

**МИНОБРНАУКИ РОССИИ**  
**федеральное государственное бюджетное образовательное учреждение**  
**высшего образования**  
**Ярославский государственный университет им. П.Г.Демидова**

**УЧЕБНО-МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ**  
**ПО ДИСЦИПЛИНЕ**

*СОВРЕМЕННЫЕ СРЕДСТВА АНАЛИЗА ДАННЫХ*

Направление подготовки (специальность):

02.04.02 Фундаментальная информатика и информационные технологии

Образовательная программа

Искусственный интеллект и компьютерные науки

**очная форма обучения**

Составитель:

**КРАСНОВ МИХАИЛ ВЛАДИМИРОВИЧ,**  
**К.Ф.-М.Н., ДОЦЕНТ Ф-ТА ИВТ**  
**ЯРГУ ИМ. П.Г. ДЕМИДОВА**

г. Ярославль

## **Перечень основной и дополнительной учебной литературы, необходимой для освоения дисциплины (модуля)**

### **Основная литература:**

- 1 Наследов А.Д. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных. - СПб.:Питер, 2013.
2. Бююль А., Цёфель П. SPSS: искусство обработки информации. Анализ статистических данных и восстановление скрытых закономерностей. – СПб.: ООО «ДиаСофтЮП», 2005.
3. Тюрин Ю.Н., Макаров А.А. Анализ данных на компьютере – М.: МЦНМО, 2016.
4. Малова А.С. Основы эконометрики в среде GRETL – М.: Проспект, 2017.

### **Дополнительная литература:**

1. Мاستицкий С.Э., Шитиков В.К. (2014) Статистический анализ и визуализация данных с помощью R. – Электронная книга, адрес доступа: <http://r-analytics.blogspot.com>
2. Chapman C., E. McDonnell Feit. R for Marketing Research and Analytics - Springer-Verlag, Switzerland, 2015.

## **Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине (модулю)**

Также для подбора учебной литературы рекомендуется использовать широкий спектр интернет-ресурсов:

1. Электронно-библиотечная система «Университетская библиотека online» ([www.biblioclub.ru](http://www.biblioclub.ru)) - электронная библиотека, обеспечивающая доступ к наиболее востребованным материалам-первоисточникам, учебной, научной и художественной литературе ведущих издательств (\*регистрация в электронной библиотеке – только в сети университета. После регистрации работа с системой возможна с любой точки доступа в Internet.).
2. Информационная система "Единое окно доступа к образовательным ресурсам" (<http://window.edu.ru/library>).

Целью создания информационной системы "Единое окно доступа к образовательным ресурсам" (ИС "Единое окно ") является обеспечение свободного доступа к интегральному каталогу образовательных интернет-ресурсов и к электронной библиотеке учебно-методических материалов для общего и профессионального образования.

Информационная система "Единое окно доступа к образовательным ресурсам" создана по заказу Федерального агентства по образованию в 2005-2008 гг. Главной разработчик проекта - Федеральное государственное автономное учреждение Государственный научно-исследовательский институт информационных технологий и телекоммуникаций (ФГАУ ГНИИ ИТТ "Информика") [www.informika.ru](http://www.informika.ru).

ИС "Единое окно" объединяет в единое информационное пространство электронные ресурсы свободного доступа для всех уровней образования в России. Разделы этой системы:

- Электронная библиотека – является крупнейшим в российском сегменте Интернета хранилищем полнотекстовых версий учебных, учебно-методических и научных материалов с открытым доступом. Библиотека содержит более 30 000 материалов, источниками которых являются более трехсот российских вузов и других образовательных и научных учреждений. Основу наполнения библиотеки составляют электронные версии учебно-методических материалов, подготовленные в вузах, прошедшие рецензирование и рекомендованные к использованию советами факультетов, учебно-методическими комиссиями и другими вузовскими структурами, осуществляющими контроль учебно-методической деятельности.
- Интегральный каталог образовательных интернет-ресурсов содержит представленные в стандартизированной форме метаданные внешних ресурсов, а также содержит описания полнотекстовых публикаций электронной библиотеки. Общий объем каталога превышает 56 000 метаописаний (из них около 25 000 - внешние ресурсы). Расширенный поиск в "Каталоге" осуществляется по названию, автору, аннотации, ключевым словам с возможной фильтрацией по тематике, предмету, типу материала, уровню образования и аудитории.
- Избранное. В разделе представлены подборки наиболее содержательных и полезных, по мнению редакции, интернет-ресурсов для общего и профессионального образования.
- Библиотеки вузов. Раздел содержит подборки сайтов вузовских библиотек, электронных каталогов библиотек вузов и полнотекстовых электронных библиотек вузов.

Для самостоятельного подбора литературы в библиотеке ЯрГУ рекомендуется использовать:

1. Личный кабинет ([http://lib.uniyar.ac.ru/opac/bk\\_login.php](http://lib.uniyar.ac.ru/opac/bk_login.php)) дает возможность получения on-line доступа к списку выданной в автоматизированном режиме литературы, просмотра и копирования электронных версий изданий сотрудников университета (учеб. и метод. пособия, тексты лекций и т.д.) Для работы в «Личном кабинете» необходимо зайти на сайт Научной библиотеки ЯрГУ с любой точки, имеющей доступ в Internet, в пункт меню «Электронный каталог»; пройти процедуру авторизации, выбрав вкладку «Авторизация», и заполнить представленные поля информации.

## 2. Электронная библиотека учебных материалов ЯрГУ

([http://www.lib.uni Yar.ac.ru/opac/bk\\_cat\\_find.php](http://www.lib.uni Yar.ac.ru/opac/bk_cat_find.php)) содержит более 2500 полных текстов учебных и учебно-методических материалов по основным изучаемым дисциплинам, изданных в университете. Доступ в сети университета, либо по логину/паролю.

## 3. Электронная картотека «Книгообеспеченность»

([http://www.lib.uni Yar.ac.ru/opac/bk\\_bookreq\\_find.php](http://www.lib.uni Yar.ac.ru/opac/bk_bookreq_find.php)) раскрывает учебный фонд научной библиотеки ЯрГУ, предоставляет оперативную информацию о состоянии книгообеспеченности дисциплин основной и дополнительной литературой, а также цикла дисциплин и специальностей. Электронная картотека «Книгообеспеченность» доступна в сети университета и через Личный кабинет.

### **Перечень ресурсов информационно-телекоммуникационной сети «Интернет», необходимых для освоения дисциплины (модуля), включая перечень информационных справочных систем (при необходимости)**

1. Электронная библиотека учебных материалов ЯрГУ ([http://www.lib.uni Yar.ac.ru/opac/bk\\_cat\\_find.php](http://www.lib.uni Yar.ac.ru/opac/bk_cat_find.php)).
2. Информационная система "Единое окно доступа к образовательным ресурсам" (<http://www.edu.ru> (раздел Учебно-методическая библиотека) или по прямой ссылке <http://window.edu.ru/library>).
3. Электронно-библиотечная система «Университетская библиотека online» ([www.biblioclub.ru](http://www.biblioclub.ru)).

### **Перечень информационных технологий, используемых при изучении дисциплины, включая программное обеспечение**

В процессе осуществления образовательного процесса используются:

- для формирования текстов материалов для промежуточной и текущей аттестации – программы Microsoft Office, издательская система LaTeX;
- компиляторы с высокоуровневых языков программирования;
- для поиска учебной литературы библиотеки ЯрГУ – Автоматизированная библиотечная информационная система "БУКИ-NEXT" (АБИС "Буки-Next").

### **Учебно-методические указания и рекомендации к изучению тем лекционных и практических занятий, самостоятельной работе студентов**

#### Содержание дисциплины

Наименование раздела	Название темы с кратким содержанием
----------------------	-------------------------------------

дисциплины	
Раздел 1.	<p>Тема 1. Работа с таблицами большой размерности в Excel</p> <p>1.1. Импорт и экспорт информации из электронных таблиц</p> <p>1.2. Связи между документами. Автофильтр и расширенный фильтр. Вычисляемые критерии</p> <p>1.3. Консолидация и объединение баз данных с помощью различных функций</p> <p>1.4. Использование условного форматирования для анализа информации</p> <p>1.5. Построение сводных таблиц с группированием данных, вычисляемыми полями и объектами</p>
Раздел 2.	<p>Тема 2. Использование Excel для анализа данных</p> <p>2.1. Применение статистических функций Excel</p> <p>2.2. Использование надстройки Excel «Пакет анализа»</p> <p>2.3. Статистические возможности Excel</p>
Раздел 3.	<p>Тема 3. Статистические пакеты Statistica и SPSS. Общая структура и интерфейс</p> <p>3.1. Общая структура и основные параметры пакетов SPSS и Statistica и способы взаимодействия с ними</p> <p>3.2. Особенности кодирования информации и формирования матрицы базы данных по результатам исследования</p> <p>3.3. Методы управления данными</p> <p>3.4. Возможности преобразования данных</p>
Раздел 4.	<p>Тема 4. Анализ основных статистик в пакетах Statistica и SPSS</p> <p>4.1. Одномерный описательный анализ</p> <p>4.2. Особенности анализа статистических взаимосвязей между различными типами переменных</p> <p>4.3. Корреляционный анализ</p>
Раздел 5	<p>Тема 5. Анализ непараметрических статистик в SPSS</p> <p>5.1. Одновыборочные тесты</p> <p>5.2. Тесты сравнения нескольких выборок</p> <p>5.3. Тесты для ранговых переменных</p>

Наименование раздела	Название темы с кратким содержанием
	5.4. Тесты для связанных выборок
Раздел 6	<p>Тема 6. Многомерный анализ в SPSS</p> <p>6.1. Факторный анализ</p> <p>6.2. Кластерный анализ</p> <p>6.3. Многомерное шкалирование</p>
Раздел 7.	<p>Тема 7. Прогнозирование и анализ временных рядов в SPSS</p> <p>7.1. Задание формата дат в SPSS. Графики sequence.</p> <p>7.2. Проверка нормальности распределения остатков. Принципы удаления тренда из данных.</p> <p>Определение формы тренда. Расчет индексов сезонности.</p> <p>7.3. Создание и редактирование временных рядов. Работа с экспоненциальным сглаживанием.</p> <p>Подбор параметра alpha</p> <p>7.4. Авторегрессия и модель скользящего среднего. ARIMA-модели.</p>

Примеры вопросов:

1. По готовой анкете в формате .doc - с помощью SPSS и SPSS Syntax создать матрицу

данных, добавить метки и значения переменных.

2. По готовой матрице данных в SPSS:

☐ Создать частотные и кросс-таблицы в готовых матрицах данных. Проверить

значимость, вывести интересные взаимосвязи.

☐ Перекодировать переменную доход в 3 категории (высокий, средний, низкий).

☐ Создать новую переменную «лояльность», состоящую из средней оценки по 3

вопросам (удовлетворенность, вероятность покупки, рекомендация).

3. По заданным таблицам с данными в Excel:

☐ Создать сводную таблицу

☐ С помощью сводной таблицы посчитать прибыль и наценку по магазинам,

товарам и группам.

4. Установить проверку значений с выводом подсказок и сообщений об ошибках:

- ☐ Введите целое число от 1 до 10
- ☐ Введите число более 2,5
- ☐ Выберите из списка Фруктов
- ☐ Введите дату 01/01/04 или позднее
- ☐ Введите Время НЕ из интервала 05:00 - 11:00
- ☐ Введите значение из 5 символов

5. По готовым данным (динамика) в Excel:

- ☐ Определить линию тренда, которая лучше всего подходит к данным,
- ☐ Спрогнозировать следующий год
- ☐ Построить график с линией тренда
- ☐ Спрогнозировать следующий год с учетом коэффициентов сезонности

Итоговый контроль по проекту:

В соответствии с учебным планом курса каждый слушатель готовит базу данных с определенным набором показателей для реализации основных видов анализа, указанных в темах курса (данные опросов, официальной статистики, коммерческая информация и т.п.). В дальнейшем данные интерпретируются и оформляются слушателями в виде отчета. Результат работы студента презентуется им и обсуждается на итоговом занятии. Процедура оценивания предусматривает оценку качества и полноты выполнения задания, правильность и полноту ответов на вопросы в ходе обсуждения проекта.

Примеры тем проектов, выполненных по итогам курса (темы из списка не являются обязательными, студент может предложить любую интересную ему область):

1. Прогноз поквартальной динамики ВВП Евросоюза в ближайшие два года
2. Влияние визуальных элементов стратегии продвижения на успешность музыкальных проектов
3. UNECE. Анализ данных по европейским странам
4. Анализ аудитории российского сегмента Интернета
5. Анализ маркетингового бюджета
6. Статистический анализ показателей ТУРИЗМ (Россия, Италия, Франция)
7. Планирование и оценка результатов коммерческих операций
8. Анализ данных по продажам алкогольной продукции магазина
9. Анализ трафика на сайте
10. Анализ взаимосвязи кассовых сборов и рейтинга фильмов

11. Анализ медийной поддержки категории «Соки»
12. Анализ данных по результатам проведения акции
13. Применение экспоненциального сглаживания динамических рядов для прогнозирования ВВП
14. Изучение особенностей восприятия телевизионных рекламных роликов в зависимости от характеристик аудиовизуального ряда
15. Факторы, влияющие на ставку в P2P-кредитовании
16. Изучение влияния факторов на размер доходов населения.
17. Анализ нелинейности и взаимосвязи регрессии
18. Влияние факторов на вероятность угона автомобиля
19. Анализ воздействия промоакции на продажи компании
20. Влияние стоимости бренда компании на инвестиционные качества её акций