Data Analysis Guide

Version 2.7
2024-07-12

InterSystems Corporation 1 Memorial Drive Cambridge MA 02142 www.intersystems.com

Data Analysis Guide
InterSystems Version 2.7 2024-07-12
Copyright © 2024 InterSystems Corporation
All rights reserved.

InterSystems®, HealthShare Care Community®, HealthShare Unified Care Record®, IntegratedML®, InterSystems Caché®, InterSystems Ensemble® InterSystems HealthShare®, InterSystems IRIS®, and TrakCare are registered trademarks of InterSystems Corporation. HealthShare® CMS Solution Pack™ HealthShare® Health Connect Cloud™, InterSystems IRIS for Health™, InterSystems Supply Chain Orchestrator™, and InterSystems TotalView™ For Asset Management are trademarks of InterSystems Corporation. TrakCare is a registered trademark in Australia and the European Union.

All other brand or product names used herein are trademarks or registered trademarks of their respective companies or organizations.

This document contains trade secret and confidential information which is the property of InterSystems Corporation, One Memorial Drive, Cambridge, MA 02142, or its affiliates, and is furnished for the sole purpose of the operation and maintenance of the products of InterSystems Corporation. No part of this publication is to be used for any other purpose, and this publication is not to be reproduced, copied, disclosed, transmitted, stored in a retrieval system or translated into any human or computer language, in any form, by any means, in whole or in part, without the express prior written consent of InterSystems Corporation.

The copying, use and disposition of this document and the software programs described herein is prohibited except to the limited extent set forth in the standard software license agreement(s) of InterSystems Corporation covering such programs and related documentation. InterSystems Corporation makes no representations and warranties concerning such software programs other than those set forth in such standard software license agreement(s). In addition, the liability of InterSystems Corporation for any losses or damages relating to or arising out of the use of such software programs is limited in the manner set forth in such standard software license agreement(s).

THE FOREGOING IS A GENERAL SUMMARY OF THE RESTRICTIONS AND LIMITATIONS IMPOSED BY INTERSYSTEMS CORPORATION ON THE USE OF, AND LIABILITY ARISING FROM, ITS COMPUTER SOFTWARE. FOR COMPLETE INFORMATION REFERENCE SHOULD BE MADE TO THE STANDARD SOFTWARE LICENSE AGREEMENT(S) OF INTERSYSTEMS CORPORATION, COPIES OF WHICH WILL BE MADE AVAILABLE UPON REQUEST.

InterSystems Corporation disclaims responsibility for errors which may appear in this document, and it reserves the right, in its sole discretion and without notice, to make substitutions and modifications in the products and practices described in this document.

For Support questions about any InterSystems products, contact:

InterSystems Worldwide Response Center (WRC)
Tel: +1-617-621-0700
Tel: +44 (0) 844 854 2917
Email: support@InterSystems.com

Table of Contents

1 Welcome, Data Analysts 1
1.1 Use Modes 1
1.2 See Also 1

2 Using the SQL Explorer 3
2.1 Introduction 3
2.2 Viewing the Contents of a Table 3
2.3 Viewing Fields in a Table 4
2.4 Running a Custom Query 4
2.5 Show History 4
2.6 See Also 5

3 Defining Cubes 7
3.1 Outline of Steps 7
3.2 See Also 7

4 Viewing Cube Build Information 9
4.1 Viewing Cube Build Information 9
4.2 Filtering the Cube Build Information 9
4.3 See Also 10

5 Accessing Your Data via JDBC 11
5.1 Outline of Steps 11
5.2 See Also 11

6 Accessing Your Data via ODBC 13
6.1 Outline of Steps 13
6.2 See Also 13

1

Welcome, Data Analysts

InterSystems® Data Fabric Studio™ is designed to bring together disparate data sources and transform them into a single source of truth—actionable, decision-ready data. Data engineers set up the automation to load data, and then you can take advantage of the data in any way that you choose.

Data Fabric Studio was designed for a specific set of overlapping use cases, and you can tailor your solution to include any parts of these cases that apply to you.

1.1 Use Modes

Data Fabric Studio provides a built-in analytics option, which enables you to build cubes that you can then use to create dashboards. You can also connect external analytics or reporting tools to these cubes.

To complement that option, it is also possible to connect directly to the tables in Data Fabric Studio via JDBC or ODBC.

In either case, it is generally necessary to be able to easily view the tables within Data Fabric Studio before proceeding. For this task, you can use the built-in SQL Explorer feature.

1.2 See Also

- Using the SQL Explorer
- Defining Cubes
- Accessing Your Data via JDBC
- Accessing Your Data via ODBC
- Data Engineering Guide
- About Your Solution: What Is Not Documented

2

Using the SQL Explorer

To help you make use of the tables in the system, InterSystems® Data Fabric Studio™ provides the SQL Explorer. You can use this to examine tables, as well as run custom SQL queries.

Note that the user roles govern both the tables that you have permission to see and the SQL queries you have permission to run.

2.1 Introduction

To access the SQL Explorer, click the SQL Explorer icon in the application menu. This opens a new browser or window, which contains a page of the Management Portal for InterSystems IRIS®, the product embedded within InterSystems® Data Fabric Studio™.

For your use cases, two areas are of primary interest:

- The tree on the left, where the **Tables** folder displays all the tables you have permission to see.
- The **Execute Query** tab on the right, which you use to execute SQL queries as described on this page.

The other SQL Explorer options are not necessarily helpful, but if you are interested, see Using the Management Portal SQL Interface in the latest InterSystems IRIS documentation.

2.2 Viewing the Contents of a Table

To view the contents of a table:

1. Click the **Tables** folder to expand it. The left area then displays all the tables that you have permission to see.

   For example:

2. Drag and drop a table name into the large box on the **Execute Query** tab.

In this box, you will then see an SQL SELECT query that lists all the fields in the table. For example:

```
SELECT  %IRISRowID, fruit, "count", comment, %T_roundcount,
%V_CountValidate, %BatchId, %StagingAction
FROM Staging_FileDir_step1_v1.sampledatacsv
```

3. Click **Execute**.

The system then executes that query, internally applying a cutoff so that only the first 1000 records are shown. By default, the system shows data in display mode. For example, this means that a date field is shown in human-readable form. Also, the length of data in a column is restricted to provide a manageable display.

You have the following additional options to refine what you see:

- Select a different mode; the options are **Display Mode** (the default), **ODBC Mode**, and **Logical Mode**.
- Specify a different value for the **Max** field, which limits how many rows of data to return from a query. This can be any non-negative integer, and the maximum is 100,000. (You can also limit the number of rows of data to return by using a TOP clause within the query.) Then click **Execute**.
- Edit the query, for example, by reordering or removing fields. Then click **Execute**.

2.3 Viewing Fields in a Table

In addition to simply viewing the data, to see the fields that are available in any table that you have permission to see, click the table name in the left area. The field names are then shown as in the following example:

2.4 Running a Custom Query

To run a custom query, you view a table as described above and then simply modify the query. Or you can type or paste a query into the large box on the **Execute Query** tab and then click **Execute**.

2.5 Show History

To rerun a query that has been run previously within the current session:

1. Click **Show History**.

   The page then displays a list of queries.

2. Click **Execute** in the row for the query you want to rerun.

2.6 See Also

- Defining User Roles
- Using the Management Portal SQL Interface (in the latest InterSystems IRIS documentation)

3

Defining Cubes

InterSystems® Data Fabric Studio™ includes InterSystems IRIS® Adaptive Analytics, a Business Intelligence tool powered by co-development with AtScale. This means that data analysts and data engineers can define cubes based on data in Data Fabric Studio, and then use those cubes for analytics.

You can also connect reporting tools or other analytics tools to these cubes, as described in the AtScale documentation.

3.1 Outline of Steps

You will need the AtScale documentation for details, but the general process is as follows:

1. Log on to the AtScale system, which has a different user interface.
2. Add a data warehouse that consists of the tables in Data Fabric Studio.
3. Define cubes based on those tables.
4. When appropriate, publish those cubes so that they can be used.
5. Then, within the main user interface for Data Fabric Studio, use the Business Scheduler to schedule the building of those cubes. Because it is best for the cubes to always display the most recent data, make sure to build any cube immediately after any necessary data is loaded; see Managing Task Dependencies.

3.2 See Also

- AtScale documentation
- Configuring the Adaptive Analytics Connection
- Scheduling and Running Tasks (for scheduling cube builds)
- Viewing Cube Build Information

4

Viewing Cube Build Information

If InterSystems® Data Fabric Studio™ has been configured to have a connection to InterSystems IRIS® Adaptive Analytics (AtScale), the Business Scheduler can build cubes. There is also an easy way to see the cube build information.

4.1 Viewing Cube Build Information

To view all the available cubes and the most recent build information for each:

1. Click the Analytics icon in the application menu.
2. Click **Cubes**.

The system then displays a table listing all the available cubes. This table provides the following information for each cube:

- **Status**—Status of the most recent cube build.
- **Organization**—Name of the organization that owns the project to which cube belongs.
- **Project**—Name of the project to which this cube belongs.
- **Cube**—Name of the cube.
- **Publish Date**—Date and time when the project was published.
- **Build Start**—Date and time when the cube build was started.
- **Build End**—Date and time when the cube build was completed.
- **Build Time**—Length of time for the cube build to complete. If the cube build is underway, this is the current duration.
- **Error Details**—Information about any errors during the cube build.
- **Visual Trace**—Provides a link to a trace session (for internal use).
- **Build Aggregates**—Button for the user to click to trigger a build. This button is disabled if the cube is currently being built.

4.2 Filtering the Cube Build Information

If this page has a large number of cubes, you may want to use the filters on the top of the page. These filters work as follows:

- **Organization**—Displays only the cubes that belong to a specific organization.
- **Project** —Displays only the cubes that belong to a specific AtScale project.
- **Cube**—Displays only the cubes whose name match the given string.

4.3 See Also

- AtScale documentation
- Scheduling and Running Tasks (for scheduling cube builds)
- Configuring the Adaptive Analytics Connection

5

Accessing Your Data via JDBC

InterSystems® Data Fabric Studio™ supports JDBC access to its tables.

5.1 Outline of Steps

To access tables within Data Fabric Studio via JDBC:

1. Obtain the JDBC connection string for the Data Fabric Studio database.
2. Use the connection string to connect just as you would with any other JDBC-compliant database.

5.2 See Also

- JDBC for Relational Access

6

Accessing Your Data via ODBC

InterSystems® Data Fabric Studio™ supports ODBC access to its tables.

6.1 Outline of Steps

To access tables within Data Fabric Studio via ODBC:

1. Obtain the appropriate ODBC driver. You can download drivers for Windows, Linux, and macOS from the InterSystems IRIS Drivers page.
2. Define a DSN to refer to the Data Fabric Studio database.

   The details depend upon the platform; see available links at Using the InterSystems ODBC Driver.
3. Use the DSN to connect to that database as you would with any other ODBC-compliant database.

6.2 See Also

- InterSystems IRIS Drivers page
- Getting Started: ODBC Connections to InterSystems Databases