

Темы практических работ

по курсу «Программные системы статистического анализа»

1. Исследование оценок метода наименьших квадратов с использованием функции `lm{stats}` и матричных вычислений.
2. Исследование оценок итерационного взвешенного метода наименьших квадратов с использованием функции `glm{MASS}` и матричных вычислений.
3. Сравнение функций, реализующих факторный анализ, предоставляемых пакетами `stats`, `psych`.
4. Сравнение реализаций дисперсионного анализа (ANOVA) с помощью функций `lm{stats}` и `aov{stats}`.
5. Исследование результатов кластерного анализа с использованием базовых функций (например, `kmeans{stats}`) и специализированных пакетов.
6. Сравнение способов визуализации результатов кластерного анализа с помощью базовой графики и пакета `ggplot`.
7. Сравнение способов визуализации результатов дисперсионного анализа с помощью базовой графики и пакета `ggplot`.
8. Сравнение способов визуализации результатов дисперсионного анализа с помощью базовой графики и пакета `ggplot`.
9. Сравнение способов визуализации результатов корреляционного анализа с помощью базовой графики и пакета `ggplot`.
10. Сравнение способов визуализации результатов метода главных компонент (PCA) с помощью `ggplot` и пакета `FactoMineR`.
11. Исследование способов извлечения текста из HTML-документа с помощью пакетов `XML` и `xml2`.
12. Исследование способов извлечения информации из текстового документа с помощью пакета `stringr` и базовых функций (например, `strsplit`).
13. Сравнение способов первичной обработки данных с помощью циклов и векторизованных функций.
14. Сравнение способов генерации случайных данных с помощью функции `replicate` и с помощью циклов.
15. Сравнение способов построения таблиц сопряженности с помощью функции `table{base}`, циклов и условных операторов.
16. Сравнение реализаций метода главных компонент (PCA) с помощью функций `prcomp{stats}`, `princomp{stats}`, `PCA{FactoMineR}`.
17. Сравнение методов классификации с учителем – деревья решений – с помощью пакетов `tree`, `rpart`.
18. Сравнение методов классификации с учителем – логистическая регрессия – с помощью функции `glm{MASS}`, `glmnet{glmnet}`.
19. Сравнение методов классификации с учителем – метод опорных векторов – с помощью функции `svm{e1071}` и путем решения оптимизационной задачи.
20. Сравнение методов классификации с учителем – метод ближайших соседей – с помощью функции `knn{class}` и путем реализации алгоритма.

21. Сравнение методов регуляризации для обобщенных линейных моделей с помощью пакета `glmnet` и путем решения оптимизационной задачи.
22. Исследование методов локального регрессионного сглаживания с помощью функций `loess{stats}`, `lm{stats}` (взвешенный метод наименьших квадратов).
23. Сравнение способов визуализации результатов локального регрессионного сглаживания с помощью базовой графики и пакета `ggplot`.
24. Сравнение способов графического представления одномерного распределения с помощью базовой графики и пакета `ggplot`.
25. Сравнение способов графического представления двумерного распределения с помощью базовой графики и пакета `ggplot`.
26. Сравнение способов визуализации прогнозов временных рядов с помощью базовой графики и специализированных пакетов.
27. Сравнение методов ресэмплинга, реализованных с помощью пакета `boot` и с помощью циклов и векторизованных функций.
28. Исследование реализаций процедуры перекрестной проверки на базе пакета `caret` и с помощью базовых функций.
29. Исследование оценок полихорического коэффициента корреляции с использованием функций `polychoric{psych}`, `polychor{polycor}`.
30. Исследование оценок полисерийного коэффициента корреляции с использованием функций `polyserial{psych}`, `polyserial{polycor}`.
31. Исследование методов прогнозирования временных рядов (модель Хольта-Уинтерса) на основе пакетов `stats`, `forecast`.
32. Исследование методов прогнозирования временных рядов (скользящее среднее) с помощью функций `SMA{TTR}`, `ma{forecast}`.
33. Исследование методов прогнозирования временных рядов (модель ARIMA) на основе пакетов `stats`, `forecast`.
34. Сравнение методов оценки автокорреляционной функции для временного ряда с помощью функций `acf{stats}`, `Acf{forecast}`.
35. Сравнение методов оценки частной автокорреляционной функции для временного ряда с помощью функций `pacf{tseries}`, `Pacf{forecast}`.
36. Сравнение способов визуализации (частной) автокорреляционной функции для временного ряда с помощью базовой графики и специализированных пакетов.
37. Сравнение способов построения облака слов с помощью функций `wordcloud{wordcloud}`, `wordcloud2{wordcloud2}`.
38. Сравнение методов иерархического кластерного анализа с помощью функций `hclust{stats}`, `agnes{cluster}`.
39. Сравнение способов построения дендрограмм с помощью стандартной графики и пакета `dendextend`.
40. Исследование методов построения матриц попарных расстояний на данных большой размерности с помощью пакетов `parallelDist`, `largeVis`.
41. Исследование методов линейного дискриминантного анализа с помощью функции `lda{MASS}` и путем реализации алгоритма.
42. Исследование методов построения самоорганизующихся карт Кохонена с помощью пакета `kohonen` и путем реализации алгоритма.
43. Сравнение способов визуализации результатов дискриминантного анализа с помощью базовой графики и специализированных пакетов.

44. Сравнение методов классификации – градиентный бустинг – с помощью Python.
45. Исследование оценок метода взвешенных наименьших квадратов с использованием функции `lm{stats}` и матричных вычислений.
46. Сравнение методов классификации с учителем – случайный лес – с помощью пакетов `randomForest` и `randomForestSRC`.
47. Исследование способов отправки HTTP запросов с помощью пакетов `RCurl` и `rvest`.
48. Исследование способов работы с кодировками с помощью функций `guess_encoding` и `repair_encoding` R-пакета `rvest`.
49. Исследование методов ресэмплинга для оценки коэффициента корреляции с помощью пакета `boot` и путем реализации алгоритма.
50. Сравнение алгоритмов оптимизации модулярности графов, реализованных в пакете `igraph`.
51. Сравнение алгоритмов сортировки с помощью функции `sort`.
52. Сравнение метрик, реализованных для вычисления расстояния между строками в пакете `stringdist`.
53. Исследование методов предобработки данных в пакете `caret`.
54. Сравнение методов оценки важности регрессоров в линейных моделях с помощью функции `varImp()` пакета `caret` и пакета `relaimpro`.
55. Сравнение методов описательной статистики, реализованных в пакетах `Hmisc` и `psych` с помощью функции `describe`.
56. Исследование методов визуализации коэффициентов модели с доверительными интервалами с помощью пакета `coefplot`.
57. Исследование способов извлечения ключевых фраз из текста с помощью пакета `RKEA`.
58. Исследование метода построения квантильной регрессии с помощью пакета `quantreg`.
59. Исследование методов оценивания моделей со структурными уравнениями с помощью пакета `lavaan`.
60. Исследование методов оценивания моделей со смешанными эффектами с помощью пакета `mle4`.
61. Исследование алгоритмов нечеткого сопоставления строк с помощью пакета `fuzzywuzzyR`.
62. Исследование оценок межклассовой корреляции с использованием функции `ICC{psych}`.