

ОБЩЕСТВО С ОГРАНИЧЕННОЙ ОТВЕТСТВЕННОСТЬЮ «БИСТАДИ»
(ООО «БИСТАДИ»)

Утверждаю
Директор ООО «БИСТАДИ»



Б.А. Сафонов

» _____ 2026 г.

**Дополнительная общеразвивающая программа
«Data Science: методы анализа и обработки данных
на Python с использованием SQL»**

Форма обучения: с применением исключительно электронного обучения,
дистанционных образовательных технологий

Трудоёмкость: 42,35 академических часа

Срок обучения: 11 недель

г. Москва, 2026 г.

СОДЕРЖАНИЕ

РАЗДЕЛ 1. ОБЩАЯ ХАРАКТЕРИСТИКА ДОПОЛНИТЕЛЬНОЙ ОБЩЕРАЗВИВАЮЩЕЙ ПРОГРАММЫ

- 1.1. Нормативно-правовые основания разработки программы
- 1.2. Цель и планируемые результаты обучения
- 1.3. Категория слушателей
- 1.4. Форма обучения
- 1.5. Трудоёмкость программы. Структура программы
- 1.6. Особенности организации образовательной деятельности по программе
- 1.7. Содержание разделов

РАЗДЕЛ 2. СОДЕРЖАНИЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

- 2.1. Учебно-тематический план
- 2.2. Календарный учебный график
- 2.3. Рабочая программа

РАЗДЕЛ 3. КОНТРОЛЬ И ОЦЕНКА РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ПРОГРАММЕ

- 3.1. Формы аттестации
- 3.2. Контрольно-оценочные средства текущего и итогового контроля

РАЗДЕЛ 4. ОРГАНИЗАЦИОННО-ПЕДАГОГИЧЕСКИЕ УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ

- 4.1. Требования к материально-техническому обеспечению
- 4.2. Информационно-методическое обеспечение обучения
- 4.3. Кадровое обеспечение реализации программы

РАЗДЕЛ 1. ОБЩАЯ ХАРАКТЕРИСТИКА ДОПОЛНИТЕЛЬНОЙ ОБЩЕРАЗВИВАЮЩЕЙ ПРОГРАММЫ

1.1. Нормативно-правовые основания разработки программы

Разработка программы осуществлена в соответствии с действующими нормативными правовыми актами Российской Федерации, регулирующими реализацию дополнительных образовательных программ:

1. Федеральный закон от 29 декабря 2012 г. № 273-ФЗ «Об образовании в Российской Федерации» (в действующей редакции);
2. Постановление Правительства Российской Федерации от 18 сентября 2020 г. № 1490 «О лицензировании образовательной деятельности»;
3. Приказ Минпросвещения России от 27 июля 2022 г. № 629 «Об утверждении Порядка организации и осуществления образовательной деятельности по дополнительным общеобразовательным программам»;
4. Устав ООО «БИСТАДИ».

1.2. Цель и планируемые результаты обучения

Цель: сформировать у обучающихся комплекс базовых компетенций в области анализа данных на языке Python и работы с реляционными базами данных на языке SQL — от подготовки рабочей среды и получения данных из различных источников до извлечения, обработки, объединения, анализа и визуализации данных с использованием современных инструментов Data Science, необходимых для решения прикладных аналитических задач и дальнейшего профессионального развития.

Программа объединяет две самостоятельные, но методически связанные части, каждая из которых завершается собственной итоговой аттестацией: «Data Science»; «SQL».

Планируемые результаты обучения:

Будут знать:

- основные понятия и этапы работы с данными в Data Science;
- виды источников данных и способы их получения;
- принципы работы с табличными данными, массивами и матрицами;
- базовые подходы к визуализации и представлению данных;
- назначение инструментов Python, используемых в анализе данных;
- основные принципы работы с базами данных и роль SQL в анализе данных;
- базовый синтаксис SQL и структуру запросов, способы фильтрации, сортировки, агрегации и группировки данных;
- принципы объединения таблиц и работы с несколькими источниками данных;
- возможности интеграции SQL с Python для работы с данными и автоматизации.

Будут уметь:

- импортировать и создавать наборы данных из различных источников;
- обрабатывать и анализировать табличные данные с использованием Pandas;
- извлекать данные из веб-ресурсов с помощью web-parsing;
- выполнять численные и матричные вычисления с использованием NumPy;

- строить визуализации для анализа и интерпретации данных;
- формировать отчётные материалы в виде Excel-файлов и презентаций;
- настраивать окружение и инструменты для работы с базами данных;
- писать SQL-запросы с нуля и извлекать необходимые данные;
- объединять данные из нескольких таблиц, использовать подзапросы и сложные конструкции;
- интегрировать SQL с Python, загружать и обрабатывать данные, автоматизировать рутинные процессы.

Будут владеть:

- базовыми навыками программирования на Python для анализа данных;
- инструментами работы с данными (Pandas, NumPy, библиотеки визуализации);
- методами структурирования, анализа и визуального представления данных;
- практическими навыками подготовки и оформления результатов аналитической работы;
- практическими навыками написания SQL-запросов различного уровня сложности;
- подходами к использованию SQL в связке с Python для анализа данных и автоматизации;
- опытом создания и презентации завершённых итоговых проектов.

1.3. Категория слушателей

Программа предназначена для лиц, заинтересованных в освоении программирования на языке Python, работы с базами данных на языке SQL и основ анализа данных с нуля, включая студентов, начинающих специалистов, офисных сотрудников, аналитиков, маркетологов, тестировщиков, предпринимателей, преподавателей, а также слушателей без профильного образования, планирующих использовать программирование и анализ данных в учебной, исследовательской или профессиональной деятельности.

Программа ориентирована на слушателей без профильного технического образования и не требует предварительного опыта программирования. Обучение построено на практических примерах и задачах: от первых строк кода и работы с типами данных до обработки табличных данных, написания SQL-запросов и выполнения законченных итоговых проектов.

Требования к поступающим: для зачисления на обучение специальные знания и профессиональная подготовка не требуются. Достаточно базовой компьютерной грамотности, уверенного владения персональным компьютером, навыков работы с файлами и документами, а также мотивации к освоению языка Python, языка запросов SQL и анализа данных.

1.4. Форма обучения

Заочная, с применением исключительно электронного обучения и дистанционных образовательных технологий (ЭО и ДОТ). Освоение программы происходит полностью удалённо на образовательной платформе.

1.5. Трудоемкость программы. Структура программы

Общий объем программы: 42,35 академических часа.

Продолжительность: 11 недель.

Режим занятий: 2–3 раза в неделю по 2–3 академических часа.

Указанные сроки освоения учебного материала носят рекомендательный характер и рассчитаны на комфортное и последовательное прохождение программы. При освоении полного объема учебного материала и успешном прохождении предусмотренных форм контроля программа может быть завершена досрочно.

Программа состоит из следующих частей:

Часть	Наименование части программы	Объем, ак. ч.	Продолжительность	Итоговая аттестация
I	Data Science	25,05	5 недель	Сквозной проект «Создание мини web-сервиса анализа данных селлера»
II	SQL	17,3	6 недель	Итоговый проект «Аналитика маркетплейса»
Итого		42,35	11 недель	

1.6. Особенности организации образовательной деятельности по программе

Академический час составляет 45 минут. Объем учебной нагрузки в учебный день не превышает 8 академических часов. Образовательный процесс осуществляется с использованием электронной информационно-образовательной среды bstudy.ru.

Программа имеет модульную структуру. Части программы являются методически самостоятельными и могут осваиваться как последовательно (в порядке, приведенном в пункте 1.5), так и параллельно. Часть «SQL» опирается на навыки работы с данными средствами Python, поэтому её рекомендуется осваивать после части «Data Science».

Календарный учебный график (пункт 2.2) составлен исходя из последовательного освоения частей программы. При параллельном освоении частей суммарная учебная нагрузка в учебный день не должна превышать 8 академических часов.

1.7. Содержание разделов

№ п/п	Наименование раздела	Всего часов	Цель и краткое содержание	Формы контроля
ЧАСТЬ I. DATA SCIENCE (25,05 ак. ч.)				
1	Модуль 1. Введение. Руководство для прохождения обучения	–	Организационное введение: знакомство с образовательной платформой, структурой программы, порядком прохождения обучения, системой текущего контроля и итоговой аттестации.	–
2	Модуль 2. Генерация данных	3,15	Создание данных вручную и программно. Формирование списков, словарей, таблиц и простых наборов данных. Генерация тестовых и учебных данных, работа с типами данных и структурами, подготовка данных к анализу.	Тестирование, ДЗ
3	Модуль 3. Импорт данных	3,1	Импорт данных из файлов различных форматов (CSV, Excel, текстовые файлы). Загрузка данных из внешних источников и открытых ресурсов. Первичный анализ и проверка структуры загруженных данных.	Тестирование, ДЗ
4	Модуль 4. Работа с таблицами Pandas	3,0	Создание и структура DataFrame. Фильтрация, сортировка и группировка данных. Очистка и преобразование данных. Расчёт агрегированных показателей и подготовка данных к анализу и визуализации.	Тестирование, ДЗ
5	Модуль 5. Web-parsing	2,4	Основы работы с HTML-страницами. Извлечение данных с веб-сайтов. Преобразование полученной информации в таблицы и структуры данных, пригодные для дальнейшего анализа.	Тестирование, ДЗ
6	Модуль 6. Матрицы NumPy	2,5	Основы библиотеки NumPy. Создание массивов и матриц. Базовые математические операции, векторизация и матричные вычисления. Применение NumPy для численного анализа данных.	Тестирование, ДЗ
7	Модуль 7. Визуализация в Python	2,9	Построение базовых графиков и диаграмм. Выбор типа визуализации в зависимости от задачи. Интерпретация графиков и использование визуализации для объяснения аналитических выводов.	Тестирование, ДЗ
8	Модуль 8. Оформление итогов	2,0	Создание и форматирование Excel-файлов средствами Python. Формирование таблиц и отчётов. Автоматизированное создание презентаций с результатами анализа данных для учебных и прикладных задач.	Тестирование, ДЗ
9	Модуль 9. Итоговая аттестация	6,0	Выполнение сквозного практического проекта по созданию мини web-сервиса анализа данных селлера.	Итоговый проект
ЧАСТЬ II. SQL (17,3 ак. ч.)				
1	Модуль 1. Введение. Руководство для	–	Организационное введение: знакомство с образовательной платформой, структурой части программы, порядком прохождения обучения и системой аттестации.	–

№ п/п	Наименование раздела	Всего часов	Цель и краткое содержание	Формы контроля
	прохождения обучения			
2	Модуль 2. Настройка окружения	2,5	Настройка окружения и подготовка всего необходимого для работы с базами данных: установка инструментов, работа с интерфейсами СУБД, быстрый старт работы без технических сложностей.	Тестирование, ДЗ
3	Модуль 3. Базовые структуры запросов. Синтаксис	2,8	Базовые структуры SQL-запросов и синтаксис: написание запросов с нуля, извлечение данных из таблиц. Преобразование данных: фильтрация, сортировка, агрегация и группировка для превращения «сырых» данных в понятную информацию для анализа.	Тестирование, ДЗ
4	Модуль 4. Преобразование таблиц	2,9	Преобразование данных и работа с несколькими таблицами. Различные способы объединения данных: соединение таблиц, выборка нужных данных из разных источников, формирование целостной картины для анализа.	Тестирование, ДЗ
5	Модуль 5. Сложные запросы	2,6	Написание сложных запросов, включая подзапросы и объединения таблиц: решение реальных бизнес-задач, работа с несколькими источниками данных, построение более глубокой аналитики.	Тестирование, ДЗ
6	Модуль 6. Реализация SQL внутри Python	2,2	Использование SQL внутри Python для анализа данных и автоматизации: соединение SQL с Python, загрузка и обработка данных, автоматизация рутинных задач.	Тестирование, ДЗ
7	Модуль 7. Итоговый проект	4,3	Выполнение итогового проекта «Аналитика маркетплейса»: разбор учебной базы интернет-магазина, документирование модели данных, построение набора бизнес-отчётов на SQL, размещение проекта на GitHub.	Итоговый проект

РАЗДЕЛ 2. СОДЕРЖАНИЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

2.1. Учебно-тематический план

№ п/п	Название раздела, темы	Всего часов	Теоретических (видеолекция + материалы)	Практических (семинар, тест, ДЗ)	Формы аттестации
ЧАСТЬ I. DATA SCIENCE					
2	МОДУЛЬ 2. ГЕНЕРАЦИЯ ДАННЫХ	3,15	1,4	1,75	Тест, ДЗ
2.1	Общая информация о Data Science	0,2	0,1	0,1	Тест
2.2	Генерация данных. Теоретический материал	0,2	0,1	0,1	Тест
2.3	Генерация данных. Практический материал	2,75	1,2	1,55	ДЗ
3	МОДУЛЬ 3. ИМПОРТ ДАННЫХ	3,1	1,5	1,6	Тест, ДЗ
3.1	Импорт данных. Теоретический материал	0,2	0,1	0,1	Тест
3.2	Импорт данных. Практический материал	2,9	1,4	1,5	ДЗ
4	МОДУЛЬ 4. РАБОТА С ТАБЛИЦАМИ PANDAS	3,0	1,4	1,6	Тест, ДЗ
4.1	Работа с таблицами Pandas. Теоретический материал	0,2	0,1	0,1	Тест
4.2	Работа с таблицами Pandas. Практический материал	2,8	1,3	1,5	ДЗ
5	МОДУЛЬ 5. WEB-PARSING	2,4	0,8	1,6	Тест, ДЗ
5.1	Web-parsing. Теоретический материал	0,2	0,1	0,1	Тест
5.2	Web-parsing. Практический материал	2,2	0,7	1,5	ДЗ
6	МОДУЛЬ 6. МАТРИЦЫ NUMPY	2,5	0,9	1,6	Тест, ДЗ
6.1	Матрицы NumPy. Теоретический материал	0,2	0,1	0,1	Тест
6.2	Матрицы NumPy. Практический материал	2,3	0,8	1,5	ДЗ
7	МОДУЛЬ 7. ВИЗУАЛИЗАЦИЯ В PYTHON	2,9	1,3	1,6	Тест, ДЗ
7.1	Визуализация в Python. Теоретический материал	0,2	0,1	0,1	Тест
7.2	Визуализация в Python. Практический материал	2,7	1,2	1,5	ДЗ
8	МОДУЛЬ 8. ОФОРМЛЕНИЕ ИТОГОВ	2,0	0,5	1,5	Тест, ДЗ
8.1	Создание Excel-файлов и презентаций в Python	2,0	0,5	1,5	ДЗ

№ п/п	Название раздела, темы	Всего часов	Теоретических (видеолекция + материалы)	Практических (семинар, тест, ДЗ)	Формы аттестации
9	МОДУЛЬ 9. ИТОГОВАЯ АТТЕСТАЦИЯ	6,0	1,0	5,0	Итоговый проект
9.1	Инструктаж по проекту	1,0	1,0	–	–
9.2	Сквозной проект «Создание мини web-сервиса анализа данных селлера»	5,0	–	5,0	Защита проекта
Итого по части I		25,05	8,8	16,25	
ЧАСТЬ II. SQL					
2	МОДУЛЬ 2. НАСТРОЙКА ОКРУЖЕНИЯ	2,5	0,6	1,9	Тест, ДЗ
2.1	Введение. Подготовка и настройка	1,0	0,6	0,4	Тест
2.2	Лабораторная работа	1,5	–	1,5	ДЗ
3	МОДУЛЬ 3. БАЗОВЫЕ СТРУКТУРЫ ЗАПРОСОВ. СИНТАКСИС	2,8	0,9	1,9	Тест, ДЗ
3.1	Пишем запросы и получаем данные	1,3	0,9	0,4	Тест
3.2	Лабораторная работа	1,5	–	1,5	ДЗ
4	МОДУЛЬ 4. ПРЕОБРАЗОВАНИЕ ТАБЛИЦ	2,9	1,0	1,9	Тест, ДЗ
4.1	Преобразовываем данные и объединяем	1,4	1,0	0,4	Тест
4.2	Лабораторная работа	1,5	–	1,5	ДЗ
5	МОДУЛЬ 5. СЛОЖНЫЕ ЗАПРОСЫ	2,6	0,7	1,9	Тест, ДЗ
5.1	Запросы с подзапросами для глубокой аналитики	1,1	0,7	0,4	Тест
5.2	Лабораторная работа	1,5	–	1,5	ДЗ
6	МОДУЛЬ 6. РЕАЛИЗАЦИЯ SQL ВНУТРИ PYTHON	2,2	0,3	1,9	Тест, ДЗ
6.1	Автоматизируем в Python	0,7	0,3	0,4	Тест
6.2	Лабораторная работа	1,5	–	1,5	ДЗ
7	МОДУЛЬ 7. ИТОГОВЫЙ ПРОЕКТ	4,3	0,3	4,0	Итоговый проект
7.1	Инструктаж по проекту. Оформление итогового проекта	3,3	0,3	3,0	Инструктаж по публикации на GitHub
7.2	Итоговая аттестация	1,0	–	1,0	Формирование отчёта
Итого по части II		17,3	3,8	13,5	
ВСЕГО ПО ПРОГРАММЕ		42,35	12,6	29,75	

2.2. Календарный учебный график

Неделя	Форма занятия	Контактная работа (ч.)	Самостоятельная работа (ч.)	Раздел	Форма контроля
ЧАСТЬ I. DATA SCIENCE					
1	Самостоятельная	1,4	1,75	Модуль 2. Генерация данных	Тест, ДЗ
1	Самостоятельная	1,5	1,6	Модуль 3. Импорт данных	Тест, ДЗ
1	Самостоятельная	1,4	1,6	Модуль 4. Работа с таблицами Pandas	Тест, ДЗ
2	Самостоятельная	0,8	1,6	Модуль 5. Web-parsing	Тест, ДЗ
2	Самостоятельная	0,9	1,6	Модуль 6. Матрицы NumPy	Тест, ДЗ
2	Самостоятельная	1,3	1,6	Модуль 7. Визуализация в Python	Тест, ДЗ
3	Самостоятельная	0,5	1,5	Модуль 8. Оформление итогов	Тест, ДЗ
3–5	Самостоятельная	1,0	5,0	Модуль 9. Итоговая аттестация	Итоговый проект
ЧАСТЬ II. SQL					
6	Самостоятельная	0,6	1,9	Модуль 2. Настройка окружения	Тест, ДЗ
7	Самостоятельная	0,9	1,9	Модуль 3. Базовые структуры запросов. Синтаксис	Тест, ДЗ
8	Самостоятельная	1,0	1,9	Модуль 4. Преобразование таблиц	Тест, ДЗ
9	Самостоятельная	0,7	1,9	Модуль 5. Сложные запросы	Тест, ДЗ
10	Самостоятельная	0,3	1,9	Модуль 6. Реализация SQL внутри Python	Тест, ДЗ
11	Самостоятельная	0,3	4,0	Модуль 7. Итоговый проект «Аналитика маркетплейса»	Итоговый проект

Примечание. Нумерация недель приведена для последовательного освоения частей программы. При параллельном освоении частей календарный график может быть скорректирован при сохранении общего объёма программы в 42,35 академических часа.

2.3. Рабочая программа

ЧАСТЬ I. DATA SCIENCE

Модуль 2. Видеолекции «Общая информация о Data Science» и «Теоретический материал по роли данных в Data Science, типам данных и способам их генерации». Практический семинар по созданию наборов данных в Python, работе с базовыми структурами данных и подготовке данных к анализу. Самостоятельная работа: тест, домашнее задание по генерации данных (1,75 ак. ч.).

Модуль 3. Теоретическая видеолекция по источникам данных и форматам хранения информации. Практическая видеолекция по импорту данных из CSV, Excel, текстовых файлов и внешних источников в Python. Самостоятельная работа: тесты, задания по импорту данных (1,6 ак. ч.).

Модуль 4. Теоретическая видеолекция по структуре DataFrame и возможностям библиотеки Pandas. Практическая видеолекция по очистке, преобразованию, фильтрации, агрегации и анализу табличных данных. Самостоятельная работа: тесты, задания по работе с таблицами Pandas (1,6 ак. ч.).

Модуль 5. Теоретическая видеолекция по основам web-parsing и работе с веб-данными. Практическая видеолекция по извлечению данных с веб-страниц и преобразованию неструктурированных данных в аналитический формат. Самостоятельная работа: тесты, задания по web-parsing (1,6 ак. ч.).

Модуль 6. Теоретическая видеолекция по численным методам и матричным вычислениям в Data Science. Практическая видеолекция по работе с массивами NumPy, выполнению вычислений и подготовке данных к анализу. Самостоятельная работа: тесты, задания по матрицам NumPy (1,6 ак. ч.).

Модуль 7. Теоретическая видеолекция по принципам визуального анализа данных и автоматизации подготовки отчётных материалов. Практическая видеолекция по построению графиков и диаграмм, интерпретации результатов и выбору подходящих способов визуализации. Самостоятельная работа: тесты, задания по визуализации в Python (1,6 ак. ч.).

Модуль 8. Практическая видеолекция по созданию Excel-файлов, формированию таблиц, отчётов и презентаций средствами Python. Самостоятельная работа: тесты, задания по оформлению итогов (1,5 ак. ч.).

Модуль 9. Инструктаж по итоговому проекту (1 ак. ч.). Самостоятельная работа над сквозным проектом «Создание мини web-сервиса анализа данных селлера» (5 ак. ч.).

ЧАСТЬ II. SQL

Модуль 2. Практическая видеолекция «Настройка окружения. Введение. Подготовка и настройка»: установка необходимого программного обеспечения, настройка среды разработки, работа с терминалом и запуск первого тестового проекта. Самостоятельная работа: тест, лабораторная работа по подготовке данных для работы с запросами SQL (1,9 ак. ч.).

Модуль 3. Практическая видеолекция «Базовые структуры запросов. Синтаксис. Пишем запросы и получаем данные»: написание простых запросов, работа с основными синтаксическими конструкциями, извлечение и фильтрация данных. Самостоятельная работа: тесты, задания на отработку базовых запросов SQL (1,9 ак. ч.).

Модуль 4. Практическая видеолекция «Преобразование таблиц. Преобразовываем данные и объединяем»: изменение структуры таблиц, применение операций

преобразования данных и объединение таблиц с использованием различных типов соединений. Самостоятельная работа: тесты, лабораторная работа по работе с таблицами (1,9 ак. ч.).

Модуль 5. Практическая видеолекция «Сложные запросы. Запросы с подзапросами для глубокой аналитики»: написание сложных SQL-запросов, использование подзапросов, вложенных конструкций и анализ данных с их помощью. Самостоятельная работа: тесты, задания на отработку глубокой аналитики (1,9 ак. ч.).

Модуль 6. Практическая видеолекция «Реализация SQL внутри Python. Автоматизируем в Python»: интеграция SQL-запросов в Python, работа с библиотеками для подключения к базам данных и автоматизация обработки данных. Самостоятельная работа: тесты, лабораторная работа по интеграции SQL и Python (1,9 ак. ч.).

Модуль 7. Практическая видеолекция: инструктаж по оформлению итогов на GitHub. Самостоятельная работа: подготовка отчёта по итоговому проекту «Аналитика маркетплейса» (4 ак. ч.).

РАЗДЕЛ 3. КОНТРОЛЬ И ОЦЕНКА РЕЗУЛЬТАТОВ ОБУЧЕНИЯ ПО ПРОГРАММЕ

3.1. Формы аттестации

Текущий контроль: тестирование после каждой лекционной темы (проходной балл — 75 %), проверка домашних заданий, практических и лабораторных работ.

Итоговая аттестация проводится по каждой части программы в форме защиты итогового практического проекта с выставлением оценки «зачтено» / «не зачтено»:

- часть I «Data Science» — сквозной проект «Создание мини web-сервиса анализа данных селлера»;
- часть II «SQL» — итоговый проект «Аналитика маркетплейса» с публикацией результата на GitHub.

Для получения сертификата необходимо успешно пройти все тесты, выполнить все практические работы и получить оценку «зачтено» по итоговому проекту каждой части программы. В сертификате указываются наименования освоенных частей программы и их объём в академических часах.

3.2. Контрольно-оценочные средства текущего и итогового контроля

Типовые тестовые вопросы. Формат — выбор одного правильного ответа из нескольких.

Часть I. Data Science

- 1) Что может произойти, если анализировать только то, что удобно измерить?
 - а) Визуализация данных станет проще.
 - б) Ускорится процесс анализа.
 - в) Можно упустить важные факторы и сделать неверные выводы.
 - г) Результаты будут более точными.
- 2) Что является первым шагом в любом аналитическом проекте?
 - а) Формулировка гипотез.
 - б) Написание кода для анализа данных.
 - в) Создание и импорт данных.
 - г) Визуализация полученных результатов.
- 3) Что важно понимать при работе с сырыми данными?
 - а) Они всегда содержат всю необходимую информацию.
 - б) Они всегда чистые и структурированные.
 - в) Они могут быть неполными, противоречивыми и «грязными».
 - г) Их достаточно просто очистить стандартными инструментами.
- 4) Какие основные источники данных выделяются в программе?
 - а) Внутренние отчёты и статистика продаж.
 - б) Социальные сети и новостные сайты.
 - в) Пассивные данные и активно создаваемые данные.
 - г) Данные с датчиков и метеостанций.

Часть II. SQL

- 1) Какая команда используется для фильтрации данных в SQL?

- а) SELECT
 - б) WHERE
 - в) GROUP BY
 - г) ORDER BY
- 2) Что делает команда SELECT * FROM таблица?
- а) Выводит информацию о структуре таблицы.
 - б) Выбирает первую строку из таблицы.
 - в) Выбирает все столбцы из указанной таблицы.
 - г) Выбирает только столбцы с именем * из таблицы.
- 3) Какая команда используется для выбора определённых столбцов из таблицы в SQL?
- а) FROM
 - б) SELECT
 - в) WHERE
 - г) ORDER BY
- 4) Для чего используется библиотека Pandas при работе с базами данных?
- а) Для выполнения SQL-запросов
 - б) Для подключения к базе данных
 - в) Для работы с полученными таблицами
 - г) Для сохранения таблиц в Excel

Примеры итоговых заданий

Часть I. Data Science. Сквозной практический проект «Создание мини web-сервиса анализа данных селлера» (6 ак. ч.). Этапы: импорт библиотек; загрузка таблиц в базу данных; работа с кодом для пополнения базы данных; работа с кодом для расчёта аналитической таблицы; работа с кодом для web-приложения; презентация результатов.

Часть II. SQL. Итоговый практический проект «Аналитика маркетплейса» (4,3 ак. ч.). Этапы: разбор учебной базы интернет-магазина; документирование модели данных; построение набора бизнес-отчётов на SQL; размещение проекта на GitHub; презентация результатов.

Критерии оценки итогового проекта (по каждой части программы): полнота выполнения этапов работы — 0–20 баллов; корректность применения методов — 0–20 баллов; качество оформления и презентации результатов — 0–15 баллов; обоснованность выводов — 0–15 баллов. Проходной балл — 48 из 80 (60 %).

РАЗДЕЛ 4. ОРГАНИЗАЦИОННО-ПЕДАГОГИЧЕСКИЕ УСЛОВИЯ РЕАЛИЗАЦИИ ПРОГРАММЫ

4.1. Требования к материально-техническому обеспечению

При реализации данной образовательной программы созданы условия для функционирования электронной информационно-образовательной среды, включающей в себя электронные информационные ресурсы, электронные образовательные ресурсы, совокупность информационных технологий, телекоммуникационных технологий, соответствующих технологических средств, которые обеспечивают освоение слушателями образовательной программы в полном объеме независимо от места нахождения слушателей.

Формирование информационной среды осуществляется с помощью программной системы дистанционного обучения.

В условиях реализации образовательной программы в дистанционном формате с использованием дистанционных образовательных технологий слушатели сами обеспечивают создание необходимых материально-технических условий для осуществления учебной деятельности.

Необходимым минимальным условием использования дистанционных образовательных технологий является наличие интернет-браузера и подключения к сети Интернет. На компьютере также должен быть установлен комплект соответствующего программного обеспечения. Для работы с использованием аудиоканала, в том числе видеоконференций, необходимо наличие веб-камеры, микрофона и динамиков (наушников).

Слушатель должен иметь возможность использовать канал связи с пропускной способностью не ниже 512 Кбит/с, для более комфортной связи рекомендовано 1 Мбит/с. Требования к скорости доступа в сеть Интернет носят рекомендательный характер и должны соблюдаться в целях беспрепятственного и своевременного освоения образовательной программы.

Для эффективного внедрения дистанционных образовательных технологий и использования электронных образовательных ресурсов имеется качественный доступ педагогических работников и слушателей к информационно-телекоммуникационной сети Интернет (далее — сеть Интернет) с использованием установленных программно-технических средств для слушателей и педагогических работников на скорости не ниже 512 Кбит/с; в труднодоступных районах, подключаемых к сети Интернет с использованием спутниковых каналов связи, скорость прямого канала должна быть не ниже 512 Кбит/с, обратного — не ниже 128 Кбит/с; обеспечен порт доступа в сеть Интернет со скоростью не ниже 10 Мбит/с и возможностью установления не менее 20 одновременных сессий по 512 Кбит/с.

Услуга подключения к сети Интернет должна предоставляться в режиме 24 часа в сутки 7 дней в неделю без учёта объёмов потребляемого трафика за исключением перерывов для проведения необходимых ремонтных и профилактических работ при обеспечении совокупной доступности услуг не менее 99,5 % в месяц.

Для использования дистанционных образовательных технологий каждому слушателю и преподавателю предоставляется свободный доступ к средствам информационных и коммуникационных технологий.

В наличии необходимое программное обеспечение, серверное оборудование, обеспечивающее функционирование электронной информационно-образовательной среды, и высокоскоростной канал доступа к электронной информационно-образовательной среде:

1. Офисная мебель — стол и стул (рабочее место педагогического работника).
2. Демонстрационная доска (1 шт.).
3. Ноутбук (процессор с тактовой частотой от 2,0 ГГц, оперативная память — не менее 8 ГБ, твердотельный накопитель (SSD) — не менее 256 ГБ) (1 шт.).
4. Оборудование для подключения к сети Интернет — роутер Wi-Fi (1 шт.).
5. Наушники — гарнитура с микрофоном (1 шт.).
6. Установлено программное обеспечение, необходимое для осуществления учебного процесса общего назначения (антивирус Kaspersky Free для Windows, браузер Яндекс.Браузер, Яндекс.Документы).

Обучение проводится посредством электронной образовательной среды (платформы) bstudy.ru. Платформа позволяет:

- размещать обучающие материалы и задания;
- загружать слушателям выполненные письменные, фото- и видеозадания, а также вопросы в адрес педагога;
- проводить вебинары педагогами, а также в процессе их проведения задавать вопросы в форме текстовых сообщений;
- осуществлять контроль прогресса изучения учебных материалов, количества выполненных слушателями заданий, а также проверять выполненные ими задания.

Требования к персональному компьютеру слушателя:

- операционная система: Windows 7 и выше, macOS 10.12 и выше, Linux;
- оперативная память: 4 ГБ (рекомендуется 8 ГБ);
- доступ в сеть Интернет: не менее 2 Мбит/с;
- веб-камера и микрофон;
- необходимое программное обеспечение (всё бесплатное): Python 3.8+ (через дистрибутив Anaconda), Jupyter Notebook, среда разработки на выбор (VS Code, PyCharm Community, Google Colab), СУБД SQLite / PostgreSQL и клиент для работы с базой данных (DBeaver, pgAdmin или аналог), веб-браузер (Chrome, Firefox, Safari), учётная запись GitHub.

4.2. Информационно-методическое обеспечение обучения

Рекомендованная литература и информационные ресурсы:

Часть I. Data Science

1. McKinney, Wes — Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython. Полное практическое руководство по работе с табличными данными, NumPy и Pandas, написанное создателем библиотеки Pandas. Рекомендуется как основной источник для освоения анализа данных в Python.
2. VanderPlas, Jake — Python Data Science Handbook: Essential Tools for Working with Data. Комплексное справочное руководство по ключевым инструментам экосистемы

Python для Data Science: IPython, NumPy, Pandas, Matplotlib и др. Доступен также в виде онлайн-версии (Jupyter-ноутбуки).

3. Grus, Joel — Data Science from Scratch: First Principles with Python. Книга для начинающих, объясняющая основные принципы Data Science «с нуля» через реализацию ключевых алгоритмов и подходов на Python.
4. Gagolewski, Marek — Minimalist Data Wrangling with Python. Онлайн-доступное пособие по подготовке, очистке и обработке данных на Python (бесплатная PDF-версия).
5. Nelli, Fabio — Python Data Analytics: With Pandas, NumPy, and Matplotlib. Всесторонний источник по аналитике данных на Python с детальным разбором Pandas, NumPy и Matplotlib; полезен для углубления практических навыков.
6. Васильев Ю. — «Python для Data Science». Русскоязычное пособие по анализу данных с Python, подходящее для слушателей, изучающих Data Science на русском языке.
7. «Python для Data Science» / Издательский дом «Питер» — обзорное руководство по основам анализа данных с Python: NumPy, Pandas, Matplotlib и другие библиотеки экосистемы Python для задач анализа и визуализации.

Часть II. SQL

1. Бьюли А. — «Изучаем SQL» (Learning SQL). Практическое введение в язык SQL: создание таблиц, выборка данных, фильтрация, агрегатные функции и соединения. Подходит для начинающих как основной учебник.
2. Молинаро Э. — «SQL. Сборник рецептов» (SQL Cookbook). Сборник практических решений типовых задач на SQL. Полезен для закрепления навыков и быстрого поиска решений в реальных задачах.
3. Форта Б. — «SQL за 10 минут в день». Доступное руководство для быстрого освоения SQL с нуля; материал разбит на короткие уроки, удобные для самостоятельного изучения.
4. Силбершатц А., Корт Г., Сударшан С. — «Базы данных. Полный курс». Фундаментальный учебник по теории баз данных: архитектура СУБД, нормализация, транзакции, индексы.
5. Дейт К. Дж. — «Введение в системы баз данных». Классический академический труд по реляционной модели данных и теории баз данных.
6. Клеппман М. — «Высоконагруженные приложения. Проектирование, масштабирование, поддержка». Современное руководство по проектированию систем хранения данных, включая SQL и NoSQL.
7. Вихренко В. А. — «SQL для профессионалов». Русскоязычное практическое пособие, охватывающее продвинутые возможности SQL и оптимизацию запросов.
8. Кузнецов С. Д. — «Основы баз данных». Отечественный учебник, подробно раскрывающий основы реляционных баз данных и язык SQL.
9. PostgreSQL Global Development Group — «Документация PostgreSQL». Официальная документация одной из самых популярных СУБД; рекомендуется как справочник для практической работы.
10. Oracle — «SQL Language Reference». Официальная документация SQL от Oracle, содержащая полное описание синтаксиса и возможностей языка.

4.3. Кадровое обеспечение реализации программы

Квалификация руководящих и научно-педагогических работников организации соответствует квалификационным характеристикам, установленным действующим законодательством Российской Федерации в области образования.

Требования к квалификации преподавателей:

- наличие высшего образования и/или опыта профессиональной деятельности в сфере анализа данных, разработки программного обеспечения, биоинформатики, ИТ или цифровой трансформации;
- стаж работы в образовательной организации не менее 1 года либо наличие учёной степени (кандидата или доктора наук) без предъявления требований к стажу работы;
- регулярное повышение квалификации (не реже одного раза в три года), в том числе через программы по применению дистанционных образовательных технологий (ДОТ) и электронного обучения;
- владение современными методами преподавания, включая разработку и использование электронных образовательных ресурсов.

Преподаватели, участвующие в реализации программы с применением дистанционных технологий, проходят повышение квалификации в области информационных и образовательных технологий не реже одного раза в 5 лет. Для поддержания актуальности профессиональных компетенций организуются внутренние семинары, конференции и стажировки.