

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ
ФЕДЕРАЦИИ

ПЕНЗЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ФАКУЛЬТЕТ Экономики и Управления



УТВЕРЖДАЮ
Декан факультета

В.М. Володин

«28» 12 2017 г.

Рабочая программа дисциплины
Б 1.2.14 «Хранилища данных и знаний»

Направление подготовки: 38.03.05 «Бизнес-информатика»

Квалификация (степень) выпускника - бакалавр

Форма обучения: очная, заочная

Пенза 2017г.

1. Цели освоения дисциплины

Целью освоения дисциплины Б 1.2.14 «Хранилище данных и знаний» является обучение теории и практике проектирования, разработки, применения хранилищ данных для экономических информационных систем.

2. Место дисциплины в структуре ОПОП бакалавриата

Учебная дисциплина относится к вариативной части Б 1.2. Изложение материалов основано на изученных дисциплинах «Информационно-коммуникационные технологии», «Базы данных», «Моделирование бизнес-процессов», «Бизнес-архитектура предприятия», «Управление ИТ-сервисами и контентом», «Анализ и управление бизнес-процессами».

Основные положения дисциплины могут быть использованы в дальнейшем при изучении дисциплин «Управление информационными системами и Интернет-ресурсами», «Эконометрическое моделирование бизнес-процессов» и других, связанных с применением методов математического моделирования в экономике, обработкой информации с использованием современных информационных технологий, а также при подготовке и защите выпускной квалификационной работы.

3. Компетенции обучающегося, формируемые в результате освоения дисциплины.

Процесс изучения дисциплины направлен на формирование элементов следующих компетенций в соответствии с ФГОС ВО по данному направлению:

Коды компетенции		Структурные элементы компетенции (в результате освоения дисциплины обучающийся должен знать, уметь, владеть)
1	2	3
ПК-3	Выбор рациональных информационных систем и информационно-коммуникативных технологий решения для управления бизнесом.	<p>Знать:</p> <ul style="list-style-type: none">– возможности использования хранилищ данных (ХД) для оперативного анализа данных (OLAP);– принципы организации хранилища, видах данных для решения задач бизнес-аналитики;– состав и основные функции корпоративного хранилища данных, его структуре;– технологии использования различных источников данных для ХД;– о проблеме подготовки данных, необходимость использования процесса извлечения, трансформации и загрузки данных (ETL), его основной цели и задачи, подходах к реализации ETL-процесса. <p>Уметь:</p> <ul style="list-style-type: none">– использовать информационные технологии и программное обеспечение современных хранилищ данных на основе использования аналитических платформ для анализа деятельности предприятия. <p>Владеть:</p> <ul style="list-style-type: none">– методами получения, анализа и использования информации, необходимой для осуществления бизнес-анализа деятельности экономических объектов.

ПК-6	Управление контентом предприятия и Интернет-ресурсов, процессами создания и использования информационных ресурсов (контент-сервисов).	Знать: – многомерное представление данных, виды метаданных, используемых в хранилищах данных.
		Уметь: – применять на практике возможности аналитической системы Deductor для построения хранилища данных.
		Владеть: – приемами работы по созданию хранилища данных, принципами разработки для них моделей данных.

4. Структура и содержание дисциплины Б1.2.14 «Хранилище данных и знаний»

4.1. Структура дисциплины

Общая трудоемкость дисциплины для очной формы обучения составляет 6 зачетных единиц, 180 часов

№ п/п	Наименование разделов и тем дисциплины	Семестр	Недели семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)						Формы текущего контроля успеваемости (по неделям семестра)
				Аудиторная работа			Самостоятельная работа			
				Всего	Лекция	Лабораторные занятия	Всего	Подготовка к аудиторным занятиям	Подготовка к экзамену	Прием лабораторных работ*
1.	Раздел 1. Использование ИТ в экономике и управлении.	7	1-2	4	4		11	10	1	1-2
	Тема 1.1. Информационные ресурсы. Современные информационные технологии (ИТ).		1		2		5.5	5	0.5	1
	Тема 1.2. Корпоративные информационные системы (КИС).		2		2		5.5	5	0.5	2
2.	Раздел 2. Современные ИТ. OLTP и OLAP технологии. Хранилище данных.	7	3-12	54	22	32	52	45	7	3-12
	Тема 2.1. Многомерное представление данных.		3-4		8	8	17	15	2	3-4
	Тема 2.2. Многомерный анализ данных		5-8		10	12	29	26	3	5-8
	Тема 2.3. Процесс ETL (извле-		9-12		4	12	16	14	2	9-12

	чение, трансформация и загрузка данных).									
3.	Раздел 3. Аналитические платформы.	7	13-18	27	10	22	27	25	2	13-17
	Тема 3.1. Обзор современных аналитических платформ.		13-14		2	2	6	4	1	13-14
	Тема 3.2. Аналитическая платформа DEDUCTOR.		15-18		8	20	21	21	1	15-18
	Общая трудоемкость, в часах			90	36	54	90	80	10	Промежуточная аттестация
										Форма
										Семестр
										Экзамен
										7

• Прием лабораторных работ включает следующие составляющие:

- проведение собеседования для проверки уровня усвоения студентом теоретических знаний;
- выполнение практических заданий по теме лабораторной работы с целью проверки уровня умения и владения навыками работы.

4.2. Содержание дисциплины

4.2.1. Содержание лекционного курса

Разделы и их содержание:

Раздел 1. Использование ИТ в экономике и управлении.

Тема 1.1. Информационные ресурсы. Современные информационные технологии (ИТ).

Классификация ИТ по степени охвата задач управления. Классификация ИТ по степени централизации технологического процесса. Классификация информационных технологий по способу построения. Информация и информационные процессы в организационно-экономической сфере.

Роль и место автоматизированных информационных систем в экономике. Технология и методы обработки экономической информации. Проектирование автоматизированных информационных систем. Функциональные и обеспечивающие подсистемы.

Тема 1.2. Корпоративные информационные системы (КИС).

Роль и место информационных систем (ИС) в управлении экономическими объектами.

Корпоративные ИС (КИС). Основные понятия. Классификация корпоративных ИС. Структура ИС. Виды КИС.

Системы поддержки принятия решений (СППР). Системы управления ресурсами предприятия MRP, MRPII, ERP/ Системы управления взаимоотношениями с пользователями (CRM).

Раздел 2. Современные ИТ. OLTP и OLAP технологии. Хранилище данных.

Тема 2.1. Многомерное представление данных.

Определение и основные функции. Состав корпоративного хранилища данных. Источники данных для ХД. Принципы организации. Многомерное представление данных. Метаданные. Виды метаданных, используемых в хранилищах данных. Оперативный склад данных (ODS). Витрины (киоски) данных. Средства выборки, реструктуризации и доставки данных (SRD).

Тема 2.2. Многомерный анализ данных

Представление данных в виде гиперкуба.

Оперативный анализ данных (OLAP) и оперативная обработка данных (OLTP). Типы OLAP. Многомерный OLAP (MOLAP). Реляционный ROLAP (Relational OLAP). Гибридный HOLAP (Hybrid OLAP). Преимущества и недостатки.

Хранилище данных. Принципы организации хранилища. Структура хранилища данных. Таблицы фактов и измерений. Представление многомерной модели данных в виде схем «звезда» и «снежинка». Преимущества и недостатки каждой из них.

Тема 2.3. Процесс ETL (извлечение, трансформация и загрузка данных).

Основные цели и задачи процесса ETL. Подходы к реализации ETL-процесса. Проблема подготовки данных. Необходимость использования ETL. Определение и структура ETL. Подходы к реализации ETL-процесса.

Извлечение данных в ETL. Критерии оценки качества данных.

Очистка данных в ETL. Два уровня очистки данных.

Преобразование данных в ETL. Преобразование структуры данных. Агрегирование данных. Загрузка данных в хранилище.

Раздел 3. Аналитические платформы.

Тема 3.1. Обзор современных аналитических платформ.

Аналитические платформы бизнес аналитики. Назначение и особенности. Инструменты анализа Настольные и статистические пакеты. СУБД с набором алгоритмов Data Mining. Способы аналитической обработки данных.

Тема 3.2. Аналитическая платформа DEDUCTOR.

Общие сведения. Состав. Основные возможности по подготовке и очистке данных в DEDUCTOR. Понятия проекта, сценария и узла обработки Базовые операции над узлами сценария. Базовые визуализаторы. Очистка данных. Преобразование данных. Узлы сортировки, замены и фильтрации. Узлы квантования и кросс-таблицы. Узел калькулятор.

4.2.2. Перечень и содержание лабораторных занятий

№ п/п	№ разделов и тем курса	Наименование лабораторных работ	Количество часов
1	1.1, 1.2, 3.2	Развитие и назначение DEDUCTOR.	2
2	2.1, 2.2, 3.2	Освоение работы с системой DEDUCTOR.	2
3	2.2, 2.3, 3.2	Использование сценариев.	4
4	2.1, 2.2, 2.3, 3.2	Загрузка текстового файла в ХД.	8
5	3.4, 3.5	Базовые визуализаторы.	6
6	2.2, 2.3, 3.2	Очистка данных.	7
7	2.3, 3.2	Узлы Сортировка, замена и фильтрация.	4
8	2.3, 3.2	Узлы Квантование. Кросс-таблица. Преобразование данных к скользящему окну.	4
9	2.3, 3.2	Узел Калькулятор.	4
10	2.1, 2.2, 2.3, 3.1 3.2	Создание структуры хранилища данных в виде схемы «звезда» согласно концептуальной модели заданной предметной области.	13

5. Образовательные технологии

Чтение лекций по дисциплине проводится с использованием мультимедийного компьютерного проектора и проблемного подхода; проведение лабораторных занятий в компьютерном классе, использование проектного подхода в лабораторном практикуме.

При изучении материалов лабораторного практикума используются образовательные материалы, программное обеспечение и информационные размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru).

В самостоятельной работе также используются образовательные материалы, программное обеспечение и информационные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru)

В лабораторном практикуме используются приложение Deductor Academic.

Технология получения студентами теоретических знаний основана на принципах самостоятельной внеаудиторной работы с рекомендуемым учебно-методическим и информационным обеспечением учебной дисциплины.

Технология формирования и закрепления практических навыков и умений у студентов базируется на выполнении и защите лабораторных и курсовых работ, основанных на компьютерных симуляциях, связанных с созданием и разработкой экономических документов и проведением экономических расчетов.

В целях реализации индивидуального подхода к обучению студентов, осуществляющих учебный процесс по собственной траектории в рамках индивидуального рабочего плана, изучение данной дисциплины базируется на следующих возможностях: обеспечение внеаудиторной работы со студентами в том числе в электронной образовательной среде с использованием соответствующего программного оборудования, дистанционных форм обучения, возможностей интернет-ресурсов, индивидуальных консультаций и т.д.

Обучение лиц с ограниченными возможностями здоровья проводится в зависимости от их индивидуальных потребностей. При необходимости обучающимся с ограниченными возможностями здоровья предоставляется социально-психологическая помощь и сопровождение. Обучающиеся из числа лиц с ограниченными возможностями здоровья обеспечиваются печатными и электронными образовательными ресурсами в формах, адаптированных к ограничениям их здоровья.

6. Учебно-методическое обеспечение самостоятельной работы студентов. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

Целью проведения текущего контроля является оценка текущих знаний, умений, приобретенных навыков и посещения студентами занятий.

Текущий контроль знаний студентов производится на лабораторных занятиях.

Для оценки уровня освоения материала разработана система лабораторных работ, которые прилагаются. Общее количество работ составляет 10.

После выполнения лабораторной работы студент должен представить отчет по ней и защитить ее. При защите лабораторной работы оцениваются уровень знаний студентов, качество выполнения лабораторной работы, активность студентов и подготовленность к занятию. Объектом оценки являются не только знания по теоретическому содержанию данной работы, но и приобретенные навыки практической работы.

Результаты текущего контроля являются второй составляющей необходимой информации для определения интегральной оценки знаний студентов по изучаемому курсу.

Промежуточный контроль осуществляется путем проверки знаний, умений и приобретенных навыков по результатам изучения основных тем учебной дисциплины, соответствующих ее структуре.

Особое внимание обращено на промежуточный контроль по ключевым темам изучаемого курса.

По второму разделу к ним относятся темы по изучению принципов построения корпоративных информационных систем, разработки модели данных, получение знаний об использовании процесса извлечения, преобразования и загрузки (ETL) данных при создании и функционировании хранилища данных.

По третьему разделу к основным относятся темы, связанные с получением знаний об основных возможностях и принципах работы с аналитической платформы Deductor Academic для создания хранилища данных и многомерного анализа.

Для оценки уровня освоения материала проводится прием лабораторных работ. В ходе приема работы студент должен продемонстрировать полученные знания и навыки, как теоретические, так и практические.

Результаты приема являются одной из составляющих необходимой информации для определения интегральной оценки знаний студентов по изучаемому курсу.

Кроме того, промежуточная оценка умений и приобретенных навыков производится при контроле этапов выполнения самостоятельной работы, которая также отражает основное содержание изучаемых тем курса. Умения и навыки демонстрируются путем представления отчета по самостоятельной работе.

Результаты защиты самостоятельных работ по этапам их выполнения являются второй составляющей необходимой информации для определения интегральной оценки знаний студентов по изучаемому курсу.

Общая оценка по изучаемому курсу отражает результаты промежуточного, текущего контроля и знания, показанные непосредственно на экзамене.

Результаты промежуточного и текущего контроля фиксируются в рабочей тетради преподавателя.

6.1. План самостоятельной работы студентов

План самостоятельной работы студентов очной формы обучения

№ нед.	Тема	Вид самостоятельной работы	Задание	Рекомендуемая литература	Количество часов
1	Тема 1.1. Информационные ресурсы. Современные информационные технологии (ИТ).	Подготовка к аудиторным занятиям по темам лекционных занятий (см. п. 4.2.1) и лабораторных занятий (см. п. 4.2.2)	Изучить классификацию ИТ и их использование в экономике.	Учебно- методические материалы и электронные учебные ресурсы, размещенные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература.	5

2	Тема 1.2. Корпоративные информационные системы (КИС).	Подготовка к аудиторным занятиям	Ознакомиться с возможностями КИС (системы управления ресурсами предприятия).	Учебно- методические материалы и электронные учебные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература.	5
3-4	Тема 2.1. Многомерное представление данных.	Подготовка к аудиторным занятиям	Проанализировать предметную область. Определить сущности, атрибуты и отношения.	Учебно- методические материалы и электронные учебные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература.	15
5-8	Тема 2.2. Многомерный анализ данных	Подготовка к аудиторным занятиям	Ознакомиться с типами OLAP. Разработать ненормализованную модель «снежинка».	Учебно- методические материалы и электронные учебные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература	26
9-12	Тема 2.3. Процесс ETL (извлечение, трансформация и загрузка данных).	Подготовка к аудиторным занятиям	Изучить основные цели и задачи процесса ETL и его структуру.	Учебно- методические материалы и электронные учебные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература	14

13-14	Тема 3.1. . Обзор современных аналитических платформ.	Подготовка к аудиторным занятиям	Ознакомиться с современными аналитическими платформами бизнес аналитики.	Учебно- методические материалы и электронные учебные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература	4
15-18	Тема 3.2. Аналитическая платформа DEDUCTOR.	Подготовка к аудиторным занятиям	Изучить состав, основные возможности по подготовке и очистке данных приложения DEDUCTOR.	Учебно- методические материалы и электронные учебные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература	11
1-18		Подготовка к экзамену		Учебно- методические материалы и электронные учебные ресурсы, размещенные и/или указанные в разделе дисциплины на учебном портале ПГУ (moodle.pnzgu.ru) Основная и дополнительная литература	10

6.2. Методические указания по организации самостоятельной работы студентов

В ходе изучения данной дисциплины предусмотрена самостоятельная подготовка к сдаче лабораторных работ в форме собеседования и выполнения практических заданий на основе использования аналитической платформы Deductor Academic.

Каждый студент получает индивидуальное задание по разработке хранилища данных для хранения и обработки информации в заданной предметной области.

При сдаче лабораторной работы необходимо продемонстрировать знания, умения и навыки, полученные при изучении учебного материала.

Перечень тем для выбора предметной области:

1. Студенческая библиотека
2. Страховая фирма
3. Агентство недвижимости

4. Отдел кадров производственного предприятия
5. Фирма покупки и продажи автомобилей
6. Гостиница
7. Железнодорожные кассы
8. Турагентство
9. Продуктовый магазин
10. Учет материалов на складе
11. Кредитование
12. Учет оплаты услуг предприятия - автосервиса
13. Учет перевозок грузов на автопредприятии
14. Учет расчетов с арендаторами
15. Учет исполнения договоров строительной фирмы
16. Учет труда и начисление заработной платы
17. Больница
18. Расчет квартплаты ТСЖ

6.3. Материалы для проведения текущего и промежуточного контроля знаний студентов

Контроль освоения компетенций

Контроль освоения компетенций проводится в такой форме текущего контроля успеваемости, как прием лабораторных работ. Этот процесс включает следующие составляющие:

- проведение собеседования для проверки уровня усвоения студентом теоретических знаний;
- выполнение практических заданий по теме лабораторной работы с целью проверки уровня умения и владения навыками работы в области разработки хранилища данных в соответствии с заданным вариантом на основе использования аналитической платформы Deductor Academic.

№ п\п	Вид контроля	Контролируемые темы (разделы)	Компетенции, компоненты которых контролируются
1.	Прием лабораторных работ.	Раздел 1. Использование ИТ в экономике и управлении. Основные понятия. Лабораторная работа № 1	ПК-3, ПК-6
2.	Прием лабораторных работ.	Раздел 2. Современные ИТ. OLTP и OLAP технологии. Хранилище данных. Лабораторные работы № 2, 3, 4, 6, 7, 8, 9, 10	ПК-3, ПК-6
3.	Прием лабораторных работ.	Раздел 3. Аналитические платформы. Лабораторные работы № 1, 2, 3, 4, 5, 6, 7, 8, 9, 10	ПК-3, ПК-6
4.	Подготовка к зачету и экзамену	Разделы 1-3	ПК-3, ПК-6

Вопросы и задания для приема лабораторных работ:

Лабораторная работа № 1. Развитие и назначение DEDUCTOR.

1. Из каких частей состоит Deductor?
2. Какие варианты поставки Deductor существуют?
3. Чем отличается версия Professional от Academic?
4. Имеются ли ограничения по количеству обрабатываемых записей в версии Deductor Academic?
5. Сколько категорий пользователей Deductor можно выделить?
6. Перечислите функции аналитика.
7. Кто обычно занимается проектированием и наполнением хранилища данных?
8. Каким образом лицензируется Deductor?

Лабораторная работа № 2. Освоение работы с системой DEDUCTOR.

1. Сколько вкладок на панели управления Deductor Studio?
2. Что такое проект в Deductor Studio?
3. Какое расширение имеет файл проекта?
4. Как создать новый проект?
5. Как сохранить текущий проект под другим именем?
6. Как отредактировать свойства проекта?
7. Сколько проектов можно одновременно открыть в Deductor Studio?
8. Сколько мастеров имеется в Deductor Studio?

Лабораторная работа № 3. Использование сценариев.

1. Сколько вкладок на панели управления Deductor Studio?
2. Что такое «проект» в Deductor Studio?
3. Какое расширение имеет файл проекта?
4. Как создать новый проект?
5. Как сохранить текущий проект под другим именем?
6. Сколько проектов можно одновременно открыть в Deductor Studio?
7. Сколько мастеров имеется в Deductor Studio?
8. Как скопировать ветвь сценария при помощи механизма drag & drop?
9. Что означает пиктограмма «!» напротив узла сценария?

Лабораторная работа № 4. Загрузка текстового файла в ХД.

1. Какие шаги мастера импорта нужно пройти для импорта текстового файла?
2. Как установить некоторые параметры форматирования?
3. При импортировании текстового файла был создан узел импорта. После чего обнаружено, что неправильно заданы параметры импорта. Как легче всего исправить ошибку?
4. Что позволяет сделать обработчик Настройка набора данных?
5. Как происходит автоматическое определение типа столбца при импорте?

Лабораторная работа № 5. Базовые визуализаторы.

1. Какие характеристики набора данных показывает визуализатор Статистика?
2. Что означает красный заголовок столбца в визуализаторе Таблица?
3. Как обнаружить, имеются ли в столбце пропущенные значения?
4. Для чего предназначен визуализатор Сведения?
5. Как скрыть столбец в визуализаторе Таблица?
6. К существующему в сценарии узлу импорта необходимо добавить еще один визуализатор. Что предпринять?

Лабораторная работа № 6. Очистка данных.

1. Назначения и ход выполнения Парциальной обработки
2. Назначение и ход выполнения Спектральной обработки
3. Назначение и ход выполнения Корреляционного анализа
4. Назначение и ход выполнения Выявления дубликатов и противоречий

Лабораторная работа № 7. Узлы Сортировка, замена и фильтрация.

1. Как работает обработчик Сортировка?
2. Можно ли отсортировать набор данных по нескольким полям?
3. Для чего предназначен узел Замена данных?
4. Как определить в мастере обработки, что для поля настроена замена?
5. Как работает Замена данных?
6. Какие существуют способы заполнить таблицу подстановок?
7. Для чего предназначен узел Фильтр?
8. Какие условия фильтрации существуют?
9. Сколько записей будет отфильтровано в результате фильтра «([Размер ссуды, руб] в интервале [2000.5000]) И ([Цель ссуды] = 'Покупка товара') И ([Цель ссуды] =

'Иное')»)?

10. Что делать, если нужно поставить фильтр по значению, которого в данный момент нет в рассматриваемом наборе данных?

Лабораторная работа № 8. Узлы Квантование. Кросс-таблица. Преобразование данных к скользящему окну.

1. Назначения и ход выполнения Квантования данных
2. Назначение обработчика Кросс таблица, порядок выполнения
3. Назначение обработчика Преобразование данных к скользящему окну, ход выполнения

Лабораторная работа № 9. Узел Калькулятор.

1. Для чего предназначен обработчик Калькулятор?
2. Как добавить новый столбец?
3. Какой символ используется для разделения параметров в функциях калькулятора?
4. Как ввести формулу для расчета значений столбца?
5. Как вывести подсказку для функции в окне создания выражений?
6. Чем отличаются функции IF и IFF?
7. Что делает функция ISNULL?
8. Как добавить существующее имя поля в формулу?
9. Как посмотреть описание той или иной функции?
10. Что делают следующие функции: NOW(), TODAY(), ROUND(), POW()?
11. Что будет, если в Калькуляторе создать новый столбец вещественного типа и написать для него формулу 15/0?

Лабораторная работа № 10. Создание структуры хранилища данных в виде схемы «звезда» согласно концептуальной модели заданной предметной области.

1. Как создать новое хранилище данных с помощью мастера подключений?
2. Что представляет многомерная модели данных «звезда» и «снежинка»?
3. Создание фактов, измерений и процессов.

Контроль освоения компетенции выполняется:

- для компетенций (ПК-3, ПК-6) путем оценки степени способности студента осваивать методики использования программных средств для решения практических задач в объеме данного курса.

6.4. Вопросы для подготовки к экзамену

1. Обработка транзакций в реальном времени (OLTP).
2. Система поддержки принятия решений.
3. Оперативный анализ данных (OLAP).
4. OLAP и OLTP. Сравнение.
5. Хранилище данных. Определение и основные функции. Принципы организации.
6. Аналитические платформы. Назначение и особенности.
7. Аналитическая платформа DEDUCTOR. Общие сведения. Состав.
8. Аналитическая платформа DEDUCTOR. Понятия проекта, сценария и узла обработки
9. Аналитическая платформа DEDUCTOR. Базовые операции над узлами сценария.
10. Аналитическая платформа DEDUCTOR. Базовые визуализаторы.
11. Аналитическая платформа DEDUCTOR. Очистка данных.
12. Аналитическая платформа DEDUCTOR. Узлы сортировки, замены и фильтрации.
13. Аналитическая платформа DEDUCTOR. Узлы квантования и кросс-таблицы.
14. Аналитическая платформа DEDUCTOR. Узел калькулятор.
15. Многомерный анализ данных (OLAP).
16. Многомерное представление данных.
17. Типы OLAP. Преимущества и недостатки.
18. Представление данных в виде гиперкуба.

19. Многомерный OLAP (MOLAP).
20. ROLAP (Relational OLAP).
21. HOLAP (Hybrid OLAP).
22. Схема «звезда». Преимущества и недостатки.
23. Схема «снежинка». Преимущества и недостатки.
24. Структура хранилища данных. Таблица фактов.
25. Структура хранилища данных. Таблица измерений.
26. Технические аспекты многомерного хранения данных.
27. Проблема подготовки данных. Необходимость использования ETL.
28. Определение и структура ETL.
29. Основные цели и задачи процесса ETL.
30. Подходы к реализации ETL-процесса.
31. Структура процесса ETL. Извлечение данных в ETL.
32. Очистка данных в ETL. Два уровня очистки данных.
33. Критерии оценки качества данных.
34. Преобразование данных в ETL. Преобразование структуры данных. Агрегирование данных.
35. Очистка данных.
36. Загрузка данных в хранилище.
37. Состав корпоративного хранилища данных.
38. Источники данных для ХД.
39. Метаданные. Виды метаданных, используемых в хранилищах данных.
40. Оперативный склад данных (ODS).
41. Витрины (киоски) данных.
42. Средства выборки, реструктуризации и доставки данных (SRD).
43. ИТ в управлении. Основные понятия.
44. Корпоративные информационные системы.
45. Какие задачи решают СППР?
46. На какие типы делятся задачи анализа данных?
47. Из чего состоит архитектура СППР?
48. Каковы требования к системам OLTP и СППР?
49. Что такое хранилище данных?
50. Что такое виртуальное и физическое хранилище данных?
51. Что такое витрина данных?
52. Из чего состоит хранилище данных?
53. Какие потоки данных имеются в хранилище данных?
54. Какие есть приемы оптимизации хранилищ данных?
55. Для чего нужна денормализация данных?
56. Какие операции можно производить над гиперкубом?
57. Поясните все составные части теста FASMI
58. Достоинства и недостатки MOLAP
59. Достоинства и недостатки ROLAP
60. Основные задачи процесса ETL
61. Извлечение данных в ETL
62. Выбор используемых источников данных в ETL. Особенности организации процесса извлечения данных
63. Очистка данных в ETL
64. Основные виды проблем в данных, из-за которых они нуждаются в очистке
65. Преобразование данных в ETL
66. Преобразование структуры данных
67. Агрегирование данных
68. Перевод значений

69. Создание новых данных
70. Преимущества и недостатки отказа от хранилищ данных
71. Квантование (дискретизация)
72. Что такое Data Mining?
73. Основные задачи Data Mining
74. Этапы интеллектуального анализа данных
75. Методы интеллектуального анализа данных

7. Учебно-методическое и информационное обеспечение дисциплины

а) основная литература:

1. Информационные системы предприятия: Учебное пособие / А.О. Варфоломеева, А.В. Коряковский, В.П. Романов. - М.: НИЦ ИНФРА-М, 2016. - 283 с.
<http://znanium.com/catalog.php?item=bookinfo&book=536732>
ЭБС Знаниум

2. Методы, модели, средства хранения и обработки данных: учебник / Э.Г. Дадян, Ю.А. Зеленков. — М.: Вузовский учебник: ИНФРА-М, 2017. — 168 с.
<http://znanium.com/catalog.php?bookinfo=543943>
ЭБС Знаниум

б) дополнительная литература:

1. Информационные технологии в менеджменте: Учебное пособие / В.И. Карпузова, Э.Н. Скрипченко, К.В. Чернышева, Н.В. Карпузова. - 2-е изд., доп. - М.: Вузовский учебник: НИЦ ИНФРА-М, 2014. - 301 с.
<http://znanium.com/catalog.php?bookinfo=410374>
ЭБС Знаниум

г) программное обеспечение и Интернет-ресурсы:

1. Сайт разработчика и поставщика программных продуктов и решений в области анализа данных BaseGroup Labs. <http://www.basegroup.ru/>.
2. Deductor Academic - <http://www.basegroup.ru/download/deductor/>
3. Интернет-университет информационных технологий - www.intuit.ru

8. Материально-техническое обеспечение дисциплины

Классы, оборудованные персональными компьютерами, объединёнными в локальную сеть, операционная система Windows версии XP и выше, пакет офисных прикладных программ Microsoft Office версии 2003 и выше, аналитическая платформа Deductor Academic, доступ к Интернет-ресурсам.

Рабочая программа дисциплины составлена в соответствии с требованиями ФГОС ВО.

Программу составил:

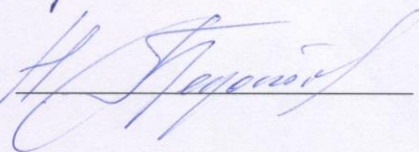
к.т.н., доцент Петренко А.Г.



Программа одобрена на заседании кафедры «Экономическая кибернетика»

Протокол № 1 от «1» сентября 2017 года

Зав. кафедрой ЭК, д.т.н., профессор



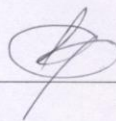
Федотов Н.Г.

Программа одобрена методической комиссией факультета ЭиУ

Протокол № 1 от «28» сентября 2017 года

Председатель методической

комиссии факультета ЭиУ



Еремина Е. В.

