#1

```
# Read the dataset and store it in an object with your initials
TS_dat <- read.csv("https://ahj.hi.is/spurningar_um_lifid_24.csv", sep = ";", dec =
",")
```

#2

```
# Taking a random sample of size 200
TS_sample <- sample_n(TS_dat, 200)
```

#3

```
mean_commute_time <- mean(TS_sample$commute_time, na.rm = TRUE)
```

```
## Warning in mean.default(TS_sample$commute_time, na.rm = TRUE): argument is not
## numeric or logical: returning NA
```

```
se_commute_time <- sd(TS_sample$commute_time, na.rm = TRUE) / sqrt(nrow(TS_sample))

alpha <- 0.05
z_score <- qnorm(1 - alpha / 2)
lower_bound <- mean_commute_time - z_score * se_commute_time
upper_bound <- mean_commute_time + z_score * se_commute_time

cat(sprintf("95%% Confidence Interval for the mean commute time: %.2f < μ < %.2f",
lower_bound, upper_bound))
```

```
## 95% Confidence Interval for the mean commute time: NA < μ < NA
```

#4

```
breakfast_yes <- TS_sample$commute_time[TS_sample$breakfast == 1]
breakfast_no <- TS_sample$commute_time[TS_sample$breakfast == 0]

if (length(breakfast_yes) > 0 & length(breakfast_no) > 0) {
    t_test_result <- t.test(breakfast_yes, breakfast_no)

    cat("The mean commute time of students who eat breakfast is different from thos
e who don't.\n")
    cat(sprintf("Test Statistic: %.3f\n", t_test_result$statistic))
    cat(sprintf("P-value: %.3f\n", t_test_result$p.value))
    if(t_test_result$p.value < alpha) {
      cat("Reject the null hypothesis, there is a significant difference in commute
times.\n")
    } else {
      cat("Fail to reject the null hypothesis, there is no significant difference i
n commute times.\n")
    }
} else {
    cat("Insufficient data for one or both groups regarding breakfast consumption.\
n")
}
```

```
## Insufficient data for one or both groups regarding breakfast consumption.
```

#5

```
median_activity <- median(TS_sample$physical_activity, na.rm = TRUE)
high_activity <- TS_sample$sleep_quality[TS_sample$physical_activity > median_activ
ity]
low_activity <- TS_sample$sleep_quality[TS_sample$physical_activity <= median_activ
ity]

# Ensure data exists for both groups before proceeding with t-test
if (length(high_activity) > 0 & length(low_activity) > 0) {
    t_test_activity <- t.test(high_activity, low_activity)

    cat("Sleep quality differs between students with high and low physical activity
levels.\n")
    cat(sprintf("Test Statistic: %.3f\n", t_test_activity$statistic))
    cat(sprintf("P-value: %.3f\n", t_test_activity$p.value))
    if(t_test_activity$p.value < alpha) {
      cat("Reject the null hypothesis, physical activity level affects sleep qualit
y.\n")
    } else {
      cat("Fail to reject the null hypothesis, no significant effect of physical ac
tivity level on sleep quality.\n")
    }
} else {
    cat("Insufficient data for one or both groups regarding physical activity level
s.\n")
}
```

```
## Insufficient data for one or both groups regarding physical activity levels.
```