Analysis Services

The journey is ever onward

Gabi Münster

- BI consultant at oh22data AG since 2010
- Experience with SQL Server since version 2000
- Main topics: SSAS (Multidimensional/Tabular), SSRS, MDS, Power BI
- Speaker at chapter meetings, SQL Saturdays, Pass Summit, SQL Bits
- Chapter lead PASS RG Ruhrgebiet
- Data Platform MVP since 2017

g.muenster@oh22.net
@SQLMissSunshine
https://www.linkedin.com/in/gabimuenster/

Agenda

- What do we need Analysis Services for? And why different modes?
- How did SSAS develop over time?
- What are the latest developments and what can be expected?
- Which languages do we need for which purpose?
- Q&A

What do we need Analysis Services for?

- Original purpose:
  - Analytical data engine
  - What for:
    - Quick analysis of big amounts of data
    - Forecasting and prediction
    - Data mining
  - Semantic model for user-driven analysis
    - Business Logic Layer
    - Access and permission management

What do we need Analysis Services for?

- Today:
  - Analytical data engine *in a broader sense*
  - What for:
    - *Data transformation and management*
    - Quick analysis of big amounts of data
    - Forecasting and prediction/Data mining
    - *Data visualization*
    - *Real-time scenarios*
  - Semantic model for user-driven analysis
    - Business Logic Layer
    - Access and permission management

What do you need Analysis Services for?

- Multidimensional and Data Mining Mode
  - Traditional Datawarehouse approach
  - Still biggest spectrum of analytical functionality
  - Optimized for highly aggregated calculations
  - Significant complexity of implementation
- Tabular
  - Similar user handling
  - Increasing spectrum of functionality
  - Optimized for leaf level based calculations
  - Easier to get started with

What do you need Analysis Services for?

- Power BI
  - Self-service approach
  - Based on tabular mode
  - Planned to become the new superset for analysis services
  - Part of the Power Platform → Shift from BI tool to Business application

And why different modes?

- Multidimensional:
  - Most improved analytical engine
  - Unique navigational approach to handle data
  - Covers a huge area of analytical functionality including writeback, datamining, ...
- But:
  - Multidimensional mode needs IT department support for implementation and enhancement
  - Increasing necessity to close gap between business need and implementation endurance
  - Complexity of application model makes enhancements almost impossible

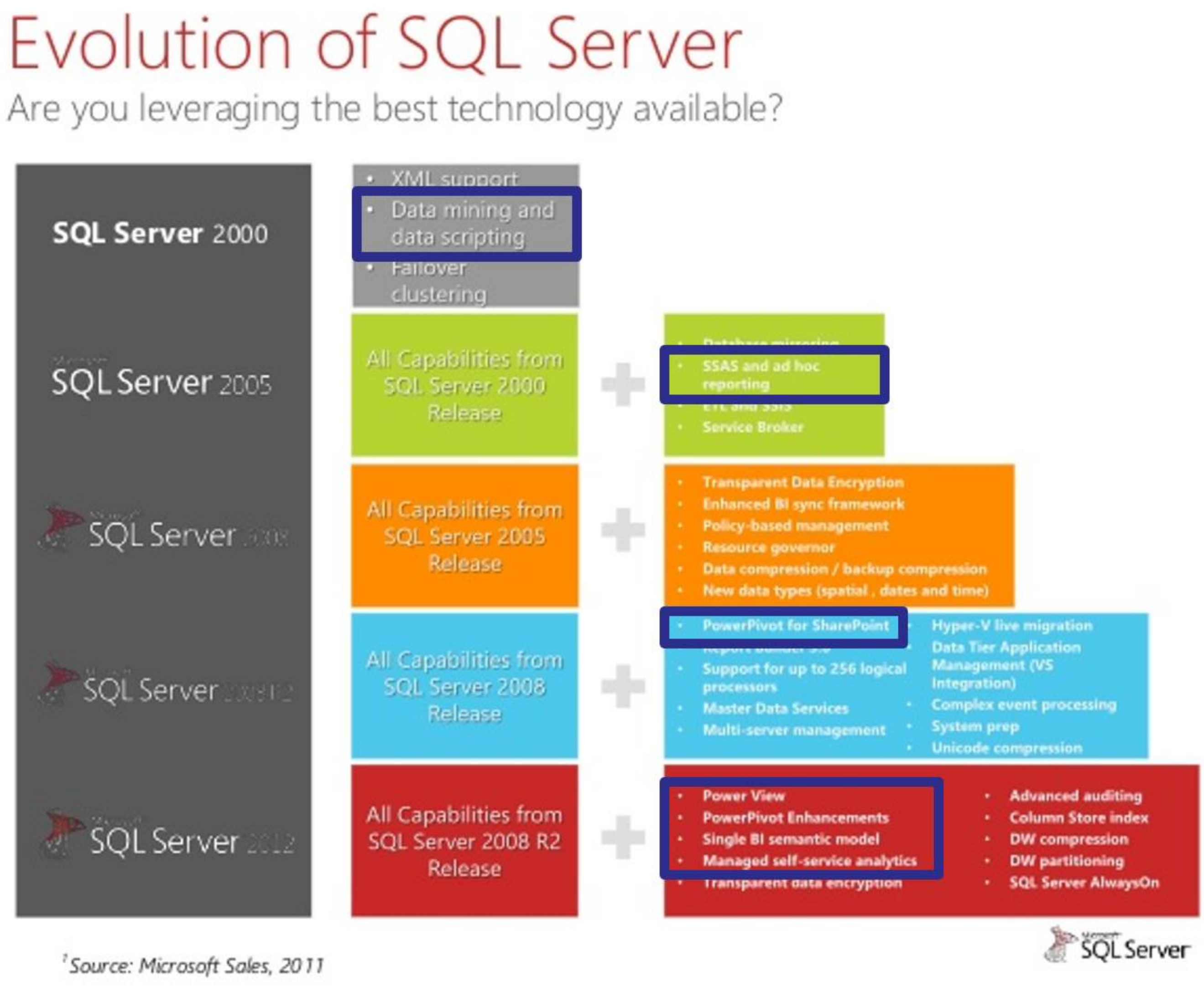
And why different modes?

- Tabular Mode:
  - In-Memory
    - Technological and price development of memory
  - Columnar Index
    - Well known concept (~ 1960s) to store registry of distinct values per column and rows as bitmap of column index positions
    - Allows for much higher compression than row based indexes
    - Tabular data compression typically 10x and more!
    - Ideal for leaf level operations
  - Data model is easier to create and manage
  - Lean JSON based object model
- But:
  - Complexity lies in realization of complex calculations
  - Some functionality is still missing and probably will never be realized

And why different modes?

- Power BI
  - Based on Tabular mode
  - Adds visualization features
  - In newer versions also ETL functionality
  - Complete integration into Business application landscape → Power Platform
- But:
  - Support for enterprise requirements (DevOps, ALM) still not satisfactory
  - On-Premise option months behind Cloud solution
  - Not cheap!

Evolution of Analysis Services



Evolution of SQL Server

Are you leveraging the best technology available?

SQL Server 2000

- XML support
- Data mining and data scripting
- Failover clustering

SQL Server 2005

All Capabilities from SQL Server 2000 Release

+

- Database mirroring
- SSAS and ad hoc reporting
- ETL and SSIS
- Service Broker

SQL Server 2008

All Capabilities from SQL Server 2005 Release

+

- Transparent Data Encryption
- Enhanced BI sync framework
- Policy-based management
- Resource governor
- Data compression / backup compression
- New data types (spatial , dates and time)

SQL Server 2008 R2

All Capabilities from SQL Server 2008 Release

+

- PowerPivot for SharePoint
- Report Builder 3.0
- Support for up to 256 logical processors
- Master Data Services
- Multi-server management
- Hyper-V live migration
- Data Tier Application Management (VS Integration)
- Complex event processing
- System prep
- Unicode compression

SQL Server 2012

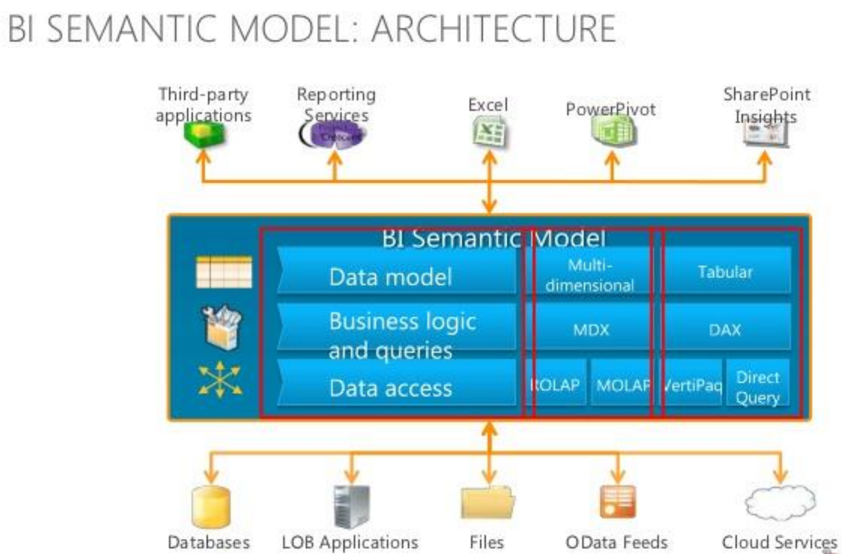
All Capabilities from SQL Server 2008 R2 Release

+

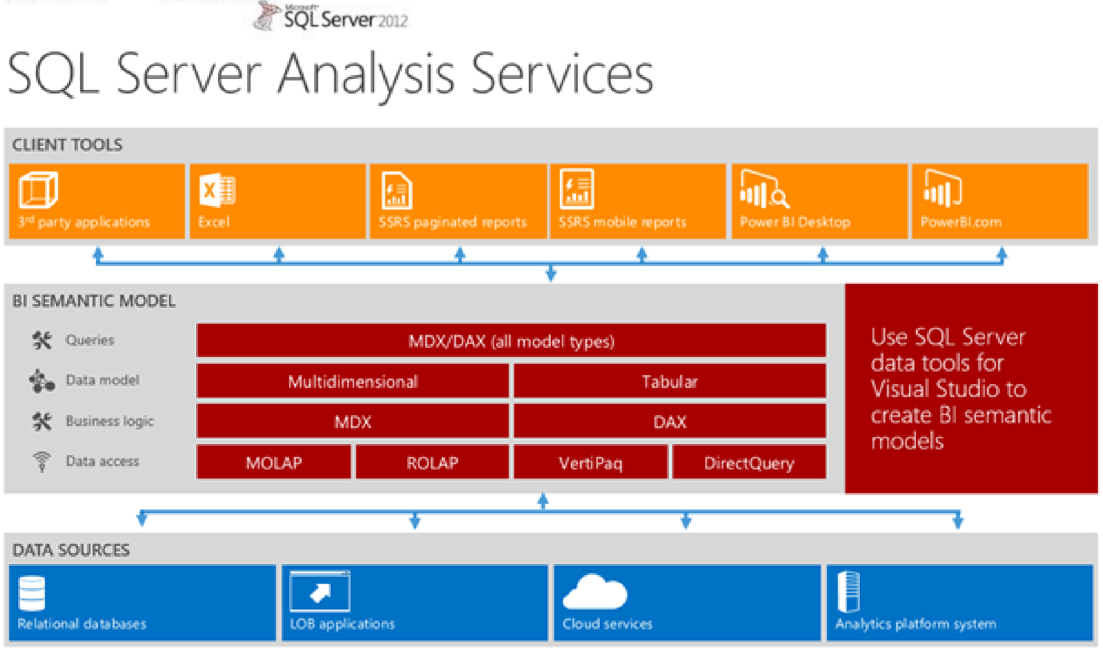
- Power View
- PowerPivot Enhancements
- Single BI semantic model
- Managed self-service analytics
- Transparent data encryption
- Advanced auditing
- Column Store index
- DW compression
- DW partitioning
- SQL Server AlwaysOn

[1] *Source: Microsoft Sales, 2011*

BI Semantic Model 2012 -> 2016



BI SEMANTIC MODEL: ARCHITECTURE



SQL Server Analysis Services

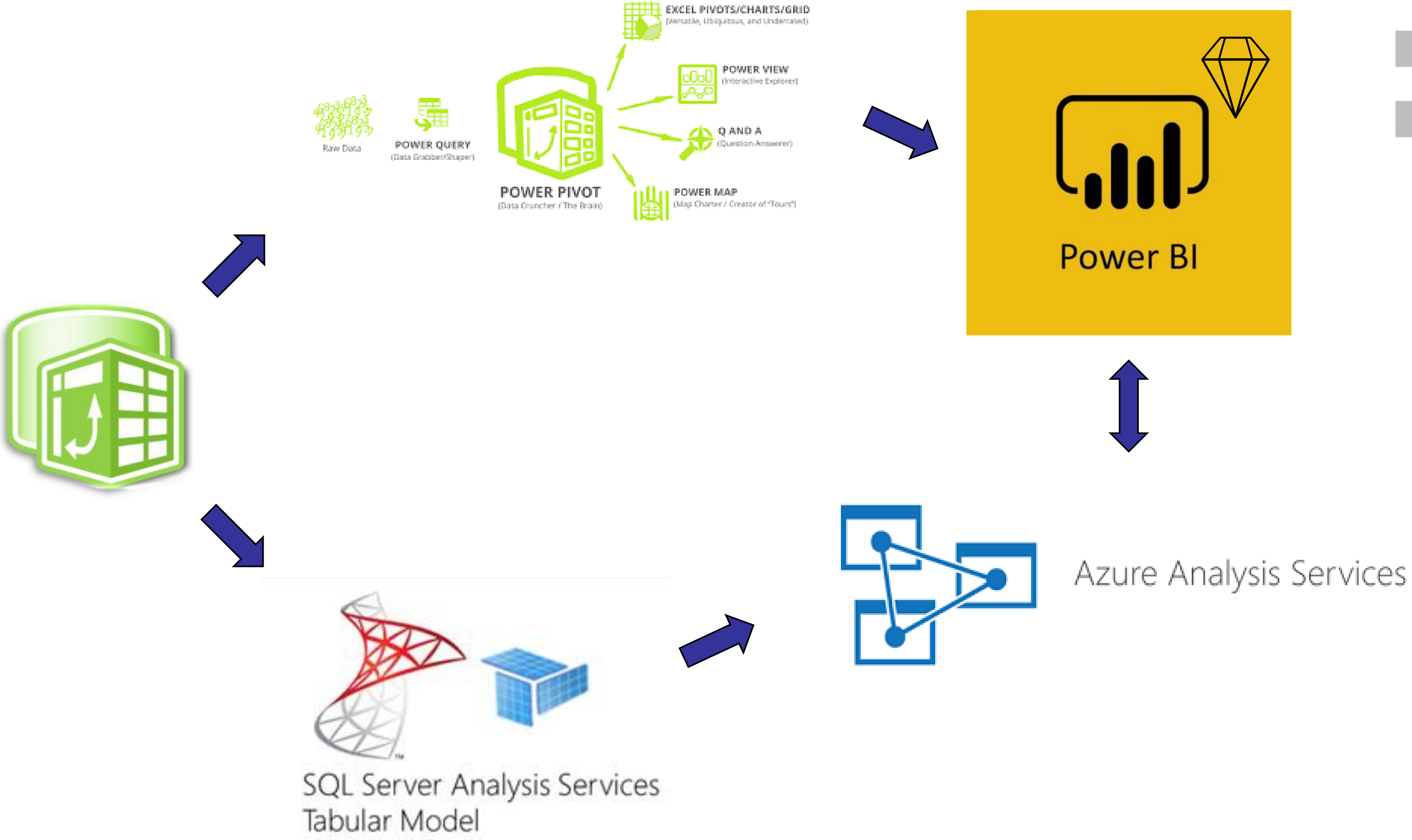
CLIENT TOOLS

BI SEMANTIC MODEL

Use SQL Server data tools for Visual Studio to create BI semantic models

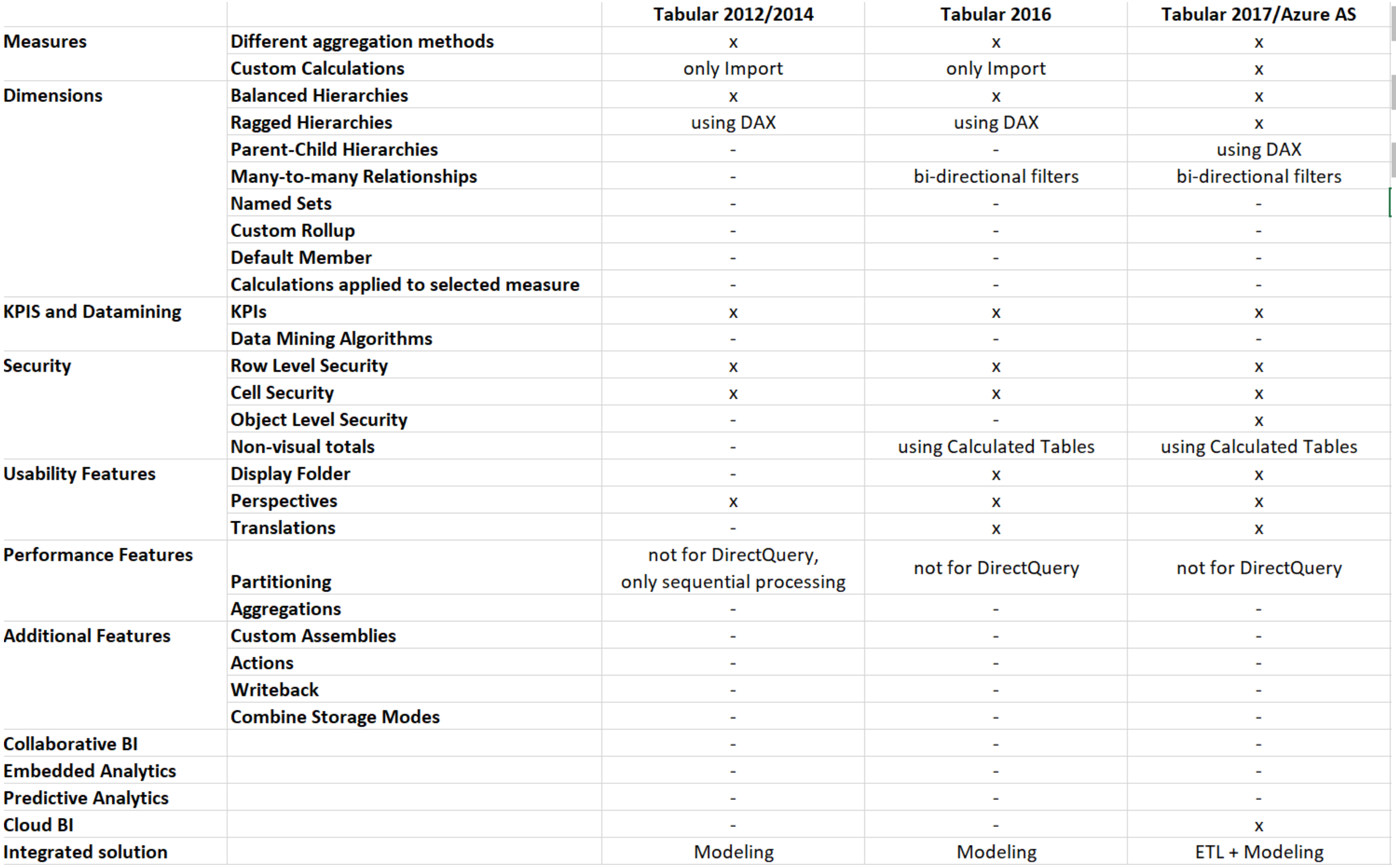
DATA SOURCES

How did SSAS Tabular mode develop over time?



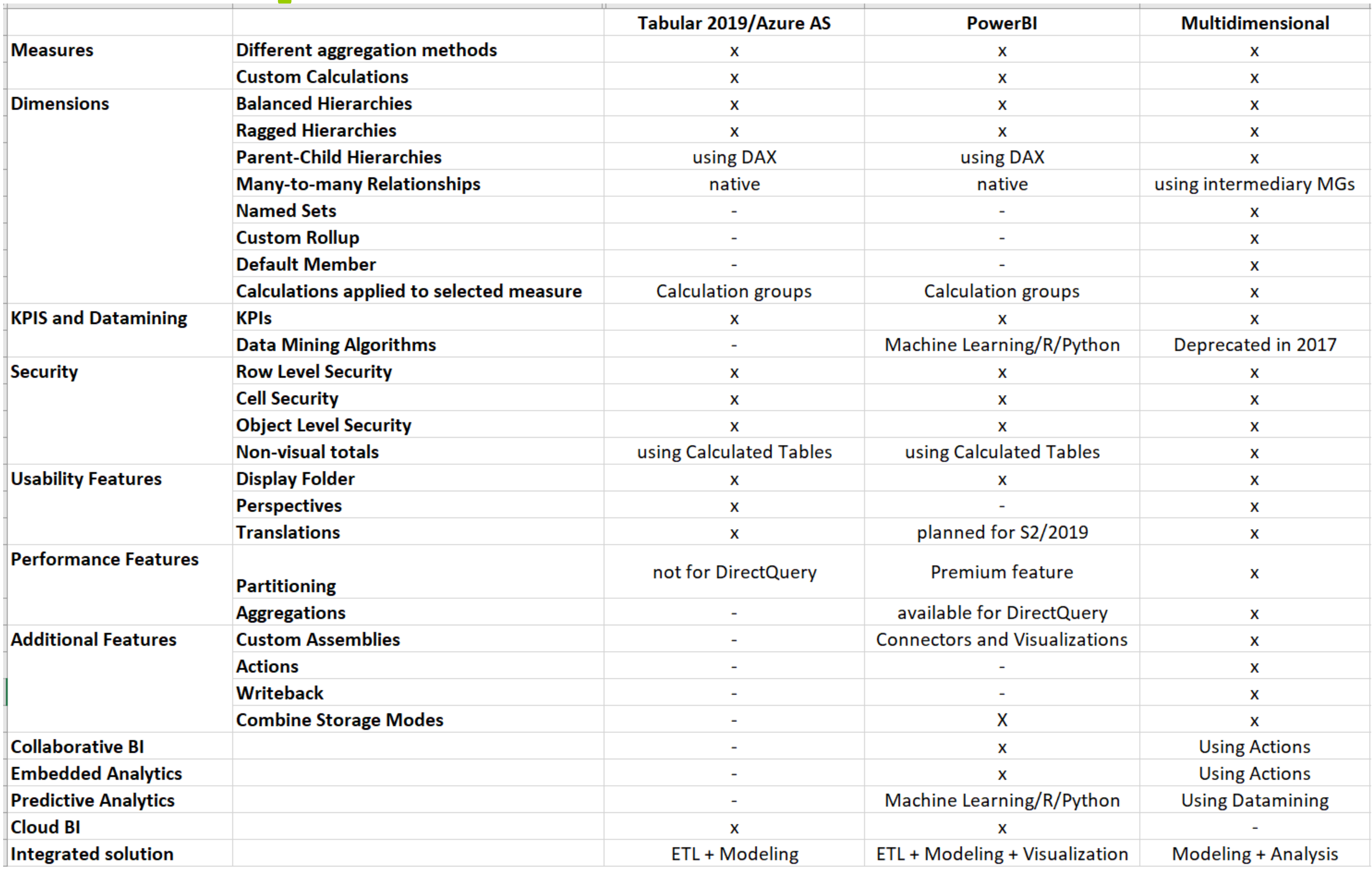
© ppt by OH22

How did SSAS Tabular mode develop over time?



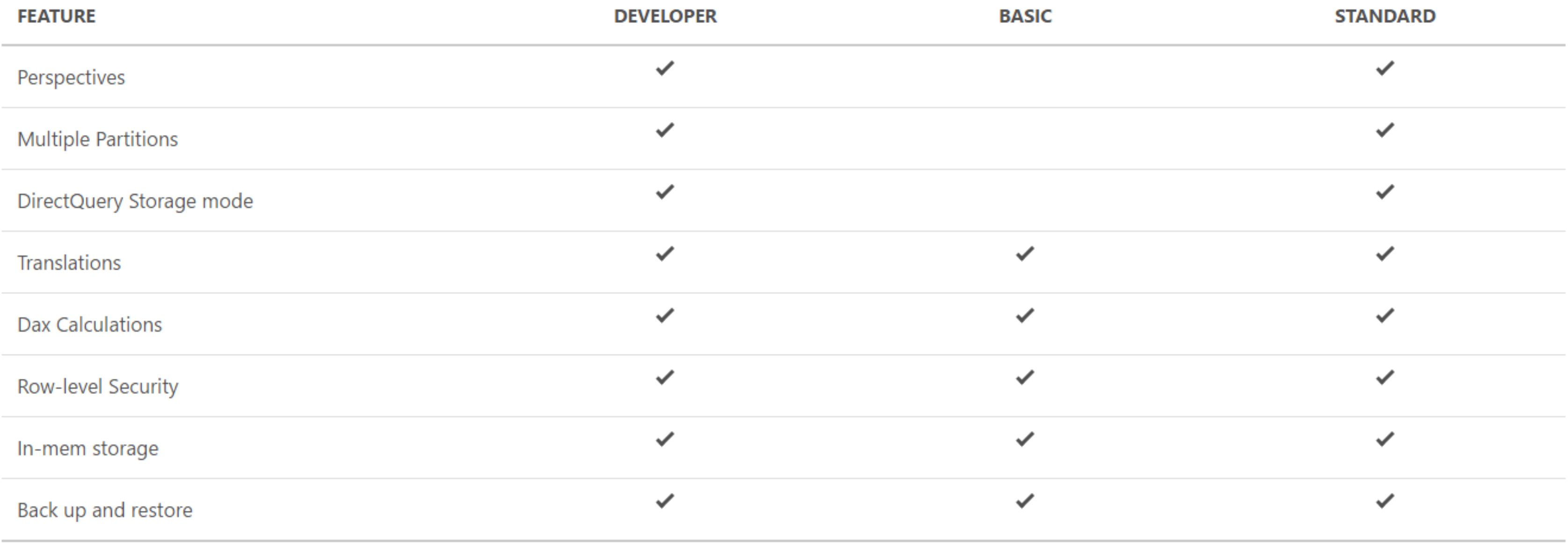
		Tabular 2012/2014	Tabular 2016	Tabular 2017/Azure AS
Measures	Different aggregation methods	x	x	x
	Custom Calculations	only Import	only Import	x
Dimensions	Balanced Hierarchies	x	x	x
	Ragged Hierarchies	using DAX	using DAX	x
	Parent-Child Hierarchies	-	-	using DAX
	Many-to-many Relationships	-	bi-directional filters	bi-directional filters
	Named Sets	-	-	-
	Custom Rollup	-	-	-
	Default Member	-	-	-
	Calculations applied to selected measure	-	-	-
KPIS and Datamining	KPIs	x	x	x
	Data Mining Algorithms	-	-	-
Security	Row Level Security	x	x	x
	Cell Security	x	x	x
	Object Level Security	-	-	x
	Non-visual totals	-	using Calculated Tables	using Calculated Tables
Usability Features	Display Folder	-	x	x
	Perspectives	x	x	x
	Translations	-	x	x
Performance Features	Partitioning	not for DirectQuery, only sequential processing	not for DirectQuery	not for DirectQuery
	Aggregations	-	-	-
Additional Features	Custom Assemblies	-	-	-
	Actions	-	-	-
	Writeback	-	-	-
	Combine Storage Modes	-	-	-
Collaborative BI		-	-	-
Embedded Analytics		-	-	-
Predictive Analytics		-	-	-
Cloud BI		-	-	x
Integrated solution		Modeling	Modeling	ETL + Modeling

How did SSAS Tabular mode develop over time?



		Tabular 2019/Azure AS	PowerBI	Multidimensional
Measures	Different aggregation methods	x	x	x
	Custom Calculations	x	x	x
Dimensions	Balanced Hierarchies	x	x	x
	Ragged Hierarchies	x	x	x
	Parent-Child Hierarchies	using DAX	using DAX	x
	Many-to-many Relationships	native	native	using intermediary MGs
	Named Sets	-	-	x
	Custom Rollup	-	-	x
	Default Member	-	-	x
	Calculations applied to selected measure	Calculation groups	Calculation groups	x
KPIS and Datamining	KPIs	x	x	x
	Data Mining Algorithms	-	Machine Learning/R/Python	Deprecated in 2017
Security	Row Level Security	x	x	x
	Cell Security	x	x	x
	Object Level Security	x	x	x
	Non-visual totals	using Calculated Tables	using Calculated Tables	x
Usability Features	Display Folder	x	x	x
	Perspectives	x	-	x
	Translations	x	planned for S2/2019	x
Performance Features	Partitioning	not for DirectQuery	Premium feature	x
	Aggregations	-	available for DirectQuery	x
Additional Features	Custom Assemblies	-	Connectors and Visualizations	x
	Actions	-	-	x
	Writeback	-	-	x
	Combine Storage Modes	-	X	x
Collaborative BI		-	x	Using Actions
Embedded Analytics		-	x	Using Actions
Predictive Analytics		-	Machine Learning/R/Python	Using Datamining
Cloud BI		x	x	-
Integrated solution		ETL + Modeling	ETL + Modeling + Visualization	Modeling + Analysis

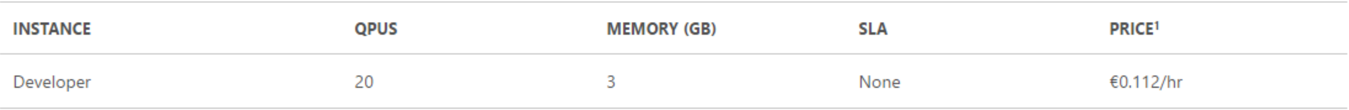
Features and pricing tiers



FEATURE	DEVELOPER	BASIC	STANDARD
Perspectives	✓		✓
Multiple Partitions	✓		✓
DirectQuery Storage mode	✓		✓
Translations	✓	✓	✓
Dax Calculations	✓	✓	✓
Row-level Security	✓	✓	✓
In-mem storage	✓	✓	✓
Back up and restore	✓	✓	✓

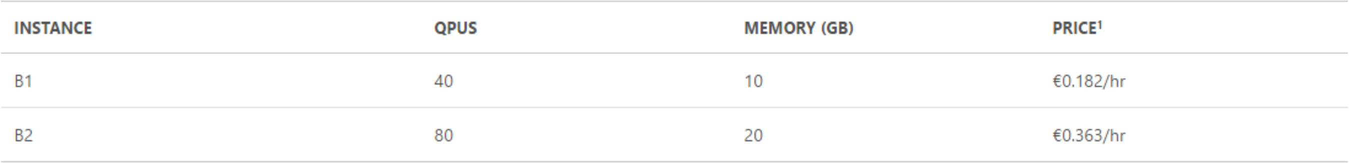
Pricing tiers

Developer



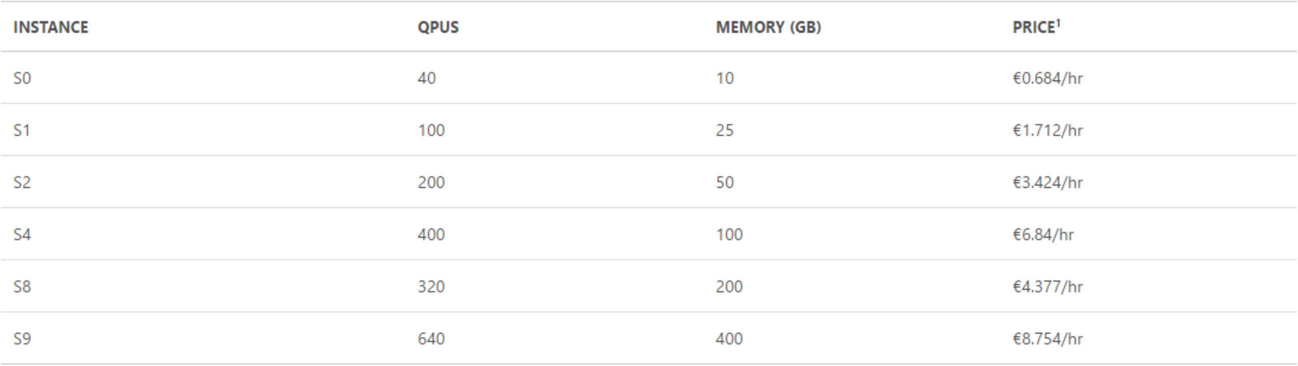
INSTANCE	QPUS	MEMORY (GB)	SLA	PRICE[1]
Developer	20	3	None	€0.112/hr

Basic



INSTANCE	QPUS	MEMORY (GB)	PRICE[1]
B1	40	10	€0.182/hr
B2	80	20	€0.363/hr

Standard



INSTANCE	QPUS	MEMORY (GB)	PRICE[1]
S0	40	10	€0.684/hr
S1	100	25	€1.712/hr
S2	200	50	€3.424/hr
S4	400	100	€6.84/hr
S8	320	200	€4.377/hr
S9	640	400	€8.754/hr

Latest releases and roadmap

- GA:
  - Governance for cache refresh settings
- Public Preview:
  - Many-to-many relationships
  - Calculation Groups (see also Calculation groups (Preview))
  - Dynamic Format Strings (CTP 3.0)
- Roadmap:
  - Query interleaving

Power BI roadmap

- Power BI in the 2nd semester of 2019:
  - Data lineage
  - Data protection
  - Monitoring features
  - Customization features
  - AI/ML features
  - …
  - https://www.youtube.com/watch?v=fQOPX--k2i8
  - https://aka.ms/ROGPBI19RW2RP

Comparison to PowerBI (current)

	PowerBI Pro	PowerBI Premium/Embedded	Azure Analysis Services
Model/Data Set Size	1 GB	Limited by capacity node size Current max: 100 GB (initial upload restricted to 10 GB)	based on pricing tier (current max: 400 GB)
Scheduled refresh times per day	8	48	not restricted
Partitioning/Incremental refresh	No	Yes	Yes (not Basic)
Scalability	No	Yes	Yes
Cost management by pausing service	No	Premium: No Embedded: Yes	Yes
Usage of development toolset	Power BI Desktop	Power BI Desktop	SSDT
Integrate into development strategies like continuous integration	No	No	Yes
Extend modelling to functionality of multidimensional	No	No	No

Comparison to PowerBI (Roadmap)

	PowerBI Pro	PowerBI Premium/Embedded	Azure Analysis Services
Model/Data Set Size	1 GB	Limited by capacity node size Current max: 100 GB (initial upload restricted to 10 GB)	based on pricing tier (current max: 400 GB)
Scheduled refresh times per day	8	48	not restricted
Partitioning/Incremental refresh	No	Yes	Yes (not Basic)
Scalability	No	Yes	Yes
Cost management by pausing service	No	Premium: No Embedded: Yes	Yes
Usage of development toolset	Power BI Desktop	Power BI Desktop SSDT will allow publishing of models	SSDT
Integrate into development strategies like continuous integration	?	Yes	Yes
Extend modelling to functionality of multidimensional	?	?	?

WHICH LANGUAGES DO WE NEED FOR WHICH PURPOSE?

Languages

- Multidimensional
  - Expressions
    - MDX
  - Querying
    - MDX
    - DAX
  - Modeling
    - XMLA (XML based)
    - AMO (C#)

- Tabular / Power BI
  - Expressions
    - DAX
  - Querying
    - MDX
    - DAX
  - Modeling
    - TMSL (JSON based)
    - TOM (C#)
    - Power Query (M)

What is DAX? And why do you need it for Tabular mode?

- DAX = Data Analysis Expressions
- Similar to Excel Functions
- Used for querying as well as programming
- Instead of navigating through a cube like MDX, DAX is working on tables and relationships and therein evaluating contexts (row and filter)

DAX Expression Syntax

- Always start with an =
- You are working with
  - Tables: Sales, ‘Internet Sales’
  - Columns: [Name], Country[Name]
  - Measures: [Profit], ‘Internet Sales’[Profit]
  - Functions: VALUES(), ALL()
  - Operators: +, >, ||
- You can define parameters for queries
- SSAS 2016 introduced variables

Basic DAX expressions

- = 3 *(scalar constant)*
- = Sales *(reference to a table)*
- = Sales[Amount] *(if used in context of sales table = value of column amount for current row)*
- = Sales[Sales Amount] * 0.3 *(if used in context of sales table = value of column amount for current row multiplied by 0.3)*

DAX query syntax

```
[DEFINE {  MEASURE <tableName>[<name>] = <expression> }

EVALUATE <table>

[ORDER BY {<expression> [{ASC | DESC}]}[, …]
[START AT {<value>|<parameter>} [, …]]]
```

Basic DAX queries

-
  ```
  EVALUATE
  ADDCOLUMNS(
     VALUES(DimChannel[ChannelName])
     ,"Sum Of Sales Amount",
  [SumOfSalesAmount]
     )
  ```

Tools & Blogs

- Tools:
  - Since SQL 2017 you‘ll have syntax support in SSDT
  - DAX editor (syntax and formatting support, query plans, execution times etc.)

- Blog recommendations:
  - http://www.sqlbi.com/
  - https://blog.crossjoin.co.uk/

Important DAX concepts

- Filter context
  - Take a pivot table showing years on columns, products on rows and sales amount as value. Each value cell is showing the sales amount for the given combination of year and product.
  - Filter context ist defined by rows, columns, slicers and filters
- Row context
  - E.g. for expressions used in calculated columns
  - Tabular identifies values by evaluating the current row context
  - Like a cursor running over all rows and doing calculations based on current position

Important DAX concepts

- Calculated columns/tables
  - Calculations based on row context/table definition
  - Calculations are executed after processing and are therefore security relevant!
  - Calculated columns can be used to slice and dice
- Measures
  - Measures are calculation rules executed during runtime and based on the user-defined context
  - Can only be used as values
  - Allow for calculation of percentages and ratios based on given context

Important DAX concepts

- Data Types (Related SQL data types)
    - Whole number (INT 64)
    - Decimal number (Float)
    - Currency (Currency)
    - Date (DateTime)
    - Boolean (TRUE/FALSE)
    - Text (String)
    - Binary large object (BLOB)
- Be aware of type-handling and type casts
    - Operator overloading
        - Date value + Number value = Date value
    - Operator related type casts
        - 5 & 3 => 53 as Text
        - „5“ + „3“ => 8

Important DAX concepts

- Operators
    - Parenthesis: ()
    - Arithmetic: +, -, *, /
    - Comparison: =, <>, >, …
    - Text concatenation: &
    - Logical: &&, ||

Important DAX concepts

- Aggregations and Iterators
  - Aggregation functions calculate the aggregated value for a given filter context
    - Beware: SUM(Sales[Sales Amount] as expression in calculated column has no filter context and therefore will return the total Sales Amount for each row.
    - SUM, MAX, MIN, AVG
  - Iterators are used to create a row context for the aggregation function
    - SUMX(Sales[Sales Amount] as expression in calculated column will regard the current row context as filter context and therefore will return the related Sales Amount for each row.
    - SUMX, MAXX, MINX, AVGX

Kinds of DAX functions

- Aggregate functions (SUM, MIN, MAX, …)
- Logical functions (AND, NOT, IF, IFERROR, …)
- Information functions (ISNUMBER, ISBLANK, …)
- Mathematical functions (ABS, MOD, FLOOR, …)
- Trigonometric functions (COS, SINH, RADIANS,..)
- Text functions (CONCATENATE, LEFT, LEN, …)
- Conversion functions (FORMAT, VALUE, INT, …)
- Date and time functions (Date, Day, Today, …)
- Relational functions (Related, Relatedtable)
- Table functions (CALCULATETABLE, VALUES, …)

The developer perspective

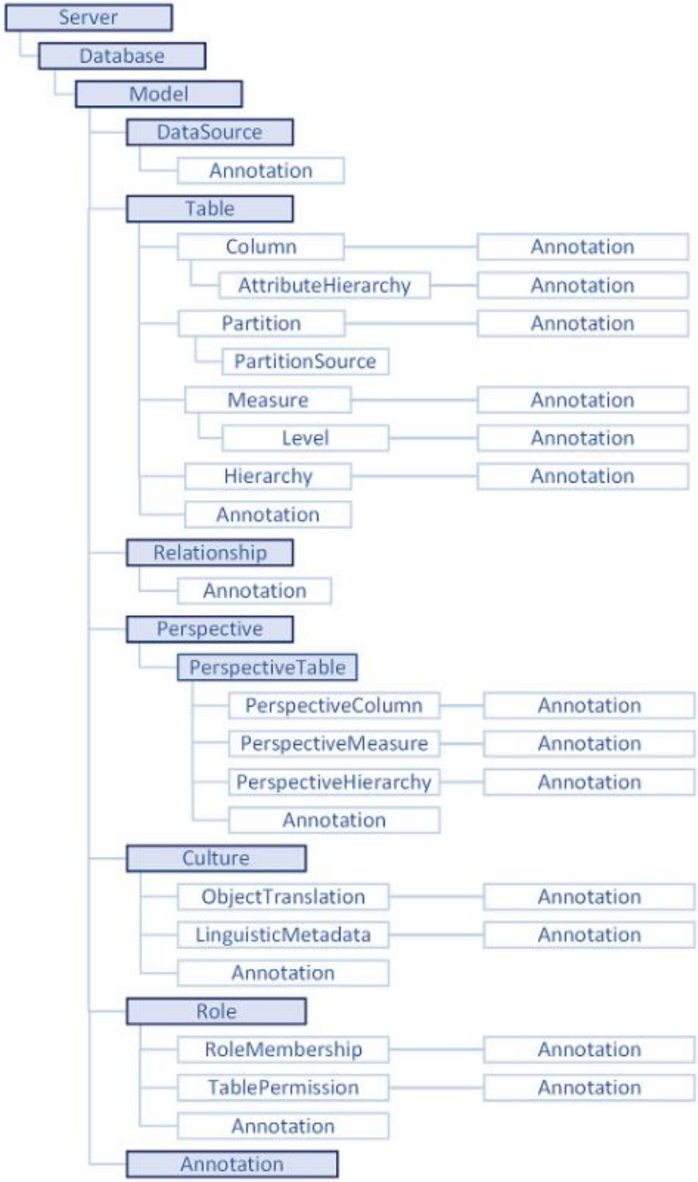
- TMSL (Tabular Model Scripting Language) $ AMO-TOM (Analysis Management Objects – Tabular Object Model)
- AMO is a C# library for creating and managing Multidimensional databases. With SQL Server 2016 a new library representing the Tabular Object Model (TOM) has been added to the AMO library.
- For the communication with the server TOM actions are translated into TMSL.

The developer perspective: TMSL

- JSON based
- Object operations:
  - Create, CreateOrReplace,Alter…
- Data refresh operations:
  - MergePartitions, Refresh
- Database management:
  - Attach/Detach, Backup/Restore, Synchronize

The developer perspective: TOM

- C# library
- Requires .NET 4.0 runtime
- Possible use cases:
  - Schema compare (BISM normalizer)
  - Continuous integration
  - Dynamic deployment



Server
Database
Model
DataSource
Annotation
Table
Column
Annotation
AttributeHierarchy
Annotation
Partition
Annotation
PartitionSource
Measure
Annotation
Level
Annotation
Hierarchy
Annotation
Annotation
Relationship
Annotation
Perspective
PerspectiveTable
PerspectiveColumn
Annotation
PerspectiveMeasure
Annotation
PerspectiveHierarchy
Annotation
Annotation
Culture
ObjectTranslation
Annotation
LinguisticMetadata
Annotation
Annotation
Role
RoleMembership
Annotation
TablePermission
Annotation
Annotation
Annotation

Useful links (I)

- PowerPivot:
  - Overview: https://technet.microsoft.com/en-us/library/gg413471(v=sql.105).aspx
  - SharePoint: https://technet.microsoft.com/en-us/library/ee210692(v=sql.105).aspx
  - Tutorial: https://technet.microsoft.com/en-us/library/gg413497(v=sql.105).aspx and http://powerpivotsdr.codeplex.com/releases/view/35438

- Tabular Mode
  - Overview: https://technet.microsoft.com/en-us/library/hh212945(v=sql.110).aspx
  - Tutorial: https://docs.microsoft.com/en-us/sql/analysis-services/tabular-modeling-adventure-works-tutorial

Useful links (II)

- DAX links and books:
  - Reference: https://msdn.microsoft.com/en-us/library/gg413422.aspx
  - Parameters: http://sqlblog.com/blogs/marco_russo/archive/2012/01/05/use-parameters-in-your-dax-queries.aspx
  - Variables: https://www.sqlbi.com/articles/variables-in-dax/
  - Detail Rows: https://blog.crossjoin.co.uk/2017/01/15/more-detail-on-detail-rows-expressions-in-ssas-tabular-v-next/
  - Book recommendation: https://www.amazon.de/Definitive-Guide-DAX-Intelligence-Microsoft/dp/073569835X

Useful links (III)

- SQL Server 2016
  - What's New in Analysis Services 2016
  - TechNet Virtual Lab: Exploring What's New in SQL Server 2016 Analysis Services
  - SQL Server 2016 Release -tes
  - Editions and Supported Features for SQL Server 2016
  - Analysis Services Features Supported by the Editions of SQL Server 2016

- SQL Server 2016 SP1
  - SQL Server 2016 Service Pack 1 (SP1) released !!!
  - SQL Server 2016 SP1: K-w your limits
  - SQL Server 2016 SP1 Editions Datasheet

Useful links (IIII)

- Azure Analysis Services
  - Azure Analysis Services Preview
- SQL Server vNext
  - What's new in SQL Server vNext?
  - What's new in SQL Server Analysis Services vNext
  - Encoding Hints and SQL Server Analysis Services vNext CTP 1.3
  - SQL Server vNext Release -tes