

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«Национальный исследовательский ядерный университет «МИФИ»
Обнинский институт атомной энергетики –
филиал федерального государственного автономного образовательного учреждения высшего образования
«Национальный исследовательский ядерный университет «МИФИ»
(ИАТЭ НИЯУ МИФИ)

Утверждено на заседании
УМС ИАТЭ НИЯУ МИФИ
Протокол №2-8/2024 От 30.08.2024

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Продвинутый SQL для работы с большими данными
Шифр, название дисциплины

01.04.02 «Прикладная математика и информатика»
Шифр, название специальности/направления подготовки

Математическое моделирование и прикладной анализ данных
Название программы магистратуры

магистр
(Квалификация (степень) выпускника)

Форма обучения: очная

г. Обнинск 2024 г.

Программа составлена в соответствии с требованиями образовательного стандарта высшего образования национального исследовательского ядерного университета «МИФИ» по направлению подготовки 01.04.02 – Прикладная математика и информатика. (квалификация (степень) магистр).

Программу составил:

_____ С.В. Ермаков, доцент, к.ф.-м.н, доцент

Рецензент:

_____ Г.Е. Деев, доцент, к.ф.-м.н, доцент

Программа рассмотрена на заседании ОИКС

(протокол № 5/7 от «30» июля от 2024 г.)

Руководитель направления подготовки 01.04.02
«Прикладная математика и информатика»

_____ Ермаков С.В.

« ____ » _____ 2024 г.

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

В результате освоения ООП магистратуры обучающийся должен овладеть следующими результатами обучения по дисциплине:

Коды компетенций	Результаты освоения ООП Содержание компетенций*	Перечень планируемых результатов обучения по дисциплине**
УК-2	Способен управлять проектом на всех этапах его жизненного цикла	З-УК-2 Знать: этапы жизненного цикла проекта; этапы разработки и реализации проекта; методы разработки и управления проектами У-УК-2 Уметь: разрабатывать проект с учетом анализа альтернативных вариантов его реализации, определять целевые этапы, основные направления работ; объяснить цели и сформулировать задачи, связанные с подготовкой и реализацией проекта; управлять проектом на всех этапах его жизненного цикла В-УК-2 Владеть: методиками разработки и управления проектом; методами оценки потребности в ресурсах и эффективности проекта
ОПК-4	Способен комбинировать и адаптировать существующие информационно-коммуникационные технологии для решения задач в области профессиональной деятельности с учетом требований информационной безопасности	З-ОПК-4 Знать основные методики и технологии использования ИКТ в профессиональной деятельности с учетом основных требований информационной безопасности. У-ОПК-4 Уметь решать типовые задачи профессиональной деятельности с использованием ИКТ, комбинировать и адаптировать существующие ИКТ для решения задач в области профессиональной деятельности с учетом основных требований информационной безопасности. В-ОПК-4 Владеть навыками использования и адаптации ИКТ в профессиональной деятельности с учетом основных требований информационной безопасности.

2. Место дисциплины в структуре ООП магистратуры

Дисциплина реализуется в рамках общенаучного модуля.

Для освоения дисциплины необходимы компетенции, сформированные в рамках изучения следующих дисциплин: SQL для анализа данных, Clickhouse и хранилища данных DWH

Дисциплина изучается на 1 курсе в 2 семестре.

3. Объем дисциплины в зачетных единицах с указанием количества академических часов, выделенных на контактную работу обучающихся с преподавателем (по видам занятий) и на самостоятельную работу обучающихся

Общая трудоемкость (объем) дисциплины составляет 2 зачетных единиц (з.е.), 72 академических часа.

3.1. Объем дисциплины по видам учебных занятий (в часах)

	Семестр		
	№ 1	№ 2	Всего
	Количество часов на вид работы:		
Контактная работа обучающихся с преподавателем			
Аудиторные занятия (всего)		64	64
В том числе:			
лекции		32	32
практические занятия		32	32
лабораторные занятия			
Промежуточная аттестация			
В том числе:			
зачет		Зачет	Зачет
экзамен		-	-
Самостоятельная работа обучающихся (всего)		40	40
В том числе:			
проработка учебного (теоретического) материала		5	5
выполнение индивидуальных заданий		5	5
подготовка ко всем видам контрольных испытаний текущего контроля успеваемости (в течение семестра)		10	10
подготовка ко всем видам контрольных испытаний промежуточной аттестации (по окончании семестра)		20	20
Всего (часы):		72	72
Всего (зачетные единицы):		2	2

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

4.1. Разделы дисциплины и трудоемкость по видам учебных занятий (в академических часах)

№ п/п	Наименование раздела /темы дисциплины	Общая трудоёмкость всего (в часах)	Виды учебных занятий, включая самостоятельную работу обучающихся и трудоемкость (в часах)	Формы текущего контроля успеваемости
-------	---------------------------------------	------------------------------------	---	--------------------------------------

			Аудиторные учебные занятия			СРО	
			Лек	Сем/Пр	Лаб		
1.			16	16	-	40	
1.1.	Введение в PL/pgSQL	7	2	2		3	
1.2.	Основные возможности PL/pgSQL	7	2	2	-	3	
1.3.	Анонимные блоки и переменные	7	2	2	-	3	
1.4.	Условные операторы и циклы	7	2	2	-	3	
1.5.	Курсоры	7	2	2		3	
1.6.	Функции и процедуры	7	2	2	-	3	
1.7.	Обработка ошибок	5	1	1	-	3	
1.8.	Триггеры	6	1	1	-	4	
1.9.	ddl базы	6	1	1		4	
1.10.	Итоговый проект	6	1	1		8	

4.2. Содержание дисциплины, структурированное по разделам (темам)

Лекционный курс

№	Наименование раздела /темы дисциплины	Содержание
1.1.	Введение в PL/pgSQL	Обзор языка PL/pgSQL, его роль в PostgreSQL для написания хранимых процедур, функций и триггеров, а также его преимущества для оптимизации работы с данными и выполнения логики на уровне базы данных.
1.2.	Основные возможности PL/pgSQL	Интеграция с SQL, работа с переменными, выражениями и типами данных. Основные особенности синтаксиса и возможности языка для создания сложных операций в базе данных.
1.3.	Анонимные блоки и переменные	Понимание анонимных блоков — фрагментов кода, не привязанных к именованным объектам. Объявление и использование переменных для хранения и управления состоянием в процессе выполнения.
1.4.	Условные операторы и циклы	Использование условных операторов (IF, CASE) для ветвления логики и циклических конструкций (LOOP, WHILE, FOR) для многократного выполнения кода в зависимости от условий.
1.5.	Курсоры	Работа с курсорами для построчной обработки результатов SQL-запросов, их типы, создание и использование для обхода больших наборов данных.
1.6.	Функции и процедуры	Различие между функциями и процедурами, передача параметров, возвращаемые значения, использование для инкапсуляции бизнес-логики в PL/pgSQL.
1.7.	Обработка ошибок	Перехват и обработка исключений с помощью конструкции BEGIN ... EXCEPTION, обеспечение стабильности работы программы при возникновении ошибок.
1.8.	Триггеры	Создание и использование триггеров для автоматического выполнения действий при изменении данных в базе, их типы

		и механизмы привязки к таблицам и событиям.
1.9	ddl базы	Работа с DDL-операциями (создание, изменение, удаление объектов базы данных) через PL/pgSQL, использование динамического SQL для изменения структуры базы данных во время выполнения.
1.10.	Итоговый проект	Итоговый проект

Практические/семинарские занятия

№	Наименование раздела /темы дисциплины	Содержание
1.1.	Введение в PL/pgSQL	Обзор языка PL/pgSQL, его роль в PostgreSQL для написания хранимых процедур, функций и триггеров, а также его преимущества для оптимизации работы с данными и выполнения логики на уровне базы данных.
1.2.	Основные возможности PL/pgSQL	Интеграция с SQL, работа с переменными, выражениями и типами данных. Основные особенности синтаксиса и возможности языка для создания сложных операций в базе данных.
1.3.	Анонимные блоки и переменные	Понимание анонимных блоков — фрагментов кода, не привязанных к именованным объектам. Объявление и использование переменных для хранения и управления состоянием в процессе выполнения.
1.4.	Условные операторы и циклы	Использование условных операторов (IF, CASE) для ветвления логики и циклических конструкций (LOOP, WHILE, FOR) для многократного выполнения кода в зависимости от условий.
1.5.	Курсоры	Работа с курсорами для построчной обработки результатов SQL-запросов, их типы, создание и использование для обхода больших наборов данных.
1.6.	Функции и процедуры	Различие между функциями и процедурами, передача параметров, возвращаемые значения, использование для инкапсуляции бизнес-логики в PL/pgSQL.
1.7.	Обработка ошибок	Перехват и обработка исключений с помощью конструкции BEGIN ... EXCEPTION, обеспечение стабильности работы программы при возникновении ошибок.
1.8.	Триггеры	Создание и использование триггеров для автоматического выполнения действий при изменении данных в базе, их типы и механизмы привязки к таблицам и событиям.
1.9	ddl базы	Работа с DDL-операциями (создание, изменение, удаление объектов базы данных) через PL/pgSQL, использование динамического SQL для изменения структуры базы данных во время выполнения.
1.10.	Итоговый проект	Итоговый проект

Лабораторные занятия

Не предусмотрены.

5. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

В качестве учебно-методических материалов используется рекомендованная литература.

6. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

6.1. Паспорт фонда оценочных средств по дисциплине

№ п/п	Контролируемые разделы (темы) дисциплины (результаты по разделам)	Код контролируемой компетенции (или её части) / и ее формулировка	Наименование оценочного средства
1.3	Функции и процедуры	УК-2	Контрольная работа № 1
2.1.	Условные операторы и циклы	УК-	Контрольная работа № 1
2.3.	Триггеры	ОПК-4:	Контрольная работа № 2
2.4.	Условные операторы и циклы	ОПК-4.	Контрольная работа № 2

6.2. Типовые контрольные задания или иные материалы

6.2.1. Зачет

В билете два теоретических вопроса и один практический

Теоретические вопросы билета:

1. Какие основные возможности PL/pgSQL позволяют интегрировать его с SQL?
2. Какие типы данных поддерживаются в PL/pgSQL?
3. Как PL/pgSQL позволяет улучшить производительность работы с данными?
4. Что такое анонимный блок в PL/pgSQL?
5. Как объявляются и используются переменные в анонимных блоках?
6. В чем отличие анонимных блоков от функций и процедур?
7. Как в PL/pgSQL реализуются условные операторы? Приведите примеры.
8. Какие циклические конструкции поддерживаются в PL/pgSQL? Как они отличаются друг от друга?
9. Как использовать оператор **CASE** в PL/pgSQL для реализации сложных условий?
10. Что такое курсоры в PL/pgSQL и как они используются для обработки данных?
11. Какие существуют типы курсоров в PL/pgSQL?
12. Как объявляются и закрываются курсоры в PL/pgSQL?
13. В чем отличие между функциями и процедурами в PL/pgSQL?
14. Как передаются параметры в функции и процедуры?
15. Как возвращаются значения из функции в PL/pgSQL?
16. Как осуществляется обработка ошибок и исключений в PL/pgSQL?
17. В чем заключается отличие между конструкциями **EXCEPTION** и **RAISE**?
18. Как реализуется логирование ошибок в PL/pgSQL?
19. Что такое триггер в PL/pgSQL и как он работает?
20. Какие типы триггеров существуют в PL/pgSQL?

21. Как создаются триггеры для обработки событий на уровне таблицы?
22. Как в PL/pgSQL выполняются DDL-операции (создание, изменение и удаление объектов)?
23. Как используется динамическое SQL для изменения структуры базы данных?
24. Какие особенности работы с DDL-операциями в контексте PL/pgSQL?

Критерий оценки – правильность и полнота ответа на вопросы. Оценка выставляется по шкале от 0 до 40 баллов: теоретические вопросы –30 баллов, 10 баллов– дополнительные вопросы. Зачет считается сданным при оценке не ниже 25 баллов.

6.2.2. Контрольная работа № 1

Напишите SQL-запрос(опираясь на данные), который должен вернуть следующий результат: Для каждого департамента, в котором есть хотя бы один сотрудник, выведите:

1. Название департамента
2. Общее количество сотрудников в департаменте
3. Среднюю зарплату сотрудников в департаменте
4. Средний стаж сотрудников в департаменте (в годах) — стаж считается с даты приёма на работу (используйте текущую дату для расчёта).

Дополнительные требования:

Результат должен быть отсортирован по названию департамента в алфавитном порядке. Если в департаменте нет сотрудников, он не должен попадать в результат.

6.2.2. Контрольная работа № 2

Напишите SQL-запрос, (опираясь на данные), который должен вернуть следующие данные:

1. Для каждого клиента, который сделал хотя бы один заказ:
 - customer_id
 - Общее количество заказов (количество записей в таблице orders для каждого клиента)
 - Общая сумма покупок (сумма total_amount всех заказов этого клиента)
 - Средний чек (средняя сумма заказа для этого клиента)
 - Самая дорогая покупка (максимальная сумма одного заказа этого клиента)
2. Для каждого клиента, включите только те заказы, которые были сделаны в период с 2024-01-10 по 2024-02-10.
3. Отсортируйте результат по customer_id.

б) критерии оценивания компетенций (результатов) – правильная работа кода программы, понимание алгоритма метода оптимизации, умение вывести необходимые для алгоритма формулы.

в) описание шкалы оценивания:

Каждая задача оценивается по шкале от 0 до 10 баллов.

Контрольная работа считается выполненной успешно при суммарной оценке не ниже 18 баллов.

6.3. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

Форма аттестации	Наименование оценочного средства	Баллы
Зачет (100 баллов)	Контрольная работа № 1	30
	Контрольная работа № 2	30
	Ответы на билет	40

7. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины

1. Дэвид М. Берк и Джеффри А. Эдвардс — "PostgreSQL: Up and Running" — 2020 — 300 страниц
2. Кристофер Д. Дей и Хэнк У. Куйпер — "SQL Performance Explained" — 2019 — 350 страниц
3. К. Р. Твэйн — "Beginning PostgreSQL on the Cloud" — 2021 — 250 страниц
4. Алессандро К. — "Mastering PostgreSQL in Application Development" — 2021 — 500 страниц

б) дополнительная учебная литература:

1. Иванов С. — "PL/pgSQL для разработчиков" — 2022 — 280 страниц
 2. Роберт М. Лифшиц — "PostgreSQL for Data Architects" — 2019 — 420 страниц
 3. Хуан Р. — "PostgreSQL 13 Administration Cookbook" — 2021 — 350 страниц
 4. Ф. Ромеро — "PostgreSQL: от начинающего до эксперта" — 2021 — 320 страниц
 5. Томас К. — "PostgreSQL Server Programming" — 2020 — 400 страниц
- Грег С. — "Advanced SQL: Understanding SQL in PostgreSQL" — 2019 — 300 страниц

8. Перечень ресурсов* информационно-телекоммуникационной сети «Интернет» (далее - сеть «Интернет»), необходимых для освоения дисциплины

-

9. Методические указания для обучающихся по освоению дисциплины

Вид учебного занятия	Организация деятельности студента
Лекция	Написание конспекта лекций: кратко, схематично, последовательно фиксировать основные положения, выводы, формулировки, обобщения; помечать важные мысли, выделять ключевые слова, термины. Проверка терминов, понятий с помощью энциклопедий, словарей, справочников с выписыванием толкований в тетрадь. Обозначить вопросы, термины, материал, который вызывает трудности, пометить и попытаться найти ответ в рекомендуемой литературе. Если самостоятельно не удастся разобраться в материале, необходимо сформулировать вопрос и задать преподавателю на консультации, на практическом занятии.
Практические занятия	Проработка рабочей программы, уделяя особое внимание целям и задачам, структуре и содержанию дисциплины. Работа с конспектом лекций, просмотр рекомендуемой литературы. Изучение выбранной предметной области на примерах решения задач семинарских занятий,

	индивидуальных домашних заданий.
Курсовая работа	Не предусмотрена
Контрольная работа	Ознакомиться с основной и дополнительной литературой, включая справочные издания, зарубежные источники, основополагающие термины. Попрактиковаться в решении аналогичных домашних задач по всем темам контрольных работ.
Лабораторная работа	Не предусмотрена.
Подготовка к зачету	При подготовке к зачету необходимо ориентироваться на конспекты лекций и рекомендуемую литературу.

10. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем (при необходимости)

Издательская система LaTeX для подготовки докладов, презентаций и учебного материала.

11. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине

Видеопроектор, компьютер, издательская система LaTeX для подготовки докладов, презентаций и учебного материала.

12. Иные сведения и (или) материалы

12.1. Перечень образовательных технологий, используемых при осуществлении образовательного процесса по дисциплине

Часов в интерактивной форме – 8.

В ходе практических занятий происходит публичное обсуждение каждой решаемой задачи. При этом студенты высказывают свои мнения по выбору наиболее простого способа поиска оптимального решения.

После решения домашних работ на консультациях проводится разбор допущенных студентами ошибок.

12.2. Формы организации самостоятельной работы обучающихся (темы, выносимые для самостоятельного изучения; вопросы для самоконтроля; типовые задания для самопроверки)

Некоторые темы изучаются студентами самостоятельно. Для изучения используется приведённая в списке основная и дополнительная литература. Контроль освоения материала осуществляется при проверке контрольных работ, домашнего задания и на зачете.

№	Тема и часть, изучаемая (осваиваемая) самостоятельно
1.1	Оптимизация запросов в PostgreSQL
1.2	Производительность и индексы в PostgreSQL
1.3	Управление транзакциями в PostgreSQL
1.4	Работа с JSON и XML в PostgreSQL
1.5	Безопасность и управление доступом в PostgreSQL
1.6	Параллельные запросы и асинхронные операции

№	Тема и часть, изучаемая (осваиваемая) самостоятельно
1.7.	Создание и использование представлений (views)
1.8.	Расширенные функции и операторные перегрузки
1.9.	Импорт и экспорт данных в PostgreSQL
1.10	Мониторинг и администрирование PostgreSQL

Вопросы и задания для самоконтроля по всем темам:

1. Что такое PL/pgSQL и в каких случаях он используется?
2. Какие основные возможности предоставляет PL/pgSQL для работы с базами данных?
3. Как объявляются и используются переменные в анонимных блоках PL/pgSQL?
4. В чем разница между условными операторами и циклами в PL/pgSQL?
5. Как создать и использовать курсоры в PL/pgSQL?
6. В чем отличие между функциями и процедурами в PL/pgSQL?
7. Как обрабатываются ошибки в PL/pgSQL и какие конструкции для этого используются?
8. Что такое триггеры и как они применяются в PostgreSQL?
9. Как создать и использовать DDL (Data Definition Language) в PostgreSQL для управления схемами и таблицами?
10. Каковы лучшие практики для написания и тестирования PL/pgSQL кода?

12.3. Краткий терминологический словарь

PL/pgSQL	язык процедурного программирования, встроенный в PostgreSQL, который позволяет создавать сложные функции и процедуры для работы с данными.
Анонимные блоки	фрагменты кода PL/pgSQL, которые могут выполняться без предварительного объявления и обычно используются для выполнения временных операций.
Курсоры	объекты, позволяющие перебирать строки результирующего набора данных по одной строке за раз, что удобно для работы с большими объемами данных.
Функции и процедуры	блоки кода, которые могут быть вызваны для выполнения определенной задачи. Функции возвращают значение, в то время как процедуры могут выполнять действия, но значения не возвращают.
Триггеры	специальные процедуры, которые автоматически выполняются в ответ на определенные события в базе данных, такие как вставка, обновление или удаление данных.