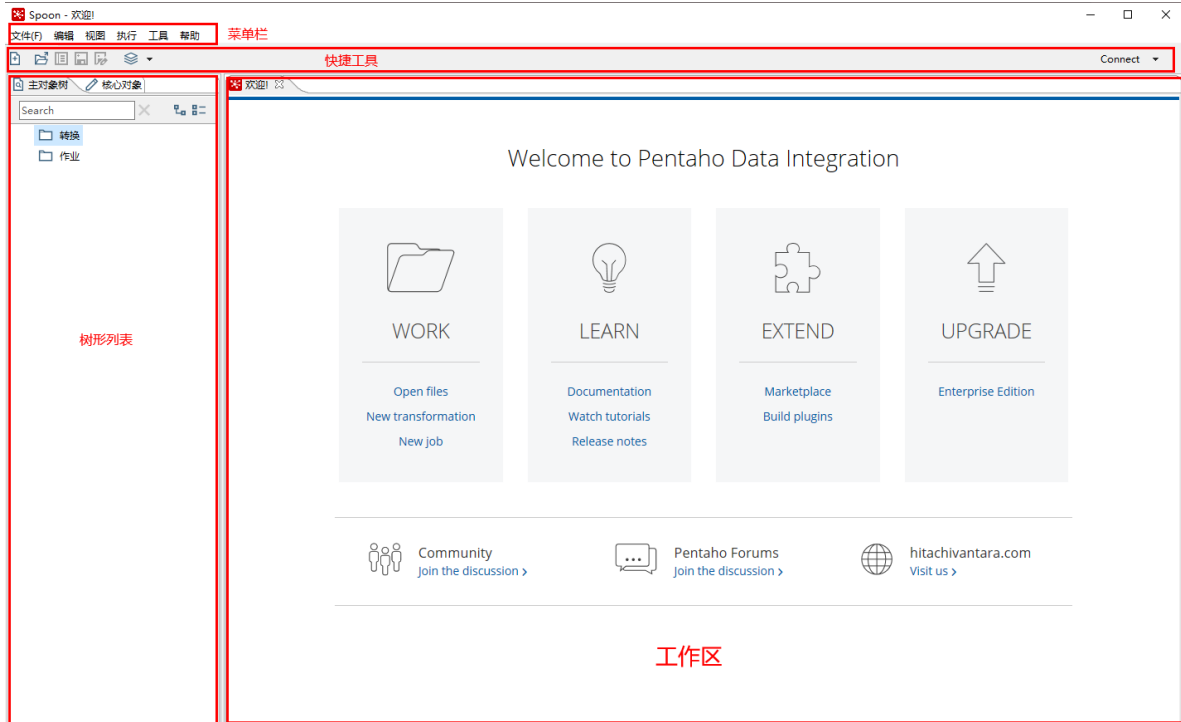


2. 第一个kettle例子

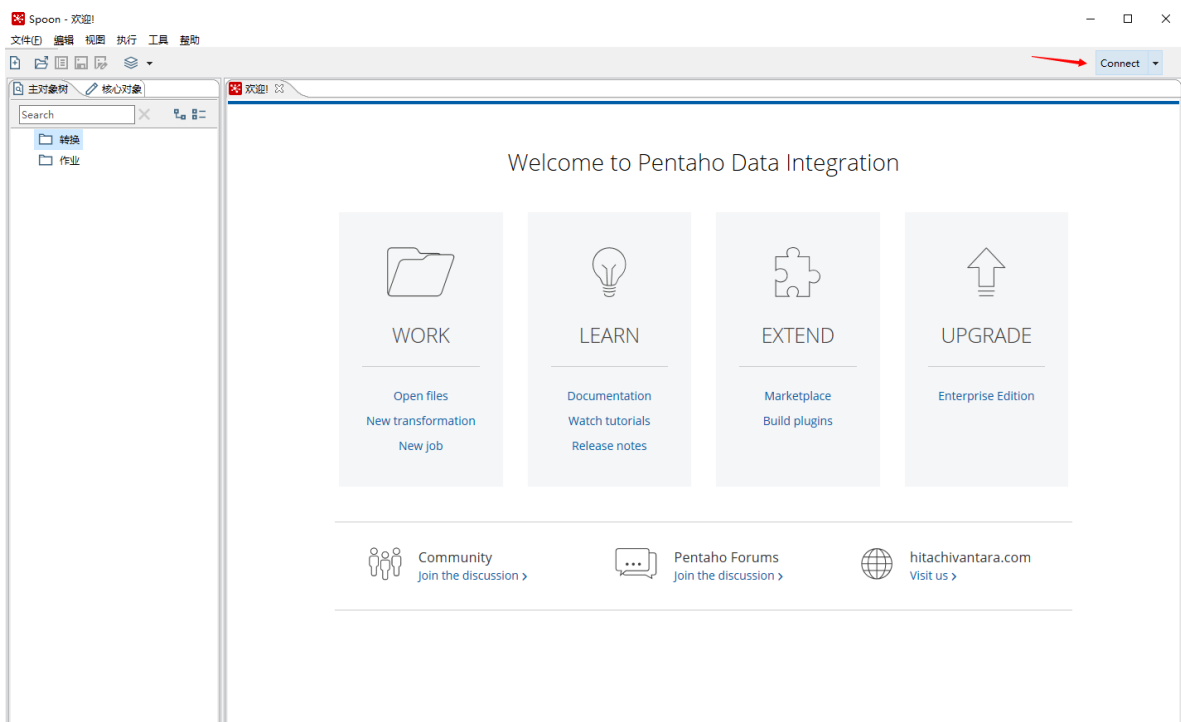
2.1 启动图形化界面

运行Spoon.bat/spoon.sh

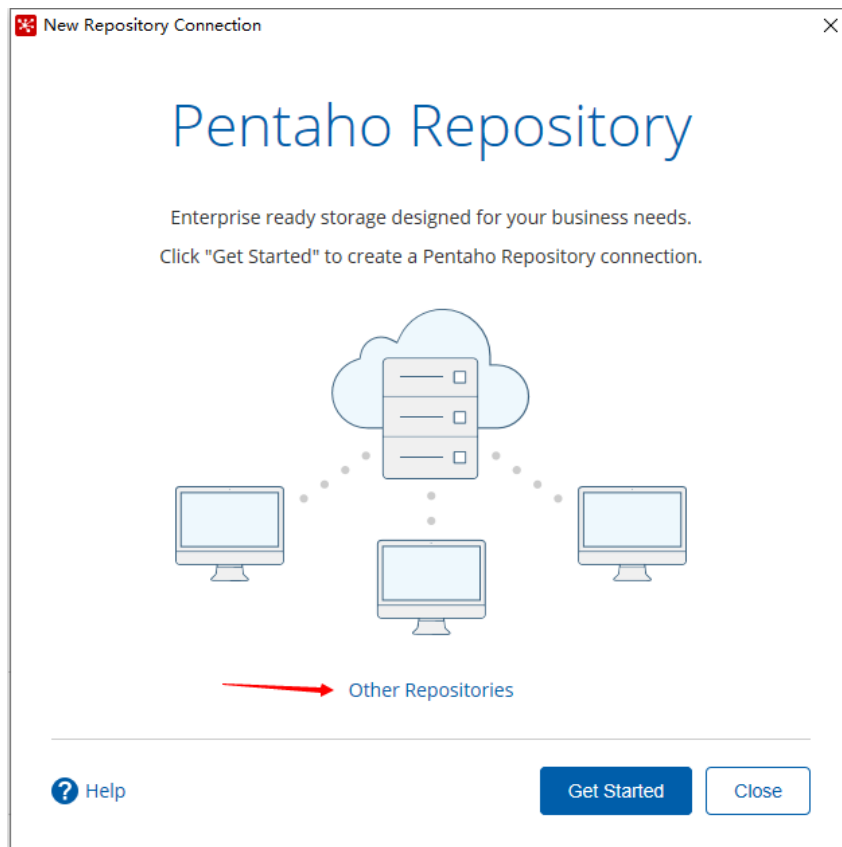


2.1 创建一个资源库

点击connect

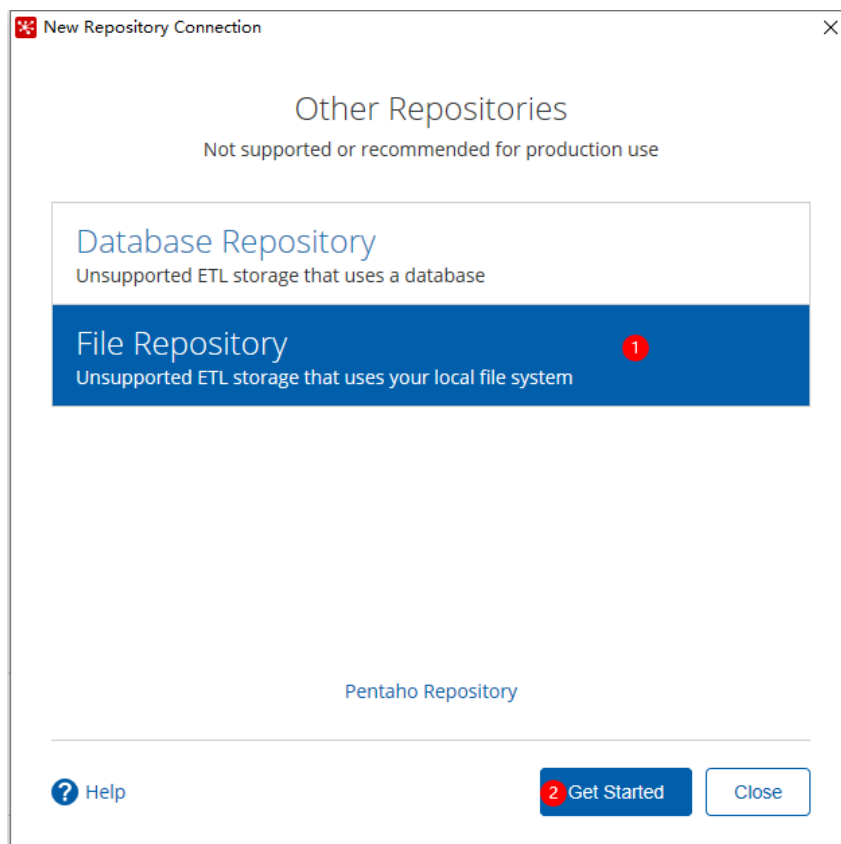


点击Other Repositories



- 文件资源库

选择文件资源库



根据需求填写连接信息，填完后点击Finish

New Repository Connection

Repository Details

Display Name
sample 资源库名称

Location
E:\pdi-ce-8.2.0.0-342\data-integration\samp Browse 资源库位置

☐ Don't modify items in this location 是否设置该资源库为只读

☒ Show hidden folders 是否显示该文件夹下隐藏文件夹

Description
File repository 资源库描述

☐ Launch connection on startup 是否启动spoon就连接到此资源库

? Help Back Finish

- 数据库资源库

选择数据库资源库

New Repository Connection

Other Repositories

Not supported or recommended for production use

Database Repository 1
Unsupported ETL storage that uses a database

File Repository
Unsupported ETL storage that uses your local file system

Pentaho Repository

? Help 2 Get Started Close

点击下图位置配置一个数据库连接

New Repository Connection

×

Connection Details

Display Name

Database Connection

None >

Description

Database repository

☐ Launch connection on startup

? Help

Back

Finish

New Repository Connection

×

Select a database connection

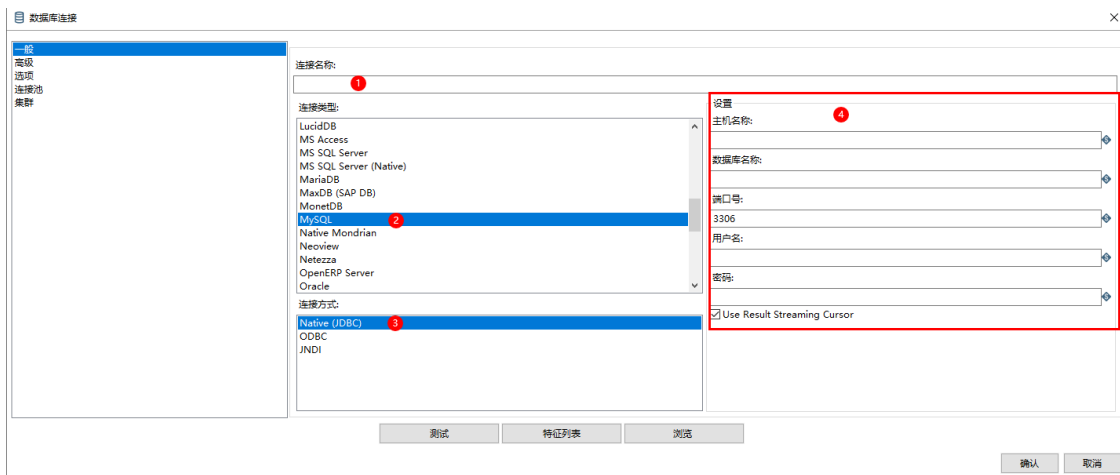
You do not have a database connection for this repository. Please create one.

Create New Connection

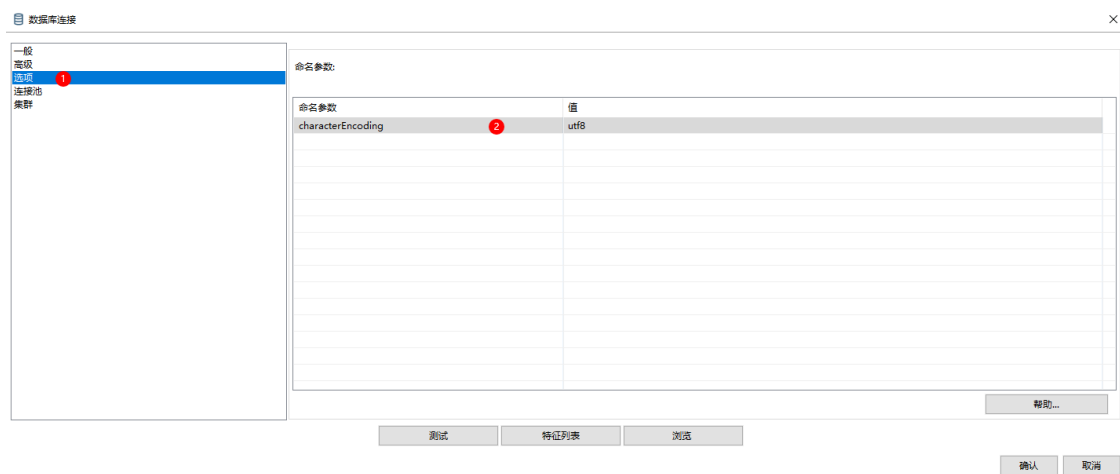
? Help

Back

根据需求填写数据库连接信息



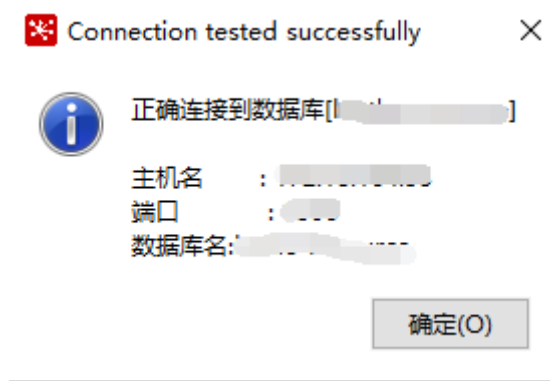
点击高级，输入set names utf8;。点击选项，设置characterEncoding为utf8。如此设置可确保sql抽取数据不出现中文乱码现象。



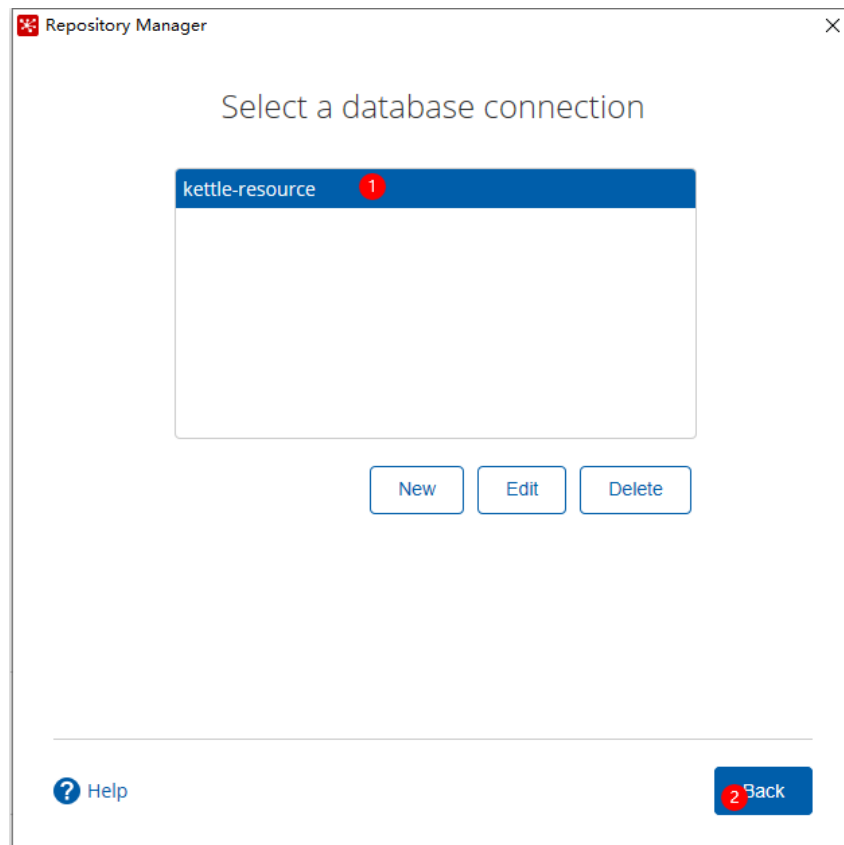
点击测试，如果报了如图所示错误，说明没有mysql的驱动包。一般新装的kettle都没有mysql连接的驱动包，所以会报这个错误，只需要将mysql的连接驱动放入lib文件夹重启后即可解决。



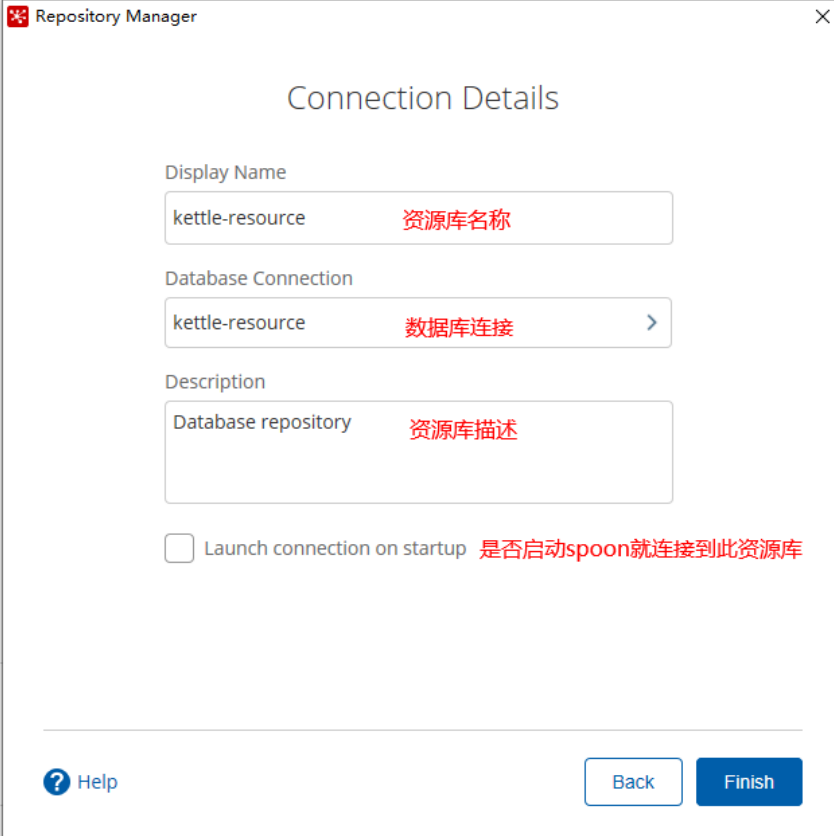
点击测试，出现如下窗口，则说明连接创建成功。



点击确认后，返回到上一级页面，选择刚建好的数据库连接，点击Back



根据需求填好信息，点击Finish



Repository Manager

Connection Details

Display Name
kettle-resource 资源库名称

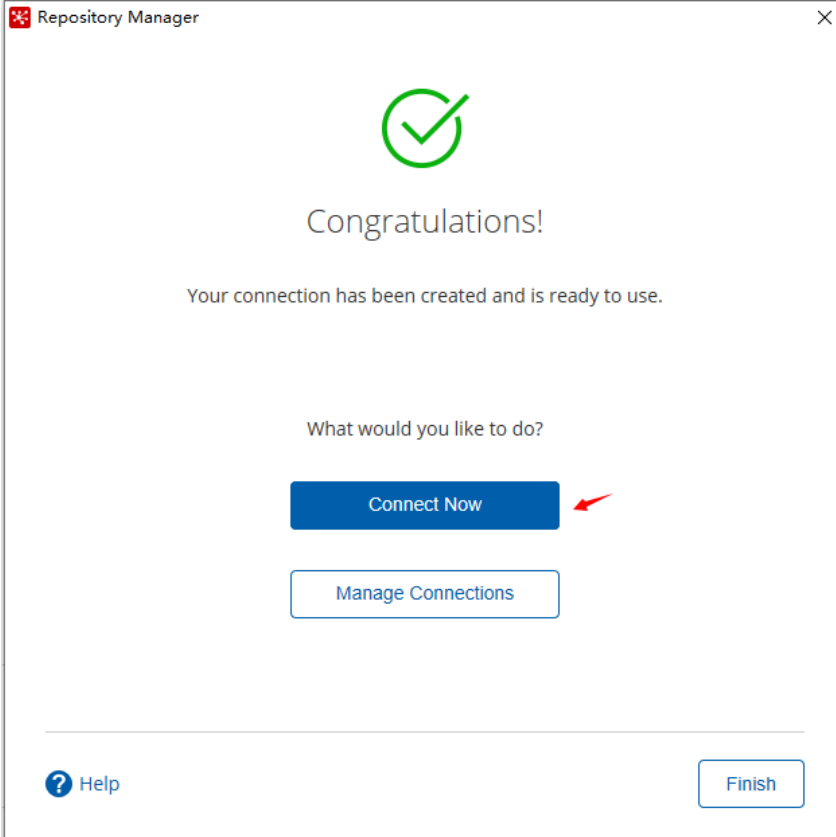
Database Connection
kettle-resource 数据库连接 >

Description
Database repository 资源库描述


☐ Launch connection on startup 是否启动spoon就连接到此资源库

? Help Back Finish

点击connect now即可连上刚创建好的资源库。



Repository Manager



Congratulations!

Your connection has been created and is ready to use.

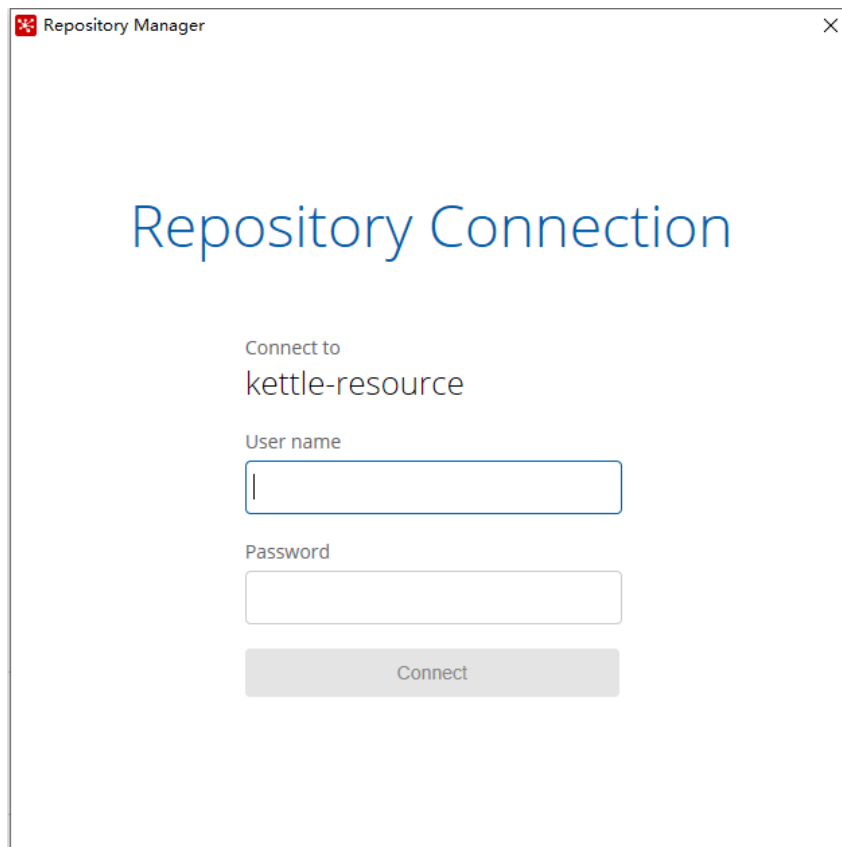
What would you like to do?

Connect Now

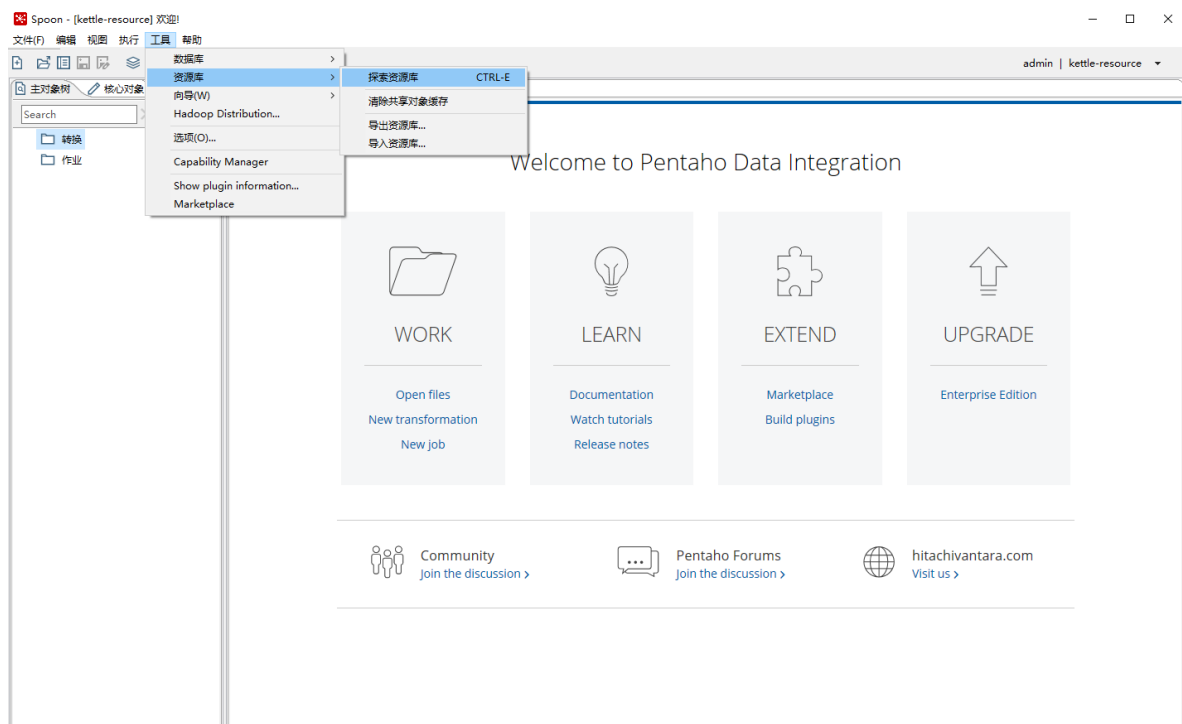
Manage Connections

? Help Finish

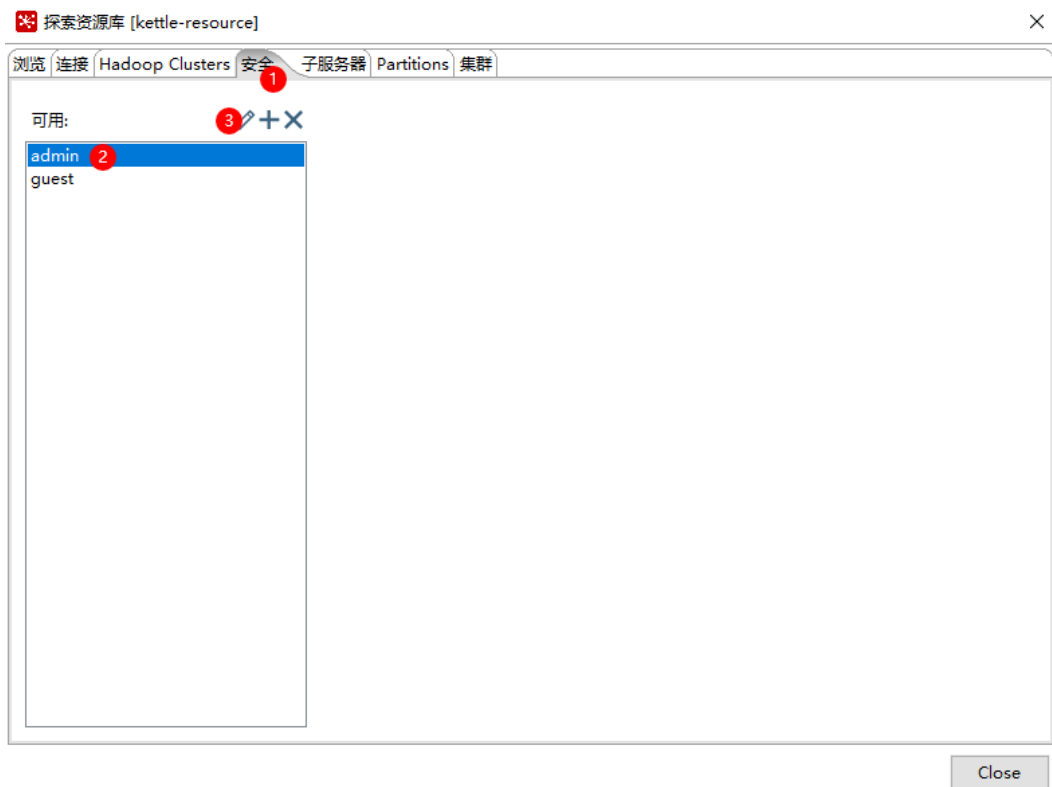
如果资源库类型是数据库资源库，则需要填写用户名和密码，注意不是数据库的用户名和密码，默认的用户名/密码是admin/admin和guest/guest。



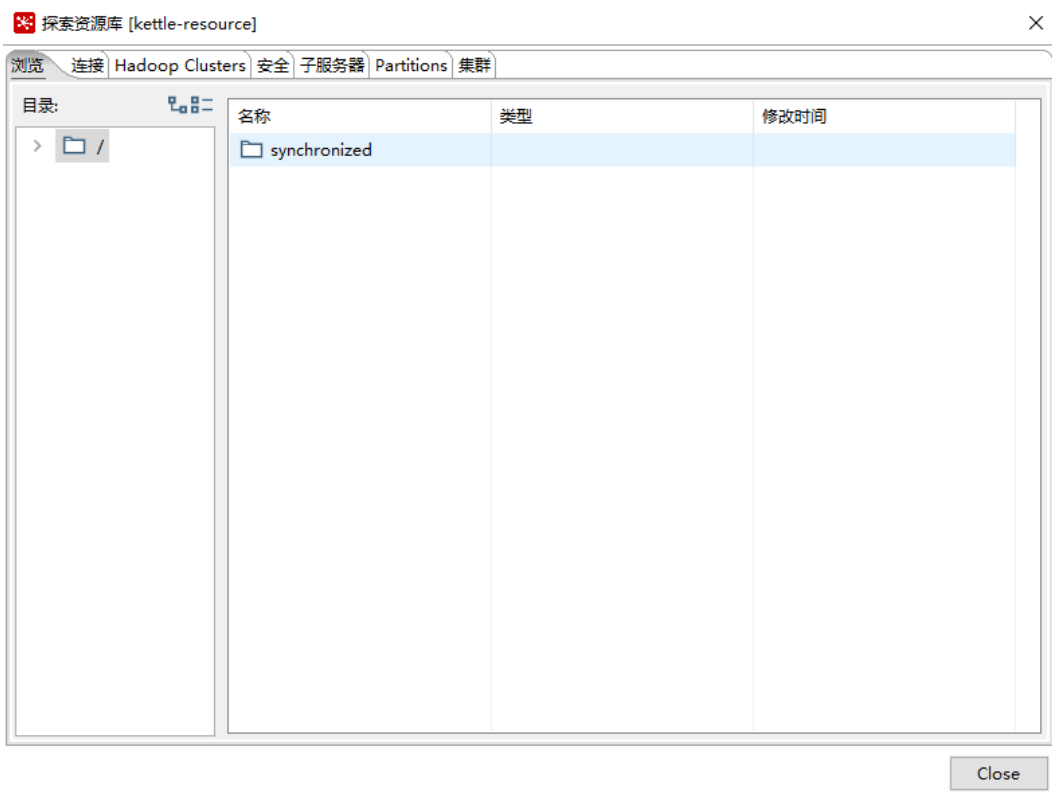
如果需要修改默认用户名和密码，点击菜单栏 工具>资源库>探索资源库



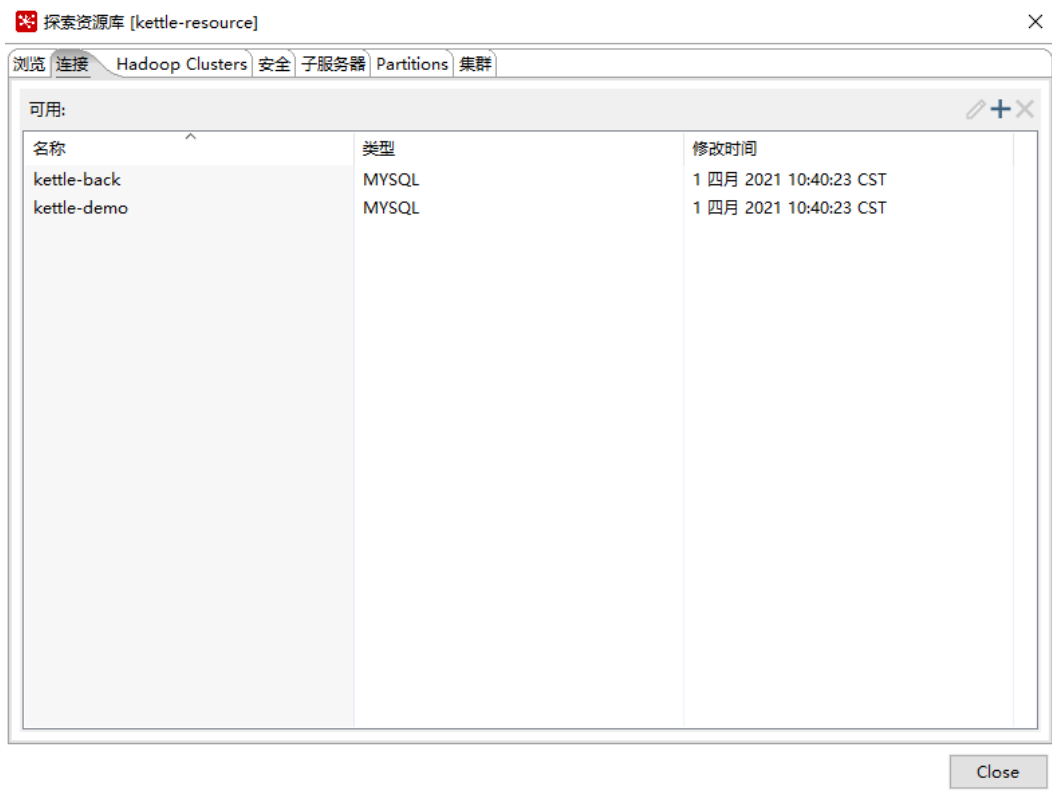
选中需要修改的用户，即可修改密码



探索资源库，浏览选项卡，就是存储转换和作业的目录，可以新建目录

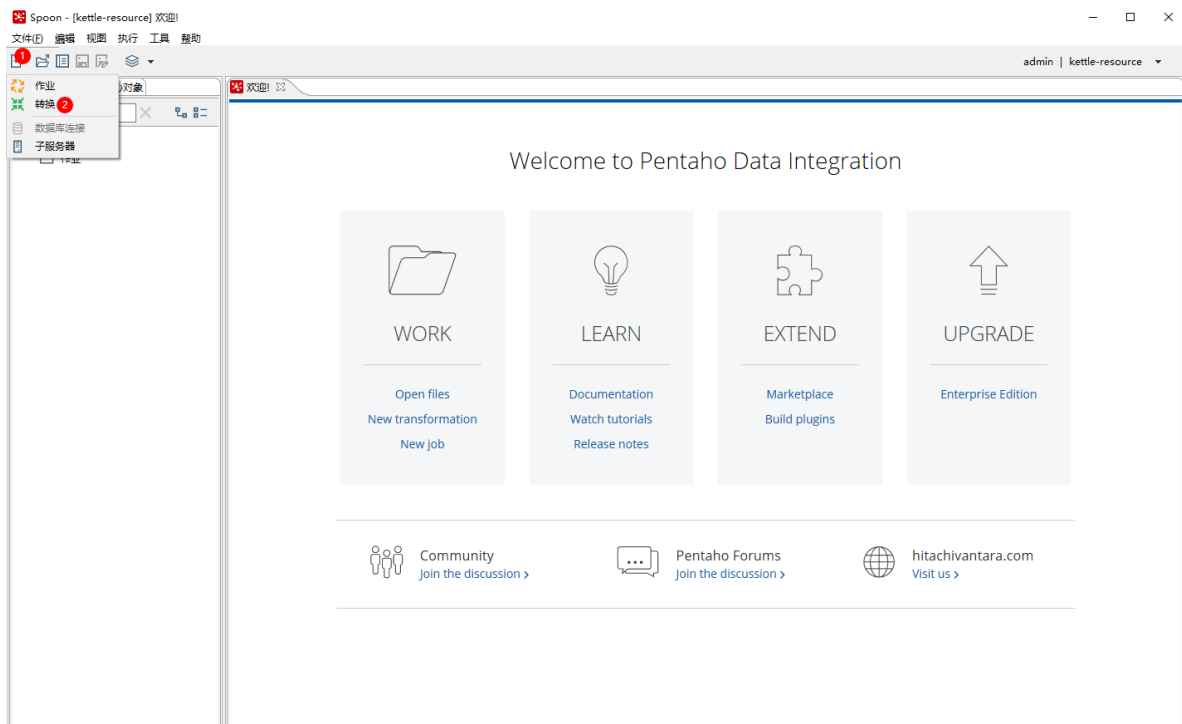


连接选项卡可以新建数据库连接，全资源库转换、作业可以公用。

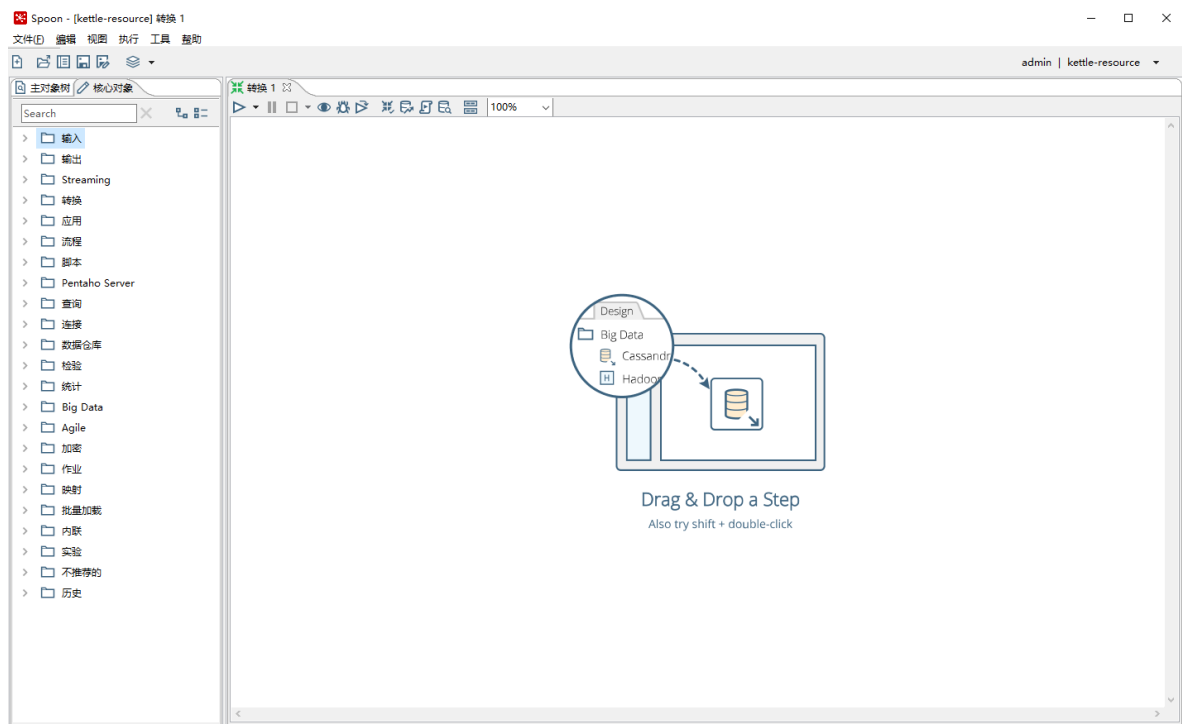


2.2 新建第一个转换

点击新建转换

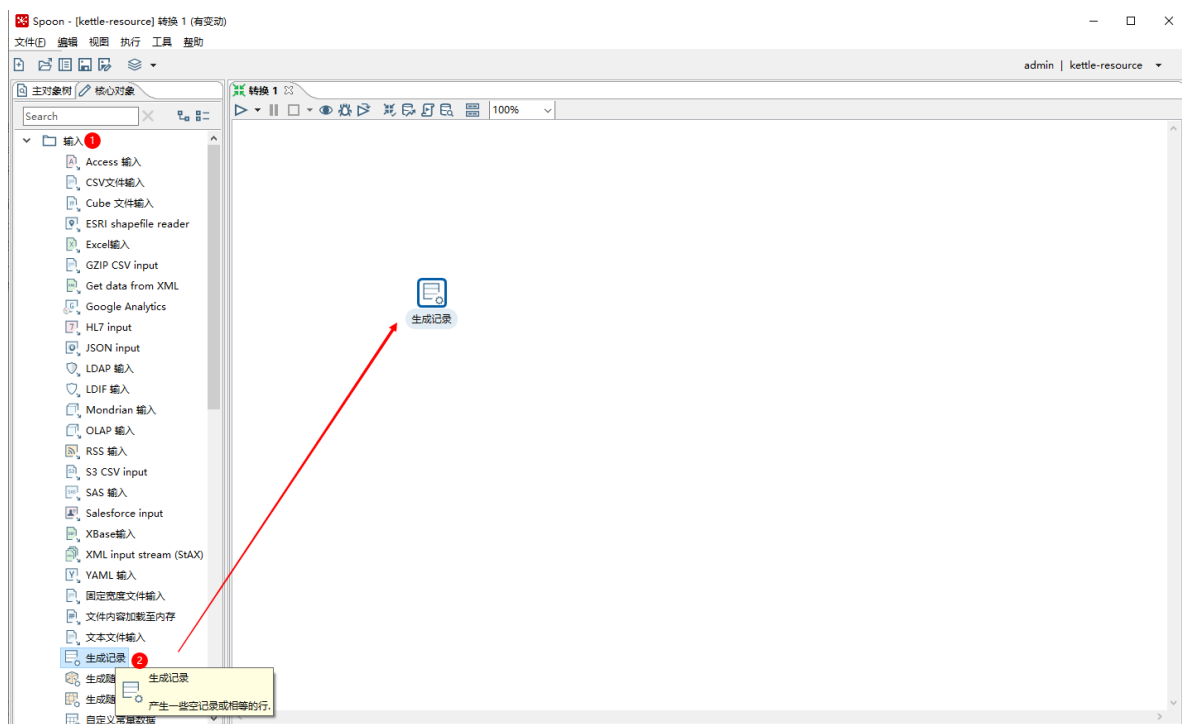


转换建好后会在工作区新建一个tab，左侧为步骤的对象，右侧是转换的工作区，秩序将左侧的步骤对象拖入工作区即可。

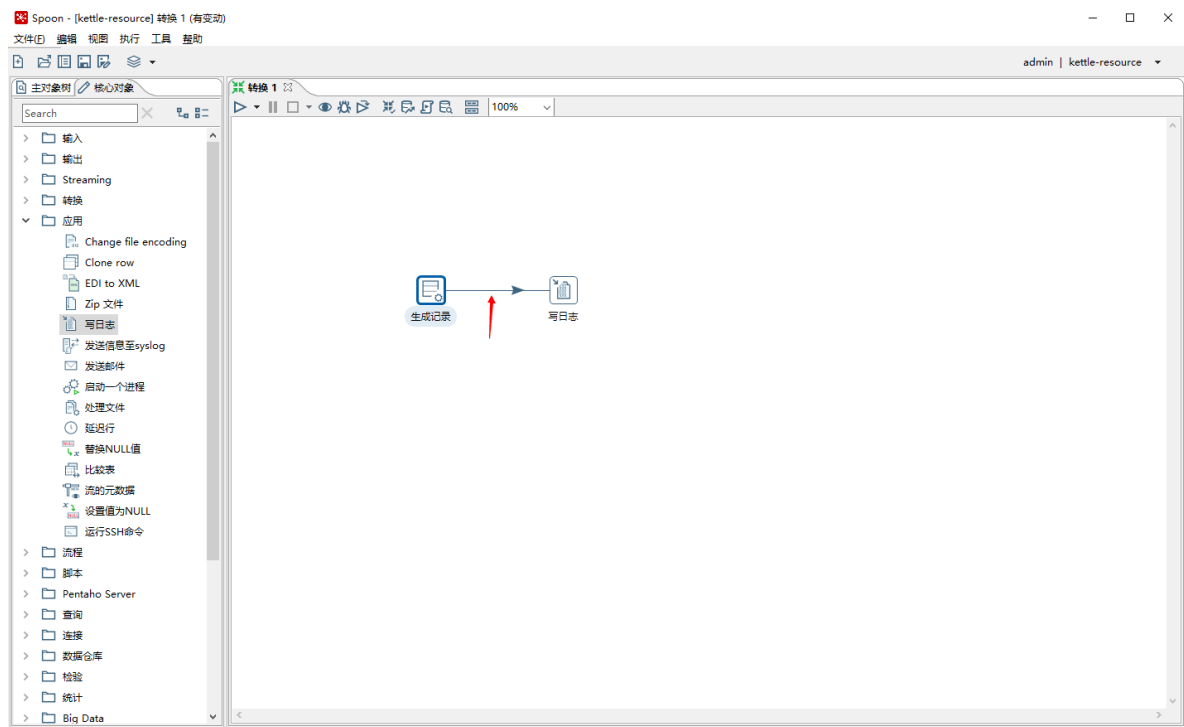


下面来演示怎么输出一个hello world

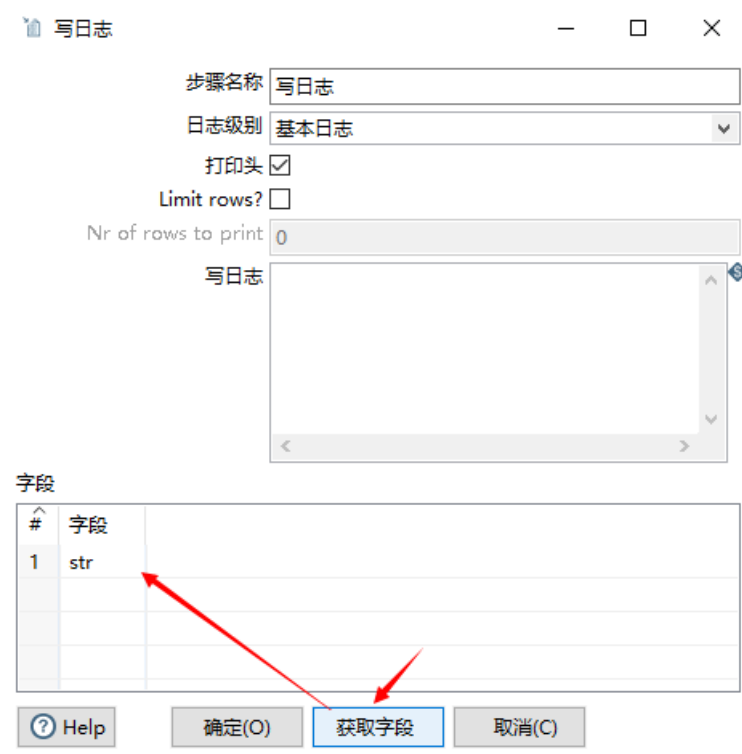
首先在输入一栏找到生成记录，拖拽到工作区



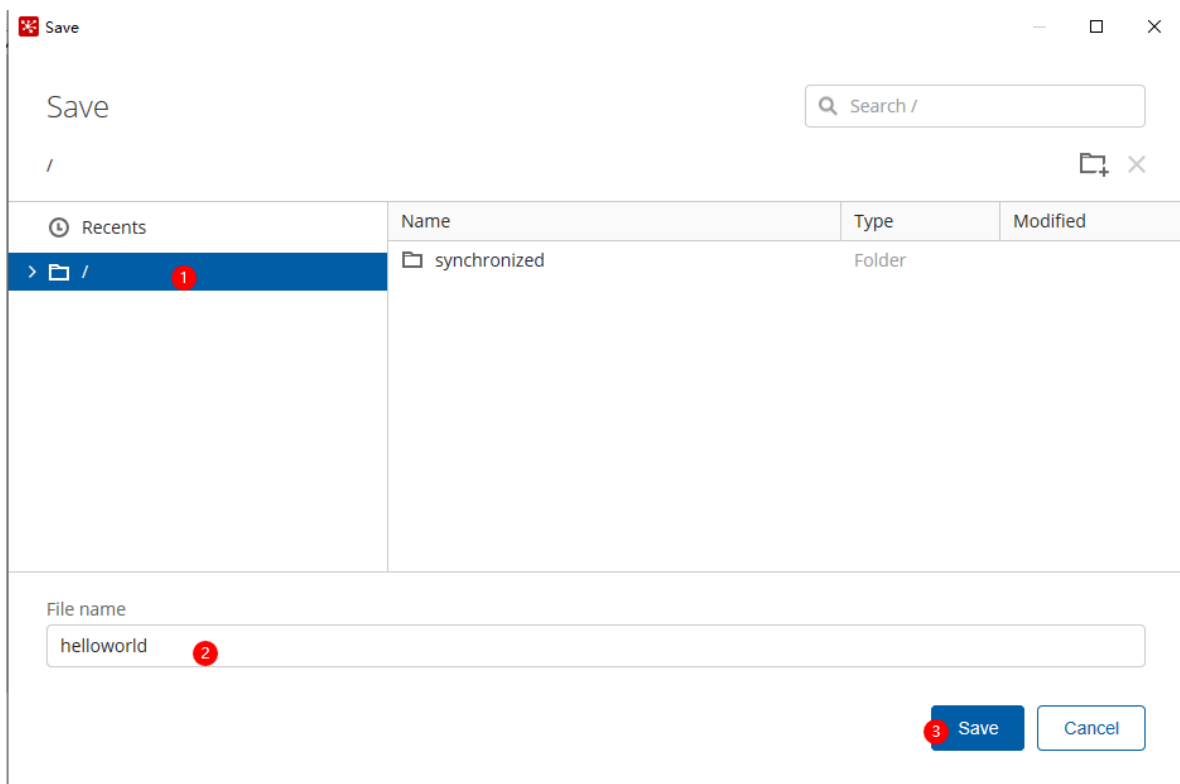
双击生成记录图标，按照如下设置。



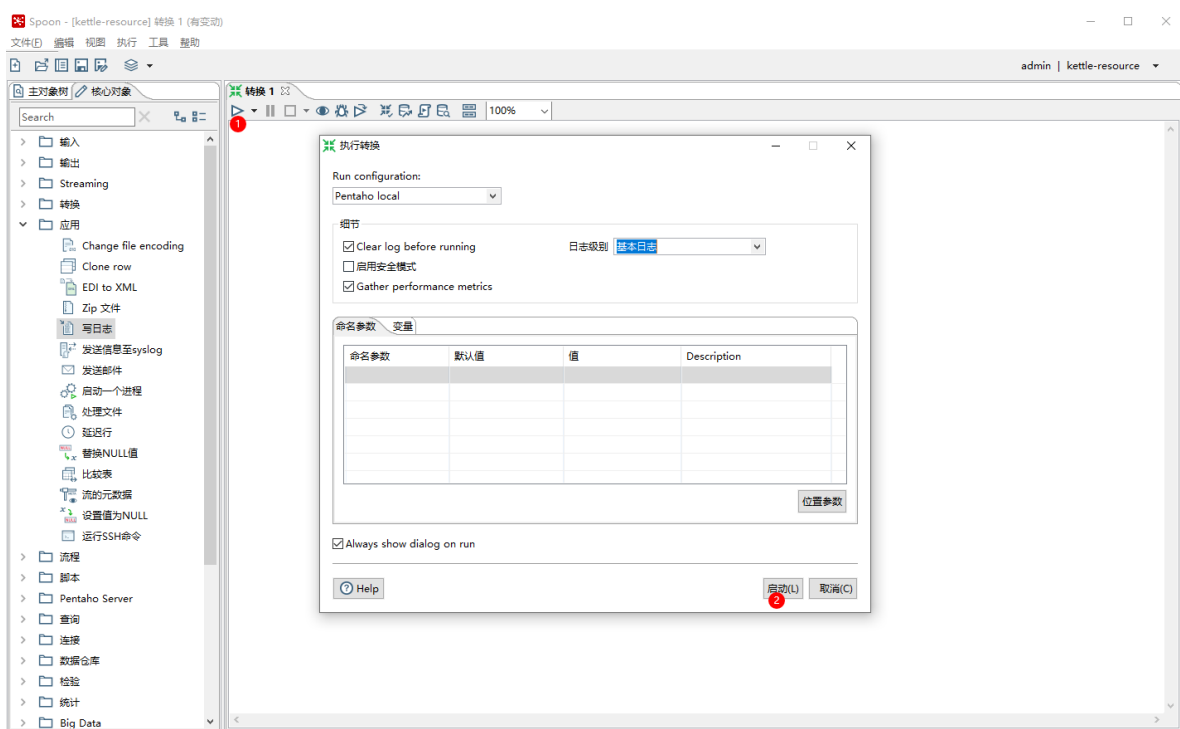
双击写日志步骤，点击获取字段即可获得到上一个步骤的字段。点击确定。



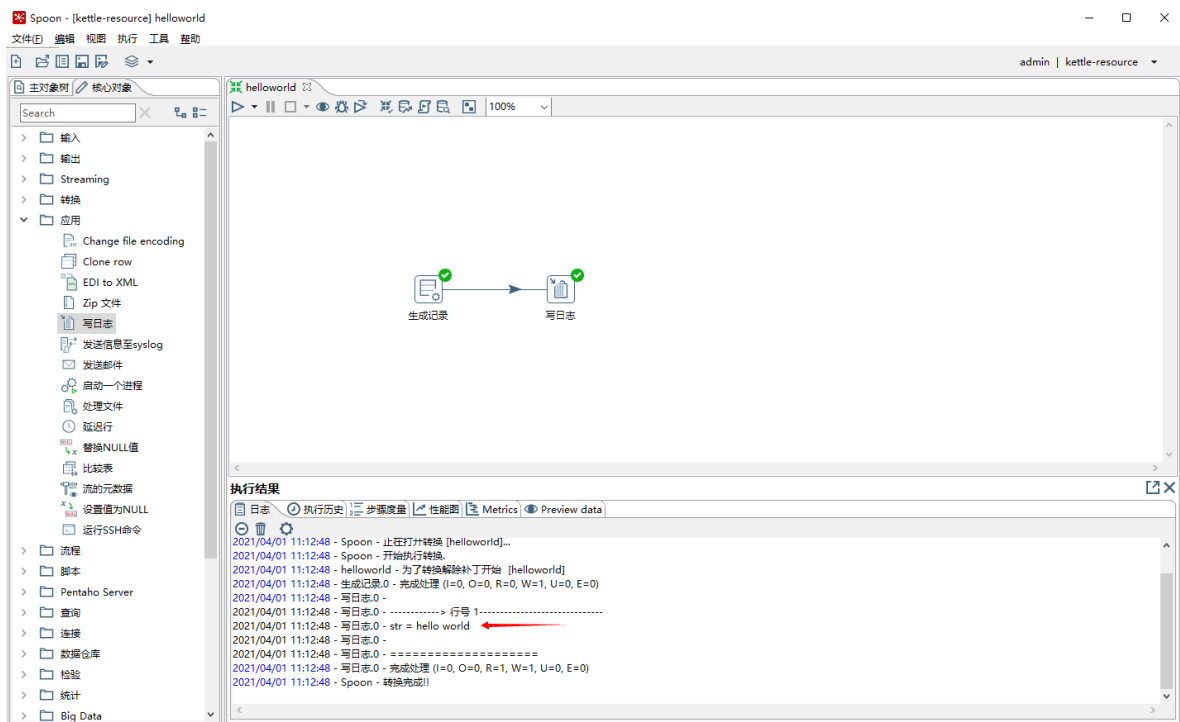
ctrl+s保存转换，可以选择保存位置，更改转换名。



点击运行

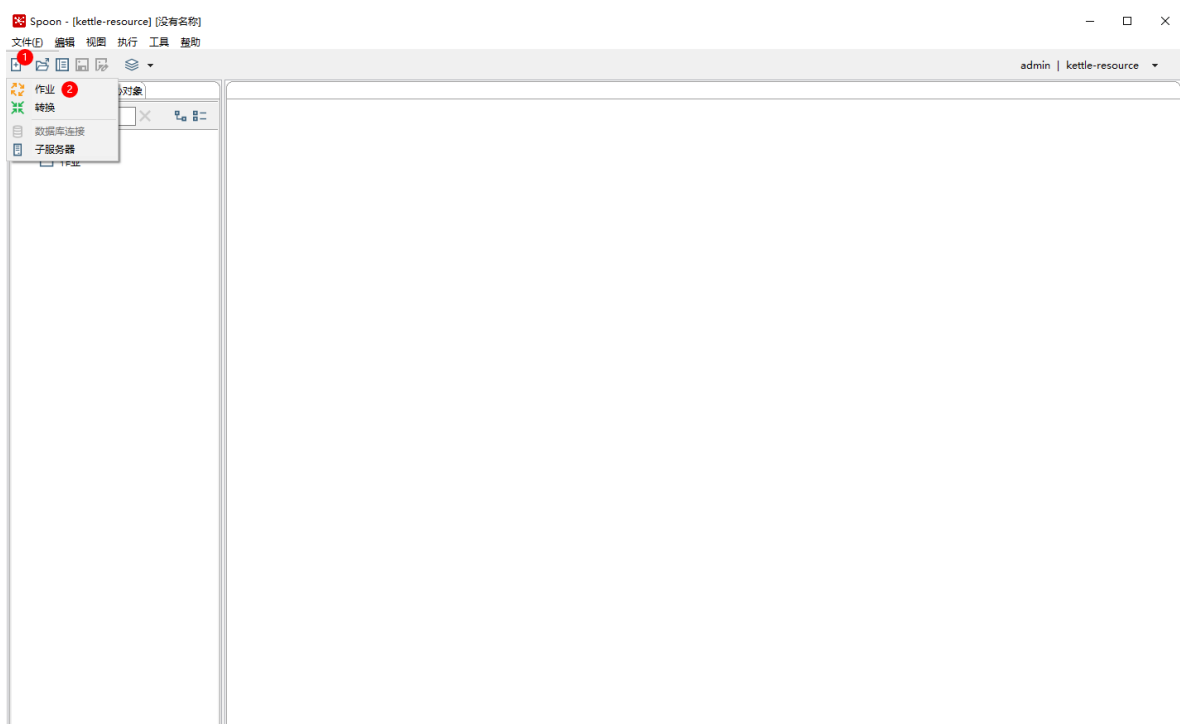


至此，第一个kettle转换就已经完成了

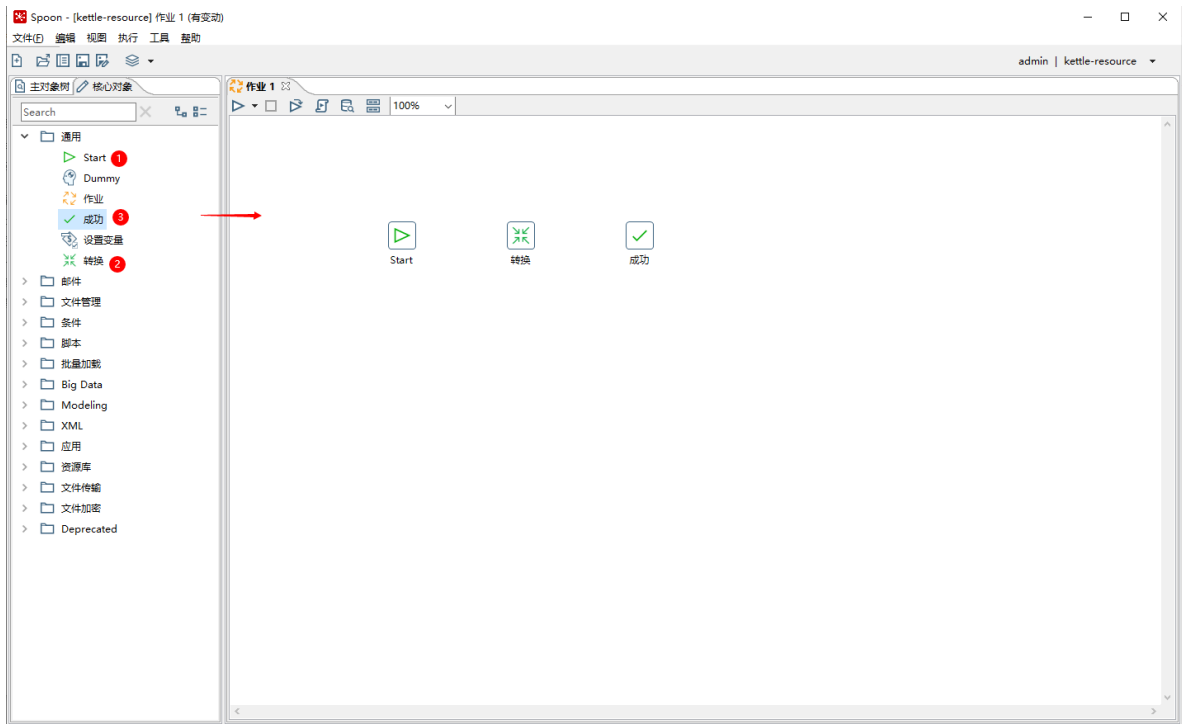


2.3 新建第一个作业

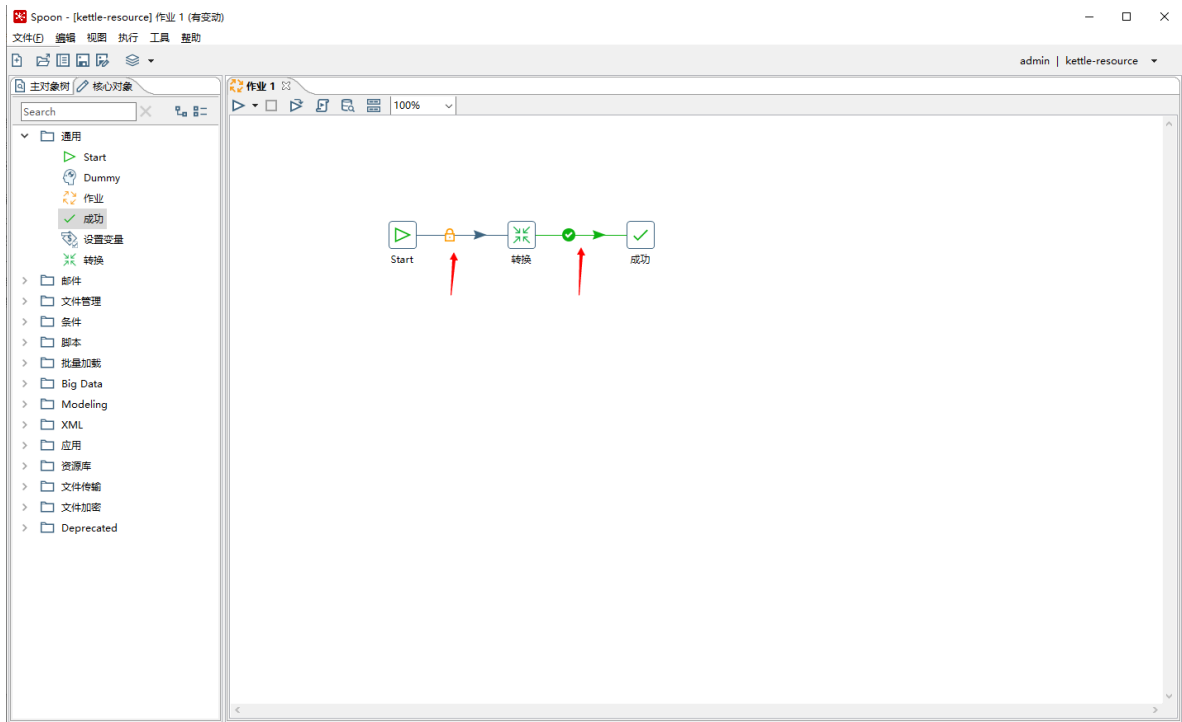
点击新建作业



将以下三个作业实体拖入



按住shift加鼠标新建两个作业跳，建好后可以点击跳更改作业跳的执行条件。



双击start，按照如图所示设置，可以每5s发起一次调用。

作业定时调度

Job entry name: Start

重复: ☒

类型: 时间间隔

以秒计算的间隔: 5

以分钟计算的间隔: 0

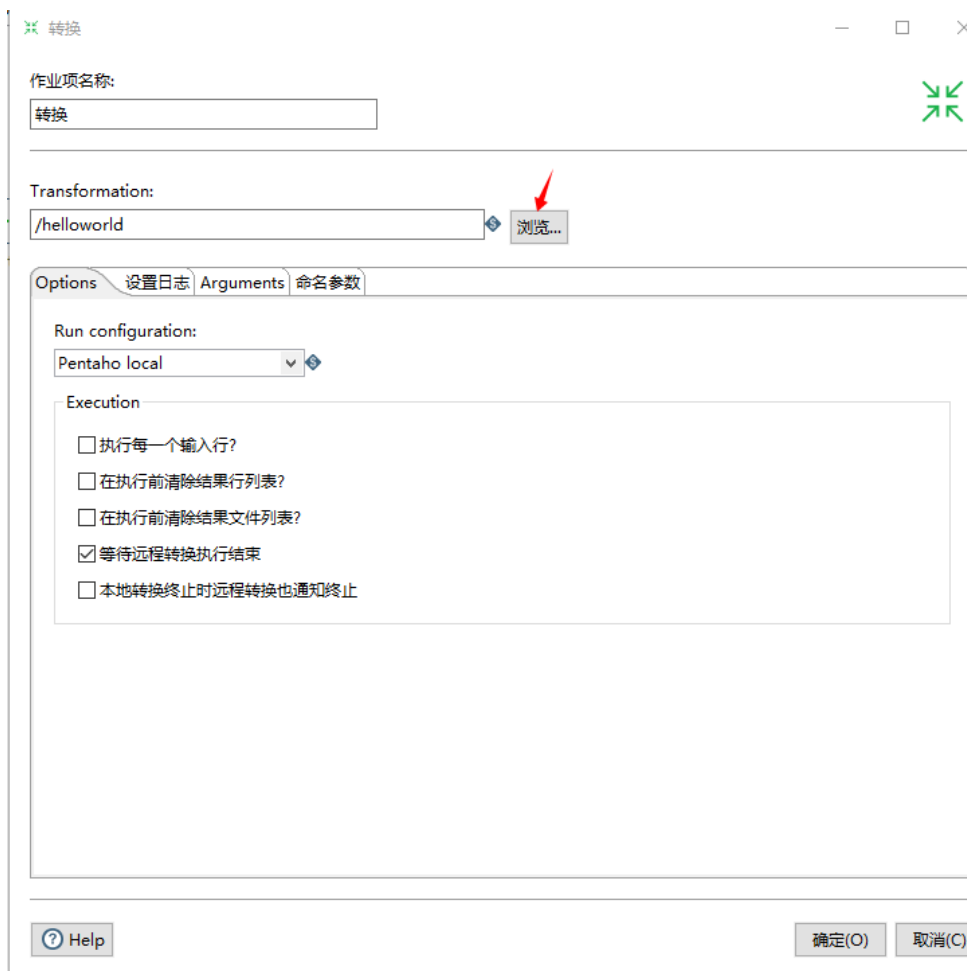
每天: 12 0

每周: 星期一

每月: 1

Help 确定(O) 取消(C)

双击转换，点击浏览选择刚才创建的helloworld转换



保存一下作业，点击运行开始运行，就可以看到在控制台每5s会有一次调用转换。

