Titanic数据集可视化报告

**一、故事链接**

第一个版本：[https://public.tableau.com/profile/ivan.wu#!/vizhome/titanic\_108/1\_1](https://public.tableau.com/profile/ivan.wu" \l "!/vizhome/titanic_108/1_1)

最终版本：[https://public.tableau.com/profile/ivan.wu#!/vizhome/titanicfinal\_1/sheet0](https://public.tableau.com/profile/ivan.wu" \l "!/vizhome/titanicfinal_1/sheet0)

**二、总结**

本可视化主要针对泰坦尼克号沉没事件中两性的生还率开展研究，先从总体观察，接着分别从年龄段、船舱等级、上船城市和亲人数量等角度分别观察，可以看出：大于10岁的女性生还率比男性高得多；从一等舱到三等舱，生还率逐渐下降，但都是女性高于男性，即使三等舱的女性生还率也高于一等舱的男性。同时，是否生还受船舱位置、亲人属性、人群性格等多种综合因素影响。但总体来看，在这次灾难中，人们把生还的机会更多地让给了女性和儿童。

**三、设计**

该可视化通过故事《女性生还率高于男性》表述：

（在所有元素中，蓝色表示女性，黄色表示男性；红色表示死亡，绿色表示生还）

***故事点一：总体来看，女性的生还率比男性高很多。***

上图：用饼图展示两性生存和死亡的人数；

下图：用柱状图展示两性的生还率对比。这里，生还率用维度“Survived”的平均值计算，下同。

***故事点二：在不同的年龄段中，女性的生还率均高于男性。（70岁以上没有女性乘客）10岁以下的儿童中，男女性存活率差别最小。***

上图：用饼图展示不同年龄段两性人数；

下图：用柱状图表示不同年龄段两性生还率。

***故事点三：随着船舱级别的降低（1到3），整体生还率下降。但在三个船舱中，女性的生还率都高于男性，即使三等舱的女性生还率也高于一等舱的男性。***

上图：用热图表示不同船舱人数；

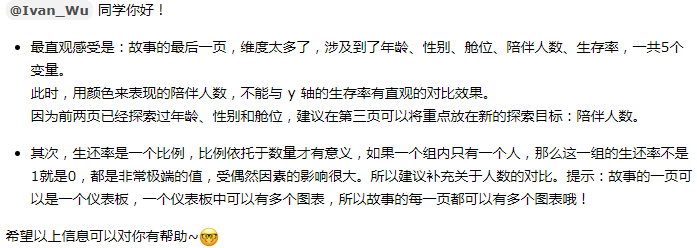
下图：用柱状图表示不同船舱两性生还率。

***故事点四：始发站南安普顿上船的女性旅客似乎更喜欢结伴旅行，但生还率最低，进一步通过筛选年龄段发现，这部分人群只有在30到50岁之间生还率排名第二，排除有些年龄段样本太少造成的误差，生存率应该受到不同城市人群性格、船舱具体位置、亲人属性等因素综合影响。***

上图：用柱状图表示不同登船口女性生还率；

下图：用柱状图表示不同登船口生还和遇难女性平均亲人数量。这里，亲人数量为计算字段，计算方法为：[SibSp] + [Parch]。同时该图表加入筛选器：年龄段，用以观察不同年龄段人群行为。

**四、反馈**

[优达学城论坛关于我征求意见的帖子链接](http://discussions.youdaxue.com/t/tableau/64318)

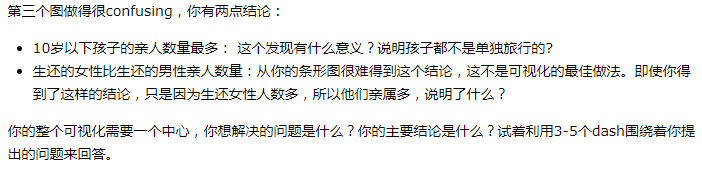
**（一）来自mentor ivylee的反馈一：**

将意见整理后如下：

1. 第三页重点放在亲人人数变量，减少图表中的变量数。

2. 在生还率旁增加人数的图表，以突出偶然因素的影响。

3. 可以在故事的每一页添加仪表盘。

**（二）来自mentor Elib的反馈二：**

将意见整理后如下：

1. 第三页关于10岁以下孩子亲人数表达的意义没有凸显出来。

2. 第三页关于生还的女性比生还的男性亲人数量多，这一点很难看出来。

3. 试着利用仪表盘组建故事。

（三）**结合上述两个反馈，做出如下完善和更改：**

1. 将年龄段和船舱相关的两个可视化改为仪表盘，均加入基础数量图表。

2. 第三页改为对不同登船城市女性生还率和亲人数量研究。

3. 加入总体情况仪表盘。

**五、资源**

N / A.