



Prof. Carlos J. Costa, PhD

Associate Professor of Information Systems and Operation Management

Carlos J. Costa (2022)



Carlos J. Costa (2022)



M



Power Query M Formula Language.



Language used in the context of PowerQuery



It's a functional, case sensitive language like F#.



## DAX



### Data Analysis Expressions



DAX is primarily a formula or a query language.

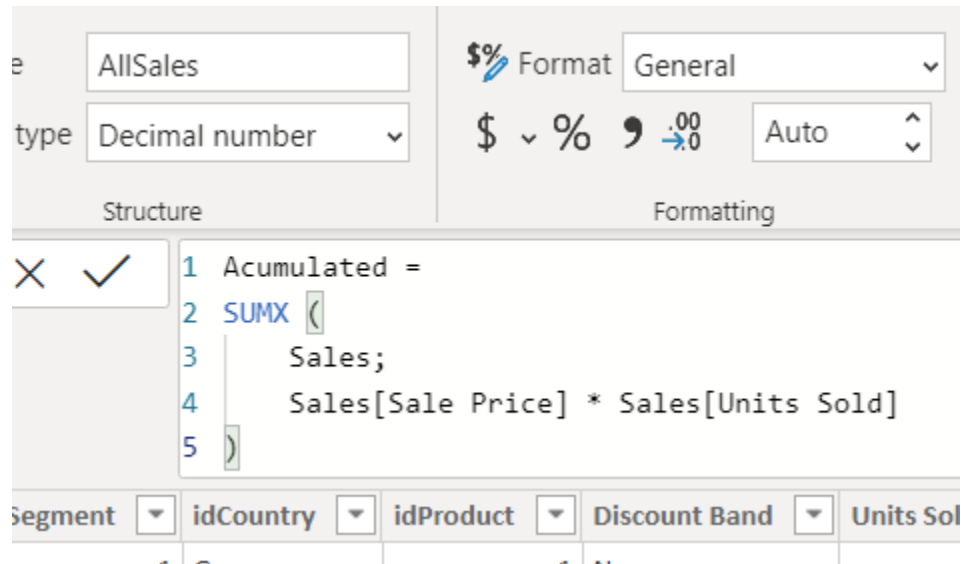


DAX was developed in 2009 by Microsoft



## DAX

- Creating Columns



The screenshot shows the DAX editor interface. The 'Name' field is set to 'AllSales' and the 'Data type' is 'Decimal number'. The 'Format' is set to 'General'. The 'Structure' tab is active, showing the DAX formula:

```
1 Acumulated =  
2 SUMX (  
3     Sales;  
4     Sales[Sales Price] * Sales[Units Sold]  
5 )
```

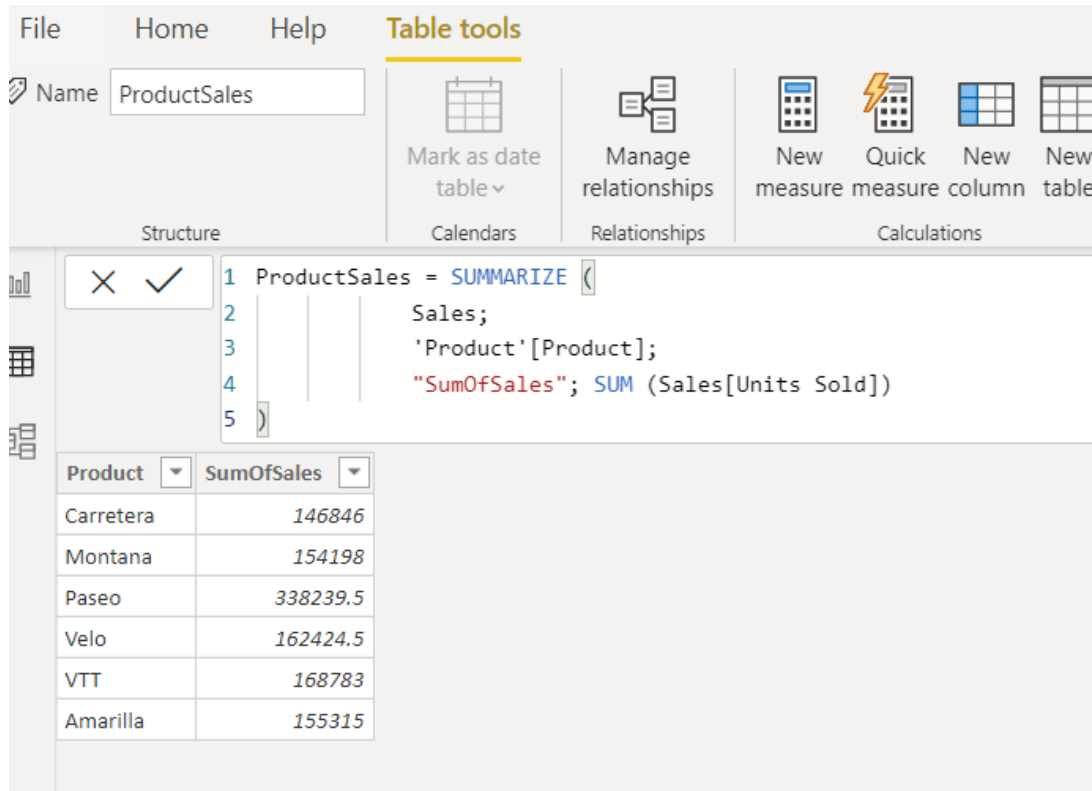
The 'Formatting' tab is also visible, showing options for currency symbols, percentage, and decimal places, with 'Auto' selected for the decimal places.

Segment	idCountry	idProduct	Discount Band	Units Sol



## DAX

- Creating Table



The screenshot shows the Power BI DAX editor interface. The 'Table tools' ribbon is active, with the 'Name' field set to 'ProductSales'. The 'Calculations' group contains the 'New table' button, which is highlighted. The DAX formula bar shows the following code:

```
1 ProductSales = SUMMARIZE (  
2     Sales;  
3     'Product'[Product];  
4     "SumOfSales"; SUM (Sales[Units Sold])  
5 )
```

Below the formula bar, a preview table is displayed with the following data:

Product	SumOfSales
Carretera	146846
Montana	154198
Paseo	338239.5
Velo	162424.5
VTT	168783
Amarilla	155315



## DAX

- Arithmetic Operations and Conditions:

```
Sales[SalesAmount] = Sales[ProductPrice] *  
Sales[ProductQuantity]
```

```
comment = if(Sales[Profit]>0;"Profit";"Loss")
```

```
Margin = if(ISERROR(Sales[profit]/Sales[value]));  
          BLANK();  
          Sales[profit]/Sales[value]  
        )
```

```
country= SWITCH ( Sales[idCountry],  
                  "Fr", "France",  
                  "Ca", "Canada",  
                  "Me", "Mexico",  
                  "Un", "USA",  
                  "Other"  
                )
```



## DAX

- LOOKUPVALUE

```
DiscountValue =  
LOOKUPVALUE (Discount [Discount];  
             Discount [Discount Band];  
             Sales [Discount Band])
```

```
TotalCosts = LOOKUPVALUE (  
    'Product' [Manufacturing cost];  
    'Product' [idProduct];  
    Sales [idProduct]  
) * Sales [Units Sold]
```





## DAX

- Summarizing

```
AllSales = sum(Sales[TotalSales_no_Discount])
```

- SUMX allows to iterate line by line

```
AllSales =  
SUMX (  
    Sales;  
    Sales[Sales Price] * Sales[Units Sold]  
)
```



## DAX

- iterate sales but with conditions

```
AllSales =  
SUMX (  
    FILTER (  
        Sales;  
        Sales[idCountry]="fr"  
    ) ;  
    Sales[Sales Price] * Sales[Units  
Sold]  
)
```



## DAX

- SQL:

```
Select Product.Product, Sum([Units Sold])  
From Sales, Product  
Where Product.idProduct=Sales.idProduct  
Group by Product.Product
```

*<- inner join*

- Corresponding in DAX:

```
Table = SUMMARIZE (  
    Sales;  
    'Product' [Product];  
    "SumOfSales"; SUM (Sales[Units Sold])  
)
```



## SQL

### SQL Server database



Server 

Database (optional)

Data Connectivity mode 

- Import
- DirectQuery

Advanced options

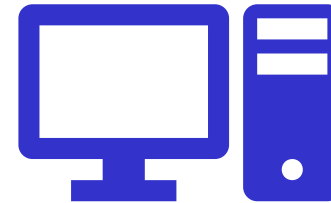
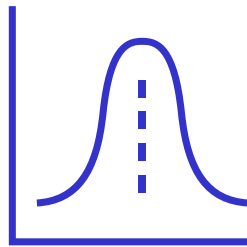
Command timeout in minutes (optional)

SQL statement (optional, requires database)

- Include relationship columns
- Navigate using full hierarchy
- Enable SQL Server Failover support

OK

Cancel

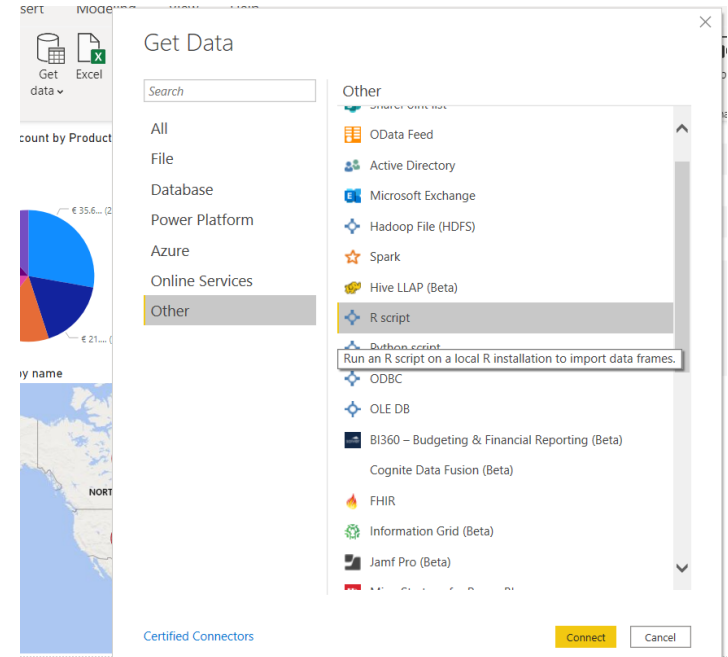
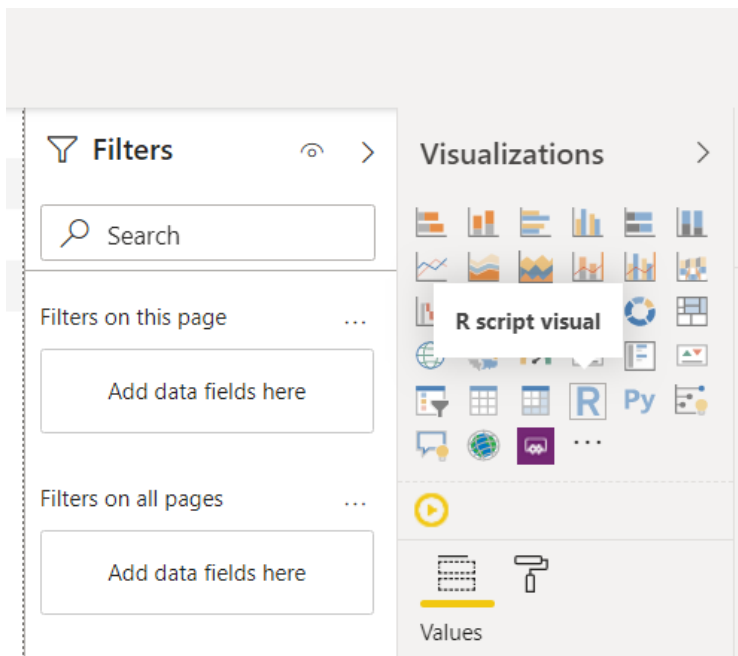


R is a free software environment for statistical computing and graphics.

It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS.



- Getting Data using script
- Customizing Visualizations





- Generic Language
- License: compatible with GNU-GPL
- Guido van Rossum
- <https://www.python.org/>





## Python script



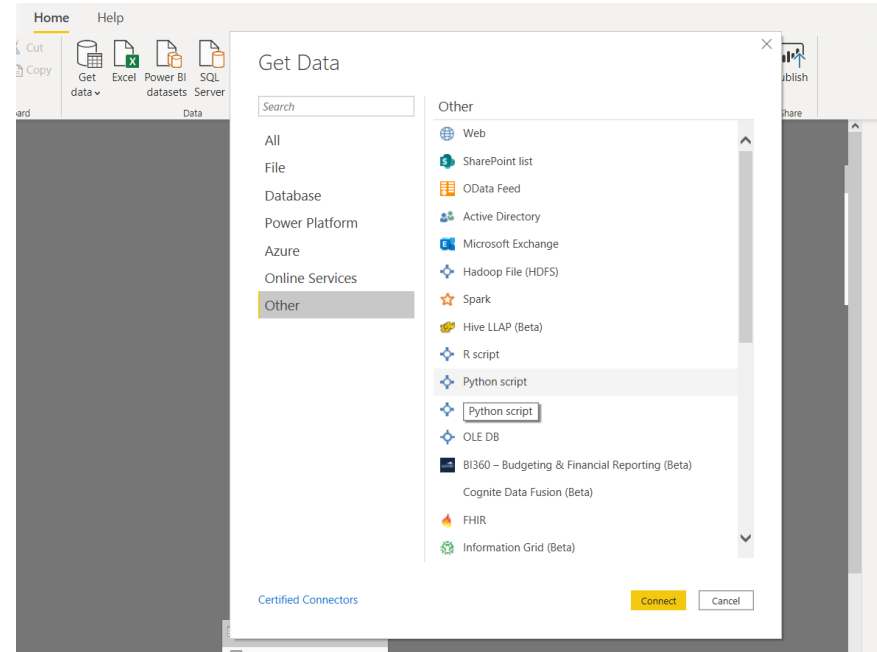
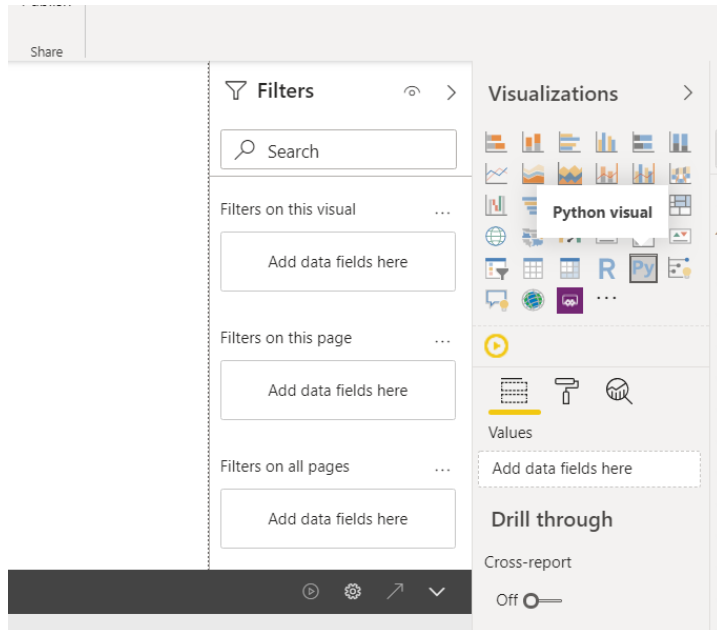
Script

```
import pandas as pd
data = [['Ana',20],['Margarida',20],['Silvia',19]]
df = pd.DataFrame(data,columns=['Name','Age'],dtype=float)
print (df)
```





- Getting Data using script
- Customizing Visualizations



# References

- <https://docs.microsoft.com/en-us/powerquery-m/>
- <https://docs.microsoft.com/en-us/dax/>
- <https://docs.microsoft.com/en-us/power-bi/desktop-python-scripts>
- <https://docs.microsoft.com/en-us/power-bi/desktop-r-visuals>