

Модуль 1. Вводный курс. Аналитика данных.

- Современная экосистема данных
- Позиция аналитика данных, кто такой, чем занимается.
- Какие виды аналитиков есть
- Классические задачи аналитика. Роли в команде.
- Сферы применения аналитика данных

Модуль 2. Python

- Python для анализа данных в Jupyter Notebook
- Основы работы в Python и знакомство с интерфейсом Jupyter Notebook
- Основы работы с библиотекой **pandas**
- Продвинутый pandas и знакомство с визуализацией на python
- Основы работы с библиотекой **numpy**
- Работа с грязными данными, чистка данных
- EDA – исследование данных и подготовка их к анализу

Модуль 3. Статистика

- Основы статистики
- Разумный сбор данных и графические методы представления
- Математические инструменты, применяемые в статистике
- Численные методы описания данных
- Вероятность
- Случайные переменные и распределение вероятностей
- Оценка с использованием одной выборки
- Выборка и генеральная совокупность.
- Проверка гипотез.
- Статистический вывод: ошибка первого и второго рода.
- Сравнение средних значений.
- Дисперсионный анализ.
- Корреляционный и регрессионный анализ.
- Множественный регрессионный анализ.
- A/B тестирование, A/A тестирование, bootstrap.
- Изучение рядов динамики, методы прогнозирования динамических рядов

Модуль 4. Продуктовая аналитика

- Методологии определения потребностей пользователя (Business Model Canvas)
- Определение unit-экономики
- Метрики продукта, построение дерева-метрик
- Сегментация клиентов и целевой аудитории
- ABC-XYZ анализ
- RFM-анализ
- Когортный анализ

Модуль 5. GIT и управление цифровым проектом

- Что такое контроль версий и для чего он нужен
- Основные команды git
- Создание веток, объединение веток, работа в команде
- Что такое управление проектами.
- Ключевые фазы проекта.
- Методологии управления проектами.
- ТЗ на внедрение, ключевые показатели эффективности (KPI)
- Порядок тестирования, типы решений, процесс принятия решения

Модуль 6. Теории Баз данных, хранилища данных

- Понятие данных и баз данных
- Различные типы данных, компоненты БД
- Модели данных
- Ключи таблиц
- Нормализация \ Денормализация данных

Модуль 7. Основы SQL

- Что такое SQL
- Основные команды SQL
- Функции агрегации
- Объединение таблиц
- Подзапросы, представления, создание таблиц
- Оконные функции

Модуль 8. ETL/ELT

- Что такое ETL/ELT процессы
- Обзор инструментов (Airflow и др.).
- Python+SQL

Модуль 9. Визуализация данных

- Зачем нужна визуализация данных, что такое дашборды.
- Виды визуализаций.
- Визуализация в Python (знакомство с библиотеками matplotlib, seaborn, pygwalker)
- Визуализация данных в **Tableau**: знакомство с интерфейсом, принципы работы, построение визуализаций, построение дашбордов)
- Визуализация данных в **Power BI** (знакомство с интерфейсом, принципы работы, построение визуализаций, построение дашбордов)

Модуль 10. MS Excel и Google Sheets.

- Сбор и преобразование данных в Excel.
- Работа с функциями и формулами

- Визуализация в Excel
- Пакет Анализ данных
- Сбор и преобразование данных в Google Sheets
- Работа с функциями и формулами
- Визуализация в Google Sheets
- Google API

Модуль 11. Google-аналитика и Яндекс-метрика.

- Что такое Google-аналитика и Яндекс-метрика и для чего они нужны.
- Работа с тестовыми аккаунтами Google-аналитика и Яндекс-метрика

Модуль 12. Решение тестовых заданий и требования к написанию итогового проекта.

Итоговый проект