Inferential analysis

Research Question: What are the health risks associated with excessive smartphone and social media use leading to isolation, and how does this impact different demographic groups in terms of social interactions and behaviors?

Findings:

The study conducted by Matar Boumosleh and Daniel Jaalouk (2017) explored the relationship between excessive smartphone usage, addiction metrics, and mental health outcomes among university students.The study found a significant association between excessive smartphone usage and symptoms of depression and anxiety among university students. This suggests that students who reported spending more time on their smartphones were more likely to experience higher levels of depression and anxiety.The researchers used the Smartphone Addiction Inventory (SPAI) Scale to measure smartphone addiction among participants. They found that higher scores on the SPAI Scale, indicating greater levels of smartphone addiction, were correlated with higher levels of depression and anxiety.

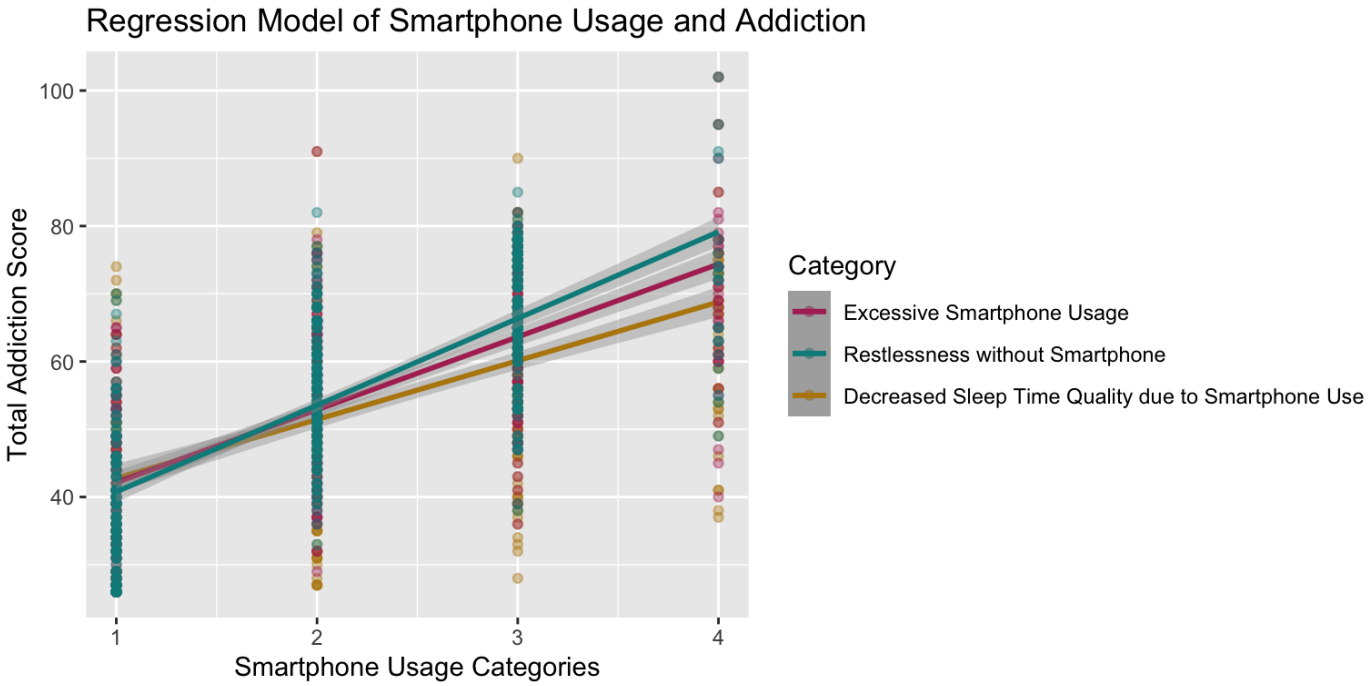
To test this we have a variety of features in our data which we can use to create linear regression models such as excessive smartphone usage, restlessness symptoms without smartphone, and decreased sleep time, and quality due to smartphone use.

```
w <- "https://raw.githubusercontent.com/ryanfernald/MATH167R-Final-Project/main/DATA/Depr
wdata <- read.csv(w)
wdata[wdata == "."] <- NA
```

```
ggplot(wdata, aes(y = as.numeric(TotAddiction_Score),
                  x = as.numeric(ExcessveSmrtPhn_Use))) +
  geom_point(aes(color = "maroon"), alpha = 0.4) +
  geom_smooth(aes(color = "maroon"), method = "lm") +

  geom_point(aes(x = as.numeric(Rstlss_NoSmrtphn),
                 color = "darkcyan"), alpha = 0.4) +
  geom_smooth(aes(x = as.numeric(Rstlss_NoSmrtphn),
                  color = "darkcyan"), method = "lm") +

  geom_point(aes(x = as.numeric(DcreasdSlpTimeQulty_SmrtPhnUse),
                 color = "darkgoldenrod"), alpha = 0.4) +
  geom_smooth(aes(x = as.numeric(DcreasdSlpTimeQulty_SmrtPhnUse),
                  color = "darkgoldenrod"), method = "lm") +

  labs(
    title = "Regression Model of Smartphone Usage and Addiction",
    x = "Smartphone Usage Categories",
```

```
    y = "Total Addiction Score",
    color = "Category"
  ) +
  scale_color_manual(values = c("maroon", "darkcyan", "darkgoldenrod"),
                     labels = c("Excessive Smartphone Usage",
                                "Restlessness without Smartphone",
                                "Decreased Sleep Time Quality due to Smartphone Use")) +
  theme(legend.position = "right")
```

```
`geom_smooth()` using formula = 'y ~ x'
`geom_smooth()` using formula = 'y ~ x'
`geom_smooth()` using formula = 'y ~ x'
```




```
# lm_model_1 <- lm(as.numeric(TotAddiction_Score) ~
#                  as.numeric(ExcessveSmrtPhn_Use),
#                  data = wdata, na.action = na.exclude)
# summary(lm_model_1)
#
# lm_model_2 <- lm(as.numeric(TotAddiction_Score) ~
#                  as.numeric(Rstlss_NoSmrtphn),
#                  data = wdata, na.action = na.exclude)
# lm_model_2
#
# lm_model_3 <- lm(as.numeric(TotAddiction_Score) ~ as.numeric(DcreasdSlpTimeQulty_SmrtPh
#                  xdata = wdata, na.action = na.exclude)
# lm_model_3

lm_model_combined <- lm(as.numeric(TotAddiction_Score) ~
                        as.numeric(ExcessveSmrtPhn_Use) +
                        as.numeric(Rstlss_NoSmrtphn) +
```

```
                        as.numeric(DcreasdSlpTimeQulty_SmrtPhnUse),
                        data = wdata, na.action = na.exclude)
summary(lm_model_combined)
```

```
Call:
lm(formula = as.numeric(TotAddiction_Score) ~ as.numeric(ExcessveSmrtPhn_Use) +
    as.numeric(Rstlss_NoSmrtphn) + as.numeric(DcreasdSlpTimeQulty_SmrtPhnUse),
    data = wdata, na.action = na.exclude)

Residuals:
     Min       1Q   Median       3Q      Max
-25.4707  -5.0014  -0.0699   4.8364  22.6955

Coefficients:
                                        Estimate Std. Error t value Pr(>|t|)
(Intercept)                              13.6330     1.1687  11.665   <2e-16
as.numeric(ExcessveSmrtPhn_Use)           3.8259     0.3954   9.675   <2e-16
as.numeric(Rstlss_NoSmrtphn)              6.1905     0.4083  15.163   <2e-16
as.numeric(DcreasdSlpTimeQulty_SmrtPhnUse)   8.6597     0.4419  19.599   <2e-16

(Intercept)                             ***
as.numeric(ExcessveSmrtPhn_Use)         ***
as.numeric(Rstlss_NoSmrtphn)            ***
as.numeric(DcreasdSlpTimeQulty_SmrtPhnUse) ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.341 on 456 degrees of freedom
  (228 observations deleted due to missingness)
Multiple R-squared:  0.7644,    Adjusted R-squared:  0.7628
F-statistic: 493.2 on 3 and 456 DF,  p-value: < 2.2e-16
```

Why we chose this model

We chose a linear regression model because our data has categorical data represented on a scale from 1 to 4 where 1 is the least extreme, and 4 is the most. We can interpret this as higher levels of X_i leads to higher levels in our Response Variable Y.

Given our response Y and our predictors $X_1, \ldots, X_p$ where,

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \epsilon$ and,

$\epsilon \sim N(O, \sigma^2)$

β_0 is the intercept. β_1 is ExcessveSmrtPhn_Use β_2 is Rstlss_NoSmrtphn β_3 is DcreasdSlpTimeQulty_SmrtPhnUse

What is the model, What x is and y

The model is a linear regression model that predicts the total addiction score (y) based on one or more independent variables (x), such as excessive smartphone usage, restlessness without smartphone, and decreased sleep time quality due to smartphone use.

X is our Independent Variables, where Y is our Dependent Variable, or Response Variable.

Interpretation of the Parameters

β_0, the intercept is 13.8067, meaning when our β_i's are 0 we expect the value of our Total addiction score to be 13.8. β_1, ExcessveSmrtPhn_Use, has a slope coefficients of 3.2675. β_2, Rstlss_NoSmrtphn, has a slope coefficients of 6.2581. β_3, DcreasdSlpTimeQulty_SmrtPhnUse, has a slope coefficients of 9.1674.

Hypothesis Testing

For each of our variables, we can get the information from hypothesis testing form the summary.

Each of our factors is being tested against the dependent variable of Total addiction score, and we can set up the H_0 and H_A as the following: H_A^* = H_0^*, means our independent variables have no effect on the dependent variable. $H_A^* \neq H_0^*$, means our independent variables have a statistically significant effect on the dependent variable. For each of our tests we have the following result:

1. β_1, ExcessveSmrtPhn_Use has a p value of 6.65e-09, and thus we can reject H_0 in favor of H_A.
2. β_2, Rstlss_NoSmrtphn, has a p value of < 2e-16, and thus we can reject H_0 in favor of H_A.
3. β_3, DcreasdSlpTimeQulty_SmrtPhnUse, has a p value of < 2e-16, and thus we can reject H_0 in favor of H_A.
4. Finally the p value for our combined linear regression model with each parameter is < 2.2e-16, and thus we can reject H_0 in favor of H_A.

Refrences

Matar Boumosleh J, Jaalouk D (2017) Depression, anxiety, and smartphone addiction in university students- A cross sectional study. PLOS ONE 12(8): e0182239. https://doi.org/10.1371/journal.pone.0182239

Simó-Sanz C, Ballestar-Tarín MªL, Martínez-Sabater A. Smartphone Addiction Inventory (SPAI): Translation, adaptation and validation of the tool in Spanish adult population. PLoS One. 2018 Oct 17;13(10):e0205389. doi: 10.1371/journal.pone.0205389. PMID: 30332481; PMCID: PMC6192628.