# Import swimming\_pools.csv: pools

pools = read.csv("swimming\_pools.csv")

# Print the structure of pools

str(pools)

==

# Import swimming\_pools.csv correctly: pools

pools = read.csv("swimming\_pools.csv", stringsAsFactors=FALSE)

# Check the structure of pools

str(pools)

==

# Import hotdogs.txt: hotdogs

hotdogs = read.delim("hotdogs.txt",header=FALSE)

# Summarize hotdogs

summary(hotdogs)

==

# Path to the hotdogs.txt file: path

path <- file.path("data", "hotdogs.txt")

# Import the hotdogs.txt file: hotdogs

hotdogs <- read.table(path,

sep = "\t",

col.names = c("type", "calories", "sodium"))

# Call head() on hotdogs

head(hotdogs)

==

# Finish the read.delim() call

hotdogs <- read.delim("hotdogs.txt", header = FALSE, col.names = c("type", "calories", "sodium"))

# Select the hot dog with the least calories: lily

lily <- hotdogs[which.min(hotdogs$calories), ]

# Select the observation with the most sodium: tom

tom <- hotdogs[which.max(hotdogs$sodium), ]

# Print lily and tom

print(lily)

print(tom)

==

# Previous call to import hotdogs.txt

hotdogs <- read.delim("hotdogs.txt", header = FALSE, col.names = c("type", "calories", "sodium"))

# Display structure of hotdogs

str(hotdogs)

# Edit the colClasses argument to import the data correctly: hotdogs2

hotdogs2 <- read.delim("hotdogs.txt", header = FALSE,

col.names = c("type", "calories", "sodium"),

colClasses = c("factor", "NULL", "numeric"))

# Display structure of hotdogs2

str(hotdogs2)

==

# Load the readr package

library(readr)

# Import potatoes.csv with read\_csv(): potatoes

potatoes <- read\_csv("potatoes.csv")

==

# readr is already loaded

# Column names

properties <- c("area", "temp", "size", "storage", "method",

"texture", "flavor", "moistness")

# Import potatoes.txt: potatoes

potatoes = read\_tsv("potatoes.txt", col\_names=properties)

# Call head() on potatoes

head(potatoes)

==

# readr is already loaded

# Column names

properties <- c("area", "temp", "size", "storage", "method",

"texture", "flavor", "moistness")

# Import potatoes.txt using read\_delim(): potatoes

potatoes = read\_delim('potatoes.txt', delim = '\t', col\_names = properties)

# Print out potatoes

print(potatoes)

==

# readr is already loaded

# Column names

properties <- c("area", "temp", "size", "storage", "method",

"texture", "flavor", "moistness")

# Import potatoes.txt using read\_delim(): potatoes

potatoes = read\_delim('potatoes.txt', delim = '\t', col\_names = properties)

# Print out potatoes

print(potatoes)

==

# readr is already loaded

# Column names

properties <- c("area", "temp", "size", "storage", "method",

"texture", "flavor", "moistness")

# Import all data, but force all columns to be character: potatoes\_char

potatoes\_char <- read\_tsv("potatoes.txt", col\_types = "cccccccc", col\_names = properties)

# Print out structure of potatoes\_char

print(str(potatoes\_char))

==

# readr is already loaded

# Import without col\_types

hotdogs <- read\_tsv("hotdogs.txt", col\_names = c("type", "calories", "sodium"))

# Display the summary of hotdogs

summary(hotdogs)

# The collectors you will need to import the data

fac <- col\_factor(levels = c("Beef", "Meat", "Poultry"))

int <- col\_integer()

# Edit the col\_types argument to import the data correctly: hotdogs\_factor

hotdogs\_factor <- read\_tsv("hotdogs.txt",

col\_names = c("type", "calories", "sodium"),

col\_types = list(fac,int,int))

# Display the summary of hotdogs\_factor

summary(hotdogs\_factor)

==

# load the data.table package using library()

library(data.table)

# Import potatoes.csv with fread(): potatoes

potatoes <- fread('potatoes.csv')

# Print out potatoes

print(potatoes)

==

# fread is already loaded

# Import columns 6 and 8 of potatoes.csv: potatoes

potatoes=fread('potatoes.csv',select=c(6,8))

# Plot texture (x) and moistness (y) of potatoes

plot(potatoes$texture, potatoes$moistness)

==

# Load the readxl package

library(readxl)

# Print the names of all worksheets

excel\_sheets('urbanpop.xlsx')

==

# The readxl package is already loaded

# Read the sheets, one by one

pop\_1 <- read\_excel("urbanpop.xlsx", sheet = 1)

pop\_2 <- read\_excel("urbanpop.xlsx", sheet = 2)

pop\_3 <- read\_excel("urbanpop.xlsx", sheet = 3)

# Put pop\_1, pop\_2 and pop\_3 in a list: pop\_list

pop\_list=list(pop\_1, pop\_2, pop\_3)

# Display the structure of pop\_list

str(pop\_list)

==

# The readxl package is already loaded

# Read all Excel sheets with lapply(): pop\_list

pop\_list = lapply(excel\_sheets('urbanpop.xlsx'), read\_excel, path='urbanpop.xlsx')

# Display the structure of pop\_list

str(pop\_list)

==

# The readxl package is already loaded

# Import the first Excel sheet of urbanpop\_nonames.xlsx (R gives names): pop\_a

pop\_a = read\_excel('urbanpop\_nonames.xlsx',col\_names=FALSE)

# Import the first Excel sheet of urbanpop\_nonames.xlsx (specify col\_names): pop\_b

cols <- c("country", paste0("year\_", 1960:1966))

pop\_b = read\_excel('urbanpop\_nonames.xlsx',col\_names=cols)

# Print the summary of pop\_a

summary(pop\_a)

# Print the summary of pop\_b

summary(pop\_b)

==

# The readxl package is already loaded

# Import the second sheet of urbanpop.xlsx, skipping the first 21 rows: urbanpop\_sel

urbanpop\_sel = read\_excel('urbanpop.xlsx', sheet=2, col\_names=FALSE, skip=21)

# Print out the first observation from urbanpop\_sel

urbanpop\_sel[1,]

==

# Load the gdata package

library(gdata)

# Import the second sheet of urbanpop.xls: urban\_pop

urban\_pop = read.xls('urbanpop.xls', sheet=2)

# Print the first 11 observations using head()

head(urban\_pop, 11)

==

# The gdata package is alreaded loaded

# Column names for urban\_pop

columns <- c("country", paste0("year\_", 1967:1974))

# Finish the read.xls call

urban\_pop <- read.xls("urbanpop.xls", sheet = 2,

skip = 50, header = FALSE, stringsAsFactors = FALSE,

col.names = columns)

# Print first 10 observation of urban\_pop

head(urban\_pop,10)

==

# Add code to import data from all three sheets in urbanpop.xls

path <- "urbanpop.xls"

urban\_sheet1 <- read.xls(path, sheet = 1, stringsAsFactors = FALSE)

urban\_sheet2 <- read.xls(path, sheet = 2, stringsAsFactors = FALSE)

urban\_sheet3 <- read.xls(path, sheet = 3, stringsAsFactors = FALSE)

# Extend the cbind() call to include urban\_sheet3: urban

urban <- cbind(urban\_sheet1, urban\_sheet2[-1], urban\_sheet3[-1])

# Remove all rows with NAs from urban: urban\_clean

urban\_clean = na.omit(urban)

# Print out a summary of urban\_clean

summary(urban\_clean)

==

# urbanpop.xlsx is available in your working directory

# Load the XLConnect package

library(XLConnect)

# Build connection to urbanpop.xlsx: my\_book

my\_book = loadWorkbook('urbanpop.xlsx')

# Print out the class of my\_book

class(my\_book)

==

# XLConnect is already available

# Build connection to urbanpop.xlsx

my\_book <- loadWorkbook("urbanpop.xlsx")

# List the sheets in my\_book

getSheets(my\_book)

# Import the second sheet in my\_book

readWorksheet(my\_book, sheet=2)

==

# XLConnect is already available

# Build connection to urbanpop.xlsx

my\_book <- loadWorkbook("urbanpop.xlsx")

# Import columns 3, 4, and 5 from second sheet in my\_book: urbanpop\_sel

urbanpop\_sel <- readWorksheet(my\_book, sheet = 2, startCol=3, endCol=5)

# Import first column from second sheet in my\_book: countries

countries = readWorksheet(my\_book, sheet=2, startCol=1, endCol=1)

# cbind() urbanpop\_sel and countries together: selection

selection = cbind(countries, urbanpop\_sel)

==

# XLConnect is already available

# Build connection to urbanpop.xlsx

my\_book <- loadWorkbook("urbanpop.xlsx")

# Add a worksheet to my\_book, named "data\_summary"

createSheet(my\_book, name = 'data\_summary')

# Use getSheets() on my\_book

getSheets(my\_book)

==

# XLConnect is already available

# Build connection to urbanpop.xlsx

my\_book <- loadWorkbook("urbanpop.xlsx")

# Add a worksheet to my\_book, named "data\_summary"

createSheet(my\_book, "data\_summary")

# Create data frame: summ

sheets <- getSheets(my\_book)[1:3]

dims <- sapply(sheets, function(x) dim(readWorksheet(my\_book, sheet = x)), USE.NAMES = FALSE)

summ <- data.frame(sheets = sheets,

nrows = dims[1, ],

ncols = dims[2, ])

# Add data in summ to "data\_summary" sheet

writeWorksheet(my\_book, summ, sheet='data\_summary')

# Save workbook as summary.xlsx

saveWorkbook(my\_book, file='summary.xlsx')

==

# my\_book is available

# Rename "data\_summary" sheet to "summary"

renameSheet(my\_book, 4, "summary")

# Print out sheets of my\_book

getSheets(my\_book)

# Save workbook to "renamed.xlsx"

saveWorkbook(my\_book, file='renamed.xlsx')

==

# Load the XLConnect package

library(XLConnect)

# Build connection to renamed.xlsx: my\_book

my\_book = loadWorkbook('renamed.xlsx')

# Remove the fourth sheet

removeSheet(my\_book, 4)

# Save workbook to "clean.xlsx"

saveWorkbook(my\_book, file='clean.xlsx')

==