

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

ИНСТИТУТ ЭМИТ
Кафедра Системного анализа и информатики

УТВЕРЖДЕНА
на заседании кафедры
Системного анализа и информатики
Протокол от «19» июня 2021 г. № 7

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

Индекс Б1.В.ДВ.04.02 «Хранилища данных»

по направлению подготовки 38.04.01 «Экономика»

направленность «Системы больших данных в экономике»

квалификация магистр

очная форма обучения

Год набора - 2021

Москва, 2021 г.

Автор(ы)—составитель(и):

к.т.н., доцент кафедры Системного анализа и информатики Стефановский Д.В.

Заведующий кафедрой

Системного анализа и информатики, к.т.н., доцент, Маруев С.А.

СОДЕРЖАНИЕ

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы.....	4
2. Объем и место дисциплины в структуре образовательной программы.....	5
3. Содержание и структура дисциплины.....	5
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине.....	7
5. Методические указания для обучающихся по освоению дисциплины	13
6. Учебная литература и ресурсы информационно-телекоммуникационной сети «Интернет», учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине	16
6.1. Основная литература	16
6.2. Дополнительная литература	16
6.3. Учебно-методическое обеспечение самостоятельной работы	17
6.4. Нормативные правовые документы	17
6.5. Интернет-ресурсы	17
6.6. Иные источники	17
7. Материально-техническая база, информационные технологии, программное обеспечение и информационные справочные системы	17

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

1.1. Дисциплина Б1.В.ДВ.04.01 «Хранилища данных» обеспечивает совершенствование следующими компетенциями:

Код компетенции	Наименование компетенции	Код этапа совершенствования компетенции	Наименование этапа освоения компетенции
ПКс-1	Способен подготовить данные для проведения аналитических работ по исследованию больших данных	ПКс-1.У	Способен к обоснованному отбору и формированию состава собственных и приобретаемых данных для проведения аналитических работ по исследованию больших данных

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

ОТФ/ТФ (при наличии профстандарта)	Код этапа освоения компетенции	Результаты обучения
ведение аналитической деятельности в области экономики	ПК-9.1У	на уровне знаний: <ul style="list-style-type: none"> • Методы разработки отчетной аналитической документации
		на уровне умений: <ul style="list-style-type: none"> • Адаптировать и развертывать модели в предметной среде • Проведение сравнительного анализа и выбора методов и методик анализа больших данных и составление рекомендаций по их использованию, оценивать условия их приобретения и использования • Разработка спецификации и описания разрабатываемых методов и методик анализа больших данных, инструментальных средств или их компонент в соответствии с требованиями технического задания
		на уровне знаний: <ul style="list-style-type: none"> • Современный опыт использования анализа больших данных • Современные методы и инструментальные средства анализа больших данных • Рекомендации и опыт использования методов анализа больших данных • Рекомендации по использованию, опыт использования и интеграции современных инструментальных средств сбора, хранения, обработки и анализа больших данных • Методы интерпретации и визуализации больших данных • Источники информации и условия их использования • Методы извлечения информации и знаний из гетерогенных, мульти структурированных, неструктурированных источников, в том числе при потоковой обработке • Современная технологическая инфраструктура высокопроизводительных и распределенных вычислений • Методы обеспечения и оценки качества информации

2. Объем и место дисциплины в структуре ОП ВО

Общий объем дисциплины 4 ЗЕ (144 ак.ч./108 астр.ч.). В целях планирования, при расчете структуры дисциплины применены академические часы (ак.ч.).

Количество академических часов, выделенных на занятия лекционного типа – 16 ак.ч., занятия семинарского типа – 16 ак.ч.; на самостоятельную работу обучающихся – 108 ак.ч. и промежуточную аттестацию – 4 ак.ч.

Место дисциплины в структуре образовательной программы:

- дисциплина Б1.В.ДВ.04.01 «Хранилища данных» изучается в 4 семестре на 2 курсе;
- дисциплина реализуется после изучения дисциплин:
 - математический анализ (в объеме бакалавриата),
 - линейная алгебра (в объеме бакалавриата),
 - экономическая информатика (в объеме бакалавриата),
 - дискретная математика (в объеме бакалавриата),
 - теория вероятности и математическая статистика (в объеме бакалавриата),
 - алгоритмизация и программирование распределенных вычислений;
- дисциплина может реализоваться частично или полностью с применением ЭО и/или ДОТ. Учебные материалы дисциплины размещаются по адресу lms.ranepa.ru
- форма промежуточной аттестации – зачет с оценкой.

3. Содержание и структура дисциплины (модуля)

Таблица 2.

№ п/п	Наименование тем (разделов)	Объем дисциплины, ак.час./час						Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
				Л	ЛР	ПЗ		КСР
Очная форма обучения								
Тема 1	Введение и общие понятия.	18	2		2		14	Опрос
Тема 2	Модель типового проекта создания хранилища данных.	18	2		2		14	Опрос
Тема 3	Метод моделирования «сущность-связь»	18	2		2		14	Опрос
Тема 4	Физическая модель хранилища данных	18	2		2		14	Опрос
Тема 5	Проектирование и разработка процесса ETL	18	2		2		14	Опрос
Тема 6	Метод моделирования «Свод данных»	18	2		2		14	Опрос
Тема 7	Проектирование производительности	18	2		2		14	Опрос
Тема 8	SQL в хранилищах данных	18	2		2		14	Опрос, КР (1-8)
Промежуточная аттестация					-		-	Зачет с оценкой
Всего:		144/108	16/12		16/12		112/84	

Примечание: * формы текущего контроля успеваемости: опрос (О), контрольная работа (КР)

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

Содержание дисциплины

Тема 1. Введение и общие понятия.

Хранилища данных. Концепция систем складирования данных. Концепция хранилищ данных. Типы хранилищ данных. Архитектура хранилищ данных. Основные типы программно-аппаратной архитектуры хранилища данных. Подходы в организации работ по созданию хранилища данных. Характеристика решений ведущих производителей. Типовые программно-аппаратные решения (технологические решения). Корпоративные информационные фабрики. Хранилища данных с архитектурой шины данных Объединенное (федеративное) ХД

Тема 2. Модель типового проекта создания хранилища данных.

Жизненный цикл разработки хранилища данных. Модель жизненного цикла хранилища данных. Бизнес-модель типового проекта создания хранилища данных. Системы деловой осведомленности (Business Intelligence Systems) и хранилища данных. Определение систем деловой осведомленности. Информационная безопасность систем бизнес-аналитики. Технологии обработки данных. Архитектура систем бизнес-уровня. Построение систем бизнес-аналитики: проблемы и решения. Хранилища данных и системы бизнес-аналитики.

Тема 3. Метод моделирования «сущность-связь».

Понятие предметной области и архитектура данных. Моделирование методом «сущность-связь». Моделирование темпоральных (временных) данных в хранилищах данных. Темпоральные данные и базы данных. Введение в моделирование темпоральных данных. Основные приемы моделирования темпоральных данных. Метод многомерного моделирования. Основные понятия метода многомерного моделирования. Многомерная модель. Факты. Ключи в таблицах фактов. Таблицы фактов. Измерения. Основные схемы многомерной модели. Моделирование таблиц фактов. Моделирование таблиц измерений.

Тема 4. Физическая модель хранилища данных.

Создание физической модели хранилища данных. Объекты физической модели данных. Моделирование объектов физической модели хранилища данных. Разработка скрипта для создания объектов физической модели хранилища данных. Метаданные в хранилищах данных. Метаданные. Функции метаданных в хранилище данных. Состав метаданных в хранилище данных. Логическая структура метаданных хранилища данных. Стандарты метаданных. Выбор метамодели при проектировании хранилища данных. Проектирование логической модели метаданных хранилища данных.

Тема 5. Проектирование и разработка процесса ETL.

Введение. Подходы к реализации ETL-процесса. Разработка ETL-процесса. Элементы ETL-процесса. Проектирование процессов преобразования данных с помощью CASE-инструментов. Создание модели хранилища данных на основе корпоративной модели данных. Корпоративная модель данных. Алгоритм преобразования корпоративной модели данных в модель хранилища данных. Использование CASE-

инструментов для создания многомерной модели на основе корпоративной модели данных. Секционирование таблиц многомерной модели данных в PowerDesigner.

Тема 6. Метод моделирования «Свод данных».

Метод моделирования «Свод данных» (Data Vault). Пример проектирования модели «Свод данных». Заполнение данными объектов «Свода данных». Физическая модель хранилища данных: учет влияния транзакций, денормализация таблиц. Определение транзакций базы данных. Денормализация. Методы разбиения таблиц. Денормализация колонок.

Тема 7. Проектирование производительности.

Повышение производительности запросов: индексы. Повышение производительности запросов: секционирование. Повышение производительности запросов: кластеры. SQL в хранилищах данных: агрегация и суммирование. Введение. Оператор SELECT и схема «звезда». Расширение оператора SELECT для обработки данных. Функция GROUPING. Управление иерархией в предложениях ROLLUP и CUBE.

Тема 8. SQL в хранилищах данных.

SQL в хранилищах данных: аналитическая обработка данных. SQL для анализа данных. Агрегатные и статистические функции. Медианы. Оконные функции. Функции ранжирования. Функции, генерирующие отчеты. CASE-выражения и создание гистограмм. Статистические функции в других диалектах SQL. Проектирование кубов данных. OLAP на клиенте и на сервере. Технические аспекты многомерного хранения данных. Основные понятия OLAP. Проектирование кубов данных с использованием CASE-инструментов. Генерация куба данных.

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

4.1. Формы и методы текущего контроля успеваемости.

4.1.1. В ходе реализации дисциплины Б1.В.ДВ.04.01 «Хранилища данных» используются следующие методы текущего контроля успеваемости обучающихся: домашние задания и контрольные работы.

Тема	Методы текущего контроля успеваемости
Тема 1	Опрос
Тема 2	Опрос
Тема 3	Опрос
Тема 4	Опрос
Тема 5	Опрос
Тема 6	Опрос
Тема 7	Опрос
Тема 8	Опрос, контрольная работа по темам 1-8

4.1.2. Зачет проводится с применением следующих методов (средств):

Зачет с оценкой проводится в виде письменной контрольной работы.

4.2. Материалы текущего контроля успеваемости обучающихся

Типовые оценочные материалы по теме 1

Опрос:

- Опишите основные особенности концепции систем складирования данных. Поясните на примере.
- Перечислите основные типы хранилищ данных и опишите архитектуру хранилищ данных. Поясните на примере их особенности.
- Перечислите основные типы программно-аппаратной архитектуры хранилища данных. Поясните на примере их особенности.
- Перечислите основные подходы в организации работ по созданию хранилища данных. Поясните на примере их особенности.

Типовые оценочные материалы по теме 2

Опрос:

- Опишите этапы жизненного цикла разработки хранилища данных. Поясните на примерах.
- Перечислите основные особенности и элементы модели жизненного цикла. Поясните на примерах.
- Приведите бизнес-модель типового проекта создания хранилища данных. Поясните на примерах.
- Опишите задачи, архитектуру и особенности: систем деловой осведомленности (Business Intelligence Systems) и хранилища данных. Поясните на примерах их использование
- Опишите основные технологии обработки данных и архитектуру систем бизнес-уровня. Поясните на примерах.
- Перечислите основные решения и проблемы построения систем бизнес-аналитики. Поясните их суть на примере.

Типовые оценочные материалы по теме 3

Опрос:

- Опишите понятие предметной области и архитектура данных. Поясните на примерах.
- Перечислите основные этапы моделирование методом «сущность-связь». Поясните на примерах их суть.
- Опишите понятие темпоральных данных и баз данных.
- Поясните на примерах основные приемы моделирования темпоральных данных.
- Опишите метод многомерного моделирования. Поясните на примерах суть многомерной модели.

Типовые оценочные материалы по теме 4

Опрос:

- Опишите основные этапы и особенности создания физической модели хранилища данных. Поясните на примерах.
- Опишите основные этапы и особенности моделирование объектов физической модели хранилища данных. Поясните на примерах.
- Опишите функции и состав метаданных в хранилище данных. Поясните на примерах.
- Опишите основные этапы проектирование логической модели метаданных хранилища данных. Поясните на примерах.

Типовые оценочные материалы по теме 5

Опрос:

- Опишите основные этапы и задачи ETL-процесса. Поясните на примерах.

- Опишите особенности проектирование процессов преобразования данных с помощью
- CASE-инструментов. Поясните на примерах.
- Опишите основные этапы и задачи создания модели хранилища данных на основе корпоративной модели данных. Поясните на примерах.
- Опишите алгоритм преобразования корпоративной модели данных в модель хранилища данных. Поясните на примерах.
- Опишите особенности секционирования таблиц многомерной модели данных в PowerDesigner. Поясните на примерах.

Типовые оценочные материалы по теме 6

Опрос:

- Опишите метод моделирования «Свод данных» (Data Vault). Поясните на примерах.
- Опишите следующие основные элементы физической модели хранилища данных: учет влияния транзакций, денормализация таблиц. Поясните на примерах.
- Дайте определение следующих понятий: транзакции базы данных, денормализация. Поясните на примере.

Типовые оценочные материалы по теме 7

Опрос:

- Опишите основные способы повышение производительности запросов. Поясните на примерах способы организации индексов.
- Секционирование как способ повышение производительности запросов. Поясните на примере.
- Опишите кластерный способ повышение производительности. Поясните на примере.
- Опишите следующие особенности SQL в хранилищах данных: агрегация и суммирование. Поясните на примерах.

Типовые оценочные материалы по теме 8

Опрос:

- Опишите агрегатные и статистические функции. Поясните на примерах.
- Перечислите основные статистические функции в диалектах SQL. Поясните их работу на примере.
- Опишите основные этапы проектирование кубов данных. Поясните их суть на примере.
- Опишите основные технические аспекты многомерного хранения. Поясните их суть на примере.

Контрольная работа:

По предложенному преподавателем описанию предметной области создайте логическую модель предметной области базы данных. Дайте формальные определения и представьте модель в виде схемы сущность-связь.

4.3. Оценочные средства для промежуточной аттестации.

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

Код компетенции	Наименование компетенции	Код этапа совершенствования компетенции	Наименование этапа освоения компетенции
ПКс-1	Способен подготовить данные для проведения аналитических работ по исследованию больших данных	ПКс-1.У	Способен к обоснованному отбору и формированию состава собственных и приобретаемых данных для проведения аналитических работ по исследованию больших данных

Этап освоения компетенции	Показатель оценивания	Критерий оценивания
ПКс-1.У	Способен к обоснованному отбору и формированию состава собственных и приобретаемых данных для проведения аналитических работ по исследованию больших данных	Понимает и обосновывает состав собственных и приобретаемых данных и информации с указанием источников данных и условий их получения и доставки

4.3.2. Типовые оценочные средства.

Вопросы к зачету:

- Опишите основные особенности концепции систем складирования данных. Поясните на примере.
- Перечислите основные типы хранилищ данных и опишите архитектуру хранилищ данных. Поясните на примере их особенности.
- Перечислите основные типы программно-аппаратной архитектуры хранилища данных. Поясните на примере их особенности.
- Перечислите основные подходы в организации работ по созданию хранилища данных. Поясните на примере их особенности.
- Опишите этапы жизненного цикла разработки хранилища данных. Поясните на примерах.
- Перечислите основные особенности и элементы модели жизненного цикла. Поясните на примерах.
- Приведите бизнес-модель типового проекта создания хранилища данных. Поясните на примерах.
- Опишите задачи, архитектуру и особенности: систем деловой осведомленности (Business Intelligence Systems) и хранилища данных. Поясните на примерах их использование
- Опишите основные технологии обработки данных и архитектуру систем бизнес-уровня. Поясните на примерах.
- Перечислите основные решения и проблемы построения систем бизнес-аналитики. Поясните их суть на примере.
- Опишите понятие предметной области и архитектура данных. Поясните на примерах.
- Перечислите основные этапы моделирование методом «сущность-связь». Поясните на примерах их суть.
- Опишите понятие темпоральных данных и баз данных.
- Поясните на примерах основные приемы моделирования темпоральных данных.

- Опишите метод многомерного моделирования. Поясните на примерах суть многомерной модели.
- Опишите основные этапы и особенности создания физической модели хранилища данных. Поясните на примерах.
- Опишите основные этапы и особенности моделирование объектов физической модели хранилища данных. Поясните на примерах.
- Опишите функции и состав метаданных в хранилище данных. Поясните на примерах.
- Опишите основные этапы проектирование логической модели метаданных хранилища данных. Поясните на примерах.
- Опишите основные этапы и задачи ETL-процесса. Поясните на примерах.
- Опишите особенности проектирование процессов преобразования данных с помощью CASE-инструментов. Поясните на примерах.
- Опишите основные этапы и задачи создания модели хранилища данных на основе корпоративной модели данных. Поясните на примерах.
- Опишите алгоритм преобразования корпоративной модели данных в модель хранилища данных. Поясните на примерах.
- Опишите особенности секционирования таблиц многомерной модели данных в PowerDesigner. Поясните на примерах.
- Опишите метод моделирования «Свод данных» (Data Vault). Поясните на примерах.
- Опишите следующие основные элементы физической модели хранилища данных: учет влияния транзакций, денормализация таблиц. Поясните на примерах.
- Дайте определение следующих понятий: транзакции базы данных, денормализация. Поясните на примере.
- Опишите основные способы повышение производительности запросов. Поясните на примерах способы организации индексов.
- Секционирование как способ повышение производительности запросов. Поясните на примере.
- Опишите кластерный способ повышение производительности. Поясните на примере.
- Опишите следующие особенности SQL в хранилищах данных: агрегация и суммирование. Поясните на примерах.
- Опишите агрегатные и статистические функции. Поясните на примерах.
- Перечислите основные статистические функции в диалектах SQL. Поясните их работу на примере.
- Опишите основные этапы проектирование кубов данных. Поясните их суть на примере.
- Опишите основные технические аспекты многомерного хранения. Поясните их суть на примере.
- По предложенному преподавателем описанию предметной области создайте логическую модель предметной области базы данных. Дайте формальные определения и представьте модель в виде схемы сущность-связь.

Шкала оценивания.

Оценка определяется по формуле: $1/3 \text{ опрос} + 1/3 \text{ к.р.} + 1/3 \text{ сдача зачета.}$

10- бальная шкала	Традиционн ая шкала	«Зачтено»/ «Не зачтено»	Определение
-------------------------	------------------------	-------------------------------	-------------

10- бальная шкала	Традиционн ая шкала	«Зачтено»/ «Не зачтено»	Определение
10	Отлично	Зачтено	Полные, глубокие и систематические знания, полный и правильный ответ на теоретический вопрос, полное и правильное решение задачи.
9	Отлично	Зачтено	Глубокие и систематические знания, правильный ответ на теоретический вопрос, правильное решение задачи.
8	Отлично	Зачтено	Систематические знания, правильный ответ на теоретический вопрос, правильное решение задачи.
7	Хорошо	Зачтено	Систематические знания, правильный ответ на теоретический вопрос с незначительными неточностями, правильное решение задачи.
6	Хорошо	Зачтено	Систематические знания, правильный ответ на теоретический вопрос с незначительными неточностями, правильное решение задачи с незначительными неточностями.
5	Удовлетвори тельно	Зачтено	Ответ на теоретический вопрос неполный, правильное решение задачи с незначительными неточностями.
4	Удовлетвори тельно	Зачтено	Ответ на теоретический вопрос неполный, решение задачи содержит арифметические ошибки, не влияющие на правильность хода решения задачи.
3	Неудовлетво рительно	Не зачтено	Ответ на теоретический вопрос неполный, решение задачи содержит идеологические ошибки.
2	Неудовлетво рительно	Не зачтено	Ответ на теоретический вопрос неверный и/или решение задачи содержит идеологические ошибки.
1	Неудовлетво рительно	Не зачтено	Ответ на теоретический вопрос неверный и решение задачи отсутствует.
0	Неудовлетво рительно	Не зачтено	Ответ на теоретический вопрос отсутствует и решение задачи отсутствует.

4.4. Методические материалы по проведению промежуточной аттестации

Зачет проводится в аудитории. Отсчет времени, отведенного на письменную работу, идет по завершении процедуры размещения студентов и раздачи заданий.

Студент обязан являться на письменный контроль в указанное в расписании время. В случае опоздания время, отведенное на письменный контроль знаний, не продлевается.

При себе студенты могут иметь только письменные принадлежности. Необходимую для выполнения работы бумагу выдает преподаватель.

Преподаватель раздает варианты работы, содержащий 2 вопроса. Листы с заданиями должны быть повернуты текстом вниз, чтобы студенты до окончания процедуры раздачи не могли начать выполнение работы. По окончании раздачи вариантов студентам разрешается перевернуть текст задания и одновременно приступить к выполнению работы. По окончании отведенного времени студенты одновременно заканчивают

выполнение работы. Если работа завершена существенно раньше срока, то по разрешению преподавателя студент может покинуть аудиторию досрочно.

Мобильные телефоны должны быть выключены и убраны со столов, допускается использование калькуляторов, выполняющих только простые арифметические вычисления.

Во время проведения письменного контроля знаний студентам не разрешается пользоваться учебными программами, справочниками и прочими источниками информации.

Использование материалов, а также попытка общения с другими студентами или иными лицами, в том числе с применением электронных средств связи, несанкционированные перемещения и т.п. являются основанием для удаления студента из аудитории и последующего проставления в ведомость оценки «неудовлетворительно».

Во время проведения письменного контроля знаний студентам разрешается покинуть аудиторию только при условии сдачи работы в объеме, выполненном к моменту выхода из аудитории. Дальнейшее продолжение работы запрещается.

Ответы в работе без объяснений не засчитываются. Рисунки должны быть четкими, все линии графиков, используемых при ответах на вопросы задач, должны быть подписаны.

5. Методические указания для обучающихся по освоению дисциплины

Любой вид занятий, создающий условия для зарождения самостоятельной мысли, познавательной и творческой активности студента связан с самостоятельной работой. В широком смысле под самостоятельной работой понимают совокупность всей самостоятельной деятельности студентов как в учебной аудитории, так и вне ее, в контакте с преподавателем и в его отсутствие. Самостоятельная работа может реализовываться: непосредственно в процессе аудиторных занятий – на лекциях, практических и семинарских занятиях, при выполнении контрольных и лабораторных работ и др.; в контакте с преподавателем вне рамок аудиторных занятий – на консультациях по учебным вопросам, в ходе творческих контактов, при ликвидации задолженностей, при выполнении индивидуальных заданий и т.д.; в библиотеке, дома, в общежитии, на кафедре и других местах при выполнении студентом учебных и творческих заданий.

Лекции

Главное в период подготовки к лекционным занятиям – научиться методам самостоятельного умственного труда, сознательно развивать свои творческие способности и овладевать навыками творческой работы. Для этого необходимо строго соблюдать дисциплину учебы и поведения. Четкое планирование своего рабочего времени и отдыха является необходимым условием для успешной самостоятельной работы. В основу его нужно положить рабочие программы изучаемых в семестре дисциплин. Каждому студенту следует составлять еженедельный и семестровый планы работы, а также план на каждый рабочий день. С вечера всегда надо распределять работу на завтрашний день. В конце каждого дня целесообразно подводить итог работы: тщательно проверить, все ли выполнено по намеченному плану, не было ли каких-либо отступлений, а если были, по какой причине это произошло. Нужно осуществлять самоконтроль, который является необходимым условием успешной учебы. Если что-то осталось невыполненным, необходимо изыскать время для завершения этой части работы, не уменьшая объема недельного плана.

Самостоятельная работа на лекции. Слушание и запись лекций – сложный вид вузовской аудиторной работы. Внимательное слушание и конспектирование лекций предполагает интенсивную умственную деятельность студента. Краткие записи лекций,

их конспектирование помогает усвоить учебный материал. Конспект является полезным тогда, когда записано самое существенное, основное и сделано это самим студентом. Не надо стремиться записать дословно всю лекцию. Такое «конспектирование» приносит больше вреда, чем пользы. Запись лекций рекомендуется вести по возможности собственными формулировками. Желательно запись осуществлять на одной странице, а следующую оставлять для проработки учебного материала самостоятельно в домашних условиях. Конспект лекции лучше подразделять на пункты, параграфы, соблюдая красную строку. Этому в большой степени будут способствовать пункты плана лекции, предложенные преподавателям. Принципиальные места, определения, формулы и другое следует сопровождать замечаниями «важно», «особо важно», «хорошо запомнить» и т.п. Можно делать это и с помощью разноцветных маркеров или ручек. Лучше если они будут собственными, чтобы не приходилось присить их у однокурсников и тем самым не отвлекать их во время лекции. Целесообразно разработать собственную «маркографию» (значки, символы), сокращения слов. Не лишним будет и изучение основ стенографии. Работая над конспектом лекций, всегда необходимо использовать не только учебник, но и ту литературу, которую дополнительно рекомендовал лектор. Именно такая серьезная, кропотливая работа с лекционным материалом позволит глубоко овладеть знаниями.

Семинар и проведение опроса

Каждый студент должен начать с ознакомления с планом занятия, который отражает содержание предложенной темы. Тщательное продумывание и изучение вопросов плана основывается на проработке текущего материала лекции, а затем изучения обязательной и дополнительной литературы, рекомендованную к данной теме. На основе индивидуальных предпочтений студенту необходимо самостоятельно выбрать тему доклада по проблеме семинара и по возможности подготовить по нему презентацию. Если программой дисциплины предусмотрено выполнение практического задания, то его необходимо выполнить с учетом предложенной инструкции (устно или письменно). Все новые понятия по изучаемой теме необходимо выучить наизусть и внести в глоссарий, который целесообразно вести с самого начала изучения курса. Результат такой работы должен проявиться в способности студента свободно ответить на теоретические вопросы семинара, его выступлении и участии в коллективном обсуждении вопросов изучаемой темы, правильном выполнении практических заданий и контрольных работ.

Работа с литературными источниками.

В процессе подготовки к семинарским занятиям, студентам необходимо обратить особое внимание на самостоятельное изучение рекомендованной учебно-методической (а также научной и популярной) литературы. Самостоятельная работа с учебниками, учебными пособиями, научной, справочной и популярной литературой, материалами периодических изданий и Интернета, статистическими данными является наиболее эффективным методом получения знаний, позволяет значительно активизировать процесс овладения информацией, способствует более глубокому усвоению изучаемого материала, формирует у студентов свое отношение к конкретной проблеме. Более глубокому раскрытию вопросов способствует знакомство с дополнительной литературой, рекомендованной преподавателем по каждой теме семинарского или практического занятия, что позволяет студентам проявить свою индивидуальность в рамках выступления на данных занятиях, выявить широкий спектр мнений по изучаемой проблеме.

Методические указания по выполнению самостоятельной работы:

Тема 1. Введение и общие понятия.

Хранилища данных. Концепция систем складирования данных. Концепция хранилищ данных. Типы хранилищ данных. Архитектура хранилищ данных. Основные типы программно-аппаратной архитектуры хранилища данных. Подходы в организации работ по созданию хранилища данных. Характеристика решений ведущих

производителей. Типовые программно-аппаратные решения (технологические решения). Корпоративные информационные фабрики. Хранилища данных с архитектурой шины данных Объединенное (федеративное) ХД

Тема 2. Модель типового проекта создания хранилища данных.

Жизненный цикл разработки хранилища данных. Модель жизненного цикла хранилища данных. Бизнес-модель типового проекта создания хранилища данных. Системы деловой осведомленности (Business Intelligence Systems) и хранилища данных. Определение систем деловой осведомленности. Информационная безопасность систем бизнес-аналитики. Технологии обработки данных. Архитектура систем бизнес-уровня. Построение систем бизнес-аналитики: проблемы и решения. Хранилища данных и системы бизнес-аналитики.

Тема 3. Метод моделирования «сущность-связь».

Понятие предметной области и архитектура данных. Моделирование методом «сущность-связь». Моделирование темпоральных (временных) данных в хранилищах данных. Темпоральные данные и базы данных. Введение в моделирование темпоральных данных. Основные приемы моделирования темпоральных данных. Метод многомерного моделирования. Основные понятия метода многомерного моделирования. Многомерная модель. Факты. Ключи в таблицах фактов. Таблицы фактов. Измерения. Основные схемы многомерной модели. Моделирование таблиц фактов. Моделирование таблиц измерений.

Тема 4. Физическая модель хранилища данных.

Создание физической модели хранилища данных. Объекты физической модели данных. Моделирование объектов физической модели хранилища данных. Разработка скрипта для создания объектов физической модели хранилища данных. Метаданные в хранилищах данных. Метаданные. Функции метаданных в хранилище данных. Состав метаданных в хранилище данных. Логическая структура метаданных хранилища данных. Стандарты метаданных. Выбор метамодели при проектировании хранилища данных. Проектирование логической модели метаданных хранилища данных.

Тема 5. Проектирование и разработка процесса ETL.

Введение. Подходы к реализации ETL-процесса. Разработка ETL-процесса. Элементы ETL-процесса. Проектирование процессов преобразования данных с помощью CASE-инструментов. Создание модели хранилища данных на основе корпоративной модели данных. Корпоративная модель данных. Алгоритм преобразования корпоративной модели данных в модель хранилища данных. Использование CASE-инструментов для создания многомерной модели на основе корпоративной модели данных. Секционирование таблиц многомерной модели данных в PowerDesigner.

Тема 6. Метод моделирования «Свод данных».

Метод моделирования «Свод данных» (Data Vault). Пример проектирования модели «Свод данных». Заполнение данными объектов «Свода данных». Физическая модель хранилища данных: учет влияния транзакций, денормализация таблиц. Определение транзакций базы данных. Денормализация. Методы разбиения таблиц. Денормализация колонок.

Тема 7. Проектирование производительности.

Повышение производительности запросов: индексы. Повышение производительности запросов: секционирование. Повышение производительности запросов: кластеры. SQL в хранилищах данных: агрегация и суммирование. Введение.

Оператор SELECT и схема «звезда». Расширение оператора SELECT для обработки данных. Функция GROUPING. Управление иерархией в предложениях ROLLUP и CUBE.

Тема 8. SQL в хранилищах данных.

SQL в хранилищах данных: аналитическая обработка данных. SQL для анализа данных. Агрегатные и статистические функции. Медианы. Оконные функции. Функции ранжирования. Функции, генерирующие отчеты. CASE-выражения и создание гистограмм. Статистические функции в других диалектах SQL. Проектирование кубов данных. OLAP на клиенте и на сервере. Технические аспекты многомерного хранения данных. Основные понятия OLAP. Проектирование кубов данных с использованием CASE-инструментов. Генерация куба данных.

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

6.1. Основная литература.

1. Туманов В.Е. Проектирование хранилищ данных для систем бизнес-аналитики: учебное пособие / В.Е. Туманов. — М.: Интернет-Университет Информационных Технологий: БИНОМ. Лаборатория знаний, 2010. — 615 с: ил., табл. — (Основы информационных технологий).

6.2. Дополнительная литература.

1. Постреляционные хранилища данных. Учебное пособие. Юрий Парфенов

6.3. Учебно-методическое обеспечение самостоятельной работы.

Отдельное обеспечение не предусмотрено.

6.4. Нормативные правовые документы.

Не предусмотрены.

6.5. Интернет-ресурсы.

1. <http://citeseer.ist.psu.edu/> — основной источник знаний по Computer Science, по многим статьям есть полные тексты
2. <http://citeseer.cs.msu.su/> — русскоязычная электронная библиотека научных статей
3. <http://arxiv.org/> — библиотека электронных публикаций, в основном по физике, но доля «Computer Science» в последнее время стремительно увеличивается
4. <http://rexa.info/> — библиографическая поисковая система по статьям, авторам и грантам
5. <http://elibrary.ru/> — российская научная электронная библиотека
6. <http://liinwww.ira.uka.de/bibliography/index.html>
7. библиографическая база данных для работы с BibTeX
8. <http://www.gotai.net/> -- русскоязычный сайт об искусственном интеллекте
9. Math-Net.ru -- общероссийский математический портал

6.6. Иные источники.

Не предусмотрены.

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

Для лекций:

1. Персональный компьютер
2. Мультимедийный проектор
3. Доска, мел или маркеры

Для лабораторных занятий:

1. Компьютерный класс
2. Виртуальная машина Ubuntu 15.04 b выше с установленным Postgresql и MongoDB
3. Мультимедийный проектор
4. Доска, маркеры
5. Компилятор R-2.15.1 – GNU - <http://www.r-project.org/> либо интегрированная среда разработки RStudio – GNU AGP - <http://www.rstudio.com/ide/>.
6. Программный комплекс анализа новостного сайта - "Crawler-Persona"
7. База данных "Централизация государственных закупок в 2014 г".
8. База данных учебно-методических материалов по дисциплине "Макроэкономика".
9. База данных Бюджетная и социально-экономическая статистика субъектов Российской Федерации.