

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

для обучающихся по организации и
проведению самостоятельной работы по дисциплине
«Научный семинар «Современные исследования информационной бизнес-аналитики»»
для студентов направления подготовки 38.04.05 «Бизнес-информатика»
(направленность (профиль) «Интеллектуальный анализ данных и цифровые бизнес-
технологии»)

Ставрополь, 2026

Методические рекомендации по дисциплине «Научный семинар «Современные исследования информационной бизнес-аналитики»» содержат задания для студентов, необходимые для организации самостоятельной работы студентов.

Предназначены для студентов направления подготовки 38.04.05 «Бизнес-информатика» (направленность (профиль) «Интеллектуальный анализ данных и цифровые бизнес-технологии»).

Общая характеристика самостоятельной работы студента при изучении дисциплины «Научный семинар «Современные исследования информационной бизнес-аналитики»»

Целью самостоятельной работы студентов является овладение фундаментальными знаниями, профессиональными умениями и навыками деятельности по профилю, опытом творческой, исследовательской деятельности. Самостоятельная работа студентов способствует развитию самостоятельности, ответственности и организованности, творческого подхода к решению проблем учебного и профессионального уровня.

Задачами СР являются:

- систематизация и закрепление полученных теоретических знаний и практических умений студентов;
 - углубление и расширение теоретических знаний;
 - формирование умений использовать нормативную, правовую, справочную документацию и специальную литературу;
 - развитие познавательных способностей и активности студентов: творческой инициативы, самостоятельности, ответственности и организованности;
 - формирование самостоятельности мышления, способностей к саморазвитию, самосовершенствованию и самореализации;
 - развитие исследовательских умений;
- использование материала, собранного и полученного в ходе самостоятельных занятий на семинарах, на практических и лабораторных занятиях, при написании курсовых и выпускной квалификационной работ, для эффективной подготовки к итоговым зачетам и экзаменам.

Методические рекомендации по подготовке к лабораторным занятиям

Работа по подготовке к лабораторным занятиям и активное в них участие — одна из форм изучения программного материала курса. Она направлена на подготовку высококвалифицированных магистров. Подготовку к занятиям следует начинать с внимательного изучения соответствующих разделов учебных пособий и учебников, далее следует изучать специальную литературу и источники, работать с картами, таблицами, схемами, наконец, написать доклад, если студент получил такое задание. Готовясь к занятиям и принимая активное участие в их работе, студент проходит школу работы над источниками и литературой, получает навыки самостоятельной работы над письменным и устным сообщением (докладом), учится участвовать в дискуссиях, отстаивать свою точку зрения, формулировать и аргументировать выводы. Форма лабораторных занятий во многом определяется его темой. Практика показывает, что основные формы занятий следующие: беседа на основе составленного преподавателем плана (она наиболее приемлема при обсуждении одного из теоретических вопросов по проблемам темы или монографии), коллоквиум по разделу учебника или одной из монографий (коллоквиум предполагает прежде всего проверку знаний по определенной теме, источникам, разделу курса); подготовка письменного доклада студентом, его устный доклад и обсуждение его на занятии.

В планы лабораторных занятий включены основные вопросы общего курса. В ходе занятий возможна их конкретизация и корректировка. При подготовке сообщений и докладов следует широко использовать опубликованные источники, мемуарную и исследовательскую литературу. Учебники и учебные пособия студент использует по своему выбору. Каждому студенту в течение года следует прочитать не менее двух

монографий (или одну монографию и одну работу мемуарного характера), которые указаны в списке литературы или рекомендовано преподавателем из числа новых публикаций, составить краткий доклад и быть готовым к беседе по ним с преподавателем.

Метод конкретных ситуаций (case method, case-study), кейс-метод, метод кейсов, метод ситуационного анализа используется как в образовательной, так и исследовательской деятельности.

Кейс - это описание реальных ситуаций. Суть метода заключается в том, что слушателям предлагается готовая ситуация, которая в той или иной степени имитирует реальную, жизненную. Чаще всего она излагается письменно в виде готовой «истории», причем финал остается «открытым». В качестве учебной задачи участникам предлагают ее проанализировать и предложить свое решение.

Иногда кейс подразумевает использование приём инсценировки или принятие на себя роли. В других учебных ситуациях более подходящим является использование так называемого «инцидента». Под «инцидентом» понимается наиболее простая конфликтная ситуация, возникающая в ходе реальной деятельности и требующая от руководителя оперативного решения.

Оценка работы над кейсами происходит по следующим критериям:

- степень проработки проблемы (обоснованность альтернативных решений, наличие рисков и ограничений);
- оригинальность, креативность подхода участников группы при разработке альтернатив;
- применимость решения на практике

Темы для докладов

1. Моделирование и информатизация принятия решений
2. Основные понятия системного анализа и теории принятия решений
3. Принятие решений в условиях определенности
4. Принятие решений в условиях полной неопределенности
5. Принятие решений в условиях частичной неопределенности
6. Матричные игры
7. Кооперативные и коалиционные игры
8. Использование комбинированных методов принятия решений в среде ЭСППР
9. Имитационное и визуальное компьютерное моделирование в принятии управленческих решений

Методические указания по подготовке к экзамену

В период подготовки к экзамену студенты вновь обращаются к пройденному учебному материалу. При этом они не только скрепляют полученные знания, но и получают новые. Подготовка студента к экзамену включает в себя три этапа:

- самостоятельная работа в течение семестра;
- непосредственная подготовка в дни, предшествующие экзамену по темам курса;
- подготовка к ответу на вопросы, содержащиеся в билетах.

Литература для подготовки к экзамену рекомендуется преподавателем либо указана в учебно-методическом комплексе. Для полноты учебной информации и ее сравнения лучше использовать не менее двух учебников. Студент вправе сам

придерживаться любой из представленных в учебниках точек зрения по спорной проблеме (в том числе отличной от преподавателя), но при условии достаточной научной аргументации.

Основным источником подготовки к экзамену является конспект лекций, где учебный материал дается в систематизированном виде, основные положения его детализируются, подкрепляются современными фактами и информацией, которые в силу новизны не вошли в опубликованные печатные источники. В ходе подготовки к экзамену студентам необходимо обращать внимание не только на уровень запоминания, но и на степень понимания излагаемых проблем.

Экзамен проводится по билетам, охватывающим весь пройденный материал. По окончании ответа экзаменатор может задать студенту дополнительные и уточняющие вопросы. На подготовку к ответу по вопросам билета студенту дается 30 минут с момента получения им билета. Положительным также будет стремление студента изложить различные точки зрения на рассматриваемую проблему, выразить свое отношение к ней, применить теоретические знания по современным проблемам жилищного права.

Результаты экзамена объявляются студенту после окончания ответа в день сдачи.

Вопросы к зачету:

1. Понятие проекта.
2. Основные характеристики проекта.
3. Жизненный цикл проекта.
4. Классификация проектов.
5. Понятие терминального (конечного) проекта.
6. Понятие развивающегося проекта.
7. Понятие открытого проекта.
8. Организационные структуры управления проектом.
9. Команда в рамках функциональных структур.
10. Матричные организационные структуры.
11. Виды матричных организационных структур.
12. Проектно-целевые организационные структуры.
13. Выбор организационной структуры управления проектом.
14. Взаимосвязь команды и участников проекта.
15. Управление командой проекта.
16. График Ганта.
17. Элемент сетевого графика
18. Проектный треугольник.
19. Структура декомпозиции работ (WBS).
20. Методы построения структуры декомпозиции работ.
21. Матрица ответственности.
22. Взаимосвязь структуры декомпозиции работ (WBS), структуры ресурсов проекта (RBS) и организационной структуры проекта (OBS).
23. Информационные системы управления проектами.
24. Классификация информационных систем управления проектами (прединвестиционный анализ проекта).
25. Классификация информационных систем управления проектами (планирование проекта).

26. Классификация информационных систем управления проектами (выполнение проекта).
27. Классификация информационных систем управления проектами (завершение проекта).
28. Сравнительная характеристика информационных систем управления проектами.
29. Понятие портфеля проектов.
30. Сравнительная характеристика проектов, программ и портфелей проектов.
31. Классификация проектов в зависимости от стадии жизненного цикла компании.
32. Управление портфелем проектов.
33. Количественные показатели для отбора проектов.
34. Оценочная матрица проектов.
35. Сравнение проектов по параметрам риска, стоимости и ценности для бизнеса.
36. Понятие и классификация рисков проекта.
37. Управление рисками проекта (экспертный анализ рисков, метод аналогов).
38. Управление рисками проекта (анализ показателей предельного уровня, анализ чувствительности проекта).
39. Управление рисками проекта (анализ сценариев развития проекта, метод построения дерева решений проекта).
40. Управление рисками проекта (метод имитационного моделирования).
41. Методы теории игр в процессе принятия решений в условиях риска и неопределенности (методы, основанные на критерии абсолютного оптимизма и пессимизма).
42. Методы теории игр в процессе принятия решений в условиях риска и неопределенности (методы, основанные на критерии Севиджа, на критерии Гурвица, на критерии Байеса-Лапласа).

Список рекомендуемой литературы

Основная литература:

1. Основы бизнес-анализа. Учебное пособие / Под ред. Бариленко В.И. - М.: КноРус, 2019. - 56 с.
2. Кошик Авинаш. Веб-аналитика 2.0 на практике. Перевод Ю.И. Корниенко. – М.: Диалектика, 2019. - 528 с.
3. Бернард Марр. Ключевые инструменты бизнес-аналитики. Перевод Егоров В.Н. М.: Лаборатория знаний, 2018. – 339 с.

Дополнительная литература:

1. Еремеева, Н.В. Планирование и анализ бизнес-процессов на основе построения моделей управления конкурентоспособности продукции / Н.В. Еремеева. - М.: Русайнс, 2018. - 16 с.
2. Карлберг, К. Бизнес-анализ с использованием Excel / К. Карлберг. - М.: Диалектика, 2019. - 576 с.

Интернет-ресурсы:

1. www.biblioclub.ru - Электронно-библиотечная система «Университетская библиотека онлайн»
2. www.iprbookshop.ru - Электронно-библиотечная система «IPRbooks»

3.МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«СЕВЕРО-КАВКАЗСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

МЕТОДИЧЕСКИЕ УКАЗАНИЯ
по выполнению лабораторных работ
по дисциплине «Научный семинар «Современные исследования информационной
бизнес-аналитики»»
для студентов направления подготовки 38.04.05 «Бизнес-информатика»
(направленность (профиль) Интеллектуальный анализ данных и цифровые бизнес-
технологии)

Ставрополь, 2026

Методические указания по дисциплине «Научный семинар «Современные исследования информационной бизнес-аналитики»» содержат задания для студентов, необходимые для проведения лабораторных занятий.

Предназначены для магистров направления подготовки 38.04.05 «Бизнес-информатика» (направленность (профиль) Интеллектуальный анализ данных и цифровые бизнес-технологии).

Введение

Целью данной дисциплины является формирование набора компетенций будущего магистра по направлению подготовки 38.04.05 «Бизнес- информатика».

Задачами дисциплины «Научный семинар «Современные исследования информационной бизнес-аналитики»» являются:

- профессиональное освоение нормативных моделей и математических методов выбора лучшего варианта действий, технологий поддержки принятия решений;
- удовлетворение потребностей заказчиков в кадрах, которые умеют применять научно обоснованные методы принятия решений и компьютерные системы для их поддержки, владеют эффективными методами обработки, формализации и структурирования выбираемых решений, понимают мировые тенденции в области экспертной поддержки принятия решений в процессе разработки на предприятиях корпоративных автоматизированных систем;
- совершенствование профессиональной компоненты образования по направлению информатика и вычислительная техника по профилю программное обеспечение средств вычислительной техники и автоматизированных систем путем применения методов теории принятия решений в конкретной предметной области.

Дисциплина изучается в первом семестре реализации магистерской программы и входит в часть, формируемую участниками образовательных отношений учебного плана.

Лабораторные работы № 1

Тема: Задачи принятия решений в условиях неопределенности

Теоретическая часть

Business analysis body of knowledge (BABOK)

В текущей практике, если опираться на BABOK, бизнес-анализ и бизнес-анализ в IT выступают как синонимичные понятия.

Задачи бизнес-аналитики:

- выявление потребностей;
- обоснование изменений;
- разработка решений.

В третьей версии BABOK акцент сделан на изменениях. Улучшения в организации реализуются за счет контролируемых изменений, которые удовлетворяют потребностям заинтересованных лиц и позволяют достигать поставленных целей. Центральные идеи стандарта BABOK – шесть ключевых понятий, которые необходимо учитывать при бизнес-анализе в организации (рисунок 1).

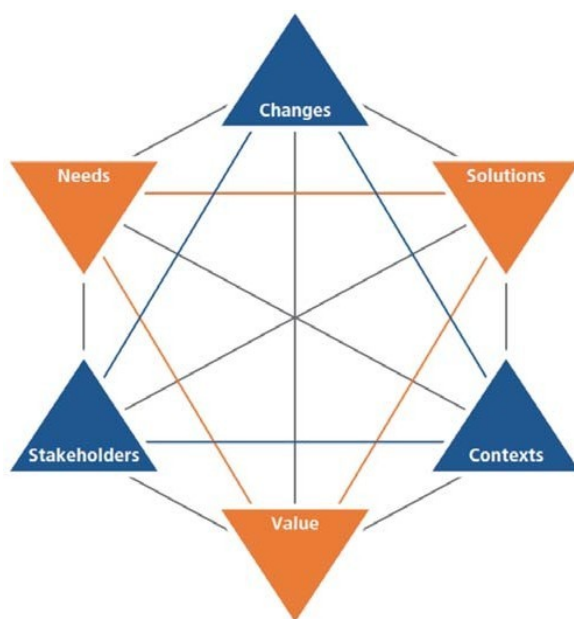


Рисунок 1 – Ключевые понятия BABOK

Во главу модели поставлены «*changes*» – «изменения», т.е. преобразования деятельности организации для повышения производительности (эффективности).

Изменения вызываются «*needs*» – «потребностями» – это проблемы и возможности, которые мотивируют действовать заинтересованных лиц, изменяют ценность продукта (услуги), разрушая или повышая ее.

Потребности требуют «*solution*» – «решения», выбора способа действовать, удовлетворяющего потребностям заинтересованных сторон и учитывающего контекст.

Изменения инициируются «*stakeholder*» – «заинтересованной стороной» на основе имеющихся потребностей, влияющих на выбор решения и воздействующих на изменения.

Критерием отбора решений и проведения изменений выступает

«value» – «ценность», потенциальная или реализованная стоимость, значимость или полезность для удовлетворения потребностей заинтересованных сторон в конкретной ситуации (контексте).

Выбор решений и изменения учитывают «context» – «контекст», обстоятельства окружающей среды, которые влияют на понимание решений и изменений.

BAВОК выделяет следующие стадии анализа:

- понять текущее состояние КАК ЕСТЬ;
- определить будущее состояние КАК ДОЛЖНО БЫТЬ;
- определить действия по переходу от «как есть» к «как должно быть».

Бизнес-анализ напрямую связан с бизнес-архитектурой компании, бизнес-процессами и IT-решениями для их поддержки. Одной из важных составляющих бизнес-анализа является анализ данных. Управление чем-либо возможно на основе измеримых показателей, которые как раз и позволяют на основе исторических данных понять – как есть, определить как должно быть и выработать программу планового изменения показателей на основе выявленных зависимостей между ними для достижения этих значений.

Благодаря текущему мониторингу данных можно отслеживать эффективность принятых решений и корректировать программу изменений, своевременно реагируя на динамично меняющийся контекст.

В BAВОК определяются пятьдесят методов бизнес-анализа, но в бизнес-аналитике на основе данных акцент будет сделан на четырех из них:

- Data Mining («Интеллектуальный анализ данных»);
- Data Modelling («Моделирование данных»);
- Data Dictionary («Словарь данных»);
- Data Flow Diagrams («Диаграммы потоков данных»).

Наборы данных для анализа. Открытые данные

Для выработки решений необходимо знать контекст. В отслеживании текущего контекста эффективны BI-системы (англ. *business intelligence*), которые обрабатывают большие массивы данных, регистрируемых в цифровой форме.

На сегодняшний день существует тенденция бесплатного и свободного раскрытия цифровых данных как со стороны государства, так и со стороны бизнеса.

Организуются конкурсы или хакатоны по анализу данных (*hackathon* – первоначально соревнование разработчиков-программистов, на текущий момент любых разработчиков, связанных с цифровыми решениями):

- Kaggle – площадка, на которой различные компании проводят спортивные соревнования по анализу данных (построение прогнозных моделей на близких к реальным данными) <https://www.kaggle.com/>;
- «Лаборатория Касперского» проводит хакатоны по анализу данных, посвященные индустриальной безопасности;

The Media в 2018 г., цель которого найти новые идеи и решения по обработке огромного массива медиаданных.

И это только малая часть возможностей попробовать свои силы в анализе реальных данных, получить практический опыт использования современных

инструментов и платформ анализа.

В реальном бизнесе данные составляют значительную часть конкурентных преимуществ, а значит, являются коммерческой тайной. Такие данные не будут публичными и результаты обработки таких данных останутся собственностью компании с ограниченным доступом.

Для учебных целей в бизнес-аналитике используются публичные данные. В них включаются открытые данные и разделяемые данные.

Существует масса способов объявления статуса данных и открытой информации: при распространении открытого программного кода часто используются лицензии GPL (General Public License), LGPL, лицензия FreeBSD, лицензия Apache. Для данных разработаны специальные лицензии, одна из этих лицензий Open Database Commons (ODBC) разработана Open Knowledge Foundation и лицензия, которая сейчас становится все более общепринятой, это Creative Commons Zero, это последняя версия лицензии Creative Commons, которая фактически позволяет использовать данные для любых целей, в том числе коммерческих.

Последний вариант помогает компании также в налаживании социальных связей, завоевании репутации и формировании канала подбора сотрудников, которые до начала официального найма доказали свою полезность компании.

Данные, которые публикуют коммерческие организации, удовлетворяют обычно не всем условиям открытых данных. Такие данные называют «разделяемые».

Публикатор разделяемых данных не обязан делиться всеми подробностями получения данных. Однако он позволяет воспользоваться накопленным объемом цифровых данных третьей стороне, не неся никакой ответственности за то, что эти данные могут оказаться не действительными, некорректными. Публикатор не обязан поддерживать данные в актуальном состоянии.

Существует несколько площадок, на которых известные компании публикуют кейсы вместе с данными для их решения. Такой площадкой является, например, <https://www.kaggle.com>.

Сайт «Открытые данные в России» предоставляет доступ к аннотированным наборам данных, многие из которых удовлетворяют условиям открытых данных (<http://data.gov.ru/opendata>).

Сайт Евростата (Eurostat) предоставляет доступ к очень обширной статистике по странам Европы и регионам (<http://ec.europa.eu/eurostat/data/database>).

Сайт Международного банка (World Bank) предоставляет доступ к данным по различным странам (<http://databank.worldbank.org/data/home.aspx>).

Сайт Организации экономического сотрудничества и развития (Organisation for Economic Cooperation and Development – OECD) публикует данные об экономических и демографических показателях, входящих в нее развитых стран (stats.oecd.org).

Сайт «Европейское социальное исследование» (European Social Survey ESS) предоставляет доступ к данным сравнительного исследования изменения установок, взглядов, ценностей и поведения населения Европы, для анализа можно использовать данные восьми волн 2002, 2004, 2006, 2008, 2010, 2012, 2014 и 2016 гг. (<http://www.europeansocialsurvey.org/data/>).

Контрольные вопросы и задания

1. Ключевые понятия ВАВОК

2. Наборы данных для анализа.
3. Открытые данные

Лабораторная работа №2. Язык программирования R. Среда разработки RStudio

Для анализа данных будем использовать язык R, который наряду с Python активно используется для решения задач анализа данных.

Для установки необходимо загрузить последнюю версию R с сайта <https://cran.r-project.org/> (имеются установщики для всех ОС).

Версия R для Windows поставляется с графической оболочкой RGui (рисунок 1.2).

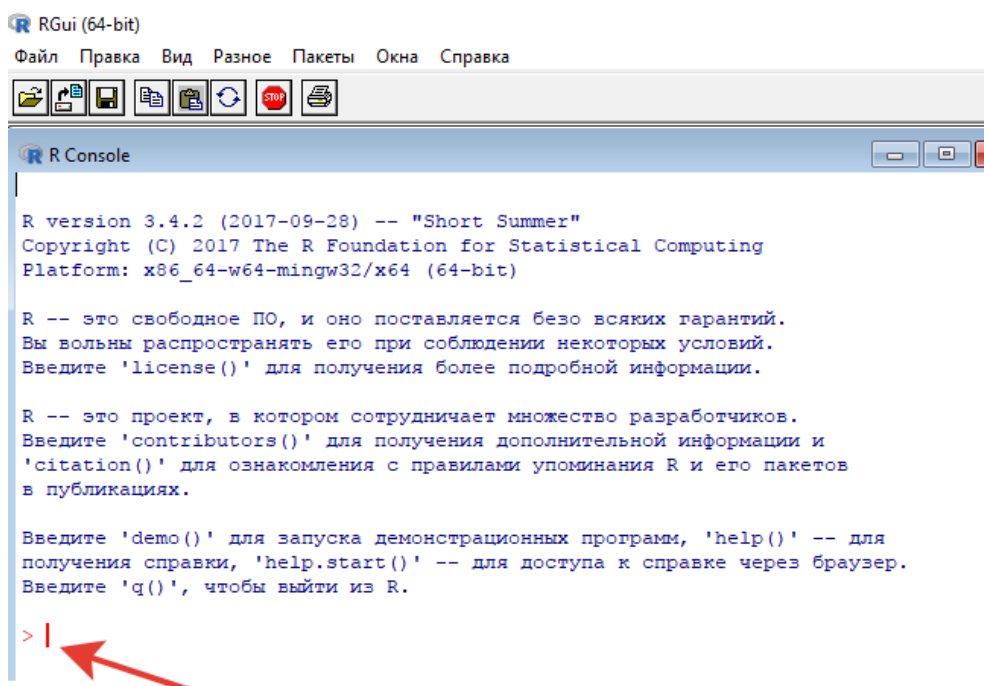


Рисунок 1.2 – Графический интерфейс RGui (окно консоли с приглашением «>» для набора кода)

Графический интерфейс поддерживает три вида окон:

- консоль, в которой после ввода кода (клавиша Enter) он сразу передается на исполнение;
- редактор скриптов (открывается с помощью меню Файл > Новый скрипт), в котором код набирается, редактируется и при необходимости сохраняется с расширением .R, на исполнение код можно посылать построчно или блоками, предварительно выделив;
- графические окна, в которых строятся и сохраняются графические объекты.

В базовой установке R имеется некоторое количество предустановленных пакетов. Просмотреть все установленные пакеты можно с помощью функции

`library()`

дополнительные пакеты находятся по адресу: <https://cran.r-project.org/>.

Замечание: код, который можно набирать в консоли или скрипте и запускать на исполнение будет набираться шрифтом Calibri на сером фоне.

Параллельно с R рекомендуется установить среду разработки (IDE) RStudio, дистрибутив которой можно загрузить по адресу: <https://www.rstudio.com/>.

Как и язык R, эта среда относится к open source.

Начальные сведения для работы в RStudio. Начальное окно среды разработки представлено на рисунке 1.3:

- Source editor and data view (редактор кода и обзор данных);
- R console (консоль);
- служебная панель 1: Environment, history and connections окружение (данные, функции, история команд и доступные внешние ресурсы);
- служебная панель 2: Files, plots, packages, help and viewer (файлы, графики, пакеты, помощь и просмотрщик).

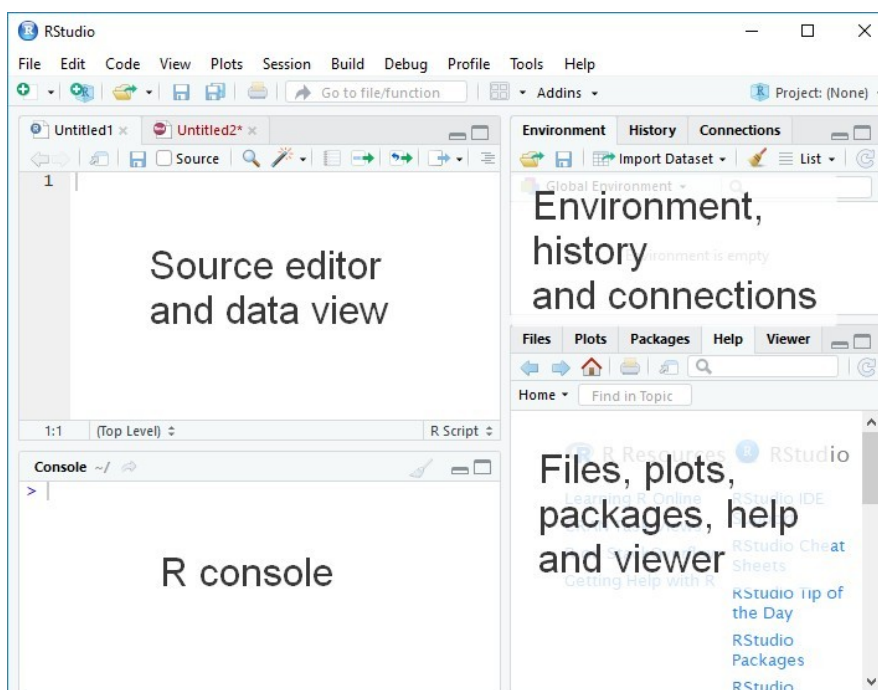


Рисунок 1.3 – IDE RStudio

В меню Tools > Global options... можно настроить расположение панелей на вкладке Pane Layout: панели Source, Console только поменять местами, а вот вкладки на служебных панелях можно перемещать с одной панели на другую по желанию (рисунок 1.4).

Код предпочтительнее набирать в панели Source, открыв новый файл с помощью меню File > New file > R script. На выполнение код посылается комбинацией клавиш CTRL+ENTER или кнопкой в правом верхнем углу панели  Run .

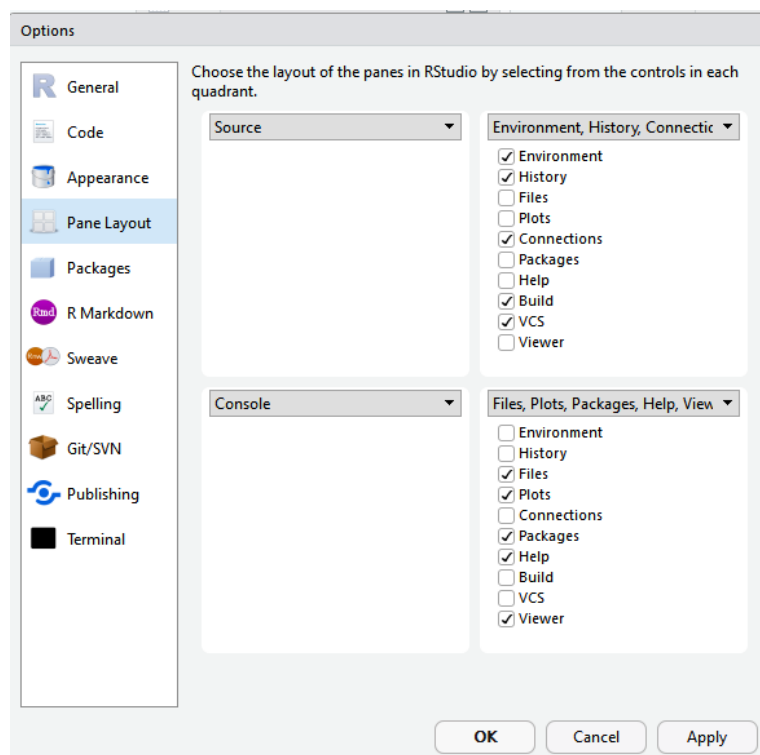


Рисунок 1.4 – Настройка панелей

RStudio поддерживает автоматическое завершение кода в окне Source при помощи клавиши TAB (рисунок 1.5). Такие подсказки можно получить и при наборе аргументов, если используются ранее определенные переменные.

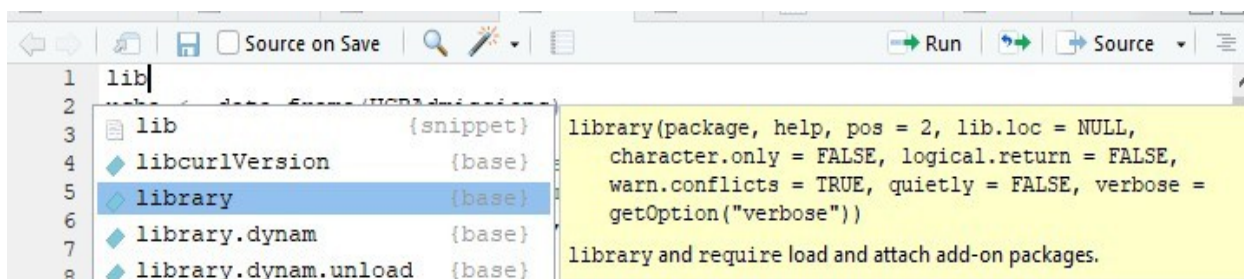


Рисунок 1.5 – Завершение функции library() с помощью клавиши TAB

Горячие клавиши

CTRL + 1 – перемещает курсор в окно Source; CTRL + 2 – перемещает курсор в Console; CTRL + L – очищает окно Console от текста; ESC – прерывает вычисления.

Открыть существующий файл можно с помощью File > Open или кнопки  . Открыть файл, с которым недавно работали, можно с помощью File > Recent Files или кнопки раскрыть список на панели инструментов (рисунок 1.6).

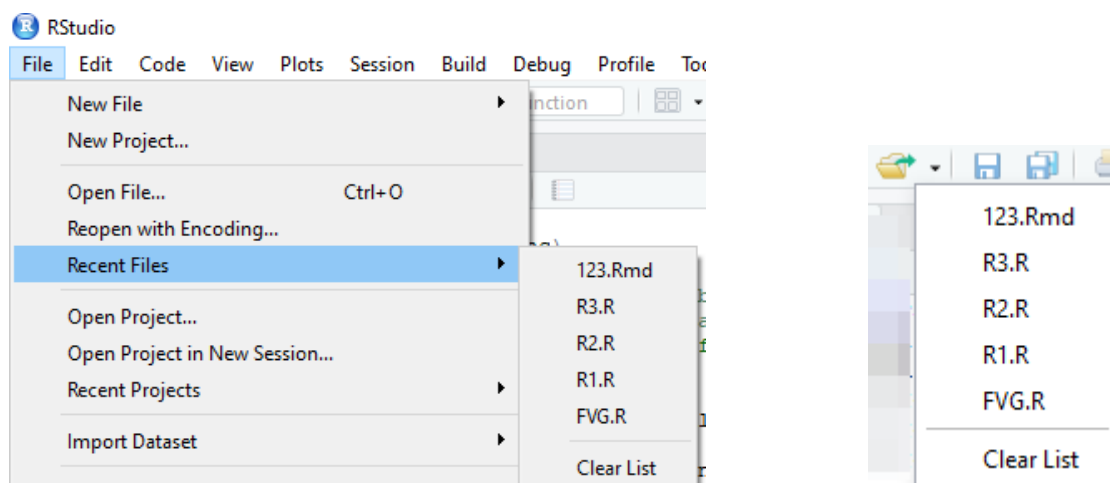


Рисунок 1.6 – Список недавно использовавшихся файлов

В окне Source есть возможность найти и заменить необходимые части текста: с помощью Edit > Find and Replace или кнопки . Объект, внесенный в поле для поиска, будет подсвечен в коде. В соседнем окне Replace можно ввести замену (рисунок 1.7).

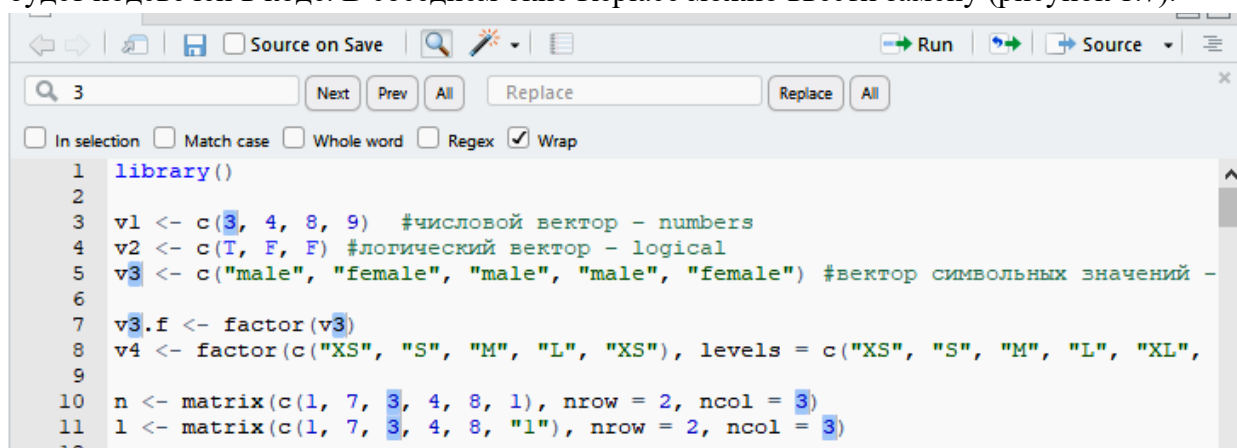


Рисунок 1.7 – Поиск и замена кода

При закрытии RStudio программа предлагает сохранить код в открытых файлах текущей сессии и рабочее пространство (Workspace), которое включает в себя созданные переменные, историю команд (рисунок 1.8).

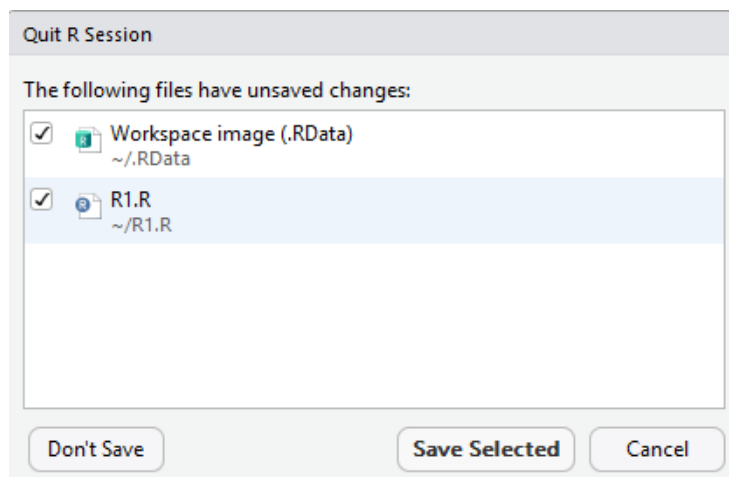


Рисунок 1.8 – Завершение работы RStudio

При открытии RStudio загружается сохраненное рабочее пространство со всеми незакрытыми в прошлой сохраненной сессии файлами, созданными переменными, функциями и историей команд.

Вкладка History (рисунок 1.9), которую можно подключить в одну из служебных панелей, содержит журнал всех ранее выполненных в рабочем пространстве команд (в окне Console).

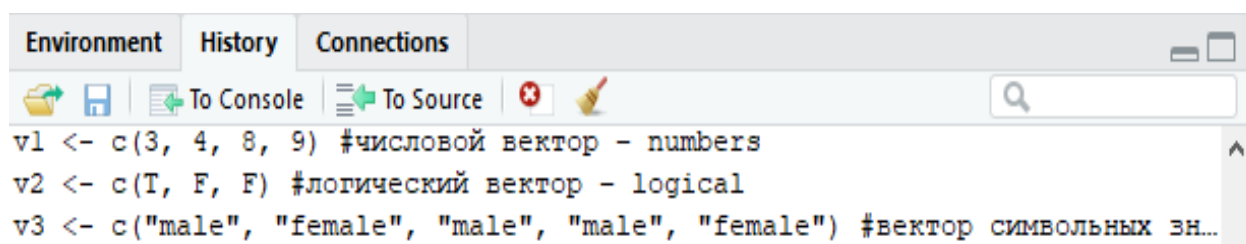
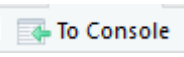
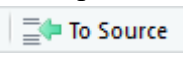
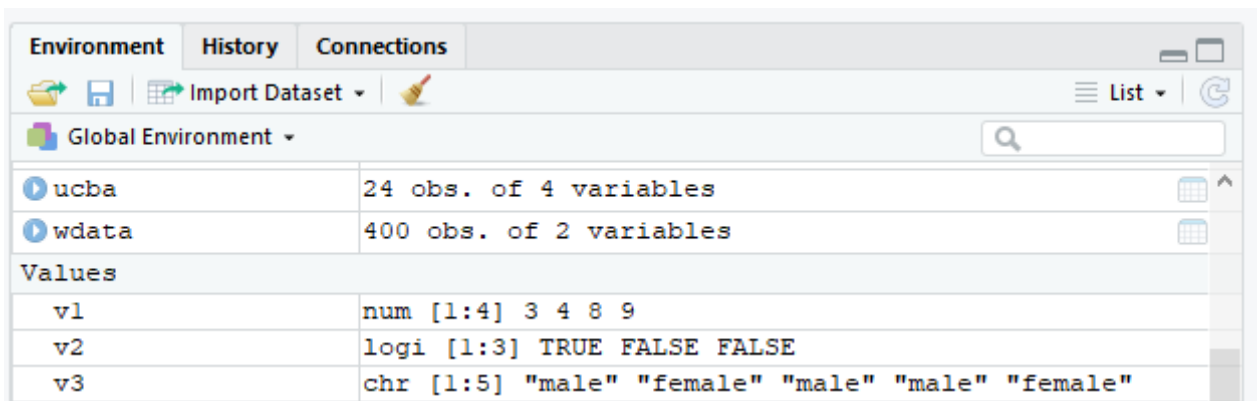


Рисунок 1.9 – История команд

Команды сохраняются и после очистки окна Console. Эти команды можно передавать на выполнение в Console кнопкой , они выполняются после нажатия ENTER, или скопировать в открытый файл в окне Source (если открытого файла нет, то создается новый) кнопкой . Для работы с несколькими командами их нужно предварительно выделить.

Вкладка Environment (рисунок 1.10), которую можно подключить в одну из служебных панелей, содержит описание всех объектов (это могут быть пользовательские функции, переменные), которые были созданы ранее в рабочем пространстве.

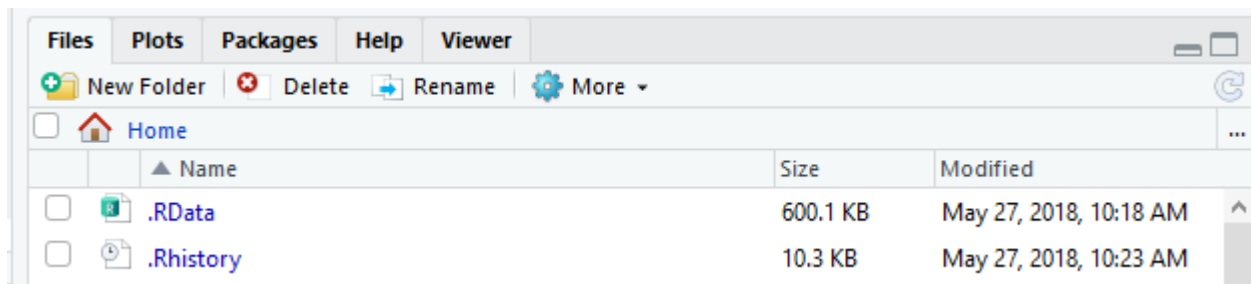


Global Environment	
ucba	24 obs. of 4 variables
wdata	400 obs. of 2 variables
Values	
v1	num [1:4] 3 4 8 9
v2	logi [1:3] TRUE FALSE FALSE
v3	chr [1:5] "male" "female" "male" "male" "female"

Рисунок 1.10 – Завершение функции library() с помощью клавиши TAB

Текущую рабочую директорию (папку) можно проверить с помощью функции `getwd()`

Файлы текущей рабочей директории отображаются в окне вкладки Files (рисунок 1.11)



Name	Size	Modified
.RData	600.1 KB	May 27, 2018, 10:18 AM
.Rhistory	10.3 KB	May 27, 2018, 10:23 AM

Рисунок 1.11 – Файлы текущей рабочей директории

Текущую рабочую папку можно поменять, например, с помощью кнопки на вкладке Files (рисунок 1.12) или с помощью меню Tools > Change Working Dir...

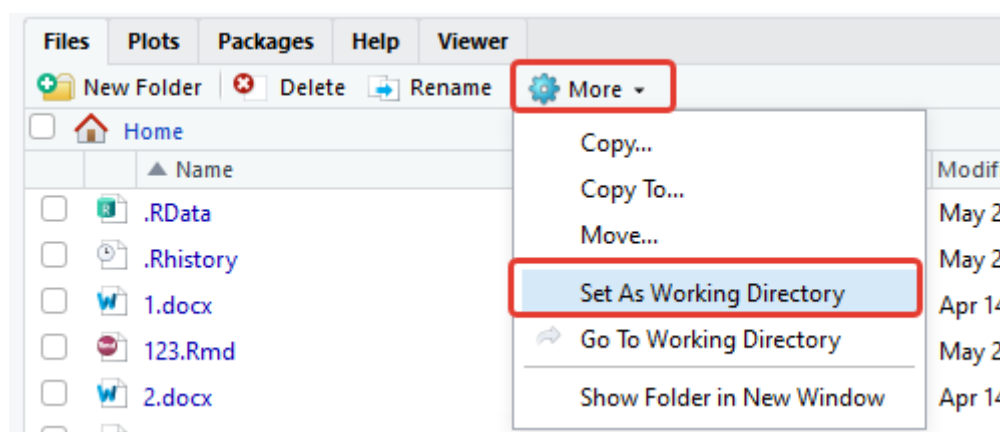


Рисунок 1.12 – Смена рабочей папки

Возможные последствия изменения рабочей папки:

- имеющиеся в коде относительные ссылки на файлы станут недействительными;
- при завершении работы файл .RData будет сохранен в новую папку.

Установка новых пакетов, необходимых в данной рабочей сессии, возможна на вкладке Packages с помощью кнопки Install (рисунок 1.13). При наборе имени пакета подключается автозаполнение.

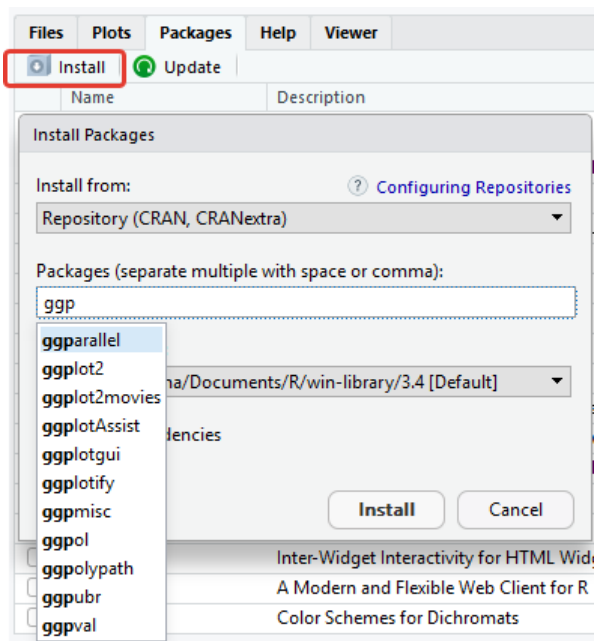


Рисунок 1.13 – Установка пакетов

Флаг на опции Install dependencies означает, что установятся и все связанные пакеты, функции которых использует нужный пакет (рисунок 1.14).

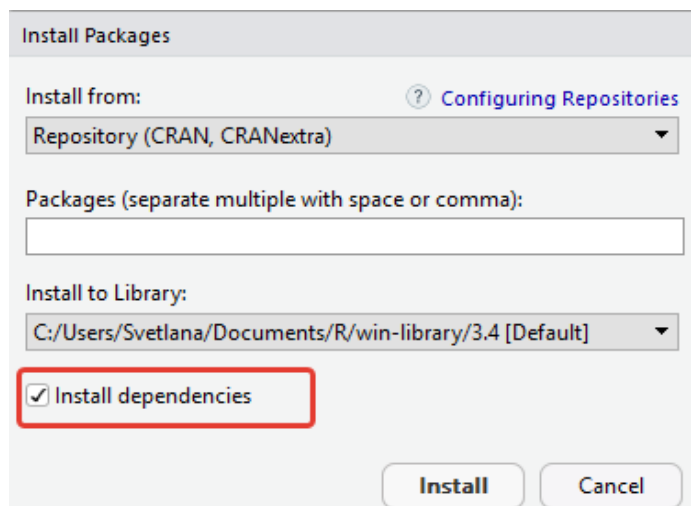


Рисунок 1.14 – Установка всех связанных пакетов

Установка новых пакетов возможна и с помощью команды, аргументы которой – это имя пакета и значение true для переменной `dependencies`, чтобы установились связанные пакеты:

```
install.packages('ggplot', dependencies = T)
```

Установленные пакеты нужно подключить в текущей рабочей сессии, чтобы можно было использовать определенные в них функции и имеющиеся в пакетах наборы данных. Это можно сделать на вкладке **Packages**, установив флаг на нужном пакете (рисунок 1.15).

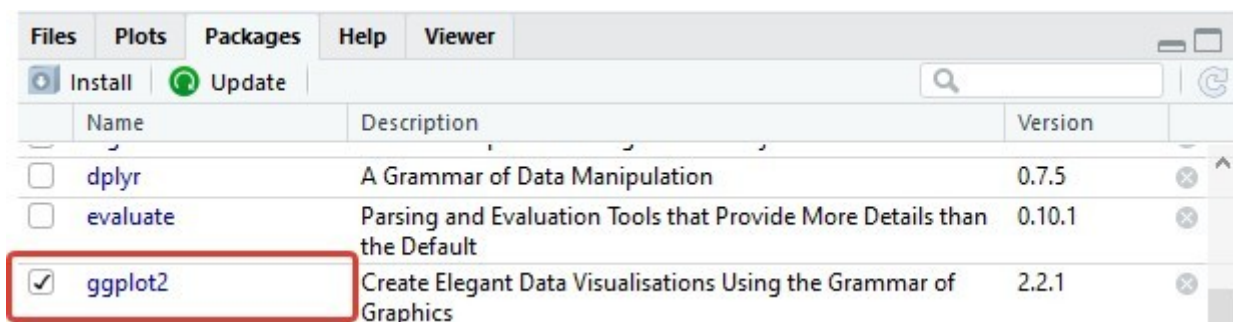


Рисунок 1.15 – Подключение пакета в текущей сессии

Или с помощью команды

```
library(ggplot2)
```

Но подключение пакета с помощью команды не гарантирует, что такой пакет уже был установлен (и подключение не выполнится), тогда как в первом варианте недоступные для подключения пакеты просто не отображаются.

На вкладке **Help** можно получить справку по пакетам, функциям, наборам данных и т.п. (рисунок 1.16).

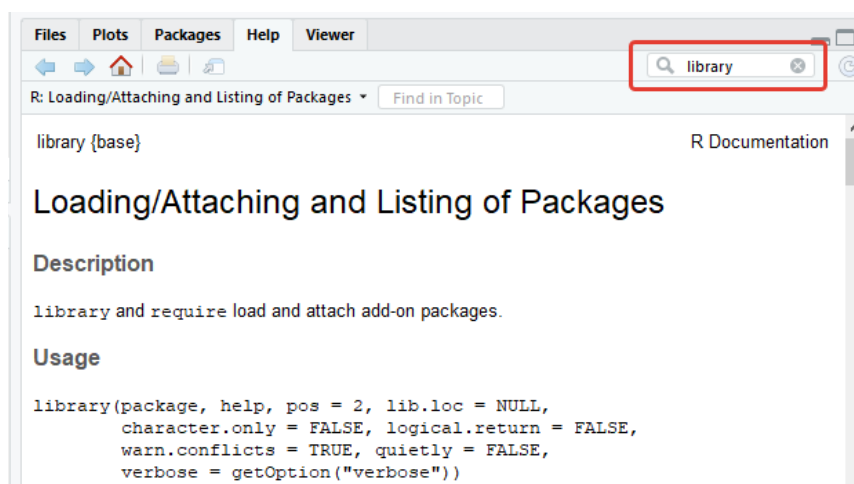


Рисунок 1.16 – Справка по функции `library()`

Или с помощью знака вопрос перед объектом для получения о нем справки:

?library()

На вкладке Plots отображаются построенные в текущей сессии графики (рисунок 1.17).

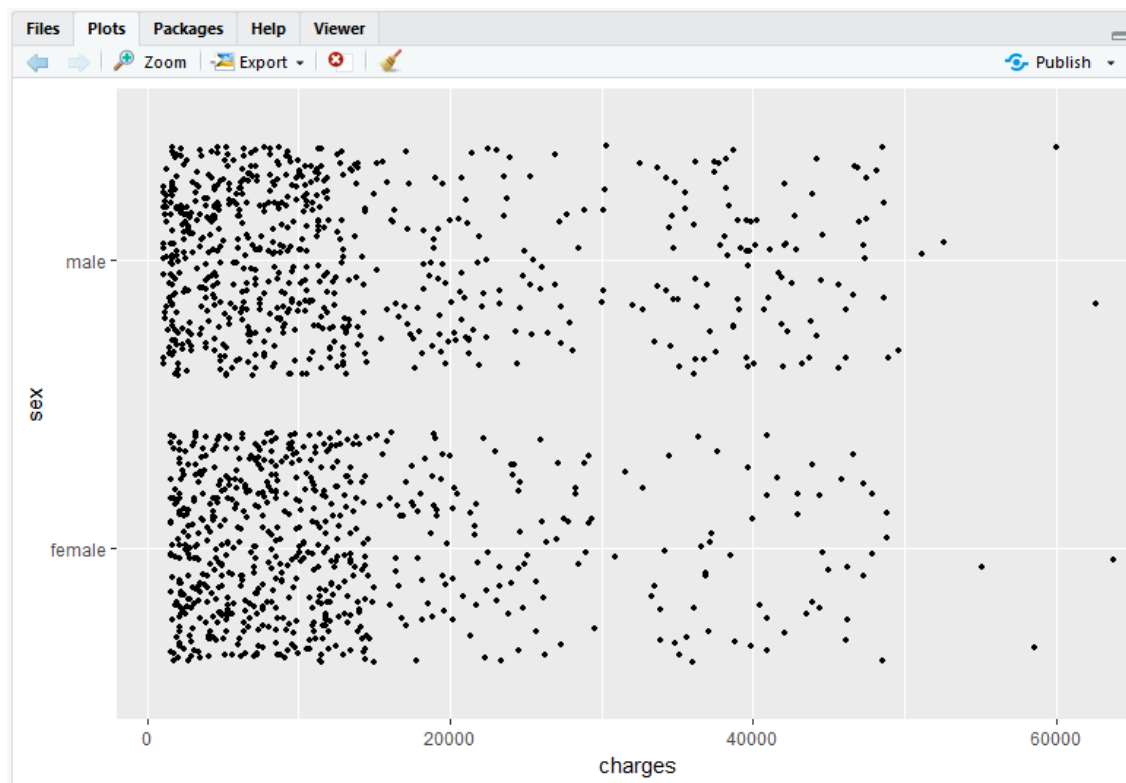


Рисунок 1.17 – График

На вкладке Plots (рисунок 1.18): 1 – между графиками можно перемещаться, 2 – увеличивать размер графика, 3 – сохранять график в доступном формате (как рисунок или файл .pdf).

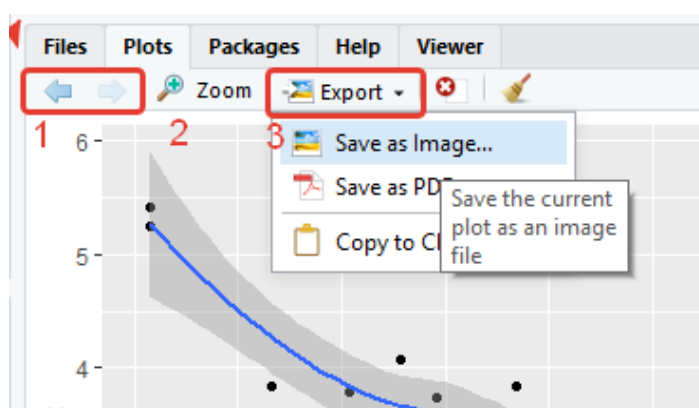


Рисунок 1.18 – Операции с графиками

Типы данных в R

Основные типы данных, которые поддерживает язык R:

- логические данные (*logical*);
- вещественные числа (*numeric*);
- комплексные числа (*complex*);
- текстовые данные или символьные (*character*).

Объекты данных: векторы, факторы, массивы, матрицы, таблицы данных и списки.

Присваивание переменной значений с помощью `<-`, несколько значений передаются с помощью команды `c()`, в которой отдельные значения разделяются запятой, а текстовые значения заключаются в кавычки. Комментируется код с помощью `#`.

Базовый объект данных в R – вектор. Создание вектора:

```
v1 <- c(T, F, F) #логический вектор – logical
v2 <- c(3, 4, 8, 9) #числовой вектор – numeric
v3 <- c("male", "female", "male", "male", "female") #вектор символьных
значений – character
```

После выполнения кода на вкладке Environment отобразятся все три переменные с указанием типа данных элементов вектора (рисунок 1.19).

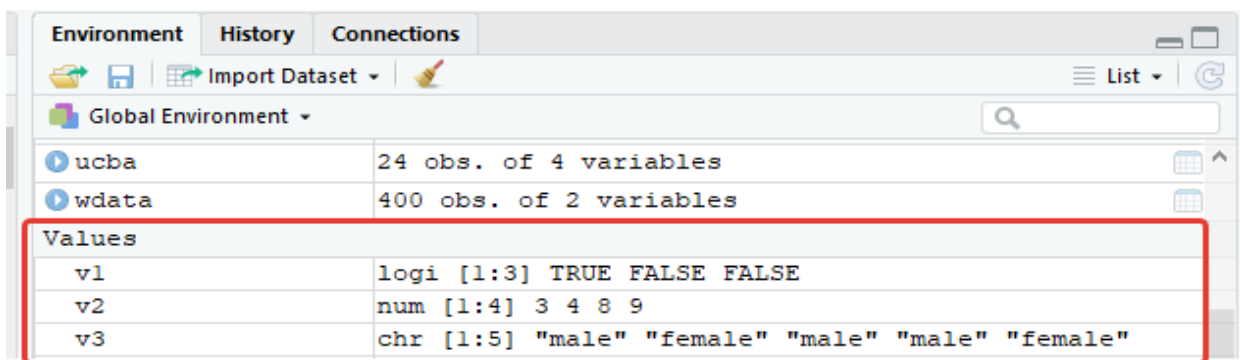


Рисунок 1.19 – Переменные (в служебной вкладке RStudio)

Вектор может содержать данные только одного типа (рисунок 1.20).

Самый общий тип данных – символьный.



logical



numbers



character

Рисунок 1.20 – Векторы

Если среди логических значений будет хотя бы одно число, то вектор определится как числовой, а если одно из значений символьное, то и весь вектор определится как вектор символьных значений (рисунок 1.21):

```
v1 <- c(T, F, F, 8) #числовой вектор – numeric
v2 <- c(T, F, F, "female") #вектор символьных значений – character
```

Рисунок 1.21 – Переменные (в служебной вкладке RStudio)

Values	
v1	num [1:4] 1 0 0 8
v2	chr [1:4] "TRUE" "FALSE" "FALSE" "female"

Частный вид векторной переменной – фактор. Фактор это группирующая переменная. Так переменная v3 может быть преобразована в фактор:

```
v3.f <- factor(v3)
```

Можно создать факторную переменную v4, значения которой будут упорядочены:

Отличие этих факторов в том, что в первом случае числовые метки присваиваются при лексикографическом упорядочении (как в словарях), а во втором – согласно следованию

```
v4 <- factor(c("S", "XS", "XXL", "L", "XS"), levels = c("XS", "S", "M", "L", "XL", "XXL", "XXXL"))
```

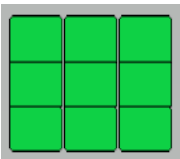
значений в переменной levels (рисунок 1.22).

v3.f	Factor w/ 2 levels "female","male": 2 1 2 2 1
v4	Factor w/ 7 levels "XS","S","M","L",...: 2 1 6 4 1

Рисунок 1.22 – Факторы (в служебной вкладке RStudio)

Матрица – это частный случай массива (двумерный массив), как, впрочем, и вектор – это одномерный массив. На матрицу также накладывается ограничение на тип значений: у всех значений в матрице один тип данных (рисунок 1.23).

Рисунок 1.23 – Матрица



Создание матриц:

```
n <- matrix(c(1, 7, 3, 4, 8, 1), nrow = 2, ncol = 3)
l <- matrix(c(1, 7, 3, 4, 8, "1"), nrow = 2, ncol = 3)
```

Результат во вкладке Environment (рисунок 1.24).

l	chr [1:2, 1:3] "1" "7" "3" "4" "8" "1"
n	num [1:2, 1:3] 1 7 3 4 8 1

Рисунок 1.24 – Матрицы (в служебной вкладке RStudio)

Массив в практике анализа данных используется достаточно редко, у всех значений в массиве – один тип данных (рисунок 1.25).

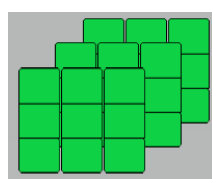


Рисунок 1.25 – Массив

Создание массива, где аргумент dim задает размерность (три матрицы размерности 2*2), а имя переменной вызывает ее содержимое в консоль:

Результат выполнения кода в Console (рисунок 1.26).

```
ar<-array(data=1:12, dim=c(2,2,3))
> ar
, , 1
      [,1] [,2]
[1,]     1     3
[2,]     2     4

, , 2
      [,1] [,2]
[1,]     5     7
[2,]     6     8

, , 3
      [,1] [,2]
[1,]     9    11
[2,]    10    12
```

Рисунок 1.26 – Вывод значений созданного массива ar в консоль

Таблица данных или двумерный набор данных встречается в анализе данных особенно часто. Колонки в таблице могут иметь различный тип данных, но в пределах одной колонки (столбца) – все данные одного типа (рисунок 1.27).

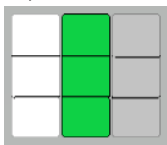


Рисунок 1.27 – Таблица данных

Создание таблицы данных:

```
sale<-data.frame(IDcustomer=c(1101, 1102, 1105), gender=
c("male", "female", "male"), Totalsum=c(18000, 1920, 1078))
```

Результат выполнения кода во вкладке Environment (рисунок 1.28).

Environment	History	Connections
Global Environment		
sale		3 obs. of 3 variables
IDcustomer: num 1101 1102 1105		
gender : Factor w/ 2 levels "female","male": 2 1 2		
Totalsum : num 18000 1920 1078		

Рисунок 1.28 – Таблица данных sale

Также таблицу данных можно отобразить в окне Source, щелкнув по имени таблицы данных во вкладке Environment (рисунок 1.29).

	IDcustomer	gender	Totalsum
1	1101	male	18000
2	1102	female	1920
3	1105	male	1078

Рисунок 1.29 – Таблица данных sale

Импорт данных в R

Язык R имеет множество возможностей для импорта данных из самых разных источников. Кроме того, устанавливаемые пакеты так- же содержат наборы данных.

Рассмотрим встроенные в пакеты наборы данных. В базовой установке есть пакет datasets, в котором большое количество именованных наборов данных. Информацию о наборах данных можно получить по команде:

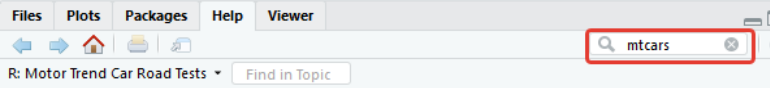
```
library(help = "datasets")
```

В результате выполнения кода в окне Source появится список наборов данных (рисунок 1.30).

Информация по пакету 'datasets'	
Определение:	
Package:	datasets
Version:	3.4.2
Priority:	base
Title:	The R Datasets Package
Author:	R Core Team and contributors worldwide
Maintainer:	R Core Team <R-core@r-project.org>
Description:	Base R datasets.
License:	Part of R 3.4.2
Built:	R 3.4.2; ; 2017-09-28 09:44:47 UTC; windows
Индекс:	
AirPassengers	Monthly Airline Passenger Numbers 1949-1960
BJSales	Sales Data with Leading Indicator
BOD	Biochemical Oxygen Demand
CO2	Carbon Dioxide Uptake in Grass Plants
ChickWeight	Weight versus age of chicks on different diets
DNase	Elisa assay of DNase
EuStockMarkets	Daily Closing Prices of Major European Stock Indices, 1991-1998
Formaldehyde	Determination of Formaldehyde
HairEyeColor	Hair and Eye Color of Statistics Students
Harman23.cor	Harman Example 2.3
Harman74.cor	Harman Example 7.4

Рисунок 1.30 – Список наборов данных в пакете datasets

Более подробно о каждом наборе можно узнать во вкладке Help (рисунок 1.31).



mtcars {datasets} R Documentation

Motor Trend Car Road Tests

Description

The data was extracted from the 1974 *Motor Trend* US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

Usage

```
mtcars
```

Format

A data frame with 32 observations on 11 variables.

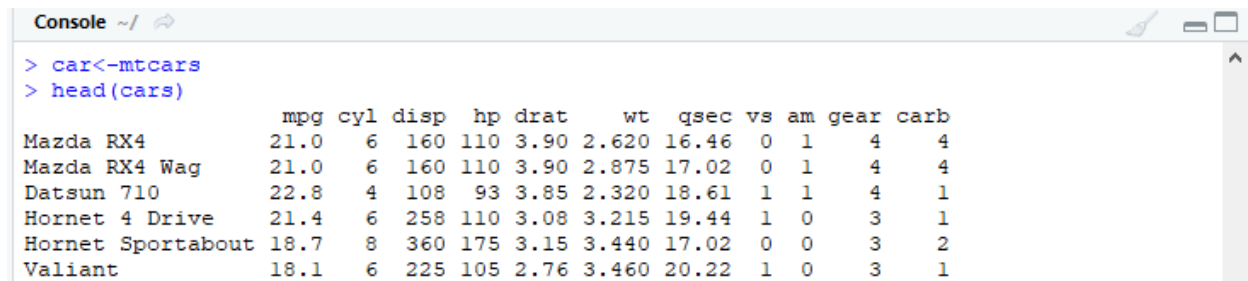
- [1] mpg Miles/(US) gallon
- [2] cyl Number of cylinders
- [3] disp Displacement (cu.in.)
- [4] hp Gross horsepower
- [5] drat Rear axle ratio
- [6] wt Weight (1000 lbs)

Рисунок 1.31 – Информация о наборе mtcars

Передать набор в новую переменную и посмотреть начало таблицы:

```
car<-mtcars  
head(cars)
```

Результат выполнения кода в консоли (рисунок 1.32).



	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

Рисунок 1.32 – Первые шесть строк таблицы данных car

Наборы данных, размещенные на сайтах в сети Интернет, можно загружать несколькими способами. Можно предварительно скачать файл данных, разместить в рабочей папке и использовать в функции `read.table()` в качестве аргумента `file` непосредственно имя файла. Можно в качестве аргумента `file` использовать гиперссылку на файл в Интернете. Не все варианты размещения данных позволяют обращаться к ним по ссылке, поэтому используем первый способ.

На сайте Kaggle (<https://www.kaggle.com/>) необходимо пройти регистрацию, после чего станут доступны наборы данных для скачивания (рисунок 1.33).

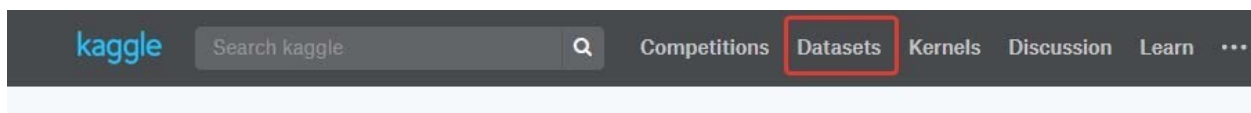


Рисунок 1.33 – Платформа kaggle

Доступны публичные наборы данных (рисунок 1.34), которые отфильтрованы по объему (<10 Mb), так как без специальных ухищрений R не позволяет напрямую из консоли анализировать большие объемы данных.

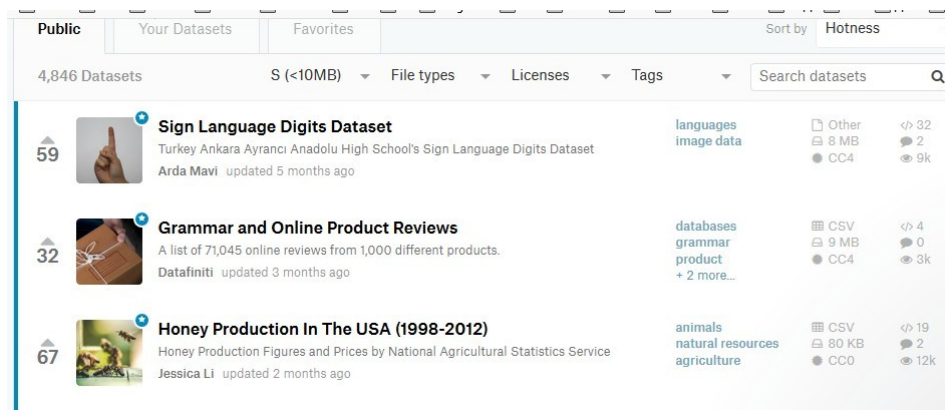


Рисунок 1.34 – Наборы данных

Среди наборов выберем набор <https://www.kaggle.com/ulabox/ulabox-orders-with-categories-partials-2017/data> (рисунок 1.35). И скачаем его.



Рисунок 1.35 – Набор данных по заказам в супермаркете Ulabox (Испания) в 2017 г.

Далее перенесем его в рабочую папку и сократим название, чтобы было легче работать с импортом (рисунок 1.36).

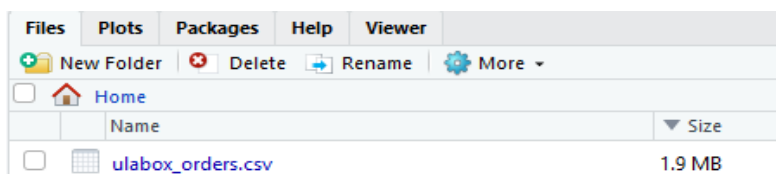


Рисунок 1.36 – Рабочая директория

Импортируем данные в переменную Orders с помощью функции `read.table()`, аргументы которой имя файла в кавычках (файл в рабочей директории), `header=T` (так как в первой строке заголовки) и `sep` – разделитель данных (в нашем случае – запятая):

```
Orders<-read.table("ulabox_orders.csv", header=T, sep=",")
```

Результат выполнения кода в двух окнах Source и Environment на рисунке 1.37.

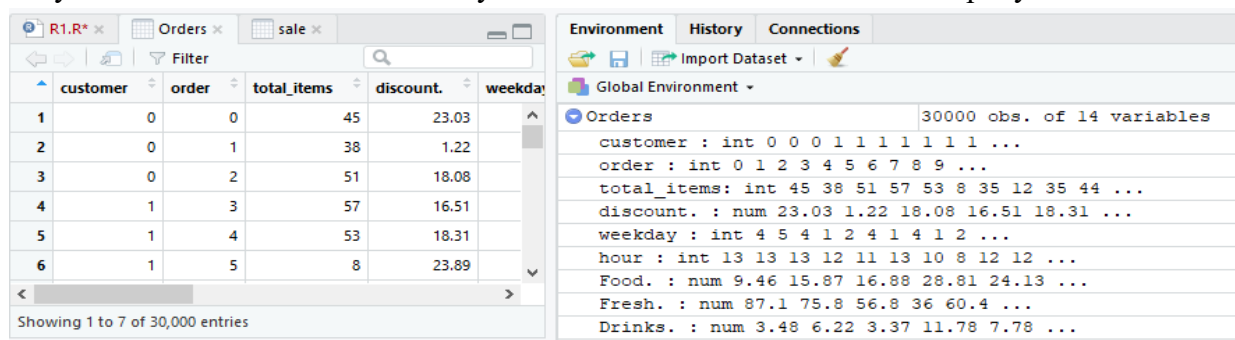


Рисунок 1.40 – Набор данных Orders

Краткое описание импортированного набора данных

Набор данных онлайн-супермаркета Ulabox 2017, отступ из <https://www.github.com/ulabox/datasets>. Ulabox – это супермаркет в Испании.

Переменные набора:

- customer: анонимизированный идентификатор клиента;
- order: ID заказа, начиная с нуля;
- total_items: количество приобретенных товаров в заказе;
- discount%: процент от общей полученной скидки;
- weekday: день недели: 1=понедельник, 7=воскресенье;
- hour: час дня, когда была совершена покупка от 00 до 23.

Categories' partials процент денег, потраченных в каждой из восьми основных категорий веб-сайта:

- food%: не скоропортящиеся продукты, например, рис, снеки, со усы и т.п.;
- fresh% : свежие и замороженные продукты, например: фрукты, овощи, замороженное мясо и т.п.;
- drinks% : все виды напитков, например: вода, соки, вино, алкогольные напитки, молоко, соевые напитки;
- home% : товары для дома, от туалетной бумаги до мелкой бытовой техники;
- beauty% : парфюмерия, косметика, уходовые средства;
- health% : лекарственные средства, которые можно продавать в Испании без рецепта врача: диетические таблетки, презервативы, зуб- ная паста;
- baby% : товары по уходу за ребенком, детское питание и т.п.;
- pets% : товары для домашних животных.

БАЗОВЫЕ ВОЗМОЖНОСТИ ЯЗЫКА R ДЛЯ ВИЗУАЛИЗАЦИИ ДАННЫХ

Функция plot()

plot() – основная графическая команда, она распознает тип объекта, который подлежит рисованию, и строит соответствующий график.

Основные аргументы функции plot()

- main – название диаграммы;
- pch (plotting character [0:25]) – вид маркера;
- cex (character extension) – размер;
- col (colour) – цвет;
- xlab – название горизонтальной оси;
- ylab – название вертикальной оси.

Для построения графиков скачаем и импортируем набор [https:// www.kaggle.com/mirichoi0218/insurance](https://www.kaggle.com/mirichoi0218/insurance) – расходы на медицинское обслуживание.

Описание переменных набора:

- age: возраст основного бенефициара;
- sex: пол застрахованного;
- bmi: индекс массы тела;
- children: число детей, охваченных медицинским страхованием / число иждивенцев;
- smoker: курит ли застрахованный;
- region: жилой район получателя в США, Северо-Восток, Юго-Восток, Юго-Запад, Северо-Запад;
- charges: индивидуальные медицинские расходы, оплачиваемые страховкой.

Рассмотрим набор данных – расходы на медицинское обслуживание тех, кто имеет медицинскую страховку. Первые строки набора на рисунке 2.1.

	age	sex	bmi	children	smoker	region	charges
1	19	female	27.900	0	yes	southwest	16884.924
2	18	male	33.770	1	no	southeast	1725.552
3	28	male	33.000	3	no	southeast	4449.462
4	33	male	22.705	0	no	northwest	21984.471
5	32	male	28.880	0	no	northwest	3866.855
6	31	female	25.740	0	no	southeast	3756.622
7	46	female	33.440	1	no	southeast	8240.590

Рисунок 2.1 – Набор данных insurance

Оценим визуально, каковы пропорции женщин и мужчин в наборе. Обращение к переменной происходит по имени таблицы до знака \$ и названию столбца после: insurance\$sex.

Результат выполнения команды на рисунке 2.2, аргумент main задает заголовок.

```
plot(insurance$sex, main = "Половой состав застрахованных", col = c("red","blue"))
```

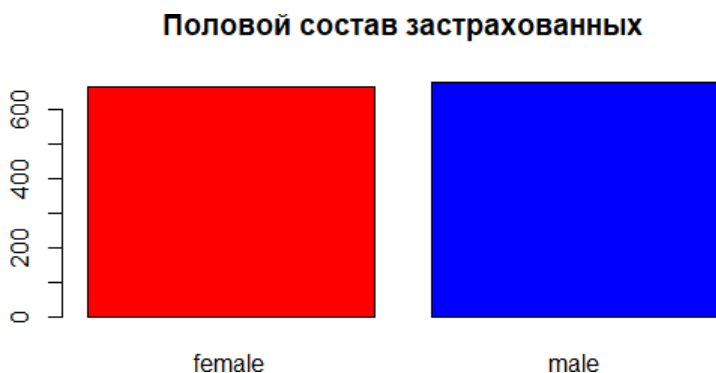


Рисунок 2.2 – Разделение набора данных по полу

Аналогично проведем оценку пропорций курящих и некурящих в наборе данных: Результат выполнения команды на рисунке 2.3.

```
plot(insurance$smoker, main = "Количество курящих среди  
застрахованных", col = c("green", "red"))
```



Рисунок 2.3 – Разделение набора данных по наличию вредной привычки

Проанализируем зависимость индивидуальных медицинских расходов в зависимости от пола, возраста и наличия вредных привычек.

Числовые переменные: возраст и расходы – построим на графике. Возраст – независимая переменная, она откладывается по оси x и записывается первой:

Результат выполнения команды на рисунке 2.4, аргумент `main` задает заголовок, а `xlab`,

```
plot(insurance$age, insurance$charges, main = "Медицинские расходы", xlab =  
"Возраст застрахованного", ylab = "Расходы")
```

`ylab` – подписи осей.

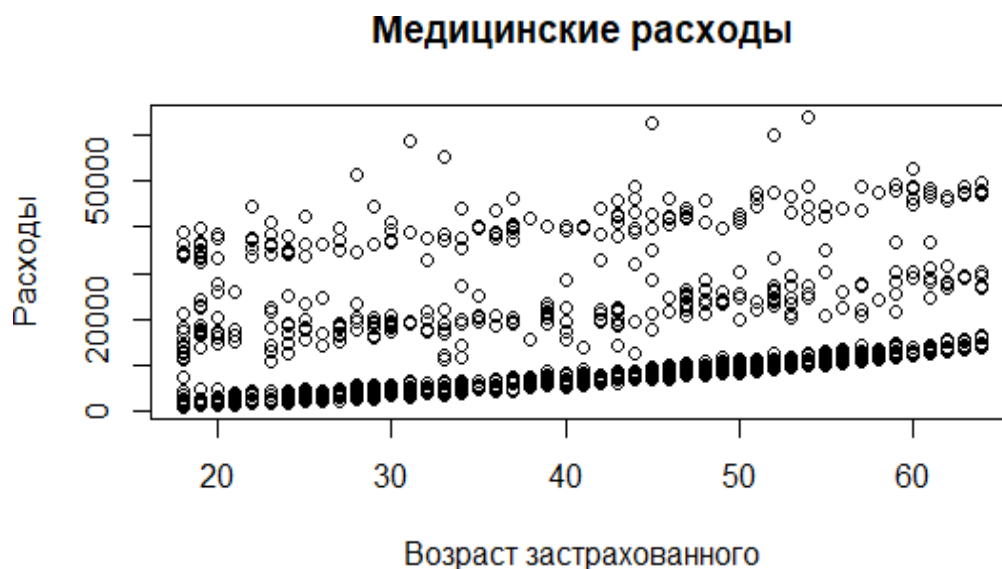


Рисунок 2.4 – Индивидуальные медицинские расходы в зависимости от возраста

Очевидно, что с возрастом индивидуальные медицинские расходы растут. Но интересно узнать, влияет ли пол или скажем наличие вредной привычки на медицинские расходы.

Изменим цвет точек графика в зависимости от пола с помощью функции `ifelse(insurance$sex=="female", "red", "blue")`, если значение пола – женщина, то цвет точек будет красным, иначе – синим:

```
plot(insurance$age, insurance$charges, col = ifelse(insurance$sex == "female", "red", "blue"), main = "Медицинские расходы", xlab = "Возраст застрахованного", ylab = "Расходы")
```

И добавим легенду: первый аргумент задает ее расположение “top(bottom)right(left)”, второй размер (лучше брать меньше 1, чтобы не перекрывался график), третий вид маркера для точек, четвертый цвет точек и последний текст подписей точек

```
legend("topleft", sex = .5, pch = 1, col = c("red", "blue"), c('женщины', 'мужчины'))
```

в легенде:

Результат выполнения команд на рисунке 2.5.

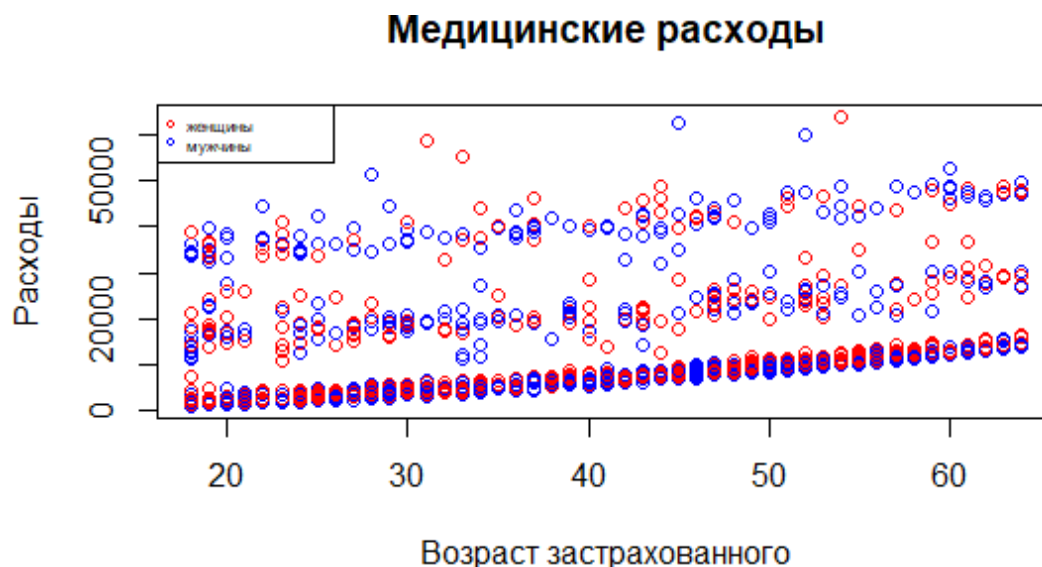


Рисунок 2.5 – Индивидуальные расходы в зависимости от возраста с учетом пола

Визуально разница между индивидуальными расходами мужчин и женщин не очень бросается в глаза. Попробуем добавить данные о курении, изменив вид маркера для курящих: `pch = as.numeric(insurance$smoker)`. Необходимо добавить преобразование к числовому виду переменной `insurance$smoker` с помощью функции `as.numeric()`:

И в легенду теперь добавляется вид маркера и дополнительные значения подписей точек легенды:

```
plot(insurance$age, insurance$charges, col = ifelse(insurance$sex == "female",  
"red", "blue"), pch = as.numeric(insurance$smoker), main = "Медицинские  
'некурящие женщины', 'некурящие мужчины',  
'курящие женщины', 'курящие мужчины')", ylab = "Расходы")  
pch = c(1, 1, 2, 2), col = c("red", "blue"))
```

Теперь наш график имеет совсем подробный вид (рисунок 2.6).

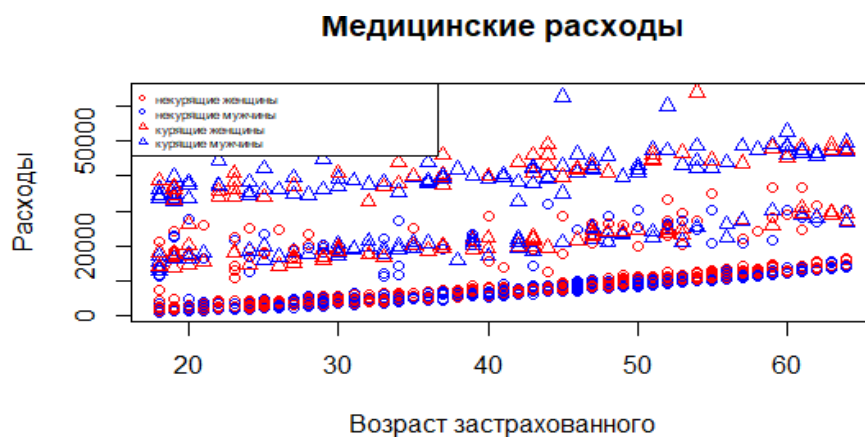


Рисунок 2.6 – Индивидуальные расходы в зависимости от возраста с учетом пола и вредной привычки

И хорошо заметно, что большее число треугольников сосредоточено в зоне высоких медицинских расходов, а большее число кружочков в зоне небольших индивидуальных медицинских расходов.

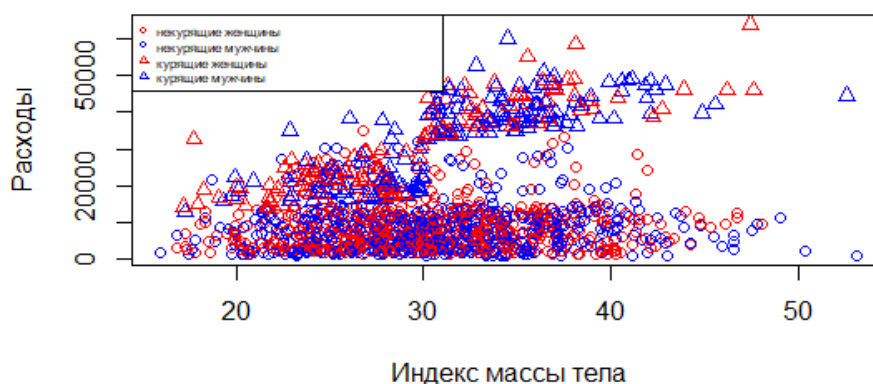
Построим аналогичным образом график зависимости медицинских расходов от индекса массы тела с учетом пола и наличия вредной привычки:

```
plot(insurance$bmi, insurance$charges, col = ifelse(insurance$sex == "female",
"red", "blue"), pch=as.numeric ( insurance$smoker), main = "Медицинские
расходы", xlab = "Индекс массы тела", ylab = "Расходы")

legend("topleft",sex=.5, c('некурящие женщины',
'некурящие мужчины', 'курящие женщины', 'курящие мужчины'), pch = c(1, 1, 2,
2), col = c("red", "blue"))
```

Теперь четко прослеживается, что повышенный индекс массы тела при наличии вредной привычки в совокупности существенно повышают медицинские расходы, независимо от пола (рисунок 2.7).

Рисунок 2.7 – Индивидуальные расходы в зависимости от индекса массы тела с учетом пола и вредной привычки



пола и вредной привычки

Функция boxplot()

boxplot() – диаграмма размахов или «ящики с усами» (англ. *box- whisker plots*) имеет характерный вид ящика включающего в себя 50 % срединных значений (между квантилями 0,25

и 0,75). На диаграмме размахов также визуально определяется наличие в данных выбросов (рисунок 2.8).

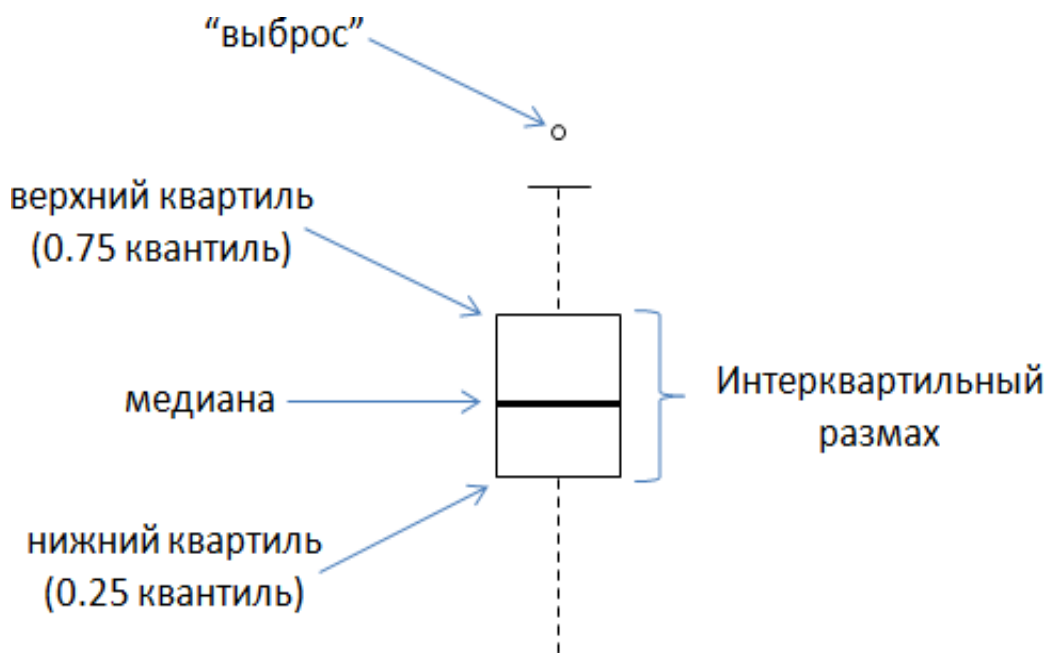


Рисунок 2.8 – Диаграмма размаха

Основные аргументы функции `boxplot()`:

- `formula` – формула, такая как `y ~ group`, где `y` – числовой вектор значений данных, которые должны быть разбиты на группы в соответствии с группирующей переменной `group` (обычно фактор);

- `data` – набор данных, переменные из которого входят в формулу.

Построим диаграмму размаха для индивидуальных медицинских расходов:

```
boxplot(insurance$charges, main = "Медицинские расходы")
```

Как видим (рисунок 2.9), в данных довольно много выбросов: это тяжелые заболевания у отдельных застрахованных, которые сопряжены с очень высокими расходами, нехарактерными для основной части наблюдений.



Рисунок 2.9 – Диаграмма размаха для медицинских расходов

Построим диаграммы размахов по медицинским расходам для двух частей выборки, разделенной по наличию вредной привычки. В качестве первого аргумента для `boxplot()` теперь используется функция: слева расходы – зависимая величина, справа наличие вредной привычки – независимая величина, второй аргумент – имя набора данных:

Разница в медицинских расходах между курящими и некурящими очень существенная (рисунок 2.10).

```
boxplot(charges~smoker, insurance, main = "Медицинские расходы в зависимости от курения")
```

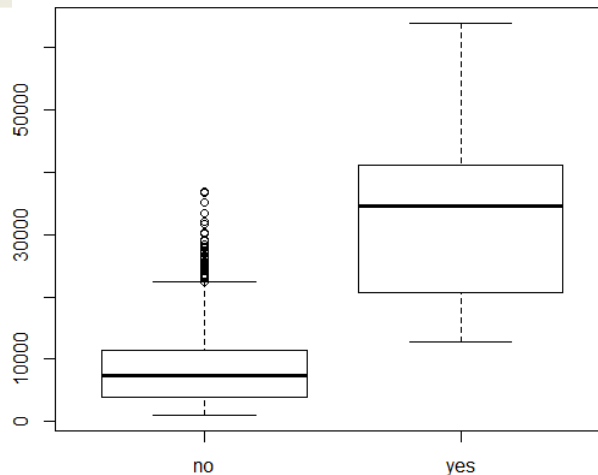


Рисунок 2.10 – Диаграмма размаха медицинских расходов для курящих и некурящих

Построим диаграммы размаха для медицинских расходов в зависимости от количества детей (рисунок 2.11):

```
boxplot(charges~children, insurance, main = "Медицинские расходы в зависимости от количества детей")
```

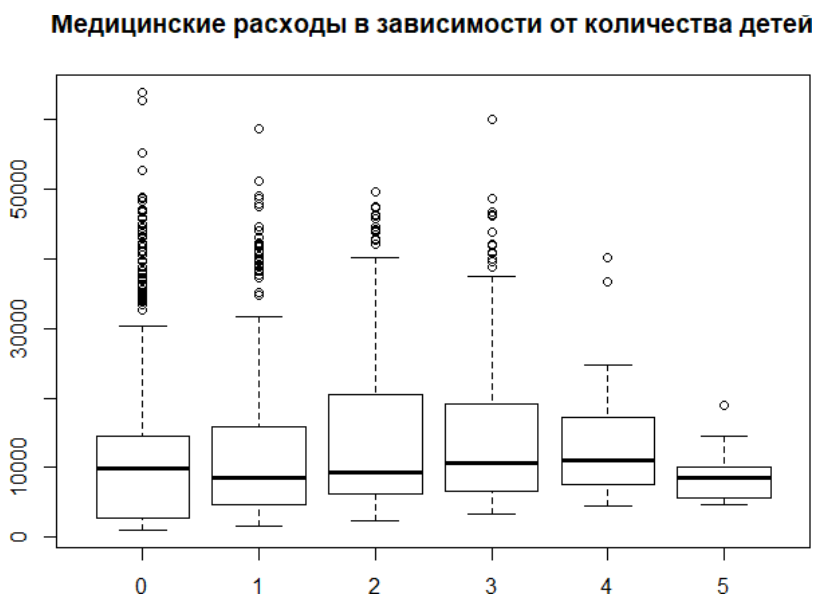


Рисунок 2.11 – Диаграмма размаха медицинских расходов при различном количестве

детей

Построим диаграммы размахов по медицинским расходам для сочетания двух факторов: наличия вредной привычки и пола. Для этого введем в исходный набор данных две новые переменные: `sex1` и `smoker1`, в которых запишем вместо исходных значений переменных их метки: 1 – женский, 2 – мужской и, соответственно, 1 – некурящий, 2 – курящий:

```
insurance$smoker1 <- as.numeric(factor(insurance$smoker))
insurance$sex1 <- as.numeric(factor(insurance$sex))
```

Далее представим значения этих переменных как факторы двумя уровнями, которым соответствуют определенные метки:

```
insurance$smoker1 <- factor(insurance$smoker1, levels=c(1, 2),
labels=c("некурящий", "курящий"))
insurance$sex1 <- factor(insurance$sex1, levels=c(1, 2), labels=c("женский", "мужской"))
```

Теперь можно построить график, в котором диаграммы размаха для медицинских расходов построены для различных сочетаний значений факторов: пол и наличие вредной привычки (рисунок 2.12):

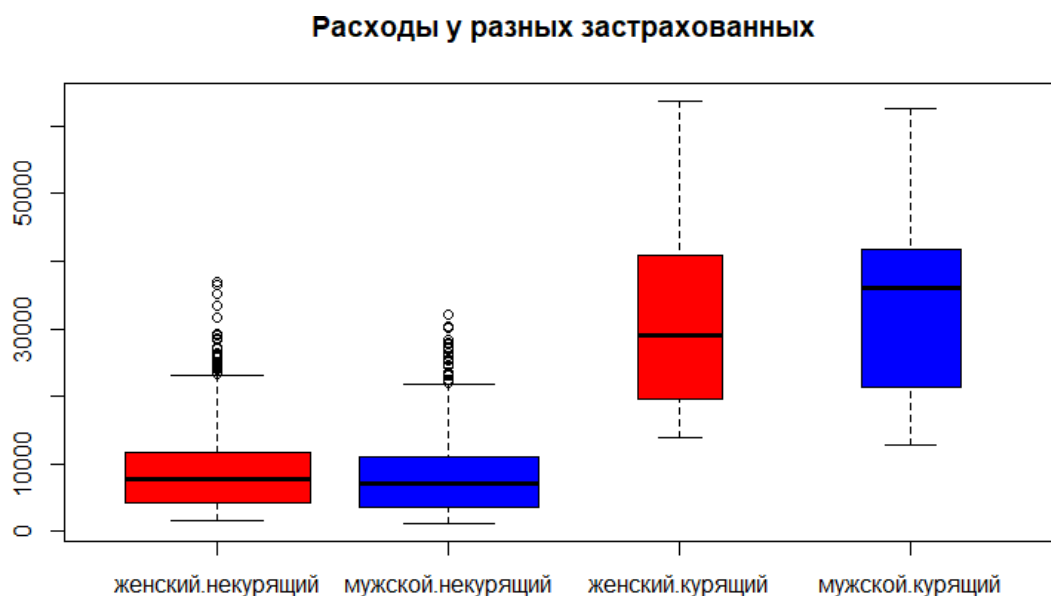


Рисунок 2.12 – Диаграмма размаха медицинских расходов при различном сочетании значений факторов: пол и наличие вредной привычки

Функция mosaicplot()

mosaicplot() – это графический метод визуализации данных из двух или более качественных переменных.

Основные аргументы функции **mosaicplot()**

– **x** – таблица сопряженности в виде массива с дополнительными категориями, указанными в **dimnames(x)**. Таблицу лучше всего создавать командой **table()**.

Построим таблицу сопряженности для факторов: пол и наличие вредной привычки:

```
x <- table(insurance$sex, insurance$smoker)
```

Таблица сопряженности разбивает все множество наблюдений на четыре группы: курящие женщины, курящие мужчины, некурящие женщины и некурящие мужчины. Если между этими показателями есть связь, то ее наличие можно обнаружить с помощью графического представления таблицы сопряженности благодаря применению к таблице функции **mosaicplot()**:

```
mosaicplot(x, color = T, shade = T, main = "Сопряженность признаков пол  
и курение", xlab = "пол", ylab = "курение")
```

На рисунке 2.12 видно, что связи между этими показателями нет, хотя среди женщин, застрахованных в компании, меньший процент курящих, чем среди мужчин.



Рисунок 2.12 – Сочетание в наборе данных признаков: пол и наличие вредной привычки

ВОЗМОЖНОСТИ ПАКЕТА GGPLOT2

Первоначально подключим пакет **ggplot2**

```
library(ggplot2)
```

Базовой графической командой пакета **ggplot2** является команда **qplot()**.

Основные аргументы функции **qplot()**:

- **x, y** – переменные набора данных;
- **data** – набор данных, переменные из которого входят в формулу;
- **color, shape, size, fill** – определяет соответствие цвета, формы и размера символов уровням переменной, для диаграмм плотности и размахов параметр **fill**

определяет соответствие заполняющего цвета значениям переменной;

- объекты, которые задают тип диаграммы. Этот параметр задается в формате текстового вектора с одним или более элементами: "point", "smooth", "boxplot", "line", "histogram", "density", "bar" и "jitter";

- facets — используется для построения категоризованных диаграмм: если используются две факторные переменные для разбиения набора данных, то между ними ставится ~: facets = sex1~smoker1, если одна переменная, то либо facets=children~, либо facets= . ~ children;

- method — можно использовать следующие методы сглаживания (параметр formula): "lm" (регрессия), "gam" (обобщенные аддитивные модели) и "glm" (устойчивая регрессия).

Построим точечный график зависимости индивидуальных медицинских расходов от возраста (рисунок 3.1):

```
qplot(x = age, y = charges, data = insurance, geom = "point", xlab = "Возраст", ylab = "Расходы")
```

Построим диаграммы размахов медицинских расходов для курящих и некурящих клиентов страховой компании (рисунок 3.2):

```
qplot(smoker1, charges, data = insurance, geom = c("boxplot", "jitter"),
fill=smoker1, xlab="Наличие вредной привычки",
ylab="Расходы",main="Диаграммы размахов")
```

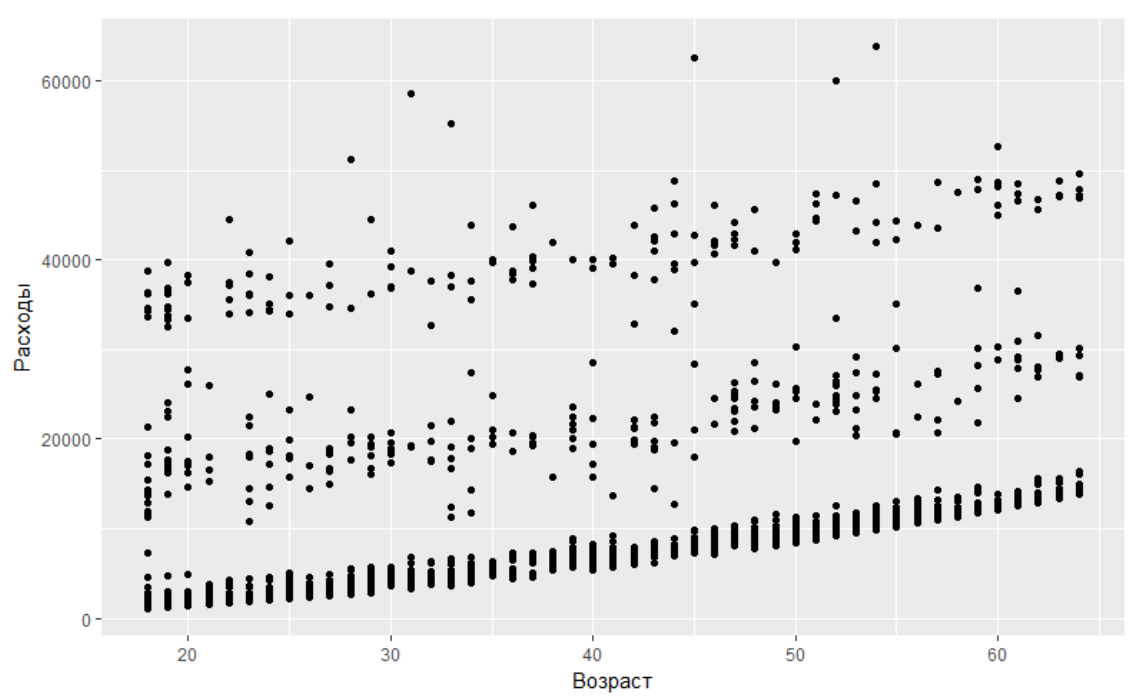


Рисунок 3.1 – Точечный график расходов от возраста застрахованного

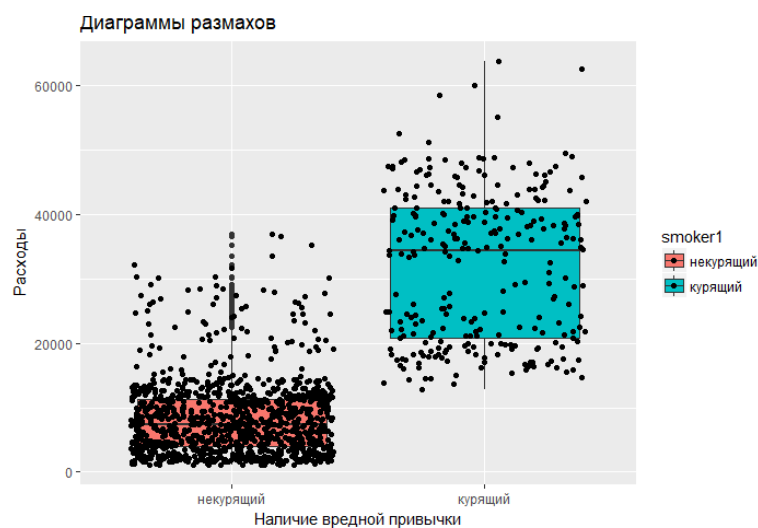


Рисунок 3.2 – Диаграммы размахов расходов для курящих и некурящих клиентов

`qplot(ages, data = insurance, facets = sex~., fill=sex, xlab = "Расходы")` – точечный график расходов от возраста застрахованного (в зависимости от сочетания значений факторов: пол и наличие вредной привычки (рисунок 3.3):

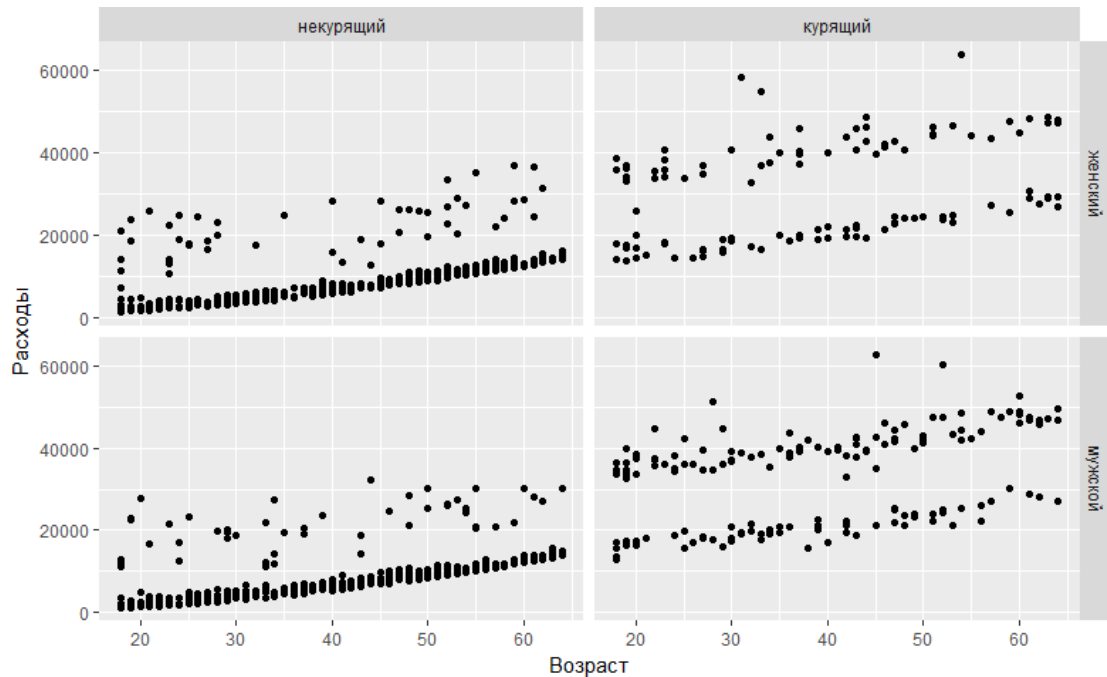
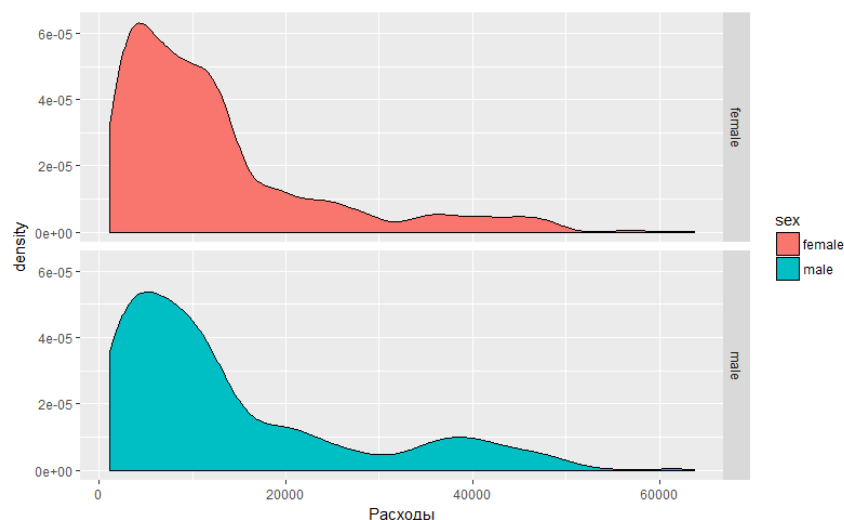


Рисунок 3.3 – Точечный график расходов от возраста застрахованного (в зависимости от сочетания значений факторов: пол и наличие вредной привычки)

Построим график плотности (`geom="density"`) распределения расходов для выборок разделенных по полу (рисунок 3.4):

`qplot(charges, data = insurance, geom="density", facets=sex~., fill=sex, xlab = "Расходы")`

Рисунок 3.4 – График плотности распределения расходов



(для мужчин и женщин)

Построим график плотности распределения расходов для выборок, разделенных по количеству детей (рисунок 3.5):

```
qplot(charges, data = insurance, geom="density",  
facets=children~., fill=children, xlab = "Расходы")
```

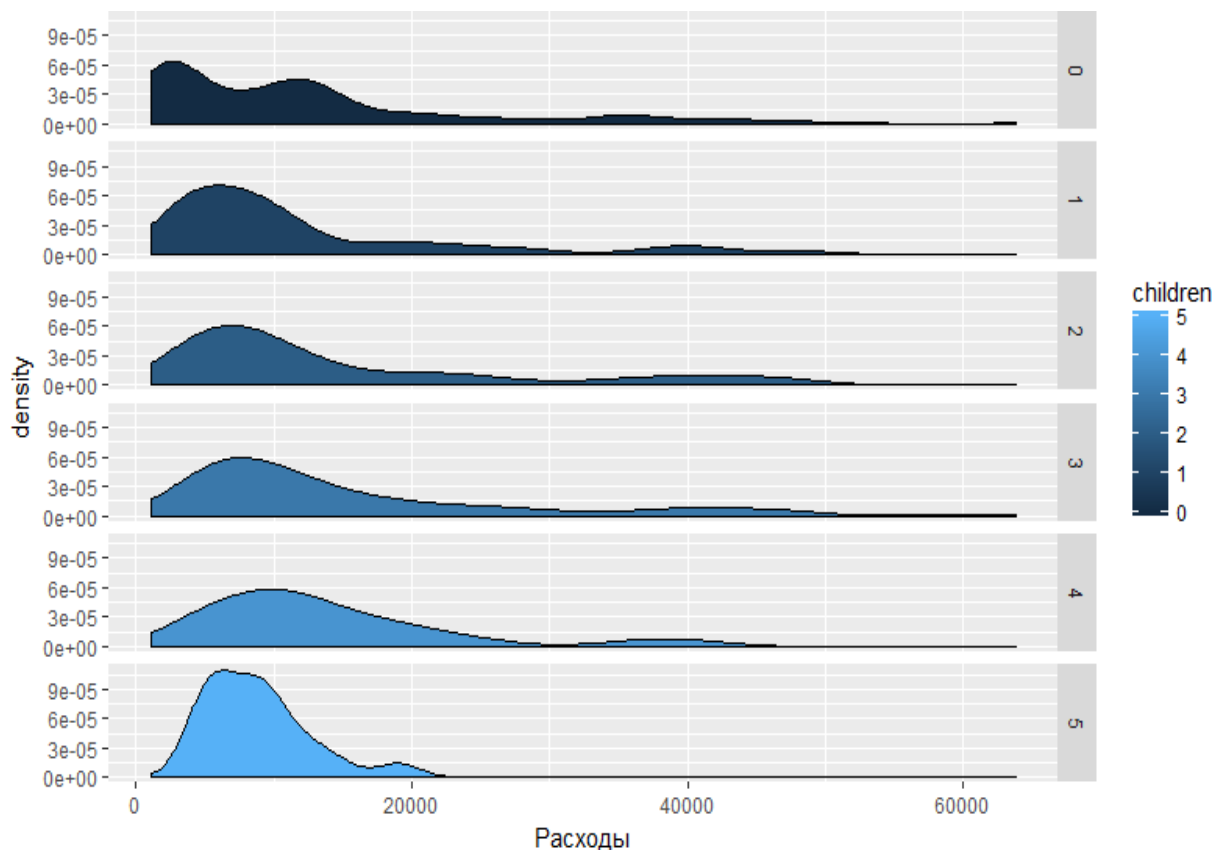


Рисунок 3.5 – График плотности распределения расходов
(в зависимости от числа детей)

Построим две регрессии на одной координатной плоскости (регрессия – это функциональная зависимость между случайными величинами, одна из которых рассматривается как независимая, другая зависимая). Первая регрессия для некурящих, вторая для курящих клиентов страховой компании (рисунок 3.6). Очевидно, что медицинские расходы у этих групп клиентов существенно различаются (в пользу некурящих):

```
qplot(age, charges, data = insurance, color = smoker1,  
shape = smoker1, geom = c("point", "smooth"), method = "lm", formula = y~x,  
xlab = "Возраст", ylab = "Расходы",  
main = "Пример с регрессией")
```

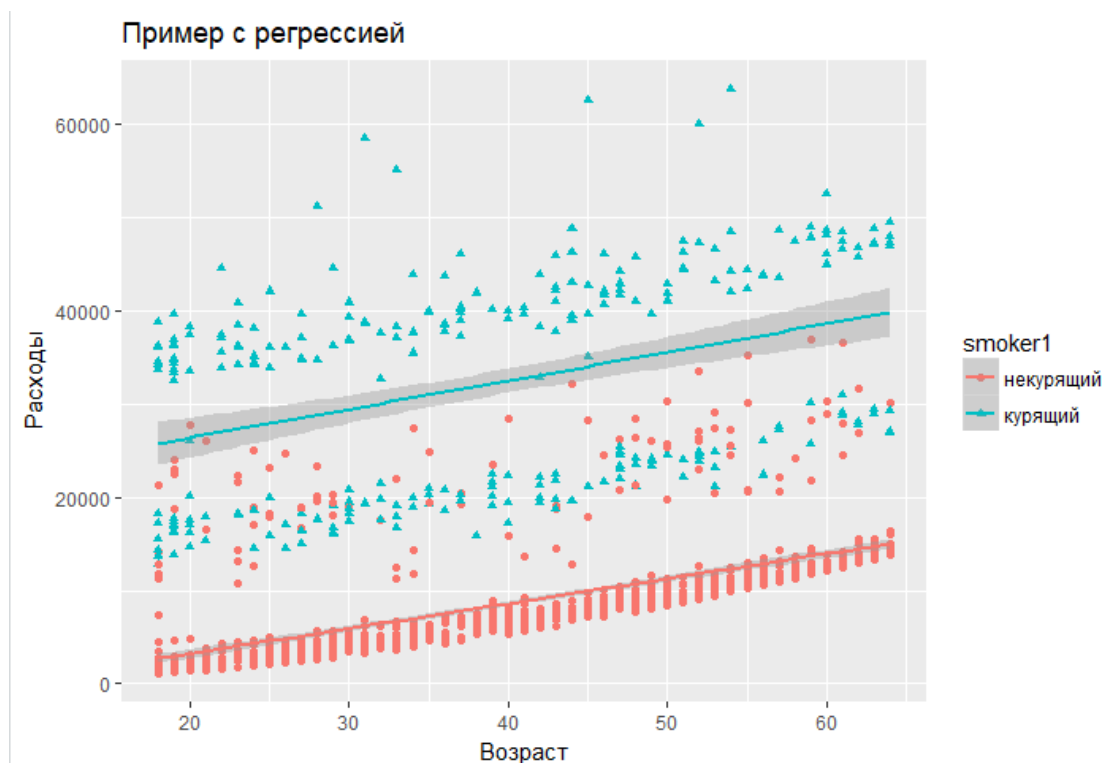


Рисунок 3.6 – Линейные регрессии расходов в зависимости от возраста отдельно для

курящих и некурящих клиентов.

Задание 1.

«Визуализация данных: базовая графика»

Цель работы: научиться импортировать данные и применять к ним базовые графические функции языка R для визуализации и выдвижения гипотез о наличии зависимостей.

Формируемые знания, умения и навыки: изучить базовые графические возможности языка R, уметь импортировать данные из доступных источников.

Необходимо:

1. С сайта Kaggle (<https://www.kaggle.com/>) импортировать один из наборов, включающий столбцы с числовыми значениями размером <5 Mb.
2. Используя базовые графические возможности языка R, провести визуальный анализ данных и выдвинуть гипотезы о наличии зависимостей между переменными набора.

Контрольные вопросы и задания

1. Какая функция используется для импорта таблиц данных в формате .csv? Какие аргументы у этой функции? Какие значения могут принимать аргументы этой функции?
2. Какие аргументы у функции plot()? Какие значения могут принимать эти аргументы?
3. Какие аргументы у функции legend()? Какие значения они могут принимать?
4. Опишите выбранный набор данных: контекст данных, переменные из набора, зависимости между переменными, выявленные визуально.

Задание 2.

«Визуализация данных: пакет ggplot2»

Цель работы: научиться использовать расширенные графические возможности, предоставляемые пакетом ggplot2.

Формируемые знания, умения и навыки: изучить возможности визуализации данных для разведочного анализа.

Необходимо:

4. На том же наборе данных провести построение графиков с помощью функции `qplot()` пакета ggplot2.

5. Самостоятельно изучить возможности визуализации данных с помощью функции `ggplot()` пакета ggplot2.

Контрольные вопросы и задания

1. Какие аргументы могут быть у функции `qplot()`?
2. Какие значения может принимать аргумент `geom`?
3. Как определяется `facet`?
4. Какие значения может принимать `method`?
5. Как определяется функция `ggplot()`?

Лабораторные работы № 3-4

Тема: Решение задач по бизнес-анализу

Формируемые компетенции или их части: ИД-1.2 Осуществляет поиск алгоритмов решения поставленной проблемной ситуации на основе доступных источников информации. Определяет в рамках выбранного алгоритма вопросы (задачи), подлежащие дальнейшей детальной разработке. Предлагает способы их решения

ИД-1.1 Выполняет разработку методик выполнения аналитических работ

ИД-5.1 Обладает навыками управления исследованиями новых рынков

Практические задания для контрольной работы

Задача 1

Используя данные таблицы 1 (вместо __ __ поставьте 2 последние цифры зачетной книжки): определите удельный вес отдельных видов продукции в общем объеме выпуска, абсолютные и относительные изменения в выпуске продукции в отчетном периоде по сравнению с базисным периодом.

По результатам расчетов сделайте выводы.

Таблица 1 - Анализ выпуска продукции по ассортименту

Вид продукции	Объем выпуска продукции								Отклонение				Отчетный период к базисному периоду, %	
	базисный период				отчетный период				в количестве		в сумме		по количеству	по сумме
			сумма		количество		сумма		тыс.шт.	удель- ный вес, %	тыс.руб.	удель- ный вес, %		
	тыс. шт.	удель- ный вес, %	тыс.руб.	удель- ный вес, %	тыс.шт.	удель- ный вес, %	тыс.руб.	удель- ный вес, %						
A	780		6 1__		800		8 7__							
B	1 3 _ _		8 1__		1 2__		12 __3							
C	__ 0		3 3 __		540		5 __9							
D	__ 3		__61		250		2 __1							
Итого														

Задача 2

Определить объемы отгрузки и продажи продукции, отклонения и темпы изменения показателей за год, влияние факторов на изменение объема продажи продукции (вместо __ __ поставьте 2 последние цифры зачетной книжки, данные выпуск продукции за год взять из таблицы 1).

По результатам расчетов сделайте выводы. Таблица

2 - Анализ продажи продукции

Показатель	Базисный год, тыс. руб.	Отчетный год, тыс. руб.	Отклонение, тыс. руб.	Отчетный период к предшествующему периоду, %
Остаток готовой продукции на начало года	590	577		
Выпуск продукции за год				
Остаток готовой продукции на конец года	5 __	5 __		
Отгрузка продукции				
Прочее выбытие	2	3		
Продажа продукции за год				

Задача 3

Данные об итоговых объемах выпусках продукции за год взять из таблицы 1. Исчислите выпуск продукции по месяцам и нарастающим итогом, рассчитайте темпы роста (снижения) выпуска продукции, рассчитайте коэффициент ритмичности работы организации по выпуску продукции

$$K_{рит} = \frac{V_{факт}}{V_{план}}$$

среднеквадратическое отклонение и коэффициенты вариации за год

$$K_s = \frac{\sqrt{\sum \Delta^2 : n}}{\bar{X}_{пл}}$$

проанализируйте исчисленные показатели.

Таблица 2 - Анализ данных о выпуске продукции по месяцам

Месяцы	Базисный год			Отчетный год			Отчетный год — по плану		
	за месяц		нарастающ им итогом, тыс. руб.	за месяц		нарастающ им итогом, тыс. руб.	за месяц		нарастающ им итогом, тыс. руб.
	тыс . руб .	Удель -ный вес, %		тыс . руб .	Удель -ный вес, %		тыс. руб.	Уде ль- ный вес, %	
Январь		7,97			7,81		2 300		
Феврал ь		8,23			7,99		2 350		
Март		7,78			8,17		2 400		
Апрель		8,38			8,27		2 430		
Май		7,92			8,10		2 380		
Июнь		8,59			8,52		2 500		
Июль		7,13			8,34		2 450		
Август		7,86			8,45		2 480		
Сентябр ь		8,83			8,34		2 450		
Октябр ь		8,33			8,41		2 470		
Ноябрь		8,31			8,52		2 500		
Декабрь							2 654		
Итого		100			100		29364	100	

Задача 4

На основании приведенных данных рассчитать показатели таблицы 4, коэффициент сортности продукции и общий индекс качества продукции. Проанализируйте исчисленные показатели и укажите основные пути повышения качества продукции организации.

n

$$K_{сорт} = \frac{\sum_{i=1}^n Q_i * P_i}{P_{\max} * \sum_{i=1}^n Q_i} - \text{коэффициент сортности продукции}$$

Q_i — количество товаров i -го сорта;

P_i — цена единицы товара i -го сорта;

$P_{\max}$ — цена единицы продукции высшего сорта;

n — количество сортов данного вида продукции.

$$\Delta K_{сорт} = \frac{K_{сорт}^{\phi}}{K_{сорт}^{баз}} - \text{общий индекс качества продукции}$$

(вместо __ поставьте 2 последние цифры зачетной книжки)

Таблица 4 - Анализ качества продукции

Продукция	Базисный год				Отчетный год				Отклонение		
	количество		цена, руб.	сумма, тыс. руб.	количество		цена, руб.	сумма, тыс. руб.	в количестве , шт.	в цене, руб.	в сумме, тыс. руб.
	тыс. шт.	структур а, %			тыс. шт.	структу ра, %					
А	780			6 1__	800			8 7__			
высший сорт	260			2300	320			4300			
первый сорт	300			2400	310			3600			
второй сорт											
В	13 __			8 1__	12__			12 __3			
высший сорт	720			6000	790			9000			
первый сорт											
С	__ 0	100,0		3 3 __	540	100,0		5 __9			
первый сорт		60,0		2 300		55,0		3300			
второй сорт		40,0				45,0					
Д	__ 3	100,0		_ _61	250	100,0		2 __1			
первый сорт		72,0	11,80			74,0	*				
второй сорт		28,0				26,0					
Итого											

* на 1 руб. больше чем средняя цена по продукции Д

Задача 5

Рассчитайте влияние изменения цен и количества проданной продукции на изменение выручки от реализации, определите процент изменения за счет цены и количества реализованной продукции;

Проанализируйте исчисленные показатели и разработайте соответствующие предложения.
(данные для заполнения таблицы 5 взять из таблицы 4)

Таблица 5 - Реализация продукции и влияние факторов на изменение выручки

Продукция	Количество реализованной продукции, тыс. шт.		Средняя цена реализации, руб.		Выручка от реализации, тыс. руб.			Изменение выручки от реализации, тыс. руб.		
	базисный год	отчетный год	базисный год	отчетный год	базисный год	отчетный год	объем отчетный цены базисного года	всего	в т.ч. за счет изменения	
									объема реализации	средней цены реализации
1	2	3	4	5	6	7	8=3*4	гр.7-гр.6	гр.8-гр.6	гр.7-гр.8
А										
высший сорт										
первый сорт										
второй сорт										
В										
высший сорт										
первый сорт										
С										
первый сорт										
второй сорт										
Д										
первый сорт										
второй сорт										
Итого и в среднем										

Задача 6

Определить влияние влияния структурных сдвигов, объема и цены реализации продукции А, В, С, D

$$N = Q^n \sum_i D_i^n * p_i^n, \text{ где}$$

N – выпуск продукции в стоимостном выражении

Q - выпуск продукции в натуральном выражении

D – удельный вес каждого наименования (коэф)

P – цена за единицу

i – наименование продукции

n – 0-базисный период, 1-отчетный период

Исходные данные (вместо двоеточия последние цифры зачетки): (данные для заполнения таблицы 6 взять из таблицы 5)

Таблица 6 - Влияние структурных сдвигов, объема и цены реализации продукции

Продукция	Количество реализованной продукции, тыс. шт.		Структура реализованной продукции (коэф)		Средняя цена реализации, руб.		Изменение выручки от реализации, тыс. руб.			
	базисный год	отчетный год	базисный год	отчетный год	базисный год	отчетный год	всего	в т.ч. за счет изменения		
								объема реализации	структуры реализованной продукции	средней цены реализации
A							X	X	X	X
B										
C										
D										
Итого и в среднем			1,000	1,000						

Задача 7

Фирма производит два вида мороженого - сливочное и шоколадное. Для его изготовления используются два исходных продукта: молоко и наполнители, расходы которых на 1 кг готового продукта и их суточные запасы приведены в таблице:

Исходный продукт	Расход исходных продуктов на 1 кг мороженого, кг		Запас, кг
	Сливочное	Шоколадное	
Молоко	0,8	0,5	400
Наполнители	0,4	0,8	365

Суточный спрос на сливочное мороженое превышает спрос на шоколадное не более чем на 100 кг. Кроме того, известно, что спрос на шоколадное мороженое не превышает 350 кг в сутки. Отпускная цена 1 кг сливочного мороженого 16 руб., шоколадного - 14 руб.

Определить в каком количестве мороженого каждого вида должна производить фирма, чтобы доход от реализации продукции был максимальным.

Задача 8

Требуется спланировать перевозку строительных материалов с трех заводов к четырем строительным площадкам по железной дороге. В течение каждого квартала на четырех площадках требуется, соответственно, 5, 10, 20, 15 вагонов строительных материалов. Возможности заводов, соответственно, равны 10, 15 и 25 вагонов в квартал. В таблице приведены стоимости перевозки одного вагона от различных заводов к различным строительным площадкам.

Завод	Стоимость перевозки 1 вагона, тыс. руб. от завода до строительной площадки			
	П1	П2	П3	П4
А	8	3	5	2
В	4	1	6	4
С	1	9	4	3

Задача 9

Организация рассматривает инвестиционный проект - приобретение новой техники. Стоимость линии – 15 млн. руб. Срок эксплуатации – 5 лет, износ оборудования исчисляется по методу прямолинейной амортизации. Суммы, вырученные от ликвидации оборудования в конце срока эксплуатации, покрывают расходы по его демонтажу.

Выручка от реализации продукции прогнозируется по годам в следующих объемах: 10200 тыс. руб.; 11100 тыс. руб.; 12300 тыс. руб.; 12000 тыс. руб.; 9000 тыс. руб.

Текущие расходы по годам осуществляются следующим образом: 5100 тыс. руб. в первый год эксплуатации; ежегодно эксплуатационные расходы увеличиваются на 4 %. Ставка налога на прибыль составляет 20 %.

Цена авансируемого капитала – 14 %. Стартовые инвестиции проводятся за счет собственных средств.

Определить является ли данный проект эффективным по показателю NPV (чистая

приведенная стоимость).

Задача 10

Необходимо определить размер необходимой процентной ставки для следующих двух ситуаций при условии, что платежи равного размера вносятся в конце года на протяжении 5 лет, а проценты начисляются один раз в год.

1. Создать фонд, равный 1 млн. руб., перечисляя ежегодно по 170,456 тыс. руб.
2. Погасить текущую задолженность, равную 1 млн. руб., выплачивая ежегодно по 250,456 тыс. руб.

Задача 11

Компания планирует приобрести новое оборудование стоимостью 7000 тыс. руб. и сроком эксплуатации 5 лет. Компания будет получать дополнительный денежный приток в размере 2500 тыс. руб. ежегодно. Известно, что на третьем году эксплуатации оборудованию потребуется плановый ремонт стоимостью 300 тыс. руб. Определить внутреннюю норму рентабельности.

Задача 12

Рассчитайте показатели, характеризующие автоматизацию работ. По результатам расчетов сделайте выводы.

Данные об итоговых объемах производства продукции за год взять из таблицы 1. (вместо __ поставьте 2 последние цифры зачетной книжки)

Таблица - Показатели, характеризующие степень автоматизации в организации

Показатель	Базисный год	Отчетный год	Отклонение
Объем выпуска продукции, тыс. в том числе произведенной на автоматизированном оборудовании,	18 0 __	26 __ 5	
Отработано часов — всего	15 __ 78	165 __ 9	
в том числе время автоматизированных работ, часы	10 __ 13	105 __ 1	
Коэффициент автоматизации работ, доли единицы:			
а) в организации			
б) на отраслевом уровне	0,890	0,914	
Коэффициент автоматизации труда, доли единицы:			
а) в организации	0,860	0,870	
б) на отраслевом уровне			
Коэффициент автоматизации по			

трудоемкости, (стр. 4 : стр. 3} : а) в организации			
--	--	--	--

Задача 13

Рассчитать недостающие показатели и оценить эффективность управления организаций. По результатам расчетов сделайте выводы.

Данные об итоговых объемах выпуска продукции за год взять из таблицы 2. (вместо __ поставьте 2 последние цифры зачетной книжки)

Таблица – Показатели эффективности управления

Показатель	Базисный год	Отчетный год	Отклонение (+; -)	Отчетный % базисному	в к
Объем проданной продукции, тыс. руб.					
Полная себестоимость про- данной продукции, тыс. руб.	14 777	22 862			
Прибыль до налогообложения, тыс. руб.	4 320	4 926			
Среднегодовая стоимость основных средств, тыс. руб.	12 __ 5	13 5 __			
Среднегодовая численность ППП, чел.	126	137			
Численность работников АУП, чел.	23	21			
Затраты на управление, тыс. руб.	15 __	19 __			
Приходится на одного работника АУП: а) проданной продукции, тыс. руб. (стр. 1 : стр. 6) б) прибыли, тыс. руб. (стр. 3 : стр. 6} в) стоимости основных производственных средств, тыс. руб. (стр. 4 : стр. 6) г) численности ППП, чел. (стр. 5 : стр. 6)					
Удельный вес затрат на управление в себестоимости, % (стр. 7 : стр. 2) x 100%					
Коэффициент эффективности управления: стр. 7 : стр. 1					

Задача 14

Рассчитайте недостающие показатели и оцените организационный уровень производства. По результатам расчетов сделайте выводы.

Данные об итоговых объемах выпуска продукции за год взять из таблицы 1. (вместо __ поставьте 2 последние цифры зачетной книжки)

Таблица - Показатели организационного уровня производства

Показатель	Базисный год	Отчетный год	Отклонение
Объем выпуска продукции, тыс. ед.			
Выпущено продукции на поточных линиях, тыс. ед.	2125	2450	
Коэффициент поточности (стр. 2 : стр. 1)			
Стоимость выпущенной продукции, тыс. руб.			
Стоимость полуфабрикатов, тыс. руб.	974	1 975	
Коэффициент кооперирования (стр. 5 : стр.			
Стоимость продукции, соответствующей производственному	16 __ 1	23 4 __	
Коэффициент специализации, доли единицы (стр. 7 : стр. 4)			

Задача 15

Определите производительность труда, фондоотдачу и материалоотдачу. Выявите влияние факторов на производство продукции в действующих и в сопоставимых ценах (при цене отчетного года и количестве выпуска продукции базисного года).

По результатам расчетов сделайте выводы.

Данные об итоговых объемах выпуска продукции за год взять из таблицы 1. (вместо __ __ поставьте 2 последние цифры зачетной книжки).

Таблица - Показатели экстенсивности и интенсивности использования ресурсов в производстве продукции

Показатели	Базисный год	Отчетный год	Откло нение	Отчетн ый год в % к
Объем выпуска продукции в действующих ценах, тыс. руб.				
Объем выпуска продукции в сопоставимых ценах, тыс. руб.				
Среднегодовая стоимость основных средств, тыс. руб.	12 __ 5	13 5 __		
Фондоотдача в действующих ценах (стр. 1: стр.3), руб.				
Фондоотдача в сопоставимых ценах (стр. 2 : стр.3), руб.				
Среднесписочная численность работников, чел.	149	158		
Среднесписочная численность ППП, чел.	126	137		
Производительность труда работников:				
в действующих ценах, тыс. руб.				
в сопоставимых ценах, тыс. руб.				
Производительность труда рабочих:				
в действующих ценах, тыс. руб.				
в сопоставимых ценах, тыс. руб.				
Материальные затраты, тыс. руб.	6 1 __	10 5 __		
Материалоотдача:				
в действующих ценах, руб.				
в сопоставимых ценах, руб.				

Задача 16

Организация приобрела объект основных средств стоимостью 600 тыс. руб. Амортизация начисляется линейным способом исходя из срока службы объекта в 10 лет. Ежемесячная сумма амортизации вносится на счет в банке под 12% годовых (проценты сложные, начисление 1 раз в месяц). Хватит ли организации накопленной на банковском счете суммы для приобретения через 10 лет аналогичного объекта основных средств по цене 1 млн. руб.?

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 17

Предполагается, что станок будет служить 3 года, принося ежегодный доход в 60 тыс. руб. В качестве альтернативы потенциальный покупатель станка рассматривает вложение денег на депозит под ставку 8% годовых (проценты сложные, начисление 1 раз в год). Определите верхний предел цены для покупателя станка.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 18

Студент 1 курса Петров очень хочет к моменту окончания обучения в ВУЗе «С» купить себе квартиру (хотя бы однокомнатную). В данный момент времени такая квартира стоит 900 тыс. руб. Петров предполагает, что за 5 лет, которые он будет учиться в ВУЗе, недвижимость вырастет в цене на 10%. Он планирует открыть счет в банке и ежемесячно перечислять из стипендии и иных доходов некоторую сумму средств. По условиям договора банковского счета на вклад начисляются сложные проценты по ставке 12% годовых при ежемесячном начислении процентов. Определите какую сумму необходимо перечислять Петрову, чтобы он смог купить себе квартиру в случае если его прогноз относительно роста цен оправдается.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 19

Необходимо определить размер необходимой процентной ставки для следующих двух ситуаций при условии, что платежи равного размера вносятся в конце года на протяжении 5 лет, а проценты начисляются один раз в год.

1. Создать фонд, равный 1 млн. руб., перечисляя ежегодно по 170,456 тыс. руб.

2. Погасить текущую задолженность, равную 1 млн. руб., выплачивая ежегодно по 250,456 тыс. руб.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 20

Компания планирует приобрести новое оборудование стоимостью 7000 тыс. руб. и сроком эксплуатации 5 лет. Компания будет получать дополнительный денежный приток в размере 2500 тыс. руб. ежегодно.

Известно, что на третьем году эксплуатации оборудованию потребуется плановый ремонт стоимостью 300 тыс. руб. Необходимо обосновать целесообразность приобретения оборудования (по критерию чистой приведенной стоимости), если стоимость капитала по проекту составляет 20%.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 21

Организация рассматривает инвестиционный проект - приобретение новой техники. Стоимость линии – 15 млн. руб. Срок эксплуатации – 5 лет, износ оборудования исчисляется по методу прямолинейной амортизации. Суммы, вырученные от ликвидации оборудования в конце срока эксплуатации, покрывают расходы по его демонтажу.

Выручка от реализации продукции прогнозируется по годам в следующих объемах: 10200 тыс. руб.; 11100 тыс. руб.; 12300 тыс. руб.; 12000 тыс. руб.; 9000 тыс. руб.

Текущие расходы по годам осуществляются следующим образом: 5100 тыс. руб. в первый год эксплуатации; ежегодно эксплуатационные расходы увеличиваются на 4 %. Ставка налога на прибыль составляет 20 %.

Цена авансируемого капитала – 14 %. Стартовые инвестиции проводятся за счет собственных средств.

Определить является ли данный проект эффективным по показателю NPV (чистая приведенная стоимость).

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 22

Компания планирует 1 июля 20_____г. приобрести новое оборудование стоимостью 7000 тыс. руб. и сроком эксплуатации 10 лет. Ожидаемая прибыль: через 1 год – 6 млн. руб., через 2,5 года – 4 млн. руб., через 7 лет – 1,5 млн. руб., через 9 лет – 0,5 млн. руб. Необходимо обосновать целесообразность приобретения оборудования, если стоимость капитала по проекту составляет 20%.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 23

Компания планирует приобрести новое оборудование стоимостью 7000 тыс. руб. и сроком эксплуатации 5 лет. Компания будет получать дополнительный денежный приток в размере 2500 тыс. руб. ежегодно.

Известно, что на третьем году эксплуатации оборудованию потребуется плановый ремонт стоимостью 300 тыс. руб. Определить внутреннюю норму рентабельности.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 24

Компания планирует 1 июля 20_____г. приобрести новое оборудование стоимостью 7000 тыс. руб. и сроком эксплуатации 10 лет. Ожидаемая прибыль: через 1 год – 6 млн. руб., через 2,5 года – 4 млн. руб., через 7 лет – 1,5 млн. руб., через 9 лет – 0,5 млн. руб. Необходимо

определить внутреннюю норму рентабельности.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 25

Компания планирует приобрести новое оборудование стоимостью 7000 тыс. руб. и сроком эксплуатации 5 лет. Компания будет получать дополнительный денежный приток в размере 2500 тыс. руб. ежегодно.

Известно, что на третьем году эксплуатации оборудованию потребуется плановый ремонт стоимостью 300 тыс. руб. Необходимо обосновать целесообразность приобретения оборудования (по критерию чистой приведенной стоимости), если стоимость капитала по проекту составляет 20%.

Решить задачу с помощью финансовых функций Microsoft Excel.

Описать и объяснить ход решения.

Задача 26

Рассчитайте материалоемкость и материалоотдачу в целом по всем материальным ресурсам и их видам. Выясните, каково влияние на материальные ресурсы объема выпуска продукции, материалоотдачи и материалоемкости.

По результатам расчетов сделайте выводы.

Таблица – Материалоемкость и материалоотдача по организации

Показатели	200_г.	200_г.	200_г.	отклонение 200_г. в % к	
				200_г.	200_г.
Выпуск продукции в действующих ценах, тыс. руб.					
Материальные затраты, тыс. руб.	1 __ 15	12 __ 1	13 1 __		
в том числе					
сырье и материалы (60%)					
полуфабрикаты (5%)					
топливо (7%)					
электроэнергия (16%)					
другие материалы (12%)					
Материалоемкость общая, руб.					
в том числе					
сырье и материалы					
полуфабрикаты					
топливо					
электроэнергия					
другие материалы					
Материалоотдача общая, руб.					
в том числе					
сырье и материалы					

полуфабрикаты					
топливо					
электроэнергия					
другие материалы					

Задача 27

Определите прибыль от продаж на 1 руб. материальных затрат, рассчитайте стоимость переработанных материалов на одного работника, исчислите производительность труда. Определите влияние факторов на прибыль на 1 руб. материальных затрат рентабельности продаж и материалоотдачи.

$$\frac{\Pi}{МЗ} = \frac{\Pi}{ВП} \times \frac{ВП}{МЗ} = P_{об} \times M_{отд}$$

По результатам расчетов сделайте выводы.

Таблица – Показатели эффективности использования материальных затрат

Показатели	200_г.	200_г.	200_г.	отклонение 200_г. от (+,-)	
				200_г.	200_г.
Выпуск продукции в действующих ценах (ВП), тыс. руб.					
Прибыль от реализации продукции (П), тыс. руб.					
Материальные затраты (МЗ), тыс. руб.					
Среднесписочная численность работников, чел.					
Стоимость переработанных материалов на одного работника, тыс. руб.					
Производительность труда, тыс. руб.					
Прибыль от реализации продукции на 1 рубль материальных затрат, руб.					
Материалоотдача (Мотд), руб.					
Рентабельность оборота (Роб), %					
Изменение прибыли от реализации продукции на 1 рубль материальных затрат, всего, руб.	-	-	-		
в том числе за счет изменения Рентабельности оборота, руб.	-	-	-		
Материалоотдачи, руб.	-	-	-		

Задача 28

Имеются следующие данные по предприятию: выручка от реализации продукции за отчетный год - 4200 тыс. руб., за предыдущий год - 3830 тыс. руб.; оплата труда с начислениями в отчетном году - 510 тыс. руб., в базисном году - 480 тыс. руб. Определите относительное отклонение в затратах трудовых ресурсов (средствах на оплату труда).

Задача 29

Оцените долю экстенсивности и интенсивности в использовании средств труда на основе следующих данных: стоимость продукции в отчетном году - 625 млн руб., в предыдущем году - 573 млн руб.;

среднегодовая величина основных средств производства в отчетном году - 448 млн руб., в базисном году - 420 млн руб.

Задача 30

Имеются следующие данные по предприятию: выручка от реализации продукции за отчетный год - 4200 тыс. руб., за предыдущий год - 3830 тыс. руб.; оплата труда с начислениями в отчетном году - 510 тыс. руб., в базисном году - 480 тыс. руб. Определите темп прироста платоотдачи.

Задача 31

Определите относительное отклонение в затратах материальных ресурсов на основе следующих данных: выручка от реализации продукции за отчетный год - 98,4 млн руб., за предыдущий год - 92,7 млн руб.; затраты предметов труда в отчетном году - 64,4 млн руб., в базисном году - 62,1 млн руб.

Задача 32

Определите влияние экстенсивного (величины основных средств) и интенсивного (фондоотдачи) факторов в использовании средств производства на приращение продукции на основе следующих данных. Стоимость продукции в отчетном году - 625 млн руб., в предыдущем году - 573 млн руб. Среднегодовая величина основных средств производства в отчетном году - 448 млн руб., в базисном году - 420 млн руб.

Задача 33

Определите влияние факторов (выпуска продукции и материальных затрат) на материалоемкость. По результатам расчетов сделайте выводы.

Таблица – Факторный анализ материалоемкости продукции

Продукция	Выпуск продукции в действующих ценах (ВП), тыс. руб.		Материальные затраты (МЗ), тыс. руб.		Изменение материалоемкости, руб.		
	базисный год	отчетный год	базисный год	отчетный год	всего	в т.ч. за счет изменения	
						объема реализации	средней цены реализации
A	6 1 __	8 7 __	3 2 __	4 6 __			
B	8 1 __	12 __ 3	3 6 __	5 1 __			
C	3 3 __	5 __ 9	1 7 __	2 __ 8			
D	__ 61	2 __ 1	__ 00	1 __ 1			
Итого							

Задача 34

Определите влияние приращения объема продаж на изменение капиталоотдачи на основе следующих данных. Стоимость реализованной продукции в отчетном году - 27800 тыс. руб., в предыдущем году - 24300 тыс. руб., средняя хронологическая величина актива баланса в отчетном году - 22100 тыс. руб., в предыдущем году - 21700 тыс. руб.

Задача 35

Определите влияние интенсификации использования материальных ресурсов на изменение прибыли на основе следующих данных. Выручка от реализации продукции за отчетный год - 98,4 млн руб., за предыдущий год - 92,7 млн руб. Затраты предметов труда в отчетном году - 64,4 млн руб., в базисном году - 62,1 млн руб. Описать и объяснить ход решения.