tidyingAgeAndCourse

Lore Faith Sabio

2024-05-25

```
library(readxl)
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
```

```
library(tidyr)
library(knitr)
library(openxlsx)
library(ggplot2)
library(plotly)
```

```
##
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
##
##     last_plot
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
## The following object is masked from 'package:graphics':
##
##     layout
```

main files

```
orig <- read_excel("/cloud/project/DataCleaning(3G)/questionnaire3g.xlsx")
deets <- read_excel("/cloud/project/DataCleaning(3G)/questionnaire3g.xlsx")
```

AGE

```r
#getting the age average

age <- c(deets$Age)
```

```r
#converting age to numeric form

age <- as.numeric(gsub("[^0-9]", "", as.character(age)))
```

```r
#removes zeros

age <- as.numeric(gsub("0$", "", as.character(age)))
```

```r
#neglecting NAs in solving average
ave_age <- mean(age, na.rm = TRUE)

ave_age
```

```
## [1] 20.70103
```

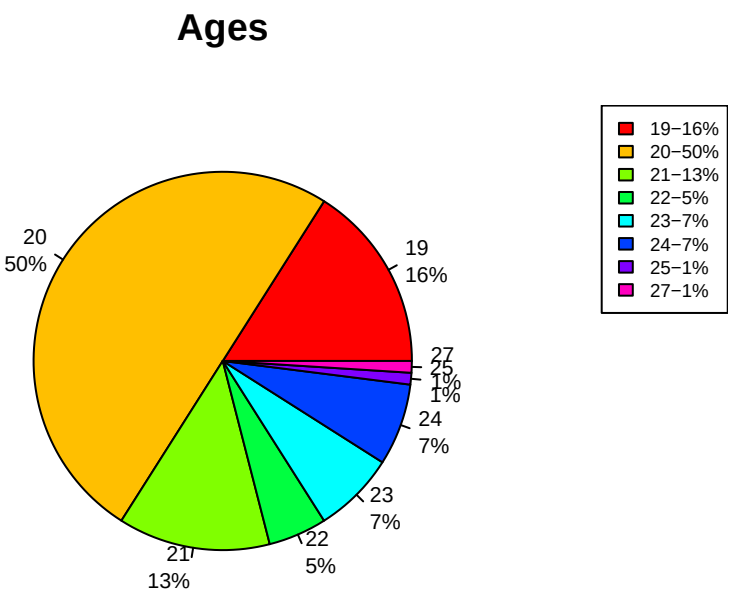
```r
# removing decimals on ave_age
(ave_age <- floor(ave_age))
```

```
## [1] 20
```

```r
#fills in the NAs with the ave_age
age[is.na(age)] <- ave_age
```

```r
# syncing to data
deets$Age <- age
```

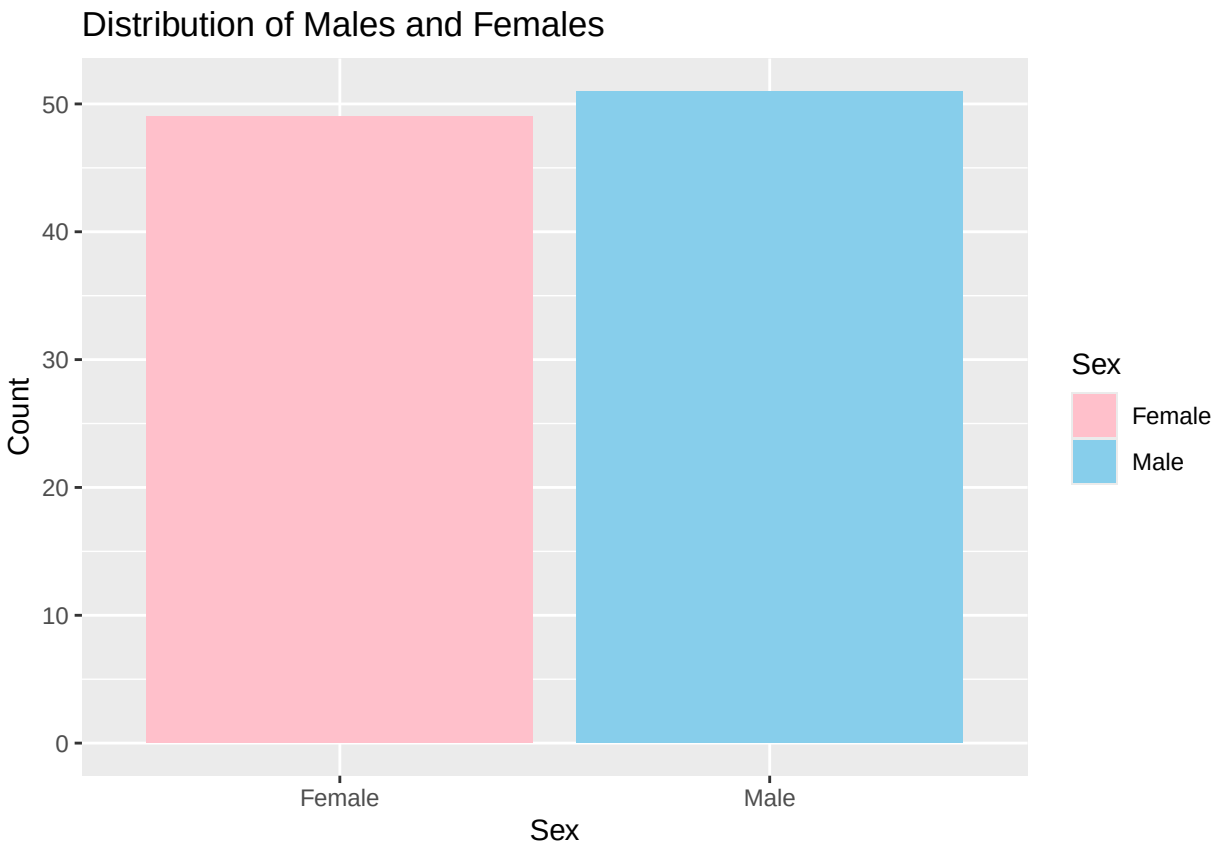
```r
# graphs for age ( PIE CHART )
age_counts <- table(age)
percentages <- round(prop.table(age_counts) * 100, 1)
pie_age <- pie(age_counts, main = "Ages", labels = paste(names(age_counts), "\n", percentages, "%", sep
legend("topright", legend = paste(names(age_counts), "-", percentages, "%", sep = ""), cex = 0.6, fill =
```



The predominance of 19-20s in Adobe application usage, with half of respondents falling within their 20s. This demographic trend highlights the importance of targeting and catering to the needs and preferences of younger users in the creative software industry.

SEX

```
males <- subset(deets, Sex == "Male")
females <- subset(deets, Sex == "Female")

sex_counts <- data.frame(
  Sex = c("Male", "Female"),
  Count = c(nrow(males), nrow(females))
)

ggplot(sex_counts, aes(x = Sex, y = Count, fill = Sex)) +
  geom_bar(stat = "identity") +
  labs(title = "Distribution of Males and Females",
       x = "Sex",
       y = "Count") +
  scale_fill_manual(values = c("Male" = "skyblue", "Female" = "pink"))
```



The data indicates a relatively balanced usage of Adobe applications between genders, suggesting widespread adoption across both male and female demographics. Regardless of gender the balanced usage between genders highlights the utility of Adobe software.

COURSES

```
# show courses
print(deets$Course[25:60])
```

```
##  [1] "BS in Accountancy"
```

```
##  [2] "BSN"
##  [3] "BSIT"
##  [4] "BS in Accountancy"
##  [5] "BS in HM"
##  [6] "BSIT"
##  [7] "BSIT"
##  [8] "BSIT"
##  [9] "BSHM"
## [10] "BSIT"
## [11] "BSIT"
## [12] "BSIT"
## [13] "BSIT"
## [14] "BSIT"
## [15] "BTVTED BCW"
## [16] "BS in Agriculture"
## [17] "BSCS"
## [18] "BSMar-E"
## [19] "BSIT"
## [20] "BS PHARMACY"
## [21] "BSIT"
## [22] "Bachelor of Science in Information Technology"
## [23] "Avionics Technology"
## [24] "BSIT"
## [25] "BSIT"
## [26] "BSIT"
## [27] "BSIT"
## [28] "BSIT"
## [29] "BSIT"
## [30] "BSIT"
## [31] "BIT-MT 2A"
## [32] "Bachelor of Science in Information Technology (BSIT)"
## [33] "BSIT"
## [34] "BSIT"
## [35] "BSIT-2B"
## [36] "BSIT"
```

We can see here that there are messy inputs we receive from the respondents. Therefore, we will clean all that by replacing it to its acronyms.

```
# replacing long terms to acronyms

# to BSIT
to_bsit <- c(1, 14, 46, 56, 59, 61, 68)
deets$Course[to_bsit] <- "BSIT"

# to BSCE
to_bsce <- c(2)
deets$Course[to_bsce] <- "BSCE"

#to BS Agri
to_bsagri <- c(17, 40)
deets$Course[to_bsagri] <- "BS Agri"

# to BS Chem
to_bschem <- c(16, 80)
```

```r
deets$Course[to_bschem] <- "BS Chem"

# to bsa
to_bsa <- c(25, 28)
deets$Course[to_bsa] <- "BSA"

# to bshm
to_bshm <- c(31, 98)
deets$Course[to_bshm] <- "BSHM"

# to bs phar
to_bsphar <- c(44)
deets$Course[to_bsphar] <- "BS Phar"
```

```r
# checks the changes
deets$Course
```

```
##   [1] "BSIT"               "BSCE"                "BSIT"
##   [4] "BSIT"               "BSIT"                "BSIT"
##   [7] "BSIT"               "BSIT"                "BSHM"
##  [10] "BSIT"               "BS Crim"             "BSIT"
##  [13] "BSIT"               "BSIT"                "BSMT"
##  [16] "BS Chem"            "BS Agri"             "BSIT"
##  [19] "BSIT"               "BSIT"                "BSIT"
##  [22] "BSIT"               "BSIT"                "BSIT"
##  [25] "BSA"                "BSN"                 "BSIT"
##  [28] "BSA"                "BS in HM"            "BSIT"
##  [31] "BSHM"               "BSIT"                "BSHM"
##  [34] "BSIT"               "BSIT"                "BSIT"
##  [37] "BSIT"               "BSIT"                "BTVTED BCW"
##  [40] "BS Agri"            "BSCS"                "BSMar-E"
##  [43] "BSIT"               "BS Phar"             "BSIT"
##  [46] "BSIT"               "Avionics Technology" "BSIT"
##  [49] "BSIT"               "BSIT"                "BSIT"
##  [52] "BSIT"               "BSIT"                "BSIT"
##  [55] "BIT-MT 2A"          "BSIT"                "BSIT"
##  [58] "BSIT"               "BSIT"                "BSIT"
##  [61] "BSIT"               "BSHM"                "BSIT"
##  [64] "BSIT"               "BSIT"                "BSIT"
##  [67] "BSIT"               "BSIT"                "BSED"
##  [70] "BSBA MM"            "BSIT"                "BSN"
##  [73] "BSHM"               "BSIT"                "BSED"
##  [76] "BSIT"               "BTVTED-CP"           "BSN"
##  [79] "BSIT"               "BS Chem"             "BSIT"
##  [82] "BSBA-FM"            "BSIT"                "BSCD"
##  [85] "BSCS"               "BSIT-Electronics"    "BSBA-FM"
##  [88] "BSED-Science"       "BSED-VE"             "BSED-English"
##  [91] "BSIS"               "BSIT"                "BSIT"
##  [94] "BSIT"               "BSA"                 "BTVTED CP"
##  [97] "BSIT"               "BSHM"                "BSHM"
## [100] "BSED-FIL"
```

We didn't change the 'Avionics Technology' since we do not know what's the exact acronym for it.

CATEGORIZING TO DEPARTMENTS

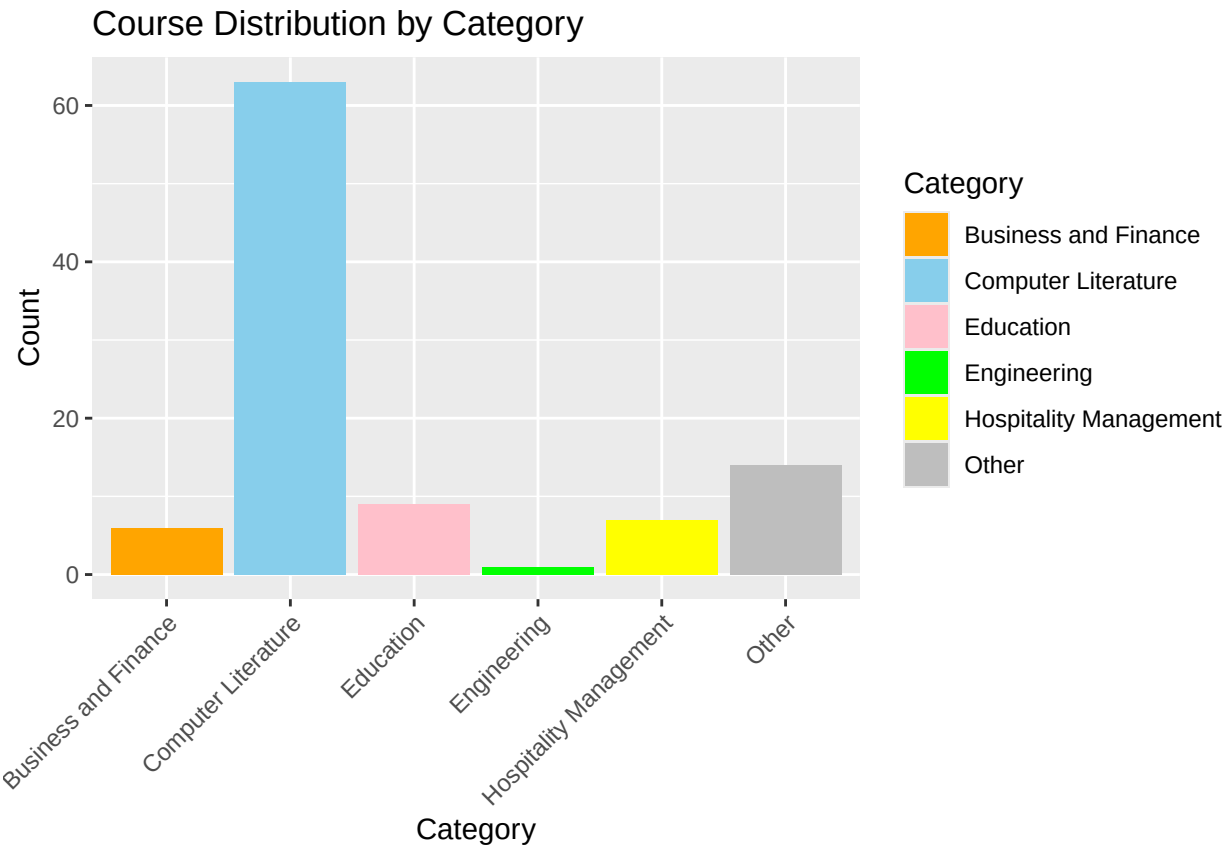
```
# Here, we will now categorize all the 27 courses we got from our data into its designated departments.

course <- deets$Course
course_factor <- factor(course) #27 levels
```

```
# here's how we categorized it
categorize_course <- function(course) {
  if (grepl("IT|CS|BSIT|IS", course)) {
    return("Computer Literature")
  }
  else if (grepl("ED|BSED", course)) {
    return("Education")
  }
  else if (grepl("BSA|BSBA", course)) {
    return("Business and Finance")
  }
  else if (grepl("CE", course)) {
    return("Engineering")
  }
  else if (grepl("BSHM", course)) {
    return("Hospitality Management")
  }
  else {
    return("Other")
  }
}

course_categories <- sapply(course, categorize_course)
```

```
# count the data of each category
category_counts <- table(course_categories)
```

```
# convert the table to a data frame
category_counts_df <- as.data.frame(category_counts)
names(category_counts_df) <- c("Category", "Count")
```

```
# create the bar plot for the department
ggplot(category_counts_df, aes(x = Category, y = Count, fill = Category)) +
  geom_bar(stat = "identity") +
  labs(title = "Course Distribution by Category",
       x = "Category",
       y = "Count") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_fill_manual(values = c("Computer Literature" = "skyblue",
                               "Education" = "pink",
                               "Business and Finance" = "orange",
                               "Engineering" = "green",
                               "Hospitality Management" = "yellow",
                               "Other" = "grey"))
```



The majority of respondents were from the computer literature department, which is unsurprising given that Adobe applications are software tools. However, the presence of respondents from diverse fields such as business and finance management, education, and engineering indicates the widespread applicability of Adobe software across various disciplines. This diversity suggests that Adobe applications are valued not only within technical domains but also in fields where creative and multimedia skills are increasingly essential.

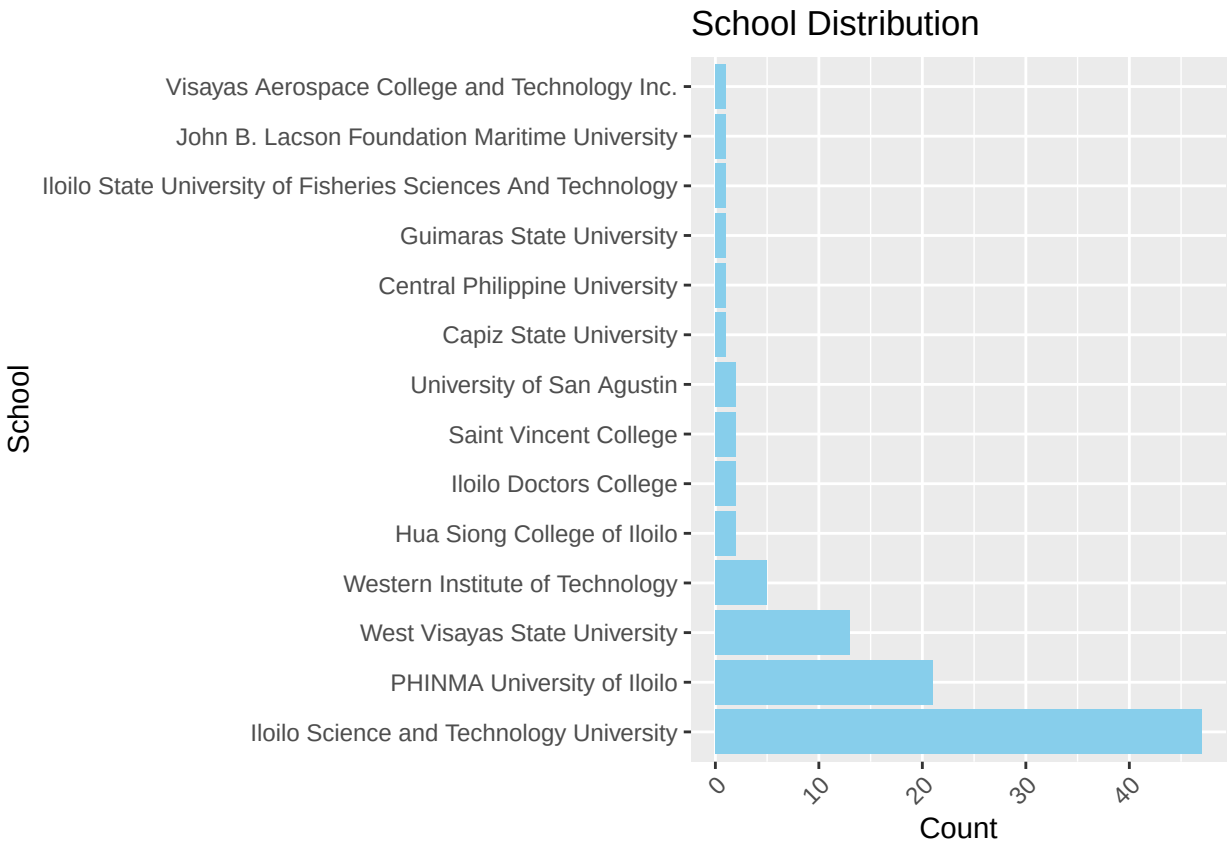
SCHOOL

```
school <- deets$School

school_factor <- factor(school)

school_counts <- table(school_factor)

school_counts_df <- as.data.frame(school_counts)
names(school_counts_df) <- c("School", "Count")
```

```
# sort the data frame by Count in descending order
school_counts_df <- school_counts_df[order(-school_counts_df$Count), ]
```

```
# creeate the bar plot for school
ggplot(school_counts_df, aes(x = reorder(School, -Count), y = Count)) +
  geom_bar(stat = "identity", fill = "skyblue") +
  labs(title = "School Distribution",
       x = "School",
       y = "Count") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
```

```
  coord_flip()
```



The primary universities where students use Adobe applications are Iloilo Science and Technology University, followed by PHINMA University of Iloilo, and then West Visayas State University, all these have a majority of students pursuing a computer literature track. This data suggests a strong correlation between the use of Adobe software and educational programs focused on computer literacy.

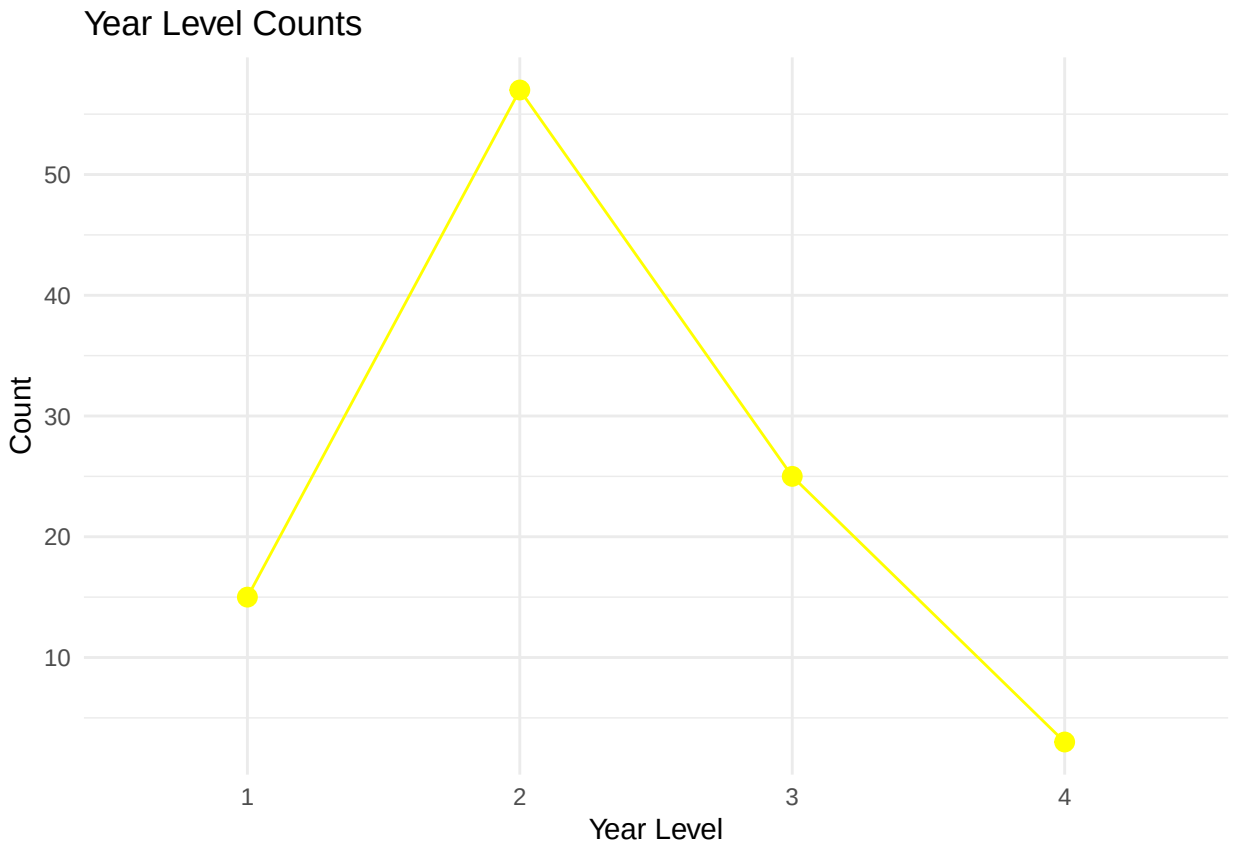
YEAR LEVEL

```
yrlvl <- deets$`Year-level`
yrlvl_factor <- factor(yrlvl)
```

```
yrlvl_counts <- table(yrlvl_factor)
```

```
yrlvl_df <- as.data.frame(yrlvl_counts)
yrlvl_df$Year_Level <- as.factor(rownames(yrlvl_df))
```

```
# on this rmd we simply use the ggplot function(needed to install
# phantomjs()),
# yet on the R file we used plot_ly function.

ggplot(yrlvl_df, aes(x = Year_Level, y = Freq, group = 1)) +
  geom_line(color = "yellow") +
  geom_point(color = "yellow", size = 3) +
  labs(x = "Year Level", y = "Count", title = "Year Level Counts") +
  theme_minimal()
```



This shows that the majority of users of Adobe Applications are in their second year, followed by third years come in second, with first years coming next , and fourth years being the least represented.

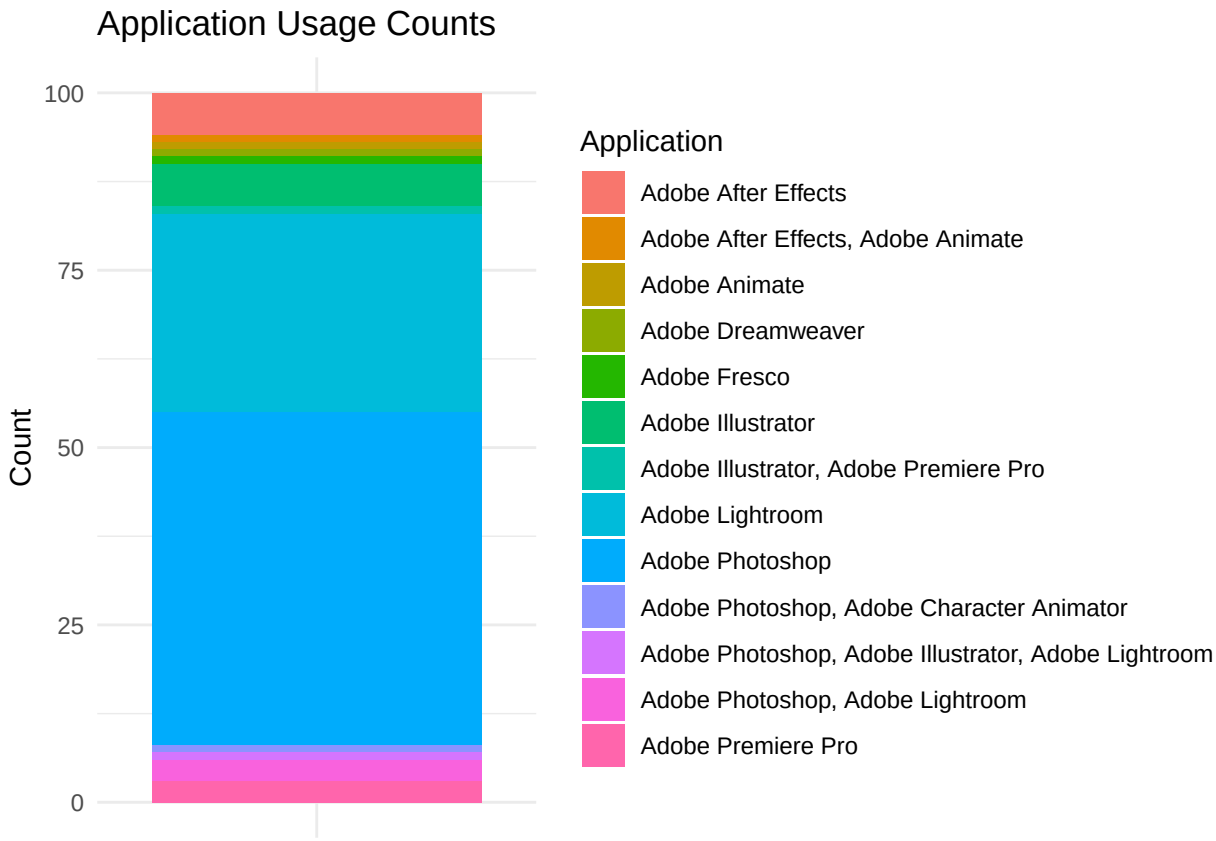
APPLICATIONS USED

```
app <- deets$`Here are some list of Adobe Cloud which of these applications do you use?`
```

```
app_factor <- factor(app)
```

```
app_counts <- table(app_factor)
app_counts_df <- as.data.frame(app_counts)
names(app_counts_df) <- c("Application", "Count")
app_counts_df <- app_counts_df[order(-app_counts_df$Count), ]
```

```
# plotting with stacked bar plot
ggplot(data = app_counts_df, aes(x = "", y = Count, fill = Application)) +
  geom_bar(stat = "identity") +
  labs(title = "Application Usage Counts",
       x = NULL,
       y = "Count") +
  theme_minimal() +
  theme(axis.text.x = element_blank(),
        legend.position = "right") +
  guides(fill=guide_legend(title="Application"))
```



We used the stacked bar plot on this demographic, so we can all see which color of the application is most used.

The most used application is Adobe Photoshop, followed by a tie between Adobe Premiere Pro and After Effects. Some applications, such as Dreamweaver and Fresco, appear to be used very little. Some of the respondents list several combination, that may indicate these applications are frequently used together in workflows. For example, the graph shows that Photoshop is often used in conjunction with Illustrator and Lightroom.

GETTING THE MEAN AND STANDARD DEVIATION FOR UTUAT

```
#Read the excel file
#questionnaire <- read_excel("questionnaire3g.xlsx", sheet = "Form Responses 1")
```

```
conversion <- function(response) {
  if (response == "Strongly Disagree") {
    return(1)
  } else if (response == "Disagree") {
    return(2)
  } else if (response == "Neutral") {
    return(3)
  } else if (response == "Agree") {
    return(4)
  } else {
    return(5)
```

```
  }
}
```

```
columns_to_convert <- c(10, 11, 12, 14, 17, 19, 20, 22, 23, 25, 26, 27, 28, 29, 30, 31, 32)

for (column_index in columns_to_convert) {
  deets[[column_index]] <- sapply(deets[[column_index]], conversion)
}
```

```
conversion1 <- function(response) {
  if (is.na(response)) {
    return(NA)
  } else if (response == "Yes") {
    return(1)
  } else if (response == "No") {
    return(2)
  } else if (response == "Possibly") {
    return(3)
  } else {
    return(NA)
  }
}
```

```
columns_to_convert1 <- c(9,15,16,18,21,24,33,34,35)

for (column in columns_to_convert1) {
 deets[[column]] <- sapply(deets[[column]], conversion1)
}
```

```
# Calculate the mean and sd for Performance Expectancy
performanceExpectancy <- deets[, c(9:11)]
```

```
# getting the mean and sd
means1 <- colMeans(performanceExpectancy, na.rm = TRUE)
means1
```

```
##                       Do you find the Adobe application useful in my studies?
##                                                                          1.22
## By using the Adobe application enables you to accomplish tasks more quickly.
##                                                                          3.90
##                  By using the Adobe application it increases my productivity.
##                                                                          3.79
```

```
sd_pe <- sapply(performanceExpectancy, sd, na.rm = TRUE)
sd_pe
```

```
##                       Do you find the Adobe application useful in my studies?
##                                                                     0.4163332
## By using the Adobe application enables you to accomplish tasks more quickly.
##                                                                     0.8468599
##                  By using the Adobe application it increases my productivity.
##                                                                     0.9133687
```

```
# getting the average
avg_mean1 <- round(mean(means1), 2)
avg_mean1
```

```
## [1] 2.97
```

```
avg_sd1 <- round(mean(sd_pe), 2)
avg_sd1
```

```
## [1] 0.73
```

```
# calculate the mean and sd for Effort Expectancy
effortExpectancy <- deets[, c(12:14)]
```

```
# getting the mean and sd
means2 <- colMeans(effortExpectancy, na.rm = TRUE)
means2
```

```
##    My interaction with the Adobe application would be clear and understandable.
##                                                                            5.00
## It would be easy for me to become more skillful at using the Adobe application.
##                                                                            3.96
##                       Learning to operate the Adobe application is easy for me.
##                                                                            5.00
```

```
sd_ee <- sapply(effortExpectancy, sd, na.rm = TRUE)
sd_ee
```

```
##    My interaction with the Adobe application would be clear and understandable.
##                                                                       0.0000000
## It would be easy for me to become more skillful at using the Adobe application.
##                                                                       0.9419516
##                       Learning to operate the Adobe application is easy for me.
##                                                                       0.0000000
```

```
# average
avg_mean2 <- round(mean(means2), 2)
avg_mean2
```

```
## [1] 4.65
```

```
avg_sd2 <- round(mean(sd_ee), 2)
avg_sd2
```

```
## [1] 0.31
```

```
# Calculate the mean and sd for Social Influence
socialInfluence <- deets[, c(19:21)]
```

```
# getting the mean and sd
means3 <- colMeans(socialInfluence, na.rm = TRUE)
means3
```

```
##                    My peers influence necessary to use the Adobe application.
##                                                                          3.59
## People who are important to me think that I should use the Adobe application.
##                                                                          3.65
##         In general, the school has supported the use of the Adobe application
##                                                                          1.17
```

```
sd_si <- sapply(socialInfluence, sd, na.rm = TRUE)
sd_si
```

```
##                    My peers influence necessary to use the Adobe application.
```

```
##                                                                    0.6830561
## People who are important to me think that I should use the Adobe application.
##                                                                    0.7961397
##         In general, the school has supported the use of the Adobe application
##                                                                    0.3775252
```

```
# average
avg_mean3 <- round(mean(means3), 2)
avg_mean3
```

```
## [1] 2.8
```

```
avg_sd3 <- round(mean(sd_si), 2)
avg_sd3
```

```
## [1] 0.62
```

```
# Calculate the mean and sd for Facilitating Conditions
facilitatingConditions <- deets[, c(22:25)]
```

```
# getting the mean and sd
means4 <- colMeans(facilitatingConditions, na.rm = TRUE)
means4
```

```
##                I have the resources necessary to use the Adobe application.
##                                                                        3.53
##                I know the knowledge necessary to use the Adobe application.
##                                                                        5.00
##           The Adobe application is not compatible with other systems I use.
##                                                                        2.18
## My peers is available for assistance with Adobe application's difficulties.
##                                                                        3.77
```

```
sd_fc <- sapply(facilitatingConditions, sd, na.rm = TRUE)
sd_fc
```

```
##                I have the resources necessary to use the Adobe application.
##                                                                   0.7028801
##                I know the knowledge necessary to use the Adobe application.
##                                                                   0.0000000
##           The Adobe application is not compatible with other systems I use.
##                                                                   0.7704125
## My peers is available for assistance with Adobe application's difficulties.
##                                                                   0.8391313
```

```
avg_mean4 <- round(mean(means4), 2)
avg_mean4
```

```
## [1] 3.62
```

```
avg_sd4 <- round(mean(sd_fc), 2)
avg_sd4
```

```
## [1] 0.58
```

```
# Calculate the mean and sd for Behavioral Intention to use the system
behavioralIntention <- deets[, c(33:35)]
```

```
# getting the mean and sd
means5 <- colMeans(behavioralIntention, na.rm = TRUE)
```

```
means5
```

```
##      I intend to use the Adobe application in the next 2 months.
##                                                             1.17
## I predict I will use the Adobe application in the next 2 months.
##                                                             1.18
##                   I plan to use the system in the next 2 months.
##                                                             1.16
```

```
sd_bi <- sapply(behavioralIntention, sd, na.rm = TRUE)
sd_bi
```

```
##      I intend to use the Adobe application in the next 2 months.
##                                                        0.3775252
## I predict I will use the Adobe application in the next 2 months.
##                                                        0.3861229
##                   I plan to use the system in the next 2 months.
##                                                        0.3684529
```

```
# average
avg_mean5 <- round(mean(means5), 2)
avg_mean5
```

```
## [1] 1.17
```

```
avg_sd5 <- round(mean(sd_bi), 2)
avg_sd5
```

```
## [1] 0.38
```

```
#####Make a data frame of all factors######
pe <- data.frame(
  Description = "Performance Expectancy",
  Mean = means1,
  SD = sd_pe,
  Average_Mean = avg_mean1,
  Average_SD = avg_sd1
)
```

```
ee <- data.frame(
  Description = "Effort Expectancy",
  Mean = means2,
  SD = sd_ee,
  Average_Mean = avg_mean2,
  Average_SD = avg_sd2
)
```

```
si <- data.frame(
  Description = "Social Influence",
  Mean = means3,
  SD = sd_si,
  Average_Mean = avg_mean3,
  Average_SD = avg_sd3
)
```

```
fc <- data.frame(
  Description = "Facilitating Conditions",
  Mean = means4,
```

```
  SD = sd_fc,
  Average_Mean = avg_mean4,
  Average_SD = avg_sd4
)
```

```
bi <- data.frame(
  Description = "Behavioral Intention",
  Mean = means5,
  SD = sd_bi,
  Average_Mean = avg_mean5,
  Average_SD = avg_sd5
)
```

```
summary <- rbind(pe, ee, si, fc, bi)
kable(summary)
```

	Description	Mean	SD	Average_Mean	Average_SD
Do you find the Adobe application useful in my studies?	Performance Expectancy	1.22	0.4163332	2.97	0.73
By using the Adobe application enables you to accomplish tasks more quickly.	Performance Expectancy	3.90	0.8468599	2.97	0.73
By using the Adobe application it increases my productivity.	Performance Expectancy	3.79	0.9133687	2.97	0.73
My interaction with the Adobe application would be clear and understandable.	Effort Expectancy	5.00	0.0000000	4.65	0.31
It would be easy for me to become more skillful at using the Adobe application.	Effort Expectancy	3.96	0.9419516	4.65	0.31
Learning to operate the Adobe application is easy for me.	Effort Expectancy	5.00	0.0000000	4.65	0.31
My peers influence necessary to use the Adobe application.	Social Influence	3.59	0.6830561	2.80	0.62
People who are important to me think that I should use the Adobe application.	Social Influence	3.65	0.7961397	2.80	0.62
In general, the school has supported the use of the Adobe application	Social Influence	1.17	0.3775252	2.80	0.62
I have the resources necessary to use the Adobe application.	Facilitating Conditions	3.53	0.7028801	3.62	0.58
I know the knowledge necessary to use the Adobe application.	Facilitating Conditions	5.00	0.0000000	3.62	0.58
The Adobe application is not compatible with other systems I use.	Facilitating Conditions	2.18	0.7704125	3.62	0.58
My peers is available for assistance with Adobe application's difficulties.	Facilitating Conditions	3.77	0.8391313	3.62	0.58
I intend to use the Adobe application in the next 2 months.	Behavioral Intention	1.17	0.3775252	1.17	0.38
I predict I will use the Adobe application in the next 2 months.	Behavioral Intention	1.18	0.3861229	1.17	0.38
I plan to use the system in the next 2 months.	Behavioral Intention	1.16	0.3684529	1.17	0.38