



链滴

# MySQL字符集与排序规则

作者: [wangsch](#)

原文链接: <https://ld246.com/article/1467368350395>

来源网站: 链滴

许可协议: [署名-相同方式共享 4.0 国际 \(CC BY-SA 4.0\)](#)

<p>什么是字符集 (Character Set) 和排序规则 (collation) </p>  
<p><a href="http://dev.mysql.com/doc/refman/5.6/en/charset-general.html">http://dev.mysql.com/doc/refman/5.6/en/charset-general.html</a> </p>  
<p>&nbsp;</p>  
<p>查看MySQL支持的字符集和默认排序规则</p>  
<p>mysql> show character set;</p>  
<p>mysql> show collation;</p>  
<p>mysql> show collation where Charset like 'utf8mb4';</p>  
<p>&nbsp;</p>  
<p>排序规则名称规则</p>  
<p>MySQL collation names follow these rules:<br />&bull; A name ending in \_ci indicates a case-insensitive collation.<br />&bull; A name ending in \_cs indicates a case-sensitive collation<br />&bull; A name ending in \_bin indicates a binary collation. Character comparisons are based on character<br />binary code values.<br />&bull; Unicode collation names may include a version number to indicate the version of the Unicode Collation<br />Algorithm (UCA) on which the collation is based. UCA-based collations without a version number in the<br />name use the version-4.0.0 UCA weight keys: http://www.unicode.org/Public/UCA/4.0.0/allkeys-4.0.0.txt.<br />A collation name such as utf8\_unicode\_520\_ci is based on UCA 5.2.0 weight keys: http://<br />www.unicode.org/Public/UCA/5.2.0/allkeys.txt.</p>  
<p>&nbsp;</p>  
<p>utf8mb4字符集</p>  
<p>The utf8mb4 Character Set (4-Byte UTF-8 Unicode Encoding)<br />The character set named utf8 uses a maximum of three bytes per character and contains only<br />BMP characters. The utf8mb4 character set uses a maximum of four bytes per character supports<br /><br />supplemental characters:<br />&bull; For a BMP character, utf8 and utf8mb4 have identical storage characteristics: same code values,<br />same encoding, same length.<br />&bull; For a supplementary character, utf8 cannot store the character at all, while utf8mb4 requires four<br />bytes to store it. Since utf8 cannot store the character at all, you do not have any supplementary<br />characters in utf8 columns and you need not worry about converting characters or losing data when<br />upgrading utf8 data from older versions of MySQL.<br /><br />utf8mb4 is a superset of utf8, so for an operation such as the following concatenation, the result is<br />character set utf8mb4 and the collation of utf8mb4\_col:<br />SELECT CONCAT(utf8\_col, utf8mb4\_col);<br /><br />Similarly, the following comparison in the WHERE clause works according to the collation of utf8mb4\_col:<br />SELECT \* FROM utf8\_tbl, utf8mb4\_tbl WHERE utf8\_tbl.utf8\_col = utf8mb4\_tbl.utf8mb4\_col;<br /><br />Tip: To save space with utf8mb4, use VARCHAR instead of CHAR. Otherwise, MySQL must reserve four<br />bytes for each character in a CHAR CHARACTER SET utf8mb4 column because that is the maximum<br />possible length. For example, MySQL must reserve 40 bytes for a CHAR(10) CHARACTER SET utf8mb4<br />column.</p>  
<p>utf8mb4字符集与emoji</p>  
<p><a href="http://www.tuicool.com/articles/zAnEV3">http://www.tuicool.com/articles/zAnEV3</a> </p>  
<p>&nbsp;</p>  
<p>字符集详解</p>  
<p>http://www.jb51.net/article/29960.htm</p>  
<p>&nbsp;</p>  
<p>MySQL处理一个请求，编解码流程</p>  
<p>http://stackoverflow.com/questions/1566602/is-set-character-set-utf8-necessary</p>  
<p>&nbsp;</p>  
<p>character\_set\_connection系统变量的作用</p>  
<p>http://stackoverflow.com/questions/16082480/what-is-the-purpose-of-character-set-connection</p>  
<p>&nbsp;</p>  
<p>官方手册关于字符集的介绍</p>

<http://dev.mysql.com/doc/refman/5.7/en/charset.html>

&nbsp;

这个帖子告诉我，支持emoji不只是修改表的字符集那么简单

[http://blog.csdn.net/e\\_laoye/article/details/51544161](http://blog.csdn.net/e_laoye/article/details/51544161)

&nbsp;

编辑my.cnf

```
character-set-server=utf8  
collation-server=utf8_bin  
init-connect='SET NAME utf8'
```

&nbsp;

# 数据库默认字符集

```
default-character-set=utf8
```

&nbsp;

重启

&nbsp;

**字符集查看与修改**

```
show variables like 'character_set_%';
```

```
+-----+-----+  
| Variable_name | Value |  
+-----+-----+  
| character_set_client | utf8 |  
| character_set_connection | utf8 |  
| character_set_database | utf8 |  
| character_set_filesystem | binary |  
| character_set_results | utf8 |  
| character_set_server | utf8 |  
| character_set_system | utf8 |  
| character_sets_dir | /usr/share/mysql/charsets/ |  
+-----+-----+
```

8 rows in set (0.00 sec)

</pre>

&nbsp;

```
show variables like 'collation_%';
```

```
+-----+-----+  
| Variable_name | Value |  
+-----+-----+  
| collation_connection | utf8_general_ci |  
| collation_database | utf8_bin |  
| collation_server | utf8_bin |  
+-----+-----+
```

3 rows in set (0.00 sec)

</pre>

&nbsp;

&nbsp;

&nbsp;

修改数据库字符集

```
ALTER DATABASE db_name DEFAULT CHARACTER SET character_name [COLLATE ...];
```

&nbsp;

把表默认的字符集和所有字符列 (CHAR,VARCHAR,TEXT) 改为新的字符集

```
ALTER TABLE tbl_name CONVERT TO CHARACTER SET character_name [COLLATE ...];
```

&nbsp;

修改表默认字符集

```
ALTER TABLE tbl_name DEFAULT CHARACTER SET character_name [COLLATE...];
```

<p>&nbsp;</p>

<p>修改字段字符集</p>

<p>ALTER TABLE tbl\_name CHANGE c\_name c\_name CHARACTER SET character\_name [COLLATE ...];</p>

<p>&nbsp;</p>

<p>查看数据库编码</p>

<p>SHOW CREATE DATABASE db\_name;</p>

<p>&nbsp;</p>

<p>查看表编码</p>

<p>SHOW CREATE TABLE tbl\_name;</p>

<p>&nbsp;</p>

<p>查看字段编码</p>

<p>SHOW FULL COLUMNS FROM tbl\_name;</p>

<p>&nbsp;</p>

<p>&nbsp;</p>

<p>&nbsp;</p>

<p>&nbsp;</p>