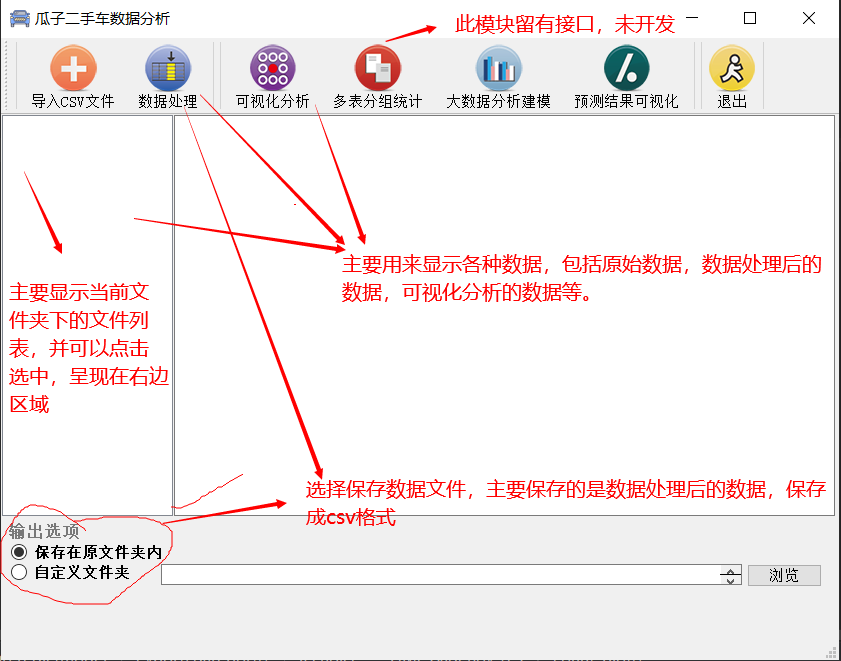
## 使用说明

文件目录说明

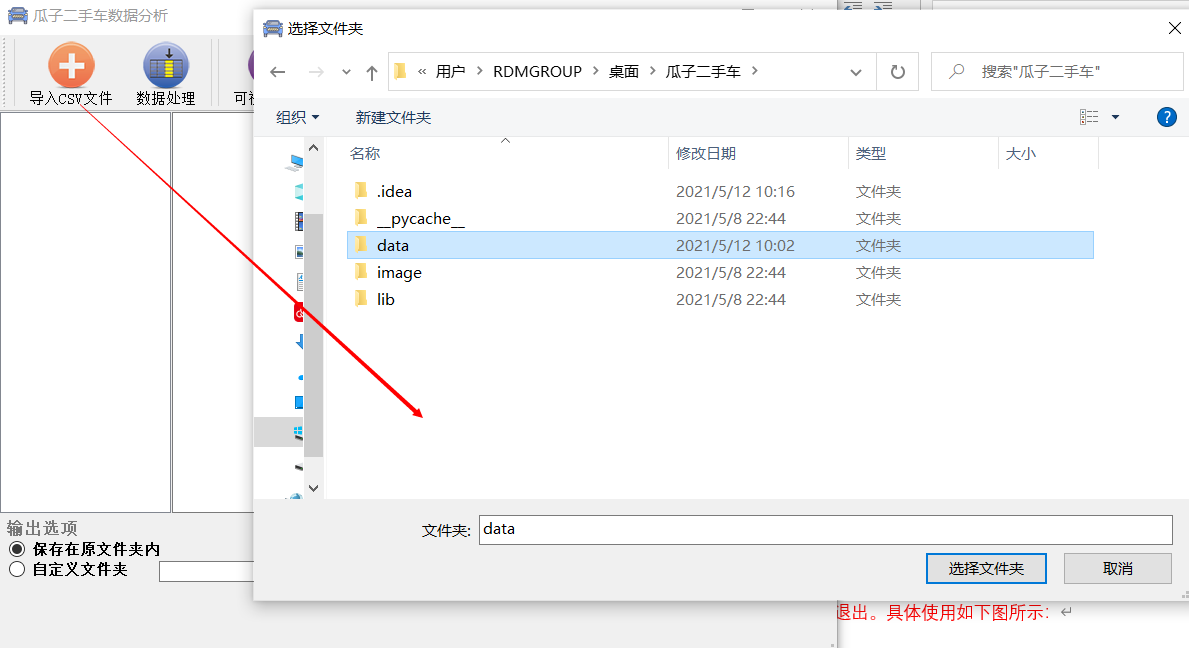
1. data文件夹 主要用来存放数据dataCar\_Columns.csv，当然此数据可以在任何地方在软件里面打开
2. image文件夹 主要用来存放制作软件界面所需的图标图片文件，不可更改，不可动，否则软件内相关图标会显示不正常
3. lib文件夹 主要用来存放程序编写中主窗口和子窗口公用的数据，已类的方式创建的，为py文件，不可动。
4. datacsv.ui 编写的软件主界面窗口的GUI程序。
5. viewresult.ui 可视化分析子窗口界面GUI程序。
6. ml\_predict\_model.ui 二手车车主报价预测模型子窗口界面GUI程序。
7. predict\_resu\_view.ui 预测结果可视化分析子窗口界面GUI程序。
8. main.py 软件主程序，打开通过运行子程序开始。
9. pachong.py 爬虫程序，爬虫瓜子二手车数据网站。
10. xiecheng\_pachong.py 异步爬虫程序，爬瓜子二手车网站，前期测试版本，使用时主要还是使用pachong.py 文件。

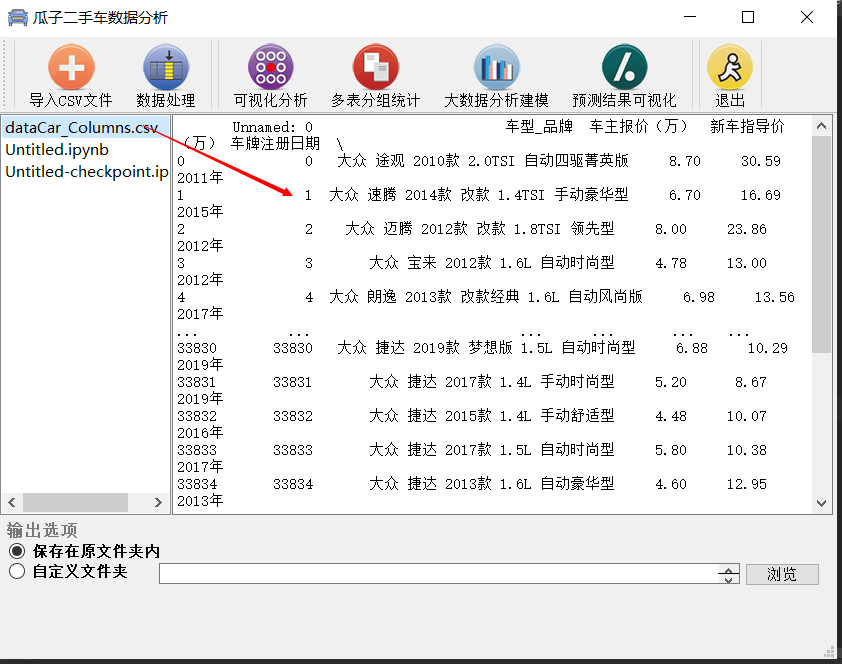
使用说明

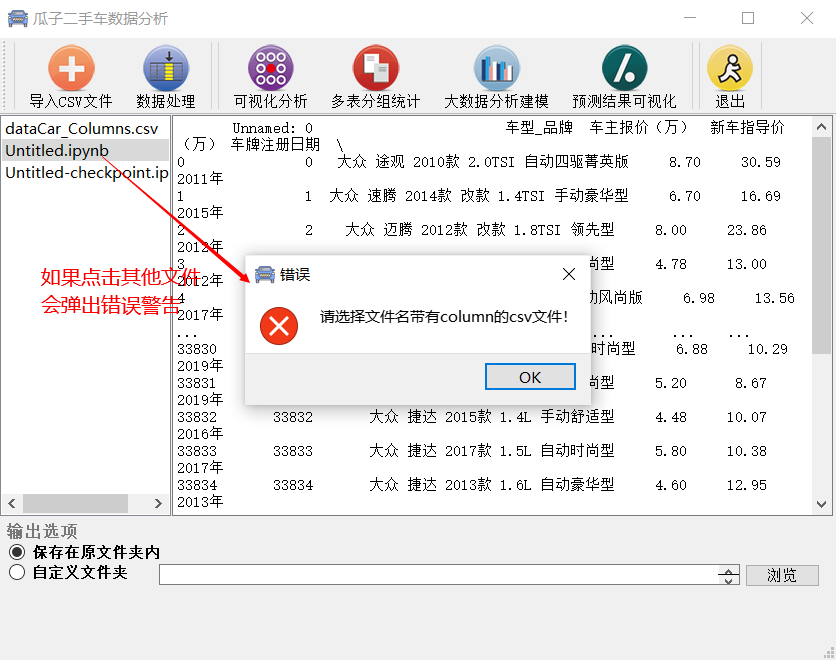
1. 首先配置好虚拟环境，如果本身有能开的环境也可以，但建议最好自己创建个专门的虚拟环境，装有pandas pyqt matplotlib numpy seaborn scipy sklearn lightgbm xgboost pandas\_profiling等库。
2. 运行main.py 库打开软件主界面，如下图所示：



1. 点击导入csv文件工具栏菜单，选择data目录下名为’ dataCar\_Columns.csv’的文件，此文件为经过爬虫程序爬下来的瓜子二手车原始数据。注意：只能选择这个文件，选错了软件可能会因所选数据格式不对不能正常运行，异常退出。具体使用如下图所示：

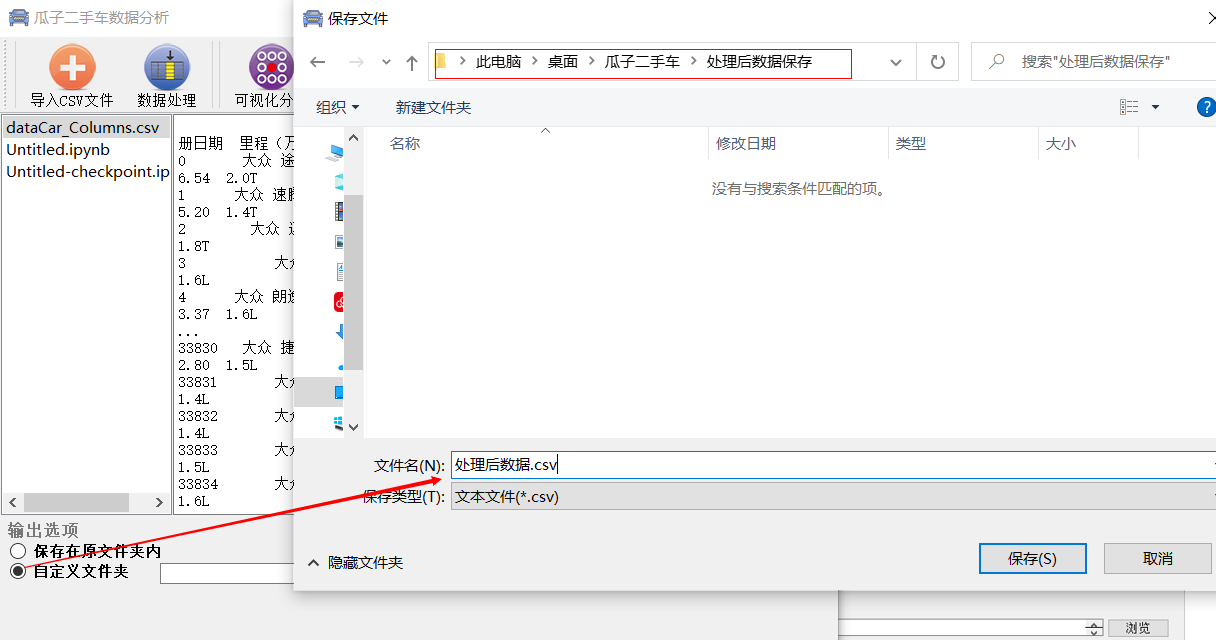


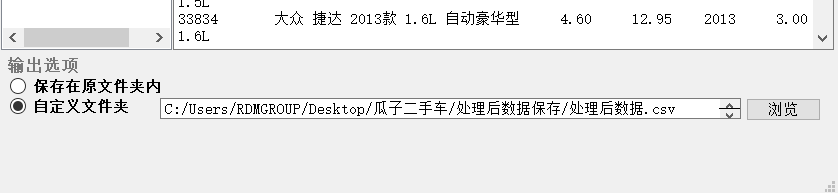




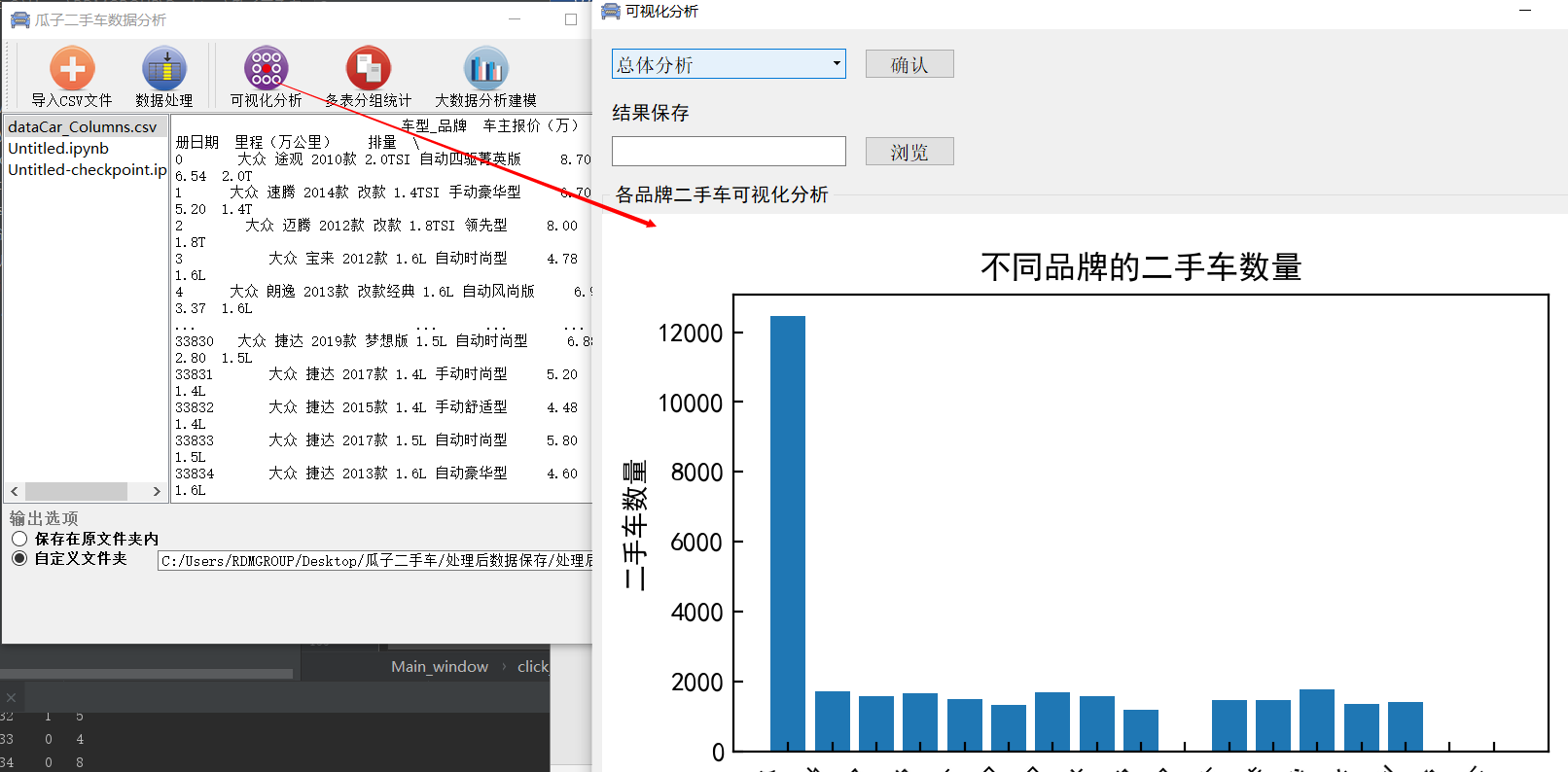
1. 在选中数据文件后，点击数据预处理，此时程序后台对数据进行处理，将处理好的数据更新在右边区域。



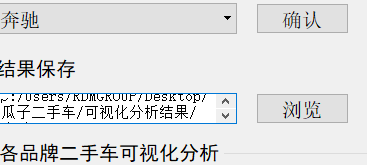
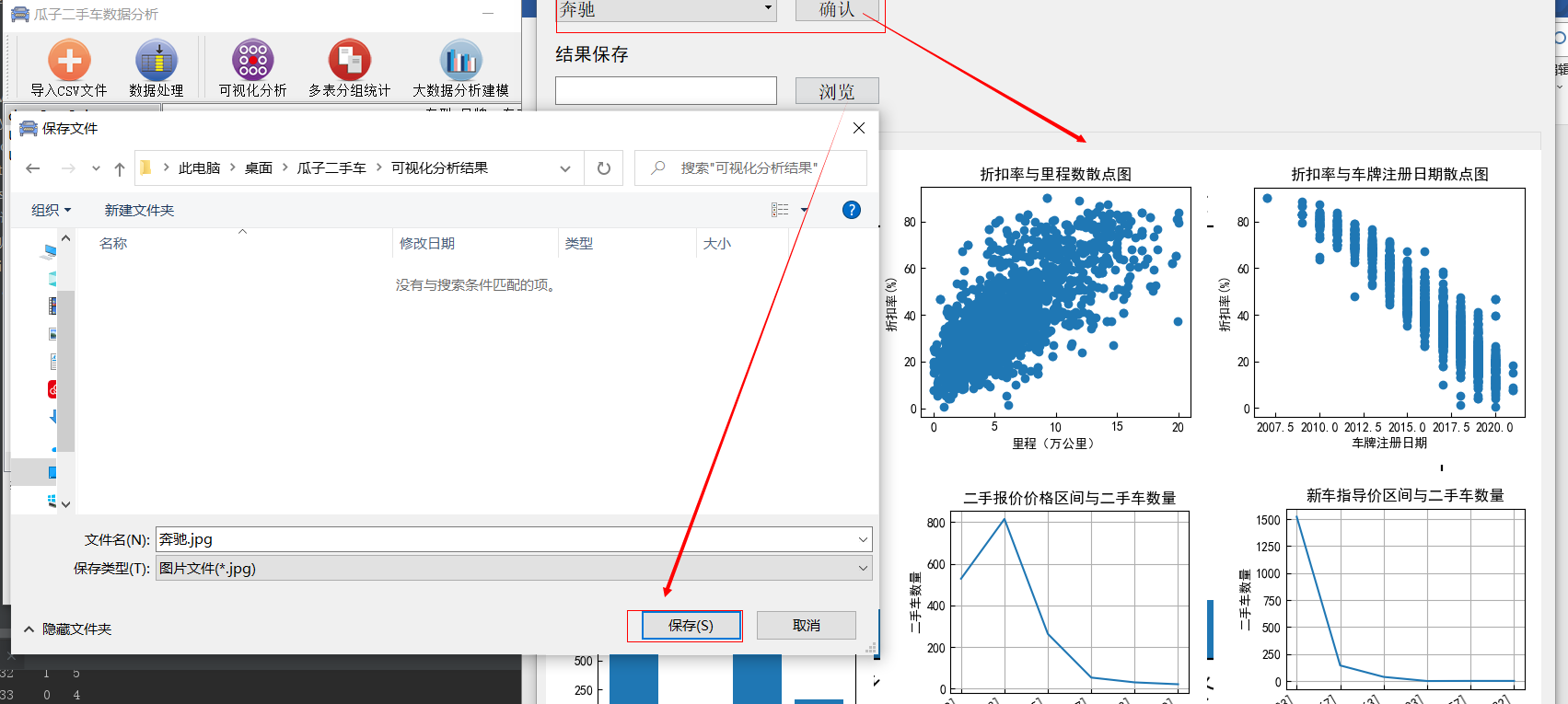


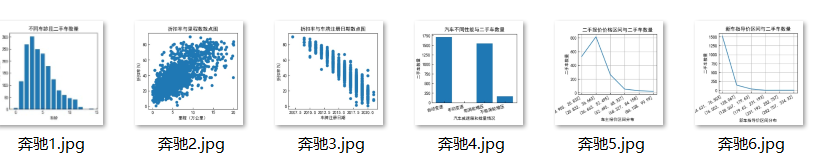


1. 点击可视化分析菜单，弹出可视化分析子页面，如下图：

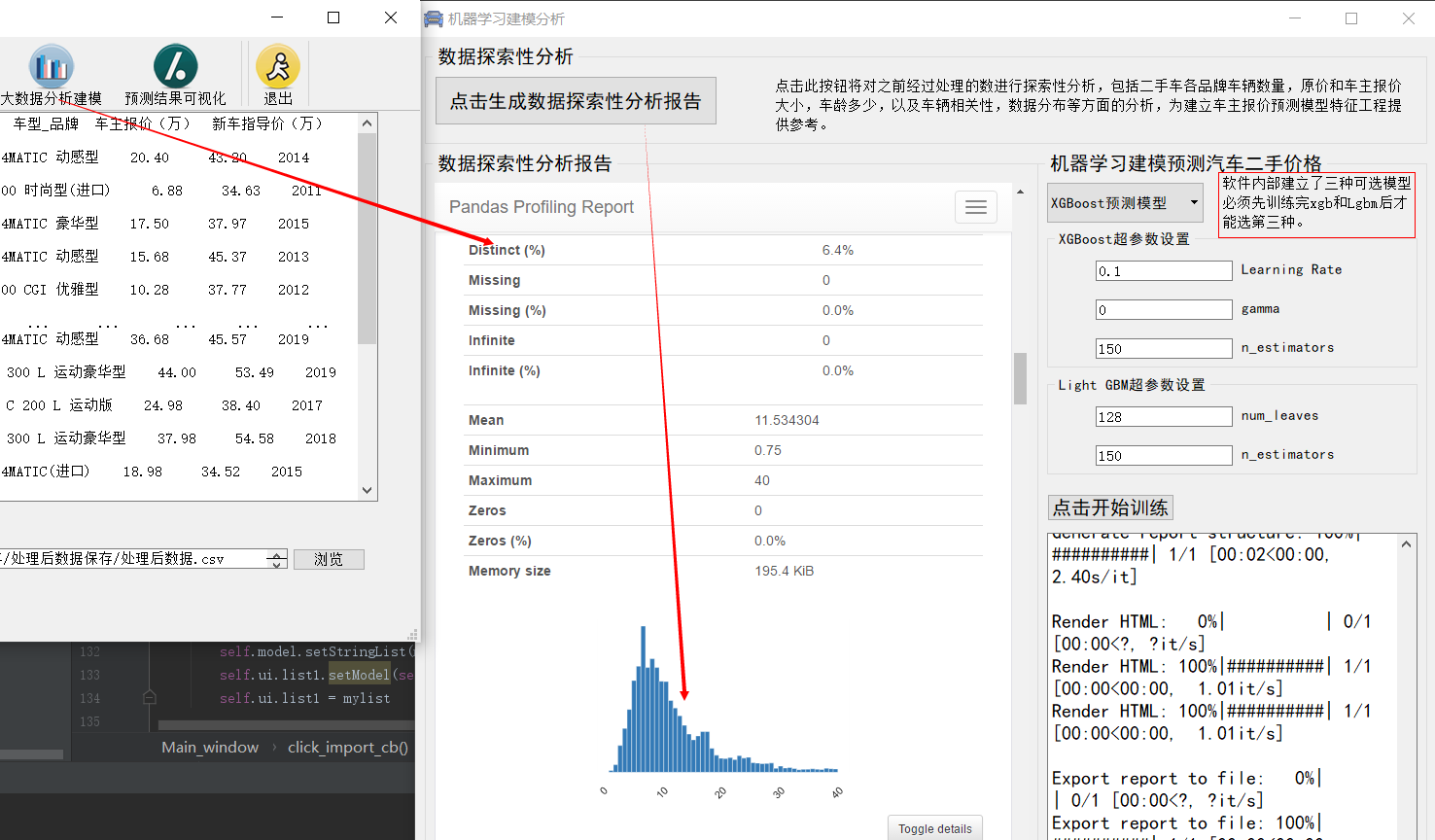


选择总体分析，即可对各种品牌车作总分析，也可选择单独品牌车进行分析，并保存。如下图：

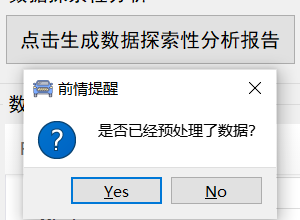


输入你想保存的文件名，软件默认将这6张图片已1 2 3 4 5 6的顺序保存在所选文件目录。

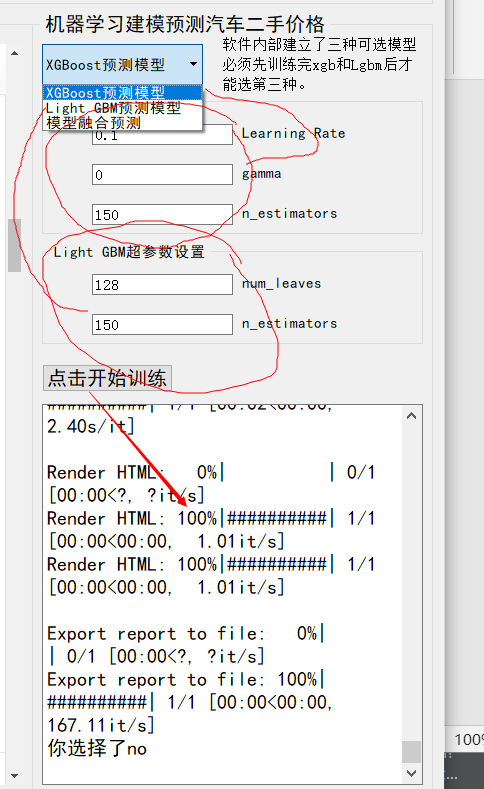
1. 选择大数据分析建模菜单，会弹出分析建模界面，如下图：

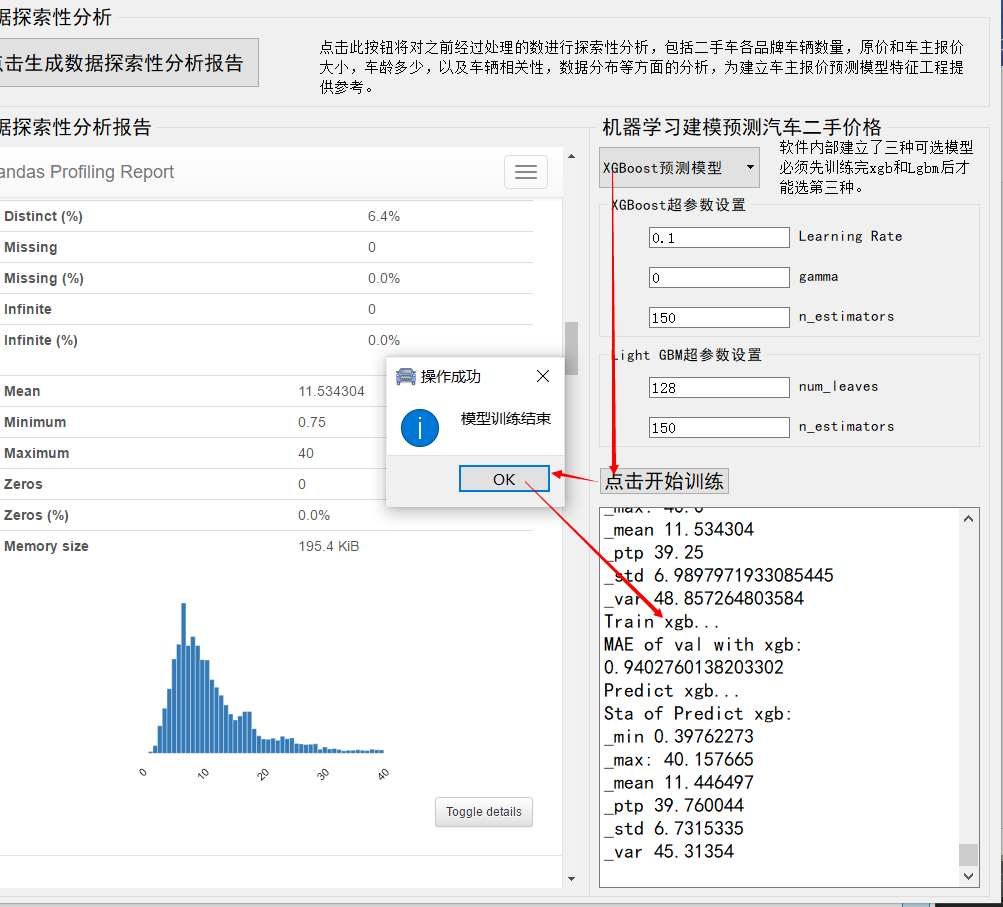


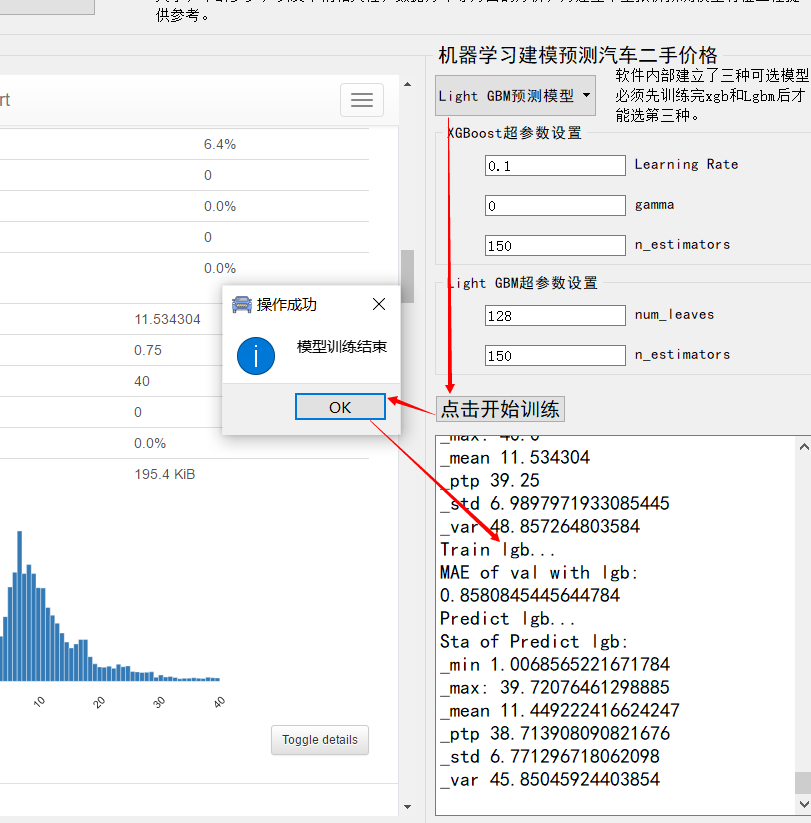
点击‘点击生成数据探索性分析报告’按钮软件在后台运用处理后的数据对数据进行各种探索性分析，此步骤为建模前必须步骤，所选数据也是得经过数据处理步骤处理后的数据，否则会弹出错误警告：

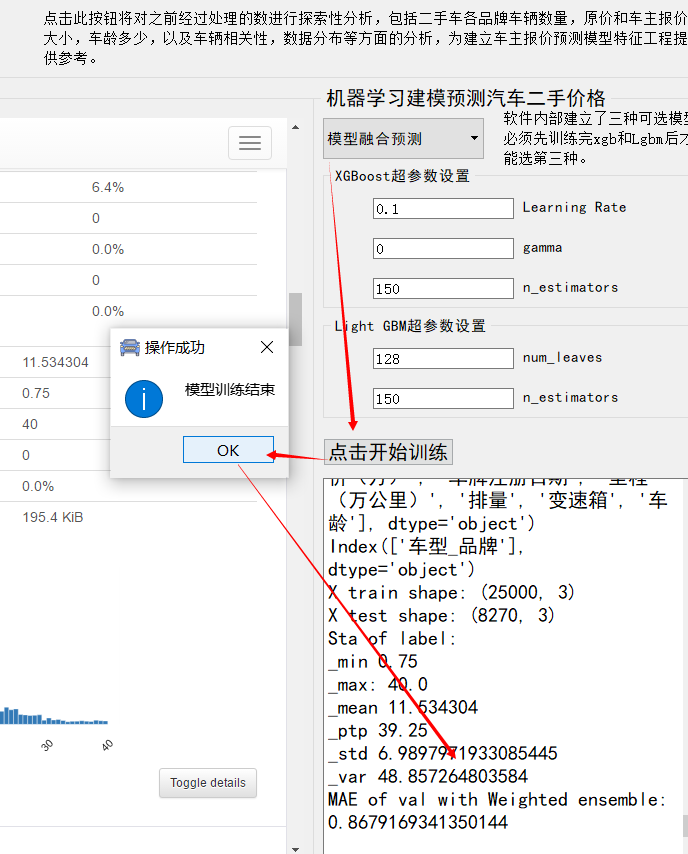
，就是这样。在此步骤运行完成后，该页面会弹出如上图所示的分析报告页面，同时，软件在本地main.py所在文件夹同级目录下生成了该报告，名为‘二手车数据分析EDA报告.html’。

1. 如下图所示，为预测方面，软件内置了三种预测模型，xgboost模型，lightbm模型，以及将上述两者融合的模型。点击下拉菜单可选择对应的模型，在选择好预测模型和设置好参数后，点击开始训练，系统会进行模型训练并对测试集数据进行结果预测。需要注意的是：你得先选择xgb或者lgb模型进行训练完成后，才能选择融合模型，因为这个是已前两者为基础的，同时设置参数时，选哪个模型，就在对应的参数模块下设置。

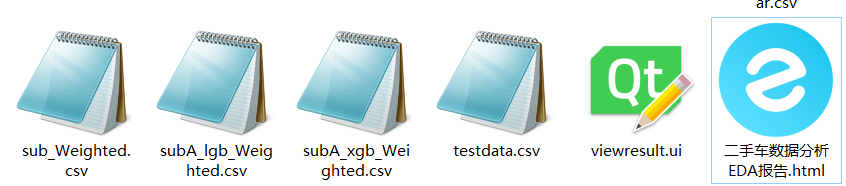






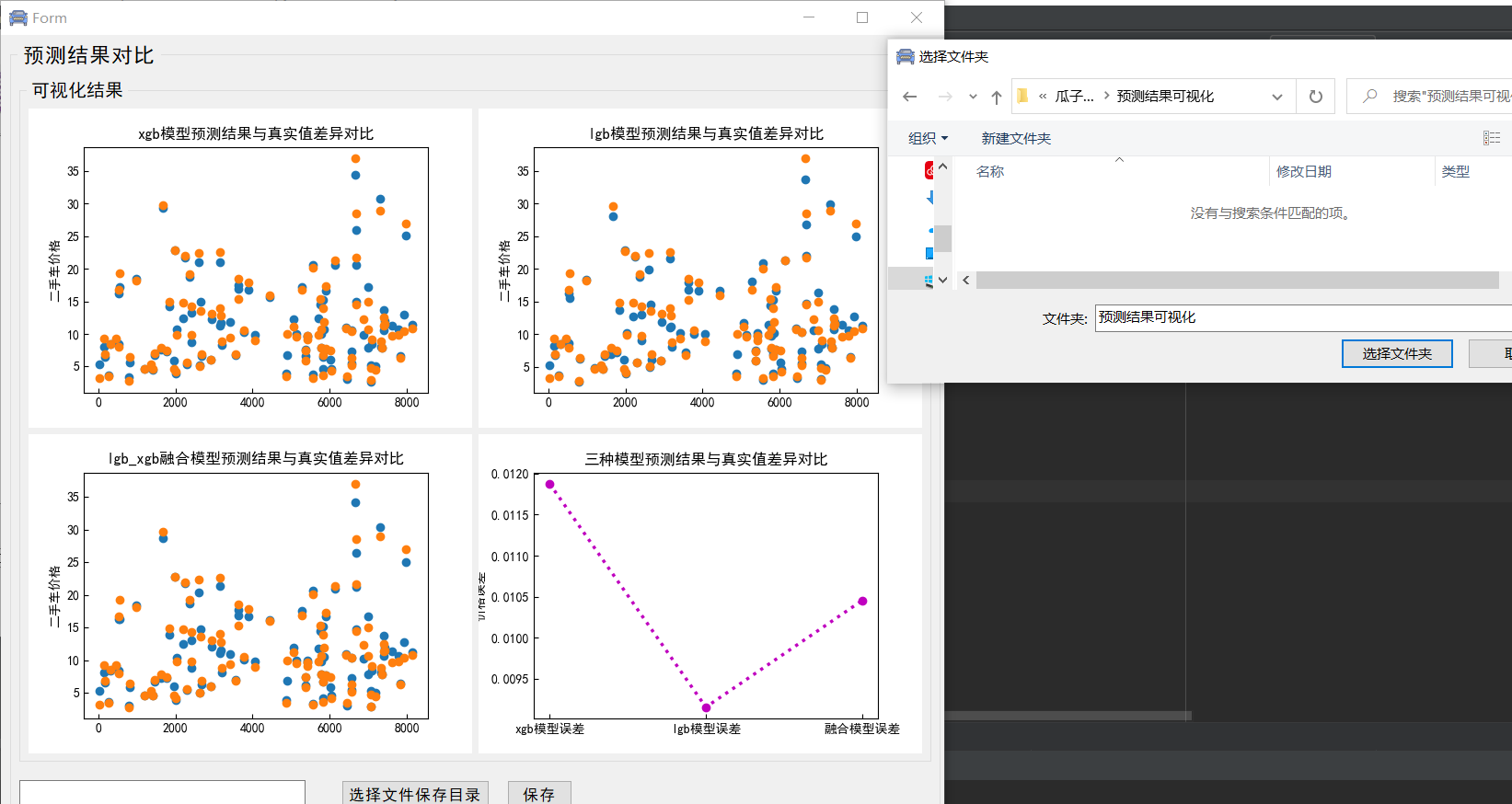


1. 在运行完此程序后，软件将在main.py文件同级目录下生成四个预测结果文件，分别为预测数据集文件，lgb模型预测文件，xgb模型预测文件，融合模型预测文件，如下图：

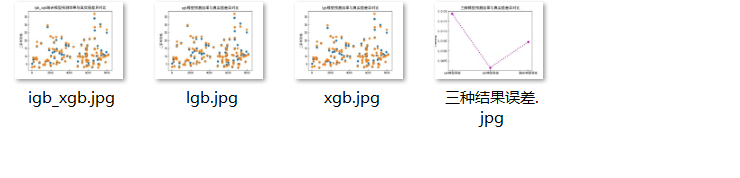


此文件在最后一个预测结果可视化模块将用到，所以在进行预测结果可视化模块时，需要先生成上述四个文件。

1. 点击预测结果可视化模块菜单，弹出该子页面，如下图：



会直接弹出分析好的结果呈现在页面，同时，用户还可以选择指定位置保存。软件以默认文件名保存。如下图文件名：



由上图结果可以看出，预测值与真实值比较接近，并且LGB模型的平均误差最小。