**2021瓜子二手车在售车辆热度建模和比较分析**

**2021瓜子二手车在售车辆热度建模和比较分析**

**摘要**

目前，瓜子二手车依托350万车辆信息及超过2亿车主和潜在买家基因库，建立了强大的基于大数据智能算法的“瓜子大脑”，以此来构建自身积木式创新”的核心壁垒，目前已经应用到获客效率提升，股价和残值预估，成交概率预估，车源个性化匹配，金融、延保服务等各个环节中。

过去一年里，瓜子二手车打造了一支强大的超过6000人的线下团队，实现了对全国20个省市200多个城市的覆盖，并在各个交易环节取得了突破性进展。

国内新车销售在近15年里一直保持飞速发展，09年起中国汽车的产销量均占到世界第一位.二手车市场也将迎来交易高峰期，据中国汽车流通协会的不完全统计，二手车在2014年虽有605万辆的交易量，而新车却达到2300多万辆，是前者的4倍。拥有亿万车主用户的瓜子二手车平台数据显示，43%的汽车车龄在三年以上，汽车市场即将进入换购潮，届时将会为二手市场提供丰富的车源。

二手车买卖渠道的利和弊：

目前国内二手车交易有三个渠道：一是卖4S店，交易便捷但价格低；二是卖车贩，不仅压价，还无法保证车质等相关问题；三是互联网交易，因彼此不了解容易产生信任和资金问题。

如何在有效范围内提高二手车成交率：事实上，二手车交易过程并不亚于新车买卖，甚至需要更加精准的车辆数据来促成交易，而仅是数据这一项，就已经难倒了90%的买主。庞大且真假难辨的网络信息量令实施个人交易的买卖双方都无从下手。

2017年3月，瓜子二手车在北京、上海、重庆等城市实现交易额过亿元，与去年同期相比，交易额成倍增长。瓜子自营金融渗透率在主流城市达20%以上发布会上，瓜子二手车直卖网CEO杨浩涌表示，经过一年多的发展，瓜子二手车已经建立了足以推动行业革新的Al大数据能力，目前阶段的战略目标是为用户提供丰富的与车相关的产品和良好的服务体验.此轮融资将主要用在布局新业务、产品服务升级、品牌打造、团队及人才培养等方面。

目前，瓜子二手车以服务于车主的平台所带来的高效率，提升用户习惯和扩展业务覆盖，是现阶段的重点目标，让我们期待更多精彩！

**正文目录**

一、**背景意义与目标**…………………………………………………………3

二、**数据来源和数据处理**……………………………………………………3

三、**数据建模**…………………………………………………………………5

四、**数据可视化分析**…………………………………………………………6

五、**总结与建议**………………………………………………………………10

**一、背景意义与目标**

2016年之前我国二手车市场一直处于不瘟不火的状态，但自2018年交易量突破千万大关后，二手车销量开始保持两位数增长。据我国的中国汽车流通协会的预测，2022年二手车销量预计会实现15%-20%的增长，交易量达到1500万量。

和发达国家二手车的交易量为新车3到4倍的市场份额相比，国内尚处于起始阶段的二手车行业无疑拥有广阔的发展前景。拥有亿万车主用户的瓜子二手车平台数据显示，43%的汽车车龄在三年以上，汽车市场即将进入换购潮，届时将会为二手市场提供丰富的车源。另一方面，消费者对二手车认可度逐步提高，国内限迁政策的全面取消，金融贷款的大力支持，这些有利因素都会促使二手车行业进入高速发展阶段。

根据相关数据，对二手车市场在手车辆从价格、车龄、燃油种类、表显里程等方面进行分析，然后为二手车购买者提供合理的建议。相信看完我的分析报告你可以调到心仪的二手车。

c

**二、数据来源和数据处理**

**1、数据来源**

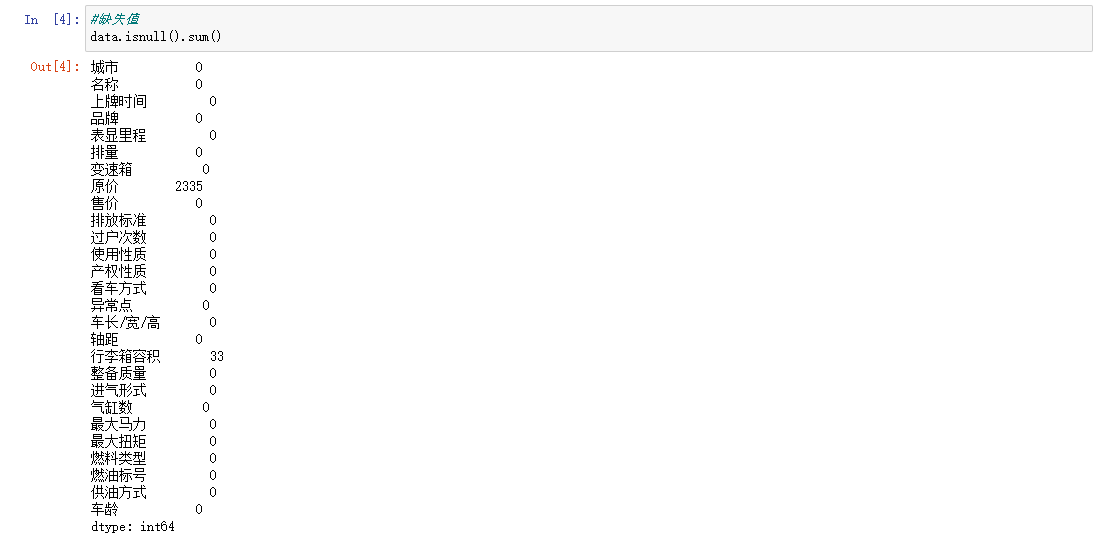
瓜子二手车5月在售车辆的数据一共有33720条，网站首页包含的数据标签有div,h1,ul,li,span等，右键检查，可以发现车辆信息都在ul->li中，查看网页源码，发现属于静态网站，然后通过BeautifulSoup获取字段：城市、名称、上牌时间、品牌、表显里程、排量、变速箱、原价、售价、排放标准、过户次数、使用性质、产权性质、看车方式、异常点、车长/宽/高、轴距、行李箱容积、整备质量、进气形式、气缸数、最大马力、最大扭矩、燃料类型、燃油标号、供油方式、车龄等字段信息用于之后的数据分析。

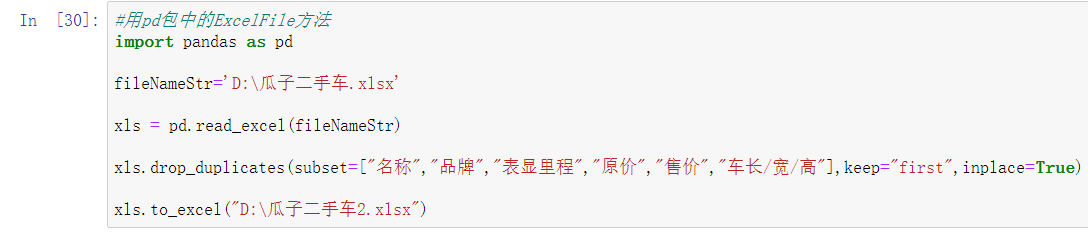
1. **数据预处理：**

原始数据的读取展示：



缺失值的检测





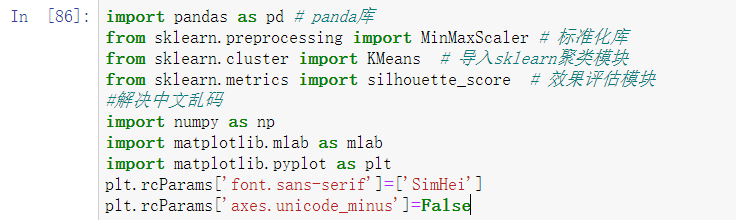
清洗数据并输出到一个新的excel文档。



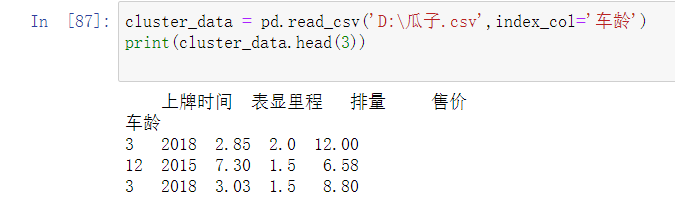
包括重复、缺失值、异常值的检测和处理

**三、数据建模**

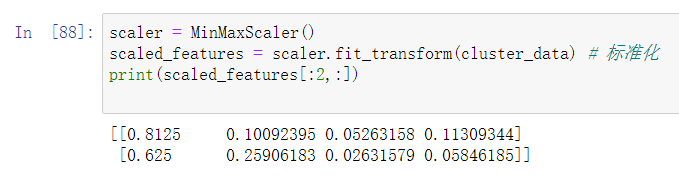
先导入相关的库

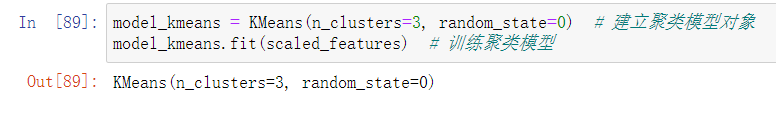


读取csv文件，然后输出表

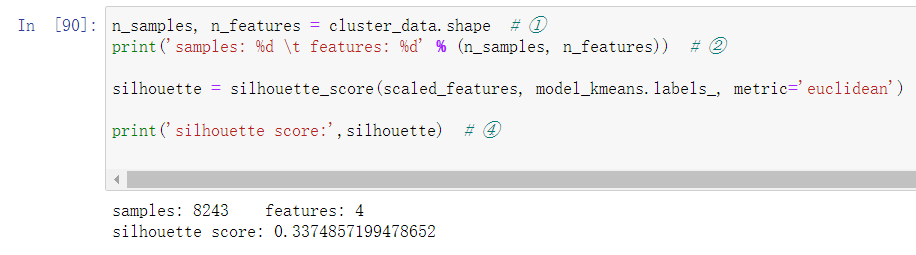


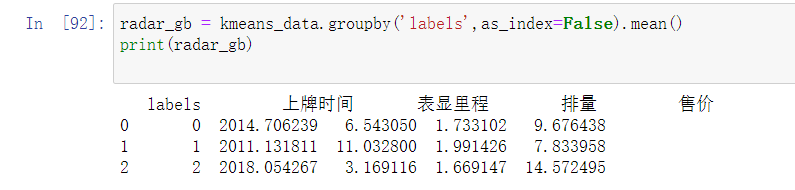
数据预处理



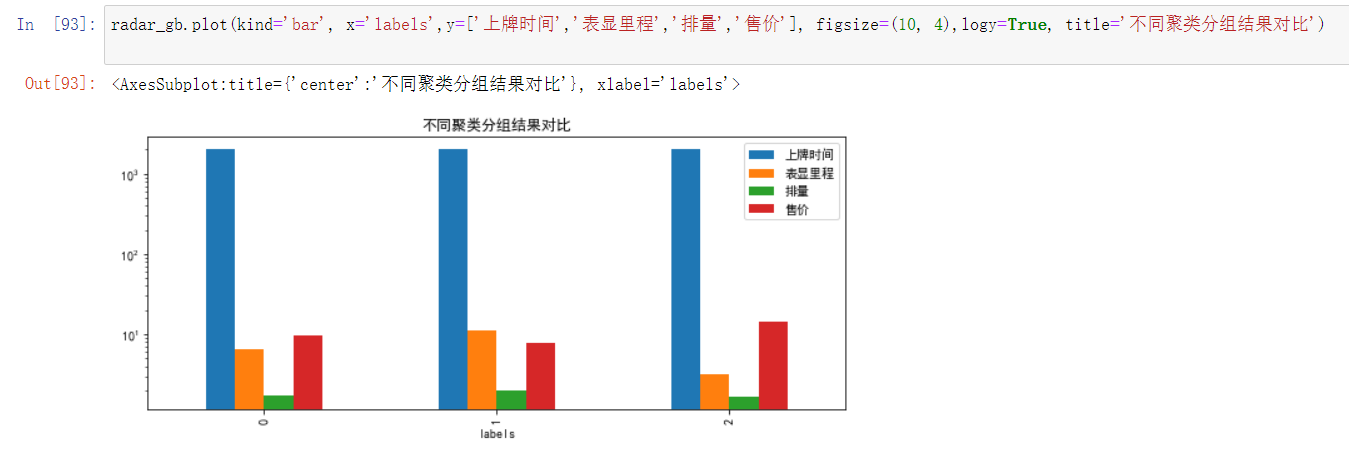
建立聚类模型对象

模型评估



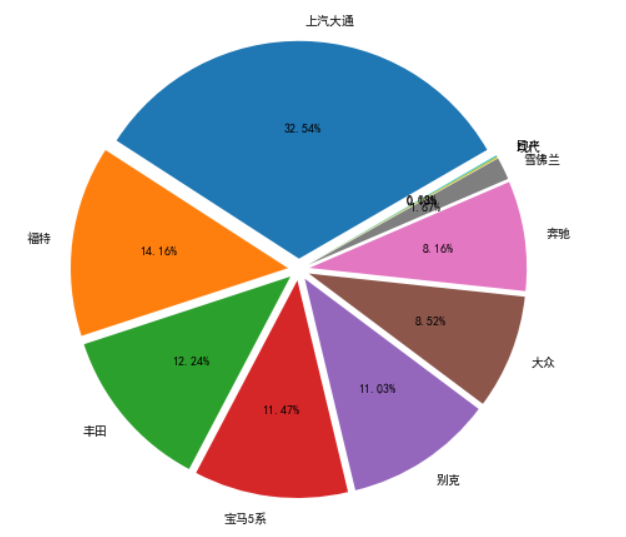


基于聚类群组汇总特征



**四、数据可视化分析**

**二手车在售前十占比图**



根据上图所示，在前十排名中在售二手车**上汽大通**的占比最大，达到总占比的32.54%，占比最小的是**现代**，达到总占比的0.03%。

卖家角度分析：

市场上很多车型新，品牌硬，保值率高。保值率高的车型，质量、技术一般都是比较好的，但买二手车图的就是便宜、好用，保值率高的二手车如果价格高了，性价比就没那么高了，甚至还不如去买个新的。

买家角度分析：

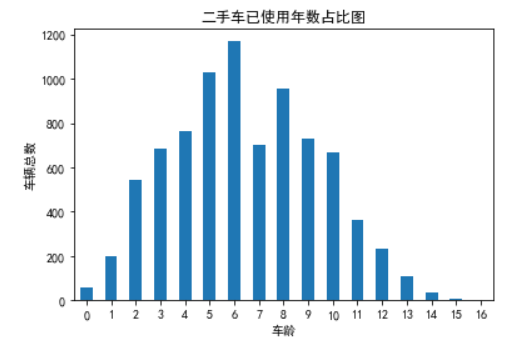
1.越是保值率高的二手车，价格越是偏高买二手车主要是为了省钱，如果你买保值率高的二手车，那就省不了多少钱，也不划算。

2.保值率越是高、价格却低的二手车，一般车龄更长价格低的二手车，车龄普遍也比一般的长。有的人会说，保值率高的二手车，技术、质量肯定更好，反正都二手车。但其实车龄越长的车子，车况也就越差，越是不适合入手。

现代二手车角度：

随着新款IX35的上市，价格直接来了个斩杀，直接从16.98~21.88跳到了11.99~16.19的价格区间，这个价格都可以和同级别国产SUV比较了。

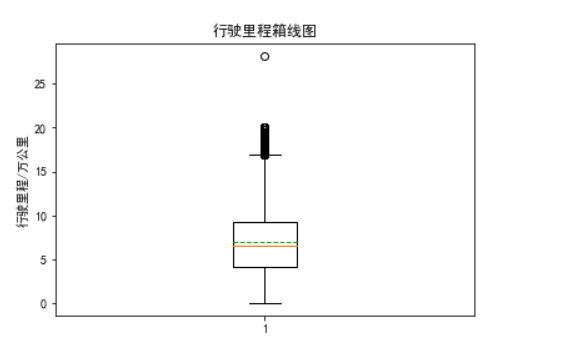
这种对热门车系大幅度的调价，个人认为是对品牌的一个提前消费，调价以后，确实让 IX35的新车销量大幅度提高，但是给大家一个极其强烈的信号就是现代汽车不保值，这也导致了一个现象，二手车商不敢收现代的车了，除非价格压得很低。毕竟二手车商还是以盈利为目的的，亏本的事情，谁都不愿意做。



根据上图所示，可以看出在售二手车车龄为6年的数量最多，最少的车龄为15年的。0年的车行业称为准新车，出售的原因可能是出现经济上的状况不得已而为之，并非是异常值。

买家角度：现在买二手车的人也比较多，毕竟二手车实在，价格实惠，可以省下一大笔钱。买二手车最佳车龄当然是越靠前越好，特别是准新车那买过来简直是太划算了。

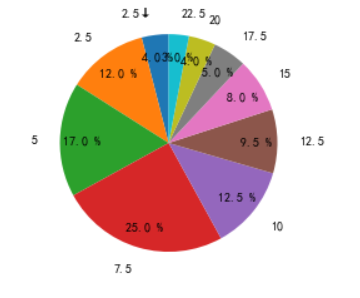
如果非要给个具体的时间的话，那么5到8年的车基本上是最好入手的时间了，这个时间点基本上磨合是在最佳的状态，但具体还要看车况的。一般来说，5-8年车辆的故障率会达到40%到60%左右，价格相对来说也比较优惠。所以二手车龄在5年到8年之间的二手车价格大幅度降低，不容易出现小毛病，同时这个时候的零件磨合是最佳的状态，换个角度来讲，5年到8年左右的二手车，价格往往会变得变得低廉了，也就是最适合入手的时机。



这是二手车在售车龄的行驶里程的箱线图，从图中可以看出上边缘和下边缘分别是17（万公里）和0，上四分位数和下四分位数分别是9（万公里）和4（万公里），中位数是7（万公里）。还可以看出，均值和中位数都是比较靠近7（万公里）附近。还有1辆开了27（万公里）以上的老车。上外栏1.5\*IQR约等于18（万公里）。

对于二手车的里程，要是说起来，最大的影响就是车价，然后就是车辆的易损件的消耗，内饰磨损，基本上没有听说过有人把车开坏的，车辆使用里程可能7万公里和10万公里没有什么区别的，一辆车发动机的设计寿命应该是不少于60万公里的，因此里程不是影响二手车的最大因素，车况才是，最多里程影响一点价格。

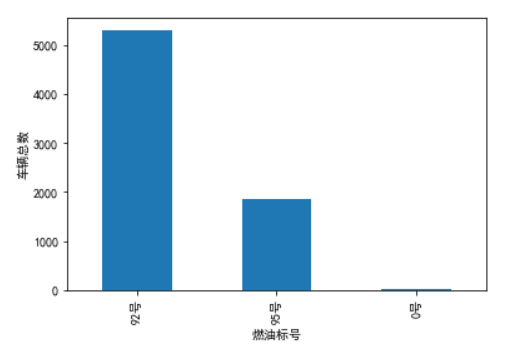
### 二手车在售销售价格占比图



根据上图可以看出在售二手车的销售价格在7.5万的占比最多，达到了25%。 卖家角度：对二手车定价的影响因素大致分为销量排名，车型换代，使用成本。

买家角度：对二手车价格的判定的方法1.市场价格比较法2.现值收益测算法3.重置成本核算法4.部件总成折旧计算法5.里程贬损核算法

**燃油标号使用车龄的直方图**



根据上面这幅燃油标号直方图可以看出，使用92号汽油的最多，最少的是0号（柴油）

导致使用92号汽油比95号汽油多的直接原因就是：实惠。

对于这款发动机来说，使用95号汽油油耗更低！根据不同的行驶路况和驾驶条件，综合下来，使用92号汽油和使用95号汽油百公里油耗会相差0.2升左右。比如使用92号汽油百公里油耗是8升，那么使用95号汽油百公里油耗大约是7.8升。我们以这个参数来算算这款车型使用92号汽油和使用95号汽油百公里价格差是多少：

使用92号汽油百公里油钱：6.87×8=54.96元

使用95号汽油百公里油钱：7.31×7.8=57.02元

一般来说都是小排量是推荐92号汽油，而对于大排量和涡轮增压发动机来说，都是推荐95号以上的汽油。个人认为，92也好，95也罢，前提你要保证油品的质量。不然你就算加95、98的汽油，也不会有多大的效果，甚至会损害发动机。

**五、总结与建议**

总结：通过上面的一些列图表，可以看出对二手车的选择要从几个总要因素去考虑，首先要考虑的就是车龄，看车辆的同时要结合车况和表显里程，车龄在5-8年之间的性价比最高，这时候的车不仅价格实惠，而且零部件磨合都达到了最佳状态。然后就是保值率，保价率（保价率=现价/原价）与使用年限和行驶里程呈现高度负相关；不同品牌车保价率随使用年限变化曲线略有不同，在列出的8个基数比较大的二手车品牌中，保价率随使用年限呈现3个梯队，以使用5年为参考基准：

保价率在0.5以上的：本田（约0.65）、日产（约0.55），均为日系车

保价率在0.5左右的：奔驰、奥迪、现代，两个德系，一个韩系

保价率在0.5以下的：宝马（约0.45）、长安（约0.45）、哈弗（约0.4），一个德系，两个国产。

二手车市场上，在品牌和质量方面齐名的"奥宝驰"不仅数量差别明显，在保价率上也是高下立判；日系车保价率居然最高；国产车的质量发展和品牌打造仍然任重而道远。

建议：我们在选择二手车时优先考虑品牌（比如：上汽大通，福特，丰田等等），这样有助于后期的维护，其二我们在选择二手车的时候选择车龄在5年左右的二手车，这个时间的二手车磨合是在最佳的状态，其三我们权衡车况和价格，做出大致判断。