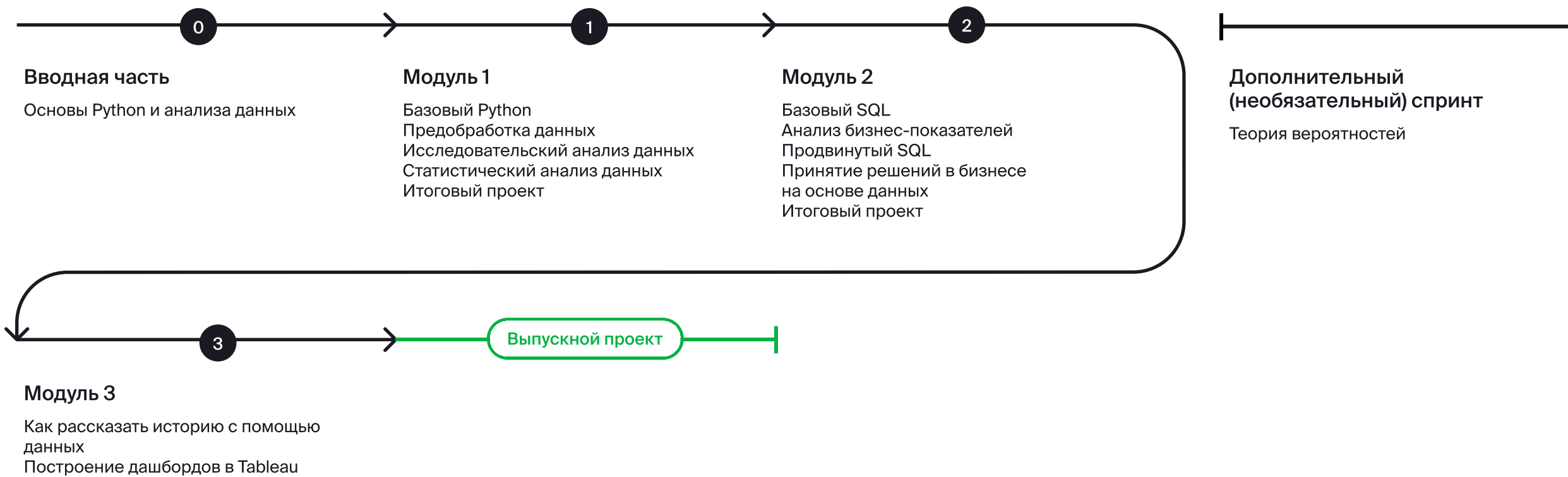


Аналитик данных



0 Вводный курс: основы Python и анализа данных

Узнаете основные концепции анализа данных и поймёте, чем занимаются аналитики данных и специалисты по Data Science. Решите пять кейсов по работе с данными из разных областей: **2 недели, от 20 часов**

- выясните причину массовой поломки гаджетов;
- проверите окупаемость рекламы мобильного приложения;
- найдёте лучшее место для нового магазина;
- поможете выбрать стратегию развития ИИ-стартапа;
- оцените эффективность роботов в службе поддержки.

Решая кейсы, вы изучите азы Python и библиотеки pandas, научитесь строить некоторые графики и верно их трактовать.

Темы

- Moscow Catnamycs
- Ошибки, переменные и гипотезы
- Что делают специалисты в области данных
- Списки в Python
- Цикл for
- Условия и булева логика
- Машинное обучение
- Финальный проект

1 Базовый Python

Глубже погрузитесь в язык программирования Python и работу с библиотекой pandas.

Темы

- Переменные и типы данных. Вывод данных и арифметические операции
- Строки
- Списки
- Цикл for
- Вложенные списки
- Условный оператор. Цикл while
- Функции
- Словари
- Pandas для анализа данных
- Предобработка данных
- Анализ данных и оформление результатов
- Jupyter Notebook — тетрадь в ячейку
- Проект: музыка больших городов

2 недели, от 30 часов

Проект

Сравните данные пользователей Яндекс.Музыки по городам и дням недели.

2 Предобработка данных

Научитесь очищать данные от выбросов, пропусков и дубликатов, а также преобразовывать разные форматы данных.

Темы

- Введение в предобработку данных
- Работа с пропусками
- Изменение типов данных
- Поиск дубликатов
- Категоризация данных
- Системное и критическое мышление в работе аналитика

2 недели, от 35 часов

Проект

Проанализируете данные о клиентах банка и определите долю кредитоспособных.

3 Исследовательский анализ данных

Изучите основы теории вероятностей и статистики. Примените их для исследования основных свойств данных, поиска закономерностей, распределений и аномалий. Познакомитесь с библиотеками SciPy и Matplotlib. Отрисуете диаграммы, поупражняетесь в анализе графиков.

Темы

- Введение в исследовательский анализ данных
- Первые графики и выводы
- Изучение срезов данных
- Работа с несколькими источниками данных
- Взаимосвязь данных
- Валидация результатов

2 недели, от 30 часов

Проект

Исследуете архив объявлений о продаже объектов недвижимости в Санкт-Петербурге и Ленинградской области.

4 Статистический анализ данных

Научитесь анализировать взаимосвязи в данных методами статистики. Узнаете, что такое статистическая значимость, гипотезы и доверительные интервалы.

2 недели, от 20 часов

Темы

- Введение в статистический анализ данных
- Описательная статистика
- Теория вероятностей
- Проверка гипотез

Проект

Оптимизируете воронки продаж для ускорения работы отдела маркетинга.

Теория вероятностей. Дополнительный курс

Вспомните или узнаете базовые термины в теории вероятностей: независимые, противоположные, несовместные события и т. д. На простых примерах и забавных задачах потренируетесь работать с числами и выстраивать логику решения.

Это необязательный спринт. Значит, каждый студент сам выбирает один из вариантов прохождения:

- Освоить дополнительный курс из десяти коротких уроков, освежить в памяти теорию и решить задачи.
- Открыть только блок с задачами для собеседований, вспомнить практику без теории.
- Пропустить курс совсем или вернуться к нему, когда будет время и необходимость.

5 Итоговый проект первого модуля

Научитесь предварительному исследованию данных, сформулируете и проверите гипотезы.

1 неделя, от 15 часов

Проект

Найдёте закономерности в данных о продаже игры.

Каникулы: 1 неделя

6 Базовый SQL

Изучите основы структурированного языка запросов SQL и реляционной алгебры для работы с базами данных. Познакомитесь с особенностями работы в PostgreSQL — популярной системе управления базами данных (СУБД). Научитесь писать запросы разного уровня сложности и переводить бизнес-задачи на язык SQL.

2 недели, от 25 часов

Вы будете работать с базой данных онлайн-магазина, который специализируется на фильмах и музыке.

Темы

- Введение в базы данных
- Срезы данных в SQL
- Агрегирующие функции. Группировка и сортировка данных
- Взаимоотношения между таблицами. Типы объединений таблиц
- Подзапросы и временные таблицы

Проект

Напишете ряд запросов разной сложности к базе данных, в которой хранятся данные по венчурным инвесторам, стартапам и инвестициям в них.

7 Анализ бизнес-показателей

Узнаете что такое метрики в бизнесе. Научитесь использовать инструменты для анализа данных в бизнесе: когортный анализ, воронка продаж и юнит-экономика.

Темы

- Введение в анализ бизнес-показателей
- Метрики и воронки
- Когортный анализ
- Юнит-экономика
- Пользовательские метрики

2 недели, от 35 часов

Проект

На основе данных изучите поведение пользователей, а также проанализируете доходность клиентов и окупаемость рекламы, чтобы предложить рекомендации для отдела маркетинга.

8 Продвинутый SQL

Пройдёте дополнительный курс по работе с базами данных и станете ещё ближе к бизнесу. С помощью языка SQL разберёте подсчёт основных бизнес-метрик, с которыми вы познакомились в курсе «Анализ бизнес-показателей». Рассмотрите работу с таким сложным инструментом, как оконные функции. Научитесь изменять содержимое баз данных локально, без тренажёра, используя специальные программы-клиенты и библиотеки для Python.

2 недели, от 35 часов

Проект

С помощью Python и SQL подключитесь к базе данных, посчитаете и визуализируете ключевые метрики сервис-системы вопросов и ответов о программировании.

9 Принятие решений в бизнесе на основе данных

Узнаете что такое A/B-тестирование и поймете в каких случаях его используют. Научитесь проектировать A/B-тестирование, производить оценку его результатов.

2 недели, от 20 часов

Темы

- Введение в принятие решений в бизнесе на основе данных
- Основы проверки гипотез в бизнесе
- Выбор метода проведения эксперимента
- Приоритизация гипотез
- Подготовка к проведению A/B-теста
- Анализ результатов A/B-теста
- Поведенческие алгоритмы: умение внятно объяснить свою позицию

Проект

Проанализируете результаты A/B-тестирования в крупном интернет-магазине.

10 Итоговый проект второго модуля

Научитесь проверять статистические гипотезы в рамках A/B-тестирования и готовить выводы и рекомендации в формате аналитического отчёта

1 неделя, от 15 часов

Проект

Исследуете воронку продаж и проанализируете результаты A/B-тестирования в мобильном приложении.

11 Как рассказать историю с помощью данных

Узнаете как правильно презентовать результаты своего исследования, оперируя графиками, важнейшими цифрами и их правильной интерпретацией. Познакомитесь с библиотеками Seaborn и Plotly.

2 недели, от 15 часов

Темы

- Как рассказать историю с помощью данных
- Подготовка презентации
- Библиотека seaborn
- Библиотека plotly

Проект

Подготовите исследование рынка на основе открытых данных о заведениях общественного питания Москвы.

12 Построение дашбордов в Tableau

Узнаете что такое автоматизация процессов анализа данных. Научитесь превращать рутинные и постоянные задачи в скрипты. Создадите дашборды для разных аудиторий и нужд компании.

2 недели, от 30 часов

Темы

- Введение в автоматизацию
- Основы запуска скриптов
- Что такое дата-пайплайны и зачем они нужны
- Проектирование и разработка дашбордов в dash
- Tableau

Проект

Разработаете дашборд для анализа пользовательского взаимодействия в сервисе Яндекс.Дзен.

Основы машинного обучения. Дополнительный курс.

Познакомитесь с основами машинного обучения и узнаете об основных задачах машинного обучения в бизнесе.

2 недели, от 20 часов

Темы

- Введение в прогнозы и предсказания
- Задачи машинного обучения в бизнесе
- Алгоритмы машинного обучения
- Процесс решения задач машинного обучения

Проект

Разработаете модель машинного обучения для прогнозирования оттока клиентов в фитнес-центре.

Каникулы: 1 неделя

13 Выпускной проект

В последнем проекте подтвердите, что освоили новую профессию. Уточните задачу заказчика, пройдёте все стадии анализа данных. Теперь без уроков и домашних заданий — всё как на реальной работе.

2 недели, от 20 часов