

АНАЛИТИКА ДАННЫХ

В центре внимания программы находится одна из основных задач цифровой экономики – сбор, хранение, анализ и управление технологическими данными, принятие бизнес-решений и разработка стратегии на основе данных.

Продолжительность обучения: 2 семестра, 252 часа

Присваиваемая квалификация: Концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности

Состав программы:

1. Основы аналитики

(сбор данных, основы статистики, основы работы с табличными редакторами, визуализация данных)

2. SQL и получение данных

(базы данных, язык запросов SQL, сложные запросы SQL, оптимизация запросов SQL, работа с PostgreSQL)

3. Метрики данных

(продуктовые метрики, когортный анализ)

4. Python для анализа данных

(основы Git, основы Python, библиотека Numpy, библиотека Pandas, основы парсинга)

5. Статистика в Python

(визуализация с python, методы визуализации данных, практика по matplotlib, seaborn, дескриптивная статистика и разведочный анализ данных в Python. Корреляция. Практикум по SciPy, A/B тестирование, работа с временными рядами в Python)

6. Машинное обучение

(Введение в машинное обучение, обучение с учителем/без учителя, библиотека Scikit-learn, NoSQL-подход к работе с большими данными, кластеризация, построение модели и метрики, Hadoop&Spark, PySpark)