Часть 1. 50 задач, связанных с созданием, группировкой, сортировкой, удалением, добавлением и редактированием строк и столбцов в фрейме данных:

1. Создайте пустой фрейм данных.

2. Загрузите данные из CSV-файла в фрейм данных.

3. Создайте новый столбец на основе существующих столбцов.

4. Удалите определенный столбец из фрейма данных.

5. Удалите строки, содержащие пустые значения.

6. Отсортируйте фрейм данных в порядке возрастания определенного столбца.

7. Отсортируйте фрейм данных в порядке убывания определенного столбца.

8. Сгруппируйте строки по значениям определенного столбца.

9. Посчитайте количество уникальных значений в определенном столбце.

10. Измените значение определенной ячейки в фрейме данных.

11. Измените название определенного столбца.

12. Измените порядок столбцов в фрейме данных.

13. Добавьте новую строку в фрейм данных.

14. Добавьте новый столбец в фрейм данных.

15. Объедините два фрейма данных по общему столбцу.

16. Извлеките определенные строки на основе условия.

17. Извлеките определенные столбцы на основе условия.

18. Примените функцию к каждому значению определенного столбца.

19. Замените пустые значения в определенном столбце на другое значение.

20. Удалите дубликаты строк в фрейме данных.

21. Измените тип данных определенного столбца.

22. Измените формат даты в определенном столбце.

23. Переименуйте индексы строк в фрейме данных.

24. Измените индексы столбцов в фрейме данных.

25. Примените агрегирующую функцию к каждому столбцу.

26. Примените условную функцию к каждому значению фрейма данных.

27. Измените размерность фрейма данных путем добавления или удаления строк и столбцов.

28. Создайте сводную таблицу на основе фрейма данных.

29. Замените все значения в определенном столбце на другое значение.

30. Отфильтруйте строки на основе условия, примененного к нескольким столбцам одновременно.

31. Переведите все значения в определенном столбце в верхний регистр.

32. Переведите все значения в определенном столбце в нижний регистр.

33. Выполните операцию сложения для двух столбцов и сохраните результат в новый столбец.

34. Выполните операцию вычитания для двух столбцов и сохраните результат в новый столбец.

35. Выполните операцию умножения для двух столбцов и сохраните результат в новый столбец.

36. Выполните операцию деления для двух столбцов и сохраните результат в новый столбец.

37. Выполните операцию возведения в степень для определенного столбца и сохраните результат в новый столбец.

38. Выполните операцию модуля для определенного столбца и сохраните результат в новый столбец.

39. Примените функцию к каждому значению двух столбцов и сохраните результат в новый столбец.

40. Переведите значения в определенном столбце в бинарный формат (0 и 1).

41. Посчитайте сумму значений в определенном столбце.

42. Посчитайте среднее значение значений в определенном столбце.

43. Посчитайте медианное значение значений в определенном столбце.

44. Посчитайте минимальное значение значений в определенном столбце.

45. Посчитайте максимальное значение значений в определенном столбце.

46. Посчитайте суммарное значение значений в каждом столбце фрейма данных.

47. Посчитайте количество значений, удовлетворяющих условию, в определенном столбце.

48. Посчитайте произведение значений в определенном столбце.

49. Посчитайте разность между максимальным и минимальным значениями в определенном столбце.

50. Посчитайте среднеквадратичное отклонение значений в определенном столбце.

Часть 2. Запросы

1. Отобразите все строки фрейма данных.
2. Отобразите первые 10 строк фрейма данных.
3. Отобразите последние 10 строк фрейма данных.
4. Отобразите значение столбца "Имя" для всех строк.
5. Отобразите значения столбца "Возраст" для всех строк.
6. Отобразите строки, в которых значение столбца "Возраст" больше 30.
7. Отобразите строки, в которых значение столбца "Имя" равно "Анна".
8. Отобразите строки, в которых значение столбца "Имя" содержит подстроку "ан".
9. Отобразите строки, в которых значение столбца "Возраст" находится в диапазоне от 25 до 35.
10. Отобразите строки, в которых значение столбца "Имя" начинается с буквы "А" и значение столбца "Возраст" больше 30.
11. Отобразите строки, в которых значение столбца "Имя" не равно "Дмитрий".
12. Отобразите строки, в которых значение столбца "Имя" равно "Анна" или "Катя".
13. Отобразите строки, в которых значение столбца "Имя" равно "Анна" и значение столбца "Возраст" больше 25.
14. Отобразите строки, в которых значение столбца "Имя" содержит подстроку "ан" или "ат".
15. Отобразите строки, в которых значение столбца "Возраст" является пропущенным (NaN).
16. Отобразите строки, в которых значение столбца "Имя" не является пропущенным (NaN).
17. Отобразите строки, в которых значение столбца "Имя" начинается с буквы "А" и значение столбца "Возраст" является пропущенным (NaN).
18. Отобразите строки, в которых значение столбца "Имя" является пропущенным (NaN) или значение столбца "Возраст" больше 30.
19. Отобразите строки, в которых значение столбца "Возраст" уникально.
20. Отобразите строки, в которых значение столбца "Имя" уникально.
21. Посчитайте количество строк во фрейме данных.
22. Посчитайте количество уникальных значений в столбце "Имя".
23. Посчитайте сумму значений в столбце "Возраст".
24. Посчитайте среднее значение значений в столбце "Возраст".
25. Посчитайте минимальное значение в столбце "Возраст".
26. Посчитайте максимальное значение в столбце "Возраст".
27. Посчитайте медиану значений в столбце "Возраст".
28. Посчитайте моду значений в столбце "Возраст".
29. Посчитайте сумму значений в каждом столбце фрейма данных.
30. Посчитайте среднее значение значений в каждом столбце фрейма данных.
31. Отобразите строки, в которых значение столбца "Имя" начинается с буквы "А" и значение столбца "Возраст" находится в диапазоне от 25 до 35, отсортированные по возрастанию значения столбца "Возраст".
32. Отобразите количество строк, в которых значение столбца "Возраст" является пропущенным (NaN).
33. Отобразите значения столбца "Имя" в верхнем регистре.

Часть 3.

Конечно! Вот 55 задач, связанных с обработкой и очисткой данных для датасайнс в фрейме данных:

1. Удалите дубликаты строк в фрейме данных.

2. Удалите строки, содержащие пустые значения.

3. Замените пустые значения в определенном столбце на среднее значение.

4. Замените пустые значения в определенном столбце на медианное значение.

5. Замените пустые значения в определенном столбце на модальное значение.

6. Измените тип данных определенного столбца.

7. Измените формат даты в определенном столбце.

8. Измените регистр символов в определенном столбце.

9. Удалите лишние пробелы в определенном столбце.

10. Удалите знаки препинания в определенном столбце.

11. Удалите числа из определенного столбца.

12. Извлеките числа из определенного столбца.

13. Извлеките текст из определенного столбца.

14. Извлеките дату из определенного столбца.

15. Извлеките время из определенного столбца.

16. Извлеките URL-адреса из определенного столбца.

17. Извлеките электронные адреса из определенного столбца.

18. Измените названия столбцов на более понятные.

19. Объедините два фрейма данных по общему столбцу.

20. Разделите один столбец на несколько на основе разделителя.

21. Объедините несколько столбцов в один.

22. Найдите выбросы в определенном столбце и удалите их.

23. Найдите выбросы в определенном столбце и замените их на среднее значение.

24. Найдите выбросы в определенном столбце и замените их на медианное значение.

25. Найдите выбросы в определенном столбце и замените их на модальное значение.

26. Удалите выбросы, находящиеся за пределами интерквартильного размаха.

27. Замените выбросы, находящиеся за пределами интерквартильного размаха, на среднее значение.

28. Замените выбросы, находящиеся за пределами интерквартильного размаха, на медианное значение.

29. Замените выбросы, находящиеся за пределами интерквартильного размаха, на модальное значение.

30. Нормализуйте значения в определенном столбце.

31. Стандартизуйте значения в определенном столбце.

32. Измените масштаб значений в определенном столбце на заданный диапазон.

33. Посчитайте корреляцию между двумя столбцами.

34. Посчитайте коэффициент детерминации между двумя столбцами.

35. Постройте график распределения значений в определенном столбце.

36. Постройте гистограмму распределения значений в определенном столбце.

37. Постройте ящик с усами для определенного столбца.

38. Постройте точечную диаграмму для двух столбцов.

39. Постройте линейный график для двух столбцов.

40. Постройте график зависимости между двумя категориальными переменными.

41. Постройте график распределения значений для каждого уникального значения категориальной переменной.

42. Сгруппируйте строки по значениям определенного столбца и посчитайте среднее значение другого столбца для каждой группы.

43. Сгруппируйте строки по значениям определенного столбца и посчитайте медианное значение другого столбца для каждой группы.

44. Сгруппируйте строки по значениям определенного столбца и посчитайте модальное значение другого столбца для каждой группы.

45. Сгруппируйте строки по значениям нескольких столбцов и посчитайте среднее значение другого столбца для каждой группы.

46. Сгруппируйте строки по значениям нескольких столбцов и посчитайте медианное значение другого столбца для каждой группы.

47. Сгруппируйте строки по значениям нескольких столбцов и посчитайте модальное значение другого столбца для каждой группы.

48. Посчитайте общее количество строк, содержащих уникальные значения в нескольких столбцах.

49. Посчитайте общее количество строк, содержащих дубликаты значений в нескольких столбцах.

50. Посчитайте общее количество строк, содержащих пустые значения в нескольких столбцах.

51. Разделите фрейм данных на две части: обучающую и тестовую выборки в заданной пропорции.

52. Разделите фрейм данных на кластеры на основе значений нескольких столбцов.

53. Примените алгоритм машинного обучения к фрейму данных и оцените его точность.

54. Определите оптимальное количество кластеров для фрейма данных с помощью метода локтя или метода силуэта.

55. Определите важность каждого признака в фрейме данных с помощью алгоритма случайного леса или метода главных компонент (PCA).