

# CORRR PACKAGE



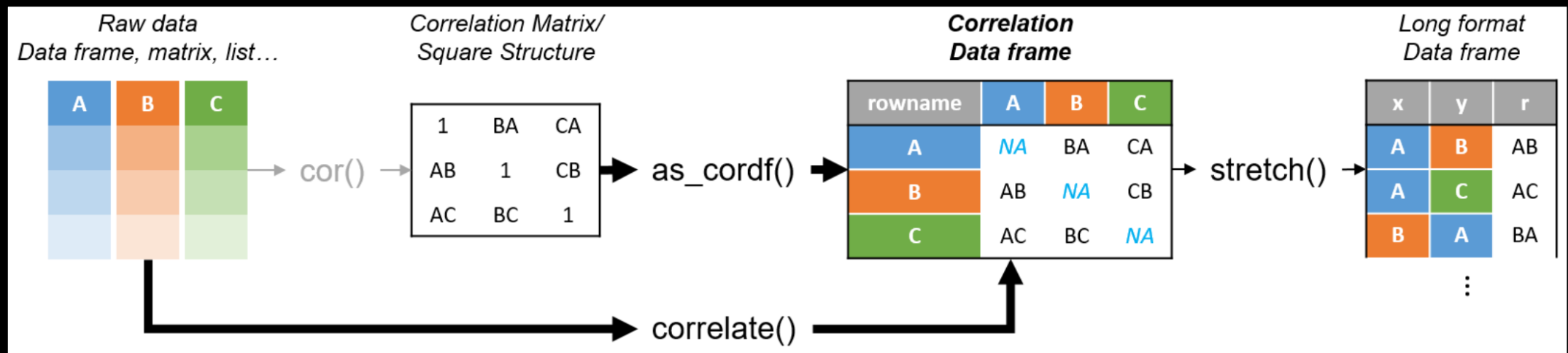
DITA SIŘÍNKOVÁ



ANDREA  
STAŠKOVÁ

# ZÁKLADNÍ CHARAKTERISTIKA

- práce s korelacemi (korelační matice ve formátu data.frame)
- úprava dat a vizualizace korelací





## ÚPRAVA DAT

- **shave()** = vybírá jednu z polovin korelační matice, kterou zobrazí
- **rearrange()** = zpřehází hodnoty korelační matice podle síly korelace (vysoké budou blízko)
  - **method**
  - **absolute** = rozlišení pozitivních a negativních korelací
- **focus()** = výběr zadaných sloupců a řádků
- **stretch()** = úprava dat na dlouhý formát

```
shave(x, upper = TRUE)
```

```
rearrange(x, method = "PC", absolute = TRUE)
```

```
focus(x, mpg, cyl) # Focus on correlations  
of mpg and cyl with all other variables
```

```
stretch(x, na.rm = FALSE) # omit all NAs, therefore keeping each  
# correlation only once.
```

# VIZUALIZACE

- **fashion()** = korelační matice v konzoli
- **rplot()** = korelační matice vizualizovaná v grafické podobě dle parametrů
- **network\_plot()** = síťový graf

```
fashion(x)
#>  rowname mpg drat  am gear qsec carb  hp  wt disp
#>  1      mpg
#>  2      drat  .68
#>  3      am   .60  .71
#>  4      gear  .48  .70  .79
#>  5      qsec  .42  .09 -.23 -.21
#>  6      carb -.55 -.09  .06  .27 -.66
#>  7      hp  -.78 -.45 -.24 -.13 -.71  .75
#>  8      wt  -.87 -.71 -.69 -.58 -.17  .43  .66
#>  9      disp -.85 -.71 -.59 -.56 -.43  .39  .79  .89
```

```
rplot(rdf, legend = TRUE, shape = 16, colours = c("indianred2",
"white", "skyblue1"), print_cor = FALSE, colors)
```

```
network_plot(rdf, min_cor = 0.3, legend = TRUE,
colours = c("indianred2", "white", "skyblue1"), repel = TRUE,
curved = TRUE, colors)
```

