

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ  
ДОНЕЦКОЙ НАРОДНОЙ РЕСПУБЛИКИ  
ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ  
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ  
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»



УТВЕРЖДАЮ:

Первый проректор

А.А. Каракозов

(подпись)

20 21 года

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
**Б1.В2 Аналитика больших данных для бизнеса**  
(код и наименование дисциплины согласно учебному плану)

Направление подготовки:

38.04.05 Бизнес-информатика

(код и наименование направления / специальности)

Магистерская программа:

IT инновации в бизнесе

(наименование профиля / магистерской программы / специализации)

Программа:

магистратура

(бакалавриат, магистратура, специалитет)

Форма обучения:

очная, заочная

(очная, заочная, очно-заочная)

| Форма обучения:                            | Очная   | Заочная |
|--------------------------------------------|---------|---------|
| Семестр(ы)                                 | 1       | 1       |
| Общая трудоёмкость в з.е./часах            | 3 (108) | 3 (108) |
| Контактная работа (час.), в том числе      | 53      | 16      |
| лекции (час.)                              | 17      | 4       |
| практические (семинарские) занятия (час.)  | -       | -       |
| лабораторные работы (час.)                 | 34      | 6       |
| Самостоятельная работа (час.), в том числе | 57      | 98      |
| Курсовой проект(работа) (семестр/час.)     | -       | -       |
| Индивидуальное задание (кол./час.)         | -       | 1/9     |
| Контроль (экзамен, час./зачёт)             | зачет   | зачет   |

Донецк, 2021 г.

Рабочая программа дисциплины «Аналитика больших данных для бизнеса» составлена в соответствии с учебными планами по направлению подготовки 38.04.05 «Бизнес-информатика», магистерская программа «IT инновации в бизнесе» для 2021 года приёма.

Составитель:

Профессор кафедры экономической кибернетики,  
доктор экономических наук, профессор

(подпись)

Тимохин В. Н.  
(Ф.И.О.)

Рабочая программа **рассмотрена и принята** на заседании кафедры экономической кибернетики.

Протокол от « 7 » 05 2021 года № 9

Заведующий кафедрой

(подпись)

Коломыцева А.О.  
(Ф.И.О.)

Рабочая программа **одобрена учебно-методической комиссией** ДОННТУ по направлению подготовки 38.04.05 «Бизнес-информатика»

Протокол от « 19 » 05 2021 года № 4

Председатель

(подпись)

Коломыцева А.О.  
(Ф.И.О.)

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры экономической кибернетики.

Протокол от « \_\_\_\_ » \_\_\_\_ 20\_\_ года № \_\_\_\_

Заведующий кафедрой

(подпись)

(Ф.И.О.)

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры экономической кибернетики.

Протокол от « \_\_\_\_ » \_\_\_\_ 20\_\_ года № \_\_\_\_

Заведующий кафедрой

(подпись)

(Ф.И.О.)

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры экономической кибернетики.

Протокол от « \_\_\_\_ » \_\_\_\_ 20\_\_ года № \_\_\_\_

Заведующий кафедрой

(подпись)

(Ф.И.О.)

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры экономической кибернетики.

Протокол от « \_\_\_\_ » \_\_\_\_ 20\_\_ года № \_\_\_\_

Заведующий кафедрой

(подпись)

(Ф.И.О.)

Рабочая программа **продлена** для 20\_\_ года приёма на заседании кафедры экономической кибернетики.

Протокол от « \_\_\_\_ » \_\_\_\_ 20\_\_ года № \_\_\_\_

Заведующий кафедрой

(подпись)

(Ф.И.О.)

## 1 ОБЪЕКТ, ЦЕЛЬ И ЗАДАЧИ ОСВОЕНИЯ ДИСЦИПЛИНЫ

Дисциплина «Аналитика больших данных для бизнеса» рассматривает методологию применения современных математических методов и производительных информационных технологий для сбора и анализа массивов первичных данных о деятельности субъектов в профессиональной сфере.

**Цель дисциплины** – формирование у студентов теоретических знаний и практических навыков по методологии сбора, анализа и моделирования больших данных для принятия управленческих решений.

**Учебные задачи дисциплины** – сформировать у студентов теоретические знания в области современной методологии и инструментальных средств сбора и обработки цифровой информации, моделирования и обработки больших данных, планирования и организации исследования больших данных; умения самостоятельной постановки задачи и организации сбора и анализа больших данных с использованием современных информационных технологий, разработки программных средств для анализа больших данных, интеграции больших данных объекта исследования для цели эффективного выполнения аналитических работ в ИТ-проекте; навыки практической деятельности в области поиска и использования эффективных стратегий для решения проблем сбора и анализа больших данных с использованием методов моделирования и математического анализа, эффективных цифровых средств и средств информационной безопасности, пакетов прикладных программ, применения методологии сбора и анализа данных на основе современных информационных технологий, моделирования на основе анализа данных реальных объектов, в том числе на основе разработки программных комплексов, принятия решений на основе анализа больших данных, представления и оформления отчётов в сфере науки и техники.

В результате освоения дисциплины студент должен

**Знать:** - основные методы системного подхода и критического анализа проблемных ситуаций (УК-1.1);

- этапы разработки стратегии действий и методы решения проблемных ситуаций (УК-1.2);

- современные методы, алгоритмы и средства сбора и обработки информации (РО 1-3 ПК-1);

- основные подходы к применению программных средств для моделирования и обработки больших данных (РО 2-3 ПК-1);

- методы планирования и организации исследований больших данных (РО 1-3 ПК-3);

- классификацию систем поддержки разработки и сопровождения требований к анализу больших данных (РО 1-3 ПК-6).

**Уметь:** - выявлять проблемные ситуации, используя методы системного подхода и критического анализа (УК-1.3);

- обосновывать выбор стратегии для достижения поставленной цели с учетом ограничений, рисков и моделируемых результатов (УК-1.4);

- анализировать проблемную ситуацию, выявлять и определять способы ее разрешения (УК-1.5);

- осуществлять сбор и обработку данных с использованием современных информационных технологий (РО 1-У ПК-1);
- разрабатывать программные средства для анализа больших данных (РО 2-У ПК-1);
- организовать исследование больших данных реального объекта (РО 1-У ПК-3);
- осуществить сбор и интеграцию больших данных объекта исследования для цели эффективного выполнения аналитических работ в ИТ-проекте (РО 2-У ПК-3);
- систематизировать информацию для решения задач управления инфраструктурой разработки и сопровождения требований к анализу больших данных (РО 1-У ПК-6).

**Владеть:** - использованием эффективных стратегий действий для решения проблемной ситуации с учетом оценки ограничений, рисков и моделируемых результатов (УК-1.6);

- использованием методов критического анализа и системного подхода в разработке стратегии действий для решения проблемных ситуаций (УК-1.7);
- аналитическими способностями и критическим мышлением (УК-1.8);
- методологией сбора и анализа данных на основе современных информационных технологий (РО 1-В ПК-1);
- методологией моделирования на основе анализа данных реальных объектов, в том числе на основе разработки программных комплексов (РО 2-В ПК-1);
- способностями по самостоятельной постановке задач исследования (РО 3-В ПК-1);
- методологией принятия решений на основе анализа больших данных (РО 1-В ПК-3);
- навыками разработки модели консолидирования данных из больших массивов информации (РО 1-В ПК-6).

Перечисленные результаты обучения являются основой для формирования следующих компетенций в соответствии с ФГОС ВО по данному направлению подготовки на основе определенных индикаторов их достижения:

**УК-1** - Способен осуществлять критический анализ проблемных ситуаций на основе системного подхода, вырабатывать стратегию действий.

**ПК-1** - Способен организовать исследования новых рынков информационно-коммуникационных технологий.

**ПК-3** - Способен осуществлять принятие решений в профессиональной деятельности на основе использования современных методов и программного инструментария сбора, обработки и анализа данных, в том числе больших данных.

**ПК-6** - Способен проектировать и совершенствовать архитектуру и ИТ-инфраструктуру предприятия в соответствии с потребностями развития бизнеса.

## 2 МЕСТО ДИСЦИПЛИНЫ В ОСНОВНОЙ ОБРАЗОВАТЕЛЬНОЙ ПРОГРАММЕ

Дисциплина «Аналитика больших данных для бизнеса» относится к части, формируемой участниками образовательных отношений, Блока 1.

Базируется на знаниях, умениях и навыках, которые студент приобрёл при освоении программы бакалавриата. Для успешного освоения дисциплины студент должен знать математические основы разработки моделей, условия анализа и сбора данных.

Знания и умения, приобретенные при освоении данной дисциплины, реализуются студентом при выполнении НИР, прохождении преддипломной практики, подготовке магистерской диссертации.

### 3 СТРУКТУРА И СОДЕРЖАНИЕ ДИСЦИПЛИНЫ

#### 3.1 Распределение учебных часов по темам дисциплины и видам занятий

| №<br>те<br>мы | Наименование тем<br>(содержательных модулей)                    | Количество часов (очная/заочная форма) |             |                         |        |         |
|---------------|-----------------------------------------------------------------|----------------------------------------|-------------|-------------------------|--------|---------|
|               |                                                                 | Всего                                  | В том числе |                         |        |         |
|               |                                                                 |                                        | Лекции      | Практ.<br>(Се-<br>мин.) | Лабор. | СР      |
| 1             | Тема 1. Введение в анализ дан-<br>ных                           | 10 (9)                                 | 2 (1)       |                         |        | 8 (8)   |
| 2             | Тема 2. Технологии анализа<br>данных                            | 10 (8)                                 | 2 (0)       |                         |        | 8 (8)   |
| 3             | Тема 3. Консолидация данных                                     | 24 (21)                                | 4 (1)       |                         | 10 (2) | 10 (18) |
| 4             | Тема 4. Трансформация данных                                    | 19 (18)                                | 3 (1)       |                         | 6 (1)  | 10 (16) |
| 5             | Тема 5. Визуализация данных                                     | 19 (18)                                | 2 (0)       |                         | 6 (1)  | 11 (17) |
| 6             | Тема 6. Оценка качества,<br>очистка и предобработка дан-<br>ных | 26 (25)                                | 4 (1)       |                         | 12 (2) | 10 (22) |
|               | Индивидуальное задание                                          | 0 (9)                                  | -           | -                       | -      | 0 (9)   |
|               | Курсовая работа (проект)                                        | -                                      | -           | -                       | -      |         |
|               | Итого по видам занятий                                          | 108 (108)                              | 17 (4)      |                         | 34 (6) | 57 (89) |
|               | Контроль                                                        | 0 (0)                                  |             |                         |        |         |
|               | ИТОГО                                                           | 108 (108)                              |             |                         |        |         |

#### Формирование компетенций в результате освоения тем дисциплины

| Компетенция | Темы дисциплины, нацеленные на формирование компетенций |
|-------------|---------------------------------------------------------|
| УК-1        | Тема 1, Тема 2, Тема 3                                  |
| ПК-1        | Тема 5                                                  |
| ПК-3        | Тема 2, Тема 4, Тема 5, Тема 6                          |
| ПК-6        | Тема 2, Тема 3, Тема 4, Тема 5, Тема 6                  |

#### 3.2 Лекции

Тема 1. Введение в анализ данных

Содержание темы 1:

Основные понятия анализа больших данных. Анализ специфики источников данных в предметной области. Цели, задачи и инструменты анализа данных

Литература к теме 1: [\[1\]](#), [\[3\]](#), [\[4\]](#)

## **Тема 2. Технологии анализа данных**

Содержание темы 2:

Аналитический и информационный подходы к моделированию. Технология KDD.

Литература к теме 2: [\[1\]](#), [\[2\]](#)

## **Тема 3. Консолидация данных**

Содержание темы 3:

Общая характеристика OLTP-систем. Предпосылки появления систем поддержки принятия решений. Понятие хранилища данных. Реляционные хранилища данных. Многомерные хранилища данных. Гибридные хранилища данных. Виртуальные хранилища данных. Процесс ETL. Извлечение данных в ETL. Преобразование данных в ETL. Загрузка данных в хранилище. Особенности загрузки данных из локальных источников. Обогащение данных.

Литература к теме 3: [\[1\]](#), [\[2\]](#), [\[4\]](#), [\[5\]](#)

## **Тема 4. Трансформация данных**

Содержание темы 4:

Понятие трансформации данных. Особенности трансформации временных рядов. Группировка и разгруппировка данных. Слияние данных. Квантование. Нормализация и кодирование данных.

Литература к теме 4: [\[1\]](#), [\[2\]](#)

## **Тема 5. Визуализация данных**

Содержание темы 5:

Сущность визуализации данных. Визуализаторы общего назначения. Визуализаторы OLAP-анализа. Визуализаторы, применяемые для оценки качества моделей.

Литература к теме 5: [\[1\]](#)

## **Тема 6. Оценка качества, очистка и предобработка данных**

Содержание темы 6:

Понятие и задачи оценки качества данных. Технологии и методы оценки качества данных. Очистка и предобработка данных. Фильтрация данных. Обработка дубликатов и противоречий. Выявление аномальных значений. Восстановление пропущенных значений. Основы сокращения размерности. Сокращение числа признаков. Сокращение числа значений признаков и записей. Сэмплинг.

Литература к теме 6: [\[1\]](#), [\[5\]](#)

### **3.3 Практические занятия учебным планом не предусмотрены.**

### **3.4 Лабораторные занятия**



| №<br>п/п | Тема занятия         | Объем часов (очн/ заочн) | Литература   |
|----------|----------------------|--------------------------|--------------|
| 1        | Консолидация данных  | 10 (2)                   | [1, 5, 8]    |
| 2        | Трансформация данных | 6 (1)                    | [1, 2, 8]    |
| 3        | Визуализация данных  | 6 (1)                    | [1, 8]       |
| 4        | Улучшение данных     | 12 (2)                   | [1, 2, 5, 8] |
| Итого:   |                      | 34(6)                    |              |

### 3.5 Самостоятельная работа студента

| №<br>п/п | Виды самостоятельной работы студента         | Объем, час. |
|----------|----------------------------------------------|-------------|
| 1        | Изучение лекционного материала               | 27 (49)     |
| 2        | Подготовка к лабораторным занятиям           | 30 (40)     |
| 3        | Выполнение курсового проекта                 | -           |
| 4        | Выполнение курсовой работы                   | -           |
| 5        | Выполнение индивидуального задания (реферат) | 0(9)        |
| Итого:   |                                              | 57 (98)     |

### 3.6 Курсовой проект (работа), индивидуальное задание

Курсовой проект (работа), индивидуальное задание по дисциплине учебным планом у студентов очной формы обучения не предусмотрены.

Для студентов заочной формы обучения предусмотрены индивидуальные задания в форме реферата по следующей тематике [9]:

#### Темы для индивидуальных заданий в форме реферата:

1. Потребность в данных при принятии решений
2. Агрегирование и анализ информации
3. Роль вычислений в анализе данных
4. Методы статистической обработки данных
5. Модели объектов и модели данных
6. Понятие знания в количественных методах
7. Проблемы детерминизма и вероятности
8. Понятие «Большие данные» в управлении
9. Методы анализа больших данных
10. Программные продукты для анализа больших данных

## 4 ФОНД ОЦЕНОЧНЫХ СРЕДСТВ

### 4.1 Критерии и шкалы для интегрированной оценки уровня сформированности компетенций

#### *Составляющая компетенции – полнота знаний*

- нулевой уровень: неверные, не аргументированные, с множеством грубых ошибок ответы на вопросы. Уровень знаний ниже минимальных требований;

- минимальный уровень: даны не полные, неточные и неаргументированные ответы на вопросы. Допущено много грубых ошибок. Уровень знаний ниже минимальных требований;
- пороговый уровень: даны недостаточно полные, точные и аргументированные ответы на вопросы. Плохо знает термины, определения и понятия; основные закономерности, соотношения, принципы. Допущено много негрубых ошибок;
- средний уровень: даны достаточно полные, точные и аргументированные ответы на вопросы. В целом знает термины, определения и понятия; основные закономерности, соотношения, принципы. Допущено несколько негрубых ошибок;
- продвинутый уровень: даны полные, точные и аргументированные ответы на вопросы. Знает термины, определения и понятия; основные закономерности, соотношения, принципы. Допущено несколько негрубых ошибок;
- высокий уровень: даны полные, точные и аргументированные ответы на вопросы. Знает термины, определения и понятия; основные закономерности, соотношения, принципы. Допущено несколько неточностей.

### ***Составляющая компетенции – умения***

- нулевой уровень: полное отсутствие понимания сути методики решения задачи, допущено множество грубейших ошибок / задания не выполнены вообще;
- минимальный уровень: слабое понимание сути методики решения задачи, допущены грубые ошибки. Решения не обоснованы. Не умеет использовать нормативно-техническую литературу. Не ориентируется в специальной научной литературе;
- пороговый уровень: достаточное понимание сути методики решения задачи, допущены ошибки. Решения не всегда обоснованы. Умеет использовать нормативно-техническую литературу. Слабо ориентируется в специальной научной литературе;
- средний уровень: в целом понимает суть методики решения задачи, допущены ошибки. Решения не всегда обоснованы. Умеет использовать нормативно-техническую и специальную научную литературу;
- продвинутый уровень: в целом понимает суть методики решения задачи, допущены неточности. Способен обосновать решения. Умеет использовать нормативно-техническую и специальную научную литературу;
- высокий уровень: понимает суть методики решения задачи. Способен обосновать решения. Умеет использовать нормативно-техническую и специальную научную литературу, передовой опыт.

### ***Составляющая компетенции – владение навыками***

- нулевой уровень: не демонстрирует владение навыками выполнения профессиональных задач. Не может выполнить задания;
- минимальный уровень: не демонстрирует владение навыками выполнения профессиональных задач. Испытывает существенные трудности при выполнении отдельных заданий;
- пороговый уровень: владеет навыками выполнения профессиональных задач на пороговом уровне. Задания выполняет медленно и некачественно;
- средний уровень: владеет навыками выполнения профессиональных задач. Задания выполняет на среднем уровне по скорости и качеству;



- продвинутый уровень: владеет уверенными навыками выполнения профессиональных задач. Быстро и качественно выполняет задания, иногда допуская незначительные погрешности;
- высокий уровень: владеет уверенными навыками выполнения профессиональных задач. Быстро и качественно выполняет задания, при необходимости демонстрируя творческий подход.

### ***Обобщенная оценка сформированности компетенций***

- нулевой уровень: на нулевом уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- минимальный уровень: на минимальном уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- пороговый уровень: на пороговом уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- средний уровень: на среднем уровне сформированы: все составляющие; одна или две из трёх, остальные – на более высоком уровне;
- продвинутый уровень: на продвинутом уровне сформированы: все составляющие; одна или две из трёх, остальные – на высоком уровне;
- высокий уровень: на высоком уровне сформированы все составляющие компетенций.

## **4.2 Учебным планом экзамен не предусмотрен**

## **4.3 Критерии оценивания результатов освоения программы**

### **Критерии оценивания общей успеваемости (формирование итоговой оценки)**

Общая оценка знаний студентов по дисциплине проводится по 100-балльной шкале согласно критериям:

| <b>Вид работы (очное отделение)</b>                            | <b>Баллы</b> |
|----------------------------------------------------------------|--------------|
| Организационно-учебная работа студента в аудитории             | 5            |
| Индивидуальная работа студента (выполнение лабораторных работ) | 25           |
| Самостоятельная работа                                         | 10           |
| Результаты устного опроса / итоговая контрольная работа        | 60           |
| <b>Количество баллов по результатам текущего контроля</b>      | <b>100</b>   |
| <b>Общий итог</b>                                              | <b>100</b>   |

| <b>Вид работы (заочное отделение)</b>                                                                              | <b>Баллы</b> |
|--------------------------------------------------------------------------------------------------------------------|--------------|
| <b>Индивидуальная работа</b> (выполненное задание (50 баллов), оформление (20 баллов) и защита (30 баллов) работы) | <b>100</b>   |
| <b>Общий итог</b>                                                                                                  | <b>100</b>   |

*Организационно-учебная работа студента* в аудитории оценивается на основе таких критериев как посещаемость занятий, активность во время проведения лекционных и лабораторных занятий (вопросы лектору по теме лекционного ма-

териала, участие в обсуждении пройденного материала, самостоятельность в выполнении этапов лабораторных работ и т.п.).

### Критерии оценивания самостоятельной работы.

Самостоятельная и индивидуальная работа (включая выполнение СРС и ИРС) максимально оценивается в 35 баллов. В разрезе отдельных видов работ оценивание осуществляется следующим образом.

#### Оценивание СРС и ИРС по дисциплине «Аналитика больших данных для бизнеса»

| Вид работы                                                                          | Плановые сроки выполнения                    | Формы контроля и отчетности                                                                                                                             | Максимальное количество баллов |
|-------------------------------------------------------------------------------------|----------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------|
| <b>Индивидуальная работа (обязательные виды работ)</b>                              |                                              |                                                                                                                                                         |                                |
| 1. Выполнение лабораторных работ по дисциплине                                      | Один раз в неделю (для очной формы обучения) | Защита лабораторных работ                                                                                                                               | 15                             |
| 2. Письменное оформление расчетно-аналитической части*                              | Один раз в течение зачетного модуля          | Проверка правильности выполненных заданий                                                                                                               | 5*2=10                         |
| <i>Итого по ИРС</i>                                                                 |                                              |                                                                                                                                                         | <b>25</b>                      |
| <b>Самостоятельная работа (обязательные виды работ)</b>                             |                                              |                                                                                                                                                         |                                |
| 1. Подготовка аннотированного списка литературы по теме                             | Один раз в семестр                           | Обсуждение подготовленных материалов во время аудиторных занятий                                                                                        | 2                              |
| 2. Разработка таблиц и графиков результирующих параметров                           | Один раз в семестр                           |                                                                                                                                                         | 1                              |
| 3. Выполнение расчетных заданий                                                     |                                              |                                                                                                                                                         | 2                              |
| <i>Итого по СРС (обязательные виды работ)</i>                                       |                                              |                                                                                                                                                         | <b>5</b>                       |
| <b>Самостоятельная работа (выборочные виды работ)</b>                               |                                              |                                                                                                                                                         |                                |
| 1. Написание научных работ, участие в научных студенческих конференциях и семинарах | Один раз в семестр                           | Обсуждение с преподавателем подготовленных материалов, представление в печать, выступление с докладами на научных студенческих конференциях и семинарах | 5                              |
| <i>Итого по СРС (выборочные виды работ)</i>                                         |                                              |                                                                                                                                                         | <b>5</b>                       |
| <i>Всего по ИРС и СРС</i>                                                           |                                              |                                                                                                                                                         | <b>35</b>                      |

\* – данный вид работы является обязательной индивидуальной работой студента, однако с целью получения дополнительных баллов предоставляется возможность выполнения данного вида работы как одного из видов СРС.

### Критерии оценивания итогового контроля по шкале.

Полученная итоговая оценка по 100 бальной шкале определяет оценку по государственной шкале и шкале ECTS:

| Сумма баллов по 100-бальной шкале | Оценка по шкале ECTS | Оценка по государственной шкале |
|-----------------------------------|----------------------|---------------------------------|
| 90-100                            | A                    | Зачтено                         |
| 80-89                             | B                    | Зачтено                         |
| 75-79                             | C                    |                                 |
| 70-74                             | D                    | Зачтено                         |
| 60-69                             | E                    |                                 |
| 35-59                             | FX                   | Не зачтено                      |
| 0-34                              | F*                   |                                 |

\* – с обязательным повторным изучением дисциплины.

#### 4.4 Перечень контрольных заданий и иных материалов, необходимых для оценки знаний, умений и навыков

### ЗНАНИЕ – ПОНИМАНИЕ

#### Устный опрос по теоретическому материалу:

##### Тема 1. Введение в анализ данных

1. Дайте определение понятиям данные, информация, знание, мудрость.
2. Охарактеризуйте понятие «большие данные».
3. В чем состоит проблема анализа больших данных?
4. Каковы источники данных в предметной области?
5. Каковы цели и задачи анализа данных?
6. Перечислите основные инструменты анализа данных.
7. Какие технологии анализа данных получили наибольшее распространение?

##### Тема 2. Технологии анализа данных

1. В чем состоят аналитический и информационный подходы к моделированию?
2. Охарактеризуйте понятие «модель данных».
3. Какие виды моделей данных существуют?
4. Какие технологии поддерживают какие модели данных?
5. В чем суть технологии KDD?

##### Тема 3. Консолидация данных

1. Дайте характеристику OLTP-систем.
2. Каковы предпосылки появления СППР?
3. Опишите понятие хранилища данных.
4. Что собой представляют реляционные хранилища данных?
5. В чем специфика многомерных хранилищ данных?
6. Охарактеризуйте понятие гибридных хранилищ данных.
7. Что собой представляют виртуальные хранилища данных?
8. Охарактеризуйте процесс ETL.
9. Как извлекаются данные в ETL?

10. Как преобразовываются данные в ETL?
11. Что собой представляет загрузка данных в хранилище?
12. Каковы особенности загрузки данных из локальных источников?

#### **Тема 4. Трансформация данных**

1. Дайте понятие трансформации данных
2. Каковы особенности трансформации временных рядов?
3. Что собой представляет группировка (разгруппировка) данных?
4. В чем смысл слияния данных?
5. Что собой представляет квантование?
6. Когда необходимо применять нормализацию и кодирование данных?

#### **Тема 5. Визуализация данных**

1. В чем сущность визуализации данных?
2. Дайте понятие визуализаторов общего назначения.
3. Охарактеризуйте визуализаторы OLAP-анализа.
4. Какие визуализаторы применяют для оценки качества моделей?

#### **Тема 6. Оценка качества, очистка и предобработка данных**

1. В чем заключается задача оценки качества данных?
2. Какие технологии и методы оценки качества данных существуют?
3. В чем смысл очистки и предобработки данных?
4. Как осуществляется фильтрация данных?
5. Охарактеризуйте подходы к обработке дубликатов и противоречий?
6. Как выявляют аномальные значения?
7. Какие методы восстановления пропущенных значений существуют?
8. Охарактеризуйте подходы к сокращению размерности.
9. Как осуществляется сокращение числа признаков?
10. Как осуществляется сокращение числа значений признаков и записей?
11. Дайте понятие сэмплинга.

### **ПРИМЕНЕНИЕ**

#### **Типовые задания по лабораторным работам**

##### **Лабораторная работа № 1. Консолидация данных**

**Задание:** Изучить подходы к консолидации данных из различных источников с использованием языка запросов.

##### **Вопросы:**

1. Установка сервера MS SQL Server.
2. Разворачивание учебных баз на базе локального сервера.
3. Основные команды языка запросов MySQL.
4. Формирование запросов для консолидации данных из различных источников

##### **Лабораторная работа № 2. Трансформация данных.**

**Задание:** Изучить инструменты трансформации данных.

**Вопросы:**

1. Использование методов группировки и фильтрации в запросах MySQL.
2. Экспорт сырых данных для обработки.
3. Установка Python Anaconda 3.0 на локальную рабочую станцию.
4. Изучение основных операторов и синтаксиса языка Python.
5. Обработка сырых данных средствами языка Python и реимпорт в базу данных.

**Лабораторная работа № 3. Визуализация данных**

**Задание:** Изучить инструменты визуализации данных

**Вопросы:**

1. Визуализация стандартными средствами языка запросов MySQL.
2. Визуализация данных стандартными средствами MS Excel.

**Лабораторная работа № 4. Улучшение данных**

**Задание:** Изучить методы повышения качества данных.

**Вопросы:**

1. Осуществить слияние конфликтующих данных из различных источников с использованием языка запросов MySQL.
2. Произвести интерполяцию пропущенных данных в результирующей базе.
3. Провести сокращение признаков и записей в учебной базе.

**ТВОРЧЕСТВО****Темы для индивидуальных заданий (рефератов)**

1. Методы статистической обработки информации
2. Алгоритмы сглаживания
3. Методы прогнозирования временных рядов
4. Эволюционные алгоритмы обработки данных
5. Применение нейронных сетей в обработке данных
6. Методы Data Engineering
7. Сплайн-функции
8. Адаптивные методы прогнозирования
9. Метод интерполяции
10. Инструменты распознавания образов в анализе данных

**ОЦЕНИВАНИЕ****Контрольные вопросы для проведения текущего контроля:**

1. Понятия «данные», «информация», «знание», «мудрость» и их соотношение.
2. Понятие «большие данные» и его характеристики.
3. Проблема анализа больших данных.
4. Источники данных в предметной области.

5. Цели и задачи анализа данных
6. Основные инструменты анализа данных.
7. Технологии анализа данных.
8. Аналитический и информационный походы к моделированию?
9. Понятие «модель данных» и их виды.
10. Информационные технологии моделей данных.
11. Технология KDD.
12. Предпосылки появления СППР.
13. Понятие хранилища данных.
14. Реляционные хранилища данных.
15. Многомерные хранилища данных.
16. Гибридные хранилища данных.
17. Виртуальные хранилища данных
18. Процесс ETL.
19. Извлечение данных в ETL.
20. Преобразование данных в ETL.
21. Загрузка данных в хранилище.
22. Особенности загрузки данных из локальных источников.
23. Методы трансформации данных.
24. Особенности трансформации временных рядов.
25. Группировка (разгруппировка) данных.
26. Процесс слияния данных.
27. Квантование данных.
28. Нормализация и кодирование данных.
29. Визуализация данных.
30. Визуализаторы общего назначения.
31. Визуализаторы OLAP-анализа.
32. Визуализаторы оценки качества моделей.
33. Оценки качества данных.
34. Технологии и методы оценки качества данных.
35. Очистка и предобработка данных.
36. Фильтрация данных.
37. Подходы к обработке дубликатов и противоречий.
38. Выявление аномальных значений.
39. Методы восстановления пропущенных значений.
40. Подходы к сокращению размерности.
41. Сокращение числа признаков.
42. Сокращение числа значений признаков и записей.
43. Понятие сэмплинга.

**4.5 Учебным планом курсовое проектирование не предусмотрено**

## **5 РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА**

### **I Основная литература**



1. Шнарева, Г. В. Анализ данных : учебно-методическое пособие / Г. В. Шнарева, Ж. Г. Пономарева. — Симферополь : Университет экономики и управления, 2019. — 129 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/89482.html>. — Режим доступа: для авторизир. пользователей

2. Глотова, М. Ю. ИКТ и математические методы обработки данных : учебное пособие / М. Ю. Глотова, Е. А. Самохвалова. — Москва : Московский педагогический государственный университет, 2019. — 244 с. — ISBN 978-5-4263-0767-4. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/94642.html>. — Режим доступа: для авторизир. пользователей.

3. Маккинли, Уэс Python и анализ данных / Уэс Маккинли ; перевод А. Слинкина. — 2-е изд. — Саратов : Профобразование, 2019. — 482 с. — ISBN 978-5-4488-0046-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/88752.html>. — Режим доступа: для авторизир. пользователей.

4. Интеллектуальный предиктивный мультимодальный анализ слабоструктурированных больших данных / Н. Г. Ярушкина, И. А. Андреев, Г. Ю. Гуськов [и др.]. — Ульяновск: Ульяновский государственный технический университет, 2020. — 221 с. — ISBN 978-5-9795-2088-9. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/106136.html>. — Режим доступа: для авторизир. пользователей.

5. Железнов, М. М. Методы и технологии обработки больших данных: учебно-методическое пособие / М. М. Железнов. — Москва: МИСИ-МГСУ, ЭБС АСВ, 2020. — 46 с. — ISBN 978-5-7264-2193-3. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/101802.html>. — Режим доступа: для авторизир. пользователей.

## **II Дополнительная литература**

6. Терминологический словарь по предметам кафедры «Бизнес-информатика» / составители Я. А. Донченко [и др.]. — Симферополь : Университет экономики и управления, 2020. — 240 с. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <https://www.iprbookshop.ru/108063.html>. — Режим доступа: для авторизир. пользователей.

7. Интеллектуальный предиктивный мультимодальный анализ слабоструктурированных больших данных / Н. Г. Ярушкина, И. А. Андреев, Г. Ю. Гуськов [и др.]. — Ульяновск: Ульяновский государственный технический университет, 2020. — 221 с. — ISBN 978-5-9795-2088-9. — Текст: электронный // Электронно-библиотечная система IPR BOOKS: [сайт]. — URL: <https://www.iprbookshop.ru/106136.html>. — Режим доступа: для авторизир. пользователей.

## 6 УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Учебно-методические издания, разработанные в ДОННТУ:

8. Методические рекомендации для проведения лабораторных занятий по дисциплине «Аналитика больших данных для бизнеса» [Электронный ресурс]. для обучающихся по направлению подготовки 38.04.05 «Бизнес-информатика» всех форм обучения / ГОУВПО «ДОННТУ», каф. экон. кибернетики; сост.: В.Н. Тимохин. – 5,14 Мб. – Донецк, ГОУВПО «ДОННТУ», 2021. – 1 файл. – Систем. требования: Acrobat Reader.

<http://ed.donntu.org/books/21/m6773.pdf>

9. Методические рекомендации к выполнению самостоятельной работы и контрольных работ по дисциплине «Аналитика больших данных для бизнеса» [Электронный ресурс]: для обучающихся по направлению подготовки 38.04.05 «Бизнес-информатика» всех форм обучения / ГОУВПО «ДОННТУ», каф. экономической кибернетики; сост.: В.Н. Тимохин. – 174 Кб. – Донецк: ДОННТУ, 2021. – 1 файл. – Систем. требования: Acrobat Reader.

<http://ed.donntu.org/books/21/m7238.pdf>

### Электронно-информационные ресурсы

ЭБС ДОННТУ – <http://donntu.org/library>

IPR BOOKS <https://www.iprbookshop.ru/>

### Internet-ресурсы

УГТУ-УПИ. – URL: <http://library.ustu.ru>

Портал информационно-образовательных ресурсов УрФУ. – URL: <http://study.urfu.ru>. (СК № 11639  
[http://study.urfu.ru/view/aid\\_view.aspx?AidId=11639](http://study.urfu.ru/view/aid_view.aspx?AidId=11639); УМКД № 11096:  
[http://study.urfu.ru/view/aid\\_view.aspx?AidId=11096](http://study.urfu.ru/view/aid_view.aspx?AidId=11096))

## 7 МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

1. Учебная аудитория № 11.208, учебный корпус 11, для проведения лекционных и лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля. Специализированная мебель: доска аудиторная, парты, столы. Оборудование: ноутбук на базе процессора Intel Core I5; мультимедийный проектор, экран; подключение к сети Internet по Wi-Fi. Программное обеспечение: MS Windows 10 (лицензия OEM); OpenOffice (бесплатная версия 4.1.6).

2. Компьютерный класс № 11.204, учебный корпус 11, для проведения лабораторных занятий, групповых и индивидуальных консультаций, текущего контроля. Специализированная мебель: доска аудиторная, парты, столы. Оборудование: ноутбук на базе процессора Intel Core I5; стационарные компьютеры на базе процессоров Intel Celeron; мультимедийный проектор, экран; подключение к сети Internet по Wi-Fi. Программное обеспечение: MS Windows 10 (лицензия OEM), MS Windows 7 (лицензия OEM); OpenOffice (бесплатная версия 4.1.6); Microsoft Office 2007 Professional (лицензия Microsoft № 00045-577-942-543); AnyLogic

8.6.0. PLE (ограниченная лицензия для обучения); PowerSim Express 10 (ограниченная лицензия для обучения); Python Anaconda 3.0 (открытая лицензия); MS SQL Server (открытая лицензия); MS Visual Studio 2010 Professional (лицензия MSDN AA и VMware AP); ARIS (ограниченная лицензия для обучения); 1С Предприятие 8.3 (ограниченная лицензия для обучения); Business Studio 3.0 (демонстрационная версия).

3. Помещения для самостоятельной работы с возможностью подключения к сети Internet и обеспечением доступа в электронную информационно-образовательную среду организации: читальные залы, учебные корпуса 2,3 (компьютерная техника с возможностью подключения к сети Internet и обеспечением доступа в электронную информационно-образовательную среду (ЭИОС ДОН-НТУ) и электронно-библиотечную систему (ЭБС IPRbooks), а также возможностью индивидуального неограниченного доступа обучающихся в ЭБС и ЭИОС посредством Wi-Fi с персональных мобильных устройств. ОС - Alt Linux (лицензия GNU LGPL), Libreoffice 5.3.4 (лицензия GNU LGPL) - общественная лицензия MPL 2.0/ Grub loader for ALT Linux - лицензия GNU LGPL v3/ Mozilla Firefox - лицензия MPL2.0, Moodle (Modular Object-Oriented Dynamic Learning Environment) - лицензия GNU GPL.