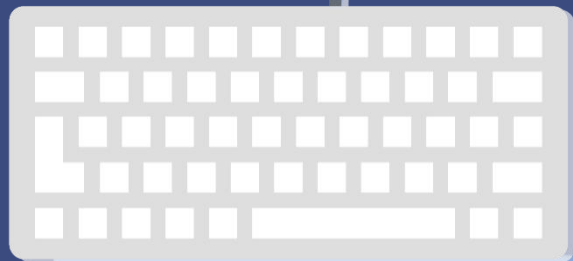




数据可视化导论





初识数据可视化技术

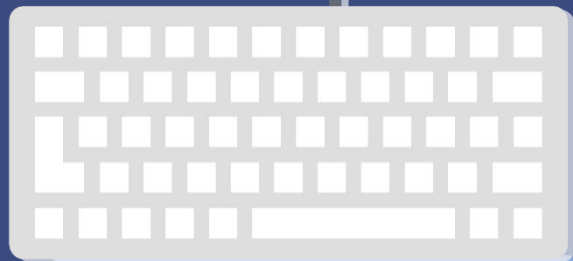
目录

- ◆ 数据可视化定义
- ◆ 数据可视化发展历程
- ◆ 数据可视化作用
- ◆ 数据可视化流程
- ◆ 典型应用场景
- ◆ 数据可视化操作工具



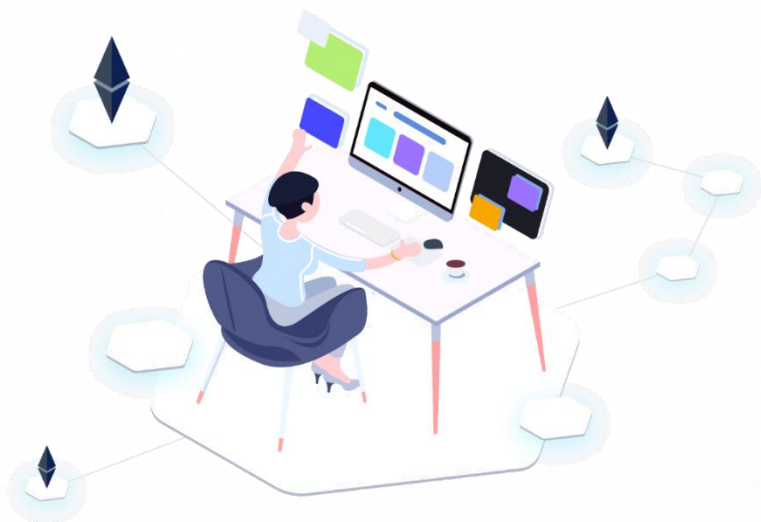


01 数据可视化定义





>>> 数据的视觉表现形式



◆ 图形与图像处理

在数据可视化中，图形与图像处理技术扮演着至关重要的角色。通过将复杂的数据集转换为直观的图表和图像，这些技术帮助用户快速理解数据模式和趋势。

◆ 计算机视觉的应用

计算机视觉技术使得机器能够“看”懂图像内容，进而分析和解读数据。这项技术广泛应用于面部识别、物体检测等领域，极大地丰富了数据可视化的表现力。

◆ 用户界面设计的重要性

优秀的用户界面设计是数据可视化成功的关键。它不仅需要美观吸引，更重要的是要提供清晰、直观的数据展示方式，确保用户能够轻松地获取信息并作出决策。



初识数据可视化技术

科学 & 艺术结合

01

数据可视化的科学性

数据可视化基于统计学、心理学和计算机科学等多学科知识，通过科学的方法和工具，将复杂的数据集转化为易于理解的视觉形式，帮助用户发现数据背后的规律和趋势。

02

数据可视化的艺术性

数据可视化不仅仅是展示数据，更是一种艺术表达。设计师通过色彩、布局、形状等视觉元素，创造出既美观又富有表现力的图表，使数据故事更加生动和吸引人。

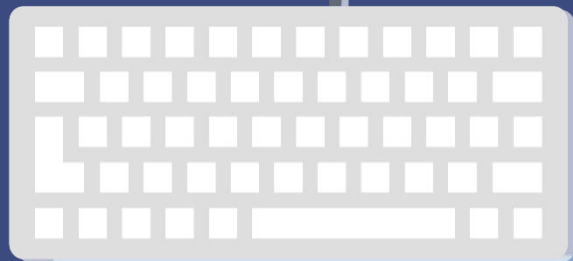
03

科学与艺术的融合

在数据可视化中，科学与艺术的结合体现在如何平衡准确性和美感。一方面，确保数据的真实性和精确性；另一方面，通过创意设计提升用户体验，使数据传达更加高效和愉悦。



02数据可视化发展历程





初识数据可视化技术

>>> 古代地图绘制

01

古代地图绘制的起源

古代地图绘制起源于人类对地理空间认知的需求，通过简单的线条和符号记录地形地貌，为早期文明的探索与交流提供了基础。

02

手工制图技术的发展

随着科学技术的进步，手工制图技术得到显著提升，人们开始使用更精确的工具和方法来制作地图，使得地图的准确性和实用性大大增强。

03

可视化工具的应用

在16世纪，几何图表和地图成为可视化的主要形式，这些工具不仅用于展示地理信息，还广泛应用于商业数据分析，促进了信息的传递和理解。



16世纪几何图表

01

几何图表的兴起

在16世纪，随着商业和航海活动的增加，人们开始寻求更有效的方式来展示地理信息和商业数据。几何图表因其直观性和易于理解的特点而逐渐流行起来。

02

商业数据的可视化

随着商业活动的发展，商人和企业开始利用几何图表来分析和展示商业数据，如贸易路线、商品流通等，这有助于他们做出更明智的商业决策。

03

地图制作的进步

16世纪的地图制作技术得到了显著的提升，制图师们开始使用更精确的测量工具和方法，使得地图更加准确和详细，为航海和探险提供了重要的支持。



初识数据可视化技术

>>> 17世纪物理测量

01

物理测量的革新

17世纪，随着科学革命的兴起，物理测量技术得到了显著的发展。科学家们开始使用更加精确的仪器和方法来测量自然现象，为数据可视化奠定了基础。

02

图表和图形的应用

17世纪，科学家们开始使用图表和图形来记录和分析实验数据。这种直观的方式使得复杂的数据更容易被理解和解释，推动了科学研究的进步。

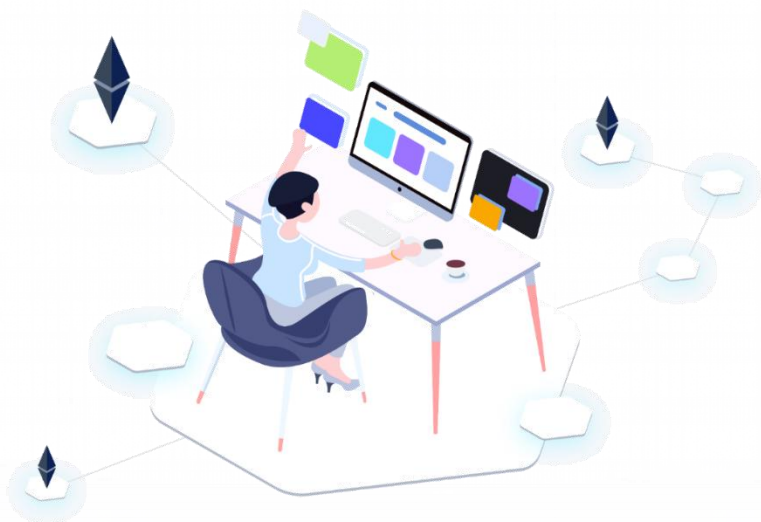
03

数据可视化的发展

17世纪的物理测量不仅推动了数据可视化的发展，也为后来的统计学和数据分析提供了重要的工具。这些工具至今仍在各个领域发挥着重要作用。



20世纪计算机支持



◆ 计算机技术的崛起

20世纪中叶，随着电子计算机的发明和普及，数据处理能力得到极大提升，为数据可视化提供了必要的技术支撑，开启了信息时代的新篇章。

◆ William S. Cleveland的贡献

William S. Cleveland在美国国家标准局的工作期间，对数据可视化进行了深入研究，提出了一系列理论和方法，为该领域的发展奠定了坚实的基础。

◆ 数据可视化理论的形成

在20世纪60年代，随着计算机技术的发展，数据可视化开始形成独立的研究领域，William S. Cleveland等人的研究为这一领域提供了重要的理论基础。



初识数据可视化技术

1970年代科学研究应用

01

数据可视化在天文学中的应用

1970年代，随着计算机技术的发展，天文学家开始利用数据可视化技术来分析宇宙数据，通过图表和图像的形式更直观地理解星系、行星等天体的运动规律。

02

生物信息学的兴起

在生物学领域，数据可视化帮助科学家解析复杂的遗传信息，通过图形化展示DNA序列、蛋白质结构等，促进了基因工程和分子生物学的研究进展。

03

地球科学的数据解读

地质学家和气象学家使用数据可视化工具来分析地震波、气候模型等数据，这些工具帮助他们更好地预测自然灾害，提高了对地球系统变化的理解。



初识数据可视化技术

1980年代计算机图形学发展

计算机图形学的兴起

在1980年代，随着计算机硬件性能的显著提升和软件算法的进步，计算机图形学开始迅速发展，为数据可视化提供了强大的技术支持。



数据可视化技术的革新

这一时期，数据可视化技术经历了重大革新，从简单的图表和图形发展到复杂的三维模型和动画，极大地丰富了数据的表现形式。



精确度与表现力的提升

随着计算机图形学的发展，数据可视化的精确度和表现力得到了显著提升，使得研究人员能够更准确地分析和解释复杂数据集。





初识数据可视化技术

1990年代互联网兴起

01



互联网的普及与影响

1990年代，随着互联网技术的飞速发展，全球范围内实现了信息的快速传播和共享，极大地改变了人们的生活方式和模式。

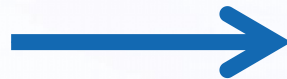
02



数据可视化的兴起

在互联网技术的支持下，数据可视化作为一种新兴的信息表达方式开始流行，它通过图形化手段帮助人们更直观地理解和分析数据。

03



在线数据分析的发展

随着互联网技术的发展，在线数据分析成为可能，企业和研究机构能够实时收集、处理和分析大量数据，为决策提供支持。



初识数据可视化技术

2000年代大数据时代

01 大数据技术的兴起

2000年代，随着互联网和信息技术的飞速发展，大数据技术应运而生。它通过收集、存储、分析海量数据，为企业决策提供有力支持，成为新时代的重要生产力。

03 大数据技术的兴起

2000年代以来，大数据技术逐渐渗透到各行各业。从金融、医疗到教育、交通等领域，大数据都发挥着重要作用，推动行业创新和发展，为社会带来巨大变革。

02 数据可视化的重要性

在大数据时代，数据可视化技术发挥着至关重要的作用。它将复杂的数据以直观、易懂的方式呈现，帮助人们快速理解数据背后的信息，提高数据分析的效率和准确性。



初识数据可视化技术

2010年代人工智能应用

自动驾驶技术的突破

2010年代，自动驾驶技术通过深度学习算法实现了对交通标志、行人和其他车辆的识别，谷歌Waymo项目成为该领域的标志性成果，标志着自动驾驶向实用化迈进。



智能家居系统的普及

利用人工智能技术，智能家居系统能够实现家居设备的自动化控制和管理，智能音箱等设备通过语音识别与用户交互，极大地提高了生活的舒适度和便利性。



医疗健康领域的AI应用

人工智能在医疗健康领域展现出巨大潜力，通过图像识别技术辅助疾病诊断，同时在药物研发、患者管理和远程医疗服务中发挥作用，显著提升了医疗服务效率和质量。





03数据可视化的作用





提高理解和分析能力



◆ 数据可视化的定义与重要性

数据可视化是一种将复杂数据通过图形、图表等视觉元素转化为易于理解的信息的技术，它在现代信息处理中发挥着重要作用，帮助用户更好地洞察数据的内涵与规律。

◆ 提高理解和分析能力的方式

通过使用数据可视化工具，我们可以将复杂的数据转化为易于理解的图形和图表，从而提高我们的理解和分析能力，使我们能够更快地从大量数据中提取有价值的信息。

◆ 数据可视化的实际应用

在实际应用中，数据可视化已经被广泛应用于各个领域，如商业分析、科学研究等，它不仅提高了我们的工作效率，也帮助我们做出了更准确的决策。



➤➤➤ 加快决策速度

01

直观呈现复杂数据

通过将复杂的数据集转化为图表和图形，决策者可以迅速识别模式、趋势和异常点，从而在较短时间内做出基于数据的决策。

02

提高信息处理效率

数据可视化工具能够高效地处理和分析大量数据，帮助决策者快速筛选关键信息，减少手动分析的时间和错误，提升决策效率。

03

促进团队协作与沟通

通过共享易于理解的视觉化数据报告，团队成员之间可以更有效地交流想法和见解，共同参与决策过程，加快决策速度并提高决策质量。



初识数据可视化技术

促进团队合作

01

迅速传达信息

数据可视化通过图表和图形的形式，能够快速向团队成员展示数据背后的故事和趋势，有效减少信息传递过程中的误解和歧义。

02

降低沟通成本

直观的图形和图表帮助团队成员更容易理解复杂数据，减少了解释和讨论所需的时间，从而降低了团队内部的沟通成本。

03

增强协作效率

当团队成员能够清晰地看到数据和分析结果时，他们可以更有效地协同工作，共同解决问题，从而提高了团队的整体协作效率。



初识数据可视化技术

提升数据驱动思维

01



鼓励以数据为基础的思维方式

数据可视化通过将复杂数据转化为直观的图表和图形，极大地促进了基于数据的思维方式，使人们能够更轻松地分析和理解数据背后的信息。

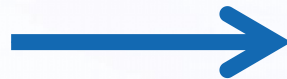
02



促进问题解决和业务优化

利用数据可视化展示数据的趋势、模式和关系，帮助用户快速识别问题根源，从而制定出更有效的解决方案和业务优化策略，提升整体效率。

03



增强决策依据

数据可视化提供了清晰、直观的数据证据，支持决策者做出基于数据的决策，减少主观判断的影响，从而提高决策的准确性和可靠性。

04数据可视化流程





数据收集

01

数据库查询

数据库查询是数据收集中常用的方法之一，通过编写SQL语句直接从数据库中检索所需数据，这种方法可以快速准确地获取大量结构化数据。

02

API调用

API调用允许开发者利用应用程序接口（API）获取实时或定期更新的数据，这种方式适用于需要频繁更新数据的应用场景，能够保证数据的时效性和准确性。

03

网络爬虫技术

网络爬虫技术使用自动化脚本从互联网上抓取公开信息，适用于大规模数据采集任务，但需要注意遵守相关法律法规和网站的爬取规则。



初识数据可视化技术

数据清洗

01 缺失值处理

缺失值处理是数据清洗中的关键步骤，涉及识别并处理数据集中的空白或不完整记录。通过删除、填充或插值等方法，确保数据的完整性和准确性，为后续分析打下坚实基础。

03 异常值检测

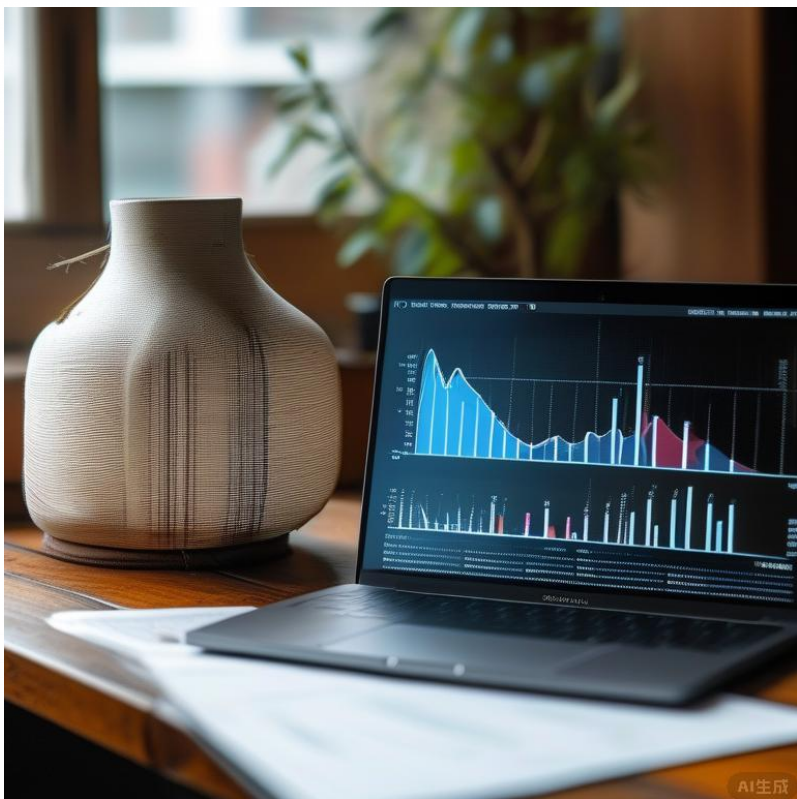
异常值检测旨在识别并处理数据集中不符合预期的极端值。这些异常值可能由错误或特殊情况造成，通过适当的处理方法，可以保证数据分析的准确性和可靠性。

02 格式标准化

格式标准化是将数据转换为统一格式的过程，包括日期、时间、数值等的统一表示。这一步骤有助于消除数据格式差异，提高数据处理效率，确保数据分析的一致性和准确性。



数据分析师



◆ 描述性统计分析

描述性统计分析是数据分析的基础，通过对数据集的均值、中位数和标准差等基本统计量的计算，帮助我们快速了解数据的整体分布情况。

◆ 探索性数据分析（EDA）

探索性数据分析通过图表和图形的形式，揭示数据中的模式、趋势和异常值，为深入分析提供直观依据，是理解数据特性的重要步骤。

◆ 假设检验与回归分析

假设检验用于验证数据背后的假设是否成立，而回归分析则探究变量间的关联性，预测因变量的变化，两者共同支持数据的科学解读和应用。



初识数据可视化技术

数据可视化设计

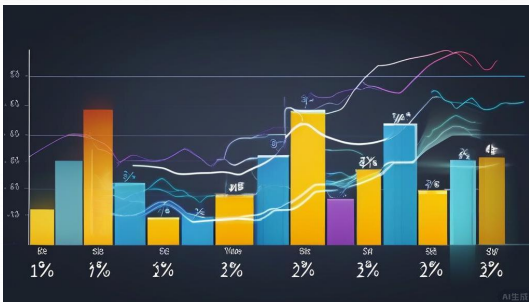
确定数据类型

在数据可视化设计之初，明确展示的数据类型至关重要。这包括区分定量数据、定性数据或时间序列数据等，以确保后续的图表选择和设计能够准确反映数据特性。



选择图表类型

根据数据的特性和分析目的，选择合适的图表类型是关键步骤。例如，柱状图适用于比较不同类别的数据，而折线图则更适合展示数据随时间的变化趋势。



故事叙述

通过图表讲述数据背后的故事，清晰地传达数据分析的结论和洞察，使观众能够快速抓住重点。





数据可视化实现

01

数据可视化工具选择

选择合适的可视化工具是实现有效数据展示的关键一步。市面上流行的工具如 Tableau、Power BI 提供了丰富的图表类型和交互功能，适合快速开发和部署。

02

自定义可视化的实现

对于特定的数据分析需求，使用编程语言如 Python、R 或 JavaScript 进行自定义可视化开发，可以更灵活地控制图表的样式和功能，满足个性化需求。

03

数据可视化目的

数据可视化的主要目的是将复杂的数据分析结果转化为直观易懂的图形表示，帮助用户更容易理解和解释数据，从而做出更加明智的决策。



初识数据可视化技术

数据解释与应用

01

数据解释的重要性

数据解释是数据可视化流程中的关键步骤，它帮助观众理解图表中的信息和发现。通过有效的解释，观众可以从数据中获得洞见，并据此做出决策或采取行动。

02

数据应用的方式

数据应用可以通过多种方式进行，包括报告、演示、文章等。这些方式可以帮助观众更好地理解 and 利用数据，从而从中获得价值。

03

数据解释与应用的目标

数据解释与应用的目标是使观众能够从数据中获得洞见，并据此做出决策或采取行动。这是实现数据价值的重要途径，也是数据可视化的最终目标。



05数据可视化应用场景





科学研究的



数据可视化技术在科研中的应用

数据可视化技术通过将复杂的科学数据转化为直观的图形和图表，帮助研究人员更好地理解 and 解释数据，从而推动科学发现和创新。

实验结果的展示与分析

利用数据可视化技术，研究人员可以将实验结果以直观的方式呈现出来，便于对实验数据进行深入分析和比较，提高研究效率。

模拟分析的可视化表达

数据可视化技术在模拟分析中发挥着重要作用，通过将模拟结果以图形化的形式展示，有助于研究人员更直观地理解模型的行为和特性。



初识数据可视化技术

交通运输

01



交通疏解可视化技术

在交通运输领域，数据可视化技术被用于监测和管理交通状况，通过将复杂的交通数据以图形、图表等形式直观展示，帮助交通管理部门和公众更好地理解交通流量、拥堵情况等信息。

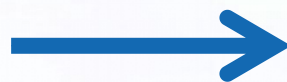
02



提高交通安全

数据可视化技术可以帮助交通管理部门及时发现事故频发区域，通过对这些区域的深入分析，可以采取有效的措施来预防交通事故的发生，从而提高整体的交通安全水平。

03



优化交通流

通过数据可视化技术，交通管理部门可以实时监控交通流量，根据交通状况调整信号灯配时、路线规划等，有效缓解交通拥堵，优化交通流，提高道路使用效率。



初识数据可视化技术

资源管理

01

数据可视化技术在资源管理中的应用

数据可视化技术通过将复杂的地理和资源数据转化为易于理解的视觉形式，使资源管理和开发更加直观，帮助决策者更有效地监控、分析和规划资源的使用和管理。

02

数字地图的形成与作用

数字地图是数据可视化技术的一种重要应用，它能够清晰地展示资源的分布情况，为资源管理提供直观的参考，有助于提高资源利用效率和管理水平。

03

提升资源管理决策效率

通过数据可视化技术，决策者可以快速获取并理解大量复杂的资源数据，从而做出更准确、更及时的决策，有效提升资源管理的决策效率和质量。



初识数据可视化技术

医疗健康

01



数据可视化在医疗健康中的应用

数据可视化技术通过将复杂的医疗数据转化为直观的图表和图形，帮助医生快速识别患者的健康状况、疾病发展趋势以及治疗效果，从而提高医疗服务的效率和质量。

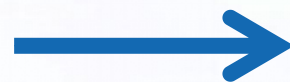
02



提高诊断效率与准确性

利用数据可视化技术，医生可以更直观地了解病人的健康数据，从而及时进行诊断和治疗。这种技术的应用有助于提高医疗服务的效率和质量，减少误诊率。

03

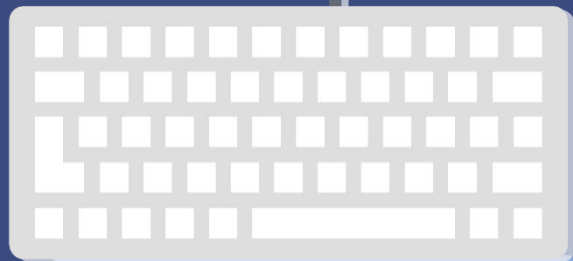


促进个性化治疗方案制定

数据可视化技术可以帮助医生更好地理解病人的病情，从而制定出更加个性化的治疗方案。这种方案能够更好地满足病人的需求，提高治疗效果。



06数据可视化工具





数据可视化工具介绍

➤➤➤ Matplotlib



◆ Matplotlib的起源与设计

Matplotlib作为Python中数据可视化的先驱，其设计理念深受20世纪80年代MATLAB的影响，旨在为科学家和工程师提供强大的绘图工具。

◆ 丰富的图表类型支持

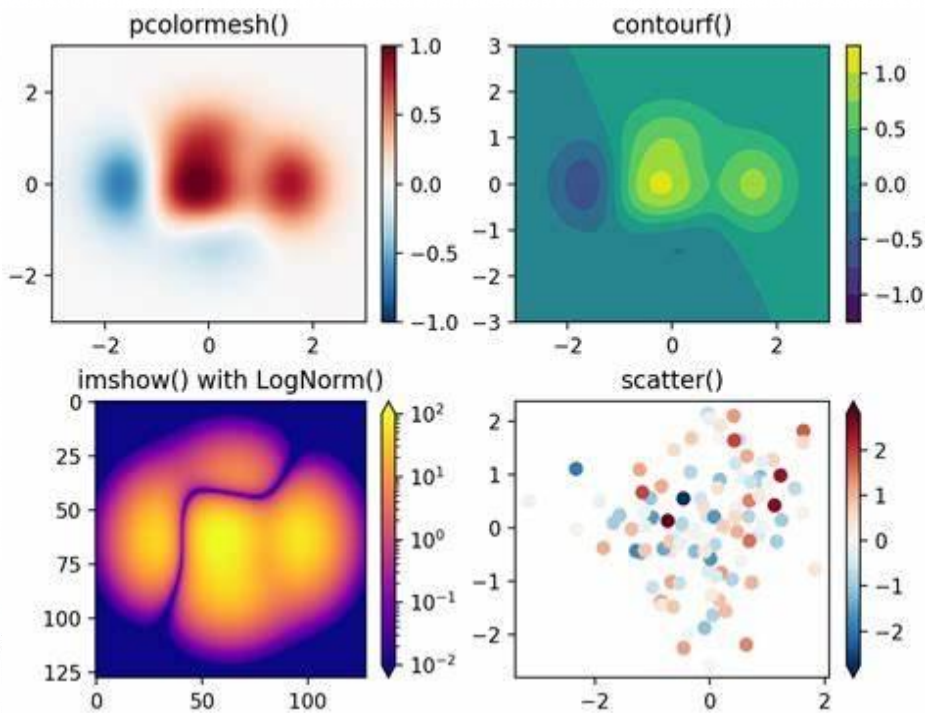
Matplotlib支持多种图表类型，包括线图、散点图、条形图等，满足不同数据展示需求，使数据分析更加直观易懂。

◆ 高度自定义与扩展性

Matplotlib提供了丰富的自定义选项，用户可以根据需要调整图表的样式、颜色、标签等，同时支持与其他Python库结合使用，增强功能。



➤➤ Matplotlib



◆ Matplotlib的起源与设计

Matplotlib作为Python中数据可视化的先驱，其设计理念深受20世纪80年代MATLAB的影响，旨在为科学家和工程师提供强大的绘图工具。

◆ 丰富的图表类型支持

Matplotlib支持多种图表类型，包括线图、散点图、条形图等，满足不同数据展示需求，使数据分析更加直观易懂。

◆ 高度自定义与扩展性

Matplotlib提供了丰富的自定义选项，用户可以根据需要调整图表的样式、颜色、标签等，同时支持与其他Python库结合使用，增强功能。



数据可视化工具介绍

>>> Seaborn

01

Seaborn简介

Seaborn是一个基于Matplotlib的高级可视化库，它提供了简洁的语法和美观的色彩方案，使得数据可视化变得更加简单和吸引人。

02

交互式界面支持

Seaborn支持交互式界面，用户可以通过简单的操作来探索和理解数据，这使得数据分析过程更加直观和高效。

03

统计数据可视化专注

Seaborn专注于统计数据的可视化，它提供了丰富的统计图形类型，可以有效地揭示数据集中的复杂关系，帮助用户深入理解数据。



数据可视化工具介绍

>>> Plotly

01

Plotly的交互式绘图功能

Plotly提供了强大的交互式绘图功能，用户可以通过缩放、平移和悬停在数据点上以获取更多信息，这使得数据分析变得更加直观和深入。

02

Plotly的可视化效果丰富性

Plotly支持丰富的可视化效果，包括科学图表、3D图表、统计图表和金融图表等，能够满足各种复杂的数据可视化需求。

03

Plotly的开源性和全面性

Plotly是一个开源且全面的工具，它不仅提供了丰富的可视化效果，还支持在Python中创建声明性可视化，使得用户可以更方便地进行数据分析和可视化。



数据可视化工具介绍

Bokeh

01



Bokeh的核心特性

Bokeh是一款Python交互数据可视化库，它以现代网络浏览器为呈现目标，提供优雅、简洁的新颖图形结构和高性能交互性。这使得用户能够轻松地创建出既美观又实用的图表。

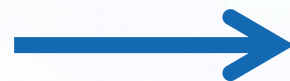
02



Bokeh的共享与集成功能

Bokeh支持在线共享选项以及与Jupyter Notebooks的集成，这意味着用户可以方便地将他们的可视化作品分享给他人，或者在Jupyter Notebooks中直接使用Bokeh进行数据可视化。

03



Bokeh的API设计

Bokeh提供了高级和低级API以满足不同需求。高级API使得初学者可以快速上手，而低级API则为有经验的开发者提供了更大的灵活性和控制力。



数据可视化工具介绍

>>> Pygal



Pygal 简介

Pygal 是一个开源的 Python 数据可视化库，它允许用户创建高度互动的图表和图形，支持 SVG 和 PNG 格式下载，适用于多种应用场景。

高度互动性

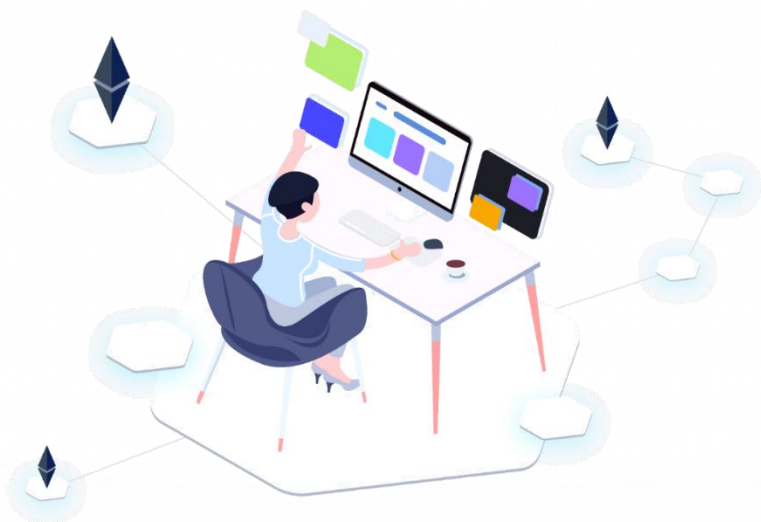
Pygal 生成的图表具有高度的互动性，用户可以通过鼠标悬停、点击等操作与图表进行交互，从而更深入地了解数据背后的信息。

定制化选项

Pygal 提供了丰富的定制化选项，用户可以根据自己的需求调整图表的颜色、样式、布局等，以创造出符合特定审美或功能要求的可视化作品。



>>> Pyecharts



◆ Pyecharts简介

Pyecharts是一个基于Python的数据可视化库，它为百度ECharts提供了Python接口，使得在Python环境中创建动态、交互式图表变得简单快捷。

◆ 支持的图表类型

Pyecharts支持多种图表类型，包括但不限于条形图、树状图、饼图和3D散点图等，这些图表类型能够满足不同场景下的数据统计和展示需求。

◆ 应用场景

Pyecharts广泛应用于数据统计和数学计算领域，通过其丰富的图表类型和强大的功能，可以帮助用户更直观地理解和分析数据。



谢谢!

主讲: XX

