What are the significant parts of the SSIS package?

Answer: The significant parts in the SSIS package are: Data flow,   Package Explorer, Event handler and Control flow.

#### What is Solution Explorer in SSIS?

Answer: Solution Explorer in SSIS stands for a screen where you can take a look and access all the data sources, projects, and other sorts of related files.

#### What is the meaning of data flow in SSIS?

Answer: Data flow in SSIS stands for the flow of data from the matching sources to the anticipated destinations.

#### What is an SSIS package?

Answer: An SSIS package is a prearranged set of connections like data flow elements, event handlers, variables, control events, and configurations. You collect them by the tools provided by SSIS.

#### What are different types of connections that support SSIS?

Answer: Different types of connection that support SSIS are: ODBC, Flat File, XML, Excel, OLEDB, .net SQLClient.

#### How many kinds of containers are present in SSIS?

Answer: In SSIS, a container represents a reasonable grouping of tasks, and it allows dealing with the scope of a task collectively. The kinds of containers in SSIS are Sequence container and Task host container.

#### What is the meaning of Precedence Constraint in SSIS?

Answer: Precedence Constraint in SSIS allows you to identify the reasonable sequence of tasks, and in the arrangement, they should be implemented.

#### What are the kinds of variables in SSIS?

Answer: Variable in SSIS are utilized to store values as the system variable and user variable.

#### What does a checkpoint in SSIS mean?

Answer: Checkpoint in SSIS permits the project to start again from the point of breakdown. Checkpoint stores the details about the package execution; if the package runs effectively the checkpoint file is removed.

#### What are connection managers in SSIS?

Answer: While collecting data from different sources and writing it for a purpose, connection managers are useful.  Connection manager eases the connection to the system that incorporates informations like data provider, server name, etc.

#### What does SSIS breakpoint mean?

Answer: SSIS breakpoint allows you to suspend the implementation of the package in the business intelligence expansion stage.

#### What is event logging in SSIS?

Answer: Event logging in SSIS permits you to choose any particular event of a job and it is advantageous at what time you are getting trouble in your package.

#### What does logging mode property mean?

Answer: SSIS packages and all the connected jobs have a property known as logging mode.   This property allows three possible values in the shape of Enabled, Disabled and Parent Setting.

#### Data checkpoint data is not saved for what?

Answer: Checkpoint data is not saved for Loop containers.

#### What is data flow buffer?

Answer: SSIS operates by means of buffers; it is a sort of an in-memory practical table to hold data.

#### What do conditional split transactions in SSIS mean?

Answer: Conditional split transformation in SSIS verifies the prearranged condition, on the basis of the condition assessment.

#### What are the different types of Data viewers in SSIS?

Answer: Different types of data viewers are Grid, Column Chart, Histogram, and Scatter Plot.

#### What are the possible locations for saving the SSIS package?

Answer: SSIS package can be saved at SQL Server, File System or Package Store.

#### What is the Event Handlers tab in SSIS?

Answer: With the aid of the Event Handlers tab in SSIS, workflows can be easily configured to act in response to package events.

#### How you can inform your workforce about package failure?

Answer: You can add a Send Mail Task in the event handlers. Also, you can set the warning in the SQL Agent at what time the package fails to run.

#### How to perform incremental load?

Answer: The greatest and top way to perform incremental load is by utilizing a Timestamp column in the supply table and storing the final ETL timestamp.

#### How logging is done in SSIS?

Answer: Logging in SSIS can be done by means of events like on Warning, on Error, etc. to the manifold options like a SQL server table, XML, etc.

#### How is the SSIS package deployed on production?

Answer: In order to deploy the SSIS package, we have to implement the manifest files and need to settle on whether to deploy this onto SQL Server of into File System.

#### What are three data flow constituents in SSIS?

Answer: Three data flow constituents are SSIS are Source, Transformation, and Destination.

#### Why check Points are utilized in SSIS?

Answer: Checkpoint utilized in SSIS permits a package to start again at the point of breakdown.

#### What are different alternatives for dynamic configuration is SSIS?

Answer: Different alternatives for dynamic configuration are   Customer variables, XML file, and Database per set with the variables

#### What is Data conversion Transformation?

Answer: Data conversion is the most excellent method to transmit the data from one category to another.  On the other hand, you need to ensure that you have well-matched data in the column.

#### Q31) What other tasks can SSIS handle?

Answer: Some of the other tasks that SSIS handles are data profiling, effectively streamline batch operations and system filing.

#### Q32) What do you understand by SSIS control flow system integration?

Answer: This factor consists of using the logical connectors between the tasks and thus graphically control the flow of data in the system.

#### Q33) What do you mean by data transformation?

Answer: When a series of rules or functions are applied to the extracted data from the source so that it can be loaded onto the end result, it is known as data transformation.

#### Q34) Can variables be created in SSIS?

Answer: Yes, variables can be created in SSIS.

#### Q35) Mention the two types of variables that can be created in SSIS?

Answer: The two variables that can be created in SSIS are task level variables and global variables.

#### Q36) What types of containers can be used with SSIS packages?

Answer: The types of containers that can be used are sequence containers and loop containers.

#### Q37) What do you mean by the Deployment process in SSIS?

Answer: When a package is transformed from the development mode into the executables mode, it is known as the Deployment process.

#### Q38) What do you understand by the Manifiest file in SSIS?

Answer: When the package is transferred using wizard on the file system and the SQL server, it is known as the Manifiest file.

#### Q39) What is known as the data flow task in SSIS?

Answer: When the data is moved from its source to its destination, this mechanism is known as the data flow task.

#### Q40) What is meant by data profiling task?

Answer: When source data is analyzed to understand it better in terms of hygiene, patterns, numbers, etc it is known as data profiling task.

[Get SSIS Certification with Expert](https://www.gangboard.com/microsoft-training/microsoft-ssis-training?utm_source=GB_blog)

#### Q41) When is the task of data profiling usually done?

Answer: The task of data profiling is generally carried out at the start of the development cycle so that any teething concerns can be immediately addressed.

#### Q42) What do you understand by a transaction in the SSIS package?

Answer: In order to maintain data integrity, packages use processes called transactions to bind the database actions.

#### Q43) What is the purpose of the config file in SSIS?

Answer: This is mainly used to provide different input to the connection manager about the different properties.

#### Q44) What is the difference between control flow and data flow?

Answer: Control flow is what determines the flow or process of the package. Data flow is a subset of the control flow. Without a control flow, the data flow cannot function.

#### Q45) What is a Checkpoint in SSIS?

Answer: The functionality or property in SSIS which allows the project (in case of failure) to restart from that point itself, is known as Checkpoint.

#### Q46) What is the main use of the SSIS package?

Answer: This is an upgraded version that allows the seamless merging of data from numerous sources. It is also used to clear up and analyze the existing data.

#### Q47) Name 4 important components of the SSIS package?

Answer: The four important components of the SSIS package are… Data Flow, Control Flow, Package Explorer, Event Handler

#### Q48) What is the solution explorer in SSIS?

Answer: The solution explorer in SSIS is a screen in which all the data, data sources, projects, etc. can be viewed.

#### Q49) What do you mean by a Task in SSIS?

Answer: Just as in any other programming language, a Task is what executes an individual unit of work

#### Q50) How many categories of tasks are there in SSIS?

Answer: There are two units of Tasks in SSIS… Control Flow tasks and Database Maintenance Tasks

#### Q51) What is Precedence Constraint in SSIS?

Answer: This process enables users to define the logical sequence of tasks to be executed.

#### Q52) What do you mean by connection managers in SSIS?

Answer: As the name suggests connection managers help facilitate the seamless connection of data from its source to the destination.

#### Q53) What do you understand by SSIS breakpoint?

Answer: When a package is getting uploaded in the business intelligence system and is suddenly stopped midway, that point is known as the breakpoint.

#### Q54) What is event logging in SSIS?

Answer: This is the term given when a specific task or event needs to be logged or selected. This is especially useful during any troubleshooting issues.

#### Q55) What is data flow buffer in SSIS?

Answer: In SSIS, this is a virtual, in-memory table that holds data.

#### Q56) Name the 4 data viewers in SSIS?

Answer: The four data viewers in SSIS include Grid, Histogram, Scatter Plot, Column Chart.

#### Q57) Mention three locations where the SSIS package can be saved?

Answer: The three locations where the SSIS package cane be potentially saved are SQL Server, Package Store, File System.

#### Q58) Which are the the three data flow components in SSIS?

Answer: The three data flow components in SSIS are Source, Transformation and Destination

#### Q59) Name three options for dynamic configuration in SSIS?

Answer: The three options for dynamic configuration in SSIS are… XML File, Customer Variables and one Database per environment in sync with the variables.

#### Q60) Mention some disadvantages of SSIS?

Answer: Some disadvantages of SSIS are that the vision and strategy are sometimes unclear. It can create problems in a non-Windows environment. It also does not yet provide support for alternative data integration methods.

[Get SSIS Training](https://www.gangboard.com/microsoft-training/microsoft-ssis-training?utm_source=GB_blog)

#### Q61) Different types of debugging in SSIS?

There are 3 types of debugging

* Execute Package partially
* By using breakpoints
* By using Data Viewers

#### Q62) what is a checkpoint?

Using checkpoint we can restart the package from a failure of the task

#### Q63) what is package configuration?

The configuration provides user interaction to the package, so a user can control the package without  
editing package from solution explorer.

#### Q64) What are the types of configurations in SSIS?

* XML file configuration
* SQL server DB configuration
* Parent package variable configuration
* Environment variable
* Registry entry.

#### Q65) What are the different types of command-line utilities in SSIS?

* DTEXEC
* DTEXECUI
* DTUTIL

#### Q66) What is the difference between merge and Union All?

* Merge will combine only two sorted dataset values.
* Union ALL will combine data from multiple data sets

#### Q67) How many types of containers?

For loop container – we can loop only for a specific number of times  
For each loop container – Loop will be repeated based on the number of files  
in the folder and we can traverse to subfolders with this container.  
Sequence container – to group the certain the tasks  
How to implement transaction control?  
By setting the Transaction Option property to Required or Supported

#### Q68) How to schedule the package?

By using SQL Server Agent.

#### Q69) Difference between Full cache, partial cache and no cache?

* FULL cache – gets the complete reference data set from the database and places in memory and compares it with source data.
* Partial Cache – Here we can specify how much amount of memory that the reference dataset has to occupy.
* No Cache – Here data will not be cached. It is not recommended.

#### Q70) Difference between DTS and SSIS?

Everything is different except both are a product of Microsoft.

#### Q71) How to achieve parallelism in SSIS?

Using the MaxConcurrent Executable property of the package. Its default is -1 and it is calculated as a number of processors +2

#### Q72) Types of Loggings?

* Flat file
* SQL Server table
* XML
* SQL Profiler

#### Q73) What is CDC Splitter?

After data has read out of a table with CDC enable, this transform send data that  
should be deleted inserted and updated down a different path.

#### Q74) What is OLEDB Command T\R?

Execute an OLEDB command for each row in the data flow.

#### Q75) What is the Dimension table?

it is a table which contains details data.

#### Q76) Different types of transformation in SSIS?

there are two different types of transformations

* Synchronous transformation’s
* Asynchronous transformation’s

#### Q77) What is the use CDC control task?

Using CDC we can Maintains and interacts with the change capture feature form SQL Server.

#### Q78) Difference between Merge join and lookup?

In Merge join, we can implement left outer join, right outer join and full outer. Whereas in lookup, we can perform inner join only.

#### Q79) What is the event handler?

An event handler is like a trigger in our SQL databases which will be executed automatically when an action takes place.

* Onerror
* Onwarning
* OnPreexecute
* OnPostExecute
* OnVariableValueChanged
* OnQueryCancled

#### Q80) Difference between control flow and data flow?

* Control flow is process-oriented and Data flow is information oriented.
* Data will not be moved from one task to the other task in control flow whereas data will be moved from one transformation to another transformation in the data flow.
* In control flow, we have Precedence constraints to set the task execution whereas in data flow we have data flow paths.

#### Q81) Different types of package storage mechanisms

* File system
* MSDB

#### Q82) List out different types of source systems you have used in your current project

* Flat file
* Excel
* Ole DB source
* ADO.NET

#### Q83) How would you do Logging in SSIS?

Logging configuration provides log the details of various events an inbuilt feature.

#### Q84) What is Integration Services?

It is a platform including packages that provide extract, building high-performance data integration solutions transformation and load processing for DWH.

#### Q85) What is the difference between union and merge in SSIS?

Merge: can take only 2 data sets, it requires the data to be sorted  
Union: can take multiple inputs and doesn’t need the data to be sorted

#### Q86) How does sort transformation work in the ssis, how does it impact the performance?

Sort transformation sorts the data by reading all the rows and sorting them completely. When  
performed on a large data set it slower. It is better to have the data sorted by the “order by” in the  
SQL query.

#### Q87) What is Multithreading?

It is a process of processing the data-parallel to reduce the time of execution of the package and  
to process a large amount of data faster.

#### Q88) In which isolation level phantom read does not occur?

Snapshot and Serializable

#### Q89) How to change the isolation level in SQL server?

This can be achieved by using the SQL statement

SET TRANSACTION ISOLATION LEVEL

{ READ UNCOMMITTED

| READ COMMITTED

| REPEATABLE READ

| SNAPSHOT

| SERIALIZABLE

}

#### Q90) When is the activation of the rcsi take place?

The activation of this has to be performed manually. It will not be created during the creation of the Database.  
It can be done by using the query

USE Master;

  GO

  ALTER DATABASE demo\_db SET READ\_COMMITTED\_SNAPSHOT ON;

  GO

  USE demo\_db;

  GO

#### Q91) What does the master database hold?

It has information such as logins, linked servers, endpoints, system configuration settings.  
information about the other databases on their instances and their physical file locations.

#### Q92) How to perform data validation in SSIS?

This can be performed by using a script task, and writing code in c# or VB to validate the data  
and give the result.

#### Q93) Can the SSRS RDL in a report server be exported to EXCEL in from a SSIS package .?

Yes, it can be performed by using a Script task.

#### Q94) Define the limitation and advantage in using the SQL command as an input source when input in the SQL data source in a data flow task?

Using a SQL command to extract the data from the SQL server is efficient in reducing the load on the site and reduces the execution time. it helps us prevent the use of separate ssis tasks for merging and joining and union. Its limitation is that the typing space is limited to a certain number and the query cannot be bigger. In that instance, it can be defined as a stored procedure and can be called in the SQL command.

#### Q95) How do we create an excel file and the sheet with defined column and data type in SSIS?

It can be done by using the create table statement in execute the SQL task component and connecting it to an empty excel connection.

#### Q96) Name a few blocking transformations?

Sorting and aggregation.

#### Q97) Can the data be viewed as between data tasks?

Yes data can be viewed by using “enable data viewer”

#### Q98) Name the ssis data types for the varchar and nvarchar data types of SQL server?

Varchar – &gt; dt\_str  
Nvarchar – &gt; dt\_wstr

#### Q99) What are the deployment models in SSIS?

Package deployment and project deployment models.

**1. Define SSIS?**  
**Answer:**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.

**2. Name a few SSIS components?**  
**Answer:** 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

**3. What is a project and Package in SSIS?**  
**Answer:** Project is a container for developing packages. The package is nothing but an object. It implements the functionality of ETL — Extract, Transform and Load — data.

**4. What are the 4 elements (tabs) that you see on a default package designer in BIDS?**  
**Answer:** Control Flow, Data Flow, event Handler, and package explorer. (Parameters – 2012 Data Tools)  
For More Enroll For SQL Server Online Training.

**5. What are a Control flow and Data Flow elements in SSIS?**  
**Answer:** Control Flow: Control flow element is one that performs any function or provides structure or controls 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. Control-Flow links together our modular data-flows as a series of operations in order to achieve the 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. ( [oracle apex training online](https://svrtechnologies.com/apex-training/) )

**6. What are the 3 different types of control flow elements in SSIS?**  
**Answer:**  
Structures provided by Containers  
The functionality provided by Tasks  
Precedence constraints that connect the executables, containers, and tasks into an ordered control flow.

**7. What are the 3 data flow components in SSIS?**  
**Answer:**  
Source  
Transformation  
Destination

**8. What is a breakpoint in**[**SSIS**](https://en.wikipedia.org/wiki/SQL_Server_Integration_Services)**? How is it setup? How do you disable it?**  
**Answer:** A breakpoint is a stopping point in the code. The breakpoint can give the Developer\DBA an  
Opportunity to review the status of the data, variables and the overall status of the SSIS package.  
10 unique conditions exist for each breakpoint. Breakpoints are setup in BIDS. In BIDS, navigate to the control flow interface. Right-click on the  
Object where you want to set the breakpoint and select the ‘Edit Breakpoints…’ option. ( [oracle apex training online](https://svrtechnologies.com/apex-training/) )

**9. What is the use of the Bulk Insert Task in SSIS?**

**Answer:** Bulk Insert Task is used to upload a large amount of data from flat files into Sql Server. It supports only OLE DB connections for the destination database.

**10. What is Conditional Split transformation in SSIS?**  
**Answer:** This is just like IF condition which checks for the given condition and based on the condition evaluation, the output will be sent to the appropriate OUTPUT path. It has ONE input and MANY outputs. Conditional Split transformation is used to send paths to different outputs based on some conditions. For example, we can organize the transform for the students in a class who have marks greater than 40 to one path and the students who score less than 40 to another path.  ( [tableau training videos](https://svrtechnologies.com/tableau-training/)  )

**11. How do you eliminate quotes from being uploaded from a flat file to SQL Server?**  
**Answer:** This can be done using TEXT QUALIFIER property. In the SSIS package on the Flat File Connection Manager Editor, enter quotes into the Text qualifier field then preview the data to ensure the quotes are not included.

**12. Can you explain how to set up a checkpoint file in SSIS?**  
**Answer:** The following items need to be configured in the properties tab for the SSIS package:  
Checkpoint Filename – Specify the full path to the Checkpoint file that the package uses to save the value of package variables and log completed tasks. Rather than using a hard-coded path, as shown above, it’s a good idea to use an expression that concatenates a path defined in a package variable and the package name.

Checkpoint Usage – Determines if/how checkpoints are used. Choose from these options: Never (default), If Exists, or Always. Never indicates that you are not using Checkpoints. If Exists is the typical setting and implements the restart at the point of failure behavior. If a Checkpoint file is found it is used to restore package variable values and restart at the point of failure. If a Checkpoint file is not found the package starts execution with the first task. The Always choice raises an error if the Checkpoint file does not exist.

Save Checkpoints – Choose from these options: True or False (default). You must select True to implement the Checkpoint behavior

**13. What are the command line tools to execute SQL Server Integration Services packages?**  
**Answer:**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. ( [data science training](https://svrtechnologies.com/data-science-training/) )

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

**14. Name Transformations available in SSIS?**  
**Answer:** DATA CONVERSION: 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 an 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.

**CHARACTER MAP**: 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 the column to the output we can later transform the copy keeping the original for auditing.

**IMPORTCOLUMN**: Reads image specific column from a 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.

**MERGE JOIN**: 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

**15. What is Ignore Failure option in SSIS?**  
**Answer:**In Ignore Failure option, the error will be ignored and the data row will be directed to continue on the next transformation. Let’s say you have some JUNK data (wrong type of data or JUNK data) flowing from source, then using this option in SSIS we can REDIRECT the junk data records to another transformation instead of FAILING the package. This helps to MOVE only valid data to destination and JUNK can be captured into a separate file

**16. What are the containers? What are the different types of containers in SSIS?**  
**Answer:**Containers are objects that provide structures to packages and extra functionality to tasks. There are four types of containers in SSIS, they are: For Each Loop Container, For Loop Container, Sequence Container and Task Host Container  ( [devops training online](https://svrtechnologies.com/devops-training/))

**17. What are the different types of data sources available in SSIS?**  
**Answer:**There are 7 types of data sources provided by SSIS: a.) Data Reader source b.) Excel source c.) Flat file sourced.) OLEDB source e.) Raw file source f.) XML source g.) Script component.

**18. What is the function of the control flow tab in SSIS?**  
**Answer:** On the control flow tab, the tasks including dataflow task, containers and precedence constraints that connect containers and tasks can be arranged and configured.

**19. What is the function of the Event handlers tab in SSIS**?  
**Answer:**On the Event handlers tab, workflows can be configured to respond to package events.  
For example, we can configure Work Flow when ANY task Fails or Stops or Starts.

**20. What is the function of the Package explorer tab in SSIS?**  
**Answer:** This tab provides an explorer view of the package. You can see what is happening in the package. The Package is a container at the top of the hierarchy.

**21. What is Solution Explorer?**  
**Answer:**It is a place in SSIS Designer where all the projects, Data Sources, Data Source Views, and other miscellaneous files can be viewed and accessed for modification.

**22. How do we convert data type in SSIS?**  
**Answer:** The Data Conversion Transformation in SSIS converts the data type of an input column to a different data type.

**23. Explain Conditional split Transformation?**  
**Answer:** It functions as if…then…else construct. It enables send input data to a satisfied conditional branch. For example, you want to split product quantity between less than 500 and greater or equal to 500. You can give the conditional a name that easily identifies its purpose. Else section will be covered in Default Output Column name.  
After you configure the component, it connects to subsequent transformation/destination, when connected, it pops up a dialog box to let you choose which conditional options will apply to the destination transformation/destination. ( [oracle apex training online](https://svrtechnologies.com/apex-training/) )

**24. Explain Copy column Transformation**?  
**Answer:** This component simply copies a column to another new column. Just like ALIAS Column in T-Sql.

**25. Explain Data conversion Transformation?**  
**Answer:** This component does conversion data type, similar to TSQL function CAST or CONVERT. If you wish to convey the data from one type to another then this is the best bet. But please make sure that you have COMPATIBLE data in the column.

**26. Explain Data mining query Transformation?**  
**Answer:** This component does prediction on the data or fills the gap in it. Some good scenarios use this component is:  
1. Take some input columns as a number of children, domestic income, and marital income to predict whether someone owns a house or not.  
2. Take prediction of what a customer would buy based analysis buying pattern on their shopping cart.  
3. Filling blank data or default values when a customer doesn’t fill some items in the questionnaire. ( [hadoop online training](https://svrtechnologies.com/hadoop-training/))

**27. Explain Derived column Transformation?**  
**Answer:** Derived column creates a new column or puts manipulation of several columns into a new column. You can directly copy existing or create a new column using more than one column also. ( [data science online training](https://svrtechnologies.com/data-science-training/) )

**28. Explain Merge Transformation?**  
**Answer:**  
Merge transformation merges two paths into a single path. It is useful when you want to break out data into a path that handles errors after the errors are handled, the data are merged back into downstream or you want to merge 2 data sources. It is similar to Union All transformation, but Merge has some restrictions:  
1. Data should be in sorted order  
2. Data type, data length, and another Metadata attribute must be similar before merged.

**29. Explain Merge Join Transformation?**  
**Answer:** Merge Join transformation will merge output from 2 inputs and doing INNER or OUTER join on the data. But if you the data come from 1 OLEDB data source, it is better you join through SQL query rather than using Merge Join transformation. Merge Join is intended to join 2 different data source.

**30. Explain Multicast Transformation?**  
**Answer:** This transformation sends output to multiple output paths with no conditional as Conditional Split does. Takes ONE Input and makes the COPY of data and passes the same data through many outputs. In simple give one input and take many outputs of the same data. ( [python training](https://svrtechnologies.com/python-training/) )

# **SSIS Scenario Based Interview Questions and Answers Pdf**

**1. Explain the Percentage and row sampling Transformations?**  
**Answer:** This transformation will take data from the source and randomly sampling data. It gives you 2 outputs. First is selected data and the second one is unselected data. It is used in a situation where you train a data mining model. These two are used to take the SAMPLE of data from the input data.

**2. Explain Sort Transformation?**  
**Answer:** This component will sort data, similar in TSQL command ORDER BY. Some transformations need sorted data.

**3. Explain Union all Transformation?**  
**Answer:** It works in the opposite way to merge transformation. It can take the output from more than 2 input paths and combines into a single output path.

**4. What are SSIS Connection Managers?**  
**Answer:** When we talk of integrating data, we are actually pulling data from different sources and writing it to a destination. But how do you get connected to the source and destination systems? This is where the connection managers come into the picture. Connection manager represents a connection to a system which includes data provider information, the server name, database name, authentication mechanism, etc. For more information check out the [SQL Server Integration Services (SSIS)](https://en.wikipedia.org/wiki/SQL_Server_Integration_Services) Connection Managers and Connection Managers in SQL Server 2005 Integration Services SSIS tips.

**5. What is the Retain Same Connection property and what is its impact?**  
**Answer:** Whenever a task uses a connection manager to connect to the source or destination database, a connection is opened and closed with the execution of that task. Sometimes you might need to open a connection, execute multiple tasks and close it at the end of the execution. This is where Retain Same Connection property of the connection manager might help you. When you set this property to TRUE, the connection will be opened on the first time it is used and remain open until execution of the package completes.

**6. What are a source and destination adapters?**  
**Answer:** A source adaptor basically indicates a source in Data Flow to pull data from. The source adapter uses a connection manager to connect to a source and along with it, you can also specify the query method and query to pull data from the source.  
Similar to a source adaptor, the destination adapter indicates a destination in the Data Flow to write data to. Again like the source adapter, the destination adapter also uses a connection manager to connect to a target system and along with that you also specify the target table and writing mode, i.e. write one row at a time or do a bulk insert as well as several other properties.  
Please note, the source and destination adapters can both use the same connection manager if you are reading and writing to the same database.

**7. What is the Data Path and how is it different from a Precedence Constraint?**  
**Answer:**Data Path is used in a Data Flow task to connect to different components of a Data Flow and show the transition of the data from one component to another. A data path contains the Meta information of the data flowing through it, such as the columns, data type, size, etc. When we talk about differences between the data path and precedence constraint; the data path is used in the data flow, which shows the flow of data. Whereas the precedence constraint is used in control flow, which shows control flow or transition from one task to another task

**8. What is SSIS event logging?**  
**Answer:**Like any other modern programming language, SSIS also raises different events during package execution life cycle. You can enable or write these events to trace the execution of your SSIS package and its tasks. You can also write your custom message as a custom log. You can enable event logging at the package level as well as at the tasks level. You can also choose any specific event of a task or a package to be logged. This is essential when you are troubleshooting your package and trying to understand a performance problem or root cause of a failure. Check out this tip about Custom Logging in SQL Server Integration Services SSIS.

**9. What is the Logging Mode property?**  
**Answer:** SSIS packages and all of the associated tasks or components have a property called Logging Mode. This property accepts three possible values: Enabled – to enable logging of that component, Disabled – to disable logging of that component and Use Parent Setting – to use parent’s setting of that component to decide whether or not to log the data.

**10. Can we add our custom code in SSIS?**  
**Answer:** To add your own code:-  
1. In the control flow tab, drag and drop Script Task from the toolbox.  
2. Double click on script task to open and select edit to open Script task editor.  
3. In the script task editor, there are 3 main properties  
i.) General – Here you can specify the name and description  
ii.) Script – through this we can add our code by clicking on Design Script button. The scripting language present is VB.Net only.  
iii.) Expression.

**11. Explain the use of containers in SSIS and also their types?**  
**Answer:**Containers can be defined as objects that store one or more tasks. The primary purpose of the container is grouping logically related tasks. Once the task is placed into the containers, we can perform various operations such as looping on container level until the desired criterion is met. Nesting of the container is allowed. The container is placed inside the control flow.  
There are 4 types of Container:-  
1. Task Host container- Only one task is placed inside the container. This is the default container.  
2. Sequence Container – This container can be defined as a subset of package control flow.  
3. For loop container – Allows looping based on condition. Runs a control flow till the condition is met.  
4. for each loop container – Loop through container-based on the enumerator.

**12. Error Handling in SSIS?**  
**Answer:** An error handler allows us to create flows to handle errors in the package is quite an easy way. Through event handler tab, we can name the event on which we want to handle errors and the task that needs to be performed when such an error arises. We can also add sending mail functionality in event of any error through SMTP Task in the Event handler. This is quite useful in event of any failure in office non-working hours. In the Data flow, we can handle errors for each connection through following failure path or red arrow.

**13. Explain why variables called the most powerful component of SSIS?**  
**Answer:**Variable allows us to dynamically control the package at runtime. Example: You have some custom code or script that determines the query parameter’s value. Now, we cannot have a fixed value for the query parameter. In such scenarios, we can use variables and refer the variable to the query parameter. We can use variables for like:  
1. Updating the properties at runtime,  
2. Populating the query parameter value at runtime,  
3. used in a script task,  
4. Error handling logic and  
5. with various looping logic.

**14. What is precedence constraint?**  
**Answer:** A precedence constraint is a link between 2 control flow tasks and lays down the condition on which the second task is run. They are used to control the workflow of the package. There are 3 kinds of precedence constraint – success (green arrow), failure (red arrow) or Completion script task (blue arrow). By default, when we add 2 tasks, it links by a green arrow. The way the precedence constraint is evaluated can be based on the outcome of the initial task. Also, we can add expression to evaluate such an outcome. Any expression that can be judged as true or false can be used for such purpose. The precedence constraint is very useful in error handling in SSIS package.

**15. How can you ensure your package works when you move it to another server?**  
**Answer:**In SSIS 2005 till 2008 R2 (also exists in 2012) we can use Package Configuration which is commonly used with SQL Server table (or XML) and environment variable. To be honest Package Configuration was a pain in the next and I much prefer SSIS 2012 Project Deployment Model with Parameters and SSIS Catalog. For more info visit our SSIS 2012 Configuration Tutorial.

**16. What is the data flow buffer?  
Answer:**SSIS operates using buffers which is kind of an in-memory virtual table to hold data. We are able to set max rows/size but this gets more interesting with blocking transformations.

**17. What is the difference between Union All and Merge Transformation?**  
**Answer:**Union All and Merge are similar but Merge requires sorted input and accepts only 2 inputs where we Union All can accept more than 2.

**18. How to send an SSRS report from SSIS?**  
**Answer:**Often there is a requirement to be able to send an SSRS report in Excel, PDF or another format to different users from an SSIS package once it has finished performing a data load. In order to do this, first, you need to create a subscription to the report. You can create an SSRS report subscription from Report Manager. At the report subscription, you can mention the report format and the email address of the recipient. When you create a schedule for the SSRS report, a SQL Server Agent Job will be created. From the SSIS, by using sp\_start\_job and passing the relevant job name you can execute the SSRS report subscription.

**19. What is Attunity Driver and why do we user in SSIS?**  
**Answer:**Attunity provides 2 high-speed connectors. One for Oracle and one for Teradata. They have been selected by Microsoft to be included with SQL Server 2008 Integration Services (SSIS) SQL 2008 Enterprise Edition. These drivers are highly optimized and very easy to use

**20. How checkpoint works in for loop?**  
**Answer:** For each Loop, the container is another atomic unit of work that can be restarted. However, the checkpoint file does not contain information about the work completed by the child containers, and the for each Loop container and its child containers run again when the package restarts.

**. What Is A Container And How Many Types Of Containers Are There?**  
**Answer**: A container is a logical grouping of tasks which allows you to manage the scope of the tasks together.  
These are the types of containers in SSIS:  
Sequence Container – Used for grouping logically related tasks together  
For Loop Container – Used when you want to have repeating flow in a package  
For Each Loop Container – Used for enumerating each object in a collection; for example a recordset or a list of files.  
Apart from the above-mentioned containers, there is one more container called the Task Host Container which is not visible from the IDE, but every task is contained in it (the default container for all the tasks).

**2. What Is A Precedence Constraint And What Types Of Precedence Constraint Are There?**  
**Answer:** [SSIS](https://en.wikipedia.org/wiki/SQL_Server_Integration_Services) allows you to place as many tasks you want to be placed in the control flow. You can connect all these tasks using connectors called Precedence Constraints. Precedence Constraints allow you to define the logical sequence of tasks in the order they should be executed. You can also specify a condition to be evaluated before the next task in the flow is executed.  
These are the types of precedence constraints and the condition could be either a constraint, an expression or both Success (next task will be executed only when the last task completed successfully) or Failure (next task will be executed only when the last task failed) or Complete (next task will be executed no matter the last task was completed or failed).

**3. What Is A Transformation?**  
**Answer**: A transformation simply means bringing in the data in the desired format. For example, you are pulling data from the source and want to ensure only distinct records are written to the destination, so duplicates are removed. Another example is if you have master/reference data and want to pull only related data from the source and hence you need some sort of lookup. There are around 30 transformation tasks available and this can be extended further with custom-built tasks if needed.

**4. If You Want To Send Some Data From Access Database To Sql Server Database. What Are Different Component Of Ssis Will You Use?**  
**Answer:** In the data flow, we will use one OLE DB source, data conversion transformation, and one OLE DB destination or SQL server destination. OLE DB source is data source is useful for reading data from Oracle, SQL Server, and Access databases. Data Conversion transformation would be needed to remove datatype abnormality since there is a difference in datatype between the two databases (Access and SQL Server) mentioned. If our database server is stored on and the package is run from the same machine, we can use SQL Server destination otherwise we need to use OLE DB destination. The SQL Server destination is the destination that optimizes the SQL Server.

**5. How To Handle Late Arriving Dimension Or Early Arriving Facts?**  
**Answer:** Late-arriving dimensions sometimes get unavoidable ‘coz delay or error in Dimension ETL or may be due to a logic of ETL. To handle Late Arriving facts, we can create a dummy Dimension with natural/business key and keep the rest of the attributes as null or default. And as soon as the Actual dimension arrives, the dummy dimension is updated with Type 1 change. These are also known as Inferred Dimensions.

**6. What is a workflow in SSIS 2014?**  
**Answer:** Workflow is a set of instructions on to specify the Program Executor on how to execute tasks and containers within SSIS Packages

**7. What is the difference between Multicast and Conditional Split?**  
**Answer:** The Multicast transformation distributes its input to one or more outputs. This transformation is similar to the Conditional Split transformation. Both transformations direct input to multiple outputs. The difference between the two is that the Multicast transformation directs every row to every output, and the Conditional Split directs a row to a single output.

**8. What is Manifest file in SSIS?**  
**Answer**: The manifest file is the utility which can be used to deploy the package using the wizard on the file system and SQL Server database.

**9. How to back up or retrieve the SSIS packages?**  
**Answer**: If your package is deployed on SQL Server then you can back up the MSDB database as all the package on SQL server deploys at MSDB.

**10. What kind of variables can you create?**  
**Answer:** You can create global variables and task level variables in SSIS. For programmers, these variables are the same as global and function level variables. A global variable is available to all tasks across the entire job. Variables created in tasks are only available within that task.