

Submit2

2024-08-08

```
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)
```

```
genes <- read.csv("QBS103_GSE157103_genes.csv")
matrix <- read.csv("QBS103_GSE157103_series_matrix.csv")
```

```
plot_create <- function(genes, matrix, item){
  genes_long <- genes %>%
    filter(X==item) %>%
    gather(key = "participant_id")
  genes_long <- genes_long[-1,]
  matrix_selected <- matrix %>% select("participant_id", "age", "icu_status", "sex")
```

```
plot_data <- merge(matrix_selected, genes_long, by = "participant_id")
plot_data$value <- as.numeric(as.character(plot_data$value))
plot_data$age <- as.numeric(as.character(plot_data$age))
```

```
frequency_plot <-
  ggplot(plot_data, aes(x=value)) +
  geom_histogram(binwidth=1, fill="#c8a2c8", color="#d9d9d9") +
  ggtitle(paste("Histogram of Values of ", item)) +
  xlab("Value") +
  ylab("Frequency")
```

```
continuous_plot <-
  ggplot(plot_data, aes(x=age, y=value)) +
  geom_point(aes(color=age)) +
  scale_color_gradient(low="#1E3D59", high="#F5C518") +
  ggtitle(paste("Scatterplot of ", item, " gene vs. Age")) +
  xlab("age") +
  ylab("Value") +
  theme_minimal()
```

```
categorical_plot <-
  ggplot(plot_data, aes(x = icu_status, y = value, fill = sex)) +
```

```

geom_boxplot(notch = TRUE) +
ggtitle(paste("Gene Value of ", item," by icu status and sex") )+
xlab("icu status") +
ylab("Gene Value") +
theme_minimal()

return (list(frequency_plot, continuous_plot, categorical_plot))
}

```

```

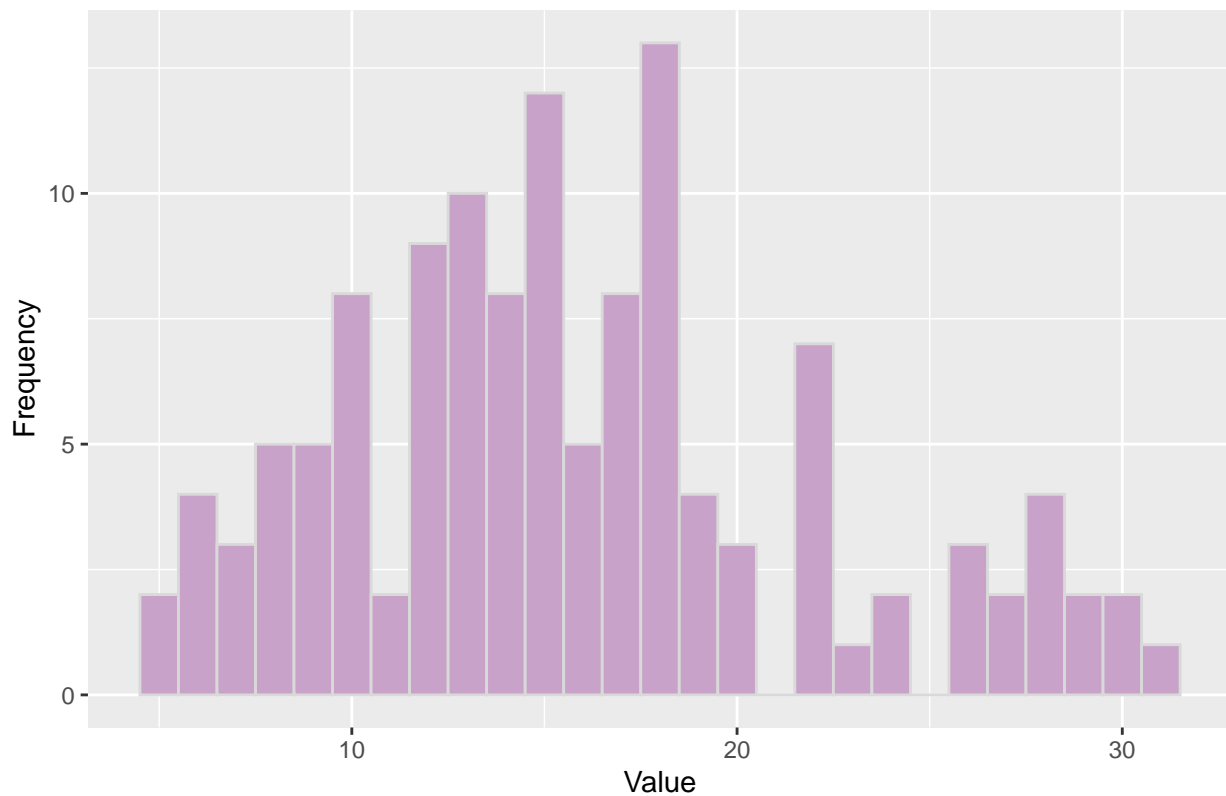
lst = list("AAGAB", "AACS", "ABI1")
for (gene in lst){
  plots = plot_create(genes, matrix, gene)
  print(plots)
}

```

Warning in plot_create(genes, matrix, gene): NAs introduced by coercion

[[1]]

Histogram of Values of AAGAB

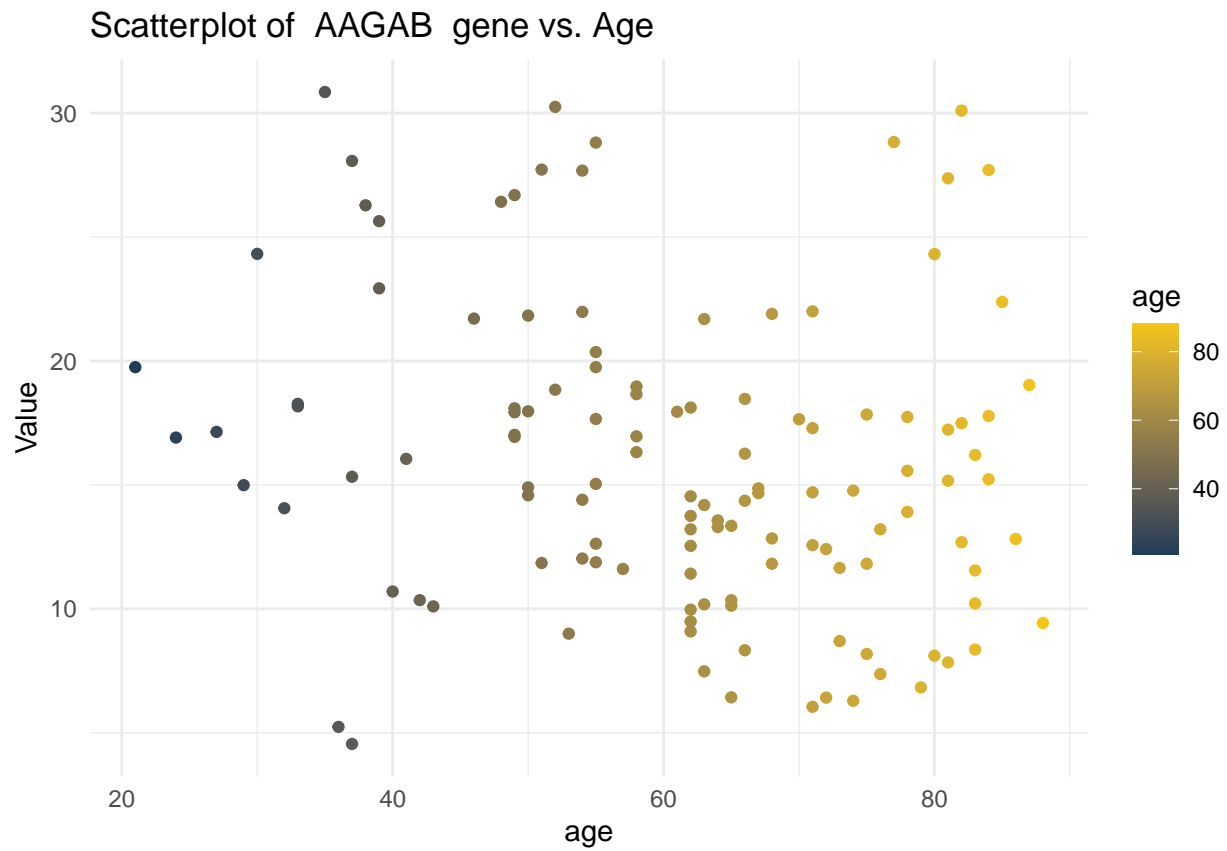


##

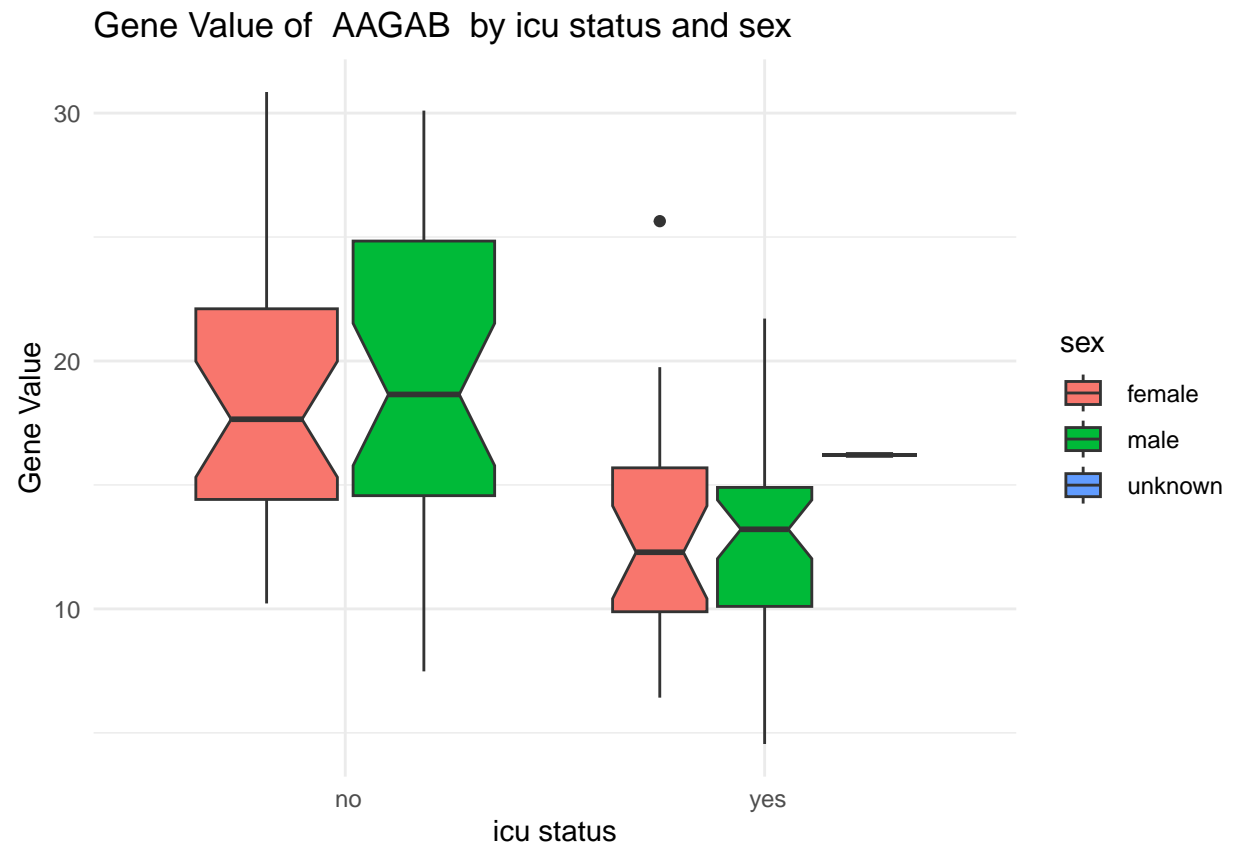
[[2]]

Warning: Removed 2 rows containing missing values or values outside the scale range

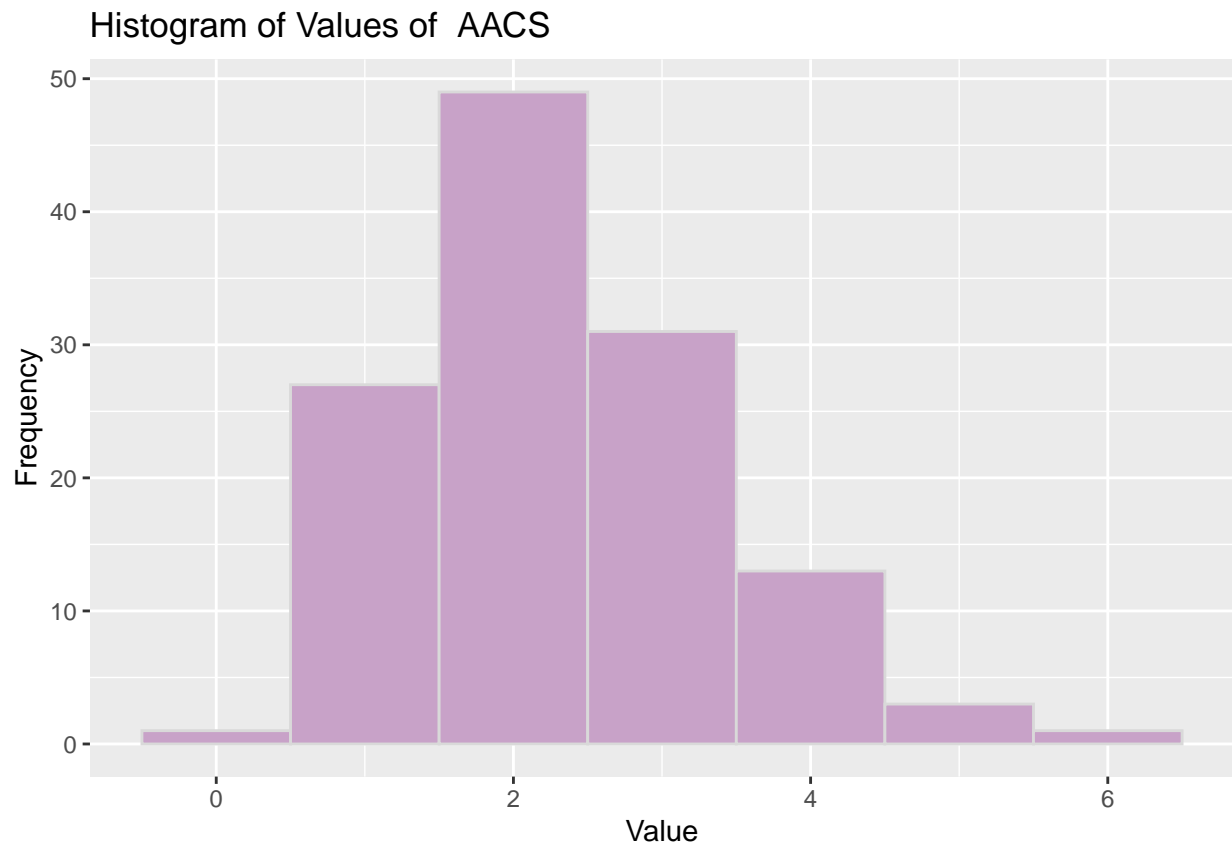
(`geom_point()`).



```
##  
## [[3]]  
## Warning in plot_create(genes, matrix, gene): NAs introduced by coercion
```



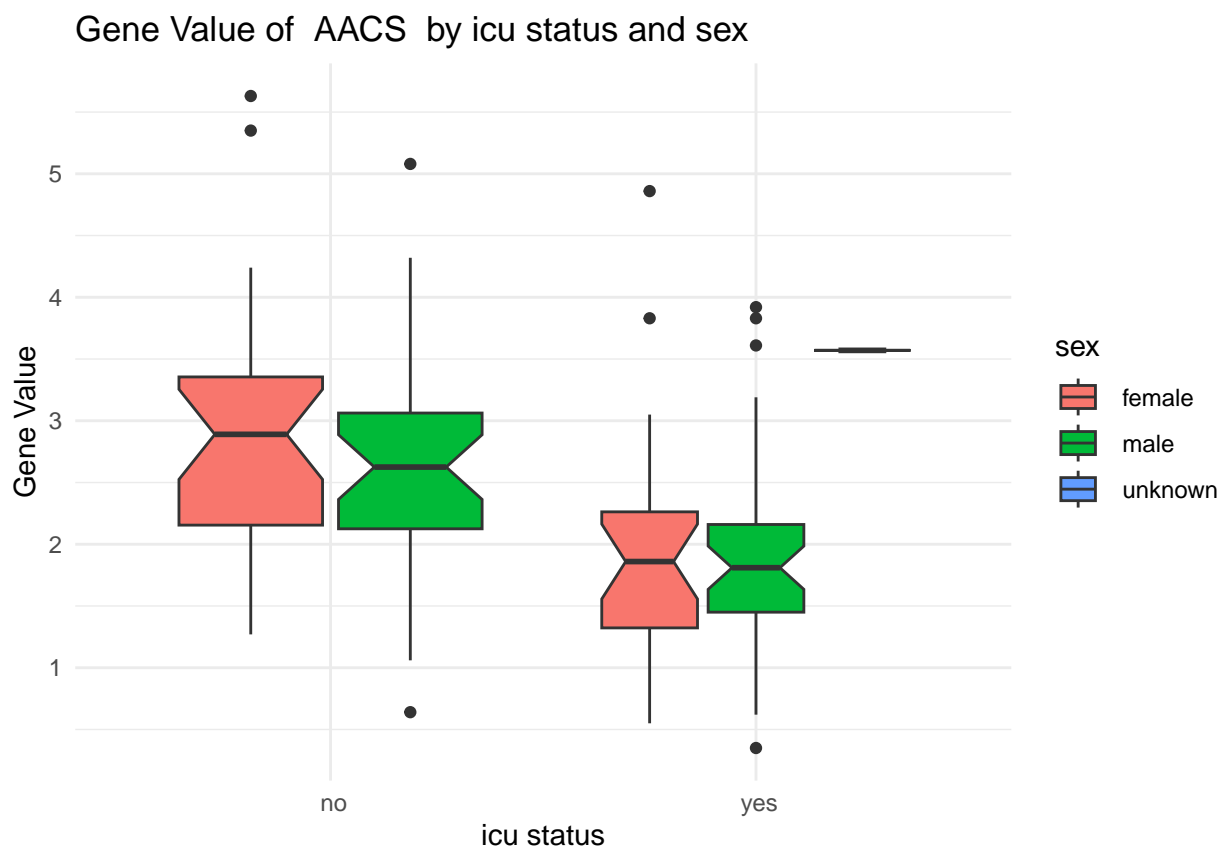
[[1]]



```
##  
## [[2]]  
  
## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom_point()`).
```

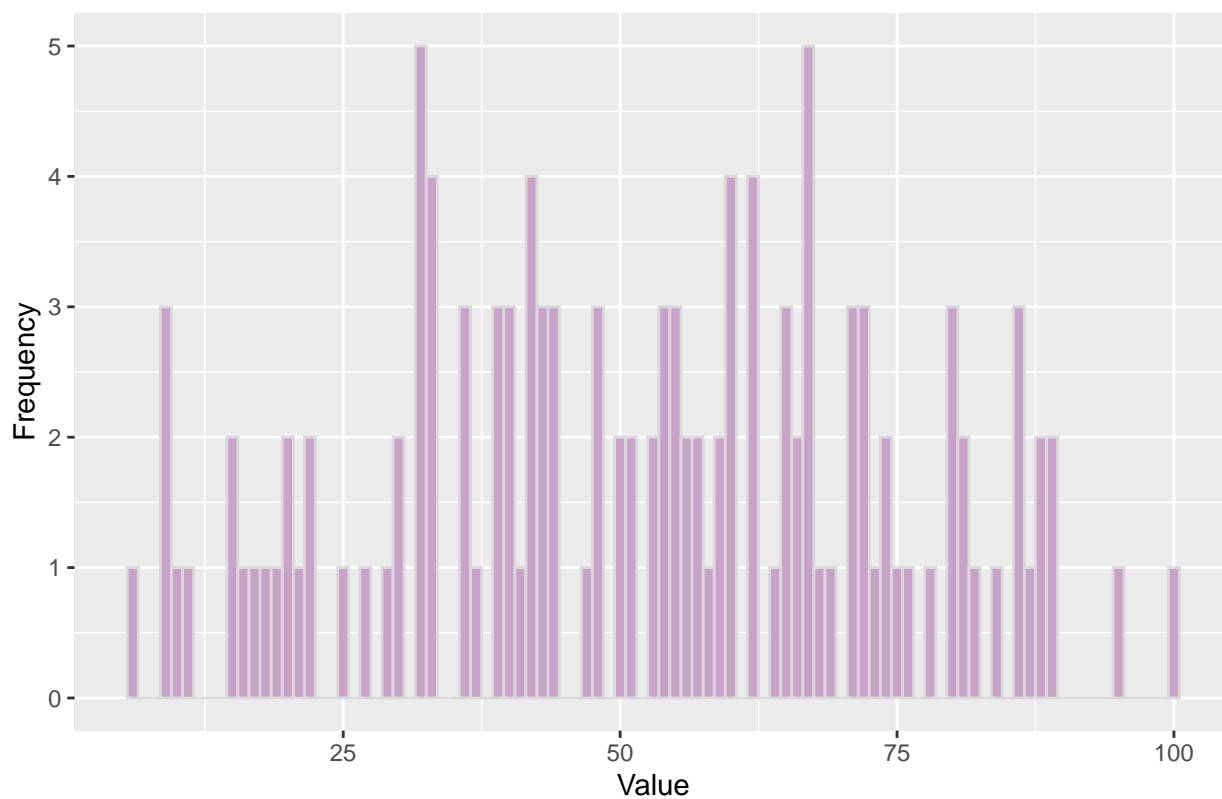


```
##  
## [[3]]  
## Warning in plot_create(genes, matrix, gene): NAs introduced by coercion
```



[[1]]

Histogram of Values of ABI1



```
##  
## [[2]]  
  
## Warning: Removed 2 rows containing missing values or values outside the scale range  
## (`geom_point()`).
```




```
##  
## [[3]]
```

