**ШИНЖЛЭХ УХААН ТЕХНОЛОГИЙН ИХ СУРГУУЛЬ**

**Мэдээлэл холбооны технологийн сургууль**

****

**ЛАБОРАТОРИЙН АЖЛЫН**

**ТАЙЛАН**

**Их Өгөгдлийн Шинжилгээ (**[F.CSC343-24/25A](https://elearn.sict.edu.mn/course/view.php?id=2368)**)**

**2023-2024 оны хичээлийн жил**

**намар**

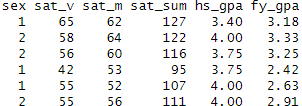
**Лабораторийн ажлын дугаар, нэр: Шалгалт**

**Хичээл заасан багш:** А. Алтангэрэл

**Лабораторийн ажил гүйцэтгэсэн:** Оюутан: С. Тэмүүжин /B221960002/

1. satgpa.csv файлыг уншиж data нэртэй өгөгдлийн фрейм үүсгэ.

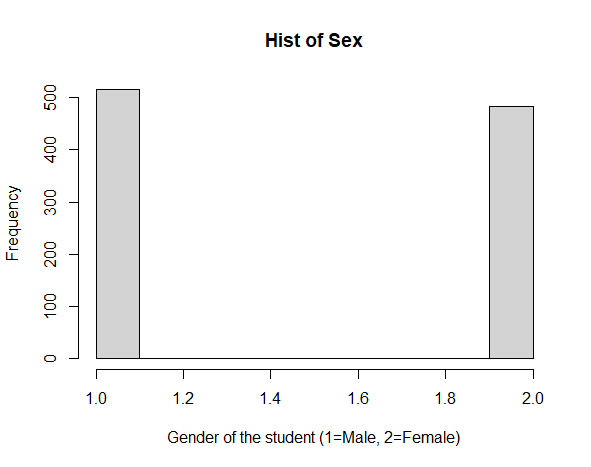
data <- read.csv("satgpa.csv")  
head(data)



1. Хувьсагч бүрээр гистограмм байгуулж үзүүл.

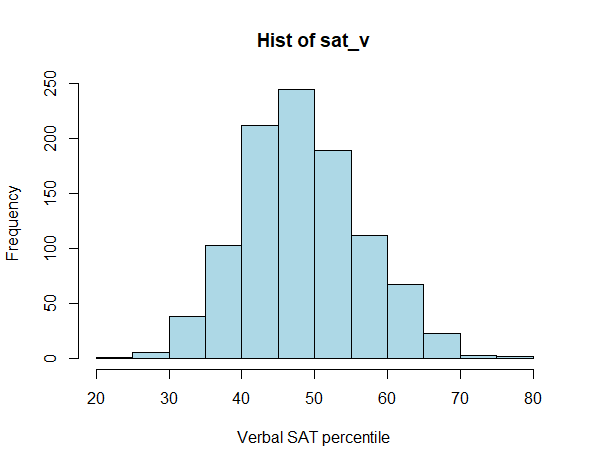
hist(data$sex, main = "Hist of Sex",

xlab = "Sex (1=Male, 2=Female)", col = "lightgray", border = "black")



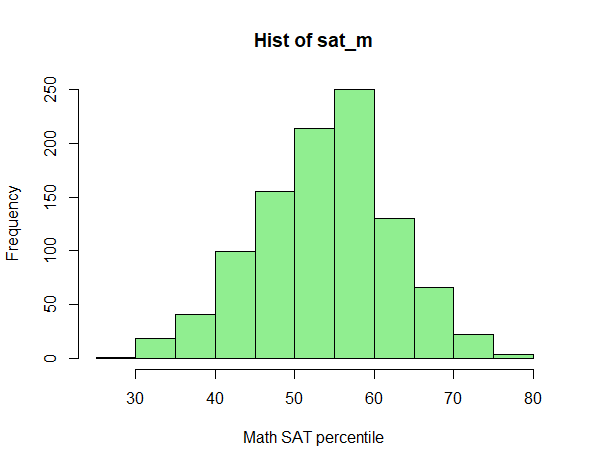
hist(data$sat\_v, main = "Hist of sat\_v",

xlab = "Percentile", col = "lightblue", border = "black")



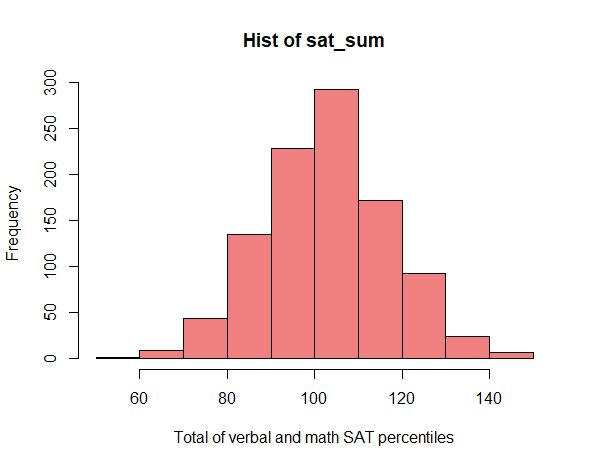
hist(data$sat\_m, main = "Hist of sat\_m",

xlab = "Percentile", col = "lightgreen", border = "black")



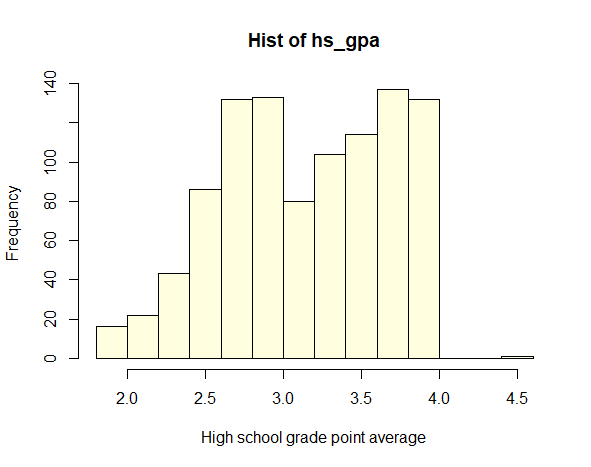
hist(data$sat\_sum, main = "Hist of sat\_sum",

xlab = "Total Percentile", col = "lightcoral", border = "black")



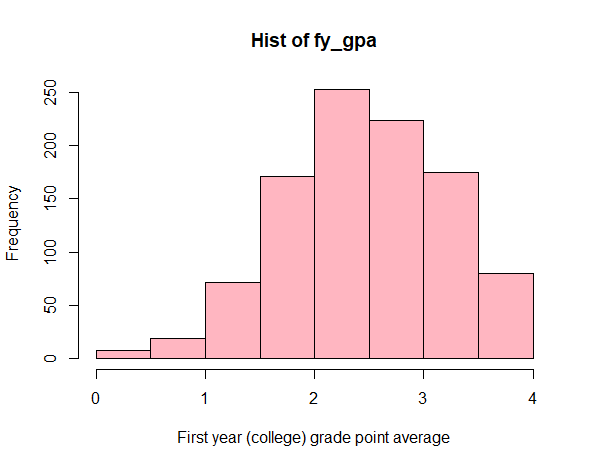
hist(data$hs\_gpa, main = "Hist of hs\_gpa",

xlab = "GPA", col = "lightyellow", border = "black")



hist(data$fy\_gpa, main = "Hist of fy\_gpa",

xlab = "FY GPA", col = "lightpink", border = "black")

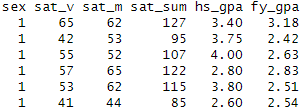


1. sex\_m, sex\_f нэртэй өгөгдлийн фрейм үүсгэ.

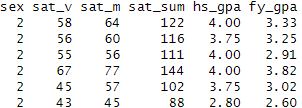
sex\_m <- subset(data, sex == 1)

sex\_f <- subset(data, sex == 2)

head(sex\_m)



head(sex\_f)



1. Фрейм бүрийн хувьд оюутнуудын тоог ол.

num\_male\_students <- nrow(sex\_m)

num\_female\_students <- nrow(sex\_f)

cat("Number of male students:", num\_male\_students, "\n")



cat("Number of female students:", num\_female\_students, "\n")



1. Фрейм бүрийн хувьд sat\_sum хувьсагчийн хувьд дундаж, дисперси, медианыг **ол.**

mean\_male\_sat\_sum <- mean(sex\_m$sat\_sum)

var\_male\_sat\_sum <- var(sex\_m$sat\_sum)

median\_male\_sat\_sum <- median(sex\_m$sat\_sum)

mean\_female\_sat\_sum <- mean(sex\_f$sat\_sum)

var\_female\_sat\_sum <- var(sex\_f$sat\_sum)

median\_female\_sat\_sum <- median(sex\_f$sat\_sum)

cat("Male students - SAT Sum:\n")

cat("Mean:", mean\_male\_sat\_sum, "\n")

cat("Variance:", var\_male\_sat\_sum, "\n")

cat("Median:", median\_male\_sat\_sum, "\n\n")









cat("Female students - SAT Sum:\n")

cat("Mean:", mean\_female\_sat\_sum, "\n")

cat("Variance:", var\_female\_sat\_sum, "\n")

cat("Median:", median\_female\_sat\_sum, "\n")









1. Эрэгтэй, эмэгтэй оюутнуудын **sat\_sum** хувьсагчийн хувьд дунджийг харгалзан μ1, μ2 гэж үзье. Өмнөх үр дүнг ашиглан μ1 -μ2 зөрүүний 90% -ийн итгэх интервалыг байгуул.

mean\_male <- mean(sex\_m$sat\_sum)

mean\_female <- mean(sex\_f$sat\_sum)

var\_male <- var(sex\_m$sat\_sum)

var\_female <- var(sex\_f$sat\_sum)

n\_male <- nrow(sex\_m)

n\_female <- nrow(sex\_f)

se\_diff <- sqrt((var\_male / n\_male) + (var\_female / n\_female))

df <- ( (var\_male / n\_male) + (var\_female / n\_female) )^2 /

( ((var\_male / n\_male)^2 / (n\_male - 1)) + ((var\_female / n\_female)^2 / (n\_female - 1)) )

alpha <- 0.10

t\_critical <- qt(1 - alpha / 2, df)

margin\_error <- t\_critical \* se\_diff

lower\_bound <- (mean\_male - mean\_female) - margin\_error

upper\_bound <- (mean\_male - mean\_female) + margin\_error

cat("90% Confidence Interval for the difference (μ1 - μ2):\n")

cat("Lower bound:", lower\_bound, "\n")

cat("Upper bound:", upper\_bound, "\n")





1. Эрэгтэйчүүдийн SAT дундаж оноо нь эмэгтэйчүүдийнхээс илүү гэсэн таамаглалыг 5% -ийн итгэх түвшинтэй шалга.

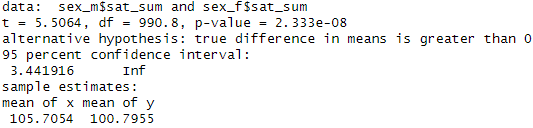
t\_test\_result <- t.test(sex\_m$sat\_sum, sex\_f$sat\_sum,

alternative = "greater",

var.equal = FALSE,

conf.level = 0.95)

print(t\_test\_result)



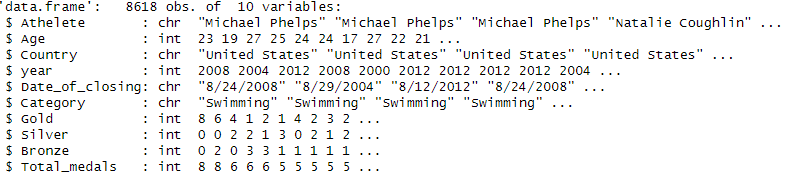
1. satgpa.csv өгөгдлийг ашиглан даалгаврыг гүйцэтгэ.

library(dplyr)

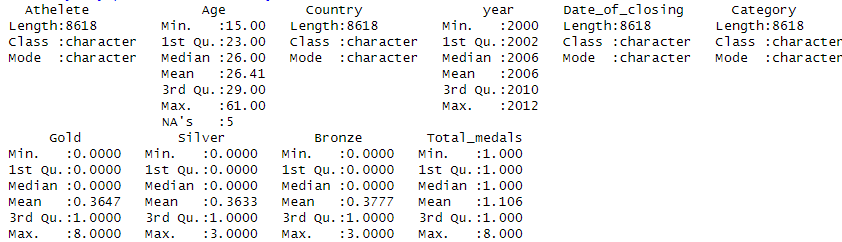
library(ggplot2)

olympic\_data <- read.csv("olympic\_data.csv")

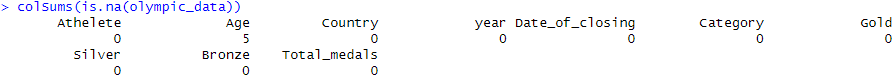
str(olympic\_data) # Structure of the dataset



summary(olympic\_data) # Summary statistics



colSums(is.na(olympic\_data))



olympic\_data\_clean <- olympic\_data %>%

filter(!is.na(Age))

medals\_by\_country <- olympic\_data\_clean %>%

group\_by(Country) %>%

summarise(Total\_Gold = sum(Gold, na.rm = TRUE),

Total\_Silver = sum(Silver, na.rm = TRUE),

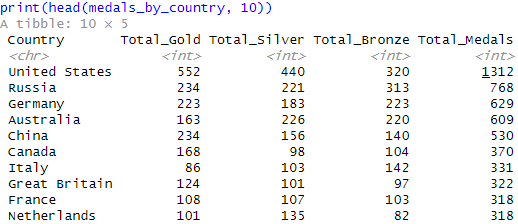
Total\_Bronze = sum(Bronze, na.rm = TRUE),

Total\_Medals = sum(Total\_medals, na.rm = TRUE),

.groups = 'drop') %>%

arrange(desc(Total\_Medals))

print(head(medals\_by\_country, 10))



ggplot(medals\_by\_country, aes(x = reorder(Country, Total\_Medals), y = Total\_Medals, fill = Country)) +

geom\_bar(stat = "identity") +

coord\_flip() +

labs(title = "Total Medals Won by Country", x = "Country", y = "Total Medals") +

theme\_minimal()

