ALIBABA CLOUD

Alibaba Cloud

DataWorks

DataAnalysis

Document Version: 20210125

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.

2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.

3. The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.

4. This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization, company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

5. By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.

6. Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	**Danger:** Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	**Warning:** Restarting will cause business interruption. About 10 minutes are required to restart an instance.
Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	**Notice:** If the weight is set to 0, the server no longer receives new requests.
Note	A note indicates supplemental instructions, best practices, tips, and other content.	**Note:** You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click **Settings**> **Network**> **Set network type**.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
`Courier font`	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.Overview

DataWorks DataAnalysis allows you to analyze, edit, and share data online.

Benefits

Compared with offline data analysis, DataAnalysis has the following benefits:

- High efficiency: DataAnalysis analyzes data in a database by using online data analysis tools such as pivot tables. For example, you can create a pivot table for the September partition of a user profile table. Then, you can update the source data with the October partition and reuse the configuration of the pivot table to avoid repeated operations.
- High capacity: DataAnalysis efficiently analyzes large amounts of data with the help of compute engines.
- Shared data: DataAnalysis can analyze data obtained from databases of different business systems. DataAnalysis allows you to export data to MaxCompute tables. It also allows you to share data with specified users and grant them access permissions. Therefore, data can be shared between different systems and different users.
- High security: DataAnalysis allows you to analyze data online without downloading data. It also allows you to manage the permissions that allow users to analyze and share data.

Features

- **Workbook**

  Workbooks are the core feature of DataAnalysis. A workbook is a workspace in which you can obtain, explore, and analyze data. A workbook is in the form of an online table and offers common table features. After you import data from data stores or import local data to workbooks, you can perform data pivoting and profiling. For more information, see Workbook.

- **MaxCompute table**

  The dimension table feature allows you to create MaxCompute tables without writing SQL code and collaboratively edit MaxCompute tables with other users online. The dimension table feature also allows you to import data to MaxCompute tables in a visualized manner. For more information, see MaxCompute table.

  Assume that data analysts and operations engineers need to manually maintain an online MaxCompute table. If the MaxCompute table is not created by using the dimension table feature, the following procedure must be performed whenever data changes occur:

  i. Operations engineers describe their requirements to data analysts.

  ii. Data analysts submit the requirements to developers.

  iii. Developers edit SQL statements by using DataWorks, create an Excel table, synchronize data to the Excel table, and then deliver the Excel table to operations engineers.

  If the dimension table feature is used, operations engineers can create, modify, and save a MaxCompute table in DataAnalysis. This improves efficiency.

- **Report**

  You can create and design reports by dragging and configuring controls without executing SQL statements. For more information, see Report.

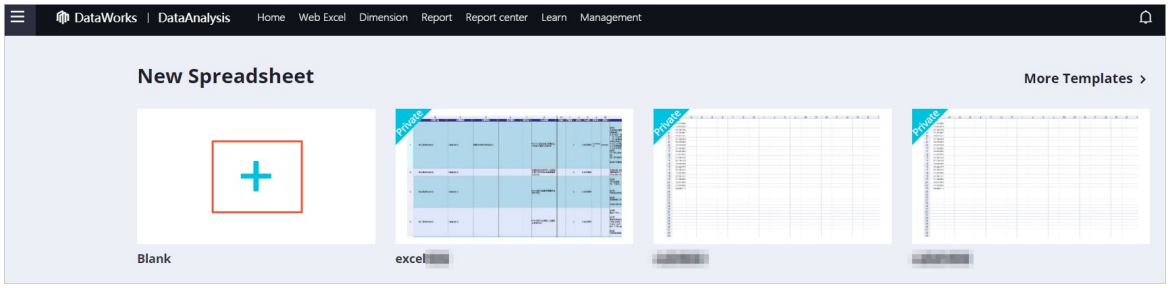
2.Workbook

2.1. Create and manage a workbook

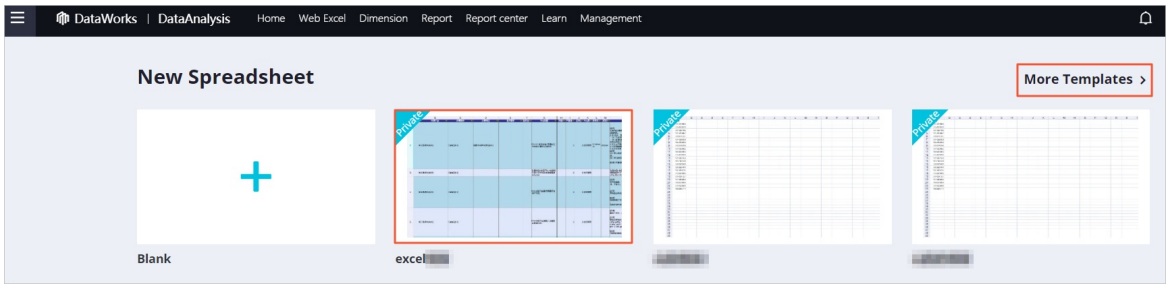
Before data analysis, you must create a workbook to store the data to be analyzed. This topic describes how to create, view, and manage a workbook.

Create a workbook

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. On the **Web Excel** page, click the icon in the **New Spreadsheet** section.



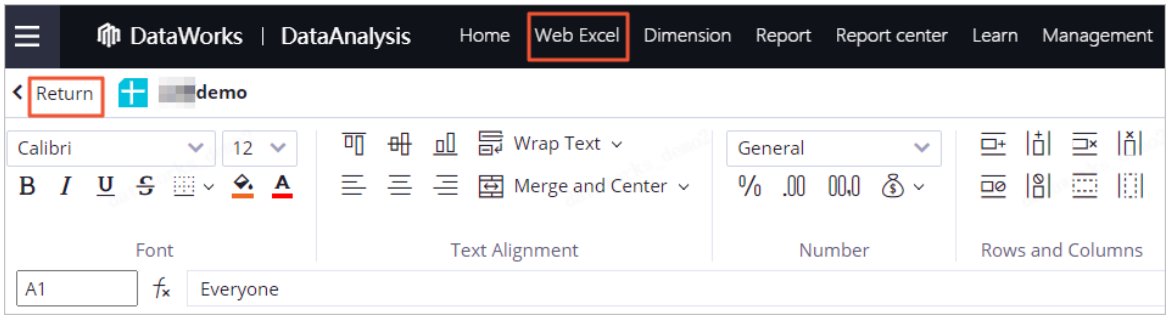
If you have templates under your account, you can click a template to create a workbook based on the template. To view all templates under your account, click **More Templates** in the upper-right corner of the New Spreadsheet section. For more information about how to create and apply a template, see Manage a workbook template.



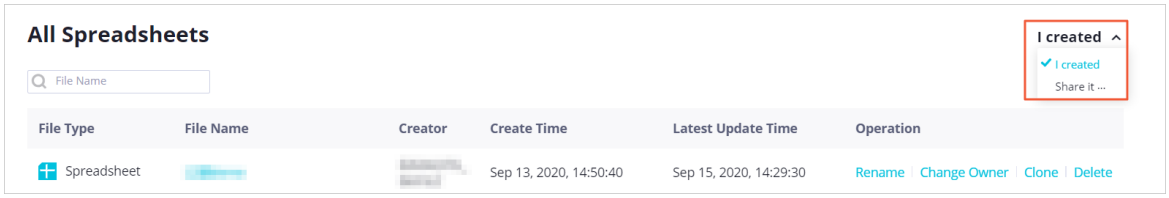
5. In the **New spreadsheet** dialog box, enter a name in the **File Name** field.
6. Click **OK** to go to the workbook editing page. For more information, see Analyze data.

View and manage a workbook

1. On the workbook editing page, click **Return** in the upper-left corner or **Web Excel** in the top navigation bar to go to the Web Excel page.



2. In the **All Spreadsheets** section of the **Web Excel** page, select **I created** or **Share it with me** from the drop-down list in the upper-right corner to view the workbooks in the corresponding category.



   You can also share workbooks with specific users. For more information, see Share a workbook.

3. Click the file name of a workbook to go to the workbook editing page.On the Web Excel page, you can perform the following operations to manage a workbook:
   - To rename a workbook, perform the following steps: Find the workbook and click **Rename** in the Operation column. In the **Rename** dialog box, enter the new name in the **File Name** field and click **OK**.
   - To change the owner of a workbook, perform the following steps: Find the workbook and click **Change Owner** in the Operation column. In the **Change Owner** dialog box, select an owner from the New Owner drop-down list and click **OK**.
   - To clone a workbook, perform the following steps: Find the workbook and click **Clone** in the Operation column. The cloned workbook appears in the workbook list. The name of the cloned workbook contains the _copy suffix.
   - To delete a workbook, perform the following steps: Find the workbook and click **Delete** in the Operation column. In the **Delete** message, click **OK**.

What to do next

After you create a workbook, go to the workbook editing page and import data to the workbook. For more information, see Import data to a workbook.

2.2. Import data to a workbook

After a workbook is created, you can write data to the workbook for data analysis. You can also import data from a data store or import local data to the workbook. This topic describes how to import data to a workbook.

Prerequisites

A blank workbook is created. For more information, see Create a workbook.

Limits

The **Query mode** feature supports the following data store types and regions:

- Supported data store types: MaxCompute, MySQL, PostgreSQL, DRDS, SQL Server, Oracle, AnalyticDB for MySQL V2.0, AnalyticDB for PostgreSQL, Hologres, EMR Hive, and EMR Spark SQL.
- Supported regions: China (Shanghai), China (Beijing), China (Hangzhou), China (Shenzhen), China (Chengdu), China (Hong Kong), China (Zhangjiakou), China North 2 Ali Gov, Singapore, Indonesia (Jakarta), and Japan (Tokyo).

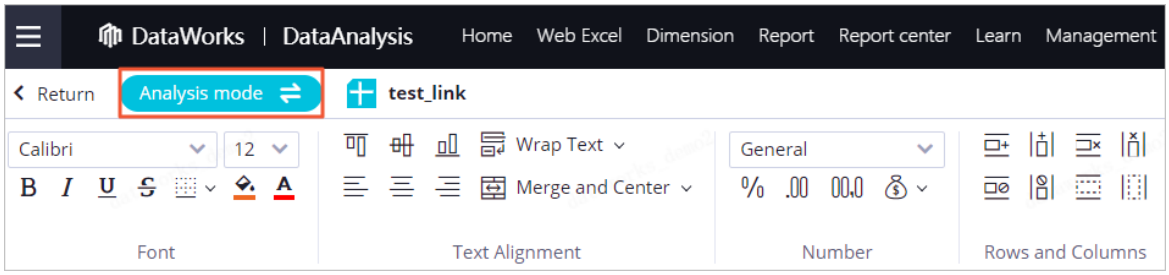
Go to the workbook editing page

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the **All Spreadsheets** section of the **Web Excel** page, click the name of the workbook that you want to edit in the **File Name** column to go to the workbook editing page. If you create a workbook in this step, the workbook editing page appears after the workbook is created. For more information, see Analyze data.

Import data from a data store

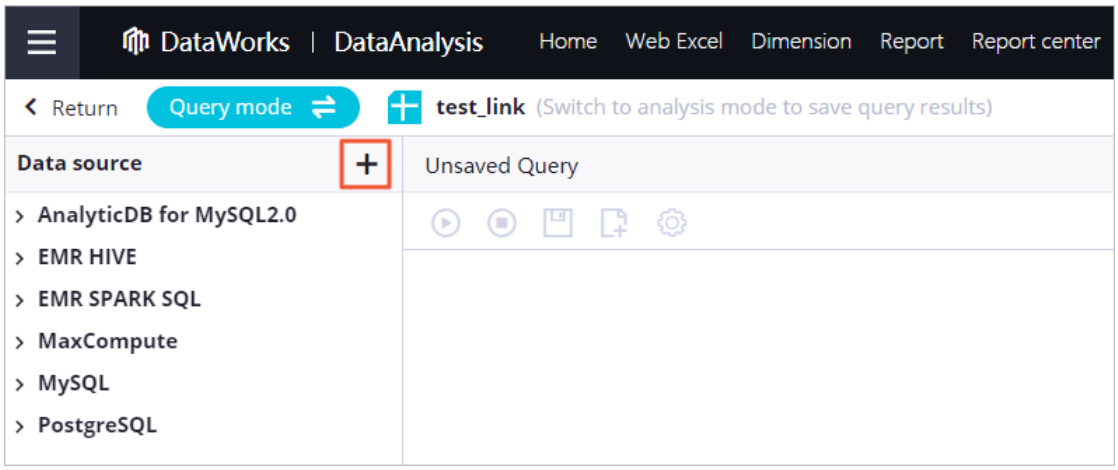
You can use the **Query mode** feature to import data from a data store to a workbook for data analysis.

1. In the upper-left corner of the workbook editing page, click **Analysis mode** to switch to **Query mode**.



2. Add a data store.If the data store from which you want to import data is already in the data store list, you can click the data store type and double-click the data store.

i. In the **Data source** section, click the + icon.



ii. In the **Select Datasource** dialog box, select a data store type and set the **Data Source** or **Engine Instance** parameter as prompted.

iii. Click **OK**.

3. Import data from the data store to the workbook.

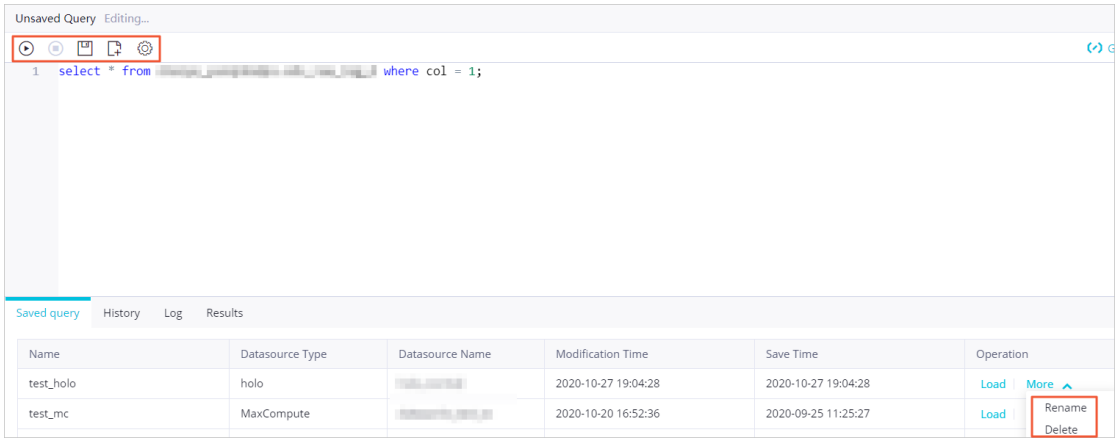
i. Compile the code for querying data from the data store.You can compile the code for querying data from the data store by using one of the following methods:

- Enter a query statement in the code editor.

> **Note** The query statement must follow the syntax that is specified by the data store type.

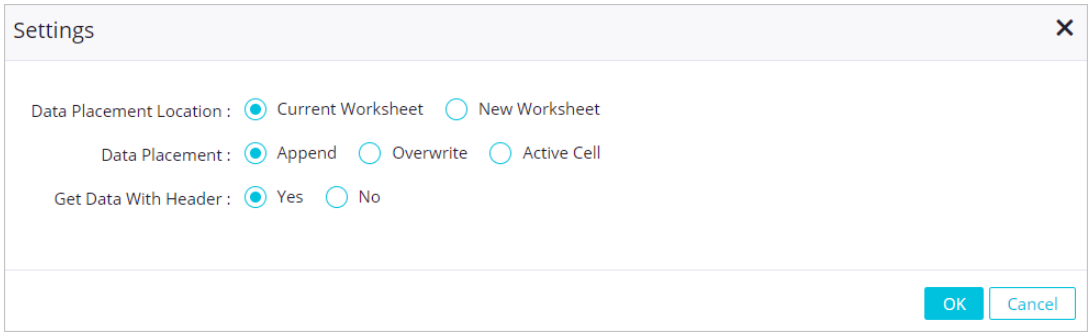
- Double-click the name of the data store to view all tables in the data store. Double-click the name of a table or field to generate a query statement.

ii. Click the ▶ icon in the top toolbar. After the query is complete, you can view the imported data in the workbook.

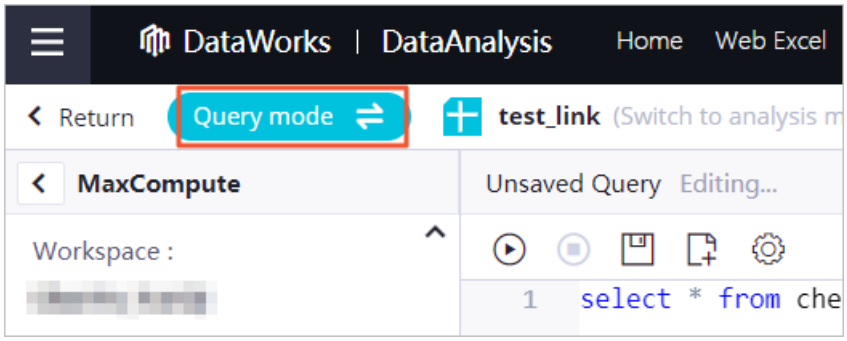


iii. Save and manage query statements.In the lower part of the right-side pane, you can view the query statements that you have saved and run on the Saved query and History tabs. You can load, rename, or delete query statements as needed. You can also perform the following operations in the code editor:

- Click the icon in the top toolbar. In the **Save** dialog box, enter a name for the query statement in the **File Name** field and click **OK**.
- Click the icon in the top toolbar to clear the query statement in the code editor. After the query statement is cleared, you can enter a new query statement.
- Click the icon in the top toolbar. In the **Settings** dialog box, set the **Data Placement Location**, **Data Placement**, and **Get Data With Header** parameters and click **OK**.



4. After you query data from the data store, you can click **Query mode** to switch to **Analysis mode** in the upper-left corner of the workbook editing page. This way, you can perform data pivoting and profiling. For more information, see Edit a workbook and analyze data.



Import local data

On the workbook editing page, move the pointer over **Import** in the upper-right corner and select **Local File** or **File Data** to import local data.

- If you select **Local File**, you can import only Excel files from an on-premises machine. Data in all sheets of a selected Excel file is imported.

  Move the pointer over **Import** and select **Local File**. Select an Excel file and click **Open** to import data in all sheets of the Excel file to the workbook.

- If you select **File Data**, you can import data from workbooks or import CSV files or Excel files from an on-premises machine. If you import data from a workbook or an Excel file, you can specify the sheet from which the data is to be imported.

  Move the pointer over **Import** and select **File Data**. In the **Import** dialog box, select one of the following types of source data based on your needs:

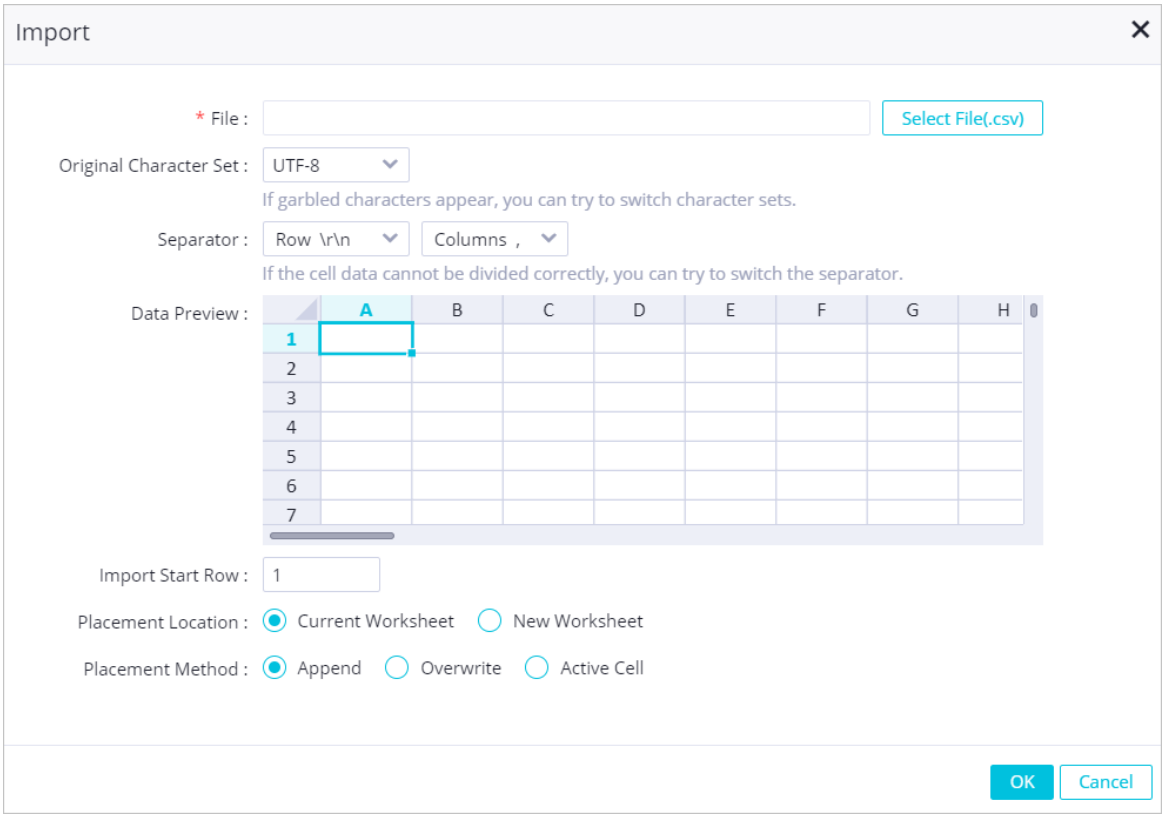
- **Spreadsheet**

  In the **Import** dialog box, click **Spreadsheet**, set the parameters, and then click **OK**.

Parameter	Description
Spreadsheet	The workbook from which the data is to be imported. Select a workbook from the **Spreadsheet** drop-down list.
Sheet	The sheet from which the data is to be imported. Select a sheet from the **Sheet** drop-down list.
Data Preview	A section that displays the data in the selected sheet.
Import Start Row	The row from which the data is to be imported. Default value: 1.
Placement Location	The location where the imported data is placed. Valid values: **Current Worksheet** and **New Worksheet**.
Placement Method	The way in which the imported data is placed. Valid values: **Append**, **Overwrite**, and **Active Cell**.

- **Local CSV File**

  In the **Import** dialog box, click **Local CSV File**, set the parameters, and then click **OK**.

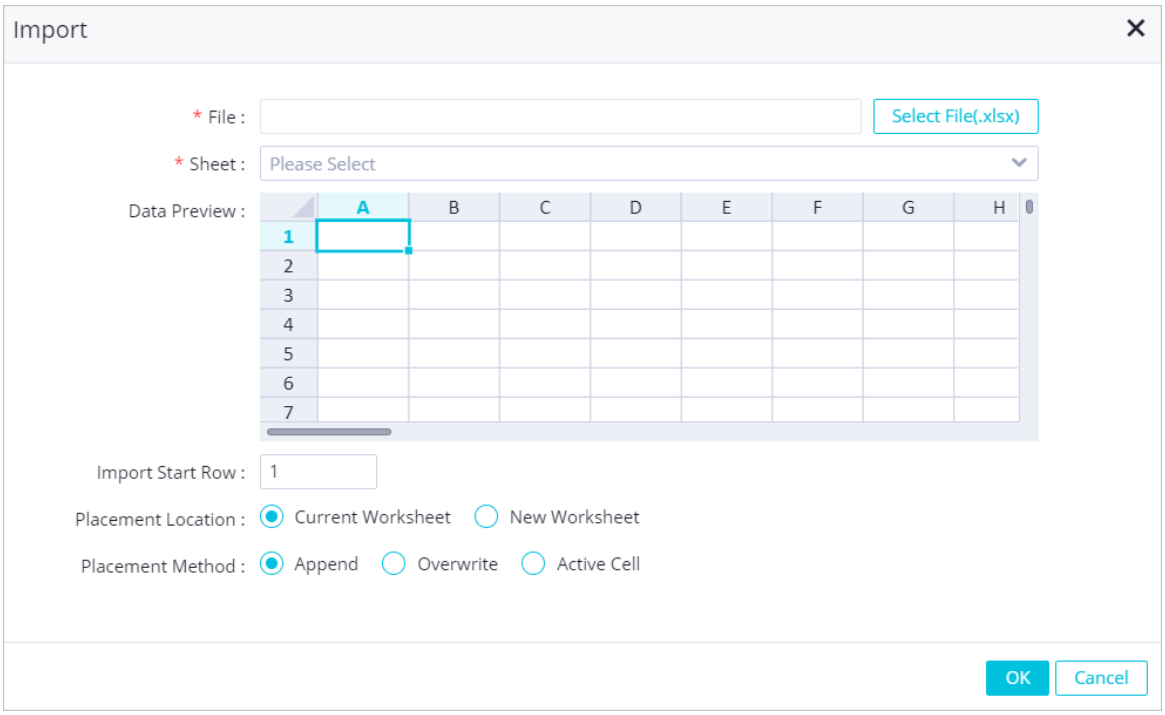


Parameter	Description

Parameter	Description
File	The CSV file from which the data is to be imported. Click **Select File(.csv)**, select a CSV file from the on-premises machine, and then click **Open**.
Original Character Set	The character set that is used by the selected CSV file. Valid values: **UTF-8** and **GBK**. If garbled characters appear, you can change the character set.
Separator	The row delimiter and column delimiter. ■ Valid values of row delimiters: \r\n, \n, and \r. ■ Valid values of column delimiters: ,, ;, and \t. If the cell data cannot be correctly divided, you can change the delimiters.
Data Preview	A section that displays the data in the selected CSV file.
Import Start Row	The row from which the data is to be imported. Default value: 1.
Placement Location	The location where the imported data is placed. Valid values: **Current Worksheet** and **New Worksheet**.
Placement Method	The way in which the imported data is placed. Valid values: **Append**, **Overwrite**, and **Active Cell**.

- **Local Excel File**

  In the **Import** dialog box, click **Local Excel File**, set the parameters, and then click **OK**.



Parameter	Description
File	The Excel file from which the data is to be imported. Click **Select File(.xlsx)**, select an Excel file from the on-premises machine, and then click **Open**.
Sheet	The sheet from which the data is to be imported. Select a sheet from the **Sheet** drop-down list.
Data Preview	A section that displays the data in the selected sheet.
Import Start Row	The row from which the data is to be imported. Default value: 1.
Placement Location	The location where the imported data is placed. Valid values: **Current Worksheet** and **New Worksheet**.
Placement Method	The way in which the imported data is placed. Valid values: **Append**, **Overwrite**, and **Active Cell**.

2.3. Analyze data

DataWorks workbooks allow you to perform multiple data analysis operations with ease. These operations are highly consistent with those in Excel. This greatly reduces learning costs.

Prerequisites

A workbook is created, and data is imported to the workbook. For more information, see Create a workbook and Import data to a workbook.

Context

On the workbook editing page, you can specify the font, text alignment, number format, rows and columns, conditional formatting, and style of a workbook. You can also perform data pivoting and profiling on the workbook. For more information, see Data pivoting and Data profiling.

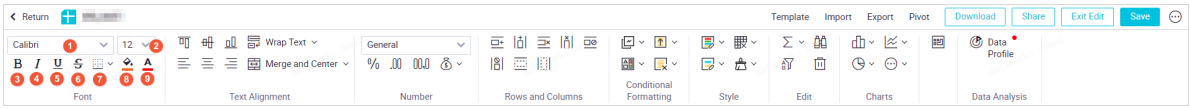
Go to the workbook editing page

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the **All Spreadsheets** section of the **Web Excel** page, click the name of the required workbook to go to the workbook editing page.If you create a workbook, the workbook editing page appears after the workbook is created. For more information, see Create a workbook.

Edit the workbook

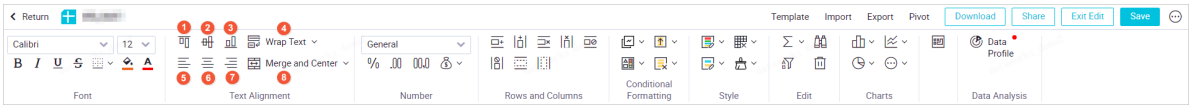
On the workbook editing page, you can specify the following settings:

- **Font**



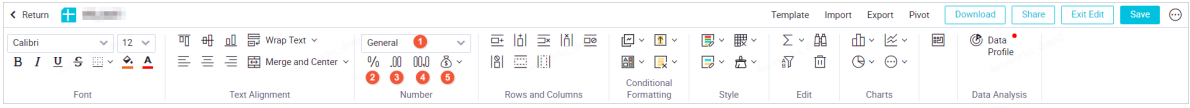
No.	Feature	Description
1	**Font**	Specifies a font.
2	**Font Size**	Specifies a font size.
3	**Bold**	Sets the text in bold.
4	**Italic**	Sets the text in italic.
5	**Underline**	Underlines the text.
6	**Strikethrough**	Adds a strikethrough to the text.
7	**Borders**	Adds borders to the text.
8	**Fill Color**	Specifies the background color of the text.
9	**Font Color**	Specifies the color of the text.

- **Text Alignment**



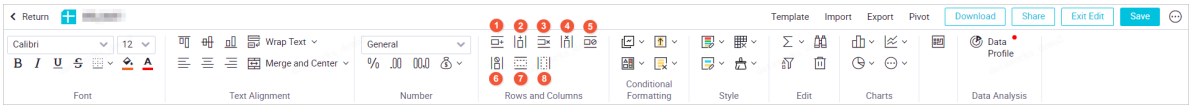
No.	Feature	Description
1	**Top Align**	Aligns the text to the top.
2	**Middle Align**	Aligns the text vertically to the center in the cell.
3	**Bottom Align**	Aligns the text to the bottom.
4	**Wrap Text**	Displays long text in multiple lines to make it easy to view all the text.
5	**Align Left**	Aligns the text to the left.
6	**Center**	Aligns the text horizontally to the center.
7	**Align Right**	Aligns the text to the right.
8	**Merge and Center**	Merges multiple cells to one cell and centers the content in the cell.

- **Number**



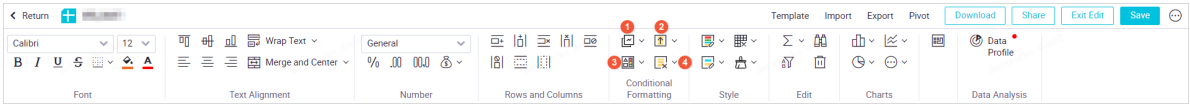
No.	Feature	Description
1	**Number Format**	Specifies the number format for selected cells. You can select General, Number, Currency, Short Date, Long Date, Time, Percentage, Fraction, Scientific, or Text.
2	**Percentage**	Applies the percentage format to numbers.
3	**Two Decimal Places**	Rounds numbers to two decimal places.
4	**1000 Separator**	Displays numbers with thousands separators, for example, **1,005**.
5	**Currency**	Adds a currency sign to numbers. The following currency signs are supported: yuan sign (¥), dollar sign ($), pound sign (£), euro sign (€), and franc sign (Fr).

- **Rows and Columns**



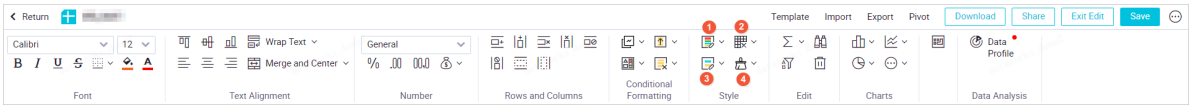
No.	Feature	Description
1	**Insert Row**	Inserts one or more rows to the workbook.
2	**Insert Column**	Inserts one or more columns to the workbook.
3	**Delete Row**	Deletes one or more selected rows from the workbook.
4	**Delete Column**	Deletes one or more selected columns from the workbook.
5	**Lock Row**	Locks the rows before the selected row in the workbook.
6	**Lock Column**	Locks the columns before the selected column in the workbook.
7	**Hide Row**	Hides one or more selected rows in the workbook.
8	**Hide Column**	Hides one or more selected columns in the workbook.

- **Conditional Formatting**



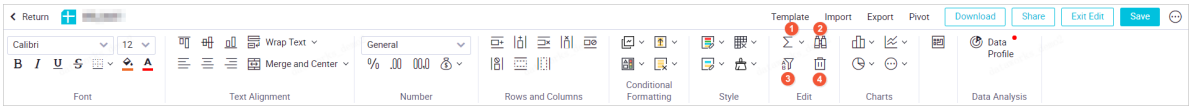
No.	Feature	Description
1	**Highlight cell rules**	Includes rules in the **Highlight Cells Rules** and **Top/Bottom Rules** categories.
2	**Data Bar/Color Scale**	Includes styles in the **Gradient Fill**, **Solid Fill**, and **Color Scales** categories.
3	**Icon Set**	Includes icons in the **Directional**, **Shapes**, **Indicators**, and **Ratings** categories.
4	**Clear Rule**	Includes **Clear Rules from Selected Cells** and **Clear Rules from Entire Sheet**.

- **Style**



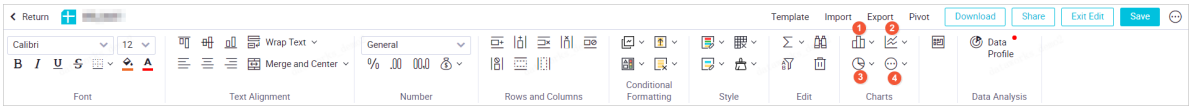
No.	Feature	Description
1	**Apply table style**	Applies a table style.
2	**Delete**	Removes the applied table style.
3	**Cell Style**	Applies a cell style.
4	**Clear**	Includes **Clear All**, **Clear Content**, and **Clear Style**.

- **Edit**



No.	Feature	Description
1	**AutoSum**	Includes **Sum**, **Average**, **Count Numbers**, **Max**, and **Min**.
2	**Search**	Displays the search box after you click **Search** or press **Ctrl+F**.
3	**Sort and Filter**	Allows you to filter data and sort data in ascending or descending order.
4	**Clear**	Clears the selected content.

- **Charts**



No.	Feature	Description
1	**Column Chart**	For more information, see Column charts.
2	**Line Chart**	For more information, see Line charts.
3	**Pie Chart**	For more information, see Pie charts.

No.	Feature	Description
4	**More**	Allows you to select one of the following chart types after you click the **More** icon: ○ Area charts ○ Horizontal bar charts ○ Scatter charts ○ Stock charts

- **Plug-in**

  The following plug-ins are supported: **Type Conversion** and **Intelligent Chart Recommendation**.

  - **Type Conversion**: To convert the selected data to numbers or strings, click **Type Conversion** and select **Convert to Numeric** or **Convert to String**.
  - **Intelligent Chart Recommendation**: To remove or retain the title of a chart, click **Intelligent Chart Recommendation** and select **Remove the Title** or **Retain the Title**.

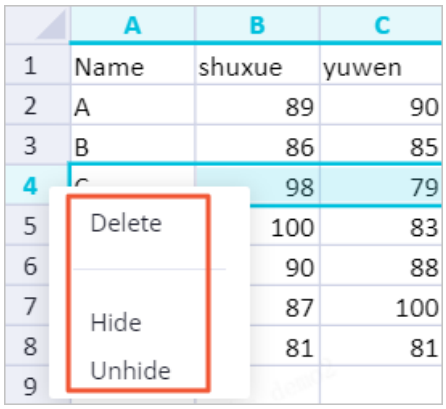
    > **Note** The **Intelligent Chart Recommendation** plug-in is supported in the following regions: China (Shanghai), China (Beijing), China (Hangzhou), China (Shenzhen), China (Chengdu), China (Hong Kong), China (Zhangjiakou), China North 2 Ali Gov, Singapore, Indonesia (Jakarta), and Japan (Tokyo).

- **List of Shortcut Keys**

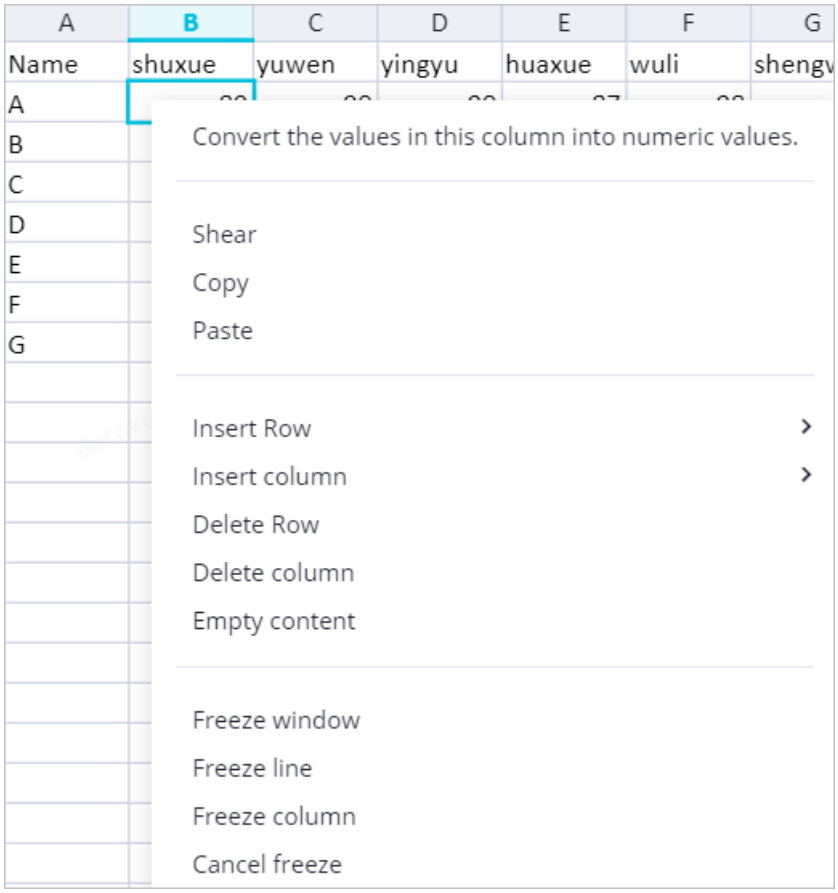
  Click the icon to view the shortcut keys for different features.

In a cell on the workbook editing page, you can enter content or enter a formula that references values in other cells. The column headings are labeled with letters, which start from A and increase alphabetically from left to right. The row headings are labeled with numbers, which start from 1 and increase from top to bottom.

To delete, hide, or show a specific row or column, right-click the row or column heading and select **Delete**, **Hide**, or **Unhide**.



To manage specific cells, select and right-click the cells, and then select an option, such as **Shear**, **Copy**, and **Paste**.

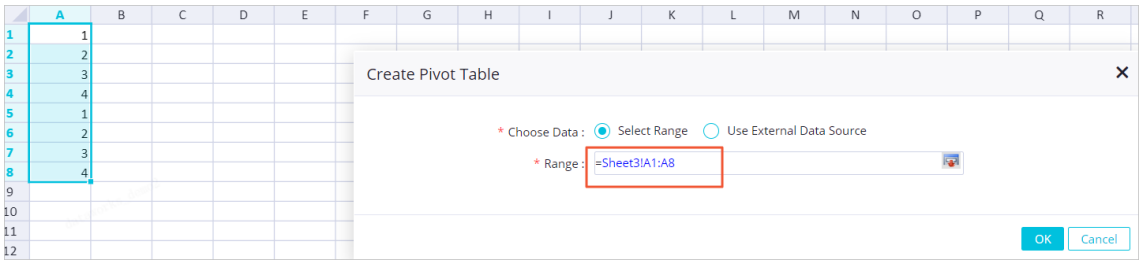


Data pivoting

Note The **Pivot** feature is supported in the following regions: China (Shanghai), China (Beijing), China (Hangzhou), China (Shenzhen), China (Chengdu), China (Hong Kong), China (Zhangjiakou), and China North 2 Ali Gov.

1. On the workbook editing page, select the data for which you want to create a pivot table and click **Pivot** in the upper-right corner.
2. In the **Create Pivot Table** dialog box, specify the data to be analyzed.You can select **Select Range** or **Use External Data Source**.
   - **Select Range**

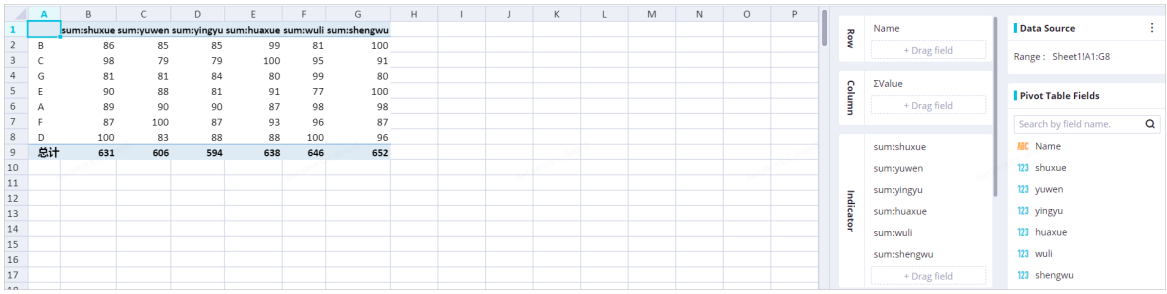
     Select the cells for which you want to create a pivot table. The value of the **Range** parameter changes based on the selected cells.



- **Use External Data Source**

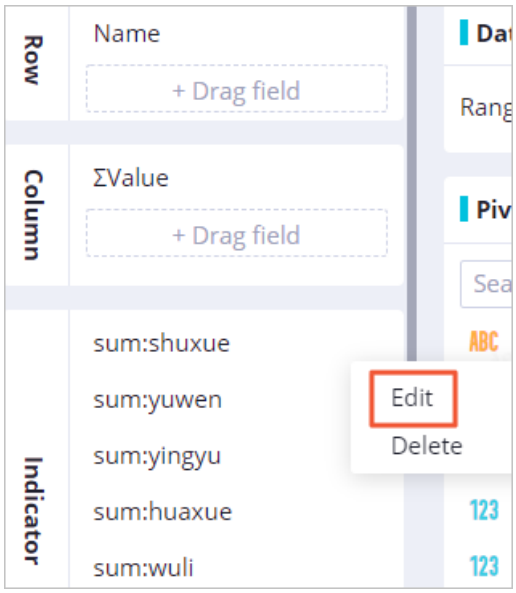
  Select a data store of the **MaxCompute**, **Mysql**, **Data Services**, or **PostgreSQL** type. If you use an external data store, prepare the required connection or API in advance and select the connection or API based on the business requirements. For more information about how to configure a connection, see Connection configuration.

3. Click **OK**. The pivot table editing page appears.

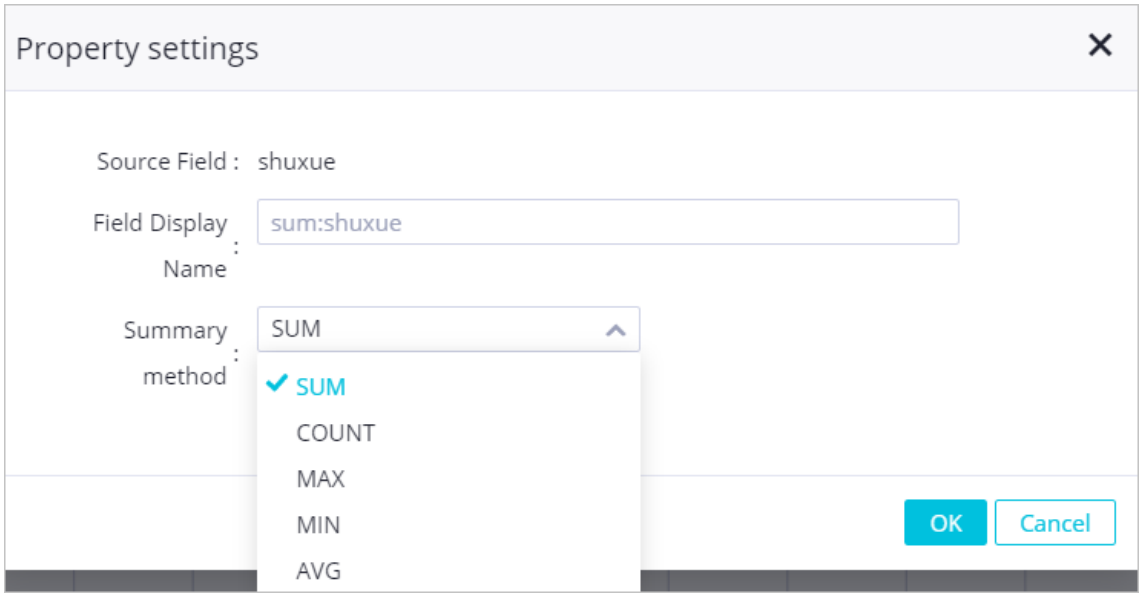


In this example, select **Select Range**.

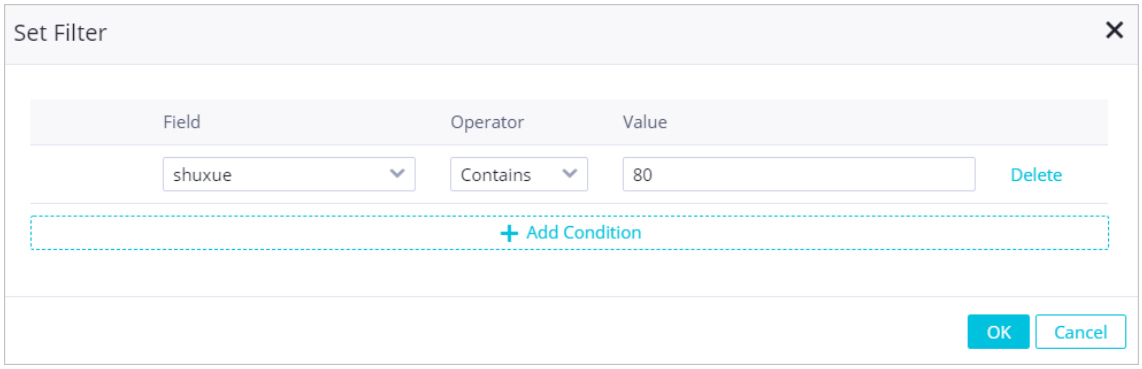
- **Data Source**: the range of the selected data in the workbook.
- **Pivot Table Fields**: the names of the fields that you selected.
- **Row**: You can drag a field to the **Row** section. Each value of the field added to the **Row** section occupies a row in the pivot table.
- **Column**: You can drag a field to the **Column** section. Each value of the field added to the **Column** section occupies a column in the pivot table.
- **Indicator**: To modify the settings of an indicator, move the pointer over the indicator, click the ⋮ icon, and then select **Edit**.



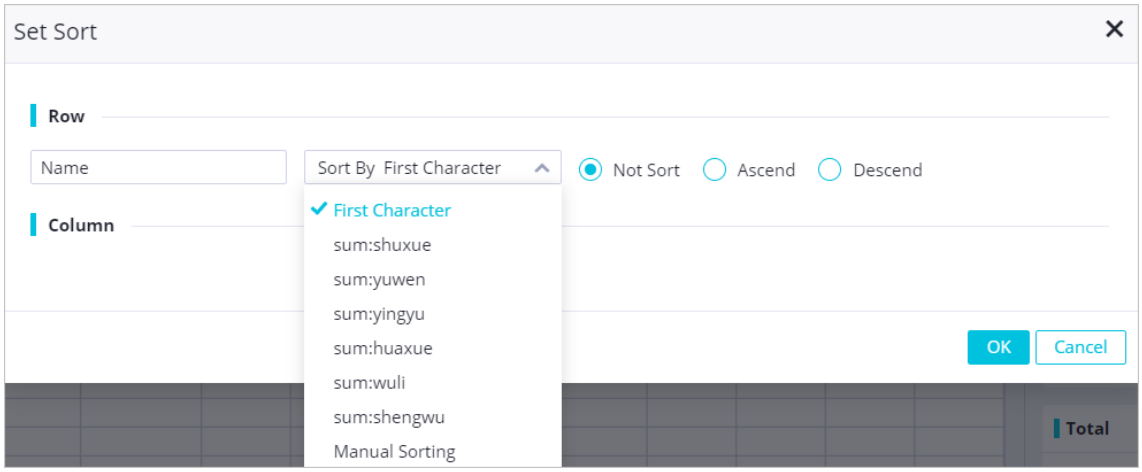
In the **Property settings** dialog box, set the **Summary method** parameter and click **OK**. By default, the name indicated by the **Field Display Name** parameter is in the format of **Summary method:Source field name** and cannot be changed.



- **Filters**: To filter data, click **Set Filter**. In the **Set Filter** dialog box, click **Add Condition**, specify the filter conditions, and then click **OK**.



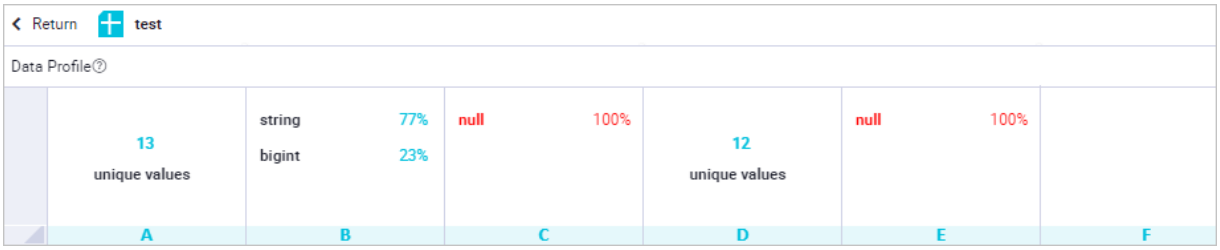
- **Sort**: To sort data based on a field specified in the **Row** section, click **Set Sort**. In the **Set Sort** dialog box, specify the sorting rule and click **OK**.



Data profiling

The Data Profile feature allows you to analyze the quality, structure, distribution, and statistics of data. It also allows you to preview, profile, process, analyze, and visualize data. The Data Profile feature analyzes data based on columns and allows you to view the data types and value distribution of each column.

Select the data to be analyzed and click **Data Profile** in the menu bar. Then, you can view the data types and value distribution of each column at the top of the workbook in the form of charts and rich text.



The simple mode of data profiling has the following features:

- For the STRING or DATE data type: displays the values ranking top 2 and their respective percentages, and the percentage of other values in the form of rich text. If the number of unique values exceeds 50% of the total number of values, the number of unique values is displayed.
- For the INTEGER or FLOAT data type: displays the value distribution in the form of a histogram.
- For the BOOLEAN data type: displays the proportions of different values in the form of pie charts.
- For mixed data types that involve two or more data types: displays the proportions of different data types in the form of pie charts. The system reminds you that the current column has dirty data. After the dirty data is cleared, the simple mode displays value distribution in one of the preceding forms based on the data type.
- For null values: displays the percentage of null values in red.

Click **Detailed Mode** in the upper-right corner. In the **Data Profile** dialog box, you can view the profiling result, including the field name, field data type, field description, and security level of each column.

The detailed mode of data profiling has the following features:

- For the STRING or DATE data type: displays basic information and the values ranking top 5 based on frequency. The basic information includes the percentage of null values and the numbers of fields, unique values, and valid values.
- For the INTEGER or FLOAT data type: displays basic information, the values ranking top 5 based on frequency, statistics, and a histogram. The basic information includes the percentage of null values and the numbers of fields, unique values, and zeros.
- For the BOOLEAN data type: displays basic information, the values ranking top 5 based on frequency, and a pie chart. The basic information includes the percentage of null values and the numbers of fields, unique values, and zeros.

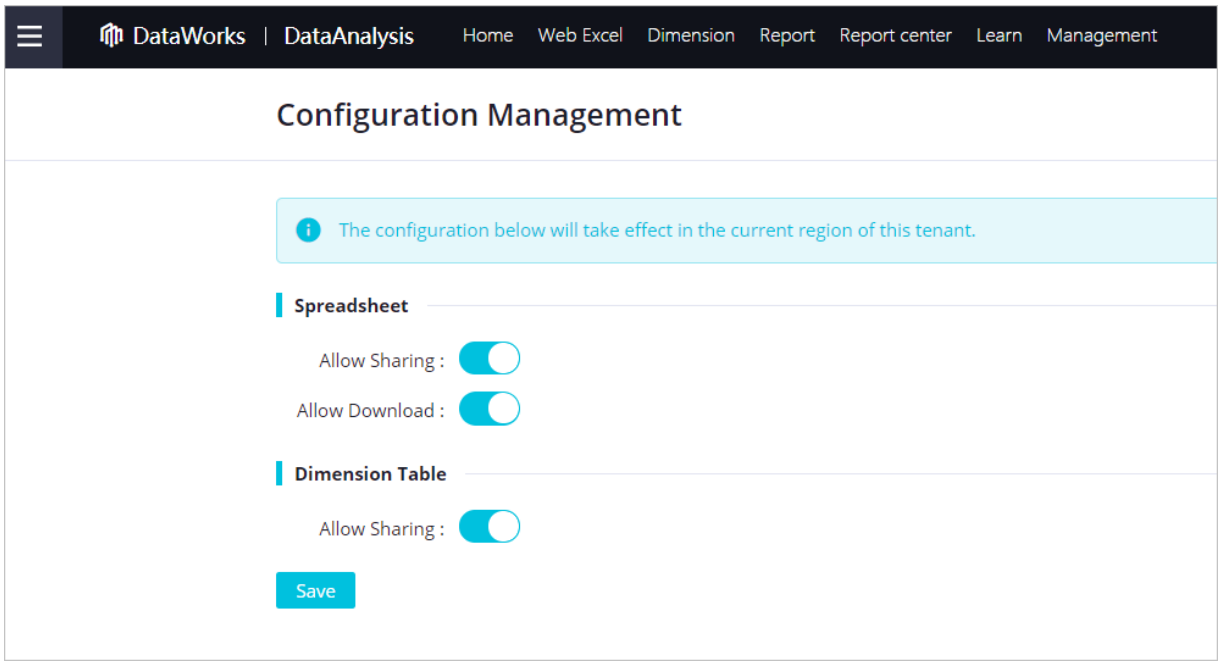
> **Note** The system considers the true and false strings and the 0 and 1 integers as values of the BOOLEAN type.

2.4. Export, share, and download a workbook

This topic describes how to share and download a workbook after data analysis is completed.

Prerequisites

Sharing and download are allowed. You can perform the following steps: On the Configuration Management page, turn on **Allow Sharing** and **Allow Download** in the Spreadsheet section.



Note You must use your Alibaba Cloud account to access the **Configuration Management** page and configure relevant settings.

Context

Workbook sharing is applicable to the following scenarios:

- Collaboratively edit a workbook with multiple users.

  For example, a workbook is used to collect personal information of team members and information about whether they enroll for an event. You can share the workbook with the team members and grant them permissions to edit the workbook.

- Share analysis results with other users.

  You can share a workbook with other users and grant them permissions to view the workbook.

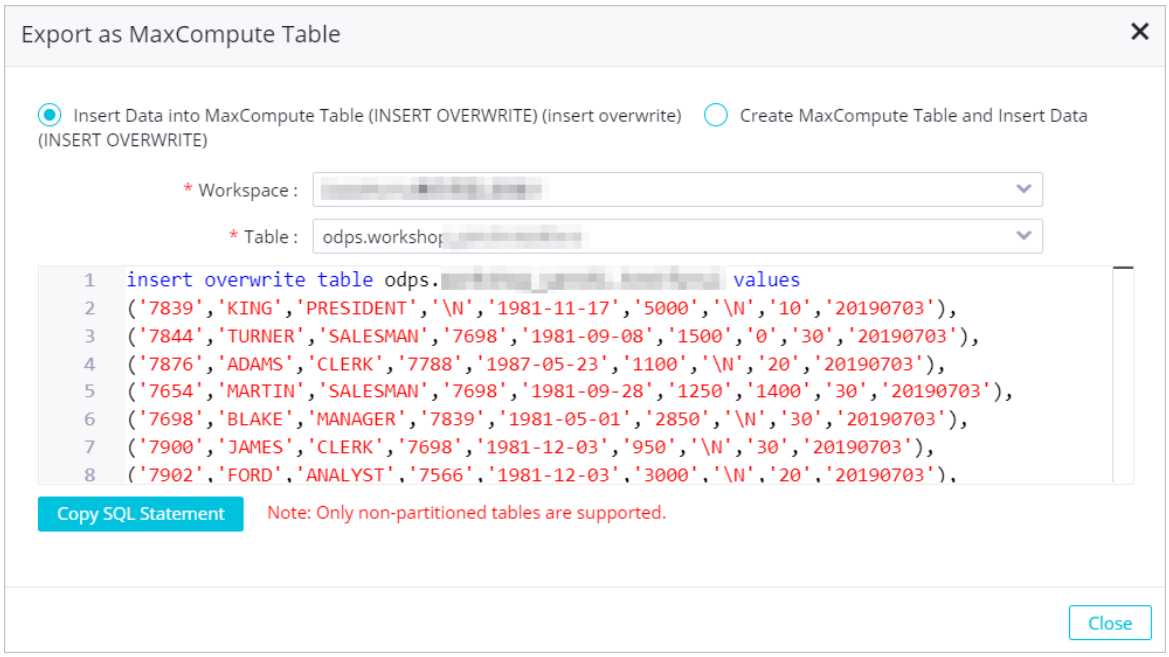
Go to the workbook editing page

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.

3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the **All Spreadsheets** section of the **Web Excel** page, click the name of the workbook that you want to edit in the **File Name** column to go to the workbook editing page. If you create a workbook in this step, the workbook editing page appears after the workbook is created. For more information, see Analyze data.

Export a workbook

1. In the upper-right corner of the workbook editing page, choose **Export > Generate MaxCompute Build Table Statement**.
2. In the **Export as MaxCompute Table** dialog box, set relevant parameters.



Insert mode	Parameter	Description
Insert Data into MaxCompute Table (INSERT OVERWRITE) (insert overwrite)	**Workspace**	The workspace to which the MaxCompute table belongs.
	Table	The MaxCompute table to which you want to insert data.
Create MaxCompute Table and Insert Data (INSERT OVERWRITE)	**Workspace**	The workspace to which the MaxCompute table belongs.
	Table Name	The name of the MaxCompute table. Make sure that the table name is unique. You can click **Check Duplicate Names** to check whether the table name exists.

3. After the parameters are set, click **Copy SQL Statement** and then click **Close**.

> **Notice** Only non-partitioned tables are supported.

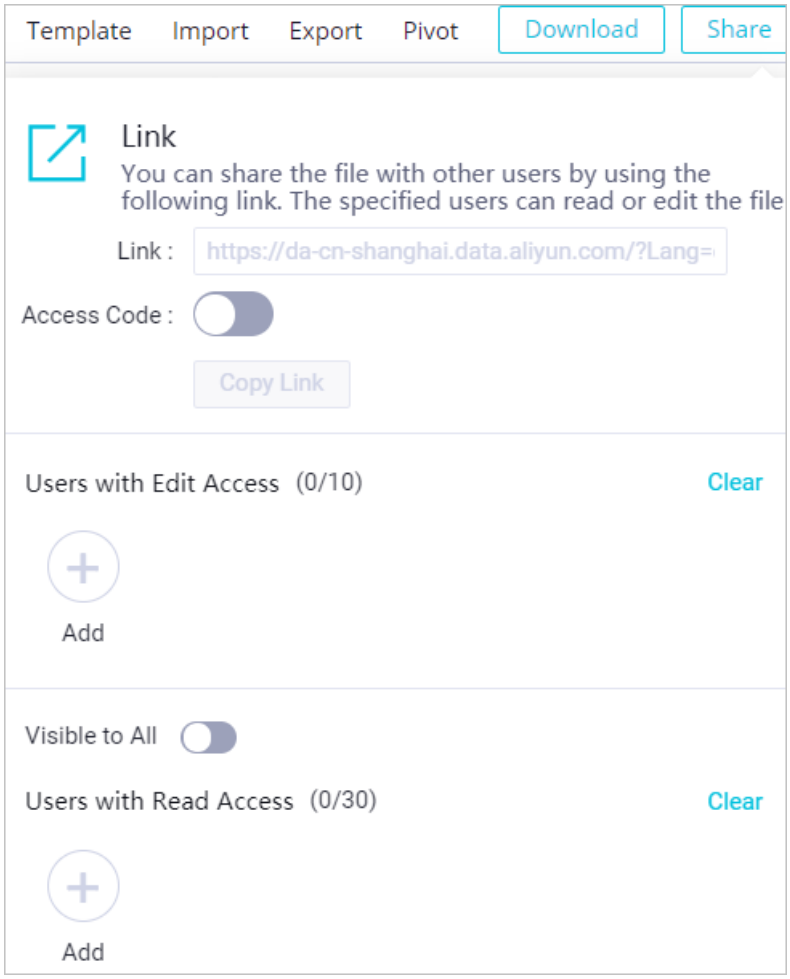
4. On the workbook editing page, click the icon in the upper-left corner and choose All Products > **Data Development > DataStudio**.
5. Insert data to an existing MaxCompute table or create a MaxCompute table and insert data to the created table.
   - **Insert Data into MaxCompute Table (INSERT OVERWRITE) (insert overwrite)**

     If you select **Insert Data into MaxCompute Table (INSERT OVERWRITE) (insert overwrite)** as the insert mode, go to the editing page of the MaxCompute table to which you want to insert data. On the editing page of the MaxCompute table, click **DDL mode**. In the DDL mode dialog box, paste the copied SQL statement and click **Generate table structure**.
   - **Create MaxCompute Table and Insert Data (INSERT OVERWRITE)**

     If you select **Create MaxCompute Table and Insert Data (INSERT OVERWRITE)** as the insert mode, create a MaxCompute table and execute the copied SQL statement.
6. Click **Submit to development environment** and **Submit to production environment** in sequence.If you are using a workspace in basic mode, you only need to click **Submit to production environment**.

Share a workbook

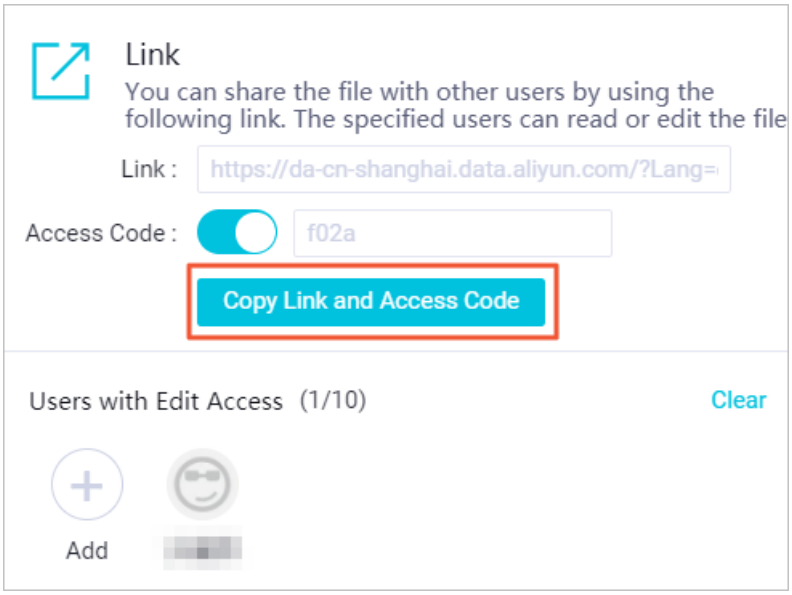
In the upper-right corner of the workbook editing page, click **Share**. In the dialog box that appears, set the sharing method as needed.



You must configure the following information before you can share a workbook with other users:

- **Link**: After you specify Users with Edit Access and Users with Read Access or turn on Visible to All, click **Copy Link** and send the copied URL to other users as needed.

  If you turn on **Access Code**, a URL with an access code is generated.



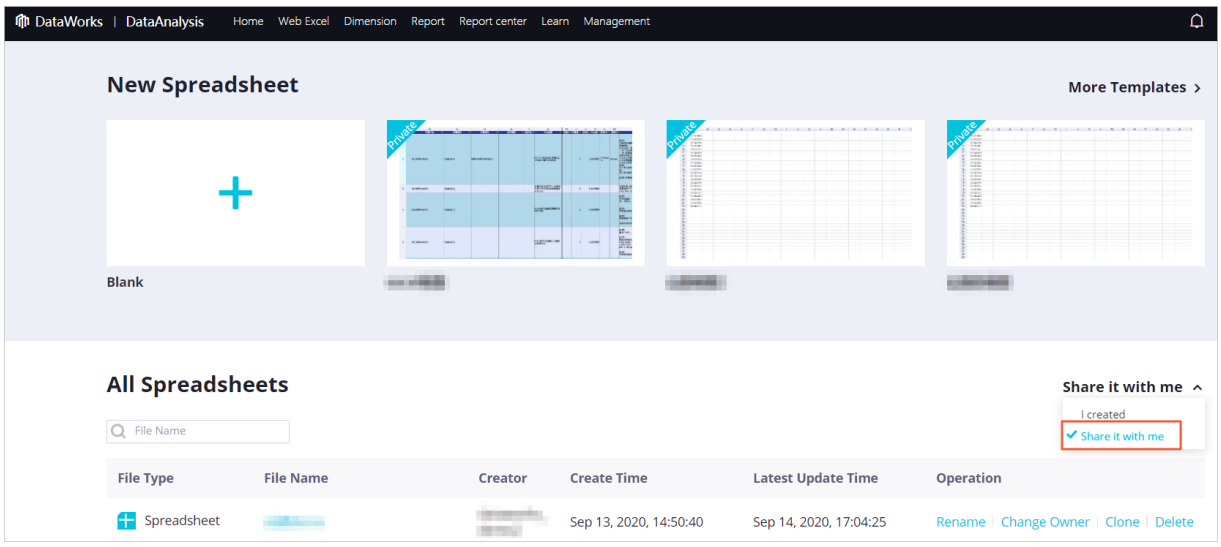
- **Users with Edit Access**: To specify users with permissions to edit the workbook, click **Add** in the **Users with Edit Access** section. In the Share File with These Users dialog box, enter and select the names of the users to be granted the edit permissions, and click **OK**.

  > **Note** You can grant the edit permissions to up to 10 users.

- **Visible to All**: To allow all users to view the workbook, turn on **Visible to All**.
- **Users with Read Access**: To specify users with permissions to view the workbook, turn off **Visible to All** and click **Add** in the **Users with Read Access** section. In the Share File with These Users dialog box, enter and select the names of the users to be granted the read permissions, and click **OK**.

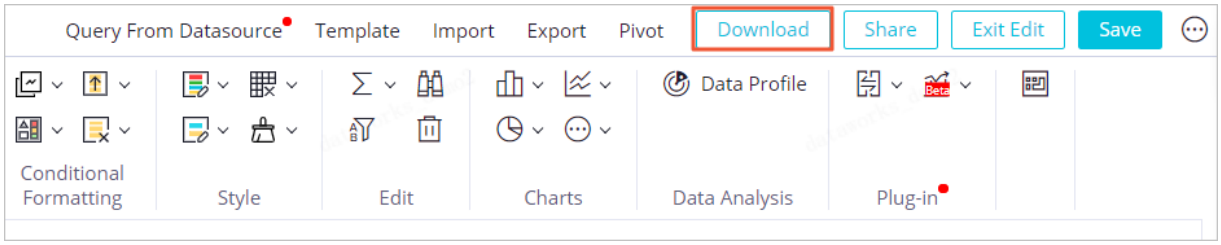
  > **Note** You can grant the read permissions to up to 30 users.

After the sharing method is set, you can send the URL to other users. The users can access the workbook through the URL. On the **Web Excel** page, you can also view the workbooks that are shared with you.



Download a workbook

In the upper-right corner of the workbook editing page, click **Download** to download the workbook to a local directory.



2.5. Manage a template

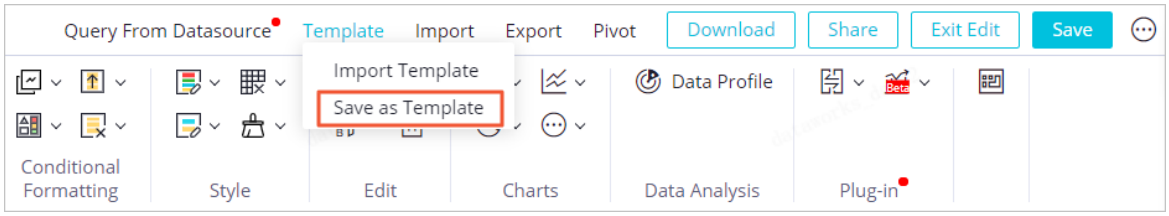
You can save the current workbook as a template or apply a saved template.

Go to the workbook editing page

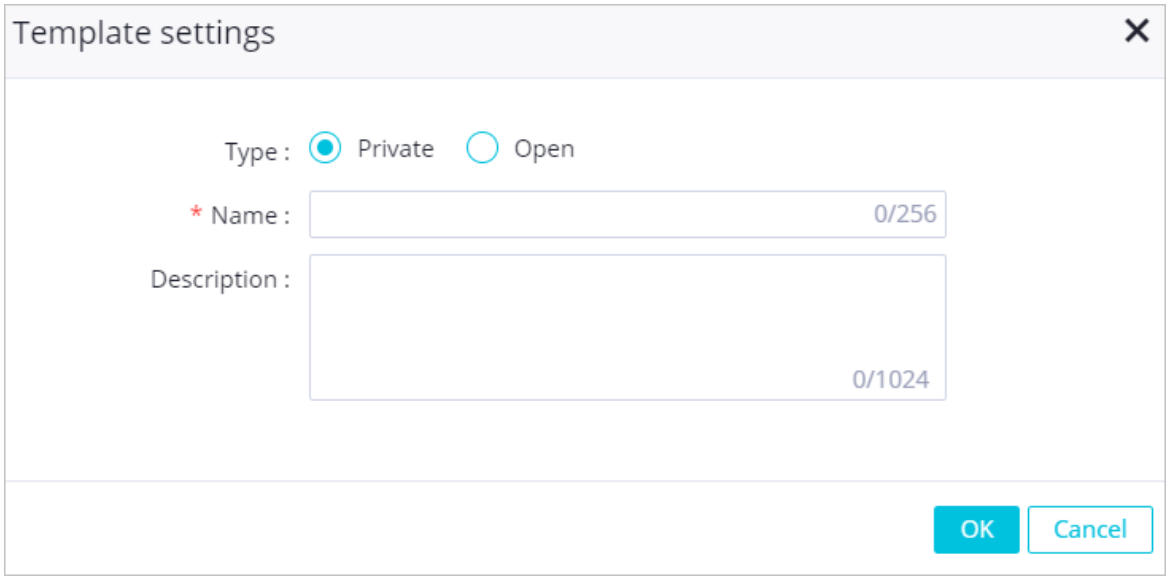
1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the **All Spreadsheets** section of the **Web Excel** page, click the name of the workbook that you want to edit in the **File Name** column to go to the workbook editing page. If you create a workbook in this step, the workbook editing page appears after the workbook is created. For more information, see Analyze data.

Save the workbook as a template

1. In the upper-right corner of the workbook editing page, choose **Template** > **Save as Template**.



2. In the **Template settings** dialog box, set the parameters.

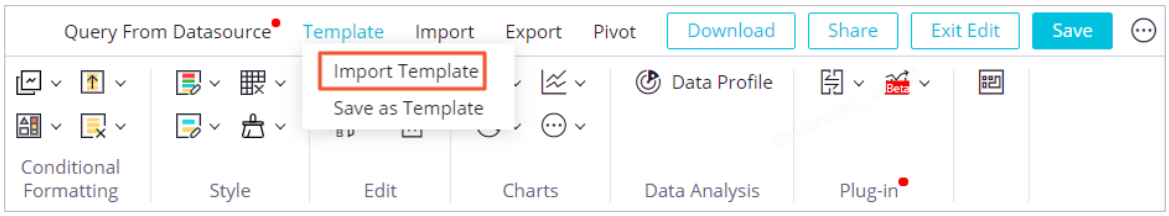


Parameter	Description
Type	Specifies whether to show or hide the template for other users. Valid values: **Private** and **Open**.
Name	The name of the template. The name can be up to 256 characters in length.
Description	The description of the template. The description can be up to 1,024 characters in length.

3. Click **OK**.

Apply a template

1. In the upper-right corner of the workbook editing page, choose **Template** > **Import Template**.



2. In the **Import Template** dialog box, select a template for the current workbook.

> **Note** The data of the selected template will overwrite that of the current workbook.

3. Click **OK**.

3.MaxCompute table

3.1. Create and manage a MaxCompute table

The dimension table feature allows you to create MaxCompute tables, import local data to MaxCompute tables, and edit MaxCompute tables in a visualized manner.

Prerequisites

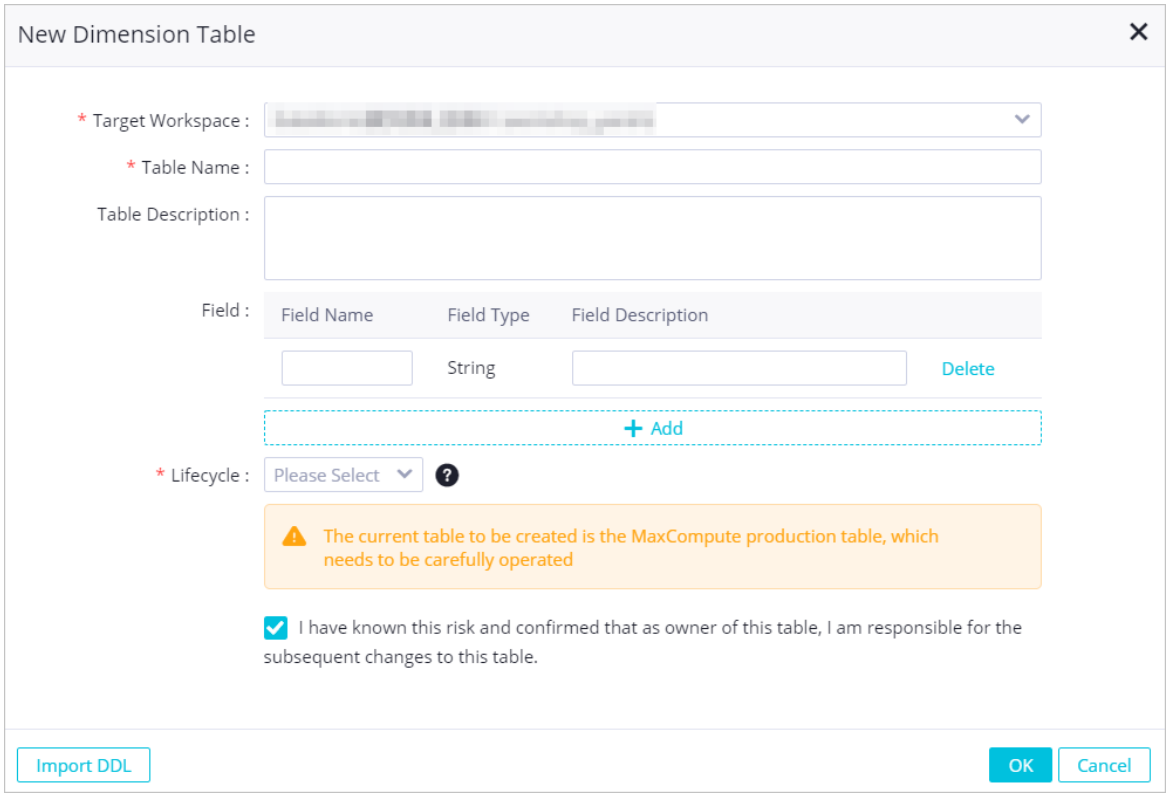
1. MaxCompute is activated. For more information, see Activate MaxCompute.
2. A MaxCompute compute engine is bound to a DataWorks workspace. For more information, see Configure a workspace.
3. A MaxCompute table is created. For more information, see Create a MaxCompute table.

Limits

- To create a MaxCompute table in DataAnalysis, you must be an administrator, an owner, or a developer of a DataWorks workspace. For more information, see Add workspace members.
- For a MaxCompute table that is created by using the dimension table feature, all fields in the MaxCompute table are of the STRING type. If you need to use fields of other data types, execute Data Definition Language (DDL) statements to create a MaxCompute table on the **DataStudio** page. For more information, see the "Create a table" section in the Table-level operations topic.

Create a MaxCompute table

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the top navigation bar, click **Dimension**.
5. On the **Dimension** page, click the icon under **New Dimension Table**.
6. In the **New Dimension Table** dialog box, set the parameters as required.

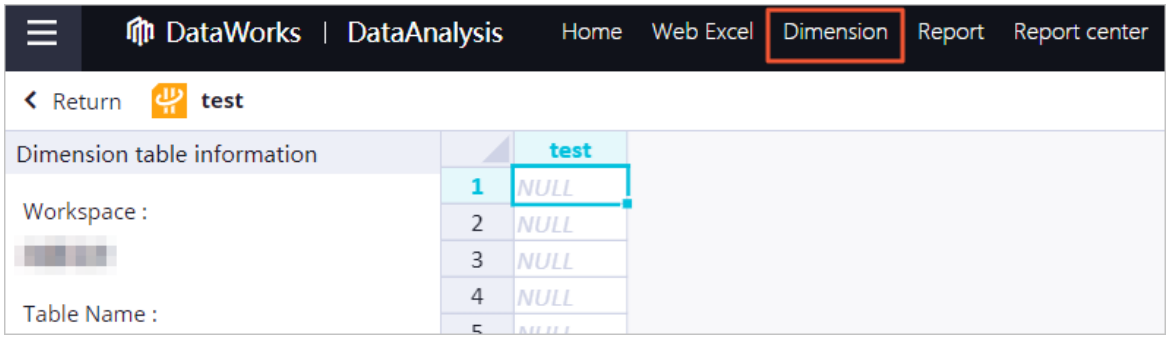


Parameter	Description
Target Workspace	The DataWorks workspace to which the MaxCompute table belongs.
Table Name	The name of the MaxCompute table. The MaxCompute table will be used in the production environment. **Note** The table name can contain only letters, digits, and underscores (_), and must start with a letter.
Table Description	The description of the MaxCompute table, such as the purpose or features.
Field	The fields in the MaxCompute table. Only fields of the STRING type can be added.
Lifecycle	The lifecycle of the MaxCompute table. The MaxCompute table occupies storage resources in MaxCompute. To make sure that the resources can be recycled, select a proper lifecycle for the MaxCompute table from the drop-down list. If the specified lifecycle expires, the table is deleted.

7. Select **I have known this risk and confirmed that as owner of this table, I am responsible for the subsequent changes to this table.** Click **OK** to go to the MaxCompute table editing page to view and modify information about the table. For more information, see Edit a MaxCompute table.The MaxCompute table created in DataAnalysis is maintained in the production environment. As the owner of the table, you are responsible for the maintenance of the table.

View and manage a MaxCompute table

1. Click **Return** in the upper-left corner of the MaxCompute table editing page or **Dimension** in the top navigation bar to go to the Dimension page.



2. In the **All Dimension Tables** section of the **Dimension** page, select **I created** or **Share it with me** from the drop-down list in the upper-right corner to view the MaxCompute tables in the corresponding category.You can also share MaxCompute tables with specific users. For more information, see Share a MaxCompute table.
3. Click the file name of a MaxCompute table, or click **Edit** in the Operation column of the MaxCompute table to go to the MaxCompute table editing page.On the Dimension page, you can perform the following operations to manage a MaxCompute table:
   - To change the owner of a MaxCompute table, perform the following steps: Find the MaxCompute table and click **Change Owner** in the Operation column. In the **Change Owner** dialog box, select an owner from the New Owner drop-down list and click **OK**.
   - To delete a MaxCompute table, perform the following steps: Find the MaxCompute table and click **Delete** in the Operation column. In the **Delete** message, click **OK**.

What's next

After you create a MaxCompute table, go to the MaxCompute table editing page and import data to the MaxCompute table. For more information, see Import data to a MaxCompute table.

3.2. Import data to a MaxCompute table

After you create a MaxCompute table, you can write data to the table for data analysis. You can also import data from a workbook, local CSV file, or local Excel file to the table.

Prerequisites

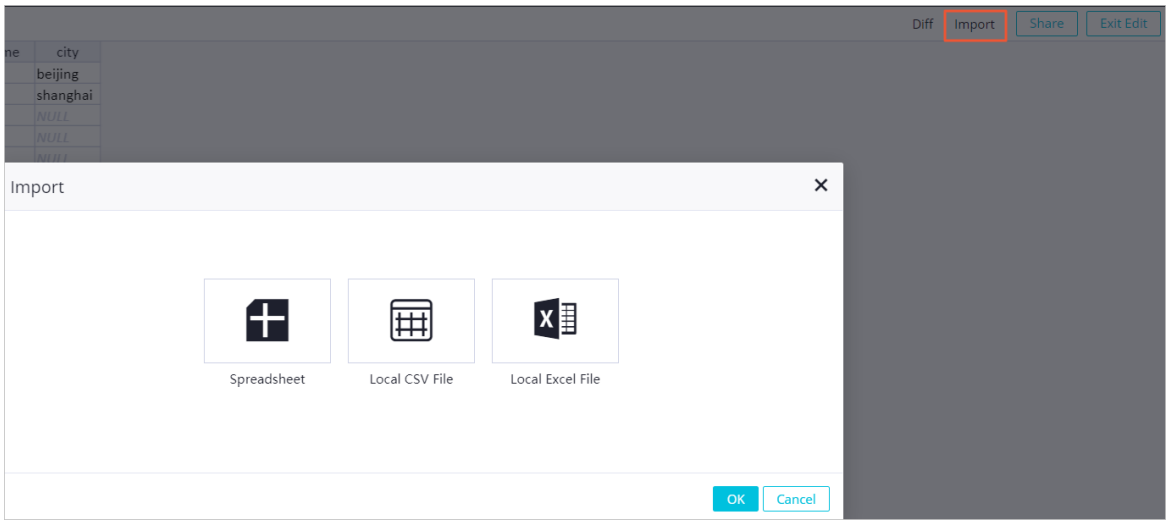
A MaxCompute table is created. For more information, see Create and manage a MaxCompute table.

Procedure

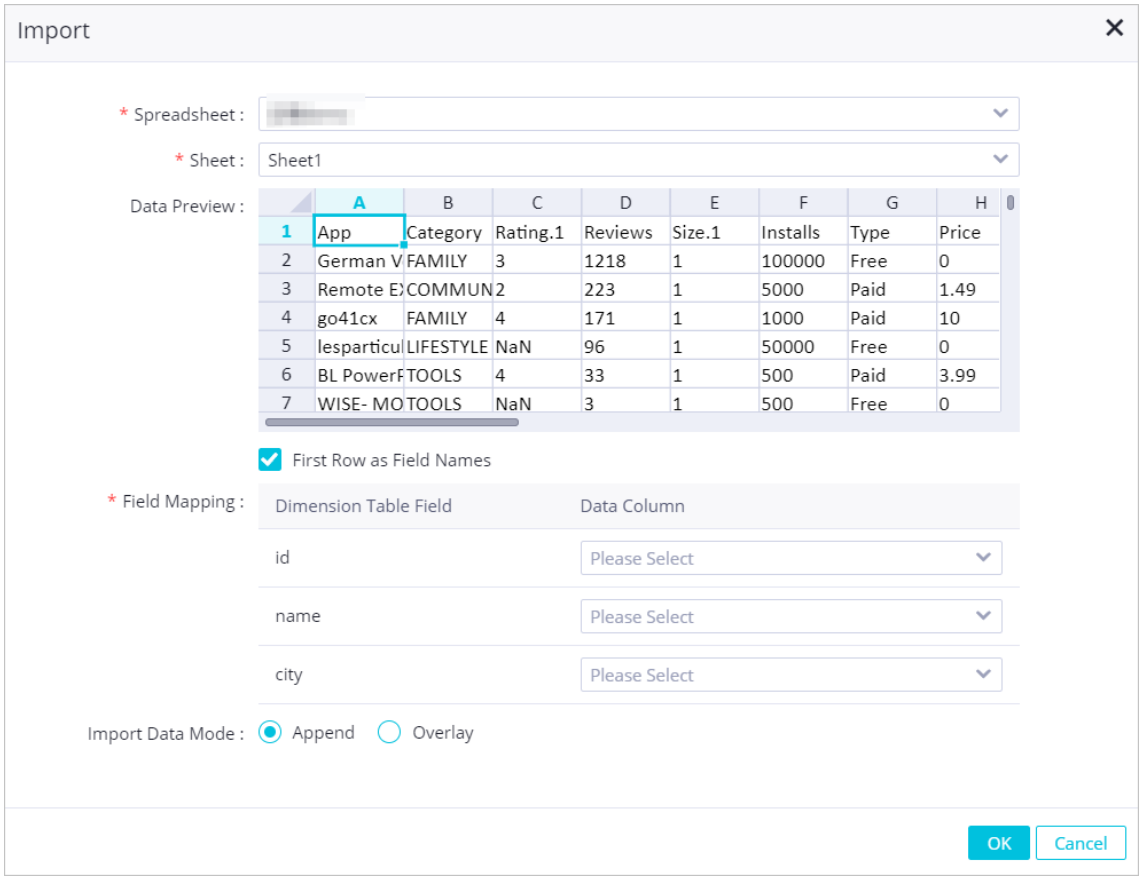
1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.

3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the top navigation bar, click **Dimension**.
5. In the **All Dimension Tables** section of the **Dimension** page, click the name of the MaxCompute table that you want to edit in the **File Name** column to go to the MaxCompute table editing page.If you create a MaxCompute table in this step, the MaxCompute table editing page appears after the MaxCompute table is created. For more information, see Edit a MaxCompute table.
6. On the MaxCompute table editing page, click **Import** in the upper-right corner.
7. In the **Import** dialog box, select the type of the file to be imported and set the parameters.

   > **Note** Only data of the STRING type can be imported to a MaxCompute table. Data that is not of the STRING type will be converted to the STRING type automatically when it is imported.

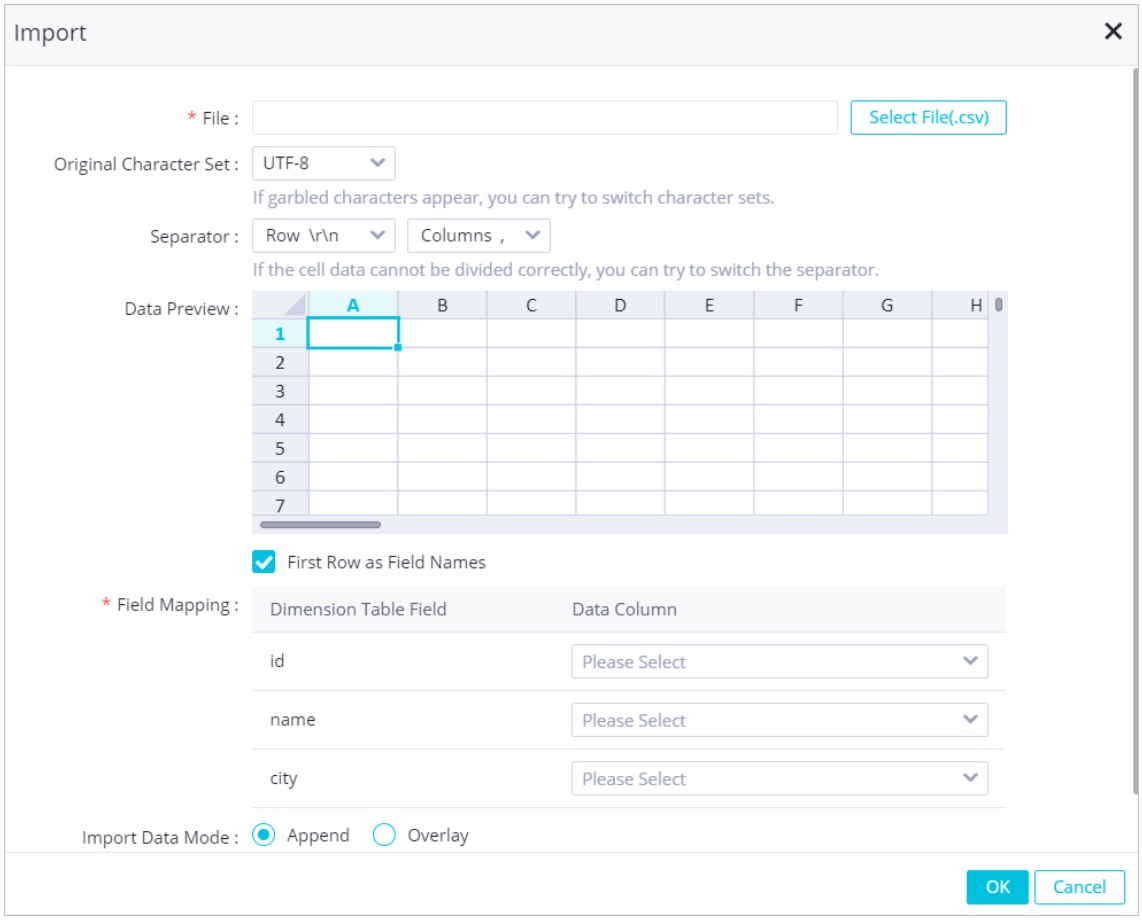
   

   - **Workbook**



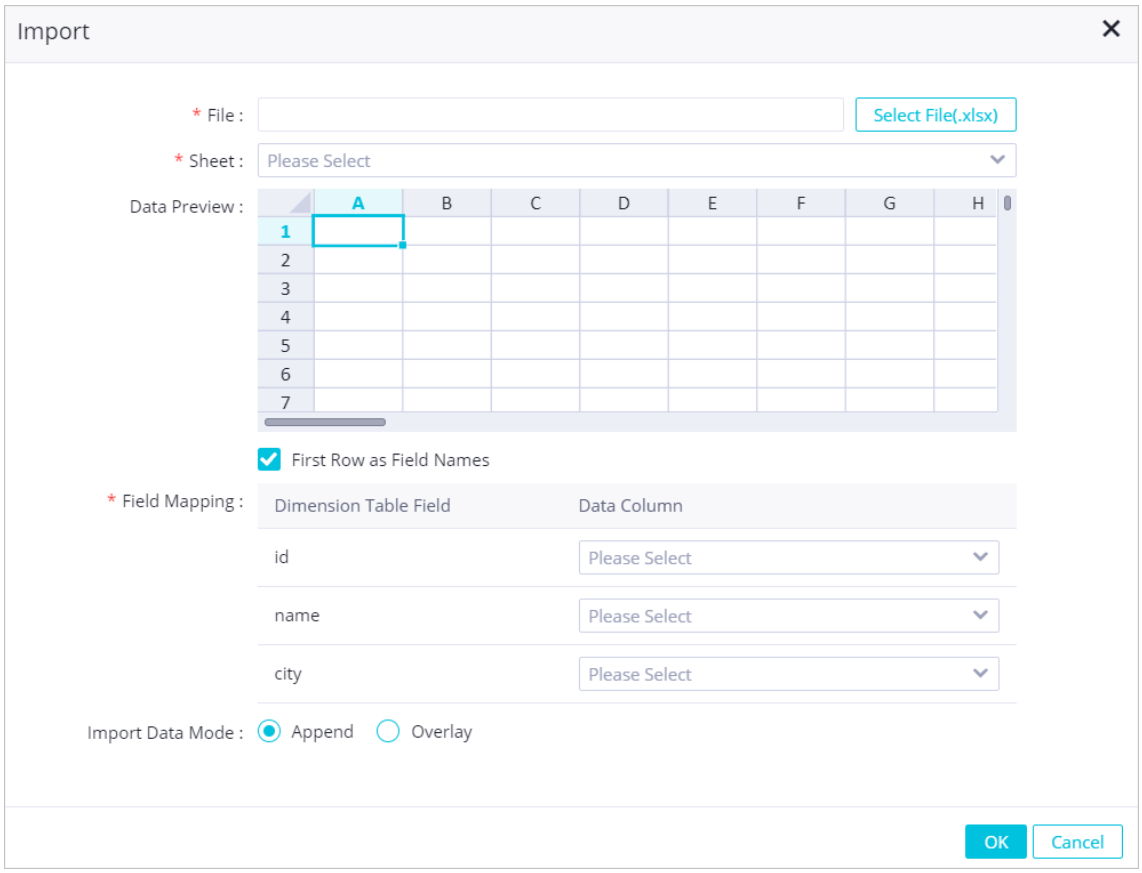
Parameter	Description
Spreadsheet	The workbook from which the data is to be imported. Select a workbook from the **Spreadsheet** drop-down list.
Sheet	The sheet from which the data is to be imported. Select a sheet from the **Sheet** drop-down list.
Data Preview	A section that displays the data in the selected sheet. When you preview the data in the selected sheet, you can determine whether to use the values in the first row as the column names by selecting or clearing **First Row as Field Names**.
Field Mapping	The mappings between the columns in the selected sheet and the fields in the MaxCompute table.
Import Data Mode	The mode used to import data. Valid values: **Append** and **Overlay**.

- **Local CSV file**



Parameter	Description
File	The local CSV file from which the data is to be imported. Click **Select File(.csv)**, select a local CSV file, and then click **Open**.
Original Character Set	The character set that is used by the selected CSV file. Valid values: **UTF-8** and **GBK**. If garbled characters appear, you can change the character set.
Separator	The row separator and column separator. ■ Valid values of row separators: \r\n, \n, and \r. ■ Valid values of column separators: ,, ;, and \t. If the cell data cannot be correctly divided, you can change the separators.
Data Preview	A section that displays the data in the selected CSV file. When you preview the data in the selected CSV file, you can determine whether to use the values in the first row as the column names by selecting or clearing **First Row as Field Names**.
Field Mapping	The mappings between the columns in the selected CSV file and the fields in the MaxCompute table.
Import Data Mode	The mode used to import data. Valid values: **Append** and **Overlay**.

- **Local Excel file**



Parameter	Description
File	The local Excel file from which the data is to be imported. Click **Select File(.xlsx)**, select a local Excel file, and then click **Open**.
Sheet	The sheet from which the data is to be imported. Select a sheet from the **Sheet** drop-down list.
Data Preview	A section that displays the data in the selected sheet. When you preview the data in the selected sheet, you can determine whether to use the values in the first row as the column names by selecting or clearing **First Row as Field Names**.
Field Mapping	The mappings between the columns in the selected sheet and the fields in the MaxCompute table.
Import Data Mode	The mode used to import data. Valid values: **Append** and **Overlay**.

8. Click **OK**.
9. Click **Save** in the upper-right corner of the page.After you save the MaxCompute table, you can use the Diff feature to check whether the changes are as expected to avoid misoperations.

3.3. Edit a MaxCompute table

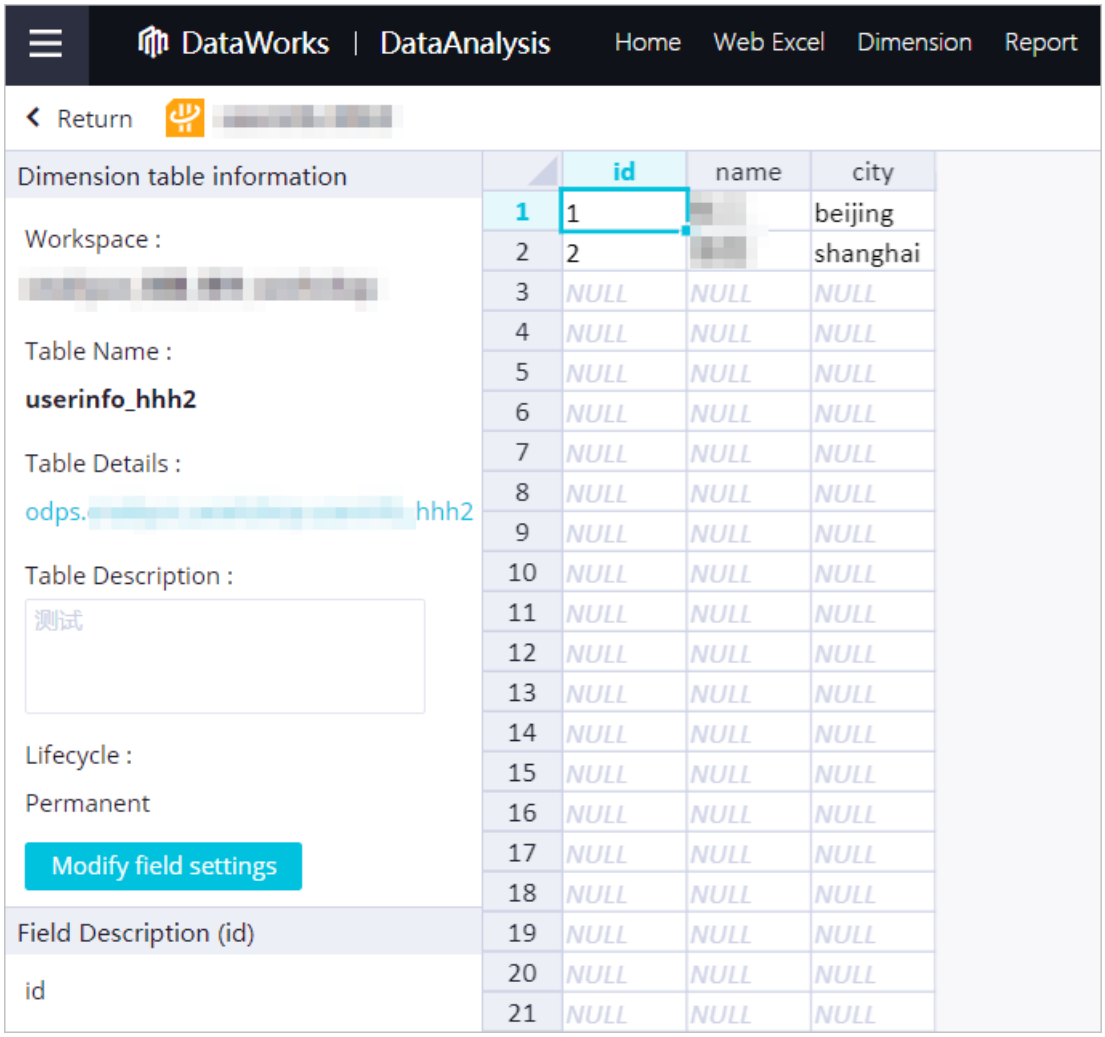
The dimension table feature allows you to edit MaxCompute tables without writing SQL code. Instead, you can edit MaxCompute tables in a visualized manner.

Prerequisites

A MaxCompute table is created. For more information, see Create and manage a MaxCompute table.

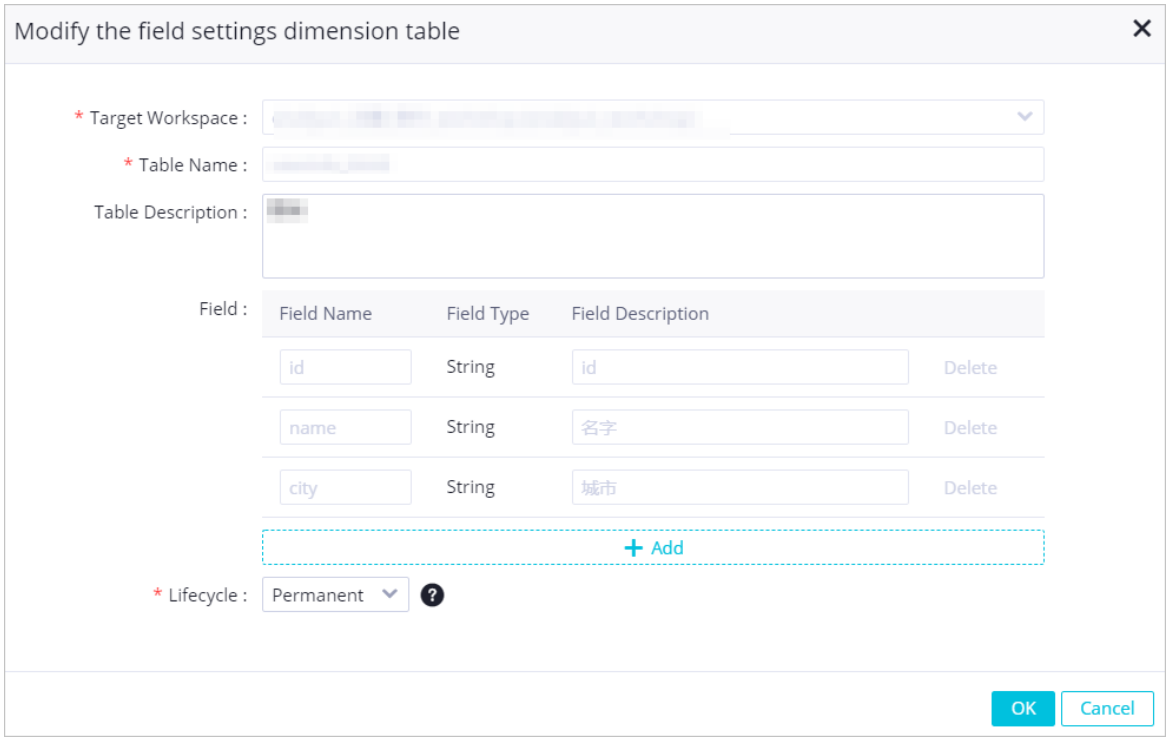
Procedure

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the top navigation bar, click **Dimension**.
5. In the **All Dimension Tables** section of the **Dimension** page, click the name of the MaxCompute table that you want to edit in the **File Name** column to go to the MaxCompute table editing page.
6. On the MaxCompute table editing page, view and modify the information about the MaxCompute table.



In the left side of the MaxCompute table editing page, you can view the MaxCompute table information, such as the workspace, table name, table description, lifecycle, and field description. To view the details of the MaxCompute table, click the link under **Table Details** to go to the **Data Map** page. For more information, see View the details of a table.

To modify the settings of the MaxCompute table, perform the following steps: Click the **Modify field settings** icon. In the **Modify the field settings dimension table** dialog box, modify **Table Description** and **Lifecycle**. You can also add fields to the MaxCompute table in this dialog box.



The right side of the MaxCompute table editing page displays all the data in the MaxCompute table as a workbook. The values in the first row are used as field names. You can double-click a cell to modify the content of a field in the corresponding row.

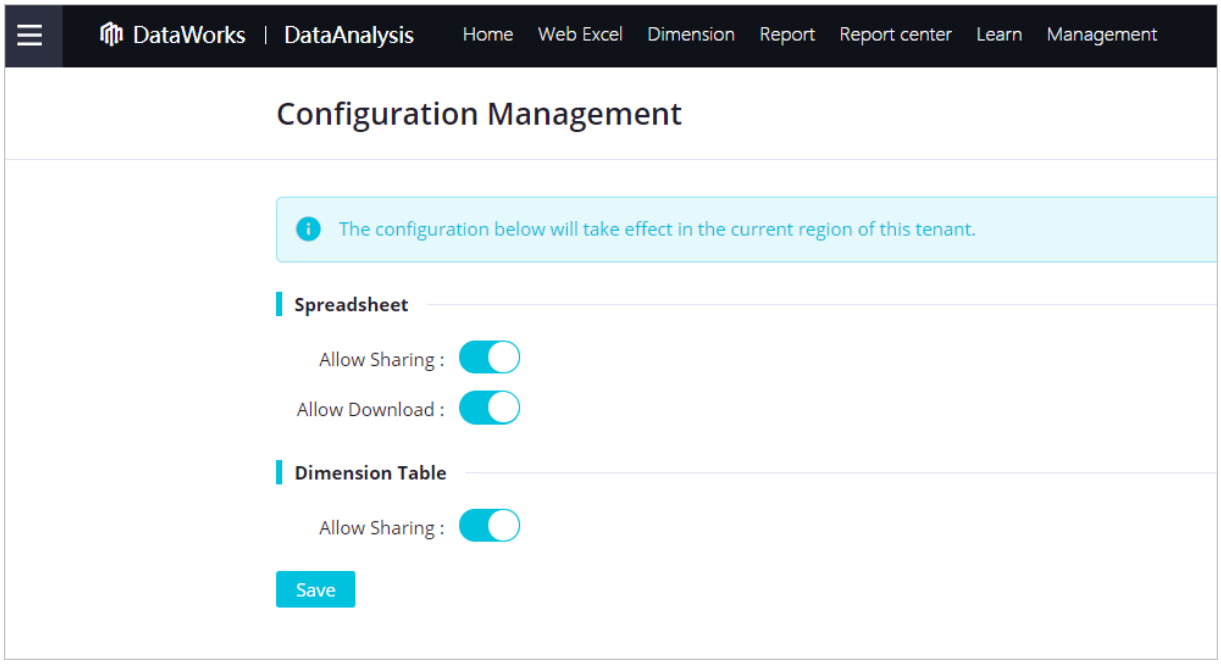
7. Click **Save** in the upper-right corner of the page to save the changes.After you save the MaxCompute table, you can view all the data in the table. You can also click **Diff** in the upper-right corner of the page to view all the data in the **Diff From the Previous Version** dialog box.

3.4. Share a MaxCompute table

If you need to collaboratively edit a MaxCompute table with multiple users, you can share the MaxCompute table and grant the users permissions to edit the MaxCompute table. This topic describes how to share a MaxCompute table and grant edit or read permissions to specified users.

Prerequisites

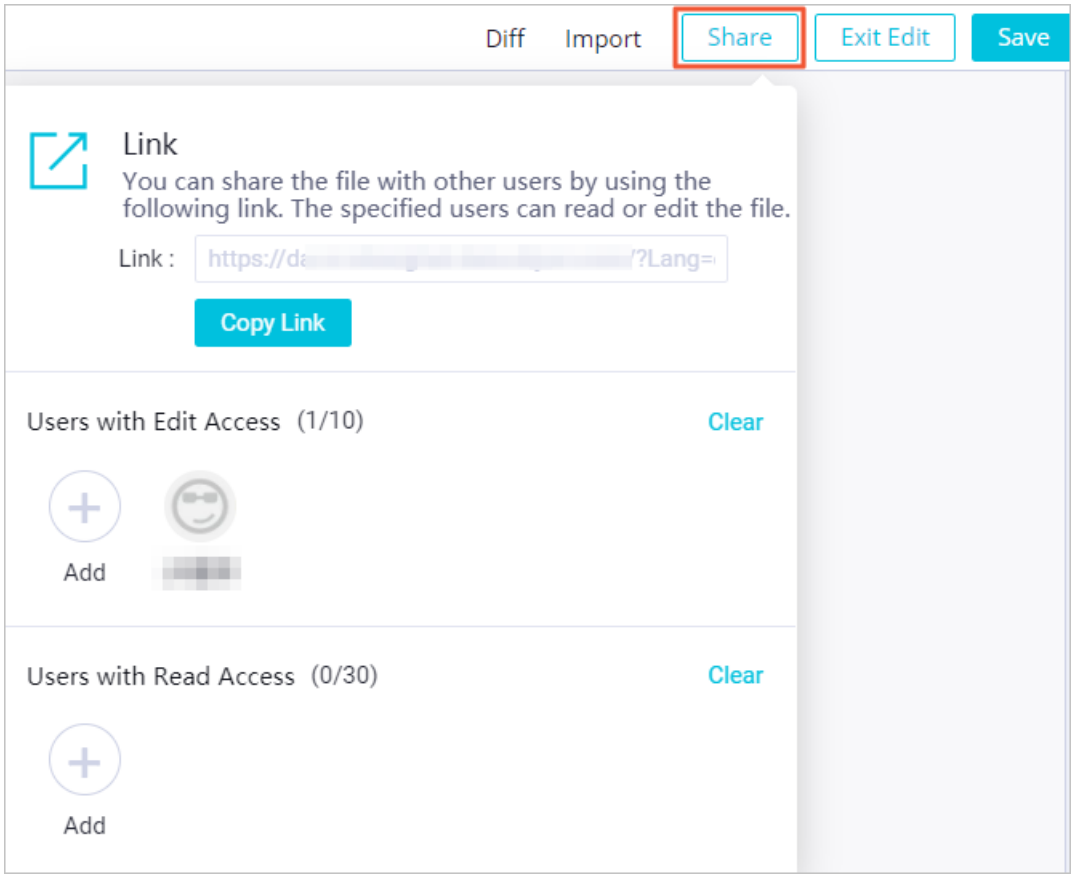
Sharing is allowed. You can perform the following steps: On the Configuration Management page, turn on **Allow Sharing** in the Dimension Table section.



Note You must use your Alibaba Cloud account to access the **Configuration Management** page and configure relevant settings.

Procedure

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the top navigation bar, click **Dimension**.
5. In the **All Dimension Tables** section of the **Dimension** page, click the name of the MaxCompute table that you want to edit in the **File Name** column to go to the MaxCompute table editing page.If you create a MaxCompute table in this step, the MaxCompute table editing page appears after the MaxCompute table is created. For more information, see Edit a MaxCompute table.
6. In the upper-right corner of the MaxCompute table editing page, click **Share**. In the dialog box that appears, set the sharing method as needed.



You must configure the following information before you can share a MaxCompute table with other users:

- **Link**: After you specify Users with Edit Access and Users with Read Access or turn on Visible to All, click **Copy Link** and send the copied URL to other users as needed.
- **Users with Edit Access**: To specify users with permissions to edit the MaxCompute table, click **Add** in the **Users with Edit Access** section. In the Share File with These Users dialog box, enter and select the names of the users to be granted the edit permissions, and click **OK**.

  > **Note** You can grant the edit permissions to up to 10 users.

- **Users with Read Access**: To specify users with permissions to view the MaxCompute table, click **Add** in the **Users with Read Access** section. In the Share File with These Users dialog box, enter and select the names of the users to be granted the read permissions, and click **OK**.

  > **Note** You can grant the read permissions to up to 30 users.

After the sharing method is set, you can send the URL to other users. The users can access the MaxCompute table through the URL. On the **Dimension** page, you can also view the MaxCompute tables that are shared with you.

4.Report

4.1. Create and manage a report

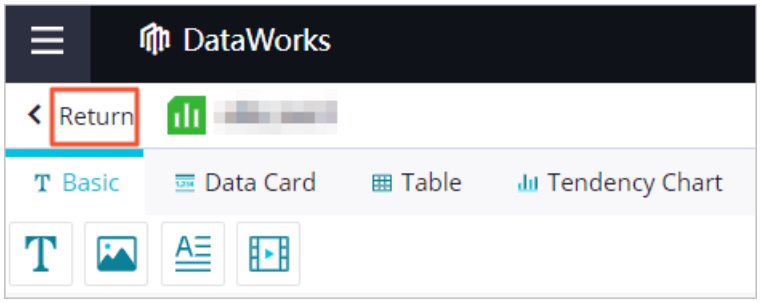
DataAnalysis allows you to explore data and create reports in a visualized manner. You can create reports by dragging and configuring controls without writing SQL code.

Create a report

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the top navigation bar, click **Report**.
5. On the **Report** page, click the icon in the **New Report** section.If you have templates under your account, you can click a template to create a report based on the template. For more information about how to create a template, see Save a report as a template.
6. In the **New Report** dialog box, set the **Report Name** and **Report Description** parameters.
7. Click **OK** to go to the report editing page. For more information, see Edit a report.

View and manage a report

1. On the report editing page, click **Return** in the upper-left corner to go to the Report page.



2. In the **All Reports** section of the **Report** page, view all reports.You can also share reports with specific users. For more information, see Share a report.
3. Click the file name of a report to go to the report editing page.On the Report page, you can perform the following operations to manage a report:
   - To rename a report, perform the following steps: Find the report and click **Rename** in the Operation column. In the **Rename** dialog box, enter the new name in the **File Name** field and click **OK**.
   - To delete a report, perform the following steps: Find the report and click **Delete** in the Operation column. In the **Delete** message, click **OK**.

What's next

After you create a report, go to the report editing page and edit the report. For more information, see Edit a report.

4.2. Edit a report

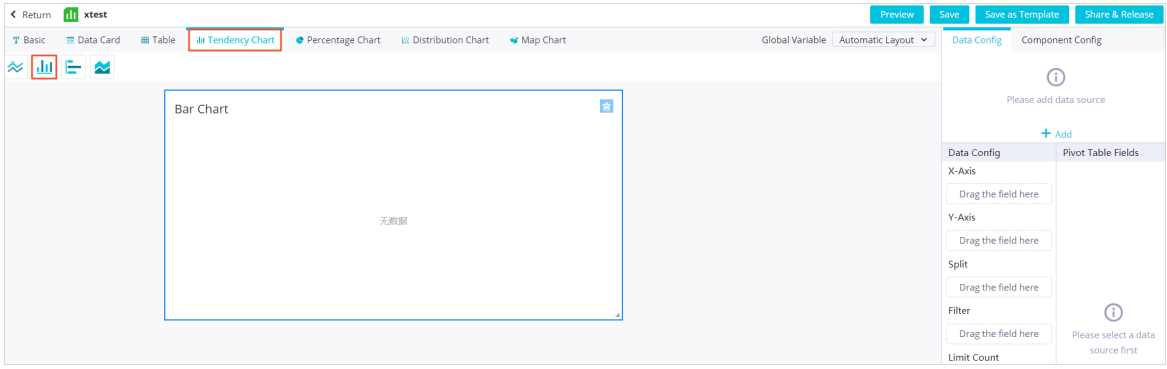
DataAnalysis provides 22 controls categorized into 7 types. You can choose the controls to use based on your business requirements. This topic describes how to add a line chart to a report, configure a data store for the line chart, and edit the line chart.

Go to the report editing page

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the **All Reports** section of the **Report** page, click the name of the report that you want to edit in the **File Name** column to go to the report editing page.

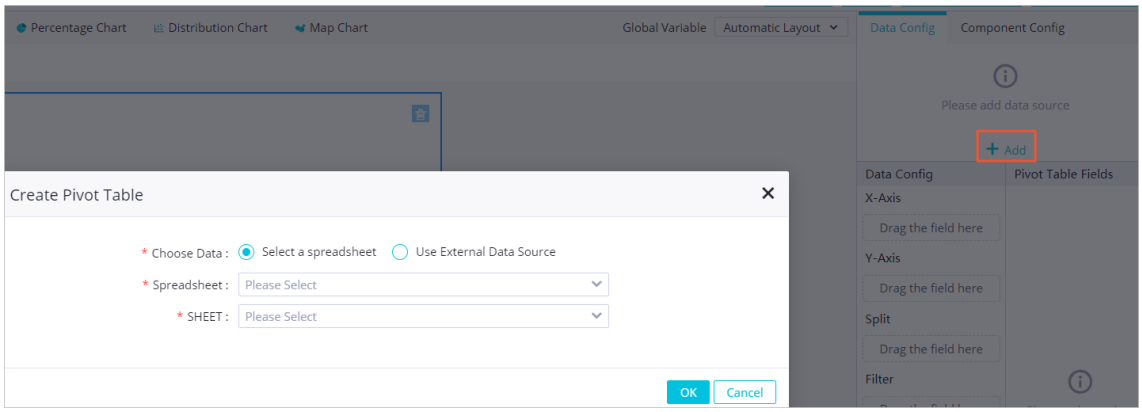
Add a line chart and configure a data store

1. On the report editing page, click **Tendency Chart** and drag the **Line Chart** control to the canvas.



2. On the **Data Config** tab on the right side of the report editing page, click **Add**.
3. In the **Create Pivot Table** dialog box, specify the range of the data to be analyzed.You can set Choose Data to **Select a spreadsheet** or **Use External Data Source** as needed.
   - **Select a spreadsheet**

You can specify a sheet from a workbook under the current account as the data store.



> **Note** A sheet of a pivot table cannot be used as the data store.

- **Use External Data Source**

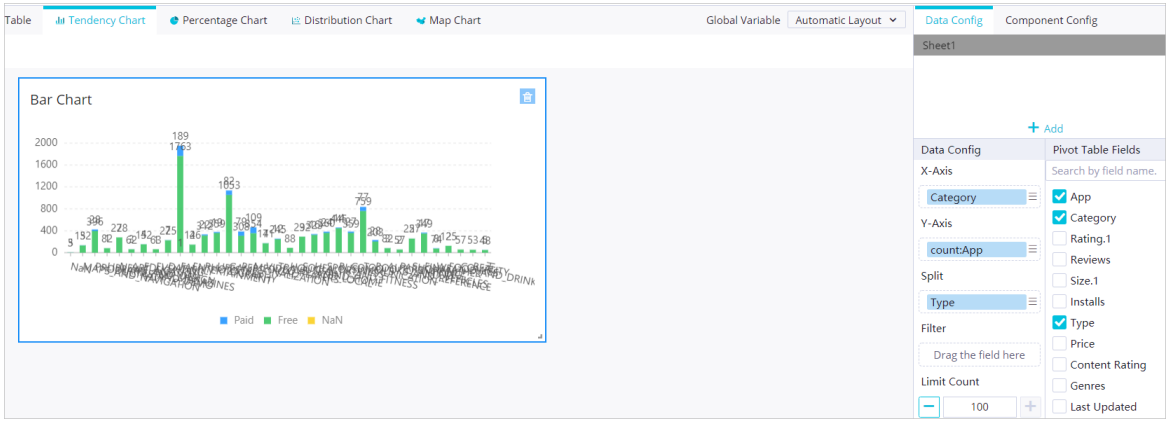
  Select a data store of the **MaxCompute**, **Mysql**, **Data Services**, **OSS**, or **PostgreSQL** type as needed. If you set Choose Data to Use External Data Source, make sure that the connection to the data store or the API for obtaining the source data has been configured. For more information, see Connection configuration.

Multiple charts can use the same data store in different ways without affecting each other. One chart can use only one data store. After you select a chart and drag fields from the Pivot Table Fields section to the Data Config section, the chart is associated with the data store.

Configure data for the line chart

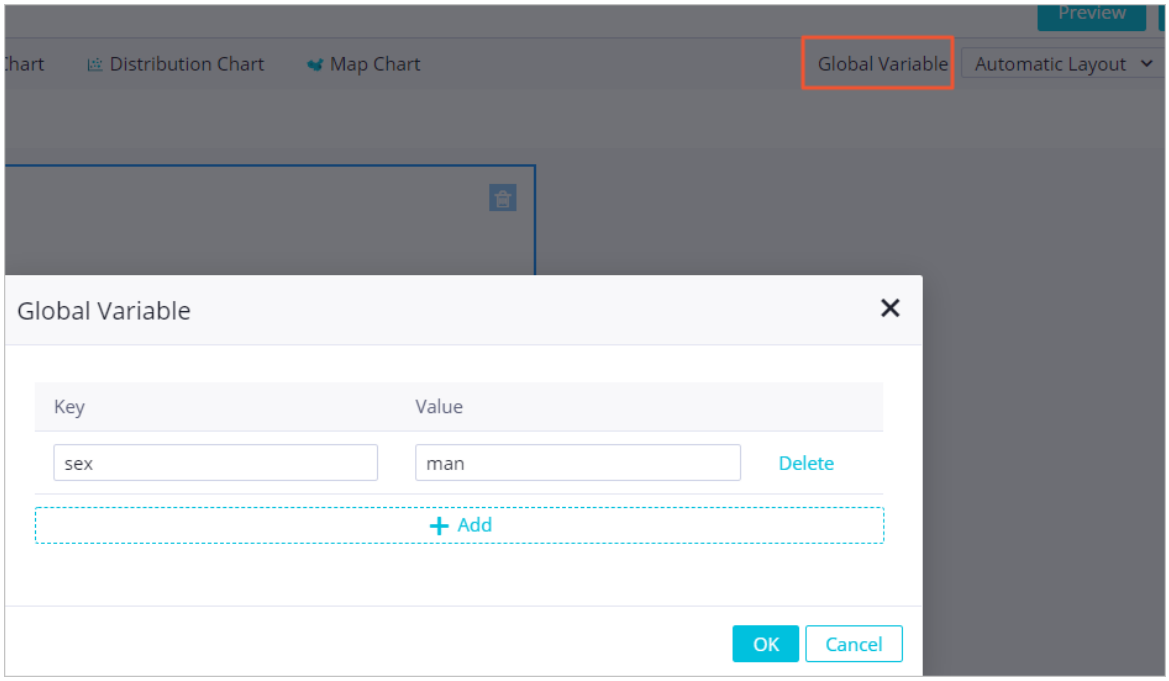
The parameters that need to be set vary with the chart type. For more information, see Charts. For example, you must specify the X-Axis and Y-Axis parameters for a line chart.

1. On the right side of the report editing page, drag fields from the **Pivot Table Fields** section to **X-Axis** and **Y-Axis** in the **Data Config** section.



   You can choose whether to set the **Split** parameter based on your business requirements.

2. Click **Global Variable** in the menu bar. In the **Global Variable** dialog box, specify **Key** and **Value**.In the Global Variable dialog box, you can specify custom filter conditions. For example, set **Key** to sex and **Value** to male. That is, set sex=male as the filter condition.



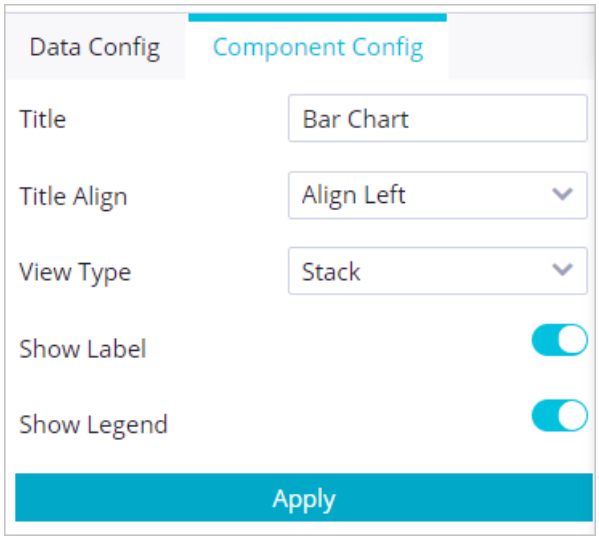
3. Click **OK**.
4. On the right side of the report editing page, drag fields from the **Pivot Table Fields** section to **Filter** in the Data Config section.
5. In the **Screening** dialog box, enter the variable name.

   > **Note** Variable names are in the ${} format.

6. Click **OK**. Only field values that meet the specified filter condition are displayed in the line chart.

Configure the line chart settings

To display the chart information more clearly, you can click the **Component Config** tab on the right side of the report editing page. On this tab, you can set the **Title**, **Title Align**, **Show Polyline Points**, **Show Label** and **Show Legend** parameters.



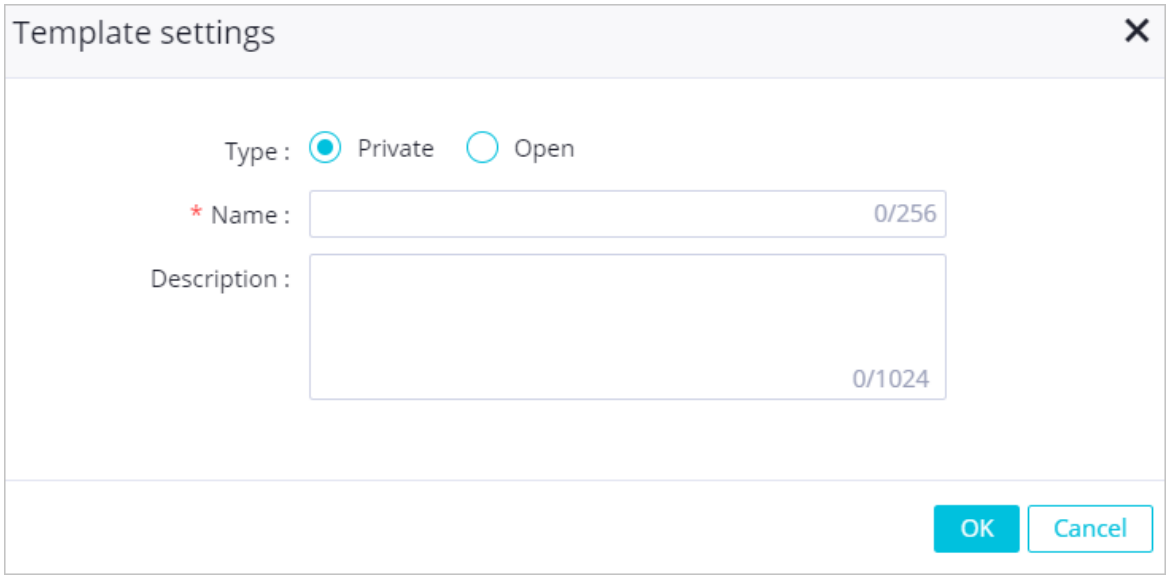
The parameters that need to be set vary with the control type. You can set the parameters based on the content to be displayed.

4.3. Save a report as a template

You can save an edited report as a template and use the template to create reports.

Procedure

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the **All Reports** section of the **Report** page, click the name of the report that you want to edit in the **File Name** column to go to the report editing page. If you create a report in this step, the report editing page appears after the report is created. For more information, see .
5. In the upper-right corner of the report editing page, click **Save as Template**.
6. On the **Preview** page, click **Next Step (Template settings)**.
7. In the **Template settings** dialog box, set the parameters.



Parameter	Description
Type	Specifies whether to show or hide the template for other users. Valid values: **Private** and **Open**.
Name	The name of the template. The name can be up to 256 characters in length.

Parameter	Description
Description	The description of the template. The description can be up to 1,024 characters in length.

8. Click **OK**.

4.4. Share a report

You can share your reports with all or specific users.

Procedure

1. Go to the **DataStudio** page.
   i. Log on to the DataWorks console.
   ii. In the left-side navigation pane, click **Workspaces**.
   iii. In the top navigation bar, select the region where your workspace resides, find the workspace, and then click **Data Analytics** in the Actions column.
2. On the DataStudio page, click the icon in the upper-left corner and choose **All Products > Data Development > DataAnalysis**.
3. On the DataAnalysis homepage, click **Experience Now**. The **Web Excel** page appears.
4. In the **All Reports** section of the **Report** page, click the name of the report that you want to edit in the **File Name** column to go to the report editing page. If you create a report in this step, the report editing page appears after the report is created. For more information, see Edit a report.
5. In the upper-right corner of the report editing page, click **Share & Release** to specify the users who can view this report.
   - **Visible to All**: To allow all users to view the report, turn on **Visible to All**.
   - **Users with Read Access**: To allow only specific users to view the report, turn off **Visible to All** and click **Add** in the **Users with Read Access** section. In the Share File with These Users dialog box, enter and select the names of the users to be granted the read permissions, and click **OK**.

     > **Note** You can grant the read permissions to up to 30 users.
6. Click **Share** in the dialog box.

5.Charts

5.1. Column charts

This topic describes the types of column charts and their examples.

A column chart is one of the most commonly used chart types in data analysis. It can be used to graphically display data that is arranged in columns or rows in a workbook.

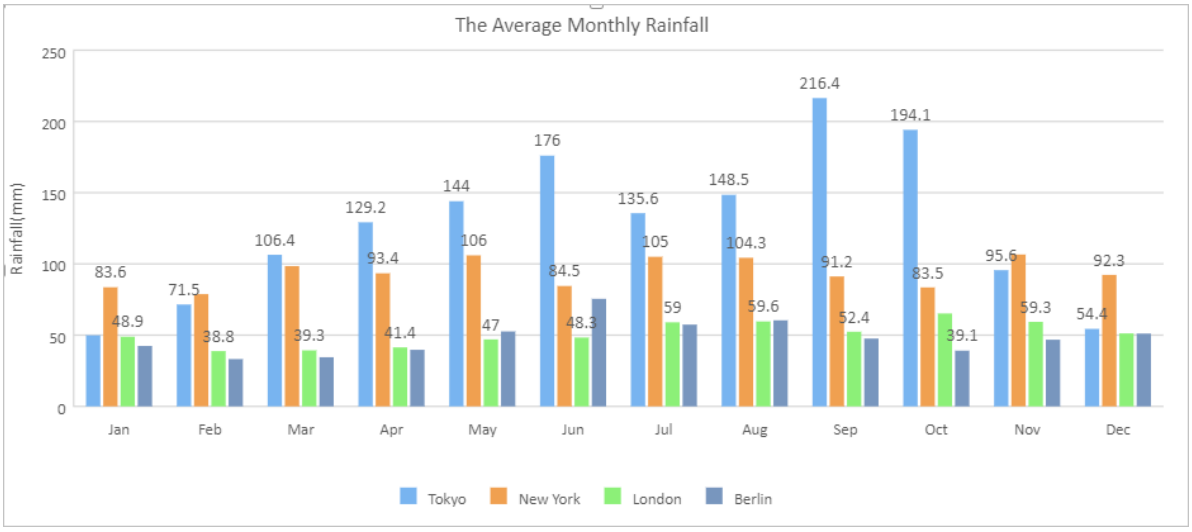
A column chart is used to show data comparisons among categories. For example, you can use a column chart to display the distribution of employees of different ages in a company. Then, you can check the number of the employees under 25 and the number of employees between 25 and 35 to analyze the ageing situation. In addition, you can use a column chart to show the change trend of data comparisons among several categories.

Clustered column chart

- Sample data

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Tokyo	49.9	71.5	106.4	129.2	144	176	135.6	148.5	216.4	194.1	95.6	54.4
New York	83.6	78.8	98.5	93.4	106	84.5	105	104.3	91.2	83.5	106.6	92.3
London	48.9	38.8	39.3	41.4	47	48.3	59	59.6	52.4	65.2	59.3	51.2
Berlin	42.4	33.2	34.5	39.7	52.6	75.5	57.4	60.4	47.6	39.1	46.8	51.1

- Sample chart

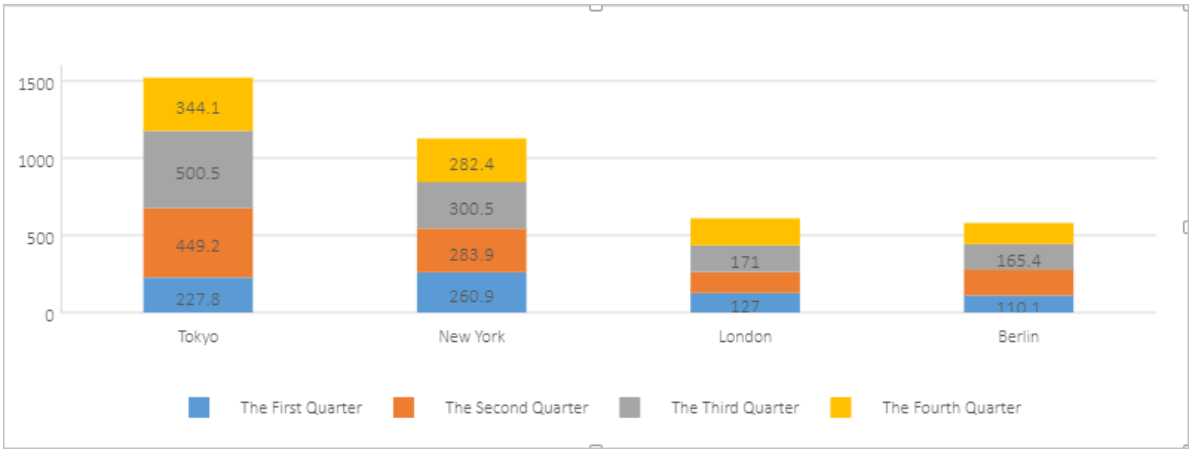


Stacked column chart

- Sample data

	Tokyo	New York	London	Berlin
The First Quarter	227.8	260.9	127	110.1
The Second Quarter	449.2	283.9	136.7	167.8
The Third Quarter	500.5	300.5	171	165.4
The Fourth Quarter	344.1	282.4	175.7	137

- Sample chart

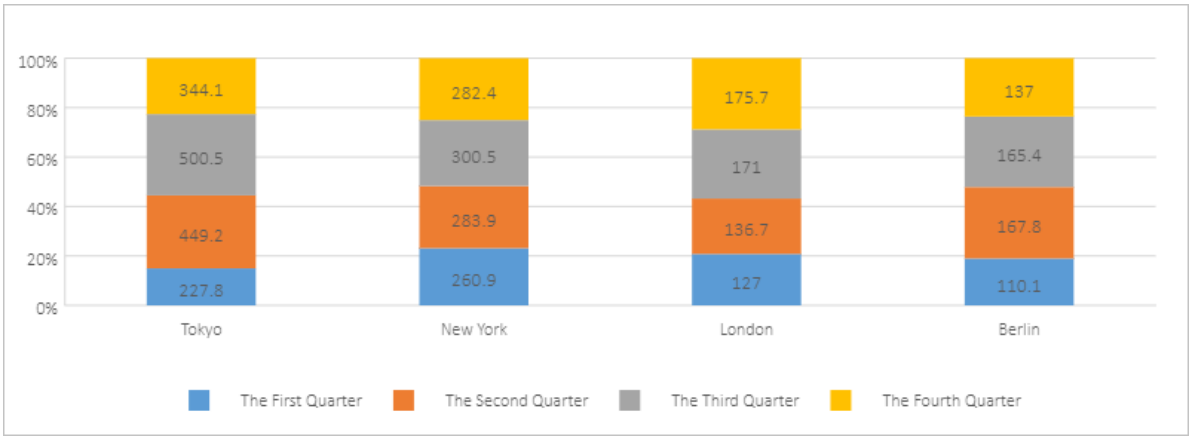


100% stacked column chart

- Sample data

	Tokyo	New York	London	Berlin
The First Quarter	227.8	260.9	127	110.1
The Second Quarter	449.2	283.9	136.7	167.8
The Third Quarter	500.5	300.5	171	165.4
The Fourth Quarter	344.1	282.4	175.7	137

- Sample chart



5.2. Line charts

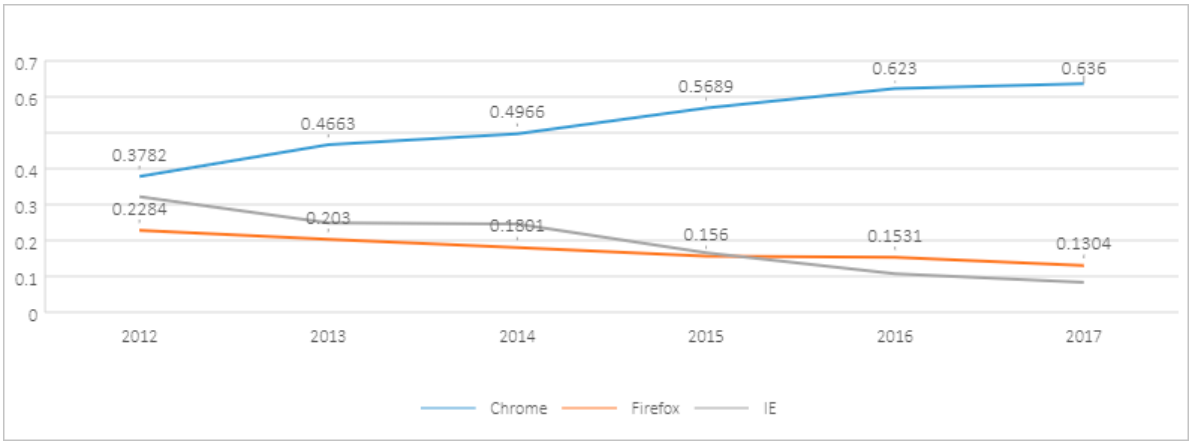
This topic describes the types of line charts and their examples.

A line chart consists of a horizontal axis that represents categories and a vertical axis that represents values. A line chart can be used to analyze the trend of changes over time, especially in the scenario where the data trend is more important than data values. Therefore, a line chart is suitable to show the trend of changes over equal intervals, for example, months, quarters, and fiscal years.

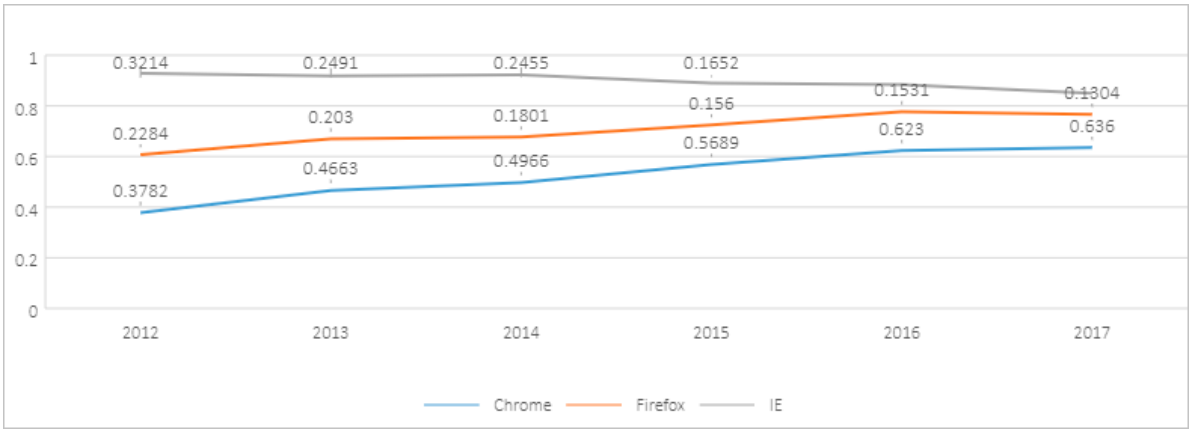
This topic uses the sample data listed in the following table to show different types of line charts.

	2012	2013	2014	2015	2016	2017
Chrome	0.3782	0.4663	0.4966	0.5689	0.623	0.636
Firefox	0.2284	0.203	0.1801	0.156	0.1531	0.1304
IE	0.3214	0.2491	0.2455	0.1652	0.1073	0.0834

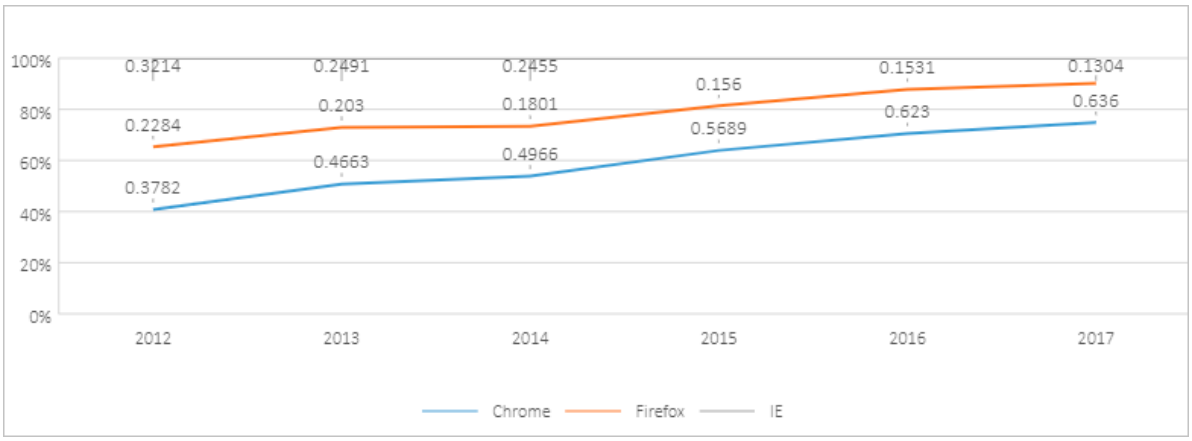
- **Line chart**



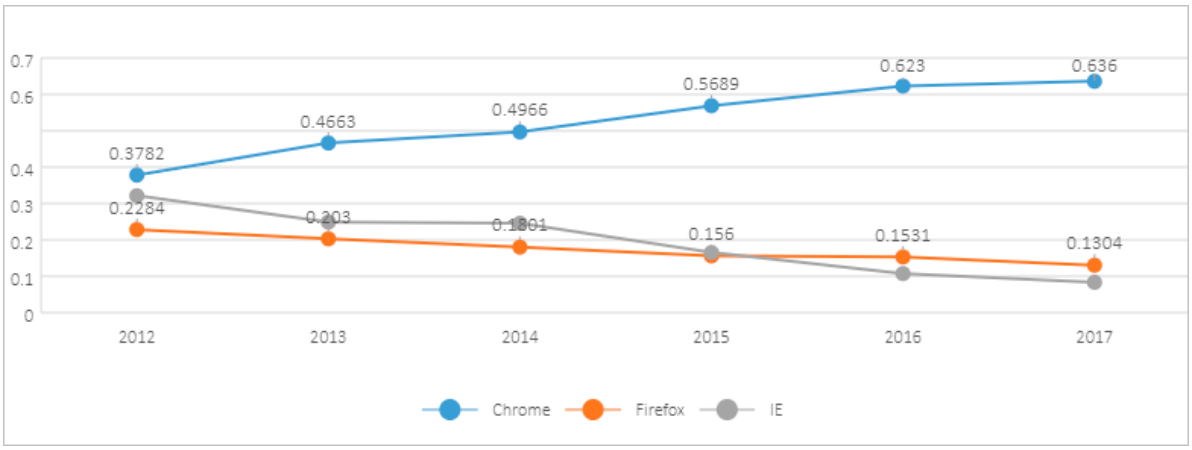
- **Stacked line chart**



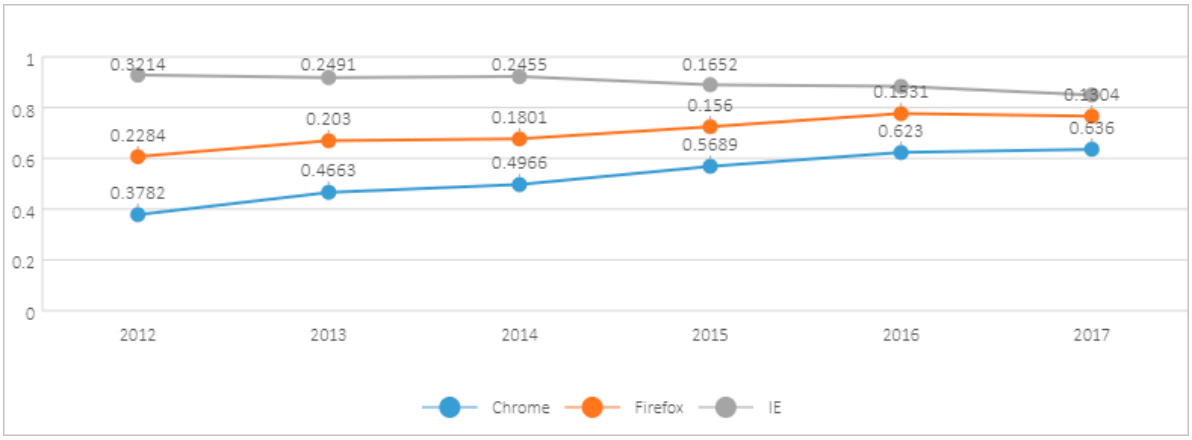
- 100% stacked line chart



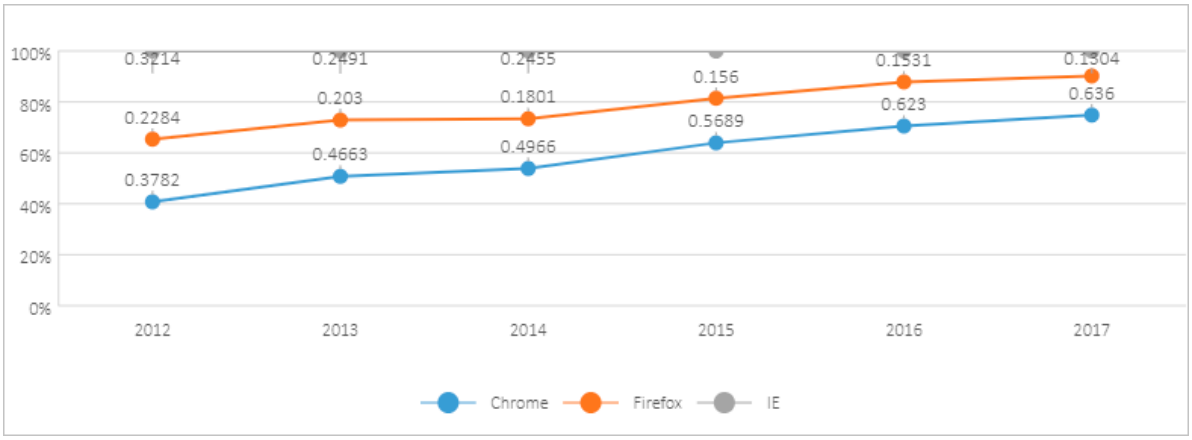
- Line chart with markers



- Stacked line chart with markers



- **100% stacked line chart with markers**



5.3. Pie charts

This topic describes the types of pie charts and their examples.

Pie chart

A pie chart can be used to graphically display data that is arranged in columns or rows in a workbook.

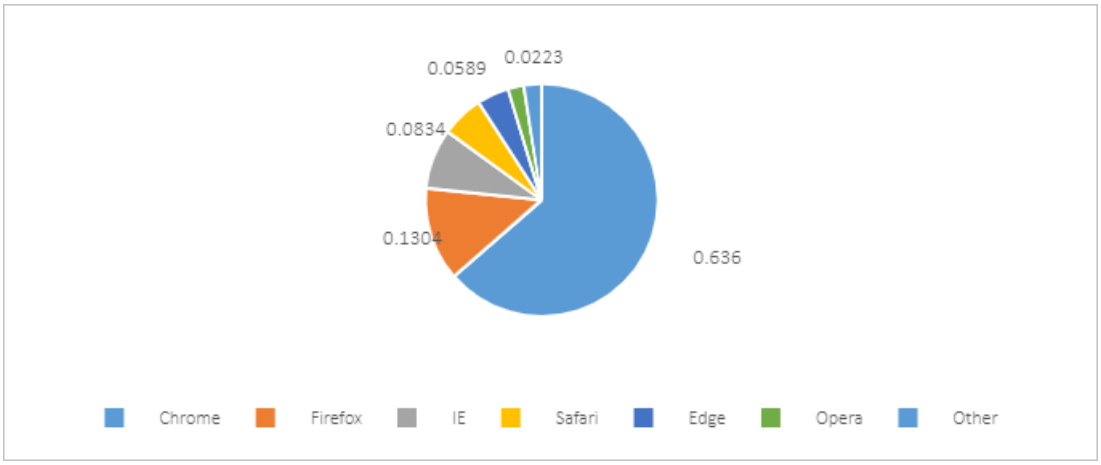
We recommend that you use a pie chart to show the ratios of different data categories to the total amount. For example, you can use a pie chart to show expected sales of products. You may find that the expected sales volume of product A accounts for the largest share of all product sales.

To maximize the effect of a pie chart, do not add over seven categories to a pie chart. For easy observation, we recommend that you place the categories in the clockwise direction and place the most important category near 12 o'clock in the pie chart. If all categories are equally important, you can sort the data categories in a descending order.

- Sample data

	Chrome	Firefox	IE	Safari	Edge	Opera	Other
2017	0.636	0.1304	0.0834	0.0589	0.0443	0.0223	0.0246

- Sample chart



Doughnut chart

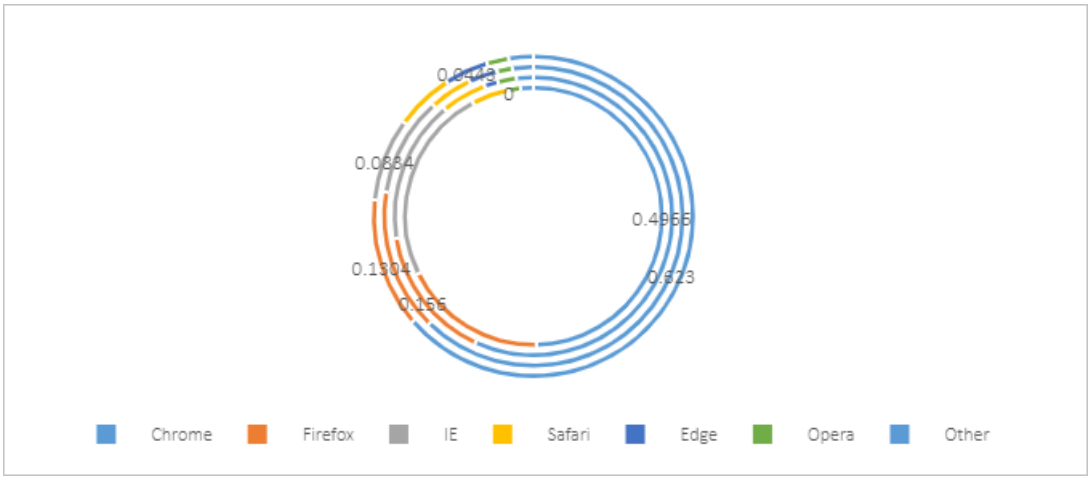
If each category has an almost equal share in a pie chart, it is difficult to compare these categories based on the area size. To show data comparisons more clearly, you can use a doughnut chart.

DataAnalysis supports both pie charts and doughnut charts.

- A pie chart can only display one data series.
- A doughnut chart displays data distribution in rings. A doughnut chart can display multiple data series. Each ring represents a data series.
- Sample data

	Chrome	Firefox	IE	Safari	Edge	Opera	Other
2014	0.4966	0.1801	0.2455	0.047	0	0.015	0.0158
2015	0.5689	0.156	0.1652	0.0529	0.0158	0.022	0.0192
2016	0.623	0.1531	0.1073	0.0464	0.0311	0.0166	0.0225
2017	0.636	0.1304	0.0834	0.0589	0.0443	0.0223	0.0246

- Sample chart



5.4. Area charts

This topic describes the types of area charts and their examples.

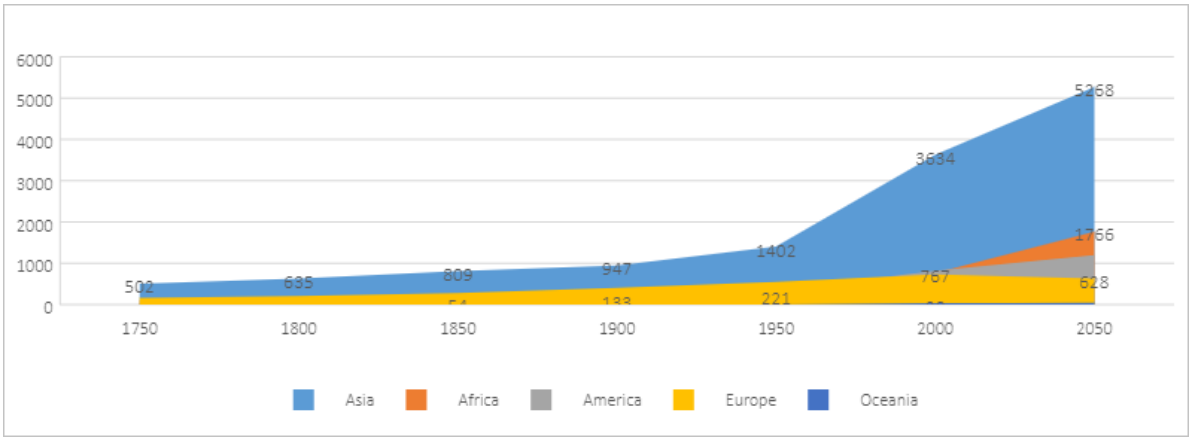
An area chart is a line chart with the areas below the lines filled with colors.

Similar to a line chart, an area chart emphasizes the changes over time, and is used to highlight the change trend. An area chart displays the trend through areas. We recommend that an area chart consist of up to five areas.

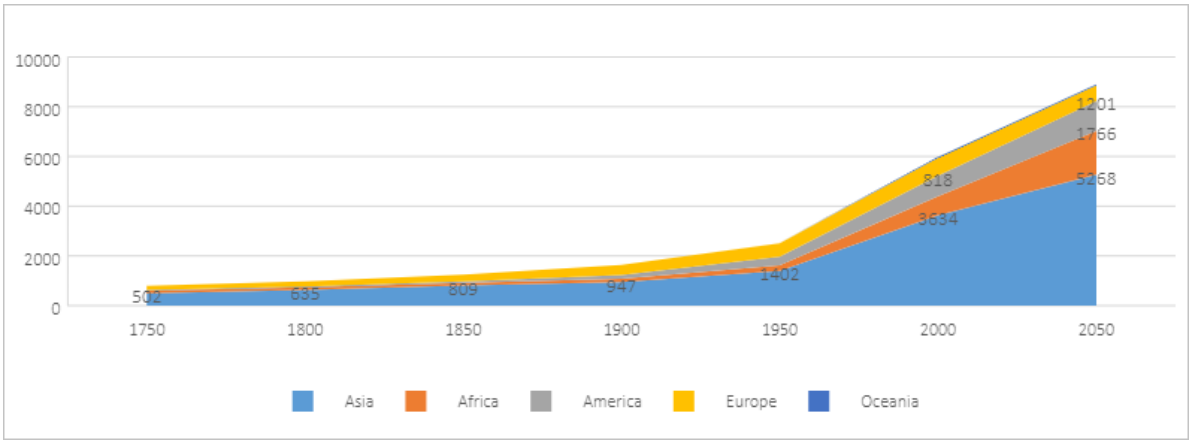
This topic uses the sample data listed in the following table to show different types of area charts.

	1750	1800	1850	1900	1950	2000	2050
Asia	502	635	809	947	1402	3634	5268
Africa	106	107	111	133	221	767	1766
America	18	31	54	156	339	818	1201
Europe	163	203	276	408	547	729	628
Oceania	2	2	2	6	13	30	46

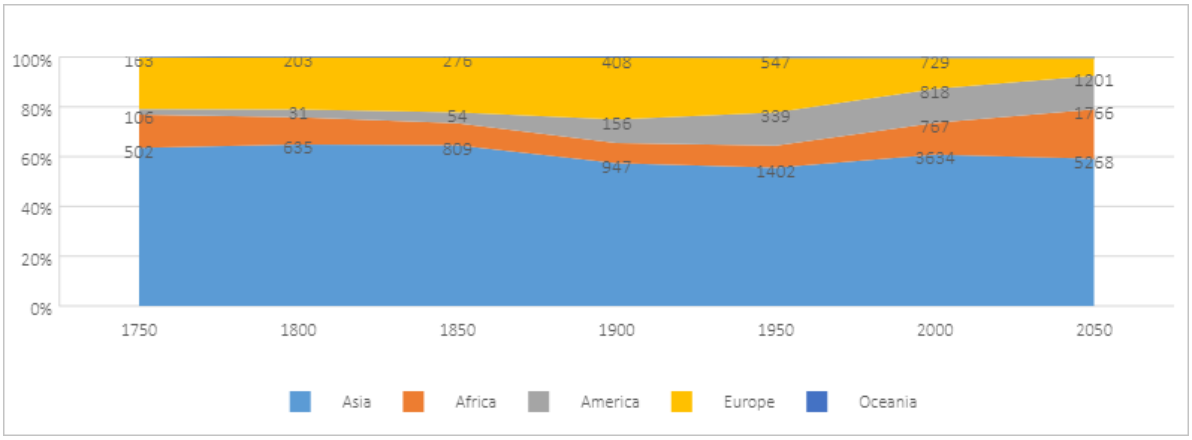
- Area chart



- Stacked area chart



- 100% stacked area chart



5.5. Horizontal bar charts

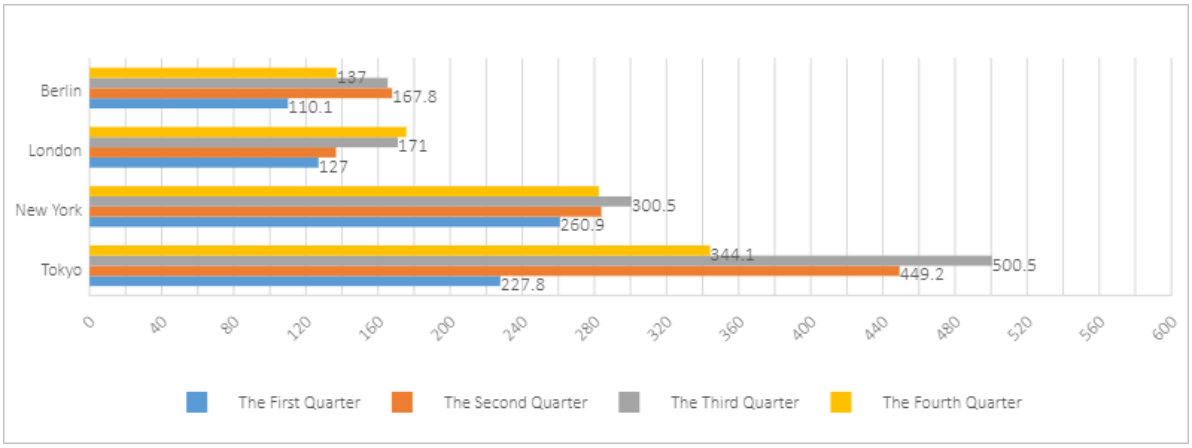
This topic describes the types of horizontal bar charts and their examples.

A horizontal bar chart shows comparisons between categories. A horizontal bar chart consists of a vertical axis that represents categories and a horizontal axis that represents values. You can rank the categories based on a specified order and highlight the top three or five categories. For example, you can use a horizontal bar chart to display the sales of popular commodities in the retail industry.

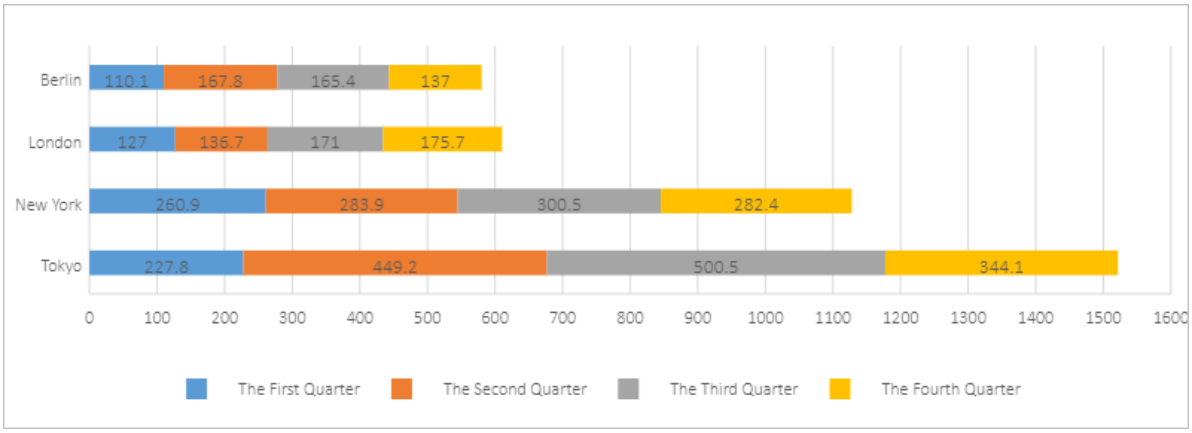
This topic uses the sample data listed in the following table to show different types of horizontal bar charts.

	Tokyo	New York	London	Berlin
The First Quarter	227.8	260.9	127	110.1
The Second Quarter	449.2	283.9	136.7	167.8
The Third Quarter	500.5	300.5	171	165.4
The Fourth Quarter	344.1	282.4	175.7	137

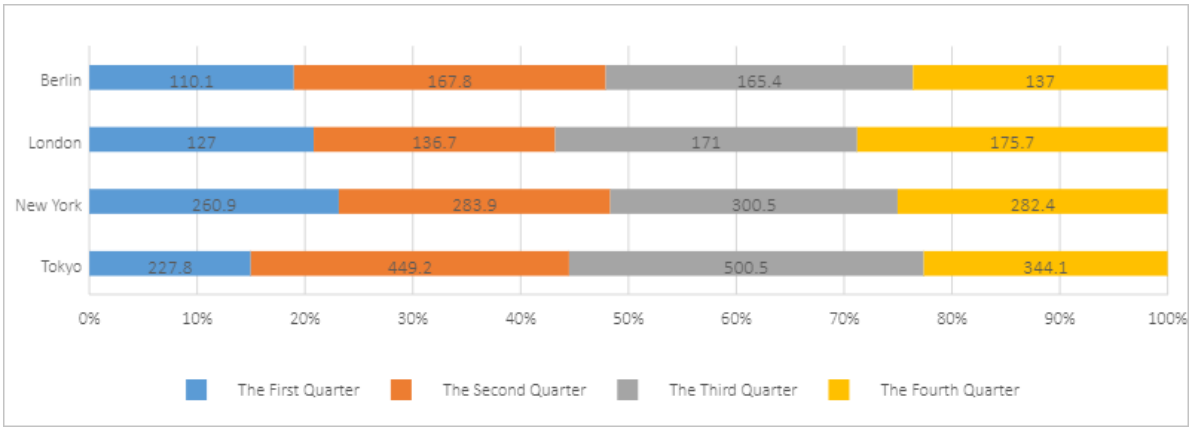
- Horizontal bar chart



- Stacked horizontal bar chart



- **100% stacked horizontal bar chart**



5.6. Scatter charts

This topic describes the types of scatter charts and their examples.

A scatter chart is often used to display the relationship between the values of variable X and those of variable Y. In a line chart, the X-axis represents different categories. In a scatter chart, the X-axis represents the values of a variable.

A scatter chart consists of a horizontal value axis for variable X and a vertical value axis for variable Y. It combines the values of variable X and those of variable Y to data points and shows them in irregular intervals or clusters. A scatter chart is typically used to show and compare numeric values in scientific, statistical, and engineering data.

You can use a scatter chart to compare values of typically two variables from five dimensions. The color, label, or size of each data point can be used to represent a dimension of values.

Scatter chart

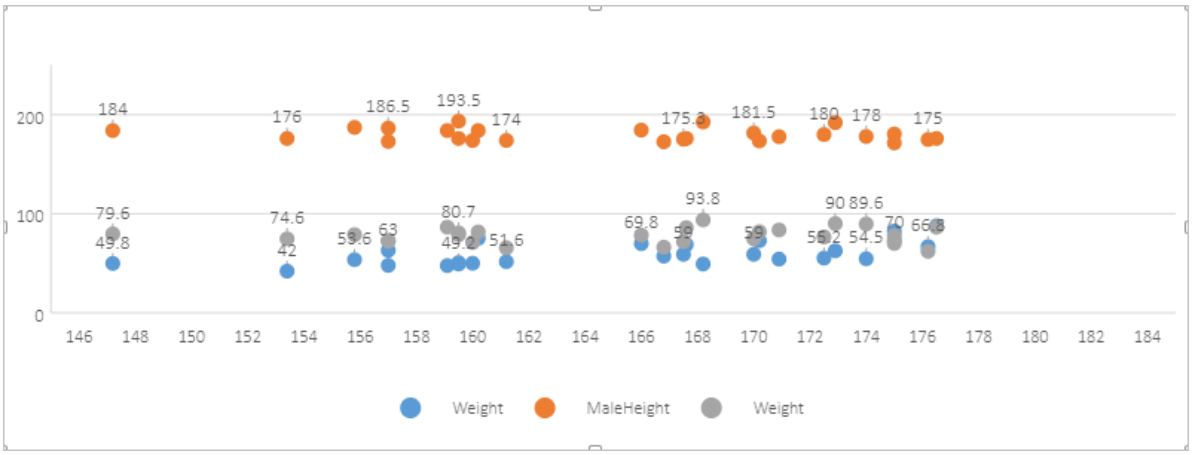
You can use a scatter chart to find out the relationship between variable X and variable Y.

- Sample data

Female		Male	
Height	Weight	Height	Weight
161.2	51.6	174	65.6

Female		Male	
Height	Weight	Height	Weight
167.5	59	175.3	71.8
159.5	49.2	193.5	80.7
157	63	186.5	72.6
155.8	53.6	187.2	78.8
170	59	181.5	74.8
159.1	47.6	184	86.4
166	69.8	184.5	78.4
176.2	66.8	175	62
160.2	75.2	184	81.6
172.5	55.2	180	76.6
170.9	54.2	177.8	83.6
172.9	62.5	192	90
153.4	42	176	74.6
160	50	174	71
147.2	49.8	184	79.6
168.2	49.2	192.7	93.8
175	73.2	171.5	70
157	47.8	173	72.4
167.6	68.8	176	85.9
159.5	50.6	176	78.8
175	82.5	180.5	77.8
166.8	57.2	172.7	66.2
176.5	87.8	176	86.4
170.2	72.8	173.5	81.8
174	54.5	178	89.6

- Sample chart



Scatter chart with smooth lines and markers

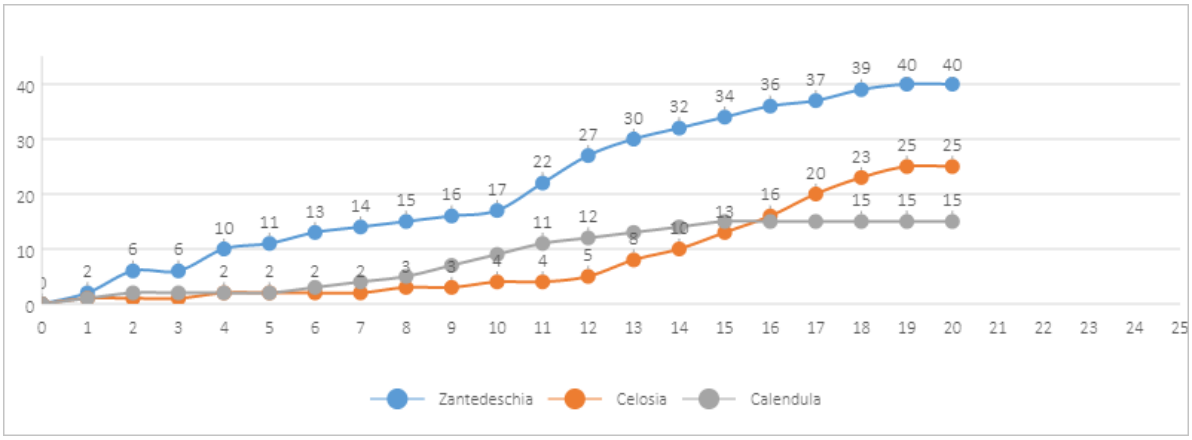
A scatter chart with smooth lines and markers displays smooth curves that connect data points.

- Sample data

Period	Zantedeschia	Celosia	Calendula
0	0	0	0
1	2	1	1
2	6	1	2
3	6	1	2
4	10	2	2
5	11	2	2
6	13	2	3
7	14	2	4
8	15	3	5
9	16	3	7
10	17	4	9
11	22	4	11
12	27	5	12
13	30	8	13
14	32	10	14
15	34	13	15

Period	Zantedeschia	Celosia	Calendula
16	36	16	15
17	37	20	15
18	39	23	15
19	40	25	15
20	40	25	15

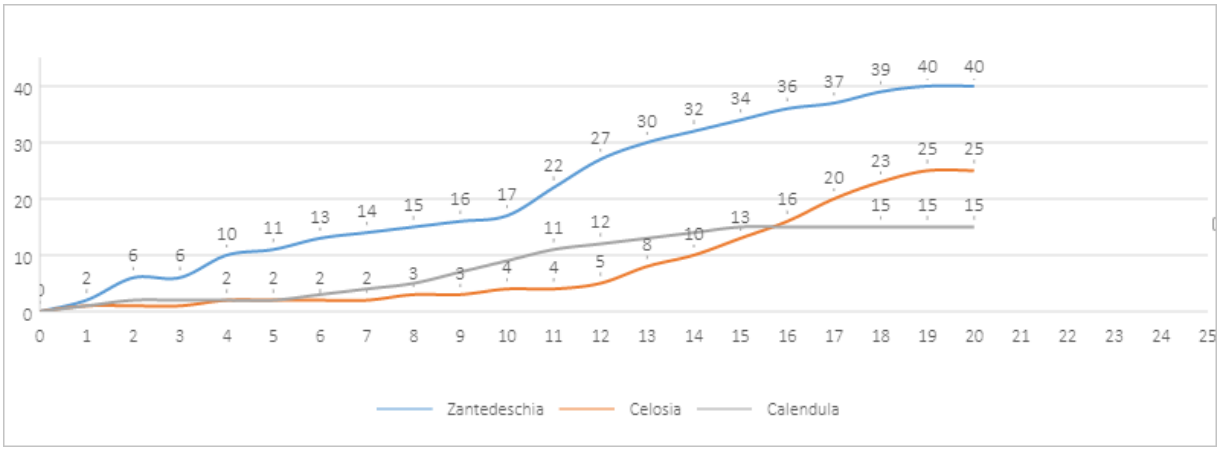
- Sample chart



Scatter chart with smooth lines

A scatter chart with smooth lines displays smooth curves that connect data points but does not display markers.

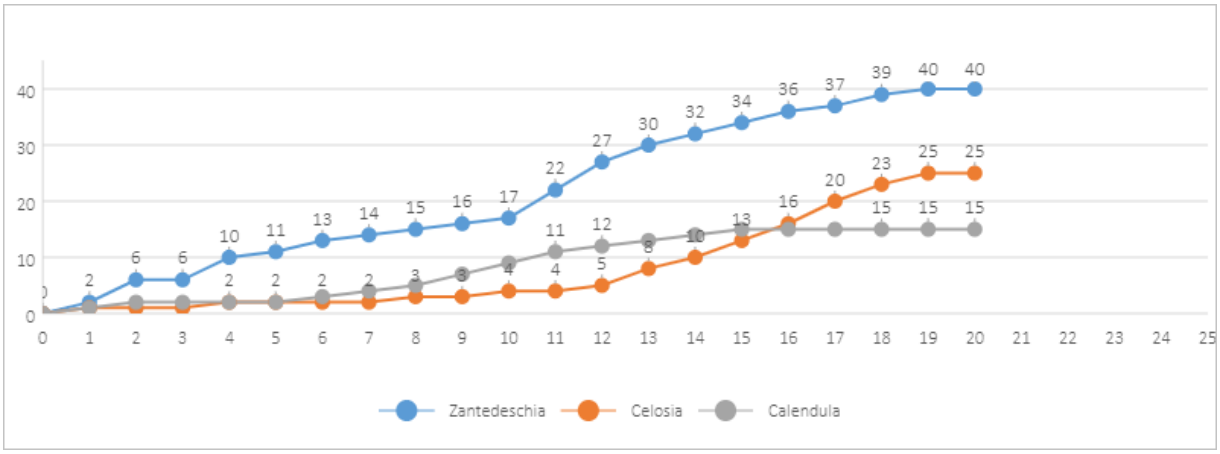
The following figure shows a scatter chart with smooth lines, which displays the same data as the preceding scatter chart with smooth lines and markers.



Scatter chart with straight lines and markers

A scatter chart with straight lines and markers displays straight lines that connect data points.

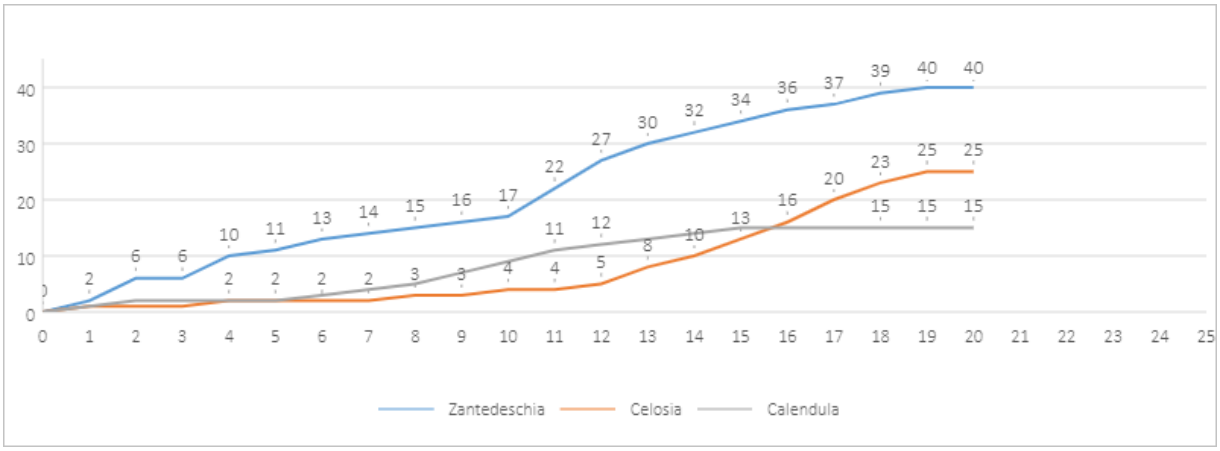
The following figure shows a scatter chart with straight lines and markers, which displays the same data as the preceding scatter chart with smooth lines and markers.



Scatter chart with straight lines

A scatter chart with straight lines displays straight lines that connect data points but does not display markers.

The following figure shows a scatter chart with straight lines, which displays the same data as the preceding scatter chart with smooth lines and markers.



Bubble chart

A bubble chart is a variation of a scatter chart. In a bubble chart, bubbles represent data points and the size of the bubbles represents an additional dimension of data. Similar to a scatter chart, a bubble chart consists of a horizontal value axis for variable X and a vertical value axis for variable Y. Different from the scatter chart, the bubble chart also displays the values of another variable, variable Z, which represents the size of the bubbles.

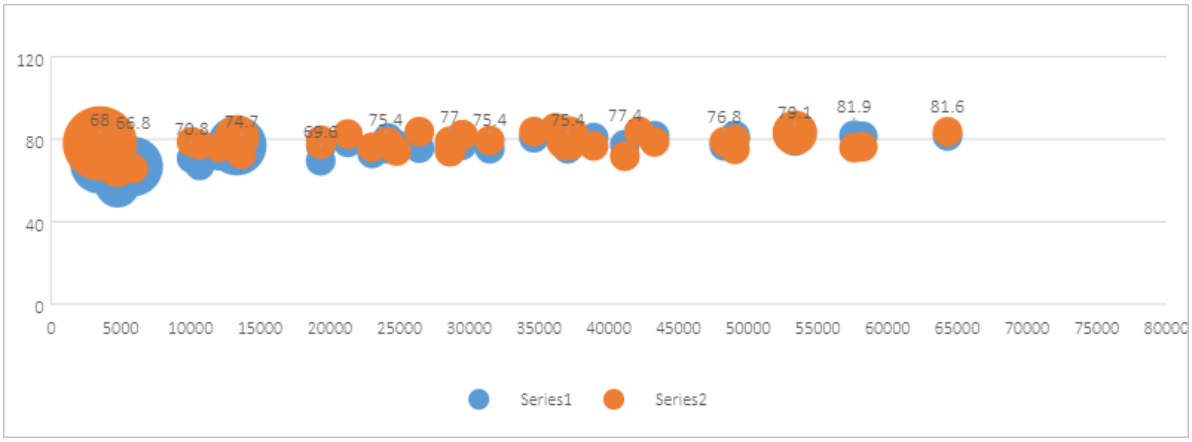
If your data contains three data series and each series contains a set of values, you can use a bubble chart instead of a scatter chart to display the data. The size of the bubbles is specified by the values in the third data series. A bubble chart is typically used to display financial data. The bubbles in different size bring strong visual impact.

- Sample data

28604	77	17096869	74	67096869
41163	77.4	27662440	71.8	47662440
3516	68	1154605773	78	1654605773
13670	74.7	10582082	72.7	69582082
28599	75	4986705	79	1986705
29476	77.1	56943299	82.1	26943299
31476	75.4	78958237	79.4	98958237
28666	78.1	254830	74.1	954830
4777	57.7	870601776	67.6	570601776
29550	79.1	122249285	82.1	22249285
5076	67.9	20194354	64.9	40194354
12087	72	42972254	76	342972254
24021	75.4	3397534	78.4	1397534
48296	76.8	4240375	78.8	14240375
1088	70.8	38195258	78.7	18195258
19349	69.6	147568552	77.6	234568552
10670	67.3	53994605	77.3	83994605
26424	75.7	57110117	83.7	86110117
37062	75.4	252847810	80.4	652847810
49056	81.8	23968973	79.8	63968973
43294	81.7	35939927	78.7	15939927
13334	76.9	1376048943	80.9	976048943
21291	78.5	11389562	82.5	151389562
38923	80.8	5503457	76.8	1503457
57599	81.9	64395345	75.9	34395345
49053	81.1	80688545	75.1	20688545
42182	82.8	329425	83.8	1329425

28604	77	17096869	74	67096869
5903	66.8	1311050527	65.8	311050527
36162	83.5	126573481	85.5	326573481
4390	71.4	25155317	77.4	55155317
34644	80.7	50293439	83.7	20293439
24186	80.6	4528526	78.6	13528526
64304	81.6	5210967	83.6	3210967
24787	77.3	38611794	74.3	88611794
23038	73.13	143456918	76.13	83456918
19360	76.5	78665830	79.5	58665830
58225	81.4	64715810	76.4	84715810
53354	79.1	321773631	83.1	721773631

- Sample chart



5.7. Stock charts

This topic describes the types of stock charts and their examples.

A stock chart can be used to show fluctuations of the data, such as stock prices and daily rainfall.

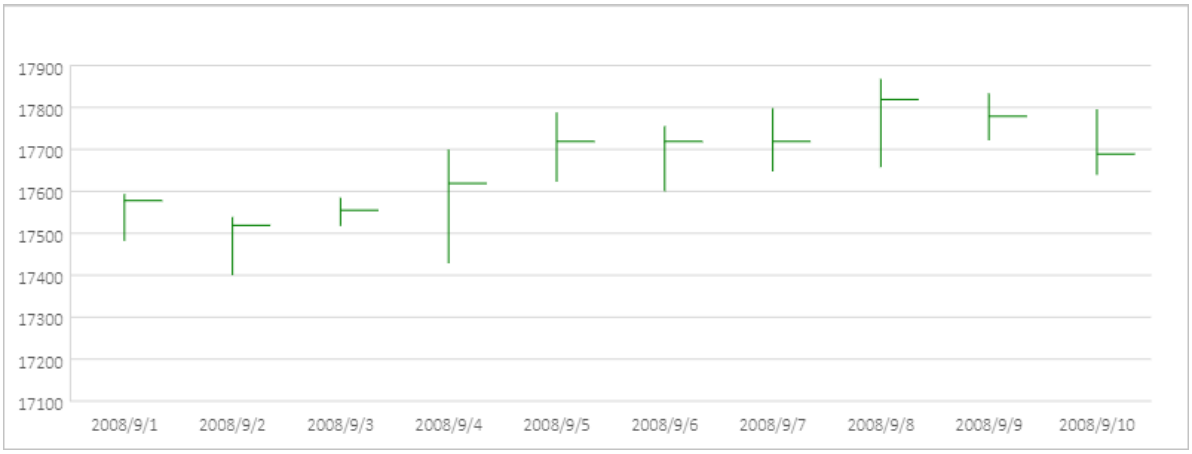
High-low-close chart

To create a high-low-close chart, arrange your data with the column headings in the following order: High, Low, and Close. Use dates and stock names as the labels of the chart.

- Sample data

Date	High	Low	Close
2008/9/1	17592.76	17482.76	17577.94
2008/9/2	17538.76	17400.76	17518.94
2008/9/3	17584.76	17517.76	17554.94
2008/9/4	17698.76	17428.76	17618.94
2008/9/5	17786.76	17623.76	17718.94
2008/9/6	17754.71	17600.76	17718.94
2008/9/7	17797.76	17647.76	17718.94
2008/9/8	17867.76	17657.76	17818.94
2008/9/9	17832.76	17721.76	17778.94
2008/9/10	17795.76	17639.76	17688.94

- Sample chart



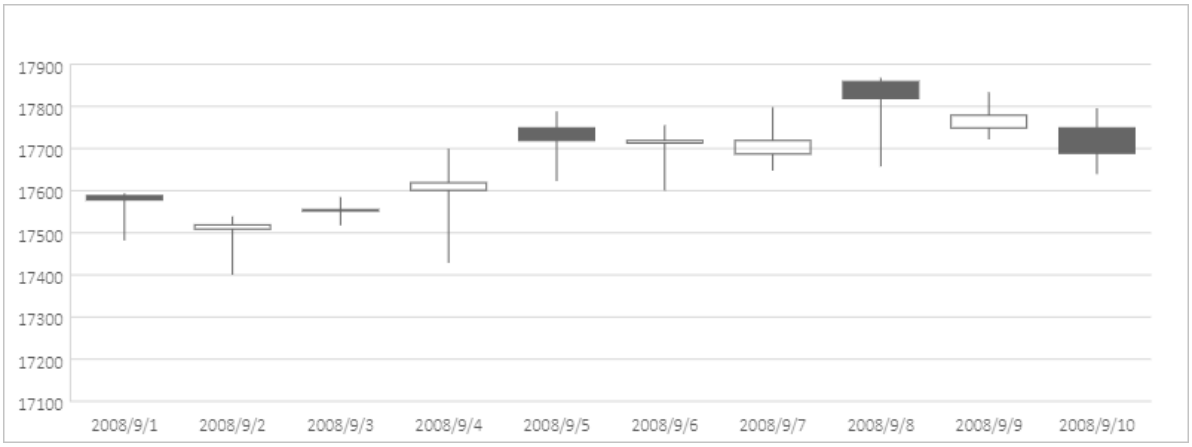
Open-high-low-close chart

- Sample data

Date	Open	High	Low	Close
2008/9/1	17587.94	17592.76	17482.76	17577.94
2008/9/2	17508.94	17538.76	17400.76	17518.94
2008/9/3	17551.94	17584.76	17517.76	17554.94
2008/9/4	17600.94	17698.76	17428.76	17618.94
2008/9/5	17748.94	17786.76	17623.76	17718.94

Date	Open	High	Low	Close
2008/9/6	17712.94	17754.71	17600.76	17718.94
2008/9/7	17686.94	17797.76	17647.76	17718.94
2008/9/8	17858.94	17867.76	17657.76	17818.94
2008/9/9	17748.94	17832.76	17721.76	17778.94
2008/9/10	17748.94	17795.76	17639.76	17688.94

- Sample chart

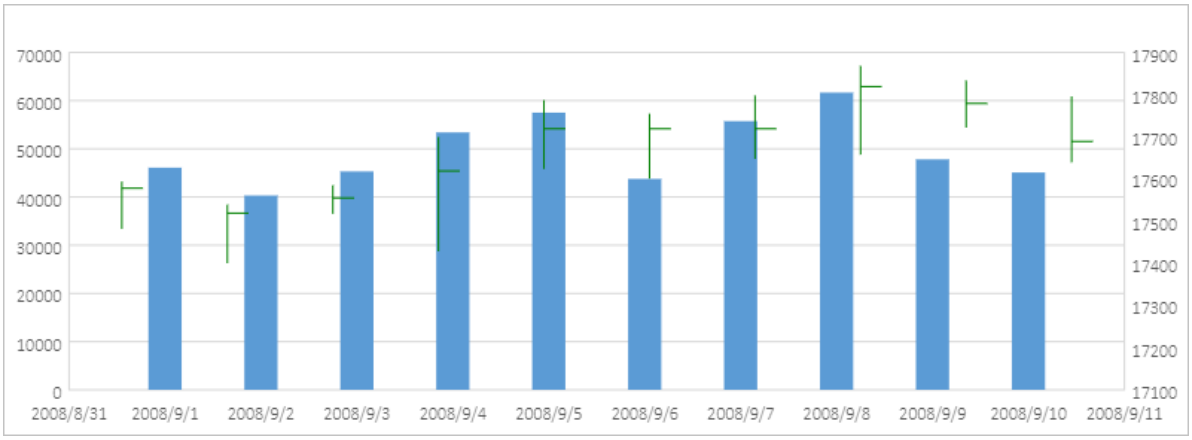


Volume-high-low-close chart

- Sample data

Date	Volume	High	Low	Close
2008/9/1	46085	17592.76	17482.76	17577.94
2008/9/2	40314	17538.76	17400.76	17518.94
2008/9/3	45308	17584.76	17517.76	17554.94
2008/9/4	53401	17698.76	17428.76	17618.94
2008/9/5	57500	17786.76	17623.76	17718.94
2008/9/6	43756	17754.71	17600.76	17718.94
2008/9/7	55737	17797.76	17647.76	17718.94
2008/9/8	61668	17867.76	17657.76	17818.94
2008/9/9	47815	17832.76	17721.76	17778.94
2008/9/10	45085	17795.76	17639.76	17688.94

- Sample chart

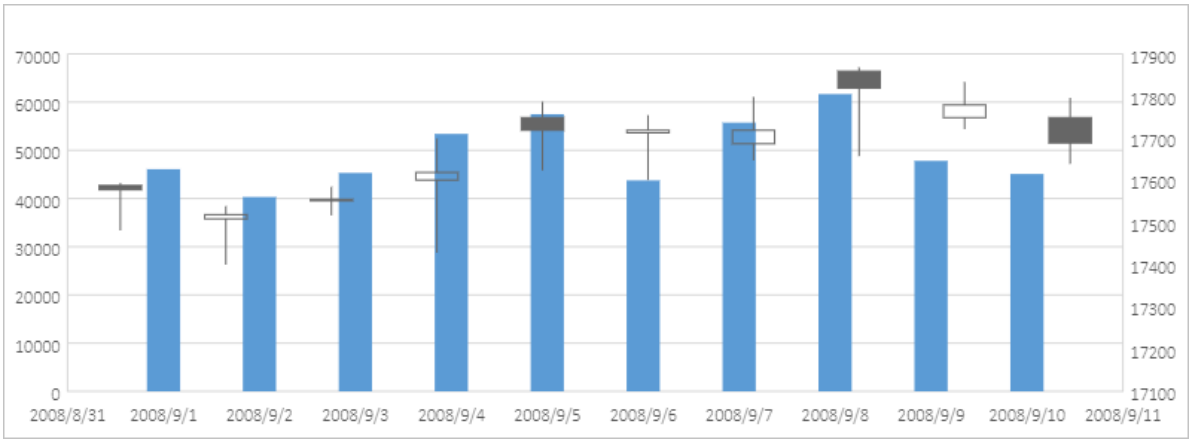


Volume-open-high-low-close chart

- Sample data

Date	Volume	Open	High	Low	Close
2008/9/1	46085	17587.94	17592.76	17482.76	17577.94
2008/9/2	40314	17508.94	17538.76	17400.76	17518.94
2008/9/3	45308	17551.94	17584.76	17517.76	17554.94
2008/9/4	53401	17600.94	17698.76	17428.76	17618.94
2008/9/5	57500	17748.94	17786.76	17623.76	17718.94
2008/9/6	43756	17712.94	17754.71	17600.76	17718.94
2008/9/7	55737	17686.94	17797.76	17647.76	17718.94
2008/9/8	61668	17858.94	17867.76	17657.76	17818.94
2008/9/9	47815	17748.94	17832.76	17721.76	17778.94
2008/9/10	45085	17748.94	17795.76	17639.76	17688.94

- Sample chart


70000
60000
50000
40000
30000
20000
10000
0
17900
17800
17700
17600
17500
17400
17300
17200
17100
2008/8/31
2008/9/1
2008/9/2
2008/9/3
2008/9/4
2008/9/5
2008/9/6
2008/9/7
2008/9/8
2008/9/9
2008/9/10
2008/9/11

6.Analyze data

DataWorks allows you to analyze data in online mode. In this topic, the MaxCompute public dataset is used to describe how to perform this operation.

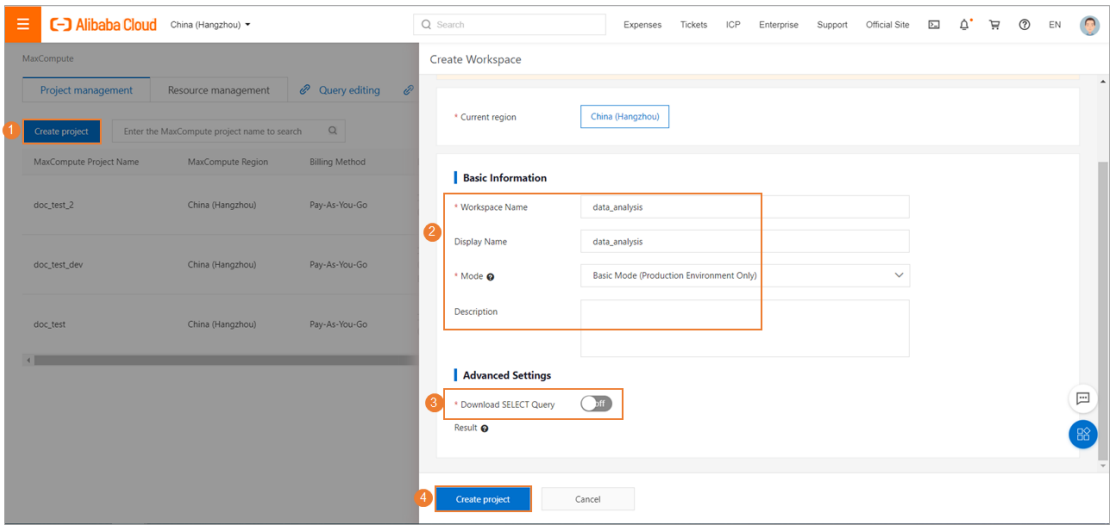
Procedure

1. Preparations: Prepare an environment. Before you prepare an environment, make sure that you understand all the requirements on the environment for data queries and analytics.
2. Query data: Query data from two tables in the MaxCompute public dataset.
3. Data analytics and sharing: Use web Excel to sort query results and perform pivoting for the results.

Preparations

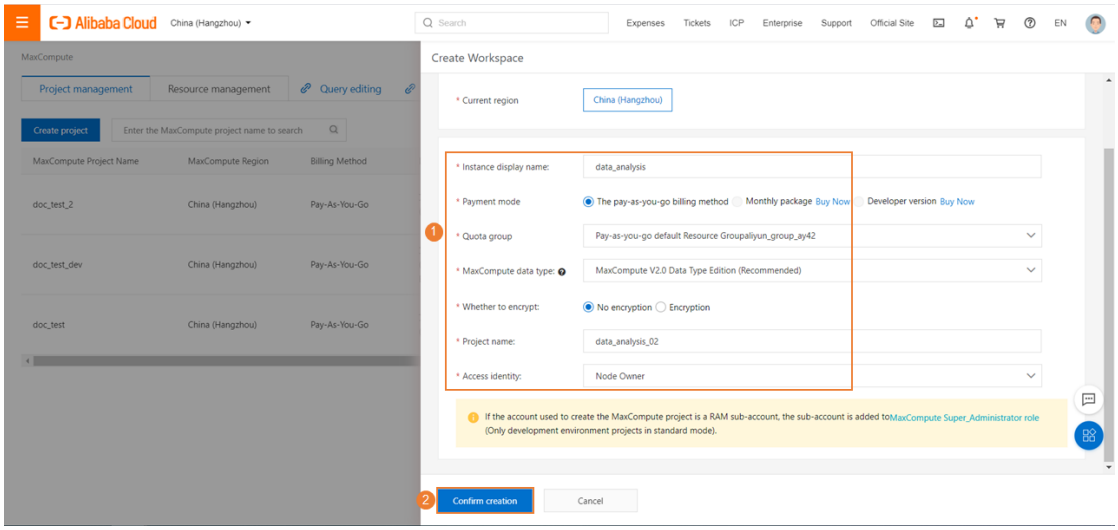
Activate MaxCompute and DataWorks that are deployed in the same region, and create a DataWorks workspace and a MaxCompute project. Skip this step if you already have an environment that meets the preceding requirements.

1. Activate MaxCompute and DataWorks.
   i. Log on to Alibaba Cloud, go to the product page of Alibaba Cloud MaxCompute, and then click **Buy Now**.
   ii. Configure parameters such as Region, read and select the terms of service, and complete payment as prompted.
2. Create a DataWorks workspace and a MaxCompute project, and associate them.
   i. Log on to the MaxCompute console and click **Create Project** on the Project management tab.
   ii. Create a DataWorks workspace.In the Create Workspace panel, configure parameters and click **Create project** in the lower-left corner.



- **Mode**: In this topic, Basic Mode (Production Environment Only) is selected. You can set this parameter based on your requirements.
- **Advanced Settings**: If you do not need to download data to your on-premises machine, we recommend that you turn off **Download SELECT Query** to prevent downloaded data from being forwarded. This improves data security.

iii. Create a MaxCompute project.Configure parameters and click **Confirm creation** in the lower-left corner.



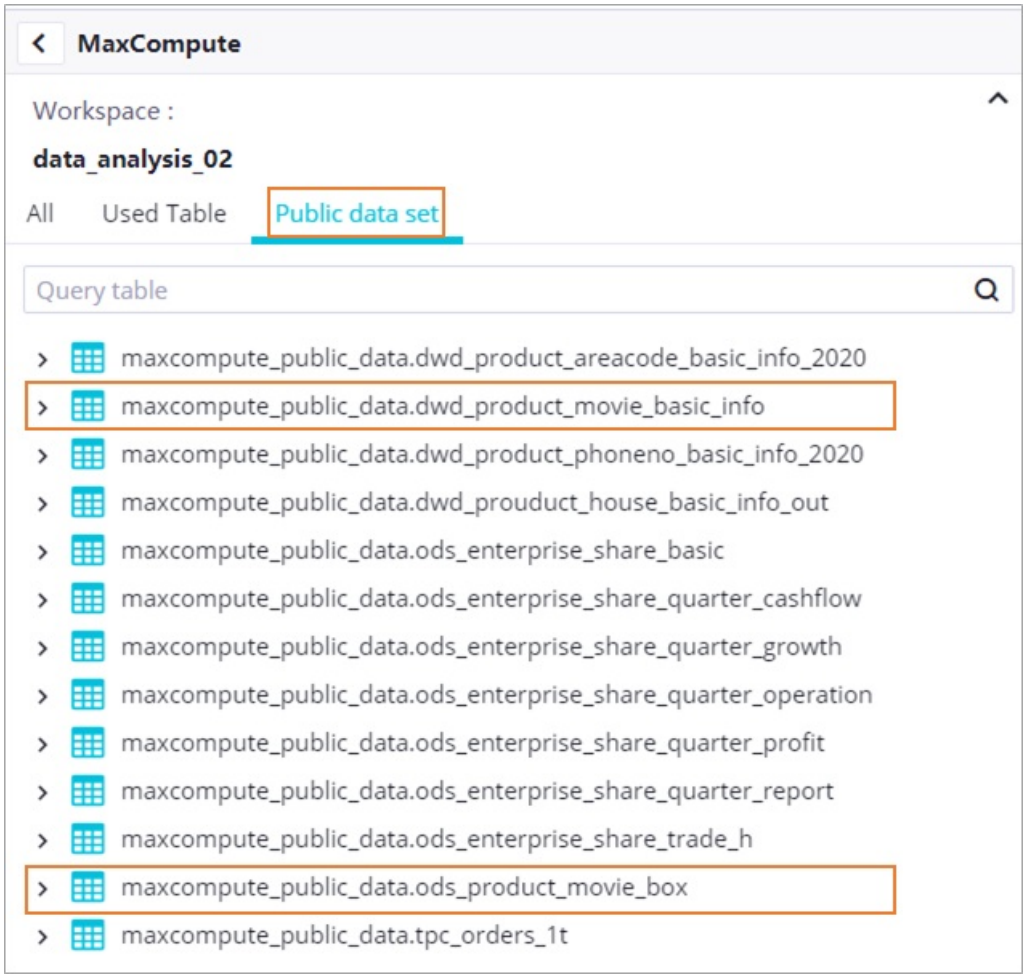
- **Payment mode**: You can use the default mode or enable another mode based on your requirements. Default value: The pay-as-you-go billing method.
- **Access identity**: For a data analytics project, we recommend that you select **Node Owner**. This indicates that the system verifies the permissions of the logon accounts of all members before the members can commit nodes.

iv. After the system displays the "Created successfully" message, click **Go to list view** to view the created workspace.

Query data

This section describes how to use the online data analytics feature provided by DataWorks to query data from the MaxCompute public dataset.

The MaxCompute public dataset provides a variety of public data for verification. This topic uses the maxcompute_public_data.dwd_product_movie_basic_info and maxcompute_public_data.ods_product_movie_box tables to collect statistics on the daily box office charts during the Spring Festival (Chinese New Year) in 2017.



- maxcompute_public_data.dwd_product_movie_basic_info: stores the basic information of movies, including the movie title, director, screenwriter, leading actors, and movie type.
- maxcompute_public_data.ods_product_movie_box: stores the basic information of the box office for movies, including the movie title, daily box office gross, and total box office gross.

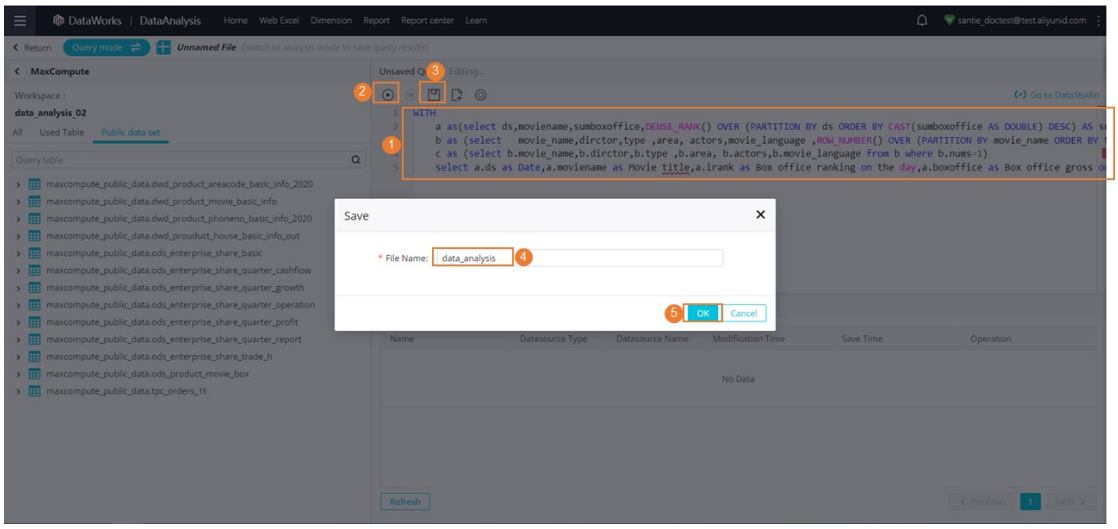
To collect statistics on the daily box office charts during the Spring Festival, you must associate the two tables. The following description provides the related operations and sample code.

1. In the MaxCompute console, click **Query editing**.
2. In the Select Datasource dialog box, select a data source.Select MaxCompute for Type and the workspace created in Preparations for Workspace. Then, click OK.
3. In the left-side navigation pane of the page that appears, find the maxcompute_public_data.dwd_product_movie_basic_info and maxcompute_public_data.ods_product_movie_box tables on the Public data set tab and view the fields of the tables.
    i. Expand a table and move the pointer over a field name to view the description of the field.
    ii. Right-click the table and select **Data Preview**. In the Table Details dialog box, 20 data records are randomly displayed.

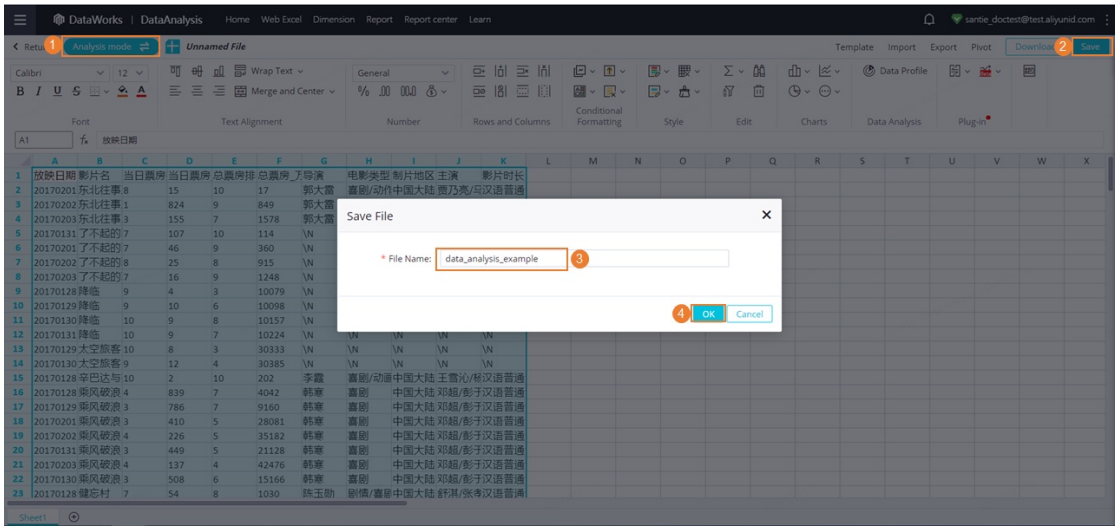
4. In the code editor, edit code to query data.In this example, the period from January 28, 2017 to February 3, 2017 is used as the official period of the Spring Festival. The daily box office gross of movies released during this period and the key information of the movies are queried. The query results are stored in a query result table, which is used for subsequent online data analytics.
   i. Edit query code in the code editor on the right.Sample code:

```
WITH
  a as(select ds,moviename,sumboxoffice,DENSE_RANK() OVER (PARTITION BY ds ORDER BY CAST(sumboxoffice AS DOUBLE) DESC) AS srank,boxoffice,irank from maxcompute_public_data.ods_product_movie_box WHERE ds>='20170128' and ds<='20170203' ),
  b as (select  movie_name,dirctor,type ,area, actors,movie_language ,ROW_NUMBER() OVER (PARTITION BY movie_name ORDER BY type DESC) AS nums from maxcompute_public_data.dwd_product_movie_basic_info where ds>='20170128' and ds<='20170203' ),
  c as (select b.movie_name,b.dirctor,b.type ,b.area, b.actors,b.movie_language from b where b.nums=1)
  select a.ds as 放映日期,a.moviename as 影片名,a.irank as 当日票房排名,a.boxoffice as 当日票房_万 ,a.srank as 总票房排名,a.sumboxoffice as 总票房_万,c.dirctor as 导演,c.type as 电影类型,c.area as 制片地区,c.actors as 主演,c.movie_language as 影片时长 from a LEFT join c on a.moviename=c.movie_name;
```

   ii. Click the Run icon to run the query code.
   iii. After the code is successfully run, click the Save query icon to save the query code.



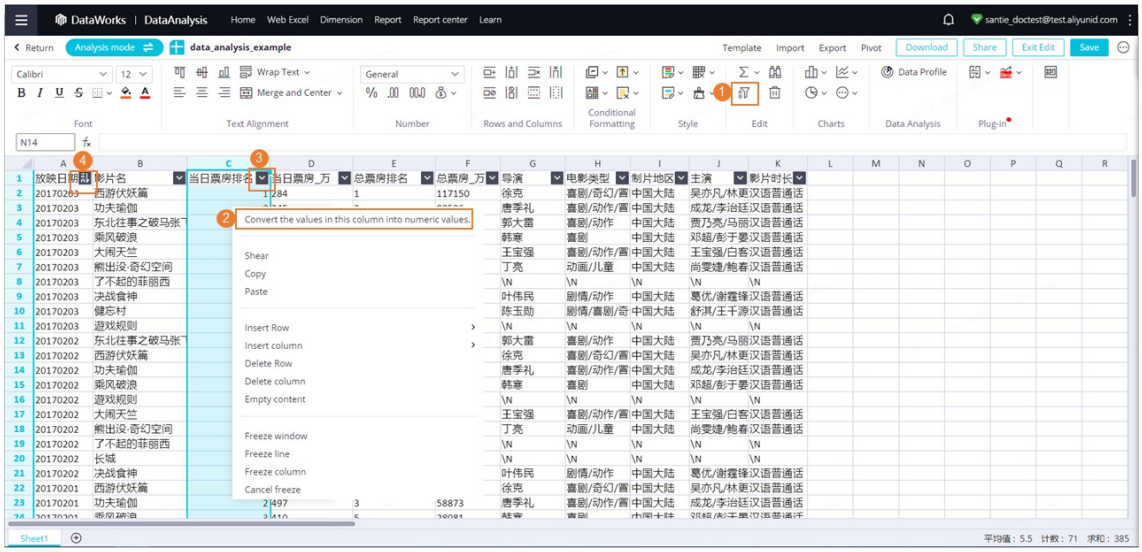
iv. Click Query mode in the upper-left corner to switch to **Analysis mode**. On the page that appears, click **Save** in the upper-right corner to save the result table for subsequent data analytics and sharing.



Data analytics and sharing

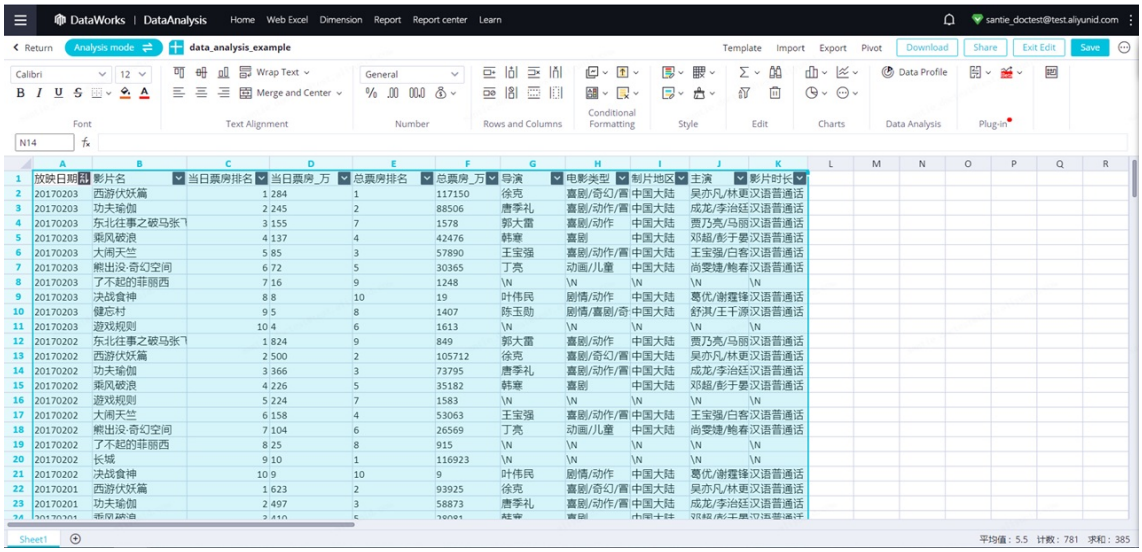
You can use the analysis mode to perform simple secondary analytics and measurements based on web Excel. You can also share the results with others.

1. Analyze data.
   - View the box office chart of movies each day.



     a. Select a cell in the result table and click the Sort and Filter icon.
     b. Right-click the **Box office ranking on the day** column and select **Convert the values in this column into numeric values**.
     c. Sort the values in the **Box office ranking on the day** column in ascending order and the values in the **Date** column in descending order.
   - Use the pivoting feature to collect statistics on the box office distribution of movies throughout the Spring Festival. Select the column on which you want to perform pivoting and click **Pivot** in the upper-right corner. In the dialog box that appears, leave the parameters at their default

values and click OK.



a. Drag the Movie title field from the Pivot Table Fields section to the Row section and the Box office gross on the day (tens of thousands) field to the Indicator section. Then, move the pointer over the field in the Indicator section, click the More icon, and select Edit. In the Property settings dialog box, select SUM for Summary method.

b. Select the columns and click the Pie Chart icon.

You can also perform more diversified analytics and measurements on data based on your business requirements. You must save all the statistics and analysis results.

2. Share data.After you save the pie chart, click **Share** in the upper-right corner to share the results with RAM users that belong to your Alibaba Cloud account. The RAM users can access the result table based on the shared link or access code. You can specify whether the RAM users have edit or read permissions on the table.