Ćwiczenie nr 1 – wstęp do systemu R

- 1. Pobierz i zainstaluj pakiet R http://www.r-project.org
- 2. Pomoc
 - Jeśli znamy nazwę funkcji, np.
 ?solve lub help(solve)
 - Jeśli nie znamy nazwy funkcji: help.start() lub help.search("szukany tekst")
- 3. Przypisywanie wartości zmiennym Zmiennym można przypisywać wartości poprzez operatory przypisania ->, <-, = lub poprzez funkcje assign(). Na przykład: a<-5; 4->b; c=b; assign("x",10) a; b; c; x
- 4. Podstawowe struktury danych
 - \bullet vector

Do tworzenia wektora zawierającego konkretne elementy używa się funkcji c(). Na przykład:

```
a<-c(10,-15,4,-2,8,7)
```

```
а
```

Operator : służy do generowania ciągu kolejnych liczb całkowitych, a funkcja seq() do generowania liczb z zakresu z określonym krokiem. Przykłady: -10:5

```
seq(-5, 5, by=.2)
```

Ciąg o powtarzających się elementach tworzy się za pomocą funkcji **rep()**, np.: **rep(c(1,2),4)**

Na wektorach można wykonywać działania lub podstawowe funkcje, np.: x <-c(2,4,5); y <-c(1,3,2)

```
x+y; x*y+2
log(x)
Indeksowanie wektorów:
a[2] # drugi element
a[-2] # wszystkie oprócz drugiego
a[c(1,5)] # pierwszy i piąty
a>0 # wektor logiczny - które większe od zera
a[a>0] # elementy większe od zera
Często łatwiej zapamiętać nazwy niż wartości liczbowe. Tekstowe nazwy przypad-
ków można nadać w np. w następujący sposób:
owoce<-c(1,5,8,10)
names(owoce)<-c("jablko","gruszka","banan","sliwka")
obiad<-owoce[c("jablko","banan")]</pre>
```

 $\bullet \ factor$

Faktor jest strukturą przechowującą informacje o powtórzeniach takich samych wartości, np,:

```
vect<-c(1,2,1,3,1,2,4,3,2); vect
fact<-factor(vect); fact
summary(fact)
plot(fact)
```

• array

Tablicę można utworzyć z wektora poprzez określenie wymiarów funkcją dim(), np.:

```
tbl=1:20
dim(tbl)=c(4,5) # wektor staje sie tablica o wymiarach 4,5
tbl
tbl[2,3] # element na pozycji 2,3
tbl[2,] # drugi wiersz
tbl[,3] # trzecia kolumna
Można również dodawać kolumny:
tbl<-cbind(tbl,3); tbl
lub wiersze:
tbl<-rbind(log(tbl[1,]),tbl); tbl</pre>
```

• data frame

Można utworzyć na kilka sposobów:

- wczytując z pliku tekstowego read.table("nazwa pliku"),
- zamieniając macierz poprzez funkcję as.data.frame(),
- poprzez funkcję data.frame(), np.: dat<-data.frame("x"=c(2,4,1,2), "y"=c(1,1,1,1))</pre>
- interaktywnie zmieniając istniejącą data.frame za pomocą funkcji fix() lub edit().

Do poszczególnych zmiennych odwołujemy się poprzez nazwa.obiektu\$nazwa.zmiennej. Aby do zmiennych można było odwołać się bezpośrednio można dany obiekt przypisać przez funkcję attach().

- 5. Przykłady i ćwiczenia
 - Zainstaluj bibliotekę faraway, otwórz ją i otwórz dane pima z tej biblioteki: Z górnego paska wybierz Packages i Instal packages, następniej wybierz serwer i bibliotekę. Następnie wpisz:

library(faraway); data(pima)

Wszystkie dostępne bazy danych uzyskujemy przez komendę data(), a opis danych poprzez ?pima. Dane te dotyczą badań nad cukrzycą u indianek Pima. Zawierają one ilość ciąży (pregnant), koncentracja glukozy (glucose), ciśnienie rozkurczowe (diastolic), grubość fałdy skórnej (triceps), stężenie insuliny (insulin), indeks masy do wzrostu (bmi), funkcja cukrzycy (diabetes), wiek (ages), czy pacjent wykazuje symptomy cukrzycy (test). Przeglądnij dane i wyświetl podstawowe statystyki dotyczące tych danych:

summary(pima)

Czy dostrzegasz coś niepokojącego?

```
W R zmienne brakujące oznacza się przez NA. Nanieś niezbędne poprawki:

pima$diastolic[pima$diastolic == 0] <- NA

pima$glucose[pima$glucose == 0] <- NA

pima$triceps[pima$triceps == 0] <- NA

pima$insulin[pima$insulin == 0] <- NA

pima$bmi[pima$bmi == 0] <- NA

Czy zmienna test jest zmienną ilościową czy jakościową? Wskazówka:

pima$test <- factor(pima$test)

levels(pima$test) <- c("negative","positive")

Przeglądnij ponownie podsumowanie danych.
```

• Stwórz obiekt westwood typu data frame i wypełnij go danymi:

LotSize	ManHours
30	73
20	50
60	128
80	170
40	87
50	108
60	135
30	69
70	148
60	132

Zapisz do pliku "westwood.txt":

```
westwood<-data.frame()
fix(westwood)
Ustaw katalog roboczy z górnego paska.
write.table(westwood,"westwood.txt")</pre>
```

- Otwórz dane zawierające średnie roczne temperatury z pliku "temperature.txt". Przeglądnij je i zastanów się czy wszystko jest w porządku. Spróbuj wczytać je poprawnie.
- Zapisz wyniki dzisiejszej pracy. Z górnego paska wybierz *File* oraz *Save Workspace, Save History* i *Save to File.*
- 6. Więcej na temat systemu R
 możesz znaleźć w "An Introduction to R" dostępnym na stronie projektu R.