

1 Zadanie 8.2

```
# 8.2

m <- read.table("http://www.ipipan.eu/~teisseyrep/TEACHING/SAR/DANE/Miasta.txt",
  header = TRUE)
head(m)

##           Work Price Salary
## Amsterdam   1714   65.6   49.0
## Athens      1792   53.8   30.4
## Bogota       2152   37.9   11.5
## Bombay       2052   30.3    5.3
## Brussels    1708   73.8   50.5
## BuenosAires 1971   56.1   12.5

# a) standaryzacja zmiennych

mm <- scale(m)

# inny sposob (taki po kolei)

sr <- numeric(3)
sd <- numeric(3)

for (i in 1:3) {
  sr[i] <- mean(m[, i])
  sd[i] <- sd(m[, i])
}

for (i in 1:3) {
  m[, i] <- (m[, i] - sr[i])/sd[i]
}

# b) dla Work i Price wyznaczyc kierunki wzdluz ktorych wystepuje najwieksza
# zmiennosc; wykres rozproszenia

pc1 <- princomp(~., cor = FALSE, data = as.data.frame(mm[, 1:2]))

# wybieramy albo macierz korelacji albo kowariancji - mamy zestandaryzowane
# dane, wiec juz korelacja nie jest potrzebna

names(pc1)

## [1] "sdev"      "loadings" "center"   "scale"    "n.obs"    "scores"
## [7] "call"

head(pc1$scores) # Comp.1 to z1, a Comp.2 to z2

##           Comp.1  Comp.2
## Amsterdam   0.5242  0.82168
## Athens      -0.1823  0.89543
## Bogota       -2.1680 -0.03904
## Bombay       -2.0137  0.61779
## Brussels     0.8196  0.57493
## BuenosAires -0.8323  0.09339

pc1$loadings # wektory ladunkow (pierwsza kolumna to a1, a druga to a2) ->

##
## Loadings:
##           Comp.1 Comp.2
```

```
plot(mm[, 1], mm[, 2])  
abline(c(0, -1), col = "red")  
abline(c(0, 1), col = "blue")
```

