

Министерство науки и высшего образования  
Российской Федерации

Федеральное государственное бюджетное  
образовательное учреждение высшего образования  
«Донецкий государственный университет»

Учетно-финансовый факультет  
Кафедра бизнес-информатики

УТВЕРЖДАЮ  
проректор

\_\_\_\_\_ П.А. Машаров  
«17» апреля 2025 г.  
МП

## **РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

### **АНАЛИЗ БОЛЬШИХ ДАННЫХ**

Укрупненная группа направлений подготовки	38.00.00 Экономика и управление
Программа высшего образования	Программа бакалавриата
Направление подготовки	38.03.05 Бизнес-информатика
Направленность (профиль) образовательной программы	Аналитика и управление данными
Квалификация	Бакалавр
Форма обучения	Очная

Рабочая программа адаптирована для лиц  
с ограниченными возможностями здоровья и инвалидов

Донецк 2025

Рабочая программа дисциплины **«Анализ больших данных»** для обучающихся по направлению подготовки 38.03.05 Бизнес-информатика (Профиль: Аналитика и управление данными) составлена на основании Федерального государственного образовательного стандарта высшего образования – бакалавриат по направлению подготовки 38.03.05 Бизнес-информатика, утвержденного приказом Министерства науки и высшего образования Российской Федерации от «29» июля 2020 г. № 838 (с изм. и доп.), Порядка организации и осуществления образовательной деятельности по образовательным программам высшего образования – программам бакалавриата, программам специалитета, программам магистратуры, утвержденного приказом Министерства науки и высшего образования Российской Федерации от 06 апреля 2021 г. № 245 (с изм. и доп.), в соответствии с учебным планом, утвержденным Ученым советом ФГБОУ ВО «ДонГУ» для набора 2024 года.

Разработчик:

доцент кафедры бизнес-информатики,  
канд. экон. наук

В.В. Гридина

Рабочая программа одобрена на заседании кафедры бизнес-информатики.  
Протокол от 10.04.2025 г. №8а

Заведующий кафедрой

Т.О. Загорная

СОГЛАСОВАНО:

Декан учетно-финансового факультета  
16.04.2025 г.

Н. В. Алексеенко

Учебно-методическая комиссия учетно-финансового факультета.  
Протокол от 15.04.2025 г. № 6.  
Председатель

А. А. Блажевич

Руководитель основной профессиональной  
образовательной программы,  
д-р экон. наук, проф.  
10.04.2025 г.

Т.О. Загорная

## 1. МЕСТО ДИСЦИПЛИНЫ В СТРУКТУРЕ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ

1.1. Требования к предварительной подготовке обучающихся, предшествующие и сопутствующие дисциплины, на которых основывается изучение данной:

Информационные технологии и инструменты программирования, Программирование на Python, Теория систем и системный анализ, Анализ данных на Python, Имитационное моделирование.

1.2. Дисциплины, курсовые работы и практики, для которых освоение данной дисциплины необходимо как предшествующее:

Производственная практика: по получению профессиональных умений и опыта профессиональной деятельности, Производственная практика: преддипломная.

## 2. ОПИСАНИЕ ДИСЦИПЛИНЫ

### 2.1. Общая характеристика

Наименование показателя	Значение показателя
Название образовательной программы	38.03.05 Бизнес-информатика (Профиль: Аналитика и управление данными)
Шифр и название в соответствии с учебным планом	Б1.В.ДВ.7.2 Анализ больших данных
Часть образовательной программы	Вариативная часть: выбор обучающихся
Количество зачетных единиц / всего часов	3 / 108

В случае предъявления от обучающегося или его родителя (законного представителя) заявления на обучение по адаптированной образовательной программе высшего образования, подкрепленного заключением психолого-медико-педагогической комиссии (ПМПК) или медико-социальной экспертизы (МСЭ) с рекомендациями создания индивидуальной программы реабилитации и абилитации (ИПРА), данная рабочая программа может быть адаптирована с учетом индивидуальных особенностей здоровья обучающегося.

### 2.2. Распределение часов по формам и периодам обучения

Форма обучения	курс	семестр	Общее количество часов					Форма контроля
			лекционных	лабораторных	практических	самостоятельной работы	всего	
Очная	4	8	20	30	0	58	108	экзамен

## 3. ЦЕЛИ ДИСЦИПЛИНЫ

Сформировать у студентов теоретические знания и практические навыки по методологии сбора, анализа и моделирования больших данных для принятия управленческих решений.

**4. КОМПЕТЕНЦИИ ОБУЧАЮЩЕГОСЯ, ФОРМИРУЕМЫЕ В РЕЗУЛЬТАТЕ  
ОСВОЕНИЯ КОМПОНЕНТА ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЫ, ИХ ИНДИКАТОРЫ  
И ПЛАНИРУЕМЫЕ РЕЗУЛЬТАТЫ ОБУЧЕНИЯ**

Компетенции	Индикаторы	Результаты обучения
ПК- 14. Способен выявлять естественнонаучную сущность проблем, возникающих в ходе профессиональной деятельности в области моделирования и анализа сложных естественных и искусственных систем.	ПК-14.1. Применяет интеллектуальные программные решения для анализа больших данных.	ПК-14.1.1. Знает основные подходы к применению программных средств для моделирования и обработки больших данных. ПК-14.1.2. Умеет осуществлять автоматизированный сбор и алгоритмическую обработку больших данных в профессиональной области с использованием современных программных средств. ПК-14.1.3. Владеет методологией сбора и анализа данных на основе методов и программного инструментария.
	ПК-14.2. Использует анализ и обработку больших данных для решения проблем, возникающих в ходе профессиональной деятельности.	ПК-14.2.1 Знает методы планирования и организации исследований больших данных, а также методологию принятия решений на основе анализа больших данных. ПК-14.2.2 Умеет организовать исследование больших данных реального объекта, осуществить сбор и интеграцию больших данных объекта исследования. ПК-14.2.3. Владеет современными методами, подходами, алгоритмами, средствами и инструментами анализа больших данных.

**5. ПРОГРАММА ДИСЦИПЛИНЫ**

Название темы	Краткое содержание темы (вопросы темы)
<b>Раздел 1. Определение больших данных. Технологии хранения больших данных.</b>	
Введение в большие данные	История и причины появления термина Big Data. Характеристики и источники Big Data. Четыре основных типа данных. Аналитика данных. Задачи, решаемые Big Data.
Классификация задач анализа данных	DataMining. Интеллектуальный анализ данных, его, отличия и задачи. Text Mining. Web Mining. Web Content Mining. Web Usage Mining. Social media mining. RapidMiner.
Жизненный цикл аналитики данных	Business Intelligence (BI). ETL (Extract, Transform, Load)-процесс. Средства BI. Online Analytical Processing (OLAP). Инструменты анализа BI. Понятие жизненного цикла аналитики данных.
Высокопроизводительные вычисления	История Hadoop и MapReduce. Hadoop Distributed File System. Технология Map Reduce. Архитектура Hadoop.

Масштабирование и многоуровневое хранение данных	NoSQL (NotOnlySQL). Масштабируемость. Репликация. CAP теорема. MongoDB.
Раздел 2. Инструменты анализа и обработки больших данных.	
Визуализация данных и результатов анализа	Типы, задачи и виды визуализации. Графики, диаграммы, инфографика. Интерактивный сторителлинг, дашборды.
Язык R и готовые инструментальные средства анализа больших данных	Язык R. Синтаксис языка R, основные типы и структуры данных. Интерфейс Rstudio.
Статистические методы анализа данных	Статистические гипотезы и критерии. Машинное обучение. Метрический и линейных классификаторы ROC–кривая. Кластерный анализ.

## 6. СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

### 6.1. Форма обучения – очная, курс – 4, семестр – 8

Наименования разделов и тем	Количество часов				
	Лекц.	Лабор.	Практ.	СРС	Всего
Раздел 1. Определение больших данных. Технологии хранения больших данных.	<b>12</b>	<b>12</b>	<b>0</b>	<b>35</b>	<b>59</b>
Введение в большие данные	2	2	0	7	11
Классификация задач анализа данных	4	4	0	7	15
Жизненный цикл аналитики данных	2	2	0	7	11
Высокопроизводительные вычисления	2	2	0	7	11
Масштабирование и многоуровневое хранение данных	2	2	0	7	11
Раздел 2. Инструменты анализа и обработки больших данных.	<b>8</b>	<b>18</b>	<b>0</b>	<b>23</b>	<b>49</b>
Визуализация данных и результатов анализа	2	4	0	7	13
Язык R и готовые инструментальные средства анализа больших данных	4	8	0	8	20
Статистические методы анализа данных	2	6	0	8	16
<b>ИТОГО ЗА СЕМЕСТР</b>	<b>20</b>	<b>30</b>	<b>0</b>	<b>58</b>	<b>108</b>

## 7. ОЦЕНОЧНЫЕ МАТЕРИАЛЫ (СРЕДСТВА) ДЛЯ ТЕКУЩЕГО КОНТРОЛЯ УСПЕВАЕМОСТИ, ПРОМЕЖУТОЧНОЙ АТТЕСТАЦИИ

### 7.1. Контрольные вопросы

Раздел 1. Определение больших данных. Технологии хранения больших данных.

1. История и причины появления термина Big Data.
2. Характеристики и источники Big Data.
3. Четыре основных типа данных.
4. Аналитика данных.
5. Задачи, решаемые Big Data.
6. DataMining.
7. Интеллектуальный анализ данных, его, отличия и задачи.
8. Text Mining.
9. Web Mining.
10. Web Content Mining.
11. Web Usage Mining.
12. Social media mining.

13. RapidMiner.
  14. Business Intelligence (BI).
  15. ETL (Extract, Transform, Load)-процесс.
  16. Средства BI.
  17. Online Analytical Processing (OLAP).
  18. Инструменты анализа BI.
  19. Понятие жизненного цикла аналитики данных.
  20. История Hadoop и MapReduce.
  21. Hadoop Distributed File System.
  22. Технология Map Reduce.
  23. Архитектура Hadoop.
  24. NoSQL (NotOnlySQL).
  25. Масштабируемость.
  26. Репликация.
  27. CAP теорема.
  28. MongoDB.
- Раздел 2. Инструменты анализа и обработки больших данных.
29. Типы, задачи и виды визуализации.
  30. Графики, диаграммы, инфографика.
  31. Интерактивный сторителлинг, дашборды.
  32. Язык R.
  33. Синтаксис языка R, основные типы и структуры данных.
  34. Интерфейс Rstudio.
  35. Статистические гипотезы и критерии.
  36. Машинное обучение.
  37. Метрический и линейных классификаторы ROC-кривая.
  38. Кластерный анализ.

## 7.2. Образец содержания экзаменационного билета.

### Экзаменационный билет № 1

1. Четыре основных типа данных.
2. Графики, диаграммы, инфографика.

В случае ведения учебного процесса с использованием электронного обучения и дистанционных образовательных технологий, содержание билета может отличаться от приведенного.

## 8. РАСПРЕДЕЛЕНИЕ БАЛЛОВ, КОТОРЫЕ ПОЛУЧАЮТ ОБУЧАЮЩИЕСЯ

Общая оценка знаний обучающихся по дисциплине проводится по 100-балльной шкале исходя из максимума, приведенного в таблице ниже.

Организационно-учебная работа в аудитории оценивается на основе таких критериев как посещаемость занятий, своевременное и качественное выполнение домашних заданий, активность во время проведения лекционных и лабораторных занятий (участие в обсуждении текущего и пройденного материала, решение задач и т.п.).

Самостоятельная работа оценивается на основе предоставленных на проверку выполненных домашних, индивидуальных заданий с учетом своевременности их предоставления и соответствия требованиям к их выполнению.

Количество баллов за контрольную работу вычисляется как сумма баллов за все входящие в её состав задания. Каждое задание оценивается исходя из максимально возможного количества баллов с учетом правильности выполнения задания, полноты приводимых обоснований.

По результатам работы в семестре обучающийся, набравший не менее 60 баллов, имеет право получить оценку. Те, кто претендует на более высокий балл, проходят промежуточную аттестацию. Максимальное количество баллов на промежуточной аттестации – 100. Общее количество баллов за семестр вычисляется как максимальная из полученных за семестр и на промежуточной аттестации и выставляется согласно принятому порядку.

#### 8.1. Форма обучения – очная.

Номера разделов	Виды работ	Максимальное количество баллов
1-2	Организационно-учебная работа в аудитории	15
	Самостоятельная работа	15
ИТОГО		60
Экзамен		40
Общий итог за семестр		100

#### Соответствие баллов оценке

Количество баллов из 100	ECTS	Оценка по пятибалльной шкале	
		Экзамен, дифференцированный зачет	Зачет
90-100	A	отлично	зачтено
80-89	B	хорошо	зачтено
75-79	C		зачтено
70-74	D	удовлетворительно	зачтено
60-69	E		зачтено
35-59	FX	неудовлетворительно	не зачтено
0-34	F		не зачтено

### 9. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ УЧЕБНОГО ПРОЦЕССА

Учебные занятия проводятся в 8-м учебном корпусе ДонГУ (г. Донецк, ул. Челюскинцев, д. 198 а). Для проведения лабораторных занятий требуется аудитория, оборудованная меловой или маркерной доской, мультимедийный проектор и экран, ноутбук, комплект учебной мебели для студентов, рабочее место преподавателя, выход в Интернет – проводной или с использованием Wi-Fi.

Для самостоятельной работы используются текстовые и электронные ресурсы Научной библиотеки университета и других электронных библиотечных баз данных, учебно-методическое обеспечение, представленное в учебно-методическом кабинете 8-го учебного корпуса (ауд. 105), материально-техническую базу учебных лабораторий кафедры бизнес-информатики (ауд. 101-103).

Обучающиеся имеют возможность использовать учебные материалы по дисциплине, размещенные на платформе Moodle Центра дистанционного образования ФГБОУ ВО «ДонГУ». При изучении дисциплины применяются электронное обучение и дистанционные образовательные технологии.

С использованием ресурсов платформы дистанционного образования осуществляется текущий контроль знаний обучающихся на основе тестирования и проверки результатов самостоятельной работы.

## 10. РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

### 10.1. Основная литература

1. Федин Ф.О. Анализ данных. Часть 2. Инструменты Data Mining [Электронный ресурс]: учебное пособие/ Федин Ф.О., Федин Ф.Ф.— Электрон. текстовые данные. — Москва: Московский городской педагогический университет, 2017.— 308 с.— URL: <https://1lib.domains/?redirectUrl=/book/3590439/181f68>
2. Воронов В.И. Data Mining - технологии обработки больших данных [Электронный ресурс]: учебное пособие/ Воронов В.И., Воронова Л.И., Усачев В.А.— Электрон. текстовые данные. — Москва: Московский технический университет связи и информатики, 2018.— 47 с. — URL: <https://knigogid.ru/books/1888855-big-data-metody-i-sredstva-analiza-uchebnoe-posobie/download/pdf>

### 10.2. Дополнительная литература

3. Крутиков, В. Н. Анализ данных [Электронный ресурс] : учебное пособие по направлению подготовки 010400.62 «Прикладная математика и информатика» / В. Н. Крутиков, В. В. Мешечкин ; Кемеровский государственный университет. - Кемерово : Кемеровский государственный университет, 2014. - 137 с.
4. Радченко, И.А. Технологии и инфраструктура Big Data [Электронный ресурс] / И.А. Радченко, И.Н. Николаев. – СПб: Университет ИТМО, 2018. – 52 с. – URL: <https://books.ifmo.ru/file/pdf/2326.pdf>
5. Мастицкий С.Э. Статистический анализ и визуализация данных с помощью R / С.Э. Мастицкий, В.К. Шитиков. – URL: [https://r-analytics.blogspot.com/p/blog-page\\_46.html](https://r-analytics.blogspot.com/p/blog-page_46.html)
6. Мастицкий С.Э. Анализ временных рядов с помощью R / С.Э. Мастицкий. – URL: [https://r-analytics.blogspot.com/p/blog-page\\_46.html](https://r-analytics.blogspot.com/p/blog-page_46.html)

## 11. ИНФОРМАЦИОННЫЕ РЕСУРСЫ

1. **Национальная электронная библиотека (НЭБ):** федеральная государственная информационная система / Министерство Культуры РФ; Российская государственная библиотека. – Москва, 2019- . – URL: <https://rusneb.ru/> (дата обращения: 01.09.2023). – Режим доступа: свободный, подписка. Необходима установка программного обеспечения. – Текст: электронный.
2. **eLIBRARY.RU:** научная электронная библиотека: сайт. – Москва, 2000- . – URL: <https://elibrary.ru> (дата обращения: 01.09.2023). – Режим доступа: для авторизов. пользователей. –Текст: электронный.
3. Научная электронная библиотека **«КиберЛенинка»:** сайт / Ассоциация «Открытая наука». – Москва, 2014- . – URL: <https://cyberleninka.ru/>. – Режим доступа: свободный. – Текст: электронный.
4. Электронно-библиотечная система **«Лань»:** [сайт]. – URL: <https://e.lanbook.com> (дата обращения: 01.09.2023). – Режим доступа: для авторизов. пользователей. – Текст: электронный.
5. **ЭБС Юрайт:** электронная библиотечная система: сайт. – Москва, 2013. – URL: <https://biblio-online.ru> (дата обращения: 01.09.2023). – Режим доступа: для авторизов. пользователей. – Текст: электронный.
6. **Электронно-библиотечная система ДонГУ:** сайт / ФГБОУ ВО «ДонГУ». – Донецк, 2016- . – URL: <http://library.donnu.ru/> (дата обращения: 01.09.2023). – Режим доступа: свободный. – Текст: электронный.



7. **Электронный каталог** Научной библиотеки ДонГУ: раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://library.donnu.ru/catalog/> (дата обращения: 01.09.2023). – Режим доступа: поиск свободный, электронные документы – для пользователей ДонГУ.

8. **Электронный архив ДонГУ**: раздел сайта / НБ ДонГУ. – Текст: электронный // ЭБС ДонГУ: сайт. – URL: <http://repo.donnu.ru/> (дата обращения: 01.09.2023). – Режим доступа: свободный.

## 12. ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ

1. Windows 7 PRO (корпоративная лицензия ДонГУ № 46484614)
2. Microsoft Office (корпоративная лицензия ДонГУ № 46472919)
3. Microsoft Visual Studio (лицензия программы Dream Spark для высших учебных заведений)
4. Антивирус Касперского, Adobe Acrobat Reader, xPDF (лицензии GPL, Apache, BSD для свободного программного обеспечения).