Exercies

Natali Tckvitishvili

2024-05-19

Applied Statistics in R

Natali Tckvitishvili

Load libraries

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts --------------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
```

Exercise 1

    a. load data & add new variable - good

```
wine <- read.csv("winequality-white.csv", sep = ";")
print(wine)

wine <- mutate(wine, good = ifelse(quality > 5, 1, 0))
```

    b. residual.sugar analysis

      • histograms for good and bad quality wines

```
ggplot(wine, aes(x = residual.sugar, fill = as.factor(good))) +
geom_histogram(position = "identity", alpha = .5, bins = 100) +
scale_fill_manual(values=c("#69b3a2", "#404080")) +
facet_wrap(vars(good))
```