#install.packages("readxl")  
library(readxl)  
library(tinytex)

## Warning: package 'tinytex' was built under R version 4.3.2

#Read the excel data file (.xlsx) and call it "myData"  
myData <- read\_excel("C:/Users/leile/OneDrive/School-Kent/BAI/mydata/Module3/Birth\_Life.xlsx")  
View(myData)  
str(myData)

## tibble [10 × 5] (S3: tbl\_df/tbl/data.frame)  
## $ Country Name: chr [1:10] "Congo, Dem. Rep." "India" "South Africa" "China" ...  
## $ Life\_Exp : num [1:10] 50 62.6 56.4 72 62.8 ...  
## $ Birth\_Rate : num [1:10] 46 26.5 24.3 14 32 ...  
## $ GNI : num [1:10] 130 440 3070 940 490 ...  
## $ Development : chr [1:10] "Developing" "Developing" "Developing" "Developing" ...

dim(myData)

## [1] 10 5

# Get the summary statistics  
summary (myData)

## Country Name Life\_Exp Birth\_Rate GNI   
## Length:10 Min. :50.00 Min. : 9.30 Min. : 130.0   
## Class :character 1st Qu.:62.63 1st Qu.:11.68 1st Qu.: 602.5   
## Mode :character Median :73.15 Median :19.27 Median : 4440.0   
## Mean :69.33 Mean :21.10 Mean :15285.0   
## 3rd Qu.:77.60 3rd Qu.:25.93 3rd Qu.:33605.0   
## Max. :81.08 Max. :45.96 Max. :43460.0   
## Development   
## Length:10   
## Class :character   
## Mode :character   
##   
##   
##

# Descriptive statistics for quantitative variables  
summary (myData$Birth\_Rate)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 9.30 11.68 19.27 21.10 25.93 45.96

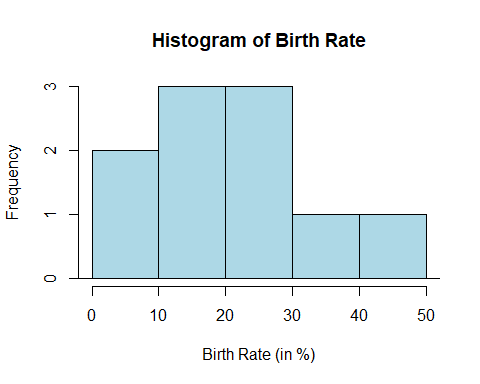
# Descriptive statistics for categorical variable  
table(myData$'Country Name')

##   
## China Congo, Dem. Rep. Germany India   
## 1 1 1 1   
## Japan Mexico Pakistan South Africa   
## 1 1 1 1   
## Switzerland United States   
## 1 1

# Transform a variable (e.g., square root transformation)  
Transform\_Life\_Exp <- sqrt(myData$Life\_Exp)  
summary (Transform\_Life\_Exp)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 7.071 7.914 8.552 8.303 8.809 9.004

# Plot a quantitative variable  
hist(myData$Birth\_Rate,   
 main = "Histogram of Birth Rate",   
 xlab = "Birth Rate (in %)",   
 col = "lightblue")  
 abline (h=0)



hist

## function (x, ...)   
## UseMethod("hist")  
## <bytecode: 0x000002797d436e30>  
## <environment: namespace:graphics>

# Do the scatter plot for Birth Rate and Life Expectancy  
plot ( myData$Birth\_Rate ~ myData$Life\_Exp,   
 main = "Scatter plot for Birth Rate against Life Expectancy",   
 xlab="Life Expectancy (in years)",   
 ylab = "Birth Rate (in %)",   
 pch=16,   
 col = ifelse ( myData$Development == "Developing", 20, 26) )  
legend ("topright", legend = c("Developing", "Developed"), pch=16, col = c(20, 26) )

