**Глава 9. Знакомство с функциями массива. ТРАНСП, МОДА.НСК и ТЕНДЕНЦИЯ**

Это глава из книги: Майкл Гирвин. Ctrl+Shift+Enter. Освоение формул массива в Excel.

[Предыдущая глава](http://baguzin.ru/wp/?p=7409) [Оглавление](http://baguzin.ru/wp/?p=7153) Следующая глава

Два типа формул массива могут возвращать более одного значения сразу в несколько ячеек:

* Созданные вами из различных элементов (о них вы узнали в [главе 8](http://baguzin.ru/wp/?p=7409)).
* Встроенные функции Excel (вы узнаете о них в этой главе).

Оба типа формул массива возвращают несколько элементов, подчиняясь одним и тем же правилам, подробно описанным в главе 8. Вот они вкратце:

1. Если формула массив вернет *n* элементов, и вы хотите отобразить *n* элементов в *n* ячейках, вы должны выбрать *n* ячеек до ввода формулы. Кроме того, необходимо учитывать, является ли массив одно- или двумерным. В первом случае вы выберите ряд или столбец ячеек, во втором – прямоугольную область.
2. Создайте свою формулу массива в активной ячейке.
3. Введите формулу массива нажатием Ctrl+Shift+Enter.

Функции массива являются встроенными функциями Excel, которые специально предназначены для возврата более чем одного элемента в более чем одну ячейку. Всего их восемь:



Первые три функции будут рассмотрены в этой главе; оставшиеся пять – в следующей.

**Функция ТРАНСП**

Функция преобразует вертикальный массив в горизонтальный (и наоборот), сохраняя ссылку на источник данных; работает как с одномерными, так и с двумерными массивами. На рис. 9.1 исходные данных двумерного массива размещены в диапазоне A2:D4. Если цель состоит в том, чтобы перенести данные в другое место и повернуть их на 90°, выполните следующее:

1. Подсчитайте число столбцов (4) и строк (3) в исходном диапазоне (А2:D4).
2. Выделите **3** столбца и **4** строки в новом месте, например, A7:C10.
3. В активной ячейке напечатайте формулу =ТРАНСП(A2:D4)
4. Введите формулу, с помощью Ctrl+Shift+Enter (рис. 9.2).

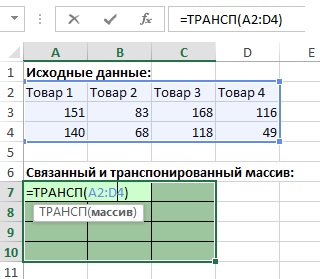


Рис. 9.1. Выделите диапазон, который имеет перекрестное количество строк и столбцов по отношению к исходному диапазону, и в активной ячейке напечатайте формулу

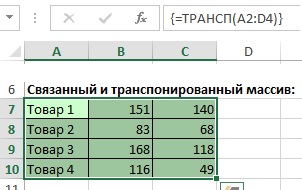


Рис. 9.2. Нажмите Ctrl+Shift+Enter и функция ТРАНСП создаст новый массив, ссылающийся на исходный; вы видите, что в строке формул появились фигурные скобки

Красота функция ТРАНСП (в отличие от *Специальной вставки* с опцией *транспонирования*) заключается в том, что в случае изменения исходных данных, данные в новом диапазоне тут же изменятся (рис. 9.3), в то время, как *Специальная вставка* вклеивает сами исходные данные (а не ссылку на них).[[1]](#footnote-1)

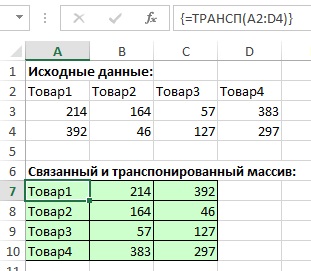


Рис. 9.3. При изменении исходных данных, транспонированные данные тут же обновятся

Как показано на рис. 9.4, иногда источник данных содержит пустые ячейки. Это приводит к тому, что функция ТРАНСП показывает ноль. Такие функции, как ТРАНСП, ВПР, ИНДЕКС и некоторые другие трактуют пустую ячейку как нулевую. Функции МИН, МАКС, СРЗНАЧ – наоборот, пропускают пустые ячейки (не берут их в расчет). Эти нюансы связаны с тем, как изначально запрограммированы функции Excel. Чтобы понимать поведение какой-то функции, читайте Help!

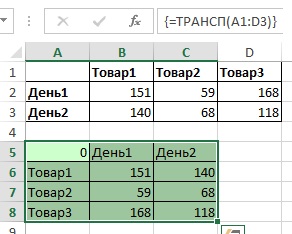


Рис. 9.4. Функция ТРАНСП трактует пустую ячейку, как нулевую

В [главе 4](http://baguzin.ru/wp/?p=7215) вы узнали, как использовать функцию ЕСЛИ, чтобы отфильтровать элементы массива внутри формулы. Сейчас же мы воспользуемся функцией ЕСЛИ, чтобы заменить ноль на пустую текстовую строку (способ показать «ничего» вместо нуля). Поскольку диапазон А1:D3 содержит пустую ячейку, следует поместить ЕСЛИ внутрь ТРАНСП (рис. 9.5). ЕПУСТО() возвращает значение ИСТИНА, если ячейка пуста, и значение ЛОЖЬ, в противном случае. Таким образом, функция ЕСЛИ() для ячейки А1 вернет значение нулевой текстовой строки (""), а для остальных ячеек в диапазоне А1:D3 их исходное значение. Чтобы увидеть работу функции ЕСЛИ, выделите ее целиком и нажмите F9 (рис. 9.6). Чтобы ввести формулу =ТРАНСП(ЕСЛИ(ЕПУСТО(A1:D3);"";A1:D3)) нажмите Ctrl+Shift+Enter (рис. 9.7).

Примечание. Нулевая текстовая строка – строка текста, которая имеет нулевую длину. Чтобы создать ее, введите две двойные кавычки, одна после другой, без пробела между ними. Ячейка, содержащая формулу с нулевой текстовой строкой (=""), показывается как пустая. В тоже время, если сделать такую ячейку активной, в строке формул мы увидим ="", а функция ЕПУСТО() при ссылке на такую ячейку вернет ЛОЖЬ, свидетельствуя, что ячейка не пуста.

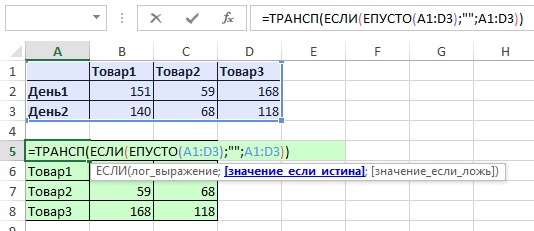


Рис. 9.5. Конструкция ЕСЛИ внутри ТРАНСП позволяет избежать нулей

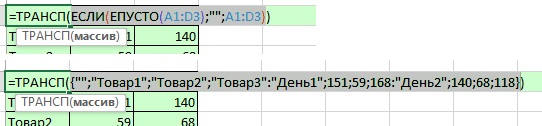


Рис. 9.6. Выделите элементы, как показано на верхней части рисунка, и нажмите F9

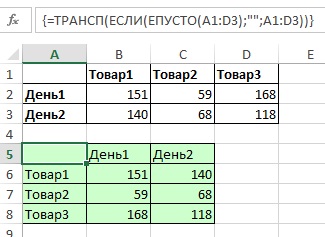


Рис. 9.7. Введите формулу, нажав Ctrl+Shift+Enter

**Вычисление моды**

Мода – статистический термин, определяющий элемент который наиболее часто встречается в выборке. Как правило, мода используется на номинальных данных (неколичественных, например, возрастные группы: до 20 лет, 20–30 лет и т.д.). Однако, мода также может быть использована и на количественных данных, как в этой главе.

На рис. 9.12 приведен список результатов заездов. Цель – найти результат, встречающийся наиболее часто. В Excel 2007 или более ранней версии для этого используется функция МОДА; в Excel 2010 г. или позже воспользуйтесь функцией МОДА.ОДН (ОДН обозначает, что функция вернет одно наиболее часто встречающееся значение; если два значения встречаются одинаково часто, функция вернет то, что появляется в выборке первым). МОДА и МОДА.ОДН дублируют друг друга, и вы можете использовать любую из них. На рис. 9.12 в диапазоне А2:А8 значение 24 встречается наиболее часто. Оно и является модой выборки.

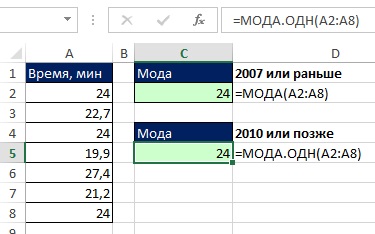


Рис. 9.12. Функции МОДА и МОДА.ОДН возвращают единственное значение моды выборки

Примечание: если начать вводить в ячейку формулу =МОДА, раскроется список (рис. 9.13). Рядом с функцией МОДА вы увидите желтый треугольный, указывающий, что эта функция оставлена в Excel 2010 для совместимости, т.е. электронные таблицы, созданные в более ранних версиях будут работать и в более поздних версиях. Зачем же тогда используется новая функция МОДА.ОДН? Начиная с Excel 2010 в ряде статистических функций применяется новый синтаксис: функция, «точка», суффикс. Разработчики Excel посчитали, что старые названия не очень хорошо описывают суть функций. К тому же, возможно, через некоторое время Microsoft перестанет поддерживать старые функции, и останутся только новые с «красивыми» названиями.

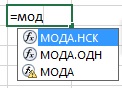


Рис. 9.13. Начиная с Excel 2010 в ряде статистических функций применяется новый синтаксис: функция, «точка», суффикс; желтый треугольник указывает, что функция оставлена в Excel 2010 для совместимости с более ранними версиями

Наиболее часто встречаемое значение в выборке может быть не одно. На рис. 9.14 результаты 23,7 и 25,3 встречаются по три раза. Обратите внимание, что и функция МОДА, и МОДА.ОДН возвращают только первое значение. Чтобы обнаружить все значения моды в выборке, используйте функцию массива МОДА.НСК. Выделите две ячейки, введите формулу и нажмите Ctrl+Shift+Enter. Часто вы не знаете, сколько значений мод вернет функция (к тому же, в зависимости от исходных данных число мод может меняться), поэтому перед вводом формулы просто выберите побольше ячеек (рис. 9.15). Если вас не раздражают значения ошибки #Н/Д, то на этом можно и остановиться. Такой подход можно назвать «слепым».

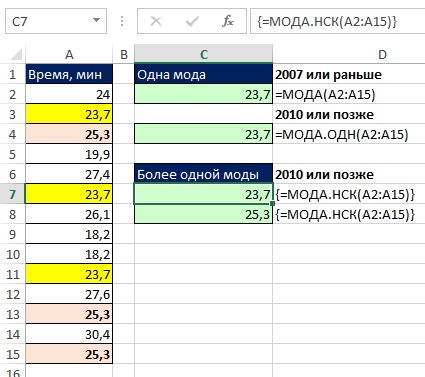


Рис. 9.14. Функция массива МОДА.НСК выявит все моды в выборке

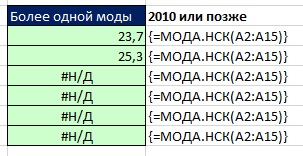


Рис. 9.15. Если вы не знаете, сколько значений мод вернет функция, выделите побольше ячеек прежде, чем ввести формулу

Причина, по которой часть ячеек заполнена ошибками #Н/Д в том, что вы выделили слишком много ячеек. На рис. 9.16 показано, что если вычислить формулу в режиме редактирования, вы точно увидите, сколько элементов массива вернет функция МОДА.НСК. Еще одним недостатком «слепого» метода является то, что в случае единственной моды, МОДА.НСК вернет одинаковые значения во все ячейки (рис. 9.17).

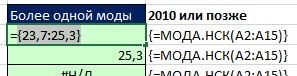


Рис. 9.16. Функция МОДА.НСК возвращает вертикальный массив чисел

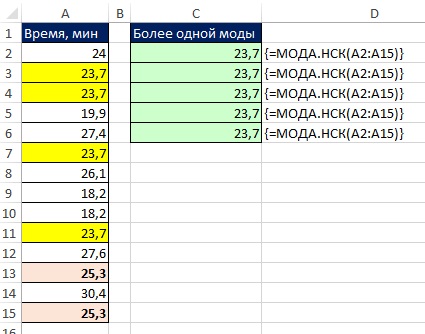


Рис. 9.17. В случае единственной моды, МОДА.НСК вернет одинаковые значения во все ячейки

Таким образом, вместо «слепого» метода гораздо лучшим решением было бы использовать сначала формулу, которая подсчитывает количество мод, а потом вторую формулу, которая отображает значения мод. Как показано на рис. 9.18, введите в ячейке C2 функцию СЧЁТ с аргументом *значение1* равным МОДА.НСК(A2:A10). Функция СЧЁТ вернет число мод.

Примечание. Обратите внимание, что при вводе функции МОДА.НСК в несколько ячеек требуется нажатие Ctrl+Shift+Enter. Когда же функция МОДА.НСК используется в качестве аргумента другой функции (за исключением функций СЧЁТЕСЛИ, СУММЕСЛИ и т.п.), нажатия Ctrl+Shift+Enter не требуется. Так ведут себя все функции массива за исключением функция ТРАНСП.

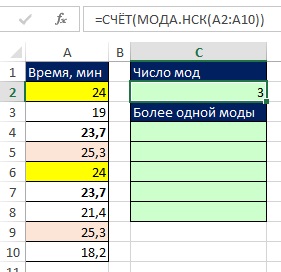


Рис. 9.18. Функция МОДА.НСК возвращает массив чисел, попадающий в аргумент функции СЧЁТ, которая, в свою очередь, возвращает количество чисел в массиве

Теперь, чтобы поместить значение первой моды в ячейку С4 начните с функции ИНДЕКС (рис. 9.19). Если выделить элементы, относящиеся к функции МОДА.НСК, и нажать F9, можно увидеть вертикальный массив, состоящий из трех значений мод.



Рис. 9.19. Введите функцию МОДА.НСК в качестве аргумента *массив* функции ИНДЕКС

Поскольку вы планируете извлекать последовательно элементы вертикального массива при «протаскивании» формулы вдоль столбца, вам нужен счетчик порядковых номеров 1, 2, 3 и т.д. В [главе 8](http://baguzin.ru/wp/?p=7409) вы использовали для этой цели функцию ЧСТРОК (рис. 9.20). В ячейке С4 функция ЧСТРОК(C$4:C4) вернет единицу, в ячейке С5 – двойку и т.д. Если подставить этот фрагмент в аргумент *номер\_строки* функции ИНДЕКС, функция вернет соответствующее значение моды: в ячейке С4 – первое значение моды, в ячейке С5 – второе значение моды и т.д.



Рис. 9.20. ЧСТРОК(C$4:C4) используют в качестве аргумента *номер\_строки* функции ИНДЕКС

Как показано на рис. 9.21 после ввода формулы (вам не нужно использовать Ctrl+Shift+Enter) и копирования ее вниз по колонке, вы получаете ошибки в части ячеек.

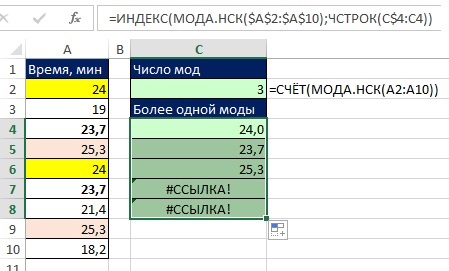


Рис. 9.21. Ошибка #ССЫЛКА! в ячейке С7 связана с тем, что функция ЧСТРОК(C$4:С7) возвращает четверку, а функция ИНДЕКС не может найти четвертый элемент в массиве, возвращенном МОДА.НСК, так как элементов в массиве всего три

Вы можете внести изменения в формулу, добавив условие: «показывать нулевую текстовую строку», если номер строки в диапазоне больше, чем количество мод (рис. 9.22). Когда ЧСТРОК(C$4:C4) вернет значение больше трех, аргумент *лог\_выражение* функции ЕСЛИ примет значение ИСТИНА, и функция ЕСЛИ вернет нулевую текстовую строку. Вы можете ввести формулу в ячейку C4 (вам не нужно использовать Ctrl+Shift+Enter) и скопировать ее вниз по диапазону С4:С8 (рис. 9.23). Красота этой формулы заключается в том, что при изменении данных, формула по-прежнему будет показывать только значимые величины, оставляя остальные ячейки как-бы пустыми (помните, что реально в них содержится текстовая строка нулевой длины).



Рис. 9.22. Поместите в аргумент *лог\_выражение* функции ЕСЛИ проверку на то, что номер строки в диапазоне С4:С8 больше числа мод

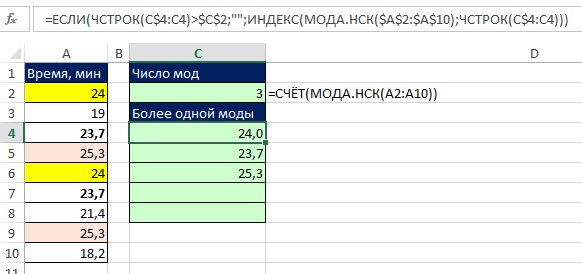


Рис. 9.23. Итоговая формула для отображения всех мод выборки

В качестве альтернативы для извлечения мод можно использовать формулу, приведенную на рис. 9.24. Функция ЕСЛИОШИБКА появилась в Excel с версии 2007. Функция возвращает аргумент *значение*, если оно имеет смысл; а если аргумент *значение* выдает ошибку, то функция возвращает значение, хранящееся в аргументе *значение\_если\_ошбка*; в нашем случае это текстовая строка нулевой длины. Эту функцию проще создать, чем предыдущую (см. рис. 9.23), но и работает она медленнее. В нашем случае это не критично, но на больших массивах данных, вы должны учитывать, что функция ЕСЛИ работает быстрее, чем ЕСЛИОШИБКА.



Рис. 9.24. Неплохая альтернатива – конструкция с ЕСЛИОШИБКА(…)

Несмотря на более медленную работу, функция ЕСЛИОШИБКА может быть очень полезна в следующей ситуации (рис. 9.25). Функция ВПР может возвращать ошибку, если ее аргумент *искомое\_значение* не найдет в аргументе *таблица*. В версиях Excel до 2007 вы должны были использовать функцию ЕСЛИ в комбинации с ЕОШИБКА, чтобы заменить значение ошибки на что-то удобоваримое, например, текстовую строку: «Отсутствует в таблице». Начиная с Excel 2007 функция ЕСЛИОШИБКА справляется с работой существенно прощу и быстрее (потому что вы используете ВПР только один раз).

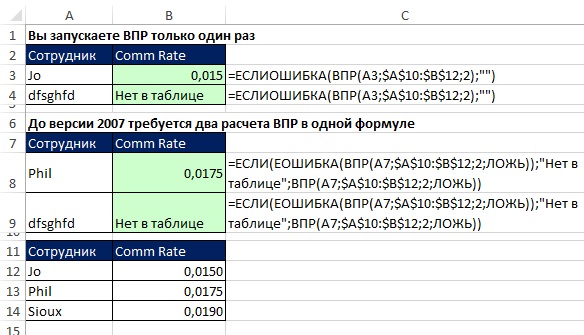


Рис. 9.25. ЕСЛИОШИБКА отлично подходит для ВПР, так как требует только одного вхождения ВПР

Если у вас есть много данных, и вы хотите быстро и легко определить моду, используйте сводную таблицу (рис. 9.26):

1. Щелкните любую ячейку в столбце исходных данных (А1:А18).
2. Перейдите на закладку *Вставка* и в области *Таблица* щелкните на *Сводная таблица*; в открывшемся окне переключитесь в положение *на существующий лист* и в области *Диапазон* введите С1. Нажмите Ok.
3. В появившейся области *Поля сводной таблицы* перетащите поле *Значение* в область *Строки*.
4. Еще раз перетащите поле *Значение* в область *Значения*.
5. Щелкните правой кнопкой мыши на любой ячейке колонки D сводной таблицы, и в контекстном меню выберите *Параметры полей значений*. В открывшемся окне в области *Операция* выберите *Количество*. Нажмите Ok.
6. Еще раз щелкните правой кнопкой мыши на любой ячейке колонки D сводной таблицы, и в контекстном меню выберите *Сортировка* → *Сортировка по убыванию*.

Одно или несколько (равных) первых значений и есть мода. В нашем пример это 28 и 23, встретившиеся в выборке по три раза.

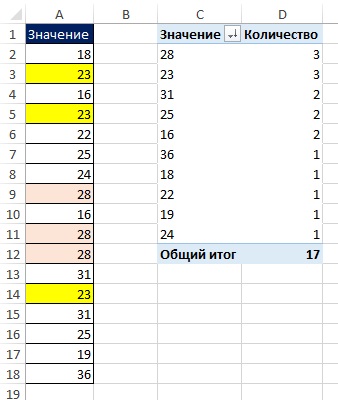


Рис. 9.26. Используйте сводную таблицу для быстрого вычисления моды (нескольких мод)

**Функция ТЕНДЕНЦИЯ**

В статистике принято использовать исходные данные, для прогнозирования будущих значений на основе линейной модели (подробнее см., например, [Простая линейная регрессия](http://baguzin.ru/wp/?p=6078)). Возможно, вы помните, как выглядит уравнение прямой *у = f(x) = mx+b*. Здесь *m* – угол наклона прямой к оси *Ох*, или, другими словами, насколько увеличится *y*, если *х* увеличится на 1; *b* – значение *у* в точке, в которой, линия пересекает ось *у* (т.е. значение *х* = 0).

На рис. 9.27–9.29 показаны три метода оценки баллов студентов, которые они могут получить за тест. Оценка основана на количестве часов, которое студенты посвятили подготовке к тестированию. Все четыре функции (НАКЛОН, ОТРЕЗОК, ПРЕДСКАЗ и ТЕНДЕНЦИЯ) обсчитывают известные *у* и известные *x* (данные из прошлого). Если вы используете функции НАКЛОН (*m*) и ОТРЕЗОК (*b*), вы раздельно вычисляете коэффициенты уравнения *у = mx+b* (рис. 9.27). Если вы используете функции ПРЕДСКАЗ и ТЕНДЕНЦИЯ, вы рассчитываете оценочное значение *y* для нового значения *x* (рис. 9.28 и 9.30), при этом расчет *m* и *b* явно не фигурирует. Для небольшого набора данных, наверное лучше подойдет функция ПРЕДСКАЗ, так как она не требует нажатия Ctrl+Shift+Enter. В то же время функция массива ТЕНДЕНЦИЯ рассчитывает *m* и *b* только один раз, и, следовательно, работает быстрее, что может пригодиться, когда вы имеете дело с большими наборами данных.

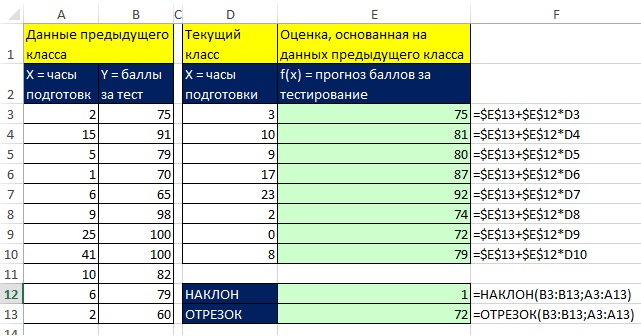


Рис. 9.27. Функции НАКЛОН и ОТРЕЗОК раздельно вычисляете коэффициенты уравнения *у=mx+b*, а затем отдельная формула вычисляет предсказание для теста

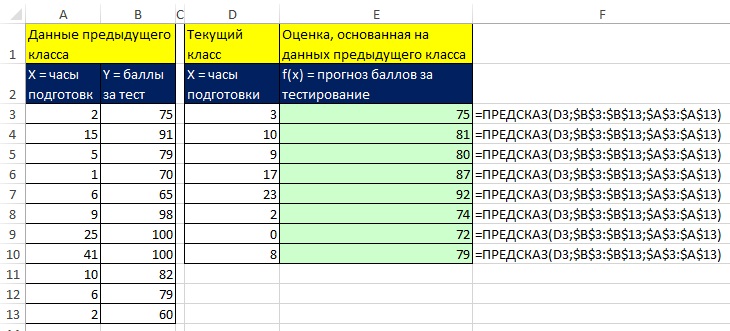


Рис. 9.28. Для небольшого набора данных хорошо подходит функция ПРЕДСКАЗ, так как она не требует нажатия Ctrl+Shift+Enter



Рис. 9.29. Функция массива ТЕНДЕНЦИЯ рассчитывает *m* и *b* только один раз, и, следовательно, работает быстрее, что может пригодиться, когда вы имеете дело с большими наборами данных

1. Любопытно, что инструмент Excel *Специальная вставка* имеет опцию *Вставить связь*, но… одновременно *Транспонировать* и *Вставить связь* не работают. [↑](#footnote-ref-1)