

آشنایی با محیط SSAS Tabular

ارائه دهنده: سینا سبزواری

بهار ۱۴۰۲

آشنایی با ارائه دهنده دوره



✓مدیر هوشمندی کسب و کار بخش فروش و بازاریابی دیجی کالا

✓معمار زیرساخت هوش تجاری شرکت اسنپ

✓مدیر پروژه داده در شرکت فرابورس ایران

✓مشاور شرکت های بیمه ای، فولادی و توریستی

مروری بر این بخش

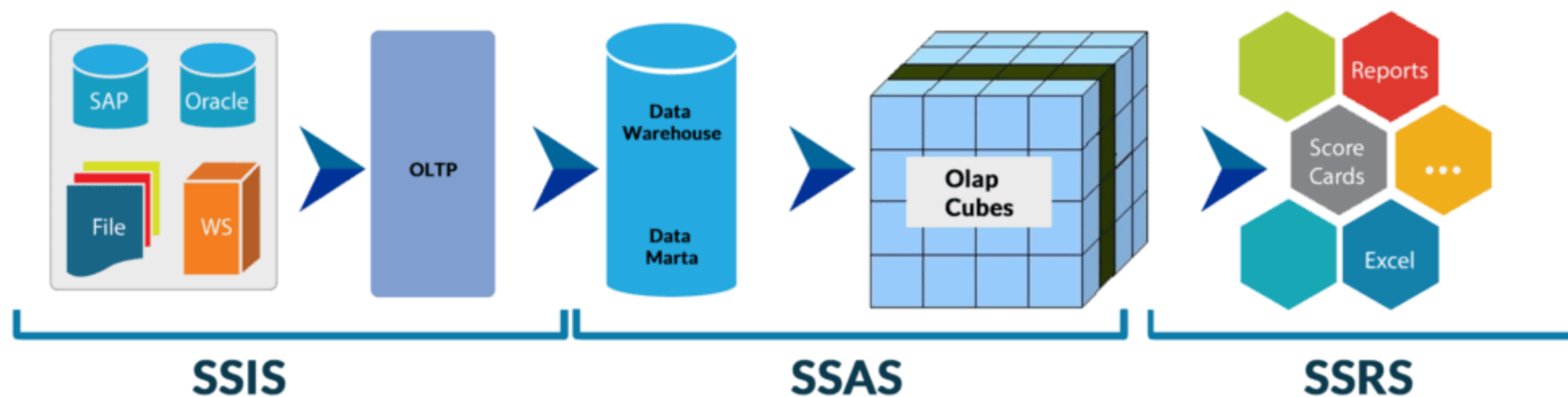
✓ آشنایی با SSAS

✓ آشنایی با نحوه نصب و راه اندازی SSAS Tabular

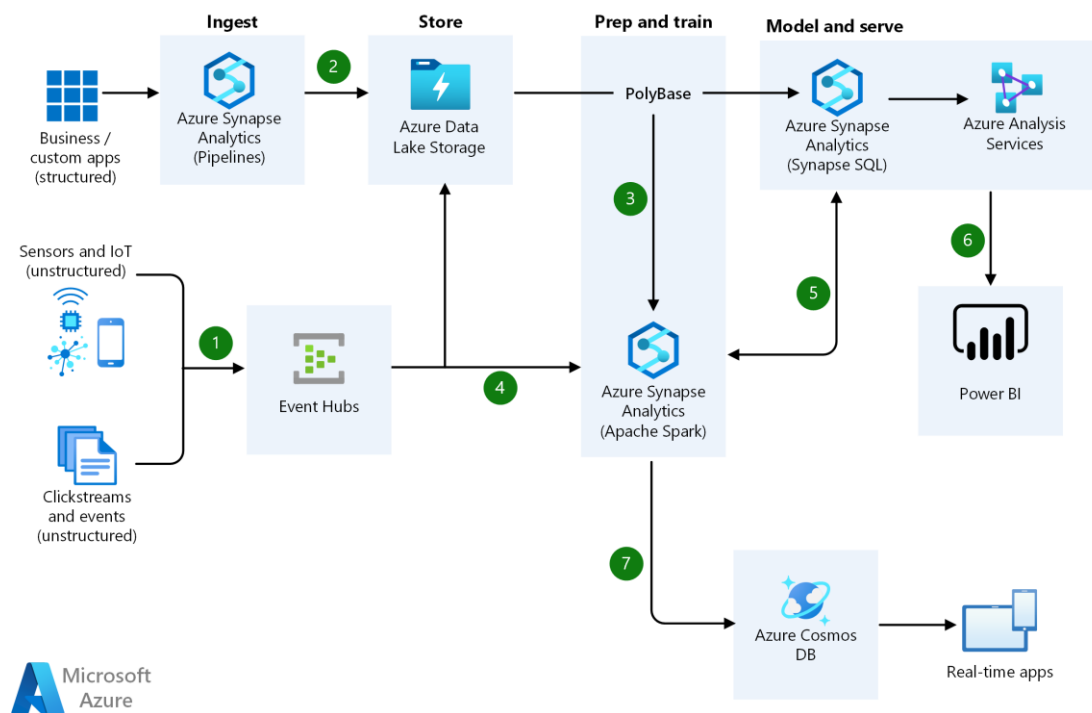
✓ آشنایی با پروژه، تنظیمات و ساختار فایل های پروژه و مدل



مرور کلی بروی معماری قدیم Microsoft BI



مرور کلی بروی معماری جدید Microsoft BI



آشنایی با SQL Server Analysis Service



✓ پایگاه داده مخصوص گزارش گیری

✓ ساختار داده مناسب برای گزارش گیری مانند Hierarchy

✓ ایجاد زیرساخت های مناسب برای افزایش سرعت در گزارش گیری

✓ امکان دادن به کاربران کسب و کار برای ایجاد گزارشات

✓ زبانی مناسب برای کندوکاو در داده ها (MDX & DAX)

✓ دارای ساختارهای توکار برای افزایش سرعت گزارش گیری (ایندکس) و فشرده سازی

آشنایی با SQL Server Analysis Service

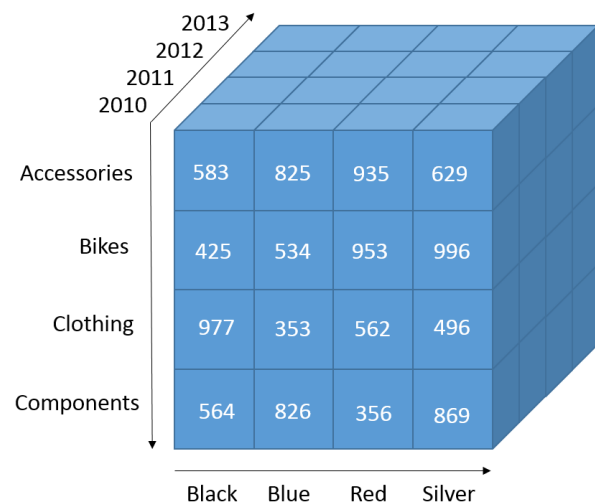
✓ ارائه در سال ۱۹۹۸ در SQL Server 7

✓ ارائه زبان MDX بر اساس ساختار سلسه مراتبی و Slice & Dice

✓ پرفروش ترین ابزار تحلیل داده در سال ۲۰۰۵

✓ رسیدن به حجم ۲۴ ترابایت در یاهو (در سال ۲۰۱۰)

✓ قابلیت استفاده در سیستم با یک یا دو هسته پردازشی و کمتر از یک گیگابایت RAM



مشکلات Multidimensional



✓ پیچیدگی بیش از اندازه زبان MDX

✓ پیچیدگی زیاد مدل سازی

✓ کاهش زیاد قیمت سخت افزار در طول چندین سال

✓ ظهور رقابایی مانند Qlik و Tableau و فراگیر شدن مفاهیم Self Service BI

آشنایی با Tabular



✓ ارائه نسخه ابتدایی Analysis Service Tabular در سال ۲۰۱۲

✓ معرفی Engine جدید Vertipaq

✓ ذخیره سازی داده به صورت Column Oriented

✓ فشرده سازی بسیار زیاد (۳ تا ۱۰ برابری برای داده Analytical)

✓ معرفی زبان DAX (Data Analysis Expressions)

✓ ساده سازی مدل سازی (کار با Table به جای Cube) مانند سایر ابزارهای Self Service BI

✓ یکپارچه سازی با Power Query برای تمیزسازی داده

✓ برای کوئری های تحلیلی (Analytical) ساخته شده است. برای کوئری با جزئیات زیاد از SQL Server استفاده شود

مشکلات سیستم Tabular

✓ سیستم تبولار در حجم داده ای کم (زیر ۱۰۰ میلیون رکورد) نیاز به بهینه سازی کمتری دارد
اما در حجم های بالاتر بهینه سازی کم شایانی به افزایش عملکرد آن می کند

✓ زبان DAX نوشتن کوئری با منطق ساده مانند جمع، میانگین و بسیار ساده است
اما در کوئری های پیچیده مانند دسترسی به نقاط مختلف درخت نسبتا پیچیده می شود.

✓ شناخت زبان M زمانبر است.



ابزارهای اولیه برای شروع کار



نصب ابزارهای زیر:

Visual Studio 2022 ✓

SQL Server 2022 ✓

Analysis Services project Extension ✓

SQL Server Management Studio ✓

DAX Studio ✓

بازیابی دیتابیس های زیر

Contoso ✓

AdventureWorksDW2017 ✓

نکات نصب در SQL Server



✓ ساخت Managed Service Account برای سرویس Analysis Service

✓ تنظیمات مربوط به Data, Log, Temp, Backup

در هنگام نصب

در Sever Properties

در فایل کانفیگ msmdsrv

✓ تغییر Account مربوط به سرویس Analysis Service

دمو نصب ابزارها



✓ نصب SQL Server

✓ نصب Visual Studio

✓ نصب SSMS

✓ نصب DAX Studio

✓ ایجاد دیتابیس Contoso

✓ ساخت Managed Service Account برای سرویس Analysis Service

✓ بررسی ساخت Managed Account برای Analysis Service

✓ بررسی سمت Server پوشه های Analysis Service

بررسی اولیه سرویس سمت سرور

✓ بررسی پوشه های Analysis Services

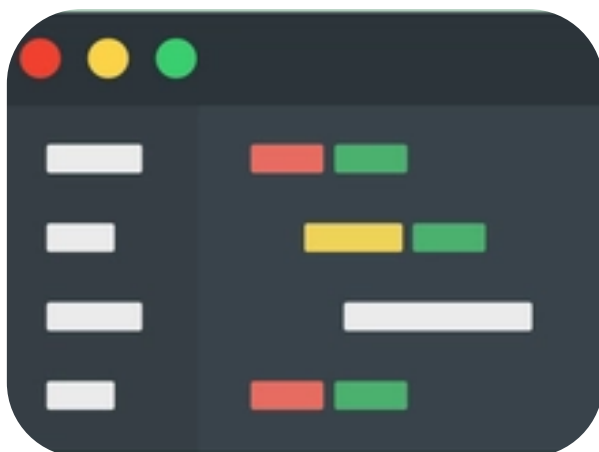
✓ ساخت Managed Service Account برای سرویس Analysis Service

✓ بررسی ساخت Managed Account برای Analysis Service

✓ بررسی سمت Server پوشه های Analysis Service



آشنایی با محیط SSDT در محیط توسعه



✓ دیتابیس محیط توسعه

❖ Integrated

❖ Workspace Server

➤ امکان دسترسی به سایر افراد برای بررسی پروژه

➤ منابع بیشتر بروی سرور و امکان ارتقا ساده تر سخت افزار بروی سرور

➤ امکان Process کردن مدل همزمان کار بروی آن

➤ امکان اتصال ابزارهای Client مانند PowerBI بروی آن

✓ WYSIWYG (What you see is what you get)

دمو ایجاد پروژه



✓ آشنایی با ایجاد پروژه

✓ آشنایی با ویژگی های پروژه

✓ آشنایی با ویژگی های مدل

✓ آشنایی با فایل های ایجاد شده سمت پروژه

✓ آشنایی با ساختار فایل ایجاد شده سمت Analysis Service

Workspace Retention

پسوند فایل	توضیحات
Keep in memory	بعد از بسته شدن همچنان پروژه را در حافظه و دیسک نگه می دارد. (سریعترین روش بارگزاری)
Unload from memory	بعد از بسته شدن پروژه داده را از روی حافظه برمی دارد اما همچنان روی دیسک نگه می دارد.
Delete workspace	داده را از روی حافظه و هم دیسک حذف می کند

خلاصه آشنایی با فایل های پروژه

پسوند فایل	توضیحات
bim	متا دیتای پروژه
asdatabase	متادیتای پروژه
deploymentoptions	Deployment Option
deploymenttargets	سروری که پروژه روی آن Deploy می شود
layout	اندازه و سایز پنجره ها در Visual Studio
abf	پسوند Backup

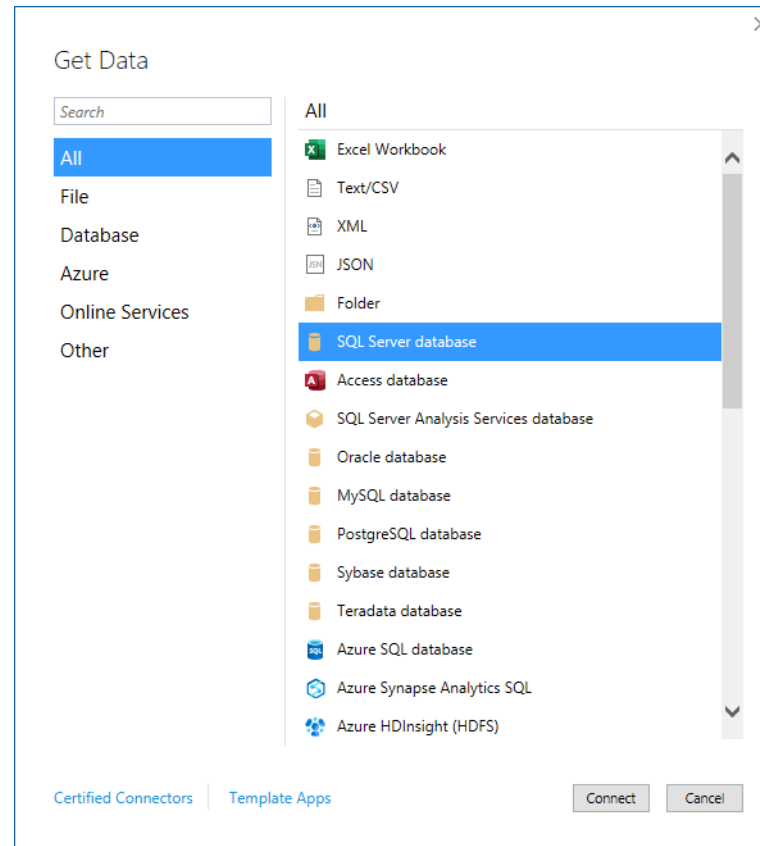
دمو آشنایی با محیط Visual Studio

✓ آشنایی با Grid

✓ آشنایی با Diagram



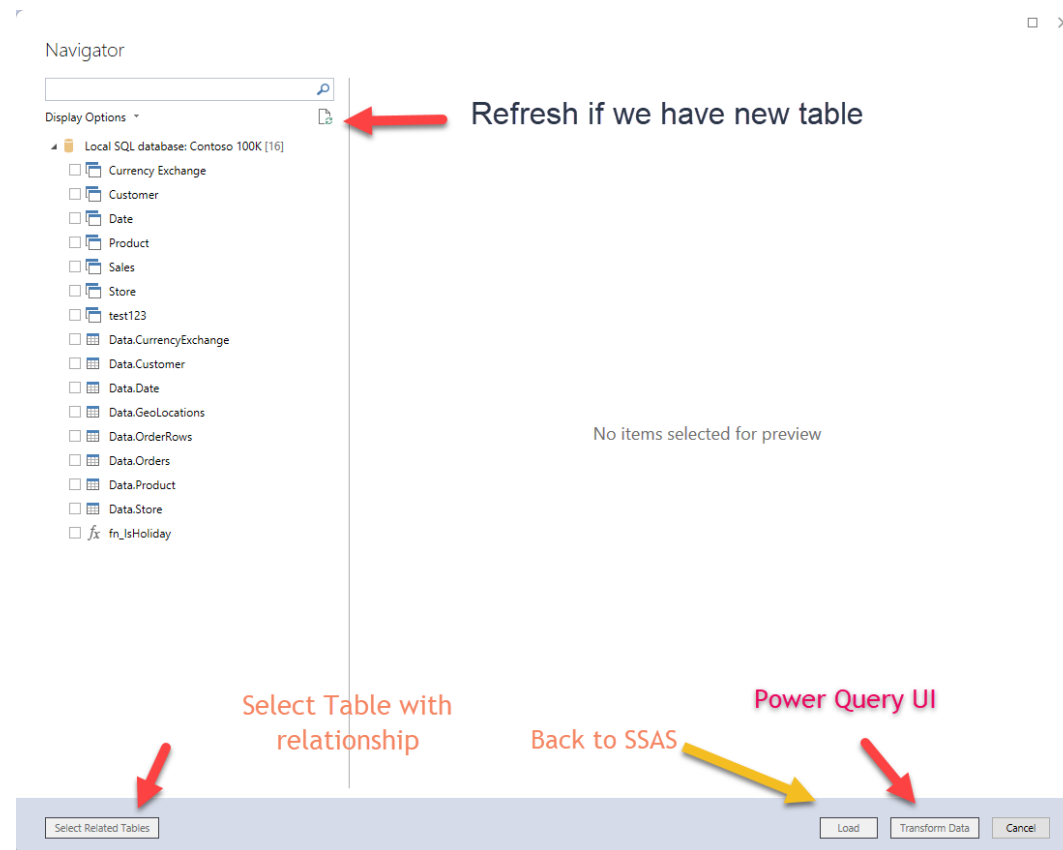
Import Data



Impersonation

پسوند فایل	توضیحات
Impersonate Account	از طریق دسترسی یک اکانت ویندوزی وارد می شود
Impersonate Identity	مربوط به دسترسی از طریق Kerberos است.
Impersonate Service Account	از طریق دسترسی سرویس Analysis Service به سورس داده متصل میشود
Impersonate Unattended Service Account	مربوط به دسترسی از طریق Kerberos است.

Navigator



آشنایی با آماده سازی داده سمت انبار داده



✓ برای وارد کردن اطلاعات سعی کنید بروی جداول حتما View بسازید:

✓ فقط ستون هایی که لازم هستند را در View بنویسید

✓ تنها قسمتی از داده مربوط به جداول Fact را لود کنید تا سرعت افزایش یابد.

✓ داده های رشته ای را به صورت Unicode تبدیل کنید

دمو وارد کردن داده از SQL Server



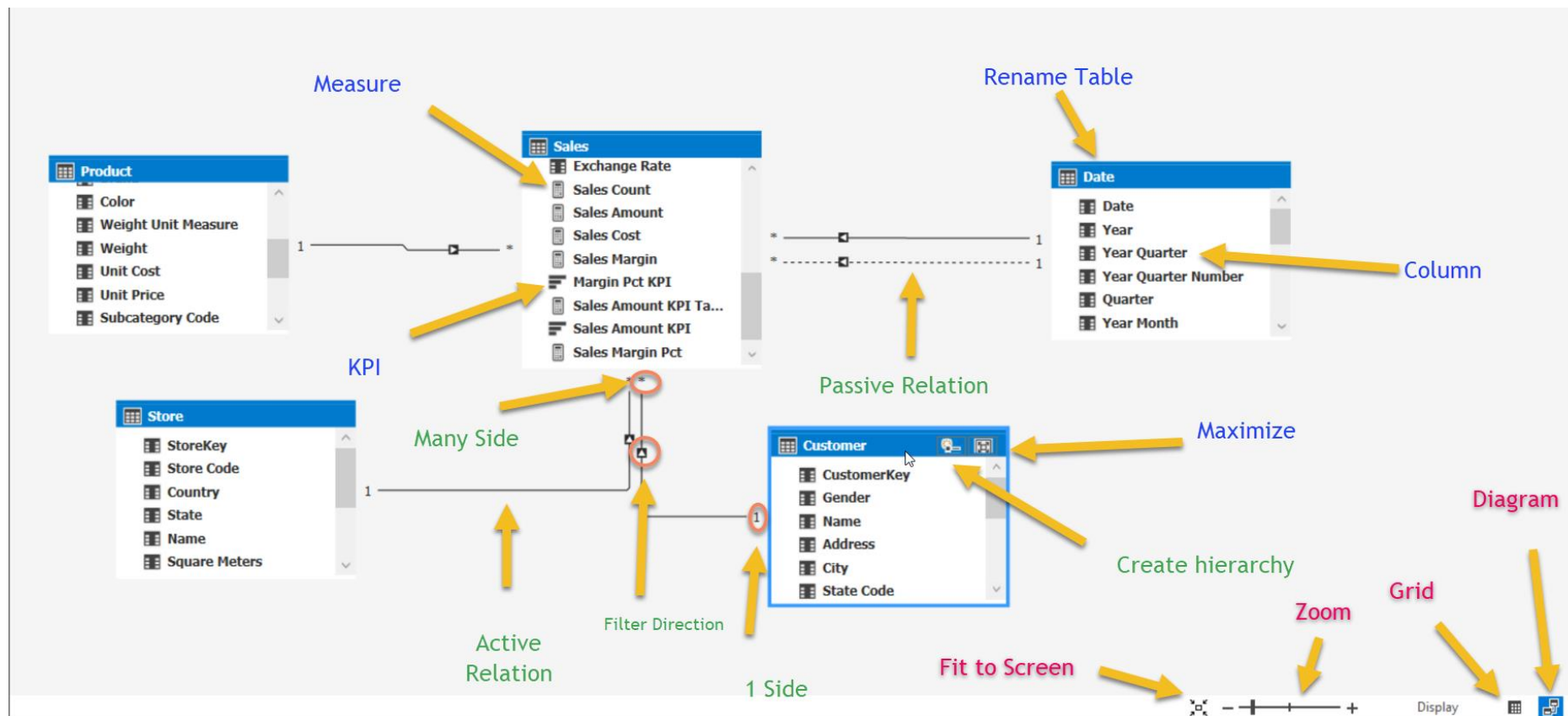
✓ آشنایی با Import کردن داده

✓ آشنایی با قسمت های مختلف Navigator

✓ آشنایی با انواع Impersonation

✓ آشنایی با آماده سازی داده سمت انبار داده

آشنایی با Diagram



آشنایی با Grid

The screenshot displays the SSAS Tabular Grid interface. At the top, a 'Formula' bar shows '[StoreKey]' and '999999'. Below it, a 'Column List' is visible. The main area is a data grid with columns: Order Number, Line Number, Or..., Delivery Date, Custom..., Stor..., ProductKey, Quantity, Unit Price, Net Price, Currency Code, Exchange..., and Add Column. A context menu is open over the grid, showing options like 'Sort Smallest to Largest', 'Sort Largest to Smallest', 'Clear Sort From "StoreKey"', 'Clear Filter From "StoreKey"', 'Number Filters' (with a sub-menu showing checkboxes for values like 10, 20, 30, 35, 40, 50, 60), 'Equals...', 'Does Not Equal...', 'Greater Than...', 'Greater Than Or Equal To...', 'Less Than...', 'Less Than Or Equal To...', 'Between...', and 'Custom Filter...'. Annotations include: 'Sort' pointing to the context menu, 'Filter' pointing to the 'Number Filters' sub-menu, 'Add calculated Column' pointing to the 'Add Column' button, 'Measure' pointing to the 'Sales Margin Pct' cell, 'KPI' pointing to the 'Margin Pct KPI: 65.85%' cell, 'Calculated Table' pointing to the 'Sales Amount KPI Target: 305891102.92' cell, 'Add Calculated Table' pointing to the 'Sales Amount KPI' cell, 'Grid' pointing to the main data grid, and 'Navigate rows' pointing to the 'Record: 1 of 208,969' navigation bar at the bottom.

Order Number	Line Number	Or...	Delivery Date	Custom...	Stor...	ProductKey	Quantity	Unit Price	Net Price	Currency Code	Exchange...	Add Column	
1	1134016	0	2/7/2013 ...	2/13/2013		1790	1	\$38.70	\$38.70	USD	1		
2	1329023	0	8/21/201...	8/24/2013		1793	1	\$38.70	\$38.70	USD	1		
3	1355006	0	9/16/201...	9/21/2013		1791	1	\$38.70	\$38.70	USD	1		
4	1774003	0	11/9/201...	11/13/2014		1783	1	\$38.70	\$38.70	USD	1		
5	1785006	0	11/20/20...	11/29/2014		1791	1	\$38.70	\$38.70	USD	1		
6	2063000	0	8/25/201...	8/27/2015				\$38.70	\$38.70	USD	1		
7	2119007	0	10/20/20...	10/23/2015				\$38.70	\$38.70	USD	1		
8	2148006	0	11/18/20...	11/24/2015				\$38.70	\$38.70	USD	1		
9	2258005	0	3/7/2016 ...	3/11/2016				\$38.70	\$38.70	USD	1		
10	2347004	0	6/4/2016 ...	6/6/2016				\$38.70	\$38.70	USD	1		
11	2404000	0	7/31/201...	8/4/2016				\$38.70	\$38.70	USD	1		
12	2442023	0	9/7/201...	9/9/2016				\$38.70	\$38.70	USD	1		
13	2531011	0	12/5/201...	12/10/2016				\$38.70	\$38.70	USD	1		
14	2589011	0	2/1/2017 ...	2/3/2017				\$38.70	\$38.70	USD	1		
15	2780029	0	8/11/201...	8/19/2017	1200000	999999		\$38.70	\$38.70	USD	1		
16	2788019	0	8/19/201...	8/20/2017 12:00:00 AM	1454163	999999	1776	1	\$38.70	\$38.70	USD	1	
17	2805001	0	9/5/2017 ...	9/7/2017 12:00:00 AM	1453776	999999	1774	1	\$38.70	\$38.70	USD	1	
18	2813039	0	9/13/201...	9/17/2017 12:00:00 AM	1432379	999999	1796	1	\$38.70	\$38.70	USD	1	
19	2863051	0	11/2/201...	11/5/2017 12:00:00 AM	1701609	999999	1797	1	\$38.70	\$38.70	USD	1	
20	2921055	0	12/30/20...	1/3/2018 12:00:00 AM	1761910	999999	1773	1	\$38.70	\$38.70	USD	1	
21	2925034	0	1/3/2018 ...	1/4/2018 12:00:00 AM	1238868	999999	1776	1	\$38.70	\$38.70	USD	1	
22	2976042	0	2/23/201...	2/27/2018 12:00:00 AM	1696695	999999	1781	1	\$38.70	\$38.70	USD	1	

Sales Count: 208969
Sales Margin Pct: 65.85%
Sales Amount: 223908740
Sales Cost: 76472776
Sales Margin: Margin Pct KPI: 65.85%
Sales Amount KPI: Sales Amount KPI Target: 305891102.92

Product Sales Customer Store Date Deliver Date
Record: 1 of 208,969

آشنایی با Measure

✓ Measure: یک محاسبه پویا ست که نتایج آن براساس context تغییر می کند.



✓ نتیجه یک Measure یک تک مقدار است و جدول نیست.

✓ محاسبه هنگام زمان کوئری

✓ نمی تواند هم نام ستون جدول باشد.

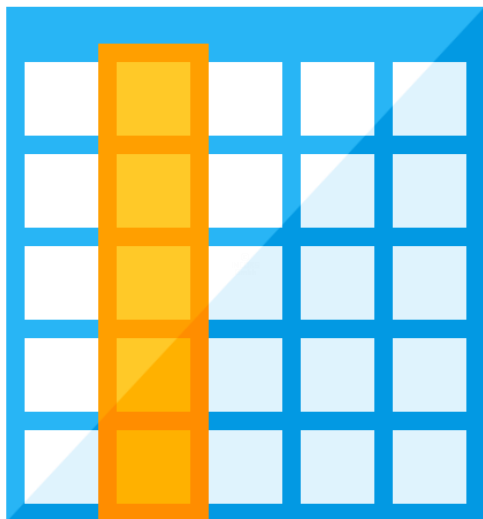
✓ محدودیتی در نام گذاری وجود ندارد.

✓ بایستی زیرمجموعه یک جدول باشد اما لزومی ندارد در محاسبات از آن جدول استفاده شود.

✓ بهتر است فرمت و پوشه بندی مشخصی برای آن در نظر بگیریم.

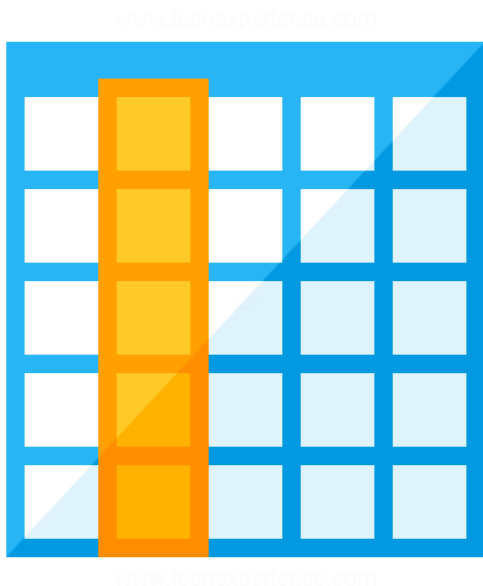
✓ در محیط توسعه نیازی به ذکر اسم جدول نیست اما در زمان اجرای کوئری بایستی حتما نامجدول را ذکر کنیم.

آشنایی با Calculated Column



- ✓ به ازای هر سطر یک مقدار محاسبه می شود.
- ✓ این مقدار هنگام پردازش محاسبه و در حافظه ذخیره می شود.
- ✓ در محیط توسعه، هنگام ورود فرمول مقدار محاسبه می شود.
- ✓ در جداول بزرگ، توصیه به استفاده از آن نمی شود و بهتر است این ستون ها در انباره داده ایجاد شود. علت این موضوع پیچیدگی و زمانبر شدن در هنگام پردازش است.
- ✓ در محاسبه می تواند از سایر Calculated Column یا Measure ها استفاده کرد.
- ✓ در محیط توسعه می توان برای دیباگ کردن از آن استفاده کرد و بهتر است محیط عملیاتی استفاده نشود.

آشنایی با Calculated Column



- ✓ در محاسبات غیرپیچیده، بهتر است Calculated برای فیلترینگ استفاده شود. بویژه جایگزینی با نامساوی اگر از bit استفاده شود کوئری به صورت بهینه تری اجرا می شود.
- ✓ بهتر است هم نام با measure نباشد هر چند ممکن است.
- ✓ اگر رابطه بین جداول پاک شود و از فرمولی استفاده شود که در آن از رابطه بین استفاده شود باعث غیر معتبر شدن آن می شود.
- ✓ در هنگام محاسبه نمی توان از Circular Dependency, Self Depedency استفاده شود.

آشنایی با Calculated Table



- ✓ این نوع جداول در پردازش مدل محاسبه و نتایج آن در حافظه ذخیره می شود.
- ✓ برای محاسبات ساده و محاسباتی که خروجی آن جداول کوچکی است می توان از آن استفاده کرد.
- ✓ بیشترین مورد استفاده به عنوان جداول کمکی، برای پویا سازی Measure ها و Dimension ها یا وارد کردن مقادیر ثابت از این نوع جداول استفاده می شود.
- ✓ از این نوع جداول می توان برای سناریوهای Role Playing Dimension در جداول کوچک استفاده کرد.

آشنایی با Data Type ها

Data type in model	Data type in DAX	Description
Whole Number	A 64 bit (eight-bytes) integer value ^{1, 2}	Numbers that have no decimal places. Integers can be positive or negative numbers, but must be whole numbers between -9,223,372,036,854,775,808 (-2^{63}) and 9,223,372,036,854,775,807 ($2^{63}-1$).
Decimal Number	A 64 bit (eight-bytes) real number ^{1, 2}	<p>Real numbers are numbers that can have decimal places. Real numbers cover a wide range of values:</p> <p>Negative values from -1.79E +308 through -2.23E -308</p> <p>Zero</p> <p>Positive values from 2.23E -308 through 1.79E + 308</p> <p>However, the number of significant digits is limited to 17 decimal digits.</p>
Boolean	Boolean	Either a True or False value.
Text	String	A Unicode character data string. Can be strings, numbers or dates represented in a text format.
Date	Date/time	<p>Dates and times in an accepted date-time representation.</p> <p>Valid dates are all dates after March 1, 1900.</p>
Currency	Currency	Currency data type allows values between -922,337,203,685,477.5808 to 922,337,203,685,477.5807 with four decimal digits of fixed precision.
N/A	Blank	A blank is a data type in DAX that represents and replaces SQL nulls. You can create a blank by using the BLANK function, and test for blanks by using the logical function, ISBLANK.