Code

* Show All Code
* Hide All Code
* Download Rmd

# WNBA Self-Presentation Data Analysis

#### September 7, 2021

Let’s start by loading in the data

── Column specification ─────────────────────────────────────────────────────────────────────────

cols(

Name = col\_character(),

`Last Name` = col\_character(),

`First Name` = col\_character(),

Position = col\_character(),

`Height (cm.)` = col\_double(),

`Weight (kg)` = col\_double(),

Birthday = col\_character(),

Age = col\_double(),

Nationality = col\_character(),

`Domestic/International` = col\_character(),

College = col\_character(),

`Current WNBA Team (or Last)` = col\_character(),

`Draft Pick` = col\_double(),

`Draft Year` = col\_double(),

`Year Started` = col\_double(),

`Seasons in League` = col\_double(),

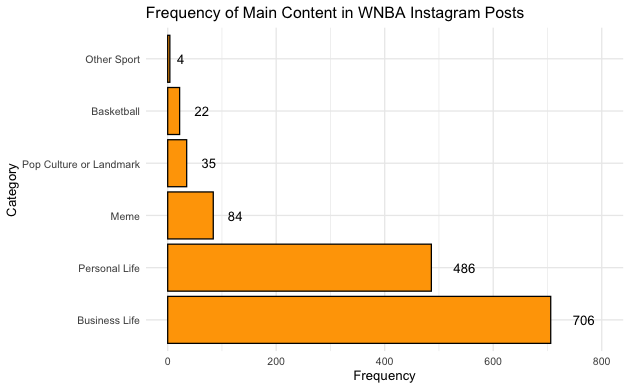
`Active/2021` = col\_character(),

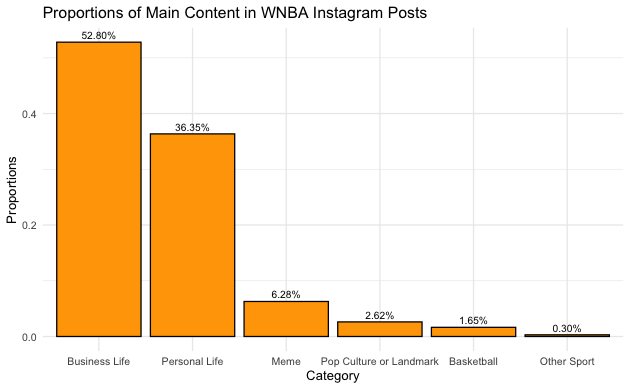
`Played in 2020?` = col\_character(),

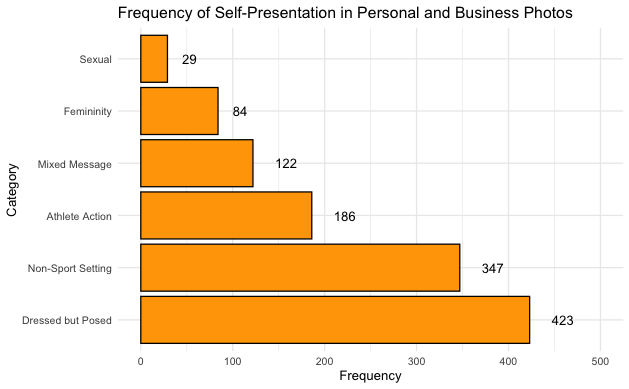
Use = col\_double()

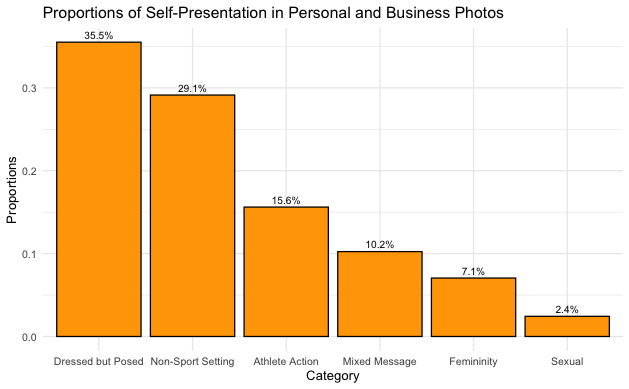
)

Warning: NAs introduced by coercion









# (Pegoraro et al., 2018)

# THIS IS THE FUN STUFF

# Engagement between Step two photos and likes (ANOVA), bivariate correlation

# between likes and type of photos

Step.2.anova <-

summary(aov(Likes ~ Step.2, data = Step\_2\_merged))

Step.2.anova

Df Sum Sq Mean Sq F value Pr(>F)

Step.2 5 1.336e+11 2.672e+10 41.45 <2e-16 \*\*\*

Residuals 1179 7.600e+11 6.446e+08

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

6 observations deleted due to missingness

# Tukey comparison

Step.2.tukey <- TukeyHSD((aov(Likes ~ Step.2, data = Step\_2\_merged)))

Step.2.tukey

Tukey multiple comparisons of means

95% family-wise confidence level

Fit: aov(formula = Likes ~ Step.2, data = Step\_2\_merged)

$Step.2

diff lwr upr p adj

Dressed but Posed-Femininity -580.7839 -9511.083 8349.515 0.9999696

Mixed Message-Femininity -1930.7713 -12437.134 8575.591 0.9952079

Athlete Action-Femininity 1070.3085 -8705.728 10846.345 0.9996034

Non-Sport Setting-Femininity -7095.9833 -16177.263 1985.296 0.2245246

Sexual-Femininity 63248.7485 47486.789 79010.708 0.0000000

Mixed Message-Dressed but Posed -1349.9875 -8797.526 6097.551 0.9955074

Athlete Action-Dressed but Posed 1651.0924 -4724.883 8027.067 0.9770143

Non-Sport Setting-Dressed but Posed -6515.1995 -11764.181 -1266.218 0.0054922

Sexual-Dressed but Posed 63829.5323 49918.331 77740.734 0.0000000

Athlete Action-Mixed Message 3001.0799 -5442.052 11444.212 0.9131705

Non-Sport Setting-Mixed Message -5165.2120 -12793.136 2462.712 0.3824789

Sexual-Mixed Message 65179.5198 50207.715 80151.325 0.0000000

Non-Sport Setting-Athlete Action -8166.2919 -14752.069 -1580.515 0.0055627

Sexual-Athlete Action 62178.4399 47709.778 76647.102 0.0000000

Sexual-Non-Sport Setting 70344.7318 56336.130 84353.334 0.0000000

# Running a linear model

lm.model <- lm(Step\_2\_merged$Likes ~ Step\_2\_merged$Step.2)

summary(lm.model)

Call:

lm(formula = Step\_2\_merged$Likes ~ Step\_2\_merged$Step.2)

Residuals:

Min 1Q Median 3Q Max

-72625 -9265 -4671 -1098 174826

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 11930.3 2874.8 4.150 3.56e-05 \*\*\*

Step\_2\_merged$Step.2Dressed but Posed -580.8 3128.6 -0.186 0.8528

Step\_2\_merged$Step.2Mixed Message -1930.8 3680.8 -0.525 0.6000

Step\_2\_merged$Step.2Athlete Action 1070.3 3424.9 0.313 0.7547

Step\_2\_merged$Step.2Non-Sport Setting -7096.0 3181.5 -2.230 0.0259 \*

Step\_2\_merged$Step.2Sexual 63248.7 5522.0 11.454 < 2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 25390 on 1179 degrees of freedom

(6 observations deleted due to missingness)

Multiple R-squared: 0.1495, Adjusted R-squared: 0.1459

F-statistic: 41.45 on 5 and 1179 DF, p-value: < 2.2e-16

# Percent athletes appeared in the photos themselves

# 52.8 business, 36.35 personal, 79.15% photos feature the athletes (INDICATE

# INTERNATIONAL AND DOMESTIC PERCENTAGES)

# t-test to see if when athlete appeared more like. ANOVA checked the number of

# likes to each type of photo (sexually suggestive)

appear\_anova <- merged %>%

drop\_na(Athlete.Appear)

summary(aov(Likes ~ Athlete.Appear, data = appear\_anova))

Df Sum Sq Mean Sq F value Pr(>F)

Athlete.Appear 1 5.197e+09 5.197e+09 7.472 0.00635 \*\*

Residuals 1329 9.244e+11 6.955e+08

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

6 observations deleted due to missingness

# Running a linear model

lm.model\_aa <- lm(appear\_anova$Likes ~ appear\_anova$Athlete.Appear)

summary(lm.model\_aa)

Call:

lm(formula = appear\_anova$Likes ~ appear\_anova$Athlete.Appear)

Residuals:

Min 1Q Median 3Q Max

-11045 -9496 -7981 -2834 175127

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 11153.8 765.8 14.565 < 2e-16 \*\*\*

appear\_anova$Athlete.AppearDoes not appear -6342.0 2320.2 -2.733 0.00635 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 26370 on 1329 degrees of freedom

(6 observations deleted due to missingness)

Multiple R-squared: 0.005591, Adjusted R-squared: 0.004842

F-statistic: 7.472 on 1 and 1329 DF, p-value: 0.006351

# RUN AN ANALYSIS TO SEE IF NUMBER OF SEASONS/AGE/POSITION/NATIONALITY IMPACTS LIKES (INSTEAD OF GE GENDER

# DIFFERENCES)

ANOVA is highly significant, thus the categories in Step 2 are statistically different in their engagement

# ANOVA WAS CONDUCTED TO COMPARE THE EFFECT OF SELF-PRESENTATION CATEGORIES OF

# ATHLETES’ PHOTOS ON THE NUMBER OF LIKES RECEIVED (I.E. ENGAGEMENT). THERE WAS

# A SIGNIFICANT DIFFERENCE IN NUMBER OF LIKES RECEIVED ACROSS SELF-PRESENTATION

# CATEGORIES AT THE p < .05 SIGNIFICANCE LEVEL. NOTE: STEP 2 ONLY CATEGORIZES THE

# PICTURES IN WHICH ATHLETES WERE PRESENT

# NEED TO LOOK UP HOW TO REPORT TUKEY’S

# THERE ARE STATISTICALLY SIGNIFICANT DIFFERENCES AT THE p < 0.05 LEVEL BETWEEN

# SEXUAL AND FEMININITY (0.0000), NON-SPORT SETTING AND DRESSED BUT POSED (0.0055),

# SEXUAL AND DRESSED BUT POSED (0.0000), SEXUAL AND MIXED MESSAGE (0.0000),

# NON-SPORT SETTING AND ATHLETE ACTION (0.0055), SEXUAL AND ATHLETE ACTION (0.0000),

# AND SEXUAL AND NON-SPORT SETTING (0.0000)

# INTERESTINGLY, THE CATEGORY WITH MORE LIKES (I.E., ENGAGEMENT) IS SEXUALLY SUGGESTIVE. # ALONG WITH THAT CATEGORY, FEMINITY AND NON-SPORT SETTING ARE STATISTICALLY SIGNIFICANT # AT THE 0.05 SIGNIFICANCE LEVEL.

# THERE IS A SIGNIFICANT DIFFERENCE AT THE p < .05 LEVEL IN THE NUMBER OF LIKES

# A POST RECEIVES IF THE ATHLETE IS PRESENT (0.0135). A LINEAR MODEL INDICATED

# THAT THERE IS A NEGATIVE CORRELATION BETWEEN THE ATHLETE NOT APPEARING IN THE

# PHOTO AND THE NUMBER OF LIKES THAT A PHOTO RECEIVES AT THE p < .05 CONFIDENCE

# LEVEL (0.00635).

filter(Step\_2\_merged, Step.2 == "Sexual")

filter(Step\_2\_merged, Account == "ecambage") %>%

select(Step.2, Likes)

filter(Step\_2\_merged, Account == "tea.cooper2") %>%

select(Step.2, Likes)

NA

