Alibaba Cloud

MaxCompute Tools and Downloads

Document Version: 20210713

C-J 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.

- You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloudauthorized 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 and/or its affiliates 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				
A 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.				
O 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.				
C) 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.				
⑦ 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 listinstanceid 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.Query editor	06
2.MaxCompute client	10
3.MaxCompute Studio	21
3.1. What is MaxCompute Studio?	21
3.2. Tools Installation and version history	25
3.2.1. Install IntelliJ IDEA	25
3.2.2. Install MaxCompute Studio	27
3.2.3. View and update the version of MaxCompute Studio	32
3.3. Configure MaxCompute Studio	33
3.4. Manage project connections	36
3.5. Manage data and resources	41
3.5.1. Manage project data	41
3.5.2. Import and export table data	44
3.5.3. Manage tables in a visualized manner	47
3.6. Develop SQL procedure	49
3.6.1. Create a MaxCompute script module	49
3.6.2. Develop and submit an SQL script	52
3.7. Developing Java	55
3.7.1. Overview	55
3.7.2. Create a MaxCompute Java module	59
3.7.3. Develop a UDF	59
3.7.4. Develop a MapReduce program	62
3.7.5. Develop Graph	66
3.7.6. Query unstructured data	69
3.7.7. Package, upload, and register a Java program	71
3.8. Develop a Python program	74

3.8.1. Configure a Python development environment	74
3.8.2. Develop a Python UDF	76
3.8.3. Develop a PyODPS script	79
3.9. Manage MaxCompute jobs	80
3.9.1. View jobs	80
3.9.2. Job instance details	82
3.10. Tool integration	85
3.10.1. Integrate with MaxCompute client	85
4.Downloads	89

1.Query editor

The MaxCompute console provides a query editor for you to execute SQL statements and analyze data. This topic describes how to use the query editor.

Overview

The query editor of MaxCompute is integrated into the DataAnalysis tool of DataWorks. You can use the query editor to edit MaxCompute SQL statements, query data, analyze data in workbooks, share data online, and download data. You can use the query editor to perform the following operations:

- Edit and execute SQL statements and authorization commands, such as access control list (ACL)based authorization commands.
- Enable the MaxCompute Query Acceleration (MCQA) feature. After this feature is enabled, the time that is used to run small- and medium-sized query jobs is reduced from minutes to seconds. For more information about MCQA, see MCQA overview.
- Use and test MaxCompute based on its public datasets. By default, the public datasets are open to users.
- Analyze, download, and share query results by using Web Excel online.

Scenarios

The query editor can be used in the following scenarios:

- If you use and test MaxCompute for the first time, you can use the query editor to experience the core features of MaxCompute based on the public datasets.
- If you are a data analyst, you can use the query editor to query data. Then, you can switch to the analysis mode to analyze the query results by using Web Excel online. To reduce the frequency at which data is transferred and ensure data security, you can also download the query results to your computer for analysis.
- If you are a security administrator, you can find a required project and click **Project permission management** in the Actions column to manage role permissions. However, this feature is in trial use. In many scenarios, you must run commands to manage permissions. The query editor allows you to run most of the security commands without the need to perform additional configurations.

Open the query editor

1. Log on to the MaxCompute console, select a region in the top navigation bar, and then click Query editing to open the query editor.

≡	C-) Alibab	a Cloud	China (Hong Kong) 👻			
Max	Compute					
	Project manage	ment	Resource management	Ð	Query	editing
c	Create project	Enter t	he MaxCompute project name to se	arch	Q	
N	/laxCompute Projec	t Name	MaxCompute Region		Billing	g Method

2. In the **Select Datasource** dialog box, select **MaxCompute** for **Type** and select an existing project from the **Workspace** drop-down list.

Type :		6	()	0	\$
	\sim	MysQL	¢\$	8	ð
	MaxCompute	Mysql	PostgreSQL	DRDS	SQLServer
	ORACLE	0	0	19	\$
	Oracle	AnalyticDB for MySQL2.0	AnalyticDB for PostgreSQL	HOLO	EMR Hive
	Spark				
Workspace :					×

? Note If you select a project that is associated with a **workspace** in standard mode and submit a job by using the query editor, the job is submitted in a development project. The name of such a project is ended with _dev.

3. Click OK. The SQL Query page appears.

Query mode and analysis mode

The query editor can work in query mode or analysis mode. This feature is integrated into the DataAnalysis tool of DataWorks. For more information, see DataAnalysis. The components displayed in query mode are different from those displayed in analysis mode.

• Query mode

- MaxCompute	0 🗉 🗳	22		(/) Go to	o DataStu		
			e, city from	<pre>MAXCOMPUTE_PUBLIC_DATA.dwd_product_phoneno_basic_info_2020;</pre>			
'orkspace :					3		
Used Table Public data set		T.U. D	D 1				
	2 Saved query Hist	ory Table Details L B	og Results				
ery table			city	×			
maxcompute_public_data.dwd_produ		上海	上海				
maxcompute_public_data.dwd_produ	uct_movie 3 三五互联	上海	上海				
🗰 maxcompute_public_data.dwd_produ	- 4 世纪互联 uct_phone 5 中兴视通	上海 上海	上海上海				
maxcompute_public_data.dwd_prou		上海	上海				
maxcompute_public_data.ods_enterp		上海	上海				
maxcompute_public_data.ods_enterp		上海	上海				
maxcompute_public_data.ods_enterp	prise_shar 9 中期移动	上海	上海				
maxcompute_public_data.ods_enterp	prise_shar 10 中麦通信 11 丰信通信	上海 上海	上海				
maxcompute_public_data.ods_enterg		上海	上海				
maxcompute_public_data.ods_enterg		上海	上海				
maxcompute_public_data.ods_enterp	(D.71.4 hol)	上海	上海				
maxcompute_public_data.ods_produ	ct_movie_15 全民优打	上海	上海				
maxcompute_public_data.tpc_orders	- 16 分享通信 1t //	上海	上海				
	- ¹⁷ 华云互联 18 华翔联信	上海 上海	上海 上海				
	19 国美极信	上海	上海				
	20 天音移动	上海	上海				
	21 小米移动	上海	上海				
	22 平安通信	上海	上海				
	23 无限互联	上海	上海				
	24 星美生活 25 普泰移动	上海 上海	上海上海				
	25 普泰移动 26 朗玛信息	上/母 上海	上海				
	27 民生通讯	上海	上海				
	28 海航通信	上海	上海				
	Result1 ×			- Analyze in S	Spreads		
No.		Descriptio	n				
1)		The name	ofthe	workspace.			
_				•			
2		Tables in datasets.	the wor	kspace, tables that have been used, and public			
3		The editor, which allows you to execute SQL statements and run authorization commands, such as ACL-based authorization commands. On the SQL Query page, you can edit a script without the					

•	Analysis mode

4

After you execute SQL statements in query mode, you can click **Analyze** in **Spreadsheet** in the lower-right corner of the page to enter the analysis mode.

and results.

need to create a file. You can also save the script as a file.

The tabs that allow you to view the saved query files, history, logs,

eadsheet	: / 🔟 Unsa	ved File													Template	Import	Export	Pivot	Download	S
~	12 🗸 🗖	<u>00</u> 10 7	🖶 Wrap Te	xt v	Gene	eral	~	프 네	⊒× Iñi	~ ~	<u>↑</u> ~		Σ	- 60	ம்∽⊯	~ (Data Profi	e 🗄	v ≝v	5 2
<u>u</u> s 🗉	× 💁 🔺 🔳	diddle Align	🛱 Merge a	nd Center 🗸	%	0,00 00.	<u>&</u> ~		III III	* ~	. ~	, × ≜ ×	1	Û	⊕ × ⊙	~				
		-						_		Condit		-			_					
Font fx		le	xt Alignment			Number		Rows and	Columns	Forma	tting	Style	E	dit	Charts)ata Analysis		Plug-in	
A	В		с	D	E	F	G	Н	1	J	К	L	M	N	0	P	Q	R	S	
p	province	 city 			E	F	G	п		J	r.	L	IVI	IN	0	٢	ų	n	5	
r 63网络通信	上海	上海																		
五互联	上海	上海																		
北互联	上海	上海																		
兴视通	上海	上海																		
国电信	上海	上海																		
回移动	上海	上海																		
国联通	上海	上海		a)																
朝移动	上海	上海		a value																
っ麦通信	上海	上海																		
F信通信	上海	上海																		
R语通信	上海	上海																		
京东通信	上海	上海																		
尤酷移动	上海	上海																		
全民优打	上海	上海																		
う享通信	上海	上海																		
#云互联	上海	上海																		
料联信	上海	上海																		
国美极信	上海	上海		·																
lt1 🕒																				
																				_
-																				
lo.						Desc	ripti	on												
\mathbf{D}						The 1	tool	bar th	nat yo	ou us	e to	edit a	nd n	nana	age d	ata i	n a w	orkb	ook.	

2.MaxCompute client

You can use the MaxCompute client to access MaxCompute projects and run commands. This topic describes how to install, configure, and run the MaxCompute client and provides other related instructions.

Prerequisites

Before you use the MaxCompute client, make sure that the following conditions are met:

- Java 8 or a later version is installed on the device on which you want to install the MaxCompute client.
- A MaxCompute project is created.

For more information about how to create a MaxCompute project, see Create a MaxCompute project.

• A RAM user as which you want to use the MaxCompute client is added to the DataWorks workspace to which the MaxCompute project belongs.

For more information about how to add members to a workspace, see Add a workspace member and configure roles.

Context

After the MaxCompute client is installed and configured, you can start the MaxCompute client by using the script file under the installation path or by running commands in the command-line interface (CLI) of the system. The MaxCompute client allows you to perform the following operations:

- Obtain the information about the current logon user. For more information, see Obtain information of the current logon user.
- Exit the MaxCompute client. For more information, see Exit the MaxCompute client.
- Obtain help information about commands. For more information, see Obtain help information about commands.
- If you want to start the MaxCompute client by running commands in the CLI, you must specify parameters in the commands. For more information, see Specify startup parameters.

Limits

MaxCompute client V0.28.0 and later support Java Development Kit (JDK) 1.9. The MaxCompute client of a version earlier than V0.28.0 supports only JDK 1.8. You can view the client version in the CLI after you start the MaxCompute client. For more information about how to start the MaxCompute client, see Run the MaxCompute client.

Usage notes

- The output format of the MaxCompute client may not be forward compatible. The command syntax and execution rules of the client vary based on the client version. We recommend that you do not rely on the output format of the client to parse data. For more information about client versions, see aliyun-odps-console.
- Two consecutive minus signs (--) are used to comment out a command line on the MaxCompute client.

Install and configure the MaxCompute client

Note MaxCompute client V0.27.0 and later support the MaxCompute V2.0 data type edition. We recommend that you use the MaxCompute V2.0 data type edition. For more information about the supported data types, see MaxCompute V2.0 data type edition.

To install and configure the MaxCompute client, perform the following steps:

- 1. Download the MaxCompute client installation package.
- 2. Decompress the downloaded package to obtain the bin, conf, lib, and plugins folders.
- 3. Open the conf folder and configure the odps_config.ini file.

The following example shows the content in the odps_config.ini file.

project_name= access_id= access_key= end_point= log_view_host= https_check= # confirm threshold for query input size(unit: GB) data_size_confirm= # this url is for odpscmd update update_url= # download sql results by instance tunnel use_instance_tunnel= # the max records when download sql results by instance tunnel instance_tunnel_max_record= **#IMPORTANT:** # If leaving tunnel_endpoint untouched, console will try to automatically get one from odps service, w hich might charge networking fees in some cases. # Please refer to Configure endpoints # tunnel_endpoint= # use set.<key>= # e.g. set.odps.sql.select.output.format=

In the odps_config.ini file, lines that start with a number sign (#) are comments. The following table describes the parameters in the file.

Parameter	Required	Description	Example
-----------	----------	-------------	---------

Parameter	Required	Description	Example
project_name	Yes	The name of the MaxCompute project that you want to access. If you create a workspace in standard mode, pay attention to the differences of the project names between the production environment and development environment when you specify this parameter. The names of the projects in the development environment end with _dev. For more information, see Basic mode and standard mode. You can log on to the MaxCompute console and view the MaxCompute project names on the Project Management tab.	doc_test_dev
access_id	Yes	The AccessKey ID of your Alibaba Cloud account or a RAM user within the Alibaba Cloud account. You can obtain the AccessKey ID from the Security Management page.	None
access_key	Yes	The AccessKey secret that corresponds to the AccessKey ID. You can obtain the AccessKey secret from the Security Management page.	None

Parameter	Required	Description	Example
end_point	Yes	The endpoint of MaxCompute. You must set this parameter based on the region and network connection method you selected when you create the MaxCompute project. For more information about the endpoints that correspond to each region and network, see Endpoints.	http://service.cn- hangzhou.maxcompute.aliyun.c om/api
		Notice If the endpoint that you configured is invalid, an error occurs when you access MaxCompute.	
log_view_hos		The Logview Uniform Resource Locator (URL). You can view the detailed runtime information of a job by using this URL. This information helps you locate job errors. Set the value to http://logview.odps.aliyun.com	http://logview.odps.aliyun.com
t		Note We recommend that you set this parameter. If you do not set this parameter, you cannot locate the cause of job errors.	
https_check	No	 Specifies whether to enable HTTPS access. If HTTPS access is enabled, requests to access MaxCompute projects are encrypted. Valid values: True: HTTPS access is used. False: HTTP access is used. Default value: False. 	True
data_size_co nfirm	No	The maximum size of input data, in GB. The value range is unlimited. We recommend that you set this parameter to 100.	100

Parameter	Required	Description	Example
update_url	No	A reserved parameter.	None
use_instance_ tunnel	No	 Specifies whether to use InstanceT unnel to download the results of SQL statements. Valid values: True: InstanceT unnel is used to download the results of SQL statements. False: InstanceT unnel is not used to download the results of SQL statements. Default value: False. 	True
instance_tunn el_max_recor d	No	The maximum number of SQL execution results that can be returned by the client. You must specify this parameter if the use_instance_tunnel parameter is set to True. Maximum value: 10000.	10000
tunnel_endpo int	No	The public endpoint of MaxCompute T unnel. If you do not specify this parameter, traffic is automatically routed to the T unnel endpoint that corresponds to the network where MaxCompute resides. If you specify this parameter, traffic is routed to the specified endpoint and automatic routing is not performed. For more information about the T unnel endpoints that correspond to each region and network, see Endpoints.	http://dt.cn- hangzhou.maxcompute.aliyun.c om
set. <key></key>	No	The properties of the MaxCompute project. For more information about the properties of MaxCompute projects, see Properties.	set.odps.sql.decimal.odps2=tru e

Run the MaxCompute client

You can start the MaxCompute client by using one of the following methods:

• Method 1: In the bin folder under the installation path of the MaxCompute client, double-click the o

dpscmd.bat file to start the MaxCompute client. If information shown in the following figure is returned, the MaxCompute project is connected.



• Method 2: In the CLI of the system, go to the bin folder under the installation path of the MaxCompute client and run the odpscmd command to start the MaxCompute client. If information shown in the following figure is returned, the MaxCompute project is connected.

D:\)\bin>odpscmd
(=X,=,//,=),=X,=,////////////////////////
Aliyun ODPS Command Line Tool
Version 0.35.1-public
@Copyright 2020 Alibaba Cloud Computing Co., Ltd. All rights reserved.
Connecting to http://servicealiyun.com/api, project:dev
Project timezone: Asia/Shanghai
Connected!
odps@dev>

You can enter a command at the cursor shown in the preceding figure and press Enter to run the command. The following figure shows the sample command and returned results.

MaxComput e

odps@dev>
>select abs(-1);
ID = :
Log view:
http://logview.odps.aliyun.com/logview/?h=http://service.cn-hangzhou.maxcompute.a
liyun.co :W1b2JhMn
A3WGdCR2 /JTdGF0ZW1
1bnQiO1t [jZSI6WyJh
Y3M6b2Rw 1171NLND4; of L-L-C
jV3eHJzYSJdfV0sI1Z1cnNpb24i0iIxIn0=
Job Queueing
Summary: resource cost: cpu 0.00 Core * Min, memory 0.00 GB * Min
resource cost. cpu 0.00 core ≁ min, memory 0.00 dD ≁ min inputs:
outputs:
Job run time: 0.000
Job run mode: service job 2.0
Job run engine: execution engine
bubble: 0
instance count: 1
run time: 0.000
instance time:
min: 0.000, max: 0.000, avg: 0.000
input records:
output records:
AdhocSink1: 1 (min: 1, max: 1, avg: 1)
metrics_output_count:
AdhocSink1: 1 (min: 1, max: 1, avg: 1)
Project1: 1 (min: 1, max: 1, avg: 1)
Values1: 1 (min: 1, max: 1, avg: 1)
metrics_inner_time_ms:
AdhocSink1: 49 (min: 49, max: 49, avg: 49) MaxInstance: 0
GlobalInit: 41 (min: 41, max: 41, avg: 41) MaxInstance: 0
Project1: 0 (min: 0, max: 0, avg: 0) MaxInstance: 0
Values1: 0 (min: 0, max: 0, avg: 0) MaxInstance: 0
++
c0
++
++
1 records (at most 10000 supported) fetched by instance tunnel.

Note For more information about the command syntax supported by the MaxCompute client, see Common commands or SQL commands and functions.

Obtain information of the current logon user

You can run the following command in the CLI to obtain the information of the current logon user:

odps@project_name>whoami;

The following figure shows the returned result.



- Name: the account of the current logon user.
- Source IP: the IP address of the device where the MaxCompute client is located.

- End_Point: the endpoint of MaxCompute.
- Project: the name of the MaxCompute project.

Exit the MaxCompute client

You can run the following command in the CLI to exit the MaxCompute client:

odps@project_name>quit; -- The preceding command is equivalent to the following command.

odps@project_name>q;

Obtain help information about commands

You can obtain help information about the commands of the MaxCompute client by using one of the following methods:

- Method 1: View the help information about commands on the MaxCompute client.
 - View help information about all commands.

```
odps@project_name>help;
-- The preceding command is equivalent to the following command.
odps@project_name>h;
```

• Specify a keyword to view the help information about the related commands.

Example: Obtain the help information about the commands related to table operations.

```
odps@project_name>help table;
-- Returned results:
Usage: alter table <tablename> merge smallfiles
Usage: export table <tablename>
Usage: show tables [in <project_name>] [like '<prefix>']
list|ls tables [-p,-project <project_name>]
Usage: describe|desc [<project_name>.]<tablename> [partition(<spec>)]
Usage: read [<project_name>.]<table_name> [(<col_name>[,..])] [PARTITION (<partition_spec>)] [line_n
um]
```

 Method 2: In the CLI of the system, go to the bin folder under the installation path of the MaxCompute client, and run the following command to view the help information about all commands. If you start the MaxCompute client by running commands in the CLI, you can specify a series of parameters in the commands. For more information about these parameters, see Specify startup parameters.

..\odpscmd\bin>odpscmd -h

Specify startup parameters

In the CLI of the system, you can specify a series of parameters to run a command. The following code shows the usage of these parameters.

Usage: odpscmd [OPT where options include	-	
help	(-h)for help	
config= <config_file< td=""><td>> specify another config file</td><td></td></config_file<>	> specify another config file	
project= <prj_name< td=""><td>> use project</td><td></td></prj_name<>	> use project	
endpoint= <http: l<="" td=""><td>ost:port> set endpoint</td><td></td></http:>	ost:port> set endpoint	
-k <n></n>	will skip begining queries and start from	n specified position
-r <n></n>	set retry times	
-f <"file_path;">	execute command in file	
-e <"command;[com	mand;]"> execute command, inclu	ıde sql command

The following table describes the startup parameters.

Parameter	Description	Sample command		
help or -h	Obtains the help information about all commands of the MaxCompute client.	odpscmdhelp		
config	Specifies the directory where the configuration file odps_config.ini is saved. The default directorv is odpscmd_public/conf/odps_config.ini	odpscmd config=D:/odpscmd/conf/odps_config.in i		
project	Specifies the name of the MaxCompute project that you want to access.	odpscmdproject=doc_test		
endpoint	Specifies the endpoint of MaxCompute. For more information about endpoints, see Configure endpoints.	odpscmdendpoint=http://service.cn- shanghai.maxcompute.aliyun.com/api		
-k	Executes the statement from the specified location. If n is set to a value that is less than or equal to 0, the execution starts from the first statement. Multiple statements are separated by semicolons (;).	Run the following command to ignore the first two statements and start to execute from the third statement: odpscmd -k 3 -e "drop table table name:create table table name (dummy string):insert overwrite table table name select count(*) from table_name;"		
-r	Specifies the number of retries allowed when the job fails to run.	odpscmd -r 2 -e "select * from sale_detail;select * from table_test;"		

Parameter	Description	Sample command	
	f Specifies the file to read.	 Prepare a script file named script.txt. In this example, the file is stored in drive D and contains the following data: 	
		drop table if exists test_table_mj; create table test_table_mj (id string, name string); drop table test_table_mj;	
-f		 In the CLI, go to the bin folder under the installation path of the MaxCompute client and run the following command: 	
		\odpscmd\bin>odpscmd -f D:/script.txt;	
-е	Specifies the command that you want to run.	odpscmd -e "select * from sale_detail;"	

The dynamic return value of an odpscmd -e command may be called by a shell script that is run in a shell window or the Command Prompt in Windows. A shell variable may obtain the return value and use it in subsequent jobs. In this scenario, only field values are required. Other information, such as runtime information and headers, must not be returned. To help vou call shell commands more easily, you can run the set odps.sql.select.output.format={"needHeader":false,"fieldDelim":""}; command to disable header display.

For example, a table named noheader has one column and three rows of data. The field values are 1, 2, and 3. After you run the following command to redirect the standard output of the query result to the destination handle, the output contains only field values.

```
odpscmd -e "set odps.sql.select.output.format={"needHeader":false,"fieldDelim":""};select * from noheader
;" >/temp/test.txt
-- Returned results:
1
2
3
```

Version updates

The following table describes the latest version updates of the MaxCompute client. For more information, click the URL of a specific version.

Version	Change type	Description

Version	Change type	Description			
New feature	New feature	An external project can be created to connect to Data Lake Formation (DLF). This implements the Alibaba Cloud LakeHouse feature.			
Fixed issue		The issue that the nanosecond part of data of the TIMESTAMP type is incorrectly processed is fixed.			

3.MaxCompute Studio

3.1. What is MaxCompute Studio?

MaxCompute IDEA

MaxCompute Studio is a big data integrated development environment (IDE) tool that is provided by Alibaba Cloud MaxCompute. It is a development plug-in based on the widely adopted integrated development environment Intellij IDEA. You can install this plug-in on the clients of developers to facilitate data development. This topic describes the functional interfaces and common application scenarios of MaxCompute Studio.

Basic user interface

MaxCompute Studio is a plug-in on the Intellij IDEA platform. It shares basic development interfaces with Intellij IDEA. For more information, see Guided Tour around the User Interface.

MaxCompute Studio provides the following features based on Intellij IDEA:

• SQL Editor: provides features such as SQL syntax highlighting, code completion, real-time error prompts, local compilation, and job submission.

Compiler View displays the prompts and error messages of statements that are locally compiled, and locates the code in the editor.

• **Project Explorer**: connects to a MaxCompute project and allows you to browse table schemas, user defined functions (UDFs), and resource files in the project. For more information, see Manage project connections.

Table Details View displays the details and sample data of tables, views, and other resources.

- Job Explorer: allows you to browse and search the MaxCompute job history.
 - Job Details View: displays the running details about a job, including the execution plan and details about each execution task. You can view all information about a job in Logview.
 - Job Output View: displays output information for jobs that are running.
 - Job Result View: displays output results for SELECT jobs.
- MaxCompute Console: integrated with the MaxCompute client. You can enter and run MaxCompute client commands in the MaxCompute console.

Connect to a MaxCompute project

Before you use the features of MaxCompute Studio, you must create a project connection. For more information, see Manage project connections. After you create a project connection, you can view related data structure and resource information in **Project Explorer**. MaxCompute Studio creates a local metadata backup task for each project. This increases the access frequency to MaxCompute metadata and reduces latency.

You must specify a MaxCompute project connection so that you can use MaxCompute Studio to modify SQL scripts, submit jobs, view job information, open the MaxCompute console, and implement other features. For more information about MaxCompute projects, see Project. For more information about how to manage projects in MaxCompute Studio, see Manage project connections.

Manage data

You can use **Project Explorer** in MaxCompute Studio to browse table schemas, UDFs, and resource files in projects. You can use the tree view hierarchy to list data tables, columns, partition key columns, virtual views, UDFs, function signatures, resource files, and types of resource files for all project connections. You can also use the tree view hierarchy to locate an object.

You can double-click a data table to open **Table Details View** to view the metadata, data schema, and sample data of the data table. If you do not have the permissions on the connected project, MaxCompute Studio displays an error message.

MaxCompute Studio is integrated with MaxCompute Tunnel to allow you to upload and download local data. For more information, see Import and export data.

Write SQL scripts

You can write MaxCompute SQL scripts in MaxCompute Studio.

- 1. Open MaxCompute Studio and choose File > New > Project or File > New > Module....
- 2. Create a MaxCompute Studio project or a MaxCompute Studio module.
- 3. Choose File > New > MaxCompute SQL Script or right-click the menu and choose New > MaxCompute SQL Script to create a MaxCompute SQL script file.

(?) Note When you create a MaxCompute SQL script, MaxCompute Studio prompts you to select a MaxCompute project to associate with. You can also use the **Project Selector** on the right of the toolbar of the SQL Editor to modify the associated project. The SQL Editor checks the metadata such as table schemas based on the project that is associated with the SQL script and reports errors in the script. When an SQL statement is submitted, the SQL Editor also sends the SQL statement to the associated project to execute. For more information, see Write SQL.

Intelligent error prompts in SQL code

The SQL Editor of MaxCompute Studio intelligently checks the code that you enter. It can display prompts for syntax errors, type matching errors, or warnings of SQL statements and mark code errors in real time.



MaxCompute Studio provides the code completion feature. It can display the project name, table, column, function, type, and code keywords based on context, and complement code based on your selections.



Compile and submit a job

• Compile a job

Click the



icon on the toolbar of the SQL Editor to locally compile an SQL script. If a syntax or semantic error occurs, the SQL Editor reports it.

7	
8	select clause in the front
9	<pre>select * from table_test;</pre>
10	
11	from clause in the front
12	<pre>from table_test table_alias select *;</pre>
13	
14	— table name with project prefix
MaxCom	pute Compiler
× Ξ	Information: Parsing
	Information: Type checking
1 🛧 😤	Information: Latency.compiler_parse_error : 44170
+ 🧐	Information: Build failed(2)
1 🛨	▼
	Error:(9, 15) table meta.table_test cannot be resolved
?	Error: (12, 6) table meta.table_test cannot be resolved

• Submit a job

Click the

icon on the toolbar of the SQL Editor to submit a compiled SQL script to the queue of the project that is specified by MaxCompute.

View job history

Open Job Explorer to view jobs that have been recently run in a specified project.

? Note This list o	only displays jobs that are submitted by the user ID of the current connection.
MaxCompute Job Explorer	- *- *-
Project: sql_optimizer	▼ Days: 2 ▼ 「∫
	Q (40/47883) Table 🔻
Instanceld Status Ov	wner StartTime EndTime
20170 SUCCESS OD	DPS 2017– 2017–
20170 SUCCESS OD	DPS 2017– 2017–
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 SUCCESS OD	DPS 2017 2017
20170 FAILED OD	DPS 2017– 2017–

Double-click a job to view the job details.

B Job ID: 201702170653179gu5qrzsk2 ×					
\odot	Refresh				
	SQL_0_0job_0				
	Task Name: sql_optimizer_201702170653179gu5qrzsk2_SQL_0_0_0_job_0				
M1	Task I/O Records Status Progress StartTime EndTime M1 1/1 CONTRACT 100.0 2017-02 2017-02				
TableScan_REL3692					
Data Source: sql_optimizer.dual TS: alias: dual					
	о М1 ×				
	Instance I/O Recor Status FinishedP StartTime EndTime IP & Path LogId				
ADHOC_SINK_3693	M1#0_0 1/1 TERMI 100.0 2017 2017 10.10 d01U				
FS: output: None					
2017-02-17 14:53:17 14:52017-02-17 14:53:35					
可视化 概要 (JSON) 概要 (文本) 结果 SQL					

If you have the Logview URL of a job, choose **MaxCompute > Open Logview** from the menu to go to the details page of the job.

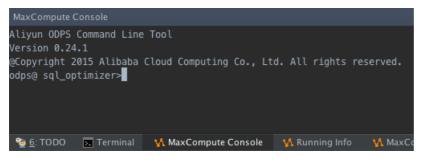
Develop MapReduce and UDF programs

MaxCompute Studio allows you to develop MapReduce and Java UDF programs. For more information, see Develop MapReduce and Develop UDFs.

Connect to the MaxCompute client

MaxCompute Studio is integrated with the latest version of the MaxCompute client. You can also specify the path of a locally installed MaxCompute client on the configuration option page of MaxCompute Studio. For more information, see Configure MaxCompute Studio.

On the **Project Explorer** tool window, right-click a project and select **Open in Console**. The **MaxCompute Console** window appears.



What to do next

After you are familiar with the functional interfaces and common application scenarios of MaxCompute Studio, you can proceed to learn how to install MaxCompute Studio. For more information, see Tools Installation and version history.

3.2. Tools Installation and version history

3.2.1. Install IntelliJ IDEA

This topic describes how to install the basic platform Intellij IDEA of MaxCompute Studio. MaxCompute Studio is a plug-in that runs on Intellij IDEA.

Prerequisites

Intellij IDEA can be installed on the Windows, MacOS, and Linux operating systems. For more information about requirements on hardware and operating systems, see Requirements for Intellij IDEA.

Procedure

1. Click Intellij IDEA to download Intellij IDEA based on your operating system, such as Windows, MacOS, or Linux. The Windows operating system is used in the following example, and supports Intellij IDEA 14.1.4 or later.

🕐 Note Intellij IDEA includes PyCharm Edition, Ultimate Edition, and free Community Edition.

2. After Intellij IDEA is downloaded, double-click the installer to start the setup wizard, and click Next.



3. Specify the destination folder and click Next.

	munity Edition Setup	
	Choose Install Location	
10	Choose the folder in which to inst Edition.	all IntelliJ IDEA Community
	liJ IDEA Community Edition in the following and select another folder. Click Next to con	
Destination Folder		
	Brains\IntelliJ IDEA Community Edition 2018	8.1.4 Browse
		8.1.4 Browse
rogram Files\Jeti	.9 MB	B. 1. 4 Browse
rogram Files\Jet	.9 MB	B.1.4 Browse

4. Select your operating system and click **Next**.

IntelliJ IDEA Communit	y Edition Set	up		
1	Installation Configure yo	-	Community Edition i	installation
Create Desktop Shortcut	54-bit launcher			
Create Associations	.kt			
		< Back	Next >	Cancel

5. Click Inst all.

IntelliJ IDEA Com	munity Edition Setup		
	Choose Start Menu Fol Choose a Start Menu fold Edition shortcuts.		A Community
	u folder in which you would like to e to create a new folder.	create the program's	shortcuts. You
Accessories Administrative Tools Cisco Eclipse Games Git Java Maintenance Microsoft Office 20 MySQL-Front Notepad++ Oxygen XML Editor	3		
	< E	Back Install	Cancel

6. After Intellij IDEA is installed, click Finish.

IntelliJ IDEA Community Edition Setup			
IJ	Completing IntelliJ IDEA Community Edition Setup		
	IntelliJ IDEA Community Edition has been installed on your computer.		
	Click Finish to close Setup.		
	Run IntelliJ IDEA Community Edition		
	< Back Finis	h Cancel	

What's next

Install MaxCompute Studio. For more information, see Install MaxCompute Studio.

3.2.2. Install MaxCompute Studio

This topic describes how to install MaxCompute Studio in Intellij IDEA.

Prerequisites

The environment in which you want to install MaxCompute Studio must meet the following requirements:

- The client runs a Windows, macOS, or Linux operating system.
- Intellij IDEA 2018.2.4 or later is installed on the client. The Ultimate, PyCharm, and free Community

editions are supported.

- Java Runtime Environment (JRE) 1.8 is installed. The latest version of Intellij IDEA is bundled with JRE 1.8. If you are using Intellij IDEA of the latest version, you do not need to separately install JRE 1.8
- Java Development Kit (JDK) 1.8 is installed. JDK 1.8 is required only if you want to develop and debug user-defined functions (UDFs) in Java.

? Note MaxCompute Studio version 0.28.0 and later support JDK 1.9. The earlier versions of MaxCompute Studio support only JDK 1.8.

Context

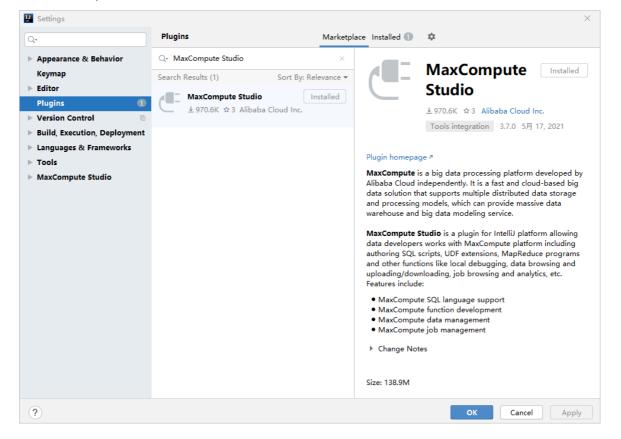
MaxCompute Studio is a plug-in of Intellij IDEA. You can use one of the following methods to install MaxCompute Studio:

- Install MaxCompute Studio online by using the plug-in library. We recommend that you use this method.
- Install MaxCompute Studio by using local files.

(Recommended) Install MaxCompute Studio online

MaxCompute Studio is available to all Internet users. You can install it by using the official Intellij plug-in library.

- 1. In the top navigation bar, choose File > Settings.
- 2. In the left-side navigation pane of the **Settings** dialog box, click **Plugins**.
- 3. In the **Plugins** dialog box, search for MaxCompute Studio.
- 4. Find MaxCompute Studio and click Inst all.

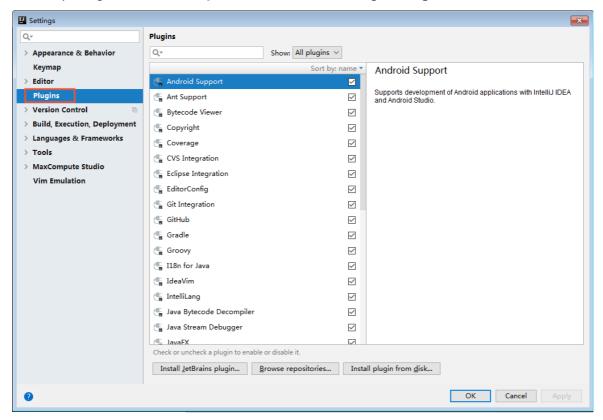


5. Restart Intellij IDEA to complete the installation.

Install MaxCompute Studio by using local files

You can download MaxCompute Studio to your computer and install it.

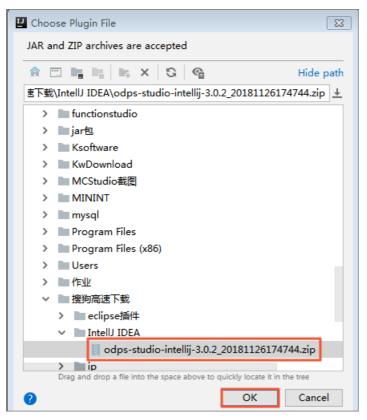
- 1. Go to the MaxCompute Studio page to download the plug-in package.
- 2. In the top navigation bar of Intellij IDEA, choose File > Settings > Plugins.



3. In the Plugins dialog box, click Inst all plugin from disk.

		×
Plugins		
Q- Show: All plug	gins \vee	
So	ort by: name 🔻	Android Support
🗲 Android Support		
📲 Ant Support		Supports development of Android applications with IntelliJ IDEA and Android Studio.
📲 Bytecode Viewer		
🗲 Copyright		
📲 Coverage		
📲 CVS Integration		
📲 Eclipse Integration		
📲 EditorConfig		
📑 Git Integration		
📑 GitHub		
📲 Gradle		
📲 Groovy		
📲 I18n for Java	\checkmark	
📲 IdeaVim	\checkmark	
🕞 IntelliLang	\checkmark	
🕞 Java Bytecode Decompiler	\checkmark	
📲 Java Stream Debugger	\checkmark	
IavaEX		
Check or uncheck a plugin to enable or disable it.		
Install JetBrains plugin Browse repositori	es Insta	ll plugin from <u>d</u> isk
		OK Carrel
		OK Cancel Apply

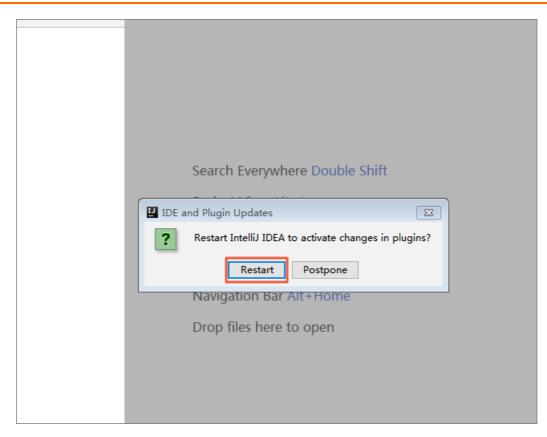
4. In the Choose Plugin File dialog box, select the downloaded plug-in file and click OK.



5. Go to the Plugins dialog box and click **OK**.

U Settings			8			
Q_	Plugins		Reset			
> Appearance & Behav	or Q	Show: All plugins \lor				
Keymap	Coverage	Sort by: name 🔻	MaxCompute Studio			
> Editor	CVS Integration		C Restart IntelliJ IDEA			
Plugins	C Faliana Internation	_	New version will be available after restart			
> Version Control			MaxCompute is a big data processing platform developed by			
> Build, Execution, Dep			Alibaba Cloud independently. It is a fast and cloud-based big			
> Languages & Framew			data solution that supports multiple distributed data storage and processing models, which can provide massive data			
> Tools	GitHub		warehouse and big data modeling service.			
> MaxCompute Studio	Gradle		MaxCompute Studio is a plugin for IntelliJ platform allowing			
Vim Emulation	Groovy		data developers works with MaxCompute platform including authoring SQL scripts, UDF extensions, MapReduce			
	📑 I18n for Java		programs and other functions like local debugging, data browsing and uploading/downloading, job browsing and			
	📑 IdeaVim	\checkmark	analytics, etc. Features include:			
	📲 IntelliLang	\checkmark	MaxCompute SQL language support			
	📑 Java Bytecode Decomp	iler 🗹	MaxCompute data management			
	📑 Java Stream Debugger	\checkmark	MaxCompute data management MaxCompute job management			
	📲 JavaFX	\checkmark	Change Notes			
	📲 JUnit		Changes in Version 3.0.0:			
	📲 Kotlin		New Feature: support write authorization statement, show user			
	📲 Maven Integration					
	🚝 MaxCompute Studio		privileges and diagnosis auth exception message. • use web view to draw sql pot graph and performance			
	Check or uncheck a plugin to	Check or uncheck a plugin to enable or disable it.				
	Install JetBrains plugin	Browse repositories Insta	Install plugin from <u>d</u> isk			
0			OK Cancel Apply			

6. After the installation is complete, click **Restart** in the message that prompts you to restart Intellij IDEA.



What's next

After MaxCompute Studio is installed, you must configure connections to MaxCompute projects to use the data and resources. For more information, see Manage project connections.

3.2.3. View and update the version of

MaxCompute Studio

This topic describes how to view and update the version of MaxCompute Studio.

View the version of MaxCompute Studio

- 1. In the top navigation bar, choose File > Settings.
- 2. On the Settings page, click MaxCompute Studio in the left-side navigation pane.
- 3. On the **MaxCompute Studio** page, view the current version and latest version of MaxCompute Studio in the **Updates Checking** section.

Update the version of MaxCompute Studio

• MaxCompute Studio automatically detects new versions. If a new version is detected, you are notified to update the version.

After you receive an update message, you can perform the following operations:

- Install: In the update message, click **Install**. The new version is automatically downloaded and installed. After the installation is complete, restart Intellij IDEA.
- Settings: In the update message, click **Settings**. You can configure whether to automatically detect new versions.

- If the automatic update feature is disabled, you can perform the following steps to check the latest version of MaxCompute Studio and install it:
 - i. In the top navigation bar, choose File > Settings.
 - ii. On the Settings page, click MaxCompute Studio in the left-side navigation pane.
 - iii. On the MaxCompute Studio page, view both the current version and latest version of MaxCompute Studio in the Updates Checking section.
 - iv. Click Check new versions to check the latest available version and install it.

Onte You can also select Automatically checks for new versions to enable automatic detection of new versions.

3.3. Configure MaxCompute Studio

This topic describes how to configure MaxCompute Studio.

After MaxCompute Studio is installed, choose File > Settings... in the main menu bar of Intellij IDEA to go to the Settings dialog box to configure MaxCompute Studio.

MaxCompute Studio configuration items

In the **Settings** dialog box, click **MaxCompute Studio** in the left-side navigation pane. The following configuration items appear on the right side of the dialog box:

- General Settings
 - Local meta store location: the storage path of the local metadatabase. This parameter specifies the local path used to store the metadata of a MaxCompute project. By default, MaxCompute Studio stores the metadata in the *.odps.studio\meta* directory for users.
 - Table preview rows: the maximum number of rows that can be previewed in a table.
 - Local directory for jobs: the local path used to store the data of MaxCompute Studio jobs. By default, MaxCompute Studio stores the data in the *.odps.studio**job* directory for users.
 - Python path to resolve UDF: the local installation path of Python.
- Updates Checking
 - Automatically checks for new version: specifies whether to enable MaxCompute Studio to automatically check for a new version. By default, this parameter is selected to support automatic updates.
 - **Check new versions**: You can click this button to manually check for a new version. After you click this button, the **Install latest version** button appears if a new version is available. You can click the **Install latest version** button to install the new version. After the new version is installed, you must restart Intellij IDEA.

SDK & Console configuration items

In the **Settings** dialog box, choose **MaxCompute Studio** > **SDK & Console** in the left-side navigation pane.

The **Installed Location** parameter appears on the right side of the dialog box. This parameter specifies the local installation path of the MaxCompute client. MaxCompute Studio automatically detects the version of the MaxCompute client installed in the path. If no client is found, an error message is returned.

Note In MaxCompute Studio versions later than V2.6.1, you do not need to specify the path because the latest MaxCompute client is installed by default. If you need to use the MaxCompute client of a specific version, you can specify the path.

MaxCompute SQL configuration items

In the **Settings** dialog box, choose **MaxCompute Studio** > **MaxCompute SQL** in the left-side navigation pane.

The following configuration items appear on the right side of the dialog box:

Color and Annotator

You can select **Enable syntax coloring** to enable the syntax highlighting feature.

- Code Completion
 - You can select **Enable code completion** to enable the automatic code completion feature.
 - If you select **Invoke code completion when you enter a space**, the automatic code completion feature is disabled when you enter a space.
- Code Formatting

You can select **Enable code formatting** to enable the code formatting feature.

- Script Submit
 - If you select **Show job detail when script submitted**, job details are displayed when you submit a script.
 - If you select **Record sql history when script submitted**, the execution history of SQL statements is recorded when you submit a script.
 - If you select **Show sql cost confirm dialog when script submitted**, a message to confirm the execution cost of SQL statements appears when you submit a script.
 - If you select **Pin sql execution result tab by default**, the SQL Execution Result tab is locked by default.
 - Job Name:
 - If you select use script file name as default, the name of the script file is used as the job name by default.
 - If you select input job name when first submit, you must enter a job name the first time you submit a job.
 - If you select input job name for every submit, you must enter a job name each time you submit a job.
- Compile | Submit

This section provides the default settings of global compiler parameters. You can separately configure the parameters for each file in the toolbar of the SQL compiler:

- Compiler mode
 - Statement Mode: In this mode, the compiler compiles and submits each statement of an SQL file as a unit.
 - Script Mode: In this mode, the compiler compiles and submits an entire SQL file as a unit. This
 mode enables the compiler and optimizer to optimize the execution plan and improve the
 overall execution efficiency.

- Data type edition
 - Legacy TypeSystem: The original data type edition of MaxCompute is used.
 - MaxCompute TypeSystem: The new data type edition introduced from MaxCompute V2.0 is used.
 - Hive Compatible TypeSystem: The Hive-compatible data type edition introduced from MaxCompute V2.0 is used.
- Compiler
 - Default : The default compiler of MaxCompute Studio is used.
 - UseShortQueryAcceleration: A compiler that can accelerate queries is used.
 - FallbackWhenFailed: A compiler that can run jobs again if query acceleration fails is used.

MaxCompute Job configuration items

In the **Settings** dialog box, choose **MaxCompute Studio > MaxCompute Job** in the left-side navigation pane.

The following configuration items appear on the right side of the dialog box:

- Maximum jobs shown in job explorer table list: the maximum number of jobs that can be displayed on the Job Explorer page. Default value: 1000.
- Reset filter when switch MaxCompute project: If you select this parameter, the job configurations are reset when you switch between MaxCompute projects.
- Cache history instance in local database: You can select this parameter to cache historical instance information in the local data warehouse.

Accounts configuration items

In the **Settings** dialog box, choose **MaxCompute Studio** > **Accounts** in the left-side navigation pane.

You can perform the following operations to manage the accounts that are used to access MaxCompute. For more information about accounts, see User authentication.

- Add an account
 - i. In the right-side navigation pane, choose + > Aliyun Account By AccessKey.
 - ii. In the Add MaxCompute Account dialog box, configure the following parameters:
 - Account Name: the display name of an account on MaxCompute Studio. Both Alibaba Cloud accounts and RAM users can be added. For example, if you add the Alibaba Cloud account us er test@test.alivun.com , set this parameter to user_test . If you add the RAM user user ram @xxxxx.onaliyun.com that belongs to the Alibaba Cloud account user_test@test.alivun.com , set this parameter to user_test.alivun.com .
 - Using properties file: specifies whether to read an AccessKey ID and an AccessKey secret from the configuration file. If you select this parameter, you must make sure that the *conf/od ps_config.ini* file is uploaded. For more information, see User authentication.

Onte MaxCompute Studio uses the configured accounts to access MaxCompute projects and run jobs.

- Using properties: specifies whether to manually enter an AccessKey ID and an AccessKey secret. You must select this parameter or Using properties file.
 - Access Id: Enter the AccessKey ID of the Alibaba Cloud account or RAM user that you want to add.
 - Access Key: Enter the AccessKey secret of the Alibaba Cloud account or RAM user that you want to add.
- iii. Click OK. The added account is displayed in the account list in the Settings dialog box.
- Remove an account

You can perform the following steps to remove an account from MaxCompute Studio. This operation does not affect the real Alibaba Cloud account or RAM user.

- i. Select the account that you want to remove from the account list.
- ii. In the right-side navigation pane, click -.
- iii. In the message that appears, click **OK**.
- Modify the information of an account
 - i. Select the account whose information you want to modify in the account list.
 - ii. In the right-side navigation pane, click the 🗾 icon.
 - iii. In the Edit MaxCompute Account dialog box, edit the following account information:
 - Access Id: Enter a new AccessKey ID.
 - Access Key: Enter a new AccessKey secret.

iv. Click OK.

3.4. Manage project connections

You can view information about your MaxCompute project on MaxCompute Studio only after you connect MaxCompute Studio to your project. The project information includes tables, views, user-defined functions (UDFs), and resources. This topic describes how to create and modify MaxCompute project connections.

Step 1: Create a MaxCompute Studio project

- 1. Start Intellij IDEA. In the main menu bar, choose File > New > Project.
- 2. In the left-side navigation pane of the **New Project** dialog box, click **MaxCompute Studio**. Then, click **Next**.

Java	Additional Libraries and Frameworks:
■ Java FX ▲ Android	
 IntelliJ Platform Plugin Maven Gradle 	Nothing to show
G Groovy	
K Kotlin	
Empty Project	
🔨 MaxCompute Java	
<mark>M</mark> MaxCompute Studio	

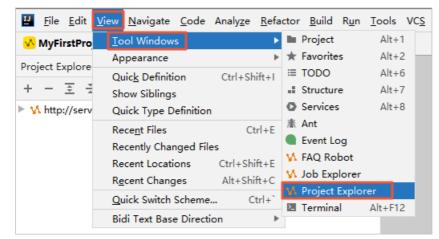
3. Specify **Project name** and click **Finish**.

Project		×
Project name: MyFire	stProject	
Project location: C:\Use	ars\\IdeaProjects\MyFirstProject	
 More Settings 		
Module name:	MyFirstProject	
Content root:	C:\Users \IdeaProjects\MyFirstProject	
Module file location:	C:\Users \\IdeaProjects\MyFirstProject	
Project format:	.idea (directory based)	•
	Previous Finish Cancel H	Help

Note If a project is open, a dialog box appears and asks whether you want to open the new project in a new window or the current window. In this case, click **This Window**.

Step 2: Connect to a MaxCompute project

1. In the main menu bar, choose View > Tool Windows > Project Explorer.



2. In the upper-left corner, click the plus sign (+) and select Add project from accessId/Key.

🛂 <u>F</u> ile <u>E</u> di	it <u>V</u> iew I	<u>N</u> avigate	<u>C</u> ode	Analy <u>z</u> e	<u>R</u> efactor	<u>B</u> uild	R <u>u</u> n	<u>T</u> o
<mark> My</mark> FirstP	roject							
Project Explo	rer						\$	-
+ - 主	ž 🗊	q 🛯	S 🕷	🕘 (i	?			
🕤 Add proj	ject from d	lataworks						
🖉 Add proj	ject from a	ccessId/K	ey					
🕤 Add proj	ject from c	lataworks		a 2	?			

3. In the Add MaxCompute project dialog box, configure the parameters on the Connection tab.

💾 Add MaxCom	pute project		×
Connection Sett	ing		
Properties File:			1
AK Account:		•	+
* Access Id:			
* Access Key:			
* Project Name:			
* End Point:	The second se		•
			~
?	ок		Cancel

? Note

- In the lower-left corner, click the ricon to view online documentation.
- If a timeout error occurs, click the **Setting** tab and modify the following parameters:
 - sync one table timeout(s): the timeout period for table synchronization.
 Default value: 5. Unit: seconds.
 - **sync one function timeout(s)**: the timeout period for function synchronization. Default value: 30. Unit: seconds.

• Use the configuration file to automatically configure the parameters.

Parameter	Description
Properties File	The directory where the configuration file <i>odps_config.ini</i> of the MaxCompute client (odpscmd) is saved. For more information, see Install and configure the odpscmd client. The configuration file is used to initialize the settings of the Access Id, Access Key, Project Name, and End Point parameters.

Parameter	Description		
	The Alibaba Cloud account that is used to connect to the MaxCompute project. You can click the plus sign (+) and select an existing account on the Accounts page.		
AK Account	Note If no accounts exist, you must click the plus sign (+) and select Aliyun Account By AccessKey to add an account on the Accounts page. You can manually add an account or use the configuration file to automatically add it. For more information about the Accounts page, see Accounts configuration items.		

Parameter	Description			
Access Id	The AccessKey ID that is used to connect to the MaxCompute project.			
Access Key	The AccessKey secret that is used to connect to the MaxCompute project.			
Project Name	The name of the MaxCompute project. Note This parameter specifies the name of your MaxCompute project rather than that of the DataWorks workspace to which the project corresponds. Log on to the MaxCompute console and view the name of your project and the DataWorks workspace to which the project corresponds on the Project management tab.			
End Point	The endpoint of the MaxCompute project.			

• Manually configure the parameters.

4. Click OK. The Project Explorer page shows information about the MaxCompute project, including the tables, views, UDFs, and resources.

Step 3: Modify the connection to a MaxCompute project

1. On the **Project Explorer** page, right-click the MaxCompute project whose connection you want to modify and select **Modify project properties**.

Project Explor	rer	☆ -
+ - \Xi	ː ʃ đ 🖻 🕏 🕷 🕘 🛈 ?	
🔻 🚺 http:// 👘	the second s	
Ĩ .	 + Add project Remove project Modify project properties S User Privilege 	

2. In the **Modify MaxCompute project** dialog box, modify the configurations of the MaxCompute project.

What's next

After the MaxCompute project is connected, you can manage and view the data and resources of the project. For more information, see Manage data and resources.

3.5. Manage data and resources

3.5.1. Manage project data

This topic describes how to use MaxCompute Studio to view tables, views, functions, and resources in projects.

Prerequisites

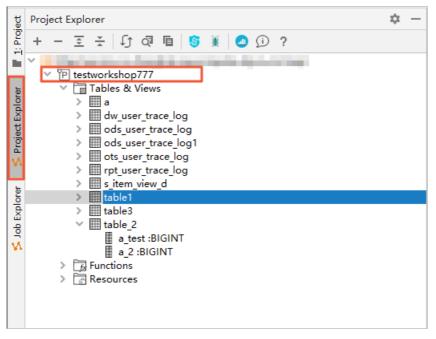
A MaxCompute project is connected. For more information, see Manage project connections.

Context

In the **Project Explorer** tool window, you can view tables, views, functions, and resources in MaxCompute projects that are connected.

Browse and update project data

1. Go to the MaxCompute Studio page. In the left-side navigation pane, click **Project Explorer** to view your connected MaxCompute project.



2. Click the drop-down arrow on the left side of the **Tables & Views**, **Functions**, or **Resources** node to view all tables, views, functions, or resources in the project.

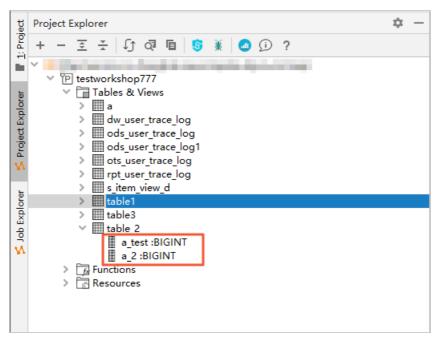
? Note To perform this operation, you must have the **DESCTABLE** permission on the MaxCompute project.

3. In the toolbar, click the 🗊 icon to update the local metadata.

? Note MaxCompute Studio downloads the metadata of MaxCompute projects to a local device. When the metadata of a MaxCompute project is updated, you must trigger a refresh to reload the changed metadata to the local device.

View details about tables or views

1. In the **Tables & Views** node tree, click a table name to expand the node. You can then view the columns and column types of the table.



2. Right-click the table name and select **Show table detail** to view the table details. The table details include the owner, size, column information, and schema of the table.

+ - 프 곳 [] 경 🖻 🧕 🖲 🗊 ?	Table information				Table schema				
V M http://service.cn-shanghai.maxcompute.aliyun.com/api	key		value		Field	Туре	Label Nullable	Default	Comment
() Tetherwitching 77 minimized in prevailing including a set of the set	Project: Name: Owner: Create Time: LastDUTime: LastDUTime: Lifecycle: TableID: IsArchived: Size: PhysicalSize: FileNum:	testworkshop777 table 2 2020-03-28 15:41:3 2020-03-28 15:51:4 2020-03-28 15:51:4 2020-03-28 15:51:4 2020-03-28 15:41:30 0 0 0	5 4		a_test a_2	BIGINT BIGINT	Y		बिधर बिधर 1
	A V								
	Preview rows: 200	(Max 5000) Data Pr	eview Import data	E	xport data				
	1 1	a_te:	at		2		a_2		
	2 3				4				
Run: 🛛 🖓 SQLtest.osql ×									\$ -
▶ 日志 结果									

3. Right-click the table name and select Find usages to query the scripts in which the table is used.

View function details

1. In the **Functions** node tree, click the **UserDefined** node to view the functions that you have created.

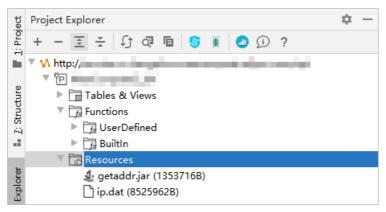
Project Explorer 🌣 -	🔹 🗑 odps_config.ini 🗶 🔄 SQLtest.osql 🗶 👔 2020032809182261gq5h5mim 👋 🛔 testworkshop777 🔧 table_2 X 🖉 adv_user_trace_log X 🔍 🚳 GetAddr.class X
+ — 포 포 印 @ 盾 🧐 🕷 🖉 🛈 ?	Decompiled .class file, bytecode version: 52.0 (Java 8)
 ✓ http://service.on-shanghal.mascompute.aligun.com/api ✓ U testworkshop?77 ✓ Table & Wiews > ✓ Muser trace.log > ✓ ods.user_trace.log > ✓ ods.user_trace.log > ✓ ods.user_trace.log > ✓ trace.log > ✓ Table3 ✓ Table3 ✓ ✓ Sourcions ✓ ✓ Sourcions ✓ Ø UderOnfind ✓ Window > ✓ Window > ✓ Apgregation > ✓ Apgregation > ✓ Other 	<pre>1</pre>

? Note The Built in node under the Functions node tree classifies the built-in functions of the system. You can double-click a function to display the usage notes of the function. For more information about the built-in functions, see Built in functions.

2. Double-click a function to display the code of the function.

View resource details

Expand the Resources node tree to view the created resources.



3.5.2. Import and export table data

MaxCompute Studio allows you to import local files in the CSV or TSV format to MaxCompute tables. It also allows you to export data from MaxCompute tables to local files. MaxCompute Studio uses MaxCompute Tunnel to import and export data.

Prerequisites

- MaxCompute Tunnel is configured for your project in MaxCompute Studio. For more information, see Install and configure the MaxCompute client.
- The account has the permissions to manage tables in your MaxCompute project.

Import data

1. In the **Project Explorer** tool window, click the drop-down arrow on the left side of **Tables & Views** under your MaxCompute project. In the table list, right-click the table to which you want to import data and select **Import data into table**.

🖳 Local Settings [C:\Users\Administrator\Local Settings] - IntelliJ IDI	EA		
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>N</u> avigate <u>C</u> ode Analy <u>z</u> e <u>R</u> efactor <u>B</u> uild R <u>u</u> n	<u>T</u> ools	VC <u>S</u>	<u>W</u> indow
Local Settings			
Project Explorer	₩- №		
+ - ∫ ≣ § ¥ Ξ ∻ 🙆 Ø ?			
 Tables & Views bank_data emp_test_new ods_user_info_d result_tat table_1 wc_in1 Functions Resources Show table detail Open table editor Generate select statement Generate select statement Generate table data Drop table from server Import data into table Export data from table 			

2. In the Importing data to table_name dialog box, configure the parameters.

Importing data t	o result_tabl	X
Input File:	A DESCRIPTION OF	
File charset:	UTF-8	~
Column Separator	: Comma(',')	~
Record Limit:	100	Size(MB) Limit: 10
🗹 Include Column	Header	Error Record Limit: 0
0		OK Cancel

- Input File: the path of the local file to import.
- File charset: the encoding format of the input file. Supported encoding formats include UTF-8, UTF-16, UTF-16BE, UTF-16LE, ISO-8859-1, US-ASCII, and GBK. The default encoding format is UTF-8.
- Column Separator: the column delimiter. Supported column delimiters include Comma(','),
 Space(' '), and Tab('\t'). The default column delimiter is Comma(',').
- **Record Limit**: the maximum number of data rows that can be imported.
- Size(MB) Limit: the maximum size of data that can be imported. Unit: MB.
- Error Record Limit : the maximum number of data rows that contain errors.
- Include Column Header: specifies whether to import the column header.

- 3. Click **OK** to import the data.
- 4. The **Success** message appears. This indicates that the data is imported. You can then view the data in the table.

Export data

1. In the **Project Explorer** tool window, click the drop-down arrow on the left side of **Tables & Views** under your MaxCompute project. In the table list, right-click the table from which you want to export data and select **Export data from table**.

+ - 🗘 🖬 💲 🕷	≆ 粪 🕘 🗊 ?		
 ✓ M ✓ P ✓ Tables & Views 			
> bank_data			
> emp_test_new > ds_user_info			
> result_tabl	🔍 Find usages 🜀 User Privilege		
> III table_1 > III wc_in1 > III Functions	∫∫ Refresh table meta ॡ Show table detail		
> 🚡 Resources	Open table editor Generate select statement Generate DDL statement		
	Truncate table data Drop table from server 		
	Import data into table		
	Export data from table		
		•	

2. In the Exporting data from table_name dialog box, configure the parameters.

₽ Exporting data f	rom result_tabl		8
Output File:		Contraction of the local division of the loc	
File charset:	UTF-8		\sim
Column Separator	: Comma(',')		~
Record Limit:	100	Size(MB) Limit: 10	
Include Column	n Header	Error Record Limit: 0	
2		OK Car	icel

- Output File: the path of the local file that stores the output data.
- File charset: the encoding format of the output file. Supported encoding formats include UTF-8, UTF-16, UTF-16BE, UTF-16LE, ISO-8859-1, US-ASCII, and GBK. The default encoding format is UTF-8.
- Column Separator: the column delimiter. Supported column delimiters include Comma(','),
 Space(' '), and Tab('\t'). The default column delimiter is Comma(',').

- Record Limit : the maximum number of data rows that can be exported.
- Size(MB) Limit : the maximum size of data that can be exported. Unit : MB.
- Error Record Limit : the maximum number of data rows that contain errors.
- Include Column Header: specifies whether to export the column header.
- 3. The **Success** message appears. This indicates that the data is exported. You can view the data in the output file.

3.5.3. Manage tables in a visualized manner

Project Explorer in MaxCompute Studio provides a table schema editor that allows you to manage tables in a visualized manner. This topic describes how to use Project Explorer to create, modify, and delete tables in a visualized manner.

Create a table

- 1. In the **Project Explorer** tool window, right-click **Tables & Views** under your MaxCompute project, and select **Create new table**.
- 2. In the Create New Table/View dialog box, configure the parameters.

创建表/视图 -			
建表:		□ 仅当不存在时创建 □ 外部表 □ 视图	<u>ସ</u>
释:			
命周期:		日	
以 AliOrc 格式存	字储 🗌 以后表为模板创建		
姻子 如何	码 AS SELECT 附加属性		
name	type	not null comr	ment ·
ł	BIGINT		
►	E TABLE	.`` (
2 `id		• • • • • • • • • • • • • • • • • • •	
L 🕨 🖯 CREAT	E TABLE	· · · · (
CREAT	E TABLE	.`` (
CREAT	E TABLE	• • • • • • • • • • • • • • • • • • •	
CREAT	E TABLE	· · · · · (
CREAT	E TABLE	••••••••••••••••••••••••••••••••••••••	
CREAT	E TABLE	· · · · (
CREAT	E TABLE	• • • • • • • • • • • • • • • • • • •	
CREAT	E TABLE	• • • • • • • • • • • • • • • • • • • •	
CREAT	E TABLE	• • • • • • • • • • • • • • • • • • •	
CREAT	E TABLE	.`` (
CREAT	E TABLE	.`` (
CREAT	E TABLE	.`` (
CREAT	E TABLE BOOM BOOM BOOM BOOM BOOM BOOM BOOM BOO		

For more information about how to configure the parameters, see Table operations.

? Note

You cannot set flags for tables that are created in a visualized manner. The system uses the following flags by default:

- odps.sql.submit.mode=script
- odps.sql.type.system.odps2=true
- 3. Click Execute. If the SUCCESS message appears, the table is created.

Modify a table

- 1. In the **Project Explorer** tool window, click the drop-down arrow on the left side of **Tables & Views** under your MaxCompute project. In the table list, right-click the table that you want to modify and select **Open table editor**.
- 2. In the **Modify Table** dialog box, modify the table.

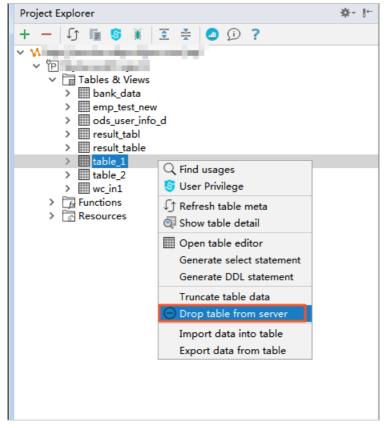
You can add columns to the table and change the table name, table comments, table lifecycle, column name, and column description. For more information, see Table operations.

修改表 - [MySecondProject2.table_2]					
表名:	A: table_2 □转为存档				
注释:	主释:				
生命周期:	-1	E	3		
数据列 分区列	索引附加	加属性			
name		type	not null	comment	+
col1		BIGINT			
col2		BIGINT			
col4		BIGINT			1
col3		BIGINT			+
col5		BIGINT			×
SQL DDL SCRIP	Т		um		
		able_2 ADD COLUMNS (θ
	5 BIGIN	();			_
操作: 提交到 Max	操作: 提交到 MaxCompute 执行 V 创建编辑过的对象 V 取消 执行 打开编辑器				

3. Click **Execute** to apply the modifications.

Delete a table

1. In the **Project Explorer** tool window, click the drop-down arrow on the left side of **Tables & Views** under your MaxCompute project. In the table list, right-click the table that you want to delete, and select **Drop table from server**.



2. In the **Confirmation Required** message, click **OK** to delete the table.

3.6. Develop SQL procedure

3.6.1. Create a MaxCompute script module

Before you use MaxCompute Studio to develop SQL programs, you must create a MaxCompute script module. This topic describes how to create a MaxCompute script module.

Context

Create a MaxCompute script module in one of the following scenarios:

- If no local script file exists, use Intellij IDEA to create a module.
- If a local script file exists, you do not need to create a module. Instead, open a module and use MaxCompute Studio to edit the script.

Create a module if no local script file exists

- 1. Start Intellij IDEA. In the top navigation bar, choose File > New > Project.
- 2. In the left-side navigation pane of the **New Project** dialog box, click **MaxCompute Studio**. Then, click **Next**.

Java	Additional Libraries and Frameworks:	
➡ Java FX ➤ Android ● IntelliJ Platform Plugin		
m Maven ☞ Gradle	Nothing to show	
G Groovy		
🗲 Kotlin		
Empty Project		
🔨 MaxCompute Java		
🔥 MaxCompute Studio		

3. Specify the **Project name** parameter and click **Finish**.

🖳 New Project		×
Project name: MyFire	stProject	
Project location: C:\Use	<pre>visit line a line</pre>	
 More Settings 		
Module name:	MyFirstProject	
Content root:	C:\Users \IdeaProjects\MyFirstProject	
Module file location:	C:\Users \IdeaProjects\MyFirstProject	
Project format:	.idea (directory based)	-
	Previous Finish Cancel He	p

? Note

If a project is opened, a dialog box appears and asks whether you want to open the new project in a new window or the current window. In this case, click **This Window**.

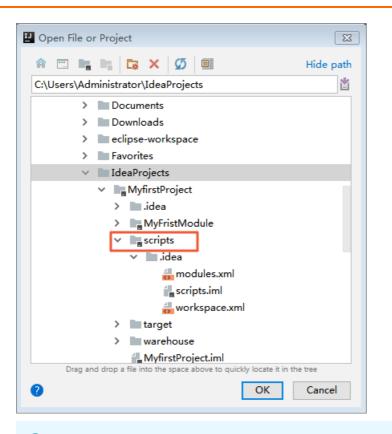
Create a module if a local script file exists

If a local script file exists, you do not need to create a module. You need only to add a MaxCompute connection configuration file to the existing module directory.

1. In the .*IdeaProjects\MaxCompute_Studio_Project_Name\scripts* folder, create a MaxCompute connection configuration file *odps_config.ini* and configure the following authentication information. This authentication information is used to connect to MaxCompute.

```
# The name of the MaxCompute project.
project_name=xxxxxxx
# The AccessKey ID of your Alibaba Cloud account.
access_id=xxxxxxxxx
# The AccessKey secret of your Alibaba Cloud account.
access_key=xxxxxxxxx
# The endpoint of the region where your MaxCompute service is deployed.
end_point=xxxxxxxx
```

2. Open Intellij IDEA. In the top menu bar, choose File > Open, and select the .*IdeaProjects\MaxCom pute_Studio_Project_Name\scripts* folder. Then, select the *odps_config.ini* file.



? Note MaxCompute Studio automatically searches for the *odps_config.inifile* in the folder and captures the metadata on the MaxCompute server based on the configuration information in the file. MaxCompute Studio then compiles all scripts in the folder.

3.6.2. Develop and submit an SQL script

This topic describes how to develop and submit an SQL script in MaxCompute Studio. The SQL script development includes writing and running an SQL script.

Prerequisites

- A MaxCompute project connection is created. For more information, see Manage project connections.
- A MaxCompute Script module is created. For more information, see Create a MaxCompute script module.

Write an SQL script

 In the Project tool window, click a project name, right-click scripts, and then choose New > MaxCompute SQL Script. .

₩ MyfirstProject [C:\L	Jsers\Administrator\IdeaProjects\My	/firstProject]	- In	telliJ IDEA	4		
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>N</u> av	rigate <u>C</u> ode Analy <u>z</u> e <u>R</u> efactor <u>F</u>	<u>B</u> uild R <u>u</u> n	Too	ols VC <u>S</u>	<u>W</u> indow	MaxCompute	<u>H</u> elp
<mark> MyfirstProject</mark>	scripts						
🗊 Project 🔻	⊕ ≑ ∳-⊩						
🗸 <mark>v My</mark> firstProject	C:\Users\Administrator\IdeaProjects	5					
> 🖿 .idea							
> MyFristMod	lule						
> scripts	New	>	С	Java Clas	ss		
target	₭ Cu <u>t</u>	Ctrl+X		Kotlin Fil	e/Class		
MyfirstProje	<u>∎</u> <u>С</u> ору	Ctrl+C		File	-11		
> External Librarie	Copy Path C	trl+Shift+C		Scratch F Package		trl+Alt+Shift+In	sert
> 👩 Scratches and C		Alt+Shift+C		FXML File			
	Paste	Ctrl+V			-info.java		
	Find <u>U</u> sages Find in Path C	Alt+F7	20 J	module-	info.java		
		trl+Shift+R	9	MaxCom	pute SQL		
	Analyze	>	5	MaxCom	pute 交互	式查询	
	<u>R</u> efactor	>	5	从 D2 加			
	Add to Favorites	>	H	HTML Fil			
	Show Image Thumbnails C	trl+Shift+T	J 41	Singleton	oplication		
	<u>R</u> eformat Code	Ctrl+Alt+L	G			Build Script	
	Optimi <u>z</u> e Imports	Ctrl+Alt+O	G	Gradle K	otlin DSL 9	Settings	
	<u>D</u> elete	Delete	4 <>	XSLT Sty	lesheet		
-	Build <u>M</u> odule 'MyfirstProject'			Edit File	Templates		
Event Log	R <u>e</u> build ' <default>' Ct</default>	rl+Shift+F9	4	GUI Form	n		
9 5	Show in Explorer		倡	Dialog			
Q	Open in terminal Set MaxCommute preject			Form Sn	· ·		
<u>5</u>	Set MaxCompute project		il.	Resource			
THE STATE	Local <u>H</u> istory	>	C	Plugin D	evKit		>

2. In the New MaxCompute SQL Script dialog box, specify the required parameters and click OK.

Prew MaxCompute SQL Script			
Script Name: MyFirstSQL			
MaxCompute Project: MySecondProject2 (service.odps.aliyun.com) V +			
OK Cancel			
	_		

- **Script Name**: the script name.
- **MaxCompute Project**: the MaxCompute project in which you want to write an SQL script. You can click the + button to create a MaxCompute project connection. For more information, see Manage project connections.
- 3. Write an SQL script in the compiler. For more information about the SQL syntax, see MaxCompute SQL overview.

? Note

- You can configure cross-project resource sharing. For example, you can bind a script to Project A and grant access permissions on table1 in Project B to the script.
- MaxCompute Studio allows you to set an SQL script compiler. For more information, see Overview.

Submit an SQL script

Before you submit an SQL script, you must configure relevant settings as required. MaxCompute Studio provides multiple compiler features. You can set these features in the toolbar at the top of the compiler. You can set the following compilation parameters:

- Compiler Mode:
 - **Statement Mode**: In this mode, the compiler separates the SQL statements in the script with semicolons (;), and commits the statements one by one to MaxCompute for execution.
 - Script Mode: In this latest development mode, the compiler submits a whole script to MaxCompute at a time. We recommend that you use this mode because this mode improves the overall execution efficiency.
- Type System: You can set this parameter to avoid compatibility issues related to SQL statements. Valid values:
 - Legacy TypeSystem: indicates the legacy type system of MaxCompute.
 - MaxCompute TypeSystem: indicates the new type system introduced by MaxCompute V2.0.
 - **Hive Compatible TypeSystem**: indicates the type system in the Hive compatibility mode introduced by MaxCompute V2.0.
- Execution Mode:
 - Default Version: indicates a stable version.
 - **MaxCompute Query Acceleration**: indicates the new MaxCompute query acceleration (MCQA) feature.
 - Rerun When Acceleration Fails: reruns the task when the query acceleration fails.
 - 1. In the toolbar or side bar, click the 🕟 icon to submit an SQL script to MaxCompute.

<u></u> MyFir	rstSQLosql ×
	単句模式 ~ MaxCompute类型系统 ~ 默认编译器 ~ ↓ ↓ ↓ ↓ 👔 🔤 🖻
1	name:MyFirstSQL
2	author: Administrator
3	create time:2018-12-06 17:36
4 🕨 🤅	CREATE TABLE MySecondProject2'. table_5' (
5	col1 BIGINT,
6	col2 BIGINT,
7	col3 BIGINT)
8 (
١. The second se	insert into table_2 values (1, ' a' , ' 201908908' , ' 20180907'); select * from ods_user_info_d where ds=' \$ (bizdate)' ;
	perecc ⊕ riom ogpTgpetTuroTg muele d2- å[orzgøre) '
	Specify script extra info
	bizdate : 20181206
	bizdate : 20181206
	? OK Cancel

? Note If a variable exists in the SQL script, such as \${bizdate} in the preceding figure, a dialog box appears, which prompts you to enter the variable value.

2. Before the SQL job runs, Intellij IDEA reminds you of the estimated cost of the SQL job. Confirm the estimated cost and click **OK** in the **Confirmation** message.

? Note

• In the toolbar, click the 😝 icon to update metadata that is used in the SQL script, such

as tables and user-defined functions (UDFs). This feature applies when MaxCompute Studio cannot detect the tables and functions that exist in MaxCompute.

- SQL scripts are compiled based on the metadata that you add in the **Project Explorer** window. If no errors are detected, the scripts are submitted to MaxCompute for execution.
- Operational logs are generated during the SQL script execution. If SQL scripts are executed in MaxCompute, the job details tab is opened. You can view the basic information about the execution.
- 3. On the **Result** tab, view the execution results.

If multiple statements are executed one by one, the execution result for each statement is displayed.

3.7. Developing Java

3.7.1. Overview

This topic describes how to develop a Java program by using MaxCompute Studio and describes related directories.

Procedure

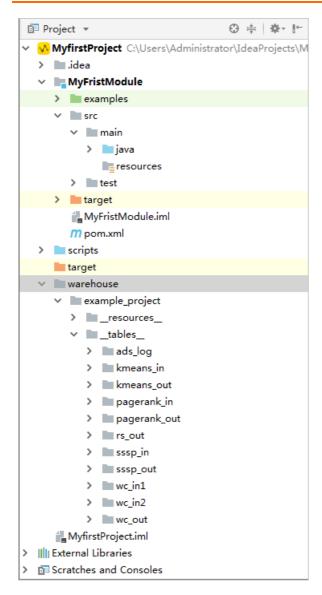
Follow these steps to develop a Java program in MaxCompute Studio:

- 1. Create a MaxCompute Java module.
- 2. Develop a Java program based on one of the following topics:
 - Develop a UDF
 - Develop a MapReduce program
 - Develop Graph
 - Query unstructured data
- 3. Package the Java program, upload the JAR package, and register a function.

Module directory

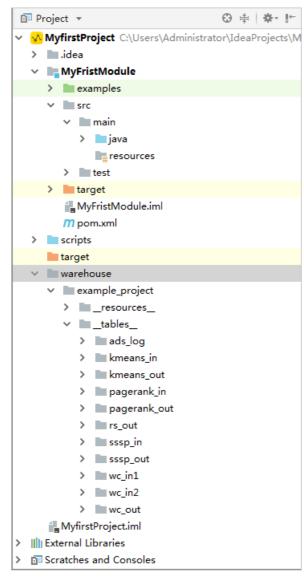
MaxCompute Studio automatically creates a module after you create a MaxCompute Java module. The module directory includes the following subdirectories:

- *examples*: stores sample code. You can reference the sample code to develop a script for unit testing.
- *src/main/java*: stores source code for Java program development.
- warehouse: stores schemas and data of tables and resources for MaxCompute projects.



warehouse directory

The *warehouse* directory stores the schemas and data of tables and resources for MaxCompute projects. These tables and resources are used to execute user-defined functions (UDFs) or MapReduce tasks.



- The *warehouse* directory contains three levels of subdirectories. Level-1 subdirectories are named by project name. Level-2 subdirectories include *_resources_* and *_tables_*. Each level-3 subdirectory under _tables_ is named by a table name and contains the *_schema_* and *data* files.
- In a *_schema_* file, configure the project name, table name, and names and data types of columns. Separate each pair of column name and data type with a colon (:). If a partitioned table is used, you must specify partition key columns. In the preceding figure, *wc_in1* is a non-partitioned table, and *wc_in2* is a partitioned table.
- In a *data* file, table data is stored in the CSV format based on the following rules:
 - In special characters, only commas (,), double quotation marks ("), and line breaks (\n or \r\n) are supported.
 - \circ A comma (,) is used as a column delimiter. \ln or \ln is used as a row delimiter.
 - If a value in a column contains special characters, enclose the value in a pair of double quotation marks ("). For example, write 3,No as "3,No".

- If a value in a column contains double quotation marks ("), use a pair of double quotation marks (") to escape each double quotation mark ("). For example, write a"b"c as "a""b""c".
- \circ \N indicates a NULL value. For a column of the STRING data type, write \N as """\N""".
- Characters in the file are encoded in UTF-8.

3.7.2. Create a MaxCompute Java module

MaxCompute Studio allows you to develop Java user-defined functions (UDFs), MapReduce programs, and Graph programs. Before the development, you must create a MaxCompute Java module. This topic describes how to create a MaxCompute Java module.

Prerequisites

A MaxCompute project is connected. For more information, see Manage project connections.

Procedure

- 1. Start Intellij IDEA. In the top navigation bar, choose File > New > Module.
- 2. In the left-side navigation pane of the **New Module** dialog box, click **MaxCompute Java**.
- 3. Configure the directory of the Module SDK file and click Next.
- 4. Specify Module name and click Finish.

Result

After the preceding steps are complete, MaxCompute Studio automatically creates a Maven module and completes the following actions:

- Introduces MaxCompute-related dependencies. For more information, see the POM file that contains the dependencies.
- Creates the *examples* folder to store sample code. For more information, see Module directory.
- Creates the *warehouse* folder to store the data required for local debugging. For more information, see warehouse directory.

What's next

After you create a MaxCompute Java module, you can develop Java programs. For more information, see the following topics:

- Develop a UDF
- Develop a MapReduce program
- Query unstructured data
- Develop Graph

3.7.3. Develop a UDF

This topic describes how to develop a user-defined functions (UDF) in MaxCompute Studio. This includes how to write and debug a UDF.

Prerequisites

Before you begin, make sure that you have completed the following operations:

Manage project connections

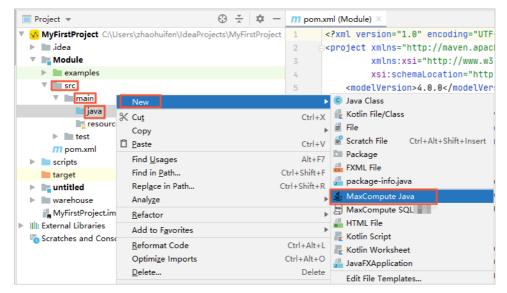
• Create a MaxCompute Java module

Context

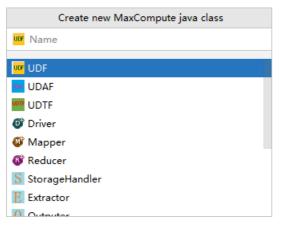
You can follow the instructions in this topic to develop a UDF, or choose **MaxCompute > Create UDF** to directly create a UDF.

Write a UDF

1. In the **Project** tool window, expand your MaxCompute Java module and choose **src** > **main** > **java**. Then, right-click java and choose **New** > **MaxCompute Java**.



2. Specify Name, select the UDF type, and press Enter.



- **Name**: the name of the MaxCompute Java class. If no package is created, enter packagename.classname. The system automatically creates a package.
- Select the UDF type.
- 3. After you create a MaxCompute java class, develop a Java program in the editor.

💀 MyfirstProject 👌 📭 MyFristModule 👌 🖿 👷 😁	iain) 🖿 java) 🖸 MyfirstUDF)	↓ VMyFirst
Project ▼ ③ ≑ ♣	r I ⁺⁻ 圏 MyFirstSQLosql × C MyFirstUDF.java ×	
 MyfirstProject C:\Users\Administrator\IdeaProject 	cts\M Project SDK is not defined	
> Lidea • MyFristModule	1 import com.aliyum.odps.udf.UDF;	
> 🖿 examples	3 public class MyFirstUDF extends UDF {	
✓ Imsrc	4 // TODO define parameters and return type, e.g. public String evaluate (String a,	String b)
✓ ■ main	5 Public String evaluate(String s) { return "hello world: " + s; }	
✓ iiijava	8 }	
> C HelloDriver		
V C MyFirstUDF	-	
m evaluate(String):String	-	
resources		

Perform a local run to debug the UDF

Perform a local run to test the UDF and check whether the results meet your expectations.

- 1. Right-click the compiled Java script and select **Run**.
- 2. In the Run/Debug Configurations dialog box, configure the runtime parameters.

– <u>□</u> □ <i>P</i> ∧ ▼ m ↓ ²	Name: MyFirstUDF	Share through VCS ② Allow parallel
MaxCompute Java	Main glass: MyFirstUDF	
MaxCompute Python MaxCompute SQL	⊻M options:	1 "Provided" scope L_UDF module)
	Enable capturing form snapshots	
	*MaxCompute project: local	✓ example_project ✓ +
	*MaxCompute table: wc_in1	
	*Table columns: col1 Download Record limit: 100	iesc
	 ▼ gefore launch: Build, Activate tool window + - ▲ <!--</td--><td></td>	

- MaxCompute project: the MaxCompute project in which the UDF runs. To perform a local run, select **local**.
- MaxCompute table: the name of the MaxCompute table in which the UDF runs.
- Table columns: the columns in the MaxCompute table in which the UDF runs.
- 3. Click OK to run the UDF.

? Note

- The system reads data from the specified table in *warehouse* as the input during the local run. You can view the log output in the console.
- The system downloads the table data from the specified MaxCompute project to the *w* arehouse directory. If the data is already downloaded, the system does not perform this step.

Perform unit testing to debug the UDF

<mark>🐼 MyfirstProject 👌 📭 MyFristModule 👌 🖿 examples 👌 🖿 com 👌 🖿 aliyun 👌 🖿 odps 👌 🖿 examples 👌 🖿 udf 👌 🖿 test 🔪 💞 UDFTest 🔪</mark> 😳 ≑ 🗰 📂 🖿 🔤 MyFirstSQL.osql × 🛛 😋 MyFirstUDF.java × 💽 UDFTest.java Project -MyfirstProject C:\Users\Administrator\IdeaProjects\M Project SDK is not defined > 📄 .idea package com. aliyun. odps. examples. udf. test; MyFristModule examples import ... > com.aliyun.odps.examples > 🛅 com.aliyun.odps.examples.graph 14 🍺 public class UDFTest { > com.alivun.odps.examples.mr @BeforeClass > 🖿 com.aliyun.odps.examples.mr.test public static void initWarehouse() { TestUtil. initWarehouse(); } > 🖿 com.aliyun.odps.examples.udf ✓ Immodel com.aliyun.odps.examples.udf.test @Test > 💣 UDAFTest 22 🕨 public void simpleInput() throws Exceptic 💑 Cut Ctrl+X ✓ C^tUDFTest BaseRunner runner = new UDFRunner (odps: 🕞 Copy Ctrl+C 📠 initWarehouse():void runner.feed(new Object[] { "one", "one" feed(new Object[] { "four", "four" Copy as Plain Text 🕷 🕤 inputFromTable():void Ctrl+Alt+Shift+C Copy Reference 🕷 🚡 resourceTest():void List<Object[]> out = runner.yield(); E Paste Ctrl+V 💕 🚡 simpleInput():void Paste from History... Ctrl+Shift+V Assert.assertEquals(3, out.size()); > CUDTFTest Paste Simple Ctrl+Alt+Shift+V Assert. assertEquals ("ss2s: one, one", Tes > 🖿 com.aliyun.odps.examples.unstructured Column Selection Mode Alt+Shift+Insert Assert. assertEquals ("ss2s: three, three" > 🛅 com.aliyun.odps.examples.unstructured.test Assert. assertEquals ("ss2s: four, four". 1 Find <u>U</u>sages Alt+F7 > 🖿 data.ambulance_csv } Refactor 5 > 🖿 data.speech_wav Foldina > 🗸 🖿 src 34 @Test Analy<u>z</u>e 🗸 📄 main 35 🕨 public void inputFromTable() throws Exc Search with Google iava BaseRunner runner = new UDFRunner (TestUti String project = "example_project"; ✓ C HelloDriver Go To String table = "wc_in2"; main(String[]):void Generate... Alt+Insert String[] partitions = new String[] { "p2 Of Recompile 'UDFTestjava' ✓ C MyFirstUDF Ctrl+Shift+F9 String[] columns = new String[] { "colc 40 m evaluate(String):String InputSource inputSource = new TableInputS 41 resources <u>
 Debug</u> 'simpleInput()' Object[] data: 42 ✓ ■ test while ((data = inputSource.getNextRow()) Run 'simpleInput()' with Coverage 43 🖿 java 44 runner. feed (data) ; Treate 'simpleInput()'... > 🖿 target } Show in Explorer HyFristModule.im 46 List<Object[]> out = runner.yield(); Open in terminal m pom.xml Assert.assertEquals(3, out.size()); 47 Local History scripts UDFTest > simpleInput() Compare with Clipboard IDE and Plugin Updates: IntelliJ IDEA is ready to update. (today 16:45)

Refer to the unit testing examples in the *examples* directory and write your test case.

What's next

After you write and debug a UDF, you must package, upload, and register UDF code. For more information, see Package, upload, and register a Java program.

3.7.4. Develop a MapReduce program

This topic describes how to use MaxCompute Studio to develop a MapReduce program. This includes how to write, debug, package, upload, and run a MapReduce program.

Prerequisites

Before you begin, make sure that you have completed the following operations:

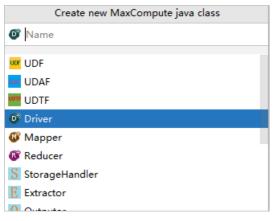
- Manage project connections
- Create a MaxCompute Java module

Write a MapReduce program

1. In the **Project** tool window, expand your MaxCompute Java module and choose **src** > **main** > **java**. Then, right-click java and choose **New** > **MaxCompute Java**.

Project 👻	0 ÷ 🕈 -	m pom.xml (Module) ×
🕨 🖿 .idea	ers\zhaohuifen\IdeaProjects\MyFirstProject	<pre>2</pre>
Module examples src		<pre>3 xmlns:xsi="http://www.w3 4 xsi:schemaLocation="http 5 <modelversion>4.0.0</modelversion></pre>
► test	New K Cut Copy Deaste	 Carl+X Ctrl+X Kotlin File/Class File Ctrl+V Scratch File Ctrl+Alt+Shift+Insert
scripts target untitled	Find <u>U</u> sages Find in <u>P</u> ath (C Repl <u>a</u> ce in Path	Alt+F7 Ctrl+Shift+F Ctrl+Shift+R MaxCompute Java
 MyFirstProject.im Illi External Libraries Scratches and Consc 	Analy <u>z</u> e <u>R</u> efactor Add to F <u>a</u> vorites	MaxCompute SQL
	<u>R</u> eformat Code Optimize Imports <u>D</u> elete	Ctrl+Alt+L Ctrl+Alt+O Delete Edit File Templates

2. Specify Name, select the Driver type, and then press Enter.



- **Name**: the name of the MaxCompute Java class. If no package is created, enter packagename.classname. The system automatically creates a package.
- Select the Driver, Mapper, or Reducer type.
- 3. After you create a MaxCompute java class, develop a Java program in the editor.

The Java template is automatically filled with the framework code. You need only to configure the input table, output table, and the Mapper and Reducer classes.

MaxComput e

C HelloDriver.java × HelloDriver 1 package mymr.myudf; 2 3 import ... 11 public class HelloDriver { 12 13 14 🕨 public static void main(String[] args) throws OdpsException { 15 JobConf job = new JobConf(); 16 17 // TODO: specify map output types 18 19 job.setMapOutputKeySchema(SchemaUtils.fromString(?)); 20 job.setMapOutputValueSchema(SchemaUtils.fromString(_?)); 21 22 // TODO: specify input and output tables 23 InputUtils.addTable(TableInfo.builder().tableName(_?).build(), job); 24 OutputUtils.addTable(TableInfo.builder().tableName(_?).build(), job); 25 26 // TODO: specify a mapper 27 job.setMapperClass(..?); 28 // TODO: specify a reducer 29 job.setReducerClass(_?); 30 RunningJob rj = JobClient.runJob(job); 31 rj.waitForCompletion(); 32 33 34 } 35 } 36

Perform a local run to debug a MapReduce program

Perform a local run to test your MapReduce program and check whether the results meet your expectations.

- 1. Right-click the compiled Java script and select **Run**.
- 2. In the **Run/Debug Configurations** dialog box, select the name of the MaxCompute project where the MapReduce program runs.

₽ Run/Debug Configurations		23				
+ - ≣ ≝ % ↑ »	Name: WordCountTest Single instance of	only				
✓ ■ JUnit	Configuration Code Coverage Logs	_				
UDFTest.simpleInput	Iest kind: Class V Fork mode: none V Repeat: Once V 1					
WordCountTest	<u>C</u> lass: com.aliyun.odps.examples.mr.test.WordCountTest					
 MaxCompute Java MaxCompute SQL 		_				
> 🌮 Defaults	⊻M options: -ea Program arguments:	2				
	Program arguments: Working directory: Working directory: %MODULE_WORKING_DIR%					
	Environment variables:					
	Use classpath of module:	~				
	JRE: Default (<no jre=""> - module not specified)</no>					
	Shorten command line: user-local default: none - java [options] classname [args]	~				
	Before launch: Build, Activate tool window					
+ − / ↑ + I the Build						
	□ Show this page ☑ Activate tool window					
	Warning: No JDK specified for module 'MyFristModule'					
0	OK Cancel Apply	/				

3. Click OK to run the UDF.

? Note

- The system reads data from the specified table in *warehouse* as the input during the local run. You can view the log output in the console.
- The system downloads the table data from the specified MaxCompute project to the *w* arehouse directory. If the data is already downloaded, the system does not perform this step.

Perform unit testing to debug a MapReduce program

Project -😳 🛊 | 🏘 - I⁺ 🕐 HelloDriver.java × 🖉 WordCount.java × 💕 WordCountTest.java × MyfirstProject C:\Users\Administrator\IdeaProjects\M Project SDK is not defined > idea package com. aliyun. odps. examples. mr. test; MyFristModule examples import ✓ ➡ com.aliyun.odps.examples public class WordCountTest extends MF 19 Þ > C TestUtil Ctrl+X > 🖿 com.aliyun.odps.examples.graph // 定义输入输出表的 schema private final static String INPUT Ctrl+C > 🛅 com.aliyun.odps.examples.mr private final static String OUTPUT Copy as Plain Text com.aliyun.odps.examples.mr.test Copy Reference Ctrl+Alt+Shift+C private JobConf job: > C WordCountTest fin Paste Ctrl+V 24 > 🖿 com.aliyun.odps.examples.udf Ctrl+Shift+V public WordCountTest() throws Except Paste from History ... > 🛅 com.aliyun.odps.examples.udf.test TestUtil. initWarehouse(); Paste Simple Ctrl+Alt+Shift+V > 🛅 com.aliyun.odps.examples.unstructured // 准备作业配置 Column Selection Mode Alt+Shift+Insert > 🛅 com.aliyun.odps.examples.unstructured.test job = new JobConf(); Find Usages Alt+F7 > data.ambulance csv Refactor > 🛅 data.speech_wav job. setMapperClass (WordCount. Tokeni Folding job.setCombinerClass(WordCount.SumC > = src job. setReducerClass (WordCount. SumRe Analyze > target MyFristModule.iml Search with Google 34 job. setMapOutputKeySchena (SchemaUti) m pom.xml Go To job. setMapOutputValueSchema(SchemaU scripts Alt+Insert Generate... > idea InputUtils. addIable(TableInfo. builde Recompile 'WordCountTest.java' Ctrl+Shift+F9 MyFirstSQL.osql OutputUtils. addTable(TableInfo. buile > Run 'WordC ountTest' Ctrl+Shift+F10 target } MordCountTest warehouse 40 Run 'WordCountTest' with Coverage example_project 41 @SuppressWarnings ("deprecation") Save 'WordCountTest' OTest > m_resources_ Show in Explore 43 1 public void testMap () throws IOExce ✓ ■ tables 44 MapUIContext mapContext = new MapUI 🗵 Open in terminal ✓ ■ ads_log 45 mapContext.setInputSchema(INPUT_SC. Local History 🚽 __schema__ nspContext.setOutputSchema(OUTPUT_ 46 😸 data 47 // 准备测试数据 File Encoding > 🖿 kmeans_in Record record = mapContext.createIng 48 > kmeans_out record set (new Text[] {new Text (st () Create Gist ...

Refer to the WordCount unit testing in the examples directory and write your test case.

Package and upload a MapReduce program

After you debug a MapReduce program, compress the MapReduce program into a JAR package and upload the package to the MaxCompute server as a resource. For more information, see Package, upload, and register.

Run a MapReduce program

Use the MaxCompute client to run a MapReduce program.

- 1. In the left-side navigation pane, click Project Explorer.
- 2. Right-click the name of your project and select **Open in Console**.
- 3. In the **Console** tool window, run the following command to start your MapReduce program.

For more commands, see JAR command.

jar-libjars wordcount.jar -classpath D:\odps\clt\wordcount.jar com.aliyun.odps.examples.mr.WordCou nt wc_in wc_out;

3.7.5. Develop Graph

This topic describes how to use MaxCompute Studio to develop Graph programs. This includes how to write, debug, package, upload, and run Graph programs.

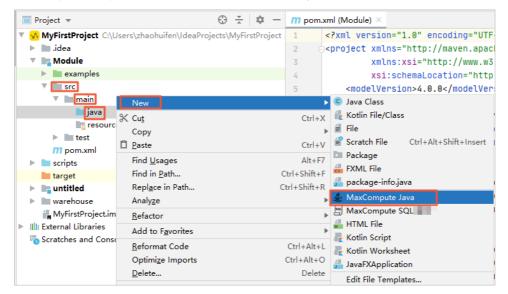
Prerequisites

> Document Version: 20210713

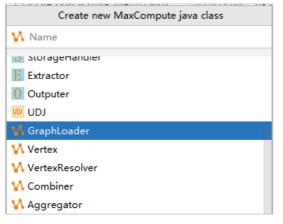
- A MaxCompute project is connected. For more information, see Manage project connections.
- A MaxCompute Java module is created. For more information, see Create a MaxCompute Java module.

Write a Graph program

1. In the **Project** tool window, expand your MaxCompute Java module and choose **src** > **main** > **java**. Then, right-click java and choose **New** > **MaxCompute Java**.



2. Specify Name, select GraphLoader or Vertex, and then press Enter.



- **Name**: the name of the MaxCompute Java class that you want to create. If no package is created, enter packagename.classname. The system automatically creates a package.
- Select GraphLoader or Vertex as the class type.
- 3. After the class is created, develop a Graph Java program in the editor.

Debug the Graph program

Locally run the Graph program and check whether the results meet your expectation.

- 1. Right-click the compiled Java script and select Run.
- 2. In the Run/Debug Configurations dialog box, configure the runtime parameters.

Run/Debug Configurations					×
+ - @ @ ≁ ▲ ▼ ₩ ↓ª	Name: graph		Share S	through VCS 💿 🔲 Allow	parallel r <u>u</u> n
✓ MaxCompute Java V ^X MyFirstUDF	Main <u>c</u> lass:	graph			
₩ graph	<u>V</u> M options:				$+ e^{\pi}$
 MaxCompute Python MaxCompute SQL 	Program a <u>rg</u> uments:				+ 2*
> 🖉 Templates	Working directory:	C:\Users\liyang\MaxCompute Studio			— ~
	Environment variables:				Ξ
	Use classpath of module:	test_UDF			~
		Include dependencies with "Provided" scope			
	JRE:	Default (1.8 (3) - SDK of 'test_UDF' module)			— ~
	Shorten command line:	user-local default: none - java [options] classna	me [args]		~
	<u>Enable capturing form</u>	snapshots			
	*MaxCompute project: loc	al	~	example_project ~	+
	Download Record limit:	100			
	 <u>Before launch: Build, Acti</u> <u>-</u> <u>-</u> 				
0			[OK Cancel	Apply

- MaxCompute project : the MaxCompute project in which you want to run the Graph program.
- **Download Record limit**: the maximum number of records that can be downloaded. Default value: 100.
- 3. Click OK to run the UDF.

? Note

- The system reads data from the specified table in *warehouse* as the input during the local run. You can view the log output in the console.
- The system downloads the table data from the specified MaxCompute project to the *w* arehouse directory. If the data is already downloaded, the system does not perform this step.

Package and upload the Graph program

After you debug the Graph program, compress the Graph program into a JAR package and upload the package to the MaxCompute server as a resource. For more information, see Package, upload, and register.

Run the Graph program

Run the Graph program on the MaxCompute client.

- 1. In the left-side navigation pane, click **Project Explorer**.
- 2. Right-click the name of your project and select **Open in Console**.
- 3. In the **Console** tool window, run the following command to start your Graph program. For more commands, see JAR command.

jar -libjars xxx.jar -classpath /Users/home/xxx.jar com.aliyun.odps.graph.examples.PageRank pagerank_ in pagerank_out;

3.7.6. Query unstructured data

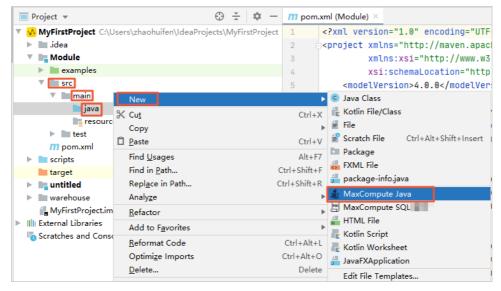
MaxCompute V2.0 allows you to use external tables to access Object Storage Service (OSS) and Tablestore. MaxCompute Studio provides code templates to help you query unstructured data. This topic describes how to use MaxCompute Studio to query unstructured data.

Prerequisites

- A MaxCompute project is connected. For more information, see Manage project connections.
- A MaxCompute Java module is created. For more information, see Create a MaxCompute Java module.

Write a StorageHandler, Extractor, or Outputer program

1. In the **Project** tool window, expand your MaxCompute Java module and choose **src** > **main** > **java**. Then, right-click java and choose **New** > **MaxCompute Java**.



2. Specify Name, select Extractor, StorageHandler, or Outputer, and then press Enter.

Create new MaxCompute java class		
E Name		
🞯 Driver		
🕼 Mapper		
🚯 Reducer		
S StorageHandler		
E Extractor		
0 Outputer		
UDJ LOD		
🕻 GraphLoader		
11 Vortov		

- **Name**: the name of the MaxCompute Java class that you want to create. If no package is created, enter packagename.classname. The system automatically creates a package.
- Select Extractor, StorageHandler, or Outputer as the class type.

3. After the class is created, develop a Java program in the editor. The Java template is automatically filled with framework code. You need only to compile the logic code based on your requirements.

Debug the Extractor or Outputer program

Write your test cases to debug your Extractor or Outputer program based on the unit test examples in the *examples* directory.

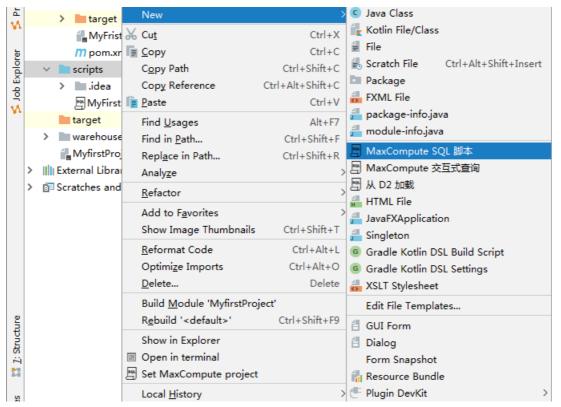
🗊 Project 🔹 😳 🕂	😴 ExtractorTest.java 🗙
🔻 <mark>🏑 studioTest</mark> ~/ldeaProjects/studioTest	ExtractorTest ambulanceFullSchema
idea	3 #import
🔻 🖿 jdev	18
examples	19 G pulic class ExtractorTest {
com.aliyun.odps.examples	<pre>20 private String ambulanceFullSchema =</pre>
► D mr	<pre>21 "vehicle:bigint;id:bigint;patient:bigint;calls:bigint;latitude:d</pre>
▶ D∎ udf	<pre>22 private String speechDataFullSchema = "sentence_snr:double;id:string</pre>
	23 24 @Test
unstructured	24 @Test 25 ♀ ⊖ public void testTextExtractor() throws Exception {
🔻 🛅 test	
💣 🖌 ExtractorTest	27 * Equivalent to the following SOL:
💣 🔓 OutputerTest	28 CREATE EXTERNAL TABLE ambulance data external
C SpeechSentenceSnrExtractor.java	29 (vehicle bigint, id bigint, patient bigint, calls bigint,
C & SpeechStorageHandler	30 Latitude double, Longitude double, time string, direction string)
C & TextExtractor	31 STORED BY 'com.aliyun.odps.udf.example.text.TextStorageHandler'
	32 LOCATION 'oss:///data/ambulance_csv/'
🙂 🚡 TextOutputer	<pre>33 USING 'jar_file_name.jar';</pre>
💿 🚡 TextStorageHandler	34 35 SELECT * FROM ambulance data external:
с 🔓 TestUtil	35 SELECT * FROM ambulance_data_external; 36 */
🕨 🖿 data	<pre>30 */ 37 Column[] externalTableSchema = UnstructuredUtils.parseSchemaString</pre>

Package and upload the program

After you debug the program, compress the program into a JAR package and upload the package to the MaxCompute server as a resource. For more information, see Package, upload, and register.

Query unstructured data

 In the Project tool window, right-click scripts under your MaxCompute project and choose New > MaxCompute SQL Script.



2. Enter the name of an SQL script in the Script Name field, select a MaxCompute project from the

MaxCompute Project drop-down list, and then click OK.

New MaxCompute	e SQL Script	×
Script Name:		
MaxCompute Project:	my_project700_dev (service.cn-shanghai.maxcompute.aliyun.com)	+
0	ОК	Cancel

- 3. In the editor, enter the SQL statement that is used to create an external table and click the ricon.
- 4. Create a MaxCompute SQL script, enter the following query statement, and then click the ▶ icon to query data.



3.7.7. Package, upload, and register a Java

program

After a Java program is developed, you must package and publish it to MaxCompute. This topic describes how to package, upload, and register a Java program.

Context

To publish a Java program such as a UDF, MapReduce program, or Graph program to MaxCompute for production use, you must package, upload, and register the Java program in sequence. You can use the one-click publish function to complete these procedures. MaxCompute Studio runs the **mvn clean package** command, uploads a JAR package, and registers a UDF in one stop.

Package the code

- 1. Right-click the compiled Java code and select Deploy to server....
- 2. In the Package a jar and submit resource dialog box, set the required parameters.

Package a jar, submit resource and register function		>	<
*MaxCompute project: doc_test_dev (service.cn-hangzhou.maxcompute.aliyun.com)	•	+	
*Resource file: C:\Users\zhaohuifen\IdeaProjects\Project2\mc_java\target\mc_java-1.0-SNAPSHOT.jar			
*Resource name: mc_java-1.0-SNAPSHOT.jar			
Extra resources:			
*Main class: com.aliyun.odps.examples.udf.UDTFResource			
*Function name:			
Force update if already exists			
? ОК		Cancel	

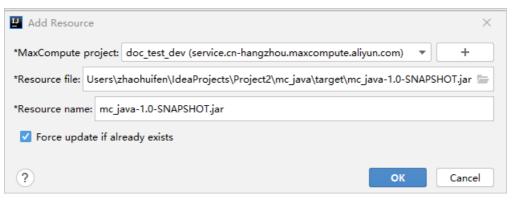
- MaxCompute project: specifies the name of the MaxCompute project.
- **Resource name:** specifies the name of the resource that you want to package.
- Function name: specifies the name of the function that you want to package.
- Force update if already exists: specifies whether to force the update when the resource or function already exists.
- 3. Click **OK** to complete the packaging.

(?) Note You can modify relevant settings in the pom.xml file based on your packaging requirements.

Upload the JAR package

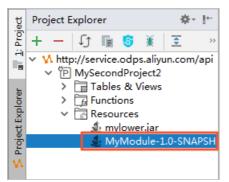
After the JAR package is prepared, you must upload it to MaxCompute.

- 1. In the top navigation bar, choose MaxCompute > Add Resource.
- 2. In the Add Resource dialog box, set the required parameters and click OK.



- MaxCompute project: specifies the name of the MaxCompute project.
- **Resource file:** specifies the path of the JAR package.
- Resource name: specifies the name of the resource that you want to upload.
- Force update if already exists: specifies whether to force the update when the resource or function already exists.
- 3. In the left-side navigation pane, click **Project Explorer**.

4. After the Java package is uploaded, you can view the resource under the **Resources** node of the **Project Explorer** window.



Register the UDF

You must register the UDF before you can call it.

- 1. In the top navigation bar, click MaxCompute and select Create Function.
- 2. On the Create Function page, set the following parameters and click OK.

Create Function
*MaxCompute project: MySecondProject2 (service.odps.aliyun.com) v +
*Function name:
mylower.jar
MyModule-1.0-SNAPSHOT.jar
*Using resources:
*Main class: MyFirstUDF ~
✓ Force update if already exists
function name can't be empty
OK Cancel

- MaxCompute project: specifies the name of the project to which you want to upload the UDF.
- Function name: specifies the name of the UDF.
- Using resources: specifies the name of the JAR package on which the UDF depends.
- Main class: specifies the main class of the JAR package.
- Force update if already exists: specifies whether to force the update when the resource or function already exists.
- 3. In the left-side navigation pane, click **Project Explorer**.
- 4. After the UDF is registered, you can view the UDF under the **Functions** node of the **Project Explorer** window.



3.8. Develop a Python program3.8.1. Configure a Python developmentenvironment

MaxCompute Studio allows you to develop Python programs, such as user defined functions (UDFs) and MaxCompute SDK for Python (PyODPS) scripts, in Intellij IDEA. This topic describes how to configure a Python development environment.

Install PyODPS

- 1. Install PyODPS. For more information, see Installation guide and limits.
- 2. Start Intellij IDEA. In the top navigation bar, choose File > Settings.
- 3. On the **Settings** page, click **MaxCompute Studio** in the left-side navigation pane.
- 4. Set the Python path to resolve UDF parameter to the local path of the Python.exe file.

E Settings	nsingay (lights, (IOC) - Instit EEA	×
Q	MaxCompute Studio	
Appearance & Behavior	General Settings	
Keymap ► Editor		
Plugins ▶ Version Control	Table preview rows: 200 (max preview 5000 rows)	
 Version Control Build, Execution, Deployment 		
 Languages & Frameworks Tools 	Python path to resolve UDF: C:\Python27\python.exe	
MaxCompute Studio		
SDK & Console MaxCompute SQL MaxCompute Job Accounts		
	Updates Checking	
	✓ Automatically checks for new version	
	Current version: 3.0.3 Latest version:	
?		

Install the Python plug-in

In the plug-in repository of Intellij IDEA, search for Python Community Edition and install the Python plug-in.

- 1. In the top navigation bar of Intellij IDEA, choose File > Settings.
- 2. On the **Settings** page, click **Plugins** in the left-side navigation pane.
- 3. In the search box, enter Python Community Edition to search for the Python plug-in. Then, install the Python plug-in.

Q _*	Plugins	Marketplace	Installed	\$	
> Appearance & Behavior	Q+ Python Community Edition	×			
Keymap	Search Results (1)		ر الش	Python	Disable \lor
Editor	Python Community Edition		Б	Community	
Plugins Version Control	2019.2.192.6603.23 JetBrains			Edition	
Build, Execution, Deployment					
Languages & Frameworks				JetBrains 2019.2.192.6603.23	
Tools				2019.2.192.0005.25	
MaxCompute Studio			Plugin homepa		
			PyCharm blog Discussion for Issue tracker	um	
0				ОК	Cancel Apply

Configure the Python dependency

After you configure the Python dependency for the MaxCompute script module, you can develop Python programs by using MaxCompute Studio.

- 1. In the top navigation bar, choose File > Project Structure.
- 2. Add the Python SDK.
 - i. In the left-side navigation pane, click **SDKs** under Platform Settings.

ii. Click the plus sign (+) at the top and add the Python SDK.

Project Structure			
⇐ ⇒	+ -	Name: Python 2.7	
Project Settings	1.8		
Project	Kotlin SDK	Python SDK home path: C:\Python27\python.exe	
Modules	Python 2.7	Classpath Packages	
Libraries			_
Facets		C:\Python27\DLLs	
Artifacts		C:\Python27\Lib\lib-tk	
Platform Settings		C:\Python27	
		C:\Python27\Lib\site-packages	
SDKs		C:\Users\Administrator\.IdeaIC2018.1\system\python_stubs\-762174762	
Global Libraries		C:\Users\Administrator\.IdeaIC2018.1\config\plugins\python-ce\helpers\python-skeletons	
		C:\Users\Administrator\.IdeaIC2018.1\config\plugins\python-ce\helpers\typeshed\stdlib\2	
Problems		C:\Users\Administrator\.IdeaIC2018.1\config\plugins\python-ce\helpers\typeshed\stdlib\2and3	
		C:\Users\Administrator\.IdeaIC2018.1\config\plugins\python-ce\helpers\typeshed\third_party\2 C:\Users\Administrator\.IdeaIC2018.1\config\plugins\python-ce\helpers\typeshed\third_party\2and3	
		C:\Users\Administrator\.ideatC2016.1\config\plugins\python-ce\neipers\typesned\third_party\zand5	
0		OK Cancel A	
2		OK Cancel Ar	pply

- iii. Click Apply.
- 3. Configure the Python dependency for the MaxCompute script module.
 - i. In the left-side navigation pane, click **Modules** under Project Settings.
 - ii. Click the plus sign (+) at the top and add the dependency on the Python SDK for the MaxCompute script module.

Project Structure				
(= = →	+ - 1	Python Interpreter:	Python 2.7 (MySecondProject) C:\Python27\python.exe	~
Project Settings	MyFirstModule			
Project	Python			
Modules	🗸 🗸 MySecondProject			
	🥐 Python			
Libraries				
Facets				
Artifacts				
Platform Settings				
SDKs				
Global Libraries				

iii. Click Apply.

3.8.2. Develop a Python UDF

MaxCompute Studio allows you to develop Python user defined functions (UDFs). This topic describes how to develop, test, and publish a Python UDF.

Prerequisites

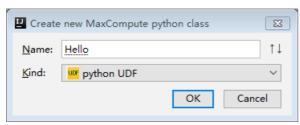
> Document Version: 20210713

The following operations are complete:

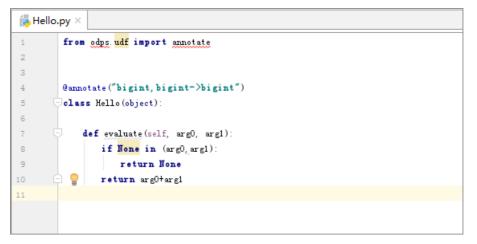
- Manage project connections
- Configure a Python development environment

Develop a Python UDF

- In the Project section, right-click scripts under the MaxCompute script module and choose New > MaxCompute Python.
- 2. In the **Create new MaxCompute python class** dialog box, enter a class name in the **Name** field, select **python UDF** from the Kind drop-down list, and then click **OK**.



3. Write the UDF code in the code editor.



Test the Python UDF

After the UDF is developed, you must test whether the UDF code works as expected. MaxCompute Studio supports local runs for you to locally test code. You can download some sample data from a table to your local computer, run the code on the sample data, and debug the code.

- 1. Right-click the developed Python UDF and select RUN.
- 2. In the Edit configuration dialog box, set the parameters and click OK.

Edit configuration			×			
<u>N</u> ame: test		Share	e through VCS 💿 🗌 Allow parallel r <u>u</u> n			
Configuration Log	s					
*Script:	C:\Users\liyang\MaxCompute	Studio\scripts\test.py	1			
Script Parameter:			× ۲			
*MaxCompute project:	local	~	example_project ~ +			
*MaxCompute table:	wc_in1		~			
*Table columns:	col1		ie:c1,c2			
Download Record limi	it: 100					
 Environment 						
Environment variables:	PYTHONUNBUFFERED=1					
Python Interpreter:	● Use SDK of module:	<u> Max</u> Compute Studio	~			
	○ <u>U</u> se specified interpreter:	<project default=""></project>	~			
Interpreter options:			ж ^ж			
Working directory:	C:\Users\liyang\MaxCompute	e Studio\scripts				
Add content roots t	o PYTHONPATH					
Add source roots to	PYTHONPATH					
 ■ <u>B</u>efore launch: Activate + -	tool window					
There are no tasks to run before launch						
□ Show this page 🗹 A	ctivate tool window					
Error: Please select a	module with a valid Python SD	к				
		l	Run Cancel <u>A</u> pply			

- **MaxCompute project**: the MaxCompute project that is used to run the UDF. For a local run, select **local**.
- MaxCompute table: the name of the MaxCompute table that is used to run the UDF.
- Table columns: the columns in the MaxCompute table that are used to run the UDF.
- **Download Record limit**: the maximum number of records that can be downloaded from the table. Default value: 100.

? Note

- If the specified data is already downloaded, MaxCompute Studio does not download the data again. To download the data again, run the tunnel command on the MaxCompute client.
- By default, 100 records are downloaded. To download more sample data, run the tunnel command on the MaxCompute client or use the table download feature of MaxCompute Studio.
- After the sample data is downloaded, you can view the data in the *data* file of the table under the *warehouse* directory.
- 3. MaxCompute Studio obtains the sample data from the *data* file based on the columns you specified and calls the UDF to locally run it on the sample data.

(?) Note Local runs are implemented by using the pyou script that the MaxCompute SDK for Python (PyODPS) provides. The command for starting a local run is pyou hello.Plus<data . After you install PyODPS, you can run the corresponding command to check whether the pyou script exists:

- For Windows, run the *\${python}/../Scripts/pyou* command.
- For macOS, run the **\${python}/../pyou** command.
- 4. View the output on the Console tab, check whether the UDF works as expected, and debug the code if required.

0	Project • ⊕ ≑ ∯• I**	🔒 hello.py 🗙
h y	studioTest ~/IdeaProjects/studioTest idea scripts leftello.py target studioTest.iml External Libraries	<pre>from odps.udf import annotate amoutate("bigint,bigint->bigint") class Plus(object): def evaluate(self, arg0, arg1): self: <hello.plus 0x113530f90="" at="" object=""> arg0: 1 arg1: None if None in (arg0, arg1): return None return arg0 + arg1 </hello.plus></pre>
Debu	g 壳 hello	
	g 🔩 hello Debugger 📧 Console 📲 🔚 💽 🗎 🋬 🗎	Ž ‰ ¥1
		Ž ‰ ×I □ → ≣ Variables

Publish the Python UDF

After the Python UDF is tested, you can publish it to the production environment. For more information, see Package, upload, and register a Java program.

3.8.3. Develop a PyODPS script

This topic describes how to develop a MaxCompute SDK for Python (PyODPS) script.

Prerequisites

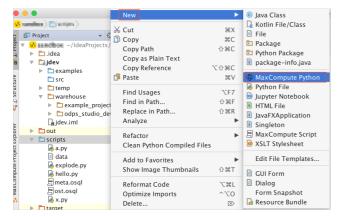
The Python development environment is configured. For more information, see Configure a Python development environment.

Context

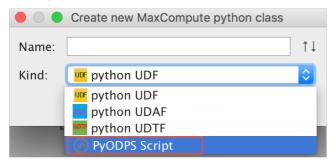
PyODPS supports basic operations on MaxCompute objects and allows you to use DataFrames. You can use PyODPS to analyze data in MaxCompute.

Procedure

1. In the Project section, right-click scripts and choose New > MaxCompute Python.

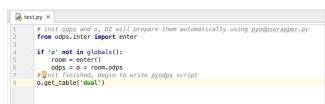


2. In the **Create new MaxCompute python class** dialog box, enter a class name in the **Name** field, select **PyODPS Script** from the **Kind** drop-down list, and then click OK.



3. Write the script code in the code editor. When you create a PyODPS script, the odps and o objects are automatically initialized in the PyODPS script template by using a PyODPS room.

When you develop a PyODPS script in DataWorks, a room is automatically created. When you develop a PyODPS script in Intellij IDEA, you must create a room yourself. For more information, see the PyODPS documentation.



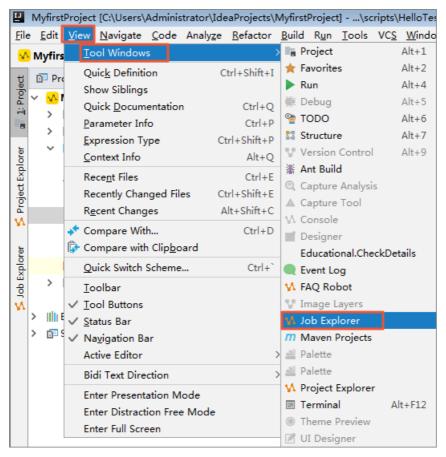
3.9. Manage MaxCompute jobs

3.9.1. View jobs

This topic describes how to use Job Explorer in MaxCompute Studio to view MaxCompute job instances that you submitted. You can view the running status, type, and start and stop time of job instances.

Open Job Explorer

To open **Job Explorer**, choose **View > Tool Windows > Job Explorer** in the top navigation bar.



View all job instances in a project

Job Explorer allows you to query the list of submitted jobs by status, user, date, and duration. Examples:

- If you want to query the list of submitted jobs by status, click **Status** and select Failed to view the jobs that failed within the last 24 hours.
- If you want to query the list of submitted jobs by date, adjust the sliders in the **Date** drop-down list to specify the time range of the jobs that you want to query.

? Note

- By default, only the first 1,000 jobs that meet the conditions are displayed. If more than 1,000 jobs meet the conditions, update the filtering conditions.
- You can click ID and Name and Submitted At to sort the jobs.

View a job queue

If a job in the running state is waiting for scheduling in a queue, click **Queue Location** to obtain the location in the queue, and click **Priority** to obtain the global priority.

? Note The job status and queue location on the Running Instances tab are automatically updated. After a job finishes, this information is removed from the list.

Save job logs

By default, Logview logs of a job are saved for seven days. If you want to save some important Logview logs for a longer period and view them in the future, you can save them locally.

- 1. Double-click a job in the list to display the job details on the right.
- 2. Click Save in the toolbar to save the logs to your local host.

You can set the path for saving the log file on the **Settings** page of MaxCompute Studio.

3.9.2. Job instance details

This topic describes how to view job instance details in MaxCompute Studio.

View job instance details

MaxCompute Studio provides the following methods to view the details about a MaxCompute job instance:

• Use a Logview URL or a local offline Logview file to open the details about a job instance in read-only mode.

Logview is a common method to view details about job instances. It allows you to view the status of tasks that are submitted by other users in other projects. You can view details about a job instance by entering a valid Logview URL in MaxCompute Studio.

In the top menu bar of Intellij IDEA, choose **MaxCompute > Open Logview**. In the **Open job detail by logview** dialog box, you can copy a valid Logview URL or export a local offline Logview file.

• In Job Explorer, double-click a MaxCompute job instance to view its details.

Job details view

The job details view page consists of a toolbar at the top, a properties bar on the left pane, and details views on the right pane. The details view pane contains the following four views:

- Execution view: displays the overall information for a job in the form of a directed acyclic graph (DAG). You can view the dependencies and detailed execution plans for each subtask.
- **Details**: displays the details about a job as a table. These details include the subtask list, worker list for each subtask, amount of data that is processed by the workers, execution time, and job status.
- Script: displays the SQL statement and parameter settings that are configured for a job that is submitted.
- **Result** : displays the execution results of a job.

The job details view page also provides the following information:

- **Timeline view**: displays the execution timeline of a job. This view provides multiple filters and allows you to view the timeline based on different granularities.
- Summary (JSON): displays the running details about a job in the JSON format.
- Analysis: provides scatter charts, long tail distributions, and skewed data charts to show the execution results of a job.

Toolbar

- < or >: used to expand or collapse views on the left and right panes. These icons allow you to focus on a specific view.
- S: used to stop a job that is running. Only the project owner or administrator has the permission to stop running jobs.
- : used to refresh details. The basic information such as status and quota of jobs is automatically refreshed. However, the details views on the right-side pane are not automatically refreshed. If you want to view the up-to-date details, you must manually refresh them.
- 👩: used to copy Logviews.
- 🛱: used to open job details in a browser. You can click this icon to generate a Logview URL and open the URL in a browser.
- 🖹: used to save job details as a local file.
- Enable auto refresh : used to determine whether to enable the auto refresh feature. If this feature

is enabled, MaxCompute Studio automatically refreshes all details about a running job on a scheduled basis.

Basic information page

This page displays the basic information about a job, including its ID, owner, status, start time and end time, computing resource usage, input tables, and output tables. The basic information about a running job is automatically refreshed on a scheduled basis.

You can double-click a table name to view the basic information about the table.

Execution view

The **Execution view** is used to display the dependencies in the Job Scheduler Job, Job Scheduler Task, and Operation layers. It also provides a series of auxiliary tools such as job playback, progress view, and heat map view. The Execution view is useful for troubleshooting.

The following items describe the meaning of the marked elements in the Execution view:

- 1. Click to go to another layer.
- 2. Zoom in and zoom out.
- 3. The dependency table.
- 4. The Job Scheduler task worker.
- 5. The thumbnail.
- 6. Click to display the job dependencies.
- 7. The dependencies in the Job Scheduler Task layer are displayed by default.

The Execution view displays job dependencies in three dimensions: Job Scheduler Job layer, Job Scheduler Task layer, and Operation layer. Click the up arrow to switch between dimensions. By default, dependencies in the Job Scheduler Task layer are displayed.

• Job Scheduler Job layer

Click **MaxCompute Job** to open the Job Scheduler Job layer. Workers in the Job Scheduler Job layer include the name, start time, and end time of each Job Scheduler task.

• Job Scheduler Task layer

Double-click a Job Scheduler job worker to go to the Job Scheduler Task layer.

If multiple Job Scheduler jobs exist, the Job Scheduler Task layer of the last Job Scheduler job is opened by default. This layer displays the dependencies, input tables, output tables, and partitions of Job Scheduler tasks. When a job is completed, you can click the drop-down box in the toolbar to switch between **Progress Chart**, **Input Heat Chart**, **Output Heat Chart**, **TaskTime Heat Chart**, and **InstanceTime Heat Chart**. The progress chart displays the task progress for the worker. The heat chart uses different colors to distinguish between the heat sources of different workers.

A Job Scheduler task worker shows the following information:

- Instance Count: expressed in the format of a/b/c. This expression indicates the status of the subtask instances at a specific point in time. The value of a represents the number of running subtask instances, the value of b represents the number of completed subtask instances, and the value of c represents the total number of subtask instances.
- I/O Records: displays the numbers of input and output records at a specific point in time.
- Percentage and orange progress bar: indicates the running status of the task. The percentage is obtained based on the analysis of running subtask instances.
- Line that connects subtasks: shows the number of output records. The arrow indicates the direction of data flow.
- Operation layer

Double-click a Job Scheduler task to open the Operation layer.

The Operation layer reveals how Job Scheduler tasks run internally. You can click a worker to view all Operation layer information.

Note For non-SQL jobs, the system displays only jobs in the Job Scheduler Job and Job Scheduler Task layers. Jobs in the Operation layer cannot be displayed.

Details view

The Details view is used to display details for SQL data manipulation language (DML) jobs. This view displays the Job Scheduler task lists and compute worker lists of SQL DML jobs on the compute cluster. Typically, each job is associated with one or more Job Scheduler jobs. Each Job Scheduler job is divided into multiple Job Scheduler tasks (stages), and each Job Scheduler task contains multiple Job Scheduler instances (workers). You can right-click a Job Scheduler instance to view the standard output, standard errors, and debugging information.

The following items describe the meaning of the marked elements in the Details view:

- 1. The Job Scheduler job tab.
- 2. The Job Scheduler task list.
- 3. The details and compute worker list of each Job Scheduler task.

Job playback

MaxCompute Studio provides the job playback feature. You can view the execution history of a job within 12 seconds in the same manner as playing a media file. This feature helps you understand the running status of a MaxCompute job instance at different time points. You can determine the order and the amount of time to run subtasks and obtain the key path to run the job. This allows you to optimize subtasks that run slowly.

Click the > icon to start playing and click the icon again to pause. You can also drag the progress bar.

? Note The playback feature only estimates the amount of I/O data at a specific point in time based on the execution time to determine the job progress. This does not represent the actual amount of I/O data at the point in time. Jobs in the Running state do not support the playback feature.

Timeline view

The Timeline view displays the details for the distributed execution of a job in a Gantt chart. You can adjust the display granularity to show all compute workers in a Gantt chart. Gantt charts display the time bottlenecks and long-tail workers of running jobs in a clear manner. You can also use multiple filters to select the key path, the largest data worker, and the worker that consumes the longest amount of time during the job execution.

Analysis view

The Analysis view displays the long-tail workers and skewed data workers of jobs. It also displays worker scatter charts and column charts to help diagnose job execution bottlenecks. Scatter charts and column charts allow you to click a worker to open the details view page and view the details about Job Scheduler instances.

Result view

The Result view displays information based on the job type and the parameters that are set when you submit the job.

- For SELECT statements that contain odps.sql.select.output.format = HumanReadable , this view displays the result in the text format.
- For SELECT statements that do not specify the output format parameter, this view displays the result in the table format.
- For scripts that are run to export data to a table, this view displays the name of the output table and the link that redirects to the table details.
- For abnormal jobs, this view displays the error details.

3.10. Tool integration

3.10.1. Integrate with MaxCompute client

MaxCompute Studio is integrated with the MaxCompute client program. You can open the client on MaxCompute Studio.

Configure the client installation path

 MaxCompute Studio contains the MaxCompute client of the latest version, which is specified as the default client. You can also install the client of another version by selecting Settings > MaxCompute Studio > SDK & Console on Intellij IDEA and adding the client program and path. Console download address

Settings	AD INTOXICIANTIN	Nethel +		X
Q) MaxCompute Studio > S	DK & Console		Reset
▶ Appearance & Behavior	MaxCompute Console Det	ails		
Keymap				
▶ Editor	Installed Location:			
Plugins	Version:	No valid Client installed in the specified path. Please check or <u>download</u>		
Version Control				
Build, Execution, Deployment				
Languages & Frameworks				
▶ Tools				
MaxCompute Studio				
SDK & Console				
MaxCompute SQL				
Accounts				
Maven Helper				
		ок	Cancel Apply H	lelp
		UK	Cancer Apply	cip

2. After setting is successful, the MaxCompute client version is displayed.

Settings	and an increase and the	withd +			23
٩	MaxCompute Studio > SD	K & Console	•		Reset
Appearance & Behavior Keymap	MaxCompute Console Deta	ils			
► Editor	Installed Location:		elop\odpscmd_public		
Plugins > Version Control > Build, Execution, Deployment > Languages & Frameworks > Tools Y MaxCompute Studio SDK & Console MaxCompute SQL Accounts Maven Helper	Version:	0.23.3			
				OK Cancel	Apply Help

Open the MaxCompute client

After the MaxCompute client installation path is set, you can open the client program on MaxCompute Studio.

1. In the project browsing list, right-click a project to be opened and select **Open in Console**.

Tools and Downloads• MaxCompute Studio

MaxComput e

MaxCompute Project Explorer 🕸 🗸	← 🔄 hello_word.osql ×	compiled:hello_word_osql_sql ×	🔄 test_demo.osql 🗙	📑 studio-script-demo.iml 🛪
+ - 〔7 王 ≍ ♥ [P] meta dev ■ ☐ Tal Add project ■ ☐ Ful Remove project ■ ☐ Re ♥ [P] meta Refresh meta Show Modify project properties Open in Console	hello_word	.osql		
MaxCompute Console Aliyun ODPS Command Line Tool Version 0.22.2 @Copyright 2015 Alibaba Cloud Computing Co., Ltd odps@ meta_dev>]	. All rights reserved	l.		

2. You can open multiple client programs by following the preceding steps.



4.Downloads

This topic provides the download addresses of the tools and plug-ins that may be required when you use MaxCompute.

- SDK: Maven users can search odps-sdk from the Maven repository to obtain different versions of the SDK for Java.
- MaxCompute client: Download the selected version of the MaxCompute client from Git.
- Intellij IDEA: Download Intellij IDEA and the MaxCompute Studio plug-in.
- Java Database Connectivity (JDBC): MaxCompute provides an open source JDBC driver. You can download the latest JDBC driver from Git Hub.
- SDK for PHP: You can download the SDK for PHP from Git Hub.