CAMBRIDGE SEMANTICS

Anzo® 5.3 Getting Started Guide

Last Updated: 4/6/2023

Online documentation is available at docs.cambridgesemantics.com

© 2023 Cambridge Semantics, Inc.

Table of Contents

Getting Started 4

Anzo Concepts and Vocabulary 5

- Glossary 7
- Graph Data Storage Reference 12

Introduction to the Anzo Application 15

- General Interface Elements 17
- Onboard 22
- Model 29
- Blend 30
- Access 32

Tutorials 36

- Building a Knowledge Graph with the Automated ETL Workflow 37
  - Prepare the Environment 38
  - Download the Sample Data 39
  - Create a CSV Data Source and Import the Sample Data 40
  - Ingest the Imported Data Set 46
  - Run the ETL Pipeline 48
  - Create and Activate the Top Movie Graphmart 49
  - Explore the Data in a Hi-Res Analytics Dashboard 52
  - Review the New Artifacts 70
- Building a Knowledge Graph with the Direct Data Load Workflow 71
  - Prepare the Environment 72
  - Download the Sample Data 73

Add a CSV Data Source for the Sample Data 74
Create a Graphmart from the Data Source 80
Explore the Data in a Hi-Res Analytics Dashboard 83

Getting Started

The Getting Started Guide helps you get started with Anzo. It includes an introduction to Anzo concepts and the user interface and provides tutorials that guide you through the process of building a sample knowledge graph from start to finish.

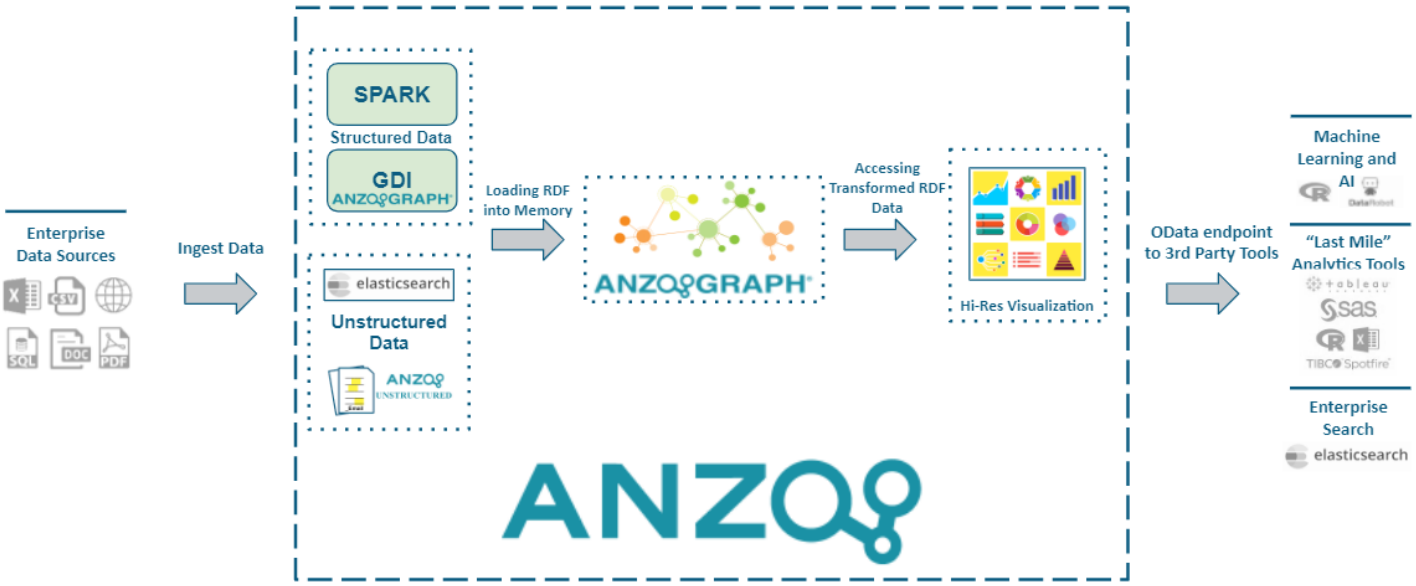
Anzo Concepts and Vocabulary

Introduction to the Anzo Application

Tutorials

Anzo Concepts and Vocabulary

This topic introduces you to key features and concepts to know when working with Anzo. The diagram below shows a high-level overview of Anzo components and concepts. Details about the components in the image are described below.



Enterprise Data Sources

Anzo supports ingesting data from structured, semi-structured, and unstructured data sources. Anzo connects to relational database sources via ODBC and JDBC drivers and supports loading data from CSV, JSON, XML, SAS, and Parquet files. Unstructured data sources such as documents, PDFs, text snippets, web pages, and content from knowledgebases are also supported.

Ingest Data

Structured and semi-structured data sources like relational databases and flat files can be directly loaded using AnzoGraph's Graph Data Interface (GDI) or Anzo's built-in ETL pipelines.

Unstructured data sources are onboarded using configurable unstructured pipelines. These pipelines onboard unstructured source files, integrate metadata about those files, and make the full text in those files as well as key facts available as part of graph data models.

Loading RDF into Memory

Whether data is ingested with the GDI, ETL pipelines, or unstructured pipelines, it is converted from its original format to a new format that describes the data as a graph model. This format, Resource Description Framework (RDF), simplifies access to complex data and flexibly accommodates new data sources and use cases. The RDF data is added to Graphmarts and loaded to AnzoGraph, Anzo's in-memory massively parallel processing (MPP) graph OLAP engine, for further transformation and analytics.

Accessing Transformed RDF Data

Users have several options for accessing and analyzing knowledge graphs. Anzo's Hi-Res Analytics application enables users to create dashboards for exploring and visualizing the data without needing to have specialized query knowledge. And, in line with Anzo's open standard architecture, Graphmarts can be accessed using modern application program interfaces (APIs) as well as SPARQL-compliant query endpoints. Anzo also offers standards-compliant Open Data Protocol (OData)-based REST endpoints as part of its Data on Demand service. The Data on Demand service provides access to data from business intelligence tools.

Related Topics

Glossary

Introduction to the Anzo Application

Glossary

This topic defines commonly used Anzo terms and phrases.

- Anzo Data Store
- AnzoGraph
- Blend
- Data Layers
- Datasets Catalog
- ELT
- ETL
- File-Based Linked Data Set
- File Store
- Graph Data Interface
- Graphmarts
- Hi-Res Analytics
- IRI
- Journal or Volume
- Linked Data Set
- Model
- Managed Model
- NLP
- OData
- OSGi
- Onboard
- Registry
- URI

Anzo Data Store

An Anzo data store defines an endpoint for writing data. It specifies the directory on the file store where Anzo can generate file-based linked data sets for unstructured pipelines (see File-Based Linked Data Set). It can also be used as a designated a directory to use for saving data sets created from graphmart exports. The data store also defines write properties such as the maximum file size and whether files should be compressed.

AnzoGraph

AnzoGraph is Anzo’s in-memory massively parallel processing (MPP) graph OLAP engine.

Blend

Using semantic models, separate data sets or new data sources can be blended into a knowledge graph. Anzo can combine and align any data set as well as apply data cleansing and/or transformation steps. Anzo delivers blending and access through Graphmarts, which give users the flexibility to combine and analyze any subset of data in Anzo.

Data Layers

Data layers enable users to enhance graphmarts dynamically. Users can create layers to load additional data sets, clean, conform, or transform data, infer new information, or export data to a file-based linked data set (FLDS).

Datasets Catalog

Anzo’s Dataset catalog combines traditional technical, operational, and business metadata with a semantic layer to describe all aspects of enterprise data elements. The catalog enables Anzo’s unique use of semantics and graph models and is the system of record for data in Anzo.

Anzo collects and generates metadata at every stage in the data discovery and integration process. Metadata in the catalog documents how data is converted during the onboarding process from its original format into a graph model. Subsequent data blending, transformation, and preparation steps are captured as additional metadata. Anzo also captures new metadata to describe all actions taken against data within Anzo. The metadata enables users to visualize their data, understand business contexts, identify connections, and blend and prepare data.

ELT

Structured and semi-structured data sources are onboarded using an extract, load, and transform (ELT) workflow as opposed to a traditional ETL flow. Data Layers are Anzo's mechanism for flexibly transforming data in memory.

ETL

The Spark extract, transform, and load (ETL) process takes source data and converts it to the graph data model using a source to target mapping. Anzo’s mapping tool enables users to define field-level transformations, including type casting, date conversions, unit conversions, etc., as data is onboarded to Anzo.

File-Based Linked Data Set

When the unstructured pipeline workflow is used to onboard data, Anzo creates a data set in the Dataset catalog. The data set in the catalog is registered in the Anzo system data source (see Journal or Volume) and includes metadata about the data, including a pointer to the data store location for the RDF files generated by the ETL pipeline. The catalog data set and the files on disk are known as a file-based linked data set (FLDS).

A file-based linked data set can also be generated by including an Export Steps in a graphmart.

File Store

The Anzo platform components, AnzoGraph, Anzo Unstructured, and Elasticsearch share a file system for maintaining onboarded data and supporting files. A file store is the shared file storage system, such as NFS, HDFS, or cloud storage, that is shared between the servers.

Graph Data Interface

The Graph Data Interface (GDI) (sometimes called the Data Toolkit) is a flexible AnzoGraph extension that enables users to access a variety of data sources via SPARQL queries. The GDI has built-in, native support for various file format types, HTTP/REST endpoints, and JDBC connections to common database sources. For more information about the GDI, see Introduction to the GDI in the User Guide.

Graphmarts

Graphmarts are collections of knowledge graphs that users can blend and enhance. Graphmarts can combine any subset of data in Anzo for analysis. For information about graph storage, see Graph Data Storage Reference.

Hi-Res Analytics

Hi-Res Analytics enable users to explore and ask questions across all of their data. Using model-guided dashboards, users can perform computations across multi-dimensional data. Hi-Res Analytics dashboards generate complex graph queries dynamically based on user input.

IRI

An Internationalized Resource Identifier (IRI) is similar to URI but allows a greater range of characters. URI and IRI are often used interchangeably.

Journal or Volume

A journal, also known as a volume, refers to data that is stored in Anzo's embedded graph store. The graph store is transactional and is used to persist metadata, which is written to disk in a .jnl file. The system volume (or system data source) is the default, required volume where Anzo stores ontologies as

well as system configuration, dataset, catalog, registry, and access control metadata. Users can create secondary local volumes that are used for more compartmentalized data and can be created and deleted without affecting the core system.

Linked Data Set

A linked data set (LDS) is a fundamental concept. Anzo organizes all data, including system data, into linked data sets. An LDS is associated with a data model and can be searched, discovered, shared, and protected with access control. For example, graphmarts are organized in a linked data set or registry of graphmarts, the Activity Log is a linked data set, data source configurations exist in a linked data set, and so on.

Model

Anzo establishes the semantic layer by enabling users to convert diverse enterprise data models into graph data models and then enhance the data by adding new business definitions, names, and tags. Further insight is added when data from separate sources are linked, connecting shared business definitions across previously siloed sources. Anzo employs open World Wide Web Consortium (W3C) standards, including Web Ontology Language (OWL), RDF, and SPARQL to model, connect, and query interconnected graphs.

Managed Model

A managed model is a model that is generated by the Graph Data Interface (GDI) and owned and managed by a data layer with a Direct Load Step. Any manual changes made to a managed model outside of a Direct Load step, such as from the Model editor, will be overwritten any time the graphmart or layer is refreshed or reloaded. For more information, see Introduction to Managed Models in the User Guide.

NLP

Anzo performs named-entity recognition (NER) using knowledge bases and can interface with natural language processing (NLP) tools. It serves as a platform that enables text analytics through interplay between best-of-breed NLP tools.

OData

Open Data Protocol (OData) facilitates the creation of interoperable RESTful APIs. The Anzo Data on Demand service provides OData-based feeds that can be used to query graphmart data from third-party business intelligence tools.

OSGi

The Open Service Gateway Initiative (OSGi) is the open-standard architecture upon which Anzo is built. It is a Java framework for developing and deploying software programs and libraries. OSGi enables Cambridge Semantics to compartmentalize Anzo into "bundles" that can be deployed, activated, and removed independently without affecting other bundles in the system.

Onboard

When data is ingested from its source platform to Anzo, it is converted from its original format to a new format that describes the data as a graph data model. This format, Resource Description Framework (RDF), captures each data value and relationship. RDF data is loaded to Anzo's in-memory graph engine, AnzoGraph for transformation and analysis.

Registry

Anzo manages configurations in system-level registries. Each registry is a collection of application and system component configurations of the same type. Like data, registries are stored and managed with RDF named graphs according to ontologies. Technically, a registry is a Linked Data Set.

URI

A Uniform Resource Identifier (URI) is a globally unique identifier for a piece of information. A URL (Uniform Resource Locator) is a URI that specifies a location, such as a web address.

Related Topics

Introduction to the Anzo Application

Graph Data Storage Reference

This topic describes the way onboarded graph data is shared between and stored in the Anzo and AnzoGraph graph stores.

The onboarding process generates different types of graph data artifacts. Storage of the artifacts differs based on the type of data that is being stored and the purpose of the data. The list below describes the artifacts and storage methods:

- The metadata, such as data models, data source configuration details, dataset catalog entries, registries, and access control definitions, are stored in Anzo's embedded graph store. The Anzo graph store is a transaction-oriented store that is built for processing many updates to small amounts of data. Data is persisted to disk in a journal, also known as a volume. The system volume (or system data source) is the default, required volume where Anzo stores models as well as system configuration, dataset, registry, and access control metadata. Users can create secondary local volumes that are used for more compartmentalized data and can be created and deleted without affecting the core system.
- When an unstructured pipeline is run or data is exported from a graphmart, the instance data and copies of the data models are written to a file-based linked data set (FLDS) on the shared file store. Each FLDS is represented as a dataset in Anzo's Datasets catalog. The catalog entry includes a pointer to the RDF files on disk.
- When a dataset from the catalog is added to a graphmart and the graphmart is activated, Anzo loads the data from the FLDS into the AnzoGraph graph store. Once the data is in memory, the rest of the graphmart's data layer steps are executed by AnzoGraph (known as the ELT process). Each data layer becomes a graph in AnzoGraph, and each layer graph includes the instance data created by that layer as well as the related models.
- Anzo system ontologies and metadata remain in Anzo's graph store, the system data source. They are not loaded to AnzoGraph unless the system data is added to a graphmart and the graphmart is activated.

Example

An Anzo instance has two active graphmarts. Each graphmart has two data layers, one for loading datasets into memory and another for creating views and running ELT queries. When the following query is run against AnzoGraph to return a list of all distinct graphs, the results show that there are five graphs:

```
SELECT DISTINCT ?graph
WHERE {
  GRAPH ?graph {
    ?s ?p ?o
  }
}
```

```
graph
-------------------------------------------------------------------------------
http://cambridgesemantics.com/Layer/546fb89ac6d245f8bea2777a52077bc9
http://cambridgesemantics.com/Layer/1162fb0d0b724a18b4133c10d69f16b7
http://cambridgesemantics.com/Layer/12c7eedddff9449ab4b133373b56e65c
http://cambridgesemantics.com/Layer/b69bb3295ba3434e846b1ed372039416
http://cambridgesemantics.com/GqeDatasource/guid_
10492203b5aa4a54f217ababb3dc6dee
5 rows
```

The first four graphs are the data layers for the two graphmarts. The graph URIs match the data layer URIs in Anzo. How do I find the graph URI for a Data Layer in a Graphmart?

The last graph in the results above is the AnzoGraph data source graph. This graph contains one triple that records a timestamp for the last time the data source was updated. If Anzo loses the connection to AnzoGraph, it checks this timestamp when it reconnects. The last updated time is used to determine whether the Anzo and AnzoGraph graph stores are in sync or if the graphmarts need to be reloaded to AnzoGraph.

> **Note**
> AnzoGraph does not have a "graphmart" construct, and graphmart URIs do not exist in the database. Though a graphmart acts as a container for data layers and its metadata can be queried in Anzo's embedded graph store, it does not include instance data that is needed by AnzoGraph.

Typically organizations manage all data with Anzo, i.e., data is onboarded to Anzo through unstructured pipelines or it is dynamically blended into data layers with the Graph Data Interface. Anzo then loads the data to AnzoGraph for analytics. When data is loaded to AnzoGraph through Anzo, Anzo manages the reloading of graphmarts if AnzoGraph is restarted. Though users can load data and create named graphs directly in AnzoGraph, AnzoGraph is not configured by default to persist the data in memory to disk. Graphs that do not originate in Anzo must be reloaded manually any time AnzoGraph is restarted. If you want to work with named graphs directly in AnzoGraph, consider configuring AnzoGraph to save data to disk. For more information, see Using AnzoGraph Persistence in the Administration Guide.

Related Topics

Onboarding Structured Data in the User Guide

Onboarding Unstructured Data in the User Guide

Introduction to the Anzo Application

This topic provides instructions for accessing the Anzo user interface and gives a general overview of the design and layout. For information about the Anzo concepts and procedures that influence the user interface's design and use, see Anzo Concepts and Vocabulary.

To access the user interface:

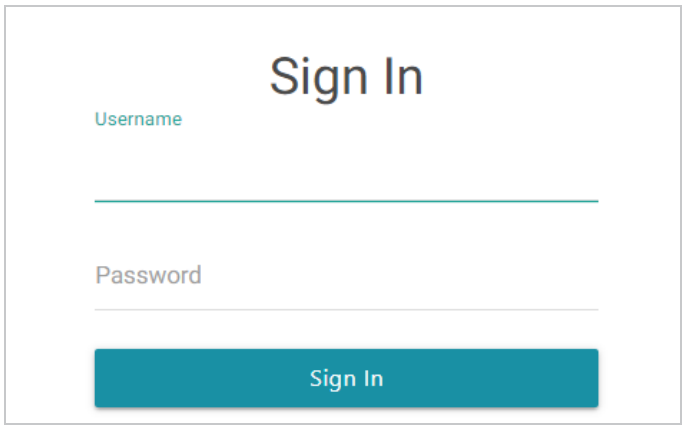
1. Go to the following URL in your browser:

   ```
   https://hostname
   ```

   Where **hostname** is the Anzo server DNS name or IP address.

   > **Note**
   > Your browser may warn you that the server's SSL certificate has not been signed by a trusted authority. This is normal behavior. To continue, click the **Advanced** link at the bottom of the page and then click the **Proceed** link. If you require a trusted site certificate, you can obtain one from a Certificate Authority and add it to the Anzo server.

2. On the Sign In screen, type your username and password and then click **Sign In**.



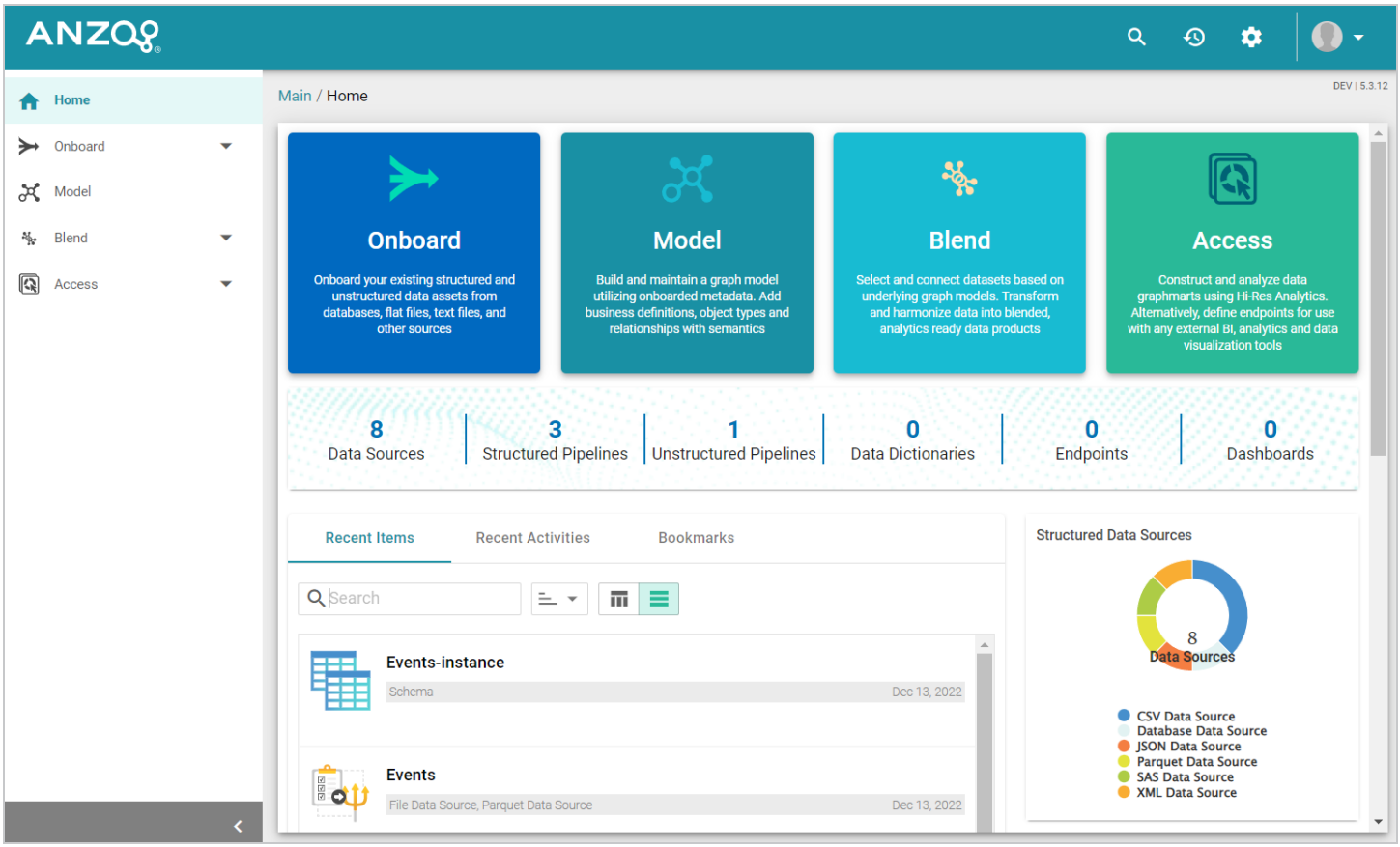
The user interface is organized and compartmentalized by the processes involved with building solutions. The design accommodates use cases where users with different permissions and responsibilities build various parts of a solution. The sections below introduce you to each of the user interface elements.

Note
The following sections show an administrator view of the user interface. In addition, the sections show a view of the user interface when the Spark component is installed with Anzo. Depending on the configuration of the environment and their assigned system role, some users might not see all components. For information about default Anzo roles and permissions, see Predefined Anzo Roles and Permissions in the Administration Guide.

- General Interface Elements
- Onboard
- Model
- Blend
- Access

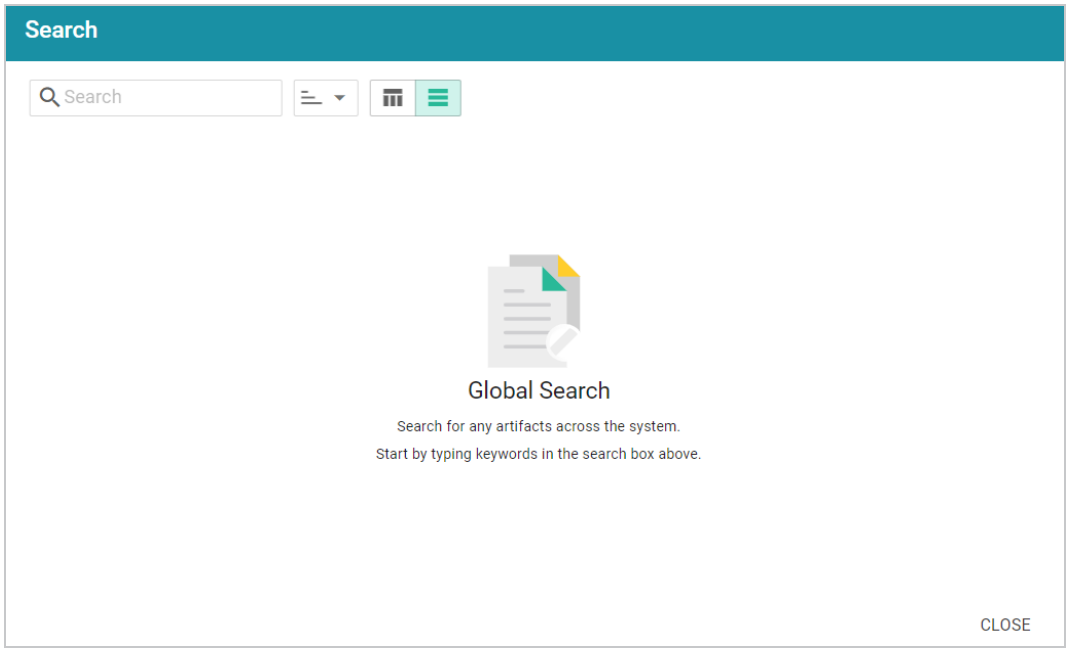
General Interface Elements

Each time you log into the application, the Home page is displayed and provides links to the features the logged in user has access to. Below the icons is a dashboard that displays an overview of the system artifacts and recently updated or bookmarked items. The dashboard becomes populated once you start onboarding data sources.

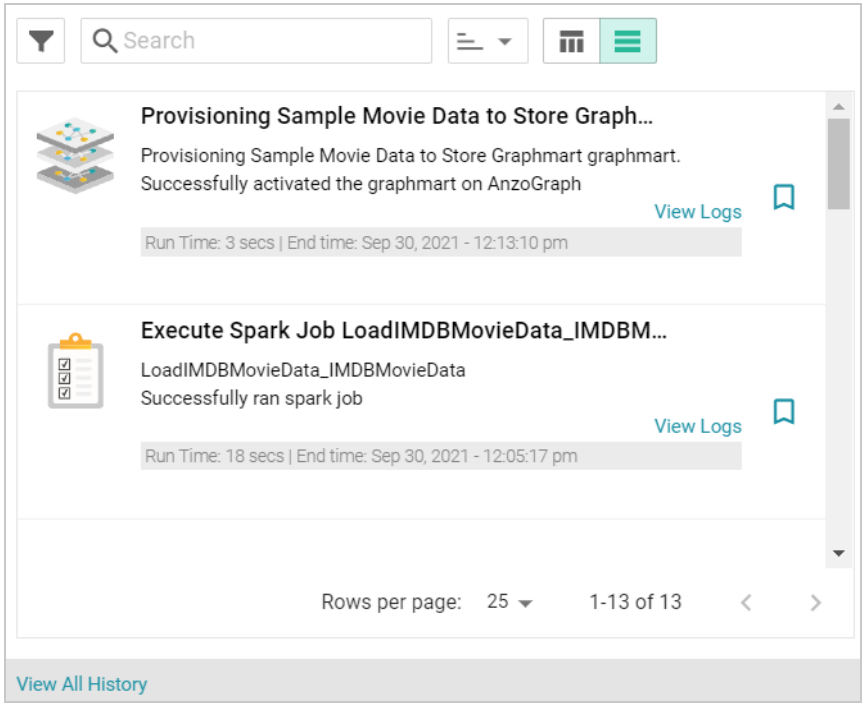


On the bottom left, you can click the collapse icon (**<**) to expand and collapse the left navigation menu. The right side of the top menu bar includes the following options:

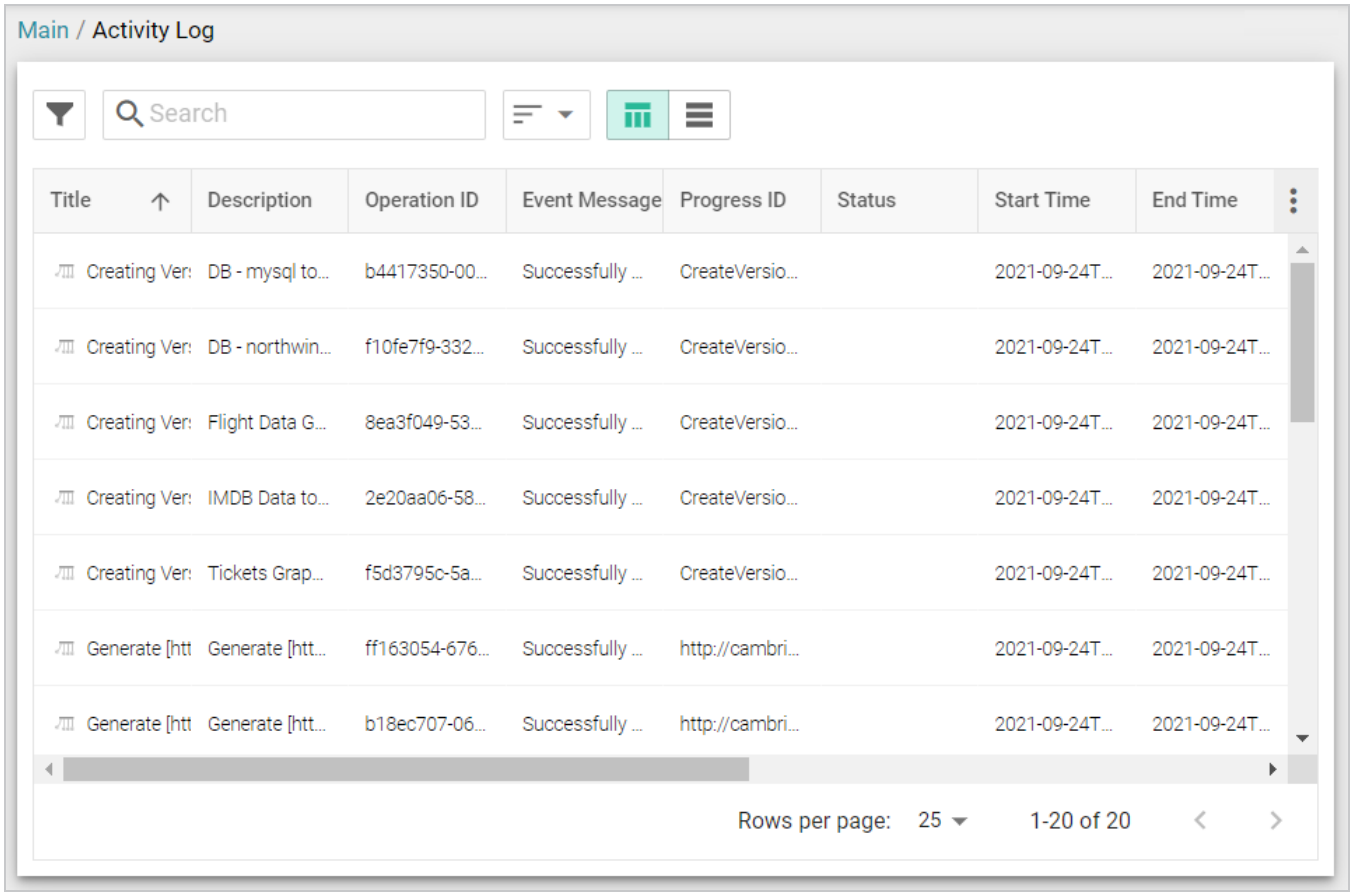
- The Search icon ( ) opens the Search dialog box where you can do a global search for a resource, such as a data source, dataset, graphmart, or schema, by title. The Search field accepts wildcard characters.



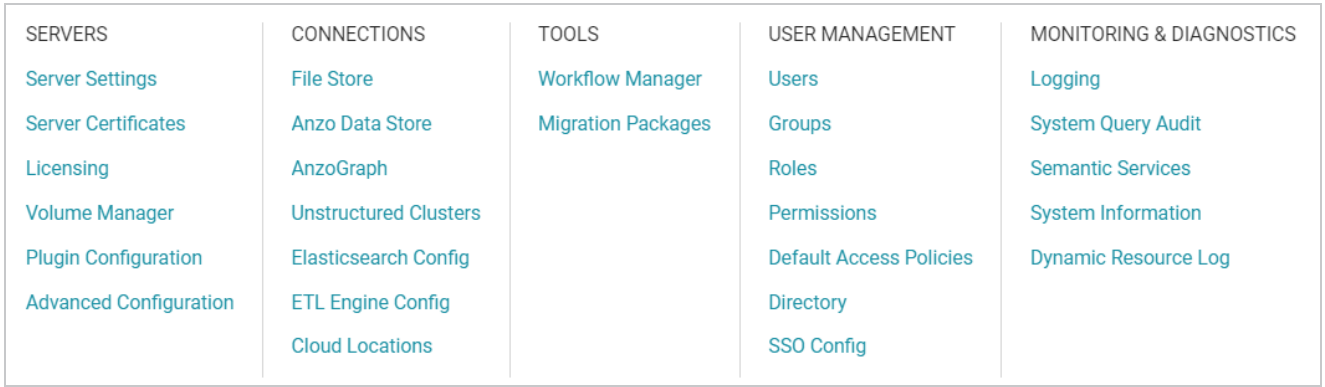
- The Activity Log icon ( ) opens the Activity Log, which offers a high-level view of individual activities, such as ETL engine and AnzoGraph usage and pipeline activity. The ability to filter activities based on their status and type enables users to drill down to activities of interest as well as profile user behavior. Additionally, the ability to include system data expands the scope of monitoring and aids in troubleshooting.

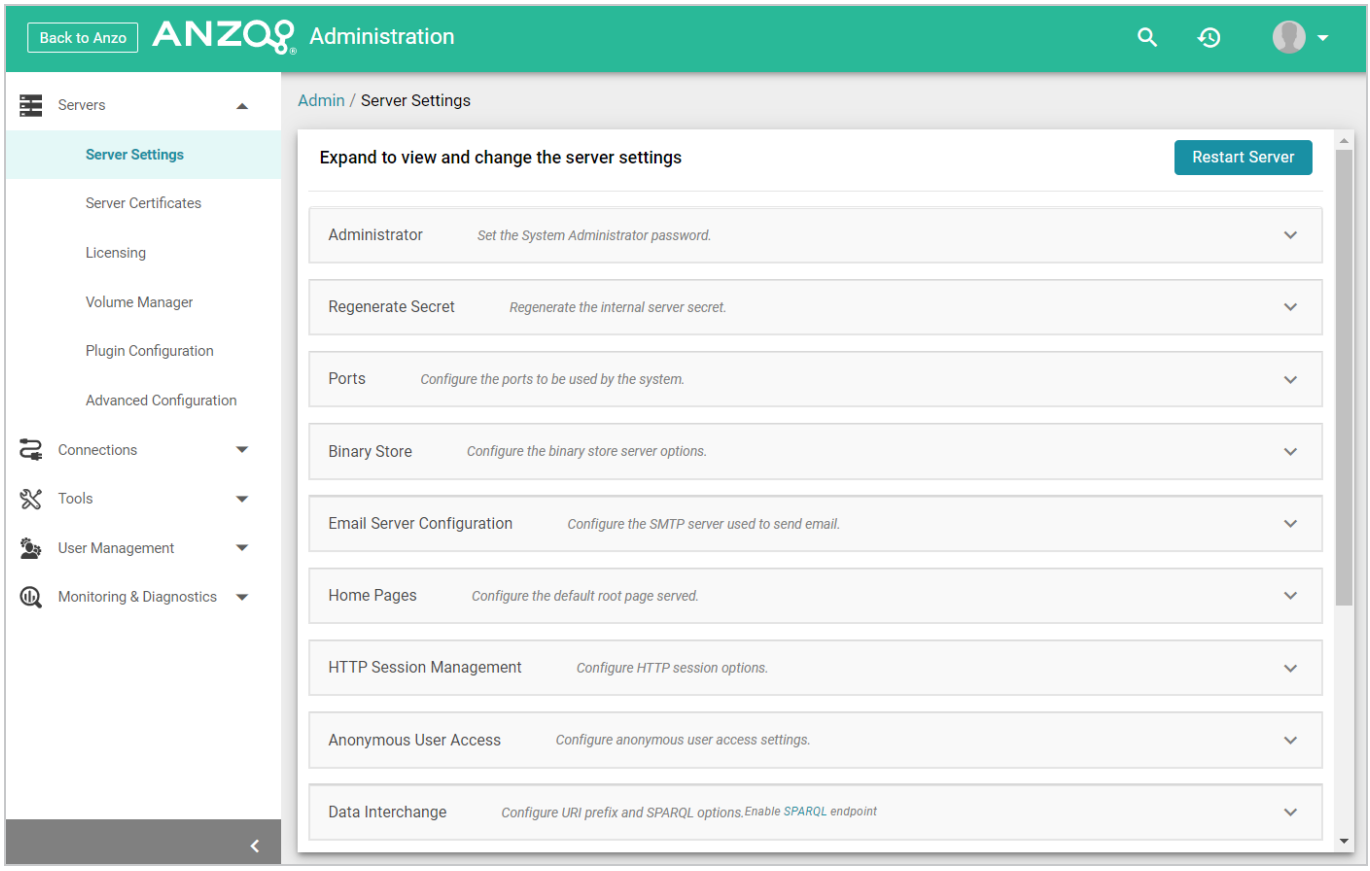


Once you start using Anzo and there are events that show up in the log, you can click **View All History** to open the full log for searching and filtering.



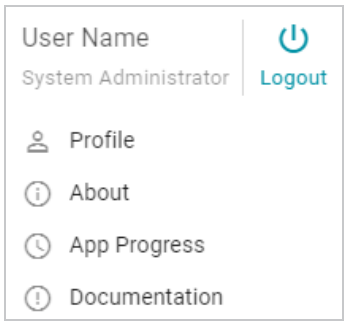
- The Administration icon ( ) opens the Administration menu. Clicking a menu option opens the Administration user interface for managing Anzo server settings, connections, users, and other components.



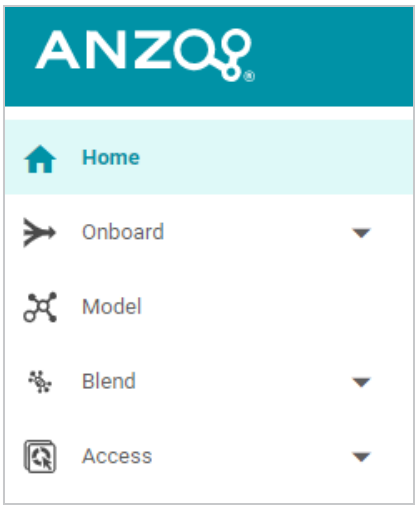


For information about the Administration application, see the Administration Guide.

- The User menu ( ) provides access to your user profile, the About screen, and the Application Progress window, which lists recent application activity. It also includes the ability to log out of the application and a Documentation link that opens this guide.

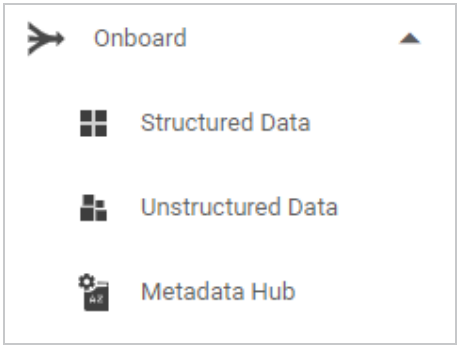


The navigation menu on the left provides access to all of the Anzo features.


ANZOG
Home
Onboard
Model
Blend
Access

Onboard

The Onboard menu provides access to the components that users configure to ingest data from various sources.

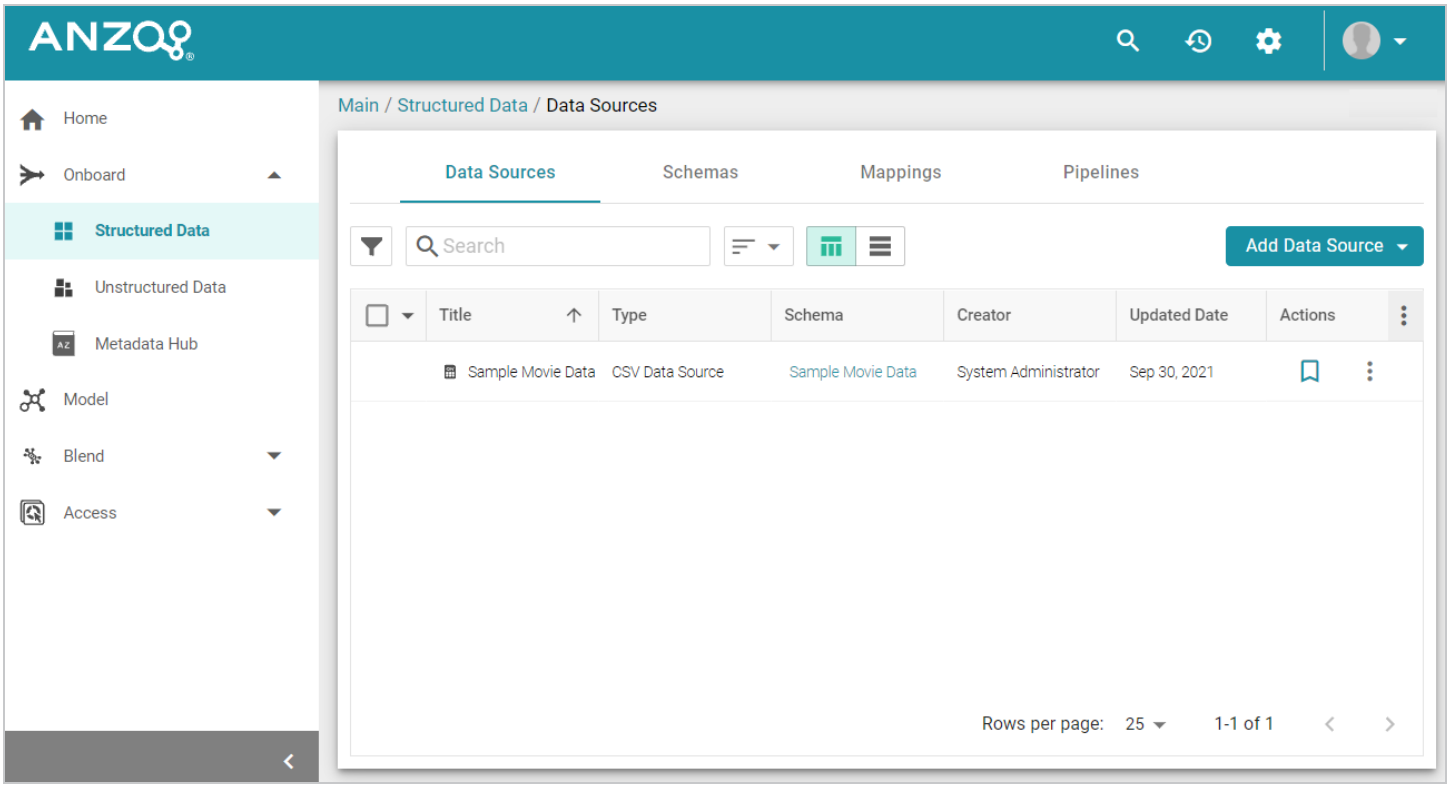


Structured Data

From **Structured Data**, users access data sources, schemas, mappings, and pipelines for structured data sources: database connections, CSV, JSON, XML, Parquet, and SAS files.

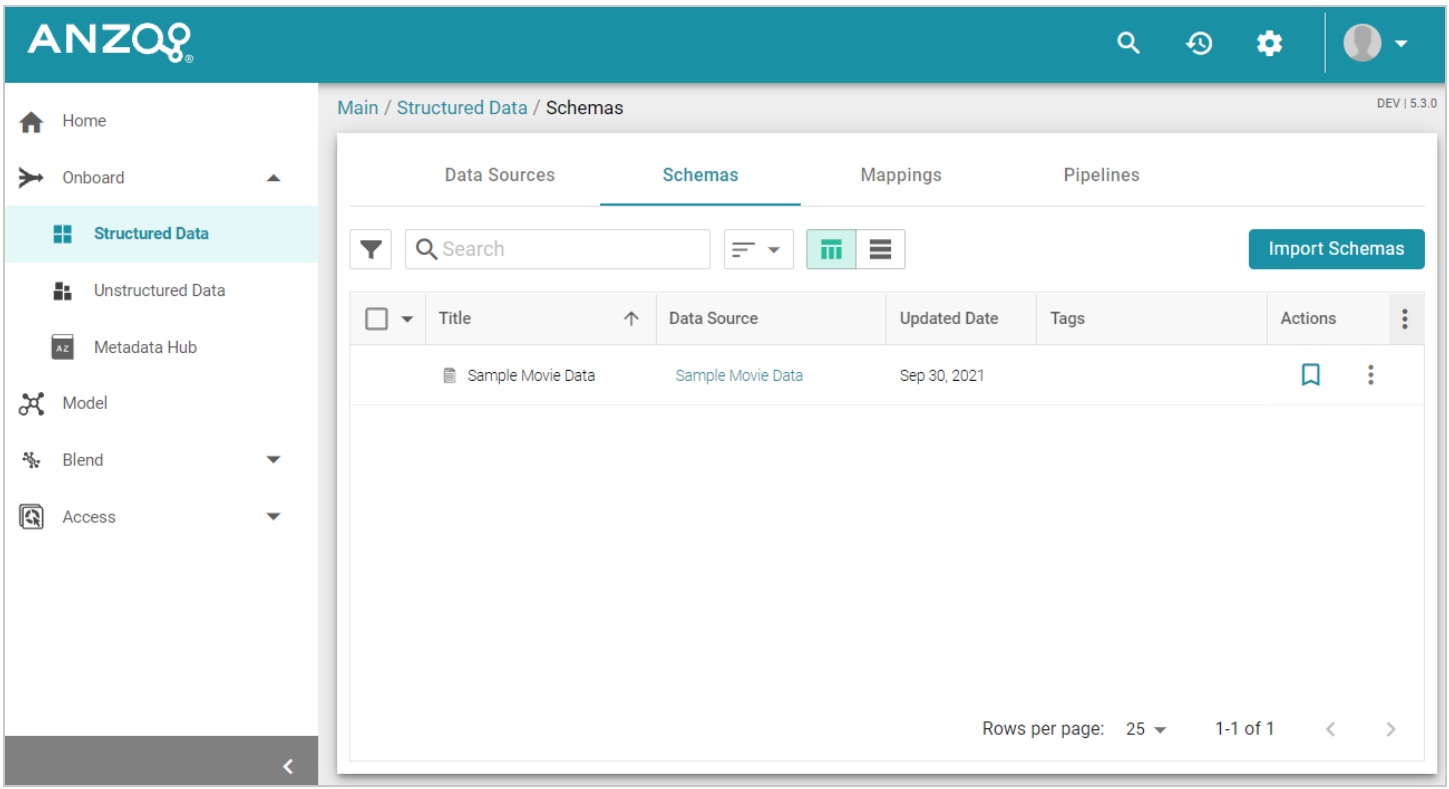
Data Sources

From the **Data Sources** tab, users connect to the files and databases that contain the data to onboard:



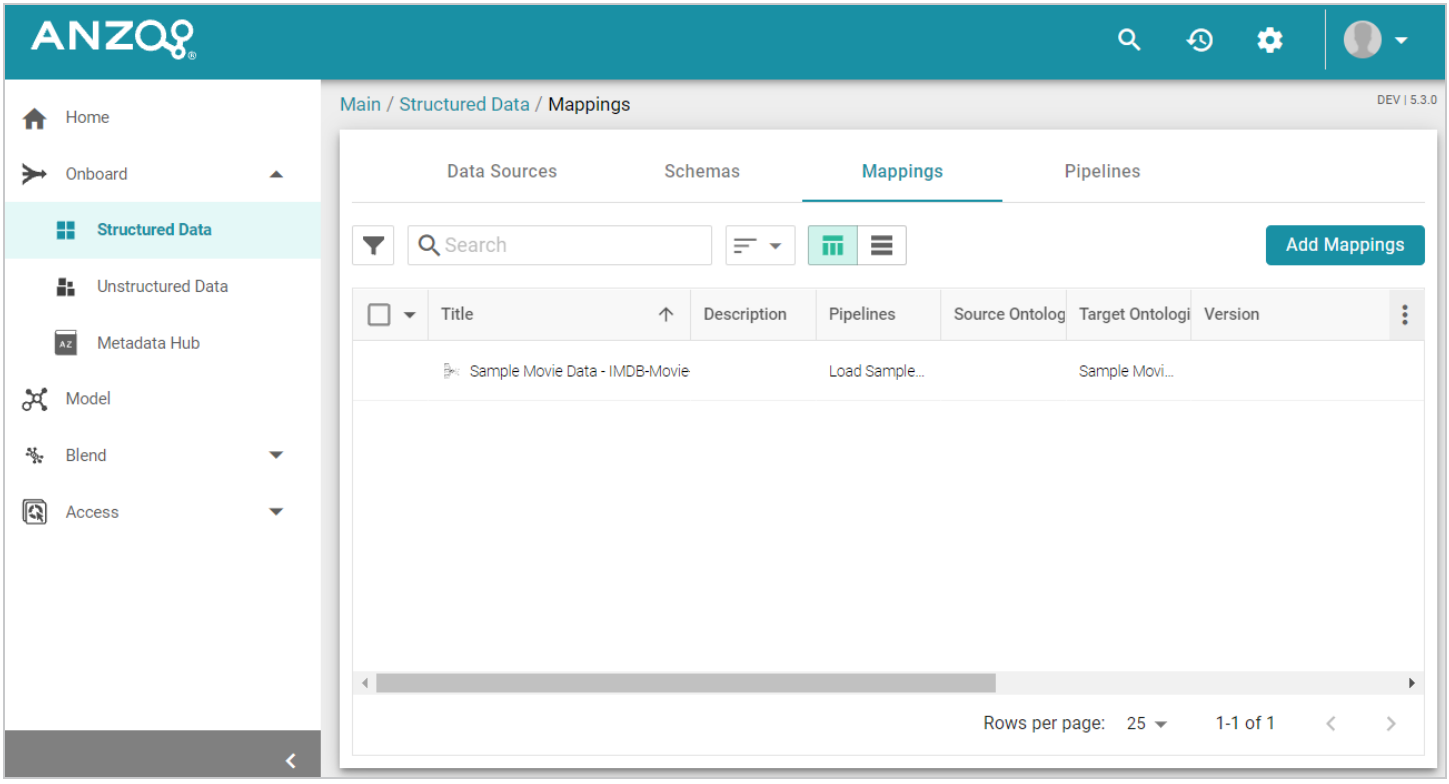
Schemas

From the **Schemas** tab, users view, create, and manage the schemas that define the data to onboard:



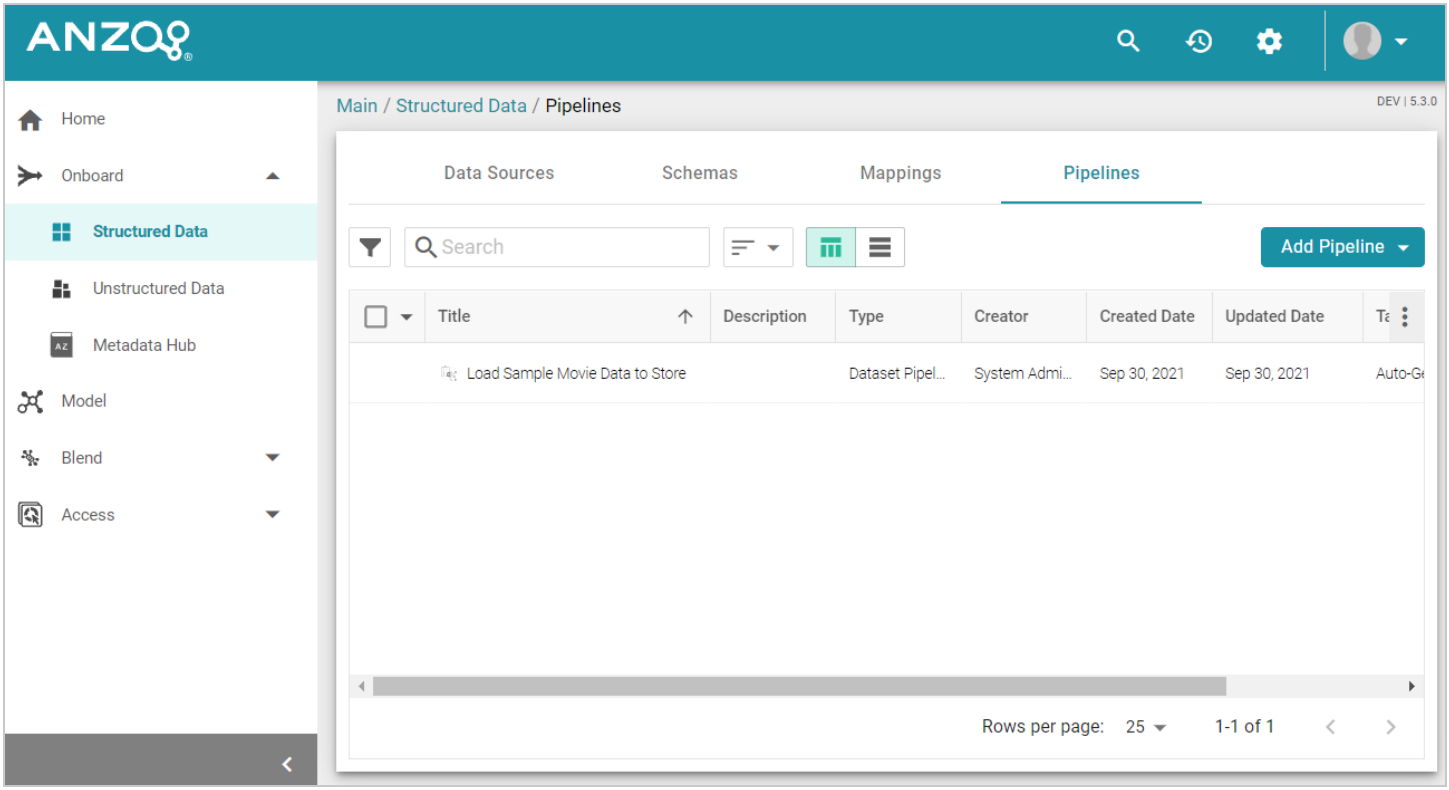
Mappings

From the **Mappings** tab, users view, create, and manage the mappings that describe the relationships between schemas and data models as well as perform optional transformations on the source data:



Pipelines

From the **Pipelines** tab, users view, create, and manage the pipelines that run ETL jobs to onboard data into Anzo:

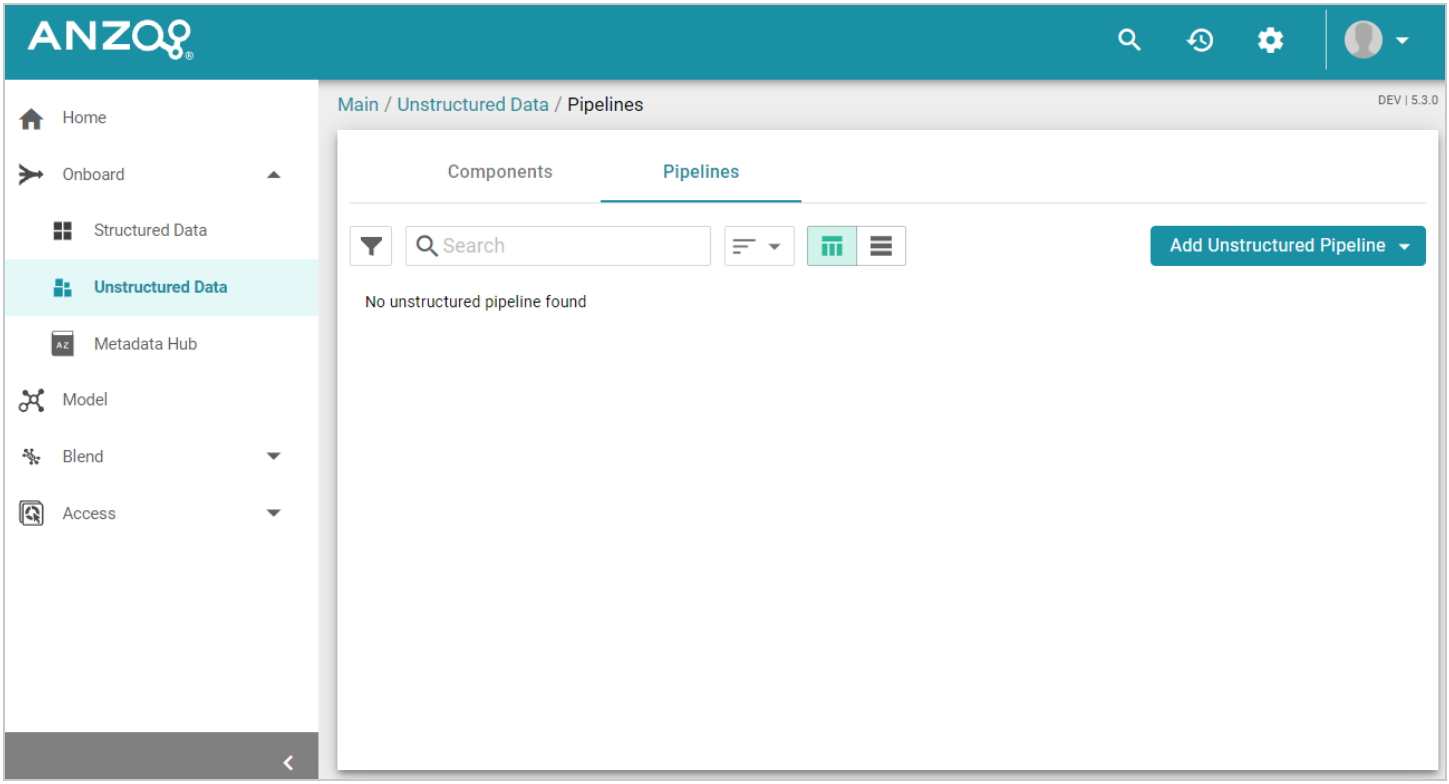


Unstructured Data

From **Unstructured Data**, users access and create pipelines for onboarding data from unstructured sources such as Office documents, PDFs, web pages, email messages, and knowledgebases.

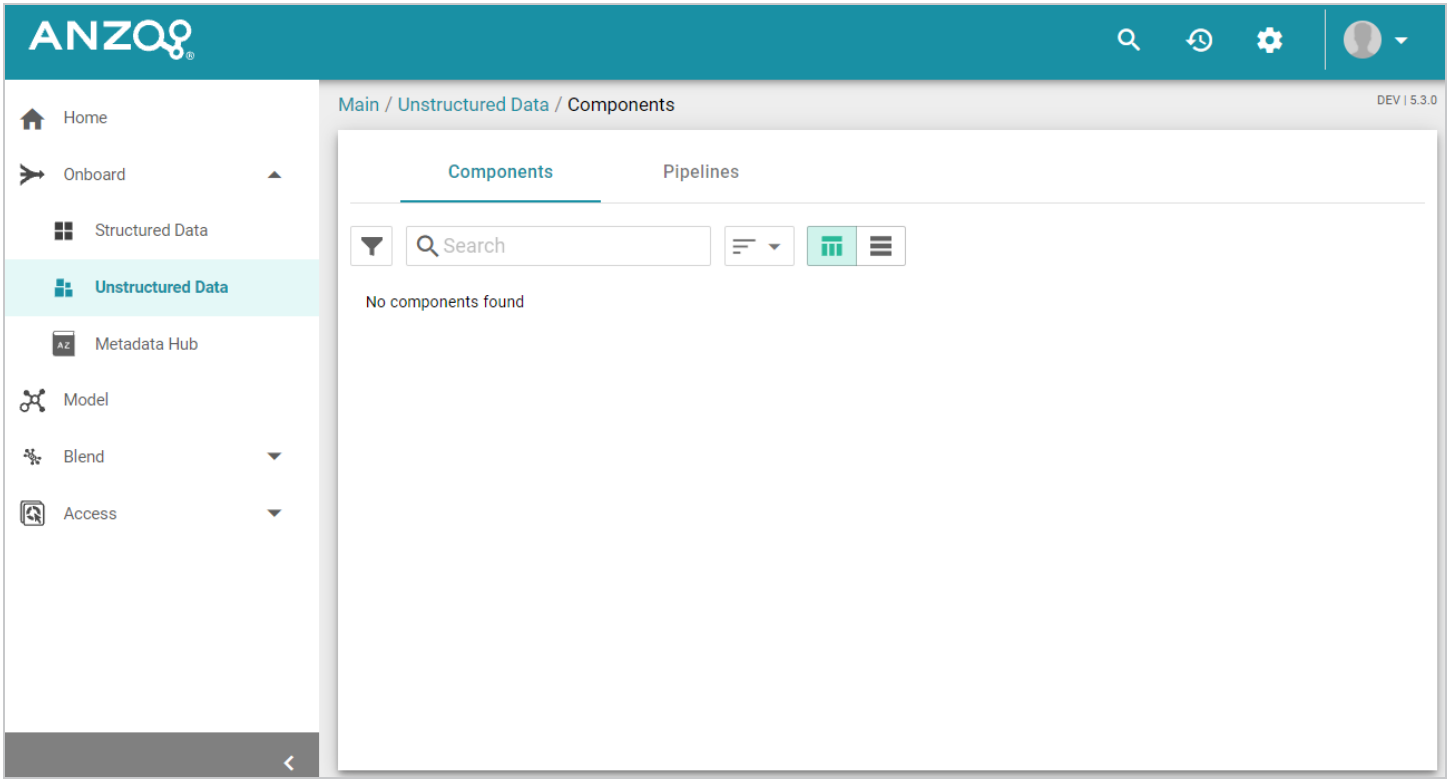
Pipelines

From the **Pipelines** tab, users can create and manage unstructured pipelines:



Components

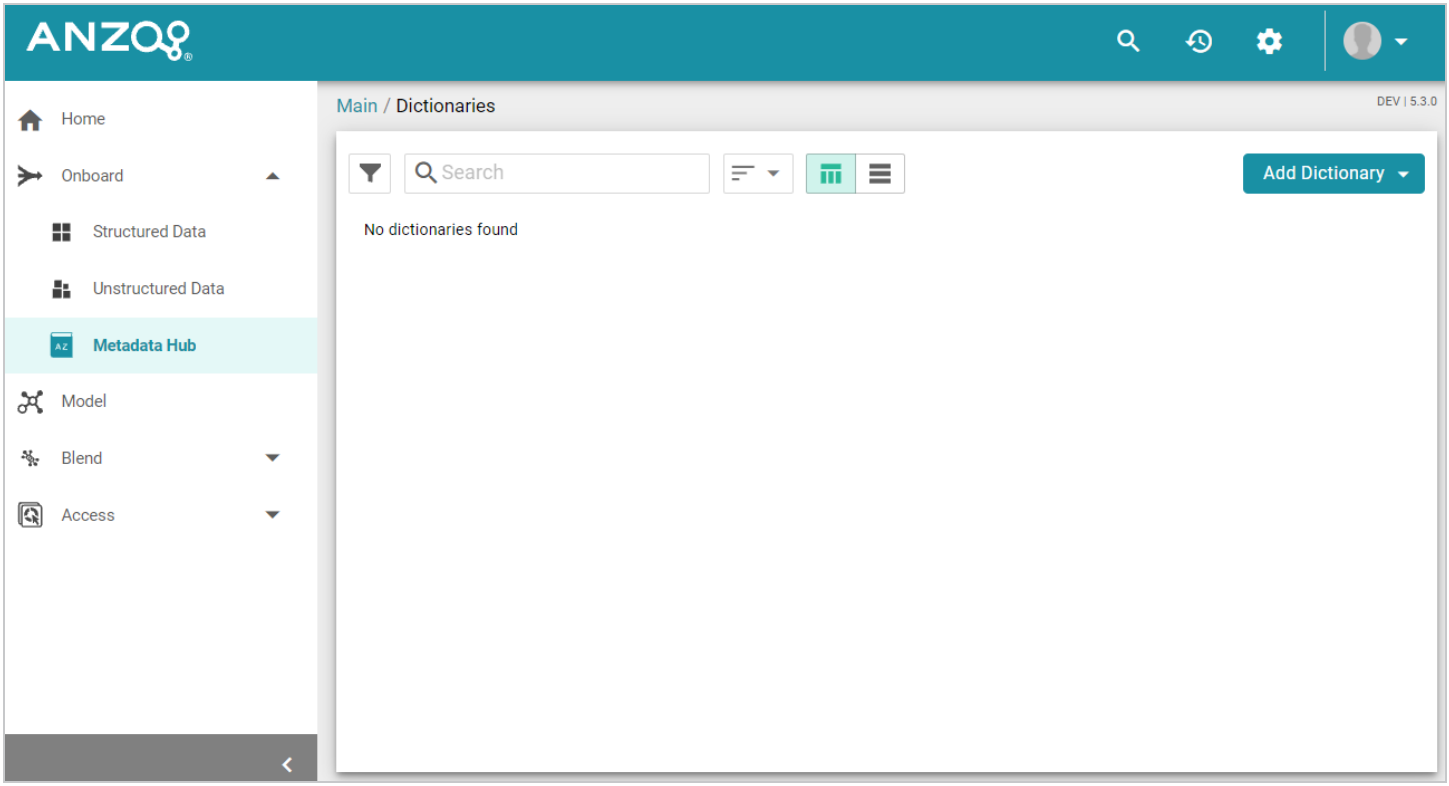
From the **Components** tab, users can view and configure unstructured pipeline components like crawlers and annotators:



Metadata Hub

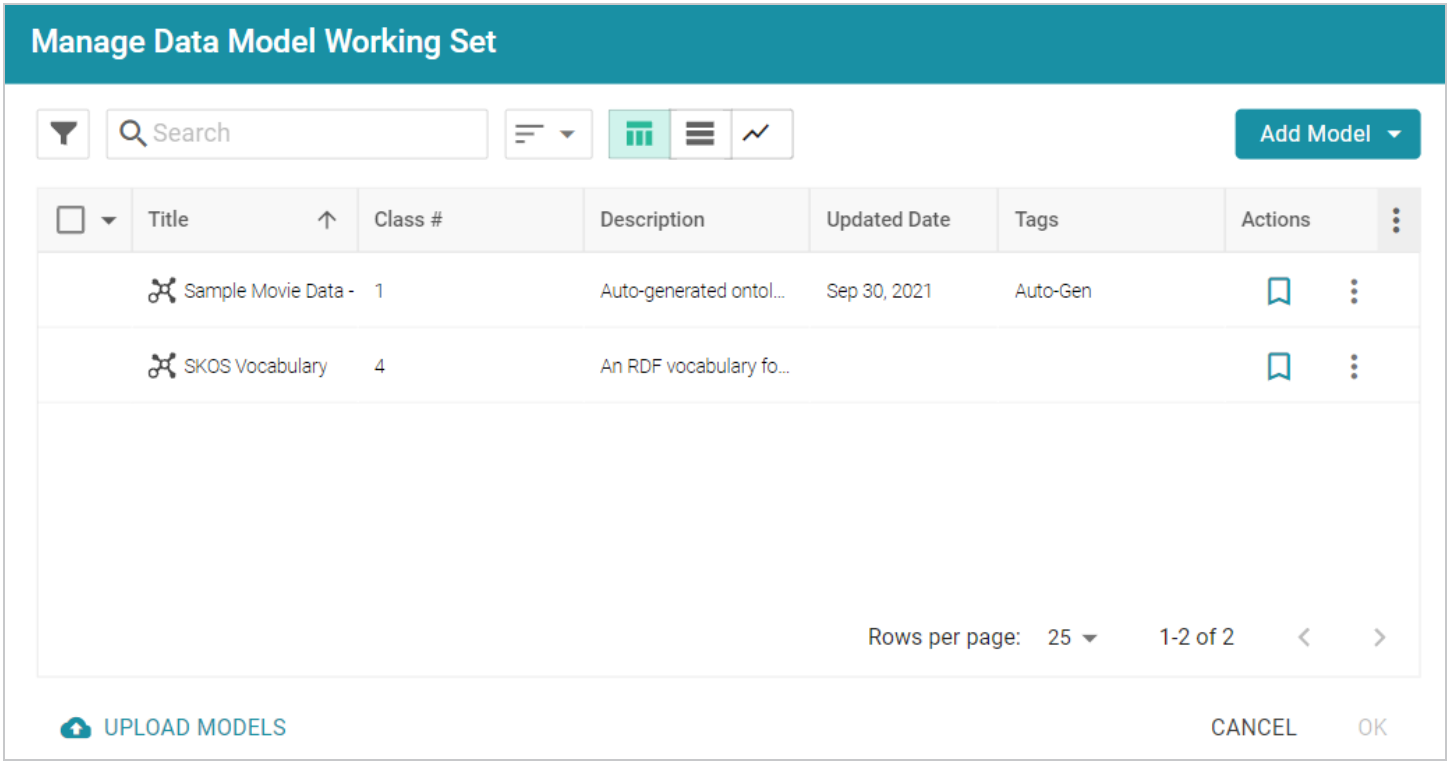
> **Note**
> By default, the Metadata Hub is available only to the sysadmin user. Whether you can access Metadata Hub depends on whether the sysadmin user has enabled the feature for additional users or groups.

From **Metadata Hub**, users view, create, and manage metadata dictionaries, which define and normalize concepts across data sources.


ANZO
Home
Onboard
Structured Data
Unstructured Data
Metadata Hub
Model
Blend
Access
Main / Dictionaries
DEV | 5.3.0
Search
Add Dictionary
No dictionaries found

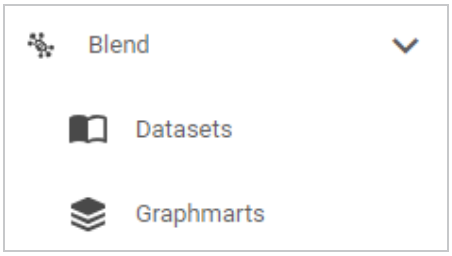
Model

The **Model** manager enables users to view, create, and manage data models, which describe the concepts, attributes, and relationships in or across data sets.



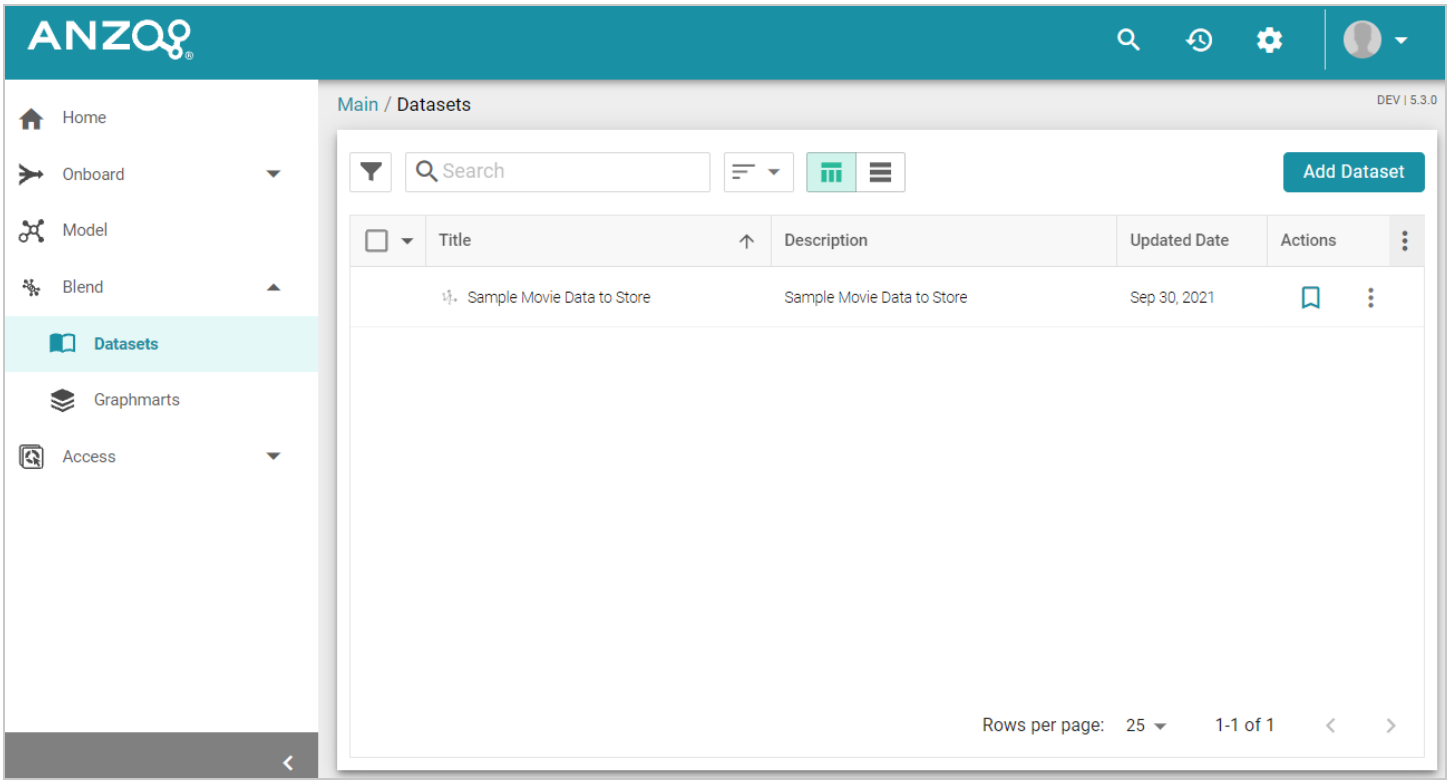
Blend

The **Blend** menu provides access to features that combine various datasets from different sources.



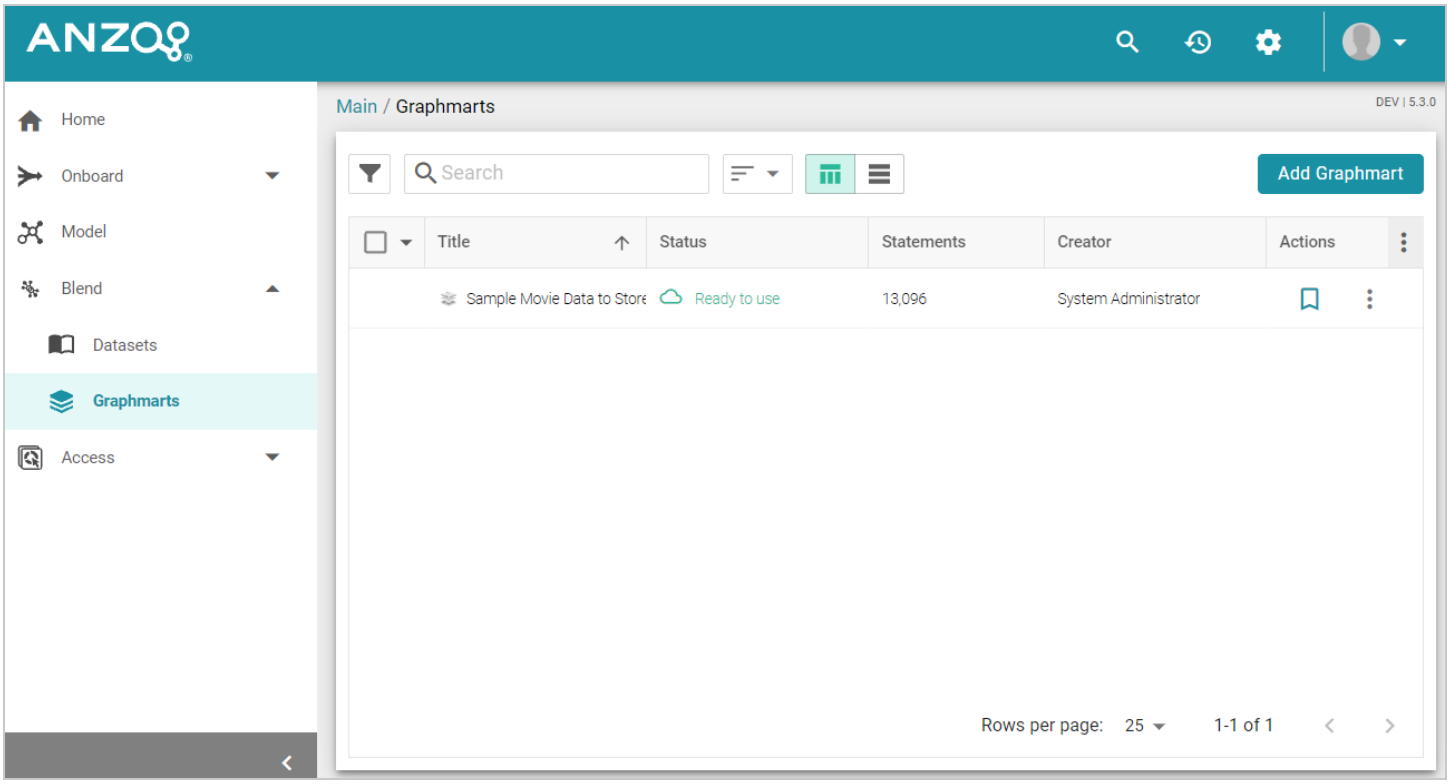
Datasets

The **Datasets** catalog is an inventory of all of the data in Anzo. Users can sort the list by tags, classes, and dataset creators. Users can also add datasets to graphmarts for loading into AnzoGraph and then designing Hi-Res Analytics dashboards.



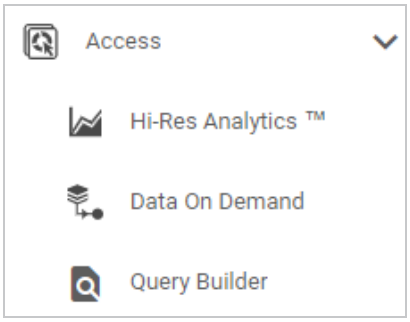
Graphmarts

The **Graphmarts** page lists all of the existing graphmarts. Users can designate favorites and sort the list by tags and graphmart creators. Users can click a graphmart in the list to view details such as the datasets included in the graphmart, data layer details, and associated dashboards.



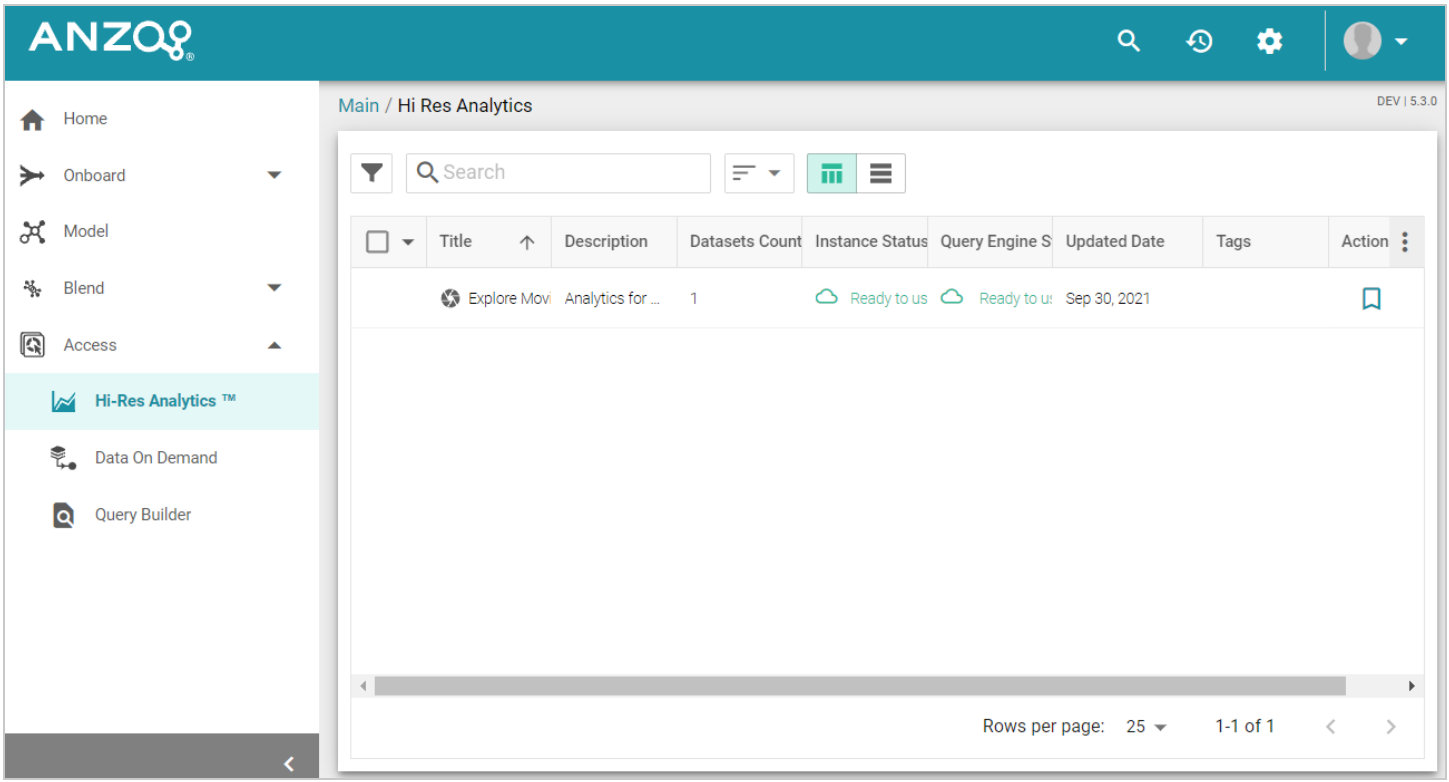
Access

The **Access** menu provides quick access to commonly used features.



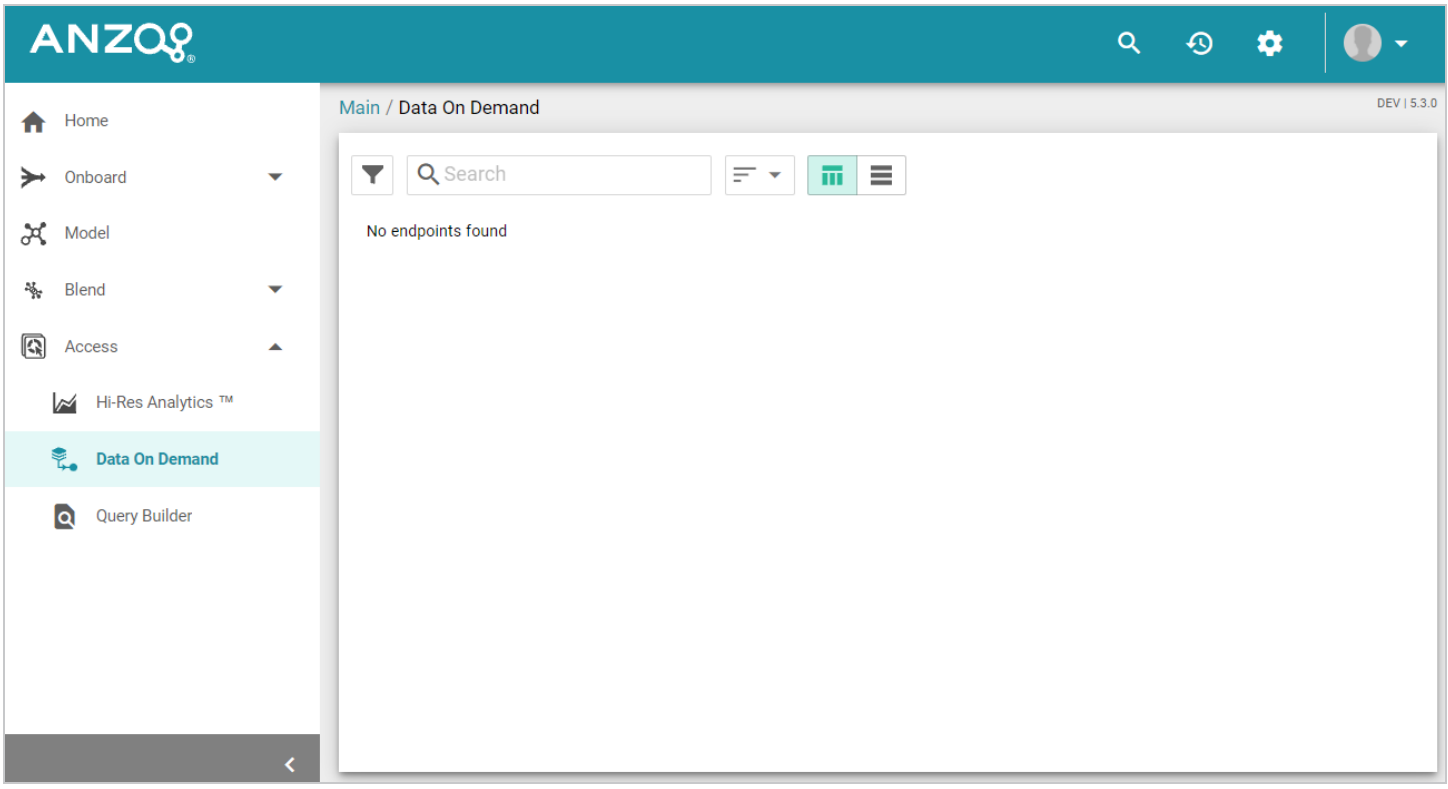
Hi-Res Analytics

The **Hi-Res Analytics** page lists all of the existing dashboards. Users can click a dashboard in the list to open it in the Hi-Res Analytics application. Users can designate favorites and sort the list by tags, graphmarts, and dashboard creators.



Data On Demand

The **Data On Demand** page provides a list of the Data on Demand endpoints that have been created.

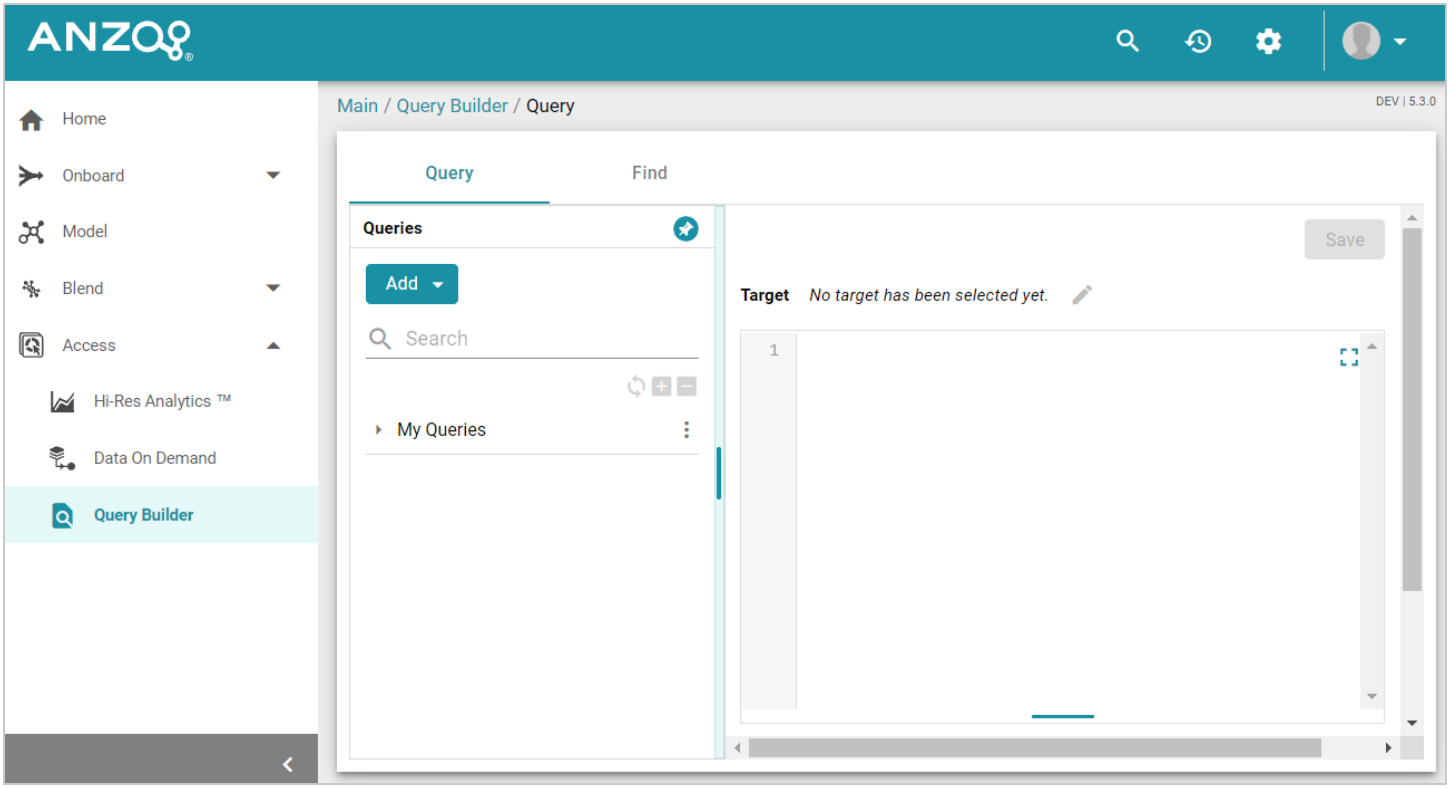


Query Builder

The **Query Builder** provides options to write and run SPARQL queries or find quads.

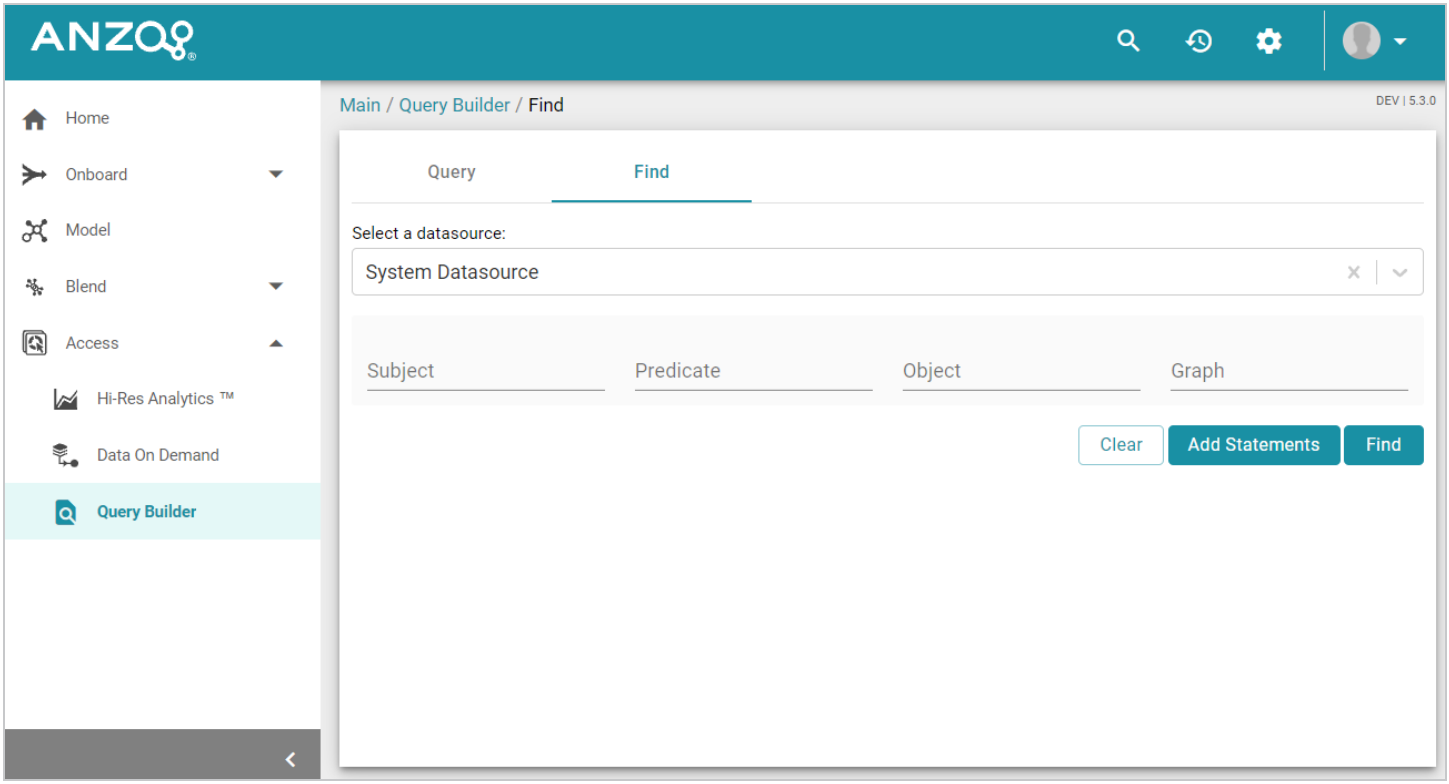
Query

The **Query** editor provides syntax assistance, type-ahead suggestions for model entity names, and automated prefix creation and query formatting for readability. It also includes the option to save queries for later use.



Find

The **Find** page enables users to search for quads by specifying a single subject, object, predicate, or graph name.



Related Topics

Tutorials

Tutorials

There are three ways to onboard structured and semi-structured data to Anzo:

Automated ETL Pipeline Workflow

If the Spark component is installed, you can onboard data using Anzo's built-in pipelines that use an automated workflow and follow the traditional extract, transform, and load (ETL) process. When a pipeline is generated, Anzo automatically creates a data model, mappings, and the ETL jobs required to ingest the source. These Spark ETL pipelines natively support CSV, JSON, XML, SAS, and Parquet files, along with all common database connections.

Automated Direct Data Load Workflow

If Spark is not installed or you do not want to use the ETL pipeline workflow, you can use another automated workflow that follows an extract, load, and transform (ELT) process. In the ELT workflow, data sources are onboarded directly to graphmarts. Data layers with SPARQL queries are automatically generated to transform and blend the data to an analytics-ready knowledge graph. The AnzoGraph Graph Data Interface (GDI) Java plugin (sometimes called the Data Toolkit) is used to connect to the sources, create a model, and generate the data layer queries. The automated direct data load workflow supports all of the data sources that the automated ETL ingestion process supports.

Manual Ingestion with the Graph Data Interface

For advanced users who are familiar with SPARQL, the GDI can also be invoked by writing queries from scratch. The GDI is extremely flexible, allowing you to connect directly to sources via queries and control all aspects of the extract, load, and transform process. In addition to the data sources that the two automated workflows support, you can also ingest raw data and data from HTTP/REST endpoints with manually written SPARQL queries.

This section includes tutorials for helping you get started with the automated ETL and direct data load ingestion methods. For information about the advanced method of writing GDI SPARQL queries for ingestion, see Ingesting Data Manually with the Graph Data Interface in the User Guide.

- Building a Knowledge Graph with the Automated ETL Workflow
- Building a Knowledge Graph with the Direct Data Load Workflow

Building a Knowledge Graph with the Automated ETL Workflow

If the Spark component is installed, you can onboard data using Anzo's built-in, automated ingestion pipelines that follow the traditional extract, transform, and load (ETL) process. When a pipeline is generated, Anzo automatically creates a data model, mappings, and the ETL jobs required to ingest the source data. This tutorial helps get you started with Anzo by teaching you the basic steps to follow when onboarding data using the automated ETL process. The instructions below provide a sample CSV file and guide you through the steps required from importing the file to visualizing the data in a Hi-Res Analytics dashboard.

1. Prepare the Environment
2. Download the Sample Data
3. Create a CSV Data Source and Import the Sample Data
4. Ingest the Imported Data Set
5. Run the ETL Pipeline
6. Create and Activate the Top Movie Graphmart
7. Explore the Data in a Hi-Res Analytics Dashboard
8. Review the New Artifacts

> **Tip**
> This procedure uses the same sample data as the Building a Knowledge Graph with the Direct Data Load Workflow tutorial. If you previously added the Top Movies source and imported the data, you can skip the first three steps and start with Ingest the Imported Data Set.

Prepare the Environment

To give the tutorial continuity and make it straightforward to follow, it helps to make sure that a few general setup steps are completed before you start building the solution. This section provides background information and describes the environment preparation.

This tutorial guides you through importing data from a file, running an ETL pipeline to generate RDF files, creating a Graphmart, and loading the data into AnzoGraph. Completing those steps requires a designated location (Anzo Data Store) on the shared File Store to place the RDF files that the ETL process generates and a connection to AnzoGraph for loading the data.

Confirm that the following setup steps are complete before proceeding:

1. An Anzo Data Store has been created for the RDF files that are generated during the ETL process. For instructions on creating a Data Store, see Creating an Anzo Data Store in the Administration Guide.
2. Anzo has a connection to a static AnzoGraph instance or a Cloud Location is set up for dynamic Kubernetes-based AnzoGraph deployments. For instructions on configuring the connection to a static instance, see Connecting to AnzoGraph in the Administration Guide. For information about Kubernetes deployments, see Using K8s for Dynamic Deployments of Anzo Components in the Deployment Guide.

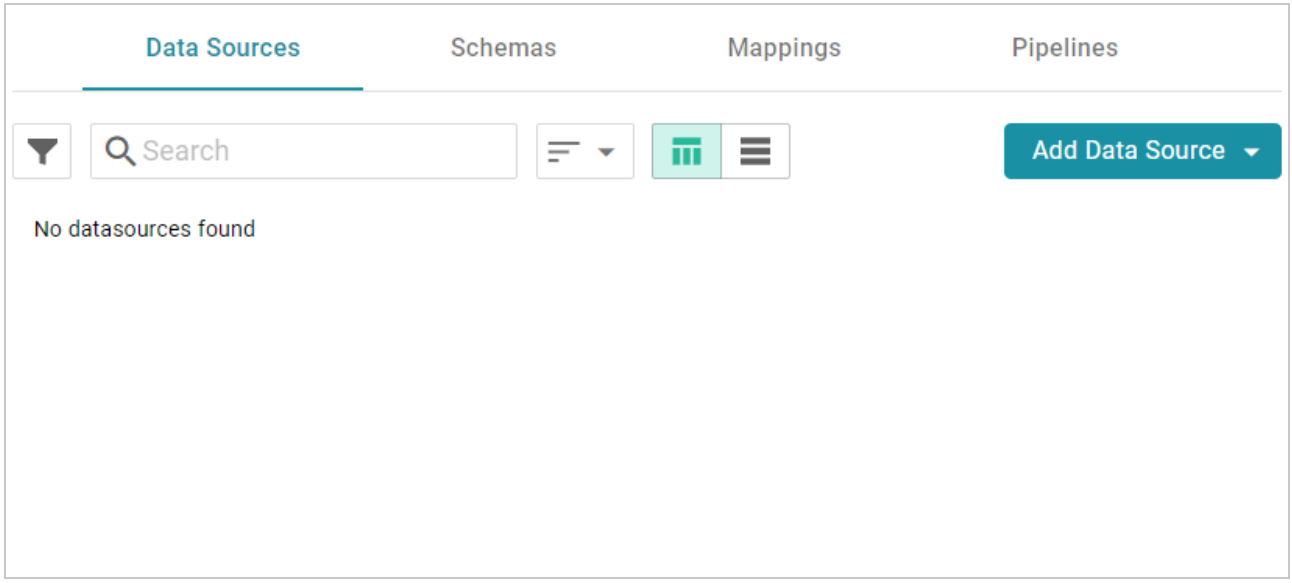
Download the Sample Data

This tutorial uses a small sample data set from Kaggle. The data is in CSV format and includes information about the top 200 movies according to IMDB members. Click the link below to download Top_200_Movies.zip to your computer. Then extract the ZIP file to make **Top_200_Movies.csv** available for importing.

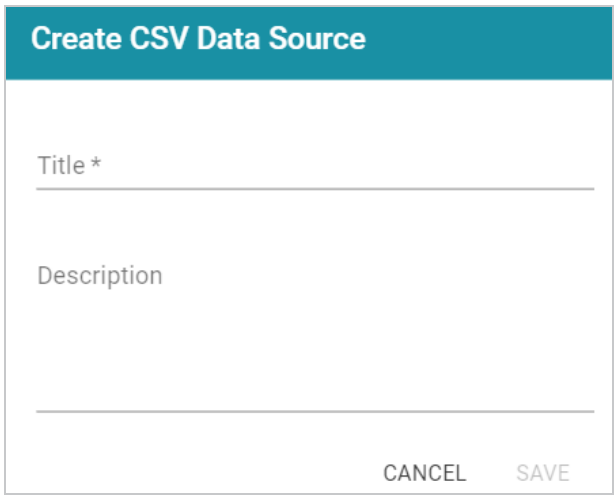
Top_200_Movies.zip

Create a CSV Data Source and Import the Sample Data

1. In the Anzo application, expand the **Onboard** menu and click **Structured Data**. Anzo displays the Data Sources screen, which lists any existing data sources:

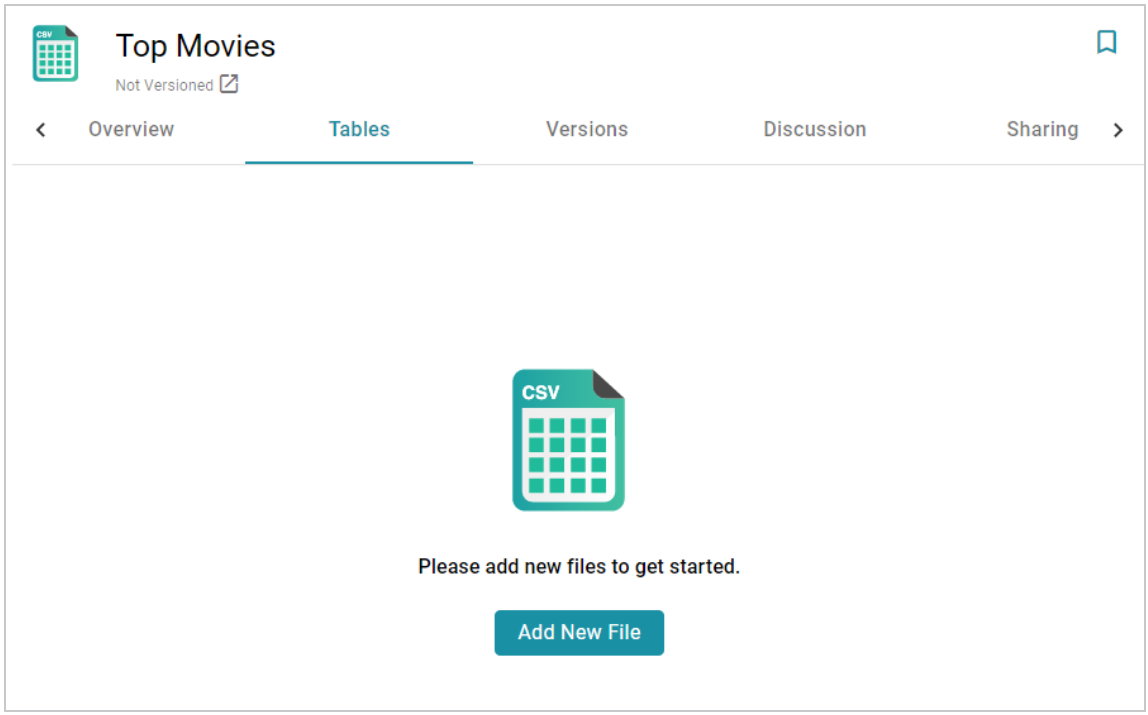


2. Click the **Add Data Source** button and select **File Data Source** > **CSV Data Source**. Anzo opens the Create CSV Data Source screen.

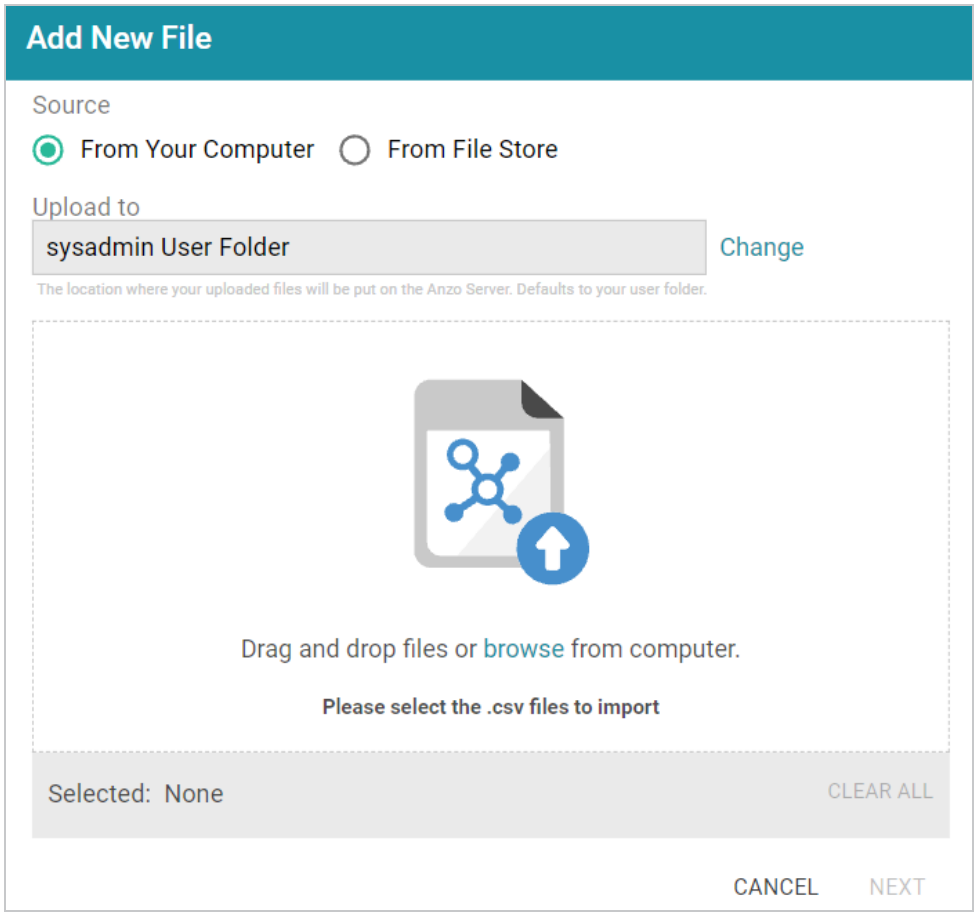


3. Type a name for the source in the **Title** field. For example, "Top Movies." Then type an optional description in the **Description** field. For example:

4. Click **Save**. Anzo saves the source and displays the Tables tab for the new Data Source.

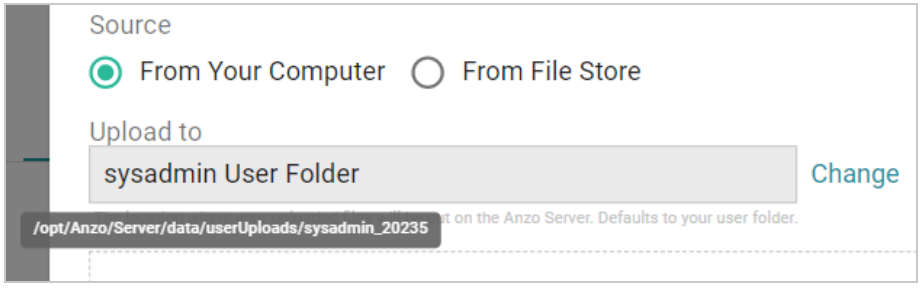


5. Click the **Add New File** button. Anzo displays the Add New File dialog box, and the Source is set to **From Your Computer**.

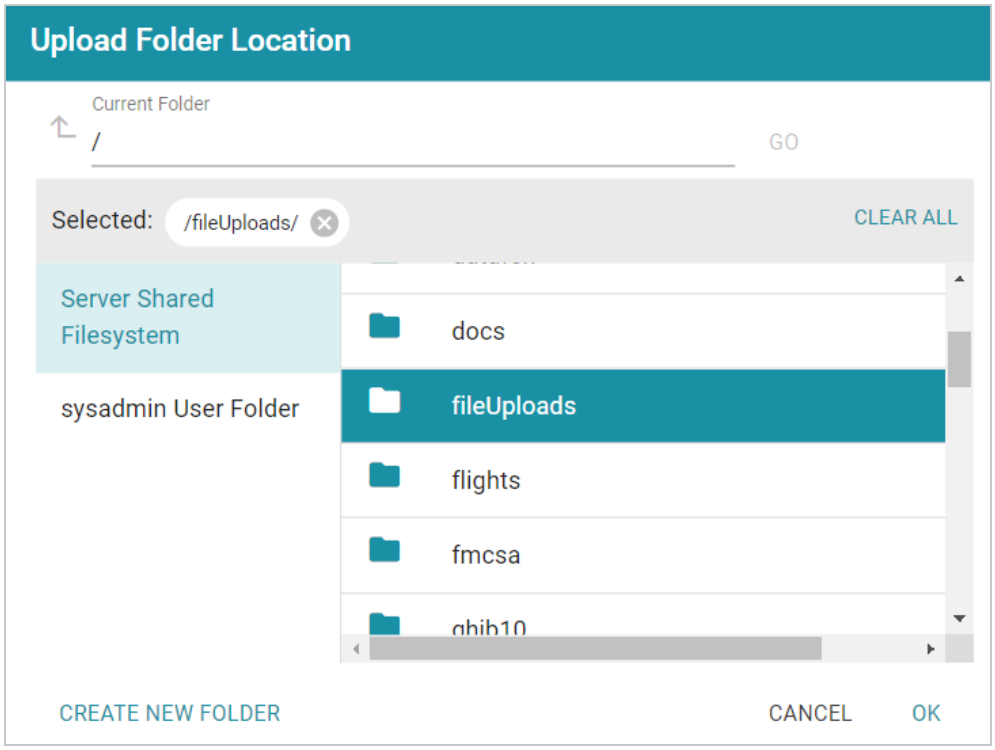


6. As a best practice, check the upload location that is listed in the **Upload To** field by hovering your pointer over the value to view the tooltip. Make sure the upload location is a directory on the shared File Store and not a location in the server installation path. If the file is not uploaded to the shared File Store it is not accessible by applications like AnzoGraph or Spark. In addition, other users cannot publish pipelines for the Data Source because they typically do not have access to the file.

   For example, viewing the Upload To location for the screen above shows that the file will be uploaded to the server installation path, `/opt/Anzo/Server/data...`

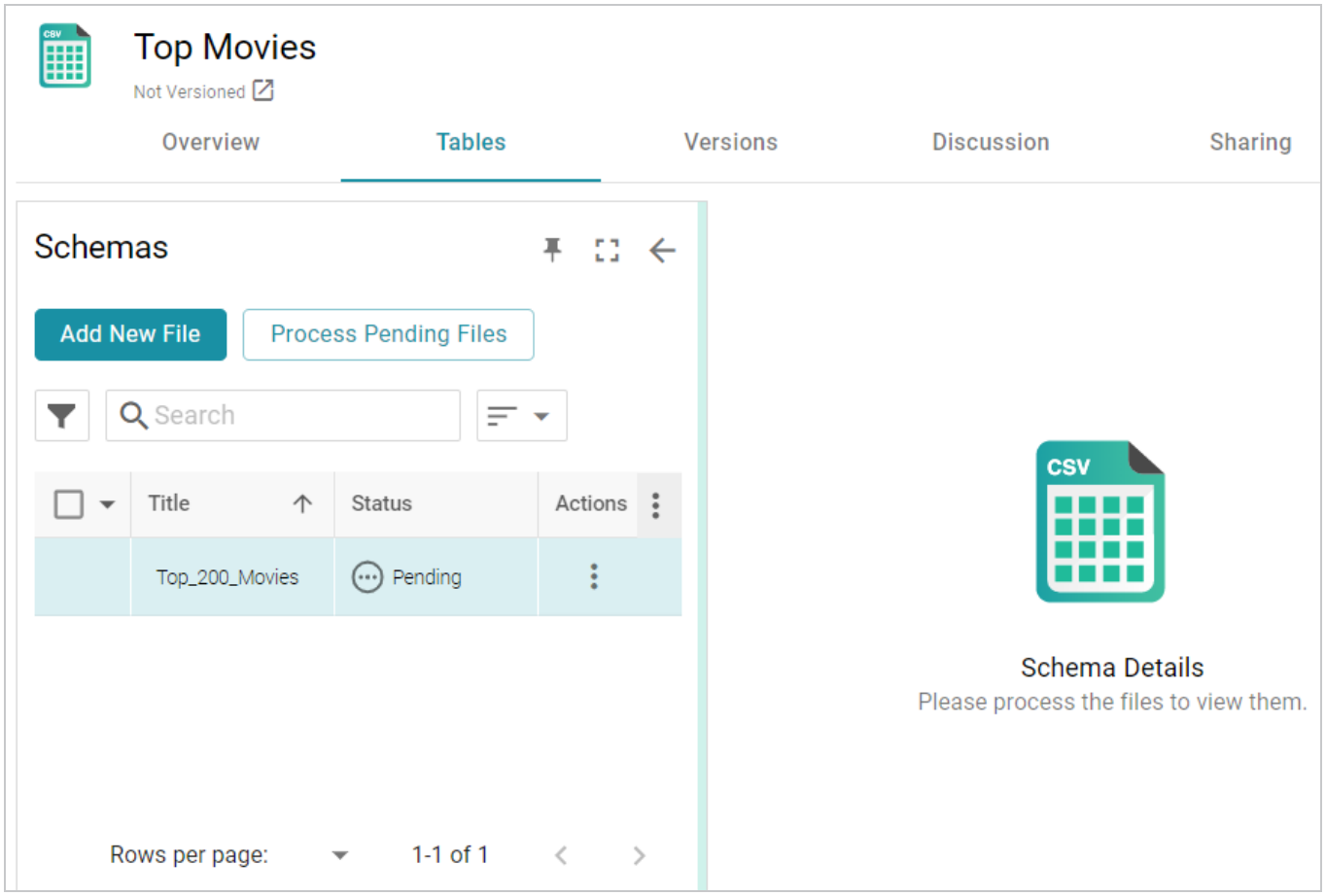


If your Upload To location is configured to upload the file to the server installation path, it is important to click **Change** and select another upload location that is on the shared File Store. For example, the image below shows the Upload Folder Location dialog box that is presented after clicking Change. A folder called **fileUploads** is selected on the File Store.

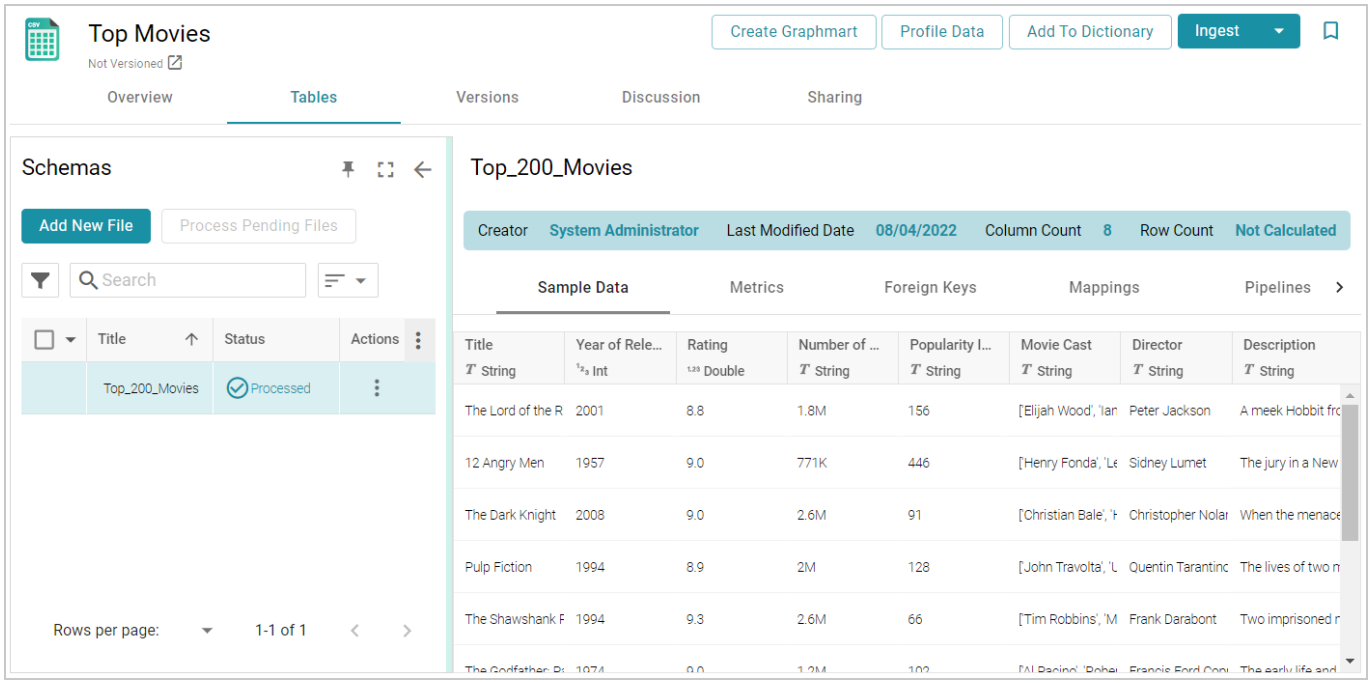


> **Tip**
> An administrator can configure the base upload path so that it points to a location on the File Store by default. For more information, see Setting the Default Base File Store Path for File Uploads in the Administration Guide.

7. Next, drag and drop the Top_200_Movies.csv file onto the screen or click **browse** and navigate to the file and double-click it to select it. Anzo attaches the file and the **Next** button becomes active. Click **Next**. Anzo lists the file on the left side of the screen with a status of Pending. For example:



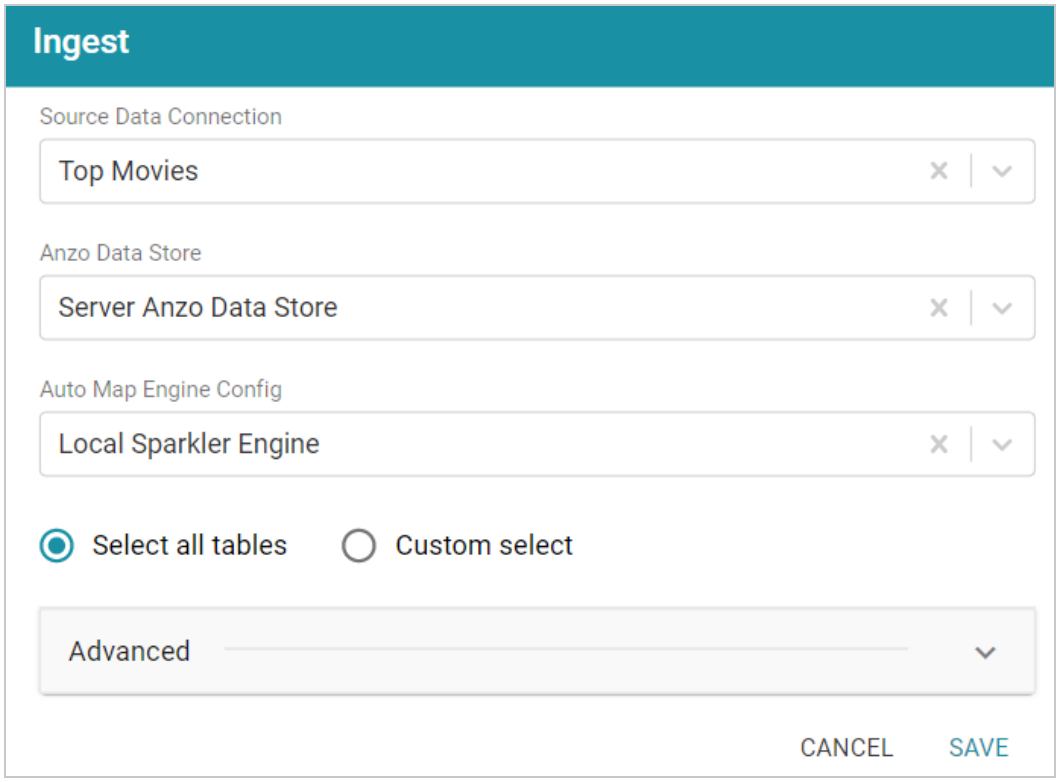
8. Click the **Process Pending Files** button to import the data to Anzo. Anzo imports the data and the status changes to **Processed**. Additional buttons are also displayed on the right side of the screen. You can click the table row for Top_200_Movies to display the schema details.



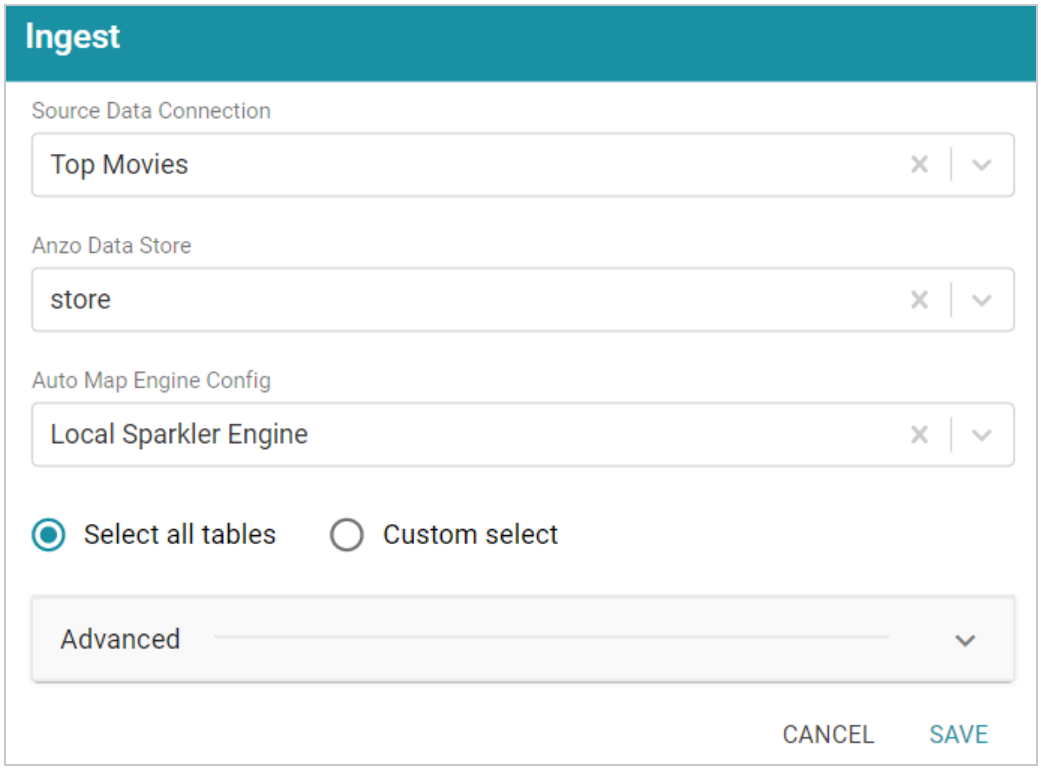
The data is now imported. Proceed to the next step to ingest the data and automatically derive a model, generate a mapping, and create an ETL pipeline for transforming the data to the RDF graph model.

Ingest the Imported Data Set

1. On the Tables screen (from the last step of the procedure above), click the **Ingest** button. Anzo opens the Ingest dialog box and automatically populates the Source Data Connection value. The Anzo Data Store field is also populated with the default data store value, and the Auto Map Engine Config field is populated with the default ETL Engine value. For example:



2. If necessary, click the **Anzo Data Store** field and select the data store for this pipeline. Make sure it is a location on the file store that AnzoGraph has access to.
3. If necessary, click the **Auto Map Engine Config** field and select the ETL Engine to use. For this tutorial, select **Local Sparkler Engine**. That means the pipeline will use the SPARQL-driven Sparkler ETL compiler to process the source data for the Spark engine.
4. Leave the **Select all tables** radio button selected. For example, the image below shows the completed screen:

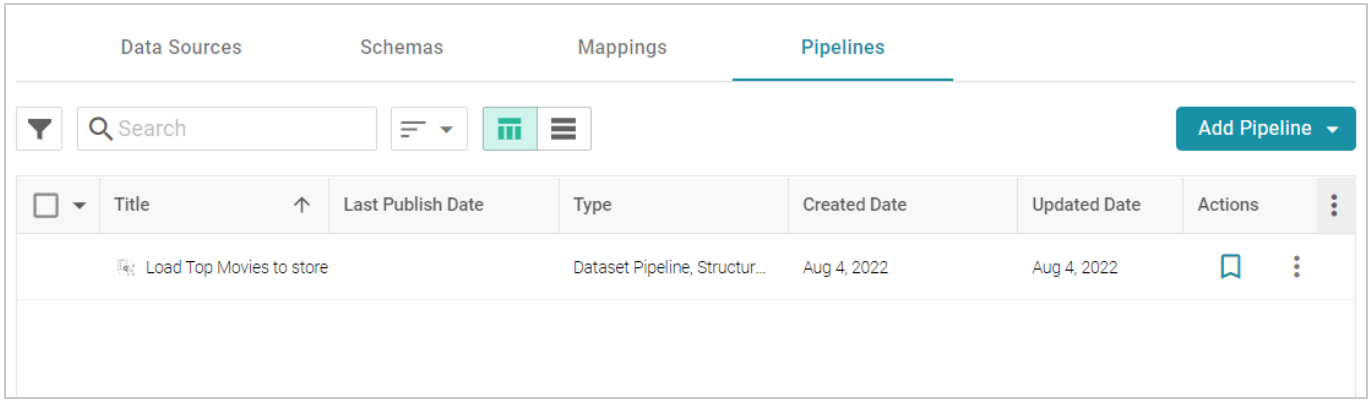


5. Click **Save**. Anzo creates a pipeline for the data source and displays a confirmation message that a pipeline was created.

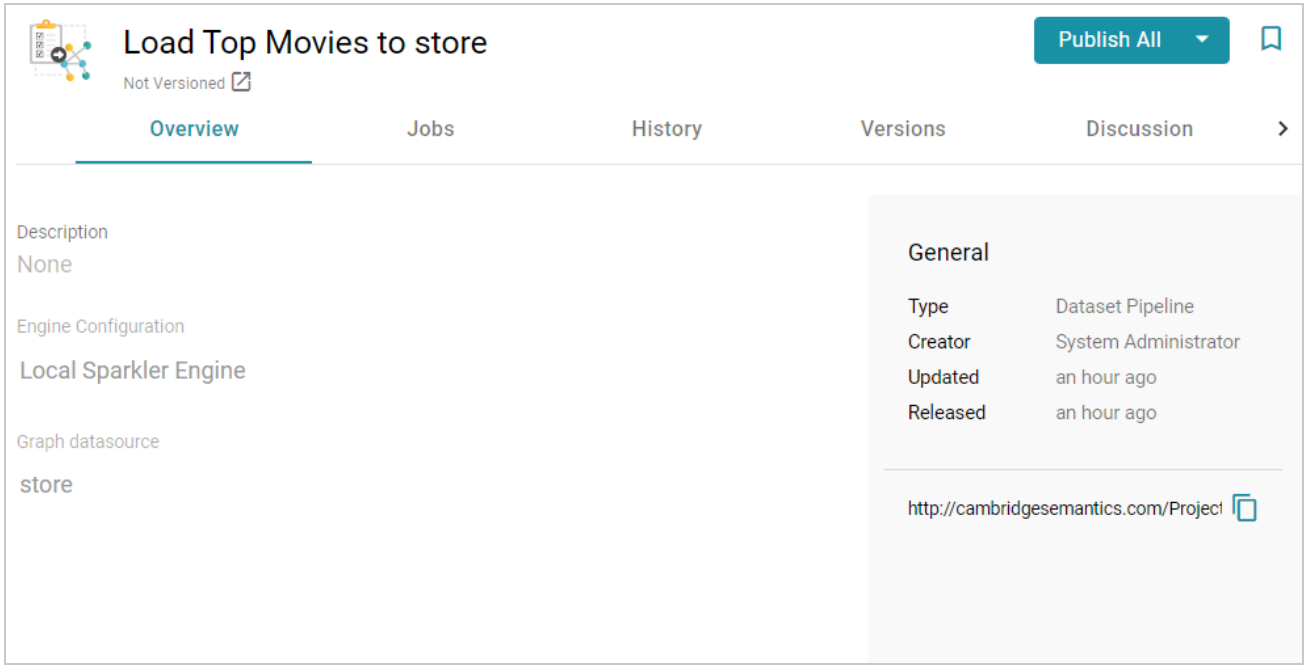
As part of creating the pipeline, Anzo also generates a model and mapping for the sample data set. The last step in this tutorial (Review the New Artifacts) describes each of the artifacts. Proceed to the next step to run the pipeline and generate the RDF File-Based Linked Data Set (FLDS).

Run the ETL Pipeline

1. In the Onboard menu, click **Structured Data**. Then click the **Pipelines** tab. Anzo displays the Pipelines screen, which lists the pipeline for the sample data set. For example:



2. Click the title to view the pipeline. For example:

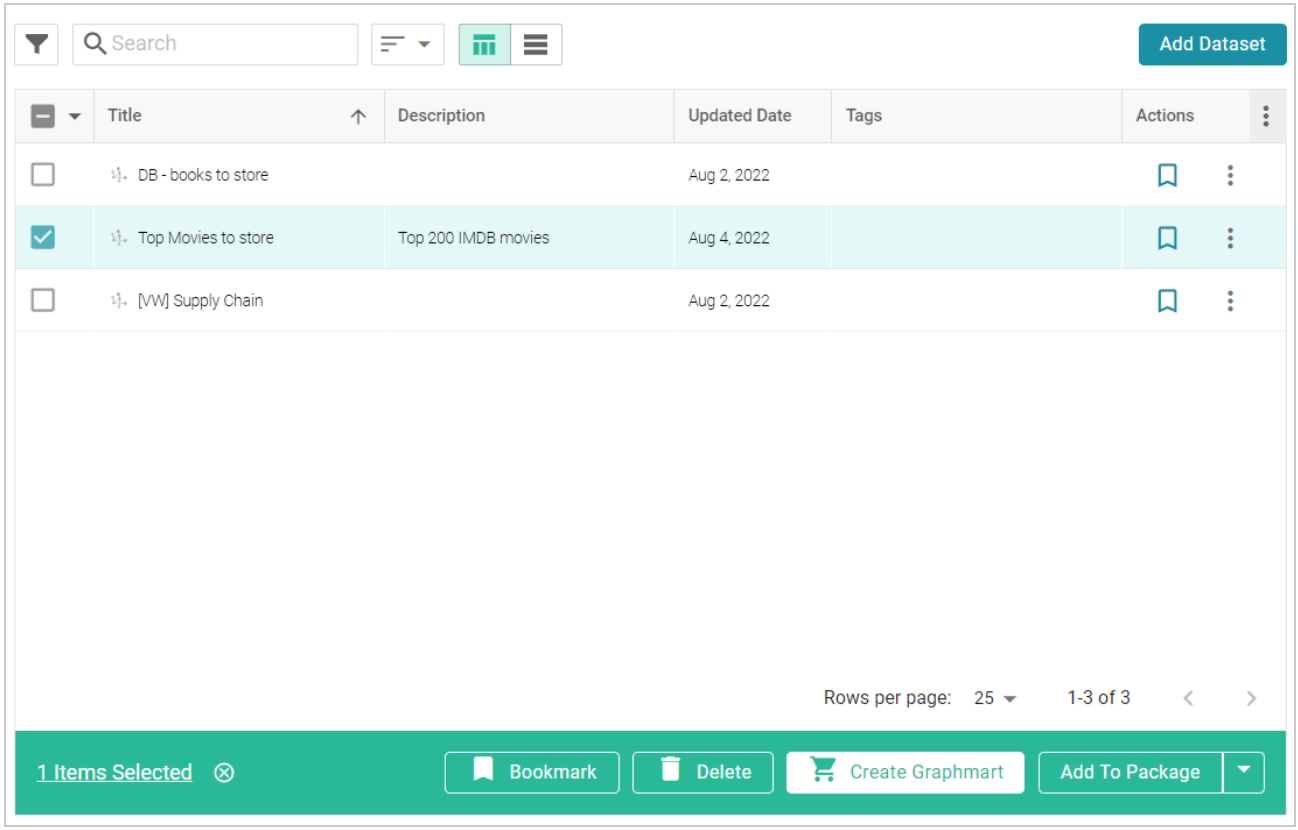


3. Click the **Publish All** button to run all the jobs in the pipeline and generate the FLDS in the specified data store.

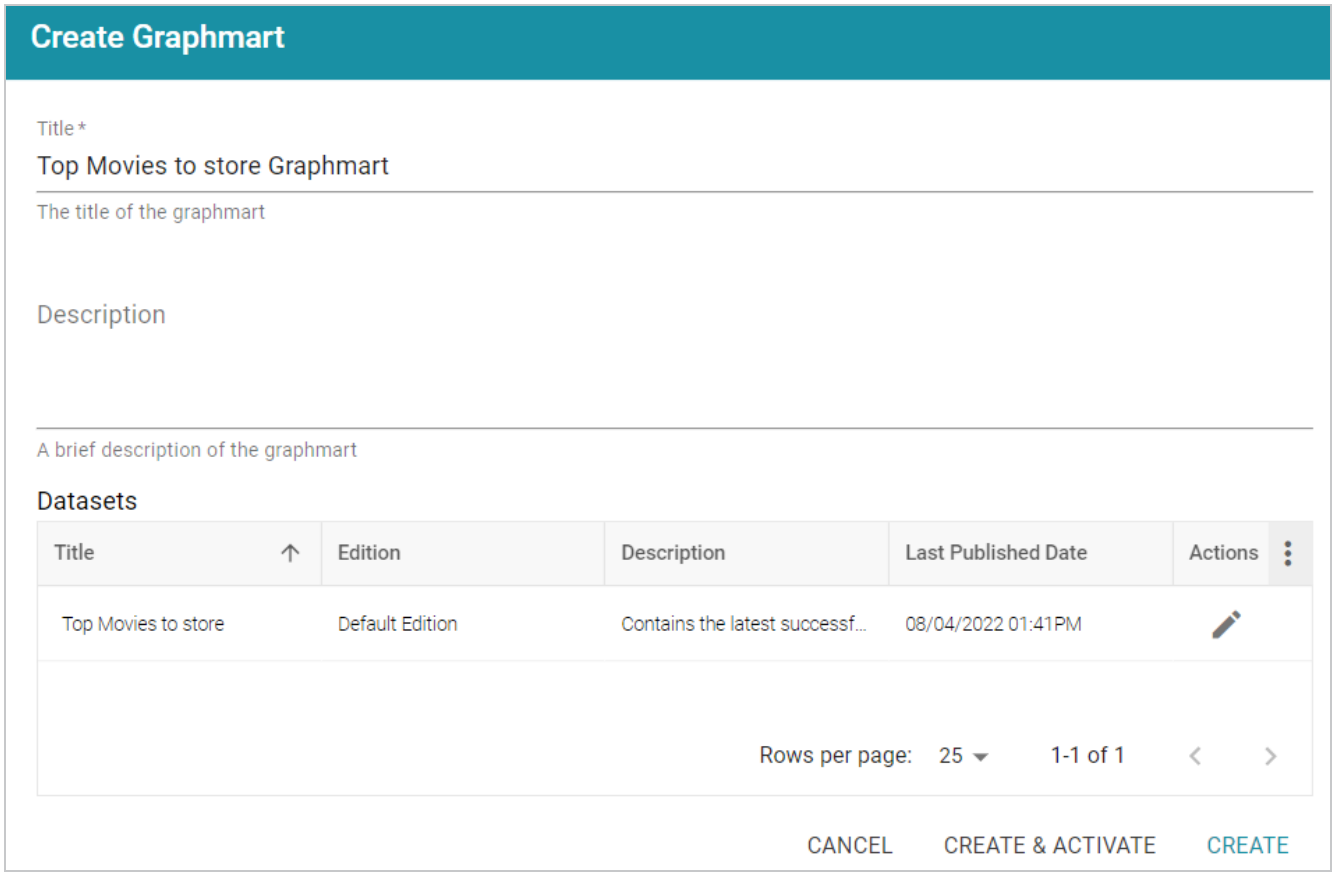
Anzo publishes the ETL files and displays confirmation messages as each step in the process completes. When the job execution is finished, the Top Movies data set becomes available in the **Dataset** catalog. Proceed to the next step to create a Graphmart for the new data set and load it into AnzoGraph.

Create and Activate the Top Movie Graphmart

1. Expand the **Blend** menu and click **Datasets**. Anzo displays the Datasets catalog, which shows an inventory of all of the onboarded data in Anzo. It lists the Top Movies data set.
2. Hover the pointer over the Top Movies item to display a checkbox in the left column, and then select the checkbox. Additional buttons become available at the bottom of the screen. For example:

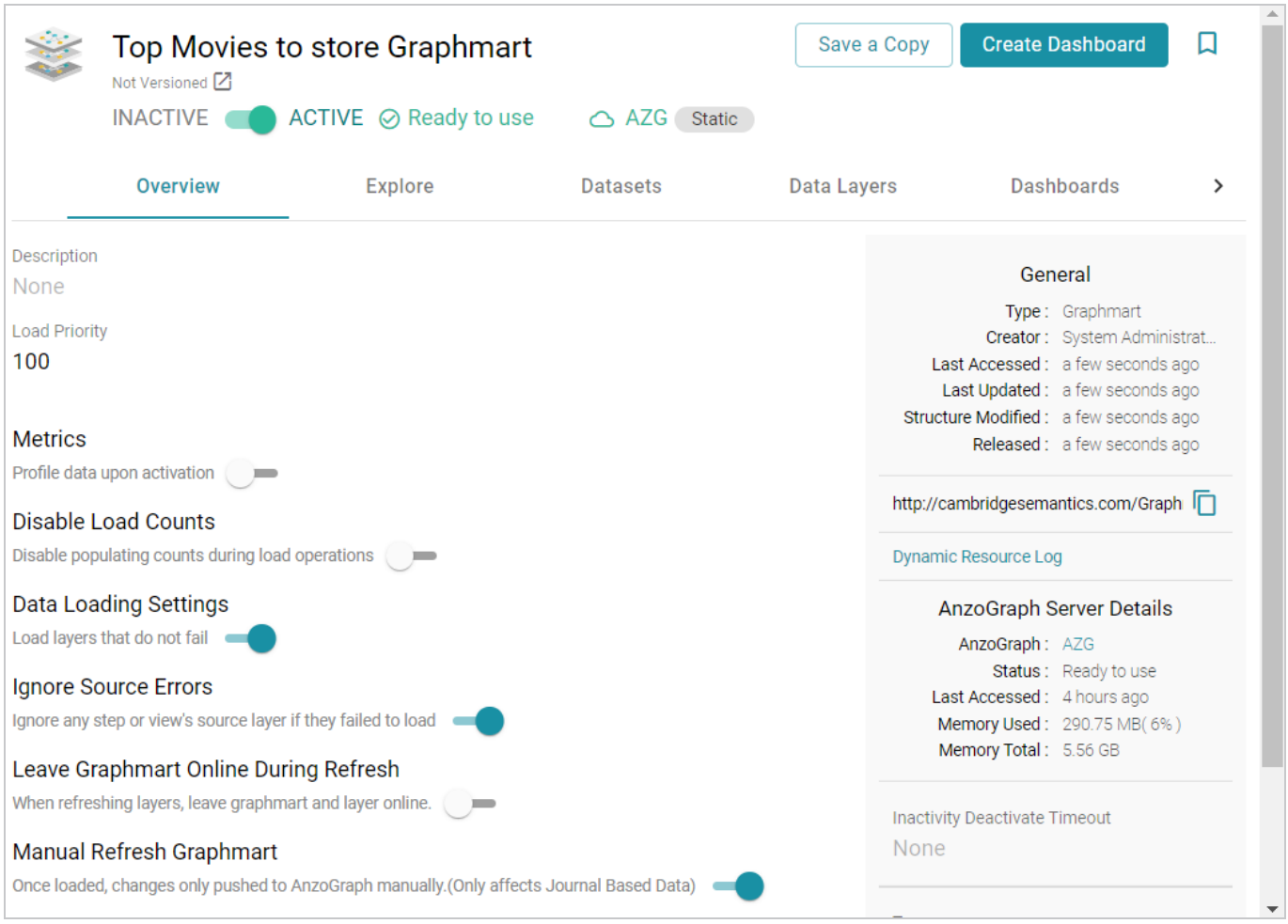


3. Click the **Create Graphmart** button. Anzo displays the Create Graphmart screen. Anzo populates the **Title** field by appending "Graphmart" to the data set name. For example:



> **Tip**
>
> By default, the current working edition of the data set, called the Default Edition, is selected for the Graphmart. The Default Edition is managed by Anzo (the Managed Edition) and always contains the Data Components from the last successful run of the pipeline. Since this is a new source and the first run of the pipeline, no other editions exist. To learn more about Dataset Editions, see Managing Dataset Editions in the User Guide.

4. On the Create Graphmart screen, you have the option edit the Title and type an optional Description. When you are ready to create the graphmart and load the data to AnzoGraph, click **Create & Activate** at the bottom of the screen. Anzo creates and activates the graphmart and displays the Overview screen for the new graphmart. For example:

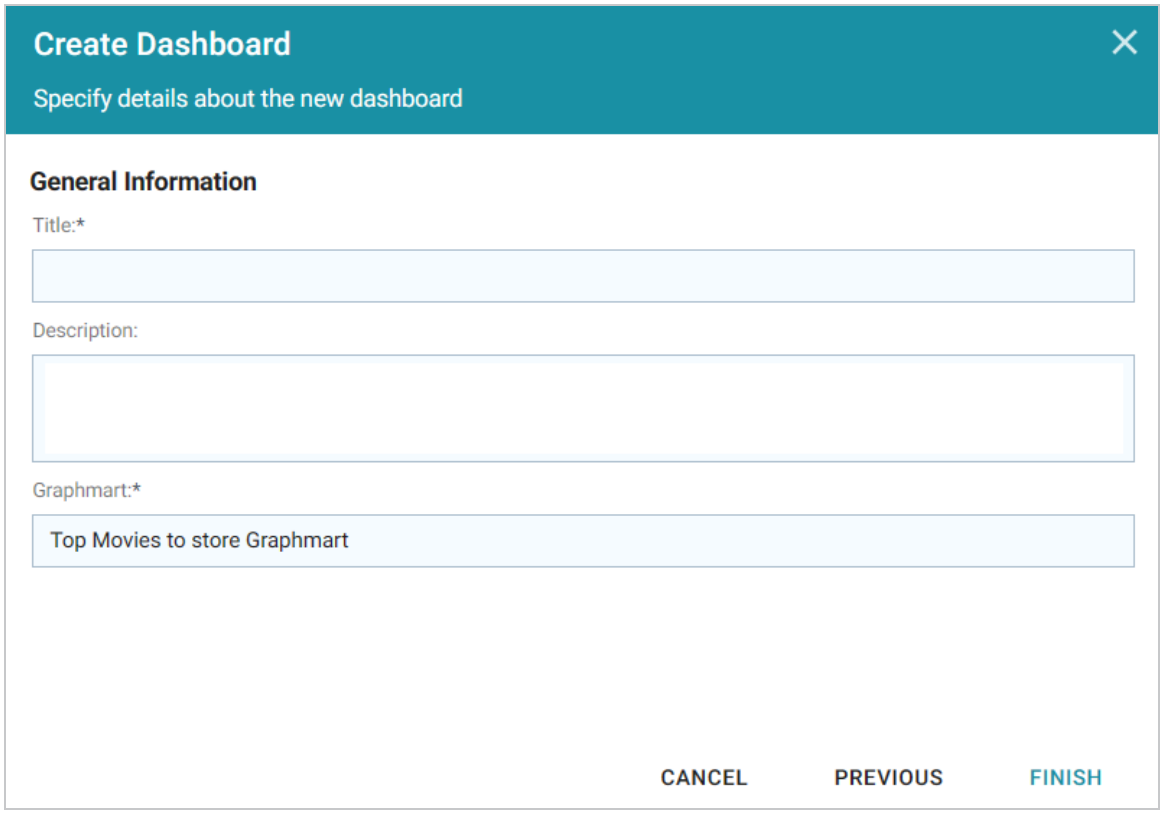


> **Tip**
>
> If you want to learn about the graphmart settings that are shown in the image above, see Graphmart Settings Reference in the User Guide.

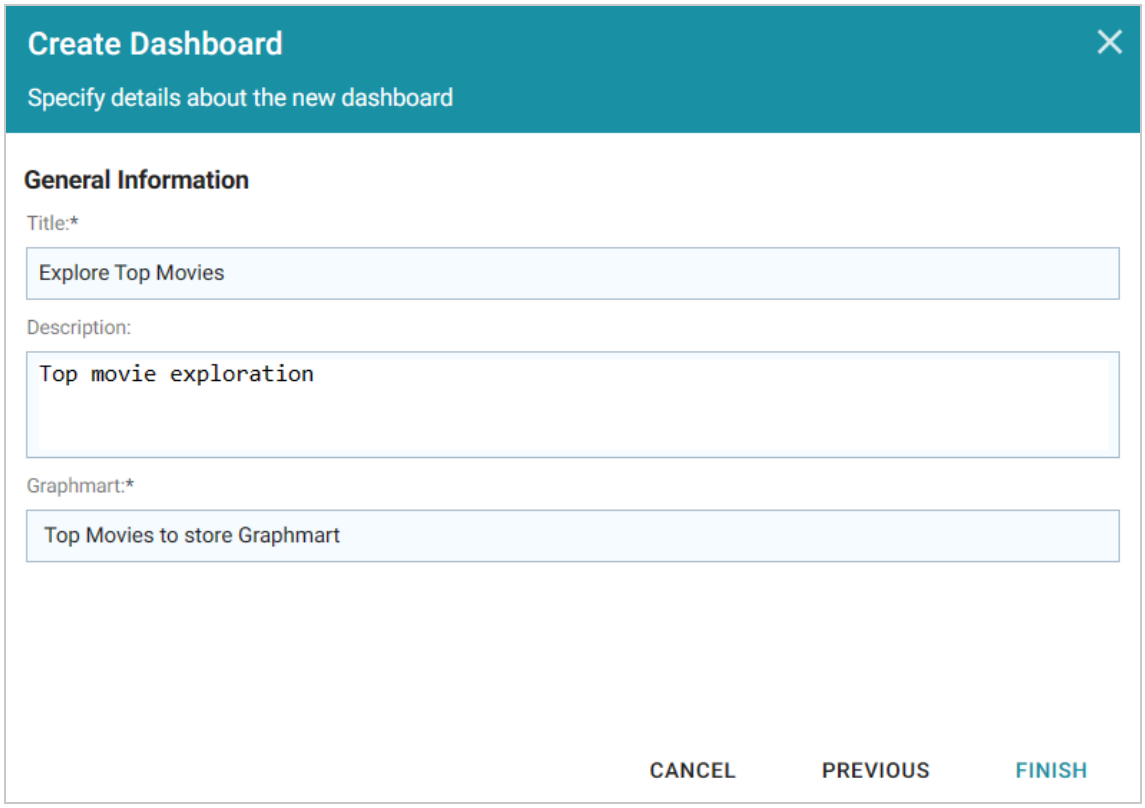
Now that the data is loaded into AnzoGraph, it is available to explore and analyze in dashboards. Proceed to the next step to explore the sample data in the Hi-Res Analytics application.

Explore the Data in a Hi-Res Analytics Dashboard

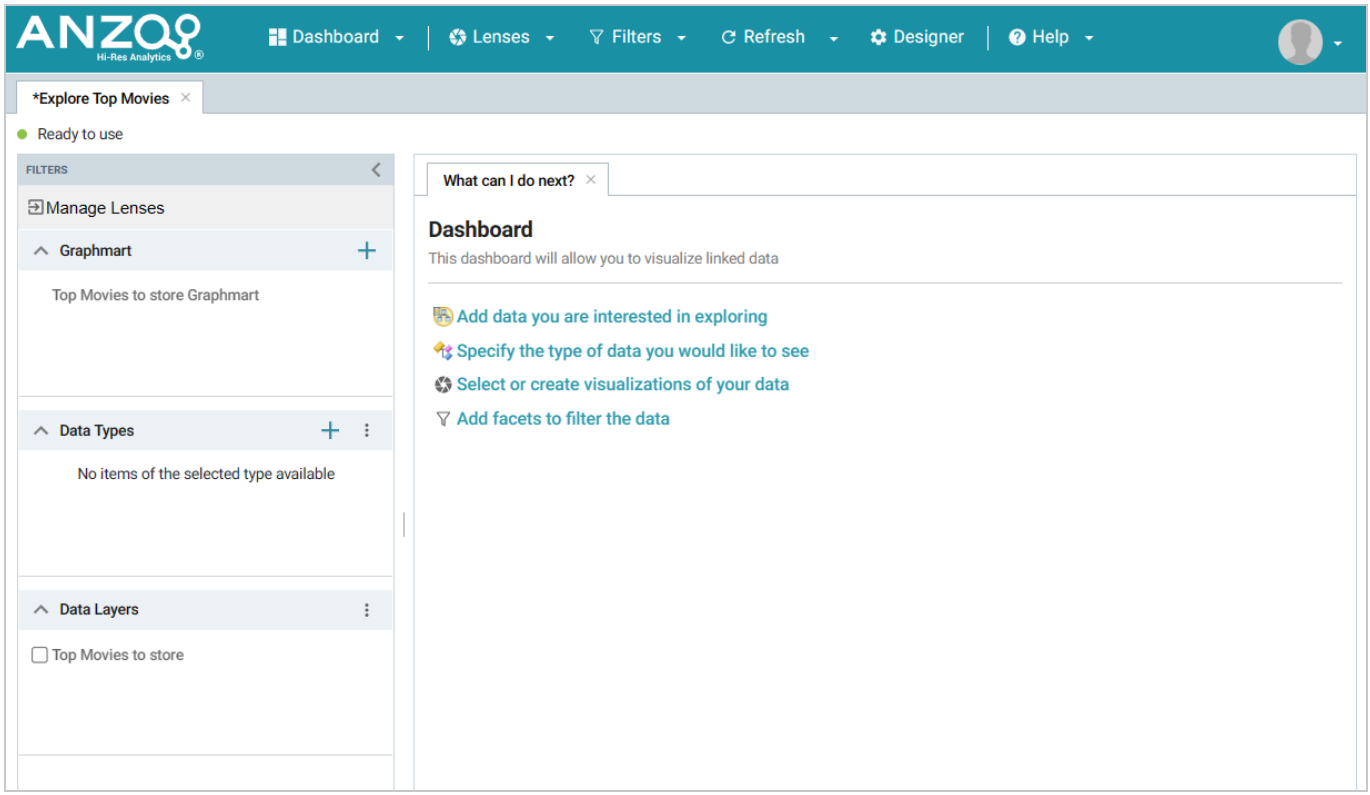
1. On the Graphmart screen (as shown in the last step above), click the **Create Dashboard** button. Anzo opens the Hi-Res Analytics application and displays the Create Dashboard dialog box.
2. Select **Graphmart Dashboard** and click **Next**. The Create Dashboard dialog box is displayed:



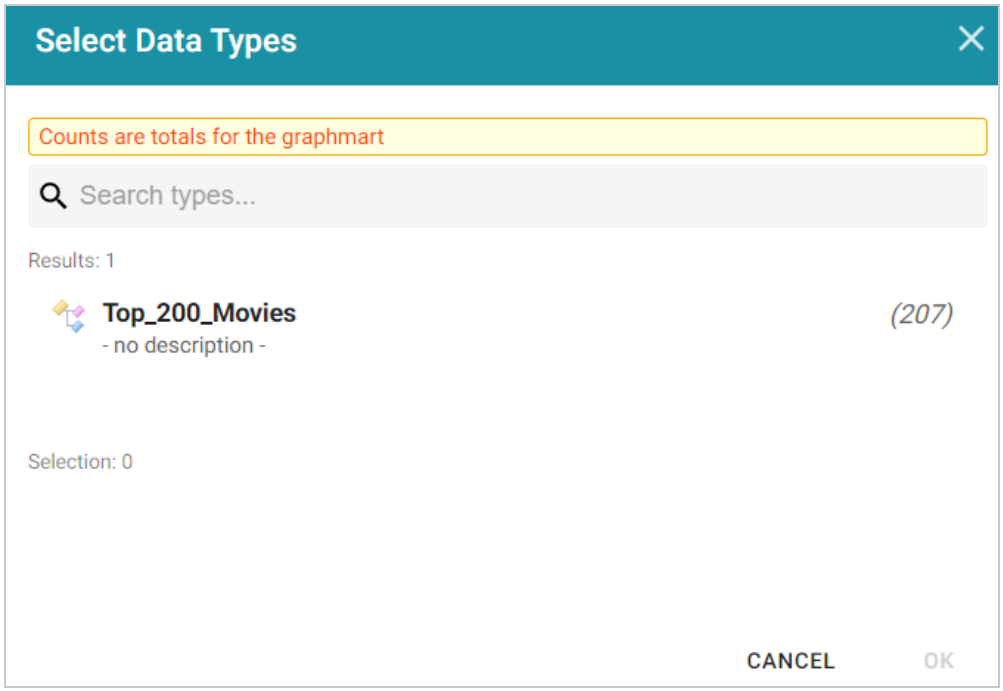
3. On the Create Dashboard screen, specify a **Title** for the dashboard and add an optional **Description**. Leave the default **Graphmart dashboard** value in the **Type** field. For example:



4. Click **Finish** to create the dashboard. The new Dashboard appears as a new tab on the screen and contains a sub-tab titled **What can I do next?**. This tab acts as a wizard to guide you through the initial Dashboard creation. Anzo populates the Graphmart and Data Layers panels with the sample movie data Graphmart name and default Data Layer for the Graphmart.



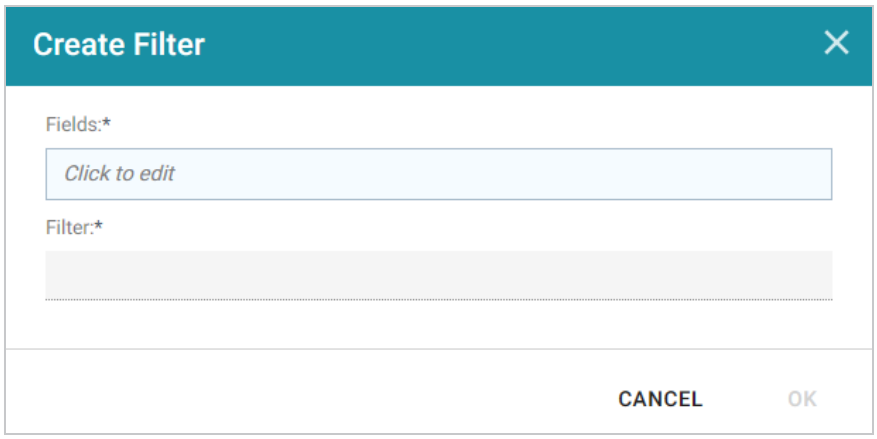
5. To define the type of data to display on the dashboard, click the **Specify the type of data you would like to see** link on the What can I do next tab. The Select Data Types dialog box appears and displays the available data types. Since the sample data set has one type or class of data, only **Top_200_Movies** is listed. The value in parentheses shows the total number of instances of that type.



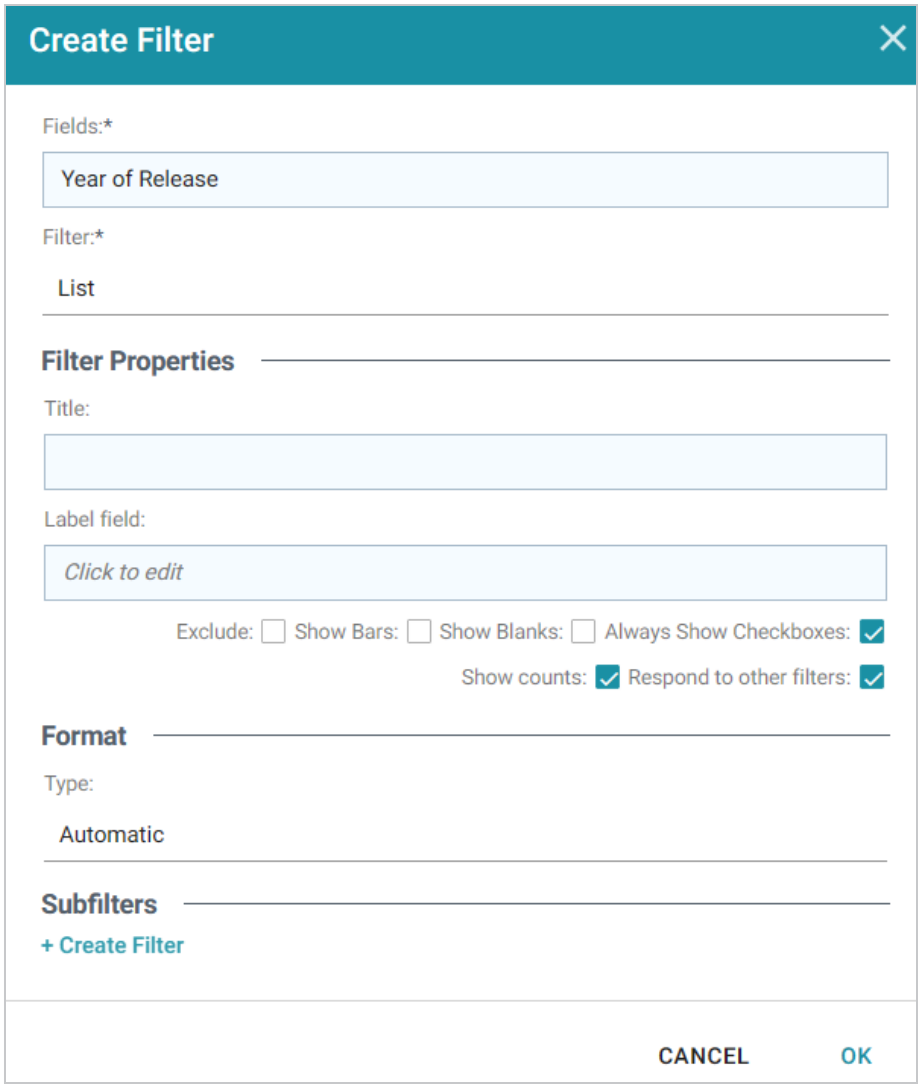
6. Click **Top_200_Movies** to select it, and then click **OK** to add the data type to the dashboard. The selected type is added to the Data Types panel.
7. In the main Hi-Res Analytics toolbar, click the **Dashboard** button and select **Save** to save the Dashboard.

> **Tip**
> To start to explore the data and see what values exist for the properties in the Top_200_ Movies class, it can help to add a filter to the dashboard. Filters reveal the values associated with properties. Learning more about the values enables you to start making decisions about what properties to group on, for example, what properties have relationships, and what results you want to visualize.

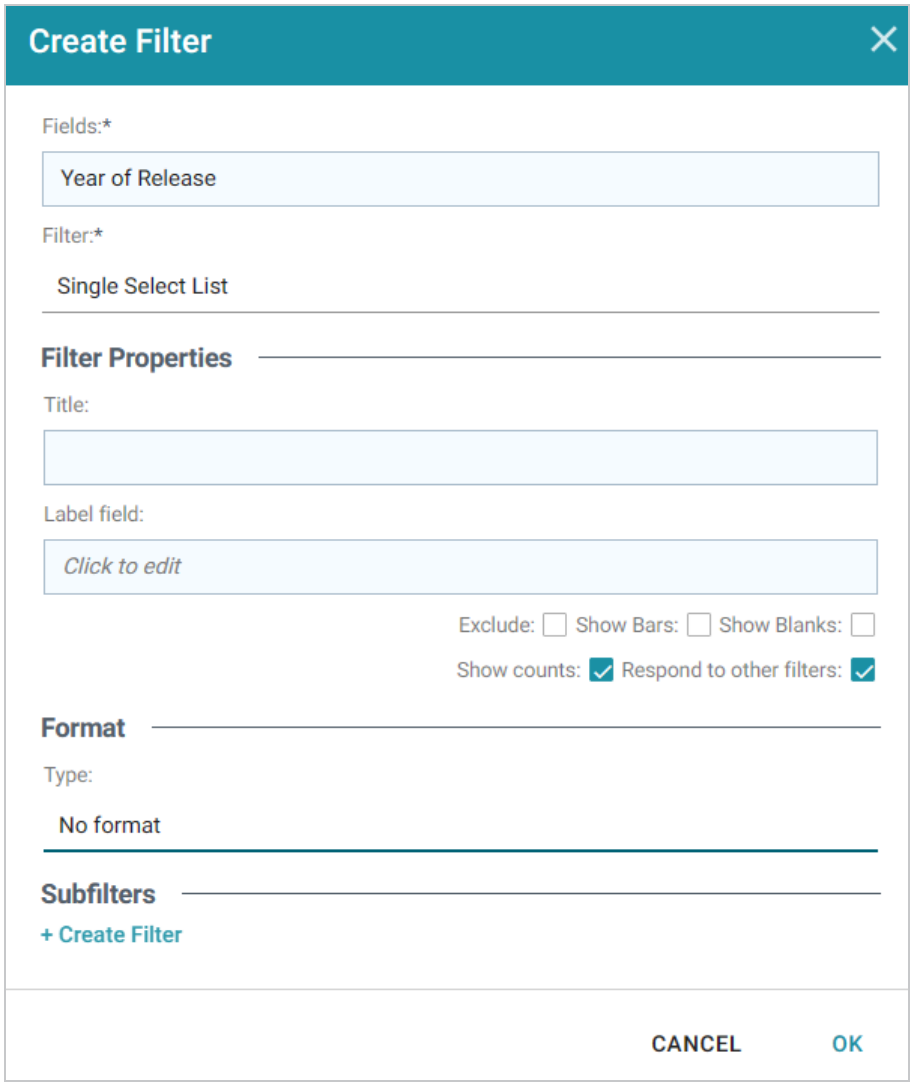
8. To create a filter, click the **Add facets to filter the data** link on the What can I do next tab. Anzo displays the Create Filter dialog box:



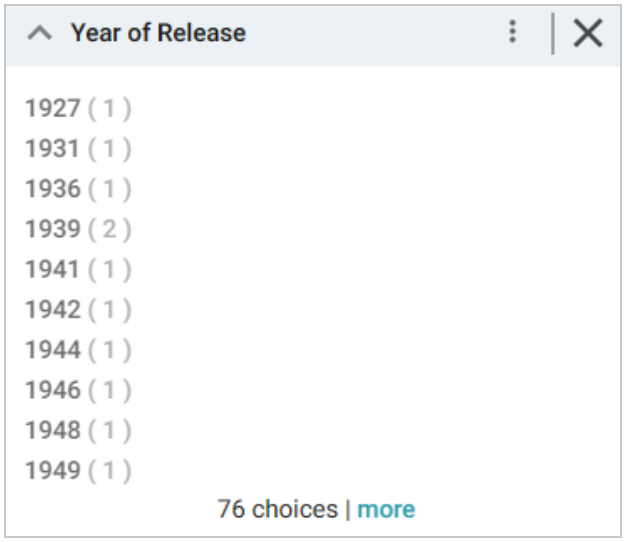
9. In the Create Filter dialog box, click the **Fields** field and browse the available properties. For this tutorial, select **Year of Release**, and then click **Close** to close the Fields drop-down list. Additional options are displayed:



10. Click the **Filter** drop-down list and browse the available filter types. For this tutorial, select **Single Select List**.
11. At the bottom of the screen under **Format**, click the **Type** drop-down list and select **No format**. Removing formatting from the filter ensures that commas are not displayed for the Year values. The image below shows the completed dialog box.

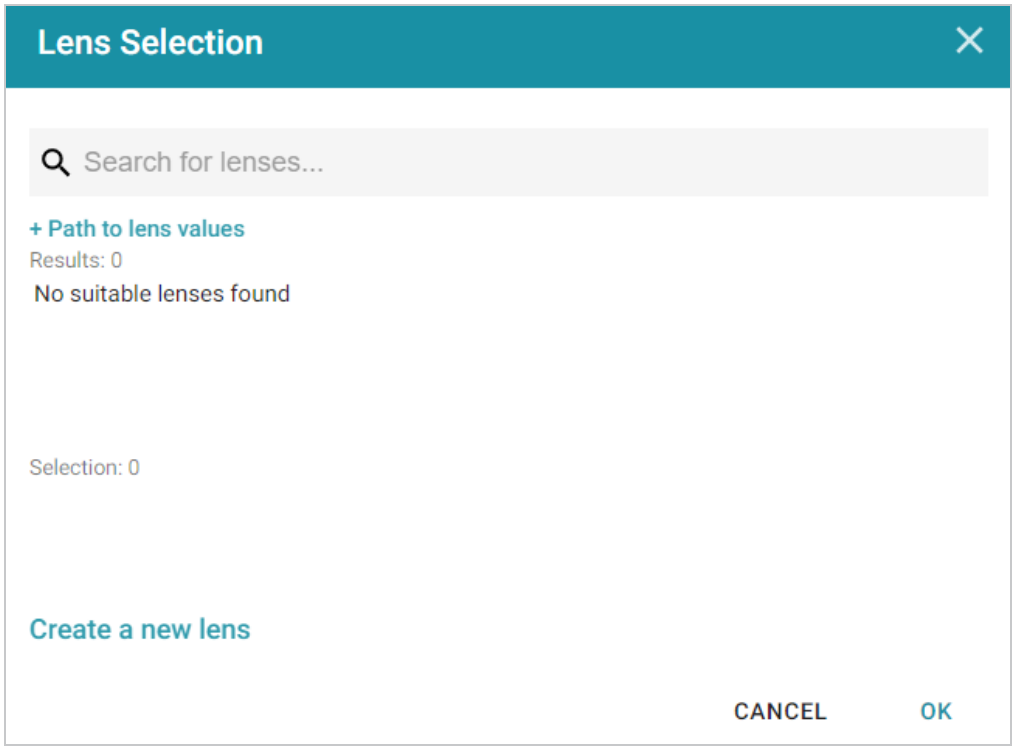


12. Click **OK** to close the dialog box and add the new filter to the dashboard. Anzo adds the Year of Release filter to the left panel. The new filter will enable you to narrow the scope of the data when it is visualized.

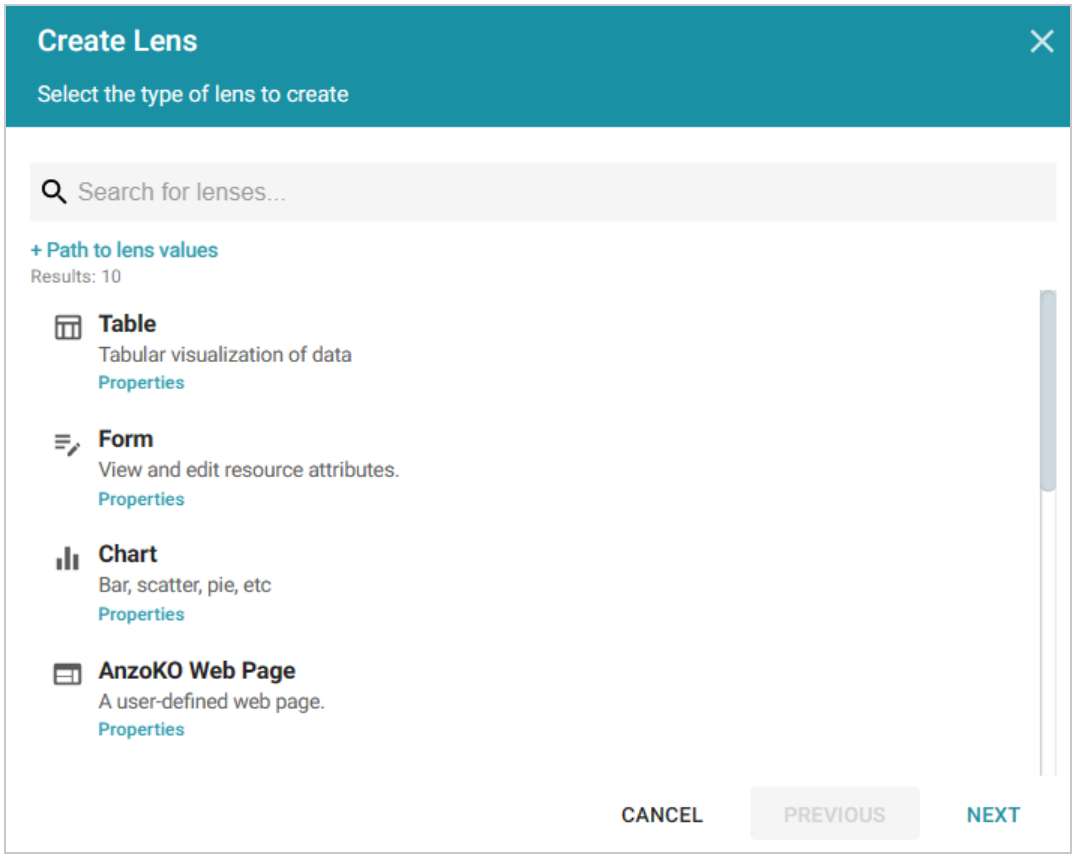


> **Tip**
> Once you have a good understanding of the values and relationships that exist in the data set, you can experiment with lenses and decide on the most appropriate way to display the data. Creating a Table lens is a quick way to get started with learning about the values and relationships in the data and to see how the new filter works.

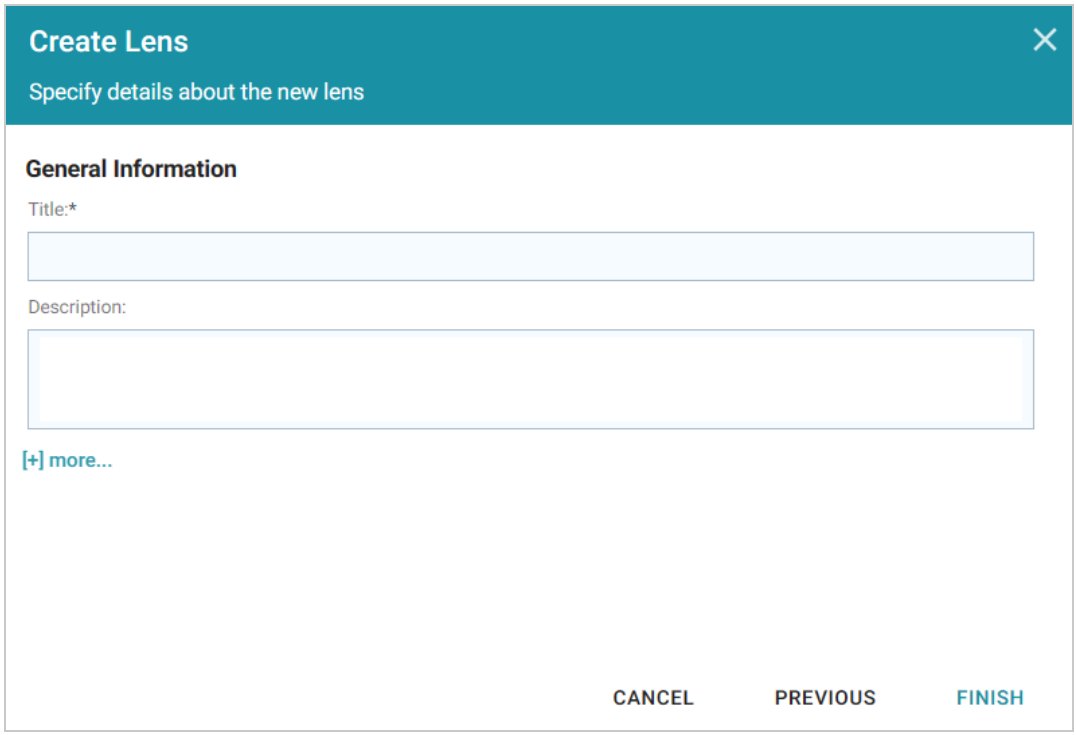
13. To create a Table lens, click the **Select or create visualizations of your data** link on the What can I do next tab. Anzo displays the Lens Selection dialog box. Since lenses have not been created previously, the list of lenses is empty:



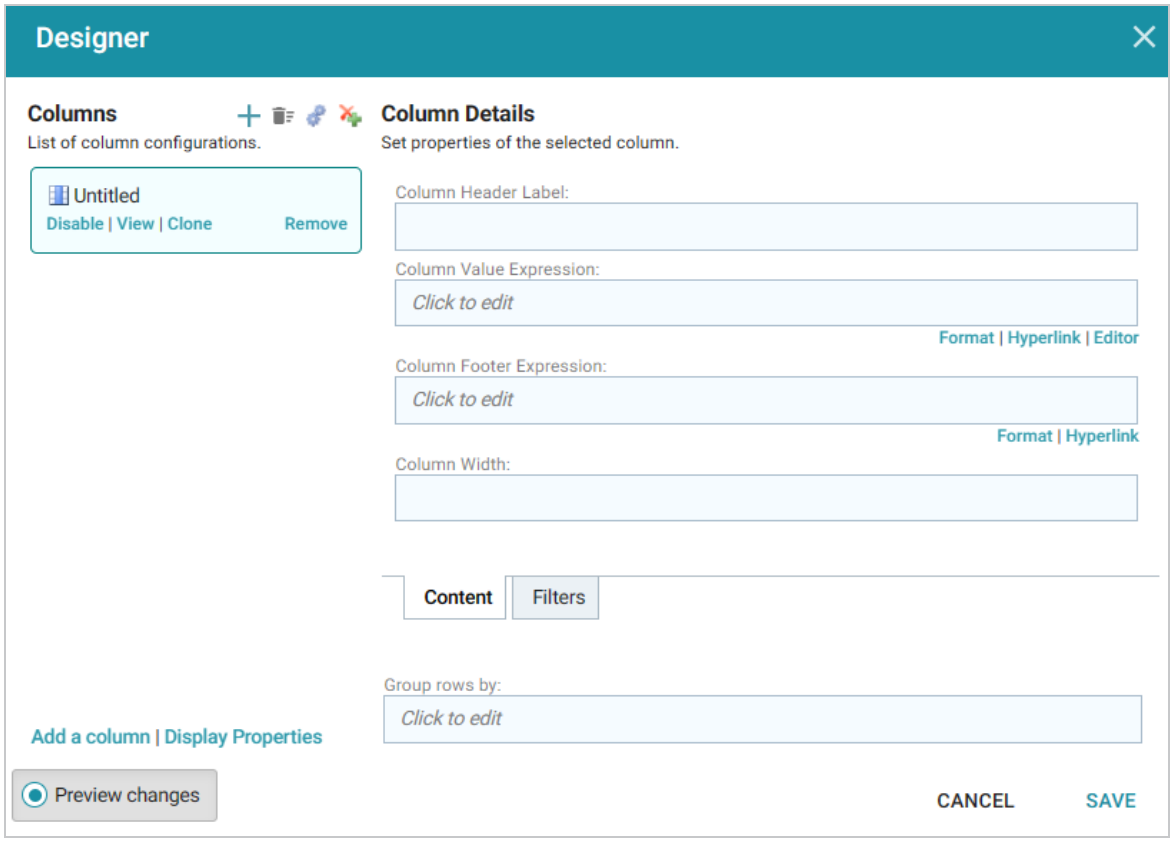
14. Click the **Create a new lens** link at the bottom of the screen. Anzo displays the Create Lens dialog box:



15. Select the **Table** lens and then click **Next**. Anzo displays the Create Lens dialog box:



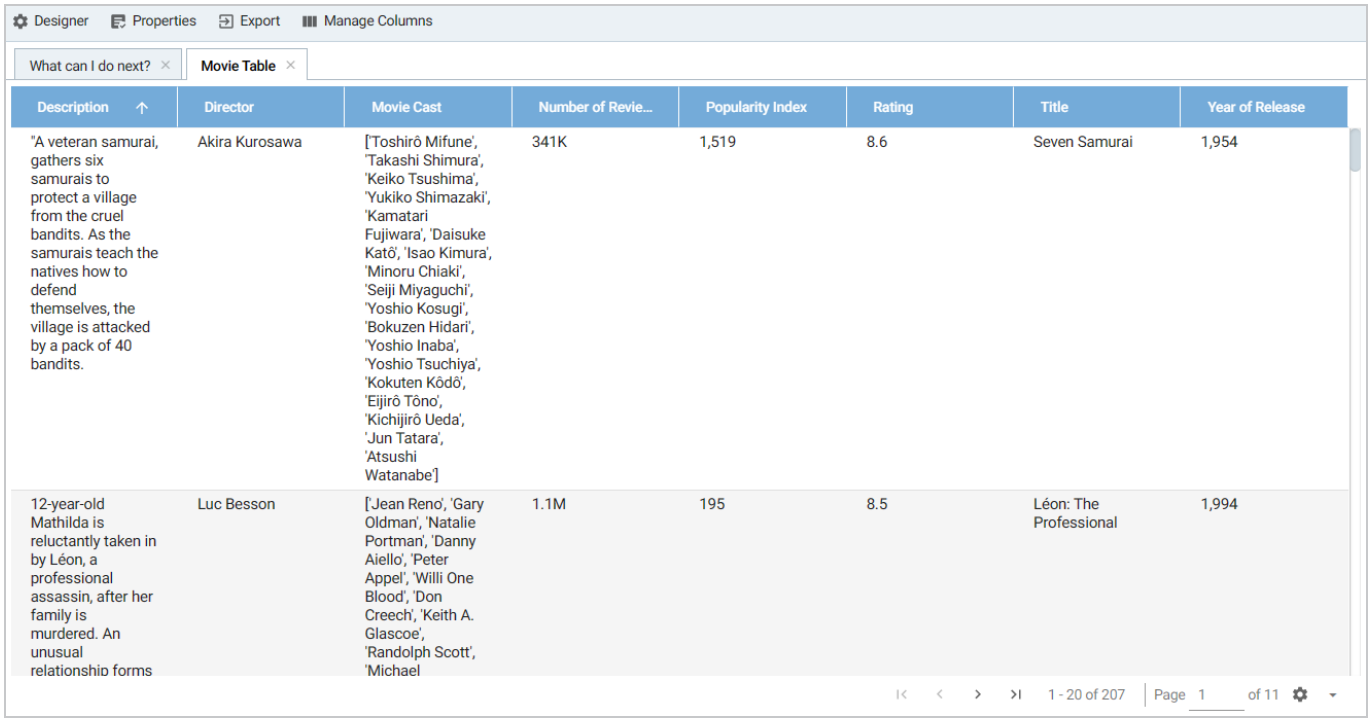
16. Type a **Title** for the lens, for example "Movie Table," and then click **Finish**. Anzo adds the lens as a new sub-tab on the dashboard and opens the Table Designer:



17. In the Designer, click the **Auto-generate columns** icon ( ) to add all available columns to the table.

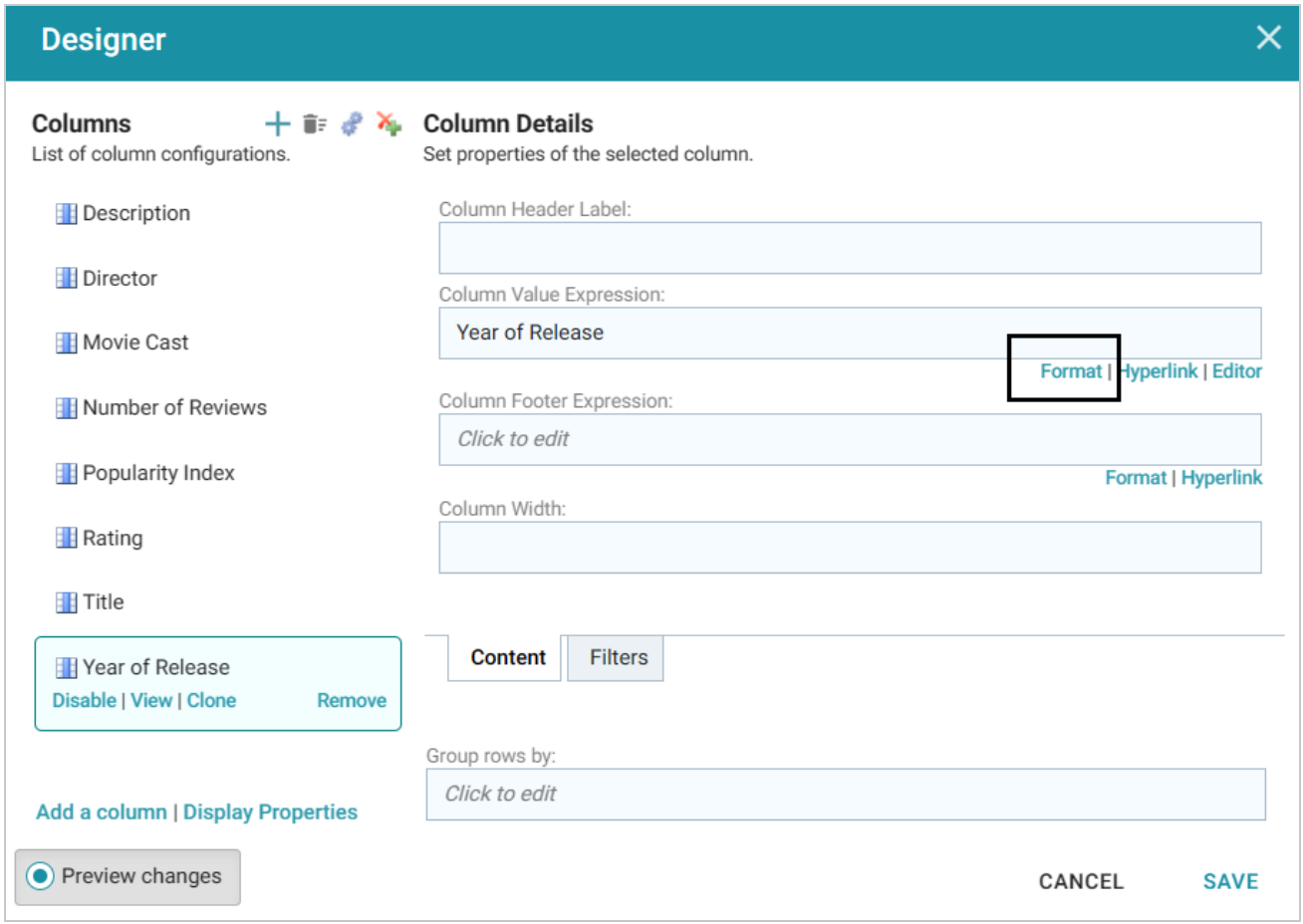
> **Tip**
> Since the sample data set includes only 8 columns, it is not overwhelming to view all columns at once. For larger data sets, you might want to be more selective when adding columns to table lenses.

18. Click **Save**. The table lens is added to the dashboard and is populated with data. For example:



Description	Director	Movie Cast	Number of Revie...	Popularity Index	Rating	Title	Year of Release
"A veteran samurai, gathers six samurais to protect a village from the cruel bandits. As the samurais teach the natives how to defend themselves, the village is attacked by a pack of 40 bandits.	Akira Kurosawa	['Toshirô Mifune', 'Takashi Shimura', 'Keiko Tsushima', 'Yukiko Shimazaki', 'Kamatari Fujiwara', 'Daisuke Katô', 'Isao Kimura', 'Minoru Chiaki', 'Seiji Miyaguchi', 'Yoshio Kosugi', 'Bokuzen Hidari', 'Yoshio Inaba', 'Yoshio Tsuchiya', 'Kokuten Kôdô', 'Eijirô Tôno', 'Kichijirô Ueda', 'Jun Tatara', 'Atsushi Watanabe']	341K	1,519	8.6	Seven Samurai	1,954
12-year-old Mathilda is reluctantly taken in by Léon, a professional assassin, after her family is murdered. An unusual relationship forms	Luc Besson	['Jean Reno', 'Gary Oldman', 'Natalie Portman', 'Danny Aiello', 'Peter Appel', 'Willi One Blood', 'Don Creech', 'Keith A. Glascoe', 'Randolph Scott', 'Michael	1.1M	195	8.5	Léon: The Professional	1,994

19. If you want to make adjustments to the table, such as to reformat the **Year of Release** column to remove the commas from the values, click the **Designer** button for the lens. In the Designer, select the **Year of Release** column, and then click **Format** under Column Value Expression in the Column Details:

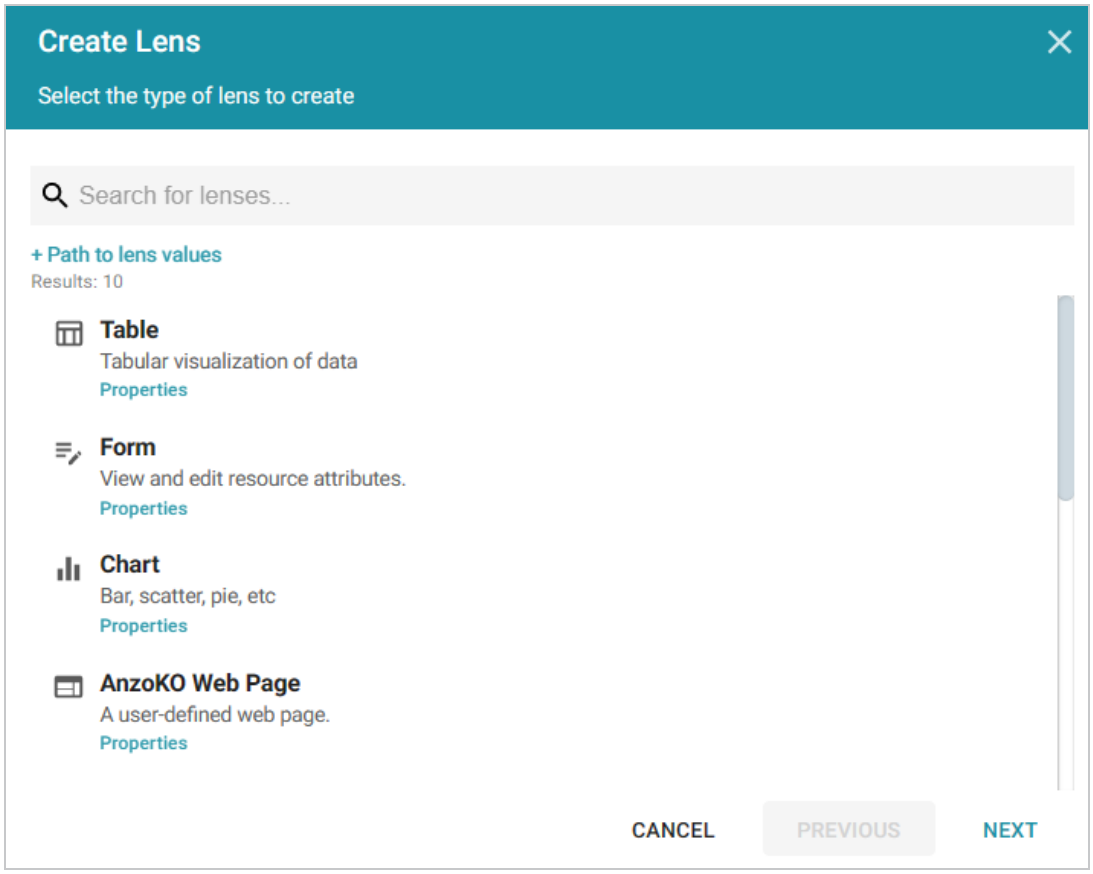


In the Format Column dialog box, click the **Type** drop-down list and select **Year** or **No format**, and then click **Close**. Click **Save** to change the lens change and close the Designer.

20. Now, you can filter the data and reduce the number of results that are displayed by clicking a year value in the **Year of Release** filter. In the filter, click any one of the years in the list to display only the data from that year. You can click another year to filter the table again, and so on.
21. In the Hi-Res Analytics main toolbar, click the **Dashboard** button and select **Save** to save the changes. Remember to save dashboards periodically.

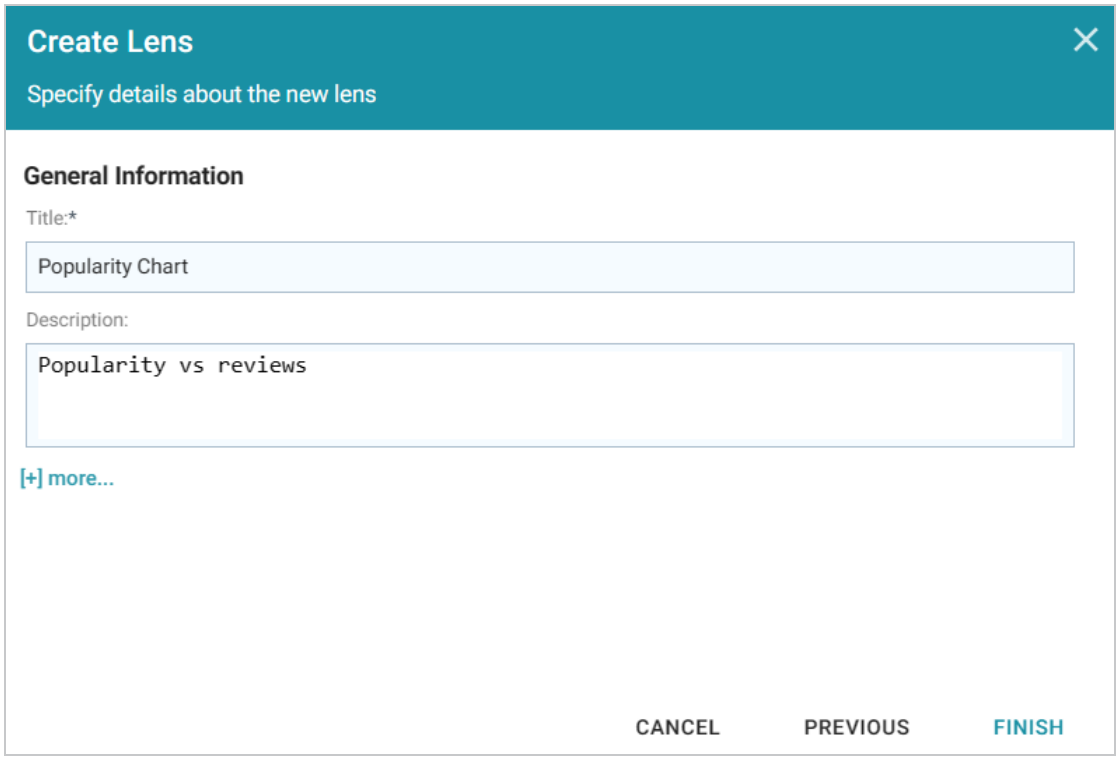
Now that you can view a summary of the data in a table, it can help you determine how to further narrow or expand the results by adding, changing, or removing filters. In addition, you can experiment by adding other lenses to the dashboard to find the ideal way to display the data to answer the questions that you have. For example, the table includes a **Popularity Index** column. It might be interesting to see which movies were the most popular. The next steps guide you through creating a chart lens to display movie popularity.

1. In the Hi-Res Analytics main toolbar, click **Lenses** and select **New**. The Create Lens dialog box opens.

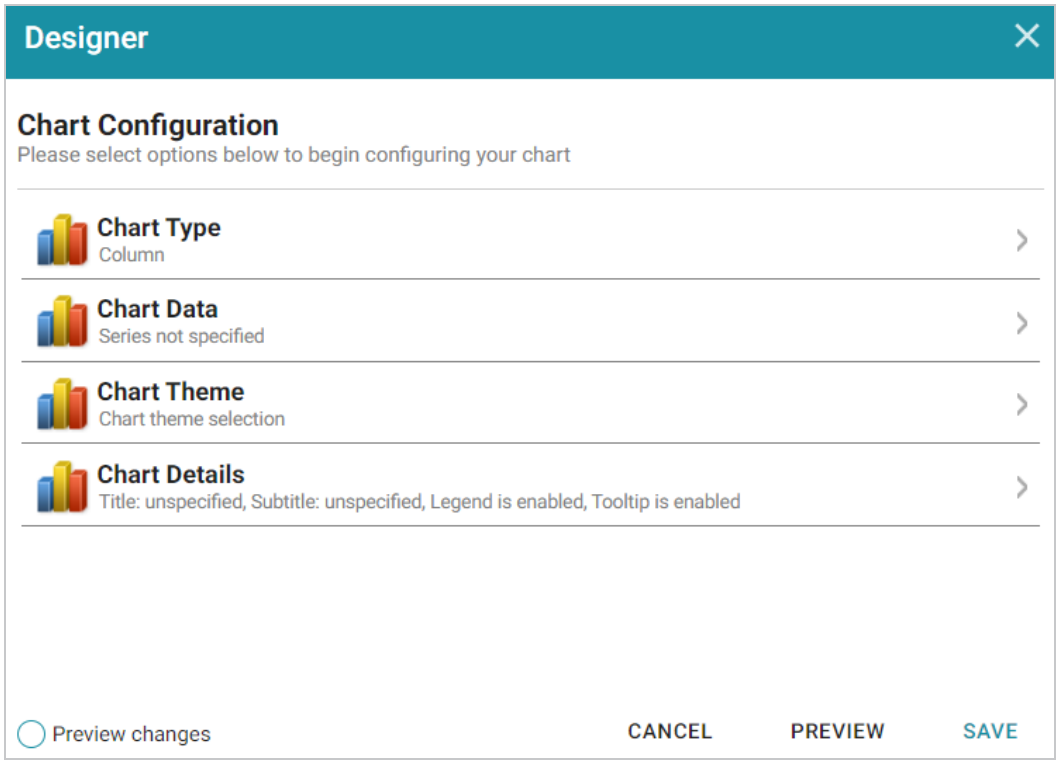


2. Select **Chart** and click **Next**. Anzo displays the Create Lens dialog box.

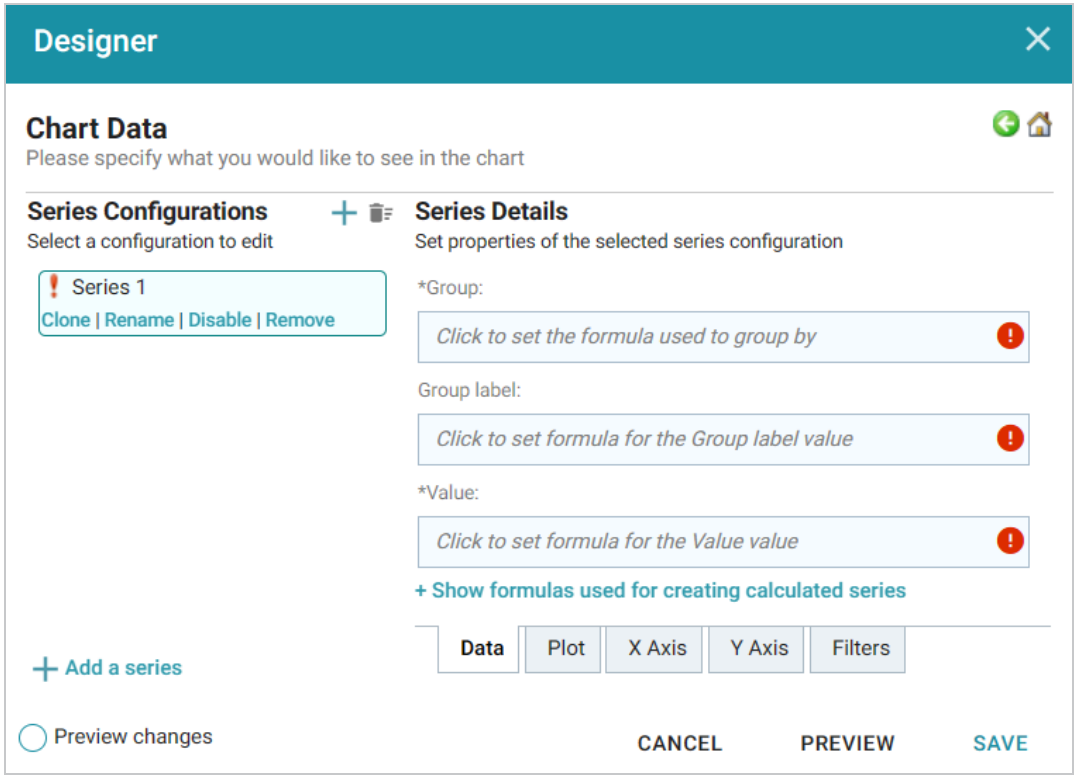
3. Specify a **Title** and add an optional **Description** for the new lens. For example:



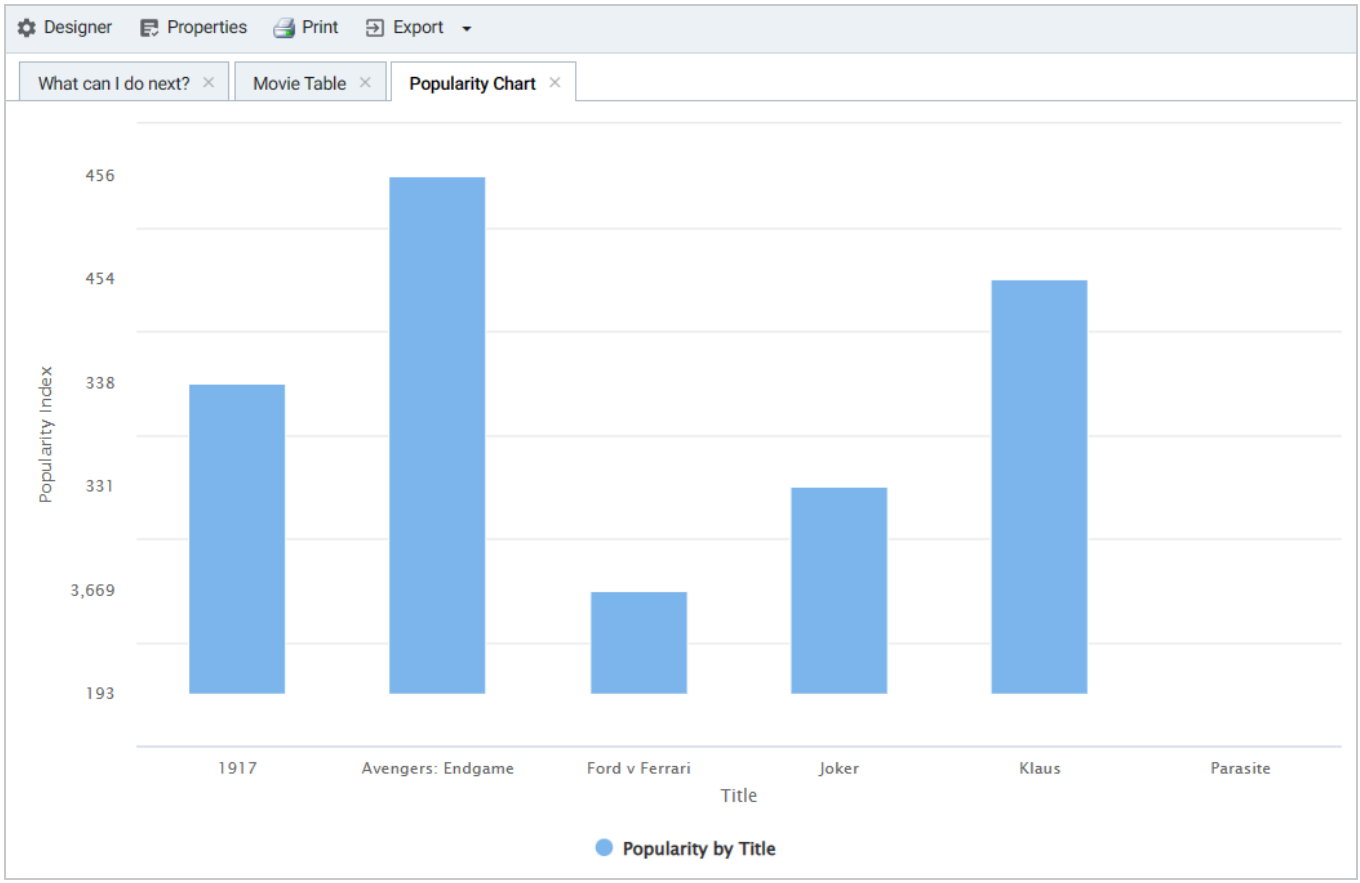
4. Click **Finish**. Anzo displays the Chart Designer, which enables you to configure the type of chart to create, such as column, pie, or line, specify the data that will populate the chart, and customize details such as the chart's theme and fonts, legend, title, and tooltip formats.



5. For this tutorial, accept the default Chart Type of **Column**, and click **Chart Data** to view the Chart Data screen and configure the data to display.



6. On the left side of the screen, click **Rename** under **Series 1**. Specify a new name in the Rename Series field, for example, "Popularity by Title," and then click **OK**.
7. On the right side of the screen, click the **Group** field and select the **Title** property from the drop-down list. The Group field sets the values for the X Axis on the column chart. Then click **Close** to close the drop-down list. Anzo also populates the Group label with the Title property.
8. Click the **Value** field and select the **Popularity Index** property from the drop-down list. The Value field sets the Y Axis values. Click **Close** to close the drop-down list.
9. Click **Save**. Anzo adds the new Chart lens to the dashboard. To narrow the results that are displayed and filter on one year at a time, click a year in the Year of Release filter. For example, by clicking **2019** in the filter, the chart lists the popularity for titles that were released in 2019 and shows that "Avengers: Endgame" had the highest popularity index for that year:



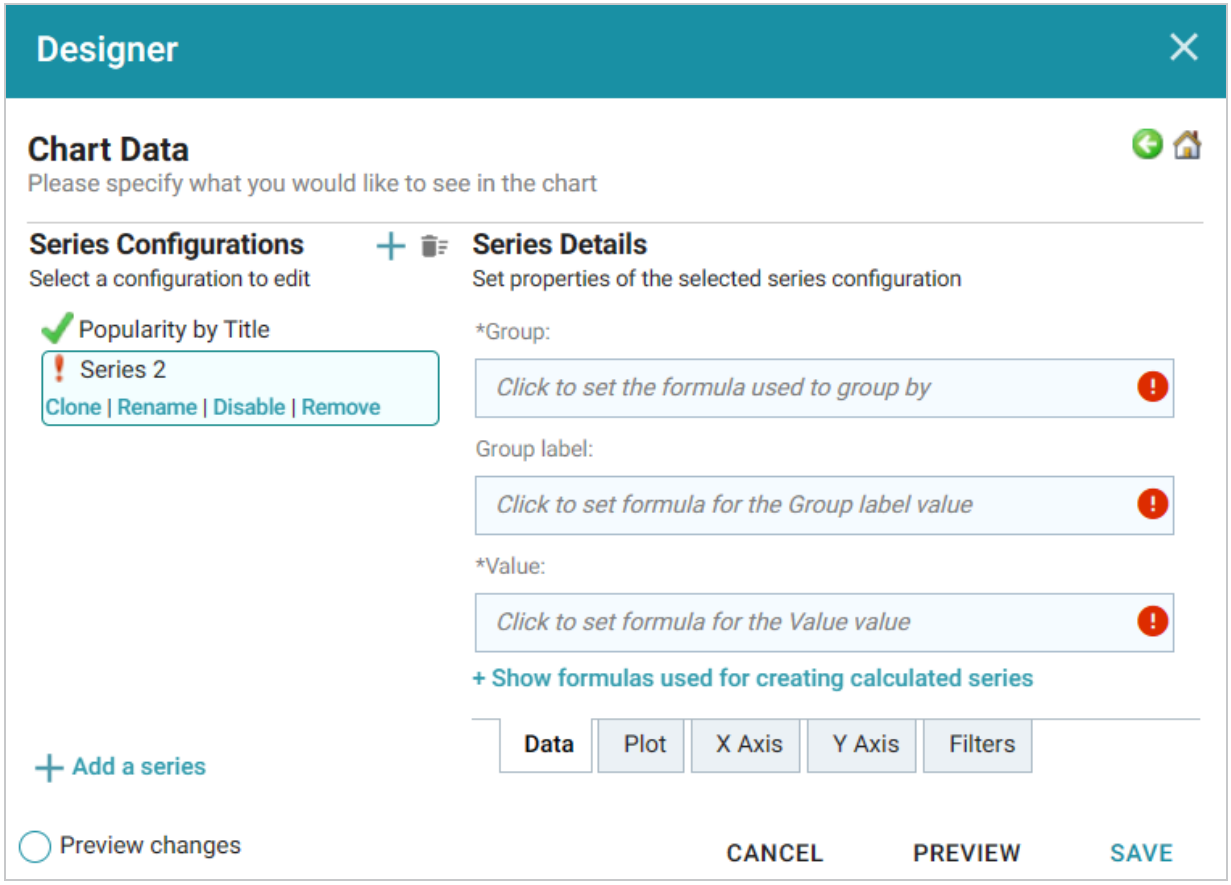
> **Tip** You can hover the pointer over a bar in the chart to view details.

10. If you want to add more data to the Chart, such as to compare Popularity versus Reviews for each Title, follow these steps:

a. Open the Chart Designer by clicking the **Designer** button above the lens tabs.

b. In the Designer, click **Chart Data**.

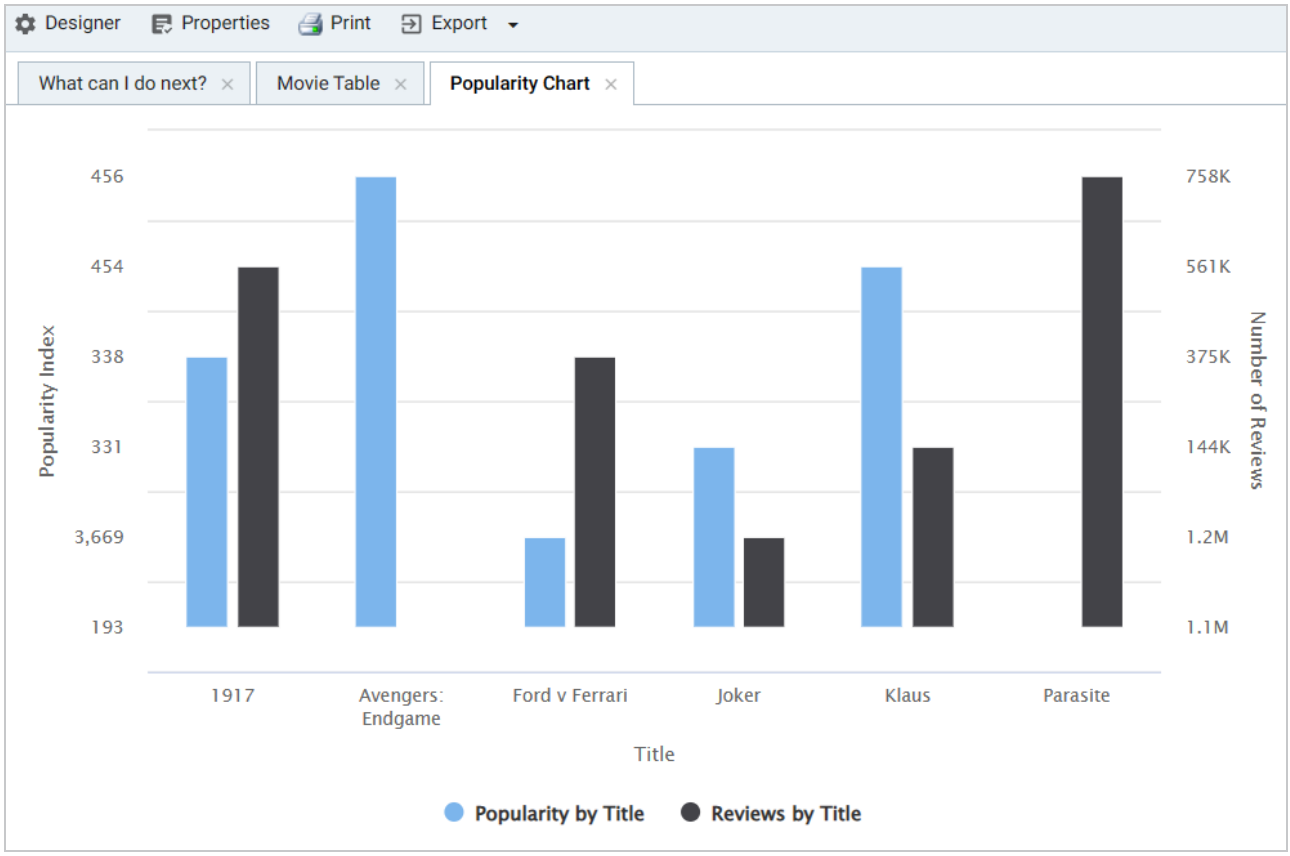
c. On the Chart Data screen, click the **Add a series** link at the bottom of the screen. A new series, named Series 2, is added to the list.



d. Under **Series 2**, click **Rename** and specify a name for the series, such as "Reviews by Title." Then click the **Group** field on the right side of the screen and select the **Title** property from the drop-down list. Click **Close** to close the list.

e. Click the **Value** field and select the **Number of Reviews** property from the drop-down list. Close the list.

f. Click **Save** to save the changes. The Chart is updated to display the number of reviews in addition to the popularity. For example:



Note

If you filtered the data in a previous step, the lens remains filtered on the chosen Year of Release. You can de-select the year in the filter or choose an alternate year to explore the reviews versus popularity for additional titles.

Tip

For more information about working with Dashboards, see Analyzing Data with Hi-Res Analytics in the User Guide.

Review the New Artifacts

This section briefly reviews the artifacts that were created as part of the new sample solution and provides references to more details about each of the artifacts.

During the process of automatically ingesting the sample movie data from the supplied CSV file, Anzo generated the following components:

- A model that describes the structure and type of data that the data set contains. For more information and instructions for viewing the derived model, see Introduction to Models in the User Guide.
- A mapping that is input to the ETL job and maps the source data to the target elements defined in the data model. For more information and instructions for viewing the generated mapping, see Working with Mappings in the User Guide.
- A dataset pipeline that contains the ETL job that generates the target data set. For more information and instructions for viewing the generated pipeline, see Working with Pipelines in the User Guide.

Related Topics

Building a Knowledge Graph with the Direct Data Load Workflow

Building a Knowledge Graph with the Direct Data Load Workflow

This tutorial helps get you started with Anzo by teaching you the basic steps to follow when onboarding data using the automated direct data load workflow. The instructions below provide a sample CSV file and guide you through the steps required from creating the data source to visualizing the data in a Hi-Res Analytics dashboard.

1. Prepare the Environment
2. Download the Sample Data
3. Add a CSV Data Source for the Sample Data
4. Create a Graphmart from the Data Source
5. Explore the Data in a Hi-Res Analytics Dashboard

> **Tip**
> This procedure uses the same sample data as the Building a Knowledge Graph with the Automated ETL Workflow tutorial. If you previously added the Top Movies source and imported the data, you can skip the first three steps and start with Create a Graphmart from the Data Source.

Prepare the Environment

This tutorial requires a connection to AnzoGraph. Before starting the procedure, ensure that Anzo has a connection to a static AnzoGraph instance or is configured for dynamic Kubernetes-based AnzoGraph deployments. For instructions on configuring the connection to a static instance, see Connecting to AnzoGraph in the Administration Guide. For information about Kubernetes deployments, see Using K8s for Dynamic Deployments of Anzo Components in the Deployment Guide.

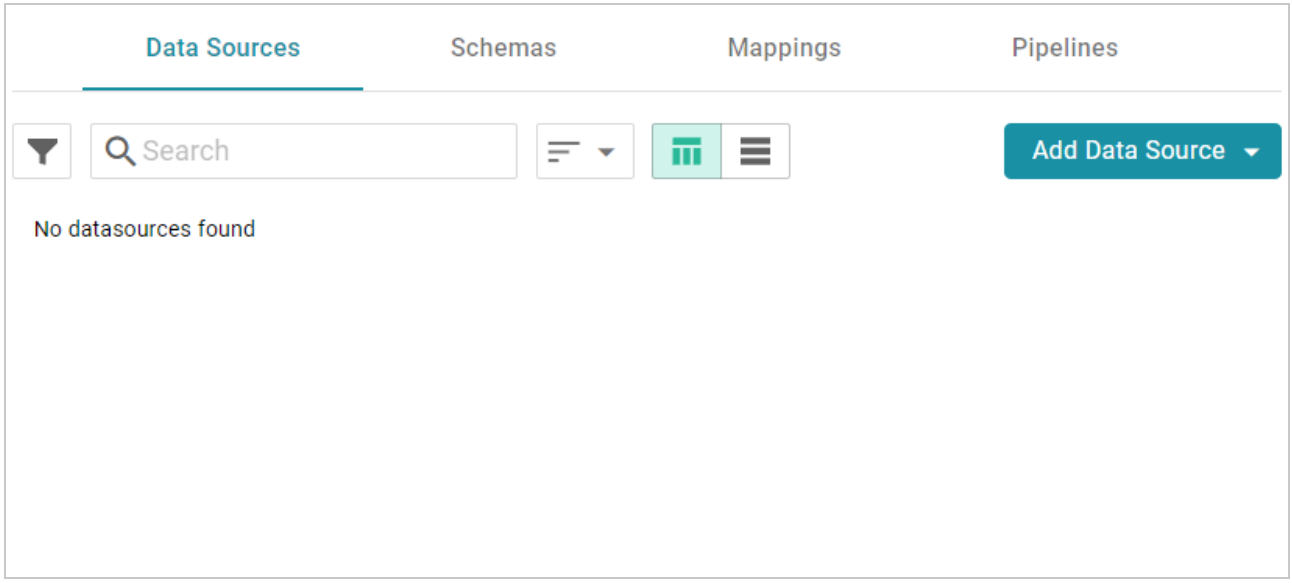
Download the Sample Data

This tutorial uses a small sample data set from Kaggle. The data is in CSV format and includes information about the top 200 movies according to IMDB members. Click the link below to download Top_200_Movies.zip to your computer. Then extract the ZIP file to make **Top_200_Movies.csv** available for importing.

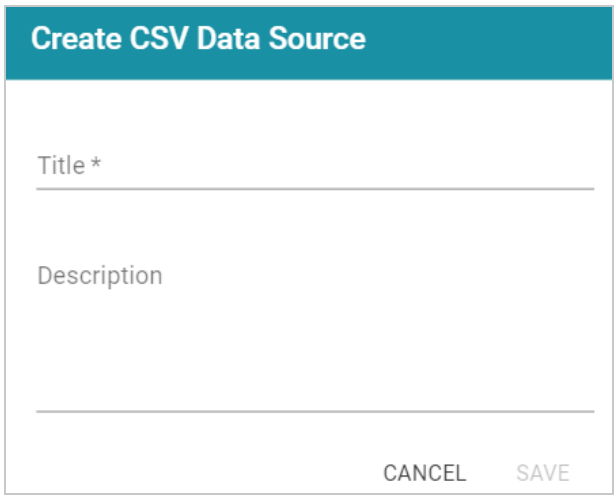
Top_200_Movies.zip

Add a CSV Data Source for the Sample Data

1. In the Anzo application, expand the **Onboard** menu and click **Structured Data**. Anzo displays the Data Sources screen, which lists any existing data sources:

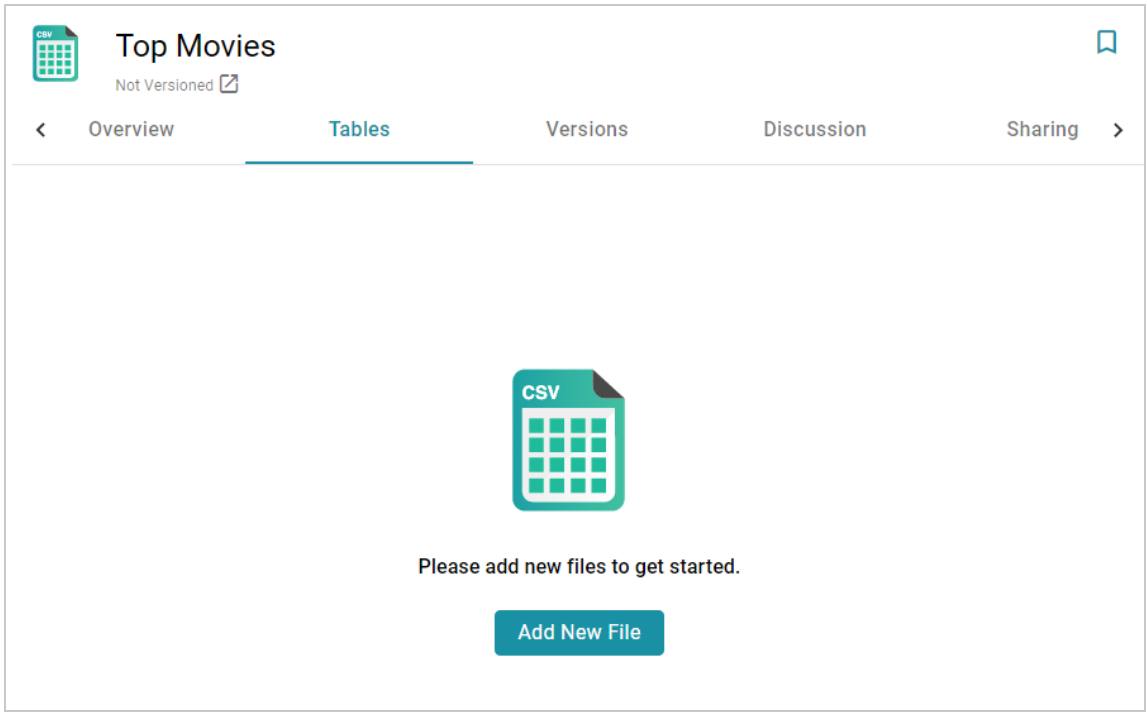


2. Click the **Add Data Source** button and select **File Data Source** > **CSV Data Source**. Anzo opens the Create CSV Data Source screen.

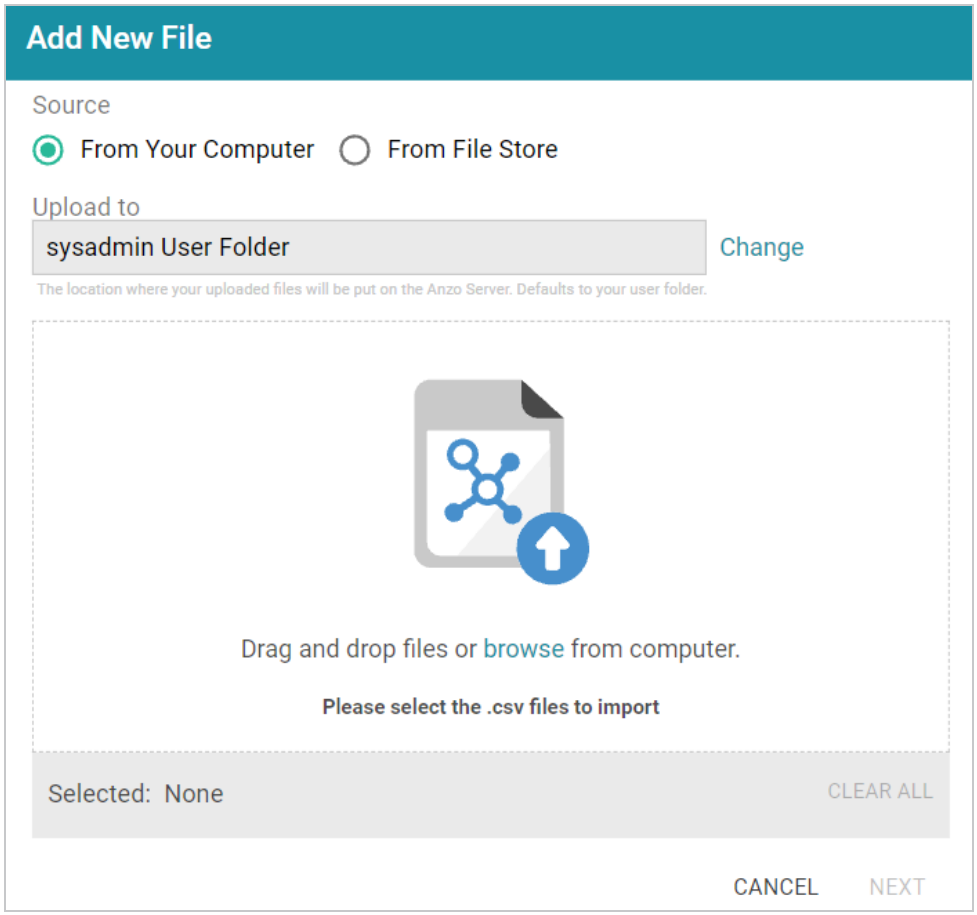


3. Type a name for the source in the **Title** field. For example, "Top Movies." Then type an optional description in the **Description** field. For example:

4. Click **Save**. Anzo saves the source and displays the Tables tab for the new Data Source.

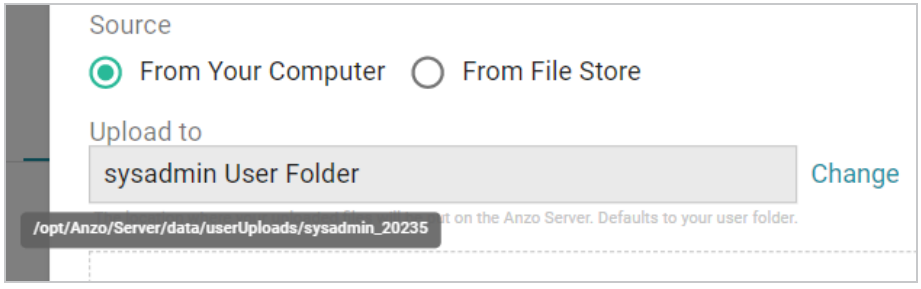


5. Click the **Add New File** button. Anzo displays the Add New File dialog box, and the Source is set to **From Your Computer**.

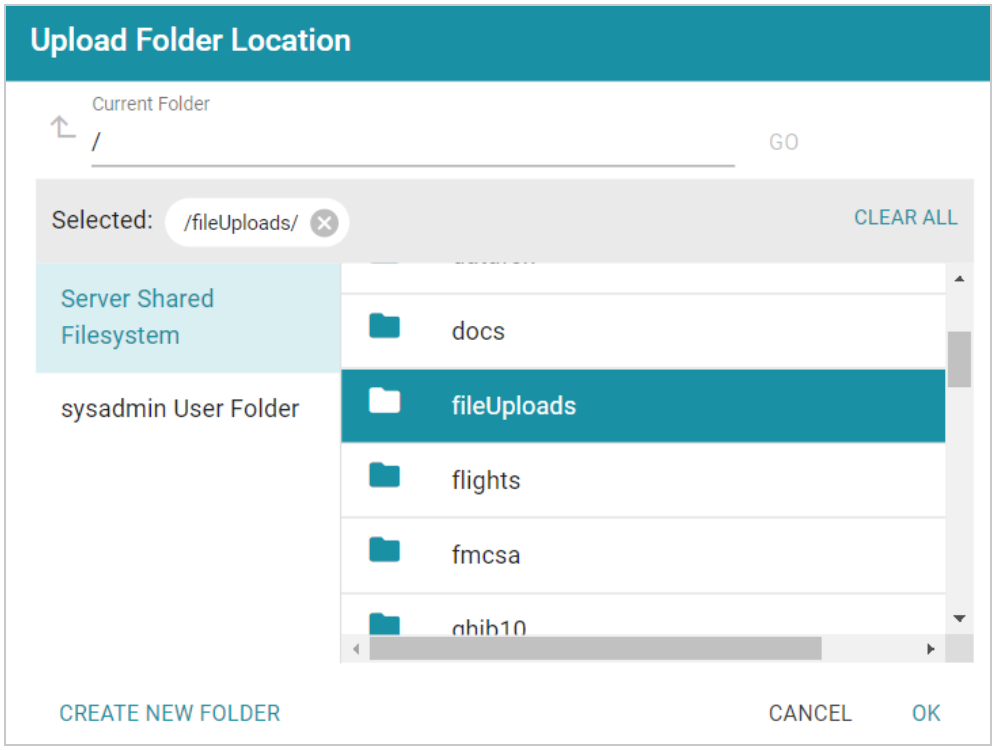


6. As a best practice, check the upload location that is listed in the **Upload To** field by hovering your pointer over the value to view the tooltip. Make sure the upload location is a directory on the shared File Store and not a location in the server installation path. If the file is not uploaded to the shared File Store it is not accessible by applications like AnzoGraph or Spark. In addition, other users cannot publish pipelines for the Data Source because they typically do not have access to the file.

   For example, viewing the Upload To location for the screen above shows that the file will be uploaded to the server installation path, `/opt/Anzo/Server/data...`

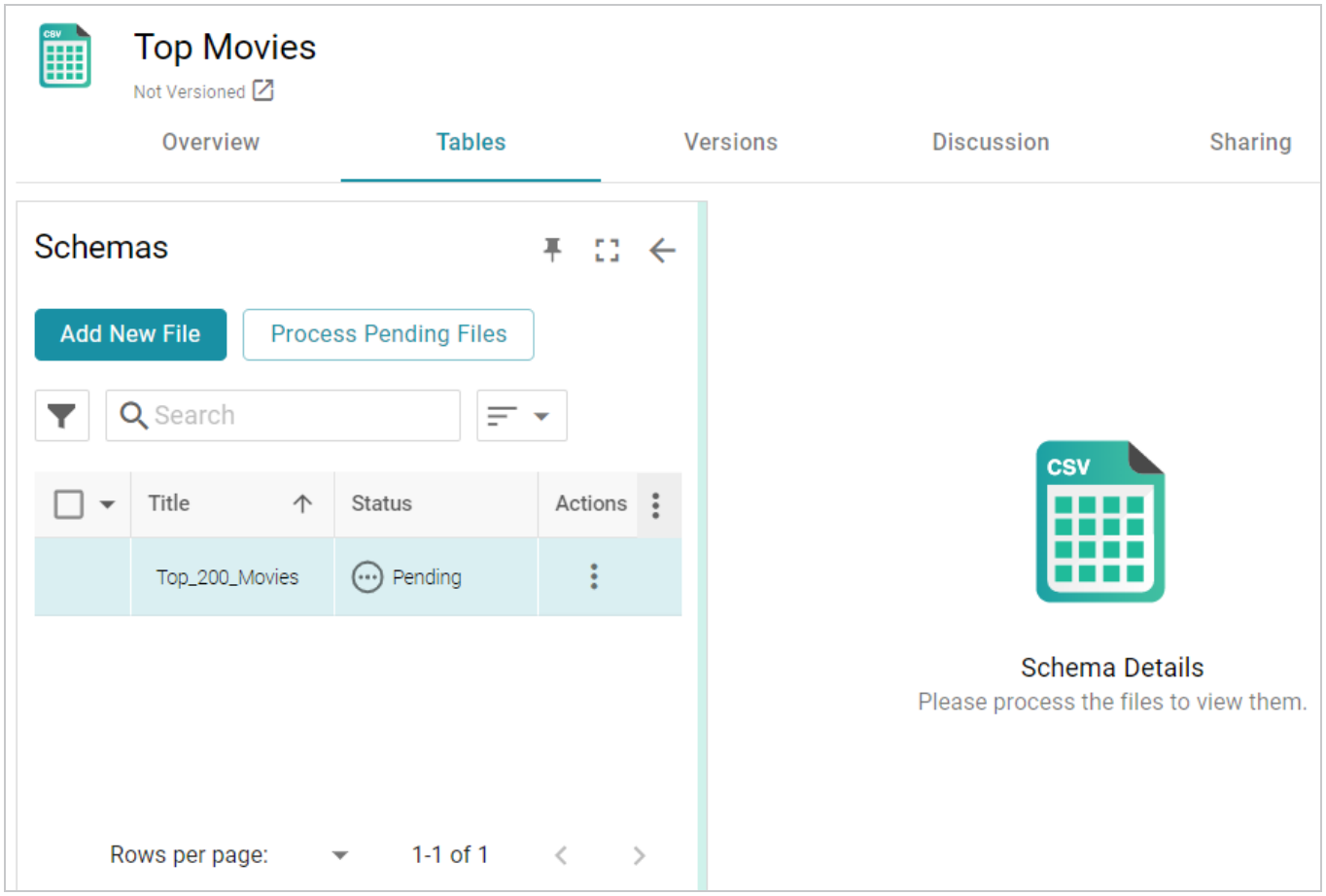


If your Upload To location is configured to upload the file to the server installation path, it is important to click **Change** and select another upload location that is on the shared File Store. For example, the image below shows the Upload Folder Location dialog box that is presented after clicking Change. A folder called **fileUploads** is selected on the File Store.

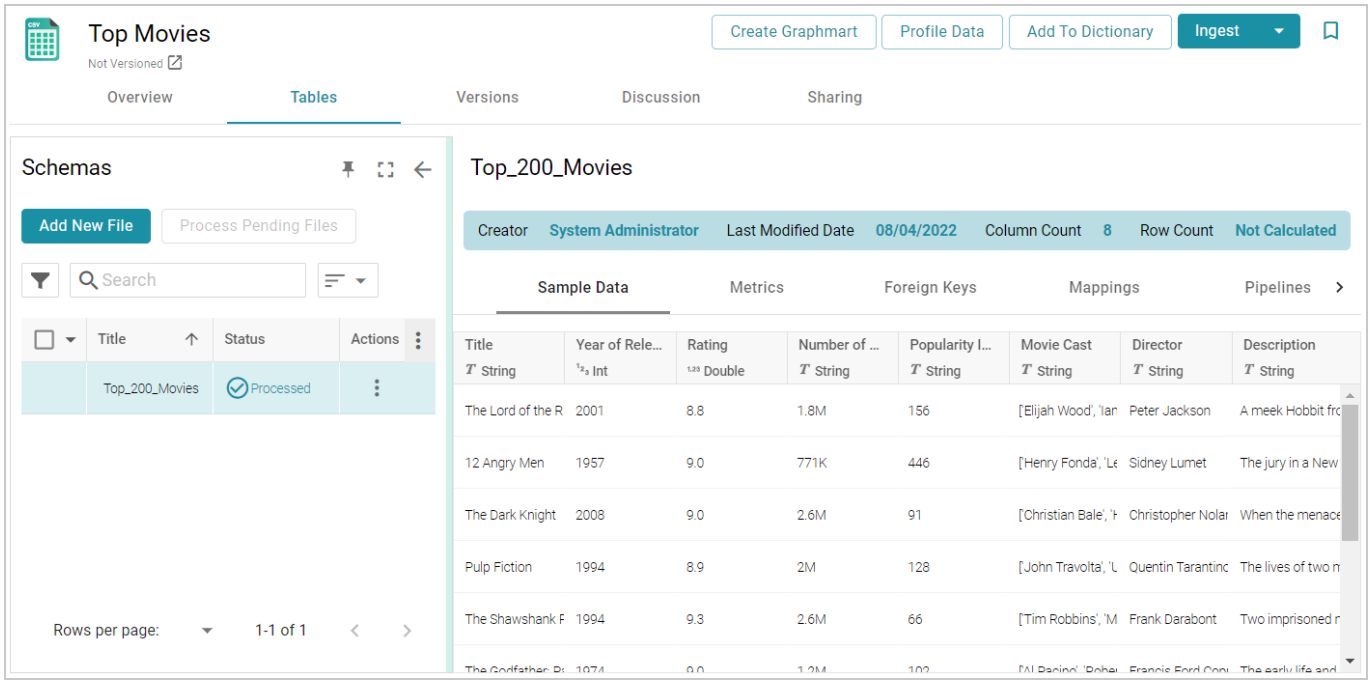


> **Tip**
> An administrator can configure the base upload path so that it points to a location on the File Store by default. For more information, see Setting the Default Base File Store Path for File Uploads in the Administration Guide.

7. Next, drag and drop the Top_200_Movies.csv file onto the screen or click **browse** and navigate to the file and double-click it to select it. Anzo attaches the file and the **Next** button becomes active. Click **Next**. Anzo lists the file on the left side of the screen with a status of Pending. For example:



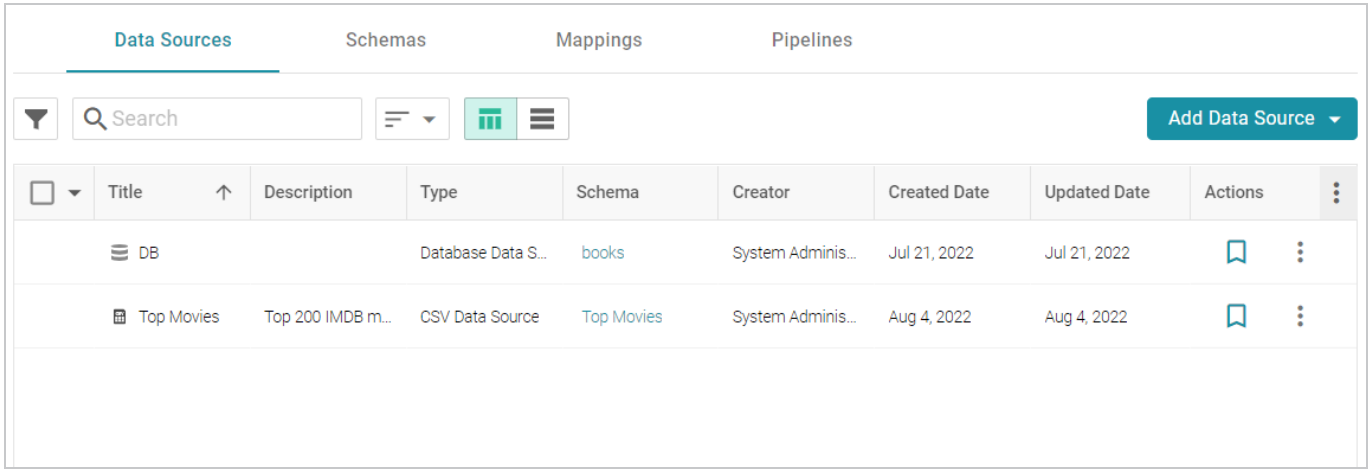
8. Click the **Process Pending Files** button to import the data to Anzo. Anzo imports the data and the status changes to **Processed**. Additional buttons are also displayed on the right side of the screen. You can click the table row for Top_200_Movies to display the schema details.



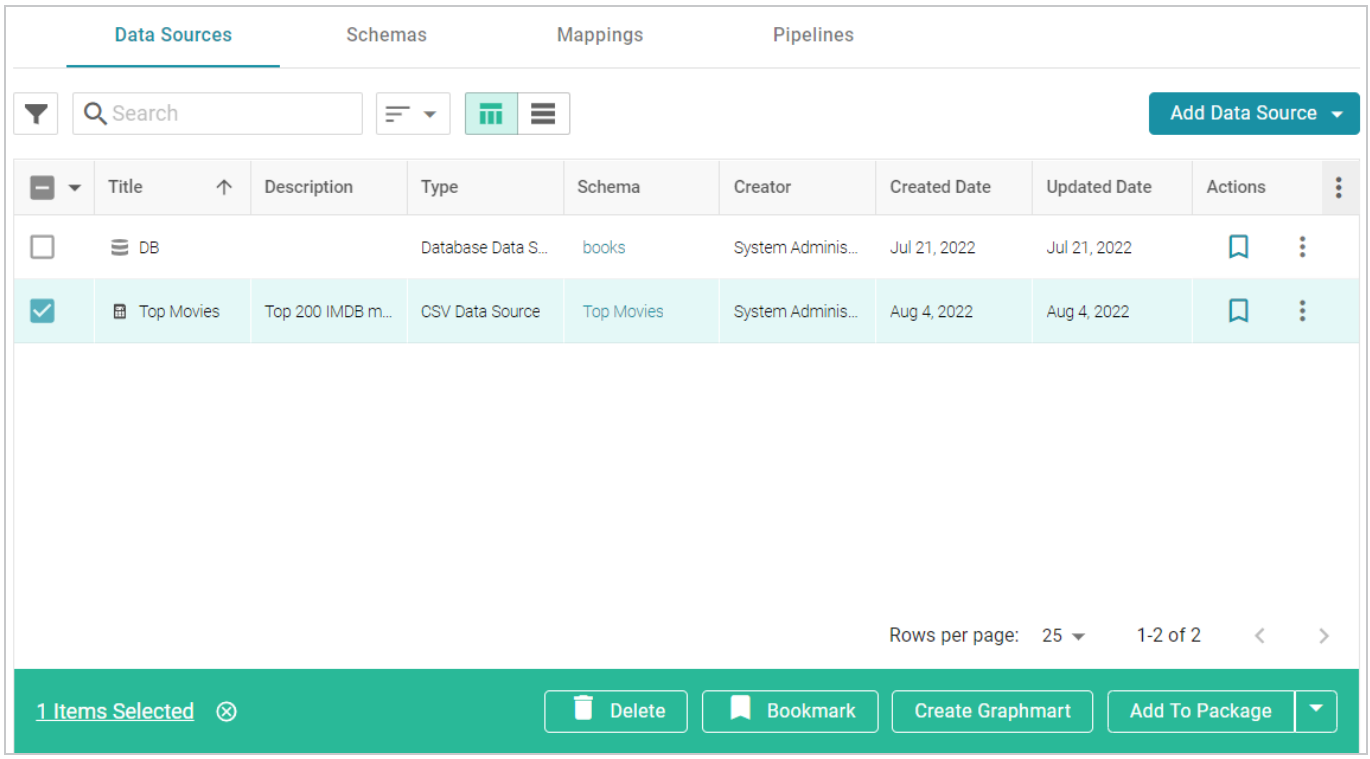
Now that the data source has been added, proceed to the next step to create a Graphmart and load the data to AnzoGraph.

Create a Graphmart from the Data Source

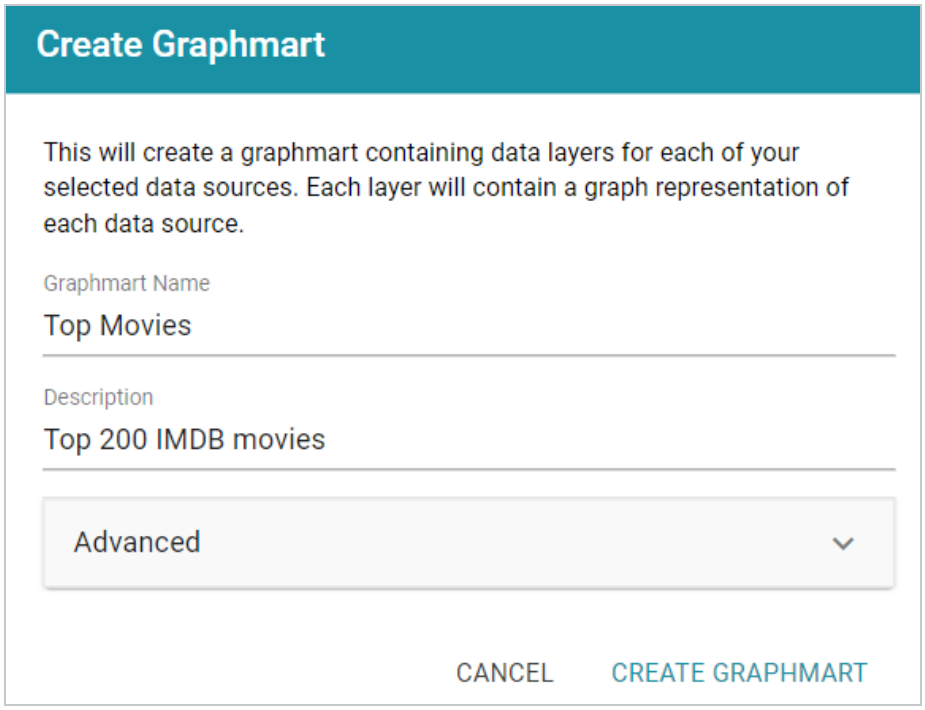
1. From the Tables screen (from the last step of the procedure above), click **Structured Data** in the left menu. Anzo displays the Data Sources screen, which lists any existing sources. For example:



2. Select the checkbox next to the Top Movies data source. Options are enabled at the bottom of the screen. For example:

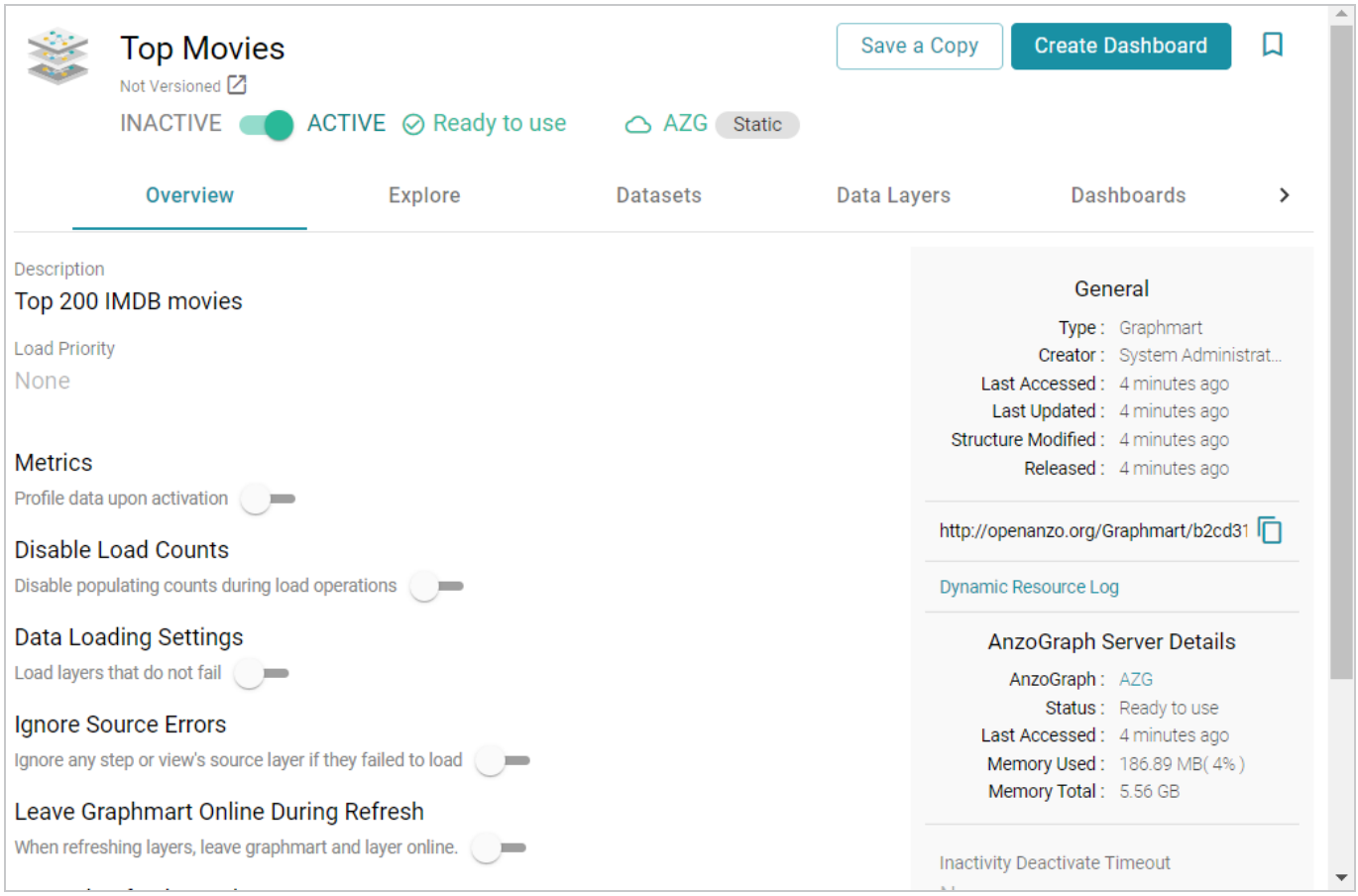


3. Click the **Create Graphmart** button. The Create Graphmart dialog box is displayed:



4. On the Create Graphmart dialog box, the **Graphmart Name** is populated with the name of the data source. You can edit the **Graphmart Name** if necessary and add an optional **Description**. No further configuration is required. For details about the Advanced options that are on this screen, refer to Direct Data Load Advanced Settings Reference in the User Guide.

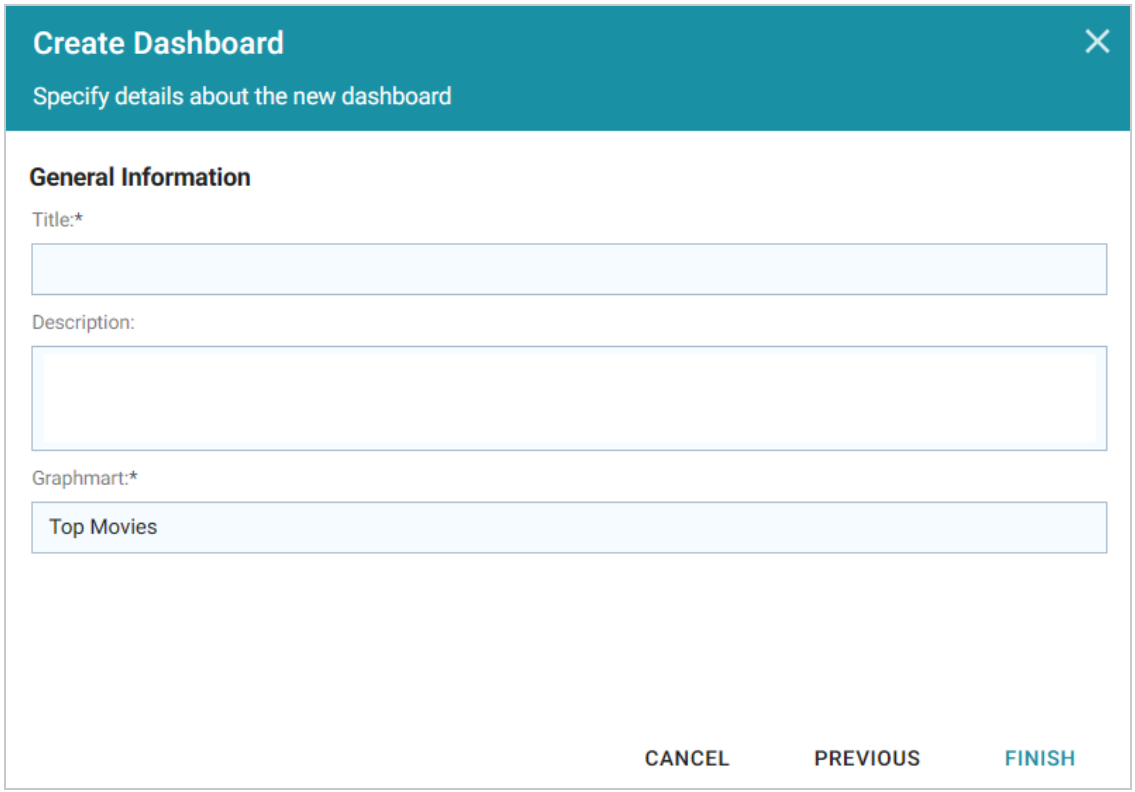
5. Click **Create Graphmart**. The new Graphmart is created and activated.



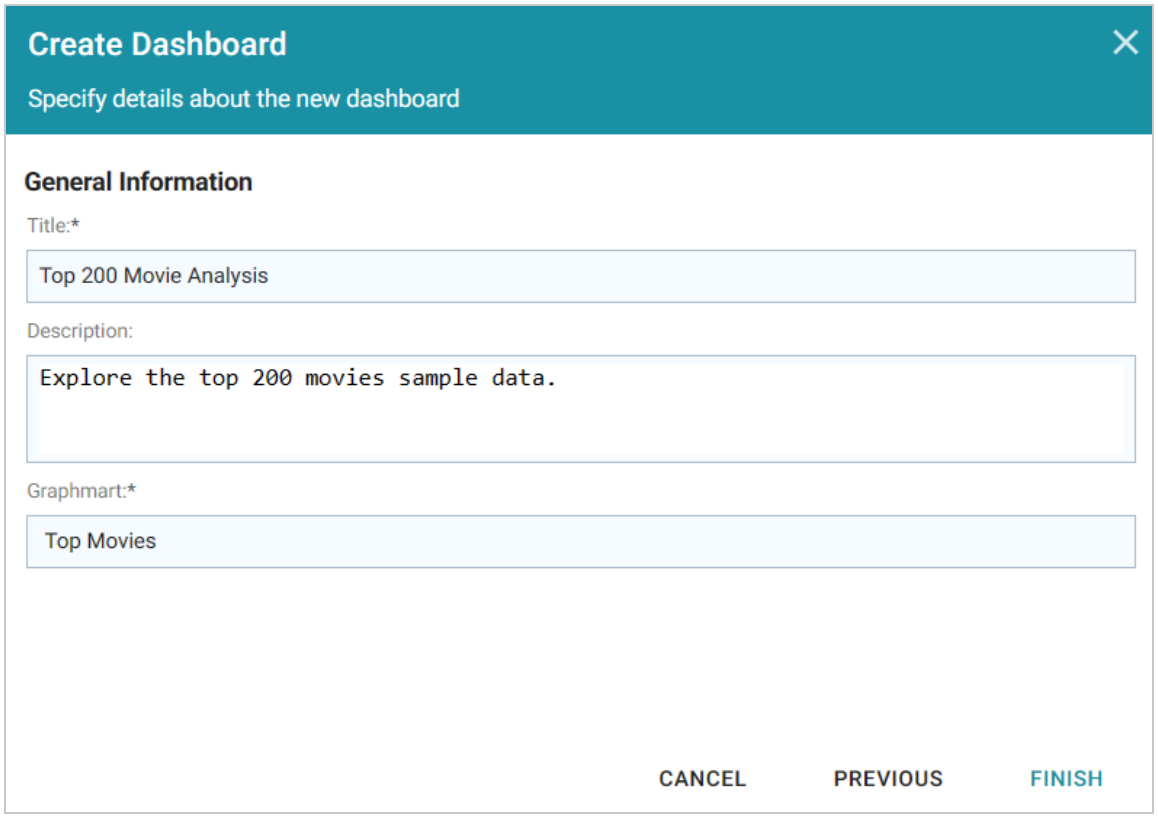
When the activation is complete, you can click the **Data Layers** tab to review the layer and step that were created. In addition to the layer and step, a Managed Model was also created. Now that the data is loaded into AnzoGraph, it is available to explore and analyze in dashboards. Proceed to the next step to explore the sample data in the Hi-Res Analytics application.

Explore the Data in a Hi-Res Analytics Dashboard

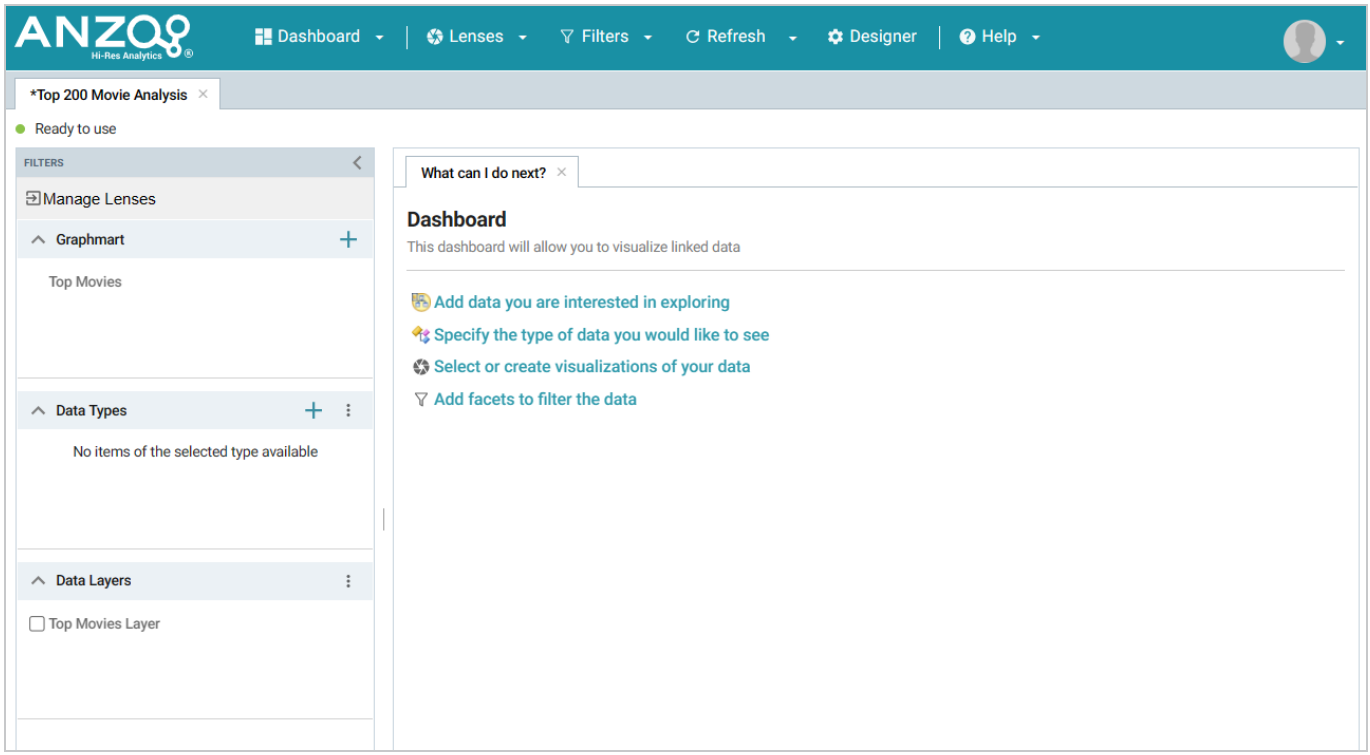
1. On the Graphmart screen (as shown in the last step above), click the **Create Dashboard** button. Anzo opens the Hi-Res Analytics application and displays the Create Dashboard dialog box.
2. Select **Graphmart Dashboard** and click **Next**. The Create Dashboard dialog box is displayed:



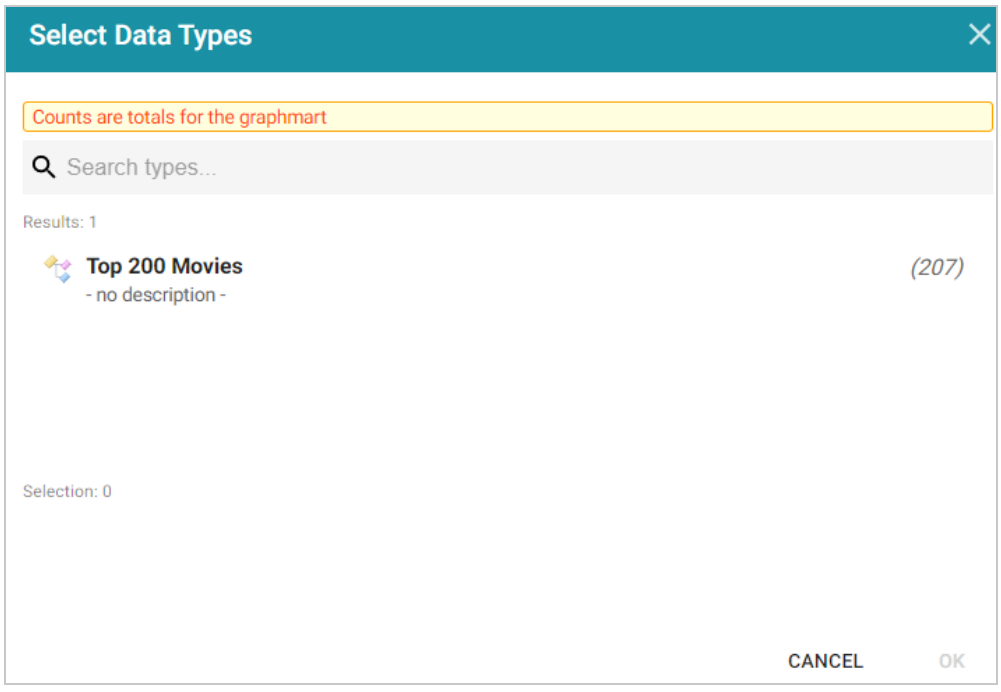
3. On the Create Dashboard screen, specify a **Title** for the dashboard and add an optional **Description**. Leave the default **Graphmart dashboard** value in the **Type** field. For example:



4. Click **Finish** to create the dashboard. The new Dashboard appears as a new tab on the screen and contains a sub-tab titled **What can I do next?**. This tab acts as a wizard to guide you through the initial Dashboard creation. Anzo populates the Graphmart and Data Layers panels with the sample movie data graphmart name and default data layer for the graphmart.



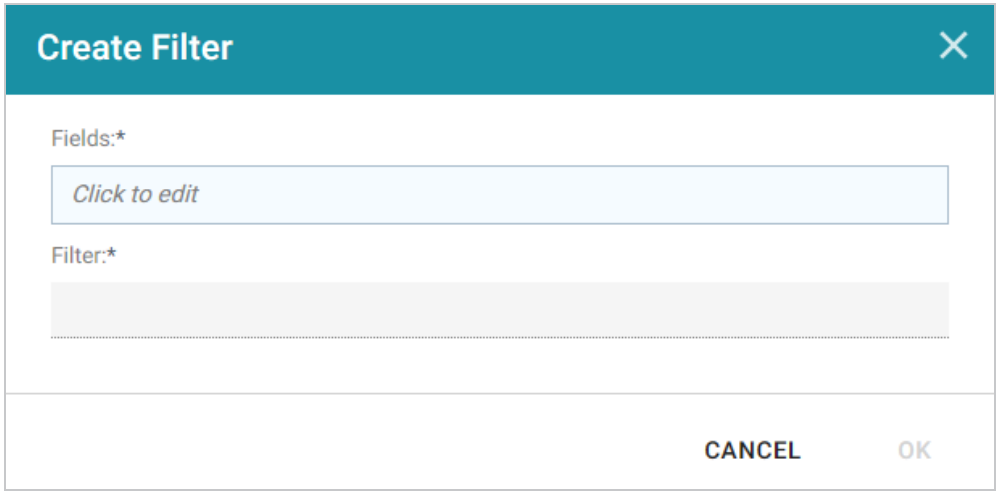
5. To define the type of data to display on the dashboard, click the **Specify the type of data you would like to see** link on the What can I do next tab. The Select Data Types dialog box appears and displays the available data types. Since the sample data set has one type or class of data, only **Top 200 Movies** is listed. The value in parentheses shows the total number of instances of that type.



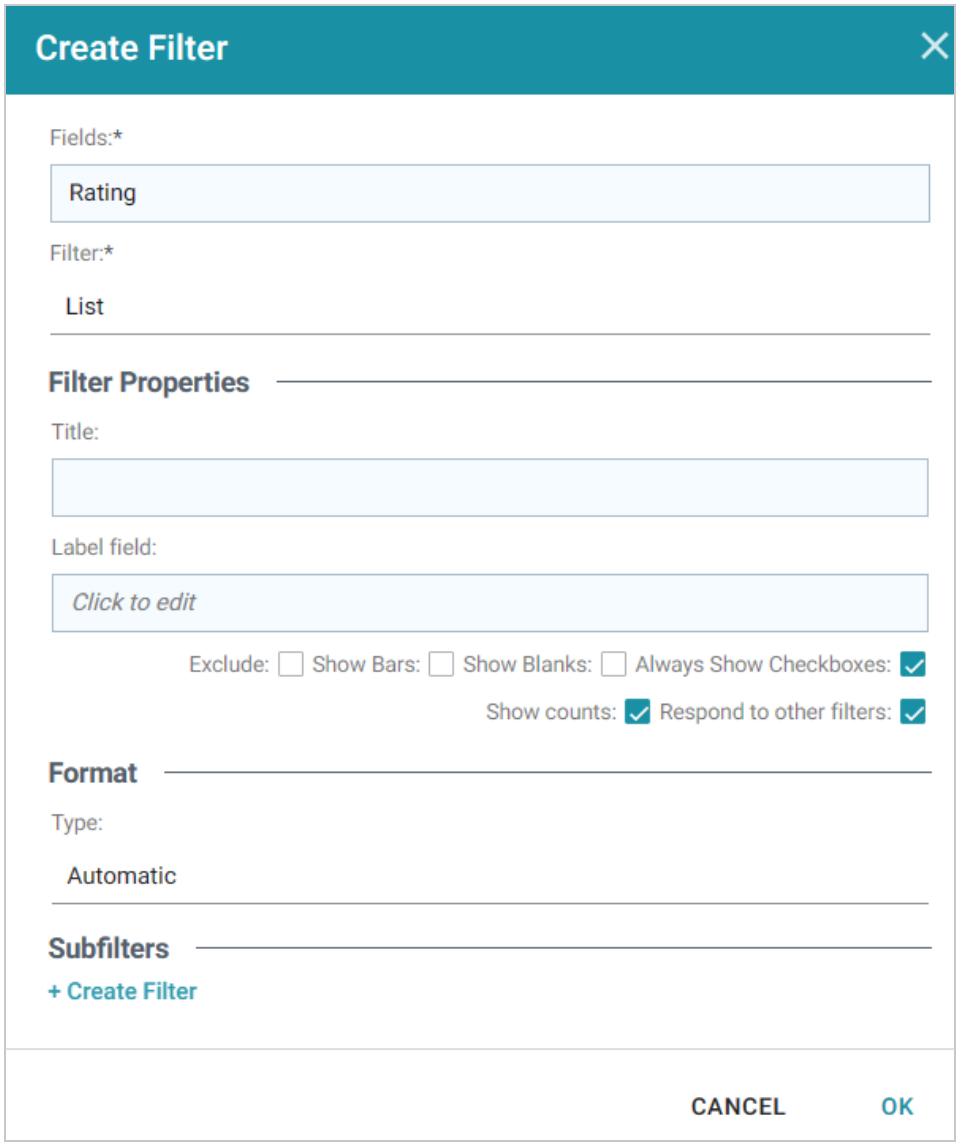
6. Click **Top 200 Movies** to select it, and then click **OK** to add the data type to the dashboard. The selected type is added to the Data Types panel.
7. In the main Hi-Res Analytics toolbar, click the **Dashboard** button and select **Save** to save the Dashboard.

> **Tip**
> To start to explore the data and see what values exist for the properties in the Top_200_ Movies class, it can help to add a filter to the dashboard. Filters reveal the values associated with properties. Learning more about the values enables you to start making decisions about what properties to group on, for example, what properties have relationships, and what results you want to visualize.

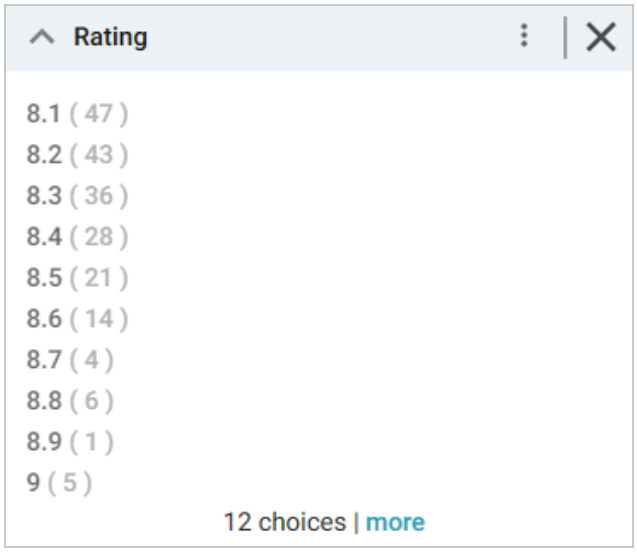
8. To create a filter, click the **Add facets to filter the data** link on the What can I do next tab. Anzo displays the Create Filter dialog box:



9. In the Create Filter dialog box, click the **Fields** field and browse the available properties. For this tutorial, select **Rating**, and then click **Close** to close the Fields drop-down list. Additional options are displayed:

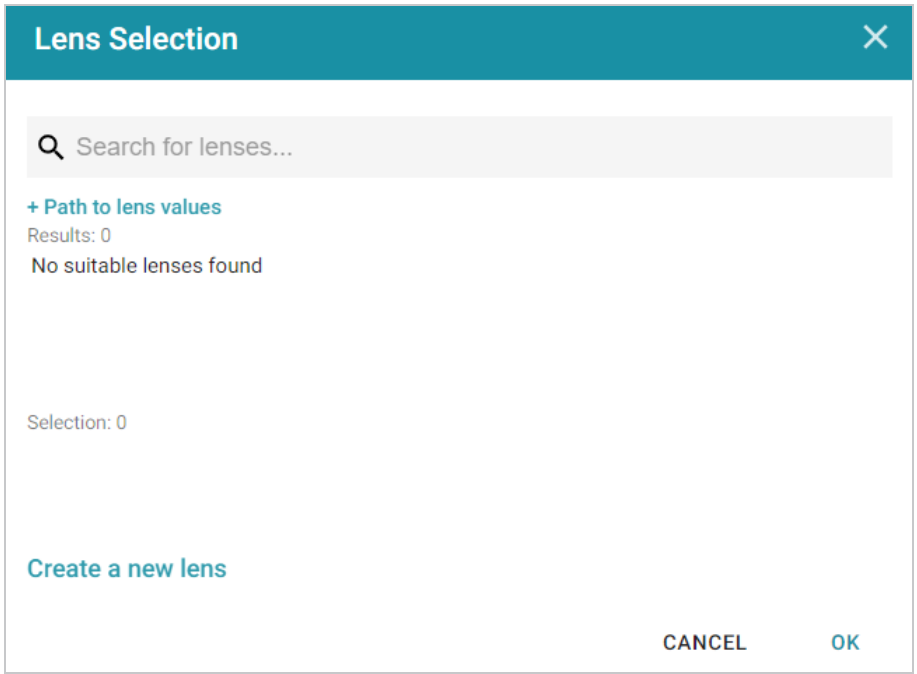


10. Click the **Filter** drop-down list and browse the available filter types. For this tutorial, select **Single Select List**.
11. Click **OK** to close the dialog box and add the new filter to the dashboard. Anzo adds the Rating filter to the left panel. The new filter will enable you to narrow the scope of the data when it is visualized.

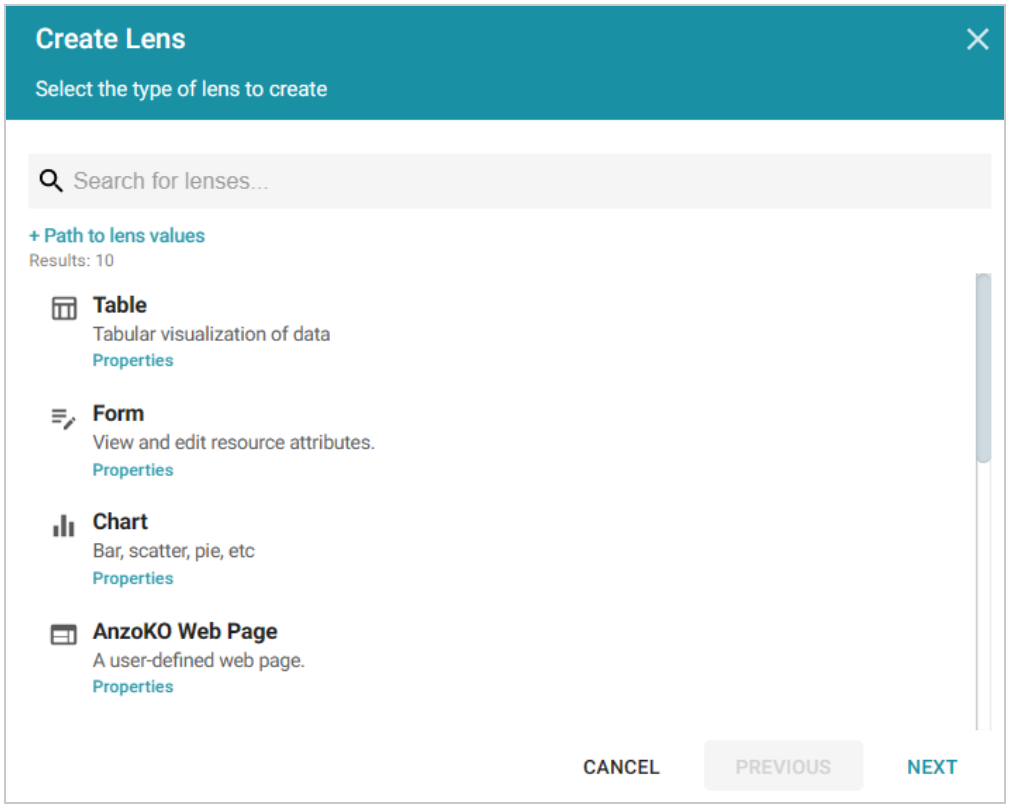


> **Tip**
> Once you have a good understanding of the values and relationships that exist in the data set, you can experiment with lenses and decide on the most appropriate way to display the data. Creating a Table lens is a quick way to get started with learning about the values and relationships in the data and to see how the new filter works.

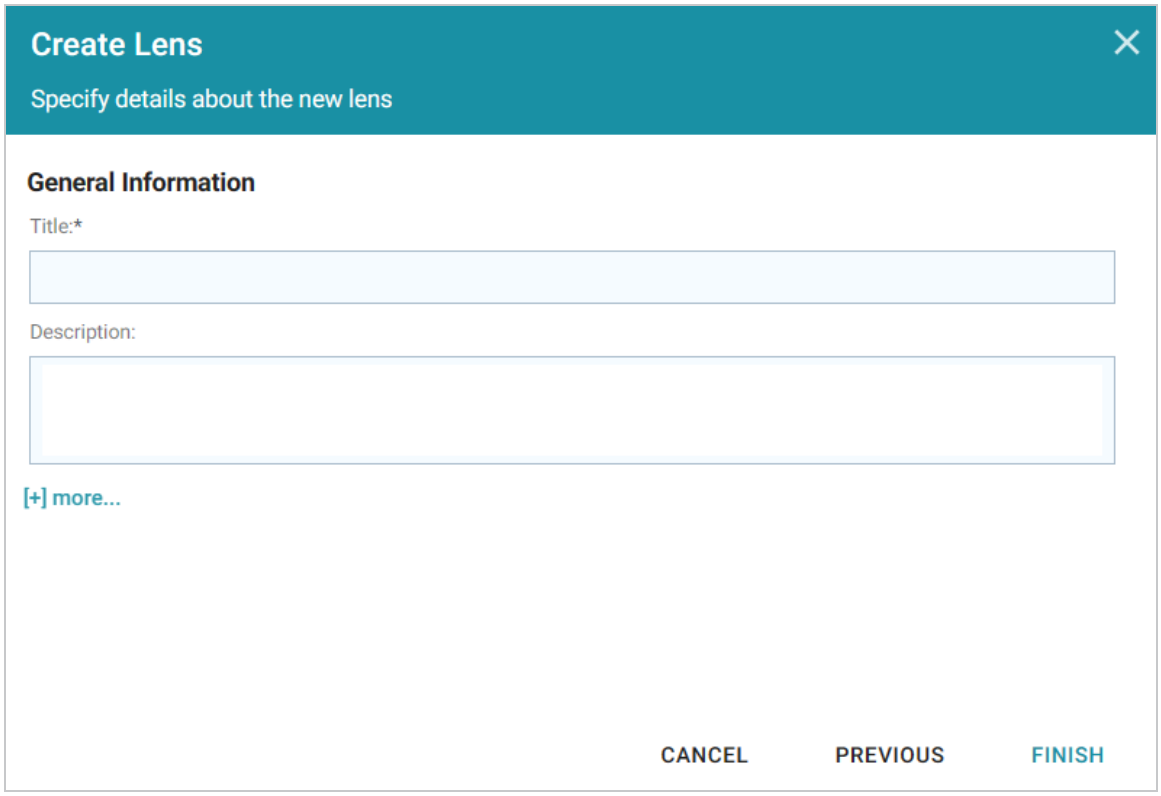
12. To create a Table lens, click the **Select or create visualizations of your data** link on the What can I do next tab. Anzo displays the Lens Selection dialog box. Since lenses have not been created previously, the list of lenses is empty:



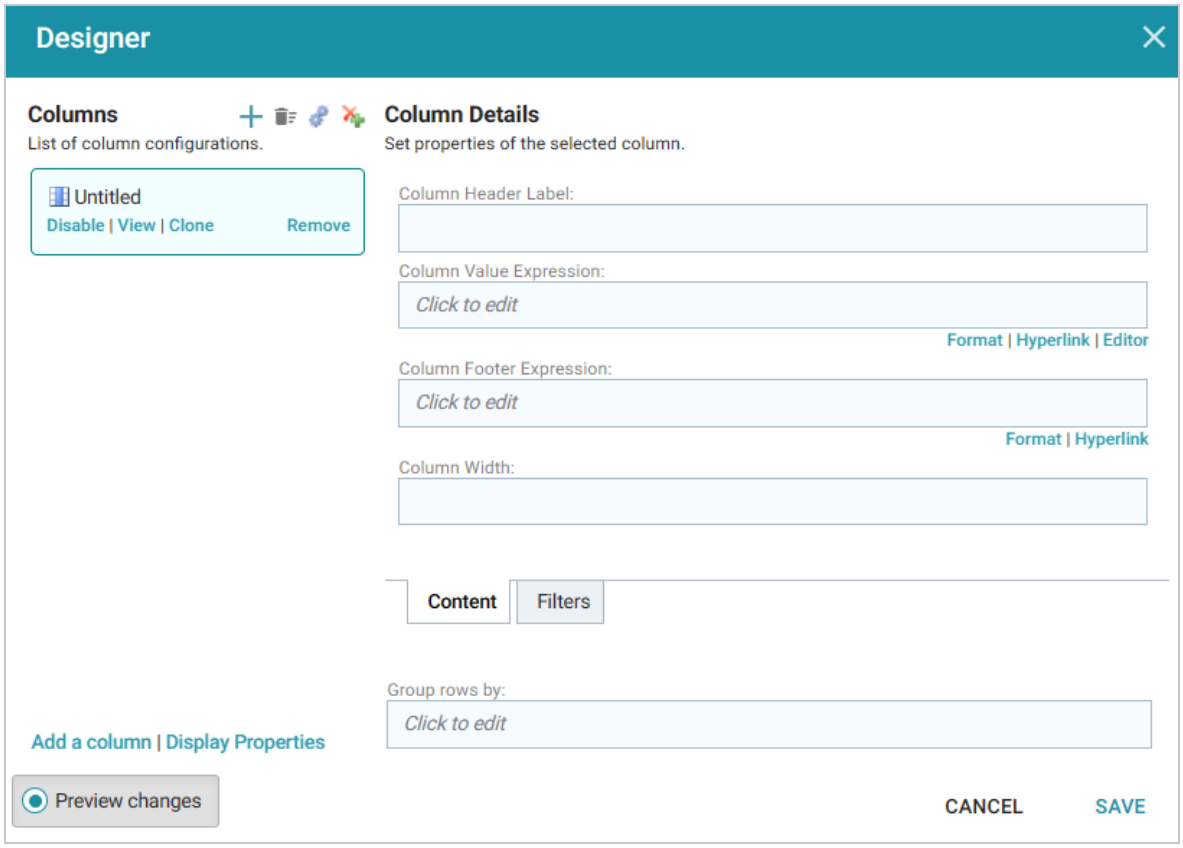
13. Click the **Create a new lens** link at the bottom of the screen. Anzo displays the Create Lens dialog box:



14. Select the **Table** lens and then click **Next**. Anzo displays the Create Lens dialog box:



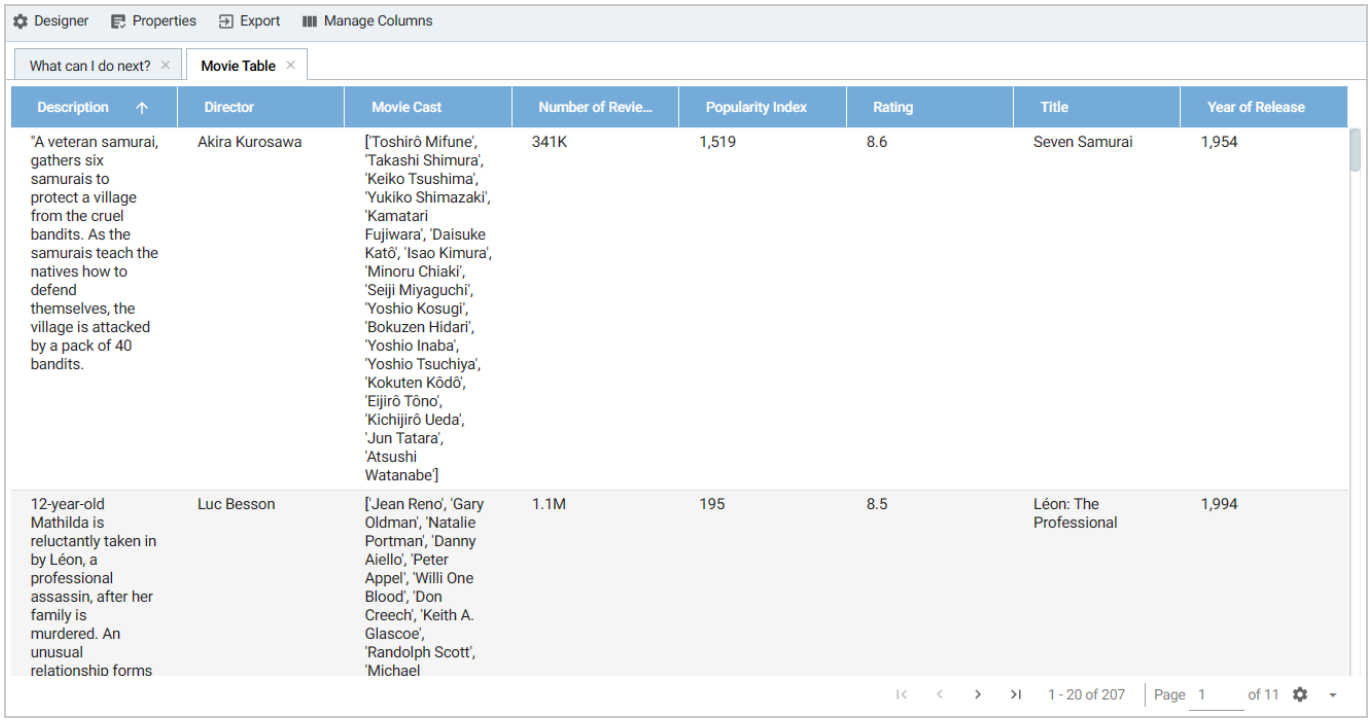
15. Type a **Title** for the lens, for example "Movie Table," and then click **Finish**. Anzo adds the lens as a new sub-tab on the dashboard and opens the Table Designer:



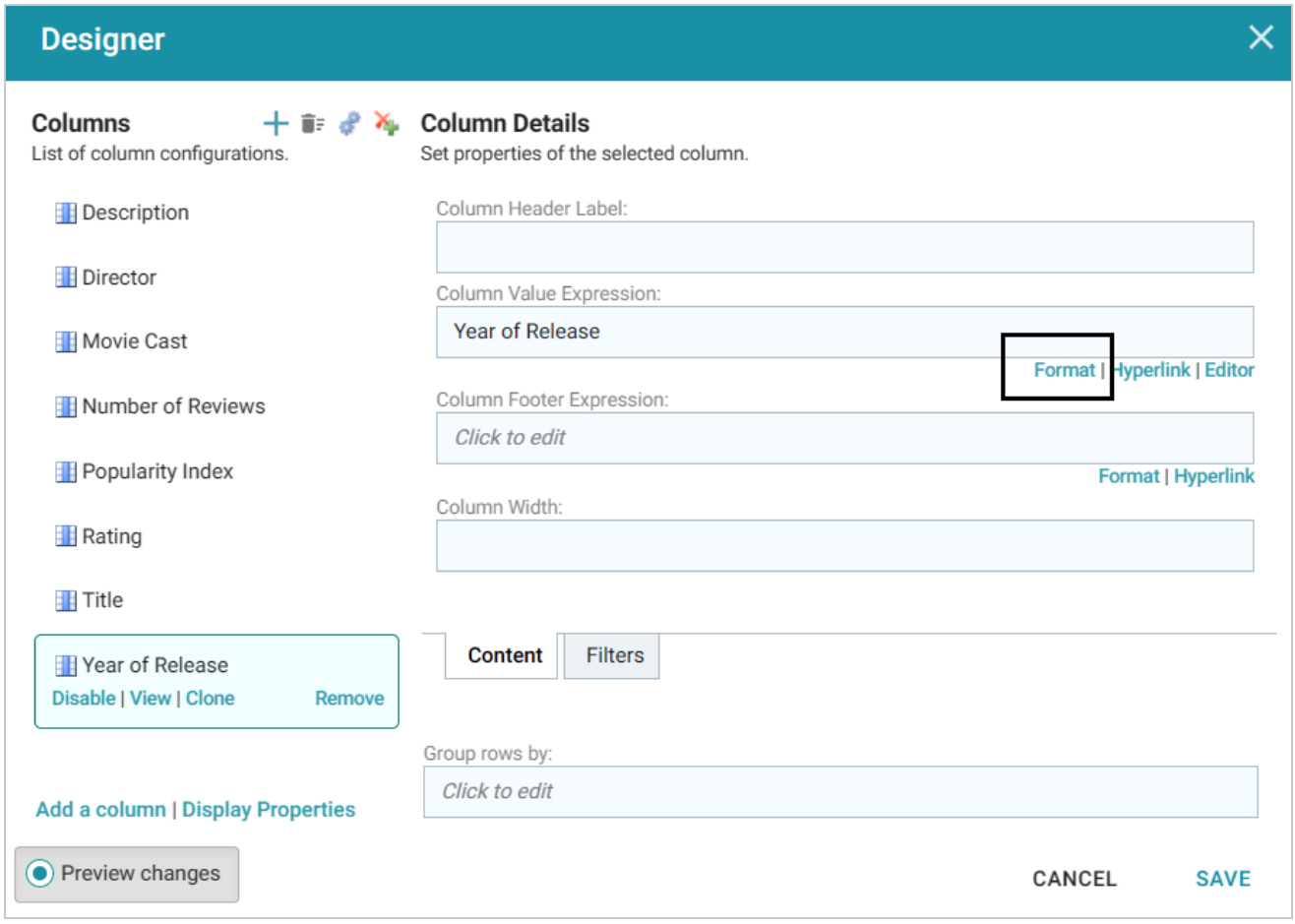
16. In the Designer, click the **Auto-generate columns** icon ( ) to add all available columns to the table.

> **Tip**
> Since the sample data set includes only 8 columns, it is not overwhelming to view all columns at once. For larger data sets, you might want to be more selective when adding columns to table lenses.

17. Click **Save**. The table lens is added to the dashboard and is populated with data. For example:



18. If you want to make adjustments to the table, such as to reformat the **Year of Release** column to remove the commas from the values, click the **Designer** button for the lens. In the Designer, select the **Year of Release** column, and then click **Format** under Column Value Expression in the Column Details:

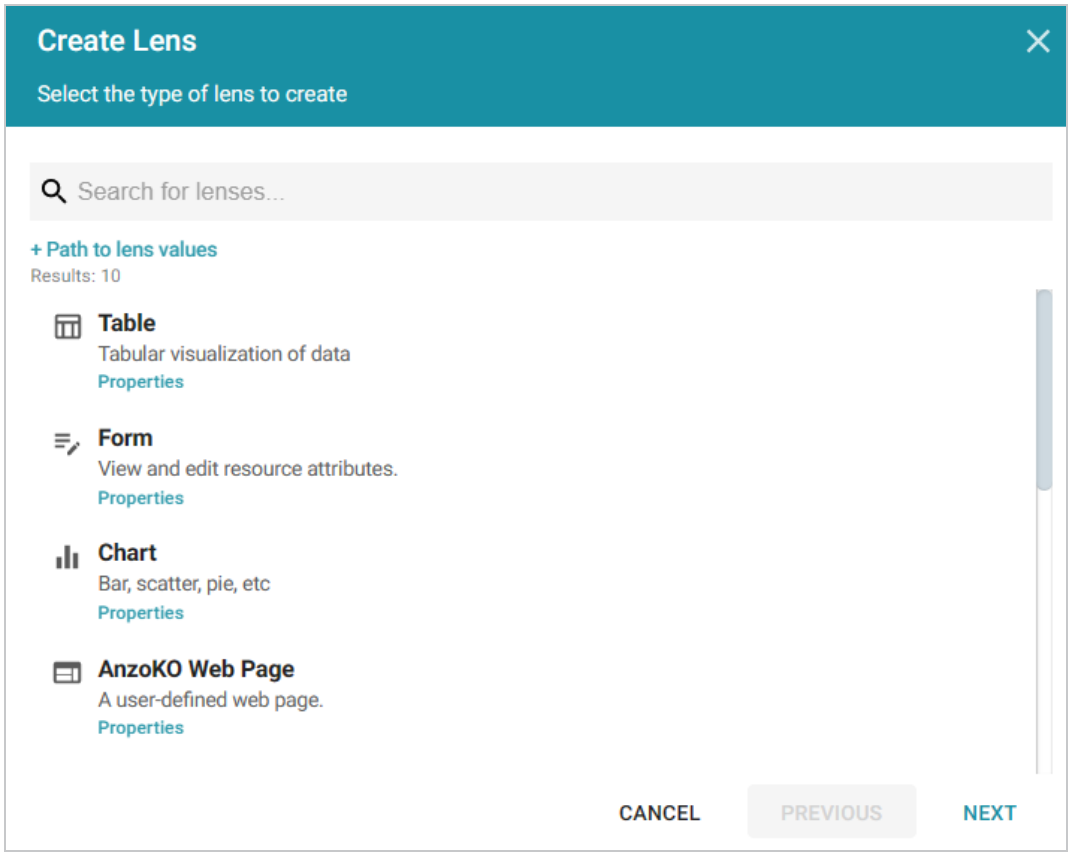


In the Format Column dialog box, click the **Type** drop-down list and select **Year** or **No format**, and then click **Close**. Click **Save** to change the lens change and close the Designer.

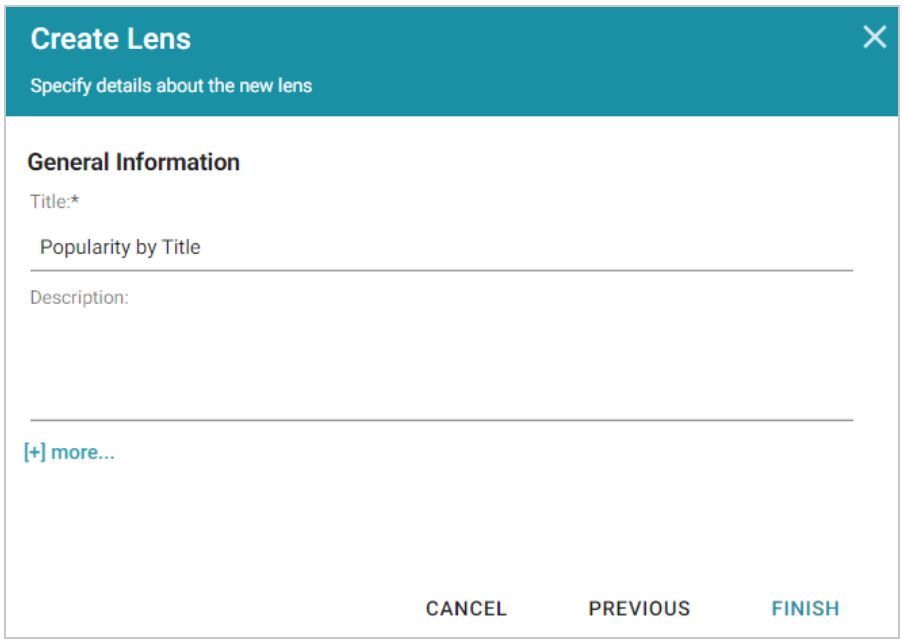
19. Now, you can filter the data and reduce the number of results that are displayed by clicking a value in the **Rating** filter. In the filter, click any one of the ratings in the list to display only the data about movies with that rating. You can click another rating to filter the table again, and so on.
20. In the Hi-Res Analytics main toolbar, click the **Dashboard** button and select **Save** to save the changes. Remember to save dashboards periodically.

Now that you can view a summary of the data in a table, it can help you determine how to further narrow or expand the results by adding, changing, or removing filters. In addition, you can experiment by adding other lenses to the dashboard to find the ideal way to display the data to answer the questions that you have. For example, the table includes a **Popularity Index** column. It might be interesting to see which movies were the most popular and how the popularity compares to the chosen rating in the filter. The next steps guide you through creating a chart lens to display movie popularity.

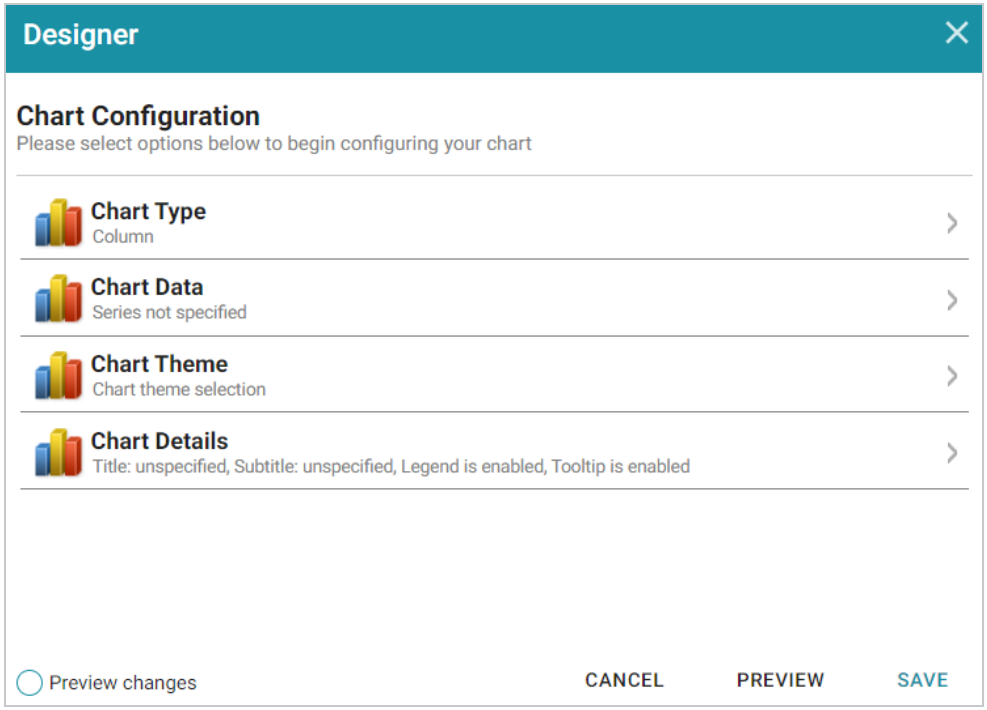
1. In the Hi-Res Analytics main toolbar, click **Lenses** and select **New**. The Create Lens dialog box opens.



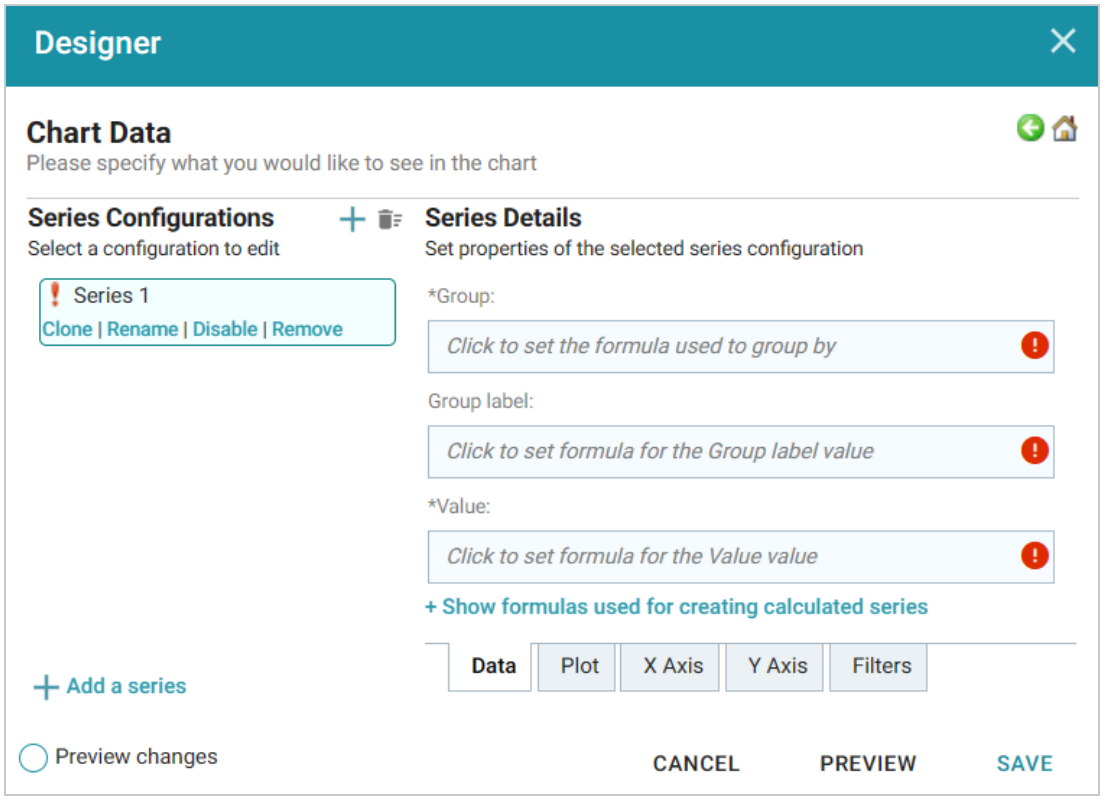
2. Select **Chart** and click **Next**. Anzo displays the Create Lens dialog box.
3. Specify a **Title** and add an optional **Description** for the new lens. For example:



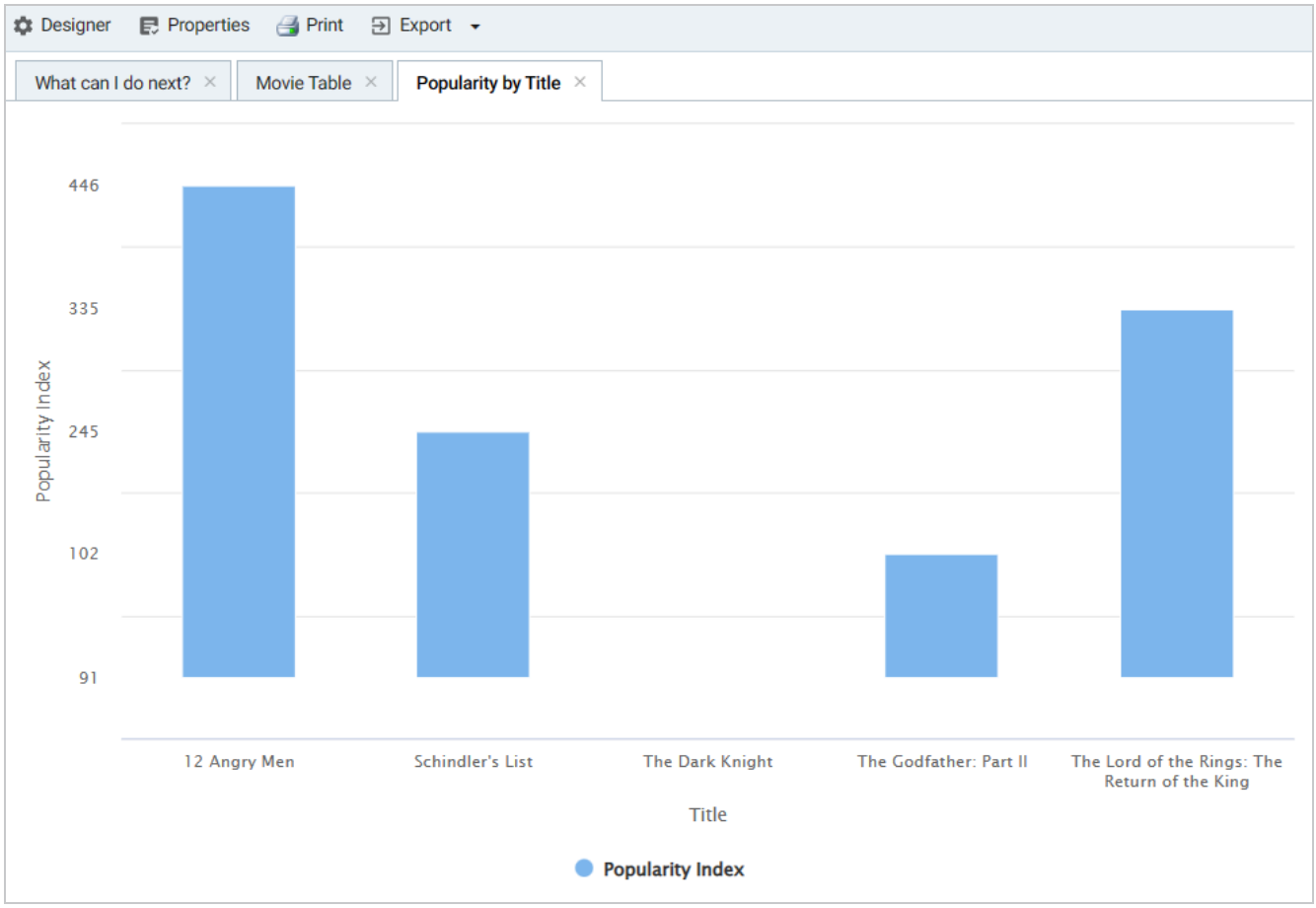
4. Click **Finish**. Anzo displays the Chart Designer, which enables you to configure the type of chart to create, such as column, pie, or line, specify the data that will populate the chart, and customize details such as the chart's theme and fonts, legend, title, and tooltip formats.



5. For this tutorial, accept the default Chart Type of **Column**, and click **Chart Data** to view the Chart Data screen and configure the data to display.

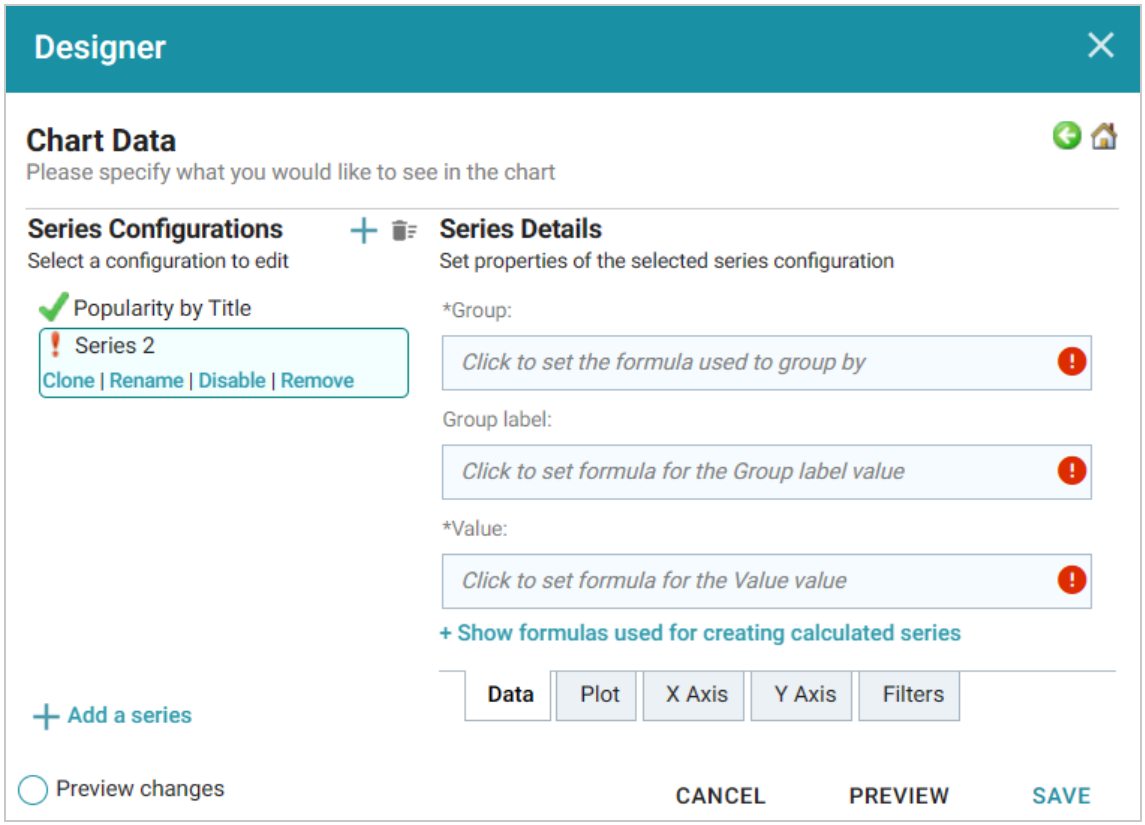


6. On the left side of the screen, click **Rename** under **Series 1**. Specify a new name in the Rename Series field, for example, "Popularity Index," and then click **OK**.
7. On the right side of the screen, click the **Group** field and select the **Title** property from the drop-down list. The Group field sets the values for the X Axis on the column chart. Then click **Close** to close the drop-down list. Anzo also populates the Group label with the Title property.
8. Click the **Value** field and select the **Popularity Index** property from the drop-down list. The Value field sets the Y Axis values. Click **Close** to close the drop-down list.
9. Click **Save**. Anzo adds the new Chart lens to the dashboard. To narrow the results that are displayed and filter on one rating at a time, click a rating in the filter. For example, by clicking **9** in the filter, the chart lists the popularity index for the titles that earned the highest rating. Of the highest rated movies, "12 Angry Men" had the highest popularity index:



> **Tip** You can hover the pointer over a bar in the chart to view details.

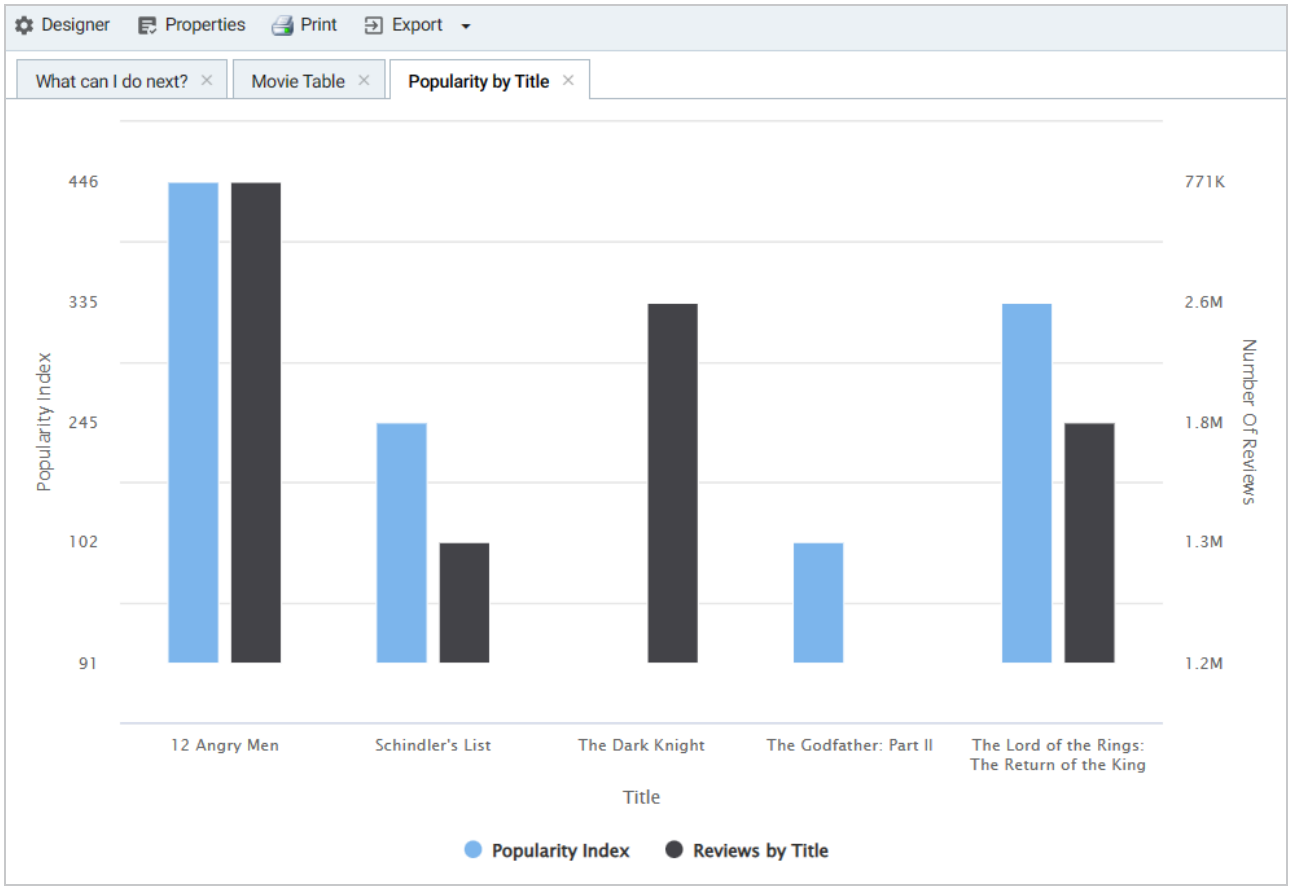
10. If you want to add more data to the Chart, such as to compare Popularity versus Reviews for each Title, follow these steps:
    a. Open the Chart Designer by clicking the **Designer** button above the lens tabs.
    b. In the Designer, click **Chart Data**.
    c. On the Chart Data screen, click the **Add a series** link at the bottom of the screen. A new series, named Series 2, is added to the list.



d. Under **Series 2**, click **Rename** and specify a name for the series, such as "Reviews by Title." Then click the **Group** field on the right side of the screen and select the **Title** property from the drop-down list. Click **Close** to close the list.

e. Click the **Value** field and select the **Number of Reviews** property from the drop-down list. Close the list.

f. Click **Save** to save the changes. The Chart is updated to display the number of reviews in addition to the popularity. For example:



> **Note**
> If you filtered the data in a previous step, the lens remains filtered on the chosen rating. You can de-select the rating in the filter or choose an alternate rating to explore the reviews versus popularity for additional titles.

> **Tip**
> For more information about working with Dashboards, see Analyzing Data with Hi-Res Analytics in the User Guide.

Related Topics

Building a Knowledge Graph with the Automated ETL Workflow