京东商城家具<http://www.jd.com/>

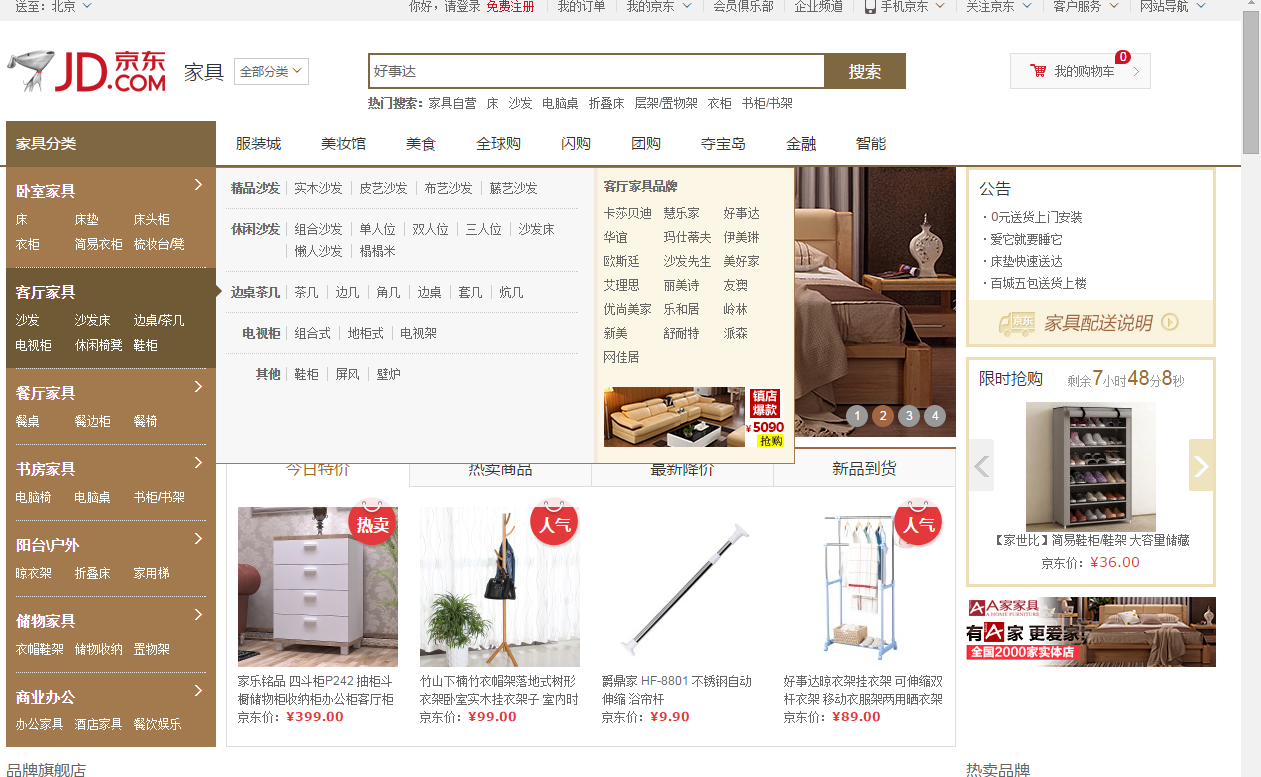
1. 选择运送到“北京”



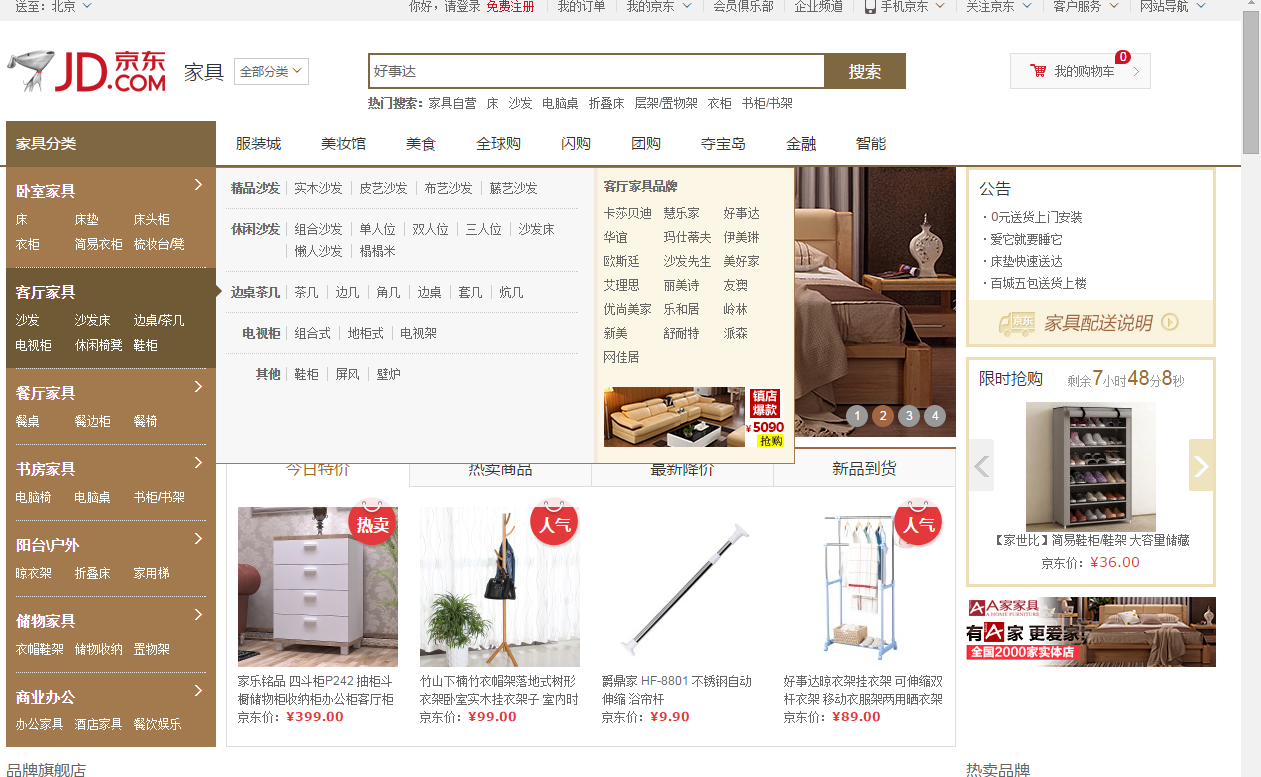
2. 点击选择“家具”



3. 进入页面如下



4.需要保存左边的层次结构，如“客厅家具”下面有六个子类（红色框），**不需要**考虑右边的展开菜单的内容。



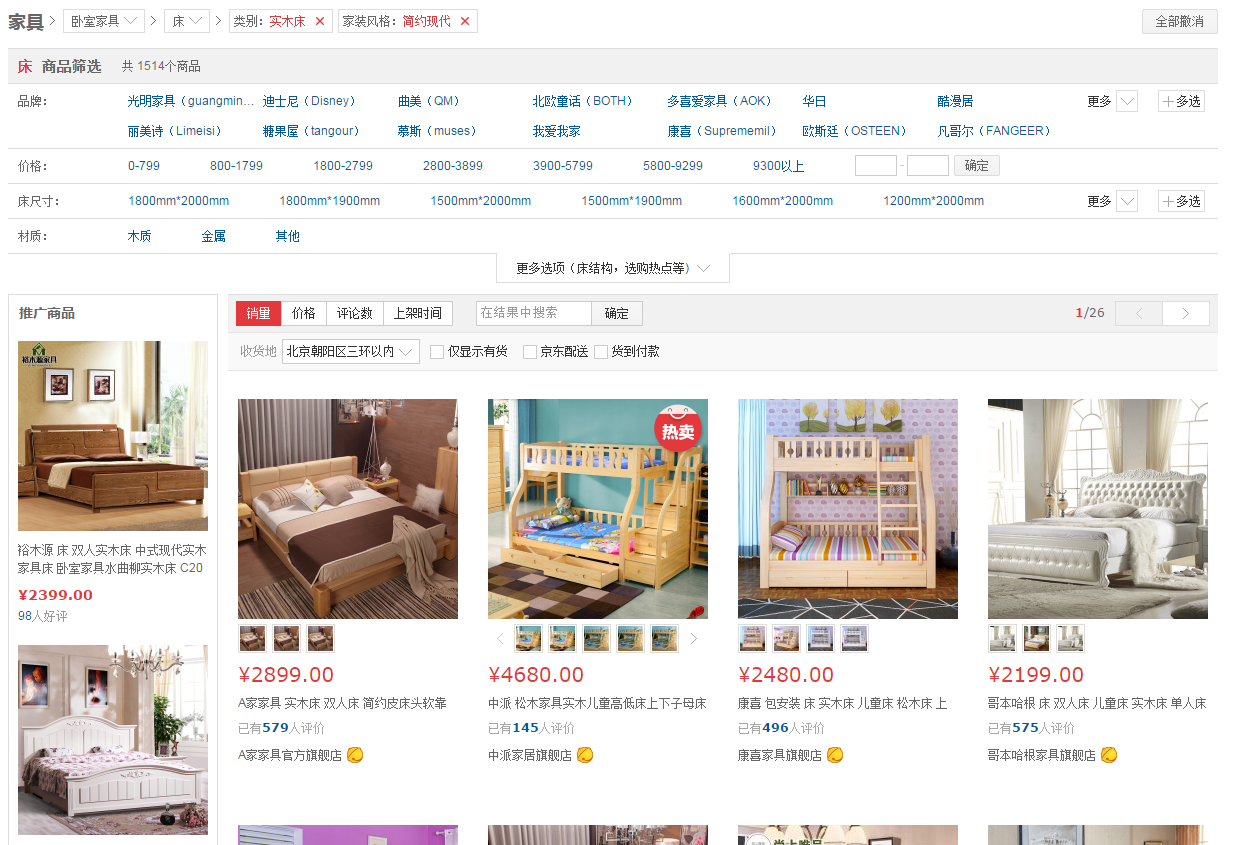
5. **对于上面的大类，即：卧室、客厅、餐厅、书房、阳台、储物（不需要下载商业办公）。请点击每个大类下面的小类，如卧室的小类包括：床、床垫、床头柜、衣柜、简易衣柜、梳妆台**

**总共24小类**

6. 下面以床为例



需要按照“**类别”**与“**风格”**进行分类下载，如“实木床+简约现代”。按照“**销量**”默认排序，下载前**500**条记录。



7. 如点击第一条记录，需要下载**图片（大图）**、**描述**、**价格、商品介绍、最佳组合里的内容（包括每一个组合的图片、描述、价格、商品介绍信息；不需要嵌套的下载最佳组合）。**



<http://item.jd.com/1354671914.html#product-detail>**请以该记录链接中的ID命名，方便后续查重处理。因为同一个商品可能是多个商品的最佳组合。**以ID命名也能避免出现文件夹乱码的问题。

大概估计总共有400万数据左右。