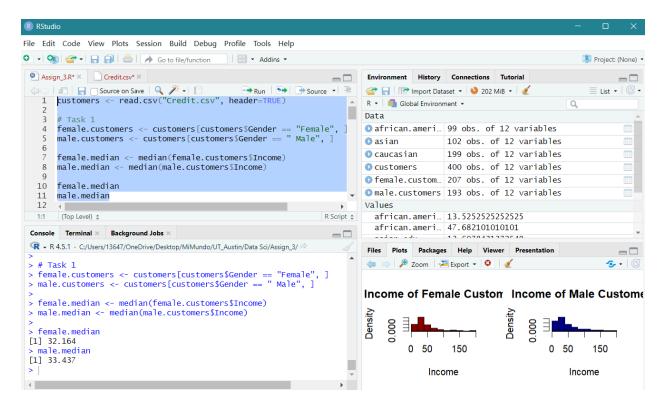
## **Assignment 3**

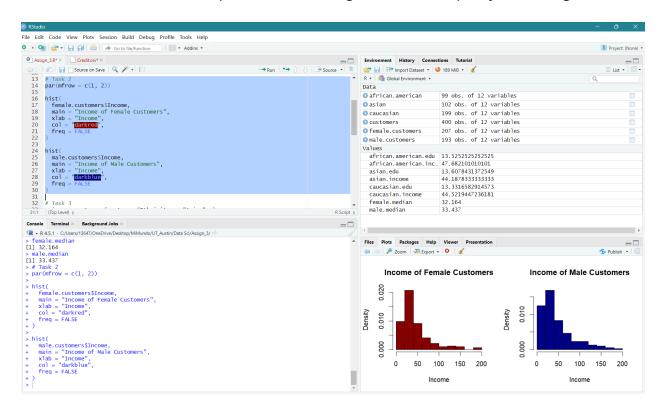
Rafael Ignacio Gonzalez Chong & Zhaowen Fan
UT Austin GSI

**Task 1:** Calculate the median income for male and female customers in the dataset. Is there a difference between the two groups in this sample dataset?



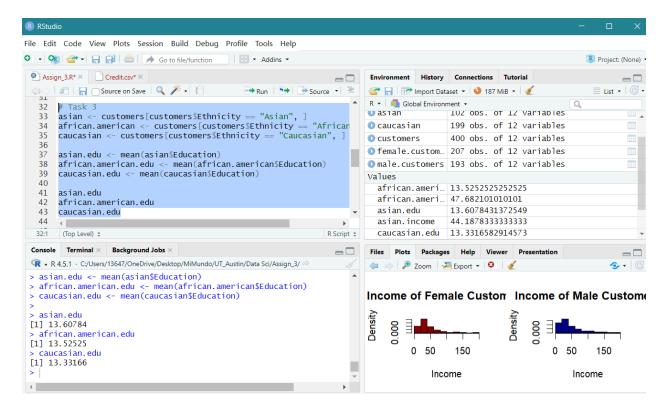
As we can see from the console, the median of male customers is 33.437, and the median of female customers is 32.164. Obviously, there is a difference between male customers and female customers. Male customers have a higher income median and the female customers have a lower income median.

**Task 2:** Create separate histograms showing the income distribution for male and female customers. Compare the two histograms and interpret your findings.



Above is the income histogram for the male and female customers respectively. As we can see from the histogram, female customers' income distribution are more centralized towards 25-50, and male customers' income distribution are more centralized towards 25-50 as well. But the lower income proportion of male are more significant as well, in other words, male customers' distribution are smoothier.

**Task 3:** Do different ethnic groups in the dataset have different average education levels?



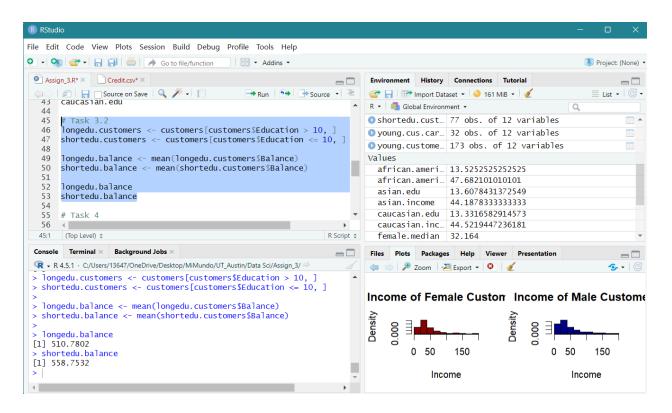
Yes, different ethnic groups have different education levels.

File Edit Code View Plots Session Build Debug Profile Tools Help ○ - On to file/function Project: (None) • Assign\_3.R\* × Credit.csv\* × Source on Save Q / > CauCas1an.edu R • Global Environment • 45 46 # Task 3.1 asian.income 44.18783333333333 "Yes", student.customers <- customers[customers\$Student == caucasian.edu 13.3316582914573 "No", ] adult.customers <- customers[customers\$Student == 44.5219447236181 caucasian.inc... 49 50 51 female.median 32.164 student.edu <- mean(student.customers\$Education)
adult.edu <- mean(adult.customers\$Education)</pre> longedu.balan... 510.780185758514 33.437 male.median 52 53 student.edu 2.97282608695652 old.cards adult.edu old.income 46.9986740088106 shortedu.bala... 558.753246753247 # Task 3.2 55 student.edu 14.125 56 sudent.edu 45:1 (Top Level) \$ R Script \$ NA real Background Jobs Plots Packages Help Viewer Presentation R 4.5.1 · C:/Users/13647/OneDrive/Desktop/MiMundo/UT\_Austin/Data Sci/Assign\_3/ 🗼 🔑 Zoom 🛮 😕 Export 🕶 🚨 🛮 🎻 > student.customers <- customers[customers\$Student == "Yes" > adult.customers <- customers[customers\$Student == "No", ]</pre> Income of Female Custom Income of Male Custome > student.edu <- mean(student.customers\$Education)
> adult.edu <- mean(adult.customers\$Education)</pre> [1] 14.125 > adult.edu 150 150 [1] 13.375 Income Income

Task 3.1: Do Students have a higher number of Education years?

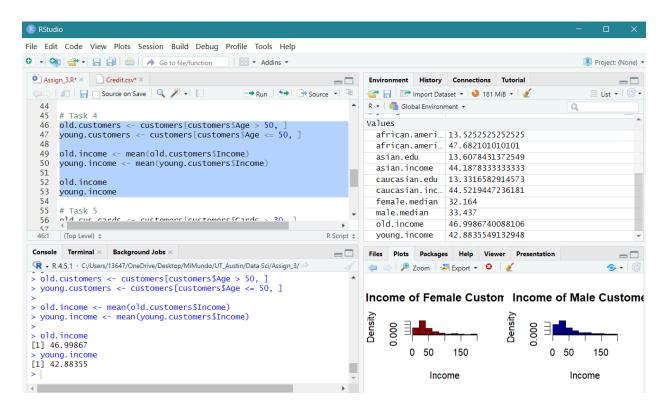
According to the output, students' average education years are 14.125, and adults' education years are 13.375. Obviously, students have a higher number of education years.

**Task 3.2:** Do People with more than 10 years of education have a higher balance on their account?



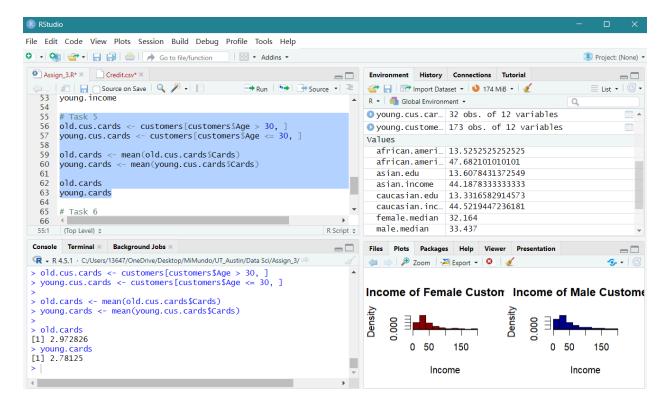
According to the output, people with more than 10 years of education do not have a higher balance on their account.

**Task 4:** Do people older than 50 years have a higher average income compared to the rest of the dataset?



According to the output in the console, I would like to say, YES!

**Task 5:** Do people younger than 30 years old have more credit cards on average?



According to the output in the console, people younger than 30 years old have less credit cards on average.

> caucasian.income [1] 44.52194

RStudio File Edit Code View Plots Session Build Debug Profile Tools Help O → Go to file/function □ - Addins -Rroject: (None) Environment History Run Source • ☐ Import Dataset ▼ 
○ 174 MiB ▼ young.caras <- mean(young.cus.caras) R • Global Environment • old.cards young.cus.car... 32 obs. of 12 variables 62 63 young.cards young.custome... 173 obs. of 12 variables Values 65 african.ameri... 13.5252525252525 asian.income <- mean(asian\$Income) 66 african.american.income <- mean(african.american\$Income)</pre> african.ameri... 47.682101010101 68 caucasian.income <- mean(caucasian\$Income)</pre> 13.6078431372549 asian.edu 69 44.18783333333333 asian.income 70 asian.income caucasian.edu 13.3316582914573 african.american.income caucasian.inc... 44.5219447236181 caucasian.income female.median 32.164 64:1 (Top Level) \$ R Script ¢ Background Jobs Plots Packages Help R 4.5.1 · C:/Users/13647/OneDrive/Desktop/MiMundo/UT\_Austin/Data Sci/Assign\_3/ \Rightarrow 🔑 Zoom 🛮 🛂 Export 🕶 🚨 🧹 > asian.income <- mean(asian\$Income)
> african.american.income <- mean(african.american\$Income)</pre> > caucasian.income <- mean(caucasian\$Income)</p> Income of Female Custom Income of Male Custome > asian.income Density > african.american.income [1] 47.6821

0 50

150

Income

0 50

150

Income

**Task 6:** Do different ethnic groups have different average incomes?

According to the output, different ethnic groups do have different average incomes. For Asians, 44.18783; For African Americans, 47.6821; For Caucasians, 44.52194.