PENTAHO COURSE CONTENT

Introduction of Pentaho BI

- What is Pentaho BI?

- The history of Pentaho BI

- Why choose Pentaho?

- Installation overview

- Additional Tools of Pentaho

Steps of Pentaho ETL

- What is ETL?

- What is Database and Data Warehouse?

- What is BI?

- BI tools available in industry

- About Pentaho Data Integration

- Different types of files

- Components in Pentaho DI

- What is Transformation and Job?

- Hops

- Transformation Folders

- Transformation Steps

- Input Folder and Output Folder steps

- How to create a connection to DB?

- How to move data from CSV file input to table output?

- How to move data from CSV file input to Microsoft excel output?

- Read data from multiple files at a time with same metadata

- Add date/time in output file name

- Perform DML Operations

- How to write data to log

- Transform Folder Steps

- Add constants

- Add sequence

- Difference between "Add sequence" and "Add value fields changing sequence"

- Calculator

- Concat Fields

- Number Range

- Replace in String

- Row Flattener

- Select Values

- Set Field Value

- Set Field Value to a Constant

- Sort Rows

- Split Field to Rows

- Split Fields

- String Operations

- String Cut

- Unique Rows

- Unique Rows (HashSet)

- Value Mapper

- Utility Folder Steps

- Clone row

- Delay row

- If Field Value is Null

- Null If

- Write to Log

- Table Compare

- Append Streams

- Block this step until steps finish

- Blocking Step

- Dummy

- Filter Rows

- Prioritize Streams

- Switch / Case

- Formula

- Modified Java Script Value

- User Defined Java Expression

- What is Variable?

- How to create and use Variables?

- Lookup Folder

- Check if a column exists

- Database lookup

- File lookup

- File Exists

- Stream lookup

- Table Exists

- Join Rows (Cartesian Product)

- Merge Join

- Merge Rows

- Multiway Merge Join

- Sorted Merge

- Dimension Lookup/Update

- How to get Previous/Next row values at current row

- Group By

- How to call a transformation inside of another transformation

- Get Variable

- Set Variable

- Jobs

- Create a Repository

- Schedule a transformation

- Run transformations via command line

- Write a batch script to execute transformation

- Write a shell script to execute a transformation

- Monitoring a job/transformation on server