

汽车数据分析报告

数据分析团队

2025-09-07

目录

1	执行摘要	1
2	数据概况	1
3	分析结果	1
3.1	价格分布	1
3.2	车龄与价格的关系	3
3.3	行驶里程与价格的关系	3
3.4	马力与价格的关系	3
3.5	发动机排量与价格的关系	7
3.6	重量对价格的影响	7
3.7	车门数量对价格的影响	7
3.8	金属漆对价格的影响	7
3.9	变量相关性分析	7
3.10	回归模型分析	7
4	结论与建议	12

1 执行摘要

本报告分析了汽车价格与各种因素（如车龄、行驶里程、马力等）之间的关系。基于对 1435 辆汽车的数据分析，我们发现车龄和行驶里程是影响汽车价格的最重要因素。

2 数据概况

本次分析共包含 1435 条汽车记录，价格范围从 4350 到 3.25×10^4 美元。数据的基本统计摘要如下：

表 1：数值变量统计摘要

age	km	weight	hp	cc	doors	price
Min. : 1.00	Min. : 1	Min. :1000	Min. : 69.0	Min. :1300	Min. :2.000	Min. : 4
1st Qu.:44.00	1st Qu.: 43000	1st Qu.:1040	1st Qu.: 90.0	1st Qu.:1400	1st Qu.:3.000	1st Qu.:
Median :61.00	Median : 63451	Median :1070	Median :110.0	Median :1600	Median :4.000	Median :
Mean :56.03	Mean : 68572	Mean :1072	Mean :101.5	Mean :1567	Mean :4.033	Mean :10
3rd Qu.:70.00	3rd Qu.: 87042	3rd Qu.:1085	3rd Qu.:110.0	3rd Qu.:1600	3rd Qu.:5.000	3rd Qu.:
Max. :80.00	Max. :243000	Max. :1320	Max. :192.0	Max. :2000	Max. :5.000	Max. :32

3 分析结果

3.1 价格分布

汽车价格呈右偏分布，大部分汽车价格集中在较低水平。

3.2 车龄与价格的关系

车龄与价格呈现明显的负相关关系，车龄越大，价格越低。这表明汽车折旧是影响价格的重要因素。

3.3 行驶里程与价格的关系

行驶里程与价格也呈现负相关关系，行驶里程越高，价格越低。这表明行驶里程是影响二手车价格的重要因素。

3.4 马力与价格的关系

马力与价格呈现正相关关系，马力越大，价格越高。这表明汽车性能对价格有积极影响。

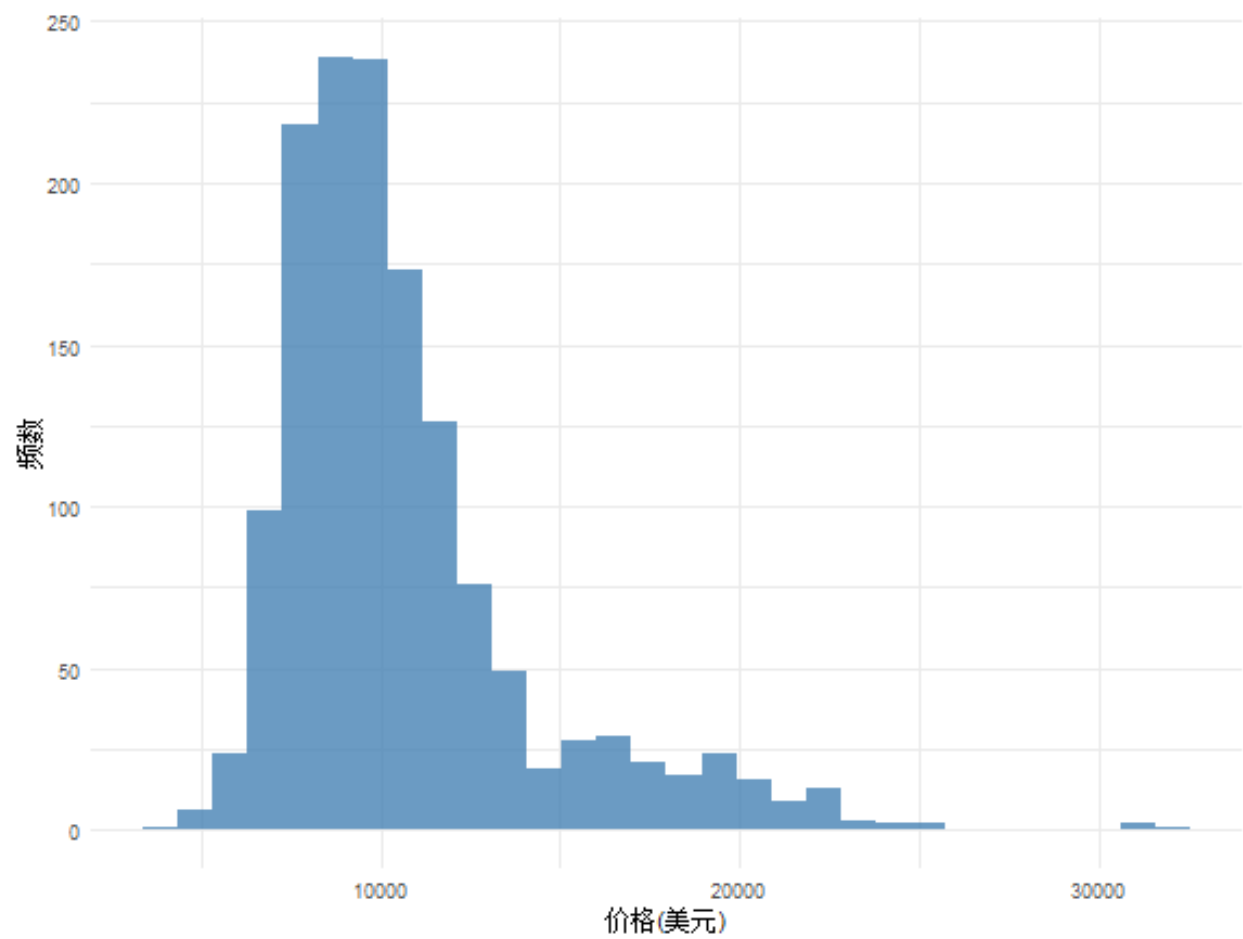


图 1：汽车价格分布直方图

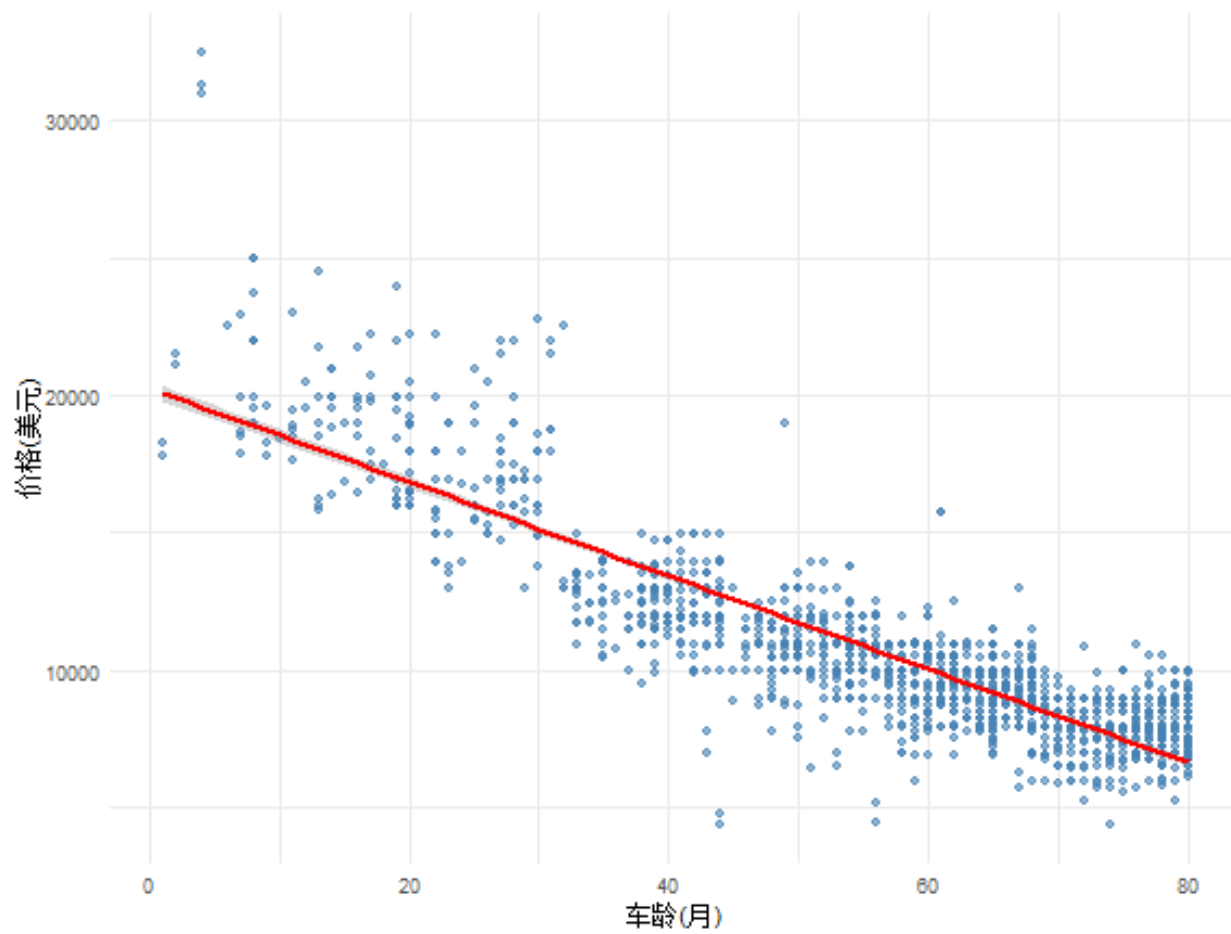


图 2：车龄与价格的关系散点图

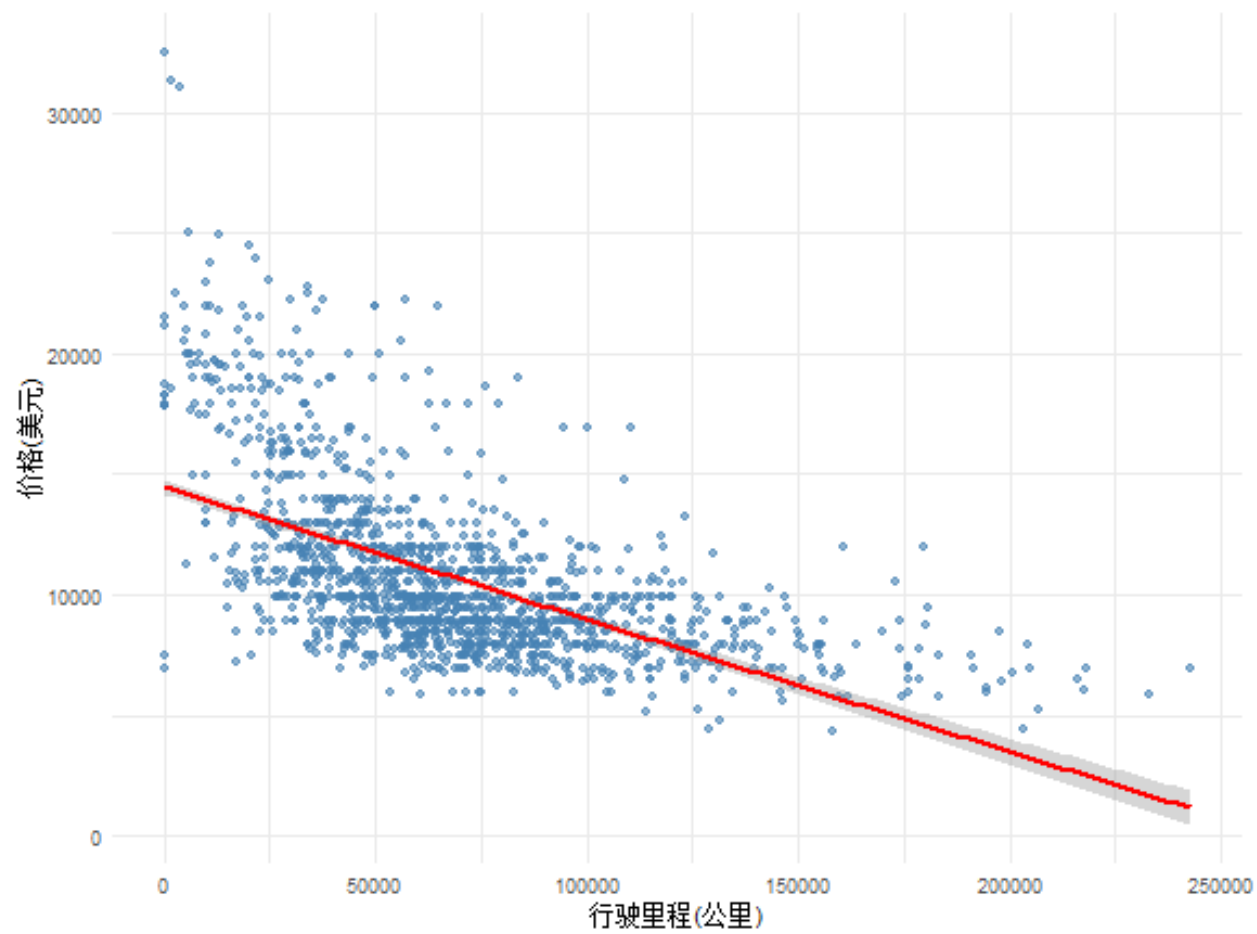


图 3：行驶里程与价格的关系散点图

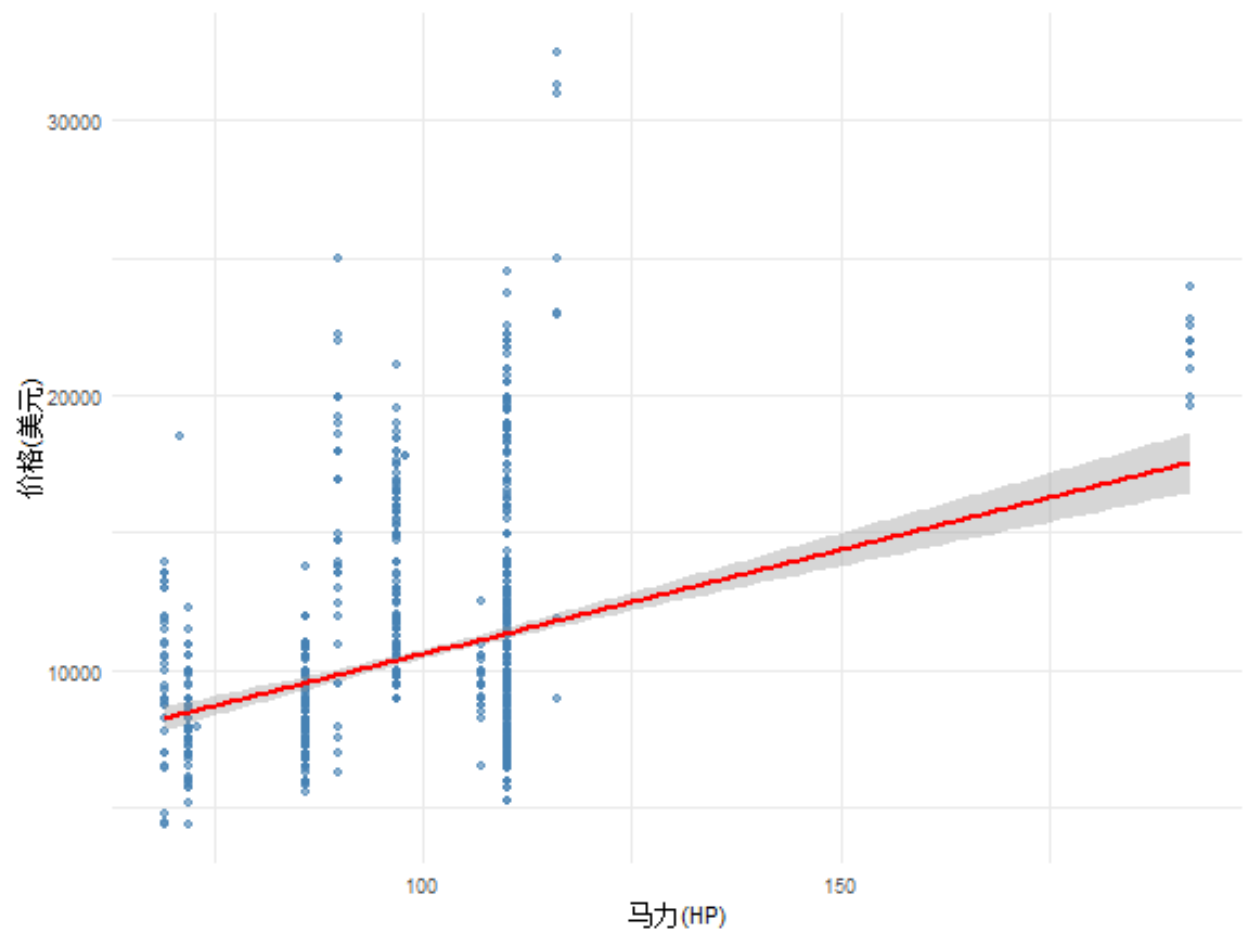


图 4: 马力与价格的关系散点图

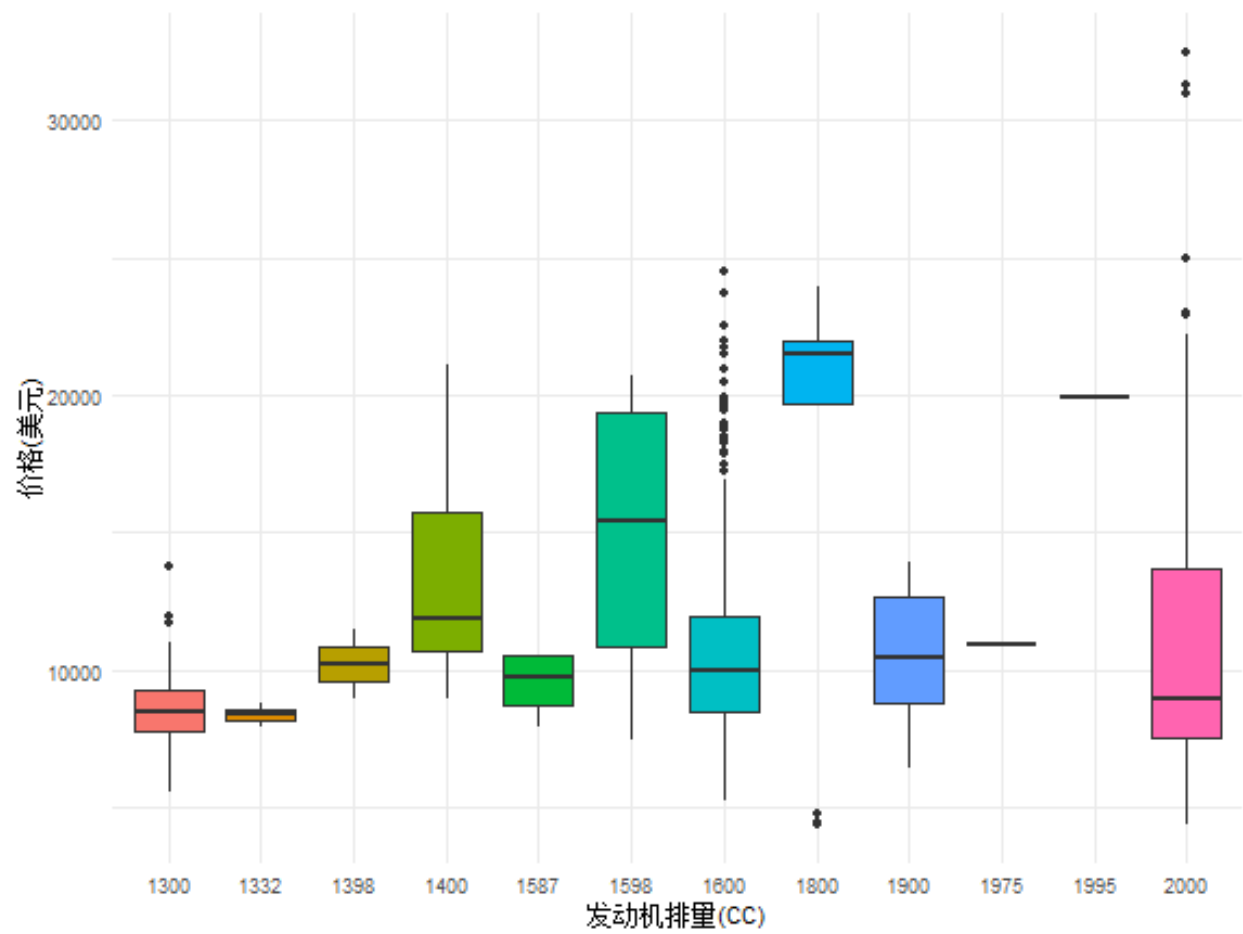


图 5：发动机排量与价格的关系箱线图

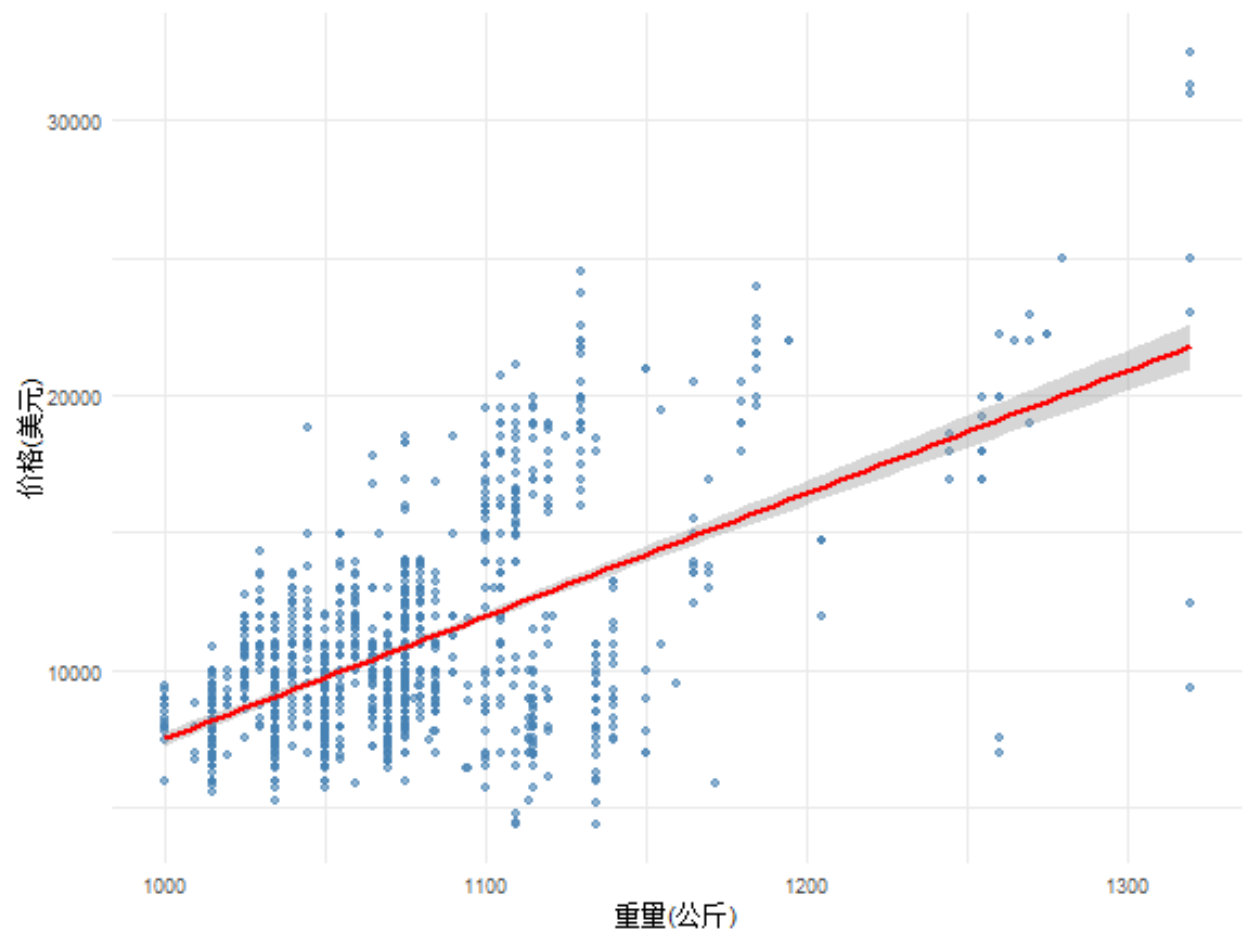


图 6: 重量对价格的影响散点图

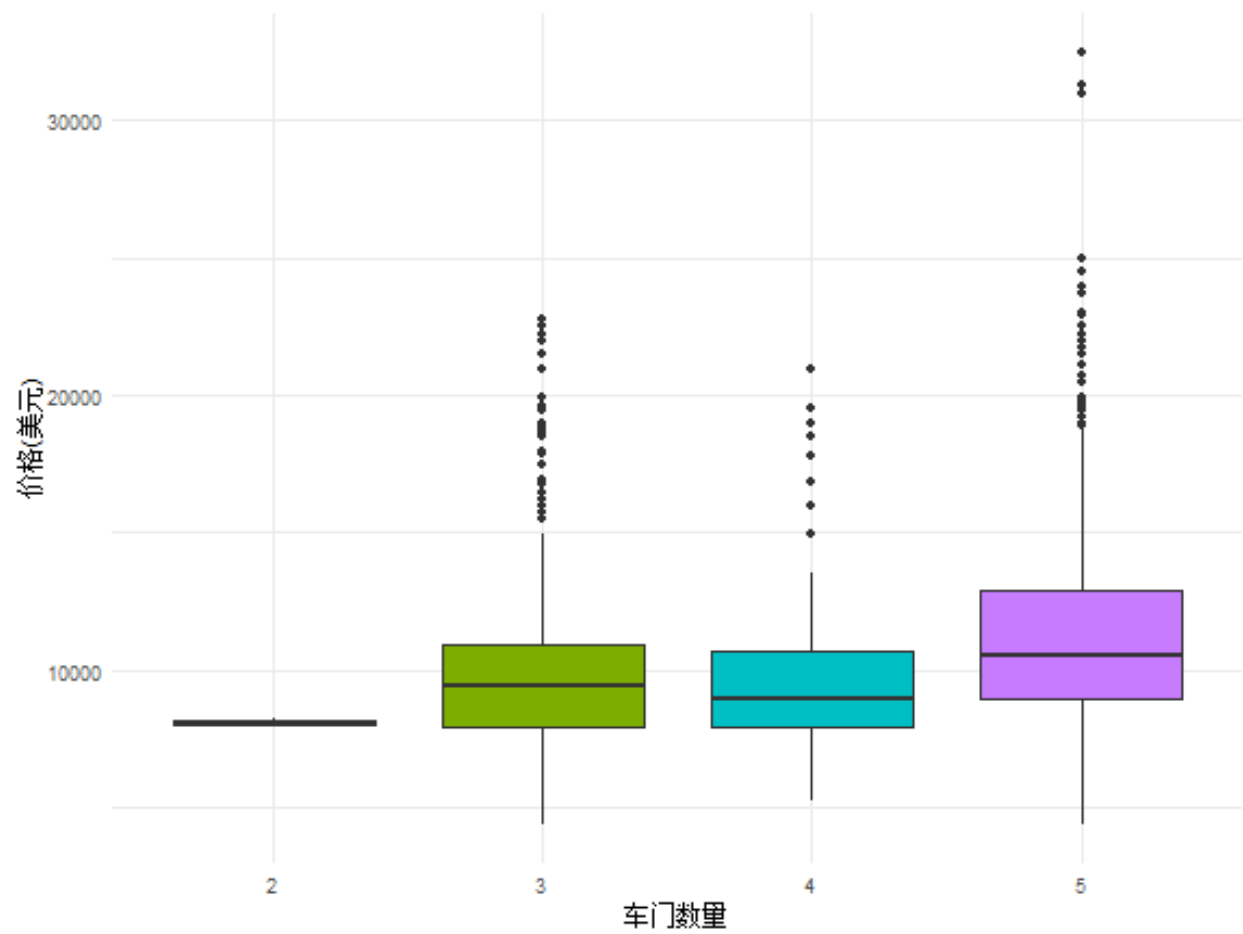


图 7： 车门数量对价格的影响箱线图

3.5 发动机排量与价格的关系

3.6 重量对价格的影响

3.7 车门数量对价格的影响

3.8 金属漆对价格的影响

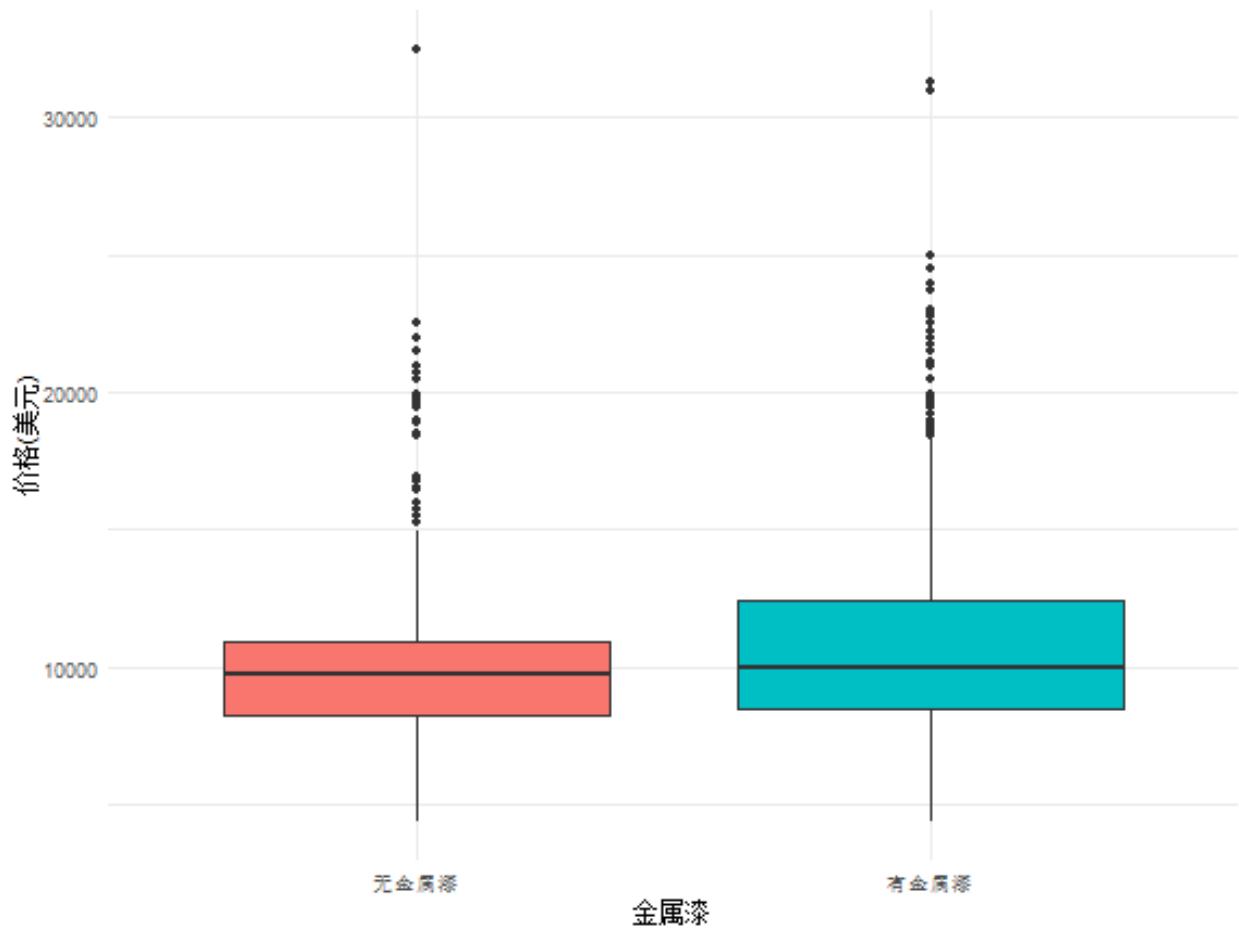


图 8: 金属漆对价格的影响箱线图

3.9 变量相关性分析

3.10 回归模型分析

我们建立了一个多元线性回归模型来预测汽车价格:

##

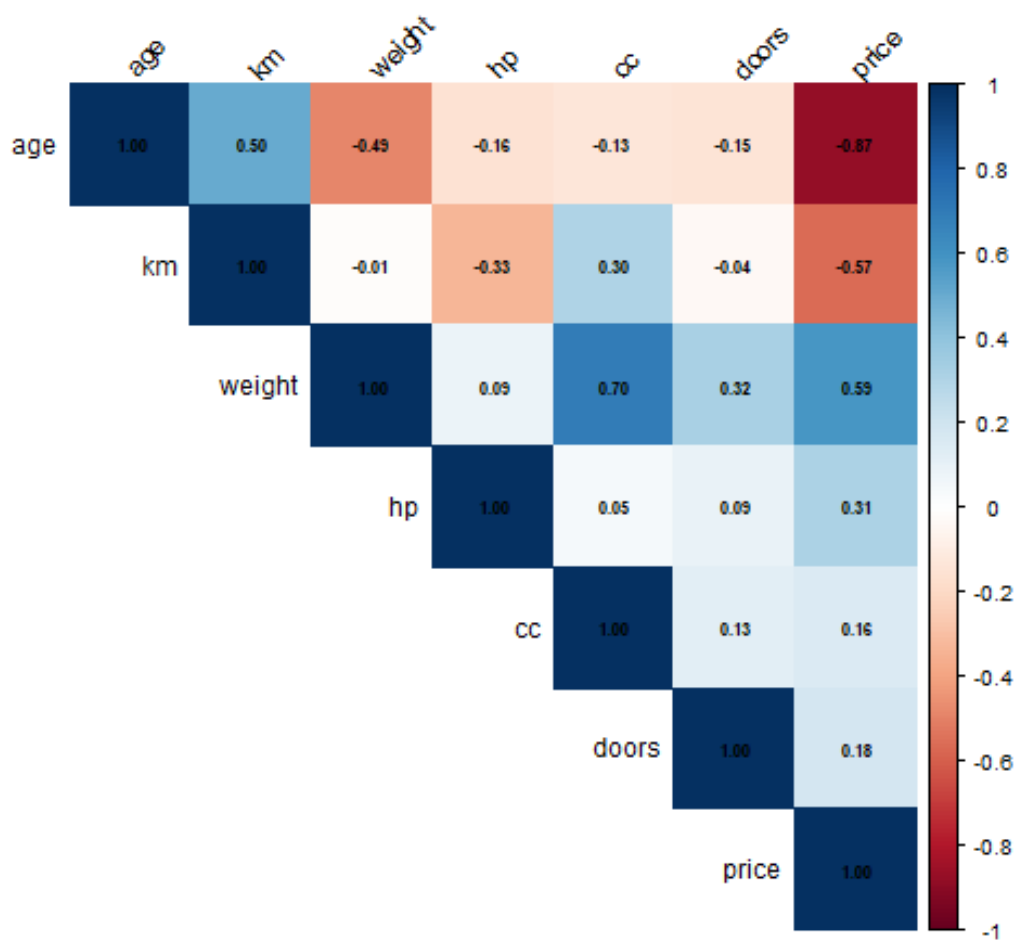


图 9: 变量相关性矩阵

```
## Call:
## lm(formula = price ~ age + km + weight + hp + met_color + cc +
##     doors, data = car_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9258.1  -771.3    -5.4    739.8   8060.0
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.209e+04  1.175e+03 -10.285  < 2e-16 ***
## age          -1.157e+02  2.647e+00 -43.729  < 2e-16 ***
## km           -1.705e-02  1.270e-03 -13.424  < 2e-16 ***
## weight       2.973e+01  1.295e+00  22.955  < 2e-16 ***
## hp           3.292e+01  2.515e+00  13.087  < 2e-16 ***
## met_color    4.612e+01  7.500e+01   0.615  0.53874
## cc          -2.741e+00  2.990e-01  -9.170  < 2e-16 ***
## doors       -1.181e+02  3.930e+01  -3.006  0.00269 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1318 on 1427 degrees of freedom
## Multiple R-squared:  0.8672, Adjusted R-squared:  0.8666
## F-statistic: 1332 on 7 and 1427 DF,  p-value: < 2.2e-16
```

表 2: 回归系数表

term	estimate	std.error	statistic	p.value
截距	12085.92	1175.06	-10.29 <	0.001
车龄	-115.75	2.65	-43.73 <	0.001
行驶里程	-0.02	0.00	-13.42 < 0	001
重量	29.73	1.30	22.96 <	0.001
马力	32.92	2.52	13.09 <	0.001
金属漆	46.12	75.00	0.61 0.	39
发动机排量	-2.74	0.30	-9.17 < 0.	01
车门数量	118.14	39.30	-3.01 0.0	3

表 3: 模型整体统计量

r.squared	adj.r.squared	sigma	statistic	p.value	df	logLik	AIC	BIC	deviance
0.867	0.867	1318.165	1331.541	0	7	-12341.2	24700.4	24747.82	2479496205

4 结论与建议

1. **主要发现:** 车龄和行驶里程是影响汽车价格的最重要因素, 两者与价格均呈负相关关系。
2. **性能影响:** 马力和发动机排量对价格有积极影响, 表明汽车性能是消费者考虑的重要因素。
3. **其他因素:** 重量、车门数量和金属漆也对价格有一定影响。
4. **建议:**
 - 购买二手车时, 应重点关注车龄和行驶里程
 - 卖车时, 应合理定价, 考虑车龄、行驶里程和车辆性能
 - 对于有金属漆、更多车门的车辆, 可以适当提高定价

模型解释力: 我们的回归模型解释了价格变异的 86.7% ($R^2 = 0.867$), 表明模型具有较好的预测能力。