

This code is part of NCRM Online Tutorial:

**Producing Automated Outputs (using R)** <https://www.ncrm.ac.uk/resources/online/all/?id=20832>

#Install and load required packages

install.packages("gtsummary")

install.packages("flextable")

install.packages("officer")

install.packages("car")

library(gtsummary)

library(flextable)

library(officer)

library(car)

#Create a dataset called 'cars' by loading the built-in 'mtcars' dataset

#Some variables are then removed to create a smaller dataset

cars <- mtcars

head(cars)

cars$drat<-NULL

cars$qsec<-NULL

cars$hp<-NULL

cars$gear<-NULL

cars$carb<-NULL

cars$disp<-NULL

head(cars)

summary(cars)

#Examine distribution of variables

hist(cars$mpg) # Approximately normally distributed

table(cars$cyl) # Values of 4, 6, 8

hist(cars$wt) # Continuous, not normally distributed

table(cars$vs) # Values of 0, 1

table(cars$am) # Values of 0, 1

#Convert weight from 1000s of pounds to tonnes

cars$wt <- cars$wt \* 0.454

#Declare categorical variables as factor variables and label the categories

cars$vs <- factor(cars$vs, levels = c(0,1), labels = c("V-shaped", "Straight"))

cars$am <- factor(cars$am, levels = c(0,1), labels = c("Automatic", "Manual"))

cars$cyl <- factor(cars$cyl, levels = c(4,6,8))

#Re-order dataset columns and examine final dataset

cars <- cars[, c("am", "wt", "vs", "cyl", "mpg")]

head(cars)

summary(cars)

#Base this table on the 'cars' dataset

#Specify which variables are continuous and categorical

#Specify the name to use in the table for each variable

#State the descriptive statistics to use for each variable

#Specify the file path and file name for the Word document

tbl\_summary(cars,

 type = list(c(wt, mpg) ~ "continuous", c(am, vs, cyl) ~ "categorical"),

 label = list(

 am ~ "Transmission type",

 wt ~ "Weight (tonnes)",

 vs ~ "Engine configuration",

 cyl ~ "Number of cylinders",

 mpg ~ "Fuel economy (miles per gallon)"),

 statistic = list(

 c(mpg) ~ "{mean} ({sd})",

 c(wt) ~ "{median} ({p25}, {p75})",

 c(am, vs, cyl) ~ "{n} ({p}%)")

) %>%

as\_flex\_table() %>%

flextable::save\_as\_docx(path="C:/Users/ldw1c13/Documents/Descriptives Table 1.docx")

#Stratify the above table according to transmission type

#Specify the stratification variable using ‘by = am’

tbl\_summary(cars,

 by = am,

 type = list(c(wt, mpg) ~ "continuous", c(vs, cyl) ~ "categorical"),

 label = list(

 wt ~ "Weight (tonnes)",

 vs ~ "Engine configuration",

 cyl ~ "Number of cylinders",

 mpg ~ "Fuel economy (miles per gallon)"),

 statistic = list(

 c(mpg) ~ "{mean} ({sd})",

 c(wt) ~ "{median} ({p25}, {p75})",

 c(vs, cyl) ~ "{n} ({p}%)")

) %>%

as\_flex\_table() %>%

flextable::save\_as\_docx(path="C:/Users/ldw1c13/Documents/Descriptives Table 2.docx")

#Define the model ‘m1’ and use it in the regression table

#Specify the name to use in the table for each exposure

#Do not display the intercept in the table

#Only show one row for the binary exposure engine configuration

#Modify column headings and footnotes

m1 <- lm(mpg ~ wt + vs + relevel(cyl, ref = "8"), data=cars)

tbl\_regression(m1,

 label = list(

 wt ~ "Weight (tonnes)",

 vs ~ "Engine configuration (straight vs v-shaped)",

 'relevel(cyl, ref = "8")' ~ "Number of cylinders"),

 intercept = FALSE,

 show\_single\_row = c(vs)

) %>%

modify\_header(label = "\*\*Exposure\*\*", estimate = "\*\*Estimate\*\*", ci = "\*\*95% CI\*\*", p.value = "\*\*P-value\*\*") %>%

modify\_footnote(ci = "CI: Confidence interval", abbreviation = TRUE) %>%

modify\_footnote(estimate ~ "Difference in fuel economy (miles per gallon) according to exposure; exposures were included simultaneously in the model") %>%

as\_flex\_table() %>%

flextable::save\_as\_docx(path="C:/Users/ldw1c13/Documents/Regression Table 1.docx")

#Define the logistic regression model ‘m2’ and use it in the regression table

#Display exponentiated coefficients (odds ratios) in the table

m2 <- glm(am ~ mpg + vs, data=cars, family=binomial)

tbl\_regression(m2, exponentiate=TRUE,

 label = list(

 mpg ~ "Fuel economy (miles per gallon)",

 vs ~ "Engine configuration (straight vs v-shaped)"),

 intercept = FALSE,

 show\_single\_row = c(vs)

) %>%

modify\_header(label = "\*\*Exposure\*\*", estimate = "\*\*OR\*\*", ci = "\*\*95% CI\*\*", p.value = "\*\*P-value\*\*") %>%

modify\_footnote(estimate = "OR: Odds ratio", abbreviation = TRUE) %>%

modify\_footnote(ci = "CI: Confidence interval", abbreviation = TRUE) %>%

modify\_footnote(estimate ~ "Odds ratios for manual transmission according to exposures; exposures were included simultaneously in the model") %>%

as\_flex\_table() %>%

flextable::save\_as\_docx(path="C:/Users/ldw1c13/Documents/Regression Table 2.docx")