Федеральное государственное бюджетное образовательное учреждение   
высшего образования

Московский государственный университет имени М. В. Ломоносова

Механико-математический факультет

|  |  |
| --- | --- |
|  | УТВЕРЖДАЮ  Декан механико-математического факультета МГУ  \_\_\_\_\_\_\_\_\_\_\_\_\_\_/А. И. Шафаревич /  «\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_2021 г. |
|  | М.П. |

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

|  |  |
| --- | --- |
| Наименование дисциплины (модуля): | Работа с данными в индустрии |
| Уровень высшего образования: | Бакалавриат, магистратура, специалитет |
| Направление подготовки / специальность: | Межфакультетский,по выбору студента |
| Направленность (профиль)/специализация ОПОП: | Междисциплинарный общеобразовательный |
| Форма обучения: | Очная |
| Язык преподавания: | Русский |
| Автор (авторы) программы: | Заплетин Максим Петрович,  Попов Олег Владимирович |

Рабочая программа рассмотрена и одобрена

*На заседании кафедры* математических и компьютерных методов анализа

(протокол № 7/22 от 24 ноября 2022 г.)

Москва 2021

Рабочая программа дисциплины разработана в соответствии с самостоятельно установленным МГУ образовательным стандартом (ОС МГУ) для реализуемых основных профессиональных образовательных программ высшего образования по направлению подготовки / специальности «\_\_\_\_\_\_\_\_\_\_\_\_\_» (*программы бакалавриата, магистратуры, реализуемых последовательно по схеме интегрированной подготовки; программы специалитета; программы магистратуры)* в редакции приказа МГУ от \_\_\_\_\_\_\_\_\_\_\_\_20\_\_\_ г.

**Содержимое**

[1. Место дисциплины (модуля) в структуре ОПОП ВО 3](#_Toc76748145)

[2. Объем дисциплины (модуля) 3](#_Toc76748146)

[3. Формат обучения 3](#_Toc76748147)

[4. Входные требования для освоения дисциплины (модуля) 3](#_Toc76748148)

[5. Результаты обучения по дисциплине (модулю) 3](#_Toc76748149)

[6. Содержание дисциплины (модуля) 4](#_Toc76748150)

[6.1. Структура дисциплины (модуля) 4](#_Toc76748151)

[7. Ресурсное обеспечение 4](#_Toc76748152)

[7.1. Список основной литературы 4](#_Toc76748153)

[7.2. Список дополнительной литературы (при наличии) 4](#_Toc76748154)

[7.3. Список программного обеспечения 4](#_Toc76748155)

[7.4. Список баз данных и информационных справочных систем 5](#_Toc76748156)

[7.5. Список ресурсов сети «Интернет» 5](#_Toc76748157)

[7.6. Материально-техническое обеспечение 5](#_Toc76748158)

[8. Фонд оценочных средств 5](#_Toc76748159)

[8.1. Текущий контроль успеваемости 5](#_Toc76748160)

[8.2. Промежуточная аттестация 5](#_Toc76748161)

1. Место дисциплины (модуля) в структуре ОПОП ВО

Является дисциплиной по выбору, избираемой в обязательном порядке.

1. Объем дисциплины (модуля)

Объем дисциплины (модуля) составляет 1 з.е., 36 академических часов, в том числе 24 академических часа, отведенных на контактную работу обучающихся с преподавателем и 12 академических часов на самостоятельную работу обучающихся.

1. Формат обучения

Дисциплина реализуется в очной форме с использованием электронного обучения

и дистанционных образовательных технологий (для выполнения домашних заданий).

1. Преподаватели

Дисциплину ведут доцент кафедры математических и компьютерных методов анализа Попов Олег Владимирович и доцент каф. общих проблем управления Заплетин Максим Петрович

1. Входные требования для освоения дисциплины (модуля)

Предварительные условия (если есть)

Для того чтобы формирование данной компетенции было возможно, обучающийся, приступивший к освоению программы, должен

Знать: основные направления, проблемы, теории и методы современной математики, элементарные навыки программирования на языке Python.

Уметь: решать стандартные задачи по программированию и применять идеи, использованные в их решениях, для решения аналогичных задач.

Владеть: основными понятиями и теоремами из этих разделов математики.

1. Результаты обучения по дисциплине (модулю)

|  |  |  |
| --- | --- | --- |
| **Компетенции выпускников** | **Индикаторы достижения компетенций, реализуемые в настоящей дисциплине (модуле)[[1]](#footnote-1)** | **Планируемые результаты обучения по дисциплине (модулю), соотнесенные с индикаторами достижения компетенций** |
| **Компетенция СПК-1**  Знание теоретических основ и умение применять основные навыки программирования на языке Python для фундаментальных и прикладных задач при работе с большими данными. | **Индикатор СПК-1.1**  Знает теоретические основы построения алгоритмов и использования компьютерных программ. | Знать основные термины, базовые элементы и синтаксис изучаемых языков программирования.  Знать и уметь использовать перечисленные инструменты при работе с большими данными. |
| **Индикатор СПК-1.2**  Умеет решать нестандартные задачи по программированию и применять идеи, использованные в их решениях, для решения аналогичных задач СУБД. | Уметь алгоритмизировать и решать задачи при работе с данными.  Владеть методами работы с СУБД. |
| *УК-1*  *УК-3* | З1 (УК-1) ЗНАТЬ:методы критического анализа и оценки современных научных достижений, а также методы генерирования новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях  У1 (УК-1) УМЕТЬ: анализировать альтернативные варианты решения исследовательских и практических задач и оценивать потенциальные выигрыши/проигрыши реализации этих вариантов  У2 (УК-1) УМЕТЬ: при решении исследовательских и практических задач генерировать новые идеи, поддающиеся операционализации исходя из наличных ресурсов и ограничений  З1 (УК-3) ЗНАТЬ: особенности представления результатов научной деятельности в устной и письменной форме при работе в российских и международных исследовательских коллективах  У1 (УК-3) УМЕТЬ: следовать нормам, принятым в научном общении при работе в российских и международных исследовательских коллективах с целью решения научных и научно-образовательных задач  У2 (УК-3) УМЕТЬ: осуществлять личностный выбор в процессе работы в российских и международных исследовательских коллективах, оценивать последствия принятого решения и нести за него ответственность перед собой, коллегами и обществом | |
| *ОПК-1* | З1 (ОПК-1) ЗНАТЬ: основные понятия, результаты и задачи фундаментальной математики и механики.  У1 (ОПК-1) УМЕТЬ: применять основные математические методы и алгоритмы для решения стандартных задач математики.  В1 (ОПК-1) ВЛАДЕТЬ: методами математического моделирования. | |
| *ПК-1* | В1 (ПК-1) ВЛАДЕТЬ:Современными навыками научно-исследовательской работы в области информационной безопасности  У1 (ПК-1) УМЕТЬ: Применять современные методы защиты информации в научно-исследовательской работе  З1 (ПК-1) ЗНАТЬ: наиболее успешные методы защиты информации  З (ПК-1)-2 ЗНАТЬ: Основные определения и формулировки наиболее важных результатов современных методов защиты информации | |

1. Содержание дисциплины (модуля)

Структура дисциплины (модуля) по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий (в строгом соответствии с учебным планом)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Наименование разделов и тем дисциплины (модуля),**  **Форма промежуточной аттестации по дисциплине (модулю)** | **Номинальные трудозатраты обучающегося** | | | | |
|  | **Контактная работа  (работа во взаимодействии с преподавателем)**  **Виды контактной работы, ак. ч.** | | | Самостоятельная работа,  ак. ч. |
| Всего,  ак. ч. | Ауд.,  ак. ч. | Лекции\*,  ак. ч. | Семинары\*, прак.,  ак. ч. |
| **Тема 1** Знакомство с SQL.  для чего нужен SQL и как писать запросы,  структура баз данных. Ключи, историчность  откуда берутся данные, источники, как применять SQL в бизнесе | 3 | 2 | 2 |  | 1 |
| **Тема 2** Python  как построить функцию, про библиотеки и работу с данными | 3 | 2 | 2 |  | 1 |
| **Тема 3** Визуализация данных.  для чего нужна визуализация, типы визуализации, хорошие и плохие примеры визуализаций | 3 | 2 | 2 |  | 1 |
| **Тема 4** Метрики общий контекст работы, количественные пользовательские метрики, относительные и сложные метрики, сегментация когорты, статика и динамика | 2 | 2 | 2 |  | 1 |
| **Тема 5** Основы математической статистики оценка параметров, предсказательные и доверительные интервалы, основы теории вероятностей и математической статистики | 3 | 2 | 2 |  | 1 |
| **Тема 6** Как проводить А/Б тесты: проверка гипотез  точки роста и гипотезы — как они появляются,  что такое АБ-тест и зачем он нужен,  статистическая и практическая значимость  практика на Python: как запускаются АБ-тесты | 3 | 2 | 2 |  | 1 |
| **Тема 7** Как проводить А/Б тесты: подготовка и проведение.  проверка статистических гипотез, основные понятия, критерии и способы сравнения критериев, поправка на множественные сравнения | 3 | 2 | 2 |  | 1 |
| **Тема 8** Как делать выводы из данных  кейсы с изменениями метрик. Поиск  причин изменения метрики, метрики продукта и метрики роста, смещенность выборки  что делать, если невозможно принять решение по цифрам, целеполагание | 3 | 2 | 2 |  | 1 |
| **Тема 9** Модели  Какие используют аналитики и зачем (с примерами)  Простейшие прогнозирующие  Выбор параметров модели | 3 | 2 | 2 |  | 1 |
| **Тема 10** Мобильная аналитика  введение и особенности данных, основные метрики, работа в Amplitude: юзер-кейсы, предиктивный интерфейс | 3 | 2 | 2 |  | 1 |
| **Тема 11** Регрессии  линейные и логистические регрессии  градиентный спуск и корреляция | 2 | 2 | 2 |  | 1 |
| **Тема 12** Случайный лес  введение в машинное обучение, что такое дерево решений, обучение деревьев — метрики качества  подготовка данных для обучения, примеры использования в Тинькофф, рассмотрение практических примеров в Python | 3 | 2 | 2 |  | 1 |
| Промежуточная аттестация (зачет(ы) и (или) экзамен(ы)) | 2 |  |  |  |  |
| **Итого** | *36* | *24* | *24* |  | *12* |

*Примечание: Формировать таблицу лучше сначала в xlsx файле с автоматическим подсчетом часов по темам.*

1. Ресурсное обеспечение
   1. Список основной литературы

Энтони Молинаро «SQL. Сборник рецептов» (2009)

Mark Lutz, David Ascher "Learning Python"

Kahled, SalehandShukairy, Ayat (2011). Conversion Optimization: The Art and Science of Converting Prospects into Customers, p. 2. O'ReillyMedia, Sebastopol. ISBN978-1-449-37756-4

DeClerck, JP. "Customer intelligence: why and how a data-driven business is a customer-centric business". Retrieved 3 April 2012.

О.О. Замков, А.В. Толстопятенко, Ю.Н. Черемных "Математические методы в экономике"

* 1. Список дополнительной литературы (при наличии)
  2. Список программного обеспечения

*WINDOWS, PYTHON, SQL*

* 1. Список баз данных и информационных справочных систем

Общие информационные ресурсы.

* 1. Список ресурсов сети «Интернет»

www.mathnet.ru, http://lib.mexmat.ru.

* 1. Материально-техническое обеспечение

Аудитория с доской или проектором;

Электронные презентации, программы курса и методического пособия;

Язык преподавания – русский язык.

1. Фонд оценочных средств

*Описание шкал оценивания*: зачтено/не зачтено

*Форма итоговой аттестации:* зачет, оценивание основывается на результатах выполненных домашних работ – тестов после лекций и домашних проектов.

* 1. Текущий контроль успеваемости

Типовые контрольные задания или иные материалы для проведения текущего контроля успеваемости, критерии и шкалы оценивания.

Вопросы к зачету:

1. Что такое Big Data.
2. Что такое machine learning.
3. Методы обработки данных.
4. Методы обработки больших данных.
5. Сходство и различие методов работы с данными и большими данными.
6. Базовые элементы и основные термины языка SQL.
7. Базовые элементы и основные термины языка Python.
8. Базовые элементы и основные термины теории вероятностей и математической статистики
9. Структуры БД. Способы предоставления информации. Диаграммы и графики.
10. Excel. Визуализация, работа с данными, ввод/вывод.
11. Тесты А-Б, статистические критерии, значимость.
12. Методы измерений, метрики.
13. Мониторинг изменений.
14. Модели вероятностные и финансовые.
15. Линейная и логистическая регрессии.
    1. Промежуточная аттестация

Зачет

1. В настоящем столбце должны быть указаны только те индикаторы достижения компетенций, которые связаны с данной дисциплиной (модулем) согласно таблице 4.1. Общей характеристики ОПОП. [↑](#footnote-ref-1)