

**УТВЕРЖДЕНА**

Решением Ученого совета  
АНО ВО «Центральный университет»  
«24» июня 2025 г.  
Протокол № 2

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

Направление подготовки: 02.04.01 Математика и компьютерные науки

Направленность (профиль) подготовки: Продуктовая аналитика

Квалификация (степень) выпускника: магистр

Форма обучения: очная

Срок освоения программы: 2 года

Год набора: 2025

**Москва  
2025**

## Содержание

1. Краткая характеристика дисциплины (модуля) .....	3
2. Перечень планируемых результатов обучения.....	5
3. Тематический план.....	7
4. Содержание дисциплины (модуля).....	7
5. Учебно-методическое обеспечение .....	8
6. Материально-техническое обеспечение .....	8
7. Методические и оценочные материалы .....	10

## 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 специальности 02.04.01 Математика и компьютерные науки, профиль Продуктовая аналитика, утвержденный приказом Министерства науки и высшего образования Российской Федерации № 810 от 23.08.2017 года.

Изучение дисциплины (модуля) «Python для анализа данных» имеет огромное значение в современном мире, где данные стали ключевым ресурсом для принятия решений в бизнесе, науке и технологиях, позволяя специалистам эффективно обрабатывать большие объемы информации и извлекать из них ценную информацию. Освоение Python и его библиотек, таких как pandas, numpy и matplotlib, открывает широкие возможности для карьерного роста в областях data science, машинного обучения и аналитики, делая выпускников востребованными на рынке труда.

### Место дисциплины (модуля) в структуре образовательной программы

Настоящая дисциплина (модуль) включена в учебный план по программе подготовки магистратуры по направлению 02.04.01 Математика и компьютерные науки, профиль Продуктовая аналитика и входит в обязательную часть Блока 1.

Дисциплина (модуль) изучается на 1 курсе в 1 семестре, доступна для прохождения при условии успешного завершения дисциплины (модуля) «Основы Python».

**Цель изучения дисциплины (модуля):** формирование знаний и навыков в применении языка Python и его специализированных библиотек для эффективной обработки, анализа и визуализации данных в профессиональной деятельности.

### Задачи изучения дисциплины (модуля):

- освоить основные объекты и методы библиотек pandas и numpy для выполнения операций с датафреймами и многомерными массивами;
- научиться применять инструменты matplotlib, seaborn и plotly для создания разнообразных типов графиков и диаграмм;
- изучить способы взаимодействия с внешними источниками данных через API, системы контроля версий Git и базы данных;
- практиковать базовые техники анализа данных, включая факторный анализ, когортный анализ и методы прогнозирования;
- развить навыки использования сред разработки, таких как Jupyter Notebook, JupyterHub и Google Colab, для автоматизации рассылки метрик и решения прикладных задач.

### 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;
- основные методы и библиотеки для работы с API, Git и базами данных;
- базовые методы анализа данных: факторный анализ, когортный анализ, методы прогнозирования.

#### **уметь:**

- обрабатывать данные в pandas и numpy;
- проводить вычисления из линейной алгебры с помощью многомерных объектов numpy;
- строить различные виды графиков с помощью библиотек matplotlib, seaborn и plotly;
- 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:***

- навыками решения прикладных задач с использованием python;
- навыками анализа данных с использованием python;
- навыками подготовки и обработки данных с помощью python;
- навыками выбора корректного типа визуализации и построения визуализаций в python;
- навыками настройки 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 самостоятельно создавать прикладные программные средства на основе современных информационных технологий и сетевых ресурсов, в том числе отечественного производства	ОПК-3.1.	Знает основные принципы программирования, архитектуры программного обеспечения и современные языки программирования, а также особенности отечественных информационных технологий и сетевых ресурсов
		ОПК-3.2.	Умеет разрабатывать прикладные программные средства, используя современные инструменты и технологии, а также интегрировать их с сетевыми ресурсами для решения конкретных задач
		ОПК-3.3.	Имеет практический опыт разработки программных средств, используемых при построении математических моделей в естественных науках
ПК-2.	Способен математически корректно ставить естественнонаучные и прикладные задачи, связанные с анализом данных продуктовой аналитикой	ПК-2.1.	Знает основные методы и подходы к математическому моделированию, а также теоретические основы естественных и прикладных наук, необходимые для корректной формулировки задач в рамках продуктовой аналитики
		ПК-2.2.	Умеет анализировать практические ситуации и формулировать на их основе математические модели, включая выбор адекватных методов решения и формулировку условий задачи
		ПК-2.3.	Имеет практический опыт в разработке и решении математических задач в рамках проектов или научных исследований, где были успешно поставлены и решены естественнонаучные и прикладные задачи
ПК-6.	Способен разрабатывать	ПК-6.1.	Знает основные языки

	программное обеспечение для решения прикладных задач в сфере продуктовой аналитики		программирования, методы разработки программного обеспечения, а также принципы проектирования и архитектуры программных систем, применяемых в продуктовой аналитике
		ПК-6.2.	Умеет анализировать прикладные задачи, разрабатывать алгоритмы и реализовывать их в виде программного обеспечения, используя современные инструменты и технологии, а также проводить тестирование и отладку созданных решений
		ПК-6.3.	Имеет практический опыт разработки программного обеспечения в рамках реальных проектов, включая участие в командах, где были успешно реализованы решения для конкретных прикладных задач

### 3. Тематический план

№ п/п	Наименование раздела дисциплины (модуля)	Трудоемкость, академические часы				ТКУ (текущий контроль успеваемости)
		<i>Очная форма</i>				
		Аудиторная работа		Контроль	Самостояте льная работа	
Лекции	Семинары (практичес кие занятия)					
1	Введение и настройка среды разработки	2	2		10	Домашние задания
2	Pandas	10	10		50	Домашние задания
3	NumPy	6	6		30	Домашние задания
4	Визуализация	4	4		22	Домашние задания
5	API and Virtual Machine	4	4		22	Домашние задания
	<i>Зачет с оценкой</i>			4		Защита проекта
	<b>Итого:</b>	<b>26</b>	<b>26</b>	<b>4</b>	<b>134</b>	
	<i>Объем дисциплины (модуля) (в ак. ч.)</i>	<b>190</b>				
	<i>Объем дисциплины (модуля) (в зач. ед.)</i>	<b>5</b>				

### 4. Содержание дисциплины (модуля)

№п/п	Наименование раздела дисциплины (модуля)	Содержание дисциплины (модуля) по темам
1	Введение и настройка среды разработки	Онбординг
2	Pandas	Pandas Факторный анализ Маппинг и лямбда-функции Когортный анализ
3	NumPy	NumPy Прогнозирование с NumPy
4	Визуализация	Matplotlib Seaborn & Plotly
5	API and Virtual Machine	API

## 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. Масис, С. Интерпретируемое машинное обучение на Python : практическое руководство / С. Масис. - Санкт-Петербург : БХВ-Петербург, 2023. - 640 с. - ISBN 978-5-9775-1735-5. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2123391>.

3. 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>.

### *Дополнительная литература:*

1. Васильев, Ю. Python для data science : практическое руководство / Ю. Васильев. - Санкт-Петербург : Питер, 2023. - 272 с. - (Библиотека программиста). - ISBN 978-5-4461-2392-6. - Текст : электронный. - URL: <https://znanium.com/catalog/product/2123381>.

2. Ричардс, Т. Streamlit для Data Science. Создаем интерактивные приложения в Python : практическое руководство / Т. Ричардс, А. Груздев ; пер. с англ. А. В. Груздева. – Москва : ДМК Пресс, 2024. - 356 с. – ISBN 978-5-93700-275-4. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2205064>.

## 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	<a href="https://elibrary.ru/defaultx.asp">https://elibrary.ru/defaultx.asp</a>
2.	База данных для IT-специалистов	<a href="https://habr.com">https://habr.com</a>
3.	База данных ScienceDirect	<a href="https://www.sciencedirect.com">https://www.sciencedirect.com</a>
4.	Официальный сайт Министерства науки и высшего образования Российской Федерации	<a href="https://minobrnauki.gov.ru/">https://minobrnauki.gov.ru/</a>
5.	Федеральный портал «Российское образование»	<a href="https://www.edu.ru/">https://www.edu.ru/</a>
6.	Информационная система "Единое окно доступа к образовательным ресурсам"	<a href="http://window.edu.ru/">http://window.edu.ru/</a>
7.	Единая коллекция цифровых образовательных ресурсов	<a href="http://school-collection.edu.ru/">http://school-collection.edu.ru/</a>
8.	Федеральный центр информационно - образовательных ресурсов	<a href="http://fcior.edu.ru/">http://fcior.edu.ru/</a>

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
<b>Операционные системы:</b>		
Microsoft Imagine (Windows Client, Server)	зарубежное	лицензионное
<b>Браузеры:</b>		
Яндекс.Браузер	отечественное	свободно распространяемое
Google Chrome	зарубежное	свободно распространяемое
<b>Офисные приложения:</b>		
Microsoft Imagine (Visio, OneNote)	зарубежное	лицензионное
TeXstudio	зарубежное	свободно распространяемое
Adobe Acrobat Reader	зарубежное	свободно распространяемое
<b>Программное обеспечение для планирования и учета времени:</b>		
Toggle app	зарубежное	свободно распространяемое
<b>Системы управления проектами:</b>		
Microsoft Imagine (Project)	зарубежное	лицензионное
<b>Системы управления базами данных:</b>		
Microsoft Imagine (SQL Server)	зарубежное	лицензионное
<b>Системы резервного копирования (backup):</b>		
Acronis Backup Advanced for HyperV	зарубежное	лицензионное
<b>Справочно-правовые системы:</b>		

КонсультантПлюс: справочно-правовая система	отечественное	лицензионное
<b>Средства антивирусной защиты:</b>		
Kaspersky Endpoint Security для бизнеса Стандартный Russian Edition	отечественное	лицензионное
<b>Среды разработки:</b>		
Visual Studio Code	зарубежное	свободно распространяемое
Bash (Unix shell)	зарубежное	свободно распространяемое
Anaconda	зарубежное	свободно распространяемое
Robotic Operating System	зарубежное	свободно распространяемое
CopelliaSim	зарубежное	свободно распространяемое
Google Colaboratory	зарубежное	свободно распространяемое
<b>Пакеты программных средств и библиотек:</b>		
AutoPsy	зарубежное	свободно распространяемое
Interactive Disassembler (IDA)	зарубежное	свободно распространяемое
<b>Системы управления библиографической информацией:</b>		
Zotero	зарубежное	свободно распространяемое
<b>Сервисы и службы:</b>		
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, проект, 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.

*Проект* – исследовательская работа по дисциплине (модулю) и презентация результатов.

Для успешной подготовки к проекту рекомендуется: четко определить цели и задачи проекта; составить план работы, разбив проект на этапы с указанием сроков выполнения каждого из них; использовать разнообразные источники информации и инструменты для исследования темы; регулярно проверять прогресс и вносить коррективы в план, если это необходимо.

*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 *зачета с оценкой*, 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:

<b>Десятибалльная оценка</b>	<b>Пятибалльная оценка</b>	<b>Оценка за зачет</b>	<b>Общая характеристика результата обучения по дисциплине (модулю)</b>
10	Отлично	Зачтено	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 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
9	Отлично	Зачтено	
8	Отлично	Зачтено	

Десятибалльная оценка	Пятибалльная оценка	Оценка за зачет	Общая характеристика результата обучения по дисциплине (модулю)
			практической деятельности, и умеет связывать теоретические аспекты дисциплины (модуля) с практическими задачами.
7	Хорошо	Зачтено	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	Зачтено	
5	Удовлетворительно	Зачтено	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	Зачтено	
3	Не сдан	Не зачтено	Студент не овладел обязательным минимумом знаний по предмету и не может ответить на вопросы, даже если преподаватель задает дополнительные наводящие вопросы.
2	Не сдан	Не зачтено	
1	Не сдан	Не зачтено	

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

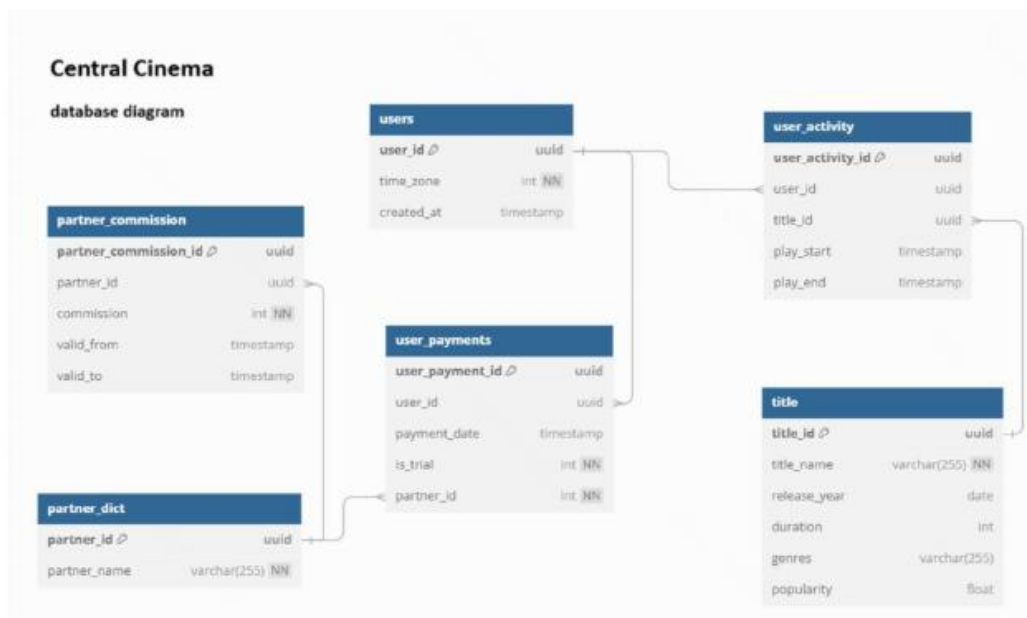
Активность	Вес	Количество	Описание
Домашние задания	80%	7	Набор задач по темам недели
Зачет с оценкой	20%	1	Защита проекта – исследовательская работа по дисциплине (модулю) и презентация результатов

**Формула расчёта итоговой оценки по дисциплине (модулю) «Python для анализа данных»:** « $0,8 \times$  среднее за домашние задания +  $0,2 \times$  зачет с оценкой».

**Текущий контроль успеваемости обучающихся по дисциплине (модулю)**

**Примерные домашние задания**

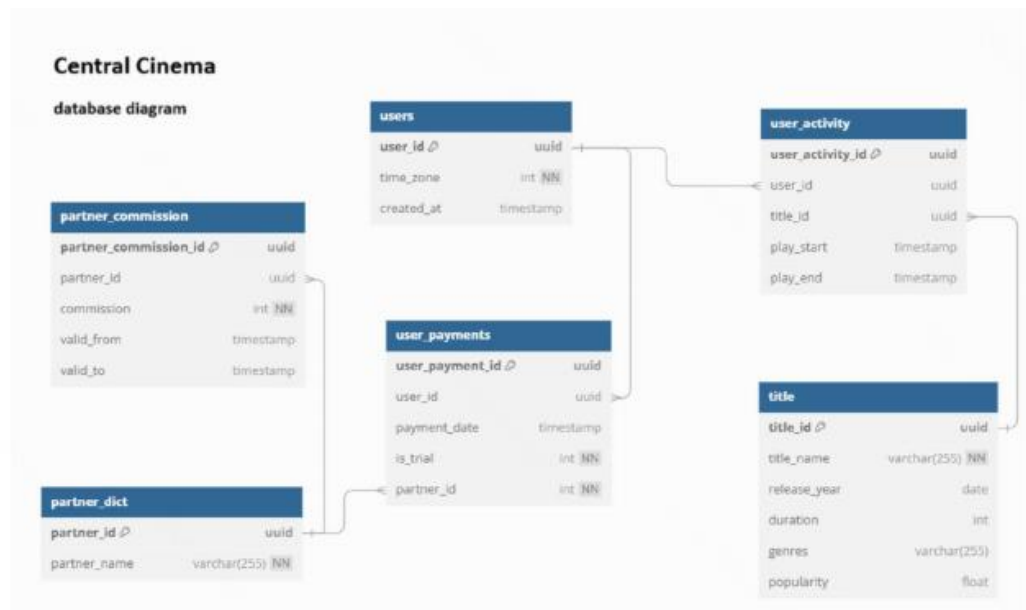
**Домашнее задание 1**



### Описание таблиц

<b>users</b>	<b>справочник Пользователи</b>
user_id	уникальный идентификатор клиента
time_zone	тайм-зона клиента
created_at	дата и время регистрации
<b>user_payments</b>	<b>сведения об оплатах</b>
user_payment_id	уникальный идентификатор платежа
user_id	идентификатор клиента
payment_date	дата оплаты
is_trial	признак триальности (подписка в подарок)
partner_id	идентификатор партнера, который продал подписку
<b>user_activity</b>	<b>сведения о просмотрах фильмов</b>
user_activity_id	уникальный идентификатор просмотра
user_id	идентификатор клиента
title_id	идентификатор фильма
play_start	дата и время начала просмотра
play_end	дата и время окончания просмотра
<b>partner_dict</b>	<b>справочник Партнеры (кто продал подписку)</b>
partner_id	уникальный идентификатор Партнера
partner_name	наименование
<b>partner_commission</b>	<b>справочник Стоимость подписки</b>
partner_commission_id	уникальный идентификатор
partner_id	идентификатор Партнера
commission	стоимость подписки в месяц в рублях
valid_from	дата действия цены С
valid_to	дата действия цены ПО
<b>title</b>	<b>справочник Фильмы</b>
title_id	уникальный идентификатор Фильма
title_name	наименование
release_year	дата выхода
duration	продолжительность в минутах
genres	жанры
popularity	популярность

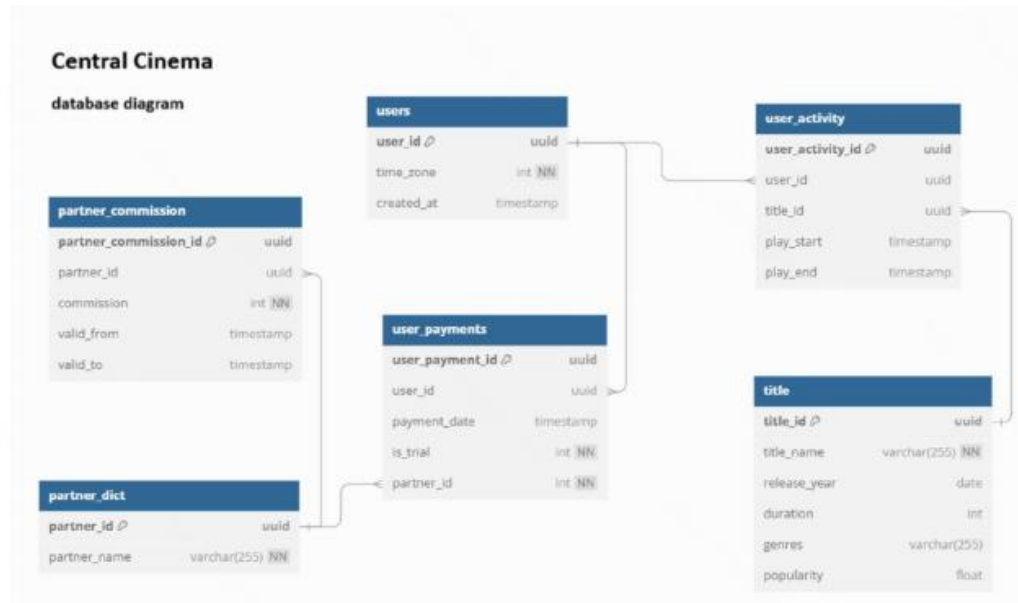
## Домашнее задание 2



### Описание таблиц

<b>users</b>	<b>справочник Пользователи</b>
<b>user_id</b>	уникальный идентификатор клиента
<b>time_zone</b>	тайм-зона клиента
<b>created_at</b>	дата и время регистрации
<b>user_payments</b>	<b>сведения об оплатах</b>
<b>user_payment_id</b>	уникальный идентификатор платежа
<b>user_id</b>	идентификатор клиента
<b>payment_date</b>	дата оплаты
<b>is_trial</b>	признак триальности (подписка в подарок)
<b>partner_id</b>	идентификатор партнера, который продал подписку
<b>user_activity</b>	<b>сведения о просмотрах фильмов</b>
<b>user_activity_id</b>	уникальный идентификатор просмотра
<b>user_id</b>	идентификатор клиента
<b>title_id</b>	идентификатор фильма
<b>play_start</b>	дата и время начала просмотра
<b>play_end</b>	дата и время окончания просмотра
<b>partner_dict</b>	<b>справочник Партнеры (кто продал подписку)</b>
<b>partner_id</b>	уникальный идентификатор Партнера
<b>partner_name</b>	наименование
<b>partner_commission</b>	<b>справочник Стоимость подписки</b>
<b>partner_commission_id</b>	уникальный идентификатор
<b>partner_id</b>	идентификатор Партнера
<b>commission</b>	стоимость подписки в месяц в рублях
<b>valid_from</b>	дата действия цены С
<b>valid_to</b>	дата действия цены ПО
<b>title</b>	<b>справочник Фильмы</b>
<b>title_id</b>	уникальный идентификатор Фильма
<b>title_name</b>	наименование
<b>release_year</b>	дата выхода
<b>duration</b>	продолжительность в минутах
<b>genres</b>	жанры
<b>popularity</b>	популярность

## Домашнее задание 3



### Описание таблиц

<b>users</b>	<b>справочник Пользователи</b>
user_id	уникальный идентификатор клиента
time_zone	тайм-зона клиента
created_at	дата и время регистрации
<b>user_payments</b>	<b>сведения об оплатах</b>
user_payment_id	уникальный идентификатор платежа
user_id	идентификатор клиента
payment_date	дата оплаты
is_trial	признак триальности (подписка в подарок)
partner_id	идентификатор партнера, который продал подписку
<b>user_activity</b>	<b>сведения о просмотрах фильмов</b>
user_activity_id	уникальный идентификатор просмотра
user_id	идентификатор клиента
title_id	идентификатор фильма
play_start	дата и время начала просмотра
play_end	дата и время окончания просмотра
<b>partner_dict</b>	<b>справочник Партнеры (кто продал подписку)</b>
partner_id	уникальный идентификатор Партнера
partner_name	наименование
<b>partner_commission</b>	<b>справочник Стоимость подписки</b>
partner_commission_id	уникальный идентификатор
partner_id	идентификатор Партнера
commission	стоимость подписки в месяц в рублях
valid_from	дата действия цены С
valid_to	дата действия цены ПО
<b>title</b>	<b>справочник Фильмы</b>
title_id	уникальный идентификатор Фильма
title_name	наименование
release_year	дата выхода
duration	продолжительность в минутах
genres	жанры
popularity	популярность

## Примерные задания по проекту

### Задание 1. Расширение отчётности

Текущую отчетность необходимо обогатить новыми метриками, пользуясь доступными данными.

- Придумай минимум две дополнительные метрики, которые можно рассчитать на имеющихся данных, и предоставь код на Python с вычислением данных метрик на имеющихся таблицах.

- Добавь эти метрики в текущую версию репортинга с телеграм-ботом.

#### В результате

- Предоставь `irunb` файл с описанием и реализацией новых метрик.
- Обнови репортинг через телеграм таким образом, чтобы расчёт и визуализация новых метрик там появились.

#### Критерии

1. Описано (словами в маркдаунах) минимум две метрики, которые имеют отношение к бизнесу онлайн-кинотеатра (1 балл за каждую метрику, максимум 2)

2. Для данных метрик представлены примеры кода на Python для их расчёта в динамике (1 балл за каждую метрику, максимум 2)

1. Сравнительные значения данных метрик, а также графики рассылаются телеграм-ботом (1 балл)

### Задание 2. Презентация результатов

Необходимо презентовать результаты вашей работы

- Сделай слайды `pptx` с демонстрацией результатов твоей работы, а также с демонстрацией

финансовых и продуктовых результатов онлайн-кинотеатра (не забудь использовать материалы из предыдущего задания).

- Сними видео (OBS/Loom/etc.) со скринкастом слайдов и устным объяснением содержимого на 5-7 минут.

#### В результате

- Предоставь слайды в формате `pptx`
- Предоставь видео с объяснением слайдов

#### Критерии

1. Длина видео незначительно отличается от установленного коридора 5-7 мин. Ваше лицо не обязано быть на видео, но с ним, конечно же, лучше. (1.5 балла)

2. Слайды оформлены с использованием визуализаций и должны объяснять выполненную работу и бизнес онлайн-кинотеатра. (1.5 балла)

3. Первое задание, наравне с другими метриками, учтено в презентации и в видео (0.5 балла)

### Задание 3. Техническое задание для DWH

#### Общая рекомендация по сдаче задания

Необходимо обратиться к дата инженерам и запросить доп. данные

Придумай метрики, которые невозможно рассчитать на имеющихся таблицах, но которые были бы полезны для мониторинга

#### В результате

Предоставь текст (не более 2000 знаков) с описанием новых метрик и данных (витрин), которые нужны для их расчета

#### Критерии

1. В тексте указаны необходимые таблицы (0.75 балла)

- Набор колонок
- Тип данных
- Смысловое наполнение колонок
- Временной промежутков, который хранится в таблице
- Частоту обновления таблицы

2. В тексте названы и описаны метрики, которые будут строиться на новых данных, а также описан их бизнес-смысл и интерпретация (0.75 балла)

### Задания для промежуточной аттестации по дисциплине (модулю)

№ п/п	Задание	Ответ	Компетенция
1.	<p>Сопоставьте типы данных и их описания:</p> <p>1 - list (список)                  2 - tuple (кортеж)                  3 - set (множество)                  4 - dict (словарь)</p> <p>А - Изменяемый тип данных, который определяется парами ключ-значение                  Б - Изменяемый тип данных, который может хранить разные типы данных и использует квадратные скобки                  В - Неизменяемый тип данных, который может хранить разные типы данных и использует круглые скобки                  Г - Изменяемый тип данных, который хранит только уникальные значения</p> <p><i>Ответ запиши в последовательности букв без пробелов и знаков препинания, пример: АБВГ.</i></p>	<p>БВГА</p> <p>1Б 2В 3Г 4А</p> <p>1=Б 2=В 3=Г 4=А</p> <p>1б, 2в, 3г, 4а</p> <p>1Б/2В/3Г/4А</p>	ПК-6
2.	<p>Пусть есть следующий датафрейм:</p> <pre>df = pd.DataFrame({'name':['Иванов', 'Петров', 'Иванов'], 'amt_debt':[50,200,300]})</pre> <p>Какое из следующих выражений отработает без ошибки?</p> <p>1 - df[df['name']=='Иванов' &amp; df['amt_debt']&gt;100]                  2 - df[df['name']== 'Иванов' &amp; df['amt_debt']&gt;100]                  3 - df[(df['name']=='Иванов') &amp; (df['amt_debt']&gt;100)]                  4 - df[(df['name']== 'Иванов') &amp; (df['amt_debt']&gt;100)]</p>	4	ОПК-3
3.	<pre>df = pd.DataFrame({'id_client':[111,112,113], 'name':['Иванов', 'Петров', 'Иванов'], 'amt_debt':[50,200,300]}) df.set_index('id_client', inplace = True)</pre> <p>Какие из следующих выражений не вернут ошибку?                  Укажите подходящие ответы через запятую без пробелов.</p> <p>1 - df.loc[0, 'name']                  2 - df.iloc[0, 'name']                  3 - df.loc[111, 'name']                  4 - df.iloc[0,1]</p>	3,4 / 4,3	ПК-6
4.	<p>Какое из следующих высказываний верно относительно выражения:</p> <pre>df.groupby('name').agg(all_debt = ('amt_debt','sum')).reset_index()</pre> <p>1 - Мерами группировки являются name и amt_debt                  2 - Измерением группировки является name</p>	2 / два / второй / второе	ОПК-3

	3 - В результирующем датафрейме будет три столбца: name, sum и all debt		
5.	<p>В какой последовательности должны быть расставлены следующие строки, чтобы код отработал?</p> <p>Укажите цифры в нужной последовательности через тире, без пробелов.</p> <ol style="list-style-type: none"> <li>1. <code>def plus_one_func(x):</code>     <code>return x + 2 + b</code></li> <li>2. <code>d = c + 15</code></li> <li>3. <code>c = plus_one_func(b)</code></li> <li>4. <code>b = 7</code></li> </ol>	4-1-3-2	ПК-6
6.	<pre>a = 4 b = -5  if (a==3 &amp; (b &gt; 2   b &lt; -2)) &amp; a &gt; 3.5:     my_int = 10 else:     my_int = 12  print(my_int)</pre> <p>Что выдаст данный код? (<i>ответ запиши целым числом с 0 знаков после запятой</i>)</p>	12 / 12.0 / 12,0	ОПК-3
7.	<pre>df = pd.DataFrame({'name':['Иванов', 'Петров', 'Иванов'], 'amt_debt':[50,200,300]}) df[df['name']=='Иванов']['amt_debt'].mean()</pre> <p>Что выдаст данный код? (<i>ответ запиши целым числом с 0 знаков после запятой</i>)</p>	175 / 175.0 / 175,0	ПК-6
8.	<pre>import numpy as np a = np.arange(12).reshape((3, 4)) a[1,3]</pre> <p>Что выдаст данный код? (<i>ответ запиши целым числом с 0 знаков после запятой</i>)</p>	7 / 7.0 / 7,0	ОПК-3
9.	<pre>import numpy as np np.array([[1, 2, 3], [4, 5, 6]]).shape[0] + np.array([[1, 2], [3, 4], [5, 6]]).shape[1]</pre> <p>Что выдаст данный код? (<i>ответ запиши целым числом с 0 знаков после запятой</i>)</p>	4 / 4.0 / 4,0	ПК-6

10.	<pre>import numpy as np a = (2, 3) x = np.ones(a) y = np.zeros(a) z = np.full(a, a[0]) print(sum(sum(x+y+z)))</pre> <p>Что выдаст данный код? (ответ запиши целым числом с 0 знаков после запятой)</p>	18 / 18.0 / 18,0	ОПК-3
11.	<pre>import matplotlib.pyplot as plt gridsize = (3, 2) fig = plt.figure(figsize=(10, 6)) ax1 = plt.subplot2grid(gridsize, (0, 0), colspan=X, rowspan=Y) ax2 = plt.subplot2grid(gridsize, (2, 0)) ax3 = plt.subplot2grid(gridsize, (2, 1))</pre> <p>Чему должны быть равны X и Y в данном коде, чтобы на дашборде получилось три графика, не пересекающихся друг с другом, а также без свободных зон на дашборде? <b>В ответе укажите сумму чисел X и Y.</b></p>	4 / 4.0 / 4,0	ОПК-3
12.	<p>Наша суммарная выручка формируется как произведение цены продукта на количество проданной продукции.</p> <p>В прошлом месяце мы продали 120 единиц товара по 150 рублей каждая.</p> <p>В этом месяце мы продали 140 единиц товара по 160 рублей каждая.</p> <p>С помощью факторного анализа рассчитайте, какую долю от общего изменения выручки занимает изменение цены товара.</p> <p><b>Результат округлите до целого процента (до 0 знаков после запятой в процентном выражении, например, ответ 0.1759 приводится к формату 18).</b></p>	30 / 30.0 / 30,0	ПК-2
13.	<p>В сентябре 2023 года на нашу платформу в платную подписку зашло 15300 пользователей.</p> <ul style="list-style-type: none"> <li>- Две трети из них продлились на следующий период</li> <li>- Из них 8000 пользователей продлились затем на третий период</li> <li>- От этих 8000 пользователей на четвёртый период продлилась половина</li> <li>- На пятый период не продлился никто</li> </ul> <p>Чему равен лайфтайм пользователя из данной когорты (в месяцах)? <b>Округлите ответ до двух знаков после запятой.</b></p>	2.45 / 2,45	ПК-2
14.	<p>Допустим, лайфтайм клиента составляет 3.5 месяца.</p> <p>Стоимость подписки за один месяц составляет</p>	200 / 200.0 / 200,0	ПК-2

	<p>100 рублей, а затраты на одного клиента (за весь период его жизни) составляют 150 рублей. Сколько составляет LTV (т.е. Lifetime Revenue за вычетом выручки) из расчёта на одного клиента?</p> <p><b>Ответ округлите до нуля знаков после запятой.</b></p>		
15.	<p>Каждый день с понедельника по субботу включительно на сервис заходила одна и та же тысяча пользователей. В воскресенье из этой тысячи не зашли пятеро, но зашли ещё трое пользователей, которых не было в первые шесть дней. Чему равно значение метрики WAU (Weekly Active Users) за эту неделю?</p>	1003 / 1003.0 / 1003,0	ПК-2
16.	<p>DAU = 4200 WAU = 7600 MAU = 22000 Чему равны Sticky Factor Weekly и Sticky Factor Monthly? В качестве ответа укажите сумму этих двух значений.</p> <p><b>Ответ округлите до двух знаков после запятой.</b></p>	0.74 / 0,74	ПК-2