Read in data

read_csv(): comma separated (CSV) files

read_tsv(): tab separated files

read delim(): general delimited files

read_fwf(): fixed width files

read_table(): tabular files where columns

are separated by white-space.

read_log(): web log files

Writing to file

write_csv()

write_tsv()

 $\mbox{write_excel_csv()} - \mbox{this writes a special}$

character (a byte order mark) at the

start of the file which tells Excel that you're using the UTF-8 encoding.

eg.write_csv(challenge, challenge.csv)

Plot graph by basic R

- 1. plot() scatter plot
- 2. pairs() scatter plot matrices
- 3. boxplot() box plots
- 5. lines() Line plot
- 6. pie() pie plot
- 7. plotix::pie3D() 3D pie chart
- 8. hist() histogram
- 9. density() dentistry plot
- 10. qqPlot() Quantile-Quantile plots, check normality
- 11. dotchart() dot plot

Parameters:

- 1. Add labels: xlab=, ylab=,
- 2. Add title: main=
- 3. Add subtitle: sub =
- 4. Add label of points: label=
- 5. Change colour: col=
- 6. Remover frame: frame=FALSE
- 7. Change the symbol of points : pch =
- 8. Change line width: lwd =
- 9. Change line type: Ity =
- 10. Change break for histogram: breaks =

ggplot2 (check in R for more information)

ggplot (data =) +

<Geom_function>(mapping = aes(<Mapping>), stat =

<STAT>, position = <Position>) +

<Coordinate_Function> +

- <Facet_Function> +
- <Scale Function>+
- <Labels Function>+
- <Theme_Function>

Geom function:

- 1. geom_abline(aes(intercept=0, slope=1))— line
- 2. geom_hline(aes(yintercept = lat)) horizontal line
- 3. geom_vline(aes(xintercept = long)) vertical line
- 4. geom_density() density curve
- 5. geom_dotplot() dot plot
- 6. geom_histogram() histogram plot
- 7. geom_boxplot() box plot
- 8. geom_violin() violin plot
- 9. geom bar() bar chart for discrete variables

Coordinate Function:

- 1. coord_fixed(ratio) fixed aspect ratio for x and y
- 2. coord flip() flip coordinates
- 3. coord_cartesian() default coordinate system

Facet Function:

- facet_grid(cols=variable) facet into columns based on a variable
- facet_grid(rows = variable) facet into rows based on a variable
- facet_wrap(variable) wrap facet into rectangle layout based on a variable

Scale Function:

- 1. scale * continuous() map cont' values to visual
- 2. scale_*_discrete() map discrete values to visual
- 3. scale_*_identity() use data values as visual ones

Theme_Function

- 1. theme_gray() grey theme
- 2. theme_bw() black and white theme
- 3. theme_light() light theme
- 4. theme_minimal() minimal theme
- 5. theme_void() empty theme

Remove ggplot components:

rrmove(Object), object list:

- both x and y grids: grid
- 2. x axis grids: x.grid
- 3. y axis grids: y.grid
- 4. both x and y axes: axis
- 5. x axis : x.axis
- 6. y.axis for y axis
- 7. x axis label: xlab
- 8. y axis label: ylab
- 9. both x and y axis labels: xylab
- 10. x axis texts: x.text
- 11. y axis texts: y.text

Dplyr:

- Mutate() adds new variables that are functions of existing variables
- 2. Select() picks variables based on their names.
- 3. Filter() picks cases based on their values.
- 4. Summarise() reduces multiple values down to a single summary.
- 5. Arrange() changes the ordering of the rows.

Change text style in graph

Font and colour:

ggpar(<graph>,

font.title = c(14, "bold.italic", "red"), #change title's style font.subtitle = c(10, "orange"), # change subtitle's style font.caption = c(10, "orange"), # change caption's style font.x = c(14, "blue"), # change x label's style font.y = c(14, "#993333") # change y label's style)

Position:

text direction: las = 1 (horizontal) justification: adj = 0 .5 1 (left, center, right)

Style

font face: font = 1 (plain) 2 (bold) 3 (italic) 4 (bold italic)