



# Deeper insights across data with SQL Server 2016

Technical white paper

Published June 2016



## Copyright

The information contained in this document represents the current view of Microsoft Corporation on the issues discussed as of the date of publication. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information presented after the date of publication.

This white paper is for informational purposes only. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED, OR STATUTORY, AS TO THE INFORMATION IN THIS DOCUMENT.

Complying with all applicable copyright laws is the responsibility of the user. Without limiting the rights under copyright, no part of this document may be reproduced, stored in, or introduced into a retrieval system, or transmitted in any form or by any means (electronic, mechanical, photocopying, recording, or otherwise), or for any purpose, without the express written permission of Microsoft Corporation.

Microsoft may have patents, patent applications, trademarks, copyrights, or other intellectual property rights covering subject matter in this document. Except as expressly provided in any written license agreement from Microsoft, the furnishing of this document does not give you any license to these patents, trademarks, copyrights, or other intellectual property.

© 2016 Microsoft Corporation. All rights reserved.

Microsoft, Microsoft Azure, SQL Server, Visual Studio, Windows, and Windows PowerShell are trademarks of the Microsoft group of companies.

All other trademarks are property of their respective owners.

# Contents

Introduction .....	4
Microsoft SQL Server evolution .....	4
SQL Server 2016 new data management enhancements .....	5
Deeper insights across data with SQL Server .....	6
Store your data and make it accessible .....	6
PolyBase .....	7
Azure HDInsight .....	8
Data warehousing .....	8
Azure Data Factory .....	10
Integration services .....	11
Hybrid data storage .....	11
Data connectors .....	11
Unstructured data .....	11
Built-in JSON support .....	12
Master Data Services .....	12
Analyze your data .....	13
Real-time operational analytics .....	13
Analysis Services .....	14
SQL Server Analysis Services tabular models .....	15
SQL Server R Services .....	17
Deliver insights to more people with Reporting Services .....	19
Modern paginated reports .....	20
Responsive mobile reports .....	21
A modern web portal to view all your reports .....	21
Hybrid BI .....	22
Pin SSRS report items to Power BI dashboards .....	22
Power BI gateways .....	23
Conclusion .....	24

# Introduction

Today, data comes from relational and non-relational sources, from on-premises environments and the cloud, and from big data and other sources. Microsoft SQL Server 2016 offers solutions designed to:

- Store data in more formats using scalable technologies
- Make data more accessible to developers and users
- Provide the services to analyze your data in highly performing, prescriptive, and predictive ways
- Share your insights across organizations to users—on-premises, via the web, or through mobile devices

## Microsoft SQL Server evolution

The growing volume, velocity, diversity, and locations of enterprise data make it increasingly challenging to discover, connect to, move, transform, integrate, and analyze it all. In fact, with the compound annual growth rate of data from 2013 to 2020 estimated at 41 percent<sup>1</sup>, this can be termed a data explosion. This is the result of more and more devices combined with a new hunger from businesses for more data to better understand and predict customer needs to make smarter decisions with data.

Yet, without the ability to analyze it, data loses its value. Microsoft SQL Server and the related Microsoft data platform can deliver the tools you need to gain deeper insights from all of your data. Microsoft customers and industry analysts agree that SQL Server is a top leader in these efforts (Figures 1 and 2).



Figure 1: Microsoft positioned as leader in Gartner Magic Quadrant for operational database management systems

<sup>1</sup> IDC, "Digital Universe," December 2012.



Figure 2: Gartner Magic Quadrant for business intelligence and analytics platforms

# SQL Server 2016 new data management enhancements

SQL Server has evolved along with the explosion in data sources and continues to innovate to facilitate data management (Figure 3).

The evolution of SQL Server continues...

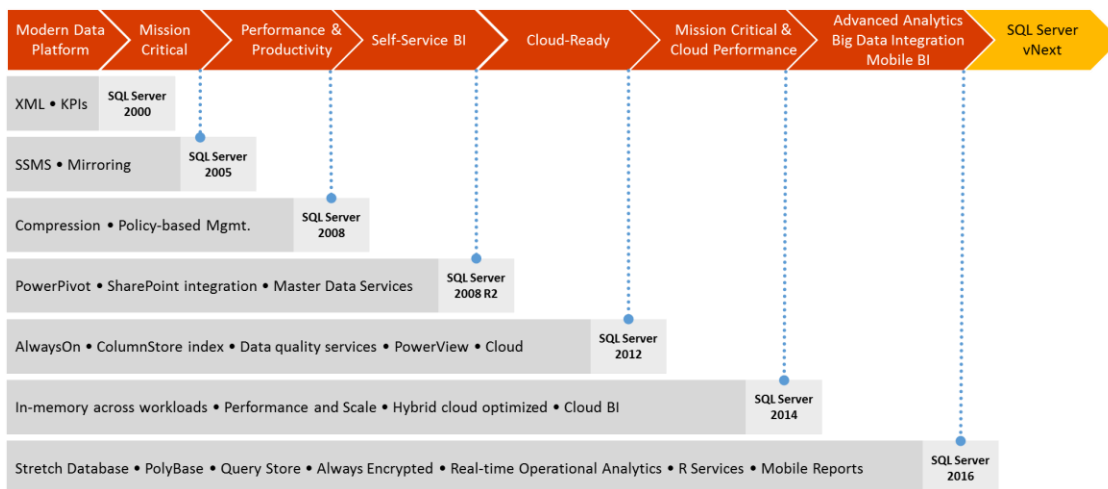


Figure 3: Major SQL Server functionalities added across releases

SQL Server 2016 introduces many new features and enhancements, including:

- SQL Server PolyBase technology to query relational SQL Server and Apache Hadoop data through a single Transact-SQL (T-SQL) query
- Support for unstructured data and data warehousing solutions—both symmetric multi-processing (SMP) and massively parallel processing (MPP)—for structured data storage
- Hybrid data storage with Stretch DB
- Connection support for more data sources
- Support for unstructured data streams
- Built-in Java Script Object Notation (JSON) support
- Enhanced Master Data Services (MDS) with Microsoft Excel add-in and speeds 15 times faster, as well as more granular security roles
- Real-time operational analytics using columnstore indexes
- Improvements to SQL Server Analysis Services (SSAS), including enterprise-grade tabular models, more efficient parallel processing, and in-memory data management
- Integration of the R language, bringing predictive analytic capabilities to your relational database
- SQL Server development tools in Microsoft Visual Studio for building business intelligence (BI) models, as well as Analysis Services models and SQL Server Reporting Services (SSRS) reports
- Ability to publish mobile-optimized reports to SSRS with Mobile Report Publisher
- New SSRS web portal that supports modern web browsers (HTML5, CSS3)

## Deeper insights across data with SQL Server

Several key trends are impacting the design of the Microsoft data platform, including big data (both in the cloud and on-premises), new types of non-relational data, and the continuing importance of data warehouses and transactional systems.

This technical white paper examines these issues in the context of accessing more types of data, scaling resources to handle the increasing data volume, analyzing data with powerful tools to gain robust insights for your business, and using technology to quickly share those insights across organizations and geographical locations.

### Store your data and make it accessible

The reality of modern data warehousing is complex. Monolithic single stores for the enterprise's data are becoming more rare. Instead, it is increasingly likely that enterprises will have multiple relational databases, Hadoop data, document-oriented NoSQL databases, and so on.

SQL Server 2016 offers new and enhanced features to support these ever-growing data storage needs. Users can now query relational and non-relational data together with PolyBase, and also take advantage of cloud-based big data capabilities with HDInsight for Azure. In addition, SQL Server 2016 includes advanced solutions for both on-premises and cloud-based data warehousing, as well as improved support for connecting, integrating, and storing data.

## PolyBase

PolyBase allows users to query non-relational data in Hadoop, Azure Storage blobs, and files—and combine it anytime, anywhere with their existing relational data in SQL Server. It also provides the option for users to import Hadoop data for persistent storage in SQL Server as well as export aged relational data into Hadoop.

PolyBase also lets users access and query data that is either on-premises or in the cloud and run analytics and business intelligence (BI) on that data. Consequently, SQL Server 2016 and PolyBase can help you build out a hybrid solution that delivers insights into your data, wherever—or however—it is stored (Figure 4).

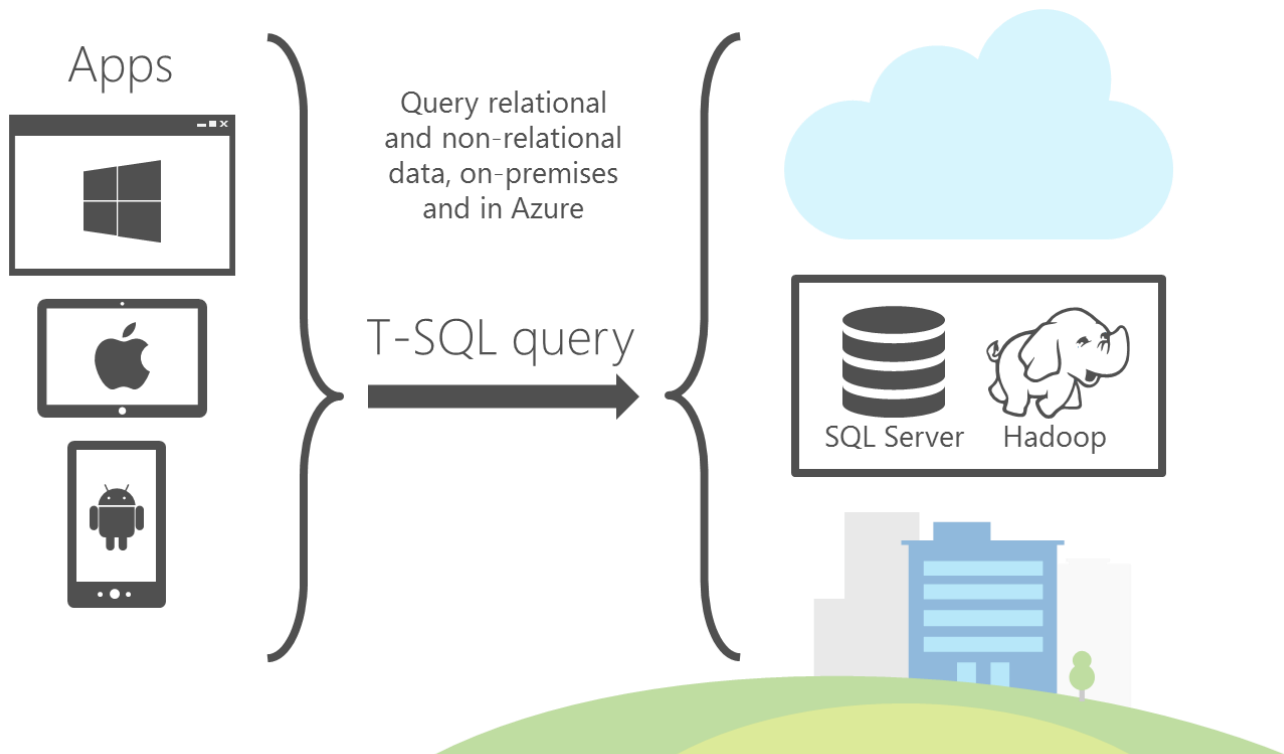


Figure 4: PolyBase access to distributed data sets

While PolyBase does allow you to move data in a hybrid scenario, it is also common to leave data where it resides and query it wherever it may be. This ties into the concept of a data lake. You can think of a data lake as providing full access to raw big data without moving it. This may be viewed as an alternate approach to processing big data to make its analysis easier, and then moving and synchronizing it into a data warehouse.

There are several benefits to not moving data. It generally means that beyond setting up the connectivity in the data lake, no additional development is required. Also, organizational limits to moving or modifying the data may become irrelevant with this approach. Finally, data processing and synchronization can be complex operations, and you may not know in advance how to process the data to deliver the best insights. SQL Server 2016 and PolyBase can be an important component in setting up a data lake, combining it with your relational data, and performing analysis and BI on it.

PolyBase can be used with Microsoft BI tools as a data source, and it can be used by many third-party BI tools as well, such as Tableau Software, IBM Cognos, and the like. PolyBase is also integrated with SQL Server AlwaysOn and failover. You can also scale out PolyBase by adding multiple SQL Server 2016 instances to a PolyBase farm.

## Azure HDInsight

Microsoft HDInsight for Azure is an enterprise-class implementation or distribution of Hadoop that is integrated with SQL Server, Active Directory, and Microsoft System Center. This makes it dramatically easier, more efficient, and more cost-effective for organizations to capitalize on the opportunity big data offers. HDInsight is the Hadoop distribution from and supported by Microsoft that is 100-percent compatible with Apache. HDInsight empowers organizations with new insights into previously untouched, unstructured data, while also connecting to widely used BI tools.

HDInsight incorporates a series of tools designed to facilitate working with big data. These tools include:

- **Sqoop.** Provides compatibility with structured data (such as that in a SQL Server database or a data warehouse) and imports it into, or exports it out of, HDInsight clusters
- **Apache HBase.** NoSQL database for unstructured and semi-structured data
- **Oozie.** Workflow management
- **Hive.** SQL-like querying of big data
- **PIG.** Scripting tools for Hadoop MapReduce transformations
- **Storm.** Data processing in real time

## Data warehousing

The traditional data warehouse is a centralized data store into which data is loaded from multiple operational databases and then transformed and stored for analytical use. The traditional data warehouse reduces strain on operational systems and enables analysis of data from multiple sources to build insights. The predictable data structure and quality in traditional data warehouses optimize processing and reporting.

The emergence of big data solutions (such as immense data volumes, unstructured data, and high-velocity streams), along with the pressing need to draw insights in real time, has given rise to the modern data warehouse. The modern data warehouse brings horizontal scalability, distributed parallel processing, and support for non-relational data storage and querying in addition to massive relational data stores. The modern data warehouse also provides the means to go beyond periodic reporting and on to advanced predictive and prescriptive analytics.

Microsoft helps organizations address the challenges of big data with its family of data warehouse solutions—SQL Server, SQL Server Fast Track Data Warehouse, Azure SQL Data Warehouse, and Analytics Platform System—that provides a robust and scalable platform for storing and analyzing data in traditional and modern data warehouses. SQL Server provides enhanced features, such as remote blob storage (RBS) and partitioned tables that scale to 15,000 partitions, to support large, sliding-window scenarios. (In a sliding-window scenario, partitioned tables are managed for efficiency to maintain the same number of partitions over time by adding a new partition to accommodate the newest data and removing the partition that contains the oldest data.)

Data warehousing solutions can be categorized by the processing architecture and whether the infrastructure is based in Azure or on-premises.



## SMP

- Scale to terabytes
- Add capacity as you grow
- Support for multi-purpose database needs



SQL Server 2016



SQL Server in an Azure VM



Azure SQL Database

## MPP

- Seamless scale to petabytes
- Optimum price/performance
- Support for active DW across many users



APS



Azure SQL Data Warehouse



Azure Data Lake

Figure 5: SQL Server 2016 SMP and MPP capabilities

Symmetric multi-processing (SMP) refers to the parallel processing capabilities of servers. SMP-optimized solutions, such as SQL Server 2016 (for on-premises) and SQL Server Fast Track Data Warehouse (for Azure), leverage multi-processor systems to run operations in parallel (Figure 5).

Massively parallel processing (MPP) refers to distributed task processing across multiple distinct servers. Azure SQL Data Warehouse and Parallel Data Warehouse (on Analytics Platform System) are MPP solutions for on-premises and cloud data warehousing.

The SQL Server in-memory columnstore index stores and manages data by using column-based data storage and batch-mode query processing. Columnstore indexes can transform the data warehouse experience for users by enabling faster performance for common data warehouse queries such as filtering, aggregating, grouping, and star-join (and snowflake) queries. SQL Server 2016 brings several improvements to columnstore indexes<sup>2</sup>:

- Support is available for primary and foreign keys on columnstores (SMP only)
- AlwaysOn supports querying a columnstore index on a readable secondary replica
- Multiple Active Result Sets (MARS) supports columnstore indexes
- A new, dynamic management view provides performance troubleshooting information at the row group level
- Single-threaded queries on columnstore indexes can run in batch mode
- The SORT operator runs in batch mode
- Multiple DISTINCT operations run in batch mode

SQL Server 2016 creates new functionality and enhances existing services to facilitate moving to a hybrid cloud. Scalability, availability, security, identity, backup and restore, and replication now have enhancements for working with SQL Server in Azure Virtual Machines, Azure SQL Database, and other Azure services.

<sup>2</sup> [Columnstore Indexes for Data Warehousing](#)

## Azure Data Factory

Azure Data Factory (ADF) enables you to process on-premises data from a database—such as SQL Server— together with cloud data from sources like Azure SQL Database, blobs, and tables. These data sources can be composed, processed, and monitored through simple, highly available, fault-tolerant data pipelines. ADF supports Apache Hive, Apache Pig, and C# processing, along with key processing features such as automatic Hadoop (HDInsight) cluster management, re-tries for transient failures, configurable timeout policies, and alerting.

Data Factory works across on-premises and cloud data sources and software as a service (SaaS) to ingest, prepare, transform, analyze, and publish your data (Figure 6). Use Data Factory to compose services into managed dataflow pipelines and transform your data. You can use services like HDInsight (Hadoop) and Microsoft Azure Batch for your big data computing needs and Azure Machine Learning to operationalize your analytics solutions. The rich visualizations of Data Factory take you beyond just a tabular monitoring view to quickly display the lineage and dependencies among your dataflow pipelines. Monitor all of your pipelines from a single, unified view to easily pinpoint issues and set up monitoring alerts.

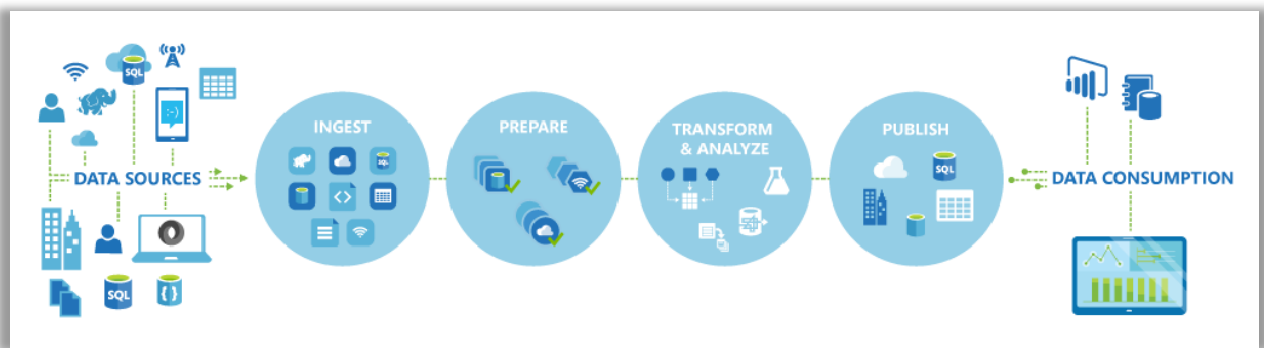


Figure 6: Azure Data Factory monitoring of dataflow pipelines from a single, unified view

## Integration services

SQL Server 2016 contains a number of enhancements that can improve the development, management, and monitoring of your SQL Server Integration Services (SSIS) data packages while benefiting your on-premises and cloud operations with cloud integration, connectivity improvements, and product improvements.

ADF can now orchestrate on-premises SSIS execution. SSIS can also read from ADF as a data source via the ADF dataflow task. In addition, SSIS developers can leverage the Azure Storage Connector to move data from on-premises to Azure Storage, or vice versa. SSIS developers can also trigger HDInsight jobs directly from SSIS, so they can better integrate with HDInsight to process data already in the cloud—without needing to move the unprocessed cloud data on-premises.

## Hybrid data storage

With SQL Server 2016 Stretch Database, you can stretch large operational tables from on-premises to Azure with the ability to query with near-infinite capacity and lower total cost of (TCO) storage. Applications continue to work without code changes, while existing database administrator (DBA) skills and processes remain relevant, and developers can continue using current tools and APIs.

For more information: [Stretch Database](#)

## Data connectors

SQL Server 2016 also has a wide range of new and enhanced data connectors, including Hadoop File System (HDFS), JavaScript Object Notation (JSON), and Oracle/Teradata connector V4 by Attunity. Support is also included for OData V3 protocol for ATOM and JSON data formats and OData V4 protocol for JSON data formats. These data formats can also be used in SSIS Data Flow and loaded in any destination or applied to any data transformation.

Hadoop connectors for SQL Server and Analytics Platform System (APS) are available for download to organizations that have licenses for SQL Server and APS. These connectors enable bidirectional data movement across SQL Server and Hadoop, so users can work effectively with both structured and unstructured data.

## Unstructured data

SQL Server supports a growing number of types and volumes of complex data with FILESTREAM, RBS, and spatial support—enhancements on top of the already robust and built-in foundation that extends beyond relational capabilities.

SQL Server FILESTREAM enables SQL Server-based applications to store unstructured data, such as documents and images, on the file system. Applications can leverage the rich streaming APIs and performance of the file system and, at the same time, maintain transactional consistency between the unstructured data and corresponding structured data.

With SQL Server, complex data types are handled with the same attention as common data types. Organizations can use FILESTREAM to store and manage complex data in a variety of ways, as if it were part of the database. Additionally, with SQL Server, organizations can enjoy the high-availability benefits of AlwaysOn for complex data managed through FILESTREAM—even when they take advantage of RBS and SQL Server FileTable.

## Built-in JSON support

SQL Server 2016 offers built-in JSON support. JSON is a serialization format that has become widely adopted in a variety of scenarios. It is important to note that built-in support is not the introduction of a native data type. JSON is represented as NVARCHAR in SQL Server 2016.

A number of operations support JSON:

- FOR JSON, similar to FOR XML, will format query results as JSON text
- OPEN JSON is a table-valued function that will parse JSON text (an array of JSON objects) and return the objects as a table
- A set of built-in functions will allow for querying JSON text with T-SQL (ISJSON, ISJSON\_VALUE, JSON\_QUERY)

Since JSON is stored as NVARCHAR, built-in JSON support works with other SQL Server features such as memory-optimized tables, Row-Level Security (RLS), and encryption, to name a few.

For more information: [JSON to Table](#)

For more information: [Built-in functions](#)

## Master Data Services

Master Data Services (MDS) continues to make it easier for organizations to manage master data structures (object mapping, reference data, dimensions, and hierarchies) used in data integration operations. With entity-based staging, DBAs can load all members and attribute values for an entity at one time. Additionally, the Explorer and integration management functional areas of the Master Data Manager web application have been updated with a new look and feel based on the Microsoft Silverlight browser development tool. DBAs can add and delete members more quickly and can move them into a hierarchy more easily.

The MDS add-in for Excel democratizes data management, so information workers have the ability to build data management applications directly in Excel. Information workers can use this add-in to load a filtered data set from the MDS database, work with data in Excel, and then publish changes back to the MDS database. Administrators also can use the add-in to create new entities and attributes.

Enhancements in MDS can be organized around performance and scale, manageability and monitoring, and security.

In the areas of performance and scale, MDS supports larger models, loads data more efficiently, and enables up to 15-times faster performance for Excel. For scalability, the MDS model deployment has been enhanced to support bigger models and scale entities to 100 million members (and beyond). Also, optional row-level compression per entity is now supported.

In the areas of manageability and monitoring, configurable retention settings for the transaction logs and reuse of entities across models is now supported. Each attribute has a viewable display name to provide more control over the names displayed for a given object, including the code and name attributes. Enhancements in hierarchy management have simplified various hierarchy types, making it easier to find and manage unused members.

In the area of security, MDS in SQL Server 2016 has more granular security permissions around read, write, delete, and create. It also supports multiple system administrators and an explicit model administrator permission property.

In the MDS add-in for Excel, you can now manage business rules, including creating and editing rules. You can create a derived hierarchy that displays many-to-many relationships. In the previous version, if you tried to publish data that was changed by another user, the publishing failed with a conflict error. To resolve this error, you can now perform merge conflicts and republish the changes.

## Analyze your data

Data that isn't analyzed quickly and accurately loses its value. SQL Server continues to improve analytics functions for today's ever-expanding stores of data.

### Real-time operational analytics

SQL Server 2016 introduces real-time operational analytics, which provides the ability to concurrently run analytical workloads on the same database tables used by operational workloads (Figure 7). This is made possible with the use of columnstore indexes.

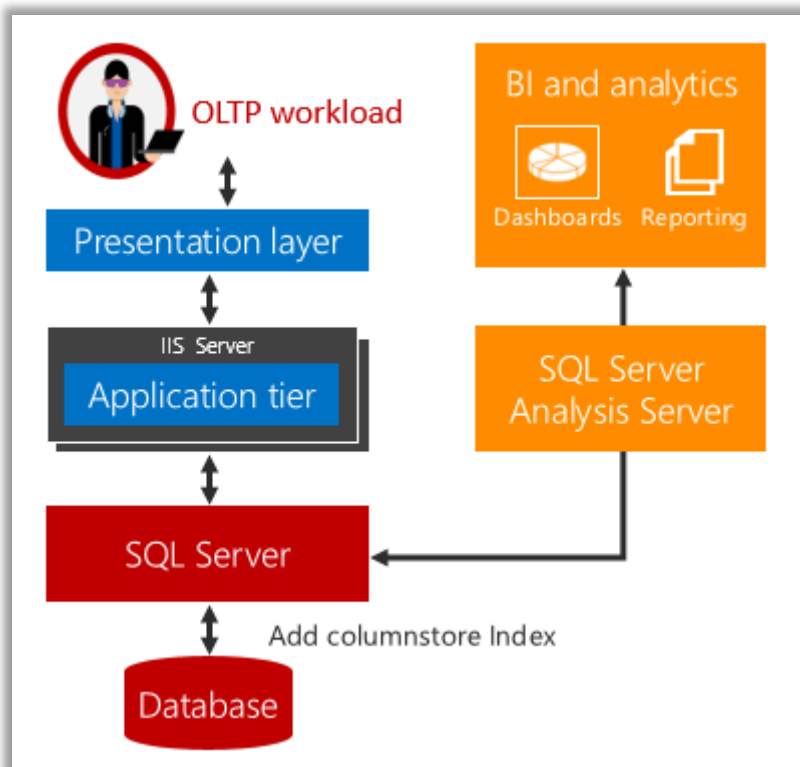


Figure 7: [Real-time operational analytics with in-memory technology](#)

SQL Server lets users create updatable, nonclustered columnstore indexes on traditional rowstore tables. Transactional workloads run against the rowstore, while analytics workloads run against the columnstore index. SQL Server automatically maintains all changes to the indexes, so the transactional changes are always up-to-date

for analytics. The impact of maintaining the columnstore index can be minimized or eliminated using compression delay or by creating a filtered columnstore index on colder data. Real-time operational analytics is supported both on disk-based and memory-optimized tables.

The use of non-clustered columnstore indexes (NCCI) can be further abstracted from the online transaction processing (OLTP) workload through the use of SQL Server AlwaysOn Availability Groups, which now supports readable secondary replicas. This enables analytics queries to run on secondaries with zero impact on the workload of the primary replica (Figure 8).

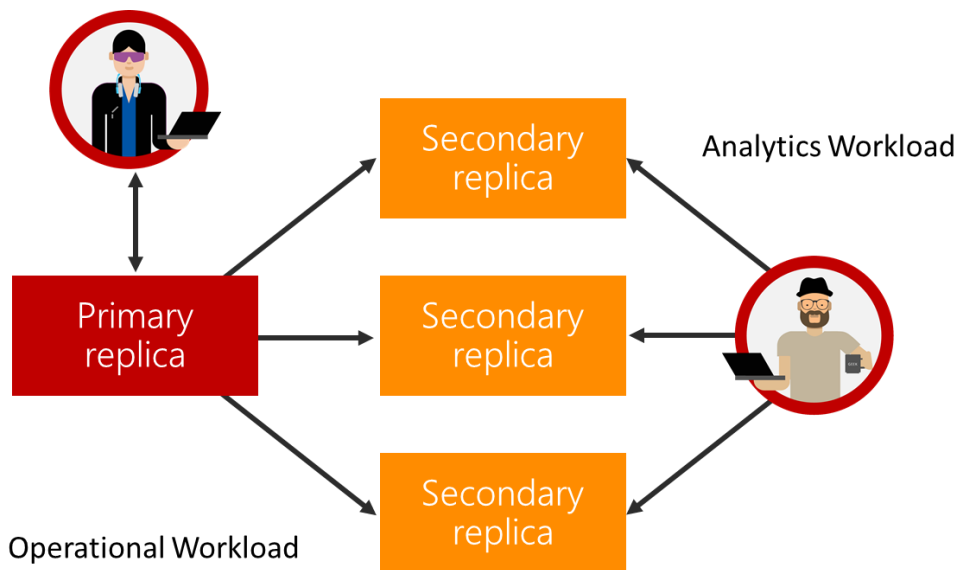


Figure 8: [Support in AlwaysOn Availability Groups for readable secondary replicas](#)

These factors make it possible and practical to run analytics in real time as opposed to traditional data warehouse implementations that require data migration/transformation from a transactional system to a dedicated data warehouse.

For more information: [Real-time operational analytics](#)

For more information: [Columnstore indexes](#)

## Analysis Services

Microsoft SQL Server Analysis Services (SSAS) is an online analytical processing (OLAP) and data mining engine used in BI applications. It provides analytical data for business reports and client applications, such as Excel, Reporting Services reports, Power BI, and other third-party data visualization tools used to define measures and key performance indicators (KPIs) for reports (Figure 9).

This engine can be used to create powerful BI semantic models that provide BI professionals with a more intuitive abstraction in either a traditional multidimensional model or a simpler tabular model. The BI Semantic model also provides fast access to data and the capability to directly connect models to the data sources without having to

move data from the source.

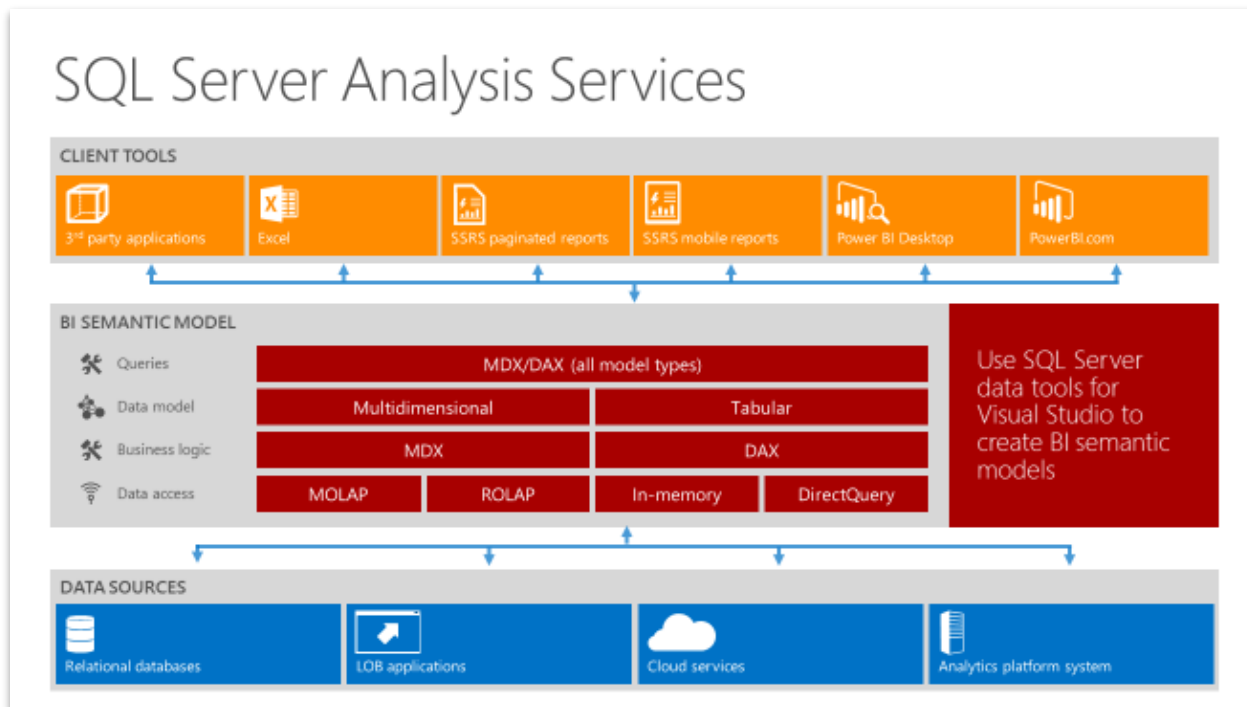


Figure 9: Capabilities of SQL Server Analysis Services

SQL Server 2016 includes several enhancements to Analysis Services in the areas of enterprise readiness, modeling platform, BI tools, SharePoint integration, and hybrid BI support. The improved Tabular model allows BI developers to solve increasingly complex business problems in an agile and straightforward manner with out-of-the-box support for modeling and calculation scenarios such as percentile and many-to-many patterns. Enhancements to the DirectQuery storage mode allow BI developers to utilize data sets from different data sources directly without the need to move the data. Improvements to the high-performing in-memory storage mode can enable analytics on large data sets at the speed expected from today's fast-paced business users.

## SQL Server Analysis Services tabular models

### Bi-directional cross filters

Bi-directional cross filtering allows modelers to determine how filters flow. With bi-directional cross filters, the filter context is propagated to the second related table on the other side of a table relationship. In essence, bi-directional cross filters enable support for many-to-many relationships—without the need for complex DAX expressions.

### Calculated tables

A calculated table is a model-only construction based on Data Analysis Expressions (DAX) or a query in SQL Server Data Tools (SSDT). When deployed in a database, a calculated table is indistinguishable from regular tables. The classic use for a calculated table is duplicating a date table for use in multiple contexts (order date, ship date, and so forth). By creating a calculated table for a given role, you can now activate a table relationship to facilitate queries or data interaction using the calculated table. Another use for calculated tables is to combine parts of existing tables into an entirely new table that exists only in the model.

For more information: [Calculated tables](#)

## DAX improvements

In addition to over 50 additional DAX functions, DAX now supports the use of named variables in expressions, which is particularly useful when working with complex, lengthy expressions. Named variables can store the result of an expression, which can then be passed as an argument to other measure expressions, in which case the variable expression is executed once but can be used multiple times if needed.

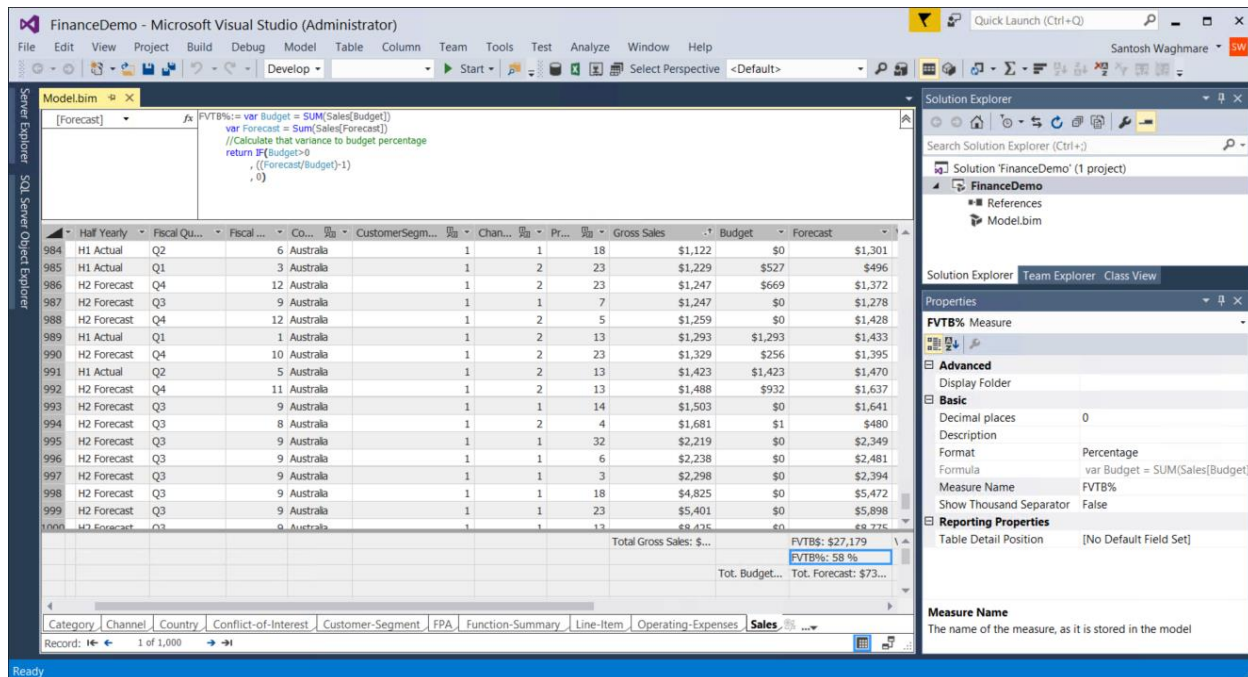


Figure 10: DAX support of named variables

## Parallel partition processing

With parallel partition processing, the partitions of a single table are processed in parallel instead of sequentially. Partitions are used to manage data in a table. With parallel partition processing, each defined partition can be processed in parallel. This means that two expensive operations per partition—that is, loading the data from the data source and compressing the data into VertiPaq—are parallelized. This will put more load onto your data source and your SSAS server, but it will also reduce the processing time significantly.

## DirectQuery

DirectQuery in SQL Server 2016 can be used to connect directly to SQL Server, SQL Server Parallel Data Warehouse (Microsoft Analytics Platform System), Oracle, and Teradata. Additionally, Excel PivotTables or any other Microsoft Multidimensional Expressions (MDX) client tool is able to query a tabular model in DirectQuery mode. This will allow you to expose more data sources directly to your business users through the semantic layer of an Analysis Services data model, without caching the data in Analysis Services.

Calculated columns and RLS are supported for DirectQuery models. Previously, the presence of calculated columns prevented customers from running a tabular model in DirectQuery mode.

Because reduced chattiness helps improve performance, the DAX query enhancements listed above help with DirectQuery optimization. This is especially important in DirectQuery mode because fewer queries are sent to the



SQL Server data source. Further optimizations help generate simpler SQL queries, and SQL Server query performance benefits from simple queries.

For more information: [Business insights with SQL Server 2016 Analysis Server](#)

### SQL Server Data Tools in Visual Studio 2015

SQL Server Data Tools (SSDT) is available for Visual Studio 2015 in a simplified setup experience.

SSDT is a modern database development, data modeling, and design tool for SQL Server, Azure SQL Database, Analysis Services data models, Integration Services packages, and Reporting Services reports. SSDT for Analysis Services is available as part of SQL Server Data Tools Visual Studio 2015, providing a single simplified installation experience for all SQL Server data tools for Visual Studio 2015. The previous SSDT-BI installation was approximately 1 GB in size; the new setup is small and will only download what you need. For example, if you already have Visual Studio installed, it will not be downloaded again. This will make incremental updates much easier and faster to consume. This version of SSDT can be used to develop both Multidimensional and Tabular models for supported versions of SSAS.

## SQL Server R Services

The most popular language for predictive analytics available today is the powerful R language. However, R as an open source programming language has not scaled well for big data analytics. With the Microsoft purchase of Revolution Analytics (the leading provider for commercial software and services built on top of R), the Microsoft data platform now has this functionality.

### Predictive analytics

Predictive analytics is a key big data capability. R allows you to bridge the gap between the database and data science. SQL Server 2016 lets you manage R models in SQL Server. This will help you use the power of R and data science to unlock big data insights with advanced analytics (Figure 10). SQL Server 2016 makes advanced analytics more accessible and increases performance for advanced analytic workloads by bringing R processing closer to the data and building advanced analytic capabilities right into SQL Server.

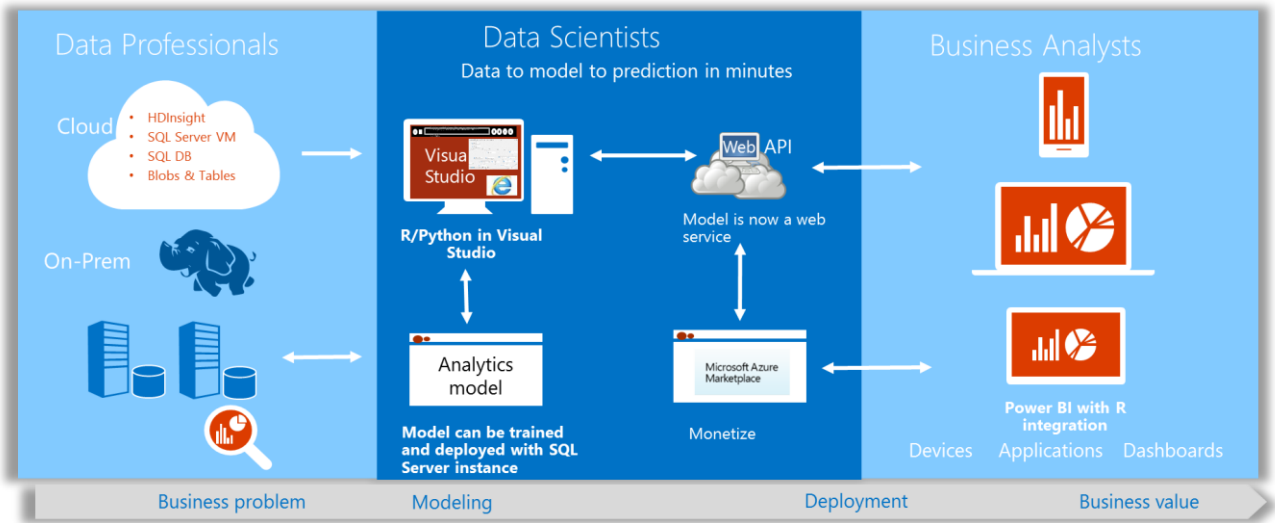


Figure 10: Management of R models in SQL Server for big data insights

### Integrated with SQL Server 2016

This integration of R with SQL Server means that database professionals can use T-SQL for advanced analytics on operational data and models, and they can secure and ensure their availability. With SQL Server 2016, data scientists no longer need to extract data from SQL Server via Microsoft Open Database Connectivity (ODBC) to analyze it with R. Instead, they can use the Revolution R Enterprise (RRE) IDE for R language to analyze large data sets and build predictive and embedded models with the compute happening on the SQL Server machine. This eliminates the time and storage required to move the data, and it gives data scientists the power of packages to apply to their database.

Database professionals can now work in their favorite analytics environment, such as R or Python in Visual Studio, while taking advantage of the computational power, memory, and parallelism of the database engine and increasing model fidelity (Figure 11). Developers can operationalize R script/model over SQL Server data by using T-SQL constructs. DBAs can manage resources and secure and govern the R runtime execution in SQL Server.

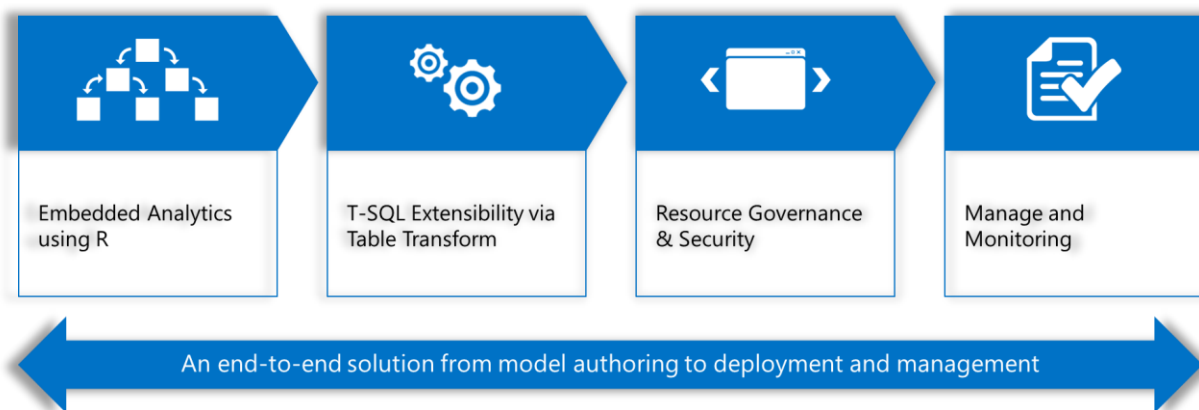


Figure 11: Advanced analytics capabilities in SQL Server 2016 for end-to-end solutions

### Big data analytics

This integration of R will facilitate many big data scenarios, such as using big data for better audience targeting, churn forecasting, anomaly detection, and fraud and risk analysis. Business users can access results from anywhere

and on any device. Further, once models have been developed and trained, they can be deployed as web services to the Microsoft Azure Marketplace. Integrating scalable R scripts and distribution into HDInsight and Azure Machine Learning makes it much easier and faster to analyze big data and to operationalize R code for production application. Azure Machine Learning pioneered this integration by offering its core machine learning functionality with the option to extend as needed. With this integration, data scientists can develop near-database analytic scenarios like fraud detection, customer churn analysis, product recommendations, and the like (Figure 12).

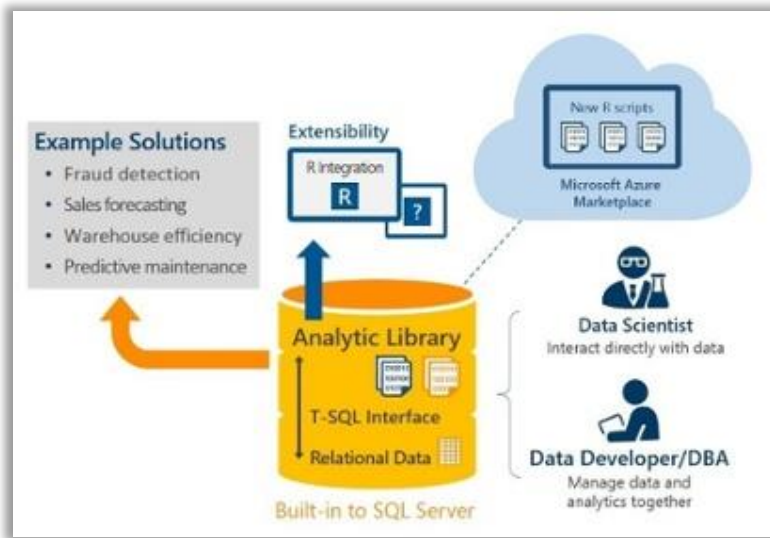


Figure 12: Near-database analytic scenarios

## Deliver insights to more people with Reporting Services

SQL Server Reporting Services (SSRS) provides a modern, on-premises solution for creating, deploying, and managing reports within your organization. With the release of SQL Server 2016, Reporting Services brings modern design to enterprise reporting. You can create mobile reports optimized for smartphones and tablets to bring the power of data to more users in the organization as well as create modern paginated reports.

This server-based reporting platform includes a complete set of tools to create, manage, and deliver reports and APIs that enable developers to integrate or extend data and report processing in custom applications. These tools work within the Microsoft Visual Studio environment and are fully integrated with SQL Server tools and components. Report Builder enables report authors to design paginated reports and publish them to a report server. Mobile Report Publisher enables report authors to design mobile reports and publish them to a report server. SQL Server Data Tools (SSDT) integrates into Visual Studio 2015 and enables developers to design paginated reports within the Visual Studio integrated development environment and take advantage of projects, source control, build, and deployment.

SSRS provides a new web portal built on HTML5 that enables users to browse, search, view, and manage reports (both paginated and mobile) using a modern browser to access all your reports in one place. A scheduling and delivery agent refreshes data sets and executes reports on a schedule, and delivers paginated reports to users via email and other means. The report server database, built on the SQL Server Database Engine, stores and manages the catalog of content, including data sources, data sets, paginated reports, mobile reports, and KPIs. The database can be on the report server or on a different server running SQL Server.

An enterprise-scale processing and rendering engine connects to and queries various types of data sources; processes data (filtering, sorting, grouping, aggregation); evaluates expressions and custom code; and renders paginated reports to various formats. The same web APIs that serve the web portal, Power BI Mobile apps, and report design apps also enable third-party and custom apps to integrate with Reporting Services and report server content. Also, multiple report servers can operate in a scale-out configuration for additional scale and availability.

## Reporting Services architecture

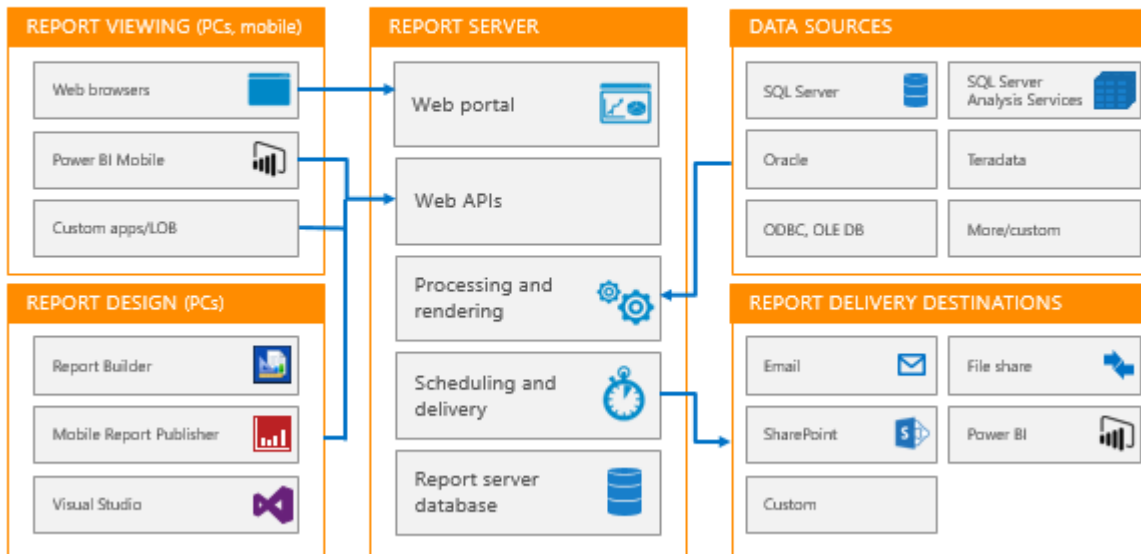


Figure 13: Reporting Services architecture

### Modern paginated reports

SQL Server 2016 Reporting Services modernizes and enhances **paginated reports** in several ways. As a report designer, you'll find modern styles for charts, gauges, maps, and other data visualizations, enabling you to create modern reports more quickly and easily than ever. In addition to the pre-existing chart types, you'll find two new ones—Treemap and Sunburst charts—which are ideal for visualizing hierarchical information. When you design parameterized reports, you now have direct [control of the position of each parameter](#) so you can arrange them the way you like, including across several columns to make the best use of wider screens.

To design reports, you'll find modern versions of familiar tools. For example, Report Builder now sports a modern look-and-feel. If you're a developer who prefers designing reports in Visual Studio, you can now do so in Visual Studio 2015. In Visual Studio 2015, you can take advantage of Visual Studio projects, source control, and more. When developing report logic or custom extensions to the platform, you can now write or integrate with code using the .NET Framework 4.6.

You'll find a number of new features when viewing reports as well. In addition to exporting reports to Word, Excel, PDF, and other formats, you can also [export them to PowerPoint presentations](#). Report items become individual PowerPoint objects, so you can move and resize them to customize your presentation. Likewise, in addition to monitoring important metrics and trends by delivering reports to your email inbox, you can now [pin report charts, gauges, and maps to your Power BI dashboards](#).

## Responsive mobile reports

SQL Server 2016 Reporting Services [introduces mobile reports](#) for on-premises implementations to address the need for **responsive-layout** reports that adapt to different screen shapes, sizes, and orientations. Mobile reports dynamically adjust the content depending on whether you're using a phone, tablet, or PC, and even as you rotate your device.

Mobile reports are built on [Datazen technology that Microsoft acquired in 2015](#) and on the idea that a "mobile-first" approach, designed for mobile devices, delivers the optimal experience for viewing reports on phones and tablets. You can create mobile reports using the SQL Server [Mobile Report Publisher](#) app and view them using either the [Power BI mobile app](#) or your browser.

Paginated reports and mobile reports are complementary; you can choose the type of report that best fits your needs on a case-by-case basis.

**Microsoft SQL Server Mobile Report Publisher** is the single point for creation of mobile reports. Connect to SQL Server Reporting Services to access data sources and easily create stunning reports, and then publish them to Reporting Services for access by others in the organization via a unified web experience for mobile devices.



Figure 14: SQL Server Mobile Report Publisher interface

Standard paginated reports are laid out using fixed page sizes, which is problematic for mobile devices. While a mobile screen may support high resolutions, a report formatted for a computer screen becomes difficult to view on a mobile device. With Mobile Report Publisher, content is dynamically adjusted to fit the screen size and orientation.

Mobile Report Publisher allows businesses to create mobile reports from data in Microsoft Excel as well as other enterprise and cloud database sources. These tools can connect to enterprise data sources, integrate with Active Directory for user authentication, deliver live data updates to mobile devices, and personalize data queries for each user. Reports can also be published to a Reporting Services web portal and viewed in a browser or in the Power BI for iOS mobile app on an Apple iPhone or iPad.

## A modern web portal to view all your reports

SQL Server Reporting Services also includes a modern web portal experience, allowing users to access KPIs, paginated and mobile reports, Power BI Desktop files, and Excel files in one centralized location. The web portal is built from the ground up on HTML5 and designed specifically to support modern browsers, including the latest versions of Internet Explorer, Chrome, Firefox, and Safari. The ActiveX print experience on the ReportViewer

toolbar has been replaced with a modern, PDF-based experience that works across the matrix of supported browsers, including Microsoft Edge.

To facilitate presenting reports to an audience, an entire report can be exported to a set of Microsoft PowerPoint slides.

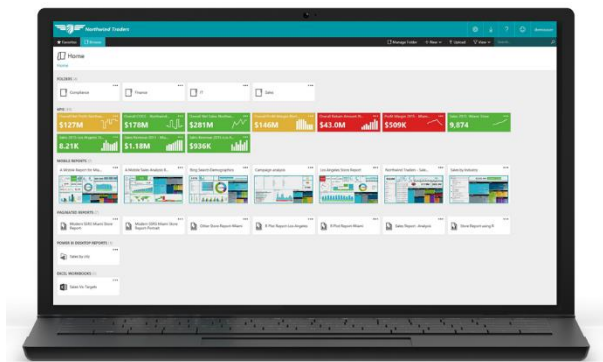


Figure 15: Modern web portal for KPIs, paginated and mobile reports, Power BI Desktop files, and Excel files

## Hybrid BI

As your organization transitions to the cloud to benefit from flexibility and scale, you may still have many data sources that reside on-premises. Microsoft BI tools support this trend so that you can transition to the cloud at your pace and take advantage of a hybrid BI solution that will let you continue to benefit from existing on-premises investments.

### Pin SSRS report items to Power BI dashboards

In addition to using Reporting Services on-premises, you can use Power BI dashboards to monitor metrics and trends that matter. You can harness your on-premises reporting investments and extend them to the cloud by simply pinning a Reporting Services paginated report item to a Power BI dashboard and view all your information in one place.

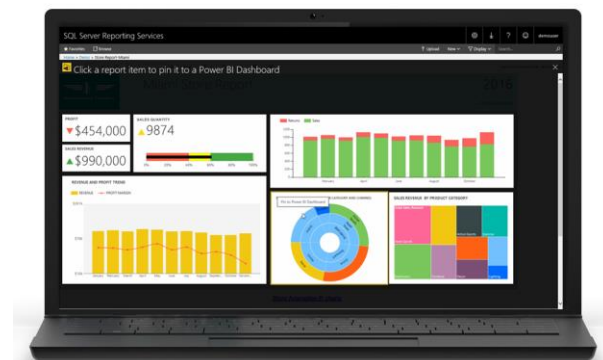


Figure 16: Power BI dashboard view

## Power BI gateways

With Power BI gateways, you can realize the benefits of a cloud-based BI solution without having to move your data. You can keep your data fresh by connecting to on-premises data sources such as SQL Server Analysis Services (tabular and multidimensional models) without needing to move data to the cloud. You can also query large data sets while utilizing your existing investments. The gateways provide the flexibility you need to meet both the individual needs of your users and the needs of your organization as a whole.

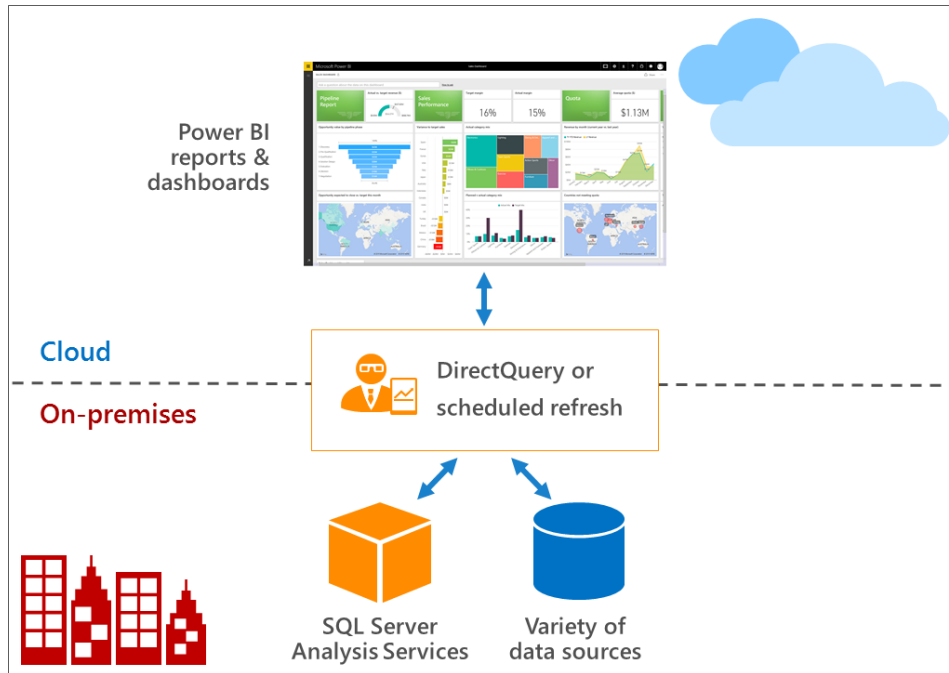


Figure 17: Combined benefits of Power BI and SQL Server Analysis Services

The Power BI gateway for enterprise deployments allows for the use of DirectQuery to an on-premises SQL Server, SAP HANA, Oracle, or Teradata, as well as a live connection to Analysis Services. You can use a live connection to either a tabular or multidimensional model in Analysis Services. The enterprise gateway can also be used with scheduled refresh for data that you have imported into Power BI.

For more information: [Power BI gateways](#)

# Conclusion

Today, data comes from relational and non-relational sources, from on-premises environments and the cloud, and from big data and other sources. As shown in Figure 15, SQL Server 2016 offers solutions designed to:

- Store data in more formats using scalable technologies
- Make data more accessible to developers and users
- Provide services to analyze your data in highly performing, prescriptive, and predictive ways
- Share your insights across organizations with users—on-premises, via the web, or through mobile devices

With technologies such as PolyBase, unstructured data can be accessed from HDInsight alongside structured data. Warehousing solutions—such as Azure Fast Track Data Warehouse or Azure SQL Data Warehouse (or APS for on-premises)—provide the means to efficient, highly available data. Real-time operational analytics brings insights to users faster. Enhancements to Analysis Services brings better modeling support and highly performing BI models. R Services enables powerful prescriptive and predictive analytics directly on the data platform. Reporting services—retooled for modern browser support and mobile reports—and Power BI ensure that insights are available on demand to users, regardless of location.



## More information

The following websites offer more information about topics discussed in this white paper:

- [SQL Server](#)
- [Microsoft business intelligence](#)
- [Microsoft big data solutions](#)

## Feedback

Did this paper help you? Please give us your feedback by telling us on a scale of 1 (poor) to 5 (excellent) how you rate this paper and why you have given it this rating. More specifically:

- Are you rating it highly because of relevant examples, helpful screen shots, clear writing, or another reason?
- Are you rating it poorly because of examples that don't apply to your concerns, fuzzy screen shots, or unclear writing?

This feedback will help us improve the quality of white papers we release.

Please send your feedback to: <mailto:sqlsrvwpfeedback@microsoft.com>