

**федеральное государственное автономное образовательное  
учреждение высшего образования  
«Национальный исследовательский университет ИТМО»  
(Университет ИТМО)**

УТВЕРЖДАЮ  
Проректор по научной работе  
\_\_\_\_\_ В.О. Никифоров  
« \_\_\_\_ » \_\_\_\_\_ 2022 г.  
м.п.

**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**

**Основы анализа данных**

**Группа научной специальности:** все группы научных специальностей

**Научной специальности:** все научные специальности

**Форма обучения:** Очная

Санкт-Петербург  
2022 г.

Рабочая программа составлена на основании Требований к программам подготовки научных и научно-педагогических кадров в аспирантуре, самостоятельно устанавливаемых федеральным государственным автономным образовательным учреждением высшего образования «Национальный исследовательский университет ИТМО»:

Код и наименование группы научной специальности	Реквизиты приказа об утверждении Требований Университета ИТМО
1.1.Математика и механика 1.2.Компьютерные науки и информатика 1.3.Физические науки 1.4.Химические науки 1.5.Биологические науки 1.6.Науки о Земле и окружающей среде 2.1.Строительство и архитектура 2.2.Электроника, фотоника, приборостроение и связь 2.3.Информационные технологии и телекоммуникации 2.4.Энергетика и электротехника 2.5.Машиностроение 2.6.Химические технологии, науки о материалах, металлургия 2.7.Биотехнологии 4.3.Агроинженерия и пищевые технологии 5.2.Экономика 5.5.Политические науки 5.7.Философия	Приказ ректора №325-од от 31.03.2022  Решение Научно-Технического совета №12 от 28.03.2022

Программу разработала: Болдырева Е.А., к.т.н.

Программа одобрена на заседании НТС Университета ИТМО протокол №12 от 28 марта 2022 года.

**Место в структуре программы аспирантуры:**

Блок 2, Образовательный компонент

**Форма обучения:** очная, дистанционная

**Год обучения:** 2

**Семестр:** 4

**Форма аттестации:** зачет

Вид деятельности	Семестр
	4
Занятий в контактной форме, час.	18
из них лекции, час.	12
из них научно-практических занятий, час.	4
из них промежуточной аттестации (включая консультации), час.	2
Самостоятельная работа, час.	90
Всего часов	108
Всего зачетных единиц	3

## Аннотация к рабочей программе дисциплины «Основы анализа данных»

Дисциплина «Основы анализа данных» реализуется в рамках образовательных программ высшего образования – программ подготовки научно-педагогических кадров в аспирантуре по очной форме обучения.

### Разделы рабочей программы

1. Место дисциплины в структуре программы аспирантуры.
2. Планируемые результаты обучения по дисциплине.
3. Структура и содержание дисциплины.
4. Текущий контроль и промежуточная аттестация.
5. Учебно-методическое и информационное обеспечение дисциплины.
6. Материально-техническое обеспечение дисциплины.
7. Фонды оценочных средств текущего контроля и промежуточной аттестации.

#### 1. Место дисциплины в структуре программы аспирантуры

Дисциплина «Основы анализа данных» реализуется в четвертом семестре в рамках образовательного компонента Блока 2, поскольку знания, умения и навыки организации и осуществления научно-исследовательской деятельности с пониманием принципов анализа, визуализации и интерпретации данных с применением языка программирования Python для различных исследовательских научных и прикладных задач.

#### 2. Планируемые результаты обучения по дисциплине

Дисциплина «Основы анализа данных» направлена на формирование **компетенции УК-1:** Способность к критическому анализу и оценке современных научных достижений, генерированию новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях, **компетенции ОПК-1:** Способность самостоятельно осуществлять научно-исследовательскую деятельность в соответствующей профессиональной области с использованием современных методов исследования и информационно-коммуникационных технологий и **компетенции ОПК-3:** Готовность организовать работу исследовательского коллектива в области профессиональной деятельности и смежных наук *в части следующих результатов обучения:*

Формируемые компетенции (код компетенции)	Планируемые результаты обучения по дисциплине (модулю)
--	--

УК-1	<p><b>Знать:</b> 33 (УК-1) основные проблемы внедрения научных разработок в практическую деятельность</p> <p><b>Уметь:</b> У2 (УК-1) оценивать подходы к реализации различных вариантов решения исследовательских и практических задач; У3 (УК-1) формулировать альтернативные варианты решения исследовательских и практических задач.</p> <p><b>Владеть (иметь опыт):</b> В2 (УК-1) навыками выбора методов и средств решения задач исследования.</p>
ОПК-1	<p><b>Знать:</b> 31 (ОПК-1) современные способы использования информационно-коммуникационных технологий в области (областях) профессиональной деятельности, к осуществлению научно-исследовательской и иной профессиональной деятельности в которых готовится выпускник</p> <p><b>Уметь:</b> У1 (ОПК-1) выбирать и применять в научно-исследовательской деятельности современные информационно-коммуникационные технологии</p> <p><b>Владеть:</b> В1 (ОПК-1) навыками использования информационно-коммуникационных технологий при поиске и анализе информации, в том числе по тематике проводимых научных исследований</p>
ОПК-3	<p><b>Уметь:</b> У1 (ОПК-3) планировать научную работу, формировать состав рабочей группы и оптимизировать распределение обязанностей между членами исследовательского коллектива</p> <p><b>Владеть:</b> В1 (ОПК-3) навыками планирования и распределения работы между членами исследовательского коллектива В2 (ОПК-3) навыками коллективного обсуждения планов работ, получаемых научных результатов, согласования интересов сторон и урегулирования конфликтных ситуаций в команде</p>

### Способы формирования планируемых результатов обучения

Результаты изучения дисциплины по уровням освоения (знать, уметь, владеть)	Формы организации занятий		
	Лекции	Научно-практические занятия	Самостоятельная работа

<b>УК-1: Способность к критическому анализу и оценке современных научных достижений, генерированию новых идей при решении исследовательских и практических задач, в том числе в междисциплинарных областях</b>			
33 (УК-1) основные проблемы внедрения научных разработок в практическую деятельность	+		+
У2 (УК-1) оценивать подходы к реализации различных вариантов решения исследовательских и практических задач; У3 (УК-1) формулировать альтернативные варианты решения исследовательских и практических задач.	+	+	+
В2 (УК-1) навыками выбора методов и средств решения задач исследования.		+	+
<b>ОПК-1: способность самостоятельно осуществлять научно-исследовательскую деятельность в соответствующей профессиональной области с использованием современных методов исследования и информационно-коммуникационных технологий</b>			
31 (ОПК-1) современные способы использования информационно-коммуникационных технологий и технологий машинного обучения в области (областях) профессиональной деятельности, к осуществлению научно-исследовательской и иной профессиональной деятельности, в которых готовится выпускник	+		+
У1 (ОПК-1) выбирать и применять в научно-исследовательской деятельности современные информационно-коммуникационные технологии	+	+	+
В1 (ОПК-1) навыками использования информационно-коммуникационных технологий при поиске и анализе информации, в том числе по тематике проводимых научных исследований		+	+
<b>ОПК-3: готовность организовать работу исследовательского коллектива в области профессиональной деятельности и смежных наук</b>			
У1 (ОПК-3) планировать научную работу, формировать состав рабочей группы и оптимизировать распределение обязанностей между членами исследовательского коллектива	+	+	+

В1 (ОПК-3) навыками планирования и распределения работы между членами исследовательского коллектива		+	+
В2 (ОПК-3) навыками коллективного обсуждения планов работ, получаемых научных результатов, согласования интересов сторон и урегулирования конфликтных ситуаций в команде		+	+

### 3. Структура и содержание дисциплины

Изучение курса «Основы анализа данных» включает в себя лекции, на которых рассматривается теоретическое содержание курса; научно-практические занятия, предусматривающие углубленное изучение и обсуждение вопросов, обозначенных в темах дисциплины; самостоятельную работу, заключающуюся в подготовке к лекционным и научно-практическим занятиям. Темы, рассматриваемые на лекциях и изучаемые самостоятельно, закрепляются на научно-практических занятиях, по вопросам, вызывающим затруднения, проводятся консультации.

#### Структура дисциплины:

№ раздела	Наименование раздела дисциплины	Виды учебной нагрузки и их трудоемкость, часы					Форма текущего контроля успеваемости, промежуточной аттестации
		Всего часов	Лекции	Научно-практические занятия	Семинары	Самостоятельная работа	
1.	Введение в интеллектуальный анализ данных.	34	4	-	-	30	Практические задания по итогам освоения модулей (разделов) курса.
2.	Предварительная обработка данных.	36	4	2	-	30	
3.	Операции над данными. Визуализация данных.	36	4	2	-	30	
4.	Промежуточная аттестация	2	-	-	-	-	Зачет в письменной форме с последующим собеседованием
<b>ИТОГО:</b>		<b>108</b>	<b>12</b>	<b>4</b>	<b>-</b>	<b>90</b>	

### Содержание дисциплины:

№ п/п	Наименование раздела дисциплины	Содержание	Ссылки на результаты обучения
1	Введение в интеллектуальный анализ данных.	Введение в интеллектуальный анализ данных, Data mining. Методы интеллектуального анализа данных. Основы теории вероятности и статистики для анализа данных. Основные этапы анализа данных. Области применения. Особенности анализа данных на языке программирования Python.	33 (УК-1) У2 (УК-1) У3 (УК-1) В2 (УК-1) 31 (ОПК-1) У1 (ОПК-1) В1 (ОПК-1) У1 (ОПК-3) В1 (ОПК-3) В2 (ОПК-3)
2	Предварительная обработка данных.	Структуры данных: создание, свойства, срезы данных, добавление, фильтрация и сортировка данных. Работа с индексами (иерархическая индексация). Чистка данных (пропущенные значения, повторяющиеся значения, замена значений, применение функций).	
3	Операции над данными. Визуализация данных.	Математические операции. Сводка статистик, арифметические операции. Преобразование данных, оконные функции. Операции объединения данных. Агрегирование данных. Визуализация данных. Настройка графиков и их основные типы.	

### Виды учебной и самостоятельной работы

Виды учебной работы	Ссылки на результаты обучения	Часы
На основе изучения литературы по теме лекционных и научно-практических занятия готовится к ответу с места на предложенные вопросы, к участию в дискуссиях, к тестированию по изученному материалу	33 (УК-1) У2 (УК-1) У3 (УК-1) В2 (УК-1) 31 (ОПК-1) У1 (ОПК-1) В1 (ОПК-1) У1 (ОПК-3)	18

	V1 (ОПК-3) V2 (ОПК-3)	
<b>Виды самостоятельной работы</b>	<b>Ссылки на результаты обучения</b>	<b>Часы на выполнение</b>
Самостоятельная подготовка к лекционным и научно-практическим занятиям	33 (УК-1) У2 (УК-1) У3 (УК-1) V2 (УК-1) 31 (ОПК-1) У1 (ОПК-1) V1 (ОПК-1) У1 (ОПК-3) V1 (ОПК-3) V2 (ОПК-3)	90
Успешное освоение материала, изучаемого в ходе лекционных и научно-практических занятия, требует дополнительного самостоятельного изучения. По каждому разделу учебной дисциплины предусмотрено изучение теоретического материала с использованием компьютерных технологий; самостоятельное изучение теоретического материала дисциплины с использованием Internet-ресурсов, информационных баз, методических разработок, специальной учебной и научной литературы.		

#### **4. Текущий контроль и промежуточная аттестация**

Текущая аттестация по дисциплине «Основы анализа данных» осуществляется на лекциях и научно-практических занятиях и заключается в оценке активности и качества участия в опросах и собеседованиях по проблемам, изучаемых в рамках тем лекционных занятий, аргументированности позиции; в форме тестирования оценивается широта используемых теоретических знаний.

Промежуточная аттестация по дисциплине «Основы анализа данных» в четвертом семестре (итоговая аттестация по дисциплине) проводится в форме собеседования. Оценка по результатам собеседования выставляется в формате «зачтено» - «не зачтено». Оценка «зачтено» подразумевает удовлетворительное освоение дисциплины и выставляется аспиранту в случае, если он в результате собеседования демонстрирует сформированность знаний, умений и навыков использования современных информационных технологий и методов научной коммуникации в образовательной и научно-исследовательской деятельности.

#### **5. Учебно-методическое и информационное обеспечение дисциплины**

Учебно-методическое и программное обеспечение дисциплины размещено на образовательном портале Университета ИТМО.

**Профессиональные базы данных, интернет-ресурсы, электронные библиотеки и информационные справочные системы:**

№	Ссылка на информационный ресурс	Наименование разработки в электронной форме	Доступность
1.	<a href="http://e.lanbook.com/">http://e.lanbook.com/</a>	ЭБС на платформе «Лань». Учебники и учебные пособия для университетов издательства «Лань»	Индивидуальный неограниченный доступ
2.	<a href="http://lib.ifmo.ru/">http://lib.ifmo.ru/</a>	Электронная библиотека НИУ ИТМО.	Индивидуальный неограниченный доступ

### Основная литература:

1. Барсегян, А.А., Куприянов М. С. и др. Методы и модели анализа данных : OLAP и Data Mining : учеб. пособие - СПб. : БХВ-Петербург, 2004.
2. Барсегян, А.А., Куприянов М. С. и др. Технологии анализа данных : Data Mining, Visual Mining, Text Mining, OLAP : учеб. - 2-е изд., перераб. и доп. - СПб. : БХВ-Петербург, 2007.
3. Маккэндлесс Д. Инфографика. Самые интересные данные в графическом представлении / пер. с англ. Валентины Соколовой.— М. : Манн, Иванов и Фербер, 2014.
4. Бессмертный И.А.. Искусственный интеллект – СПб: СПбГУ ИТМО, 2010. – 132 с.

### Дополнительная литература:

1. Ian H. Witten, Eibe Frank "Data Mining: Practical Machine Learning Tools and Techniques"
2. Орлов А.И. Прикладная статистика. Учебник. / А.И.Орлов.- М.: Издательство «Экзамен», 2004. - 656 с.

Аспирант может дополнить список использованной литературы иными современными источниками, не представленными в списке рекомендованной литературы.

**Средства, обеспечивающие адаптацию электронных и печатных образовательных ресурсов для обучающихся из числа лиц с инвалидностью и ограниченными возможностями здоровья:**

1) Доступ к изданиям электронно-библиотечной системы «Издательство «Лань» (<https://e.lanbook.com>), в адаптированных форматах для лиц с инвалидностью и ОВЗ.

2) Специальные технические средства обучения коллективного и индивидуального пользования для лиц с инвалидностью и ОВЗ:

а) В библиотеке по адресам Кронверкский пр., д.49 и ул. Ломоносова, д.9 обучающимся, имеющим нарушения зрения, предоставляется компьютерное место с клавиатурой, маркированной шрифтом Брайля, и увеличительные лупы нового поколения с подсветкой и семикратным увеличением (лупы настольные с подсветкой PowerLux).

3) Услуги по адаптации учебно-методического материала для лиц с инвалидностью и ОВЗ:

а) обучающиеся с нарушениями зрения по запросу могут получить специальную учебную, научную литературу и периодические издания на основании

действующего договора о сотрудничестве между Университетом ИТМО и Государственной библиотекой для слепых и слабовидящих; для обучающихся с нарушениями зрения учебные материалы могут быть предложены на шрифте Брайля.

б) обучающимся с нарушениями слуха по запросу предоставляются услуги сурдопереводчика на основании договора между Университетом ИТМО и «Всероссийским обществом глухих» (СПб РО ОООИ ВОГ).

## 6. Материально-техническое обеспечение дисциплины

Программное обеспечение:

Для обеспечения реализации дисциплины используется стандартный комплект программного обеспечения (ПО), включающий регулярно обновляемое лицензионное ПО Windows и MS Office.

Использование специализированного программного обеспечения для изучения дисциплины не требуется.

Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения.
<b>Лекции:</b>		
Учебная аудитория (для проведения лекций, научно-практических занятий, семинаров, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, самостоятельной работы)	мультимедийное оборудование (проектор, экран, звуковоспроизводящее оборудование)	Операционная система Windows 10 (ключ Windows 7 Professional на системном блоке, договор на поставку аппаратно-программных мультимедийных комплексов № А 58-13 от 19 июня 2013г.)
<b>Научно-практические занятия:</b>		
Компьютерная аудитория (для проведения лекций, научно-практических занятий, семинаров, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, самостоятельной работы)	аудитория, оснащенная компьютерами с возможностью подключения к сети «Интернет» и обеспечением доступа к ЭИОС на 12 посадочных мест	Операционная система Microsoft WINHOME 10 (лицензия №69794626), Microsoft Office (лицензия №69794626)
<b>Самостоятельная работа:</b>		

Компьютерная аудитория (для проведения лекций, научно-практических занятий, семинаров, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации, самостоятельной работы)	12 персональных компьютеров в составе локальной вычислительной сети, подключенной к Internet (30 Мбит/с).	Операционная система Microsoft WINHOME 10 (лицензия №69794626), Microsoft Office (лицензия №69794626)
---	---	---

## 7. Фонды оценочных средств текущего контроля и промежуточной аттестации

### Требования к структуре и содержанию фонда оценочных средств текущего контроля и промежуточной аттестации по дисциплине

Перечень оценочных средств, применяемых на каждом этапе проведения текущего контроля и промежуточной аттестации по дисциплине, представлены в таблице

Наименование оценочного средства	Краткая характеристика оценочного средства	Представление оценочного средства в фонде
<b>Оценочные средства текущего контроля</b>		
Тематический опрос (в форме ответов на вопросы)	Средство контроля, организованное как специальная беседа по тематике предыдущей лекции и рассчитанное на выяснение объема и качества знаний, усвоенных обучающимися по определенному разделу, теме, проблеме.	Перечень тем, изучаемых в рамках дисциплины
Собеседование (в форме беседы, дискуссии по теме)	Средство контроля, организованное как свободная беседа, дискуссия по тематике изучаемой дисциплины, рассчитанное на выяснение объема знаний обучающегося по всем изученным разделам, темам; свободного использования терминологии для аргументированного выражения собственной позиции.	Перечень тем, изучаемых в рамках дисциплины
Тестирование	Средство контроля, позволяющее получить оценку уровня фактических знаний аспиранта по изученной теме.	Образцы тестов
<b>Оценочные средства промежуточной аттестации</b>		

<p>Собеседование</p>	<p>Средство, позволяющее получить экспертную оценку умения использовать положения и категории современных методов и технологий научной коммуникации для оценивания и анализа различных фактов и явлений в своей профессиональной области; способности успешно применять навыки анализа основных методологических проблем своей профессиональной области в историческом контексте</p>	<p>Перечень вопросов к зачету</p>
----------------------	--	-----------------------------------

### Критерии сформированности компетенций

Код формируемой компетенции	Планируемые результаты обучения по дисциплине	Не сформировано	Сформировано
<b>УК-1</b>	З3 (УК-1) основные проблемы внедрения научных разработок в практическую деятельность	Отсутствие знаний об основных проблемах внедрения научных разработок в практическую деятельность	Сформированные знания об основных проблемах внедрения научных разработок в практическую деятельность
	У2 (УК-1) уметь оценивать подходы к реализации различных вариантов решения исследовательских и практических задач	Отсутствие умения оценивать подходы к реализации различных вариантов решения исследовательских и практических задач	Сформированные умения оценивать подходы к реализации различных вариантов решения исследовательских и практических задач
	У3 (УК-1) уметь формулировать альтернативные варианты решения исследовательских и практических задач	Отсутствие умения формулировать альтернативные варианты решения исследовательских и практических задач.	Сформированные умения формулировать альтернативные варианты решения исследовательских и практических задач.
	В2 (УК-1) владеть навыками выбора методов и средств решения задач исследования	Отсутствие владения навыками выбора методов и средств решения задач исследования	Сформированные навыки выбора методов и средств решения задач исследования
<b>ОПК-1</b>	З1 (ОПК-1) современные способы использования информационно-коммуникационных технологий в области (областях) профессиональной деятельности, к осуществлению научно-исследовательской и иной профессиональной деятельности, в которых готовится выпускник	Отсутствие знаний о современных способах использования информационно-коммуникационных технологий в области (областях) профессиональной деятельности, к осуществлению научно-исследовательской и иной профессиональной деятельности, в которых готовится выпускник	Сформированные знания о современных способах использования информационно-коммуникационных технологий в области (областях) профессиональной деятельности, к осуществлению научно-исследовательской и иной профессиональной деятельности, в которых готовится выпускник
	У1 (ОПК-1) выбирать и применять в научно-исследовательской деятельности современные информационно-коммуникационные технологии	Отсутствие умения выбирать и применять в научно-исследовательской деятельности современные информационно-коммуникационные технологии	Сформированные умения выбирать и применять в научно-исследовательской деятельности современные информационно-коммуникационные технологии
	В1 (ОПК-1) владеть навыками использования информационно-коммуникационных технологий при поиске и анализе информации, в том числе по тематике проводимых научных исследований	Отсутствие владения навыками использования информационно-коммуникационных технологий при поиске и анализе информации, в том числе по тематике проводимых научных исследований	Сформированные владения использованием информационно-коммуникационных технологий при поиске и анализе информации, в том числе по тематике проводимых научных исследований
<b>ОПК-3</b>	У1 (ОПК-3) планировать научную работу, формировать состав рабочей группы и	Отсутствие умения планировать научную работу, формировать состав рабочей группы и	Сформированные умения планировать научную работу, формировать состав рабочей группы и оптимизировать

	оптимизировать распределение обязанностей между членами исследовательского коллектива	оптимизировать распределение обязанностей между членами исследовательского коллектива	распределение обязанностей между членами исследовательского коллектива
	V1 (ОПК-3) навыками планирования и распределения работы между членами исследовательского коллектива	Отсутствие владения навыками планирования и распределения работы между членами исследовательского коллектива	Сформированные владения навыками планирования и распределения работы между членами исследовательского коллектива
	V2 (ОПК-3) навыками коллективного обсуждения планов работ, получаемых научных результатов, согласования интересов сторон и урегулирования конфликтных ситуаций в команде	Отсутствие владения навыками коллективного обсуждения планов работ, получаемых научных результатов, согласования интересов сторон и урегулирования конфликтных ситуаций в команде	Сформированные владения навыками коллективного обсуждения планов работ, получаемых научных результатов, согласования интересов сторон и урегулирования конфликтных ситуаций в команде

## **Требования к структуре и содержанию оценочных средств**

Требования к структуре и содержанию тестов:

Тестирование проводится с применением тестов открытого и закрытого типа. Тест выполняется письменно. Время выполнения теста 15-20 минут.

Требования к порядку проведения зачета:

Зачет проводится в форме собеседования. Подготовка к собеседованию выполняется по билетам, в билете 2 вопроса. Время подготовки - 30 минут. Собеседование проводится преподавателем по темам билета. Аспиранту могут быть заданы дополнительные вопросы в рамках изученного курса.

### **Критерии выставления оценки:**

Знания, умения и навыки обучающихся при промежуточном контроле в форме зачета определяются оценками «зачтено» и «не зачтено».

Оценка «зачтено» подразумевает удовлетворительное освоение дисциплины и выставляется аспиранту в случае, если он в результате собеседования демонстрирует усвоение программного материала, способен грамотно изложить его, может правильно применять теоретические положения и владеет необходимыми умениями и навыками при выполнении заданий по дисциплине.

Оценка «не зачтено» подразумевает неудовлетворительное освоение дисциплины и выставляется аспиранту в случае, если он в результате собеседования допускает существенные ошибки, с большими затруднениями выполняет задания и задачи по дисциплине.

### **Примеры контрольных вопросов для самостоятельной работы и подготовки к зачету:**

1. Основные задачи анализа, интерпретации и визуализации данных.
2. Назовите основные проблемы обработки данных.
3. Матрица данных. Структура матрицы данных.
4. Матрица объект-объект. Матрица признак-признак.
5. Методы измерения признаков и основные проблемы измерения признаков.
6. Классификация случайных событий. Классическое определение вероятности. Свойства вероятности.
7. Статистическое и геометрическое определения вероятности события и условия их применимости.
8. Совместные и несовместные события. Зависимые и независимые события.
9. Понятие случайной величины и ее описание.
10. Дискретная случайная величина и закон ее распределения.
11. Математическое ожидание дискретной случайной величины и его свойства.
12. Библиотеки Pandas и NumPy языка программирования Python для анализа данных.
13. Структуры данных библиотеки Pandas. Создание структур из списка, из словаря, при помощи функций, из файла. Примеры.
14. Основные свойства структур данных библиотеки Pandas.
15. Определение среза данных. Создание среза данных.
16. Работа с индексами. Иерархическая индексация данных.
17. Методы чистки данных. Повторяющиеся значения, пропущенные значения, замена значений.

18. Объединение и агрегирование данных.
19. Визуализация данных. Цели и задачи визуализации данных
20. Основные типы и настройка графиков.
21. Одномерная, двумерная и многомерная визуализация данных.
22. Сферы применения технологии интеллектуального анализа данных.
23. Системы поддержки и принятия решений и управленческие информационные системы. Роль данных в принятии решений.

### Пример практического задания.

#### Практическое задание 1. Работа с библиотекой Pandas

Pandas - это программная библиотека на языке Python для обработки и анализа данных. Основными структурами данных в Pandas являются классы Series и DataFrame.

Цель работы – изучение основных функций и методов библиотеки Pandas для анализа данных.

Задания:

1. Установите программное обеспечение (ПО) для работы с языком программирования и его библиотеками. Рекомендованное ПО – пакет Anacondas и его «подсервис» Jupiter Notebook.
2. Скачайте предложенный набор данных (датасет) data.csv.
3. Прочтите данные из файла data.csv с помощью функции: `pd.read_csv(..., delimiter=',')`. Что изменится, если в качестве параметра функции указать `delimiter=';'`?
4. Установите первый столбец датафрейма в качестве индексного столбца.
5. Выведите описание прочтенных данных с помощью функции `.describe()`. Опишите полученные измерения.
6. Отобразите несколько первых и несколько последних записей датафрейма с помощью функций `.head()`, `.tail()`. Как вывести 10 первых и 10 последних записей? Как вывести 10 случайных записей?
7. Выведите названия столбцов датафрейма. Выведите типы данных столбцов. Какому типу принадлежит каждый столбец (вещественный, целый, категориальный)?
8. Найдите столбец, который содержит неправдоподобные данные (перемешаны абсолютные и относительные значения). Сделайте все значения данного столбца абсолютными (необходимо перемножить значения столбцов без использования циклов).
9. Поменяйте названия столбца.
10. Вычислите среднее значение столбца и поставьте вычисленное значение на места пропущенных значений.
11. Используя метод `groupby` оцените «картину» для различных категорий данных. Сделайте выводы.
12. \* Визуализация полученных данных. Постройте график по результатам работы метода `groupby`.