Showing High-dimensional Correlation with *Color*

Ivan Zhao
a general visualization technique,
no problem domain
a general visualization technique,
no problem domain

a. high-dimensional data (tabular)
b. correlations (between values)
previous work #1: Parallel Coordinates

- focus on individual documents (each as a edge)
  high level of detail
previous work #1: Parallel Coordinates

- focus on individual documents (each as a edge)
  high level of detail

- annoying edges, given our limited attention
previous work #1: Parallel Coordinates

in essence, we are connecting neighbouring dimensions
previous work #1: Parallel Coordinates

in essence, we are connecting neighbouring dimensions

we have many other options (besides edges/lines)
e.g. texture, animation, color, area, shapes...
my proposed design:
my proposed design:

- Each block represents a value (e.g., phase of flight).
my proposed design:

Each block is a value (e.g. phase of flight)

height = number of documents
my proposed design:

height = number of documents

each block is a value (e.g. phase of flight)

taller, due to the gaps
my proposed design:

- adjustable bins for numerical dimension
- height = number of documents
- each block is a value (e.g. phase of flight)
- taller, due to the gaps
my proposed design:

- Adjustable bins for numerical dimension
- Height = number of documents
- Each block is a value (e.g., phase of flight)
- How do we show the connections b/w dimensions?
- Taller, due to the gaps
my proposed design:
my proposed design:

percentage of documents has phase of flight = ascent
my proposed design: percentage of documents has phase of flight = ascent

many-to-one relationship (v.s neighbouring pairs in Parallel Coordinates)
my proposed design:

percentage of documents has \textit{phase of flight} = \textit{ascent}

many-to-one relationship (v.s neighbouring pairs in Parallel Coordinates)

outliers still visible
my proposed design:

stacked bar charts turned around, and normalized
my proposed design:

stacked bar charts turned around, and normalized

correlations shown through “color domination”
my proposed design:

stacked bar charts turned around, and normalized correlations shown through “color domination”

one interested dimension at a time, but shifts to others quickly (more in tune with how attention works?)
my proposed design:

freedom in layout
my proposed design:

known issues so far:

- limited # of colors
  but, ordering helps
- correlations in numerical dimensions
previous work #2: Mosaic Display

- correlation = area
- two dimensions at a time
- rigid layout
previous work #3:
Parallel Set

- correlation = edges + color
- messy edges
questions?