#####task 1##### ##Which variabeles are there?## ##how many observations are there?## View(earnings) count(earnings) library(tidyverse) library(ggplot2) library(rstanarm) library(devtools) library(rmarkdown) library(rosdata) data(earnings)

View(earnings) count(earnings)

####task 2#### ##make a density plot or a histogram of the variable earn## hist(earningsearn) plot(earningdens)

plotty <- ggplot(earnings, aes(x= earn))+ geom\_density() plotty + geom\_vline(aes(xintercept = mean(earn)), color = “blue”, linetype =“dashed”, size =1) ##evaluate if it’s nesscary to transform the variable and make a new plot## earningsearn) hist(earnings$earn\_log2) plotty <- ggplot(earnings, aes(x =earn\_log2))+ geom\_density() ###task 3### make a new plot with gender as a factor and show the median. ### what does this plot tell us? earnings$male <- factor(c(“male”, “female”)) levels(earningsmale) nlevels(earnings$male)

ggplot(earnings, aes(x=earn, color=male, fill=male)) + geom\_density(alpha=0.3)

median <- earnings %>% group\_by(male) %>% summarize(median=median(earn))

ggplot(earnings, aes(x=earn, color=male, fill=male)) + geom\_density(alpha=0.3)+ geom\_vline(data = median, aes(xintercept = median, color = male), size=0.5) ################this density plot tells us that men have a higher median income than women







