

УЧЕБНО-ТЕМАТИЧЕСКИЙ ПЛАН
 программы повышения квалификации
«Инженер данных: python, sql, ml-модели»
 (Data Engineer)

№ п/п	Наименование дисциплины, модуля	Трудоемкость	В том числе				Форма контроля
			В часах	Контактная работа		Самостоятельная работа	
		Всего		Лекции	Практич. занятия		
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>
Входное тестирование		2	2		2		Зачет
М 1.	Инструменты и методы бизнес-аналитики	30	14	4	10	16	Зачет
1.1.	Примеры аналитических задач. Сбор данных для анализа данных	4	2	2		2	Тестирование
1.2.	Инструмент работы с таблицами в режиме онлайн на примере Google-таблицы.	6	2		2	4	Решение ситуационных задач
1.3.	Статистика и тестирование гипотез	4	2	2		2	Тестирование
1.4.	Бизнес-показатели и точки роста. Анализ продуктовых метрик	6	2		2	4	Решение практических задач
1.5.	Финансовые метрики и Маркетинговые метрики.	4	2		2	2	Решение практических задач
1.6.	Иерархия метрик и разработка отчетности. Дизайн тестов, проведение и анализ. Построение простых моделей	4	2		2	2	Решение практических задач
Промежуточная аттестация		2	2		2		Зачет
М 2.	Python и анализ данных	42	22	8	14	20	Зачет
2.1.	Python: Jupiter Notebook, комментарии в коде, арифметические операции. Управляющие конструкции и коллекции	4	2	2		2	Тестирование
2.2.	Функции и работа с данными, основы парсинга. Работа с файловой системой и модули	4	2	2		2	Тестирование
2.3.	Регулярные выражения и основы синтаксического разбора. Исключения и обработка ошибок	8	4	2	2	4	Решение практических задач

2.4.	Функции и понятие класса. Библиотека NumPy: вычислительные задачи	8	4		4	4	Решение практических задач
2.5.	Библиотека Pandas: возможности и оптимизация. Методы оптимизации Pandas	8	4	2	2	4	Решение практических задач
2.6.	Корреляция и корреляционный анализ. Статистическая проверка гипотез, А/В-тесты	8	4		4	4	Решение практических задач
Промежуточная аттестация		2	2		2		Зачет
М 3.	SQL и получение данных	38	20	8	12	18	Зачет
3.1.	Введение в SQL. Установка и знакомство. Основы и работа с базами данных	4	2	2		2	Тестирование
3.2.	Основы SQL. Базовые запросы SQL. Фильтрация данных	6	2	2		4	Тестирование
3.3.	Углубление в SQL. Агрегация данных. Группировка данных. Join. Оконные функции	8	4		4	4	Решение практических задач
3.4.	Работа с PostgreSQL	4	2		2	2	Решение практических задач
3.5.	"NoSql & MongoDB"	6	2	2		4	Тестирование
3.6.	Примеры создания дашбордов и анализ продуктовых метрик с применением SQL	8	6	2	4	2	Решение практических задач
Промежуточная аттестация		2	2		2		Зачет
М 4.	Визуализация данных	48	24	8	16	24	Зачет
4.1.	Анализ данных средствами Microsoft Excel	8	4	2	2	4	Решение практических задач
4.2.	Визуализация данных: основные BI-инструменты	4	2	2		2	Тестирование
4.3.	DataLens: быстрый анализ данных с визуализацией	4	2	2		2	Решение практических задач
4.4.	Power BI. Построение моделей данных из разных неструктурированных источников: таблиц, сайтов и баз данных	8	4		4	4	Решение практических задач
4.5.	Tableau. Основы работы с расчётными полями, фильтрами, множествами и группировками. Основные виды визуализаций. Лучшие практики визуализации	12	6	2	4	6	Решение практических задач
4.6.	Создание интерактивных дашбордов	10	4		4	6	Решение практических задач
Промежуточная аттестация		2	2		2		Зачет
М 5.	Основные методы машинного обучения	38	20	6	14	18	Зачет
5.1.	Понятие машинного обучения. Регрессионный анализ. Линейная, полиномиальная и логарифмическая регрессия	8	4	2	2	4	Решение практических задач
5.2.	Классификация. Логистическая регрессия и SVM	6	4		4	2	Решение практических задач

5.3.	Функции потерь и оптимизация Оценка точности модели, переобучение, регуляризация	6	2		2	4	Решение практических задач
5.4.	Проблема качества данных. Работа с пропусками Работа с переменными	4	2		2	2	Решение практических задач
5.5.	Деревья решений Ансамблирование Поиск выбросов и генерация новых признаков	6	2	2		4	Тестирование
5.6.	Алгоритмы кластеризации Улучшение качества модели.	6	4	2	2	2	Решение практических задач
Промежуточная аттестация		2	2		2		Зачет
М 6.	Глубокое машинное обучение и обработка больших данных	46	20	4	16	26	Зачет
6.1.	Машинное обучение и Data Science Мотивация и технологии работы с большими данными. Основные инструменты обработки больших данных. HADOOP & SPARK	10	6	2	4	4	Решение практических задач
6.2.	Рекомендательные системы	6	2		2	4	Решение практических задач
6.3.	Методы анализа временных рядов, модели ARIMA, GARCH	8	4	2	2	4	Решение практических задач
6.4.	Обработка естественного языка	6	2		2	4	Решение практических задач
6.5.	Компьютерное зрение, сегментация и детекция объектов	6	2		2	4	Решение практических задач
6.6.	Машинное обучение, подготовка модели по практическому кейсу	8	2		2	6	Решение практических задач
Промежуточная аттестация		2	2		2		Зачет
Итоговая аттестация		6	6		6		Защита итоговой работы
ИТОГО		250	128	38	90	122	