Rozdział 8 Współpraca R z bazami danych - ćwiczenia

1. Połaczenie z bazą danych MS Access przez pakiet DBI

rm(list=ls()) #usuwanie danych  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(DBI)  
library(knitr)  
polaczenie <- dbConnect(odbc::odbc(), .connection\_string = "Driver={Microsoft Access Driver (\*.mdb, \*.accdb)};Dbq=C:/OneDrive/OneDrive - Norbert Mielczarek/Projekty/Altkom/R - Analiza, przetwarzanie i wizualizacja danych/Skrypty R/Dane/09 - hurtownia.accdb", timeout = 10,encoding="CP1250")  
dbListTables(polaczenie) %>% as.data.frame() %>% rename(.,Tabele=.) %>%   
 filter(grepl("MSys",Tabele)==F) %>% kable()

|  |
| --- |
| Tabele |
| tblKategorie |
| tblKlienci |
| tblOpisSprzedazy |
| tblPracownicy |
| tblSprzedaz |
| tblTowary |

kwerenda<-"SELECT TOP 3 \* FROM tblKlienci"  
Dane<-dbGetQuery(polaczenie,kwerenda) %>% kable()  
dbListFields(polaczenie,"tblKlienci") %>% as.data.frame() %>% rename(.,Pole=.) %>%   
 kable()

|  |
| --- |
| Pole |
| KodKlienta |
| NazwaFirmy |
| NIP |
| Telefon |
| TelefonKomorkowy |
| Ulica |
| Kod |
| Miasto |
| Aktywny |

dbDisconnect(polaczenie)

1. Połaczenie z bazą danych MS Access przez pakier RODBC

rm(list=ls()) #usuwanie danych  
library(RODBC)  
library(dplyr)  
#działa w wersji sesji R 32-bitowej  
polaczenie <- odbcConnectAccess2007("./Dane/09 - hurtownia.accdb")  
sqlTables(polaczenie) %>% filter(grepl("MSys",TABLE\_NAME)==F) %>% kable()

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| TABLE\_CAT | TABLE\_SCHEM | TABLE\_NAME | TABLE\_TYPE | REMARKS |
| .\09 - hurtownia.accdb | NA | tblKategorie | TABLE | NA |
| .\09 - hurtownia.accdb | NA | tblKlienci | TABLE | NA |
| .\09 - hurtownia.accdb | NA | tblOpisSprzedazy | TABLE | NA |
| .\09 - hurtownia.accdb | NA | tblPracownicy | TABLE | NA |
| .\09 - hurtownia.accdb | NA | tblSprzedaz | TABLE | NA |
| .\09 - hurtownia.accdb | NA | tblTowary | TABLE | NA |

kwerenda<-"SELECT TOP 3 \* FROM tblKlienci"  
Dane<-sqlQuery(polaczenie,kwerenda,max=0,stringsAsFactors=F) %>% kable()  
close(polaczenie)

1. Test szybkości

rm(list=ls()) #usuwanie danych  
library(nycflights13)

## Warning: package 'nycflights13' was built under R version 3.6.2

#działa w wersji sesji R 32-bitowej  
polaczenie <- odbcConnectAccess2007("C:/OneDrive/hurtownia2.accdb")  
Stat<-summary(system.time(  
 sqlSave(polaczenie, flights[,-19], tablename = "loty",  
 append=FALSE,rownames = FALSE)))  
Stat<-rbind(Stat,summary(system.time(  
 Dane<-sqlQuery(polaczenie,"SELECT \* FROM loty"))))  
sqlDrop(polaczenie,"loty")  
close(polaczenie)  
  
polaczenie <- dbConnect(odbc::odbc(), .connection\_string = "Driver={Microsoft Access Driver (\*.mdb, \*.accdb)};Dbq=C:/OneDrive/hurtownia2.accdb")  
Stat<-rbind(Stat,summary(system.time(  
 dbWriteTable(polaczenie, "loty", flights[,-19],batch\_row=1,overwrite=T))))  
Stat<-rbind(Stat,summary(system.time(  
 Dane<-dbGetQuery(polaczenie,"SELECT \* FROM loty"))))  
dbSendQuery(polaczenie,"DROP TABLE loty")

## <OdbcResult>  
## SQL DROP TABLE loty  
## ROWS Fetched: 0 [complete]  
## Changed: 0

dbDisconnect(polaczenie)

## Warning in connection\_release(conn@ptr): There is a result object still in use.  
## The connection will be automatically released when it is closed

rownames(Stat)<-c("RODBC-zapis","RODBC-odczyt","DBI-zapis","DBI-odczyt")  
kable(Stat)

|  |  |  |  |
| --- | --- | --- | --- |
|  | user | system | elapsed |
| RODBC-zapis | 200.30 | 53.00 | 256.30 |
| RODBC-odczyt | 4.20 | 0.17 | 4.52 |
| DBI-zapis | 251.36 | 54.19 | 310.67 |
| DBI-odczyt | 5.35 | 0.11 | 5.44 |

# Ćwiczenia

1. Podłącz się do dowolnego serwera (MS SQLServer,Oracle…) i wykonaj test szybkości zapisu i odczytu porównując opisane biblioteki,
2. Wyślij na serwer wszystkie tabele bazy danych ***09 - hurtownia.accdb***,
3. Wykonaj analizę i wizualizację danych wysłanych na serwer związanych z obrotami klientów,