

摩拜单车用户画像（上海）

Tableau 故事 1 链接如下：

https://public.tableau.com/profile/candis.li#!/vizhome/MobikeShanghai/1_2

最新版故事链接如下：

<https://public.tableau.com/profile/candis.li#!/vizhome/mobikedataofShanghai-20180830/story>

【总结】

使用 Tableau 完成上海摩拜单车用户画像（2016 年 8 月），用户最喜欢在距离低于 3km 时使用摩拜，上下班高峰会有更多的人使用摩拜出行。

【设计】

1. 首先使用 Python 对原数据进行一定清理，并将骑行时长按照 0-10, 10-20, 20-30, 30-45, 45-60, 60-120, >120 分钟进行分类。
2. 添加计算字段“距离”，根据所给出的经纬度计算出起点和终点的距离，并添加计算字段“距离段”，将其按照 0-1km, 1-3km, 3-5km, 5-10km, 10-20km, >20km 进行分类。
3. 使用聚合函数计算“人均订单数（订单数/用户数）”，“人均拥有车辆数（用户数/单车数）”，“车均订单数（订单数/单车数）”的值，并用条形图显示。
4. 使用地图显示骑行（起点-终点）的路径，显示密集程度，用白天和夜晚做筛选，并用不同的颜色表示，会发现白天的订单远远多于夜晚（与常识是相符的），城市中心的订单远远多于城市周边。
5. 使用折线图表示每日订单数量和一天当中的每个小时的变化，此外，在同一张图上，使用不同的颜色表示一周内不同天的订单数量变化，观察周末和工作日的订单随时间的变化情况。观察发现，随着时间的推移，从 1 号到 31 号，整体趋势上，摩拜单车的订单数量稳定上升，1 号仅有 1890 单，31 号已有 5384 单。早上 8 点是早高峰，晚上 6 点是晚高峰，这和人们上班出行时间是一致的。下午 5 点至晚上 8 点的订单数量也维持在一个比较高的水平，猜想这可能是这个时间段不同的下班时间，以及很多人下班会选择骑单车休闲锻炼造成的。周末没有工作日 8 点的早高峰，整体骑行量少于工作日，可以理解，人们会想要在周日有更多的休息，以投入到新一周的工作当中。
6. 使用条形图+百分比的形式表示距离段和骑行时长的分布情况，并降序排列。会发现，目的地距离在 1-3km 的骑行数量最多，占到 47%，少于 1km 的骑行数量紧随其后，占到 42%，二者相加，发现目的地距离在 3km 以下的订单占到总订单数量的 90%。与此相应的是骑行时长，0-10 分钟的订单最多，占到 45%，骑行时长在 0-30 分钟的订单占到近 90%。对产生这种情况的猜想如下：如果选择单车，当目的地距离大于 3km，会需要一定的体力，并耗费一定的时间，这种情况下，人们可能更乐意选择其他的交通

工具，比如打车、公交、地铁等。

7. 使用条形图分组表示用户骑行数量的变化，降序排列，不同的距离段用不同的颜色表示，并使用距离段作为筛选器，观察用户骑行的距离段的特点。可以发现经常骑行的人基本上在白天（7 a.m.-7 p.m.）和夜晚（7 p.m.-7 a.m.）均有骑行订单，在订单的目的地距离上和整体趋势是一致的，1-3km 最多，其次是 0-1km。

【反馈】

1. 之前一版的故事 1 的条形图和故事 2 的饼图所述内容是过于重叠，完全可以通过条形图+百分比达到饼图的效果。
2. 细节上说，如果饼图%字体的颜色改成统一的黑色，是否会更利于阅读？采纳建议，已做修改。
3. 地图上的颜色过多且过于复杂了（一般而言，整个作品使用的颜色<5，考虑使用集/组或把颜色放在其他维度上），即使分别点击不同时段，我也很难比较出各个小时出行轨迹的区别。如果高峰和非高峰期有区别，请更突出地展示出来。如果天与天之间没有明显的区别，为什么要用不同的颜色标识周一-周日呢？不如将颜色用在别的字段上（譬如白天/黑夜，或者高峰期/非高峰期）。
4. 建议丰富内容。
5. 骑行时长分布和一周的不同天都选择了用颜色区分，建议使用同一色系渐变区分骑行时长分类，彩虹色系区分不同天数，而不应该都用彩虹色系。

此版本已根据审阅意见和他人建议，对 1/2/3/4 点进行了修改。

第五条，考虑到这两张视图是出现在不同的故事点，综合视觉呈现效果，使用色盲系颜色。

【资源】

N/A。