

# Assignment 8

---

## 1.概述

本项目基于Flask后端框架与JS前端编写，使用d3对数据进行可视化，其中图表之间需要联动配合，并且在可视化系统对数据进行分析，完成一系列分析的任务，任务及解决思路具体情况如下：

- 找到酒店淡季、旺季的时间段。使用数据中的入住率Occupancy和价格Price作为参考，将两者相乘得到评估营收的参数Average Income，并使用柱状图进行可视化。为分析产生淡/旺季的原因，考查价格Price，是否举办大型活动Conventions与入住率Occupancy的关系，使用散点图进行表示，其中不同的Conventions表示为不同颜色的散点。
- 分析酒店入住人员的特点。从数据分布上看，使用百分比堆积柱状图更合适，分别从地域分布、性别分布、年龄分布、所属人群四个方面分析：
  - 地域：Local、USA、SA、EU、MEA、ASIA；
  - 年龄：U20、20to35、35to55、m55；
  - 居住目的：Businessmen、Tourist；

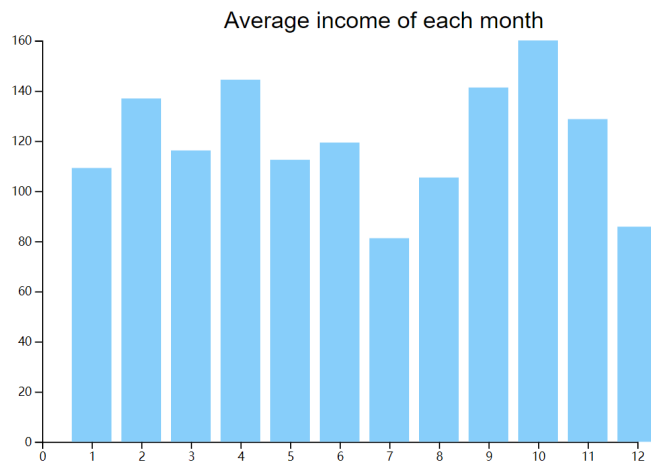
并且为更清晰看出某一指标随时间的变化关系，在百分比堆积柱状图的基础上，使用折线图进行展示；

- 分析季节因素对酒店的影响。季节对于酒店的影响是多方面的，本项目将从季节对淡/旺季的影响，对酒店价格，对大型活动的举办，和对酒店入住人员特点的影响多方面进行分析。

## 2.数据分析结果

### 2.1 酒店淡/旺季分析

综合使用表格中提供的入住率Occupancy和价格Price指标，相乘获得Average Income，使用柱状图进行表示。

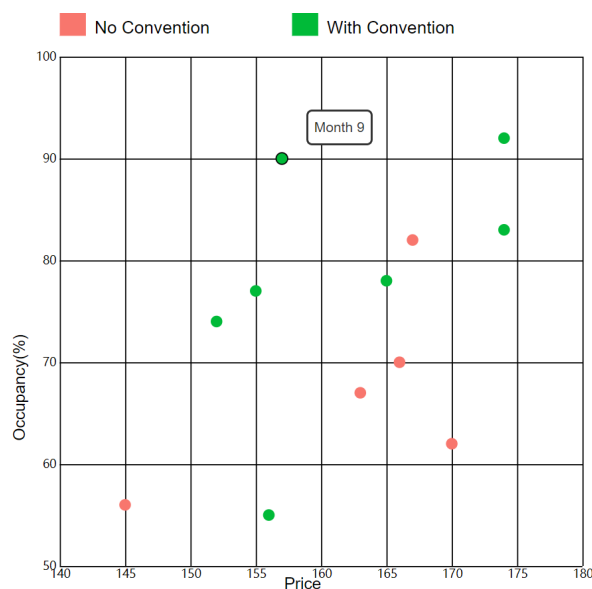


根据柱状图中各月的相对大小，可以粗略判断淡季与旺季的分布：

- 淡季：12-1月（冬天）、7-8月（夏天）；
- 旺季：2-6月（春天），9-11月（秋天）；

接下来分考查价格Price，是否举办大型活动Conventions与入住率Occupancy的关系。如下图所示，使用散点图进行表示，横轴表示价格，纵轴为入住率以反映生意好坏，可以看出：

- 大型活动的举办与入住率存在显著的正相关关系，在举办大型活动的月份，除去12月入住率较低，其余月份总体来说生意好于无大型活动的月份；
- 价格与入住率没有明显的相关性，在一些月份（4月和10月），尽管酒店的价格相对较高，生意依然较好，原因可能是当地人流量大，酒店为提高营收而提高价格；而在另一些月份（9月），较低的价格与较高的入住率相关联，原因可能是当地酒店竞争相对激烈，较低的价格能够吸引更多的住户。价格的选择是商家根据客流量等信息综合考量的结果，又反过来影响客流量。

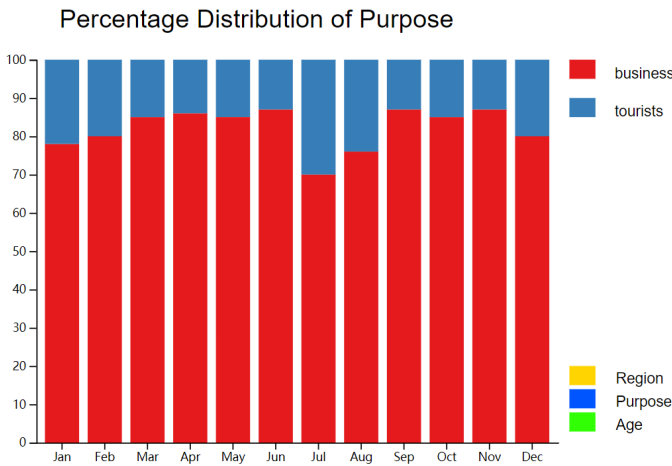


## 2.2 酒店入住人员特点分析

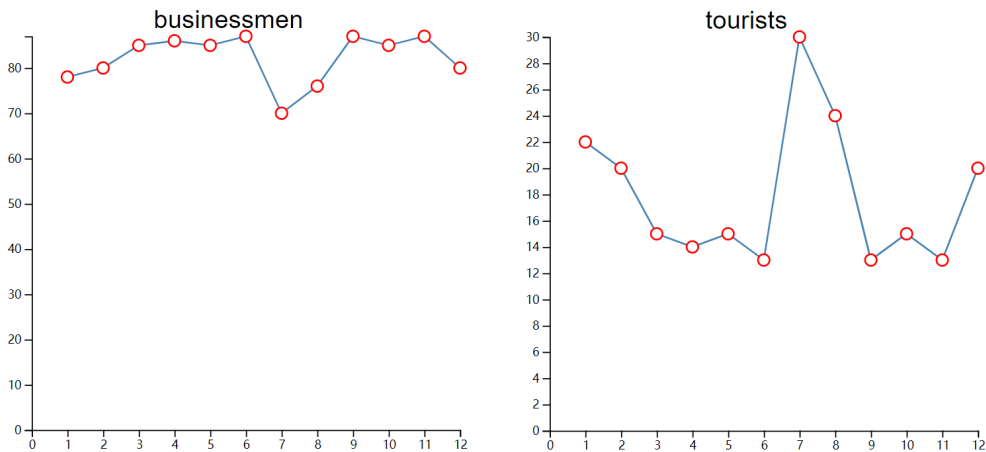
在入住人员特点的分析上，本项目主要从地域分布、年龄分布、住户目的三方面进行分析。使用百分比堆积柱状图作为总览，使用折线图展示某一指标随时间的变化关系。其中表项具有联动关系，点击百分比柱状图中的任意块，显示对应指标的折线图，在将鼠标移动到图表的信息块上时，均有悬浮窗显示详细信息（详见Demo）。

### 2.2.1 住户目的分布

从住户目的上看，一年中商业目的的比例维持在70%-90%之间，其中在相对的旺季商业比例更高，在相对淡季旅游比例更高。可以推测酒店附近没有名胜景点，但是地理位置较好，距离商圈较近。

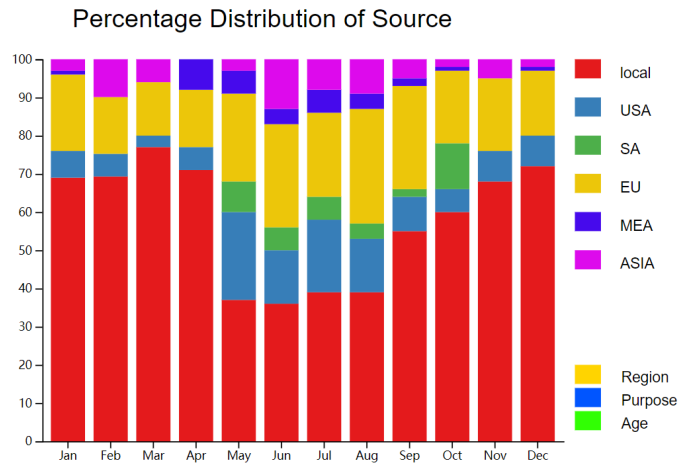


更进一步看，旅游目的的比例在7-8月达到高峰，结合地域分布看，此时处于暑假，外国游客显著增加，目的以旅游为主，因此住户目的特征主要受到季节的影响。

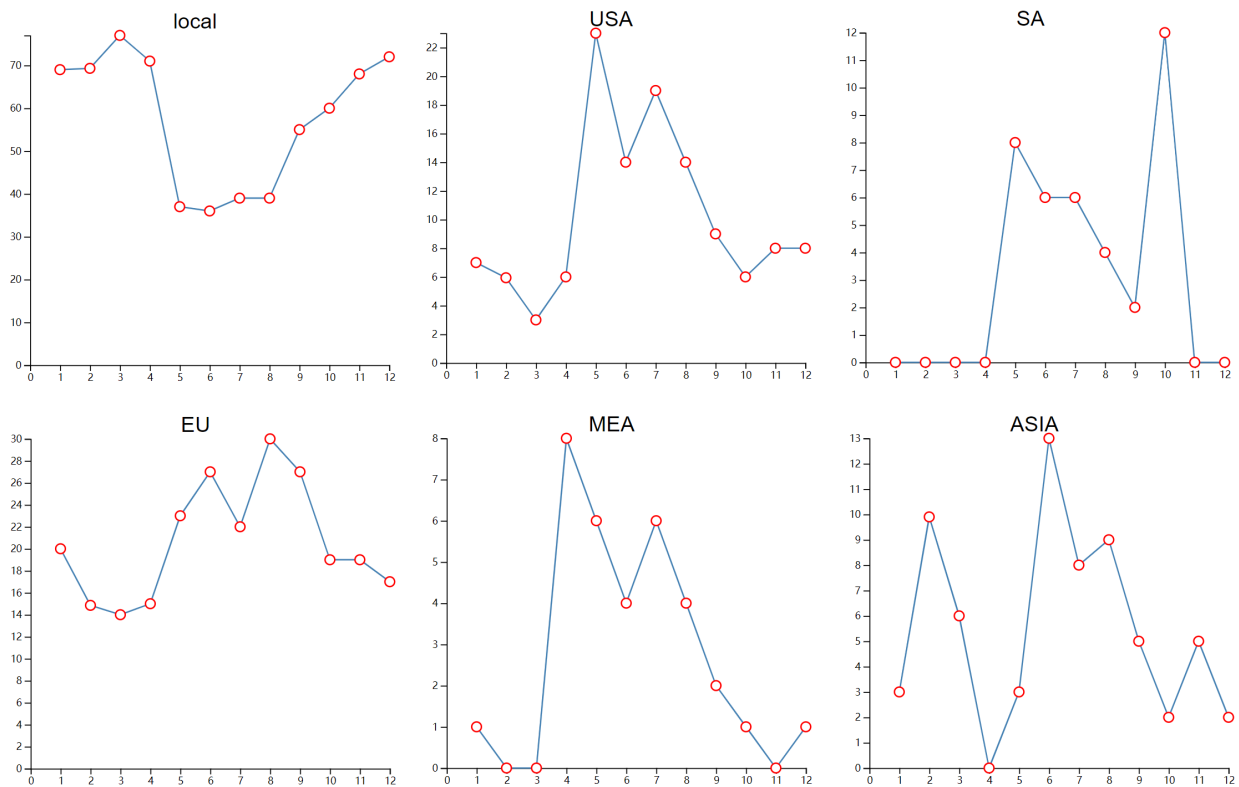


### 2.2.2 住户地域分布

从住户来源上看，本地居民基本占据大部分，其次是欧洲与美国，而来自中东与非洲、中美及南美、亚洲的游客相对较少，因此可以推测酒店的位置位于欧洲或美国。

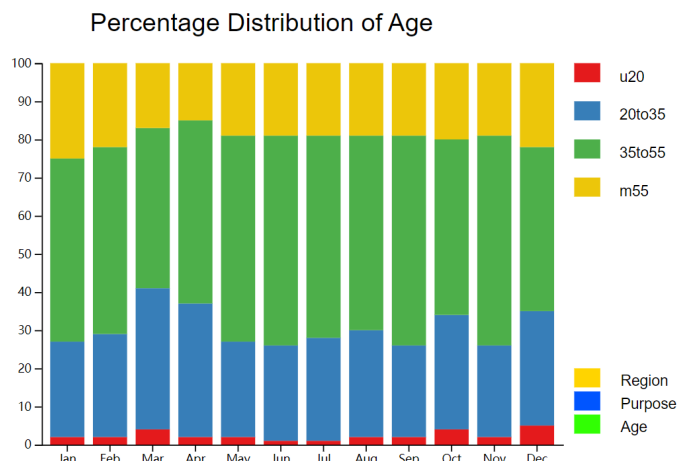


接下来展示各指标随时间的变化，以本地住户为例。从折线图中可以看出，本地住户更倾向于在春天、秋天、冬天入住，而其他地区的住户更倾向于在春末、夏天入住。推测入住目的，本地住户多为商业目的，而其他地区住户多为旅游目的，住户来源的特征主要受到季节的影响。但结合淡季与旺季的分布，可以看出其他地区游客比例增长的时间段，正处在当地的淡季，因此可以看出客流量很依赖于本地住户，淡季产生的原因可能来自于暑期休假导致的商业活动减少。

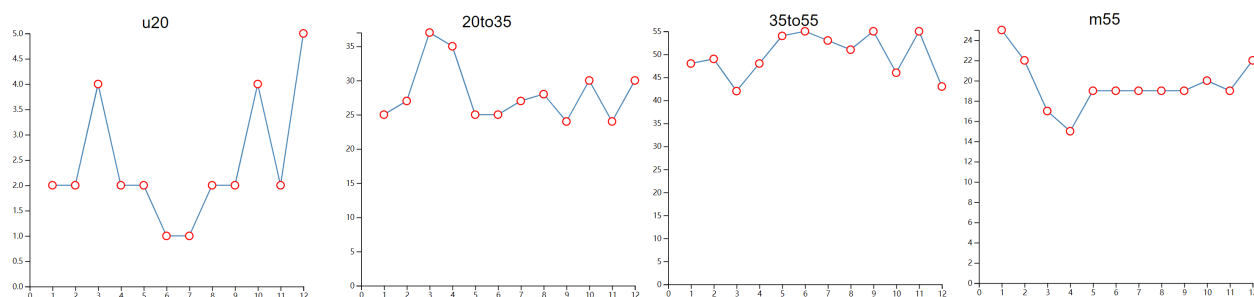


### 2.2.3 住户年龄分布

从住户年龄的分布上看，约一半的住户处于中年(35-55)，青年(20-35)和老年(>55)各占约1/4，儿童(<20)的占比很少，因此酒店较为吸引中年人群体，结合前文中提到的住户以商业目的为主，可以推测酒店住户的主体为商务人士。



从各年龄段住户的入住时间看，青年住户(20-35)更倾向于在3月和4月入住，老年住户(>55)更倾向于在 入住，中年住户(35-55)的入住比例除去3月与10月较低，在其他月份相对稳定。但是由于10月的入住率较高，因此就入住人数来说，中年人的数量并没有明显的变化。结合大型活动举办等信息，推测4月可能举办了对青年住户吸引力更大的非商业活动。综合来说，中年人的客流量相对稳定，因此住户年龄分布特征主要受到非商业因素的影响，例如大型活动的举办等。



## 2.3 季节对酒店的影响

- 季节对淡/旺季的影响：淡季为夏季与冬季，旺季为春季与秋季；
- 季节对酒店房价的影响：酒店在淡季会适当降低房价，在旺季适当提高房价；
- 季节对大型活动的影响：春天与秋天时当地更倾向于举办大型活动；
- 季节对住户特征的影响：
  - 目的：夏季住户中以旅游为目的的住户比例相对提高；
  - 地域：夏季来自于其他地区的住户比例相对提高；
  - 年龄：春季青年住户的比例相对提高；
- 具有较相似季节性变化规律的因素：
  - 住户目的 & 住户地域
  - 入住率/平均收入 & 是否举办大型活动

## 2.4 其他结论

1. 从住户的平均入住时间上看，一年内都维持在1-2天的水平，代表几乎都是1-2天的时间。对于游客来说，代表酒店周边景点少，可以在短时间内遍历；对于商务人士来说，代表商务活动的时间很短，规模不大。

2. 从住户的预定方式上看，现场预订(dr)的比例占据大部分，其次是代理人预定(agency)，最后是空勤预定(ac)。而在现代酒店的预定方式应当为代理人预定为主，现场预定的比例应当很小，可以推测数据统计年份较为久远，或当地数字化程度低下。

3. 从住户性别上看，女性住户所占比例一直很小，这可能与住户的商业目的为主有关。