Q. Define SSIS?

Ans:

SQL Server Integration Services — commonly known as SSIS is the new platform that was introduced in SQL Server 2005, for data transformation and data integration solutions. This replaced the DTS in SQL Server 2000.

Q. Name a few SSIS components?

Ans:

* Integration Services Projects
* Integration Services Packages
* Control Flow Elements
* Data Flow Elements
* Integration Services Connections
* Integration Services Variables
* Integration Services Event Handlers
* Integration Services Log Providers

Q. What is a project and Package in SSIS?

Ans:

Project is a container for developing packages. Package is nothing but an object. It implements the functionality of ETL — Extract, Transform and Load — data.

Q. What are the 4 elements (tabs) that you see on a default package designer in BIDS?

Ans:

Control Flow, Data Flow, event Handler and package explorer. (Parameters – 2012 Data Tools)

Q. What is a Control flow and Data Flow elements in SSIS?

Ans:

Control Flow:

Control flow element is one that performs any function or provides structure or control the flow of the elements. There must be at least one control flow element in the SSIS package. In SSIS a workflow is called a control-flow. A control-flow links together our modular data-flows as a series of operations in order to achieve a desired result.

A control flow consists of one or more tasks and containers that execute when the package runs. To control order or define the conditions for running the next task or container in the package control flow

Data Flow:

All ETL tasks related to data are done by data flow elements. It is not necessary to have a data flow element in the SSIS package. A data flow consists of the sources and destinations that extract and load data, the transformations that modify and extend data, and the paths that link sources, transformations, and destinations. Before you can add a data flow to a package, the package control flow must include a Data Flow task. The Data Flow task is the executable within the SSIS package that creates, orders, and runs the data flow. A separate instance of the data flow engine is opened for each Data Flow task in a package.

Q. What are the 3 different types of control flow elements in SSIS?

Ans:

* Structures provided by Containers
* Functionality provided by Tasks
* Precedence constraints that connect the executables, containers, and tasks into an ordered control flow.

Q. What are the 3 data flow components in SSIS?

Ans:

* Source
* Transformation
* Destination

Q. What are connections and connection managers in SSIS?

Ans:

Connection as its name suggests is a component to connect to any source or destination from SSIS — like a sql server or flat file or lot of other options that SSIS provides. Connection manager is a logical representation of a connection.

Q. What is the use of Check Points in SSIS?

Ans:

SSIS provides a Checkpoint capability which allows a package to restart at the point of failure.

Q. What are the command line tools to execute SQL Server Integration Services packages?

Ans:

DTSEXECUI – When this command line tool is run a user interface is loaded in order to configure each of the applicable parameters to execute an SSIS package.

DTEXEC – This is a pure command line tool where all of the needed switches must be passed into the command for successful execution of the SSIS package.

Q. Can you explain the SQL Server Integration Services functionality in Management Studio?

Ans:

You have the ability to do the following:

* Login to the SQL Server Integration Services instance
* View the SSIS log
* View the packages that are currently running on that instance
* Browse the packages stored in MSDB or the file system
* Import or export packages
* Delete packages
* Run packages

Q. Can you name some of the core SSIS components in the Business Intelligence Development Studio you work with on a regular basis when building an SSIS package?

Ans:

* Connection Managers
* Control Flow
* Data Flow
* Event Handlers
* Variables window
* Toolbox window
* Output window
* Logging
* Package Configurations

Q. Name Transformations available in SSIS?

Ans:

DATACONVERSION: Converts columns data types from one to another type. It stands for Explicit Column Conversion.

DATAMININGQUERY: Used to perform data mining query against analysis services and manage Predictions Graphs and Controls.

DERIVEDCOLUMN: Create a new (computed) column from given expressions.

EXPORTCOLUMN: Used to export a Image specific column from the database to a flat file.

FUZZYGROUPING: Used for data cleansing by finding rows that are likely duplicates.

FUZZYLOOKUP: Used for Pattern Matching and Ranking based on fuzzy logic.

AGGREGATE: It applies aggregate functions to Record Sets to produce new output records from aggregated values.

AUDIT: Adds Package and Task level Metadata: such as Machine Name, Execution Instance, Package Name, Package ID, etc..

CHARACTERMAP: Performs SQL Server column level string operations such as changing data from lower case to upper case.

MULTICAST: Sends a copy of supplied Data Source onto multiple Destinations.

CONDITIONALSPLIT: Separates available input into separate output pipelines based on Boolean Expressions configured for each output.

COPYCOLUMN: Add a copy of column to the output we can later transform the copy keeping the original for auditing.

IMPORTCOLUMN: Reads image specific column from database onto a flat file.

LOOKUP: Performs the lookup (searching) of a given reference object set to a data source. It is used for exact matches only.

MERGE: Merges two sorted data sets into a single data set into a single data flow.

MERGEJOIN: Merges two data sets into a single dataset using a join junction.

ROWCOUNT: Stores the resulting row count from the data flow / transformation into a variable.

ROWSAMPLING: Captures sample data by using a row count of the total rows in dataflow specified by rows or percentage.

UNIONALL: Merge multiple data sets into a single dataset.

PIVOT: Used for Normalization of data sources to reduce anomalies by converting rows into columns

UNPIVOT: Used for de-normalizing the data structure by converts columns into rows in case of building Data Warehouses.

Various Types of Transformations in SSIS-2012  
  
In SSIS-2012 there are nearly 30 types of transformations.  
  
Depending on the functionality transformations are divided into the following five groups:

1. Business Intelligence Transformations
2. Split And Join Transformations
3. Row Transformations
4. Row-Set Transformations
5. Other Transformations.

Business Intelligence Transformations  
  
This transformation is classified into the following six types:

1. Fuzzy Group Transformation: The Fuzzy Group Transformation is used do data cleansing by finding rows that are likely duplicates and reduce the number of duplicates with in a dataset based on the Matching Decision. This Transformation will accept only string data types while reducing the amount of duplicate data.
2. Fuzzy Lookup Transformation: The Fuzzy Lookup Transformation is used to return close matches of reference data for the incoming data stream. It can be matches and standardizes the data based on fuzzy logic.
3. Term Extraction Transformation: Extracts terms (nouns and noun phrases) from the input text into the transformation output column.
4. Term Lookup Transformation: Extracts terms from the input column with TEXT data type and match them with the same or similar terms found in the lookup table. Each term found in the lookup table is scanned for in the input column. If the term is found then the transformation returns the value as well as the number of times it occurs in the row. You can configure this transformation to do a case-sensitive search.
5. Data Mining Query Transformation: Queries a data mining model. Includes a query builder to assist you with the development of Data Mining eXpressions (DMX) prediction queries.
6. Data Cleansing Transformation: This transformation is used to do automating data cleansing and monitoring the overall status of the data cleansing process.

Split and Join Transformations  
  
This transformation is classified into the following seven types:

1. Cache Transformation: This transformation is used to store the data as a file or in memory for use in a look up transformation
2. Conditional Split: This transformation is used to accepts an input and determine which destination to pipe the data into based on the result of an expression. It redirect rows of data that meet specific conditions to different outputs
3. Look-Up Transformation: This Look Up Transformation is used to join the input data set to the reference table, view or row set created by a SQL statement to lookup corresponding values. If some rows in the input data do not have corresponding rows in the lookup table then you must redirect such rows to a different output.
4. The **Fuzzy Lookup transformation** performs data cleaning tasks such as standardizing data, correcting data, and providing missing values. The **Fuzzy Lookup transformation** differs from the **Lookup transformation** in its use of **fuzzy** matching.
5. Merge Transformation: This transformation is used to merge two sorted inputs into a single output based on the values of the key columns in each data set. Merged columns must have either identical or compatible data types
6. Merge-Join Transformation: This transformation is used to merge two datasets into a single dataset using a JOIN function.
7. Multicast Transformation: This transformation is used to send a copy of the data to an additional path in the workflow. It duplicates the data in the dataflow that again we can send the data in parallel, or when we want to send the data to multiple destinations simultaneously.
8. Union-All Transformation: Combines multiple inputs into a single output. Rows are sorted in the order they're added to the transformation. You can ignore some columns from each output, but each output column must be mapped to at least one input column.

Row Transformations  
  
This transformation is classified into the following six types.

1. Character Map Transformation: The Character Map transformation allows you to do character operations on string columns. It makes common string data changes for you.
2. Copy Column Transformation: This transformation is used to add a copy of column to the transformation output. You can later transform the copy. Makes a copy of a single or multiple columns that will be further transformed by subsequent tasks in the package
3. Data Conversion Transformation: This transformation is used to convert a column data type to a new (another) column data type.
4. Derived Column Transformation: This transformation is used to apply expression to a data column and create a new derived column calculated from an expression.
5. OLEDB Command Transformation: Runs a SQL command for each input data row. Normally your SQL statement will include a parameter (denoted by the question mark)
6. Script Component Transformation: This transformation is used to do a custom transformation.It uses a script to transform the data and you can apply specialized business logic to your data flow.

Row Set Transformations  
  
This transformation is classified into the following six types:

1. Aggregate Transformation: This transformation is used to aggregates the data from transformation or source and it aggregates the values by group.
2. Row Sampling Transformation: This transformation is used to capture a sampling of the data from the data flow by using a row count of the data flow's total rows. It Loads only a subset of your data, defined as the number of rows and it randomly selects the data and is delivered to somewhere.
3. Percentage Sampling Transformation: Loads only a subset of your data, defined as the percentage of all rows in the data source. It randomly selects percentage of rows.
4. Sort Transformation: This transformation is used to sort the data in the data flow by a given column and discard with duplicate values (optionally eliminating duplicates).
5. Pivot Transformation: This transformation is used to pivot the data on a column into a more non-relational form. It converts rows into columns.
6. UnPivot Transformation: This transformation is used to unpivot the data from a non-normalized format to a relational format.

Other Transformations  
  
This transformation is classified into the following five types:

1. Audit Transformation: This transformation is used to expose auditing information from the package to the data pipe, such as package execution and Execution Time.
2. Row Count Transformation: This transformation is used to count the rows in the data flow and stores them as a variable.
3. SCD Transformation: This transformation maintains the historical values of the dimension members when new members are introduced. It automatically generates transformations for TYPE1 and TYPE2 SCD's.
4. Export Transformation: This transformation is used to export the column from the data flow to the system.
5. Import Transformation: This transformation is used to read data from files and appends it to the data flow.

Finally we can see the List of SSIS-2012 Transformations alphabetically as below:

1. SSIS Aggregate Transformation
2. SSIS Audit Transformation
3. SSIS Cache Transform Transformation
4. SSIS Character Map Transformation
5. SSIS Conditional Split Transformation
6. SSIS Copy Column Transformation
7. SSIS Data Conversion Transformation
8. SSIS Data Mining Query Transformation
9. SSIS Derived Column Transformation
10. SSIS DQS Cleansing Transformation
11. SSIS Export Column Transformation
12. SSIS Fuzzy Grouping Transformation
13. SSIS Fuzzy LookUp Transformation
14. SSIS Import Column Transformation
15. SSIS Look Up Transformation
16. SSIS Merge Transformation
17. SSIS Merge Join Transformation
18. SSIS Multicast Transformation
19. SSIS OLEDB Command Transformation
20. SSIS Percentage Sampling Transformation
21. SSIS PIVOT Transformation
22. SSIS Row Count Transformation
23. SSIS Row Sampling Transformation
24. SSIS Script Component Transformation
25. SSIS Slowly Changing Dimension Transformation
26. SSIS Sort Transformation
27. SSIS Term Extraction Transformation
28. SSIS Term Look Up Transformation
29. SSIS Union All Transformation
30. SSIS UnPivot Transformation

1. What are the two authentication modes in SQL Server?

There are two authentication modes –

* Windows Mode
* Mixed Mode

Modes can be changed by selecting the tools menu of SQL Server configuration properties and choose security page.