

# 2016年8月上海市摩拜数据分析

## 文件

- [初始文件](#)
- [修改文件](#)

## 总结

在本项目中，选用了2016年8月份上海市摩拜单车的用户骑行数据进行可视化及分析，并将分析结果与三位好友进行了分享，获得了他们珍贵的意见，最终得以完成项目。得出结论如下：

- 整个8月，摩拜上海的订单量与用户量都随着日期的向后推移而不断增加。
- 在工作日，6-9点和16-21点为订单量的两个高峰期；但在周末，只有在16点-21点有一个相对高峰期，订单量远少于工作日的；
- 就骑行时间而言，6-9点的最短，16-21的较长；同时段相比，周末的要比工作日的稍长；
- 订单量最多的几个起点分别为：市光路地铁站、仁德路桥和江湾体育场附近；
- 订单量最多的几个终点分别为：市光路地铁站、殷高西路地铁站和江湾镇地铁站附近；
- 作为起点的基本上都是地铁站附近，路径较为发散；而驶向终点的路径则相对较为集中，主要来自附近的地铁站方向。

现将项目设计初衷、反馈及参考资料分享如下。

## 设计

### 变量

在数据集中，包含了订单ID、自行车ID、用户ID、骑行的起始时间和位置坐标、终止时间和终止坐标以及骑行过程中采集到的坐标。

### 提出问题

根据以上变量，提出以下问题：

- 在这一个月內，日订单量和日用户使用量的变化是怎样的？
- 在一天24小时内，订单量的变化是怎样的？这样的变化在工作日和周末有什么区别吗？
- 在一天24小时内，骑行时间的变化是怎样的？这样的变化在工作日和周末有什么区别吗？
- 在这一个月內，出现频次最高的骑行的起始位置有哪些？终止位置呢？在不同时段有什么区别？
- 在这一个月內，出现频次最高的起始位置骑行路径是怎样的？终止位置呢？

### 进行可视化！

- 对于第一个问题，观察的是在一段时间内观察变量的变化，所以选用曲线图，并采用移动平均值计算，让曲线更平滑；
- 对于第二个问题，因为要比较的是各个小时之间订单量的不同，所以采用曲线图，并用颜色的深浅表示订单量的多少，这样能使结果更加直观；至于工作日与周末，做成两个集就可以了，

通过添加筛选条件，观察之间的区别；

- 对于第三个问题，骑行时间的差异可能不是很明显，若用曲线图的话可能不够直观，所以改用柱状图，并且添加具体的时间标示；
- 对于第四个问题，可以用圆点的大小表示订单数量的多少，用颜色区分工作日与周末，再添加一个可筛选的时间，这样就可以观察不同时间节点，工作日和周末的最热起始和终止位置了；
- 对于最后一个问题，需要对原始数据进行处理，先筛选出最热位置，并找出与其相连的坐标点，最后绘制路径图。在路径图中，添加位置的筛选条件，并用颜色区分起始点与终点。

## 分享与反馈

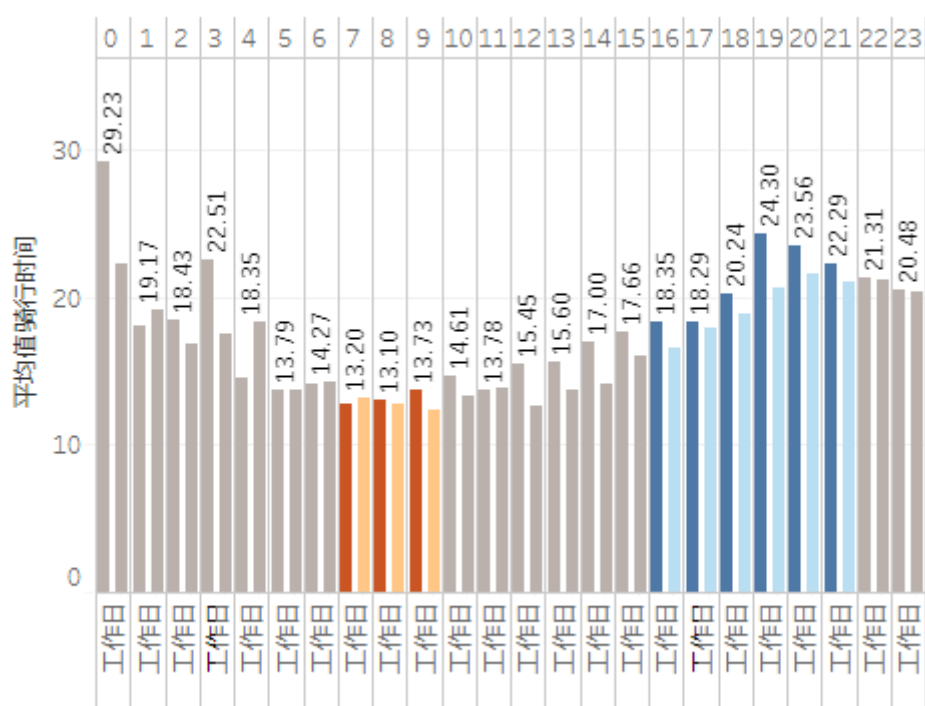
### 被分享人一

- 被分享人角色：uda数据分析进阶同学
- 反馈意见：
  - 故事框中的文字偏多，阅读不是很方便，可以将文字添加到仪表盘中；

早晚高峰与周末图例



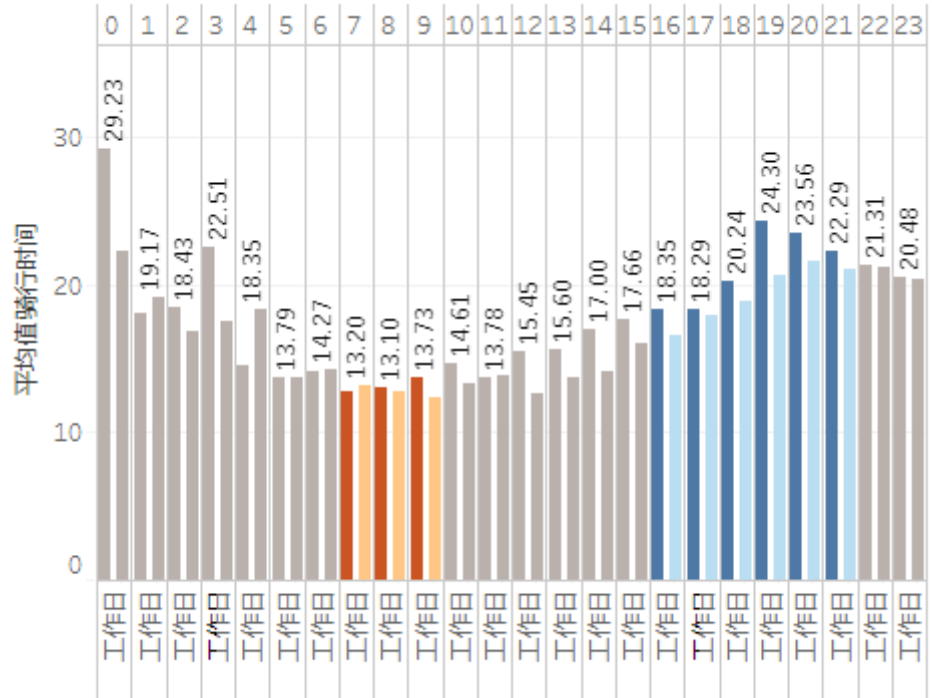
- 图例太长，后面文字都被隐藏了，请进行精简
- 下图中，坐标轴不够清晰，不知道此图想表达的是什么意思



- 在最热起终点路径图中，多了一个毫无意义的station。

## 被分享人二

- 被分享人角色：商务英语专业大四生
- 反馈意见：



- 看不懂该图
- 在最热路径图中，能否突出标记终点或起点位置？

## 被分享人三

- 被分享人角色：C++工程师
- 反馈意见：
  - 第一个图中，有两点疑惑：1.单位（虽然文字描述中采用了'个'，但还是觉得在图例或者轴标签也附上更好）；2.图形上方的start\_time貌似没什么用？
  - 第二个图中：1.筛选条件标题为“是否为工作日”，但下面的选项是“内/外”，不是很理解；2.不知道orderid计数的具体含义；3.因为图表中涉及到了平均值，所以平均值具体是多少，最好也能给出标注；4，下面的图混乱不堪，不知所云
  - 第三个图中：同样工作日内外问题；筛选能不能设置成滚动条？
  - 第五个图，不知所云

## 反馈后修改

### 修改

- 将每页故事框中文字提炼成十字左右的标题，把解释性文字放在可视化图中；
- 第一页：添加单位个，删除无关项；
- 第二页：转换描述词汇；标出平均值大小；优化骑行时间条形图（将冗余的横坐标去掉，添加颜色图例以区分周末和工作日）

- 第三页/第四页：转换描述词汇；将时间筛选改为列表形式
- 第五页：删除station行数据

## 反思

- 图中涉及到的变量名称，应该在图例以及坐标轴中转换为生活中的描述，而不是数据建模中的描述；
- 可视化以及图例一定要本着“少即是多”的原则，必要时添加文字解释，最好能让客户一眼知道这幅可视化要解决的问题是什么；
- 后续要尝试在地图上添加最新的包含地名、路名、地铁站名的图层，这样最后三页中有关位置的可视化结果就能一目了然了。

## 审阅

### 审阅意见

- 发布到Tableau Public前可以先将不需要展示的工作表、面板隐藏，这样作品更简洁美观；
- 将得出的结论进行提炼，放在故事点的文字框。

## 资源

- [在 Tableau 中创建显示起点和终点之间的路径的地图](#)
- [在 Tableau 中创建集](#)