

Lista 1

Mikołaj Langner, Marcin Kostrzewa

31.3.2021

1 Etap I

Wczytajmy dane z pliku i przeprowadźmy ich wstępną analizę i obróbkę:

```
df <- read.csv('churn.txt', stringsAsFactors = TRUE)
```

- poznajmy rozmiar naszych danych:

```
dim(df)

## [1] 3333 21
```

- sprawdźmy ich typ:

	Typ zmiennej
State	factor
Account.Length	integer
Area.Code	integer
Phone	factor
Int.l.Plan	factor
VMail.Plan	factor
VMail.Message	integer
Day.Mins	numeric
Day.Calls	integer
Day.Charge	numeric
Eve.Mins	numeric
Eve.Calls	integer
Eve.Charge	numeric
Night.Mins	numeric
Night.Calls	integer
Night.Charge	numeric
Intl.Mins	numeric
Intl.Calls	integer
Intl.Charge	numeric
CustServ.Calls	integer
Churn.	factor

Tabela 1: Hello

- sprawdźmy czy pojawiają się wartości brakujące:

```
sum(sapply(df, function(x) sum(is.na(x))))  
## [1] 0
```

- usuńmy dane pełniące rolę identyfikatora:

```
df <- subset(df, select=-Phone)
```

2 Etap II

3 Etap III

4 Etap IV