###Read data

d1 = read.csv("D:\\input1.csv", stringsAsFactors = FALSE)

str(d1)

d2 = read.csv("D:\\input2.csv", stringsAsFactors=FALSE)

#删除不必要的第一列

d1 = d1[, -1]

d2 = d2[, -1]

#设置正确的数据框列名

names(d1)

names(d1) = d1[1, ]

names(d2)

names(d2) = d2[1, ]

d1 = d1[-1, ]

d2 = d2[-1, ]

#Combine two data frame

c1 = rbind(d1, d2)

c2 = c1[!duplicated(c1), ] #去除重複

#将salary列设置为数值型

class(c2$salary)

c2$salary = as.numeric(c2$salary)

#将start\_date列设置为日期型

class(d1$start\_date) #Check start\_date

d1$start\_date = as.Date(d1$start\_date)

#Use loop to create the column

comment = list()

for(i in 1:length(c2$name))

{

p = paste(c2$name[i], "entered", c2$dept[i], "at", c2$start\_date[i], sep = " ")

print(p)

comment[i] = p

}

print(comment)

c2$comment = comment

#Flag

Flag = list()

for(i in 1:length(c2$name))

{

if(T %in% is.na(c2[i,]))

{

Flag[i] = T

}

else

{

Flag[i] = F

}

print(Flag[i])

}

c2$Flag = Flag

#CSV

write.csv(c2, "D:\\HW2.csv")