Маме к юбилею
и с благодарностью за лучшее в мире детство!
Спасибо, что помогала, когда могла,
и не мешала, когда не могла помочь!

Содержание

От издательства 8

Об авторе 9

Предисловие к первому изданию 10

Предисловие ко второму изданию 11

Введение 13

Глава 1. Что такое Yandex DataLens? 15

Способы развертывания DataLens 16
Облако 16
Open source 17
DataLens бесплатный, но за деньги да 19
Возможности, пришедшие в единый тариф из тарифа Business 20
Входит и выходит. Замечательно выходит! 21
DataLens Галерея 21
Мы будем жить теперь по-новому (воркбуки и коллекции) 23
Навигация по воркбукам и коллекциям 24
Создание воркбуков и коллекций 26
Особенности коллекций и воркбуков 26
Миграция объектов в воркбуки 27
Экспорт и импорт воркбуков 31
Настройка доступа к коллекциям и воркбукам 33
Настройки сервиса 35
Публичный API DataLens 36

Глава 2. Подключения – получение данных 38

Файлы и сервисы 40
Файлы 41
Yandex Query и проект № 1 43
Яндекс Документы 106
Google Sheets 107
Metrica 109
AppMetrica 110
Базы данных 111
Подключение к внешним базам данных 112
Подключение к базам данных в облаке Yandex 114

Глава 3. Датасеты – переход от подключений к чартам 190
Вкладки датасета 191
Источники 191
Поля 198
Параметры 208
Фильтрация 211
Настройки датасета 212
Версионирование 214

Глава 4. Формулы и функции – кирпичики вычислений 217
Синтаксис формул 218
Агрегатные функции 219
Прямая агрегация 221
Условная агрегация 226
Статистическая агрегация 229
Приблизительная агрегация 236
Строковая агрегация 236
Логические функции 238
Операторы 241
Математические операторы 241
Логические операторы 242
Математические функции 245
Функции округления 245
Функции сравнения 246
Арифметические функции 247
Тригонометрические функции 248
Строковые функции 248
Функции поиска в строке 248
Функции строковых преобразований 253
Функции для работы с регулярными выражениями 256
Функции преобразования типов 260
Функции даты и времени 263
Функции работы с датами 264
Функции возвращения части даты 266
Функции для работы с временными рядами 266
Функции для работы с массивами 269
Функции разметки 273
Нативные функции 275
Проект № 4. Кластеризация вин при помощи метода DBSCAN 276
Заключение 305

Глава 5. Оконные функции и LOD-выражения – просто о сложном 306
Оконные функции 307
Агрегатные оконные функции 310
LOD-выражения 349

Фиксированный набор измерений (FIXED) ... 350
Расширенный и уменьшенный набор измерений (INCLUDE и EXCLUDE) ... 355
Примеры использования LOD-выражений ... 356
Проект № 5. ROC-анализ ... 365
Введение в бинарную классификацию ... 365
Метрики бинарной классификации ... 367
ROC-анализ ... 371
Подготовка данных ... 374
Реализация проекта в DataLens ... 387
Использование дашборда и анализ разных наборов данных ... 414
Заключение ... 417

Глава 6. Чарты – основа дашбордов ... 418

Чарты на основе датасета ... 418
Создание чарта на основе датасета ... 419
Разновидности чартов ... 420
Линейная диаграмма ... 421
Диаграмма с областями ... 428
Нормированная диаграмма с областями ... 430
Столбчатая диаграмма ... 431
Нормированная столбчатая диаграмма ... 435
Линейчатая диаграмма ... 436
Нормированная линейчатая диаграмма ... 436
Точечная диаграмма ... 437
Круговая диаграмма ... 439
Кольцевая диаграмма ... 440
Индикатор ... 443
Древовидная диаграмма ... 445
Комбинированная диаграмма ... 445
Таблица ... 446
Сводная таблица ... 448
Карта ... 450
QL-чарты ... 452
Создание QL-чарта ... 453
Привязка QL-чартов к виджетам на дашборде ... 456
Настройки чартов ... 462
Общие настройки ... 462
Настройки измерений и показателей ... 463
Настройки секций ... 465
Настройки секций на карте ... 468
Параметры уровня чарта ... 470
Иерархии в чартах ... 471
Диагностика чартов ... 473
Публикация чартов ... 474
Встраивание публичных и непубличных чартов ... 476
Версионирование ... 480

Экспорт данных из чартов ... 480
Заключение ... 482

Глава 7. Дашборды и отчеты – ваши холсты и ваши краски ... 483
Проект № 6. Прогнозирование временных рядов ... 484
Обработка исходных данных и построение моделей в Python ... 485
Создание подключения в DataLens ... 498
Создание датасетов ... 499
Создание чартов ... 499
Создание дашборда ... 506
Текстовые виджеты и заголовки ... 520
Селекторы JS ... 523
Дополнительный функционал дашбордов ... 528
Настройки дашборда ... 532
Отчеты в DataLens ... 534
Заключение ... 540

Глава 8. Чарты и селекторы в Editor – чистое творчество ... 541
Интерфейс Editor ... 541
Вкладки в Code editor на примере создания таблицы ... 543
Meta ... 543
Params ... 544
Sources ... 545
Config ... 546
Prepare ... 546
Controls ... 548
Создание графика в Editor ... 548
Создание составного селектора в Editor ... 554
Создание маркдауна в Editor ... 559
Создание Advanced-чартов в Editor ... 560
Создание дашборда на основе чартов из Editor ... 561
Заключение ... 564

Предметный указатель ... 565

От издательства

Отзывы и пожелания

Мы всегда рады отзывам наших читателей. Расскажите нам, что вы думаете об этой книге – что понравилось или, может быть, не понравилось. Отзывы важны для нас, чтобы выпускать книги, которые будут для вас максимально полезны.

Вы можете написать отзыв на нашем сайте **www.dmkpress.com**, зайдя на страницу книги и оставив комментарий в разделе «Отзывы и рецензии». Также можно послать письмо главному редактору по адресу **dmkpress@gmail.com**; при этом укажите название книги в теме письма.

Если вы являетесь экспертом в какой-либо области и заинтересованы в написании новой книги, заполните форму на нашем сайте по адресу **http://dmkpress.com/authors/publish_book/** или напишите в издательство по адресу **dmkpress@gmail.com**.

Список опечаток

Хотя мы приняли все возможные меры для того, чтобы обеспечить высокое качество наших текстов, ошибки все равно случаются. Если вы найдете ошибку в одной из наших книг, мы будем очень благодарны, если вы сообщите о ней главному редактору по адресу **dmkpress@gmail.com**. Сделав это, вы избавите других читателей от недопонимания и поможете нам улучшить последующие издания этой книги.

Нарушение авторских прав

Пиратство в интернете по-прежнему остается насущной проблемой. Издательство «ДМК Пресс» очень серьезно относится к вопросам защиты авторских прав и лицензирования. Если вы столкнетесь в интернете с незаконной публикацией какой-либо из наших книг, пожалуйста, пришлите нам ссылку на интернет-ресурс, чтобы мы могли применить санкции.

Ссылку на подозрительные материалы можно прислать по адресу электронной почты **dmkpress@gmail.com**.

Мы высоко ценим любую помощь по защите наших авторов, благодаря которой мы можем предоставлять вам качественные материалы.

Об авторе



Если вы читали первое издание книги, значит, мы уже заочно знакомы. Если же нет, позвольте представиться. Меня зовут Александр Гинько, и я более пяти лет занимаюсь переводом книг в сотрудничестве с издательством «ДМК Пресс». За это время успел перевести уже почти 30 книг и продолжаю переводить новые. Общий стаж переводов – более 15 лет.

Среди переведенных мной книг – все бестселлеры известных итальянских специалистов и популяризаторов Power BI и DAX Марко Руссо и Альберто Феррари, а также книги по Tableau, Excel, Power Query, Python, R, SQL, машинному и глубокому обучению.

Каждая новая книга для меня – в первую очередь возможность узнать что-то новое, а во вторую – помочь узнать это другим. В процессе перевода я стараюсь полностью погружаться в предметную область и проверять все приведенные в книгах примеры. Могу сказать, что до сих пор мне не было стыдно ни за один переведенный мною абзац.

Написание первого издания книги по *Yandex DataLens* стало для меня пробой пера в качестве автора, и, судя по отзывам, проба удалась. Теперь осталось не уронить марку и доходчиво рассказать читателям о *DataLens* так, как вижу этот инструмент я.

Именно как вижу его я. Не стоит относиться к этой книге как к инструкции или руководству от самих разработчиков *Yandex DataLens*. И второе издание, как и первое, ни в коем случае не является рекламой! Ни Павел Дубинин, которого я безмерно уважаю, ни кто-либо другой никоим образом не влияли на концепцию и тональность написанного материала. Иначе бы я за это не взялся.

Как и в случае с первым изданием, я просто попытался сделать эту книгу такой, какую сам хотел бы почитать, если бы мне понадобилось изучить новый инструмент. Надеюсь, у меня снова получилось.

Всю информацию, анонсы и промокоды со скидками на покупку моих книг вы можете найти на моем канале в Telegram, подписывайтесь и следите за новостями: **https://t.me/alexanderginko_books**. Также у меня есть собственноручно написанный бот, который поможет вам узнать подробнее обо всех моих книгах и наблюдать онлайн за переводом новых книг: **https://t.me/alexanderginko_books_bot**.

В добрый путь!

Предисловие к первому изданию

«Современный мир, это данные! – Данные, это современный мир!» Впрочем, знакомить их уже не нужно, они давно знакомы. Сегодня ни одну область жизни невозможно представить без данных, их анализа и визуального представления.

Кто более или менее следит за тенденциями в области обработки и потребления информации, меня поймет. Еще совсем недавно – буквально два десятилетия назад – визуальному представлению данных уделялось не так много внимания. Я лично, будучи в ту пору разработчиком, программистом и внедренцем проектов на базе 1С:Предприятия, прекрасно помню, что львиная доля задач при работе даже с крупными организациями сводилась к реализации требуемой бизнес-логики на уровне расчетов и хранения данных, а о визуальной части никто даже не задумывался. Многие из вас не поверят, но тогда (совсем недавно) были огромной редкостью даже самые простые графики и диаграммы в отчетах, предназначавшихся для руководителей разного уровня. Вместо этого большая часть информации представлялась в виде таблиц, пределом ухищрений в которых были раскраска строк в разные цвета и нехитрое (и крайне неудобное) объединение заголовков столбцов.

«Да, были люди в наше время»… Не то чтобы я имел что-то против нынешнего племени, но вряд ли сегодня можно представить себе руководителя даже самого незначительного уровня, перебирающего сшитые в папку листы А4 с бесконечной таблицей в поисках нужной ему продажи. Прогресс клипового мышления, при котором человек воспринимает информацию фрагментарно, добрался и до анализа данных, и в результате прежние таблицы (так называемые «портянки») превратились в сжатые и очень лаконичные визуальные представления, получившие название дашборды.

Кстати, за названием долго ходить не пришлось. Дашборд (dashboard) в переводе с английского – это приборная панель. А где еще, как не в автомобиле, нам нужно иметь возможность мгновенно оценивать информацию при беглом взгляде на панель? Если бы вам пришлось прокручивать таблицу в поисках вашей текущей скорости или уровня топлива, далеко бы вы не уехали.

Современные информационные дашборды не уступают по степени информативности приборной панели автомобиля, и с их помощью руководители отделов и предприятий могут очень быстро и в сжатом виде получать «слепок» деятельности подразделений, не тратя на это много времени.

Данные для анализа в современном мире должны быть представлены в виде удобных и красочных дашбордов, это всеми признанный факт. А что нужно, чтобы строить лаконичные визуальные представления без лишних усилий? Средств создания богатых визуализаций на рынке масса, и в этой книге мы подробно, насколько это возможно, поговорим об одном из них, а именно о *Yandex DataLens*…

Предисловие ко второму изданию

5 сентября 2022 года. Именно в этот день, как оказалось, я внес заключительную правку в текст своей первой авторской книги под названием «Анализ и визуализация данных в Yandex DataLens». Сегодня 7 июля 2025 года, и с тех пор прошло уже почти три года. За это время я успел добавить к своим переводам еще книг 12–13, а за развитием Yandex DataLens наблюдал вполглаза, изредка читая новости в их пабликах. И вот 13 декабря 2024 года я получаю сообщение от руководителя продуктового развития DataLens Павла Дубинина, который всячески поддерживал меня при написании первой книги. Начиналось сообщение так: «Александр, добрый день! Как твои дела? Как в итоге оцениваешь свой эксперимент с книгой DataLens? Не думал о переиздании? У нас многие спрашивают...» Признаюсь, до того момента не думал. Но потом ка-а-ак подумал...

Вспомнив опыт написания первого издания и все те сложности, с которыми пришлось столкнуться в процессе, мне захотелось вновь окунуться вот в это вот все. А подробно ознакомившись со всеми изменениями, появившимися в DataLens за последние три года, я понял, что это уже совершенно новый продукт, а значит, первое издание книги действительно требует обновления. Следующие полгода ушли у меня на разработку концепции новой книги (за это время я успел перевести еще две большие книги), оценку предстоящей работы, поиск и анализ подходящих наборов данных, построение ETL-процессов и разработку дашбордов, которые легли в основу второго издания книги.

Могу сказать, что я на своем веку переводчика видел немало вторых изданий книг. Часто они представляли собой попытку «выехать на старом материале», а иногда были почти полной копией первого издания с несколькими изменениями. В случае с этой книгой все будет не так. Я решил полностью переработать концепцию и подачу, а повторы (о них я буду сообщать) будут касаться только вещей, не претерпевших изменений и, как мне кажется, хорошо изложенных в первом издании. К тому же мне не хотелось, чтобы второе издание можно было рассматривать лишь в качестве дополнения к первому. Я ставлю себе целью, чтобы любой читатель, не державший в руках мое первое вытворение, мог без труда освоить DataLense с абсолютного нуля.

В то же время на этот раз я решил не ограничиваться одним лишь описанием инструментов DataLens. Конечно, я досконально расскажу про все новинки в DataLens, произошедшие с момента написания первого издания книги, а также плотно зайду на территорию революционной новинки в виде Editor с его кастомными диаграммами. Но мы с вами пойдем дальше. Современный мир анализа данных предъявляет серьезнейшие требования, связанные со

знанием десятков смежных дисциплин, начиная от базовой статистики и заканчивая построением полноценных конвейеров обработки сырых данных и применением к ним методов машинного и глубокого обучения. Именно поэтому в процессе чтения книги вы познакомитесь со способами предварительной обработки данных с помощью популярного языка Python и увидите несколько примеров применения несложных методов статистики и машинного обучения. В частности, мы реализуем в DataLens визуализацию на основе прогрессивного алгоритма кластеризации данных DBSCAN, построим в DataLens прогнозы временных рядов с использованием всех современных методов, включая SARIMA, градиентный бустинг и XGBoost, а также освоим бинарную классификацию и собственноручно с нуля выполним ROC-анализ моделей, построенных с помощью логистической регрессии, случайного леса и градиентного бустинга. Не стоит воспринимать это как отхождение от темы. Все это напрямую связано с анализом данных и когда-нибудь обязательно вам пригодится! Учиться стоит всему, что плохо лежит. А если оно еще и не приколочено...

Ну что ж, погнали!

Введение

Для кого предназначена эта книга

Если вы хотите приобрести навыки работы с *Yandex DataLens* и научиться самостоятельно выполнять все действия по извлечению, обработке, преобразованию и визуализации данных, эта книга точно не отяготит вашу полку. Мы очень подробно пройдем по всем этапам и досконально погрузимся в тонкости работы всех инструментов и механизмов *DataLens*.

Есть у вас опыт работы в других системах бизнес-аналитики или нет – абсолютно не важно, поскольку все темы мы будем обсуждать с самого нуля, что позволит вам уверенно чувствовать себя во всех без исключения аспектах *DataLens* – от загрузки и обработки сырых исходных данных при помощи богатого спектра функций и до размещения на вашем красочном дашборде очередного чарта с динамическим выбором измерений и цветными подписями, которые можно еще и отключить, просто сняв флажок.

Если вы уже читали первое издание книги, вам будет не менее интересно, поскольку у вас, я надеюсь, ни разу не возникнет ощущения, что вы уже где-то это читали (за исключением повторений, о которых я будут сообщать). Кроме того, я буду особенно фокусировать внимание на новинках, появившихся с момента выхода первого издания.

Конечно, было бы неплохо иметь хотя бы базовые навыки работы с Python, поскольку мы будем активно использовать этот популярный язык в процессе подготовки данных. Если же вас интересует только визуализация, вы можете пропускать фрагменты кода и приступать к чтению, когда повествование переходит к построению чартов и дашбордов. Вряд ли Лев Толстой был в обиде на тех, кто в его «Войне и мире» прочитал только о мире. А я кто такой, чтобы этому противиться?

В общем, если вы уже дочитали до этого места, эта книга точно для вас! Иначе бы мы с вами сейчас не разговаривали...

Структура книги

Книга, которую вы держите в руках, является сугубо практической, в ней вы встретите множество рекомендаций по поводу того, какими инструментами и техниками можно воспользоваться при реализации того или иного механизма. Если кратко говорить о содержании глав, оно будет таким:

- в главе 1 мы узнаем, что из себя представляет Yandex DataLens, какие способы его развертывания существуют, а также обсудим новую систему навигации при помощи коллекций и воркбуков, которая пришла на смену традиционным и не очень удобным папкам;

- глава 2 будет целиком и полностью посвящена различным видам подключения к источникам в DataLens. Но этим мы не ограничимся и реализуем сразу три несложных проекта из шести, которые вы увидите в данной книге;
- в главе 3 мы поговорим о том, что такое датасеты, модель данных, вычисляемые поля, измерения и показатели. Также мы остановимся на параметризации и фильтрации датасетов;
- глава 4 подробно расскажет вам о формулах и функциях, применяющихся при написании вычисляемых полей в DataLens. Помимо этого, мы реализуем очередной проект, в котором воспользуемся многими из описанных функций;
- в главе 5 мы отдельно поговорим об оконных функциях и LOD-выражениях. Это очень большая и для многих сложная тема, и я постараюсь досконально в нее погрузиться и вас с собой погрузить. Также мы вместе реализуем пятый проект в книге, который будет посвящен ROC-анализу;
- глава 6 будет посвящена чартам в DataLens как основному инструменту создания отдельных визуализаций. Мы обсудим все разновидности чартов, включая QL-чарты, их настройки, параметризацию, версионирование, публикацию и т. д.;
- в главе 7 мы перейдем к обсуждению дашбордов, объединяющих в себе все множество чартов. При этом мы будем учиться на примере очередного, заключительного проекта в этой книге, посвященного прогнозированию временных рядов. Также я вас немного напугаю добавлением на дашборд селектора на основе JavaScript, чтобы вы могли подготовиться к заключительной главе книги;
- в главе 8 будут описаны механизмы создания чартов и селекторов в мощнейшем инструменте Editor при помощи языка JavaScript, а также процедура объединения их в дашборды.

Сопроводительные материалы и исходные файлы

Все исходные коды и файлы, которые будут использованы в книге, можно загрузить на странице книги на сайте издательства «ДМК Пресс».

Глава 1

Что такое Yandex DataLens?

Что такое *Yandex DataLens*? Если просто, это сервис для бизнес-аналитики, с помощью которого можно легко и просто подключаться к различным источникам, строить на их основе модели данных, создавать красивые чарты-диаграммы, наполнять ими богатые по функционалу интерактивные дашборды и делиться результатами с коллегами и друзьями.

Конечно, сегодня BI-инструментов существует великое множество, и у каждого есть свои достоинства и недостатки. У каких-то из них более богатый список коннекторов для внешних источников, другие отличаются широким спектром инструментов для моделирования данных, третьи – разнообразием предлагаемых из коробки или в качестве плагинов диаграмм и графиков, четвертые – возможностью выгружать аналитические выводы в разных форматах.

Ни одну систему нельзя назвать безусловным лидером на рынке, да и для разных задач могут подходить разные инструменты. У кого-то весь процесс обработки сырых данных производится в источнике, а в BI-инструмент поступают уже очищенные данные, готовые к визуализации. Другие возлагают на свой инструмент гораздо большую ответственность в отношении преобразования и моделирования данных. Кому-то нужна потоковая обработка в режиме реального времени, а кому-то достаточно ежедневного или даже еженедельного обновления исходных данных. К визуализации также могут предъявляться совершенно разные требования. Одним нужна высокая степень интерактивности, чтобы пользователи могли в реальном времени выбирать индивидуальные фильтры для отдельных графиков и пользоваться преимуществами кросс-фильтрации между чартами, другие вполне удовлетворяются довольно статичными диаграммами и таблицами, которые можно распечатывать и анализировать на бумаге.

Как видите, единых универсальных требований к BI-инструментам просто не существует, вследствие чего такой разрозненный спрос породил весьма вариативное предложение. Если текстовые редакторы, табличные процессоры и графические редакторы, как правило, похожи как родные братья, то

сервисы для бизнес-аналитики от разных разработчиков могут даже отдаленно не напоминать друг друга.

Также разные BI-инструменты отличаются по порогу входа. И если тот же Power BI от Microsoft требует знания определенных отдельных вспомогательных инструментов и языков запросов, таких как Power Query, M и DAX, без которых ничего серьезного не построишь, то *DataLens* обладает одним из самых низких порогов входа. Фактически никаких сторонних знаний для его освоения вам не понадобится. Достаточно понимать, что вы хотите получить на выходе, обладать критическим мышлением и разбираться в общих принципах визуализации данных. Кстати, многие из этих принципов прекрасно описаны в классическом бестселлере «The Big Book of Dashboards», который я перевел и который можно приобрести на сайте «ДМК Пресс» по адресу **https://dmkpress.com/catalog/computer/data/978-5-93700-387-4** (скидки есть на моем канале в Telegram).

Лично мне время от времени требуется визуализировать какие-то ad hoc данные – будь то для учета переведенных знаков и книг или в процессе подготовки к новому переводу, – и если раньше я открывал Excel и лепил какие-то графики из того, что было, то теперь мне бывает удобнее подключиться куда надо из *DataLens* и быстро набросать чарты на дашборд, объединив их селекторами. С учетом того, что в Excel каждый график приходится допиливать напильником, чтобы увидеть то, что хочешь, *DataLens* позволяет сэкономить драгоценное время.

Но, конечно же, это лишь один из рецептов приготовления этой BI-системы – так называемый *DataLens* по-домашнему. Вы считаете кулинарную тему неуместной? Тогда вспомните, что слово Lens произошло от латинского lēns, что означает чечевица, и все встанет на свои места.

А как используют *DataLens* во взрослых компаниях? Варианта на самом деле два: облачный и… безоблачный. Давайте поговорим о каждом из них в отдельности.

Способы развертывания DataLens

Облако

Изначально *DataLens* вышел на рынок как исключительно *облачная система*, и это неудивительно с учетом того, насколько развитую экосистему в облаке развернул Yandex. Именно этим он привлек как небольшие стартапы, так и очень крупные компании, которые могли задействовать в своей работе самые разные сервисы признанного технологического гиганта – от хранения информации в управляемых базах данных до использования инструментов внедрения машинного и глубокого обучения, и все это не выходя из общего облачного контура.

Если вспомнить, еще несколько лет назад облачные инфраструктуры пугали владельцев бизнеса своей неустойчивостью и ненадежностью, но сегодня эти опасения по большей части ушли в прошлое. Облако – это как саксофон,

на котором очень легко научиться играть очень плохо, – его тоже очень просто сделать невыносимо душным, и достучаться до небес будет проще, чем до такого недооблака. Однако с опытом Yandex в выстраивании облачных систем и принятыми ими самими стандартами качества и гарантиями отказоустойчивости своих сервисов об этом беспокоиться уже не приходится.

Плюсов же от использования облачных решений гораздо больше, и вот лишь некоторые из них:

- *разделение ответственности*: в случае с использованием облачных ресурсов в процессе работы вам практически не приходится заботиться о технической стороне вопроса, включая работоспособность и отказоустойчивость серверов, количество свободного места в памяти и на дисках и пр. Все это входит в зону ответственности облачного провайдера, а вам остается лишь сосредоточиться на работе;
- *масштабируемость проекта*: под масштабируемостью мы подразумеваем гибкость, предоставляемую облачным провайдером в плане задействованных ресурсов. К примеру, при использовании облака Yandex вы можете с учетом сезонных требований снижать и повышать объем используемых ресурсов, что позволит сэкономить немало средств. При использовании собственного оборудования любое горизонтальное и вертикальное масштабирование проекта потребует от вас гораздо бóльших вложений;
- *безопасность данных*: ответственность за хранение и конфиденциальность ваших данных целиком и полностью возлагается на облачного провайдера, обычно обладающего гораздо бóльшими ресурсами, чем сами компании, для обеспечения безопасности;
- *скорость запуска продуктов*: при использовании облачных ресурсов время, требующееся на полный цикл разработки и внедрение нового цифрового продукта (так называемый показатель *Time to Market*), значительно снижается, что объясняется соблюдением стандартов и снятием с вас нагрузки по обеспечению технической поддержки решения;
- *экспертные инструменты и поддержка*: специалисты вашей растущей компании могут не обладать всем богатством навыков, необходимых для самостоятельного развертывания полноценной информационной системы, включающей множество сервисов. Облачные провайдеры предлагают вам необходимую техническую поддержку на протяжении всего жизненного цикла решения.

Но в *Yandex DataLens* решили пойти дальше и в конце 2023 года приподняли заветный засов, заявив во всеуслышание: «Входите, открыто!» Это событие ознаменовало начало новой эпохи под названием…

Open source

Итак, в конце сентября 2023 года было объявлено о публикации на GitHub под открытой лицензией Apache 2.0 исходного кода *Yandex DataLens*. С тех пор *опенсорс-версию* продукта можно использовать в абсолютно любой инфраструктуре. Пользователи, будь то небольшие стартапы или крупные компании, полу-

чили возможность разворачивать *DataLens* на своем аппаратном обеспечении, не опасаясь привязки к облачному вендору, а также выстраивать собственные экосистемы на основе нескольких продуктов с открытым исходным кодом.

Поскольку на момент первого издания книги *DataLens* еще не вышел в «открытое плавание», я остановлюсь на этой теме чуть более подробно. Кроме того, это действительно очень важная веха для продукта, которая делит историю на до и после как в плане фактического привлечения сообщества к его разработке, так и в отношении общего развития. Если вы романтик и не интересуетесь ничем, кроме облаков, можете просто пропустить следующие несколько абзацев.

Развертывание *DataLens* в локальном окружении выполняется посредством запуска нескольких контейнеров через Docker Compose. Если у вас не установлен Docker, вы можете легко поставить его на MacOS (**https://docs.docker.com/desktop/setup/install/mac-install**), Linux (**https://docs.docker.com/engine/install**) или Windows (**https://docs.docker.com/desktop/setup/install/windows-install**).

Для установки *DataLens* вам достаточно будет склонировать репозиторий и перейти в него с помощью команды, показанной ниже:

```
git clone https://github.com/datalens-tech/datalens && cd datalens
```

Далее просто запустите контейнеры, необходимые для старта *DataLens*, следующей командой:

```
HC=1 docker compose up
```

Переменная `HC` здесь относится к библиотеке Highcharts, которая позволит вам использовать большее количество чартов на дашбордах. При отключенной переменной `HC` будет использоваться исключительно библиотека D3.js, на которую *DataLens* в будущем планирует перейти полностью.

После запуска необходимых контейнеров вы сможете открыть интерфейс *DataLens* на локальной машине по адресу `http://localhost:8080` под пользователем и паролем `admin`. Для использования другого порта, в случае если порт 8080 у вас занят, вы можете воспользоваться переменной окружения `UI_PORT`, как показано ниже:

```
UI_PORT=8081 docker compose up
```

Для развертывания *DataLens* по-взрослому рекомендуется сгенерировать файл конфигурации с предварительно заполненными кредами, как показано ниже:

```
./init.sh --hc
docker compose -f ./docker-compose.production.yaml up -d
./init.sh --hc --up
```

Конец ознакомительного фрагмента.
Приобрести книгу можно
в интернет-магазине
«Электронный универс»
e-Univers.ru