



PowerBIUG

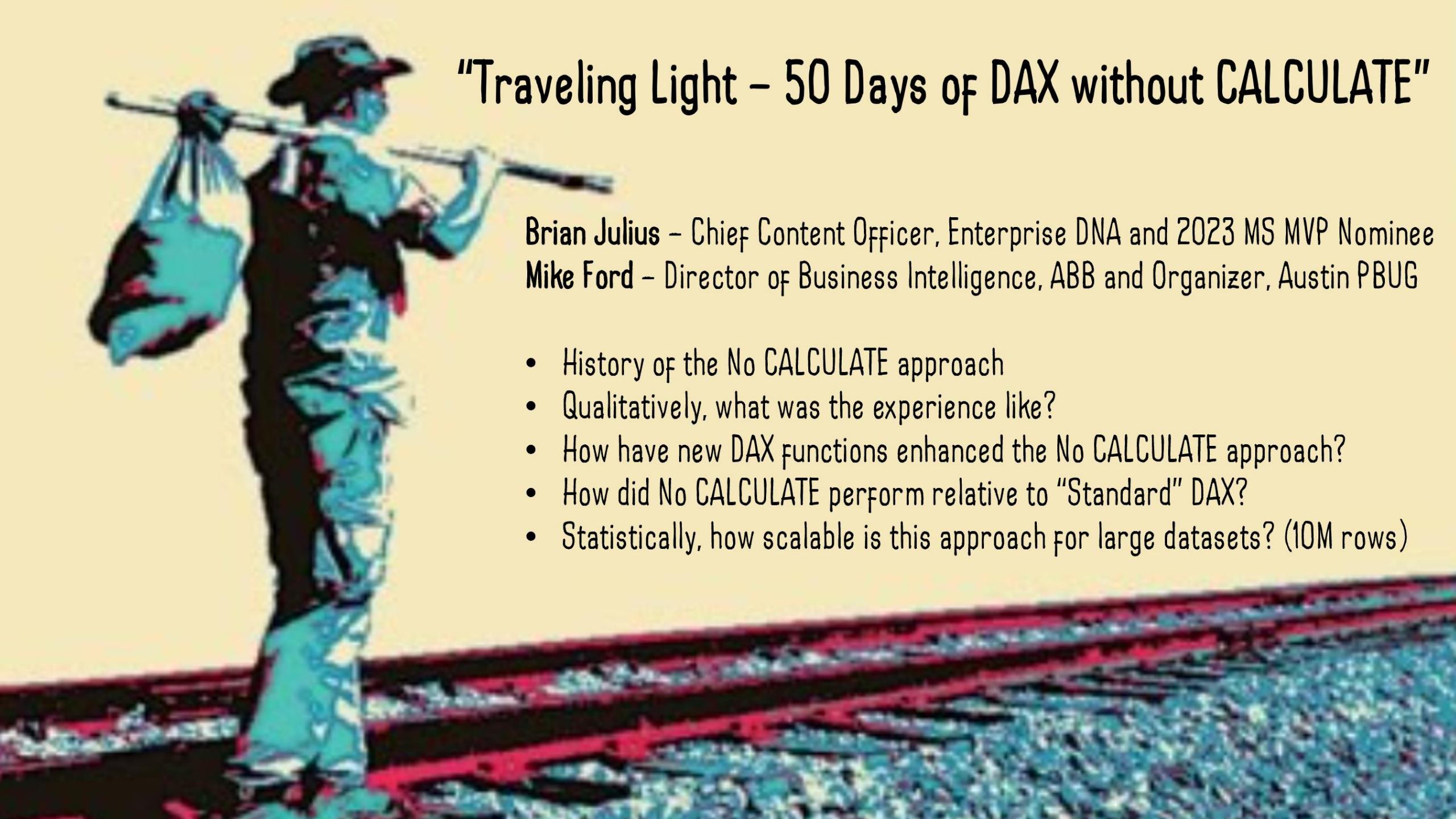


Welcome!

**Austin Power BI User
Group (PUG)**

www.pbiusergroup.com/Austin

@Power_BI_ATX



"Traveling Light - 50 Days of DAX without CALCULATE"

Brian Julius - Chief Content Officer, Enterprise DNA and 2023 MS MVP Nominee
Mike Ford - Director of Business Intelligence, ABB and Organizer, Austin PBUG

- History of the No CALCULATE approach
- Qualitatively, what was the experience like?
- How have new DAX functions enhanced the No CALCULATE approach?
- How did No CALCULATE perform relative to "Standard" DAX?
- Statistically, how scalable is this approach for large datasets? (10M rows)

**SILLY MEME,
SERIOUS QUESTION...**

PowerBIUG

Learning DAX



Basic DAX
starts making
sense

Learning
Advanced
DAX

 You and 949 others



75 comments • 22 shares

ADDRESSING THE MYTHS AND MISCONCEPTIONS

PowerBIUG

Not trying to convert anyone here

- If CALCULATE is just fine for you now – rock on...

Not the refuge of lazy simpletons who can't/won't learn DAX

Does not remove the need to learn foundational concepts – evaluation/filter/row context, context transition, etc.

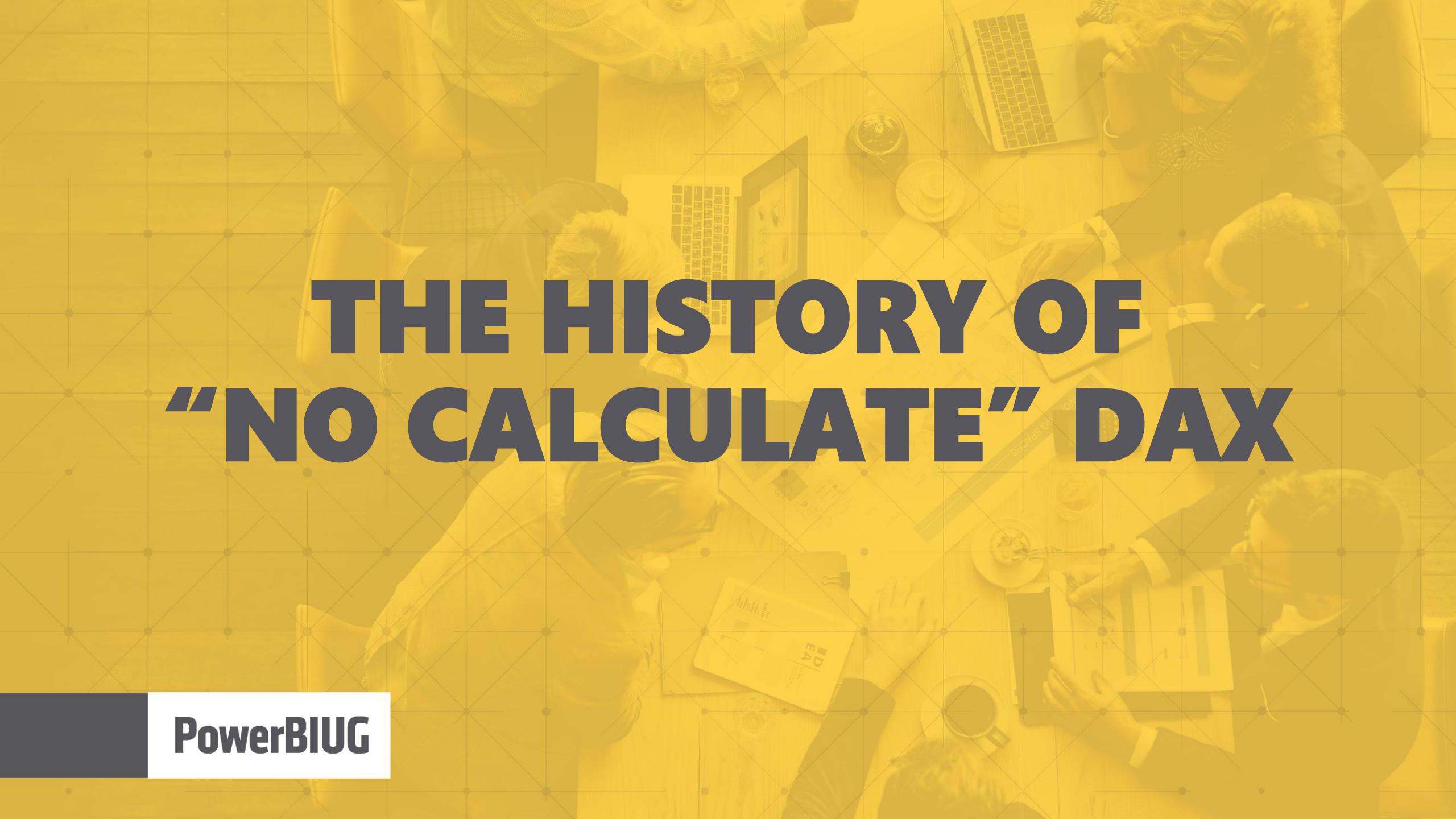
- In fact, makes some of these concepts EASIER to learn IMO

Not an absolute

- You don't get kicked out of the club for using CALCULATE

Not limited to those using small and simple datasets

- MUCH more on this later

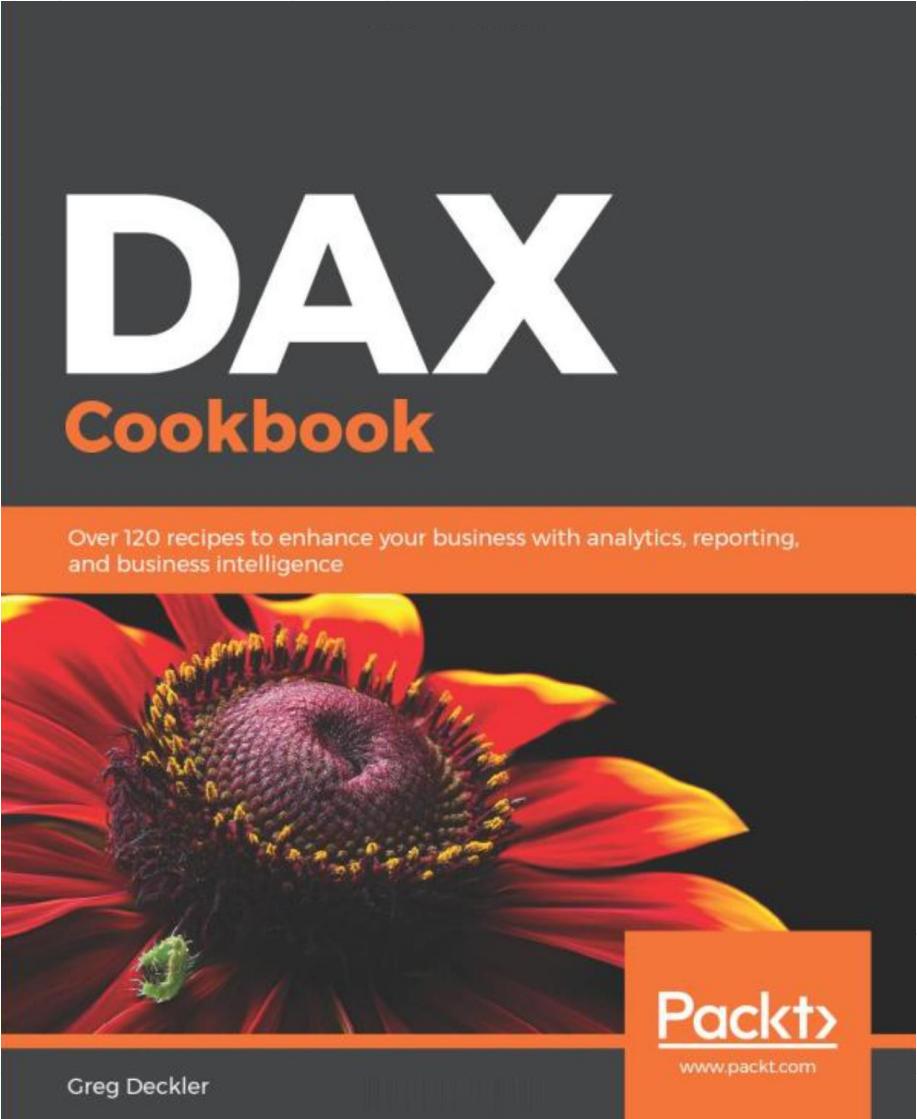
A collage of various people working on laptops, tablets, and phones, with a network grid overlay.

THE HISTORY OF “NO CALCULATE” DAX

PowerBIUG

THE FIRST RUMBLINGS (MARCH 2020)

PowerBIUG



You purchased this item on July 1, 2020.
[View this order](#)

THE BIG BANG (JULY 24, 2020)

PowerBIUG

Microsoft Power BI Community > Community Blog > Power BI Community Blog

CALCUHATE - Why I Don't Use DAX's CALCULATE Function

Article Options ▾



Greg_Deckler 07-24-2020 11:13 AM

CALCUHATE - Why I Don't Use DAX's CALCULATE Function

⋮

Introduction

OK, first off, I don't hate, hate the CALCULATE function. At least not to the same level that I despise the time "intelligence" functions. I mean, I realize that I have made reference to DAX's time "intelligence" functions being one of the four horsemen of the apocalypse but just so we are 100% clear, I truly despise them. But that's for another blog article.

So, I do not hate DAX's CALCULATE to anywhere near that level but it made for a catchy title for a blog post. In reality, I honestly just don't find myself thinking about CALCULATE most of the time any longer. Over the 5 or 6 years that I have been coding DAX I have simply developed a coding style and technique where CALCULATE is just generally useless. So let me explain how I got there and then I would love to hear everyone else's thoughts on this subject because I think it is an interesting topic with a lot of far reaching implications. In reality, the use of CALCULATE or not CALCULATE really represents two wildly different programming approaches and philosophies around DAX coding.

Variables

IS THIS SOMETHING?...

PowerBIUG



BrianJ Enterprise DNA Expert

Jul '20

Forum Members,

I'm a big fan of [Greg Deckler's work](#), and on Friday he posted an incredibly interesting and thought-provoking piece entitled "CALCUHATE - Why I Don't Use DAX's CALCULATE Function".

community.powerbi.com – 24 Jul 20

[CALCUHATE - Why I Don't Use DAX's CALCULATE Function](#) 36

Introduction OK, first off, I don't hate, hate the CALCULATE function. At least not to the same level that I despise the time "intelligence" functions. I mean, I realize that I have made reference to DAX's time "intelligence" functions being one of...

I initially thought this was a joke, since I couldn't imagine how you could do anything useful in DAX without CALCULATE, but after reading through this twice now, I think he makes some very compelling points. Having just spent the better part of two days debugging a really complex report with layers and layers of branched measures, his points really resonated with me. I'm not ready to throw CALCULATE on the scrap heap just yet, but this post definitely got my wheels turning.

Will be very interested to hear what you all think.

• Brian

IN THE INTERIM... (DECEMBER 2021)

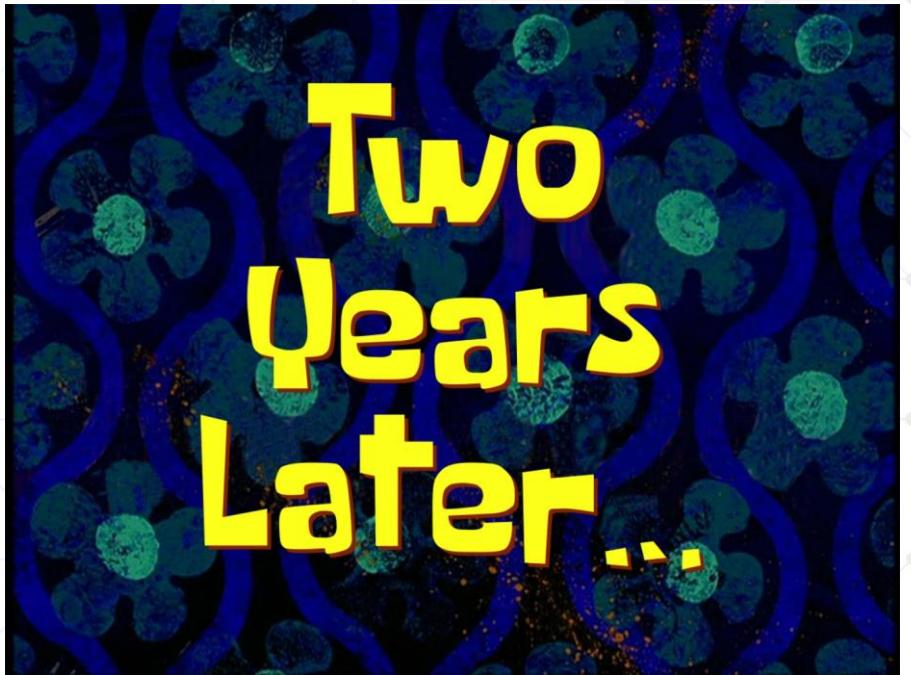
PowerBIUG

curbal

25 DAYS OF
DAX FRIDAYS!

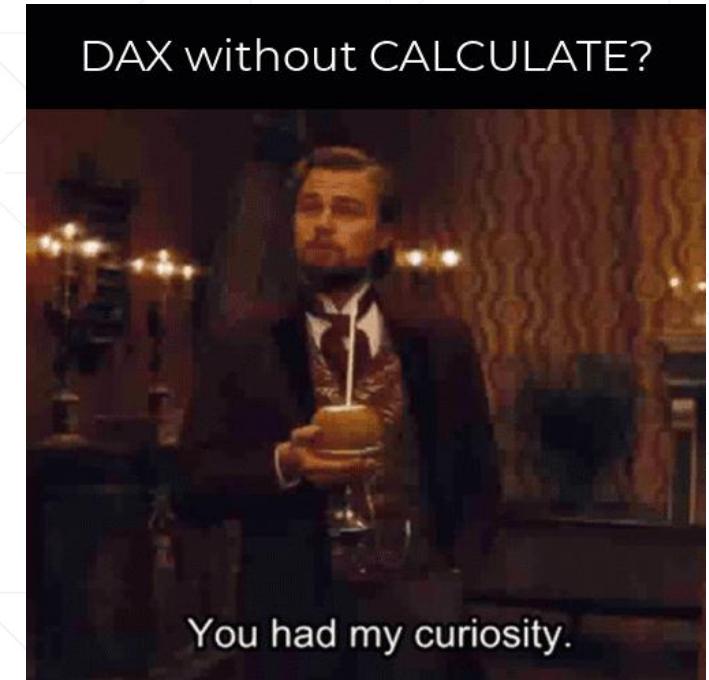
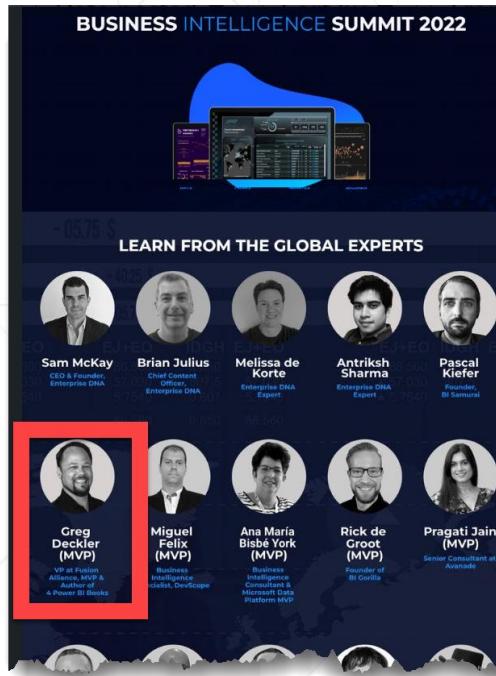
37 Day 1	Côte de Blaye Day 2	28.87 Day 3	25 Day 4	25 Day 5
61.83 Day 6	\$27,444 Day 7	137 Day 8	293.46 Day 9	417 Day 10
1 Day 11	(Blank) Day 12	89 Day 13	59 Day 14	5 Day 15
5 Day 16	18 Day 17	14 Day 18	\$4,453 Day 19	Aux joyeux éclésiastiques Day 20
56% Day 21	8 Day 22	Janet Leverling Day 23	4 Day 24	3 Day 25

25 Days Of DAX Fridays! Challenge
From: Curbal
To: PBI Enthusiasts
<https://25 Days of DAX Challenge>



PowerBIUG

- SEPTEMBER 2022 – Greg Deckler's presentation at Enterprise DNA Global Summit on "DAX Counterculture"



SEP 17, 2022

 Brian Julius (He/Him) • 12:40 PM

Mike,

Hope you're having a good weekend. I have a quick favor to ask ...

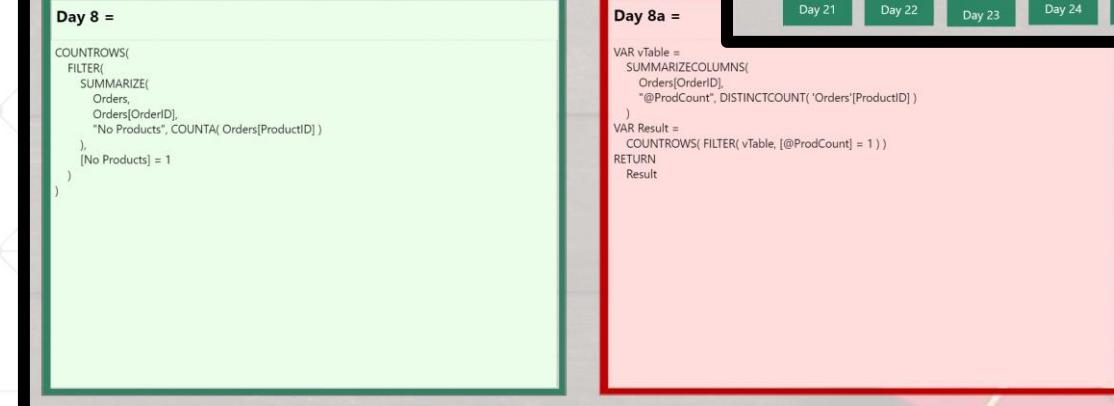
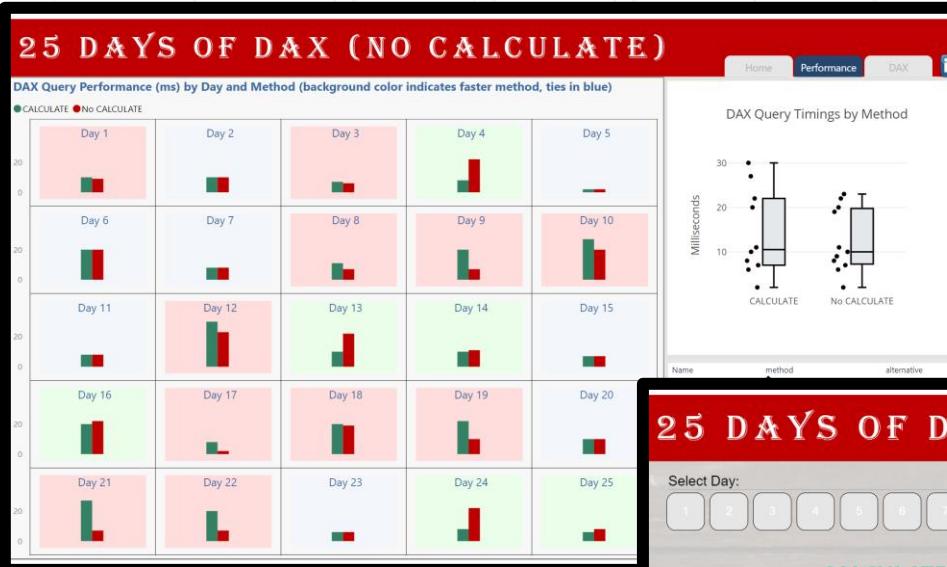
I seem to recall you put together a really nice report in response to Ruth's 25 Days of DAX last year. Would you be willing to share the PBIX with me?

SEP 18, 2022

 Mike Ford • 1:33 AM

Hi Brian. Absolutely! I would be happy to share it with you. Just let me know the best way to send and I'll do it.

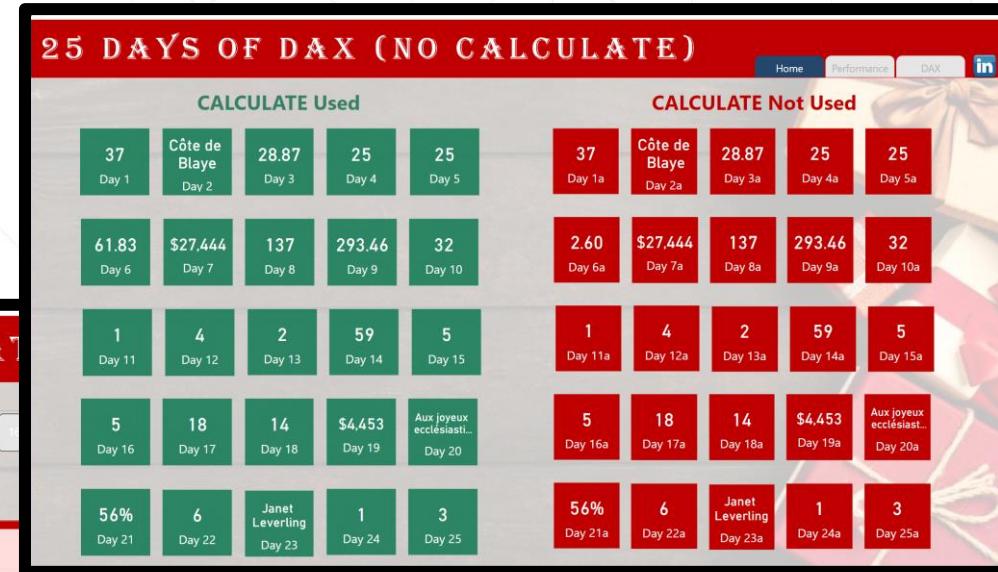
"NO CALCULATE SEPTEMBER"



Two initiatives:

Complete "cold turkey" on stopping all use of CALCULATE for the entire month of September

Re-doing Ruth Pozuelo Martinez's (Curbal) 25 Days of DAX Advent Calendar exercise, but with no CALCULATE (or TI) functions



<https://www.novapro.com/project/brianj25days>

PowerBIUG

"NO CALCULATE" REDUX

- WORLD CUP EDITION

(NOVEMBER 2022)



Huge thanks to Cristian Angyal for allowing us to use his base report!

25 Days Of DAX Fridays! Challenge (Ed 2) by #curbal

<https://curbal.com/25-days-of-dax-fridays-challenge-ed2-world-cup-data>

Day 1 Brazil had most tournaments won	Day 2 Argentina, England, France, Italy, Uruguay, West Germany Host countries that won	Day 3 12 years Longest gap in years	Day 4 Brazil, France, Italy, Mexico highest no. hosted tournaments	Day 5 Argentina, Netherlands, West Germany Most second place finishes
Day 6 Brazil Most WC appearances	Day 7 Brazil Most finished in top two	Day 8 44 years Longest gap in years bet. titles	Day 9 Brazil, Most consecutive WC won	Day 10
Day 11 Antonio Carbajal, Gianluigi Buffon, Lothar Matthäus, Rafael Márquez Players with most tournaments (5)	Day 12 Lothar Matthäus Player with most matches at WC (25)	Day 13 Norman Whiteside Youngest Player	Day 14 Giuseppe Bergomi Youngest Player in Final	Day 15
Day 16 Rafael Márquez Most tournaments as captain	Day 17 Miroslav Klose Most goals scored (16)	Day 18 Hungary v El Salvador, Hungary v South Korea, Yugoslavia v Brazil, Yugoslavia v Mexico Matches with biggest goal margin	Day 19 10 goals Most goals by one player	Day 20
Day 21 England v Belgium, Soviet Union v Colombia Matches with highest scoring draw	Day 22 Brazil v France, Brazil v Italy, Brazil v Sweden Finals with biggest score margin	Day 23 171 goals, 70 goals Most and least goals in a tournament	Day 24 2018 FIFA World Cup	Day 25

PowerBIUG

#25daysofdaxfridays - Edition 2

curbal

25 days of DAX FRIDAYS!

Leaderboard

25 Days Of DAX Fridays! Challenge (Ed 2) by #curbal

<https://curbal.com/25-days-of-dax-fridays-challenge-ed2-world-cup-data>

Day 1 Brazil had most tournaments won	Day 2 West Germany, Uruguay, Italy, France, England, Argentina Host countries that won	Day 3 12 years Longest gap in years	Day 4 Brazil, France, Italy, Mexico highest no. hosted tournaments	Day 5 Argentina, Netherlands, West Germany Most second place finishes
Day 6 Brazil Most WC appearances	Day 7 Brazil Most finished in top two	Day 8 44 years Longest gap in years bet. titles	Day 9 Brazil, Most consecutive WC won	Day 10
Day 11 Antonio Carbajal, Gianluigi Buffon, Lothar Matthäus, Rafael Márquez Players with most tournaments (5)	Day 12 Lothar Matthäus Player with most matches at WC (25)	Day 13 Norman Whiteside Youngest Player	Day 14 Giuseppe Bergomi Youngest Player in Final	Day 15 Essam El Hadary Oldest Captain
Day 16 Rafael Márquez Most tournaments as captain	Day 17 Miroslav Klose Most goals scored (16)	Day 18 Hungary v El Salvador, Hungary v South Korea, Yugoslavia v Brazil, Yugoslavia v Mexico Matches with biggest goal margin	Day 19 10 goals Most goals by one player	Day 20 Austria v Switzerland Day 20 NC
Day 21 England v Belgium, Soviet Union v Colombia Matches with highest scoring draw	Day 22 Brazil v France, Brazil v Italy, Brazil v Sweden Finals with biggest score margin	Day 23 171 goals, 70 goals Most and least goals in a tournament	Day 24 2018 FIFA World Cup	Day 25 5.38 goals Day 25 NC

Last Refresh (UTC) = 11/12/2022 9:50:20 AM

WHY “25 DAYS OF DAX” AS BASIS FOR EXPERIMENT?

Highly recognizable, lots of community participation

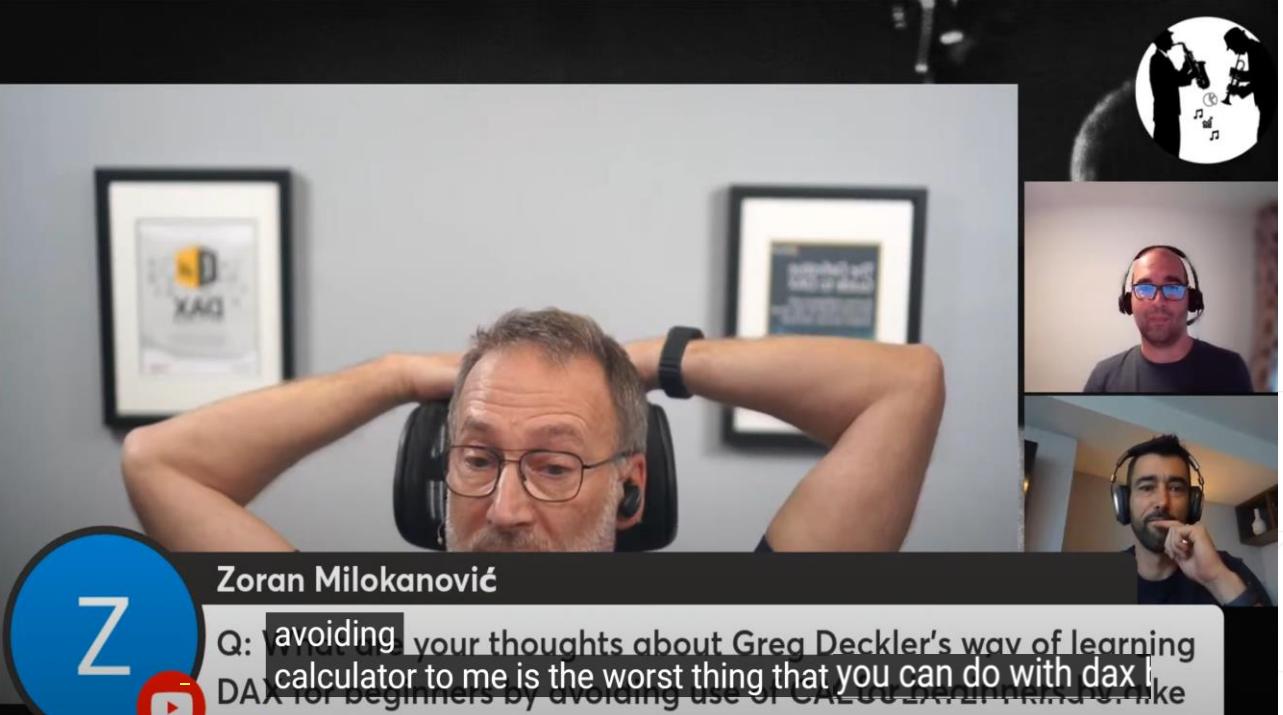
“Fair test” – not jury-rigged to produce any particular result

Easy to validate accuracy of measure results

Wide range of measures, many reflective of common business scenarios/questions

Great initiative, worth promoting

NOT EVERYONE IS A FAN...

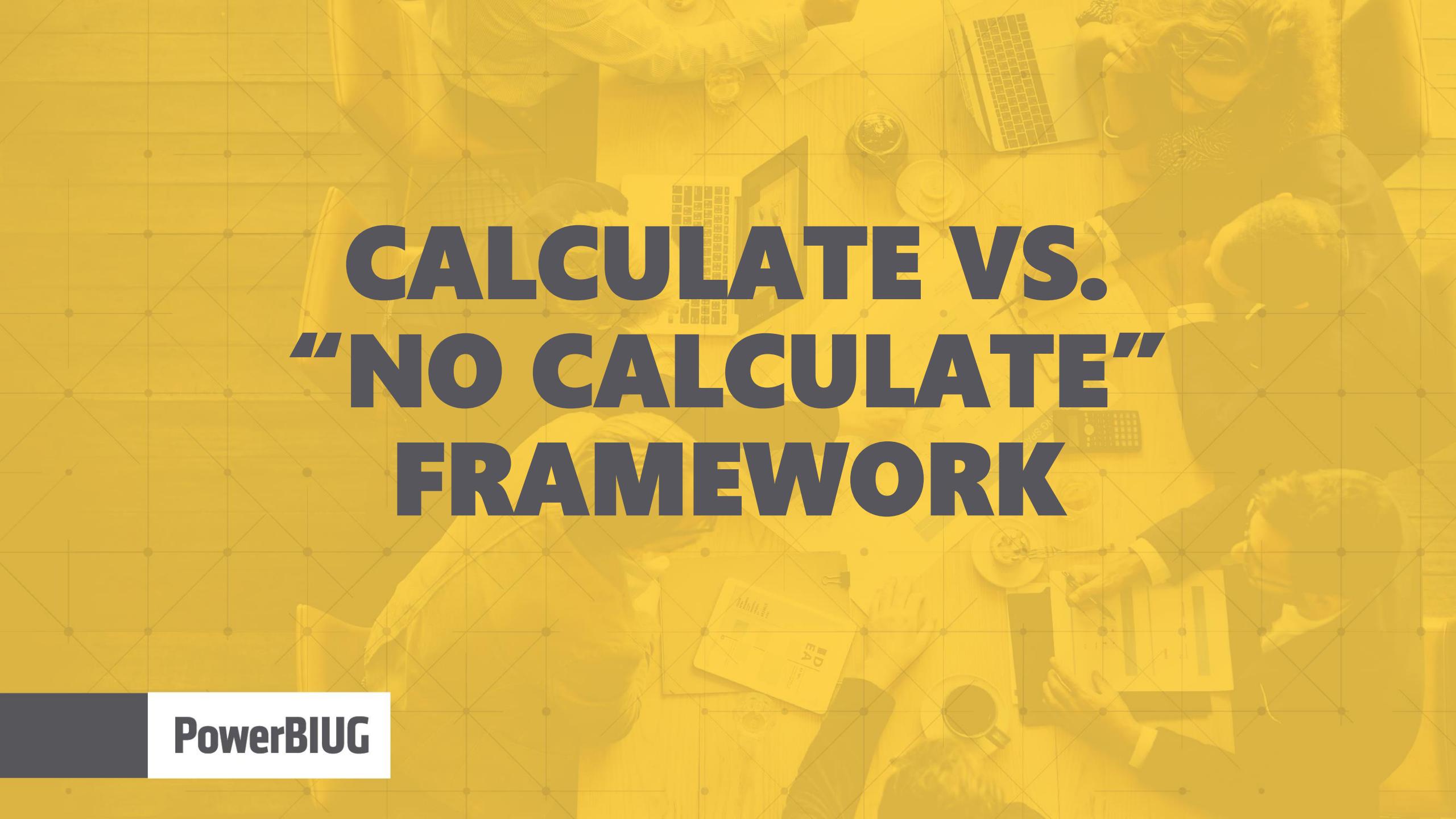


Alberto Ferrari on Data Jazzers podcast

To be abundantly clear, the no-`CALCULATE` approach is incorrect and will manifest performance pathologies in larger models. I am not endorsing the approach, but I must admit that casual dismissiveness is an unhelpful response. In fact, there are useful things to learn from the movement, just not any of its actual DAX.

PowerBIUG

Greg Baldini blog
<https://www.antifound.com/posts/on-dax-and-calculate/>

A collage of various business-related images including people working at desks, a globe, charts, and a keyboard.

CALCULATE VS. “NO CALCULATE” FRAMEWORK

PowerBIUG

CALCULATE VS NO CALCULATE FRAMEWORK

Calculate - General pattern of measures using Calculate is they are evaluated "Outside In"

- Calculations are based on arguments that are nested together inside the Calculate function

No Calculate - General pattern of measures that are written using the No Calculate method is they follow a sequential order

- Table variables
- X aggregation functions
- Return results

25 DAYS OF WORLD CUP EXAMPLE – DAY 4

```
Day 04 = VAR SummarizedTable =
    SUMMARIZECOLUMNS(
        'Tournaments'[host_country],
        "Count",
        COUNTROWS( 'Tournaments' )
    )
    VAR MaxWins =
        MAXX(
            SummarizedTable,
            [Count]
        )
    VAR MaxWinList =
        FILTER(
            SummarizedTable,
            [Count] = 2
        )
    VAR Result =
        CALCULATE(
            CONCATENATEX(
                VALUES(
                    'Tournaments'[host_country]
                ),
                'Tournaments'[host_country],
                ", ",
                'Tournaments'[host_country], ASC
            ),
            MaxWinList
        )
    RETURN
    Result
```

No Calculate Method

```
Day 04 NC = VAR _vTable =
    ADDCOLUMNS(
        DISTINCT(
            'Tournaments'[host_country]
        ),
        "@hostings", [Row Count Tournaments]
    )
    VAR _Result =
        CONCATENATEX(
            TOPN(
                1,
                _vTable,
                [@hostings], DESC
            ),
            'Tournaments'[host_country],
            ", ",
            'Tournaments'[host_country], ASC
        )
    RETURN
    _Result
```

- Table variable

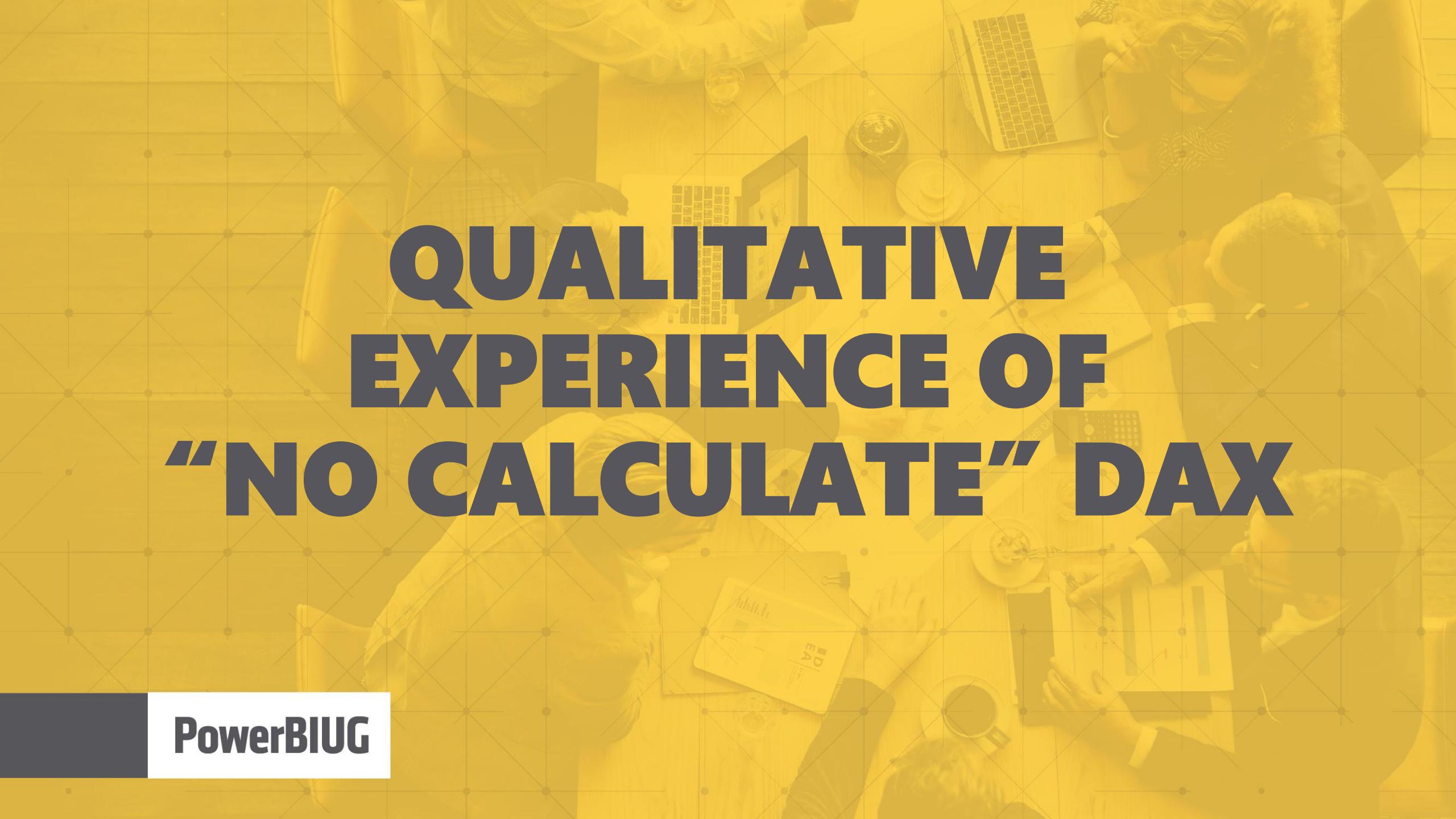
- X aggregation function

- Results

Results

Day 4
Brazil, France, Italy,
Mexico
highest no. hosted tournaments

- Both methods provide the same result



QUALITATIVE EXPERIENCE OF “NO CALCULATE” DAX

PowerBIUG

BEST ARGUMENT IN FAVOR OF “NO CALCULATE” DAX FROM ONE OF ITS MOST VOCAL CRITICS

PowerBIUG

I applaud the marketing acumen on display in this movement:

- "DAX Counterculture"
- "Ignore the complexity of CALCULATE "
- "Make your DAX easier to understand"

These are great bits of marketing and you can see this signature across all the no-CALCULATE writings that pop up, regardless of author. The proponents of this approach have done a great job of crafting a message that resonates and largely self-perpetuates.

And you know what? I agree that a series of tables makes for an understandable computational model. You know who else thought that was a good idea? The Power Query folks. Power Query generates a series of steps where each step is a table representing a partial result. These partial results are refined until they add up to the result table we want. Doing the same thing in DAX lets you keep the same mental model of calculation. It is a good mental model of analytical computation!

And ultimately, the no-CALCULATE movement has probably helped more people solve actual problems in DAX than I have in my career. Because it is more approachable to a novice. It allows a more imperative style, which is usually a more intuitively understandable programming paradigm than functional. This is especially true when the language is DAX, a functional and relational language with a concept of filter context that does not exist in the vast majority of programming languages. And it matches the way you should think about transformations in Power Query, the other data engine that goes with Tabular. And it does not matter for so very many models that will never grow beyond a few hundred thousand rows of fact data.

But it is incorrect.

Greg Baldini blog
<https://www.antifound.com/posts/on-dax-and-calculate/>

LEARNINGS FROM TRAVELING WITHOUT CALCULATE (BRIAN)

- It takes a little time to overcome the “muscle memory” of using CALCULATE
- Once I got used to it, it felt very natural, intuitive and in line with the linear way in which I think – much more natural to me than typical “outside-in” DAX
- Mimics what I like about working in Power Query
- There was NEVER a point the entire month when I was unable to accomplish what I needed to without the use of CALCULATE
- Writing DAX becomes very formulaic – pretty much every measure gets handled the same basic way
- Much easier to visualize and debug
- Most surprisingly – In the majority of cases, “No CALCULATE” DAX performed equally to or better than “traditional” DAX (statistically not significant difference between query timings)

LEARNINGS FROM TRAVELING WITHOUT CALCULATE (MIKE)

The sequential order of the No Calculate framework may appeal to people with a SQL background since SQL is written using a similar pattern which might make it more intuitive

```
SELECT AlbumName, ArtistName, ReleaseDate  
  
FROM Albums  
  
    INNER JOIN Artists  
  
        ON Albums.ArtistId =  
  
Artists.ArtistId  
  
    INNER JOIN Genres  
  
        ON Albums.GenreId =  
  
Genres.GenreId  
  
WHERE Genres.Genre = 'Rock'
```

- Select
- From
- Where

```
Day 04 NC = VAR _vTable =  
ADDCOLUMNS(  
DISTINCT(  
    'Tournaments'[host_country]  
,  
    "@hostings", [Row Count Tournaments]  
)  
VAR _Result =  
CONCATENATEX(  
TOPN(  
    1,  
    _vTable,  
    [@hostings], DESC  
,  
    'Tournaments'[host_country],  
    ", ",  
    'Tournaments'[host_country], ASC  
)  
RETURN  
_Result
```

- Table variable
- Iterator
- Function
- Result



COOL STORIES, BRO NOW SHOW ME THE DATA

PowerBIUG

25 DAYS OF DAX ADVENT CALENDAR

Calculate and No Calculate examples are based on results from the DAX Fridays Advent Calendar designed by Curbal.

Remember to refresh the dataset before answering :)



Day 1: How many current products cost less than \$20?	Day 2: Which product is the most expensive?	Day 3: What is the average unit price for our products?	Day 4: How many products are above the average unit price?	Day 5: How many products cost between \$15 and \$25? (inclusive)
Day 6: What is the average number of products (not qty) per order?	Day 7: What is the order value in \$ of open orders? (Not shipped yet)	Day 8: How many orders are "single item" (only one product ordered)?	Day 9: Average sales per transaction (orderID) for "Romero y Tomillo"	Day 10: How many days since "North/South" last purchase?
Day 11: How many customers have ordered only once?	Day 12: How many New customers (first purchase in current year) in 2022?	Day 13: How many Lost customers (no purchases in current year) in 2022?	Day 14: How many customers have NEVER purchased Queso Cabrales?	Day 15: How many customers have purchased only Queso Cabrales (per OrderID)?
Day 16: How many products are out of stock?	Day 17: How many products need to be restocked? (based on restock levels)	Day 18: How many products on order we need to restock?	Day 19: What is the stocked value of the discontinued products?	Day 20: Which vendor has the highest stock value?
Day 21: How many employees (%) are female?	Day 22: How many employees are 60 years old or over?	Day 23: Which employee had the highest sales in 2022?	Day 24: How many employees sold over \$100k in 2022?	Day 25: How many employees got hired in 1994?

All DAX solutions

PowerBIUG

THE CALCULATE AND NO CALCULATE RESULTS

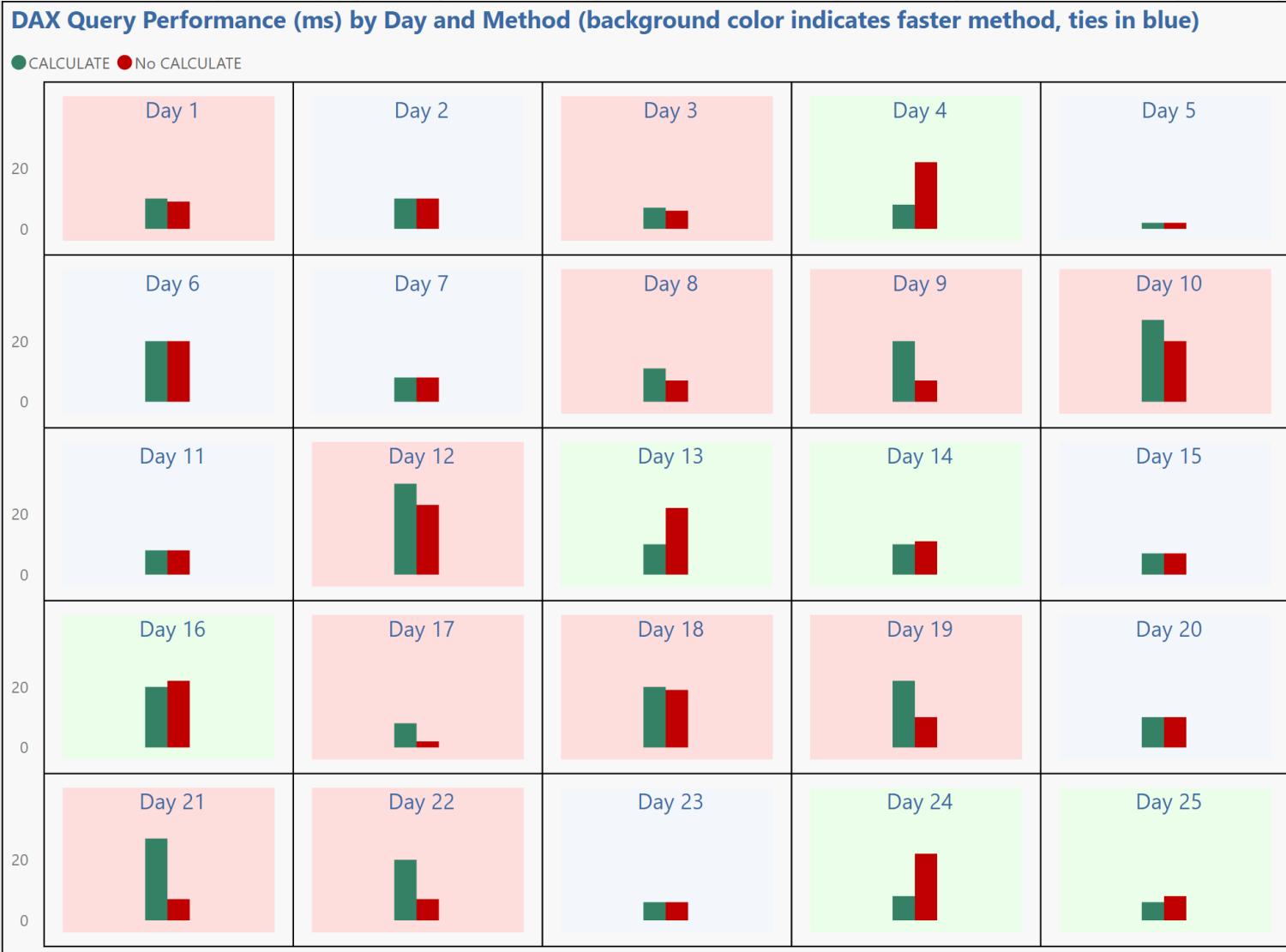
25 DAYS OF DAX (NO CALCULATE)									
CALCULATE Used					CALCULATE Not Used				
37 Day 1	Côte de Blaye Day 2	28.87 Day 3	25 Day 4	25 Day 5	37 Day 1a	Côte de Blaye Day 2a	28.87 Day 3a	25 Day 4a	25 Day 5a
61.83 Day 6	\$27,444 Day 7	137 Day 8	293.46 Day 9	7 Day 10	61.83 Day 6a	\$27,444 Day 7a	137 Day 8a	293.46 Day 9a	7 Day 10a
1 Day 11	(Blank) Day 12	1 Day 13	59 Day 14	5 Day 15	1 Day 11a	(Blank) Day 12a	3 Day 13a	59 Day 14a	5 Day 15a
5 Day 16	18 Day 17	14 Day 18	\$4,453 Day 19	Aux joyeux ecclésiasti... Day 20	5 Day 16a	18 Day 17a	14 Day 18a	\$4,453 Day 19a	Aux joyeux ecclésiasti... Day 20a
56% Day 21	8 Day 22	Janet Leverling Day 23	(Blank) Day 24	3 Day 25	56% Day 21a	6 Day 22a	Janet Leverling Day 23a	(Blank) Day 24a	3 Day 25a

Note: the rollover to the new year broke measure 12 in both versions, but the performance results were generated prior to the year rollover.

THE CALCULATE AND NO CALCULATE PERFORMANCE RESULTS

(SMALL MULTIPLES)

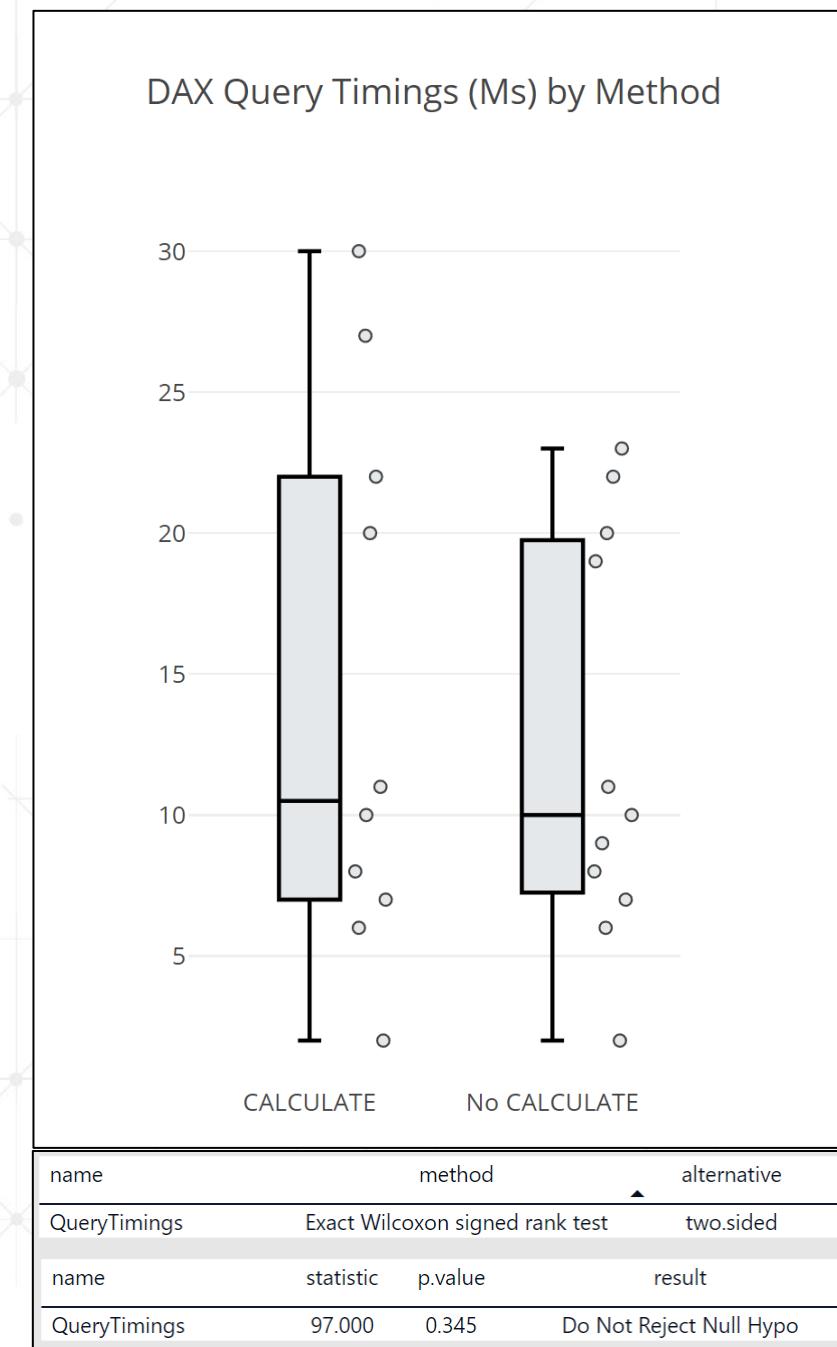
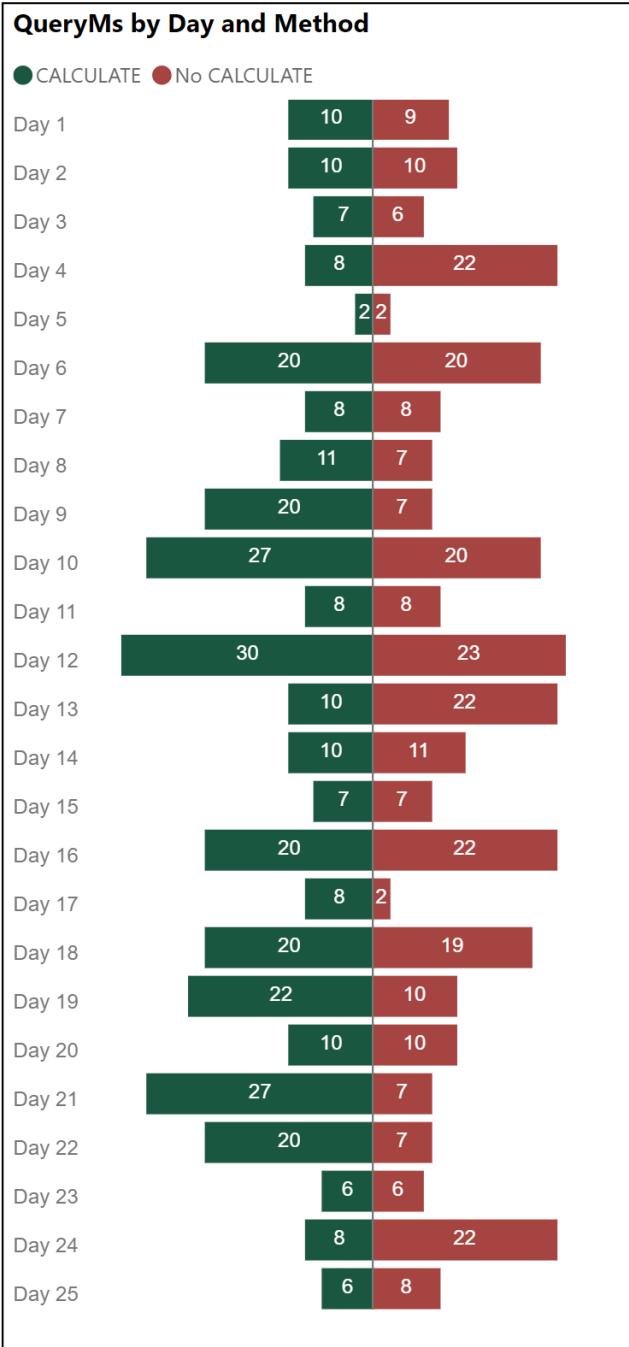
PowerBIUG



STATISTICAL ANALYSIS OF CALCULATE VS “NO CALCULATE” PERFORMANCE

TL;DR: No evidence to support difference in performance between the two methods.

PowerBIUG



25 DAYS OF DAX WORLD CUP CHALLENGE

Calculate and No Calculate examples are based on results from the DAX Fridays World Cup Challenge designed by Curbal. Link to the challenge can be found [HERE](#)

#25daysofdaxfridays - Edition 2

curbal

Day 1: Country/s with most tournaments won	Day 2: Host country/s that won	Day 3: Longest gap in years between tournaments	Day 4: Country(s) with highest number of hosted tournaments	Day 5: Country/s with most second place finnishes
Day 6: Country/s with most world cup appearances	Day 7: Country/s with most finnishes in the top two	Day 8: Longest gap in years between titles	Day 9: Country/s with most consecutive championship wins	Day 10: Country/s with most finals played but never lost
Day 11: Players with most tournaments played	Day 12: Player/s with most matches played	Day 13: Youngest player/s when they first played	Day 14: Youngest player/s on a final	Day 15: Oldest captain/s
Day 16: Player/s with most tournaments as captain	Day 17: Player/s with Most goals scored	Day 18: Match/s with biggest goal margin	Day 19: Most goals scored in a match for one team	Day 20: Match/es with most goals scored for both teams
Day 21: Match/es with highest scoring draw	Day 22: Match/es with biggest goal margin in a final	Day 23: Most and least goals in a tournament	Day 24: Tournament with highest number of own goals	Day 25: Highest average goals per match

Leaderboard

THE CALCULATE AND NO CALCULATE RESULTS

PowerBIUG

25 Days Of DAX Fridays! Challenge (Ed 2) by #curbal

<https://curbal.com/25-days-of-dax-fridays-challenge-ed2-world-cup-data>

 @cristian_angyal

Last Refresh (UTC) = 11/12/2022 9:50:20 AM

Day 1	Day 2	Day 3	Day 4	Day 5
Brazil had most tournaments won	Argentina, England, France, Italy, Uruguay, West Germany Host countries that won	12 years Longest gap in years	Brazil, France, Italy, Mexico highest no. hosted tournaments	Argentina, Netherlands, West Germany Most second place finishes
Brazil Most WC appearances	Brazil Most finished in top two	44 years Longest gap in years bet. titles	Brazil, Italy Most consecutive WC won	Uruguay Most finals won never lost
Antonio Carbajal, Gianluigi Buffon, Lothar Matthäus, Rafael Márquez Players with most tournaments (5)	Lothar Matthäus Player with most matches at WC (25)	Norman Whiteside Youngest Player	Giuseppe Bergomi Youngest Player in a Final	Essam El Hadary Oldest Captain
Rafael Márquez Most tournaments as captain	Miroslav Klose Most goals scored (16)	Hungary v El Salvador, Hungary v South Korea, Yugoslavia... Matches with highest goal margin	10 goals Most goals by one team in a match	Austria v Switzerland Match with most goals for both teams
England v Belgium, Soviet Union v Colombia Matches with highest scoring draw	Brazil v France, Brazil v Italy, Brazil v Sweden Finals with biggest score margin	171 goals, 70 goals Most and least goals in a tournament	2018 FIFA World Cup Tournament with most Own Goals	5.38 goals Highest Avg goals per match
Last Refresh (UTC) = 11/12/2022 9:50:20 AM				

25 Days Of DAX Fridays! Challenge (Ed 2) by #curbal

<https://curbal.com/25-days-of-dax-fridays-challenge-ed2-world-cup-data>

 Brian Julius

Last Refresh (UTC) = 11/12/2022 9:50:20 AM

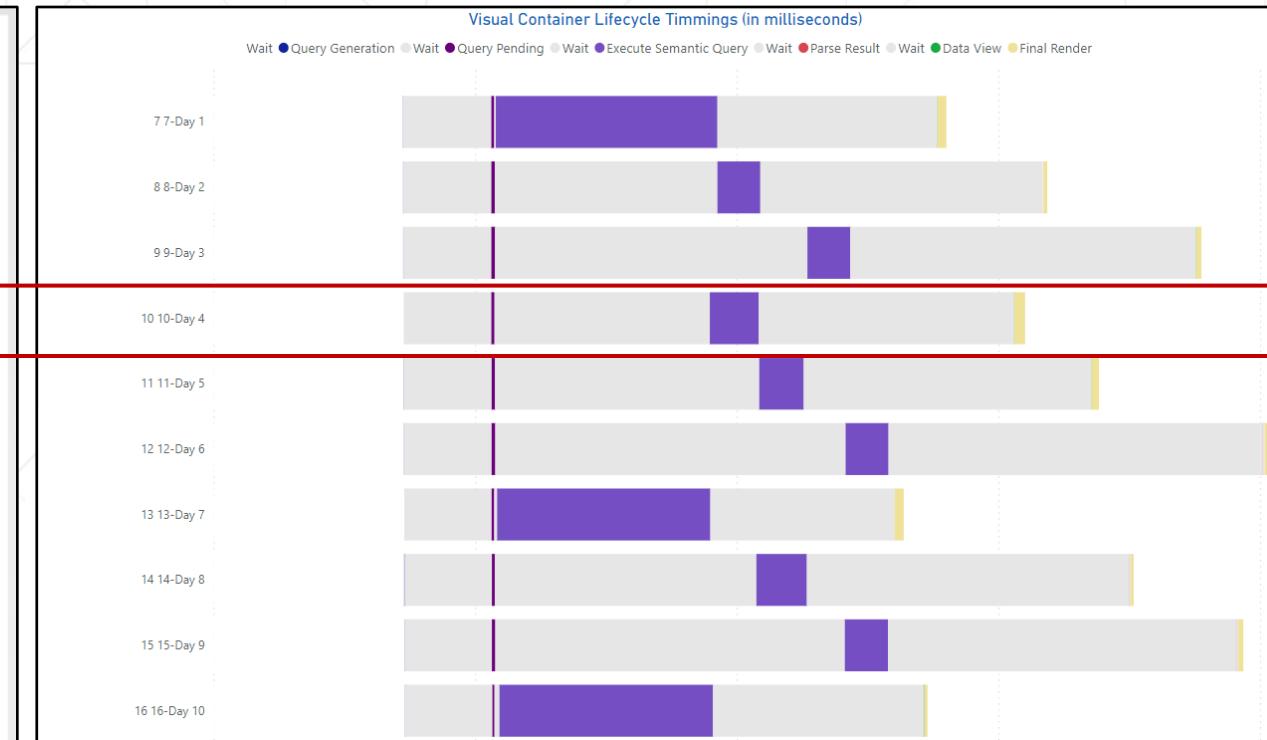
Day 1	Day 2	Day 3	Day 4	Day 5
Brazil had most tournaments won	West Germany, Uruguay, Italy, France, England, Argentina Host countries that won	12 years Longest gap in years	Brazil, France, Italy, Mexico highest no. hosted tournaments	Argentina, Netherlands, West Germany Most second place finishes
Brazil Most WC appearances	Brazil Most finished in top two	44 years Longest gap in years bet. titles	Brazil, Italy Most consecutive WC won	Uruguay Day 10 NC
Antonio Carbajal, Gianluigi Buffon, Lothar Matthäus, Rafael Márquez Players with most tournaments	Lothar Matthäus Day 12 NC	Norman Whiteside Youngest Player	Giuseppe Bergomi Youngest Player in Final	Essam El Hadary Oldest Captain
Rafael Márquez Day 16 NC	Miroslav Klose Day 17 NC	Hungary v South Korea, Yugoslavia v Zaire, Hungary v El Salvador Day 18 NC	10 goals Day 19 NC	Austria v Switzerland Day 20 NC
England v Belgium, Soviet Union v Colombia Day 21 NC	Brazil v Sweden, Brazil v Italy, Brazil v France Day 22 NC	171 goals, 70 goals Day 23 NC	2018 FIFA World Cup Day 24 NC	5.38 goals Day 25 NC
Last Refresh (UTC) = 11/12/2022 9:50:20 AM				

DAX WORLD CUP REPORT RESULTS

Calculate Results – Measures for the first 10 days of challenge



No Calculate Results – Measures for the first 10 days of challenge

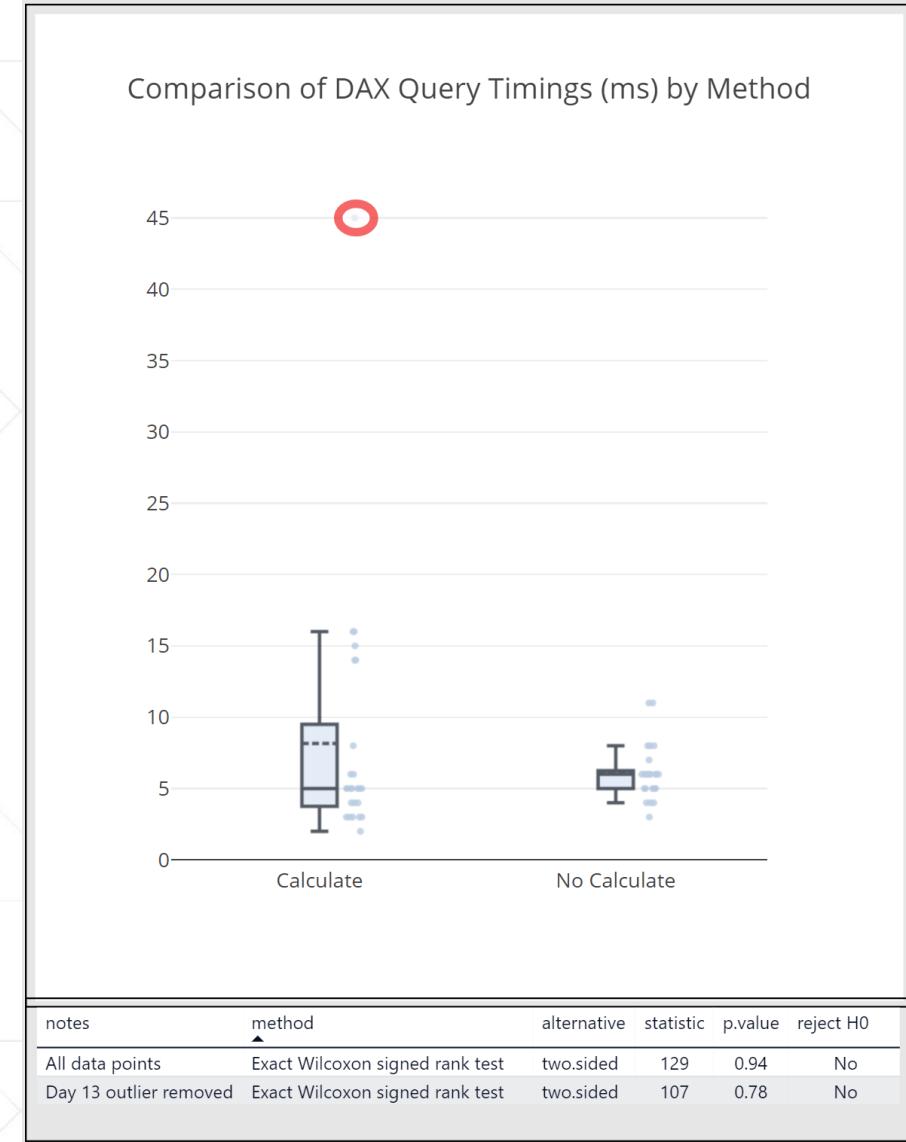
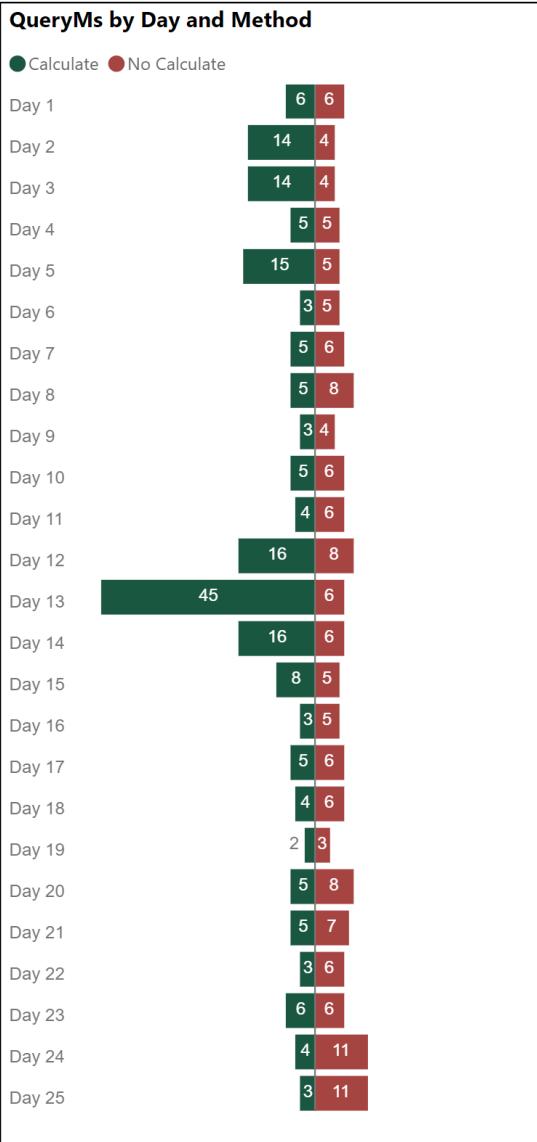


- Utilizing the Power BI Advanced Performance Analyser by SmartPowerBI. It was created by Alex J. Garcia. [Link](#) to blog article with GitHub repository and instructions.
- Show results from the 25 Days of World Cup report using measures that were written using Calculate and equivalent measures written with No Calculate framework
- Query execution time for each measure that was written for each day of the challenge is shown in Purple
- Many instances where No Calculate measures ran faster. The Day 4 measure is outlined above.
 - Calculate Measure - Ran in 1,070 milliseconds
 - No Calculate Measure – Ran in 187 milliseconds

STATISTICAL ANALYSIS OF CALCULATE VS “NO CALCULATE” PERFORMANCE

TL;DR: No evidence to support difference in performance between the two methods.

PowerBIUG





BUT DOES IT SCALE?

PowerBIUG

WHAT IS A “LARGE” DATASET IN POWER BI?

PowerBIUG



Brian Julius (He/Him) • You
Lifelong data geek | Data Science/Econometrics | Passionate about creating, d...
3mo • 

Power BI Users - for some research I'm doing, I would really appreciate your response to the following question.
[...see more](#)

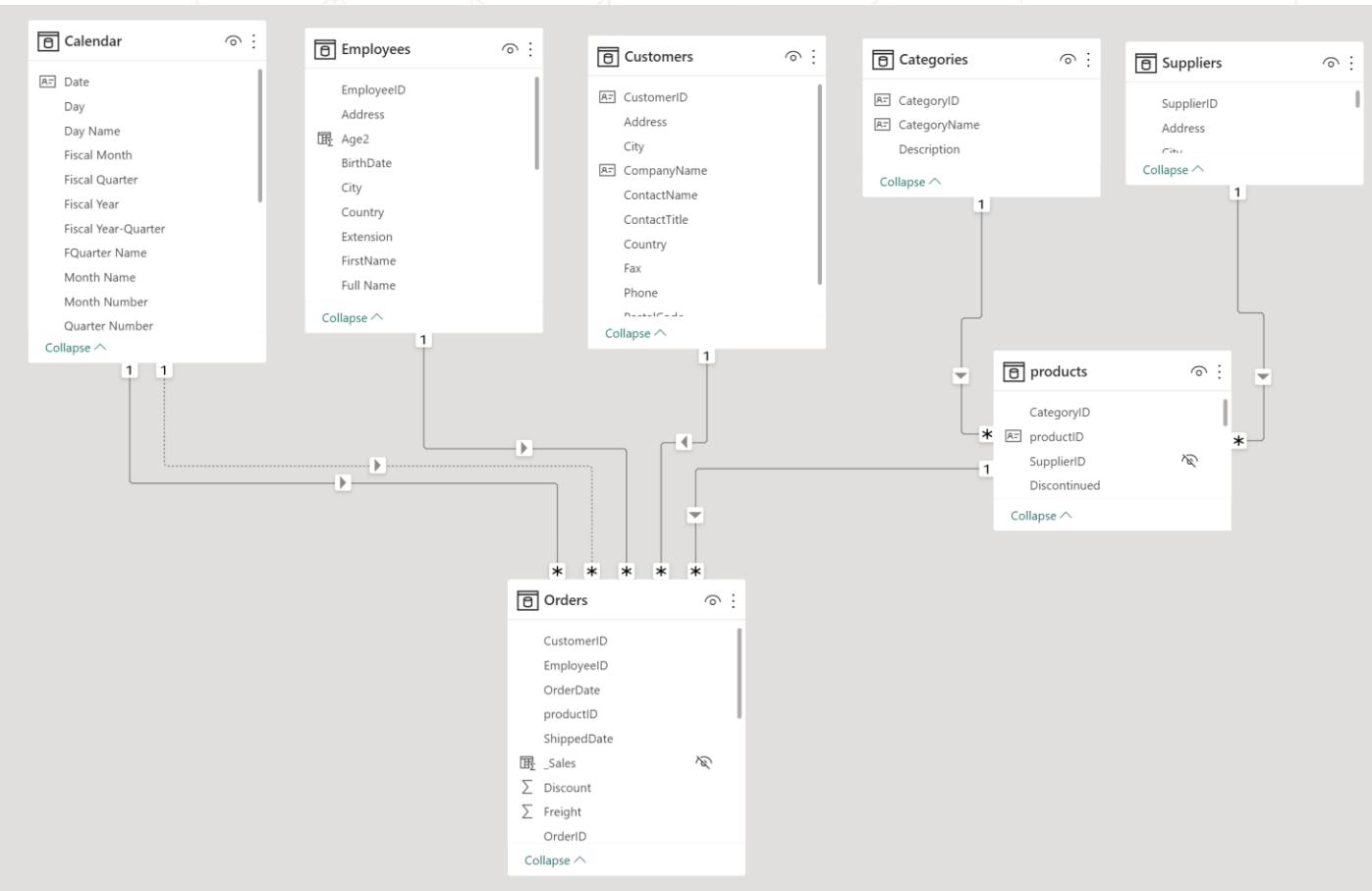
In the course of a TYPICAL month, how big is the largest dataset you likely would work with in Power BI?

You can see how people vote. [Learn more](#)

Response	Percentage
Less than 100,000 rows	18%
100,000 to 999,999 rows	26%
1,000,000 to 9,999,999 rows	32%
Greater than 10 million rows	24%

597 votes • Poll closed

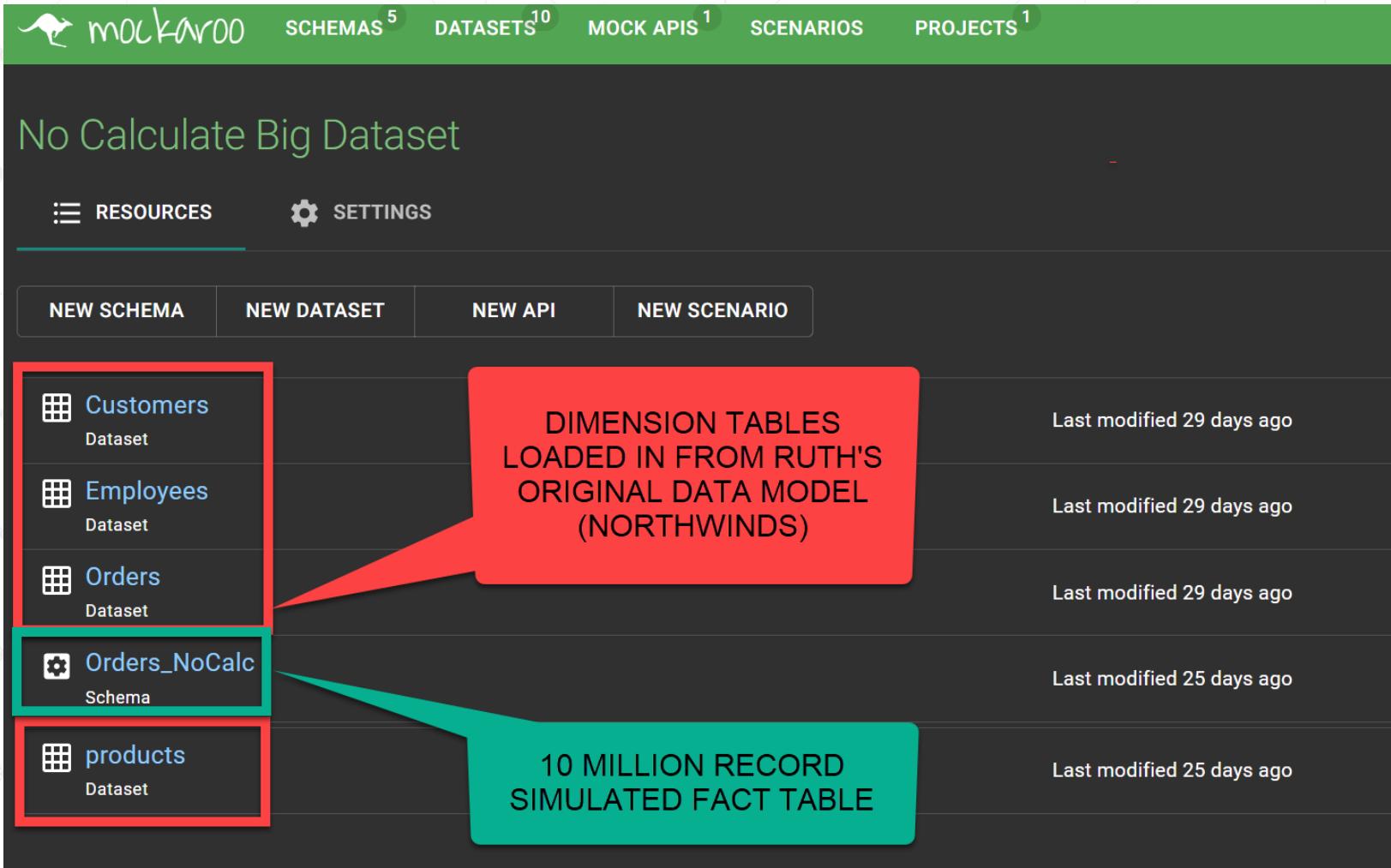
MOCKING UP THE DATA MODEL (2,155 TO 10M RECORDS)



CHALLENGE:
DRAMATICALLY EXPANDING FACT TABLE WHILE MAINTAINING
INTEGRITY OF DIMENSION TABLES AND RELATIONSHIPS

MOCKAROO DATASET/ SCHEMA APPROACH

PowerBIUG



The screenshot shows the Mockaroo web interface with a dark theme. At the top, there's a navigation bar with the Mockaroo logo, 'SCHEMAS 5', 'DATASETS 10', 'MOCK APIs 1', 'SCENARIOS', and 'PROJECTS 1'. Below the navigation, the title 'No Calculate Big Dataset' is displayed. Underneath it, there are tabs for 'RESOURCES' and 'SETTINGS', with 'RESOURCES' being active. Below these are four buttons: 'NEW SCHEMA', 'NEW DATASET', 'NEW API', and 'NEW SCENARIO'. The main content area lists datasets and schemas:

Type	Name	Last modified
Dataset	Customers	29 days ago
Dataset	Employees	29 days ago
Dataset	Orders	29 days ago
Schema	Orders_NoCalc	25 days ago
Dataset	products	25 days ago

A red box highlights the 'Customers', 'Employees', and 'Orders' datasets. A green box highlights the 'Orders_NoCalc' schema. A red callout bubble points to the 'Customers', 'Employees', and 'Orders' entries with the text: 'DIMENSION TABLES LOADED IN FROM RUTH'S ORIGINAL DATA MODEL (NORTHWINDS)'. A green callout bubble points to the 'products' entry with the text: '10 MILLION RECORD SIMULATED FACT TABLE'.

MOCKAROO DETAILED SCHEMA SETUP

PowerBIUG

mockaroo SCHEMAS⁵ DATASETS¹⁰ MOCK APIs¹ SCENARIOS PROJECTS¹ MOVE TO PROJECT... GOLD

No Calculate Big Dataset / Orders_NoCalc

Orders_NoCalc

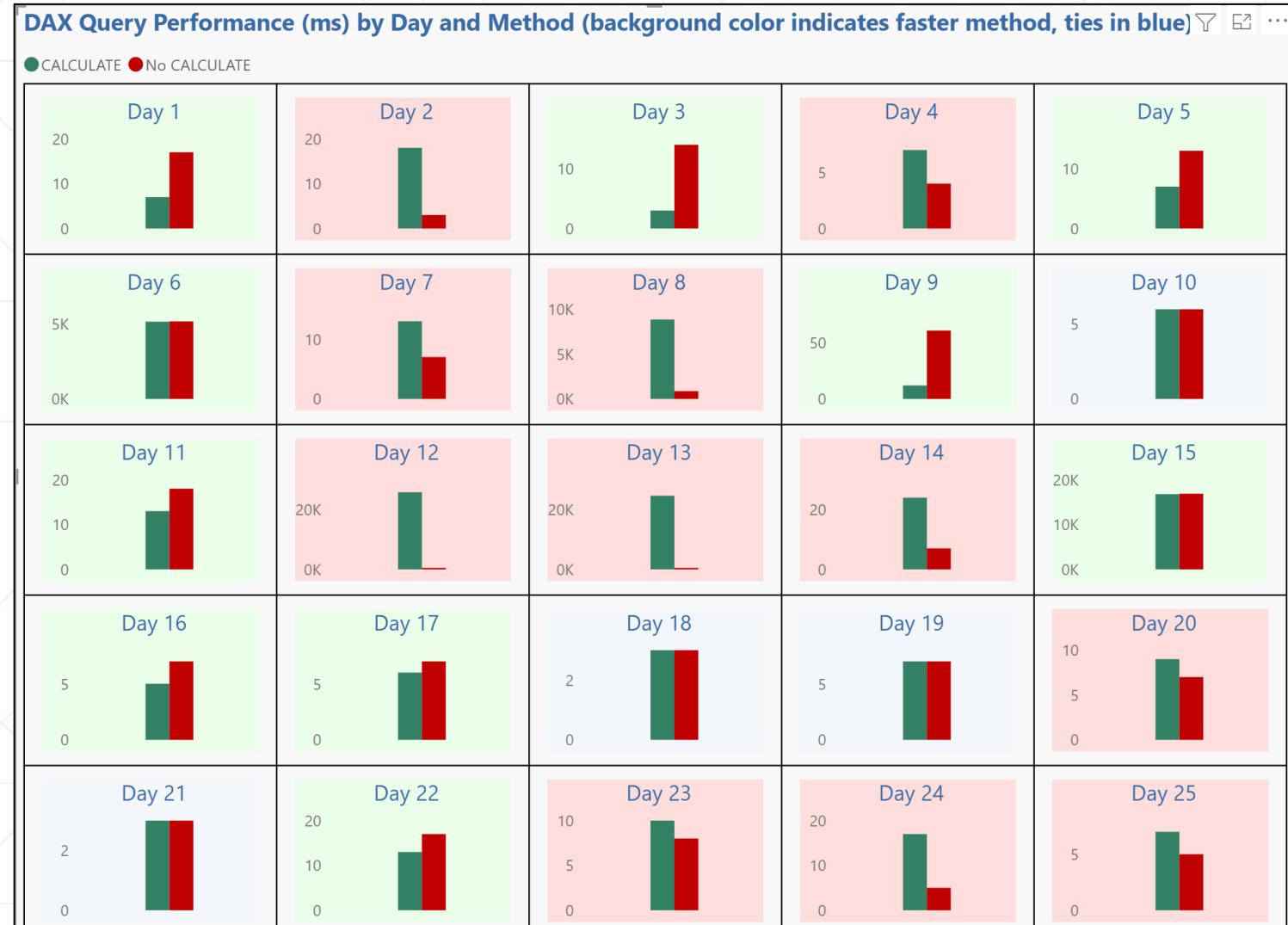
Field Name	Type	Options
OrderID	ULID	blank: 0 % Σ X
CustomerID	Dataset Column	Customers random blank: 0 % Σ X
productID	Dataset Column	products random blank: 0 % Σ X
EmployeeID	Dataset Column	Employees random blank: 0 % Σ X
RequiredDate	Datetime	01/12/2022 to 01/12/2023 format: m/d/yyyy blank: 0 % Σ X
ShippedDate	Datetime	01/12/2022 to 01/12/2023 format: m/d/yyyy blank: 37 % Σ X
ShipVia	Number	min: 1 max: 3 decimals: 0 blank: 0 % Σ X
Freight	Number	min: 2 max: 100 decimals: 0 blank: 0 % Σ X
ShipName	Dataset Column	Orders random blank: 0 % Σ X
ShipAddress	Dataset Column	Orders random blank: 0 % Σ X
ShipCity	Dataset Column	Orders random blank: 0 % Σ X
ShipRegion	Dataset Column	Orders random blank: 60 % Σ X
ShipPostalCode	Dataset Column	Orders random blank: 2 % Σ X
ShipCountry	Dataset Column	Orders random blank: 0 % Σ X
OrderDate	Datetime	03/03/2021 to 12/31/2022 format: m/d/yyyy blank: 0 % Σ X
Discount	Custom List	0, 5, 10, 15, 20, 25 random blank: 0 % Σ X
UnitPrice	Number	min: 2 max: 263 decimals: 0 blank: 0 % Σ X
Quantity	Number	min: 1 max: 130 decimals: 0 blank: 0 % Σ X

ADD ANOTHER FIELD

Rows: 10000000 Format: CSV Line Ending: Unix (LF) Include: header BOM

Public URL: https://www.mockaroo.com/Orders_NoCalc.csv

THE CALCULATE AND NO CALCULATE PERFORMANCE RESULTS (SMALL MULTIPLES)

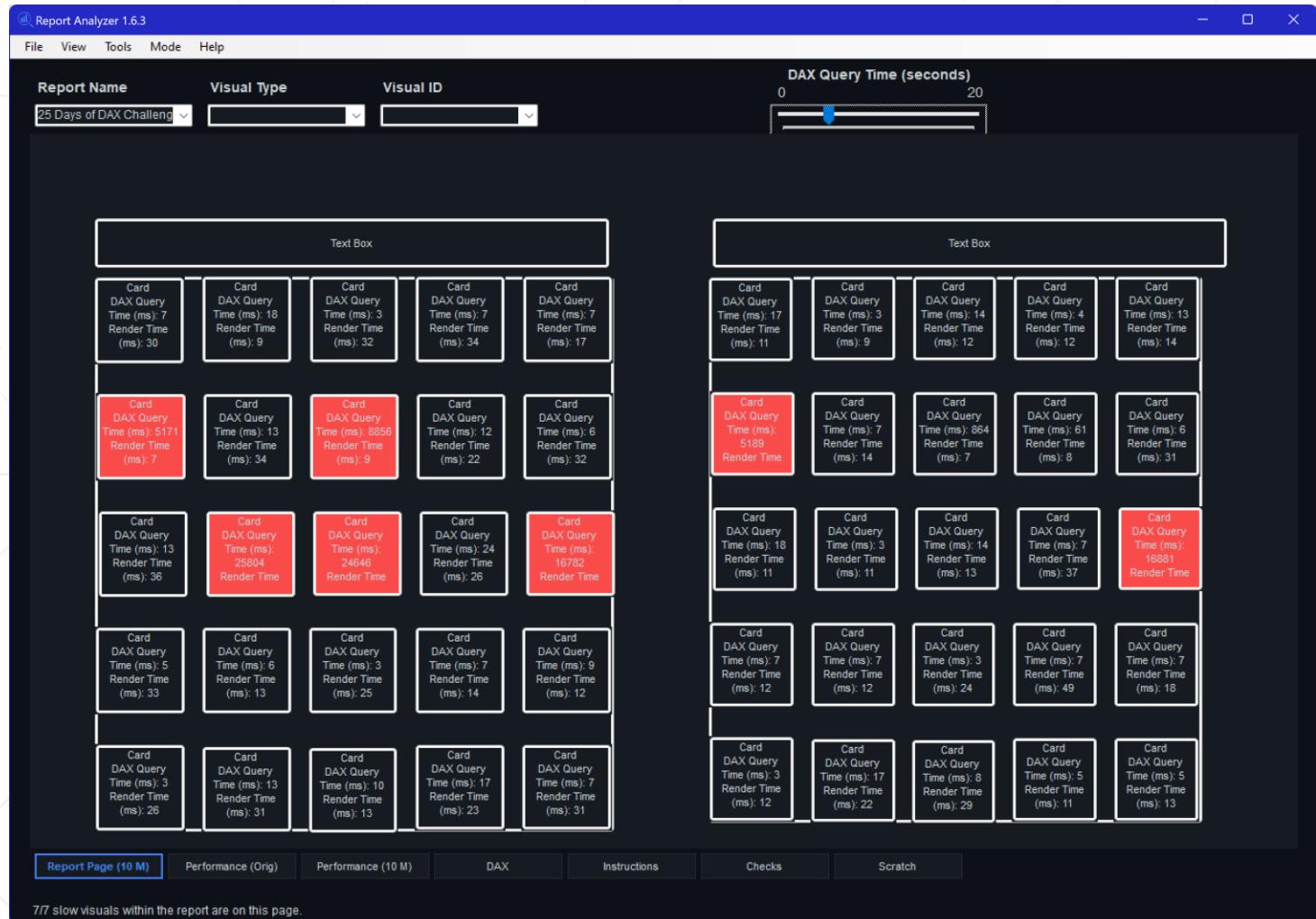


Note: Y axis scale varies by Day

THE CALCULATE AND NO CALCULATE PERFORMANCE RESULTS

(MICHAEL KOVALSKY'S REPORT
ANALYZER TOOL)

PowerBIUG



THE CALCULATE AND NO CALCULATE PERFORMANCE RESULTS

(MICHAEL KOVALSKY'S REPORT
ANALYZER TOOL)

PowerBIUG

The screenshot shows a window titled "Report Analyzer Recommendations" with a dark theme. At the top, there are "Tools" and "Help" menu options. Below the title, it says "Report Name: 25 Days of DAX Challenge - Mike Ford". The main content area lists recommendations with some items highlighted in yellow and one in blue:

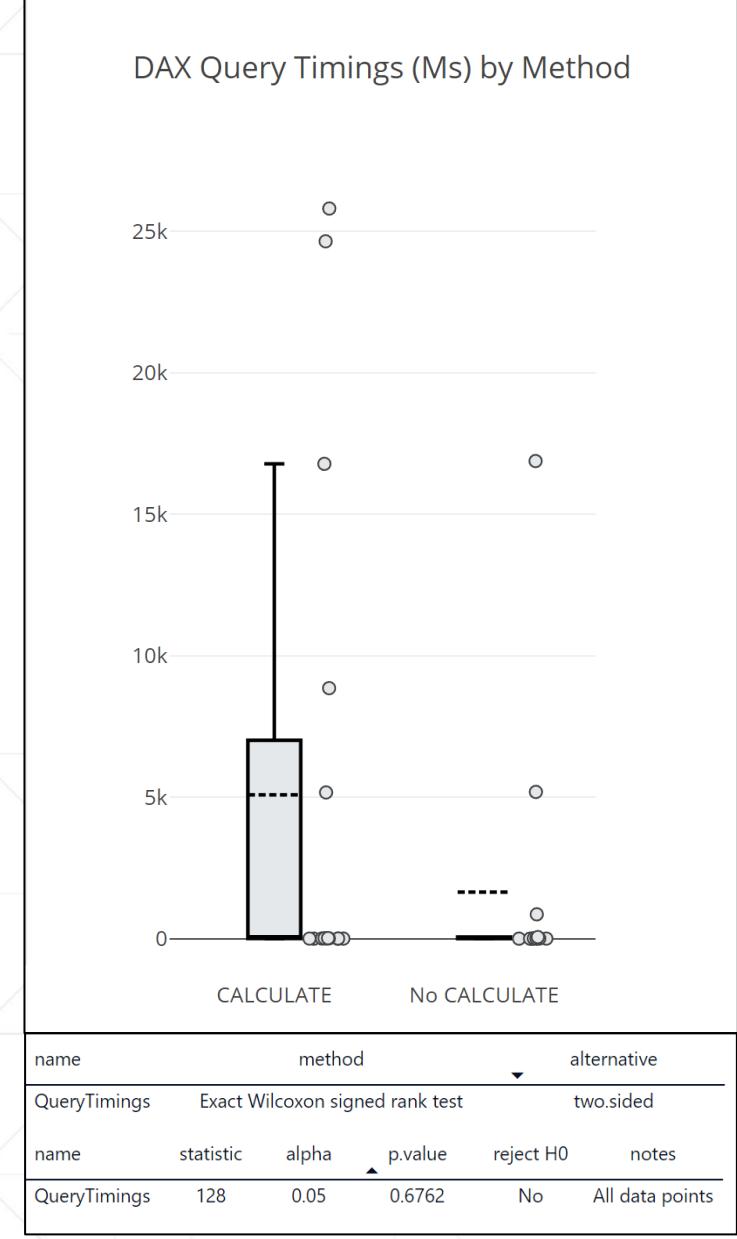
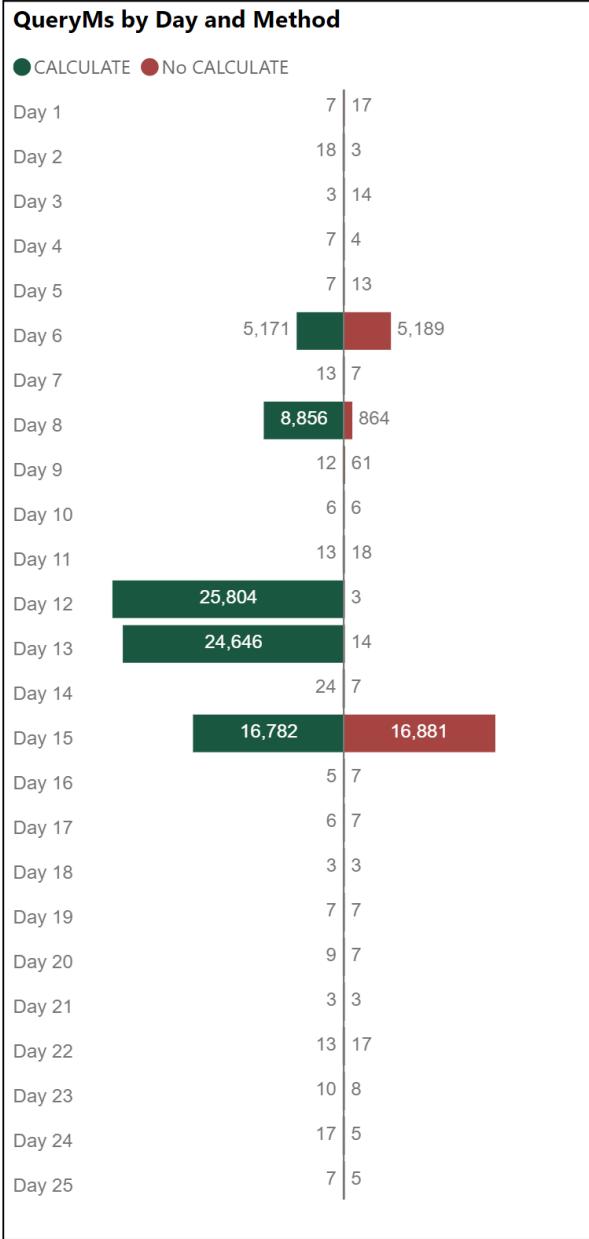
- Remove custom visuals which are not used in the report (2 objects)
 - Power Automate for Power BI
 - Microsoft Charticulator (Preview)
- Reduce the number of visible visuals on the page (1 object)
 - Report Page (10 M) (54)
- Reduce the number of objects within visuals (0 objects)
- Reduce usage of TopN filtering within visuals (0 objects)
- Reduce usage of filters on measures (0 objects)
- Avoid setting 'Show items with no data' on columns (1 object)
 - Checks - 17e6fd04b9221bbe6b74
- Avoid tall report pages with vertical scrolling (0 objects)
- Reduce usage of custom visuals (0 objects)
- Set 'Edit Interactions' for non-data visuals to 'none' (0 objects)**

At the bottom, a message states "4 objects in violation of 3 Report Analyzer recommendations."

STATISTICAL ANALYSIS OF CALCULATE VS “NO CALCULATE” PERFORMANCE

TL;DR: No evidence to support difference in performance between the two methods.

PowerBIUG



NEXT STEPS...

Extend out to 100M+ records

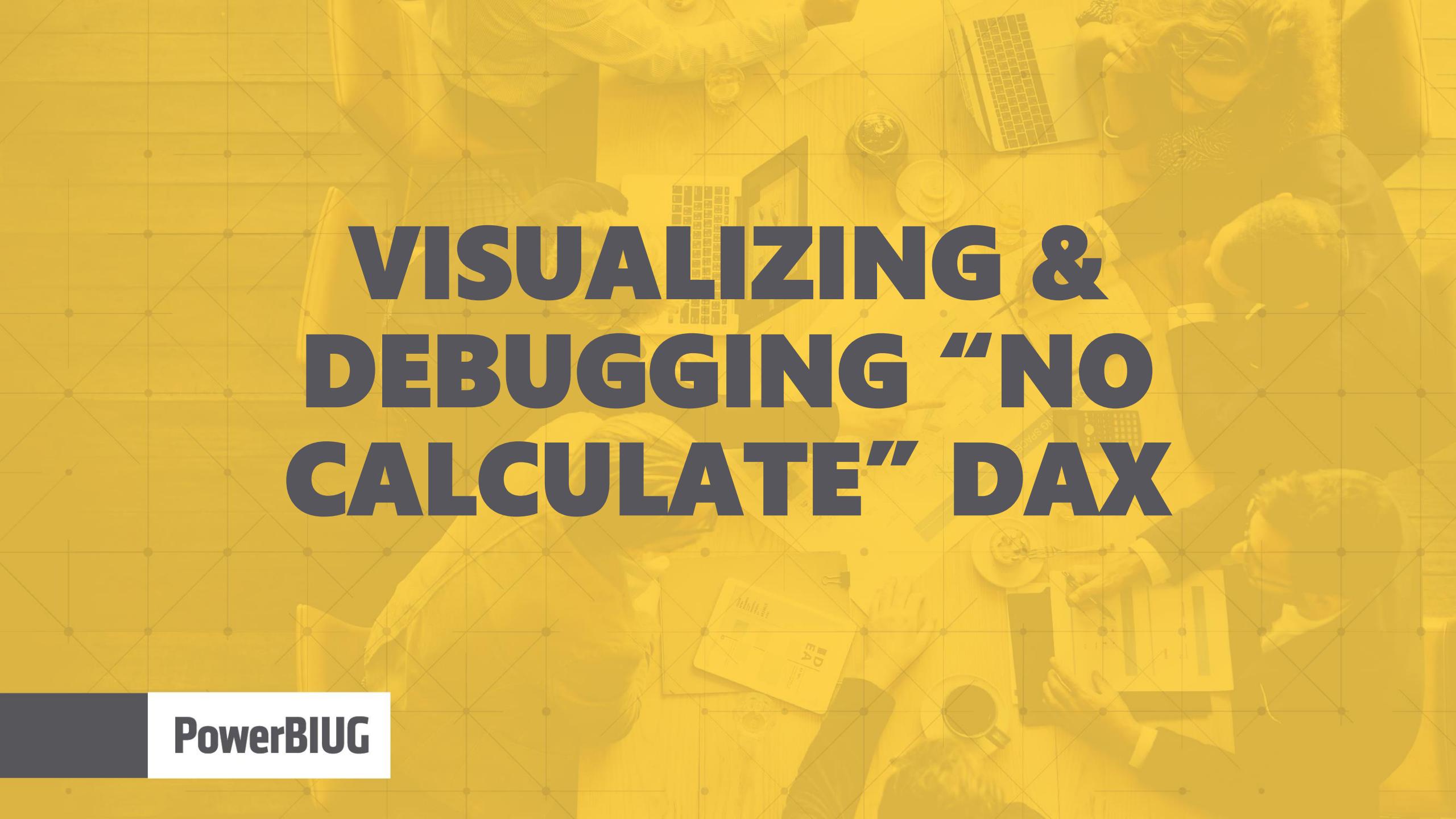
Larger sample size of measures

Test hypothesis that “No Calculate” results in smaller variance of query timings

Publish/peer review statistical analysis

Examine query plans to understand better how No Calculate and traditional DAX perform at large-scale

Develop dedicated training materials for this approach



VISUALIZING & DEBUGGING “NO CALCULATE” DAX

PowerBIUG

THE BEST ADDITION TO POWER BI IN 2022

PowerBIUG

TOCSV DAX Function (Text)

[Syntax](#) | [Return values](#) | [Examples](#) | [Related](#)

Converts the records of a table into a CSV (comma-separated values) text.

Syntax

```
TOCSV ( <Table> [, <MaxRows>] [, <Delimiter>] [, <IncludeHeaders>] )
```

PARAMETER	ATTRIBUTES	DESCRIPTION
Table		The table to be converted.
MaxRows	Optional	The maximum number of rows to be converted. A negative number means all rows are converted. Default is 10.
Delimiter	Optional	The field separator. Must be a non-empty constant string. Default is ",".
IncludeHeaders	Optional	If true, the header row is included. Default is true.

Return values

SCALAR A single [string](#) value.

Result in CSV format.



Brian Julius • You

Lifelong data geek | Data Science/Econometrics | Passionate ...
2mo •

...

A few days ago, I showed how you could use the new TOJSON DAX function plus a short R script to dynamically display and debug virtual tables. Here's an even easier way (no R required) using the TOCSV function.....
[...see more](#)

Even Easier Dynamic DAX Debugging with New TOCSV Function

```
CSVDAx =  
    VAR _variable =  
        ADDCOLUMNS(  
            FILTER(  
                SUMMARIZE(  
                    Teams, Team[team], Team[stage], Team[result]  
                ), Team[stage] = "final"  
            ), "#appear", [How Count TeamAppearances]  
        )  
    VAR _initialData =  
        TOPN(1, _variable, DESC)  
    VAR Result =  
        CONCATENATEX(_initialData,  
            TOCSV(_variable, 100, REPT(UNICHAR(127), 1), FALSE)  
        )  
    RETURN IntermediateCSV
```

The TOCSV statement dynamically converts the current state of the virtual table (`_variable`).

team
Argentina
Brazil
Croatia
Czechoslovakia
England
France
Germany
Hungary
Italy
Netherlands
Spain
Sweden
Uruguay
West Germany

Brazil	final	win	5
Brazil	final	lose	1

Brian Julius, 2022

LIVE EXAMPLES USING:

TOCSV

TABULAR EDITOR 3

DAX STUDIO

PowerBIUG

A collage of various people working on laptops, tablets, and phones, with a network grid overlay.

Q&A / Open Forum

PowerBIUG