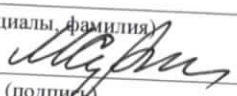


МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ  
 ФЕДЕРАЦИИ  
 федеральное государственное автономное образовательное учреждение высшего  
 образования  
 "САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
 АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ"

Кафедра № 14


УТВЕРЖДАЮ  
 Руководитель направления  
 д.т.н., проф.  
 (должность, уч. степень, звание)  
 М.Б. Сергеев  
 (инициалы, фамилия)  
  
 (подпись)  
 «24» июня 2021 г

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ


«Управление большими данными»  
 (Наименование дисциплины)


Код направления подготовки/ специальности	09.03.01
Наименование направления подготовки/ специальности	Информатика и вычислительная техника
Наименование направленности	Интегрированные автоматизированные информационные системы
Форма обучения	очная

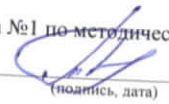
Лист согласования рабочей программы дисциплины

Программу составил (а)  
 проф. д.т.н., проф.  
 (должность, уч. степень, звание)  Ф.А. Таубин  
 (подпись, дата) (инициалы, фамилия)

Программа одобрена на заседании кафедры № 14  
 «15» июня 2021 г, протокол № 11

Заведующий кафедрой № 14  
 к.т.н., доц.  
 (уч. степень, звание)  В.Л. Оленев  
 (подпись, дата) (инициалы, фамилия)

Ответственный за ОП ВО 09.03.01(03)  
 доц., к.т.н., доц.  
 (должность, уч. степень, звание)  А.В. Шахомиров  
 (подпись, дата) (инициалы, фамилия)

Заместитель директора института №1 по методической работе  
 ст. преподаватель  
 (должность, уч. степень, звание)  В.Е. Таратун  
 (подпись, дата) (инициалы, фамилия)

## Аннотация

Дисциплина «Управление большими данными» входит в образовательную программу высшего образования по направлению подготовки/ специальности 09.03.01 «Информатика и вычислительная техника» направленности «Интегрированные автоматизированные информационные системы». Дисциплина реализуется кафедрой «№14».

Дисциплина нацелена на формирование у выпускника следующих компетенций:

ПК-2 «Способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности»

ПК-4 «Способен разрабатывать технические спецификации на программные компоненты и их взаимодействие, способен проектировать программное обеспечение»

Содержание дисциплины охватывает круг вопросов, связанных с управлением большими данными.

Преподавание дисциплины предусматривает следующие формы организации учебного процесса: лекции, лабораторные работы, самостоятельная работа студента, консультации.

Программой дисциплины предусмотрены следующие виды контроля: текущий контроль успеваемости, промежуточная аттестация в форме экзамена.

Общая трудоемкость освоения дисциплины составляет 3 зачетных единицы, 108 часов.

Язык обучения по дисциплине «русский»

## 1. Перечень планируемых результатов обучения по дисциплине

### 1.1. Цели преподавания дисциплины

Цель – изучение и приобретение студентами навыков управления большими данными.

1.2. Дисциплина входит в состав части, формируемой участниками образовательных отношений, образовательной программы высшего образования (далее – ОП ВО).

1.3. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения ОП ВО.

В результате изучения дисциплины обучающийся должен обладать следующими компетенциями или их частями. Компетенции и индикаторы их достижения приведены в таблице 1.

Таблица 1 – Перечень компетенций и индикаторов их достижения

Категория (группа) компетенции	Код и наименование компетенции	Код и наименование индикатора достижения компетенции
Профессиональные компетенции	ПК-2 Способен осуществлять концептуальное, функциональное и логическое проектирование систем среднего и крупного масштаба и сложности	ПК-2.3.1 знать методы концептуального, функционального и логического проектирования, принципы разработки технико-экономических характеристик вариантов концептуальной архитектуры ПК-2.У.1 уметь разрабатывать технико-экономическое обоснование, определять ключевые свойства системы, определять ограничения системы, варианты концептуальной архитектуры системы ПК-2.В.1 владеть навыками определения ключевых свойств и ограничений системы, навыками определения вариантов концептуальной архитектуры системы, навыками описания технико-экономического обоснования
Профессиональные компетенции	ПК-4 Способен разрабатывать технические спецификации на программные компоненты и их взаимодействие, способен проектировать программное обеспечение	ПК-4.3.1 знать типовые решения, библиотеки программных модулей, шаблоны, классы объектов, используемые при разработке программного обеспечения; знать методы и средства проектирования программного обеспечения, баз данных и программных интерфейсов ПК-4.У.1 уметь проводить оценку и обоснование рекомендуемых решений; вырабатывать варианты реализации программного обеспечения; применять методы и средства проектирования программного обеспечения, баз данных и программных интерфейсов ПК-4.В.1 владеть технологией программирования, методами и средствами проектирования программного обеспечения, баз данных и программных интерфейсов

## 2. Место дисциплины в структуре ОП

Дисциплина может базироваться на знаниях, ранее приобретенных обучающимися при изучении следующих дисциплин:

- Базы данных

Знания, полученные при изучении материала данной дисциплины, имеют как самостоятельное значение, так и могут использоваться при изучении других дисциплин:

- Выпускная квалификационная работа

### 3. Объем и трудоемкость дисциплины

Данные об общем объеме дисциплины, трудоемкости отдельных видов учебной работы по дисциплине (и распределение этой трудоемкости по семестрам) представлены в таблице 2.

Таблица 2 – Объем и трудоемкость дисциплины

Вид учебной работы	Всего	Трудоемкость по семестрам
		№8
1	2	3
<b>Общая трудоемкость дисциплины, ЗЕ/ (час)</b>	3/ 108	3/ 108
<b>Из них часов практической подготовки</b>	20	20
<b>Аудиторные занятия, всего час.</b>	40	40
в том числе:		
лекции (Л), (час)	20	20
практические/семинарские занятия (ПЗ), (час)		
лабораторные работы (ЛР), (час)	20	20
курсовой проект (работа) (КП, КР), (час)		
экзамен, (час)	63	63
<b>Самостоятельная работа, всего (час)</b>	5	5
<b>Вид промежуточной аттестации:</b> зачет, дифф. зачет, экзамен (Зачет, Дифф. зач, Экз.**)	Экз.	Экз.

Примечание: \*\* кандидатский экзамен

### 4. Содержание дисциплины

4.1. Распределение трудоемкости дисциплины по разделам и видам занятий.

Разделы, темы дисциплины и их трудоемкость приведены в таблице 3.

Таблица 3 – Разделы, темы дисциплины, их трудоемкость

Разделы, темы дисциплины	Лекции (час)	ПЗ (СЗ) (час)	ЛР (час)	КП (час)	СРС (час)
Раздел 1. Введение.	5		4		1
Раздел 2. Методы многомерного статистического анализа и анализа нечисловой информации.	5		4		1
Раздел 3. Программные средства для управления большими данными.	5		8		2
Раздел 4. Системы Data Mining	5		4		1
Итого в семестре:	20		20		5
Итого	20	0	20	0	5

4.2. Содержание разделов и тем лекционных занятий.

Содержание разделов и тем лекционных занятий приведено в таблице 4.

Таблица 4 – Содержание разделов и тем лекционного цикла

Номер раздела	Название и содержание разделов и тем лекционных занятий
1	Раздел 1. Введение. Тема 1.1. Простейшие статистические характеристики Тема 1.2. Приведение к нормальной форме Тема 1.3. Оцифровка нечисловых данных
2	Раздел 2. Методы многомерного статистического анализа и анализа нечисловой информации. Тема 2.1. Роль и сущность многомерных статистических методов Тема 2.2 Многомерное нормальное распределение как основная модель современных многомерных статистических методов. Тема 2.3. Практическое применение многомерных методов Тема 2.4. Методы статистического оценивания многомерных параметров и проверки гипотез. Тема 2.5. Особенности анализа количественных и качественных признаков. Тема 2.6. Кластерный анализ. Тема 2.7. Компонентный анализ.
3	Раздел 3. Программные средства для управления большими данными. Тема 3.1. NoSQL СУБД . Тема 3.2. Технология MapReduce. Тема 3.3. Распределенные файловые системы. Тема 3.4 Google Bigtable.
4	Раздел 4. Системы Data Mining Тема 4.1. Понятие о технологии Data Mining. Тема 4.2. Реализация в пакетах прикладных программ. Тема 4.3. Сетевые технологии Data Mining.

4.3. Практические (семинарские) занятия

Темы практических занятий и их трудоемкость приведены в таблице 5.

Таблица 5 – Практические занятия и их трудоемкость

№ п/п	Темы практических занятий	Формы практических занятий	Трудоемкость, (час)	№ раздела дисциплины
Учебным планом не предусмотрено				

4.4. Лабораторные занятия

Темы лабораторных занятий и их трудоемкость приведены в таблице 6.

Таблица 6 – Лабораторные занятия и их трудоемкость

№ п/п	Наименование лабораторных работ	Трудоемкость, (час)	№ раздела дисциплины
Семестр 8			
1	Приведение к нормальной форме	4	1
2	Кластерный анализ	4	2

3	Проектирование реляционной базы данных	4	3
4	Реализация спроектированной базы данных в СУБД	4	3
5	Реализация технологии Data Mining в пакетах прикладных программ	4	4
Всего		20	

4.5. Курсовое проектирование/ выполнение курсовой работы  
Учебным планом не предусмотрено

4.6. Самостоятельная работа обучающихся  
Виды самостоятельной работы и ее трудоемкость приведены в таблице 7.

Таблица 7 – Виды самостоятельной работы и ее трудоемкость

Вид самостоятельной работы	Всего, час	Семестр 8, час
1	2	3
Изучение теоретического материала дисциплины (ТО)	2	2
Курсовое проектирование (КП, КР)		
Расчетно-графические задания (РГЗ)		
Выполнение реферата (Р)		
Подготовка к текущему контролю успеваемости (ТКУ)	2	2
Домашнее задание (ДЗ)		
Контрольные работы заочников (КРЗ)		
Подготовка к промежуточной аттестации (ПА)	1	1
Всего:	5	5

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

Учебно-методические материалы для самостоятельной работы обучающихся указаны в п.п. 7-11.

Шифр/ URL адрес	Библиографическая ссылка	Количество экземпляров в библиотеке (кроме электронных экземпляров)
	Tom White. Hadoop: The Definitive Guide, 3rd Edition. O'Reilly Media, 2013. 630p. ISBN 978-1-4493-1152-0	
	E. Capriolo, D. Wampler, J. Rutherglen. Programming Hive: Data Warehouse and Qiuery Language for Hadoop. O'reilly Media, 2012 – 352p. ISBN 978-1-4493-1933-5.	
	Дейт К. Дж. Введение в с системы баз данных. – М.: Вильямс, 2005 – 1328с.	

	Trevor Hastie, Elements of statistical learning, Springer, 2009. <a href="http://statweb.stanford.edu/~tibs/ElemStatLe_arn/">http://statweb.stanford.edu/~tibs/ElemStatLe_arn/</a>	
	Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data // EMC Education Services. 2015. - 432p. - ISBN: 978-1-118-87613-8.	
	C. B. B. D. Manyika, “Big Data: The Next Frontier for Innovation, Competition, and Productivity,” McKinsey Global Institute, 2011. URL: <a href="http://www.mckinsey.com/~media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx">http://www.mckinsey.com/~media/McKinsey/dotcom/Insights%20and%20pubs/MGI/Research/Technology%20and%20Innovation/Big%20Data/MGI_big_data_full_report.ashx</a>	

6. Перечень печатных и электронных учебных изданий

Перечень печатных и электронных учебных изданий приведен в таблице 8.

Таблица 8– Перечень печатных и электронных учебных изданий

7. Перечень электронных образовательных ресурсов  
информационно-телекоммуникационной сети «Интернет»

Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины приведен в таблице 9.

Таблица 9 – Перечень электронных образовательных ресурсов информационно-телекоммуникационной сети «Интернет»

URL адрес	Наименование
<a href="http://e.lanbook.com/books/">e.lanbook.com/books/</a>	Электронно-библиотечная система
<a href="http://www.gpntb.ru/">www.gpntb.ru/</a>	Государственная публичная научно-техническая библиотека.
<a href="http://www.nlr.ru/">www.nlr.ru/</a>	Российская национальная библиотека.
<a href="http://www.nns.ru/">www.nns.ru/</a>	Национальная электронная библиотека.
<a href="http://www.rsl.ru/">www.rsl.ru/</a>	Российская государственная библиотека
<a href="http://www.microinform.ru/">www.microinform.ru/</a>	Учебный центр компьютерных технологий «Микроинформ»
<a href="http://www.intuit.ru/">www.intuit.ru/</a>	Образовательный сайт
<a href="http://www.window.edu.ru/">www.window.edu.ru/</a>	Библиотека учебной и методической литературы
<a href="http://www.osp.ru/">www.osp.ru/</a>	Журнал «Открытые системы»
<a href="http://www.ihitika.lib.ru/">www.ihitika.lib.ru/</a>	Библиотека учебной и методической литературы

8. Перечень информационных технологий

8.1. Перечень программного обеспечения, используемого при осуществлении образовательного процесса по дисциплине.

Перечень используемого программного обеспечения представлен в таблице 10.

Таблица 10– Перечень программного обеспечения

№ п/п	Наименование
	Не предусмотрено

8.2. Перечень информационно-справочных систем,используемых при осуществлении образовательного процесса по дисциплине

Перечень используемых информационно-справочных систем представлен в таблице 11.

Таблица 11– Перечень информационно-справочных систем

№ п/п	Наименование
	Не предусмотрено

#### 9. Материально-техническая база

Состав материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине, представлен в таблице 12.

Таблица 12 – Состав материально-технической базы

№ п/п	Наименование составной части материально-технической базы	Номер аудитории (при необходимости)
1	Лекционная аудитория	
2	Компьютерный класс	

#### 10. Оценочные средства для проведения промежуточной аттестации

10.1. Состав оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине приведен в таблице 13.

Таблица 13 – Состав оценочных средств для проведения промежуточной аттестации

Вид промежуточной аттестации	Перечень оценочных средств
Экзамен	Список вопросов к экзамену

10.2. В качестве критериев оценки уровня сформированности (освоения) компетенций обучающимися применяется 5-балльная шкала оценки сформированности компетенций, которая приведена в таблице 14. В течение семестра может использоваться 100-балльная шкала модульно-рейтинговой системы Университета, правила использования которой, установлены соответствующим локальным нормативным актом ГУАП.

Таблица 14 – Критерии оценки уровня сформированности компетенций

Оценка компетенции 5-балльная шкала	Характеристика сформированных компетенций
«отлично» «зачтено»	<ul style="list-style-type: none"> <li>– обучающийся глубоко и всесторонне усвоил программный материал;</li> <li>– уверенно, логично, последовательно и грамотно его излагает;</li> <li>– опираясь на знания основной и дополнительной литературы, тесно привязывает усвоенные научные положения с практической деятельностью направления;</li> <li>– умело обосновывает и аргументирует выдвигаемые им идеи;</li> <li>– делает выводы и обобщения;</li> <li>– свободно владеет системой специализированных понятий.</li> </ul>
«хорошо» «зачтено»	<ul style="list-style-type: none"> <li>– обучающийся твердо усвоил программный материал, грамотно и по существу излагает его, опираясь на знания основной литературы;</li> <li>– не допускает существенных неточностей;</li> <li>– увязывает усвоенные знания с практической деятельностью направления;</li> <li>– аргументирует научные положения;</li> <li>– делает выводы и обобщения;</li> <li>– владеет системой специализированных понятий.</li> </ul>

Оценка компетенции 5-балльная шкала	Характеристика сформированных компетенций
«удовлетворительно» «зачтено»	<ul style="list-style-type: none"> <li>– обучающийся усвоил только основной программный материал, по существу излагает его, опираясь на знания только основной литературы;</li> <li>– допускает несущественные ошибки и неточности;</li> <li>– испытывает затруднения в практическом применении знаний направления;</li> <li>– слабо аргументирует научные положения;</li> <li>– затрудняется в формулировании выводов и обобщений;</li> <li>– частично владеет системой специализированных понятий.</li> </ul>
«неудовлетворительно» «не зачтено»	<ul style="list-style-type: none"> <li>– обучающийся не усвоил значительной части программного материала;</li> <li>– допускает существенные ошибки и неточности при рассмотрении проблем в конкретном направлении;</li> <li>– испытывает трудности в практическом применении знаний;</li> <li>– не может аргументировать научные положения;</li> <li>– не формулирует выводов и обобщений.</li> </ul>

#### 10.3. Типовые контрольные задания или иные материалы.

Вопросы (задачи) для экзамена представлены в таблице 15.

Таблица 15 – Вопросы (задачи) для экзамена

№ п/п	Перечень вопросов (задач) для экзамена
1	Простейшие статистические характеристики, приведение к нормальной форме и оцифровка нечисловых данных
2	Роль и сущность многомерных статистических методов, многомерное нормальное распределение как основная модель современных многомерных статистических методов
3	Практическое применение многомерных методов
4	Методы статистического оценивания многомерных параметров и проверки гипотез, особенности анализа количественных и качественных признаков.
5	Кластерный анализ.
6	Компонентный анализ.
7	NoSQL СУБД.
8	Технология MapReduce.
9	Распределенные файловые системы, Google Bigtable.
10	Понятие о технологии Data Mining.
11	Реализация в пакетах прикладных программ.
12	Сетевые технологии Data Mining.

Вопросы (задачи) для зачета / дифф. зачета представлены в таблице 16.

Таблица 16 – Вопросы (задачи) для зачета / дифф. зачета

№ п/п	Перечень вопросов (задач) для зачета / дифф. зачета
	Учебным планом не предусмотрено

Перечень тем для курсового проектирования/выполнения курсовой работы представлены в таблице 17.

Таблица 17 – Перечень тем для курсового проектирования/выполнения курсовой работы

№ п/п	Примерный перечень тем для курсового проектирования/выполнения курсовой работы
	Учебным планом не предусмотрено

Вопросы для проведения промежуточной аттестации в виде тестирования представлены в таблице 18.

Таблица 18 – Примерный перечень вопросов для тестов

№ п/п	Примерный перечень вопросов для тестов
	Учебным планом не предусмотрено

Перечень тем контрольных работ по дисциплине обучающихся заочной формы обучения, представлены в таблице 19.

Таблица 19 – Перечень контрольных работ

№ п/п	Перечень контрольных работ
	Не предусмотрено

10.4. Методические материалы, определяющие процедуры оценивания индикаторов, характеризующих этапы формирования компетенций, содержатся в локальных нормативных актах ГУАП, регламентирующих порядок и процедуру проведения текущего контроля успеваемости и промежуточной аттестации обучающихся ГУАП.

## 11. Методические указания для обучающихся по освоению дисциплины

11.1. Методические указания для обучающихся по освоению лекционного материала

Основное назначение лекционного материала – логически стройное, системное, глубокое и ясное изложение учебного материала. Назначение современной лекции в рамках дисциплины не в том, чтобы получить всю информацию по теме, а в освоении фундаментальных проблем дисциплины, методов научного познания, новейших достижений научной мысли. В учебном процессе лекция выполняет методологическую, организационную и информационную функции. Лекция раскрывает понятийный аппарат конкретной области знания, её проблемы, дает цельное представление о дисциплине, показывает взаимосвязь с другими дисциплинами.

Планируемые результаты при освоении обучающимися лекционного материала:

- получение современных, целостных, взаимосвязанных знаний, уровень которых определяется целевой установкой к каждой конкретной теме;
- получение опыта творческой работы совместно с преподавателем;
- развитие профессионально-деловых качеств, любви к предмету и самостоятельного творческого мышления.
- появление необходимого интереса, необходимого для самостоятельной работы;
- получение знаний о современном уровне развития науки и техники и о прогнозе их развития на ближайшие годы;
- научиться методически обрабатывать материал (выделять главные мысли и положения, приходиться к конкретным выводам, повторять их в различных формулировках);
- получение точного понимания всех необходимых терминов и понятий.

Лекционный материал может сопровождаться демонстрацией слайдов и использованием раздаточного материала при проведении коротких дискуссий об особенностях применения отдельных тематик по дисциплине.

11.2. Методические указания для обучающихся по выполнению лабораторных работ

В ходе выполнения лабораторных работ обучающийся должен углубить и закрепить знания, практические навыки, овладеть современной методикой и техникой эксперимента в соответствии с квалификационной характеристикой обучающегося. Выполнение лабораторных работ состоит из экспериментально-практической, расчетно-аналитической частей и контрольных мероприятий.

Выполнение лабораторных работ обучающимся является неотъемлемой частью изучения дисциплины, определяемой учебным планом, и относится к средствам, обеспечивающим решение следующих основных задач обучающегося:

- приобретение навыков исследования процессов, явлений и объектов, изучаемых в рамках данной дисциплины;
- закрепление, развитие и детализация теоретических знаний, полученных на лекциях;
- получение новой информации по изучаемой дисциплине;
- приобретение навыков самостоятельной работы с лабораторным оборудованием и приборами.

### Структура и форма отчета о лабораторной работе

- Титульный лист.
- Содержание.
- Задание. Номер варианта.
- Постановка задачи.
- Теоретические выкладки.
- Расчеты.
- Листинг программы (если требуется).
- Выводы.
- Список литературы.

### Требования к оформлению отчета о лабораторной работе

- Изложение текста и оформление работ следует выполнять в соответствии с требованиями ГОСТ 7.32 – 2001.
- [http://guap.ru/guap/standart/prav\\_main.shtml](http://guap.ru/guap/standart/prav_main.shtml)

11.3. Методические указания для обучающихся по прохождению самостоятельной работы

В ходе выполнения самостоятельной работы, обучающийся выполняет работу по заданию и при методическом руководстве преподавателя, но без его непосредственного участия.

Для обучающихся по заочной форме обучения, самостоятельная работа может включать в себя контрольную работу.

В процессе выполнения самостоятельной работы, у обучающегося формируется целесообразное планирование рабочего времени, которое позволяет им развивать умения и навыки в усвоении и систематизации приобретаемых знаний, обеспечивает высокий уровень успеваемости в период обучения, помогает получить навыки повышения профессионального уровня.

Методическими материалами, направляющими самостоятельную работу обучающихся являются:

- учебно-методический материал по дисциплине;
- методические указания по выполнению контрольных работ (для обучающихся по заочной форме обучения).

11.4. Методические указания для обучающихся по прохождению текущего контроля успеваемости.

Текущий контроль успеваемости предусматривает контроль качества знаний обучающихся, осуществляемого в течение семестра с целью оценивания хода освоения дисциплины.

11.5. Методические указания для обучающихся по прохождению промежуточной аттестации.

Промежуточная аттестация обучающихся предусматривает оценивание промежуточных и окончательных результатов обучения по дисциплине. Она включает в себя:

– экзамен – форма оценки знаний, полученных обучающимся в процессе изучения всей дисциплины или ее части, навыков самостоятельной работы, способности применять их для решения практических задач. Экзамен, как правило, проводится в период экзаменационной сессии и завершается аттестационной оценкой «отлично», «хорошо», «удовлетворительно», «неудовлетворительно».

Лист внесения изменений в рабочую программу дисциплины

Дата внесения изменений и дополнений. Подпись внесшего изменения	Содержание изменений и дополнений	Дата и № протокола заседания кафедры	Подпись зав. кафедрой