

Social Media Impact on Education

Introduction

Social media's influence on education is a vital area of study, exploring its effects on student behavior and academic engagement.

```
library(wesanderson)
library(ggplot2)
library(dplyr)
```

Warning: package 'dplyr' was built under R version 4.2.3

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(ggalt)
```

Registered S3 methods overwritten by 'ggalt':

method	from
grid.draw.absoluteGrob	ggplot2
grobHeight.absoluteGrob	ggplot2
grobWidth.absoluteGrob	ggplot2

```
grobX.absoluteGrob      ggplot2
grobY.absoluteGrob      ggplot2
```

```
library(tidyverse)
```

Warning: package 'stringr' was built under R version 4.2.3

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
```

```
v forcats   1.0.0      v stringr   1.5.1
v lubridate 1.9.3      v tibble    3.2.1
v purrr     1.0.2      v tidyr     1.3.0
v readr     2.1.4
```

```
-- Conflicts ----- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
```

```
x dplyr::lag()     masks stats::lag()
```

```
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
library(mosaic)
```

Warning: package 'mosaic' was built under R version 4.2.3

Registered S3 method overwritten by 'mosaic':

```
method          from
fortify.SpatialPolygonsDataFrame ggplot2
```

The 'mosaic' package masks several functions from core packages in order to add additional features. The original behavior of these functions should not be affected by this.

Attaching package: 'mosaic'

The following object is masked from 'package:Matrix':

```
mean
```

The following object is masked from 'package:purrr':

```
cross
```

The following objects are masked from 'package:dplyr':

count, do, tally

The following object is masked from 'package:ggplot2':

stat

The following objects are masked from 'package:stats':

binom.test, cor, cor.test, cov, fivenum, IQR, median, prop.test,
quantile, sd, t.test, var

The following objects are masked from 'package:base':

max, mean, min, prod, range, sample, sum

```
library("ggpubr")  
library(latticeExtra)
```

Attaching package: 'latticeExtra'

The following object is masked from 'package:ggplot2':

layer

```
library(sysfonts)  
library("showtext")
```

Loading required package: showtextdb

```
library(PairedData)
```

Loading required package: MASS

Attaching package: 'MASS'

The following object is masked from 'package:dplyr':

select

Loading required package: gld

Loading required package: mvtnorm

Warning: package 'mvtnorm' was built under R version 4.2.3

Attaching package: 'PairedData'

The following object is masked from 'package:Matrix':

summary

The following object is masked from 'package:base':

summary

```
library(rstatix)
```

Attaching package: 'rstatix'

The following object is masked from 'package:MASS':

select

The following objects are masked from 'package:mosaic':

cor_test, prop_test, t_test

The following object is masked from 'package:stats':

filter

```
library(XML)
```

Warning: package 'XML' was built under R version 4.2.3

```
library(reshape2)
```

Attaching package: 'reshape2'

The following object is masked from 'package:tidyr':

```
smiths
```

```
library(pyramid)
library(cowplot)
```

Attaching package: 'cowplot'

The following object is masked from 'package:ggpubr':

```
get_legend
```

The following object is masked from 'package:mosaic':

```
theme_map
```

The following object is masked from 'package:lubridate':

```
stamp
```

Methods

We conducted an experimental survey via Qualtrics and mTurk, examining students' interaction with educational content.

```
blue <- "#0171CE"
red <- "#DE4433"
font_add_google("Lato")
showtext_auto(enable = TRUE, record = TRUE)

demo<-read.csv2(file="SocialMedia_Cleaned.csv", sep = ",")
prepost<-read.csv2(file="prepost.csv", sep = ",")
```

Results

Analysis of 229 respondents revealed insights into age, gender, internet usage, and educational levels impacting social media engagement.

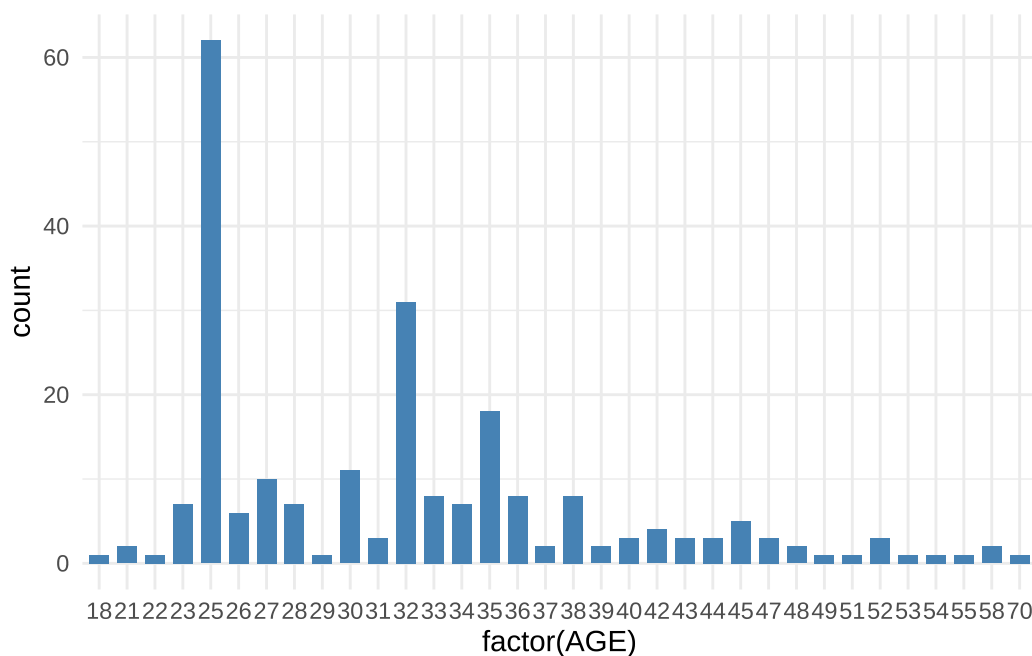
#—————Age & Gender—————

```
favstats(demo$AGE)
```

min	Q1	median	Q3	max	mean	sd	n	missing
18	25	32	35	70	32.09607	8.126167	229	0

```
demo$AGE<-as.numeric(demo$AGE)
```

```
ggplot(demo, aes(x=factor(AGE)))+  
  geom_bar(width=0.7, fill="steelblue")+  
  theme_minimal()
```

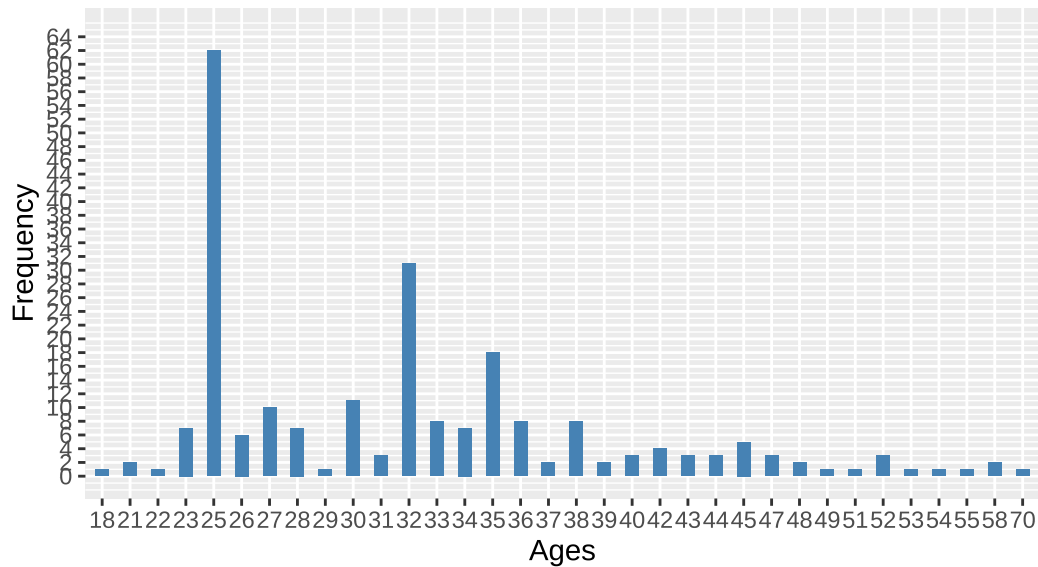


```
ggplot(demo, aes(x=as.factor(AGE), fill=as.factor(AGE) )) +  
  geom_bar(width=0.5, fill = "steelblue") +  
  ggtitle("Participant Ages\nMean = 32.06; Standard Deviation = 8.12") +
```

```
xlab("Ages") + ylab("Frequency")+
scale_y_continuous(limits = c(0, 65), breaks = seq(0, 65, by = 2))
```

Participant Ages

Mean = 32.06; Standard Deviation = 8.12



```
demo %>% count(GENDER_1)
```

	GENDER_1	n
1	Female	74
2	Male	150
3	Non-Binary	4
4	Other (please specify)	1

```
demo %>% group_by(GENDER_1) %>% count(AGE)
```

```
# A tibble: 57 x 3
# Groups:   GENDER_1 [4]
  GENDER_1 AGE      n
  <chr>    <dbl> <int>
1 Female    21         1
2 Female    23         2
```

```

3 Female      25      21
4 Female      26       3
5 Female      27       5
6 Female      28       5
7 Female      29       1
8 Female      30       5
9 Female      31       1
10 Female     32       3
# i 47 more rows

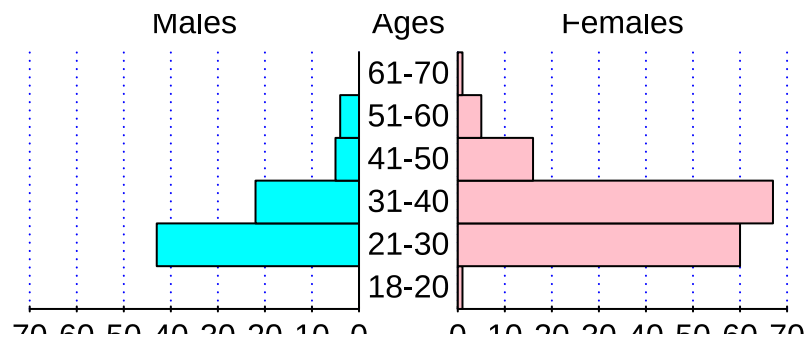
```

```

# ag_2<-demo %>%
#   group_by(GENDER_1) %>%
#   count(age_group)

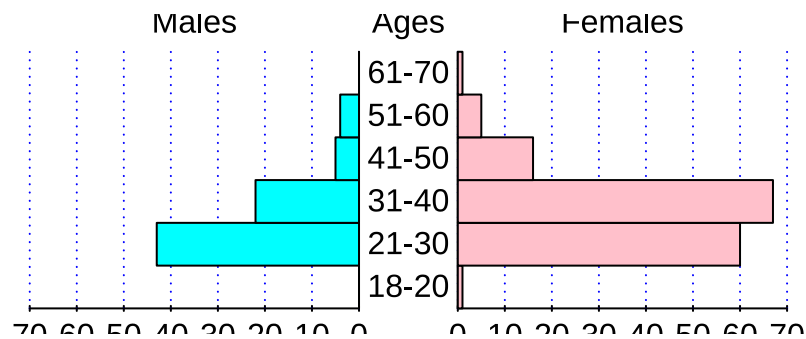
demo["age_group"] = cut(demo$AGE, c(18, 20, 30, 40, 50, 60, 70, Inf), c("18-20", "21-30",
ages <- c("18-20", "21-30", "31-40", "41-50", "51-60", "61-70")
males <- c(0,43,22,5,4,0)
females <- c(1,60,67,16,5,1)
data <- data.frame(males,females,ages)
pyramid(data)

```




```
age<-pyramids(Left=males, Llab="Males", Right=females, Rlab="Females", Center = ages,
  Laxis=c(0,10,20,30,40,50,60,70), main="Participants Age - Mean = 32.06; Standard
```

Participants Age - Mean = 32.06; Standard Deviation = 8.



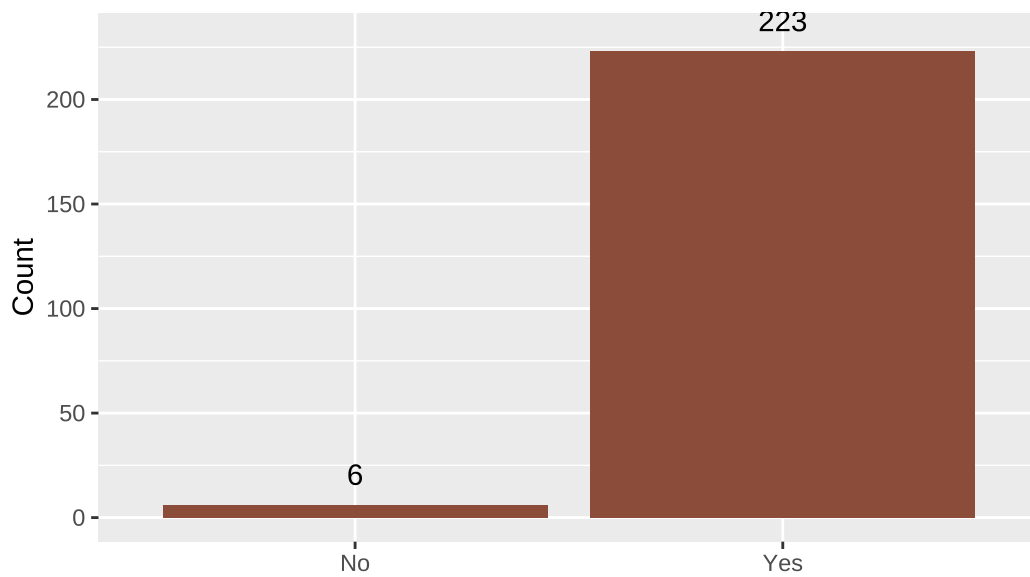
#-----Internet to study-----

```
demo$IN_EDU<-as.factor(demo$IN_EDU)

edu<-demo %>%
  count(IN_EDU)

ggplot(edu, aes(x=IN_EDU, y=n)) +
  geom_bar(stat="identity", fill = "salmon4") +
  ggtitle("Respondents that use internet to study") +
  geom_text(aes(label=n, vjust=-1))+
  xlab("") + ylab("Count")+
  scale_y_continuous(limits = c(0, 230), breaks = seq(0, 230, by = 50))
```

Respondents that use internet to study



#-----Income-----

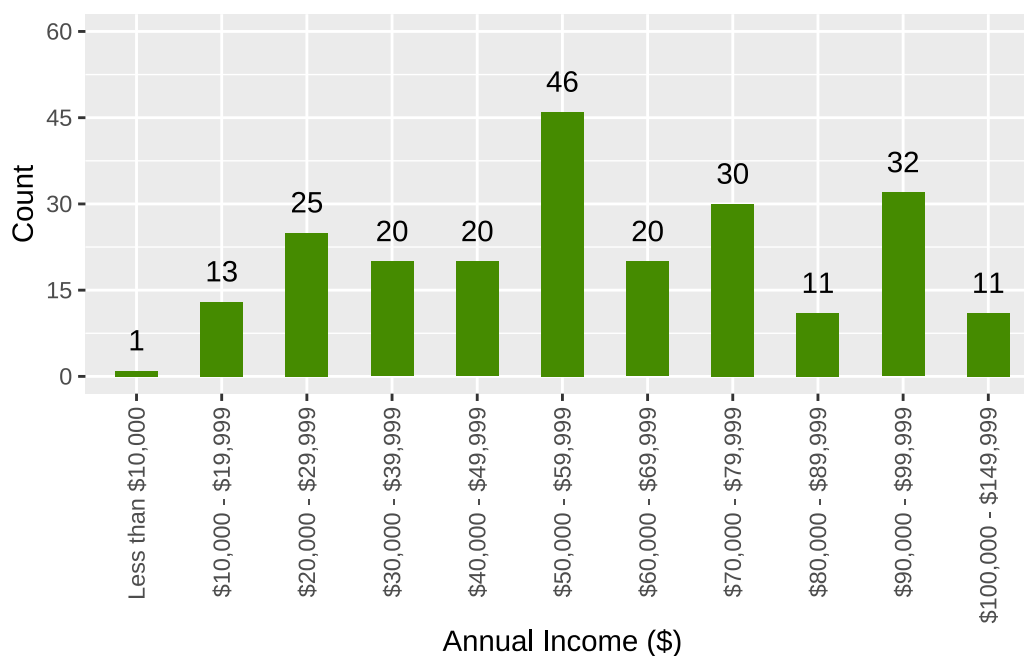
```
demo$INCOME<-as.factor(demo$INCOME)
demo %>% count(INCOME)
```

	INCOME	n
1	\$10,000 - \$19,999	13
2	\$100,000 - \$149,999	11
3	\$20,000 - \$29,999	25
4	\$30,000 - \$39,999	20
5	\$40,000 - \$49,999	20
6	\$50,000 - \$59,999	46
7	\$60,000 - \$69,999	20
8	\$70,000 - \$79,999	30
9	\$80,000 - \$89,999	11
10	\$90,000 - \$99,999	32
11	Less than \$10,000	1

```
demo$INCOME <- factor(demo$INCOME, levels=c("Less than $10,000", "$10,000 - $19,999", "$20,000 - $29,999", "$30,000 - $39,999", "$40,000 - $49,999", "$50,000 - $59,999", "$60,000 - $69,999", "$70,000 - $79,999", "$80,000 - $89,999", "$90,000 - $99,999", "$100,000 - $149,999"))

ggplot(demo, aes(x=INCOME, fill=INCOME)) +
```

```
geom_bar(width=0.5, fill = "chartreuse4") +
#ggtitle("Annual Income ($)") +
xlab("Annual Income ($)") + ylab("Count")+ geom_text(stat='count', aes(label=after_stat(
scale_y_continuous(limits = c(0, 60), breaks = seq(0, 60, by = 15))+
theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))+
scale_x_discrete(limits = c("Less than $10,000", "$10,000 - $19,999", "$20,000 - $29,999",
```



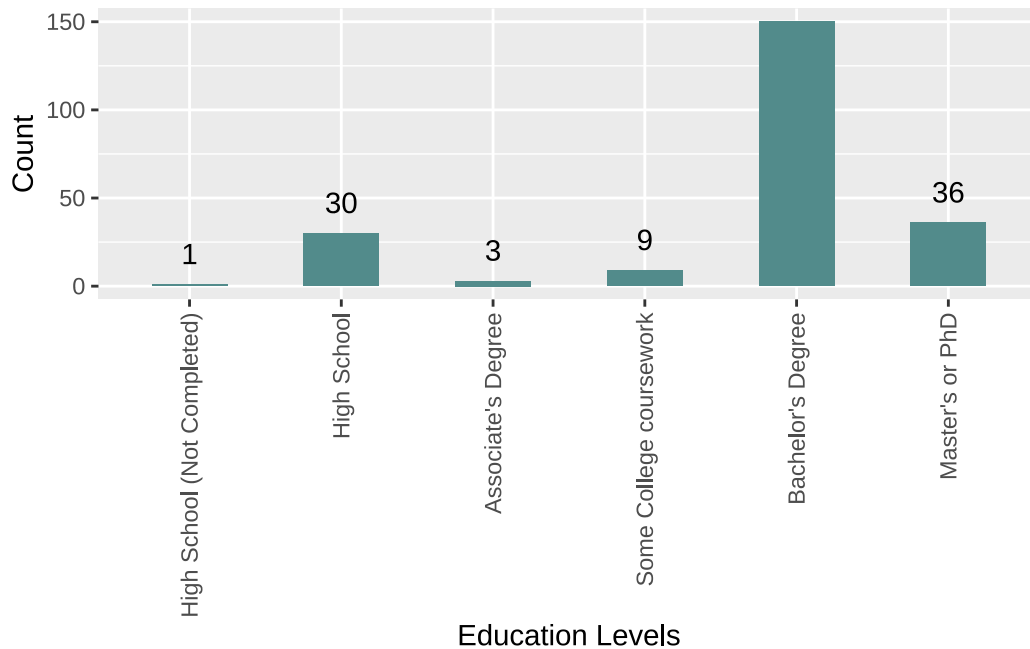
#-----Education-----

```
demo %>% count(EDU_1)
```

	EDU_1	n
1	Associate's Degree	3
2	Bachelor's Degree	150
3	High School	30
4	High School (Not Completed)	1
5	Master's or PhD	36
6	Some College coursework	9

```
demo$EDU_1 <- factor(demo$EDU_1, levels=c("High School (Not Completed)", "High School", "A

ggplot(demo, aes(x=EDU_1, fill=EDU_1)) +
  geom_bar(width=0.5, fill = "darkslategray4") +
  #ggtitle("Annual Income ($)") +
  xlab("Education Levels") + ylab("Count")+ geom_text(stat='count', aes(label=after_stat(c
  scale_y_continuous(limits = c(0, 150), breaks = seq(0, 150, by = 50))+
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



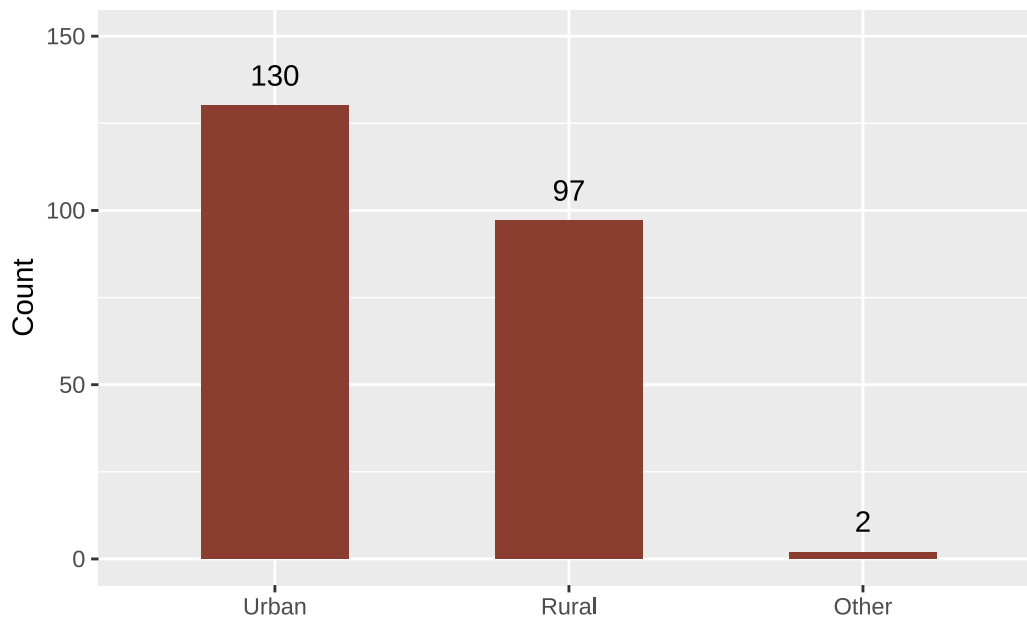
#-----Living Environment-----

```
demo %>% count(LOC_1)
```

LOC_1	n
1 Other	2
2 Rural	97
3 Urban	130

```
demo$LOC_1 <- factor(demo$LOC_1, levels=c("Urban", "Rural", "Other"))
```

```
ggplot(demo, aes(x=LOC_1, fill=LOC_1)) +
  geom_bar(width=0.5, fill = "coral4") +
  #ggtitle("Annual Income ($)") +
  xlab("") + ylab("Count")+ geom_text(stat='count', aes(label=after_stat(count)), vjust=-1)
  scale_y_continuous(limits = c(0, 150), breaks = seq(0, 150, by = 50))
```



#-----Hours spent on internet-----

```
favstats(demo$HOURS)
```

	min	Q1	median	Q3	max	mean	sd	n	missing
1	3	4	5	6	3.659389	1.066706	229	0	

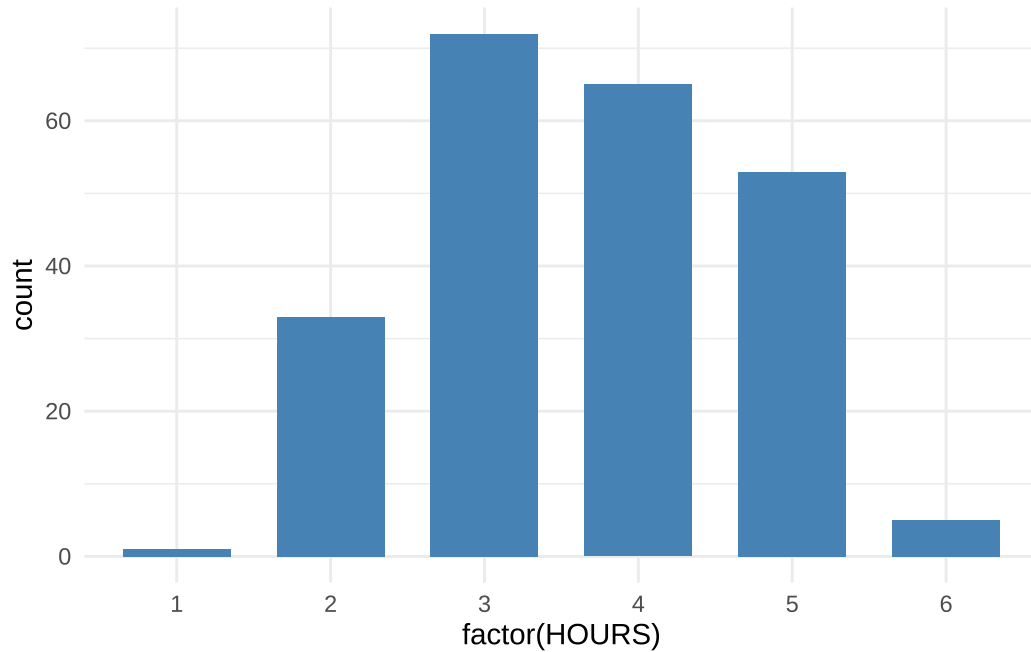
```
demo %>% count(HOURS)
```

	HOURS	n
1	1	1
2	2	33
3	3	72
4	4	65

```
5      5 53
6      6 5
```

```
demo$AGE<-as.numeric(demo$HOURS)

ggplot(demo, aes(x=factor(HOURS)))+
  geom_bar(width=0.7, fill="steelblue")+
  theme_minimal()
```



```
#-----PRE-POST-----
```

```
favstats(demo$PRE)
```

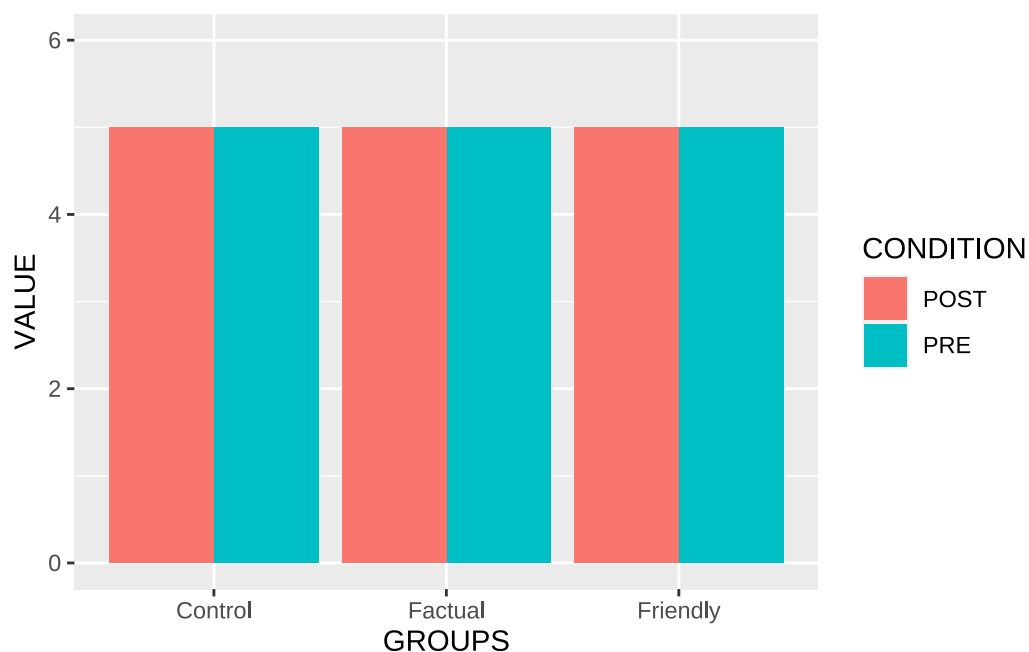
min	Q1	median	Q3	max	mean	sd	n	missing
1	4	4	4	5	3.855895	0.9556952	229	0

```
favstats(demo$POST)
```

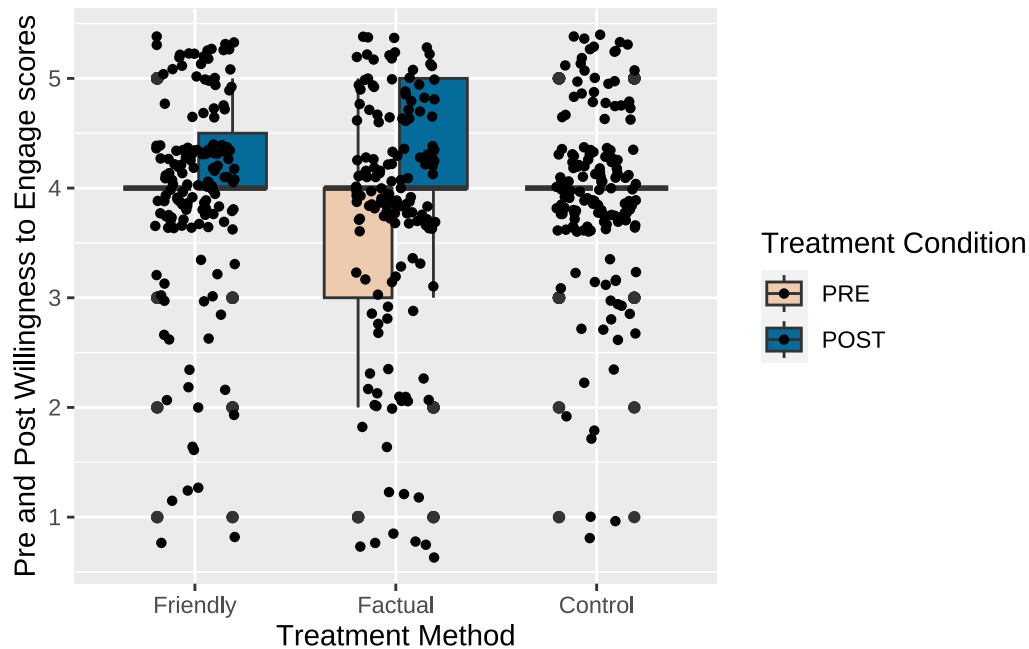
min	Q1	median	Q3	max	mean	sd	n	missing
1	4	4	4	5	3.938865	0.9392609	229	0

```
#prepost$PID<-as.factor(prepost$PID)
prepost$GROUP<-as.factor(prepost$GROUP)
prepost$VALUE<-as.numeric(prepost$VALUE)
prepost$CONDITION<-as.factor(prepost$CONDITION)

ggplot(prepost, aes(x=GROUPS, y=VALUE, fill=CONDITION)) +
  geom_bar(position = "dodge", stat = "identity")+
  ylim(0,6)
```



```
ggplot(prepost, aes(x=reorder(GROUPS,desc(GROUPS)), y=VALUE, fill=reorder(CONDITION, desc(
  geom_boxplot()+
  geom_jitter(shape=16, position=position_jitter(0.2))+
  scale_fill_manual(values= wes_palette("Darjeeling2"))+
  xlab("Treatment Method")+
  ylab("Pre and Post Willingness to Engage scores")+
  guides(fill=guide_legend(title="Treatment Condition"))
```



```
res_aov<-aov(VALUE ~ CONDITION+GROUPS, data = prepost)
```

```
summary(res_aov)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
CONDITION	1	0.8	0.7882	0.879	0.349
GROUPS	2	2.4	1.1931	1.331	0.265
Residuals	454	407.0	0.8965		

```
ggplot(data = demo,
  aes(
    x = EDU_1,
    y = prop.table(stat(count)),
    fill = factor(IN_EDU), width = -6,
    label = scales::percent(prop.table(stat(count)))
  )) +
  geom_bar(position = position_dodge(), width = 0.4) + theme(axis.text = element_text(size
  geom_text(
    stat = "count",
    position = position_dodge(.8),
    vjust = -1,
```

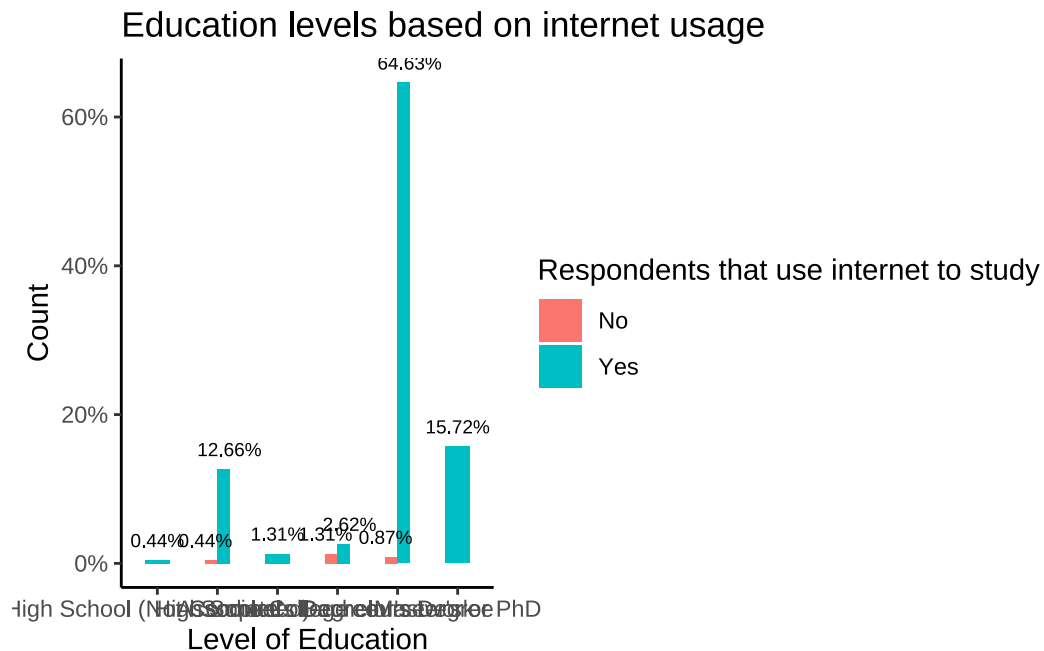


```

    size = 2.5
  ) + scale_y_continuous(labels = scales::percent) +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1)) +
  labs(title = "Education levels based on internet usage",
       x = "Level of Education",
       y = "Count") +
  theme_classic() +
  scale_fill_discrete(
    name = "Respondents that use internet to study",
    labels = c("No", "Yes")
  )
)

```

Warning: `stat(count)` was deprecated in ggplot2 3.4.0.
 i Please use `after_stat(count)` instead.



Discussion

The study's findings emphasize the need for tailored educational strategies considering social media's ubiquitous presence in students' lives.

Conclusion

This research provides valuable insights for educators and policymakers to leverage social media positively in education.

Repository Name: Social-Media-Impact-on-Education-Research