Demographics

Joshua Somosera

2024-05-04

```r
library(readr)
library(ggplot2)
data <- read_csv("/cloud/project/Group Project (Survey)/data.csv")
```

```
## New names:
## Rows: 50 Columns: 37
## -- Column specification
## -------------------------------------------------------- Delimiter: "," chr
## (37): Timestamp, Username, Name (First Name, Last Name):, Age:, SEX:, Ge...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `` -> `...37`
```

```r
print(data)
```

```
## # A tibble: 50 x 37
##    Timestamp              Username Name (First Name, La~1 `Age:` `SEX:` `Gender:`
##    <chr>                  <chr>    <chr>                  <chr>  <chr>  <chr>
##  1 2024/03/10 10:49:41 ~ primero~ Ellema, Prime Rose     20     Female Straight
##  2 2024/03/10 10:50:38 ~ keilapa~ Keila, Palmos          19     Female Straight
##  3 2024/03/10 10:56:18 ~ reneero~ Renee Rose Flogoso     21     Female Straight
##  4 2024/03/10 10:56:26 ~ armonio~ Mechaila Armonio       19     Female Straight
##  5 2024/03/10 10:56:45 ~ talong7~ Christian Dave Magno   21     Male   Straight
##  6 2024/03/10 10:59:55 ~ camango~ ARGIE CAMANGON         19     Male   Straight
##  7 2024/03/10 11:00:30 ~ ventila~ Roleah Anne            20     Female Straight
##  8 2024/03/10 11:02:12 ~ katemar~ Kayt                   11     Female Straight
##  9 2024/03/10 11:09:54 ~ brillan~ Meryll Joy Mana-ay     19     Female Straight
## 10 2024/03/10 11:20:05 ~ opino.a~ Arabella Kristel ,Opi~ 20     Female Straight
## # i 40 more rows
## # i abbreviated name: 1: `Name (First Name, Last Name):`
## # i 31 more variables: `School Name:` <chr>, `SECTION:` <chr>,
## #   `Course: (Type only the name e.g Information technology)` <chr>,
## #   `How do you use canva? (Check all boxes that apply)` <chr>,
## #   `Have you used Canva to help with any assignments or projects connected to your studies?` <chr>,
## #   `using Canva for school-related work helps ME complete assignments more quickly.` <chr>, ...
```

```r
#Removing the unnecessary columns (Timestamp, School Name, Section, and Course)
CleanedData <- data[, -c(1,7, 8, 9)]
```

```r
CleanedData$`Gender:`
```

```
##  [1] "Straight"          "Straight"          "Straight"
##  [4] "Straight"          "Straight"          "Straight"
##  [7] "Straight"          "Straight"          "Straight"
## [10] "Straight"          "Bisexual"          "Straight"
## [13] "Prefer Not to Say" "Bisexual"          "Straight"
## [16] "Prefer Not to Say" "Straight"          "Prefer Not to Say"
## [19] "Straight"          "Straight"          "Bisexual"
## [22] "Straight"          "Straight"          "Prefer Not to Say"
## [25] "Transgender"       "Straight"          "Straight"
## [28] "Straight"          "Straight"          "Straight"
## [31] "Straight"          "Straight"          "Straight"
## [34] "Straight"          "Straight"          "Straight"
## [37] "Prefer Not to Say" "Straight"          "Straight"
## [40] "Straight"          "Prefer Not to Say" "Straight"
## [43] "Straight"          "Straight"          "Straight"
## [46] "Straight"          "Straight"          "Straight"
## [49] "Straight"          "Straight"
```

```
#Factor Gender
CleanedData$`Gender:`[is.na(CleanedData$`Gender:`)] <- "Prefer Not To Say"
genderfactor <- factor(CleanedData$`Gender:`, levels = c("Straight", "Bisexual", "Gay", "Lesbian", "Tran
summary(genderfactor)
```

```
##                  Straight                  Bisexual                       Gay
##                        40                         3                         0
##                   Lesbian               Transgender Non-binary/non-conforming
##                         0                         1                         0
##         Prefer Not To Say                      NA's
##                         0                         6
```

```
#Factor Sex
sexfactor<-factor(CleanedData$`SEX:`, levels = c("Male", "Female"))
summary(sexfactor)
```

```
##   Male Female
##     23     27
```

```
#Factor Age
#The data has "$1" as a value, converted it to "21"
CleanedData$`Age:`[CleanedData$`Age:` == "$1"] <- 21
# Convert Age: column to numeric
CleanedData$`Age:` <- as.numeric(CleanedData$`Age:`)
agefactor <- factor(CleanedData$`Age:`, levels = 11:23)
summary(agefactor)
```

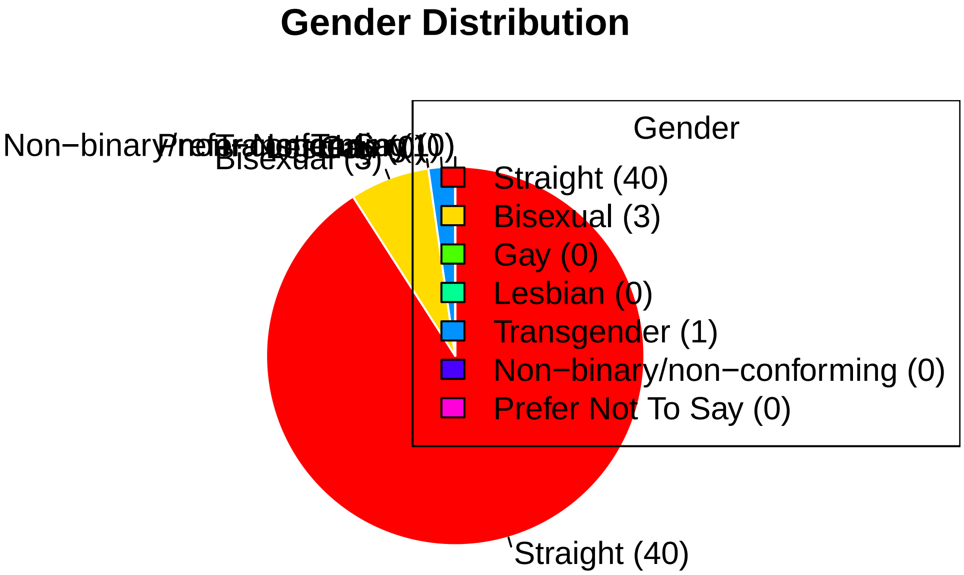
```
## 11 12 13 14 15 16 17 18 19 20 21 22 23
##  1  0  0  0  0  1  1  4 14 20  8  0  1
```

```
#Getting the mean for Age
age <- c(CleanedData$`Age:`)
average <- mean(age, na.rm = TRUE)
avg <- paste("The mean age of the respondents is", average)
print(avg)
```

```
## [1] "The mean age of the respondents is 19.46"
```

```
gender_counts <- table(genderfactor)

# Plot a pie chart
pie(gender_counts,
    main = "Gender Distribution",
    col = rainbow(length(gender_counts)),
```

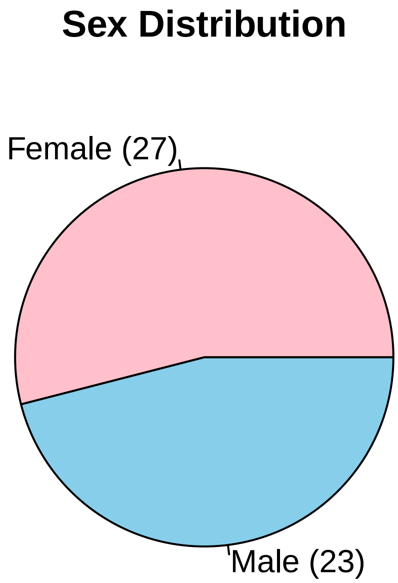
```r
    labels = paste(names(gender_counts), " (", gender_counts, ")", sep = ""),
    clockwise = TRUE,
    density = NULL,
    angle = 45,
    init.angle = 90,
    border = "white"
)

# Add a legend with counts
legend("topright",
       legend = paste(names(gender_counts), " (", gender_counts, ")", sep = ""),
       fill = rainbow(length(gender_counts)),
       title = "Gender"
)
```



```r
"This bar plot visualizes the gender distribution among respondents in the survey.
Each bar represents the count of respondents on who identified with a particular gender category.
The height of each bar indicates the number of respondents, with labels on top of each bar showing
the exact count.The legend on the top-right corner provides a color key for each gender category."
```

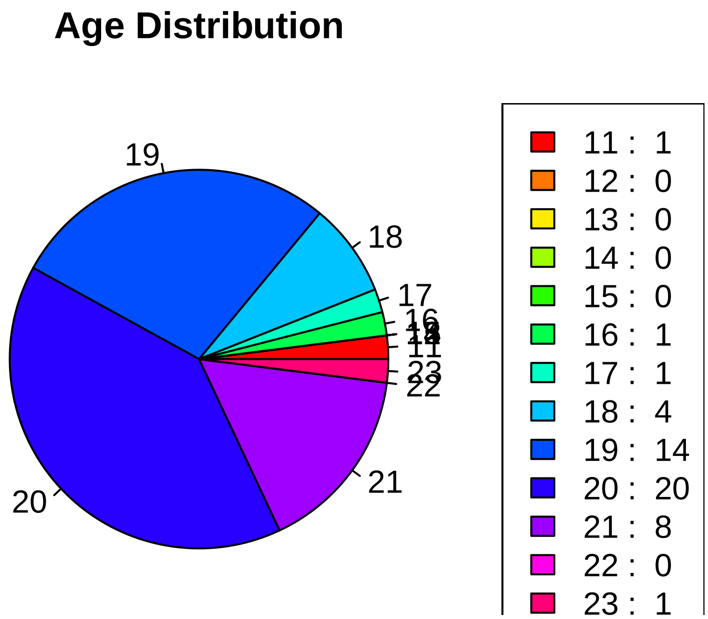
```
## [1] "This bar plot visualizes the gender distribution among respondents in the survey. \nEach bar re
```

```r
sex_table <- table(CleanedData$`SEX:`)

sex_colors <- c("pink", "skyblue")

pie(sex_table,
    main = "Sex Distribution",
    labels = paste(names(sex_table), " (", sex_table, ")", sep = ""),
    col = sex_colors
)
```



```
#
```

```
age_counts <- table(agefactor)
age_labels <- names(age_counts)

pie(age_counts, labels = age_labels, col = rainbow(length(age_counts)), main = "Age Distribution")
legend("topright", legend = paste(age_labels, ": ", age_counts), fill = rainbow(length(age_counts)))
```


Age Distribution
11 : 1
12 : 0
13 : 0
14 : 0
15 : 0
16 : 1
17 : 1
18 : 4
19 : 14
20 : 20
21 : 8
22 : 0
23 : 1