

DAP4: 创建一个 Tableau 故事

Tableau Public 工作簿的链接

Tableau Public 工作簿的链接：

v2 版本： <https://public.tableau.com/profile/qingzi#!/vizhome/titanic-data-v2/titanic?publish=yes>

v1 版本： https://public.tableau.com/shared/NN9M234X4?:display_count=yes

～ 背景：泰坦尼克号上2224名乘客及船员的数据子集，包括了人口统计资料和乘客信息。点击查看更多信息[Titanic: Machine Learning from Disaster | Kaggle](https://www.kaggle.com/c/titanic)。
～ 目标：创建数据可视化，展现幸存者和遇难者的人口统计资料或乘客信息。

总结

不超过四句话，简要介绍你的数据可视化，并添加可帮助读者理解的上下文信息

泰坦尼克号遇难者人数过半，幸存者与性别、年龄、舱位等有较大关系。女性幸存率高于男性，在三种舱位都呈现出该规律；小孩的幸存率高，其次是 30.5-40 岁的年轻人；坐在头等舱的人幸存率高，三等舱的人幸存率低。无论是男性还是女性，是否携带亲属对幸存率无明显影响。

设计

解释你所做的任何设计选择，包括收集反馈后对可视化进行的更改

故事分为五篇：

- 01 幸存者与性别的关系
 - 饼状图：幸存者和遇难者的整体比例
 - 饼状图：幸存者和遇难者的不同性别所占百分比
 - 筛选器：Sex
- 02 幸存者与年龄的关系
 - 柱形图：幸存者和遇难者的不同年龄所占百分比（每 10 年划分一个年龄段）
 - 柱形图：不同年龄段的幸存率和遇难率比较
 - 筛选器：Pclass、Survived
- 03 幸存者与身份地位的关系
 - 条形柱状图：幸存者和遇难者在各舱位级别的分布

- 堆叠柱状图：不同舱位级别的幸存率比较
 - 箱线图：舱位级别和票价的关系
 - 筛选器：舱位级别（Pclass）、Survived，浮动的下拉菜单
 - 突出显示操作，运行操作方式为悬停
- 04 舱位、性别与幸存率的关系
 - 柱状图：不同舱位和性别的幸存及遇难人数比较
 - 堆叠柱状图：不同舱位和性别的幸存及遇难人数比较
 - 突出显示操作，运行操作方式为悬停
- 05 性别、有无携带亲属与幸存率的关系
 - 表格：不同性别在有无携带亲属状况下的幸存人数比较
 - 堆叠柱状图：不同性别在有无携带亲属状况下的幸存率比较
 - 突出显示操作，运行操作方式为悬停

反馈

包含从第一份草图到最终可视化期间，你从他人那里获得的针对你的可视化的所有反馈

第一次提交

01 在“幸存者和身份地位关系”分析中，加入乘客在各舱位级别分布图，方便显示不同舱位级别的人数。

已增加该条形图

02 在“幸存者和性别关系”分析中，仅有两个图例，空白区域有点大，不够美观。

增加了一张图片，填充空白区域不足。

第二次提交

03 既然你呈现的最终结果为一个故事，其他部分已经包含在故事中，那么可以将其隐藏，使得工作簿更加简洁。

隐藏除故事以外的工作表、仪表盘

04 在故事中，你的文字叙述部分较多，而文字框内容较小，影响了可视化效果。请对文字框的边框适当拖拽调整尺寸，使得可视化效果更好。

调整了文字框尺寸

05 你的仪表板以及故事的尺寸较小，导致可视化内容比较拥挤，请在故事和各仪表板的Size中选择Automatic以解决此问题（对于故事及所有故事点中的仪表板，都应当选择此选项）。选择“Automatic”还有一个重要理由，即每个人所使用的设备分辨率不尽相同，如果使用固定尺寸，会导致可视化在某些设备上无法合理展示。因此请对此部分进行适当调整。

将故事和仪表盘的大小，由固定大小改为自动。

06 故事点1的图片没有意义，无法为分析提供帮助，从布局上来看也并不合适，建议删除。针对图表尺寸过小的情况，建议合理增加尺寸，或增加更多可视化解决；

删除了故事点 1 的图片

07 故事点2，下方散点图没有太大意义，难以直观看出生还情况。建议直接通过折线图或条形图展示分组别的生还率信息或改为其他更有意义的可视化，而非散点图；

删除了散点图，增加了分组别生还率信息的条形图。

08 故事点4和5，没有看出意义，从业务分析的角度而言没有价值。观众不会具体关注某一个遇难者的姓名或是寻找一个特定的人（如果有需求，也并不应当由可视化完成），可视化的重点应当放在对整体数据的分析上而不是展示所有行数据的具体信息或寻找特定的某行数据上。请对故事结构进行调整。

删除了故事 4 和故事 5，重新调整了故事结构

09 目前，针对生还率的分析仅仅停留在单变量的因素分析，而未结合至少2个变量综合分析，分析角度较为缺乏。请补充多变量因素分析图表（例如，各舱位分性别的生还率、各年龄段人群在有无携带亲属状况下的生还率比较等），充实故事的内容。

补充了多变量因素分析图表。各舱位分性别的生还率、不同性别在有无携带亲属状况下的生还率比较

10 读者对图形的总结将与文件中的书面摘要紧密匹配，或者读者可以识别图形尝试传达的至少一个主要论点或关系。

调整故事结构后，重新丰富了书面摘要。

11 请注意，你在此部分还存在一些问题。请确保你的每幅图表（请注意是图表，而非仪表板）都有：合适的标题（以规范的中文/英文表述概括可视化所展示的信息）；坐标轴（显示数值）；经过清理的坐标轴标签（显示为清晰易懂的中文或英文名称；如果标题已经明显提示了变量，则可以不加）。你有部分图表至少缺失了这三者之一（部分图表缺乏合适标题），请仔细检查并修改。

检查并调整了每幅图表的三要素：标题、坐标轴、坐标轴标签。尤其是图表标题。

12 请注意，你在部分图表中使用了红绿（或近似红绿）调色板（红色和绿色作为主要部分出现在同一幅图表中），对红绿色盲不友好。红绿色盲在人群中占比较高（占全球男性人口约8%，女性人口约0.5%），因此为请修改可视化中的颜色表示，使其对色盲更加友好。

修改了调色板颜色，使用色盲调色板，尤其是其中的蓝色和橙色，对色盲更加友好。

资源

列出你创建可视化时参考的任何来源

- [Save Workbooks to Tableau Public - Tableau](#)
- [Titanic: Machine Learning from Disaster | Kaggle](#)
- [如何构建视图 URL - Tableau](#)
- [Dashboard Interactivity Using Actions](#)

CHANGELOG

- 190414 修订 2.0
- 190329 晴子创建 1.0