Research\_Team\_2\_Progress\_Update\_Report\_Knight

Michael Knight

11/21/2021

## Importing and Formatting the data

## # A tibble: 4,654 × 190  
## QKEY NEW\_Device\_Type\_W14 LANG\_W14 LOCALRATING\_W14 LOCALIMPACT\_W14 LIVE3\_W14  
## <dbl> <fct> <fct> <fct> <fct> <fct>   
## 1 100197 Tablet English Good A small impact 10-25 mi…  
## 2 100260 Tablet English Good A moderate imp… Less tha…  
## 3 100314 Desktop English Good A moderate imp… More tha…  
## 4 100363 Desktop English Good A moderate imp… 26-50 mi…  
## 5 100446 Desktop English Excellent A moderate imp… 51-100 m…  
## 6 100514 Desktop English Excellent A big impact More tha…  
## 7 100588 Mobile phone English Only Fair A small impact 26-50 mi…  
## 8 100598 Desktop English Excellent A big impact More tha…  
## 9 100734 Desktop English Excellent A big impact More tha…  
## 10 100803 Desktop English Excellent A big impact 51-100 m…  
## # … with 4,644 more rows, and 184 more variables: WORKCLOSE\_W14 <fct>,  
## # NEIGHBORS\_W14 <fct>, COMATTACH\_W14 <fct>, SOCTRUST\_W14 <fct>,  
## # COMTRUST\_W14 <fct>, NEWS\_LEVELA\_W14 <fct>, NEWS\_LEVELB\_W14 <fct>,  
## # NEWS\_LEVELC\_W14 <fct>, NEWS\_LEVELD\_W14 <fct>, NEWS\_TOPICA\_W14 <fct>,  
## # NEWS\_TOPICB\_W14 <fct>, NEWS\_TOPICC\_W14 <fct>, NEWS\_TOPICD\_W14 <fct>,  
## # NEWS\_TOPICE\_W14 <fct>, NEWS\_TOPICF\_W14 <fct>, NEWS\_TOPICG\_W14 <fct>,  
## # NEWS\_TOPICH\_W14 <fct>, NEWS\_DEVICEA\_W14 <fct>, NEWS\_DEVICEB\_W14 <fct>, …

## Rename The Columns We will be Observing

names(atp\_as\_factor)[names(atp\_as\_factor) == 'NEWS\_PREFER\_W14'] <- 'News\_Preference'  
names(atp\_as\_factor)[names(atp\_as\_factor) == 'DEBBOTH2\_W14'] <- 'Debate\_Preference'  
names(atp\_as\_factor)[names(atp\_as\_factor) == 'SOCTRUST\_W14'] <- 'Trust\_in\_Others'  
names(atp\_as\_factor)[names(atp\_as\_factor) == 'GROUP\_TRUSTD\_W14'] <- 'Trust\_in\_News'  
names(atp\_as\_factor)[names(atp\_as\_factor) == 'DIGNEWSFACT\_W14'] <- 'Fact\_Check'  
names(atp\_as\_factor)[names(atp\_as\_factor) == 'F\_PARTY\_FINAL'] <- 'Party'  
names(atp\_as\_factor)[names(atp\_as\_factor) == 'NEWS\_PLATFORMG\_W14'] <- 'How\_often\_social\_media\_for\_news'  
  
  
  
atp\_as\_factor %>%   
 select('News\_Preference', 'Debate\_Preference', 'Trust\_in\_Others', 'Trust\_in\_News', 'Fact\_Check', 'Party', 'How\_often\_social\_media\_for\_news')

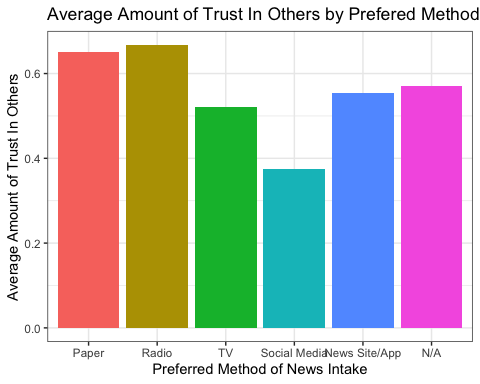
## # A tibble: 4,654 × 7  
## News\_Preference Debate\_Preference Trust\_in\_Others Trust\_in\_News Fact\_Check  
## <fct> <fct> <fct> <fct> <fct>   
## 1 Watching news on … Debates for both… Most people ca… Not too much Sometimes   
## 2 Listening to news… Debates for the … Most people ca… Not at all Often   
## 3 Watching news on … <NA> Can't be too c… Not too much Hardly ev…  
## 4 Reading news in a… <NA> Most people ca… Some Sometimes   
## 5 Watching news on … <NA> Most people ca… Not too much Sometimes   
## 6 Watching news on … Debates for the … Most people ca… Not at all Sometimes   
## 7 Getting news from… <NA> Can't be too c… A lot Sometimes   
## 8 Getting news from… <NA> Most people ca… Not at all Sometimes   
## 9 Watching news on … Debates for both… Most people ca… Not at all <NA>   
## 10 Watching news on … <NA> Can't be too c… Not too much Hardly ev…  
## # … with 4,644 more rows, and 2 more variables: Party <fct>,  
## # How\_often\_social\_media\_for\_news <fct>

The hypotheses for this research are: 1. People who get their news from social media feel that most people cannot be trusted. News\_Preference vs. Trust\_in\_Others

# atp\_as\_factor %>%   
# select('News\_Preference', 'How\_often\_social\_media\_for\_news', 'Trust\_in\_Others')  
  
  
# 26 nulls for News\_Preference  
#sum(is.na(atp\_as\_factor$News\_Preference))  
# 0 nulls for Trust\_in\_Others  
#sum(is.na(atp\_as\_factor$Trust\_in\_Others))  
  
  
#unique(atp\_as\_factor$News\_Preference)  
#unique(atp\_as\_factor$Trust\_in\_Others)  
  
# only 26 instances where user refused to fill in answer for Trust\_in\_Others  
sum(atp\_as\_factor$Trust\_in\_Others == "Refused")

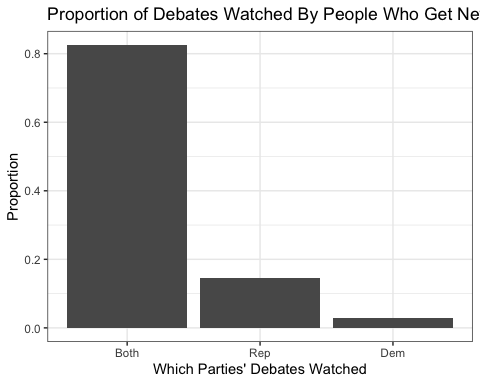
## [1] 26

# Recode values  
atp\_as\_factor %>%   
 select('News\_Preference', 'Trust\_in\_Others') %>%   
 mutate(Trust\_in\_Others = recode(Trust\_in\_Others,  
 "Can't be too careful" = 0,  
 "Most people can be trusted" = 1,  
 "Refused" = 2),  
 News\_Preference = recode(News\_Preference,  
 "Refused" = "N/A",  
 "Watching news on television" = "TV",  
 "Listening to news on the radio" = "Radio",  
 "Reading news in a print newspaper" = "Paper",   
 "Getting news from a social networking site (such as Facebook or Twitter)" = "Social Media",  
 "Getting news from a news website or app" = "News Site/App")) -> social\_med\_vs\_trusts\_people  
  
  
#drop 64 rows that had null value for Trusts\_in\_others (out of 4,654 total entries)  
social\_med\_vs\_trusts\_people <- na.omit(social\_med\_vs\_trusts\_people)  
  
  
# notice that after dropping the null values, there are 23 entries where there was a refusal to answer Trust others  
#sum(social\_med\_vs\_trusts\_people$Trust\_in\_Others == 2)  
  
# drop the 23 entries where people refused to answer the trust question  
social\_med\_vs\_trusts\_people %>%   
 filter(Trust\_in\_Others != 2) -> social\_med\_vs\_trusts\_people  
#sum(social\_med\_vs\_trusts\_people$Trust\_in\_Others == 2)  
  
  
social\_med\_vs\_trusts\_people %>%   
 group\_by(News\_Preference) %>%   
 summarise(mean\_trust = mean(Trust\_in\_Others)) %>%  
 ggplot(aes(x=News\_Preference, y=mean\_trust, fill = News\_Preference)) +   
 geom\_col() +  
 xlab("Preferred Method of News Intake") +   
 ylab("Average Amount of Trust In Others") +  
 ggtitle("Average Amount of Trust In Others by Prefered Method of Getting News") +   
 theme\_bw() +  
 theme(legend.position="none")



1. Most people who prefer to get their news from social media only watched either the republican or only the democratic debates.

# 64 null entries  
#sum(is.na(atp\_as\_factor$News\_Preference))  
# 2532 null entries! I recommend we not do this hypothesis  
#sum(is.na(atp\_as\_factor$Debate\_Preference))  
  
#unique(atp\_as\_factor$News\_Preference)  
#unique(atp\_as\_factor$Debate\_Preference)  
  
# recode variables to be less long and clunky  
atp\_as\_factor %>%   
 select('News\_Preference', 'Debate\_Preference') %>%   
 mutate(Debate\_Preference = recode(Debate\_Preference,  
 "Debates for both party's candidates" = "Both",  
 "Debates for the Republican party's candidate" = "Rep",  
 "Debates for the Democratic party's candidate" = "Dem",  
 "Refused" = "N/A"),  
 News\_Preference = recode(News\_Preference,  
 "Refused" = "N/A",  
 "Watching news on television" = "TV",  
 "Listening to news on the radio" = "Radio",  
 "Reading news in a print newspaper" = "Paper",   
 "Getting news from a social networking site (such as Facebook or Twitter)" = "Social Media",  
 "Getting news from a news website or app" = "News Site/App")) -> debates\_only\_one\_party  
  
  
#drop 2532 rows that had null value for Trusts\_in\_others (out of 4,654 total entries)  
debates\_only\_one\_party <- na.omit(debates\_only\_one\_party)  
  
  
# hypothesis clearly very false  
debates\_only\_one\_party %>%   
 filter(News\_Preference == "Social Media") %>%   
 ggplot() +   
 geom\_bar(aes(x=Debate\_Preference, y=stat(prop), group = 1)) +   
 xlab("Which Parties' Debates Watched") +  
 ylab("Proportion") +  
 ggtitle("Proportion of Debates Watched By People Who Get News From Social Media") +   
 theme\_bw()



1. People who do not trust the news they see online are more likely to fact check.

# 315 null values  
sum(is.na(atp\_as\_factor$Trust\_in\_News))

## [1] 315

# 580 null values  
sum(is.na(atp\_as\_factor$Fact\_Check))

## [1] 580

# unique(atp\_as\_factor$Trust\_in\_News)  
# unique(atp\_as\_factor$Fact\_Check)  
#   
# atp\_as\_factor %>%   
# select('Trust\_in\_News', 'Fact\_Check')  
# atp\_numerical %>%   
# select('Trust\_in\_News', 'Fact\_Check')  
  
atp\_as\_factor %>%   
 select('Trust\_in\_News', 'Fact\_Check') %>%   
 mutate(Fact\_Check = recode(Fact\_Check,  
 "Never" = 0,  
 "Hardly ever" = 1,  
 "Sometimes" = 2,  
 "Often" = 3,  
 "Refused" = 4),  
 Trust\_in\_News = recode(Trust\_in\_News,  
 "Not at all" = 0,  
 "Not too much" = 1,  
 "Some" = 2,  
 "A lot" = 3,   
 "Refused" = 4)) -> news\_trust\_v\_fact\_check\_numerical  
  
  
news\_trust\_v\_fact\_check\_numerical <- na.omit(news\_trust\_v\_fact\_check\_numerical)  
  
# only 10 values where Refused to answer about Fact Checking  
#sum(news\_trust\_v\_fact\_check\_numerical$Fact\_Check == 4)  
  
# only 20 values where Refused to answer about Trust\_in\_News  
#sum(news\_trust\_v\_fact\_check\_numerical$Trust\_in\_News == 4)  
  
# drop the entries where people refused to answer the trust question and the fact check question  
news\_trust\_v\_fact\_check\_numerical %>%   
 filter(Trust\_in\_News != 4 & Fact\_Check != 4) -> news\_trust\_v\_fact\_check\_numerical  
  
  
  
cor(news\_trust\_v\_fact\_check\_numerical)

## Trust\_in\_News Fact\_Check  
## Trust\_in\_News 1.0000000 0.0389359  
## Fact\_Check 0.0389359 1.0000000

atp\_as\_factor %>%   
 select('Trust\_in\_News', 'Fact\_Check') -> news\_trust\_v\_fact\_check\_as\_factor  
  
news\_trust\_v\_fact\_check\_as\_factor <- na.omit(news\_trust\_v\_fact\_check\_as\_factor)  
  
# only 10 values where Refused to answer about Fact Checking  
#sum(news\_trust\_v\_fact\_check\_as\_factor$Fact\_Check == "Refused")  
  
# only 20 values where Refused to answer about Trust\_in\_News  
#sum(news\_trust\_v\_fact\_check\_as\_factor$Trust\_in\_News == "Refused")  
  
  
  
# drop the entries where people refused to answer the trust question and the fact check question  
news\_trust\_v\_fact\_check\_as\_factor %>%   
 filter(Trust\_in\_News != "Refused" & Fact\_Check != "Refused") -> news\_trust\_v\_fact\_check\_as\_factor  
  
#news\_trust\_v\_fact\_check\_as\_factor  
  
# hypothesis inconclusive  
news\_trust\_v\_fact\_check\_as\_factor %>%   
 ggplot() +   
 geom\_bar(aes(x=Fact\_Check, y=stat(prop), group = 1)) +   
 facet\_wrap("Trust\_in\_News") +  
 xlab("How Often Fact Check") +  
 ylab("Proportion") +  
 ggtitle("Proportion of Fact Checking by How Much People Trust Online News") +  
 theme\_bw()

