность среди тех, кто ответил.

ПРЕДЕЛ ПОГРЕШНОСТИ (ОШИБКИ В ВЫБОРКЕ)

Мы часто публикуем информацию о пределах допустимой погрешности или ошибках в выборке в тексте или в графиках. После этого мы моментально забываем об этой информации и игнорируем её как в материале, так и в заголовках. Этого делать не нужно.

Предел погрешности говорит о том, насколько социолог уверен в корректности предоставленных результатов. Погрешность в 2,5 процента означает, что предоставленные ответы находятся диапазоне 5 процентных пунктов от возможного истинного числа.

ПРИМЕР

В ходе опроса 45 процентов респондентов заявили, что будут голосовать за мэра Джонса, по голосам перегнавшего комиссара Смита, которого поддержали лишь 40 процентов опрошенных. Пятнадцать

процентов опрошенных затруднились ответить. Размер погрешности составил ± 3 процента. Это означает, что Джонса фактически поддерживают от 42 до 48 процентов респондентов, а Смита – от 37 до 43 процентов.

Неизвестно, кто бы вышел победителем, если бы выборы проводились сегодня. Есть шанс, что на самом деле Смит победит Джонса, от 43 до 42 процентов. Существует также вероятность (обычно в 5 процентов), что размер погрешности, заявленной в опросе, на деле значительно выше.

ЭЛЕМЕНТЫ ПОГРЕШНОСТИ ВЫБОРКИ

При определении допустимой погрешности используются три следующих элемента.

1. ЧИСЛО РЕСПОНДЕНТОВ ОПРОСА

Минимально допустимое число респондентов опроса составляет 385 человек. Многие опросы, которые охватывают респондентов по всей стране, выполняют это требование, поскольку число опрошиваемых составляет приблизительно 1500 человек.

Тем не менее предел допустимой погрешности должен быть использован от общего числа опрошенных внутри каждой группы, а не от общего числа всех респондентов, принявших участие в опросе. Если, например, в вашем опросе приняли участие только 50 афро-американцев, то он неадекватно отражает различия во мнениях этой группы лиц.

2. РАЗЛИЧИЯ В ОТВЕТАХ

Предел погрешности существенно уменьшается, когда 90 процентов респондентов придерживаются одного мнения, нежели, когда только 55 процентов выбирают схожие ответы.

Многие агентства не переживают по этому поводу: они просто предоставляют самую консервативную оценку предела погрешности – ту, которая соответствует распределению ответов 50 на 50. Таким

образом, составители опроса могут обобщить результаты опроса одной цифрой, нежели дифференцировать ответы респондентов по отдельным вопросам.

В фигурирующих в опросах цифрах имеет смысл уточнить, являются ли предоставленные оценки консервативными или агрессивными, или на каких ответах базируются эти оценки. Также имейте в виду, что вопросы с большим разбросом внушают больше доверия.

3. НАСКОЛЬКО В ТОЧНЫХ ОТВЕТАХ ЗАИНТЕРЕСОВАНЫ СОСТАВИТЕЛИ ОПРОСА

Как правило, социологи используют 95%-ный доверительный интервал. Другими словами, они оперируют исходя из предположения, что один из каждых 20 результатов будет выходить за рамки погрешности.

ВАШ САМОСТОЯТЕЛЬНЫЙ ПОДСЧЕТ ДОПУСТИМОГО ПРЕДЕЛА ПОГРЕШНОСТИ

Для выведения предела погрешности социологи могут использовать сложные вычисления с целью учета нетипичных выборок и других технических проблем. Не всегда возможно повторить эти расчеты. Но вы можете проверить основные результаты при помощи простой формулы, которая ничуть не уступает сложным математическим расчетам. Не переживайте, если ваш результат незначительно отличается от того, что предоставлен в исследовании. Однако, если результат сильно разнится (например, на более чем одну процентную единицу), то лучше уточнить, каким образом были проведены расчеты.

Можно воспользоваться следующей формулой:

$$1,96 \times \sqrt{((\text{«Да»} \times \text{«Нет»}) / (\text{Число респондентов}))}$$

Коэффициент 1,96 используется в тех случаях, когда применяется 95-процентный доверительный интервал. Это означает, что вы исходите из предположения, что одна из 20 анкет ошибочна и находится

вне диапазона погрешности. Увеличьте этот коэффициент до 3, если вы хотите быть на 99 процентов уверены, что полученные ответы респондентов не выходят за рамки допустимой погрешности.

«Да» – это доля (или процент / 100) респондентов, которые ответили «Да», или каким-то одним образом. «Нет» – это доля тех, кто ответил «Нет», или другим образом.

Рассмотрим три следующих примера:

ПРИМЕР 1	<p>В опросе о медицинском страховании приняло участие 535 жителей штата Миссури. Десять процентов заявили, что не имеют страховки. Остальные 90 процентов либо назвали тип своей страховки, либо затруднились его указать.</p> <p>Для этого конкретного вопроса предел погрешности составляет:</p> <p>Шаг 1: «Да» × «Нет» = $0,9 \times 0,1 = 0,09$</p> <p>Шаг 2: («Да» × «Нет») / число респондентов = $0,09 / 535 =$ (на калькуляторе или в таблице Excel отображается длинный ответ со знаком «е», который приблизительно составляет 0,000168)</p> <p>Шаг 3: Квадратный корень этого числа = 0,01297</p> <p>Шаг 4: Умножьте на 1,96 = 0,0254, или ±2.5процента. Так, в зависимости от того, насколько точен ваш опрос, вы можете утверждать, что от 7,5 до 12,5 процентов жителей штата Миссури не имеют медицинской страховки.</p> <p>Ввиду специфики подобной информации лучше придерживаться консервативных расчетов. Порой статистики предоставляют слишком консервативные цифры, или чересчур заниженные показатели, особенно тогда, когда они не хотят, чтобы их цитировали журналисты.</p>
-----------------	--

<p>ПРИМЕР 2</p>	<p>Предлагаем другой способ решить ту же задачу. Однако, в этот раз ваша цель – представить наиболее консервативную оценку, которая обобщит результаты всего исследования, а не какого-либо отдельного вопроса. Представленный способ является наиболее распространенным способом проведения таких расчетов.</p> <p>Шаг 1: Вместо «Да» × «Нет», используйте соотношение 50:-50:</p> <p>«Да» × «Нет» = $0,5 \times 0,5 = 0,25$ (здесь отражается консервативная оценка)</p> <p>Шаг 2: («Да» × «Нет») / Число респондентов = $0,25 / 535 = 0,000467$</p> <p>Шаг 3: Квадратный корень из этого числа = $0,0216$</p> <p>Шаг 4: Умножьте на 1,96 = $0,042$, или $\pm 4,2$ процентных пунктов. Результат надо выразить следующим образом: от 6 до 14 процентов людей не имеют медицинского страхования.</p>
<p>ПРИМЕР 3:</p>	<p>Придерживаясь консервативных расчетов, вы, наверняка, хотите заявить результат с точностью до 99 процентов:</p> <p>Шаг 1: «Да» × «Нет» = $0,5 \times 0,5 = 0,25$</p> <p>Шаг 2: («Да» × «Нет») / Число респондентов = $0,25 / 535 = 0,000467$</p> <p>Шаг 3: Квадратный корень этого числа = $0,0216$</p> <p>Шаг 4: Вместо коэффициента в 1,96, умножьте на коэффициент 3 (здесь отражается большая точность): $= 0,065$, или $\pm 6,5$ процентов. Теперь, с чуть большей уверенностью, вы можете утверждать, что примерно от 3,5 до 16,5 процентов жителей штата Миссури не имеют медицинского страхования. Однако подобный доверительный интервал используется в опросах редко.</p>

РАЗМЕРЫ ВЫБОРКИ И ИХ КОНСЕРВАТИВНЫЕ ОЦЕНКИ ПРЕДЕЛОВ ПОГРЕШНОСТИ

Существует один аспект в составлении опросов, который часто непонятен для новичков: численность населения, мнение которого необходимо оценить, не имеет значения. Не имеет значения: оцениваете ли вы настроение жителей всей страны или какого-либо отдельного города. Размеры необходимой выборки одинаковы в обоих случаях.

(Существует, однако, определенная корректировка, применяемая социологами в том случае, когда опрошен большой процент населения. Такое случается редко).

Самое важное, что нужно запомнить – это то, что при маленькой выборке респондентов невозможно измерить различия во мнениях, если даже таковые существуют.

С другой стороны, со слишком большой выборкой вы можете измерить «статистически значимые» различия, хотя они фактически тривиальны.

РАЗМЕР ВЫБОРКИ	ДОСТОВЕРНОСТЬ	
	95%	99%
385	5,0	7,6
500	4,4	6,7
750	3,6	5,5
1 000	3,1	4,7
1 500	2,5	3,9
3 500	1,7	2,5

Эта таблица показывает, почему многие опросы с охватом на всю страну, основываются на выборке примерно в 1500 респондентов. Такая выборка необходима, чтобы сократить размер погрешности до +/- 2,5 процентных пунктов.

ОПРОСЫ: ДОПОЛНИТЕЛЬНЫЕ ПОДСКАЗКИ

- «Не знаю» или «Нет мнения» являются ответами на вопросы, которые не стоит скидывать со счетов; особенно, если речь идет о предвыборных кампаниях, где число сомневающихся может быть важнее, нежели число тех, кто уже определился с выбором.
- Пусть Вас не пугает небольшой размер выборки. Лучше проанализируйте предел погрешности. Тысяча респондентов в опросе является вполне адекватным размером выборки, если только вопросы не разбиты по малым подгруппам. Для локальных опросов стандартный размер выборки составляет 350-400 респондентов.

В последние годы задача выявления тех граждан, которые намерены принять участие в голосовании, стала одной из наиболее сложных при проведении опросов. Множество различий среди результатов государственного голосования зависят от этого аспекта.

- Если в исследовании общественного мнения разработчик умалчивает размер выборки, методологию определения этой выборки, размер допустимой погрешности, доверительный интервал и точные формулировки задаваемых респондентам вопросов, то стоит ли опубликовывать результаты такого опроса? Доступность каких-либо из вышеперечисленных данных определяет степень профессиональности вашего источника.
- Убедитесь, что ваш материал не вращается вокруг бессмысленных различий или изменений. Помните, что заявленные изменения считаются статистически значимыми лишь в том случае, если они превышают размер установленной допустимой погрешности в два раза.
- Информация, представленная в опросах, может быть ошибочной! Предположим, что размер допустимой погрешности в ± 3 процентные единицы означает, что с вероятностью 95 процентов ответы достоверны в пределах 6 процентных единиц, (или, ска-

жем, от 41 до 47 процентов). Это означает, что с вероятностью в 5 процентов, возможно, что размер погрешности на деле гораздо выше. Если какая-либо информация в опросе не внушает вам доверия, не доверяйте его результатам!

ГЛАВА 6. 10 САМЫХ РАСПРОСТРАНЕННЫХ «ЛЯПОВ» В НОВОСТЯХ: ОТ ПРОСТЫХ МАТЕМАТИЧЕСКИХ ОШИБОК ДО ОШИБОК В СУЖДЕНИИ

Если до сих пор мы концентрировали внимание исключительно на корректном отражении чисел в новостях, то теперь настало время обсудить самые распространенные ошибки, которые допускаются при публикациях или выходах в эфир.

Вот список 10, на мой взгляд, наиболее часто встречающихся ошибок.

№ 10. ОБРАТНАЯ ВЕРОЯТНОСТЬ, ИЛИ КАКОВЫ ШАНСЫ ОБНАРУЖИТЬ КЛАСТЕР БОЛЬНЫХ РАКОМ?

Обратной вероятностью называют ту вероятность, которая была подсчитана от противного.

Вполне достоверно утверждать, что вероятность найти группу из семерых онкобольных в одном здании составляет один на миллиард. Но также достоверно и то, что вероятность обнаружить семь случаев заболевания раком в любом здании мира составляет 50 на 50.

Эта концепция сложна к восприятию и может быть лучше всего продемонстрирована на примере национальной лотереи. Если было продано 60 миллионов лотерейных билетов из 80 миллионов возможных комбинаций чисел, то вероятность вашего выигрыша составляет один из 80 миллионов. Однако вероятность любого выигрыша составляет 60 миллионов из 80 миллионов, или 75 процентов.

Другими словами, вы вполне уверены, что кто-то непременно выиграет, но вы просто не знаете заранее, кто это будет. Аналогичный принцип применим к раковым заболеваниям и ударам молнии: в этом году вы, вероятно, не заболеете раком, но это может произойти с кем-то, проживающим в вашем районе.

Чтобы выяснить, является ли редкое событие невозможным, вам надо вычислить вероятность того, что это событие не произойдет нигде, нежели, наоборот, произойдет где-то. Такие расчеты могут оказаться сложными, но вам могут помочь соответствующие компьютерные программы. Прежде чем опубликовывать данные по редким событиям, проконсультируйтесь с грамотными статистиками или эпидемиологами.

№ 9. ТРЮК С ВОЛШЕБНОЙ РЕЗИНКОЙ, ИЛИ ОШИБКА В ПОНИМАНИИ «ЗАКОНА СРЕДНИХ ЧИСЕЛ»

Представление о том, как после оттягивания волшебная резинка автоматически возвращается на исходную позицию, является одной из самых распространенных ошибок математического анализа происходящего, свойственных многим СМИ. В самом деле, часто слыша о так называемом «законе средних чисел», мы полагаем, что речь идет именно о таком событии. Математики часто называют этот эффект эффектом ошибки игрока. Он заключается в уверенности в том, что при следующей игре кубик выпадет в вашу пользу, исходя из того, что невыигрышные для Вас комбинации выпадали слишком часто.

Но на деле наблюдаемая ситуация объясняется понятием «регрессия к среднему», заключающемся в том, что за каждым нестандартным событием ожидается происхождение стандартного события. Лучшее описание этого явления, на мой взгляд, приведено в книге «Беседы о журналистике» Уильяма Зинсера («Speaking of Journalism» by William Zinsser). В этой книге автор Кевин Маккин (Kevin McKean), пишущий на научную тематику, приводит случай с психологом, который работал с летными инструкторами.

Один из летных инструкторов отказывался верить исследованиям, показывающим, что для повышения работоспособности поощрение дает лучшие результаты, чем наказание. Он аргументировал это тем, что всякий раз, когда он хвалил летчика за красиво выполненный маневр, следующий его маневр оказывался хуже преды-

дущего. Но после наказания за небрежно выполненную работу, результат был лучше.

Ответ здесь кроется в «регрессии к среднему»: при присутствии в событиях элемента случайности за неординарным событием, как правило, следует ординарное. Таким образом, как и за великолепным, так и за из рук вон плохим маневром зачастую следует более обычный.

И 8. СРЕДНЕЕ ОТ СРЕДНЕГО ЧИСЛА – ЛУЧШЕ НЕ ПЕРЕУСЕРДСТВОВАТЬ

Как правило, осреднение набора средних чисел не приводит к каким-либо осмысленным результатам. Избегайте подобных расчетов. Они не только могут ввести в заблуждение, но могут оказаться и в корне ошибочными.

Этот подход подобен излишнему перевариванию овощей – в полученной смеси сложно определить вкусовые ингредиенты и сложно понять, что было приготовлено. Хотя бывают и исключения из правил, при которых рассматриваются взвешенные средние значения, либо применяются другие способы.

Порой журналисты, работая с бюджетом, пытаются вычислить среднее арифметическое от «среднего процентного изменения». Это может привести, например, к тому, что 140-процентное увеличение в бюджете нового небольшого отдела департамента рассматривается эквивалентно 2-процентному увеличению бюджета во всей сфере образования.

Для дополнительной информации по осреднению обратитесь к разделу «Понимание средних – среднее арифметическое, медиана, мода» в Главе 2.

И 7. К ЧЕМУ ЭТО ПРИМЕНИМО? ИЛИ НЕ УПУСКАЙТЕ ИЗ ВИДУ ИСХОДНЫЕ ОБЪЕМЫ

Работа с малыми числами приводит к большим процентам. Работа с большими числами – к малым. Это означает, что отслеживать

фактическое число смертельных случаев, арестов, преступлений или каких-либо других событий не менее важно, чем освещать результирующие показатели этих событий или степень их изменения.

Как правило, сложно указывать процентное значение, если число рассматриваемых предметов, будь то людей, домов или долларов, меньше 100. Это происходит потому, что в этом случае любое изменение в процентном соотношении будет составлять больше одного процентного пункта. Поэтому представьте информацию, исходя из начальных данных, как, например, 1 из 12, чтобы сделать цифры более осязаемыми для восприятия. Если и это не поможет, то обратитесь к округлению, описывая свой результат как «более половины» или «почти две третьих».

Н 6. ОШИБОЧНАЯ ТОЧНОСТЬ

По сути, любые числа являются обобщениями, в какой-то степени построенными на предположениях. Минуту спустя после сбора данных, они перестают быть актуальными.

Из этого не следует, что нужно избегать доступных нам чисел, но все же доля предосторожности при работе с ними не повредит. Это означает, что лучше не увлекаться чрезмерной точностью, которой грешат некоторые из журналистских материалов.

Вы можете проверить, присутствует ли эффект чрезмерной точности в вашем материале, путем округления любого присутствующего в нем числа.

Некоторые отделы новостей придерживаются стилистических правил, налагающих ограничения на допустимое число знаков, которые можно оставлять после запятой. Вам, возможно, придется следовать этим правилам. Но публикацию стоит переосмыслить, если идея вашего материала зависит от точности представления чисел с запятой.

Количество чисел после запятой может быть, например, важно, если речь идет об уровне безработицы или уровне преступности. Из предыдущих примеров мы знаем, как изменение исходных данных может обратить малое число в большое. Конечно, если при этом результаты математически аккуратны, то можно не переживать.

Но, зачастую, данные не соответствуют уровню аккуратности, который предполагается в материале. Вместо этого, многие результаты основаны на небольших выборках, округленных данных и всевозможных предположениях.

Вспомните приведенный ранее пример расчетов корректировки на инфляцию. Используя один тип Индекса Потребительских Цен (ИПЦ) (Consumer Price Index, CPI), мы можем вычислить зарплату учителя в размере \$53234 в год. Однако если выразить ИПЦ в других масштабных рамках, то можно получить \$53226 – с общей разницей в \$8. Но, если округлить результат до сотых, то в обоих случаях это будут те же \$53200.

Создавая материал, избегайте зависимости от бессмысленных отклонений в числовых данных. Такие отклонения будут оспорены, и эта критика будет вполне правомерна.

Н 5. ЕСЛИ БЫ, ТО ТОГДА ...: ПУТАНИЦА МЕЖДУ ПРИЧИНОЙ И СЛЕДСТВИЕМ

Критики, как правило, преувеличивают степень того, насколько журналисты смешивают понятие корреляции с понятием причинно-следственной связи. Тем не менее повторное освещение этого вопроса может оказаться полезным, так как многие критики твердо уверены в том, что журналисты не понимают различие между корреляцией – какой-либо взаимосвязи между случайными величинами – и реально существующей причинно-следственной связью между случайными величинами, освещение которой и является целью наших материалов.

Корреляция – это математическая концепция. Например, если с одной стороны уровень бедности повышается, а с другой стороны

одновременно снижается уровень успеваемости среди студентов, то две эти случайные величины коррелируют. Таким образом, можно предположить, что, зная уровень благосостояния студента, возможно предсказать его успеваемость. Но таким же образом можно предположить и обратное: на основе уровня успеваемости студента можно предсказать уровень его благосостояния.

В учебниках собрано множество примеров, описывающих путаницу между корреляцией и причинно-следственной связью, но эти примеры мало применимы в журналистике. Одним из таких примеров является пример о взаимосвязи между числом священнослужителей и числом проданных бутылок ликера по городам. Здесь, безусловно, присутствует недостающее звено – число горожан, проживающих в каждом городе.

Ричард М. Джагер, автор книги «Статистика: Зрелищный вид спорта», («Statistics: A Spectator Sport» by Richard M. Jaeger) приводит примеры того, как сами исследователи определяют наличие причинно-следственных связей.

а). НАСТУПАЕТ ЛИ ПРИЧИНА РАНЬШЕ СЛЕДСТВИЯ?

Этот вопрос кажется более очевидным, чем он есть на самом деле. Уровень успеваемости может зависеть от числа преподавателей в классе. Расходы на преподавателей могут, в свою очередь, зависеть от успеваемости только в прошлом году. Однако уровень успеваемости обычно так быстро не изменяется.

б). СУЩЕСТВЕТ ЛИ ТРЕТИЙ ФАКТОР, ВЛИЯЮЩИЙ НА ВЗАИМОСВЯЗЬ?

Опять же, уместно рассмотреть данный аспект на примере школы. Большинство журналистов согласится с тем, что в школах, где обучаются представители расовых меньшинств, наблюдается низкий уровень успеваемости. Однако при этом не стоит забывать о третьем факторе, который оказывает влияние как и на причину, так и на следствие в этом случае, а именно - бедности и низком уровне образования родителей. Хотя этническая принадлежность играет наиболее важную роль, немногие смогут утверждать, что именно

она является причиной трудностей, с которыми сталкиваются расовые меньшинства в школах.

с). ЕСТЬ ЛИ ДРУГОЕ ОБЪЯСНЕНИЕ?

Можете ли вы (или кто-нибудь другой) создать опровергающую теорию, одинаково хорошо объясняющую происходящее. На этом этапе вместо того, чтобы обращаться к статистике, приложите больше усилий к сбору дополнительной информации по материалу.

d). СУЩЕСТВУЕТ ЛИ ЧТО-ТО, ЧТО ОПРОВЕРГАЕТ ВАШУ ТЕОРИЮ?

Есть ли какие-либо другие исследования, которые не подтверждают ожидаемую причинно-следственную связь? Почему ваша теория работает лишь в частном случае, но не в общем? Почему результаты являются верными в одном году, но могут оказаться ошибочными в другом?

И 4. ПРЕНЕБРЕЖЕНИЕ ВНУТРЕННИМ ГОЛОСОМ

Журналисты порой буквально доверяют числам, игнорируя все то, что было достоверно известно заранее. Наиболее часто такое происходит в случае с опросами, которые могут быть ошибочными. Как советует один автор, пропускайте каждую цифру через так называемый «детектор ерунды».

Если цифра не соответствует здравому смыслу, либо вам известна информация, которая ей противоречит, то она может быть в действительности неверной.

И 3. НЕАДЕКВАТНЫЕ ИСТОЧНИКИ

Признаюсь честно: мне самой не раз приходилось в отчаянии искать то самое или хоть какое-то число, которое могло бы удовлетворить моего редактора, стремящегося выстроить материал в определенном ракурсе.

Зачастую мы смиряемся с некачественными данными, поскольку нам неловко оспаривать наши источники информации. Или

поскольку данные скучны, а нам важнее написать материал, не углубляясь при этом в дебри исследования. Или поскольку эти данные будоражат воображение, гарантируя этим более интересную историю.

В начале 1990-х годов в администрации президента Клинтона, среди экспертов по вопросам труда, а также среди всех тех лиц, которые сетовали на низкий уровень корпоративной лояльности, была популярна одна цифра. Все цитировали одну и ту же статистику, согласно которой работающие люди меняют место работы шесть раз за период своей профессиональной деятельности. Как и многие другие журналисты, освещающие эту тему, я задалась целью найти первоисточник этой статистической информации. Как выяснилось, источником явилось плохо проведенное исследование, посвященное совсем другой тематике. Но цифры, некогда процитированные, стали уже частью фольклора.

Не вздыхайте попросту с облегчением, когда цифра, необходимая для материала, получена. Поинтересуйтесь у своего источника: «Откуда у Вас эта информация? Кто её может подтвердить?». Включите свой здравый смысл и найдите первоисточник.

N 2. ПЕРЕПОЛНЕНИЕ МАТЕРИАЛА ЧИСЛАМИ

Одна из наших самых больших ошибок – это стремление уместить в материал как можно больше чисел, нежели выбрать лишь те, которые лучше раскрывают суть материала. Я думаю, что это делается в попытке справедливо отразить происходящее – если каждое число представляет отдельное мнение, то точки зрения должны быть сбалансированы.

Но теперь, когда вы более виртуозны в обращении с цифрами, попробуйте существенно сократить их присутствие в своем материале. Как писал Уильям Бланделл (WilliamBlundell), переполение материала числами является тем самым бесом, который приводит к большому количеству непрочитанной прозы, чем какие-либо другие причины. Не позволяйте этому случиться с Вами.

Н 1. ПОЗВОЛИТЬ ПРИЧИНАМ ОТ 10 ДО 2 ВВЕСТИ НАС В СТУПОР

Порой мы позволяем нашему непониманию цифровых данных фактически свести нас до положения стенографистов, в особенности в материалах на тему медицины или образования.

Я надеюсь, что, благодаря этой книге, вы теперь с большей уверенностью сможете читать отчеты экспертов, вместо того, чтобы просто цитировать их пресс-релизы. Я также надеюсь, что теперь вы понимаете, что цифры – это просто один из способов выражения мнения, резюмирования какой-то информации или раскрытия материала. Если прочитав это пособие, вы сделали для себя какие-то выводы, то я надеюсь, что вы осознали то, что большая часть работы с числами основывается на здравом смысле. Используйте это знание.

ГЛАВА 7: ЛОТЕРЕИ, УДАРЫ МОЛНИИ, ПРОДОЛЖИТЕЛЬНОСТЬ ЖИЗНИ – КОММЕНТАРИИ НА ТЕМУ ВЕРОЯТНОСТИ

Большинство книг по математике, в особенности книги профессора математики Университета Темпл в Филадельфии Джона Аллена Паулоса (JohnAllenPaulos), рассматривают понятие вероятности – понятие о том, что на основе либо исторических тенденций, либо известных количественных данных возможно предсказать вероятность наступления какого-либо события. Эта глава будет краткой, поскольку в новостях вероятности затрагиваются редко. Тем не менее, важно знать, как с ними работать.

ВЫЧИСЛЕНИЕ ВЕРОЯТНОСТИ

Вероятность вычисляется путем использования отношений. Рассмотрим следующие примеры:

<p>ВЕРОЯТНОСТЬ УДАРА МОЛНИЕЙ</p>	<p>В регионе Тампа Бэй (Тампа Bay), где проживает около 4,3 млн. людей, каждый год приблизительно 20 человек попадают под удары молнии.</p> <p>Так, вероятность пострадать от удара молнии составляет $20 / 4\,300\,000$, или $0,000005$ в год.</p> <p>Поскольку это число сложно для восприятия, многие его упрощают путем рассматривания «одного случая из» определенного количества случаев.</p> <p>Дабы получить это определенное количество случаев разделите единицу на ваш результат: $1/0,000005 = 200000$</p> <p>Таким образом, вероятность удара молнией составляет один случай из 200 000.</p>
---	---

**ВЕРОЯТНОСТЬ
СМЕРТИ ОТ
РАКА**

Уровень смертности от рака зависит от возрастной категории пациентов, поэтому в качестве примера давайте выберем какую-то определенную возрастную группу. Так, в 2009 году приблизительно 6 000 женщин в возрасте от 45 до 54 лет умерли от рака молочной железы. 22,6 млн. женщин по всей стране принадлежали тогда к этой возрастной категории.

Чтобы получить уровень смертности, разделите число людей, умерших от рака, на число людей, находящихся в данной возрастной категории:

$$6000 / 22\,600\,000 = 0,000\,265$$

Получившееся число и есть вероятность смертности от рака молочной железы среди женщин в возрасте от 45 до 54 лет. (Вероятность выражается пропорцией и всегда находится в промежутке от 0 (нет шансов) до 1 (абсолютная уверенность)).

Чтобы получить уровень смертности, умножьте результат на 100 000:

$0,000\,265 \times 100\,000 = 26,5$, что означает, что уровень смертности составляет приблизительно 27 человек из 100 000 женщин.

Иногда те же результаты можно выразить гораздо проще, разделив единицу на полученную вероятность:

$$1 / 0,000\,265 = 3\,774.$$

Это означает, что для женщин, находящихся в рассмотренной возрастной категории, шансы умереть от рака молочной железы составляют один из 3 800.

	<p>В большинстве случаев вероятность смертности выражается на основе ежегодных подсчетов. Будьте осторожны с вычислением риска заболеваемости на протяжении всей жизни. Подобные вычисления намного сложнее и зависят от предположений об общей продолжительности жизни человека.</p>
--	---

ВЫИГРЫШ В ЛОТЕРЕЮ

На оборотной стороне лотерейного билета можно найти информацию о вероятности выигрыша в эту лотерею. В этом разделе вы узнаете, как вычисляются такие вероятности.

Шансы на выигрыш в лотерею основаны на «комбинаторике» – области математики, в которой производятся подсчеты объектов. Это те самые факториалы, комбинации и перестановки, которые послужили причиной нашего отторжения от математики где-то в районе пятого класса.

Чтобы выяснить выигрышные шансы лотерейного билета, для начала вам необходимо узнать, из какого диапазона и какое количество чисел выбираются в игре.

Шанс выигрыша в лотерею всегда составляет один шанс из числа всех возможных комбинаций чисел лотерейного билета. Поэтому, если выпущено в общей сложности 2 млн. всевозможных лотерейных комбинаций, то ваш шанс на выигрыш составляет 1 из 2 млн.

Поймите, что шансы на выигрыш всегда одинаковы от одного выбранного числа к другому и от одной лотереи к другой. Если какое-то число было выбрано прежде, это не означает, что его не выберут снова. Само число, за исключением тех случаев, когда имеет место обман или мошенничество, не имеет значения при расчете шансов.

**ВЫЧИСЛЕНИЕ
ЛОТЕРЕИ**

Итак, рассмотрим типичный пример, позаимствованный из книги «200% of Nothing» («200% из ничего»): в лотерее «Pick 6» выбирается комбинация из шести чисел в диапазоне от 1 до 36. Выбранные числа не должны повторяться, и их последовательность не играет роли.

На месте первого числа вы можете выбрать любое из 36 доступных чисел. Но при выборе второго числа у Вас останется только 35 чисел. При выборе третьего – только 34. Другими словами, для каждой из 36 комбинации выбора первого числа существует 35 комбинаций выбора второго числа и так далее.

Всякий раз, когда вы слышите фразу «для каждой», знайте, что речь идет об умножении. Таким образом:

$36 \times 35 \times 34 \times 33 \times 32 \times 31 =$ число возможных комбинаций выбора 6 чисел из 36.

Но этот расчет не учитывает одного важного аспекта: многие комбинации учитываются дважды, поскольку, например, числа 1, 17, 27, 2, 18, 30 являются в этом расчете комбинацией, отличной от 18, 2, 17, 30, 27, 1. Нам, однако, безразличен порядок выбранных чисел, нам важны только сами числа на лотерейном билете.

Таким образом, мы сократим полученный результат, разделив его на:

$6 \times 5 \times 4 \times 3 \times 2 \times 1$.

Поскольку существует шесть способов для выбора первого числа; пять способов – для второго числа и т.д.

Окончательный результат составляет

$(36 \times 35 \times 34 \times 33 \times 32 \times 31) / (6 \times 5 \times 4 \times 3 \times 2 \times 1)$

	или $1\,402\,410\,240 / 720 = 1\,947\,792$ возможных лотерейных билетов
<p>ДРУГОЙ ПРИМЕР КОМБИНАЦИЙ</p>	<p>Рассмотренную концепцию следует использовать при вычислении количества возможных событий.</p> <p>Как-то в пятницу вечером в библиотеку IRE обратился журналист со следующим вопросом:</p> <p>Каким образом можно рассчитать число решений, которые принимает менеджер по бейсболу при расстановке команды на поле? Другими словами, сколько возможных комбинаций можно составить из девяти игроков, если в команде их 25?</p> <p>Исходя из рассуждений, приведенных в предыдущем примере, мы получаем:</p> $(25 \times 24 \times 23 \times 22 \times 21 \times 20 \times 19 \times 18 \times 17) / (9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1).$ <p>Вы можете рассчитать точное число. Оно составляет более 2 млн. возможных комбинаций игроков без учета их позиции. (Если вы хотите рассчитать каждую возможную позицию как отдельную комбинацию, то пропустите деление. В результате вы получите около 740 миллионов возможных ответов.)</p> <p>Из этого можно сделать следующий вывод: простые комбинации чисел могут быстро возрасти до огромных числовых значений за счет рассмотренных нами умножений.</p>

СОВЕТЫ ДЛЯ ПОЛЬЗОВАТЕЛЕЙ ЭЛЕКТРОННЫХ ТАБЛИЦ EXCEL

В Excel есть пара полезных функций, которые облегчат вам жизнь. Посредством функции COMBIN (# возможных чисел в выборке, # выбранных чисел) можно осуществить рассмотренные нами расчеты. Так, в примере с бейсболом мы используем COMBIN(25,9), а в примере с лотереей – COMBIN(36,6).

Если вы хотите учитывать последовательность выбранных чисел, то используйте функцию PERMUT (# возможных чисел в выборке, # выбранных чисел). Если же вы хотите произвести расчеты, необходимые для получения знаменателя, то используйте функцию FACT(#), которая возвращает факториал числа.

МАЛЕНЬКИЕ ВЕРОЯТНОСТИ X БОЛЬШИЕ ЧИСЛА = ГРУБЫЕ ОШИБКИ

В статистике большие числа приводят к большим числам, даже при небольшом процентном соотношении. Классическими примерами таких ситуаций являются примеры с допинг контролем и судебные экспертизы.

Давайте рассмотрим наиболее простой пример по экспертизе на наличие наркотических средств в крови. Помните, что, если эксперты говорят, что вероятность происшествия того или иного события составляет один случай из какого-то неисчислимого количества случаев, то непременно важно знать, сколько всего этих самых неисчислимых случаев было рассмотрено.

Следующий пример вы встретите и во многих других математических пособиях:

Крупная компания, где работает 5 000 человек, планирует протестировать всех сотрудников на употребление наркотиков. Около 10 процентов сотрудников в действительности употребляют наркотики. Точность результатов теста составляет 98 процентов, это означает, что в 98 процентах случаев при «положительных» резуль-

татах протестированные действительно употребляют наркотики, и наоборот.

При таком наборе исходных чисел, получается, что один из пяти протестированных с положительным результатом на самом деле не употребляет наркотики:

1. Из 5 000 человек 4 500 не употребляют наркотики. Тест корректно прогнозирует, что 4410 человек не употребляют наркотики и ошибочно прогнозирует, что оставшиеся 90 человек употребляют.
2. Из 5000 человек 500 употребляют наркотики. Тест верно определяет, что 490 человек употребляют наркотики, и что 10 – нет.
3. При «положительном» результате тестирования вероятность отсутствия наркотиков в крови составляет: $90 / (490+90) = 0,18$, или почти 1/5.

СОВПАДЕНИЯ С ТЕЧЕНИЕМ ВРЕМЕНИ – КОМБИНИРОВАНИЕ ВЕРОЯТНОСТЕЙ

На мой взгляд, большие совпадения всегда являются поводом для новостного материала, и нет ничего ошибочного в рассмотрении их подобным образом. В этом аспекте мое мнение расходится с мнением экспертов, которые считают, что совпадения нужно игнорировать до тех пор, пока нет весомых доказательств тому, что они являются чем-то большим.

Тем не менее, полезно знать, насколько совпадение является на самом деле совпадением.

Чтобы выяснить вероятность на протяжении какого-то времени, необходимо перемножить вероятности между собой. Но будьте внимательны – зачастую не имеет значения, где и когда событие может произойти, а только может ли оно произойти вообще.

БИРЖЕВЫЕ БРОКЕРЫ

Одно из условий торгов на фондовом рынке заключается в том, что вся необходимая информация должна быть доступна для всех участников рынка – было бы только время и силы ее изучать. Цены на акции, предположительно, являются результатом коллективной мудрости самого рынка. Поэтому никто из участников не должен иметь преимущества по отношению к другим, за исключением преимуществ в виде умения понимать эту мудрость лучше, чем другие.

Когда управляющий фондом виртуозно «обходит» рынок пять лет подряд, то его стратегии и «секретам» мы посвящаем целые материалы. Но все это может быть просто слепой удачей.

Как насчет того, чтобы начать с предположения, что событие обыгрывания рынка в текущем году является результатом слепой удачи. Вероятность вашей победы составляет 0,5, или 1 из 2, или 50 процентов.

Таким образом, в первый год 50 биржевых брокеров из 100, смогут обыграть рынок, и, соответственно, 50 из них останутся за бортом. Далее возьмём за основу следующий год: 50 процентов из предыдущей группы победителей выигрывают снова. Таким образом, число тех, кто одержал победу в течение двух лет подряд, будет составлять 25 процентов. Еще год спустя, т.е. уже на третий год, число трехкратных победителей будет составлять 12,5 процентов. И еще через два года – их останется только 3 процента.

Таким образом, если исходить из предположения слепой удачи, 3 из 100 гуру рынка способны выигрывать на протяжении пяти лет подряд.

ЧИСЛО СТРАЙКОВ

В истории профессионального бейсбола 1998 год был годом многочисленных «страйков» (мяч, поданный в зону удара и не затрону-

тый битой отбивающего). Марк Макгвайр (MarkMcGwire) и Самми Соса (Sammy Sosa) побили все предыдущие рекорды по выбиванию хоум-ранов.

Но из года в год репортеры, освещая бейсбол, чаще всего фокусируются на мелких страйках: пять игр подряд с хоум-ранами или 10 игр без единого хоум-рана.

Для простоты подсчетов и восприятия, предположим, что игрок выбивает 35 хоум-ранов в 140 играх. Это означает, что за каждую игру он должен выбивать 0,25 хоум-рана, или один хоум-ран в каждой из четырех игр.

Давайте сначала решим простую задачу: каковы шансы бейсболиста выбить хоум-ран на протяжении 10 игр подряд. Решением является простое умножение 0,25 самого на себя 10 раз, что приводит к вероятности, равной одному из более миллиона.

Обратите внимание, что случаи, когда 10 игр подряд проходят без единого хоум-рана, более распространены. Чтобы произвести эти расчеты, необходимо определить вероятность отсутствия хоум-рана на протяжении нескольких игр подряд. Это необходимо, поскольку нам не важно, в какой из игр может случиться хоум-ран. Имеет значение только его отсутствие.

В первой игре шансы отсутствия хоум-рана составляют 0,75 или 3/4. Во второй игре эта вероятность составит $0,75 \times 0,75$. И так далее – пока вы не умножите 0,75 само на себя 10 раз, что, в итоге, даст результат в 0,05, что эквивалентно одному из 20 случаев того, что во всех 10 играх не будет ни одного хоум-рана.

ВСПЫШКИ ОНКОЗАБОЛЕВАНИЙ

Пример со вспышками онкозаболеваний я оставила напоследок, поскольку логика вычисления вероятности этих событий схожа с идеей расчета крайне редких «холодных» страйков (когда подающий не наносит удара при подаче).

Но прежде чем продолжить, повторите вслух несколько раз: «Вспышки онкозаболеваний не будет!»

Вполне возможно, что в одном и том же городе / регионе вы обнаружите множество случаев заболевания раком, но это не является подтверждением наличия кластера раковых заболеваний. Документирование вспышки онкозаболеваний – это очень сложная процедура, и эксперты не берутся давать оценку до тех пор, пока они не владеют следующей информацией для анализа ситуации:

- Четкая информация о тех лицах, которые были подвержены влиянию рассматриваемого вредного вещества. Это означает, что вам нужно знать всех подверженных влиянию людей, а не только тех, кто заболел.
- Тип рака должен быть одинаковым среди всех случаев, либо источник заражения должен быть единым. Вам будет затруднительно исследовать кластер больных раком, который одновременно содержит больных с раком молочной железы и больных с детской лейкемией.
- Должны существовать веские медицинские доказательства того, что рак был спровоцирован каким-то из вредных веществ, влиянию которого были подвержены люди. Недостаточно знать, что вспышка лейкемии произошла неподалеку от свалки токсичных отходов. Вместо этого нужно также уметь доказать, что в этих отходах содержатся вредные вещества, вызывающие рак.
- Наблюдаемый уровень заболеваемости раком на деле значительно выше, чем ожидаемый уровень в масштабах рассматриваемого населения этого города / региона.

Большинство из этих аспектов журналистам доказать весьма непросто. Случаи из книги «Гражданский иск» («ACivilAction») являются этому подтверждением.

Как правило, мы, журналисты, способны осветить только последний аспект в предложенном списке: является ли наблюдаемый уровень заболеваемости раком выше ожидаемого? Зачастую производимые подсчеты аналогичны подсчетам для вычисления уровня относительного риска, которые были приведены ранее в этой книге.

В этих случаях статистики, как правило, проводят вероятностный тест для того, чтобы установить, являются ли наблюдаемая разница результатом случайности.

Но они также производят другие расчеты, как, например, расчет вероятности того, что в городе Уоберн штата Массачусетс никогда не случится такое большое количество случаев заболевания раком одновременно. Эти подсчеты основываются на вероятностном распределении, которую статистики называют «Таблицей средней продолжительности жизни». Есть специальные компьютерные программы, которые производят эти расчеты. Но концептуально в расчетах заложена та же идея, что и при расчетах вероятности «холодного» страйка в бейсболе. Вместо того, чтобы выяснять, какова вероятность возникновения 25 случаев детской лейкемии в городе Уоберн, они вычисляют, каковы шансы наличия стольких же случаев этого заболевания в любом другом городе с той же численностью населения. Убедитесь, что вы понимаете, о чем конкретно идет речь.

ГЛАВА 8. ДОПОЛНИТЕЛЬНЫЕ РЕСУРСЫ ДЛЯ ЖУРНАЛИСТОВ

Об Ассоциации журналистов-расследователей и редакторов

Ассоциация журналистов-расследователей и редакторов (Investigative Reporters and Editors, IRE) и Национальный институт по журналистике данных (National Institute for Computer-Assisted Reporting, NICAR) предоставляют огромный спектр обучающих материалов для журналистов.

ОБУЧЕНИЕ

ИНТЕНСИВНЫЕ КУРСЫ

Помимо семинаров, где обучают журналистов как работать в электронных таблицах, базах данных и с различными интернет-ресурсами, IRE ежегодно проводит продвинутый курс по статистическому анализу данных.

ОБУЧЕНИЕ РЕДАКЦИЙ СМИ

IRE может отправить тренерский десант в вашу редакцию, пренеся в ее работу новые знания, технологии и профессиональные навыки: от журналистики данных до статистического анализа и визуализации информации.

СЕРИЯ ПРАКТИЧЕСКИХ МАТЕРИАЛОВ

Для тех, кто ранее проходил обучение на наших тренингах, открыт доступ к различным практическим материалам по журналистике данных. Есть примеры базы данных для Excel, Access, FoxPro и SPSS.

РЕСУРСНЫЙ ЦЕНТР**ИСТОРИИ ПРАКТИЧЕСКИЕ СОВЕТЫ**

Ресурсный центр IRE (<http://ire.org/resource-center/>) является крупнейшей научной библиотекой, содержащей порядка 26000 журналистских расследований из периодической печати, ТВ и онлайн-СМИ. К этому перечню еще добавьте более 4 000 докладов, выступлений, практических советов спикеров – участников конференций IRE, и вы поймете, что в ваших руках – кладезь информации, которую вы можете использовать при написании своих материалов. Вся эта информация также доступна через Интернет.

БИБЛИОТЕКА ИНФОРМАЦИОННЫХ БАЗ ДАННЫХ

IREсовместно с «NICAR» постоянно обновляют библиотеку правительственных данных. Журналисты могут приобрести целые базы данных или запросить фрагменты уже отформатированных баз данных, информацию с которых они могут использовать в своих материалах. Например, это могут быть: статистика преступности, данные по авиации, сведения об авариях на дорогах и даже – федеральные контракты с частными компаниями. Наши библиотекари всегда готовы помочь вам решить любые головомки с цифрами.

ПУБЛИКАЦИИ**UPLINK**

Uplink – онлайн-журнал, выпускаемый «NICAR», где журналисты делятся своим опытом использования баз данных, электронных таблиц, статистических данных, технологий геолокации и др. Информация для подписчиков: <http://ire.org/blog/uplink/>

ЖУРНАЛИРЕ

IRE выпускает ежеквартальный журнал. Члены Ассоциации получают его бесплатно, но подписку может оформить каждый, даже не имея специального членства в организации. В журнале публикуются материалы тех авторов, которые использовали методы журналистских расследований для создания своих материалов, чтобы тем самым передать свои знания коллегам по цеху.

КНИЖНАЯ СЕРИЯ IRE

Книга, которую вы сейчас читаете, является частью серии справочных пособий для журналистов по математике и работе с цифрами.

ОНЛАЙН-РАССЫЛКИ

«NICAR-L» и «Census-L» – это электронные рассылки, благодаря которым журналисты помогают друг другу в вопросах с цифрами, – как правило, в режиме реального времени, ввиду сжатых сроков на подготовку материалов. Подписаться можно на сайте IRE: <http://www.ire.org/resource-center/listservs/>.

В ИНТЕРНЕТЕ

Интернет расширяет доступ журналистов к информационным справочникам по математике и статистике – особенно в тех случаях, когда ответ на вопрос нужно получить в максимально сжатые сроки. Вот список полезных ресурсов:

www.census.gov/popest

Бюро переписи населения США содержит статистическую информацию по населению США и Пуэрто Рико.

pwww.robertniles.com

Отличный сайт в помощь при решении математических задач, где также содержатся практические рекомендации и ссылки на статистические данные.

<https://www.cia.gov/library/publications/the-world-factbook/appendix/appendix-g.html>

Справочник фактов Центрального Разведывательного Управления США (ЦРУ)

ДРОБИ И ПРОЦЕНТЫ – СРАВНИТЕЛЬНАЯ ТАБЛИЦА

Дробь	Числовой эквивалент	Словесный эквивалент	Процент	Лучше для восприятия читателей
1/10		1 из 10	10%	Десятая часть
1/9		1 из 9	11%	Девятая часть
1/8		1 из 8	13%	Восьмая часть
1/7		1 из 7	14%	Седьмая часть
1/6		1 из 6	17%	Шестая часть
1/5	2/10	1 из 5	20%	Пятая часть
2/9		2 из 9	22%	Почти четверть
1/4	2/8	1 из 4	25%	Четверть
2/7		2 из 7	29%	Порядка четверти
3/10		3 из 10	30%	Почти одна треть
1/3	2/6, 3/9	1 из 3	33%	Треть
3/8		3 из 8	38%	Порядка одной трети
2/5	4/10	2 из 5	40%	
3/7		3 из 7	43%	
4/9		4 из 9	44%	Почти половина
1/2	5/10; 4/8; 3/6; 2/4	1 из 2	50%	Половина
5/9		5 из 9	56%	Более половины
4/7		4 из 7	57%	Более половины
3/5	6/10	3 из 5	60%	
5/8		5 из 8	63%	Почти две трети
2/3	4/6, 6/9, 6/8	2 из 3	67%	Две трети
7/10		7 из 10	70%	
5/7		5 из 7	71%	
3/4		3 из 4	75%	Три четверти
7/9		7 из 9	78%	Более чем три четверти
4/5	8/10	4 из 5	80%	
5/6		5 из 6	83%	

6/7		6 из 7	85%	
7/8		7 из 8	88%	
8/9		8 из 9	89%	
9/10		9 из 10	90%	
1		100%		Ровно столько же
2		200%		Дважды
3		300%		Трижды

ОБ АВТОРЕ

Сара Коэн (Sarah Cohen) – редактор отдела по журналистике данных в «Нью-Йорк Таймс» (The New York Times). Ранее преподавала в Университете Дьюка (Duke University), будучи обладателем стипендии им. Джона Найта, также работала редактором по информационным данным в газете «Вашингтон Пост» (The Washington Post). Была удостоена премии Пулитцера (Pulitzer Prize) за коллективный проект, является обладателем премии Голдсмита (The Goldsmith Awards) в области расследовательской журналистики и премии IRE (IRE Awards). Она также работала в качестве репортера в штате Флорида, была директором IRE по тренинговым программам с 1996 по 1998 годы. С 2010 года входила в совет директоров IRE.

СОВЕТ ДИРЕКТОРОВ IRE (2014 Г.)

Дэйвид Кей Джонстон, Сиракузский Университет
(David Cay Johnston, Syracuse University)Президент

Сара Коэн, «Нью-Йорк Таймс»
(Sarah Cohen, The New York Times)Вице-президент

Эллен Гэблер, «Милуоки Джорнал Сентинель»
(Ellen Gabler, Milwaukee Journal Sentinel).....Секретарь

Эндрю Донохью, Центр Журналистских Расследований
(Andrew Donohue, The Center for Investigative Reporting) Казначей

Мак Нелли Торрес, «Эн-Би-Си-Майами»
(Mc Nelly Torres, NBC Miami)..... Член исполнительного комитета

Зива Бранстеттер, «Талса»
(Ziva Branstetter, Tulsa)

Ворлд Роберт Крибб, «Торонто Стар»
(World Robert Cribb, Toronto Star)

Леонард Доуни мл., «Вашингтон Пост» / Университет штата Аризона
(Leonard Downie Jr., The Washington Post / Arizona State University)

Мэнни Гарсия, «Нэйплс Дэйли Ньюс»
(Manny Garcia, Naples Daily News)

Мэтт Голдберг, «Эн-Би-Си», г. Сан-Франциско
(Matt Goldberg, NBC Bay Area)

Джош Мейер, Журналистская инициатива по национальной
безопасности «Medill»
(Josh Meyer, Medill National Security Journalism Initiative)

Кристиан Миллер, «ПроПублика» (Christian Miller, ProPublica)
Стюарт Уотсон, «Дабл-Ю-Си-Эн-Си», г. Шарлотт (Stuart Watson,
WCNC-Charlotte)

ШТАТ РУКОВОДИТЕЛЯ IRE

Марк Хорвит (Mark Horvit) Исполнительный директор

КОНТАКТНАЯ ИНФОРМАЦИЯ

IRE и библиотека базы данных «NICAR» 573-884-7711

Ресурсный центр IRE 573-882-3364

Вопросы подписки..... 573-882-2042

или info@ire.org