

时序数据可视化

大数据可视分析导论

陈晴

<https://idvxlab.com>

同济大学

课程大纲

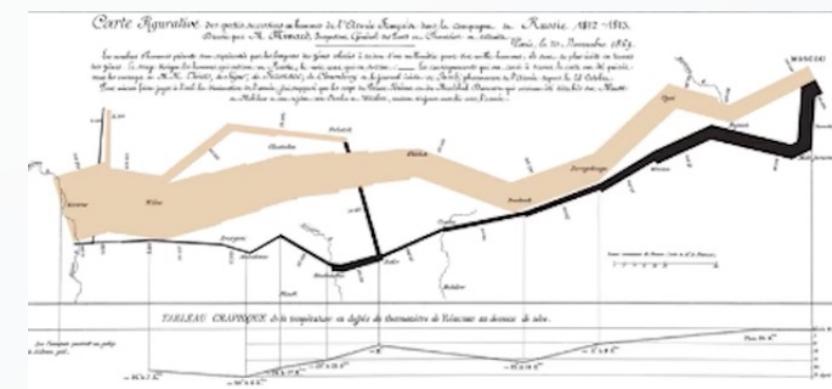
- 什么是时序数据？
- 时序数据可视化
- 多变量时序数据可视化
- 流数据可视化
- 时空数据可视化

课程大纲

- 什么是时序数据？
- 时序数据可视化
- 多变量时序数据可视化
- 流数据可视化
- 时空数据可视化

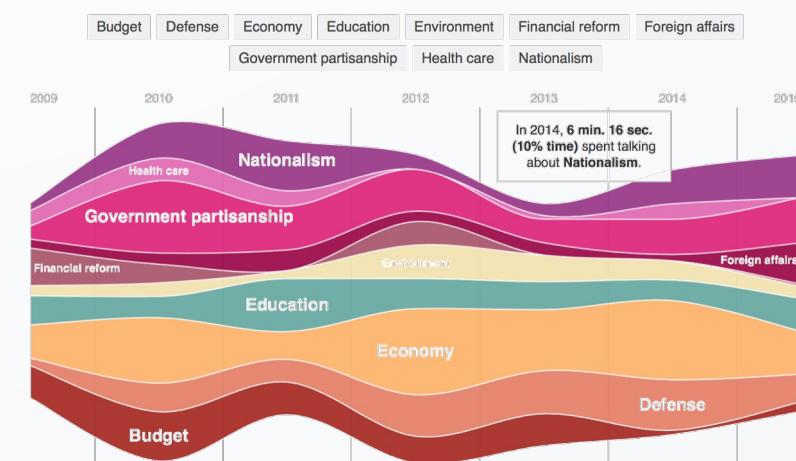
时序数据

- 时变数据 (time-series data; time-varying data)
 - 股票数据
 - 温度数据
- 序列数据 (sequential data)
 - 文本数据
 - 医学数据 (DNA, 蛋白质序列)



时序数据

- 时变数据 (time-series data; time-varying data)
 - 股票数据
 - 温度数据
- 序列数据 (sequential data)
 - 文本数据
 - 医学数据 (DNA, 蛋白质序列)



时序数据

- 时变数据 (time-series data; time-varying data)
 - 股票数据
 - 温度数据
- 序列数据 (sequential data)
 - 文本数据
 - 医学数据 (DNA, 蛋白质序列)
- 特点：量大、维数多、变量多、类型丰富、分布范围广

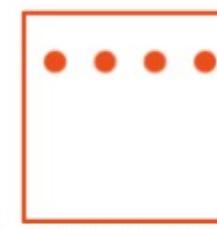
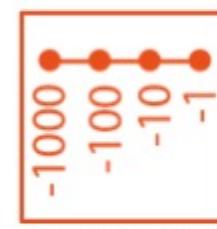
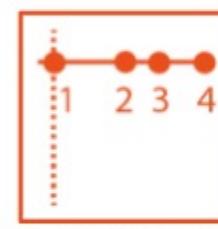
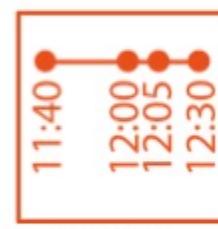
课程大纲

- 什么是时序数据？
- **时序数据可视化**
- 多变量时序数据可视化
- 流数据可视化
- 时空数据可视化

时序数据可视化的设计维度

表现形式

Representation



比例

Scale

布局

Layout



时序数据可视化的设计维度

表现形式

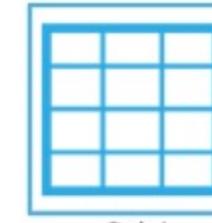
Representation



Linear



Radial



Grid



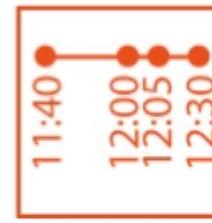
Spiral



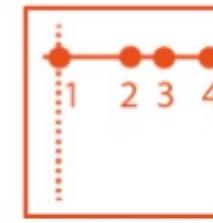
Arbitrary

比例

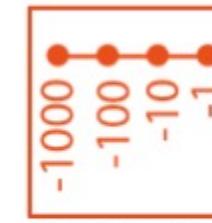
Scale



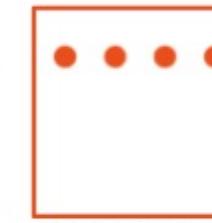
Chrono-
logical



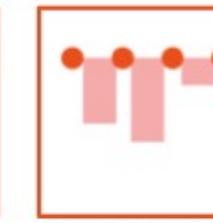
Relative



Logarithmic



Sequential



Sequential +
Interim
Duration

布局

Layout



Unified
(single
timeline)



Faceted
(multiple
timelines)



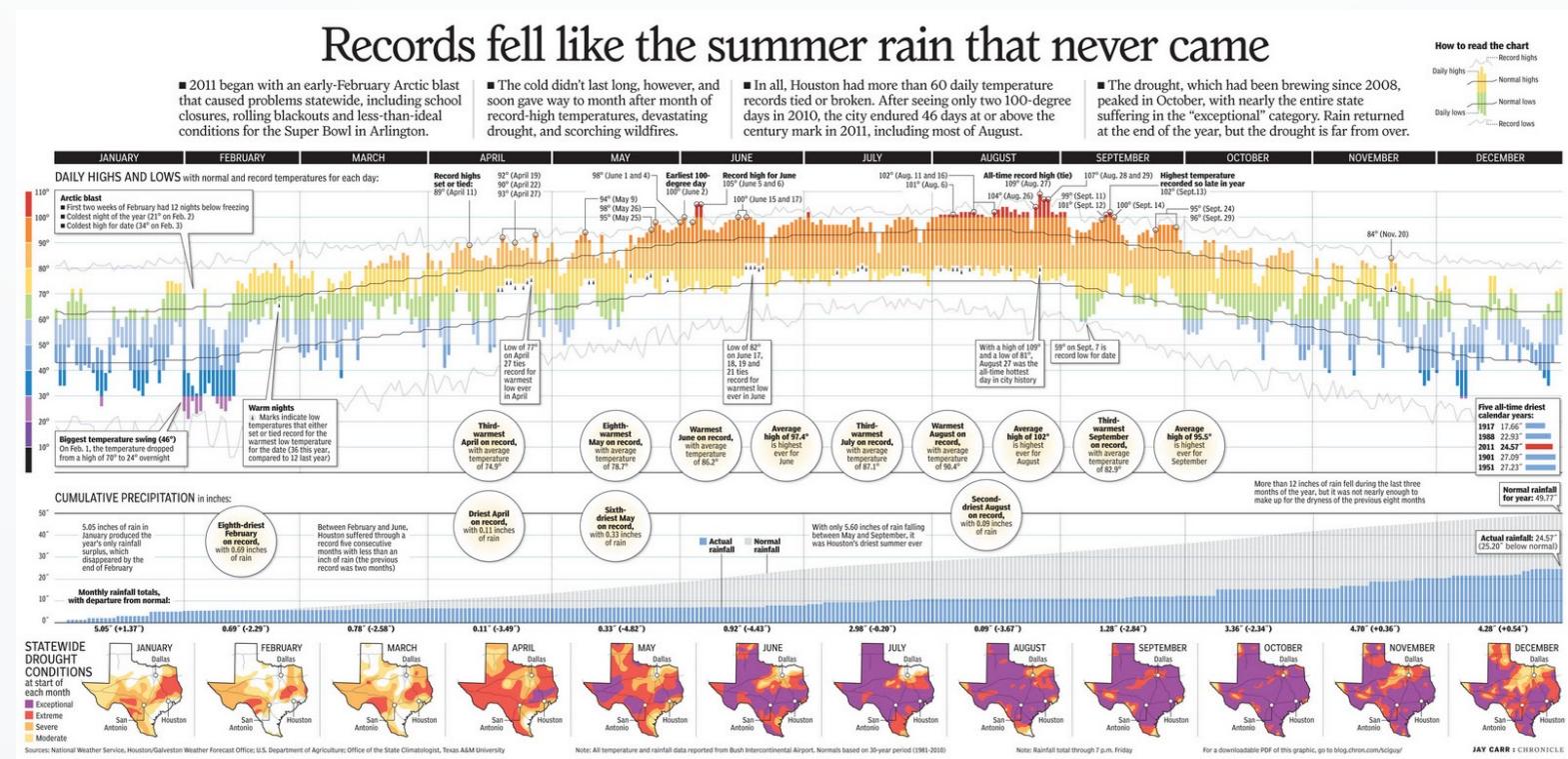
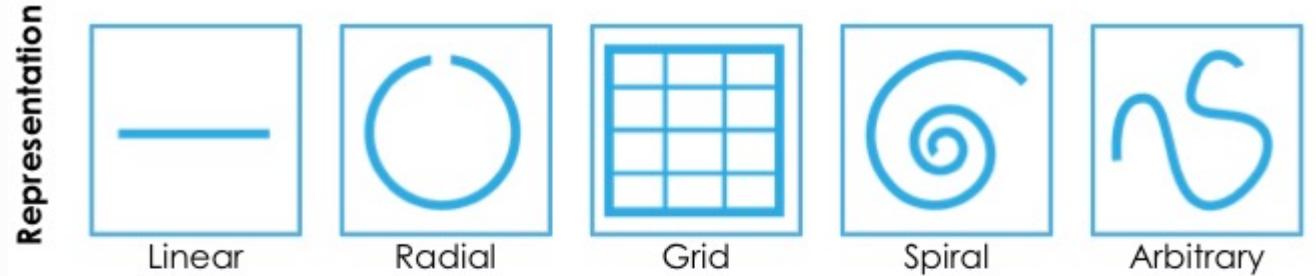
Segmented
timeline



Faceted +
Segmented

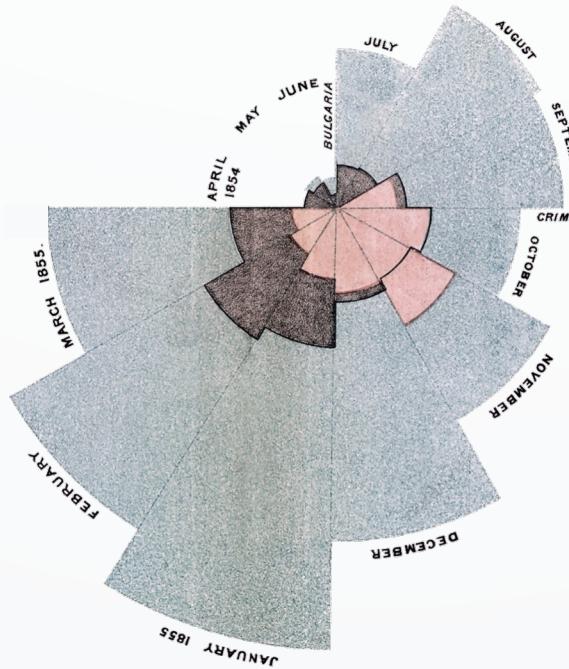
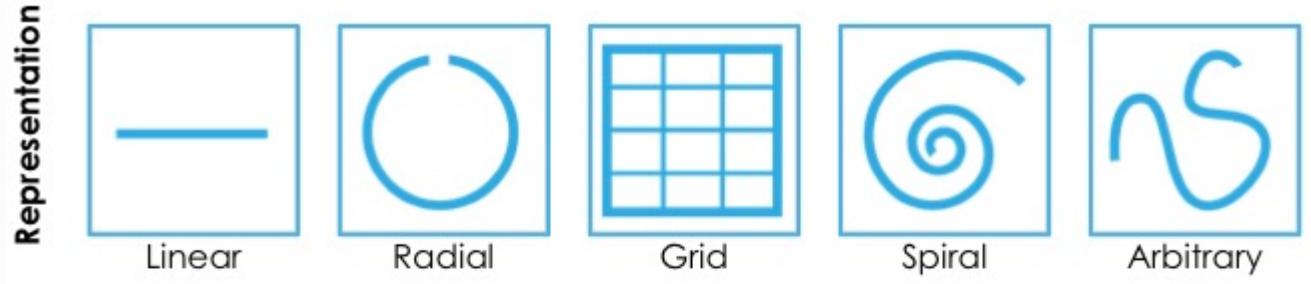
时序数据可视化

Linear 线性：将时间数据作为二维的线图表示，x轴表示时间，y轴表示其他变量。



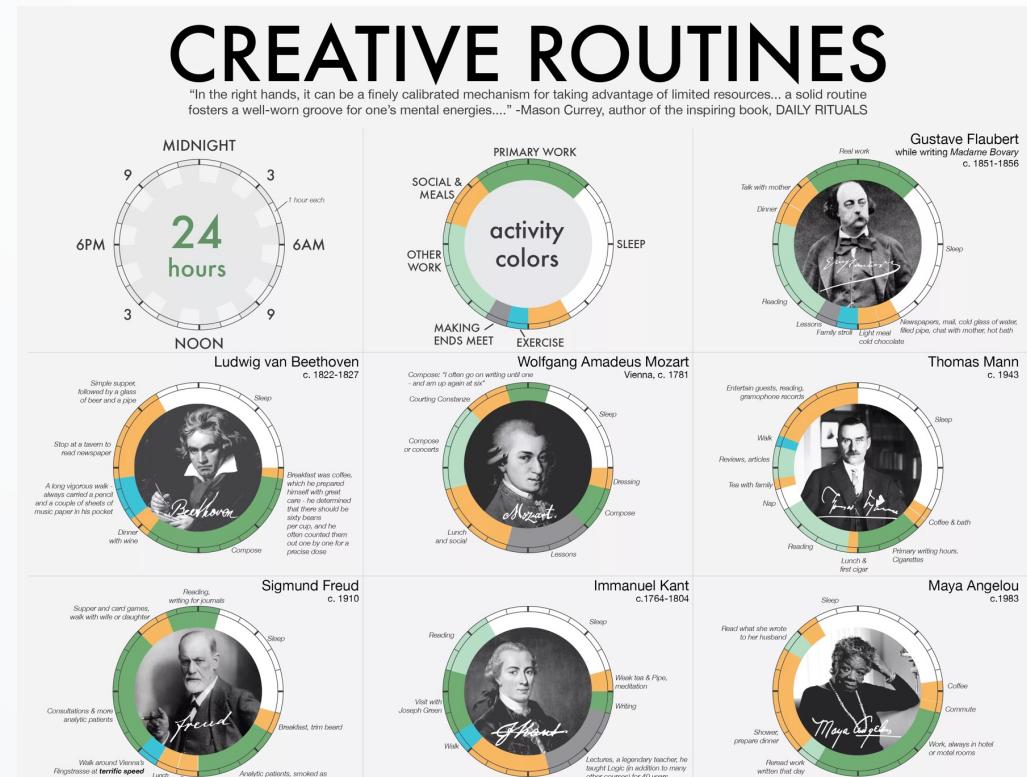
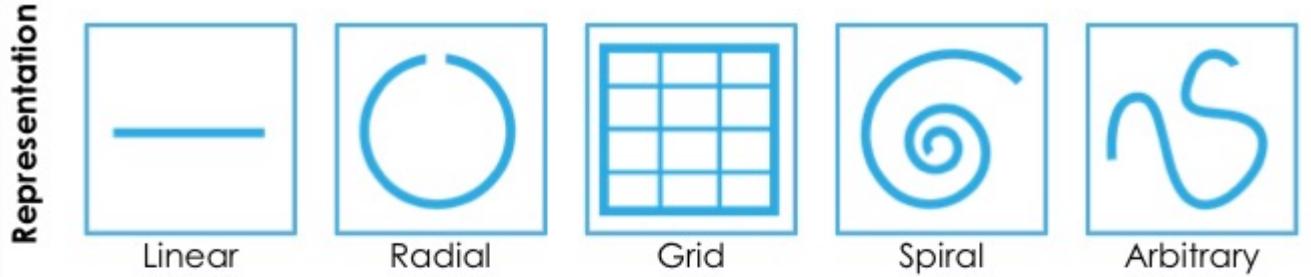
时序数据可视化

Radial 径向：沿圆周排列，将时间序列编码成弧形，合适呈现周期性的时变型数据。



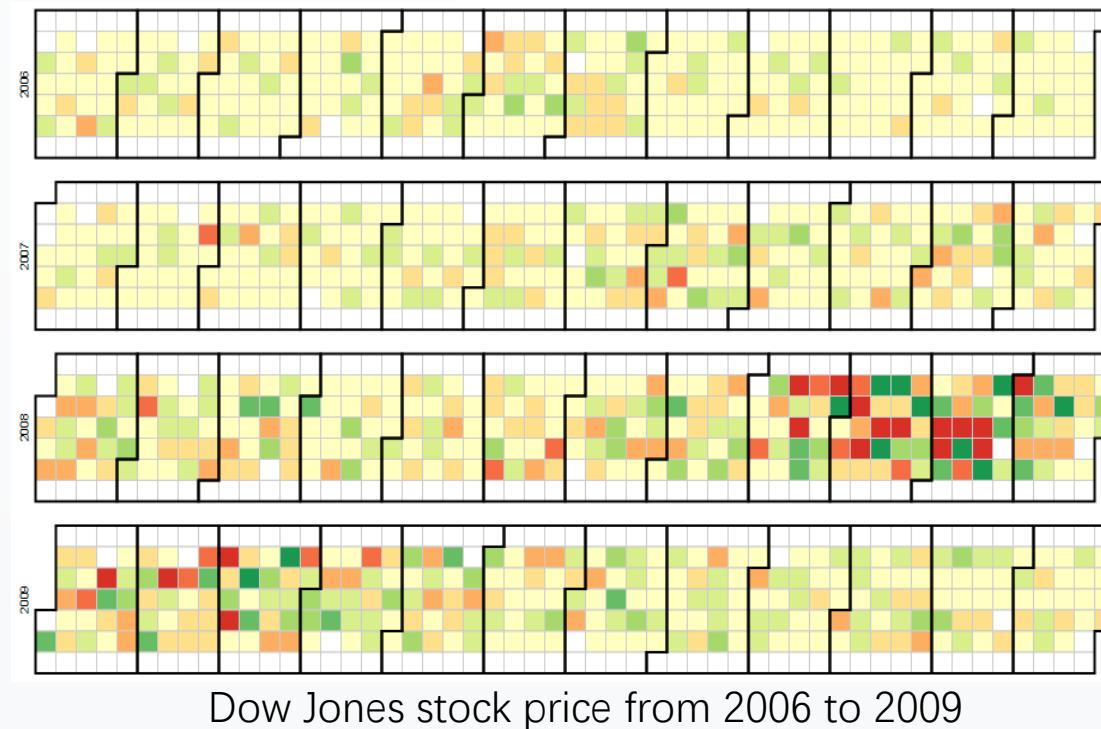
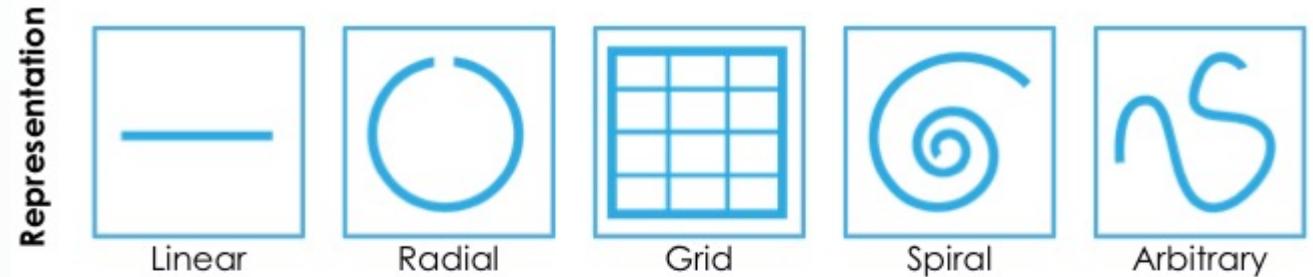
时序数据可视化

Radial 径向：沿圆周排列，将时间序列编码成弧形，合适呈现周期性的时变型数据。



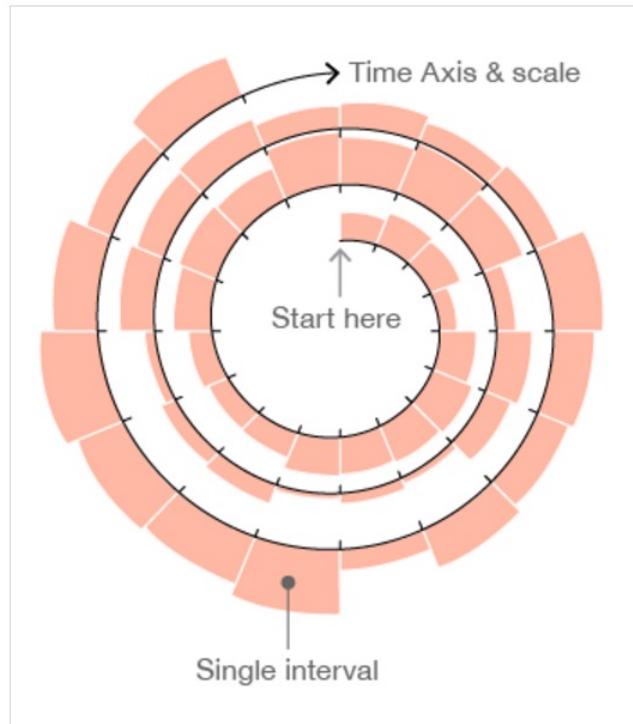
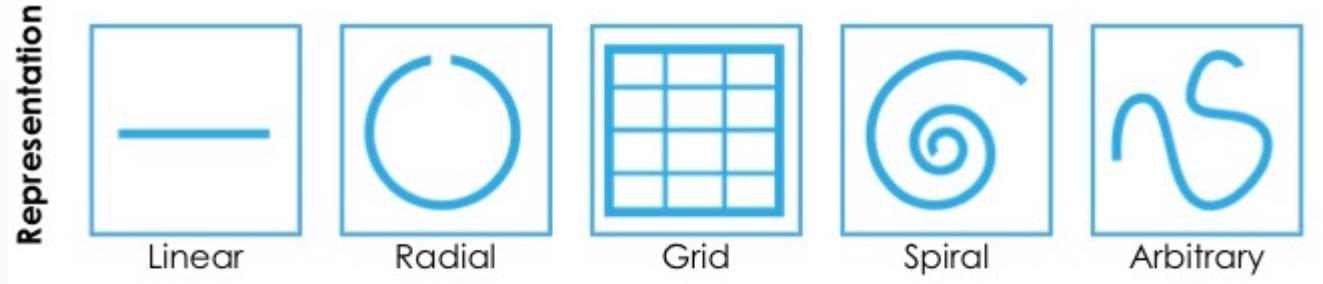
时序数据可视化

Grid 网格：和日历对应，一般采用表格映射的方式。



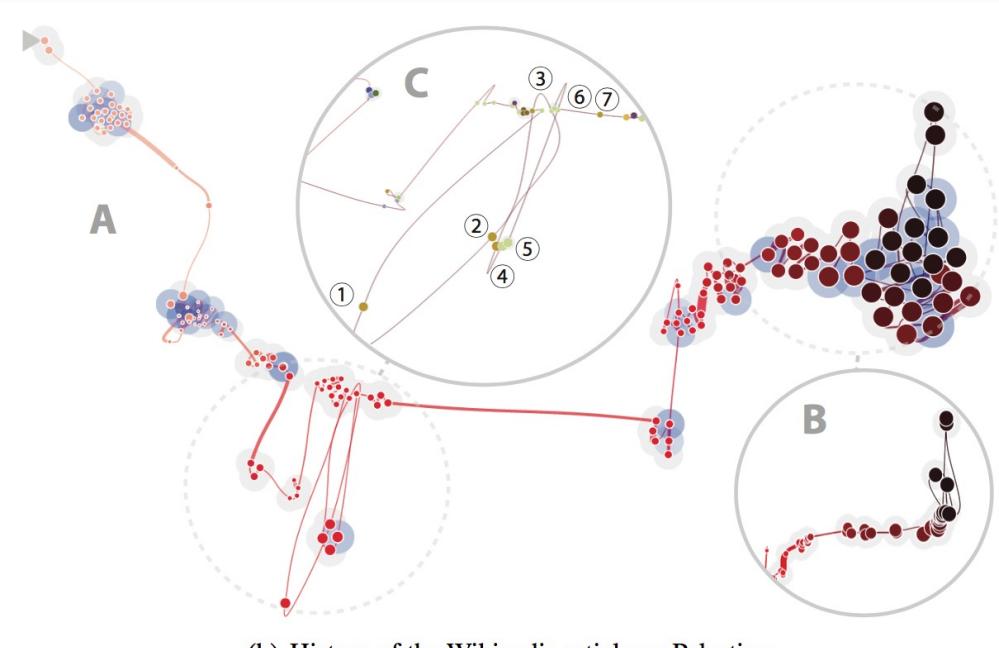
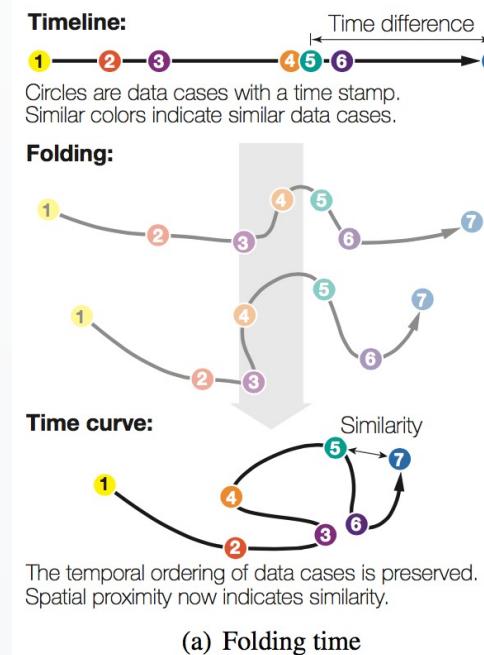
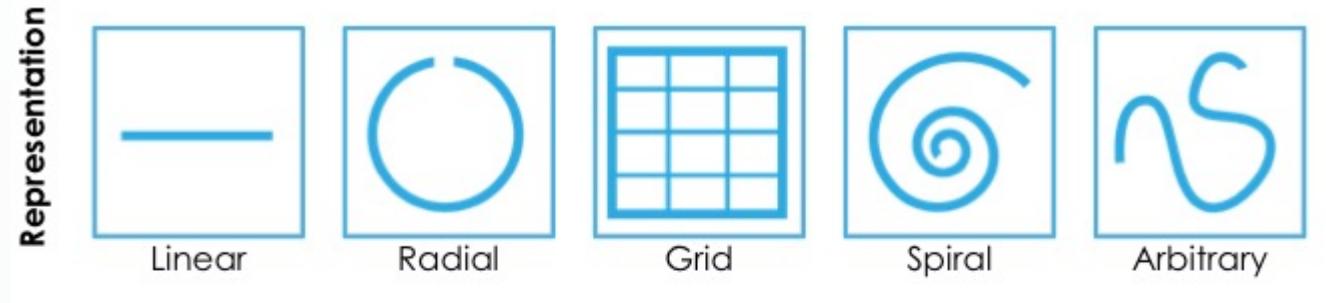
时序数据可视化

Spiral 螺旋：用条形、线条或数据点，沿着螺旋路径展示。



时序数据可视化

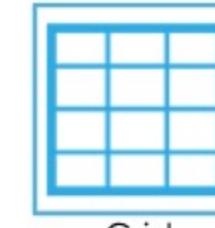
Arbitrary 随机：基于排版形式的随机、时间曲线等。



时序数据可视化的设计维度

表现形式

Representation



Linear

Radial

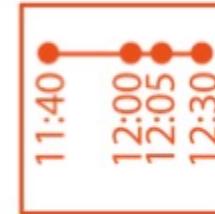
Grid

Spiral

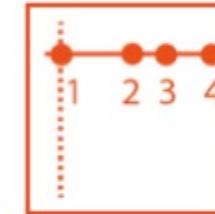
Arbitrary

比例

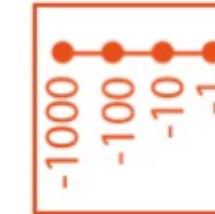
Scale



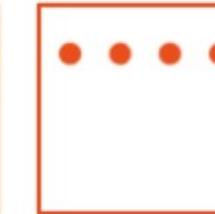
Chrono-
logical



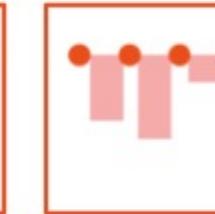
Relative



Logarithmic



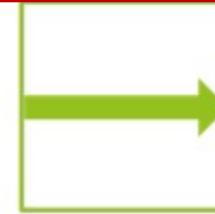
Sequential



Sequential +
Interim
Duration

布局

Layout



Unified
(single
timeline)



Faceted
(multiple
timelines)



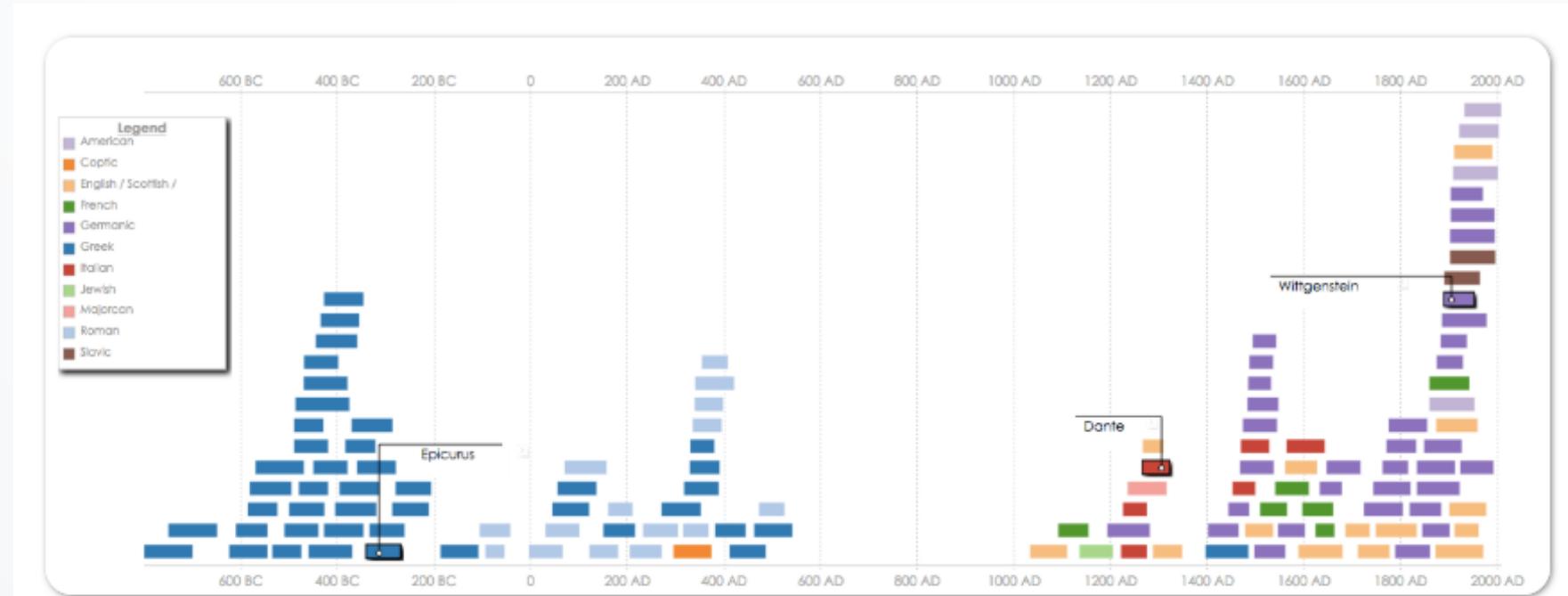
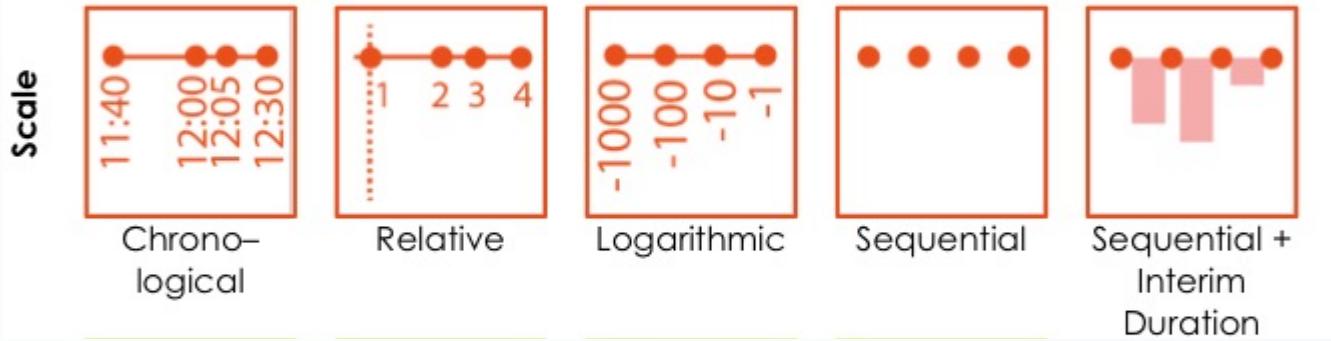
Segmented
timeline



Faceted +
Segmented

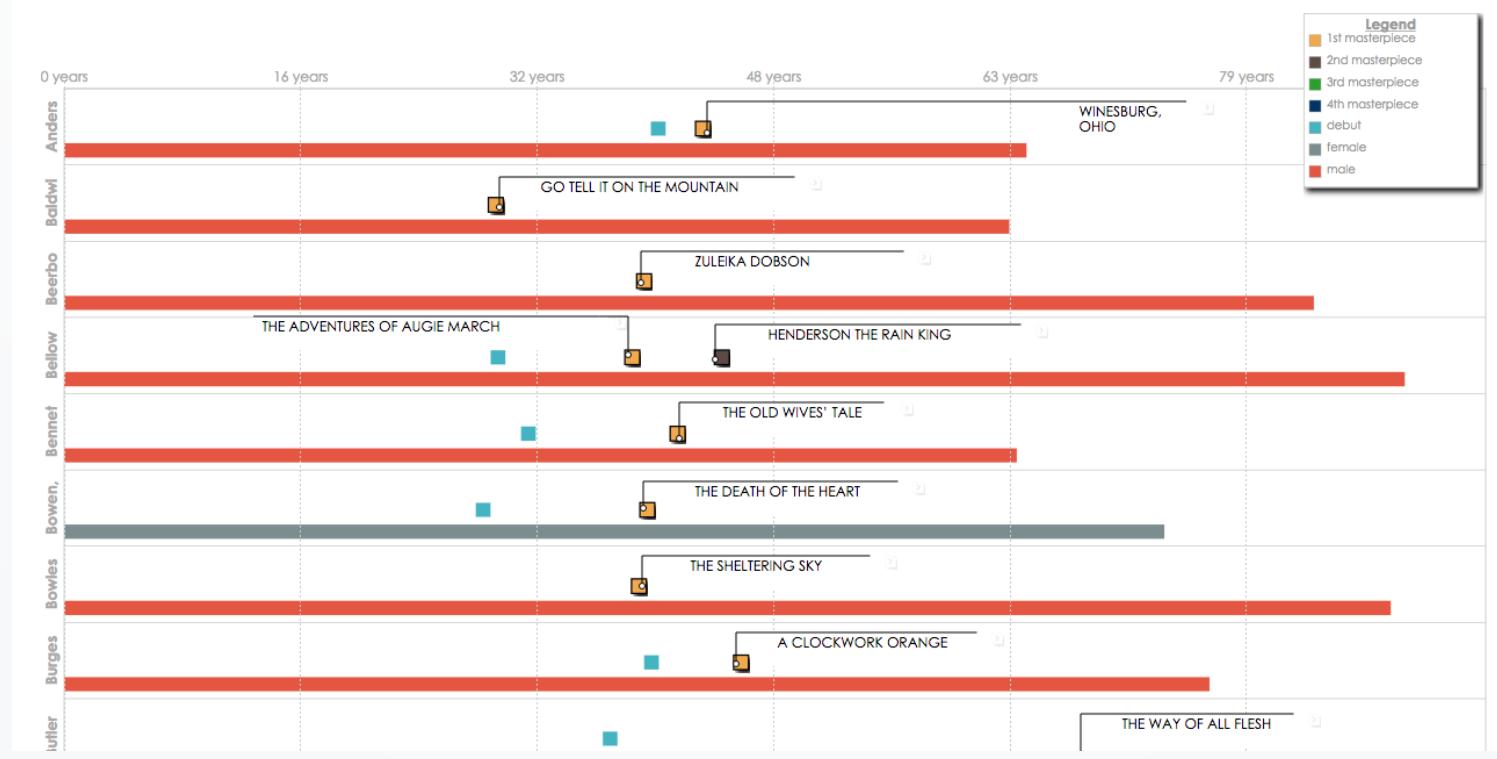
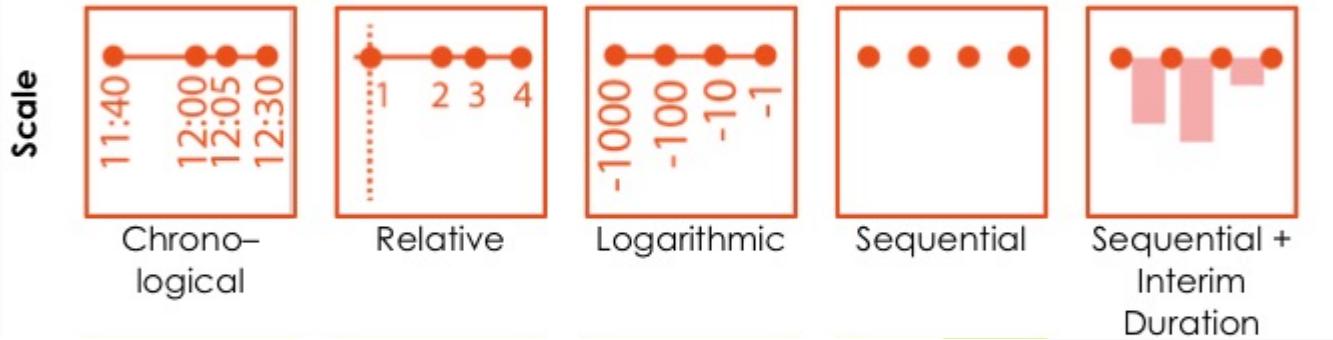
时序数据可视化

Chronological 按时间顺序：
可以被用来表示事件之间的
距离，事件的持续时间。



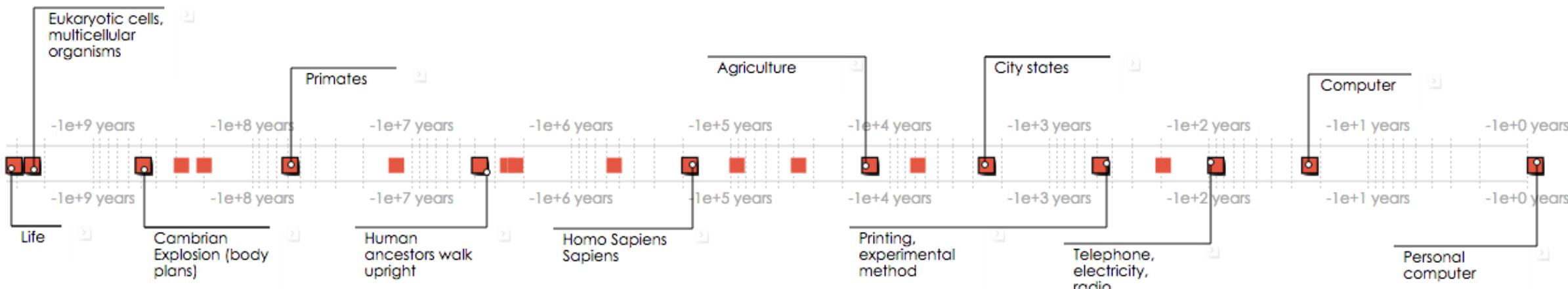
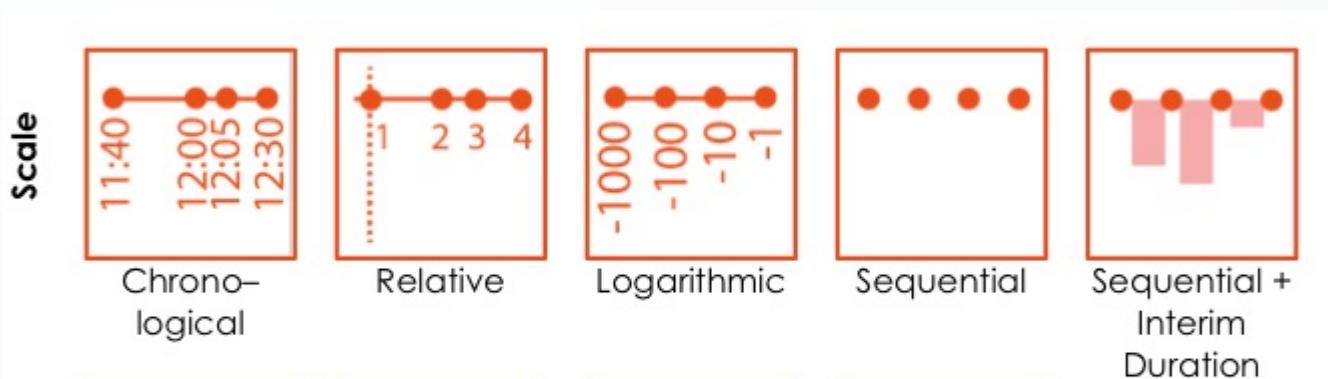
时序数据可视化

Relative 相对顺序：存在一个基线时间的时间零点，可以用在多时间线的对比。



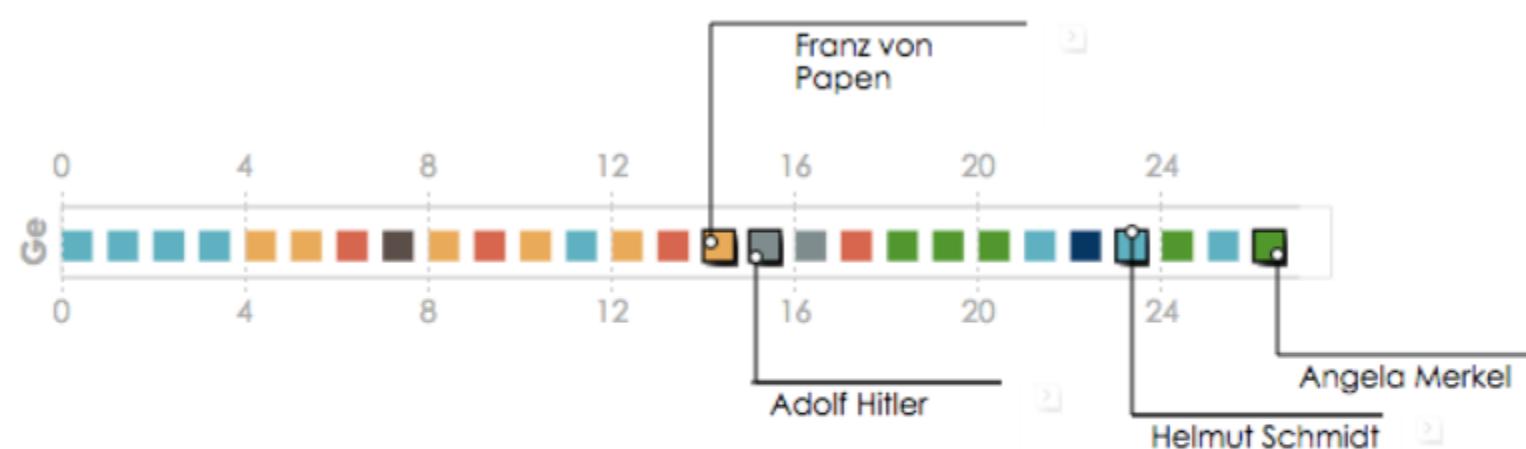
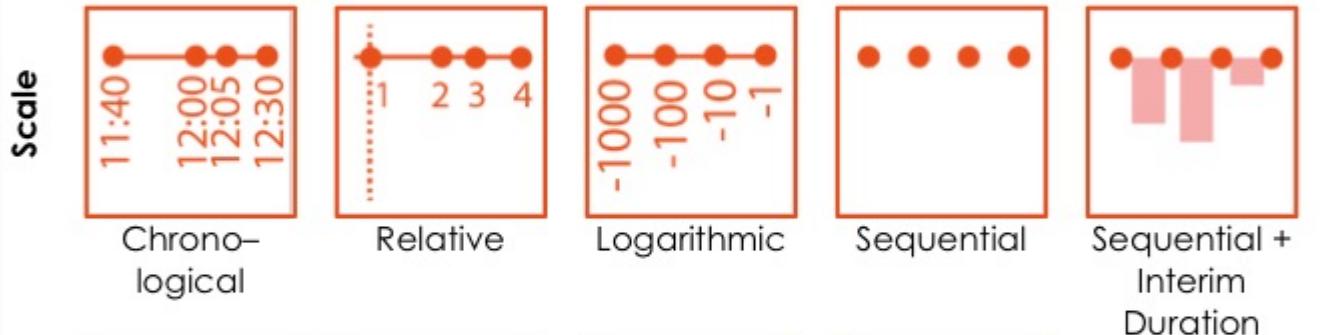
时序数据可视化

Logarithmic 对数比例：从时间先后顺序排列的比例转换而来，强调最早或最近的事件，适用于长范围或不均匀的事件布局。



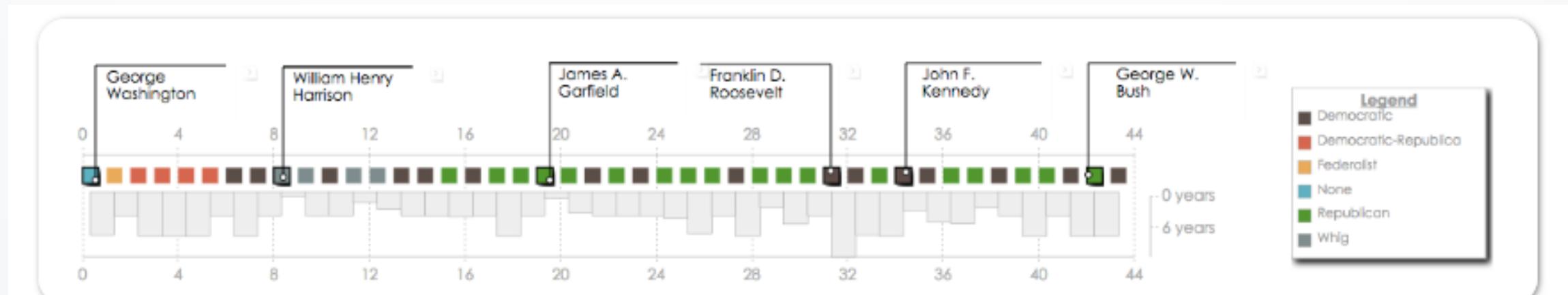
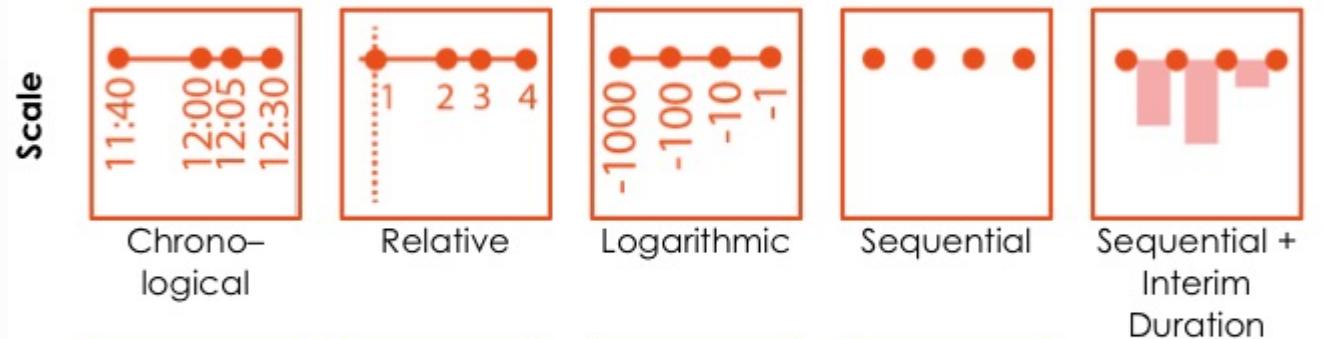
时序数据可视化

Sequential 次序：该比例中连续事件之间的距离是相等的，只表达事件的顺序。



时序数据可视化

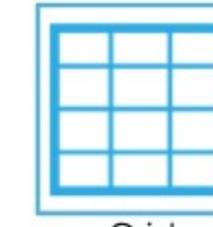
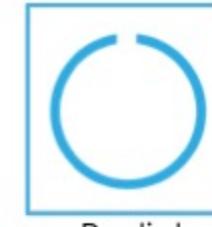
Sequential+Interim Duration
次序+中间时长：可以用来表达
长时间和不均匀分布的事件。



时序数据可视化的设计维度

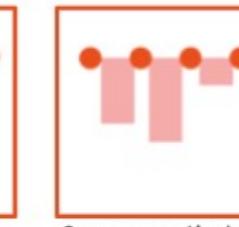
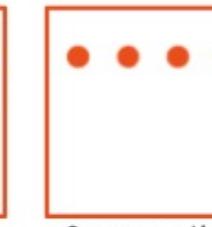
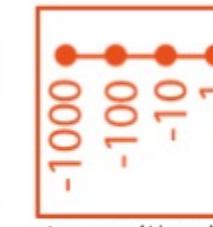
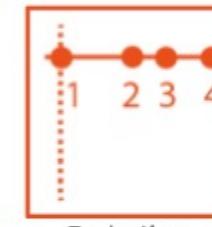
表现形式

Representation



比例

Scale



布局

Layout



课程大纲

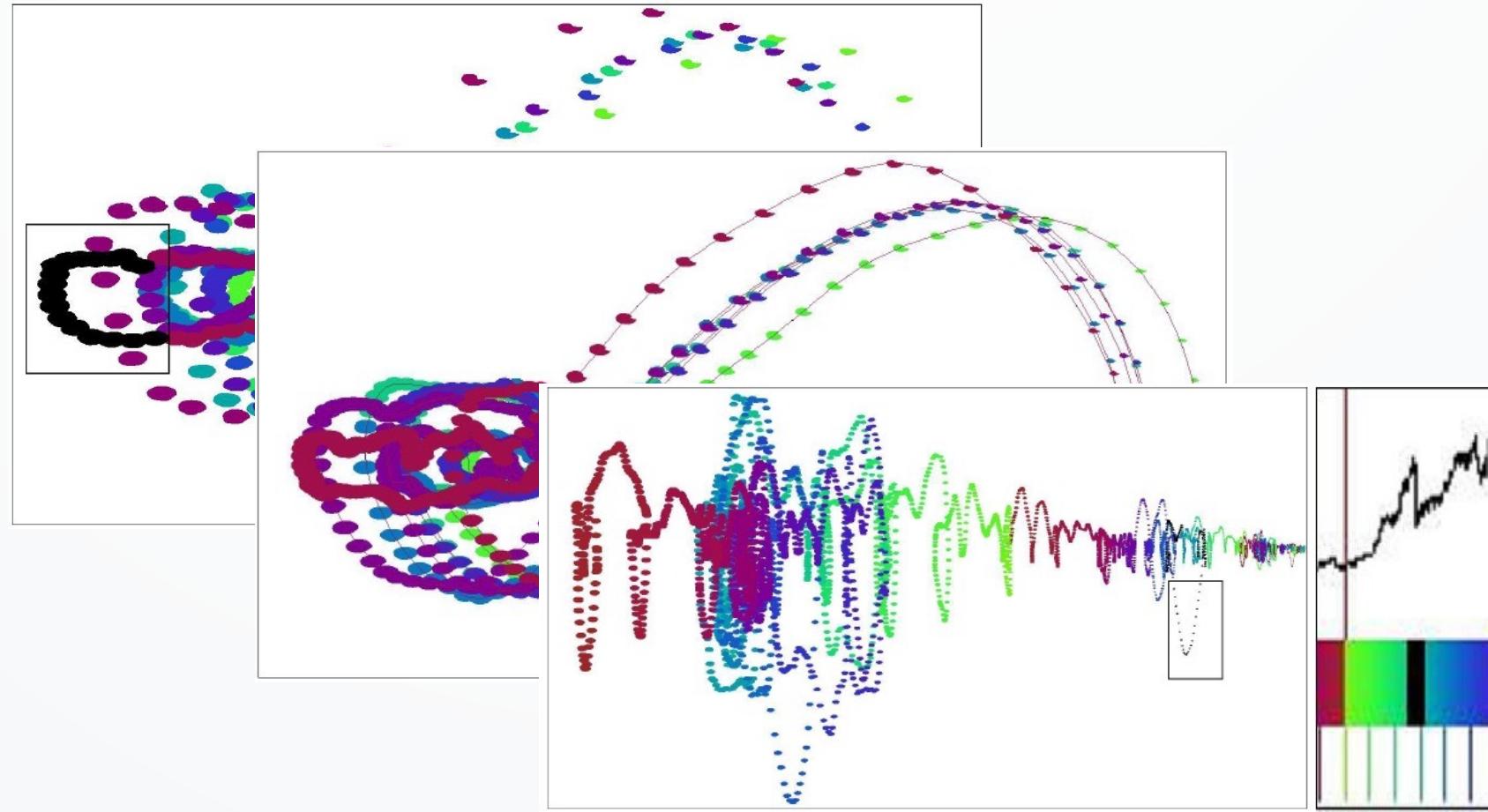
- 什么是时序数据？
- 时序数据可视化
- 多变量时序数据可视化
- 流数据可视化
- 时空数据可视化

多变量时序数据可视化

- **数据抽象 Data Abstraction**
 - 数据降维、特征选取和数据简化
- **数据聚类 Data Clustering**
 - 核心在于定义恰当的距离或相似性度量
- **特征分析 Feature Analysis**
 - 特征抽取、语义分析
 - 事件：事件定义、事件抽取、基于事件的语义分析

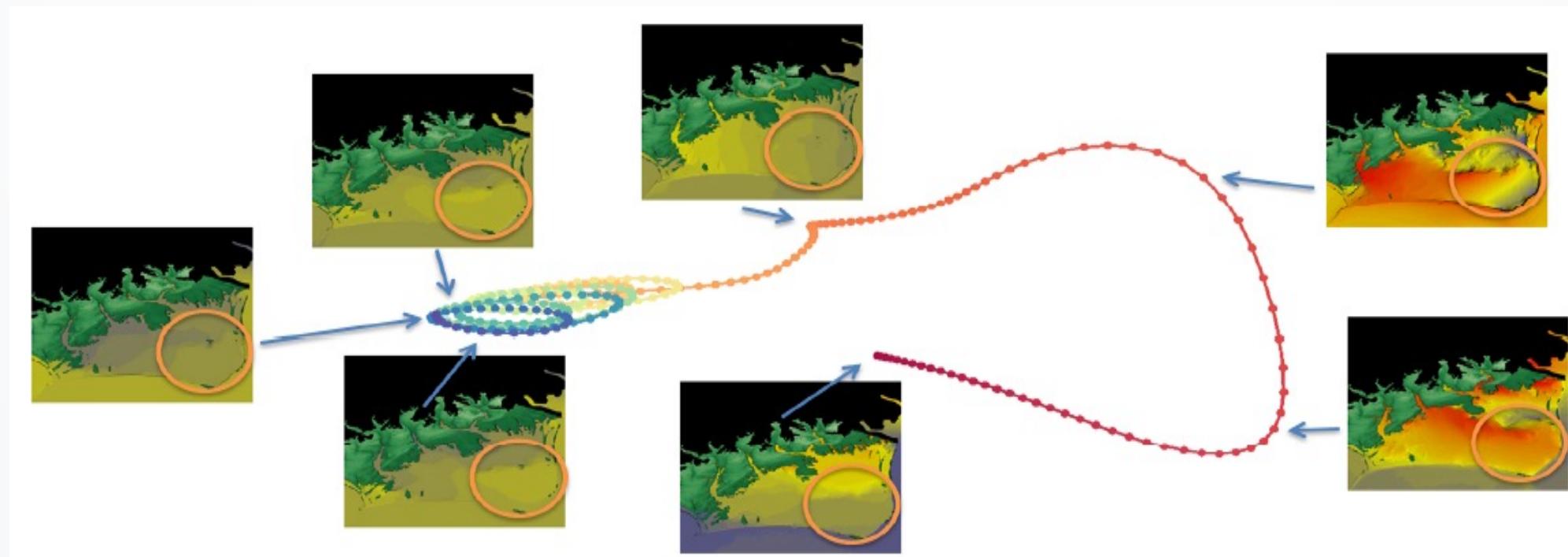
多变量时序数据可视化

- 基于线表示的可视化 Line-based Representation



多变量时序数据可视化

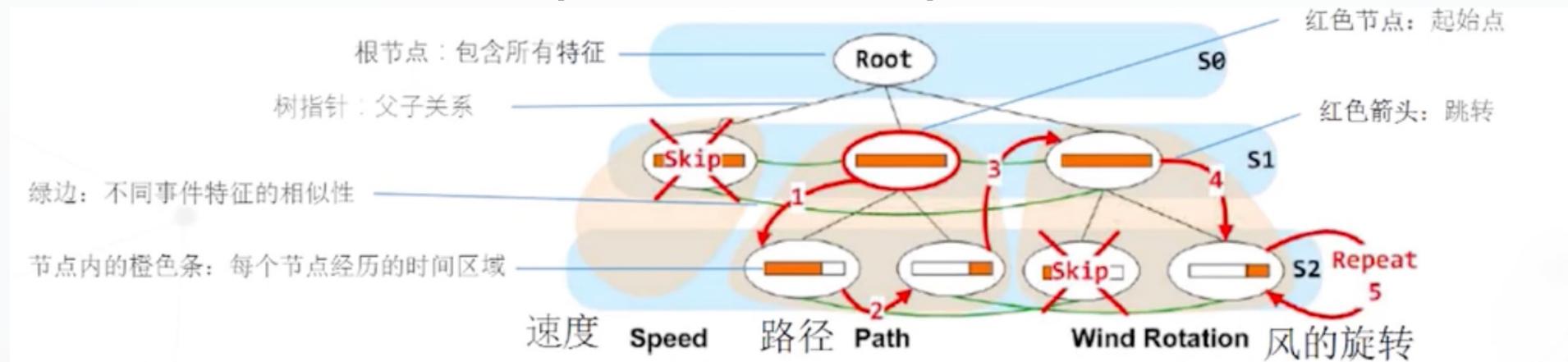
- 基于线表示的可视化 Line-based Representation



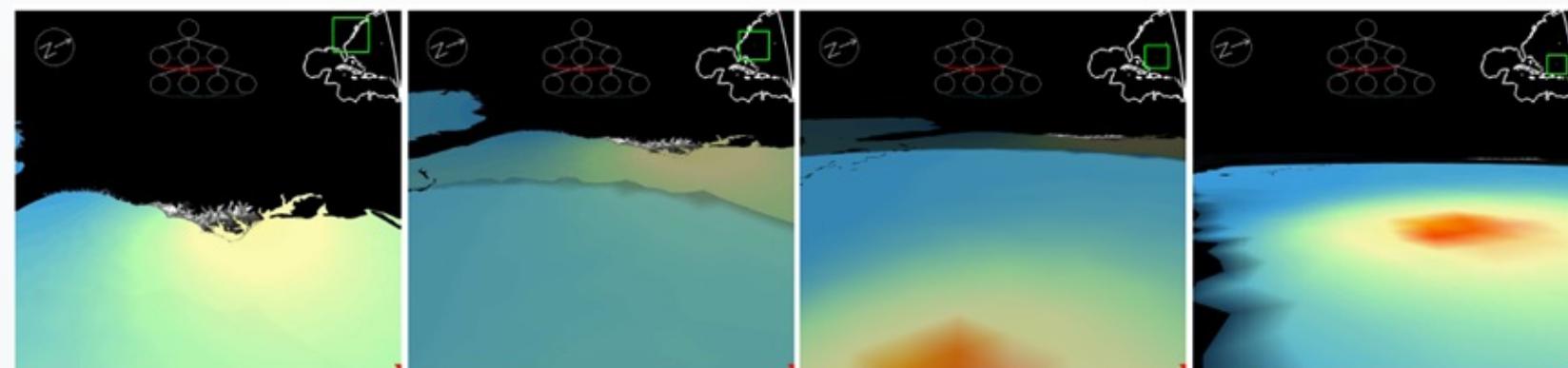
三维时变模拟飓风体数据集可视化。左边蓝色部分呈现了由于潮汐影响每12.4小时的周期演化；而右边从橙色变为深红的部分则揭示了海洋表面的巨大变化。

多变量时序数据可视化

• 基于图结构的可视化 Graph-based Representation



Interactive depth-first search 交互式深度优先探索

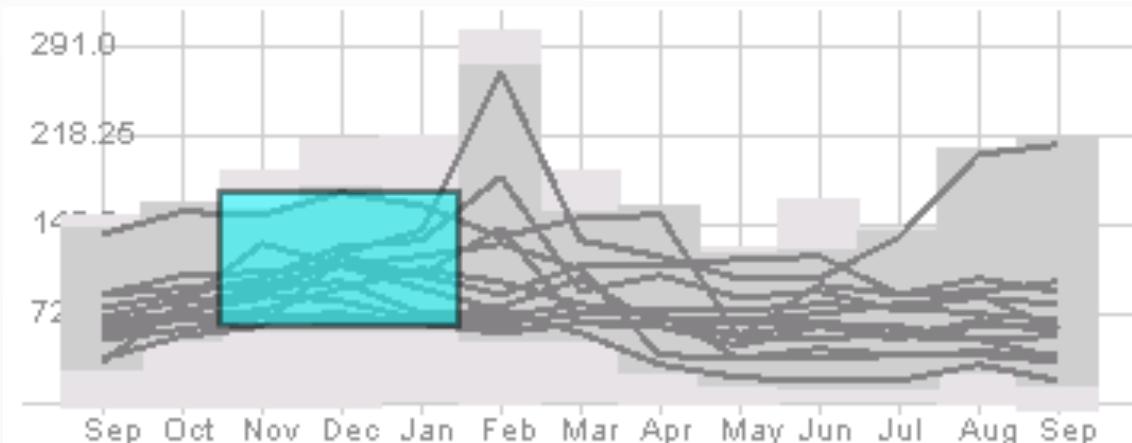
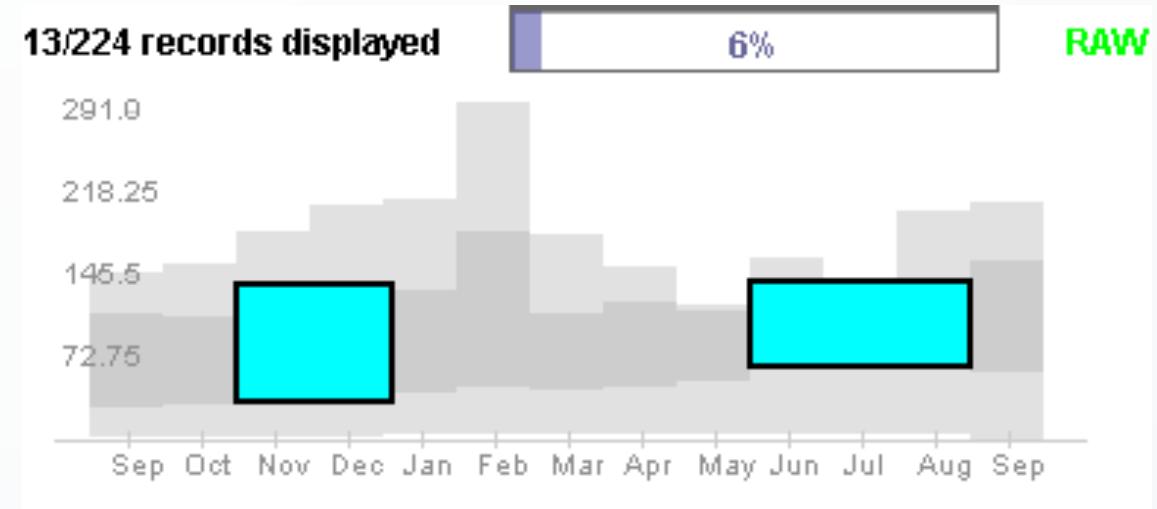


事件图的第一个层次上从风的旋转过渡到移动速度的效果演示

多变量时序数据可视化

• 交互 – 动态查询

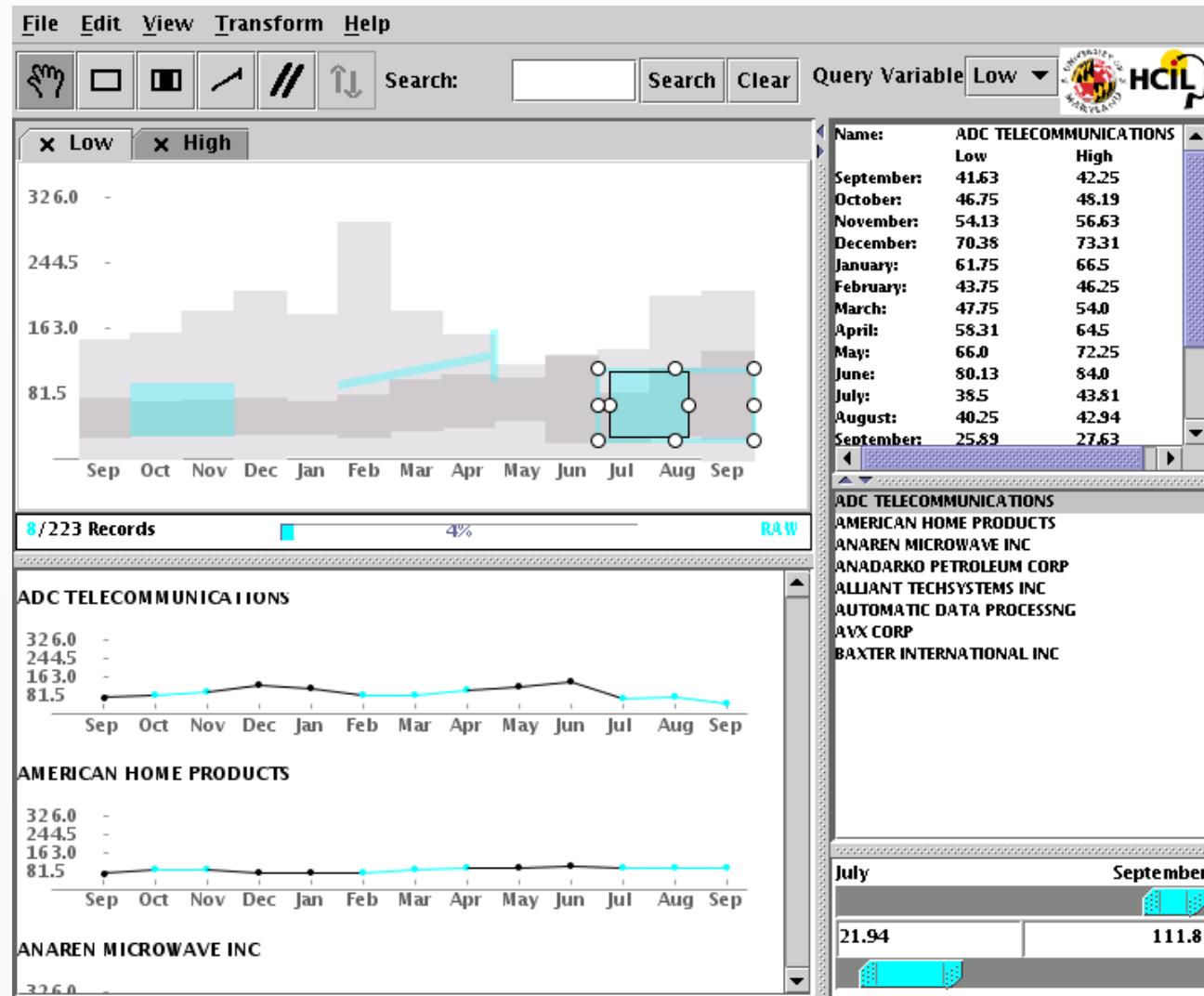
- 概览+上下文
- 避免视觉上的突变
- 用户交互要有直接的视觉反馈
- 在上下文中同时展现不同层次的信息



H. Hochheiser, B. Shneiderman. Dynamic Query Tools for Time Series Data Sets, Timebox Widgets for Interactive Exploration, Journal of Information Visualization 3(1), pp.1-18, 2004

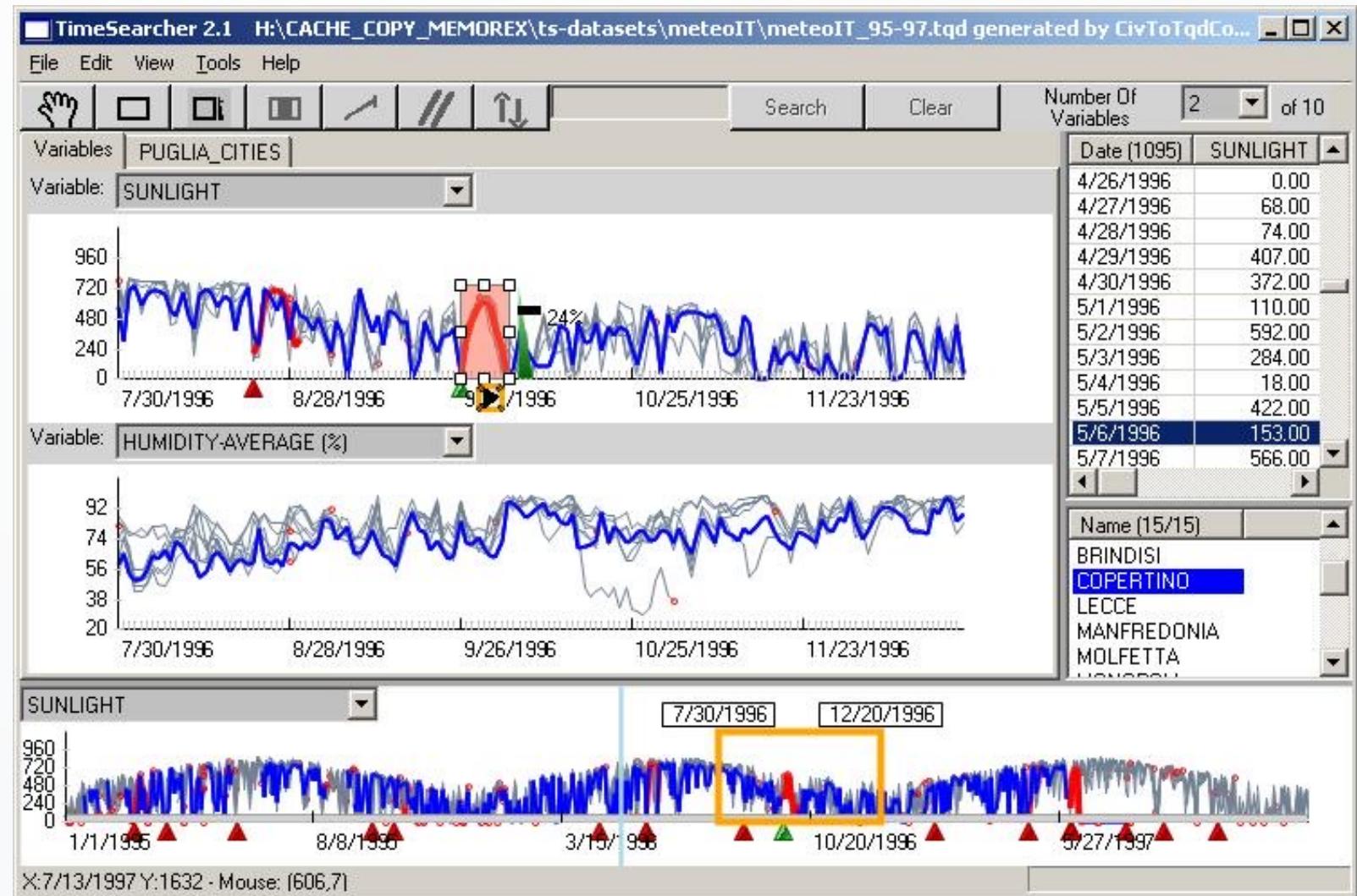
- 交互 - 动态查询
- TimeSearcher1.0

多变量时序数据可视化



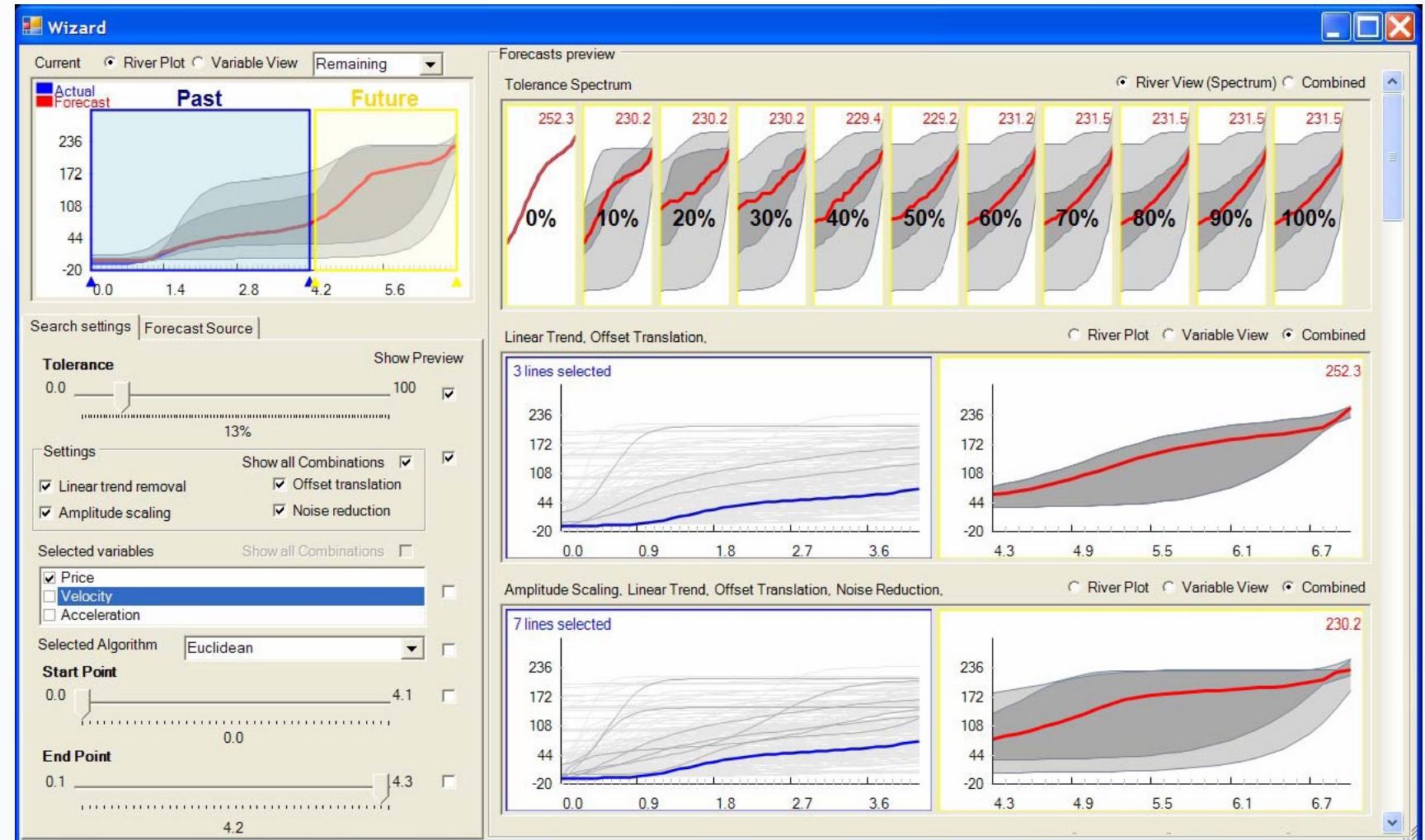
多变量时序数据可视化

- 交互 - 动态查询
- TimeSearcher2.0



多变量时序数据可视化

- 交互 - 动态查询
- TimeSearcher3.0



课程大纲

- 什么是时序数据？
- 时序数据可视化
- 多变量时序数据可视化
- 流数据可视化
- 时空数据可视化

流数据可视化

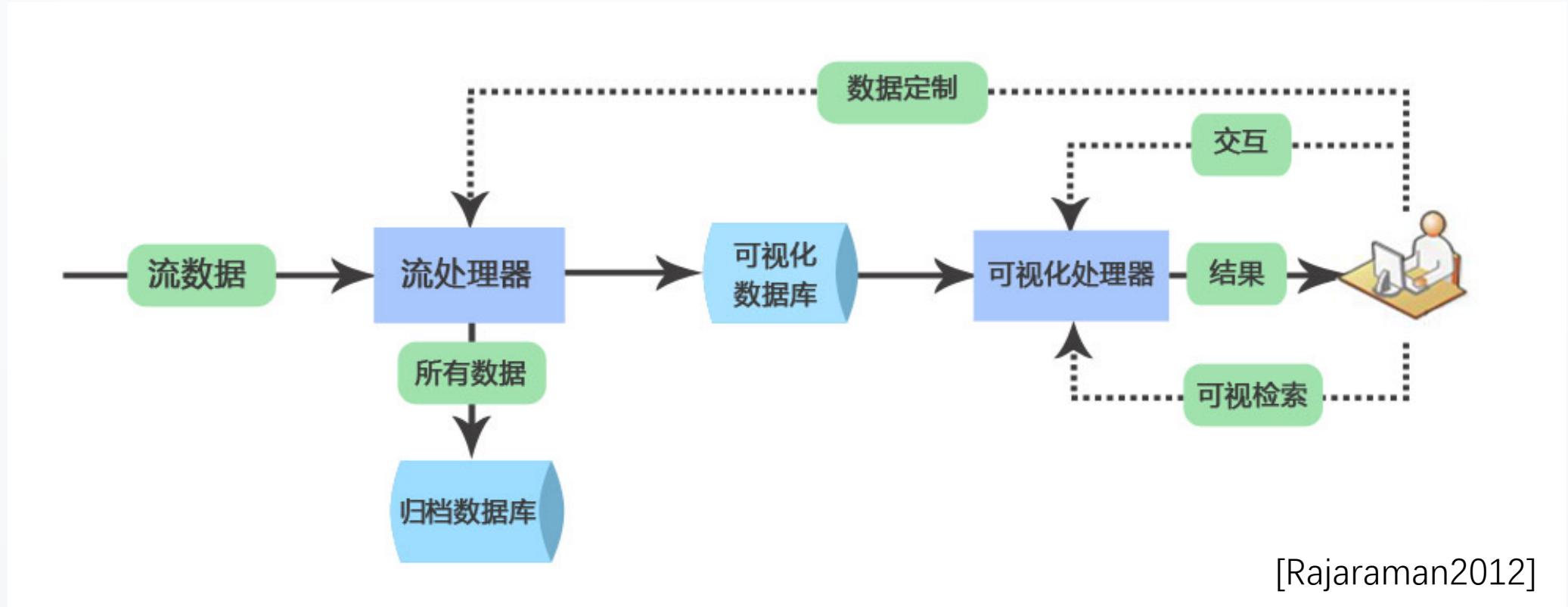
- 流数据 Streaming Data
- 动态实时更新的时序数据
 - 移动日志、网络数据包、高性能集群日志、传感器日志、金融市场数据(股票数据)、社交网络数据
- 思考：相应可视化上会有什么区别？
 - 哪些可视化可以直接用？
 - 需要设计哪些新的可视化？

流数据可视化

- 流数据与普通时序数据的异同
- 数据流的**潜在数量**也许是无限的。
- 数据元素在线到达，需要实时处理，否则数据的价值随时间的流逝可能降低。
- 无法控制数据元素的**到达顺序和数量**，每次流入的数据顺序都可能不一致，数量时多时少。
- 某个元素被处理后，要么被丢弃，要么被归档存储。
- 对于流数据的查询异常情况和相似类型比较耗时，人工检测日志相当乏味且易出错。

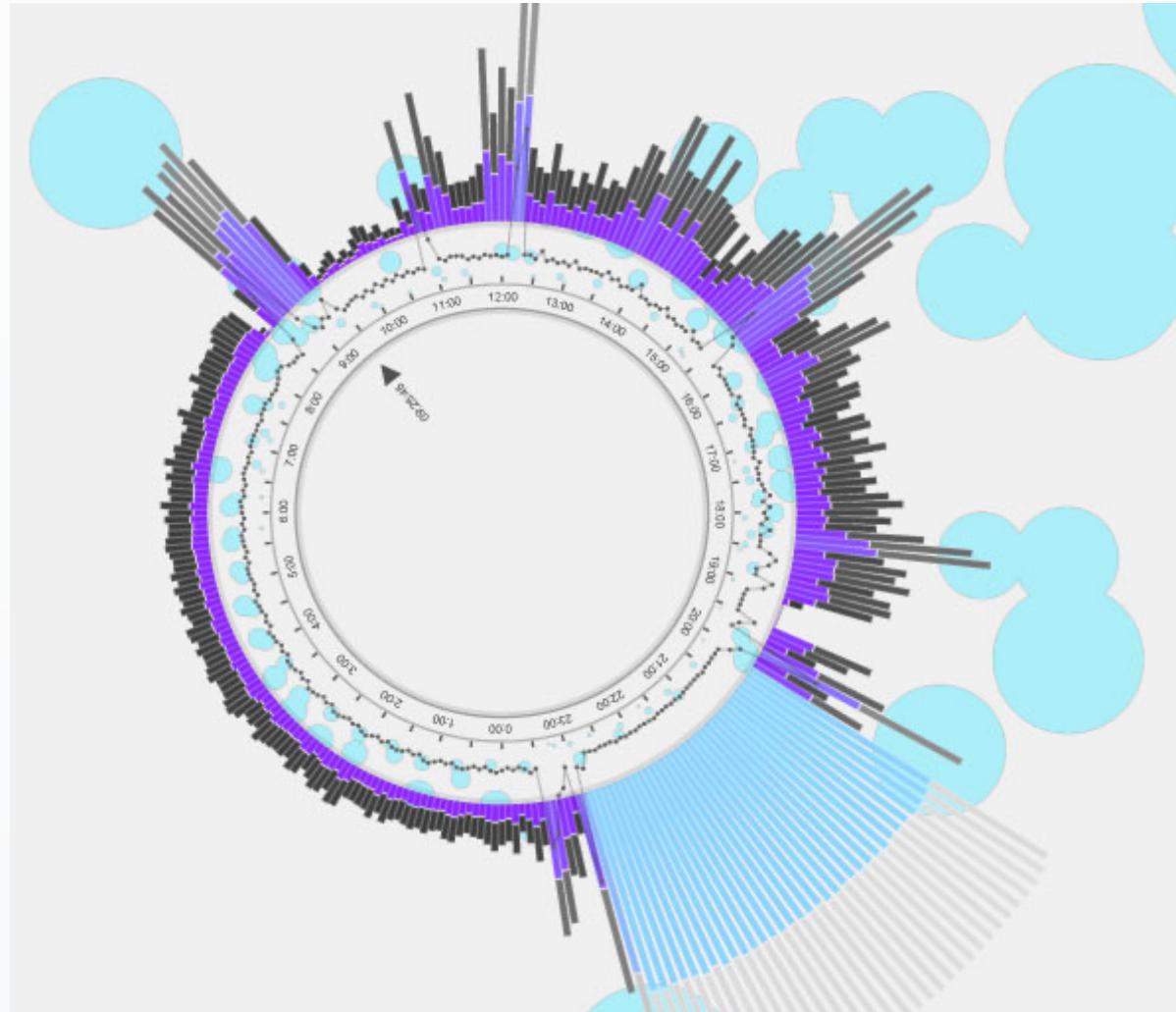
流数据可视化

- 流数据可视化流程



流数据可视化

- 流数据可视化案例
 - 网络日志数据
 - 线上用户行为可视分析工具



课程大纲

- 什么是时序数据？
- 时序数据可视化
- 多变量时序数据可视化
- 流数据可视化
- **时空数据可视化**

时空数据可视化

- 时空数据同时包含时间变量和位置变量。
- 时空的对象包含不同的语义信息。
- 通常用动画或3D显示来表示。

时空数据可视化

- **空间-时间序列 Spatial time series:**
- **空间事件数据 Spatial event data:** 描述空间事件 spatial events (i.e., discrete objects)
- **移动数据 Movement data (trajectories):** 描述移动物体 moving objects

时空数据可视化

- **空间-时间序列 Spatial time series:**
不同地点在不同时间上的属性变化 - space and time are references
- **空间事件数据 Spatial event data:** 描述空间事件 spatial events (i.e., discrete objects)
- **移动数据 Movement data (trajectories):** 描述移动物体 moving objects

时空数据可视化

- 空间-时间序列 **Spatial time series**:
不同地点在不同时间上的属性变化 - space and time are references
- 空间事件数据 **Spatial event data**: 描述空间事件 spatial events (i.e., discrete objects)
不同物体对象在不同时空上的属性 : space and time are attributes
- 移动数据 **Movement data (trajectories)**: 描述移动物体 moving objects

时空数据可视化

- **空间-时间序列 Spatial time series:**
不同地点在不同时间上的属性变化 - space and time are references
- **空间事件数据 Spatial event data:** 描述空间事件 spatial events (i.e., discrete objects)
不同物体对象在不同时空上的属性 : space and time are attributes
- **移动数据 Movement data (trajectories):** 描述移动物体 moving objects
空间是移动物体的属性，时间是reference，记录了对象在不同时刻的位置，
(object identifier, time, location, <thematic attributes>)

时空数据可视化

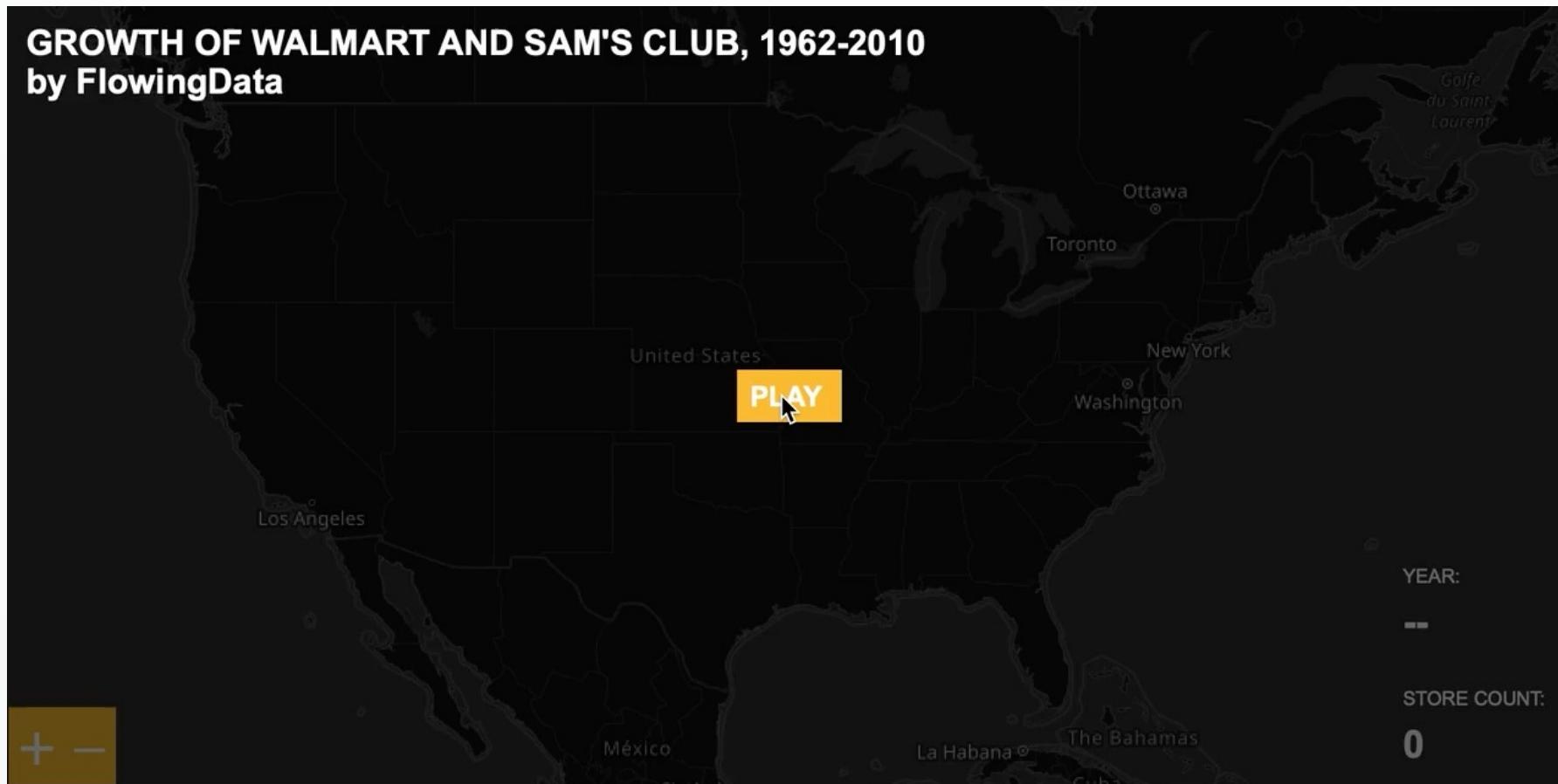
- LA Traffic Radio



<http://www.youtube.com/watch?v=72-blGAuh6s>

时空数据可视化

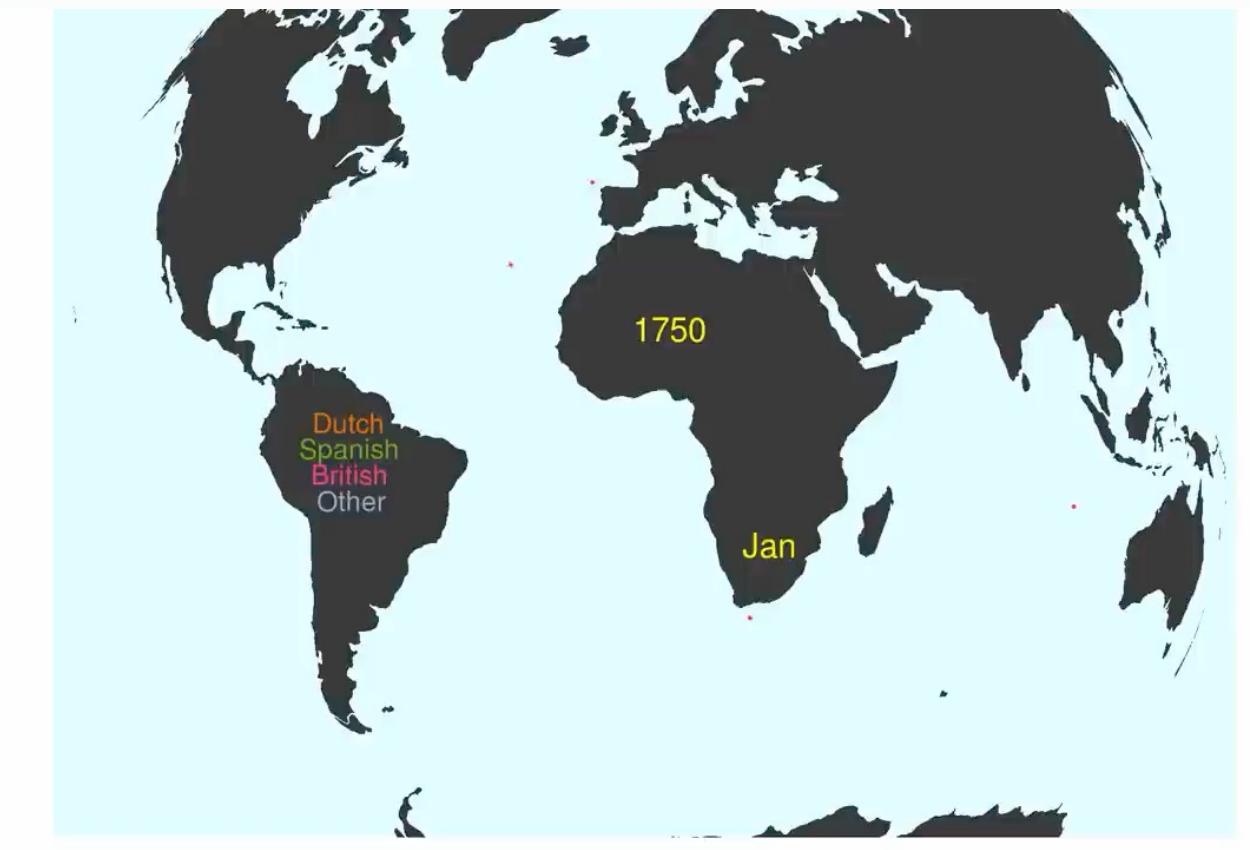
- Growth of Walmart



<http://projects.flowingdata.com/walmart/>

时空数据可视化

- Global Ocean Shipping History



时变数据可视化技术汇总

The TimeViz Browser
A Visual Survey of Visualization Techniques for Time-Oriented Data
by Christian Tominski and Wolfgang Aigner

of Techniques: 115

Search:

How to use:
Want - I want to see.
? - I'm neutral.
Hide - Don't show me.

Data

Frame of Reference
 Abstract Want ? Hide
 Spatial Want ? Hide

Number of Variables
 Univariate Want ? Hide
 Multivariate Want ? Hide

Time

Arrangement
 Linear Want ? Hide
 Cyclic Want ? Hide

Time Primitives
 Instant Want ? Hide
 Interval Want ? Hide

Visualization

Mapping
 Static Want ? Hide
 Dynamic Want ? Hide

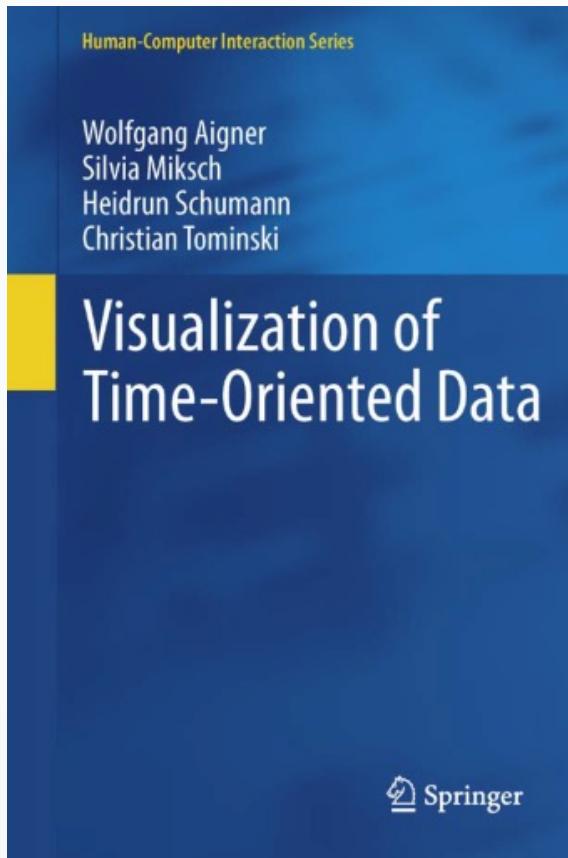
Dimensionality
 2D Want ? Hide



事件序列数据可视化技术汇总

	Plaisant et al. 1996	Wongsuphasawat et al. 2001	Cao et al. 2011	Wongsuphasawat et al. 2011	Golz et al. 2014	Pere et al. 2014	Kwon et al. 2016	Wang et al. 2016	Cappers et al. 2017	Guo et al. 2017	Chen et al. 2017	Law et al. 2018	Chen et al. 2018	Kwon et al. 2020	Nguyen et al. 2020	Du et al. 2016	Krause et al. 2016	Du et al. 2017	Jin et al. 2018	Kwon et al. 2018	Wongsuphasawat et al. 2009	Malik et al. 2015	Zhao et al. 2015	Qi et al. 2019	Guo et al. 2020	Fischer et al. 2012	Zhao et al. 2014	Xu et al. 2016	Nguyen et al. 2018	Guo et al. 2019	Mu et al. 2019	Jin et al. 2021	Aijun et al. 2021	Sum
	Summarization								Prediction & Recommendation				Comparison			Anomaly Detection			Causality															
DS-Single event																													5					
DS-Subsequence																													4					
DS-Sequence																													9					
DS-Sequence corpus																													19					
ASA-Pattern discovery																													14					
ASA-Prediction																													6					
ASA-Inference																													7					
VR-Chart-based																													14					
VR-Timeline-based																													19					
VR-Sankey-based																													10					
VR-Hierarchy-based																													6					
VR-Matrix-based																													4					
IT-Segment																													2					
IT-Emphasis																													30					
IT-Scaling																													21					
IT-Filter/Query																													33					
IT-Aggregation																													8					
IT-Alignment																													7					
IT-Editing																													6					

推荐书籍



Visualization of Time-Oriented Data

Aigner, W., Miksch, S., Schumann, H., Tominski, C. Springer

课程大纲

- 什么是时序数据？
- 时序数据可视化
- 多变量时序数据可视化
- 流数据可视化
- 时空数据可视化