Rozdział 2 - ćwiczenia

## Ćwiczenia

1. Wczytaj plik *01 - adult.data* oraz z pliku *01 - adult.names* pobierz nazwy kolumn i przypisz do wiersza nagłówkowego. Wybierz nazwy dla każdej kolumny.

rm(list=ls())  
Dane<-read.table("./Dane/01 - adult.data",  
 stringsAsFactors = T, sep = ",")  
Naglowek<-read.table("./Dane/01 - adult.names",  
 stringsAsFactors = F, sep = ":",comment.char = "|")  
colnames(Dane)<-rownames(Naglowek)

1. Jakie typy danych ma każda kolumna?
2. Zmień odpowiednią kolumnę na *Płeć* z polskojęzycznymi danymi.

colnames(Dane)[colnames(Dane)=="sex"]<-"Płeć"  
levels(Dane$'Płeć')<-c("Kobieta","Mężczyzna")

1. Zmień odpowiednią kolumnę na *Kraj* i wiersze ze znakiem zapytania na *brak danych*.

colnames(Dane)[colnames(Dane)=="native-country"]<-"Kraj"  
levels(Dane$Kraj)[levels(Dane$Kraj)==" ?"]<-"brak danych"

1. Podaj średni wiek, ilość osób w grupach wynikających z pochodzenia (*Kraj*) i płci.

tabela1<-aggregate(age~Kraj+`Płeć`,Dane,length)  
tabela2<-aggregate(age~Kraj+`Płeć`,Dane,function(x) round(mean(x),2))  
tabela<-merge(tabela1,tabela2,by.x=c("Kraj","Płeć"),by.y=c("Kraj","Płeć"))  
colnames(tabela)<-c("Kraj","Płeć","Ilość osób","Średni wiek")  
tail(tabela,n=8)

## Kraj Płeć Ilość osób Średni wiek  
## 76 United-States Kobieta 9682 36.78  
## 77 United-States Mężczyzna 19488 39.59  
## 78 Vietnam Kobieta 23 30.48  
## 79 Vietnam Mężczyzna 44 35.93  
## 80 Yugoslavia Kobieta 3 40.33  
## 81 Yugoslavia Mężczyzna 13 38.46  
## 82 brak danych Kobieta 163 37.23  
## 83 brak danych Mężczyzna 420 39.31

1. Wynik poprzedniego zadania zapisać w Excelu

library(writexl)  
write\_xlsx(tabela,"./Dane/01 - adult - raport.xlsx")

1. Oblicz ile osób pracuje więcej niż 40 godzin w tygodniu z podziałem na rasę.

aggregate(age~race,subset(Dane,Dane$`hours-per-week`>40),length)

## race age  
## 1 Amer-Indian-Eskimo 71  
## 2 Asian-Pac-Islander 236  
## 3 Black 461  
## 4 Other 53  
## 5 White 8760

1. Porównać rodowitych amerykanów z obywatelami USA o korzeniach polskich pod kątem ilości godzin pracy w tygodniu.