Predicting Trust in Media Across Countries

Jack Miller

#Loading Packages and Data

```
library(broom)
library(inspectdf)
library(foreign)
library(tidyverse)
library(patchwork)
library(knitr)
```

```
hkwave1 <- read.spss("data/Hong Kong v4.2.sav", to.data.frame=TRUE)
japanwave1 <- read.spss("data/Japan v4.2.sav", to.data.frame=TRUE)
```

These are the packages we will be using for data analysis as well as the datasets we will be working with from various Asian Barometer Survey waves.

Data Cleaning

```
japanwave1 %>% inspect_types()
```

```
## # A tibble: 2 x 4
##   type      cnt  pcnt col_name
##   <chr>   <int> <dbl> <named list>
## 1 factor    259 93.5  <chr [259]>
## 2 numeric    18  6.50 <chr [18]>
```

```
hkwave1 %>% inspect_types()
```

```
## # A tibble: 2 x 4
##   type      cnt  pcnt col_name
##   <chr>   <int> <dbl> <named list>
## 1 factor    282 92.5  <chr [282]>
## 2 numeric    23  7.54 <chr [23]>
```

Most of our categorical data is already in factor form so we can continue.

```
hkwave1 <- hkwave1 %>%
  drop_na(q016) %>%
  mutate(TV_Trust = as.factor(case_when(
    q016 == "QUITE A LOT OF TRUST" | q016 == "A GREAT DEAL OF TRUST" ~ 1,
    q016 == "NONE AT ALL" | q016 == "NOT VERY MUCH TRUST" ~ 0
  ))) %>%
  drop_na(level3) %>%
  mutate(urban = as.factor(case_when(
    level3 == "urban" ~ 1,
    level3 == "rural" ~ 0
  ))) %>%
```

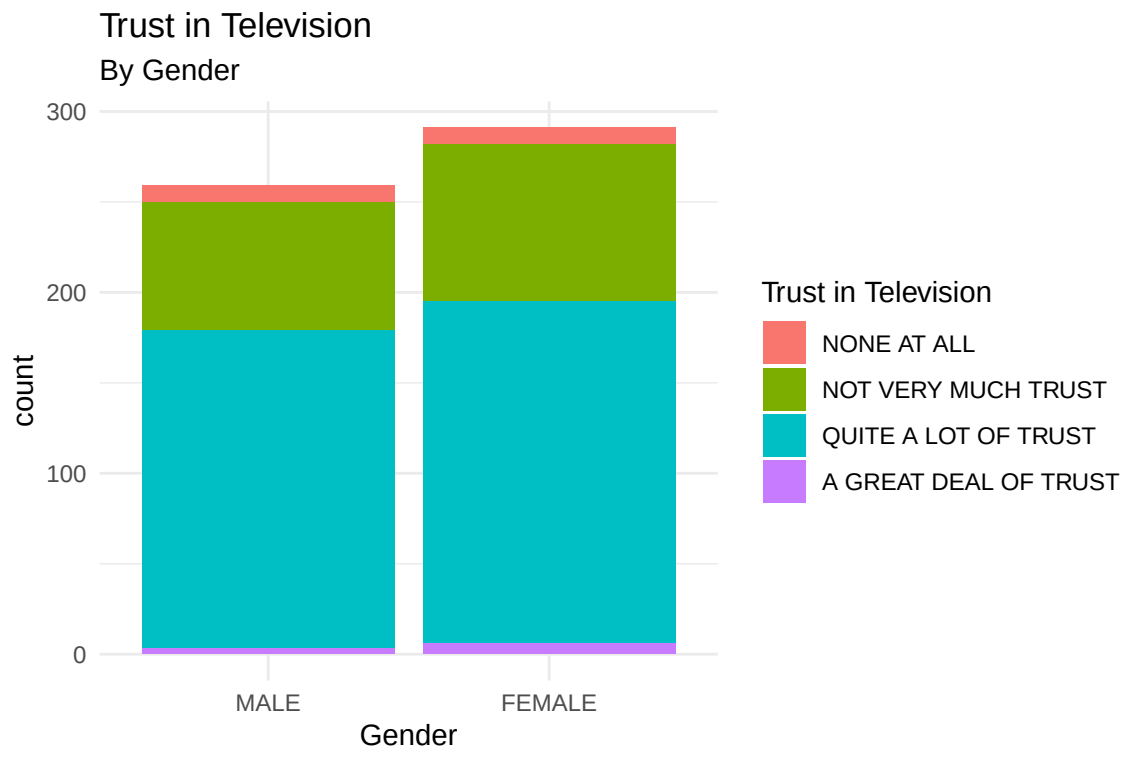
```
  drop_na(se003a) %>%
  mutate(age = as.numeric(se003a)) %>%
  drop_na(se005b) %>%
  mutate(educated = as.factor(case_when(
    se005b == "tertiary education" | se005b == "secondary education" ~ 1,
    se005b == "primary education" | se005b == "illiterate" ~ 0
  ))) %>%
  drop_na(se002) %>%
  mutate(male = as.factor(case_when(
    se002 == "MALE" ~ 1,
    se002 == "FEMALE" ~ 0
  ))) %>%
  drop_na(se004a) %>%
  mutate(married = as.factor(case_when(
    se004a == "Yes" ~ 1,
    TRUE ~ 0
  ))) %>%
  drop_na(q024) %>%
  mutate(trusting = as.factor(case_when(
    q024 == "MOST PEOPLE CAN BE TRUSTED" ~ 1,
    TRUE ~ 0
  ))) %>%
  drop_na(se009) %>%
  mutate(income = as.factor(se009)) %>%
  drop_na(q008) %>%
  mutate(govtrust = as.factor(case_when(
    q008 == "QUITE A LOT OF TRUST" | q008 == "A GREAT DEAL OF TRUST" ~ 1,
    q008 == "NONE AT ALL" | q008 == "NOT VERY MUCH TRUST" ~ 0)))
```

```
japanwave1 <- japanwave1 %>%
  drop_na(q016) %>%
  mutate(TV_Trust = as.factor(case_when(
    q016 == "Quite a lot of trust" | q016 == "A great deal of trust" ~ 1,
    q016 == "None at all" | q016 == "Not very much trust" ~ 0
  ))) %>%
  drop_na(level3) %>%
  mutate(urban = as.factor(case_when(
    level3 == "urban" ~ 1,
    level3 == "rural" ~ 0
  ))) %>%
  drop_na(se003a) %>%
  mutate(age = as.numeric(se003a)) %>%
  drop_na(se005b) %>%
  mutate(educated = as.factor(case_when(
    se005b == "tertitary education" | se005b == "secondary education" ~ 1,
    se005b == "primary education" | se005b == "illiterate" ~ 0
  ))) %>%
  drop_na(se002) %>%
  mutate(male = as.factor(case_when(
    se002 == "male" ~ 1,
    se002 == "female" ~ 0
  ))) %>%
  drop_na(se004a) %>%
  mutate(married = as.factor(case_when(
```

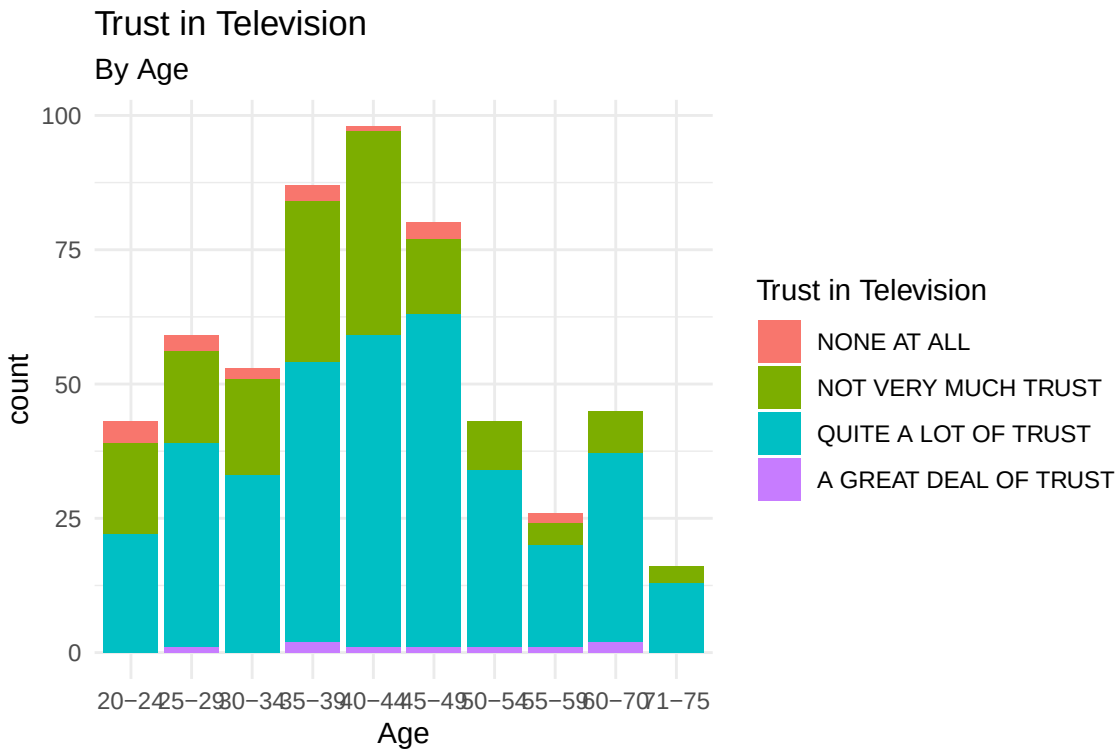
```
    se004a == "Yes" ~ 1,
    TRUE ~ 0
  ))) %>%
  drop_na(q024) %>%
  mutate(trusting = as.factor(case_when(
    q024 == "Most people can be trusted" ~ 1,
    TRUE ~ 0
  ))) %>%
  drop_na(se009) %>%
  mutate(income = as.factor(se009)) %>%
  drop_na(q008) %>%
  mutate(govtrust = as.factor(case_when(
    q008 == "Quite a lot of trust" | q008 == "A great deal of trust" ~ 1,
    q008 == "None at all" | q008 == "Not very much trust" ~ 0)))
```

Exploratory Data Analysis

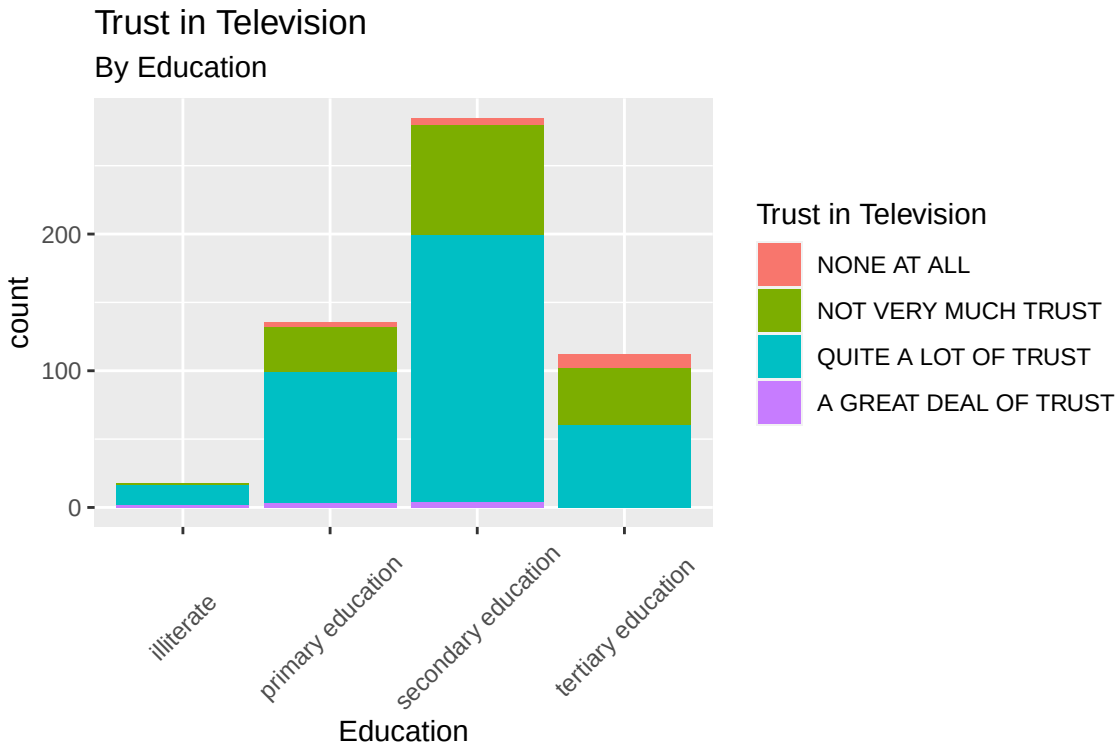
```
ggplot(data = hkwave1, aes(x = se002, fill = q016)) + geom_bar() +
  labs(title = "Trust in Television", subtitle = "By Gender", x = "Gender") +
  guides(fill=guide_legend(title="Trust in Television")) + theme_minimal()
```



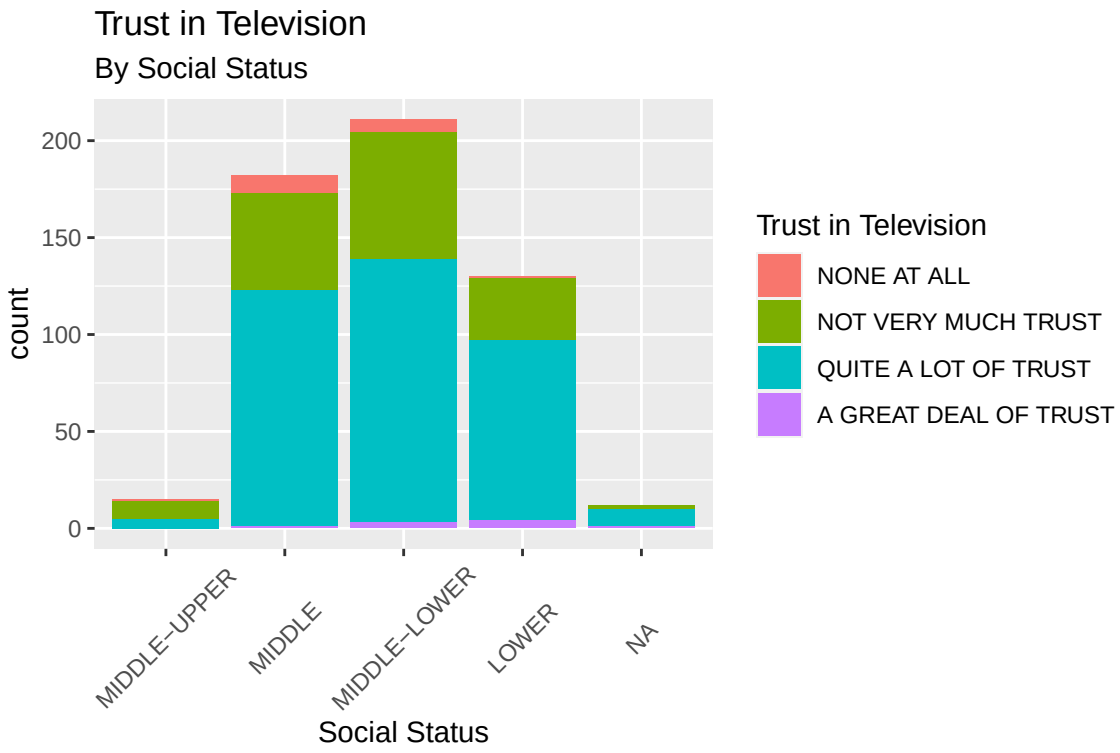

```
ggplot(data = hkwave1, aes(x = se003, fill = q016)) + geom_bar() +
  labs(title = "Trust in Television",subtitle = "By Age", x = "Age") +
  guides(fill=guide_legend(title="Trust in Television")) + theme_minimal()
```



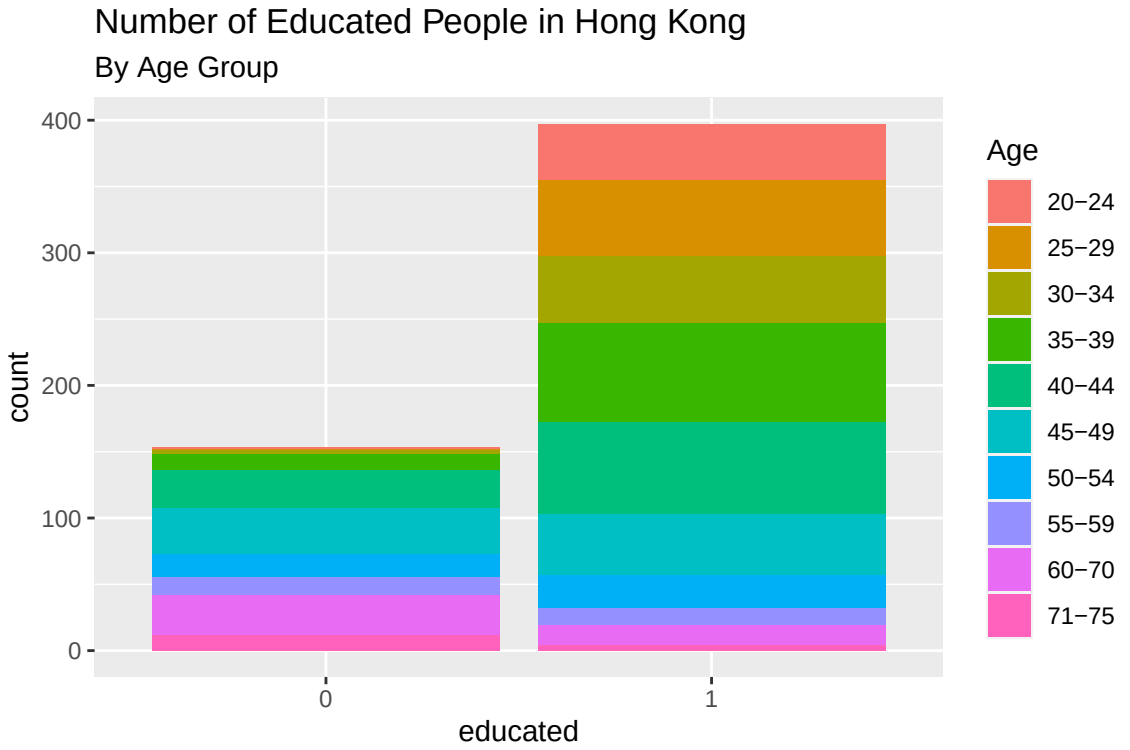

```
ggplot(data = hkwave1, aes(x = se005b, fill = q016)) + geom_bar() +
  labs(title = "Trust in Television", subtitle = "By Education",
       x = "Education") + guides(fill=guide_legend(title="Trust in Television")) +
  theme(axis.text.x = element_text(angle = 45, vjust = 0.5))
```



```
ggplot(data = hkwave1, aes(x = se017, fill = q016)) + geom_bar() +
  labs(title = "Trust in Television", subtitle = "By Social Status",
       x = "Social Status") +
  guides(fill=guide_legend(title="Trust in Television")) +
  theme(axis.text.x = element_text(angle = 45, vjust = 0.5))
```




```
hkwave1 %>%
  group_by(age, educated) %>%
  ggplot(aes(x = educated, fill = se003)) +
  geom_bar() +
  labs(title = "Number of Educated People in Hong Kong",
       subtitle = "By Age Group",
       ) +
  guides(fill=guide_legend(title="Age"))
```



While more people are eduacted than not (secondary or tertiary education), older age groups appear to be much more unlikely to be educated whereas there is a much greater difference in the number of people educated and uneducated among younger ages.

Research Question

Because of what we noticed in the difference in education levels and age, we wanted to explore a relationship that potentially had a lot of correlation with these two attributes. One thing our group was extremely interested in was the level of trust people had in their government and media and trying to predict this based on factors such as age, gender, and education. Our underlying motivation comes from two articles found while researching the Asian Barometer Survey, mainly one article from WNG about a

```
modelhkw1 <- glm(TV_Trust ~ educated + age + male + married + trusting + income + govtrust, data = hkwa
tidy(modelhkw1)
```

```
## # A tibble: 11 x 5
##    term                      estimate std.error statistic p.value
##    <chr>                        <dbl>     <dbl>     <dbl>   <dbl>
##  1 (Intercept)                 0.153      0.454     0.337   0.736
##  2 educated1                   0.0630     0.263     0.240   0.811
##  3 age                         0.0215     0.0100    2.15    0.0317
##  4 male1                       0.00839    0.192     0.0436  0.965
##  5 married1                    0.298      0.253     1.18    0.238
##  6 trusting1                   0.398      0.217     1.84    0.0665
##  7 income8000-14999 HK        -0.461      0.335    -1.38    0.169
##  8 income15000-24999 HK       -0.468      0.345    -1.36    0.175
##  9 income25000-39999 HK       -0.612      0.375    -1.63    0.103
## 10 income40000 HK AND ABOVE -1.00        0.392    -2.56    0.0105
## 11 govtrust1                  0.387      0.193     2.01    0.0449
```

```
modeljw1 <- glm(TV_Trust ~ educated + age + male + married + trusting + income + govtrust, data = japanw
tidy(modeljw1)
```

```
## # A tibble: 11 x 5
##    term                                estimate std.error statistic     p.value
##    <chr>                                  <dbl>     <dbl>     <dbl>       <dbl>
##  1 (Intercept)                          -0.187     0.375     -0.499      6.18e-1
##  2 educated1                            -0.462     0.218     -2.12       3.41e-2
##  3 age                                   0.00757   0.00601    1.26       2.08e-1
##  4 male1                                -0.0759    0.145     -0.523      6.01e-1
##  5 married1                              0.0393    0.266      0.148      8.83e-1
##  6 trusting1                             0.0652    0.154      0.424      6.72e-1
##  7 incomeGE 2,500,000 LT 4,500,000 yen   0.258     0.259      0.997      3.19e-1
##  8 incomeGE 4,500,000 LT 6,500,000 yen   0.214     0.266      0.805      4.21e-1
##  9 incomeGE 6,500,000 LT 10,000,000 yen  0.0118    0.266      0.0441     9.65e-1
## 10 incomeGE 10,000,000 yen               0.192     0.311      0.616      5.38e-1
## 11 govtrust1                             0.941     0.175      5.37       7.97e-8
```

```
new_hk <- glm(TV_Trust ~ age + trusting + govtrust, data = hkwave1, family = binomial)
tidy(new_hk) %>%
  kable(digits = 3)
```

term	estimate	std.error	statistic	p.value
(Intercept)	-0.239	0.218	-1.095	0.273
age	0.030	0.008	3.792	0.000
trusting1	0.336	0.212	1.584	0.113
govtrust1	0.402	0.191	2.103	0.035

Our backward select produces a model that has predictor variables age, whether or not the individual is trusting in general, and whether they trust the government. This tells us older people and people who are more trusting of the government and of people in general are more likely to be trusting of the television and media.

```
newjap <- glm(TV_Trust ~ educated + govtrust + govtrust*educated, data = japanwave1, family = binomial)
tidy(newjap) %>%
  kable(digits = 3)
```

term	estimate	std.error	statistic	p.value
(Intercept)	0.021	0.205	0.103	0.918
educated1	-0.235	0.223	-1.052	0.293
govtrust1	2.176	0.514	4.233	0.000
educated1:govtrust1	-1.448	0.547	-2.646	0.008

Unfortunately when we run backwards select on our Japan data, we see that the two biggest predictors for them are trust in the government and whether or not the person is educated.

Appendix

"Hong Kong's Defiant Filmmaker", WNG, September 21, 2021, https://wng.org/articles/hong-kongs-defiant-filmmaker-1632248931

"Almost half of journalists in Hong Kong considering leaving the city" https://thehill.com/policy/international/asia-pacific/580343-almost-half-of- journalists-in-hong-kong-media-group