# Alibaba Cloud MaxCompute

**Tools and Downloads** 

Issue: 20190701



## <u>Legal disclaimer</u>

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

- 1. You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
- 2. No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company, or individual in any form or by any means without the prior written consent of Alibaba Cloud.
- 3. The content of this document may be changed due to product version upgrades , adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall 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 the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults " and "as available" basis. 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 bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, 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 content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, 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 the Alibaba Cloud website, product programs, or content shall be used, modified , reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates . The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates).
- 6. Please contact Alibaba Cloud directly if you discover any errors in this document.

## **Generic conventions**

#### Table -1: Style conventions

Style	Description	Example
-	This warning information indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	<b>Danger:</b> Resetting will result in the loss of user configuration data.
	This warning information 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 restore business.
	This indicates warning informatio n, supplementary instructions, and other content that the user must understand.	• Notice: Take the necessary precautions to save exported data containing sensitive information.
	This indicates supplemental instructions, best practices, tips, and other content that is good to know for the user.	Note: You can use Ctrl + A to select all files.
>	Multi-level menu cascade.	Settings > Network > Set network type
Bold	It is used for buttons, menus , page names, and other UI elements.	Click OK.
Courier font	It is used for commands.	Run the cd / d C :/ windows command to enter the Windows system folder.
Italics	It is used for parameters and variables.	bae log list instanceid Instance_ID
[] or [a b]	It indicates that it is a optional value, and only one item can be selected.	ipconfig [-all -t]

Style	Description	Example
{} or {a b}	It indicates that it is a required value, and only one item can be selected.	<pre>swich {stand   slave}</pre>

## Contents

Legal disclaimerI
Generic conventions
1 Client
2 MaxCompute Studio5
2.1 What is Studio
2.2 Project space connection management10
2.3 Tools Installation and version history15
2.3.1 Install IntelliJ IDEA15
2.3.2 Installation procedure
2.3.3 View and upgrade the version
2.4 Manage data and resources
2.4.1 View tables and UDF
2.4.2 Import and export data
2.4.3 Visualization of operating the tables
2.5 Develop SQL procedure
2.5.1 Create MaxCompute Script module
2.5.2 Write SQL
2.5.3 Submit SQL scripts
2.6 Developing Java
2.6.1 Create MaxCompute Java Module
2.6.2 Develop and debug UDF
2.6.3 Develop MapReduce
2.6.4 Unstructured development
2.6.5 Develop Graph
2.6.6 Package, Upload and Register
2.7 Manage MaxCompute jobs
2.7.1 Job viewing
2.7.2 JOD Instance
2.8 1001 integration
2.8.1 Integrate with MaxCompute client
2.9 Configure options
2.9.1 Configure MaxCompute Studio
3 Eclipse Plugins 90
3.1 Install
3.2 Create a MaxCompute project
3.3 MapReduce97
3.4 UDF 113
3.5 Graph120
4 Downloads124

## 1 Client

This article describes how to use the basic functions of the MaxCompute using client command line tool. Before using the MaxCompute client, you must *install and configure the client*.



- Do not perform the analysis operation based on the output format of the client.
   The output format of the client is not ensured for forward compatibility. Clients of different versions are different in their command formats and behaviors.
- · For more information about basic commands of the client, see *Basic commands*.
- · *Click here* to download the new version of MaxCompute client.
- The client supports JDK 1.9 from the 0.28.0 version, and the previous version can only use JDK 1.8.
- The client supports MaxCompute 2.0 from the 0.27.0 version *New data type*.

After the client is installed and configured, you can use a command line to perform the following operations.

#### Get Help

To view the help information of the console, the command format is as follows:

odps @>./ bin / odpscmd - h ;

You can also input h; or help; (case-insensitive) in an interactive mode.

The console also provides the help [keyword]; command to get the command prompts related to the keyword. For example, input help table ; to get command prompts related to the table operation as follows:

```
odps @
                            table ;
          odps >
                   help
                                      smallfiles
Usage :
          alter
                   table
                            merge
                  tables
Usage :
          show
                            [ in ]
                 of
                       tables | ls [- p ,- project ]
deserves mention | [.] [ partition ()]
           list
       а
Usage :
          describe
```

Usage : read [.] [ partition ()] [ line\_num ]

#### Start parameters

When start the console, you can specify a series of parameters as follows:

```
Usage : odpscmd
                   [ option ]...
      options
                   include :
where
   -- help (- h ) for
-- project = use p
                          help
   -- project = use project
-- endpoint = set endpoint

u - p user name and
k will skip begining

                                    password
         will skip
                                                               from
                         begining
                                    queries
                                                and
                                                      start
specified position
   - r
        set retry
                        times
   - f <" file_path ;"> execute
                                                  in
                                                       file
                                       command
   - e <" command ; [ command ;]..."> execute
                                                      command ,
                                                                  include
        command
  sql
                 display
   – C
         will
                            job
                                  counters
```

Take the -f parameter as an example, the operation is as follows:

1. Prepare the local script file script.txt. Suppose that the file is located in the disk D, and the content is shown as follows:

DROP test\_table TABLE ΙF EXISTS \_mj ; CREATE test\_table TABLE \_mj (id string ); string , name DROP TABLE test\_table \_mj ;

2. Run the following command:

odpscmd \ bin > odpscmd - f D :/ script . txt ;

Interactive mode

Run the console to directly enter the interactive mode:

```
[ admin : ~]$ odpscmd
         ODPS Command
 Aliyun
                         Line
                               Tool
Version
        1.0
@ Copyright
            2012
                   Alibaba
                             Cloud
                                    Computing
                                                Co ., Ltd .
                                                             All
       reserved .
 rights
                                   TABLE
                                           DUAL
odps @
        odps > INSERT
                        OVERWRITE
                                                  SELECT *
                                                            FROM
DUAL ;
```

Enter the command at the cursor position (use a semicolon as a statement terminator ), and press Enter to run.

#### **Continuous running**

- When using -e or -f option to run a command, if there are multiple statements, and you want to start running from a middle statement, you can specify the parameter -k, indicating to ignore the previous statements and to start running from the specified position. When the parameter <= 0 is specified, the execution starts from the first statement.
- Each statement separated by a semicolon is considered as a valid statement. The statements which run successfully or fail to run are printed out at runtime.

For example,

suppose there are three SQL statements in the file /tmp/dual.sql:

```
drop table dual;
create table dual ( dummy string );
insert overwrite table dual select count (*) from dual;
```

To ignore the first two statements, and start running from the third statement, the command format is as follows:

odpscmd - k 3 - f dual.sql

Get current logon user

To get current logon user, the command format is as follows:

whoami ;

Use example:

```
odps @ hiveut > whoami ;
Name : odpstest @ aliyun . com
End_Point : http :// service . odps . aliyun . com / api
Project : lijunsecur itytest
```

Use the preceding command to get the current logon user Alibaba Cloud account, endpoint configuration, and project name.

#### Exit

#### To exit the console, the command format is as follows:

odps @ > quit ;

#### You can also use the following command to exit the console:

odps@>q;

## 2 MaxCompute Studio

### 2.1 What is Studio

MaxCompute Studio is a big data integrated development environment (IDE) tool that is provided by the Alibaba Cloud MaxCompute platform and installed on the developer's client. It is a development plug-in based on the popular integrated development platform *IntelliJ IDEA*, helping users develop data conveniently. This article describes functional interfaces and common application scenarios of MaxCompute Studio.

**Basic user interface** 

MaxCompute Studio is a plug-in on the IntelliJ IDEA platform, which shares basic development interfaces with IntelliJ IDEA. For more information about the IntelliJ IDEA interfaces, see *the Interface operation guide*.

Based on the IntelliJ IDEA interfaces, MaxCompute Studio provides the following functional interfaces.

• SQL Editor: Provides features such as SQL syntax highlighting, code complementing, real-time error prompting, local compilation, and job submission.

Compiler View: Displays locally compiled prompts and error messages, and locates the code in the editor.

• Project Explorer: Connects to a MaxCompute project, and browses table structures, custom functions, and resource files in the project.

Table Details View: Displays details and sample data of tables, views, and other resources.

- · Job Explorer: Browses and searches for historical jobs of MaxCompute.
  - Job Details View: Displays running details of a job, including the execution plan and details of each execution task.
  - Job Output View: Displays output information of a running job.
  - Job Result View: Displays the output result of the SELECT job.
- MaxCompute Console: Integrates the *MaxCompute client*, on which MaxCompute client commands can be input and executed.

#### Connect to MaxCompute project

Before using most features of MaxCompute Studio, you must *Create a project connection*. After the project connection is created, you can view related data structures and resource information in the Project Explorer. MaxCompute Studio automatically creates a local metadata backup task for each project to increase the access frequency to MaxCompute metadata and reduce the latency.

## Note:

- You must specify the target project connection to modify SQL scripts, submit jobs, view job information, open the MaxCompute console, and implement other functions using MaxCompute Studio. Therefore, creating a connection to the MaxCompute project is necessary.
- · For more information about MaxCompute projects, see Project.
- For more information about project management using MaxCompute Studio, see *Project space connection management*.

#### Manage data

You can use the Project Explorer of MaxCompute Studio to quickly browse table structures, custom functions, and resource files in the project. The tree control can be used to list data tables, columns, partition columns, virtual views, custom functions, function signatures, and resource files and types of all project connections. It also supports fast locating.

You can double-click a data table to open the Table Details View and view metadata, structure, and sample data of the data table. If you do not have the permission for a project, an error message is prompted.

MaxCompute Studio integrates *MaxCompute Tunnel* and supports local data upload and download. For more information, see *Import and export data*.

#### Write SQL scripts

You can easily compile a MaxCompute SQL script on MaxCompute Studio.

- 1. Open MaxCompute Studio and select File > New > Project or File > New > Module….
- 2. Create a MaxCompute Studio project or module.

3. Select File > New > MaxCompute Script or right-click the menu and select New > MaxCompute Script , to create a maxcompute SQL script file.

## Note:

When a MaxCompute SQL script is created, MaxCompute Studio prompts you to select an associated MaxCompute project. You can also modify the associated project using the project selector on the right of the toolbar on the SQL editor. The editor automatically checks metadata (such as the table structure) and reports errors of an SQL statement based on the project associated with the SQL script. The editor also sends the SQL statement to the associated project for execution when it submits the SQL statement for running. For more information, see *Compile an SQL script*.

#### SQL code intelligent prompt

After you enter the code, the SQL editor provided by MaxCompute Studio intelligently prompts the syntax errors, type matching errors, or warnings of SQL statements, and marks them on the code in real time. as shown in the following figure.



By using the code complementing function, MaxCompute Studio prompts you the name, table, field, function, type, and code keyword of a project based on the code context, and automatically complements the code based on your selections. as shown in the following figure.



#### 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 syntax or semantic errors occur, the editor reports it.

		— select clause in the front
		<pre>select * from table_test;</pre>
		— from clause in the front
		<pre>from table_test table_alias select *;</pre>
		— table name with project prefix
Max	Comp	oute Compiler
x	Ŧ	Information: Parsing
	-	Information: Type checking
	Ť	Information: Latency.compiler parse error : 44170
Ŧ	6	Information: Build failed(2)
<b>F</b> *	-1-	▼ 🕾 /Users/xueming.xm/IdeaProjects/MyUDF/Script/scripts/
Ы	~	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 **sector** icon on the toolbar of the SQL editor to submit an SQL script to the

queue of the project specified by MaxCompute.

#### View history jobs

Open Job Explorer to view recently executed jobs in the specified project.



List only displays jobs submitted by the user ID of the current connection.

MaxCompute Job Explorer 🎄 🕂 📔								
Project: sql_optimizer V Days: 2 V J								
Q (40/47883) Table 🔻								
Instanceld	Status	Owner	StartTime	EndTime				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	SUCCESS	ODPS	2017	2017				
20170	FAILED	ODPS	2017	2017				

Double-click a job to view the job details. as shown in the following figure.



If you have the Log view URL of a job, you can select MaxCompute > Open Logview from the menu to go to the details page of the job.

#### Develop a MapReduce program and UDF program

MaxCompute Studio also allows you to develop MapReduce and Java UDF programs.

Connect to a MaxCompute client

MaxCompute Studio is integrated with the MaxCompute Client of the latest version.

Alternatively, you can specify the path of the locally installed MaxCompute client on the *Configuration page* of MaxCompute Studio.

On the Project Explorer, right-click a project and select Open in Console to open the MaxCompute Console window.



#### Next step

Now, you know the functional interfaces and common application scenarios of MaxCompute Studio. Continue to the next tutorial. In this tutorial, you will learn how to install MaxCompute Studio. For more information, see *Install Intellij IDEA*.

## 2.2 Project space connection management

MaxCompute One of the core features of MaxCompute Studio is to browse resources of a MaxCompute project, including Table, UDF, and Resource. To realize this feature, create a project connection first.

#### Prerequisites

To display Tool Windows of IntelliJ IDEA, you must open an IntelliJ IDEA project, and the MaxCompute Project must be configured on IntelliJ IDEA using MaxCompute Project Explorer in Tool Windows. Therefore, before creating a MaxCompute Project connection, add or import an IntelliJ IDEA project. This document uses adding a project under Windows as an example.

## Open IntelliJ IDEA, click New > Project, select MaxCompute Studio on the displayed page, and click Next.

MyFirstProject [C: ] June 1 and 11 HTTC: [IdeaProjects \MyFirstProject] - IntelliJ IDEA									
<u>File E</u> dit <u>V</u> iew <u>N</u> avigate <u>C</u> ode	Analy <u>z</u> e <u>R</u> e	efactor <u>B</u> uild	R <u>u</u> n	<u>T</u> ools	VC <u>S</u>	<u>W</u> indow	MaxCompute	<u>H</u> elp	
New	>	<u>P</u> roject							
<u>ो О</u> реп		Project from	Existing	Source	s				
Open <u>R</u> ecent	>	Project from	Version	Contro	d	>			
Close Project		Module		_					
Settings Ci	trl+Alt+S	Module from	n Existing	g Sourc	es	_			
Project Structure Ctrl+Alt	+Shift+S	Scratch File	Ctrl	+Alt+Sł	nift+Ins	sert			
Other Settings	0	FXML File	_	_	_	- 11			
Import Settings									
Export settings						-			
Export to Zip File						Searc	h Everywhe	re Double Shift	
Settings Repository						Golto	File Ctrl+S	hift+N	
<mark>≣ S</mark> ave All	Ctrl+S					00 11			
😳 Synchronize Ci	trl+Alt+Y					Recei	nt Files Ctrl-	۰E	
Invalidate Caches / Restart							1' D 4		
Print						Navię	Jation Bar A	lt+Home	
Line Separators	>					Drop	files here to	o open	
Power Save Mode									
E <u>x</u> it									
_								572	
Project								22	
Pe lava	Additiona	Libraries a	nd Fram	nework	(5:				

Java	Additional Libraries and <u>F</u> rameworks:
Java FX	
👾 Android	
뜬 IntelliJ Platform Plugin	
<i>M</i> Maven	Nothing to show
📀 Gradle	
G Groovy	
🧐 Griffon	
【 Kotlin	
Empty Project	
🔥 MaxCompute Java	
<mark>v M</mark> axCompute Studio	
	Previous <u>N</u> ext Cancel Help

Enter the project name, and click Finish.

🖳 New Project		×
Project n <u>a</u> me: MyF	irstProject	
Project location: C:\U	Jsers\Administrator\IdeaProjects\MyFirstProject	
More Settings		
Module na <u>m</u> e:	MyFirstProject	
Content <u>r</u> oot:	C:\Users\Administrator\IdeaProjects\MyFirstProject	
Mod <u>u</u> le file locatio	n: C:\Users\Administrator\IdeaProjects\MyFirstProject	
Project <u>f</u> ormat:	.idea (directory based)	$\sim$
	Previous Einish Cancel Help	

#### Create a MaxCompute Project Connection

We recommend that you configure the MaxCompute project connection according to your region strictly. Otherwise, some errors may occur.

Procedure

1. Click view, select Tool Windows.

iew <u>N</u> avigate <u>C</u> ode Analy <u>z</u> e <u>R</u> efactor	<u>Build Run T</u> ools VC	<u>S W</u> indow MaxCompute <u>H</u> elp
<u>T</u> ool Windows >	Project	Alt+1
Recent Files     Ctrl+E       Recently Changed Files     Ctrl+Shift+E       Recent Changes     Alt+Shift+C       Quick Switch Scheme     Ctrl+`       Tool Bart     Tool Buttons	<ul> <li>★ Favorites</li> <li>▶ Run</li> <li>₩ Debug</li> <li>™ TODO</li> <li>₩ Structure</li> <li>♥ Version Control</li> <li>₩ Act Build</li> </ul>	Alt+2 Alt+4 Alt+5 Alt+6 Alt+7 Alt+9
/ <u>S</u> tatus Bar / Na <u>v</u> igation Bar Bidi Text Direction >	<ul> <li>Ant Build</li> <li>Capture Analysis</li> <li>Capture Tool</li> <li>Console</li> </ul>	irch Everywhere Double Sh
Enter Presentation Mode Enter Distraction Free Mode Enter Full Screen	Designer     Event Log     Image Layers     Job Explorer	to File Ctrl+Shift+N ent Files Ctrl+E
	Maven Projects         Palette         Palette         Project Explorer	vigation Bar Alt+Home
	<ul> <li>Terminal</li> <li>Theme Preview</li> <li>UI Designer</li> </ul>	Alt+F12

#### 2. Click Project Explorer.

3. Click plus sign + at the upper left corner to add a MaxCompute project.

IJ	MyFirst	Project	[C:\U	di valiti	PRINCIPAL A	rojects\l	MyFirst	Projec	t] - Intel	lij IDE	A	
<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	MaxCompute
<b>~</b>	MyFir	stProje	ect $ angle$									
rer	Projec	t Explo	rer						<b>☆</b> - I+			
xplo	+ -	Ĵ	<b>₹</b>	🕘 (	D ?							
gt	🕤 Ad	d proje	ct from dat	taworks	/api							
Proj	🖉 Ad	d proje	ect from acc	:essId/K	ey							
M												

#### 4. In the Add MaxCompute Project dialog box, set configuration options.

P Add MaxCompu	ite project
Connection Set	ting
Properties File:	Initialize AK/Endpoint and other configuration
AK Account: To i	hitialize AK with AK account, click on the +
* Access Id:	Configuration information such as AK
* Access Key:	Endpoint when connecting to MaxCompute project is required. It can be filled in manually
* Project Name:	or initialized by the above Properties File or AK Account.
* End Point:	
	~
0	OK Cancel

Note:

- Click question mark (?) at the lower left corner of the dialog box to go to the online document page.
- If the synchronization times out, you can consider increasing the time-out duration for synchronizing metadata to the local host on the Setting tab.
- 5. After the preceding settings, click OK. Information about the MaxCompute project is displayed on the left of MaxCompute Project Explorer. You can click Tables, Views, Functions, and Resources to view tables, views, functions, and resources of the project.



#### View and modify a MaxCompute connection

In MaxCompute Project Explorer, right-click a MaxCompupte project and select Show| Modify project properties.

ч	MyFirst	Project	[C:\	4-438	19890 Me	aProjects\/	MyFirst	Projec	t] - Intel	lij IDE	Α
<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
<u>~</u>	MyFir	stProje	ect $ angle$								
rer	Projec	t Explo	rer						<b>☆</b> - ⊮		
Splo	+ -	- 1	<b>₹</b>		D ?						
t	~ <u>M</u> h	ttp://se	ervice.odps	.aliyun.c	om/api						
Proj	~ `	12 <mark>Myf</mark> ≻ 급	FirstProjectV Tables & Vi	Vdh ew: Ad	d project			>			
n		> 👍	Functions	Rei	move proj	ject					
Б		s-	Resources	Sho	ow Modify	/ project pr	opertie	es			
plor				Ret	fresh meta	3					
ΡE				Ор	en in Con	sole					
-				Ne	w sql edit	or					
¥1				Cre	eate a new	v table					
📷 1: Project											

In the displayed dialog box, you can view or modify connections and settings of the MaxCompute project.

#### Subsequent operations

Now, you know how to create and manage a project connection. You can continue to the next tutorial. In the tutorial, you will learn how to query metadata, clear data, and upload and download data to manage data and resources. For more information, see *Manage data and resources*.

## 2.3 Tools Installation and version history

## 2.3.1 Install IntelliJ IDEA

This document describes how to install the basic platform IntelliJ IDEA of MaxCompute Studio.

#### Context

#### Procedure

#### Procedure

1. Click *here* to download the IntelliJ IDEA of the version corresponding to your operating system (Windows, macOS, or Linux). The following assumes that the Windows operating system is used.

Download IntelliJ IDEA 14.1.4 or a later version. (The Ultimate version, PyCharm version, and free Community version are supported.)

2. After the download is complete, double-click the installation program to enter the installation page, and click Next, as shown in the following figure.



3. Specify the installation directory, and click Next, as shown in the following figure.

🛂 IntelliJ IDEA Communi	ty Edition Setup	)		
<b>(11)</b>	Choose Install Choose the fold Edition.	Location ler in which to inst	tall IntelliJ IDEA	Community
Setup will install IntelliJ IDE folder, dick Browse and se	EA Community Editi lect another folder	on in the following . Click Next to cor	) folder. To inst Itinue.	all in a different
Destination Folder rogram Files\JetBrains	\IntelliJ IDEA Comn	nunity Edition 201	8.1.4 Br	owse
Space required: 926.9 MB Space available: 121.8 GB				
		< Back	Next >	Cancel

4. Select the 32-bit or 64-bit IntelliJ IDEA based on the version of the local operating system.

You can query the local operating system version by following these steps:

- a) Open Windows Resource Manager, right-click Computer and select Properties from the shortcut menu. as shown in the following figure.
- b) In the displayed window, check the type of the operating system.

5. Select the corresponding system type and click Next , as shown in the following figure.

IntelliJ IDEA Community Edition Setup	
Installation Options Configure your IntelliJ IDEA Community Edition	installation
Create Desktop Shortcut 32-bit launcher 4-bit launcher	
Create Associations	
< Back Next >	Cancel

### 6. Click Install to start installation, as shown in the following figure.

LT & IPUTPEA C				
Intellij IDEA Con	nmunity Edition Setup			
<b>!!</b>	Choose Start M Choose a Start Edition shortcut	<b>fenu Folder</b> Menu folder for ti s.	he IntelliJ IDE/	A Community
Select the Start Mer can also enter a nar	nu folder in which you wou me to create a new folder.	ld like to create t	he program's s	shortcuts. You
JetBrains				
Accessories Administrative Too Cisco Eclipse Games Git Java Maintenance Microsoft Office 20 MySQL-Front Notepad++ Oxygen XML Editor	ls 013 r 18.1 (64-bit)			
		< Back	Install	Cancel
IntelliJ IDEA Com	munity Edition Setup			
IntelliJ IDEA Com	nmunity Edition Setup <b>Installing</b> Please wait while installed.	Intellij IDEA Cor	nmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed.	Intellij IDEA Cor	mmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed. _off_focused_holo_dark.9	IntelliJ IDEA Cor .png 100%	nmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed. _off_focused_holo_dark.9	IntelliJ IDEA Cor .png 100%	mmunity Edition	n is being
IntelliJ IDEA Com	nmunity Edition Setup Installing Please wait while installed. _off_focused_holo_dark.9	: IntelliJ IDEA Cor .png 100%	mmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed. _off_focused_holo_dark.9	IntelliJ IDEA Cor .png 100%	mmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed. _off_focused_holo_dark.9	IntelliJ IDEA Cor .png 100%	mmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed. _off_focused_holo_dark.9	IntelliJ IDEA Cor .png 100%	mmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed. _off_focused_holo_dark.9	e Intellij IDEA Cor .png 100%	mmunity Edition	n is being
IntelliJ IDEA Com	Installing Please wait while installed. _off_focused_holo_dark.9	: IntelliJ IDEA Cor .png 100%	mmunity Edition	n is being

7. After the installation is complete, click Finish.



#### What's next

Now you know how to install IntelliJ IDEA. Continue to the next tutorial. In this tutorial, you will learn how to install the MaxCompute Studio plugin. For more information, see *Install the MaxCompute Studio plugin*.

## 2.3.2 Installation procedure

#### **Environment requirements**

IntelliJ IDEA can be installed on *Windows*, *Mac*, *Linux*. For more information about the hardware and system environment requirements, click *here*. IntelliJ IDEA-based MaxCompute Studio can also be installed on clients running these operating systems.

MaxCompute Studio has the following requirements on the your environment:

- · A client running Windows, macOS, or Linux.
- IntelliJ IDEA 14.1.4 or a later version is installed. (The Ultimate version, PyCharm version, and free *Community version* are supported.)
- · JRE 1.8 is installed. (JRE 1.8 has been bound to the latest IntelliJ IDEA.)

• JDK 1.8 is installed. (*Optional*: JDK is required if you need to develop and debug Java UDF.)

## Note:

The client supports JDK 1.9 from the 0.28.0 version. The previous version only supports JDK 1.8.

#### Installation method

MaxCompute Studio is a plugin of IntelliJ IDEA, which can be installed using either of the following two methods:

- · Online installation using the plugin library (recommended)
- · Installation using a local file

#### Online installation (recommended)

MaxCompute Studio The MaxCompute Studio plugin has been opened for all users on the Internet. You can install MaxCompute Studio using the official IntelliJ IDEA plugin library.

Procedure

- Open the plugin configuration page on IntelliJ IDEA. (If you are a Windows/Linux user, choose File > Settings > > Plugins. If you are a macOS user, choose IntelliJ IDEA > Preferences > Plugins ).
- 2. Click Browse repositories... and search for MaxCompute Studio .
- 3. On the MaxCompute Studio plugin page, click Install.

#### dministrator\Local Settings] - IntelliJ IDEA - - 2 <u>C</u>ode Analyze <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 MaxCompute <u>H</u>elp U Settings Browse Repositories × Q Plugins 💋 Category: All 🗸 Q > Appearance & Behavior Q, Show: All plugins $\lor$ Sort by: name 🔻 Keymap env files support. > Editor FRAMEWORK IN1 3 m 🖺 Android S 🛃 Install Plugins 8,364,740 \*\*\*\* ignore 📲 Ant Support $\checkmark$ Ċ. ★★★★★ 962252 downloads VCS INTEGRATI > Version Control 4 months ago 📲 Bytecode Viewer $\checkmark$ Updated 2018/8/25 v0.7 .NR Null Object > Build, Execution, Deployment 882 \*\*\*\*\* 📲 Copyright $\checkmark$ Homepage | Github | Issues ANDROID > Languages & Frameworks one month ago 101 Header Coverage $\checkmark$ Features 3,341 \*\*\*\*\* > Tools Environment variables completion for PHP, JavaScript, Python, Go and Ruby languages Dockerfile and docker-compose.yml files support. Co to declaration(in.env file) and urconsolition and by Crit/Completicities CVS Integration $\checkmark$ FORMATTING 7 months ago > MaxCompute Studio $\checkmark$ 🕞 Eclipse Integration (1024 Tools CODE TOOLS 5,592 \*\*\*\* 🖷 EditorConfig $\checkmark$ 3 months ago 360 FireLine Plugin 35,236 ★★★★☆ 🕞 Git Integration usages(in code), by Ctrl(Cmd)+click or hot key(Ctrl(Cmd)-B, etc.) .env file commenter $\checkmark$ CODE TOOLS 🖷 GitHub $\checkmark$ one month ago A Powerful Java Stri 7,780 🗲 Gradle $\checkmark$ Change Notes OBFUSCATION one year ago 🖷 Groovy $\checkmark$ A8Translate .env syntax highlighting Go support .env.\* files support Javascript indexing optimization(remove /node\_modules from index) 15,184 \*\*\*\* 📲 I18n for Java $\checkmark$ MISC 11 months ago A8Translate 雪 IntelliLang $\checkmark$ 13.775 \*\*\*\* Vendor LANGUAGES 🕞 Java Bytecode Decompiler $\checkmark$ one vear ago Adel Fayzrakhmanov http://adelf.ru adel.faiz@gmail.com Aardvark $\checkmark$ 📲 Java Stream Debugger 11,606 \*\*\*\*\* 🗲 JavaFX $\checkmark$ BUILD 5 years ago Size ABACUS Plugin 909 \*\*\*\*\* Check or uncheck a plugin to enable or disable it. 20C 98.8 K FRAMEWORK INTEGF 7 months ago Install JetBrains plugin... Browse repositories... Ins HTTP Proxy Settings... Manage repositories... <u>C</u>lose 2

#### 4. After the installation is confirmed, restart IntelliJ IDEA to complete installation.

#### Local installation

MaxCompute Studio MaxCompute Studio can also be installed in a local environment.

#### Procedure

1. Go to the MaxCompute Studio plugin page to download the plugin package.

#### 2. Run IntelliJ IDEA.

• If you access IntelliJ IDEA for the first time, a welcome page is displayed. Click Configure and select Plugins from the shortcut menu, as shown in the following figure.



• If you have accessed IntelliJ IDEA before, choose File > Settings > Plugins to enter the same page, as shown in the following figure.

MySecondProjectil [C:\Users\Administrato	r∖.odps.stu	idio\me	ta\serv	ice.odp	s.aliyı	un.com\M	ytes		hoje	mt2] -	- Intel	alij ide	EA (Adr	dministra	rator)											x
File Edit View Navigate Code Analyze	Refactor	Build	Run	Iools	VCS	Window	M	ахСо	mput	ute <u>H</u>	Help															
New ⇒ <u>Open</u> Open <u>Becent</u> Close Project	D ?			\$	-													<b>x</b>	₩ wc	in1.osc	► 1	5	 D <sub>1</sub>	Q	COSI	an in # An
					_																					Build
Den Project Structure Ctrl+Alt+Shift+S Other Settings																										m
Import Settings Export Settings Settings Repository Export Project to Eclipse																										laven Projects
Export to Zip File  Save All Ctrl+S Ctrl+S Synchronize Ctrl+Alt+Y Invalidate Caches / Restart								Sea Proj	irch ject	t Viev	erywł ew Al	here Nt+1	e Dout	ible St	hift											🗲 FAQ Robo
Eprint Associate with File Type Power Save Mode								Go Rec	to F ent	File C	Ctrl+ es Cti	+Shift trl+E	ift+N													~
Egit								Nav Dro	viga' op fi	ation iles h	n Bar here	r Alt+ e to o	+Hom	me												L
2. Structure																										

U Settings			
Q	Plugins		
> Appearance & Behavior	Q,•	Show: All plugins ${\scriptstyle\checkmark}$	
Keymap		Sort by: name 🔻	Android Support
> Editor	🖺 Android Support		
Plugins	📲 Ant Support		Supports development of Android applications with IntelliJ IDEA and Android Studio.
> Version Control	📑 Bytecode Viewer		
> Build, Execution, Deployment	🗲 Copyright		
> Languages & Frameworks	Coverage		
> Tools	CVS Integration		
> MaxCompute Studio	📲 Eclipse Integration		
Vim Emulation	🗲 EditorConfig		
	Git Integration		
	GitHub		
	🖫 Gradle		
	Groovy		
	📲 I18n for Java		
	🖳 IdeaVim		
	🖳 IntelliLang		
	Lava Bytecode Decompiler		
	Lava Stream Debugger		
	LavaEX		
	Check or uncheck a plugin to enable	or disable it.	
	Install JetBrains plugin B	rowse repositories Instal	ll plugin from <u>d</u> isk
0			OK Cancel Apply

## 3. On the Plugins page, click Install plugin from disk…, as shown in the following figure.

Plugins		
Q- Show: All plugins	/	
Sort by:	name 🔻	Android Support
🗲 Android Support		
📲 Ant Support	$\checkmark$	Supports development of Android applications with IntelliJ IDEA and Android Studio.
📲 Bytecode Viewer	$\checkmark$	
📲 Copyright	$\checkmark$	
📲 Coverage	$\checkmark$	
📲 CVS Integration	$\checkmark$	
📲 Eclipse Integration	$\checkmark$	
📲 EditorConfig	$\checkmark$	
📲 Git Integration	$\checkmark$	
📲 GitHub	$\checkmark$	
📲 Gradle	$\checkmark$	
📲 Groovy	$\checkmark$	
📲 I18n for Java	$\checkmark$	
📲 IdeaVim	$\checkmark$	
📲 IntelliLang	$\checkmark$	
📲 Java Bytecode Decompiler	$\checkmark$	
🗲 Java Stream Debugger	$\checkmark$	
/LavaFX		
Check or uncheck a plugin to enable or disable it.		
Install JetBrains plugin Browse repositories	Insta	all plugin from <u>d</u> isk
		OK Cancel Apply
		OK Cancer Apply

4. In the displayed window, click the gray icon before a directory for navigation, find the plugin file, select it, and click OK.

Choose Plugin File
JAR and ZIP archives are accepted
🟫 🛄 🐂 🛤 🗙 😋 😭 Hide path
恵下载∖IntellJ IDEA∖odps-studio-intellij-3.0.2_20181126174744.zip 土
> functionstudio
> ■jar包
> Ksoftware
> KwDownload
> MCStudio截图
> MININT
> mysql
> Program Files
Program Files (x86)
> Users
> eclipse插件
Intell/ IDEA
Dups-studio-intenij-5.0.2_201011201/4/44.2ip
Drag and drop a file into the space above to quickly locate it in the tree
OK Cancel

### 5. Return to the Plugins page and click OK to install the local plugin.

= Settings				8
Q,*	Plugins			Reset
> Appearance & Behavior	Q,T	Show: All plugins $\lor$		
Keymap		Sort by: na	me 🔻	MaxCompute Studio
> Editor	Coverage			C Restart IntelliJ IDEA
Plugins				New version will be available after restart
> Version Control	Eclipse Integration			New version will be available after restart
> Build, Execution, Deployment	EditorConfig			Alibaba Cloud independently. It is a fast and cloud-based big
> Languages & Frameworks	📑 Git Integration			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		$\checkmark$	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			analytics, etc.
	📑 IntelliLang			Features include:
	📲 Java Bytecode Decompiler			MaxCompute SQL language support     MaxCompute function development
	雪 Java Stream Debugger			MaxCompute data management     MaxCompute job management
	📲 JavaFX			Change Notes
	📲 JUnit			Changes in Version 3.0.0:
	📲 Kotlin			New Feature:
	🕒 Maven Integration			<ul> <li>support write authorization statement, show user</li> </ul>
	MaxCompute Studio			privileges and diagnosis auth exception message. • use web view to draw sql pot graph and performance
	Check or uncheck a plugin to enable	e or disable it.		
	Install <u>J</u> etBrains plugin	Browse repositories	Insta	ll plugin from <u>d</u> isk
0				OK Cancel <u>A</u> pply

6. After the installation is complete, a dialog box is displayed, prompting you to restart IntelliJ IDEA. Click Restart.

Search Everywhere Double Shift
IDE and Plugin Updates       Image: Comparison of the second
# 7. After IntelliJ IDEA is restarted, the page is displayed as shown in the following figure.

<b>2</b> 1	Myllimu	a dias	mt2 [Cill]	ison's Adv	whether	of codpo at	alio) me	trioer	vice.ode	e ally	un.com/i/M	SecondProjecti	] - IntelliJ IDEA (Administrator)
File	<u>E</u> dit	View	Navigate	Code	Analyze	Refactor	<u>B</u> uild	Run	Tools	VC <u>S</u>	Window	MaxCompute	Help
ţ	Projec	t Explo	rer						\$	-			
Proj	+ -	-   fj	🖻 🌀	¥   3	ž   (	<u>)</u> () ?							
Ē	> WF	nttp://w	enine and p	a aliyana	api								
8													
plon													
ectE								_					
Proj								-	Tip of t	the Da	y		
w								1	The sho	rtcuts	such as	Ctr1+0 (View	Ouick Documentation), Ctr1+P (View
ler									Parame	eter i	Info), Ctr	1+B (Navigate	e   Declaration) and others can be used not only in
old								1	ne eulo	or bu	c in the co	de completion	popup list as well.
쉉									catch	Unsl	Listiadi	okEccor /inu	
N									60	Unsu	pportedLo	okAndFeelExce	eption (javax.s
										D	ocumentat	ion for Unsatisfi	iedLinkError 🍂
									\$	♦ 1			<b>C</b> < 1.7 > 🔅
									java.l	ang		less Vessbird	
									exter	nds 1	ava.lang.	LinkageError	Inequinkerfor
													P-3 ///
								6	Chow	tine o	n startun		Previous Tip Next Tip Close
									_ onen	0000	in startap		
Idur													
Stru													

#### Next step

Now, you know how to install the MaxCompute Studio plugin. Continue to the next tutorial. In the tutorial, you will learn how to configure a MaxCompute project connection to manage data and resources. For more information, see *Create a MaxCompute project connection*.

### 2.3.3 View and upgrade the version

View the MaxCompute Studio version

Perform the following steps to view the Studio version:

- 2. Select Plugins on the left bar of the dialog box and search for MaxCompute Studio.
- 3. View the MaxCompute Studio version number and release information.

Alternatively, you can select MaxCompute Studio on the left bar of the Settings page to view the current version number.

#### Check new versions

By default, MaxCompute Studio automatically detects new versions. If a new version is available, MaxCompute Studio automatically notifies you.

After receiving an update prompt, you can select:

- Install: Click the Install link in the update prompt. The new version is automatically downloaded and installed. After the installation is complete, restart IntelliJ IDEA.
- Configure: Click the Configure link in the update prompt. You can configure whether to detect new versions automatically.

If the automatic update function is disabled, you can perform the following steps to check and install a new version of MaxCompute Studio.

- 2. Select MaxCompute Studio on the left bar of the dialog box.
- 3. On the MaxCompute Studio configuration page, click Check new versions.
- 4. If a new available version is detected, Studio notifies you of the new version number. Click Install new version and restart IntelliJ IDEA to complete installation.

You can use the Automatically checks for new version check box to control the switch for automatic version update check.

#### Next step

Create a MaxCompute project connection

### 2.4 Manage data and resources

# 2.4.1 View tables and UDF

View tables and functionsView tables and functions

In the Project Explorer window, you can view tables, functions, and resources with connections added. For tables and functions to be viewed in the Project Explorer window, the MaxCompute project connections must be added, for more information, see *Add MaxCompute project connections*.

#### Browse tables and functions

- To browse tables and functions in the project space, follow these steps.
- 1. Open the Project Explorer window and you can view the added Project node tree.



The toolbar is displayed at the top of the node tree, and includes:

- · Add Project: Adds a connection to the MaxCompute project space.
- Delete Project: Deletes a connection from Project Explorer, which has no impact on the project space on the server end.
- Update Metadata: Updates metadata information from the project space on the server end and updates the locally buffered metadata.
- Expand Node: Expands all tree nodes.
- Fold Node: Folds all tree nodes.
- User Feedback: Submits user feedback.
- · Online Documentation: Opens online documents.
- 2. Double-click the Tables node or click the drop-down arrow to expand the Tables node to list all tables in the project (including virtual views). The table name list serves the same purpose as the show tables command. You must have the List Table permission in the project. The methods for the Functions and Resources nodes are similar to that of the Tables node.

P odps_studio_dev
Tables & Views
alitrip_dest_relation
array_map_test2
haoping_sdk_log_copy
haoping_test_part
hy_test1
🕨 🏢 key
table test
▶ <u>□ Functions</u>
▶ Resources

3. MaxCompute Studio downloads project metadata on the server to the local device . When metadata on the server end is updated, for example, a new table is added,

you must manually trigger a refresh to reload changed metadata to the local device

. The refresh can be performed at the Project or Table level. The procedure is as follows:

- a. Select a node.
- b. Click the Refresh icon on the toolbar or right-click the node and select Refresh meta.



View table details

You can view data table information in Table Details View of MaxCompute Studio.

1. In the node tree, expand a table node to view the column name and type.



2. Double-click a table or right-click a table and choose Show Table Detail to view the table details. The table details include metadata, such as owner, size, and column, table structure information, and data preview.

- +j 主 🚖	Table information		Table	schema				
meta dev	(key	value		name	type		comments	
a 0101020223 storontrol 395	Name:	b_hour	hour		STRING			
studiotest1 test1	Comments:		c		STRING			
Tables & Views	Owner:	1830820614534790	d		STRING			
Tables & views	Created time:	2016-02-18 15:20:46	e		STRING			
▼bnour	Meta Modified time:	2016-04-18 16:48:18	f		STRING			
hour :STRING	Data Modified time:	2016-08-05 10:18:33	9		STRING			
c :STRING	Size:	63048	h		STRING			
d :STRING	Physical size:	189144						
e :STRING	Partitions:	0						
f :STRING	Charding							
E a STRING	Sharding columns:							
	sharding cordinits.							
E H STRING	Data and fair							
P III D_second	Data preview							_
P III DIG	nour	c		d	e		9	
▶ m big2	00		Mall	1001	1011	null	null	
⊨ III cc	01		null	India India	Inuit	NUI	nui	
▶ 🋄 dual	02		Tripine .	and the second s	and the second s	THE REAL PROPERTY AND ADDRESS OF THE REAL PROPERTY	THE REAL PROPERTY AND ADDRESS OF	
Image: haoping_aaa2	04		Sull.			ault		
haoping test4	05		2011	12011	12011	ault		
b ■ haoping test part	06		17711	177111	1500 U	avil 1	I THE REAL PROPERTY IN THE REAL PROPERTY INTERNAL	
b m hy test	07		null	null	null	null	null	
b III liming testake	08		null	Inull	Inull	Inull	null	
P miniao_testabe	09		null	null	null	null	null	
mydual	10		null	Inull	Inull	Inull	null	
skew	11		null	Inull	Inull	null	Inull	
▶	12		null	Inull	null	null	null	
▶ III t_view	13		null	Inull	null	null	null	
III t vuvuan	14		null	Inull	null	null	null	
▶ III test1	15		null	null	null	null	null	
tast2	16		Inull	Inuli	Inull	null	null	
r III wate and	17		Inuli	mult	Inull	null	null	
munit_test_sak	18		null	Inull	null	null	null	

3. Right-click Tables & Views and select Open specific entity to display the details of the specific table. Note that the complete table name must be specified. If you do not have the List permission on the project and only have the permission on a specific table, you can also view details of the table using this method. The methods for the Functions and Resources nodes are similar to that of the Tables node.

MaxCompute Project Explorer
• - ↓ ₹ ₹ 0 0 ?
P odps_studio_dev
▶ Tabl Add project
Resc Refresh meta
<ul> <li>▶ P wpp_dg Open specific entity</li> <li>▶ P cdo_meta</li> <li>▶ P studiotest1_test1</li> </ul>
Open specific table
Name:
ОК Сапсеі

Intellij IDEA supports searching by default. After a table is expanded, you can directly press keys on the keyboard to perform fuzzy match.

4. MaxCompute Studio also supports quick search for the table, you can use the shortcut key (Windows: Ctrl + Alt + Shift + N, macOS: # + 光 + O) to call the navigation bar, then enter the name of the table and press Enter.

Enter symbol name: Include non-project symbo	Is (て第O) 🍸 💉
Q hy	8
hy_partition_test odps_studio_dev	
<pre>hy_test odps_studio_dev</pre>	
■ hy_view odps_studio_dev	년区 yq.aliyun.com

You can narrow the search by using the pre-keyword (table:, function:, or resource :). For example, to search for the function count, enter function:count.

5. To know the scripts in which the table is used, right-click the table and select Find usages.



View function details

1. Expand a function node under the UserDefined node of Functions to display the method signature of this function. Double-click a function node under the Functions node. Alternatively, double-click the source code resource of the function under the Resources node. In this case, codes of this function are displayed.

J,	MaxCompute Project Explorer — — — — — — — — — — — — — — — — — —	C JavaMethodFunction.java × 📴 PlusHello.class ×	
nu.	+ - 1) 포 축	Decompiled .class file, bytecode version: 50.0 (Java 6)	
די מו חרוחו ב 🔲 🔻	<ul> <li>P meta_dev</li> <li>P a_0101020223_stgcontrol_395</li> <li>P studiotest1_test1</li> <li>Tables &amp; Views</li> <li>Functions</li> <li>fx divideerror</li> <li>fx myudtf</li> </ul>	<pre>1</pre>	DEA
ri ujeci Expiri el	<ul> <li><i>fx</i> plushello</li> <li><i>fx</i> plushello(STRING):STRING</li> <li><i>fx</i> timestwo</li> <li><i>fx</i> Resources</li> </ul>	<pre>10 public class PlusHello extends UDF { 11     public PlusHello() { 12     } 13 14     public String evaluate(String a) { return "hello" + a, 17     } 18</pre>	; }

# Note:

The Java code is obtained by decompiling JAR, which is not the source code. To enable the Python UDF to parse the signature, install PyODPS (MaxCompute Python SDK) first. Install *pip* : *sudo* / *usr* / *bin* / *python get* - *pip* . *py* (Download *get* - *pip* . *py* from Google manually) and then *PyODPS* : *sudo* / *usr* / *bin* / *python* - *m pip install PyODPS* . Note that the Mac operating system has Python, which is stored in /usr/bin/python. Install PyODPS in this directory.

2. The classification under the BuiltIn node of Functions shows the built-in functions of the system, expand it to display signature and double-click it to display function document.

# 2.4.2 Import and export data

MaxCompute Studio can import local data files in CSV or TSV format to MaxCompute tables and export MaxCompute table data to local files. MaxCompute Studio completes data import and export by using Batch data Tunnel provided by the MaxCompute platform.

Usage instructions

- The MaxCompute Tunnel service must be used for data import and export.
   Therefore, the MaxCompute project added in Studio must be configured with the Tunnel service.
- · Related permissions must be granted for table import and export.

#### Import data

1. Open the Project Explorer window, right-click a table name or a field attribute in Data preview of Table details and select Import Data Into Table.

Uccal Settings [C:\User	rs\Administ	trator\Loc	al Settings]	- Intel	IJ IDE	А		
<u>F</u> ile <u>E</u> dit <u>V</u> iew <u>N</u> aviga	<u>File Edit View N</u> avigate <u>C</u> ode Analyze <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						<b>*</b> - I+		
+ - 🗊 🖬 😏	¥ 主 🗄	ŧ 🕘 🤉	J ?					
M http://service.odps          MySecondProje         Tables & Vie         Image: Tables & Vie <td< td=""><td>aliyun.com, ct2 ews a _new info_d Q Find u S User F I Refres I Show Gener Gener Gener Trunca Drop f Impor</td><td>(api sages Privilege h table m table deta table deta table edit ate select ate DDL s ate table c table from t data into t data from</td><td>eta il or statement tatement lata o server o table m table</td><td></td><td></td><td></td><td></td><td></td></td<>	aliyun.com, ct2 ews a _new info_d Q Find u S User F I Refres I Show Gener Gener Gener Trunca Drop f Impor	(api sages Privilege h table m table deta table deta table edit ate select ate DDL s ate table c table from t data into t data from	eta il or statement tatement lata o server o table m table					

2. In the Import Data dialog box that appears, select the path of the imported data file, column separator, size limitation, and number of lines for an error tolerance, and click OK .



3. If Import Data Success is displayed, data import is successful and imported data can be viewed in the table.

#### Export data

- 1. Two methods are provided for table data export.
  - Right-click a table name and select Export Data From Table.



• Right-click a field attribute in Data Preview of Table details and select Export Data From Table.

key     value       Name:     hy_test1       Comments:     ALIVUN\$dsp_06894684@aliyun.com       Created time:     2016-12-29 11:50:22       Meta Modified time:     2016-12-29 11:50:22       Data Modified time:     0       Physical size:     3840       Partition columns     0       Sharding:     0       Sharding:     0       Sharding:     0       Sharding:     1	key Jame: Comments: Swner: Treated time: Aeta Modified time:	value hy_test1 ALIVUN\$dxp_06894684@aliyun.com 2016-12-29 11:50:22 2016-12-29 11:50:22	id name age	type BIGINT STRING PIGINT	comments
Name:       hy_test1         Comments:       ALIVUN\$Sdxp 06894684@aliyun.com         Owner:       ALIVUN\$5dxp 06894684@aliyun.com         Created time:       2016-12-29 11:50:22         Data Modified time:       2017-01-22 0:42:37         Size:       1280         Physical size:       3840         Partitions:       0         Partitions:       0         Sharding:	lame: comments: Dwner: reated time: Aeta Modified time:	hy_test1 ALIYUN\$dxp_06894684@aliyun.com 2016-12-29 11:50:22 2016-12-29 11:50:22	id name age	BIGINT STRING PIGINT	
Comments: ALIYUN\$dxp_06894684@aliyun.com Created time: 2016-12-29 11:50:22 Data Modified time: 2017-01-22 20:42:37 Size: 1280 Physical size: 3840 Partition columns Sharding columns: Sharding columns: Alignment of the second sec	Comments: Dwner: Created time: Aeta Modified time:	ALIYUN\$dxp_06894684@aliyun.com 2016-12-29 11:50:22 2016-12-29 11:50:22	name age	STRING	
Owner: ALIYUN\$dxp_06894684@aliyun.com Created time: 2016-12-29 11:50:22 Meta Modified time: 2016-12-29 11:50:22 Data Modified time: 2017-01-22 20:42:37 Size: 1280 Physical size: 3840 Partition: 0 Partition columns Sharding: 0 Sharding: 4 Data preview id a Export data from table 1 Import data into table 1 Import data into table 1	Owner: Created time: Aeta Modified time:	ALIYUN\$dxp_06894684@aliyun.com 2016-12-29 11:50:22 2016-12-29 11:50:22	age	DICINIT	
Created time: 2016-12-29 11:50:22 Data Modified time: 2017-01-22 20:42:37 Size: 1280 Physical size: 33840 Partitions: 0 Partition columns Sharding: Columns:	reated time: Aeta Modified time:	2016-12-29 11:50:22	8	DIGINI	
Meta Modified time: 2016-12-29 11:50:22   Data Modified time: 2017-01-22 20:42:37   Size: 1280   Physical size: 3840   Partitions: 0   Partition columns 1     Sharding: 1     Ata preview a     id a   Import data from table 11	Aeta Modified time:	2016-12-29 11:50:22	B		
Data Modified time: 2017-01-22 20:42:37 Physical size: 1280 Physical size: 3840 Partition columns Partition columns Sharding: Sharding: Sharding: A start of the		2010 12 23 11.30.22			
Size: 1280 Physical Size: 3840 Partitions: 0 Partition columns Sharding: Sharding: Sharding columns:	ata woolfied time:	2017-01-22 20:42:37			
Physical size: 3840 Partitions: 0 Partition columns Partition columns Sharding:  Shardin	ize:	1280			
Partition columns Sharding: Sharding columns: a Export data from table a Import data into table 11 12 13 14 14 14 15 15 15 15 15 15 15 15 15 15	hysical size:	3840			
Artition columns	artitions:	0			
Sharding: Sharding columns:	artition columns				
Sharding columns:	harding:				
id age age and a more that a from table age and a more that a from table age age and a more that a into table age age and a more that a into table age age and a more that a into table age age and a more that a into table age age and a more that a	harding columns:				
a Import data into table 1	ata preview id	a Export data from tab	le <sup>1e</sup>	11	age
		a Import data into tabl	A	1	
		Import data into tabi	-		

2. In the Export Data dialog box that appears, select the path for saving the exported data file, column separator, size limitation, and number of lines for an error tolerance, and click OK.

• - ↓ = ₹ (i) (i) ?	
P odps_studio_dev	
Tables & Views	
alitrip_dest_relation	
array_map_test2	
haoping_sdk_log_copy	
haoping_test_part	
▶ <mark>IIII hy test1</mark>	
ke Add project	
tal tal Refresh table meta	
Functi Show table details	
► fx adc	
fx ex2 Delete table from server	
fx hac Generate select statement	
▶ <i>f×</i> hac	
▶ fx hell import data into table	
fx par Export data from table	
TY parcoulsors tastic	

3. If Export Data Success is displayed, data export is successful and exported data can be viewed in the target file.

You can also right-click Data Preview of Table and choose Export Grid Data to export data.

≜hy_test1 ×								
Table information		🚦 Tab	Table schema					
key		value		name	type	comments		
Name:	hy_test1		id		BIGINT			
Comments:		nam	e	STRING				
Owner:	ALIYUN\$dxp_06894	1684@aliyun.com	age		BIGINT			
Created time:	2016-12-29 11:50:2	2						
Meta Modified time:	2016-12-29 11:50:2	2						
Data Modified time:	2017-01-22 20:42:3	7						
Size:	1280							
Physical size:	3840		- 8					
Partitions:	0		- 8					
Partition columns								
Sharding:								
sharding columns:			- 3					
Data preview								
id		r	ame			age		
1		a			11			
1		a Copy cell content Export grid data View cell text			1			



Note:

The Data Preview function in Data preview is used only to export data displayed in Data sample instead of all data in the table.

#### New Type Import Export

Simply generate text in the agreed format and store it in CSV or TSV format, you can import to table through studio.

Conversion rules for each data type are described in detail below.

#### **Basic Type**

- 1. Tinyint, smallint, Int, bigint is stored directly as an integer string, and the numeric value exceeds the type boundary is reported as wrong
- Float, double to store fractional strings or floating-point forms, such as: 2.342 1x +
- 3. Varchar is stored directly as a string, which is automatically truncated above the upper limit, and no errors will be reported
- 4. String is stored directly as a string
- 5. Decimal string that supports shaping or floating-point
- 6. Binary needs to encode binary data to base64 string
- 7. Datetime date time requires that the format specified in import dialog is consistent and that the format mismatches will be reported incorrectly
- 8. Timestamp timestamp needs to follow yyyy-[m] M-[d] d hh \: mm \: ss [. f...] Format is stored as a string
- 9. Boolean true or false string?

**Composite Type** 

- 1. Array needs to be stored as a JSON array. The array elements are converted to strings according to the rules agreed in this article, array elements support any type.
- 2. Map needs to be stored as a JSON object, and map key, value are converted to a string according to the rules agreed in this article, value supports any type of nesting.
- 3. Struct needs to be stored as a JSON object, the struct field name is string, and the key converted to a JSON object, strect field values converted to JSON The value of the object, the value of the field that defines the rule transformation in this article.

Example

Array type

For Table structures as follows:

Column name	Column data types
c_1	ARRAY <tinyint></tinyint>
c_2	ARRAY <int></int>
c_3	ARRAY <float></float>

Column name	Column data types
c_4	ARRAY <datetime></datetime>
c_6	ARRAY <timestamp></timestamp>
c_7	ARRAY <string></string>

You can import data in the CSV format shown below:



The CSV format needs to escape double quotes, which are expressed by two double quotes, you can refer specifically to the CSV format specification.

Map Type

For Table structures as follows:

Column name	Column data types
c_1	MAP <tinyint,string></tinyint,string>
c_2	MAP <string,int></string,int>
c_3	MAP <float,string></float,string>
c_4	MAP <string,datetime></string,datetime>
c_5	MAP <string,string></string,string>
c_6	MAP <timestamp,string></timestamp,string>

You can import data in the CSV format shown below:

#### Struct Type

For Table structures as follows:

Column name	Column data types
C_struct	<ruct<x:int,y:varchar(256),z: STRUCT<a:tinyint,b:string>&gt;</a:tinyint,b:string></ruct<x:int,y:varchar(256),z: 

You can import data in the CSV format shown below:

```
c_struct
"{"" x "":"" 1000 "","" y "":"" varchar_te st "","" z "":{"" a "":""
123 "","" b "":"" stringdemo ""}}"
"{"" x "":"" 1000 "","" y "":"" varchar_te st "","" z "":{"" a "":""
123 "","" b "":"" stringdemo ""}}"
```

### 2.4.3 Visualization of operating the tables

The Project Explorer of MaxCompute Studio provides the visualized table structure editor used to create and modify tables.

Visualization of creating a table

Procedure

1. Right-click the project that you want to create the table, and select create a new table.



2. In the dialog box that appears, enter a table name and column information. Click Generate CreateTable Statement generates the corresponding pant statement, click Execute to execute the build table.



When you set the table name, column name, type, and lifecycle, observe the related requirements of MaxCompute. For more information, see*Table operations*.

3. After the table is created, view the table metadata in table&view of the Project Explorer. If no metadata is displayed, refresh the list.

Visualization of modifying a table

Procedure

1. Intable&view of the Project Explorer, right-click the expected table and select Open table Editor:

Project Explorer		<b>☆</b> - I=
+ - 🗘 🗉 💲 🕷	፻ 🗧 🙆 🗊 ?	
MySecondProject2          MySecondProject2         Tables & Views         bank_data         emp_test_ne         ods_user_inf         result_tabl         result_table         table_1         wc_in1	<ul> <li>Q Find usages</li> <li>S User Privilege</li> <li>I Refresh table meta</li> <li>Show table detail</li> <li>Open table editor</li> </ul>	
> 🕞 Functions > 💽 Resources	Generate select statement Generate DDL statement Truncate table data <ul> <li>Drop table from server</li> </ul> <li>Import data into table</li> <li>Export data from table</li>	

2. In the dialog box that appears, edit the table. You can modify the table comments, table lifecycle, column name, and column description, and add columns. Specific

#### rules follow the MaxCompute table requirements and can be found in Table

operations.

長名:	table_2		□转为存档		
i释:					
上命周期:	-1		B		
劉塘列   分区	(列 索引 附)	加厲性			
na	ime	type	not null	comment	t +
col1		BIGINT			
col2		BIGINT			
col4		BIGINT			<b>↑</b>
col3		BIGINT			+
			11111		
SQL DDL SC	RIPT		2000		
	RIPT TER TABLE t	able_2 ADD COLUMNS (			
	RIPT TER TABLE t col5' BIGIN	able_2 ADD COLUMNS ( T):			0
SQL DDL SC	RIPT TER TABLE t col5 BIGIN	able_2 ADD COLUMNS ( T);			<u> </u>
SQL DDL SC 1 AL 2 AL 3	RIPT TER TABLE t col5 BIGIN	able_2 ADD COLUMNS ( T):	22222		9
SQL DDL SC	RIPT TER TABLE t col5 BIGIN	able_2 ADD COLUMNS ( T):			9
SQL DDL SC 1 AL 2 3	RIPT TER TABLE t col5 BIGIN	able_2 ADD COLUMNS ( T):			<u> </u>
SQL DDL SC 1 AL 2 I	RIPT TER TABLE t col5 BIGIN	able_2 ADD COLUMNS ( T):	22222		
SQL DDL SC	RIPT TER TABLE t col5 BIGIN	able_2 ADD COLUMNS ( T):			9
SQL DDL SC	RIPT TER TABLE t col5 BIGIN	able_2 ADD COLUMNS ( T):			<u> </u>

3. After completing the modifications, click Alter Table Statement Generate a specific alter statement and click Executeto perform the table modify operation. After successful execution, view the table metadata.

Visualization of deleting a table

In table&view of theProject Explorer, right-click the expected table and select Drop table from server:



Select OKin the bullet box to remove the table from the MaxCompute service.

# 2.5 Develop SQL procedure

### 2.5.1 Create MaxCompute Script module

Before developing MaxCompute Script, you need to create a MaxCompute Script module in either of the following scenarios:

No script file exists locally

If no script file exists locally, you can use Intellij IDEA to create a new module.

Procedure

1. Open or create a MaxCompute Studio project. This article uses creating a project as an example. Click File in the menu and select New > Project . as shown in the following figure.

IJ	Local Se	ettings	[C:\Users\A	Administ	trator\Loc	al Settings]	] - C:\U	sers∖A	dminist	rator\.	IdeaIC
Eile	<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> ind
	New				>	<u>P</u> roject	i				
	Open					Project	from E	Existing	g Source	es	
	Open <u>F</u>	Recent			>	Project	from \	/ersio	n Contro	ol	>
	Close F	Project				<u>M</u> odul	e				
Ŷ	Se <u>t</u> ting	s		Ct	rl+Alt+S	Modul	e from	Existir	ng Sourd	es	
Ē.	Project	Struct	ure (	Ctrl+Alt-	+Shift+S	揚 Scratch	n File	Ctr	·l+Alt+S	hift+In	sert
	Other S	Setting	s		>	븕 FXML F	File				
	Import	Setting	gs		,					5	
	<u>E</u> xport	Setting	js							6	
	Export	to Ecli	pse							7	ę
	Export	to Zip	File							8	
	Setting	s Repo	sitory							9	
H	<u>Save</u> A	I			Ctrl+S						
9	Synchro	onize		Ct	rl+Alt+Y						
	Invalida	ate Cao	ches / Resta	art							
Ð	<u>P</u> rint										
	Power	Save N	1ode								
	E <u>x</u> it										

### 2. Select MaxCompute Studio on the left-side navigation pane, and click Next.

먼 New Project		8		
Java	Additional Libraries and <u>F</u> rameworks:			
Java FX				
🖷 Android				
IntelliJ Platform Plugin				
<i>M</i> Maven	Nothing to show			
📀 Gradle				
G Groovy				
🕐 Griffon				
Kotlin				
Empty Project				
<mark> MaxCompute Java</mark>				
🔥 MaxCompute Studio				
	Previous <u>N</u> ext Cancel Help			

3. Enter the project name, and click Finish.

Project	23
Project name: MySQLProject	
Project location: C:\Users\Administrator\IdeaProjects\MySQLProject	
More Settings	
Previous <u>F</u> inish Cancel Help	



If a project has been opened before, a dialog box appears, prompting whether to open the new project in the existing window (closing the previous project). Click This Window.



4. After the project is created, the page shown in the following figure appears. You can develop SQL scripts in the project.



#### Script files exist locally

If many scripts have been stored in a local folder, MaxCompute Studio is used to edit the scripts. You can open a module directly.

#### Procedure

- 1. Create a connection configuration file *odps\_config.ini* for MaxCompute in the scripts folder, and configure authentication information for connecting to MaxCompute.
  - project\_name=xxxxxxxx
  - · access\_id=xxxxxxxxx
  - access\_key=xxxxxxxx
  - end\_point=xxxxxxxxx

2. Open IntelliJ IDEA, select File > > Open, and select the scripts folder.



3. MaxCompute Studio detects whether the odps\_config.ini file exists in the folder, captures metadata on the server based on the configuration information in the file, and compiles all scripts in the folder.

# 2.5.2 Write SQL

Table operations

# 2.5.3 Submit SQL scripts

MaxCompute Studio directly submits MaxCompute SQL scripts to the server for running, and displays detailed information about the query result and execution plan. Before submission, MaxCompute Studio compiles scripts to effectively prevent compilation errors that are detected after the scripts are submitted to the server.

#### Prerequisite

<sup>·</sup> Create a MaxCompute project connection and bind it to the target project.

- · Create a MaxCompute Studio module.
- Before submission, perform setting as required. MaxCompute Studio provides various setting features. You can perform quick setting on the toolbar at the top of the editor page. The following three types of setting can be performed:
  - Compiler Mode: It can be set to Script Mode or Statement Mode.
    - In statement mode, scripts are separated by ; and submitted to the server one by one.
    - The script mode is newly developed. A whole script can be submitted to the server immediately. The server provides overall optimization, which is more efficient. Therefore, this mode is recommended.
  - Type System: It mainly solves the compatibility problem of SQL statements, which can be set to the following values:
    - Legacy TypeSystem: Indicates the type system of original MaxCompute.
    - MaxCompute TypeSystem: Indicates the new type system introduced by MaxCompute 2.0.
    - Hive Compatible TypeSystem: Indicates the type system in Hive compatibility mode introduced by MaxCompute 2.0.
  - Compiler Version: MaxCompute Studio provides the stable compiler and experimental compiler.
    - Default Version: Indicates the stable version.
    - Flighting Version: Includes the latest features of the compiler.

# Note:

You can use Global Settings to set the submitted scripts. Select File > > Settings > > MaxCompute , select MaxCompute SQL, and choose Compiler > > Submit to set the preceding attributes.

#### Submit SQL scripts

The top toolbar of the editor provides the Synchronize and Compile and Submit features.

```
* ** Synchroniz e **: Updates
                            metadata
                                     in
                                         SQL
                                               scripts
including table names and UDFs.
                                    If
                                         MaxCompute
                                                    Studio
                                                    found ,
  prompts that a table or function
                                        cannot be
but the table or function obviously exists on
                                                    the
server , you
              can use
                       this
                              function to
                                           update
                                                   metadata
```

```
Submit **: SOL
* ** Compile and
                                    scripts
                                             are
                                                  compiled
                                        compliance
or
     submitted to
                     the
                          server
                                   in
                                                   with
                                                          pre -
 released MaxCompute
                       SQL
                           rules .
                                     Details
                                              of
                                                   compilatio n
                displayed
                           in the ** MaxCompute
                                                   Compiler **
  errors
           are
window .
```

#### Procedure

- After SQL statements are compiled, click the green running icon on the toolbar, or right-click Script Editor and select Run MaxCompute SQL Script to submit the SQL statement to the server. If a variable exists in the SQL statement (such as \${bizdate} in the following figure), a dialog box is displayed, prompting you to enter the variable value.
- 2. The script will be locally compiled (depending on the project metadata you added in the Project Explorer window). If no compilation error exists, the script is submitted to the server for execution. When the SQL script is being executed, the running logs are displayed. If the script is running on the server, the Job Details page is displayed, showing the basic information about job running and the execution diagram.
- 3. You can view SQL results on the Results page. If there are multiple statements in the single-sentence mode, the result of each statement is displayed. You can select rows or columns in the table, and copy them to the Clipboard.

### 2.6 Developing Java

### 2.6.1 Create MaxCompute Java Module

MaxCompute Studio supports Java user-defined function (UDF) and MapReduce development. First, a MaxCompute Java module must be created.

#### Create a module

ChooseFile > New > Module, set the module type to MaxCompute Java, and configure Java JDK. Click Next, enter a module name, and click Finish. MaxCompute Studio automatically creates a Maven module and introduces MaxCompute dependencies.

#### Module structure

So far, a module for developing a MaxCompute Java program has been established, that is the mDev shown in the following figure. Its main directories include:

- src/main/java: Source code for Java program development.
- examples: Sample code, including unit test (UT) examples. You can see the examples to develop or compile UT.
- warehouse: Schema and data required for running locally.



# 2.6.2 Develop and debug UDF

Once the MaxCompute Java Module has been created, udfs can be developed.

#### Procedure

1. Expand the MaxCompute Java Module Directory that you created and navigate to src > main > java > new, and click MaxCompute Java as shown in the following figure.



2. Set Name and Kind, and click OK. as shown in the following figure.



- Name: Specifies the name of the MaxCompute Java Class. If you have not created a package, you can enter packagename.classname to automatically create a package.
- Kind: Specifies the type. Supported types include custom functions (UDF/UDAF/ UDTF), MapReduce (Driver/Mapper/Reducer), and non-structural development ( StorageHandler/Extractor).
- 3. After the creation is successful, the Java program can be developed, modified, and tested.





Here's a code template that can be customized in Intellij. You can define it in Preference > Editor > File > Code Templates. Then look for the corresponding template in the Code tab.

For detailed development steps, seeJAVA UDF development.

Normally, the development of JAVA UDF can be done in the following ways:

· Use MaxCompute Studio to complete the whole process of JAVA UDF development.

• Use Develop and debug JAVA UDF using the Eclipse plug-in, export the Jar package, then Add resources through commands or DataWorks, and Register the function.

For detailed development steps, seeJAVA UDF Development.

#### Debug the UDF program

After the UDF program is developed, it can be tested using unit test (UT) or local running to check whether it meets expectations.

#### **Unit Testing**

There are various UT examples in the examples directory and you can refer to them to compile your UT.



IDE and Plugin Updates: IntelliJ IDEA is ready to update. (today 16:45)

#### **Run locally**

During local running of the UDF program, the running data source must be specified. The following two methods are provided to set the test data source:

- MaxCompute Studio uses the Tunnel Service to automatically download table data of a specific project to the warehouse directory.
- The mock project and table data are provided. You can see example\_project in warehouse to set it by yourself.

#### Procedure

1. Right-click UDF Class and select Run UDF class.main(). The Run Configuration dialog box is displayed. In normal cases, UDF/UDAF/UDTF data is used as columns in tables of a select sub-statement. The MaxCompute project, table, and column need to be configured. (The metadata is from the mock project under project explorer and warehouse.) Debugging for complex types is also supported, as shown in the following figure:

Project Explorer ♣- I+-	MyFi	rstSQL.osql × 🥃 MyFirstUDF.java ×	C UDTFTest.java ×		
+ -  ĵ 🖩 § 🕷  Ξ 😤  »	Project	SDK is not defined			
<ul> <li>▼ http://service.odps.aliyun.com/api</li> <li>♥ MySecondProject2</li> <li>♥ Tables &amp; Views</li> <li>&gt; ■ bank_data</li> <li>&gt; ■ emp_test_new</li> <li>&gt; ■ ods_user_info_d</li> </ul>	10 11 12 13   14	<pre>import org.junit.BeforeClass; import org.junit.Test; import java.util.List; @Resolve ({"array{struct{x:bigin}}})</pre>	t,y;bigint≫ =>bigint*)		
> III result_tabl	15 🕨	<pre>public class UDIFTest UDIF{</pre>			
> table_1	16	Run/Debug Configurations			<b>—</b> ×
>     table_2 >     table_3 >     table_4	18 19 20	+ - ≣ ♀ ↑ ↓ »	Name: UDTFTest		tance only
v table_6	21	✓ ¼MaxCompute Java	Wain <u>c</u> lass:	com.anyun.oops.examples.udi.test.ooTrTest	
col1 :ARRAY <struct<x:< td=""><td>22</td><td>WUDTFTest</td><td><u>V</u>M options:</td><td></td><td>×7</td></struct<x:<>	22	WUDTFTest	<u>V</u> M options:		×7
	24	> % Defaults	Program arguments:		κ <sup>2</sup>
✓ III table_7	25		Working directory:		~
col1 :BIGINT col2 :ARRAY <bigint></bigint>	27		Environment variables:		
>     table_8 >     wc_in1 >     win_c >     Functions	28 29 30 31		Use classpath of m <u>o</u> dule:	■ MyFristModule □ Include dependencies with "Provided" scope	~
> 🚡 Resources	32		JRE:	Default ( <no jre=""> - module not specified)</no>	×
	34		Shorten command line:	user-local default: none - java [options] classname [args]	~
	35 🕨 36		<u>Enable capturing form</u>	snapshots (requires JRE 5.0 or higher)	
	37		*MaxCompute project: htt	tp://service.odps.aliyun.com/api \vee MySecondProject2 🗸	+
	39		*MaxCompute table: wo	c_in1	~
	40		*Table columns: co	11	ie:c1,c2
	42		Download Record limit:	100	

#### 2. Click OK.

Project •	⊕ ≑ ♦- ⊩	A hy_test × Studiotest1.osql × C PlusHello.java × C UDFTest.java ×
Dut		<pre>package myudf;</pre>
Scripts		
🕨 🛅 target		<pre>import com.aliyun.odps.udf.UDF;</pre>
🔻 🛅 warehouse		<pre>public class PlusHello extends UDF {</pre>
🔻 🛅 meta_dev		// TODO define parameters and return type, e.g: public S
tables		<pre>public String evaluate(String s) {</pre>
🔻 🗖 hy_test		return "hello world:" + s;
📄schema		<pre>public String evaluate(Long a, String b) {</pre>
🗋 data		<b>return</b> "hello world2:" + a + " " + b;
🔓 studioTest.iml		<pre></pre>
External Libraries		3
Run 🐪 PlusHello		
▶ ⊛ 🗐 ↓⅔ ५☴ 至 😤 ♠ ∔ ᠉		
Test framework quit unexpectedly	/Library/Java/	<pre>JavaVirtualMachines/jdk1.8.0_65.jdk/Contents/Home/bin/java</pre>
0	Testing started	<pre>d at 15:02 write table scheme : meta dev by test&gt;/lisers/liuvi/IdeaProjects.</pre>
	[INF0]generate	e schema file: /Users/liuyi/IdeaProjects/studioTest/warehouse/meta_
	[INF0]Finished	d to write table scheme : meta_dev.hy_test>/Users/liuyi/IdeaProje
	[INFO]Start to	<pre>o download table: 'meta_dev.hy_test', download mode:AUTO DeveloadSection ID is . 20160220160200402572001854c05</pre>
	[INFO]Start to	o write table: meta dev.hv test>/Users/liuvi/IdeaProjects/studio
25	[INF0]Finished	d write table: meta_dev.hy_test>/Users/liuyi/IdeaProjects/studio
×	hello world2:1	La
2	hello world2:2	b
•	netto wortuzio	c,u
	Process finishe	ed with exit code 0

### Note:

- If the table data under the specified project is not downloaded into glashourse
   , You need to download the data first, default download 100 entries. If more
   data is required, use the Tunnel Command of the console or table downloading
   function of Studio.
- If the mock project is used or the table data is downloaded, directly run the program.
- The UDF local run framework uses data in specific columns in warehouse as the UDF input and run the UDF program locally. You can view log output and result display on the console.

#### Local warehouse directory

The local warehouse directory is used to store tables (including meta and data) or resources for local UDF or MR running. The following figure shows the warehouse directory.





#### Note:

- The project name, tables, table name, table scheme, and sample data are under the warehouse directory in sequence.
- The schema file is configured with the project name, table name, and column name and type (separated using a colon) in sequence. For a partition table, the partition column also needs to be configured. (For a non-partition table, refer to wc\_in1. For a partition table, refer to wc\_in2).
- $\cdot~$  The data file uses the standard CSV format to store table sample data.
  - Special characters include comma, double quotation marks, and line feed (\n or rn).
  - The column separator is comma and the line separator is \n or \r\n.
  - If the column content includes special characters, double quotation marks (") must be added before and after the column content. For example, if the column content is 3,No, it is changed to "3, No".
  - If the column content includes double quotation marks, each double quotation mark is converted to two double quotation marks. For example, if the column content is a" b" c, it is changed to "a" " b" " c".
  - \N indicates that a column is null. If the column content (string type) is \N, it must be converted to """ \N""".
  - The file character code is UTF-8.

# 2.6.3 Develop MapReduce

After the Create MaxCompute Java Module is created, MR can be developed.

#### Develop the MR program

- 1. Right-click the module source code directory src > main, select New > java, and select MaxCompute Java.
- 2. Create Driver, Mapper, and Reducer.

<u>N</u> ame:	↑↓
Kind: Oriver	~
ОК	Cancel

3. Set the input/output table and Mapper/Reducer class. The framework code is automatically filled in the template.

C He	C HelloDriver.java ×					
	HelloDriver					
1		<pre>package mymr.myudf;</pre>				
2						
3	+	import				
12	П	public class HelloDriver {				
13						
14 🕨	E	<pre>public static void main(String[] args) throws OdpsException {</pre>				
15						
16		<pre>JobConf job = new JobConf();</pre>				
17						
18		<pre>// TODO: specify map output types</pre>				
19		job.setMapOutputKeySchema(SchemaUtils. <i>tromString</i> (_?));				
20						
21		// TODO: specify input and output tables				
23		InputUtils.addTable(TableInfo.builder().tableName( ?).build(). iob):				
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						
31		RunningJob rj = JobClient.runJob(job);				
32		ij.wattroicompterion();				
34						
35						
36		}				
	1	-				

For details of developing MR, see To write MapReduce.

#### Debug the MR program

After the MR program is developed, test your code and check whether it meets the expectations. The following two methods are supported:

Unit test (UT): There are WordCount UT examples in the examples directory. You can refer to them to compile your UT.



Local MR running: During local running, the running data source must be specified. The following two methods are provided to set the test data source:

- MaxCompute Studio uses the Tunnel Service to automatically download table data of a specific MaxCompute project to the warehouse directory. By default, 100 data records are downloaded. If more data is required for testing, use the Tunnel Command of the console or table downloading function of MaxCompute Studio.
- Provide the mock project (example\_project) and table data. You can see example\_project in warehouse to set it by yourself.

1. Run the MR program. Right-click the Driver class and select Run. In the displayed Run Configuration dialog box, configure the MaxCompute project on which the MR program runs.

Run/Debug Configurations		ß
+ - 🖬 🗏 🌮 ↑ 🛛 »	Name: WordCountTest	<u>Share</u> Single instance only
JUnit	Configuration Code Cove	rage Logs
₩UDFTest.simpleInput	Test kind: Class	✓ Fork mode: none ✓ Repeat: Once ✓ 1
WordCountTest	<u>C</u> lass: com.aliy	/un.odps.examples.mr.test.WordCountTest
> MaxCompute Java		
> MaxCompute SQL	<u>V</u> M options:	-ea "
y y benuns	Program a <u>rg</u> uments:	2 <sup>3</sup>
	Working directory:	%MODULE_WORKING_DIR%
	Environment variables:	
	Use classpath of module:	MyFristModule V
	JRE:	Default ( <no jre=""> - module not specified)</no>
	Shorten command <u>l</u> ine:	user-local default: none - java [options] classname [args]
	Before launch: Build. Activate	e tool window
	+ - / + +	
	<b>↓</b> <sup>®</sup> Build	
	🗌 Show this page 🗹 Activa	ate tool window
		d for used of the following dates
0	warning: No JUK specific	OK Cancel Apply

2. Click OK. If table data of the specified MaxCompute project is not downloaded to warehouse, download data first. If a mock project is used or the MaxCompute project table data is downloaded, skip this step. Then, the MR local run framework reads specified table data in warehouse as the MR input and runs the MR program locally. You can view log output and result display on the console.

<pre>/Library/Java/Java/Java/JirualMachines/jdtl.8.0_65.jdt/Contents/Home/hir/java [INF0]Fkum appreduce job in local mode, Type: MR, Job ID: mr_2017022103325_750_3935 [INF0]Start to process input tables INF0]Start to copy table: example_project.wc_in1—&gt;/Users/liuyi/IdeaProjects/sandbox/temp/mr_2017022103325_750_3935/input/example_project/wc_in1 [INF0]Finished copy table: example_project.wc_in2[p1=2, p2=1]—&gt;/Users/liuyi/IdeaProjects/sandbox/temp/mr_2017022103325_750_3935/input/example_project/wc_in2 [INF0]Finished process input tables INF0]Finished process input tables INF0]Start to vrite table scheme : example_project.wc_ont—&gt;/Users/liuyi/IdeaProjects/sandbox/temp/mr_2017022103325_750_3935/output/_default_ INF0]Finished process output tables INF0]Finished to write table scheme : example_project.wc_out&gt;/Users/liuyi/IdeaProjects/sandbox/temp/mr_2017022103325_750_3935/output/_default_ INF0]Finished process output tables INF0]Finished process resources INF0]Finished to write table scheme : example_project.wc_out&gt;/Users/liuyi/IdeaProjects/sandbox/temp/mr_20170221193325_750_3935/output/_default_ INF0]Finished process resources INF0]Finished fill tableInfo INF0]Finished fill tableInfo INF0]Finished fill tableInfo INF0]Finished fill tableInfo INF0]Finished run mapper, num: 2 INF0]Finished run mapper, TaskId: M_000001, Input: example_project.wc_in1 INF0]Finis</pre>
#### Run the MR program in the production environment

After local debugging is complete, release the MR program to the server and run it in the MaxCompute distributed environment.

- 1. Package the MR program to a JAR package and release it to the server. For more information, see *Package, Upload, and Register*.
- 2. Use the MaxCompute console integrated with MaxCompute Studio in seamless mode, that is, in the Project Explorer window, right-click Project and select Open in Console, and input the commands similar to the following *JAR command* in the console command line:

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

# 2.6.4 Unstructured development

An *unstructured data processing framework* is added for MaxCompute 2.0, supporting access to the OSS and Table Store using external tables. Studio provides some code templates for the framework, facilitating users' fast development.

Compile StorageHandler/Extractor/Outputter

- 1. After the *MaxCompute Java Module* (Sample code is provided in the unstructured folder of the examples directory for your reference).
- 2. Right-click the module source code directory src > > main, select new, and select MaxCompute Java.
- 3. Specify Name and Kind. For example, set Name to myun.MyExtractor and Kind to Extractor. Click OK.



- 4. The framework code has been automatically filled in the template. Compile your logic code.
- 5. Compile Outputter and StorageHandler by following the preceding steps.

### **Unit Testing**

You can compile the unit test (UT) by following the examples in the examples directory to test your Extractor/Outputter.

Project ▼ ③ ÷	C ExtractorTest.java ×
🔻 <mark>V studioTest</mark> ~/ldeaProjects/studioTest	ExtractorTest ambulanceFullSchema
idea	3 <b>± import</b>
🔻 🖿 jdev	18
examples	19 G polic class ExtractorTest {
🔻 🖿 com.aliyun.odps.examples	20 private String ambulanceFullSchema =
▶ <b>D</b> mr	21 "vehicle:bigint;id:bigint;patient:bigint;calls:bigint;latitude:dc
udf	22 private String speechbataruttSchema = "Sentence_shr:double;id:String"
	24 @Test
	25 G 🖯 public void testTextExtractor() throws Exception {
	26 /**
	27 * Equivalent to the following SQL:
C & Outputer lest	28 CREATE EXTERNAL TABLE ambulance_data_external
SpeechSentenceSnrExtractor.java	29 (Venicle Digint, 10 Digint, patient Digint, calls Digint, 20 Latitude double Longitude double, time string, direction string)
с ኈ SpeechStorageHandler	31 STORED BY 'com_alivun_odps_udf_example_text_TextStorageHandler'
💿 庙 TextExtractor	32 LOCATION 'oss:///data/ambulance csv/'
с 🔓 TextOutputer	<pre>33 USING 'jar_file_name.jar';</pre>
c 🔓 TextStorageHandler	34
C TestUtil	35 SELECT * FROM ambulance_data_external;
	36 */
	5/ Column[] externation testiment = Unstructuredutits.parseschemastring(

Package and upload

After StorageHandler/Extractor/Outputter is compiled, compress the completed Java program to a JAR package, and upload the package as a resource to the server, see *Package and release*.

### **Create External Table**

1	ř			>	target		New		>	C	Java Class
	^				MyFrist	Ж	Cut		Ctrl+X	Íĸ	Kotlin File/Class
	Ъ				m pom.xr		Copy		Ctrl+C		File
	b		$\mathbf{v}$		scripts		Copy Path	Ctrl+	Shift+C	-0	Scratch File Ctrl+Alt+Shift+Insert
1	Ä			>	.idea		Copy Reference	Ctrl+Alt+	Shift+C	CI.	Package
	9				<b>MyFirst</b>	ĥ	Paste		Ctrl+V	<>	FXML File
	^				target		Find Usages		Alt+F7	4II J	package-info.java
			>		warehouse		Find in Path	Ctrl+	Shift+F	J	module-info.java
				í.	MyfirstPro		Replace in Path	Ctrl+	Shift+R	5	MaxCompute SQL 脚本
		>	IIII	Ext	ernal Libra		Analyze		>	F	MaxCompute 交互式查询
		>	<u>B</u>	Sci	atches and		Refactor		>	F	从 D2 加载
										≝∎ H	HTML File
							Add to Favorites	~ L	2 01/0 T	4II J	JavaFXApplication
							Show Image Thumbhails	Ctrl+	Shift+1	4II J	Singleton
							<u>R</u> eformat Code	Ctr	l+Alt+L	G	Gradle Kotlin DSL Build Script
							Optimi <u>z</u> e Imports	Ctrl	+Alt+O	G	Gradle Kotlin DSL Settings
							<u>D</u> elete		Delete	4	XSLT Stylesheet
							Build Module 'MyfirstProje	ect'			Edit File Templates
	nre						R <u>e</u> build ' <default>'</default>	Ctrl+9	Shift+F9	믬	GUI Form
	빌						Show in Explorer			相	Dialog
	₩ 1					2.	Open in terminal				Form Snapshot
	3					5	Set MaxCompute project			1	Resource Bundle
	s						Local <u>H</u> istory		>	C	Plugin DevKit >

1. Right-click scripts and select new > MaxCompute Script.

- 2. Enter the SQL script name. Select the MaxCompute project in which the script is to be executed for Target Project and click OK.
- 3. Select create external table live template in the editor to rapidly insert the script template for creating an external table.

Modify the external table name, column, type, StorageHanlder class path, configuration parameter, external path, and JAR name. Click Run MaxCompute SQL Script to create the external table.

4. Query the created external table.

# 2.6.5 Develop Graph

After the MaxCompute Java module is created, Overview of Graph models can be developed.

### Sample Code

There are some code examples of Graph in the examples directory, and you can refer to the example to get familiar with the structure of the Graph program.

■ Project • ③ ÷ ♦• ↓	C PageRank.java ×
<ul> <li>sandbox ~/IdeaProjects/sandbox</li> <li>idea</li> <li>graph</li> </ul>	2 3 <b>Dimport</b>
<ul> <li>examples</li> <li>com.aliyun.odps.examples</li> <li>graph</li> </ul>	19     -/**       20     * Set program arguments:       21     \$pagerank_in pagerank_out       22     *
C Kmeans C PageRank	23
SSSP  mr udf	<pre>26 public static class PageRankVertex extends 27  28 29 @Override</pre>
<ul> <li>Image: Image: Ima</li></ul>	30 ●↑ 31 ↓ 32 ↓ 32 ↓ 33 ↓ 34 ↓ 35 ↓ 36 ↓ 30 ↓ 31 ↓ 31 ↓ 31 ↓ 31 ↓ 32 ↓ 32 ↓ 32 ↓ 33 ↓ 34 ↓ 35 ↓ 36 ↓ 37 ↓ 37 ↓ 37 ↓ 38 ↓ 39 ↓ 30 ↓ 31 ↓
src     target	<pre>34 } else if (context.getSuperstep() &gt;= 1) { 35      double sum = 0; 36      for (DoubleWriteble recent processes) { 37</pre>
i∎ graph.imi <i>m</i> pom.xml	36 <b>Tor</b> (Doublewritable msg : messages) { 37 sum += msg.get();

#### **Develop a Graph Program**

- 1. Right-click the module source code directory src > > main, select new, and select MaxCompute Java.
- 2. Select the GraphLoader/Vertex type and enter the class name (package name is supported) in the Name text box. Click OK, and the frame code will be automatically filled in by the template, you can continue to modify.

🖳 Create	new MaxCompute java class	×
<u>N</u> ame:		†1
<u>K</u> ind:	UDF UDF	$\sim$
	S StorageHandler	
	Extractor	
ir ;	0 Outputer	
	non AD1	
	🔥 GraphLoader	
	🔨 Vertex	
	🔨 VertexResolver	
	🔨 Combiner	=

#### **Debug Graph Locally**

After the Graph program is developed, test your code and check whether it meets the expectations. You can run the Graph code locally.

1. Run the Graph program: Right-click the Driver class and select Run. In the displayed Run Configuration dialog box, configure the MaxCompute project on which the Graph program runs.



2. Click OK. If table data of the specified MaxCompute project is not downloaded to warehouse, download data first. If a mock project is used or the MaxCompute project table data is downloaded, skip this step. Then, the graph local run framework reads specified table data in warehouse as the Graph input and runs the Graph program locally. You can view log output and result display on the console. Each time you debug locally , a new temporary directory is created under the Intellij directory, as shown in the following figure:



For a detailed introduction to warehouse, see Local Warehouse Directory section in *Develop UDF*.

Run the Graph Program in the Production Environment

After local debugging is complete, release the Graph program to the server and run it in the MaxCompute distributed environment.

- 1. Package the Graph program to a JAR package and release it to the server. For more imformation, please see *How to package and release Graph*.
- 2. Use the MaxCompute console integrated with MaxCompute Studio in seamless mode, that is, in the Project Explorer Window, right-click Project and select Open in Console, and input the commands similar to the following *JAR command* in the console command line:

```
jar - libjars xxx . jar - classpath / Users / home / xxx . jar
com . aliyun . odps . graph . examples . PageRank pagerank_i n
pagerank_o ut ;
```

For more imformation about Graph development, please see Graph.

# 2.6.6 Package、Upload and Register

After a *user-defined function* or *MapReduce* is developed, you must package and release it to the MaxCompute system.

#### Package a UDF or MapReduce

Package a jar and submit resource	8
*MaxCompute project: MySecondProject2 (service.odps	.aliyun.com) 🗸 🕂
*Resource name: MyFristModule-1.0-SNAPSHOT.jar	
*Main class: MyFirstUDF	
*Function name:	
✓ Force update if already exists	
?	OK Cancel

### Note:

If you require special packaging, you can modify relevant settings in the pom.xml file. After packaging, follow these steps to upload the JAR package and register the UDF.

### Upload the JAR package

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

1. Select Add Resource from the MaxCompute menu.

2. Select the MaxCompute project you want to upload the resource to, the JAR file path, and the resource name you want to register. Determine whether to force update when the resource or function already exists. Then click OK.

P Add Resource	X
*MaxCompute project: MySet2 (service.odps.	aliyun.com) 🗸 🕂
*Resource file: /MyfirstProject/MyFristModule/src/main/j	ava/MyFirstUDF.java
*Resource name: MyFirstUDF.java	
✓ Force update if already exists	
2	OK Cancel

3. After uploading is successful, you can view the resource under the Resources node of the Project Explorer window.



Register the UDF

After the JAR package is uploaded, register the UDF.

1. Select Create Function from the MaxCompute menu.

2. Select the required resource JAR and JAR main class, and enter the function name. Click OK.

🖳 Create Function		<b>-</b> ×
*MaxCompute project: MySecondProject2 (service.odps.aliyun.com)	~	+
*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		
ОК	Ca	ancel

3. After the registration is successful, you can view the function under the Functions node of the Project Explorer window.

▼ 🌮 meta_dev
▶ 🛅 Tables & Views
🔻 🛱 Functions
▶ fx plushello
▶ fx ss2s
▶ 🛃 Resources

### Apply the UDF

· Apply the UDF in SQL to complete subsequent development.



# 2.7 Manage MaxCompute jobs

### 2.7.1 Job viewing

MaxCompute Studio supports viewing information of MaxCompute running instances submitted by the current user, including the running status, job type, and start and stop time.

Open Job Explorer

If Job Explorer View is not displayed on Dock on the left, open Job Explorer by choosing View > > Tool > Windows > > MaxCompute Job Explorer.

IJ	M	yfirst	Project	[C:\Users\	Adminis	trator\Ide	eaProjects∖N	Myfirs	tProjec	t]\sc	ripts\	HelloTes
<u>F</u> ile	e <u>F</u>	<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> indo
<b>N</b>	м	yfirs	<u>I</u> o	ol Window	s		)	i i i i i i i i i i i i i i i i i i i	Project			Alt+1
++	D	Dra	Qu	uic <u>k</u> Definiti	on	C	trl+Shift+I	★ F	avorite	es		Alt+2
ojec			Sh	ow Siblings	;			🕨 F	Run			Alt+4
L P	ľ	<u> </u>	Qu	uick <u>D</u> ocum	entatior	n	Ctrl+Q	i ≣∰€	Debug			Alt+5
		(	<u>P</u> a	rameter Inf	ю		Ctrl+P	1	ODO			Alt+6
		2	<u>E</u> x	pression Ty	/pe	C	trl+Shift+P	53 5	Structur	e		Alt+7
orer		Ť	<u>C</u> o	ontext Info			Alt+Q	N V	/ersion	Control		Alt+9
Ц.			Re	ce <u>n</u> t Files			Ctrl+E	* 4	Ant Buil	d		
jet			Re	cently Char	nged Fil	es Ci	trl+Shift+E		Capture	Analysis	s	
P2			Re	cent Chang	es	Д	lt+Shift+C		Capture	e Tool		
N			📌 Co	mpare Wit	h		Ctrl+D	L'	Console	9		
5			Co	mpare with	n Clip <u>b</u> o	ard			esigne	er Lot		
			0	uick Switch	Scheme		Ctrl+`		ducati	onal.Cne	скрет	ans
L A		>	<u> </u>	albar	Serieine		Curri	1.4		bg bot		
힉								P P T		avers		
l M	>	IIII E	V 10	of Buttons					ob Evr	lorer		
	>	<u> </u>	✓ <u>5</u> u	atus bar wigation Br	ar.			m	Vaven	Projects		
			ν IN2	tive Editor			,	-21	alette	riojecia		
			C						Palette			
			BIG	ai Text Dire	ction		2	M.F	roject	Explorer		
			En	ter Present	ation M	ode			ermina	al	A	t+F12
			En	ter Distract	ion Free	e Mode			heme	Preview		
			En	ter Full Scre	een				JI Desi	gner		

### View all job instances in a project

Job Explorer allows you to query submitted job lists by status.

Click the date drop-down box to select another date.

Click Refresh to obtain the job list.

📕 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.

Sort the job list

You can click the column name in the job list to sort the jobs.

Job queue

If a job in running status is waiting for scheduling in a queue, the job's location in the queue and global priority is displayed in the job list.

Note:

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

Save job logs

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

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

You can set the path for saving the log file on the Setting tab of MaxCompute Studio.

### 2.7.2 Job instance

View a job instance

Studio supports the following two ways to view MaxCompute job instances:

1. Open the details of a job in read-only mode using a Logview URL or local offline Logview file.

Users of MaxCompute will be familiar with using Logview to view the details of a job. Using Logview, you can also view the status of tasks submitted by other

users in other projects. You can also view the details of any job by entering a valid Logview URL in Studio.

In the menu bar, select MaxCompute > Open Logview. Valid Logview URLs in the Clipboard are automatically copied to the dialog box displayed. Alternatively, you can select to export the local offline Logview file.

2. In *Job Explorer*, double-click a MaxCompute instance to view its details. You can also right-click the instance and select Open.

### Job details view

The job details view page comprises a toolbar at the top, a properties bar on the left, and a detailed view on the right. The detailed view consists of seven views:

- Execution view: Displays the overall information of a job in the form of a DAG. You can view the dependencies and detailed execution plans for each subtask.
- Timeline view: Displays the execution timeline of a job, allows you to view this timeline in different granularities, and provides a number of filters.
- Details: Displays the details of a job in table view, including the subtask list, the worker list for each subtask, the volume of data processed by workers, the execution time, and the status.
- Script: Displays the corresponding SQL statement and parameter configuration for when a job is submitted.
- Summary (JSON): Displays the running details of a job in JSON format.
- Result: Displays the running results of a job.
- Analysis: Provides scatter plots, long tail distributions, and skewed data charts to show the results of a job execution.

P Job ID: 201702082324	Job ID: 20170208232404727gchxi7u1 ×									
			• • >	Refresh						
				SQL_0_0_j	ob_0u1_SQL_0_0	)_merge				
				Task Name: r	neta dev 20170208	3232404727achx	i7u1 SOL 0 0 0 io	b 0		
				Task I/O Records Status Progress				StartTime	StartTime EndTime	
				M1	10766488/	863168	TERMINATED	100.0	2017-02-09 07:24:19	2017-02-09 07:26:15
	SQL_0_0_job_0			J3_1_2	865705/96	7	TERMINATED	100.0	2017-02-09 07:24:16	2017-02-09 07:24:43
	M1	M2		R4_3	967/967		TERMINATED	100.0	2017-02-09 07:26:38	2017-02-09 07:26:46
	Instance Count: 0/82/82	Instance Count:0/3/3 I/O Records:18907361/2537								
	2017-02-09 07:24:19	2817-82-89 87:24:19								
	2017-02-09 07:26:15	2017-02-09 07:24:43								
	883168 records	2537 records								
	J3									
	Instance Cou	nt: 0/85/85								
	2017 02 09 0	7-26-16								
	2017-02-09 0	7:26:38								
	1004									
	907 rec	ords								
	R4			M1 × M2 ×	J3_1_2 * R4_3 *					
	Instance Cou	nt: 0/85/85		Instance	I/O Records	Status	FinishedPercent.	. StartTime	EndTime	Logid
	170 Records.	3077307		R4_3#0_0 R4_3#10_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	PU1UQXVNVGSX
	2017-02-09 0	7:26:46		R4_3#11_0 R4_3#12_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	eU1URXVNVGM
	100%			R4_3#13_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	ME1UQXVNVGc
				R4_3#14_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:44	PU1UQXVNVGsx
				R4_3#16_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	PUTURXVNVGM
				R4_3#17_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:44	NE1URXVNVGM
				R4_3#18_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	PU1UQXVNVGsx
				R4_3#19_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	PU1UQXVNVGsx
				R4_3#20_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	PUTURXVNVGM
>				R4_3#21_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:43	ek1UQXVNVGsx
2017 02 00 07 24 05			2017 02 00 07 27 20	R4_3#22_0	0/0	TERMINATED	100.0	2017-02-09 07	2017-02-09 07:26:44	PU1UQXVNVGc
2017-02-09 07:24:05	07:27	:28	2017-02-09 07:27:28	#R4_3#23_0	0/0	TERMINATED	100.0	2017-02-09.07	2017-02-09/07:26:43	PUTURXVNVGM
可视化 概要 (JSON) 概	婁 (文本) 结果 SQL									
6: TODO ANTER P	review 🛛 🙆 Tool Output 🛛 🔧	Running Info							V Run	ning Result Devent Log

### Toolbar

Collapse pages on left/right

# <

### >

Used to expand or collapse views on the left and right, allowing you to focus on a particular view.

Stop a job



Used to stop a job that is being executed. You need permission (owner or administrator) to do so.

**Refresh details** 

### G

The basic information of a running job, such as its status and quota, is refreshed automatically, but the detailed view on the right is not. If you want the most up-todate details, you have to refresh them manually.

**Copy Logview** 

0

Copies the corresponding Logview of a job to the Clipboard.

Open job details in a browser

### ٦

Generates a Logview URL and opens it in a browser.

### Store job details locally

B

Stores the job details as a local file.

Auto refresh

With this function enabled, Studio automatically refreshes all details of the executed job on a timed basis.

**Basic information page** 

Displays the basic information of a job, including its ID, Owner, status, start and end time, computing resource usage, input, output, and so on. The basic information of a running job is automatically refreshed on a timed basis. The input and output items list the input and output tables of the job. Double-click a table name to view its details

### **Execution view**

Execution view is the most commonly used tool to display the dependency between Fuxi Job, Fuxi Task, and Operation. It also provides a series of auxiliary tools such as job playback, progress view, heat map view, and so on. The execution view is useful for troubleshooting problems.

The execution view can display job dependencies in three dimensions: the Fuxi Job layer, Fuxi Task layer, and Operation layer. Click the breadcrumb or the up arrow to switch between dimensions. By default, the dependency in the Fuxi Task layer is displayed.

### Note:

For non-SQL jobs, only those in the Fuxi Job and Fuxi Task layers can be displayed. Operation layer jobs cannot be displayed.

• Fuxi Job layer:

Click the breadcrumb or the up arrow to switch between dimensions. By default, the dependency in the Fuxi Task layer is displayed. Fuxi To open the Fuxi Job layer , click the MaxCompute Job in the breadcrumb or the up arrow in the Fuxi Task layer. Fuxi Job workers include the names, start times, and end times of Fuxi Tasks . To go to the Fuxi Task layer, double-click any Fuxi Job worker. · Fuxi Task layer:

In the event there is more than one Fuxi Job, the Fuxi Task layer of the last Fuxi Job is opened by default. This layer can display the dependency, input/output tables, and partitions of Fuxi Tasks. Once the job has ended, click the drop-down box in the toolbar to switch between views, including those for progress, input and output heat maps, and Task time and Instance heat maps. The progress view displays the progress of completion for the worker. The heat map view uses colors to distinguis h between different worker heats. Double-click any Fuxi Task to open its Operation layer. Right-click to open the Operation layers of all Fuxi Tasks.

Contents of a Fuxi Task worker:

Count : a / b / c indicates 1 . Instance that at а certain point in time, the number running of subtask instances is a , the of number completed instances is b, and the task total number of task instances is с. 2. I / 0 records : Similarly , this displays the number of input records and output records certain at а time . point in bar : Indicates 3 . Percentage and orange progress the running status of the task. It is obtained by analyzing the running subtask instances. line connecting subtasks the number of . The shows output records . The arrow indicates the data flow direction .

• Operation layer

The Operation layer reveals how Fuxi Tasks run internally. By clicking a worker, all Operation information is displayed.

### Job playback

Studio supports the job playback function. The history of a job can be reviewed within 12s, just like playing a media file. This function helps you understand the running status of a MaxCompute instance in different time points, rapidly determine the sub-task-level running sequence and time consumed, master the key path for executing a job, and accordingly optimize sub-tasks that run slowly.

• Click > to start playing the job, and click > again to pause. You can also manually drag the progress bar.

• The start time of the job is displayed on the left side of the progress bar, the playing time in the middle, and the end time on the right.

# Note:

The playback function only estimates the volume of I/O data at a certain point in time by measuring the time, thereby determining the progress of completion. This does not represent the actual volume of I/O data. Jobs with a running status do not support playback.

### **Timeline view**

Displays in the form of a Gantt chart detailed data for the distributed execution of a job. You can adjust the display granularity to show all computing workers in a Gantt chart.

Gantt charts can be used to display the time bottlenecks and long-tail workers of running jobs. Multiple filters are also provided that help select the key paths, the largest data worker, and the longest time worker for job execution.

### Job details page

Mainly applies to SQL DML jobs, displaying their Fuxi Task lists and computing worker lists on the compute cluster.

One job typically corresponds to one or more Fuxi Jobs. Each Fuxi Job is divided into multiple Fuxi Tasks (stages), and each Fuxi Task comprises multiple Fuxi instances ( workers).

Right-click a Fuxi instance, and the displayed menu allows you to view the standard output, standard errors, and Debug Info.

### Analysis page

Displays the long-tail workers and skewed data workers of jobs. Displays worker scatter plots and column charts to help diagnose job execution bottlenecks.

Scatter plots and column charts allow you to call the details view page from their workers and check Fuxi instance details.

### **Results page**

The results page displays different pages based on the job type and the parameters configured at the time the job was submitted.

- Select statement and set odps.sql.select.output.format = HumanReadable. This displays the result in text format
- Select statement but do not set the output format. This displays the result in table format
- For scripts where data is exported to the table, the output table name and the link that redirects to the table details are displayed
- · For abnormal jobs, the results page limits the details on abnormalities

# 2.8 Tool integration

# 2.8.1 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

1. 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	de une home constitute	e turbat	X
٩	MaxCompute Studio > SI	DK & Console	Reset
Appearance & Behavior	MaxCompute Console Deta	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			
Mayen Helper			
marchineper			
		OK Cancel Apply	Help

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

Settings	D INCREMENTATION	who +			25			
٩	MaxCompute Studio > SDI	K & Console	•		Reset			
Appearance & Behavior         MaxCompute Console Details           Keymap								
► Editor	Installed Location:	D:\dev	elop\odpscmd_public					
Plugins	Version:	0.23.3						
Version Control			4					
Build, Execution, Deployment								
Languages & Frameworks								
Tools								
MaxCompute Studio								
SDK & Console								
MaxCompute SQL								
Accounts								
Maven Helper								
				OK Cancel App	ly 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.

MaxCompute Project Explorer	- I+ <b>S</b>	nello_word.osql ×	$\mathbb{R}_{compiled:hello_word_osql_sql} \times$	🔄 test_demo.osql 🗙	📑 studio-script-demo.iml 🗙
+ - ∫ĵ Ξ ÷ ♥ îP meta dev ▶ ☐ Ta Add project ▶ ☐ Fu Remove project ▶ ☐ Re ▶ îP meta Show Modify project properties Open in Console		<pre> f t: ►hello_wordk+k+k+k+k+k+k+k+k+k+k+k+k+k+k+k+k+k+k</pre>	osql ************************************		
MaxCompute Console					
Aliyun ODPS Command Line Tool Version 0.22.2 @Copyright 2015 Alibaba Cloud Computing Co., I odps@ meta_dev>[]	Ltd. All n	ights reserved			

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



# 2.9 Configure options

# 2.9.1 Configure MaxCompute Studio

After the MaxCompute Studio plug-in is installed, you can find configuration items of MaxCompute Studio on the left bar of the Settings page of IntelliJ IDEA. For more

# information about how to open the IntelliJ IDEA configuration page, see IntelliJ IDEA Documentation.

MaxCompute Studio configuration option page

The MaxCompute Studio configuration option page provides the following configurat ion items:

1. Path for storing the local metadata base

Specifies the path for locally storing metadata of a MaxCompute project. On MaxCompute Studio, the metadata is stored in the hidden directory. *odps*. *studio / meta* of the local user directory by default.

- 2. Version update options
  - You can use the Automatically checks for new version check box to control whether MaxCompute Studio automatically checks for new version updates.
  - You can use the Check new versions button to manually check new versions. After you click this button, if a new version is available, the Install new version button is displayed. You can click this button to install the new version, and restart IntelliJ IDEA after the installation is complete.

SDK and Console configuration option page

The SDK and Console configuration option page provides the following configuration items:

1. Path for installing a MaxCompute client

Specifies the path for local installation of MaxCompute client. MaxCompute Studio detects the version of the MaxCompute client installed in the path. If detection fails, an error message is prompted.

# Note:

MaxCompute Studio later than the 2.6.1 version provides the latest MaxCompute client. You do not need to specify the path. If you must use a MaxCompute client of a specific version, you can specify the path.

### MaxCompute SQL configuration option page

The MaxCompute SQL configuration option page provides the following configuration items:

### 1. Enable syntax coloring

Select Enable syntax coloring to enable the syntax highlighting feature.

2. Enable code completion

Select Enable code completion to enable the automatic code complementing feature.

3. Enable code formatting

Select Enable code formatting to enable the code formatting feature.

4. Compiler options

These are global default compiler options. The following options can be separately set for each file on the toolbar of the SQL compiler.

- Compiler Mode
  - Statement Mode: In this mode, the compiler compiles and submits a single 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. *NOTE:* Script Mode enables the compiler and optimizer to optimize the execution plan and improve the overall execution efficiency. This mode is in the test phase now.
- Type System
  - Legacy TypeSystem: Indicates the type system of original MaxCompute.
  - MaxCompute TypeSystem: Indicates the new type system introduced by MaxCompute 2.0.
  - Hive Compatible TypeSystem: Indicates the type system in Hive compatibility mode introduced by MaxCompute 2.0.
- · Compiler Version
  - Default Version: Indicates the default version of the compiler.
  - Flighting Version: Indicates the experimental version of the compiler, which includes new features of the compiler being tested.

### Account configuration option page

You can add or manage accounts used to access MaxCompute on the Account configuration option page. For more information, see *User authentication*.

You must specify an account on MaxCompute Studio to access a MaxCompute project and run or submit jobs. MaxCompute Studio currently supports the following account type:

- · Alibaba Cloud account (AccessKey)
- Add an account

On the Account configuration option page, follow these steps:

- 1. Click + or press Ctrl N .
- 2. Select the account type Aliyun Account by AccessKey.
- 3. In the displayed Add Account window, set the following items:
  - · Account Name : Indicates the name of the account on MaxCompute Studio.
  - Using properties file : Read the AccessKey ID and AccessKey Secret from the configuration file.
    - Select the configuration file conf / odps\_confi g . ini after you process User authentication.
  - Using properties : Manually enter the AccessKey ID and AccessKey Secret.
    - Access Id : Enter the AccessKey ID of your Alibaba Cloud account.
    - Access Key : Enter the AccessKey Secret of your Alibaba Cloud account.

P Add MaxComp	oute project Odps.ini files need to be
Connection	downloaded in advance to initialize Setting configuration items such as AK/Endpoint.
Properties File:	Ioad\MaxCompute'Client\conf\odps_config.ini 🗁
AK Account:	Using AK account to initialize AK, you can + click on the right + to add.
* Access Id:	LTA HINK ITZIh9
* Access Key:	oAPre 92Rr2jJEf
* Project Name:	MySecondProject2
* End Point:	http://service.odps.aliyun.com/api
	~
0	OK Cancel

4. Click OK to complete addition. Then, the account will be displayed in the Account list on the Account configuration option page.

Delete an account

On the Account configuration option page, follow these steps: (This operation only deletes the account configuration on Studio configuration, which does not affect your account.)

- 1. Select the account to be deleted in the account list.
- 2. Click -.
- 3. In the displayed dialog box, click OK.

Modify the AccessKey of an account

On the Account configuration option page, follow these steps:

- 1. Select the account to be deleted in the Account list.
- 2. Click the pencil icon.
- 3. In the displayed Edit Account window, modify the account information. The content is similar to that in the preceding section Add Account.

View the opening and connection of MaxCompute Region and the settings of Endpoint, see *Endpoints and Data Centers*.

# **3 Eclipse Plugins**

### 3.1 Install

To facilitate the development work with Java SDK of MapReduce and UDF, MaxCompute provides Eclipse Development Plug-in. This plug-in can simulate the running process of MapReduce and UDF to provide local debugging methods and simple template generation.



- Note:
- · Current versions of Eclipse Neon are likely to cause plug-in loading to fail, please use the Eclipse Luna version.
- To download this plug-in, click Here.
- · Unlike the local running mode provided by MapReduce, Eclipse plug-in cannot synchronize data with MaxCompute. Data must be manually copied to the warehouse directory of Eclipse plug-in.

After downloading the Eclipse plug-in, decompress the software package to find the following jar:

odps - eclipse - plugin - bundle - 0 . 16 . 0 . jar

Place the plug-in into the subdirectory plugins in Eclipse installation directory. Start the Eclipse plug-in, and click Open Perspective in the upper right corner. As follows.



After clicking the button, the following dialog box is displayed. As follows.

CVS Repository Exploring
🕸 Debug
🔐 Git
Java (default)
Java Browsing
Java Type Hierarchy
a cava type hierarchy
CODPS
Plug-in Development
i Resource
E <sup>0</sup> Team Synchronizing
Control OK
Cancel OK

Select ODPS and click OK. The MaxCompute icon appears in the upper right corner, indicating that the plug-in takes effect.As follows.



# 3.2 Create a MaxCompute project

This topic describes two methods you can use to create a MaxCompute project.

#### Method 1

1. In the upper left corner, choose File > New > ODPS Project to create an ODPS project (*ODPS* is used as the project name).

	New	Alt+Shift+N ►	1	ODPS Project
	Open File		12	Java Project
	Open Projects from File System			Project
	Close	Ctrl+W	Ø	MapReduce Driver
	Close All	Ctrl+Shift+W	Ø	Mapper
	Save	Ctrl+S	6	Reducer
	Save As		UDF	UDF
6	Save All	Ctrl+Shift+S	UDTF	UDTF
	Revert		ŧ	Package
	Move		0	Class
	Rename	F2	C	Interface
8	Refresh	F5	G	Enum
-	Convert Line Delimiters To	+	e	Annotation
æ	Drint	Ctrl+D	± €	Source Folder
	Fillion	Currr		Folder
è	Import			File
4	Export			Untitled Text File
	Properties	Alt+Enter		
	1 DownloadUtils.class [com.aliyun.odp]			Example
	2 _schema_ [odps/warehouse//wc_in1]		Ľ	Other Ctrl+N
	3 UDFExample.java [odps/examples/com/]		r p	roject: example project
	4 _schema_ [odps/warehouse//wc_in2]		ms	@ Javadoc 😣 Declaration 토
	Switch Workspace	۱.	les t	o display at this time.
	Restart			
	Exit			

2. In the displayed dialog box, enter the project name in the Project name field, specify the installation path of the MaxCompute client, and click Apply.

New ODPS Project Wizard		Quick Access
Create ODPS project	Preferences ODBS Settings	- • •
Project name:       odps         Image: Use default location       Location:         D:\Users\Administrator\WorkSpace\odps         Config ODPS console installation path         Image: Use default ODPS console installation path	OUPS Settings	Config ODPS console installation path Config ODPS console installation path D::::::::::::::::::::::::::::::::::::
?	?	Restore Defaults     Apply       Apply and Close     Cancel



For information about the MaxCompute client, see *Client*.

3. On the left-side Package Explorer tab, view the directory structure of the project.



### Method 2

1. In the upper left corner, click New.



### 2. In the displayed dialog box, select ODPS Project and click Next.

	New			
Select a wizard				
Wizards:				
type filter text				8
<ul> <li>Class</li> <li>Interface</li> <li>Java Project from Existing Ant Buildfile</li> <li>Plug-in Project</li> <li>General</li> <li>CVS</li> <li>Git</li> <li>Java</li> <li>ODPS</li> <li>MapReduce Driver</li> <li>ODPS Project</li> <li>Reducer</li> <li>UDAF</li> <li>UDF</li> <li>UDTF</li> <li>Plug-in Development</li> <li>Cyser Assistance</li> </ul>				
?	< Back	Next >	Cancel	Finish

#### 3. Perform the steps described in *Method* 1.

After the MaxCompute Eclipse plugin is installed, you can use it to compile a MapReduce or UDF program. For information about how to compile a MapReduce program, see *MapReduce*. For information about how to compile a UDF program, see *UDF*.

### 3.3 MapReduce

This article shows you how to use Eclipse to develop and run MapReduce programs.

Select WordCount example in MaxCompute project, as follows.



### Right-click WordCount.java and choose Run As -> ODPS MapReduce, as follows.

洋 Package Explorer 🛛	E 😫 😜 🗸 🗆 🗖	🚺 WordCount.java 🛛		
⊿ ﷺ odps ⊯ src		1 package com.aliyun.	odps.examples.mr;	
JRE System Library [Java]		3⊕ import java.io.IOException;		
Referenced Libraries		18 /*	1 =	
a 进 examples	New	•	1UCE绵细祭洞 <u>爱</u> 预烤烘渣细散渣	
tom.aliyun.odps	Open	F3	nt {	
4 🔏 com.aliyun.odps	Open With	•	s TokenizerMapper extends MapperBase {	
⊳ ∭ Kesource.jav	Open Type Hierarchy	F4		
> A com.alivun.odps	Show In	Alt+Shift+W ▶		
i de la com.aliyun.odps	Сору	Ctrl+C		
🗁 temp	Copy Qualified Name			
b b warehouse	Paste	Ctrl+V	<pre>p(TaskContext context) throws IOException {     constema 200 to the formation of the f</pre>	
×	Delete	Delete	createMapOutputValueRecord();	
8	Remove from Context	Ctrl+Alt+Shift+Down	<pre>ject[] {1L}); getCounter("MyCounters" "global counts");</pre>	
	Build Path	•		
	Source	Alt+Shift+S ▶		
	Refactor	Alt+Shift+T ►	Declaration 🗉 Console 🛛	
249	Import			
2/3	Export			
	- -			
	References	*		
	Declarations	,	-	
୍ର ଜୁନ	Refresh	F5		
	Assign Working Sets			
	Coverage As	•		
	Run As	•	I Java Application         Alt+Shift+X, J	
	Debug As	+	2 ODPS Mapreduce Graph	
	Validate		3 ODPS UDF UDTF UDAF	
	Restore from Local Histo	ory	Run Configurations	
	Team	•		
	Compare With	÷.	5	

After the dialog box is popped up, select example\_project and click Finish。As follows.

ODPS MapReduce Graph Run Configuration
ODPS Mapreduce Graph Run Configuration
Class com.aliyun.odps.examples.mr.WordCount
Run Mode Local   Remote
Select ODPS Project
MySecondProject2       Add         example_project       Edit         Remove       Remove
Resources
Program Arguments
? <u>Finish</u> Cancel

After running is completed, the following result is displayed, as follows.

Console X	- 🗶 💥 🕞 🛃 🛃 🛃 🛃 - 🗂 - 🗖 -
<terminated>WordCount [ODPS Mapreduce] /Library/Java/JavaVirtualMachines/jdk1.7.0_71.jdk/Contents/Home/bin/java (2015年1月27日下午3:</terminated>	12:38)
ARS. Retoud warehouse cable.wc_out	
Summary:	
example_project.wc_in1.example_project.wc_in2/p1=2/p2=1	
Outputs:	
example_project.wc_out	
M1_example_project_LOT_0_0_0_job0	
Worker Count: 2	
Input Records:	
Output Records:	
R2_1: 17 (min: 8, max: 9, avg: 8)	
R2_1_example_project_LOT_0_0_0_job0	
Worker Count: 1	
input seconds: input: 5 (min: 5, max: 5, ava: 5)	
Output Records:	
R2_1FS_9: 5 (min: 5, max: 5, avg: 5)	
counters: 10	
map-reduce transvors: /	
combine_input_groups=s	
map_input_bytes=87	
map_input_records=7	
<pre>map_output_records=17</pre>	
reduce_output_[example_project.wc_out]_bytes=37	
reduce_output_lexample_project.wc_out_l_records=5	
mycounters	
global_counts=22	
map_outputs=17	
reduce_outputs=5	
OK	
InstanceId: mr_20150127074239_358_27772	
	1

### Run User-defined MapReduce Program

**Right-click src directory. Select New -> > Mapper, as follows.**


After selecting Mapper, the following dialog box is displayed. Input the name of Mapper class and click Finish.as follows.

● New Mapper						
Mapper						
Create a new I	Mapper implementation.					
Source fol <u>d</u> er:	odps/examples	Br <u>o</u> wse				
Pac <u>k</u> age:	com.aliyun.odps.examples.mr	Bro <u>w</u> se				
Nama	UserManner	1				
Na <u>m</u> e:	Userwapper					
<u>S</u> uperclass:	com.aliyun.odps.mapred.MapperBase	Brows <u>e</u>				
Interfaces:		<u>A</u> dd				
		Kemove				
		]				
?	<u> </u>	Cancel				

The file UserMapper.java is generated in the src directory in Package Explorer. The content of this file is a template of Mapper class, as follows.

```
odps ;
java . io . IOExceptio n ;
com . aliyun . odps . data . Record ;
com . aliyun . odps . mapred . MapperBase ;
class UserMapper extends MapperBase
package
import
import
import
                                                    MapperBase {
public
    @ Override
              void
                          setup ( TaskContex t
                                                         context )
     public
                                                                       throws
IOExceptio n {
    @ Override
                         map ( long
     public
               void
                                           recordNum ,
                                                                       record ,
                                                           Record
               t context)
TaskContex
               throws
                           IOExceptio n {
    @ Override
                          cleanup ( TaskContex t
     public
               void
                                                           context )
                                                                          throws
IOExceptio n {
```

In the template, the configured package name defaults to odps. You can modify it according to your actual requirement. Write the template content as follows, as follows.

```
package
         odps ;
        java . io . IOExceptio n ;
import
        com . aliyun . odps . counter . Counter ;
import
       com . aliyun . odps . data . Record ;
import
import
        com . aliyun . odps . mapred . MapperBase ;
public
        class
                UserMapper extends
                                       MapperBase {
    Record
            word ;
    Record
            one ;
   Counter
             gCnt ;
   @ Override
   public
            void
                   setup ( TaskContex t
                                          context ) throws
IOExceptio n
              {
         word = context . createMapO utputKeyRe cord ();
         one = context . createMapO utputValue Record ();
one . set ( new Object [] { 1L });
         gCnt =
                 context . getCounter (" MyCounters ", "
global_cou nts ");
   @ Override
                   map ( long
    public
            void
                               recordNum ,
                                            Record
                                                     record ,
TaskContex
           t context)
           for
   i ++) {
();
             String [] words = record . get ( i ). toString ().
split ("\\ s +");
                           w : words ) {
new Object [] {
               vord . set ( new
             for
                                             w });
                        cnt = context . getCounter (" MyCounters
               Counter
", " map_output s ");
cnt . increment ( 1 );
               gCnt . increment ( 1 );
               context . write ( word ,
                                        one);
   @ Override
    public
            void
                   cleanup ( TaskContex t context ) throws
IOExceptio n {
```

Similarly, right-click src directory and select New -> > Reduce:,as follows.

New Reduce	● New Reducer								
Reducer	Reducer								
Create a new l	Reducer implementation.								
Source folder:	odps/examples	Browse							
Package:	com.aliyun.odps.examples.mr	Browse							
Name:	UserReduce								
Superclass:	com.aliyun.odps.mapred.ReducerBase	Browse							
Interfaces:		Add							
		Remove							
		]							
?	Finish	Cancel							

Input the name of Reduce class. (In this example, use UserReduce as the class name.)

In Package Explorer, a file name UserReduce.java is generated in the src directory. This file content is a template of Reduce class. Edit the template, as follows.

```
odps ;
package
import
        java . io . IOExceptio n;
import
        java . util . Iterator ;
import
        com . aliyun . odps . counter . Counter ;
import
        com . aliyun . odps . data . Record ;
import
        com . aliyun . odps . mapred . ReducerBas e ;
public
        class UserReduce
                           extends
                                       ReducerBas e
                                                     {
    private
             Record
                      result ;
    Counter
             gCnt ;
  @ Override
   public
            void
                 setup ( TaskContex t
                                           context ) throws
IOExceptio n {
         result = context . createOutp utRecord ();
         gCnt = context . getCounter (" MyCounters ",
                                                       ...
global_cou nts ");
   @ Override
   public void
                   reduce ( Record
                                     key , Iterator < Record >
values , TaskContex t context )
```

```
throws IOExceptio n {
    long count = 0;
    while (values . hasNext ()) {
        Record val = values . next ();
        count += (Long) val . get (0);
        result . set (0, key . get (0));
        result . set (1, count);
        Counter cnt = context . getCounter (" MyCounters ", "
        reduce_out puts ");
        cnt . increment (1);
        gCnt . increment (1);
        gCnt . write ( result );
        @ Override
        public void cleanup (TaskContex t context) throws
IOExceptio n {
    }
    }
    }
}
```

Create main function: right-click src and select New -> > MapReduce Driver. Enter Driver Name (in this example, use UserDriver as the name), Mapper and Reduce (in this example use UserMapper and UserReduce as corresponding names) and click Finish. The file MyDriver.java is also displayed in src directory, as follows.

e New MapRe	⊖ New MapReduce Driver 🗆 🖻 🕱								
MapReduce Driver									
Create a new l	Create a new MapReduce driver.								
Source fol <u>d</u> er:	odps/examples	Br <u>o</u> wse							
Pac <u>k</u> age:	com.aliyun.odps.examples.mr	Bro <u>w</u> se							
Na <u>m</u> e:	UserDriver								
<u>S</u> uperclass:	java.lang.Object	Brows <u>e</u>							
Interfaces:		<u>A</u> dd							
		Remove							
Ma <u>p</u> per:	UserMapper	<u>B</u> rowse							
<u>R</u> educer:	UserReduce	Browse <u>.</u>							
?	<u> </u>	Cancel							

Edit the driver content, as follows.

```
odps ;
package
import
           com . aliyun . odps . OdpsExcept ion ;
           com . aliyun . odps . data . TableInfo';
com . aliyun . odps . examples . mr . WordCount .
import
import
SumCombine r;
import
           com . aliyun . odps . examples . mr . WordCount . SumReducer
,
import
           com . aliyun . odps . examples . mr . WordCount .
TokenizerM apper;
import com . aliyun . odps . mapred . JobClient;
import
           com . aliyun . odps . mapred . RunningJob ;
           com . aliyun . odps . mapred . conf . JobConf ;
com . aliyun . odps . mapred . utils . InputUtils ;
com . aliyun . odps . mapred . utils . OutputUtil
import
import
import
                                                                            s;
           com . aliyun . odps . mapred . utils . SchemaUtil
class UserDriver {
import
                                                                           s;
                     UserDriver
public
     public
                                    main ( String [] args ) throws
                static
                          void
OdpsExcept ion
                      {
                                        JobConf ();
          JobConf
                      job =
                                 new
          job . setMapperC
                                 lass ( TokenizerM apper . class );
```

job . setCombine rClass ( SumCombine r . class ); job . setReducer Class ( SumReducer . class ); job . setMapOutp utKeySchem a ( SchemaUtil s . fromString (" word : string ")); job . setMapOutp utValueSch ema ( SchemaUtil s . fromString (" count : bigint ")); InputUtils . addTable ( TableInfo . builder (). tableName (" wc\_in1 "). cols ( new String [] { " col2 ", " col3 " }). build (), job ); InputUtils . addTable ( TableInfo . builder (). tableName (" wc\_in2 "). partSpec (" p1 = 2 / p2 = 1 "). build (), job ); OutputUtil s . addTable ( TableInfo . builder (). tableName (" wc\_out "). build (), job ); RunningJob rj = JobClient . runJob ( job ); rj . waitForCom pletion ();

Run MapReduce program. Right-click UserDriver.java and select Run As -> > ODPS MapReduce, the following dialog box is displayed,as follows.

🖨 ODPS MapReduce Graph Run Configuration 📃 📼 🖾
ODPS Mapreduce Graph Run Configuration
Class com.aliyun.odps.examples.mr.UserDriver
Run Mode Ocal      Remote
Select ODPS Project
MySecondProject2 example_project Add Edit Remove
Resources
•
Program Arguments
? Finish Cancel

Select example\_project as the MaxCompute Project and click Finish to run MapReduce program in the local,as follows.

E Console X	= 🗙 💥 🛼 🚮 🖅 🕾 🖃 - 🗂 - 🗖 -
<terminated> UserDriver [ODPS Mapreduce] /Library/Java/Java/JavaVirtualMachines/jdk1.7.0_71.jdk/Contents/Home/bin/java (2015年</terminated>	1月27日 下午4:22:42)
Summory:	
Inputs:	
example_project.wc_in1,example_project.wc_in2/p1=2/p2=1	
Outputs:	
example_project.wc_out	
M1_example_project_LOT_0_0_0_job0	
Worker Count: 2	
Input Records:	
input: 7 (min: 3, max: 4, avg: 3)	
Output Records:	
R2_1: 17 (min: 8, max: 9, avg: 8)	
R2_1_example_project_LOT_0_0_0_job0	
Worker Count: 1	
Input Records:	
unput: 5 (min: 5, max: 5, avg: 5)	
$P_2$ (ES 0) 5 (min) 5 may 5 ave 5)	
counters: 10	
map-reduce framework: 7	
combine input groups=5	
combine_output_records=5	
map_input_bytes=87	
<pre>map_input_records=7</pre>	
<pre>map_output_records=17</pre>	
<pre>reduce_output_[example_project.wc_out]_bytes=37</pre>	
reduce_output_[example_project.wc_out]_records=5	
user defined counters: 3	
mycounters	
global_counts=22	
map_outputs=17	
reduce_outputs=5	
OK .	
InstanceId: mr 20150127082243 694 27864	
	1

If the output is the same as in the preceding figure, it indicates that local operation runs successfully. The output result is saved in the warehouse directory. Refresh MaxCompute project, as follows.

		- odps/warehouse/ex	ample_project/tables	_/wc_out/R_000000	- Eclipse - /Use	ers/aliba	aba/Do	ocumer	ts/eclip	se/work	space	_
◘ •    ¶ @ @ \$• ● • <b>Q</b> •   ₩ @ •   ₫	• 🔗 • 🔲 🔳	⊈ •∦I•∜⊃⇔⇒	•				0		ЛТ	· ±	×	
🛱 Package Explorer 🛿	· ·	D WordCount.java	J UserReduce.java	🕽 UserDriver.java	R_000000			•				
▼ (10 <sup>45</sup> odps		A2,4					_			_	_	_
▼ 🕮 src		A3,4										
🔻 🌐 odps		three1,3										
UserDriver.java		three3 3										
🕨 🚺 UserMapper.java	L	chi cco,o										
UserReduce.java												
JRE System Library [Java SE 7 [1.7.0_7	1]]											
Referenced Libraries												
🔻 🚰 examples												
The com.aliyun.odps.examples.mr												
Resource.java												
WordCount.java												
VordCount												
SumCombiner												
► SumReducer												
► S TokenizerMapper												
@ main(String[]) : void												
Com.aliyun.odps.examples.udf												
test												
temp												
V warehouse												
• example_project												
resources_												
table_resource1												
File resource tet												
► C= wc in1												
▶ 🗁 wc_in2												
V Cout												
schema												
R_000000												
Readme		<terminated> LiserDriv</terminated>	er IODES Manreduce] /Lib	ary/ Java/ Java/VirtualMac	binee/idk1 7 0 71	idk/Con	tente/H	ome/bir	liava (20	15年1日2	7日 下午	4-22-42)
		温思: Keloaa warer	iouse table:wc_out	ary/ouva/ouva virtualiviac	innowpart.r.o_r1	.juit oon	101110/11	0110/01	1444 (20	10-11/12	A 14 17 1	4.66.46)
		Summary:										
		Inputs:										
		example_p	roject.wc_in1,example	e_project.wc_in2/p1	=2/p2=1							
						Writable		Ins	ert	1:	1	

wc\_out is the output directory and R\_000000 is the result file. By local debugging, the result is confirmed to be correct and you can package MapReduce program using Eclipse export function. After it is packaged, upload the jar package to MaxCompute. For more information how to run MapReduce in distributed environment, see *Quick Start*.

After the local debugging is completed, you can package the codes in jar package using Eclipse Export function, provided for subsequent distributed environment. In this example, the package name is mr-examples.jar. Select the src directory and click Export, as follows.

▼ APPS odps	A2,4	4
<ul> <li>Ferric System Library [Java SE 7 ['</li> <li>Referenced Libraries</li> </ul>	New Go Into	•
<ul> <li>examples</li> <li>com.aliyun.odps.examples.mi</li> <li>Resource.java</li> <li>WordCount.java</li> <li>WordCount</li> <li>WordCount</li> <li>SumCombiner</li> <li>SumReducer</li> <li>SumReducer</li> <li>Simain(String[]) : void</li> <li>com.aliyun.odps.examples.ud</li> <li>temp</li> <li>com.aliyun.odps.examples.ud</li> <li>temp</li> <li>warehouse</li> <li>example_project</li> <li>resources</li> </ul>	Open in New Window Open Type Hierarchy Show In て第	F4 W ►
	<ul> <li>Copy</li> <li>Copy Qualified Name</li> <li>Paste</li> <li>Delete</li> </ul>	¥C ¥V ⊗
	Build Path Source ℃第 Refactor ℃第 ≧ Import	S F
► table_resource1	🛆 Export	
<ul> <li>table_resource2</li> <li>file_resource.txt</li> <li>tables</li> <li>rs_out</li> <li>rs_out</li> <li>wc_in1</li> <li>wc_in2</li> <li>wc_out</li> <li>schema</li> <li>R_000000</li> <li>Readme</li> </ul>	Refresh Assign Working Sets	F5
	Debug As Run As Team Compare With Restore from Local Histo % Import from Repositor	• • • • • • • • • •
	Properties	жı

Select Jar File as an export mode, as follows.

Export	
Select Export resources into a JAR file on the local file system.	ß
Select an export destination:	
type filter text	
<ul> <li>Ceneral</li> <li>Install</li> <li>Java</li> <li>J</li></ul>	
(?) < Back Next > Finish	Cancel

You must only export the package in src. The Jar File name must be specified as mr

- examples . jar ,as follows.

JAR Export	
JAR File Specification (i) The export destination will be relative to ye	our workspace.
Select the resources to export:	
<ul> <li>Image: COPS</li> <li>Image: Comparison of the sector of</li></ul>	<ul> <li>classpath</li> <li>project</li> </ul>
<ul> <li>Export generated class files and resource</li> <li>Export all output folders for checked proj</li> <li>Export Java source files and resources</li> </ul>	s ects
Export refactorings for checked projects.	Select refactorings
Select the export destination: JAR file: mr-examples.jar Options:	▼ Browse

Click Next to export the jar file.

If you want to simulate new Project creation in the local, you can create a subdirectory (has same level with example\_project) in the warehouse directory. The directory hierarchy structure is shown as follows.

schema Example, as follows.

```
Non - partiton
                    table :
project = project_na me
table = table_name
columns = col1 : BIGINT , col2 : DOUBLE , col3 : BOOLEAN , col4 :
DATETIME , col5 : STRING
Partition table :
project = project_na
                          me
table = table_name
columns = col1 : BIGINT , col2 : DOUBLE , col3 : BOOLEAN , col4 :
DATETIME , col5 : STRING
partitions = col1 : BIGINT , col2 : DOUBLE , col3 : BOOLEAN , col4 :
DATETIME , col5 : STRING
Note :
Currently ,
             the following
                                     five
                                             data
                                                      formats
                                                                  are
supported : bigint , double , boolean , datetime , string , which
correspond to the data types in java : - long , double ,
boolean , java . util . Date , java . lang . String .
```

data Example, as follows.

1 , 1 . 1 , true , 2015 - 06 - 04 11 : 22 : 42 896 , hello world \ N ,\ N ,\ N ,\ N ,\ N Note :

The	time	format	is	accura	te to	o th	e	millise	con	d		
level	and	all	types	are	repres	sente	d	NULL	by	'\	Ν	۰.

# Note:

- If MapReduce program runs in the local, the default is to search corresponding tables or resources from the warehouse directory. If the tables or resources do not exist, corresponding data will be downloaded from the server and saved in warehouse. Then run MapReduce in the local.
- After running MapReduce is finished, refresh the warehouse directory to view the generated result.

### 3.4 UDF

This section describes how to develop UDF with the Eclipse plug-in and how to run UDF on local. The preparation and implement process is similar to UDF. You can see the example of UDF. MaxCompute Eclipse plug-in provides two methods to run UDF: Menu Bar and run by right-clicking it.

#### Run UDF using Menu Bar

1. Select Run > > Run Configurations… from the menu bar and the following dialog box appears, as follows.

Run Configurations	Rep-4.7	×
Create, manage, and run con	figurations	
Image: Second	Name:       Resource            • Main ODPS Config M= Arguments INE Classpath Environmer         Project:         odps project         Main class:         com.aliyun.odps.examples.mr.Resource         Include system libraries when searching for a main class         Include inherited mains when searching for a main class         Stop in main	t Common Browse Search
?	Run	Close

2. You can create a new Run Configuration. Select the UDF class and type to be executed, select MaxCompute Project and enter the information of input table. Aa follows.

000	ODPS UDF UDTF UDAF Ru	un Configuration
ODPS UDF U	JDTF UDAF Run Configuration	
Class		
com.aliyun	.odps.examples.udf.UDFExample	
Select ODP	PS Project	
example_p	roject	Add
		Edit
		Bemove
Input Table		
Table:	wc_in2	
Partitions:	p2=1,p1=2	ie: p1=1,p2=1 (default all partitions)
Columns:	colc,colb,cola	ie: c1,c2,c3 (default all columns)
?		Cancel Finish

In the preceding configuration, Table indicates the input table of UDF.

Partitions indicates the partitions from which the data is read, separated by commas. Columns indicates the columns, which are considered as the parameters of UDF to be introduced. The columns are separated by commas. 3. Click Run to run the program and the running result is displayed in the console, as

follows.

E Console 🕱	= 🗙 💥 📑 📮 🖉
<terminated>udf [ODPS UDF UDTF UDAF] /home/shihai/</terminated>	/lib/java/bin/java (Dec 12, 2014 7:32:23 PM)
<pre>sss2s:three3,three1,three2 sss2s:three3,three1,three2 sss2s:three3,three1,three2</pre>	

#### Run by right-clicking

1. Select a udf.java file (such as UDFExample.java) and right-click it. Then select Run As >> Run UDF|UDAF|UDTF:

	ODPS UDF UDTF UDAF Run	n Configuration
ODPS UDF U	JDTF UDAF Run Configuration	
Class		
com.aliyun	.odps.examples.udf.UDFExample	
Select ODF	PS Project	
example_p	project	Add Edit Remove
Table:	wc_in2	
Partitions:	p2=1,p1=2	ie: p1=1,p2=1 (default all partitions)
Columns:	colc,colb,cola	ie: c1,c2,c3 (default all columns)
?		Cancel Finish

#### 2. The configuration information is shown as follows.

	ODPS UDF UDTF UDAF Ru	n Configuration
ODPS UDF	JDTF UDAF Run Configuration	
Class		
com.aliyun	.odps.examples.udf.UDFExample	
Select ODF	PS Project	
Input Table		Add Edit Remove
Table:	wc_in2	
Partitions:	p2=1,p1=2	ie: p1=1,p2=1 (default all partitions)
Columns:	colc,colb,cola	ie: c1,c2,c3 (default all columns)
?		Cancel Finish

In the preceding configuration, Table indicates the input table of UDF. Partitions indicates the partitions from which the data is read, separated by commas. Columns indicates the columns, which are considered as the parameters of UDF to be introduced. The columns are separated by commas.

3. Click Finish to run UDF and get the output result.

#### Running customized UDF program

Right-click a project and select New >>UDF (or select the menu bar File > >New > > UDF).

Enter the UDF class name and click Finish. Generate a Java file in corresponding src directory with the same name as this UDF class. Edit this java file as follows.

```
odps ;
package
           com . aliyun . odps . udf . UDF ;
class UserUDF extends UDF {
import
public
          project : example_pr oject
       *
           table : wc_in1
       *
           columns : col1 , col2
       *
       public String evaluate ( String
  return " ss2s :" + a + "," + b ;
                                                              String
                                                       а,
                                                                          b) {
      }
```

Right-click this java file (such as UserUDF.java) and select Run As -> ODPS UDF|UDTF| UDAF:

🛱 Package Explorer 🔀 🔲 🔄	UDFExample.java	🚺 UserUDF.java 🔀
▼ ∰ src ▼ ⊕ odps	package odps import com.a	; liyun.odps.udf.UDF;
UserUDF.java	New	DF extends UDF {
<ul> <li>Alte System Library [Java SE 7 [1.7.0_71]]</li> <li>Afferenced Libraries</li> <li>Bexamples</li> <li>Com.aliyun.odps.examples.mr</li> <li>Com.aliyun.odps.examples.udf</li> <li>UDAFExample.java</li> </ul>	Open Open With Open Type Hierarchy Show In て第W	F3 example_project _in1 col1,col2
<ul> <li>▶ ↓ UDAFResource.java</li> <li>▶ ↓ UDFExample.java</li> <li>▶ ↓ UDFResource.java</li> <li>▶ ↓ UDTFExample.java</li> <li>▶ ↓ UDTFResource.java</li> </ul>	<ul> <li>Copy</li> <li>Copy Qualified Name</li> <li>Paste</li> <li>Delete</li> </ul>	<pre>#C g evaluate(String a, String b) { 2s:" + a + "," + b; #V ☑</pre>
<ul> <li>temp</li> <li>warehouse</li> </ul>	Build Path Source て#S Refactor て#T	
	≧ Import ⊿ Export	
	References Declarations	S UDF UDTF UDAF] /Library/Java/JavaVirtual
	Refresh Assign Working Sets	F5
	Debug As	•
	Run As	I ODPS UDF UDTF UDAF
	Team Compare With Replace With Restore from Local History	Run Configurations
	Properties	жі

Configure the following dialog box, as follows.

000	ODPS UDF UDTF UDAF Run	Configuration
ODPS UDF U	JDTF UDAF Run Configuration	
Class		
odps.Userl	JDF	
Select ODP	S Project	
example_p	roject	Add Edit Remove
Table:	wc_in1	ia: p1-1 p2-1 (default all partitions)
Columns:	col1,col2	ie: c1,c2,c3 (default all columns)
?		Cancel Finish

Click Finish to get the result:

ss2s : A1 , A2

Only the operation instance of UDF is described in this section, and the way of UDTF operating is basically similar to the UDF.

## 3.5 Graph

After creating a MaxCompute project, you can write Graph program and complete the local debugging according to the following steps.

In this example, *PageRank* . *java* provided by the plug-in is selected to complete the debugging. Select *PageRank* . *java* in examples, as follows.



**Right-click and select Debug As > ODPS MapReduce**|**Graph,as follows.** 



The dialog box appears, and configure it as follows.

ODPS MapReduce Graph Run Configuration	
ODPS Mapreduce Graph Run Configuration	
Class	
com.aliyun.odps.examples.graph.PageRank	
Run Mode	
Select ODPS Project	
MySecondProject2 example_project	Add
	Edit
	Remove
Resources	
	*
	*
Program Arguments	
pagerank_in pagerank_out	*
? <u>Finish</u>	Cancel

View the running result, as follows.

🖹 Problems 🏼 @	Javadoc	😟 Declaratio	n 📃 Console 🛛		×	×
<terminated> Pag</terminated>	geRank [OI	DPS Mapredu	ce Graph] D:\Prog	ram Files\Java\jre\bin\javaw.exe (2018年12月17日 上午11:45:54)		
3						
vertex edgs siz	ze: 2					
4						
2						
3						
vertex edgs siz	ze: 2					
3						
1						
2						
vertex edgs siz	ze: 2					
User defined co	ounters:	4				
com.ali	iyun.odps	.graph.loca	1.COUNTER			
	TASK_IN	PUT_BYTE=24				
	TASK_IN	PUT_RECORD=	4			
	TASK_OU	TPUT_BYTE=8	8			
	TASK_OU	TPUT_RECORD	=4			
[INFO]Reload wa	arehouse	table:pager	ank_out			
graph task fin	nish					
Job Finished in	1 0.5 sec	onds				

You can view the computing result on the local, as follows.



After the debugging is complete, you can package the program and upload it to MaxCompute as a Jar resource. Then submit Graph job.



Note:

· For the package process, see MapReduce Eclipse Plug-in Introduction.

- For the structure introduction of local result, see *MapReduce Eclipse Plug-in Introduction*.
- For the detailed introduction of uploading Jar resource, see Add Resource in *Basic Introduction*.
- For submitting the Graph job, see *Graph Function*.

# 4 Downloads

This document provides you with the download address of the relevant tools and plugins.

- SDK Downloads: Maven users can search odps-sdk from Maven library to get different versions of the Java SDK.
- MaxCompute client: *Click here* to download the new version of MaxCompute client.
- Eclipse plugin
- IntelliJ plugin, Studio
- JDBC
- PHP SDK