

ТОМАС ДЭВЕНПОРТ

КИМ ДЖИН ХО



О ЧЕМ ГОВОРЯТ ЦИФРЫ

КАК ПОНИМАТЬ И ИСПОЛЬЗОВАТЬ ДАННЫЕ



[Купить книгу на сайте kniga.biz.ua >>>](#)

Thomas H. Davenport and Jinho Kim

Keeping Up With The Quants

YOUR GUIDE TO UNDERSTANDING
AND USING ANALYTICS

HARVARD BUSINESS
REVIEW PRESS

[Купити книгу на сайті kniga.biz.ua >>>](#)

УДК 311.1
ББК 65.051.03
д94

Издано с разрешения издательства Harvard Business Review Press (США)
при участии Агентства Александра Корженевского (Россия)

На русском языке публикуется впервые

Дэвенпорт, Т.

- Д94 О чем говорят цифры. Как понимать и использовать данные / Томас Дэвенпорт, Ким Джин Хо ; пер. с англ. Э. Кондуковой — М. : Манн, Иванов и Фербер, 2014. — 224 с.

ISBN 978-5-00057-118-7

Неважно, чем вы интересуетесь, в какой сфере работаете и каких размеров ваша компания — цифры и аналитика сегодня повсюду, и всем приходится иметь с ними дело.

Эта книга в доступной форме познакомит вас с количественным анализом, его терминами и методами, поможет развить аналитические навыки и разговаривать на одном языке с количественными аналитиками.

УДК 311.1
ББК 65.051.03

Все права защищены.

Никакая часть данной книги не может быть воспроизведена в какой бы то ни было форме без письменного разрешения владельцев авторских прав.

Правовую поддержку издательства обеспечивает юридическая фирма «Вегас-Лекс»

VEGAS LEX

© 2013 Harvard Business School Publishing Corporation
© Перевод на русский язык, издание на русском языке,
оформление. ООО «Манн, Иванов и Фербер», 2014
ISBN 978-5-00057-118-7

[<< Купить книгу на сайте kniga.biz.ua >>](http://kniga.biz.ua)

ОГЛАВЛЕНИЕ

ГЛАВА 1. Почему аналитические способности нужны каждому	7
ГЛАВА 2. Формулирование проблемы.....	28
ГЛАВА 3. Решение проблемы	61
ГЛАВА 4. Результаты и необходимые меры	98
ГЛАВА 5. Креативность в количественном анализе	129
ГЛАВА 6. Развитие навыков количественного анализа	155
ГЛАВА 7. Работа с квантами	183
Итоговые выводы об аналитическом мышлении	209
Благодарности	211
Об авторах	212

ГЛАВА 1

Почему аналитические способности нужны каждому

Мы живем в мире информации, объем которой нарастает с поразительной скоростью — все больше и больше данных сваливается на наши головы, причем большую их часть собирают, чтобы улучшить качество принимаемых решений в бизнесе, государственном управлении или общественной деятельности. Если нам не удается это сделать методами количественного анализа, тогда данные пропадают впустую, а уровень эффективности нашей деятельности далек от желаемого. Цель этой книги в том, чтобы показать читателю, как работает количественный анализ (даже если у читателя нет математической подготовки) и как использовать его для улучшения принимаемых решений.

Актуальность данных и аналитических исследований

Рост объема информации характерен почти для каждой сферы общественной деятельности. Если вы занимаетесь спортом, то наверняка знаете о *бейсбольной статистике Moneyball** и революции в профессиональном бейсболе, которую позволил совершить анализ данных об эффективности действий отдельных игроков. Сейчас такая статистика внедрена практически во всех популярных видах спорта. Если вы увлекаетесь сетевыми компьютерными играми, то наверняка знаете, что разнообразные сведения о вашем игровом поведении

* Этой системе посвящена книга: *Льюис М. Moneyball. Как математика изменила самую популярную спортивную лигу в мире*. М.: Манн, Иванов и Фербер, 2013. 464 с. Прим. ред.

накапливают и анализируют компании Zynga и Electronic Arts. Любите кино? Возможно, слышали о методике, применяемой компанией Netflix для прогнозирования предпочтений в области кино. Может быть, вы не знаете, что некоторые голливудские киностудии (например, Relativity Media) используют похожие методики, принимая решение о том, какие кинопроекты финансировать.

Важно, что существуют различные типы данных. Некоторые используют для бухгалтерского учета. Например, ваша компания наверняка сохраняет информацию о том, с какого числа вы зачислены в штат или сколько дней ежегодного отпуска использовали. Но по мере накопления все большего объема данных того или иного типа у компаний возникает понятное желание найти им полезное применение, в частности использовать для принятия решений. Обычно активно пользуются базой данных о персонале. На основе этих данных руководство может задать вопрос: какое количество сотрудников, возможно, уволится в следующем году? Есть ли взаимосвязь между полным использованием отпуска за текущий год и результатами работы сотрудника за тот же период?

Но сбор данных и аналитика не просто способствуют принятию оптимальных управленческих решений. Многие интернет-компании — Google, Facebook, Amazon, eBay и прочие — используют так называемые *большие данные* о текущих онлайновых операциях не только когда нужно обосновать принимаемое решение, но и при разработке новых продуктов и их свойств. Стремитесь ли вы создать дополнительную ценность для потребителей или оптимизировать принимаемые решения — все это задачи для аналитиков. Они обобщают данные, ищут в них внутреннюю логику и на ее основе разрабатывают модели. Найти в данных логику и эффективно их использовать можно лишь с помощью методов математического или статистического анализа, обобщенно называемых *аналитикой*.

Что такое аналитика

Аналитикой мы называем всестороннее использование баз данных, статистический и количественный анализ, объяснительные и прогнозные модели, а также доказательный менеджмент, применяемые для поддержки решений и увеличения ценности для потребителей.

В зависимости от цели и методов аналитику можно разделить на описательную (дескриптивную), предсказательную (предикативную) и нормативную (прескриптивную). **Описательная аналитика** включает сбор, систематизацию, представление данных в табличной форме, а затем выделение их основных характеристик. Этот вид аналитики всегда ориентировался на информирование о характеристиках данных. Он может оказаться весьма полезным, но ничего не говорит о причинах сложившейся ситуации или о том, что произойдет в будущем.

Предсказательная аналитика выходит за рамки простого описания данных и зависимостей между переменными (в виде показателей, которые могут иметь целый ряд значений) и прогнозирует динамику показателей в будущем на основе данных за прошлые периоды. Сначала определяются связи между переменными, а затем на основе их анализа оценивается вероятность того или иного события: например, насколько вероятно, что потребитель отреагирует на рекламу и купит данный продукт. Хотя связи между переменными используются для прогнозирования будущего, явная причинно-следственная связь обнаруживается далеко не всегда. По сути, она совсем не обязательна для получения точного прогноза.

Нормативная аналитика ориентируется на более широкий круг задач и включает такие методы, как проведение экспериментов и оптимизация. Подобно тому как доктор выписывает рецепт, нормативная аналитика предлагает направление действий. **Эксперимент** призван ответить на вопросы о **причине** тех или иных явлений. Чтобы с уверенностью делать выводы о причинных связях, исследователи изменяют одну или несколько независимых переменных и наблюдают реакцию зависимой переменной, одновременно контролируя внешние по отношению к исследуемой системе факторы. Если тестовая группа, подчиняющаяся условию эксперимента, показывает существенно лучшие результаты по сравнению с контрольной группой, то ответственный менеджер может принять решение о широком внедрении этого условия.

Еще один вид нормативной аналитики — **оптимизация**. Она направлена на выявление оптимального значения конкретной переменной во взаимосвязи с другой переменной. Например, нам

необходимо рассчитать цену продукта, обеспечивающую максимальную рентабельность его продаж. В розничной торговле оптимизационный подход таким же образом позволяет выявить уровень запасов, гарантирующий отсутствие сбоев из-за временного отсутствия какого-либо товара.

В зависимости от применяемых методов и видов данных, подлежащих сбору и анализу, аналитику можно разделить на количественную и качественную. Цель **качественной аналитики** состоит в углубленном понимании причин и мотивов тех или иных явлений. Обычно для этого собирают ограниченное количество неструктурированных данных на основе нерепрезентативной выборки*, а анализ проводят нестатистическими методами. Качественная аналитика полезна при проведении поисковых исследований, то есть на первом этапе исследований аналитических. **Количественная аналитика** представляет собой систематическое изучение событий при помощи статистических, математических и вычислительных процедур. На основе сбора информации о большом количестве репрезентативных событий с последующей статистической обработкой обычно получают структурированные данные.

В зависимости от целей исследователи применяют различные аналитические методы:

статистика — сбор, систематизация, анализ, интерпретация и оглашение данных;

прогнозирование — оценка динамики той или иной переменной в определенный момент в будущем на основе данных о ее динамике в прошлом;

интеллектуальный анализ данных (Data mining) — автоматизированное или полуавтоматизированное выявление ранее неизвестных зависимостей в больших массивах данных с помощью специальных вычислительных алгоритмов или статистических методов;

интеллектуальный анализ текстов — выявление неизвестных зависимостей или тенденций в тексте методами, подобными интеллектуальному анализу данных;

* Нерепрезентативная — выборка, не отражающая характеристики генеральной совокупности. *Прим. перев.*

оптимизация — использование математических методов для того, чтобы найти оптимальные решения на основе заданных критериев и установленных ограничений.

эксперимент — формирование тестовой и контрольной групп методом случайного отбора и выявление причин и степени влияния независимых переменных на зависимую переменную.

В этом списке приведены широко известные аналитические методы, причем многие из них используют одни и те же аналитические приемы и процедуры. Например, регрессионный анализ — наиболее распространенный аналитический прием в предсказательной аналитике — не менее популярен и в статистике, прогнозировании и интеллектуальном анализе данных. Точно так же анализ временных рядов, специальная аналитическая процедура из арсенала статистики, предназначенная для анализа меняющихся во времени значений переменных, используется не только в статистике, но и в прогнозировании.

Учетные данные, помогающие принимать решения по персоналу (мы уже говорили о них), являются структурированными (легко представляются в виде таблицы), количественными и относительно небольшими по объему (не более терабайта или двух даже в очень крупных компаниях). Такие данные традиционно использовались в аналитике, поэтому назовем их *малыми данными*. Долгое время аналитики ни с чем другим дела не имели.

Но сегодня крупные компании, некоммерческие организации и даже стартапы сталкиваются с так называемыми *большими данными* — неструктуризованными массивами информации колоссальных объемов. Их источниками могут быть онлайновые дискуссии в интернете, видеоматериалы или данные анализа ДНК пациентов больницы. У данных такого рода объем намного больше — иногда тысячи петабайт*. Например, Google обрабатывает порядка 24 петабайт интернет-данных ежедневно, а AT&T** передает по телекоммуникационным

* Один петабайт равен 1 073 741 824 мегабайтам. *Прим. перев.*

** AT&T Inc — одна из крупнейших телекоммуникационных компаний в США, провайдер телефонной и беспроводной связи. *Прим. ред.*

сетям около 30 петабайт музыки и прочих данных в день. Благодаря новым прикладным компьютерным программам и техническим новшествам мы можем анализировать огромные массивы данных и извлекать из них полезную информацию.

Что такое большие данные

Термин **большие данные** применяется для обозначения данных уникально большого объема или неструктурированных данных. Приведем несколько примеров:

- За месяц 600 миллионов пользователей Facebook добавили в сеть 30 миллиардов единиц контента.
- Компания Zynga, занимающаяся сетевыми виртуальными играми, ежедневно обрабатывает более петабайта игровой информации.
- Пользователи YouTube просматривают более двух миллиардов видеоклипов в день.
- Пользователи Twitter выполняют 32 миллиарда поисковых запросов в месяц.
- Пользователи Google в 2011 году выполняли почти 5 миллиардов поисковых запросов в день.
- В 2009 году ежедневно отсылалось более 2,5 миллиарда текстовых сообщений.
- В 2010 году население планеты использовало 5 миллиардов мобильных телефонов.
- Объем файла с полной расшифровкой человеческого генома составляет около одного терабайта.
- Беспроводной датчик для контроля физических параметров одной коровы передает около 200 мегабайт данных в год.
- В 2008 году количество подключенных к интернету устройств превысило численность населения земного шара.
- По оценке компании Cisco Systems, к концу 2011 года двадцать типичных домохозяйств генерировали более интенсивный интернет-трафик, чем все пользователи интернета в 2008 году.

- McKinsey & Company считает, что почти в каждой отрасли американской экономики компании с численностью персонала более 1000 человек накапливают в среднем больший объем информации, чем Библиотека Конгресса США.

Большие данные и основанная на них аналитика способны существенно изменить практически каждую отрасль экономики и бизнес-процессы в течение следующих десяти лет. Любая организация (и любой ее сотрудник), если вовремя ознакомится с сутью и методами обработки больших данных, получит огромное конкурентное преимущество. Точно так же как компании, в свое время первыми освоившие методы обработки малых данных, опередили своих конкурентов, сейчас на позиции лидеров выйдут те, кто раньше других сумеет использовать возможности больших данных.

Потенциал больших данных можно реализовать благодаря глобальным устройствам их сбора и обработки. Сенсоры и микропроцессоры в скором времени будут везде. Практически каждое механическое или электронное устройство регистрирует свои действия, местонахождение или состояние. Эти устройства и люди, их эксплуатирующие, поддерживают связь через интернет, а это еще один колоссальный источник данных. Если добавить к этому объемы информации, проходящие через прочие средства связи (беспроводные и проводные телефонные линии, кабели, спутники и т. п.), трудно даже оценить все перспективы.

Доступность всех этих данных означает, что практически любая предпринимательская или управленческая деятельность может рассматриваться либо как проблема больших данных, либо как возможность их обработки. Производство, в котором значительная часть оборудования оснащена одним или несколькими микропроцессорами, все чаще становится средой, где функционируют большие данные. Потребительский маркетинг с мириадами покупок и историй посещений покупателей также сталкивается с проблемой их обработки. Google даже описывала свою самоходную повозку из будущего как проект, связанный с обработкой больших данных.

Гэри Лавмен, CEO^{*} компании Caesars Entertainment (известный своим выражением «Мы так полагаем или мы знаем?»), глава Amazon Джефф Безос («Мы никогда не пренебрегаем данными») и Рид Хоффман, руководящий LinkedIn («Web 3.0 — это сеть с новым типом данных»), публично заявляли, что аналитическое мышление и принятие решений — это надежный способ обеспечить успех компании и свой личный успех. Любая компания в любой отрасли заинтересована в том, чтобы извлечь пользу из вала данных. Для этого требуются люди, умеющие провести их детальный анализ. У них разные имена, но всех их называют *квантами*^{**}, и эта книга предназначена не им. А еще компаниям нужны люди, способные принимать оптимальные решения на основе анализа и воплощать их в жизнь. Именно для них написана эта книга. Это вовсе не те самые *кванты*, не аналитики, у них нет математической подготовки, но им приходится работать с количественными данными и принимать решения на основе их анализа.

Какую пользу вы извлечете из этой книги

У нас, авторов, по этому вопросу разные мнения, но общая цель: расширить применение аналитического мышления в бизнесе и обществе, в первую очередь помогая *не-квантам* извлекать больше пользы из массивов данных. Том вообще не очень-то серьезно относится к квантам, будучи социологом по образованию и не слишком хорошо разбираясь в статистике, но ему приходится заниматься аналитикой и определять направления ее применения в бизнесе. В течение двадцати лет Том вел исследования, писал книги, обучал студентов и консультировал компании по вопросам формирования аналитического потенциала. Его опыт работы с менеджерами и обучения их аналитическим исследованиям лег в основу этой книги. Кроме того, Том — автор и соавтор бестселлеров *Competing on Analytics* («Аналитика как конкурентное преимущество»)^{***} и *Analytics*

* CEO — высшее должностное лицо компании, ближайший российский аналог — генеральный директор.

** От англ. quant — специалист по применению математических методов, методов количественного анализа в решении финансовых вопросов, оценке рисков и пр. Сленговое сокращение от quantitative analyst, quantitative analysis — количественные методы анализа. Прим. ред.

*** Издана на русском языке: Харрис Д., Дэвенпорт Т. Аналитика как конкурентное преимущество. М. : BestBusinessBooks, 2010. 264 с.

at Work («Аналитика в работе»), в которых рассказывается о том, какую роль аналитика играет в разработке стратегии больших компаний. В предлагаемой вниманию читателей книге основное внимание уделено тому, как помочь сотрудникам развить аналитические навыки и способности.

Джин Хо в университете учился именно деловому администрированию и статистике и, естественно, стал убежденным квантом. Он исследовал применение аналитических методов для решения разнообразных проблем в бизнесе и общественной деятельности. Кроме того, он разработал и вел образовательный курс по развитию аналитических способностей сотрудников. Джин Хо получил в Корее звание профессора делового администрирования и статистики, написал шесть книг, в том числе 100 Common Senses in Statistics («100 здравых смыслов в статистике») и Freak Statistics («Причудливая статистика»). Главная их цель — помочь людям без статистической или математической подготовки лучше понимать и интерпретировать статистические данные и результаты.

Мы надеемся, что эта книга поможет вам разобраться в аналитике и уверенно использовать статистическую информацию. Она облегчит сотрудничество с квантами и научит на равных дискутировать с ними об аналитических процедурах и методиках. Вы заговорите языком количественного анализа и сможете задавать правильные вопросы. Может быть, она даже вдохновит вас на то, чтобы стать квантом!

Информированный потребитель аналитических данных

У менеджеров, тесно сотрудничающих с аналитиками, множество вариантов того, как можно использовать количественно-аналитические данные для обоснования своих действий. Вот, например, какие решения недавно приняла Дженифер Джой, вице-президент отделения клинических исследований и управляющая колл-центром компании Cigna, ведущей в отрасли здравоохранения. Колл-центр активно работает с клиентами, стремясь улучшить их состояние, особенно если у них хроническое заболевание (например, диабет или болезни сердца), требующее постоянного лечения и контроля. Это способствует повышению качества медицинских услуг. Джен в свое время работала медицинской сестрой и не занималась аналитикой. Но она получила

степень магистра делового администрирования и считает, что аналитика очень важна для компании. На ее примере легко убедиться в том, что два ключевых аспекта аналитического мышления — умение определять проблему и задавать правильные вопросы — помогают сберечь деньги для компании и ее клиентов.

Ключевой вопрос в работе Джой — сколько времени следует уделять консультированию клиентов Cigna через колл-центр по поводу того, как предотвратить хроническое заболевание или замедлить его развитие. Она стремится доказать, что можно одновременно создавать дополнительную ценность для потребителей и контролировать издержки компании. Ключевая итоговая переменная для Джой — это частота первоначальной и повторной госпитализации пациентов. С одной стороны, она просматривает десятки страниц ежемесячных отчетов, где отражена динамика этого показателя. Процент повторной госпитализации то повышается, то понижается, и Джой не уверена, что понимает, почему так происходит: «Мне приходится изучать множество разных показателей, но это всего лишь цифры, они не отвечают на вопрос, что из этого следует». Но ей действительно хотелось бы выяснить, помогают ли пациентам звонки в колл-центр компании и влияют ли они на частоту повторных госпитализаций.

Чтобы лучше понять причинно-следственную связь между этими показателями, Джой обратилась к экспертам по аналитике в Cigna. Группу аналитиков в компании возглавляет Майкл Казинс. Он и его коллеги решили помочь Джой разобраться в возникшей проблеме. По словам Майкла, «хотя Джен и не специалист по методике анализа причинно-следственных связей, у нее блестящая логика, пытливый ум и она умеет задавать правильные вопросы. У нее нет математической подготовки, но она придает огромное значение аналитике». Джен относится именно к тому типу лиц, принимающих решения, с которым аналитики особенно любят работать.

Группа Казинса специализируется на применении аналитики для нужд бизнеса Cigna. Майкл и его коллеги целиком разделяют мнение Дженифер о том, что показатель частоты госпитализаций не слишком полезен сам по себе, вне связи с вопросом о влиянии консультаций ее врачей на здоровье пациентов. Иными словами, предшествующие отчеты, показывавшие то повышение, то понижение частоты госпитализаций, были основаны на методике, которая не искала решения проблемы. В частности, не проводилось обоснованное сравнение

с контрольной группой, результаты которого можно было бы использовать при принятии решения. В прошлом, например, отчеты не учитывали степень тяжести заболевания тех или иных пациентов. Группа Казинса разработала методику *парной группировки пациентов* в зависимости от тяжести заболевания, демографических параметров, образа жизни, региона проживания. Один из пары пациентов, у которых эти показатели были похожими, пользовался консультационными услугами колл-центра, а второй — нет. Казинс подчеркнул, что «Джой потребовалась определенная решимость, чтобы проверить, действительно ли эффективна консультационная служба, ее любимое детище. И все же она без колебаний приступила к выяснению истины».

Полученные результаты свидетельствовали: консультации врачей колл-центра по поводу некоторых болезней оказались вовсе не столь эффективны, как ожидалось; зато по другим заболеваниям картина была обратной. Джой решила сократить время телефонных консультаций для пациентов из целевой группы с определенными заболеваниями на тот период, пока не удается выяснить, как сделать их более эффективными. Для контрольной группы предполагалось внедрить консультации, способные принести реальную дополнительную ценность.

В то же время Джой продолжала совместную работу с группой Казинса над другими аналитическими проектами. Один из них — контролируемый эксперимент с разными подходами к консультированию, например с включением в процесс консультаций персонального врача пациента. Вместо того чтобы полагаться на интуицию, решая, какой подход сработает лучше, Джой с энтузиазмом реализовывала идею структурированной аналитики, например пилотные программы типа «тестируй и учись», в соответствии с которыми в год выполнялось двадцать или тридцать тестов.

Дженнифер Джой всегда отличалась аналитическим складом ума, но благодаря сотрудничеству с аналитиками она получила инструменты, позволяющие проверять различные гипотезы и принимать обоснованные решения. И хотя так и не удалось отыскать идеальный способ лечить болезнь по телефону, но Cigna по крайней мере тратит меньше денег на неэффективные формы работы с пациентами и больше на те, которые доказали свою эффективность. С другой стороны, Майкл Казинс и его коллеги извлекли большую пользу из работы с Джой и другими аналитически мыслящими топ-менеджерами компании.

Они научились излагать результаты исследований языком бизнеса и объяснять их смысл пациентам.

Вот почему эту книгу мы писали для множества Дженифер Джой в разных уголках мира, а не для Майклов Казинсов. Мы не предлагаем вам стать квалифицированным количественным аналитиком или специалистом по базам данных. Для этого потребуется намного больше, чем просто прочесть книгу, да и не у каждого есть к этому интерес и способности. Но мы предлагаем стать компетентным пользователем данных и аналитики. Тогда вы сможете правильно ставить задачу аналитикам, использовать данные в своей работе, принимать на их основе решения и пропагандировать их применение в вашей организации. Мы хотим, чтобы о вас говорили как о человеке, глубоко увлеченном аналитикой, или как о Джен Джой из Cigna: «Сама она не аналитик, но понимает их и ценит их работу». Мы не ожидаем, что вы сами начнете проводить сложный анализ баз данных, но очертить поле поиска решения, поставить нужные вопросы о данных и методологии их анализа, интерпретировать полученные ответы и на их основе повысить эффективность действий компании вы сможете. Если перефразировать заведующего кафедрой статистики Гарвардского университета Сяо Лименга, цель этой книги не в том, чтобы сделать из вас винодела (так он называет обладателей степени PhD* по статистике), а в том, чтобы привить вкус к хорошему вину**.

В прошлом в большинстве случаев было куда трудней стать компетентным пользователем информации, не ориентируясь в методах и приемах ее получения и обработки. Но сегодня в этой области произошли большие перемены. Теперь не надо до тонкостей разбираться в устройстве двигателя внутреннего сгорания, чтобы стать хорошим водителем; точно так же не обязательно вникать в детали статистического анализа, чтобы использовать статистические данные для принятия решений. Аналитическое программное обеспечение взяло на себя черновую работу, иногда даже может выбирать методику анализа, соответствующую характеру данных и переменных. Некоторые

* Ученая степень PhD (Doctor of Philosophy) примерно соответствует степени кандидата наук в России. Прим. ред.

** Цитата из выступления Сяо Лименга на семинаре «Статистическое образование и обучение статистиков: подготовка знатоков вина и виноделов», проводившемся в Университете Миннесоты 28.10.2011 г. URL: <http://catalystsumn.blogspot.com/2011/11/statistics-education-seminar-presented.htm>.

новые программы (например, от компании SAS) имеют справочную функцию, простым и понятным языком объясняющую смысл тех или иных зависимостей или характер методов, применяемых для прогнозных расчетов.

Хотя потребность в квалифицированных потребителях аналитики высока, на текущий момент нет книг, просто и без математического сленга написанных для новичков в области количественного анализа. В этой книге говорится о том, что такое аналитика, как можно ее использовать во многих жизненных ситуациях и как развить свои аналитические способности. Это поможет вам не только лучше разбираться в аналитике, но и значительно эффективнее обсуждать со специалистами различные аналитические методы и их применение для решения проблем компании. В соответствии с отчетом международной консалтинговой компании McKinsey Global Institute о больших данных за 2011 год, экономике требуются более полутора миллионов компетентных в аналитике менеджеров, чтобы эффективно использовать данные, накапливаемые обществом*. Надеемся, что вы станете одним из них.

Роль аналитики в принятии решений

Решения в коммерческих и некоммерческих организациях принимают исходя из целого ряда факторов: опыта, интуиции, результатов экспериментов, аналитических исследований и накопленных данных. В книге *Moneyball*, посвященной применению аналитических процедур в профессиональном бейсболе, говорится, что одно это отнюдь не гарантирует неизменно положительного результата. Команда *Oakland Athletics* выигрывала далеко не каждую игру описанного в книге сезона, да и всех последующих тоже. Тем не менее аналитика способна обеспечить некоторое конкурентное преимущество тем, кто в ней разбирается. *Oakland Athletics* добивается гораздо лучших результатов, чем можно было бы ожидать с учетом ее более чем скромного бюджета.

Конечно, ответственный менеджер вполне способен принять удачное решение, руководствуясь лишь интуицией и опытом, особенно

* Big Data: The Next Frontier for Innovation, Competition, and Productivity // McKinsey Global Institute, May 2011 URL: http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation.

когда решение лежит в сфере его непосредственной компетенции. Но почти в каждой области деятельности можно найти доказательства того, что решения, принятые на основе анализа данных, более точны и эффективны, обеспечивают больший выигрыш для организации*. В настоящее время в профессиональном бейсболе практически каждая команда применяет аналитические подходы, разработанные в Oakland Athletics. Даже команда New York Yankees, некогда чуть ли не гордившаяся отказом от аналитики в вопросах подбора игроков и определения стратегии игры, сейчас пригласила на работу 21 специалиста по спортивной статистике.

В коммерческих организациях традиционная аналитика чаще всего применяется для поддержки внутренних решений компаний: «Сколько должен стоить этот продукт?» или «Как стимулировать покупателей совершать у нас покупки?» Аналитика в среде больших данных часто используется для разработки новых видов продуктов или дополнительных потребительских свойств. Например, Google создала PageRank — алгоритм ранжирования для поиска, социальная сеть LinkedIn — функцию «Люди, которых вы можете знать» или «С кем я могу связаться в сети», а компания Zynga — новые игры. Все эти продукты и свойства стали результатом управлеченческих решений или компаний, или их потребителей. (На рис. 1.1 «Типы управлеченческих решений, которые требуют аналитической поддержки» приведены еще некоторые примеры.) И это только то, что лежит на поверхности. В других отраслях и секторах экономики (государственном управлении, здравоохранении, спорте и других) можно найти массу подобных образцов.

Тот, кто принимает решения, оценивает возможные варианты решений с учетом информации как количественного, так и качественного характера. Источники качественной информации включают интуицию, опыт, здравый смысл, житейскую мудрость, слухи и предположения. Иногда они оказываются полезными, однако нужно помнить об их недостатках. Даже если у вас богатый опыт принятия решений в той или иной области, очередная ситуация может оказаться непохожей на предыдущие. Принимать решения на основе предположений всегда рискованно, а интуиция может вас обмануть. Большинство людей слишком полагаются на нее при принятии решений.

* Dawes R., Faust D., and Meehl P. Clinical Versus Actuarial Judgment // Science. March 1989. Vol. 243, no. 4899. P. 1668–1674.

Рис. 1.1. Типы управленческих решений, которые требуют аналитической поддержки

Маркетинг



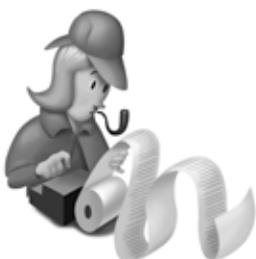
- Ценообразование
- Размещение торговых точек и региональных подразделений
- Целевое продвижение продуктов
- Дизайн сайта компании
- Размещение рекламы в электронных СМИ

Поставщики



- Объем складских запасов
- Размещение дистрибуторских центров и складов
- Маршруты доставки продуктов или движения транспорта
- Загрузка транспорта

Финансы



- Факторы финансовой деятельности
- Сбалансированная система показателей
- Различные виды прогнозов

Персонал



- Каких сотрудников нанимать
- Кто из сотрудников собирается уволиться
- Какой должна быть сумма выплат и компенсаций
- Какое образование предпочтительно для сотрудников

Исследования и разработки



- Какие потребительские свойства продукта больше всего привлекают потребителей
- Насколько эффективно производство и продажа продукта
- Какой дизайн продукта максимально привлекает потребителей

Экономическая школа «Поведенческая экономика» исходит из постулата о том, что интуиция — далеко не лучший советчик при принятии экономических решений.

Несмотря на все достоинства аналитических исследований, бывают ситуации, когда нецелесообразно опираться на них при принятии решения. Если проблема слишком незначительна, носит единовременный характер или решение зависит от личных предпочтений, нет особого смысла тратить время на сбор и анализ данных, построение аналитической модели. Если действовать нужно быстро, то заниматься анализом просто некогда. Напротив, если решения приходится принимать периодически, есть время на проведение анализа, а проблема достаточно серьезна, чтобы оправдать и затраты, и труд, то аналитический подход будет весьма полезен.

Три этапа аналитического подхода и порядок их выполнения

В центре внимания в этой книге — три основных этапа аналитического подхода, каждому из которых посвящено по главе. Приводятся примеры аналитических задач, для которых особенно важен тот или иной этап. Схематически эти этапы вместе с входящими в них шагами показаны на рис. 1.2.

В главе 2 рассказывается о первом этапе — *формулировании проблемы*. Он включает в себя постановку вопросов, на которые предстоит ответить аналитику, и поиск формулировки для проблемы, которую нужно решить. Понятно, что это очень важный этап, ведь если сформулировать проблему некорректно, никакие данные или процедуры анализа не помогут ее преодолеть. Этот этап включает два шага: *определение проблемы и изучение предыдущих поисков решения*. В главе 2 мы увидим, что после того, как проблема определена и сформулирована для решения аналитическими методами, как правило, оказывается, что некоторые ее аспекты уже рассматривались другими исследователями ранее, и это помогает уточнить формулировку.

Второй этап рассматривается в главе 3 и охватывает наиболее интересующие читателей методы *решения проблемы*. На этом этапе вам предстоит отобрать переменные для будущей модели, собрать данные об их значениях за определенные периоды, а затем провести анализ. Допустим, вы вовсе не квант и не имеете особого желания им

Рис. 1.2. Три этапа и шесть шагов количественного анализа

Формулирование проблемы

1. Определение проблемы

2. Изучение предыдущих поисков решения



Решение проблемы

3. Моделирование ситуации

4. Сбор данных

5. Анализ данных



Результаты и необходимые меры

6. Демонстрация результатов и дальнейшие действия



24 ГЛАВА 1

становиться. Тогда выход один: поручить большую часть этих работ количественному аналитику. Тем не менее очень полезно знать хотя бы в общих чертах, что и как тот собирается делать. Вам не обязательно решать всю проблему самостоятельно, но если вы будете уметь задавать правильные вопросы и понимать общие принципы анализа, то это, безусловно, поможет вам найти лучшее решение.

В главе 4 рассматривается третий и последний этап количественного анализа, ничуть не менее важный, чем остальные, но часто игнорируемый, — *оформление результатов анализа и принятие необходимых мер*. От того, как вы расскажете о результатах анализа, в определяющей степени зависит, станут ли они толчком к дальнейшим действиям. Если тот, кто принимает решения (возможно, это вы и есть), не понимает, в чем именно заключается проведенный анализ и что означают полученные результаты, то вряд ли они станут основой для управленческих решений. Тогда не стоит и тратить время на прохождение первых двух этапов. Мы живем в информационно перегруженном мире, в котором трудно привлечь к чему-то внимание. Вот почему особенно важно подать результаты анализа в интересной, захватывающей форме. Сейчас уже нельзя представить информацию в виде толстого отчета, переполненного сухими цифрами, и при этом надеяться, что для кого-то они станут стимулом к действию.

В остальных главах рассматриваются некоторые частные вопросы аналитического подхода. В частности, в главе 5 мы говорим о роли креативности в аналитической работе: оказывается, эти понятия вполне совместимы! В главе 6 описаны несколько способов развития аналитических способностей для тех, кто почувствовал вкус к аналитике. Глава 7 посвящена налаживанию конструктивных взаимодействий менеджеров и квантов для принятия более эффективных решений. Нечего и говорить, что успех этих взаимодействий зависит от обеих сторон. Мы приводим много разнообразных примеров, иллюстрирующих возможности использования аналитических методов для решения проблем (или, напротив, то, как из-за пренебрежения ими возможности были упущены; см. вставку «Почему опасно пренебрегать аналитическими методами»). Кроме того, в этой главе показаны примеры расчетов и описаны самые простые возможности применения аналитики, для внедрения которых не требуется много времени и сил.

Почему опасно пренебрегать аналитическими методами

Мы учимся как на положительных, так и на отрицательных примерах, но почему-то отрицательные обычно производят более сильное впечатление. Трудно найти более яркий образец, чем история Джо Кассано, практически в одиночку доведшего до кризиса огромную компанию, а заодно и экономику США, да черт возьми, всего мира!

Кто же такой Кассано и чего он не знал об аналитике? Возможно, вы вспомните этого человека, если назвать его должность — глава AIG Financial Products (AIGFP), подразделения с четырьмя сотнями сотрудников, входящего в состав гиганта страховой отрасли AIG. Именно он несет ответственность за потерю компанией колоссальной суммы денег, о точной ее величине все еще идут споры, но что-то около 85 миллиардов долларов. Столько американским налогоплательщикам пришлось заплатить, чтобы удержать AIG на плаву и погасить ее долги.

Не один Кассано терял деньги, но, по словам расследовавшего эту историю репортера Мэтта Тайбби из журнала *Rolling Stone*, он был «виновником № 1 мирового экономического кризиса»*. Тайбби описывал его как «толстого лысеющего коротышку с глазами-бузинками и высоким лбом, окончившего Бруклинский колледж», хотя это не имеет особого отношения к нашей истории. Уверены, если бы он зарабатывал деньги вместо того, чтобы их терять, его внешность показалась бы репортеру более привлекательной.

Итак, к чему же привело пренебрежение аналитикой и количественным анализом? AIGFP потеряла эти деньги в результате активной продажи финансового продукта, называемого **кредитным дефолтным свопом** (CDS) и представляющего собой своего рода страховку ипотечных деривативов. Репортер *New York Times* Гретчен Моргенсон вскоре после этих событий писала: «Хотя крах рынка жилой недвижимости в США часто называют непосредственной причиной кризиса, надо отметить, что экономическая система в целом отличалась неустойчивостью из-за широкого распространения загадочных ценных бумаг — кредитных деривативов, предоставлявших

* *Taibbi M. AIG and the Long Con // Rolling Stone. 2009, March 23.*

кредиторам страховку на случай неплатежеспособности должников. Их выпускали частным образом, не привлекая внимания регуляторов фондового рынка, а иногда и не слишком включая мозги менеджеров, ответственных за их выпуск»*.

Кассано уж точно оказался одним из топ-менеджеров, не перегружавших мозги размышлениями о последствиях. И деривативы, и CDS принадлежат к числу сложных финансовых инструментов, разрабатываемых методами статистического и математического анализа, но, как позже выяснилось, на этот раз математика и статистика подвели. Если ипотечные заемщики оказывались неплатежеспособными, деривативы обесценивались и AIG сталкивалась с необходимостью выплачивать их держателям страховую стоимость ценных бумаг. Стоит ли говорить, что именно это и произошло?

Майкл Левис описал суть проблем подразделения AIGFP в журнале *Vanity Fair***:

«Трейдеры AIGFP считают загадкой причины, приведшие их компанию к катастрофе. А начиналось все просто — с изменения процедуры принятия решений новым руководством. В конце 2001 года второй по счету генеральный директор компании Том Саваж ушел в отставку и на эту должность назначили его бывшего заместителя Джо Кассано. Саваж имел солидную математическую подготовку и хорошо разбирался в моделях оценки риска, применявшимся трейдерами AIG для гарантии получения достаточной компенсации. Кроме того, Саваж любил и часто проводил внутрикорпоративные обсуждения достоинств тех или иных моделей, а также сделок. Новый генеральный директор намного хуже разбирался в математике и не имел особого желания вести дискуссии по этим вопросам.

Кассано не слишком беспокоился о том, насколько точна оценка риска и адекватны ли модели, на основе которых она получена. Он не задавал лишних вопросов аналитикам AIG и Уолл-стрит,

* Morgenson G. Behind Insurer's Crisis, Blind Eye to a Web of Risk // New York Times. September 27, 2008.

** Lewis M. The Man Who Crashed the World // Vanity Fair, August 2009. URL: <http://www.vanityfair.com/politics/features/2009/08/aig200908>.

разрабатывавшим эти модели. Он просто продавал CDS всем, кто желал их купить. Мысль о том, что их цена рассчитана на основе неверных допущений, в частности о том, что люди, бравшие эти недорогие ипотечные кредиты, окажутся неспособны их выплатить, кажется, никогда не приходила ему в голову. И даже если приходила, то не слишком его беспокоила».

Развить в себе аналитические способности путем самостоятельных занятий очень сложно. Мы создали специальный сайт (<http://keepingupwiththequants.weebly.com>), где читатели могут задавать любые вопросы об аналитике, аналитических процедурах или трудностях, с которыми приходится сталкиваться при решении конкретных проблем. Не стесняйтесь обращаться туда со всем, что показалось непонятным. На сайте подробно проанализированы некоторые ситуации, описанные в этой книге. Периодически мы будем ссылаться здесь на этот сайт.