

МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»

УТВЕРЖДАЮ:

Директор ИЭФ



Ю.И. Соколов

26 июня 2019 г.



Кафедра «Информационные системы цифровой экономики»

Автор Медникова Оксана Васильевна, к.т.н.

РАБОЧАЯ ПРОГРАММА УЧЕБНОЙ ДИСЦИПЛИНЫ

Техники, методы анализа и обработки больших данных

Направление подготовки:	<u>09.03.03 – Прикладная информатика</u>
Профиль:	<u>Прикладная информатика в бизнесе</u>
Квалификация выпускника:	<u>Бакалавр</u>
Форма обучения:	<u>очная</u>
Год начала подготовки	<u>2019</u>

<p style="text-align: center;">Одобрено на заседании Учебно-методической комиссии института Протокол № 8 25 июня 2019 г. Председатель учебно-методической комиссии</p>  <p style="text-align: right;">М.В. Ишханян</p>	<p style="text-align: center;">Одобрено на заседании кафедры</p> <p>Протокол № 15 24 июня 2019 г. Заведующий кафедрой</p>  <p style="text-align: right;">О.В. Ефимова</p>
---	--

Рабочая программа учебной дисциплины (модуля) в виде электронного документа выгружена из единой корпоративной информационной системы управления университетом и соответствует оригиналу

Простая электронная подпись, выданная РУТ (МИИТ)
ID подписи: 2017
Подписал: Заведующий кафедрой Ефимова Ольга Владимировна
Дата: 24.06.2019

Москва 2019 г.

1. ЦЕЛИ ОСВОЕНИЯ УЧЕБНОЙ ДИСЦИПЛИНЫ

Целью дисциплины «Техники, методы анализа и обработки больших данных» является получение студентами знаний в области технологий хранения данных при принятии решения. Разобрать хранилища данных, как предметно-ориентированную информационную базу данных, специально разработанную и предназначенную для подготовки отчётов и бизнес-анализа с целью поддержки принятия решений в организации. Научить заполнять и использовать хранилища данных, а также представлять данные, полученные из хранилища данных с помощью запросов в удобном виде. Знания и практические навыки, полученные в курсе «Хранилища данных в системах электронного бизнеса», используются обучаемыми при разработке курсовых и дипломных работ.

2. МЕСТО УЧЕБНОЙ ДИСЦИПЛИНЫ В СТРУКТУРЕ ОП ВО

Учебная дисциплина "Техники, методы анализа и обработки больших данных" относится к блоку 1 "Дисциплины (модули)" и входит в его вариативную часть.

2.1. Наименования предшествующих дисциплин

Для изучения данной дисциплины необходимы следующие знания, умения и навыки, формируемые предшествующими дисциплинами:

2.1.1. Информационные системы и технологии:

Знания: использует современные информационные технологии и программные средства, в том числе отечественного производства при решении задач профессиональной деятельности

Умения: выбирает современные информационные технологии и программные средства, в том числе отечественного производства при решении задач профессиональной деятельности

Навыки: применяет современные информационных технологи и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности

2.2. Наименование последующих дисциплин

Результаты освоения дисциплины используются при изучении последующих учебных дисциплин:

2.2.1. Разработка программных приложений

Знания: Управление эффективностью ресурсного обеспечения проекта.

Умения: Управление взаимоотношениями с заинтересованными лицами в процессе обеспечения техническими, технологическими и человеческими ресурсами.

Навыки: Формирование и развитие команды проекта в области ИТ, планирование и управление персоналом проекта.

**3. ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ),
СООТНЕСЕННЫЕ С ПЛАНИРУЕМЫМИ РЕЗУЛЬТАТАМИ ОСВОЕНИЯ
ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ**

В результате освоения дисциплины студент должен:

№ п/п	Код и название компетенции	Ожидаемые результаты
1	ПКС-4 Способен принимать решения по управлению техническими, программно-технологическими и информационно-коммуникационными ресурсами	ПКС-4.1 Управление эффективностью ресурсного обеспечения ИТ проекта. ПКС-4.2 Управление взаимоотношениями с заинтересованными лицами в процессе обеспечения техническими, технологическими информационно-коммуникационными ресурсами.

4. ОБЪЕМ ДИСЦИПЛИНЫ (МОДУЛЯ) В ЗАЧЕТНЫХ ЕДИНИЦАХ И АКАДЕМИЧЕСКИХ ЧАСАХ

4.1. Общая трудоемкость дисциплины составляет:

2 зачетные единицы (72 ак. ч.).

4.2. Распределение объема учебной дисциплины на контактную работу с преподавателем и самостоятельную работу обучающихся

Вид учебной работы	Количество часов	
	Всего по учебному плану	Семестр 6
Контактная работа	56	56,15
Аудиторные занятия (всего):	56	56
В том числе:		
лекции (Л)	28	28
практические (ПЗ) и семинарские (С)	28	28
Самостоятельная работа (всего)	16	16
ОБЩАЯ трудоемкость дисциплины, часы:	72	72
ОБЩАЯ трудоемкость дисциплины, зач.ед.:	2.0	2.0
Текущий контроль успеваемости (количество и вид текущего контроля)	ПК1, ПК2	ПК1, ПК2
Виды промежуточной аттестации (экзамен, зачет)	ЗЧ	ЗЧ

4.3. Содержание дисциплины (модуля), структурированное по темам (разделам)

№ п/п	Семестр	Тема (раздел) учебной дисциплины	Виды учебной деятельности в часах/ в том числе интерактивной форме						Формы текущего контроля успеваемости и промежуточной аттестации
			Л	ЛР	ПЗ/ТП	КСР	СР	Всего	
1	2	3	4	5	6	7	8	9	10
1	6	Раздел 1 Технологии организации хранилищ данных	7		10			17	
2	6	Тема 1.1 Технология хранения данных при принятии решений.	1		7			8	
3	6	Тема 1.2 Хранилище данных (Data Warehousing) в виде ненормализованных баз данных.	2		2			4	
4	6	Тема 1.3 Многомерные системы управления базами данных – МСУБД.	4		1			5	
5	6	Раздел 2 Работа с хранилищами данных	7		7		0	14	
6	6	Тема 2.1 Заполнение хранилища данными.	5		1		0	6	
7	6	Тема 2.2 Создание многомерного хранилища данных (куба).	2		1			3	ПК1, ситуационные задачи
8	6	Раздел 3 Анализ данных хранилища	9		4			13	
9	6	Тема 3.1 Технология аналитической обработки данных (OLAP).	2					2	
10	6	Тема 3.2 Клиент Microsoft Data Analyzer.	7		2			9	
11	6	Тема 3.3 Аналитические возможности Analysis Manager.			2			2	
12	6	Раздел 4 Внешние средства работы с хранилищами данных	5		7		16	28	
13	6	Тема 4.1 Интеграция Web - технологии и	4		2			6	

№ п/п	Семестр	Тема (раздел) учебной дисциплины	Виды учебной деятельности в часах/ в том числе интерактивной форме						Формы текущего контроля успеваемости и промежуточной аттестации
			Л	ЛР	ПЗ/ТП	КСР	СР	Всего	
1	2	3	4	5	6	7	8	9	10
		технологии Хранилища.							
14	6	Тема 4.2 Язык MDX (Multidimensional Expressions)- непроцедурный язык для формулирования запросов к многомерным базам данных.			2		6	8	
15	6	Тема 4.3 Поддержка распределенных данных			2		5	7	ПК2, тествыборочный опрос
16	6	Тема 4.4 Специализированные средства представления отчетности.	1		1		5	7	
17		Раздел 5 Дифференцированный зачет							
18		Всего:	28		28		16	72	

4.4. Лабораторные работы / практические занятия

Лабораторные работы учебным планом не предусмотрены.

Практические занятия предусмотрены в объеме 28 ак. ч.

№ п/п	№ семестра	Тема (раздел) учебной дисциплины	Наименование занятий	Всего часов/ из них часов в интерактивной форме
1	2	3	4	5
1	6	РАЗДЕЛ 1 Технологии организации хранилищ данных Тема: Технология хранения данных при принятии решений.	Ознакомление с CASE средствами проектирования хранилища данных. Проектирование таблицы фактов и таблиц измерений. Установление связей между таблицами. Генерация хранилища.	2
2	6	РАЗДЕЛ 1 Технологии организации хранилищ данных Тема: Технология хранения данных при принятии решений.	Ознакомление с CASE средствами проектирования хранилища данных. Проектирование таблицы фактов и таблиц измерений. Установление связей между таблицами. Генерация хранилища.	2
3	6	РАЗДЕЛ 1 Технологии организации хранилищ данных	Технология хранения данных при принятии решений.	5
4	6	РАЗДЕЛ 1 Технологии организации хранилищ данных Тема: Хранилище данных (Data Warehousing) в виде ненормализованных баз данных.	Описание источников данных. Описание потоков данных. Описание преобразования данных. Выполнение DTS. Просмотр таблиц в Query Analyzer.	2
5	6	РАЗДЕЛ 1 Технологии организации хранилищ данных Тема: Многомерные системы управления базами данных – МСУБД.	Организация доступа к источнику данных – оперативной реляционной базе данных. Построение куба. Настройка измерений и фактов.	1
6	6	РАЗДЕЛ 2 Работа с хранилищами данных Тема: Заполнение хранилища данными.	Использование кросс-таблиц. Управления отображением вычисляемых значений. Операции с таблицами. Экспорт данных в другие аналитические модули.	1
7	6	РАЗДЕЛ 2 Работа с хранилищами данных Тема: Создание многомерного хранилища данных (куба).	Получение MOLAP срезов кубов. Создание сводных диаграмм с данными OLAP-кубов.	1

№ п/п	№ семестра	Тема (раздел) учебной дисциплины	Наименование занятий	Всего часов/ из них часов в интерактивной форме
1	2	3	4	5
8	6	РАЗДЕЛ 3 Анализ данных хранилища Тема: Клиент Microsoft Data Analyzer.	Возможности и область применения Microsoft Data Analyzer Средства анализа данных.	2
9	6	РАЗДЕЛ 3 Анализ данных хранилища Тема: Аналитические возможности Analysis Manager.	Средства публикации и создания отчетов. Публикация сводной таблицы на Web. Создание Web-страниц со сводными диаграммами.	1
10	6	РАЗДЕЛ 3 Анализ данных хранилища Тема: Аналитические возможности Analysis Manager.	Назначение языка MDX. Синтаксис запроса на языке MDX.	1
11	6	РАЗДЕЛ 4 Внешние средства работы с хранилищами данных Тема: Интеграция Web - технологии и технологии Хранилища.	Мастер построения модели DataMining. Выбор критерия для анализа.	2
12	6	РАЗДЕЛ 4 Внешние средства работы с хранилищами данных Тема: Язык MDX	Включение и отключение публикующего и распределительного серверов Создание транзакционных и мгновенных публикаций. Принудительная подписка. Публикация в Internet.	2
13	6	РАЗДЕЛ 4 Внешние средства работы с хранилищами данных Тема: Поддержка распределенных данных	Обеспечение доступа к данным.	2
14	6	РАЗДЕЛ 4 Внешние средства работы с хранилищами данных Тема: Специализированные средства представления отчетности.	Создание отчетов с разрывами Установка срезов (получение Кросс-таблиц). Построение графических отчетов (диаграмм). Средства доступа и генерация отчетов Business Object.	1
15	6		Работа с хранилищами данных	5
ВСЕГО:				30/0

4.5. Примерная тематика курсовых проектов (работ)

Курсовые работы (проекты) не предусмотрены.

5. ОБРАЗОВАТЕЛЬНЫЕ ТЕХНОЛОГИИ

В обучении студентов по данной дисциплине используются:

1. при проведении лекционных занятий:

вводная;

лекция-информация;

проблемная лекция;

лекция визуализация;

2. для проведения лабораторных занятий:

проектная технология;

технология учебного исследования;

техника «круглый стол»,

технология обучения в сотрудничестве и в малых группах;

технология проблемного обучения;

разбор конкретных ситуаций.

6. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДЛЯ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

№ п/п	№ семестра	Тема (раздел) учебной дисциплины	Вид самостоятельной работы студента. Перечень учебно-методического обеспечения для самостоятельной работы	Всего часов
1	2	3	4	5
1	6	РАЗДЕЛ 4 Внешние средства работы с хранилищами данных Тема 2: Язык MDX	Создание транзакционных и мгновенных публикаций. Выбор сервера и БД. Выбор сервера подписчика и таблиц для публикации. Включение в фильтрации столбцов и строк выбранных таблиц. Настройка расписания для агента мгновенной репликации. . Назначение подписки для серверов подписчиков. Управление распределением данных. Поддержка неоднородных репликаций. Публикация в Internet. Диспетчер репликации. Работа с журналом ошибок. [4, стр. 44-71]	6
2	6	РАЗДЕЛ 4 Внешние средства работы с хранилищами данных Тема 3: Поддержка распределенных данных	Установка связи с исходной Базой данных Выбор фрагмента Базы данных для включения в Юниверс. Создание структуры Юниверса. Обеспечение доступа к данным. Выбор переменных Юниверса для анализа. Замена иерархии переменных. Ввод в Юниверс расчетных переменных. [1, стр. 65-71]	5
3	6	РАЗДЕЛ 4 Внешние средства работы с хранилищами данных Тема 4: Специализированные средства представления отчетности.	Создание отчетов с разрывами Создание условий выбора. Установка срезов (получение Кросс-таблиц). Транспонирование Кросс-таблиц. Установка иерархических срезов. Использование фильтрации. Редактирование переменных в Юниверсе. Расчет суммирующих показателей по выбранному столбцу. Построение графических отчетов (диаграмм). Средства доступа и генерация отчетов Business Object. [1, стр. 234-267]	5
ВСЕГО:				16

7. ПЕРЕЧЕНЬ ОСНОВНОЙ И ДОПОЛНИТЕЛЬНОЙ ЛИТЕРАТУРЫ, НЕОБХОДИМОЙ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

7.1. Основная литература

№ п/п	Наименование	Автор (ы)	Год и место издания Место доступа	Используется при изучении разделов, номера страниц
1	Постреляционные хранилища данных : учебное пособие для вузов	Ю.П. Парфенов; под науч. ред. Н.В. Папуловской	М.: Издательство Юрайт, 2017 https://www.biblio-online.ru/viewer/628DAC6C-ECBF-45B3-BD23-F6B57148D18F#page/4	1-4 стр.4-121

7.2. Дополнительная литература

№ п/п	Наименование	Автор (ы)	Год и место издания Место доступа	Используется при изучении разделов, номера страниц
2	Методы и модели анализа данных OLAP и DataMining.	Барсегян А.А., Куприянов М.С., Степаненко В.В., Холод И.И.	Петербург, 2014	3
3	Microsoft SQL Server 2005. Библия пользователя.	А.В. Маркин	Диалектика, 2013	4
4	Программирование на sql в 2 ч. Часть 1: учебник и практикум для бакалавриата и магистратуры	А.В. Маркин	М.: Издательство Юрайт, 2017. - 362 с. - (Бакалавр и магистр. Академический курс). , 2017 https://www.biblio-online.ru/viewer/65D478FB-E9CC-444C-9015-237C4ECB0AA1#page/10	1-4 стр.10-362
5	Программирование на sql в 2 ч. Часть 2: учебник и практикум для бакалавриата и магистратуры	А.В. Маркин	М.: Издательство Юрайт, 2017. - 292 с. - (Бакалавр и магистр. Академический курс). , 0 https://www.biblio-online.ru/viewer/BCC5FE83-9878-4ED2-AB2A-DFC7E60C3847#page/5	1-4 стр.5-292

8. ПЕРЕЧЕНЬ РЕСУРСОВ ИНФОРМАЦИОННО-ТЕЛЕКОММУНИКАЦИОННОЙ СЕТИ "ИНТЕРНЕТ", НЕОБХОДИМЫЕ ДЛЯ ОСВОЕНИЯ ДИСЦИПЛИНЫ (МОДУЛЯ)

- Вузовские электронно-библиотечные системы учебной литературы.
- База научно-технической информации ВИНТИ РАН.
- Интернет-ресурсы:

<http://www54.sap.com/solutions/analytics/business-intelligence.html> – Business Intelligence Software, BI with Analytics | SAP.

<http://www.sql.ru/> – SQL.ru - все про SQL, базы данных, программирование и разработку информационных систем.

<http://msdn.microsoft.com/> – "MSDN" – информационный сервис для разработчиков.
<http://www.intuit.ru/> – ИНТУИТ – Интернет-Университет Информационных Технологий.

9. ПЕРЕЧЕНЬ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНФОРМАЦИОННЫХ СПРАВОЧНЫХ СИСТЕМ, ИСПОЛЬЗУЕМЫХ ПРИ ОСУЩЕСТВЛЕНИИ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

- Локальная сеть, состоящая из 20 рабочих мест (компьютеров), сервера, компьютера преподавателя.
- Проектор.
- Электронная доска.

10. ОПИСАНИЕ МАТЕРИАЛЬНО ТЕХНИЧЕСКОЙ БАЗЫ, НЕОБХОДИМОЙ ДЛЯ ОСУЩЕСТВЛЕНИЯ ОБРАЗОВАТЕЛЬНОГО ПРОЦЕССА ПО ДИСЦИПЛИНЕ (МОДУЛЮ)

Для проведения практических занятий:

- На сервере установлен Microsoft SQL Server 2000 (или более поздней версии) с с модулем Analysis Services компании Microsoft.
- Программное средство Data Analyser.
- Программное средство MOLAP компании Microsoft.
- Программное средство Power Designer Warehouse Architect 6.0, компании Sybase Inc., US. (или другое аналогичное Case – средство).

11. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ (МОДУЛЯ)

Самостоятельное изучение теоретического материала.

Целью самостоятельного изучения теоретического курса является закрепление лекционного материала для применения полученных знаний при выполнении лабораторных работ .

Порядок самостоятельной проработки вопросов должен совпадать с порядком чтения лекций преподавателем. Изучение всех вопросов, отведенных на самостоятельную проработку, и материала лекций контролируется посредством тестовых заданий.

Текущий контроль знаний по дисциплине выполняется два раза в семестр на контрольных неделях. Студенты должны продемонстрировать знание теоретического курса, выполнив тестовые задания (ТЗ). В тестовые задания включены вопросы, отведенные для самостоятельного изучения.