A Comparative Evaluation Of Visual Summarization Techniques For Event Sequences

Kazi Tasnim Zinat, Jinhua Yang, Arjun Gandhi, Nistha Mitra, **Zhicheng Liu**

University of Maryland, College Park

Event Sequence Data

EUROVIS Leipzig 2023

- Connected by a common entity
- Ordered temporally to form sequences
- Categorical event information along with temporal data

Example: Pediatric Trauma Unit

EUROVIS Leipzig 2023

Real-world example using EventFlow [CBM*13, MMPS13]

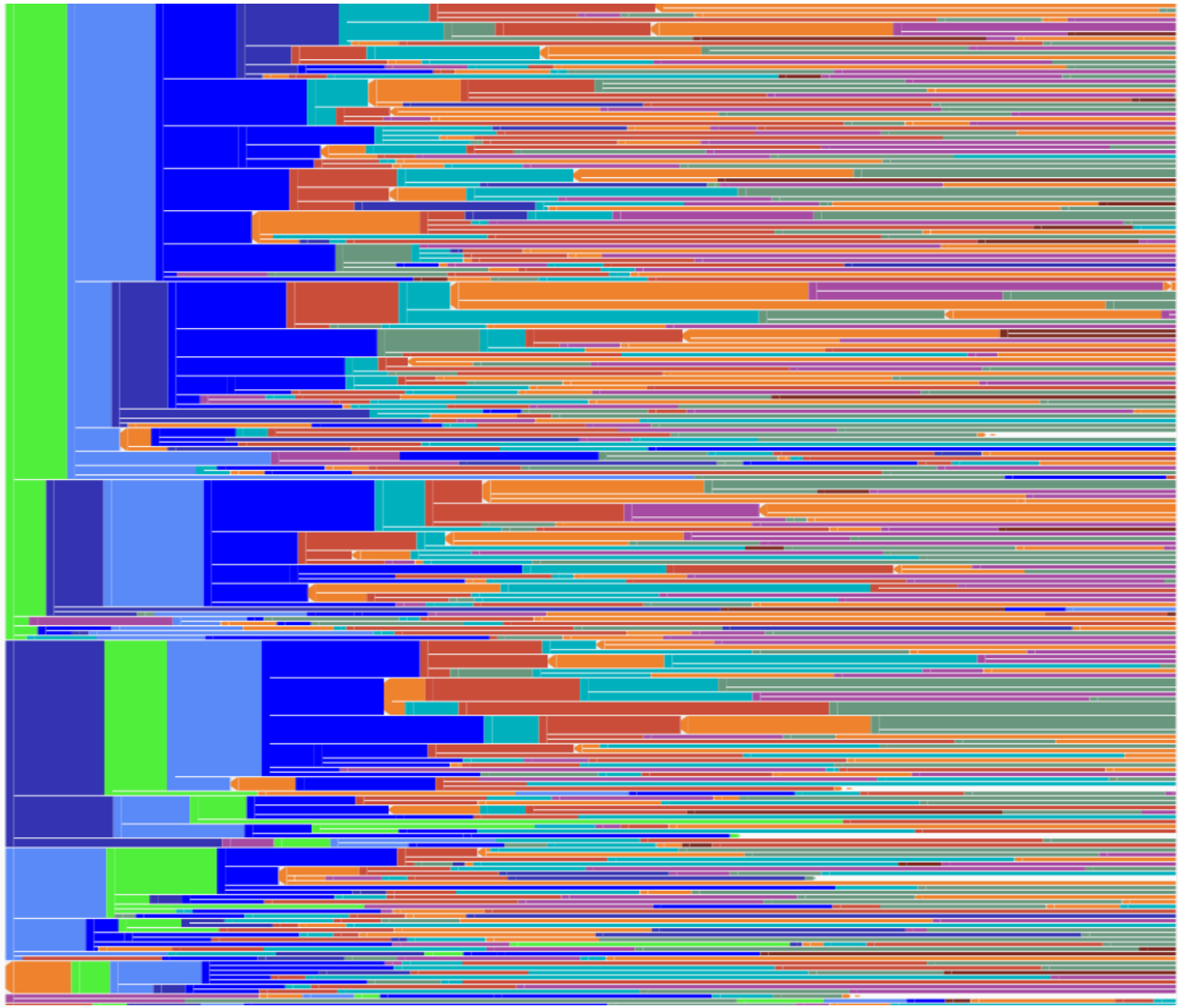
Trauma teams must follow a specific protocol for each new patient, called the abcdes

airway-> breathing-> circulation (pulse)-> disability (gcs)-> external injuries (secondary survey)

Example: Analysis Question

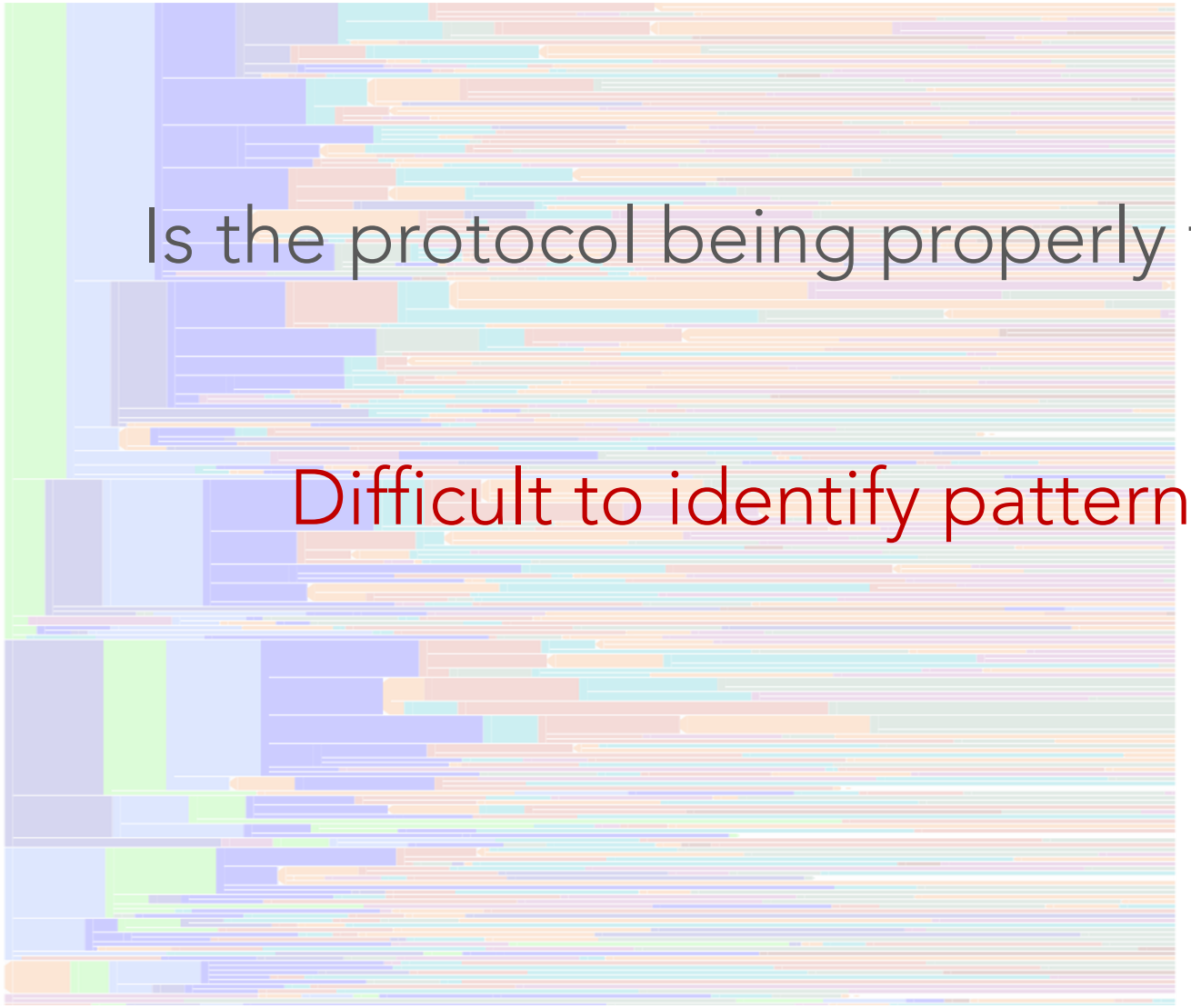
Is the protocol being properly followed in general?

Example: Simple Aggregation



https://hcil.umd.edu/eventflow/

Example: Simple Aggregation

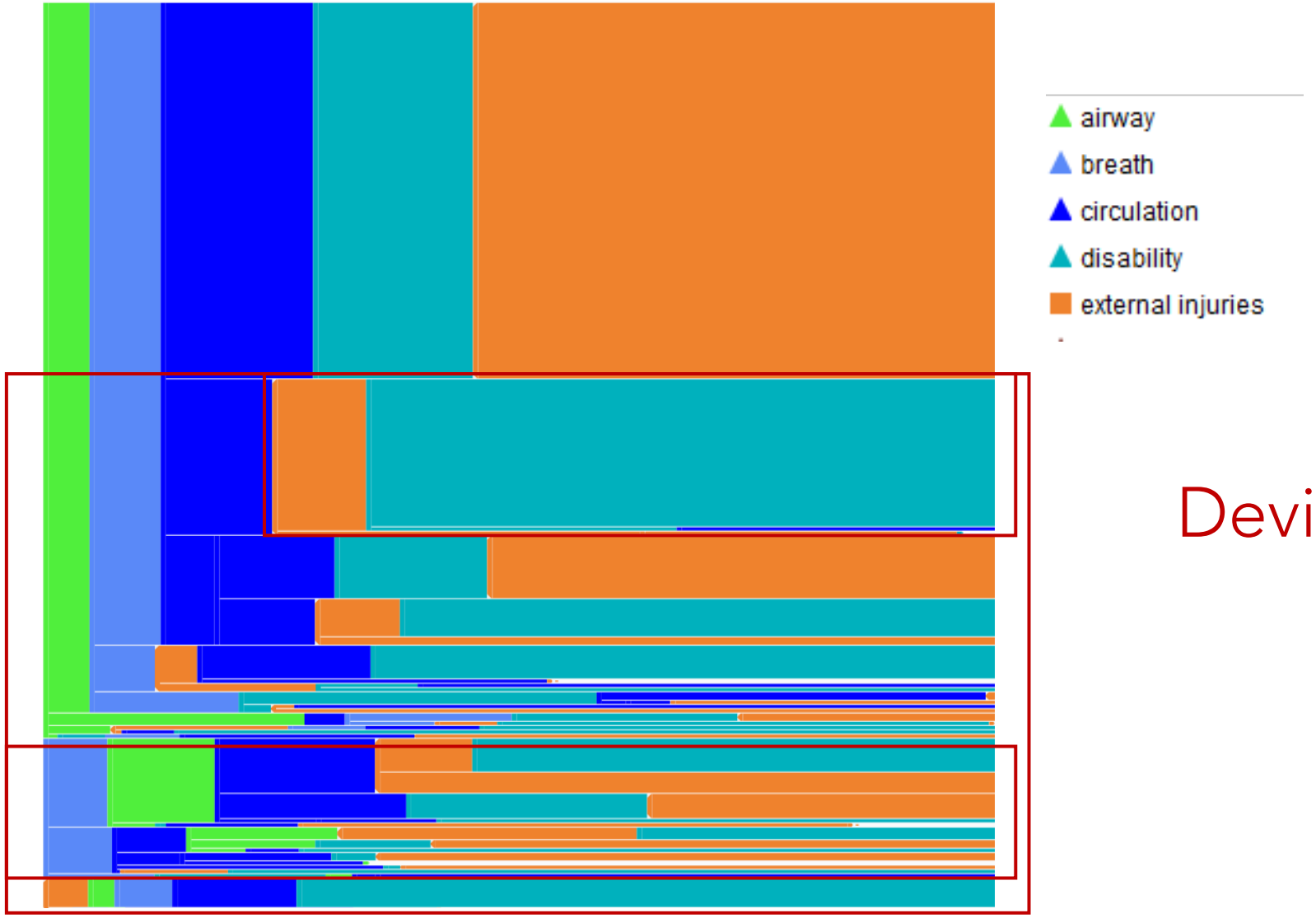


Is the protocol being properly followed in general?

Difficult to identify patterns from aggregation

https://hcil.umd.edu/eventflow/

Example: Manual Summarization



Deviation from abcde

https://hcil.umd.edu/eventflow/

Example: Manual Summarization

Manual summarization requires effort,
and is time consuming!

https://hcil.umd.edu/eventflow/

Automated Visual Summary

EUROVIS Leipzig 2023

Extensive research has focused on visual summary techniques that

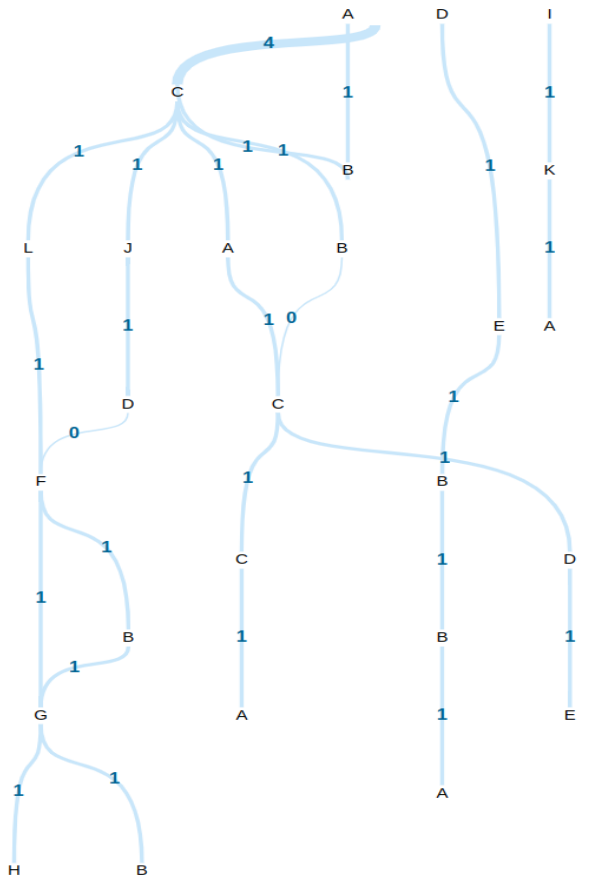
- Automatically extract patterns
- Display results as a concise “visual summary”
  - Must have a data-reduction component, where the analytical results contain much fewer events and sequences than the original dataset
- Provide an 'at-a-glance' overview


world
own
world
goal
2014
2014
world cup
2014
bad
brazil
ow
cup
own goal
cup
own
#brazilvscroatia
world
Admission
Lab & Prescription
Medical Treatment
Operation
Discharge
Acetaminophen & Aspirin
Admission
Creatinine
Blood Test
Insulin & D5W
Metoprolol
Integrelin & Heparin
Calcium Gluconate
Neosynephrine & Propofol & D5W
Bisacodyl & Acetaminophen
Blood Test
Discharge
Prescription
Lab Test
Infusion
Admission
Discharge
Care Pathway Explorer
Metformin Hydrochloride
Other And Unspecified Hyper
Pure Hypercholesterolemia
Simvastatin
Mixed Hyperlipidemia
Lisinopril
Ezetimibe
Pure Hypercho
Mixed Hyperli
Other And Unspecif
Lisinopril
accounts.techX.com:plans
plans:switch:complete
software:AnywareCheckout:checkoutLoading
software:AnywareCheckout:checkoutLoaded
software:AnywareCheckout:validateOrder
Page
Frequent Patterns
Attribute Value Pairs

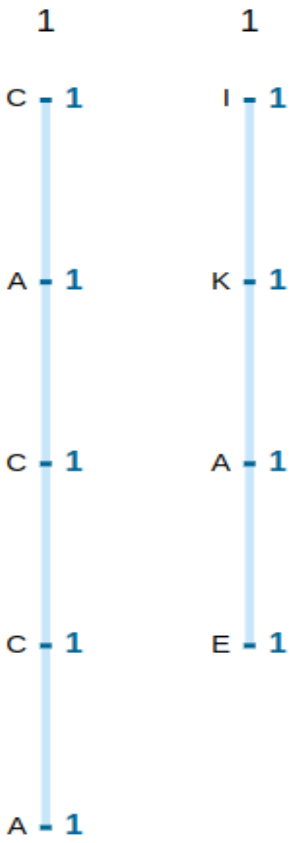
Summary Structure

EUROVIS Leipzig 2023

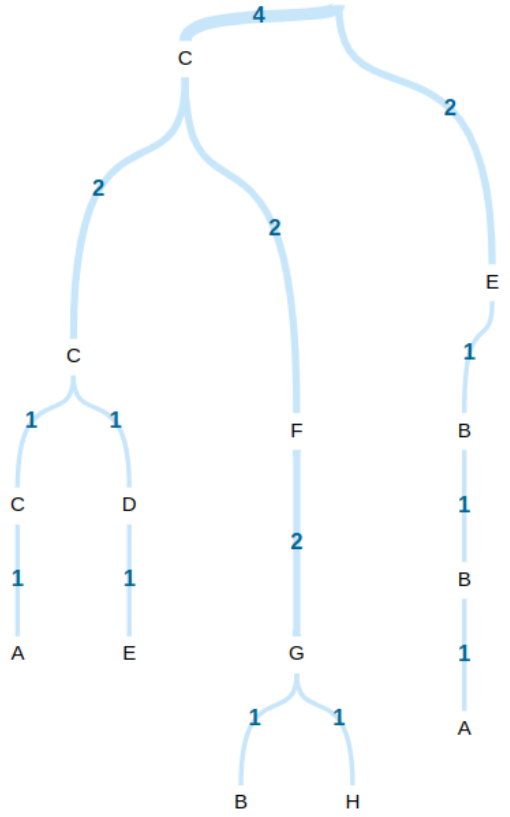
Graph



Linear Sequences



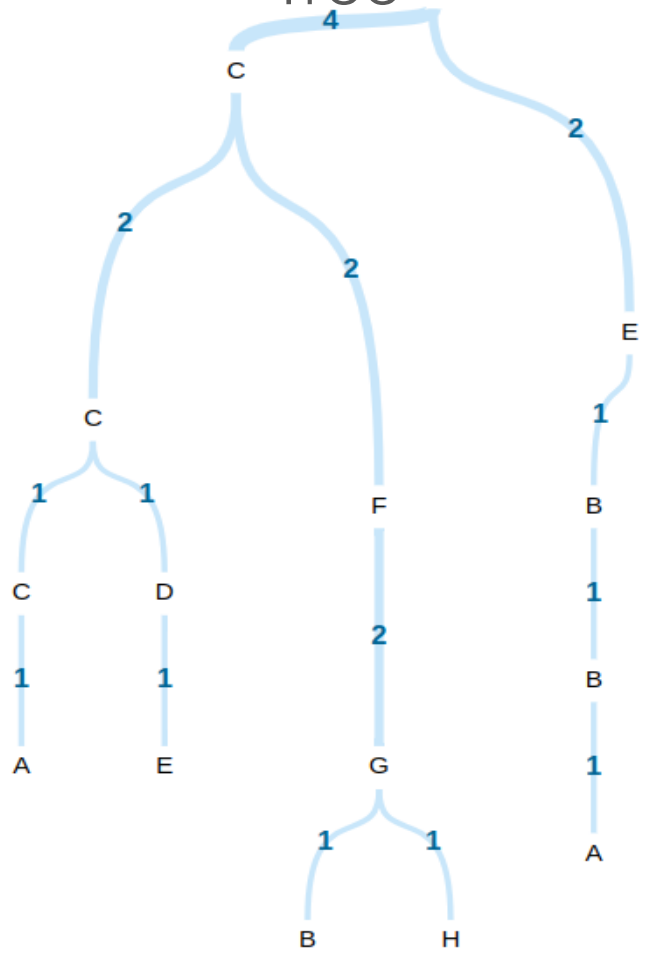
Tree



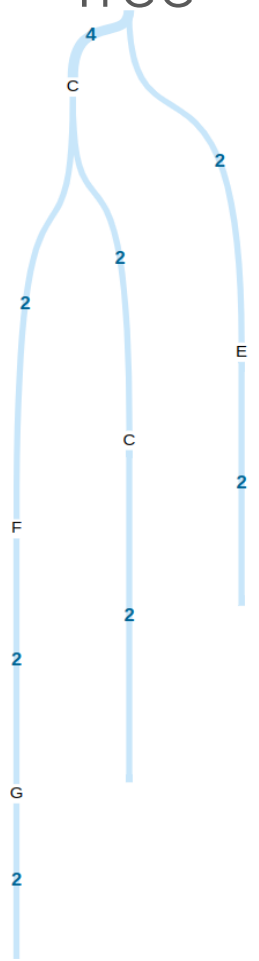
Content & Granularity

EUROVIS Leipzig 2023

Finer Granularity Tree



Coarser Granularity Tree



But How To Choose a Technique?

But How To Choose a Technique?

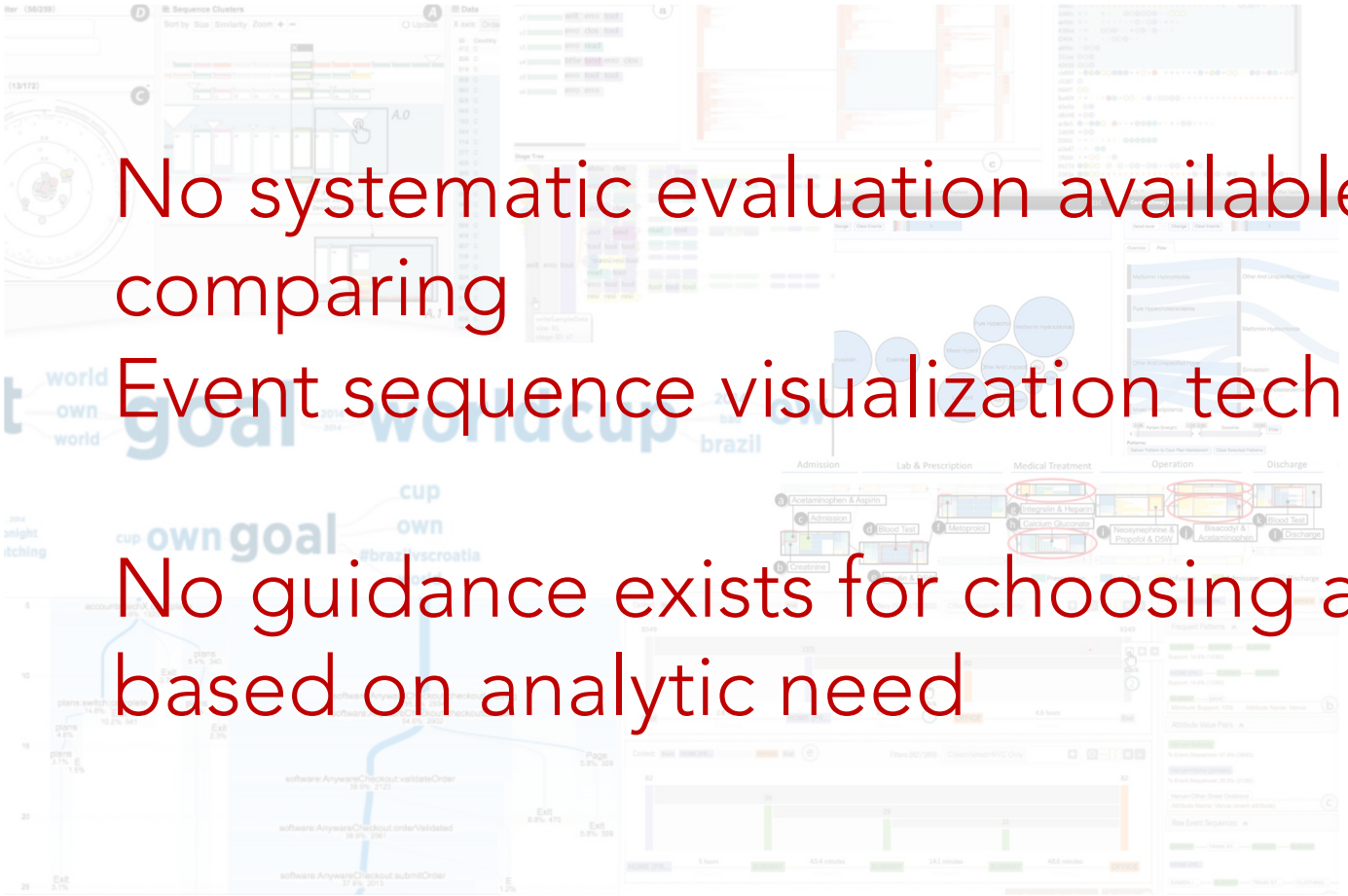
EUROVIS Leipzig 2023

To find an answer, We surveyed 14 sequence summary visualization techniques

13 out of 14 evaluated the proposed techniques through qualitative case studies with domain experts

Only CoreFlow [LKD*17] compared with others, but it only showed sample visualizations

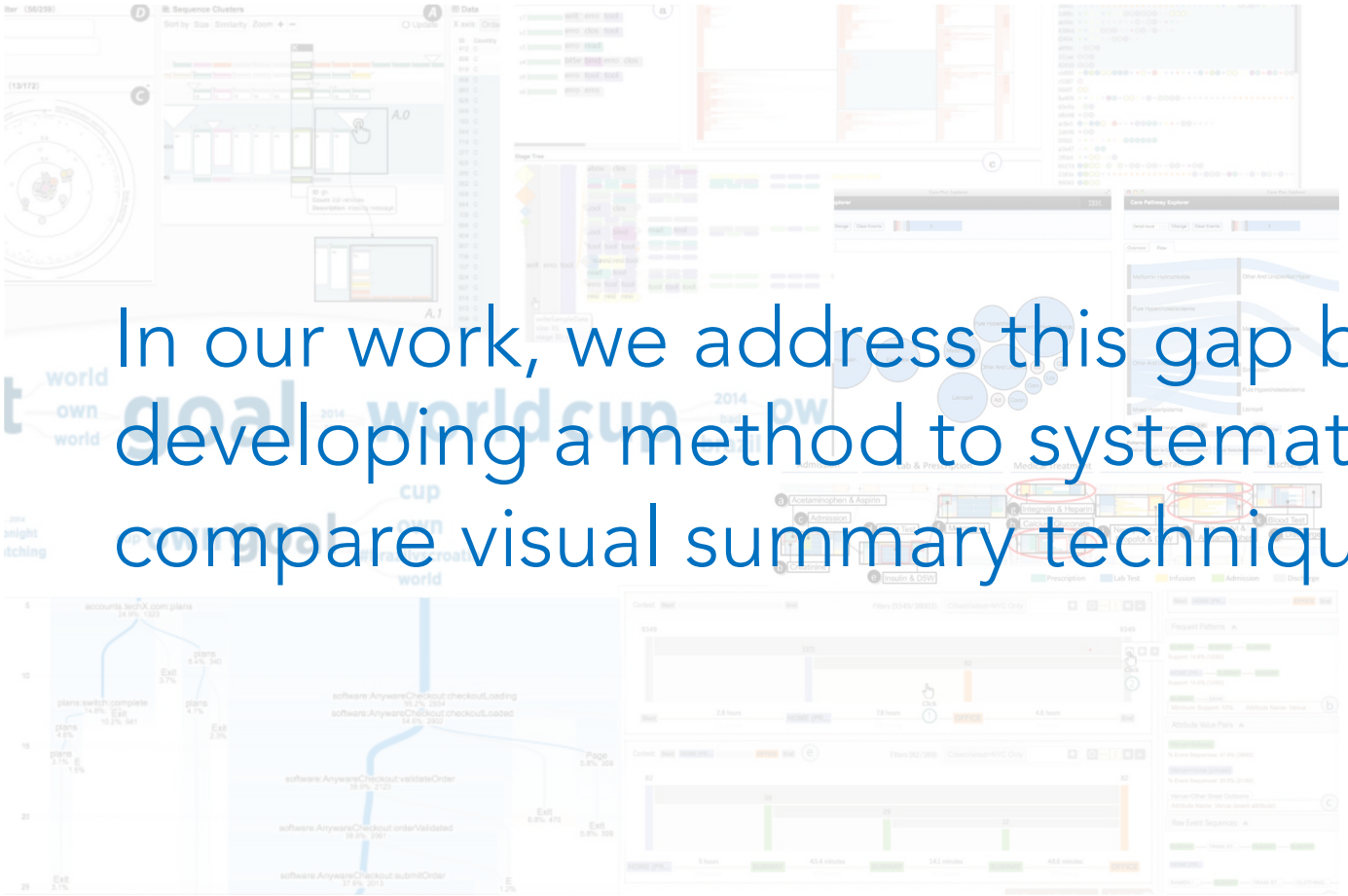
But How To Choose a Technique?



No systematic evaluation available for comparing
Event sequence visualization techniques

No guidance exists for choosing a technique based on analytic need

But How To Choose a Technique?



In our work, we address this gap by developing a method to systematically compare visual summary techniques

Comparative Evaluation: Challenges

EUROVIS Leipzig 2023

- Visualization Systems = Algorithm + Visual Design + Interactivity
- Evaluating graphical perception alone overlooks data reduction component
- Low-level tasks, such as looking up and comparing data values, do not address visual summary quality

Our Approach

EUROVIS Leipzig 2023

- Visualization Systems = Algorithm + Visual Design + Interactivity

✓Focus on algorithms with consistent visual design

- Evaluating graphical perception alone overlooks data reduction component

✓Design study to combine both **summary structure** and **content** in comparison

- Low-level tasks, such as looking up and comparing data values, do not address visual summary quality

✓Adopt **insight-based methodology**: participants assess summary quality based on pre-formulated insight

Technique Selection Criteria

EUROVIS Leipzig 2023

- Domain Agnostic
  Should be able to handle datasets from different problem domains
- Automated
  Generation of visual summaries should require minimal human input
- Data Reduction Component
  Generated summaries should consist much fewer events and sequences
- Granularity (Level of detail) Control
  Granularity of data representation in the summary structure can be controlled
- Summary Structure
  Each technique should have a different summary structure

Representative Techniques

EUROVIS Leipzig 2023

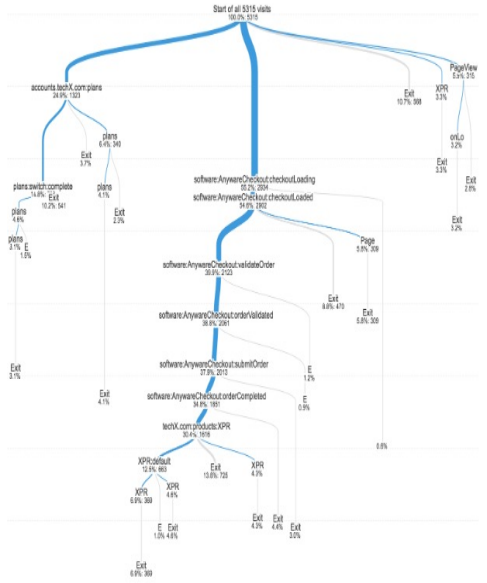
Linear sequences

Sequence Synopsis [CXR 18]

Directed acyclic graph

SentenTree [HWS 17]

Tree



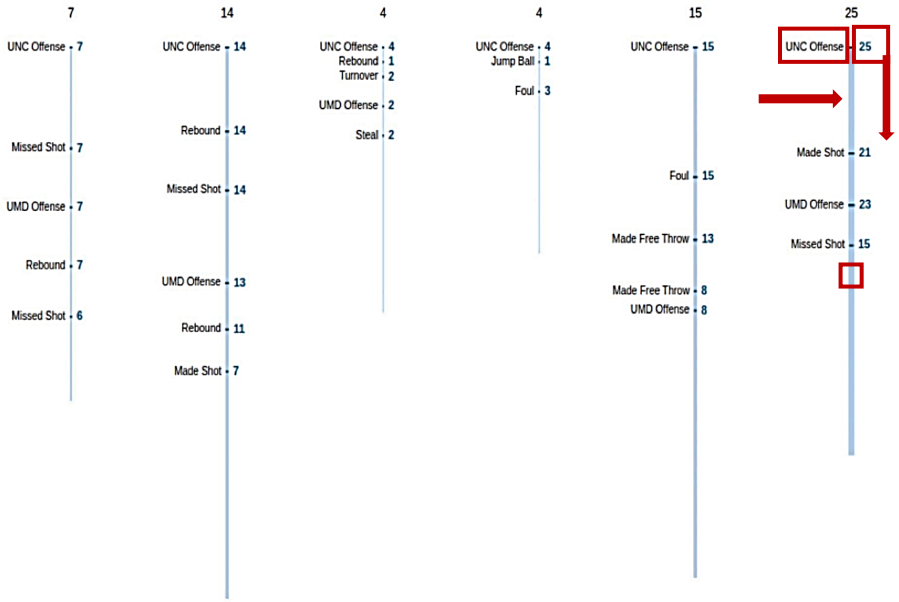
CoreFlow [LKD*17]

Re-implemented as not publicly available

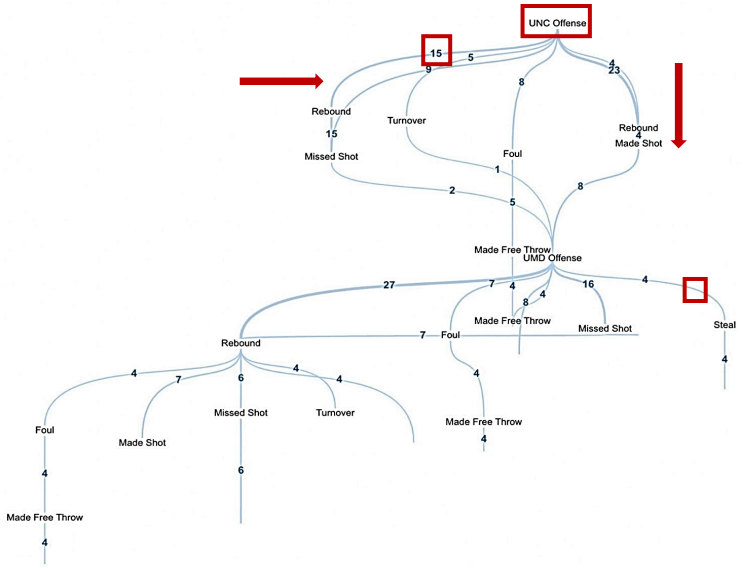
Our Visual Design

EUROVIS Leipzig 2023

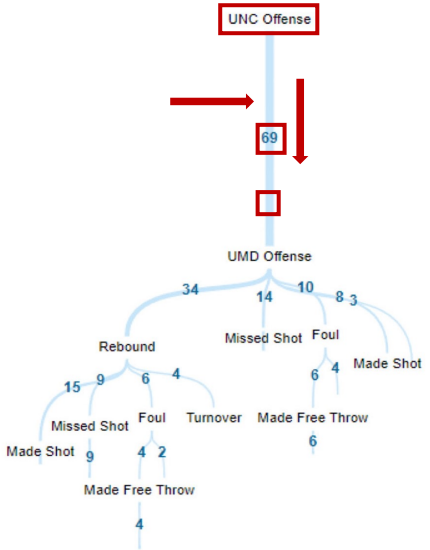
Sequence Synopsis



SentenTree



CoreFlow



Parameterized Granularity Tuning

EUROVIS Leipzig 2023

- All three techniques support parameterized tuning of summary granularity
- Decided upon using 6 granularity levels
- CoreFlow & SentenTree: Minimum support
- Sequence Synopsis: Pattern count

Dataset Selection

EUROVIS Leipzig 2023

- Application Domain
  Include datasets from different areas of application
- Insight (Ground Truth) Availability
  Include datasets with available ground truth from previous studies
- Not used in original paper
  To reduce bias, we excluded datasets used in any of the three original papers

Collected 15 Potential datasets, Selected 6

Dataset	Domain
Emergency Department Records	Medical
UMD vs. UNC Basketball	Sports
Vehicle Movements (VAST Challenge 2017)	Transport
Dev Issue Workflows	Technical
Professor Careers	Academic
Pediatric Patient Records	Medical

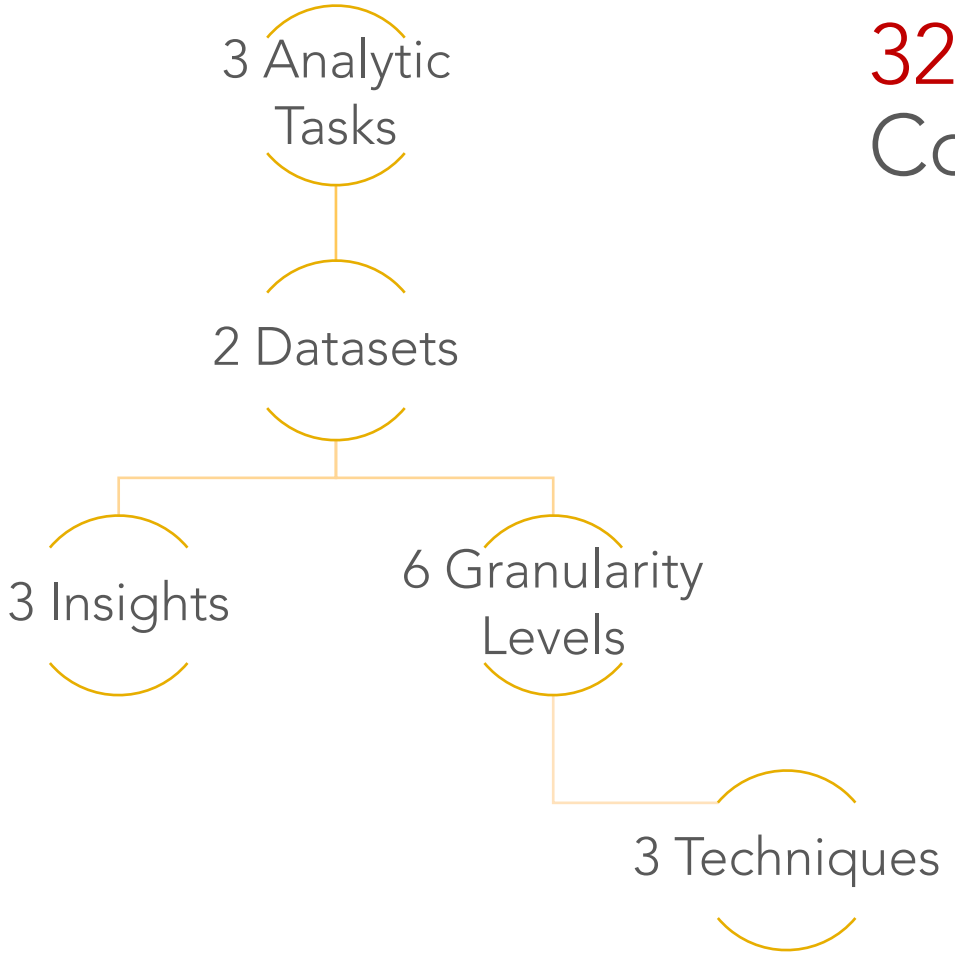
Insight Curation & Task Assignment

EUROVIS Leipzig 2023

- No established benchmark ground truth on the dataset readily available for evaluation
- Curated associated insight from corresponding publications and supplemental videos
- Identified 3 Insights per Dataset
- Curated insights lead to 3 high-level analytical tasks:
  - Common Pattern Identification
  - Clustering
  - Anomaly Detection

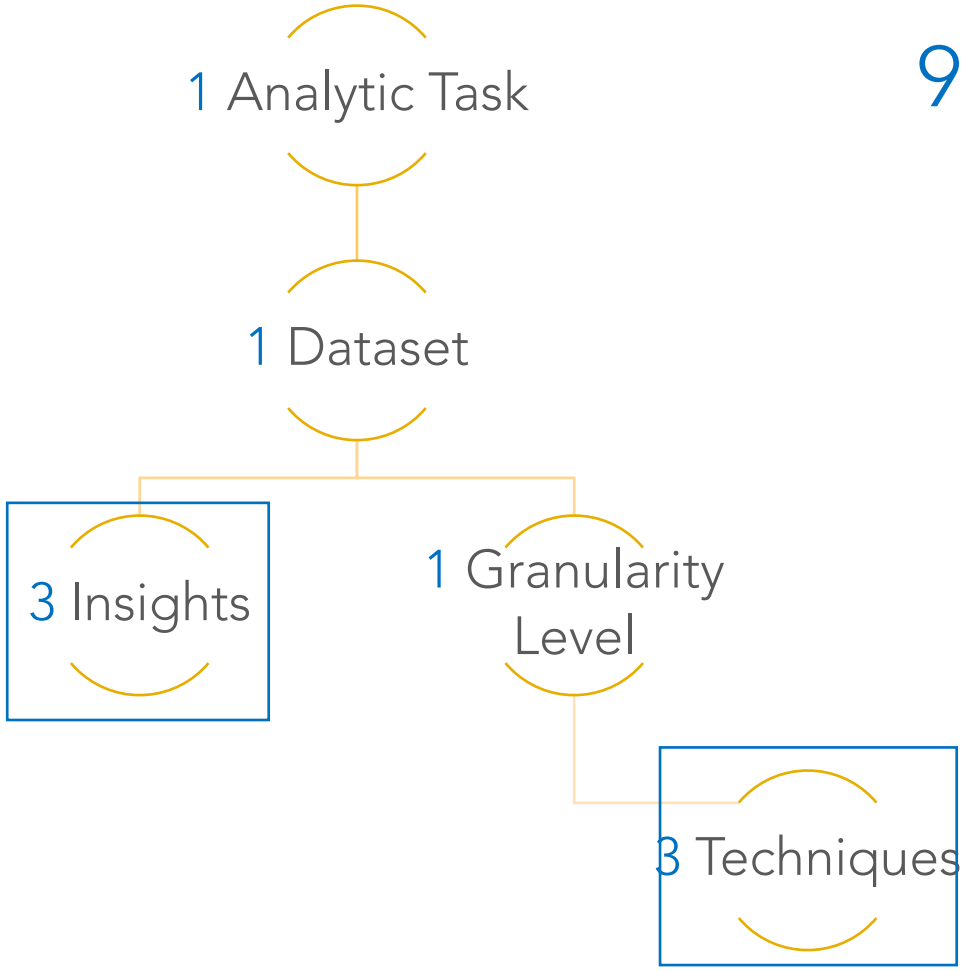
Dataset	Domain	Task
Emergency Department Records	Medical	Common Pattern Identification
UMD vs. UNC Basketball	Sports	
Vehicle Movements (VAST Challenge 2017)	Transport	Clustering
Dev Issue Workflows	Technical	
Professor Careers	Academic	Anomaly Detection
Pediatric Patient Records	Medical	

Crowd-Sourced Study Design



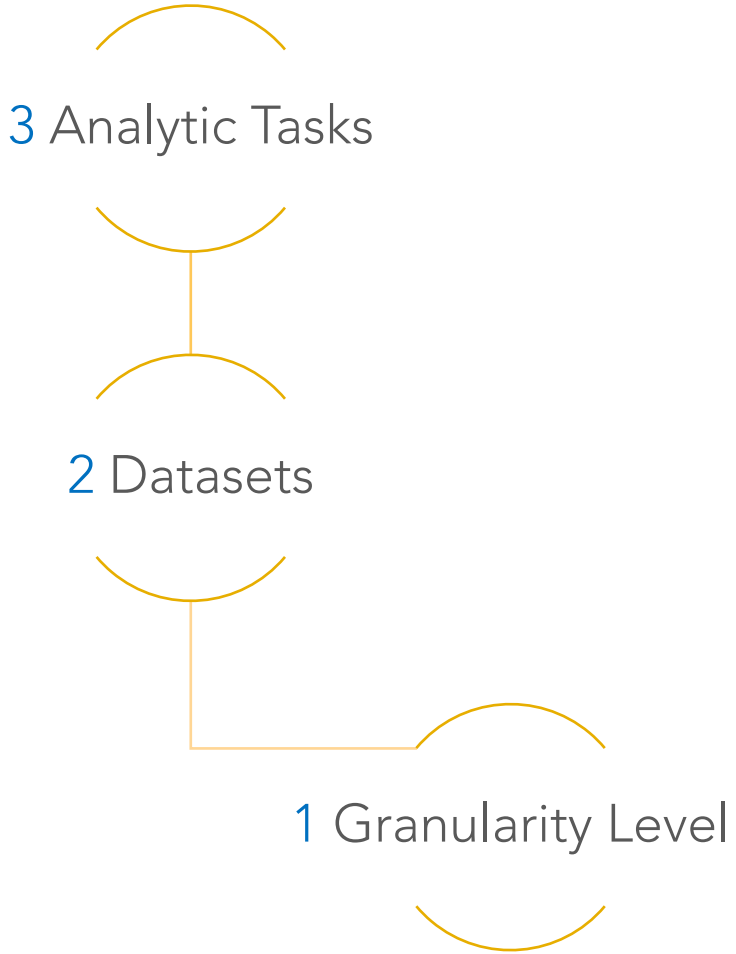
324 Unique Combinations!!

Crowd-Sourced Study Design: Assignment Per Participant



9 Combinations

Crowd-Sourced Study Design: Total Participants



36 Unique conditions

*5 Participants per condition

= 180 Total Participants

Crowd-Sourced Study Design: 3-Phase Evaluation Process

EUROVIS Leipzig 2023

Three-phase evaluation process to ensure the participants have adequate visual and data literacy

Tutorial

Pre-screening

Main Experiment

Crowd-Sourced Study Design: Tutorial

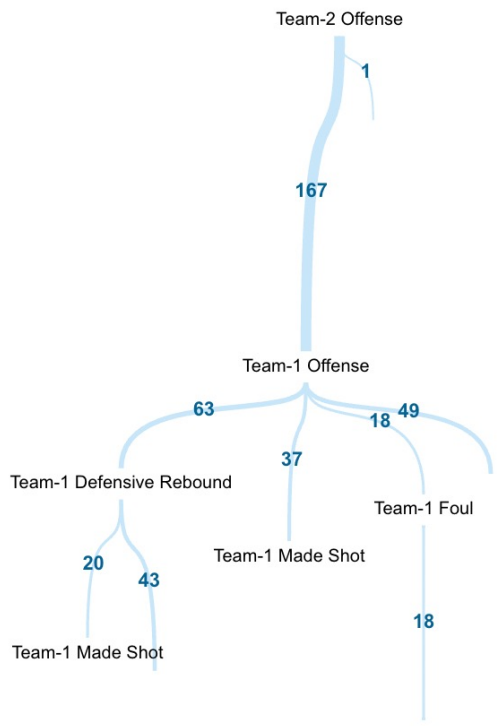
The following tutorial walks you through how to interpret the visualization images.

This visualization shows patterns of 337 + 137 + 188 = 662 people's daily activities. Starting from the top, we can see that 337 people went to a fast food place, another 137 people went to a restaurant, and the rest 188 people did not have a common activity.

The vertical position of the event represents the average time taken. For example, on average, it took longer for people to go to the fast-food place than to the restaurant.

Crowd-Sourced Study Design: Pre-screening

How many times **in total**, after Team-1 Offense starts., do we see Team-1 Made a shot?



- ○ 20
- ○ 57
- ○ 37
- ○ 67

Crowd-Sourced Study Design: Main Experiment

EUROVIS Leipzig 2023

Congratulations on Passing the Pre-screening

In this part, You will be shown different visualization images of a dataset.

The data is about a Basketball match between University of Maryland men's team against UNC. Maryland lost by three points. L 76-79.

If you have read the text above, please proceed.

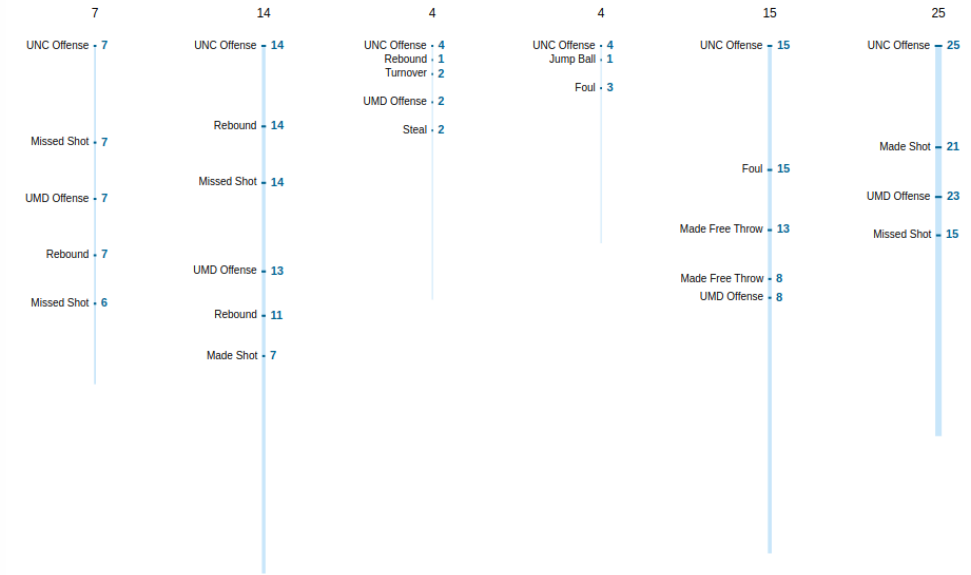
Proceed to Questions

Crowd-Sourced Study Design: Main Experiment

EUROVIS Leipzig 2023

Background information about the dataset:The data is about a Basketball match between University of Maryland men's team against UNC. Maryland lost by three points. L 76-79.

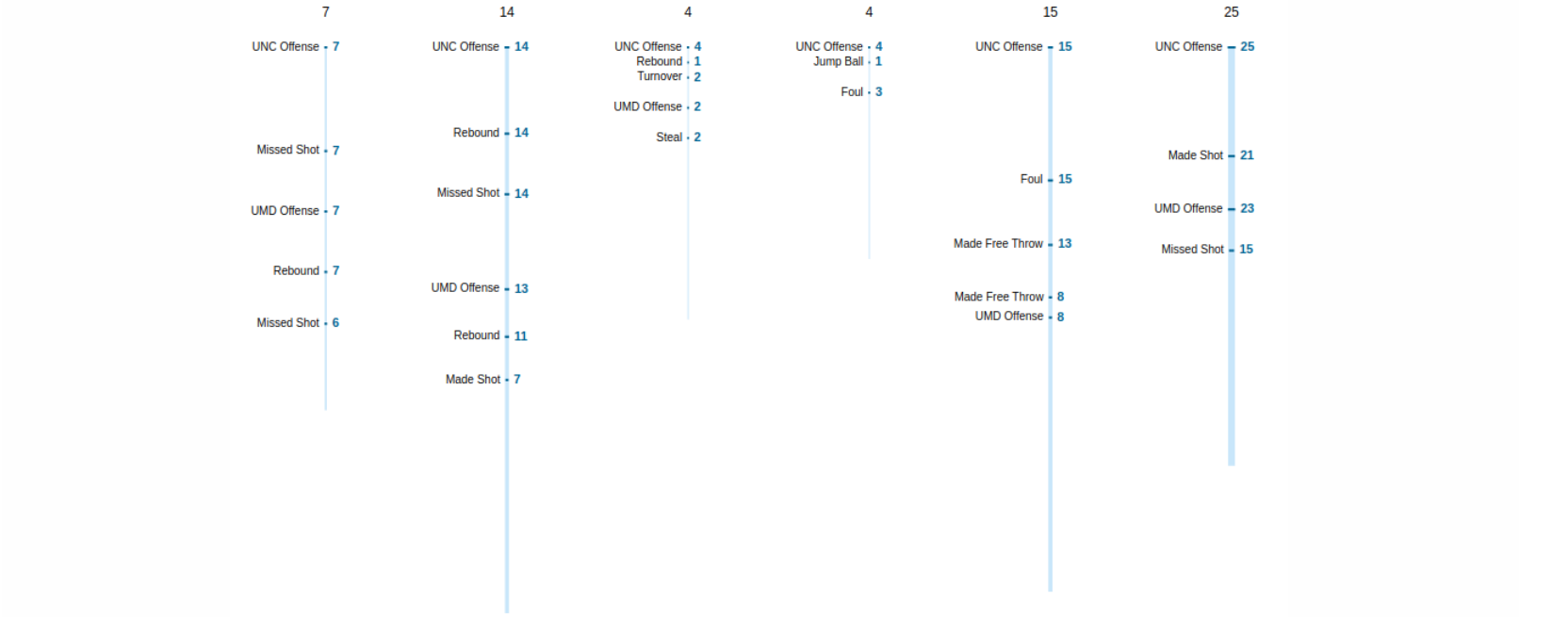
Fact: After a Made Shot or a Made Free Throw during UNC Offense, the most common sequence of events is UMD transitioned onto offense, followed by a missed shot.



The given fact can be confirmed based on the image

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree
○	○	○	○	○	○	○

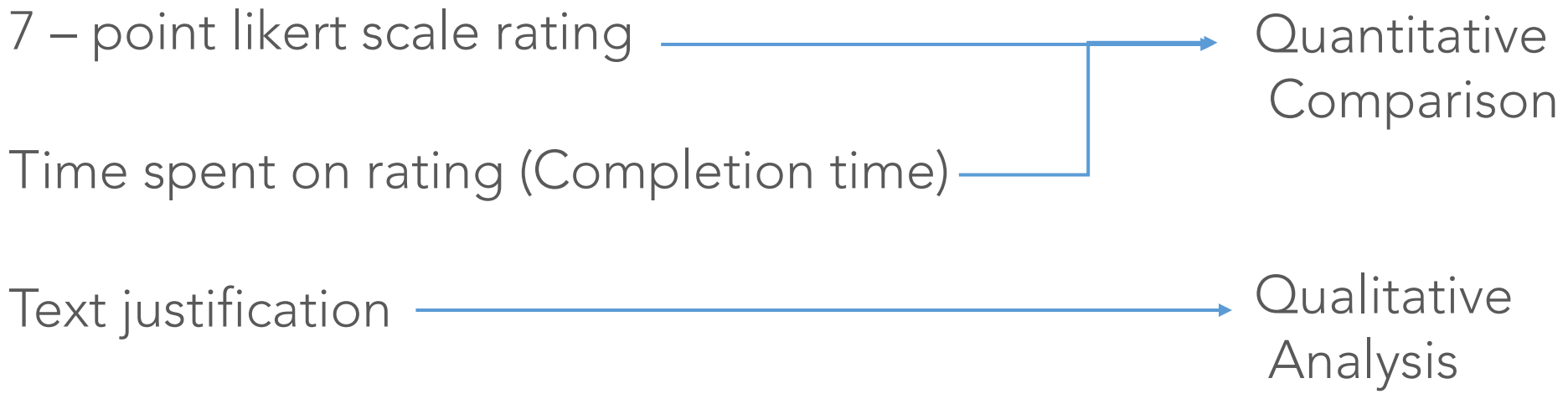
Crowd-Sourced Study Design: Main Experiment



Please explain the reasons behind your ratings for this image. Your explanations should focus on the quality of the image, for example, the image is easy/difficult to understand, the activities in the fact are not shown in the image.

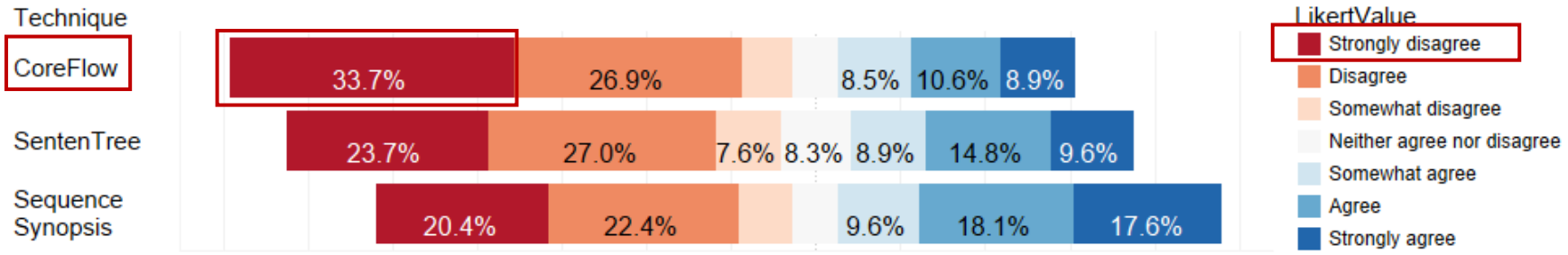
Crowd-Sourced Study Design: Collected Information

We perform our evaluation based on the following collected data



Likert-Scale Rating

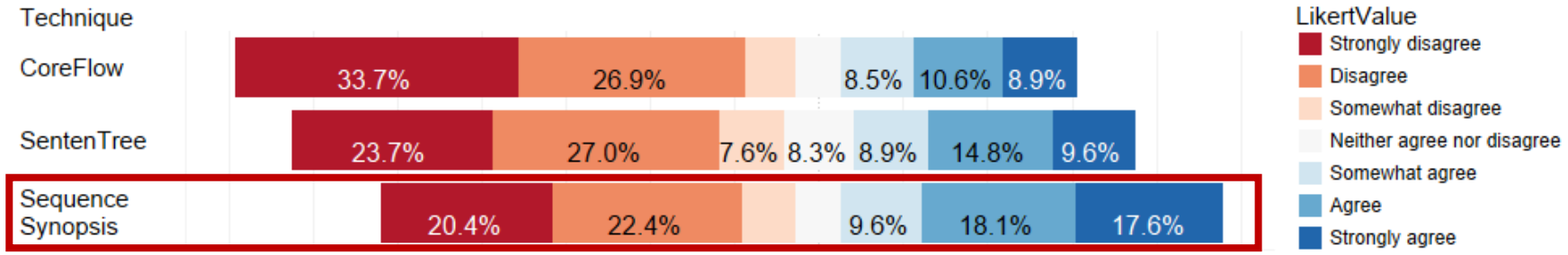
Distribution of ratings across Likert scale for each of the visualization techniques



Likert-Scale Rating: Sequence Synopsis Performs Best

EUROVIS Leipzig 2023

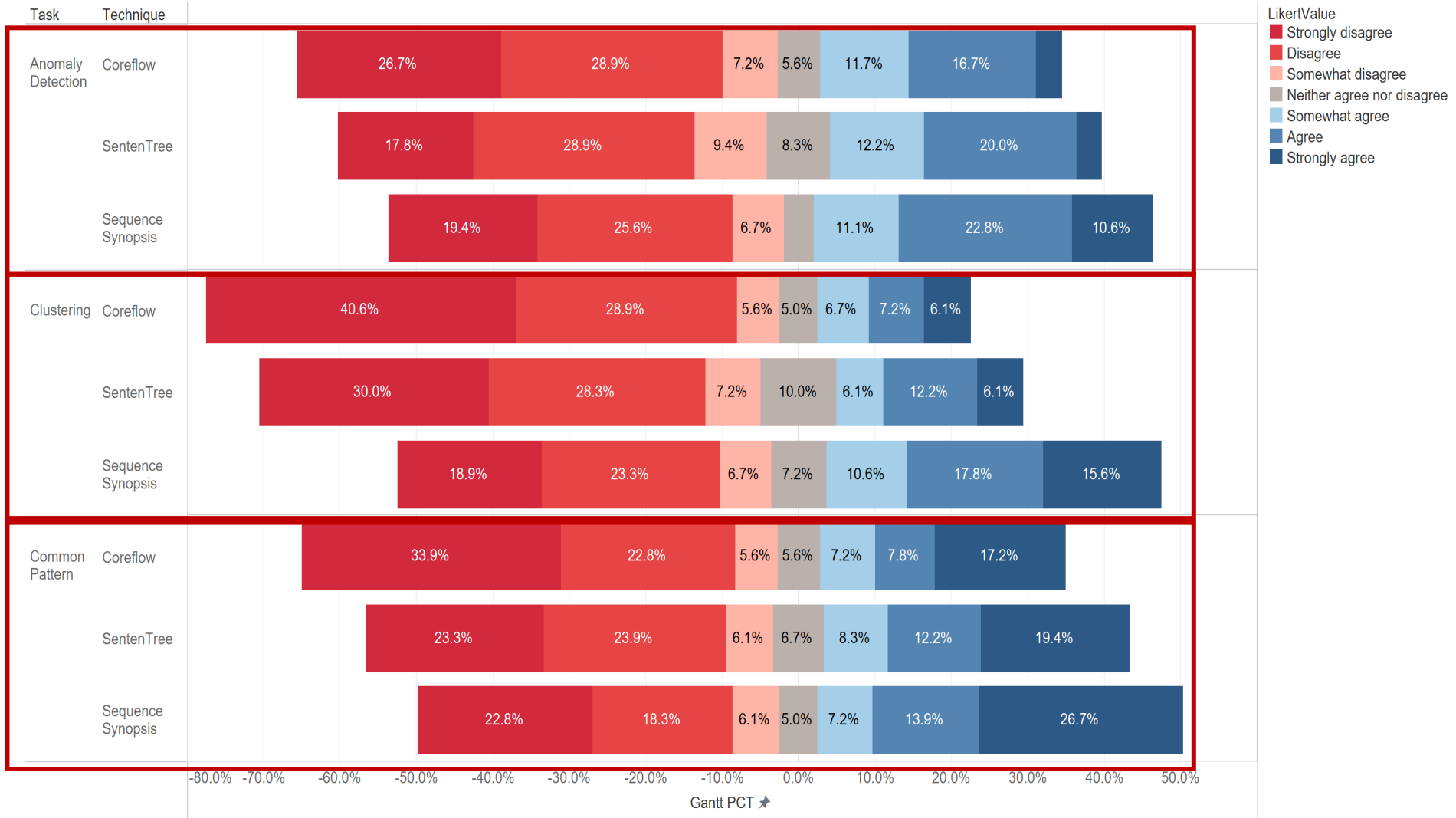
Distribution of ratings across Likert scale for each of the visualization techniques



Technique	Average Rating
CoreFlow	2.95
SentenTree	3.35
Sequence Synopsis	3.86

Likert-Scale Rating: Sequence Synopsis Performs Best

EUROVIS Leipzig 2023



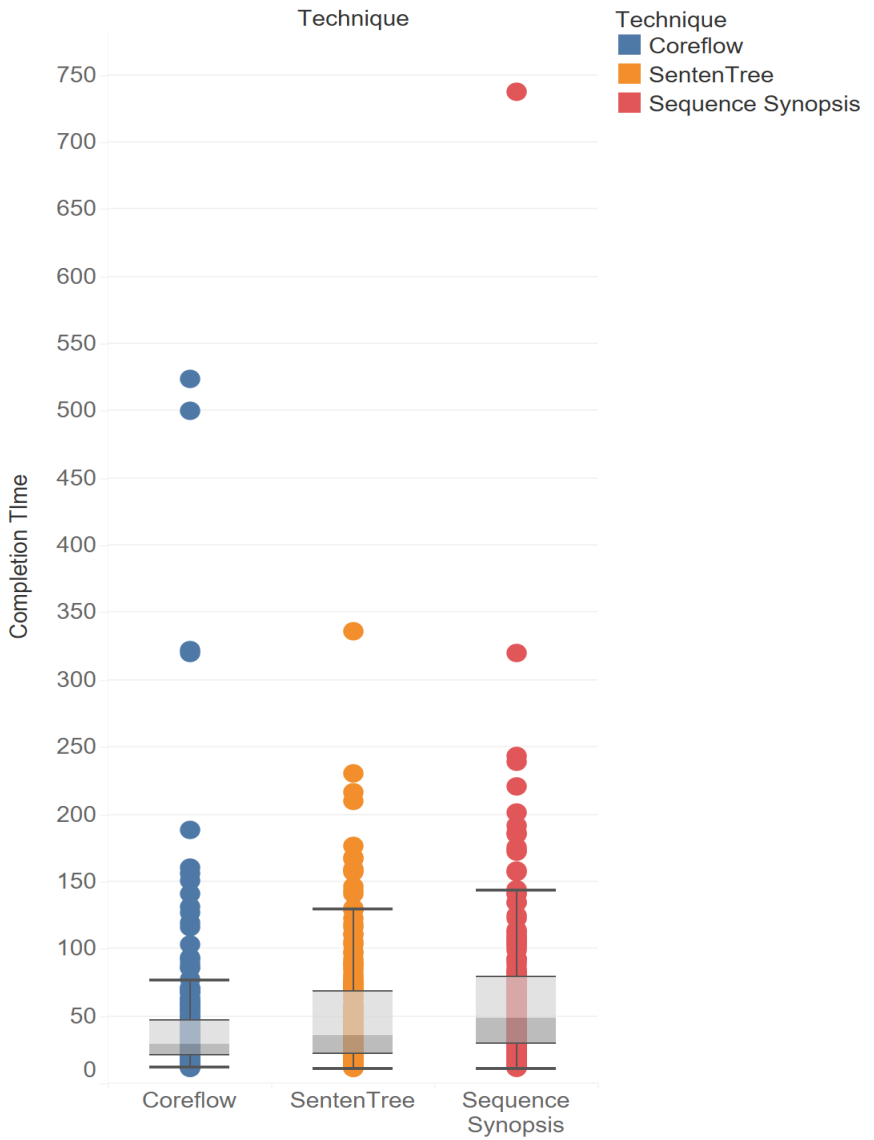
Likert-Scale Rating: Sequence Synopsis Performs Best

EUROVIS Leipzig 2023

Rank:

- 1st : Sequence Synopsis
- 2nd: SentenTree
- 3rd: CoreFlow

Completion Time

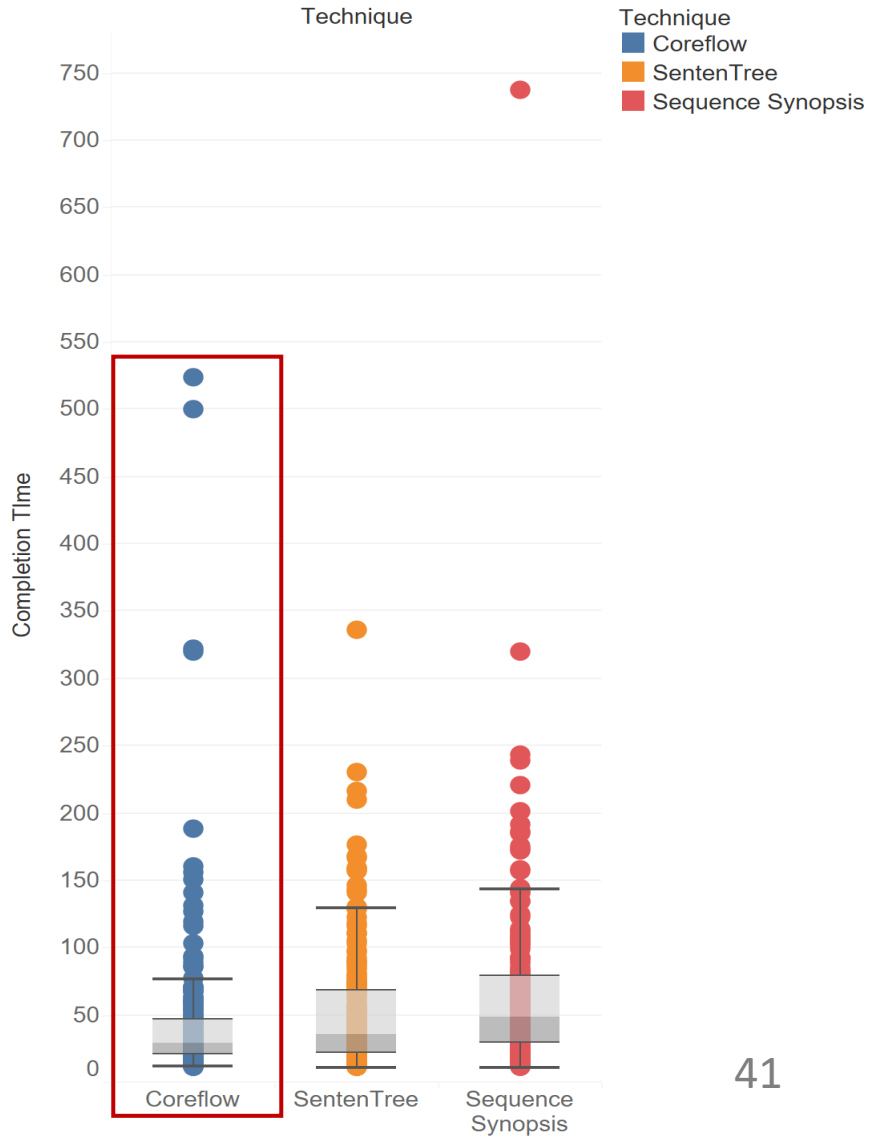


Completion Time: CoreFlow Performs Best

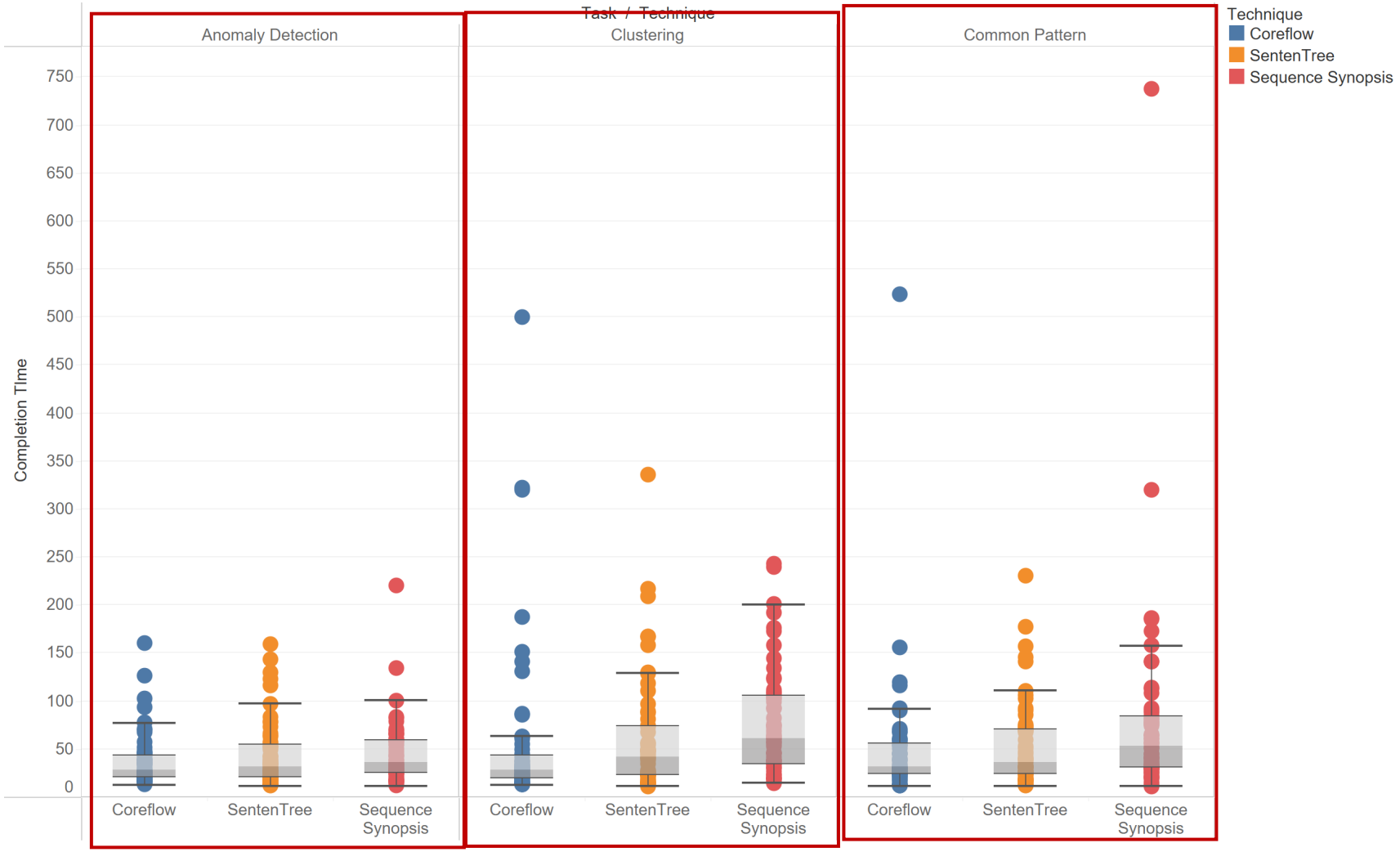
EUROVIS Leipzig 2023

For completion time, smaller value indicates better performance

Technique	Avg. Completion Time (seconds)
CoreFlow	47.61
SentenTree	52.43
Sequence Synopsis	66.92



Completion Time: CoreFlow Performs Best



Completion Time: CoreFlow Performs Best

EUROVIS Leipzig 2023

Rank:

- 1st : CoreFlow
- 2nd: SentenTree
- 3rd: Sequence Synopsis

Identifying Predictors: Technique and Granularity

EUROVIS Leipzig 2023

Built A linear mixed effect model with participant, dataset, and insight as random effects to identify predictor variables

Technique has a significance in both metric

Task is not a significant predictor of likert scale rating or completion time

Granularity is important for completion time

Variables	Technique	Task	Granularity
Likert Rating	✓	✗	✓(SentenTree)
Completion Time	✓	✗	✓

Table 5: Analysis summary of statistical significance

Text Justification Analysis

Performed open coding on the collected text justifications from 180 participants

Aids understanding the reason behind summary quality assessment

"Some of the activities in the facts were not shown in the image. The activities that were shown had numerical discrepancies to the fact."

Missing Key Event
Numbers Do Not Match Text Description

Text Justification Analysis

Performed open coding on the collected text justifications from 180 participants

Aids understanding the reason behind summary quality assessment

"This image is easy to understand and made the questions easy to answer as well because the path the numbers take is quite simple",

Easy to Understand

Text Justification Analysis

Identified 8 total tag categories

799 total tags

One comment may have multiple tags associated with it

Themes: Content & Interpretability

Two overarching themes covering the tag categories:

Content: Match between visualization pattern and insight

Whether important events and associated quantitative information are included in the visualization

Interpretability: Ease of reading the visualization

Content: Events and Numbers

4 associated tag categories:

- Contains Key Event
- Missing Key Event
- Numbers Match Text Description
- Numbers Do Not Match Text Description

Sequence Synopsis outperforms other techniques in terms of including key events and numeric information accuracy

Interpretability: Ease of Understanding

EUROVIS Leipzig 2023

3 associated tag categories:

- Easy to Understand
- Difficult to Understand
- Overlapping Branches

All techniques have almost equal share in both Easy to understand and Difficult to Understand Category

Further analysis shows Interpretability depends on dataset and granularity

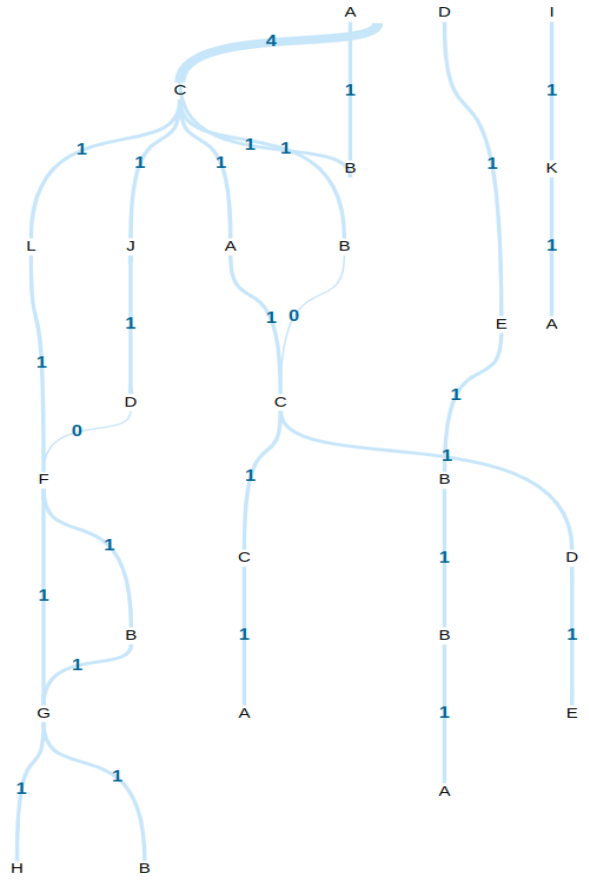
Interpretability: Mixed Reactions to Branching Patterns

EUROVIS Leipzig 2023

Some participants preferred CoreFlow for its simplicity

Some said it is easy to find anomaly following branches in SentenTree

However, overlapping Branches are a prominent issue in SentenTree, which led to confusion and difficulty of understanding

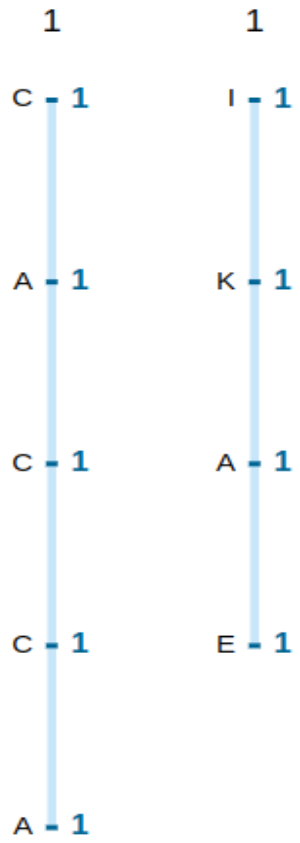


Interpretability: Mixed Reactions to Linear Sequences

EUROVIS Leipzig 2023

Some participants prefer the distinction of linear sequences in Sequence Synopsis

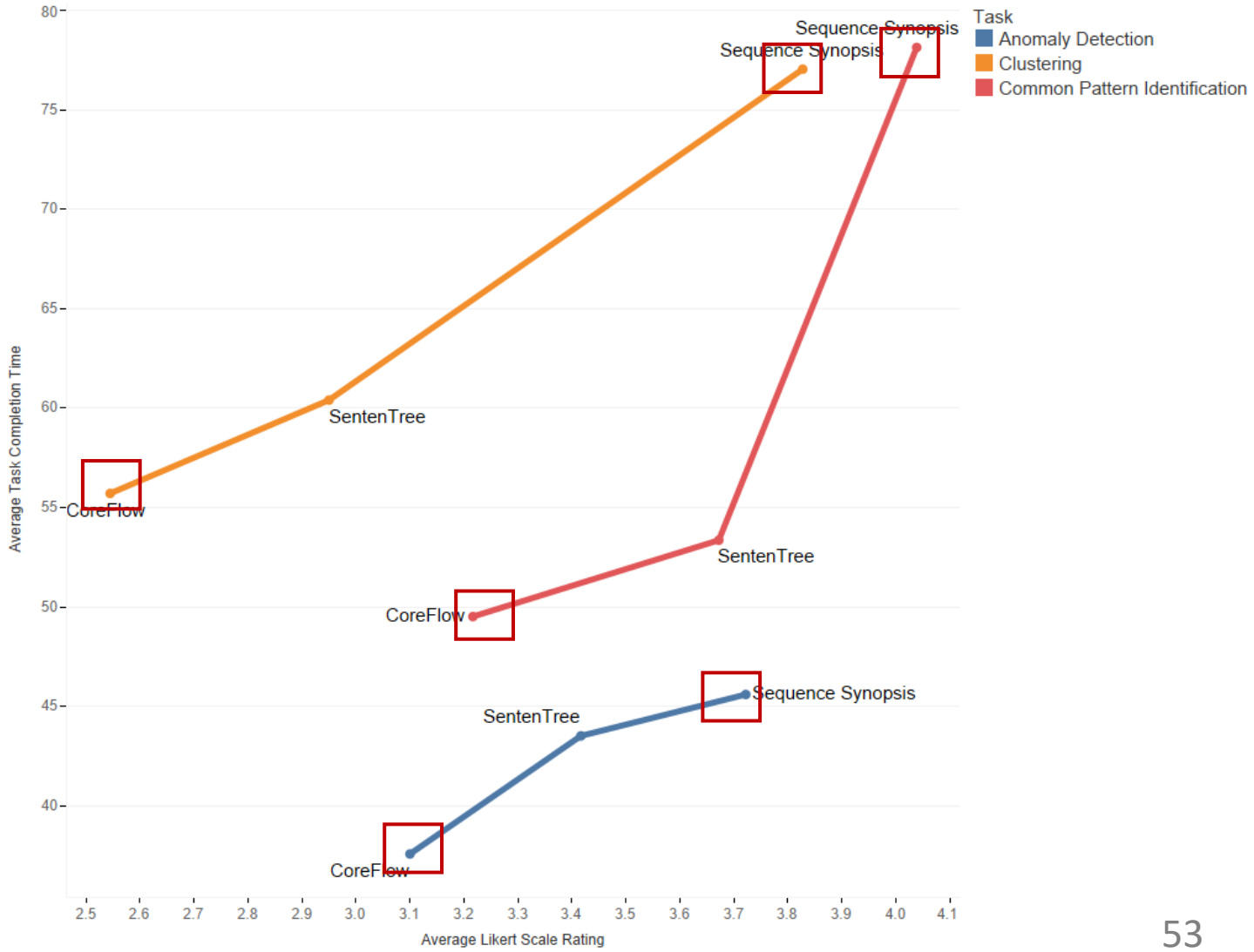
On the other hand, some individuals find it difficult to consolidate information across individual sequences



Rating/Completion Time Trade-off

EUROVIS Leipzig 2023

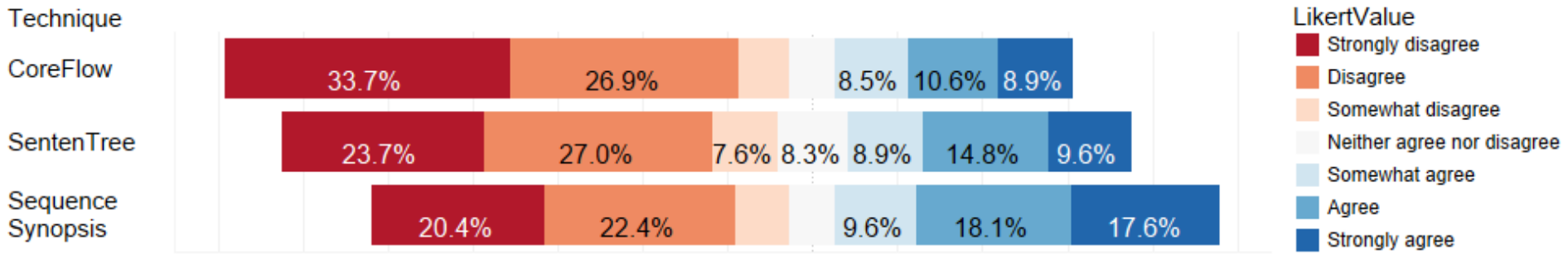
- Task completion time is inversely correlated with technique ratings
- Balancing summary complexity and accuracy is crucial for all visual summarization techniques for event sequences



Visual Summary Techniques Need Improvement

EUROVIS Leipzig 2023

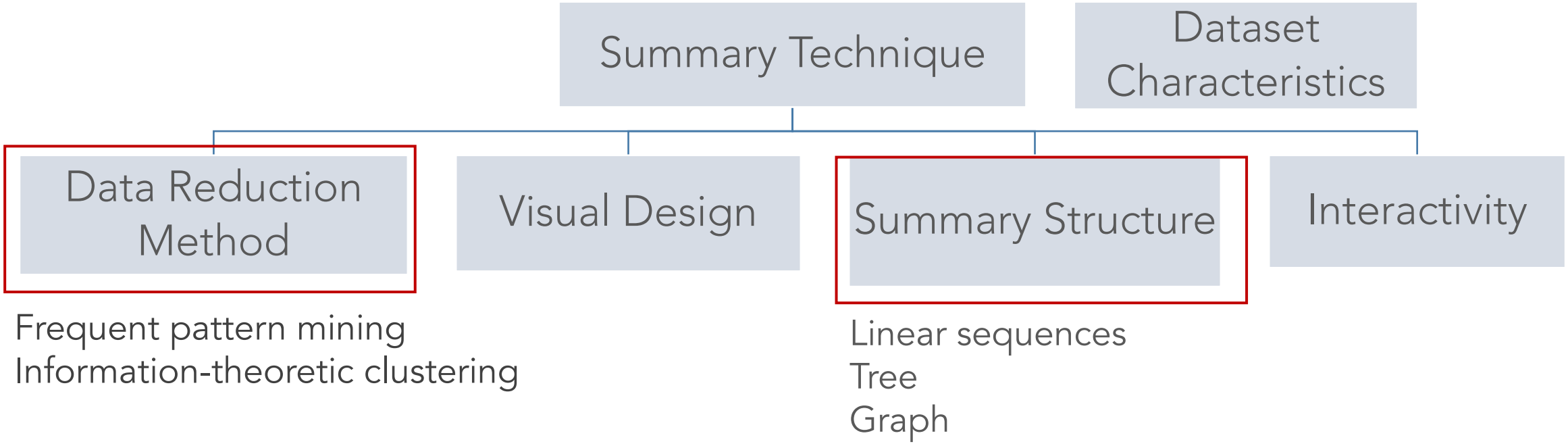
- No perfect visual summarization technique for event sequence data exists
- There is still room for improvement in accuracy, computational resources, and interpretability of visual summaries



Future Work: Assessing Factors Influencing Technique Effectiveness

EUROVIS Leipzig 2023

Further study is required to assess the effects of individual components and their interactions on visual summarization outcomes



Our Contribution

EUROVIS Leipzig 2023

- Experiment Design
  Designed the first study to comprehensively compare visual summary effectiveness
- Result Analysis
  Analysis offers understanding of technique performance, trade-offs, and areas of improvement
- System Implementation
  - Re-implemented 3 existing sequence summary techniques that we plan to Open source