
УТВЕРЖДЕНА

Решением Ученого совета
АНО ВО «Центральный университет»
«25» мая 2026 г.
Протокол № 2

**Рабочая программа дисциплины (модуля)
«Python для анализа данных»**

Направление подготовки: 38.04.05 Бизнес-информатика

Направленность (профиль) подготовки: Продуктовый менеджмент

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2026

**Москва
2026**

Содержание

1. Краткая характеристика дисциплины (модуля)	3
2. Перечень планируемых результатов обучения.....	3
3. Тематический план.....	7
4. Содержание дисциплины (модуля).....	7
5. Учебно-методическое обеспечение	9
6. Материально-техническое обеспечение	9
7. Методические и оценочные материалы	11

1. Краткая характеристика дисциплины (модуля)

Рабочая программа дисциплины (модуля) «Python для анализа данных» составлена в соответствии с федеральным государственным образовательным стандартом высшего образования – магистратура по направлению 38.04.05 Бизнес-информатика, профиль Продуктовый менеджмент, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 990 от 12.08.2020 года.

Изучение дисциплины (модуля) «Python для анализа данных» позволяет сформировать практические навыки работы с данными с использованием языка Python, включая обработку, анализ, визуализацию и автоматизацию.

Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 38.04.05 Бизнес-информатика, профиль Продуктовый менеджмент и входит в обязательную часть Блока 1.

Дисциплина (модуль) доступна к изучению на 1 курсе в 1 семестре.

Цель изучения дисциплины (модуля): подготовить обучающихся к самостоятельной работе с данными с использованием языка Python, включая навыки очистки, трансформации, анализа и визуализации данных, а также подключения к API и базам данных, настройки автоматизации и применения готовых решений для прогнозирования и аналитики.

Задачи изучения дисциплины (модуля):

- освоить основные объекты и методы библиотеки pandas для работы с датафреймами: загрузка, фильтрация, группировка, агрегация, объединение данных;
- изучить возможности библиотеки numpy для работы с многомерными массивами и математическими операциями;
- научиться использовать библиотеки matplotlib, seaborn и plotly для построения различных типов визуализаций: линейные графики, гистограммы, тепловые карты, интерактивные дашборды;
- познакомиться с инструментами для работы с большими данными в Python;
- освоить методы подключения к внешним источникам: API, SQL-базы данных, файловые хранилища;

В результате освоения дисциплины (модуля) обучающийся должен:

знать:

- объекты и методы pandas для работы с датафреймами;
- объекты и методы numpy для работы с многомерными массивами;
- объекты и методы matplotlib, seaborn и plotly для визуализации;
- объекты и методы библиотек для работы с большими данными;
- основные методы и библиотеки для работы с API и базами данных;
- базовые методы анализа данных: когортный анализ, прогнозирование временных рядов.

уметь:

- обрабатывать данные в pandas и numpy;
- обрабатывать большие данные через библиотеки python;
- строить различные виды графиков с помощью библиотек matplotlib, seaborn и plotly;
- применять коробочные решения для решения задач машинного обучения и

- временных рядов;
- подключаться к API и базам данных, а также работать с Git;
- настраивать автоматическую рассылку метрик через телеграм-бота;
- использовать основные средства разработки: jupyter notebook, jupyterhub, Google Colab.

владеть:

- практикой решения прикладных задач с использованием python;
- способами анализировать данные с использованием python;
- навыками подготавливать и обрабатывать данные с помощью python;
- навыком выбирать корректный тип визуализации и строить визуализации в python;
- навыком настраивать подключения к внешним источникам.

2. Перечень планируемых результатов обучения

Компетенции, формируемые в результате освоения дисциплины (модуля) при проведении учебных занятий в форме контактной работы обучающихся с педагогическими работниками Университета и в форме самостоятельной работы обучающихся:

Компетенция	Содержание компетенции	Индикатор компетенции	Перечень планируемых результатов обучения по дисциплине (модулю)
ОПК-3.	Способен принимать решения, осуществлять стратегическое планирование и прогнозирование в профессиональной деятельности с использованием современных методов и программного инструментария сбора, обработки и анализа данных, интеллектуального оборудования и систем искусственного интеллекта	ОПК-3.1.	Знает современные методы анализа данных, прогнозирования и поддержки принятия решений, включая технологии машинного обучения и искусственного интеллекта
		ОПК-3.2.	Умеет применять программные инструменты и интеллектуальные системы для сбора, обработки и интерпретации данных в целях стратегического планирования и управления продуктом
		ОПК-3.3.	Имеет практический опыт применения аналитических инструментов и технологий искусственного интеллекта для принятия обоснованных стратегических решений и прогнозирования развития бизнес-продуктов
ОПК-5.	Способен проводить исследования, организовывать самостоятельную и коллективную научно-исследовательскую, проектную и учебно-профессиональную деятельность для поиска, выработки и применения новых решений в области информационно-коммуникационных технологий	ОПК-5.1.	Знает методы научных исследований и проектной деятельности, применяемые для анализа и разработки инновационных решений в сфере ИКТ
		ОПК-5.2.	Умеет планировать и организовывать исследовательскую и проектную работу как индивидуально, так и в составе команды, с учётом целей и задач цифровой трансформации
		ОПК-5.3.	Имеет практический опыт участия в научно-исследовательских и прикладных проектах, направленных на создание и внедрение новых ИКТ-продуктов и сервисов, включая презентацию результатов
ПК-1.	Способен проводить анализ и формулировать требования пользователей на основе	ПК-1.1.	Знает основные методы и инструменты проведения пользовательских

	качественных и количественных исследований		исследований, включая интервью, опросы, наблюдение и анализ пользовательских данных
		ПК-1.2.	Умеет планировать и проводить исследования для выявления потребностей и болей целевой аудитории, а также интерпретировать полученные данные
		ПК-1.3.	Имеет практический опыт формулирования пользовательских инсайтов и требований к продукту на основе результатов качественного и количественного анализа
ПК-2.	Способен разрабатывать и внедрять цифровые продукты с учетом принципов устойчивого развития и цифровой инклюзивности	ПК-2.1.	Знает принципы устойчивого развития и цифровой инклюзивности, а также их применение в процессе создания цифровых продуктов
		ПК-2.2.	Умеет проектировать цифровые решения, учитывающие экологические, социальные и этические аспекты, включая доступность для различных категорий пользователей
		ПК-2.3.	Имеет практический опыт внедрения цифровых продуктов с соблюдением стандартов инклюзивного дизайна и оценки их воздействия на устойчивое развитие

3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		Очная форма				
		Аудиторная работа		Контроль	Самостояте льная работа	
		Лекции	Семинары (практичес кие занятия)			
1	Онбординг	1	3		19	Домашние задания Тест
2	Обработка данных в Pandas	2	4		20	Домашние задания Тест
3	Визуализация	2	4		20	Домашние задания Тест
4	Прикладные инструменты анализа данных	2	4		20	Домашние задания Тест
5	Библиотека NumPy	2	4		2	Домашние задания Тест
6	Машинное обучение и Большие данные	3	4		21	Домашние задания Тест
7	Интеллектуальный анализ данных (Data mining)	4	5		22	Домашние задания Тест
	Экзамен			4		
Итого:		16	28	4	142	
Объем дисциплины (модуля) (в ак. ч.)		190				
Объем дисциплины (модуля) (в зач. ед.)		5				

4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Онбординг	Онбординг
2	Обработка данных в Pandas	Pandas-1. Введение Pandas-2. Соединения и группировки
3	Визуализация	Визуализация-1. Базовая визуализация Визуализация-2. Дашборды
4	Прикладные инструменты анализа данных	Маппинг и лямбда-выражения Когортный анализ
5	Библиотека NumPy	NumPy-1. Введение NumPy-2. Связь с Pandas
6	Машинное обучение и Большие данные	Машинное обучение и коробочные решения Обработка больших данных
7	Интеллектуальный анализ данных (Data mining)	API-1. Основы data mining API-2. Пайплайны

		Рекап и подготовка инфраструктуры для проекта
--	--	---

5. Учебно-методическое обеспечение

Университет располагает полным набором лицензионного и свободно распространяемого программного обеспечения, включая продукты отечественного производства.

Каждый студент в течение всего периода обучения получает индивидуальный неограниченный доступ к электронно-библиотечной системе и электронной информационно-образовательной среде университета. Эти системы предоставляют возможность доступа к ресурсам из любой точки, где есть подключение к сети Интернет, как на территории университета, так и за его пределами.

Студентам обеспечен удаленный доступ к современным профессиональным базам данных и информационным справочным системам.

Основная литература:

1. Плас, Дж. В. Python для сложных задач: наука о данных и машинное обучение : практическое руководство / Дж. В. Плас. - Санкт-Петербург : Питер, 2021. - 576 с. - (Серия «Бестселлеры O'Reilly»). - ISBN 978-5-4461-0914-2. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1739601>.

2. Бизли, Д. Python. Исчерпывающее руководство : практическое руководство / Д. Бизли. - Санкт-Петербург : Питер, 2023. - 368 с. - (Для профессионалов). - ISBN 978-5-4461-1956-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2123386>.

Дополнительная литература:

1. Нисчал, Н. Python — это просто. Пошаговое руководство по программированию и анализу данных : практическое руководство / Н. Нисчал. - Санкт-Петербург : БХВ-Петербург, 2022. - 416 с. - ISBN 978-5-9775-6849-4. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2139457>.

6. Материально-техническое обеспечение

Университет располагает материально-технической базой, соответствующей действующим противопожарным правилам и нормам и обеспечивающей проведение всех видов дисциплинарной и междисциплинарной подготовки, практической и научно-исследовательской работ обучающихся, предусмотренных учебным планом.

Помещения, которые представляют собой учебные аудитории для проведения занятий лекционного типа, занятий семинарского (практического) типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, а также помещения для самостоятельной работы и помещения для хранения и профилактического обслуживания учебного оборудования. Помещения укомплектованы специализированной мебелью и техническими средствами обучения, служащими для представления учебной информации большой аудитории.

Изучение дисциплины (модуля) обеспечивается в учебных аудиториях, оснащенных:

- столами и стульями;
- компьютерной техникой;
- механическими калькуляторами;

— специализированным оборудованием, включая демонстрационное оборудование.

Помещения для самостоятельной работы обучающихся, в том числе приспособленные для использования инвалидами и лицами с ограниченными возможностями здоровья, оснащены компьютерной техникой с возможностью подключения к сети «Интернет» и обеспечением доступа в электронную информационно-образовательную среду Университета.

Обучающимся предоставляется доступ (в том числе удаленный) к ресурсам информационно-телекоммуникационной сети «Интернет», электронным ресурсам (в том числе электронным библиотечным системам, современным профессиональным базам данных и информационным справочным системам):

№	Наименование портала (издания, курса, документа)	Ссылка
1.	Научная электронная библиотека elibrary.ru библиотека	https://elibrary.ru/defaultx.asp
2.	База данных для IT-специалистов	https://habr.com
3.	База данных ScienceDirect	https://www.sciencedirect.com
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	https://minobrnauki.gov.ru/
5.	Федеральный портал «Российское образование»	https://www.edu.ru/
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	http://window.edu.ru/
7.	Единая коллекция цифровых образовательных ресурсов	http://school-collection.edu.ru/
8.	Федеральный центр информационно - образовательных ресурсов	http://fcior.edu.ru/

Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине (модулю), в том числе комплект лицензионного программного обеспечения, современные профессиональные базы данных и информационные справочные системы:

Наименование ПО	Производство	Лицензионное / свободно распространяемое
Операционные системы:		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
Браузеры:		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
Офисные приложения:		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
Программное обеспечение для планирования и учета времени:		
Toggle app	зарубежное	свободно распространяемое
Системы управления проектами:		
Microsoft Imagine (Project)	зарубежное	лицензионное

Системы управления базами данных:		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
Системы резервного копирования (backup):		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
Справочно-правовые системы:		
КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
Средства антивирусной защиты:		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
Среды разработки:		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CorolliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
Пакеты программных средств и библиотек:		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
Системы управления библиографической информацией:		
Zotero	зарубежное	свободно распространяемое
Сервисы и службы:		
Bind	зарубежное	свободно распространяемое
Docker	зарубежное	свободно распространяемое

7. Методические и оценочные материалы

Методические указания для обучающихся по освоению дисциплины (модуля)

В процессе изучения дисциплины (модуля) «Python для анализа данных» в рамках текущего контроля успеваемости используются такие виды учебной работы, как лекции, семинары, домашние задания, тесты, а также различные виды самостоятельной работы обучающихся по заданию преподавателя, направленные на развитие навыков профессиональной лексики, закрепление практических профессиональных компетенций, поощрение инициатив.

Лекция – систематическое, последовательное, монологическое изложение преподавателем учебного материала, как правило, теоретического характера.

В процессе лекций рекомендуется вести конспект лекций: кратко и схематично фиксировать основные идеи, выводы и обобщения лекции; выделять важные мысли, ключевые слова и термины. Необходимо отметить вопросы или материалы, которые вызывают затруднения, и попытаться найти ответы в рекомендованной литературе. Если разобраться в материале не удастся, следует сформулировать вопрос и задать его преподавателю на консультации или во время семинарского (практического) занятия.

Семинар – это форма учебной деятельности, проводимая в учебном заведении под

руководством преподавателя, где студенты активно участвуют в обсуждениях, практических заданиях и других формах взаимодействия.

Для успешной подготовки к семинару рекомендуется заранее ознакомиться с темой занятия и основными материалами, чтобы иметь возможность активно участвовать в обсуждении. Также полезно подготовить вопросы и идеи для обсуждения, что поможет глубже понять материал и продемонстрировать заинтересованность.

Домашнее задание – набор задач по темам недели.

При работе над домашними заданиями важно внимательно ознакомиться с требованиями и сроками выполнения. Рекомендуется разбивать задания на этапы, чтобы избежать перегрузки и лучше усвоить материал. Использовать различные источники информации, включая учебники и онлайн-ресурсы, для более глубокого понимания темы.

Тест – особая форма проверки знаний. Проводится после освоения одной или нескольких тем и свидетельствует о качестве понимания основных понятий изучаемого материала. Тестовые задания составлены к ключевым понятиям, основным разделам, важным терминологическим категориям изучаемой дисциплины.

Для подготовки к тесту необходимо знать терминологический аппарат дисциплины, понимать смысл научных категорий и уметь их использовать в профессиональной лексике. Владение понятийным аппаратом, включённым в тестовые задания, позволяет преподавателю быстро проверить уровень понимания студентами важных методологических категорий.

Самостоятельная работа – работа студентов, направленная на углубленное изучение отдельных тем и вопросов учебной дисциплины (модуля).

В процессе самостоятельной работы студенты взаимодействуют с рекомендованными материалами при минимальном участии преподавателя. Задачи студента включают работу с конспектами лекций (обработка текста), повторное изучение учебных материалов планов и тезисов ответов, изучение дополнительных тем, выполнение учебно-исследовательских заданий и другое.

Система оценивания результатов обучения по дисциплине (модулю)

Критерии получения уровня и оценивания сформированности компетенций по дисциплине (модулю) «Python для анализа данных»

Оценивание уровня учебных достижений обучающихся по дисциплине (модулю) осуществляется в виде текущего контроля успеваемости и промежуточной аттестации.

Промежуточная аттестация по дисциплине (модулю) осуществляется в форме **экзамена**, при этом проводится оценка компетенций, сформированных по дисциплине.

Для оценивания текущего контроля успеваемости и промежуточной аттестации используется десятибалльная шкала оценивания, которая соотносится с традиционной пятибалльной шкалой следующим образом, если иное не указано в силлабусе дисциплины:

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
10	Отлично	Студент полностью владеет знаниями, изложенными в рабочей программе, и глубоко осмысляет дисциплину (модуль). Он самостоятельно и логически последовательно отвечает на все вопросы, акцентируя внимание на наиболее
9	Отлично	
8	Отлично	

Десятибалльная оценка	Пятибалльная оценка	Общая характеристика результата обучения по дисциплине (модулю)
		важном. Умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя ключевые моменты и устанавливая причинно-следственные связи. Четко формулирует ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
7	Хорошо	Студент обладает знаниями предмета почти в полном объеме рабочей программы и самостоятельно, логически последовательно и всесторонне отвечает на все вопросы, акцентируя внимание на наиболее значимых моментах. Он умеет анализировать, сравнивать, классифицировать, обобщать, конкретизировать и систематизировать изученный материал, выделяя его ключевые аспекты и устанавливая причинно-следственные связи. Формулирует свои ответы, уверенно интерпретирует результаты анализов и других исследований, а также решает сложные ситуационные задачи. Студент хорошо знаком с методами исследования, необходимыми для практической деятельности, и умеет связывать теоретические аспекты предмета с практическими задачами.
6	Хорошо	
5	Удовлетворительно	Студент обладает базовыми знаниями по дисциплине (модулю), но испытывает трудности при самостоятельных ответах и использует неточные формулировки. В ходе ответов он допускает ошибки, касающиеся сути вопросов. Студент способен решать только самые простые задачи и владеет лишь минимальным набором методов исследования.
4	Удовлетворительно	
3	Не сдан	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	
1	Не сдан	

Дисциплина (модуль) «Python для анализа данных» оценивается следующим образом:

Активность	Вес	Количество	Описание
Домашние задания	50%	8	Набор задач по темам недели
Тесты	30%	12	Форма проверки знаний студентов через ответы на вопросы с выбором из нескольких вариантов ответов
Экзамен	20%	1	Письменная или устная работа над заданием, направленным на проверку полученных знаний и навыков по дисциплине (модулю)

Формула расчёта итоговой оценки по дисциплине (модулю) «Python для анализа данных»: « $0,5 \times \text{среднее за домашние задания} + 0,3 \times \text{среднее за тесты} + 0,2 \times \text{экзамен}$ ».

Текущий контроль успеваемости обучающихся по дисциплине (модулю)

Примерные домашние задания

Домашнее задание 1. Обработка данных в Pandas

1. Загрузите датасет (например, sales_data.csv — продажи интернет-магазина).
2. Выполните:
 - Проверку на пропущенные значения и их обработку
 - Фильтрацию: выбрать данные за последний квартал
 - Группировку: по товарам и регионам (сумма продаж, средний чек)
 - Агрегацию: количество заказов, общая выручка
3. Выполните соединение:
 - Присоедините таблицу с информацией о клиентах (customers.csv) по customer_id
 - Оставьте только активных клиентов (статус = "active")
4. Сохраните результат в новый файл processed_sales.csv.
5. Напишите краткий вывод: какие товары и регионы показали лучшие результаты?

Домашнее задание 2. Визуализация и когортный анализ

1. На основе обработанных данных из ДЗ1:
 - Постройте линейный график динамики выручки по дням
 - Постройте гистограмму распределения суммы заказов
 - Постройте тепловую карту продаж по дням недели и часам
2. Создайте когортный анализ:
 - Разбейте клиентов по месяцу первой покупки
 - Рассчитайте retention на 1, 7, 30 дни
 - Визуализируйте результат в виде таблицы (тепловая карта)
3. Используйте matplotlib и seaborn.
4. (Дополнительно) Попробуйте воссоздать одну из визуализаций в plotly (интерактивный график).
5. Оформите всё в Jupyter Notebook с заголовками и пояснениями.

Домашнее задание 3. API, большие данные и ML-решения

1. Подключитесь к публичному API:
 - Сделайте запрос с помощью requests
 - Преобразуйте JSON в pandas.DataFrame
 - Оставьте 5 самых дорогих активов
2. Симулируйте большой датасет (1 млн строк) с помощью numpy и pandas:
 - Используйте Dask или Vaex для агрегации (среднее, сумма)
 - Сравните время выполнения с обычным pandas (на выборке 10k)
3. Примените коробочное решение:
 - Прогнозирование временного ряда цен на биткоин с помощью Prophet
 - Постройте график прогноза на 7 дней вперёд
4. Настройте телеграм-бота (через python-telegram-bot):
 - Отправьте сообщение с текущей средней ценой из API
5. Сохраните код в GitHub, пришлите ссылку.

Примерные задания для теста

Тест 1. Pandas и обработка данных

1. Какой метод используется для удаления строк с пропущенными значениями?
a) fillna()
b) dropna()
c) replace()
d) isnull()

Правильный ответ: b

2. Как объединить два DataFrame по общему столбцу?
a) concat()
b) merge()
c) append()
d) join()

Правильный ответ: b

3. Что делает метод groupby()?
a) Сортирует строки
b) Объединяет строки по значению и применяет агрегацию
c) Удаляет дубликаты
d) Меняет тип данных

Правильный ответ: b

4. Как выбрать столбец price из DataFrame df?
a) df.select("price")
b) df.column["price"]
c) df["price"]
d) df.get_column("price")

Правильный ответ: c

5. Как посмотреть первые 5 строк DataFrame?
a) df.head()
b) df.first(5)
c) df.preview()
d) df.show()

Правильный ответ: a

Тест 2. Визуализация и NumPy

1. Какая библиотека используется для построения гистограмм?
a) numpy
b) requests
c) matplotlib
d) os

Правильный ответ: c

2. Как создать массив из 10 нулей с помощью NumPy?
a) np.zeros(10)
b) np.array(10)
c) np.ones(10)
d) np.empty(10)

Правильный ответ: a

3. Какой метод seaborn строит тепловую карту?
a) sns.lineplot()
b) sns.scatterplot()
c) sns.heatmap()
d) sns.barplot()

Правильный ответ: c

4. Что делает `pr.mean()`?
- Находит максимальное значение
 - Считает среднее арифметическое
 - Считает дисперсию
 - Сортирует массив

Правильный ответ: b

5. Как сделать график интерактивным?
- matplotlib
 - seaborn
 - plotly
 - pandas

Правильный ответ: c

Тест 3. API, большие данные, Машинное обучение (ML)

1. Какая библиотека используется для HTTP-запросов?
- json
 - pandas
 - requests
 - numpy

Правильный ответ: c

2. Что такое когортный анализ?
- Анализ временных рядов
 - Группировка пользователей по дате первого действия и оценка их удержания
 - Прогнозирование цен
 - Визуализация распределения

Правильный ответ: b

3. Какая библиотека подходит для обработки больших данных в Python?
- Dask
 - os
 - math
 - random

Правильный ответ: a

4. Какая модель используется для прогнозирования временных рядов?
- KMeans
 - Prophet
 - LinearRegression
 - RandomForest

Правильный ответ: b

5. Что делает лямбда-функция?
- Создаёт цикл
 - Определяет анонимную функцию в одну строку
 - Импортирует библиотеку
 - Сохраняет файл

Правильный ответ: b

Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	Что выведет код: <code>print(pd.DataFrame({'A': [1][2][3]}).['A'].sum())</code> ?	6	ПК-2

2.	Какой результат выполнения: <code>print(np.array([1][2][3]) * 2)?</code>	[2]	ПК-2
3.	Какой метод Pandas используется для объединения таблиц по ключу?	merge или join	ПК-1
4.	Что выведет код: <code>df.groupby('col')['val'].mean()</code> ?	Среднее значение по группам	ОПК-3
5.	Какая библиотека является стандартом для статической визуализации в Python?	Matplotlib	ПК-1
6.	Какой результат выполнения: <code>np.zeros((2, 3)).shape</code> ?	(2, 3)	ПК-2
7.	Для чего используется класс Pipeline в scikit-learn?	Последовательное применение трансформаций и модели	ОПК-3
8.	Что выведет код: <code>df.fillna(0)</code> при наличии пропусков?	DataFrame с заменой NaN на 0	ПК-2
9.	Какая метрика лучше подходит для несбалансированных классов?	F1-score или ROC-AUC	ПК-1
10.	Какой этап следует за Пониманием данных в CRISP-DM?	Подготовка данных	ОПК-5
11.	Что выведет код: <code>df.loc[0, 'col']</code> ?	Значение ячейки в строке 0, столбце col	ПК-2
12.	Как рассчитывается удержание в когортном анализе?	Активные пользователи / Размер когорты	ПК-1
13.	Какой инструмент используется для интерактивных дашбордов?	Plotly Dash или Streamlit	ОПК-3
14.	Какой результат выполнения: <code>np.arange(0, 5, 2)</code> ?	[0]	ПК-2
15.	Какая задача Data Mining ищет скрытые закономерности?	Классификация, кластеризация, ассоциация	ПК-1
16.	Как избежать утечки данных при валидации?	Разделение до масштабирования	ОПК-3
17.	Что выведет код: <code>df['col'].value_counts().head(1)</code> ?	Самое частое значение и его количество	ПК-2
18.	Какой принцип лежит в основе обработки больших данных MapReduce?	Разделение задачи на отображение и сведение	ОПК-5
19.	Как обнаруживается дрейф данных?	Мониторинг распределения признаков во времени	ПК-1
20.	Зачем используется виртуальное окружение <code>venv</code> или <code>conda</code> ?	Изоляция зависимостей проекта	ОПК-5