Summary of common packages

https://statsandr.com/blog/correlation-coefficient-and-correlation-test-in-r/

1. Correlation coefficient related

|  |  |  |
| --- | --- | --- |
| **purpose** | **Package** | **function** |
| Calculate correlation coefficient of two variables | stats | cor |
| correlation coefficient test of two variables | stats | cor.test |
| Combine both together (two variables) | ggstatsplot | ggscatterstats |
| Calculate correlation matrix | Hmisc | rcorr |
| Combine correlation matrix and test multiple variables  (no visualization (figure)) | correlation | correlation |
| Combine correlation matrix and test of multiple variables  with visualization | GGally  ggstatsplot | Ggpairs  ggcorrmat |

对于两个变量是否相关，只看correlation test里的 pvalue, 只要pvalue小于0.05，就相关，且相关系数R往往也会很大。

Evaluate the relationship/association between two variables

|  |  |  |
| --- | --- | --- |
| **Type of variables** | **measurement** | **R function** |
| Two continuous variables | Pearson CC | Cor.test (x,y, method = "pearson") |
| One continuous and one qualitative ordinal | Spearman CC or  Kendall’s tau-b | Cor.test (x,y, method = "spearman")  Cor.test (x,y, method = "kendall") |
| One continuous and one two-level nominal | Pearson CC after or  point-biserial correlation  two categories must be encoded as 0 and 1 numbers | Cor.test (x,y, method = "pearson")  Biserial.cor(x,y,level=2) |
| Two qualitative ordinal variables | Spearman CC or  Kendall’s tau-b | Cor (x,y, method = "spearman")  Cor (x,y, method = "kendall") |
| One qualitative ordinal and one nominal | Rank-biserial CC | Did not find the R package to implement this function |
| Two nominal variables | Chi-square test (two levels) | chisq.test from “stats” package |
| Goodman and Kruskal's lambda (multiple levels) | lambda.test function from  “rapportools” package  **Actually, the Chi-square test can be used to two nominal variables with more than two levels.** |

The info in the table above is summarized from <https://statsandr.com/blog/correlation-coefficient-and-correlation-test-in-r/> and a paper titled “Measures of association how to choose”

1. Linear regression related

https://statsandr.com/blog/multiple-linear-regression-made-simple/

|  |  |  |
| --- | --- | --- |
| **purpose** | **Package** | **function** |
| Visualize regression result | visreg | visreg |
| Visualize regression result with equation and R^2 | ggplot2+ggpubr | ggplot+ stat\_cor stat\_regline\_equation+ |
| Check the correlation between variables for modeling comparison | Ggstatsplot  jtools | Ggcoefstats  plot\_summs |

1. Comparison with statistical result

|  |  |  |
| --- | --- | --- |
| **Box plots** | | |
| **purpose** | **Package** | **function** |
| Box & Violin plots with statistical result | ggstatsplot | ggbetweenstats |
| Box plot with jitter dots and statistical result | ggpubr | ggboxplot + stat\_compare\_means |
| **Bar plots** | | |
| **purpose** | **Package** | **function** |
| Bar plot with statistical result | ggstatsplot | ggbarstats |
| Bar plot | ggplot2 | Ggplot + geom\_bar |
| **Dot plots** | | |
| Dot plot with statistical result | ggstatsplot | ggdotplotstats |
| Dot plots only | ggplot2 | ggplot+ geom\_point |
|  |  |  |