

Федеральное государственное автономное образовательное учреждение
высшего образования
«Российский государственный гуманитарный университет»
(ФГАОУ ВО «РГГУ»)

ИНСТИТУТ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ НАУК

ЭКОНОМИЧЕСКИЙ ФАКУЛЬТЕТ

Кафедра социально-экономической статистики и демографии

АНАЛИЗ БОЛЬШИХ ДАННЫХ

РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ

01.03.05 Статистика

38.03.01 Экономика

Код и наименование направления подготовки/специальности

Социально-экономическая статистика и демография

Наименование направленности (профиля)/специализации

Уровень высшего образования: *бакалавриат*

Форма обучения: *очная*

РПД адаптирована для лиц
с ограниченными возможностями
здоровья и инвалидов

Москва 2026

АНАЛИЗ БОЛЬШИХ ДАННЫХ
Рабочая программа дисциплины
Составитель:
Канд.техн.наук, доц. Сысоева Л.А.

УТВЕРЖДЕНО:

Протокол заседания кафедры социально-экономической
статистики и демографии
№4 от 10.11. 2025 года

ОГЛАВЛЕНИЕ

1	Пояснительная записка.....	4
1.1	Цель и задачи дисциплины.....	4
1.2	Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций.....	4
1.2	Место дисциплины в структуре образовательной программы.....	6
2	Структура дисциплины.....	6
3	Содержание дисциплины.....	6
4	Образовательные технологии.....	8
5	Оценка планируемых результатов обучения.....	8
5.1	Система оценивания.....	8
5.2	Критерии выставления оценки по дисциплине.....	9
5.3	Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине.....	11
6	Учебно-методическое и информационное обеспечение дисциплины.....	12
6.1	Список литературы.....	12
6.2	Перечень ресурсов информационно-коммуникационной среды “Интернет”.....	12
6.3	Профессиональные базы данных и информационно-справочные системы.....	12
7	Материально-техническое обеспечение дисциплины.....	13
8	Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов.....	14
9	Методические материалы.....	15
9.1	Планы практических занятий.....	15
9.2	Методические рекомендации по подготовке письменных работ.....	16
	Приложение 1	17

1 Пояснительная записка

1.1 Цель и задачи дисциплины

Цель дисциплины – формирование знаний, умений, навыков в области систем и технологий обработки больших объемов структурированной или неструктурированной информации.

Задачи дисциплины:

1. Изучение технологий хранения, обработки и анализа больших данных.
2. Формирование умений рационального использования систем и технологий обработки больших данных.
3. Применение статистических и математических методов для анализа больших объемов информации.

1.2. Перечень планируемых результатов обучения по дисциплине, соотнесенных с индикаторами достижения компетенций

Компетенция (код и наименование)	Индикаторы компетенций (код и наименование)	Результаты обучения
ПК-1 Способен осуществлять сводку статистических данных с целью формирования информационной базы проводимого исследования	ПК-1.1 Формирует входные массивы статистических данных ПК-1.2 Проводит сводку единиц статистического наблюдения в соответствии с принятыми методиками ПК-1.3 Формирует выходные массивы статистической информации	Знать методы решения задач обработки и анализа больших данных, возможности высокопроизводительных вычислительных систем, технологии распределенных вычислений, методы и модели Data Mining; Уметь использовать современные информационные технологии и программные средства в области анализа и обработки больших данных. Владеть основными методами Data Mining для обнаружения в данных ранее неизвестных, практически полезных и доступных интерпретаций знаний; методами статистического и математического анализа больших данных.

1.2 Место дисциплины в структуре образовательной программы

Дисциплина «Анализ больших данных» относится к обязательной части блока дисциплин учебного плана.

Для освоения дисциплины необходимы знания, умения и владения, сформированные в ходе изучения следующих дисциплин: «Теория вероятностей и математическая статистика», «Информационные технологии и системы в сфере статистики», «Теория статистики», «Математический анализ».

В результате освоения дисциплины формируются знания, умения и владения, необходимые для изучения следующих дисциплин: «Статистический анализ социологических данных», «Статистика населения», «Социологические методы в демографических исследованиях», «Интеллектуальный анализ статистических данных».

2 Структура дисциплины

Структура дисциплины для очной формы обучения

Общая трудоёмкость дисциплины составляет 3 з.е., 108 ч., в том числе контактная работа обучающихся с преподавателем 42 ч., самостоятельная работа обучающихся 66 ч.

Семестр	Тип учебных занятий	Количество часов
4	Лекция	18
4	Семинар	24
Всего:		42

3 Содержание дисциплины

№	Наименование раздела дисциплины	Содержание
1	Введение в анализ больших данных	Основные определения, термины, задачи анализа больших данных. Понятие Data Mining. Когнитивный анализ данных. Обзор источников информации для Big Data (открытые источники информации: статистические сборники, опубликованные отчеты и результаты исследований; доступ к закрытой информации). Методики сбора данных. Признаки больших данных. Источники больших данных. Характеристики VVV («объем», «скорость», «многообразие»)
2	Технологии хранения и обработки больших данных	Обзор технологий хранения больших данных. Базы данных. Системы управления базами данных. Модели данных. Подготовка исходных данных для анализа: первичная обработка и визуализация имеющихся данных. Поиск источников информации в сети Интернет: открытые и закрытые источники данных. Портал открытых данных РФ. Сохранение данных в программе MS Excel. Преобразование и первичная обработка данных.
3	Методы анализа больших данных	Обзор методов анализа больших данных.

		Смешение и интеграция данных. Краудсорсинг. Машинное обучение. Искусственные нейронные сети. Распознавание образов. Обзор методов анализа больших данных класса Data Mining. Методы классификации, моделирования и прогнозирования, генетические алгоритмы, эволюционное программирование, нечеткая логика, статистические методы.
4	Современные технологии обработки больших данных	Обзор технологий обработки больших данных. Программно-аппаратные комплексы. Базы данных новой архитектуры. Аналитические платформы. Специализированные языки программирования. Обзор современных популярных программных средств анализа данных. Платные, бесплатные программные средства: Statistica, SPSS, Excel, R-Studio и другие; их преимущества и недостатки.
5	Методы обработки, анализа и визуализации больших данных	Представление исходных данных в программе RStudio (векторы, массивы, матрицы, списки, таблицы). Статистическая обработка данных в программах MS Excel и RStudio: подсчет описательных статистик, графическое представление данных. Группировка данных, обнаружение значимых зависимостей и тенденций в результате анализа имеющейся информации, выявления отношений между данными различного типа. Применение различных методов выделения, извлечения и группировки данных, которые позволяют выявить систематизированные структуры данных и вывести из них правила для принятия решений и прогнозирования их последствий (регрессионный, дисперсионный, кластерный, дискриминантный, факторный анализы). Возможности графического представления информации в программе RStudio: графические функции отображения одномерных и многомерных данных, графический вывод с использованием графических параметров.

4 Образовательные технологии

Для проведения учебных занятий по дисциплине используются различные образовательные технологии. Для организации учебного процесса может быть использовано электронное обучение и (или) дистанционные образовательные технологии.

5 Оценка планируемых результатов обучения

5.1 Система оценивания

Форма контроля	Макс. количество баллов	
	За одну	Всего

	работу	
Текущий контроль: - защита отчета по практической работе	12 баллов	60 баллов
Промежуточная аттестация: зачет с оценкой		40 баллов
Итого за семестр		100 баллов

Полученный совокупный результат конвертируется в традиционную шкалу оценок и в шкалу оценок Европейской системы переноса и накопления кредитов (European Credit Transfer System; далее – ECTS) в соответствии с таблицей:

100-балльная шкала	Традиционная шкала		Шкала ECTS
95 – 100	отлично	зачтено	A
83 – 94			B
68 – 82	хорошо		C
56 – 67	удовлетворительно		D
50 – 55			E
20 – 49	неудовлетворительно	не зачтено	FX
0 – 19			F

5.2 Критерии выставления оценки по дисциплине

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
100-83/ A,B	«отлично»/ «зачтено (отлично)»/ «зачтено»	Выставляется обучающемуся, если он глубоко и прочно усвоил теоретический и практический материал, может продемонстрировать это на занятиях и в ходе промежуточной аттестации. Обучающийся исчерпывающе и логически стройно излагает учебный материал, умеет увязывать теорию с практикой, справляется с решением задач профессиональной направленности высокого уровня сложности, правильно обосновывает принятые решения. Свободно ориентируется в учебной и профессиональной литературе. Оценка по дисциплине выставляется обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «высокий».
82-68/ C	«хорошо»/ «зачтено (хорошо)»/ «зачтено»	Выставляется обучающемуся, если он знает теоретический и практический материал, грамотно и по существу излагает его на занятиях и в ходе промежуточной аттестации, не допуская существенных неточностей. Обучающийся правильно применяет теоретические положения при решении практических задач профессиональной направленности разного уровня

Баллы/ Шкала ECTS	Оценка по дисциплине	Критерии оценки результатов обучения по дисциплине
		сложности, владеет необходимыми для этого навыками и приёмами. Достаточно хорошо ориентируется в учебной и профессиональной литературе. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «хороший».
67-50/ D,E	«удовлетворительно»/ «зачтено (удовлетворительно)»/ «зачтено»	Выставляется обучающемуся, если он знает на базовом уровне теоретический и практический материал, допускает отдельные ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает определённые затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, владеет необходимыми для этого базовыми навыками и приёмами. Демонстрирует достаточный уровень знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции, закреплённые за дисциплиной, сформированы на уровне – «достаточный».
49-0/ F,FX	«неудовлетворительно»/ не зачтено	Выставляется обучающемуся, если он не знает на базовом уровне теоретический и практический материал, допускает грубые ошибки при его изложении на занятиях и в ходе промежуточной аттестации. Обучающийся испытывает серьёзные затруднения в применении теоретических положений при решении практических задач профессиональной направленности стандартного уровня сложности, не владеет необходимыми для этого навыками и приёмами. Демонстрирует фрагментарные знания учебной литературы по дисциплине. Оценка по дисциплине выставляются обучающемуся с учётом результатов текущей и промежуточной аттестации. Компетенции на уровне «достаточный», закреплённые за дисциплиной, не сформированы.

Критерии оценивания практических работ:

Критерии оценивания / Уровень требований к обучающемуся	Макс. кол-во баллов
---	---------------------

Текущий контроль, всего в т.ч.:	60
Практическая работа	12
Задания выполнены не полностью и (или) допущены две и более ошибки или три и более недочета	1-6
Задания выполнены полностью, но допущены два-три недочета, в т. ч. при ответе на контрольные вопросы	7-9
Задания выполнены полностью, возможна одна неточность, ответы на контрольные вопросы правильные	10-12

5.3 Оценочные средства (материалы) для текущего контроля успеваемости, промежуточной аттестации обучающихся по дисциплине

Примеры тестов

Данные – это:

- a) факты, характеризующие объекты, процессы, явления предметной области
- b) данные, рассматриваемые в каком-либо контексте, из которого пользователь может составить собственное мнение
- c) закономерности проблемной области, полученные в результате практической деятельности и профессионального опыта, позволяющие специалистам ставить и решать задачи в этой области
- d) сведения, передаваемые людьми устным, письменным или другим способом

Каким признаком не обладают большие данные?

- a) многообразие данных;
- b) достоверность данных;
- c) скорость накопления данных;
- d) типизацией исходных данных.

Слабо структурированные данные могут быть записаны в форме

- a) таблиц;
- b) отдельных векторов (строк);
- c) записей произвольной последовательности;
- d) таблиц и отношений.

Неструктурированные данные могут быть записаны в форме

- a) таблиц;
- b) записей произвольной последовательности;
- c) таблиц и отношений;
- d) записей произвольной длины.

Data Mining включает в себя

- a) один базовый метод обнаружения знаний;
- b) только статистические методы обработки данных для извлечения знаний;
- c) большое число различных методов извлечения знаний;
- d) большое число статистических методов извлечения знаний.

Целью задачи классификации можно назвать

- a) числовой зависимой переменной, основываясь на выборке непрерывных и/или категориальных переменных;

- b) объединение объектов или наблюдений, на основе близости значений их атрибутов (признаков);
- c) исследование взаимной связи между объектами и/или событиями;
- d) разбиения множества объектов или наблюдений на априорно заданные группы.

Целью задачи кластеризации можно назвать

- a) числовой зависимой переменной, основываясь на выборке непрерывных и/или категориальных переменных;
- b) объединение объектов или наблюдений, на основе близости значений их атрибутов (признаков);
- c) исследование взаимной связи между объектами и/или событиями;
- d) разбиения множества объектов или наблюдений на априорно заданные группы.

Целью формирования ассоциативных правил можно назвать

- a) числовой зависимой переменной, основываясь на выборке непрерывных и/или категориальных переменных;
- b) объединение объектов или наблюдений, на основе близости значений их атрибутов (признаков);
- c) установление взаимной связи между объектами и/или событиями;
- d) разбиения множества объектов или наблюдений на априорно заданные группы.

К классу прогнозирующих задач Data Mining относится

- a) кластеризация;
- b) поиск ассоциативных правил;
- c) регрессия;
- d) классификация.

Задания для практических работ представлены в разделе 9.

Вопросы для подготовки к зачету с оценкой

1. Определите сущность понятия «большие данные».
2. Опишите методики анализа больших данных.
3. Процесс аналитики анализа больших данных.
4. Дайте характеристику Big Data на мировом рынке.
5. Охарактеризуйте Big Data в России.
6. Определите понятие Data Mining.
7. Вопросы безопасности больших данных.
8. В чем состоит когнитивный анализ данных.
9. Основные описательные статистики.
10. Определите различия между параметрическими, непараметрическими и номинальными методами.
11. Опишите основную идею корреляционного анализа.
12. Регрессионный анализ.
13. Основная идея дисперсионного анализа.
14. Сущность кластерного анализа.
15. Дискриминантный анализ: модель и общая процедура выполнения.
16. Цели факторного анализа.
17. Программные средства анализа данных: Statistica, SPSS, Excel; их преимущества и недостатки.
18. Преимущества работа с данными в программе R-Studio.
19. Представление исходных данных в программе R-Studio.

20. Выполнение анализа данных в R-Studio.

6 Учебно-методическое и информационное обеспечение дисциплины**6.1 Список литературы****Основная**

1. Анализ данных : учебник для вузов / под редакцией В. С. Мхитаряна. — Москва : Издательство Юрайт, 2025. — 448 с. — (Высшее образование). — ISBN 978-5-534-19964-2. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://www.urait.ru/bcode/560311> (дата обращения: 23.05.2025).
2. Зарова, Е. В. Методы Data mining в обработке и анализе статистических данных (решения в R) : монография / Е.В. Зарова. — Москва : ИНФРА-М, 2021. — 232 с. : ил. - ISBN 978-5-16-016814-2. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/1240276> (дата обращения: 23.05.2025). – Режим доступа: по подписке.

Дополнительная

1. Рабчевский, А. Н. Синтетические данные и развитие нейросетевых технологий : учебник для вузов / А. Н. Рабчевский. — Москва : Издательство Юрайт, 2025. — 187 с. — (Высшее образование). — ISBN 978-5-534-17716-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://www.urait.ru/bcode/568661> (дата обращения: 23.05.2025).
2. Станкевич, Л. А. Интеллектуальные системы и технологии : учебник и практикум для вузов / Л. А. Станкевич. — 2-е изд., перераб. и доп. — Москва : Издательство Юрайт, 2025. — 478 с. — (Высшее образование). — ISBN 978-5-534-20363-9. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://www.urait.ru/bcode/560754> (дата обращения: 23.05.2025).
3. Кудрявцев, В. Б. Интеллектуальные системы : учебник и практикум для вузов / В. Б. Кудрявцев, Э. Э. Гасанов, А. С. Подколзин. — 2-е изд., испр. и доп. — Москва : Издательство Юрайт, 2025. — 165 с. — (Высшее образование). — ISBN 978-5-534-07779-7. — Текст : электронный // Образовательная платформа Юрайт [сайт]. — URL: <https://www.urait.ru/bcode/561954> (дата обращения: 23.05.2025).
4. Криволапов, С. Я. Анализ данных. Методы теории вероятностей и математической статистики на языке Python : учебное пособие / С.Я. Криволапов. — Москва : ИНФРА-М, 2025. — 678 с. — (Высшее образование). — DOI 10.12737/2034420. - ISBN 978-5-16-018616-0. - Текст : электронный. - URL: <https://znanium.ru/catalog/product/2034420> (дата обращения: 23.05.2025). – Режим доступа: по подписке.

6.2 Перечень ресурсов информационно-коммуникационной среды «Интернет»

1. Форум программистов и сисадминов Киберфорум <http://www.cyberforum.ru/>
2. Клуб программистов <https://programmersforum.ru/>
3. Форум программистов <https://programmersforum.ru/>
4. RStudio - свободная среда разработки программного обеспечения с открытым исходным кодом для языка программирования R.

6.3 Профессиональные базы данных и информационно-справочные системы
Доступ к профессиональным базам данных: <https://liber.rsuh.ru/ru/bases>

Информационные справочные системы:

1. Консультант Плюс
2. Гарант

7 Материально-техническое обеспечение дисциплины

Для материально-технического обеспечения дисциплины необходимы:

- для лекций:

- учебная аудитория,
- доска,
- проектор (стационарный или переносной),
- компьютер или ноутбук,
- программное обеспечение (ПО).

- для практических занятий:

- лаборатория,
- доска,
- проектор (стационарный или переносной),
- компьютер или ноутбук для преподавателя,
- компьютеры для обучающихся,
- выход в Интернет,
- программное обеспечение (ПО).

Перечень программного обеспечения (ПО)

- для лекций:

№п/п	Наименование ПО	Способ распространения
1	Microsoft Office 2010 Pro	лицензионное
2	Windows 10	лицензионное
3	Kaspersky Endpoint Security	лицензионное

- для практических занятий:

Наименование ПО	Способ распространения
Windows 10	лицензионное
Microsoft Office 2010 Pro	лицензионное
Mozilla Firefox	свободно распространяемое
Kaspersky Endpoint Security	лицензионное
Microsoft SQL Server 2008	лицензионное
Microsoft Visual Professional 2019	лицензионное

8 Обеспечение образовательного процесса для лиц с ограниченными возможностями здоровья и инвалидов

В ходе реализации дисциплины используются следующие дополнительные методы обучения, текущего контроля успеваемости и промежуточной аттестации обучающихся в зависимости от их индивидуальных особенностей:

- для слепых и слабовидящих:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением, или могут быть заменены устным ответом;
 - обеспечивается индивидуальное равномерное освещение не менее 300 люкс;

- для выполнения задания при необходимости предоставляется увеличивающее устройство; возможно также использование собственных увеличивающих устройств;
- письменные задания оформляются увеличенным шрифтом;
- экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

- для глухих и слабослышащих:
 - лекции оформляются в виде электронного документа, либо предоставляется звукоусиливающая аппаратура индивидуального пользования;
 - письменные задания выполняются на компьютере в письменной форме;
 - экзамен и зачёт проводятся в письменной форме на компьютере; возможно проведение в форме тестирования.
- для лиц с нарушениями опорно-двигательного аппарата:
 - лекции оформляются в виде электронного документа, доступного с помощью компьютера со специализированным программным обеспечением;
 - письменные задания выполняются на компьютере со специализированным программным обеспечением;
 - экзамен и зачёт проводятся в устной форме или выполняются в письменной форме на компьютере.

При необходимости предусматривается увеличение времени для подготовки ответа.

Процедура проведения промежуточной аттестации для обучающихся устанавливается с учётом их индивидуальных психофизических особенностей. Промежуточная аттестация может проводиться в несколько этапов.

При проведении процедуры оценивания результатов обучения предусматривается использование технических средств, необходимых в связи с индивидуальными особенностями обучающихся. Эти средства могут быть предоставлены университетом, или могут использоваться собственные технические средства.

Проведение процедуры оценивания результатов обучения допускается с использованием дистанционных образовательных технологий.

Обеспечивается доступ к информационным и библиографическим ресурсам в сети Интернет для каждого обучающегося в формах, адаптированных к ограничениям их здоровья и восприятия информации:

- для слепых и слабовидящих:
 - в печатной форме увеличенным шрифтом;
 - в форме электронного документа;
 - в форме аудиофайла.
- для глухих и слабослышащих:
 - в печатной форме;
 - в форме электронного документа.
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - в печатной форме;
 - в форме электронного документа;
 - в форме аудиофайла.

Учебные аудитории для всех видов контактной и самостоятельной работы, научная библиотека и иные помещения для обучения оснащены специальным оборудованием и учебными местами с техническими средствами обучения:

- для слепых и слабовидящих:
 - устройством для сканирования и чтения с камерой SARA CE;
 - дисплеем Брайля PAC Mate 20;
 - принтером Брайля EmBraille ViewPlus;
- для глухих и слабослышащих:
 - автоматизированным рабочим местом для людей с нарушением слуха и слабослышащих;

- акустический усилитель и колонки;
- для обучающихся с нарушениями опорно-двигательного аппарата:
 - передвижными, регулируемые эргономическими партами СИ-1;
 - компьютерной техникой со специальным программным обеспечением.

9 Методические материалы

9.1 Планы практических занятий

Практическая работа №1 Сбор данных из различных источников в сети Интернет.

Вопросы для защиты практической работы №1:

1. Портал открытых данных РФ.
2. Хранение данных в программе MS Excel.
3. Первичный анализ данных и визуализация данных в программе Excel.
4. Microsoft PowerBI – назначение, сфера применения, основные функциональные возможности.

Практическая работа №2 Загрузка и преобразование данных в Microsoft PowerBI.

Вопросы для защиты практической работы №2:

Формирование исходных данных.

Упрощение структуры данных.

Вычисление и изменение типов данных столбцов.

Объединение нескольких таблиц в одну.

Профилирование данных.

Загрузка данных в Microsoft PowerBI.

Создание модели данных в Microsoft PowerBI.

Настройка связей между таблицами.

Практическая работа №3 Знакомство с программой RStudio.

Вопросы для защиты практической работы №3:

1. Синтаксис.
2. Представление исходных данных в программе RStudio (векторы, массивы, матрицы, списки, таблицы).
3. Выборка и преобразование исходных данных в программе RStudio. Удаление пропущенных значений

Практическая работа №4 Статистическая обработка данных в RStudio

Вопросы для защиты практической работы №4:

1. Подсчет описательных статистик.
2. Законы распределения вероятностей, реализованные в R.
3. Проверка данных на нормальность распределения (критерии Шапиро-Уилка, Колмогорова-Смирнова и др.).
4. Уровень статистической достоверности.
5. Корреляционный анализ.
6. Расчет коэффициентов корреляции Пирсона, Спирмена, Кендалла.

Практическая работа №5 Возможности графического представления информации в программе RStudio.

Вопросы для защиты практической работы №5:

1. Графические функции отображения одномерных и многомерных данных.
2. Графический вывод с использованием графических параметров.

9.2 Методические рекомендации по подготовке письменных работ

Отчет по проделанной работе должен быть изложен с соблюдением правил грамматики русского и английского языков (в случаях необходимости). При этом отражаемые результаты работы должны быть информативными, тезисного порядка. В отчет входят следующие обязательные разделы:

1. Титульный лист с полным указанием ведомственной принадлежности, названия ВУЗа, института, факультета, кафедры. Кроме того, полное точное название лабораторной работы, Ф.И.О. студента, подготовившего отчет о результатах проделанной работы и Ф.И.О., должность, название кафедры преподавателя осуществляющего проверку и оценивание полученных результатов.
2. Содержание.
3. Введение.
4. Цели и задачи практической работы.
5. Методы и технологии, применяемые для решения поставленных задач оформленные в виде отдельных этапов работы.
6. Выводы по работе.
7. Приложения.

Оформление отчета выполняется с использованием компьютерной верстки LaTeX. Отчет сохраняется и представляет для проверки в виде отдельного pdf файла. В имени файла показывается фамилия студента и номер выполненной работы.

АННОТАЦИЯ ДИСЦИПЛИНЫ

Цель дисциплины – формирование знаний, умений, навыков в области систем и технологий обработки больших объемов структурированной или неструктурированной информации.

Задачи дисциплины:

1. Изучение технологий хранения, обработки и анализа больших данных.
2. Формирование умений рационального использования систем и технологий обработки больших данных.
3. Применение статистических и математических методов для анализа больших объемов информации.

В результате освоения дисциплины обучающийся должен:

Знать:

- методы решения задач обработки и анализа больших данных, возможности высокопроизводительных вычислительных систем, технологии распределенных вычислений, методы и модели Data Mining;
- современные информационные технологии и программные средства в области анализа и обработки больших данных.

Уметь:

- использовать современные информационные технологии и программные средства в области анализа и обработки больших данных для решения практических задач в реальной предметной области;
- создавать алгоритмы анализа и обработки большого объема данных с применением моделей Data Mining.

Владеть:

- основными методами Data Mining для обнаружения в данных ранее неизвестных, практически полезных и доступных интерпретаций знаний;
- методами статистического и математического анализа больших данных.

Рабочей программой предусмотрены следующие виды контроля: текущий контроль успеваемости в форме контрольной работы, промежуточная аттестация в форме зачета с оценкой.

Общая трудоемкость освоения дисциплины составляет 3 зачетные единицы.