variable naming for coding

https://www.youtube.com/watch?v=-J3wNP6u5YU

**General:**

Rstudio cheat sheet

<https://rstudio.github.io/cheatsheets/html/rstudio-ide.html>

ggplot cheat sheet:

<https://github.com/rstudio/cheatsheets/blob/main/data-visualization.pdf>

aesthetic visual property: shape(0-24), color, size, alpha = variable name

use ?geom\_point to find out definition

facet\_wrap: wrap in single variable

facet\_grid: wrap in 2 variables

How to create a dataframe

<https://www.dataquest.io/blog/how-to-create-a-dataframe-in-r/#:~:text=We%20can%20create%20a%20DataFrame%20in%20R%20by%20combining%20two,we%20pass%20the%20necessary%20DataFrames>.

how to open multiple table using dbReadTable

<https://campus.datacamp.com/courses/intermediate-importing-data-in-r/importing-data-from-databases-part-1?ex=7>

introduction to R

<https://www.youtube.com/watch?v=_V8eKsto3Ug>

at 15:45 most commonly used package

Have you installed a LaTeX distribution in your system? For rmarkdown, tinytex is recommended, you would need to install the R package and then the TinyTex distribution.

-->

install.packages('tinytex')

tinytex::install\_tinytex()

pdf output: knitr::stitch('file\_name.r'), need to set work dirictory first



**ggplot** hint:

scatter plot: geom\_point()

<https://ggplot2.tidyverse.org/reference/geom_point.html>

variables levels: size, shape, color, alpha

with regression line: the `geom\_smooth()` function

histogram plot: geom\_hitogram()

<https://ggplot2.tidyverse.org/reference/geom_histogram.html>

density plot:geom\_density()

<https://ggplot2.tidyverse.org/reference/geom_density.html>

tidyverse package

data() can show the build in dataset

View(dataset name) will see the dataset

glimpse() will show class of the element

name(dataset name) is similar to head() or str()

Aming-lists-within-for-loop-using-paste0-and-get-eval-parse

<https://stackoverflow.com/questions/76298192/naming-lists-within-for-loop-using-paste0-and-get-eval-parse>

Tidy data with **Tidyr:**

A tidy data is: a tibble, each column has only one variable.

**R Data structure  
*Vector*:**

Vector: a collection of elements of same type. create a vector by using **c()**

Factor vector: vector with levels, using **as.factor** to create. The levels do not matter but assigned to each unique elements inside vector.

Ordered factor: using **ordered=TRUE** to turn factor vector into an order, then the level matters and elements can be ranked.

***Array*** is a special vector, all elements are same type. It has row, column and outer dimension. So it could be 3 dimensions.

Using square bracket to access, like this [,1] [,2] [,3]

***Data Frames***

data.frame is like an excel. It had row and columns, row represents an observation, column is a variable. So all columns are same length, each column is a vector with same type of elements. Different column could be different class and different type.

If each elements are same type, then it is also called ***Matrices***

Created by using **data.frame()** function.

Could assign name to column by using name=vectors.

Common function: **ncol(), nrow(), dim(), head(), tail(), class()**

Access by calling out row and column name with square bracket. like DF[row, column], number or name.

***List*:**

List is like a excel but can be mix of any elements, it can also has nested list inside.

Created by using **list()** function, assign name by using **names()**

It will access using square bracket like this: [[1]], using $ to access the elements, like list[[1]]$sport

Common function: **length(), names()**