



Total Confusion



TIPS & TRICKS

3rd Edition – Solving Benchmarking Challenges (#30-32)

Carl-Olow “Kalle” Magnusson
Nordic Senior Visualization Advisor



Benchmarking is very common task in SAS Visual Analytics

I want to have percentage of total... when I add filters, how can I keep my totals?



Benchmarking - Simplified

Benchmarking = *comparing something with something else*

Comparison with Average

$$(455 - 512) / 512 = -11 \%$$

Percentage of Total

$$14\,600 / 32\,500 = 0,449231 \text{ (45 \%)}$$

Benchmarking - Simplified

Benchmarking = comparing something to a reference

Very straightforward, right?

Well, I've been involved in many benchmark challenges over the years and my conclusion is: Sometimes it is straightforward, sometimes it can be very complicated

$$14\ 600 / 32\ 500 = 0,449231 \text{ (45 \%)}$$

What functions in SAS Visual Analytics are useful?

Feature/Function	Type	VA feature	Usage
Aggregation operators, such as SUM, AVG, DISTINCT etc.	Aggregated Operator	Expression builder	Calculate different statistics with aggregation level control
AND, IF...ELSE, NOT and OR	Boolean Operator	Expression builder	Adding logical control in expressions
AggregateTable	Aggregated Operator	Expression builder	Aggregates values of a specific set of cells, such as 7 days average. No specific date or category column is required
Scoped Calculation	Data	Data Action	Allows us to create a new aggregated data set in VA, useful when handling advanced aggregations or handling nested aggregated data items
Parameters	Boolean Operator	Expression builder	parameter is set (true) or not (false).
IsSet	Boolean Operator	Expression builder	parameter is set (true) or not (false).
NumMiss	Boolean Operator	Expression builder	Is used to identify if missing values are present in your aggregation
New data from aggregation...	Data	Data Action	Allows us to create a new aggregated data set in VA, useful when handling advanced aggregations or handling nested aggregated data items
Periodic operators, such as Period, CumulativePeriod etc.	Aggregated Operator	Expression builder	Calculate comparing periodic and aggregated metrics over time, such as % difference compared with last year. A date column is required
AggregateCells	Aggregated Operator	Expression builder	Aggregates values of a specific set of cells, such as 7 days average. No specific date or category column is required

What VA capabilities to use and when?
It is always difficult to understand when to use what without a real use-case. In VA we have a lot of different functions, but when and how can they be used?

What functions in SAS Visual Analytics are useful?

Feature/Function	Type	VA feature	Usage
Aggregation operators, such as SUM, AVG, DISTINCT etc.	Aggregated Operator	Expression builder	Calculate different statistics with aggregation level control
AND, IF...ELSE, NOT and OR	Boolean Operator	Expression builder	Adding logical control in expressions
AggregateTable	Aggregated Operator	Expression builder	AggregateTable allows us to aggregate data and lock it onto a specific data/aggregation level
Scoped Calculations	Feature	Expression builder	Allows us to add multiple/different expressions into our calculated item
Parameters	Data Item	Data Item	Parameters can fetch a value to be used in a calculated item, filters, ranks etc.
IsSet	Comparison Operator	Expression builder	IsSet function is used to determine if a parameter is set (true) or not (false).
NumMiss	Aggregated Operator	Expression builder	Is used to identify if missing values are present in your aggregation
New data from aggregation...	Data	Data Action	Allows us to create a new aggregated data set in VA, useful when handling advanced aggregations or handling nested aggregated data items
Periodic operators, such as Period, CumulativePeriod etc.	Aggregated Operator	Expression builder	Calculate comparing periodic and aggregated metrics over time, such as % difference compared with last year. A date column is required
AggregateCells	Aggregated Operator	Expression builder	Aggregates values of a specific set of cells, such as 7 days average. No specific date or category column is required

My tips & tricks rating system 😊



Easy – Requires just one or few clicks



Intermediate - Requires one or few clicks and some other actions/objects/parameters



Expert - Requires multiple clicks, multiple actions/objects **parameters** and might require programming skills

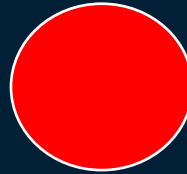
Today we have 3 different use-cases/scenarios

- How can I keep my totals intact when using filters?
- What about calculated items using Distinct Count?
- AggregateTable function, what is it good for?

Solving Benchmarking Challenges

Tips & Tricks # 30

How can I keep my totals intact when using filters?






How can I keep my totals intact when using filters?

Tips & Tricks # 30

Challenge: I have a dashboard with some sale benchmarking metrics and added some filters. However, end-users DO NOT want Sales % of Total to be affected by Product Brand and Line filters, only Transaction year (baseline)

Sales Benchmark Dashboard 

Product Brand: Transaction Year: 2000 to 2012

Product Line:

- Bead
- Figurine
- Game
- Gift
- Kiosk
- Plush
- Promo
- Store

Sales Benchmarking per Country

Facility Country	Product Sale	Sales % of Total
Spain	4 026 412	17,5 %
United Kingdom	3 876 165	16,8 %
Brazil	1 669 142	7,3 %
Germany	1 591 235	6,9 %
Sweden	1 555 995	6,8 %
Venezuela	1 389 260	6,0 %
Peru	1 368 653	5,9 %
Chile	1 070 876	4,7 %
Norway	1 017 861	4,4 %
Italy	1 004 142	4,4 %
Colombia	761 669	3,3 %
France	583 330	2,5 %
Denmark	492 724	2,1 %
Australia	479 695	2,1 %
Sum:	23 021 376	Total: 100,0 %





How can I keep my totals intact when using filters?

Tips & Tricks # 30

Challenge: I have a dashboard with some sale benchmarking metrics and added some filters. However, end-users DO NOT want Sales % of Total to be affected by Product Brand and Line filters, only Transaction year (baseline)

Product Brand: Novelty

Transaction Year: 2000 to 2012

Product Line:

- Bead
- Gift
- Kiosk
- Promo
- Store

Sales Benchmarking per Country

Facility Country	Product Sale	Sales % of Total
Spain	635 798	26,9 %
United Kingdom	614 239	26,0 %
Germany	268 801	11,4 %
Sweden	248 131	10,5 %
Italy	158 043	6,7 %
Norway	146 719	6,2 %
France	101 798	4,3 %
Australia	91 390	3,9 %
Denmark	81 484	3,4 %
New Zealand	18 197	0,8 %
Sum:	2 364 599	total: 100,0 %

According to our end-users, this is not correct!

It should be 2,8 %

$$635\,798 / 23\,021\,376$$





How can I keep my totals intact when using filters?

Tips & Tricks # 30

Challenge: I have a dashboard with some sale benchmarking metrics and added some filters. However, end-users DO NOT want Sales % of Total to be affected by Product Brand and Line filters, only Transaction year (baseline)

Product Brand: Novelty

Transaction Year: 2000

This is how Sales % of Total is calculated

(Sum (Product Sale) / Sum (Product Sale))

Product Line

- Bead
- Gift
- Kiosk
- Promo
- Store

Sales Benchmarking per Country

Facility Country	Product Sale	Sales % of Total
Spain	635 798	26,9 %
United Kingdom	614 239	26,0 %
Germany	268 801	11,4 %
Sweden	248 131	10,5 %
Italy	158 043	6,7 %
Norway	146 719	6,2 %
France	101 798	4,3 %
Australia	91 390	3,9 %
Denmark	81 484	3,4 %
New Zealand	18 197	0,8 %
Sum:	2 364 599	Total: 100,0 %

How can I keep my totals intact when using filters?

Tips & Tricks # 30

Challenge: I have a dashboard with some sale benchmarking metrics and added some filters. However, end-users DO NOT want Sales % of Total to be affected by Product Brand and Line filters, only Transaction year (baseline)



Preparing data with pre-defined columns with totals is a very common solution, but it might impact flexibility...

Is there any other solution?



How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution

We will keep the original data item Product Sale as our total and create a new data item that will be affected by our filters

$$\frac{\text{Product Sale (Filtered)}}{\text{Product Sale}} = \text{Sales \% of Total}$$

Product Line <input checked="" type="checkbox"/> Bead <input type="checkbox"/> Gift <input checked="" type="checkbox"/> Kiosk <input type="checkbox"/> Promo <input type="checkbox"/> Store	Sales Benchmarking per Country <table><thead><tr><th>Facility Country</th><th>Product Sale ▼</th><th>Product Sale (Filtered)</th><th>Sales % of Total (Filter Control)</th></tr></thead><tbody><tr><td>Spain</td><td>4 026 412</td><td>635 798</td><td>2,8 %</td></tr><tr><td>United Kingdom</td><td>3 876 165</td><td>614 239</td><td>2,7 %</td></tr><tr><td>Brazil</td><td>1 669 142</td><td>0</td><td>0,0 %</td></tr><tr><td>Germany</td><td>1 591 235</td><td>268 801</td><td>1,2 %</td></tr><tr><td>Sweden</td><td>1 555 995</td><td>248 131</td><td>1,1 %</td></tr><tr><td>Venezuela</td><td>1 389 260</td><td>0</td><td>0,0 %</td></tr></tbody></table>	Facility Country	Product Sale ▼	Product Sale (Filtered)	Sales % of Total (Filter Control)	Spain	4 026 412	635 798	2,8 %	United Kingdom	3 876 165	614 239	2,7 %	Brazil	1 669 142	0	0,0 %	Germany	1 591 235	268 801	1,2 %	Sweden	1 555 995	248 131	1,1 %	Venezuela	1 389 260	0	0,0 %
Facility Country	Product Sale ▼	Product Sale (Filtered)	Sales % of Total (Filter Control)																										
Spain	4 026 412	635 798	2,8 %																										
United Kingdom	3 876 165	614 239	2,7 %																										
Brazil	1 669 142	0	0,0 %																										
Germany	1 591 235	268 801	1,2 %																										
Sweden	1 555 995	248 131	1,1 %																										
Venezuela	1 389 260	0	0,0 %																										

How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution



Take-Off Checklist

1. Inactivate all current filters to avoid unexpected confusion 😊
2. Add two parameters and assign them
3. Create calculated items that are affected by our filters
4. Create calculated item for % of total
5. Activate necessary filters and filter dependencies



How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution Step 1: Inactivate all filters

Filter/Prompt Type	Effect on Data	Parameter Assignment	Filter/Prompt Control
Report	Always	Yes	Dependencies can be set
Page	Always	Yes	Dependencies can be set
Page Canvas/Body	Defined by the user	Yes	Filter actions is set by the user
Report Object	Defined by the user	No (Parameters can be used in filter expression)	Filter object logic is set by the user, either in simple or advanced mode
Data Filter	Always	No (parameters can be used in filter expression)	Filter object logic is set by the user
Ranks	By default	Yes (to control number of items to rank)	To keep your totals, make sure you have "All Other" selected

How can I keep my totals intact when using filters?



Tips & Tricks # 30 - Solution Step 2:1: Add two parameters

New Parameter [x]

Name:

Type:

Multiple values

Current value:

New Parameter [x]

Name:

Type:

Multiple values

Current value:



How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution

Step 2:2: Assign parameters to my control objects (Product Brand & Line)

Product Brand

Product Brand

Data Roles

Drop-down list - Product Brand 2

- Category
 - Product Brand
- Measure
 - + Add
- Parameter
 - _ProductBrandSelect
- Hidden
 - + Add

Product Line

- Bead
- Figurine
- Game
- Gift
- Kiosk
- Plush
- Promo
- Store

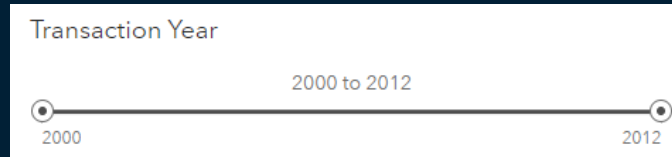
Data Roles

List - Product Line 2

- Category
 - Product Line
- Measure
 - + Add
- Parameter
 - _ProductLineSelect
- Hidden
 - + Add

How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution Step 2:2: What about the year interval slider/filter?



To demonstrate the flexibility with this solution, we will keep the year interval slider/filter as a normal filter. The year interval is used to set our benchmark baseline based on a specific year interval.



How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution

Step 3: Create calculated items that is affected by our filters, either we can do a one or two step approach. This is the one step approach:

Calculated data item: **Product Sale (Filtered)**

IF Product Brand **OR** Product Line is selected **RETURN** Product Sales **ELSE** 0

IsSet function is used to check if Product brand **OR** Product Line is selected

```
IF AND
  [
    _ProductBrandSelect IsSet
    NOT ( _ProductLineSelect IsSet )
  ]
  IF ( Product Brand = _ProductBrandSelect )
  RETURN
  RETURN Product Sale
  ELSE 0
```

```
IF OR
  [
    _ProductBrandSelect IsSet
    _ProductLineSelect IsSet
  ]
  IF Product Line In _ProductLineSelect
  RETURN
  RETURN Product Sale
  ELSE 0
ELSE
  Product Sale
```





How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution Step 3: Create calculated items that is affected by our filters, either we can do a one or two step approach. This is the one step approach:
Calculated data item: **Product Sale (Filtered)**

IF Product Brand OR Product Line is selected RETURN Product Sales ELSE 0

Product Brand: Novelty
Transaction Year: 2000 to 2012

Product Line:
 Bead
 Gift
 Kiosk
 Promo
 Store

Sales Benchmarking per Country

Facility Country	Product Sale	Product Sale (Filtered)
Spain	4 026 412	952 865
United Kingdom	3 876 165	884 371
Brazil	1 669 142	0
Germany	1 591 235	391 335
Sweden	1 555 995	363 345
Venezuela	1 389 260	0
Peru	1 368 653	0
Chile	1 070 876	0
Norway	1 017 861	234 865
Italy	1 004 142	242 187
Colombia	761 669	0
France	583 330	140 613
Denmark	492 724	120 609
Australia	479 695	121 988
Sum:	23 021 376	3 479 663



How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution

Step 4: Create Sales % of Total with filter control, based on the calculated item we created in step 3

Calculated data item: **Sales % of Total (Filter Control)**

```
( Sum  ( Product Sale (Filtered) ) / Sum  ( Product Sale ) )
```

Calculated item we created during step 3

Product Sale is our original column. As it is never affected by our Product Brand and Line filters, it will keep total sales
Make sure you set aggregation to `_ForAll_`





How can I keep my totals intact when using filters?

Tips & Tricks # 30 - Solution

Step 5: Activate necessary filters and filter dependencies

Setup a dependency filter between Product Brand and Line control/filter objects.

Note: All filtering is controlled by parameters and the calculated item we created during step 3.

Our Year slider control/filter is using normal actions filtering the list table. (Remember, we want to use year interval to set a benchmark baseline)

Product Brand: Novelty

Transaction Year: 2000 to 2012

Product Line:

- Bead
- Gift
- Kiosk
- Promo
- Store

Sales Benchmarking per Country

Facility Country	Product Sale	Sales % of Total	Sales % of Total (Filter Control)
Spain	4 026 412	17,5 %	2,8 %
United Kingdom	3 876 165	16,8 %	2,7 %
Brazil	1 669 142	7,3 %	0,0 %
Germany	1 591 235	6,9 %	1,2 %
Sweden	1 555 995	6,8 %	1,1 %
Venezuela	1 389 260	6,0 %	0,0 %
Peru	1 368 653	5,9 %	0,0 %
Chile	1 070 876	4,7 %	0,0 %
Norway	1 017 861	4,4 %	0,6 %
Italy	1 004 142	4,4 %	0,7 %
Colombia	761 669	3,3 %	0,0 %
France	583 330	2,5 %	0,4 %
Denmark	492 724	2,1 %	0,4 %
Australia	479 695	2,1 %	0,4 %

DO NOT use "Automatic actions on all objects"!!! If we do, our regain of total control will be lost 😊

How can I keep my totals intact when using filters?



Tips & Tricks # 30 - Solution

Sales Benchmark Dashboard

Product Brand:

Transaction Year: 2000 to 2012

Product Line:

- Bead
- Gift
- Kiosk
- Promo
- Store

Sales Benchmarking per Country

Facility Country	Product Sale	Sales % of Total	Sales % of Total (Filter Control)
Spain	4 026 412	17,5 %	2,8 %
United Kingdom	3 876 165	16,8 %	2,7 %
Brazil	1 669 142	7,3 %	0,0 %
Germany	1 591 235	6,9 %	1,2 %
Sweden	1 555 995	6,8 %	1,1 %
Venezuela	1 389 260	6,0 %	0,0 %
Peru	1 368 653	5,9 %	0,0 %
Chile	1 070 876	4,7 %	0,0 %
Norway	1 017 861	4,4 %	0,6 %
Italy	1 004 142	4,4 %	0,7 %
Colombia	761 669	3,3 %	0,0 %
France	583 330	2,5 %	0,4 %
Denmark	492 724	2,1 %	0,4 %
Australia	479 695	2,1 %	0,4 %
Sum:	23 021 376	Total: 100,0 %	Total: 10,3 %

Solving Benchmarking Challenges

Tips & Tricks # 31

What about calculated items using Distinct Count?





What about calculated items using Distinct Count?

Tips & Tricks # 31

Challenge: I have added a new benchmark metrics into my Dashboard and it is based on Distinct Customer Count. I have followed the steps in Tips & Tricks #30, but all my distinct counts is increased by one, why?

```
IF AND [
  _ProductBrandSelec IsSet
  NOT ( _ProductLineSelect IsSet )
]
IF ( Product Brand = _ProductBrandSelec )
RETURN Customer
ELSE " "
```

We are using the same approach as we did in Tips & Tricks #30.

In this case we are introducing an error, why? Column Customer (ID) is a character and when using our filter control approach, we are setting Customer to missing (blank).

So, when using the Distinct Count operator, it will also count the blanks (because Customer is a character).



What about calculated items using Distinct Count?

Tips & Tricks # 31

Challenge: I have added a new benchmark metrics into my Dashboard and it is based on Distinct Customer Count. I have followed the steps in Tips & Tricks #30, but all my distinct counts is increased by one, why?

Distinct Customer Count (Filtered), Distinct Count that is affected by our filters (using the same solution approach as we did in Tips & Tricks # 30)

Product Line

- Bead
- Figurine
- Game
- Gift
- Kiosk
- Plush
- Promo
- Store

Sales Benchmarking per Country

Facility Country	Distinct Customer Count	Distinct Customer Count (Filtered)
United Kingdom	6 227	1 087
Spain	6 036	1 163
Brazil	4 767	1
Germany	3 843	523
Italy	2 261	325
Venezuela	2 144	1
Sweden	2 018	460
Australia	2 008	257
France	1 740	229
Norway	1 385	258
Chile	1 229	1
Denmark	1 211	173
Russia	1 092	1
Total:	45 671	4 516



What about calculated items using Distinct Count?

Tips & Tricks # 31 - Solution The NumMiss operator is our Hero

So, when we are using Distinct Count (Character based), we need to make sure that we are not counting our missing values. We need an extra step where we are using the NumMiss function.

Calculated data item:

Distinct Customer Count
(NumMiss Filtered)

Calculated item **Customer (Filtered)** uses the same approach we did in Tips & Trick #30

```
Distinct [_ByGroup_] ( Customer (Filtered) ) - IF ( NumMiss [_ByGroup_] ( Customer (Filtered) ) > 0 ) RETURN 1 ELSE 0
```

NumMiss will look for missing values (if NumMiss is > 0) and in our use-case it will subtract one from our Customer (Filtered) distinct count 😊



What about calculated items using Distinct Count?

Tips & Tricks # 31 - Solution The NumMiss operator is our Hero

So, when we are using Distinct Count (Character based), we need to make sure that we are not counting our missing values. This is where the NumMiss function comes into play.

Distinct Customer Count (Filtered), Distinct Count that is affected by filters

Distinct Customer Count (NumMiss Filtered), Distinct Count that is affected by filters. But with the NumMiss function

Product Line

- Bead
- Figurine
- Game
- Gift
- Kiosk
- Plush
- Promo
- Store

Sales Benchmarking per Country

Facility Country	Distinct Customer Count	Distinct Customer Count (Filtered)	Distinct Customer Count (NumMiss Filtered)
United Kingdom	6 227	1 087	1 086
Spain	6 036	1 163	1 162
Brazil	4 767	1	0
Germany	3 843	523	522
Italy	2 261	325	324
Venezuela	2 144	1	0
Sweden	2 018	460	459
Australia	2 008	257	256
France	1 740	229	228
Norway	1 385	258	257
Chile	1 229	1	0
Denmark	1 211	173	172
Russia	1 092	1	0
Total:	45 671	4 516	4 515





What about calculated items using Distinct Count?

Tips & Tricks # 31 - Solution Let's wrap it up

Distinct Customer Count % of total (with filter control)

Distinct Customer Count (NumMiss Filtered) / Distinct (Customer))



Product Line

- Bead
- Figurine
- Game
- Gift
- Kiosk
- Plush
- Promo
- Store

Sales Benchmarking per Country

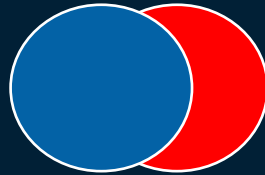
Facility Country	Distinct Customer Count	Distinct Customer Count (Filtered)	Distinct Customer Count (NumMiss Filtered)	Distinct Customer Count % of total (with filter control)
United Kingdom	6 227	1 087	1 086	2,4 %
Spain	6 036	1 163	1 162	2,5 %
Brazil	4 767	1	0	0,0 %
Germany	3 843	523	522	1,1 %
Italy	2 261	325	324	0,7 %
Venezuela	2 144	1	0	0,0 %
Sweden	2 018	460	459	1,0 %
Australia	2 008	257	256	0,6 %
France	1 740	229	228	0,5 %
Norway	1 385	258	257	0,6 %
Chile	1 229	1	0	0,0 %
Denmark	1 211	173	172	0,4 %
Russia	1 092	1	0	0,0 %
Total:	45 671	4 516	4 515	9,9 %



Solving Benchmarking Challenges

Tips & Tricks # 32

AggregateTable function, what is it good for?



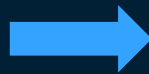
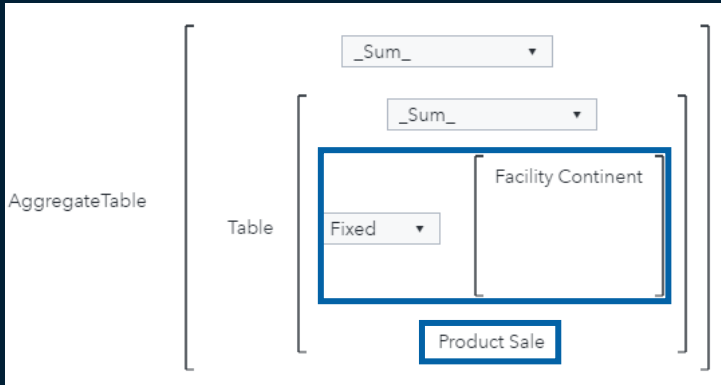


AggregateTable Function – what is it good for?

Tips & Tricks # 32

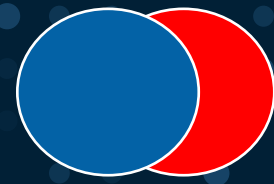
Using AggregateTable Function allows you to lock a value on a specific category (or level). In this example we will aggregate (Sum) Product Sales and lock it on a Continent Category level

Calculated Item: Product Sale (Continent)



Facility Continent	Facility Country	Product Sale	Product Sale (Continent)
Africa	Morocco	239 816	962 386
Africa	South Africa	237 676	962 386
Africa	Nigeria	277 147	962 386
Africa	Egypt	207 747	962 386
Asia	Indonesia	59 034	825 542
Asia	Russia	195 588	825 542
Asia	South Korea	63 147	825 542
Asia	China	140 487	825 542
Asia	Japan	121 247	825 542
Asia	Saudi Arabia	61 040	825 542
Asia	Singapore	63 078	825 542
Asia	India	56 696	825 542
Asia	Israel	65 226	825 542

AggregateTable Function – what is it good for?

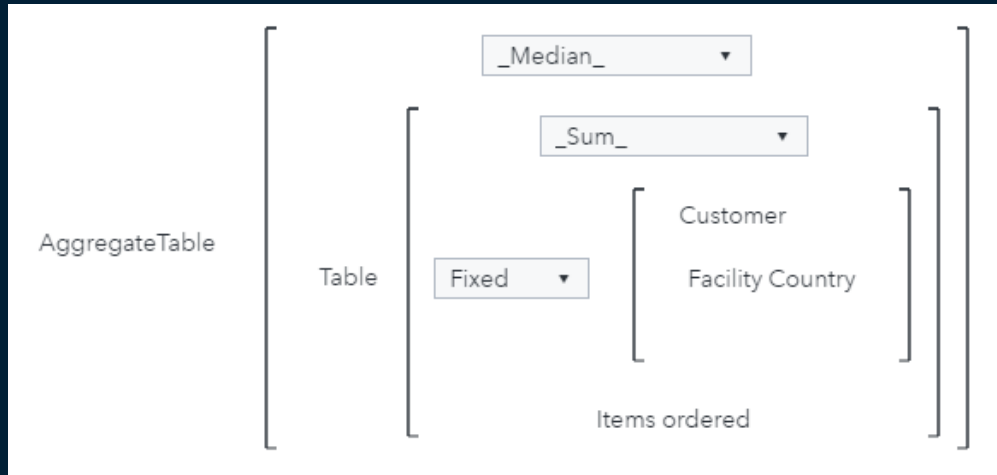


Tips & Tricks # 32 - Solution

Challenge: This is a very specific use-case that I have encountered twice over the last couple of months.

Similar use-case with my demo data: I want to count number of ordered items per customer, and then I want to calculate the Median of customer ordered items by country. Sounds straight forward, right?

Using AggregateTable might be the option to solve this challenge





AggregateTable Function – what is it good for?

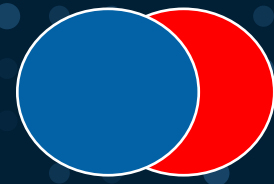
Tips & Tricks # 32 - Solution

AggregateTable is a very useful and powerful function/operator with many options

The screenshot shows the SAS AggregateTable function interface. The 'Table' dropdown is set to 'Fixed'. The 'Items ordered' section lists 'Customer' and 'Facility Country'. The 'AggregateTable' dropdown is set to '_Median_'. A list of available aggregate functions is shown on the right, with '_Median_' highlighted.

- _Avg_
- _CoeffVar_
- _Count_
- _CountDistinct_
- _CountMissing_
- _CSS_
- _Kurtosis_
- _Max_
- _Median_**
- _Min_
- _PValT_
- _Q1_
- _Q3_
- _Skew_
- _StdDev_
- _StdErr_
- _Sum_
- _TStat_
- _USS_
- _Var_

AggregateTable Function – what is it good for?



Tips & Tricks # 32 - Solution

AggregateTable is a very useful and powerful function/operator with many options

Do you want to learn more about the AggregateTable function?

Fantastic article by Renato Luppi

SAS Visual Analytics Advanced Calculations (part 2 of 4): AggregateTable

<https://communities.sas.com/t5/SAS-Communities-Library/SAS-Visual-Analytics-Advanced-Calculations-part-2-of-4/ta-p/538541>



Thank you for your time

Any Questions?