Analytics Engineer Challange

Vincenzo Grasso - vincenzo.grass@gmail.com

1/27/2022

A - Message Conversion Rates

This is the typical example of Simpon's Paradox aka a phenomenon in in which a trend appears in groups of data but disappears or reverses when the groups are combined.

```
library(tidyverse)

y <- rep(0:1, c(4, 4))
trt <- rep(c("A", "B"), 4)
platform <- rep(c("IOS", "IOS", "Android", "Android"), 2)
all_cnt <- c(174000, 540000, 526000, 160000)
subscriptions <- c(16200, 46800, 38400, 11000)
non_subscriptions <- all_cnt - subscriptions

weight <- c(non_subscriptions, subscriptions)

df <- tibble(
  y,
  trt = factor(trt, levels = c("B", "A")),
  platform,
  weight
)

print(df)
```

```
## # A tibble: 8 x 4
##       y trt   platform weight
##   <int> <fct> <chr>     <dbl>
## 1     0 A     IOS      157800
## 2     0 B     IOS      493200
## 3     0 A     Android  487600
## 4     0 B     Android  149000
## 5     1 A     IOS       16200
## 6     1 B     IOS       46800
## 7     1 A     Android   38400
## 8     1 B     Android   11000
```

Resulting in:

```
df_long <- df %>%
  rowwise() %>%
  mutate(ids = list(seq(1, weight))) %>%
  unnest(ids)
```

```r
summary(lm(y ~ trt + platform, df_long))
```

```
##
## Call:
## lm(formula = y ~ trt + platform, data = df_long)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.09231 -0.08692 -0.07327 -0.07327  0.93212
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0678837  0.0005274 128.726   <2e-16 ***
## trtA        0.0053836  0.0005386   9.995   <2e-16 ***
## platformIOS 0.0190396  0.0005387  35.342   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2716 on 1399997 degrees of freedom
## Multiple R-squared:  0.0009621,  Adjusted R-squared:  0.0009606
## F-statistic: 674.1 on 2 and 1399997 DF,  p-value: < 2.2e-16
```

Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.