## Написание сложных формул на языке DAX Power Pivot с использованием ChatGPT

Я увлекаюсь игрой [Fantasy Premier League](https://fantasy.premierleague.com/). Помимо спортивного азарта соперничества с 10М игроков по всему миру, игра создаёт почву для серьёзных статистических исследований. Особенно после того, как я открыл для себя мир [API](https://baguzin.ru/wp/?p=25833). Большие данные, получаемые с помощью API, я помещаю в Excel, в модель данных Power Pivot. Дальнейший анализ веду с использованием формул на языке DAX (Data Analysis Expressions). Я только осваиваю этот язык, а практические задачи требуют продвинутых навыков. На помощь приходит ChatGPT.

Я взял итоги выступления игроков в сезоне 2022/23, поместил их в модель данных и вывел на лист Excel в виде сводной таблицы. Ниже показан фрагмент таких данных:



Рис. 1. Фрагмент сводной таблицы

В столбцах указаны: ID игрока, его имя, команда, за которую он выступает, позиция, номер тура, стоимость и набранные очки. В исходных данных отсутствует важный параметр для анализа – форма. Строго говоря, форма – среднее количество очков, набранных игроком за последние 30 дней. Для целей анализа я упрощаю определение, и хочу рассчитать форму, как среднее число очков, набранных в предыдущих четырех турах. Таким образом в первых четырех турах форму определить нельзя, а в пятом туре форма равна среднему числу очков, набранных в турах 1–4. И т.д.

Рассчитать форму я хочу не на листе Excel, используя приведенную сводную таблицу, а в модели данных. Мало того, что я не знаю, какой использовать код, так еще и строки в модели данных не отсортированы.



Рис. 2. Фрагмент модели данных в Power Pivot

Ну что ж… обратимся к ChatGPT.

#### Вопрос

Напиши формулу на языке DAX. В столбце А номера категорий. В столбце В номера периодов. В столбце С объем продаж. Нужно добавить столбец D, в котором подсчитать скользящее среднее по столбцу С за 4 предшествующих периода. Если строк за предыдущие периоды меньше четырех, то в столбце D указываем 0. Строки в таблице не отсортированы.

#### Ответ

Для расчета скользящего среднего в столбце D по столбцу C с учетом предыдущих четырех периодов в Power BI или Excel с использованием языка DAX, вы можете использовать следующую формулу:

Moving Average =

VAR CurrentCategory = [Category] // Здесь [Category] – это ссылка на столбец с категориями

VAR CurrentPeriod = [Period] // Здесь [Period] – это ссылка на столбец с номерами периодов

VAR PreviousPeriods = FILTER(ALL('YourTable'), [Category] = CurrentCategory && [Period] < CurrentPeriod)

VAR CountPreviousPeriods = COUNTROWS(PreviousPeriods)

RETURN

 IF(

 CountPreviousPeriods < 4,

 0, // Если предыдущих периодов меньше 4, то возвращаем 0

 AVERAGEX(TOPN(4, PreviousPeriods, [Period], ASC), [Sales]) // Иначе считаем среднее за последние 4 периода

 )

В этой формуле мы сначала определяем текущую категорию и текущий период, а затем создаем временную таблицу PreviousPeriods, которая фильтруется так, чтобы включать только строки с предыдущими периодами для текущей категории. Затем мы подсчитываем количество предыдущих периодов.

Используя функцию IF, мы проверяем, если количество предыдущих периодов меньше 4, то возвращаем 0. В противном случае мы используем функцию AVERAGEX для расчета среднего значения продаж ([Sales]) за последние 4 периода, отсортированных по номерам периодов в возрастающем порядке.

Замените 'YourTable', [Category], [Period] и [Sales] на соответствующие имена вашей таблицы и столбцов данных. Эта формула будет рассчитывать скользящее среднее в столбце D для каждой строки в вашей таблице, учитывая условие на количество предыдущих периодов.

#### Доработка кода

После подстановки названий таблицы и столбцов код не заработал, но моих знаний хватило, чтобы его поправить.

=

VAR CurrentCategory = [Element]

VAR CurrentPeriod = [event]

VAR PreviousPeriods = FILTER(ALL(ElementSummary);
 [Element] = CurrentCategory && [event] < CurrentPeriod)

VAR CountPreviousPeriods = COUNTROWS(PreviousPeriods)

RETURN

 IF(

 CountPreviousPeriods < 4;

 0;

 AVERAGEX(TOPN(4; PreviousPeriods; [event];DESC); [total\_points])

 )

Я убрал название Moving Average, и начал формулу со знака равно. Заменил разделитель запятую на точку с запятой. Изменил тип сортировки периодов, заменив аргумент ASC на DESC. Вуаля!



Рис. 3. Форма игрока выведена в сводную таблицу