打开贝壳网，可以很清晰的看到每一个楼盘都有地理位置，类型，面积，单价，户型及一些附加的信息，这其中的位置，单价，类型尤为重要，因为这是我们分析所需要的。在数据采集的时候也是不可缺少的。

为了方便分析，我们采集的将数据导入到execl表格中，因为在现有的数据中，有部分数据在单价，面积的字段为待定，所以这部分数据我们就要将它们清洗掉，剩下的房源信息才是真正有效的数据。

在这些有效的数据中，我们分商业，住宅，写字楼和别墅四个种类，区域选择为成都的各个区，当然，因为地理位置和种类不同，他们的单价肯定不同。为了更直观的展示出成都各个区域的房屋价格，我们将每个区域的新房平均价格计算出来，对比下来，各个区域房屋价格孰高孰低就一目了然了。

为了有更好的用户体验，让用户更直观的看出房价，数量等信息，用什么样的图展现出来也是非常重要的，比如：新房的四个种类很明显用饼图比较直观易懂，各区域新房数量可能用柱状图比饼图更好。

画图工具的选择也很重要，开始我们选择的是用matplotlib，也可以做出我们想要的图，后面接触了Echarts发现它不仅能做出我们所需要的图而且还能是动态的，显的非常高大上，相比之下matplotlib就逊色不少，所以最终的画图工具我们选择了Echarts。

最后通过对有效数据的分析，我们得出了一个大致结论：价格方面：商业楼单价最高，住宅最低；数量上比较：商业楼最多，写字楼最少；房源数量方面：目前高新区的新房数量是最多的。从用户角度出发，相信大多数用户更关心的是住宅的相关信息，所以有必要单独说一下住宅分析结果：目前成都新房楼盘数量大概1100多个，过滤掉面积，单价待定的房源后，有600条左右的有效数据，成都住房的均价大概为12000左右，并且市区新房均价远远高于郊县。