A **Reporting Services subscription** is a configuration that delivers a report at a specific time or in response to an event, and in a file format that you specify. For example, every Wednesday, save the MonthlySales.rdl report as a Microsoft Word document to a file share.

An **Ad**-**Hoc Query** is a **query** that cannot be determined prior to the moment the **query** is issued. It is created in order to get information when need arises and it consists of dynamically constructed SQL which is usually constructed by desktop-resident **query** tools.

On Transact SQL language the sp\_helpdb is part of Database Engine Stored Procedures and shows information about a specified database or all databases.

A linked dimension is based on a dimension created and stored in another Analysis Services database of the same version and compatibility level. By using a linked dimension, you can create, store, and maintain a dimension on one database, while making it available to users of multiple databases. To users, a linked dimension appears like any other dimension.

Linked dimensions are read-only. If you want to modify the dimension or create new relationships, you must change the source dimension, then delete and recreate the linked dimension and its relationships. You cannot refresh a linked dimension to pick up changes from the source object.

A perspective is a definition that allows users to see a cube in a simpler way. A perspective is a subset of the features of a cube. A perspective enables administrators to create views of a cube, helping users to focus on the most relevant data for them. A perspective contains subsets of all objects from a cube. A perspective cannot include elements that are not defined in the parent cube.

**dbo** is the default schema in **SQL** Server. You can create your own schemas to allow you to better manage your object namespace

Data **integrity** is the maintenance of, and the assurance of the accuracy and **consistency** of, data over its entire life-cycle, and is a critical aspect to the design, implementation and usage of any system which stores, processes, or retrieves data.

Hierarchies, in tabular models, are metadata that define relationships between two or more columns in a table.

A stored procedure is nothing more than prepared SQL code that you save so you can reuse the code over and over again.  So if you think about a query that you write over and over again, instead of having to write that query each time you would save it as a stored procedure and then just call the stored procedure to execute the SQL code that you saved as part of the stored procedure.

**stored** in the database in compiled form

Views that access multiple tables can only modify one of the tables in the view. Views that use functions, specify DISTINCT, or utilize the GROUP BY clause may not be updated.

OLE DB is Microsoft's strategic low-level application program interface ([API](http://searchexchange.techtarget.com/definition/application-program-interface)) for access to different data sources. OLE DB includes not only the Structured Query Language ([SQL](http://searchsqlserver.techtarget.com/definition/SQL)) capabilities of the Microsoft-sponsored standard data interface Open Database Connectivity ([ODBC](http://searchoracle.techtarget.com/definition/Open-Database-Connectivity)) but also includes access to [data](http://searchdatamanagement.techtarget.com/definition/data) other than SQL data.

A **relational database** is a collection of data items organized as a set of formally-described tables from which data can be accessed or reassembled in many different ways without having to reorganize the**database** tables

A relational database is a set of tables containing data fitted into predefined categories. Each table (which is sometimes called a *relation*) contains one or more data categories in columns. Each [row](http://searchoracle.techtarget.com/definition/row) contains a unique instance of data for the categories defined by the columns.

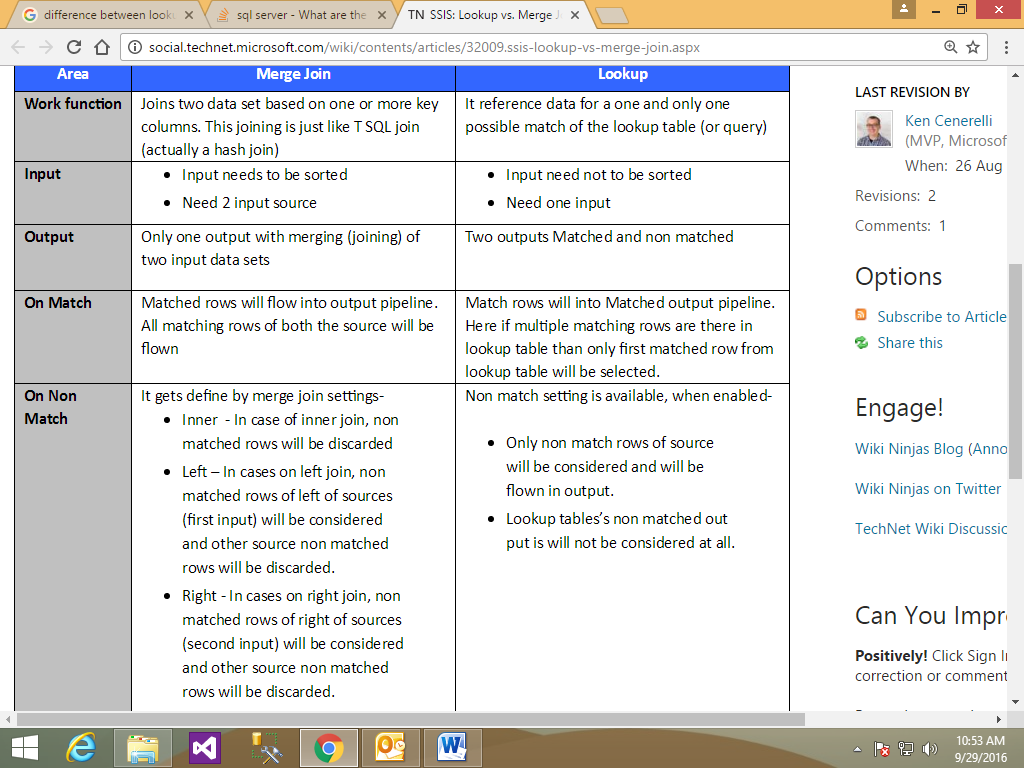
The standard user and application program interface to a relational database is the *structured query language* ([SQL](http://searchsqlserver.techtarget.com/definition/SQL)).

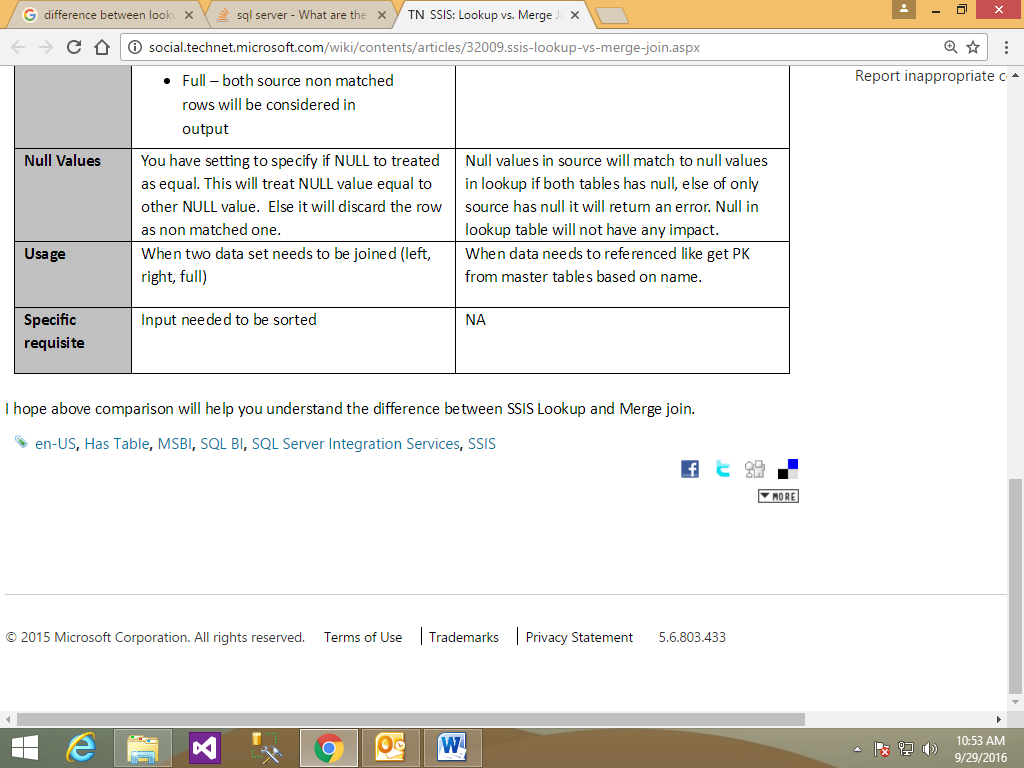
A relational database (the concept) is a data structure that allows you to link information from different 'tables', or different types of data buckets

A non-relational database just stores data without explicit and structured mechanisms to link data from different buckets to one another.

**Fragmentation** occurs in a dynamic [memory](http://ecomputernotes.com/fundamental/input-output-and-memory/what-are-the-different-types-of-ram-explain-in-detail) allocation system when many of the free blocks are too small to satisfy any request.

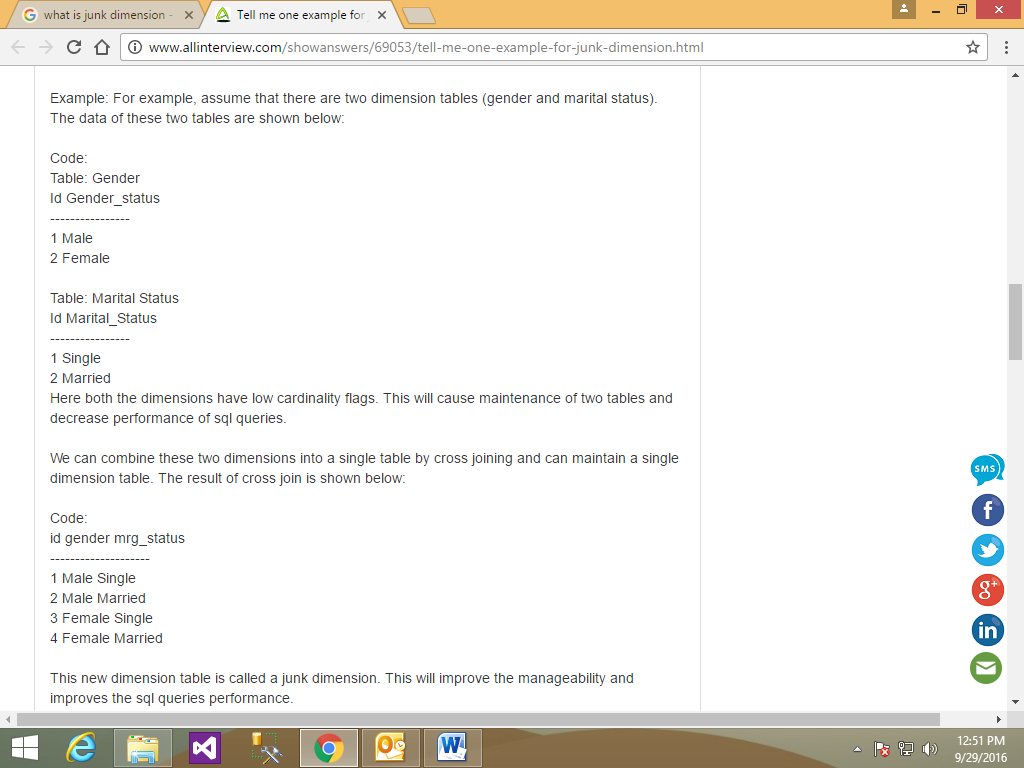
 number of columns have to be equal to the number of columns in the GROUP BY clause,





**Indicators** are minimal gauges that convey the state of a single data value at a glance and are mostly used to represent state value of Key Performance **Indicator** (KPI).

A **key performance indicator** (**KPI**) is a business metric used to evaluate factors that are crucial to the success of an organization.



drill up : One level up in the Hirearchy

Drill down: One level down in the Hirearchy

drill by : direct selection of level in the Hirearchy

drill trough : to drill data from one Hirearchy to another Hirearchy

An example to drill through is in the case of two reports that are in a master /detail relation with each other, and by clicking a master item on the master report you reach the details of the clicked item on the details report.

The ROLLUP operator allows SQL Server to create subtotals and grand totals, while it groups data using the GROUP BY clause.

In case of non-numeric data, JOINs are slower and occupies more space than an integer.

A Natural Key is a type of key in a table which uniquely identifies each record and has a business meaning attached to it.

* History cannot be maintained if used as a Primary Key.

A Surrogate Key is a type of key in a table which uniquely identifies each record, but has no business meaning attached to it. It is merely a value used to uniquely identify a record in a table.

It is useful because the natural primary key (i.e. Customer Number in Customer table) can change and this makes updates more difficult.

Another benefit you can get from surrogate keys (SID) is :

Tracking the SCD - Slowly Changing Dimension.

A Hierarchy is a collection of one or more related Attributes which are organized in a Parent-Child fashion. An Attribute at a higher level is a parent of an Attribute at the next level and so on.

A Role-Playing Dimension is a Dimension which is connected to the same Fact Table multiple times using different Foreign Keys. This helps the users to visualize the same cube data in different contexts/angles to get a better understanding and make better decisions.

*Example*: Consider a Time Dimension which is joined to the same Fact Table (Say FactSales) multiple times, each time using a different Foreign Key in the Fact Table like Order Date, Due Date, Ship Date, Delivery Date, etc. Essentially there is only one single physical dimension called Date Dimension. However, it is joined multiple times to the Fact Table to help the users to visualize the cube data in the context of different dates.

A Conformed Dimension is a Dimension which connects to multiple Fact Tables across one or more Data Marts (cubes). A Confirmed Dimension is physically implemented across multiple Data Marts with exactly the same structure, attributes, values (dimension members), meaning and definition.

*Example*: A Date Dimension has exactly the same set of attributes, same members and same meaning irrespective of which Fact Table it is connected to or to which Data Mart it belongs to as long as it belongs to the same organization. For instance, a Fiscal Calendar is exactly the same with same start and end dates across all the departments within an organization.

* It is derived from the Fact Table and does not have an underlying physical Dimension Table of its own.
* It is also called as a Fact Dimension.
* Since these dimensions are built on top of Fact Table, these are usually very large dimensions.
* The attribute of a Degenerate Dimension is not a Foreign Key in the Fact Table.

*Example*: Degenerate Dimensions having unique Order Numbers can be used to identify the various items sold as part of a particular order.

A Junk Dimension is often a collection of Unrelated Attributes like indicators, flags, codes, etc. This Dimension usually contains data which cannot be created as a separate Dimension as they tend to be too small (often) and are not worth having a separate Dimension

indicators, flags, etc

* It is also called as a Garbage Dimension.
* Junk Dimensions are usually small in size.
* It is a convenient way to consolidate many Smaller Dimensions which are not really worth having as a separate Dimension.

A Named Query is a SQL query/expression in your DSV which acts as a Table.

* Named Queries are used in various scenarios, following are some of the common scenarios:
  + Combining data from multiple Tables/Views from the underlying data source by using either simple or complex join operations.
  + Adding filter conditions for filtering out unwanted data or selecting the required data (limiting the data).

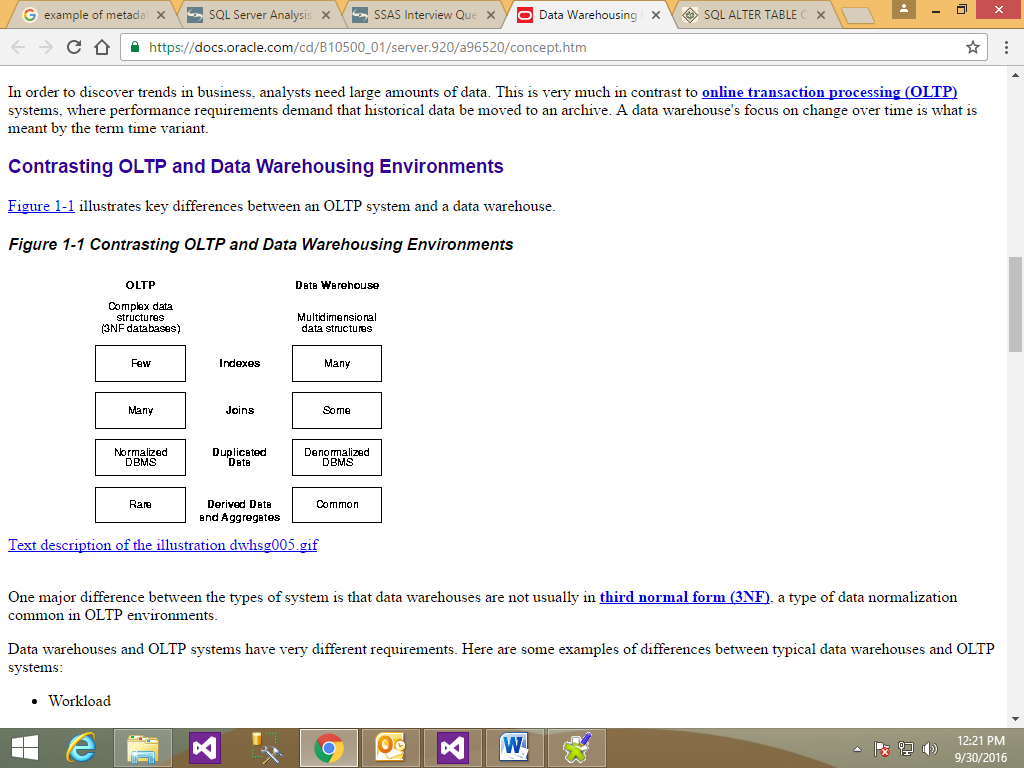
A Named Calculation is a new column added to a Table in DSV and is based on an expression.

* A [Data Source View (DSV)](https://www.mssqltips.com/sqlservertutorial/2004/creating-a-star-schema-using-a-data-source-view/) is a logical view of the underlying database schema and offers a layer of abstraction for the underlying database schema. This layer acts as a source for SSAS and captures the schema related information from the underlying database. The schematic information present in DSV includes the following:
  + Underlying database Table(s)/View(s) metadata
  + [Primary Key & Foreign Key relationships](https://www.mssqltips.com/sql-server-tip-category/74/referential-integrity/) between the underlying database Table(s)
  + Additional columns in the form of Named Calculations
  + Complex logic on the underlying Table(s)/View(s) in the form of Named Queries
* SSAS can only see the schematic information present in the DSV and it cannot see the schematic information from the underlying database.

For example, at the control-flow level, you can add a failure constraint that, if an error occurs, redirects the workflow to a specified alternate task. Similarly, in the data flow, if a row causes an error in a transformation, you can send the row out an error path. SSIS even includes event-handling capabilities that let you trap OnWarning and OnError events.

A data warehouse is a relational database that is designed for query and analysis rather than for transaction processing.

Indexes are special lookup tables that the database search engine can use to speed up data retrieval. Simply put, an index is a pointer to data in a table.



**Data integrity** is a fundamental component of information security. In its broadest use, “**data integrity**” refers to the accuracy and consistency of **data** stored in a database, **data** warehouse, **data** mart or other construct

A **measure** is also commonly called a fact. The term "**measures**" and "facts" are used interchangeably. A **Measure Group** is a collection/**group** of **measures** which belong to the same underlying fact table. In **SSAS**, typically each **Measure Group** is tied to each one of the underlying fact tables.

**Domain integrity** specifies that all columns in a relational database must be declared upon a defined **domain**. The primary unit of data in the relational data model is the data item. Such data items are said to be non-decomposable or atomic. A **domain** is a set of values of the same type.

**Pentaho Data Integration PDI (Kettle)** is the default ETL tool for the **Pentaho** ecosystem. On its very intuitive graphical editor (**Spoon**) it is easy to build Data Integration procedures. The procedures can be run by **Kettle runtime**in different ways: Using the command line utility (Pan), a small server (Carte), a database repository (Kitchen) or directly from the IDE (Spoon).

**Kettle is an interpreter of ELT procedures saved in XML format, while Talend Open Studio is a Java code generator tool.**

Talend is more focused on **data integration**, **data quality**and **data management** solutions, while Pentaho is focused on **Business Intelligence**.

One of Talend Open Studio requirements is to define the **correct schema** of the data to be processed and the IDE helps a lot on this task. Kettle is more flexible on this and the ETL procedures can be built quickly.

**Talend Open Studio** is an Eclipse based Java tool.