

# Základní funkce vektorů

- `mode()` – mód resp. typ vektoru
- `length()` – délka vektoru (počet prvků vektoru)
  - `a <- 1:10`
  - `length(a)`
  - `[1] 10`
- `names()` – pojmenování prvků vektoru
  - `b <- c(27, 20, 35)`
  - `names(b) <- c("humerus", "radius", "femur"); b`

humerus	radius	femur
27	20	35
  - `names(b)`
  - `[1] "humerus" "radius" "femur"`
- `as.numeric()`, `as.character()`

# Výběr prvků vektoru

- `x2 <- letters; x2`
- `x2[5]` – výběr pátého prvku
- `x2[1:5]` – výběr prvních pěti prvků
- `x2[c(1,5)]` – výběr prvního a pátého
- `head(x2)`
- `head(x2,3)`
- `tail(x2)`
- `tail(x2,4)`
- `x3 <- 1:20; x3`
- `x3[x3<13 & x3>5]` – výběr prvků menších než 13 a větších než 5
- `b["humerus"]`

# Operace s vektory 1

- **Přidání nového prvku a přeskládání**

- `y <- 1:7`
- `y[6] <- 7` nahrazení/přepsání prvku
- `y <- c(y[1:2], 70, y[3:7])` přeskládání vektoru
- `y[9] <- 15` přidání nového prvku pozor: `y[13] <- 13`  

```
[1] 1 2 70 3 4 5 7 7 15 NA NA NA 13
```

- **Odstranění NA hodnot**

- `y <- y[!is.na(y)]`  

```
[1] 1 2 70 3 4 5 7 7 15 13
```

# Operace s vektory 2

- `summary(b)`
- `str(b)`
- **Spojování vektorů**
- `c(a, c)` spojení vektorů do jednoho delšího vektoru
- `cbind(a, b)` spojení vektorů (nebo matic) podle sloupců
- `rbind(a, b)` spojení vektorů (nebo matic) podle řádků
- `paste(a, b, c, sep=" ")` spojení vektorů do textového řetězce
- *Note:* s logickým vektorem je možné pracovat jako s vektorem numerickým, resp. `FALSE` je ekvivalentní 0, `TRUE` je ekvivalentní 1

- *Vyzkoušejte:*

- $a+b$ ,  $a*b$ ,  $a^b$ ,  $a/b$ ,  $a+c$ ,  $a*c$ ,  $a<c$ ,  $a-b$ ,  $a-c$

- **Note:**  $a+b = a[1]+b[1]$ ,  $a[2]+b[2]$ , etc.

- `d <- c(FALSE, TRUE, TRUE, FALSE, FALSE, TRUE)`

- $a*d$

- `a[7] <- 25`

- `cbind(a, c)`

```
> a[7] <- 25  
> cbind(a, c)
```

```
      a  c  
[1,]  1 55  
[2,] 43 13  
[3,] 86 55  
[4,] 12 13  
[5,] 57 55  
[6,] 98 13  
[7,] 25 55
```

```
Warning message:
```

```
In cbind(a, c) :
```

```
number of rows of result is not a multiple of vector length (arg 2)
```

# Vektor jako řetězec textu

- `nchar()` počet znaků textového řetězce
- `sub()` výměna/nahrazení znaků v textovém řetězci
- `paste()` spojení objektů daného vektoru do jednoho (textového) řetězce
- `strsplit()` rozdělí textový řetězec na jednotlivé složky
- `substring()` vrací část textového řetězce
  - `k <- "Coffee is your only hope"`
  - `l <- c("Coffee", "is", "your", "only", "hope")`
  - `nchar(k), nchar(l)`
  - `sub("Coffee", "Caffeine", k)`
  - `paste(l, collapse = ' ')`
  - `substring(l, 1, 1)`
  - `strsplit(k, split = ' ')`

## DOMÁCÍ ÚKOL 2

- A) Z vektoru `k` vytvořte za pomoci představených funkcí (předchozí strana) vektor `C`, který bude obsahovat textový řetězec: "Coffee is your only Caffeine".
- B) Zjistěte jak se budou lišit výstupy funkce `substring` v těchto případech:
  - `substring(k, 8, 8+11)`
  - `substring(1, 2, 2+1)`

# FAKTORY

- Speciální typ vektoru s možností přiřadit názvy jednotlivým kategoriím
- Obsahují informaci `levels` (množina hodnot, kterých může **kategorická** proměnná nabývat)
  - `f <- c("humerus", "ulna", "ulna", "radius", "femur", "ulna", NA)`
  - `as.factor(f)`
- Změna pořadí `levels`:
  - `factor(c("humerus", "ulna", "ulna", "radius", "femur", "ulna", NA), levels = c("humerus", "radius", "ulna", "femur"))`



- `f2 <- as.factor(1:5)`
- `mean(f2)`
- `Warning message: In mean.default(f2) :  
argument is not numeric or logical:  
returning NA`
- `mean(as.numeric(f2))`
  
- **Note:**
  - `as.factor()`
  - `as.numeric()`
  - `as.character()`
  - `as.matrix()`
  - `as.list()`
  - `as.data.frame()`
  
- *Vyzkoušejte si jednotlivé funkce na různých typech vektorů.*

# MATICE (`matrix`)

- datová struktura skládající se z řádků a sloupců
- `matrix(x, R, C)` – tvorba matice o R řádcích a C sloupcích, přičemž prvky jsou skládány po sloupcích
- `matrix(NA, nrow=2, ncol=2)` tvorba prázdné matice
  - `m <- 1:20`
  - `matrix(m, 4, 5)`
  - `matrix(m, ncol=5)`
  - `matrix(m, nrow=4); matrix(m, 4)`
  - `matrix(m, c(4, 5))`
  - `dim(m) <- c(4, 5); m`
  - `matrix(m, byrow=TRUE)` hodnoty skládáme po řádcích

# Výběr z matice - podmatice

- `m[1,2]` první řádek druhý sloupec
- `m[1:3,]` první tři řádky
- `m[,2]` druhý sloupec
- `m[c(1,3),]` první a třetí řádek
  
- **Funkce matic**
  - `nrow(m)` počet řádků matice *m*
  - `ncol(m)` počet sloupců matice *m*
  - `diag(m)` diagonála matice *m*
  - `rowSums(m)` sumarizační funkce řádků (vrátí součet hodnot v každém řádku)
  - `colSums(m)` sumarizační funkce sloupců (vrátí součet hodnot v každém sloupci)
  - `dimnames(m)` názvy sloupců a řádků matice
  - `rownames(m)` názvy řádků matice
  - `colnames(m)` názvy sloupců matice

# Spojování matic

- obdobně jako u vektorů:
  - `cbind(matice1, matice2)` spojení podle sloupců – je nutné mít stejný počet řádků obou matic
  - `rbind(matice1, matice2)` spojení podle řádků – je nutné mít stejný počet sloupců obou matic

- `A <- matrix(1:20, ncol = 5); A`

- `B <- matrix(letters[1:20], ncol = 5); B`

- `cbind(A, B) > cbind(A, B)`

- `rbind(A, B)`

```
[1,] [1,] [2,] [3,] [4,] [5,] [6,] [7,] [8,] [9,] [10,]
[2,] "1" "5" "9" "13" "17" "a" "e" "i" "m" "q"
[3,] "2" "6" "10" "14" "18" "b" "f" "j" "n" "r"
[4,] "3" "7" "11" "15" "19" "c" "g" "k" "o" "s"
[5,] "4" "8" "12" "16" "20" "d" "h" "l" "p" "t"
```

- `> rbind(A, B)`

- ```
[1,] [1,] [2,] [3,] [4,] [5,]
[2,] "1" "5" "9" "13" "17"
[3,] "2" "6" "10" "14" "18"
[4,] "3" "7" "11" "15" "19"
[5,] "4" "8" "12" "16" "20"
[6,] "a" "e" "i" "m" "q"
[7,] "b" "f" "j" "n" "r"
[8,] "c" "g" "k" "o" "s"
[9,] "d" "h" "l" "p" "t"
```

- `> mode(rbind(A, B))`

- ```
[1] "character"
```

- `> typeof(rbind(A, B))`

- ```
[1] "character"
```

# POLE (Array)

- v podstatě se jedná o vícerozměrné matice
  - *Note: dvourozměrné pole lze tedy chápat jako matici*
- Příklad dvourozměrné pole, kde 1. dimenze = 4 a 2. dimenze = 3
  - `array (1:12, c(4, 3))`
- Příklad dvourozměrné pole (o 4 řádcích a 2 sloupcích) s názvy jednotlivých dimenzí (vzorek a sex)
  - `array (c(1,5,3,7,8,9), c(4, 2), dimnames=list(vzorek=1:4, sex=c("m", "f")))`

```
> array (c(1,5,3,7,8,9),c(4, 2),dimnames=list(vzorek=1:4,sex=c("m","f")))
```

```
      sex
vzorek m f
  1 1 8
  2 5 9
  3 3 1
  4 7 5
```

# Výběr z pole

- Máme pole `x1` a chceme vybrat například:
  - Třetí řádek druhé matice
    - `x1 [3, , 2]`
  - Celou první matici
    - `x1 [ , , 1]`
  - Prvek z prvního řádku třetího sloupce druhé matice
    - `x1 [1, 3, 2]`
  - Prvky z prvního a druhého řádku třetího sloupce první matice
    - `x1 [1:2, 3, 1]`
  - První a třetí řádek
    - `x1 [c(1, 3) , , ]`

- Mějme 3D pole ( $p \times k \times n$ ) souřadnic landmarků
  - pro následnou analýzu např. PCA, MANOVA, etc. potřebuje pole převést na 2D pole
  - Balíček geomorph
  - Funkce `two.d.array`

3D pole ( $p \times k \times n$ )



2D pole ( $n \times [p \times k]$ )

- $p$  : *měření*
- $k$  : *rozměry*
- $n$  : *sex*

```
, , sex = male
rozmary
mereni H1 H2 F1 F2
prvni 35 34 45 44
druhe 34 34 46 45
tretí 30 29 40 39
```

```
, , sex = female
rozmary
mereni H1 H2 F1 F2
prvni 33 32 47 45
druhe 35 35 46 46
tretí 34 33 45 45
```

```
prvni.H1 prvni.H2 prvni.F1 prvni.F2 druhe.H1 druhe.H2 druhe.F1 druhe.F2 tretí.H1 tretí.H2 tretí.F1 tretí.F2
male 35 34 45 44 34 34 46 45 30 29 40 39
female 33 32 47 45 35 35 46 46 34 33 45 45
```

## DOMÁCÍ ÚKOL 3

- A) Vytvořte matici `MAT` s rozměry  $7 \times 3$ , tak:
  - Aby první řádek obsahoval posloupnost:  $-2, 2, 6, 10, \dots$
  - Druhý řádek tvořily hodnoty `13` a
  - Třetí řádek měl minimální hodnotu  $-3$  a maximální `3`
- B) vypište sedmý sloupec matice `MAT` jako:
  - i) řádkový vektor `a`
  - ii) jako sloupcový vektor
- C) využitím funkce `cbind`, vytvořte matici `B` tak, aby výsledek vypadal následujícím způsobem:



# SEZNAMY (list)

- specifické objekty sdružující různé jiné objekty různých typů;
- jednotlivé položky seznamu mohou mít vlastní název (není povinný) - v rámci jednotlivých položek jsou uloženy dané hodnoty: **v každé položce pouze jeden typ**
  - ```
Doctor <- list(name = c("William Hartnell", "Patrick Troughton", "Jon Pertwee", "Tom Baker", "Peter Davison", "Colin Baker", "Sylvester McCoy", "Paul McGann", "John Hurt", "Christopher Eccleston"), year=c("55", "45", "50", "40", "29", "40", "44", "36", "73", "41"), episodes=c("134", "119", "128", "172", "69", "31", "42", "1", "2", "13")); Doctor
```

# Přidání položek do seznamu

- `Doctor <- list(name = c(Doctor$name, "David Tennat", "Matt Smith", "Peter Capaldi"), year = c(Doctor$year, "34", "27", "55"), episodes = c(Doctor$episodes, "47", "44", "26")); Doctor`

```
> Doctor <- list(name = c(Doctor$name, "David Tennat", "Matt Smith", "Peter Capaldi"), year = c(Doctor$year, "34", "27", "55"),
+               episodes = c(Doctor$episodes, "47", "44", "26")); Doctor
$name
 [1] "william Hartnell"      "Patrick Troughton"    "Jon Pertwee"          "Tom Baker"           "Peter Davison"
 [7] "sylvester McCoy"      "Paul McGann"          "John Hurt"            "Christopher Eccleston" "David Tennat"
[13] "Peter Capaldi"

$year
 [1] "55" "45" "50" "40" "29" "40" "44" "36" "73" "41" "34" "27" "55"

$episodes
 [1] "134" "119" "128" "172" "69" "31" "42" "1" "2" "13" "47" "44" "26"
```