Rozwiązania zadań z laboratoriów 3.

ZADANIE 1

Stan po wczytaniu pliku:

```
> str(df)
Classes ‘spec_tbl_df’, ‘tbl_df’, ‘tbl’ and 'data.frame':        150 obs. of  5 variables:
 $ sepal.length: chr  "5.1" "4.9" "4.7" "4.6" ...
 $ sepal.width : chr  "3.5" "3" "3.2" "3.1" ...
 $ petal.length: num  1.4 1.4 1.3 1.5 1.4 1.7 1.4 1.5 1.4 1.5 ...
 $ petal.width : num  0.2 0.2 0.2 0.2 0.2 0.4 0.3 0.2 0.2 0.1 ...
 $ variety     : chr  "Setosa" "Setosa" "Setosa" "Setosa" ...
 - attr(*, "spec")=
  .. cols(
  ..   sepal.length = col_character(),
  ..   sepal.width = col_character(),
  ..   petal.length = col_double(),
  ..   petal.width = col_double(),
  ..   variety = col_character()
  .. )
```

Po konwersji do typu liczbowego. Rekordy z pustymi danymi:

```
> glimpse(df %>%
+   filter(is.na(df$petal.width) | is.na(df$petal.length) | is.na(df$sepal.length) | is.na(df$sepal.widt
h)))
Observations: 4
Variables: 5
$ sepal.length <dbl> 6.3, NA, 6.3, NA
$ sepal.width  <dbl> NA, 3.0, 2.8, 2.7
$ petal.length <dbl> 5.6, 5.8, 5.1, 5.1
$ petal.width  <dbl> 1.8, 2.2, NA, 1.9
$ variety      <chr> "Virginica", "Virginica", "Virginica", "Virginica"
```

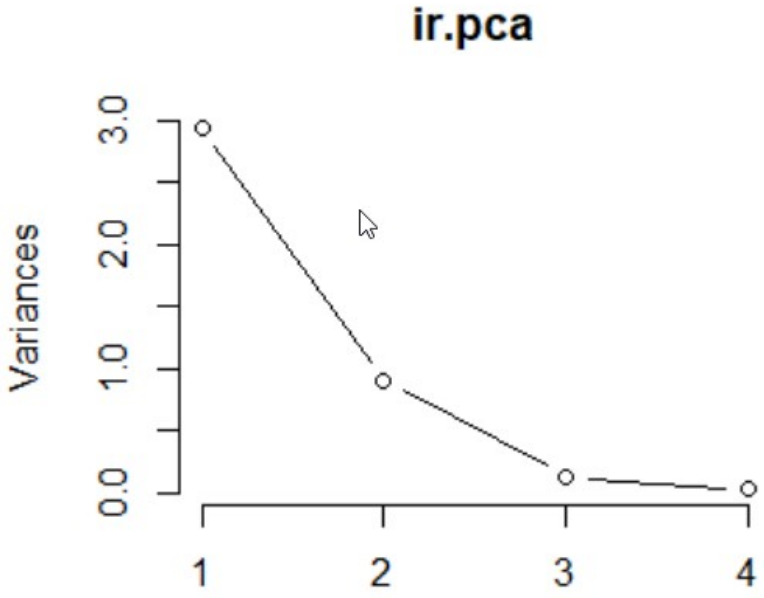
Puste rekordy zastąpiono medianą. Ostatnie poprawki polegały na poprawieniu błędów w nazwie gatunków:

```
> #show which variety has wrong format
> glimpse(df %>%
+   filter(variety != 'Virginica' & variety != 'Setosa' & variety != 'Versicolor'))
Observations: 4
Variables: 5
$ sepal.length <dbl> 4.6, 4.9, 6.0, 6.7
$ sepal.width  <dbl> 3.0, 2.4, 3.0, 3.3
$ petal.length <dbl> 1.4, 3.3, 4.8, 5.7
$ petal.width  <dbl> 0.2, 1.0, 1.8, 2.5
$ variety      <chr> "setosa", "Versicolour", "virginica", "virginica"
```

ZADANIE 2:

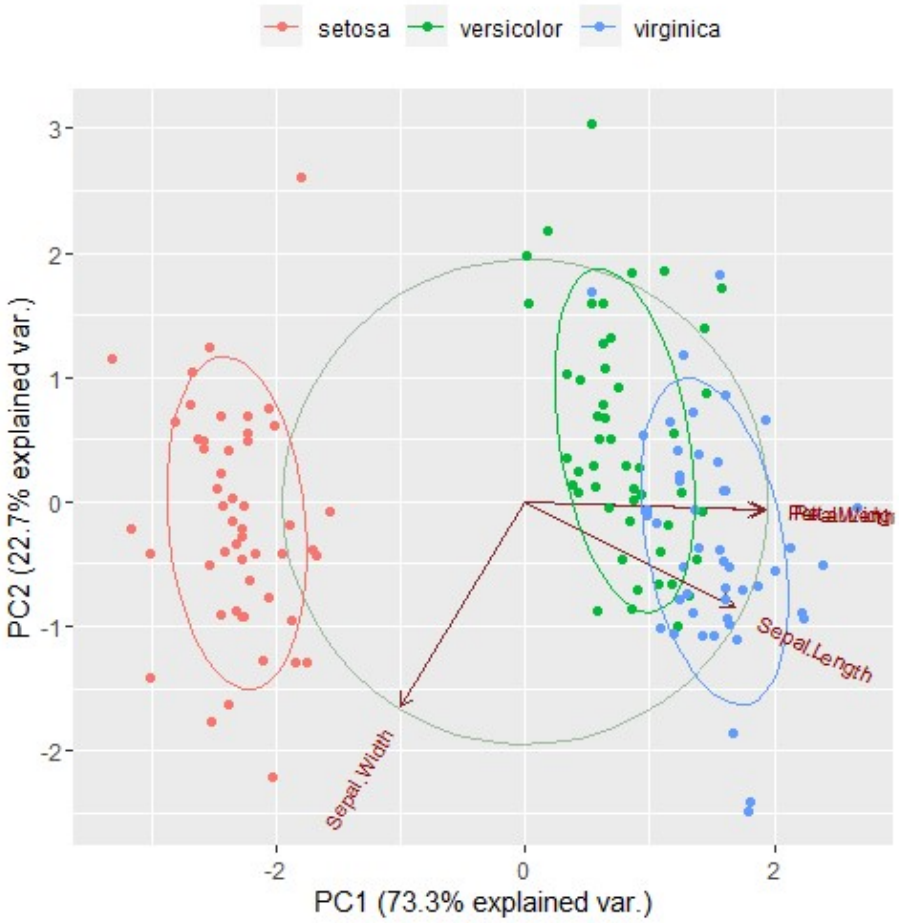
```
Standard deviations (1, .., p=4):
[1] 1.7124583 0.9523797 0.3647029 0.1656840

Rotation (n x k) = (4 x 4):
                    PC1         PC2        PC3         PC4
Sepal.Length  0.5038236 -0.45499872  0.7088547  0.19147575
Sepal.Width  -0.3023682 -0.88914419 -0.3311628 -0.09125405
Petal.Length  0.5767881 -0.03378802 -0.2192793 -0.78618732
Petal.Width   0.5674952 -0.03545628 -0.5829003  0.58044745
```




```
Importance of components:
                          PC1    PC2     PC3     PC4
Standard deviation     1.7125 0.9524 0.36470 0.16568
Proportion of Variance 0.7331 0.2268 0.03325 0.00686
Cumulative Proportion  0.7331 0.9599 0.99314 1.00000
```

```
> predict(ir.pca,
+         newdata=tail(log.ir, 2))
          PC1         PC2        PC3         PC4
149 1.0809930 -1.01155751 -0.7082289 -0.06811063
150 0.9712116 -0.06158655 -0.5008674 -0.12411524
```



```
> head(PC, 3)
        PC1         PC2
1 -2.303540 -0.4748260
2 -2.151310  0.6482903
3 -2.461341  0.3463921
> trans$rotation
                    PC1         PC2
Sepal.Length  0.5202351 -0.38632246
Sepal.Width  -0.2720448 -0.92031253
Petal.Length  0.5775402 -0.04885509
Petal.Width   0.5672693 -0.03732262
```