1. 先说一下字符串的编码，最开始字符串用ASCII编码，也就是用一个八位二进制来表示，完全可以覆盖所有英文。后来由于其他国家语言的加入，ASCII不够用了，所以就出现一种万国码Unicode，Unicode对所有语言用两个字节，部分汉语用三个字节。这就导致Unicode不仅不能兼容ASCII，而且造成了空间浪费。由此又产生了utf-8编码，utf-8编码对英文使用一个字节编码，既能够兼容ASCII，又能节约空间。
2. python3中的二进制文件，统一采用字节码读写，bytes用b’XXXX’来表示，其中的X可以是字符，也可以是ASCII。将bytes进行适当格式的编码，就可以变成字符了；当然字符通过适当形式的解码就可以变成bytes；分别使用函数encode（）和decode（）
3. python3默认使用str对字符串进行编码，使用bytes对二进制数据流进行编码。Python3表示字符序列的方式有两种：bytes和str，前者的实例包含原始的8位值，后者的实例包含Unicode字符。

4、文件操作时，如果open的时候有参数’wb’,则在写的时候必须以bytes为输入参数：write（‘ssss’）是错的，wirte（b’ssss’）是对的