>

> correlation <- cor(selected\_data$removed, selected\_data$S1901\_C01\_012E)

> print(correlation)

[1] -0.09966586

>

> overall\_mean <- mean(selected\_data$S1901\_C01\_012E)

> overall\_std\_dev <- sd(selected\_data$S1901\_C01\_012E)

>

> print(paste("Overall Mean of Medians:", overall\_mean))

[1] "Overall Mean of Medians: 109555.847723705"

> print(paste("Overall Standard Deviation of Medians:", overall\_std\_dev))

[1] "Overall Standard Deviation of Medians: 47320.4650304546"

>

>

> # Calculate the mean and standard deviation for each group

> grouped\_data <- selected\_data %>%

+ group\_by(removed) %>%

+ summarize(mean\_income = mean(S1901\_C01\_012E),

+ sd\_income = sd(S1901\_C01\_012E),

+ n = n())

>

> # Calculate the standard error

> grouped\_data$se\_income <- grouped\_data$sd\_income / sqrt(grouped\_data$n)

>

> # Print the results

> print(grouped\_data$se\_income)

[1] 1960.819 6110.347

>

> # Calculate the mean, standard deviation, sample size, and standard error

> grouped\_data <- selected\_data %>%

+ group\_by(removed) %>%

+ summarize(mean\_income = mean(S1901\_C01\_012E),

+ sd\_income = sd(S1901\_C01\_012E),

+ n = n(),

+ se\_income = sd\_income / sqrt(n))

>

> # Calculate the confidence intervals (assuming a 95% confidence level)

> grouped\_data$lower\_ci <- grouped\_data$mean\_income - 1.96 \* grouped\_data$se\_income

> grouped\_data$upper\_ci <- grouped\_data$mean\_income + 1.96 \* grouped\_data$se\_income

>

> # Print the results

> print(grouped\_data)

# A tibble: 2 × 7

removed mean\_income sd\_income n se\_income lower\_ci upper\_ci

<int> <dbl> <dbl> <int> <dbl> <dbl> <dbl>

1 0 111075. 47100. 577 1961. 107232. 114919.

2 1 94942. 47331. 60 6110. 82966. 106918.

>