Install Guide - SetupWizard - Azure SQL Server

WhereScape Enablement Pack for Azure SQL Server - RED 10

This is a guide to installing the WhereScape Enablement Pack for Azure SQL Server for WhereScape RED10

Table of Contents

Prerequisites For PostgreSQL Metadata

Before you begin the following prerequisites must be met:

- Create Database and ODBC DSN :
  - Supported* version of PostgreSQL (PostgreSQL 12 or higher)
    - A database to house the RED Metadata Repository.
    - A database for the Range Table DB (Optional)
    - A database to house scheduler (Optional)
- Software Installations
  - WhereScape RED10 with valid license key entered and EULA accepted
  - WhereScape Enablement Pack for target database version RED10
- Windows Powershell (64 bit) version 4 or higher
  - To check Windows Powershell Version:
    - Run below command in Windows Powershell

      ```
      Get-Host|Select-Object Version
      ```

    - Run below command in Command Prompt

      ```
      powershell $psversiontable
      ```

- Run the following command using PowerShell
    - The security protocol TLS 1.0 and 1.1 used by PowerShell to communicate with PowerShell gallery has deprecated and TLS 1.2 has been made mandatory

      ```
      [Net.ServicePointManager]::SecurityProtocol = [Net.ServicePointManager]::
      SecurityProtocol -bor [Net.SecurityProtocolType]::Tls12
      Register-PSRepository -Default -Verbose
      Set-PSRepository -Name "PSGallery" -InstallationPolicy Trusted
      ```

    - Progress bar placeholder info line

      ```
      Install-Module -Name PoshProgressBar -SkipPublisherCheck -Force
      ```

* : *RED supports the following versions for the metadata repository: PostgreSQL 12 or higher*

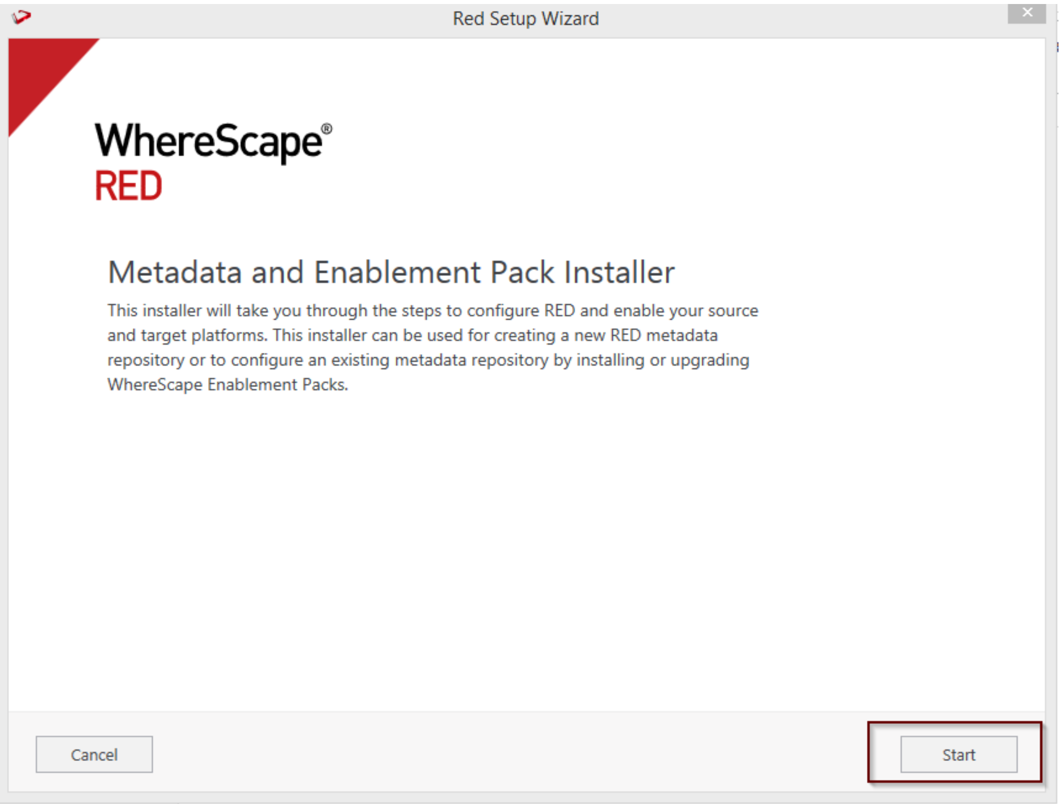
Prerequisites Target Database

Before you begin the following prerequisites must be met:

- Access to an Azure SQL Server/SQL Server (Target Environment), with the following connectivity information :
    1. Server Name
    2. Database Name
    3. User Name
    4. Password
    5. At least one schema available to use as a RED Data Warehouse Target
- Azure SQL Server software installed
    - Azure SQL Server ODBC driver (64-bit)
        1. Add ODBC DSN for Azure SQL Server Target Database
- SQL Server software installed
    - SQL Server ODBC driver (64-bit)
        1. Add ODBC DSN for SQL Server Target Database
- Python 3.8 or higher
    - Select "Add Python 3.8 to PATH" from installation Window
    - Pip Manager Install with command : *python -m pip install --upgrade pip*

Installation Through Setup Wizard

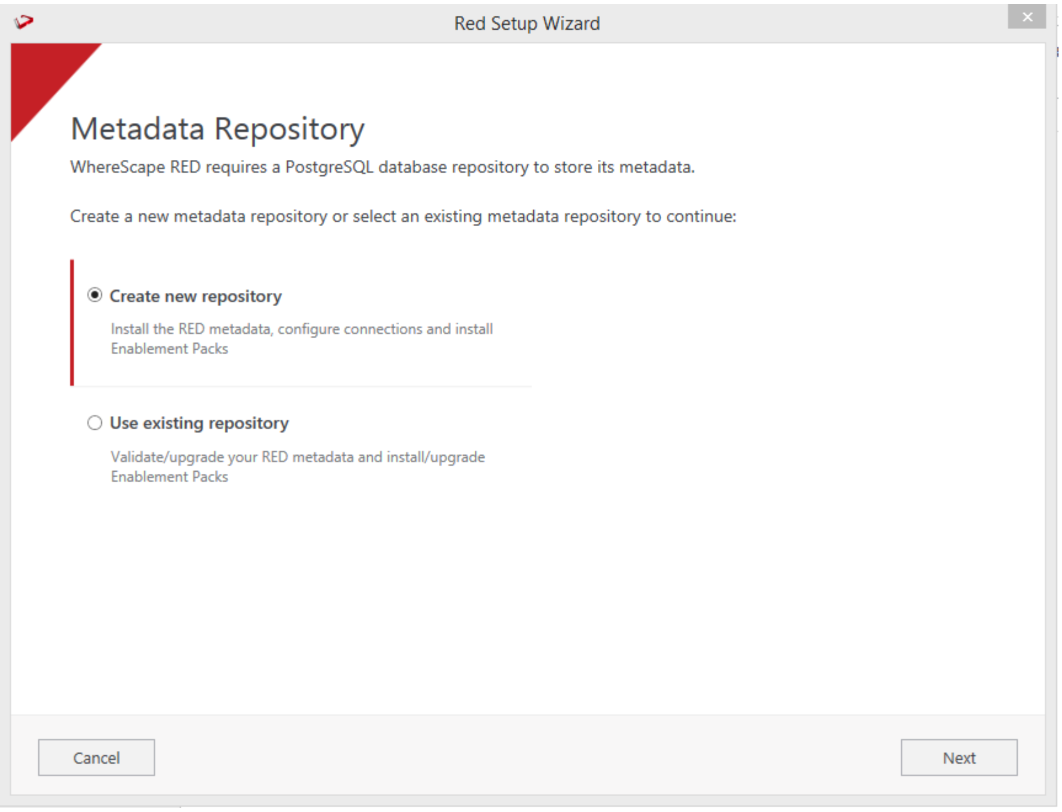
Run Setup Wizard as administrator



Metadata and Enablement Pack Installer

This installer will take you through the steps to configure RED and enable your source and target platforms. This installer can be used for creating a new RED metadata repository or to configure an existing metadata repository by installing or upgrading WhereScape Enablement Packs.

Create new repository or upgrade already existing repository.



Metadata Repository

WhereScape RED requires a PostgreSQL database repository to store its metadata.

Create a new metadata repository or select an existing metadata repository to continue:

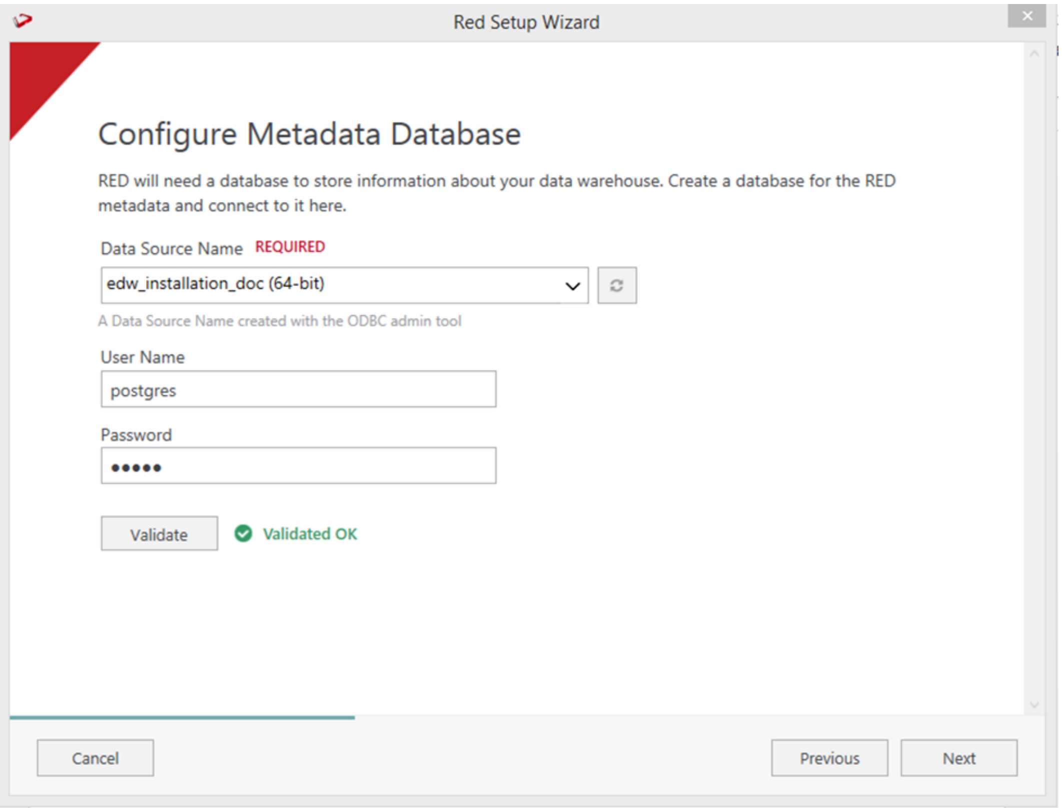
● **Create new repository**

Install the RED metadata, configure connections and install Enablement Packs

○ **Use existing repository**

Validate/upgrade your RED metadata and install/upgrade Enablement Packs

Select the created ODBC DSN, input login details and then select "Validate". Press Next



Configure Metadata Database

RED will need a database to store information about your data warehouse. Create a database for the RED metadata and connect to it here.

Data Source Name REQUIRED: edw_installation_doc (64-bit)

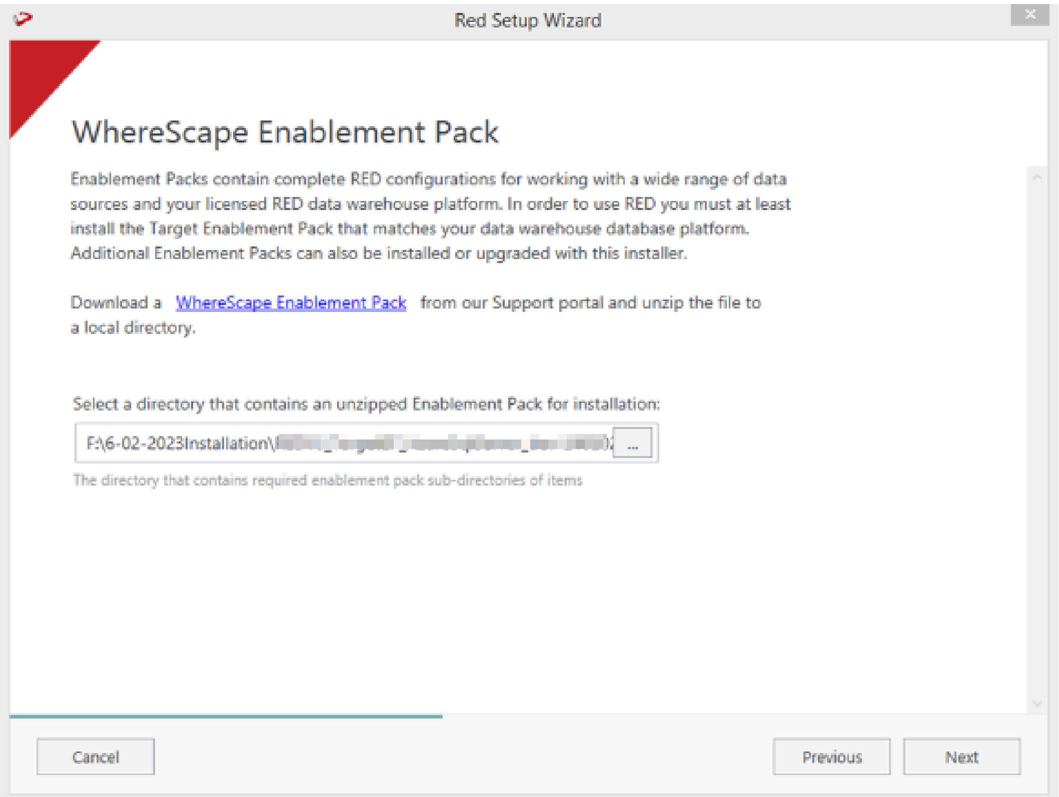
A Data Source Name created with the ODBC admin tool

User Name: postgres

Password: •••••

Validate ✓ Validated OK

Select the directory that contains unzipped Enablement Pack for installation. Press Next



WhereScape Enablement Pack

Enablement Packs contain complete RED configurations for working with a wide range of data sources and your licensed RED data warehouse platform. In order to use RED you must at least install the Target Enablement Pack that matches your data warehouse database platform. Additional Enablement Packs can also be installed or upgraded with this installer.

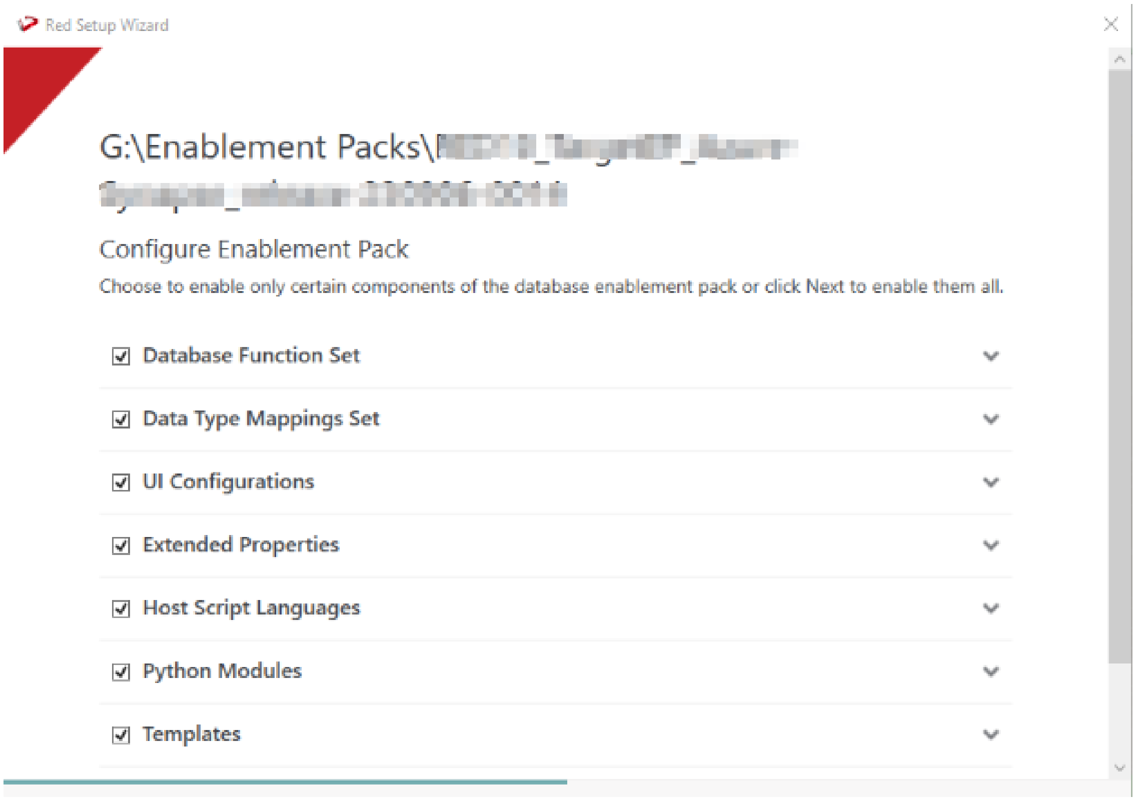
Download a WhereScape Enablement Pack from our Support portal and unzip the file to a local directory.

Select a directory that contains an unzipped Enablement Pack for installation:

F:\6-02-2023Installation\[redacted]

The directory that contains required enablement pack sub-directories of items

Using the check boxed list, include or exclude the components that are to be installed. Press Next.



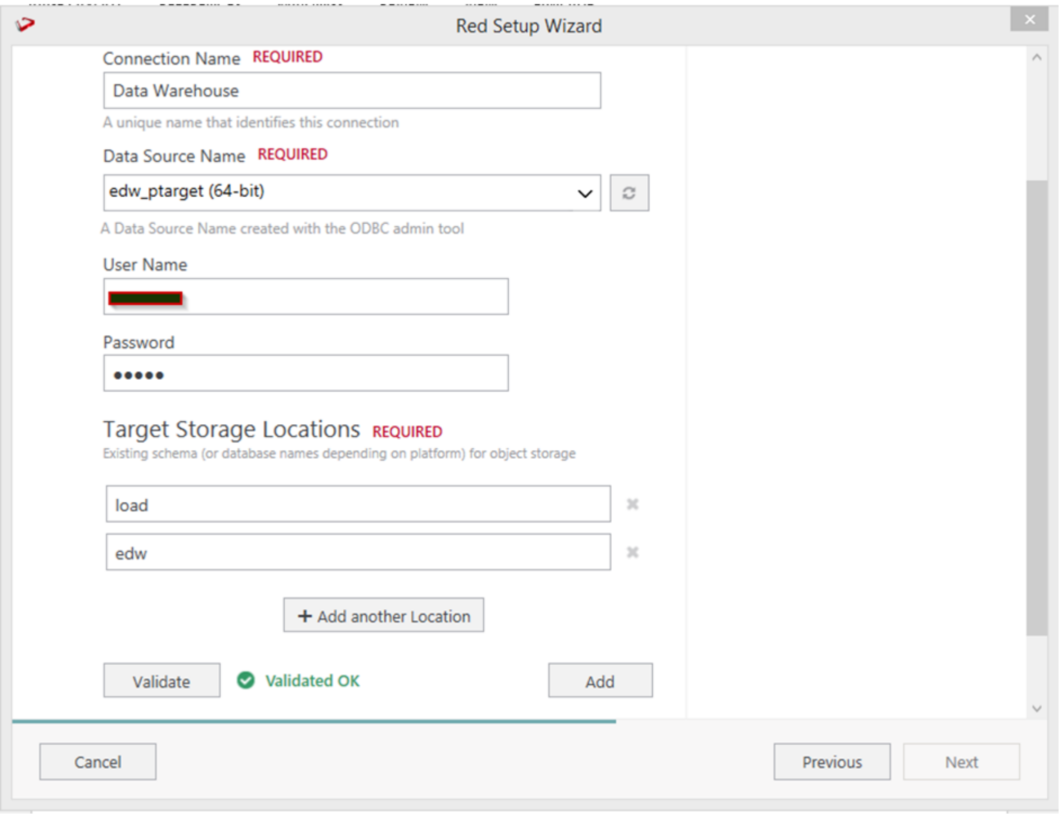
G:\Enablement Packs\[redacted]

Configure Enablement Pack

Choose to enable only certain components of the database enablement pack or click Next to enable them all.

- ☑ Database Function Set
- ☑ Data Type Mappings Set
- ☑ UI Configurations
- ☑ Extended Properties
- ☑ Host Script Languages
- ☑ Python Modules
- ☑ Templates

Configure a target connection (example, Data Warehouse) and its target locations. Validate and press ADD.


Red Setup Wizard

Connection Name REQUIRED
Data Warehouse
A unique name that identifies this connection

Data Source Name REQUIRED
edw_ptarget (64-bit)
A Data Source Name created with the ODBC admin tool

User Name
[redacted]

Password
●●●●●

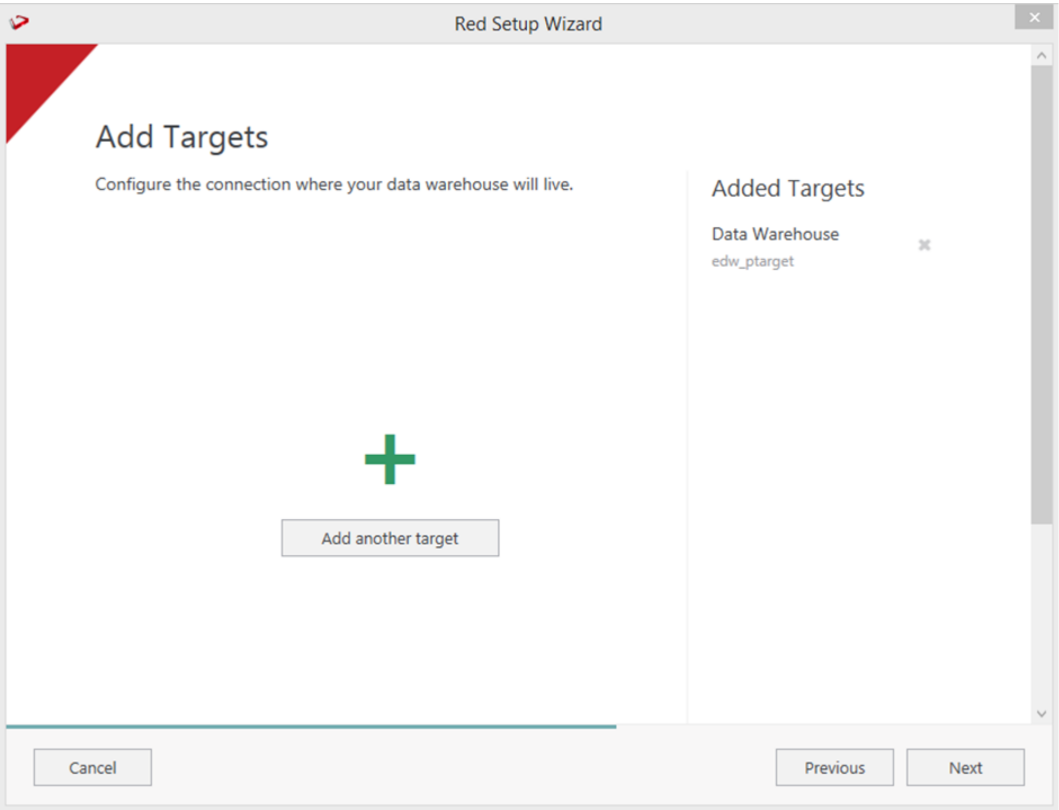
Target Storage Locations REQUIRED
Existing schema (or database names depending on platform) for object storage

load

edw

✓ Validated OK

When done, press ADD and then Press Next to advance.


Red Setup Wizard

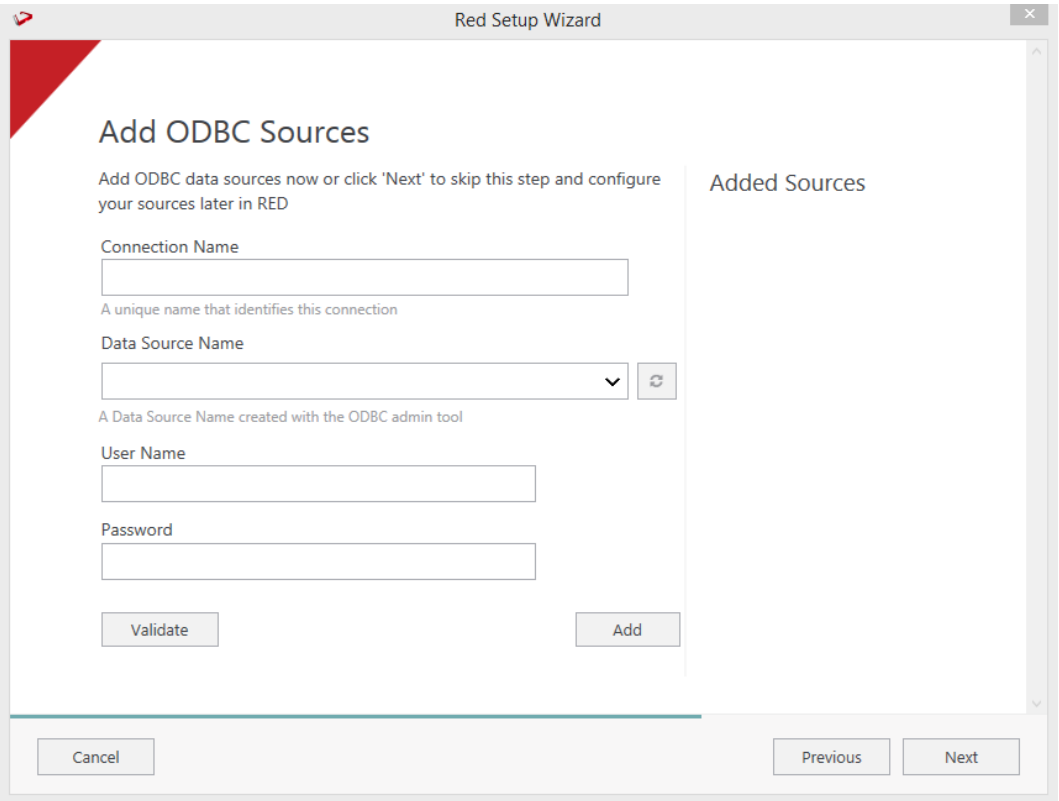
Add Targets

Configure the connection where your data warehouse will live.

Added Targets

Data Warehouse
edw_ptarget

Configure a data source connection (optional) and its target locations. Validate and press ADD. Press Next to advance.



Add ODBC Sources

Add ODBC data sources now or click 'Next' to skip this step and configure your sources later in RED

Connection Name

A unique name that identifies this connection

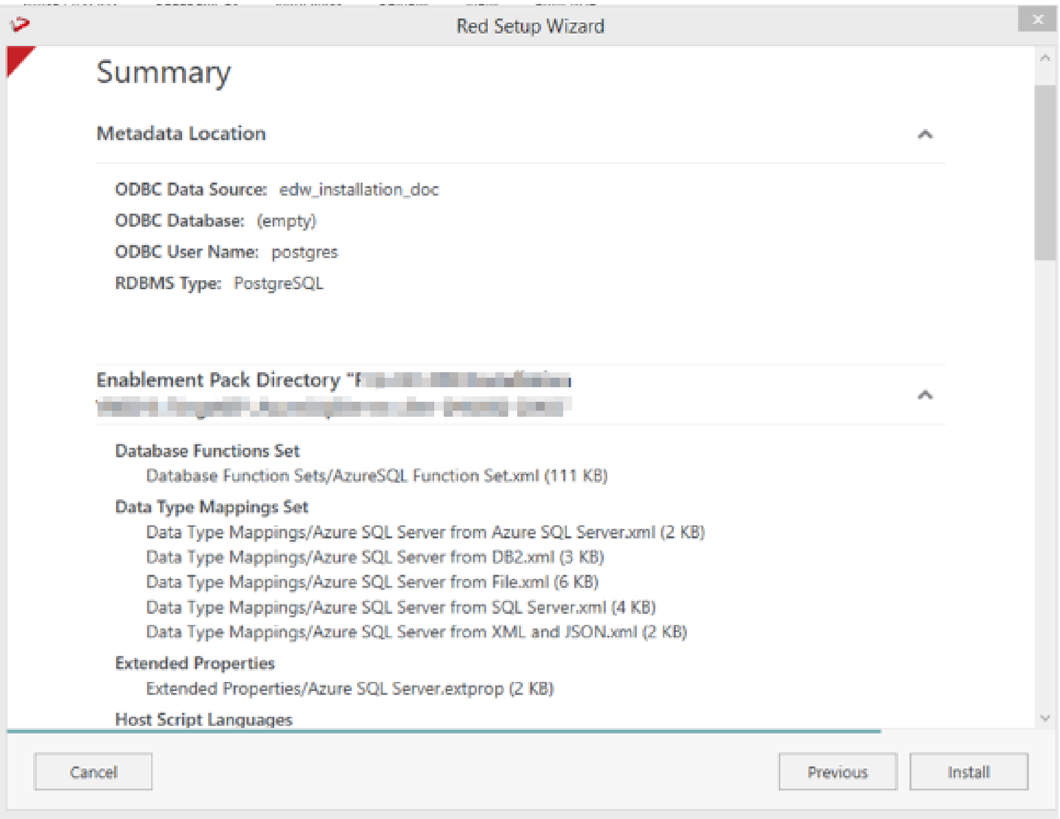
Data Source Name

A Data Source Name created with the ODBC admin tool

User Name

Password

Review the installation summary and click Install



Summary

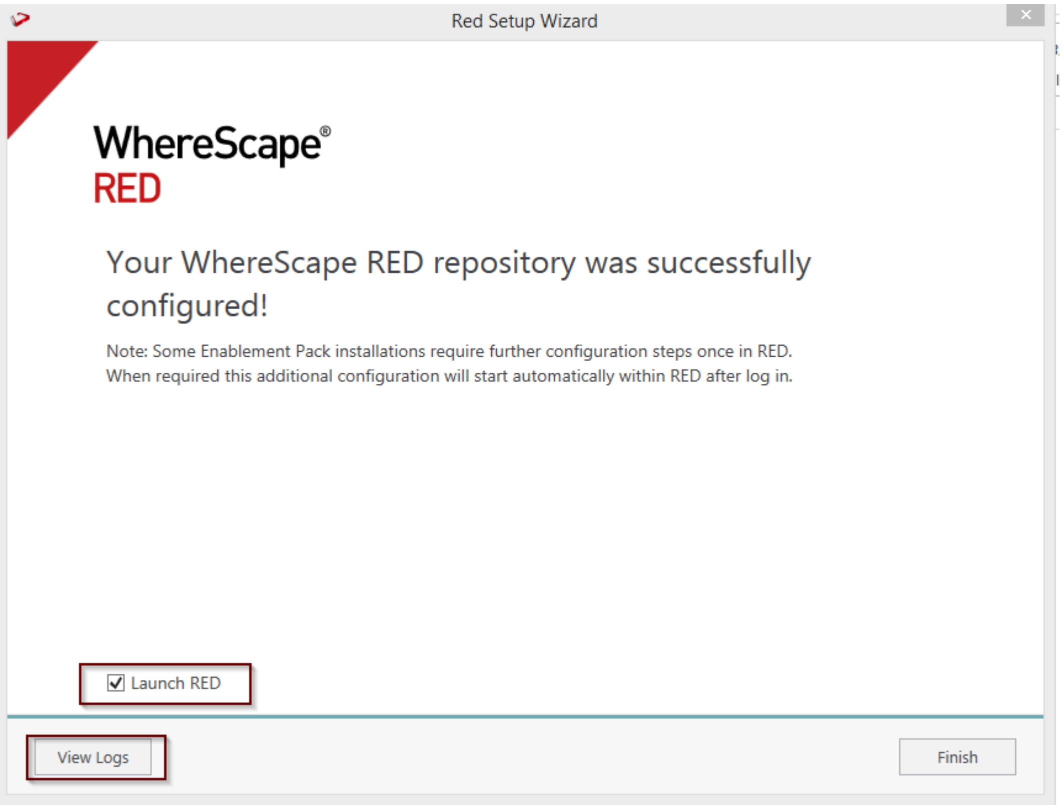
Metadata Location

ODBC Data Source: edw_installation_doc
ODBC Database: (empty)
ODBC User Name: postgres
RDBMS Type: PostgreSQL

Enablement Pack Directory "[redacted]

Database Functions Set
Database Function Sets/AzureSQL Function Set.xml (111 KB)
Data Type Mappings Set
Data Type Mappings/Azure SQL Server from Azure SQL Server.xml (2 KB)
Data Type Mappings/Azure SQL Server from DB2.xml (3 KB)
Data Type Mappings/Azure SQL Server from File.xml (6 KB)
Data Type Mappings/Azure SQL Server from SQL Server.xml (4 KB)
Data Type Mappings/Azure SQL Server from XML and JSON.xml (2 KB)
Extended Properties
Extended Properties/Azure SQL Server.extprop (2 KB)
Host Script Languages

Clicking on the View Logs will take to the installation log. Click on Finish once the installation is completed successfully.



Red Setup Wizard

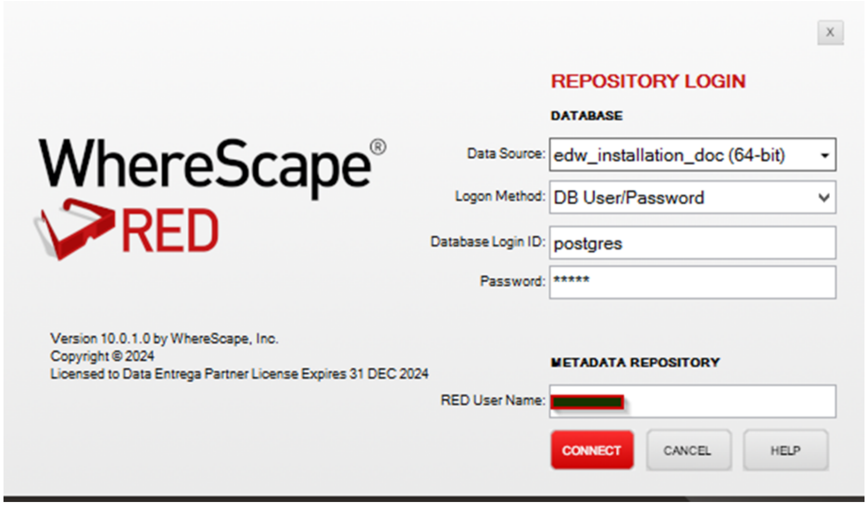
WhereScape® RED

Your WhereScape RED repository was successfully configured!

Note: Some Enablement Pack installations require further configuration steps once in RED. When required this additional configuration will start automatically within RED after log in.

☑ Launch RED

Login to WhereScape RED.



WhereScape® RED

Version 10.0.1.0 by WhereScape, Inc.
Copyright © 2024
Licensed to Data Entrega Partner License Expires 31 DEC 2024

REPOSITORY LOGIN

DATABASE

Data Source: edw_installation_doc (64-bit)

Logon Method: DB User/Password

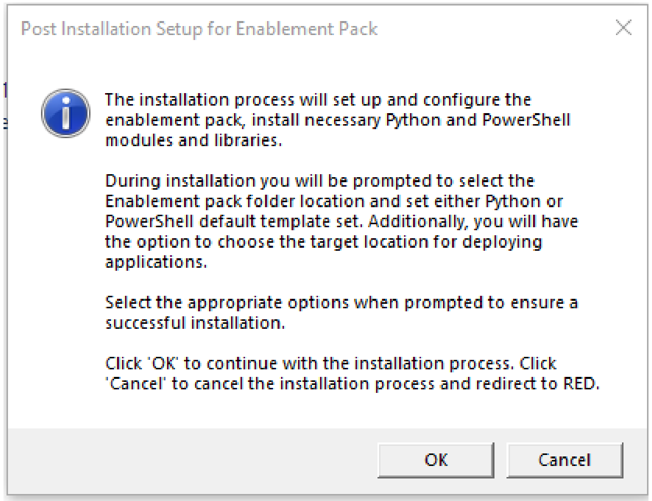
Database Login ID: postgres

Password: *****

METADATA REPOSITORY

RED User Name: [redacted]

Note: There is a post-install script that will run at the first login to RED10 to complete the post setup wizard installation process. You will be directed to below PowerShell window which will give brief explanation about post installation process.



Post Installation Setup for Enablement Pack

The installation process will set up and configure the enablement pack, install necessary Python and PowerShell modules and libraries.

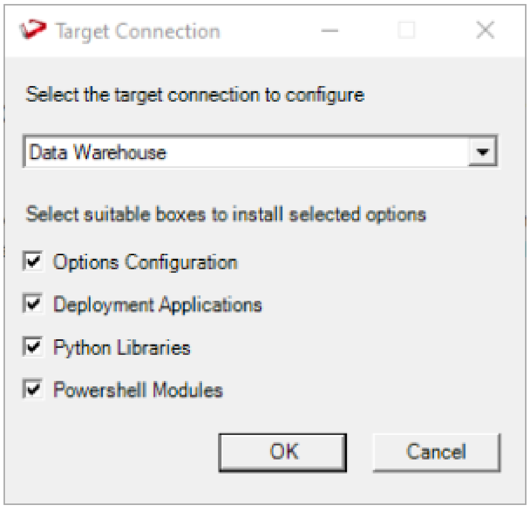
During installation you will be prompted to select the Enablement pack folder location and set either Python or PowerShell default template set. Additionally, you will have the option to choose the target location for deploying applications.

Select the appropriate options when prompted to ensure a successful installation.

Click 'OK' to continue with the installation process. Click 'Cancel' to cancel the installation process and redirect to RED.

Press OK to start the post installation. If pressed Cancel installation will stop and user will be directed to RED.

The user will be directed to the window below, where they have to select the target connection to be configured. Additionally, by deselecting the provided options, the user can choose not to install a particular option.



Target Connection

Select the target connection to configure

Data Warehouse

Select suitable boxes to install selected options

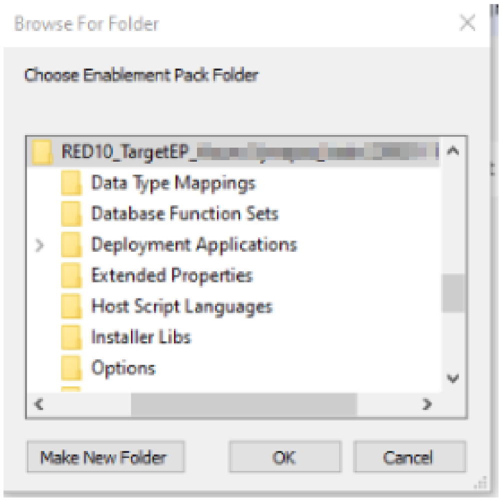
☑ Options Configuration

☑ Deployment Applications

☑ Python Libraries

☑ Powershell Modules

You will be directed to below PowerShell window. Provide the directory that contains unzipped Enablement Pack.



Browse For Folder

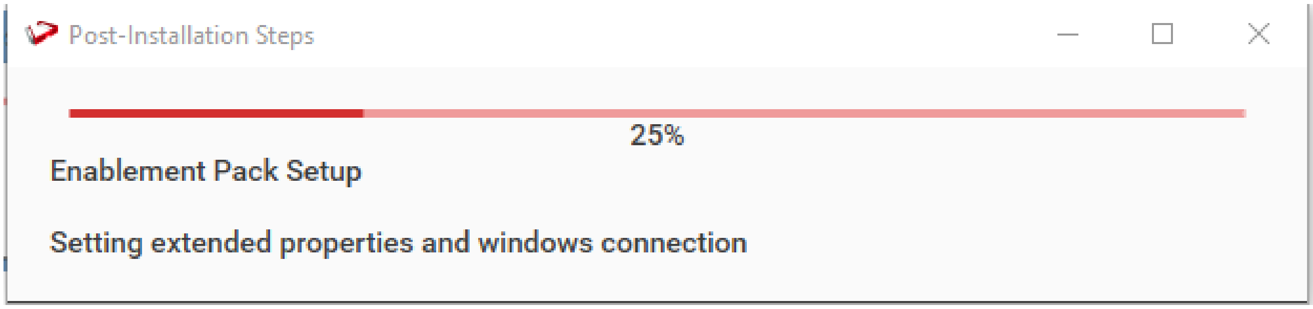
Choose Enablement Pack Folder

RED10_TargetEP_[redacted]

- Data Type Mappings
- Database Function Sets
- Deployment Applications
- Extended Properties
- Host Script Languages
- Installer Libs
- Options

Press OK

The progress bar will show the post installation progress.



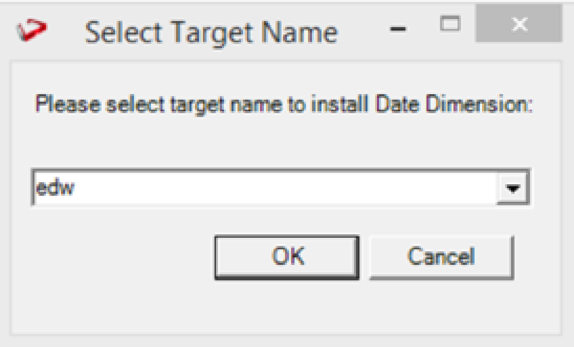
Post-Installation Steps

25%

Enablement Pack Setup

Setting extended properties and windows connection

User will have to choose the schema for the target setting that were provided. One pop up will come for setting default target schema for Date Dimension.

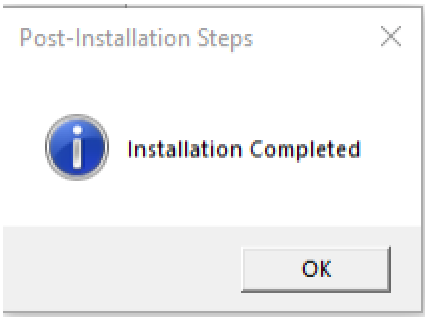


Select Target Name

Please select target name to install Date Dimension:

edw

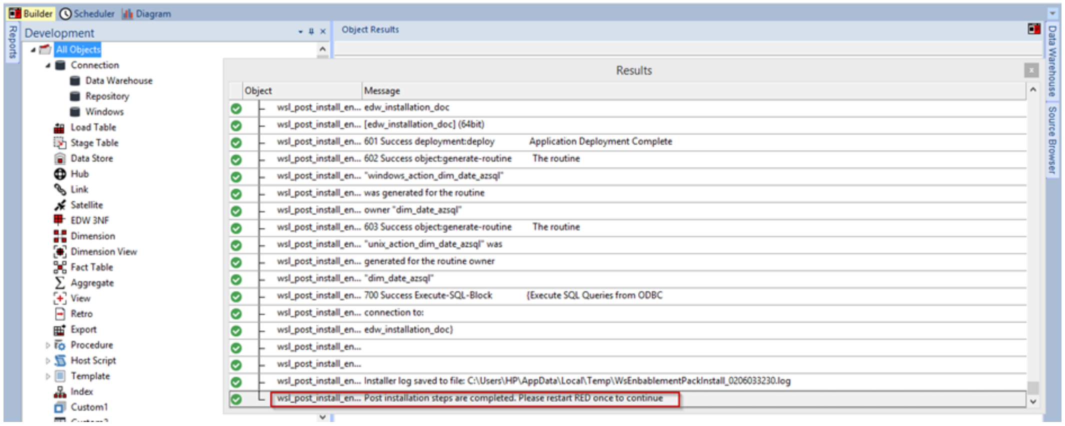
After selecting the target schema progress bar will show the progress for the installation and once it's completed, you will get the below pop up.



Post-Installation Steps

Installation Completed

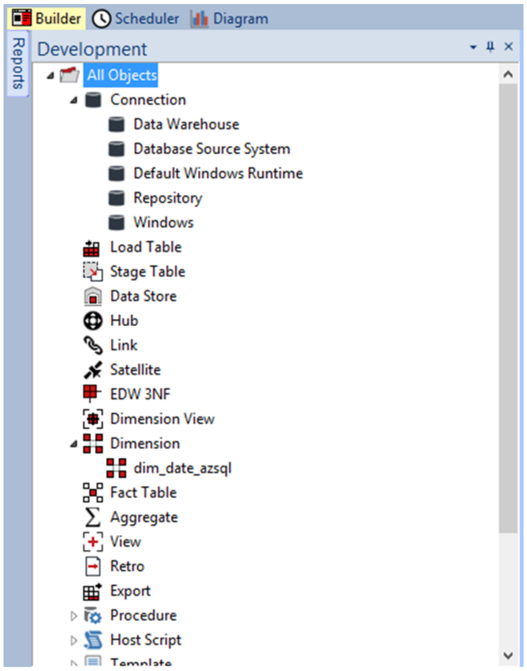
After pressing OK RED10 will open automatically.



Results

Object	Message
wsl_post_install_en...	edw_installation_doc
wsl_post_install_en...	[edw_installation_doc] (64bit)
wsl_post_install_en...	601 Success deployment:deploy Application Deployment Complete
wsl_post_install_en...	602 Success object:generate-routine The routine
wsl_post_install_en...	"windows_action_dim_date_azsql"
wsl_post_install_en...	was generated for the routine
wsl_post_install_en...	owner "dim_date_azsql"
wsl_post_install_en...	603 Success object:generate-routine The routine
wsl_post_install_en...	"unix_action_dim_date_azsql" was
wsl_post_install_en...	generated for the routine owner
wsl_post_install_en...	"dim_date_azsql"
wsl_post_install_en...	700 Success Execute-SQL-Block {Execute SQL Queries from ODBC
wsl_post_install_en...	connection to:
wsl_post_install_en...	edw_installation_doc}
wsl_post_install_en...	
wsl_post_install_en...	
wsl_post_install_en...	Installer log saved to file: C:\Users\HP\AppData\Local\Temp\WsEnbablementPackInstall_0206033230.log
wsl_post_install_en...	Post installation steps are completed. Please restart RED once to continue

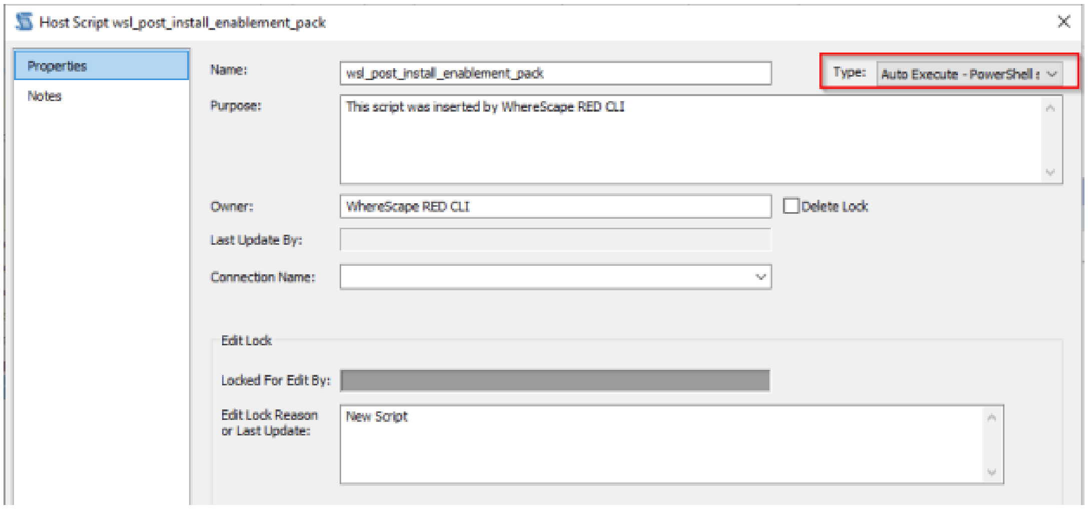
User will need to restart RED once.



Upgrade Of Existing Repository

For upgrade of existing repository

- From host script set script type of *wsl_post_install_enablement_pack* as Auto Execute - PowerShell Script



Name: wsl_post_install_enablement_pack

Type: Auto Execute - PowerShell s

Purpose: This script was inserted by WhereScape RED CLI

Owner: WhereScape RED CLI

□ Delete Lock

Last Update By:

Connection Name:

Locked For Edit By:

Edit Lock Reason or Last Update: New Script

Important Upgrade Notes

If RED upgrade the repository option is chosen.

This enablement pack will overwrite any existing Source Enablement Pack UI Configs:

Connection UI Config	Load UI Config
Amazon S3	Load From Amazon S3
Azure Data Lake Storage Gen2	Load From Azure Data Lake Storage Gen2
Google Cloud	Load From Google Cloud

To ensure existing Source Enablement Pack connections and associated Load Tables continue to browse and load:

Go into UI Configuration Maintenance in RED prior to installing this Enablement Pack and rename the affected UI Configurations. While the updated Load Template will work with previous Source Enablement Pack's we recommend moving these previous versions of Load Tables to newly created Parser based connections following this install. The earlier versions of the Source Enablement Pack will be deprecated following this release.

Post Install Steps – Optional

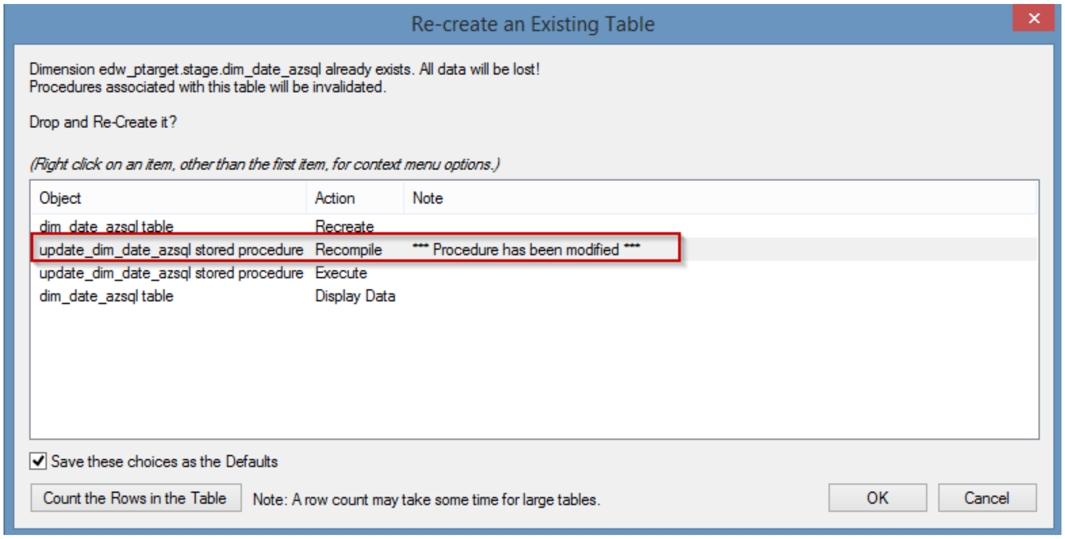
If you used the script Setup Wizard for installation then the following optional post install steps are available.

Configure Connections

There were two connections added that will optionally require your attention :

Below two are common for SQL Server and Azure SQL

1. Connection: 'Database Source System' - this connection was setup as an example source connection,
   - open it's properties and set it up for a source DB in your environment
   - or you can remove it if not required
2. Execute Update Procedure for dim_date_azsql to **Initialize Date Dimension for Target**



Re-create an Existing Table

Dimension edw_ptarget.stage.dim_date_azsql already exists. All data will be lost!
Procedures associated with this table will be invalidated.

Drop and Re-Create it?

(Right click on an item, other than the first item, for context menu options.)

Object	Action	Note
dim_date_azsql table	Recreate	
update_dim_date_azsql stored procedure	Recompile	*** Procedure has been modified ***
update_dim_date_azsql stored procedure	Execute	
dim_date_azsql table	Display Data	

☑ Save these choices as the Defaults

Note: A row count may take some time for large tables.

NOTE: Please avoid regenerating the procedure update_dim_date_azsql as it is a custom procedure.

For Azure SQL Target following extended properties will require your attention:

1. Target Connection Extended Property - Blob Storage Account
2. Target Connection Extended Property - Blob Storage Access Key
3. Target Connection Extended Property - Blob Storage Container
4. Target Connection Extended Property - Blob Data Source
5. Add exact record terminators for the files from windows source connection, please note \r\n and \n are considered as different record terminators.
   NOTE: For record terminator '\n' use hexadecimal notation`'0x0A'`

NOTE: For Blob extended properties refer to the below section

Configure Azure SQL Database and Blob Storage

Run below queries in azure sql database query editor.

```
CREATE MASTER KEY ENCRYPTION BY PASSWORD = '<MASTER_KEY>';

CREATE DATABASE SCOPED CREDENTIAL <BlobStorageCredential>  WITH IDENTITY = 'SHARED ACCESS SIGNATURE', SECRET = '<SAS_TOKEN>';

DROP EXTERNAL DATA SOURCE <Blob Data Source>;

CREATE EXTERNAL DATA SOURCE <Blob Data Source>WITH (TYPE = BLOB_STORAGE,LOCATION = 'https://<Blob Storage Account>.blob.core.windows.net/<Blob Storage Container>/<Directory>',CREDENTIAL = <BlobStorageCredential> );
```

NOTE: Keywords mentioned in <> in blue needs to be replaced with the values which can be found on azure

Source Enablement Pack Support

Source Pack Name	Supported By Azure SQL Server /SQL Server	Supported Features	Prerequisites
Cloud File Parser 1. Amazon S3 2. Azure Data Lake Storage Gen2 3. Google Cloud Storage	1. CSV 2. Excel 3. JSON 4. XML 5. AVRO 6. ORC 7. PARQUET	Bulk load	Refer to Windows Parser Guide
Windows Parser	1. CSV 2. Excel 3. JSON 4. XML 5. AVRO 6. ORC 7. PARQUET	Load Template, Source Properties will have option to select parser type to load the files.	Refer to Windows Parser Guide

Troubleshooting and Tips

Run As Administrator

Press the Windows Key on your keyboard and start typing cmd.exe, when the cmd.exe icon shows up in the search list right click it to bring up the context menu, select "Run As Administrator"

Now you have an admin prompt navigate to to the folder where you have unpacked your WhereScape Red Enablement Pack to using the 'cd' command:

```
C:\Windows\system32> cd <full path to the unpacked folder>
```

Run Powershell (.ps1) scripts from the administrator prompt by typing the Powershell run script command, for example:

```
C:\temp\EnablementPack>Powershell -ExecutionPolicy Bypass -File .\Setup_Enablement_Pack.ps1
```

Notes: In the event you can not bypass the Powershell execution policy due to group policies you can instead try "-ExecutionPolicy RemoteSigned" which should allow unsigned local scripts.

Windows Powershell Script Execution

On some systems Windows Powershell script execution is disabled by default. There are a number of workarounds for this which can be found by searching the term "Powershell Execution Policy".

Here is the most common workaround which WhereScape suggests, which does not permanently change the execution rights:

Start a Windows CMD prompt as Administrator, change directory to your script directory and run the WhereScape Powershell scripts with this command:

- cmd:>Powershell -ExecutionPolicy Bypass -File .\<script_file_name.ps1>

Restarting failed scripts

Some of the setup scripts will track each step and output the step number when there is a failure. To restart from the failed step (or to skip the step) provide the parameter "-startAtStep <step number>" to the script.

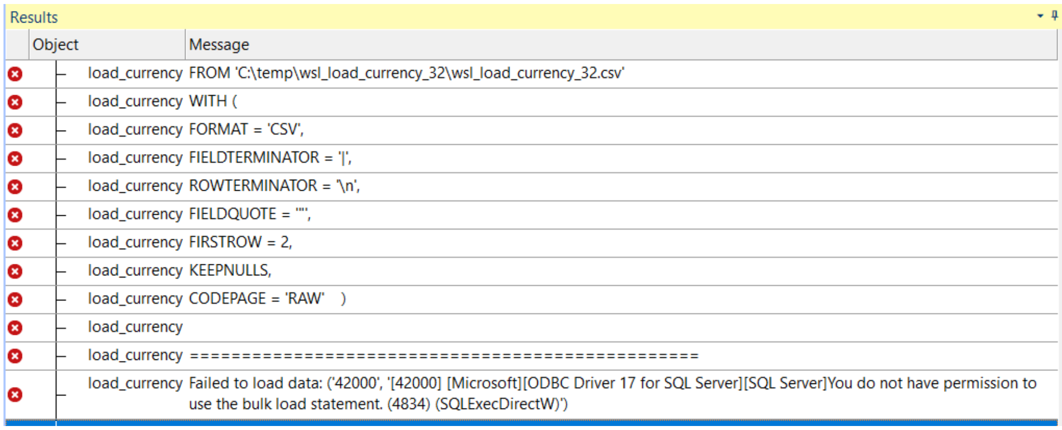
Example:

```
Powershell -ExecutionPolicy Bypass -File .\<script_file_name.ps1> -startAtStep 123
```

Tip: to avoid having to provide all the parameters again you can copy the full command line with parameters from the first "INFO" message from the beginning of the console output.

Bulk load error on SQL SERVER

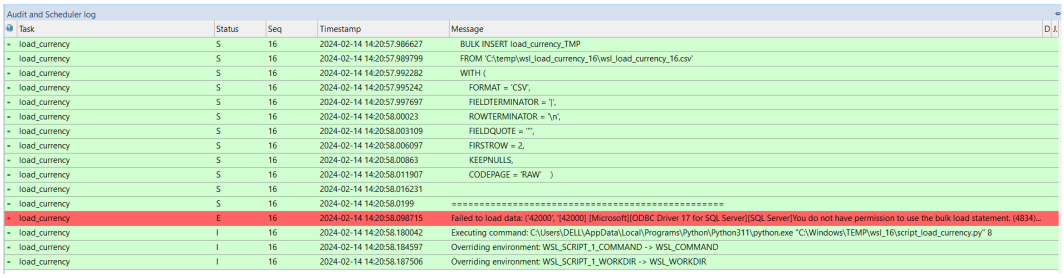
Interactive load



Results

	Object	Message
☒	load_currency	FROM 'C:\temp\wsl_load_currency_32\wsl_load_currency_32.csv'
☒	load_currency	WITH (
☒	load_currency	FORMAT = 'CSV',
☒	load_currency	FIELDTERMINATOR = '|',
☒	load_currency	ROWTERMINATOR = '\n',
☒	load_currency	FIELDQUOTE = '"',
☒	load_currency	FIRSTROW = 2,
☒	load_currency	KEEPNULLS,
☒	load_currency	CODEPAGE = 'RAW')
☒	load_currency	
☒	load_currency	==
☒	load_currency	Failed to load data: ('42000', '[42000] [Microsoft][ODBC Driver 17 for SQL Server][SQL Server]You do not have permission to use the bulk load statement. (4834) (SQLExecDirectW)')

Scheduler load

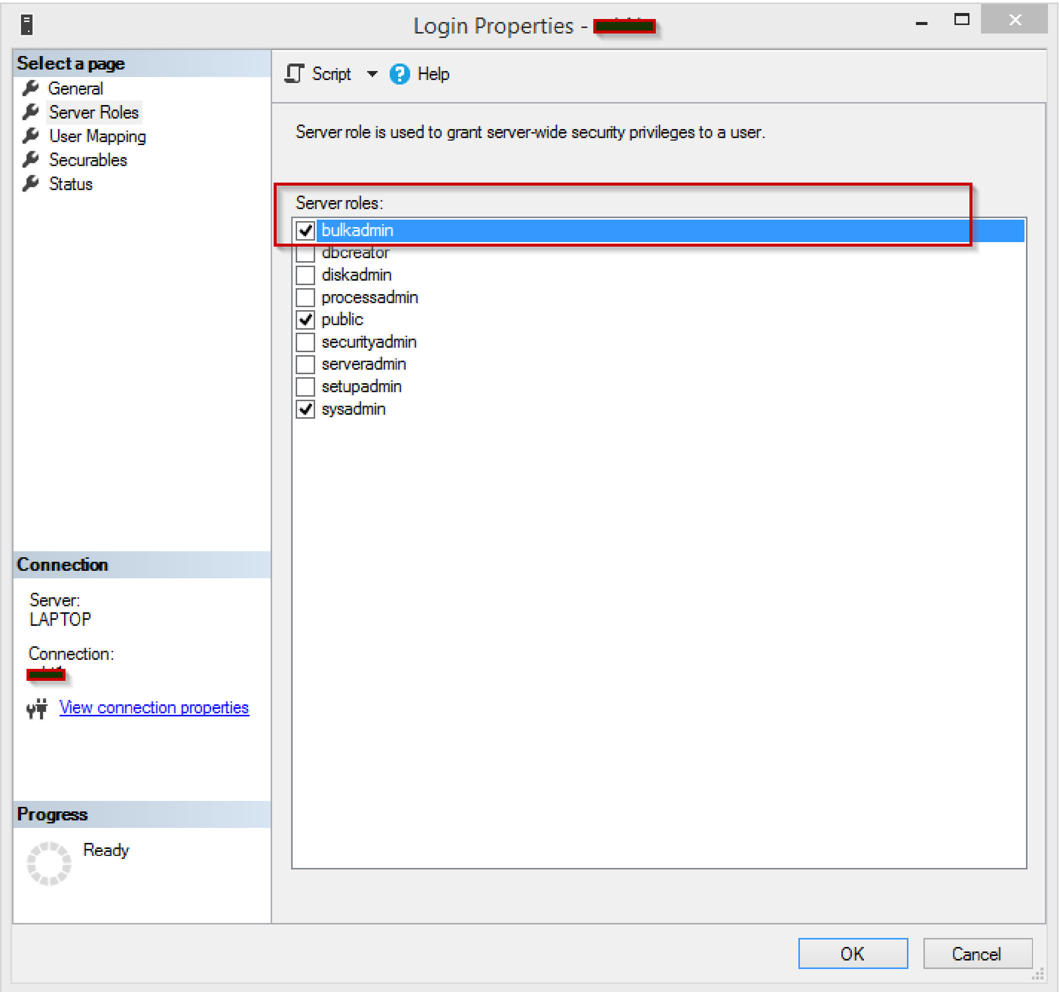


Audit and Scheduler log

Task	Status	Seq	Timestamp	Message	D	J.
load_currency	S	16	2024-02-14 14:20:57.986627	BULK INSERT load_currency_TMP		
load_currency	S	16	2024-02-14 14:20:57.989799	FROM 'C:\temp\wsl_load_currency_16\wsl_load_currency_16.csv'		
load_currency	S	16	2024-02-14 14:20:57.992282	WITH (		
load_currency	S	16	2024-02-14 14:20:57.995242	FORMAT = 'CSV',		
load_currency	S	16	2024-02-14 14:20:57.997697	FIELDTERMINATOR = '|',		
load_currency	S	16	2024-02-14 14:20:58.00023	ROWTERMINATOR = '\n',		
load_currency	S	16	2024-02-14 14:20:58.003109	FIELDQUOTE = '"',		
load_currency	S	16	2024-02-14 14:20:58.006097	FIRSTROW = 2,		
load_currency	S	16	2024-02-14 14:20:58.00863	KEEPNULLS,		
load_currency	S	16	2024-02-14 14:20:58.011907	CODEPAGE = 'RAW')		
load_currency	S	16	2024-02-14 14:20:58.016231			
load_currency	S	16	2024-02-14 14:20:58.0199	==		
load_currency	E	16	2024-02-14 14:20:58.098715	Failed to load data: ('42000', '[42000] [Microsoft][ODBC Driver 17 for SQL Server][SQL Server]You do not have permission to use the bulk load statement. (4834)...		
load_currency	I	16	2024-02-14 14:20:58.180042	Executing command: C:\Users\DELL\AppData\Local\Programs\Python\Python311\python.exe "C:\Windows\TEMP\wsl_16\script_load_currency.py" 8		
load_currency	I	16	2024-02-14 14:20:58.184597	Overriding environment: WSL_SCRIPT_1_COMMAND -> WSL_COMMAND		
load_currency	I	16	2024-02-14 14:20:58.187506	Overriding environment: WSL_SCRIPT_1_WORKDIR -> WSL_WORKDIR		

For bulk load error, the 'bulkadmin' permission is required.

SQL SERVER -> SECURITY -> LOGINS -> USER -> PROPERTIES ->SERVER ROLES



Server role is used to grant server-wide security privileges to a user.

Server roles:

☑ bulkadmin
□ dbcreator
□ diskadmin
□ processadmin
☑ public
□ securityadmin
□ serveradmin
□ setupadmin
☑ sysadmin

Server:
LAPTOP

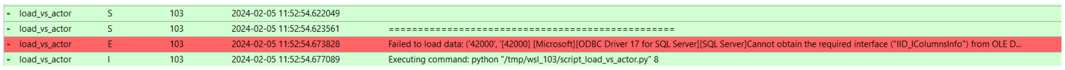
Connection:
[redacted]

For bulk load error Operating system error code 5(Access is denied)

```
Failed to load data: ('42000', '[42000] [Microsoft][ODBC Driver 17 for SQL Server][SQL Server]Cannot bulk load
because the file "C:\WhereScape\Temp\...\xxxx.csv" could not be opened. Operating system error code 5(Access
is denied.). (4861) (SQLExecDirectW)')
```

1. Grant SQL Server access to the bulk load folder by adjusting folder permissions. Navigate to the folder, right-click, select "Properties," go to the "Security" tab, click "Edit," then "Add." In the new window, select "Advanced," click "Find Now," and locate the appropriate SQL Server user (e.g., SQLServerMSSQLUser$UserName$SQLExpress). Confirm by clicking "OK" through all opened dialogs.
2. If the first fix doesn't resolve the issue, verify if permissions are applied to subfolders as well. Navigate to the advanced settings, and check the 'Inheritance' status. If it's 'Disabled,' enable it by using the "Allow Inheritance" button.

Row Terminator For Linux



```
-  load_vs_actor  S  103  2024-02-05 11:52:54.622049
-  load_vs_actor  S  103  2024-02-05 11:52:54.623561  ==================================================
-  load_vs_actor  E  103  2024-02-05 11:52:54.673828  Failed to load data: ('42000', '[42000] [Microsoft][ODBC Driver 17 for SQL Server][SQL Server]Cannot obtain the required interface ("IID_IColumnsInfo") from OLE D...
-  load_vs_actor  I  103  2024-02-05 11:52:54.677089  Executing command: python "/tmp/wsl_103/script_load_vs_actor.py" 8
```

Error Message:

```
Failed to load data: ('42000', '[42000] [Microsoft][ODBC Driver 17 for SQL Server][SQL Server]Cannot obtain
the required interface ("IID_IColumnsInfo") from OLE DB provider "BULK" for linked server "(null)". (7301)
(SQLExecDirectW)')
```

When you specify '**\n**' as a row terminator for bulk import in *UNLOAD_RECORD_CHAR* extended property or source properties, or implicitly use the default row terminator, BULK INSERT statement expects a carriage return-line feed combination (CRLF) as the row terminator. If your source file uses a line feed character only (LF) as the row terminator - as is typical in files generated on Unix and Linux computers - use hexadecimal notation to specify the LF row terminator. For example use '**0x0A**' in *UNLOAD_RECORD_CHAR* extended property or source properties for Unix and Linux machines.

Note: The hexadecimal notation is not required for Parquet file for Unix and Linux machines

Python requirements for offline install

Additionally to the base Python installation being required, the WhereScape Python Template set also requires certain additional Python libraries. The install scripts uses the PIP (package manager) to download these libraries, however for offline installs you will need to install the required libraries yourself.

Required Python libraries/add-ons:

- pywin32-ctypes
- python-tds
- pywin32
- glob2
- gzip-reader
- regex
- pyodbc

If a valid RED installation can not be found

If you have RED 10.x or higher installed but the script (Setup_Enablement_Pack.ps1) fails to find it on you system then you are most likely running PowerShell (x86) version which does not show installed 64 bit apps by default. Please open a 64 bit version of Powershell instead and re-run the script.