

**Федеральное государственное бюджетное образовательное
учреждение высшего образования
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

ИНСТИТУТ ЭКОНОМИКИ, МАТЕМАТИКИ И ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ
кафедра Эконометрики и математической экономики

УТВЕРЖДЕНА

на заседании кафедры Эконометрики и
математической экономики

Протокол от «02» июня 2021 г. № 10

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Индекс ФТД.02 «Введение в анализ данных с Python»

по направлению подготовки 38.04.01 «Экономика»

направленность «Экономика и Финансы»

квалификация магистр

очная форма обучения

Год набора - 2021

Москва, 2021 г.

Автор—составитель:

д.т.н. проф. кафедры эконометрики и математической экономики Шилин К.Ю.

Заведующий кафедрой

эконометрики и математической экономики, к. ф.-м. н, Носко В.П.

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине соотнесенных с планируемыми результатами освоения образовательной программы.....	4
2. Объем и место дисциплины в структуре образовательной программы.....	4
3. Содержание и структура дисциплины.....	4
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине.....	5
5. Методические указания для обучающихся по освоению дисциплины.....	8
6. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет», учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине	9
6.1. Основная литература	9
6.2. Дополнительная литература	9
6.3. Учебно-методическое обеспечение самостоятельной работы	10
6.4. Нормативные правовые документы	10
6.5. Интернет-ресурсы	10
6.6. Иные источники	10
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	10

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения программы

1.1. Дисциплина обеспечивает овладение следующими компетенциями:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1	способен обобщать и критически оценивать НИР и на их основе составлять план НИР, а также подобрать необходимые инструментальные методы
УК-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.2	способен самосовершенствоваться свои знания и умения

1.2. В результате освоения дисциплины у студентов должны быть сформированы:

ОТФ/ТФ профессиональные действия	Код этапа освоения компетенции	Результаты обучения
ведение аналитической работы в области экономики и финансов	УК-1.1	на уровне знаний: предметной области; критического анализа; на уровне умений: подобрать необходимые инструментальные методы; применять инструментальные методы для проведения НИР;
ведение аналитической работы в области экономики и финансов	УК-6.2	на уровне знаний: методов самосовершенствования на уровне умений: корректного распределения приоритетов.

2. Объем и место дисциплины в структуре ОП ВО

Объем дисциплины

3 ЗЕ, 48 ак. часов на контактную работу с преподавателем, 56 ак. часов на самостоятельную работу обучающихся; 4 ак. часа. На промежуточную аттестацию;

Место дисциплины в структуре ОП ВО

- ФТД.В.02 «Введение в анализ данных с Python» 1 курс, 1 семестр
 - дисциплина реализуется после изучения дисциплин:
экономическая информатика (в объеме бакалавриата)
математический анализ (в объеме бакалавриата)
алгебра (в объеме бакалавриата)
 - дисциплина может реализоваться частично или полностью с применением ЭО и/или ДОТ.
- Учебные материалы дисциплины размещаются по адресу lms.ranepa.ru
- форма промежуточной аттестации – зачет с оценкой.

3. Содержание и структура дисциплины

№ п/п	Наименование тем (разделов)	Объем дисциплины (модуля), ак. час./ час.						Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий**				СР	
			Л	ЛР	ПЗ	КСР		
Очная форма обучения								
Тема 1	Основы языка программирования Python	8	2		2		4	ДЗ
Тема 2	Основы NumPy. Вычисления с векторами и матрицами (массивами)	10	2		2		6	ДЗ
Тема 3	Введение в Pandas	12	2		4		6	ДЗ
Тема 4	Очистка и подготовка данных	12	2		4		6	ДЗ
Тема 5	Объединение, комбинация и изменение размера данных	14	2		4		8	
Тема 6	Визуализация	14	2		4		8	
Тема 7	Агрегирование и группировки	12	2		4		6	
Тема 8	Временные ряды	12	2		4		6	
Тема 9	Примеры анализа данных	12			4		6	
Промежуточная аттестация		4						Зачет с оценкой
Всего:		108/81	16/12		32/24		56/42	

Примечание* – формы текущего контроля успеваемости: домашнее задание (ДЗ)

Примечание **: в рамках указанной контактной работы с обучающимися учебные занятия могут проводиться с использованием ДОТ и/или ЭО

Содержание дисциплины

Тема 1. Основы языка программирования Python

Дистрибутив Anaconda. Jupyter Notebooks. Основы языка программирования. Структуры данных, функции, файлы.

Тема 2. Основы NumPy. Вычисления с векторами и матрицами (массивами)

Объекты NumPy ndarray. Основные функции. Векторное программирование. Математические и статистические методы в NumPy. Булевы операции. Сортировка. Поиск уникальных имен. Линейная алгебра. Генератор «случайных» чисел.

Тема 3. Введение в Pandas.

Структурированные данные: Series, DataFrame. Индексация. Основные функции. Расчет статистик таблиц. Импорт текстовых данных, таблиц csv, таблиц Excel.

Тема 4. Очистка и подготовка данных

Обработка отсутствующих данных. Преобразование данных: поиск дубликатов, преобразование по условию, замена значений по условию. Обнаружение и фильтрация выбросов. Дамми. Манипуляции со столбцами. Векторизация строковых данных.

Тема 5. Объединение, комбинация и изменение размера данных

Иерархическое индексирование. Переконфигурирование данных с сортировкой. Вычисление статистик. Объединение наборов данных. Изменение размерности данных, сводные таблицы.

Тема 6. Визуализация

Визуализация matplotlib. Визуализация seaborn.

Иерархическое индексирование. Переконфигурирование данных с сортировкой. Вычисление статистик. Объединение наборов данных. Изменение размерности данных, сводные таблицы.

Тема 7. Агрегирование и группировки

Иерархическое индексирование. Переконфигурирование данных с сортировкой. Вычисление статистик. Объединение наборов данных. Изменение размерности данных, сводные таблицы.

Тема 8. Временные ряды

Типы данных: время, даты. Конвертация времени и дат. Временной ряд. Диапазоны дат, частотность и смещение. Периоды, расчеты с периодами. Смена размерности и частотности временного ряда. Подвижное окно выбора.

Тема 9. Примеры анализа данных

Анализ имен новорожденных. Анализ результатов выборов.

4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине

4.1. Формы и методы текущего контроля успеваемости обучающихся и промежуточной аттестации.

4.1.1. В ходе реализации дисциплины ФТД.02 «Введение в анализ данных с Python» используются следующие методы текущего контроля успеваемости обучающихся:

Тема (раздел)	Методы текущего контроля успеваемости
Тема 1	Домашнее задание 1
Тема 2	Домашнее задание 2
Тема 3	Домашнее задание 3
Тема 4	Домашнее задание 4
Тема 5	Домашнее задание 5
Тема 6	Домашнее задание 6
Тема 7	Домашнее задание 7
Тема 8	Домашнее задание 8
Тема 9	Домашнее задание 9

4.1.2. Зачет проводится с применением следующих методов (средств): в виде устного собеседования по результатам выполнения заданий текущего контроля успеваемости.

4.2. Материалы текущего контроля успеваемости обучающихся

Типовые оценочные материалы по теме 1

Домашнее задание 1. Установка дистрибутива Anaconda на персональном компьютере, обновление пакет входящих в дистрибутив

Типовые оценочные материалы по теме 2

Домашнее задание 2. Вычисление определителя матрицы, нахождение корреляционной матрицы.

Типовые оценочные материалы по теме 3

Домашнее задание 3. Импорт данных из Microsoft Excel. Импорт временного ряда с сайта Yahoo Finance.

Типовые оценочные материалы по теме 4

Домашнее задание 4. Работа с данными <https://openflights.org/data.html>. Найти пропущенные данные, найти уникальные имена пропущенных данных.

Типовые оценочные материалы по теме 5

Домашнее задание 5. Объединение данных <https://openflights.org/data.html>. Отбор данных по условию.

Типовые оценочные материалы по теме 6

Домашнее задание 6. Визуализация данных <https://openflights.org/data.html> на карте мира. Визуализация статистических данных.

Типовые оценочные материалы по теме 7

Домашнее задание 7. Агрегирование данных по странам и авиакомпаниям <https://openflights.org/data.html>.

Типовые оценочные материалы по теме 8

Домашнее задание 8. Выгрузка котировок акций с Google Finance. Расчет кумулятивных процентов. Расчет 200-дневной скользящей.

Типовые оценочные материалы по теме 9

Домашнее задание 9. Анализ датасета Ирисы.

4.3. Оценочные средства для промежуточной аттестации.

4.3.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы. Показатели и критерии оценивания компетенций с учетом этапа их формирования

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
УК-1	Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий	УК-1.1	способен обобщать и критически оценивать НИР и на их основе составлять план НИР, а также подобрать необходимые инструментальные методы
УК-6	Способен определять и реализовывать приоритеты собственной деятельности и способы ее совершенствования на основе самооценки	УК-6.2	способен самосовершенствоваться свои знания и умения

Этап освоения компетенции	Показатель оценивания	Критерий оценивания
УК-1.1	способность составить план НИР на основе анализа предметной области и подобрать инструменты	доклад научному руководителю
УК-6.2	способен самосовершенствоваться свои знания и умения	уровень профессионального роста

4.3.2 Типовые оценочные средства

Используя датасет

<https://raw.githubusercontent.com/jpatokal/openflights/master/data/airports.dat>

и датасет

<https://raw.githubusercontent.com/jpatokal/openflights/master/data/routes.dat>

найти 5 крупнейших аэропортов в странах Евросоюза по количеству вылетов/прилетов авиакомпаний.

Шкала оценивания.

Оценка определяется по формуле:

$$I = \frac{\sum_{i=1}^n B_i}{n} \quad (1)$$

где n – домашних заданий, B_i – оценка за домашнее задание по 10-ти бальной шкале, I – итоговое количество баллов.

10- бальная шкала	Традиционная шкала	«Зачтено»/ «Не зачтено»	Определение
10	Отлично	Зачтено	Поставленная задача решена верно, код оптимален и работоспособен. Анализ решения проведен в полном объеме.
9	Отлично	Зачтено	Поставленная задача решена верно, код частично оптимален и работоспособен. Анализ решения проведен в полном объеме.
8	Отлично	Зачтено	Поставленная задача решена верно, код работоспособен. Анализ решения проведен в полном объеме.
7	Хорошо	Зачтено	Поставленная задача решена верно, код работоспособен. Анализ решения проведен в достаточном объеме.
6	Хорошо	Зачтено	Поставленная задача решена верно, код работоспособен. Проведен анализ решения.
5	Удовлетворительно	Зачтено	Поставленная задача решена верно, код работоспособен. Проведен неполный анализ решения.
4	Удовлетворительно	Зачтено	Поставленная задача решена верно, код работоспособен. Проведен неполный анализ решения с частичными ошибками.
3	Неудовлетворительно	Не зачтено	Поставленная задача решена с неверным ответом.
2	Неудовлетворительно	Не зачтено	Поставленная задача не решена, представленный студентом код частично работоспособен
1	Неудовлетворительно	Не зачтено	Поставленная задача не решена, представленный студентом код не работоспособен.
0	Неудовлетворительно	Не зачтено	Поставленная задача не выполнена, код не представлен

4.4. Методические материалы

Зачет с оценкой проводится в компьютерной аудитории, с установленным дистрибутивом Anaconda, с доступом в локальную сеть Академии и глобальную сеть Интернет. Собеседование проводится индивидуально с каждым студентом. В ходе собеседования подводятся итоги успеваемости студента в соответствии с формулой (1). В том случае, если полученная оценка устраивает студента данная оценка является итоговой. В том случае, если студент не согласен с оценкой, студент выполняет предложенное задание, после решения которого, со студентом проводится собеседование. Оценка за задание выставляется согласно приведённой выше шкале. Итоговая оценка выставляется, как среднее арифметическое от суммы набранного количества баллов с округлением до целого в меньшую сторону.

При условии набора 4 и более итоговых баллов, компетенция считается сформированной в части указанной в п. 1, по данной дисциплине выставляется оценка. В ином случае, компетенция считается несформированной, по данной дисциплине выставляется оценка «неудовлетворительно».

В случае, если дисциплина полностью или частично проводилась с применением технологий электронного обучения и/или дистанционных технологий, зачет может производиться с использованием системы СДО Академии и применением прокторинга.

5. Методические указания для обучающихся по освоению дисциплины

Домашние задания для самостоятельной подготовки могут быть опубликованы:

1. В блоге преподавателя на сайте economy.ranepa.ru;

2. В группе ВКонтакте;
3. Высланы на общую почту группы

Конкретный способ коммуникации со студентами определяется преподавателем.

Для подготовки к практическим занятиям необходимо ознакомиться с рекомендованной литературой, повторить материал предыдущих практических занятий. При возникновении вопросов обратиться к преподавателю по электронной почте с указанием конкретной проблемы и (или) прийти к преподавателю на консультацию в установленное время.

Для подготовки письменных домашних заданий необходимо ознакомиться с рекомендованной литературой, повторить материал предыдущих практических занятий.

В течение курса преподаватель вправе предлагать студентам дополнительные задания повышенной сложности для начисления дополнительных баллов. Правила выполнения данных заданий и начисления баллов объявляются преподавателем индивидуально для каждого задания повышенной сложности.

Самостоятельная работа студентом осуществляется для закрепления изученного материала после практических занятий, для выполнения домашних заданий, для изучения дополнительных материалов.

№ п/п	Тип занятия	Указания
Тема 1. Основы языка программирования Python		
1	Л/ПЗ	Установка Anaconda на домашний компьютер. Прочитать главы 1-2 (Wes MCKinney 2018). Прочитать и разобрать главу 3 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 2. Основы NumPy. Вычисления с векторами и матрицами (массивами)		
2	Л/ПЗ	Прочитать и разобрать главу 4 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 3. Введение в Pandas		
3	Л/ПЗ	Прочитать и разобрать главы 5-6 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 4. Очистка и подготовка данных		
4	Л/ПЗ	Прочитать и разобрать главы 5-6 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 5. Объединение, комбинация и изменение размера данных		
5	Л/ПЗ	Прочитать и разобрать главы 7-8 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 6 Визуализация.		
6	Л/ПЗ	Прочитать и разобрать главу 9 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 7. Агрегирование и группировки.		
7	Л/ПЗ	Прочитать и разобрать главу 10 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 8. Временные ряды.		
8	Л/ПЗ	Прочитать и разобрать главу 11 (Wes MCKinney 2018). Выполнить домашнее задание.
Тема 9. Примеры анализа данных.		
9	Л/ПЗ	Прочитать и разобрать главу 14 (Wes MCKinney 2018). Выполнить домашнее задание.

6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

6.1. Основная литература.

1. Wes MCKinney. Python for Data Analysis. 2nd edition. 2018

6.2. Дополнительная литература.

2. Уэс Маккинни. Python и анализ данных. 2015.

6.3. Учебно-методическое обеспечение самостоятельной работы.

Отдельное обеспечение не предусмотрено.

6.4. Нормативные правовые документы.

Не предусмотрены.

6.5. Интернет-ресурсы.

<https://www.anaconda.com/download/>

<https://pandas.pydata.org>

<https://seaborn.pydata.org>

<https://matplotlib.org>

<http://lms.ranepa.ru>

6.6. Иные источники.

Не предусмотрены.

**7. Материально-техническая база, информационные технологии,
программное обеспечение и информационные справочные системы**

1. Компьютерный класс
2. Доступ в интернет и локальную сеть Академии
3. Проекционное оборудование
4. Программное обеспечение:
 - Windows/Linux/Mac OS
 - Дистрибутив Anaconda