

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ЛУГАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ ВЛАДИМИРА ДАЛЯ»

Экономический институт
Кафедра экономической кибернетики и прикладной статистики

УТВЕРЖДАЮ

Директор

« 28 »

(подпись)

20 25

года



ФОНД ОЦЕНОЧНЫХ СРЕДСТВ
по учебной дисциплине

Анализ данных средствами языка программирования R

(наименование учебной дисциплины, практики)

38.04.05 Бизнес-информатика

(код и наименование направления подготовки (специальности))

Бизнес-аналитика

(наименование профиля подготовки (специальности, магистерской программы); при отсутствии ставится прочерк)

Разработчик (разработчики):

доцент

Спорняк С.А.

старший преподаватель

Лофиченко А.А.

ФОС рассмотрен и одобрен на заседании кафедры экономической кибернетики
и прикладной статистики от « 15 » 02 2025 г., протокол № 25

Заведующий кафедрой экономической кибернетики и прикладной статистики

Велигура А.В.

Луганск 2025 г.

**Комплект оценочных материалов по дисциплине
«Анализ данных средствами языка программирования R»**

Задания закрытого типа

Задания закрытого типа на выбор правильного ответа

1. Выберите один правильный ответ:

При анализе эффективности двух социальных программ (А и В) по уровню доходов населения исследователь провел двухвыборочный t-тест. Каким образом может быть сформулирована нулевая гипотеза?

- А) средние доходы в группах А и В равны
- Б) средние доходы в группах А и В не равны
- В) доходы в группе А распределены нормально
- Г) дисперсии доходов в группах А и В равны

Правильный ответ: А

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Выберите один правильный ответ:

При анализе данных о клиентах банка (возраст, доход, кредитная история) необходимо предсказать, вернет ли клиент кредит. Какой алгоритм машинного обучения в R следует выбрать для решения этой задачи?

- А) дерево решений (rpart())
- Б) метод k-средних (kmeans())
- В) линейная регрессия (lm())
- Г) анализ главных компонент (prcomp())

Правильный ответ: А

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. Выберите один правильный ответ:

Какой тип графика в R (пакет ggplot2) лучше всего подойдет для визуализации распределения доходов населения по регионам с разделением на городское и сельское население?

- А) гистограмма (geom_histogram())
- Б) ящик с усами (geom_boxplot())
- В) точечная диаграмма (geom_point())
- Г) линейный график (geom_line())

Правильный ответ: Б

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. Выберите один правильный ответ:

Какой пакет в R используется для автоматизации процесса анализа данных и машинного обучения, включая визуализацию и предобработку данных?

- А) dplyr

Б) `ggplot2`

В) `Rattle`

Г) `caret`

Правильный ответ: В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. *Выберите один правильный ответ:*

Какая команда в R используется для импорта данных о доходах населения из CSV-файла для последующего анализа социально-экономических показателей?

А) `read.csv("data.csv")`

Б) `install.packages("readr")`

В) `library(dplyr)`

Г) `summary(data)`

Правильный ответ: А

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. *Выберите все правильные ответы:*

Какие функции из `ggplot2` используются для аннотирования графиков?

А) `geom_text()`

Б) `labs()`

В) `theme_minimal()`

Г) `annotate()`

Правильный ответ: А, Б, Г

Компетенции (индикаторы): ПК-2 (ПК-2.2)

7. *Выберите все правильные ответы:*

Какие элементы можно добавить в Shiny-приложение для взаимодействия с пользователем?

А) слайдер (`sliderInput()`)

Б) выпадающий список (`selectInput()`)

В) кнопка действия (`actionButton()`)

Г) гистограмма (`geom_histogram()`)

Правильный ответ: А, Б, В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

8. *Выберите все правильные ответы:*

Какие пакеты R предназначены для работы с нейронными сетями?

А) `nnet`

Б) `neuralnet`

В) `dplyr`

Г) `keras`

Правильный ответ: А, Б, Г

Компетенции (индикаторы): ПК-2 (ПК-2.2)

9. Выберите все правильные ответы:

Какие критерии используются для проверки нормальности распределения данных?

- Д) тест Шапиро-Уилка (shapiro.test())
- Е) тест Стьюдента (t.test())
- Ж) тест Колмогорова-Смирнова (ks.test())
- З) критерий хи-квадрат

Правильный ответ: А, В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

10. Выберите все правильные ответы:

Какие инструменты в R позволяют создавать дашборды?

- А) shiny
- Б) shinydashboard
- В) ggplot2
- Г) plotly

Правильный ответ: А, Б

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания закрытого типа на установление соответствия

1. Установите соответствие между типами графиков в R и их назначением. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|--------------------------------------|--|
| 1) гистограмма (geom_histogram()) | А) отображение структуры распределения |
| 2) ящик с усами (geom_boxplot()) | Б) визуализация распределения данных |
| 3) точечная диаграмма (geom_point()) | В) сравнение средних значений между группами |
| 4) линейный график (geom_line()) | Г) отображение изменения данных во времени |
| | Д) отображение взаимосвязи между двумя переменными |

Правильный ответ: 1–Б, 2–В, 3–Д, 4–Г

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Установите соответствие между типами статистических тестов и их применением. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|--------------------------|--|
| 1) одновыборочный t-тест | А) сравнение средних значений двух независимых групп |
|--------------------------|--|

- | | |
|--------------------------|--|
| 2) двухвыборочный t-тест | Б) проверка связи между категориальными переменными |
| 3) ANOVA | В) сравнение среднего значения выборки с известным значением |
| 4) Хи-квадрат тест | Г) сравнение средних значений трех и более групп |
| | Д) проверка наличия тенденции в выборке |

Правильный ответ: 1–В, 2–А, 3–Г, 4–Б

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. Установите соответствие между методами анализа данных и их применением. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|------------------------------------|---|
| 1) дерево решений | А) классификация данных с использованием ансамбля деревьев |
| 2) метод случайного леса | Б) классификация с использованием случайного разделения групп |
| 3) метод опорных векторов | В) классификация данных на основе множества условий |
| 4) кластеризация методом k-средних | Г) разделение данных на группы на основе схожести |
| | Д) классификация данных с использованием гиперплоскостей |

Правильный ответ: 1–В, 2–А, 3–Д, 4–Г

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. Установите соответствие между пакетами R и их назначением. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|------------|--|
| 1) shiny | А) создание интерактивных графиков |
| 2) plotly | Б) создание интерактивных таблиц |
| 3) leaflet | В) создание деревьев решений |
| 4) DT | Г) визуализация данных на интерактивных картах |
| | Д) создание интерактивных веб-приложений |

Правильный ответ: 1–Д, 2–А, 3–Г, 4–Б

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. Установите соответствие между типами регрессионных моделей и их использованием. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|----------------------------|---|
| 1) линейная регрессия | А) прогнозирование уровня бедности на основе ВВП, инфляции и инвестиций |
| 2) логистическая регрессия | Б) определение вероятности безработицы для отдельного домохозяйства |

- | | |
|----------------------------|---|
| 3) множественная регрессия | В) анализ влияния времени на рост населения |
| 4) нелинейная регрессия | Г) моделирование зависимости ВВП от квадрата инвестиций |
| | Д) разделение населения на городское и сельское |

Правильный ответ: 1–В, 2–Б, 3–А, 4–Г

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. Установите соответствие между пакетами R и их функциями. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|-----------|---|
| 1) Rattle | А) автоматизация предобработки данных и обучения моделей |
| 2) caret | Б) создание интерактивных дашбордов для визуализации данных |
| 3) dplyr | В) упрощённый графический интерфейс для анализа данных |
| 4) shiny | Г) построение прогнозных моделей |
| | Д) фильтрация, группировка и трансформация данных |

Правильный ответ: 1–В, 2–А, 3–Д, 4–Б

Компетенции (индикаторы): ПК-2 (ПК-2.2)

7. Установите соответствие между элементами пакета shiny и их назначением. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|------------------|---|
| 1) fluidPage() | А) создание интерактивного слайдера для выбора диапазона данных |
| 2) renderPlot() | Б) отображение динамически обновляемого графика |
| 3) sliderInput() | В) реактивное вычисление данных на основе ввода пользователя |
| 4) reactive() | Г) определение структуры пользовательского интерфейса |
| | Д) создание выпадающего списка |

Правильный ответ: 1–Г, 2–Б, 3–А, 4–В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

8. Установите соответствие между функциями R и их применением. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|--------------|--|
| 1) summary() | А) стандартизация данных перед кластеризацией |
| 2) table() | Б) расчёт корреляции между ВВП и уровнем образования |

- | | |
|------------|---|
| 3) cor() | В) сводная статистика по данным о доходах населения |
| 4) scale() | Г) создание таблицы распределения регионов по климатическим зонам |
| | Д) прогнозирование уровня бедности на основе ВВП, инфляции и инвестиций |

Правильный ответ: 1–В, 2–Г, 3–Б, 4–А

Компетенции (индикаторы): ПК-2 (ПК-2.2)

9. Установите соответствие между терминами и их определениями. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|---------------------------|--|
| 1) р-значение | А) вероятность отвергнуть ложную нулевую гипотезу |
| 2) доверительный интервал | Б) диапазон, в котором с заданной вероятностью находится истинный параметр |
| 3) мощность теста | В) условие для проведения ANOVA |
| 4) гомогенность дисперсий | Г) доля дисперсии, объясняемая моделью |
| | Д) вероятность получить наблюдаемые данные при верной нулевой гипотезе |

Правильный ответ: 1–Д, 2–Б, 3–А, 4–В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

10. Установите соответствие между методами обработки пропусков и их описанием. Каждому элементу левого столбца соответствует только один элемент правого столбца.

- | | |
|---|--|
| 1) удаление строк с NA | А) восстановление данных с учётом взаимосвязей между переменными |
| 2) замена NA средним значением | Б) подход, применяемый в методах вроде случайного леса |
| 3) использование алгоритмов, устойчивых к пропускам | В) риск потери информации при большом числе пропусков |
| 4) множественное импутирование | Г) сохранение общего объёма данных, но возможное искажение распределения |
| | Д) потеря некоторых значений, существенно отличных от среднего |

Правильный ответ: 1–В, 2–Г, 3–Б, 4–А

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания закрытого типа на установление правильной последовательности

1. Установите правильную последовательность шагов для построения линейной регрессии в R. Запишите правильную последовательность букв слева направо:

- А) проверка значимости модели с помощью функции summary()
 - Б) загрузка данных с помощью функции read.csv()
 - В) визуализация результатов с помощью функции plot()
 - Г) проверка данных на мультиколлинеарность с помощью функции cor()
 - Д) построение модели линейной регрессии с помощью функции lm()
- Правильный ответ: Б, Г, Д, А, В
Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Установите правильную последовательность шагов для создания дашборда в shiny. Запишите правильную последовательность букв слева направо:

- А) написание серверной логики: server <- function(input, output) {...}
 - Б) запуск приложения: runApp()
 - В) создание UI: fluidPage(...)
 - Г) объединение UI и сервера: shinyApp(ui, server)
- Правильный ответ: В, А, Г, Б
Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. Установите правильную последовательность шагов для проведения ANOVA. Запишите правильную последовательность букв слева направо:

- А) проверка гомогенности дисперсий: bartlett.test()
 - Б) пост-хок анализ: TukeyHSD()
 - В) построение модели: aov(Result ~ Group, data)
 - Г) проверка нормальности распределения: shapiro.test()
- Правильный ответ: Г, А, В, Б
Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. Установите правильную последовательность шагов для создания диаграммы рассеяния в ggplot2. Запишите правильную последовательность букв слева направо:

- А) загрузка данных: data <- read.csv("data.csv")
 - Б) добавление слоя точек: + geom_point()
 - В) добавление заголовка: + ggtitle("Зависимость расходов от дохода")
 - Г) создание графика: ggplot(data, aes(x = Income, y = Spending))
- Правильный ответ: А, Г, Б, В
Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. Установите порядок действий для анализа распределения доходов населения. Запишите правильную последовательность букв слева направо:

- А) построение гистограммы с geom_histogram()

- Б) проверка данных на пропуски с `is.na()`
- В) очистка данных от пропусков с `na.omit()`
- Г) визуализация результатов с `ggplot()`

Правильный ответ: Б, В, Г, А

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. Установите порядок построения модели множественной регрессии для прогнозирования ВВП. Запишите правильную последовательность букв слева направо:

- А) проверка корреляции между переменными с `cor()`
- Б) построение модели с `lm(GDP ~ Investment + Education)`
- В) анализ значимости коэффициентов через `summary()`
- Г) визуализация остатков с `plot(model)`

Правильный ответ: А, Б, В, Г

Компетенции (индикаторы): ПК-2 (ПК-2.2)

7. Расположите в правильной последовательности этапы работы с *Rattle* для анализа данных о бедности. Запишите правильную последовательность букв слева направо:

- А) запуск кластеризации методом k-средних
- Б) загрузка данных в интерфейс *Rattle*
- В) интерпретация результатов кластеризации
- Г) выбор целевой переменной (уровень бедности)

Правильный ответ: Б, Г, А, В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

8. Установите последовательность построения дерева решений для прогноза банкротства компаний. Запишите правильную последовательность букв слева направо:

- А) визуализация дерева через `rpart.plot()`
- Б) построение модели с `rpart(Bankruptcy ~ .)`
- В) оценка точности на тестовых данных
- Г) разделение данных на обучающую и тестовую выборки

Правильный ответ: Г, Б, А, В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

9. Расположите в правильной последовательности этапы классификации регионов по уровню развития с помощью k-средних. Запишите правильную последовательность букв слева направо:

- А) построение модели `kmeans(data, centers=3)`
- Б) нормализация данных с `scale()`
- В) визуализация кластеров с `fviz_cluster()`
- Г) выбор оптимального числа кластеров через `fviz_nbclust()`

Правильный ответ: Б, Г, А, В

Компетенции (индикаторы): ПК-2 (ПК-2.2)

10. Установите последовательность обработки пропущенных значений в данных о доходах. Запишите правильную последовательность букв слева направо:

- А) выявление пропусков через `sum(is.na(data))`
- Б) замена пропусков медианным значением с `mutate()`
- В) визуализация пропусков через `vis_miss()`
- Г) сохранение очищенных данных в новый файл

Правильный ответ: А, В, Б, Г

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания открытого типа

Задания открытого типа на дополнение

1. Напишите пропущенное слово:

Для загрузки данных из CSV-файла в R используется функция _____("data.csv").

Правильный ответ: `read.csv`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. Напишите пропущенное слово:

Чтобы построить ящик с усами для сравнения уровня безработицы в регионах, в `ggplot2` используется функция _____().

Правильный ответ: `geom_boxplot`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. Напишите пропущенное слово:

Для проверки нормальности распределения данных о доходах применяется тест _____().

Правильный ответ: `shapiro.test`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. Напишите пропущенное слово:

Функция для построения линейной регрессии, прогнозирующей ВВП на основе инвестиций, – _____(`GDP ~ Investments, data`).

Ответ: `lm`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. Напишите пропущенное слово:

Пакет _____ предоставляет графический интерфейс для автоматизации анализа данных в R.

Правильный ответ: `Rattle`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. *Напишите пропущенное слово:*

Для разделения регионов на группы по уровню экономического развития используется метод _____().

Правильный ответ: kmeans

Компетенции (индикаторы): ПК-2 (ПК-2.2)

7. *Напишите пропущенное слово:*

Пакет _____ позволяет создать веб-приложение для визуализации динамики ВВП.

Правильный ответ: shiny

Компетенции (индикаторы): ПК-2 (ПК-2.2)

8. *Напишите пропущенное слово:*

Функция _____() в dplyr используется для фильтрации строк с уровнем бедности выше 10%.

Правильный ответ: filter

Компетенции (индикаторы): ПК-2 (ПК-2.2)

9. *Напишите пропущенное слово:*

Для сравнения средних доходов в трёх секторах экономики используется функция _____().

Правильный ответ: aov

Компетенции (индикаторы): ПК-2 (ПК-2.2)

10. *Напишите пропущенное слово:*

Чтобы вычислить корреляцию между ВВП и уровнем образования, применяется функция _____().

Правильный ответ: cor

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания открытого типа с кратким свободным ответом

1. *Дайте ответ на вопрос:*

Какая функция в R используется для чтения данных из файла в формате JSON, содержащего показатели экономического развития регионов?

Правильный ответ: fromJSON()

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. *Дайте ответ на вопрос:*

Какой тип графика в ggplot2 лучше всего подходит для отображения распределения населения по возрастным группам в разных странах?

Правильный ответ: гистограмма / geom_histogram()

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. *Дайте ответ на вопрос:*

Какой тест используется для сравнения медианных значений доходов в двух независимых группах, если данные не распределены нормально?

Правильный ответ: тест Манна-Уитни

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. *Дайте ответ на вопрос:*

Какая функция в R применяется для построения логистической регрессии, прогнозирующей вероятность экономического кризиса?

Правильный ответ: `glm()`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. *Дайте ответ на вопрос:*

Какой алгоритм используется для выявления аномалий в данных о финансовых транзакциях?

Правильный ответ: изолированный лес / Isolation Forest

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. *Дайте ответ на вопрос:*

Какая функция в `dplyr` позволяет добавлять новые переменные на основе существующих столбцов (например, рассчитывать ВВП на душу населения)?

Правильный ответ: `mutate()`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

7. *Дайте ответ на вопрос:*

Какой пакет R используется для создания интерактивных карт, отображающих уровень безработицы по регионам?

Правильный ответ: `leaflet`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

8. *Дайте ответ на вопрос:*

Какая функция в R применяется для прогнозирования будущих значений ВВП на основе исторических данных?

Правильный ответ: `arima()`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

9. *Дайте ответ на вопрос:*

Какой метод используется для визуализации нелинейных зависимостей между экономическими показателями?

Правильный ответ: график рассеяния с линией тренда / `geom_smooth()`

Компетенции (индикаторы): ПК-2 (ПК-2.2)

10. *Дайте ответ на вопрос:*

Какой пакет R позволяет анализировать текстовые данные (например, отзывы о государственных программах)?

Правильный ответ: tm / Text Mining Package

Компетенции (индикаторы): ПК-2 (ПК-2.2)

Задания открытого типа с развернутым ответом

1. *Прочитайте текст задания. Продумайте логику и полноту ответа. Запишите развернутый и обоснованный ответ.*

Опишите процесс загрузки данных о продажах компании из CSV-файла в R. Как вы проверите структуру загруженных данных и какие шаги предпримете для подготовки данных к анализу? Приведите пример кода.

Время выполнения – 15 минут

Ожидаемый результат:

Для загрузки данных из CSV-файла в R используется функция read.csv().

```
sales_data <- read.csv("sales.csv")
```

Чтобы проверить структуру данных, применяется функция str(), которая показывает типы переменных и первые значения:

```
str(sales_data)
```

Для подготовки данных к анализу необходимо:

Проверить наличие пропущенных значений с помощью summary(sales_data) или is.na(sales_data). Пропуски можно удалить (na.omit(sales_data)) или заполнить средними значениями.

Преобразовать категориальные переменные в факторы, если требуется, например:

```
sales_data$category <- as.factor(sales_data$category)
```

Критерии оценивания:

правильное использование функции для загрузки данных;

использование функции для проверки структуры данных;

упоминание проверки пропущенных значений;

упоминание преобразования типов данных.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

2. *Прочитайте текст задания. Продумайте логику и полноту ответа. Запишите развернутый и обоснованный ответ.*

Напишите код для построения графика с помощью пакета ggplot2, отображающий динамику ВВП за последние 10 лет для пяти стран (имеется набор данных data, в котором содержатся 3 переменных: Year – год, GDP – значение ВВП, Country – страна). Объясните, почему вы выбрали этот тип визуализации.

Время выполнения – 20 минут

Ожидаемый результат:

```
ggplot(data, aes(x=Year, y=GDP, color=Country)) +
```

```
geom_line() +  
labs(title="Динамика ВВП по странам")
```

Критерии оценивания:

наличие кода построение графика;

наличие обоснования выбора типа графика: линейный график показывает тренды и сравнение между странами.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

3. *Прочитайте текст задания. Продумайте логику и полноту ответа. Запишите развернутый и обоснованный ответ.*

На рисунке представлен результат проведения t-теста для проверки гипотезы о разнице между средними значениями выпуска деталей двумя машинами. Какие выводы можно сделать о разнице между средними значениями выпуска деталей двумя машинами?

```
data: machine1 and machine2  
t = 2.4396, df = 18, p-value = 0.02528  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 0.4164695 5.5835305  
sample estimates:  
mean of x mean of y  
      23      20
```

Время выполнения – 15 минут

Ожидаемый результат:

Для проверки гипотезы использовался двусторонний t-тест

Значение t и низкое значение p указывают на то, что можно отвергнуть нулевую гипотезу. Машина 2 значительно быстрее, чем Машина 1

Среднее значение выпуска деталей Машиной 1 – 23, Машиной 2 – 20

Критерии оценивания:

указан тип используемого теста;

указано, что можно отвергнуть нулевую гипотезу и обосновано почему;

указаны средние значения выпуска деталей для обеих машин.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

4. *Прочитайте текст задания. Продумайте логику и полноту ответа. Запишите развернутый и обоснованный ответ.*

Для набора данных employees (содержит 2 переменные: employee – сотрудник и income – доход сотрудника) постройте гистограмму распределения доходов сотрудников компании и опишите, что она показывает. Как вы интерпретируете скошенность распределения? Приведите пример кода.

Время выполнения – 20 минут

Ожидаемый результат:

Для построения гистограммы используется функция hist():

```
hist(employees$income, main="Распределение доходов", xlab="Доход",  
ylab="Частота")
```

Гистограмма показывает частоту встречаемости различных уровней доходов. Если она скошена вправо (длинный хвост справа), большинство сотрудников имеют доход ниже среднего. Если скошена влево – наоборот, большинство доходов выше среднего.

Критерии оценивания:

правильное использование функции для построения гистограммы;

описание того, что показывает гистограмма;

интерпретация скошенности.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

5. *Прочитайте текст задания. Продумайте логику и полноту ответа. Запишите развернутый и обоснованный ответ.*

Вы провели опрос удовлетворенности клиентов до и после внедрения новой системы обслуживания. Как вы проверите, есть ли статистически значимая разница в уровне удовлетворенности? Опишите шаги и приведите пример кода для t-теста в R.

Время выполнения – 20 минут

Ожидаемый результат:

Для проверки разницы используется парный t-тест (paired t-test), если измерения связаны (те же клиенты до и после). Шаги:

Нулевая гипотеза (H_0): средние уровни удовлетворенности равны.

Альтернативная гипотеза (H_1): средние различаются.

Выполнить t-тест:

`t.test(before, after, paired=TRUE)`

Если p-значение < 0.05 , H_0 отвергается – разница значима.

Критерии оценивания:

выбор правильного теста (парный t-тест);

формулировка гипотез;

пример кода.

Компетенции (индикаторы): ПК-2 (ПК-2.2)

6. *Прочитайте текст задания. Продумайте логику и полноту ответа. Запишите развернутый и обоснованный ответ.*

Объясните, как вы бы провели одновыборочную проверку гипотезы о среднем значении в R на примере социально-экономического показателя – среднего дохода населения. Укажите функции и интерпретацию результатов.

Время выполнения – 20 минут

Ожидаемый результат:

Для одновыборочной проверки гипотезы о среднем значении используется функция `t.test()`. Предположим, у нас есть вектор `incomes` с данными о доходах, и мы проверяем, равно ли среднее значение 50000. Шаги:

Сформулировать гипотезы:

H_0 : средний доход = 50000.

H_1 : средний доход $\neq$ 50000.

Выполнить тест:

```
result <- t.test(incomes, mu = 50000)
```

Интерпретировать результаты:

Если p -значение (из вывода `result$p.value`) меньше 0.05, H_0 отвергается, то есть средний доход статистически значимо отличается от 50000.

Если p -значение ≥ 0.05 , нет оснований отвергать H_0 .

Эта проверка помогает понять, соответствует ли наблюдаемый доход ожидаемому уровню, что важно для социально-экономического планирования.

Критерии оценивания:

упоминание функции `t.test()`;

описание формулировки гипотез;

объяснение интерпретации p -значения.

Компетенции (индикаторы): ПК-6 (ПК-6.3)

Экспертное заключение

Представленный фонд оценочных средств (далее – ФОС) по дисциплине «Анализ данных средствами языка программирования R» соответствует требованиям ФГОС ВО.

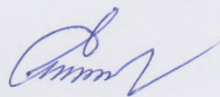
Предлагаемые формы и средства текущего и промежуточного контроля адекватны целям и задачам реализации основной профессиональной образовательной программы по направлению подготовки 38.04.05 Бизнес-информатика.

Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины представлены в полном объеме.

Виды оценочных средств, включенные в представленный фонд, отвечают основным принципам формирования ФОС.

Разработанный и представленный для экспертизы фонд оценочных средств рекомендуется к использованию в процессе подготовки обучающихся по указанному направлению.

Председатель учебно-методической
комиссии экономического института



Шаповалова Е.Н.

Лист изменений и дополнений

№ п/п	Виды дополнений и изменений	Дата и номер протокола заседания кафедры (кафедр), на котором были рассмотрены и одобрены изменения и дополнения	Подпись (с расшифровкой) заведующего кафедрой (заведующих кафедрами)