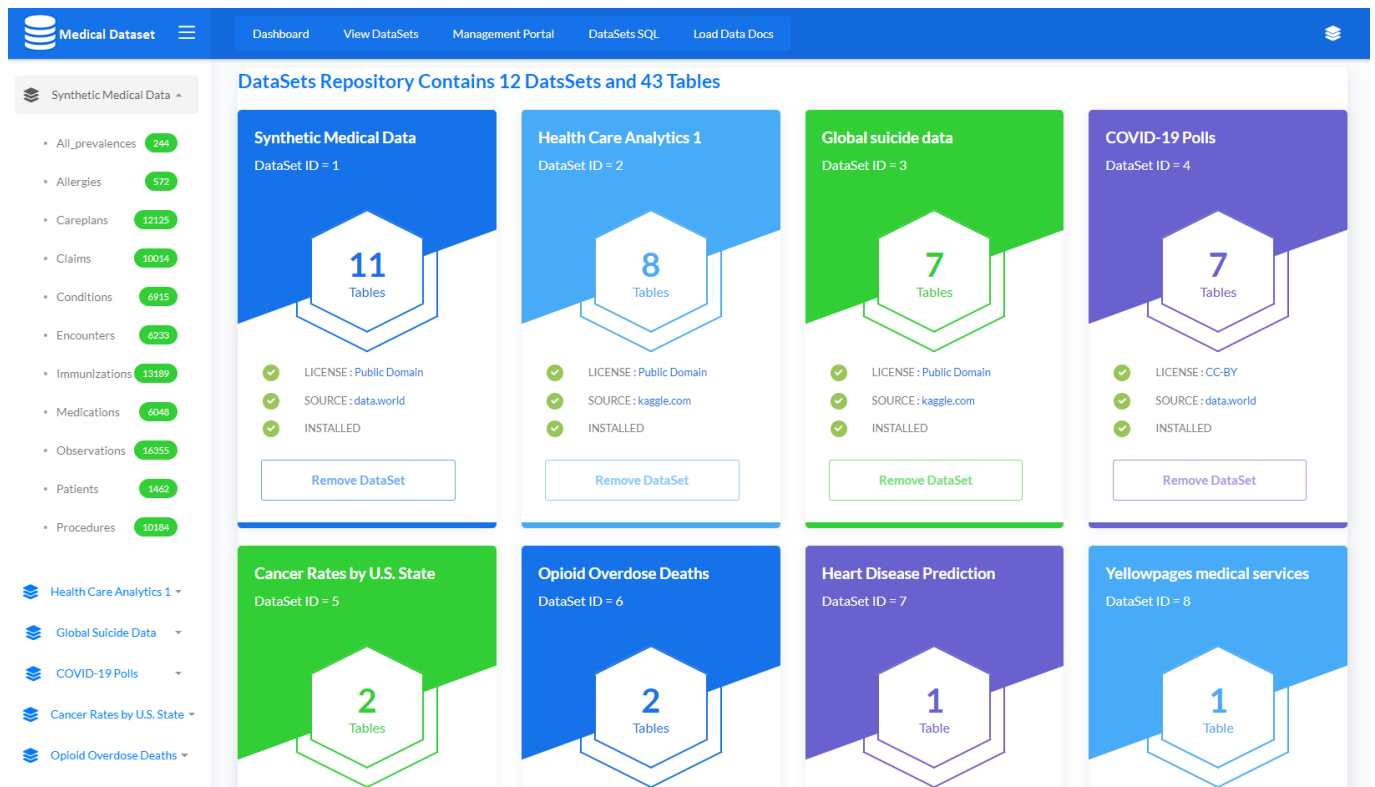


Article

[Muhammad Waseem](#) · Jan 10, 2022 4m read

[Open Exchange](#)

12 Medical Datasets along with 43 tables dynamically created by using all-new LOAD DATA SQL functionality



Application to import 12 Datasets along with 43 tables dynamically by using LOAD DATA command which loads data from a source into an IRIS SQL table.

List of Datasets

- 1 - [Synthetic Medical Data](#) : Synthetic medical record data for Introduction to Biomedical Data Science. [LICENCE:Public Domain](#)
- 2 - [Health Care Analytics - 1](#) : Part of Janatahack Hackathon in Analytics Vidhya [LICENCE:Public Domain](#)
- 3 - [Global Suicide Data](#) : World Health Organization (WHO) dataset [LICENCE:Public Domain](#)
- 4 - [COVID-19 Polls](#) : Survey data to understand how Americans view the coronavirus pandemic and the president's response [LICENCE:CC-BY](#)
- 5 - [Cancer Rates by U.S. State](#) : The rates are the numbers out of 100,000 people who developed or died from cancer each year. [LICENCE:Public Domain](#)
- 6 - [Opioid Overdose Deaths](#) : Opioid addiction and death rates in the U.S. and abroad have reached "epidemic" levels. The CDC's data reflects the incredible spike in overdoses caused by drugs containing opioids. [LICENCE:Public Domain](#)
- 7 - [Heart Disease Prediction](#) : This data set came from the University of California Irvine data repository and is used to predict heart disease [LICENCE:Public Domain](#)
- 8 - [Yellowpages medical services dataset](#) : Yellowpages dataset with more than 400K+ records from Crawl Feeds. [LICENCE:Public Domain](#)
- 9 - [Hospital ratings](#) : This are the official datasets used on the Medicare.gov Hospital Compare Website

provided by the Centers for Medicare & Medicaid Services. These data allow you to compare the quality of care at over 4,000 Medicare-certified hospitals across the country. [LICENCE:Public Domain](#)

- [10 -Heart Failure Prediction](#) : 12 clinical features for predicting death events. [LICENCE:Attribution 4.0 International \(CC BY 4.0\)](#)
- [11 -Diabetes Dataset](#) : This dataset is originally from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective is to predict based on diagnostic measurements whether a patient has diabetes. [LICENCE:Public Domain](#)
- [12 -Chronic Kidney Disease](#) : This dataset is originally from UCI Machine Learning Repository. The objective of the dataset is to diagnostically predict whether a patient is having chronic kidney disease or not, based on certain diagnostic measurements included in the dataset. [LICENCE:Public Domain](#)

Features

- Dynamically creation of table based on the CSV file data.
- Import dataset by using LOAD DATA functionality by Terminal or by web application.
- Remove dataset programmatically by terminal or by Web Application.
- Dynamically view Imported Data from Web Application.
- Functionality to Import data into CSV, Excel or PDF Format.
- Responsive web application to view the status of Dataset.
- Any or All datasets can be install or Remove with single command.
- Use of %SQLDiag.Result and %SQLDiag.Message tables with LOAD DATA

It is recommended to read related documentations [LOAD DATA \(SQL\)](#).

How to install or remove any Dataset from Terminal

Use below command to import particular Dataset by passing it's ID or pass 999 to import all DataSets

```
do ##class(dc.data.medical.utility).ImportDS(1)
```

Below is the main script to create table dynamically and load data by using LOAD DATA functionality. Please note that table is created dynamically

```
//Get file name
SET filename=tRS.Get("Name")
//Remove .csv from the file name
SET tableName=$REPLACE("dc_data_"_ds_"_"_tRS.Get("ItemName"),".csv","")
//Get columns based on the header row of csv file
Do ##class(dc.data.medical.utility).GetColTypes(filename,.coltype)
//Dynamically create table based on tablename and column types
SET qry = "CREATE TABLE "_tableName_" ("_coltype_)"
SET rset = ##class(%SQL.Statement).%ExecDirect(,qry)
//Check if table created successfully
IF rset.%SQLCODE
{
    WRITE "ERROR : ",rset.%Message,!
}
ELSE
{
    //Dynamically construct LOAD DATA statement
```

```
SET qry = "LOAD DATA FROM FILE &nbsp;'_'_filename_' INTO '_tableName_ ' '_USING { "
"from":{ "file":{ "header": "1" } } }"
SET rset = ##class(%SQL.Statement).%ExecDirect(,qry)
// Check result set sqlcode, In case of error write resultset message
IF rset.%SQLCODE
{
&nbsp;WRITE "ERROR Table : " _tableName_ " IMPORT FAILED: ",rset.%Message,!
}
ELSE
{
&nbsp;WRITE "SUCCESS table : " _tableName_ " created and "_rset.%ROWCOUNT_" Rows Impo
rted Successfully",!
}
}
```

Use below command to remove particular Dataset by passing it's ID or pass 999 to remove all DataSets

```
do ##class(dc.data.medical.utility).RemoveDS(1)
```

Below is the main script to remove table dynamically

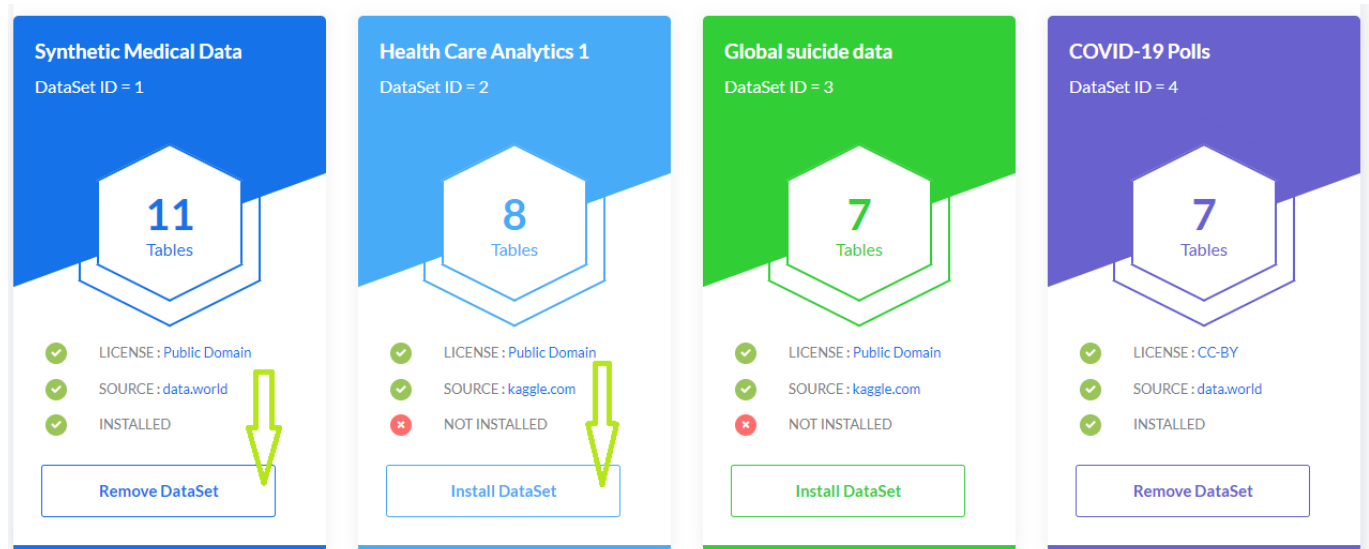
```
//Get file name
SET filename=tRS.Get("Name")
//Remove .csv from file name
SET tableName=$REPLACE("dc_data_"_ds_"_"_tRS.Get("ItemName"),".csv","")
//Drop table
SET qry = "DROP TABLE "_tableName
SET rset = ##class(%SQL.Statement).%ExecDirect(,qry)
//Check if table deleted successfully
IF rset.%SQLCODE
{
WRITE "ERROR : ",rset.%Message,!
}
ELSE
{
WRITE "Table "_tableName_ " deleted successfully",!
}
}
```

How to install or remove any Dataset from Web Application

Navigate to <http://localhost:52773/csp/datasets/index.csp> application dashboard

If Dataset is not installed then install DataSet button will be visible and if Dataset is installed then Remove Dataset button will be visible.

Just press the desire button to install or remove any data set



How to view and export data from Dataset

Navigate to <http://localhost:52773/csp/datasets/datasets.csp> View DataSets page.

Select particular dataset and then table from the list. Press Excel, CSV, PDF file button to export the data.

The screenshot shows the 'Medical Dataset' web application interface. The 'View DataSets' tab is active. On the left, a list of datasets is shown, with 'All_prevalences' selected under 'Synthetic Medical Data'. The main area displays the 'DATASET (Synthetic Medical Data) TABLE NAME (dc_data_synthetic.all_prevalences)' table. The table has columns: item, population_type, occurrences, population_count, prevalence_rate, and prevalence_percentage. The 'Excel' button is highlighted with a green arrow.

item	population_type	occurrences	population_count	prevalence_rate	prevalence_percentage
0.3 MI Epinephrine 0.5 Mg/MI Auto Injector	LIVING	92	1000	.092	9.2
0.65 MI Depo Subq Provera 160 Mg/MI Prefilled Syringe	LIVING	18	1000	.018	1.8
1 MI Depo Provera 150 Mg/MI Injection	LIVING	135	1000	.135	13.5
12 Hr Cefaclor 500 Mg Extended Release Oral Tablet	LIVING	15	1000	.015	1.5
120 Actuat Fluticasone Propionate 0.044 Mg/Actuat Metered Dose Inhaler	LIVING	42	1000	.042	4.2
200 Actuat Albuterol 0.09 Mg/Actuat Metered Dose Inhaler	LIVING	42	1000	.042	4.2
24 Hr Metformin Hydrochloride 500 Mg Extended Release Oral Tablet	LIVING	44	1000	.044	4.4
3 MI Liraglutide 6 Mg/MI Pen Injector	LIVING	15	1000	.015	1.5
Acetaminophen 160 Mg	LIVING	689	1000	.689	68.9
Acetaminophen 160 Mg Oral Tablet	LIVING	536	1000	.536	53.6

Thanks

[#Contest](#) [#Databases](#) [#InterSystems Package Manager \(IPM\)](#) [#SQL](#) [#InterSystems IRIS](#)

12 Medical Datasets along with 43 tables dynamically created by using all-new LOAD DATA SQL functionality
Published on InterSystems Developer Community (<https://community.intersystems.com>)

[Check the related application on InterSystems Open Exchange](#)

Source

URL:<https://community.intersystems.com/post/12-medical-datasets-along-43-tables-dynamically-created-using-all-new-load-data-sql>