[**UTF-8和GBK编码之间的区别(页面编码、数据库编码区别)以及在实际项目中的应用**](https://www.cnblogs.com/sochishun/p/7026762.html)

**第一节：UTF-8和GBK编码概述**

**UTF-8 (8-bit Unicode Transformation Format)**是一种针对Unicode的可变长度字符编码，又称万国码，它包含全世界所有国家需要用到的字符，是国际编码，通用性强，是用以解决国际上字符的一种多字节编码。由Ken Thompson于1992年创建。UTF-8用1到4个字节编码UNICODE字符，它对英文使用8位/8Bit（即1个字节/1Byte），中文使用24位/24Bit（3个字节/3Byte）来编码。用在网页上可以同一页面显示中文简体繁体及其它语言（如日文，韩文）。

**GBK (Chinese Internal Code Specification)**是汉字编码标准之一，全称《汉字内码扩展规范》，中华人民共和国全国信息技术标准化技术委员会1995年12月1日制订，国家技术监督局标准化司、电子工业部科技与质量监督司1995年12月15日联合以技监标函1995 229号文件的形式，将它确定为技术规范指导性文件。  
GBK是国家标准GB2312基础上扩容后兼容GB2312的标准（GB2312共收录了7445个字符，包括6763个汉字和682个其它符号；GBK共收录了27484个汉字，同时还收录了藏文、蒙文、维吾尔文等主要的少数民族文字）。GBK的文字编码是用双字节来表示的，即不论中、英文字符均使用双字节来表示（注意，GB系列编码是利用了字节中的最高位和ASCII编码区分的，可以和ASCII码混用。所以全角模式下英文是2字节，半角模式英文还是1字节）。为了区分中文，将其最高位都设定成1。GBK包含全部中文字符，是国家编码，通用性比UTF8差，不过UTF8占用的数据库比GBD大。

**简单概况就是：**  
UTF-8英文1字节中文3字节，在编码效率和编码安全性之间做了平衡，适合网络传输，是理想的中文编码方式.  
GBK英文1字节(半角1字节，全角2字节)，中文2字节,GBK的范围比GB2312广，GBK兼容GB2312。

**参考文章：**  
http://blog.csdn.net/mydriverc2/article/details/50525203  
http://blog.csdn.net/liudajiang/article/details/41133077  
http://www.cnblogs.com/xiaomia/archive/2010/11/28/1890072.html