

第一题：

1. 单引号跟双引号的普通用法是相同的，str1 = ‘abc’ str2 = “abc”
2. 单引号跟双引号的主要区别体现在当单引号抱起来的字符串含有双引号是，不需要使用转义字符，同样，在双引号抱起来的字符串含有单引号时，不需要使用转义字符
3. 三引号：’’’str‘’’, “””str”””, 核心用法体现在跨行的字符串上

Str = ‘’’abc

def’’’

python中没有多行注释，一般使用三引号

第二题：

使用global修饰变量名

第三题：

带有yield的函数不再是一个普通的函数，Python解释器沪江其视为一个generator，所以yield的作用就是把一个函数编程一个生成器

第四题：

random.randint(1, 11)

第五题：

from PIL import Image

img = Image.open(‘pic.jpg’)

img = img.resize((250, 156), Image.ANTIALIAS)

img.save(‘pic\_small.jpg’)

第六题：

1).Scrapy

Scrapy是一个为了爬取网站数据，提取结构性数据而编写的应用框架。 可以应用在包括数据挖掘，信息处理或存储历史数据等一系列的程序中。。用这个框架可以轻松爬下来如亚马逊商品信息之类的数据。

2).PySpider

pyspider 是一个用python实现的功能强大的网络爬虫系统，能在浏览器界面上进行脚本的编写，功能的调度和爬取结果的实时查看，后端使用常用的数据库进行爬取结果的存储，还能定时设置任务与任务优先级等。

3).Beautiful Soup

Beautiful Soup 是一个可以从HTML或XML文件中提取数据的Python库.它能够通过你喜欢的转换器实现惯用的文档导航,查找,修改文档的方式.Beautiful Soup会帮你节省数小时甚至数天的工作时间。

4).Portia

Portia是一个开源可视化爬虫工具

第七题：

一、通过User-Agent来控制访问：

无论是浏览器还是爬虫程序，在向服务器发起网络请求的时候，都会发过去一个头文件：headers

方案：可以自己设置一下user-agent，或者更好的是，可以从一系列的user-agent里随机挑出一个符合标准的使用

二、通过IP限制来反爬虫

方案：比较成熟的方式是：IP代理池。简单的说，就是通过ip代理，从不同的ip进行访问，这样就不会被封掉ip了。

三、通过JS脚本来防止爬虫：

方案：使用”PhantomJS“：PhantomJS是一个Python包，他可以在没有图形界面的情况下，完全模拟一个”浏览器“，js脚本验证什么的再也不是问题了

四、通过robots.txt来限制爬虫

第八题：

HTTP协议传输的数据都是未加密的，也就是明文的，因此使用HTTP协议传输隐私信息非常不安全，为了保证这些隐私数据能加密传输，于是网景公司设计了SSL（Secure Sockets Layer）协议用于对HTTP协议传输的数据进行加密，从而就诞生了HTTPS。简单来说，HTTPS协议是由SSL+HTTP协议构建的可进行加密传输、身份认证的网络协议，要比http协议安全。

HTTPS和HTTP的区别主要如下：

1、https协议需要到ca申请证书，一般免费证书较少，因而需要一定费用。

2、http是超文本传输协议，信息是明文传输，https则是具有安全性的ssl加密传输协议。

3、http和https使用的是完全不同的连接方式，用的端口也不一样，前者是80，后者是443。

4、http的连接很简单，是无状态的；HTTPS协议是由SSL+HTTP协议构建的可进行加密传输、身份认证的网络协议，比http协议安全。

第九题：

cookie(储存在用户本地终端上的数据)：指某些网站为了辨别用户身份、进行session跟踪而储存在用户本地终端上的数据（通常经过加密）。

Session技术：

是一种将会话状态保存在服务器端的技术，它可以比喻成是医院发放给病人的病历卡和医院为每个病人保留的病历档案的结合方式。客户端需要接收、记忆和回送Session的会话标识号，Session可以且通常是借助Cookie来传递会话标识号。

cookie 和session 的区别：

1、cookie数据存放在客户的浏览器上，session数据放在服务器上。

2、cookie不是很安全，别人可以分析存放在本地的COOKIE并进行COOKIE欺骗

考虑到安全应当使用session。

3、session会在一定时间内保存在服务器上。当访问增多，会比较占用你服务器的性能

考虑到减轻服务器性能方面，应当使用COOKIE。

4、单个cookie保存的数据不能超过4K，很多浏览器都限制一个站点最多保存20个cookie。

5、建议：将登陆信息等重要信息存放为SESSION，其他信息如果需要保留，可以放在COOKIE中

第十题：

InnoDB：

支持事务处理等

不加锁读取

支持外键

支持行锁

不支持FULLTEXT类型的索引

不保存表的具体行数，扫描表来计算有多少行

DELETE 表时，是一行一行的删除

InnoDB 把数据和索引存放在表空间里面

跨平台可直接拷贝使用

InnoDB中必须包含AUTO\_INCREMENT类型字段的索引

表格很难被压缩

MyISAM：

不支持事务，回滚将造成不完全回滚，不具有原子性

不支持外键

不支持外键

支持全文搜索

保存表的具体行数,不带where时，直接返回保存的行数

DELETE 表时，先drop表，然后重建表

MyISAM 表被存放在三个文件 。frm 文件存放表格定义。 数据文件是MYD (MYData) 。 索引文件是MYI (MYIndex)引伸

跨平台很难直接拷贝

MyISAM中可以使AUTO\_INCREMENT类型字段建立联合索引

表格可以被压缩