





МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
**«Дальневосточный федеральный университет»**  
(ДВФУ)

## **ШКОЛА ЭКОНОМИКИ И МЕНЕДЖМЕНТА**

СОГЛАСОВАНО  
Руководитель ОП  
  
(подпись) Сащенко А.Ю.  
(ФИО)

УТВЕРЖДАЮ  
Директор департамента прикладной экономики  
  
(подпись) Мосолова Н.А.  
(И.О. Фамилия)  
«24» ноября 2022 г.

### **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

*Наука о данных и аналитика больших объемов информации*

*Специальность 38.05.01 Экономическая безопасность*

*Специализация: Экономико-правовое обеспечение экономической безопасности*

*Форма обучения: очная*

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта по специальности 38.05.01 Экономическая безопасность, утвержденного приказом Минобрнауки России от 14 апреля 2021 г. № 293.

Директор Департамента прикладной экономики, канд. экон. наук, доцент

Н.А. Мосолова

Составитель: канд. филос. наук, доцент

А.В. Луценко

Владивосток

2022

## Оборотная сторона титульного листа РПД

1. Рабочая программа пересмотрена на заседании Департамента/кафедры/отделения (реализующего дисциплину) и утверждена на заседании Департамента/кафедры/отделения (выпускающего структурного подразделения), протокол от «\_»\_20\_г. № \_\_

2. Рабочая программа пересмотрена на заседании Департамента/кафедры/отделения (реализующего дисциплину) и утверждена на заседании Департамента/кафедры/отделения (выпускающего структурного подразделения), протокол от «\_»\_20\_г. № \_\_

3. Рабочая программа пересмотрена на заседании Департамента/кафедры/отделения (реализующего дисциплину) и утверждена на заседании Департамента/кафедры/отделения (выпускающего структурного подразделения), протокол от «\_»\_20\_г. № \_\_

4. Рабочая программа пересмотрена на заседании Департамента/кафедры/отделения (реализующего дисциплину) и утверждена на заседании Департамента/кафедры/отделения (выпускающего структурного подразделения), протокол от «\_»\_20\_г. № \_\_

5. Рабочая программа пересмотрена на заседании Департамента/кафедры/отделения (реализующего дисциплину) и утверждена на заседании Департамента/кафедры/отделения (выпускающего структурного подразделения), протокол от «\_»\_20\_г. № \_\_

# 1. Цели и планируемые результаты изучения дисциплины

## Цели освоения дисциплины

Цель изучения дисциплины «Наука о данных и аналитика больших объемов данных» направлена на формирование у обучающихся пониманий и знаний теоретических и практических аспектов технологий, в основе которых лежат принципы получения, преобразования, обработки, хранения и анализа больших объемов данных. Также целью изучения дисциплины является подготовка квалифицированных выпускников, умеющих обосновано и результативно использовать, совершенствовать, разрабатывать и внедрять современные технологии и инструментальные средства анализа и работы с большими объемами данных.

*Язык реализации: русский*

Задачи дисциплины:

- научить студентов рассчитывать экономические показатели, характеризующие деятельность хозяйствующих субъектов, формировать экономические разделы отчетов, планов, программ развития субъектов экономики на микро- и макроуровне;

- дать представление об анализе финансовой, бухгалтерской и иной информации, содержащейся в отчетности предприятий различных форм собственности, организаций, ведомств и т.д.;

- дать представление обучающимся о значении социально-экономического развития на рост территориального развития, определить показатели, которые характеризуют процессы развития территорий;

- по окончании курса обучающиеся знают принципы и методы экономико-математического моделирования.

## 2. Место дисциплины в структуре ООП

В учебном плане дисциплина Б1.В.01 «Наука о данных и аналитика больших объемов информации» относится к вариативным дисциплинам, к части, которая формируется участниками образовательных отношений.

Дисциплина изучается на 1 курсе, во 2-ом семестре и завершается сдачей зачета. Общий объем дисциплины 108 часов (3 з.ед.), из них: лекции – 8 часов, практические занятия – не предусмотрены; контрольные работы – 40 часов; текущий контроль – 16 часов; самостоятельная работа 16 часов, промежуточная аттестация – 28 часов.

Изучение дисциплины базируется на знаниях школьной программы, дисциплины «Экономическое мышление», «Основы экономической грамотности» и т.д.

По результатам изучения дисциплины студенты должны освоить следующие компетенции:

Код и наименование компетенций	Код и наименование индикатора достижения компетенций, показатели оценивания	
<p>ПК-1</p> <p>Способен рассчитывать экономические показатели, характеризующие деятельность хозяйствующих субъектов, формировать экономические разделы отчетов, планов, программ развития субъектов экономики на микро- и макроуровне</p>	<p>ПК-1.1</p> <p>Собирает, рассчитывает, анализирует данные, необходимые для формирования экономических разделов отчетов, планов, программ на основе типовых методик и действующей нормативно-правовой базы</p>	<p>Знает состав и структуру экономических отчетов, программ с учетом действующей нормативно-правовой базы</p>
		<p>Умеет собрать и рассчитать данные, необходимые для формирования экономических разделов отчетов, планов, программ на основе типовых методик и действующей нормативно-правовой базы</p>
		<p>Владеет навыками анализа данных, необходимых для формирования экономических разделов отчетов, планов, программ на основе типовых методик и действующей нормативно-правовой базы</p>
	<p>ПК-1.2</p> <p>Анализирует и интерпретирует финансовую, бухгалтерскую и иную информацию, содержащуюся в отчетности предприятий различных форм собственности, организаций, ведомств и т.д.</p>	<p>Знает формы собственности организаций, знает особенности финансовой и бухгалтерской отчетности для разных форм собственности</p>
		<p>Умеет анализировать и интерпретировать финансовую, бухгалтерскую и иную информацию, содержащуюся в отчетности предприятий различных форм собственности, организаций, ведомств и т.д.</p>
	<p>ПК-1.3</p> <p>Анализирует и интерпретирует показатели социально-экономического развития территорий</p>	<p>Знает показатели социально-экономического развития территорий</p>
	<p>Умеет анализировать и интерпретировать показатели социально-экономического развития территорий</p>	
<p>ПК-3</p> <p>Способен решать типовые профессиональные задачи с помощью правил формального анализа, математических приемов, инструментальных методов, информационных технологий и программных средств</p>	<p>ПК-3.4</p> <p>Использует теоретические концепции и инструментальные методы экономико-математического моделирования</p>	<p>Знает теоретические концепции и инструментальные методы экономико-математического моделирования</p>
		<p>Умеет применять на практике теоретические концепции и инструментальные методы экономико-математического моделирования</p>
		<p>Владеет навыками экономико-математического моделирования социально-экономических процессов</p>

#### 4.1 Разделы дисциплины и виды учебной работы

№ раздела	Разделы дисциплины, мероприятия текущего контроля	Очная форма		
		Лек, ач	Пр, ач	СР, ач
1.	Введение в большие данные			
1.1.	Определение больших данных и причины их появления	1	0	0
1.2.	Примеры возможностей для бизнеса	0	0	1
1.3.	Различие между Business Intelligence и Big Data	1	0	0
2.	Жизненный цикл аналитики данных			
2.1.	Понятие жизненного цикла аналитики данных	1	0	1
2.2.	Роли, необходимые для успешного создания проекта по аналитике данных	1	0	0
2.3.	Решение задач по организации жизненного цикла аналитики данных	0	1	0
3.	Сложные методы аналитики			
3.1.	Классификация задач анализа: Text, Data, Web, Social Mining	1	0	0
3.2.	Применение машинного обучения в аналитике	1	0	0

№ раздела	Разделы дисциплины, мероприятия текущего контроля	Очная форма		
		Лек, ач	Пр, ач	СР, ач
3.3.	К-mean и C-means кластеризация, классификация	0	1	1
3.4.	Логистическая регрессия, Ассоциации, алгоритм Априори	0	1	1
3.5.	Контрольная работа 1	0	1	0
4.	Высокопроизводительные вычисления			
4.1.	Распределенные вычисления на нескольких серверах	1	2	0
4.2.	Вычислительная парадигма MapReduce	0	0	1
4.3.	Проект Apache Hadoop и его экосистема	1	0	0
4.4.	Apache Spark и его компоненты	1	0	0
4.5.	Вычисления в реальном времени и Apache Storm	0	0	1
5.	Масштабирование и многоуровневое хранение данных			
5.1.	Теорема CAP	1	0	1
5.2.	Парадигма NoSQL	1	0	1
5.3.	Классификация NoSQL баз данных	1	0	0
5.4.	Дедупликация данных в хранилище	1	0	0
6.	Анализ текста			
6.1.	Стратегии анализа текста	1	0	0
6.2.	Поисковые механизмы: Lucene, Solr, ElasticSearch	0	2	0
6.3.	Алгоритмы Word2Vec и Glove	0	0	1
7.	Визуализация данных и результатов анализа			
7.1.	Техники визуализации данных	1	0	0
7.2.	Введение в язык R	0	1	0
7.3.	Визуализация данных в R	0	0	1
8.	Обзор пройденного материала			
8.1.	Ключевые аспекты успешного проекта в области аналитики данных	1	0	1
8.2.	Защита курсового проекта	0	4	2
8.3.	Зачет	0	2	5

№ раздела	Разделы дисциплины, мероприятия текущего контроля	Очная форма		
		Лек, ач	Пр, ач	СР, ач
8.4.	Экзамен	0	1	4
<b>Итого по видам учебной работы:</b>		16	16	22
Экзамены, ач				54
Часы на контроль, ач				54
<b>Общая трудоёмкость освоения: ач / зет</b>		108 / 3		

#### 4.2. Содержание разделов и результаты изучения дисциплины

Раздел дисциплины	Содержание
<b>1. Введение в большие данные</b>	
<b>1.1. Определение больших данных и причины их появления</b>	Предпосылки появления больших данных. Устаревание информации. Рост объемов данных по причине продвижения цифровой техники и вытеснения аналоговых средств хранения и передачи информации. Технические признаки, характеризующие большие данные. Принцип трех V - Volume (объём данных), Variety (разнообразие данных), Velocity (скорость генерации и работы с данными). Информационные потоки и их обработка
<b>1.2. Примеры возможностей для бизнеса</b>	Примеры возможностей для бизнеса, открывающиеся благодаря появлению больших объемов данных и проведению аналитических действий над ними. Предметные области, требующие применение подходов аналитики данных.
<b>1.3. Различие между Business Intelligence и Big Data</b>	Определение «бизнес-аналитики» (Business Intelligence). Недостатки традиционного подхода к анализу и обработке больших объемов данных. Различие между Business Intelligence и Big Data. Разно-структурированные данные. Их представление в моделях данных: структурированные, слабоструктурированные, неструктурированные. Понятие профессии аналитика данных.
<b>2. Жизненный цикл аналитики данных</b>	

<p><b>2.1. Понятие жизненного цикла аналитики данных</b></p>	<p>Необходимость в аналитической работе с большими данными. Процедура анализа текстов. Этапы жизненного цикла аналитики данных. Этап исследования, формулировки целей и задач, выбор эмпирических данных, выдвижение гипотез. Манипуляционный этап, на котором осуществляется определение категорий и подкатегорий данных и моделей, выбор единиц анализа, установление правил кодирования. Этап вычисления. Этап интерпретации результатов. Презентация результатов. Ошибки при проведении анализа</p>
<p><b>2.2. Роли, необходимые для успешного создания проекта по аналитике данных</b></p>	<p>Определение круга проблем для анализа предметной области и роли участников анализа.</p>
<p><b>2.3. Решение задач по организации жизненного цикла аналитики данных</b></p>	<p>Явная (выраженная) и скрытая (структурная) информация. Количественная и качественная стратегия анализа данных. Возможности и ограничения существующих подходов. Трудоемкость этапов жизненного цикла. Решение задач на каждом из этапов жизненного цикла. ETL-процесс.</p>
<p><b>3. Сложные методы аналитики</b></p>	
<p><b>3.1. Классификация задач анализа: Text, Data, Web, Social Mining</b></p>	<p>Рассмотрение традиционной классификации задач анализа данных. Примеры бизнес вариантов использования каждой из задач. Определение подходов: Быстрые Данные (Fast Data), Большая Аналитика (Big Analytics) и Глубокое Проникновение (Deep Insight). Обработка быстрых данных, подтверждение и корректировка априорных знаний и гипотез, синхронизация скорости работы с ростом объема данных. Принцип «обучения с учителем». Современные методы аналитики.</p>
<p><b>3.2. Применение машинного обучения в аналитике</b></p>	<p>Необходимость применения машинного обучения в аналитике. Варианты машинного обучения: машинное обучение с учителем (Supervised Learning), машинное обучение без учителя (Unsupervised Learning). Обзор наиболее популярных методов анализа.</p>



<b>3.3. K-mean и C-means кластеризация, классификация</b>	<p>Сценарии использования, алгоритм, определение оптимального значения для K, диагностика с целью оценки эффективности метода. Задача кластеризации, методы кластеризации, иерархическая кластеризация, алгоритм k-средних, зонтичная кластеризация, методы ненаправленного обучения. Постановка задачи классификации, подходы и применения, построение и обучение классификатора, оценка качества классификации, рубрикации тренировочных данных (Training Data Set), Базовые алгоритмы кластеризации. Адаптивная кластеризация.</p> <p>OLAP 3D куб.</p>
<b>3.4. Логистическая регрессия, Ассоциации, алгоритм Априори</b>	<p>Определение и описание логистической регрессии. Подбор параметров. Регуляризация. Сценарии использования логистической регрессии. Диагностика для проверки корректности модели логистической регрессии. Дерево принятия решений. Энтропия и получение информации.</p> <p>Правила ассоциации. Алгоритм априори. Сценарии использования правил ассоциации.</p>
<b>3.5. Контрольная работа 1</b>	Контрольная работа по разделам 1,2,3
<b>4. Высокопроизводительные вычисления</b>	
<b>4.1. Распределенные вычисления на нескольких серверах</b>	<p>Понятие распределенных вычислений (Grid Computing). История появления. Типы распределенных систем (грид-системы). Сравнение грид-систем и суперкомпьютеров. Интеграция, миграция и построение хранилищ данных. Высокопроизводительные вычисления (High Performance Computing) при выполнении аналитических исследований. Основные идеи и подходы параллельных архитектур оперирования разно-структурированными данными.</p>
<b>4.2. Вычислительная парадигма MapReduce</b>	<p>Парадигма Map-Reduce, к каким задачам эту технологию можно применять и при каких условиях она будет эффективной. Шаги Map и Reduce. Map-Reduce на примере. Вычисления наборов распределенных задач с использованием большого количества компьютеров, образующих кластер. Предварительная обработка входных данных и свёртка данных. Концепция параллелизма. Реализация алгоритмов в среде Map Reduce: матрично-векторное умножение, поддержка SQL-подобных операций и операций реляционной алгебры. Теория сложности алгоритмов для map-reduce. Создание собственной программы для map-reduce.</p>

<p><b>4.3. Проект Apache Hadoop и его экосистема</b></p>	<p>Введение в Hadoop. Экосистема Hadoop. Архитектура Hadoop. Интеграция и взаимодействие с Map Reduce. Архитектура стека компонентов Hadoop.</p> <p>Администрирование Hadoop. Запуск приложений (Jobs), мониторинг приложений, Импорт/экспорт больших данных в/из Hadoop.</p> <p>Распределенные файловые системы. Основные аспекты функционирования распределённых файловых систем. История развития. Разработка и выполнение распределённых программ, расширение вычислительных мощностей посредством добавления в кластер дополнительных узлов. Распределённые файловые системы на примере Hadoop Distributed File System (HDFS). Архитектура GFS, Apache HDFS и алгоритмы. Интеграция с NoSQL и MapReduce. Работа с HDFS, запуск простейших заданий.</p>
<p><b>4.4. Apache Spark и его компоненты</b></p>	<p>Сравнение вычислительных мощностей Spark и Hadoop. Состав проекта Spark: Core, SQL, Streaming, MLlib, GraphX.</p> <p>Механизм хранения данных в Spark. Основы RDD, операции и действия над RDD, отложенные вычисления. Загрузка и хранение данных: HDFS, AWS, Cassandra.</p> <p>Выполнение на кластере. Архитектура Spark времени выполнения. Упаковка программного кода и зависимостей. Диспетчер кластеров.</p>
<p><b>4.5. Вычисления в реальном времени и Apache Storm</b></p>	<p>Понятие вычислений в реальном времени. Технология и фреймворк построения систем анализа в реальном времени - Apache Storm.</p> <p>Топология Storm.</p> <p>Существующие системы обработки в реальном времени. Примеры использования.</p> <p>API фреймворка Storm.</p>
<p><b>5. Масштабирование и многоуровневое хранение данных</b></p>	
<p><b>5.1. Теорема CAP</b></p>	<p>Определение теоремы CAP. Типы баз данных: реляционные БД и распределенные базы данных.</p> <p>Шардинг и репликация.</p> <p>Модели развёртывания: частное облако, публичное облако, гибридное облако, общественное облако. Модели обслуживания: программное обеспечение, платформа, инфраструктура. Экономические аспекты центров обработки данных. Безопасность при хранении и пересылке данных. Проблема «последней мили».</p>

<b>5.2. Парадигма NoSQL</b>	Базы данных нетрадиционных моделей данных Not Only SQL (NoSQL). Предпосылки появления NoSQL СУБД, их разновидности, основные отличия от реляционных СУБД, и основные используемые технологии. Шаблоны доступа к данным, хеш-таблица, деревья, таксономия NoSQL, колоночные СУБД. Связь Рассмотрение базы данных HBase.
<b>5.3. Классификация NoSQL баз данных</b>	Классификация баз данных на следующие типы: ключ-значение, колоночные, документно-ориентированные, графовые.
<b>5.4. Дедупликация данных в хранилище</b>	Принципы уменьшения стоимости хранения. Онлайн и оффлайн дедупликация. Сравнение подходов дедупликации. Архитектура системы дедупликации данных. Инструменты дедупликации OpenDedup и возможности AWS по дедупликации при хранении данных в облаке.
<b>6. Анализ текста</b>	
<b>6.1. Стратегии анализа текста</b>	Понятие анализа текста. Анализ временных рядов. Количественная и качественная стратегия анализа текстов. Возможности и ограничения каждого из подходов. Процедура анализа текстов (контент-анализ). Определение круга проблем для контент-анализа.
<b>6.2. Поисковые механизмы: Lucene, Solr, Elasticsearch</b>	Понимание принципов работы поисковых механизмов. Прямой и обратные индексы. Распределенное хранение индексов. Базовая библиотеке поиска Lucene. Полноценный поисковый фреймворк для создания Enterprise приложения - Solr и Elasticsearch. Применение поисковых механизмов для проведение анализа текста.
<b>6.3. Алгоритмы Word2Vec и Glove</b>	Применение алгоритмов Word2Vec и Glove для анализа текста и поиска связей в тексте. Теоретические основы анализа текста на основе векторов. Пример использования векторных алгоритмов для системы автоматического определения уязвимостей.
<b>7. Визуализация данных и результатов анализа</b>	
<b>7.1. Техники визуализации данных</b>	Выполнение визуального анализа данных. Характеристики средств визуализации данных. Метод геометрических преобразований. Отображение иконок. Методы, ориентированные на пиксели. Иерархические образы.

<b>7.2. Введение в язык R</b>	История создания и разработки языка. Его назначение. Статистический анализ. Интерпретатор языка R. IDE для написания программ – RStudio. Типы данных, используемые в R. Базовые операции в R. Базовые статические функции. Общие функции. Преобразование данных для последующего анализа. Руководство по использованию графического интерфейса. Сбор и обработка данных в предметной области. Анализ данных и преобразование их в типы данных языка R. Использование R для извлечения данных из Web-ресурсов.
<b>7.3. Визуализация данных в R</b>	Принципы построения моделей. Визуальное отображение моделей. Оценка моделей.
<b>8. Обзор пройденного материала</b>	
<b>8.1. Ключевые аспекты успешного проекта в области аналитики данных</b>	Умение самостоятельно работать с литературой и материалами лекций.
<b>8.2. Защита курсового проекта</b>	Индивидуальное выступление с демонстрационными материалами для представления результатов по курсовому проекту.
<b>8.3. Зачет</b>	Сдача контрольных работ и отчетов по курсовому проекту. Устное обсуждение и повторение пройденного материала.
<b>8.4. Экзамен</b>	Повторение материала, прочитанного в курсе лекций и закрепленного на практических занятиях, экзамен

### **3. Образовательные технологии**

1. Лекции
2. Практические занятия
3. Самостоятельная работа

### **4. Лабораторный практикум**

Не предусмотрено

### **5. Практические занятия**

№ раздела	Наименование практических занятий (семинаров)	Трудоемкость, ач
		Очная форма
1.	Решение задач по организации жизненного цикла аналитики данных	1
2.	K-mean и C-means кластеризация, классификация	1
3.	Логистическая регрессия, Ассоциации, алгоритм Априори	1
4.	Распределенные вычисления на нескольких серверах	0
5.	Поисковые механизмы: Lucene, Solr, ElasticSearch	2
6.	Введение в язык R	1
7.	Защита курсового проекта	4
8.	Зачет	6
<b>Итого часов</b>		16

## 6. Организация и учебно-методическое обеспечение самостоятельной работы

Самостоятельная работа направлена на закрепление и углубление освоения учебного материала, развитие практических умений. Самостоятельная работа включает следующие виды работы студентов:

- работа с лекционным материалом и учебной литературой;
- поиск, изучение и презентация информации по выбранной проблеме;
- выполнение курсового проекта;
- подготовка к экзамену: включает систематизацию изученного материала, отслеживание связей между разделами программы, работу с лекционным материалом и материалом практических занятий;
- самостоятельное изучение дополнительных разделов, на которые в лекционном материале имеются только ссылки, или подробное рассмотрение которых является составной частью программ дисциплин, использующих результаты изучения данной. Такими разделами являются:
  1. Машинная обработка естественных языков.
  2. Переосмысление глобальных данных: рост объемов и ценности информации (Mapping global data: growth and value creation).

3. Преобразование потенциала больших данных в сфере здравоохранения (The transformative potential of big data in health care domain).
4. Преобразование потенциала больших данных в сфере государственного управления (The transformative potential of big data in public sector administration domain).
5. Преобразование потенциала больших данных в сфере торговли (The transformative potential of big data in retail domain).
6. Преобразование потенциала больших данных в сфере промышленности (The transformative potential of big data in manufacturing domain).
7. Преобразование потенциала больших данных в сфере навигации (The transformative potential of big data in personal location data domain).
8. Влияние «Больших данных» на руководителей организаций (Implications for organization leaders).
9. Влияние «Больших данных» на политику (Implications for policy makers).
10. Планировщик рабочего процесса системы управления рабочими местами – Oozie.
11. Система для эффективного сбора, объединения и перемещения больших объемов данных, полученных из различных источников, в централизованное хранилище данных - Flume.
12. Пользовательский интерфейс Hadoop - Hue.
13. Удаленный вызов процедур и сериализация в Avro.

Методам контроля для работы с лекционным материалом и подготовки к экзамену является контроль преподавателем результатов контрольных работ и ответов на экзаменационные вопросы.

Методами контроля самостоятельного изучения дополнительных разделов дисциплины является либо самоконтроль, либо выступления студентов с докладами на практических занятиях. В последнем случае студент, желающий выступить с докладом, в индивидуальном порядке согласовывает с преподавателем тему своего доклада, а преподаватель выделяет время для доклада на практическом занятии. Такая форма способствует формированию более углубленных знаний у студентов как слушающих, так и готовящих доклады, а готовящие доклады, в дополнение приобретают умения и опыт в педагогической деятельности. Участие студентов в подготовке и обсуждении докладов используется при оценивании качества освоения дисциплины.

Достижимые результаты: знания, умения, навыки и опыт работы с технологиями и средствами анализа больших данных.

## Примерное распределение времени самостоятельной работы студентов

Вид самостоятельной работы	Примерная трудоемкость, ач
	Очная форма
<b>Текущая СР</b>	
работа с лекционным материалом, с учебной литературой	2
опережающая самостоятельная работа (изучение нового материала до его изложения на занятиях)	3
самостоятельное изучение разделов дисциплины	3
выполнение домашних заданий, домашних контрольных работ	0
подготовка к лабораторным работам, к практическим и семинарским занятиям	0
подготовка к контрольным работам, коллоквиумам	7
<b>Итого текущей СР:</b>	15
<b>Творческая проблемно-ориентированная СР</b>	
выполнение расчётно-графических работ	0
выполнение курсового проекта или курсовой работы	1
поиск, изучение и презентация информации по заданной проблеме, анализ научных публикаций по заданной теме	6
работа над междисциплинарным проектом	0
исследовательская работа, участие в конференциях, семинарах, олимпиадах	0
анализ данных по заданной теме, выполнение расчётов, составление схем и моделей на основе собранных данных	0
<b>Итого творческой СР:</b>	7
<b>Общая трудоемкость СР:</b>	22

## 7. Учебно-методическое обеспечение дисциплины

### 7.1. Адрес сайта курса

<https://openedu.ru/course/spbstu/BIGDATA/>

## 7.2. Рекомендуемая литература

### Основная литература

№	Автор, название, место издания, издательство, год (годы) издания	Год изд.	Источник
1	Барсегян А.А. и др. Анализ данных и процессов: Санкт-Петербург: БХВ-Петербург, 2009.	2009	ИБК СПбПУ
2	Никифоров И.В. Курсовое проектирование по учебной дисциплине "Наука о данных и аналитика больших объемов информации": Санкт-Петербург: Изд-во Политехн. ун-та, 2017.	2017	ИБК СПбПУ

### 7.3. Технические средства обеспечения дисциплины

Для выполнения курсового проекта работ студенты по своему выбору могут использовать персональные компьютеры (личные или имеющиеся в учебных классах).

Учебные класс должны соответствовать следующим параметрам:

- операционная система Windows или Linux;
- стандартные пакеты прикладных программ офисного назначения, в том числе: информационные системы подготовки текстов, электронных таблиц, презентаций (Microsoft Word, Microsoft Excel, Microsoft PowerPoint);
- сетевой доступ к профессиональным информационным системам, в том числе вычислительному кластеру ВШПИ.

По курсу имеется набор электронных документов, подготовленных автором программы, представленных на сайте курса, а также выдаваемых студентам на электронных носителях, либо высылаемых по электронной почте (по индивидуальным запросам). Эти документы включают:

- электронные презентации по разделам курса;
- электронные книги и другая учебная литература;
- методические указания по практическим занятиям.

## 8. Материально-техническое обеспечение дисциплины

Мультимедийный проектор.

Компьютерный класс на 15 учебных мест

Выход в Интернет.

Используются персональный компьютер (ноутбук) и проектор для проведения лекций и практических занятий, техническое оснащение компьютерных классов.



## 9. Критерии оценивания и оценочные средства

### 9.1. Критерии оценивания

Для дисциплины «Наука о данных и аналитика больших объемов информации» формой аттестации является зачет. Оценивание качества освоения дисциплины производится с использованием рейтинговой системы.

Максимальное количество баллов: 100

Минимальный балл для получения зачета: 50 %

Оценка	Количество баллов	Описание
неудовлетворительно	0 - 49	Студент не набрал определенное количество баллов для получения зачета, материал для студента сложный, работы не подготовлены в срок и по требованиям. В данной ситуации студент не получает по дисциплине зачет и будет пересдавать в следующем семестре
удовлетворительно	50 - 69	Для получения зачета студенту достаточно набрать 50 %. Для этого необходимо выполнить часть требуемых работ правильно или выполнить все работы но с ошибками.
хорошо	70 - 89	Студент получает зачет, все задания выполнены, существуют незначительные ошибки
отлично	90 - 100	Студент получает зачет, все задания выполнены правильно без существенных замечаний

Уровень освоения дисциплины оценивается по результатам исполнения контрольных заданий, степени владения студентом понятийным аппаратом, умением применять полученные знания в конкретных ситуациях, предлагаемых для рассмотрения, активностью работы на лекциях, результатам выполнения и защиты курсового проекта

В соответствии с рабочим учебным планом, формами контроля являются две контрольные работы, курсовой проект и экзамен. Каждая из форм текущего контроля оценивается по 100-балльной шкале.

## 10. Методические рекомендации по организации изучения дисциплины

Методические рекомендации по каждому разделу приводятся в лекциях в качестве дополнительных слайдов, со ссылкой на рекомендуемую литературу для самостоятельного изучения, и предварительного списка вопросов, заданий для самостоятельной работы.

Адаптированная программа разрабатывается при наличии заявления со стороны обучающегося (родителей, законных представителей) и медицинских показаний (рекомендациями психолого-медико-педагогической комиссии). Для инвалидов адаптированная образовательная программа разрабатывается в соответствии с индивидуальной программой реабилитации.