

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ ИМЕНИ Г.В. ПЛЕХАНОВА»

УТВЕРЖДАЮ

Председатель комиссии по
дополнительному образованию




В.Г. Минашкин

(Решение Комиссии по дополнительному
образованию

от «07» 2022 г., протокол № 02)

ВЫСШАЯ ШКОЛА КИБЕРТЕХНОЛОГИЙ, МАТЕМАТИКИ И СТАТИСТИКИ

ДОПОЛНИТЕЛЬНАЯ ПРОФЕССИОНАЛЬНАЯ ПРОГРАММА
ПРОФЕССИОНАЛЬНОЙ ПЕРЕПОДГОТОВКИ

«АНАЛИЗ ДАННЫХ И МАШИННОЕ ОБУЧЕНИЕ В СРЕДЕ PYTHON»

МОСКВА, 2022

1. ОБЩАЯ ХАРАКТЕРИСТИКА ПРОГРАММЫ

1.1. Цель реализации программы.

Целью реализации программы «Анализ данных и машинное обучение в среде Python» является получение слушателями необходимых цифровых компетенций в области анализа больших данных для решения бизнес-задач компаний на стратегическом уровне управления, получение навыков использования языка разработки Python 3 для применения в области принятия управленческих решений в современных компаниях в условиях цифровой трансформации.

1.2. Характеристика нового вида профессиональной деятельности, новой квалификации

Программа профессиональной переподготовки «Анализ данных и машинное обучение в среде Python» разработана на основании ФГОС ВО 38.03.05 «Бизнес-информатика», ФГОС ВО «02.03.03 Математическое обеспечение и администрирование информационных систем», ФГОС ВО «09.03.02. Информационные системы и технологии»

Область профессиональной деятельности включает: связь, информационные и коммуникационные технологии (в сфере исследования, разработки, внедрения и сопровождения информационных технологий и систем).

Выпускники могут осуществлять профессиональную деятельность в других областях профессиональной деятельности и (или) сферах профессиональной деятельности при условии соответствия уровня их образования и полученных компетенций требованиям к квалификации работника.

Объектами профессиональной деятельности являются:

- информационные системы и технологии
- базы данных и хранилища информации
- проекты в области информационных технологий
- языки программирования, алгоритмы, библиотеки и пакеты программ, продукты системного и прикладного программного обеспечения;
- системное и прикладное программное обеспечение;
- системы управления предприятием;

Профильность программы направлена на решение профессиональных задач, связанных с исследованием, моделированием, наукоемким сопровождением разработки, включая сбор и анализ данных; а также с углубленным изучением вопросов, связанных с построением нейронных сетей.

Виды профессиональной деятельности:

- Аналитическая
- Организационно-управленческая
- Научно-исследовательская

Профессиональные задачи:

Слушатель, успешно завершивший обучение по данной программе, должен решать следующие профессиональные задачи в соответствии с видами профессиональной деятельности:

аналитическая деятельность:

анализ архитектуры предприятия, исследование и анализ рынка ИС и ИКТ, анализ и оценка применения ИС и ИКТ для управления бизнесом, анализ инноваций в экономике, управлении и ИТ, сбор, обработка и анализ информации о факторах внешней и внутренней среды проектов и организации в целом для принятия управленческих решений

организационно-управленческая деятельность:

- соблюдение кодекса профессиональной этики;

- планирование процессов и ресурсов для решения бизнес - задач в области экономики;

- разработка методов и механизмов мониторинга и оценки качества процессов производственной деятельности, связанной с созданием и использованием информационных систем.

научно-исследовательская деятельность:

поиск, сбор, обработка, анализ и систематизация данных в сфере экономики, подготовка визуализации данных и научных публикаций, разработка бизнес-планов на основе анализа данных, создание новых бизнесов на основе нейронных сетей.

В рамках данной программы подразумевается изучение принципов разработки моделей машинного обучения с использованием библиотек Python 3 для осуществления поддержки принятия управленческих решений в компаниях. Программа включает в себя изучение основных языковых конструкций языка Python 3, а также получения практических навыков работы с библиотеками: NumPy, Pandas, Scikit-learn, Keras и TensorFlow. Также в программе рассматриваются библиотеки визуализации данных Seaborn и Matplotlib. Слушатели получат навыки работы со средой разработки Spyder, а также с интерактивной облачной средой Google Colab. Слушатели будут ознакомлены с принципами стандарта CRISP-DM для понимания жизненного цикла исследования данных.

Для оценки уровня освоения полученных в процессе обучения компетенций слушатели будут выполнять проекты по экономической тематике с использованием языка Python, состоящие из следующих этапов (согласно целям стандарта CRISP-DM):

1. Постановка бизнес-задачи (определение бизнес-целей, оценка текущей ситуации, определение целей аналитики, подготовка плана анализа данных);
2. Начальное изучение данных (сбор данных, описание данных, изучение данных проверка качества данных);
3. Подготовка данных (Очистка данных, генерация данных, интеграция данных, форматирование данных)
4. Моделирование (выбор алгоритмов, подготовка плана тестирования, обучение моделей, проверка качества моделей)
5. Оценка (оценка результатов, оценка процесса, определение следующих шагов)

Внедрение (планирование мониторинга и поддержки, подготовка отчета, ревью проекта)

В результате освоения программы и в случае успешной сдачи итоговой аттестации слушателю выдается диплом о профессиональной переподготовке в соответствии с федеральным государственным образовательным стандартом, утвержденным приказом Министерства образования и науки Российской Федерации от 29.07.2020 г. № 838, с присвоением квалификации «Специалист по анализу данных и программированию».

1.3. Планируемые результаты обучения

В результате освоения данной программы слушатель должен освоить следующие компетенции:

1. Применяет языки программирования для решения профессиональных задач

Знает:

- Основы программирования на языке Python 3.
- Среды разработки программных приложений с использованием языка программирования Python 3.

Умеет: использовать среды программирования на Python 3, ориентированные на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных.

Владеет: навыками разработки программных приложений с использованием сред программирования на языке Python 3.

2. Применяет форматы обмена данными и языки разметки

Знает:

- Набор инструментов для загрузки данных из различных форматов, а также осуществления создания исходного набора данных для анализа (встроенные средства работы с файлами csv, json языка python, а также инструментарий библиотеки Pandas с использованием объектов DataFrame и Series);

Умеет:

- Использовать библиотеку Pandas и встроенные средства Python3 для загрузки данных в среду разработки;
- Использовать библиотеки Matplotlib и Seaborn для визуализации данных;
- Использовать библиотеки Keras и TensorFlow для построения моделей нейронных сетей;

Владеет:

- Навыками работы с библиотеками NumPy, Pandas, Scikit-learn, keras и TensorFlow;

3. Использует большие данные

Знает:

- Основные принципы построения моделей поддержки принятия решений на основе методов машинного обучения;
- Основные инструменты языка python для осуществления построения моделей машинного обучения (методы и объекты библиотек Scikit-learn и TensorFlow);

Умеет:

- Использовать библиотеки Scikit-learn для построения моделей машинного обучения;

Владеет:

- Навыками работы с библиотеками визуализации данных Matplotlib и seaborn.

Наименование сферы	Наименование компетенции	БАЗОВЫЙ УРОВЕНЬ РАЗВИТИЯ КОМПЕТЕНЦИЙ Способность проявляется под внешним контролем / при внешней постановке задачи/ обучающийся пользуется готовыми, рекомендованными продуктами	ПРОДВИНУТЫЙ УРОВЕНЬ РАЗВИТИЯ КОМПЕТЕНЦИЙ Способность проявляется, но обучающийся эпизодически прибегает к экспертной консультации/ самостоятельно подбирает и пользуется готовыми продуктами
Стандарты и методики в ИТ	Применяет языки программирования для решения профессиональных задач		Самостоятельно применяет языки программирования (в т.ч. скрипты) и настраиваемые программные инструменты для автоматизации процессов в профессиональной деятельности

Средства программной разработки	Применяет форматы обмена данными и языки разметки	Участвует в проектах по созданию заказного ПО в роли разработчика интеграционных решений под контролем опытных специалистов	
Большие данные	Использует большие данные	Анализирует большие данные в проектах под контролем опытных специалистов	

1.4. Категория слушателей

К освоению программы допускаются лица, имеющие высшее образование, владеющие основами экономических и математических дисциплин (в объёме базовой части, читаемой всем специальностям в вузах) и являющиеся уверенными пользователями персонального компьютера. Программа реализуется в рамках проекта «Цифровые кафедры», для студентов ВШ КМиС IT-направления подготовки.

1.5. Трудоемкость обучения

Объем программ профессиональной переподготовки составляет 252 ак. ч. Продолжительность обучения – 9 месяцев (4 часа в неделю).

1.6. Форма обучения

Форма обучения очно-заочная с применением дистанционных образовательных технологий.

2. СОДЕРЖАНИЕ ПРОГРАММЫ

2.1. Учебный план дополнительной профессиональной программы профессиональной переподготовки «Анализ данных и машинное обучение в среде Python»

№№	Наименование дисциплины	Общая трудоемкость, ч	По учебному плану с использованием дистанционных образовательных технологий, ч								СРС, ч, в т.ч. КСР, ч	Текущий контроль (при наличии)**	Промежуточная аттестация (при наличии)	
			Аудиторные занятия, ч ¹				Дистанционные занятия, ч						Зачет	Экзамен
			Всего	Из них			Всего	Из них						
				Лекции	Лабораторные работы	Практические занятия, семинары		Лекции	Лабораторные работы	Практические занятия, семинары				
1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
1.	Основы процедурного программирования с применением языка Python 3	121	33	15		18	34	15		19	54	1«Д»	1«Т»	
2.	Использование инструментария языка Python для проведения анализа данных	105	28	4		24	27	17		10	50	1«Д»	1«Т»	
3.	Практика	20	2			2					18	1«Д»		
Консультация к итоговой аттестации		4	4			4								

¹ * В учебном плане программы, реализуемой в полном объеме с использованием дистанционных образовательных технологий, графы 4-7 исключаются.

** В соответствующей графе указывается количество и технология приема:

«Т»- прием, осуществляемый по традиционной образовательной технологии;

«Д»- прием, осуществляемый с использованием дистанционных образовательных технологий.

Подготовка и проведение итоговой аттестации ²	2									2			1«Т»
Итого:	252	67	19		48	61	32		29	124			

² Количество часов указываются в столбцах «Общая трудоемкость» и СРС

2.2. Учебно-тематический план дополнительной профессиональной программы профессиональной переподготовки «Анализ данных и машинное обучение в среде Python»

№№	Наименование дисциплины	Общая трудоемкость, ч	По учебному плану с использованием дистанционных образовательных технологий, ч								СРС, ч, в т.ч. КСР, ч	Текущий контроль (при наличии) **	Промежуточная аттестация (при наличии)	
			Аудиторные занятия, ч ³				Дистанционные занятия, ч						Зачет	Экзамен
			Всего	Из них			Всего	Из них						
				Лекции	Лабораторные работы	Практические занятия, семинары		Лекции	Лабораторные работы	Практические занятия, семинары				
16.	17.	18.	19.	20.	21.	22.	23.	24.	25.	26.	27.	28.	29.	30.
1.	Основы процедурного программирования с применением языка Python 3	121	33	15		18	34	15		19	54	1«Д»	1«Т»	
1.1	Типы данных, операторы и переменные в Python	14	4	2		2	4	2		2	6	Расчетно-аналитические задания		
1.2	Управление потоком исполнения с использованием оператора if/else	15	4	2		2	5	2		3	6	Расчетно-аналитические задания		

³ * В учебном плане программы, реализуемой в полном объеме с использованием дистанционных образовательных технологий, графы 4-7 исключаются.

** В соответствующей графе указывается количество и технология приема:

«Т»- прием, осуществляемый по традиционной образовательной технологии;

«Д»- прием, осуществляемый с использованием дистанционных образовательных технологий.

1.3	Работа со строками	16	4	2		2	4	2		2	8	Расчетно-аналитические задания		
1.4	Создание циклов	18	6	2		4	4	2		2	8	Расчетно-аналитические задания		
1.5	Работа со списками, кортежами	14	4	2		2	4	2		2	6	Расчетно-аналитические задания		
1.6	Создание функций	14	4	2		2	4	2		2	6	Расчетно-аналитические задания		
1.7	Подключение модулей	15	4	2		2	5	2		3	6	Расчетно-аналитические задания		
1.8	Создание словарей	15	3	1		2	4	1		3	8	Расчетно-аналитические задания		
2.	Использование инструментария языка Python для проведения анализа данных	105	28	4		24	27	17		10	50	1«Д»	1«Т»	
2.1	Методология анализа данных CRISP-DM	15	4	2		2	5	3		2	6	Расчетно-аналитические задания		

2.2	Работа с данными с использованием библиотеки NumPy	17	4			4	5	3		2	8	Расчетно-аналитические задания		
2.3	Манипуляции над данными с использованием библиотеки Pandas	16	4	2		2	4	2		2	8	Расчетно-аналитические задания		
2.4	Визуализация данных с использованием библиотек Matplotlib и Seaborn	15	2			2	5	3		2	8	Расчетно-аналитические задания		
2.5	Машинное обучение с использованием Scikit-learn	28	12			12	4	4			12	Расчетно-аналитические задания		
2.6	Построение нейронных сетей с использованием Keras и TensorFlow	14	2			2	4	2		2	8	Расчетно-аналитические задания		
3.	Практика	20	2			2					18	1«Д»		
	Консультация к итоговой аттестации	4	4			4								
	Подготовка и проведение итоговой аттестации ⁴	2									2			1«Т»
	Итого:	252	67	19		48	61	32		29	124			

⁴ Количество часов указываются в столбцах «Общая трудоемкость» и СРС

2.3. Рабочая учебная программа дисциплины

Раздел 1. «Основы процедурного программирования с применением языка Python 3»

Целью дисциплины является формирование систематизированных знаний и навыков в области программирования: ознакомление слушателей с основами процедурного программирования с применением языка Python 3, принципами работы, с актуальными парадигмами программирования

Задачи дисциплины: изучение основных конструкций современного языка программирования на примере Python; изучение алгоритмов обработки массивов и других структурированных данных с использованием языка Python.

Тема 1.1. «Типы данных, операторы и переменные в Python»

Содержание: Общие сведения о языке Python, Типы данных, поддерживаемые в языке Python 3, основные проблемы динамической типизации. Основные операторы языка. Способы создания переменных. Функция вывода в консоль.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание первого приложения на Python «Hello World!». Выполнение практических задач для закрепления знаний о переменных, операторах и типах данных. Обучение работе в Spyder	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой, связанной с устройством языка Python	Работа с литературой, подготовка отчёта	6

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Описание языка Python.
2. Техническая документация Python.
3. Кроссплатформенность языка Python.
4. Среды разработки для работы с Python.
5. Высокоуровневый язык Python.
6. Среда IDLE.
7. Области применения Python.
8. Пустые строки и комментарии в python. Вывод в консоль

Задача по переменным

1. Создайте переменную clothes. Присвойте ей значение «домашняя одежда»
2. Выведите в консоль следующее: «У меня большой гардероб». Следующей строкой выведите: «Утром лучше всего подходит домашняя одежда» с использованием созданной переменной. Вывести ту же строку для дня, вечера и ночью.
3. Прodelать тоже упражнение для завтрака, обеда и ужина введя переменную meal и заменив первую строку на «мои предпочтения в еде».

Задача на остаток от деления

Необходимо разделить 30 человек на команды. Каждому была выдана футболка с его номером. Для определения номера команды необходимо найти остаток от деления порядкового номера на 4. Вывести в консоль список порядковых номеров и соответствующую ему команду.

Кейс

Директор цеха официального дилера автомобилей дорогой немецкой марки просит вас разработать программу, которая выводила бы на экран калькуляцию по осмотру автомобиля и объему планируемых работ.

Калькуляция должна выводиться в следующем виде:

Расчет работ по представленному автомобилю:

Замена масла (работы)..... Значение переменной Oil_work_cost руб.

Масло Castrol..... Значение переменной Oil_price руб.

Замена воздушного фильтра.... Значение переменной Filtr_work руб.

Воздушный фильтр..... Значение переменной Filtr_price руб.

Итого.....Значение переменной Cost_sum руб.

Персональная скидка.....3%

Итого с учетом скидки.....Net_sum_cost

Спасибо, что выбираете Нас!

Для выполнения задания необходимо:

Создать переменную Oil_work_price, равную стоимости единицы времени (зарплата сотрудника в час). В данном случае, необходимо установить ее равной 500 руб.

Создать переменную Oil_work_time, которая будет равна трудозатратам сотрудника на замену масла. Присвоить ей значение 0,7

Для расчета переменной Oil_work_cost необходимо выполнить умножение Oil_work_price и Oil_work_time.

Переменная Oil_price должна содержать цену на масло. Закупочная цена равна 700 рублей. Однако, нам необходимо включить в ее стоимость наценку, включив в нее затраты на хранение и транспортировку. Для расчета Oil_price предположим, что наценка составляет 5 процентов от стоимости масла.

Аналогично с переменной Oil_work_cost рассчитать переменную Filtr_work, с учетом того, что стоимость единицы времени 450 руб, трудозатраты составят 0,5.

Аналогично переменной Oil_price рассчитать переменную Filtr_price при учете, что закупочная стоимость фильтра составляет 300 руб, наценка 5 процентов.

Рассчитать значение переменной Cost_sum, как сумму всех проведенных работ.

Значение Net_sum_cost рассчитать, как Cost_sum за вычетом скидки.

Тема 1.2. «Управление потоком исполнения с использованием оператора if/else»

Содержание: Общие сведения об условных выражениях. Синтаксис и структура условных выражений с использованием конструкции if/else, elif. Логические операторы языка Python, их использование при построении логических выражений. Проверка ошибок с использованием конструкции try/except.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание приложений с использованием конструкций if/else и try/except. Решение кейса	5

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
-------	---------	------------	-------	------------------

1	Базовая	Самостоятельно поработать с литературой и электронными ресурсами, решить задачи для закрепления полученных знаний	Работа с литературой, подготовка отчёта	6
---	---------	---	---	---

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Переменные в Python.
2. Получение пользовательского ввода.
3. Применение строковых методов. Конвертация значений.
4. Генерирование случайных чисел.
5. Условные конструкции с if.
6. Использование выражения elif.

Создание приложений с использованием конструкций if/else и try/except. Решение кейса.

Примеры задач.

Задача 1.

Университет просит вас разработать программу, которая будет информировать студентов, в каком из трех спортивных залов (большой, малый, тренажерный) будут проходить занятия. Для этого необходимо создать переменную grade, которая будет содержать номер курса (1, 2 или 3).

С помощью конструкции if/elif/else создать программу, которая в зависимости от значения переменной будет выводить следующее:

Если студент младше 2го курса, то «Вам необходимо пройти в малый зал»

Если студент на 2м курсе, то «Вам необходимо пройти в большой зал»

Если студент 3го курс и старше, то «Вам необходимо пройти в тренажерный зал».

Задача 2.

К вам обращается сотрудник колл-центра с просьбой автоматизировать процесс направления звонков по сотрудникам. В колл-центре всего 4 менеджера, отвечающих на звонки, и старший менеджер, отвечающий на звонки в случае, если все остальные заняты уже в течение 2 минут, а очередь из клиентов на линии составляет более 2 человек. Необходимо разработать приложение, которое будет направлять первого в очереди клиента к одному из менеджеров по следующим условиям:

Если все менеджеры свободны, направляет к первому менеджеру;

Если первый занят, второй и третий свободны, то звонок направляют второму;

Если первый свободен, второй занят, то направляется третьему;

Если первый и второй заняты, третий свободен, то направляем звонок третьему;

Если заняты второй и третий, первый свободен, то звонок направляется первому;

Если заняты первый и третий, то направляем звонок второму;

(перебрать все возможные варианты)

Если заняты все, очередь составляет меньше двух человек, то ожидание оператора;

Если все заняты, очередь больше или эквивалентна двум человекам, то переадресация клиента главному менеджеру.

Необходима переменная manager_num, которая будет содержать либо номер менеджера, к которому адресуется данный абонент, либо запись «Главный менеджер». Обязательно использовать логические переменные, принимающие значения True или False.

В результате работы приложения в консоль должен выводиться «Ваш менеджер - manager_num».

Кейс.

В компании по продаже автомобилей необходимо приложение, которое будет рассчитывать сумму отчислений при растаможивании автомобиля. Для расчета используется несколько параметров:

Возраст авто;

Объем двигателя

Цена.

Цена учитывается только в случае, если авто старше 3х лет.

Тарифы, которые необходимо учитывать при расчете:

Возраст	Цена	Ставка
Новый (от года выпуска прошло не более 2х лет)	До 8500 евро	2,5 евро за 1 куб. сантиметр
	8500 до 16 700 евро	2,5 евро за 1 куб. сантиметр
	16 700 до 42 300 евро	5,5 евро за 1 куб. сантиметр
	42 300 до 84 500 евро	7,5 евро за 1 куб. сантиметр
	84 500 до 169 000 евро	15 евро за 1 куб. сантиметр
	Свыше 169 000 евро	не менее 20 евро за 1 куб. сантиметр
Возраст	Объем двигателя	Ставка
Автомобиль от 3х до 5ти лет	Не более 1000 куб.см	1,5 евро за 1 куб.см
	1000 – 1500 куб.см	1,7 евро за 1 куб.см
	1500 – 1800 куб.см.	2,5 евро за 1 куб.см
	1800 – 2300 куб.см.	2,7 евро за 1 куб.см
	2300 – 3000 куб.см	3 евро за 1 куб.см
	Более 3000 куб.см.	3,6 евро за 1 куб.см

Разработать приложение с использованием конструкции if/elif/else, которое позволяло бы по значению параметров рассчитывать величину таможенных отчислений. Результирующую переменную назвать result_calc. Отчисления должны выводиться в консоль по следующему примеру: «Таможенные отчисления на ваш автомобиль составят result_calc»

Тема 1.3. «Работа со строками»

Содержание: Понятие строки в программировании. Создание строк средствами языка Python. Основные встроенные функции работы со строками в Python. Индексация строк, формирование срезов. Иммутабельность строк в Python. Использование escape-последовательностей в строках.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание приложений для обработки строк. Решение кейса	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой и электронными ресурсами, решить задачи для	Работа с литературой, подготовка отчёта	8

		закрепления полученных знаний		
--	--	-------------------------------	--	--

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Строки в Python.
2. Escape-последовательности в строках.
3. Сцепление и повторение строк.
4. Работа с числами.
5. Переменные в Python.
6. Получение пользовательского ввода.
7. Применение строковых методов. Конвертация значений.

Создание приложений для обработки строк. Решение кейса.

Пример задач.

Задача 1.

Компания просит вас разработать приложение, которое позволит генерировать имя пользователей в системе как конкатенацию имени и инициалов. Необходимо создать переменные:

Name – содержит имя пользователя

Surname – содержит фамилию пользователя

Patronymic - содержит отчество.

С использованием формирования срезов строк и операции конкатенации сгенерировать имя пользователя, записать его в переменную login и вывести в консоль: «Имя пользователя login@company.ru»

Далее пользователю необходимо задать пароль. Политика компании такова, что пароль должен содержать не менее 7 символов и обязательно включать первую букву имени (обязательно заглавную) и символ «?».

Написать приложение, которое проверяет качество пароля.

Если удовлетворяет требованиям компании, то вывести в консоль сообщение «Пароль изменен!»

Если не удовлетворяет хотя бы одному из них, то «Пароль не соответствует требованиям компании».

Задача 2.

Программист, разрабатывающий приложение для хранения документов для архива, должен предусмотреть в приложении возможность вывода информации о документе в консоль.

Заголовок должен содержать только заглавные буквы, Фамилия, имя автора должны начинаться с Заглавной буквы, остальные строчные, Тип документа должен быть выведен только строчными буквами. На каждый параметр необходимо создать отдельную переменную. Вывести результат в консоль.

Кейс.

Компания по продаже одежды разрабатывает каталог товаров и просит вас помочь в разработке. Необходимо создать приложение, которое:

- Генерирует артикул товара из названия товара и первых трех букв страны производителя, все буквы должны быть заглавные (артикул должен выглядеть, например, если переменная, характеризующая название товара хранит значение «Платье», а переменная для страны производителя содержит «Россия», то артикул: платье_рос).
- Выводит в консоль следующую информацию: Наименование товара, артикул, цена. Наименование товара должно выводиться как заголовок (первая буква заглавная).

- Магазин проводит акцию, вся одежда из Германии распродается со скидкой 20 процентов. Необходимо проверить условие по стране производителя и изменять цену.

Тема 1.4. «Создание циклов»

Содержание: Изучение понятия цикла. Создание итеративных циклов с использованием конструкции for. Использование range в конструкции for. Создание циклов с использованием конструкции while.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание приложений с использованием циклов. Решение кейса	6

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой и электронными ресурсами, решение задач для закрепления полученных знаний	Работа с литературой, подготовка отчёта	8

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Создание циклов с использованием while.
2. Составные условия.
3. Алгоритмы на псевдокоде.
4. Применение циклов for.
5. Счет с помощью цикла for.
6. Операторы и функции для работы с последовательностями: применение к строкам.
7. Индексация строк. неизменяемость строк.
8. Срезы строк.

Создание приложений с использованием циклов. Решение кейса.

Пример задач.

Задача 1:

Имеется строка:

"ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz"

Необходимо разработать приложение, рассчитывающее количество уникальных символов (заглавные и строчные буквы – это разные символы). Результаты записать в переменную ges и вывести в консоль.

Задача 2.

В торговом центре проводится розыгрыш призов. Для получения приза необходимо загадать число. Случайным образом администраторы выбирают второе число в интервале от 1 до 9. Те пользовательские номера, сумма цифр которого делятся нацело на загаданное администрацией число, считаются выигрышными. Компания обратилась к вам, чтобы вы разработали приложение, которое будет осуществлять вывод в консоль выигрышных номеров. Если выигрышных номеров слишком много, отбираются первые 5, то есть как только в консоль выведено 5 игроков, розыгрыш заканчивается.

Кейс.

Модифицируем программный код из задачи темы «Работа со строками».

Политика компании изменилась. Теперь они ужесточили требования к паролям. Пароль может быть задан пользователем, однако к нему есть требования:

Не может содержать менее 10 символов

Обязательно содержит одну заглавную букву;

Все буквы должны быть латинскими;

В пароле должны содержаться символы @,~,*,(,),'

Доработайте код с использованием цикла для модернизации проверки паролей.

Тема 1.5. «Работа со списками, кортежами»

Содержание: Изучение понятия списка. Создание списка средствами языка Python.

Функции для работы со списками, встроенные в язык Python. Понятие кортежа. Отличие списка и кортежа. Назначение списков и кортежей.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание приложений с применением списков. Решение кейса	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой и электронными ресурсами, решение задач для закрепления полученных знаний	Работа с литературой, подготовка отчёта	6

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов:

1. Использование кортежей.
2. Использование списков.
3. Применение списочных методов.
4. Вложенные последовательности.
5. Распределенные ссылки.
6. Использование словарей.

Создание приложений с применением списков. Решение кейса.

Примеры задач

Задача 1.

Компания имеет выгрузки из систем, которые сохранены в виде списков. Один список содержит наименования товаров, второй артикулы. Компания просит вас написать приложение, которое позволило бы объединить эти списки в один и вывести в консоль.

Задача 2.

В компании на входе при прикладывании пропуска к турникету имя пользователя заносится в список. В конце дня рассчитывается, сколько сотрудников было в этот день на в офисе, а также при запросе Роспотребнадзора в случае выявления коронавирусной инфекции необходима возможность проверки, был ли определенный сотрудник на рабочем месте.

Вас просят реализовать данные требования. В консоль необходимо выводить, сколько сотрудников было в офисе в этот день, а также, был ли на работе сотрудник с подозрением на коронавирус.

Кейс.

Компания реализует оказание услуг в области ногтевого сервиса. Вас просят, как аналитика данных, рассчитать некоторую статистику по одной неделе. Имеются 3 списка, сгенерированных в результате работы в эту неделю:

Первый список характеризует виды маникюра, которые делает данный салон:

Nail_style=["Шеллак", "Френч", "Обычный лак", "Гель-лак", "Акрил"]

Второй список характеризует цены на эти услуги:

Price = [2000, 1500, 1000, 3000, 3500, 3200]

Третий список показывает, сколько раз за неделю заказывали тот или иной вид маникюра соответственно

Week = [4, 5, 4, 8, 6]

Хозяйка салона просит вас написать приложение, которое выводит в консоль:

Среднее значение посещений салона

Общее количество посещений салона

Выручку салона

Выручку по видам маникюра

Среднюю выручку в день по видам маникюра.

Тема 1.6. «Создание функций»

Содержание: Понятие функций в программировании. Назначение функций. Синтаксис создания функций в языке Python. Параметры функций. Возвращаемые значения функций.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание приложений с применением функций. Решение кейса	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой и электронными ресурсами, решение задач для закрепления полученных знаний	Работа с литературой, подготовка отчёта	6

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Создание функций.
2. Параметры и возвращаемые значения.
3. Именованные аргументы и значения параметров по умолчанию.
4. Использование глобальных переменных и констант.
5. Сложности использования функций в Python

Создание приложений с применением функций. Решение кейса.

Задача 1.

Усовершенствовать код из задачи про ногтевой сервис (тема «работа со списками, картами»).

Необходимо реализовать функцию descriptive statistics, которая на вход принимает 3 параметра (список наименования услуг, цены, количество посещений), а на выходе возвращает значение средней выручки за неделю, максимальное число продаж за неделю,

среднюю цену на услуги, среднее число посещений салона). Эти показатели необходимо вывести в консоль.

Задача 2.

Усовершенствовать код из задачи про пропуска и турникеты (тема «работа со списками, картами»). Необходимо создать функцию, которая на входе принимает данные о людях, пришедших в офис, и сотрудника. Которым интересуется Роспотребнадзор. Функция должна возвращать информацию о том, был ли указанный сотрудник на работе. Вывести в консоль результат работы функции.

Кейс.

Имеются данные о продажах некоторой аптеки:

Товары, которые продавала аптека на неделе: Спазмалгон (300 руб, купили 10 раз), Нурофен (500 руб, купили 6 раз), Тизин (124 руб, купили 12 раз), Назонекс (750 руб, купили 4 раза), Грипферон (567 руб, купили 5 раз), Боярышник (320 руб., купили 20 раз).

Необходимо разработать приложение (с использованием функций), чтобы на выходе можно было получить средние данные о продажах, выручке, ценах, а также максимальное и минимальное значение. Всю статистику необходимо вывести в консоль.

Тема 1.7. «Подключение модулей»

Содержание: Понятие библиотеки (модуля). Способы подключения внешних модулей и библиотек. Понятие пространства имен и его назначение. Способы избегания перенасыщения пространства имен. Уровни доступа.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание приложений с применением внешних модулей. Решение кейса	5

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой и электронными ресурсами, решение задач для закрепления полученных знаний	Работа с литературой, подготовка отчёта	6

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Классы в python.
2. Создание новых классов с помощью наследования.
3. Расширение класса через наследование.
4. Переопределение унаследованных методов.
5. Что такое полиморфизм.
6. Создание модулей.
7. Инкапсуляция в python.

Создание приложений с применением внешних модулей. Решение кейса.

Кейс.

Код для расчета описательной статистики сохранить как библиотеку (Тема: «Создание функций»). Импортировать данный модуль и с помощью него рассчитать и вывести описательную статистику для заболевших коронавирусной инфекцией за последние 7 дней

Тема 1.8. «Создание словарей»

Содержание: Понятие словаря. Синтаксис создания словаря в языке Python. Встроенные функции Python для работы со словарями

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Создание приложений с применение словарей. Решение кейса	5

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой и электронными ресурсами, решение задач для закрепления полученных знаний	Работа с литературой, подготовка отчёта	8

Форма текущего контроля расчетно-аналитические задания

Создание приложений с применение словарей. Решение кейса.

Кейс:

Реализовать функцию подсчета очков в игре Эрудит.

```
letters = ["А", "Б", "В", "Г", "Д", "Е", "Ж", "З", "И", "Й", "К", "Л", "М", "Н", "О", "П", "Р", "С", "Т", "У", "Ф", "Х", "Ц", "Ч", "Ш", "Щ", "Ъ", "Ы", "Ь", "Э", "Ю", "Я"]
```

```
points = [1, 3, 3, 2, 1, 4, 2, 4, 1, 8, 5, 1, 3, 4, 1, 3, 10, 1, 1, 1, 1, 4, 4, 8, 4, 10, 11, 10, 10, 10]
```

Форма промежуточной аттестации зачет

Примерный перечень контрольных вопросов:

1. Описание языка Python.
2. Техническая документация Python.
3. Кроссплатформенность языка Python.
4. Среды разработки для работы с Python.
5. Высокоуровневый язык Python.
6. Среда IDLE.
7. Области применения Python.
8. Пустые строки и комментарии в python. Вывод в консоль
9. Метрики кода.
10. Строки в Python.
11. Escape-последовательности в строках.
12. Сцепление и повторение строк.
13. Работа с числами.
14. Переменные в Python.
15. Получение пользовательского ввода.
16. Применение строковых методов. Конвертация значений.
17. Генерирование случайных чисел.
18. Условные конструкции с if.
19. Использование выражения elif.
20. Создание циклов с использованием while.
21. Составные условия.
22. Алгоритмы на псевдокоде.

23. Применение циклов for.
24. Счет с помощью цикла for.
25. Операторы и функции для работы с последовательностями: применение к строкам.
26. Индексация строк. неизменяемость строк.
27. Срезы строк.
28. Использование кортежей.
29. Использование списков.
30. Применение списочных методов.
31. Вложенные последовательности.
32. Распределенные ссылки.
33. Использование словарей.
34. Создание функций.
35. Параметры и возвращаемые значения.
36. Именованные аргументы и значения параметров по умолчанию.
37. Использование глобальных переменных и констант.
38. Сложности использования функций в Python
39. Классы в Python.
40. Создание новых классов с помощью наследования.
41. Расширение класса через наследование.
42. Переопределение унаследованных методов.
43. Что такое полиморфизм.
44. Создание модулей.
45. Инкапсуляция в Python.

Критерии оценивания зачета:

Зачет	От 50 до 100 баллов
Незачет	От 0 до 49 баллов

Учебно-методическое и информационное обеспечение дисциплины

Рекомендуемая литература

Основная литература

1. Python 3 и PyQt 5. Разработка приложений: [Электронный ресурс] Пособие / Прохоренок Н.А. - СПб:БХВ-Петербург, 2016. - 833 с. ISBN 978-5-9775-3648-6 - Режим доступа: <http://znanium.com/catalog/product/944629>
2. Рамальо, Л. Python. К вершинам мастерства / Лучано Рамальо ; пер. с англ. А.А. Слинкина. - Москва : ДМК Пресс, 2016. - 768 с. - ISBN 978-5-97060-384-0. - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1028052> - Текст : электронный. - URL: <http://znanium.com/catalog/product/1028052>
3. Основы алгоритмизации и программирования на Python : учеб. пособие / С.Р. Гуриков. — М. : ФОРУМ : ИНФРА-М, 2018. — 343 с. — (Высшее образование: Бакалавриат). - Режим доступа: <http://znanium.com/catalog/product/924699>

Дополнительная литература:

4. Язык программирования Python: практикум : учеб. пособие / Р.А. Жуков. — М. : ИНФРА-М, 2019. — 216 с. + Доп. материалы [Электронный ресурс; Режим доступа: <http://www.znanium.com>]. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/textbook_5cb5ca35aaa7f5.89424805.

5. Паклин Н.Б., Орешков В.И. Бизнес-аналитика, От данных к знаниям(+CO): Учебное пособие. 2е изд. испр. СПб.: Питер, 2013. 704 с.: ил.
6. Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rüdiger Wirth (2000); CRISP-DM 1.0 Step-by-step data mining guide
7. Шелудько, В. М. Основы программирования на языке высокого уровня Python : учебное пособие / В. М. Шелудько ; Южный федеральный университет. - Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2017. - 146 с. - ISBN 978-5-9275-2649-9. - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1021662>
8. Златопольский, Д.М. Основы программирования на языке Python / Д.М. Златопольский. - Москва : ДМК Пресс, 2017. - 284 с. - ISBN 978-5-97060-552-3. - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1028147>
9. Коэльо, Луис Педро Построение систем машинного обучения на языке Python / Луис Педро Коэльо, Вилли Ричарт ; пер. с англ. А. А. Слинкина. - 2-е изд. - Москва : ДМК Пресс, 2016. - 302 с. - ISBN 978-5-97060-330-7. - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1027824>
10. Python. Самое необходимое: Практическое руководство / Прохоренок Н.А. - СПб:БХВ-Петербург, 2010. - 414 с.: ISBN 978-5-9775-0614-4 - Режим доступа: <http://znanium.com/catalog/product/354989>

Нормативно-правовые документы:

1. PEP 8 Style Guide for Python Code
2. Python 3. documentation

Перечень информационно-справочных систем

1. <http://www.consultant.ru> - Справочно-правовая система Консультант Плюс;
2. <http://www.garant.ru> - Справочно-правовая система Гарант.

Перечень электронно-образовательных ресурсов

Перечень профессиональных баз данных

1. <https://github.com/> - платформа разработчиков, включающая базу данных открытых библиотек
2. <https://stackoverflow.com/questions/tagged/python> - библиотека ответов на типичные ошибки профессиональных программистов

Перечень ресурсов информационно-телекоммуникационной сети "интернет", необходимых для освоения дисциплины

1. <https://docs.python.org/>
2. <https://scikit-learn.org/stable/>
3. <http://deeplearning.net/software/theano/>

Материально-техническое обеспечение дисциплины

Дисциплина «**Основы процедурного программирования с применением языка Python 3**» обеспечена:

- учебной аудиторией, оборудованной учебной мебелью, мультимедийными средствами обучения для демонстрации лекций-презентаций;
для проведения занятий семинарского типа (*практические занятия, лабораторные занятия*);
- учебной аудиторией, оборудованной учебной мебелью и техническими средствами обучения, служащими для представления учебной информации;
- компьютерным классом;
для самостоятельной работы, в том числе для курсового проектирования;

- помещением для самостоятельной работы, оснащённым компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде университета.

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения

№ п/п	Перечень информационных технологий, программного обеспечения
1	Операционная система Windows 10, Microsoft Office Professional Plus: 2019 год (Word, MS Excel, MS Power Point, MS Access) Python 3 Anaconda

Рабочая учебная программа дисциплины

Раздел 2. «Использование инструментария языка Python для проведения анализа данных»

Целью дисциплины является развитие навыков применения встроенных средств языка Python 3, библиотеки визуализации данных Matplotlib и Seaborn, библиотеки работы с данными Pandas и NumPy, Библиотеки машинного обучения Scikit-learn, TensorFlow. Введение в автоматизированные методы работы с данными - машинное обучение и нейронные сети.

Задачи дисциплины: получение слушателями представления о Data mining, методологией исследования данных, навыков работы с библиотекой NumPy, Pandas, Matplotlib и Seaborn.

Тема 2.1. «Методология анализа данных CRISP-DM»

Содержание: Получение слушателями представления о Data mining, Применение технологий Data mining в бизнес-задачах, планирование этапов проведения анализа данных, изучение методологии исследования данных CRISP-DM и его основных фаз.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Рассмотрение практических примеров применения изученных методов и подходов в формате case-study	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Самостоятельно поработать с литературой, связанной с методами и моделями Data Mining	Работа с литературой, подготовка отчёта	6

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Представления о Data mining.
2. Применение технологий Data mining в бизнес-задачах.
3. Этапов проведения анализа данных.
4. Исследования данных CRISP-DM и его основных фаз.

Рассмотрение практических примеров применения изученных методов и подходов в формате case-study

Крупная ИТ-компания занимается сервисным обслуживанием в сфере информационных технологий. В их задачи входят обеспечение информационной безопасности, мелкие проекты в сфере доработки существующих информационных систем, поддержка пользователей. В рамках этих задач компания осуществляет планирование трудовых ресурсов на основе технологических карт заказчика.

В рамках компании планирование проводится в краткосрочных форматах (на месяц, на конец года). Каждый месяц план корректируется в соответствии с фактическим выполнением. В настоящий момент в компании все планирование осуществляется вручную интуитивно на каждого сотрудника. Планирование осуществляет начальники отделов и тратят на это 20 человеко-часов каждый месяц. В компании существует воронка продаж, которая определяет целевые значения расходов компании за год. Расходы компании рассчитываются в зависимости от затраченных часов сотрудника и его оклада.

В рамках кейса необходима определить:

1. Какая бизнес-задача существует в компании?
2. Какова цель компании для снижения затрат?
3. Возможно ли решить данную проблему с использованием Data Mining?
4. Какие задачи машинного обучения сформированы в рамках данного кейса?
5. Какие модели Data Mining Применимы для решения поставленных задач?
6. Какие данные необходимы для решения поставленных задач?

На основании Кейса 1 (тема 1 первого модуля) и проделанного в рамках него анализа необходимо подготовить план проекта Data mining по стандарту CRISP-DM

Тема 2.2. «Работа с данными с использованием библиотеки NumPy»

Содержание: Стандартные типы библиотеки NumPy, Массивы в NumPy, Выполнение вычислений над массивами и универсальные функции, Агрегирование в NumPy, Сравнения, маски и алгебра логики, Особенности индексации, Сортировка массивов, Структурированные массивы

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Выполнение заданий для закрепления навыка работы с библиотекой NumPy	6

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Решение задач с использованием языка программирования Python и библиотекой NumPy Загрузить файл audi.csv и выполнить исследование качества данных	Работа с литературой, подготовка отчёта	8

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Стандартные типы библиотеки NumPy.
2. Массивы в NumPy.
3. Выполнение вычислений над массивами и универсальные функции.

4. Агрегирование в NumPy, Сравнения, маски и алгебра логики.
5. Особенности индексации.
6. Сортировка массивов.
7. Структурированные массивы.

Выполнение заданий для закрепления навыка работы с библиотекой NumPy

Примеры задач

1. Напишите программу с использованием NumPy, чтобы создать матрицу 8x8 и заполнить ее шаблоном шахматной доски. Вывести результат в консоль. Результат должен выглядеть следующим образом:
2. Напишите программу с использованием библиотеки NumPy для преобразования значений градусов Цельсия в градусы Фаренгейта. Значения Цельсия сохраняются в массиве NumPy

Кейс.

Компания по продаже недвижимости хочет реализовать модель прогнозирования цен на квартиры. На первом этапе она обращается к Вам с просьбой подготовить данные для анализа. С использованием библиотеки NumPy загрузить данные из файла Nedvig.csv. Рассчитать описательную статистику. Есть ли в файле пропуски? Пригодны ли данные для анализа? заполнить пропуски данных средним значением
Сохранить результат в новом файле JSON

Тема 2.3. «Манипуляции над данными с использованием библиотеки Pandas»

Содержание: Объекты библиотеки Pandas. Индексация и выборка данных, операции над данными, обработка пропусков в данных, иерархическая индексация, объединение наборов данных, агрегирование и группировка, сводные таблицы, векторизованные операции над строками, обработка временных рядов, производительность библиотеки Pandas

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Выполнение заданий для закрепления навыка работы с библиотекой Pandas	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Решение задач с использованием языка программирования Python и библиотекой Pandas	Работа с литературой, подготовка отчёта	8

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Объекты библиотеки Pandas.
2. Индексация и выборка данных, операции над данными.
3. Обработка пропусков в данных, иерархическая индексация.
4. Объединение наборов данных, агрегирование и группировка, сводные таблицы, векторизованные операции над строками.
5. Обработка временных рядов, производительность библиотеки Pandas.

Решение задач с использованием языка программирования Python и библиотекой Pandas

Кейс

Компания по продажам компьютерных игр хочет адаптировать видео-игры из прошлого. Однако, они хотят отобрать только те игры, которые были наиболее популярными. Для этого они обращаются к вам, как к специалисту по анализу данных и просят вас подготовить ретро-данные для анализа.

Вам необходимо:

1. Создать объект DataFrame и загрузить данные vgsales.csv
2. Напишите программу Python с использованием библиотеки Pandas, чтобы получить столбцы DataFrame. Сохранить имена столбцов в списке.
3. Напишите программу Pandas, чтобы получить подробную информацию о третьей видеоигре в таблице DataFrame. Записать информацию о третьей видео-игре в строку.
4. Напишите программу Pandas для подсчета количества строк и столбцов в DataFrame
5. Напишите программу Pandas для сортировки DataFrame на основе поля Year
6. Отфильтровать все игры, разработчик которых является Take-Two Interactive
7. Создать новый объект DataFrame и сохранить в него игры из исходного набора, жанр которых Platform
8. Проверить качество данных. Построить описательную статистику
9. Добавить поле «Коммерчески успешнее», в котором будет значение True, если GlobalSales Выше среднего, False – если меньше.
10. Вывести итоговый DataFrame в файл формата.xlsx.

Сделать выводы, какие игры были наиболее коммерчески успешными.

Тема 2.4. «Визуализация данных с использованием библиотек Matplotlib и Seaborn»

Содержание: Общее описание библиотек, способы построения линейных графиков, Диаграммы рассеяния, визуализация погрешностей, графики плотности, гистограммы, настройки стилей и легенды.

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Выполнение заданий для закрепления навыка работы с библиотеками Matplotlib и Seaborn	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Решение задач с использованием языка программирования Python и библиотеками Matplotlib и Seaborn	Работа с литературой, подготовка отчёта	8

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Описание библиотек, способы построения линейных графиков.
2. Диаграммы рассеяния.
3. Визуализация погрешностей.
4. Графики плотности, гистограммы.
5. Настройки стилей и легенды.

Решение задач с использованием языка программирования Python и библиотеками Matplotlib и Seaborn

Ваш знакомый просит вас помочь ему с выбором фильма на выходные. Для этого он просит вас наглядно отобразить рейтинги отобранных им фильмов. Для решения задачи необходимо проанализировать следующие данные:

Фильмы: Темный рыцарь (2008), Властелин колец: Возвращение Короля (2003), Властелин колец: Братство кольца (2001), Бойцовский клуб (1999), рейтинг

Рейтинг: 9,0; 8,9; 8,8; 8,7

1. Напишите программирование на Python для отображения горизонтальной гистограммы по тем же данным
2. Раскрасить столбцы диаграммы. Каждый столбец должен быть своего цвета.
3. Выберите ширину каждого столбца и их позиции.
4. Напишите программу на Python для создания круговой диаграммы на основе тех же данных с названием фильмов
5. На случайных данных создать диаграмму рассеяния.
6. Напишите программу на Python для построения точечной диаграммы с использованием случайных распределений для создания кругов разных размеров.

Тема 2.5. «Машинное обучение с использованием Scikit-learn»

Содержание: Представление данных в Scikit-learn, Проектирование признаков, Наивная байесовская классификация, линейная регрессия, метод опорных векторов, деревья решений и случайные леса, метод главных компонент, кластеризация методом k-средних, ближайших соседей, модель многослойного персептрона, ансамбли моделей (бэггинг и бустинг), верификация

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Выполнение заданий для закрепления навыка работы с библиотекой Scikit-learn	12

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Решение задач с использованием языка программирования Python и библиотекой Scikit-learn	Работа с литературой, подготовка отчёта	12

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Представление данных в Scikit-learn.
2. Проектирование признаков.
3. Наивная байесовская классификация, линейная регрессия, метод опорных векторов.
4. Деревья решений и случайные леса, метод главных компонент.
5. Кластеризация методом k-средних,
6. Модель многослойного персептрона, ансамбли моделей (бэггинг и бустинг).

Выполнение заданий для закрепления навыка работы с библиотекой Scikit-learn

Кейс:

Кардиологическая больница нуждается в модели прогнозирования смертности от сердечной недостаточности по симптомам. Для этого они обращаются к вам и просят построить модель прогнозирования смерти от сердечной недостаточности.

1. Создайте модель для оценки вероятности смерти от сердечной недостаточности.

Это может быть использовано для помощи больницам в оценке степени тяжести пациентов с сердечно-сосудистыми заболеваниями.

Для этого используйте файл `heart_failure_clinical_records_dataset.csv`

2. Анализ производить по плану CRISP-DM

Компания, занимающаяся продажами автомобилей иностранных марок с пробегом, хочет построить модель для ценообразования для повышения эффективности своей стратегии продаж. Для этого ему необходима модель регрессии, которая позволит по имеющимся данным обобщить зависимости и позволит рассчитывать цены в зависимости от входных параметров.

Имеются данные по автомобилям. Необходимо:

Построить различные модели регрессии для прогнозирования цены на автомобили, сравнить их по точности и качеству. Оценить предсказательную силу модели. Необходимо использовать данные из наборов `audi.csv`, `bmw.csv`, `focus.csv`. Попробовать построить модель прогнозирования цены по отдельности по каждой марке и объединив все данные вместе.

Тема 2.6. «Построение нейронных сетей с использованием Keras и TensorFlow»

Содержание: Понятие нейронных сетей, понятие deep learning, введение в библиотеку Tensor Flow. Возможности библиотеки, построение сетей прямого распространения, построение рекуррентных сетей, построение моделей глубокого обучения, верификация

Перечень лабораторных работ и практических (семинарских) занятий

№ п/п	Наименование лабораторных работ/ практических (семинарских) занятий	Трудоемкость, ч.
1	Выполнение заданий для закрепления навыка работы с библиотеками Keras и TensorFlow	4

Виды самостоятельной работы слушателей (СРС)

№ п/п	Вид СРС	Содержание	Форма	Трудоемкость, ч.
1	Базовая	Решение задач с использованием языка программирования Python и библиотеками Keras и TensorFlow	Работа с литературой, подготовка отчёта	8

Форма текущего контроля расчетно-аналитические задания

Примерный перечень контрольных вопросов

1. Понятие нейронных сетей, понятие deep learning.
2. Введение в библиотеку Tensor Flow.
3. Возможности библиотеки, построение сетей прямого распространения.
4. Построение рекуррентных сетей.
5. Построение моделей глубокого обучения, верификация.

Выполнение заданий для закрепления навыка работы с библиотеками Keras и TensorFlow

Кейс пример

Банк осуществляет выдачу кредитов физическим лицам. С целью уменьшить количество невозвратов кредитов компания принимает решение создать модель, определяющую целесообразность выдачи кредита.

Используя данные `Credit.csv` построить модель нейронной сети, позволяющей определять целесообразность выдачи кредита запрашившим лицам по предоставленным ими данным.

Учебно-методическое и информационное обеспечение дисциплины

Рекомендуемая литература

Основная литература

1. Python 3 и PyQt 5. Разработка приложений: [Электронный ресурс] Пособие / Прохоренок Н.А. - СПб:БХВ-Петербург, 2016. - 833 с. ISBN 978-5-9775-3648-6 - Режим доступа: <http://znanium.com/catalog/product/944629>
2. Рамальо, Л. Python. К вершинам мастерства / Лучано Рамальо ; пер. с англ. А.А. Слинкина. - Москва : ДМК Пресс, 2016. - 768 с. - ISBN 978-5-97060-384-0. - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1028052> - Текст : электронный. - URL: <http://znanium.com/catalog/product/1028052>
3. Основы алгоритмизации и программирования на Python : учеб. пособие / С.Р. Гуриков. — М. : ФОРУМ : ИНФРА-М, 2018. — 343 с. — (Высшее образование: Бакалавриат). - Режим доступа: <http://znanium.com/catalog/product/924699>
4. Китова. О.В. Цифровой бизнес [Электронный ресурс]: учебник / под науч. ред. О.В. Китовой. — М.: Инфра-М, 2018 — 418 с. – ISBN 978-5-16-013017-0/
Режим доступа: <http://znanium.com/catalog/product/905363>

Дополнительная литература:

5. Язык программирования Python: практикум : учеб. пособие / Р.А. Жуков. — М. : ИНФРА-М, 2019. — 216 с. + Доп. материалы [Электронный ресурс; Режим доступа: <http://www.znanium.com>]. — (Высшее образование: Бакалавриат). — www.dx.doi.org/10.12737/textbook_5cb5ca35aaa7f5.89424805.
6. Паклин Н.Б., Орешков В.И. Бизнес-аналитика, От данных к знаниям(+CO): Учебное пособие. 2е изд.испр. СПб.: Питер, 2013. 704 с.: ил.
7. Pete Chapman, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, and Rüdiger Wirth (2000); CRISP-DM 1.0 Step-by-step data mining guide
8. Шелудько, В. М. Основы программирования на языке высокого уровня Python : учебное пособие / В. М. Шелудько ; Южный федеральный университет. - Ростов-на-Дону ; Таганрог : Издательство Южного федерального университета, 2017. - 146 с. - ISBN 978-5-9275-2649-9. - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1021662>
9. Златопольский, Д.М. Основы программирования на языке Python / Д.М. Златопольский. - Москва : ДМК Пресс, 2017. - 284 с. - ISBN 978-5-97060-552-3. - Текст: электронный. - URL: <https://new.znanium.com/catalog/product/1028147>
10. Коэльо, Луис Педро Построение систем машинного обучения на языке Python / Луис Педро Коэльо, Вилли Ричарт ; пер. с англ. А. А. Слинкина. - 2-е изд. - Москва : ДМК Пресс, 2016. - 302 с. - ISBN 978-5-97060-330-7. - Текст : электронный. - URL: <https://new.znanium.com/catalog/product/1027824>
11. Python. Самое необходимое: Практическое руководство / Прохоренок Н.А. - СПб:БХВ-Петербург, 2010. - 414 с.: ISBN 978-5-9775-0614-4 - Режим доступа: <http://znanium.com/catalog/product/354989>

Нормативные правовые документы:

1. PEP 8 Style Guide for Python Code
2. Python 3. documentation

Перечень информационно-справочных систем

1. <http://www.consultant.ru> - Справочно-правовая система Консультант Плюс;
2. <http://www.garant.ru> - Справочно-правовая система Гарант.

Перечень электронно-образовательных ресурсов

Перечень профессиональных баз данных

1. <https://github.com/> - платформа разработчиков, включающая базу данных открытых библиотек
2. <https://stackoverflow.com/questions/tagged/python> - библиотека ответов на типичные ошибки профессиональных программистов

Перечень ресурсов информационно-телекоммуникационной сети "интернет", необходимых для освоения дисциплины

1. <https://docs.python.org/>
2. <https://scikit-learn.org/stable/>
3. <http://deeplearning.net/software/theano/>

Материально-техническое обеспечение дисциплины

Дисциплина «Использование инструментария языка Python для проведения анализа данных» обеспечена:

- учебной аудиторией, оборудованной учебной мебелью, мультимедийными средствами обучения для демонстрации лекций-презентаций, для проведения занятий семинарского типа (*практические занятия, лабораторные занятия*);
- учебной аудиторией, оборудованной учебной мебелью и техническими средствами обучения, служащими для представления учебной информации;
- компьютерным классом; для самостоятельной работы, в том числе для курсового проектирования;
- помещением для самостоятельной работы, оснащенный компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа к электронной информационно-образовательной среде университета.

Преподаватель на базе электронной платформы Zoom организует видеоконференцию для проведения лекций в формате он-лайн (слушателям перед каждой лекцией направляется по электронной почте ссылка- приглашение принять участие в лекции).

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения

№ п/п	Перечень информационных технологий, программного обеспечения
1	Операционная система Windows XP SP3, Windows 7, Anaconda и MS Power Point, проектор, Webinar.ru, Python 3, Anaconda, Google Chrome
2	Internet Explorer 9, Google Chrome 37 и выше, Mozilla Firefox 32 и выше Adobe Flash Player Sun Java JRE Zoom

3. ОЦЕНКА КАЧЕСТВА ОСВОЕНИЯ ПРОГРАММЫ

Форма текущего контроля: расчетно-аналитические задания

Форма промежуточной аттестации: зачет, дифференцированный зачет

Практика. Дифференцированный зачет.

Форма и вид отчетности слушателей о прохождении практической подготовки при проведении практики определяются программой подготовки.

Структура отчета по практике должна включать следующие разделы:

1. титульный лист;
2. задание по практической подготовке;
3. отчет о практической подготовке;
4. дневник по практической подготовке;
5. характеристика - отзыв с базы практики;
6. отзыв руководителя практики от Университета;
7. оглавление (содержание);
8. основную часть;
9. приложения;
10. список использованных источников.

Дифференцированный зачет выставляется по результатам проверки отчетной документации, собеседования и защиты отчета с представлением презентации.

Отлично	От 85 до 100 баллов
Хорошо	От 70 до 84 баллов
Удовлетворительно	От 50 до 69 баллов
Неудовлетворительно	Менее 50 баллов

Форма итоговой аттестации: демонстрационный экзамен

Примерный перечень контрольных вопросов:

1. Описание языка Python.
2. Техническая документация Python.
3. Кроссплатформенность языка Python.
4. Среды разработки для работы с Python.
5. Высокоуровневый язык Python.
6. Среда IDLE.
7. Области применения Python.
8. Пустые строки и комментарии в python. Вывод в консоль
9. Метрики кода.
10. Строки в Python.
11. Эскапе-последовательности в строках.
12. Сцепление и повторение строк.
13. Работа с числами.
14. Переменные в Python.
15. Получение пользовательского ввода.
16. Применение строковых методов. Конвертация значений.
17. Генерирование случайных чисел.
18. Условные конструкции с if.
19. Использование выражения elif.
20. Создание циклов с использованием while.
21. Составные условия.
22. Алгоритмы на псевдокоде.
23. Применение циклов for.
24. Счет с помощью цикла for.
25. Операторы и функции для работы с последовательностями: применение к строкам.
26. Индексация строк. неизменяемость строк.
27. Срезы строк.
28. Использование кортежей.
29. Использование списков.
30. Применение списочных методов.
31. Вложенные последовательности.

32. Распределенные ссылки.
33. Использование словарей.
34. Создание функций.
35. Параметры и возвращаемые значения.
36. Именованные аргументы и значения параметров по умолчанию.
37. Использование глобальных переменных и констант.
38. Сложности использования функций в Python
39. Классы в python.
40. Создание новых классов с помощью наследования.
41. Расширение класса через наследование.
42. Переопределение унаследованных методов.
43. Что такое полиморфизм.
44. Создание модулей.
45. Инкапсуляция в python.
46. Представления о Data mining.
47. Применение технологий Data mining в бизнес-задачах.
48. Этапов проведения анализа данных.
49. Исследования данных CRISP-DM и его основных фаз.
50. Стандартные типы библиотеки NumPy.
51. Массивы в NumPy.
52. Выполнение вычислений над массивами и универсальные функции.
53. Агрегирование в NumPy, Сравнения, маски и алгебра логики.
54. Особенности индексации.
55. Сортировка массивов. Структурированные массивы.
56. Объекты библиотеки Pandas.
57. Индексация и выборка данных, операции над данными.
58. Обработка пропусков в данных, иерархическая индексация.
59. Объединение наборов данных, агрегирование и группировка, сводные таблицы, векторизованные операции над строками.
60. Обработка временных рядов, производительность библиотеки Pandas.
61. Описание библиотек, способы построения линейных графиков.
62. Диаграммы рассеяния, визуализация погрешностей, графики плотности, гистограммы, настройки стилей и легенды.
63. Представление данных в Scikit-learn.
64. Проектирование признаков.
65. Наивная байесовская классификация, линейная регрессия, метод опорных векторов.
66. Деревья решений и случайные леса, метод главных компонент.
67. Кластеризация методом k-средних,
68. Модель многослойного перцептрона, ансамбли моделей (бэггинг и бустинг).
69. Понятие нейронных сетей, понятие deep learning.
70. Введение в библиотеку Tensor Flow.
71. Возможности библиотеки, построение сетей прямого распространения.
72. Построение рекуррентных сетей.
73. Построение моделей глубокого обучения, верификация

Структура демонстрационного экзамена

Типовое задание по демонстрационному экзамену - проект по анализу данных.

Проект необходимо подготовить по плану:

1. Бизнес-анализ (Business Understanding)
2. Анализ данных (Data Understanding)
3. Подготовка данных (Data Preparation)

4. Моделирование (Modeling)
5. Оценка результата (Evaluation)
6. Внедрение (Deployment)

Итоговая аттестация. Показатели и критерии оценивания

Наименование критерия	Баллы
Актуальность сформулированной бизнес-задачи	20
Соответствие этапов проекта стандарту CRISP-DM	10
Проведение адекватной оценки и предобработки данных с использованием библиотек Python	10
Проведение обоснованного выбора алгоритмов моделирования	20
Проведение моделирования с использованием библиотек Python (оценка качества кода)	20
Проведение верификации моделей с использованием библиотек Python (оценка качества кода)	10
Интерпретация полученных результатов	10
Итого	100

Максимально возможное число баллов за аттестационную работу – 100.

Показатели и критерии оценивания планируемых результатов освоения компетенций и результатов обучения, шкала оценивания

Таблица 5

Шкала оценивания		Формируемые компетенции	Индикатор достижения компетенции	Критерии оценивания	Уровень сформированности компетенций
85 - 100	«отлично»	Применяет языки программирования	Применяет языки программирования системно на экспертном уровне. Контролирует весь цикл программной разработки в проектах. Обучает других	Знает верно и в полном объеме: основы программирования на языке Python 3; среды разработки программных приложений с использование языка программирования Python 3.	Экспертный

		<p>Применяет форматы обмена данными и языки разметки</p> <p>Большие данные</p>	<p>Применяет системно на экспертном уровне использование форматов обмена данными и языки разметки. Контролирует решение интеграционных задач в проектах по созданию заказного ПО. Обучает других</p> <p>Применяет системно на экспертном уровне. Контролирует проекты по тематике. Оценивает и применяет новые аналоги и инструменты. Обучает других</p>	<p>Умеет верно и в полном объеме: использовать среды программировать на python 3, ориентированных на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных.</p>	
70 - 84	«хорошо»	<p>Применяет языки программирования</p> <p>Применяет форматы обмена данными и языки разметки</p>	<p>Участствует в проектах по созданию заказного ПО в роли технического лидера (ведущего разработчика). Самостоятельно разрабатывает отдельные модули.</p> <p>Участствует в проектах по созданию заказного ПО в роли ведущего разработчика интеграционных решений.</p>	<p>Знает с незначительными замечаниями: основы программирования на языке Python 3; среды разработки программных приложений с использованием языка программирования Python 3.</p> <p>Умеет с незначительными замечаниями: использовать среды программировать на Python 3,</p>	Продвинутый

		Большие данные	Самостоятельно разрабатывает модули в проектах по тематике больших данных	ориентированны на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных.	
50 - 69	«удовлетворительно»	<p>Применяет языки программирования</p> <p>Применяет форматы обмена данными и языки разметки</p> <p>Большие данные</p>	<p>Участствует в проектах по созданию заказного ПО (по коммерческим договорам) под контролем опытных специалистов.</p> <p>Участствует в проектах по созданию заказного ПО в роли разработчика интеграционных решений под контролем опытных специалистов</p> <p>Владеет базовыми представлениями о тематике, участвует в проектах по большим данным под руководством опытных специалистов.</p>	<p>Знает на базовом уровне, с ошибками: основы программирования на языке Python 3; среды разработки программных приложений с использование языка программирования Python 3.</p> <p>Умеет на базовом уровне, с ошибками: использовать среды программировать на python 3, ориентированны на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных</p>	Базовый

				ых и неструктурированных данных.	
0 - 49	«неудовлетворительн о»	Применяет языки программирования Применяет форматы обмена данными и языки разметки Большие данные	Разработку ведет под контролем опытных наставников. Не привлекается в проекты по созданию заказного ПО (по коммерческим договорам) применяет языки программирования. Владеет основами обмена данными и разметкой. Не привлекается в проекты по созданию заказного ПО (по коммерческим договорам). нет	Не знает на базовом уровне: основы программирования на языке Python 3; среды разработки программных приложений с использование языка программирования Python 3. Не умеет на базовом уровне: использовать среды программирования на python 3, ориентированных на работу с большими данными: для статистической обработки данных и работы с графикой, для работы с разрозненными фрагментами данных в больших массивах, для работы с базами структурированных и неструктурированных данных	Минимальный исходный

Составители программы:

/ Трамова Азиза Мухамадияевна, д.э.н., профессор
кафедры Информатики


/ Савинова Виктория Михайловна, старший преподаватель
кафедры Информатики


/ Красиков Виталий Александрович, к. ф.-м. н., доцент
кафедры Информатики

Заведующий кафедрой информатики

**И.о. директора Высшей школы
кибертехнологий, математики и статистики**











/Китова О.В./

/Титов В.А./

Календарный учебный график
по программе профессиональной переподготовки «Анализ данных и машинное обучение в среде Python»

15.09.2022- 25.11.2022	26.11.2022- 27.12.2022	01.02.2023- 10.02.2023	10.02.2023- 30.03.2023	30.03.2023- 11.05.2023	11.05.2023- 19.05.2023	19.05.2023- 07.06.2023	07.06.2023- 19.06.2023
А	Д	ПА	А	Д	ПА	П	ИА

Условные обозначения:

А- Аудиторные занятия

Д- Дистанционные занятия

П- Практика

С- Стажировка

ПА- Промежуточная аттестация

ИА- Итоговая аттестация

К- Каникулы