DRY coding

Data Science in Biomedicine - Week 5

2024-10-17

Introduction

We'll be using an in-built data set from the MASS package. But, because the MASS package can be a bit problematic to install, I've provided the data for you. It's in .rds format, so we load it in using the `read_rds()` function.

For more info about the dataset, head to https://stat.ethz.ch/R-manual/R-devel/library/MASS/html/survey.html

```r
library(tidyverse)
library(knitr)

survey <- read_rds("./data/survey.rds")
```

Exercise 1

Explore the data - what variables are in there, what variable types, how many variables and observations?

Exercise 2

Round every numeric variable to 0 decimal places.

Exercise 3

Using the rounded dataframe you created in Exercise 2, calculate the mean of each numeric variable in the dataset. Use kable() to insert a nicely formatted table into your R Markdown document.

Exercise 4

Now make exactly the same table, using the pivot functions. Again, use kable() to insert a nicely formatted table into your R Markdown document.

Exercise 5

Students had initially expressed their height in either imperial or metric units, but someone converted everything into centimeters. Create a new variable, Height_original, which will show height in centimeters for students who expressed their height in metric units, and height in inches for students who prefer imperial units.

NOTE: This one is a bit of a challenge. You may want to check out the `case_when()` function.

Exercise 6

Create a table that shows the mean and standard deviation of height - in centimeters for students who used metric units, and in inches for students who used imperial units. Use kable() to insert a nicely formatted table into your R Markdown document.

Exercise 7

Create a contingency table to see the relationship between exercise and smoking. Pivot your table to make it easier to read, and then use kable() to insert a nicely formatted table into your R Markdown document.

Exercise 8

Create a grid of Pulse histograms - one histogram for each combination of Sex and Exercise.

Exercise 9

Create a new variable, Hnd.diff, that shows the difference in span between the writing and non-writing hand. Create a set of boxplots of this variable, with a separate boxplot for those who write with their right hand, and another one for those who write with their left hand.