

Федеральное государственное образовательное бюджетное учреждение  
высшего образования

**«ФИНАНСОВЫЙ УНИВЕРСИТЕТ ПРИ ПРАВИТЕЛЬСТВЕ  
РОССИЙСКОЙ ФЕДЕРАЦИИ»  
(Финансовый университет)**

**Департамент анализа данных и машинного обучения  
Факультета информационных технологий и анализа больших данных**

СОГЛАСОВАНО

АО «ЦПР»

Генеральный директор

\_\_\_\_\_  
Самохин А.Н.

(подпись)

18.04.2023 г.

УТВЕРЖДАЮ

Проректор по учебной  
и методической работе

\_\_\_\_\_  
Е.А. Каменева

25.04.2023 г.

**А.А. Быков**

**Методы визуализации данных**

**Рабочая программа дисциплины**

для студентов, обучающихся по направлению подготовки

09.03.03 - Прикладная информатика,

ОП «Инженерия данных»,

ОП «Прикладные информационные системы в экономике и финансах»

*Рекомендовано Ученым советом*

*Факультета информационных технологий и анализа больших данных*

*(протокол №31 от 18.04.2023г.)*

*Одобрено Советом учебно-научного*

*Департамента анализа данных и машинного обучения*

*(протокол №2 от 29.03.2023г.)*

**Москва 2023**

## Содержание

Наименование разделов РПД		Стр.
1.	Наименование дисциплины	3
2.	Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине	3
3.	Место дисциплины в структуре образовательной программы	5
4.	Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	5
5.	Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	6
5.1.	Содержание дисциплины	6
5.2.	Учебно – тематический план	7
5.3.	Содержание семинаров, практических занятий	8
6.	Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	9
6.1.	Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	9
6.2.	Перечень вопросов, заданий, тем для подготовки к текущему контролю	10
7.	Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	15
8.	Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	26
9.	Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	27
10.	Методические указания для обучающихся по освоению дисциплины	28
11.	Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	29
12.	Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	29

## 1. Наименование дисциплины

«Методы визуализации данных».

## 2. Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине

Код компетенции	Наименование компетенции	Индикаторы достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции
<b>ОП «Инженерия данных»</b>			
ПКН-5	Способность участвовать в документальном сопровождении разработки ИТ в рамках проектных групп, применять средства автоматизации управления проектами ИТ	1. Демонстрирует знание основ версионирования и управления изменениями при разработке ПО. Использует системы контроля версий для ведения совместной разработки.	<b>Знать:</b> приемы визуального представления хода разработки программных продуктов.  <b>Уметь:</b> использовать системы контроля версий для ведения совместной разработки приложений визуализации данных, работать с базами данных для хранения и извлечения необходимых данных
		2. Демонстрирует знание основ тестирования программного обеспечения, умение создавать автоматизированные модульные и интеграционные тесты	<b>Знать:</b> способы обработки результатов тестирования для их наглядного представления  <b>Уметь:</b> визуализировать результаты тестирования программного обеспечения
		3. Готовит документацию к программе, коммуницирует в пределах группы разработки и за ее пределами о значимых аспектах информационной системы и информационной инфраструктуры в письменной и устной форме.	<b>Знать:</b> требования к составу программной документации  <b>Уметь:</b> визуализировать структуру программного продукта и статистику результатов его работы
		4. Демонстрирует знание жизненного цикла информационных систем, участвует в процессе разработки ПО	<b>Знать:</b> способы визуального представления жизненных циклов информационных систем  <b>Уметь:</b> визуализировать этапы

		на разных этапах	жизненного цикла информационных систем, строить тренд состояний процесса разработки программного продукта
ПКП-2	Способность разрабатывать, согласовывать и управлять исполнением технического задания и технического проекта с использованием технологий больших данных	1.Работает со стандартами, в том числе адаптирует стандарты для специфических требований больших данных.	<p><b>Знать:</b> стандарты и их возможности адаптации под специфические требования больших данных.</p> <p><b>Уметь:</b> визуализировать наборы данных в соответствии с поставленными требованиями с целью их дальнейшего анализа</p>
		2.Разрабатывает технические задания и технические проекты для технологий больших данных.	<p><b>Знать:</b> методы составления технических заданий на реализацию программных продуктов для визуализации данных</p> <p><b>Уметь:</b> составлять технических заданий на реализацию программных продуктов для визуализации данных и проводить их верификацию поставленным требованиям</p>
		3. Реализует управление рабочими проектами технологической инфраструктуры больших данных.	<p><b>Знать:</b> как возможно визуализировать процесс управления программными проектами в рамках технологической инфраструктуры больших данных</p> <p><b>Уметь:</b> визуализировать процесс управления программными проектами в рамках технологической инфраструктуры больших данных</p>
ОП «Прикладные информационные системы в экономике и финансах»			
ПKN-8	Способность использовать современные информационные системы для решения задач предметной области, в том числе отечественного производства	1.Демонстрирует знания об основных информационных технологиях и программных средствах, позволяющих их использовать.	<p><b>Знать:</b> основные информационные технологии и программные средства визуализации данных</p> <p><b>Уметь:</b> использовать современные информационные системы для решения профессиональных задач в области визуализации данных</p>

		1.Рационально выбирает информационные технологии и реализующие их программные средства, в том числе, с учетом страны происхождения программных средств.	<p><b>Знать:</b> функциональные возможности, отличительные особенности и сферы применения различных программных средств визуализации данных</p> <p><b>Уметь:</b> выбирать программные средства визуализации данных с учетом из функциональных возможностей и отличительных особенностей, в том числе, с учетом страны происхождения программных средств</p>
		2.Использует современные информационные технологии и программные средства при решении задач разработки программного обеспечения для экономических и финансовых приложений.	<p><b>Знать:</b> возможности современных программных средств визуализации данных и методику их использования при решении задач профессиональной деятельности</p> <p><b>Уметь:</b> разрабатывать программное обеспечение визуализации данных для экономических и финансовых приложений</p>

### 3. Место дисциплины в структуре образовательных программ

Дисциплина «Методы визуализации данных» является дисциплиной Профиля «Инженерия данных» по направлению подготовки 09.03.03 – Прикладная информатика, ОП «Инженерия данных».

Дисциплина «Методы визуализации данных» является дисциплиной Цикла профиля (элективный) «Прикладные информационные системы в экономике и финансах» по направлению подготовки 09.03.03 – Прикладная информатика, ОП «Прикладные информационные системы в экономике и финансах».

### 4. Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся

Очная форма обучения / очно-заочная форма обучения / заочная форма обучения (ИОО)

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 6/7/8 (в часах)
<b>Общая трудоемкость дисциплины</b>	<b>3/108</b>	<b>108</b>
<i>Контактная работа – Аудиторные занятия</i>	<i>50/34/12</i>	<i>50/34/12</i>
<i>Лекции</i>	<i>16/16/4</i>	<i>16/16/4</i>
<i>Семинары, практические занятия</i>	<i>34/18/8</i>	<i>34/18/8</i>
<b>Самостоятельная работа</b>	<b>58/74/96</b>	<b>58/74/96</b>
Вид текущего контроля	Контрольная работа	Контрольная работа
Вид промежуточной аттестации	зачет	зачет

## 5. Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий

### 5.1. Содержание дисциплины

#### Тема1. Подключение и преобразование внешних данных

##### 1.1. Основные понятия, связанные с визуализацией данных и визуального восприятия

История визуализации данных. Формы визуального представления информации. Основные концепции визуального восприятия графиков. Сервисы для визуализации данных (Инструмент для работы со сквозной аналитикой Google Data Studio, Интерактивная визуализация данных и бизнес-аналитики Yandex DataLens, Создание интерактивных диаграмм, карт, таблиц)

##### 1.2. Источники данных.

Современные программные средства обработки и анализа данных позволяют работать с десятками источниками данных. Будут рассмотрены подключение к базам данных (sqlite, postgre), таблицам Excel, LibreOffice Calc, файлам Text/CSV, XML, pdf, doc. Загрузка и сохранение информации в разных форматах файлов.

## Тема 2. Работа с СУБД SQLite

### 1.1. Создание базы данных, создание таблиц.

Будут рассмотрены типы данных SQLite, методика создания БД, таблиц, организация связей между таблицами.

### 1.2. Работа с записями в БД

Наполнение таблиц данными. Выборка данных. Изменение и удаление данных. Работа с одиночными записями и множеством записей.

### 1.3. Организация сложных SQL запросов

Вложенные запросы. Связные запросы. Операции для работы со множеством записей (объединение, пересечение, декартовы операции и т.д.).

## Тема 3. Выбор визуализации данных для различных типов данных

### 1.1. Разбор типов данных и выбор подходящих под них графиков.

### 1.2. Работа с количественными и качественными данными.

### 1.3. Совмещение различных типов графиков на одном полотне.

## Тема 4. Улучшение визуализаций

### 1.1. Предварительная обработка данных для улучшения визуализации.

Сглаживание, фильтрация данных.

### 1.2. Выделение и анализ характерных особенностей в данных.

Тренд. Регрессия.

## 5.2. Учебно – тематический план

очная форма обучения / очно-заочная форма обучения /  
заочная форма обучения (ИОО)

п/п	Наименование тем (разделов) дисциплины	Трудоемкость в часах					Формы текущего контроля успеваемости
		Всего	Контактная работа - Аудиторная работа			Самостоятельная работа	
			Общая, в т.ч.:	Лекции	Семинары, практические занятия		
1	Подключение и преобразование внешних данных	28/28/23	12/8/3	4/4/1	8/4/2	16/20/20	Собеседования по домашним заданиям. Проверка заданий.

2	Работа с СУБД SQLite	18/16/15	8/4/3	2/2/1	6/2/2	10/12/12	Собеседования по домашним заданиям. Проверка заданий.
3	Выбор визуализации данных для различных типов данных	30/30/33	14/10/3	4/4/1	10/6/2	16/20/30	Собеседования по домашним заданиям. Проверка заданий.
4	Улучшение визуализаций	32/34/37	16/12/3	6/6/1	10/6/2	16/22/34	Собеседования по домашним заданиям. Проверка заданий.
	В целом по дисциплине	108	50/34/12	16/16/4	34/18/8	58/74/96	Согласно учебному плану: контрольная работа
	Итого в %		46/31/11	32/47/33	68/53/67	54/69/89	

### 5.3. Содержание семинаров, практических занятий

Наименование тем (разделов) дисциплины	Перечень вопросов для обсуждения на семинарах, практических занятиях, рекомендуемые источники из разделов 8,9 (указывается раздел и порядковый номер источника)	Формы проведения занятий
Подключение и преобразование внешних данных	1. История визуализации данных. 2. Формы визуального представления информации. 3. Основные концепции визуального восприятия графиков. 4. Сервисы для визуализации данных (Инструмент для работы со сквозной аналитикой Google Data Studio, Интерактивная визуализация данных и бизнес-аналитики Yandex DataLens, Создание интерактивных диаграмм, карт, таблиц) <i>Рекомендуемые источники литературы: 8. 1, 2</i>	Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Работа с СУБД SQLite	1. Создание базы данных, создание таблиц. 2. Будут рассмотрены типы данных SQLite, методика создания БД, таблиц, организация связей между таблицами. 3. Работа с записями в БД 4. Наполнение таблиц данными. Выборка данных. Изменение и удаление данных. Работа с одиночными записями и множеством записей. 5. Организация сложных SQL запросов 6. Вложенные запросы. Связные запросы. Операции для работы со множеством записей (объединение, пересечение, декартовы операции и т.д.). <i>Рекомендуемые источники литературы: 8. 4,5,6</i>	Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений



Выбор визуализации данных для различных типов данных	<ol style="list-style-type: none"> <li>1. Разбор типов данных и выбор подходящих под них графиков.</li> <li>2. Работа с количественными и качественными данными.</li> <li>3. Совмещение различных типов графиков на одном полотне.</li> </ol> <p><i>Рекомендуемые источники литературы: 8. 1-3,5-7</i></p>	Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений
Улучшение визуализаций	<ol style="list-style-type: none"> <li>1. Предварительная обработка данных для улучшения визуализации.</li> <li>2. Сглаживание, фильтрация данных.</li> <li>3. Выделение и анализ характерных особенностей в данных.</li> <li>4. Тренд. Регрессия.</li> </ol> <p><i>Рекомендуемые источники литературы: 8. 2,3,5-7</i></p>	Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений

## 6. Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине

### 6.1. Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы

Наименование тем (разделов) дисциплины	Перечень вопросов, отводимых на самостоятельное освоение	Формы внеаудиторной самостоятельной работы
Подключение и преобразование внешних данных	Загрузка данных из файлов формата xls, csv, txt	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Работа с СУБД SQLite	Создание БД. Создание таблиц. Наполнение таблиц данными вручную или из внешних файлов. Вывод данных на экран или сохранение в файл.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
Выбор визуализации данных для различных типов данных	Разбор типов данных и выбор подходящих под них графиков. Работа с количественными и качественными данными. Совмещение различных типов графиков на одном полотне. Инструменты визуализации кода программных продуктов.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.

Улучшение визуализаций	Предварительная обработка данных для улучшения визуализации. Сглаживание, фильтрация данных. Выделение и анализ характерных особенностей в данных. Тренд. Регрессия.	Работа с учебной литературой. Решение типовых задач. Разбор вопросов по теме занятия. Выполнение домашних заданий к каждому занятию.
------------------------	---	--

## 6.2. Перечень вопросов, заданий, тем для подготовки к текущему контролю

### *Примерные задания контрольной работы по теме 1. Подключение и преобразование внешних данных*

1. Импортируйте
  - 1.1. Данные из файла xls
  - 1.2. Данные из файла формата csv
  - 1.3. Данные из файла txt
2. Выведите первые 10 записей на экран. Предварительно отсортируйте по одному из столбцов. В импортированных данных найдите максимальное, минимальное среднее значения в числовых столбцах. Для максимума и минимума выведите индекс (индексы) данных элементов.
3. Напишите программу, которая считывает данные из файла CSV и выводит на экран среднее значение чисел в столбце "Цена".
4. Напишите программу, которая импортирует данные из файла Excel и выводит на экран количество строк в нем.
5. Напишите программу, которая считывает данные из файла TXT и выводит на экран общее количество слов, включая повторения.

### *Примерные задания контрольной работы по теме 2. Работа с СУБД SQLite*

1. Создайте БД с таблицей, в которую нужно занести последние 100 строк из анализируемых данных. Структура таблицы должна совпадать со структурой CSV файла (количество и название столбцов).

2. Создать БД, которая содержит информацию об учебных дисциплинах, заполнить ее данными. В программе разработать функции, которые позволяют удалять строки из таблицы в БД в соответствии с номером строки.

3. Реализовать программу, в которой пользователь может добавлять данные в БД, вводя необходимые данные в консоли. Реализовать программу, которая выводит информацию из Базы данных. База данных состоит из одной таблицы с Четырьмя полями разных типов.

4. Создайте БД с таблицей, в которую нужно занести первые 100 строк из анализируемых данных. Данные взять из файла CSV из открытого источника. Структура таблицы должна совпадать со структурой CSV файла (количество и название столбцов).

5. Создать Базу данных DataBase, с таблицей Students с полями (id, Name, age, INN). Дать соответствующие типы полям. Заполнить строки с консоли значениями. Не менее 10 записей. Считать данные из БД и вывести все сразу на экран.

6. Создать Базу данных DataBase, с таблицей Students с полями (id, Name, age, INN). Дать соответствующие типы полям. Заполнить строки с консоли значениями. Не менее 10 записей. Считать данные из БД и построить диаграмму по столбцу AGE.

7. Создать Базу данных DataBase, с таблицей Students с полями (id, Name, age, adress). Дать соответствующие типы полям. Заполнить строки с консоли значениями. Не менее 14 записей. Считать данные из БД и записать их в файл CSV.

8. Создать Базу данных DataBase, с таблицей Students с полями (id, Name, Surnave, age, group). Дать соответствующие типы полям. Заполнить строки с консоли значениями. Не менее 10 записей. Считать из БД информацию только о тех студентов, чей возраст = 20. Считанные данные вывести на экран.

9. Создать БД, которая содержит информацию об учебных дисциплинах, заполнить ее данными. В программе разработать функции, которые позволяют изменять значения полей в строках в соответствии с номером строки.

10. Создать БД, которая содержит информацию об учебных дисциплинах, заполнить ее данными. Считать данные из БД и вывести их на экран. Записать первые 3 строки из БД в файл CSV.

***Примерные задания контрольной работы по теме 3. Выбор визуализации данных для различных типов данных***

1. Постройте временные ряды и их визуализацию для данных, загруженных с файла "consumers-price-index-march-2022-seasonally-adjusted.csv", столбец "Data\_value". Должно быть построено 2 различных типа графика для анализируемых данных.

2. Постройте временные ряды и их визуализацию для данных, полученного из любого CSV-файла по двум столбцам. Должны быть использованы столбчатые диаграммы.

3. Постройте графики функций  $y_1 = x^2 + x$  и  $y = x^3 + x$ . Графики строятся на одном поле. Должна быть показана легенда.  $x$  изменяется от -10 до 25.

4. Постройте временные ряды и их визуализацию для данных, загруженных с файла "consumers-price-index-march-2022-seasonally-adjusted.csv", столбец "Data\_value". Должно быть построено 2 различных типа графика для анализируемых данных. Каждое числовое значение должно быть увеличено в два раза.

5. Из файла electronic-card-transactions-april-2022-csv-tables.csv считать столбец Data\_Value, вывести значения в виде графика на экран. Тип линии - пунктир, цвет - красный. Считанные значения записать в базу данных. Считать из базы данных первые 10 значений и вывести на экран.

6. Из файла electronic-card-transactions-april-2022-csv-tables.csv считать столбец Data\_Value, вывести значения в виде графика на экран. Тип линии - пунктир, цвет - синий. Считанные значения записать в базу данных. При этом каждое значение увеличить на 1. Считать из базы данных первые 20 значений и построить по ним график.

7. Из файла electronic-card-transactions-april-2022-csv-tables.csv считать столбец Data\_Value, вывести значения в виде графика на экран. Тип линии - пунктир, цвет - черный. Первые 30 значений сохранить в CSV-файле out.csv.

8. Из файла `electronic-card-transactions-april-2022-csv-tables.csv` считать столбец `Data_Value`, вывести значения в виде графика на экран. Тип линии - пунктир, цвет - черный. Затем построить график по этим же значениям, умноженным в 2 раза, но цвет линий - зеленый.

9. Открыть файл `wholesale-trade-survey-dec-2021-quarter-csv.csv`. Считать столбец `Data_Value` (первые 200 значений). Сохранить его содержимое в Базе Данных. Второй столбец - порядковый номер значения. С консоли изменить Строку 50 - обнулить значение.

10. Открыть файл `wholesale-trade-survey-dec-2021-quarter-csv.csv`. Считать столбец `Data_Value` (первые 100 значений). Сохранить его содержимое в Базе Данных. Второй столбец - порядковый номер значения. Считать данные из БД и вывести на экран.

11. Открыть файлы `wholesale-trade-survey-dec-2021-quarter-csv.csv` и `electronic-card-transactions-april-2022-csv-tables.csv`. Считать столбец `Data_Value` (первые 100 значений). Вывести считанные значения на графике, выводить данные Одновременно на одном поле.

12. Открыть файлы `wholesale-trade-survey-dec-2021-quarter-csv.csv` и `electronic-card-transactions-april-2022-csv-tables.csv`. Перемножить соответствующие значения двух считанных столбцов. Результат вывести на экран в виде столбца данных. При выводе подписать каждую строку порядковым номером.

13. Открыть файлы `wholesale-trade-survey-dec-2021-quarter-csv.csv` и `electronic-card-transactions-april-2022-csv-tables.csv`. Перемножить соответствующие значения двух считанных столбцов. Вывести на экран в виде графиков исходные данные и их произведение.

#### ***Примерные задания контрольной работы по теме 4. Улучшение визуализаций***

1. Считать данные из файла CSV. Построить график по данным временным рядам. Найти и удалить шумовые всплески в последовательности данных. Построить линию тренда. Спрогнозировать развитие процесса.

2. Считать данные из файла CSV. Построить график по данным временным рядам. Сгладить временные ряды методом скользящей средней и экспоненциальным сглаживанием. Восстановить пропущенные данные с помощью линейной и полиномиальной регрессии. Вычислить точность восстановления.
3. В файле имеются данные о совокупных личных расходах; располагаемых личных доходах; расходах на табак для США на период с 1959 по 1983 годы. Оцените множественную регрессию между регрессандом (эндогенной переменной) Var1 и регрессорами (экзогенными переменными) Var2, Var3 и Var4 используя данные за 25 лет. Дайте интерпретацию коэффициентам регрессии. Исследуйте степень корреляционной зависимости между переменными.
4. Требуется произвести анализ процесса разработки определенного программного проекта, используя статистические данные. Задача заключается в разработке программы, которая будет строить линию тренда, отображающую динамику процесса разработки и предоставлять прогнозируемую дату окончания проекта на основе имеющихся данных.
5. Проанализировать статистические данные о ходе разработки программного продукта и отобразить на графике участки времени, в течение которых работы с приостанавливались или велись с задержками. Для выполнения задания необходимо разработать программу, которая проведет анализ данных, выявит перерывы и отклонения в ходе работ, и затем создаст визуальное представление в виде графиков или диаграмм с использованием удобных средств визуализации данных.

*Критерии балльной оценки различных форм текущего контроля успеваемости содержатся в соответствующих методических рекомендациях Департамента анализа данных и машинного обучения.*

*Полное прохождение одного из онлайн-курсов, «Визуализация данных: от скучных графиков к интерактивным дашбордам» (Нетология) или «Дизайн инфографики и визуализации данных» (Stepik), будет учитываться при контроле текущей успеваемости.*

## 7. Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине

Перечень компетенций с указанием индикаторов их достижения в процессе освоения образовательной программы содержится в разделе 2. *«Перечень планируемых результатов освоения образовательной программы (перечень компетенций) с указанием индикаторов их достижения и планируемых результатов обучения по дисциплине».*

**Типовые контрольные задания или иные материалы, необходимые для оценки индикаторов достижения компетенций, умений и знаний**

Наименование компетенции	Наименование индикаторов достижения компетенции	Результаты обучения (умения и знания), соотнесенные с индикаторами достижения компетенции	Типовые контрольные задания
<b>ОП «Инженерия данных»</b>			
ПКН -5. Способность участвовать в документальном сопровождении разработки ИТ в рамках проектных групп, применять средства автоматизации управления проектами ИТ	1. Демонстрирует знание основ версионирования и управления изменениями при разработке ПО. Использует системы контроля версий для ведения совместной разработки.	<b>Знать:</b> приемы визуального представления хода разработки программных продуктов.	1. Определить основные этапы разработки программного продукта и задачи, которые необходимо выполнить на каждом этапе. 2. Определить зависимости между задачами и представить их в графической форме с помощью диаграммы PERT или Ганта. 3. Определить продолжительность каждой задачи, использование временных интервалов и установить сроки выполнения каждой задачи.

		<b>Уметь:</b> использовать системы контроля версий для ведения совместной разработки приложений визуализации данных, работать с базами данных для хранения и извлечения необходимых данных	4. Создать отчет о проделанной работе и представить его в письменном виде, включая диаграммы, графики и другие визуальные представления.
	2. Демонстрирует знание основ тестирования программного обеспечения, умение создавать автоматизированные модульные и интеграционные тесты	<b>Знать:</b> способы обработки результатов тестирования для их наглядного представления	1. Получить данные из источников данных (текстовый файл, база данных) 2. Проанализировать и числовые значения в столбцах таблиц, найти основные статистические характеристики. Сделать выводы по полученным данным.
		<b>Уметь:</b> визуализировать результаты тестирования программного обеспечения	3. Найти коэффициенты корреляции между рядами данных. 4. Построить линию тренда с помощью полиномиальной и линейной регрессии. Составить прогнозные оценки.
	3. Готовит документацию к программе, коммуницирует в пределах группы разработки и за ее границами о значимых аспектах информационной системы и информационной инфраструктуры в письменной и устной форме.	<b>Знать:</b> требования к составу программной документации	1. Определить требования, функциональные характеристики и особенности программного продукта по заданной тематике
		<b>Уметь:</b> визуализировать структуру программного продукта и статистику результатов его работы	2. Создать диаграмму вариантов использования, демонстрирующую взаимодействие пользователя и программного продукта. 3. Описать структуру программы с помощью диаграммы классов или другой визуализации данных.



			<p>4. Разработать диаграмму последовательности, диаграмму активности или другую визуализацию, представляющую процессы, связанные с выполнением программного продукта.</p> <p>5. Сделать выводы о преимуществах использования каждого метода визуализации данных в программной документации.</p>
	<p>4. Демонстрирует знание жизненного цикла информационных систем, участвует в процессе разработки ПО на разных этапах</p>	<p><b>Знать:</b> способы визуального представления жизненных циклов информационных систем</p>	<p>1. Выберите жизненный цикл определенной информационной системы.</p> <p>2. Составьте таблицу с этапами жизненного цикла, и для каждого этапа составьте таблицу с данными и показателями (например, бюджет, уровень сложности, качество и т.д.).</p>
		<p><b>Уметь:</b> визуализировать этапы жизненного цикла информационных систем, строить тренд состояний процесса разработки программного продукта</p>	<p>3. Используя библиотеку Pandas, выполните выгрузку данных в формат CSV.</p> <p>4. Используя библиотеку Matplotlib постройте графики для каждого этапа жизненного цикла, демонстрирующие динамику изменения показателей, а также связь между ними.</p> <p>5. Используйте Seaborn, чтобы построить графики, демонстрирующие сводную информацию по всем этапам жизненного цикла.</p>
<p>ПКП-2. Способность разрабатывать, согласовывать и управлять исполнением технического задания и технического проекта с использованием технологий больших данных</p>	<p>1. Работает со стандартами, в том числе адаптирует стандарты для специфических требований больших данных.</p>	<p><b>Знать:</b> стандарты и их возможности адаптации под специфические требования больших данных.</p>	<p>1. Определите специфические требования к большому набору данных (например, количество, разнообразность, структура, качество).</p> <p>2. Используя библиотеку Pandas, выполните загрузку данных в формате CSV.</p> <p>3. Произведите первоначальный анализ данных, включающий</p>

			<p>обработку и фильтрацию, при необходимости.</p> <p>4. Разработайте предложения по визуализации для наборов данных в соответствии с поставленными требованиями, включая выбор типа графиков, цветовой гаммы, масштабов и др.</p>
		<p><b>Уметь:</b> визуализировать наборы данных в соответствии с поставленными требованиями с целью их дальнейшего анализа</p>	<p>5. Используя библиотеки Matplotlib и Seaborn, постройте графики, демонстрирующие особенности и закономерности данных, выделяющие наиболее важные параметры.</p>
	<p>2. Разрабатывает технические задания и технические проекты для технологий больших данных.</p>	<p><b>Знать:</b> методы составления технических заданий на реализацию программных продуктов для визуализации данных</p>	<p>1. Определить цели и задачи программного продукта для визуализации больших наборов данных и его потенциальных пользователей.</p> <p>2. Графически описать алгоритм работы программного продукта</p> <p>3. Выберите библиотеки и средства для разработки продукта, укажите причины их выбора и обоснуйте их использование.</p>
		<p><b>Уметь:</b> составлять технических заданий на реализацию программных продуктов для визуализации данных и проводить их верификацию поставленным требованиям</p>	<p>4. Укажите требования к входным данным, форматам данных и методам хранения и обработки данных.</p> <p>5. Создайте прототип программного продукта, который позволит оценить функциональность продукта и соответствие требованиям заказчика.</p> <p>6. Разработайте систему тестирования продукта и опишите критерии оценки эффективности и качества продукта.</p>

	3. Реализует управление рабочими проектами технологической инфраструктуры больших данных.	<b>Знать:</b> как возможно визуализировать процесс управления программными проектами в рамках технологической инфраструктуры больших данных	1. Опишите, как можно визуально представить календарный график выполнения программного проекта по реализации автоматизированной обработки больших данных. 2. Составьте перечень ключевых этапов проекта и определите ключевые показатели достижения успеха на каждом этапе.
		<b>Уметь:</b> визуализировать процесс управления программными проектами в рамках технологической инфраструктуры больших данных	3. Создайте или импортируйте набор данных в формате cvs. 4. Примените выбранные методы визуализации данных для презентации процесса управления проектом и его текущего статуса. Используйте инструменты визуализации данных, такие как DataLens или Power BI, для создания интерактивных дашбордов и графиков. 5. Определите аномалии и вызывающие наибольшее беспокойство тренды в данных и выведите их в отчет. 6. Подготовьте отчет о проделанной работе, включающий анализ процесса управления проектами, описание созданного набора данных, графиков и дашбордов, а также выводов об аномалиях и трендах в данных.
ОП «Прикладные информационные системы в экономике и финансах»			
ПКН-8. Способность использовать современные информационные системы для решения задач предметной области, в том числе отечественного	1. Демонстрирует знания об основных информационных технологиях и программных средствах, позволяющих их использовать.	<b>Знать:</b> основные информационные технологии и программные средства визуализации данных	Используя необходимое программное обеспечение, выполните: 1. Создайте таблицу, содержащую данные о сотрудниках компании, их должностях, продолжительности рабочей недели и соответствующей производительности.

производства			<p>2. Разбейте сотрудников на группы в соответствии с продолжительностью рабочей недели для уточнения влияния продолжительности на производительность.</p> <p>3. Создайте диаграмму, отображающую зависимость производительности сотрудников от продолжительности рабочей недели, чтобы проанализировать эффекты влияния на производительность при различных уровнях интенсивности труда.</p>
		<p><b>Уметь:</b> использовать современные информационные системы для решения профессиональных задач в области визуализации данных</p>	<p>4. Используя различные типы визуализации выявите зависимости между продолжительностью рабочей недели и производительностью сотрудников.</p> <p>5. Сформируйте сводную таблицу, чтобы выявить связь между продолжительностью рабочей недели и соответствующей производительностью для каждого отдельного сотрудника в компании.</p> <p>6. Оцените результаты работы программы, сделайте выводы о применении методов визуализации для решения конкретной задачи и о их эффективности в сравнении с другими методами визуализации данных.</p>
	<p>2. Рационально выбирает информационные технологии и реализующие их программные средства, в том числе, с учетом страны</p>	<p><b>Знать:</b> функциональные возможности, отличительные особенности и сферы применения различных программные</p>	<p>Требуется разработать отчет, в котором будут проводиться сравнительные исследования программных средств визуализации данных:</p> <p>1. Определите целевую аудиторию и требования к созданию визуализаций,</p>

	происхождения программных средств.	средств визуализации данных	<p>учитывая функциональные возможности, отличительные особенности и страну происхождения программных средств.</p> <p>2. Выберите несколько программных средств визуализации данных, которые могут наилучшим образом соответствовать требованиям и условиям задачи.</p> <p>3. Оцените каждое из выбранных программных средств визуализации данных с точки зрения их функциональных возможностей, отличительных особенностей и страны происхождения.</p> <p>требуемым критериям.</p>
		<p><b>Уметь:</b> выбирать программные средства визуализации данных с учетом из функциональных возможностей и отличительных особенностей, в том числе, с учетом страны происхождения программных средств</p>	<p>4. Создайте несколько визуализаций одной и той же информации используя выбранные программные средства и сравните результаты.</p> <p>5. Проанализируйте полученную информацию на предмет соответствия требованиям и условиям задачи.</p> <p>6. Оформите отчет, содержащий объективную информацию о сравнении программных средств визуализации данных, сделайте выводы о том программном средстве, которое наилучшим образом соответствует требованиям и условиям задачи, а также функциональных возможностей, отличительных особенностей и страны происхождения.</p>

	3.Использует современные информационные технологии и программные средства при решении задач разработки программного обеспечения для экономических и финансовых приложений	<b>Знать:</b> возможности современных программных средств визуализации данных и методику их использования при решении задач профессиональной деятельности	1. В соответствии с исходными данными для визуализации и техническим заданием разработайте концептуальное описание программного обеспечения для визуализации данных на основе существующих методов и технологий. 2. Определите требования к функциональности, пользовательскому интерфейсу, типам диаграмм и графиков, совместимости программного обеспечения с системами, а также выберите язык программирования и фреймворк.
		<b>Уметь:</b> разрабатывать программное обеспечение визуализации данных для экономических и финансовых приложений	3. Разработайте и протестируйте создание данных для отображения, включающие данные из финансового сектора, экономические данные, данные транзакций и т.п. 4. Разработайте и протестируйте функционал программного обеспечения по созданию визуализации данных, включая различного рода диаграммы, графики, тепловые карты и др. с различными параметрами с целью оперативного анализа данных 5. Обеспечьте программное обеспечение возможностями импорта и экспорта данных для проведения дополнительного анализа.

### ***Примеры практико-ориентированных (ситуационных) заданий***

1. Возьмите основные показатели московского бюджета за три последних года в формате CSV, импортируйте данные в своей программе на Python. Вычислите средние, пользуясь функциями Python. Постройте график изменения показателей во времени.

2. Выберите три произвольные ценные бумаги и посмотреть их котировки на Московской фондовой бирже в последние три недели. Данные из файла CSV занесите в БД SQLite. Вычислите средние значения по каждой неделе. Постройте гистограмму.

3. Откройте БД с данными о заболеваемости за последний год. Постройте гистограмму.

4. Создайте БД. Создайте таблицу, которая будет содержать информацию о студентах университета. Занесите в таблицу данные. Постройте на основе этих данных гистограмму распределения количества студентов по курсам.

5. Напишите программу на Python, чтобы нарисовать линейные графики финансовых данных Alphabet Inc. в период с 3 октября 2016 года по 7 октября 2016 года.

Пример финансовых данных (fdata.csv):

Дата, Open, High, Low, Close

10-03-16,774.25,776.065002,769.5,772.559998

10-04-16,776.030029,778.710022,772.890015,776.429993

10-05-16,779.309998,782.070007,775.650024,776.469971

10-06-16,779,780.47998,775.539978,776.859985

10-07-16,779.659973,779.659973,770.75,775.080017.

6. Создайте набор данных в формате csv и загрузите его в Python. Используя библиотеки pandas и matplotlib, постройте график распределения данных и произведите анализ достоверности данных. Примените различные методы визуализации данных для изучения связей между данными.

7. Используя Yandex datalens, создайте интерактивную карту по выбранной теме, используя геоданные в формате csv. Примените различные методы визуализации данных, такие как тепловая карта и маркеры, для отображения различных аспектов данных.

8. Используя библиотеку Python seaborn, произведите анализ данных временных рядов. Отобразите эти данные на графике, с помощью линейных и точечных диаграмм. Изучите взаимосвязи между различными переменными и проанализируете их корреляцию.

9. Используя Python и библиотеки математических функций, проанализируйте изменение параметров моделей и динамику значений. Постройте графики, где на одной странице отображены графики зависимости значения зависимых и независимых переменных.

10. Используя Python pandas и seaborn, реализуйте приложения анализа данных, с помощью которого необходимо проанализировать основные статистические показатели, построить графики распределения данных и корреляционный анализ. В результате, определите наиболее значимые переменные и их взаимосвязи и выведите их в отчет.

### ***Примерные вопросы для подготовки к зачету***

1. Информационный дизайн как междисциплинарная область исследований и форма коммуникации. Основные вехи истории и возникновение термина.
2. Средства визуализации данных. Сравнительный анализ.
3. Информационный дизайн и инфографика. Сходства и различия. Сферы применения в экономической науке.
4. Классификационные стратегии и методы репрезентации. Особенности визуализации при многомерном и многовариантном анализе.
5. Большие данные. Особенности статистического анализа при невозможности использования всего объема данных. Параметрический анализ.
6. Типы наглядного представления численных данных. Наиболее популярные пакеты программ, использующиеся для визуализации численных данных. Краткая характеристика и особенности.
7. Работа с массивами данных с помощью языка Python. Общее описание области применимости (типов задач, для решения которых он пригоден).
8. Табличное и графическое представление данных с помощью средств языка Python
9. Понятие баз данных. СУБД. Реляционные базы данных. СУБД SQLite.
10. Работа с базами данных с помощью языка Python
11. Язык SQL и особенности его синтаксиса в Python и SQLite.
12. Отображение данных с помощью графиков в программах на языке



программирования Python. Параметры Matplotlib.

13. Файлы CSV. Структура. Загрузка и сохранение.
14. Понятие больших данных. Особенности их вывода, представления, визуализации.
15. Применение методов визуализации данных.
16. Методы визуализации, необходимые для интерпретации полученных результатов исследований.
17. Основные тенденции в области визуализации данных.
18. Как осуществляется нахождение шумов, выбросов в данных? На что они влияют?
19. Работа с файлами в Python. Методы `open()`, `close()`, `read()`, `write()`.
20. Строковые операции с Python. Назначение, примеры использования
21. Каковы основные преимущества визуализации данных при работе с большими объемами информации? Какие библиотеки Python обычно используются для визуализации данных?
22. Как `Yansex datalens` может использоваться для создания интерактивных отчетов и дашбордов на основе импортированных данных из различных источников? Возможно ли использовать данные, хранимые на сервере для этой цели?
23. Каковы преимущества использования графиков, диаграмм и тепловых карт, по сравнению с другими типами визуализации? В каких случаях наиболее целесообразно их использование?
24. Что такое импорт и экспорт данных? Какой формат данных наиболее распространен для хранения табличных данных? Как создать файл в формате csv с помощью Python?
25. Какая роль задействованных переменных в визуализации данных? Каким образом могут быть использованы номинативные и количественные переменные при построении графиков и диаграмм?
26. Каковы факторы, влияющие на результаты визуализации данных?

## **8. Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины**

### **Основная:**

1. Исаков, В. Б. Говорите языком схем: краткий справочник / В. Б. Исаков; НИУ ВШЭ. – Москва : Норма, 2019. - 144 с. - Текст : непосредственный. - То же. – 2023. - ЭБС ZNANIUM.com. – URL: <https://znanium.com/catalog/product/1915379> (дата обращения: 12.05.2023). – Текст: электронный.

2. Представление и визуализация результатов научных исследований : учебник / О. С. Логунова, П. Ю. Романов, Л. Г. Егорова [и др.] ; под ред. О. С. Логуновой. — Москва : ИНФРА-М, 2020. — 156 с. - ЭБС ZNANIUM.com. - URL: <https://new.znanium.com/catalog/product/1056236> (дата обращения: 12.05.2023). - Текст : электронный.

3. Полищук, Ю. В. Базы данных и их безопасность : учебное пособие / Ю. В. Полищук, А. С. Боровский. — Москва : ИНФРА-М, 2022. — 210 с. — (Высшее образование: Специалитет). - ЭБС ZNANIUM.com. - URL: <https://znanium.com/catalog/product/1811408> (дата обращения: 12.05.2023). – Текст : электронный.

5. Жуков, Р. А. Язык программирования Python: практикум : учебное пособие / Р. А. Жуков. — Москва : ИНФРА-М, 2022. — 216 с. — (Высшее образование: Бакалавриат). - ЭБС ZNANIUM.com. - URL: <https://znanium.com/catalog/product/1689648> (дата обращения: 12.05.2023). - Текст : электронный.

### **Дополнительная:**

6. Шелудько, В. М. Основы программирования на языке высокого уровня Python : учебное пособие / В. М. Шелудько. – Ростов-на-Дону ; Таганрог : Южный федеральный университет, 2017. – 147 с. : ил. – ЭБС Университетская библиотека ONLINE. - URL: <https://biblioclub.ru/index.php?page=book&id=500056> (дата обращения: 12.05.2023). – Текст : электронный.

7. Статистика (компьютеризированный курс) : учебник / Д. А. Ловцов, М. В. Богданова, А. В. Лобан, Л. С. Паршинцева ; под. ред. Д. А. Ловцова. - Москва : РГУП, 2020. - 400 с. - ЭБС ZNANIUM.com. - URL: <https://znanium.com/catalog/product/1689630> (дата обращения: 12.05.2023). – Текст: электронный.

**9. Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины**

1. <https://www.planetaexcel.ru/>
2. Электронная библиотека Финансового университета (ЭБ) <http://elib.fa.ru/>  
(<http://library.fa.ru/files/elibfa.pdf>)
3. Электронный учебный курс в электронной информационно-образовательной среде Финуниверситета <https://campus.fa.ru/course/view.php?id=12719>
4. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
5. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
6. Электронно-библиотечная система Znanium <http://www.znanium.com>
7. «Деловая онлайн библиотека» издательства «Альпина Паблишер» <http://lib.alpinadigital.ru/en/library>
8. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
9. Образовательная платформа "ЮРАЙТ" <https://urait.ru/>
10. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
11. Онлайн курс «Дизайн инфографики и визуализации данных» <https://stepik.org/course/70703/promo>
12. Онлайн курс «Визуализация данных: от скучных графиков к интерактивным дашбордам» <https://netology.ru/programs/analyst-bifree>

## 10. Методические указания для обучающихся по освоению дисциплины

Основные этапы работы студента по дисциплине *Методы визуализации данных*:

1. Предварительная ориентировка в подлежащем изучению учебном материале по программе.
2. Ознакомление с рекомендованной учебной литературой.
3. Слушание и конспектирование лекций, а также выполнение других видов учебной работы.
4. Планирование самостоятельной работы.
5. Обобщение и систематизация информации, почерпнутой из лекций и прочитанной литературы.
6. Выполнение контрольных работ.

Рекомендации по работе с учебным материалом:

1. Осознавайте наличный уровень полученных вами знаний.
2. В ситуации непонимания нужно выявить тот первичный уровень и факторы непонимания, которые стали препятствием понимания последующего.
3. Задавайте сами себе вопросы и пытайтесь ответить на них.

Рекомендации по работе на лекции и с лекционным материалом:

1. Основная задача на лекции – осмысление излагаемого в ней материала. Для этого необходимо слушать лекцию с самого начала, не упуская общих, ориентирующих в материале рассуждений и установок лектора.
2. Ведение записей на лекции важно и полезно для лучшего осмысливания материала, для сохранения информации, с целью ее дальнейшего использования.
3. Для облегчения записи рекомендуется применять сокращения повторяющихся терминов или хорошо известных понятий.

Рекомендации по работе с литературой:

1. Если возникли затруднения при разыскивании материала, по какому-либо конкретному вопросу, следует обратиться к предметному указателю, напечатанному, как правило, в конце каждого литературного источника.

2. Предметный указатель – это алфавитный список основных научных понятий (терминов), содержание которых раскрыто в книге, рядом с термином стоят числа, обозначающие номера страниц, на которых изложен материал, относящийся к данному понятию.

## **11. Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем**

### **11.1. Комплект лицензионного программного обеспечения:**

1. R 4.1.2, RStudio.
2. Python 3.8
3. Jupyter Notebook
4. Пакет офисных программ
2. Антивирус Kaspersky

### **11.2. Современные профессиональные базы данных и информационные справочные системы**

1. Информационно-правовая система «Гарант»
2. Информационно-правовая система «Консультант Плюс»
3. Электронная энциклопедия: <http://ru.wikipedia.org/wiki/Wiki>
4. Система комплексного раскрытия информации «СКРИН» - <http://www.skrin.ru/>
5. Язык программирования Python 3. <https://pythonworld.ru/>
6. Документация Yandex DataLens <https://cloud.yandex.ru/docs/datalens/>

### **11.3. Сертифицированные программные и аппаратные средства защиты информации**

- не предусмотрены

## **12. Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине**

Наличие аудитории, оснащенной компьютерной техникой и проектором, с возможностью подключения к сети «Интернет».