FINAL PROJECT

Dawa Oluya

2024-02-19

The dataset pertains to the learning attitudes and perceptions of 69 second-year dental students in Korea. These students participated in a clinical periodontology course delivered via YouTube in a flipped classroom setting. The course, held from September 2 to December 16, 2016, comprised 1-hour weekly sessions over a total of 14 hours. After completing the course and final examination, students responded to a survey via Google Forms on their smartphones in January 2017.The analysis aims to provide insights into the effectiveness and acceptability of online learning methods in dental education.

```
# Set the working directory to the directory containing the R Markdown file
here::i_am("code/project_code.Rmd")
```

```
## here() starts at C:/Users/dawal/OneDrive/Documents/DATA_550/final_project
```

```
absolute_path_to_data <- here::here("raw_data", "raw.xlsx")
# Read the Excel file into R as a data frame
data <- readxl::read_excel(absolute_path_to_data)
```

Cleaning dataset variable names to make it easier to use and understand

```
#cleaning names using janitor package
data <- janitor :: clean_names(data)

# Rename variables to more concise names
new_names <- c(
  "timestamp",  "gender",  "skill_proficiency", "device_used","location","watch_time",
  "companions",  "watch_before_class",  "playback_speed", "skip_button_used",
  "pause_button_used","rewatch_video","video_ease_comparison", "future_video_watch",
  "subscribe_to_channel",  "share_video","watch_recommended_videos", "youtube_advantages"
)

# Assign the new names to the dataset
names(data) <- new_names
```

```
library(knitr)
library(gtsummary)
```

```
## #Uighur
```

```r
# Create summary table
watched_video <- data %>%
  select("gender", "location", "watch_before_class") %>%
  tbl_summary(by = watch_before_class) %>%
  modify_spanning_header(c("stat_1", "stat_2") ~ "**Watch The Video Before Class**") %>%
  add_overall() %>%
  add_p()

watched_video
```

```
## Table printed with `knitr::kable()`, not {gt}. Learn why at
## https://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html
## To suppress this message, include `message = FALSE` in code chunk header.
```

Characteristic	**Overall**, N = 72	**Always**, N = 56	**Generally**, N = 11	**Occasionally**, N = 1	**Often**, N = 4	**p-value**
gender						0.4
F	17 (24%)	16 (29%)	1 (9.1%)	0 (0%)	0 (0%)	
M	55 (76%)	40 (71%)	10 (91%)	1 (100%)	4 (100%)	
location						0.5
Campus	3 (4.2%)	1 (1.8%)	1 (9.1%)	0 (0%)	1 (25%)	
Home	52 (72%)	40 (71%)	8 (73%)	1 (100%)	3 (75%)	
Home, cafe	2 (2.8%)	1 (1.8%)	1 (9.1%)	0 (0%)	0 (0%)	
Home, Campus	12 (17%)	11 (20%)	1 (9.1%)	0 (0%)	0 (0%)	
Home, Campus, cafe	1 (1.4%)	1 (1.8%)	0 (0%)	0 (0%)	0 (0%)	
Home, Campus, On the move	1 (1.4%)	1 (1.8%)	0 (0%)	0 (0%)	0 (0%)	
On the move	1 (1.4%)	1 (1.8%)	0 (0%)	0 (0%)	0 (0%)	

```r
# Save summary table as RDS file
saveRDS(
  watched_video,
  file = here::here("output/watched_video.rds")
)
```

The table provides a summary on variables gender, where they watched the videos and whether the video was watched before class. It displays descriptive statistics and p-values, organized by the "Watch Before Class" variable.

```r
# Load necessary libraries
library(dplyr)
```

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

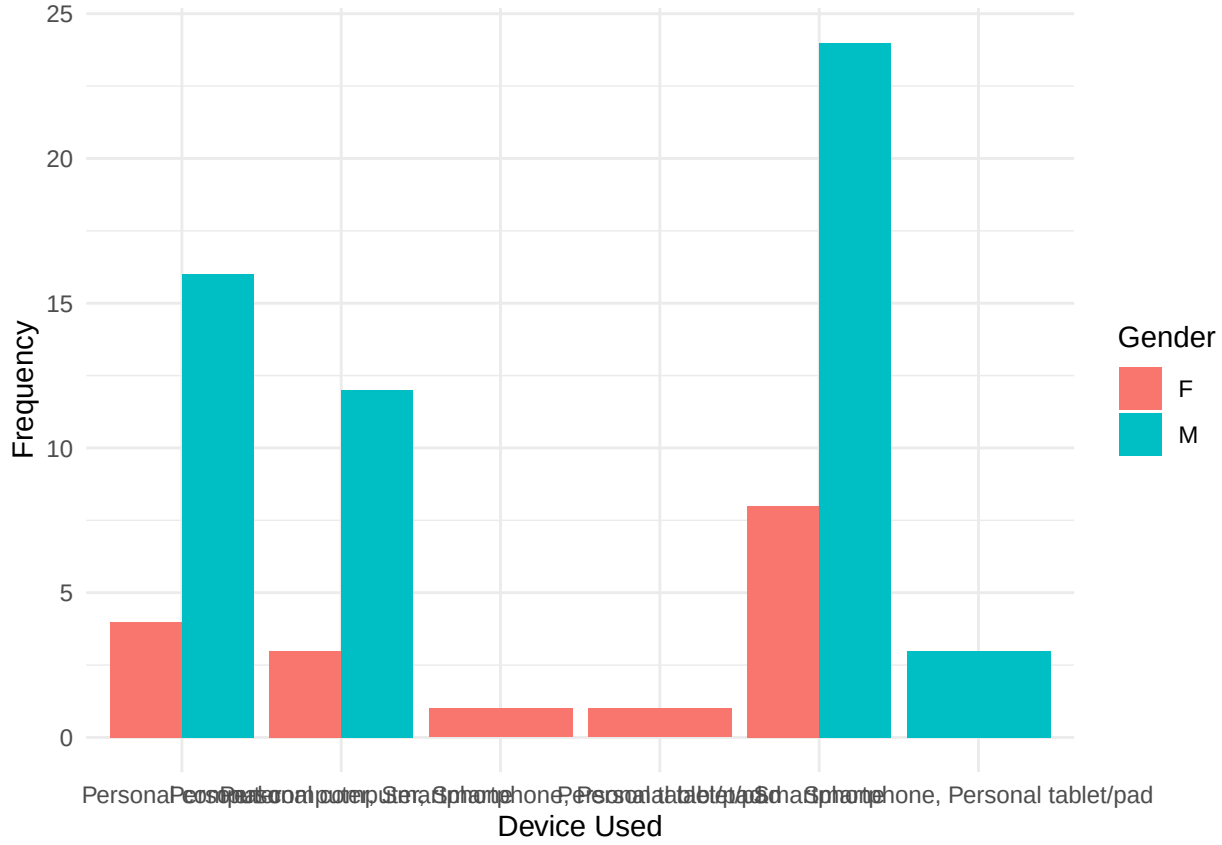
```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

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

```r
library(ggplot2)

# Count the frequency of each device used by gender
device_frequency <- data %>%
  count(gender, device_used)

# Create a bar plot using ggplot2
barplot <- ggplot(device_frequency, aes(x = device_used, y = n, fill = gender)) +
  geom_bar(stat = "identity", position = "dodge") +
  labs(x = "Device Used", y = "Frequency", fill = "Gender") +
  theme_minimal()
 print(barplot)
```




```r
# Save the plot as a PNG file
ggsave(
 here::here("output/device_frequency.png"),
  plot = barplot,
  device = "png"
)
```

```
## Saving 6.5 x 4.5 in image
```

The chart displays the frequency of devices used by gender. The x-axis shows the devices, while the y-axis represents the frequency. Each device has two bars side by side, one for each gender, with colors indicating gender. This allows for easy comparison of device usage patterns across genders.