

## Программа факультативного курса

«Решение задач эконометрики в среде R»

Кафедра Экономической информатики. Ауд.548, тел. 7548

Сайт [esinf.econ.msu.ru](http://esinf.econ.msu.ru).

**Статус дисциплины:** *Факультатив*, читается в 5 семестре на программе бакалавров по направлению «Экономика».

**Авторы программы и лекторы:**

Пересветов Сергей Борисович, [peresvetov@econ.msu.ru](mailto:peresvetov@econ.msu.ru)

### 1. Место дисциплины в структуре ООП подготовки бакалавра – Профессиональный цикл

Обеспечиваются базовые знания, умения и навыки в области обработки информации для дисциплин «Информационные системы в управлении организацией», «Экономика информации», «Информационные бухгалтерские системы».

Дисциплина обеспечивает базовые умения и навыки в области обработки информации для дисциплин «Микроэкономика-1», «Микроэкономика-2», «Эконометрика», «Статистика», «Теория финансов» и других дисциплин, использующих современные средства обработки данных.

Также обеспечивается овладение умениями и навыками сбора экономической информации из отечественных и зарубежных источников (ПК-9), выбора оптимальных инструментальных средств обработки экономических данных (ПК-5), продвинутый уровень использования для аналитических и исследовательских задач современных технических средств и пакетов прикладных программ (ПК-10), способность участвовать в постановке задач автоматизации своей предметной области (ПК-14).

Входные знания – владение знаниями информационных систем, умениями и навыками работы в прикладных офисных пакетах в объеме дисциплины «Экономическая информатика»

Предшествующие дисциплины – «Экономическая информатика», «Теория вероятностей и математическая статистика»

### 2. Объем дисциплины и виды учебной работы

Нагрузка	В часах	В кредитах
Общая трудоемкость	32	
В том числе: Аудиторная	32	
Самостоятельная		
Контактная		
Форма итогового контроля	Зачёт	

### 3. Цель и задачи дисциплины.

Цель – получение базовых знаний, умений и навыков автоматизированной обработки информации и данных в соответствии с компетенциями ОК-12 и ОК-13, базовые знания основ электронных коммуникаций (ОК-19, ПК-12) и постановки задач автоматизации (ПК-14).

Задачи:

- Получение продвинутых знаний, умений и навыков обработки данных и информации (ПК-10, ПК-5, ПК-12);
- Развитие умений и навыков использования внутренних (корпоративных) и внешних источников данных (ПК-7, ПК-9);
- Выработка знаний, умений и навыков постановки задачи автоматизации своей предметной области (ПК-14).

#### **4. Компетенции обучающегося, формируемые в результате освоения дисциплины (модуля).** В результате освоения дисциплины студент должен:

*Знать:* Основные принципы построения персональной информационной системы, обеспечивающей производительную работу экономиста

*Уметь:* Выделить в решаемой задаче необходимые ИТ-сервисы, подобрать пакет прикладных программ, поддерживающий эти сервисы, построить необходимые структуры данных и интерфейсы между пакетами, получить необходимые исходные данные из статистических библиотек или корпоративной ИС.

*Владеть:* навыками обработки данных в офисных пакетах, статистических пакетах, интеграции данных различных пакетов, поиском данных в электронных библиотеках.

#### **5. Структура и содержание дисциплины (модуля)**

##### **5.1. Содержание разделов дисциплины**

###### **Часть I. Базовые методы статистической обработки данных.**

###### **Тема 1. Язык программирования и среда выполнения R.** Базовые типы данных.

Объекты данных и методы. Работа с пакетами (расширениями) в R. Поиск, получение и загрузка необходимых пакетов расширения. Графика в R. Встроенные графические функции низкого и высокого уровня. Дополнительные графические пакеты.

###### **Тема 2. Базовые диаграммы.** Столбчатые диаграммы. Простые столбчатые диаграммы.

Столбчатые диаграммы: составные и с группировкой. Столбчатые диаграммы для средних значений. Оптимизация столбчатых диаграмм. Круговые диаграммы.

Гистограммы. Диаграммы ядерной оценки функции плотности. Диаграммы размахов.

Использование диаграмм размахов для сравнения групп между собой. Точечные диаграммы.

###### **Тема 3. Основные методы статистической обработки данных.** Описательные

статистики. Вычисление описательных статистик для групп данных. Визуализация результатов. Таблицы частот и таблицы сопряженности. Создание таблиц частот. Тесты на независимость. Показатели взаимосвязи. Визуализация результатов. Типы корреляций.

Проверка статистической значимости корреляций. Визуализация корреляций. Тесты Стьюдента. Тест Стьюдента для независимых выборок. Тест Стьюдента для зависимых выборок. Когда имеется больше двух групп. Непараметрические тесты межгрупповых различий. Сравнение двух групп. Сравнение более двух групп. Визуализация групповых различий.

###### **Часть II. Расширенные методы статистической обработки данных.**

###### **Тема 4. Регрессия.** МНК-регрессия. Подгонка регрессионных моделей при помощи

команды `lm()`. Простая линейная регрессия. Полиномиальная регрессия. Множественная линейная регрессия. Множественная линейная регрессия со взаимодействиями.

Диагностика регрессионных моделей. Общая проверка выполнения требований, предъявляемых к линейным моделям. Мультиколлинеарность. Необычные наблюдения.

Выбросы. Точки высокой напряженности. Влиятельные наблюдения. Способы корректировки. Удаление наблюдений. Преобразование переменных. Добавление или удаление переменных. Выбор регрессионной модели. Сравнение моделей. Выбор переменных. Продолжение анализа. Кросс-валидация. Введение в методы машинного обучения.

###### **Тема 4. Продвинутое графическое моделирование.** Четыре графические системы R. Пакет

`lattice`. Условные переменные. Функции для изменения формата ячеек. Группировка переменных. Графические параметры. Расположение диаграмм на странице. Пакет `ggplot2`. Интерактивная графика. Взаимодействие с диаграммами: идентификация

точек. Создание интерактивной графики. Трехмерная графика.

## **6. Образовательные технологии**

Лекции, в т.ч. с презентациями PowerPoint, решение задач в интерактивном режиме в ходе семинара, самостоятельное решение задач вне аудитории, коллоквиумы, электронные тесты. Лекции составляют 33% аудиторного времени, самостоятельная внеаудиторная работа равна аудиторному времени.

## **7. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины и учебно-методическое обеспечение самостоятельной работы студентов**

Самостоятельные работы, контрольные работы, зачет. Эссе в рамках дисциплины не предполагаются.

## **8. Учебно-методическое и информационное обеспечение дисциплины (модуля)**

### **8.1 основная литература**

1. Роберт Кабаков, R в действии. Анализ и визуализация данных на языке R. - М.: ДМК Пресс, 2016 г.
2. W.N.Venables, D.M.Smith and the R Development Core Team, An Introduction to R, 2010
3. А.Б. Шипунов, Е.М. Балдин, Анализ данных с R, часть I, 2009
4. А.Б. Шипунов, А. И. Коробейников, Е.М. Балдин, Анализ данных с R, часть II, 2009
5. С.В.Петров, Е.М. Балдин, Анализ данных с R, часть III, 2010
6. Н.Ю.Золотых, А.Н.Половинкин, Введение в систему R, М.: 2007
7. Oleg Nenadic, Walter Zucchini, Statistical Analysis with R - a quick start, 2004
8. Grant V. Farnsworth, Econometrics in R, 2008
9. Paul S.P.Cowpertwait, Andrew V.Metcalfе, Introductory Time Series with R, Springer, 2009
10. Eric Zivot and Jiahui Wang, Modelling Financial Time Series with S-PLUS, Second Edition, 2005

### **8.2 дополнительная литература**

1. Меретиллов М.А., Методические указания к лабораторным работам по курсу «Методы анализа данных», 2006
2. А.А.Барсегян, М.С.Куприянов, И.И.Холод, М.Д.Тесс, С.И. Елизаров, Анализ данных и процессов, СПб: БХВ-Петербург, 2009
3. Паклин Н.Б., Орешков В.И. Бизнес-аналитика: от данных к знаниям: Учебное пособие, 2-е изд., СПб.:Питер, 2010

### **8.3 Интернет-источники:**

1. [cran.r-project.org](http://cran.r-project.org)

## **9. Балльная система оценки знаний**

1. Самостоятельные работы – 50 баллов
2. Контрольные работы – 50 баллов
3. Зачет – 50 баллов

(условие получения итоговой оценки в соответствии с положением о балльно-рейтинговой системой оценки знаний)

## **10. Материально-техническое обеспечение дисциплины (модуля)**

Для лекций – проектор.

Для семинаров – компьютерный класс, ПО Microsoft Office Enterprise 2016, пакет R, учебная база данных в среде MS SQL, доступ к институциональной подписке ЭФ МГУ.