hci lab
SEOUL NATIONAL UNIVERSITY

VIS 2021

Augmenting Parallel Coordinates Plots with Color-coded Stacked Histograms

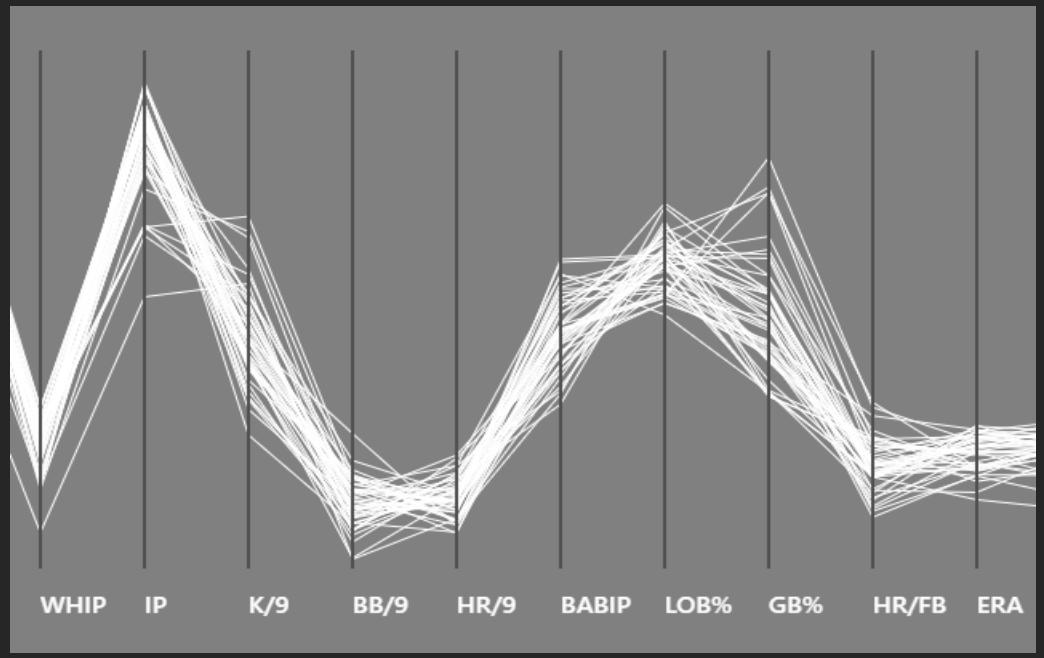
Jinwook Bok
bok@hcil.snu.ac.kr
Seoul National University

Bohyoung Kim
bkim@hufs.ac.kr
Hankook University of Foreign Studies

Jinwook Seo
jseo@snu.ac.kr
Seoul National University

Motivation

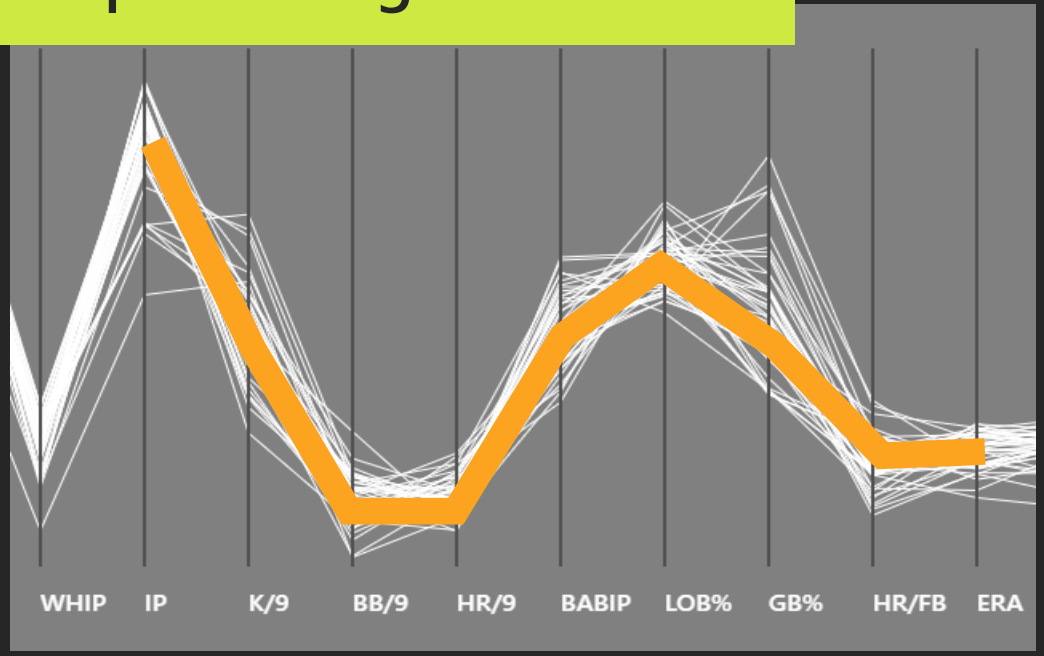
Innate advantages of parallel coordinates plot (PCP)…



Motivation

Innate advantages of parallel coordinates plot (PCP)…

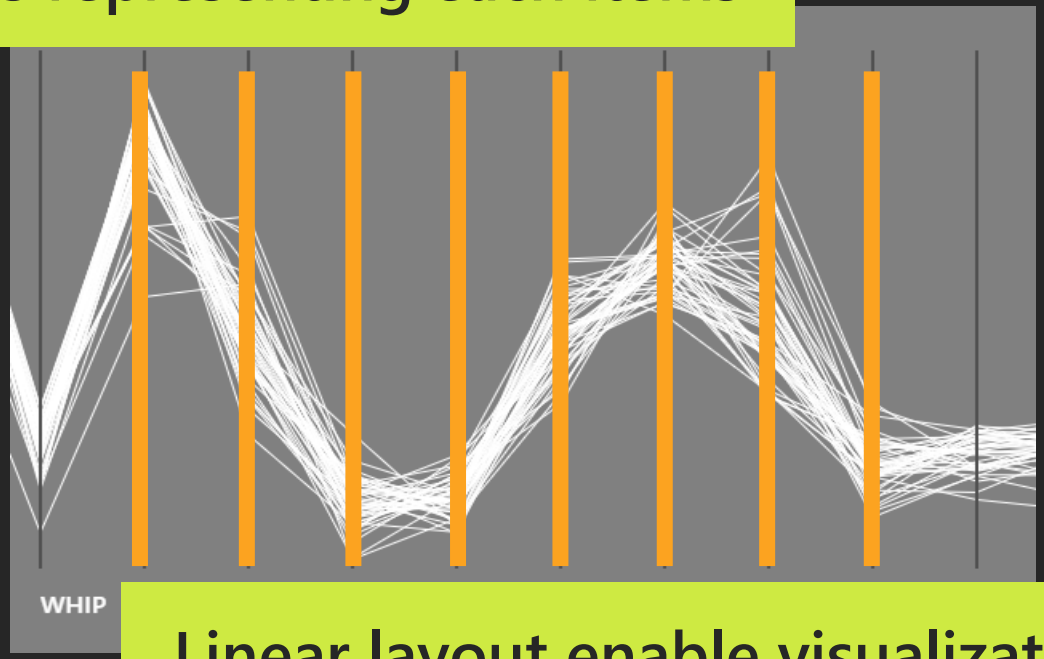
Visual patterns of polylines representing each items



Motivation

Innate advantages of parallel coordinates plot (PCP)…

Visual patterns of polylines representing each items

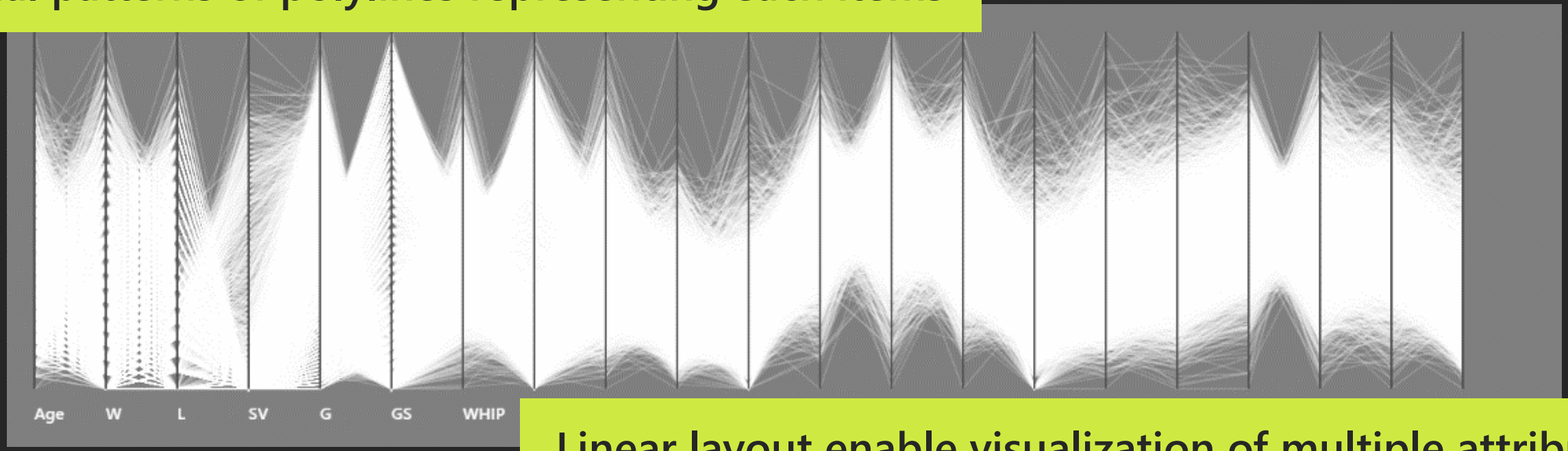


Linear layout enable visualization of multiple attributes

Motivation

Innate advantages of parallel coordinates plot (PCP)…

Visual patterns of polylines representing each items

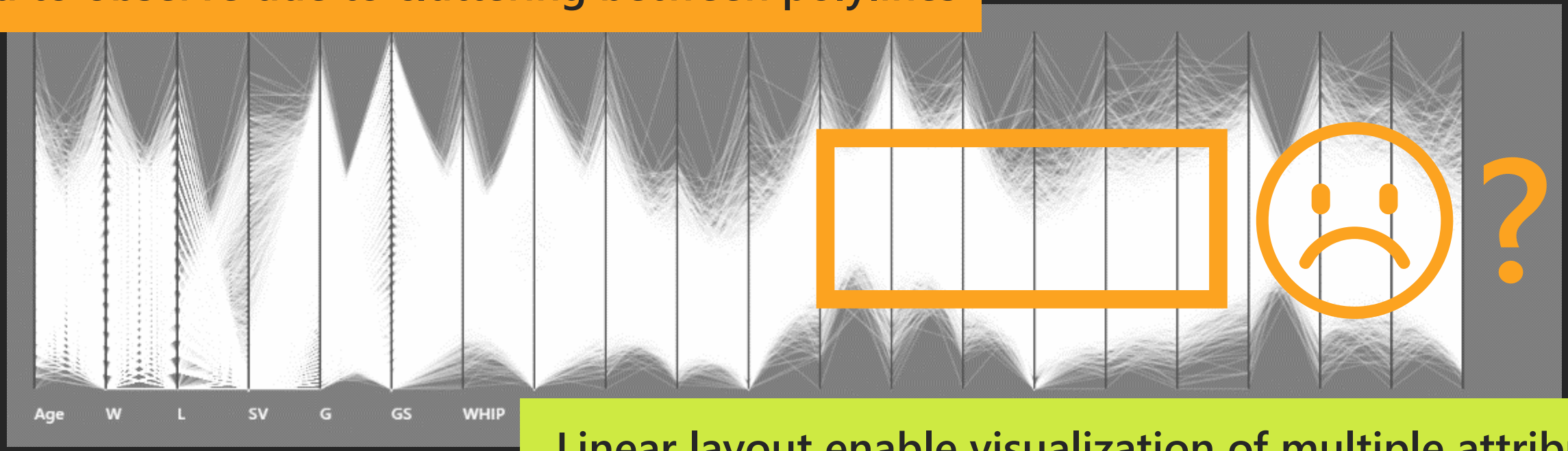


Linear layout enable visualization of multiple attributes

Motivation

…become limitations when size of data increases

Hard to observe due to cluttering between polylines



Linear layout enable visualization of multiple attributes

Motivation

…become limitations when size of data increases

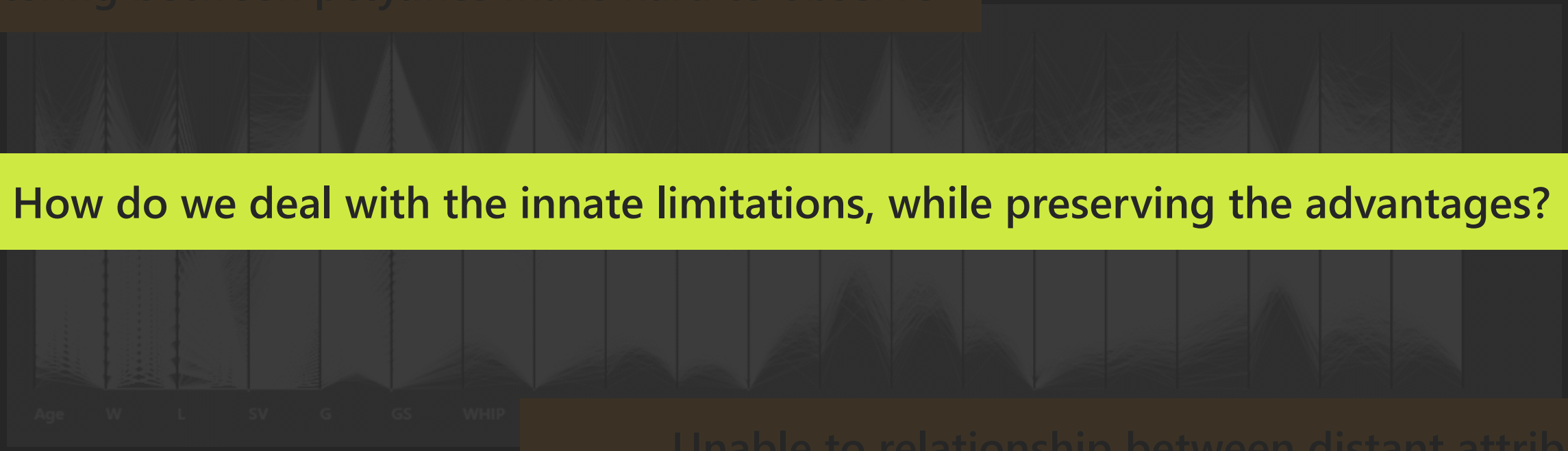
Hard to observe due to cluttering between polylines

Unable to find relationship between distant attributes

Motivation

…become limitations when size of data increases

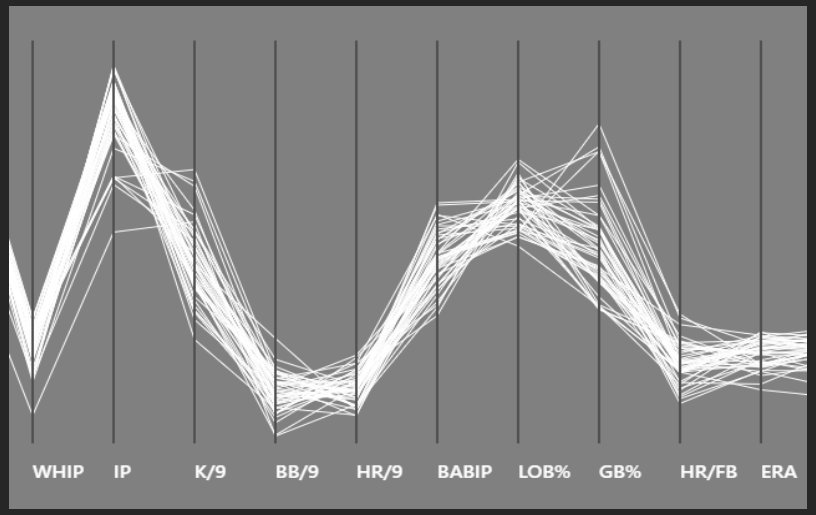
Cluttering between polylines make hard to observe

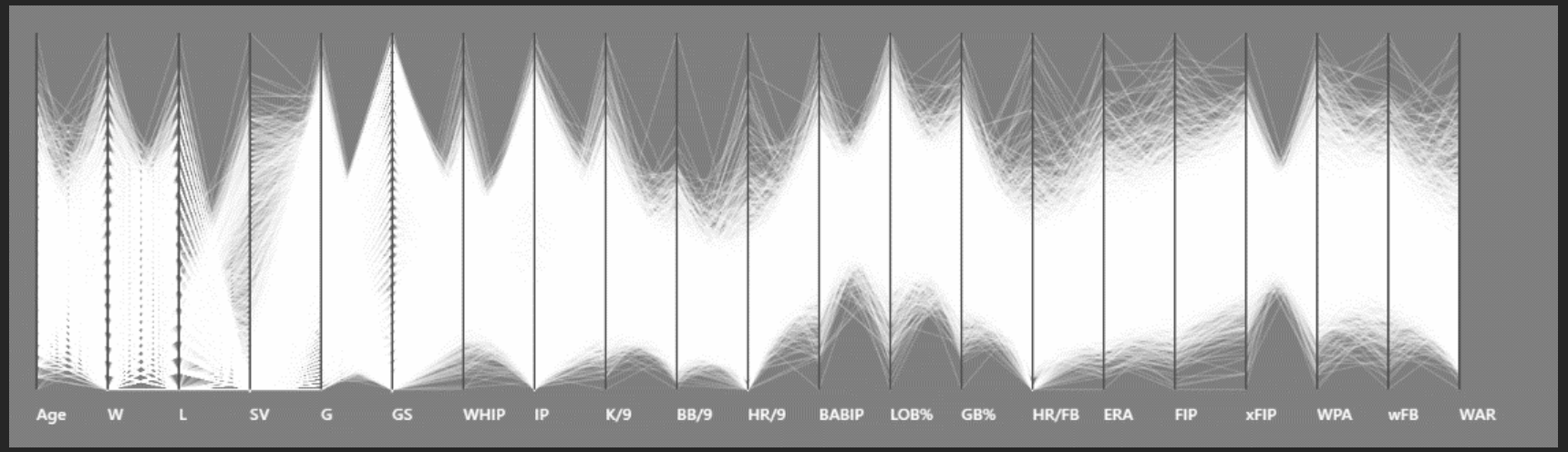


How do we deal with the innate limitations, while preserving the advantages?

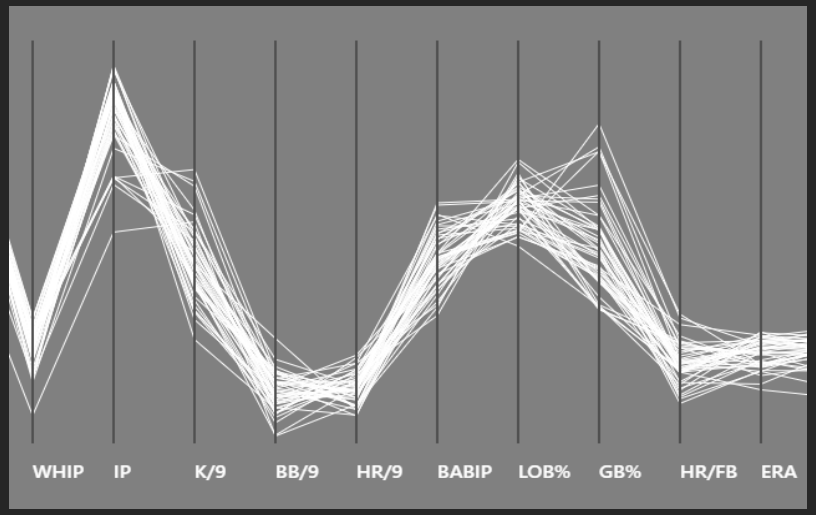
Unable to relationship between distant attributes

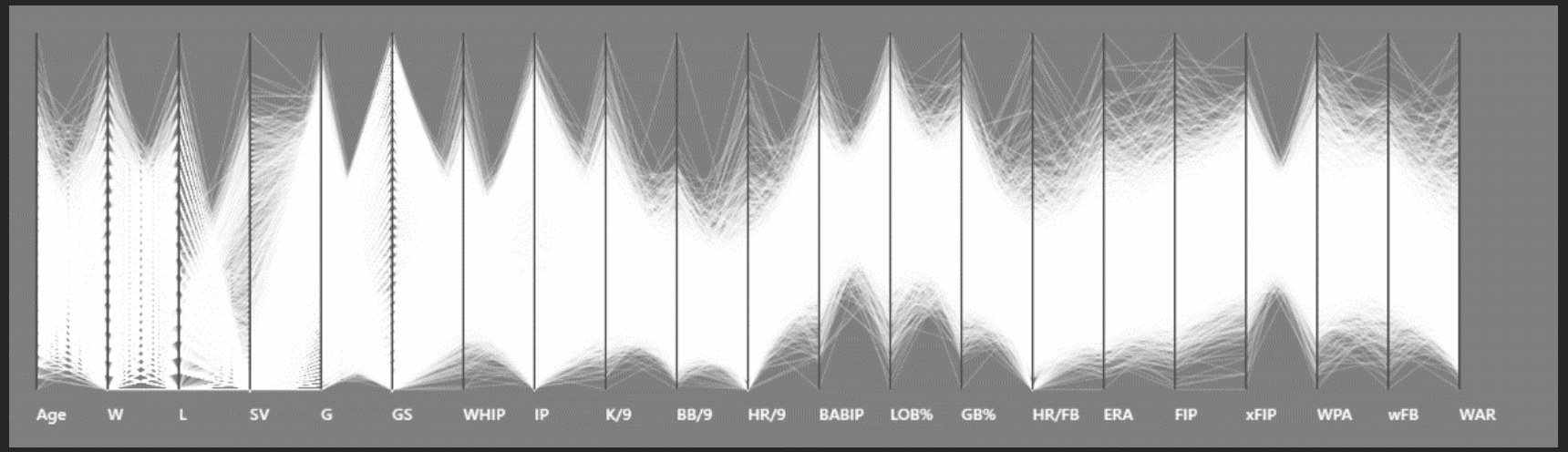
Approach


WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA

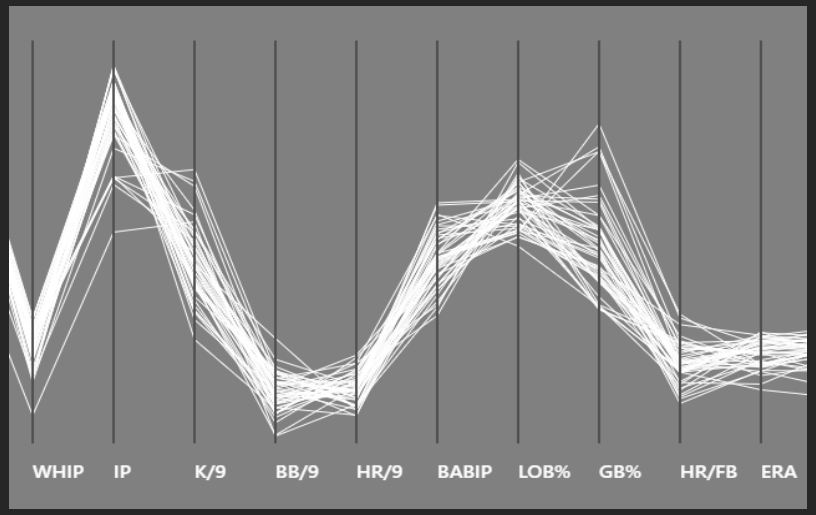

Age
W
L
SV
G
GS
WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA
FIP
xFIP
WPA
wFB
WAR

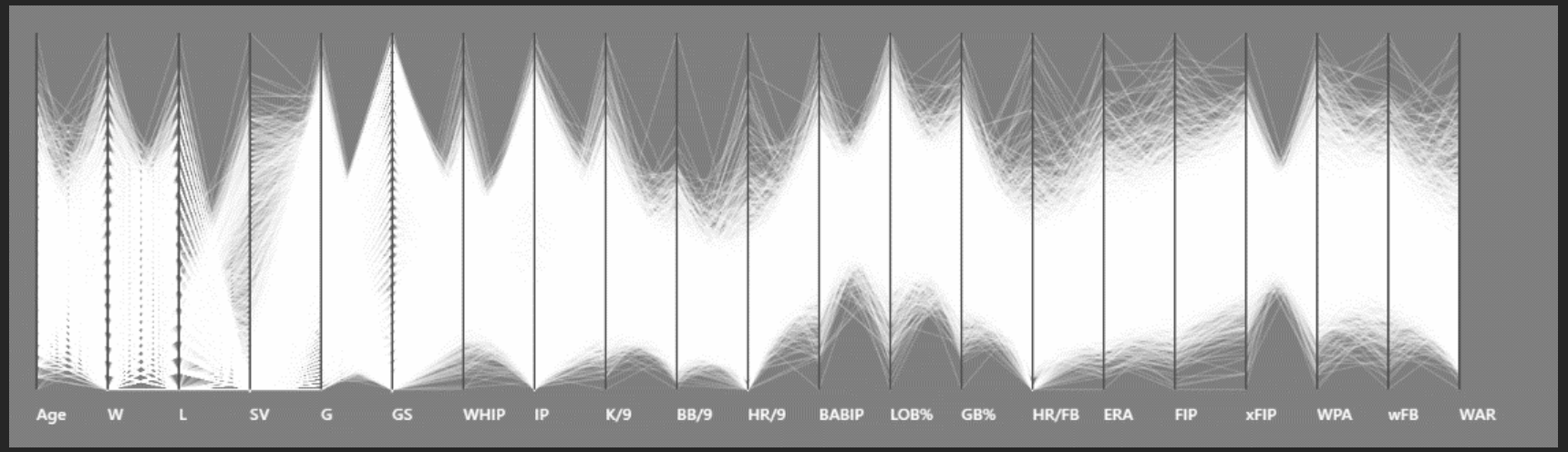
Approach


WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA

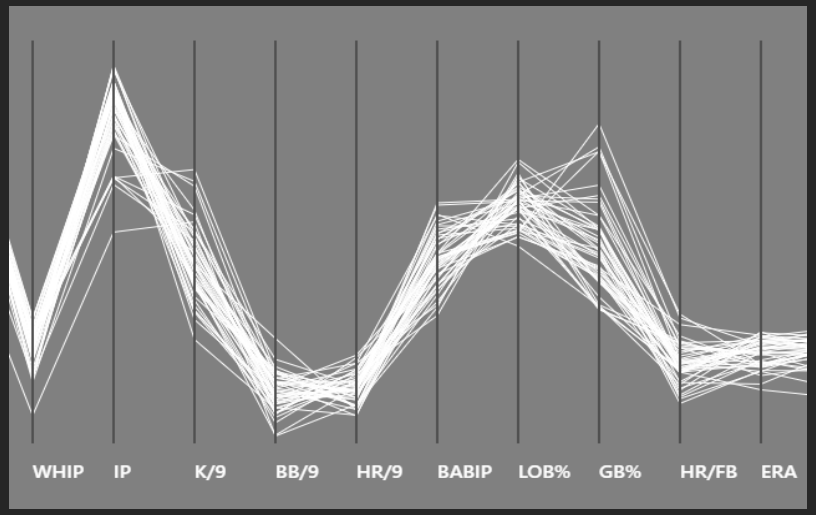

Age
W
L
SV
G
GS
WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA
FIP
xFIP
WPA
wFB
WAR

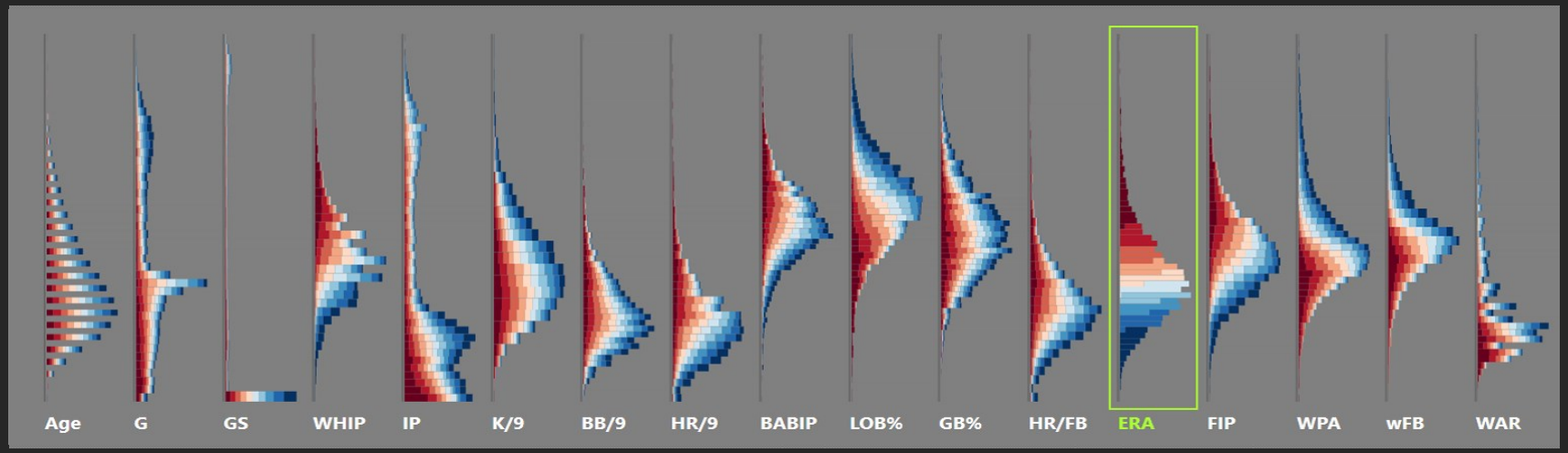
Approach


WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA


Age
W
L
SV
G
GS
WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA
FIP
xFIP
WPA
wFB
WAR

Approach


WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA


Age
G
GS
WHIP
IP
K/9
BB/9
HR/9
BABIP
LOB%
GB%
HR/FB
ERA
FIP
WPA
wFB
WAR

Approach

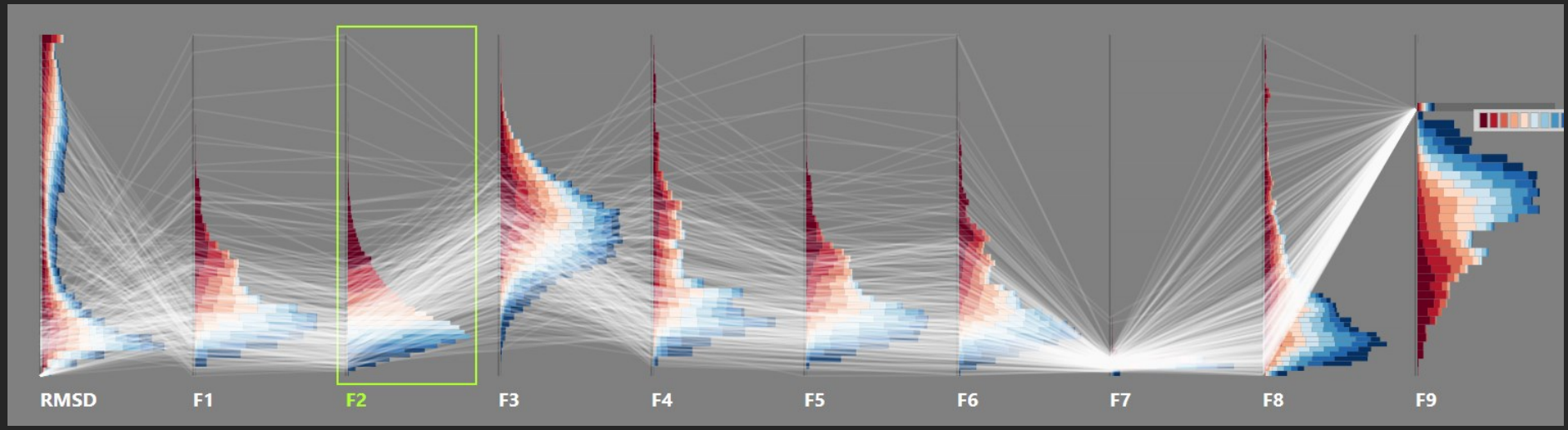
- The disadvantages of PCP occur when data size is large
  - So, what if we use another view to augment PCP when cluttering is severe?
  - Use PCP in limited conditions when cluttering subsides!
- Parallel Histogram Plots
- Following the Visual Information Seeking Mantra [Shneiderman, 1996]

  Overview first, zoom and filter, then details-on-demand
  - Overview with PHP, zoom and filter, detail with PCP

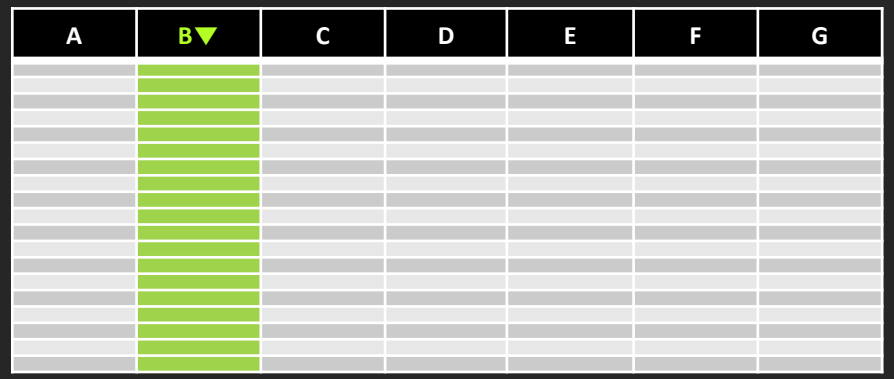
[Shneiderman, 1996] Shneiderman, B. (1996, September). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the 1996 IEEE Symposium on Visual Languages* (p. 336).

Parallel Histogram Plots [PHP]

- Attach histogram on each axis of PCP
  - Scalable distribution of each attribute
- Apply colors to histograms
  - Relationship between attributes across distance

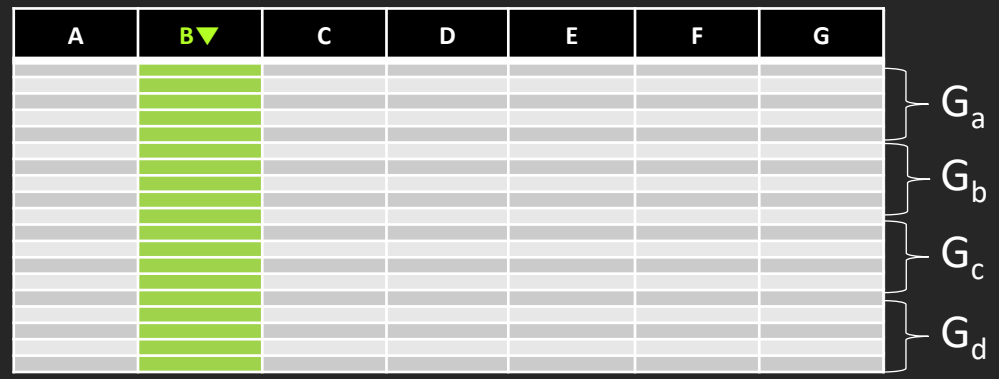


Constructing PHP



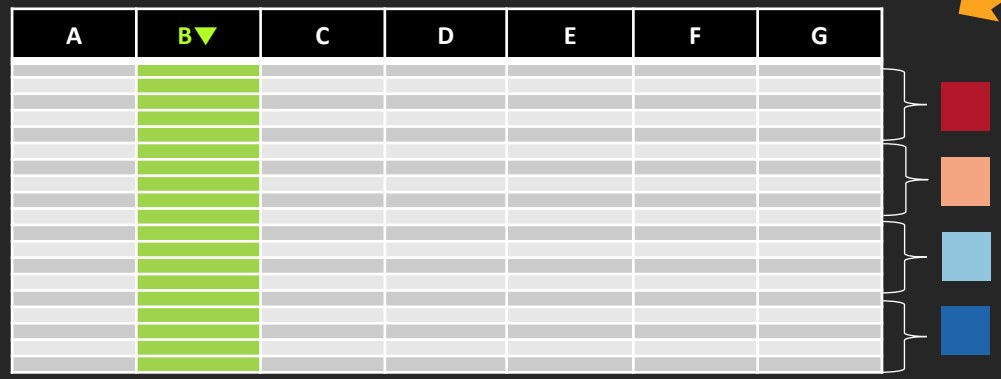
- Select one pivot attribute
- Sort data according the pivot attribute
  - Observe data from the perspective of the pivot attribute

Constructing PHP



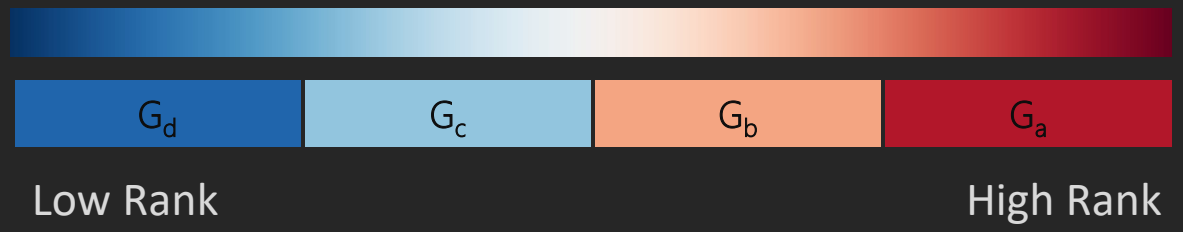
- Equally group items by sorted order
  - Each group corresponds to rank of items

Constructing PHP

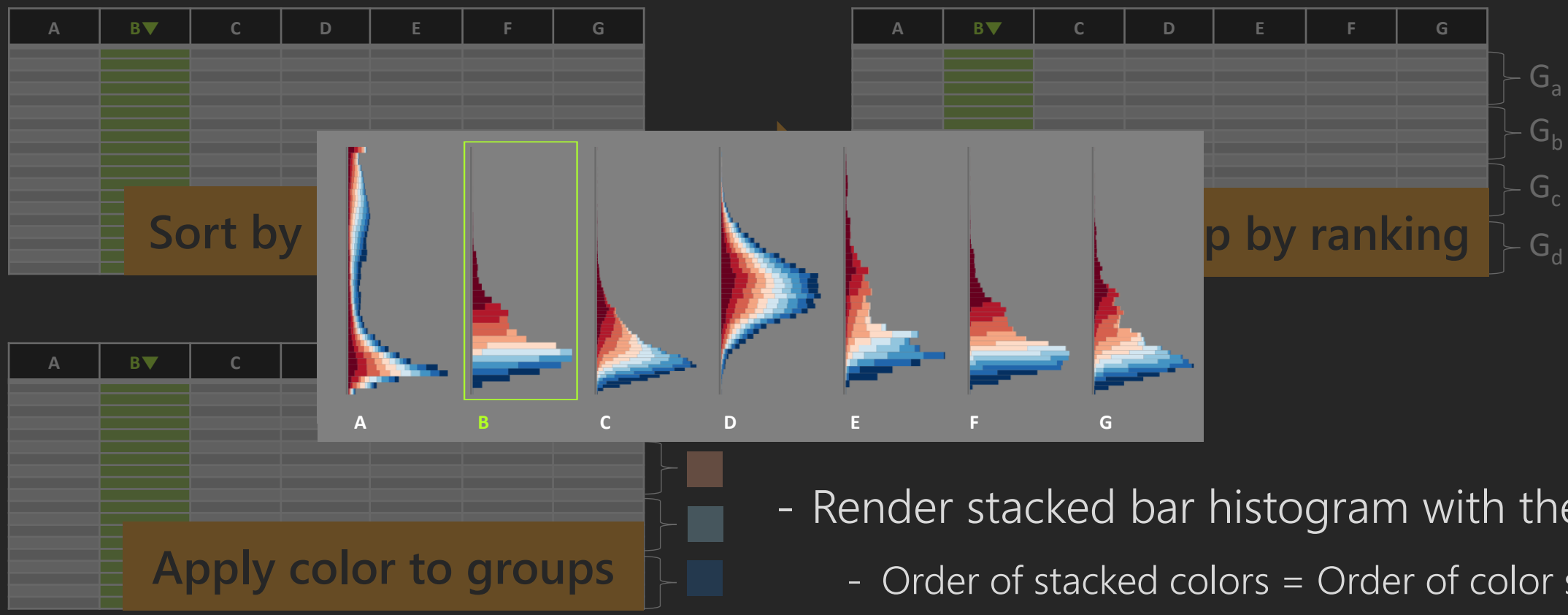


- Apply color to each group
  - Each group corresponds to rank of items

  => Each color corresponds to rank of items

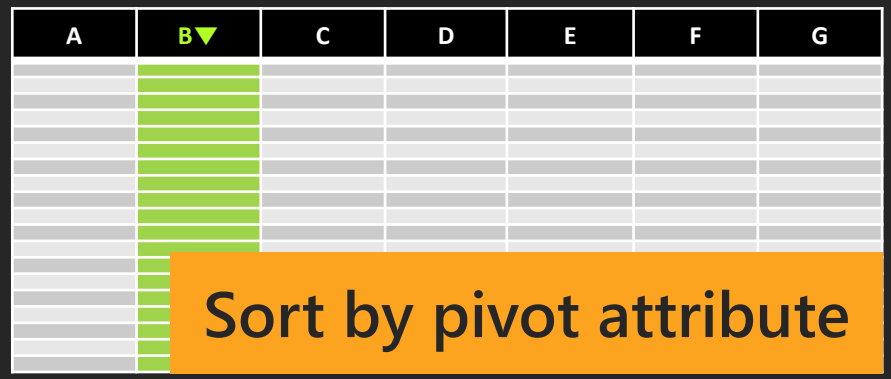


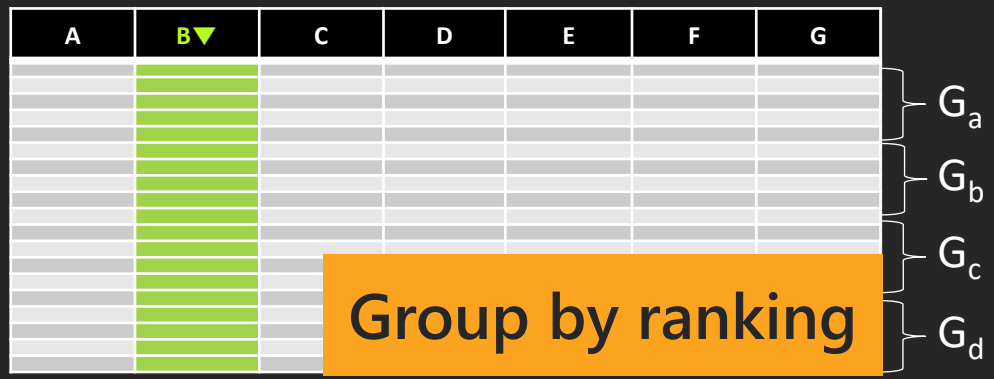
Constructing PHP

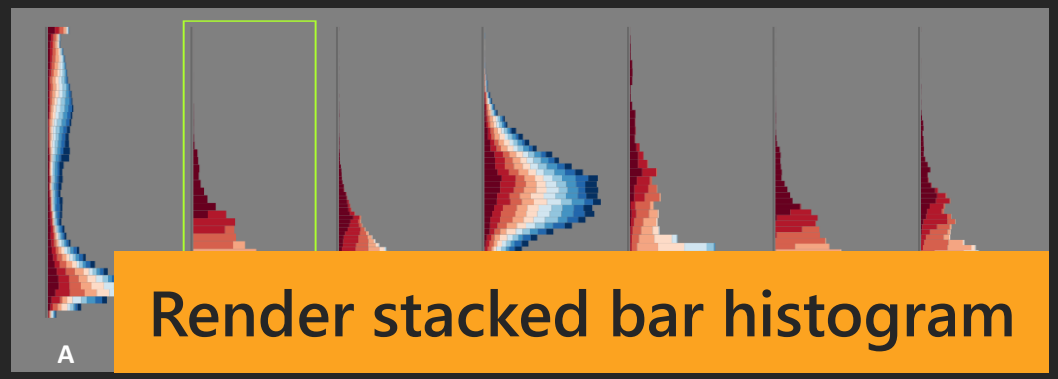


- Render stacked bar histogram with the groups
  - Order of stacked colors = Order of color scheme
  - Heatmap-like property in each of histograms

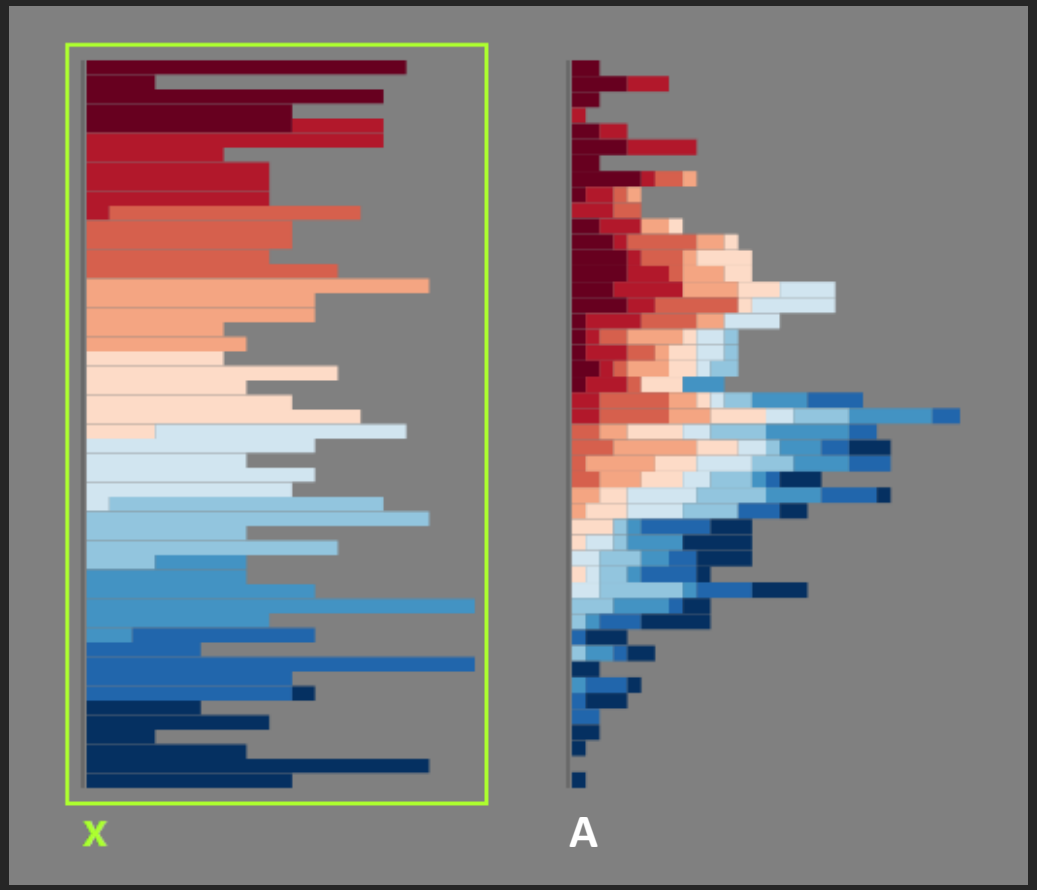
Constructing PHP



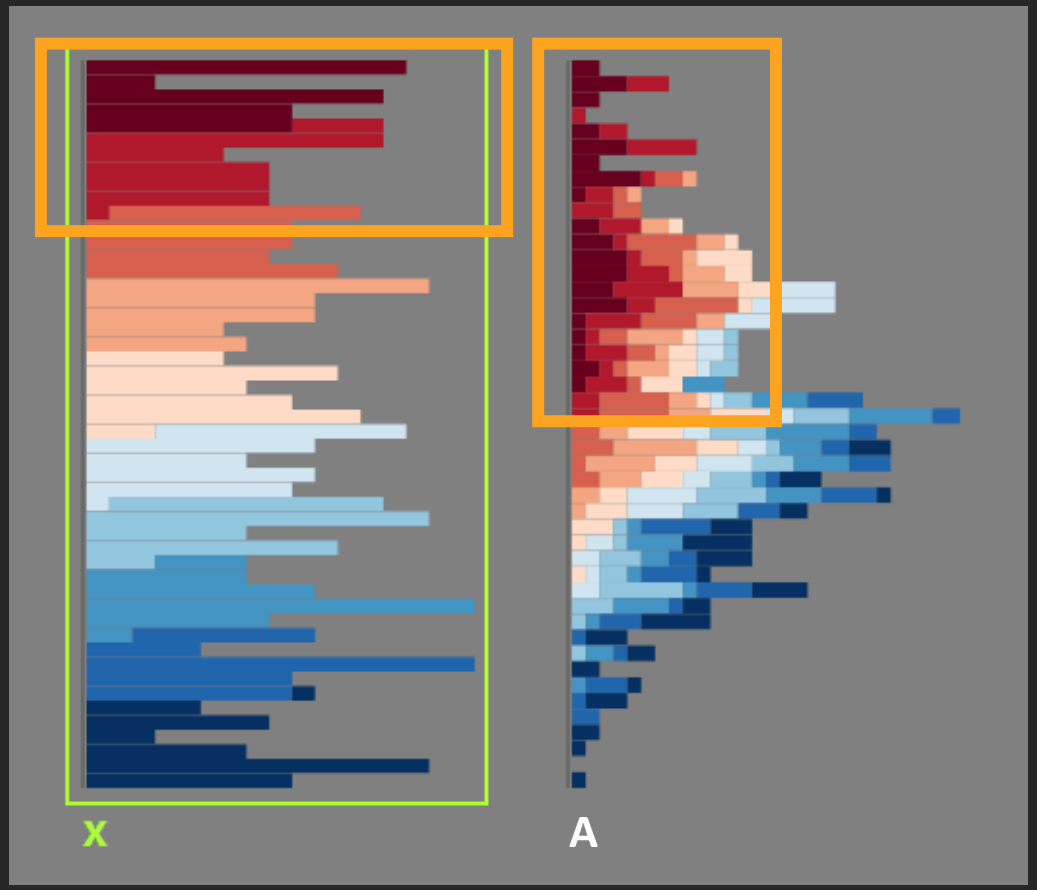




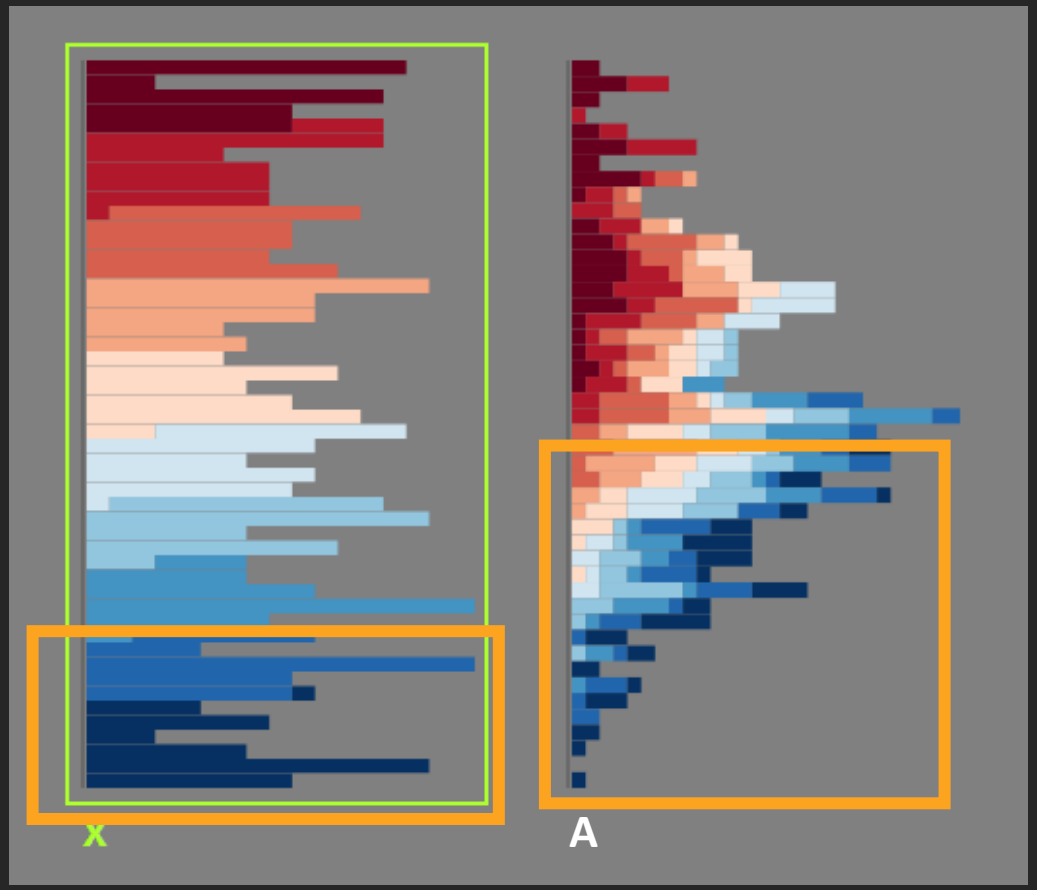
Interpreting PHP



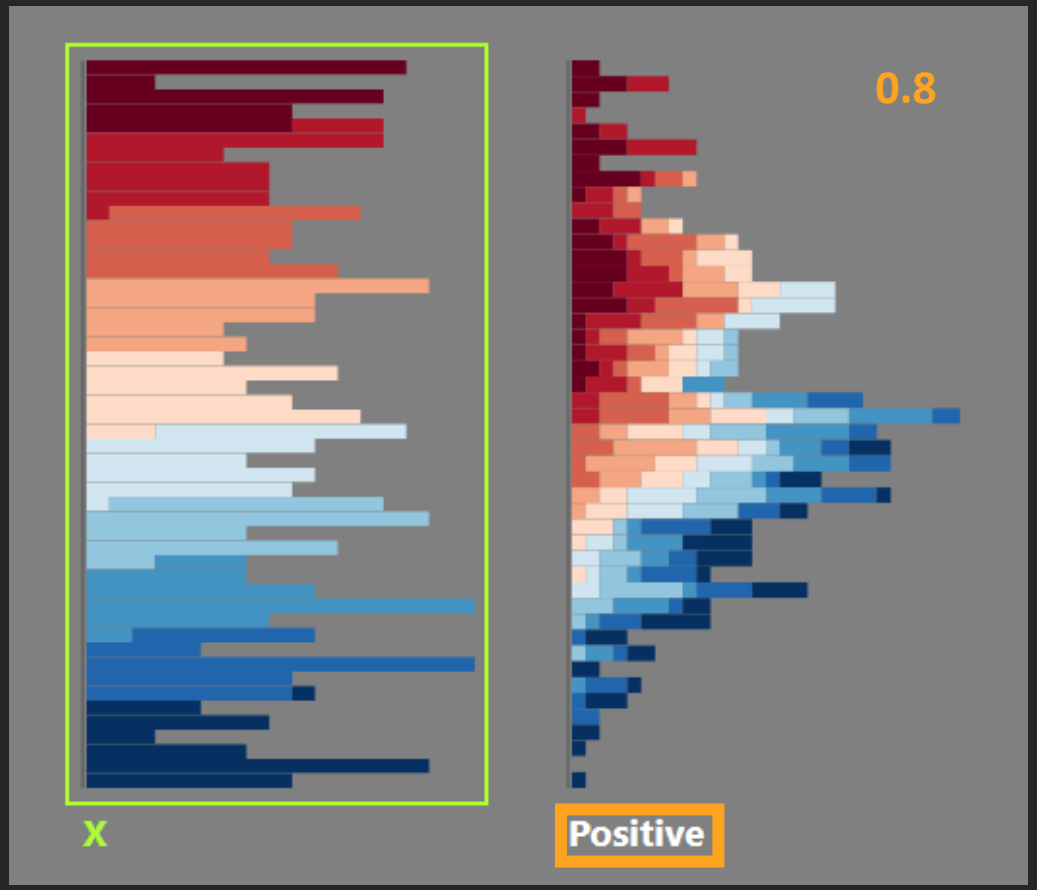
Interpreting PHP



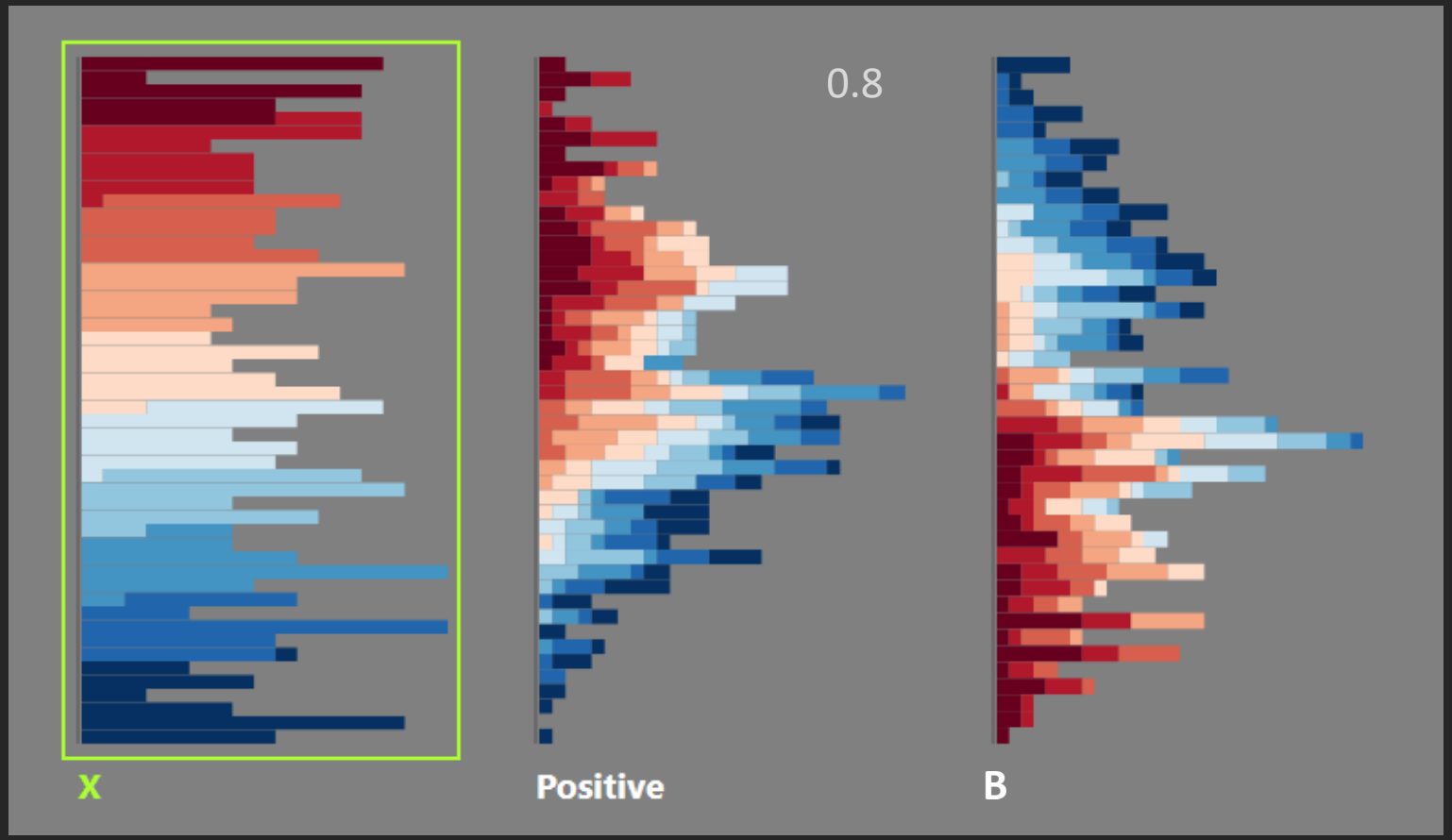
Interpreting PHP



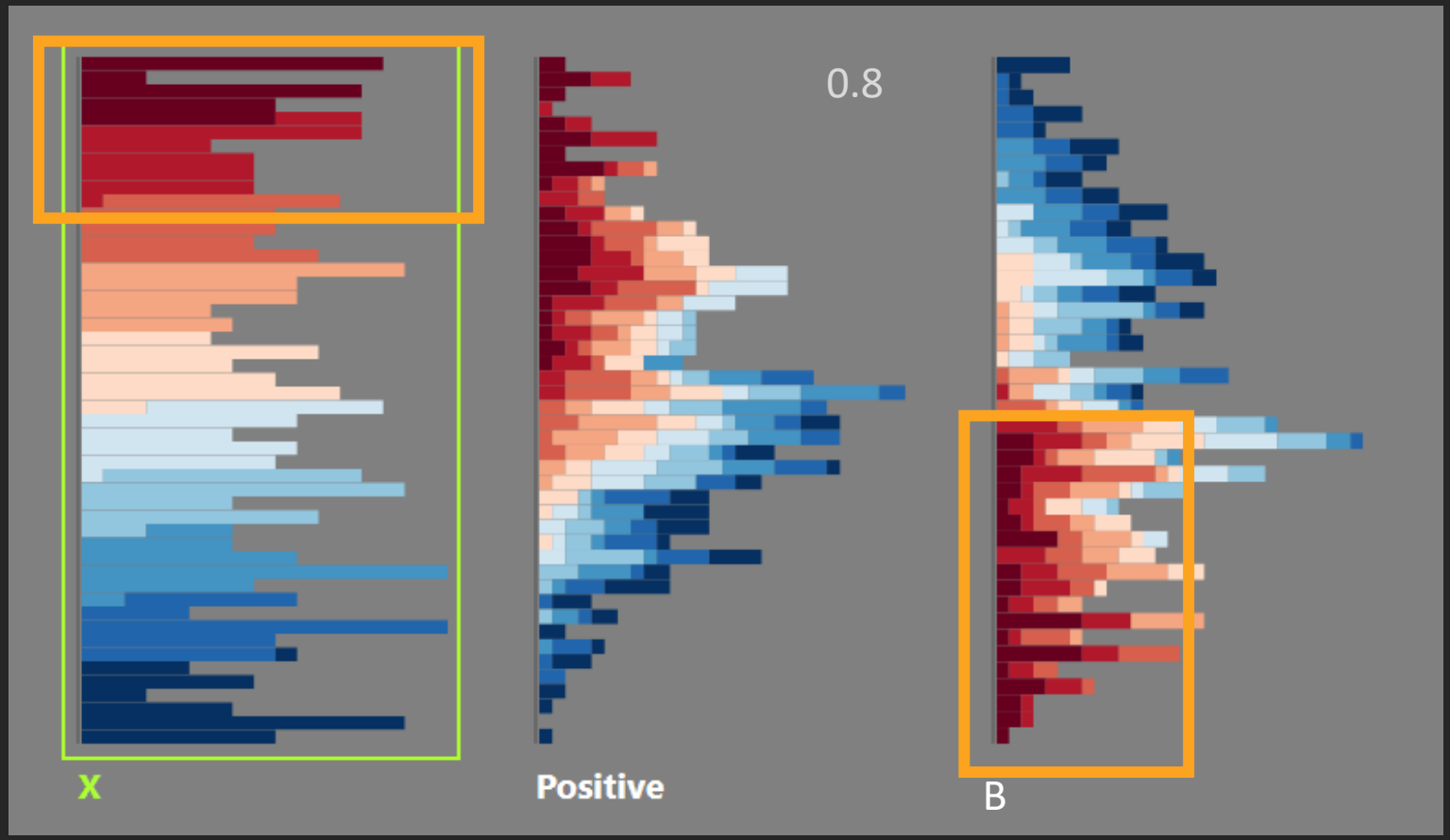
Interpreting PHP



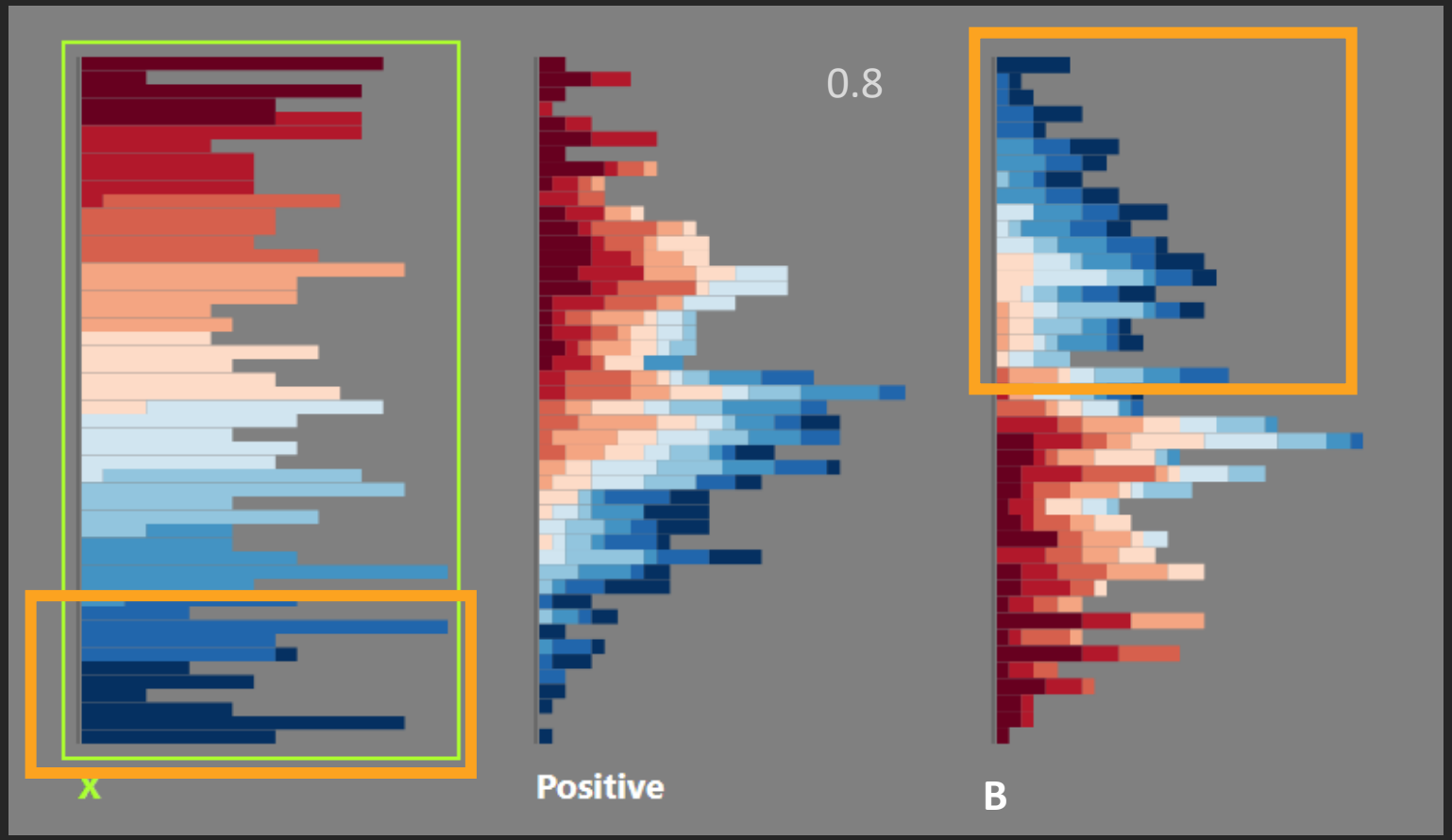
Interpreting PHP



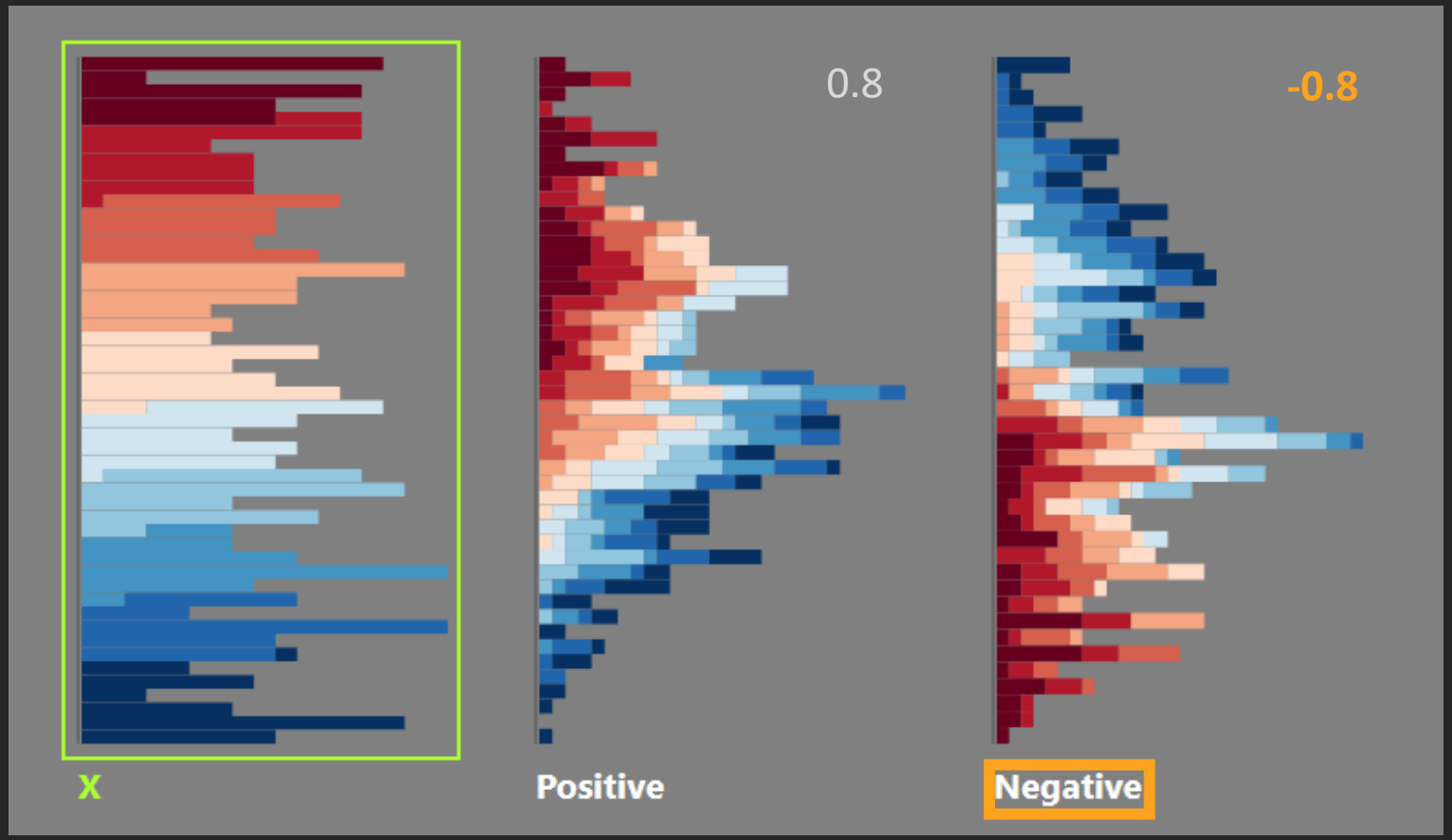
Interpreting PHP



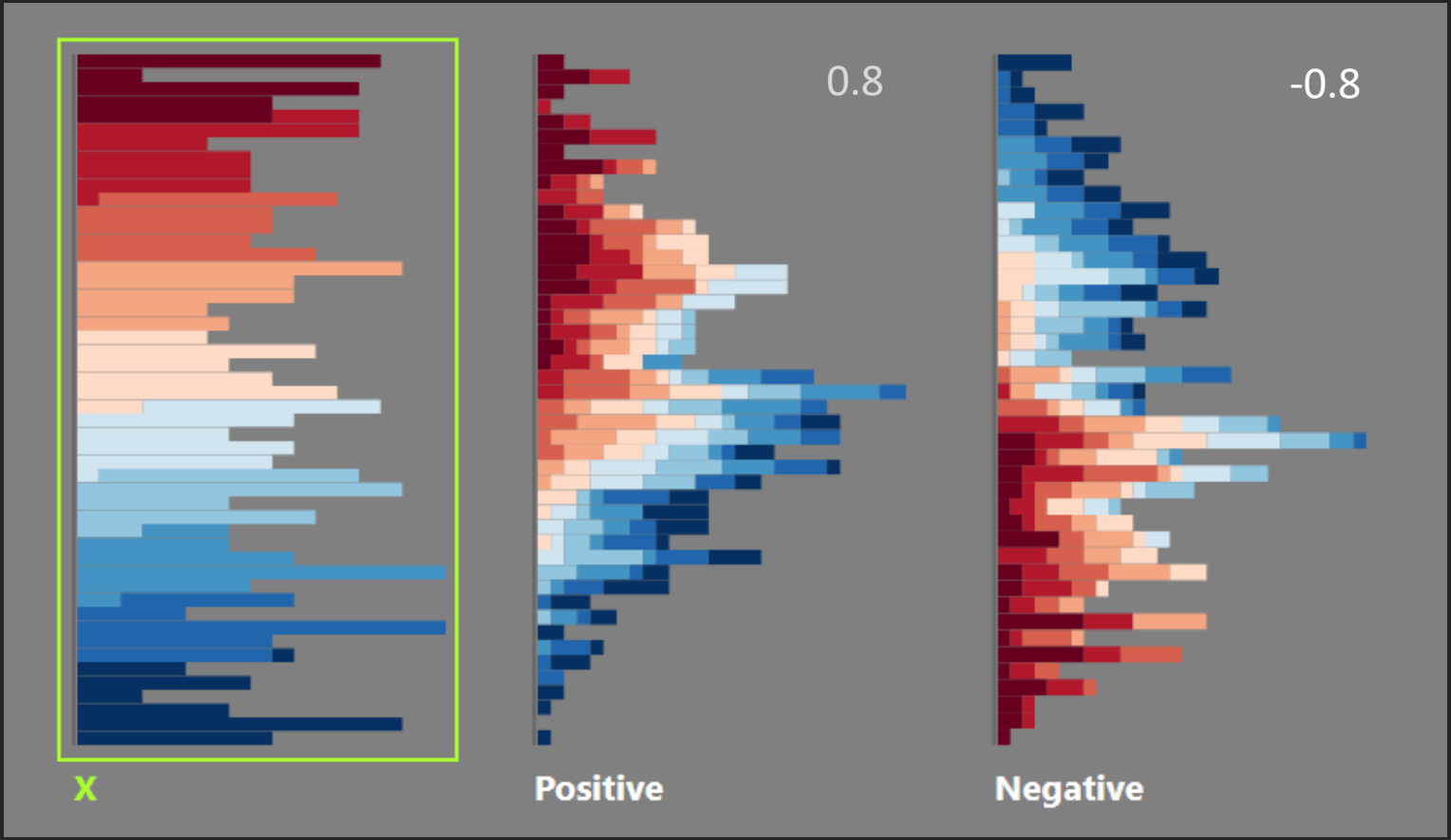
Interpreting PHP



Interpreting PHP

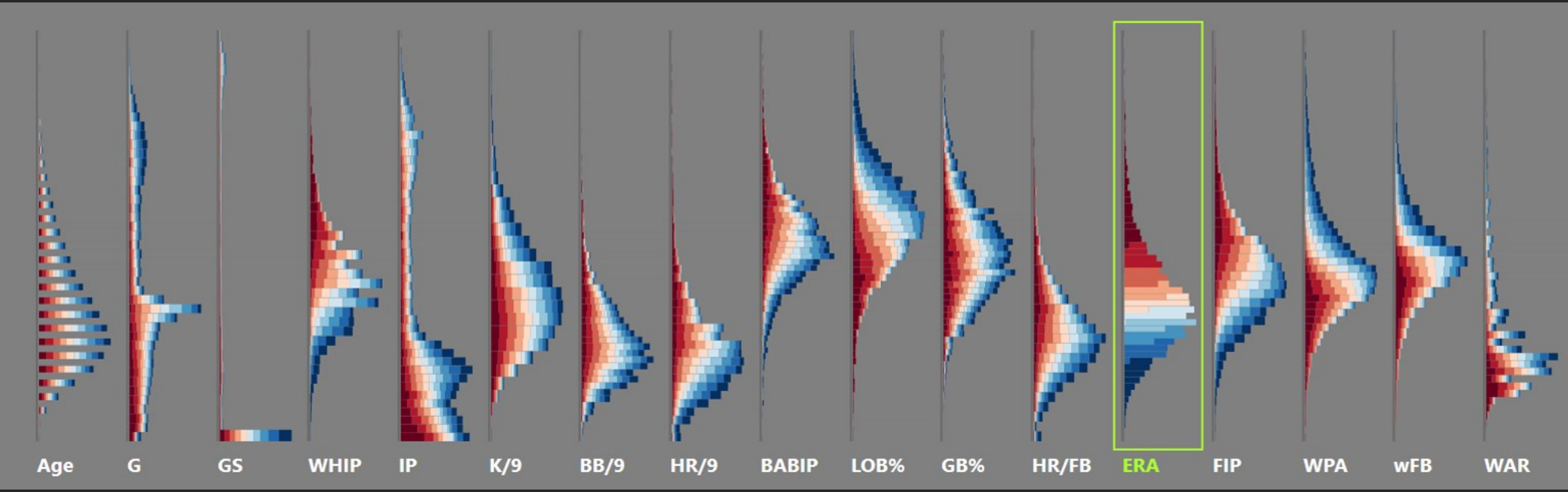


Interpreting PHP

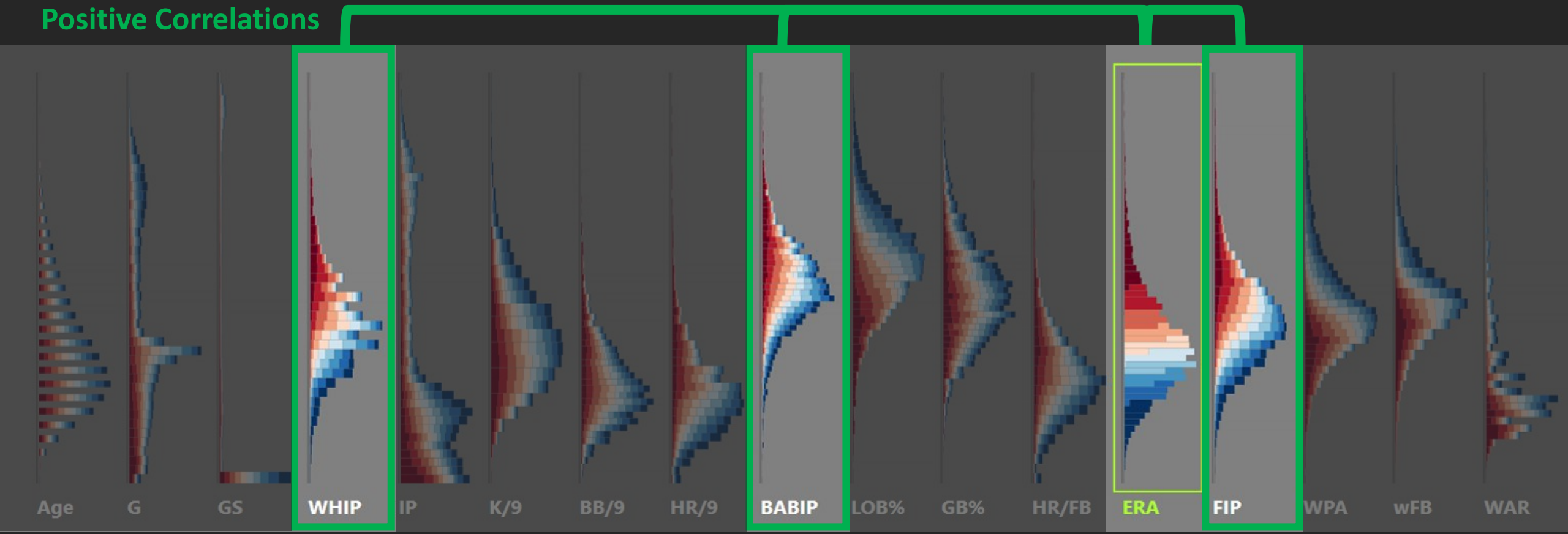


- Similar color of distribution to the pivot attribute
  => Positive correlation
- Inverted color of distribution to the pivot attribute
  => Negative correlation

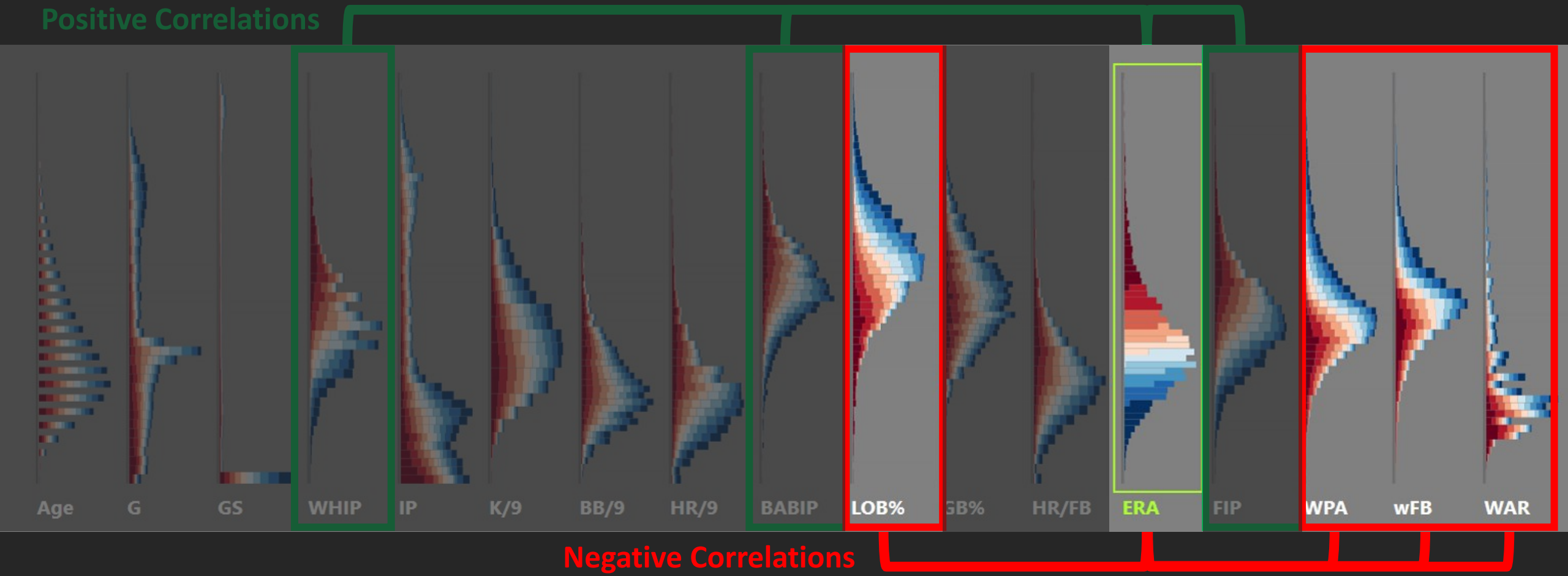
…In Many Attributes



...In Many Attributes

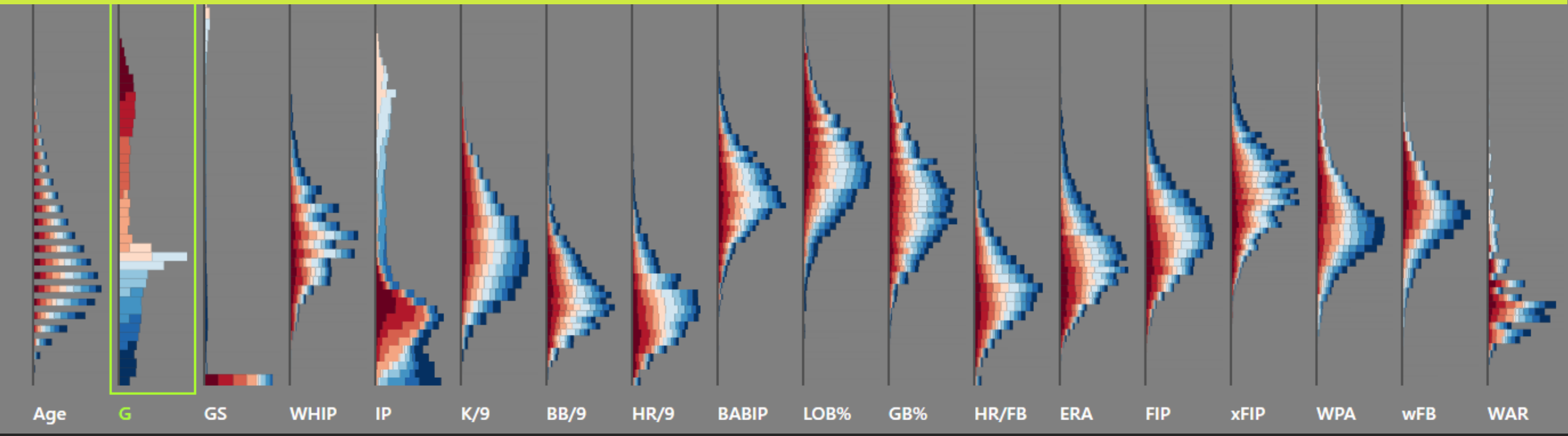


...In Many Attributes



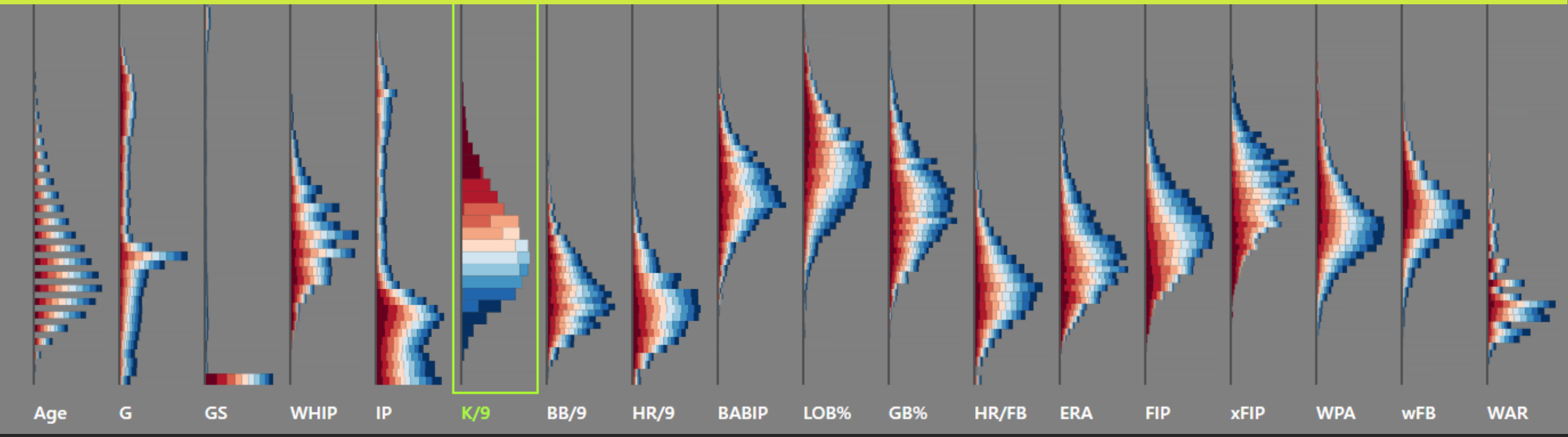
Changing Pivot Attributes

Observe data in different aspects by changing the pivot attribute



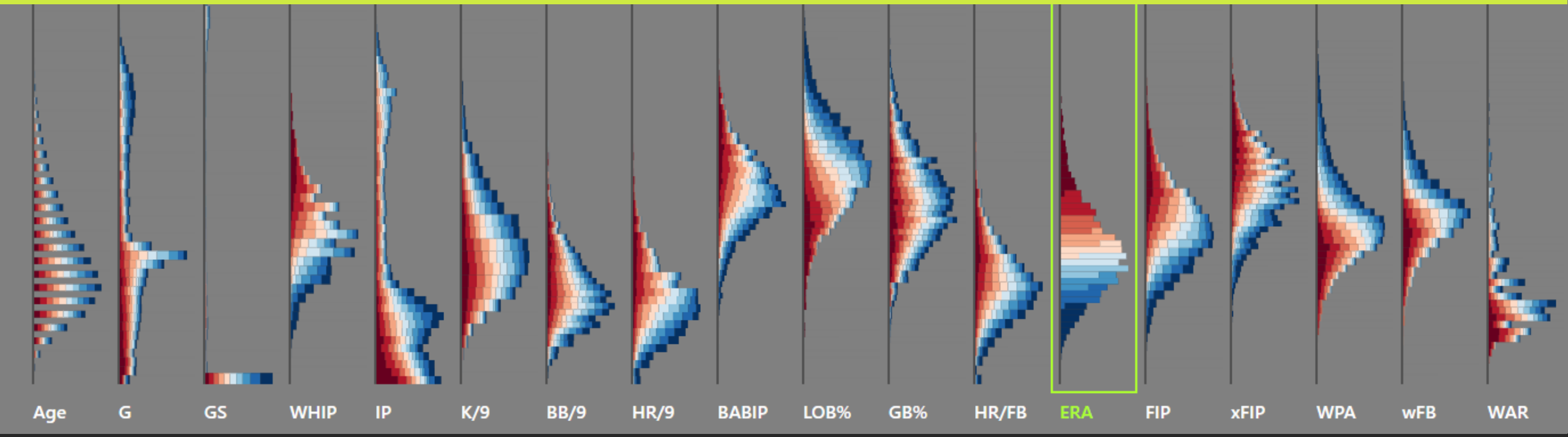
Changing Pivot Attributes

Observe data in different aspects by changing the pivot attribute



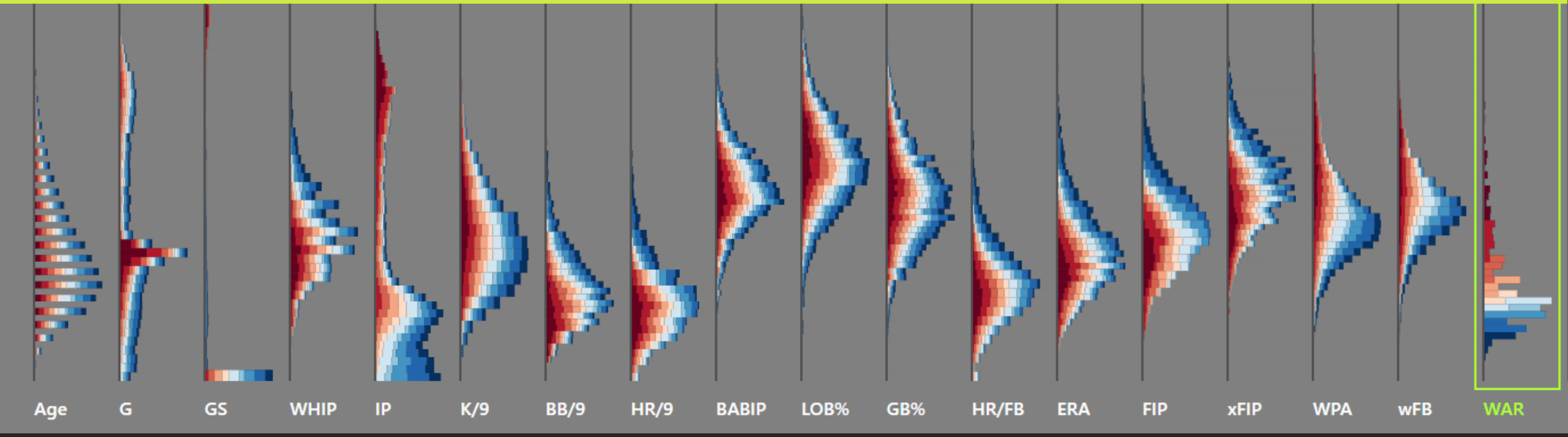
Changing Pivot Attributes

Observe data in different aspects by changing the pivot attribute

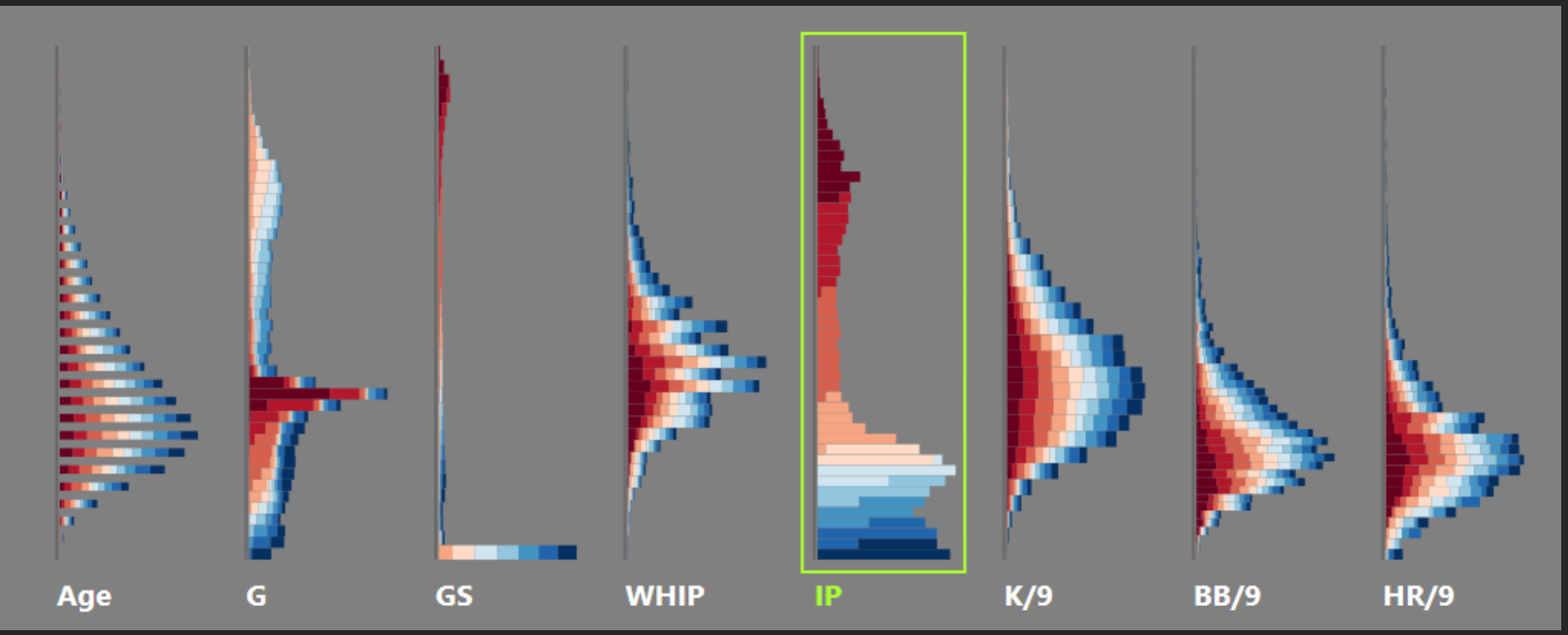


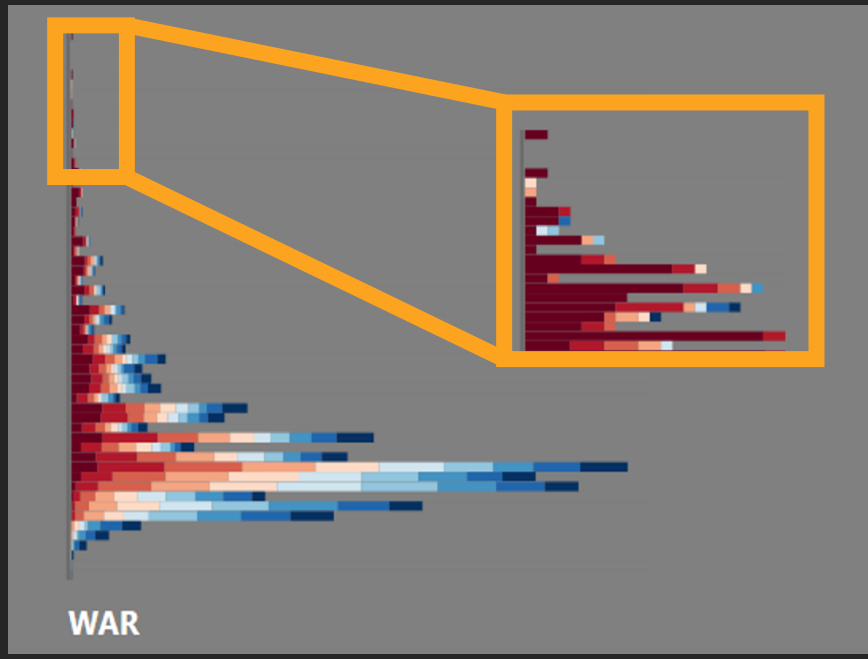
Changing Pivot Attributes

Observe data in different aspects by changing the pivot attribute

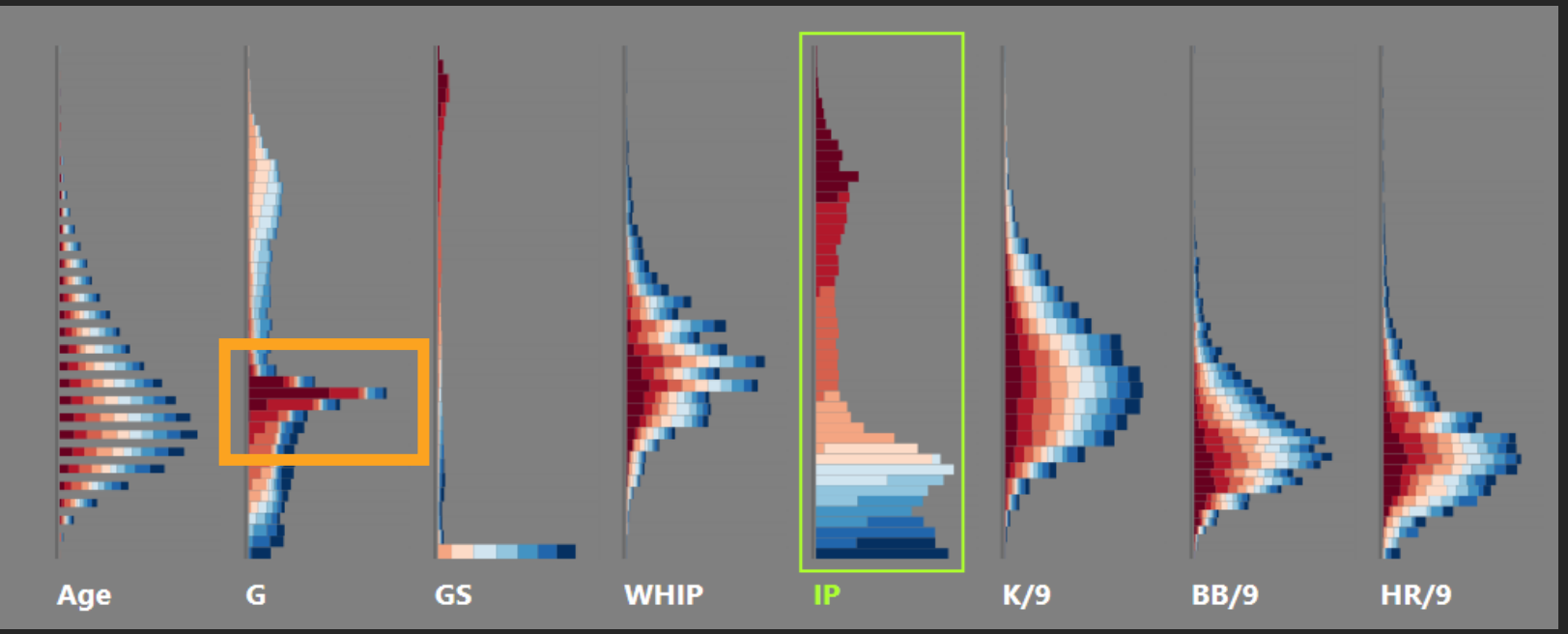


Groups and Outliers





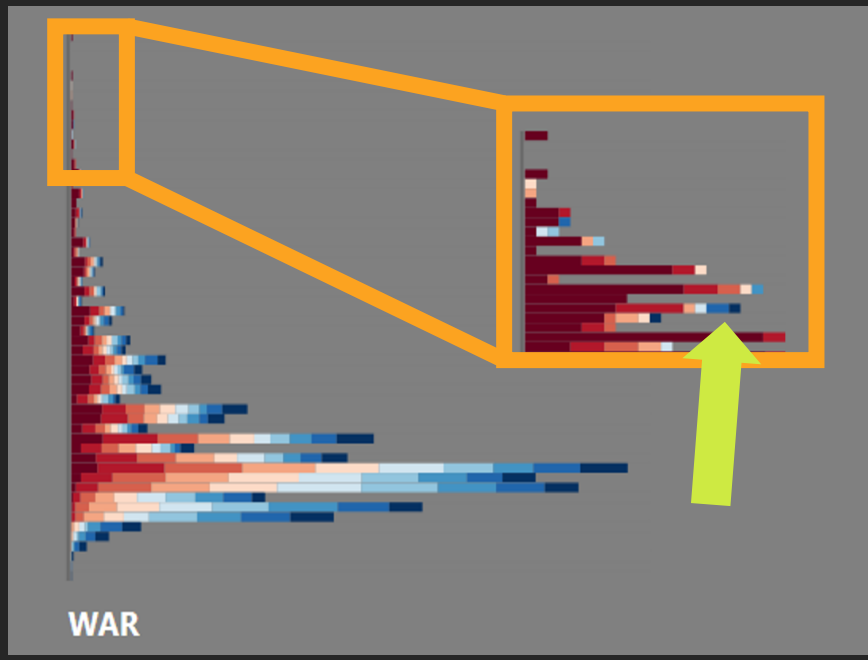
Groups and Outliers



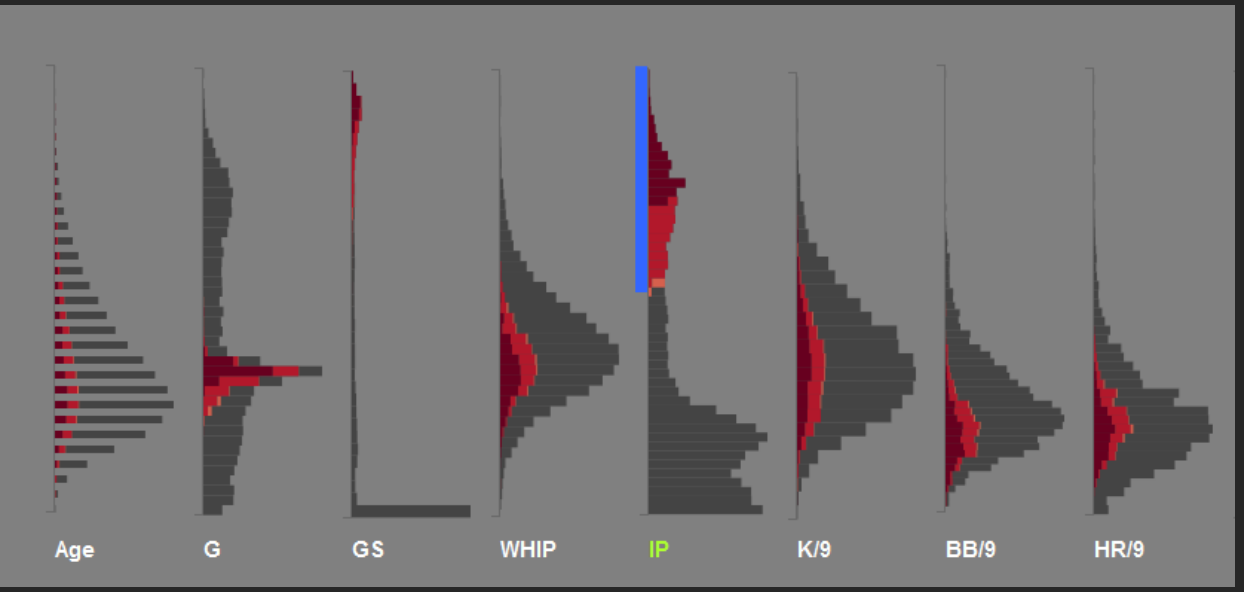
Gathered color indicate cluster of items with similar properties

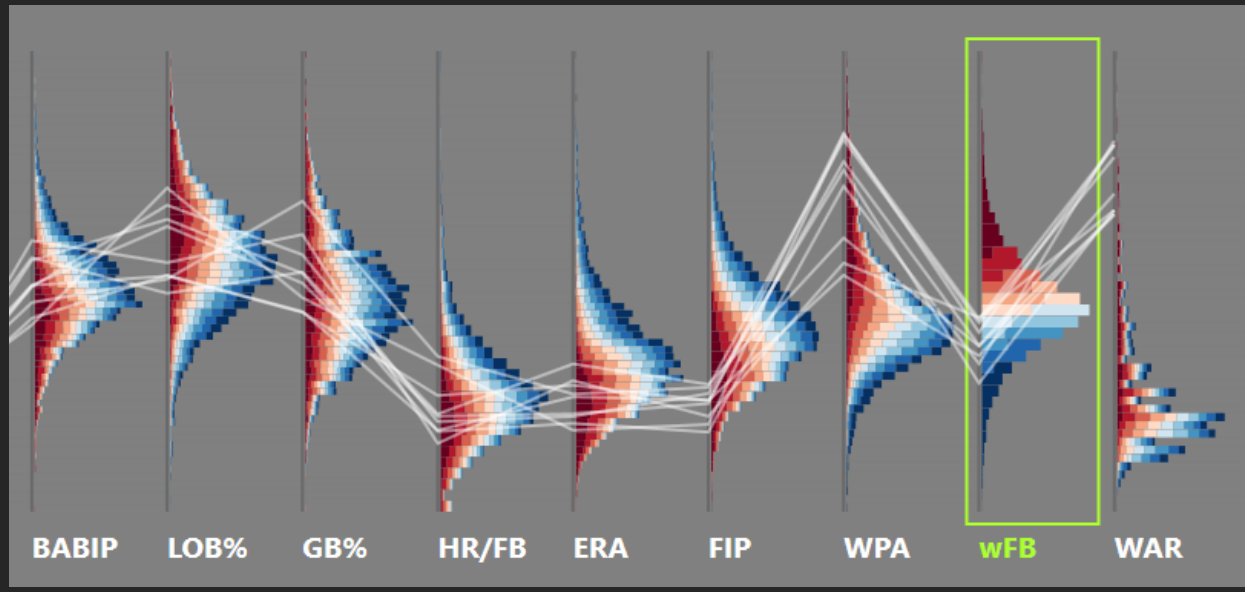
Groups and Outliers

Salient colors suggest outliers with different trends

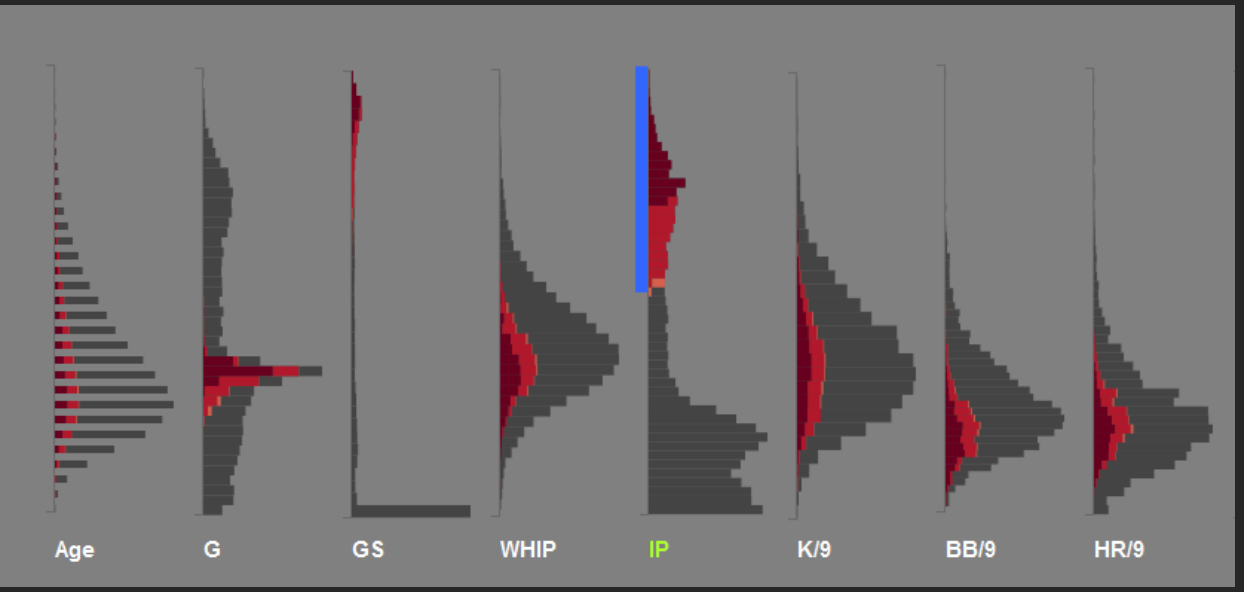


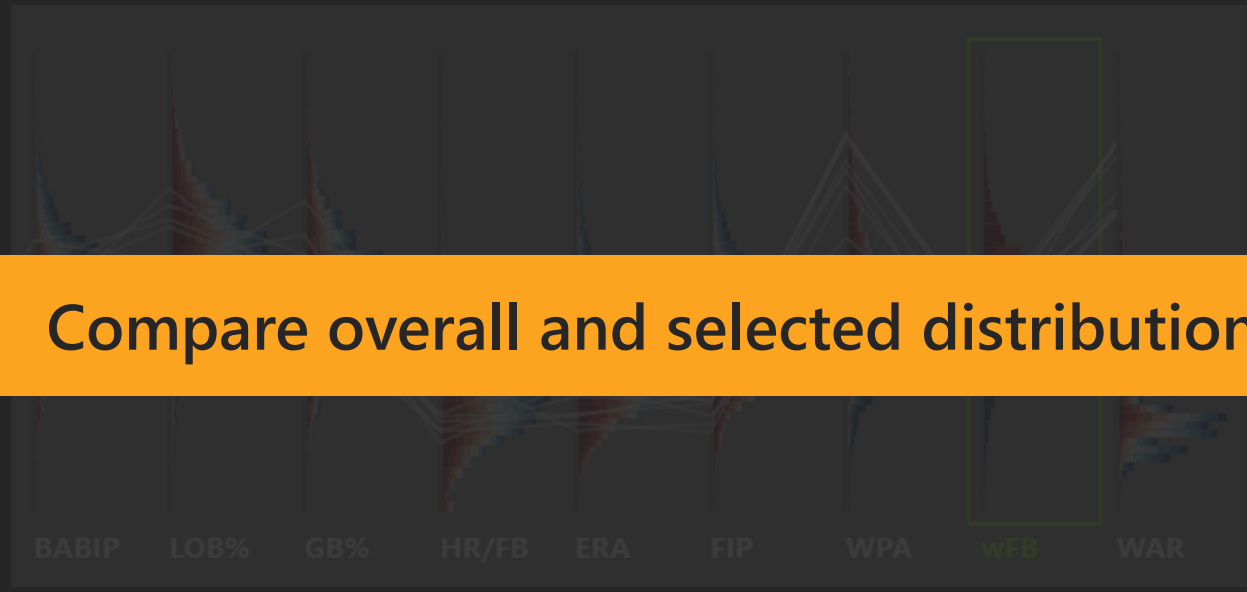
Zoom and Filter, Details on Demand





Zoom and Filter, Details on Demand

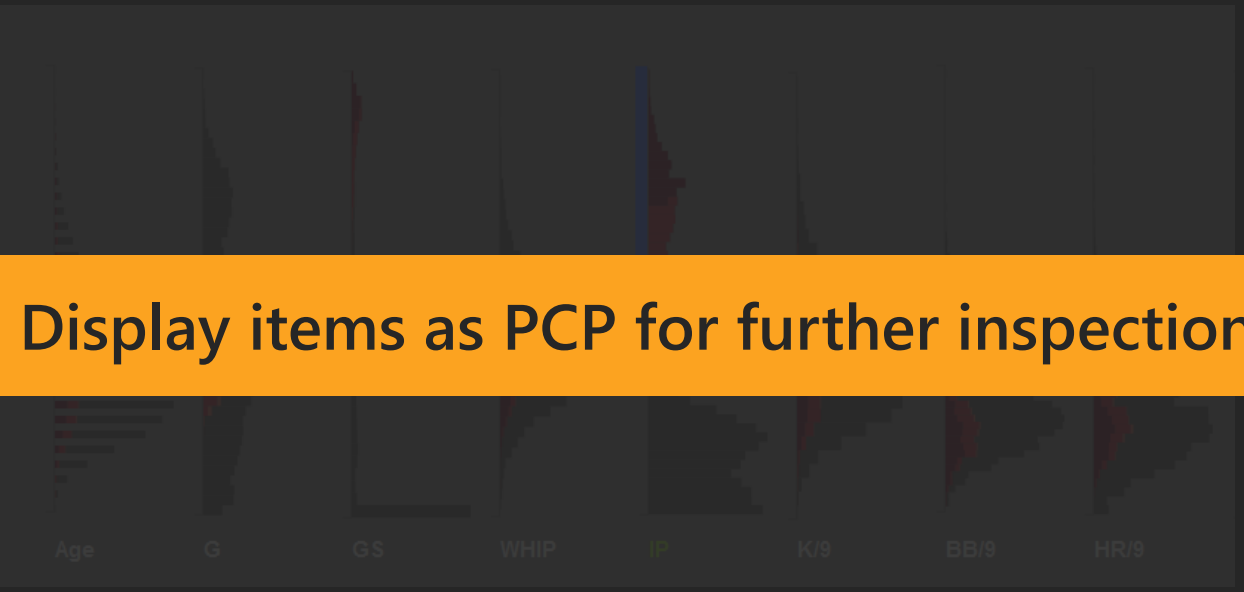


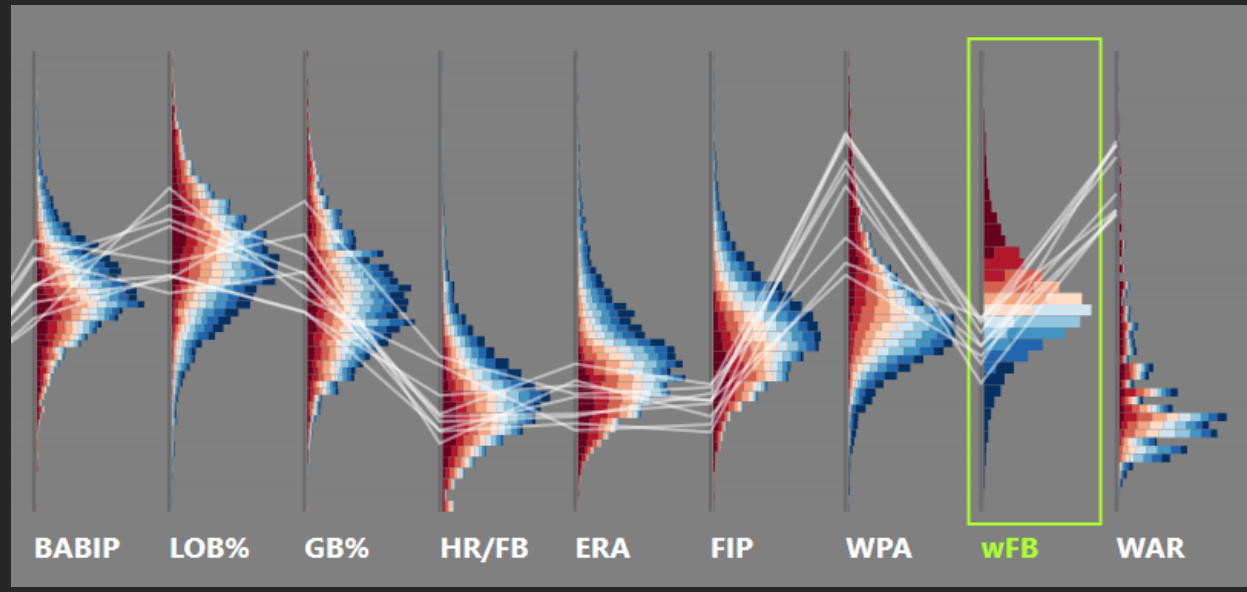


Compare overall and selected distribution

Zoom and Filter, Details on Demand

Display items as PCP for further inspection

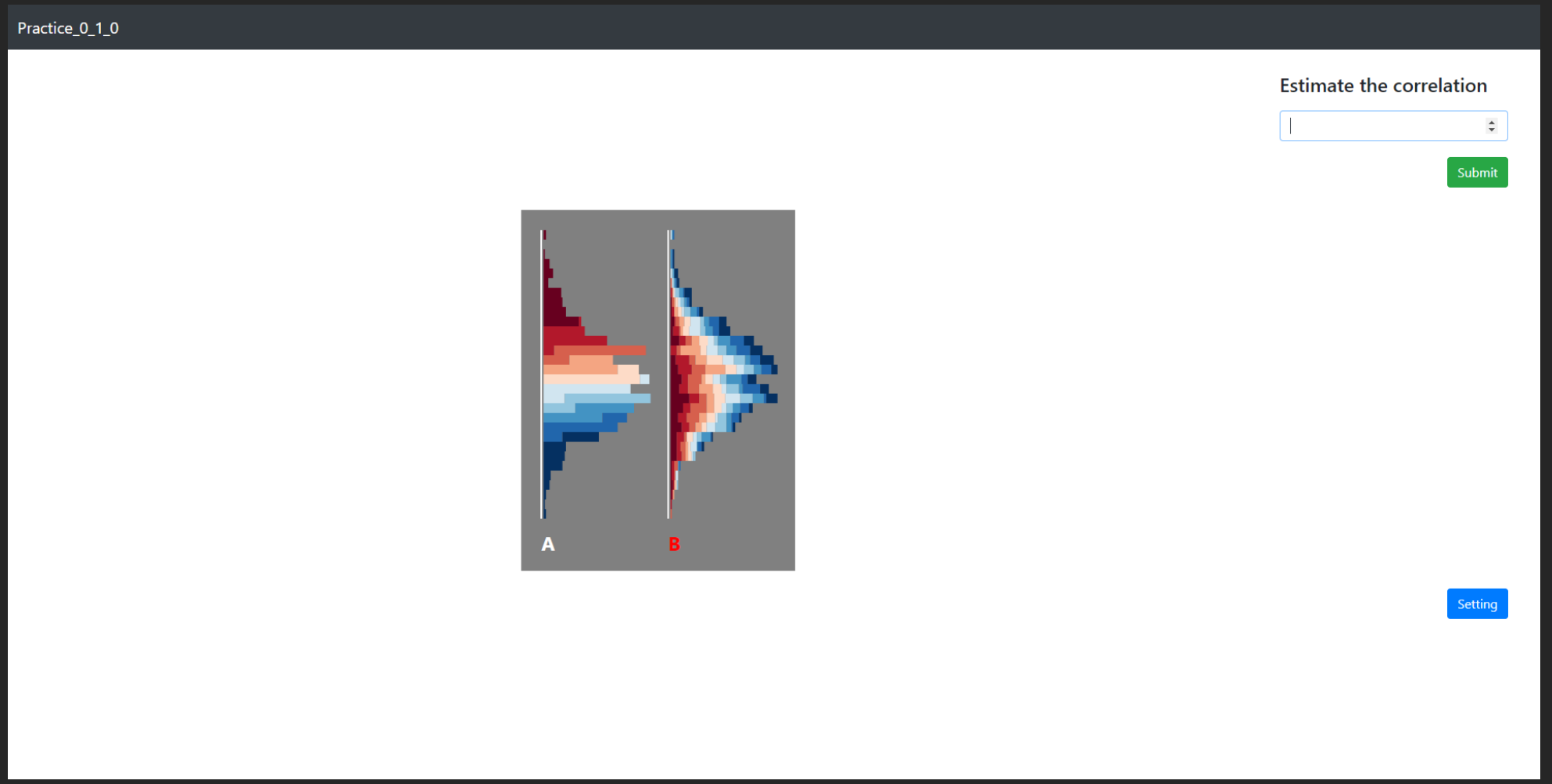




User Study

- Measure performance of PHP in terms of correlation coefficient retrieval, compared to previous techniques
- 36 participants
- 2 within-subject tasks
  - Correlation estimation task of two attributes
  - Correlation estimation task of two attributes, among multiple (4) attributes
- Measure accuracy and response time
  - Analyze with RM-ANOVA, Bonferroni's pairwise comparison for post hoc

Correlation Estimation Task

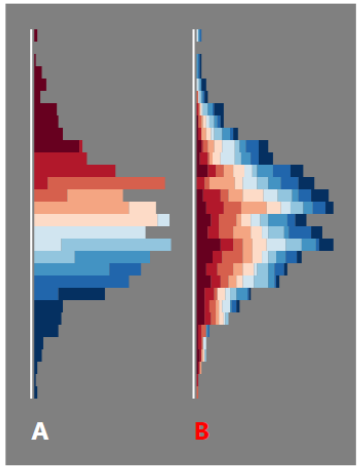


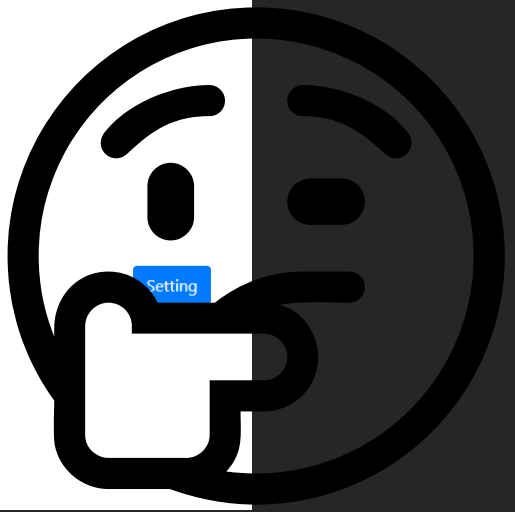
Correlation Estimation Task

Practice_0_1_0

Estimate the correlation

Submit



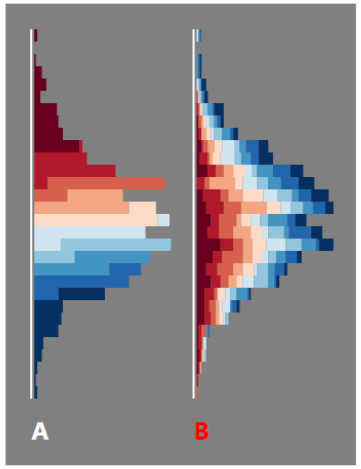


Correlation Estimation Task

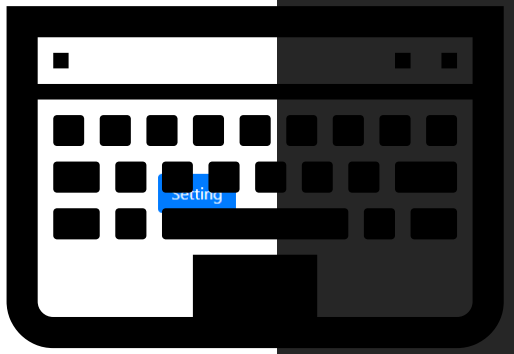
Practice_0_1_0

Estimate the correlation

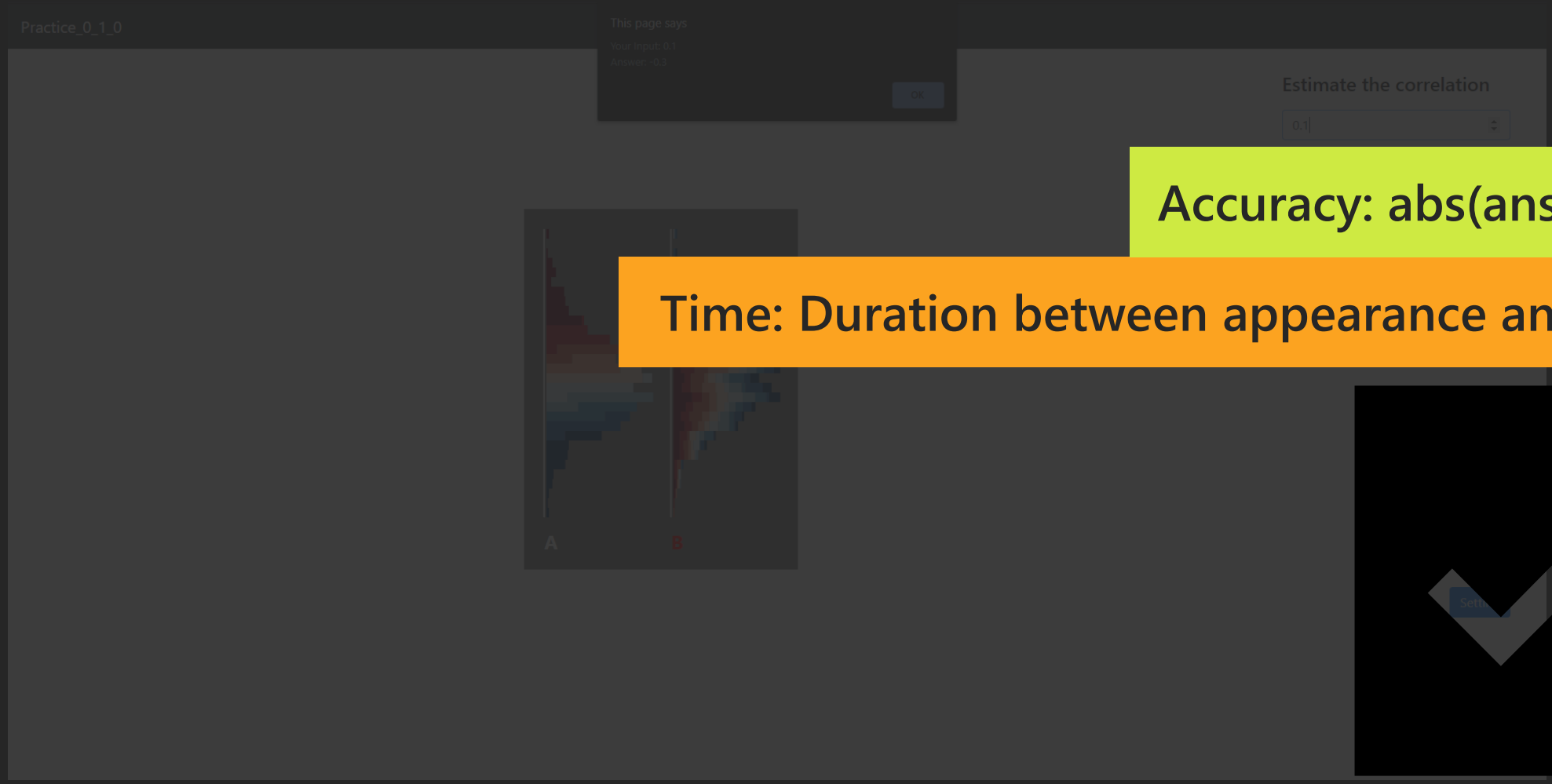
Submit



Answer by 0.1 interval

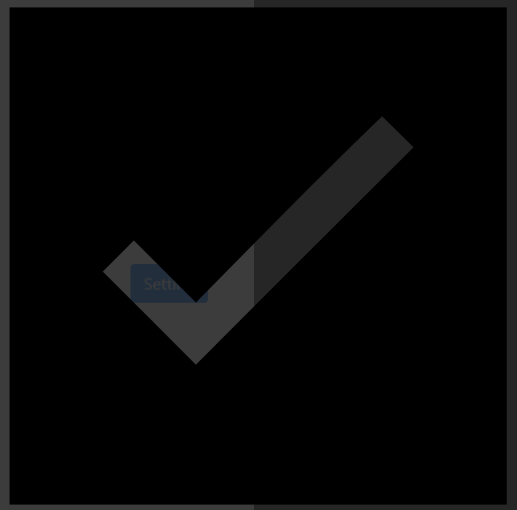


Correlation Estimation Task



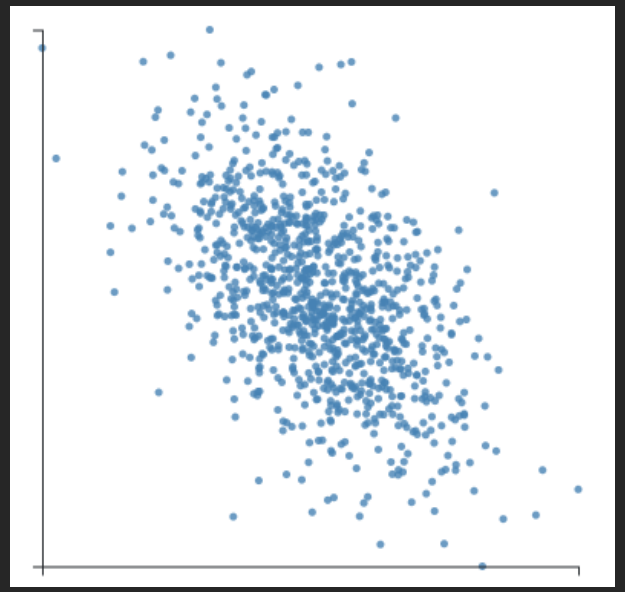
Accuracy: abs(answer – response)

Time: Duration between appearance and response (ms)

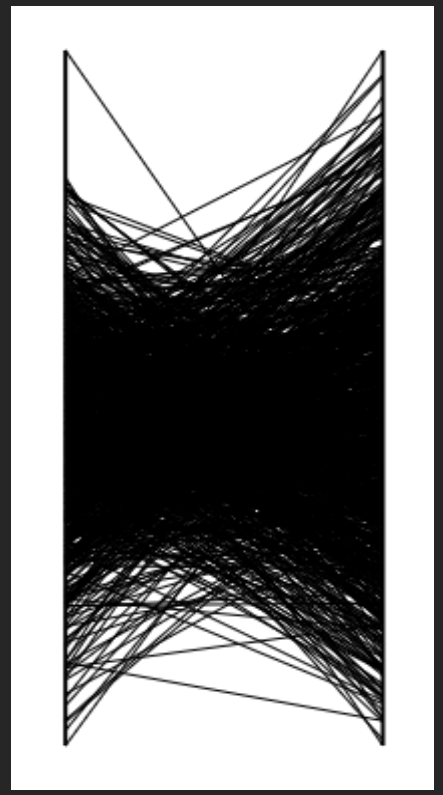


Target Visualizations

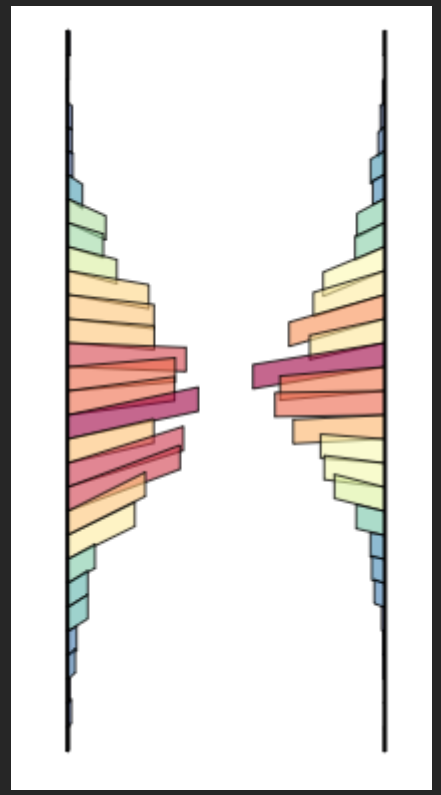
[Geng et. al, 2011] Geng, Z., Peng, Z., Laramee, R. S., Roberts, J. C., & Walker, R. (2011). Angular histograms: Frequency-based visualizations for large, high dimensional data. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2572-2580.



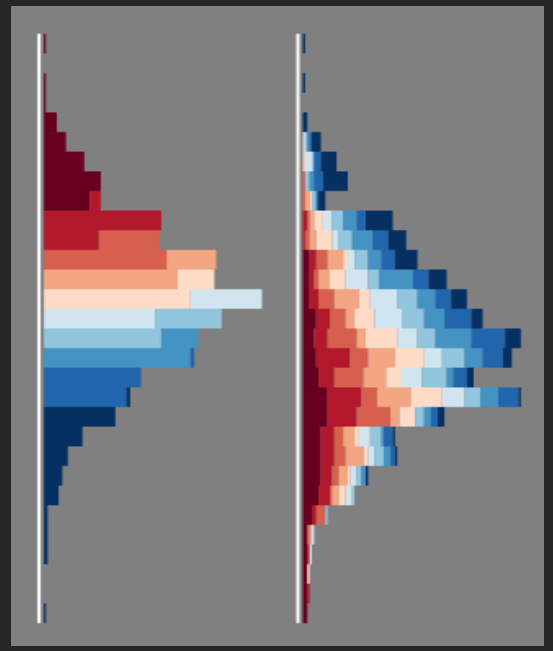
Scatterplot



PCP



Angular Histogram [Geng et. al, 2011]

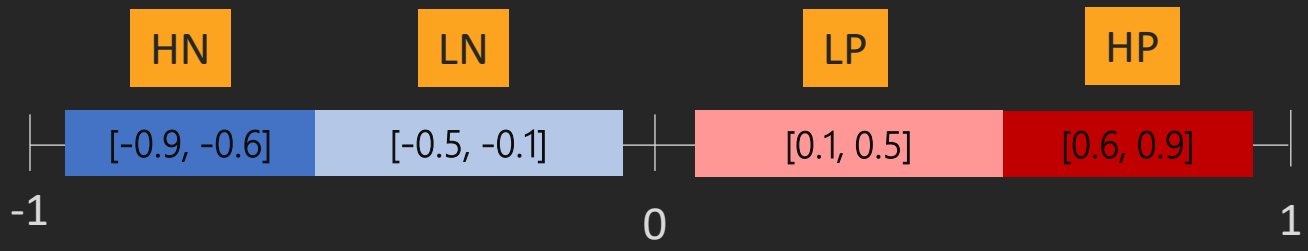


PHP

Task 1 [Two Attributes]

- Estimate correlation between two attributes, 1000 items
- 4 visualization techniques
  - PCP / Scatterplot / AH / PHP*, order balanced by Latin Square
- 4 correlation conditions
  - HN, LN, LP, HP
- Repeat each condition 5 times

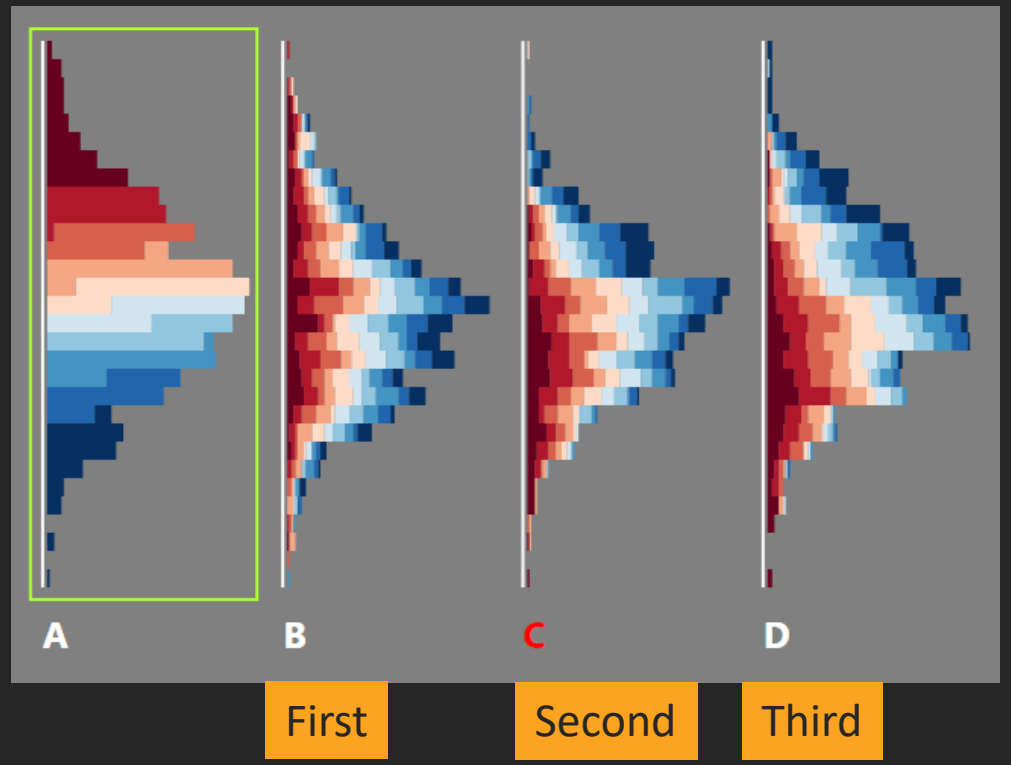
*Pivot attribute is always leftmost



Task 2 [Multiple Attributes]

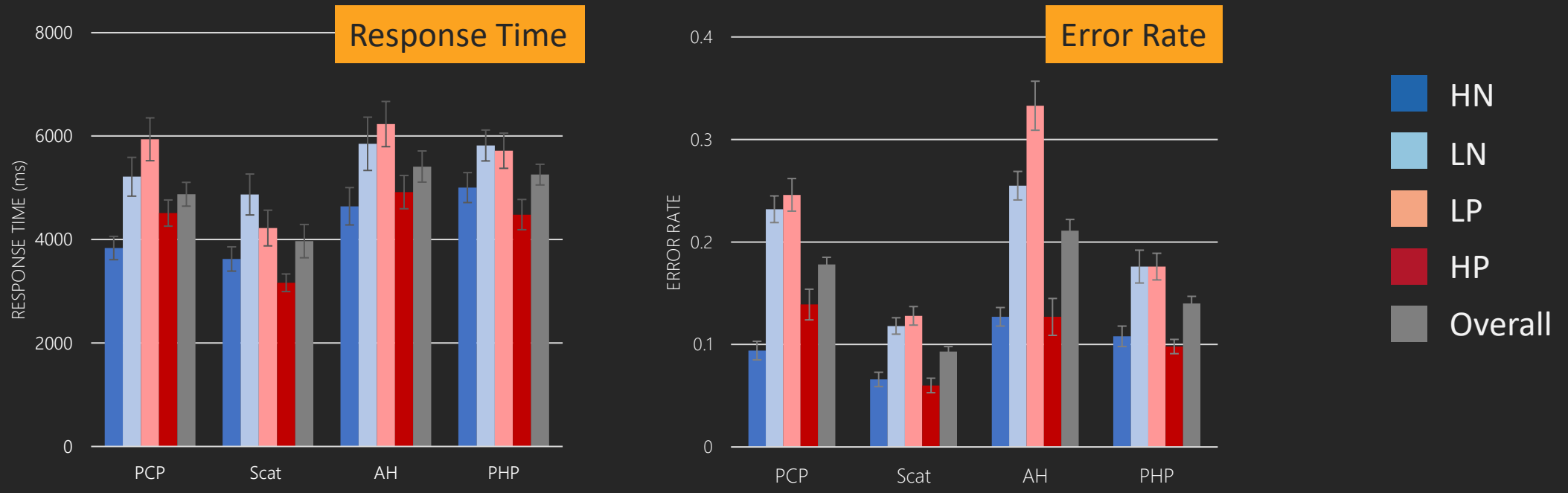
- Estimate correlation between two attributes, among four
  - 1000 items
  - Leftmost and another attribute
- 3 visualization techniques
  - PCP / AH / PHP*, order balanced by Latin Square
- 4 correlation conditions (Same as Task 1)
- 3 distance conditions
  - First, second, third
- Repeat each condition 3 times

*Pivot attribute is always leftmost



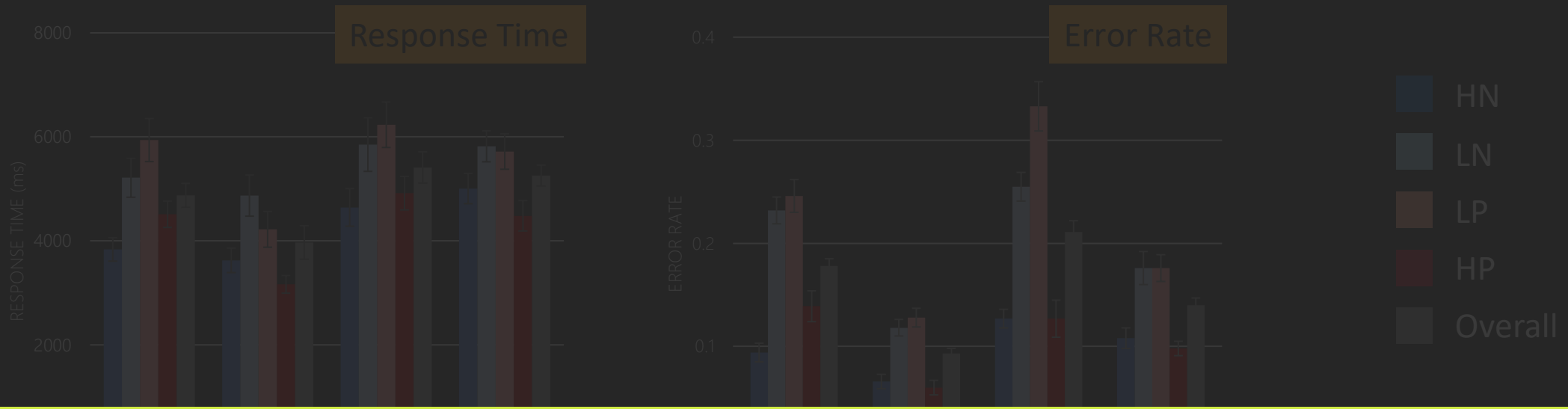
Result [Task 1]

- Time: Scatterplot > Others
- Accuracy: Scatterplot > PHP > PCP, AH



Result [Task 1]

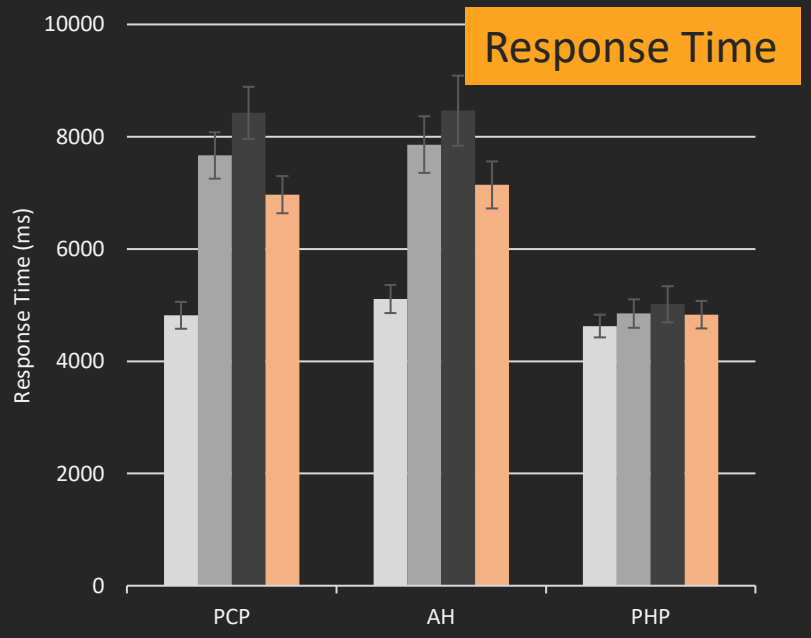
- Time: Scatterplot > Others
- Accuracy: Scatterplot > PHP > PCP, AH

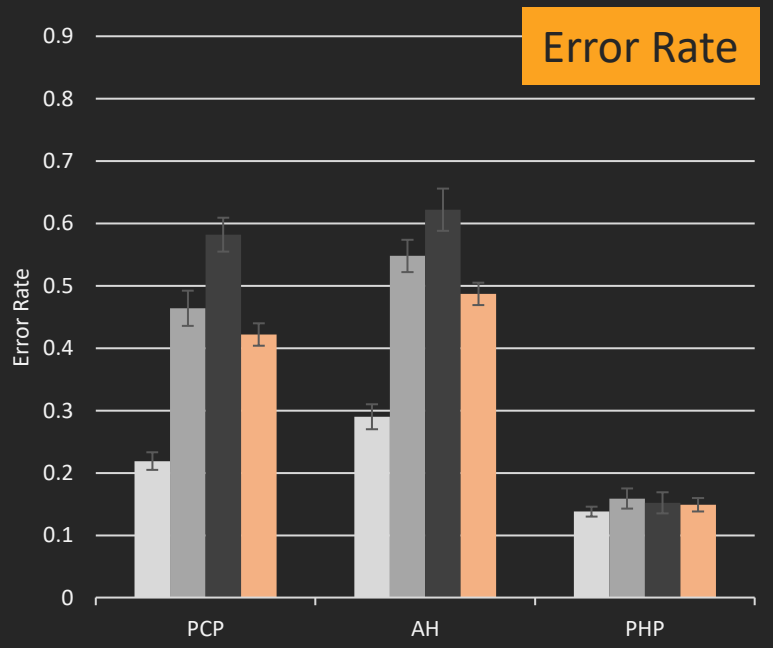


PHP can overcome the cluttering patterns of polylines!

Result [Task 2]

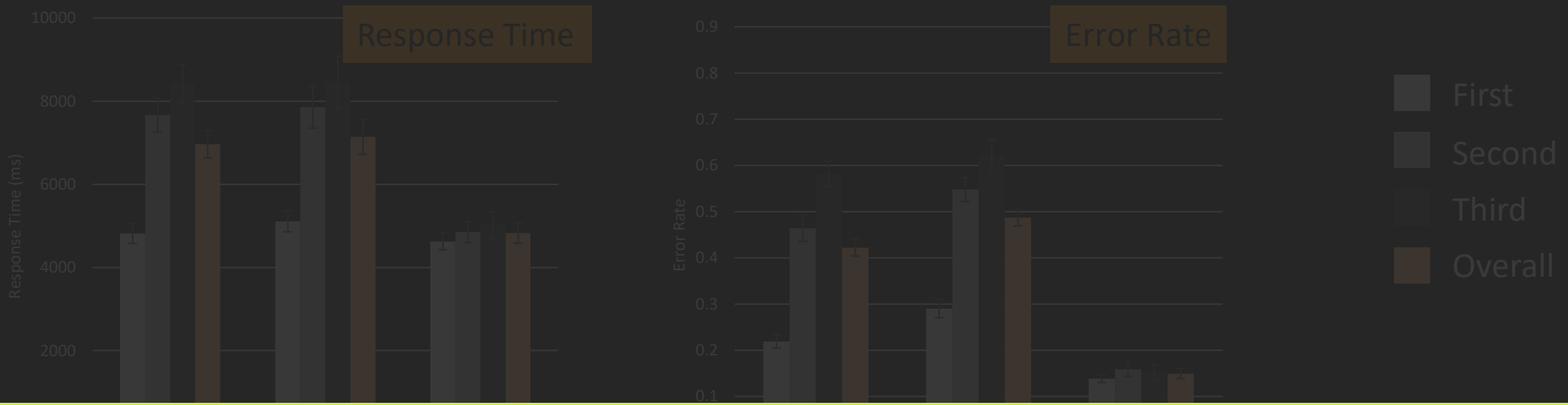
- Time: PHP > Others
- Accuracy: PHP > PCP > AH





Result [Task 2]

- Time: PHP > Others
- Accuracy: PHP > PCP > AH

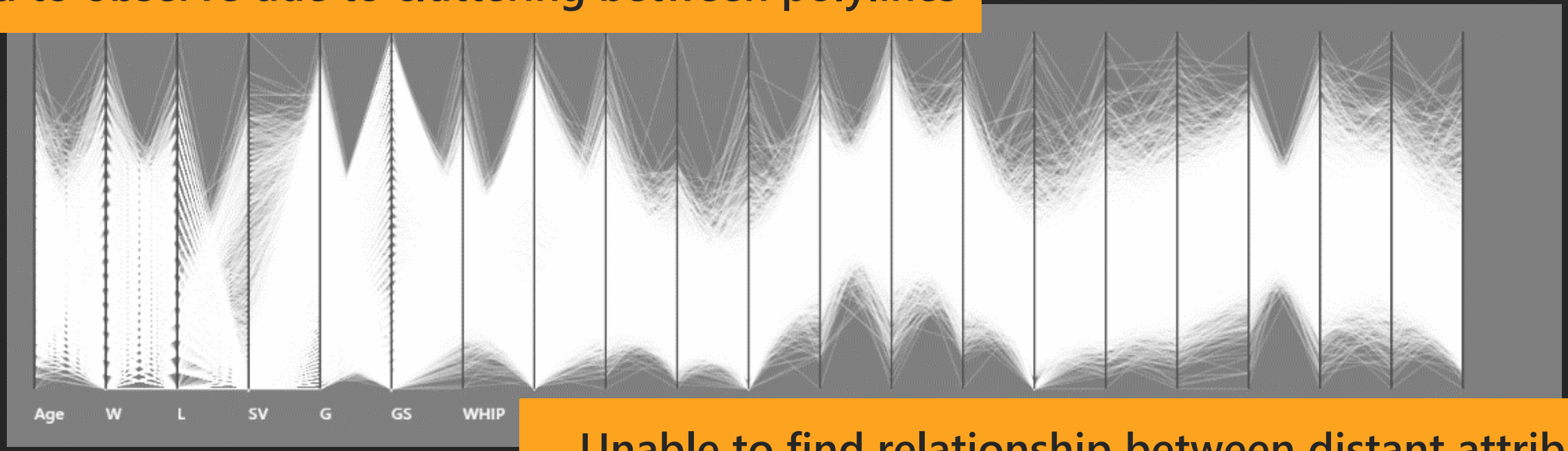


Visual patterns of PHP is effective even across distant attributes!

Discussion

Limitations of parallel coordinates plot

Hard to observe due to cluttering between polylines

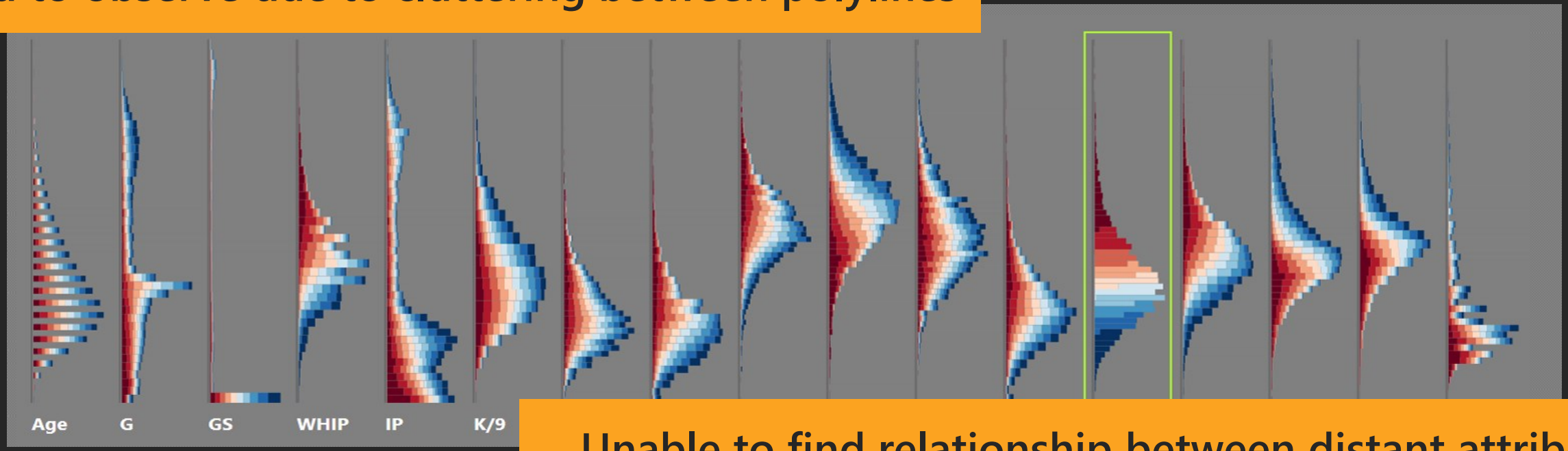


Unable to find relationship between distant attributes

Discussion

…can be overcome with Parallel Histogram Plots!

Hard to observe due to cluttering between polylines

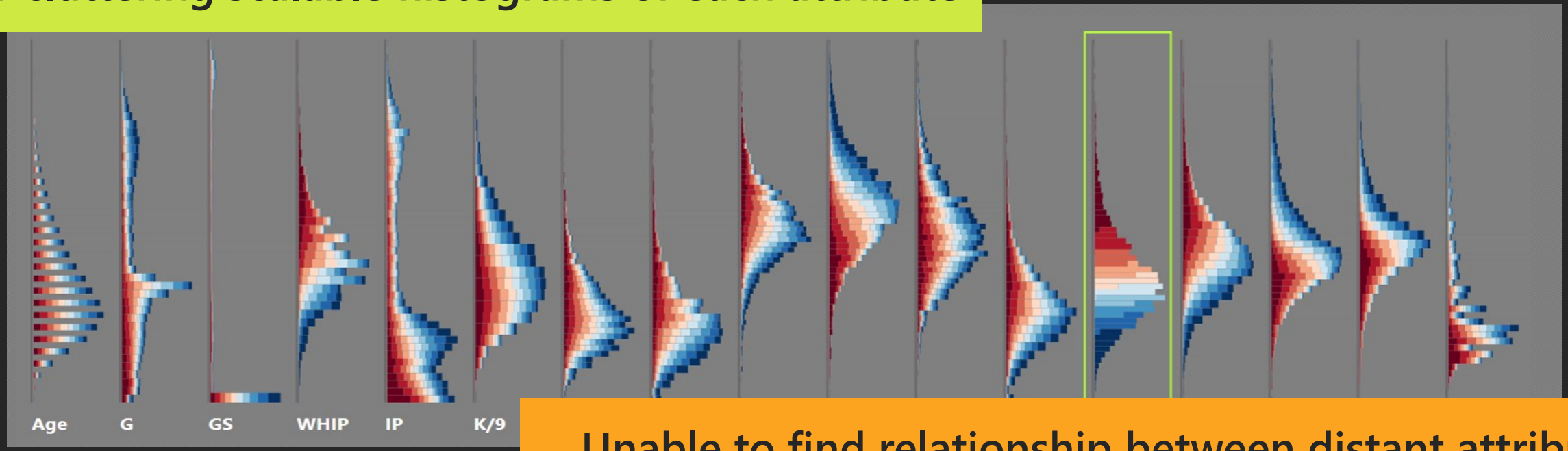


Unable to find relationship between distant attributes

Discussion

...can be overcome with Parallel Histogram Plots!

Non-cluttering scalable histograms of each attribute

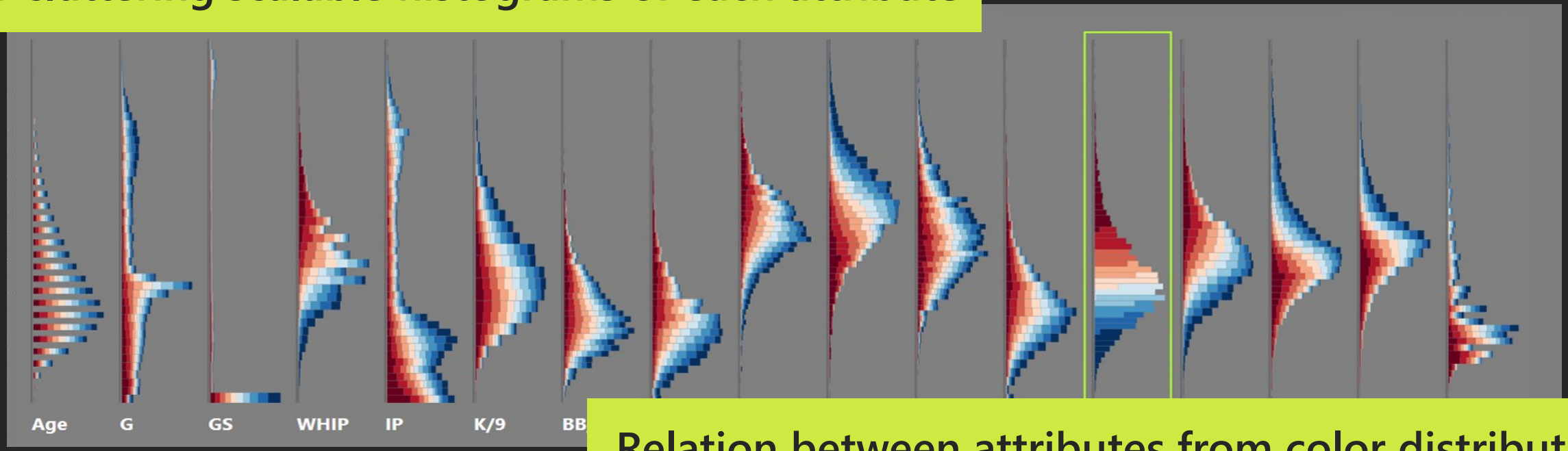


Unable to find relationship between distant attributes

Discussion

...can be overcome with Parallel Histogram Plots!

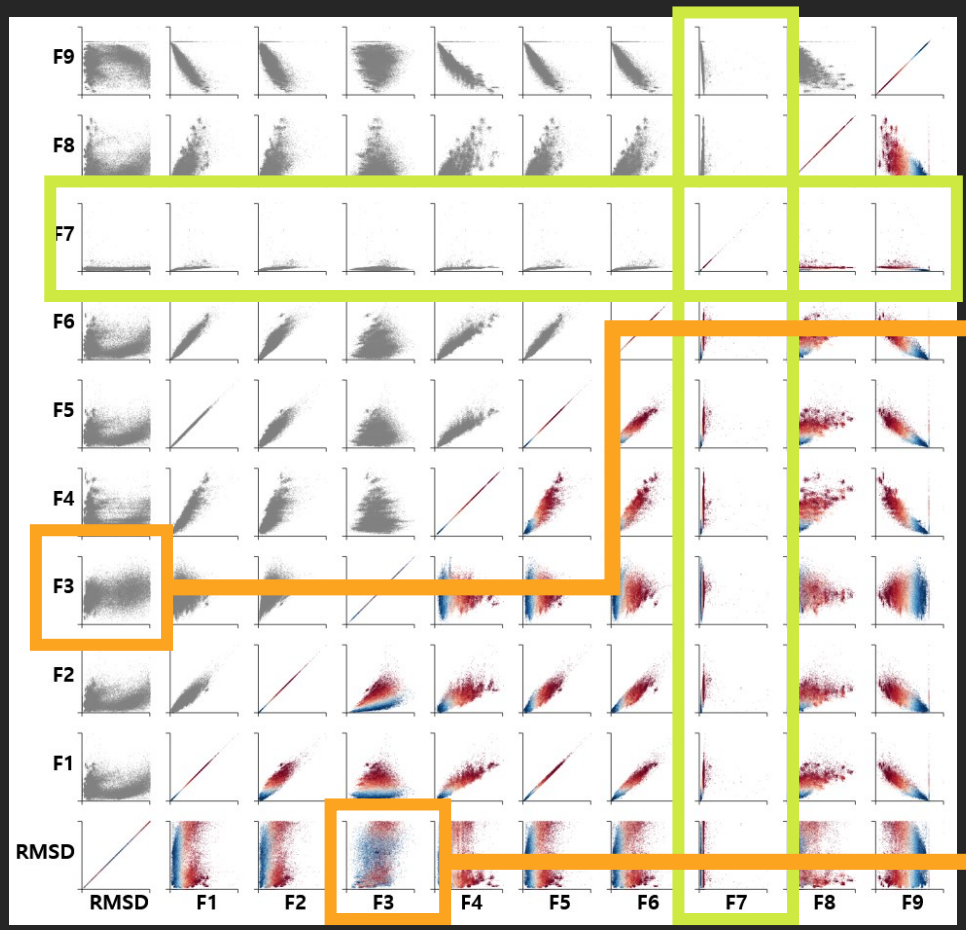
Non-cluttering scalable histograms of each attribute

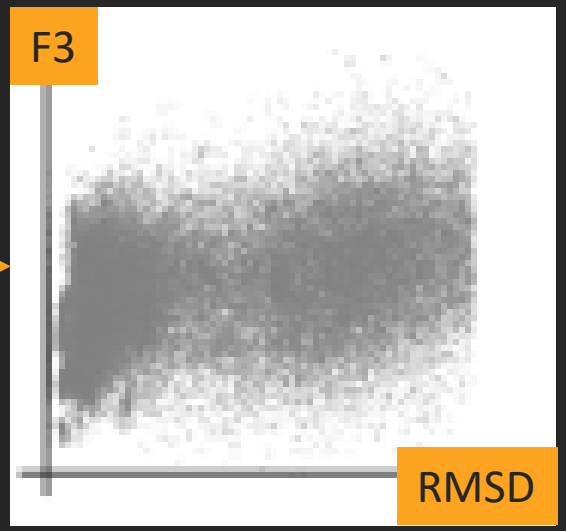


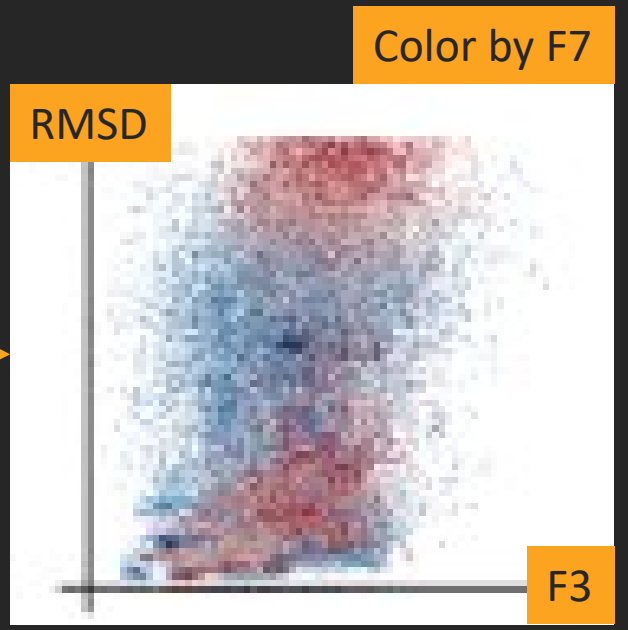
Relation between attributes from color distributions

Discussion

Extension to other visualizations







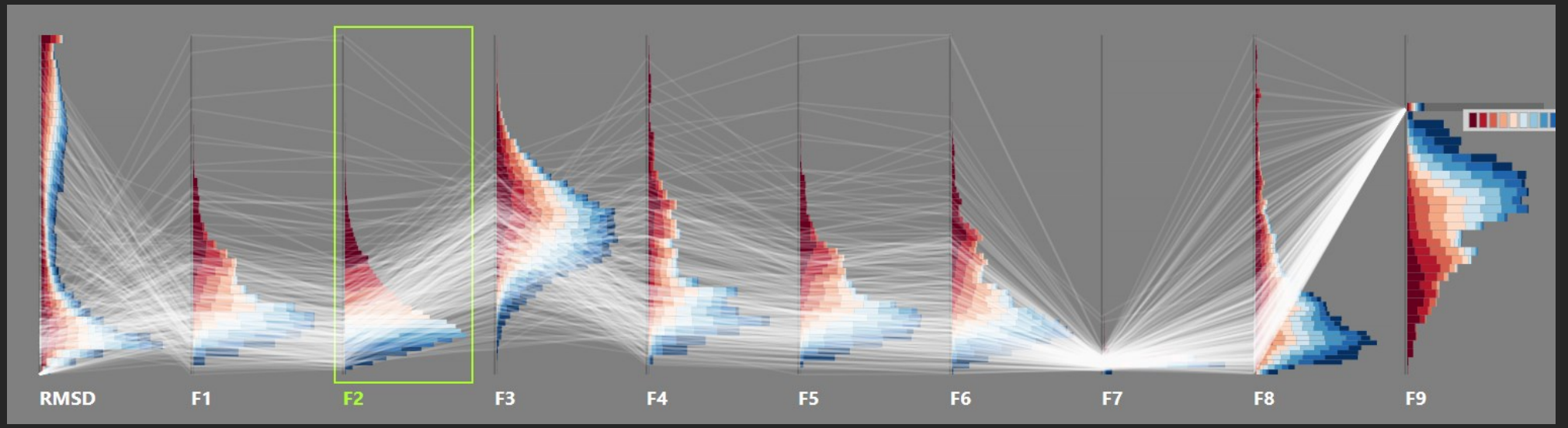
Augmenting Parallel Coordinates Plots with Color-coded Stacked Histograms

Jinwook Bok
bok@hcil.snu.ac.kr
Seoul National University

Bohyoung Kim
bkim@hufs.ac.kr
Hankook University of Foreign Studies

Jinwook Seo
jseo@snu.ac.kr
Seoul National University

https://github.com/bokjinwook/ParallelHistogramPlot



hci lab
SEOUL NATIONAL UNIVERSITY

VIS2021