

Инструкция для финального проекта

А.В. Лысов

11 января 2017 г.

1 Описание программы

Данная программа предназначена для анализа набора данных, заданного заранее в расширении «.xls».

В данной программе в качестве выбора переменных для анализов и тестов могут использоваться любые из числа возможных.

По умолчанию набор данных — «data/DataSetLA.xls», который содержится на 2 каталога выше от исполняемого файла.

Презентация, отчет, а также «.xls»-файл с регрессионным и корреляционным анализами при их наличии хранятся в папке «tmp», расположенной в каталоге исполняемого файла. Созданные изображения кластеризации хранятся там же.

2 Описание интерфейса

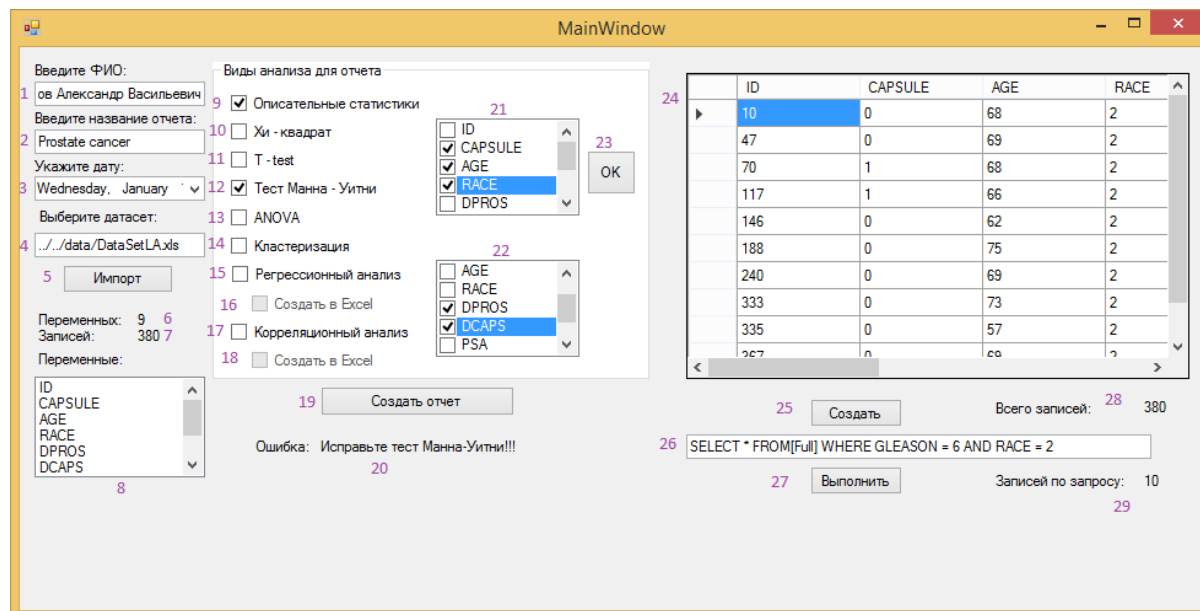


Рис. 1: Интерфейс программы

Легенда:

1. Поле для ввода ФИО автора отчета;
2. Поле для ввода названия отчета;
3. Поле для указания даты создания отчета;
4. Поле для выбора расположения набора данных (по умолчанию стоит на data/DataSetLA.xls, расположенном на 2 католога выше исполняемого файла);
5. Кнопка, которая импортирует данные из указанного в 4 набора данных;
6. Показывает, сколько переменных было в наборе данных;
7. Показывает, сколько записей было в наборе данных;
8. Выводит список переменных;
9. Чекбокс для внесения в отчет описательных статистик;
10. Чекбокс для внесения в отчет теста Хи-квадрат;

11. Чекбокс для внесения в отчет Т-теста;
12. Чекбокс для внесения в отчет теста Манна-Уитни;
13. Чекбокс для внесения в отчет теста ANOVA;
14. Чекбокс для внесения в отчет кластеризации;
15. Чекбокс для внесения в отчет регрессионного анализа;
16. Чекбокс для создания в Excel регрессионного анализа;
17. Чекбокс для внесения в отчет корреляционного анализа;
18. Чекбокс для создания в Excel корреляционного анализа;
19. Кнопка для создания отчета;
20. Поле для возможных ошибок;
21. Первое поле для переменных:
 - В описательных статистиках, кластеризации и корреляционном анализе используется по очевидному назначению;
 - В тестах используется как левая часть выражения;
 - В регрессионном анализе используется как переменная, для которого нужно найти функцию по переменным из второго поля для переменных.
22. Второе поле для переменных:
 - В описательных статистиках, кластеризации и корреляционном анализе отсутствует;
 - В тестах используется как правая часть выражения;
 - В регрессионном анализе используется как набор переменных функции, при помощи которых нужно определять значение переменной из первого поля переменных.
23. Кнопка, которая для конкретного теста/анализа выбирает переменные;
24. Поле выборки по запросу из базы данных;
25. Кнопка, которая создает базу данных на основе импортированных данных;
26. Запрос выборки из базы данных;

- 27. Кнопка, которая выполняет запрос (26);
- 28. Количество записей в исходной базе данных;
- 29. Количество записей по выборке из запроса (26).

3 Возможности программы

3.1 Создание отчета по анализу данных

По задумке запуск программы работает на одном наборе данных, то есть чтобы обработать другой набор данных, рекомендуется перезапустить программу.

Чтобы создать отчет, нужно:

1. (*) Вписать ФИО автора;
2. (*) Вписать название отчета;
3. (*) Указать дату;
4. Выбрать набор данных;
5. Импортировать данные, нажав кнопку «Импорт»;
6. Убедиться, что количество переменных и записей соответствует действительности;
7. В видах анализов выбрать нужные, в каждом выбирая необходимый набор переменных и нажимая кнопку «ОК»;
8. Нажать на кнопку «Создать отчет»;
9. Если все прошло без проблем, в поле «Ошибка» будет написано "Отчет создан!";
10. Если в поле «Ошибка» что-то отличное от "Отчет создан!", рекомендуется следовать совету этого поля, посмотрев, что может быть не так, и исправив это, предварительно закрыв Word-документ;
11. GO TO 9.

Пункты, помеченные (*) необязательные.

3.2 Выборка элементов базы данных

1. Выбрать набор данных;
2. Импортировать данные, нажав кнопку «Импорт»;
3. Убедиться, что количество записей в базе данных соответствует количеству записей в наборе данных;

4. Создать базу данных, нажав на кнопку «Создать»;
5. Написать запрос выборки элементов. Изменять что-либо от запроса по умолчанию рекомендуется только в местах:
 - (a) Между SELECT и FROM;
 - (b) После WHERE.
6. Выполнить запрос, нажав на кнопку «Выполнить»;
7. В случае, если элементы, удовлетворяющие запросу, существуют, в поле «Записей по запросу» будет выведено их количество, а в поле базы данных будут выведены сами записи. В противном случае будет выведено -1.

4 Возможные ошибки

- «Необходимо импортировать данные!!!» — возникает при попытке загрузить базу данных или создать отчет без загрузки набора данных;
- «Необходимо выбрать переменные!!!» — возникает, когда не выбраны переменные для какого-либо из тестов;
- «Необходимо загрузить данные!!!» — возникает при попытке выполнить запрос, без создания базы данных;
- «Исправьте описательные статистики!!!» — возникает при выборе переменных, которые не удовлетворяют условиям выбора переменных для описательных статистик;
- «Исправьте Хи-квадрат!!!» — возникает при выборе переменных, которые не удовлетворяют условиям выбора переменных для теста Хи-квадрат;
- «Исправьте Т-Тест!!!» — возникает при выборе переменных, которые не удовлетворяют условиям выбора переменных для Т-Теста;
- «Исправьте тест Манна-Уитни!!!» — возникает при выборе переменных, которые не удовлетворяют условиям выбора переменных для теста Манна-Уитни;
- «Исправьте ANOVA!!!» — возникает при выборе переменных, которые не удовлетворяют условиям выбора переменных для теста ANOVA;
- «Исправьте кластеризацию!!!» — возникает при выборе переменных, которые не удовлетворяют условиям выбора переменных для кластеризации;
- «Отчет создан!» — все хорошо.
- ...