

Python для анализа данных

Продолжительность курса — 3 месяца

01

Основы Python

4 недели (40 часов)
3 спринта
1 контрольная работа

02

Основы
анализа данных

4 недели (40 часов)
3 спринта
1 контрольная работа

03

Продвинутый
анализ данных

4 недели (40 часов)
3 спринта
1 контрольная работа

01

4 недели

Основы Python

Освоите основы синтаксиса python, изучите типы данных, как простые, так и сложные. Изучите основные алгоритмические инструменты, такие как циклы и условия. Научитесь создавать свои собственные подпрограммы – пользовательские функции, освоите их применение.

Научитесь импортировать библиотеки для анализа, загружать файлы данных различных форматов.

Спринт

1. Управляющие конструкции Python. Циклы и условия
2. Продвинутые концепции. Функции и модули
3. Работа с данными в Python, основы Pandas

Контрольная работа

Набор заданий на комбинирование функций, циклов и условий.

02

4 недели

Основы анализа данных

Научитесь преобразовывать данные, находить и заменять пропуски, определять аномалии. Освоите исследовательский анализ данных, научитесь выявлять выбросы, читать гистограммы. Разберетесь в основах статистики, научитесь сравнивать данные статистическими методами.

Спринт

4. Импорт и обработка данных. Трансформация
5. Визуализация данных
6. Исследовательский анализ и основы статистики

Контрольная работа

Исследовательский и статистический анализ набора данных

Продвинутый анализ данных

Научитесь проектировать и проводить АВ-тестирование. Изучите Основы машинного обучения. Узнаете основные модели и алгоритмы машинного обучения для работы с данными.

Спринт

7. АВ-тестирование
8. Регрессии и временные ряды. Машинное обучение
9. Работа с реальными данными

Финальный проект

Реальный аналитический кейс с проверкой ревьюера.

Суть проекта:

Анализ данных интернет-магазина

В этом итоговом проекте вам предстоит проанализировать реальные данные по продажам и покупателям интернет-магазина.

Вы поработаете с данными о товарах, заказах, отзывах, покупателях.

Вам нужно будет подготовить данные, провести анализ и выявить важные закономерности, которые помогут повысить продажи. На основе анализа необходимо разработать рекомендации для магазина и представить их в виде презентации с визуализацией результатов.

Технологии и инструменты

Jupyter Notebook - интерактивная тетрадь для работы с кодом

Pandas - главная библиотека в Python для работы с данными

NumPy - библиотека в Python с открытым исходным кодом

A/B-тест - эксперимент, который позволяет сравнить 2 версии чего-либо, чтобы проверить гипотезы и определить, какая версия лучше

Scikit-Learn - библиотека, которая предоставляет инструменты для машинного обучения

Вебинары

В каждом модуле у вас будут групповые встречи с наставником в формате вебинаров. На них вы сможете разобрать отдельные темы, задать вопросы и обсудить рабочие кейсы.