R6R7.r

aravi

2023-11-28

```
#Setting Directory
getwd()
```

```
## [1] "C:/Users/aravi/Desktop/StatisticsUsingR/SocialMediaStudy"
```

```
setwd("C:\\Users\\aravi\\Desktop\\StatisticsUsingR\\SocialMediaStudy")
getwd()
```

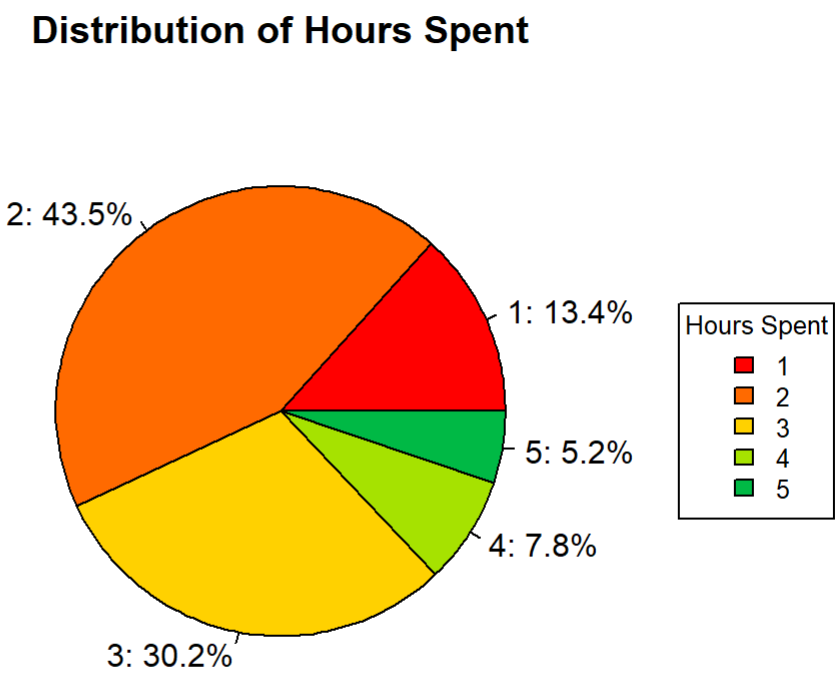
```
## [1] "C:/Users/aravi/Desktop/StatisticsUsingR/SocialMediaStudy"
```

```
#Importing Packages
library(ggplot2)
```

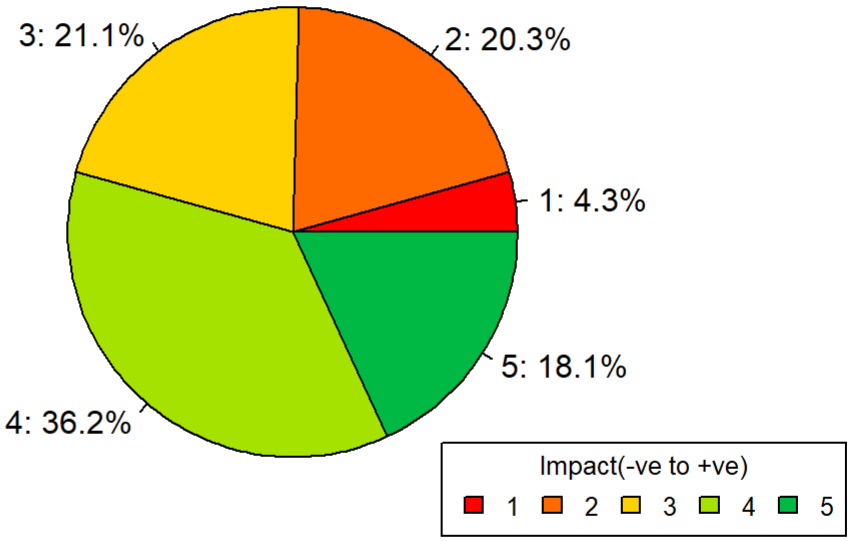
```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

```
#Colors
pos_neg_pallete<- c("#FF0000", "#FF6A00", "#FFD100", "#A6E200", "#00B945")
my_palette <- rev( c("#FF6F61", "#FFD166", "#6B5B95", "#88B04B", "#4F6367"))
blue_palette <- rev(c("#3498db", "#5dade2", "#85c1e9", "#aed6f1", "#d6eaf8"))
red_palette <- c("#e74c3c", "#ec7063", "#f1948a", "#f5b7b1", "#f9bdbb")
green_palette <- c("#2ecc71", "#58d68d", "#82e0aa", "#a9dfbf", "#d0e9c6")
yellow_palette <- c("#f39c12", "#f5b041", "#f8c471", "#f9e79f", "#fcf3cf")
purple_palette <- c("#8e44ad", "#af7ac5", "#bb8fce", "#d2b4de", "#e8daef")

#data
data <- read.csv("IPSMOI.csv")
data <- data[-c(1, 2, 3)]



#HRSPD - Hours Spend per Day
#SIEFF - Social Interaction Effects Friends and Family
#ACPA - Academic Performance Affect
#MHAF - Mental Health Affected
#QRI - Quality and Relevance of Information
#PC - Privacy Concerns in social Media
#CCF - How Often Content Creation
#SSCC - Support Community and social Cause
#UEP - Use on Education Purpose
#CASI - Creating Awareness On Social Issues
#PDDA - Purchase decisions due to advertising
#SOSP - How Often Social Media Use
#SMPF - Social Media Platform
#RPS - Review and Update Privacy Settings
#TCEM - Type of Content Engagement Most
#OLF - Online Friends
#APOC - Actively Participate in Online Community
#SMIF - Social Media Influencer Following
#PFO - Provide Feedback or opinions
#PG - Personal Growth

# Change Column Names to abbreviations
colnames(data)<-c("AgeGroup","Gender","CareerStatus","ResidentialArea",
                  "HRSPD","SIEFF","ACPA","MHAF","QRI",
                  "PC","CCF","SSCC","UEP","CASI","PDDA",
                  "SOSP","SMPF","RPS","TCEM","OLF","APOC","SMIF","PFO","PG")


#Data Conversions To numeric Ordinal Data
#data$HRSPD
data$HRSPD<-as.numeric(
  factor(data$HRSPD,
         levels = c("Less than 1 hour",
                    "1-2 hours",
                    "3-4 hours",
                    "5-6 hours",
                    "More than 6 hours")))
pie(table(data$HRSPD), labels = paste0(names(table(data$HRSPD)), ": ", round(table(data$HRSP
D) / sum(table(data$HRSPD)) * 100, 1), "%"),
    col = pos_neg_pallete, main = "Distribution of Hours Spent")
```

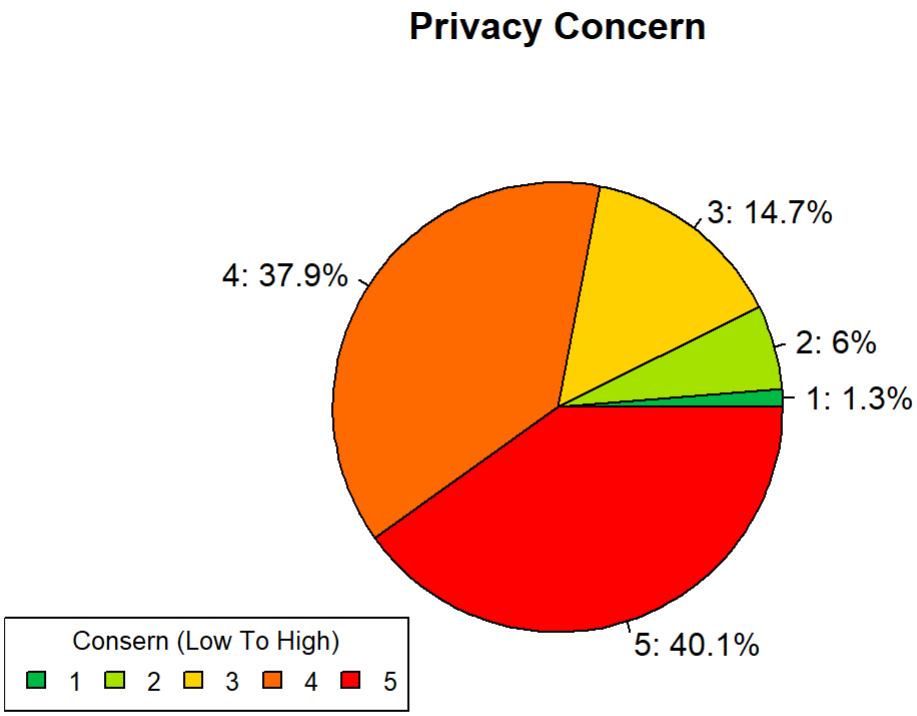
```
legend("right", legend = names(table(data$HRSPD)), fill = pos_neg_pallete, title = "Hours Spe
nt",xpd=TRUE, inset=c(0, -.15), cex=.8)
```



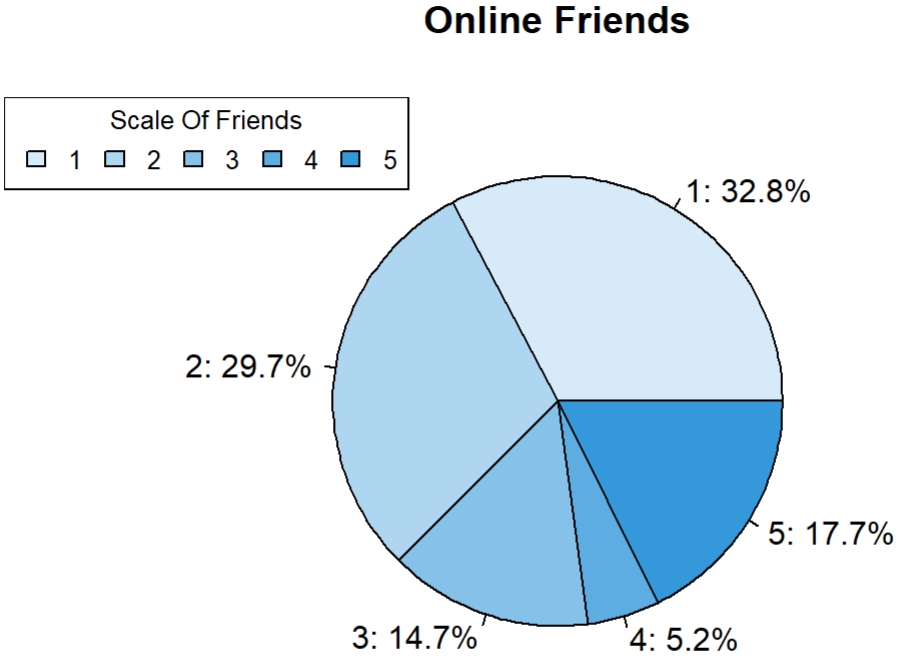

```
#data$SIEFF
data$SIEFF<-as.numeric(
  factor(data$SIEFF,
         levels = c("Significantly Negative",
                    "Slightly Negative",
                    "No Impact",
                    "Slightly Positive",
                    "Significantly Positive")))
pie(table(data$SIEFF), labels = paste0(names(table(data$SIEFF)), ": ", round(table(data$SIEF
F) / sum(table(data$SIEFF)) * 100, 1), "%"),
    col = pos_neg_pallete, main = "Distribution of Social interaction Impact")
legend("bottomright", legend = names(table(data$SIEFF)),horiz = TRUE, fill = pos_neg_pallete,
title = "Impact(-ve to +ve)",cex = 0.8)
```



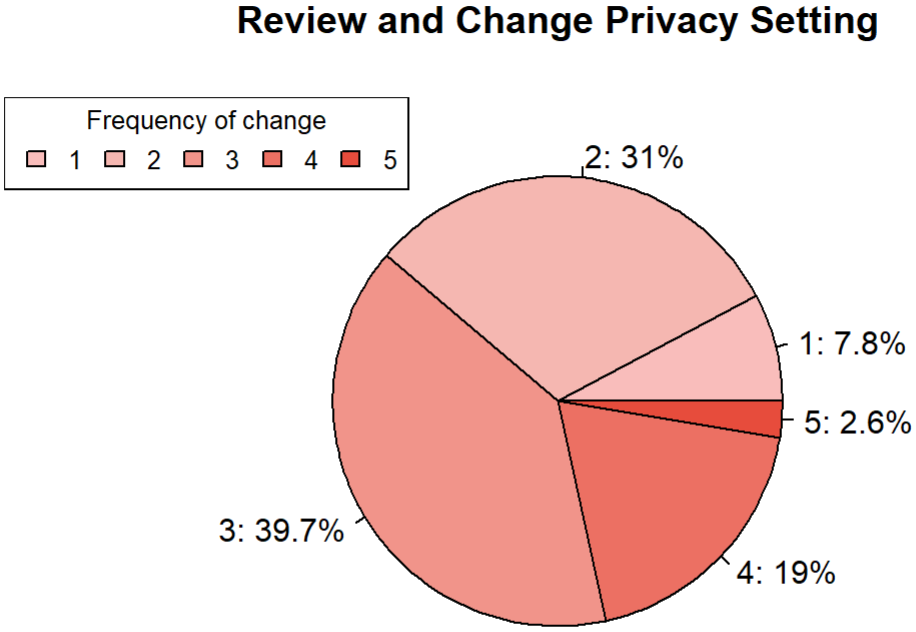

```
#data$CareerStatus
data$CareerStatus<-factor(data$CareerStatus)

#data$PC
data$PC<-as.numeric(
  factor(data$PC,
         levels = c("Not Concerned at All",
                    "Not Very Concerned",
                    "Neutral",
                    "Somewhat Concerned",
                    "Very Concerned")))
pie(table(data$PC), labels = paste0(names(table(data$PC)), ": ", round(table(data$PC) / sum(t
able(data$PC)) * 100, 1), "%"),
    col = rev(pos_neg_pallete), main = "Privacy Concern")
legend("bottomleft", legend = names(table(data$PC)), fill = rev(pos_neg_pallete),horiz = TRUE
,title = "Consern (Low To High)",cex=.8)
```



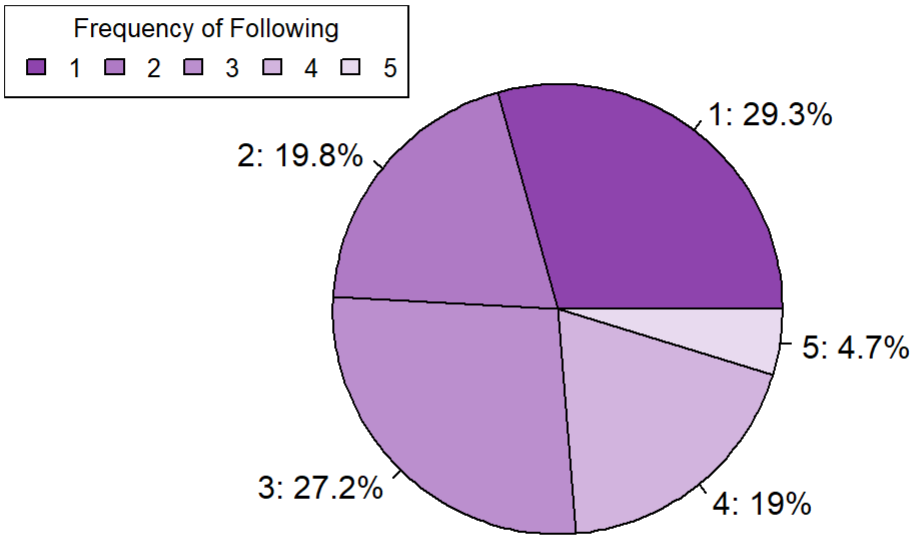

```
#data$OLF
data$OLF<-as.numeric(
  factor(data$OLF,
         levels = c("No Online Friends",
                    "Less than 10",
                    "10-30",
                    "30-50",
                    "More than 50")))

pie(table(data$OLF), labels = paste0(names(table(data$OLF)), ": ", round(table(data$OLF) / su
m(table(data$OLF)) * 100, 1), "%"),
    col = blue_palette, main = "Online Friends")
legend("topleft", legend = names(table(data$OLF)), horiz = TRUE,fill = blue_palette, title =
"Scale Of Friends",cex=.8)
```



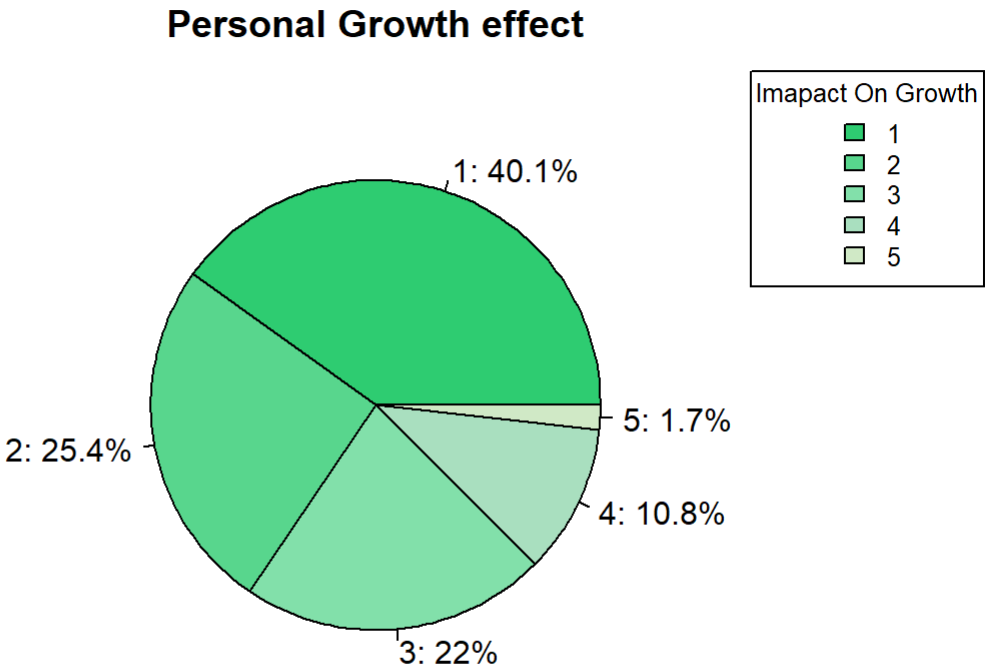

```
#data$RPS
data$RPS<-as.numeric(
  factor(data$RPS,
         levels = c("Never",
                    "Rarely",
                    "Occasionally",
                    "Frequently",
                    "All the time")))
pie(table(data$RPS), labels = paste0(names(table(data$RPS)), ": ", round(table(data$RPS) / su
m(table(data$RPS)) * 100, 1), "%"),
    col = rev(red_palette), main = "Review and Change Privacy Setting")
legend("topleft", legend = names(table(data$RPS)), fill = rev(red_palette),horiz = TRUE, titl
e = "Frequency of change",cex=.8)
```



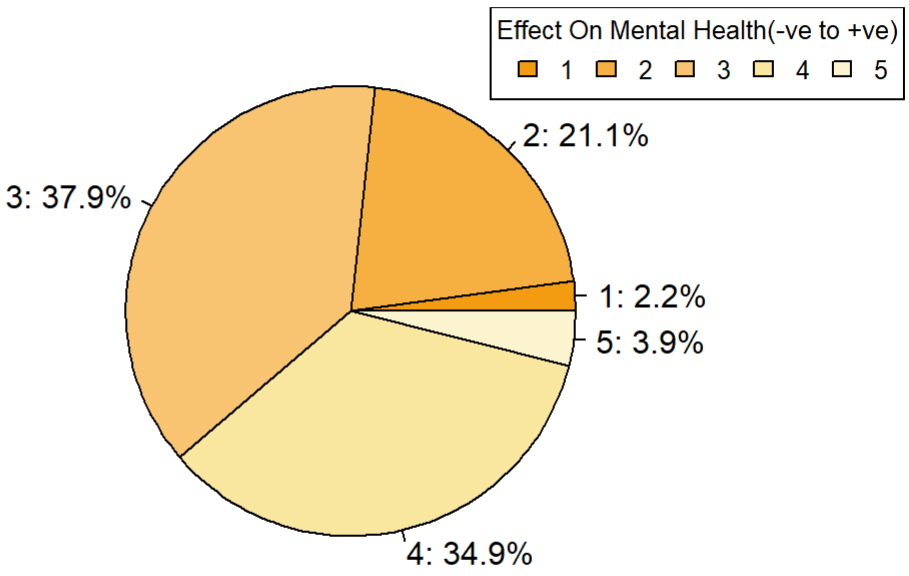

```
#data$SMIF
data$SMIF<-as.numeric(
  factor(data$SMIF,
         levels = c("Never",
                    "Rarely",
                    "Occasionally",
                    "Frequently",
                    "All the time")))
pie(table(data$SMIF), labels = paste0(names(table(data$SMIF)), ": ", round(table(data$SMIF) /
sum(table(data$SMIF)) * 100, 1), "%"),
    col = purple_palette, main = "Social Media influencer Following Frequency")
legend("topleft", legend = names(table(data$SMIF)), fill = purple_palette, horiz = TRUE,title
= "Frequency of Following",cex=.8)
```



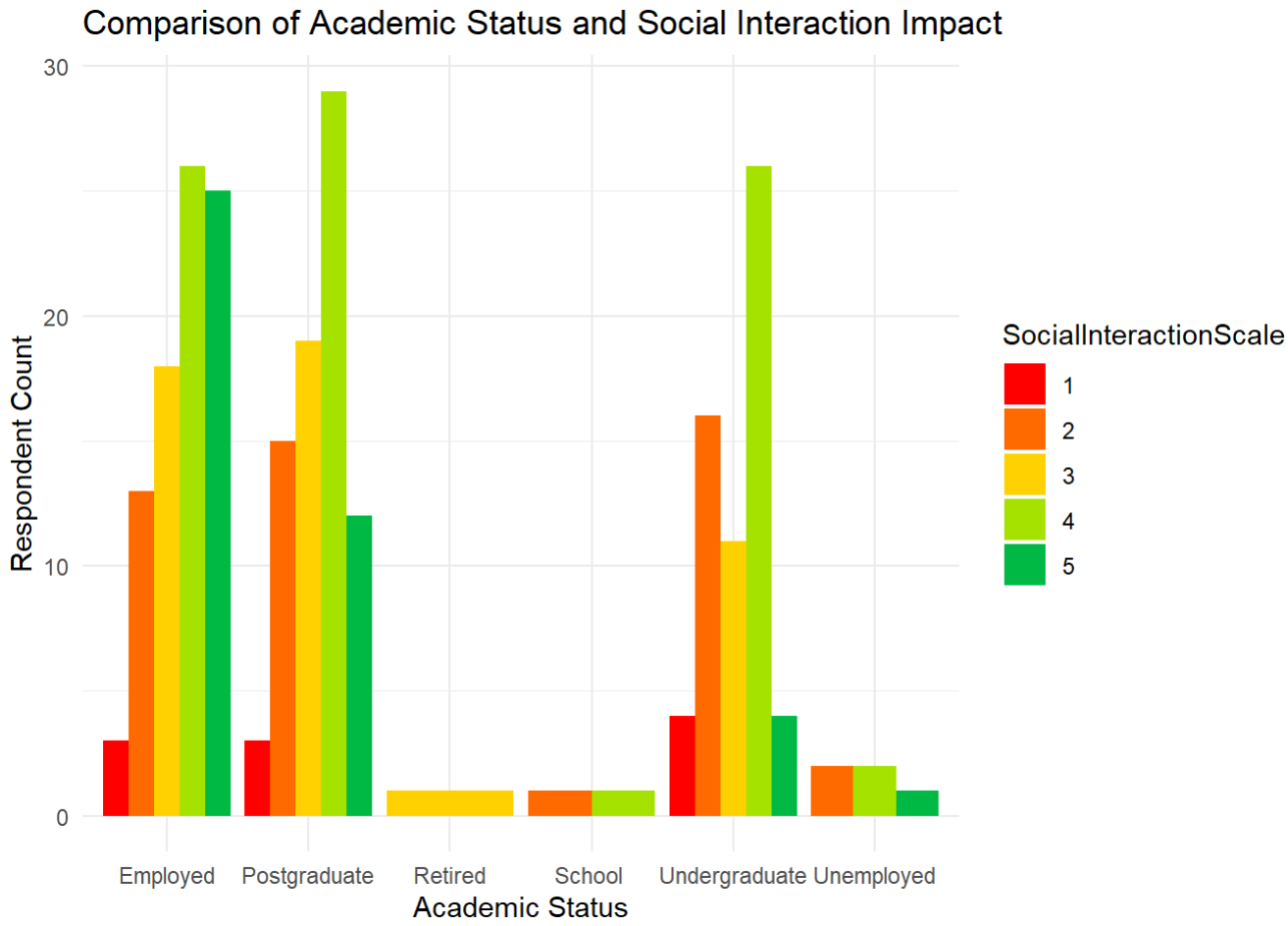

```
#data$PG
data$PG<-as.numeric(
  factor(data$PG,
         levels = c("Never",
                    "Rarely",
                    "Occasionally",
                    "Frequently",
                    "All the time")))
pie(table(data$PG), labels = paste0(names(table(data$PG)), ": ", round(table(data$PG) / sum(t
able(data$PG)) * 100, 1), "%"),
    col = green_palette, main = "Personal Growth effect")
legend("topright", legend = names(table(data$PG)), fill = green_palette, title = "Imapact On
Growth",xpd=TRUE, inset=c(-0.05,-0.05), cex=.8)
```




```
#data$MHAF
data$MHAF<-as.numeric(
  factor(data$MHAF,
         levels = c("Extremely Negative",
                    "Negative",
                    "No Impact",
                    "Positive",
                    "Extremely Positive")))
pie(table(data$MHAF), labels = paste0(names(table(data$MHAF)), ": ", round(table(data$MHAF) /
sum(table(data$MHAF)) * 100, 1), "%"),
    col = yellow_palette, main = "Personal Growth effect")
legend("topright", legend = names(table(data$MHAF)),horiz = TRUE, fill = yellow_palette, titl
e = "Effect On Mental Health(-ve to +ve)", inset=c(-0.0,-0.0), cex=.8)
```




```
#Comparison Of Academic Status and Social Interaction Impact
SocialInteractionScale<-as.factor(data$SIEFF)
ggplot(data, aes(x = CareerStatus, fill = SocialInteractionScale)) +
  geom_bar(position = "dodge") +
  labs(title = "Comparison of Academic Status and Social Interaction Impact",
       x = "Academic Status",
       y = "Respondent Count") +
  scale_fill_manual(values = pos_neg_pallete)+
  theme_minimal()
```

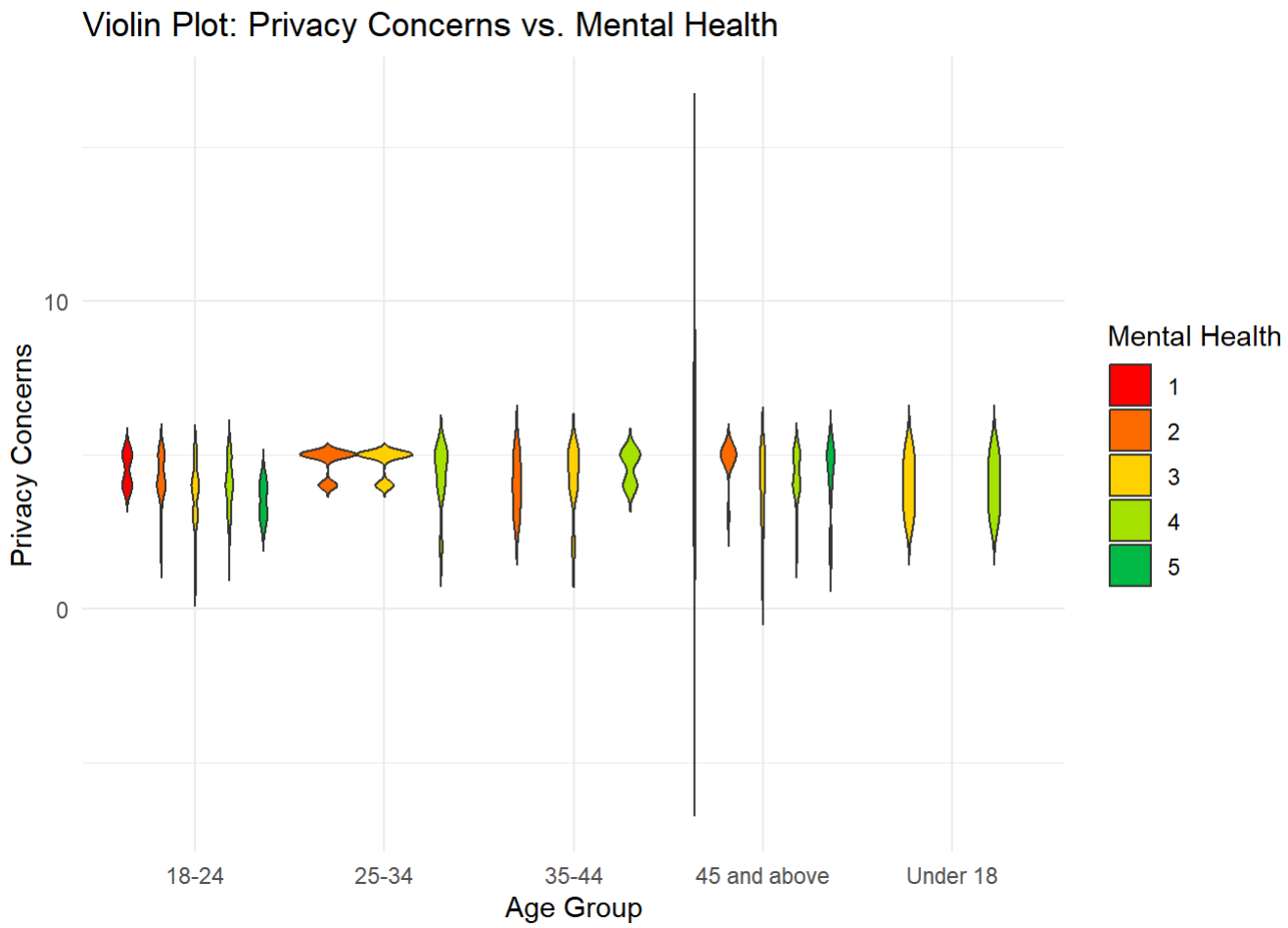



```
print("It was Seen that Employed Respondents had better
      social Interaction than other groups like unemployed and
      retired and school respondents having little negative effect.
      This can be a conclusion that people having more real life interaction
      have felt the negative effect on social interaction after social media use")
```

```
## [1] "It was Seen that Employed Respondents had better\n      social Interaction than other
groups like unemployed and\n      retired and school respondents having little negative effec
t.\n      This can be a conclusion that people having more real life interaction \n      have
felt the negative effect on social interaction after social media use"
```

```
#Comparing Privacy Concerns and Mental Health
ggplot(data, aes(x = factor(AgeGroup), y = PC, fill = factor(MHAF))) +
  geom_violin(trim = FALSE) +
  labs(title = "Violin Plot: Privacy Concerns vs. Mental Health",
       x = "Age Group",
       y = "Privacy Concerns",
       fill = "Mental Health") +
  scale_fill_manual(values = pos_neg_pallete) +
  theme_minimal()
```

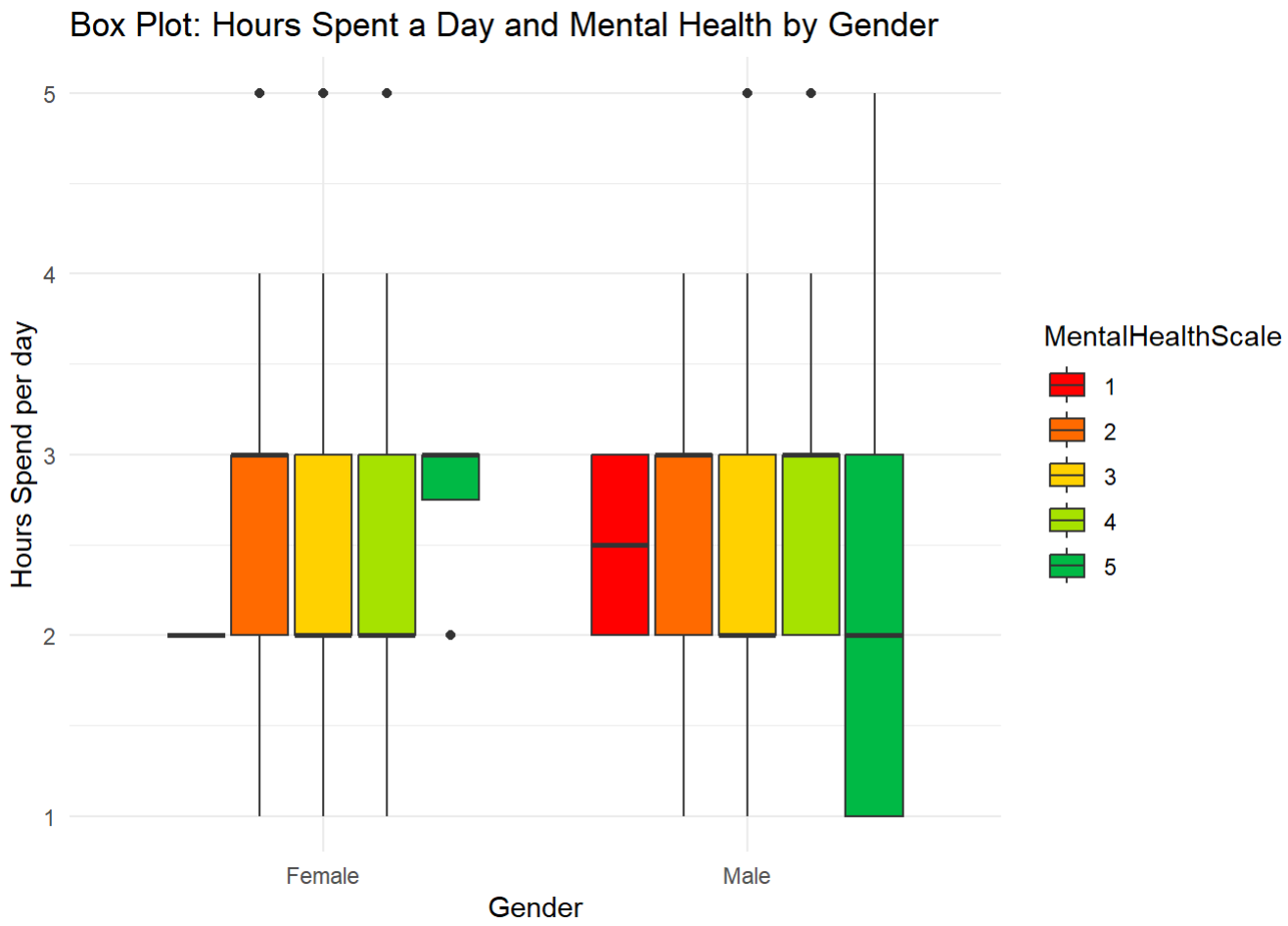
```
## Warning: Groups with fewer than two data points have been dropped.
## Groups with fewer than two data points have been dropped.
```




```
print("The data shows that spread on highly negative metal health
      effect with very less respondents for the age group 45+ .
      With under 18 having very less negative impact on mental health compared to others.
      and a small spike in the respondents on the slightly
      negative side in the age group of 25-34.")
```

```
## [1] "The data shows that spread on highly negative metal health \n      effect with very l
ess respondents for the age group 45+ .\n      With under 18 having very less negative impact
on mental health compared to others.\n      and a small spike in the respondents on the sligh
tly\n      negative side in the age group of 25-34."
```

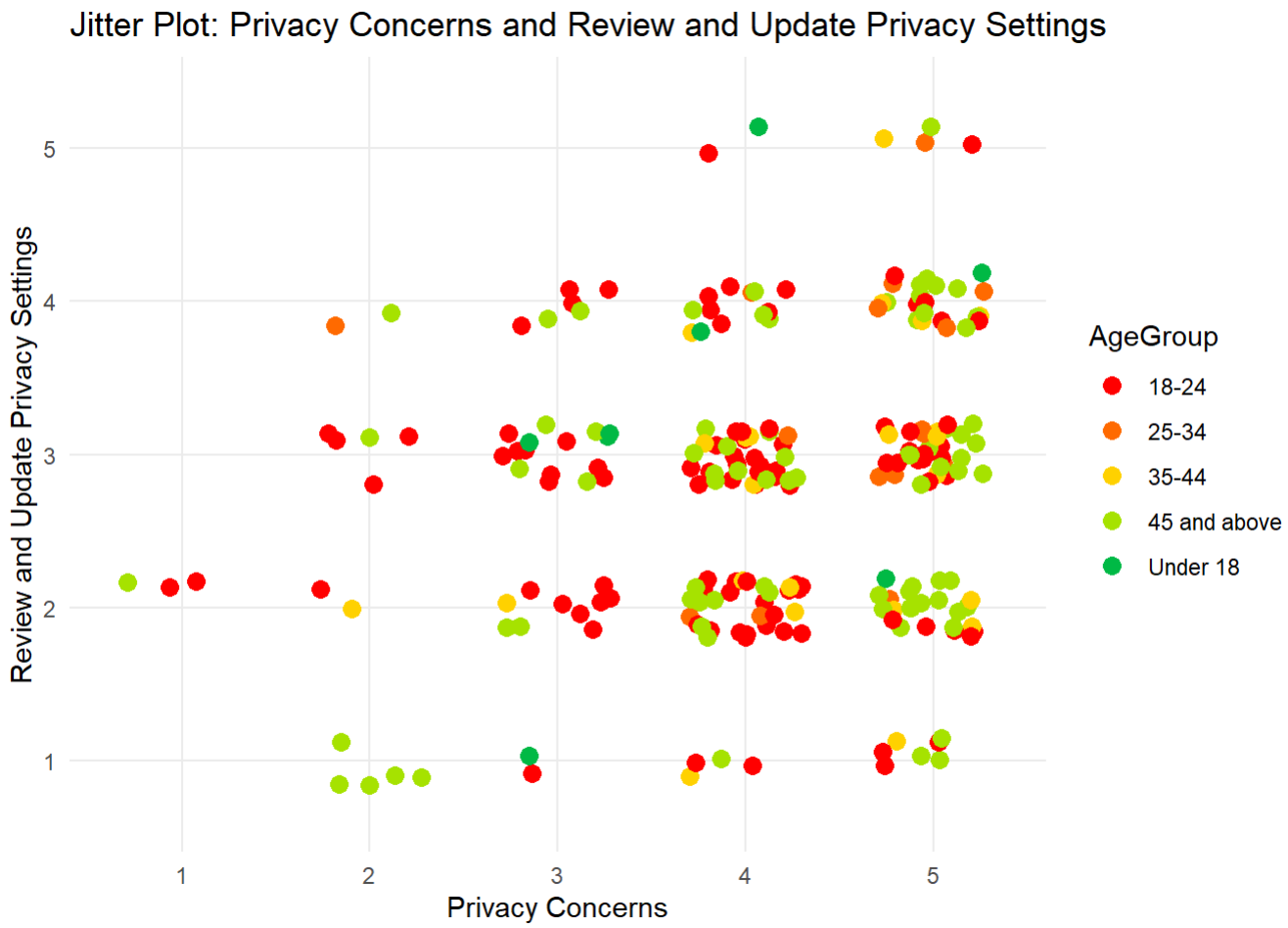
```
#Comparing Hours spent a day and Mental Health
MentalHealthScale<-factor(data$MHAF)
ggplot(data, aes(x = Gender, y = HRSPD, fill = MentalHealthScale)) +
  geom_boxplot() +
  labs(title = "Box Plot: Hours Spent a Day and Mental Health by Gender",
       x = "Gender",
       y = "Hours Spend per day") +
  scale_fill_manual(values = pos_neg_pallete) +
  theme_minimal()
```




```
print("It could be seen that both male and female use social media around the
      duration of 2-3 hours a day. but within that males who spent lesser than
      3hours felt that the mental health had a a positive impact, while majority
      of the data that is number of repondents with diffrent mental health effects
      are alternating as the medians of the diffrent scales not in same level . but it could
be seen that
      The same Gender of Male has a higher toll on mental health affecting negatively.
      ")
```

```
## [1] "It could be seen that both male and female use social media around the \n      durati
on of 2-3 hours a day. but within that males who spent lesser than\n      3hours felt that th
e mental health had a a positive impact, while majority\n      of the data that is number of
repondents with diffrent mental health effects\n      are alternating as the medians of the d
iffrent scales not in same level . but it could be seen that \n      The same Gender of Male
has a higher toll on mental health affecting negatively.\n      "
```

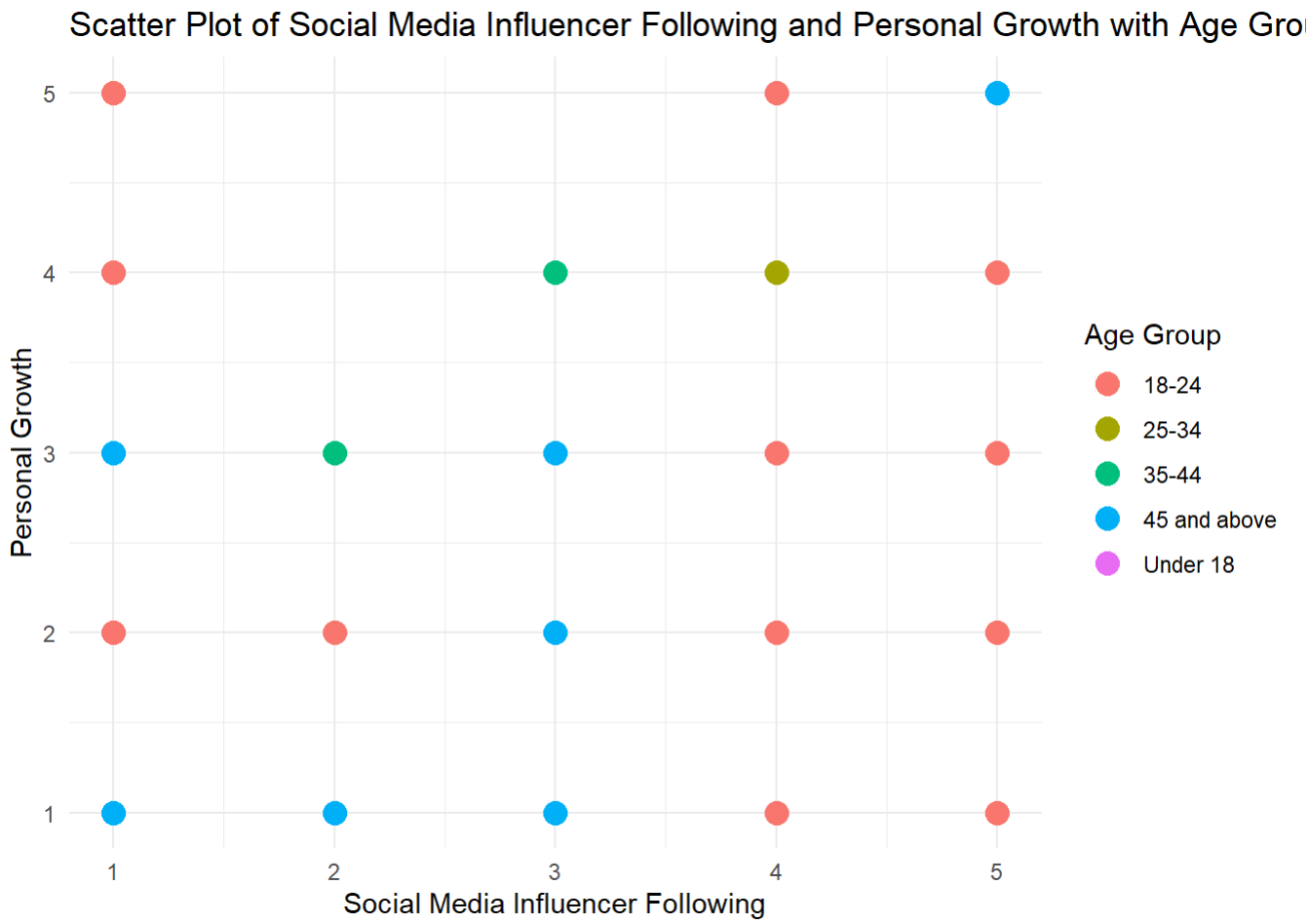
```
#Comparing Privacy Concerns and Privacy Setting Updation
ggplot(data, aes(x = factor(PC), y = factor(RPS), color = AgeGroup)) +
  geom_jitter(position = position_jitter(width = 0.3, height = 0.2), size = 3) +
  labs(title = "Jitter Plot: Privacy Concerns and Review and Update Privacy Settings",
       x = "Privacy Concerns",
       y = "Review and Update Privacy Settings") +
  scale_color_manual(values = pos_neg_pallete)+
  theme_minimal()
```




```
print("The data shows us that the Age of 18-24 have the maximum
      chances of changing their privacy settings as they are
      conserned about thesere privacy, As most of the data collected was from that group
      they seem to be present in all categories.the inference is that people who are
      conserned have or will change the settings ")
```

```
## [1] "The data shows us that the Age of 18-24 have the maximum\n      chances of changing t
heir privacy settings as they are \n      conserned about thesere privacy, As most of the dat
a collected was from that group \n      they seem to be present in all categories.the inferen
ce is that people who are \n      conserned have or will change the settings "
```

```
#Influencer Following and Personal Growth
  ggplot(data, aes(x = SMIF, y = PG, color = AgeGroup)) +
    geom_point(size=4) +
    labs(title = "Scatter Plot of Social Media Influencer Following and Personal Growth with
Age Group",
         x = "Social Media Influencer Following",
         y = "Personal Growth",
         color = "Age Group") +
    theme_minimal()
```




```
print("This visualization shows us that the age group of 18- 24 are following
      more social media  influencers and they have a mixed opinion on personal growth.But
      the age group of 35-44 have a positive personal growth if they followed influencers
      the age group of 45 and above didnt feel much of any improvement in personal growth
      and they belong to the category of people who follow very less influencers ")
```

```
## [1] "This visualization shows us that the age group of 18- 24 are following\n      more so
cial media  influencers and they have a mixed opinion on personal growth.But\n      the age g
roup of 35-44 have a positive personal growth if they followed influencers\n      the age gro
up of 45 and above didnt feel much of any improvement in personal growth\n      and they belo
ng to the category of people who follow very less influencers "
```

```
#This study was a small part of the bigger study conducted on the impact of social media
#which produced a very robust dataset with verity of datas .the aim was to find how social me
dia affected ones life
#and on what aspects they made changes and also on different demographics.
#the main take away is that the negative or positive is not significantly predictable as both
have equal weightage in this data
#but it is for sure that the new life style with internet and how it affects our mental healt
h is a real thing and
#how internet personalities or influencers shape the lifes of people and how social media has
bought a change in
#how we interact with each other . the Time we spend on social media in higher.Everyone are c
oncerned about their
#digital privacy but yet still only the people who are very conserned are doin somethings on
their side to fix it
```