

**Федеральное государственное бюджетное образовательное  
учреждение высшего образования  
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА  
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ  
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»**

Институт бизнеса и делового администрирования

Факультет международного бизнеса и делового администрирования

УТВЕРЖДЕНА

решением Ученого совета ИБДА

протокол от «13» сентября 2018 г.

№ 4

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Б1.В.01 Бизнес аналитика и анализ больших данных / Big Data &  
Business Analytics**

38.04.02 МЕНЕДЖМЕНТ

---

*(код и наименование направления подготовки)*

**«МЕЖДУНАРОДНЫЙ МЕНЕДЖМЕНТ»**

---

*направленность (профиль)*

**МАГИСТР**

---

*квалификация*

**очная**

---

*форма обучения*

Год набора – 2019

Москва, 2018

**Автор(ы)–составитель(и):**

кандидат технических наук,  
доцент кафедры Менеджмента ИБДА

А.Н. Павлов

проф, д.э.н. заведующий кафедры  
Менеджмента ИБДА

А.Л. Гапоненко

**СОДЕРЖАНИЕ**

1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы	4
2. Объем дисциплины в зачетных единицах с указанием количества академических или астрономических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся и место дисциплины в структуре образовательной программы	7
3. Содержание дисциплины, структурированное по темам (разделам) с указанием отведенного на них количества академических или астрономических часов и видов учебных занятий и структура дисциплины	7
4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине	10
5. Методические указания для обучающихся по освоению дисциплины	18
6. Основная и дополнительная учебная литература, необходимая для освоения дисциплины (модуля), ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	18
7. Материально-техническая база, информационные технологии, используемые при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем (при необходимости)	20

# 1. Перечень планируемых результатов обучения по дисциплине, соотнесенных с планируемыми результатами освоения образовательной программы

1.1. Дисциплина **Б1.В.01 Бизнес аналитика и анализ больших данных / Big Data & Business Analytics** обеспечивает овладение следующими компетенциями с учетом этапа:

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК -4	способностью использовать количественные и качественные методы для проведения прикладных исследований и управления бизнес-процессами, готовить аналитические материалы по результатам их применения	ПК-4.2	Формирование у обучающихся комплексного представления о применении принципов, инструментов и методов технологии «Большие Данные» (BIG DATA) в аналитике бизнес-процессов.

В результате освоения дисциплины у студентов должны быть сформированы:

Обобщенные трудовые функции	Код этапа освоения компетенции	Результаты обучения
Обеспечение сбора и анализа информации	ПК-4.2	<p><u>на уровне знаний</u></p> <ul style="list-style-type: none"> <li>теорию и принципы организации Больших Данных, основные этапы становления и развития концепции Больших Данных</li> <li>определяющие характеристики Больших Данных - «три V»: объём (volume), скорость (velocity), многообразие (variety)</li> <li>основные составляющие обработки структурированных и неструктурированных данных огромных объёмов ресурсного потенциала мирового хозяйства</li> <li>основные проблемы, связанные обработкой и анализом данных, размер которых превосходит возможности типичных баз данных (БД) по занесению, хранению, управлению и анализу информации</li> <li>методы категоризации новых данных на основе принципов, ранее применённых к уже наличествующим данным</li> <li>разницу между бизнес-аналитикой и Большими Данными</li> <li>особенности процесса очистки Больших Данных путём последовательного моделирования: формулирование гипотезы, статистической, визуальной или семантической модели, проверка на основании модели верности гипотезы и генерация</li> </ul>

		<p>следующей гипотезы функционирования мировой экономической системы в современных условиях</p> <ul style="list-style-type: none"> <li>• основные способы отбора данных для анализа объемных и диверсифицируемых массивов данных для получения более точных и релевантных результатов</li> <li>• методику, в которой контрольная выборка поочередно сравнивается с другими для выявления оптимальной комбинации показателей данных</li> <li>• методики для выявления взаимосвязей, т.е. ассоциативных правил, между переменными величинами в больших массивах данных структуры мировой экономики и её основных субъектов;</li> <li>• как выявлять на основе анализа Больших Данных мегатренды мирового экономического процесса и закономерности их развития</li> <li>• статистический метод классификации объектов по группам за счет выявления наперед не известных общих признаков</li> <li>• методику сбора данных из большого количества источников</li> <li>• как задействуется множество предикативных моделей за счет чего повышается качество сделанных прогнозов</li> <li>• как достигается цель создания алгоритмов самообучения на основе анализа эмпирических данных</li> <li>• численные методы для редизайна сложных систем и процессов для улучшения одного или нескольких показателей в принятии стратегических решений, например, состава выводимой на рынок продуктовой линейки, проведении инвестиционного анализа и проч.</li> <li>• набор методик, которые позволяют создать математическую модель наперед заданного вероятного сценария развития событий.</li> </ul> <p><u>на уровне умений</u></p> <ul style="list-style-type: none"> <li>• выявлять логику представления Больших Данных в сфере мировой экономики</li> <li>• применять техники и технологии, которые извлекают смысл из данных на экстремальном пределе практичности (принцип Консалтинговой Компании Forrester) для анализа экономической информации в условиях ограниченного времени</li> <li>• использовать набор техник, позволяющих интегрировать разнородные данные из разнообразных источников для возможности глубинного анализа и уверенно ориентироваться в тематическом информационном поле</li> <li>• находить альтернативные решения возникающих</li> </ul>
--	--	--

		<p>проблем</p> <ul style="list-style-type: none"> <li>• аргументированно отстаивать свою позицию</li> <li>• использовать набор методик, которые позволяют предсказать поведение потребителей в определенном сегменте рынка (принятие решений о покупке, отток, объем потребления и проч.) развитых и развивающихся стран, стран с переходной экономикой</li> <li>• определять наиболее восприимчивые для продвигаемого продукта или услуги категории потребителей, выявлять поведенческие модели потребителей</li> <li>• использовать набор методик анализа связей между узлами в социальных сетях которые позволяют анализировать взаимосвязи между отдельными пользователями, компаниями, сообществами</li> <li>• находить и анализировать информацию, относящуюся к сфере стратегического планирования международной деятельности и международных экономических связей</li> <li>• использовать набор статистических методов для выявления закономерности между изменением зависимой переменной и одной или несколькими независимыми</li> </ul>
		<p><u>на уровне навыков</u></p> <ul style="list-style-type: none"> <li>• понятийным аппаратом концепции Больших Данных</li> <li>• навыками анализа данных огромных объёмов и значительного многообразия национальной и международной статистики для получения результатов, воспринимаемых человеком</li> <li>• навыками использования моделей, построенных на базе статистического анализа, для получения комплексных прогнозов на основе базовых моделей подбора, обработки и интерпретации экономической информации</li> <li>• навыками самостоятельного мышления</li> <li>• навыками самостоятельной работы с национальной и международной статистикой, а также экономической литературой (на русском и английском языках);</li> <li>• навыками анализа экономической ситуации в отдельных странах, регионах, отраслях мирового хозяйства и адекватно презентовать результаты анализа в письменной и устной форме, в том числе с использованием графического материала.</li> <li>• инструментарием оценки экономической конъюнктуры</li> <li>• навыками анализа информации, относящейся к экономическим, экологическим, демографическим процессам</li> <li>• методиками с элементами самообучения для</li> </ul>

		<p>предсказания поведенческих моделей международных рынков и потребителей.</p> <ul style="list-style-type: none"> <li>• навыками вычленения из общего информационного потока сообщений, связанных с интересующим предметом (например, потребительским продуктом)</li> </ul>
--	--	---

**2. Объем дисциплины в зачетных единицах с указанием количества академических или астрономических часов, выделенных на контактную работу обучающихся с преподавателем (по видам учебных занятий) и на самостоятельную работу обучающихся и место дисциплины в структуре образовательной программы.**

#### **Объем дисциплины**

Учебным планом для дисциплины Б1.В.01 «Анализ Больших Данных / Big Data» установлено:

- трудоемкость дисциплины – 4 з.е.,
- контактная работа с преподавателем – 32 часа, в том числе 12 часов – лекции, 20 часов – практические занятия;
- самостоятельная работа – 76 часов.

#### **Место дисциплины в структуре ОП ВО**

Дисциплина Б1.В.01 «Анализ Больших Данных / Big Data», предназначена для студентов 2-го курса, изучается в 3 семестре.

Форма промежуточной аттестации – зачет.

**3. Содержание дисциплины, структурированное по темам с указанием отведенного на них количества академических или астрономических часов и видов учебных занятий и структура дисциплины**

(очная форма обучения)

№ п/п	Наименование тем и/или разделов	Объем дисциплины (модуля), час.					Форма текущего контроля успеваемости*, промежуточной аттестации	
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий					СР
			Л	ЛР	ПЗ	КСР		
Тема 1	Введение. Задачи комплексной аналитики. Наука о данных и интеллектуальный анализ данных (BI). Концепции технологии и экосистемы «Большие Данные» (BIG DATA). Возможности применения технологии «Большие Данные».	18	2		4		12	КР

№ п/п	Наименование тем и/или разделов	Объем дисциплины (модуля), час.						Форма текущего контроля успеваемости*, промежуточной аттестации
		Всего	Контактная работа обучающихся с преподавателем по видам учебных занятий				СР	
			Л	ЛР	ПЗ	КСР		
Тема 2	Жизненный цикла анализа Больших Данных. Структурирование бизнес проблемы как проблемы аналитической. Главные результаты аналитического проекта.	16	2		2		12	П
Тема 3	Методы анализа и технологии Больших Данных.	18	2		4		12	Р
Тема 4	Комплексная аналитика. Теория и методы.	20	2		4		14	Э
Тема 5	Комплексная аналитика. Технология и средства. Отраслевые Решения Больших Данных.	20	2		4		14	ПО
Тема 6	Разработка окончательных результатов анализа Больших Данных.	16	2		2		12	Р
Промежуточная аттестация								экзамен
Всего:		144	12	0	20	0	76	36

Примечание:

\* – формы текущего контроля успеваемости:

Опрос (О)

Кейс (К)

Групповая работа (ГР)

Презентация (Пр)

### **Тема 1. Введение. Задачи комплексной аналитики. Наука о данных и интеллектуальный анализ данных (ВІ). Концепции технологии и экосистемы «Большие Данные» (BIG DATA). Возможности применения технологии «Большие Данные»**

Определение Больших Данных. Характеристики Больших Данных: объём, сложность типов и структур, скорость создания и обновления. Анализ потоков неструктурированных, полуструктурированных и квази-структурированных Больших Данных. Репозитории Больших Данных. Необходимость применения комплексных аналитических методов. Применение методов интеллектуального анализа данных (ВІ) и науки о данных для достижения стратегических целей организаций. Архитектура науки о данных. Новые подходы в анализе Больших Данных. Возможности новых подходов в анализе экосистем Больших Данных. Ключевые роли новой экосистемы Больших Данных. Знания и компетенции необходимые в новой экосистеме Больших Данных. Профиль Аналитика Данных. Примеры применения аналитики Больших Данных в вертикальных рынках.

### **Тема 2. Жизненный цикла анализа Больших Данных. Структурирование бизнес проблемы как проблемы аналитической. Главные результаты аналитического проекта**

Определение аналитической проблемы. Ценность использования жизненного цикла анализа Больших Данных. Принципы управления проектами исследования Больших Данных. Ключевые



роли успешного аналитического проекта. Жизненный цикл проекта анализа Больших Данных: обследование, подготовка данных, планирование модели, построение модели, подготовка результатов, использование результатов. Обследование: изучение бизнес домена, изучение прошлого опыта, анализ доступных и достаточных ресурсов, структурирование проблемы, формулировка исходной гипотезы, определение источников данных, изучение данных, определение подходов к структурированию аналитической проблемы. Подготовка данных: определение качества и структуры данных, выбор данных и определение взаимосвязей массивов данных, изучение, очистка и нормализация данных, обзор и визуализация данных. Планирование модели: определение аналитических методов основанных на гипотезе, структуре и объёме данных, обеспечение достижения бизнес целей с использованием выбранных подходов и технологий, определение и документирование допущений модели, выбор переменных, выбор модели, конверсия в SQL или язык базы данных для лучшего представления данных. Построение модели: подготовка массивов данных, построение информационно – вычислительной среды, включающей высокопроизводительные вычислительные комплексы, технологии и средства параллельной обработки данных, параллельной обработки данных.

Подготовка результатов: интерпретация результатов, сравнение результатов с исходной гипотезой, определение ключевых результатов их бизнес ценности, подготовка рекомендаций для будущего использования. Использование результатов: запуск пилота, оценка бизнес выгоды, получение окончательных результатов, старт операционного использования модели, определение путей дальнейшего развития и улучшения модели. Четыре основных результата анализа Больших Данных для удовлетворения стейкхолдеров: презентация для спонсоров проекта, презентация для аналитиков, коды для технического персонала, технические спецификации для исполнения кода.

### **Тема 3. Методы анализа и технологии Больших Данных**

Использование основных аналитических методов, статистических распределений, тестов для обследования массивов данных.

### **Тема 4. Комплексная аналитика. Теория и методы**

Определение аналитических потребностей и технологий. Использование методов аналитических решений. Оценка созданных моделей с помощью диагностических методов.

Технологии решения проблем: кластеризация, ассоциативные ряды, регрессионный анализ, методы классификации, анализ временных серий, анализ больших текстов. K-means кластеризация. Правила алгоритмов ассоциаций. Алгоритм Априори. Линейный регрессионный анализ. Модель логистической регрессии. Деревья решений. Анализ временных интервалов. Методы ARMA, ARIMA. Анализ временных интервалов. Анализ текстов.

### **Тема 5. Комплексная аналитика. Технология и средства. Отраслевые Решения Больших Данных**

Аналитические методы неструктурированных данных. Модель MapReduce обработки Больших Данных и извлечения только необходимых данных. Декомпозиция задач на стадии с их последующей параллельной обработкой. Обработка естественных текстов. «Векторизация» документов путём представления тестов в виде математических объектов. Две главные компоненты Hadoop: HDFS для хранения Больших Данных и MapReduce для их анализа. Использование RHadoop для прямого доступа к HDFS и MapReduce. Функции Окна для обработки рядов данных. Двигающиеся Окна Больших Данных.

Примеры отраслевых решений Больших Данных: Банки, Телеком, Ритейл, Энергетика, Медицина.

## **Тема 6. Разработка окончательных результатов анализа Больших Данных**

Шаги получения окончательных результатов жизненного цикла анализа Больших Данных: запуск пилота, оценка выгод, получение окончательных результатов, оценка результативности модели, определение процесса улучшения модели, при необходимости. Требования основных участников аналитического проекта: конечного бизнес – пользователя, спонсора проекта, менеджера проекта, бизнес аналитика, программиста, администратора базы данных, аналитика данных. Четыре основных результата необходимых стейкхолдерам аналитического проекта: презентация для спонсора проекта, презентация для аналитиков, код для технических специалистов, технические спецификации для реализации кода. Подготовка двух презентаций для бизнес и технических аудиторий. Рекомендации по содержанию презентации для спонсора проекта: акцент на влиянии результатов проекта на бизнес организации, возврат инвестиций, управление рисками. Рекомендации по содержанию презентации для аналитиков: техники поддержки реализации и сопровождения модели данных в реальной среде. Использование различных моделей визуализации данных для презентаций спонсору проекта и аналитикам. Заключение. Подведение итогов курса.

### **4. Материалы текущего контроля успеваемости обучающихся и фонд оценочных средств промежуточной аттестации по дисциплине**

#### **4.1. Формы и методы текущего контроля успеваемости обучающихся и промежуточной аттестации**

**4.1.1.** В ходе реализации дисциплины используются следующие методы текущего контроля успеваемости обучающихся:

<b>Тема</b>	<b>Методы текущего контроля успеваемости</b>
Тема 1. Введение. Задачи комплексной аналитики. Наука о данных и интеллектуальный анализ данных (BI). Концепции технологии и экосистемы «Большие Данные» (BIG DATA). Возможности применения технологии «Большие Данные».	Контрольная Работа
Тема 2. Жизненный цикла анализа Больших Данных. Структурирование бизнес проблемы как проблемы аналитической. Главные результаты аналитического проекта.	Презентация
Тема 3. Методы анализа и технологии Больших Данных.	Реферат
Тема 4. Комплексная аналитика. Теория и методы.	Эссе
Тема 5. Комплексная аналитика. Технология и средства. Отраслевые Решения Больших Данных.	Письменный Опрос
Тема 6. Разработка окончательных результатов анализа Больших Данных.	Реферат

#### **Задания для самостоятельной работы по темам**

*Тема 1. Введение. Задачи комплексной аналитики. Наука о данных и интеллектуальный анализ данных (BI). Концепции технологии и экосистемы «Большие Данные» (BIG DATA). Возможности применения технологии «Большие Данные».*

Задание. На основании лекционного материала и материала учебников составьте таблицу анализа потоков неструктурированных, полу-структурированных и квази-структурированных Больших Данных (форма произвольная).

*Тема 2. Жизненный цикл анализа Больших Данных. Структурирование бизнес проблемы как проблемы аналитической. Главные результаты аналитического проекта.*

Задание. Самостоятельно повторите темы: Принципы управления проектами исследования Больших Данных. Ключевые роли успешного аналитического проекта. Жизненный цикл проекта анализа Больших Данных и ответьте на следующие вопросы (для самопроверки):

- 1) Каковы ключевые роли успешного аналитического проекта ?
- 2) Чем отличается роль Заказчика от роли Спонсора аналитического проекта ?
- 3) Каковы этапы жизненного цикла проекта анализа Больших Данных ?

*Тема 3. Методы анализа и технологии Больших Данных*

Задание.

- 1) Проанализируйте использование технологии R как средства представления, отчётности и визуализации аналитических данных и ответьте на вопросы :

- Как осуществляются ввод данных в R и получение данных из R ?
- Каковы типы данных, используемых в R ?
- Каковы основные операции в R ?
- Как используются функции генерации отчётов ?

Источники

[https://en.wikipedia.org/wiki/Big\\_data](https://en.wikipedia.org/wiki/Big_data)

[http://www.sas.com/en\\_us/insights/big-data/what-is-big-data.html](http://www.sas.com/en_us/insights/big-data/what-is-big-data.html)

[http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation)

*Тема 4. Механизм мирового хозяйства*

Задание. Самостоятельно изучите технологии решения проблем:

кластеризация, ассоциативные ряды, регрессионный анализ, методы классификации, анализ временных серий, анализ больших текстов и ответьте на вопросы :

- 1) Что такое K-means кластеризация ?
- 2) Каковы правила алгоритмов ассоциаций ?
- 3) Как применять алгоритм Априори ?

Ответы оформите в виде эссе (2 стр.)

Источники

[http://www.webopedia.com/TERM/B/big\\_data.html](http://www.webopedia.com/TERM/B/big_data.html)

<http://www-01.ibm.com/software/data/bigdata/>

<http://www.theguardian.com/technology/big-data>

*Тема 5. Комплексная аналитика. Технология и средства. Отраслевые Решения Больших Данных.*

Задание. Углубленное изучение темы «Две главные компоненты Hadoop: HDFS для хранения Больших Данных и MapReduce для их анализа».

Написание эссе на тему: «Использование RНadoop для прямого доступа к HDFS и MapReduce» (1 стр.)

*Тема 6. Разработка окончательных результатов анализа Больших Данных.*

Задание. Написание реферата (см. список тем)

**4.1.2. Зачет проводится в форме собеседования по подготовленному эссе.**

**4.2. Материалы текущего контроля успеваемости обучающихся**

**Примерные темы рефератов**

К теме 3.

1. Использование основных аналитических методов, статистических распределений, тестов для обследования массивов данных.
2. Использование технологии R как средства представления, отчётности и визуализации аналитических данных.
3. Использование графического интерфейса пользователя R
4. Ввод данных в R и получение данных из R. Типы данных используемых в R. Основные операции в R. Функции генерации отчётов.
5. R как язык объектно – ориентированного программирования.
6. Определение аналитических потребностей и технологий.
7. Использование методов аналитических решений.
8. Оценка созданных моделей с помощью диагностических методов.
9. Использование R и аналитических функций для оценки моделей.

К теме 6.

1. Шаги получения окончательных результатов жизненного цикла анализа Больших Данных
2. Требования основных участников аналитического проекта: конечного бизнес – пользователя, спонсора проекта, менеджера проекта, бизнес аналитика, программиста, администратора базы данных, аналитика данных.
3. Четыре основных результата необходимых стейкхолдерам аналитического проекта: презентация для спонсора проекта, презентация для аналитиков, код для технических специалистов, технические спецификации для реализации кода.
4. Рекомендации по содержанию презентации для спонсора проекта: акцент на влиянии результатов проекта на бизнес организации, возврат инвестиций, управление рисками.
5. Рекомендации по содержанию презентации для аналитиков: техники поддержки реализации и сопровождения модели данных в реальной среде.
6. Использование различных моделей визуализации данных для презентаций спонсору проекта и аналитикам.

**Примерный перечень вопросов для подготовки к опросу и контрольным работам**

1. Определение аналитической проблемы.

2. Ценность использования жизненного цикла анализа Больших Данных.
3. Принципы управления проектами исследования Больших Данных.
4. Ключевые роли успешного аналитического проекта.
5. Жизненный цикл проекта анализа Больших Данных: обследование, подготовка данных, планирование модели, построение модели, подготовка результатов, использование результатов.
6. Обследование: изучение бизнес домена, изучение прошлого опыта, анализ доступных и достаточных ресурсов, структурирование проблемы, формулировка исходной гипотезы, определение источников данных, изучение данных, определение подходов к структурированию аналитической проблемы.
7. Подготовка данных: определение качества и структуры данных, выбор данных и определение взаимосвязей массивов данных, изучение, очистка и нормализация данных, обзор и визуализация данных.
8. Планирование модели: определение аналитических методов основанных на гипотезе, структуре и объёме данных, обеспечение достижения бизнес целей с использованием выбранных подходов и технологий.
9. Определение и документирование допущений модели, выбор переменных, выбор модели, конверсия в SQL или язык базы данных для лучшего представления данных.
10. Построение модели: подготовка массивов данных, построение информационно – вычислительной среды, включающей высокопроизводительные вычислительные комплексы, технологии и средства параллельной обработки данных.
11. Подготовка результатов: интерпретация результатов, сравнение результатов с исходной гипотезой, определение ключевых результатов их бизнес ценности, подготовка рекомендаций для будущего использования.
12. Использование результатов: запуск пилота, оценка бизнес выгоды, получение окончательных результатов, старт операционного использования модели, определение путей дальнейшего развития и улучшения модели.
13. Четыре основных результата анализа Больших Данных для удовлетворения стейкхолдеров.

#### 4.3. Оценочные средства для промежуточной аттестации

##### 4.3.1. Перечень компетенций с указанием этапов их формирования в процессе освоения образовательной программы

Код компетенции	Наименование компетенции	Код этапа освоения компетенции	Наименование этапа освоения компетенции
ПК -4	способностью использовать количественные и качественные методы для проведения прикладных исследований и управления бизнес-процессами, готовить аналитические материалы по результатам их применения	ПК-4.2	Формирование у обучающихся комплексного представления о применении принципов, инструментов и методов технологии «Большие Данные» (BIG DATA) в аналитике бизнес-процессов.

##### 4.3.2 Показатели и критерии оценивания компетенций на различных этапах их формирования

Этап освоения компетенции	Показатель оценивания	Критерий оценивания
ПК-4.2	Знание концепции технологии и экосистемы «Большие Данные» (BIG DATA), роли Аналитики Данных, потоков не структурированных данных и их репозитариев	<p>Знает теорию и принципы организации Больших Данных, основные этапы становления и развития концепции Больших Данных;</p> <p>Знает определяющие характеристики Больших Данных - «три V»: объём (volume), скорость (velocity), многообразие (variety)</p> <p>Умеет выявлять логику представления Больших Данных в сфере мировой экономики</p> <p>Владеет понятийным аппаратом концепции Больших Данных</p>
	Умение использовать основные аналитические методы для исследования массивов Больших Данных и представления результатов анализа Больших Данных;	<p>Знает основные способы отбора данных для анализа объемных и диверсифицируемых массивов данных для получения более точных и релевантных результатов</p> <p>Знает набор методик, которые позволяют создать математическую модель наперед заданного вероятного сценария развития событий.</p> <p>Умеет использовать набор статистических методов для выявления закономерности между изменением зависимой переменной и одной или несколькими независимыми</p> <p>Владеет навыками вычленения из общего информационного потока сообщений, связанных с интересующим предметом (например, потребительским продуктом)</p>
	Владение навыками количественного и качественного анализа Больших Данных, характеризующих тенденции развития организации, мировых рынков и рынков отдельных стран.	<p>Демонстрирует знание основных подходов проведения статистического анализа</p> <p>Умеет использовать набор методик, которые позволяет предсказать поведение потребителей в определенном сегменте рынка</p>

		<p>(принятие решений о покупке, отток, объем потребления и проч.) развитых и развивающихся стран, стран с переходной экономикой</p> <p>Владеет навыками</p> <ul style="list-style-type: none"> <li>– самостоятельной работы с национальной и международной статистикой, а также экономической литературой (на русском и английском языках);</li> <li>– анализа экономической ситуации в отдельных странах, регионах, отраслях мирового хозяйства и адекватно презентовать результаты анализа в письменной и устной форме, в том числе с использованием графического материала</li> </ul>
--	--	--

**4.3.3 Типовые контрольные задания или иные материалы (типовые оценочные материалы), необходимые для оценки знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций в процессе освоения образовательной программы**

Промежуточная аттестация по дисциплине – зачет, который проводится в форме собеседования по подготовленному заранее эссе.

Примерная тема эссе: «Принятие управленческого решения для разрешения проблемной ситуации».

Студенту будет представлен кейс с описанием компании, и текущей ситуации в ней, включая имеющиеся проблемы и сложности.

Студенту будет предложено разработать управленческое решение для разрешения проблем, описанных в кейсе, от лица одного из топ-менеджеров данной компании.

**Список вопросов для подготовки к зачёту.**

1. Определение Больших Данных.
2. Характеристики Больших Данных: объём, сложность типов и структур, скорость создания и обновления.
3. Анализ потоков неструктурированных, полу-структурированных и квази-структурированных Больших Данных.
4. Репозитории Больших Данных.
5. Необходимость применения комплексных аналитических методов. Применение методов интеллектуального анализа данных (BI) и науки о данных для достижения стратегических целей организаций.
6. Архитектура науки о данных. Новые подходы в анализе Больших Данных. Возможности новых подходов в анализе экосистем Больших Данных.
7. Ключевые роли новой экосистемы Больших Данных. Знания и компетенции необходимые в новой экосистеме Больших Данных.

8. Профиль Аналитика Данных. Примеры применения аналитики Больших Данных в вертикальных рынках
9. Аналитические методы неструктурированных данных. Модель MapReduce обработки Больших Данных и извлечения только необходимых данных.
10. Декомпозиция задач на стадии с их последующей параллельной обработкой.
11. Платформа HDFS для хранения Больших Данных и MapReduce для реализации MapReduce модели параллельной обработки Больших Данных.
12. Обработка естественных текстов. «Векторизация» документов путём представления тестов в виде математических объектов.
13. Две главные компоненты Hadoop: HDFS для хранения Больших Данных и MapReduce для их анализа.
14. Использование RНadoop для прямого доступа к HDFS и MapReduce.
15. Функции Окна для обработки рядов данных. Двигающиеся Окна Больших Данных.
16. Примеры отраслевых решений Больших Данных: Банки, Телеком, Ритейл, Энергетика, Медицина.

### Шкала оценивания

<b>85-100 баллов</b>	Программа курса освоена полностью. Студент свободно ориентируется в проблематике всех тем курса, владеет терминологией, умеет использовать теоретический материал и научный инструментарий для решения практических задач. Понимает место данной учебной дисциплины в своем профессиональном становлении. Обладает высоким уровнем профессиональной эрудиции. Умеет рассматривать конкретные вопросы учебной дисциплины в контексте, выходящем за рамки этой дисциплины в междисциплинарное пространство. Имеет собственную аргументированную позицию по ключевым проблемам курса.
<b>84-70 баллов</b>	Программа курса освоена полностью. Студент ориентируется в проблематике практически всех тем курса. Владеет терминологией. Понимает, каким образом теоретический материал может быть использован в качестве инструментария для решения практических задач. Обладает достаточным уровнем профессиональной эрудиции.
<b>69-55 баллов</b>	Программа курса в основном освоена. Глубина понимания отдельных тем курса варьируется. Студент владеет терминологией. Имеет представление, каким образом теоретический материал может использоваться в качестве инструмента для решения практических задач. Демонстрирует базовый уровень профессиональной эрудиции.
<b>Менее 55 баллов</b>	Программа курса не освоена. Понимание основных тем курса является поверхностным. Студент плохо знаком с терминологией.  Профессиональная эрудиция отсутствует.

### Перевод баллов в традиционную систему оценки:

Баллы по 100-балльной системе	Пятибалльная система оценки	Система оценивания «зачтено-не зачтено»
85-100 баллов	отлично	Зачтено
70-84 баллов	хорошо	зачтено



55-69 баллов	удовлетворительно	зачтено
Менее 55 баллов	неудовлетворительно	незачтено

#### 4.4. Методические материалы, определяющие процедуры оценивания знаний, умений, навыков и (или) опыта деятельности, характеризующих этапы формирования компетенций

К основной форме контроля, определяющей процедуры оценивания знаний, умений, навыков, характеризующих этапы формирования компетенций по дисциплине, относится промежуточная аттестация в форме письменно-устного зачета.

Выполнение заданий, предусмотренных программой курса, является обязательным для всех студентов. Студенты, не выполнившие в полном объеме все эти задания, не допускаются к сдаче зачета.

Результаты текущего и итогового контроля формируют рейтинговую оценку работы студента. Оценки за разные виды работы выставляются исходя из следующих критериев:

Вид работы	Оценка/Процент	Описание критериев оценки
Письменная работа	(90-100%)	В работе раскрыто блестящее и всестороннее понимание лекционного материала, приведены разнообразные кейсы и ситуации, дана их оригинальная интерпретация.
	(82-89%)	В работе раскрыто всестороннее понимание лекционного материала, приведен один кейс или ситуация, дана их глубокая интерпретация.
	(75-81%)	В работе раскрыто всестороннее понимание лекционного материала, приведен кейс или ситуация, дана их достаточная и очевидная интерпретация.
	(67-74%)	В работе раскрыто поверхностное и неполное понимание лекционного материала, приведен кейс или ситуация, дана их в общем и целом приемлемая интерпретация.
	(60-66%)	В работе раскрыто поверхностное и фрагментарное понимание лекционного материала, приведен кейс или ситуация, дана их выборочная интерпретация.
	(менее 60%)	Отсутствует целостное понимание лекционного материала. Анализ кейсов неадекватен теоретической фактуре.
Работа на семинаре	(90-100%)	Оригинальная презентация самостоятельной работы, активное участие и критичное участие в дискуссии.
	(82-89%)	Оригинальная презентация самостоятельной работы при ограниченном участии в дискуссии.
	(75-81%)	Стандартная (достаточная) презентация при ограниченном участии в дискуссии.
	(67-74%)	Стандартная презентация при отсутствии участия в дискуссии.
	(60-66%)	Стандартная презентация с видимыми ошибками.
	(менее 60%)	Отсутствие понимания предмета в презентации и

## 5. Методические указания для обучающихся по освоению дисциплины

Дисциплина «Большие Данные» (BIG DATA) в Аналитике Международных процессов нацелена на формирование у обучающихся комплексного представления о применении принципов, инструментов и методов технологии «Большие Данные» (BIG DATA) в аналитике международных процессов Евразии. Формат дисциплины предполагает активное и заинтересованное участие обучающихся во всех формах освоения дисциплины, включая лекционные и практические занятия, а также самостоятельную работу обучающихся.

При подготовке к практическим занятиям обучающимся рекомендуется:

1. Внимательно изучить вопрос/вопросы (задание).
  2. Прочитать предложенную Вам литературу и материалы лекции по соответствующей теме.
  3. Найти рекомендованные источники.
  4. Сформировать свою «базу данных» по выбранному вопросу (заданию).
  5. Составить план ответа (доклада, презентации).
  6. Весь материал (статистику, информацию на сайтах, литературные источники, как на русском, так и на иностранных языках) обработать и использовать реферативно
  7. Указать авторство использованных материалов
  8. При чтении литературы кратко конспектировать основные положения статьи или главы
  8. Не использовать рефераты и курсовые, помещенные в Интернете.
  9. Планировать свой ответ (доклад, презентацию) по времени.
  10. Не пересказывать материал, а презентовать свои выводы и заключения, доказывая и/или иллюстрируя их конкретными экономическими примерами, результатами расчетов, произведенных самостоятельно на основе использования российских и международных баз экономической информации.
  11. При работе в малых группах в начале работы четко распределить обязанности.
- Основными видами самостоятельной работы по дисциплине «Большие Данные» (BIG DATA) в Аналитике Международных процессов являются:
- повторение лекционного материала и материала учебников;
  - чтение основной и дополнительной литературы
  - самостоятельное изучение разделов курса (см. «Задания для самостоятельной работы по темам»)<sup>1</sup>
  - подготовка к практическим занятиям;
  - подготовка к промежуточной аттестации;
  - написание рефератов.

## 6. Учебная литература и ресурсы информационно-телекоммуникационной сети "Интернет", включая перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

### 6.1 Основная литература

1. Миркин, Б. Г. Введение в анализ данных: учебник и практикум / Б. Г. Миркин. — М.: Издательство Юрайт, 2019. — 174 с. — (Авторский учебник). — ISBN 978-5-9916-5009-0. —

<sup>1</sup> Задания для самостоятельной работы, не предполагающие написания реферата или эссе, оцениваются в ходе промежуточной и итоговой аттестации.

Текст: электронный // ЭБС Юрайт [сайт]. — URL: <https://biblio-online.ru/bcode/432851> (дата обращения: 11.02.2020).

2. Preimesberger, Chris Hadoop, Yahoo, 'Big Data' Brighten BI Future (англ.). EWeek (15 August 2011).
3. Черняк, Леонид Большие Данные — новая теория и практика (рус.) // Открытые системы. СУБД. — М.: Открытые системы, 2011. — № 10. — ISSN 1028-7493.
4. Моррисон, Алан и др. Большие Данные: как извлечь из них информацию (рус.). Технологический прогноз. Ежеквартальный журнал, российское издание, 2010 выпуск 3. PricewaterhouseCoopers (17 декабря 2010).
5. Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data (англ.). Gartner (27 June 2011).
6. Manyika, James et al. Big data: The next frontier for innovation, competition, and productivity (англ.) (PDF). McKinsey Global Institute, June, 2011. McKinsey (9 August 2011).
7. Большие данные: современные подходы к хранению и обработке. П.А. Клеменков, С.Д. Кузнецов. Труды Института системного программирования. РАН Том 23. 2012 г. Стр. 143-158. ISSN 2220-6426 (Online), ISSN 2079-8156

#### 6.2 Дополнительная литература:

1. Couldry, Nick; Turow, Joseph (2014). "Advertising, Big Data, and the Clearance of the Public Realm: Marketers' New Approaches to the Content Subsidy". *International Journal of Communication* 8: 1710–1726.
2. Layton, Julia. "Amazon Technology". Money.howstuffworks.com. Retrieved 2013-03-05.
3. "Scaling Facebook to 500 Million Users and Beyond". Facebook.com. Retrieved 2013-07-21.
4. "Google Still Doing at Least 1 Trillion Searches Per Year". Search Engine Land. 16 January 2015. Retrieved 15 April 2015.
5. Lamb, Charles. "Oracle NoSQL Database Exceeds 1 Million Mixed YCSB Ops/Sec".
6. "FICO® Falcon® Fraud Manager". Fico.com. Retrieved 2013-07-21.
7. "eBay Study: How to Build Trust and Improve the Shopping Experience". Knowwpcarey.com. 8 May 2012. Retrieved 2015-12-20.
8. Leading Priorities for Big Data for Business and IT. eMarketer. October 2013. Retrieved January 2014.
9. Wingfield, Nick (12 March 2013). "Predicting Commutes More Accurately for Would-Be Home Buyers – NYTimes.com". Bits.blogs.nytimes.com. Retrieved 2013-07-21.
10. Alexandru, Dan. "Prof" (PDF). cds.cern.ch. CERN. Retrieved 24 March 2015.
11. Форман Дж., Много цифр: Анализ больших данных при помощи Excel [Электронный ресурс] — Электрон. дан. — Москва : Альпина Паблишер, 2016. — 461 с. — Режим доступа: . — Загл. с экрана.

#### 6.3 Интернет-ресурсы, справочные системы.

1. [http://rusrim.blogspot.ru/2015/04/blog-post\\_2.html](http://rusrim.blogspot.ru/2015/04/blog-post_2.html) Бизнес и власть Аналитические статьи, интервью, информационные заметки обо всем, что касается взаимодействия бизнеса и власти (сайт журнала «Эксперт» <http://expert.ru/dossier/story/128475> )
2. [https://www.capgemini-consulting.com/resource-file-access/resource/pdf/cracking\\_the\\_data\\_conundrum-big\\_data\\_pov\\_13-1-15\\_v2.pdf](https://www.capgemini-consulting.com/resource-file-access/resource/pdf/cracking_the_data_conundrum-big_data_pov_13-1-15_v2.pdf)
3. [http://www.eos.ru/eos\\_delopr/eos\\_delopr\\_intesting/detail.php?id=131174](http://www.eos.ru/eos_delopr/eos_delopr_intesting/detail.php?id=131174)
4. [https://en.wikipedia.org/wiki/Big\\_data](https://en.wikipedia.org/wiki/Big_data)
5. [http://www.sas.com/en\\_us/insights/big-data/what-is-big-data.html](http://www.sas.com/en_us/insights/big-data/what-is-big-data.html)
6. [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innov](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innov)

ation

7. [http://www.webopedia.com/TERM/B/big\\_data.html](http://www.webopedia.com/TERM/B/big_data.html)
8. <http://www-01.ibm.com/software/data/bigdata/>
9. <http://www.theguardian.com/technology/big-data>

**7. Материально-техническая база, информационные технологии, используемые при осуществлении образовательного процесса по дисциплине, включая перечень программного обеспечения и информационных справочных систем (при необходимости)**

Аудитории достаточной вместимости.

Ноутбук (пакет офисных приложений MS Office)

Мультимедийный проектор

Флип-чарт и маркеры