

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ
М.В.ЛОМОНОСОВА»**

ЭКОНОМИЧЕСКИЙ ФАКУЛЬТЕТ

«УТВЕРЖДАЮ»

Декан экономического факультета МГУ
профессор. _____ А.А.Аузан

«» 2021 год

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Программирование, основы алгоритмов и анализ данных в Python и R

Уровень высшего образования
Магистратура

Направление подготовки (специальность)
38.04.01 Экономика

Направленность (профиль) ОПОП
Анализ данных в экономике

Форма обучения
очная

Рабочая программа рассмотрена и одобрена
Учебно-методической комиссией экономического факультета
(протокол № _____, дата)

Москва 2021

На оборотной стороне титульного листа указывается:

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 «_____» магистратуры

ОС МГУ утвержден решением Ученого совета МГУ имени М.В.Ломоносова от _____ 20 _____ года (протокол № ____).

Год (годы) приема на обучение: 2020 и последующие

1. Место и статус дисциплины в структуре основной профессиональной образовательной программы подготовки магистра

Статус дисциплины: *обязательная (дисциплина по выбору программы)*

Триместр: 1

2. Входные требования для освоения дисциплины

Для успешного освоения данного курса требуются знания и навыки, полученные в следующих дисциплинах:

- теория вероятностей и математическая статистика (в рамках программы вступительного испытания);
- основы эконометрики (в рамках программы вступительного испытания).

3. Планируемые результаты обучения по дисциплине, соотнесенные с требуемыми компетенциями выпускников

Компетенции выпускников (коды)	Индикаторы (показатели) достижения компетенций	Планируемые результаты обучения по дисциплине (модулю), сопряженные с компетенциями
УК-1. Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий, формулировать научно обоснованные гипотезы, применять методологию научного познания в профессиональной деятельности	УК-1.И-1. Анализирует проблемную ситуацию как целостную систему, выявляя ее составляющие и связи между ними	УК-1.И-1.У-2. Умеет осуществлять сбор, систематизацию и критический анализ информации, необходимой для выработки стратегии действий по разрешению проблемной ситуации
ОПК-2. Способен применять продвинутое инструментальные методы экономического анализа в прикладных и (или) фундаментальных исследованиях.	ОПК-2.И-2. Обрабатывает статистическую информацию и получает статистически обоснованные выводы	ОПК-2.И-2.3-1. Знает основные виды статистической информации ОПК-2.И-2.У-1. Умеет обрабатывать статистическую информацию и визуализировать данные
ОПК-5. Способен использовать современные информационные технологии и программные средства при решении профессиональных задач	ОПК-5.И-1. Применяет общие или специализированные пакеты прикладных программ, предназначенные для обработки, визуализации и анализа данных, в том числе эконометрического анализа и имитационного моделирования	ОПК-5.И-1.3-1. Знает общие и специализированные пакеты прикладных программ, предназначенные для обработки, визуализации и анализа данных, в том числе эконометрического анализа и имитационного моделирования, в соответствии со своими профессиональными задачами

		ОПК-5.И-1.У-1. Умеет применять общие или специализированные пакеты прикладных программ (MS Excel, Eviews, Stata, SPSS, AnyLogic, Tableau и др.) или языки программирования (R, Python и др.), предназначенные для обработки, визуализации и анализа данных, в том числе эконометрического анализа и имитационного моделирования в, в соответствии со своими профессиональными задачами
ПК-8. Способен использовать различные инструментальные методы расчета и анализа социально-экономических показателей	ПК-8.И-2. Применяет современные инструментальные методы расчета и анализа социально-экономических показателей при решении практических и (или) исследовательских задач	ПК-8.И-2.У-1. Умеет применять инструментальные методы расчета и анализа показателей, в том числе с использованием различных платформ анализа данных и с применением языков программирования R и Python
МПК-3. Способен использовать соответствующий математический аппарат и инструментальные средства для сбора, обработки, анализа и систематизации информации	МПК-3.И-1. Использует современные программные и аналитические средства для сбора, первичной обработки и анализа данных	МПК-3.И-1.3-1. Знает современные программные и аналитические способы сбора, анализа и систематизации информации. МПК-3.И-1.У-1. Умеет применять современные программные и аналитические способы сбора, анализа и систематизации информации

4. Объем дисциплины по видам занятий

Объем дисциплины составляет 6 зачетных единицы: 216 академических часов, из которых 52 академических часов составляет аудиторная нагрузка, из них 26 академических часов — лекции и 26 академических часов — семинары, 52 академических часа — групповая контактная работа, 0 академических часов — индивидуальная контактная работа, 108 академических часов составляет самостоятельная работа магистранта.

5. Формат обучения: очный, используется электронная информационная среда экономического факультета МГУ имени М.В. Ломоносова «ON.ECON»..

6. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических часов и виды учебных занятий

Название раздела/темы	Всего, часы	В том числе			
		Контактная работа с преподавателем			Самостоятельная работа магистранта, часы
		Лекции, часы	Семинары, часы	Групповая, часы	
Тема 1. Язык программирования Python	64	10	10	20	24
Тема 2. Разработка и анализ алгоритмов	60	10	10	20	20
Тема 3. Углубленный Python	18	4	4	2	8
Тема 4. Анализ данных на Python	42	8	8	16	10
Тема 5. Основы языка R	16	4	4	4	4
Текущая аттестация: — <i>домашние задания</i>	26	—	—	—	26
Текущая аттестация: — <i>итоговый проект</i>	36	—	—	10	26
Промежуточная аттестация: <i>экзамен</i>	26	-	-	—	26
Всего	288	36	36	72	144

Краткое содержание тем дисциплины

Тема 1. Язык программирования Python

Обзор основных структур данных, используемых при программировании. Формы представления алгоритмов. Методы тестирования программ. Абстрактные типы данных. Этапы решения прикладных задач с использованием ЭВМ. Переменные. Условные операторы и циклы. Строки. Списки и кортежи. Словари. Функции. Работа с файлами. Основы объектно-ориентированного программирования.

Тема 2. Разработка и анализ алгоритмов

Список рассматриваемых тем:

1. Введение в алгоритмы, алгоритм Карацубы
2. Асимптотический анализ алгоритмов, алгоритмы сортировки
3. Рекурсивные алгоритмы, Мастер-теорема рекурсии
4. Алгоритм выбора
5. Рандомизированные алгоритмы и быстрая сортировка, $O(n)$ сортировка
6. Бинарное дерево поиска, сбалансированное дерево поиска
7. Графы, поиск в глубину и поиск в ширину
8. Хэширование
9. Динамическое программирование

В этой теме рассматриваются базовые алгоритмы, которые составляют базу для изучения любого языка программирования. Все примеры и упражнения приведены на языке Python.

Тема 3. Углубленный Python

Изучение GIL концепции в python. Многопоточность и многопроцессорность для решения задач обработки и анализа данных.

Тема 4. Анализ данных на Python

Список рассматриваемых тем:

1. Работа со неструктурированными данными
2. Работа со структурированными данными (HTML, XML, JSON, CSV, H5, и другие)
3. Работа с библиотеками Numpy, Scipy, Pandas, Matplotlib, Plotly, Seaborn
4. Библиотеки машинного обучения Sklearn, Pytorch
5. Менеджер данных DVC

В теме рассматриваются основные библиотеки, используемые для научных вычислений и визуализации результатов. А также версионирование данных для моделей машинного обучения.

Тема 5. Основы R

Типы данных. Векторы и матрицы. Списки и таблицы данных, сортировка и объединение таблиц. Импорт данных в R, временные ряды. Функции, условия, ветвления, циклы.

Векторные вычисления.

Литература:

1. Мастицкий С. Э., Шитиков В. К. Статистический анализ и визуализация данных с помощью R. – М.: ДМК Пресс, 2015. – 496 с.
2. Fischetti T. Data Analysis with R. – Birmingham, UK: Packt Publishing, 2015. – 390 p.

7. Фонд оценочных средств для оценивания результатов обучения по дисциплине

Шкала оценивания результатов (баллы) по дисциплине:

Результаты обучения по дисциплине	Виды оценочных средств
УК-1.И-1.У-2. Умеет осуществлять сбор, систематизацию и критический анализ информации, необходимой для выработки стратегии действий по разрешению проблемной ситуации	Домашние работы Итоговое проектное задание
ОПК-2.И-2.3-1. Знает основные виды статистической информации	Домашние работы Итоговое проектное задание
ОПК-2.И-2.У-1. Умеет обрабатывать статистическую информацию и визуализировать данные	Домашние работы Итоговое проектное задание
ОПК-5.И-1.3-1. Знает общие и специализированные пакеты прикладных программ, предназначенные для обработки, визуализации и анализа данных, в том числе эконометрического анализа и имитационного моделирования, в соответствии со своими	Домашние работы Итоговое проектное задание Экзамен

профессиональными задачами	
ОПК-5.И-1.У-1. Умеет применять общие или специализированные пакеты прикладных программ (MS Excel, Eviews, Stata, SPSS, AnyLogic, Tableau и др.) или языки программирования (R, Python и др.), предназначенные для обработки, визуализации и анализа данных, в том числе эконометрического анализа и имитационного моделирования в, в соответствии со своими профессиональными задачами	Домашние работы Итоговое проектное задание
ПК-8.И-2.У-1. Умеет применять инструментальные методы расчета и анализа показателей, в том числе с использованием различных платформ анализа данных и с применением языков программирования R и Python	Домашние работы Итоговое проектное задание
МПК-3.И-1.3-1. Знает современные программные и аналитические способы сбора, анализа и систематизации информации.	Домашние работы Итоговое проектное задание
МПК-3.И-1.У-1. Умеет применять современные программные и аналитические способы сбора, анализа и систематизации информации	Домашние работы Итоговое проектное задание

Виды оценочных средств	Баллы
Домашние работы (10 шт.)	100
Итоговое проектное задание	140
Экзамен	60

Оценка по дисциплине выставляется, исходя из следующих критериев:

Оценка	Минимальное количество баллов	Максимальное количество баллов
<i>Отлично</i>	255	300
<i>Хорошо</i>	195	254
<i>Удовлетворительно</i>	120	194
<i>Неудовлетворительно</i>	—	119

Примечание: в случае, если магистрант за триместр набирает менее 20% баллов от максимального количества по дисциплине, то уже на промежуточном контроле (и далее на пересдачах) действует следующее правило сдачи: «магистрант может получить только оценку «Удовлетворительно», и только если получит за промежуточный контроль, включающий весь материал дисциплины, не менее, чем 85% от баллов за промежуточный контроль».

Типовые задания, методические рекомендации по их подготовке и требования к их выполнению:

Домашние работы представлены в количестве 10 штук, их выполнение позволит набрать до 100 баллов. Домашние работы представляют собой практические задания, ориентированные на закрепление навыков программирования и анализа данных на языках Python в рамках различных тем. Задание сдается в формате .py/ipy/nb.pdf в виде аналитической записки или исполняемого кода. В тексте следует описать используемые для ответа на поставленный

вопрос данные, подробно обосновать выбор языков конструкций языка программирования, представить результаты расчётов (при необходимости используя таблицы и рисунки).

В качестве домашних заданий выдаются задачи, позволяющие отработать темы, рассмотренные в рамках аудиторных часов. Например:

1. Реализовать чтение данных (в частности, котировок акций или сведений о стоимости иных финансовых инструментов) из файла на диске и реализовать его преобразование в структуру данных на языке Python (например, список).
2. Реализовать алгоритм (например, бинарное дерево поиска), который решает поставленную задачу (например, поиск минимального элемента).
3. Реализовать несколько алгоритмов сортировки, оценить сложность полученных алгоритмов.
4. Реализовать многопроцессорное вычисление предложенной задачи (например, числа π).
5. Выполнить описательный анализ данных, используя инструменты Python.

В качестве **итогового проекта** используется комплексное задание, которое включает в себя:

1. самостоятельный поиск студентом источников данных (например, временных рядов) их обработка для последующего анализа;
2. анализ этих данных с помощью инструментов Python: выявление статистических закономерностей с помощью проверки гипотез, определение сезонности, аппроксимация регрессионными моделями, дисперсионный анализ и т.п.

В зависимости от уровня реализации проекта и широты спектра используемых интересов, обучающиеся могут получить до 140 баллов.

На **экзамене** проверяется владение теоретическими основами курса, понимание языковых структур и обоснованность выбора инструментов для решения тех или иных типов прикладных задач (до 60 баллов по сумме).

8. Ресурсное обеспечение

8.1. Перечень основной и дополнительной литературы

Основная литература:

1. Fischetti T. Data Analysis with R. – Birmingham, UK: Packt Publishing, 2015. – 390 p.
2. James G., Witten D., Hastie T., Tibshirani R. An Introduction to Statistical Learning with Applications in R. – New York: Springer, 2017. – 434 p.
3. Idris I. Python data analysis. – Birmingham, UK: Packt Publishing, 2014. – 348 p.
4. Васильев А. Н. Python на примерах. Практический курс по программированию. – СПб.: Наука и Техника, 2016. – 432 с.
5. Горшенин А.К. Методы реализации абстрактных типов данных. Учебное пособие. – М.: ФИЦ ИУ РАН, 2016. – 100 с.
6. Горшенин А.К. Принципы разработки программного обеспечения и основы языка C. Учебное пособие. – М.: ФИЦ ИУ РАН, 2016. – 100 с.
7. Лутц М. Изучаем Python. – СПб.: Символ-Плюс, 2011. – 1280 с.
8. Маккинни У. Python и анализ данных. – М.: ДМК Пресс, 2015. – 482 с.
9. Мастицкий С. Э., Шитиков В. К. Статистический анализ и визуализация

данных с помощью R. – М.: ДМК Пресс, 2015. – 496 с.

10. Мэтиз Э. Изучаем Python. Программирование игр, визуализация данных, веб-приложения. – СПб.: Питер, 2017. – 496 с.

8.2. Перечень лицензионного программного обеспечения

1. Среда для языка Python;
2. Продукты для создания отчетов в форматах DOCX и/или PDF.

8.3. Перечень профессиональных баз данных и информационных справочных систем

8.4. Перечень ресурсов информационно-телекоммуникационной сети «Интернет» (при необходимости)

8.5. Описание материально-технической базы

Для организации занятий по дисциплине необходимы следующие технические средства обучения: компьютерный класс с проектором.

1. **Язык преподавания:** русский.
2. **Преподаватель (преподаватели):** Кафтанов И. А., Руденко Д. А.
3. **Автор (авторы) программы:** Кафтанов И. А., Руденко Д. А.