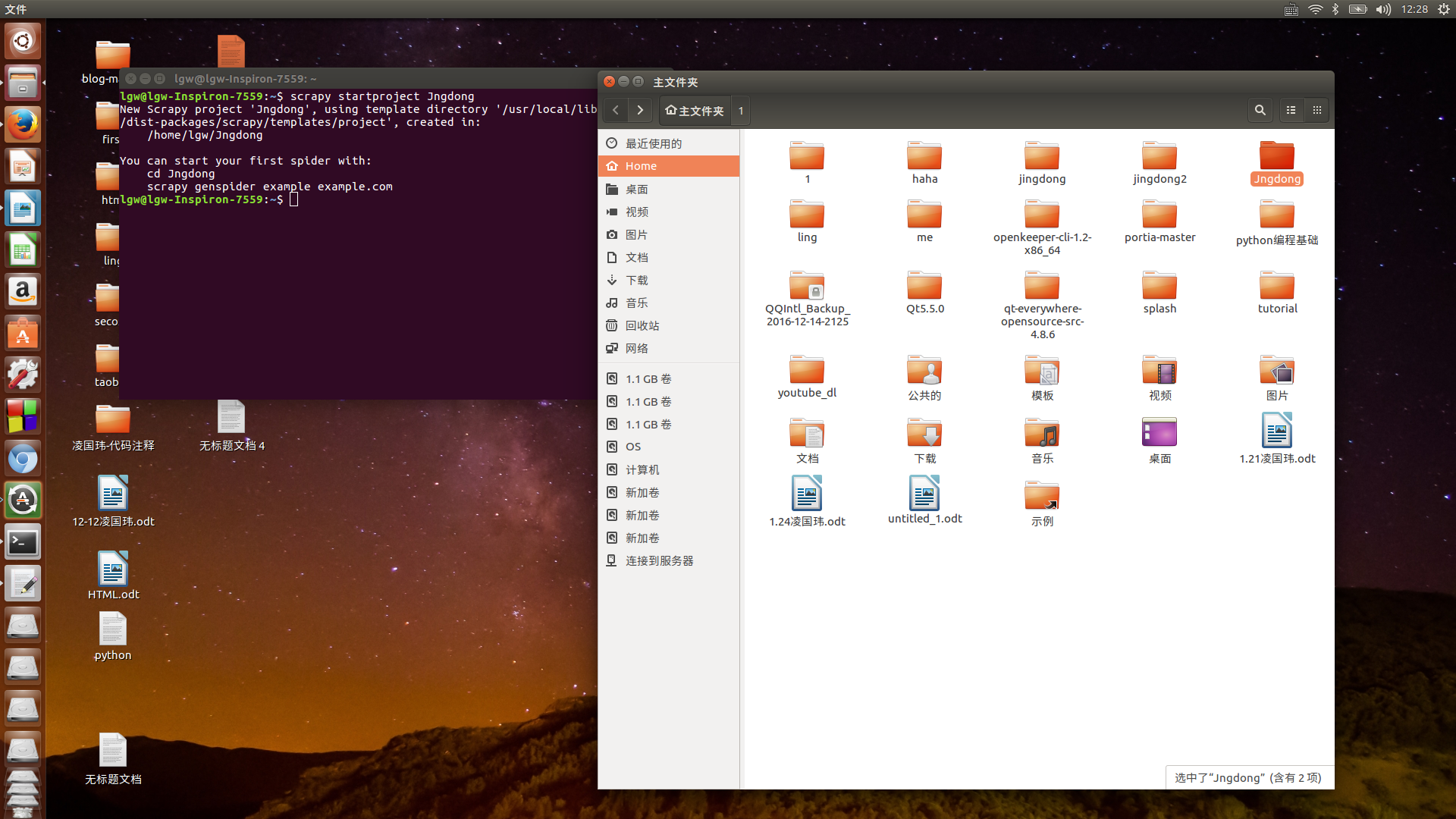
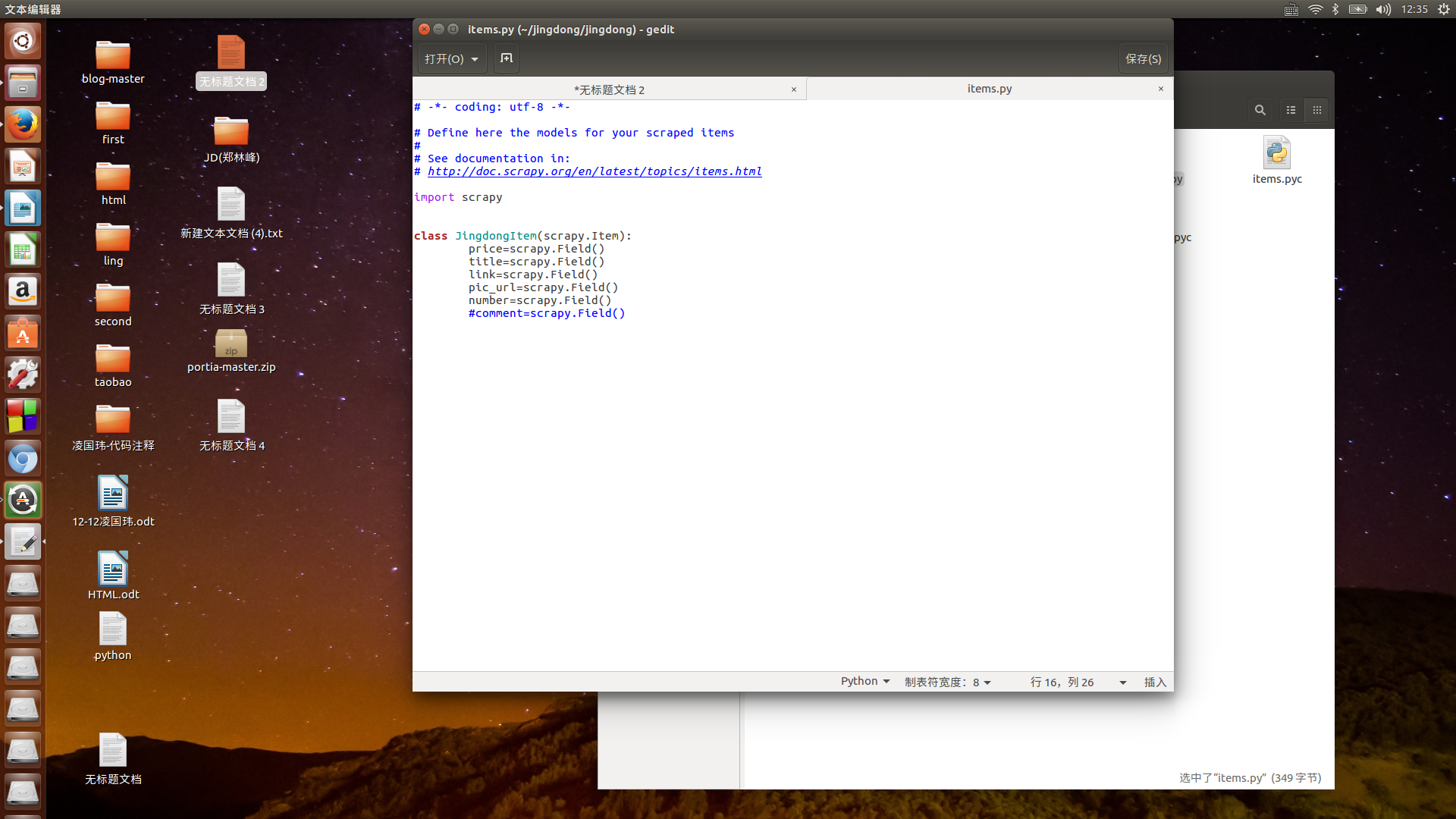
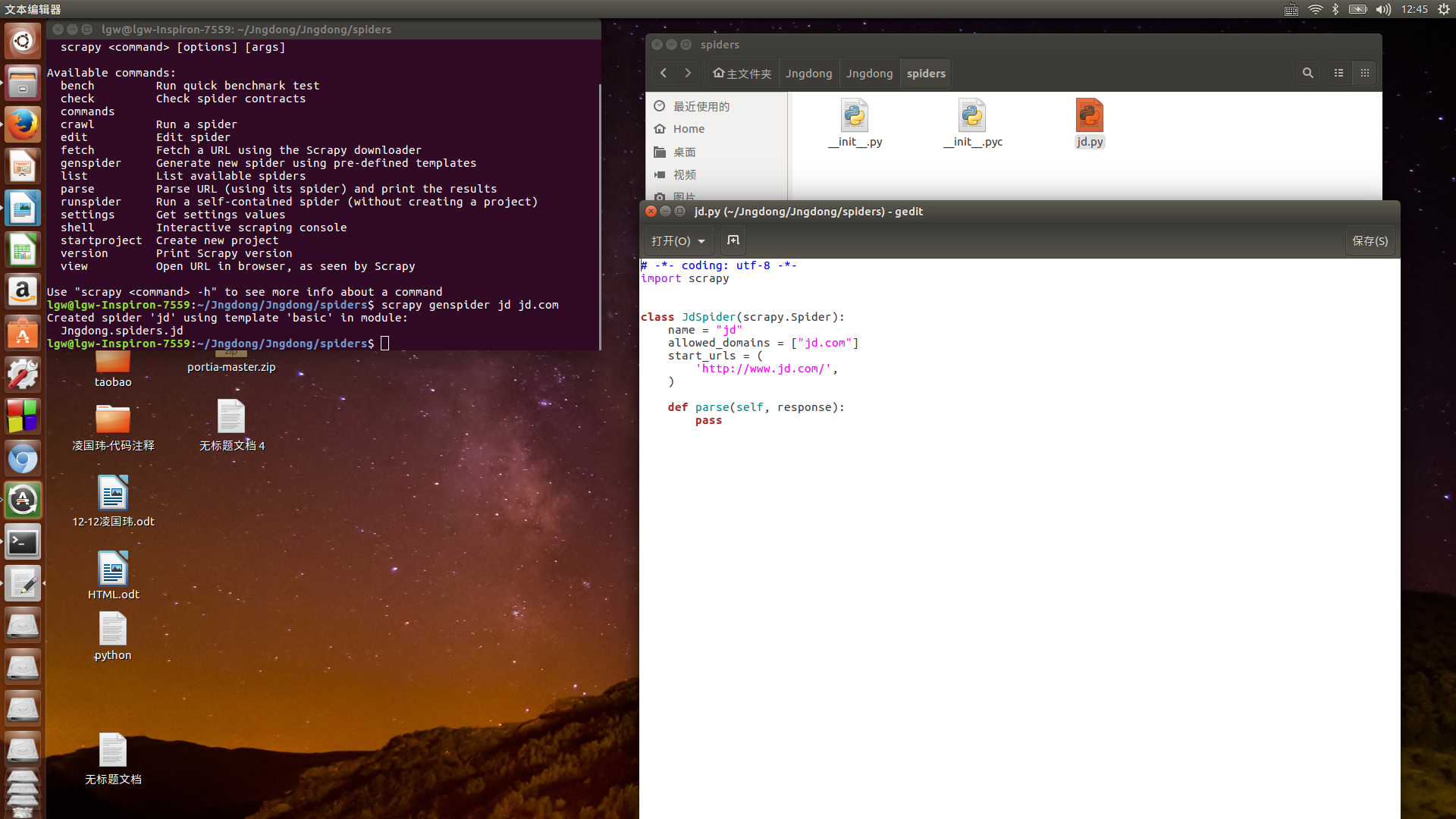
1. 首先，我们需要创建一个爬虫，无论是什么爬虫，只要是用scrapy写的，创建方法都大致相同。我们需要给爬虫的文件夹起个名字，比如，我们爬取京东的爬虫。取名为Jngdong。

可以看到我们运行scrapy startproject Jngdong，就有了这个文件夹，我们的爬虫就在这里写。

1. 现在我们打开这个文件夹，里面有一个items.py。具体来说，就是定义我们究竟要抓取的东西。

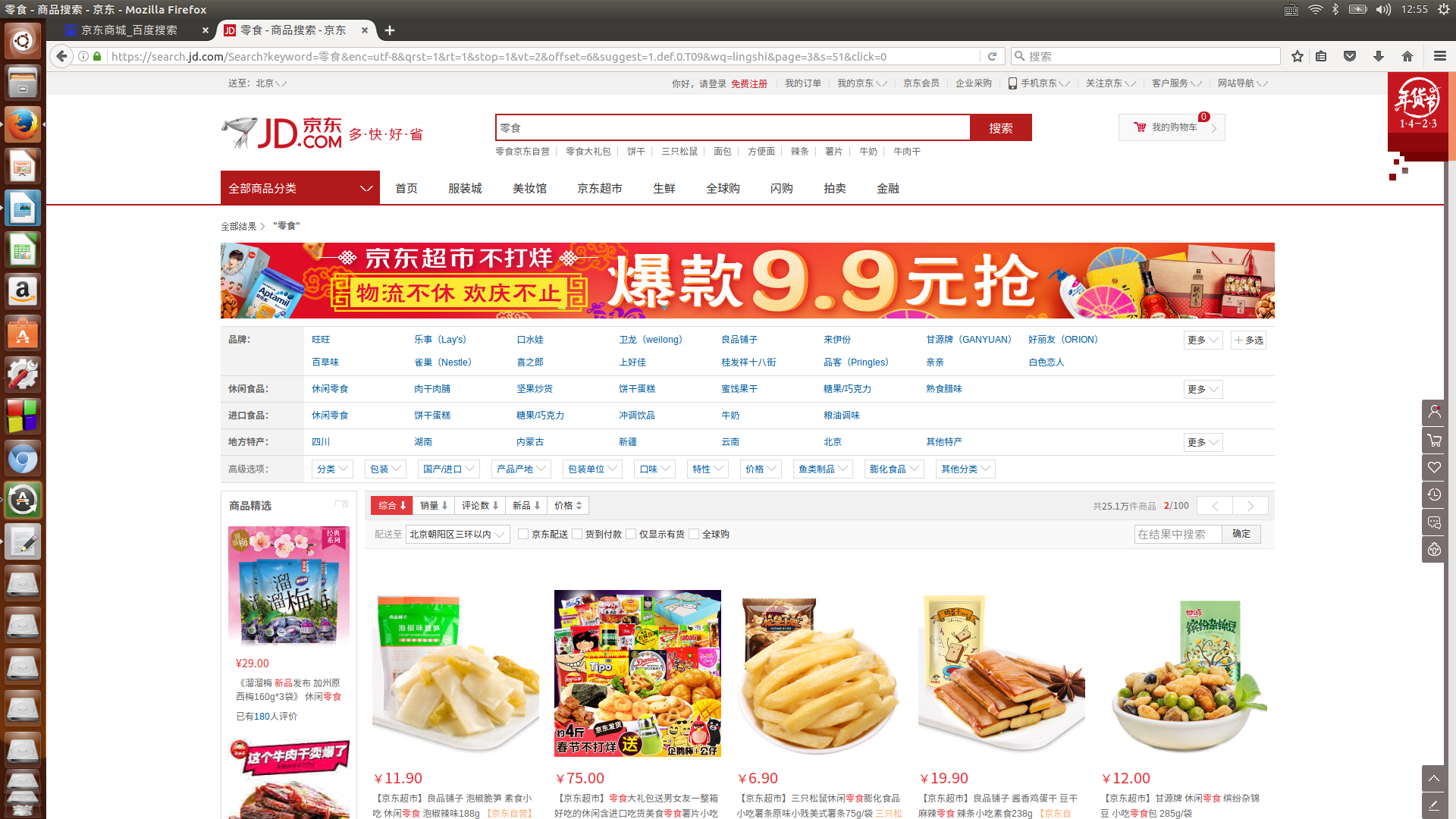
可以看到，我们需要的都用统一的格式定义。比如，我们需要价格，就是price=scrapy.Field（）

1. 现在准备工作做好了之后，我们就开始正式编写京东的爬虫。在spiders目录下，我们还需要一个真正的爬虫py。我运行scrapy genspider jd jd.com创建这个py。

如果不知道scrapy的命令。可以在终端上直接运行scrapy。我们的京东爬虫。就在这个jd.py上完成.

1. 现在开始的是网站分析，无论爬取什么网站，都必须分析这个网站。

现在开始京东的分析。

我搜索的零食，爬取有关京东商城零食的有关信息。可以看到，这个网站的地址是如此的长，其实，我们的地址，根本不需要这么长，有一些字符是用来迷惑我们眼球的。在我们的删减下，其实真正需要的地址很短。

看吧，在删减下，真正需要的地址只是这么一点。真正需要分析地址的规律，每一页的地址不同在于&page=？？？京东的是以1,3,5,7,9的规律下去的，这是第二页，所以是3.

1. 现在，进入jd.py开始写京东爬虫。（jd.py注释）

补充：对于抓包，用浏览器抓包，f12可以刷新，然后再包中找到自己需要抓取的，打开该包的网页，再用正则表达式进行匹配，配合jd.PY上的注释，应该可以理解。