

**МИНИСТЕРСТВО ТРАНСПОРТА РОССИЙСКОЙ ФЕДЕРАЦИИ**  
**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ**  
**УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ**  
**«РОССИЙСКИЙ УНИВЕРСИТЕТ ТРАНСПОРТА»**  
**(РУТ (МИИТ))**



Рабочая программа дисциплины (модуля),  
как компонент образовательной программы  
высшего образования - программы бакалавриата  
по направлению подготовки  
38.03.05 Бизнес-информатика,  
утвержденной первым проректором РУТ (МИИТ)  
Тимониным В.С.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ (МОДУЛЯ)**

**Алгоритмы работы с большими данными**

Направление подготовки: 38.03.05 Бизнес-информатика

Направленность (профиль): Цифровая экономика

Форма обучения: Очная

Рабочая программа 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: 564169  
Подписал: заведующий кафедрой Каргина Лариса Андреевна  
Дата: 01.03.2023

## 1. Общие сведения о дисциплине (модуле).

Целями освоения дисциплины являются:

- развитие общематематической культуры, логического и алгоритмического мышления студентов;
- выработка умения моделировать реальные финансово-экономические процессы;
- освоение приемов исследования и решения математически формализованных задач.

Задачами освоения дисциплины является формирование у обучающихся навыков:

- получение необходимого математического аппарата для изучения дисциплин естественнонаучного и профессионального циклов и применения этого аппарата в будущей профессиональной деятельности.

## 2. Планируемые результаты обучения по дисциплине (модулю).

Перечень формируемых результатов освоения образовательной программы (компетенций) в результате обучения по дисциплине (модулю):

**ОПК-4** - Способен понимать принципы работы информационных технологий; использовать информацию, методы и программные средства ее сбора, обработки и анализа для информационно-аналитической поддержки принятия управленческих решений;

**ПК-7** - Способен проводить сбор информации о деятельности подразделения организации с целью разработки административного регламента подразделения организации.

Обучение по дисциплине (модулю) предполагает, что по его результатам обучающийся будет:

### **Знать:**

- принципы обработки больших данных, область их практического применения;
- принципы работы реляционных и нереляционных баз данных.

### **Уметь:**

- разрабатывать концептуальную модель прикладной области, выбирать методы анализа данных, а также адаптировать базовые методы к решению прикладных задач, формулировать содержательные выводы и рекомендации

### **Владеть:**

- навыками решения конкретных задач в профессиональной области.

### 3. Объем дисциплины (модуля).

#### 3.1. Общая трудоемкость дисциплины (модуля).

Общая трудоемкость дисциплины (модуля) составляет 4 з.е. (144 академических часа(ов)).

3.2. Объем дисциплины (модуля) в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении учебных занятий:

Тип учебных занятий	Количество часов	
	Всего	Сем. №7
Контактная работа при проведении учебных занятий (всего):	64	64
В том числе:		
Занятия лекционного типа	32	32
Занятия семинарского типа	32	32

3.3. Объем дисциплины (модуля) в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации составляет 80 академических часа (ов).

3.4. При обучении по индивидуальному учебному плану, в том числе при ускоренном обучении, объем дисциплины (модуля) может быть реализован полностью в форме самостоятельной работы обучающихся, а также в форме контактной работы обучающихся с педагогическими работниками и (или) лицами, привлекаемыми к реализации образовательной программы на иных условиях, при проведении промежуточной аттестации.

### 4. Содержание дисциплины (модуля).

#### 4.1. Занятия лекционного типа.

№ п/п	Тематика лекционных занятий / краткое содержание
1	Введение в «большие данные» Рассматриваемые вопросы: -Понятие «большие данные».
2	Введение в «большие данные» Рассматриваемые вопросы: -Классические подходы. Подходы, ориентированные на данные.
3	Введение в «большие данные» Рассматриваемые вопросы: -Принципы работы с большими данными.
4	Технологии, связанные с большими данными Рассматриваемые вопросы: -Уровни в системах обработки данных. Сбор данных. Преобразование данных.
5	Технологии, связанные с большими данными Рассматриваемые вопросы: -Контроль данных. Анализ данных. Представление результатов.
6	Технологии, связанные с большими данными Рассматриваемые вопросы: -Распределённые файловые системы.
7	Технологии, связанные с большими данными Рассматриваемые вопросы: -Файловая система Hadoop
8	Технологии, связанные с большими данными Рассматриваемые вопросы: -Реляционные хранилища данных.
9	Технологии, связанные с большими данными Рассматриваемые вопросы: -NoSQL- хранилища данных. Виды NoSQL. HBase. MongoDB.
10	Прикладные инструменты для работы с данными Рассматриваемые вопросы: -Лямбда-архитектура.
11	Прикладные инструменты для работы с данными Рассматриваемые вопросы: -Каппа-архитектура.
12	Прикладные инструменты для работы с данными Рассматриваемые вопросы: -Инструментарий Map/Reduce.
13	Прикладные инструменты для работы с данными Рассматриваемые вопросы: -Apache Spark
14	Прикладные инструменты для работы с данными Рассматриваемые вопросы: -Машинное обучение

#### 4.2. Занятия семинарского типа.

##### Практические занятия

№ п/п	Тематика практических занятий/краткое содержание
1	Большие данные. На лабораторном занятии студент изучит: -Понятие «большие данные».
2	Большие данные. На лабораторном занятии студент изучит: -Примеры работы с «большими данными»
3	Большие данные. На лабораторном занятии студент изучит: -Принципы работы с данными.
4	Большие данные. На лабораторном занятии студент изучит: - Получение первичных статистик набора данных
5	Большие данные. На лабораторном занятии студент изучит: -Возможности MS Excel
6	Анализ. В результате лабораторного занятия осваиваются: -Элементы кластерного анализа
7	Анализ. В результате лабораторного занятия осваиваются: -SQL запросы.
8	Анализ. В результате лабораторного занятия осваиваются: -NoSQL запросы.
9	Анализ. В результате лабораторного занятия осваиваются: -Визуализация. Примеры визуализации
10	Анализ. В результате лабораторного занятия осваиваются: -Инструменты для визуализации данных
11	Инструменты для работы с данными. В результате лабораторного занятия изучается: -Инструментарий HDMR.
12	Инструменты для работы с данными. В результате лабораторного занятия изучается: -Язык R для анализа статистических данных.
13	Инструменты для работы с данными. В результате лабораторного занятия изучается: -Примеры применения операторов Map на языках Python и R.
14	Инструменты для работы с данными. В результате лабораторного занятия изучается: -Примеры применения операторов Reduce на языках Python и R.

#### 4.3. Самостоятельная работа обучающихся.

№ п/п	Вид самостоятельной работы
1	Подготовка к практическим занятиям

№ п/п	Вид самостоятельной работы
2	Самостоятельное изучение темы «База данных MongoDB» [1, с.49-95]
3	Подготовка к промежуточной аттестации.
4	Подготовка к текущему контролю.

5. Перечень изданий, которые рекомендуется использовать при освоении дисциплины (модуля).

№ п/п	Библиографическое описание	Место доступа
1	Постреляционные хранилища данных : учебное пособие для вузов / Ю. П. Парфенов ; под научной редакцией Н. В. Папуловской. — Москва : Издательство Юрайт, 2022. — 121 с. — ISBN 978-5-534-09837-2.	Юрайт [сайт]. — URL: <a href="https://urait.ru/bcode/492609">https://urait.ru/bcode/492609</a> (дата обращения: 03.10.2022).— Текст : электронный
2	Линейное программирование : учебное пособие для среднего профессионального образования / И. А. Палий. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2024. — 175 с. — (Профессиональное образование). — ISBN 978-5-534-18270-5. — Текст : электронный // Образовательная платформа Юрайт [сайт].	<a href="https://urait.ru/bcode/534655">https://urait.ru/bcode/534655</a> (дата обращения: 21.03.2024).

6. Перечень современных профессиональных баз данных и информационных справочных систем, которые могут использоваться при освоении дисциплины (модуля).

Обязательный набор:

Официальный сайт РУТ (МИИТ) (<https://www.miit.ru/>).

Научно-техническая библиотека РУТ (МИИТ) (<http://library.miit.ru>).

Образовательная платформа «Юрайт» (<https://urait.ru/>).

Общие информационные, справочные и поисковые системы «Консультант Плюс», «Гарант».

Электронно-библиотечная система издательства «Лань» (<http://e.lanbook.com/>).

Научно-техническая библиотека РУТ (МИИТ): <http://library.miit.ru>

Федеральная служба государственной статистики: <https://rosstat.gov.ru/>

КонсультантПлюс: <http://www.consultant.ru/>

Гарант: <http://www.garant.ru/>

7. Перечень лицензионного и свободно распространяемого программного обеспечения, в том числе отечественного производства, необходимого для освоения дисциплины (модуля).

Обязательный набор:

Офисный пакет приложений Microsoft Office;

Python и/или R;

MySQL или другая SQL-база данных;

MongoDB или другая NoSQL база данных.

8. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине (модулю).

Для проведения лекционных занятий необходима аудитория с мультимедиа аппаратурой. Для проведения практических занятий требуется аудитория, оснащенная мультимедиа аппаратурой и ПК с необходимым программным обеспечением и подключением к сети интернет.

9. Форма промежуточной аттестации:

Экзамен в 7 семестре.

10. Оценочные материалы.

Оценочные материалы, применяемые при проведении промежуточной аттестации, разрабатываются в соответствии с локальным нормативным актом РУТ (МИИТ).

Авторы:

доцент, доцент, к.н. кафедры  
«Цифровые технологии управления  
транспортными процессами»

А.С. Милевский

Согласовано:

Заведующий кафедрой ИСЦЭ  
Председатель учебно-методической  
комиссии

Л.А. Каргина

М.В. Ишханян