

Распознавание Excel таблиц с помощью C#

Маринчук Александр Сергеевич

Приамурский государственный университет им. Шолом-Алейхема

Студент

Аннотация

Существует огромное количество различных программных продуктов, которые облегчают жизнь обычным пользователям и используются ими на ежедневной основе, и без которых сложно представить обыденный рабочий день. Один из таких продуктов это Microsoft Office, в частности Microsoft Excel, зарекомендовавший себя с лучшей стороны и ставший чуть ли не монополистом в сфере работы с табличными данными. Умение работать с данной программой является одним из наиболее частых требований при приеме на работу, а профессиональная обработка Excel таблиц высоко ценится начальством. В данной статье создана обработка для распознавания Excel таблиц в Microsoft Visual Studio на языке программирования C# с использованием элемента DataGridView.

Ключевые слова: Excel, распознавание, C#, таблица, DataGridView, MVS.

Recognizing an Excel tables using C

Marinchuk Alexander Sergeevich

Sholom-Aleichem Priamursky State University

Student

Abstract

There are a huge number of different software products that make life easier for ordinary users and are used by them on a daily basis, and without which it is difficult to imagine an ordinary working day. One of these products is Microsoft Office, in particular Microsoft Excel, which has proven itself from the best side and has become almost a monopoly in the field of working with tabular data. The ability to work with this program is one of the most frequent requirements when applying for a job, and professional processing of Excel spreadsheets is highly appreciated by the bosses. This article has created processing for recognizing Excel tables in Microsoft Visual Studio in the C # programming language using the DataGridView control.

Keywords: Excel, recognizing, C #, table, DataGridView, MVS.

1. Введение

1.1 Актуальность исследования

Существует огромное количество различных программных продуктов, которые облегчают жизнь обычным пользователям и используются ими на

ежедневной основе, и без которых сложно представить обыденный рабочий день. Один из таких продуктов это Microsoft Office, в частности Microsoft Excel, зарекомендовавший себя с лучшей стороны и ставший чуть ли не монополистом в сфере работы с табличными данными. Умение работать с данной программой является одним из наиболее частых требований при приеме на работу, а профессиональная обработка Excel таблиц высоко ценится начальством. Но в современном развивающемся мире не мало важным является загрузка табличных частей в различные языки программирования и их считывание для дальнейшей обработки и преобразования. На данный момент создано множество библиотек, которые «понимают» и могут «читать» Excel таблиц, но конкретно в данной статье будет рассмотрена библиотека ExcelDataReader в составе Microsoft Visual Studio, являющаяся достаточно популярным решением для распознавания различных таблиц.

1.2 Обзор исследований

В статье Н. Н. Елизаровой рассматриваются статистические пакеты обработки данных STATISTICA, SPSS, Deductor, а также возможности MS Excel, которые могут быть использованы для задач принятия решений планирования и управления различных объектов [1]. В работе А. О. Кизянова представлено практическое создание программы на языке программирования Python для извлечения данных из электронных таблиц MSEXcel [2]. При обработке различных опросных данных часто возникают технические трудности их переноса из опросных форматов (например, excel-таблиц) в данные программы, непосредственно занимающейся их математической обработкой. В статье В. Н. Шумилова и Е. Г. Андриановой рассмотрен случай, когда надо интегрировать средства программы обчета данных, реализованной на C++ и программы выгрузки этих данных из excel-таблиц, реализованной на C#, а также описана последовательность действий по программной интеграции объектно-ориентированных библиотек. Приведены фрагменты программного кода реализации указанной последовательности действий [3]. Дается описание надстройки ExcelToR, разработанной для совместного использования Microsoft Excel и статистического пакета R в исследовании А. Б. Новаковского [4]. Описана реализация балльно-рейтинговой системы оценки знаний в системе математического образования для нематематических специальностей на основе взаимодействия Maplet с базой данных в формате Excel в статье Ю. Г. Игнатьева и А. Р. Самигуллиной [5].

1.3 Цель исследования

Целью данного исследования разработка программы для распознавания Excel таблиц с помощью C#.

2. Методы исследования

Для начала создадим пустой проект в Microsoft Visual Studio [6] и добавим на форму элементы MenuStrip, ToolStrip, DataGridView и OpenFileDialog (рис. 1).

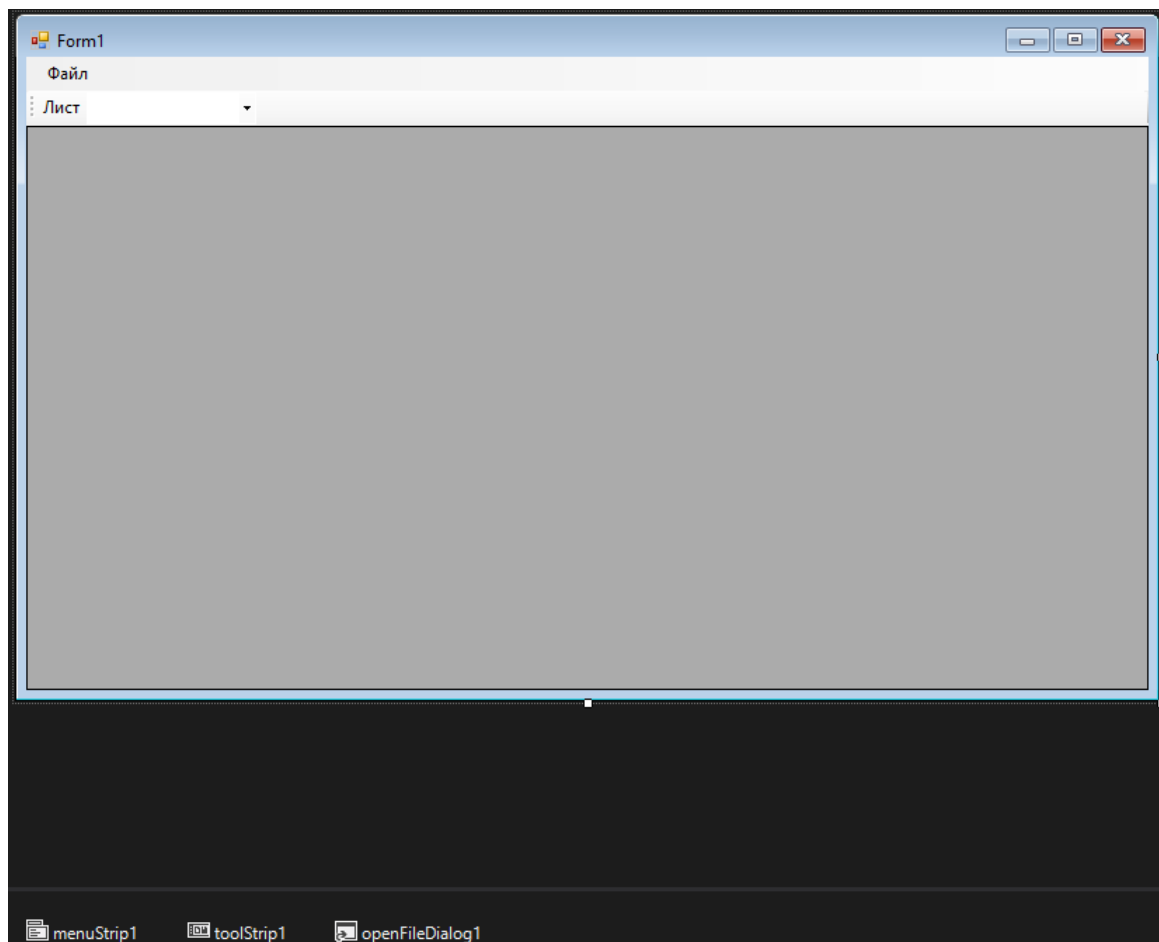


Рисунок 1 – Форма проекта

Далее установим 2 библиотеки: ExcelDataReader и ExcelDataReader.DataSet через вкладку «Управление пакетами NuGet» (рис. 2).

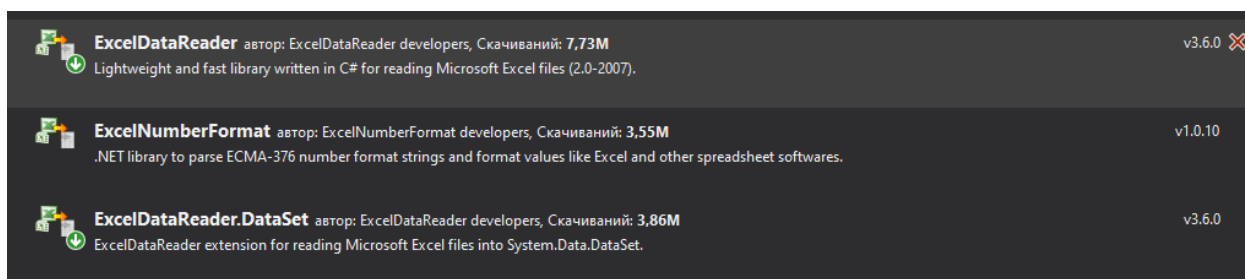


Рисунок 2 – Установка необходимых библиотек

3). Подключаем все необходимые библиотеки для работы программы (рис.

```
using System;
using System.Collections.Generic;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.IO;
using System.Linq;
using System.Text;
using System.Threading.Tasks;
using System.Windows.Forms;
using ExcelDataReader;

namespace ExcelRecognition
```

Рисунок 3 – Подключение библиотек

После создаем нужные переменные, необходимые для хранения имени открываемого файла и листов Excel таблицы (рис. 4).

```
namespace ExcelRecognition
{
    Ссылка: 3
    public partial class Form1 : Form
    {
        private string fileName = string.Empty;

        private DataTableCollection tableCollection = null;
        ссылка: 1
        public Form1()
        {
            InitializeComponent();
        }
    }
}
```

Рисунок 4 – Объявление переменных

Далее напишем обработчик нажатия на кнопку «Открыть», который при успехе будет открывать файл и записывать его название в заголовок, а в случае ошибки выводить предупреждающее окно (рис. 5).

```
ССЫЛКА: 1
private void открытьToolStripMenuItem_Click(object sender, EventArgs e)
{
    try
    {
        DialogResult res = openFileDialog1.ShowDialog();

        if(res == DialogResult.OK)
        {
            fileName = openFileDialog1.FileName;
            Text = fileName;

            OpenExcelFile(fileName);
        }
        else
        {
            throw new Exception("Файл не выбран");
        }
    }
    catch(Exception ex)
    {
        MessageBox.Show(ex.Message, "Ошибка", MessageBoxButtons.OK, MessageBoxIcon.Error);
    }
}
```

Рисунок 5 – Открытие файла и обработка ошибок

Затем создадим метод `OpenExcelFile`, отвечающий за открытие и считывание Excel таблицы. Также напишем обработчик изменения текущего листа в Excel, отвечающего за отображение таблицы в элементе `DataGridView` (рис. 6).

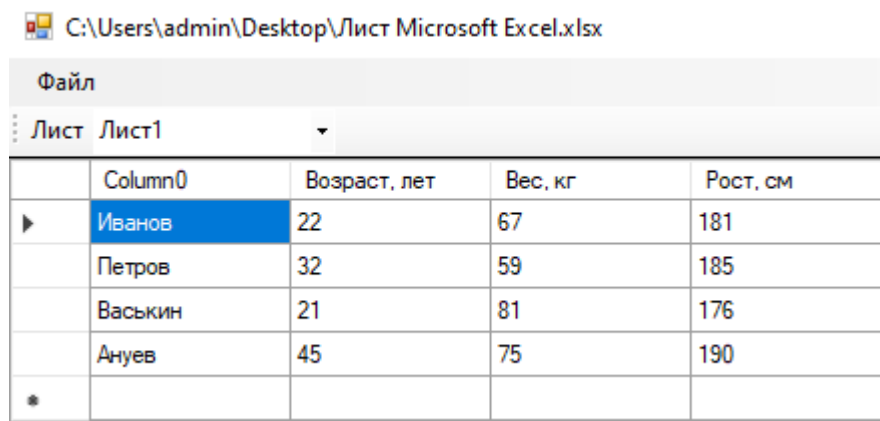
```
private void OpenExcelFile(string path)
{
    FileStream stream = File.Open(path, FileMode.Open, FileAccess.Read);
    IExcelDataReader reader = ExcelReaderFactory.CreateReader(stream);
    DataSet db = reader.AsDataSet(new ExcelDataSetConfiguration()
    {
        ConfigureDataTable = (x) => new ExcelDataTableConfiguration()
        {
            UseHeaderRow = true
        }
    });
    tableCollection = db.Tables;

    toolStripComboBox1.Items.Clear();
    foreach(DataTable table in tableCollection)
    {
        toolStripComboBox1.Items.Add(table.TableName);
    }
    toolStripComboBox1.SelectedIndex = 0;
}

ССЫЛКА: 1
private void toolStripComboBox1_SelectedIndexChanged(object sender, EventArgs e)
{
    DataTable table = tableCollection[Convert.ToString(toolStripComboBox1.SelectedItem)];
    dataGridView1.DataSource = table;
}
```

Рисунок 6 – Функция `OpenExcelFile` и обработчик события `SelectedIndexChanged`

Запустим нашу программу и откроем любой Excel файл, чтобы проверить работу приложения (рис. 7).



The screenshot shows a console window titled "C:\Users\admin\Desktop\Лист Microsoft Excel.xlsx". Below the title bar is a "Файл" (File) menu with "Лист" (Sheet) and "Лист1" (Sheet1) options. The main content is a table with the following data:

Column0	Возраст, лет	Вес, кг	Рост, см
Иванов	22	67	181
Петров	32	59	185
Васькин	21	81	176
Ануев	45	75	190

Рисунок 7 – Консольное окно с выводом нажатых клавиш

Также можно выбирать любые листы из файла, и данные из них будут автоматически отображаться в виде таблицы в элементе DataGridView.

3. Выводы

Работа с Excel таблицами стала неотъемлемой частью современного мира, где работа с большим набором данных является обыденной вещью. Продукт Excel имеет в своем арсенале огромное множество полезных инструментов, облегчающих жизнь людям. Именно поэтому возможность интегрировать существующие таблицы из Excel в другие программные продукты, а также возможность работы с ними является значимой функцией любого языка программирования. В данной статье была рассмотрена разработка программы для распознавания Excel таблиц с помощью Microsoft Visual Studio на языке программирования C#.

Библиографический список

1. Елизарова Н. Н. Использование программных средств статистической обработки данных при формировании информационного обеспечения управления // Вестник Ивановского государственного энергетического университета. 2009. № 3.
2. Кизянов А. О. Импорт данных из файлов MS Excel с помощью языка программирования python // Постулат. 2017. № 8.
3. Шумилов В. Н., Андрианова Е. Г. Программа переноса данных из Excel-таблиц в программу их обработки // ИТ-стандарт. 2018. №1. С. 32-39.
4. Новаковский А. Б. Взаимодействие Excel и статистического пакета R для обработки данных в экологии // Вестник института биологии Коми научного центра уральского отделения РАН. 2016. №3. С. 26-33.
5. Игнатъев Ю. Г., Самигуллина А. Р. Реализация балльно-рейтинговой системы оценки знаний на основе взаимодействия Maplet с базой данных

в формате Excel // Системы компьютерной математики и их приложения. 2015. №16. С. 248-250.

6. Microsoft Visual Studio URL: <https://visualstudio.microsoft.com/ru/downloads/> (дата обращения: 03.09.2020).