

38 | 如何在Excel中使用SQL语言？

2019-08-30 陈旻



在进阶篇中，我们对设计范式、索引、页结构、事务以及查询优化器的原理进行了学习，了解这些可以让我们更好地使用SQL来操作RDBMS。实际上SQL的影响力远不止于此，在数据的世界里，SQL更像是一门通用的语言，虽然每种工具都会有一些自己的“方言”，但是掌握SQL可以让我们接触其它以数据为核心的工具时，更加游刃有余。

比如Excel。

你一定使用过Excel，事实上，Excel的某些部分同样支持我们使用SQL语言，那么具体该如何操作呢？

今天的课程主要包括以下几方面的内容：

1. 如何在Excel中获取外部数据源？
2. 数据透视表和数据透视图是Excel的两个重要功能，如何通过SQL查询在Excel中完成数据透视表 and 透视图？
3. 如何让Excel与MySQL进行数据交互？

如何在Excel中获取外部数据源？

使用SQL查询数据，首先需要数据源。如果我们用Excel来呈现这些数据的话，就需要先从外部导入数据源。这里介绍两种直接导入的方式：

1. 通过OLE DB接口获取外部数据源；
2. 通过Microsoft Query导入外部数据源。

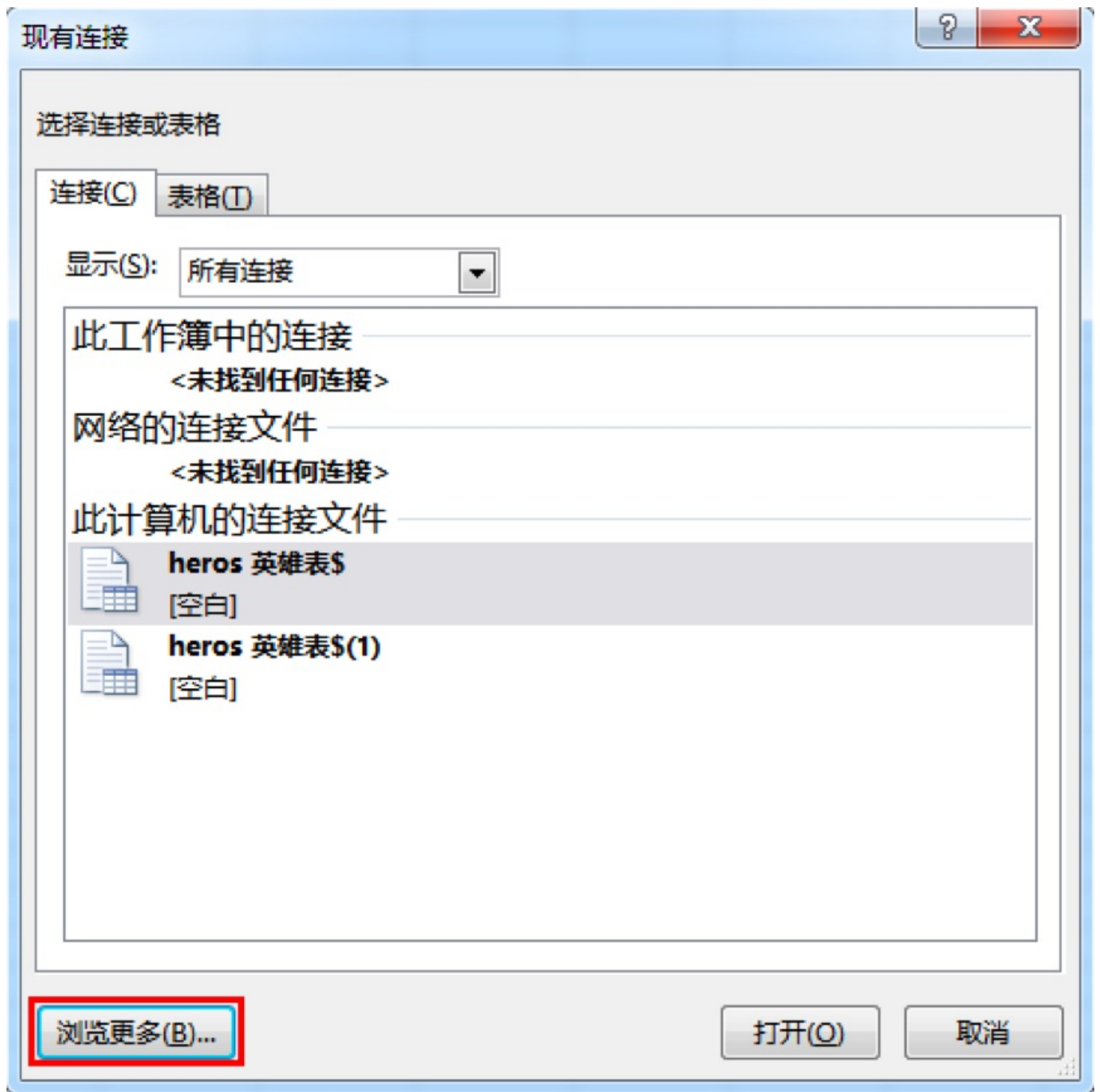
下面我们通过导入数据源heros.xlsx体验一下这两种方式，你可以从[这里](#)下载数据源。

通过OLE DB接口获取外部数据源

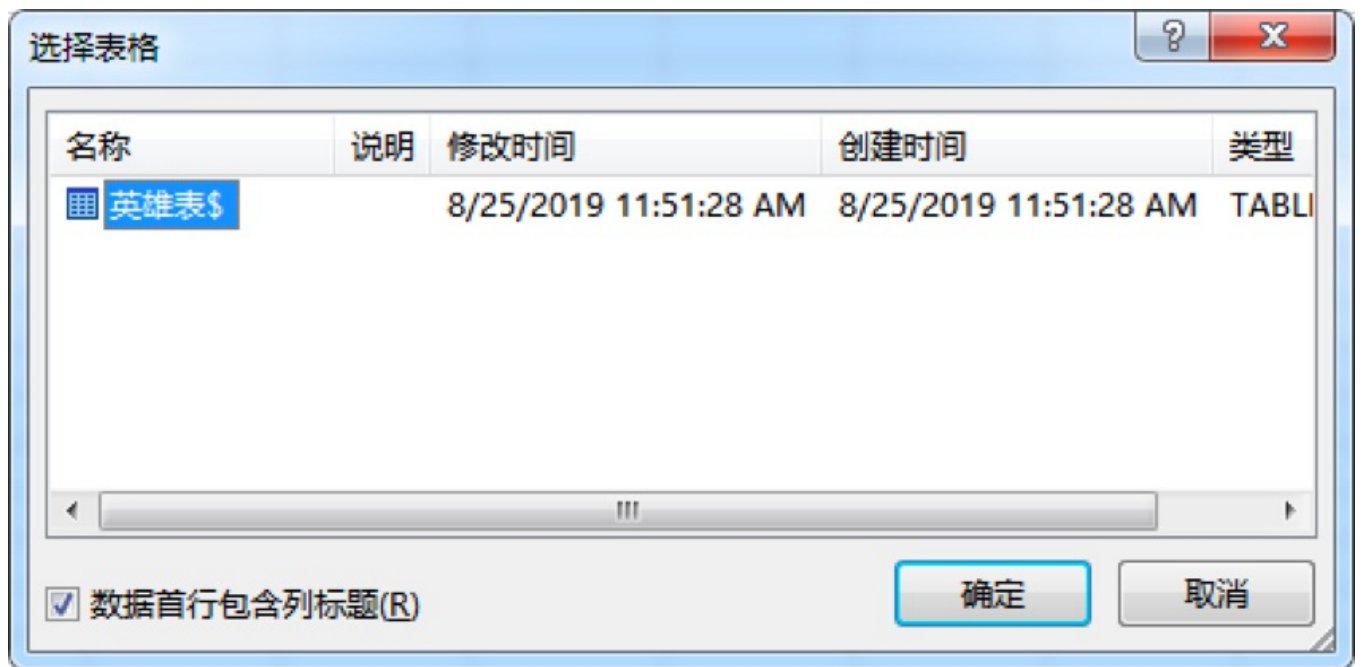
OLE的英文是Object Link and Embedding，中文意思是对象连接与嵌入，它是一种面向对象的技术。DB代表的就是数据库。OLE DB的作用就是通向不同的数据源的程序接口，方便获取外部数据，这里不仅包括ODBC，也包括其他非SQL数据类型的通路，你可以把OLE DB的作用理解成通过统一的接口来访问不同的数据源。

如果你想要在Excel中通过OLE DB接口导入数据，需要执行下面的步骤：

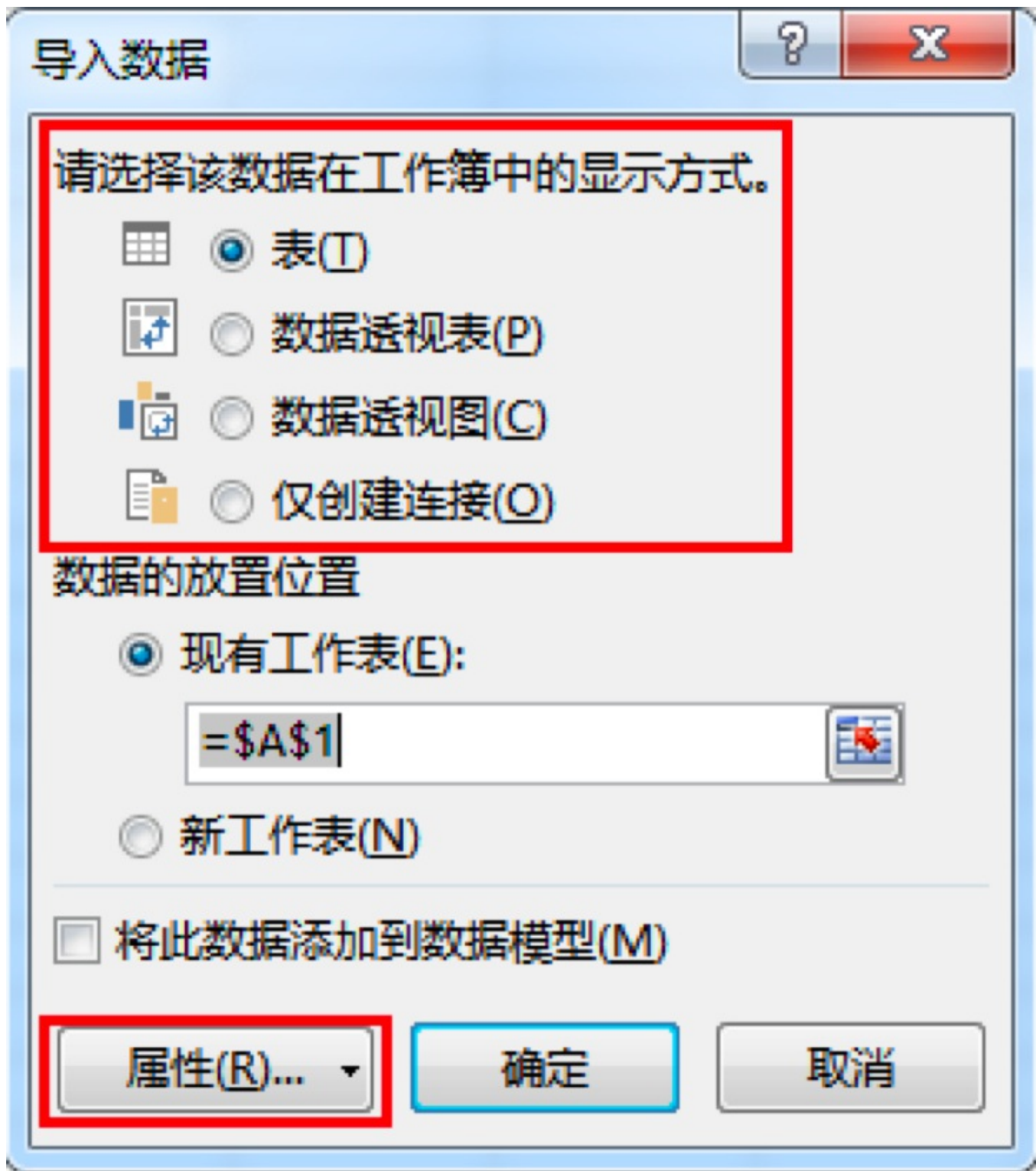
第一步，选择指定的文件。方法是通过“数据”→“现有连接”按钮选择连接。这里选择“浏览更多”，然后选择指定的xls文件。



第二步，选择指定的表格，勾选数据首行包含列标题，目的是将第一行的列名也加载进来。



第三步，通过“属性”→“定义”中的命令文本来使用SQL查询，选择我们想要的数据库，也可以将整张表直接导入到指定的位置。



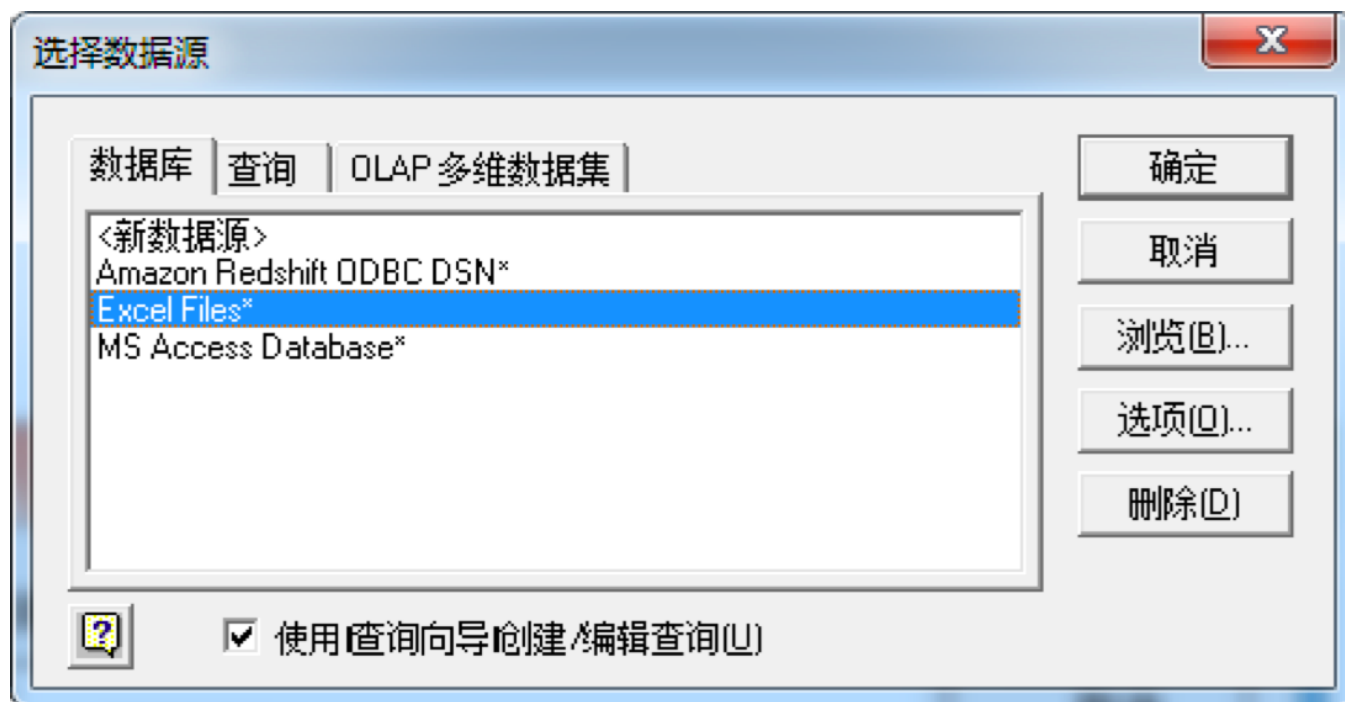
如果我们显示方式为“表”，导入全部的数据到指定的\$A\$1（代表A1单元格），那么在Excel中就可以导入整个数据表，如下图所示：

A	B	C	D	E	F	G	H	I	J	K
id	name	hp_max	hp_growth	hp_start	mp_max	mp_growth	mp_start	attack_max	attack_growth	attack_start
10000	夏侯惇	7350	288.8	3307	1746	94	430	321	11.57	159
10001	钟无艳	7000	275	3150	1760	95	430	318	11	164
10002	张飞	8341	329.4	3450	100	0	100	301	10.57	153
10003	牛魔	8476	352.8	3537	1926	104	470	273	8.775	156
10004	吕布	7344	270	3564	0	0	0	343	12.36	170
10005	亚瑟	8050	316.3	3622	0	0	0	346	13	164
10006	芈月	6164	281.5	3105	100	0	100	289	10.275	152
10007	程咬金	8611	369.6	3437	0	0	0	316	11.07	161
10008	廉颇	9328	412.1	3558	1708	92	420	286	9.225	163
10009	东皇太一	7669	319.1	3201	1926	104	470	286	9.225	163
10010	庄周	8149	345.6	3311	1694	91	420	297	9.525	170
10011	太乙真人	6835	242.3	3443	1680	90	420	284	9.75	154
10012	白起	8638	366.3	3510	1666	89	420	288	9.75	158
10013	雅典娜	6264	243	2862	1732	93	430	327	11.79	162
10014	刘邦	8073	336	3369	1940	105	470	302	10.29	158
10015	刘禅	8581	372.6	3364	1694	91	420	295	8.775	178
10016	墨子	7176	292.4	3083	1722	93	420	328	10.5	181
10017	项羽	8057	380.1	3535	1694	91	420	306	10.64	157
10018	关羽	7107	270.4	3322	10	0	10	343	12.36	170
10019	后羿	5986	200.3	3182	1784	96	440	396	16.79	161

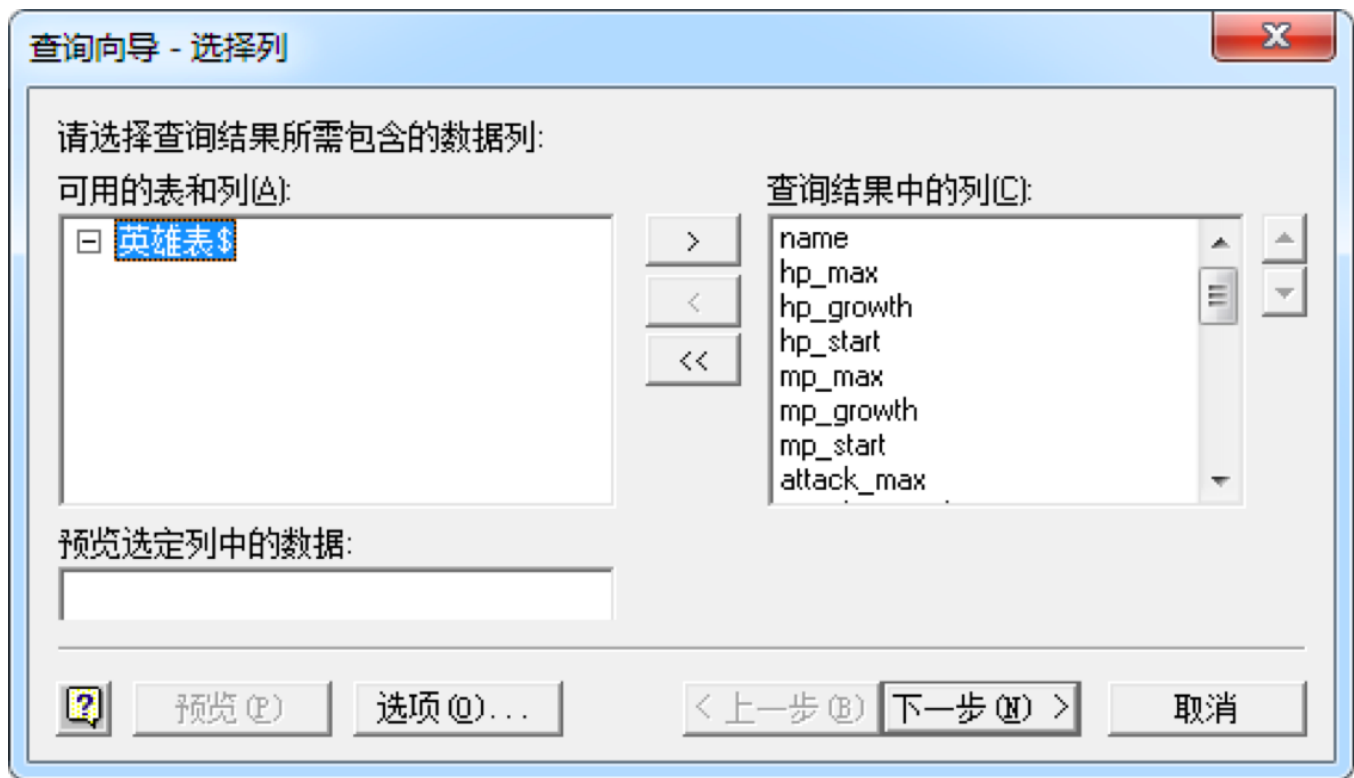
通过Microsoft Query获取外部数据源

第二种方式是利用Microsoft Query功能导入外部数据源，具体步骤如下：

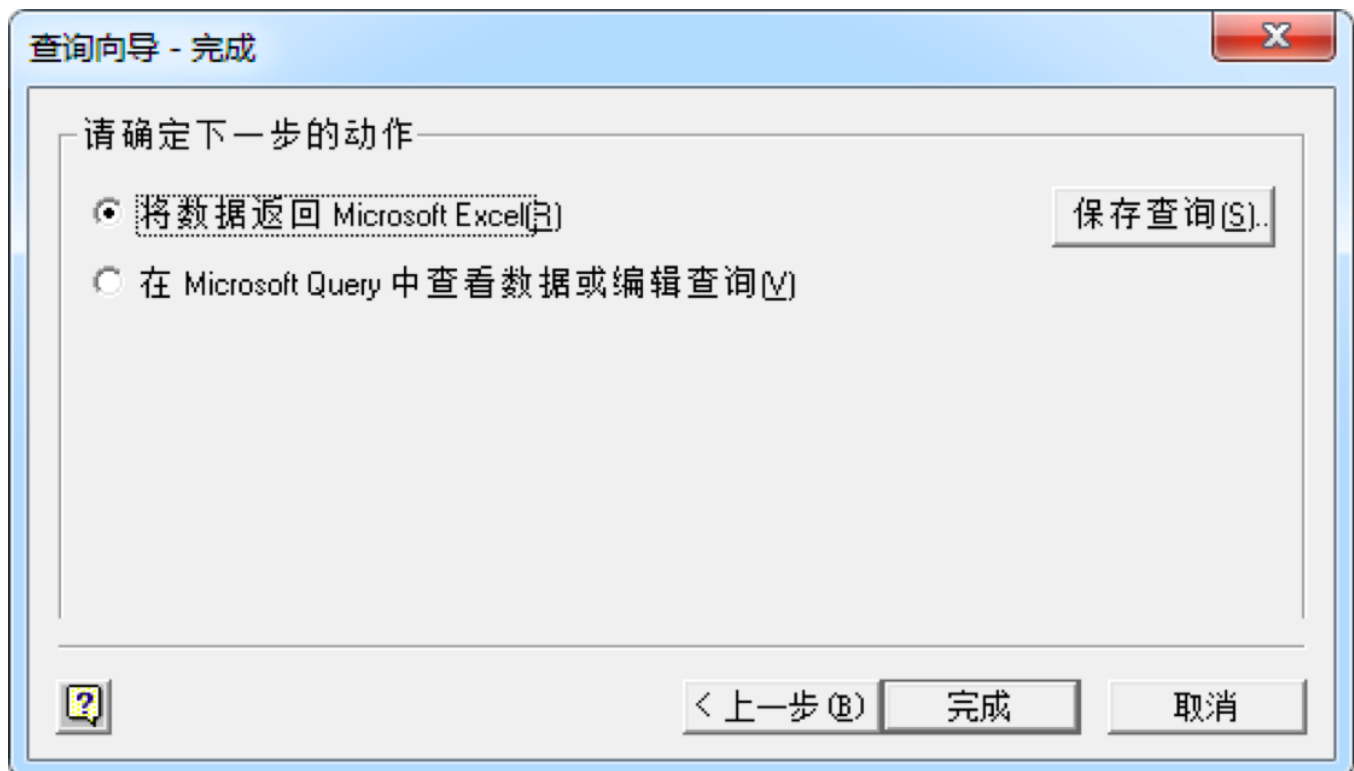
第一步，选择指定的文件。方法是通过“数据”→“获取外部数据”按钮选择数据库，这里我选择了“Excel Files”，然后选择我们想要导入的xls文件。



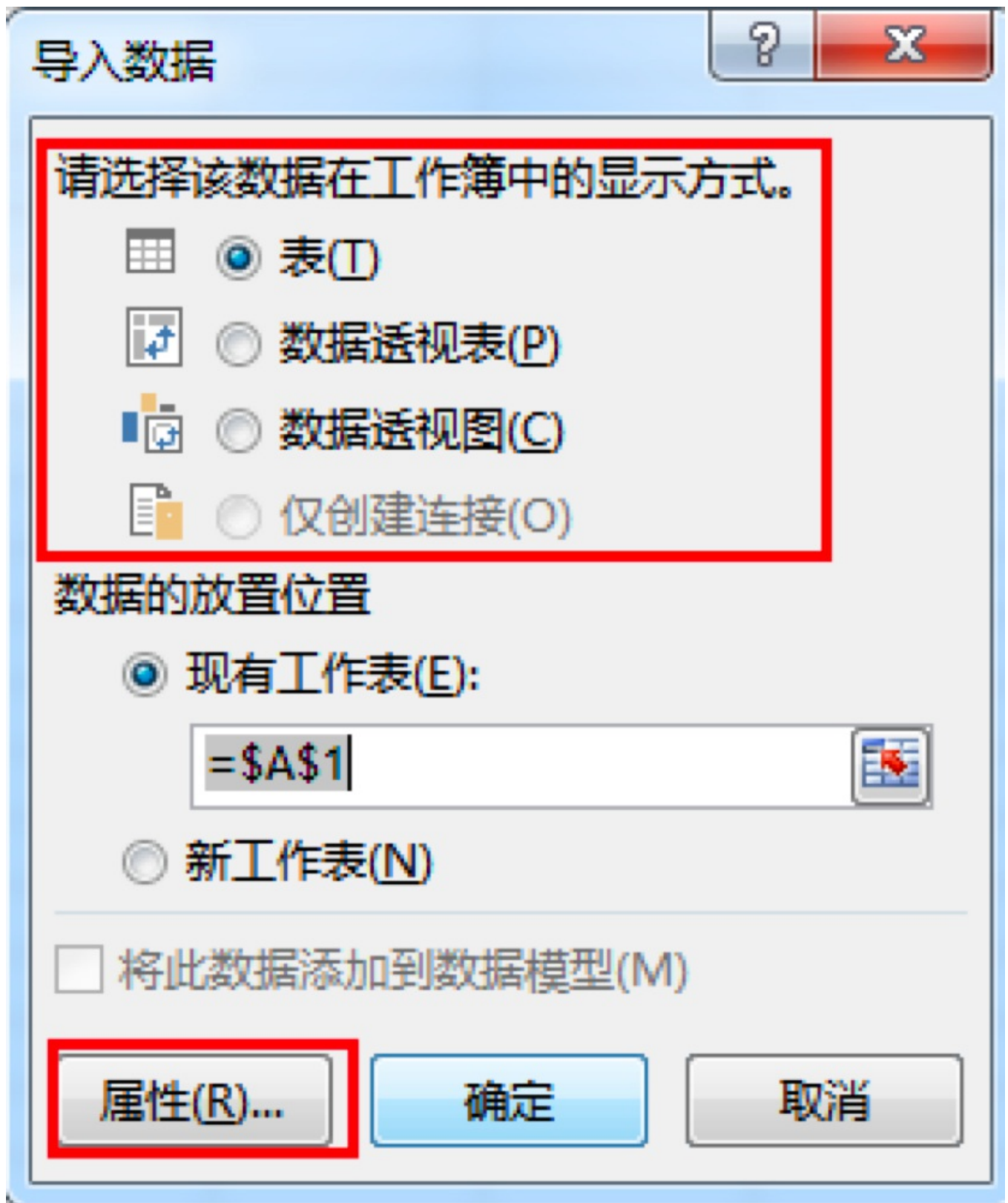
第二步。选择可用的表和列，在左侧面板中勾选我们想要导入的数据表及相应的列，点击（>）按钮导入到右侧的面板中，然后点击下一步。



最后我们可以选择“将数据返回Microsoft Excel”还是“在Microsoft Query中查看数据或编辑查询”。这里我们选择第一个选项。



当我们选择“将数据返回到Microsoft Excel”后，接下来的操作和使用OLE DB接口方式导入数据一样，可以对显示方式以及属性进行调整：



这里，我们同样选择显示方式为“表”，导入全部的数据到指定的\$A\$1(代表A1单元格)，同样会看到如下的结果：

A	B	C	D	E	F	G	H	I	J	K
id	name	hp_max	hp_growth	hp_start	mp_max	mp_growth	mp_start	attack_max	attack_growth	attack_start
10000	夏侯惇	7350	288.8	3307	1746	94	430	321	11.57	159
10001	钟无艳	7000	275	3150	1760	95	430	318	11	164
10002	张飞	8341	329.4	3450	100	0	100	301	10.57	153
10003	牛魔	8476	352.8	3537	1926	104	470	273	8.775	156
10004	吕布	7344	270	3564	0	0	0	343	12.36	170
10005	亚瑟	8050	316.3	3622	0	0	0	346	13	164
10006	芈月	6164	281.5	3105	100	0	100	289	10.275	152
10007	程咬金	8611	369.6	3437	0	0	0	316	11.07	161
10008	廉颇	9328	412.1	3558	1708	92	420	286	9.225	163
10009	东皇太一	7669	319.1	3201	1926	104	470	286	9.225	163
10010	庄周	8149	345.6	3311	1694	91	420	297	9.525	170
10011	太乙真人	6835	242.3	3443	1680	90	420	284	9.75	154
10012	白起	8638	366.3	3510	1666	89	420	288	9.75	158
10013	雅典娜	6264	243	2862	1732	93	430	327	11.79	162
10014	刘邦	8073	336	3369	1940	105	470	302	10.29	158
10015	刘禅	8581	372.6	3364	1694	91	420	295	8.775	178
10016	墨子	7176	292.4	3083	1722	93	420	328	10.5	181
10017	项羽	8057	380.1	3535	1694	91	420	306	10.64	157
10018	关羽	7107	270.4	3322	10	0	10	343	12.36	170
10019	后羿	5986	200.3	3182	1784	96	440	396	16.79	161

使用数据透视表和数据透视图做分析

通过上面的操作你也能看出来，从外部导入数据并不难，关键在于通过SQL控制想要的结果集，这里我们需要使用到Excel的数据透视表以及数据透视图的功能。

我简单介绍下数据透视表和数据透视图：

数据透视表可以快速汇总大量数据，帮助我们统计和分析数据，比如求和，计数，查看数据中的对比情况和趋势等。数据透视图则可以对数据透视表中的汇总数据进行可视化，方便我们直观地查看数据的对比与趋势等。

假设我想对主要角色（role_main）的英雄数据进行统计，分析他们平均的最大生命值（hp_max），平均的最大法力值(mp_max)，平均的最大攻击值(attack_max)，那么对应的SQL查询为：

```
SELECT role_main, avg(hp_max) AS `平均最大生命`, avg(mp_max) AS `平均最大法力`, avg(attack_max) AS `平均`
```

使用SQL+数据透视表

现在我们使用SQL查询，通过OLE DB的方式来完成数据透视表。我们在第三步的时候选择“属性”，并且在命令文本中输入相应的SQL语句，注意这里的数据表是[heros\$]，对应的命令文本为：

```
SELECT role_main, avg(hp_max) AS `平均最大生命`, avg(mp_max) AS `平均最大法力`, avg(attack_max) AS `平均`
```

连接属性

连接名称(N): heros

说明(I):

使用状况(G) 定义(D)

连接类型: Excel 文件

连接文件(F): H:\培训\SQL必知必会\L35 如何在E> 浏览(B)...

☐ 始终使用连接文件(A)

连接字符串(S): Provider=Microsoft.ACE.OLEDB.12.0;Password="";User ID=Admin;Data Source=H:\培训\SQL必知必会\L35 如何在Excel中使用SQL语言\heros.xlsx;Mode=Share Deny

☐ 保存密码(W)

命令类型(C): 表

命令文本(M): 英雄表\$

Excel Services: 验证设置(U)...

编辑查询(E)... 参数(P)... 导出连接文件(X)...

确定 取消

然后我们在右侧面板中选择“数据透视表字段”，以便对数据透视表中的字段进行管理，比如我们勾选num，role_main，平均最大生命，平均最大法力，平均最大攻击力。

数据透视表字段



选择要添加到报表的字段:



- ☒ num
- ☒ role_main
- ☒ 平均最大法力
- ☒ 平均最大攻击力
- ☒ 平均最大生命

更多表格...

最后会在Excel中呈现如下的数据透视表:

行标签	求和项:num	求和项:平均最大生命	求和项:平均最大法力	求和项:平均最大攻击力
刺客	6	6090.5	1190.333333	377.1666667
法师	19	5858.789474	1775.368421	295.1578947
辅助	6	6701.833333	1822	300.5
射手	10	5798.5	1611.6	395.5
坦克	10	8312.4	1440	297.4
战士	18	6937.388889	1081.055556	342.1666667
总计	69	39699.4117	8920.35731	2007.891228

操作视频如下:

使用SQL+数据透视图

数据透视图可以呈现可视化的形式,方便我们直观地了解数据的特征。这里我们使用SQL查询,通过Microsoft Query的方式来完成数据透视图。我们在第三步的时候选择在Microsoft Query中查看数据或编辑查询,来看下Microsoft Query的界面:

Microsoft Query

文件(F) 编辑(E) 视图(V) 格式(T) 表(B) 条件(C) 记录(R) 窗口(W) 帮助(H)

SQL

查询来自 Excel Files

heros\$

- * attack_growt
- attack_max
- attack_range
- attack_speed
- attack_start

	id	name	hp_max	hp_growth	hp_start	mp_max	mp
▶	10000.0	夏侯惇	7350.0	288.8	3307.0	1746.0	94.0
	10001.0	钟无艳	7000.0	275.0	3150.0	1760.0	95.0
	10002.0	张飞	8341.0	329.4	3450.0	100.0	0.0
	10003.0	牛魔	8476.0	352.8	3537.0	1926.0	104.0
	10004.0	吕布	7344.0	270.0	3564.0	0.0	0.0
	10005.0	亚瑟	8050.0	316.3	3622.0	0.0	0.0
	10006.0	半月	6164.0	281.5	3105.0	100.0	0.0
	10007.0	程咬金	8611.0	369.6	3437.0	0.0	0.0
	10008.0	廉颇	9328.0	412.1	3558.0	1708.0	92.0
	10009.0	东皇太一	7669.0	319.1	3201.0	1926.0	104.0
	10010.0	庄周	8149.0	345.6	3311.0	1694.0	91.0
	10011.0	太乙真人	6635.0	242.3	3443.0	1680.0	90.0
	10012.0	白起	8638.0	366.3	3510.0	1666.0	89.0
	10013.0	雅典娜	6264.0	243.0	2862.0	1732.0	93.0
	10014.0	刘邦	8073.0	336.0	3369.0	1940.0	105.0

选定“视图”菜单的“条件”，以显示/编辑条件。

然后我们点击“SQL”按钮，可以对SQL语句进行编辑，筛选我们想要的结果集，可以得到：

Microsoft Query

文件(F) 编辑(E) 视图(V) 格式(T) 表(B) 条件(C) 记录(R) 窗口(W) 帮助(H)

SQL

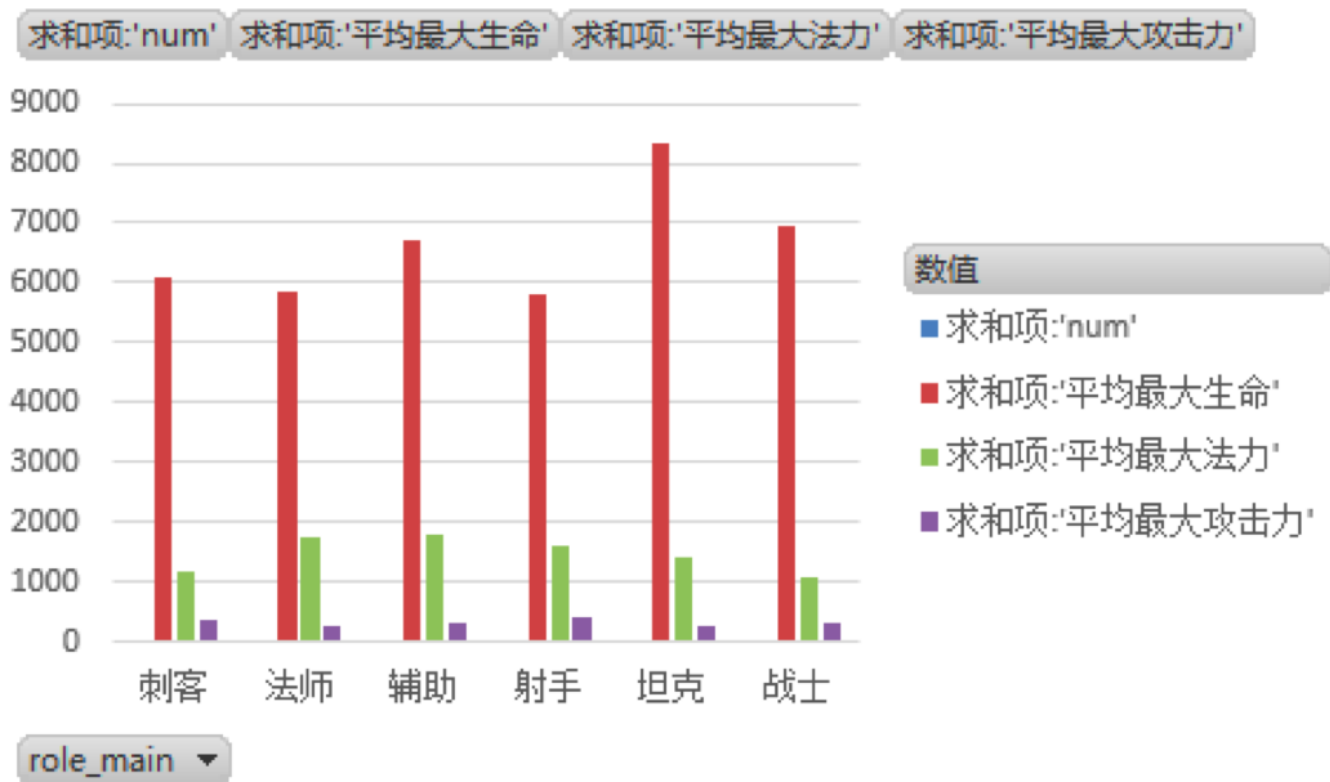
查询来自 Excel Files

heros\$

- * attack_growt
- attack_max
- attack_range
- attack_speed
- attack_start

	role_main	平均最大生命	平均最大法力	平均最大攻击力	num
▶	射手	5798.5	1611.6	395.5	10
	战士	6937.38888888889	1081.05555555556	342.166666666667	18
	刺客	6090.5	1190.33333333333	377.166666666667	6
	坦克	8312.4	1440.0	297.4	10
	辅助	6701.83333333333	1822.0	300.5	6
	法师	5858.78947368421	1775.36842105263	295.157894736842	19

然后选择“将数据返回Microsoft Excel”，在返回时选择“数据透视图”，然后在右侧选择数据透视图的字段，就可以得到下面这张图：



你可以看到使用起来还是很方便。

具体操作视频如下：

让Excel与MySQL进行数据交互


刚才我们讲解的是如何从Excel中导入外部的xls文件数据，并在Excel实现数据透视表和数据透视图的呈现。实际上，Excel也可以与MySQL进行数据交互，这里我们需要使用到MySQL for Excel插件：

下载mysql-for-excel并安装，地址：<https://dev.mysql.com/downloads/windows/excel/>

下载mysql-connector-odbc并安装，地址：<https://dev.mysql.com/downloads/connector/odbc/>

这次我们的任务是给数据表增加一个last_name字段，并且使用Excel的自动填充功能来填充好英雄的姓氏。

第一步，连接MySQL。打开一个新的Excel文件的时候，会在“数据”面板中看到MySQL for Excel的插件，点击后可以打开MySQL的连接界面，如下：



Please enter the password for the connection:

Connection: Local instance MySQL80 - Mysql@localhost:3306

User: root

Password:

☐ Store password securely?

OK

Cancel

第二步，导入heros数据表。输入密码后，我们在右侧选择想要的数据表heros，然后选择Import MySQL Data导入数据表的导入，结果如下：

A	B	C	D	E	F	G	H	I	J	K
id	name	hp_max	hp_growth	hp_start	mp_max	mp_growth	mp_start	attack_max	attack_growth	attack_start
10000	夏侯惇	7350	288.8	3307	1746	94	430	321	11.57	159
10001	钟无艳	7000	275	3150	1760	95	430	318	11	164
10002	张飞	8341	329.4	3450	100	0	100	301	10.57	153
10003	牛魔	8476	352.8	3537	1926	104	470	273	8.775	156
10004	吕布	7344	270	3564	0	0	0	343	12.36	170
10005	亚瑟	8050	316.3	3622	0	0	0	346	13	164
10006	半月	6164	281.5	3105	100	0	100	289	10.275	152
10007	程咬金	8611	369.6	3437	0	0	0	316	11.07	161
10008	廉颇	9328	412.1	3558	1708	92	420	286	9.225	163
10009	东皇太一	7669	319.1	3201	1926	104	470	286	9.225	163
10010	庄周	8149	345.6	3311	1694	91	420	297	9.525	170
10011	太乙真人	6835	242.3	3443	1680	90	420	284	9.75	154
10012	白起	8638	366.3	3510	1666	89	420	288	9.75	158
10013	雅典娜	6264	243	2862	1732	93	430	327	11.79	162
10014	刘邦	8073	336	3369	1940	105	470	302	10.29	158
10015	刘禅	8581	372.6	3364	1694	91	420	295	8.775	178
10016	墨子	7176	292.4	3083	1722	93	420	328	10.5	181
10017	项羽	8057	380.1	3535	1694	91	420	306	10.64	157
10018	关羽	7107	270.4	3322	10	0	10	343	12.36	170
10019	后羿	5986	200.3	3182	1784	96	440	396	16.79	161

第三步，创建last_name字段，使用Excel的自动填充功能来进行姓氏的填写（Excel自带的“自动填充”可以帮我们智能填充一些数据），完成之后如下图所示：

A	B	C	D	E	F	G	H	I	J	K
id	name	last_r	hp_max	hp_growth	hp_start	mp_max	mp_growth	mp_start	attack_max	attack_growth
10000	夏侯惇	夏	7350	288.8	3307	1746	94	430	321	11.57
10001	钟无艳	钟	7000	275	3150	1760	95	430	318	11
10002	张飞	张	8341	329.4	3450	100	0	100	301	10.57
10003	牛魔	牛	8476	352.8	3537	1926	104	470	273	8.775
10004	吕布	吕	7344	270	3564	0	0	0	343	12.36
10005	亚瑟	亚	8050	316.3	3622	0	0	0	346	13
10006	半月	半	6164	281.5	3105	100	0	100	289	10.275
10007	程咬金	程	8611	369.6	3437	0	0	0	316	11.07
10008	廉颇	廉	9328	412.1	3558	1708	92	420	286	9.225
10009	东皇太一	东	7669	319.1	3201	1926	104	470	286	9.225
10010	庄周	庄	8149	345.6	3311	1694	91	420	297	9.525
10011	太乙真人	太	6835	242.3	3443	1680	90	420	284	9.75
10012	白起	白	8638	366.3	3510	1666	89	420	288	9.75
10013	雅典娜	雅	6264	243	2862	1732	93	430	327	11.79
10014	刘邦	刘	8073	336	3369	1940	105	470	302	10.29
10015	刘禅	刘	8581	372.6	3364	1694	91	420	295	8.775
10016	墨子	墨	7176	292.4	3083	1722	93	420	328	10.5
10017	项羽	项	8057	380.1	3535	1694	91	420	306	10.64
10018	关羽	关	7107	270.4	3322	10	0	10	343	12.36
10019	后羿	后	5986	200.3	3182	1784	96	440	396	16.79

第四步，将修改好的Excel表导入到MySQL中，创建一个新表heros_xls。选中整个数据表（包括数据行及列名），然后在右侧选择“Export Excel Data to New Table”。这时在MySQL中你就能看到相应的数据表heros_xls了，我们在MySQL中使用SQL进行查询：

```
mysql > SELECT * FROM heros_xls
```

运行结果（69条记录）：

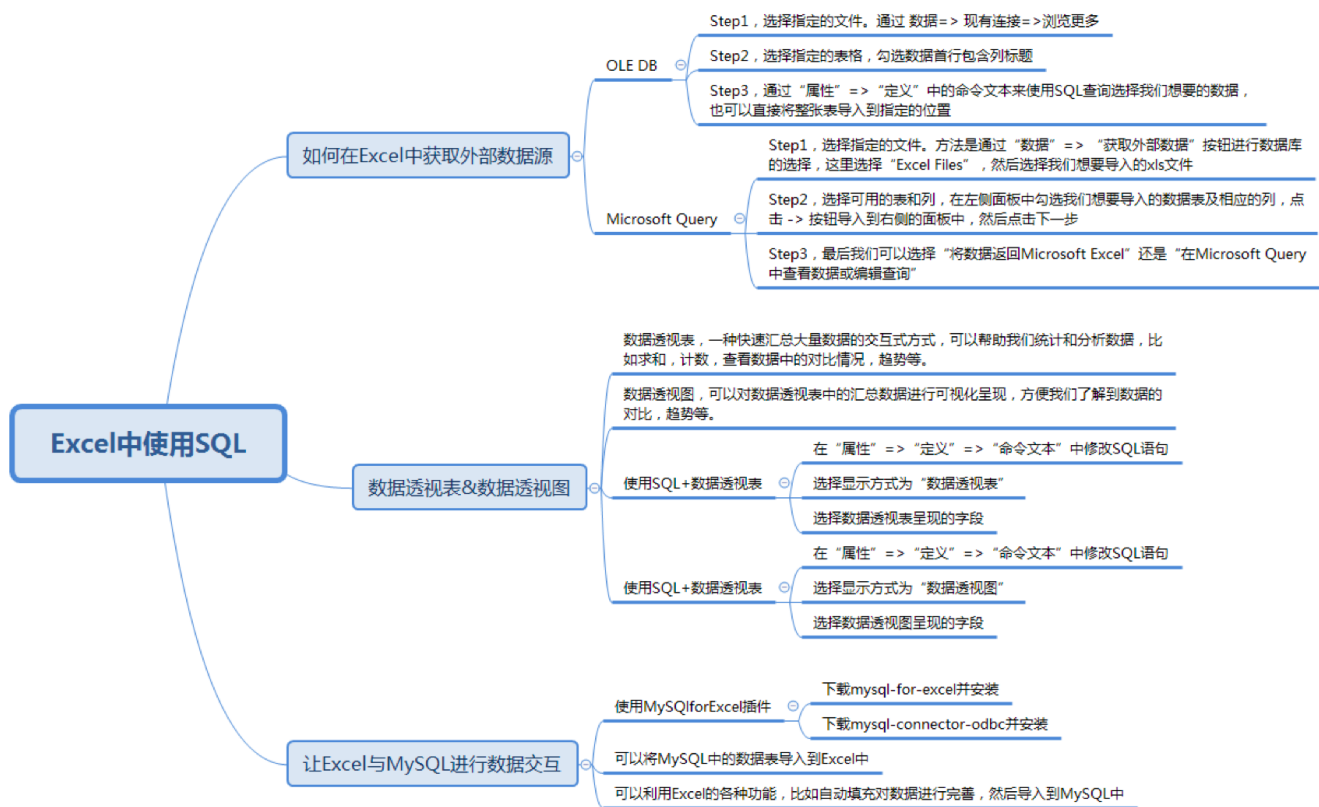
id	name	last_name	hp_max	mp_max	attack_max	defense_max
10000	夏侯惇	夏	7350	1746	321	397
10001	钟无艳	钟	7000	1760	318	409
10002	张飞	张	8341	100	301	504
.....
10068	百里守约	百	5611	1784	410	329

需要说明的是，有时候自动填充功能并不完全准确，我们还需要对某些数据行的last_name进行修改，比如“夏侯惇”的姓氏应该改成“夏侯”，“百里守约”改成“百里”等。

总结

我们今天讲解了如何在Excel中使用SQL进行查询，在这个过程中你应该对”SQL定义了查询的标准“更有体会。SQL使得各种工具可以遵守SQL语言的标准（当然也有各自的方言）。

如果你已经是个SQL高手，你会发现原来SQL和Excel还可以如此“亲密”。Excel作为使用人数非常多的办公软件，提供了SQL查询会让我们操作起来非常方便。如果你还没有使用过Excel的这些功能，那么就赶快来用一下吧。



SQL作为一门结构化查询语言，具有很好的通用性，你还在其他工具中使用过SQL语言吗？如果有的话可以分享一下你的体会。

最后留一道动手题吧。你可以创建一个新的xls文件，导入heros.xlsx数据表，用数据透视图的方式对英雄主要定位为刺客、法师、射手的英雄数值进行可视化，数据查询方式使用SQL查询，统计的英雄数值为平均生命成长hp_growth，平均法力成长mp_growth，平均攻击力成长attack_growth。

欢迎你在评论区写下你的体会与思考，也欢迎把这篇文章分享给你的朋友或者同事，一起来交流。

SQL 必知必会

从入门到数据实战

陈旻

清华大学计算机博士



新版升级：点击「👤请朋友读」，20位好友免费读，邀请订阅更有**现金**奖励。

精选留言



莫弹弹

👍 3

老师我发现安装 mysql-for-excel-1.3.8.msi 时会报错

The Microsoft Visual Studio Tools for Office Runtime must be installed prior to running this installation.

找了一下微软的说明

<https://docs.microsoft.com/en-us/visualstudio/vsto/how-to-install-the-visual-studio-tools-for-office-runtime-redistributable?view=vs-2019>

需要下载 Visual Studio 2010 Tools for Office Runtime 才能运行

下载链接在这里

<https://www.microsoft.com/zh-CN/download/confirmation.aspx?id=56961>

2019-08-31



Monday

👍 2

这个功能犹如AWT

2019-09-06

作者回复

有些类似，提供了图形化界面

2019-09-09



莫弹弹

👍 2

这个功能也是比较好的报表工具了，产品经理经常让我提取一些数据出来，例如指定时间段的

数据，指定用户的业务记录数量，如果有Excel插件和MySQL配合的话，可以实现我从MySQL导出Excel报表的功能，非常方便

2019-08-31

作者回复

是的，非常简单方便

2019-09-09



ttttt

1

Mac 用户说我太难了

2019-09-05



往事随风，顺其自然

1

为什么我在用Microsoft query导入表中数据报错，数据源中没有包含可见的表格，这个怎么解决？

2019-09-01



许童童

1

日常编程的话确实用不到，不过工作中，如果想临时生成一下可视化的图表，用这种方式那确实是很方便的，感谢老师分享。

2019-08-30

作者回复

是的 简单实用

2019-09-09



Geek_aab935

0

老师好，请教一个问题，我要建一张80万条的数据表，数据来自3张表的汇总，我用create table 表名 as 方式建表太慢，有什么好的方式快速建表？

2019-08-30



jonnypppp

0

平时基本不用，不过涨知识了

2019-08-30

作者回复

谢谢

2019-09-09