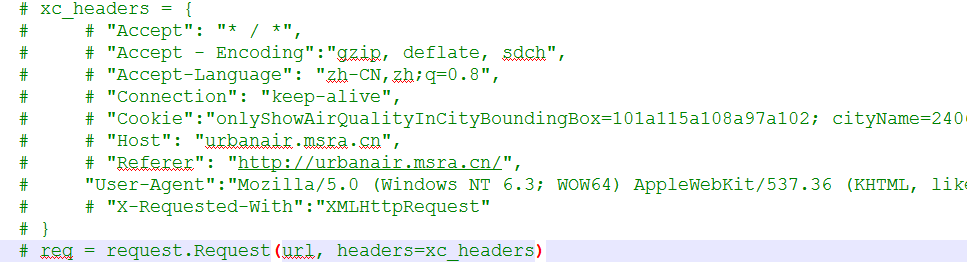
在使用python爬虫爬取数据的时候， 可能会遇到一些网站的反爬虫措施，解决方法就是针对于headers中的User-Agent进行设置，并设置一些请求信息——Reqeust Header的内容,

一条一条设置太麻烦，可以进行以下设置：



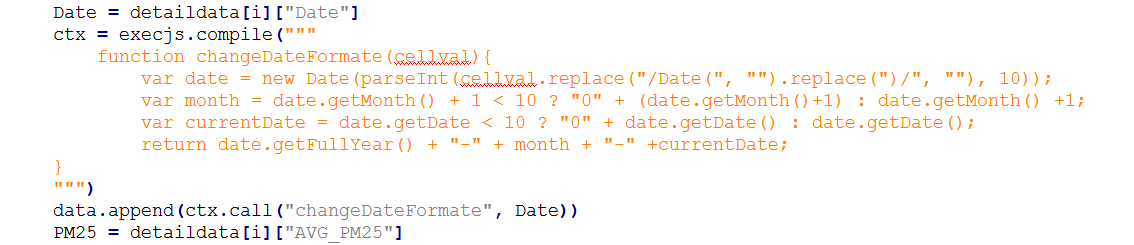
对于爬取的网页简单，页面少且数据量少任务，只设置User-Agent就行。

在这次实践中，遇到一个没见过的大问题，就是json日期格式的问题，我从网上爬下来的

日期格式为：/Date(1306418993027)/

我是第一次见到，从网上得知这是从服务器端返回的json格式的数据，从网上查找可靠的转换方式，基本都是用js代码进行转换，代码不难，难的是如何将爬下的数据用这几行代码转换。由于js代码无法嵌入到java当中，但在jsp页面中反过来却可以，想到用java代码将数据传送至jsp页面用js代码处理，但是这太麻烦。然后继续在网上找，但就是找不到一个合适的解决方案。 就在绝望的时候看到一丝希望。 经网上查得python中可以通过pyexecjs这个插件嵌入js代码，经实践确实简单又实用，最后终于把问题解决了。

代码如下：



经过这次实践，不仅学到有关python新知识，而且让我深刻的感触到python的强大，同时也明白了自己一些方面知识的欠缺。