SURVEY DATA ANALYSIS AND INSIGHTS

THE POST

ANNA WANDER, PRIYA THAKORE,
EMMA MCELWEE, LEAH SIMON

WHAT HAVE WE DONE?

1. Experience variable analysis
2. Factor Analysis
3. Linear Regression
4. Additional Analysis

THE SURVEY ITSELF

ABOUT OUR RESEARCH

- ➔ Received data from Post, a media company, with information about their consumers and the consumer experience with Post content
- ➔ Total of with 69 variables with 739 observations
- ➔ Combination of categorical and numeric variables
- ➔ Tasked with learning more about which factors may affect the frequency with which consumers interact with Post

SURVEY PURPOSE

What are we trying to find?

Managerial decision problem

How can we increase our reader frequency?

Marketing research problem

What causes some people to read the Post more often than other people?

Research questions

What factors do readers value in a newspaper?

What types of consumers value these specific factors?

How can we increase the presence of these factors in our newspaper content?

INITIAL DATA ANALYSIS

WHAT EXPERIENCE VARIABLES MOST IMPACT HOW FREQUENTLY PEOPLE READ THE POST?

1. UTILITY
2. INFORMATION
3. SOCIAL
4. CIVIC
5. SURPRISE
6. FEEL GOOD
7. FACT BIAS
8. DIVERSITY

INITIAL EXPERIENCE VARIABLE ANALYSIS

CORRELATION

```
Correlation
Data        : Post_market_survey
Method      : Pearson
Variables   : freqPost, aveeutil, aveinfo, avesocial, avecivic, avesurprise, avefeelgood, avefactbias, avediv
Null hyp.   : variables x and y are not correlated
Alt. hyp.   : variables x and y are correlated

Correlation matrix:
             freqPost aveeutil aveinfo avesocial avecivic avesurprise avefeelgood avefactbias
aveeutil     0.07
aveinfo      0.06     0.64
avesocial    0.19     0.44     0.37
avecivic     0.11     0.53     0.58     0.58
avesurprise  0.05     0.46     0.38     0.50      0.42
avefeelgood  0.07     0.57     0.44     0.57      0.56      0.57
avefactbias  0.08     0.50     0.68     0.37      0.58      0.40        0.43
avediversity 0.06     0.37     0.43     0.41      0.57      0.45        0.44        0.44

p.values:
             freqPost aveeutil aveinfo avesocial avecivic avesurprise avefeelgood avefactbias
aveeutil     0.07
aveinfo      0.09     0.00
avesocial    0.00     0.00     0.00
avecivic     0.00     0.00     0.00     0.00
avesurprise  0.16     0.00     0.00     0.00      0.00
avefeelgood  0.07     0.00     0.00     0.00      0.00      0.00
avefactbias  0.03     0.00     0.00     0.00      0.00      0.00        0.00
avediversity 0.10     0.00     0.00     0.00      0.00      0.00        0.00        0.00
```

we found that the three most correlated experience variables were **“avesocial”, “avecivic”**, and **“avefactbias”**

LINEAR REGRESSION

MULTIPLE REGRESSION (OLS)

```
             coefficient std.error t.value p.value
 (Intercept)       3.337     0.293  11.388  < .001 ***
 aveeutil          0.006     0.075   0.084   0.933
 aveinfo          -0.033     0.082  -0.397   0.692
 avesocial         0.302     0.068   4.435  < .001 ***
 avecivic          0.049     0.086   0.577   0.564
 avesurprise      -0.063     0.069  -0.914   0.361
 avefeelgood      -0.074     0.071  -1.050   0.294
 avefactbias       0.066     0.081   0.818   0.414
 avediversity     -0.021     0.064  -0.322   0.747

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.04,  Adjusted R-squared: 0.029
F-statistic: 3.797 df(8,730), p.value < .001
Nr obs: 739
```

“avesocial” was the only factor that statistically significantly affected the frequency of reading the Post (“**freqPost**”)

FACTOR ANALYSIS

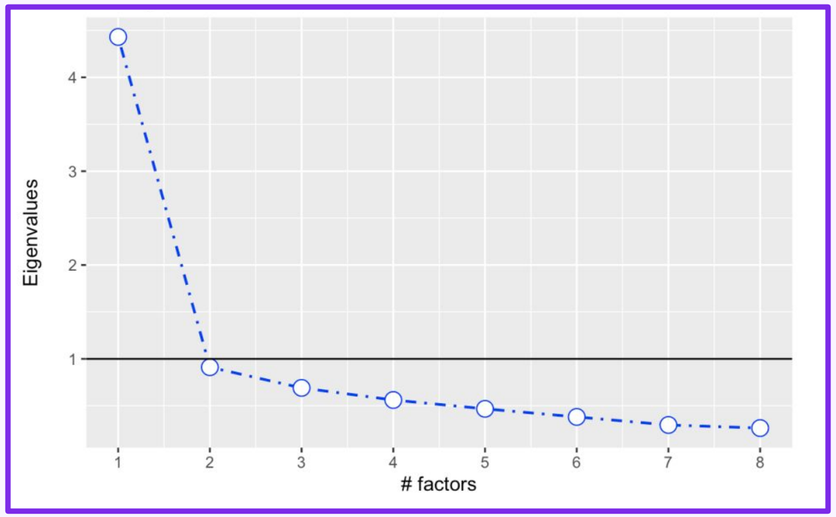
KMO & BARTLETT

```
Bartlett test
Null hyp. : variables are not correlated
Alt. hyp. : variables are correlated
Chi-square: 2816.32 df(28), p.value < .001

KMO test:  0.87

Variable collinearity:
                Rsq  KMO
Ave_Eutil      0.53 0.87
Ave_Einfo      0.60 0.82
Ave_Esocial    0.45 0.88
Ave_Ecivic     0.59 0.87
Ave_Esurprise  0.42 0.88
Ave_Efeelgood  0.53 0.89
Ave_Efactbias  0.53 0.87
Ave_Ediversity 0.39 0.90
```

Bartlett - significant p, we can do factor analysis

KMO - is between 0.5 and 1, factor analysis is appropriate

SCREE PLOT



Based on the elbow method, we decided to test for 2-3 factors

FACTORS (3):

```
Factor loadings:
                RC1  RC2  RC3
Ave_Efeelgood  0.79
Ave_Esurprise  0.76
Ave_Esocial    0.75
Ave_Einfo           0.87
Ave_Efactbias       0.76
Ave_Eutil      0.51 0.72
Ave_Ediversity           0.86
Ave_Ecivic      0.4 0.48 0.58
```

RC1 - emotional

RC2 - info

RC3 - community

FACTORS: WHAT DO OUR READERS LOOK FOR?

EMOTIONS

These readers value stories that make them feel good, allow them to connect socially with others, and are surprising - stories that keep them entertained.

INFORMATION

There readers value stories that give them important information about current events that is useful in their day-to-day life, unbiased, and relevant.

COMMUNITY

Theses readers value stories that teach them about the people in their community, celebrate diversity. These readers want to feel connected with those around them.

```
Linear regression (OLS)
Data     : Post_market_survey_1_
Response variable    : freqPost
Explanatory variables: emotional, info, community
Null hyp.: the effect of x on freqPost is zero
Alt. hyp.: the effect of x on freqPost is not zero
**Standardized coefficients shown (2 X SD)**

             coefficient std.error t.value p.value
(Intercept)        0.000     0.018   0.000   1.000
emotional          0.101     0.037   2.768   0.006 **
info               0.035     0.037   0.958   0.338
community          0.066     0.037   1.805   0.071 .

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.016,  Adjusted R-squared: 0.012
F-statistic: 3.946 df(3,735), p.value 0.008
Nr obs: 739
```

REGRESSION RESULTS WITH FACTORS

When condensed, we can see that the emotional variable as a whole has the highest impact on reader frequency.

If you increase the components of an article linked to emotion, that will increase its frequency of being read by 10.1%

FOCUS ON SOCIAL/EMOTIONAL ASPECT

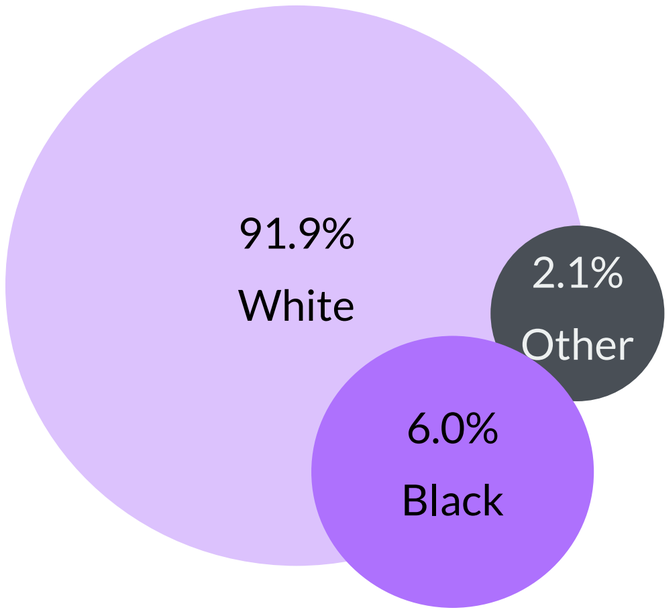
Potential Marketing Suggestions

1. Seek out local news stories, or local angles on global stories
2. Publish inspirational stories from readers
3. Add a fun facts section about topics that could change weekly

ADDITIONAL INSIGHTS

CORRELATION BETWEEN RACE WITH FACTORS

Race Distribution



Regression

```
             coefficient std.error t.value p.value
(Intercept)        0.000     0.018   0.000   1.000
emotional         -0.004     0.036  -0.109   0.913
info              -0.109     0.036  -3.000   0.003 **
community          0.112     0.036   3.068   0.002 **

Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.024,  Adjusted R-squared: 0.02
F-statistic: 6.14 df(3,735), p.value < .001
Nr obs: 739
```

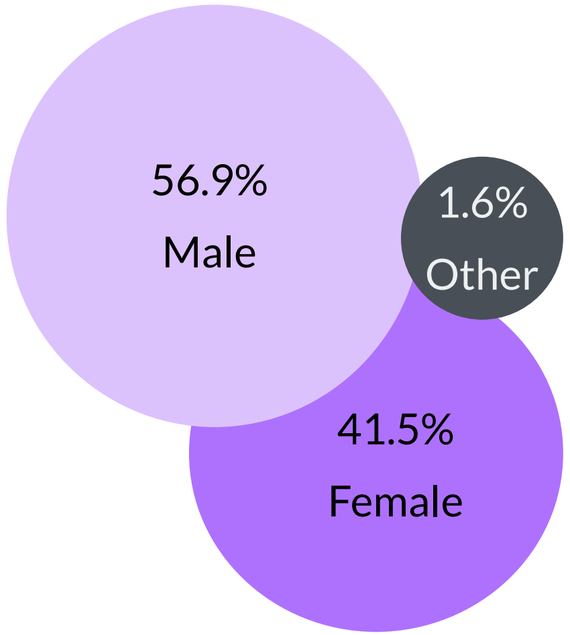
skewed data with majority white participants

white participants care more about information factors

black participants care more about community factors

CORRELATION BETWEEN GENDER WITH FACTORS

Gender Distribution



Regression

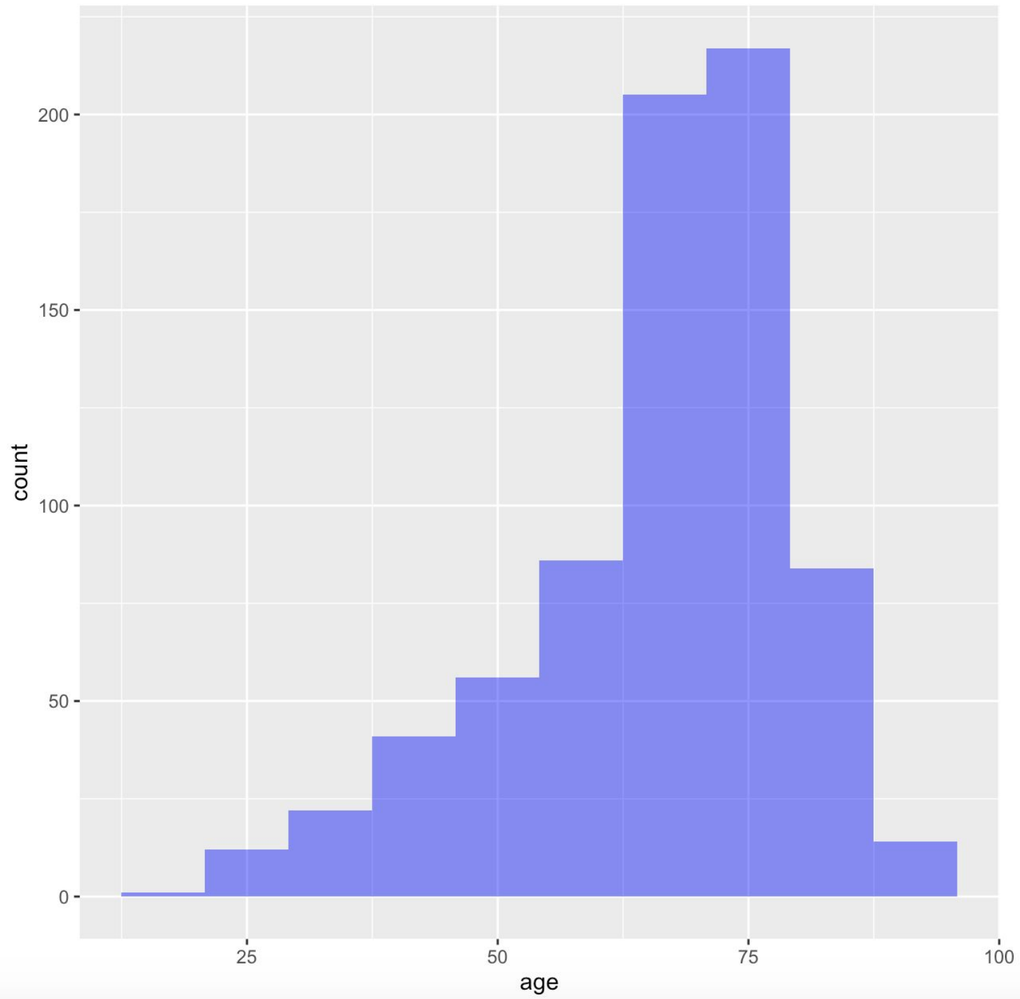
	coefficient	std.error	t.value	p.value
(Intercept)	-0.000	0.018	-0.000	1.000
emotional	0.010	0.037	0.281	0.779
info	0.034	0.037	0.937	0.349
community	0.109	0.037	2.966	0.003 **

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-squared: 0.013,  Adjusted R-squared: 0.009
F-statistic: 3.252 df(3,735), p.value 0.021
Nr obs: 739
```

Significant relationship between gender with community factors

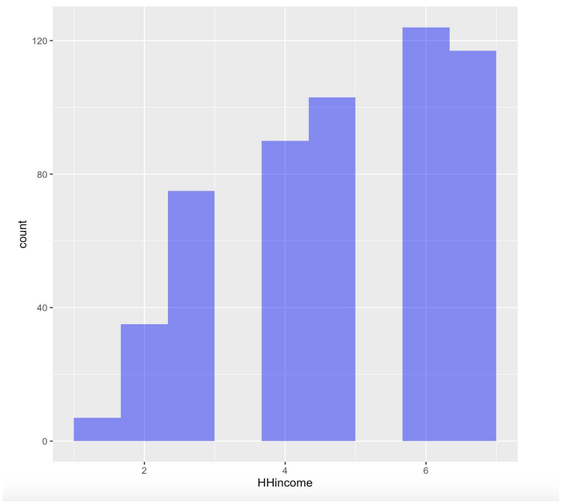
females care more community factor than males

AGE

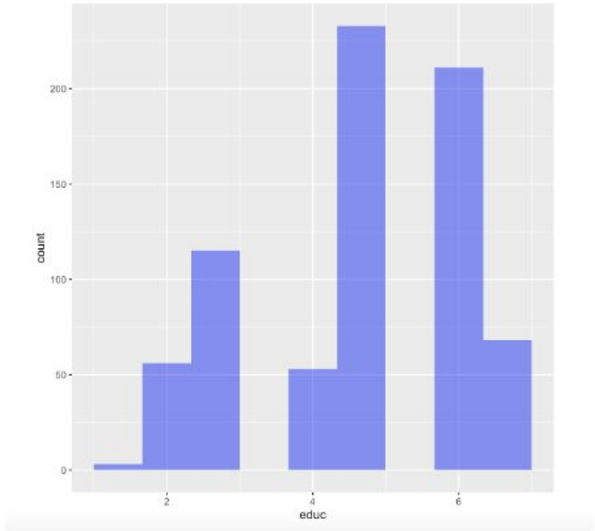


factbias - the younger they are, the more participants valued content based on verifiable facts

HOUSEHOLD INCOME



EDUCATION



<u>The more money they made and the higher level of education they got...</u>

feelgood - the less participants valued content that made them feel positively

surprise - the less participants valued surprising content

util - the less participants valued content useful for their lives

IMPLICATIONS & RECOMMENDATIONS

FURTHER RESEARCH

Broadening Audiences

Should we focus on a more specific target audience? Is it worth targeting a more diverse audience? Should we utilize non-probability sampling?

Understanding the disconnect between our initial findings and our additional findings (interest vs. disinterest in social/emotional).

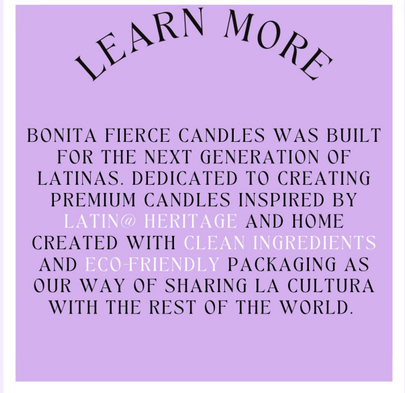
Figuring out whether more educated/wealthier audiences are solely deterred by feel good / surprise / utility content or whether they do have specific interests in media.

Marketing Campaign Suggestion

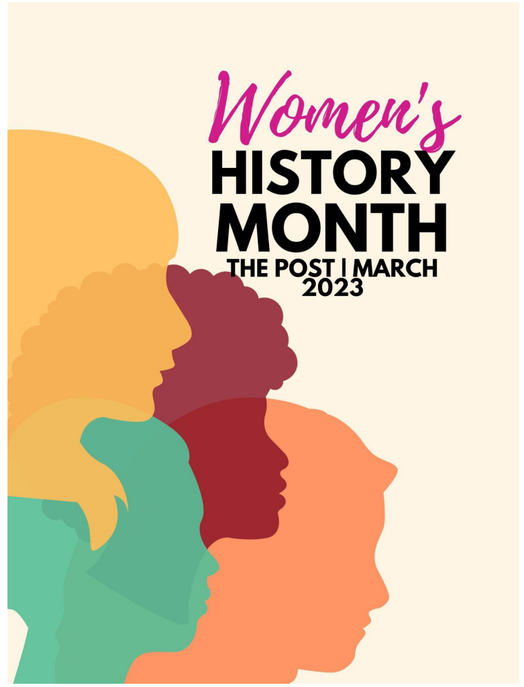
”Friday Feel-Good”

Campaign to share emotional news content on a weekly basis - can be shared via social media and in our publication

Small Business



Women's History Month



Marketing Campaign Suggestion

"Women in the Community""

Campaign to share news about women doing impactful work in our community to increase engagement from a female community

Target certain time frames such as Women's History month to draw in female subscribers through discount deals

FINAL RECOMMENDATIONS

1. Train reporters on finding the emotion behind the story and seeking out local stories about global events (talked-about in social contexts).
2. Identify our consumer demographic, and if it is worth it for us to target a more diverse audience.
3. Pending who the specific target audience is, conduct additional research to support segmented marketing campaigns.