#### NUGM 2024

### BUILDING AN INTELLIGENT SCHEMA

Schema designed to present Business Data in SQL as a preparation for analysis and Business Intelligence

Trey Miller, NovoRoi Systems, LLC



### THE PROBLEM OF UNIDATA

Data is multi-valued.

Those curly braces are the ascii character that separates multivalued data throughout the database.

Sales\_Due\_Date 10-04-17}10-11-17}10-13-17}10-14-17}10-17-17
}10-18-17}10-20-17}10-21-17}11-06-17}11-07-17}11-14-17
}11-15-17}11-16-17}11-17-17}11-20-17}11-21-17}12-08-17

Sales\_Qty

3}1}161}20}20}20}100}30}50}40}185}25}25}25}100}35

For purposes of illustration, consider UniData to organize data horizontally.



### THE PROBLEM OF UNIDATA

SQL Server, and most of the rest of the world, organizes data into tables of columns and rows.

Multiple data elements are separated into vertical rows.

Vertical vs. Horizontal Seems simple enough?

Sales Order Due Date	Sales Order Quantity
10-04-17	3
10-11-17	1
10-13-17	161
10-14-17	20
10-17-17	20
10-18-17	20
10-20-17	100
10-21-17	30
11-06-17	50
11-07-17	40
11-14-17	185
11-15-17	25
11-16-17	25



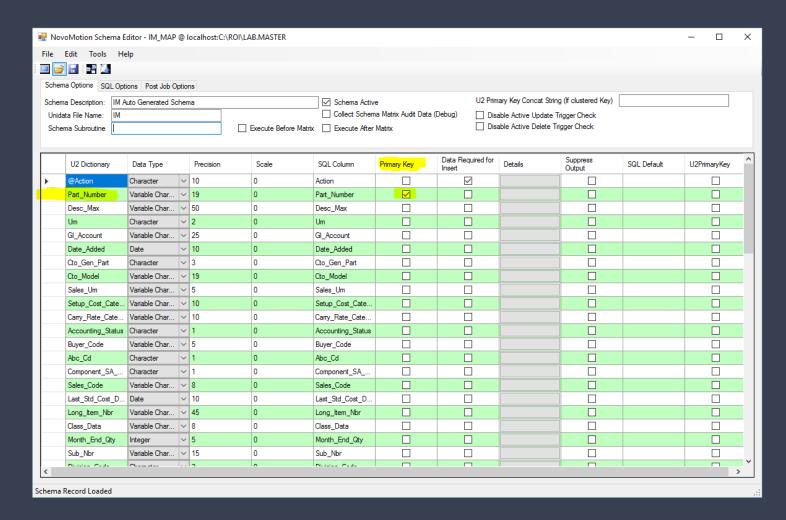
#### UNIDATA IS EXCEPTIONALLY FORGIVING

- The UniData database defines its data elements by Dictionary definition. All databases have some means of organizing the meta-data (data about data) that describes their basic building block. For UniData, this important job falls to the data Dictionary; HOWEVER, where other databases establish meta data as hard and fast rules (what type of data, how many characters long, what are the data associations), UniData treats its own Dictionary definitions as more of a *wish* than a *rule*.
- The other basic problem with UniData's forgiving nature specifically involves the data associations between its multivalued sets. Where a rigid, strongly-typed, database like SQL is completely unforgiving to data-type violations—if you say it's a Date, it better be a Date—another major conflict arises in that SQL server is a hierarchical database such that multi-valued data is organized in columns and rows, and must be associated to other multi-values which have the same number of columns and rows. The data might be "null", but there still needs to be a meta-data place holder to determine that the associated null valued is in the proper column, in the proper row.
- This mapping of horizontal UniData data to the columns and rows of SQL is called data normalization.

# NOVOMOTION IS THE NORMALIZATION SYSTEM BETWEEN UNIDATA AND SQL

The Attributes of UniData are mapped to the Columns of SQL.

The most important element to determine multi-valued association is the Primary Key (PK).

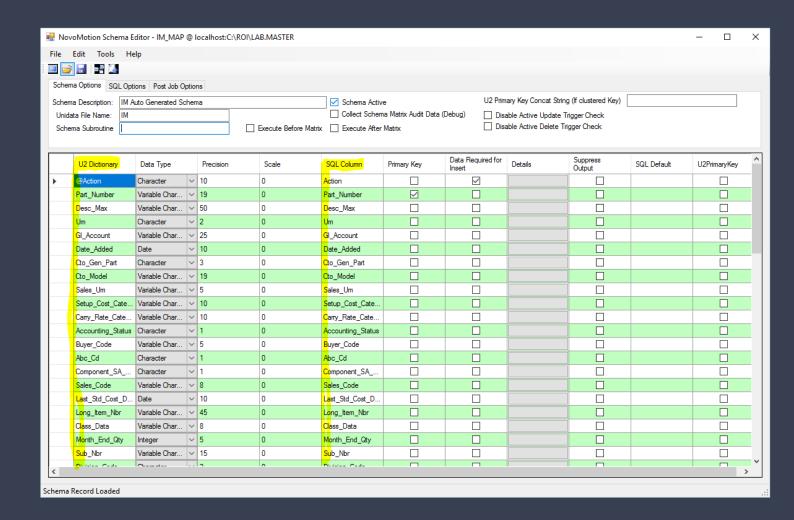




### SINGLE VALUED DATA IS EASY TO MAP

In this example a Primary Key (PK) will only ever have one of these UniData elements and so the resulting table equates perfectly to SQL columns.

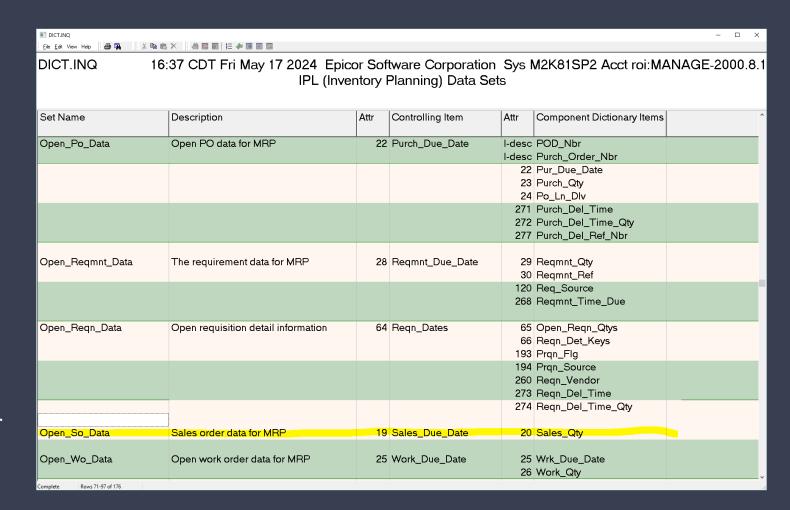
Each PK will be a row, each U2 dictionary will be a column.





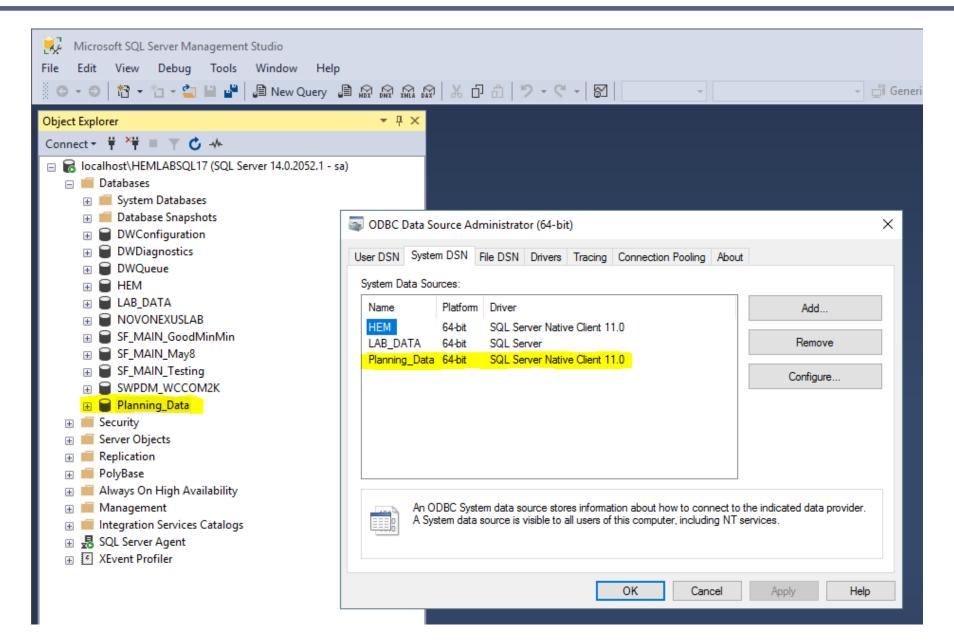
## DICT.INQ IS THE M2K TOOL FOR INVESTIGATING DATA

This is a DICT.INQ listing of sets for the IPL file. Sets are another way of saying 'associated multi-values'. The IPL is a difficult file to normalize. Our Open\_So\_Data is easy to normalize. Some of the other data sets are not.

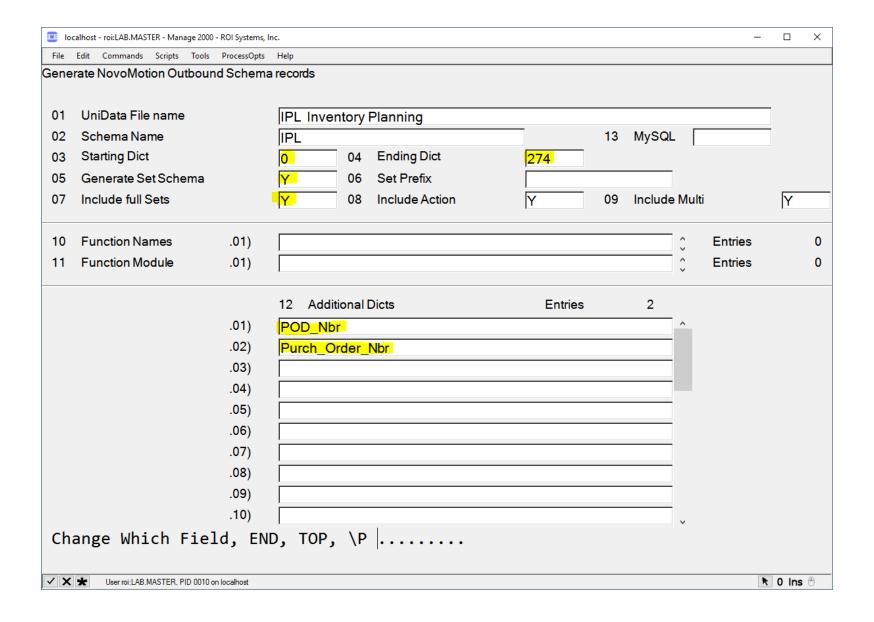




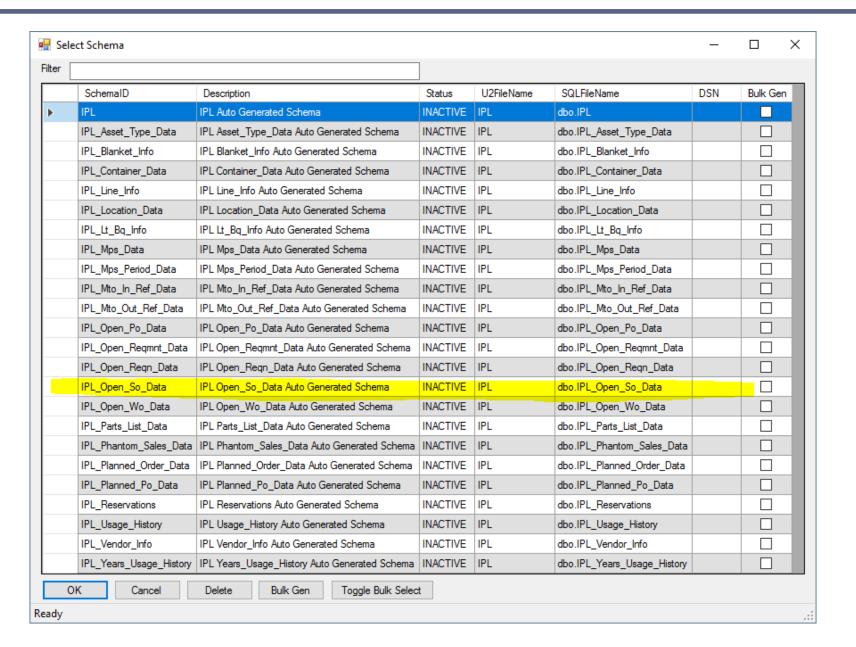
#### PREPARE YOUR SQL REPOSITORY



#### NOVO.GEN.SCHEMA - GENERATE YOUR SCHEMA



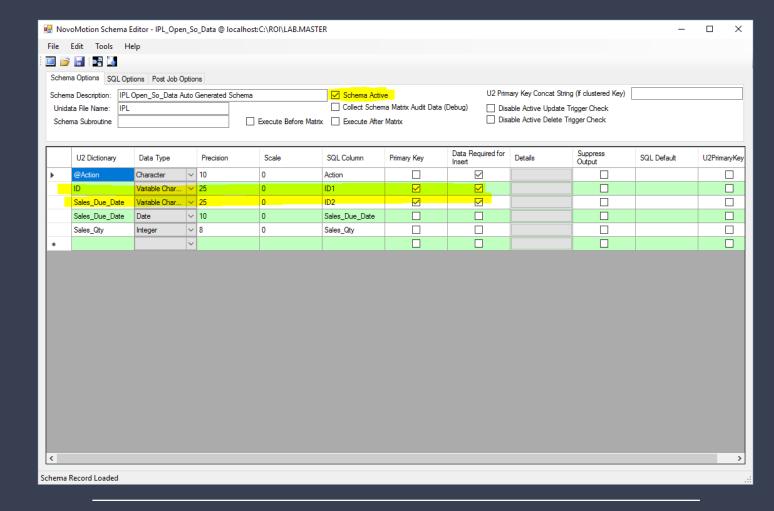
#### NOVOMOTION DIALOGUE - VIEW YOUR RESULTS





#### NOVOMOTION INVESTIGATE PRIMARY KEY CLUSTER

- Activate the schema
- Note ID1, ID2, . . . etc.
- Alter Primary ID to facilitate joins
- Remove redundant lines
- Save your results

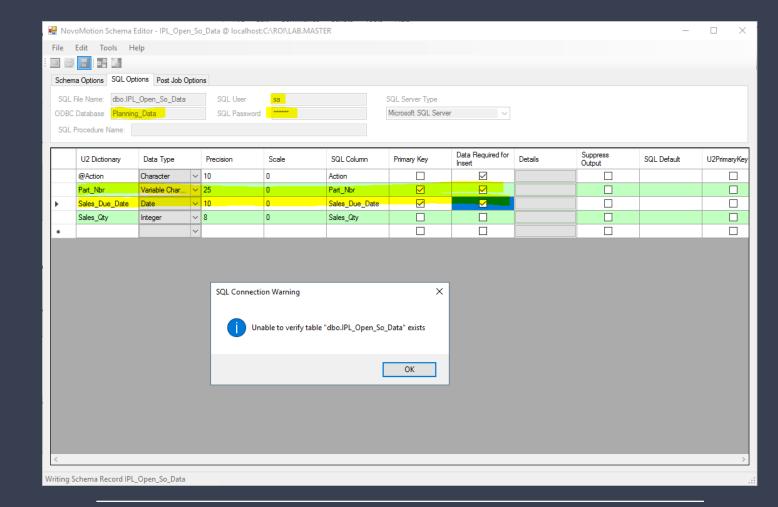




# NOVOMOTION ENTER THE ODBC CREDENTIALS ESTABLISHED FOR THE DATABASE

The credentials are correct if you get the 'Unable to verify table', message.

This means that the NovoMotion system was able to navigate the ODBC pipe and could not find the database object by that name.





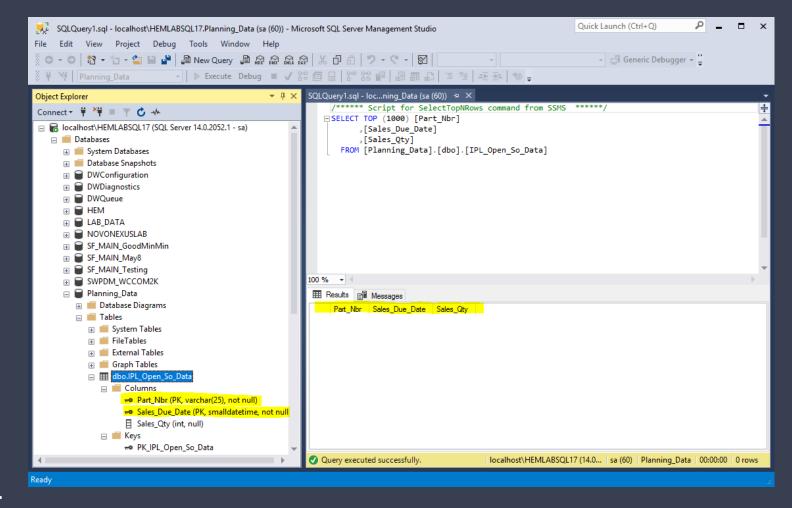
#### NOVOMOTION - GENERATE TABLES AND PROCEDURE

```
Generate SQL
                                                                                                                           Copy to Clipboard
                              Execute on SQL Server
Generate Primary Table Generate Procedure
                                ----- BEGIN for Table [dbo].[IPL Open So Data]----
2 SET ANSI NULLS ON
4 SET QUOTED IDENTIFIER ON
 if exists (select * from sysobjects where id = object id('[dbo].[IPL Open So Data]') and OBJECTPROPERTY(id, N'IsUserTable') = 1)
7 DROP TABLE [dbo].[IPL Open So Data]
9 CREATE TABLE [dbo].[IPL_Open_So_Data](
10 [Part Nbr] [varchar] (25) NOT NULL,
11 [Sales Due Date] [smalldatetime] NOT NULL,
12 [Sales Qty] [int] NULL
13 ) ON [PRIMARY]
15 ALTER TABLE [dbo].[IPL Open_So_Data] WITH NOCHECK ADD
16 CONSTRAINT [PK IPL Open So Data] PRIMARY KEY CLUSTERED
18 [Part Nbr],
19 [Sales Due Date]
20 ) ON [PRIMARY]
                    ----- BEGIN Stored Procedure Code for Table [dbo].[IPL Open So Data]---
23 if exists (select * from sysobjects where id = object id('[dbo].[sp novosys rx IPL Open So Data]'))
24 DROP PROCEDURE [dbo].[sp novosys rx IPL Open So Data]
25 GO
26 CREATE PROCEDURE [dbo].[sp_novosys_rx_IPL_Open_So_Data]
27 @ActionFlag [char] (10),
   @Part Nbr [varchar] (25),
    @Sales Due Date [smalldatetime],
      @Sales Qty [int]
32 SET XACT ABORT, NOCOUNT ON
33 BEGIN TRY
34 IF UPPER(@ActionFlag) = 'C'
36
            IF exists (select Part Nbr FROM [dbo].[IPL Open So Data] WHERE Part Nbr = @Part Nbr AND Sales Due Date = @Sales Due Date)
               HDDATE (dhol (TDI Open So Data)
```



# NOVOMOTION TABLE HAS BEEN GENERATED WITH THE PROPER PRIMARY KEYS

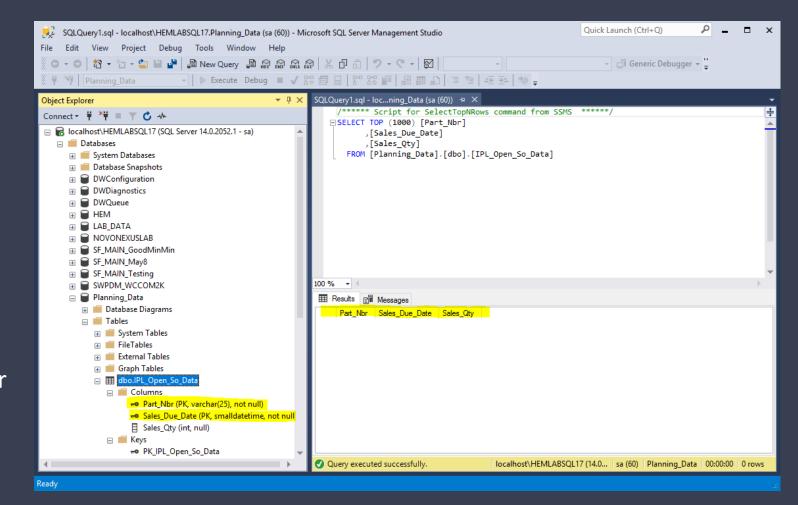
Of the three columns in this multivalued data set, two of the three are the Primary Key designation. This is critical to ensuring the important concept of uniqueness in the SQL data table. The third element is the data, which is not likely to be unique.





## NOVOMOTION LAUNCH FILE LOADER TO INITIALIZE DATA

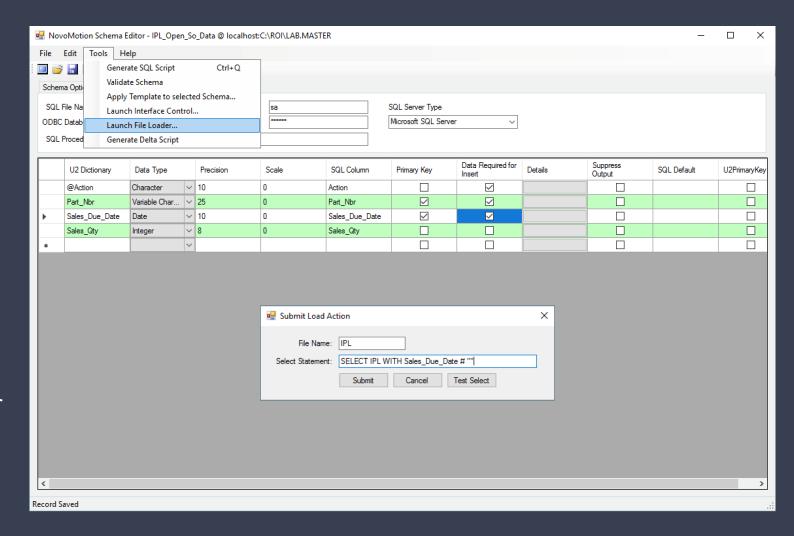
Selection Criteria must follow
UniData syntax rules and the proper
upper and lower case of the
Dictionaries must be utilized in order
to get good selection results.





## NOVOMOTION LAUNCH FILE LOADER TO INITIALIZE DATA

Selection Criteria must follow
UniData syntax rules and the proper
upper and lower case of the
Dictionaries must be utilized in order
to get good selection results.

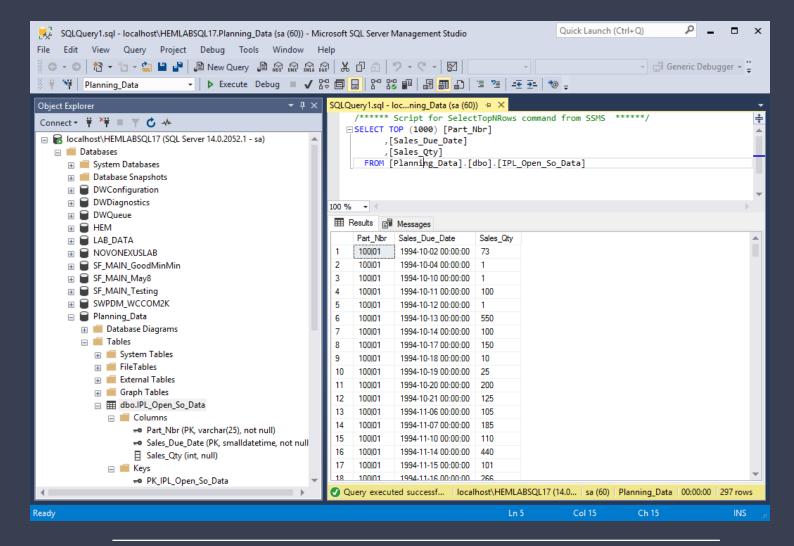




#### NOVOMOTION BACK TO SQL MANAGEMENT STUDIO

The Data is presented in the promised rows and columns.

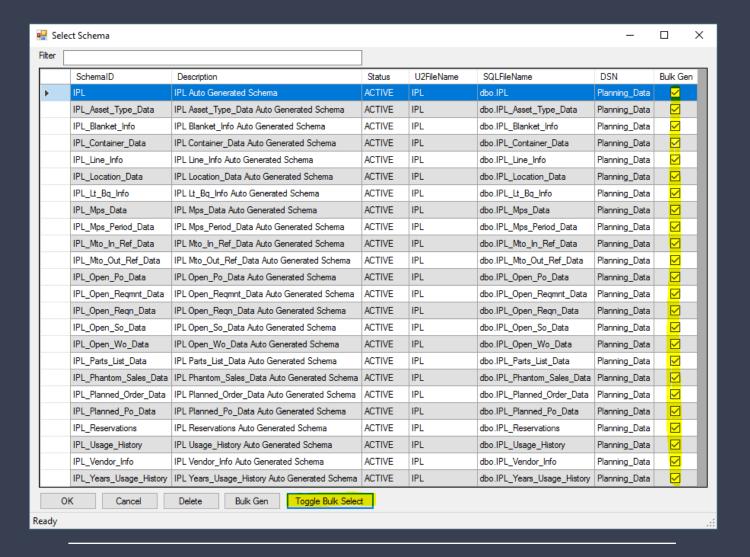
Notice that the primary key cluster was necessary to separate the multivalued data into the proper row structures.





# NOVOMOTION THE SCHEMA DIALOGUE BOX DISPLAYS ACTIVE SCHEMA

The important activity that has occurred on our single test schema is that it has been investigated and altered from the auto generated state. The "intelligence" was provided by a database designer following naming convention, with specific design intentions.

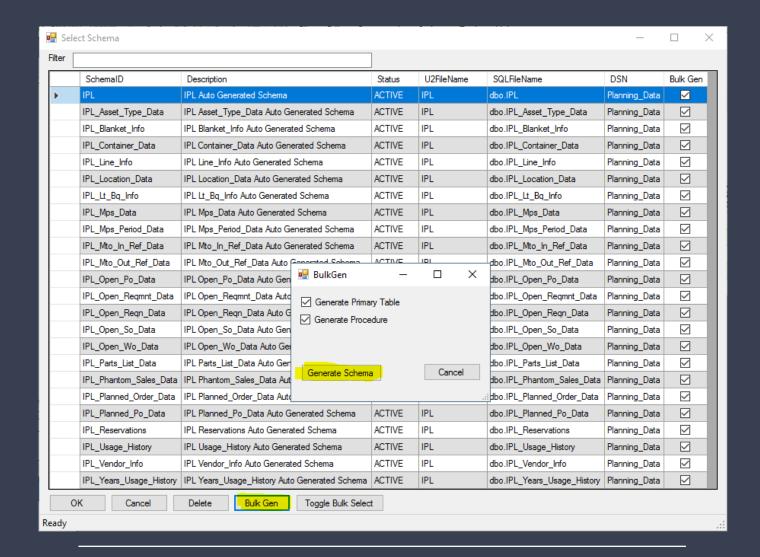




# NOVOMOTION THE SCHEMA BULK TABLE AND PROCEDURE SCRIPT

This process will only work on ACTIVE schema, regardless of being tagged for bulk generation.

A SQL script is produced, and this should be saved to a file location on your workstation. You can edit that script prior to execution in SQL Management Studio.

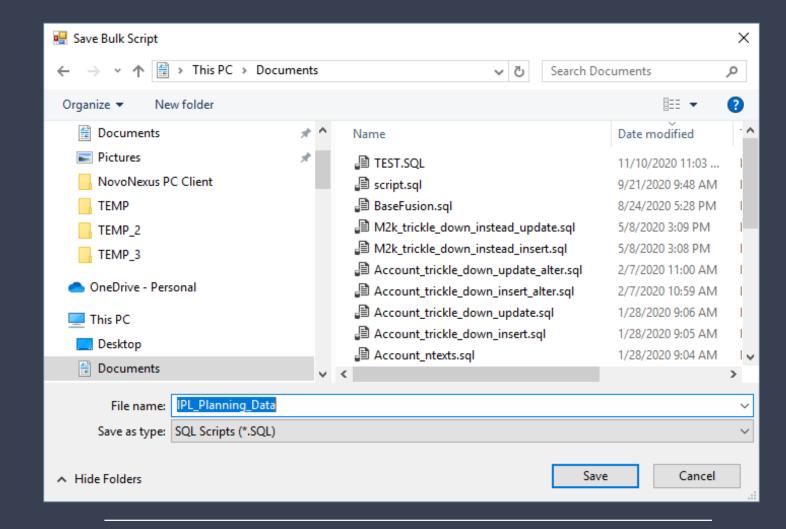




# NOVOMOTION THE SCHEMA BULK TABLE AND PROCEDURE SCRIPT

It is not necessary to save the generated bulk script forever, but you might want to keep a library of scripts to detail how your SQL database came into existence.

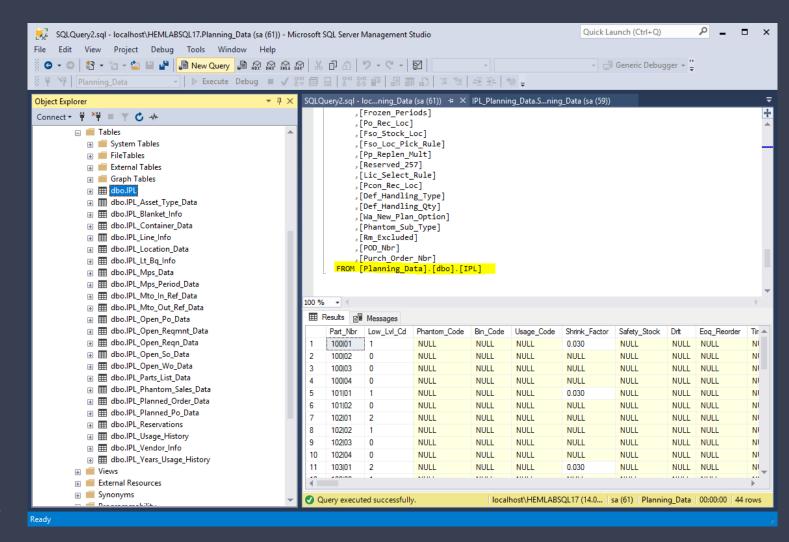
In that regard, a good naming convention is useful.





#### NOVOMOTION LOAD YOUR DATABASE FOR INITIALIZATION VALUES

- Execute the load data process
- Confirm data in new tables
- Now you have an Operational
   Data Store (ODS) ready for data
   reporting or Business Intelligence.





#### NUGM 2024

QUESTIONS?



### NUGM 2024 THANK YOU

