

A

B

C

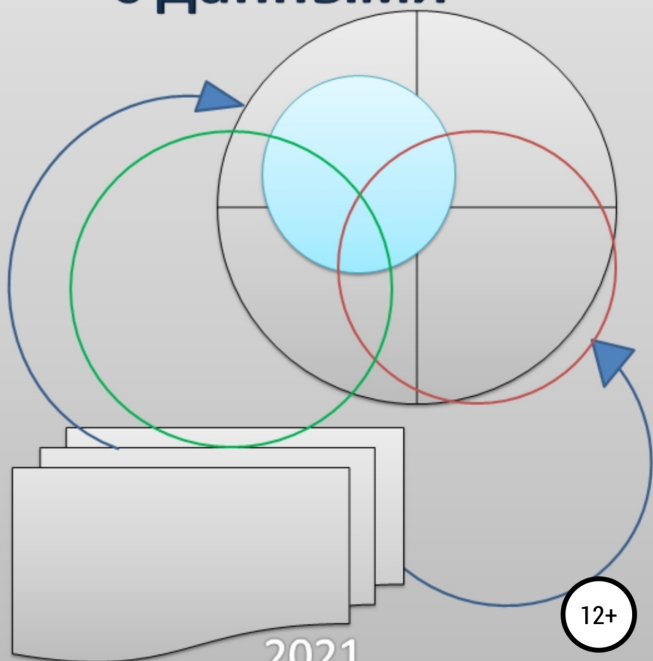
D

E

Лубягин Юрий

MS Excel

Приемы работы с данными



2021

12+

Юрий Николаевич Лубягин

MS Excel. Приемы работы с данными

http://www.litres.ru/pages/biblio_book/?art=66981378

SelfPub; 2023

Аннотация

В книге собраны приемы работы с тестовой строкой и массивами данных. Даны способы анализа массивов. Приведён пример практической реализации создания интерактивной системы автоматических расчетов. Для изучения материала необходимы базовые навыки работы в MS Excel.

Содержание

Введение	5
Глава 1. Приемы работы с текстовой строкой	7
Получение первого слова с начала текстовой строки	7
Получение любого слова из текстовой строки	10
Получение последовательности слов с начала текстовой строки	14
Получение слова с конца текстовой строки	15
Получение слова с конца строки заданной длины	17
Получение последовательности слов из середины текстовой строки	18
Глава 2. Приемы работы с массивами данных	20
Работа с одномерными массивами данных	21
Проверка наличия и удаление дублей данных в одномерном массиве	21
Поиск повторяющихся данных в массиве	24
Конец ознакомительного фрагмента.	28

Юрий Лубягин

MS Excel. Приемы работы с данными

Всегда есть способ любую работу делать быстрее и качественнее, чем вы ее делаете сейчас.

Введение

Пожалуй не найти ту область деятельности, где нельзя было бы использовать электронные таблицы MS Excel. Когда стоимость специализированного программного обеспечения несоизмеримо большая, особенно когда это касается личных потребностей или ведения небольшого бизнеса, электронные таблицы Excel подходят как никогда лучше.

В небольшой компании или индивидуальном бизнесе все данные можно хранить в электронных таблицах MS Excel, предварительно продумав систему безопасного хранения – лучше для этого использовать достаточно распространённые сервисы облачного хранения данных. Возможности Excel довольно обширные и позволяют проводить быстро и качественно анализ в разрезе большого числа параметров, построить интерактивную систему и, меняя различные параметры, смотреть, как она себя ведет, и многие другие задачи, требующие вычислительные мощности.

В книге читатель ознакомится с приемами и алгоритмами работы с массивами данных, используя табличный процессор MS Excel наработанные в течение 10 лет практической работы.

В первой главе рассматриваются, как работают функции в Excel, а их огромное количество. На примере работы с текстом, рассмотрим алгоритмы и приемы использования тек-

стовых функций.

Во второй главе, рассматривается работа с массивами данных, модификация и их анализ.

В третьей главе на практикуме рассматривается построение интерактивной локальной автоматизированной информационной системы.

Важным моментом при изучении представленного материала является его параллельная с чтением практическая реализация, так он будет более понятным и усвоится быстрее.

Глава 1. Приемы работы с текстовой строкой

Получение первого слова с начала текстовой строки

Необходимо извлечь из текстовой строки первое слово, например, из фамилии имени отчества нам необходима фамилия. Для этого можем использовать функцию «**ЛЕВ-СИМВ**», согласно описанию в справке Excel, данная функция возвращает первый символ или несколько первых символов текстовой строки на основе заданного числа символов. Но это если нами известно количество символов, которые хотим извлечь из строки, к примеру, ИНН юридического лица всегда состоит из 10 цифр. Фамилия же состоит из разного количества символов. Для решения этой задачи используем дополнительно функцию — «**НАЙТИ**».

Функция «**НАЙТИ**» находит вхождение одной текстовой строки в другую и возвращает искомую позицию искомой строки относительно первого знака второй строки.

Синтаксис функции.

НАЙТИ(*искомый_текст*, *просматриваемый_текст*,
[нач_позиция])

Аргументы функции.

Искомый_текст – обязательный аргумент. Текст, который необходимо найти.

Просматриваемый_текст – обязательный аргумент. Текст, в котором нужно найти искомый текст.

Нач_позиция – необязательный аргумент. Знак, с которого нужно начать поиск. Первый знак в тексте "**просматриваемый_текст**" имеет номер 1. Если номер опущен, он полагается равным 1.

Рассмотрим строку «Фамилия Имя Отчество» видим, что после фамилии стоит пробел. Пробел это тоже знак, хотя мы его и не видим.

Алгоритм действий – для получения фамилии из строки с ФИО необходимо вернуть количество символов до первого пробела.

Используем функцию «**ЛЕВСИМВ**»

Синтаксис функции.

ЛЕВСИМВ(текст, [число_знаков])

Аргументы функции.

Текст Обязательный. Текстовая строка, содержащая символы, которые требуется извлечь.

Число_знаков Необязательный. Количество символов, извлекаемых функцией. "**Число_знаков**" должно быть больше нуля или равно ему. Если значение "**число_знаков**" опущено, оно считается равным 1.

Полезно знать. Входной переменной аргумента любой

функции может быть и другая функция.

В нашем примере число знаков, которые необходимо вернуть мы получим с помощью функции **«НАЙТИ»**.

Условимся, текстовая строка, с которой мы работаем, всегда находится в ячейке A1, а формулы последовательно вводим в следующие по порядку ячейки в строке. В ячейку B1 вводим формулу **«=НАЙТИ(" ";A1;1)»**. Чтобы указать адрес ячейки в формуле, достаточно только при вводе нажать курсором на необходимую ячейку. Тут функция **«НАЙТИ»** возвращает позицию пробела в тексте находящимся в ячейке A1, поиск идет с первого символа. Результатом функции будет число – позиция первого пробела с начала строки, которое мы используем в качестве аргумента число знаков функции **ЛЕВСИМВ**. В ячейку C1 вводим формулу **«=ЛЕВСИМВ(A1;B1-1)»**, тут функция **«ЛЕВСИМВ»** возвращает количество знаков до первого пробела из текста, находящегося в ячейке A1. Почему мы отнимаем один от результата функции **«НАЙТИ»**, она вернула положение пробела, и если этого не сделать, то функция **ЛЕВСИМВ** вернет слово с пробелом на конце.

Можно формулы всех ячеек записать в одну, для этого в формуле **«=ЛЕВСИМВ(A1;B1-1)»** вместо адреса ячейки B1 вводим формулу из этой ячейки, только без знака **«=»**, и получим следящую формулу **«=ЛЕВСИМВ(A1;НАЙТИ(" ";A1;1)-1)»**, – результатом будет первое слово до пробела из строки ФИО.

Получение любого слова из текстовой строки

Получив фамилию в предыдущем примере, логично получить и инициалы.

Чтобы получить инициалы необходимо отдельно, имя и отчество, и используя функцию **«ЛЕВСИМВ»**, где в качестве аргумента *Число_знаков* используем 1 или опускаем данный аргумент, получим первый символы имени и отчества.

Как получить имя из текстовой строки «Фамилия Имя Отчество»?

В строке «Фамилия Имя Отчество» слово имя стоит после пробела и после имени стоит пробел.

Алгоритм действий: для получения имени, необходимо вернуть из строки с ФИО количество символов от первого пробела до второго пробела.

Внимание! В строке не должно быть между словами лишних пробелов. Для этого можно использовать функцию **«СЖПРОБЕЛЫ»** она удаляет из текста все пробелы, за исключением одиночных между словами и имеет только один аргумент – текст. Если вы сомневаетесь в количестве пробелов в используемом вами тексте, то необходимо воспользоваться данной функцией.

Для реализации вышеописанного алгоритма используем

еще одну текстовую функцию – «ПСТР».

Функция **ПСТР** возвращает заданное количество знаков из текстовой строки, начиная с указанной позиции.

Синтаксис функции.

ПСТР(*текст, начальная_позиция, число_знаков*)

Аргументы функции.

Текст Обязательный. Текстовая строка, содержащая символы, которые требуется извлечь.

Начальная_позиция Обязательный. Позиция первого знака, извлекаемого из текста. Первый знак в тексте имеет начальную позицию 1 и так далее.

Число_знаков Обязательный. Указывает, сколько знаков должна вернуть функция «ПСТР».

Переменную аргумента «**начальная позиция**» получаем, функцией «**НАЙТИ**», в ячейке В1 вводим формулу «=**НАЙТИ**(" ";А1;1)+1» находим начальную позицию слова «Имя». Для получения переменной аргумента «**число_знаков**» необходимо от позиции второго пробела отнять начальную позицию слова «Имя». Для получения позиции второго пробела вспомним о необязательном аргументе функции «**НАЙТИ**» «**Нач_позиция**» позиция, с которой нужно начать поиск, в нашем примере поиск мы должны начать с позиции начала слова «Имя», иначе функция нам вернет позицию первого пробела.

В ячейку С1 вводим формулу «=**НАЙТИ**(" ";А1;В1)» получаем позицию второго пробела.

Для получения слова «Имя» в ячейку D1 вводим формулу **«=ПСТР(A1;B1;C1-B1)»**.

Запишем все формулы в виде одной.

Выполним последовательно следующие действия:

1. В ячейке C1 формулу **«=НАЙТИ(" ";A1;B1)»** меняем на **«=НАЙТИ(" ";A1;НАЙТИ(" ";A1;1)+1)»**;

2. В ячейке D1 формулу **«=ПСТР(A1;B1;C1-B1)»** меняем на **«=ПСТР(A1;НАЙТИ(" ";A1;1)+1;НАЙТИ(" ";A1;НАЙТИ(" ";A1;1)+1)-(НАЙТИ(" ";A1;1)+1)»**.

Для корректности расчетов формулу из ячейки B1 необходимо взять в скобки.

Но можно и не усложнять формулы так сильно, а скрывать столбцы с промежуточными расчетами.

Получив слово «Имя» и используя функцию **«ЛЕВСИМВ»** вводим в ячейку E1 формулу – **«=ЛЕВСИМВ(D1)»** и получим первую букву имени.

Для добавления точки в конце инициала, используем функцию **«СЦЕПИТЬ»**. Данная функция объединяет до 255 текстовых строк в одну. Объединяемые элементы могут быть текстом, числами, ссылками на ячейки или сочетанием этих элементов. Аргументами функции являются минимум две тактовые строки, которые необходимо сцепить. В ячейку F1 вводим формулу **«=СЦЕПИТЬ(E2;"")»** получаем первый инициал с точкой. Точка, как и любая текстовая строка, в формуле берется в двойные кавычки – **""**.

Для получения второго инициала действия аналогичные,

получаем слово «Отчество» берем от него первую букву и добавляем точку в конце.

Алгоритм действий получения отчества: вернуть из строки с ФИО количество символов от второго пробела до конца строки. Для реализации алгоритма используем функцию «ПСТР». Для получения переменной аргумента «*число_знаков*» нам необходимо от числа символов строки отнять начальную позицию слова «Отчество»+1.

Для получения позиции начала слова «Отчество» нужна позиция второго пробела, воспользуемся информацией полученной ранее. Позиция второго пробела находится в ячейке C1. В ячейку G1 вводим формулу «=C1+1» и получаем позицию начала слова «Отчество».

Для получения количества символов текстовой строки используется функция «ДЛСТР». В ячейку H1, вводим формулу «=ДЛСТР(A1)».

В ячейку I1 вводим формулу «=ПСТР(A1;G1;H1-G1+1)» и получим слово «Отчество». Далее используя функцию «ЛЕВСИМВ» и «СЦЕПИТЬ» вводим в ячейку J1 формулу «=СЦЕПИТЬ(ЛЕВСИМВ(I1);".")» получаем второй инициал. И для получения Фамилии с инициалами в ячейку K1 введем формулу «=СЦЕПИТЬ(ЛЕВСИМВ(A1;НАЙТИ(" ";A1;1)-1);" ";F2;" ";J2)».

Разберем еще несколько примеров использования текстовых функций.

Получение последовательности слов с начала текстовой строки

Например, у нас есть список из одного столбца, в котором в строку записано Фамилия Имя Отчество, дата рождения, нам необходимо разделить этот список на два столбца с ФИО и датой рождения отдельно.

Рассмотрим строку «Фамилия Имя Отчество 01.01.1900» видим, что строка «Фамилия Имя Отчество» заканчивается на третьем пробеле.

Алгоритм действий: чтобы получить «ФИО» необходимо вернуть количество символов до третьего пробела.

Порядок действий:

1. Получаем позицию первого пробела, в ячейку B1 вводим формулу «=**НАЙТИ**(" ";A1;1)»;
2. Используя данные первой позиции, получим позицию второго пробела, в C1 вводим формулу «=**НАЙТИ**(" ";A1;B1+1)»;
3. Далее используя данные позиции второго слова получим позицию третьего пробела и в D1 вводим формулу «=**НАЙТИ**(" ";A1;C1+1)».

И, используя функцию «**ЛЕВСИМВ**» с результатом третьей операции, как переменную аргумента **Число_знаков-1**, получаем Фамилию Имя Отчество, введя в ячейку E1 формулу «=**ЛЕВСИМВ**(A1;D1-1)».

Получение слова с конца текстовой строки

Далее из текста необходимо получить дату рождения. В строке «Фамилия Имя Отчество 01.01.1900» дата рождения идет после третьего пробела.

Алгоритм действий: чтобы получить дату рождения из строки «Фамилия Имя Отчество 01.01.1900» необходимо вернуть количество знаков от третьего пробела до конца строки.

Порядок действий:

1. В ячейке F1 вводим формулу «**=ДЛСТР(A1)**» и получаем количество символов всей строки;
2. Позицию третьего пробела берем из расчетов предыдущего примера, ячейка D1.
3. В ячейку G1 вводим формулу «**=ПСТР(A1;D1+1;ДЛСТР(A1)-D1)**» получаем строку – дата рождения.

Разберем пример. У функции «**ПСТР**» есть три аргумента это *текст*, *начальная_позиция*, *число_знаков*. Начальную позицию, с которой функция вернет слово, берем из ячейки D1, это позиция третьего пробела и прибавляем к нему 1, получаем позицию начала слова дата рождения. Чтобы получить число знаков, которое вернет функция, мы от количества знаков всей строки отнимаем позицию третьего пробела – ДЛСТР(A1).

Важно! При возвращении слова из конца строки данным способом, необходимо учитывать, что количество пробелов в текстовой строке должно быть одинаково во всем списке обрабатываемых строк. Функция, которая бы возвращала позицию вхождения одной строки в другую, просматривая строку с конца в Excel нет.

Получение слова с конца строки заданной длины

Для получения слова с конца строки заданной длины, используется функция «**ПРАВСИМВ**», она возвращает последний символ или несколько последних символов текстовой строки на основе заданного числа символов.

Синтаксис.

ПРАВСИМВ(*текст*,[*число_знаков*])

Аргументы функции.

Текст Обязательный. Текстовая строка, содержащая символы, которые требуется извлечь.

Число_знаков Необязательный. Количество символов, извлекаемых функцией «**ПРАВСИМВ**».

В предыдущем примере дата рождения в формате «00.00.0000» всегда имеет одинаковое количество символов – 10. И наиболее рациональнее было в предыдущем примере использовать функцию «**ПРАВСИМВ**», введя в ячейку H1 формулу «=**ПРАВСИМВ**(A1;10)» мы получим то же результат.

Получение последовательности слов из середины текстовой строки

Рассмотрим еще один пример. У нас есть список из одного столбца, в котором в строку записано ИНН, Фамилия Имя Отчество, дата рождения, и необходимо получить текст содержащий Фамилию Имя Отчество. Усложним задачу тем, что первое и последнее слово могут быть разной длины. Но количество пробелов во всех словах списка одинаковое.

Рассмотрим строку «123456789100 Фамилия Имя Отчество 01.01.1900» текст «Фамилия Имя Отчество» находится между первым и четвертым пробелом.

Алгоритм действий: Чтобы получить текст «Фамилия Имя Отчество» из строки «123456789100 Фамилия Имя Отчество 01.01.1900» необходимо вернуть строку от первого пробела до четвертого, чтобы узнать длину строки, надо отнять от позиции четвертого пробела позицию второго пробела минус один.

Порядок действий:

1. Находим позицию первого пробела – в ячейку B1 вводим формулу «=**НАЙТИ**(" ";A1;1)»;
2. Имея позицию первого пробела, находим позицию второго пробела и в ячейку C1 вводим формулу «=**НАЙТИ**(" ";A1;B1+1)»;
3. Так же находим позицию третьего пробела – в ячейку

D1 вводим формулу «**=НАЙТИ(" ";A1;C1+1)**»;

4. И получаем позицию четвертого пробела – в ячейку E1 вводим формулу «**=НАЙТИ(" ";A1;D1+1)**»;

5. Для получения текстовой строки «Фамилия Имя Отчество» в ячейку G1 вводим формулу «**=ПСТР(A1;B1+1;E1-B1-1)**».

В Excel достаточный арсенал функций для работы с текстом. А их применение зависит от особенности текстовой строки, и для получения нужного результата обычно необходимо их комплексное применение.

Глава 2. Приемы работы с массивами данных

Определимся с терминологией. Массив данных это совокупность взаимосвязанных сведений, подлежащих совместной обработке и хранению. Массив данных состоит из одной или более записей – набора данных. Данные принадлежащие одному массиву записываются по общим правилам. Массив данных может быть представлен в виде одного столбца или строки – одномерный или виде нескольких столбцов и строк – двумерный (матрица). Запись столбцов (полей) данных массива может быть разных форматов, наиболее часто используемые форматы – это число, дата и текст. Причем, число и дата могут вводиться как текст.

В Excel массив данных представлен в виде диапазона или нескольких диапазонов. Диапазон в Excel состоит из одной и более ячеек. В Excel ограничения на количество столбцов и строк на листе определены версией приложения. Версия 2010, на которой рассматриваются примеры, позволяет создать на одном листе диапазон из 16384 столбца и 1048576 строк.

Работа с одномерными массивами данных

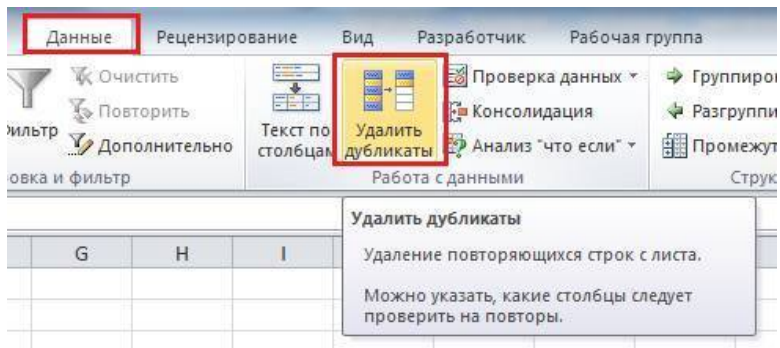
Одномерный массив данных, например, перечисление ИНН физических лиц или ФИО состоящий из одного столбца.

Операции с одномерным массивом

1. Проверка на наличие дублей записей и их удаление;
2. Поиск дублирующих записей;
3. Сравнение (поиск разности и пересечения).

Проверка наличия и удаление дублей данных в одномерном массиве

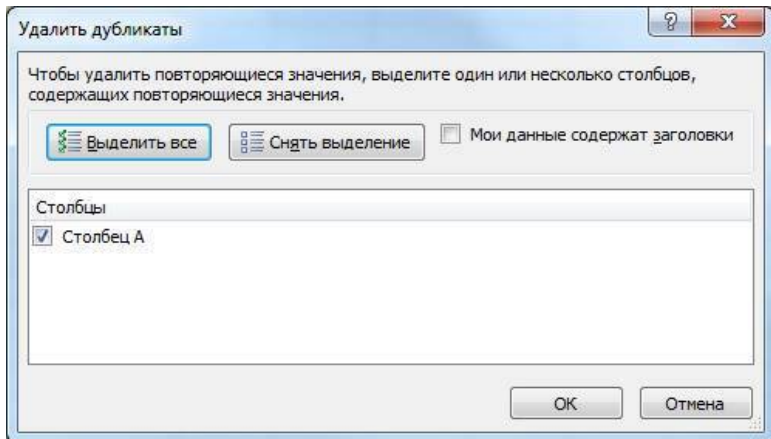
У нас есть массив данных, состоящий из 10000 записей и не известно, есть ли в нем повторяющиеся записи, а просматривать визуально потребует большие временные затраты. Для удаления дубликатов в Excel на вкладке «Данные» есть кнопка «Удалить дубликаты».



Выделяем весь диапазон данных. Для быстрого выделения необходимо выделить первую запись в списке (заголовок можно не выделять) и одновременно нажать на клавиатуре кнопки **Ctrl+Shift+↓** (стрелочка вниз). Это прием работает и с конца списка, достаточно выделить последнюю ячейку и одновременно нажать клавиши **Ctrl+Shift+↑** (стрелочка вверх). И так же и вправо и влево.

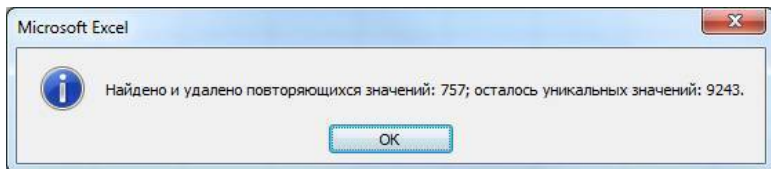
Выделив диапазон, на вкладке «Данные» нажимаем кнопку «Удалить дубликаты»

Выходит диалоговое окно.



Нажимаем «ОК»

Получаем информационное сообщение о количестве удаленных повторяющихся значений и количестве оставшихся уникальных значений.



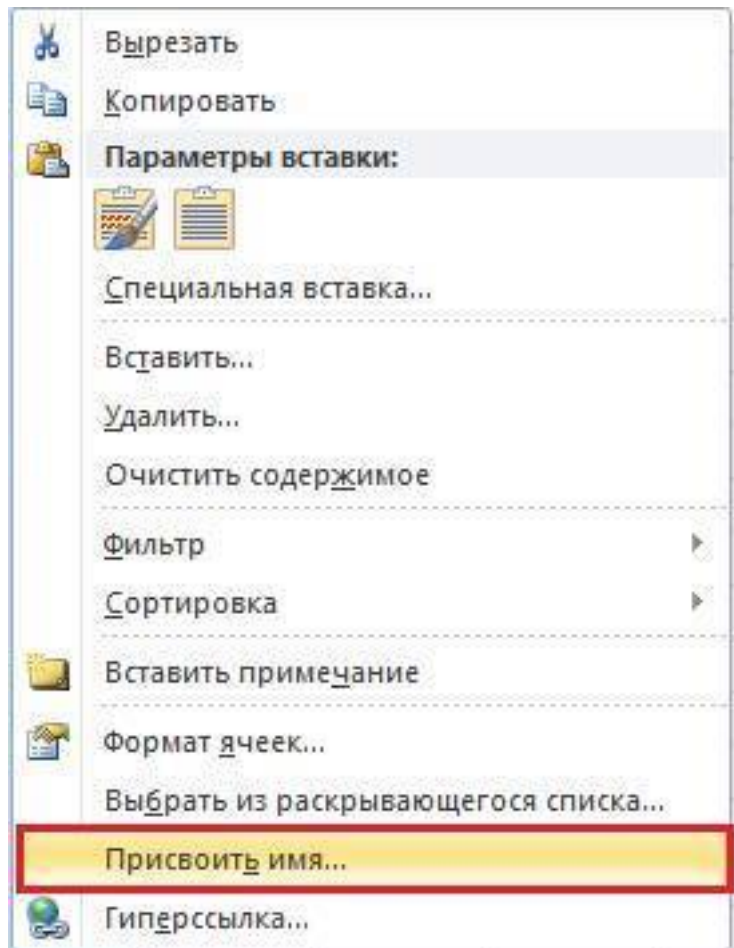
Дубли из массива удалены, все записи в нем уникальны.

Поиск повторяющих данных в массиве

Теперь необходимо узнать, какие записи и сколько раз повторяются. Вернем массив, в исходное состояние, через кнопку отмена, или используя сочетание клавиш Ctrl+Z.

Выделяем массив, копируем и вставляем рядом или на другой лист (в описываемом примере мы вставили список на этот же лист, в столбец D). И проводим над скопированным массивом операции по удалению повторяющих значений. Теперь у нас два массива, исходный и содержащий уникальные записи.

Исходному массиву присваиваем имя. Для этого выделяем его Ctrl+Shift+↓(стрелочка вниз) и не снимая выделения, нажатием правой клавиши мыши вызываем контекстное меню и выбираем пункт «Присвоить имя...»



Присваиваем имя выделенному диапазону.

Создание имени

Имя: список

Область: Книга

Примечание:

Диапазон: ='Модификация списка 3'!\$A\$1:\$A\$5

OK Отмена

При вводе недопустимого имени выходит сообщение.

Microsoft Excel

Введено недопустимое имя.

Возможные причины:

- имя не начинается с буквы или знака подчеркивания;
- имя содержит пробел или другие недопустимые знаки;
- имя совпадает со встроенным именем Excel или именем другого объекта в книге.

OK

(В примере диапазону ячеек массива присвоим имя «ПРИМЕР». В дальнейшем диапазонам рекомендую присваивать осмысленные имена.

Полезно знать. Присвоить имя можно и одной ячейке. Удобно это будет когда, используется функция « **СЦЕПИТЬ**

Конец ознакомительного фрагмента.

Текст предоставлен ООО «ЛитРес».

Прочитайте эту книгу целиком, [купив полную легальную версию](#) на ЛитРес.

Безопасно оплатить книгу можно банковской картой Visa, MasterCard, Maestro, со счета мобильного телефона, с платежного терминала, в салоне МТС или Связной, через PayPal, WebMoney, Яндекс.Деньги, QIWI Кошелек, бонусными картами или другим удобным Вам способом.