

COMP4037 Research Methods Spring

Coursework 2: Data Visualization

Number of Credits: 25% of module

Recommended hours: 20-25 hours (for distinction level)

Submission Deadline: 25 April by 16:00

Reassessment Deadline: 21 Aug by 17:00

Late submissions will incur a penalty of 10% per day including weekends and bank holidays.

Submit a **PDF** file on Moodle-a link to a submission form will be provided a few days before the deadline.

Questions: If you have any questions about this coursework, please ask them: 1) in class, 2) after class, or 3) in the Research Methods Teams chat. Individual questions received via email will be redirected towards these other options due to the class size. See:

 [Research Methods Questions Management: Asking and Answering Questions](#)

Page Limit: Your submission is limited to 2 pages. Any graphics you use must be on page 1. The description may extend to page 2. Your supplementary video is limited to 2 minutes.

The **aim** of this coursework is to:

- Help you to build an understanding of the research methods in your area of interest
- Develop and assess your ability to present research in a concise visual form
- Give you practice in the use of data as evidence to support your work

This coursework supports the following **Learning Objectives** of the module:

- To increase understanding of the research process and the application of appropriate research methods to given problems.
- To develop the ability to critically appraise and choose research methods, and justify their application to appropriate research problems.
- To enhance skills to present results as part of implementing a project.

Task Description and Context

Professor C was approached by a team of researchers from the University of Oxford studying the impact that diet and food consumption has on the environment. They have published a fascinating research paper on their research (Scarborough et al,2023), however, there is a major limitation with their work. There is a startling absence of any graphics or visual representations of their research data and findings. They need your help with this challenge. They have been collecting a massive amount of diet and food consumption data from over 55,000 consumers linked together with food production from 38,000 farms across 119 countries..

Each row in the data set from the project contains the following information:

1. Participant number

COMP4037 研究方法 Spring

课程作业 2：数据可视化

学分数：模块的 25%

建议时间：20-25 小时（优秀级别）

提交截止日期：4月 25 日 16:00 之前 **重新评估截止日期：**8月 21 日 17:00 **逾期提交将产生每天 10% 的罚款，包括周末和银行假期。** 在 Moodle 上提交 PDF 文件 - 提交表格的链接将在截止日期前几天提供。

问题：如果您对本课程有任何疑问，请问他们：1) 在课堂上, 2) 课后, 或 3) 在研究方法团队聊天中。由于班级规模, 通过电子邮件收到的个别问题将被重定向到这些其他选项。看:

 **研究方法问题管理：提问和回答问题** 页面限制：您的提交限制为 2 页。您使用的任何图形都必须位于第 1 页上。

描述可以延伸到第 2 页。您的补充视频限制为 2 分钟。

本课程的目标是：

- 帮助您了解您感兴趣的领域的研究方法
- 培养和评估您以简洁的视觉形式展示研究的能力
- 练习使用数据作为证据来支持您的工作

本课程支持本模块的以下学习目标：

- 增加对研究过程的理解以及将适当的研究方法应用于给定的问题。
- 培养批判性地评估和选择研究方法的能力，并证明它们应用于适当的研究问题。
- 提高在实施项目过程中呈现结果的技能。

任务描述和上下文

牛津大学的一组研究人员找到了 C 教授，研究饮食和食物消费对环境的影响。他们发表了一篇关于他们研究的引人入胜的研究论文 (Scarborough et al, 2023)，但是，他们的工作存在重大局限性。令人惊讶的是，他们的研究数据和发现没有任何图形或视觉表示。他们需要您的帮助来应对这一挑战。他们一直在收集超过 55,000 名消费者的饮食和食品消费数据，这些数据与 119 个国家的 38,000 个农场的食品生产有关。

项目数据集中的每一行都包含以下信息：1. 参与者编号

2. Type of Diet
3. Gender
4. Age Group
5. Mean GHG emissions (GreenHouse Gas) measured in kg
6. Mean Agricultural Land Use in square meters
7. Mean Water Scarcity
8. Mean Eutrophication¹ Potential– measured in g of PO₄e, gPO₄e
9. Mean GHG from CH₄ emissions (Methane) from livestock management measured in kg
10. Mean GHG from N₂O (Nitrous Oxide) emissions associated with fertilizer use²
11. Mean Biodiversity Impact–species extinction per day
12. Mean Agricultural Water Usage in cubic meters (1 m³ - 1,000 liters)
13. Mean Acidification Potential

It is a lot of data. In fact, it feels like data chaos. The researchers would like to know if visualization can be used to gain insight into their data.

Understanding the Data: As with real-world projects, acronyms and special terminology are used. Thus, you may need to conduct some research, like Googling various terms, in order to gain an understanding of the data and its associated terminology. *We recommend you start by reading the research paper about the dataset:*

[Vegans, vegetarians, fish-eaters and meat-eaters in the UK show discrepant environmental impacts | Nature Food](#)

Download the Data: The data can be downloaded from:

[Dataset for 'Vegans, vegetarians, fish-eaters and meat-eaters in the UK show discrepant environmental impacts' - ORA - Oxford University Research Archive](#)

Optional Link to Scripts that Generate the Figures

[Supplemental Data for 'Estimating the environmental impacts of 57,000 food products' - ORA - Oxford University Research Archive](#)

Better Than Professor C: Professor C would like to obtain and convey the insight contained in the data visually. The goal is to create images that maximize our understanding of the data.

These are some examples of questions that can be considered in the coursework.

Some obvious factual information, such as:

- Which diet is most beneficial for the environment?
- What is the disparity between the best and worst diets?
- Which age group has the best diet?

¹ What is Eutrophication?

<https://oceanservice.noaa.gov/facts/eutrophication.html>

² The Green House Gas No One Is Talking About

<https://civileats.com/2019/09/19/the-greenhouse-gas-no-ones-talking-about-nitrous-oxide-on-farms-explained/>

2. 饮食类型
3. 性别
4. 年龄组
5. 以千克为单位的平均温室气体排放量 (温室气体)
6. 平均农业用地面积 (平方米)
7. 平均缺水
8. 平均富营养化电位 – 以 g of POe、gPOe 为单位测量
9. 畜牧管理产生的 CH₄ 排放 (甲烷) 的平均温室气体, 以千克为单位
10. 与肥料使用相关的 NO_x (一氧化二氮) 排放的平均温室气体
11. 平均生物多样性影响 - 每天的物种灭绝
12. 平均农业用水量, 单位为立方米 (1 m³ - 1,000 升)
13. 平均酸化潜力

这是大量的数据。事实上, 这感觉就像数据混乱。研究人员想知道是否可以使用可视化来深入了解他们的数据。

了解数据: 与实际项目一样, 使用首字母缩略词和特殊术语。因此, 您可能需要进行一些研究, 例如在谷歌上搜索各种术语, 以便了解数据及其相关术语。我们建议您首先阅读有关数据集的研究论文: 英国的素食主义者、素食者、食鱼者和肉食者表现出不同的环境影响 | 自然美食

下载数据: 数据可以从以下位置下载:

"英国的素食主义者、素食者、食鱼者和肉食者显示差异性环境影响" 数据集 - ORA - 牛津大学研究档案

指向生成图窗的脚本的可选链接

"估计 57,000 种食品的环境影响" 的补充数据 - ORA 牛津大学研究档案

优于 C 教授: C 教授希望以可视化方式获取和传达数据中包含的见解。目标是创建能够最大限度地理解数据的图像。

这些是课程中可以考虑的一些问题示例。一些明显的事实信息, 例如:

- 哪种饮食对环境最有益?
- 最好和最差的饮食有什么区别?
- 哪个年龄段的饮食最好?

¹ 什么是富营养化?

<https://oceanservice.noaa.gov/facts/eutrophication.html>

² 没有人谈论的温室气体

<https://civileats.com/2019/09/19/the-greenhouse-gas-no-ones-talking-about-nitrous-oxide-on-farms-explained/>

- Is there a difference in diet according to gender?

Professor C would also like to find out some less obvious information, such as:

- Are there any interesting patterns or trends in the data?
- Are there any outliers?
- What insight can we gain from multivariate visualizations of the data?
- Can we compare the diets and their environmental impacts across all categories?

Professor C made some attempts to visualize the data using old-fashioned pie charts, bar charts, bubble charts, and line graphs. He was disappointed by the results, which are not very insightful, not visually uninformative, nor aesthetically impressive. *Can you do better than Professor C?*

A Real-World Challenge

Selecting a tool: You are required to select an appropriate visualization tool for depicting the dataset concerned. You may have a look at the visualization resources paper (Liu, et al 2022) for a description of available tools. See the Section called “Free, Off-the-Shelf Software Collections for Data Visualization Practitioners” in the survey paper by Liu et al (2022) in the References. You may also use Python, Java, R, Matlab or some other appropriate programming library.

Process the data: You are allowed and encouraged to further abstract the data, e.g., aggregate data records into groups or add new data or meta-data. In fact, you can process or supplement the data in many ways and you are expected to make changes, e.g., perhaps linking some spreadsheets together or creating new spreadsheets.

Your task is to produce an advanced visual design which can convey some meaningful and hopefully interesting insight into the data. You are required to supply a unique observation about the data. For your final image, fill out a description template as below. See Figure 1 for an example.

Visual Design Description Template:

Provide the following information for the image you create:

- **Image:** The visualization itself as an image
- **Visual Design Type:** The name/type of the visual design
- **Name of Tool:** The tool that was used to generate the image
- **Diet Groups:** The diet groups shown
- **Variables:** the other data attributes shown and *why you chose them*
- **Visual Mappings:** Each of the visual design mappings. Include the data mapping information about **color**, **position** (x,y axes), **shape**, **size**, **hierarchy**, and any other visual mappings.
- **Unique Observation:** Things we can learn from the visualization, e.g, from this image we can see this pattern... **Make sure you describe where and how in the imagery**

- 饮食是否因性别而异?

C 教授还想找出一些不太明显的信息, 例如:

- 数据中是否有任何有趣的模式或趋势?
- 是否有任何异常值?
- 我们可以从数据的多变量可视化中获得什么见解?
- 我们能否比较所有类别的饮食及其对环境的影响?

C 教授尝试使用老式的饼图、条形图、气泡图和折线图来可视化数据。他对结果感到失望, 这些结果不是很洞察力, 不是视觉上没有信息量, 也不是美学上令人印象深刻的。你能比 C 教授做得更好吗?

现实世界的挑战

选择工具: 您需要选择合适的可视化工具来描述相关数据集。您可以查看可视化资源论文 (Liu, et al 2022) 以获取可用工具的描述。请参阅 Liu 等人 (2022) 在参考文献中的调查论文中名为“面向数据可视化从业者的免费现成软件集合”的部分。您也可以使用 Python、Java、R、Matlab 或其他一些合适的编程库。

处理数据: 允许并鼓励您进一步提取数据, 例如, 将数据记录汇总到组中或添加新数据或元数据。事实上, 您可以通过多种方式处理或补充数据, 并且需要进行更改, 例如, 可能将一些电子表格链接在一起或创建新的电子表格。

您的任务是生成高级可视化设计, 该设计可以传达一些有意义且希望有趣的数据见解。您需要提供有关数据的唯一观察值。对于最终图像, 请填写描述模板, 如下所示。有关示例, 请参阅 图 1。

视觉设计描述模板:

为您创建的映像提供以下信息:

- **Image:** 作为图像的可视化效果本身
- **Visual Design Type:** 视觉设计的名称/类型
- **工具名称:** 用于生成图像的工具
- **饮食组:** 显示的饮食组
- **变量:** 显示的其他数据属性以及您选择它们的原因
- **Visual Mappings (视觉映射):** 每个可视化设计映射。包括有关颜色、位置 (x, y 轴) 、形状、大小、层次结构和任何其他视觉映射的数据映射信息。
- **独特观察:** 我们可以从可视化中学到的东西, 例如, 从这张图片中我们可以看到这个模式.....确保描述图像中的位置和方式

your unique observation can be seen. Is it shown in the accompanying video? If so, at what time in the video? Note that outliers are normally very interesting. They indicate that something unusual is happening.

- **Data Preparation:** Any modifications to the original data that had to be performed to generate your beautiful image.
- **(Optional) URL to screen-capture demo** showing any animation or user-interaction (any form of user input like mouse interaction, user-options, keyboard options, filtering, selection, brushing). You can also use the video to emphasize your unique observation and highlight where in the image this can be seen. You can include a link to a YouTube or Vimeo video to demonstrate any interaction or animation (similar to CW1). I recommend **two minutes maximum time**. Due to the large volume of submissions, we will not have time to look at and interact with Tableau public software web pages. Therefore, we require a video.
- **(Optional) URL to source code:** if you use a programming language like R or Python you can provide a link to GitLab or GitHub where your code is stored.

A good observation requires interpretation of the resulting image that you generate. Also, if your visual design features interaction or animation, you may upload a screen-capture demonstration to YouTube and provide the link in the PDF you submit. Note that all user-interface components should be shown in English.

Advanced Visual Designs: You are to do better than Professor C! That means you are required to use more advanced visual layouts than typical line graphs, bar charts, pie charts, and bubble charts. That means the following are not allowed: bar charts, pie charts, line charts, and bubble charts, however, you can use these simple charts to support more advanced visual designs, in other words, as complementary charts to support an interesting observation.

What is Better? Create a visualization that:

- conveys information and knowledge
- enables discovery of: patterns, deviation, hierarchy, relationships and association
- identifies relationships between data attributes
- depicts data at different scales
- separates noise from the signal
- is intuitive to laypersons and easy to learn
- is aesthetically pleasing

Describe the insight that your visualizations provide. What can we learn from your visualization? How is it better than a standard line, pie, or bar chart? You are not required to answer all of the above questions. They merely serve as examples of what you could discuss when providing a unique observation about the data through the lens of your imagery.

你独特的观察是可见的。它是否显示在随附的视频中？如果是这样，在视频中的什么时间？请注意，异常值通常非常有趣。它们表明正在发生一些不寻常的事情。

- 数据准备：为生成精美图像而必须对原始数据进行的任何修改。
 - (可选) 显示任何动画或用户交互（任何形式的用户输入，如鼠标交互、用户选项、键盘选项、筛选、选择、刷动）的屏幕捕获演示的 URL。您还可以使用视频来强调您的独特观察结果，并突出显示图像中可以看到它的位置。您可以包含指向 YouTube 或 Vimeo 视频的链接，以演示任何交互或动画（类似于 CW1）。我建议最长用时 2 分钟。由于提交的内容量很大，我们将没有时间查看 Tableau Public 软件网页并与之交互。
- 因此，我们需要一个视频。
- (可选) 源代码的 URL：如果您使用 R 或 Python 等编程语言，则可以提供指向存储代码的 GitLab 或 GitHub 的链接。

良好的观察需要解释您生成的结果图像。此外，如果您的视觉设计具有交互或动画功能，您可以将屏幕截图演示上传到 YouTube，并在您提交的 PDF 中提供链接。请注意，所有用户界面组件都应以英文显示。

高级视觉设计：你要比 C 教授做得更好！这意味着您需要使用比典型折线图、条形图、饼图和气泡图更高级的可视化布局。这意味着不允许出现以下情况：条形图、饼图、折线图和气泡图，但是，您可以使用这些简单的图表来支持更高级的视觉设计，换句话说，作为支持有趣观察的补充图表。

什么是更好的？创建符合以下条件的可视化效果：

- 传达信息和知识
- 支持发现：模式、偏差、层次结构、关系和关联
- 识别数据属性之间的关系
- 描绘不同比例的数据
- 将噪声与信号分离
- 对外行人来说很直观，易于学习
- 美观

描述您的可视化提供的见解。我们能从您的可视化中学到什么？它比标准折线图、饼图或条形图好在哪里？您不需要回答上述所有问题。它们仅用作示例，说明在通过影像镜头提供有关数据的独特观测时可以讨论的内容。

Examples: Some examples of more advanced visual designs are treemaps, parallel coordinates, matrix charts, radar charts, scatterplot matrices, and digital maps.

Very Common Mistake with Treemaps: A common mistake that students make is using a treemap to show nonhierarchical data. Treemaps are a hierarchical data visualization technique. Without a hierarchy, they may be even less effective than a bar chart. However, you can create a hierarchy from your data to visualize. You may look at the following video tutorial(s) for extra help on this topic:

- ▶ A Short Introduction to Treemaps
- ▶ Treemaps for Complete Beginners
- ▶ Visualization Lecture 5 1. Treemap Construction

You will be penalized by -20 points for this error.

Color: For a color map, you are encouraged to use this one: <https://paletton.com/>

Digital Maps: You are encouraged to explore the use of digital maps in your visualizations. Using digital maps will yield bonus points. For a good resource on finding digital maps, please refer to the survey paper on visualization resources by Liu et al 2022 in the References. You may certainly use the digital, geo-spatial maps from the Tableau tutorials in your submission. However, you will not receive credit for the bar charts and line charts demonstrated in the tutorial. Trying to introduce another dataset that links together diet, ingredients, and location is very difficult. It would require reading the research paper for CW 2 and the research paper from Workshop 5 on data visualization.

Lots of Extra Help

- **Supplementary Questions and Answers:**
 - ▶ Research Methods: Questions and Answers on Courseworks 2 & 3
- **More Questions and Answers on Courseworks 2 & 3**
<https://youtu.be/6DJYgh8B1uE>
- [▶ Research Methods Lecture on Visualization](#)
- **Supplementary Visualization Lectures:** You will see many examples in the supplementary data visualization lectures. If you'd like to get extra help, you may watch the course lectures 1-6 on YouTube: [Data Visualization Classes 2019 - YouTube](#)
- **Hands-On Tutorials:** For extra help, you may also watch either one of the following live, hands-on, video tutorials of how to create a visual design including a geo-spatial map:
 - Visualization in Practice: A Live Introduction and Hands-on Demonstration of Tableau
 - ▶ Tableau for Absolute Beginners: A Live Introduction and Hands-on Demonstration
 - Applied Visualization using Tableau: A Hands-On Tutorial and Demonstration
 - ▶ Applied Visualization using Tableau: A Hands-On Tutorial and Demonstration
- **Help from Visualization Experts:** You may also ask any visualization questions on the Vis Guides Website: <http://VisGuides.org>

示例：更高级的视觉设计的一些示例包括树状图、平行坐标、矩阵图、雷达图、散点图矩阵和数字地图。

树状图的非常常见错误：学生常犯的一个错误是使用树状图来显示非分层数据。树状图是一种分层数据可视化技术。如果没有层次结构，它们的效果可能甚至不如条形图。但是，您可以从数据创建层次结构以进行可视化。您可以查看以下视频教程以获取有关此主题的额外帮助： [树状图简介](#) [适合初学者的完整树状图可视化](#) 第 5 讲 1. 树状图构造 您将因此错误而被罚 -20 分。



颜色：对于颜色图，建议您使用以下颜色图：<https://paletton.com/>

数字地图：我们鼓励您探索在可视化中使用数字地图。使用数字地图将产生奖励积分。有关查找数字地图的良好资源，请参阅 Liu 等人 2022 年在参考文献中关于可视化资源的调查论文。您当然可以在提交的内容中使用 Tableau 教程中的数字地理空间地图。但是，您将不会获得本教程中演示的条形图和折线图的积分。试图引入另一个将饮食、成分和位置联系在一起的数据集是非常困难的。它需要阅读 CW 2 的研究论文和研讨会 5 中关于数据可视化的研究论文。

很多额外的帮助

- 补充问题和答案：研究方法：关于课程作业 2 和 3 的问题和答案
[▶](#)
- 更多问题和答案在 Courseworks 2 & 3 上
<https://youtu.be/6DJYgh8B1uE>
- [▶ 研究方法 可视化讲座](#)
- 补充观想讲座：您将在补充数据可视化讲座。如果您想获得额外帮助，可以观看该课程在 YouTube 上第 1-6 讲：2019 年数据可视化课程 - YouTube
- 动手教程：如需额外帮助，您还可以观看以下任一实时动手视频教程，了解如何创建包括地理空间地图在内的可视化设计：
- 可视化实践：面向绝对初学者的 Tableau Tableau 实时介绍和动手演示：现场介绍和动手演示
[▶](#)
- 使用 Tableau 的应用可视化：动手实践教程和演示使用 Tableau 的应用可视化：动手实践教程和演示
- 可视化专家的帮助：您也可以在 Vis Guides 网站：<http://VisGuides.org>

- **Help from Diet Data Experts:** Or you may contact any member of the Oxford University's Research Team:
<https://www.nature.com/articles/s43016-023-00795-w#citeas>

Both carefully prepared tutorials start from the very beginning with a description of a sample data set and proceed step-by-step creating a selection of visual designs.

Assessment Criteria

The following will be examined to assess the coursework

The **quality** of the image supplied
 The appropriateness of the **visual design**
 The **visual design** type
 The appropriateness of the **tool** used
 The completeness of the description of the **attributes/columns**
 The description of the **attributes/columns** conveyed in the image
 The completeness and clarity of the **Visual Mappings**
 The quality of the **Unique Observation**.
 The quality of the **Data Preparation** description.
 Adherence to the report **template** provided.
 The quality of the supplementary **video**.

See the detailed assessment criteria in the appendix after the references.

Academic Integrity: This is an individual assessment that should consist of your own unaided work. You are permitted to use any material (e.g., diagrams and quotations) from the research paper that your video is based on, but you must make it clear when you are quoting from the paper. The University has [detailed advice about academic integrity](#) and submissions that demonstrate a lack of that integrity will be treated under appropriate disciplinary procedures.

References

(Liu et al, 2022) Xiaoxiao Liu, Mohammed Alharbi, Jian Chen, Alexandra Diehl, Elif E Firat, Dylan Rees, Qiru Wang, and Robert S Laramee, **Visualization Resources: A Survey**, *Information Visualization*, Volume 22, Issue 1, pages 3-30 ([PDF file](#), [web page](#), <https://doi.org/10.1177/14738716221126992>)

(Scarborough et al 2023) Scarborough, P., Clark, M., Cobiac, L. et al. **Vegans, Vegetarians, Fish-eaters and Meat-eaters in the UK Show Discrepant Environmental Impacts**. *Nature Food* 4, 565–574 (2023). <https://doi.org/10.1038/s43016-023-00795-w>

- 饮食数据专家的帮助：或者您可以联系牛津的任何成员大学研究团队：
<https://www.nature.com/articles/s43016-023-00795-w#citeas>

这两个精心准备的教程都从样本数据集的描述开始，然后逐步创建一系列可视化设计。

评审准则

将检查以下内容以评估课程作业

所提供的图像的质量 视觉设计的适当性 视觉设计类型 所用工具的适当性 属性/列描述的完整性 图像中传达的属性/列的描述 视觉映射的完整性和清晰度 独特观察的质量。

Data Preparation 描述的质量。

遵守提供的报告模板。

补充视频的质量。

请参阅参考资料后附录中的详细评估标准。

学术诚信：这是一项个人评估，应包括您自己的独立工作。您可以使用视频所依据的研究论文中的任何材料（例如图表和引文），但在引用论文时必须明确说明。大学对学术诚信有详细的建议，证明缺乏诚信的提交将根据适当的纪律处分程序进行处理。

引用

(Liu 等人, 2022 年) Xiaoxiao Liu, Mohammed Alharbi, Jian Chen, Alexandra Diehl, Elif E Firat, Dylan Rees, Qiru Wang, and Robert S Laramee, 可视化资源：调查，信息可视化，第 22 卷，第 1 期，第 3-30 页 (PDF 文件, 网页, <https://doi.org/10.1177/14738716221126992>)

(Scarborough 等人, 2023 年) Scarborough, P., Clark, M., Cobiac, L. et al. 素食主义者, 素食主义者, 英国的食鱼者和肉食者对环境的影响不同。自然界食物 4, 565–574 (2023) 。<https://doi.org/10.1038/s43016-023-00795-w>

Carbon Independent (n.d.). **Emissions from food**. [online] www.carbonindependent.org.
Available at: [Emissions from food](#).

Clark, D. (2021). **UK: population by country**. [online] Statista. Available at:
<https://www.statista.com/statistics/294729/uk-population-by-region/>.

YouGov (2024). **Dietary choices of Brits (e.g. vegetarian, flexitarian, meat-eater etc)?**
[online] yougov.co.uk. Available at:
<https://yougov.co.uk/topics/society/trackers/dietary-choices-of-brits-eq-vegetarian-flexitarian-meat-eater-etc?crossBreak=london>.

Allès, B., Baudry, J., Méjean, C., Touvier, M., Péneau, S., Hercberg, S. and Kesse-Guyot, E. (2017). **Comparison of Sociodemographic and Nutritional Characteristics between Self-Reported Vegetarians, Vegans, and Meat-Eaters**, NutriNet-Santé Study. *Nutrients*, 9(9), p.1023. DOI:<https://doi.org/10.3390/nu9091023>

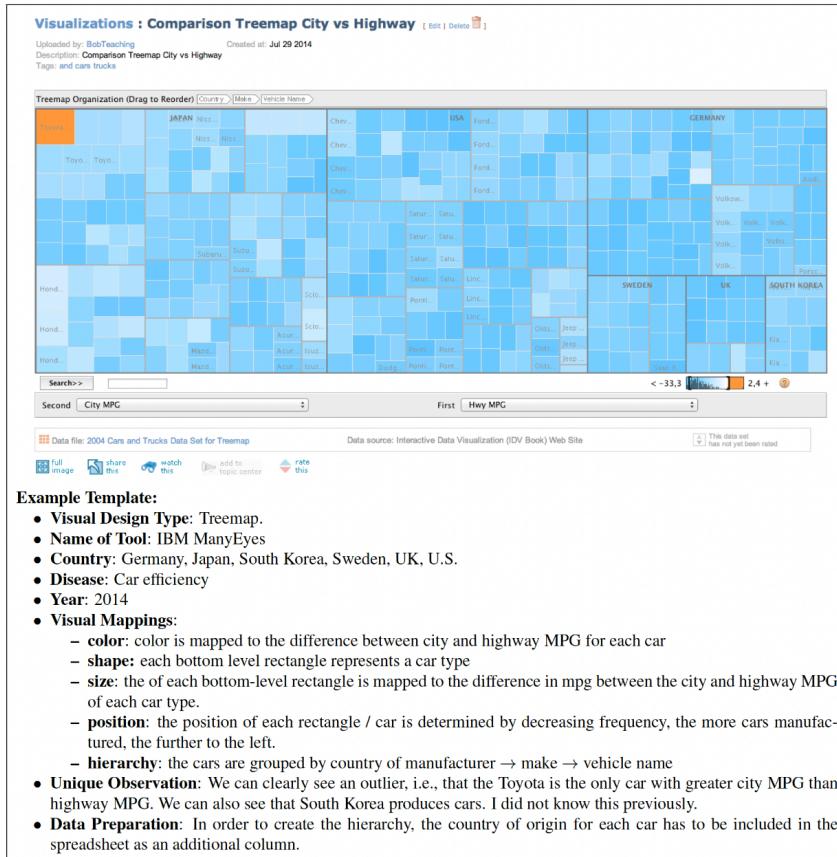
碳独立 (n.d.) 。食品排放物。[在线] www.carbonindependent.org,
可在以下网址获得：食品排放。

克拉克, D. (2021 年) 。英国: 按国家划分的人口。[在线]统计。网址:
<https://www.statista.com/statistics/294729/uk-population-by-region/>.

YouGov (2024 年) 。英国人的饮食选择 (例如素食主义者、弹性素食者、肉食者等) ?
[在线] yougov.co.uk。可在以下网址获得:
<https://yougov.co.uk/topics/society/trackers/dietary-choices-of-brits-eq-vegetarian-flexitarian-meat-eater-etc?crossBreak=london>.

Allès, B., Baudry, J., Méjean, C., Touvier, M., Péneau, S., Hercberg, S. 和 Kesse-Guyot, E. (2017) . 自我报告的素食者、纯素食者和肉食者之间的社会人口学和营养特征比较, NutriNet-Santé 研究。营养学, 9 (9) , p.1023.DOI: <https://doi.org/10.3390/nu9091023>

Example of What We Are Looking For



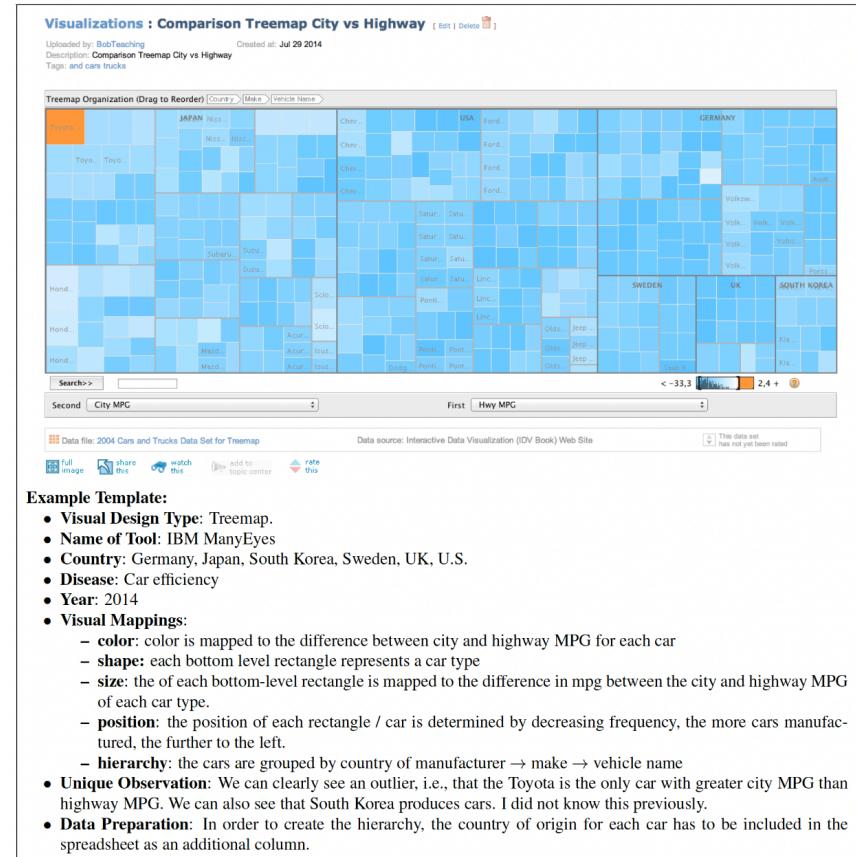
Example Template:

- **Visual Design Type:** Treemap.
- **Name of Tool:** IBM ManyEyes
- **Country:** Germany, Japan, South Korea, Sweden, UK, U.S.
- **Disease:** Car efficiency
- **Year:** 2014
- **Visual Mappings:**
 - **color:** color is mapped to the difference between city and highway MPG for each car
 - **shape:** each bottom level rectangle represents a car type
 - **size:** the of each bottom-level rectangle is mapped to the difference in mpg between the city and highway MPG of each car type.
 - **position:** the position of each rectangle / car is determined by decreasing frequency, the more cars manufactured, the further to the left.
 - **hierarchy:** the cars are grouped by country of manufacturer → make → vehicle name
- **Unique Observation:** We can clearly see an outlier, i.e., that the Toyota is the only car with greater city MPG than highway MPG. We can also see that South Korea produces cars. I did not know this previously.
- **Data Preparation:** In order to create the hierarchy, the country of origin for each car has to be included in the spreadsheet as an additional column.

Figure 1: Your task is to produce something like this. **The dataset used for this example is different from the one used for this assignment.** Therefore, your report will be different.

Also note that this example does not include a link to a supplementary video. A nice supplementary video will help in receiving more points.

我们正在寻找的示例



Example Template:

- **Visual Design Type:** Treemap.
- **Name of Tool:** IBM ManyEyes
- **Country:** Germany, Japan, South Korea, Sweden, UK, U.S.
- **Disease:** Car efficiency
- **Year:** 2014
- **Visual Mappings:**
 - **color:** color is mapped to the difference between city and highway MPG for each car
 - **shape:** each bottom level rectangle represents a car type
 - **size:** the of each bottom-level rectangle is mapped to the difference in mpg between the city and highway MPG of each car type.
 - **position:** the position of each rectangle / car is determined by decreasing frequency, the more cars manufactured, the further to the left.
 - **hierarchy:** the cars are grouped by country of manufacturer → make → vehicle name
- **Unique Observation:** We can clearly see an outlier, i.e., that the Toyota is the only car with greater city MPG than highway MPG. We can also see that South Korea produces cars. I did not know this previously.
- **Data Preparation:** In order to create the hierarchy, the country of origin for each car has to be included in the spreadsheet as an additional column.

图 1：您的任务是生成类似这样的东西。此示例使用的数据集与用于此作业的数据集不同。因此，您的报告将有所不同。

另请注意，此示例不包含指向补充视频的链接。一个很好的补充视频将有助于获得更多积分。

Assessment

Grade 80-100. Exceptional insight into the data is provided, similar to that of a good PhD student. The quality of the image and the description provided are at a publishable level. A special effort is invested into aesthetics (nice appearance) and a very clear set of labels, e.g., a labeled hierarchy. This also includes the video. The visual design is advanced. Every aspect of the template is identified and completed at a publishable level.

- The **visual design** is high quality and at a publishable level.
- The name of the **tool** and **data attributes** are provided
- The **attributes/columns** conveyed in the image are provided
- The **Visual Mappings** are completely described where applicable: color legend with labels, axis labels, position, size, shape, hierarchy
- A **Unique Observation** is provided that is unique, non-trivial, and could not be simply guessed using prior knowledge or by looking at the raw data.
- The **Data Preparation** is described at a publishable level
- The accompanying **video** is created at a publishable level
- The URL of any source code is provided, the code follows coding conventions and is commented.

Grade 70-79. Demonstrates a thorough understanding of the data and the image, explanations are very clear for all components in the image template. The visual design is advanced. Very good insight is provided. The writing has a very good structure and the work has been carried out using a professional standard.

- The **visual design** is high quality.
- The name of the **tool** and **data attributes** are provided
- The **attributes/columns** used in the image are provided
- The **Visual Mappings** are completely described where applicable: color legend with labels, axis labels, position, size, shape, hierarchy
- A Unique **Observation** is provided that is non-trivial, and could not be simply guessed using prior knowledge or by looking at the raw data.
- The **Data Preparation** is described at a very good level
- An accompanying **video** is created at a very good level
- The URL of any source code is provided

Grade 60-69. Demonstrates a decent understanding of most of the aspects of the data and the visual design, explanations are largely clear for most components of the template, the submission has a decent structure, and each part is completed, though perhaps falling short of a thorough, professional writing and image quality. The chosen visual design is advanced.

- The **visual design** is high quality.
- The name of the **tool** and **data attributes** are provided
- The **attributes/columns** used in the image are provided
- The **Visual Mappings** are described where applicable: color legend with labels, axis labels, position, size, shape, hierarchy, although falling short of a thorough and professional level.
- An **Observation** is provided that is unique, non-trivial, and could not be simply guessed using prior knowledge or by looking at the raw data.
- The **Data Preparation** is described at a good level.
- An accompanying **video** is created at a good level
- The URL of any source code is provided

评估

80-100 级。提供了对数据的非凡洞察力，类似于一个优秀的博士生。图像的质量和提供的描述处于可发布级别。在美学（漂亮的外观）和一组非常清晰的标签上投入了特别的努力，例如，标记的层次结构。这也包括视频。视觉设计很先进。模板的各个方面都在可发布级别进行识别和完成。

- 视觉设计质量高，处于可发布水平。
- 将提供工具和 data 属性的名称
- 提供了图像中传达的属性/列
- 在适用的情况下，对 Visual Mappings 进行了完整的描述：带有标签、轴标签、位置、大小、形状、层次结构的颜色图例
- 提供的唯一观察是唯一的、非平凡的，并且不能简单地使用先验知识或查看原始数据来猜测。
- 的 Data Preparation 在可发布级别进行描述
- 随附的视频是在可发布级别创建的
- 提供任何源代码的 URL，代码遵循编码约定并带有注释。

70-79 年级。展示了对数据和图像的透彻理解，图像模板中所有组件的解释都非常清晰。视觉设计很先进。提供了非常好的见解。写作结构非常好，工作是按照专业标准进行的。

- 视觉设计是高质量的。
- 将提供工具和 data 属性的名称
- 提供了图像中使用的属性/列
- 在适用的情况下，对 Visual Mappings 进行了完整的描述：带有标签、轴标签、位置、大小、形状、层次结构的颜色图例
- 提供了非同寻常的独特观察结果，不能简单地使用先验知识或查看原始数据来猜测。
- 数据准备的描述非常好
- 一个非常好的视频被制作出来
- 提供任何源代码的 URL

60-69 年级。展示了对数据和视觉设计的大部分方面的良好理解，模板的大多数组成部分的解释基本清晰，提交的结构不错，每个部分都已完成，尽管可能没有达到彻底、专业的写作和图像质量。所选的视觉设计是高级的。

- 视觉设计是高质量的。
- 将提供工具和 data 属性的名称
- 提供了图像中使用的属性/列
- 在适用的情况下描述了视觉映射：带有标签、轴标签、位置、大小、形状、层次结构的颜色图例，尽管没有达到全面和专业的水平。
- 提供的观察是唯一的、非平凡的，不能简单地使用先验知识或查看原始数据来猜测。
- 数据准备在良好的级别进行了描述。
- 制作了一个良好的视频
- 提供任何源代码的 URL

Grade 50-59. Demonstrates some understanding of the data and the visualization, the descriptions and explanations of each component of the template are generally okay but sometimes lack detail, are ambiguous, or contain errors. There is some attempt to structure the submission, and a decent (if perhaps) informal attempt has been made at the writing, image, and the presentation. The insight provided is rather obvious or trivial. The visual design could be improved.

- The **visual design** is of reasonable quality.
- The name of the **tool** and **data attributes** are provided
- The **attributes/columns** used in the image are provided
- The **Visual Mappings** are described where applicable: color legend with labels, axis labels, position, size, shape, hierarchy, at an adequate level.
- A **Unique Observation** is provided but is not unique, trivial, and could be simply guessed using prior knowledge or by looking at the raw data.
- The **Data Preparation** is described.
- An accompanying video is absent.
- The URL of any source code is absent.

Grade 40-49. There is some understanding of the data and visualization, but mostly the descriptions are vague, lack a lot of detail, or have errors. The supplied image is very basic with no special insight given. There is a vague structure to the submission but it isn't easy to follow throughout, some attempt at writing.

- The **visual design** is low quality.
- The name of the **tool** and **data attributes** are not provided.
- The **attributes/columns** used in the image are not provided
- The **Visual Mappings** are described poorly.
- A **Unique Observation** is provided but is not unique, trivial, and could be simply guessed using prior knowledge or by looking at the raw data.
- The **Data Preparation** is poorly described.
- An accompanying video is provided but needs improvement

Grade 30-39. A little bit of understanding and explanation of the data and visual design, but most of the components in the template description are not explained in much detail or contain significant errors. The submission doesn't have much of a structure, there is some attempt at writing. No significant insight is conveyed by the image.

- The **visual design** is low quality.
- The name of the **tool** and **data attributes** are not provided.
- The **attributes/columns** used in the image are not provided
- The **Visual Mappings** are described poorly or erroneously.
- A **Unique Observation** is provided but is not unique, trivial, and could be simply guessed using prior knowledge or by looking at the raw data.
- The **Data Preparation** is poorly or erroneously described.
- An accompanying video is provided but needs improvement

Grade 20-29. A few aspects of the data and visualization have been understood, but overall the submission doesn't show much understanding of the visual design nor the explanation of different components of the template description. The submission is poorly structured and contains little coherent attempt at writing. No effort is made to convey insight into the data.

The visual design is too basic.

- The **visual design** is low quality and too basic.
- The name of the **tool** and **data attributes** are not provided.

50-59 年级。展示对数据和可视化的一些理解，模板每个组件的描述和解释通常都没问题，但有时缺乏细节、模棱两可或包含错误。有一些尝试来组织提交，并且一个体面的（如果

也许）在写作、图像和演示方面已经进行了非正式的尝试。提供的见解相当明显或微不足道。视觉设计有待改进。

- 视觉设计质量合理。
- 将提供工具和 data 属性的名称
- 提供了图像中使用的属性/列
- 在适用的情况下描述了 Visual Mappings：带有标签、轴标签、位置、大小、形状、层次结构的颜色图例，处于适当的级别。
- 提供了唯一观察结果，但不是唯一的、微不足道的，可以使用先验知识或查看原始数据简单地猜测。
- 介绍了 Data Preparation。
- 没有随附的视频。
- 任何源代码的 URL 都不存在。

40-49 年级。对数据和可视化有一定的了解，但大多数描述含糊不清，缺乏很多细节，或者有错误。提供的图像非常基本，没有给出特殊的见解。

提交的结构模糊不清，但并不容易贯穿始终，一些写作尝试。

- 视觉设计质量低下。
- 未提供工具和 data 属性的名称。
- 未提供图像中使用的属性/列
- Visual Mappings 的描述很糟糕。
- 提供了唯一观察结果，但不是唯一的、微不足道的，可以使用先验知识或查看原始数据简单地猜测。
- 数据准备描述不佳。
- 提供了随附的视频，但需要改进

30-39 年级。对数据和视觉设计有一点了解和解释，但模板描述中的大多数组件没有详细解释或包含重大错误。提交的内容没有太多的结构，有一些写作的尝试。该图像没有传达任何重要的见解。

- 视觉设计质量低下。
- 未提供工具和 data 属性的名称。
- 未提供图像中使用的属性/列
- Visual Mappings 描述不佳或错误。
- 提供了唯一观察结果，但不是唯一的、微不足道的，可以使用先验知识或查看原始数据简单地猜测。
- 数据准备描述不佳或错误。
- 提供了随附的视频，但需要改进

20-29 年级。数据和可视化的一些方面已经被理解，但总的来说，提交的内容并没有显示出对视觉设计的太多理解，也没有表现出对模板描述的不同组成部分的解释。该提交的结构很差，几乎没有连贯的写作尝试。没有努力传达对数据的见解。

视觉设计太基础了。

- 视觉设计质量低下且过于基础。
- 未提供工具和 data 属性的名称。

- The **attributes/columns** used in the image are not provided.
- The **Visual Mappings** are described poorly or erroneously or are not provided.
- No serious attempt at a **Unique Observation**.
- No serious attempt describing the **Data Preparation**.
- No attempt at a quality video.

Grade 10-19. Shows some awareness of the data and the associated visualization, but doesn't go much beyond mentioning a few keywords and doesn't have any meaningful structure or attempt at writing. The supplied image is low quality or trivial.

- The **visual design** is low quality or contains errors.
- The name of the **tool** and **data attributes** are not provided or contain errors.
- The **attributes/columns** used in the image are not provided or contain errors.
- The **Visual Mappings** are described poorly or erroneously or are not provided.
- No serious attempt at a **Unique Observation**.
- No serious attempt describing the **Data Preparation**.

Grade 0-9. No or minimal attempt

- 未提供图片中使用的属性/列。
- Visual Mappings 描述不佳或错误，或者未提供。
- 没有认真尝试进行独特观察。
- 没有认真的尝试来描述数据准备。
- 不尝试提供高质量的视频。

10-19 年级。显示了对数据和相关可视化的一些认识，但除了提到几个关键字之外，没有太多内容，并且没有任何有意义的结构或编写尝试。提供的图像质量低或微不足道。

- 视觉设计质量低下或包含错误。
- 未提供工具和数据属性的名称，也不包含错误。
- 图像中使用的属性/列未提供或包含错误。
- Visual Mappings 描述不佳或错误，或者未提供。
- 没有认真尝试进行独特观察。
- 没有认真的尝试来描述数据准备。

0-9 年级。无尝试或极少尝试