

Lab Assignment 4

AUTHOR

Mary Boateng

Setup

```
library(readr)
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

```
library(ggplot2)
library(ggthemes)
library(ggstatsplot)
```

You can cite this package as:

Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach.
Journal of Open Source Software, 6(61), 3167, doi:10.21105/joss.03167

import and read data

```
dfr <- readr::read_csv("PlottingData.csv")
```

Rows: 312 Columns: 9

— Column specification —

Delimiter: ","

chr (4): Ventilation, Diabetes, Obesity, Grade

dbl (5): Age, dBP, CRP, HB, HR

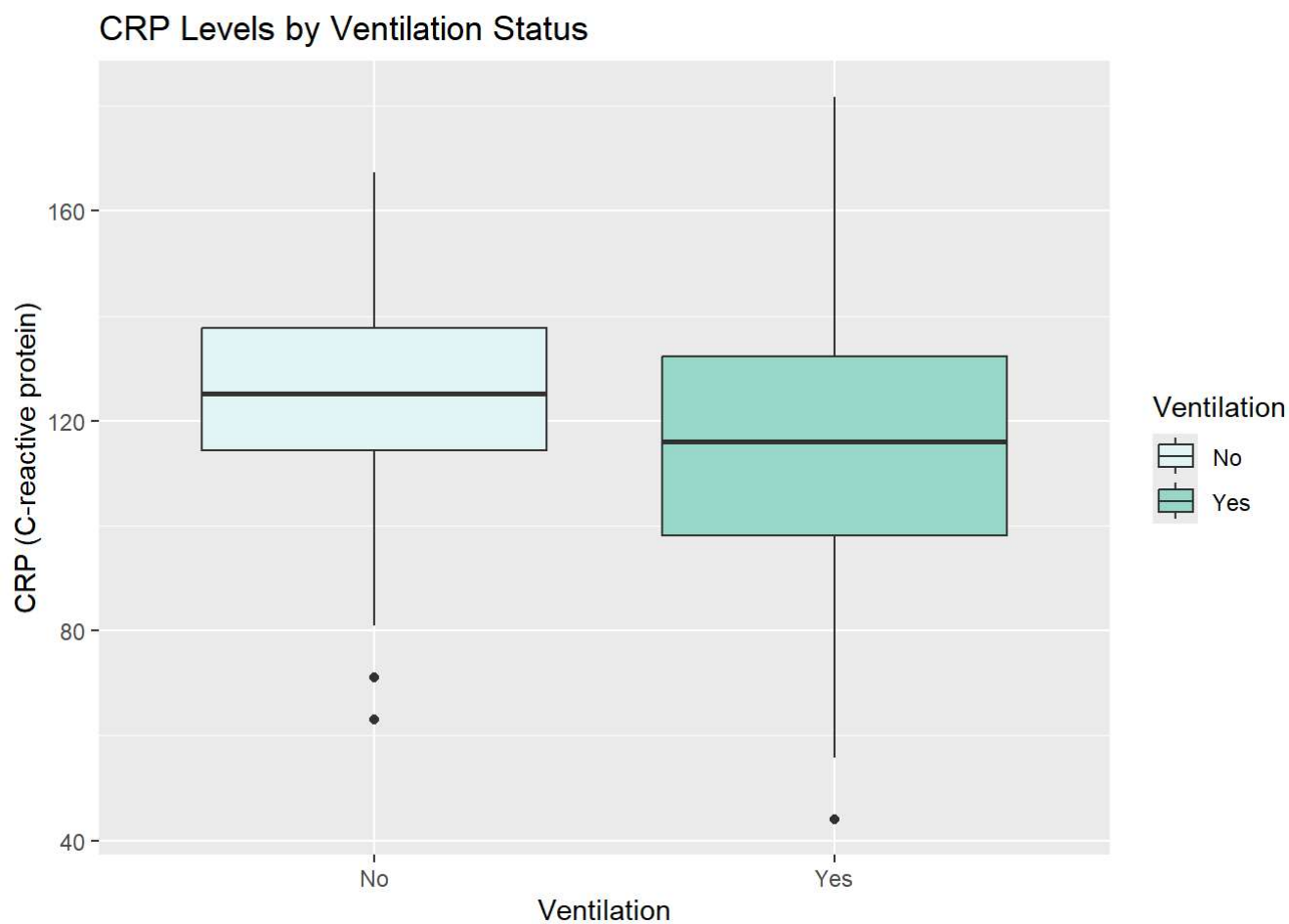
i Use `spec()` to retrieve the full column specification for this data.

i Specify the column types or set `show_col_types = FALSE` to quiet this message.

Section 1

A box-and-whisker plot with the categorical variable **Ventilation** and the numerical variable **CRP**

```
dfr %>%
  ggplot2::ggplot(
    aes(
      x = Ventilation,
      fill = Ventilation,
      y = CRP
    )
  ) +
  ggplot2::geom_boxplot() +
  ggplot2::scale_fill_brewer(
    palette = "BuGn",
    direction = 1
  ) +
  labs(
    title = "CRP Levels by Ventilation Status",
    y = "CRP (C-reactive protein)",
    x = "Ventilation"
  )
)
```

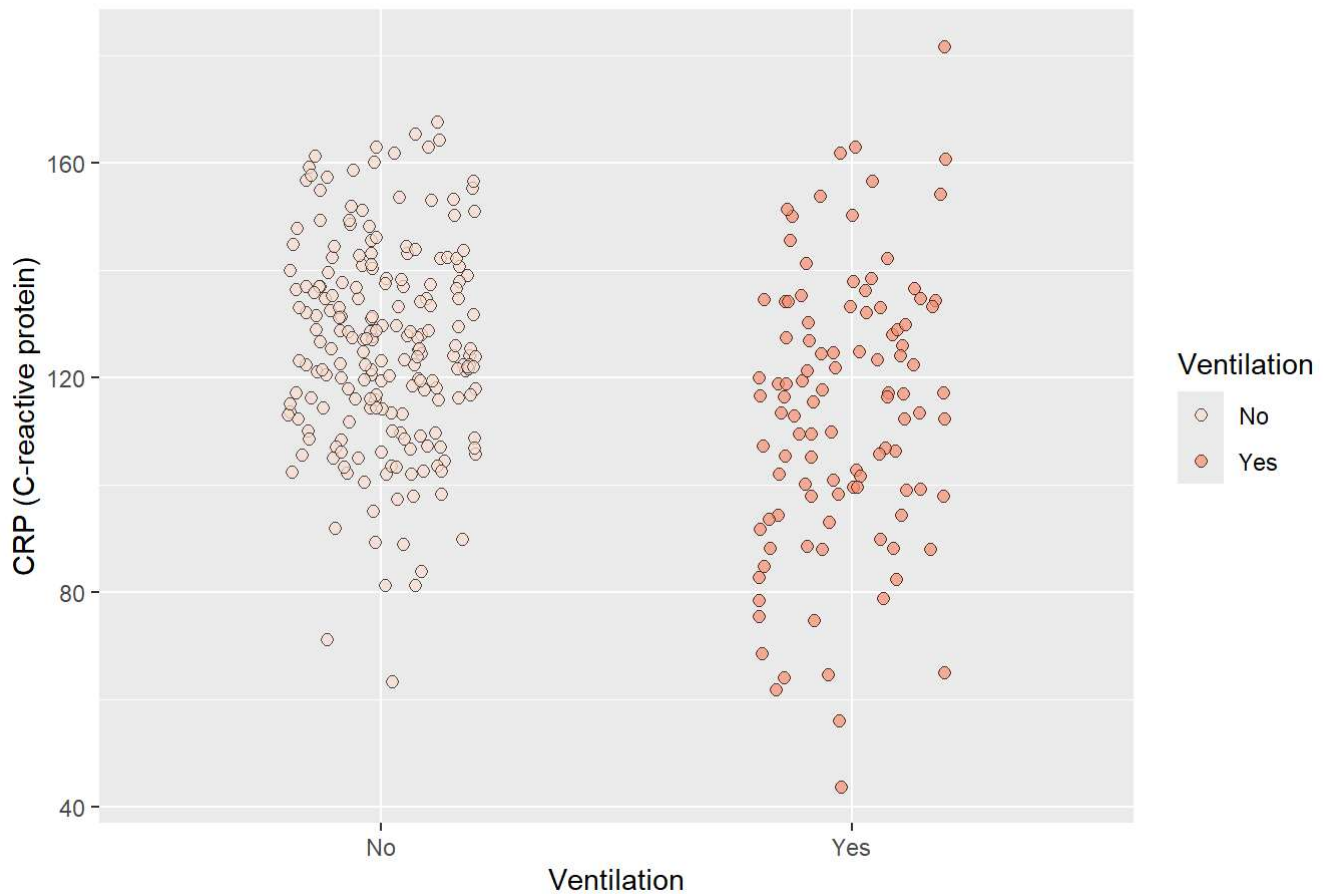


Section 2

Here is a scatter plot (grouped by the levels of the **Ventilation**) made With the a categorical variable **Ventilation** and the numerical variable **CRP**

```
dfr %>%
  ggplot2::ggplot(
    aes(
      x = Ventilation,
      fill = Ventilation,
      y = CRP
    )
  ) +
  ggplot2::geom_jitter(
    shape = 21,
    size = 2,
    width = 0.2,
    alpha = 0.7
  ) +
  ggplot2::scale_fill_brewer(
    palette = "Reds",
    direction = 1
  ) +
  labs(
    title = "CRP Levels by Ventilation Status",
    y = "CRP (C-reactive protein)",
    x = "Ventilation"
  )
```

CRP Levels by Ventilation Status



Section 3

Look up what a mosaic plot is. Import an appropriate library and create a mosaic plot of the variables

```
library(dplyr)
library(ggplot2)
library(ggmosaic)

dfr %>%
  ggplot(aes(x = Obesity, fill = Ventilation)) +
  geom_bar(position = "fill", color = "white") +
  scale_fill_brewer(palette = "Pastel2") +
  labs(
    title = "Proportion of Ventilation by Obesity",
    x = "Obesity",
    y = "Proportion",
    fill = "Ventilation"
  ) +
  theme_minimal()
```



Another way to create a mosaic plot is by using the `vcd` package.

```
library(vcd)
```

Loading required package: grid

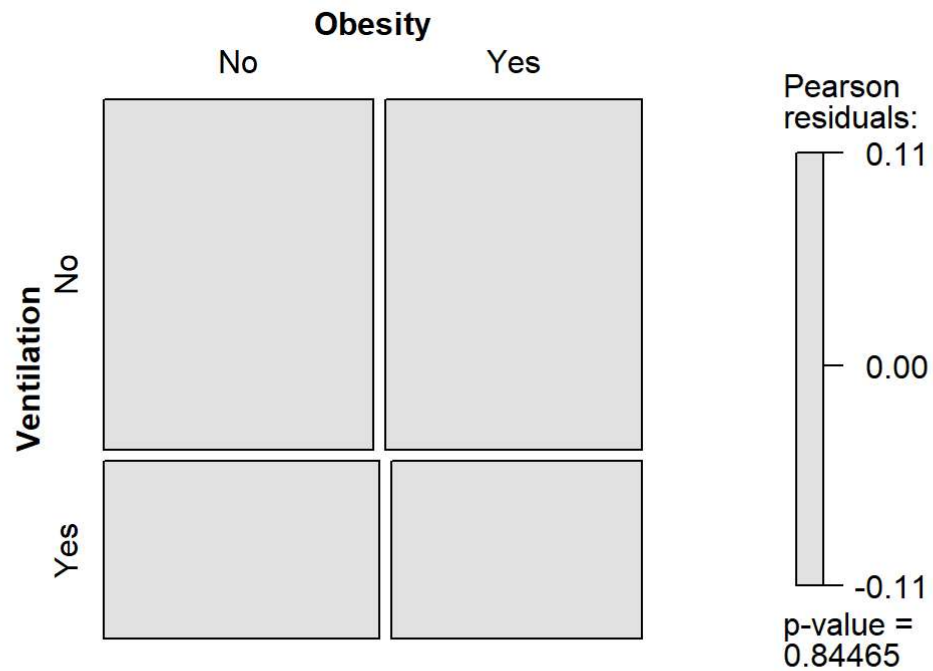
Attaching package: 'vcd'

The following objects are masked from 'package:ggmosaic':

mosaic, spine

```
# Mosaic plot of Ventilation vs Obesity
mosaic(~ Ventilation + Obesity, data = dfr,
       shade = TRUE, legend = TRUE,
       main = "Mosaic Plot of Ventilation and Obesity")
```

Mosaic Plot of Ventilation and Obesity



Add color to the mosaic plot that used the `vcd` package

```
library(vcd)
library(RColorBrewer)

mosaic(~ Ventilation + Obesity, data = dfr,
       gp = gpar(fill = brewer.pal(3, "Set3")),
       main = "Mosaic Plot of Ventilation and Obesity")
```

Mosaic Plot of Ventilation and Obesity

