

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**

УТВЕРЖДАЮ

Проектор 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  
01.03.02 - Прикладная математика и информатика,  
ОП «Анализ данных»

*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

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	4
4.	Объем дисциплины (модуля) в зачетных единицах и в академических часах с выделением объема аудиторной (лекции, семинары) и самостоятельной работы обучающихся	4
5.	Содержание дисциплины, структурированное по темам (разделам) дисциплины с указанием их объемов (в академических часах) и видов учебных занятий	4
5.1.	Содержание дисциплины	4
5.2.	Учебно – тематический план	6
5.3.	Содержание семинаров, практических занятий	6
6.	Перечень учебно-методического обеспечения для самостоятельной работы обучающихся по дисциплине	7
6.1.	Перечень вопросов, отводимых на самостоятельное освоение дисциплины, формы внеаудиторной самостоятельной работы	7
6.2.	Перечень вопросов, заданий, тем для подготовки к текущему контролю	8
7.	Фонд оценочных средств для проведения промежуточной аттестации обучающихся по дисциплине	13
8.	Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины	18
9.	Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины	20
10.	Методические указания для обучающихся по освоению дисциплины	20
11.	Перечень информационных технологий, используемых при осуществлении образовательного процесса по дисциплине, включая перечень необходимого программного обеспечения и информационных справочных систем	22
12.	Описание материально-технической базы, необходимой для осуществления образовательного процесса по дисциплине	22

## 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
ПКП-5	Способность применять методы и инструменты анализа данных и машинного обучения при подготовке аналитического обоснования финансово-экономических решений	1. Демонстрирует знание методов и инструментов анализа данных и машинного обучения	<b>Знать:</b> концептуальные положения информационных технологий, позволяющих реализовывать модели машинного обучения и анализа данных <b>Уметь:</b> использовать результаты анализа данных для построения моделей машинного обучения
		2. Демонстрирует знание принципов подготовки аналитического обоснования финансово-экономических решений	<b>Знать:</b> способы использования результатов визуализации данных в задачах разработки моделей машинного обучения <b>Уметь:</b> использовать результаты визуализации данных в задачах разработки моделей машинного обучения
		3. Владеет навыками использования методов и инструментов анализа данных и машинного обучения при подготовке аналитического обоснования финансово-экономических решений	<b>Знать:</b> методы контроля результатов сборки развертывания многокомпонентных моделей машинного обучения <b>Уметь:</b> контролировать процесс сборки и развертывания многокомпонентных моделей машинного обучения с использованием методов визуализации данных

### 3. Место дисциплины в структуре образовательной программы

Дисциплина «Методы визуализации данных» является дисциплиной Цикла профиля (элективный) по направлению подготовки 01.03.02 – Прикладная математика и информатика, ОП «Анализ данных».

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

Вид учебной работы по дисциплине	Всего (в з/е и часах)	Семестр 6 (в часах)
<b>Общая трудоемкость дисциплины</b>	<b>3/108</b>	<b>108</b>
<b>Контактная работа – Аудиторные занятия</b>	<b>34</b>	<b>34</b>
<i>Лекции</i>	<i>16</i>	<i>16</i>
<i>Семинары, практические занятия</i>	<i>18</i>	<i>18</i>
<b>Самостоятельная работа</b>	<b>74</b>	<b>74</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

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	8	4	4	20	Собеседования по домашним заданиям. Проверка заданий.
2	Работа с СУБД SQLite	18	6	2	4	12	Собеседования по домашним заданиям. Проверка заданий.
3	Выбор визуализации данных для различных типов данных	30	8	4	4	22	Собеседования по домашним заданиям. Проверка заданий.
4	Улучшение визуализаций	32	12	6	6	20	Собеседования по домашним заданиям. Проверка заданий.
	В целом по дисциплине	108	34	16	18	74	Согласно учебному плану: контрольная работа
	Итого в %		31	47	53	69	

## 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	<ol style="list-style-type: none"> <li>1. История визуализации данных.</li> <li>2. Формы визуального представления информации.</li> <li>3. Основные концепции визуального восприятия графиков.</li> <li>4. Сервисы для визуализации данных (Инструмент для работы со сквозной аналитикой Google Data Studio, Интерактивная визуализация данных и бизнес-аналитики)</li> </ol>	Практикум по решению задач по тематике занятия в малых группах (2-4 студента) и коллективное обсуждение решений

	Yandex DataLens, Создание интерактивных диаграмм, карт, таблиц) <i>Рекомендуемые источники литературы: 8.1 - 8.6</i>	
Работа с СУБД SQLite	<ol style="list-style-type: none"> <li>1. Создание базы данных, создание таблиц.</li> <li>2. Будут рассмотрены типы данных SQLite, методика создания БД, таблиц, организация связей между таблицами.</li> <li>3. Работа с записями в БД</li> <li>4. Наполнение таблиц данными. Выборка данных. Изменение и удаление данных. Работа с одиночными записями и множеством записей.</li> <li>5. Организация сложных SQL запросов</li> <li>6. Вложенные запросы. Связные запросы. Операции для работы со множеством записей (объединение, пересечение, декартовы операции и т.д.).</li> </ol> <i>Рекомендуемые источники литературы: 8.1 - 8.6</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> <i>Рекомендуемые источники литературы: 8.1 - 8.6</i>	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> <i>Рекомендуемые источники литературы: 8.1 - 8.6</i>	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 *Подключение и преобразование внешних данных*

- Импортируйте
  - Данные из файла xls
  - Данные из файла формата csv
  - Данные из файла txt
- 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.
- 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 *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 студентах, 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 *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 *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
ПКП-5. Способность применять методы и инструменты анализа данных и машинного обучения при подготовке аналитического обоснования финансово-экономических решений	1. Демонстрирует знание методов и инструментов анализа данных и машинного обучения	<b>Знать:</b> концептуальные положения информационных технологий, позволяющих реализовывать модели машинного обучения и анализа данных	Требуется обработать набор данных по определенной теме с использованием методов визуализации и анализа данных. Опишите предметную область, обоснуйте выбор методов и инструментов визуализации данных, проанализируйте полученные результаты

		<p><b>Уметь:</b> использовать результаты анализа данных для построения моделей машинного обучения</p>	<p>Напишите программу, которая позволяет обработать набор данных по определенной теме с использованием методов анализа и визуализации данных. В качестве информационных технологий рассмотрите:</p> <ul style="list-style-type: none"> <li>- Основы статистики и теория вероятности;</li> <li>- Методы анализа данных, такие как кластерный анализ, корреляционный анализ и т.д.;</li> <li>- Методы визуализации данных в машинном обучении и анализе данных, такие как круговые диаграммы, сводные таблицы, графики.</li> </ul>
	2. Демонстрирует знание принципов подготовки аналитического обоснования финансово-экономических решений.	<p><b>Знать:</b> способы использования результатов визуализации данных в задачах разработки моделей машинного обучения</p>	<p>Требуется подготовить данные, необходимые для разработки модели машинного обучения:</p> <ul style="list-style-type: none"> <li>- Выберите набор данных, подходящий для построения модели. Предложите приемлемые источники данных или найдите наборы данных в Интернете. Импортируйте данные в свое приложение.</li> <li>- Разработайте процедуры анализа данных, прежде чем строить модель, включая очистку данных, масштабирование, визуализацию и т.д.</li> </ul>
		<p><b>Уметь:</b> использовать результаты визуализации данных в задачах разработки моделей машинного обучения</p>	<ul style="list-style-type: none"> <li>- Используя полученные визуализации, определите характерные особенности в данных, наличие отдельных кластеров в данных, тенденции изменения во времени</li> </ul>

	3. Владеет навыками использования методов и инструментов анализа данных и машинного обучения при подготовке аналитического обоснования финансово-экономических решений.	<b>Знать:</b> методы контроля результатов сборки и развертывания многокомпонентных моделей машинного обучения	Исследуйте процесс развертывания моделей машинного обучения с использованием методов визуализации данных: - Создайте прототип модели машинного обучения на выбранном способе классификации, регрессии или кластеризации. Возьмите пример датасета и начните обработку данных.
		<b>Уметь:</b> контролировать процесс сборки и развертывания многокомпонентных моделей машинного обучения с использованием методов визуализации данных	- Определите метрики качества модели, которые будут использованы для оценки ее предсказательной способности. Систематизируйте и выведите метрики в графическом представлении. - Проанализируйте данные на предмет наличия аномалий, используя методы визуализации.

### ***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 посмотрите 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 проанализируйте 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, используемые 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 Yandex 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: 16.05.2023). – 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: 16.05.2023). - 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: 16.05.2023). — Текст : электронный.

4. 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: 16.05.2022). - Текст : электронный.

**Дополнительная:**

5. 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: 16.05.2023). – Текст : электронный.

6. 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: 16.05.2023). – 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/>
3. (<http://library.fa.ru/files/elibfa.pdf>)
4. Электронный учебный курс в электронной информационно-образовательной среде Финуниверситета <https://campus.fa.ru/course/view.php?id=20786>
5. Электронно-библиотечная система BOOK.RU <http://www.book.ru>
6. Электронно-библиотечная система «Университетская библиотека ОНЛАЙН» <http://biblioclub.ru/>
7. Электронно-библиотечная система Znanium <http://www.znanium.com>
8. «Деловая онлайн библиотека» издательства «Альпина Паблишер» <http://lib.alpinadigital.ru/en/library>
9. Электронно-библиотечная система издательства «Лань» <https://e.lanbook.com/>
10. Образовательная платформа "ЮРАЙТ" <https://urait.ru/>
11. Научная электронная библиотека eLibrary.ru <http://elibrary.ru>
12. Онлайн курс «Дизайн инфографики и визуализации данных» <https://stepik.org/course/70703/promo>
13. 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».