

2018.11.16 #每周干货#

同一个城市里，房价与哪些因素高度相关？如何使用 R 语言做一个回归模型？知道房屋的地理位置和周边 POI，如何大致预测房价？

本期【每周干货】《大佬，您的选房系统已上线》为你——道来~

请戳这儿: <https://mp.weixin.qq.com/s/soPSbvJJnostZAE4J0GOKQ>

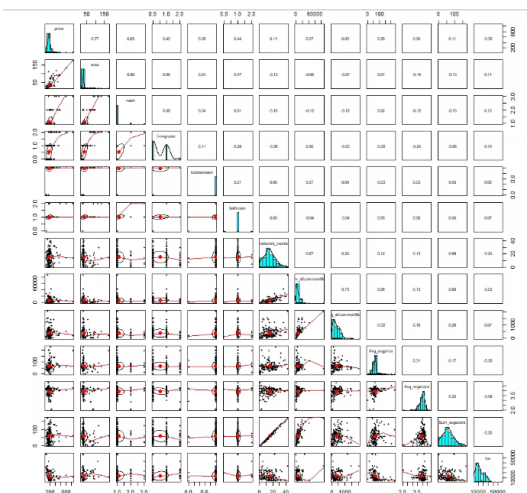
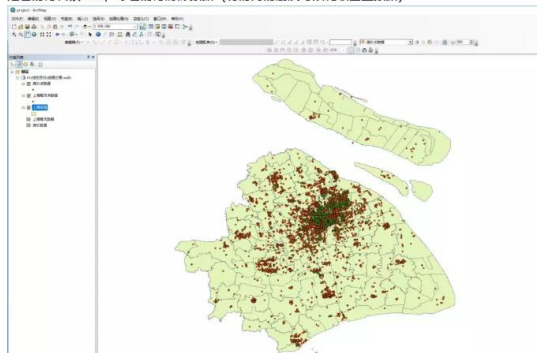
另外，我准备了另一份数据：美团美食频道1W+的餐饮POI，同样使用爬虫获取。

[illegible]

建模准备

这一部分我们会将房源数据和POI数据等信息整合在一起用于建模，可以用ArcGis完成。

红色的为餐饮POI，绿色的为房源数据（有时间的朋友可以爬取全量数据）



理解几个关键点：

1. 相关性系数，绝对值越接近1表明相关性越显著，比如房价和面积相关度为0.77；
2. 图中的直方图为每个维度的数据分布；
3. 散点图中圆圈完美表明相关度越低，椭圆形表明相关度显著。
4. 图中的线条是查看线性关系的，可以体现自变量和因变量之间的关系。

下面，我们使用全部属性构建一个回归模型。Price是因变量，“~.”表示导入所有自变量。最后打印模型结果查看。

```
model <- lm(price~.,data=model_data)
print(summary(model))
print(car::vif(model))
```

更多每周干货分享+好玩的数据分析小技巧

欢迎关注：数据团学社



Python 数据分析师微专业限时优惠，不想错过的话请戳这儿→

https://mooc.study.163.com/smartSpec/detail/1001359001.htm?utm_campaign=commission&utm_medium=cps&share=1&shareId=1135186195&hideAppEntrance=1&utm_source=400000000123092

了解往期干货，请移步群文件【每周干货】