

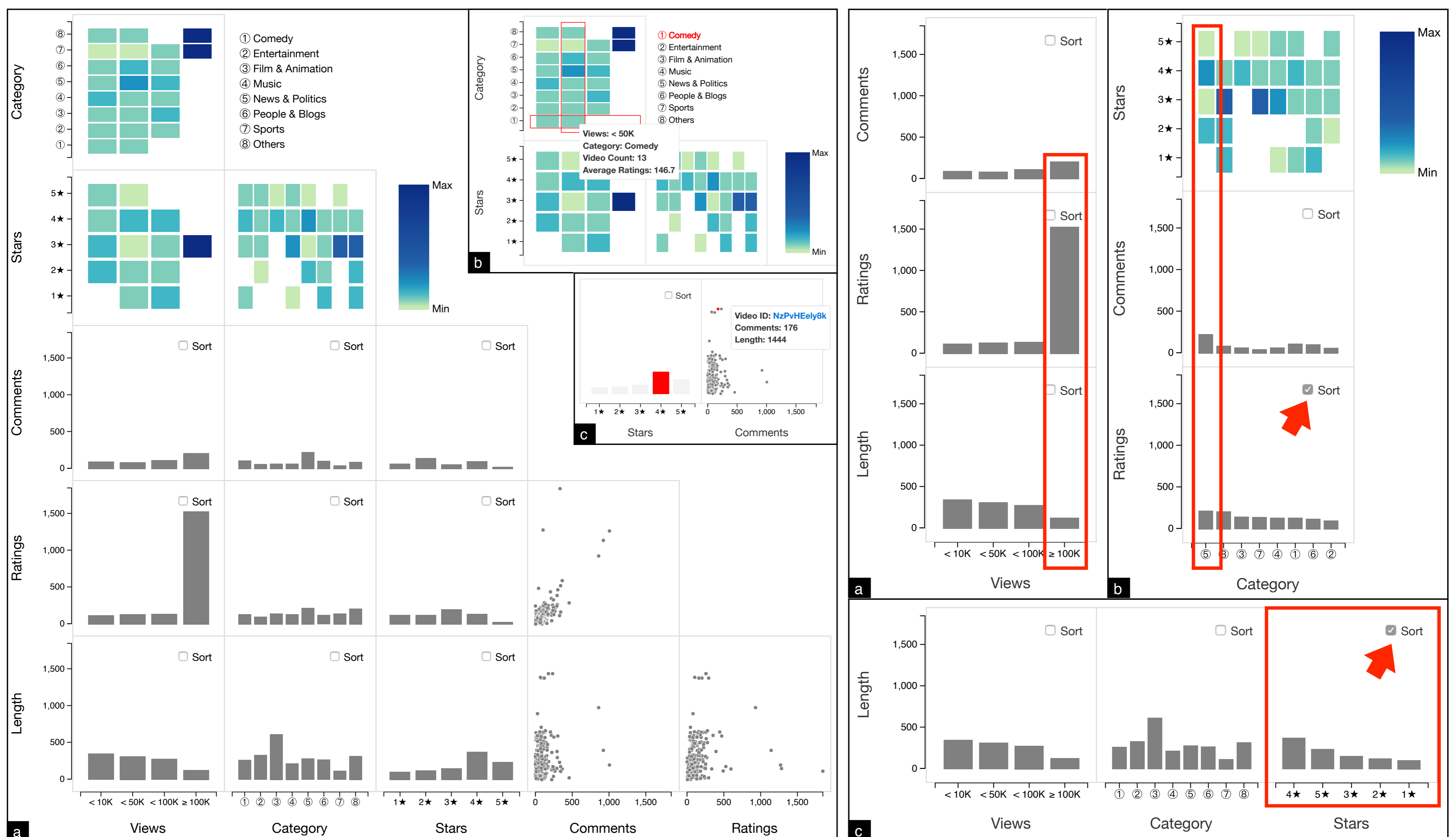
INTERACTIVE EXPLORATION OF MULTIDIMENSIONAL YOUTUBE DATA USING THE GPLOM TECHNIQUE



UNIVERSITY OF
SAN FRANCISCO

Seimei Matsusaki, Xi Han, Alark Joshi, and Sophie Engle
[smatsusaki, xhan14, apjoshi, sjengle]@usfca.edu

Visit git.io/vwDqw for a live demo.



TECHNIQUE

- Used the GPLOM technique by Im et al. [3].
- Uses a scatterplot for two quantitative attributes.
- Uses a bar chart for a categorical and quantitative attribute.
- Uses a heatmap for two categorical attributes

DATASET

- Used 202 videos from "Statistics and Social Network of YouTube Videos" project [2].
- Combined categories, like "Howto & DIY" and "Gadgets & Games", into an "Other" category.
- Binned total views into less than 10k, 10k to 50k, 50k to 100k, and 100k+ views.
- Binned average rating into ★, ★, ★★★, etc.

IMPLEMENTATION (LEFT PANEL)

- Used D3.js [1] and 27 linked SVGs total.
- Example of focus + context [4] on heatmap. The row, column, and legend entry are highlighted.
- Example of coordinated views [5] between barchart and scatterplot.

FINDINGS (RIGHT PANEL)

- Total comments and ratings increase as the total view increases, but ratings increased more drastically for videos over 100k views.
- "News and Politics" has the most comments and ratings, but not the most 5★ ratings.
- As the length increases, the ratings tend to be higher but total views tends to be lower.