

Документ подписан простой электронной подписью  
Информация о владельце:  
ФИО: Ливанов Дмитрий Викторович  
Должность: Ректор  
Дата подписания: 19.03.2022  
Уникальный программный ключ:  
с6d909c49c1d2034fa3a0156c4eaa51e7232a3a2

Федеральное государственное автономное образовательное учреждение  
высшего образования  
«Московский физико-технический институт  
(национальный исследовательский университет)»

**ВЫПИСКА ИЗ ПРОТОКОЛА № 10**  
**заседания учебно-методического совета от 26 мая 2022 г.**

ПОВЕСТКА:

Рассмотрение дополнительных общеобразовательных и профессиональных программ.

Проректор по учебной работе А. А. Воронов.

СЛУШАЛИ: Директора Физтех-школы прикладной математики и информатики  
А.М. Райгородского.

ПОСТАНОВИЛИ:

Рекомендовать к утверждению в установленном порядке дополнительную  
общеобразовательную программу «Data Analyst».

Решение принято единогласно заочным голосованием.

Форма проведения заседания: заочная.

Председатель УМС МФТИ

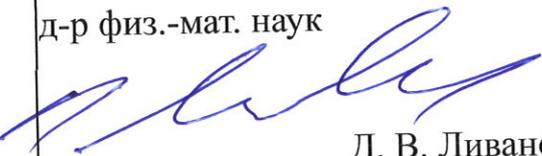
А.А. Воронов

Ученый секретарь УМС МФТИ

М.В. Березникова

Министерство науки и высшего образования Российской Федерации  
федеральное государственное автономное образовательное учреждение высшего  
образования

«Московский физико-технический институт  
(национальный исследовательский университет)»

	УТВЕРЖДАЮ
	Ректор д-р физ.-мат. наук
	 Д. В. Ливанов
	«26» мая 2022 г.

**Дополнительная общеобразовательная программа  
«Data Analyst»**

Москва 2022

## Оглавление

1. Рабочая группа	2
2. Профессиональные стандарты, учтённые в содержании программы	2
3. Квалификационные требования, учтённые в содержании программы	2
4. Требования федерального государственного стандарта высшего образования, на основании которых была разработана программа	2
5. Характеристика новой квалификации и связанных с ней видов профессиональной деятельности, трудовых функций и (или) уровней квалификации.	2
6. Характеристика компетенций, подлежащих совершенствованию, и (или) перечень компетенций, формирующихся в результате освоения программы	4
7. Цель программы	4
8. Планируемые результаты обучения по программе	4
9. Учебный план	5
10. Календарный учебный график	6
11. Рабочие программы (учебная программа)	7
12. Организационно-педагогические условия реализации дополнительной профессиональной программы	17
12.1. Требования к квалификации педагогических кадров/представителей предприятий и организаций, обеспечивающих реализацию образовательного процесса	17
12.2. Требования к материально-техническим условиям	18
12.3. Требованиям к информационным и учебно-методическим условиям*	18
12.4. Общие требования к организации образовательного процесса*	20
13. Формы аттестации и оценочные материалы по программе	21

## **1. Рабочая группа**

С целью повышения качества дополнительной общеобразовательной программы (ДОП) с учетом требований профессиональных стандартов (ПС) в группу разработчиков входили:

**Райгородский Андрей Михайлович**

Доктор физико-математических наук, директор ФПМИ МФТИ

**Благодарный Евгений Владимирович**

заведующий учебно-методической лабораторией инноватики

**Иванова Анастасия Сергеевна**

руководитель проектов учебно-методической лаборатории инноватики

## **2. Профессиональные стандарты, учтённые в содержании программы**

Настоящая дополнительная общеобразовательная программа (ДОП) разработана с учётом соответствующих ей профессиональных стандартов (ПС) из национального реестра профессиональных стандартов. Профессиональные стандарты (ПС), связанные с настоящей ДПП:

*06.042 Специалист по большим данным*

*08.037 Бизнес-аналитик*

## **3. Квалификационные требования, учтённые в содержании программы**

Квалификационные требования, учтённые в содержании программы, указаны в описании должностей единого квалификационного справочника должностей руководителей, специалистов и служащих. Причём учтены требования должностей, указанных в дополнительных характеристиках, отобранных для данной программы ОТФ (п. 5. настоящего документа), которые приведены в указанных в п. 2. настоящего документа профессиональных стандартах (ПС) в строке «ЕТКС или ЕКС». Квалификационные требования инкорпорированы в настоящий документ путём отсылки и являются его неотъемлемой частью.

## **4. Требования федерального государственного стандарта высшего образования, на основании которых была разработана программа**

Настоящей дополнительной профессиональной программе соответствует федеральный государственный стандарт высшего образования (ФГОС ВО):

*01.03.02 “Прикладная математика и информатика” (бакалавриат)*

Настоящая дополнительная общеобразовательная программа разработана на основании требований указанного ФГОС ВО.

Полный перечень требований приведён в тексте указанного ФГОС ВО, инкорпорирован в настоящий документ путём отсылки и является его неотъемлемой частью.

## **5. Характеристика новой квалификации и связанных с ней видов профессиональной деятельности, трудовых функций и (или) уровней квалификации.**

«Программа общеобразовательной подготовки направлена на получение компетенции, необходимой для выполнения нового вида профессиональной деятельности, приобретение новой квалификации (часть 5 статьи 76 Федерального закона N 273-ФЗ). Для их определения и может использоваться ПС. При этом необходимо учесть различия терминологии, используемой

в образовании и в профессиональных стандартах, о которых говорилось выше. Вид профессиональной деятельности, квалификация, упоминаемые в цитируемой статье закона, в ПС в большинстве случаев соответствуют ОТФ, иногда - ТФ.»<sup>1</sup>,

Новые квалификации (ОТФ/ТФ) и их уровни:

Вид профессиональной деятельности (ПС)	Новые квалификации (ОТФ/ТФ)	Уровень
Деятельность по выявлению бизнес-проблем, выяснению потребностей заинтересованных сторон, обоснованию решений и обеспечению проведения изменений в организации	Сбор информации о бизнес-проблемах или бизнес-возможностях	5
	Выявление истинных бизнес-проблем или бизнес-возможностей	5
Создание и применение технологий больших данных	Анализ больших данных с использованием существующей в организации методологической и технологической инфраструктуры	6

Характеристики данных ОТФ/ТФ приведены в профессиональном стандарте (ПС), указанном в п. 2. настоящего документа. инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Связанные виды профессиональной деятельности (ВПД):

*научно-исследовательская деятельность*

*проектная и производственно-технологическая деятельность*

Характеристика данных ВПД приведена в федеральном государственном стандарте высшего образования (ФГОС ВО), указанном в п. 4. настоящего документа, инкорпорирована в настоящий документ путём отсылки, и является его неотъемлемой частью. Указанные ВПД, освоение которых предусмотрено указанным ФГОС ВО, являются «сквозными» и служат основой овладения выбранной квалификацией (ОТФ/ТФ).

ТФ, связанные с указанными в настоящем пункте ОТФ (в случае их указания), а также их характеристики приведены в профессиональном стандарте (ПС), указанном в п. 2. настоящего документа. инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

<sup>1</sup> Письмо Министерства образования и науки РФ от 22 апреля 2015 г. N ВК-1032/06 "О направлении методических рекомендаций"

## **6. Характеристика компетенций, подлежащих совершенствованию, и (или) перечень компетенций, формирующихся в результате освоения программы**

Учитывая, что слушатели программы уже освоили или осваивают основную профессиональную образовательную программу, в том числе общие компетенции, соответствующего уровня профессионального образования, компетенции формирующиеся в результате освоения настоящей программы и компетенции, подлежащие совершенствованию, а также их характеристика, представлены в ФГОС, указанном в п. 4. настоящего документа, инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Данные компетенции указанного ФГОС ВО служат основой определения результатов обучения по программе.

## **7. Цель программы**

Реализация программы профессиональной переподготовки направлена на получение компетенции, необходимой для выполнения нового вида профессиональной деятельности, приобретение новой квалификации.

В соответствие с Письмом Министерства образования и науки РФ от 22 апреля 2015 г. N ВК-1032/06 «О направлении методических рекомендаций», цель представляет собой осознанное представление (предвосхищение) результата деятельности. Планируемые результаты отражены ниже.

## **8. Планируемые результаты обучения по программе**

Планируемые результаты обучения соответствуют результатам освоения соответствующей основной профессиональной образовательной программы, а также направлены на приобретение новой квалификации, требующей изменение направленности (профиля) или специализации в рамках направления подготовки (специальности) полученного ранее профессионального образования, определены на основе профессиональных компетенций соответствующего федерального государственного образовательного стандарта, указанного в п. 4. настоящего документа.

Имеющаяся квалификация и (или) уровень образования (требования к слушателям):

*нет*

Новые виды профессиональной деятельности, и ПК, соответствующие им, которые планируются к получению в результате обучения по настоящей программе, а также их характеристики, представлены в ФГОС, указанном в п. 4. настоящего документа, инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Выпускник также должен обладать прочими компетенциями в соответствие с указанным ФГОС, которые приведены и охарактеризованы в нём, инкорпорированы в настоящий документ путём отсылки, и являются его неотъемлемой частью.

Новые квалификации (ОТФ/ТФ), планируемые к приобретению в результате обучения по настоящей программе, приведены и охарактеризованы в п. 5. настоящего документа.

Приведённая информация о цели и результатах обучения является основой для разработки рабочих программ, оценочных материалов и иных компонентов дополнительной профессиональной программы.

## 9. Учебный план

№ п/п *	Наименование Модулей, дисциплин	Всего, час.	В том числе:				
			Лек-ци и	Практиче- ские занятия (семинар ы)	Самостоя-т ельная работа	Контроль- ные зада-ния	Форма контроля
1	Модуль 1 Python и инструменты машинного обучения	100	24	24	46	6	проект
2	Модуль 2 Базовый SQL	48	12	12	20	4	тестирование
3	Модуль 3 Математика для анализа данных. (Теория вероятностей и математическая статистика)	64	20	20	20	4	тестирование
3	Модуль 4 Инструменты анализа и визуализации данных	100	24	24	46	6	проект
4	Итоговая квалификационная работа	35			35		
	ИТОГО:	347	80	80	167	20	

\* Слушатели могут проходить модули в произвольном порядке

Срок освоения настоящей программы профессиональной переподготовки превышает минимально допустимый срок освоения 250 часов (либо другой установленный актуальными нормативными документами соответствующего Министерства срок).

## 10. Календарный учебный график

### Модуль 1 Python и инструменты машинного обучения

Обучение по установленной форме: очная, с применением электронного обучения и дистанционных образовательных технологий													
Количество академических часов: 100													
Порядковый № занятия (по горизонтали)	1	2	3	4	5	6	7	8	9	10	11	12	13
Лекции***	2	2	2	2	2	2	2	2	2	2	2	2	
практические	2	2	2	2	2	2	2	2	2	2	2	2	
самостоятельные	3	3	3	3	3	3	3	3	3	3	3	3	10
Контрольные задания	0	1	0	1	0	1	0	1	0	1	0	1	
Итого	7	8	7	8	7	8	7	8	7	8	7	8	10

### Модуль 2 Базовый SQL

Обучение по установленной форме: очная с применением дистанционных технологий						
Количество академических часов: 48						
Порядковый № занятия (по горизонтали)	1	2	3	4	5	6
Лекции	2	2	2	2	2	2
Практические занятия	2	2	2	2	2	2
Самостоятельная работа	2	2	4	4	4	4
Контрольные задания	0,5	0,5	0,75	0,75	0,75	0,75
Итого	6,5	6,5	8,75	8,75	8,75	8,75

### Модуль 3. Математика для анализа данных (Теория вероятностей и математическая статистика)

Обучение по установленной форме: очная с применением дистанционных технологий										
Количество академических часов: 64										
Порядковый № занятия (по горизонтали)	1	2	3	4	5	6	7	8	9	10
Лекции***	2	2	2	2	2	2	2	2	2	2
практические	2	2	2	2	2	2	2	2	2	2
самостоятельные	2	2	2	2	2	2	2	2	2	2
Контрольные задания	0	1	0	1	0	1	0	0	1	
Итого	6	7	6	7	6	7	6	6	7	6

## Модуль 4. Инструменты анализа и визуализации данных

Обучение по установленной форме: очная, с применением электронного обучения и дистанционных образовательных технологий

Количество академических часов: 100

Порядковый № занятия (по горизонтали)	1	2	3	4	5	6	7	8	9	10	11	12	13
Лекции***	2	2	2	2	2	2	2	2	2	2	2	2	
практические	2	2	2	2	2	2	2	2	2	2	2	2	
самостоятельные	3	3	3	3	3	3	3	3	3	3	3	3	10
Контрольные задания	0	1	0	1	0	1	0	1	0	1	0	1	
Итого	7	8	7	8	7	8	7	8	7	8	7	8	10

### Итоговая квалификационная работа

Обучение по установленной форме: очная с применением дистанционных технологий	
Количество академических часов: 35	
самостоятельная работа	35
Итого	35

## 11. Рабочие программы (учебная программа)

### Модуль 1 Python и инструменты машинного обучения

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем, ак.час.
1.1	Знакомство с Python (часть 1).	7
	Лекция. Введение в Python. работа в Jupyter Notebook и Google Colab, числовые типы данных, условия и циклы.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
1.2	Управление вычислениями. Контейнеры, итераторы, генераторы	8
	Лекция. Set, dict; enumerate, zip, map; list, dict, set comprehensions; itertools, collections.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3

	Выполнение контрольных заданий.	1
1.3	Введение в анализ данных с помощью pandas	7
	Лекция. Введение в анализ данных в Pandas. Фильтрация строк по условию, группировка данных. Работа с несколькими таблицами. Чтение-запись таблиц в файл, формат хранения данных .csv.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
1.4	Модули для визуализации данных	8
	Лекция. Введение в matplotlib. Визуализация данных при помощи графиков и гистограмм. Другие модули для визуализации.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
1.5	Оптимизация кода с помощью numru, обзор модуля. Работа со строками.	7
	Лекция. введение в Numru, оптимизация кода, работа со строками.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
1.6	Функциональное программирование в Python	8
	Лекция.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
1.7.	ООП в Python	7
	Лекция. Объявление класса, создание экземпляра. Атрибуты и методы класса, приватность. Наследование. Множественное наследование. Ромбовидное наследование. Приватность атрибутов. Объектно-ориентированный подход к созданию итераторов и генераторов.	2

	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
1.8.	Командная строка Unix	8
	Лекция. – Введение в командную строку. Взаимодействие с оболочкой, навигация и подключение программ. Скрипты; bash. Инструменты оболочки, поиск: файлов; содержимого; команд. Vim, режимы работы. Вставка текста, буферы, вкладки и окна. Быстрые команды и макросы. Работа с процессами. Терминальные мультиплекторы. Alias. SSH, генерация ключа, копирование файлов. Port Forwarding	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
1.9	GIT, профилирование и параллелизм.	7
	Лекция. Версионный контроль git. Модель данных git. Командная строка git. Дебаг и логирование. Профилирование по времени и памяти. Джит компиляция и параллелизм в python. Метaprogramмирование: системы сборки, зависимостей. Тестирующие системы (CI).	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
1.10	Python и web: обзор.	8
	Лекция. Протокол HTTP, Домены, URL, HTTP-сообщения, основы HTML. Парсинг данных с помощью beautiful soup, прокси. Парсинг и автоматизация взаимодействие с веб с помощью selenium. Быстрое прототипирование и деплоинг мл-моделей со streamlit на примере простейшего классификатора интенгов. Обзор примеры работы Flask, верстка с bootstrap.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
1.11	Теория вероятностей и статистика данных в Python. Понятия о методологии A/B тестирования	7

	Лекция.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	0
1.12	Пример использования Python для машинного обучения. Модули scikit-learn, Pytorch.	8
	Лекция.	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	3
	Выполнение контрольных заданий.	1
1.13	Промежуточная аттестация.	10
	Итого	100

## Модуль 2. SQL базовый

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем, ак. час.
1.1	Базовый синтаксис SQL	6,5
	Лекция. Базовый синтаксис SQL (синтаксис запроса, выборка, проекция)	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	0,5
1.2	Создание и заполнение таблиц	6,5
	Лекция. Создание и заполнение таблиц (понятие DDL и DML, создание таблиц, изменение таблиц, добавление данных, изменение данных, удаление данных)	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	1

	Выполнение контрольных заданий.	0,5
1.3	Объединение данных	8,75
	Лекция. Объединение данных (Union / Union All ,Особенности Join, Типы Join и различия между ними)	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	4
	Выполнение контрольных заданий.	0,75
1.4	Агрегация	8,75
	Лекция. Агрегация (агрегация, группировка having, применение подзапросов, оператор case)	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	4
	Выполнение контрольных заданий.	0,75
1.5	Очистка данных	8,75
	Лекция. Очистка данных (строковые методы, очистка данных, отработка очистки данных)	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	4
	Выполнение контрольных заданий.	0,75
1.6	Регулярные выражения	8,75

	Лекция.Регулярные выражения (синтаксис регулярных выражений, методы работы с регулярными выражениями, отработка очистки данных)	2
	Практическое занятие по теме лекции	2
	Самостоятельная работа.	4
	Выполнение контрольных заданий.	0,75
	Итого	48

**Модуль 3 Математика для анализа данных (Теория вероятностей и математическая статистика)**

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем, ак.час.
3.1.	Теория вероятностей.	26
3.1.1	Лекция: Конечные множества элементарных событий. Классическая вероятность. Схема Бернулли. Условная вероятность. Независимость событий.	2
	Практические занятия: 1)Решение разнообразных задач по классической и геометрической вероятности 2)Решение задач на условную вероятность	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	0
3.1.2	Лекция: Распределения вероятностей. Дискретные случайные величины и их характеристики.	2
	Практические занятия: Поиск математических ожиданий и дисперсий дискретных случайных величин	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	1
3.1.3	Лекция. Математическое ожидание и дисперсия дискретных случайных величин. Неравенство Чебышева. Закон Больших Чисел. Центральная предельная теорема	2
	Практическое занятие: Применения ЦПТ на практике, вычисление математических ожиданий в абсолютно непрерывном случае	2

	Самостоятельная работа.	2
	Выполнение контрольных заданий.	0
3.1.4	Лекция: Непрерывные случайные величины.	2
	Практическое занятие : применение формулы свёртки	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	1
3.2.	Математическая статистика и элементы аналитики	
3.2.1	Лекция. Задача математической статистики и точечные оценки.	2
	Практическое занятие: Построение оценок по методу моментов и методу максимального правдоподобия	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	0
3.2.2	Лекция: Доверительные интервалы	2
	Практическое занятие: Построение доверительных интервалов в теории и на практике	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	1
3.2.3	Лекция: Проверка гипотез и статистические критерии	2
	Практическое занятие: Построение статистических критериев, измерение качества критериев.	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	0
3.2.4	Лекция: Корреляционный анализ	2
	Практическое занятие: Вычисление корреляций. Связь с метриками машинного обучения	2
	Самостоятельная работа.	2
	Выполнение контрольных заданий.	0
3.2.5	Лекция: Линейная регрессия	2
	Практическое занятие: Построение оценок в линейной регрессионной модели	2

	Самостоятельная работа	2
	Выполнение контрольных заданий	1
3.2.6	Лекция: Некоторые вопросы современной математической статистики.	2
	Практическое занятие: решение задач по математической статистике	2
	Самостоятельная работа	2
	Выполнение контрольных заданий	0
	Итого	64

#### Модуль 4. Инструменты анализа и визуализации данных

№ п/п	Содержание обучения, наименование и тематика практических занятий (вебинаров), самостоятельных работы	Объем, ак. час.
1	Модуль 1. Визуализация данных с помощью Excel	15
1.1	Лекция 1: Анализ данных в Excel. Импорт данных из внешних источников. Структура книг и листов. Предобработка данных. Вычисления и формулы. Проверка данных, поиск ошибок. Сводные таблицы. Умные таблицы. Формулы массива. Динамические массивы. ВПР. Форматы и значения. Стилизовое и условное форматирование Фильтрация.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	0
1.2	Лекция 2: Визуализация и подготовка отчетов/дашбордов в Excel Базовые диаграммы и спарклайны Сложные диаграммы. Динамические диаграммы Оптимизация и прогнозирование. Создание дашбордов.	1
	Практическое занятие	3
	Самостоятельная работа	3
	Выполнение контрольных заданий	1
2	Модуль 2. Визуализация данных с помощью Power BI	45

2.1.	Лекция 3. Обзор возможностей Power BI. Концепция работы в Power BI, ее назначение и свойства Введение в анализ данных. Советы по созданию визуализаций Компоненты и архитектура Power BI Представление данных.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	0
2.2	Лекция 4. Загрузка и предобработка данных в Power BI Получение данных с помощью Power Query Синтаксис языка M и работа со структурными данными Сложные запросы Параметризация и работа с функциями	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	1
2.3.	Лекция 5. Основы анализа данных в Power BI Введение в язык анализа данных Data Analysis Expressions (DAX). Создание и настройка модели данных. Создание вычисляемых столбцов. Создание мер. Функции аналитики времени. Фильтрация данных.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	0
2.4.	Лекция 6. Продвинутый анализ данных в Power BI Создание модели данных. Управление контекстом. Функции преобразование таблиц и передача контекста при отсутствии прямой связи. Извлечения информации о контексте. Форматирование полученных результатов	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	1

2.5.	Лекция 7. Интерактивная визуализация данных в Power BI Виды графиков и интерактивных элементов. Простые и сложные сценарии создания визуальных элементов. Взаимодействие между визуальными элементами. Интерактивная фильтрация.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	0
2.6.	Лекция 8. Публикация данных и совместная работа с отчетами Настройка и подготовка внешнего вида дашбордов для внешних пользователей. Публикация отчетов и дашбордов. Создание отчетов в браузере. Работа с готовыми отчетами в браузере, мобильной и десктопных версиях Power BI.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	1
3.	Модуль 3. Визуализация данных с помощью Tableau	30
3.1.	Лекция 9. Знакомство с Tableau Знакомство с Tableau. Основные элементы Загрузка и предобработка данных Создание переменных и мер. Построение первых графиков. Отличие Public и Professional версий.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	0
3.2.	Лекция 10. Аналитика данных в Tableau Основы работы с расчетными полями, фильтрами, множествами и группировками Использование параметров, функции и операторы в расчетах. Базовые операции с метриками. Цвета и шрифты при работе с данными. Форматирование данных. Интерактивные фильтры.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	1

3.3.	Лекция 11. Визуализация данных в Tableau Табличные вычисления + Функции LOD. Иерархии. Интерактивность при работе с данными. Вкладка Analytics.	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	0
3.4.	Лекция 12. Построение дашбордов, публикация результатов Создание дашбордов. Сторителлинг. Адаптивный дизайн. Настройка взаимодействия между визуализациями. Совместное использование отчетов: Tableau Server, Tableau Online	2
	Практическое занятие	2
	Самостоятельная работа	3
	Выполнение контрольных заданий	1
4	Промежуточная аттестация.	10
	Итого	100

5.	<b>Итоговая квалификационная работа</b>	<b>35</b>
	Самостоятельная работа.	35
	<b>Итого</b>	<b>347</b>

Содержание дополнительной профессиональной программы направлено на достижение результатов её целей (планируемых результатов).

## **12. Организационно-педагогические условия реализации дополнительной профессиональной программы**

### **12.1. Требования к квалификации педагогических кадров/представителей предприятий и организаций, обеспечивающих реализацию образовательного процесса**

Требования к образованию педагогических и иных работников, а также (при наличии) требования к освоению ими дополнительных профессиональных программ, опыту работы в области профессиональной деятельности, соответствующей направленности ДПП:

высшее образование, трудовой и (или) педагогический стаж от полугода, наличие опыта практической работы в сфере деятельности, совпадающей с направлением преподавания

### 12.2. Требования к материально-техническим условиям

Перечень кабинетов, лабораторий, мастерских, тренажеров и др., обеспечивающих проведение всех предусмотренных программой видов занятий:

Наименование специализированных аудиторий, кабинетов, лабораторий	Вид занятий	Наименование оборудования, программного обеспечения
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Лекции	<p>Слушателю необходимо наличие доступа в сеть интернет, компьютер.</p> <p>Преподавателю курса необходимо наличие доступа администратора курса на LMS-платформе к материалам курса.</p>
Информационно-коммуникационная платформа дистанционных семинаров	Практические занятия (дистанционные семинары)	<p>Слушателю необходимо наличие доступа в сеть интернет, компьютер.</p> <p>Преподавателю курса необходимо оборудование для проведения дистанционных семинаров (вебинаров), качественный отказоустойчивый доступ в сеть интернет.</p>
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Самостоятельная работа	Наличие компьютера и доступа в сеть интернет.
Система дистанционного обучения провайдера массовых открытых онлайн курсов	Промежуточный, рубежный контроль, Итоговая аттестация	Наличие компьютера и доступа в сеть интернет.

### 12.3. Требования к информационным и учебно-методическим условиям

Список литературы:

## МОДУЛЬ 1:

1. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Springer, 2009.
2. Murphy K.P. Machine Learning: A Probabilistic Perspective. The MIT Press, 2012.
3. Sheldon A. Linear Algebra Done Right. Springer, 2015.
4. DasGupta A. Probability for statistics and machine learning. Springer, 2011.
5. Иванов Г.Е. Лекции по математическому анализу. Ч.1 - М.:МФТИ, 2004, 2011.
6. Умнов. Аналитическая геометрия и линейная алгебра (2011) — МФТИ.
7. Нестеров. Методы выпуклой оптимизации (2010)
8. Diez, Barr, Çetinkaya-Rundel, Dorazio. Advanced High School Statistics (2015)
9. DasGupta. Probability for Statistics and Machine Learning: Fundamentals and Advanced Topics (2011)
10. Enders, W., 2003, Applied Econometric Time Series, Wiley Publ., 2nd ed.
11. Vance Martin, Stan Hurn, David Harris. Econometric Modelling with Time Series.
12. Specification, Estimation and Testing. Cambridge University Press, 2013.
13. Прикладная и компьютерная лингвистика / Под ред. И. С. Николаева, О. В. Митрениной, Т. М. Ландо. — М.: URSS, 2016. — 320 с.
14. Автоматическая обработка текстов на естественном языке и компьютерная лингвистика : учеб. пособие / Большакова Е.И., Клышинский Э.С., Ландэ Д.В., Носков А.А., Пескова О.В., Ягунова Е.В. — М.: МИЭМ, 2011. — 272 с.
15. Введение в информационный поиск / Кристофер Д. Маннинг, Прабхакар Рагхаван, Хайнрих Шютце. — М.: Вильямс, 2011. — 528 с.
16. Natural Language Processing for the Working Programmer / Daniël de Kok, 2011.

## МОДУЛЬ 2:

1. Bob Ward. SQL Server 2019 Revealed: Including Big Data Clusters and Machine Learning.
2. Грофф, Джеймс . SQL : Энциклопедия : пер. с англ. / Д. Р. Грофф, П. Н. Вайнберг. — 3-е изд. — СПб. : Питер. 2016
3. Дэви Силен, Арно Мейсман, Мохамед Али. Основа Data Science и BigData. Python и наука о данных. Питер. 2017

## МОДУЛЬ 3:

4. Hastie T., Tibshirani R., Friedman J. The Elements of Statistical Learning. Springer, 2009.
5. Bishop C.M. Pattern Recognition and Machine Learning. Springer, 2006.
6. Murphy K.P. Machine Learning: A Probabilistic Perspective. The MIT Press, 2012.
7. Boyd S., Vandenberghe L. Convex Optimization. Cambridge University Press, 2009.
8. Sheldon A. Linear Algebra Done Right. Springer, 2015.
9. DasGupta A. Probability for statistics and machine learning. Springer, 2011.
10. Беклемишев Д.В. Курс аналитической геометрии и линейной алгебры - 10-е изд. - М. Наука, 2003.
11. Иванов Г.Е. Лекции по математическому анализу. Ч.1 - М.:МФТИ, 2004, 2011.
12. Тер - Крикоров А.М., Шабунин М.И. Курс математического анализа - М.Ж: МФТИ, 1997, М.: Физматлит, 2003-2007.
13. Умнов. Аналитическая геометрия и линейная алгебра (2011) — МФТИ.
14. Деммель. Вычислительная линейная алгебра. Теория и приложения (2001)
15. Ильин, Позняк, Основы математического анализа (2005) — МГУ.
16. Тер-Крикоров, Шабунин. Курс математического анализа (2001) — МФТИ

## МОДУЛЬ 4:

1. Эдвард Тафти. Наглядное отображение количественной информации (The Visual Display of Quantitative Information), Graphics Press, 1983
2. Натан Яу. Визуализируйте это: руководство по FlowingData для проектирования, визуализации и статистики (Visualize This: The FlowingData Guide to Design, Visualization, and Statistics), John Wiley & Sons, 2011
3. Стивен Фью. Дизайн информационной панели: отображение данных для мгновенного мониторинга (Information Dashboard Design: Displaying Data for at-a-glance Monitoring), Analytics Press, 2013
4. Фил Саймон. Визуальная организация: визуализация данных, большие данные и поиск лучших решений (The Visual Organization: Data Visualization, Big Data, and the Quest for Better Decisions), Wiley, 2014
5. Эйлин и Стивен МакДэниел. Случайный аналитик: покажите своим данным кто босс (The Accidental Analyst: Show Your Data Who's Boss). Независимая платформа CreateSpace, 2012

#### 12.4. Общие требования к организации образовательного процесса

Форма обучения – очная с применением дистанционных технологий

Режим обучения – 2 раз(а) в неделю по 4 академических часа.

Преподаватель проводит практические занятия дистанционно в форме вебинаров с использованием платформы ZOOM (или аналогичной).

Самостоятельная работа выполняется слушателем в удобном для слушателя режиме

В Таблице ниже описаны образовательные технологии.

№ п/п	Вид занятия	Форма проведения занятий	Цель
1	Лекция	Изложение материала посредством лекций, обсуждение общих вопросов по тематике курса	Ознакомление слушателей с базовым материалом по тематике курса
2	Практическое занятие	Выполнение практических заданий, получение обратной связи от преподавателя.	Практическое освоение теоретических знаний, Разбор ситуаций, а также углубление знаний по курсу
3	Самостоятельная работа	Самостоятельное изучение дополнительных материалов и литературы. Выполнение тренировочных тестов и заданий.	Углубление знаний по курсу. Применение знаний к своей организации / проекту
4	Выполнение контрольных заданий	Выполнение тестов, проверочных заданий.	Практическое освоение теоретических знаний, контроль освоения материалов.
5	Итоговая аттестация	Подготовка итоговой аттестационной работы.	Практическое освоение теоретических знаний, контроль освоения материалов. Получение

			практического результата в своих организациях / проектах
--	--	--	--

### 13. Формы аттестации и оценочные материалы по программе

Оценка качества освоения программы проводится по двухбалльной системе: «зачтено», «не зачтено» по результатам промежуточного контроля (тестирование, проверочные задания на взаимную оценку), контроля посещаемости практических занятий (вебинаров) и результатам итоговой аттестации.

Слушатель считается аттестованным в случае положительных результатов работы (не менее 70% баллов от итоговой оценки) в процессе обучения и успешной сдачи экзамена. При этом баллы за экзамены начисляются только при достижении 50% порога при прохождении каждого экзаменационного испытания. После аттестации слушатель получает оценку “отлично”, если набрано не меньше 80% баллов от возможного максимума, “хорошо”, если набрано не меньше 65%, “удовлетворительно”, если набрано не меньше 40%, в противном случае слушатель курс не сдает.

Результат тестирования, решения проверочных заданий и написания кода проверяется автоматически системой на образовательной платформе. Итоговая аттестационная работа проверяется преподавателем

Составляющие процесса обучения, которые оцениваются в ходе обучения, и их вклад в итоговую оценку представлены в Таблице ниже.

Таблица – Составляющие процесса обучения

	Основные показатели оценки	Вклад в итоговую оценку
1	Основной курс обучения на образовательной платформе	50%
2	Практические занятия	10%
3	Итоговая аттестация	40%

#### Оценочные материалы:

1) Пример тестового задания к модулю 1 «Python и инструменты машинного обучения»

*Создайте одномерный массив чисел от 0 до 8.*

*Сделайте это двумя способами: с помощью метода `arange` из `numpy` и с помощью преобразования списка (`list`) в `numpy array`*

2) Пример тестового задания к модулю 2 “Базовый SQL”

*Имеется база данных интернет-магазина. В этой БД имеются таблицы:*

*Покупатели – хранит информацию по клиентам (ФИО, логин, пароль и т.п.);*

*Заказы – содержит номер заказа, комментарий, дату доставки, покупателя (который сделал заказ) и т.д.;*

*Товары – `id` товара, наименование, цена, остатки и т.п.;*

*Заказанные товары – хранит информацию о заказанных товарах (заказ, товар, количество).*

*Один к одному*

*Один ко многим*

*Многие ко многим*

*Таблицы никак не связаны*

3) Пример тестового задания к модулю 3 «Теория вероятностей и математическая статистика»

*В веб-поиске при решении различных задач машинного обучения часто используется статистический метод, который называется бутстрэппинг. Суть состоит в следующем. Предположим, что у нас есть  $N$  веб-страниц. Мы хотим узнать, насколько наш алгоритм устойчив. Для этого выбираются случайно  $N$  страниц (некоторые могут совпадать) большое количество раз  $M$  (в этом случае говорят, что генерируется  $M$  выборок размера  $N$ ). Если выбор производится упорядоченно, то найдите вероятность того, что первая страница встречается в одной такой выборке  $k$  раз, а вторая —  $r$  раз.*

3) Пример тестового задания к модулю 4 «Инструменты для анализа и визуализации данных»

*Новосибирская строительная компания продаёт квартиры в жилом комплексе «Миргород».*

*Данные о продажах: Mirgorod-sales.xls.*

*1. Ответьте на вопросы. Что является элементарным кирпичиком данных? Какие закономерности проявляются на макроуровне? Какой способ представления будет наиболее естественным и наглядным? Каких параметров, количественных и качественных, не хватает в наборе?*

*2. Подготовьте отчёт для руководства. Сколько квартир продано, на какую сумму? Сколько получено средств? Какие квартиры продаются лучше всего?*

*3. Помогите покупателю выбрать квартиру: визуализируйте доступные варианты и необходимую для выбора информацию. Используйте интерактивность.*

Согласовано

Директор ФПМИ, д.ф.-м.н.

А.М. Райгородский

# Пояснительная записка к разработке и реализации дополнительной общеобразовательной программы

## «Data Analyst»

Анализ данных — это не просто обработка информации после ее получения и сбора, это средство проверки гипотез. Цель любого анализа данных — понимание исследуемой ситуации целиком (выявление тенденций, в том числе негативных отклонений от плана, прогнозирование и получение рекомендации). Для достижения этой цели ставятся следующие задачи анализа данных: сбор информации, структуризация информации, выявление закономерностей, анализ, прогнозирование и получение рекомендаций,

Специалист, который делает такую работу, называется дата-аналитиком (или Data Analyst). Data Analyst работает с информационными массивами для извлечения сведений, ценных для бизнеса для принятия оптимальных управленческих решений.

Программа нацелена на тех, кто хочет разобраться в анализе данных, получить практический опыт решения большого количества реальных задач и заинтересован в начале карьеры в этом направлении. Для обучения на программе необходимо знать основы разделов высшей математики - дискретная математика, линейная алгебра и математический анализ.

Программа начинается с изучения программирования на языке Python, изучения языка программирования структурированных запросов SQL, основ математической статистики и теории вероятности — и переходит к изучению инструментов анализа и визуализации данных таких как Excel, Power BI, Tableau. При разработке программы делается акцент на практическую работу и проектную деятельность.

Для успешного завершения программы слушателю необходимо выполнить собственный проект, где он решит прикладную задачу с использованием реальных данных и изученных инструментов для анализа данных и визуализации

Целевая аудитория программы: ИТ-специалисты, желающие повысить свои компетенции в Анализе Данных, или специалисты без опыта, желающие переквалифицироваться в специалистов, занимающихся Анализом Данных со знанием основных на сегодняшний день инструментов и библиотек.

Задачами являются:

- совершенствование и (или) получение новых компетенций, необходимых для осуществления профессиональной деятельности;
- получение новой специализации в ИТ индустрии;
- ведение новой профессиональной деятельности с учетом новых технологий, подходов и способов мышления.

Программа построена по очному принципу с применением дистанционных технологий (с применением электронного обучения и дистанционных образовательных технологий) в формате вебинаров, которые займут 160 ак. ч.

Преподавателями программы будут являться преподаватели МФТИ, имеющие опыт преподавания и консультирования по заявленным модулям, а также опыт практической работы в отечественных и/или зарубежных организациях в сфере деятельности, совпадающей с направлением преподавания.

Взаимодействие со слушателями курса будет осуществляться дистанционно с помощью системы управления образовательным процессом (далее СДО – система дистанционного образования). Посредством СДО слушатели будут получать доступ к интерактивным лекциям в формате видеолекций. Контроль изучения этих материалов будет организован, как автоматически с помощью средств СДО, так и через тестовые задания. Дистанционный цикл будет проводиться одновременно для всех участников. Задания с открытым вариантом ответа будут проверяться преподавателями курса.

### **Составители программы:**

**Райгородский Андрей Михайлович**

Доктор физико-математических наук, директор ФПМИ МФТИ

**Благодарный Евгений Владимирович**

заведующий учебно-методической лабораторией инноватики

**Иванова Анастасия Сергеевна**

руководитель проектов учебно-методической лаборатории инноватики ФПМИ

Директор Физтех-школы ПМИ



А.М. Райгородский

## Квалификация и опыт привлекаемого преподавательского состава для реализации дополнительной образовательной программы «Data Analyst»

<p>Ф.И.О. лектора, год рождения</p>	<p>Информация об образовании, полученном в соответствии с образовательными программами высшего профессионального образования, дополнительного профессионального образования (в т.ч. о наличии званий и ученых степеней) и т.д.</p>	<p>Место работы, занимаемая должность в настоящий момент, общий трудовой стаж, педагогический стаж</p>	<p>Опыт преподавания и консультирования по предмету, сопоставляемому с направлением логот (перечислить), преподавательский стаж</p>	<p>Наличие опыта практической работы в отечественных и зарубежных организациях в сфере деятельности, совпадающей с направлением преподавания</p>
<p>Склонин Илья Андреевич</p>	<p>Аспирант кафедры Информатики и вычислительной математики МФТИ, выпускник кафедры теоретической и прикладной информатики МФТИ.</p>	<p>Преподаватель “Программирование на Python” и “Основные алгоритмы” в МФТИ</p>	<p>Преподаватель курсов «Программирование на Python» и «Основные алгоритмы» в МФТИ. Тема диссертации: «Исследование способов построения графовых ядер и эмбедингов для вычислительно эффективного сравнения снимков состояния операционных систем».</p>	<p>Работал в команде компьютерного зрения TCS Group</p>
<p>Иванц Гайк Ашотович</p>	<p>ИИНТБ Защита информации, Комплексная защита объектов информатизации</p>	<p>МФТИ Должность: Преподаватель Высшая школа информационных технологий и безопасности HaskerU Должность: Преподаватель СПАО «Ингосстрах» (проект «Ингоскод») Должность: преподаватель</p>	<p>Преподавательский стаж: Более четырёх лет преподавания в компаниях и проектах: - HaskerU - IThub - МФТИ - Сбербанк - MTS - Ингосстрах - Школа СТО яндекс</p>	<p>Более пяти лет опыта работы в сфере работы с данными Работа над КХД таких банков, как: - ВТБ - Райффайзен банк - Альфа банк</p>

<p>Дайняк Александр Борисович</p>	<p>Кандидат физико-математических наук (2009) по специальности 01.01.09 (дискретная математика и математическая кибернетика), выпускник МГУ имени М.В. Ломоносова, Школа Анализа Данных Яндекса</p>	<p>Доцент кафедры дискретной математики ФПМИ МФТИ.</p>	<p>Автор нескольких онлайн-курсов по дискретной математике и информатике. Преподавательский стаж: с 2010 года</p>	<p>Выпускник Школы Данных Анализа</p>
<p>Шарапов Роман Андреевич</p>	<p>МФТИ (2021), ФРТК/ФРКТ, компьютерная безопасность, тема дипломной работы «Защищенный программно-аппаратный комплекс управления криптоваллотными активами», Специалистит</p>	<p>SberDevices (ПАО Сбербанк), инженер-аналитик графа знаний Общий трудовой стаж: 4 года Педагогический стаж: 2 года</p>	<p>Преподавательский стаж: 2 года</p>	<p>2021-2022 в TalentTech на позиции аналитика, в ходе которой занимался построением дашбордов в Tableau и PowerBI, а также визуализацией в python и Excel</p>