

ОАНО ДПО «СКАЕНГ»

Утверждаю
Генеральный директор

Айнетдинов Р.Х

«29» января 2021 г.

Дополнительная общеобразовательная (общеразвивающая) программа
«Аналитик данных»
для взрослых учащихся

Составитель:
Лисица Елена Сергеевна,
Контент-менеджер проекта «Математика»
ОАНО ДПО «СКАЕНГ»

Москва

2021 г.

Пояснительная записка

Дополнительная общеобразовательная (общеразвивающая) программа «Аналитик данных» предназначена для взрослых учащихся, желающих освоить работу и научиться решать аналитические задачи с помощью основных инструментов (SQL, Excel, Python, PowerBI) для получения данных, проведения аналитических исследований и решения бизнес-проблем.

В основе программы лежат следующие **федеральные нормативные акты**:

- Федеральный закон от 29 декабря 2012 г. №273-ФЗ «Об образовании в Российской Федерации»;
- Закон Российской Федерации от 07 февраля 1992 г. № 2300-1 «О защите прав потребителей»;
- Порядок организации и осуществления образовательной деятельности по дополнительным профессиональным программам. Приказ МОиН РФ от 01.07.2013 №499/99;
- Порядок организации и осуществления образовательной деятельности по дополнительным общеобразовательным программам. Приказ Министерства просвещения РФ от 9 ноября 2018 г. N 196
- Показатели, характеризующие общие критерии оценки качества условий осуществления образовательной деятельности организациями, осуществляющими образовательную деятельность по дополнительным профессиональным программам. Приказ Министерства науки и высшего образования РФ от 15 апреля 2019 г. N 31н

Актуальность программы определяется содержанием обучения, отвечающих потребностям обучающихся в профессиональном развитии. **Новизна программы** и ее отличительная особенность заключается в применении цифровых технологий, позволяющих построить индивидуальную образовательную траекторию обучающихся и создающих возможности более полного учета их индивидуальных потребностей, возможностей и интересов. **Педагогическая целесообразность программы** обусловлена удовлетворением персональных образовательных потребностей обучающихся.

Цель и задачи программы

Цель:

Целью реализации программы является формирование у слушателей профессиональных компетенций в области аналитики данных, включая решение аналитических задач с помощью основных инструментов (SQL,

Excel, Python, PowerBI) для получения данных, проведения аналитических исследований и решения бизнес-проблем.

Задачи программы:

1. Научить обучающихся решению аналитических задач и развить умения:
 - формализовать бизнес-проблему в аналитическую задачу
 - определить комплексную метрику, которая будет отражать результат
 - определить, какие нужны данные, собрать и преобразовать их
 - выбрать метод анализа данных и инструмент
 - провести анализ с помощью этого инструмента и представить результаты
2. Овладеть инструментами анализа данных: SQL, Excel, Python, PowerBI
3. Изучить теоретические основы статистики для аналитиков, практикуя освоенный ранее язык Python
4. Научиться применять освоенные знания и инструменты в комплексе для решения бизнес-задач

Категория слушателей К освоению дополнительной профессиональной программы повышения квалификации допускаются лица, имеющие или получающие высшее образование.

Форма обучения – заочная (дистанционная) на платформе Vimbox.

Программа рассчитана на **266 часов** и состоит из 6 модулей.

Характер занятий: все занятия носят практический характер.

Режим занятий формируется по календарному плану 1 раз в неделю.

Планируемые результаты:

Выпускник программы должен уметь решать аналитические задачи:

- формализовать бизнес-проблему в аналитическую задачу
- определить комплексную метрику, которая будет отражать результат
- определить, какие нужны данные, собрать и преобразовать их
- выбрать метод анализа данных и инструмент
- провести анализ с помощью этого инструмента и представить результаты

У обучающихся формируются следующие компетенции: способность работать с компьютером как средством управления информацией, работать с информацией из различных источников, в том числе в глобальных компьютерных сетях (ОПК-3); проведение анализа инноваций в экономике,

управлении и информационно-коммуникационных технологиях (ПК 4); способность использовать соответствующий математический аппарат и инструментальные средства для обработки, анализа и систематизации информации по теме исследования (ПК-18).

В результате освоения программы слушатель должен:

Иметь представление:

- о современных методах анализа данных;
- инструментах решения задач анализа данных;
- тенденциях технологий интеллектуального анализа данных, стандартах и инструментах.

Знать:

- основные проблемы, возникающие при анализе данных, и пути их решения;
- методы анализа данных
- процесс и этапы анализа данных
- виды и способы организации хранилищ данных;
- классификацию аналитических систем;
- состав классов программных продуктов, образующих набор Business Intelligence.

Уметь:

- использовать программные и аппаратные средства персонального компьютера для анализа данных;
- ориентироваться в современной системе источников информации;
- применять методы интеллектуального анализа данных;
- использовать современные информационные технологии в своей профессиональной деятельности;
- видеть и формулировать проблему; видеть конкретную ситуацию; прогнозировать и предвидеть;
- ставить цели и задачи.
- Приобрести навыки:
- обработки больших массивов информации (Big data); анализа данных различной природы;

- формализации бизнес-задач
- выбора инструмента для анализа данных
- выбора метода анализа данных
- реализации методов анализа данных
- презентации результатов анализа

Владеть, иметь опыт:

- применения современной терминологии в области систем поддержки принятия решений и методологии решения задач в области многомерного анализа данных;
- применения современных программных пакетов многомерного анализа.

Контроль в каждом модуле осуществляется в виде: текущего контроля (данные системы о результатах выполнения каждого задания), и итогового контроля – курсовой работы. Итоговым контролем в курсе является дипломная работа.

Материально-техническое обеспечение программы:

- 1) Персональный компьютер (стационарный компьютер/ноутбук).
- 2) Микрофон и наушник.
- 3) Установленная на персональный компьютер программа Skype.
- 4) Установленный на персональный компьютер браузер Google Chrome.
- 5) Необходимая скорость интернет-соединения - от 5 МБит/с, ping не выше 50.

Учебно-тематический план:

Модуль 1. SQL для анализа данных

Тема	Урок	Количество часов	
		Аудиторных	Самостоятельных
Введение в SQL и основы синтаксиса	Базовые запросы	2	3
Генерация новых признаков и очистка данных	Генерация новых признаков и очистка данных	2	3
Агрегатные функции сортировка, группировка и	Агрегатные функции	2	3
Операции объединения таблиц с данными	Объединение таблиц	2	3
Подзапросы и WITH	Подзапросы и WITH	2	3
Оконные функции	Оконные функции	2	3
Аналитическое исследование (workshop)	Практикум. Проведение аналитического исследования и решение бизнес-проблемы с помощью SQL	2	3
Создание, изменение, удаление данных в таблице, создание индексов	Создание, изменение, удаление данных в таблице, создание индексов	2	3

Временные таблицы	Временные таблицы	2	3
Основы оптимизации	Основы оптимизации	2	3
Курсовая работа	Проект по анализу метрик качества контента и результатов обучения студентов на датасете онлайн школы за 2021 год.		3
Итого			

Модуль 2. Анализ данных в Excel

Тема	Урок	Количество часов	
		Аудиторных	Самостоятельных
Функционал Excel	Функционал Excel	2	3
Процедуры и функции Excel	Оценка эффективности CPC рекламы в Excel: расчет метрик, применение базовых функций и формул	2	3
Процедуры и функции Excel	Оценка эффективности CPC рекламы в Excel: расчет метрик, применение продвинутых функций и формул	2	3
Процедуры и функции Excel	RFM анализ и когортный анализ в Excel: расчет метрик, применение базовых функций и формул	2	3
Визуализация в Excel	RFM анализ и когортный анализ в Excel: расчет метрик, применение продвинутых функций и формул,	2	3

	визуализация		
Прогнозирование	Анализ SEO-датасета: расчет метрик, прогнозирование	2	3
Семантический анализ	Семантический анализ отзывов: расчет метрик, применение продвинутых функций и формул	2	3
Курсовая работа	Аналитический проект		3
Итого			

Модуль 3. Python для анализа данных

Модуль	Тема	Количество часов	
		Аудиторных	Самостоятельных
Погружение в Python	Python Quick Start, переменные и типы данных	1	2
	Работа со списками и словарями	1	2
	Циклы, функции	1	2
Библиотеки Python	Библиотеки: стандартные, Pandas	1	2
Визуализация данных	Выбор и настройка визуализаций	1	2
Python для анализа стат.тестов	Статистика и АБ-тесты	1	2
	Воркшоп «Определение статистической значимости на малых	1	2

	выборках»		
Получение данных из разных источников	Подключение к API	1	2
	Запросы к удаленным базам данных	1	2
Прогнозирование	Запросы к удаленным базам данных	1	2
Дипломная работа	Одна из двух задач на выбор: 1. E-commerce: задача прогнозирования конверсии по историческим данным с учетом сезонности 2. ФинТех: задача составить рекомендации о развитии продукта, основываясь на данных удаленной БД	0	10
Итого		10 часов	30 часов
		40 часов	

Модуль 4. Визуализация в PowerBI

Тема	Урок	Количество часов	
		Аудиторных	Самостоятельных
Загрузка и преобразование данных	Загрузка данных из разных источников	2	3
Загрузка и преобразование данных	Построение связей в таблицах	2	3
Загрузка и преобразование	Объединение столбцов и создание	2	3

данных	ключа между таблицами		
Моделирование и анализ данных	RFM анализ и когортный анализ: расчет метрик, применение продвинутых функций	2	3
Визуализация данных и работа с отчетами	Визуализация данных	2	3
Визуализация данных и работа с отчетами	Построение динамического дашборда	2	3
Работа с DAX	Работа с DAX	2	3
Advanced DAX	Advanced DAX	2	3
Построение комплексных отчетов	Построение комплексных отчетов	2	3
Power BI Service и Power BI Mobile	Power BI Service и Power BI Mobile	2	3
Курсовая работа	Дашборд для отдела с ключевыми метриками		3
Итого			

Модуль 5. Статистика для аналитиков (на Python)

Тема	Урок	Количество часов	
		Аудиторных	Самостоятельных
Основы статистики	Основы статистики	2	3
Свойства	Свойства	2	3

статистических оценок	статистических оценок		
Доверительные интервалы	Доверительные интервалы	2	3
Параметрические критерии и проверка гипотез	Параметрические критерии и проверка гипотез	2	3
Непараметрические критерии, критерии согласия, бутстрап	Непараметрические критерии, критерии согласия, бутстрап	2	3
А/В тестирование	А/В тестирование	2	3
Курсовая работа	Анализ А/В теста		3
Итого			

Модуль 6. Аналитика в продукте и маркетинге

Тема	Урок	Количество часов	
		Аудиторных	Самостоятельных
Маркетинговая аналитика (воронка, конверсии, модели атрибуции)	Маркетинговая аналитика	2	3
Продуктовая аналитика (метрики продукта, конверсии, валидация продуктовых	Продуктовая аналитика	2	3

гипотез – А/В-тесты)			
Операционная аналитика	Планирование ресурсов	2	3
Дипломная работа			3
Итого			

Кадровое обеспечение образовательного процесса

Квалификация педагогических работников Организации соответствует квалификационным характеристикам, установленным в Едином квалификационном справочнике должностей руководителей, специалистов и служащих, разделе «Квалификационные характеристики должностей руководителей и специалистов высшего профессионального и дополнительного профессионального образования», утвержденного приказом Министерства здравоохранения и социального развития РФ от 11 января 2011 г. N 1н

Пример итоговой контрольной работы

Прогнозирование прибыли компании на основе ретроспективных данных и создание динамического отчета в PowerBI с результатами RFM-анализа.