



Гайд

Как изучить Data Science самостоятельно

Подборка ресурсов

В области Data Science есть 3 специальности: **аналитика данных, дата-инженерия и машинное обучение**. У всех — единая база, которую должен освоить дата-сайентист: программирование на Python, математика, статистика и, желательно, английский язык.

Многие пробуют освоить это всё самостоятельно по разным источникам. Не будем лукавить, это не так просто. В этом гайде мы постарались собрать ресурсы для тех, кто хочет разобраться в Data Science. Вы узнаете, где можно прокачать английский, какую теорию подтянуть и какие книги полезно прочитать. А заодно мы поделимся советами по самообразованию и расскажем истории успешных дата-сайентистов.



Выучите английский язык

Для работы в России вам хватит уровня Intermediate. С ним вы сможете читать статьи и следить за новостями в сфере Data Science. Если планируете работать на международном рынке, нужно стремиться к Advanced.

Можно стать дата-сайентистом и без английского. Но если вы хотите разбираться в англоязычных гайдах, читать форумы и потом устроиться в зарубежную компанию, нужно освоить технический язык. Вот что вам поможет:

1

Сайт [Lingualeo](#). Пройдите тестирование, укажите направление «IT» и изучайте английский по статьям, текстам и подкастам.

2

Книга Виная Триведи [How to Speak Tech](#). Разберёте все актуальные IT-концепции: фронтенд и бэкенд, UI, интернет вещей, ИИ, блокчейн.

3

Книга Джона МакЮэна [Oxford English for Information Technology](#). Это университетское пособие Оксфорда для тех, кто хочет прокачать языковые скиллы в IT.

Подтяните математику

Это навык, без которого в Data Science не обойтись. Будущий специалист должен идеально знать основы математики, разбираться в статистике, работать с функциями и матрицами. Всю эту информацию можно найти в пособиях по Data Science и университетских учебниках или поискать на форумах.

Самое сложное — выбрать среди доступных источников действительно полезные. Где-то может не быть нужных тем, какие-то авторы глубоко изучают одно направление и поверхностно рассматривают другие. Чтобы не потеряться в ворохе книг и ресурсов, предлагаем:

- Почитать [наш гайд по математике](#). В нём мы объяснили законы и понятия, про которые спрашивают на собеседованиях на должность Junior Data Scientist.
- Повторить все базовые темы [на сайте «Математика с нуля»](#). По порядку пройдёте все важные темы, начиная от дробей и последовательности действий, заканчивая уравнениями с модулем и квадратными уравнениями.

Когда освежите базу, можете переходить к трём ключевым направлениям:

- [«Практическая статистика и теория вероятностей»](#). Книга для будущих дата-сайентистов, в которой собраны ключевые понятия статистики: разведочный анализ данных, статистические эксперименты, распределение данных и многое другое.
- [«Математический анализ»](#). Курс лекций от преподавателей Московского государственного университета. Разберётесь с матрицами, определителями, пространствами и функциями.
- [«Линейная алгебра»](#). Учебник, в котором разбираются понятия и исследования математического анализа.

Все важные основы и специальные знания для дата-сайентистов собраны на курсе [Data Scientist PRO](#). Вы сможете изучать то, что пригодится в профессии. Например, узнаете про статистические тесты для проверки гипотез и освоите методы распределения, которые помогут составить модели машинного обучения.

Освойте и практикуйте программирование

Изучать математику можно параллельно с программированием. Когда вы начнёте делать первые практические задачи, сможете обращаться к нужным разделам математики. Самые важные языки программирования для дата-сайентиста — SQL и Python.

SQL можно освоить тут:

- [MySQL.ru](#). Простой и понятный сайт для пошагового изучения SQL. Узнаете, для чего нужен этот язык, научитесь работать с таблицами и использовать операторы. В конце каждого раздела есть практические задания.
- [«SQL. Сборник рецептов»](#). Книга с практическими советами по работе с системами баз данных: MySQL, PostgreSQL, DB2, ORACLE.

А узнать про Python — здесь:

- Маккини Уэс — [«Python и анализ данных»](#). Книга для новичков в аналитике о форматировании, очистке и обработке данных на Python.
- Охеда Тони, Билбро Ребекка — [«Прикладной анализ текстовых данных на Python»](#). Методы машинного обучения для анализа текста. Научитесь обрабатывать данные, классифицировать текст, анализировать графы и проводить глубокое обучение.
- [«Быстрое введение в Python»](#) от Skillbox. Объясним все ключевые понятия языка: переменные, строки, списки, срезы, цикл for и отступы.

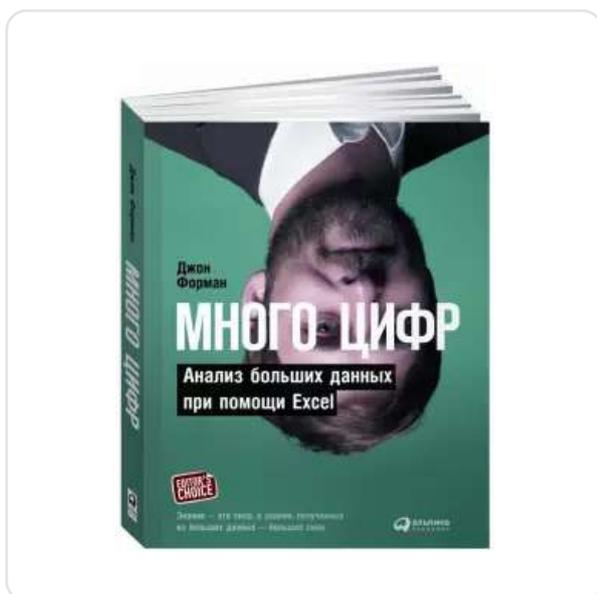
От теории — к практике! Каждую тему важно закреплять практическим заданием, чтобы развить реальные навыки и понять, какие задачи вас ждут. Так, например, и проходит обучение на платформе Skillbox. После каждого модуля вы выполняете работу, которую проверяет эксперт и даёт профессиональный фидбэк.

Уже через 6 месяцев вы накопите знаний, чтобы найти стажировку. А с дальнейшим трудоустройством вам помогут карьерные консультанты: составят резюме, подготовят к собеседованию и подберут вакансии.

С чего начать практику и как строить карьеру — рассказываем в этих статьях:

- Создаём первую модель машинного обучения — на примерах объясним, как сформировать датасет, обучить модель и написать код на Python для будущих предсказаний.
- Где новичку в Data Science искать проекты для портфолио — советы от аналитика по поиску работы: мероприятия, курсы, соревнования, сообщества.

Почитайте литературу по Data Science



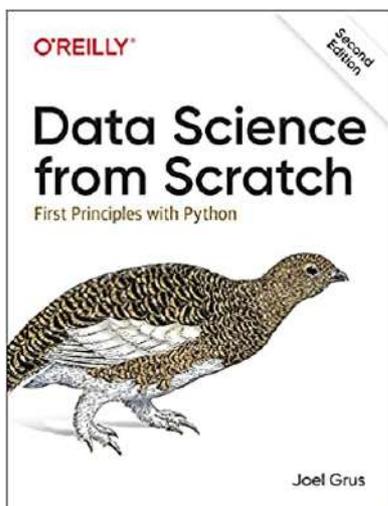
«Много цифр. Анализ больших данных при помощи Excel». Джон Форман

Понятный взгляд на Big Data. Автор расскажет, как организовать анализ данных в простом Excel.



«Роман с Data Science». Роман Зыков

Владелец международной компании по применению ИИ и автор аналитики OZON поможет освоиться в Data Science и разобраться в анализе данных.



«Data Science from Scratch». Джоэл Грус

Быстрый курс по Python для Data Science. Изучите не только библиотеки и фреймворки, но и поймёте базовые принципы анализа данных.

Отдельная рекомендация — подборка книг по машинному обучению. 15 материалов на русском и английском языках для новичков и начинающих специалистов в Data Science.

Освойте и практикуйте программирование

Освоить обширную тему с нуля непросто. Сесть и в один день начать изучение не получится — нужно подобрать авторитетные источники, составить график обучения. А потом постоянно практиковаться и следить за своим прогрессом. Многим не хватает мотивации, чтобы контролировать всё это самостоятельно, и они бросают обучение. Но есть ресурсы, которые помогут подготовиться к длительному самообразованию:

- Полезные советы от Skillbox. Честно рассказываем про долгий и сложный путь дата-сайентиста. Узнаете, какие знания вам пригодятся и как наработать опыт.
- «Думай как математик», Барбара Оакли. В книге собраны образовательные методики от преподавателей и психологов. Они помогут вам научиться фокусироваться и быстро включаться в работу, натренировать память и выбрать режим работы.
- «Суперобучение», Скотт Янг. 9 принципов, которые помогут ускорить процесс обучения. Сам автор за год выучил 4 языка и освоил четырехлетнюю программу курса Computer Science в MIT.

Масштабные темы удобнее изучать, если есть помощник, который проведёт вас по всем этапам. У вас будет такой человек на курсе Data Scientist PRO. Вы сможете задать любые вопросы эксперту в телеграм-чате и получить профессиональные советы. А проверяющие эксперты дадут фидбэк по практическим заданиям и посоветуют дополнительные материалы.

Вступите в тематические сообщества

Data Science постоянно развивается, поэтому будущему специалисту нужно быть в курсе новостей. В сообществах вы сможете изучить опыт других, обсудить вопросы, а может, даже завести полезные контакты.

- [Gonzo-обзоры ML статей](#) — Telegram-канал, где публикуют и обсуждают материалы по машинному обучению: статьи, обзоры и примеры кода.
- [Data Science Weekly](#) — бесплатная email-рассылка с новостями из мира Data Science. Можно найти не только статьи, но и вакансии.
- [Kaggle Blog — Medium](#) — медиа про Kaggle. Заодно можете зарегистрироваться на самой платформе, чтобы участвовать в соревнованиях и решать задачи. Если не знаете, что такое Kaggle, [читайте нашу статью](#).

Вдохновляйтесь историями тех, кто нашёл призвание в Data Science после курса Skillbox



[Артём Лёвкин](#)

Ближе к концу курса Data Scientist PRO Артём начал искать работу на неполный рабочий день. В Центре карьеры ему помогли найти вакансии. В итоге Артём попал в крупнейшую аудиторскую компанию Deloitte, не успев закончить курс.

[Читать историю Артёма](#)



[Жираслан Шеджем](#)

Жираслан всегда любил кодить, но программистом не стал. Он также увлекался экономикой и бизнесом, поэтому решил уйти в дата-аналитику. Теперь Жираслан работает в «Тинькофф Банке».

[Читать историю Жираслана](#)



Ксения Титова

После переезда в Грецию Ксения два с половиной года работала в продажах, но не любила это дело. Она решила поменять свою жизнь, заинтересовалась Big Data и теперь работает бизнес-аналитиком в «Тинькофф Банке».

[Читать историю Ксении](#)



Владимир Андриенко

Владимир с детства знал, что будет работать в IT. Он окончил радиотехнический вуз и какое-то время работал на фрилансе. После одного заказа Владимир решил уйти в машинное обучение и уже разработал собственную нейросеть.

[Читать историю Владимира](#)



Даша Бокарева

Даша изучала бизнес-информатику в университете, но решила сменить сферу и ушла в Data Science. После хакатона она ведёт собственный проект и работает в стартапе.

[Читать историю Даши](#)

Все эти ресурсы помогут вам окунуться в Data Science самостоятельно. Первые челленджи, которые вас ждут, — выбрать подходящие материалы и составить график обучения. А в процессе вы будете контролировать свой прогресс, самостоятельно находить ошибки и искать ответы на вопросы на форумах и в сообществах.

А для тех, кто хочет сразу получить структурированные знания и помощь экспертов, мы сделали курс Data Scientist PRO. Все материалы подготовлены для постепенного погружения в профессию с нуля до уровня профессионала. Каждую тему подкрепляют полезные материалы, а сам курс постоянно обновляется с учётом новых требований и трендов в Data Science.

Вас ждут 3 уровня: базовая подготовка, специализация и повышение квалификации. Начнёте с математики и статистики, попробуете себя в каждой сфере Data Science и сможете углубиться в интересующую специализацию. А в конце курса карьерные специалисты помогут найти работу.