```r
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
data <- readxl::read_excel("/Users/ankushidutta/Downloads/Group5 (1).xlsx")
```

CSSAS vs graduate

```r
data %>%
  group_by(graduate) %>%
  summarise(
    count = n(),
    mean_CSSAS = mean(CSSAS, na.rm = TRUE),
    sd_CSSAS = sd(CSSAS, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   graduate count mean_CSSAS sd_CSSAS
##      <dbl> <int>      <dbl>    <dbl>
## 1        0  4502       19.7     4.09
## 2        1   960       24.1     3.79
```

```r
t.test(CSSAS ~ graduate, data = data, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  CSSAS by graduate
## t = -32.449, df = 1474.8, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -4.705335 -4.168878
## sample estimates:
## mean in group 0 mean in group 1
##        19.67748        24.11458
```

CSSAS vs mplan

```r
data %>%
  group_by(mplan) %>%
  summarise(
    count = n(),
    mean_CSSAS = mean(CSSAS, na.rm = TRUE),
    sd_CSSAS = sd(CSSAS, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   mplan count mean_CSSAS sd_CSSAS
##   <dbl> <int>      <dbl>    <dbl>
## 1     0  2766       21.2     4.73
## 2     1  2696       19.7     3.83
```

```r
t.test(CSSAS ~ mplan, data = data, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  CSSAS by mplan
## t = 13.504, df = 5282.5, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  1.342404 1.798343
## sample estimates:
## mean in group 0 mean in group 1
##        21.23247        19.66209
```

T-Test shows a statistically significant difference in mean CSSAS scores, with non-meal-plan students reporting higher stress ($p < 2.2e\text{-}16$). 95% confidence interval for the difference in means was [1.34, 1.80], indicating a reliable difference between the groups.

CSSAS vs. loans

```r
data %>%
  group_by(loans) %>%
  summarise(
    count = n(),
    mean_CSSAS = mean(CSSAS, na.rm = TRUE),
    sd_CSSAS = sd(CSSAS, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   loans count mean_CSSAS sd_CSSAS
##   <dbl> <int>      <dbl>    <dbl>
## 1     0  2266       20.4     4.40
## 2     1  3196       20.5     4.36
```

```r
t.test(CSSAS ~ loans, data = data, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  CSSAS by loans
## t = -0.55937, df = 4849.6, p-value = 0.5759
## alternative hypothesis: true difference in means between group 0 and group 1 is not equal to 0
## 95 percent confidence interval:
##  -0.3035172  0.1687630
## sample estimates:
## mean in group 0 mean in group 1
##        20.41792        20.48529
```

CSSAS vs. grade (Low vs High GPA)

```r
grade_data <- data %>%
  filter(grade %in% c(1, 4)) %>%
  mutate(gpa_group = ifelse(grade == 1, "low", "high"))

grade_data %>%
  group_by(gpa_group) %>%
  summarise(
    count = n(),
    mean_CSSAS = mean(CSSAS, na.rm = TRUE),
    sd_CSSAS = sd(CSSAS, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   gpa_group count mean_CSSAS sd_CSSAS
##   <chr>     <int>      <dbl>    <dbl>
## 1 high        471       20.5     4.51
## 2 low        1440       20.2     4.19
```

```r
t.test(CSSAS ~ gpa_group, data = grade_data, var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  CSSAS by gpa_group
## t = 1.0215, df = 752.82, p-value = 0.3073
## alternative hypothesis: true difference in means between group high and group low is not equal to 0
## 95 percent confidence interval:
##  -0.2215733  0.7023058
## sample estimates:
## mean in group high  mean in group low
##           20.48620           20.24583
```

FI vs graduate

```
data %>%
  group_by(graduate) %>%
  summarise(
    count = n(),
    FI_mean = mean(FI, na.rm = TRUE),
    FI_sd = sd(FI, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   graduate count FI_mean FI_sd
##      <dbl> <int>   <dbl> <dbl>
## 1        0  4502   0.269 0.443
## 2        1   960   0.585 0.493
```

```
table(data$graduate, data$FI) %>% chisq.test()
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  .
## X-squared = 360.54, df = 1, p-value < 2.2e-16
```

FI vs mplan

```
data %>%
  group_by(mplan) %>%
  summarise(
    count = n(),
    FI_mean = mean(FI, na.rm = TRUE),
    FI_sd = sd(FI, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   mplan count FI_mean FI_sd
##   <dbl> <int>   <dbl> <dbl>
## 1     0  2766   0.409 0.492
## 2     1  2696   0.237 0.426
```

```
table(data$mplan, data$FI) %>% chisq.test()
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  .
## X-squared = 183.22, df = 1, p-value < 2.2e-16
```

FI vs loans

```
data %>%
  group_by(loans) %>%
  summarise(
    count = n(),
    FI_mean = mean(FI, na.rm = TRUE),
    FI_sd = sd(FI, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   loans count FI_mean FI_sd
##   <dbl> <int>   <dbl> <dbl>
## 1     0  2266   0.328 0.470
## 2     1  3196   0.322 0.467
```

```
table(data$loans, data$FI) %>% chisq.test()
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  .
## X-squared = 0.18624, df = 1, p-value = 0.6661
```

FI vs grade (Low vs High GPA)

```
grade_FI <- data %>%
  filter(grade %in% c(1, 4)) %>%
  mutate(gpa_group = ifelse(grade == 1, "low", "high"))

grade_FI %>%
  group_by(gpa_group) %>%
  summarise(
    count = n(),
    FI_mean = mean(FI, na.rm = TRUE),
    FI_sd = sd(FI, na.rm = TRUE)
  )
```

```
## # A tibble: 2 x 4
##   gpa_group count FI_mean FI_sd
##   <chr>     <int>   <dbl> <dbl>
## 1 high        471   0.312 0.464
## 2 low        1440   0.312 0.464
```

```
table(grade_FI$gpa_group, grade_FI$FI) %>% chisq.test()
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  .
## X-squared = 3.2633e-30, df = 1, p-value = 1
```