Descriptions for dashboard_df.csv

Since the view pair (A, B) is directional, i.e., some features of (A, B) are different from that of (B, A). Thus, in a view pair (A, B), we call A as **Source** view, which means that it is a view to raise comparison or raise coordination; and we call B ad **Target** view, representing a view to receive comparison and coordination.

Specifically, the data table includes both the attributes of each view pair and the features extracted based on these attributes.

In the following, we will introduce the meanings of the features first, and then the meanings of other fields in the table.

Single-view features

Data and encoding features

feature name	Meaning	type
qkColNumS	the number of used fields are quantitative	Number
okColNumS	the number of used fields are ordinal	Number
nkColNumS	the number of used fields are nominal	Number
xColNumS	the number of the fields used on X-axis of source view	Number
yColNumS	the number of the fields used on Y-axis of source view	Number
colNumS	the number of the used fields, and regarding *sum(col1)* is same as *count(col1)*	Number
colNumS_agg	the number of the used fields, and regarding *sum(col1)* is different from *count(col1)*	Number
colwithAggPerS	the percentage of the fields using aggregation to all used fields	Number
...	those features replace 'S' with 'T' has same meaning	

Layout Arrangement features

feature name	Meaning	type
Sgrid*m_n*	boolean, whether the source view has area on (m,n) grid larger than the threshold (like more than 90% of that grid)	Boolean
Tgrid*m_n*	boolean, whether the target view has area on (m,n) grid larger than the threshold	Boolean
heightS_	the **normalized** height of the **source** view	Number
widthS_	the normalized width of the source view	Number
xS_	the normalized x position of the source view	Number
yS_	the normalized y position of the source view	Number
sizeS_	the normalized size of the source view	Number
...	those features replace 'S' with 'T' has same meaning	

Pairwise-view features

Data and encoding relationship features

feature name	Meaning	type
equalMark	whether two views are in the same mark	Boolean
withSameX	boolean, whether two views use the shared fields on X-axis	Boolean
withSameY	boolean, whether two views use the shared fields on Y-axis	Boolean
withSameSize	boolean, whether two views use the shared fields on size encoding	Boolean
withSameShape	boolean, whether two views use the shared fields on shape encoding	Boolean
withSameColor	boolean, whether two views use the shared fields on color encoding	Boolean
withSameEncoding	boolean, whether two views use the shared field on size, or shape, or color encoding	Boolean
equalX	whether two views used the exact same fields on X-axis	Boolean
equalY	whether two views used the exact same fields on Y-axis	Boolean
equalXCol	whether two views use the same number of fields on X-axis	Boolean

equalYCol	whether two views use the same number of fields on Y-axis	Boolean
SwithMoreXColNum	whether source view uses more fields on X-axis	Boolean
SwithMoreYColNum	whether source view uses more fields on Y-axis	Boolean
SwithMoreCols	whether source view uses more fields	Boolean
equalColNum	whether two views use the same number of fields	Boolean
isOverlap	whether two views use the shared fields when regarding *sum(col1)* is same as *count(col1)*	Boolean
overlapNum	how many shared fields	Number
dataOverlapS2T	the percentage of the number of the shared fields occupy the the number of the used fields of the target views, i.e., the number of the shared fields/ number of the used fields of the target views	Number
dataOverlapT2S	the percentage of the number of the shared fields occupy the the number of the used fields of the source views, i.e., the number of the shared fields/ number of the used fields of the source views	Number
dataR	define the relationship into 5 types, i.e., No, Equal, SasSubset (source view as subset), TasSubset (target view as subset), Intersection	String
SwithMoreAggCols	whether source views use more fields using aggregation	Boolean
equalAggCols	whether two views have the same number of fields using aggregation	Boolean
SwithMoreAggColPer	whether source views use more percentage of fields using aggregation	Boolean
equalAggColPer	whether two views use the same percentage of fields using aggregation	Boolean
...	those features adding the *_agg* suffix represent *sum(col1)* is different to *count(col1)*	

Layout arrangement relationship features

feature name	Meaning	type
shortest_view_dis	the minimum distance of two views	Number
isNeighbor	whether the minimum distance of two views are in the threshold	Boolean
angle	the angle of the center of two views	Number
direction	we divided the 360 degree to 8 directions and each with 45 degrees, e.g., [-22.5, 22.5) as right, [22.5,67.5] as top right	String
SwithLargerHeight	Boolean, whether the height of source view is larger than that of target view	Boolean
SwithLargerWidth	Boolean, whether the width of source view is larger than that of target view	Boolean
SwithLargerSize	Boolean, whether the size of source view is larger than that of target view	Boolean
equalHeight	whether two views have the same height	Boolean
equalWidth	whether two views have the same width	Boolean
equalSize	whether two views have the same size	Boolean

Coordination features

feature name	Meaning	type
coordination	the coordination type between two views, including *filter, brush, and no*	String

Others

Column name	Meaning
fileIndex	the index of the collected twbx files
imageIdx	the name of the dashboard image
dashboardId	the id of the dashboard in a file (a file may have several dashboards)
dashboardW	the width of the dashboard (not true width)
dashboardH	the height of the dashboard (not true height)
datasource_colNum	the num of the fields of the used data source
datasource_colName	the array of the field names of the used data source

datasource_calColName&Caption	the array of the new-defined filed names of the used data source, e.g., define a new attribute avg(age)
dashboard_viewNameList	the array of the view names to composite a dashboard
dashboard_viewNum	the num of the views in one dashboard
nameS	the view name of the **Source** view
dataCoverageS	the percentage of the number of the used fieds of the whole data source
encodingS	the json describe how a view is composited
colsS	the array to describe which filed names are used on the X-axis
rowsS	the array to describe which filed names are used on the Y-axis
colInPlotS	the array to describe which filed names are used and their corresponding aggregation and data type, including those used on X-axis, Y-axis, color encoding, shape encoding, and so on
colInPlotAggregationS	the array to describe the arregation of each used field, like 'sum' and 'count'
colInPlotNameS	the array of the filed name
colInPlotTypeS	the array of the data type
colNumS_agg	the number of the used fields, and regarding *sum(col1)* is different from *count(col1)*
colNumS	the number of the used fields, and regarding *sum(col1)* is same as *count(col1)*
heightS	the height of the **source** view
widthS	the width of the source view
xS	the x position of the source view
yS	the y position of the source view
markS	the mark type used in the source view
...	those attributes replace 'S' with 'T' has same meaning