变量和类型

整型(int)：支持二进制(0b1011)/八进制(0o13)/十进制(11)/十六进制(0xa)

浮点型(float)：小数，有数学表示法(12.34)和科学表示法(1.234e1)

字符串(str)：使用单引号或双引号限定

布尔型(bool)：使用True或False表示真假

复数型(complex)：表现形式为2+3j。2为实数部分，3j为虚数部分

变量命名

一、硬性规则

1.有字母、数字和下划线组成，不能以数字作为开头

2.大小写敏感

3.不跟关键字和系统保留字重合

二、PEP8要求

1.用小写字母拼写，单词之间用下划线连接

2.受保护的实例属性以单个下划线开头

3.私有的实例属性以两个下划线开头

数字运算符：

+：加法运算符

-：减法运算符

\*：乘法运算符

/：除法运算符，结果为浮点数(含小数部分)

//：除法运算符，结果为整数

%：求余数

\*\*：幂运算符

\_：最后一个表达式的值

字符串：

str是不可变的Unicode代码序列点(可encode为bytes)

str默认在内存中以Unicode(2字节)方式存储，在文本档案中以UTF-8(1~4字节，RFC 3629)为具体实现方式

ANSI编码规则：使用本地编码规则(如：简体中文的GBK)，其中GBK的0x00~0x7f依旧使用一个字节，兼容于ASCII

UTF-8编码规则：

Unicode 7bit编码类似于ASCII。8bit以上时，首字节连续1的个数代表UTF字节，后续紧跟一个0分开实质编码内容；后续的每个字节都已10开头，后面紧跟实质编码内容。如下图：



[u]“ ”：限定单行字符串

[u]‘ ’：限定单行字符串

‘‘‘ ’’’：限定多行字符串

“““ ”””：限定多行字符串

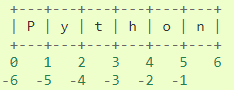
r‘ ’：原始字符串(所有字符串按编辑内容打印)

b‘ ’：bytes类型字符串(被编码的内容)

\*‘ ’：重复字符串内容次数

+‘ ’：叠加字符串

索引：通过str[i]的形式，对字符串内的字符进行读取，i正向由0起始，反向由-1起始，如图



切片：通过str[i:j]的形式，对字符串内的多个连续字符进行读取，读取范围为左闭右开区间[i,j)

built-in的Unicode单字符与int转换方法：

ord()：将Unicode单字符转化为int

chr()：将int转化为对应Unicode单字符内容

\*\*字符串内容为不可更改，str[i]=”a”为错误操作，但str=”abc”可行，因为直接修改了str的内容指向

str与bytes区别：

str属于Text Sequence Type，bytes属于Binary Sequence Type(编码后)

方法一览：

capitalize()：字符串首字母大写

casefold()：将所有字母转化为小写。相较于lower，casefold还能将其他语言的字符转化为小写英文字母

center(width[,fillchar])：将字符串宽度拓展为width(小于len(str)时，该值为len值)，自身置于中间范围，可指定剩余空间的填充字符(默认为空格)

count(substr[,start[,end]])：substr在字符串中出现的次数，可指定字符串的起始和结束搜索范围

encode(encoding="utf-8", errors="strict")：将字符串编码为bytes，可指定编码和错误类型。

startswith(prefix[, start[, end]])：是否以指定字符串集prefix起始，并且可指定起始和结束位置

endswith(suffix[, start[, end]])：是否以指定字符串集suffix结尾(suffix可以为元组)，可指定匹配的起始和结束搜索范围

expandtabs(tabsize=8)：将\t替换为空格，并可指定单个\t替换为几个空格

find(sub[, start[, end]])：从字符串中查找sub子串，并返回子串出现的首字母位置(无匹配返回-1)，可指定字符串起始和结束搜寻位置

format(\*args, \*\*kwargs)：在字符串中以数字{n}标记参数(或命名{name})，使用参数替换字符串中的内容。例：”I love {me}”.format(me=”you”)

format\_map(mapping)：类似于format的format(\*\*kwargs)，但mapping为dict的子类

index(sub[, start[, end]])：类似于find，但无匹配时导致ValueError异常

isalpha()：判断是否为字母

isdecimal()：判断是否为十进制数字

isdigit()：判断是否为十进制数字或者其他特殊意义数字(如上标数字)

isnumeric()：判断是否为数字

isalnum()：判断是否为字母或者数字

isascii()：判断是否为ascii字符，Unicode范围U+0000-U+007F

isidentifier()：判断是否为identifier。identifier内容[跳转](https://docs.python.org/3/reference/lexical_analysis.html#identifiers)

isprintable()：是否为可打印字符

islower()：是否为小写字母

isupper()：是否为大写字母

isspace()：是否仅包含空格符

istitle()：是否为每个单词首字母大写形式

join(iterable)：将字符串作为iterable内容的分隔符

ljust(width[,fillchar])：将字符串的以左对齐排列(当width小于len(str)时，返回原字符串)，并可以提供填充字符

rjust(width[,fillchar])：将字符串的以右对齐排列(当width小于len(str)时，返回原字符串)，并可以提供填充字符

lower()：将字符串的字符全部转化为小写

upper()：将字符串的字符全部转化为大写

lstrip([chars])：删除前置的指定字符集，默认为空格

rstrip([chars])：删除后置的指定字符集，默认为空格

strip([chars])：删除前置和后置的指定字符集，默认为空格

static maketrans(x[,y[,z]])：返回一个转化table(x为转化前的字符集，y为转化前对应字符的转化后字符集，z为需要转化为None的字符集)，用于translate方法

translate(table)：实际的字符映射转化，table为maketrans返回的内容

partition(sep)：以匹配内容sep第一次出现的地方进行分割，返回三元元组，第一部分为分割符前的内容，第二部分为分割符，第三部分为分割符后的内容。若无法匹配sep，则第一部分为整体字符串，第二和第三部分为空字符串

rpartition(sep)：以匹配内容sep最后一次出现的地方进行分割，返回三元元组，第一部分为分割符前的内容，第二部分为分割符，第三部分为分割符后的内容。若无法匹配sep，则第一和第二部分为空字符串，第三部分为整体字符串

replace(old,new[,count])：将匹配的old内容替换为new，可指定替换的次数count

rfind(sub[,start[,end])：从右边开始匹配sub，并可指定起始和结束位置，匹配失败返回-1

rindex(sub[,start[,end])：类似于rfind，但是匹配失败导致ValueError异常

split(sep=None, maxsplit=-1)：将字符串以sep为分割符字符集从左边开始进行分割(默认为空白字符)，最大分割次数为maxsplit

rsplit(sep=None, maxsplit=-1)：将字符串以sep为分割符字符集从右边开始进行分割(默认为空白字符)，最大分割次数为maxsplit

splitlines([keepends])：以换行符作为分割(\n、\r、\r\n)

swapcase()：转换字符串中的大小写

title()：将每个单词的首字母转化为大写

zfill(width)：将字符串延长为width宽度，并以前置”0”填充，当width小于len(str)时，返回原字符串

printf格式字符串

“%(NAME)FLAG-WIDTH-PRECISION-TYPE”%{KEY:VALUE}

FLAG：

‘#’ 前置填充空格，并为十六进制和八进制自动添加前缀(0x或0o)，适用于字符串和数字

‘0’ 前置填充0，适用于数字

‘-’ 后置填充空格，适用于字符串和数字(覆盖’0’FLAG)

‘ ’ 前置填充空格，适用于字符串和数字

‘+’ 前置填充空格，并为数字添加前置sign符号，适用于字符串和数字(覆盖’ ’FLAG)

TYPE:

‘d’ 符号十进制数字

‘i’ 符号十进制数字

‘o’ 符号八进制数字

‘x’ 符号十六进制数字(小写)

‘X’ 符号十六进制数字(大写)

‘e’ 指数形式浮点数(小写)

‘E’ 指数形式浮点数(大写)

‘f’ 十进制浮点数

‘F’ 十进制浮点数

‘g’ 十进制浮点数，当指数小于-4或大于5时，使用指数形式(小写)

‘G’ 十进制浮点数，当指数小于-4或大于5时，使用指数形式(大写)

‘c’ 单一字符

‘r’ 字符串(类似于repr())

‘s’ 字符串(类似于str())

‘a’ 字符串(类似于ascii())

‘%’ 无NAME字段，只返回%