总概：

Histogram用来分析某一个因素在一年中不同月份的变化。

Heatmap饼图，和堆叠柱状图用来分析Area,Group,Reservation,Clients'age,这四大类，用不同的形式来展现。

Gender和occupancy只有一个变量和月份，是一维的，因此不需要饼图和热点图等高维图像，使用Histogram查看即可。

交互联动方式，通过点击四种按钮，图一起更新。

尝试了<http://bl.ocks.org/NPashaP/96447623ef4d342ee09b>

的dashboard，但是d3版本不一致以及时间不够了没有实现出来，应该把api改成v4版本的就能出图了，但是修改很多而且繁杂，时间不够用，后续有时间可以实现。

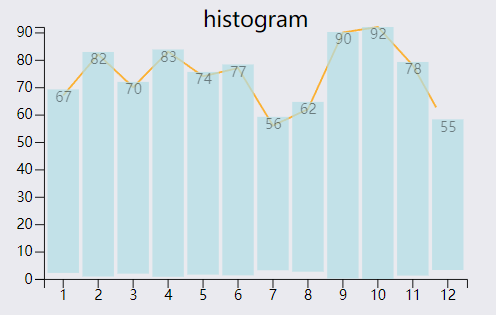
练习

* 基于hotel数据的进行**完整的**可视分析系统的设计和实现，要求各个视图模块可以交互联动，互相配合。同时提交1到2页的数据分析结果。回答下列问题，并阐述用于得到和验证结论的视图和分析方法。

1. 找到酒店淡季、旺季的时间段

假定都是两个月

根据histogram occupancy变量找到 旺季9，10 淡季7，8

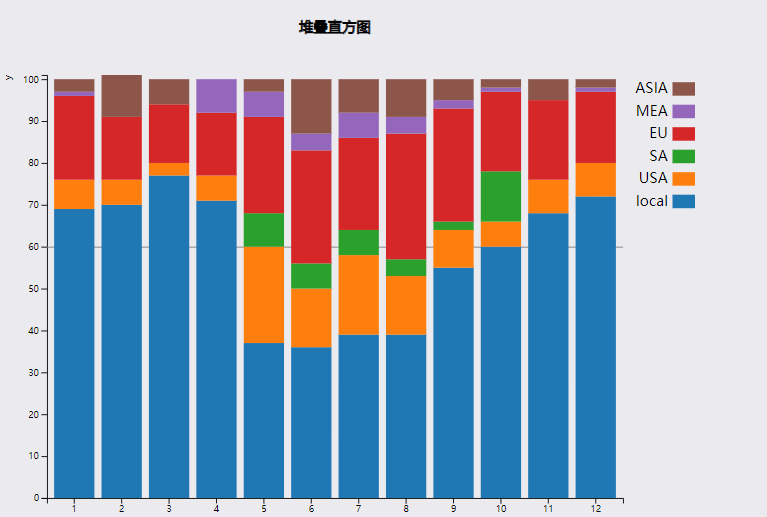


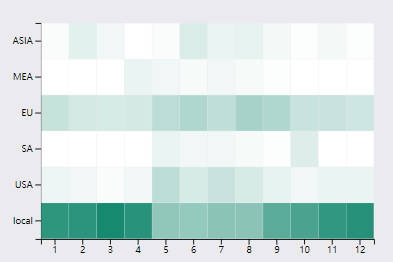
1. 分析酒店入住人员的特点，尝试回答：酒店住客有什么特征？12个月中住客特点是否发生过变化？如果是，分析哪些因素可能导致了变化。

特点类型：Area,Group,Reservation,Clients'age

（1）

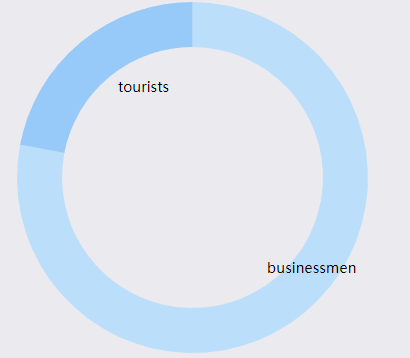
特点1：Area：

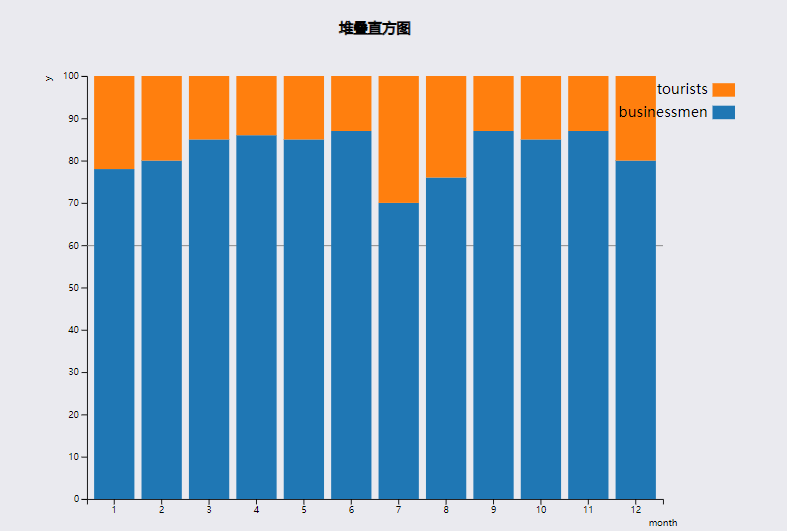


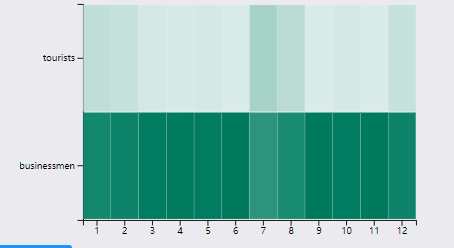


可以看到本地人占大头，其次是EU的，其余的很少。

特点2：Group：







可以看到商人比游客约为4：1。

特点3：Reservation：



可以看到通过DR方式订购的占比70%左右，agency约为20%,AC 10%.

特点4：Clients'age：



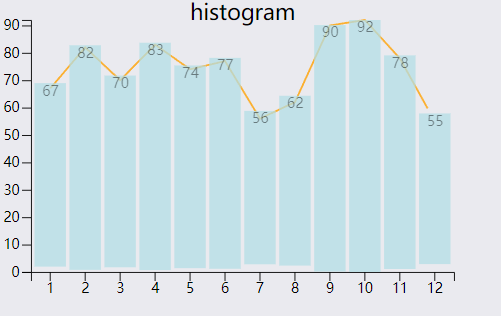
小于20岁的几乎没有，大于55和20-35之间的分别占比20%左右，35-55占比50%左右。

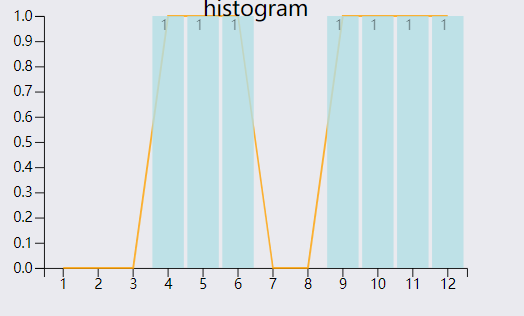
综上，大多是本地35-55岁的商人通过DR订购的酒店。

（2）发生过变化。

注意到拥挤度和当地是否有传统节日直观上会有关系，查看直方图：

拥挤度：

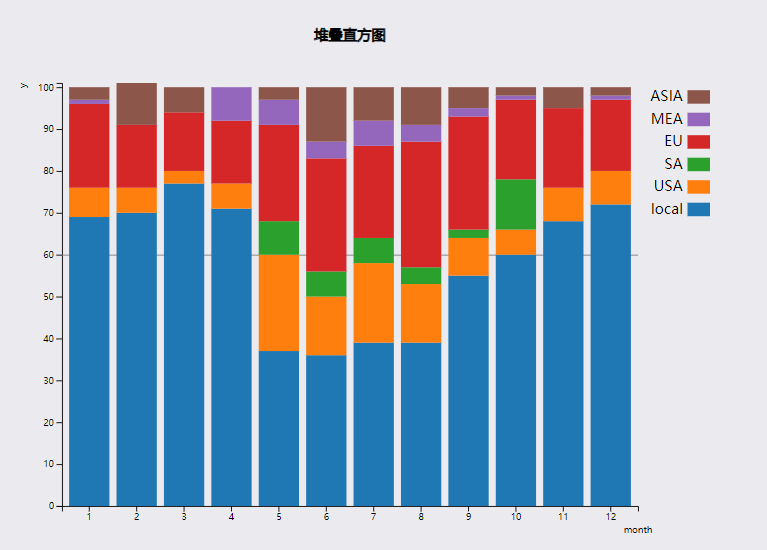


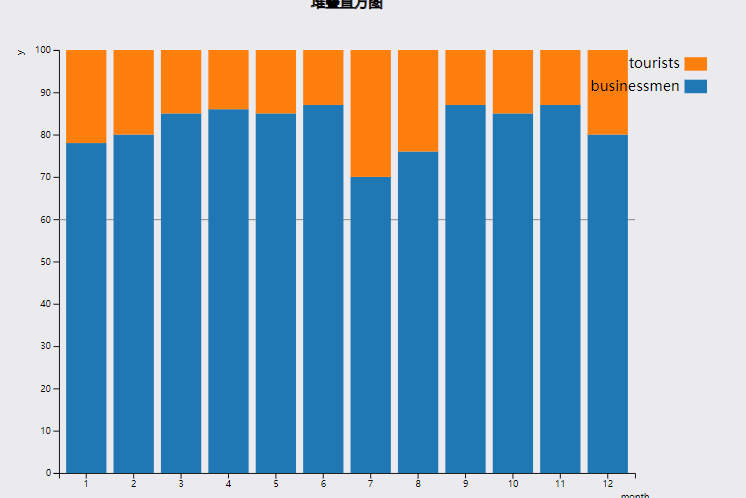
是否有传统节日：  


可以看到有传统节日的生意更好。

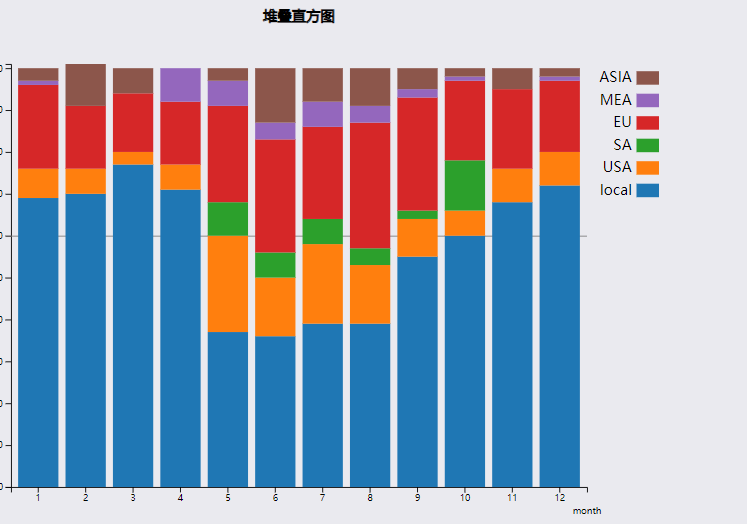
季节的因素也影响了游客商人占比等变量，具体分析见第三问。

1. 分析季节因素对酒店的影响。并分析哪些因素有具有相似的季节性变化规律。

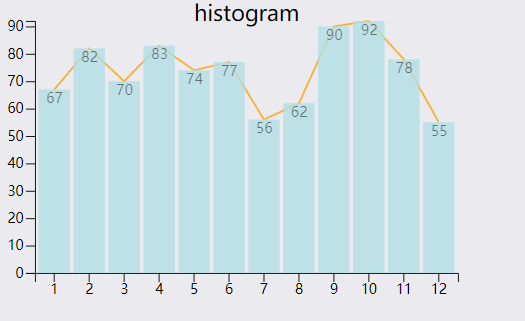


area和group栏目相似，显然地域是其他国家越多，游客占比越多。

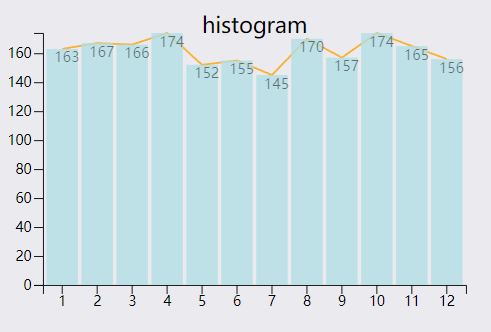
从area一栏可以看到，5，6，7，8月份明显本地人减少，以及拥挤度也减少。



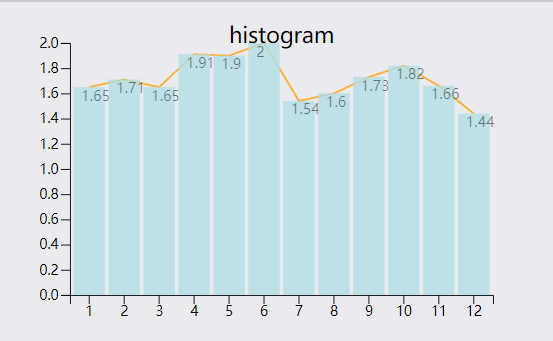
拥挤度：

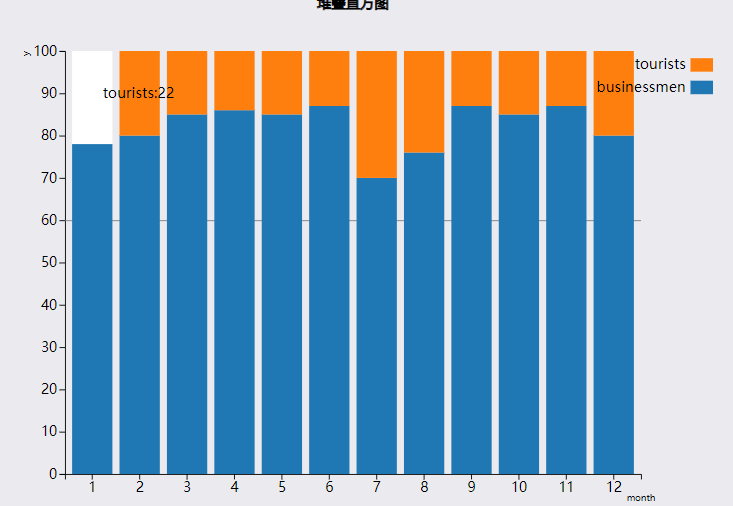


个人猜测可能是暑假本地人出去玩，游客来玩的相对占比多了， 但是又不是著名景点，来的没有走的多，因此拥挤度下降。既然拥挤度下降了，那自然而然酒店需要降低价格。价格如图，符合猜测。

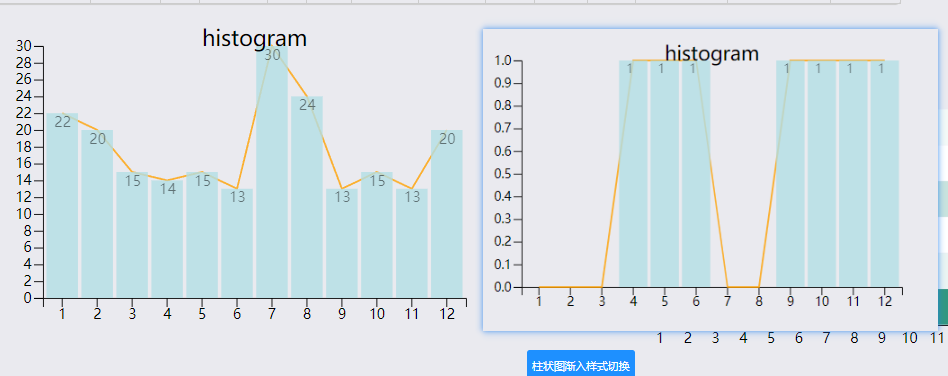


住宿天数和游客商人占比：





可以看到，商人占比和住宿天数成正比，这也好理解，商人大多做生意，停留时间长，游客住的时间短。

1. 除了以上问题外，写出自己分析过程中得到的其他结论。

左图为游客占比，右图为传统节日，我本以为传统节日和游客占比正相关，但是不然，说明游客并不是因为传统节日而吸引来的。