

1 этап
PYTHON 3

1. Операции
2. Типы данных
3. Циклы
4. Работы с данными (Pandas, EXCEL, CSV)
5. Работа с вэб (API, парсинг)
6. Работа с сервером (после LINUX)
7. Методы
8. ООП (лайт уровень)

2 этап
LINUX

1. Структура Linux
2. Linux-command
3. SSH
4. Права доступа
5. Bash (опционально)

3 этап
SQL (Mysql)

1. Поднятие сервера
2. Создание архитектура через диаграммы
3. DDL и DML

4 этап
HADOOP

1. Архитектура кластера
2. Инфраструктура (HIVE, HUE, AIRFLOW, HDFS)
3. Типы данных
4. Разворачивание кластера и знакомство с ним
5. Beeline (не наш), HIVE
6. Pyspark
7. ETL (сырые данные -> Hadoop)
8. ETL (Mysql -> Hadoop)
9. ETL (Hadoop <- сырые данные)
10. ETL (Hadoop <- Mysql)

ПРОЕКТ