SparkGrid

This guide describes the use of the SparkGrid service for implementing graphical editing for Snowflake, MySQL, and PostgreSQL-powered relational databases. The images shown reflect the product's usage in a Snowflake environment; however, the functionality will be identical unless otherwise specified. We appreciate feedback to continue to improve your experience!

Logging In

🋞 Sj	barkGrid					
	Snowflake 🗸					
	Account:					
	Username:			Password:		
	SparkGrid uses	commands that requ	ire certain permiss	ions. By using Sp	varkGrid, you agree to our <u>Terms of Ser</u>	vice.

When you first load SparkGrid, you will be greeted with the login page. At the bottom of the page, the footer states that using the app requires certain permissions and means you agree to the app's terms of use. Click the underlined words "Terms of Use" to open 1a small modal that includes terms and conditions and privacy information. After reading these points, scroll down the modal (if applicable) and then click "I Agree." This will close the modal, allowing you to enter your login information.

Start by specifying your RDBMS of choice from the top left dropdown. After this, you will need to enter three fields to log in: your database host name (or account name for Snowflake), your username for that database, and your password. Depending on the release, you may also be asked to enter a valid e-mail address before or during this process. Once you have entered all of these, click "Log in."

If you are using MySQL or PostgreSQL, the host name is likely given by your database provider and may also be listed as the database endpoint. If you are using Snowflake and can access its web browser interface, the account name (formatted as [identifier].[region]) can often be found in your URL:

- If your URL resembles <u>https://abc12345.us-east-1.snowflakecomputing.com/...</u>, copy the text between "https://" and ".snowflakecomputing.com".
- If your URL resembles <u>https://app.snowflake.com/us-east-1/abc12345/...</u>, copy <u>us-east-1/abc12345</u>, switch the order of the two terms, and replace the slash mark with a period.

If all goes well, you will briefly see a loading symbol, and then you will be taken to the next page. However, if the login is incorrect or the connection with the database endpoint was otherwise unsuccessful, red text will appear beneath the login button that states that the connection failed and gives the likely cause of the error (such as incorrect credentials or a nonexistent account name). If this occurs, make sure that each of your credentials is in place correctly; if they are and login attempts continue to fail, the database may not be reachable at that time.

Header and Footer Elements

When you successfully log in, you will briefly see a loading screen while the app calls data from Snowflake; after this, the Files view (covered in the next section) will load by default. However, several elements will persist through all "views" of the app.

🍪 SparkGrid	ٽ	Access: Editor	Warehouse: COMPUTE V	Role: PUBLIC	, Ö
					JOHNDOE
					Options
					Log out

The header is used to provide options that will impact how the entire app is used, and provides the following options:

- A reload button: clicking this will refresh the list of databases, schemas, tables, warehouses, and roles (as available) that you can access, in addition to most table metadata. This refreshes nearly the entire interface, including any open table's contents and profile results, but does not affect table filter results, page number, or selected ID columns.
- An access level dropdown: changing this will switch between Editor mode (which offers full
 access to editing tables and data therein) and Read Only mode (which allows one to view data,
 but not to delete or modify it). Due to the table loading process and the editing-driven nature of
 some features, any open table will close instantly upon changing this option and some views or
 features discussed later will disappear.
- A "person" icon: click this to bring up a box with three items: your username, an "Options" button, and a "Log Out" button. To close the box, click anywhere outside it.
 - Click the "Options" button to open the Options modal. The options are divided into two sections, "Grid" and "Appearance." "Grid" includes a "Grid rows per page" dropdown, which allows one to choose how many rows are shown in the Grid view at one time (see "The Grid View Part 2"); it also contains an "Expand grid cells to fit content" checkbox, which automatically expands columns in the Grid view for convenience when cells contain longer text. "Appearance" includes a "Glass effects" checkbox, which offers a minor visual difference, and a "Theme" dropdown that lets users choose between the default dark theme and an in-beta light theme.
 - At the bottom of the modal, you can click the "Apply" button to apply these changes for your browser session (note that the Grid view may need to be reloaded to see changes), or "Cancel" to leave the changes unapplied.
 - Click the "Log Out" button to end your session and return to the login screen.

Depending on your database system, some additional dropdowns may appear:

- A warehouse dropdown (Snowflake-only): changing this will choose which warehouse you will use to process the changes. For basic changes or new users, this is unlikely to affect the app.
- A role dropdown (Snowflake-only): changing this will reload mass amounts of the page, as it can change which warehouses, databases, schemas, and tables you can access.

 A database dropdown (PostgreSQL-only): PostgreSQL treats databases differently from some other systems and only permits working with one database at a time. SparkGrid addresses this by treating PostgreSQL databases similar to how it treats Snowflake roles, reloading available schemas and tables whenever a new database is selected.



The footer allows you to switch between "views," which resemble web pages, but offer a more integrated experience. SparkGrid contains four views, each one accessible by clicking on an icon on the footer: from left to right, these are the Files view, the Grid view, the Upload view, and the Toolkit view. The active view will be indicated by a different gradient behind its icon (the Grid view in the image), which will generally blend visually into SparkGrid's background.

The Files View



As mentioned earlier, the Files view loads by default once the loading screen is finished fetching data. This view's name comes from its resemblance to the "folder/file" structure of most operating systems, and it offers a way to manage tables at a more structural level than the grid view. The Files view is divided into two halves: a structure panel on the left side of the screen, and a details panel on the right.

The Structure Panel

When the Files view first loads, the structure panel will appear as a series of boxes: one for each database or each schema in the account, depending on the system used. Clicking on one of these boxes will expand it to reveal smaller boxes; for MySQL and PostgreSQL, these represent the parent schema's tables and/or views, while in Snowflake they represent the parent database's schemas, which similarly expand when clicked to reveal a third layer representing tables and/or views.

The structure panel offers several ways to manipulate the tables:

- Right click a table to bring up a context menu offering three options: Open as Grid, Rename, and Delete.
 - Click "Open as Grid" to either open the table in the Grid view or a display a column selector (see "The Grid View Part 1", paragraph 2).
 - Click "Rename" to enter a new name for the table, then click "Confirm" to enter the new name or "Cancel" to cancel the rename.
 - Click "Delete" to bring up a confirmation box asking if you want to delete the box. Click "Confirm" to

SITE_LIVE							
DEV_DATA]					
EXAMPLE_SCHEMA (9 tables)							
INFORMATION_SCHEMA (empty)							
PUBLIC (10 tables)							
BIG_TABLE		=					
CLONE							
IVT_DATA_PRFL_RPT		ן (≡ר					
IVT_DATA_PRFL_TBL_LST	Open as Grid						
IVT_PARM_TBL	Rename						
	Delete	\square					

proceed with the deletion or "Cancel" to cancel the operation.

- Click and hold the mouse button down on the hamburger icon at the right of a table to start dragging it. You can then drag the table to another schema or database in view.
 - If you drag the table into a schema and release the mouse button, a confirmation box will appear; click "Confirm" to proceed with the move or "Cancel" to cancel it.
 - If you drag the table into a database in Snowflake and release the mouse button, a dropdown will appear asking which schema within that database you would like to place the table in. Select a schema of your choice from the dropdown, then click "Confirm" to proceed with the move or "Cancel" to cancel it.
 - If you drop the table back over the existing schema (or any place on the screen outside the structure panel's contents), the operation will cancel.
- Click a table to display it in the Files view's details panel.

CUS									
Table									
	Database: SNOWFLAKE_SAMPLE_DATA								
	 Schema: TPCH_SF1 								
	Created: 2022-03-02 13:29:42	780 -0800							
	Rows: 150000								
	Comment: Customer data as defined by TPC-H								
	Load as Grid	Load with Filter							

The Details Panel

Once you click a table in the structure panel, table details will appear in the box to the right of that panel. Here, you can view the table's name, its parent database and/or schema, when it was created, how many rows it contains, and its comment where applicable.

 Click a property denoted with a pencil icon (at present, either the table name or comment property) to open a small input.



Enter a new value, then click the "Commit to Database" button to the right of the input to proceed or "Cancel" to cancel the operation.

• Click the "Load as Grid" button to open the table in the Grid view or display a column selector (see "The Grid View Part 1", paragraph 2).

• Click "Load with Filter" to similarly load the table or display a column selector, but specifying one or more filters or conditions for the data to be displayed (see "The Grid View Part 2," "Header Buttons" subsection).

Read Only Tables and Views

You may encounter schema entries where some of these options will not appear. These are views and read only tables, which your Snowflake account may view but not edit. Key differences may include:

- The tables may be tinted red (to indicate a read only table), or blue with the word "(view)",
- Hamburger icons may not appear at their right side,
- Right clicking them may no longer include "Rename" or "Delete" options,
- The details panel's properties no longer include a pencil icon and cannot be changed, and
- The details panel's "Open as Grid" and "Load with Filter" buttons state that the table or view will be opened as read only.

If you are working in Read Only mode (see "Header and Footer elements"), hamburger icons and pencil icons will similarly disappear, and right clicking a table in the structure panel will only offer the "Open as Grid" option.

The Grid View Part 1: Selecting the Table



If you click the grid icon in the footer after logging in, you will see a small form to select a table. The dropdown fields are dynamically updated with a top-down hierarchy; this means a database>schema>table hierarchy for Snowflake and a schema>table hierarchy for MySQL and PostgreSQL. Selecting a higher level will populate the next dropdown with the higher level's contents: in a Snowflake environment, for example, you must select a database to select a schema, and you must select both a schema and database to select a table. If you cannot find the database, schema, or table you are looking for, try selecting a different role or database from the available header dropdowns, then check if the desired choice appears on the refreshed list.

Once you have selected your database, schema, and table as available, click "Load Data" or "Load with Filter" to bring up on the table. If the table has not been loaded before during this session, a column selector will appear to let you specify which column(s) to group rows by. Depending on how you want to use this table, these columns will work as groupers or as identifiers; if one of the columns is an automatically incrementing identifier, it will be pushed to the top of the list (otherwise sorted alphabetically) to enable a quick way to edit on a row-to-row basis. All cells within these columns will be treated as read-only, with the exceptions of non-autoincrementing columns in a new row. These columns will be remembered for the session, even after refreshing the interface; however, you can easily bring up this column selector again in the next section.

If you clicked the "Load with Filter" option, a filter dialog will appear, allowing you to add various simple conditions to what data is fetched. After setting and applying the filters (see "The Grid View Part 2," "Header Buttons" subsection), the table will appear as a grid.

Please select one or more columns to group or identify table rows by. While editing the table all selected columns will be read only unless changed later.	
■ IVT_ROWID (fixed; autoincrement)	
■ IVT_LOAD_TS (timestamp_ntz)	
■ SRC_445_CALENDAR_DAY_NUM_OF_MO_ (text)	
■ SRC_454_CALENDAR_DAY_NUM_OF_MO_ (text)	
Use Selected Columns Cancel	

The Grid View Part 2: Working with the Table

	SparkGri	d			J	Access: Editor 💊	Warehouse	: Compute ~	Role: PUBLIC	, Ö
corpo	oate_data.public.calendar	_dimensions						Choose ID	Filter	Change Table
	Group/ID Details	SRC_WK_NUM	SRC_YR_MO_NUM	SRC_DATE_NUM	SRC_DATE	SRC_CLNDR_QTR	SRC_MO_NUM	SRC_MO_NAME	SRC_MO_SHRT_	N/SRC_DAY_NUM_
1	Distinct row		190001	18991231	12-31-99	QTR 1	1	January	Fred	Fred
2	Group (5 matching rows)		190101							
3	Detail (in 5-row group)	1	190101	19010105	01-05-01	QTR 1	1	January	Jan	5
4	Detail (in 5-row group)		190101	19010101	01-01-01	QTR 1		January	Jan	
5	Detail (in 5-row group)	1	190101	19010102	01-02-01	QTR 1	1	January	Jan	2
6	Detail (in 5-row group)	1	190101	19010104	01-04-01	QTR 1	1	January	Jan	4
7	Detail (in 5-row group)	1	190101	19010103	01-03-01	QTR 1	1	January	Jan	3
8	Group (3 matching rows)	1	190301						Edited	
9	Detail (in 3-row group)	1	190301	19030101	01-01-03	QTR 1	1	January	Jan	1
10	Detail (in 3-row group)	1	190301	19030102	01-02-03	QTR 1	1	January	Edited	2
11	Detail (in 3-row group)	1	190301	19030103	01-03-03	QTR 1	1	January	Jan	3
12	Group (2 matching rows)	1	190401							
13	Detail (in 2-row group)	1	190401	19040102	01-02-04	QTR 1	1	January	Jan	2
14	Detail (in 2-row group)	1	190401	19040101	01-01-04	QTR 1	1	January	Jan	1
15	Group (7 matching rows)	1	190501							
16	Detail (in 7-row group)	1	190501	19050107	01-07-05	QTR 1	1	January	Jan	7
17	Detail (in 7-row group)	1	190501	19050103	01-03-05	QTR 1	1	January	Jan	3
18	Detail (in 7-row group)	1	190501	19050102	01-02-05	QTR 1	1	January	Jan	2
19	Detail (in 7-row group)	1	190501	19050101	01-01-05	QTR 1	1	January	Jan	1
< Pre	evious Page 1 of 10	3 Next >	Jump to page		Go Of OC OF		SQL Preview	v:	Cor	nmit
									<u>j</u> rl	

Once the table loads, the table select form is replaced by a grid reflecting the table and new header and footer buttons.

Unless the table is read only (see "The Files View") or you are loading the table in Read Only mode (see "Header and Footer Elements"), one can interact with the table in a similar way to how one interacts with any simple spreadsheet: Click a cell and start typing to replace the text in that cell; right-click a column name to add, remove, or rename a column; or right-click a row number to add or remove a row. The table itself is powered by the Jspreadsheet CE library; you can view the library's full documentation here: https://bossanova.uk/jspreadsheet/v4/docs/. If the table is read only, it cannot be edited (among other changes, the "Commit" button and "SQL Preview" box will be absent to make this clear); however, one can still filter, sort, and navigate the table. SparkGrid will usually recognize text, numeric, date/time, and Boolean (true/false) values, and will set each field to its corresponding type and character limit.

Grouping

SparkGrid tracks changes to the table by checking the column(s) you specified in the column selector. The names of these columns appear in italics, are typically the furthest regular columns to the left, and—unless the loaded table is a read-only table or view—the columns' fields will be a slightly different shade from the rest of the table do clarify that they cannot be edited regularly. (The only exception to this final part is new rows, and even these may only be edited if the column isn't an automaticallyincrementing column.) For the purposes of this documentation, these will be called ID columns.

If two or more rows have ID columns with the same values, these will be grouped together with a special "grouper row" above them, which doesn't exist in the database but serves as a visual convenience. Editing any row in one of these groups, or editing the grouper row itself, will visually change all cells in that column of that group to clarify that editing one of these values will change all of them in the database; similarly, deleting one row of a group will hide the entire group. The number of rows in the group can be seen in the leftmost "grouper column;" if the row's ID columns are unique, however, the column will simply label it with "Distinct row." If you want to collapse a particular group, click the group's "corner"—the field on the grouper row's grouper column—to hide all rows with the same ID columns except for the grouper row. Clicking it again will restore that group.

Header Buttons: Changing ID Columns, Filtering, Sorting, and Returning to Table Select Below the header are three buttons in the Grid view: "Choose ID," "Filter" and "Change Table."

Click "Choose ID" to bring up the column selector from Part 1 and change which columns you would like to identify or group rows by.

Click "Change Table" to go back to the previous table select form. If you have made changes to the grid and have not yet committed them to the database, you will be warned that these changes will be lost; click "Yes" to proceed with moving back to the table select form and "Cancel" to cancel the operation.

Click "Filter" to open a form offering a filter tool and a sort tool to refine table results. The filter tool includes three fields by default: column, filter type, and value. The filter type has eight total options, with an "is" and "is not" variant of four factors:

- 1. Equals—used for exact values, such as finding data from a specific city,
- 2. Includes—used for parts of any column, such as finding customers' full names with the same last name,
- 3. Is between—used for finding a range of numbers, such as finding data between two years, and
- 4. Is null—used for finding whether a data field is empty or not.

"Is between," "Is null," and their negative counterparts will change the form: "Is between" adds a second text field, with the first text field acting as a minimum value and the second acting as a maximum. "Is null," on the other hand, disables the text field, as no additional input is needed. Please note that these filters are case-sensitive.

The sort tool is simpler than the filter tool, but works on a similar principle: it includes one field to select a column to sort by, and a choice of whether the sort order should be ascending (lowest value to

SparkGrid Documentation



highest) or descending (highest to lowest). If you want to sort the table without filtering it, ensure that the filter value is set to "N/A (No filter)."

By default, only one filter and one sort option are visible; however, below either field is a small "plus" button. Click this button to add another filter or sort layer, respectively. Filters will work together regardless of their order, while sorting will follow a hierarchy from top to bottom (for instance, sorting a list of purchases by NAME and then DATE will primarily sort by NAME and only by DATE when two or more NAME fields are the same). To the left of every additional filter or sort layer is a "minus" button; just as the "plus" button adds a layer, clicking a layer's "minus" button will remove it.

Click "Cancel" to close the modal without affecting the grid; alternatively, click "Apply" after setting the filter and/or sort options to begin the process. This will lose any unsaved table changes, so please ensure that you have committed all changes first (covered in the "Previewing and Committing Changes" subsection).

If the table contains no rows that match the requested filter(s), a small window will appear stating that the filter failed. If rows were found and the filter was successful, however, a filtered and/or sorted grid will replace the default one and a new "Remove Filter" button will appear to the left of the "Filter" and "Change Table" buttons. Clicking "Filter" again will allow you to replace the current filter with a separate one (the filters do not stack), while clicking "Remove Filter" will reload the grid by default.

Navigation



Many database tables are too large to realistically load at once; SparkGrid resolves this by splitting tables into pages. The default number of rows for each page is 500, but this number for the session can be changed from the Options modal (see "Header and Footer Elements") in incremental steps ranging from 50 to 5,000. At the footer, you will see text showing both the current page and the total number of pages in the table. Click "Previous" or "Next" to move to the previous or next page. (If one cannot move further in a given direction, the button corresponding to that direction will be disabled.) You can also enter a page number in the text input to the right of those buttons, and then press "Go" to move to that page.

Moving to another page will lose any changes that have not been committed to the database. If you have any uncommitted changes and try to change your page (or to go back to the table selector through the header's "Change Table" button), a brief warning screen will appear asking if you want to continue. Pressing "Yes" will proceed with your



previous action and will lose the changes you made earlier, while pressing "Cancel" will close the box and allow you to finish your work.

New Columns



To add a new column, begin with Jspreadsheet's standard column addition procedure: right-click on one of the grid's columns, then click "Add a new column before" or "Add a new column after." (The order here is merely visual; the database itself adds the column to the right using a standard ADD COLUMN command.) This will raise a popup modal with at least two fields: one for the new column's name, and a dropdown for one of six data types. Depending on the type, other fields may appear:

- The default Varchar/Text field, which uses the VARCHAR data type, has a numeric field to specify the maximum allowed length of the new column.
- The Integer field has a dropdown to specify a subtype: INTEGER, BYTEINT, TINYINT, SMALLINT, or BIGINT.
- "Number with Decimal" uses the DECIMAL data type and offers two numeric fields: one to specify the column's precision (max length, in essence) and another for its scale (how many of the previous number come after the decimal point).
- True/False (BOOLEAN), Date, and Timestamp all lack options.

Once you are finished, click "Add column" to proceed with the addition. You can cancel the addition and return to the Grid view at any type using the Cancel button.

Previewing and Committing Changes

SparkGrid does not require any knowledge of SQL, but it provides an optional feature to view how the changes will reflect for the database. To the right of the navigation buttons is the text "SQL Preview:" and a box icon. Click the text or the box to expand this text into a small window just above the "Commit" button. To close the window, click "SQL Preview:" or the symbol to its right (which should now resemble a "minimize" button) a second time. This window will usually stay open or closed based on where you last left it in your browser session.



The preview window comes with a straightforward interface: a blank text box, and a button with the text "Generate SQL Commands." If you click this button without making any changes (or directly after committing), the text box will merely show "(No changes found)." If changes have been made, however, the text box will populate with SQL equivalents to the changes made to the table.

To avoid potential conflicts, SparkGrid organizes your changes into the following order:

- 1. Add and/or delete columns,
- 2. Rename columns,

- 3. Insert and/or delete rows, and
- 4. Edit cells.

To commit your changes, press the "Commit" button at the bottom right corner of the Grid view. Once you have committed your changes, the table history will reset and you will not be able to undo any changes directly, so ensure you are happy with your changes first. Note that some spreadsheet changes, such as sorting an already-loaded page or changing column width, do not directly affect the database and so do not reflect in the final commit.

The Upload View

CORPORATE_DATA V PUBLIC		✓ MARKER_RESU	TS Browse	narkers.csv	Create Table		
Please choose a data type and further information where applicable for each column, or leave them as the default values below (the larger of 255 or the largest value found in the file, up to 16777216 characters). DECIMAL columns with a precision higher than what the database supports (38 digits) will be converted to FLOAT columns.							
id	type	resolved	text	author	users		
Туре:	Туре:	Туре:	Туре:	Туре:	Type:		
Integer	✓ Varchar / Text	✓ True/False	✓ Varchar / Text ✓	Varchar / Text 🗸 🗸 🗸 🗸 🗸 🗸	Varchar / Text		
Integer type:	Maximum length:		Maximum length:	Maximum length:	Maximum lenç		
INTEGER	✓ 255		255 ᅌ	60 🖸	255		
<							

The Upload view is based on creating entirely new tables from comma-delimited CSV files, which can be easily exported from mainstream spreadsheet programs or created with a basic text editor. The interface resembles the Grid view's table select form; the two main differences are that the table dropdown is now a text input field—allowing one to input a custom name for the new table—and a file upload field. These instructions show how to upload a new table from a CSV file:

- 1. Select a database and/or schema from the dropdown.
- 2. Enter a new table name. The name must be a valid SQL: It must not start with a digit or contain spaces, periods, or similarly invalid characters.
- 3. Click the next field to browse for a file on your local machine. This file must end with ".csv" and be correctly formatted. Navigate to the folder containing your file, then double click the file or click the file once and then click the "Open" button.
- 4. Set column data types and further details (optional).
- 5. Click "Create Table."

After selecting a CSV file, SparkGrid will parse the file and display a list of the columns, offering a way to set data types for each. Outside column names, each column's settings are identical to the New Column box in the Grid view ("The Grid View Part 2", "New Columns" section). By default, each column will be viewed as a text string 255 characters long, but will expand up to Snowflake's maximum supported length if some of the data in the CSV cell exceeds this.

As the wording suggests, "Create Table" will automatically process the CSV data; if the data is valid and SQL-compatible (meaning the columns follow most or all of the same rules used for table names in step

3), a new table will be created in the database. Once the table is fully processed, the ID/Grouper modal described earlier will open ("The Grid View Part 2", "Grouping" section); clicking "Cancel" here will cancel the table load, but will keep the created table. After you confirm which columns would want to group or identify your rows by, the Grid view will open your new table, where you can refine and edit the data as you see fit.

The Toolkit View

While the previous views listed are more focused on visualizing standard SQL commands, the Toolkit view allows for more advanced forms data evaluation and manipulation, primarily on a table- and column-level. The Toolkit View is the newest and most subject to change of SparkGrid's features, so this text is especially subject to updates as new features are added.

Profiling Table Data

Among this view's features is the ability to generate a data profile for a table. This feature gathers data about the content of each column in the profiled table, measuring (among other values):

- How many values in the column are null,
- The column's average value (if the column's data type is set to be a number),
- The minimum and maximum string lengths in the column,
- The minimum and maximum values in the column (number or otherwise), and

•	How many rows qua	lify as valid dates,	timestamps, and numbers.
---	-------------------	----------------------	--------------------------

CORPORATE_DATA		PUBLIC			CALENDA	R_DIMENSIONS		V Load Colu	umns Profile	Table
CORPORATE_DATA.PUBLIC.CALENDAR_DIMENSIONS										
Column Name	NOT NULL Count	Approx. Distinct	Max Text Value	Min Text Value	Max String Length	Min String Length	Avg. String Length	Default, NULL, or Space	Valid Number Count	V (DD
IVT_LOAD_TS	51135	1	2022-10-10 16:53:49.259	2022-10-10 16:53:49.259	23	23	23.000	0	0	
IVT_ROWID	51135	51135	9999		5		4.783	0	51135	
SRC_454_MO_NUM	51135	12	9	1	2	1	1.253	0	51135	
SRC_MO_SHRT_NAME	51135	12	Sep	Apr	3	3	3.000	0	0	
SRC_454MO_YR	51135	1680	Sep-2039	Apr-1900	8	8	8.000	0	0	
SRC_4_4_5_MO_NUM	51135	12	9				1.253	0	51135	
SRC_4_4_5MO_YR	51135	1680	Sep-2039	Apr-1900	8	8	8.000	0	0	
SRC_CLNDR_QTR	51135	8	Qtr 4	QTR 1			5.000	0		
SRC_DATE	51135	36525	12-31-99	01-01-00	8	8	8.000	0	0	

To use the data profiling tool, begin by selecting a table from the view's dropdown lists. As with the Grid view's table selector, this must be done through a top-down hierarchy. After setting these, click on "Profile Table" to generate the data profile.

Depending on the table size, profiling can take anywhere from a few seconds to several minutes, during which the application cannot be interacted with. If you suspect the application is stuck, it may be helpful to check your database provider's monitor/activities section if available to verify that the command is still running. Note that some very large tables may cause a timeout error. If this occurs, the query may still be running in the database; in the case of Snowflake, we recommend checking the query history through the Snowsight web interface and canceling the statement if it is still running. Since the results are evaluative by nature and do not use a traditional SQL table, finished profiles appear in a read-only table separate from the Grid view.

Profile Results

At this time of writing, these are the profile's result columns and their functions:

- NOT NULL Count: How many cells in this column are null.
- Approx. Distinct: How many cells are unique, as opposed to duplicates.
- Max Text Value: The column's largest value (numerical or text).
- Min Text Value: The column's smallest value (numerical or text).
- Max String Length: The longest value in the column (treated as text).
- Min String Length: The shortest value in the column (treated as text).
- Avg. String Length: The average length of the column's values (treated as text).
- Default, NULL, or Space: How many cells in this column are null, " "/space only, or the column's default value.
- Valid Number Count: How many cells can be viewed and processed as numbers.
- Valid Date: Three columns showing how many cells can be viewed and processed as dates that follow the DD/MM/YYYY (12/25/2023), MM/DD/YYYY (25/12/2023), and YYYY/MM/DD (2023/12/25) date formats, respectively.
- Valid Timestamp: How many cells can be viewed and processed as timestamps.

Editing Columns

The Toolkit view also provides a way to view and edit column metadata. Select a table with the same hierarchical dropdown system described previously, then click "Load Columns." This will load a table separate from both the previous profile table and the Grid view, showing each column's name and several properties:

- Type (the column's data type),
- Nullable (whether the column allows NULL values),
- Maxlength (the maximum length of text-based columns),
- Precision (the maximum length of a number-based columns),
- Scale (how many digits in the precision should be after a decimal point), and
- Comment (a generic text comment for the applicable column).

Each of these factors can be edited, although some of these options are limited by the column's data type: "maxlength" can only be modified in VARCHAR columns, while "precision" and "scale" can only be specified in DECIMAL columns.

The "type" column contains the same six data types as the Grid and Upload views' column selectors ("The Grid View Part 2", "New Columns"), as well as the column's original data type if it wasn't one of the default options.

After you have changed the columns to your preference, press the "Commit Changes" button to bring up a box that will inform you about any issues with the changes: changing a column to INTEGER, for example, would make it incompatible with VARCHAR values. Note that minor changes, such as rounding off decimals when decreasing a column's scale or converting a column to INTEGER, may not be listed here. After reviewing the changes, click "Cancel" if you wish to go back to revise the change list or "Proceed" to commit the changes. When changing a column's data type, cells with values incompatible with the column's new data type will usually be set to NULL (or its closest equivalent, if the column is not nullable). However, in some circumstances SparkGrid will attempt to convert the values:

- Changing a timestamp column to a numeric type ("Integer" or "Number with Decimal") will interpret the time values as UNIX timestamps (a number measured in seconds from Jan 1, 1970) and vice versa;
- Changing a Boolean column to a numeric type will change TRUE and FALSE values to 1 and 0, respectively;
- Changing any column to a Boolean type will change all cell values to TRUE unless a cell is NULL or matches certain text values like "0," "no," or "false" (non-case sensitive).

Error Handling

SparkGrid's error messages aim to be as detailed as possible, with a user-friendly description of what

went wrong at the top and more technical details for SparkGrid's support team at the bottom. Usually, an error will consist of three aspects:

- A user-friendly error message,
- One of the following:
 - A numeric code at the

header (usually if the error was SQL-related), or

• A "stack" object below the error message that describes specific lines of code that went wrong (usually if the error was Node.js-related), and

An error occurred (001003)

SQL compilation error: syntax error line 1 at position 65 unexpected '-'

Full context for debugging: {"function":"renameTable","purpose":"Rename table

from context menu","request": {"role":"PUBLIC","type":"table","name":"ABCDDD","ext":"CMS_DATA.PUBLIC","c hangeto":"ABCD-&DD","warehouse":"COMPUTE_WH"},"misc":{}}

• A JSON-styled "Context" section that gives the request's front end details, including the function name and the request itself. Depending on the request, this section may be extremely long, but it is mainly meant for SparkGrid's team and may not be helpful for most users.

OK

The cause for these errors may come down to human error (the image uses an example of renaming a table in an incorrect format), but it could also point to an issue with SparkGrid. If you encounter such an error, especially on a consistent basis, please send an email to sparkgridsupport@sparksoftcorp.com with the contents of the error message (sans any sensitive information the context section may have) and whatever other details you think would be helpful.

Appendix A: Technical Details

SparkGrid is built with an <u>EJS</u> frontend that makes extensive use of HTML, CSS, JavaScript, and the JavaScript library <u>iSpreadsheet CE</u>, which is released by a third party under the MIT license. Its backend is primarily built with <u>Node.js</u> and <u>ExpressJS</u>. It can be customized for use in a serverless environment, but it can also work on a standard server or even a local Node execution. Its SQL commands are specialized for Snowflake's syntax, and the application itself is stored in AWS' ECS service.

Appendix B: Requirements

To fully take advantage of SparkGrid, you will need the following:

- Hardware:
 - A non-public desktop, laptop, or workstation computer with a minimum screen resolution of 1024x768 or 1280x720 pixels
 - A stable Internet connection
- Software:
 - A Snowflake, MySQL, or PostgreSQL account that your user account has sufficient authorization for
 - o A reasonably up-to-date web browser (from at least April 2022) with JavaScript enabled