

Джордан Морроу

как вытащить из данных максимум

навыки
аналитики
для неспециалистов

 альпина
ПАБЛИШЕР

бизнес

Джордан Морроу

**Как вытащить из данных
максимум. Навыки аналитики
для неспециалистов**

«Альпина Диджитал»

2021

УДК 004.6
ББК 65.291.213

Морроу Д.

Как вытащить из данных максимум. Навыки аналитики для неспециалистов / Д. Морроу — «Альпина Диджитал», 2021

ISBN 978-5-9614-7563-0

Дата-грамотность, то есть способность ориентироваться в мире данных, – ключевой навык сегодняшнего дня. Ежедневно в соцсетях публикуются миллиарды сообщений, электронные почтовые ящики по всей планете гудят от писем, а каждый подключенный к интернету автомобиль производит терабайты данных, не говоря уже об онлайн-магазинах, платежных системах и государственных цифровых сервисах. Однако работать с данными, анализировать их и использовать их для бизнеса по-прежнему умеет меньшинство, а специалистов катастрофически не хватает. Для тех, кто хочет научиться говорить на языке данных уверенно, признанный эксперт в области дата-грамотности Джордан Морроу и написал свою книгу. Это практическое руководство позволит даже неспециалисту освоить четыре базовых уровня аналитики и узнать, как принимать эффективные решения на основе данных, чтобы извлекать максимум из информации и быть успешным в быстро меняющемся цифровом мире.

УДК 004.6
ББК 65.291.213

ISBN 978-5-9614-7563-0

© Морроу Д., 2021
© Альпина Диджитал, 2021

Содержание

Об авторе	6
Предисловие	7
01	8
Данные: мир, в котором мы живем	9
Данные: нехватка навыков	12
Данные: в чем причина нехватки навыков?	14
Данные: что дальше?	17
Краткое содержание главы	18
02	19
Данные и их анализ – целых четыре уровня?	19
Четыре уровня аналитических методов	20
Конец ознакомительного фрагмента.	22

Джордан Морроу

Как вытащить из данных максимум.

Навыки аналитики для неспециалистов

Переводчик *М. Кульнева*

Редактор *Л. Макарина*

Главный редактор *С. Турко*

Руководитель проекта *А. Деркач*

Корректоры *Т. Редькина, Е. Аксенова*

Компьютерная верстка *А. Абрамов*

Художественное оформление и макет *Ю. Буга*

© Jordan Morrow, 2021

This translation of Be Data Literate is published by arrangement with Kogan Page

© Издание на русском языке, перевод, оформление. ООО «Альпина Паблишер», 2022

Все права защищены. Данная электронная книга предназначена исключительно для частного использования в личных (некоммерческих) целях. Электронная книга, ее части, фрагменты и элементы, включая текст, изображения и иное, не подлежат копированию и любому другому использованию без разрешения правообладателя. В частности, запрещено такое использование, в результате которого электронная книга, ее часть, фрагмент или элемент станут доступными ограниченному или неопределенному кругу лиц, в том числе посредством сети интернет, независимо от того, будет предоставляться доступ за плату или безвозмездно.

Копирование, воспроизведение и иное использование электронной книги, ее частей, фрагментов и элементов, выходящее за пределы частного использования в личных (некоммерческих) целях, без согласия правообладателя является незаконным и влечет уголовную, административную и гражданскую ответственность.

* * *

Моей прекрасной жене и пятерым чудесным детям.

Спасибо за поддержку в моем путешествии к дата-грамотности.

Об авторе

Джордана Морроу называют крестным отцом дата-грамотности. Он один из пионеров движения за дата-грамотность и постоянно стремится помогать руководителям и сотрудникам различных организаций раскрывать потенциал данных и аналитики. Вне мира данных Джордан Морроу – счастливый семьянин, у него пятеро детей. Кроме того, он увлеченный любитель бега по пересеченной местности и ультрамарафона.

Предисловие

«Есть три типа лжи – ложь, наглая ложь и статистика».

В этой цитате, которая приписывается самым разным людям, в частности Марку Твену, отражена вся наша сегодняшняя жизнь. Мы живем в мире, где данные – статистика и многое другое – находятся в свободном доступе. Они что-то нам рассказывают, мы черпаем из них информацию. Легко ли нам разобраться во всех данных, которые регулярно к нам поступают? Увы, зачастую они используются и интерпретируются совершенно неправильно. Так можем ли мы что-то с этим сделать?

Мое путешествие в мир дата-грамотности официально началось в июне 2016 года, но первые идеи по этой теме появились у меня еще раньше. Меня называли крестным отцом дата-грамотности и Главным Ботаником (второе прозвище мне и правда нравится). Я долго развивал и совершенствовал свои представления. В этой книге я хочу поделиться своими наблюдениями и помочь вам совершить путешествие в мир данных и аналитики.

Мой подход, возможно, отличается от традиционного, и кое-что в книге покажется вам неожиданным. Для большинства из нас данные стоят далеко не на первом месте в списке самых интересных тем на свете, но лично я ими просто очарован. Мне бы очень хотелось донести до вас правильное представление о мире данных и показать вам, что вы можете сделать, чтобы стать частью этого удивительного мира. Спасибо за то, что решили уделить мне время. Надеюсь, что смогу зажечь в вас искру любопытства и любви к данным: возможно, они смогут изменить вашу жизнь к лучшему, как это произошло со мной.

01

Мир данных

Задумывались ли вы когда-нибудь о том, какое оно – будущее? Каким, например, станет рынок труда? Неужели роботы и техника действительно монополизируют его и отберут у нас работу? Какие нас ждут открытия и когда уже можно будет заказать летающий автомобиль? Будущее всегда туманно, и мы пока не знаем, какие именно изобретения будут определять наше завтра. Бесспорно, в будущем появятся такие профессии, о которых мы пока даже не догадываемся. Но при всей этой неопределенности можно назвать одну вещь, которая точно станет частью будущего и уже присутствует в нашей жизни, – это данные. Будущее обещает нам множество изобретений и удивительных новых профессий, но этого мы только ждем – а вот с властью данных уже столкнулись. И это навсегда.

Мир данных поразителен, огромен и дает нам всем безграничные возможности для развития. Слишком долго попытки людей и организаций добиться успеха с помощью данных были неубедительными. Это неправильно, и мы должны помочь каждому научиться пользоваться всеми возможностями, которые предоставляет этот неисчерпаемый источник.

Данные называют новой нефтью, говорят, что они необходимы как воздух, – избитых штампов и дешевых преувеличений хватает. Но на самом деле данные – это актив, который при правильном использовании может помочь всему человечеству двигаться вперед. С ними легче получать и усваивать знания, они не только готовят нас к будущему, но и закладывают прочное (и в то же время гибкое) основание для него. Такое комплексное понимание мира данных в наше время просто необходимо.

Данные: мир, в котором мы живем

Не секрет, что мы живем в мире, поглощенном технологиями и данными. Вряд ли можно пройти по улице какого-нибудь крупного города вроде Лондона или Нью-Йорка, не увидев людей, уткнувшихся в телефоны, вместо того чтобы смотреть на удивительный мир вокруг и на его обитателей. Убедитесь сами: в следующий раз, выйдя на улицу, посчитайте, сколько людей глядят в экраны, а сколько – смотрят по сторонам. Можете даже прибавить к последним тех, кто болтает друг с другом или приветствует кого-то. Да и сами не забывайте смотреть по сторонам, чтобы избежать неприятных сюрпризов: можно споткнуться о бордюр или не увидеть машину.

В последние лет пятьдесят, а в особенности в последние лет тридцать – с приходом в нашу жизнь интернета, персональных компьютеров, смартфонов и т. д. – мы стали свидетелями невероятного прогресса в развитии технологий и работе с данными. Задумайтесь об этом прогрессе хотя бы на минуту. Вселенная существует примерно 13,8 миллиарда лет¹, Земля – примерно 4,5 миллиарда², а мы говорим всего лишь о тридцати – пятидесяти годах. Но за эти несколько десятилетий у нас на глазах персональные компьютеры и сотовые телефоны стали обычным делом. В масштабах вечности это миг. Но сейчас, в наше время, нам уже трудно представить жизнь без этих технологических новшеств. И каждое из таких устройств генерирует данные. А интернет? Он еще «моложе» – проник в нашу жизнь в начале 1990-х³ и повсеместно распространился лишь к началу 2000-х. Но теперь это неотъемлемая часть нашей жизни – и на работе, и в быту. Чем больше становилось персональных компьютеров, тем быстрее развивались технологии: темп роста не падает и по сей день. Мы постоянно наблюдаем инновации, видим, как развиваются самые разные аспекты цифрового мира. Все это напрямую влияет на нашу жизнь. И в первую очередь – учитывая тему нашей книги – это влияет на работу с данными и на их мощь.

Вернемся к тем временам, когда интернет стал массовым достоянием и начал использоваться более активно. Организации, учебные заведения и отдельные люди теперь живут и работают иначе. И наша жизнь благодаря ему не просто изменилась, но и значительно улучшилась: интернет позволяет совершенствоваться, учиться и развиваться с невиданной скоростью. Когда к интернету стало можно подключить ПК, а теперь и смартфоны, мы смогли принести всю компьютерную мощь к себе домой. Теперь мы можем почти мгновенно получить столько же информации, сколько во всей «Британской энциклопедии»: больше сведений на нас вывалит только торговый агент, постучавшийся к нам в дверь. Мы гораздо быстрее получаем ответы на вопросы. Из стремления найти ответы и вырос Google – более того, это слово (google) даже получило статус глагола в словарях!⁴

Прогресс компьютерных технологий привел к развитию электронной коммерции и появлению Amazon и других компаний, которые смогли изменить потребительские привычки и захватить рынок. У нас на глазах лопнул пузырь доткомов: сначала стоимость IT-компаний, не производящих никакой продукции, вдруг стала огромной – а затем они обанкротились. Примером может служить компания Pets.com. Она возникла в 1998 году и закрылась в 2000-м⁵. Это время благополучно миновало, появилось множество других разнообразных сайтов, а затем началась эпоха социальных сетей. Соцсети открыли двери к профилям потребителей и пуб-

¹ Redd, N. (2017) How Old is the Universe, Space.com, 8 June. <https://www.space.com/24054-how-old-is-the-universe.html>.

² Там же.

³ Zimmerman, K. & Emspak, J. (2017) Internet History Timeline: ARPANET to the World Wide Web, Live Science, 27 June. <https://www.livescience.com/20727-internet-history.html>.

⁴ Словарь Merriam-Webster, определение Google. <https://www.merriam-webster.com/dictionary/google>.

⁵ Aune, S. (2010). Five Dot-Coms That Didn't Survive the Bubble, TechnoBuffalo, 25 January. <https://www.technobuffalo.com/five-dot-coms-that-didnt-survive-the-bubble>.

личным данным, так что самая разная информация о нас – от селфи и фото ужинов до товаров, которые нам нравятся, – стала доступна всем желающим, причем как отдельным людям, так и бизнесу. Ведь так приятно, когда контекстная реклама решает за нас, что мы хотим сегодня на ужин!

Наряду с соцсетями, предоставляющими массу занимательных данных, в 2000-х появилась и новая технология, которая быстро вышла на передний план в области связи и сбора данных, – интернет вещей, или IoT. Что это такое? Все очень просто – это связь всего со всем через интернет. Возьмем, к примеру, датчики в автомобиле или самолете, которые собирают данные обо всем, что происходит в моторе или других частях механизма. Но знаете ли вы, когда на самом деле появился интернет вещей? Думаете, в 2000-х? А вот и нет, хотя многие об этом не знают. Термин появился в 1999 году, но одним из первых примеров интернета вещей можно считать давно знакомый нам... торговый автомат Coca-Cola. Все мы не раз видели такие автоматы, но конкретный автомат, о котором идет речь, находился в Университете Карнеги – Меллона. Он позволял покупателю связаться с холодильником через интернет и выяснить, холодный ли лимонад⁶. Об использовании данных для принятия более взвешенных решений благодаря «связи всего со всем» (то есть интернету вещей) люди задумывались уже давно: как собранные данные могут помочь нам делать правильный выбор в личной и профессиональной жизни? Достаточно вспомнить такие компании, как Amazon или Netflix: насколько часто они собирают наши данные, чтобы «рекомендовать» то, что нам может быть нужно? Очень часто... И знаете что? Эти рекомендации нередко оказываются верными!

Хотя интернет вещей берет начало в 1980-х, в силу он входит только теперь. Представьте, к примеру, любителя бегать ультрамарафоны. Еще несколько лет назад практически не было способов сбора данных, которые могли бы помочь бегуну развиваться... а сейчас они есть. Но нужны ли нам на самом деле часы, которые показывают буквально все аспекты бега – от уклона трассы до скорости на разных участках? Сегодня бегун может получить от своих часов больше информации, чем ему, вероятно, требуется, – за все время бега она накапливается тоннами! Конечно, данные с часов потом интересно рассматривать и обсуждать, но ключевой момент заключается в том, что благодаря «связи всего со всем» и технологическому прогрессу современного мира генерируется все больше и больше данных о нашей жизни. Вы наверняка можете привести и другие примеры того, как работает интернет вещей. Но давайте вспомним именно те случаи, когда прогресс действительно определяет нашу жизнь и помогает нам.

Одна из компаний, культивирующих сетевое взаимодействие, цифровизацию и производство данных, – знаменитый Rolls-Royce. Сейчас это не просто мощная инженерная компания, производящая замечательные двигатели. Rolls-Royce – эффективная организация, управляемая на основе данных: она использует технологию интернета вещей и возможности связи для предоставления и получения данных, которые стали ценным активом компании⁷. Один из примеров использования Rolls-Royce возможностей данных – прогностический метод мониторинга техобслуживания двигателей⁸. С помощью датчиков, собирающих данные, Rolls-Royce успешно предсказывает возможные проблемы с авиационными двигателями и гарантирует, что они не откажут прямо в воздухе. Другой пример того, как «связь всего со всем» и интернет вещей улучшают нашу жизнь, можно найти в сфере здравоохранения. Использование данных, генерируемых оборудованием для физиотерапии, позволяет разрабатывать более эффектив-

⁶ Foote, K. (2016). A Brief History of the Internet of Things, Dataversity.net, 16 August. <https://www.dataversity.net/brief-history-internet-things>.

⁷ Choudhury, A. R. and Mortleman, J. (2018). How IoT is Turning Rolls-Royce into a Data Fuelled Business, CIO, January. <https://www.i-cio.com/innovation/internet-of-things/item/how-iot-is-turning-rolls-royce-into-a-data-fuelled-business>.

⁸ RTInsights Team (2016). How Rolls-Royce Maintains Jet Engines With the IoT, RT insights.com, 11 October. <https://www.rtinsights.com/rolls-royce-jet-engine-maintenance-iot>.

ные программы для тех, кто нуждается в физиотерапевтическом лечении⁹. Учитывая растущую стоимость медицинских услуг, такие более персонализированные программы будут помогать людям следить за здоровьем и реже попадать в больницу.

Еще одна сфера, в которой анализ данных способствует росту и развитию, – это спорт. Помните фильм «Человек, который изменил все» с великолепным Брэдом Питтом в главной роли? В нем говорится, что данные и их анализ могут очень серьезно влиять на судьбу спортивных команд, помогая добиваться побед. Речь, конечно, о бейсболе, но примеры работы с данными можно найти и в баскетболе, включая НБА (Национальную баскетбольную ассоциацию) – а это уже совсем другой масштаб. У большинства команд НБА (возможно, даже у всех) есть свои аналитики и эксперты по данным. Их задача – обнаруживать тенденции и закономерности в данных, которые они собирают: например, поиск недооцененных игроков и повышение их стоимости для продажи или обмена. Кроме того, команды НБА используют данные и технологии для отслеживания уровня утомляемости и качества сна своих игроков, что позволяет корректировать режим тренировок, предотвращать травмы и т. д. НБА даже проводит свой собственный ежегодный «хакатон», чтобы найти и привлечь новых талантливых аналитиков. Знаете ли вы, что количество трехочковых бросков в лиге выросло не в последнюю очередь благодаря анализу данных?¹⁰

Но довольна не только НБА. Всю мощь данных заключают и вещи, которыми мы пользуемся в быту: смарт-часы, смартфоны, посудомоечные машины, холодильники, системы обогрева и кондиционирования, автомобили и прочие транспортные средства... и многое, многое другое. В прочих сферах нашей жизни данные тоже производятся в невероятном количестве. Только представьте себе, сколько информации проходит через соцсети, торговые сайты вроде Amazon и eBay, платежные системы и т. д. Невероятные цифры. Давайте взглянем на статистику, предоставленную Всемирным экономическим форумом. В 2019 году¹¹:

- ежедневно публиковалось 500 миллионов твитов;
- было отправлено 294 миллиарда электронных писем;
- каждый подключенный к интернету автомобиль произвел 4 терабайта данных;
- к 2025 году каждый день в мире будет производиться ориентировочно 463 эксабайта данных.

Если бы мы до сих пор смотрели DVD, а не Netflix, то такое количество данных было бы эквивалентно 212 765 957 дискам! Какое-то немыслимое число. Что это означает для нас? Это означает, что мы за всю жизнь не сможем просмотреть все DVD с данными, произведенными всего лишь за один день. И что нам делать со всей этой информацией?

Конечно же, организации и частные лица могут воспользоваться этим удивительным богатством. Согласитесь, любая организация ухватится за эту идею – как можно применить знания, находящиеся в ее распоряжении, для принятия взвешенных, подкрепленных данными решений. Однако в реальности все оказывается совершенно не так. Исследования демонстрируют нам истинное положение дел: в мире данных налицо чудовищная нехватка навыков, что и мешает организациям успешно применять этот ценнейший ресурс.

⁹ Medical Device Network (2018). Bringing the Internet of Things to Healthcare, MedicalDevice-Network.com, 3 September. <https://www.medicaldevice-network.com/comment/bringing-internet-things-healthcare>.

¹⁰ McLaughlin, M. (2018). How Data Analytics in Sports is Revolutionizing the Game, Biztechmagazine, 13 December. <https://biztechmagazine.com/article/2018/12/how-data-analytics-revolutionizing-sports>.

¹¹ Desjardins, J. (2019). How Much Data is Generated Each Day? World Economic Forum, 17 April. <https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f>.

Данные: нехватка навыков

Чтобы понять причины такой нехватки, необходимо иметь представление о том, как обстоят дела с навыками дата-грамотности. Аналитическая компания Qlik помогает разобраться в общей картине и указывает на конкретные аспекты, где могут наблюдаться пробелы. Одно из исследований (август 2017-го – февраль 2018 года) обрисовало современное состояние сферы дата-грамотности и необходимых навыков, а также предоставило ценнейшую информацию относительно уровня квалификации и уровня комфорта людей при взаимодействии с данными¹². Результаты исследования просто поражают. Выяснилось, что всего 24 % людей, ответственных за принятие важных решений, чувствуют себя уверенно, работая с данными. Всего лишь 24 %. Это крайне малое количество для тех, кто отвечает за будущее организации. В ряде случаев именно им приходится принимать решения именно на основании данных. Как же можно доверять таким решениям, принятым при нехватке навыков?

Когда организации определяют стратегию в области данных и их анализа (если это вообще происходит, потому что тенденции говорят об обратном), подразумевается, что команда топ-менеджеров должна сформулировать такую стратегию, задать направление и создать план использования данных для развития бизнеса. А знаете, какова доля руководителей высшего звена, уверенно пользующихся данными? Согласно результатам того же исследования – 32 %! То есть приблизительно одна треть, причем, судя по всему, это субъективная оценка и она завышена: на самом деле еще меньше топ-менеджеров способны правильно использовать данные и обладают подлинной дата-грамотностью. Итак, если руководители высшего звена определяют представление организации о работе с данными и ее стратегию в этой области, то получается, что лишь 24 % из них принимают правильные решения. Может ли стратегия быть эффективной, если за ее реализацию отвечают люди, не уверенные в своем умении обращаться с данными? Надеюсь, вы уже осознали масштабы проблемы нехватки навыков.

Интересно, что можно сказать о молодежи, которая только сейчас выходит на рынок труда? Все то же исследование Qlik выявило, что среди молодых людей – от 16 до 24 лет – уверены в своих способностях правильно обращаться с данными лишь 21 %! Возникает закономерный вопрос: почему молодые люди настолько не уверены в своих способностях? Разве они не даны им от природы или не формируются на более раннем этапе? Это очень интересный вопрос, требующий более пристального рассмотрения с демографической точки зрения. Люди, которым в 2017 году (когда проводилось исследование) было от 18 до 24 лет, родились в цифровом мире, когда интернет и персональные компьютеры уже стали частью повседневной жизни. Молодежь выросла в интернете, в соцсетях... но значит ли это, что ее учили правильно пользоваться данными и анализировать их? Да, она разбирается в компьютерах, но при этом не разбирается в данных.

Итак, в целом только один из пяти участников исследования уверен в своих навыках обращения с данными, так что для руководителей организаций должен быть очевиден огромный пробел, который необходимо заполнить. Здесь и кроется корень проблемы: если организации хотят извлечь выгоду из данных и аналитики, но при этом нужных специалистов критически не хватает, то как, собственно, извлечь эту выгоду? И какое влияние нехватка дата-грамотности и уверенности в умении обращаться с данными оказывает на организации? Не отражается ли этот недостаток на прибыли?

Влияние человеческого фактора на недостаток знаний о данных трудно переоценить. В исследовании 2019 года было выявлено, что лишь 32 % опрошенных топ-менеджеров утверждают, что способны извлечь измеримую пользу из данных, и 27 % – что их проекты в сфере

¹² Qlik (2018). How to Drive Data Literacy in the Enterprise. <https://www.qlik.com/us/bi/data-literacy-report>.

данных и аналитики «дают им применимые на практике знания»¹³. Конечно же, это тоже следствие недостаточной дата-грамотности. Когда вспоминаешь, что миллионы, а то и миллиарды долларов вкладываются в аналитические проекты, ПО и технологии, связанные с данными, то поневоле задумываешься, сколько из них потрачено впустую. И если среди людей в целом лишь каждый пятый уверен в своих навыках обращения с данными, а топ-менеджеры не осознают ценность этих навыков, то потери действительно огромны.

Какое же влияние нехватка дата-грамотности оказывает на нас, отдельно взятых людей? В упомянутом исследовании приведена лишь общая количественная оценка соответствующих навыков. Но если задуматься о сути исследования, то становится ясно: из-за отсутствия необходимых навыков люди совершенно не справляются с новыми технологиями и потоком данных. Более трети участников исследования признались: если бы перед ними стояла задача, которую можно решить с помощью данных, они предпочли бы искать иные способы ее решения. А целых 14 % в таком случае постарались бы вообще уклониться от этой задачи. Картина всеобщей перегруженности данными представляется еще более убедительной, если обрисовать ее в виде рабочих часов, потерянных в результате неумения пользоваться данными и технологиями: более 5 рабочих дней (43 рабочих часа) на сотрудника за год. Сколько это в реальном денежном выражении? Очень немало! Согласно исследованию, для американской экономики эти потери составляют около 100 миллиардов долларов в год. По-вашему, это огромная цифра? Тогда давайте задумаемся, почему люди не справляются с данными и каковы причины такой нехватки навыков.

¹³ Desjardins, J. (2019). How Much Data is Generated Each Day? World Economic Forum, 17 April. <https://www.weforum.org/agenda/2019/04/how-much-data-is-generated-each-day-cf4bddf29f>.

Данные: в чем причина нехватки навыков?

У такого огромного пробела в знаниях и навыках наверняка есть какие-то причины. Каковы же они? Очевидно, что факторов множество: от недостатков в системе образования до проблем с технологиями и программным обеспечением и даже до производства данных как такового. Сейчас мы подробно рассмотрим эти возможные причины, а вы задумайтесь, как они могли повлиять на вашу карьеру и умение извлекать пользу из данных (будь то на личном уровне или во взаимодействии с людьми, с которыми вы работаете).

Программное обеспечение и технологии

Возможно, вы задумаетесь: как ПО и технологии могут быть причиной нехватки навыков дата-грамотности? Разве все не должно быть наоборот? Разве они не уменьшают этот пробел? Что ж, ответ вроде бы очевиден: да, должно быть наоборот; да, уменьшают. Задача ПО и технологий – помогать нам решать задачи, связанные с данными и их анализом, и получать реальные бизнес-результаты. Они должны дополнять человека, если он достаточно образован и обучен, чтобы ими пользоваться.

Проблема заключается в том, как именно вложения в технологии и ПО осуществляются людьми и организациями. Представьте себе, что вы руководите компанией и пытаетесь выстроить стратегию работы с данными, которая должна помочь вам добиться успеха в эпоху цифровой революции. К вам приходит замечательный торговый агент и заявляет: «Наше новое программное обеспечение предназначено для того, чтобы расширить возможности для решения задач по данным и аналитике». Или обещает: «Наша программа решит все проблемы с данными и их анализом». Словом, чтобы убедить вас купить его ПО, он засыплет вас самыми разнообразными убедительными аргументами. Он покажет вам примеры из практики и данные исследований. Полюбовавшись на все это, вы решите вложить свои деньги в ПО и внедрить его в своей организации. Когда организация обеспечивает всем своим работникам равный доступ к аналитическим программам, это называется «демократизация данных». Открою небольшой секрет. С одной стороны, демократизация данных – именно то, что должна проделать каждая компания. С другой – это проблема. Давайте рассмотрим ее немного подробнее.

Исторически данные принадлежали сфере информации и технологий (или какой-то другой отдельной части бизнеса), так что лишь немногие сотрудники работали с ними, составляли отчеты и делали анализ. Таким образом, организация доверяла получение достоверных, практически применимых результатов этой ограниченной группе сотрудников. По мере того, как компании, предоставляющие инструменты бизнес-аналитики, такие как Qlik или Tableau, прогрессировали и превращались в гиганты разработки ПО, организации стали искать пути широкого распространения и использования данных. Таким образом они демократизировали данные и информацию, рассчитывая получить полезные знания и результаты. Но налицо одна большая проблема: многих ли в школе и университете учили работе с данными и их анализу? В последние годы количество таких людей растет, но что делать с теми, кто не получил соответствующего образования?

Может показаться, будто мы хотим сказать, что демократизация данных – это не решение. Не хотим! Демократизация данных полезна – благодаря инвестициям в данные и аналитику организации раскрывают свой потенциал. Демократизация данных позволяет использовать уникальные таланты и способности сотрудников, получать больше отдачи от вложений в ПО, данные и технологии. Причина, по которой демократизация данных может увеличивать недостаток знаний в организации, – образовательный фундамент ее коллектива. Если от людей, которые не обучались использованию данных и технологий, требуют освоить новое ПО и новые

технологии или извлекать пользу из данных и информации, они никак не смогут проделать это эффективно. Как вы думаете, многие ли из них с готовностью (и радостью) бросятся использовать эти новые технологии? Часто ли вы сами с восторгом соглашаетесь учиться чему-то новому, тому, чем вас внезапно нагружают по работе?

Вернемся к исследованию 2019 года о влиянии человеческого фактора на дата-грамотность и вспомним, какую проблему оно выявило: 36 % участников «предпочли бы искать иные методы решения задачи вместо использования данных», а 14 % – «постарались бы полностью уклониться от задачи», если бы им нужно было использовать данные. Это прекрасная демонстрация нехватки навыков работы с данными. А те, кто чувствует себя в этой области комфортно и уверенно, с большей охотой станут осваивать технологии. То же исследование показало, что почти три четверти участников (74 %) чувствуют себя расстроенными или несчастными при работе с данными. Этот показатель свидетельствует: внедрение новых технологий и демократизация данных утомляет сотрудников. Корень проблемы именно в этой всеобщей усталости и нехватке навыков использования данных.

Производство данных

Как связаны производство данных и отсутствие умения с ними обращаться? Выше уже говорилось, сколько данных ежегодно производится в мире. Технологический прогресс ускоряется, количество данных растет, а в итоге организации и их сотрудники оказываются не готовы к обработке всего этого объема. Организации, которые возникли уже в цифровую эпоху, подготовлены к лавине данных несколько лучше, поскольку именно данные во многом определяют их рабочие процессы. Но что делать более старым компаниям? Их руководители отчаянно пытаются создать возможности для эффективного использования данных – и обнаруживают, что это не так-то просто. Речь не о том, что у них не выходит получать и использовать данные. Они обнаруживают, что нельзя просто инвестировать в ПО и технологии, собрать данные... И вот ты уже царь горы – горы данных, доступных для анализа. В то же время, как показывают исследования, сотрудники просто не в состоянии достаточно быстро приспосабливаться к среде, которая становится все динамичнее. Отсюда и нехватка навыков работы с данными.

Отсутствие стратегии в сфере данных и их анализа

Еще одна причина нехватки навыков работы с данными – отсутствие у организаций соответствующей стратегии. Каков механизм этого? Для начала сядьте и подумайте: есть ли у вашей организации четкая и конкретная стратегия обращения с данными? Увы, многие из вас ответят, что нет. Отсутствие стратегии – это лишняя нагрузка на сотрудников, которые выбиваются из сил, пытаясь понять, как использовать ПО и технологии, в которые компания вложила деньги. (Помните исследование 2019 года? Люди действительно расстраиваются и устают!)

В чем важность для предприятия четкой и конкретной стратегии работы с данными и их анализом? Давайте вернемся к примеру с бегуном, которому также необходима определенная стратегия. Представьте себе начинающего (или не совсем начинающего, но в любом случае непрофессионального) спортсмена, который участвует в забеге.

Допустим, этот спортсмен – вы. Вы давно не бегали, но кто-то из ваших друзей, родных или коллег участвует в забегах. Вы видите, как это их воодушевляет, постоянно слушаете их разговоры (я сам бегун и прекрасно понимаю, что меня сложно заставить замолчать, если уж я оседлал любимого конька). И вы решаете тоже записаться на забег. До него остается несколько месяцев. Вы записались, но у вас нет стратегии тренировок, вы не изучили трассу, вы не представляете, как правильно питаться и сколько нужно пить жидкости, – однако тренируетесь и

едите, зная, что это необходимо. Кроме того, не имея стратегии, вы не знаете, какие вложения требуются для осуществления вашей мечты.

И вот наступил день забега. Вы стоите у стартовой линии, чувствуя себя более или менее подготовленным. У вас с собой, возможно, запас пищи и воды, и вы надеетесь, что тренировки помогут вам дойти до финиша. Но на самом деле вы не готовились должным образом – хотя, конечно, кое-как тренировались и даже купили спортивную форму. Если вы бежите полумарафон, то, возможно, справитесь с забегом, пусть и с трудом. Но и в полном, и в ультрамарафоне вам, скорее всего, придется сойти с дистанции.

А теперь представьте другой подход. Вы записываетесь на забег уже после того, как выработали стратегию, которая поможет вам добиться результата. Вы точно знаете, какое снаряжение вам необходимо, сколько питательных веществ и жидкости нужно вашему организму. Вы обратились к хорошему тренеру, который проследит за процессом вашей подготовки, давая необходимые указания. Итак, тренировки завершены, и вы выходите на старт. Теперь вы дойдете до финиша на любой дистанции – через боль, через усталость, но дойдете. Вы знали, что нужно делать, и достигли успеха именно благодаря стратегии и правильной подготовке.

В целом эти примеры наглядно показывают, что нужно организации от стратегии в сфере данных и их анализа. Слишком долго руководители полагались на более бессистемный подход. Они знали, что нужно инвестировать в данные и аналитику, но делали недостаточно, не разобравшись, какое снаряжение и зачем они приобретают (и подходит ли оно им вообще). Они не обращались к «тренерам», которые могли бы помочь им разработать и воплотить стратегию. А потом обнаруживалось, что инвестиции – порой миллионы долларов – не дают ожидаемых результатов. К несчастью, таких руководителей и компаний очень много.

Организация должна построить стратегию получения и использования данных, чтобы расширять видение, ставить цели и решать задачи. На сегодня большинству организаций во всем мире не хватает такой стратегии.

При отсутствии стратегии навыки обращения с данными у сотрудников только ухудшаются. Вместо того чтобы выработать четкую стратегию, определяющую и обеспечивающую поток инвестиций в ПО, технологии и обучение персонала, компании просто покупают ПО и технологии, рассчитывая, что все как-то само собой заработает. Налицо подмена понятий: руководители думают, что технологии – это и есть стратегия, и навязывают их сотрудникам. Но в таком случае стратегия не определяет, какую технологию использовать. В результате сотрудники отвергают новшества и возвращаются к старым способам решения задач. А новая технология как ненужная игрушка, которая лежит на полке и пылится.

Это приводит к двойной проблеме. Во-первых, ПО, купленное компанией, не внедряется или не используется сколько-нибудь эффективно. Во-вторых, вместо совершенствования навыков работы с данными сотрудники компании лишь еще больше отстают, потому что не желают осваивать программы и технологии, которые были приобретены, казалось бы, для их удобства.

Данные: что дальше?

Итак, мы выяснили, что при нехватке у сотрудников соответствующих навыков данные не используются достаточно эффективно. И что же делать? Действительно ли необходимо бороться с этим недостатком и ликвидировать пробелы – или все-таки можно продолжать работать как раньше? Ответ очевиден: ликвидировать пробелы необходимо!

Как уже было сказано, к 2025 году будет производиться примерно 463 эксабайта данных ежедневно: это наше будущее. Давайте еще нагляднее, без примеров с DVD: один эксабайт – это единица с 18 нулями. Так что представьте себе число 463 и припишите к нему 18 нулей. По другим прогнозам, к 2025 году количество производимых ежедневно данных будет равняться 175 секстибайтам, а секстибайт – это единица с 21 нулем: 1 секстибайт равен триллиону гигабайт¹⁴. Так какой же прогноз верен? Или лучше задать другой вопрос: а так ли это важно? Ведь это просто *очень много данных*, и нет никаких сомнений, что где-то среди них скрывается много ценных знаний. И мы снова возвращаемся к той же проблеме: если нам так не хватает навыков в обращении с данными, смогут ли отдельные люди и организации воспользоваться этим огромным объемом данных с выгодой для себя? Или большинству придется бессильно наблюдать, как организации, умеющие использовать данные, легко обходят конкурентов?

¹⁴ Morris, T. (2020). How Much Data by 2025? [Blog], Microstrategy, 6 January. <https://www.microstrategy.cn/us/resources/blog/bi-trends/how-much-data-by-2025>.

Краткое содержание главы

В целом можно сказать, что мир данных, в котором мы живем, – удивительный, пугающий и неизведанный. В будущем появятся новые профессии, возможности и изобретения, и мы пока не можем сказать, что повлекут за собой все эти новшества. Наверняка возникнет необходимость в новых навыках, но мы не знаем, каких именно. Одно ясно уже сейчас: данные были и будут всегда! Сегодня налицо тенденция быстрого и масштабного накопления данных, которые необходимо обрабатывать. Есть и другая тенденция: все очевиднее нехватка навыков у сотрудников организаций. Это мешает организациям добиваться успеха и эффективно использовать инвестиции в данные и их анализ. Что же мы можем сделать? Есть ли решение этой проблемы? Конечно же, да! Перед людьми и организациями, которые смогут освоить мир дата-грамотности, откроются огромные возможности.

02

Четыре уровня аналитических методов

Данные и их анализ – целых четыре уровня?

Теперь, когда мы получили представление о мире данных, нам нужно целостное понимание методов анализа данных. Только в этом случае люди и организации смогут эффективно использовать данные и извлекать выгоду из аналитики. Понимание – ключ к реализации работающей стратегии в области работы с данными. Если понимания нет, руководители могут закупать сколь угодно дорогое и мощное ПО, получать данные и демократизировать инструменты работы с ними, но им будет непонятно, действительно ли их организации нужно именно это. Основа мира аналитики – четыре уровня: **дескриптивный** (описательный), **диагностический**, **предиктивный** (предсказательный) и **прескриптивный** (предписывающий). Чтобы осознать необходимость изучения этих уровней, давайте заглянем в прошлое: как организации исторически воспринимали данные и методы их анализа? Эта картина поможет нам понять, как относиться к ним теперь и как можно использовать эти четыре уровня для достижения успеха.

Итак, организации, данные и дата-аналитика: исторический аспект. Организации всегда вкладывали много средств в технологии, программное обеспечение и инструменты работы с данными. ПО и технологии считались манной небесной, все были уверены, что с их помощью можно найти любые решения в области данных и аналитики. Руководители инвестировали в технологии, которые могли помочь им добиться целей и воплотить в жизнь самые смелые мечты из области данных и дата-аналитики (а такие мечты у руководителей есть всегда). Со времен появления самой первой версии Microsoft Excel в 1985 году¹⁵ (кстати, если вы не в курсе, то сначала эту программу выпустили для Apple Macintosh – странно, правда?) и даже более раннего ПО для электронных таблиц организации постоянно покупали либо продавали ПО, считая, что оно способно все решить, творя некую «магию». Итак, давным-давно существует ПО для хранения, использования и анализа данных и информации. По мере развития и совершенствования технологий также росли и инвестиции в них. По прогнозам, в 2019 году доходы, полученные во всем мире от анализа больших объемов данных и бизнес-информации, должны были составить 187 миллиардов долларов¹⁶. Этот рынок продолжает развиваться столь же бурно... однако существует и неприятная тенденция. Учитывая нехватку знаний у 24 % людей, ответственных за принятие бизнес-решений, и у 32 % топ-менеджеров, инвестиции в технологии сбора и анализа данных не приносят ожидаемых результатов. Эта нехватка знаний, о которой мы говорили в первой главе, мешает инвестициям приносить свои плоды.

Итак, каким же образом понимание четырех уровней аналитических методов влияет на грамотное внедрение технологий и возврат инвестиций в данные и аналитику? Когда организации и сотрудники разберутся, как работают все четыре уровня, это позволит понять, как коллектив и отдельные сотрудники со своими индивидуальными навыками при помощи технических возможностей могут совместно построить правильную стратегию работы с данными и дата-аналитикой. А затем эту стратегию можно будет применить.

¹⁵ CIS Poly. History of Microsoft Excel. http://cis.poly.edu/~mleung/CS394/f06/week01/Excel_history.html.

¹⁶ Olavsrud, T. (2016). Big Data and Analytics Spending to hit \$187 Billion, CIO, 24 May. <https://www.cio.com/article/3074238/big-data-and-analytics-spending-to-hit-187-billion.html>.

Четыре уровня аналитических методов

А сейчас, после краткого экскурса в историю, давайте попробуем разобраться в самих четырех уровнях аналитики. Затем мы покажем вам, как сотрудники на самых разных уровнях, от рядовых работников до топ-менеджеров, могут применять эти четыре уровня: 1) для продвижения стратегии работы с данными и аналитикой; 2) для принятия более разумных решений на основе данных; 3) для формирования правильного видения данных и аналитики. Понимание четырех уровней аналитики позволяет организации не стрелять по мишени вслепую, а осуществлять реальную работу по формированию стратегии.

Как уже говорилось выше, четыре уровня аналитических методов – это дескриптивный, диагностический, предиктивный и прескриптивный (см. рис. 2.1). Чтобы получить прочные базовые представления о каждом уровне, давайте для начала рассмотрим их определения и примеры. Кроме того, мы также познакомимся с конкретными программами и технологиями, соответствующими каждому уровню. Затем, разобравшись с каждым уровнем, мы увидим, как они работают в комплексе: формируют правильную аналитическую картину и помогают организации успешно освоить инвестиции в данные.



Рис. 2.1. Четыре уровня аналитической мозаики

Прежде чем перейти к первому уровню, давайте разберемся, что на самом деле в нашем контексте означает слово «аналитика». Мы слышим его постоянно, но известно ли нам его истинное значение?

Если взглянуть на определение нужного нам значения слова «аналитика», то окажется, что аналитика – это «систематический вычислительный анализ данных или статистики»¹⁷

¹⁷ Lexico.com, определение Analytics. <https://www.lexico.com/en/definition/analytics>.

Конец ознакомительного фрагмента.

Текст предоставлен ООО «ЛитРес».

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

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