

5.1 可视化入门

2023 年 9 月 18 日

5.1 可视化入门

俗话说“一图胜千言”，这恰恰体现了视觉传播的价值和效率。

1. 可视化的科学

1.1 视觉中枢

研究发现，大脑的一大任务就是视觉处理，几乎半个大脑都是直接或间接用来处理视觉功能的，详见 [1]。

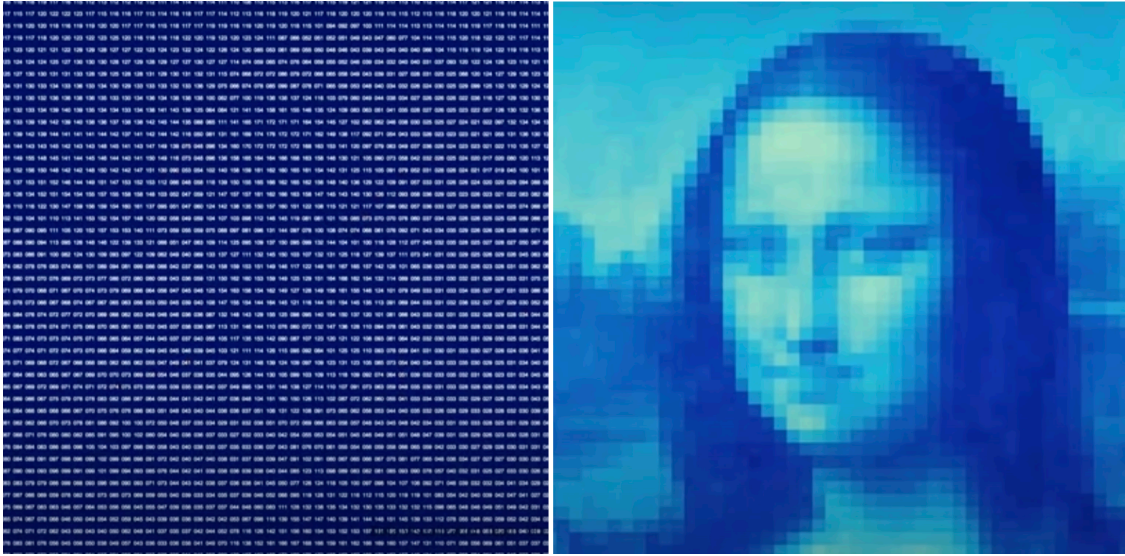
涉及这项活动的细胞、神经元、纤维的数量非常庞大。仅眼睛内部的视网膜就由超过 1.5 亿个细胞构成，它实际上是大脑的延伸部分。另外，大脑中与视觉相关的神经元占了很大的比重，将近 30% 的脑灰质由这些神经元组成。相比之下，与触觉和听觉有关的神经元只占了 8% 和 3%，参见 [2]。

1.2 视觉利于理解

基于这些视觉“网络”，我们不难理解，比起单纯文本，大脑处理信息图会更加省力。

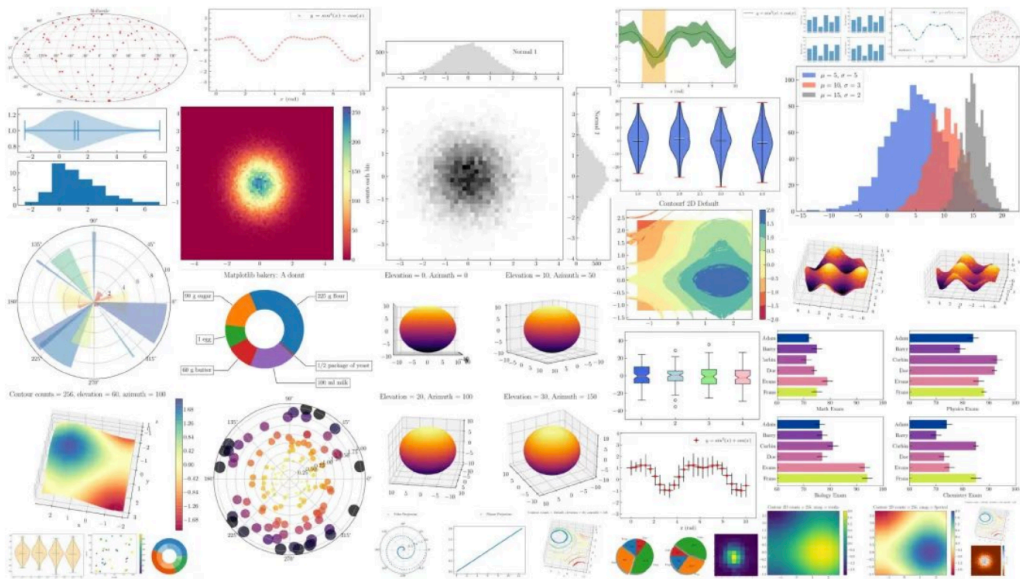
单词中的每一个字母都代表一个符号。在阅读文本的时候，大脑首先对其解码，将这些字母与记忆中存储的形状相匹配，进一步理解这些字母如何组成单词，单词再组成句子，句子最后组成段落。尽管这一解读过程只需瞬间完成，但与大脑处理图像的方式相比，还是需要消耗更多的脑力，参见 [3]。

有时，我们无法直观的理解数据，就像下图，如果不将数字进行编码和着色，我们永远猜不到图片里隐藏了一位女士。



2. 可视化的图类型

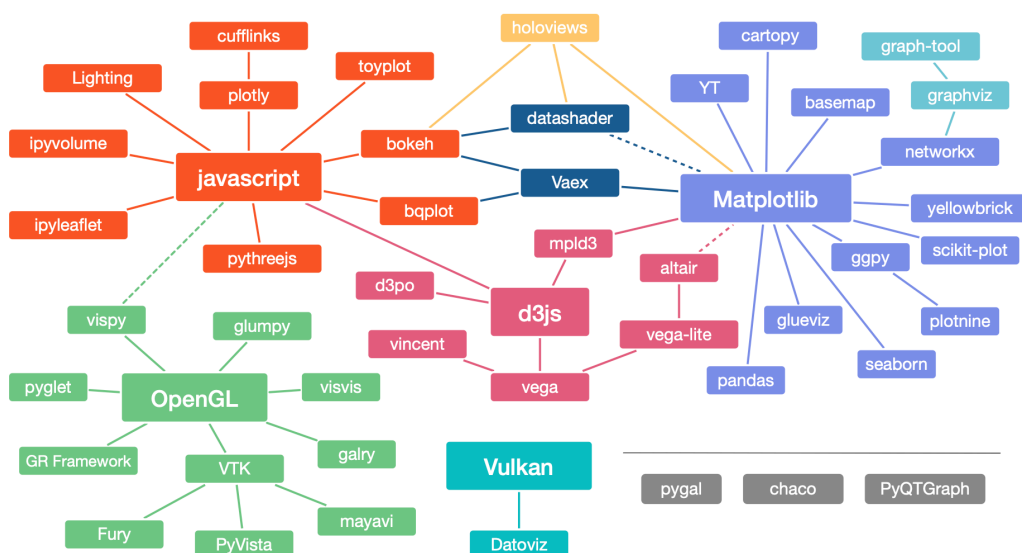
数据可视化，主要体现对于数据的呈现方式。比如使用线条、饼状、柱状、条状、环状、地理图等来呈现数据。



3. 可视化与 Python

Python 科学可视化领域是巨大的 (见下图)。它由无数的工具组成，从最通用和广泛使用的到更专业和精密的。其中一些工具是基于社区的，而另一些则由公司开发。有些是专门为网页制作的，有些

只针对桌面，有些处理 3D 和大数据，而有些则针对完美的 2D 渲染。



选择很多造成了很难找到最适合你需要的库，只是因为您可能甚至没有意识到这个或那个库的存在。为了帮助你做出选择，你可以先问自己几个问题：- 你的目标是桌面渲染还是网页渲染？- 你需要复杂的 3D 渲染吗？- 你需要出版质量吗？- 你有非常大的数据吗？- 是否有活跃的社区？- 是否有文档和教程？

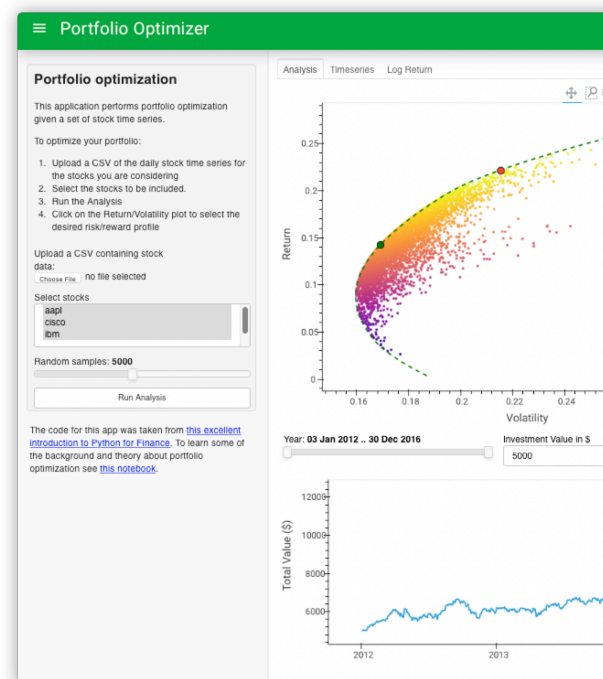
根据你的答案，你可能能够决定使用哪个 Python 库，并投入一些时间来学习它。例如，如果你需要在浏览器中与 Jupyter 无缝集成的交互式可视化，**bokeh** 可能是一个答案。如果你有非常大的数据，需要在桌面上使用 3D，**vispy** 或 **mayavi** 可能是一个选择。如果你对一个非常直观的工具感兴趣，快速展示漂亮的统计信息，那么 **seaborn** 和 **altair** 是你的朋友。然而，如果你从事的是地球科学，那么你就不能忽视 **cartopy** 等。pyviz 网站是一个很好的信息来源，它提供了许多可视化库的列表。

可视化库	官网	可交互	擅长
matplotlib	https://matplotlib.org/stable/index.html	否	基础绘图库
seaborn	https://seaborn.pydata.org	否	统计数据可视化
bokeh	https://bokeh.org	是	-

可视化库	官网	可交互	擅长
altair	https://altair-viz.github.io/gallery/index.html	是	-

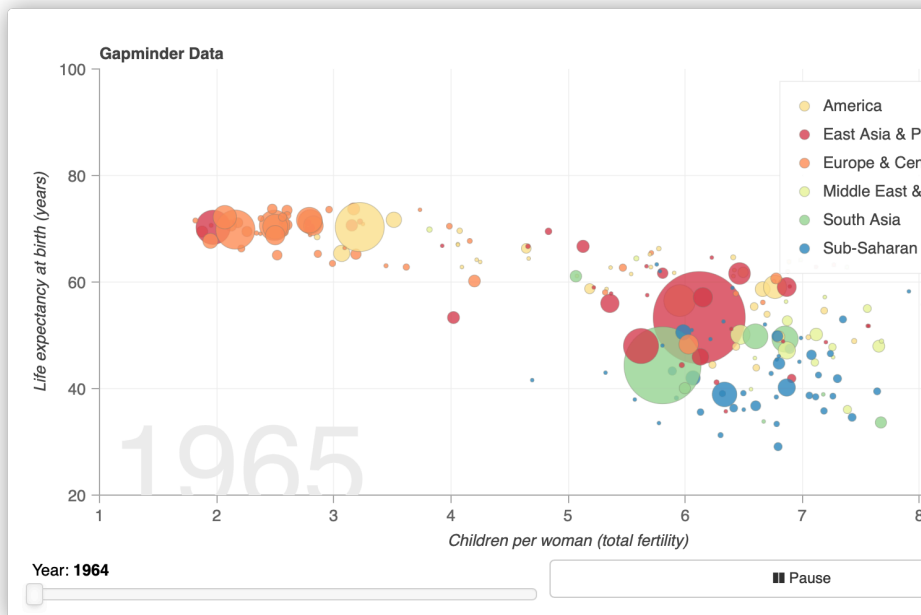
4. 一些有趣的应用

4.1 金融投资组合优化



链接: <https://portfolio-optimizer.pyviz.demo.anaconda.com/portfolio>

4.2 动态显示各个国家的生育率



链接: <https://demo.bokeh.org/gapminder>

参考

1. MIT Research - Brain Processing of Visual Information, <http://bit.ly/smIch0>
2. 参见 Denise Grady 发表在 Discover 杂志的文章: “The Vision Thing: Mainly in the Brain”, <http://bit.ly/upYVBr>
3. “Show Me! What Brain Research Says About Visuals in PowerPoint”, 链接为: <http://bitly.com/s3lseP>
4. Mark Smiciklas, 《视不可挡: 信息图与可视化传播》, 人民邮电出版社
5. Nicolas P.Rougier, Scientific Visualization: Python & Matplotlib[M], <https://github.com/rougier/scientific-visualization-book/>
6. pyviz, 帮助用户根据自己的目的决定最好的开源 (OSS) Python 数据可视化工具, 提供链接、概述、比较和示例, <https://pyviz.org/overviews/index.html>
7. 高维空间可视化: <https://www.youtube.com/watch?v=wvsE8jm1GzE&t=30s>
8. 使用 matplotlib 和 pandas 的绘图示例代码, <https://www.python-graph-gallery.com>