**DATA LOADING WITH INCREMENTAL PROCESSING**

**STEP 1 :** To create the **5 tables**

First table is students,

**create table Students (**

**studentid INT PRIMARY KEY, 🡪deltacolumn**

**name VARCHAR(100),**

**email VARCHAR(100) unique,**

**city varchar(100)**

**)**

**insert into Students values (1,'Abhi','xyz@gmail.com','tornto')**

**insert into Students values (2,'Joe','joe@gmail.com','texas')**

**insert into Students values (3,'Anna','anna@gmail.com','chicago')**

**insert into Students values (4,'Sam','sam@gmail.com','la')**

**insert into Students values (5,'Ken','ken@gmail.com','alberta')**

A screenshot of a computer

AI-generated content may be incorrect.

Second table is departments,

**create table Departments (**

**DepartmentID INT PRIMARY KEY, 🡪deltacolumn**

**DepartmentName VARCHAR(100)**

**)**

**insert into Departments values (1,'ADE')**

**insert into Departments values (2,'CYS')**

**insert into Departments values (3,'ADA')**

**insert into Departments values (4,'ADE')**

**insert into Departments values (5,'AI')**

A screenshot of a computer

AI-generated content may be incorrect.

Third table is Courses,

**create table Courses (**

**cid INT PRIMARY KEY,**

**coursename VARCHAR(100),**

**courseupdate datetime, -->deltacolumn**

**DepartmentID INT,**

**FOREIGN KEY (DepartmentID) REFERENCES Departments(DepartmentID)**

**)**

**insert into Courses values (1,'azure data engineer','2020-01-01 00:00:00',1)**

**insert into Courses values (2,'cyber security','2019-01-01 00:00:00',2)**

**insert into Courses values (3,'azure data analyst','2021-01-01 00:00:00',3)**

**insert into Courses values (4,'azure data engineer','2018-01-01 00:00:00',4)**

**insert into Courses values (5'artificial intelligence','2024-01-01 00:00:00',5)**

A screenshot of a computer

AI-generated content may be incorrect.

Fourth Table is Enrollment

**CREATE TABLE Enrollments (**

**enrollid INT PRIMARY KEY,**

**studentid INT,**

**cid INT,**

**enrollDate datetime, -->deltacolumn**

**FOREIGN KEY (studentid) REFERENCES Students(studentid),**

**FOREIGN KEY (cid) REFERENCES Courses(cid)**

**)**

**insert into Enrollments values (100,1,1,'2020-01-01 00:00:00')**

**insert into Enrollments values (101,2,2,'2019-01-01 00:00:00')**

**insert into Enrollments values (102,3,3,'2021-01-01 00:00:00')**

**insert into Enrollments values (103,4,4,'2018-01-01 00:00:00')**

**insert into Enrollments values (104,5,5,'2024-01-01 00:00:00')**

A screenshot of a computer

AI-generated content may be incorrect.

Fifth Tbale is Professors

**create table Professors (**

**professorid INT PRIMARY KEY,**

**name VARCHAR(100),**

**email VARCHAR(100) UNIQUE,**

**DepartmentID INT,**

**joindate datetime, -->deltacolumn**

**FOREIGN KEY (DepartmentID) REFERENCES Departments(DepartmentID)**

**)**

**insert into Professors values (1001,'Emily','emily@gmail.com',1,'2017-01-01 00:00:00')**

**insert into Professors values (1002,'Sarah','sarah@gmail.com',3,'2023-01-01 00:00:00')**

**insert into Professors values (1003,'David','david@gmail.com',5,'2014-01-01 00:00:00')**

A screenshot of a computer

AI-generated content may be incorrect.

**STEP 2 :** To create the **watermark table**

**create table watermark**

**(**

**id int identity(1,1),**

**tablename varchar(100),**

**schemaname varchar(100),**

**foldername varchar(100),**

**lpv varchar(100),**

**deltacolumn varchar(100)**

**)**

A screenshot of a computer

AI-generated content may be incorrect.

Inserting the values in the watermark table

**insert into watermark values('Students','dbo','dbo/Students',0,'Studentid')**

**insert into watermark values('Departments','dbo','dbo/Departments',0,'DepartmentID')**

**insert into watermark values('Courses','dbo','dbo/Courses','1900-01-01 00:00:00','courseupdate')**

**insert into watermark values('Enrollments','dbo','dbo/Enrollments','1900-01-01 00:00:00','enrollDate')**

**insert into watermark values('Professors','dbo','dbo/Professors','1900-01-01 00:00:00','joindate')**

A screenshot of a computer

AI-generated content may be incorrect.

Stored procedure

**create procedure watermark\_update**

**@tablename varchar(100),**

**@lpv varchar(100)**

**as**

**update watermark**

**set lpv=@lpv**

**where tablename=@tablename**

**TO CREATE INCREMENTAL PIPELINE IN THE SYNAPSE**

**Create lookup activity, from source get data from sqldatabase (ssms)**

**Step 1 : Getdata from sources (sqlserver)**

A screenshot of a computer

AI-generated content may be incorrect.

**I didn’t select specific table because will use parameter to fetch the table dynamically**

A screenshot of a computer

AI-generated content may be incorrect.

**To do dynamically follow this steps**

**Step 2: click on the open from sources**



**Step 3: add parameters**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 4: to add schemaname and tablename dynamically**

**Click on add dynamic content and fetch the values**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 5 : in my getwatermark sources the tablename and schemaname will dynamically fetch**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 6 : By clicking preview data, can see this table list**

A screenshot of a computer

AI-generated content may be incorrect.

**Create foreach activity**

**Step 1 : link the lookup and foreach activity**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 2 : in setting, items click on add dynamic content and select the getwatermark data value array**

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

**Step 3 : Inside the foreach activity, add one lookup activity**

**The lookup activity will give max value from source (my sources is the 5 table)**

**Click setting and add dynamic content in tablename and schemaname**

A screenshot of a computer

AI-generated content may be incorrect.

**Add tablename**

A screenshot of a computer

AI-generated content may be incorrect.

**Add schemaname**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 5 : Value is passed inside the foreach (lookup activity)**

A screenshot of a computer

AI-generated content may be incorrect.

**By clicking first row only**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 6 : get the max value in each table (click query and add dynamic content)**

A blue rectangle with text

AI-generated content may be incorrect.

**Query to get max value in each table**

**select max(@{item().deltacolumn}) as maxvalue from @{item().schemaname}.@{item().tablename}**

A screenshot of a computer program

AI-generated content may be incorrect.

**Inside the foreach already we have created lookup and now we are creating copydata activity**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 1 : by giving ‘’ it will get only incremental rows**

A close-up of a computer screen

AI-generated content may be incorrect.

**Step 2 : pass the query**

**select \* from @{item().tablename} where @{item().deltacolumn}>'@{item().lpv}'**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 3 : in sink I wanna load the data in ADLS gen 2**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 4 : in sink create the parameter**

A screenshot of a computer

AI-generated content may be incorrect.

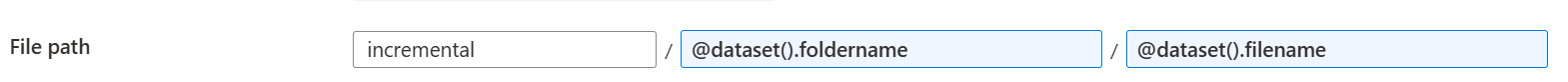
**Step 5 : click on connection and in directory click add dynamic contect and give foldername and do as for filename**

A screenshot of a computer

AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

**Step 6 : both done**

**Step 7 : give the foldername in sink**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 8 : for file name use concat function**

**@concat(item().tablename,'\_',utcNow(),'.csv')**

A screenshot of a computer

AI-generated content may be incorrect.

**Inside the foreach already we have created lookup and now we are creating Stored procedure activity**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 1 : by clicking on linked service, click the existing db which you have created in ssms and in stored procedure give the name which you have created.**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 2 : By clicking import will get this**

A screenshot of a computer

AI-generated content may be incorrect.

**Step 3 : Click on add dynamic conent give the query @activity('getmaxvaluefromsouce').output.firstRow.maxvalue**

A white rectangular object with a black border

AI-generated content may be incorrect.

**PUBLISH THE PIPELINE**

A screenshot of a computer

AI-generated content may be incorrect.

**After the pipeline is successfully completed**

**Check on ssms**

A screenshot of a computer

AI-generated content may be incorrect.

**IN ADLS gen 2 the csv file will be there**

A screenshot of a computer

AI-generated content may be incorrect.

**By inserting the new rows in each table**

**insert into Students values (6,'Robert','robert@gmail.com','la')**

**insert into Departments values (6,'AI')**

**insert into Courses values (6,'artificial intelligence','2025-02-28 00:00:00',6)**

**insert into Enrollments values (105,6,6,'2025-02-28 00:00:00')**

**insert into Professors values (1004,'Danish','danish@gmail.com',3,'2025-02-28 00:00:00')**

A computer code with red text

AI-generated content may be incorrect.

**Output in SSMS**

A screenshot of a computer

AI-generated content may be incorrect.

**Output in Gen2**

