R Notebook

Loading of Data

```
library(tidyverse)
beauty <- read_csv("beauty.csv")
```

Filtering to only female respondents:

```
beauty_data <- beauty %>% filter(gender == "Female")
```

Demographic characteristics:

```
fivenum(beauty_data$age)
```

```
## [1] 18 20 20 21 30
```

```
unique(beauty_data$faculty)
```

```
## [1] "Business" "CHS"      "CDE"      "Law"      "SOC"      "Medicine"
```

```
fivenum(beauty_data$tiktok_hours)
```

```
## [1] 0.0 0.5 1.5 2.0 6.0
```

```
beauty_data %>% mutate(beauty_index = tiktok_hours*frequency) -> beauty_data
```

T-tests between IV and DVs

First, split respondents into two groups: those who do not consume beauty content on Tiktok (ie. Watched Hours = 0 + Tiktoks related to beauty answer "not at all"), and those who do.

```
beauty_false <- beauty_data %>% filter(tiktok_hours == 0 | frequency == 1)
beauty_true <- setdiff(beauty_data, beauty_false)
```

Then, compare means of dependent variables.

For self-esteem, our hypothesis is that the self-esteems of those who consume beauty content are lower than those who do not. $t = 1.2579$, $p = 0.8907$. Hence, the longer-term self-esteems of those who consume beauty content is not significantly lesser than those who do not.

```r
t.test(beauty_true$high_self_esteem, beauty_false$high_self_esteem, alternative = "less")
```

```
##
##  Welch Two Sample t-test
##
## data:  beauty_true$high_self_esteem and beauty_false$high_self_esteem
## t = 1.2579, df = 28.488, p-value = 0.8907
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf 0.8336102
## sample estimates:
## mean of x mean of y
##  3.259259  2.904762
```

For appearance anxiety, our hypothesis is that those who consume beauty content experience more appearance anxiety than those who do not. t = 1.0299, , p-value = 0.156. Hence, the longer-term appearance anxiety of those who consume beauty content is not significantly greater than those who do not.

```r
t.test(beauty_true$anxiety, beauty_false$anxiety, alternative = "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  beauty_true$anxiety and beauty_false$anxiety
## t = 1.0299, df = 27.477, p-value = 0.156
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  -0.1876606        Inf
## sample estimates:
## mean of x mean of y
##  3.049383  2.761905
```

For mood, our hypothesis is that those who consume beauty content have their moods more negatively affected by such content than those who do not. t = 2.9167, p-value = 0.003428. Hence, the moods of those who consume beauty content are significantly more negatively affected than those who do not.

```r
t.test(beauty_true$negative_mood, beauty_false$negative_mood, alternative = "greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  beauty_true$negative_mood and beauty_false$negative_mood
## t = 2.9167, df = 28.284, p-value = 0.003428
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.5463918       Inf
## sample estimates:
## mean of x mean of y
##  3.691358  2.380952
```

Simple Linear Regression between IV and DVs

Beauty Index on Mood: correlation is estimated to be 0.11257, p-value is 0.0003816, and there is a significant positive correlation.

```
beauty_mood <- lm(negative_mood ~ beauty_index, data = beauty_data)
summary(beauty_mood)
```

```
##
## Call:
## lm(formula = negative_mood ~ beauty_index, data = beauty_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.2078 -1.3990 -0.2912  1.2431  3.7088
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.85986    0.24223  11.806  < 2e-16 ***
## beauty_index  0.10784    0.03372   3.198  0.00185 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.685 on 100 degrees of freedom
## Multiple R-squared:  0.09279,    Adjusted R-squared:  0.08372
## F-statistic: 10.23 on 1 and 100 DF,  p-value: 0.001853
```

Beauty Index on Anxiety: correlation is estimated to be 0.04991, p-value is 0.0137, and there is a non-significant positive correlation.

```
beauty_anxiety <- lm(anxiety ~ beauty_index, data = beauty_data)
summary(beauty_anxiety)
```

```
##
## Call:
## lm(formula = anxiety ~ beauty_index, data = beauty_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.35412 -0.82380  0.07101  0.87049  2.26979
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.73021    0.14289  19.107   <2e-16 ***
## beauty_index  0.04991    0.01989   2.509   0.0137 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9938 on 100 degrees of freedom
## Multiple R-squared:  0.05924,    Adjusted R-squared:  0.04983
## F-statistic: 6.297 on 1 and 100 DF,  p-value: 0.0137
```

Beauty Index on Self-esteem: correlation is estimated to be 0.1155, p-value is 0.592, and there is a extremely non-significant positive correlation.

```r
beauty_esteem <- lm(high_self_esteem ~ beauty_index, data = beauty_data)
summary(beauty_esteem)
```

```
##
## Call:
## lm(formula = high_self_esteem ~ beauty_index, data = beauty_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.2994 -0.9223 -0.1261  0.8277  1.8623
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3.12612    0.15427  20.264   <2e-16 ***
## beauty_index  0.01155    0.02147   0.538    0.592
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.073 on 100 degrees of freedom
## Multiple R-squared:  0.002884,   Adjusted R-squared:  -0.007087
## F-statistic: 0.2892 on 1 and 100 DF,  p-value: 0.5919
```

Confounders

We have identified consumption of beauty content from other platforms outside of Tiktok as a possible confounder for the correlations between our IV and DVs. Our confounder hence becomes the variable of the number of other platforms used outside of Tiktok, found through mutating a new column from the data in beauty_data$platforms.

```r
mutate(beauty_data, other_platforms = str_remove_all(beauty_data$platforms, "Tiktok/ Douyin;")) -> beau
mutate(beauty_data, other_platforms = str_remove_all(beauty_data$other_platforms, "Tiktok/ Douyin")) ->
mutate(beauty_data, other_platforms = str_count(beauty_data$other_platforms, ";")) -> beauty_data
```

The effect of consuming beauty Tiktok content on appearance anxiety, when controlled for consumption of content from other platforms, shows no change in terms of significance. Estimated correlation = 0.05009, p = 0.04532.

```r
beauty_anxiety_conf <- lm(anxiety ~ beauty_index + other_platforms, data = beauty_data)
summary(beauty_anxiety_conf)
```

```
##
## Call:
## lm(formula = anxiety ~ beauty_index + other_platforms, data = beauty_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.35195 -0.81773  0.06876  0.87344  2.27413
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.66747    0.21959  12.147   <2e-16 ***
```

```
## beauty_index     0.05009    0.01998   2.507   0.0138 *
## other_platforms  0.05840    0.15471   0.377   0.7066
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9981 on 99 degrees of freedom
## Multiple R-squared:  0.06059,    Adjusted R-squared:  0.04161
## F-statistic: 3.193 on 2 and 99 DF,  p-value: 0.04532
```

However, when examining the effect of consuming beauty Tiktok content on self-esteem, when controlled for consumption of content from other platforms, shows that while beauty Tiktok content does not have a significant effect, consumption of content from other platforms indeed has a significant effect at a p = 0.05 level.

```
beauty_esteem_conf <- lm(high_self_esteem ~ beauty_index + other_platforms, data = beauty_data)
summary(beauty_esteem_conf)
```

```
##
## Call:
## lm(formula = high_self_esteem ~ beauty_index + other_platforms,
##     data = beauty_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2.6716 -0.7129  0.1972  0.8263  1.8339
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.71396    0.23085  11.756   <2e-16 ***
## beauty_index     0.01269    0.02101   0.604   0.5470
## other_platforms  0.38363    0.16264   2.359   0.0203 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.049 on 99 degrees of freedom
## Multiple R-squared:  0.05594,    Adjusted R-squared:  0.03687
## F-statistic: 2.933 on 2 and 99 DF,  p-value: 0.05787
```

We do not repeat the test for confounders on the single linear regression model between beauty Tiktok consumption and mood because the question for that set of observations surveys for a cause and effect relationship between Tiktok consumption specifically and mood ("negative changes in your mood (e.g., increased sadness or self consciousness) *after* watching beauty TikTok video") for non-neutral answers.

Conclusions from t-tests and linear regressions: