

```

sink("C:/Users/khoongwh/Desktop/Wei Hao/NUS/Applied Mathematics Major/Modules/Year 4 Semester 2/ST2137 Computer Aided Data Analysis/Tutorials/Code Outputs Tutorial 2.txt")
#####
##### Tutorial 2 #####
#####

# Author: Khoong Wei Hao
# Matric: A0140425U
# Tutorial group: T03

### Question 1 ###
htwt2 <- read.fwf("C:/Users/khoongwh/Desktop/Wei Hao/NUS/Applied Mathematics Major/Modules/Year 4 Semester 2/ST2137 Computer Aided Data Analysis/Data/tut2htwtfixed.txt",
width=c(3,1,3,2,1))
names(htwt2) <- c("id", "gender", "height", "weight", "siblings")
htwt2

### Question 2 ###
htwt2m <- htwt2[htwt2$gender=="M",]
htwt2m
num_males <- nrow(htwt2m)
paste("Number of males is ", num_males, collapse = "")

### Question 3 ###
tut2test <- read.csv("C:/Users/khoongwh/Desktop/Wei Hao/NUS/Applied Mathematics Major/Modules/Year 4 Semester 2/ST2137 Computer Aided Data Analysis/Data/tut2test.csv", sep=",")
tut2test

# merge tut2test and htwt2 by id
htwttest2 <- merge(htwt2, tut2test, by="id")
htwttest2

# get those with heights > 182
ht_geq182 <- htwttest2[htwttest2$height > "182",]
paste("The individuals are 267, 271, 285 and their respective test scores are 55, 76 & 54.", collapse = "")

### Question 4 ###
htwttest2remo <- htwttest2[htwttest2$id != "210",]

### Question 5 ###
weight_column_num <- which( colnames(htwttest2)=="weight" )
htwttest2[htwttest2$id == "210", weight_column_num] <- 68

### Question 6 ###
# get only females
htwt2testf <- htwttest2[htwttest2$gender=="F",]

# sort data frame by height in ascending order
htwt2testf <- htwt2testf[order(htwt2testf$height),]
id_col_num <- which( colnames(htwt2testf)=="id" )
height_col_num <- which( colnames(htwt2testf)=="height" )
score_col_num <- which( colnames(htwt2testf)=="test" )

sec_tallest_id <- htwt2testf[nrow(htwt2testf)-1, id_col_num]
sec_tallest_height <- htwt2testf[nrow(htwt2testf)-1, height_col_num]
sec_tallest_weight <- htwt2testf[nrow(htwt2testf)-1, weight_column_num]
sec_tallest_score <- htwt2testf[nrow(htwt2testf)-1, score_col_num]

paste("The 2nd tallest female is of id ", sec_tallest_id, ", height ", sec_tallest_height, ", weight ",
      sec_tallest_weight, ", test score ", sec_tallest_score, ".", collapse = "")

### Question 7 ###
htwttest2$grade <- NULL
htwttest2$grade[htwttest2$test < 50] <- "F"
htwttest2$grade[htwttest2$test >= 50 & htwttest2$test < 60] <- "D"
htwttest2$grade[htwttest2$test >= 60 & htwttest2$test < 70] <- "C"
htwttest2$grade[htwttest2$test >= 70 & htwttest2$test < 80] <- "B"
htwttest2$grade[htwttest2$test >= 80] <- "A"
num_F <- length(htwttest2$grade[htwttest2$grade=="A"])

paste("The number of subjects with F grade is ", num_F, ".", collapse = "")

### Question 8 ###
# rep replicates elements of vectors. e.g. rep(1,5) gets a vector (1,1,1,1,1)
x <- cbind(rep(1,5),c(1,3,4,7,11))
y <- c(4,6,13,15,18)

# inverse of a matrix A is solve(A), transpose of A is t(A), matrix multiplication: %*%
betahat <- solve(t(x)%*%x) %*% t(x) %*% y
betahat

### Question 9 (i) ###
# numeric(n) creates a double-precision vector of the specified length with each element equal to 0
x <- numeric(3)
x[1] <- 0
x[2] <- 3

for(i in 3:28){
  x[i] <- 2*x[i-1] - x[i-2]
}

x
x[18]
paste("The 18th term of the series is", x[18], ".", collapse = "")

### Question 9 (ii) ###
sum_first_20 <- sum(x[1:20])
paste("The sum of the first 20 terms is", sum_first_20, ".", collapse = "")

### Question 10 ###
# The function cenmom finds the mean, the 2nd, 3rd & 4th central moments
cenmom <- function(x){
  n <- length(x)
  s <- numeric(4)
  # mean
  s[1] <- mean(x)
  # 2nd moment
  s[2] <- sum((x-s[1])^2)/n
  # 3rd moment
  s[3] <- sum((x-s[1])^3)/n
  # 4th moment
  s[4] <- sum((x-s[1])^4)/n
  return(s)
}

central_moments <- cenmom(htwt2$height)
paste("The four central moments are respectively",
      central_moments[1],",",
      central_moments[2],",",
      central_moments[3],",",
      central_moments[4],
      ".", collapse = "")

sink()

```