

# 清华大学

# 数据可视化: 作业一

清华大学 可视媒体研究中心

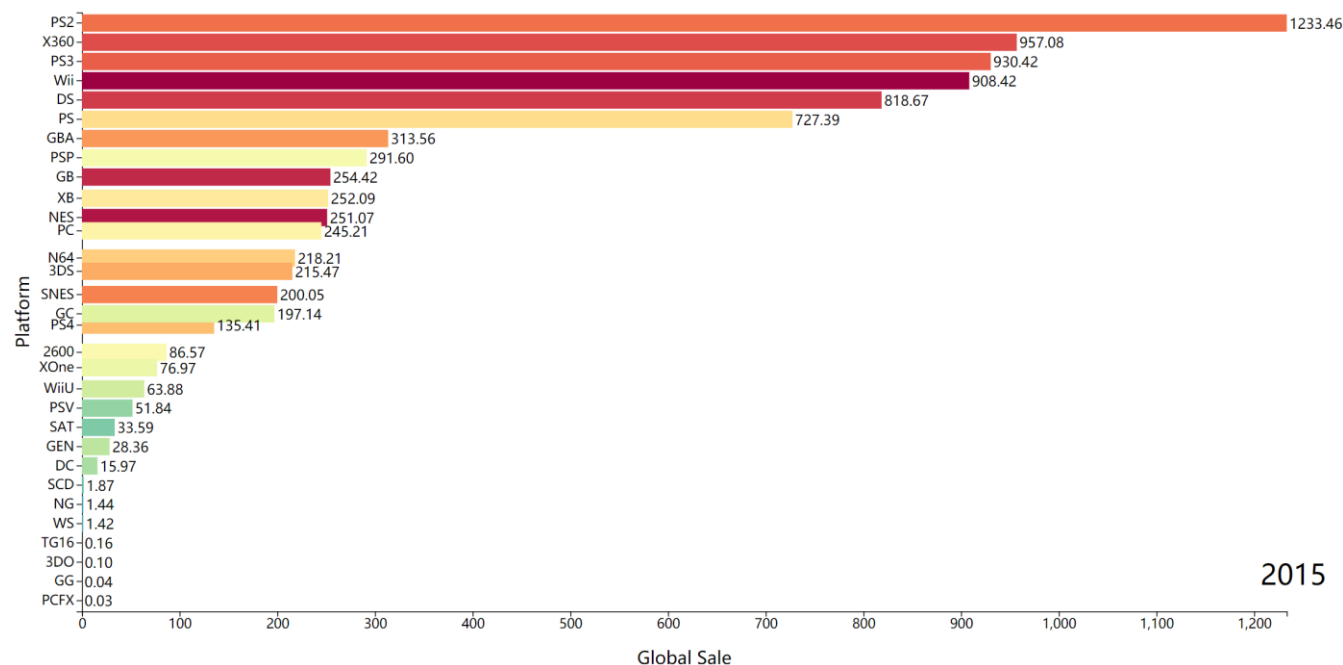
张松海 张少魁 周文洋

2021年3月18日

请在后文的任务中二选一完成: 编程 OR 文献阅读

# 任务1：电玩行业数据可视化。

- 目标：用柱状图的形式展示出近40年来，电玩行业各个平台的全球销售量。
  - 各个平台：如PSP、PS4、GBA、X360等。
  - 全球销售量：以百万为单位，精度不限。



2015

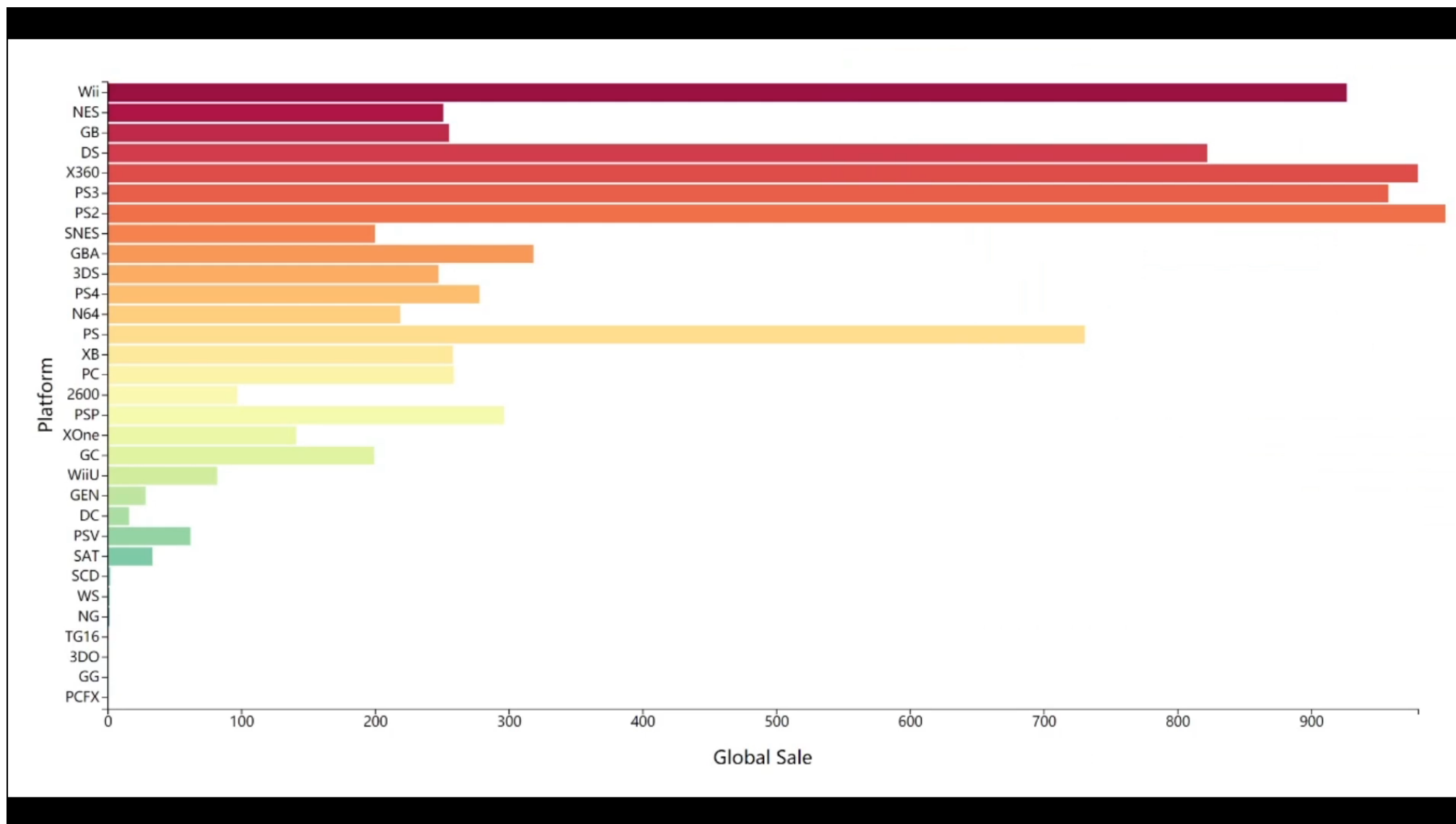
# 数据来源。

- Kaggle数据集：
  - <https://www.kaggle.com/gregorut/videogamesales>
- 本次小作业只需要用到Platform、Global\_Sales、Years。
- 原本数据已经过清洗与预处理，提供给大家。
- 参考：platform\_globalsale.csv、pgy.csv
- 最原始的数据：vgsales.csv，如果倾向于用自己习惯的方式转换数据。
- 公平起见，小作业一统一使用提供的数据。

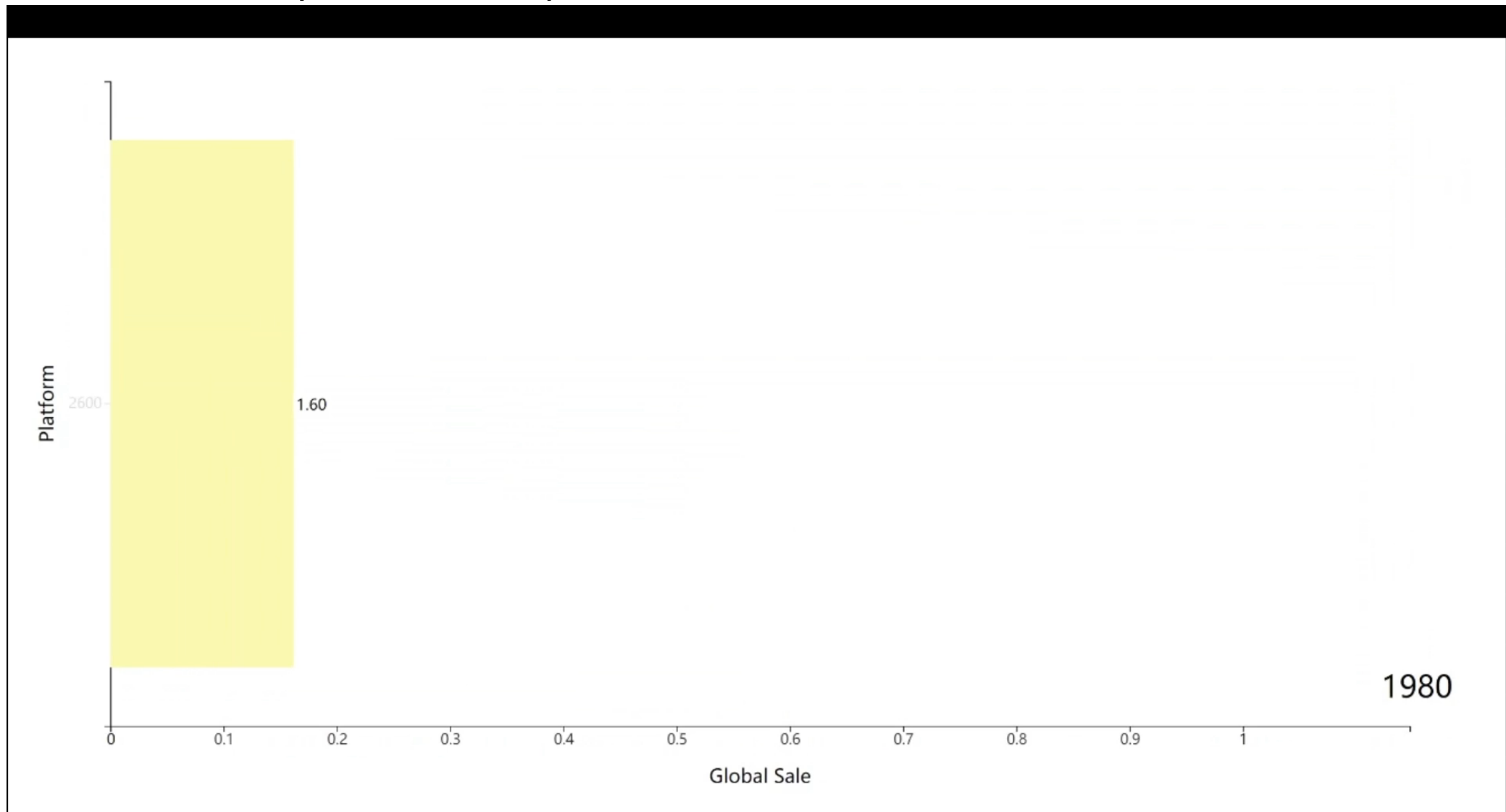
# 任务与给分。

- 用红色标出的内容会主观给分。（主观打分不会过分严格）
- 完成基础的柱状图。（50分）
- 美观。（15分）
  - 整体的美观性。
  - 选取合适、美观的颜色，为图上色。
  - 附文字说明你的设计。（可选）
- 坐标轴与字体的合适尺寸与大小、坐标轴标题。（10分）
- 柱状图开启时的动态效果。（10分）
- 根据年份的Racing与‘动态文本’。（20分）
- 评分严格按照遵照上述条目，请勿‘攀比’编程。

示例。(80分+)



示例。(95分+)

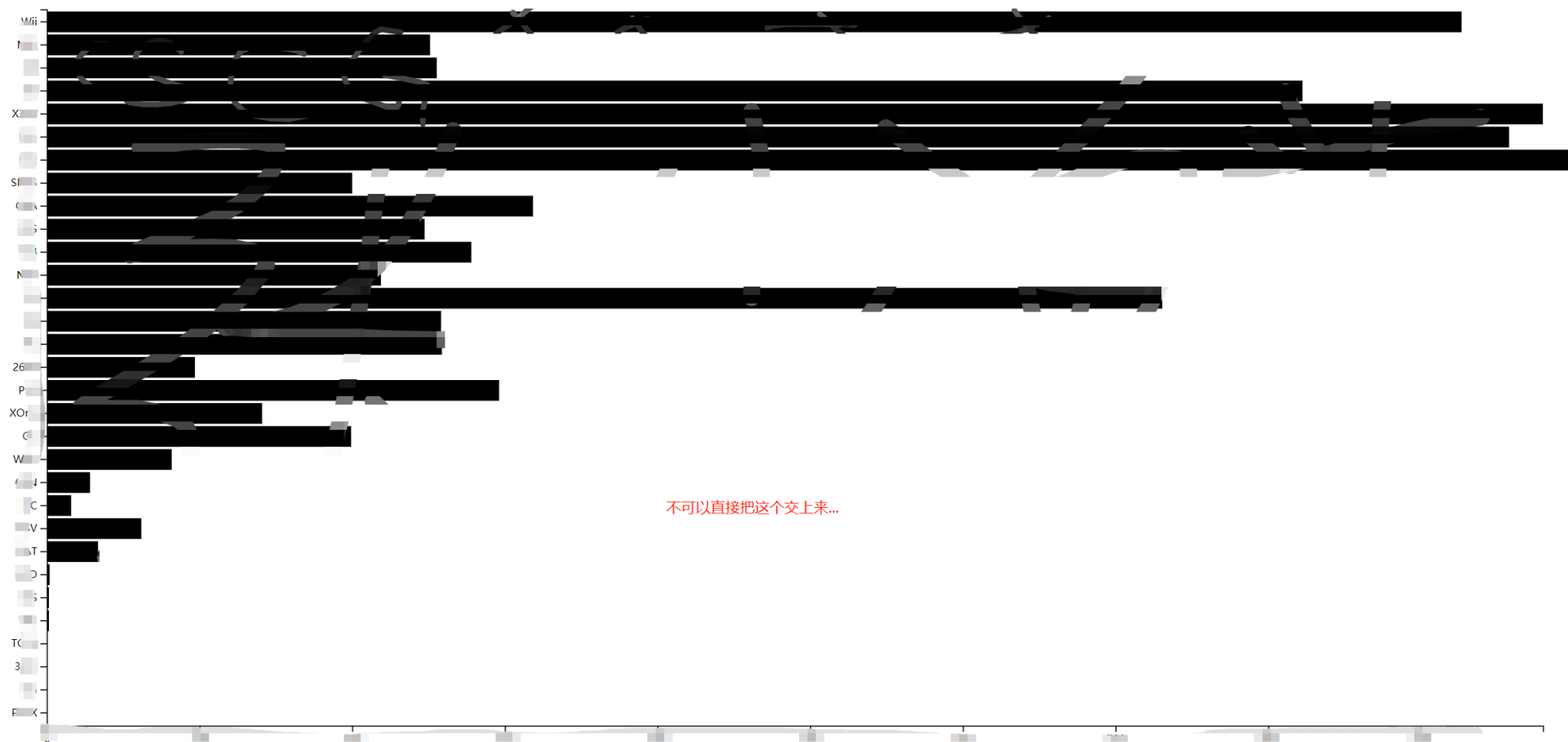


# 需要提交什么？

- 单人完成、单人提交。
- [源代码] + [可视化效果的截图OR录屏] + [报告]。
- 报告：
  - 长度不限，附一个.doc, .pdf or .txt 等文件即可。
  - 主要包含：
    - 前后端分别的框架说明，如前端是D3.js、后端是Node.js Simple HTTP Server。如前端是Echarts、后端是VSCode Live Server。
    - 设计说明。（可选）
  - 若最终代码有BUG导致其无法正常运行，则提交实验报告说明做了哪写尝试，可根据过程给分。
- 所有需要提交的内容，以压缩包的形式上交到网络学堂。

# 技术支持与说明 (1) 。

- 若使用D3.js编程， 如下脚本的源代码共享给大家：
  - 参考main.html





# 技术支持与说明（2）。

- 不限制实现方式，但一些库禁止使用！
  - 本次作业包括：matplotlib、seaborn。
  - 所有动态效果要求真动态，不可以使用静态的结果拼出视频、git等。
- 关于Racing与‘动态文本’需要参考的技术，以D3.js为例：
  - Data-Join。
  - Transition的自定义。
  - JavaScript异步编程。
- 任何技术问题都可以找助教讨论。

## 任务2：文献阅读题

- 阅读论文《VisCode: Embedding Information in Visualization Images using Encoder-Decoder Network》并完成如下任务：
  - 用自己的语言概括论文，要求逻辑清晰、内容充实，涉及算法介绍的要包含公式，每一小节都不得遗漏（如摘要、介绍、相关工作等…），不得直接谷歌翻译，总字数不得少于3000字。（50分）
  - 你认为本文提到了诸多系统实现细节中有哪些不合理或者值得改进的地方？要求字数不少于300字。（10分）你能想到什么方法进行改进？请尽可能详细地描述算法细节，要求字数不少于700字。（10分）
  - 本文将可视化图表进行二次加工，对其进行美化、交互等操作，你能想到其他什么有趣、有价值的想法？要求字数不少于300字。（10分）如果要实现你的想法，请给出一套实现方案，画出整个方案的架构图并进行每个模块的详细介绍，要求字数不少于700字。（20分）
- 提交：
  - 报告即可，格式不限，如.pdf, .txt, .doc等。
  - 字数长度参考题目。

# 迟交策略...

- 如有**正当、很正当且极其正当**的理由，请先联系助教。
- 否则：
- 晚交一天，单次作业成绩扣除1分，最高扣除10分。
- 如延迟到期末提交，单次作业扣除20分。
- 注：单次作业均为百分制，总成绩会在最终上分时核算。
- 迟交会为助教整理作业造成大量负担，同时也会设置‘宽裕’的dealline，故请大家合理安排时间，尽快动手。
- 编程题预估作业所需时间：
  - 完全无编程经验 & 80+成绩：20小时（包含上手时间）。
  - 有很强的可视化开发经验 & 80+成绩：2小时。
  - 有很强的可视化开发经验 & 95+成绩：10小时。

# 注意事项！

- 单人完成。
- 可以使用任何网上的现有代码，但必须标明出处！
- 有过程分，所有提交的作业会被审查并运行。
- **禁止使用下列库完成任务：**
  - matplotlib
  - seaborn
- 若使用被禁止的库完成任务，最终作业不计入总成绩！
- 助教的联系方式：
  - 张少魁 [zhangsk18@mails.tsinghua.edu.cn](mailto:zhangsk18@mails.tsinghua.edu.cn) 东主楼 9-209
  - 周文洋 [zhouwy19@mails.tsinghua.edu.cn](mailto:zhouwy19@mails.tsinghua.edu.cn) Fit楼3-507
- 有任何问题（技术、非技术均可），请随时与助教沟通！！！！