Assignment 2

Your Name and Surname

2025-10-13

```r
# Installing necessary package(s)
install.packages("ggplot2")
```

```
## Installing package into '/cloud/lib/x86_64-pc-linux-gnu-library/4.5'
## (as 'lib' is unspecified)
```

```r
library(ggplot2)
```

Dataset

The `BreastCancer` dataset contains cytological measurements of breast tissue samples, used to classify tumors as benign or malignant. It includes numerical features describing cell characteristics derived from microscopic examination.

To load the dataset to your environment, just run the code below:

```r
install.packages("mlbench")
install.packages("dplyr")
library(mlbench)
data("BreastCancer")
BreastCancer$Cl.thickness <- as.numeric(BreastCancer$Cl.thickness)
```

Drawing a plot for distribution

1. Please draw a plot to visualize the distribution of tissue thickness (`Cl.thickness`) of the patients according to `Class` indicating whether the tumor in the breast tissue is benign or malignant. (40 pts).

```r
ggplot(BreastCancer, aes(x = Cl.thickness, y = Class)) +
  geom_bar(stat = "identity")
```
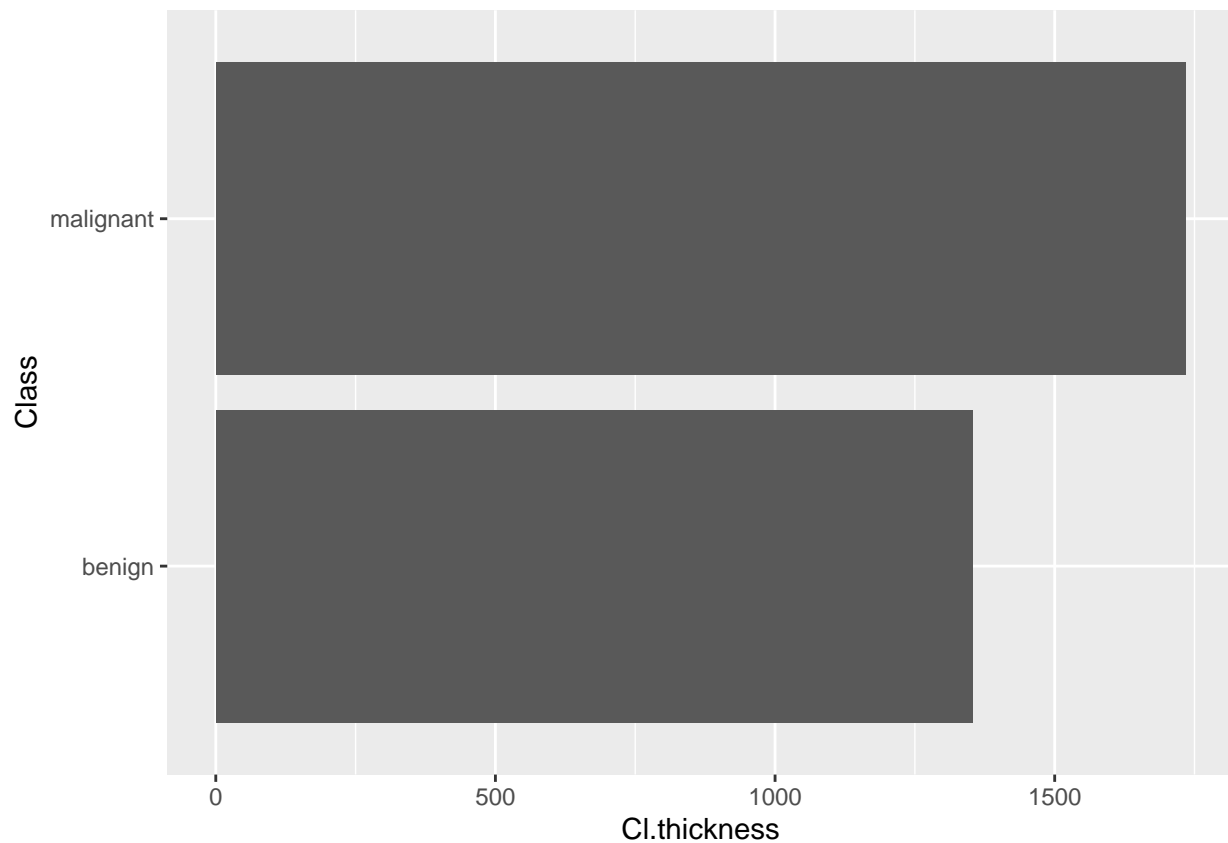
2. Please solve the problems in the plot and make it better by adding some information (30 pts).

```r
library(ggplot2)
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
```

```r
ggplot(BreastCancer, aes(x=Cl.thickness, y=Class, fill=Class)) +
  geom_violin()
```
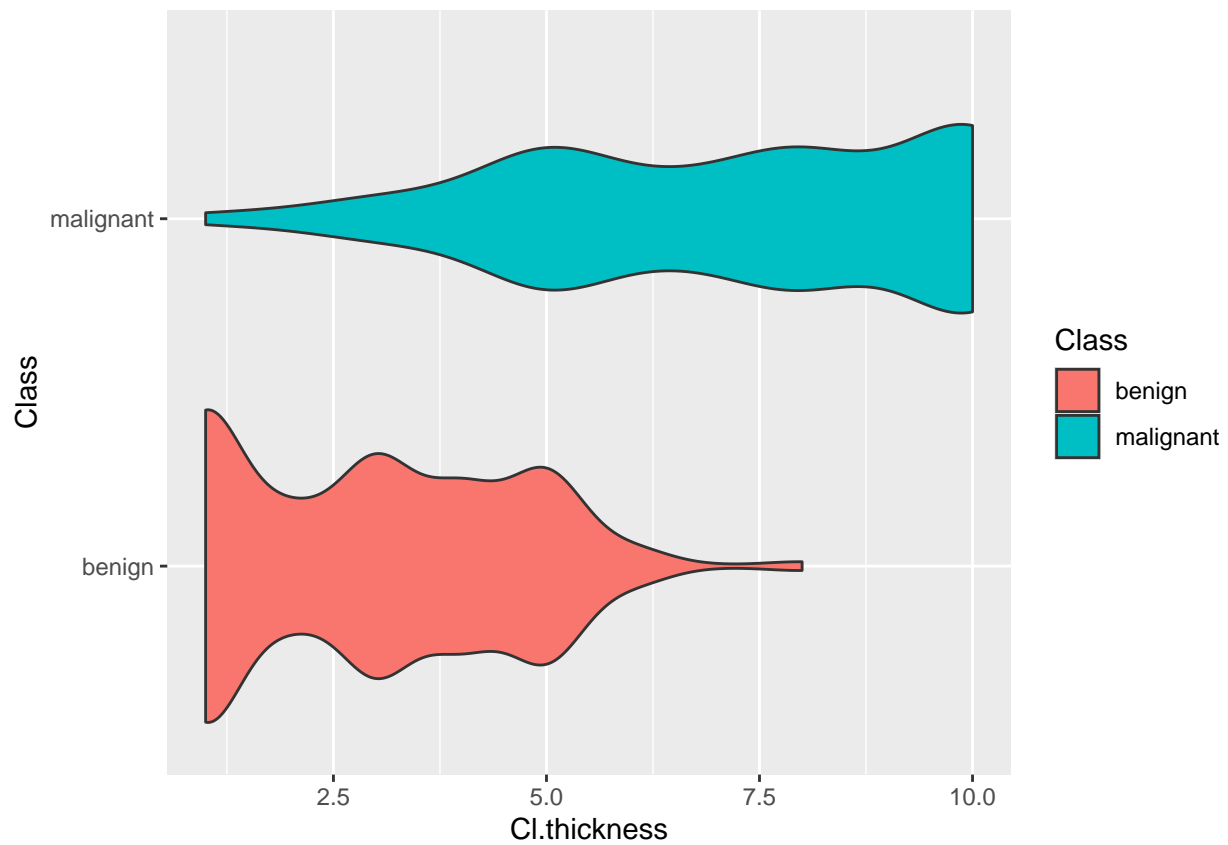
3. Interpret the plot (30 pts).

. . .

the plots shows how many patients has which class tumor that they have and what kind of tumors are more by counts.