Cheatography

Data Visualization in R for GR5293 Cheat Sheet

by Icey_cey via cheatography.com/145663/cs/31390/

Histogram(L4)

```
hist(x, col = "lightblue", ylim = c(a,b), xlim=c(a,b), xlab = "Lab for x axis", right = TRUE, main="Title for the histogram", breaks = seq(m,n,p))
```

x: the vector to visualize
col=: change the color of the histogram
xlim=/ylim=: define the range of x/y axis
xlab=/ylab=: rename the label for x/y axis
right=TRUE/FALSE: "TRUE" stands for the right-closed (left-opened) interval. "FALSE" stands for the right-opened (left-closed) interval
main=: name the title for the histogram
breaks=: set up the value of x axis

Single Boxplot(L5)

```
boxplot(x, horizontal=TRUE, log="x")
```

x: the vector to visualize
horizontal=TRUE/FALSE: make the boxplot horizontally or vertically
log=: if the x value is in the log scale

Multiple Boxplot(L5)

```
ggplot(dataset, aes(x= ,y=))
+geom_ box plot()
+labs()
+ theme( leg end.po sition = " - bot tom ")
```

dataset: the dataset to visualize
aes(x= ,y=): plot by x & y
labs(): label the element in the boxplot
theme(legend.position): assign the position of the legend

Violin Plot(L5)

```
ggplot(dataset, aes(x= ,y= ))
+geom_ vio lin()
+coord _flip()
+labs()
+theme()
```

dataset: the dataset to visualize
aes(x= ,y=): plot by x & y
geom_violin: get the violin plot
coord_flip(): flip the x and y coordinate
theme(): customize the non-data component

Ridgeline Plot(L5)

```
ggplot(dataset, aes(x= ,y= ))+ geom_d ens ity _ri dge s(f ill - ="bl ue", alpha= ,scale= )
```

dataset: the dataset to visualize
aes(x= ,y=): plot by x & y
geom_density_ridges(): get the Ridgeline plot
fill= : fill the Ridgeline with specific color
alpha= : set the transparency of the area under the Ridgeline

Q-Q plot (Quantile-Quantile)(L6)

```
qqnorm(x)
qqline(x, col="re d")
```

qqnorm(): produce a normal QQ plot of the values in x
qqline(): add a line to a “theoretical”, by default normal, quantile-quantile plot

Types of data(L8)

Numerical data	Categorical data
~Discrete	~Nominal - no fixed category order
~Conti-nuous	~Ordinal - fixed category order

Tidy Data(L10)

pivot_longer(data, cols = , names_to = ,values_to =): move selected columns' name to "name" column, and move values to a single "value" column

pivot_wider(data, names_from = , values-_from =): use the name from a column as the column name, and use the value from select column to be the value in the final Dataframe

row names_to_column(): add the column name to the rowname in the Dataframe

Parallel Coordinates(L13)

```
ggparcoord(dataset, columns = ,scale = ,alphaLines= ,splineFactor= ,groupColumn = )
```

dataset: the dataset to visualize
columns= : select columns of data that will include in the plot
scale= : method to scale the data (default is "std")
alphaLines= : value of alpha scaler for the lines of the parcoord plot or a column name of the data
splineFactor= : logical or numeric operator indicating whether spline interpolation should be used
groupColumn = : a single variable to group (color) by

Data Visualization in R for GR5293 Cheat Sheet

by Icey_cey via cheatography.com/145663/cs/31390/

Biplot (L14)

```
pca<- prcomp(dataset)
biplot (pca)
draw_b ipl ot( dat aset)
```

prcomp(): perform a principal components analysis on the given data matrix
draw_biplot():perform PCA on a data frame and draw a biplot

Cleveland dot plot(L15)

```
ggplot(dataset, aes(x = , y =
fct_reorder()))
+geom_ poi nt( color = )
+theme _li ned raw()
```

fct_reorder(): reorder factor levels by sorting along the variables
geom_point(): create scatterplots
theme_linedraw(): add black lines of various widths on white backgrounds

Multivariate Data(L15)

Stacked bar chart	Grouped bar chart	Mosaic plot (two variables)
ggplot(data, aes(x= , fill =))+geo- m_bar()+s- cale_fill- _manual()	ggplot(data, aes(x= ,fill=))+geom_bar(- position = "dod- ge")+scale_fill_- manual()	mosaic- (x~y, direction = c("v","h"), highligh- ting_fill=)

Multivariate Data(L15) (cont)

~plot x with different fill in different color	~bar plot grouped x filling with different color	~direction stands for the direction of different variables. highlighting_fill used for distinguish different group

Alluvial diagram(L16)

```
ggplot(dataset, aes(axis1 = ,
axis2 = , y = )) +
    geo m_a llu viu m(color = )
+
    geo m_s tra tum() +
    geo m_t ext (stat = " str -
atu m", aes(label = paste( aft -
er_ sta t(s tra tum), " \n",
after_ sta t(c oun t)))) +
    sca le_ x_d isc ret e(l -
imits = )
```

geom_alluvium(): plot both the lodes themselves, using geom_lode(), and the flows between them, using geom_flow()
geom_stratum(): plot rectangles for these strata of a provided width
geom_text(): add only text to the plot
scale_x_discrete(): set the values for discrete x scale aesthetics

Heatmap(L17)

```
ggplot(dataset, aes(x= , y= )) +
    geo m_t ile (ae s(fill = ),
color = ) +
    coo rd_ fixed()
```

geom_rect(): use the locations of the four corners (xmin, xmax, ymin and ymax)
geom_tile(): use the center of the tile and its size (x, y, width, height)
geom_raster(): a high performance special case for when all the tiles are the same size
coord_fixed(): a fixed scale coordinate system forces a specified ratio between data units on the axes

Time series(L20)

```
ggplot(dataset, aes(x= ,y=
,color= ))
+geom_ line()+
geom_s moo th( method= ,span= )
```

ggplot(dataset, aes(x= ,y= ,color=)): plot multiple time series by different colors
geom_smooth(): add a smooth line according to the data
method= : smoothing method (function) to use
span= : control the amount of smoothing for the default loess smoother

Factor in R

fct_recode(): change the name of the factor

fct_inorder(): display by each factor in the original order

fct_relevel(x, "G1", "G2", after = 3): move the factor "G1", "G2" after the third item in factor x

fct_reorder(color, count, .desc=TRUE): order by decreasing frequency count

fct_infreq(): display by number of observ- ations with each level (default is decreasing order of frequency)

fct_rev(): reverse the order of factor levels

fct_explicit_na(): turn NAs into a real factor level